Kirkwood-Dirac representations beyond quantum states and their relation to noncontextuality

David Schmid ®,[1,*] Roberto D. Baldijão,[1] Yìlè Yīng ®,[2,3,†] Rafael Wagner ®,[4,5,‡] and John H. Selby ®[1]

[1]*International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland*
[2]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*
[3]*Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[4]*INL–International Iberian Nanotechnology Laboratory, Av. Mestre José Veiga s/n, 4715-330 Braga, Portugal*
[5]*Centro de Física, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal*

Kirkwood-Dirac representations of quantum states are increasingly finding use in many areas within quantum theory. Usually, representations of this sort are only applied to provide a representation of quantum states (as complex functions over some set). We show how standard Kirkwood-Dirac representations can be extended to a fully compositional representation of all of quantum theory (including channels, measurements, and so on), and prove that this extension satisfies the essential features of functoriality (namely, that the representation commutes with composition of channels), linearity, and quasistochasticity. Interestingly, the representation of a positive operator-valued measure (POVM) element is uniquely picked out to be the collection of weak values for it relative to the bases defining the representation. We then prove that if one can find any Kirkwood-Dirac representation that is everywhere real and non-negative for a given experimental scenario or fragment of quantum theory, then the scenario or fragment is consistent with the principle of generalized noncontextuality, a key notion of classicality in quantum foundations. We also show that the converse does not hold: even if one verifies that all Kirkwood-Dirac representations (as defined herein) of an experiment require negativity or imaginarity, one cannot generally conclude that the experiment witnesses contextuality.

I. INTRODUCTION

It is often useful to represent quantum systems as functions over some kind of phase space. If one wishes to represent all of standard quantum theory, however, well-known no-go theorems demonstrate that such representations must be quasiprobabilistic rather than probabilistic [1,2]. Often, the term "quasi" refers to the fact that these functions take negative values, with well-known examples including Wigner quasiprobability distributions [3] or discrete versions thereof [4]. However, one intriguing family of quasiprobability distributions of quantum states that is defined over the complex numbers is the Kirkwood-Dirac (KD) quasiprobability distribution [5–7]. KD distributions have been used to study postselected quantum metrology [8–10], quantum fluctuation theorems [11–13], work extraction [14–17], heat flow [18], weak value theory and quantum coherence estimation [19,20], indefinite causal order [21], incompatibility [22–24], and scrambling of quantum information [25–27]. Moreover, KD distributions can be experimentally measured in various ways [15,28]. Perhaps surprisingly, given the recent surge of interest in such representations, KD distributions have so far only been described for quantum states, in contrast to the majority of other established representations (such as Wigner's

that can be used to represent quantum dynamics, quantum measurements, and indeed arbitrary quantum processes.

The first contribution of this work is to extend KD distributions as typically defined—for quantum states only—to full representations of quantum theory, applicable to arbitrary quantum processes—states, channels, measurements, instruments, and so on. Our extension is designed to satisfy a powerful and elegant mathematical property: that the representation commutes with composition. Formally, the representation is a symmetric strict monoidal functor [29] (i.e., is diagram preserving [30]). Our arguments here are the natural extension of work regarding real-valued quasiprobability representations, most notably Refs. [1,2,30–32], to the complex-valued family of KD representations. Indeed, all of these constructions are special cases of what are known as frame representations [33], and this fact immediately unifies much of the study of KD representations and other representations (such as the Wigner function).

Our representation may be of independent interest to research on weak values [34–36]. In particular, we show that the representation of a positive operator-valued measure (POVM) element in a particular KD representation is given by the collection of weak values for that operator with respect to the basis vectors defining the KD representation. This implies a formal duality between KD representations for states and weak values.

It is sometimes suggested that when a quantum state's KD distribution has negative real components or nonzero imaginary components, respectively termed negativity and

---
*Contact author: davidschmid10@gmail.com
†Contact author: yying@pitp.ca
‡Contact author: rafael.wagner@inl.int

imaginarity, then that state is somehow nonclassical. But (as has been argued before [2]) it is meaningless to talk about the nonclassicality of a state without specifying more information—e.g., the measurements one can do on the state or the experiment one will embed the state in. By considering the representation of arbitrary processes (not merely quantum states), we begin to derive direct connections between KD representations and an important foundational notion of classicality, namely generalized noncontextuality. Generalized noncontextuality [37] (henceforth simply noncontextuality) is a principled [37–39], useful [40–53], and operational [54–62] notion of classicality, relative to which one can classify any theory or experiment as classically explainable or not. Indeed, this classification has been studied for phenomena ranging from computation [63,64], state discrimination [40,65–67], interference [68–70], compatibility [61,71], uncertainty relations [72], metrology [53], thermodynamics [53,73], weak values [41,42], coherence [74,75], quantum Darwinism [76], information processing and communication [43–46,51], cloning [47], and (as mentioned above) Bell [77,78] and Kochen-Specker correlations [54,57,79–83].

The second main contribution of this paper is to prove that if an experiment or theory admits of any Kirkwood-Dirac representation that is everywhere real and non-negative, then that theory or experiment is consistent with noncontextuality. We also prove that the converse does not hold—so that even if every KD representation (as defined herein) requires negativity or imaginarity, one has not necessarily proven the failure of noncontextuality.

II. KIRKWOOD-DIRAC REPRESENTATIONS OF STATES, CHANNELS, AND MEASUREMENTS

Given a finite-dimensional Hilbert space $\mathcal{H}_A$ associated to a single quantum system A, pick any two bases $\{|a_i\rangle\}_i$ and $\{|a'_{i'}\rangle\}_{i'}$ for the Hilbert space, under the constraint that $\langle a_i|a'_{i'}\rangle \neq 0$ for all i and i'. From these, define a (nonovercomplete) basis $F := \{F_{i,i'}\}_{i,i'}$ for the space of bounded linear operators $\mathcal{B}(\mathcal{H}_A)$ as

$$F_{i,i'} := |a'_{i'}\rangle\langle a_i|\langle a'_{i'}|a_i\rangle. \tag{1}$$

Such a basis is an example of a frame [33]. Next, define the dual frame $D := \{D_{i,i'}\}_{i,i'}$ as the unique basis of operators satisfying

$$\mathrm{Tr}[F_{\bar{i},\bar{i}'}D_{i,i'}] = \delta_{\bar{i},i}\delta_{\bar{i}',i'}, \tag{2}$$

namely,

$$D_{i,i'} := \frac{|a_i\rangle\langle a'_{i'}|}{\langle a'_{i'}|a_i\rangle}. \tag{3}$$

The frame and the dual frame operators satisfy

$$\mathrm{Tr}[D_{i,i'}] = 1, \tag{4}$$

$$\mathrm{Tr}[F_{i,i'}] = |\langle a'_{i'}|a_i\rangle|^2. \tag{5}$$

Using these, we construct the representations of state ρ and POVM element E, respectively, as

$$\mu(i, i'|\rho) := \mathrm{Tr}[F_{i,i'}\rho] \tag{6}$$

and

$$\xi(E|i, i') := \mathrm{Tr}[ED_{i,i'}]. \tag{7}$$

Given a channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$, one defines a frame $F_{j,j'}$ and its dual $D_{j,j'}$ for $\mathcal{H}_B$ in the same manner as above (but relative to two bases for $\mathcal{H}_B$, denoted $\{|b_j\rangle\}_j$ and $\{|b'_{j'}\rangle\}_{j'}$, satisfying $\langle b_j|b'_{j'}\rangle \neq 0, \forall j, j'$), in addition to those introduced above for system $\mathcal{H}_A$, with which the channel is represented by

$$\Gamma(j, j'|i, i', \mathcal{E}) := \mathrm{Tr}[F_{j,j'}\mathcal{E}(D_{i,i'})]. \tag{8}$$

It is often useful to interpret $\Gamma(j, j'|i, i', \mathcal{E})$ as a matrix mapping vectors indexed by (i, i') to vectors indexed by (j, j'). Composition within the representation is then given by matrix multiplication. The family of all KD representations is generated by ranging over possible basis choices for each Hilbert space in question. Note that while we use the suggestive notation typically used in the literature on ontological models for epistemic states, response functions, and stochastic maps, these are more general, as they are complex-valued and not bounded by 0 or 1.

Such representations are faithful (injective).[1] That is, given some representation μ, ξ, or Γ, one can reconstruct the quantum processes ρ, E, or $\mathcal{E}(\cdot)$, respectively, by

$$\rho = \sum_{i,i'} \mu(i, i'|\rho)D_{i,i'}; \tag{9}$$

$$E = \sum_{i,i'} \xi(E|i, i')F_{i,i'}; \tag{10}$$

$$\mathcal{E}(\cdot) = \sum_{j,j',i,i'} \Gamma(j, j'|i, i', \mathcal{E})\mathrm{Tr}[F_{i,i'} \cdot]D_{j,j'}, \tag{11}$$

as we prove in Appendix A.

To see that the quantum predictions are recovered by the representation, consider an experiment where a channel $\mathcal{E}$ is carried out on state ρ, after which a measurement outcome corresponding to POVM element E is obtained. The representation of this scenario is

$$\sum_{j,j',i,i'} \xi(E|j, j')\Gamma(j, j'|i, i', \mathcal{E})\mu(i, i'|\rho)$$

$$= \sum_{j,j',i,i'} \xi(E|j, j')\mathrm{Tr}[F_{j,j'}\mathcal{E}(D_{i,i'})]\mu(i, i'|\rho)$$

$$= \mathrm{Tr}\left[\left(\sum_{j,j'} \xi(E|j, j')F_{j,j'}\right)\mathcal{E}\left(\sum_{i,i'} \mu(i, i'|\rho)D_{i,i'}\right)\right]$$

$$= \mathrm{Tr}[E\mathcal{E}(\rho)], \tag{12}$$

where we used the linearity of the trace operation and of the channel for the second equality, and used Eq. (9) and Eq. (10) for the last equality.

---

[1]One could consider representations that are not faithful, but such representations generally lose information; moreover, faithfulness is necessary for reproducing the experimental predictions arising in scenarios where the states and measurement operators span the space $\mathcal{B}(\mathcal{H})$.

The representation of the identity channel $\mathcal{I}$ is given by Eq. (8) to be the identity matrix:

$$\Gamma(\bar{i}, \bar{i}'|i, i', \mathcal{I}) = \text{Tr}\big[F_{\bar{i},\bar{i}'}\mathcal{I}(D_{i,i'})\big]$$
$$= \text{Tr}\big[F_{\bar{i},\bar{i}'}D_{i,i'}\big] = \delta_{\bar{i},i}\delta_{\bar{i}',i'}, \qquad (13)$$

as per Eq. (2). Finally, the representation of the identity POVM element is given by Eq. (7) to be the all-ones vector: for all i, i',

$$\xi(\mathbb{1}|i, i') = \text{Tr}[\mathbb{1}D_{i,i'}] = \frac{\text{Tr}[|a_i\rangle\langle a_{i'}'|]}{\langle a_{i'}'|a_i\rangle} = 1. \qquad (14)$$

This condition is the reason it is meaningful to think of these representations as quasistochastic (i.e., where the sum of the components in each column of one's matrix is less than or equal to one), as opposed to some arbitrary complex linear representation. To see this, note that Eq. (14) implies that the representation of every trace-preserving channel $\mathcal{E}$ satisfies

$$\sum_{j,j'} \Gamma(j, j'|i, i', \mathcal{E}) = 1 \quad \forall i, i' \qquad (15)$$

and, for the special case of a state ρ, satisfies

$$\sum_{i,i'} \mu(i, i'|\rho) = 1. \qquad (16)$$

The proof is simple, and given in Appendix A.

Our choice of normalization factors for the frame in Eq. (1) might have seemed arbitrary, but in fact it is fixed, since [by Eq. (2)] our choice fixes the normalization for the dual, which in turn is fixed by Eq. (14). This is one reason why it must be the case that $\langle a_{i'}'|a_i\rangle \neq 0$ for all i and i'. (See Appendix A for more.) Note in particular that this rules out taking the bases $\{|a_i\rangle\}_i$ and $\{|a_{i'}'\rangle\}_{i'}$ to be the same basis, which would lead to the standard density matrix representation of quantum states [for which Eq. (14) is not satisfied].

A. Multiple systems

For composite systems, we take the operator basis to be the tensor product of the bases of the space of operators on each Hilbert space in the composite. For example, for the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, we pick any two bases $\{|a_i\rangle\}_i$ and $\{|a_{i'}'\rangle\}_{i'}$ satisfying $\langle a_i|a_{i'}'\rangle \neq 0, \forall i, i'$ for $\mathcal{H}_A$ and any two bases $\{|b_j\rangle\}_j$ and $\{|b_{j'}'\rangle\}_{j'}$ satisfying $\langle b_j|b_{j'}'\rangle \neq 0, \forall j, j'$ for $\mathcal{H}_B$. We introduce the frame

$$F_{i,i';j,j'} := |a_{i'}'\rangle\langle a_i| \otimes |b_{j'}'\rangle\langle b_j|\langle a_{i'}'|a_i\rangle\langle b_{j'}'|b_j\rangle$$
$$= F_{i,i'} \otimes F_{j,j'} \qquad (17)$$

and the dual frame

$$D_{i,i';j,j'} := \frac{|a_i\rangle\langle a_{i'}'| \otimes |b_j\rangle\langle b_{j'}'|}{\langle a_{i'}'|a_i\rangle\langle b_{j'}'|b_j\rangle}$$
$$= D_{i,i'} \otimes D_{j,j'} \qquad (18)$$

and define the representation of quantum processes exactly as was done for single systems.

B. Recovering standard Kirkwood-Dirac distributions for states

For the special case of quantum states, our representation recovers the standard KD distribution [7,25,28], since

$$\mu(i, i'|\rho) := \text{Tr}[F_{i,i'}\rho] = \langle a_i|\rho|a_{i'}'\rangle\langle a_{i'}'|a_i\rangle.$$

Consequently, our representation satisfies all known properties of preexisting KD representations. For instance, for some values of i, i', and some ρ we can have $\text{Im}[\mu(i, i'|\rho)] \neq 0$, since our frame operators are not necessarily Hermitian, unlike other frame representations usually considered in the literature [1,2,30–32]. As another example, the distribution associated to any state is bounded by one and has maximal negativity and imaginarity given by $\min_{\{a_i\}_i, \{a_{i'}'\}_{i'}, \rho} \text{Re}[\mu(i, i'|\rho)] = -1/8$ and $\max_{\{a_i\}_i, \{a_{i'}'\}_{i'}, \rho} \text{Im}[\mu(i, i'|\rho)] = 1/4$ [84]; see Appendix A for more. Consequently, existing techniques to witness KD negativity or imaginarity, or that investigate the geometry of real non-negatively represented states with respect to fixed KD bases [23,24,28,85–87] also remain applicable. Our work opens the possibility of extending these investigations beyond quantum states.

C. Connections with weak values and tomography

While $\mu(i, i'|\rho)$ is bounded, the representation $\xi(E|i, i')$ is not, possibly having arbitrarily large positive, negative, or imaginary values. This is immediate from noticing that, for any dual frame D given by Eq. (3), each value $\xi(E|i, i')$ is exactly what is known as the weak value of E with respect to the vectors $|a_i\rangle$ and $|a_{i'}'\rangle$ [35,36]. Indeed, KD distributions for states and these particular weak values must be dual representations of states and POVM elements in our representation, as the dual basis is uniquely singled out by Eq. (2).

Prior work in Ref. [88] also found this connection between weak values and KD representations of states (although this work did not establish that this connection is necessary in the sense mentioned in the previous paragraph).

A similar connection is discussed for instance in Refs. [15,25]. There, the connection between weak values and KD distributions is made via defining the so called conditional KD distribution $\mu(i, i'|\rho)/\sum_i \mu(i, i'|\rho)$. In the case where $\rho = |\psi\rangle\langle\psi|$ is a pure state, one has

$$\mu(i, i'|\rho) = \langle a_i|\psi\rangle\langle\psi|a_{i'}'\rangle\langle a_{i'}'|a_i\rangle = |\langle\psi|a_{i'}'\rangle|^2 \frac{\langle a_{i'}'|a_i\rangle\langle a_i|\psi\rangle}{\langle a_{i'}'|\psi\rangle}.$$

Thus, the conditional KD distribution is equivalent to the weak value of the rank-1 projections $|a_i\rangle\langle a_i|$ with respect to $|a_{i'}'\rangle$ and the pure state $|\psi\rangle$. In contrast, in our work, weak values are obtained directly through the representation of measurements, and this procedure is valid for any quantum state, any choice of bases $\{|a_i\rangle\}_i$ and $\{|a_{i'}'\rangle\}_{i'}$, and any POVMs.

The KD representation Γ of a channel $\mathcal{E}$, given by Eq. (8), can also have arbitrarily large absolute values due to the inner products $\langle a_{i'}'|a_i\rangle$ in the denominator. Still, the quasistochastic behavior from Eq. (15) continues to hold. A natural question is how these representations relate to physical quantities of interest in the same way that those for states have been shown to, or as we have just argued that those for POVM elements relate to weak values.

As a final remark, since our representation is faithful and so contains complete information about one's quantum process, any estimation of our KD representation for a state, channel, or measurement is a form of state tomography, process tomography, or measurement tomography. For any process, the representation can be written as a multivariate trace, and so can be estimated using polynomial [89,90] or even constant [91] depth quantum circuits.

D. Compositionality and functoriality

A useful property for any representation to have is functoriality, or more specifically, that the representation is a symmetric strict monoidal functor [29] (sometimes called a diagram-preserving map [30]), which means that the representation commutes with composition of processes. Functoriality also requires that identity and swap operations are represented by the identity and the swap, respectively, in the representation. Functoriality is arguably essential if that representation is to carry physical meaning (e.g., to provide a physical explanation) rather than merely be a mathematical simulation. Most representations studied to date are functorial, although some [92–94] are not [30]. In particular, those based on minimal (not overcomplete) and factorizing frame representations are functorial; those based on overcomplete frames will not represent the identity transformation as the identity, while those based on nonfactorizing frames [i.e., those not satisfying Eqs. (17) and (18)] will not represent the swap as the swap.

As we have defined it, every KD representation is functorial. We prove this in Appendix C, by showing that the representation of a composite process (whether composed in parallel or in sequence) is the same as the composite of the representations of each component of it, and by proving that the identity and swap operations are represented correctly. For example, the representation of two channels composed in sequence, e.g., $\mathcal{E}_2 \circ \mathcal{E}_1$, must be the composition (matrix product) of the representations of each individually, namely

$$\Gamma(k, k'|i, i', \mathcal{E}_2 \circ \mathcal{E}_1) = \sum_{j,j'} \Gamma(k, k'|j, j', \mathcal{E}_1)\Gamma(j, j'|i, i', \mathcal{E}_2).$$

(19)

III. CONNECTION TO GENERALIZED NONCONTEXTUALITY

With Kirkwood-Dirac representations for all of quantum theory in hand, it is straightforward to establish connections with noncontextuality. Consider any fragment of quantum theory. In the simplest case, this means a set $\{\rho_i\}_i$ of states, a set $\{\mathcal{E}_j\}_j$ of channels, and a set $\{\{E_{[k|m]}\}_k\}_m$ of measurements. Or, one may have some fixed scenario or circuit (such as a particular Bell scenario) and a given set of states, channels and effects that can be realized in that scenario or circuit (such as different choices of entangled states and measurements in a Bell scenario). These sets may arise from the particular capabilities of some particular hardware, from a specific experiment, or from some theoretical considerations. They may (but need not) have further structure, such as being finite or convexly closed.

An ontological model for a fragment of quantum theory is any linear and functorial mapping from the quantum processes

to (the monoidal category of) substochastic matrices [30], such that the trace operation is mapped to marginalization (i.e., the all-ones vector). See Ref. [95] for an introduction to ontological models, and Ref. [30] for the first fully precise and compositional definition. Many fragments of quantum theory do not admit of any ontological model, but some fragments do—even very rich fragments such as the stabilizer subtheory in all odd dimensions [4,63].

Under the formulation we have introduced herein, any KD representation that happens to be real and non-negative for all processes in a given fragment also satisfies all of these properties, and so constitutes an ontological model of that fragment. Clearly the KD representation in this case is a mapping from quantum processes to non-negative real-valued matrices; as we prove in Appendix B, these matrices are moreover substochastic. And as we saw in Eqs. (13) and (14) (respectively), the identity is represented by the identity matrix and the representation of the trace operation is equal to marginalization.

The connection to generalized noncontextuality now follows immediately from the results of Refs. [30,38], which show that the predictions of a fragment of quantum theory can be reproduced by an ontological model if and only if any laboratory procedures whose quantum description is given by this fragment admits of a generalized noncontextual ontological model.[2] In other words, a scenario is consistent with generalized noncontextuality if and only if its description within standard quantum theory admits of an ontological model. But every real and non-negative KD representation constitutes an ontological model, so the existence of any real and non-negative KD representation implies noncontextuality.

A. Negativity or imaginarity of a KD representation is not sufficient for nonclassicality

However, negativity or imaginarity of a KD representation does not necessarily imply the failure of noncontextuality. Most obviously, there may be another KD representation that manages to represent all processes in the scenario or theory with real and non-negative numbers, in which case this other representation constitutes a perfectly fine classical model, consistent with noncontextuality.

Moreover, even if one shows that for a given scenario or theory, all KD representations (as defined herein) require negativity or imaginarity, it is still possible that an ontological model of an entirely different form is possible.

Indeed, this sometimes happens, as we prove by example. Consider the stabilizer subtheory of quantum mechanics in odd dimensions, which has a unique [63] ontological model, namely that given by Gross and by Spekkens [4,97]. This model is not of the form of any Kirkwood-Dirac representation, as one can see by the fact that the frame operators defining it are Hermitian, while those of a KD representation are not. But by uniqueness, there can be no other ontological

---

[2]Formally, by laboratory procedures we mean processes within an unquotiented operational theory, while by the quantum description of the fragment, we mean a quotiented operational theory (with density operators, quantum channels, and the like) [30,96].

model of it, so all KD representations of it must have negativity or imaginarity, even though the subtheory is noncontextual.

That said, there is a natural (although rather broad) generalization of the family of KD representations considered here and in the literature for which one can make a tighter connection with noncontextuality. This is the natural class of linear, functorial, and empirically adequate quasistochastic representations of quantum processes as complex-valued functions: essentially, it is what one obtains if one relaxes the frame in Eq. (1) to be an arbitrary nonovercomplete frame (basis) for the space of all operators on $\mathcal{H}$. As we show in future work, one can prove the failure of noncontextuality by proving that every representation in this considerably broader class has negativity or imaginarity, as a natural extension of analogous arguments for real-valued functions in Ref. [30].

A final note is that for any given state $|\psi\rangle$, one can always choose an appropriate frame such that the KD representation of that state exhibits negativity or imaginarity. Consequently, not only is negativity or imaginarity insufficient to establish the failure of noncontextuality, it also is no basis for any reasonable notion of nonclassicality.

B. A fragment of quantum theory with a non-negative and real KD representation

Let us illustrate some simple aspects of our findings with an example: a nontrivial fragment of quantum theory that admits of a real and non-negative KD representation for its states, measurements, and transformations.

The states and measurement effects in the fragment are taken to be the operators in ConvHull($\{|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |-\rangle\langle -|\}$), and the transformations are taken to be $\{\mathrm{id}, \mathcal{D}_{\nu,Z}^{\mathrm{deph}}, \mathcal{D}_{\nu,X}^{\mathrm{deph}}\}$, where id is the identity channel, and $\mathcal{D}_{\nu,Z}^{\mathrm{deph}}, \mathcal{D}_{\nu,X}^{\mathrm{deph}}$ are dephasing channels with respect to the bases $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$, respectively. This fragment has a real and non-negative KD representation of any process if we consider the frames (and dual frames) defined with respect to the two mutually unbiased bases $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$.

It is simple to see that the representation μ of these quantum states and the representation ξ of these quantum effects take values that are always real and non-negative. Moreover $\xi \in [0, 1]$ since

$$\xi(E|i, i') = \frac{\mathrm{Tr}(|i\rangle\langle i|i'\rangle\langle i'|E)}{|\langle i|i'\rangle|^2}$$
$$= \frac{1}{2}\sum_{j=0,1,-,+} \alpha_j \mathrm{Tr}(\Pi_i \Pi_{i'} \Pi_j) \leqslant \frac{1}{2}\sum_j \alpha_j = \frac{1}{2},$$

where $E = \sum_{j=0,1,-,+} \alpha_j \Pi_j$ with $\Pi_j \equiv |j\rangle\langle j|$, $\sum_j \alpha_j = 1$, $0 \leqslant \alpha_j \leqslant 1$ for every j. We have used that for every i, i' fixed and every j, $\mathrm{Tr}(\Pi_i \Pi_{i'} \Pi_j) \in [0, 1]$, since Π_j either commutes with $\Pi_{i'}$ or with Π_i. For the representation of the channels, we first write them as

$$\mathcal{D}_{\nu,Z}^{\mathrm{deph}}(M) = (1-\nu)M + \nu(\langle 0|M|0\rangle\Pi_0 + \langle 1|M|1\rangle\Pi_1);$$

$$\mathcal{D}_{\nu,X}^{\mathrm{deph}}(M) = (1-\nu)M + \nu(\langle +|M|+\rangle\Pi_+ + \langle -|M|-\rangle\Pi_-),$$

where M is a handle for a generic operator. Then, specifying for $D_{i,i'}$ we get

$$\mathcal{D}_{\nu,Z}^{\mathrm{deph}}(D_{i,i'}) = (1-\nu)D_{i,i'} + \nu(\langle 0|D_{i,i'}|0\rangle\Pi_0 + \langle 1|D_{i,i'}|1\rangle\Pi_1),$$

and using Eq. (8) we get

$$\Gamma(j, j'|i, i', \mathcal{D}_{\nu,Z}^{\mathrm{deph}}) = (1-\nu)\delta_{i,j}\delta_{i',j'} + \nu(\xi(\Pi_0|i, i') \times \mu(j, j'|\rho_0) + \xi(\Pi_1|i, i')\mu(j, j'|\rho_1)),$$

where $\rho_0 \equiv \Pi_0 = |0\rangle\langle 0|$ and $\rho_1 \equiv \Pi_1 = |1\rangle\langle 1|$. Since ξ, μ in these specific ranges are valid classical response functions and probability distributions, Γ is just a stochastic matrix. The same analysis holds for $\Gamma(j, j'|i, i', \mathcal{D}_{\nu,X}^{\mathrm{deph}})$, where one changes the projectors accordingly. Consequently, and from the fact that every KD representation is functorial, this representation serves as an ontological model for this fragment.

Of course, other representations exist that are not real and non-negative. For instance, choosing a KD representation with respect to $\{|0\rangle, |1\rangle\}$ and $\{|+i\rangle, |-i\rangle\}$ would not lead to a real non-negative representation, showcasing that this property is dependent on the choice of frame.

IV. CONCLUSIONS

We have shown how to generalize standard KD representations to any quantum process, including measurements and channels. The representation satisfies key properties such as respecting composition and linearity. This allows us to clarify the relationship between KD distributions, ontological models, and noncontextuality. It also leads to new avenues of study, such as the connection of the KD representation of POVM elements with weak values.

Let us now compare our findings with previous literature. It is fairly common knowledge that one can represent generic Hermitian observables using frame operators $D_{i,i'}$ and recover the information of any such observable (when the choice of frame is faithful) using Eq. (9), substituting ρ by any Hermitian observable [7,25,88,98]. Notably, Ref. [88] has not only considered representations of generic observables and made a connection with weak values equivalent to the one we have drawn here, but also investigated how one's representation changes when moving from a frame $F_{i,i'}$ to another frame $\tilde{F}_{\tilde{i},\tilde{i}'}$. In our view, these were all important previous steps converging to the concrete, organized, and comprehensive notion of a (functorial) KD representation, that our work introduces.

Unlike previous works, our work does not focus on distributions representing a given kind of process. It rather focuses on giving a full-fledged representation of quantum theory as a whole theory. In other words, rather than merely representing quantum states, measurements, and transformations, our aim is to provide a fully compositional representation, which moreover allows one to reproduce the empirical predictions of quantum theory. Or, when applied to an experiment or fragment of quantum theory, our work asks the reader not only to represent individual components of the experiment, but rather to provide a consistent and complete representation of every process in the experiment (or fragment), such that taken together the quantum predictions are reproduced. This constraint is encoded in Eq. (12), which in some sense crystallizes the novelty of our work. Although not every ap-

plication of KD distributions requires providing a complete and consistent representation of a complete theory or of every process in a given experiment, many applications do—most notably the important question of what quantum phenomena are nonclassical. This point has been argued before, e.g., in Ref. [37].

We have focused on finite-dimensional Hilbert spaces, resulting in finite (hence discrete) phase space representations. In the case of continuous phase space representations and infinite-dimensional Hilbert spaces, the KD distribution—also known as the Kirkwood-Rihaczek distribution—has been shown to be related to the Wigner distribution [99,100]. One can then presumably infer the form of KD representations using their connections with the Wigner representation. To the best of our knowledge, this has not been studied, but it would be interesting for future work. Note that our results do not follow immediately from the connection between KD representations and the Wigner representation. For one thing, there are many definitions of discrete Wigner distributions in finite-dimensional systems; for another, functoriality of a representation is not granted by its connection with the Wigner distribution (as it is clear from the similar connections between the nonfunctorial P and Q- representations and the functorial Wigner representation). However, we do suspect that the functoriality of the KD representation we derived in the finite-dimensional case will extend to the infinite-dimensional case.

Other extensions of KD distributions for quantum states have been proposed [9,12,25]. However, we are not aware of a fully general definition of a representation of quantum theory (or even of quantum states) in complex-valued functions.

In forthcoming work, we extend the present work to define a very general class of representations and discuss their relation with some other generalized KD representations one can find in the literature. We then show that every ontological model of a quantum scenario has the form of some such representation that is real and non-negative on all processes in the scenario, which implies that one can prove nonclassicality (the failure of generalized noncontextuality) if one can prove that no such representation is real and non-negative. We also show that these general representations and results (which subsume all the ones in this paper) apply to arbitrary generalized probabilistic theories [101,102], not just quantum theory. Finally, we note that working with representations such as these is typically dramatically simpler when one uses diagrammatic notation, as we do in the forthcoming work, and as is done in Ref. [30].

ACKNOWLEDGMENTS

APPENDIX A: USEFUL MATHEMATICAL FACTS

We now present some useful mathematical facts regarding the family of Kirkwood-Dirac representations that we have defined. In the main text, we showed how arbitrary quantum processes could be represented and then reconstructed. In fact, arbitrary operators on a Hilbert space can be represented and reconstructed in an exactly analogous fashion. There are two ways to do so, depending on whether the representation or the reconstruction is done using the frame (as opposed to the dual). One can define the representation of an arbitrary operator O acting on Hilbert space $\mathcal{H}_A$ as

$$\mu(i, i'|O) := \text{Tr}[F_{i,i'}O], \tag{A1}$$

$$= \langle a_i|O|a'_{i'}\rangle\langle a'_{i'}|a_i\rangle, \tag{A2}$$

where Eq. (A2) has the form of standard KD representations commonly seen in the literature. In this case, one reconstructs O via

$$O = \sum_{i,i'} \mu(i, i'|O)D_{i,i'}. \tag{A3}$$

The above was already noticed in, for example, Ref. [25], Eq. (19). Alternatively, one can represent O as

$$\xi(O|i, i') := \text{Tr}[OD_{i,i'}] \tag{A4}$$

$$= \frac{\langle a'_{i'}|O|a_i\rangle}{\langle a'_{i'}|a_i\rangle}; \tag{A5}$$

in which case the reconstruction equation is

$$O = \sum_{i,i'} \xi(O|i, i')F_{i,i'}. \tag{A6}$$

Here we are using the convention that representations in terms of the frame are denoted by μ, while those in terms of the dual frame are represented by ξ, consistent with the notation in the main text.

We now prove the validity of the reconstruction equations Eqs. 9 to 11 in the main text, of which the proof for Eq. (9) and Eq. (10) also works for the respective reconstruction equations for a general operator, namely, Eqs. (A3) and (A6).

For Eq. (9), i.e., $\rho = \sum_{i,i'} \mu(i, i'|\rho)D_{i,i'}$:

$$\sum_{i,i'} \mu(i, i'|\rho)D_{i,i'} = \sum_{i,i'} \langle a_i|\rho|a'_{i'}\rangle\langle a'_{i'}|a_i\rangle \frac{|a_i\rangle\langle a'_{i'}|}{\langle a'_{i'}|a_i\rangle}$$

$$= \sum_{i,i'} |a_i\rangle\langle a_i|\rho|a'_{i'}\rangle\langle a'_{i'}| = \rho. \tag{A7}$$

When we replace ρ with a general operator O, Eq. (A7) becomes a proof for Eq. (A3).

For Eq. (10), i.e., $E = \sum_{i,i'} \xi(E|i,i')F_{i,i'}$:

$$\sum_{i,i'} \xi(E|i,i')F_{i,i'} = \sum_{i,i'} \frac{\langle a'_{i'}|E|a_i\rangle}{\langle a'_{i'}|a_i\rangle} |a'_{i'}\rangle\langle a_i|\langle a'_{i'}|a_i\rangle$$

$$= \sum_{i,i'} |a'_{i'}\rangle\langle a'_{i'}|E|a_i\rangle\langle a_i| = E. \quad \text{(A8)}$$

When we replace E with a general operator O, Eq. (A8) becomes a proof for Eq. (A6).

To prove Eq. (11), namely that

$$\mathcal{E}(\cdot) = \sum_{j,j',i,i'} \Gamma(j,j'|i,i',\mathcal{E})\text{Tr}[F_{i,i'} \cdot]D_{j,j'},$$

we have

$$\sum_{j,j',i,i'} \Gamma(j,j'|i,i',\mathcal{E})\text{Tr}[F_{i,i'} \cdot]D_{j,j'}$$

$$= \sum_{j,j',i,i'} \text{Tr}[F_{j,j'}\mathcal{E}(D_{i,i'})]\mu(i,i'|\cdot)D_{j,j'}$$

$$= \sum_{j,j'} \text{Tr}\left[F_{j,j'}\mathcal{E}\left(\sum_{i,i'}\mu(i,i'|\cdot)D_{i,i'}\right)\right]D_{j,j'}$$

$$= \sum_{j,j'} \text{Tr}[F_{j,j'}\mathcal{E}(\cdot)]D_{j,j'} = \sum_{j,j'} \mu(j,j'|\mathcal{E}(\cdot))D_{j,j'} = \mathcal{E}(\cdot),$$
$$\text{(A9)}$$

where we used the linearity of trace and channel at the second equal sign, and used the reconstruction equations at the third and fifth equal sign. One relation that will be useful in the following proofs is that $\sum_{i,i'} F_{i,i'}$ preserves the trace of any operator. That is,

$$\sum_{i,i'} \text{Tr}[F_{i,i'}O] = \text{Tr}[O]. \quad \text{(A10)}$$

This can be seen by explicit calculation:

$$\sum_{i,i'} \text{Tr}[F_{i,i'}O] = \sum_{i,i'} \text{Tr}[\langle a'_{i'}|a_i\rangle|a'_{i'}\rangle\langle a_i|O]$$

$$= \sum_{i,i'} \text{Tr}[|a'_{i'}\rangle\langle a'_{i'}|a_i\rangle\langle a_i|O]$$

$$= \text{Tr}\left[\left(\sum_{i'}|a'_{i'}\rangle\langle a'_{i'}|\right)\left(\sum_i|a_i\rangle\langle a_i|\right)O\right]$$

$$= \text{Tr}[O]. \quad \text{(A11)}$$

In particular, this immediately implies that the KD representation of an arbitrary quantum state (or indeed of an arbitrary trace-1 operator when using the frame rather than the dual representation) satisfies the normalization property in Eq. (16) of the main text,

$$\sum_{i,i'} \mu(i,i'|\rho) = 1. \quad \text{(A12)}$$

In addition, the representation of a trace-preserving channel in every KD representation satisfies Eq. (15) of the main text, namely

$$\sum_{j,j'} \Gamma(j,j'|i,i',\mathcal{E}) = 1 \quad \forall i,i',\mathcal{E} \quad \text{(A13)}$$

This follows from

$$\sum_{j,j'} \Gamma(j,j'|i,i',\mathcal{E}) = \sum_{j,j'} \text{Tr}[F_{j,j'}\mathcal{E}(D_{i,i'})]$$

$$= \text{Tr}[\mathcal{E}(D_{i,i'})] = \text{Tr}[\mathbb{1}\mathcal{E}(D_{i,i'})]$$

$$= \text{Tr}[\mathcal{E}^\dagger(\mathbb{1})D_{i,i'}] = \text{Tr}[D_{i,i'}] = 1, \quad \text{(A14)}$$

where in the second equality we used Eq. (A10), in the penultimate step we used that $\mathcal{E}$ is trace preserving, which implies its adjoint $\mathcal{E}^\dagger$ is unital, and in the last step we used Eq. (4).

A final fact one might note is that

$$|\mu(i,i'|\rho)| \in [0,1]. \quad \text{(A15)}$$

This is because from Eq. (A2), we have $|\mu(i,i'|\rho)| = |\langle a_i|\rho|a'_{i'}\rangle||\langle a'_{i'}|a_i\rangle|$. The first term in the product satisfies $|\langle a_i|\rho|a'_{i'}\rangle| \leqslant 1$, since the spectral decomposition $\rho = \sum_k p_k|\psi_k\rangle\langle\psi_k|$ gives $|\langle a_i|\rho|a'_{i'}\rangle| = \sum_k p_k|\langle a_i|\psi_k\rangle||\langle\psi_k|b_j\rangle| \leqslant \sum_k p_k = 1$. Then, since $|\langle a'_{i'}|a_i\rangle| \in [0,1]$, we get $|\mu(i,i'|\rho)| \in [0,1]$.

In fact, we can slightly generalize the results from Ref. [84] from pure states to mixed states, to note that any complex number $\mu(i,i'|\rho) = |\Delta|e^{i\phi}$ must satisfy that

$$1 - 3|\Delta|^{\frac{2}{3}} + 2|\Delta|\cos(\phi) \geqslant 0, \quad \text{(A16)}$$

for all possible $|a_i\rangle$, $|a'_{i'}\rangle$ and state ρ, with respect to any Hilbert space $\mathcal{H}_A$. This is true since $\mu(i,i'|\rho) = \text{Tr}[F_{i,i'}\rho]$ is linear with respect to ρ. Therefore, for any convex combination of pure states $\rho = \sum_\lambda \alpha_\lambda|\psi_\lambda\rangle\langle\psi_\lambda|$, we have that $\mu(i,i'|\rho)$ is also a convex combination of pure state KD phase space points $\mu(i,i'|\psi_\lambda)$. Reference [84] showed that any $\mu(i,i'|\psi_\lambda)$ must be in the region described by Eq. (A16), but since this region forms a convex set of points in $\mathbb{C}$, any convex combination of terms $\mu(i,i'|\psi_\lambda)$ must also lie inside this region. This shows that the phases and absolute values of the numbers $\mu(i,i'|\psi_\lambda)$ satisfy nontrivial constraints—constraints, which moreover relate to relevant physical facts. For example, maximal negativity is relevant for work extraction beyond classical thermodynamic limits [16,17]. Equation (A16) implies that the maximal negative and imaginary values any KD distribution can attain are $\min_{\{a_i\}_i,\{a'_{i'}\}_{i'},\rho} \text{Re}[\mu(i,i'|\rho)] = -1/8$ and $\max_{\{a_i\}_i,\{a'_{i'}\}_{i'},\rho} \text{Im}[\mu(i,i'|\rho)] = 1/4$ [84].

Finally, we comment on our choice to only define KD representations using bases $\{|a_i\rangle\}_i$ and $\{|a'_{i'}\rangle\}_{i'}$ for the Hilbert space such that $\langle a_i|a'_{i'}\rangle \neq 0$ for all i and i'. This choice is fairly standard—but not universal—in the literature. Making this choice ensures that the representation we construct is faithful. As also pointed out in the main text, only in this case does learning a KD representation of a quantum process imply that one can reconstruct the given process, in which case ways of measuring the KD representation are simply ways of carrying out tomography (as recognized for the case of states in Ref. [98]). Additionally, this property implies that the representation of the identity operator is the all-ones vector [Eq. (14)], which is crucial for interpreting these KD distributions as quasistochastic in general—i.e., for normalization to be given by summing up all elements of the state.

1. Relationship with frame representations and real-valued quasiprobability representations

The terms quasiprobability representation and frame-representation sometimes refer only to real-valued representations of quantum theory, but are sometimes rather used to subsume complex-valued representations (such as KD representations). Similarly, the term frame is sometimes used specifically for bases for the space of Hermitian operators [1,30–32]; however, the origins of the term frame allow for arbitrary bases of more general inner product spaces [33], so one can use these terms in the context of complex-valued KD representations, as we have herein.

Frame representations are strictly more general than KD representations (as we have defined them).[3] For one thing, KD representations use complete bases, while frame representations may use either complete or overcomplete frames (although the latter do not represent the identity as the quasistochastic identity map [30]). For another, the operator basis used in a standard KD representation is constructed to have a particular form [that in Eq. (1)], whereas any basis of operators may be used in a frame representation; for instance, the operators in the former case (but not in the latter) must be non-Hermitian and rank one.

APPENDIX B: PROVING STOCHASTICITY FOR ALL REAL NON-NEGATIVE KD REPRESENTATIONS

We now prove that if a KD representation is real and non-negative for some set of processes, then in fact it represents all those processes as substochastic matrices (and consequently, the representation constitutes a valid ontological model for the processes).

We begin by showing that the representation of any state is a valid probability distribution in this case. By assumption, $\mu(i, i'|\rho) \geqslant 0$ for all ρ under consideration, and by Eq. (A12) we have that $\sum_{i,i'} \mu(i, i'|\rho) = 1$. So $\mu(i, i'|\rho)$ is a valid probability distribution.

Next, we wish to show that a POVM element is in this case represented by a valid response function: that is, a real vector whose elements are between 0 and 1. By assumption, $\xi(E|i, i') \geqslant 0$ for all E under consideration. Moreover, a POVM element E can only arise in a theory or fragment of quantum theory as part of a measurement—a set of effects that sums to the identity effect $\mathbb{1}$. Consequently, the POVM element $\mathbb{1} - E$ is necessarily also in the scenario or fragment (one can measure this effect by doing any measurement containing POVM element E and coarse graining all the outcomes other than E). So (again by the assumption of real non-negativity) we have $\xi(E|i, i') \geqslant 0$ and $\xi(\mathbb{1} - E|i, i') \geqslant 0$ for every E under consideration. By linearity of the representation, the latter implies that $\xi(\mathbb{1}|i, i') - \xi(E|i, i') \geqslant 0$. By Eq. (14), i.e., $\xi(\mathbb{1}|i, i') = 1$, we further have $1 - \xi(E|i, i') \geqslant 0$ and so

$$0 \leqslant \xi(E|i, i') \leqslant 1 \tag{B1}$$

for all i, i' and for all E under consideration, as required.

---

[3]Note however that some distributions under the KD moniker are not faithful (see, e.g., Ref. [23], Appendix A), and so are not frame representations.

Finally, we show that under the assumption of real non-negativity, any KD representation of a quantum channel is a substochastic matrix. For quantum channels (which are trace preserving), this is immediately evident from the assumption of real non-negativity together with Eq. (A13), namely $\sum_{j,j'} \Gamma(j, j'|i, i', \mathcal{E}) = 1$ for all i, i'. If $\mathcal{E}$ is a trace-decreasing quantum operation, the argument is much the same as for effects: $\mathcal{E}$ can only arise in a scenario where it is complemented by some other complementary trace-decreasing quantum channel $\tilde{\mathcal{E}}$ such that $\mathcal{E} + \tilde{\mathcal{E}}$ is trace preserving (so that each corresponds to the selective update rule for two possible outcomes of some quantum instrument). By linearity, one has

$$1 = \sum_{j,j'} \Gamma(j, j'|i, i', \mathcal{E} + \tilde{\mathcal{E}}) = \sum_{j,j'} \Gamma(j, j'|i, i', \mathcal{E})$$
$$+ \sum_{j,j'} \Gamma(j, j'|i, i', \tilde{\mathcal{E}}) \tag{B2}$$

for all i, i'. By assumption, all the terms in these sums are real non-negative, so none of them can be larger than 1. So for a trace-decreasing operation $\mathcal{E}$, one has

$$\sum_{j,j'} \Gamma(j, j'|i, i', \mathcal{E}) \leqslant 1, \tag{B3}$$

as required for the representation to be substochastic.

APPENDIX C: PROOF OF FUNCTORIALITY

We now prove that every KD representation, as we have defined it, commutes with composition of quantum processes—formally, that it is a symmetric strict monoidal functor, or is diagram preserving.

First, we prove that the quantum identity is represented as the identity on phase space:

$$\Gamma(\bar{i}, \bar{i}'|i, i', \mathcal{I}) = \text{Tr}[F_{\bar{i}, \bar{i}'} \mathcal{I}(D_{i,i'})]$$
$$= \text{Tr}[F_{\bar{i}, \bar{i}'} D_{i,i'}] = \delta_{\bar{i}, i} \delta_{\bar{i}', i'}, , \tag{C1}$$

where we have used the duality of the frames, i.e., Eq. (2).

Second, we prove that the KD representation preserves sequential composition, that is, for $\mathcal{E}_1$ transforming system A to system B and $\mathcal{E}_2$ transforming system B to system C, we have

$$\Gamma(k, k'|i, i', \mathcal{E}_2 \circ \mathcal{E}_1) = \sum_{j,j'} \Gamma(k, k'|j, j', \mathcal{E}_2) \Gamma(j, j'|i, i', \mathcal{E}_1). \tag{C2}$$

Expanding the expression in the second line, we have

$$\sum_{j,j'} \Gamma(k, k'|j, j', \mathcal{E}_2) \Gamma(j, j'|i, i', \mathcal{E}_1)$$
$$= \sum_{j,j'} \text{Tr}[F_{k,k'} \mathcal{E}_2(D_{j,j'})] \text{Tr}[F_{j,j'} \mathcal{E}_1(D_{i,i'})]$$
$$= \text{Tr}\left[F_{k,k'} \mathcal{E}_2\left(\sum_{j,j'} \text{Tr}[F_{j,j'} \mathcal{E}_1(D_{i,i'})] D_{j,j'}\right)\right]$$
$$\stackrel{(A3)}{=} \text{Tr}\{F_{k,k'} \mathcal{E}_2[\mathcal{E}_1(D_{i,i'})]\} = \text{Tr}[F_{k,k'} (\mathcal{E}_2 \circ \mathcal{E}_1)(D_{i,i'})]$$
$$= \Gamma(k, k'|i, i', \mathcal{E}_2 \circ \mathcal{E}_1), \tag{C3}$$

where we used the linearity of the trace operation and the channel at the third equal sign, and the definitions of KD representations and the reconstruction equations at various places.

Third, we prove that KD representations preserve parallel composition, that is, for $\mathcal{E}_1$ transforming system A to system C and $\mathcal{E}_2$ transforming system B to system D, we have that the representation of the two channels in parallel, namely $\mathcal{E}_1 \otimes \mathcal{E}_2$, satisfies

$$\Gamma(k, k'|i, i', \mathcal{E}_1)\Gamma(l, l'|j, j', \mathcal{E}_2)$$
$$= \Gamma(k, k'; l, l'|i, i'; j, j', \mathcal{E}_1 \otimes \mathcal{E}_2). \quad (C4)$$

We expand the expression in the second line as

$$\Gamma(k, k'; l, l'|i, i'; j, j', \mathcal{E}_1 \otimes \mathcal{E}_2)$$
$$= \text{Tr}[F_{k,k';l,l'}\mathcal{E}_1 \otimes \mathcal{E}_2(D_{i,i';j,j'})]$$
$$= \text{Tr}[(F_{k,k'} \otimes F_{l,l'})\mathcal{E}_1 \otimes \mathcal{E}_2(D_{i,i'} \otimes D_{j,j'})]$$
$$= \text{Tr}[F_{k,k'}\mathcal{E}_1(D_{i,i'}) \otimes F_{l,l'}\mathcal{E}_1(D_{j,j'})]$$
$$= \text{Tr}[F_{k,k'}\mathcal{E}_1(D_{i,i'})]\text{Tr}[F_{l,l'}\mathcal{E}_1(D_{j,j'})]$$
$$= \Gamma(k, k'|i, i', \mathcal{E}_1)\Gamma(l, l'|j, j', \mathcal{E}_2), \quad (C5)$$

where the second equality follows from the definition of the frame and the dual frame for two systems [namely Eqs. (17) and (18)], and where the second to last equality uses a simple property of the trace, namely that $\text{Tr}[A \otimes B] = \text{Tr}[A]\text{Tr}[B]$.

Last, we prove that the quantum swap channel is represented as the corresponding swap channel in the representation. Consider the swap channel $\mathcal{S}$ such that

$$\mathcal{S}(O_1 \otimes O_2) = O_2 \otimes O_1, \quad (C6)$$

where O_1 and O_2 are any operators, each of which is on a (potentially different) Hilbert space. Note that the order $O_1 \otimes O_2$ is reflected on Γ via the order of the basis indices. We then need to show that in the KD representation of the swap, the order i, i', j, j' of the input indices is swapped in the output to $\bar{j}, \bar{j}', \bar{i}, \bar{i}'$. Mathematically,

$$\Gamma(\bar{j}, \bar{j}'; \bar{i}, \bar{i}'|i, i'; j, j', \mathcal{S}) = \delta_{\bar{i},i}\delta_{\bar{i}',i'}\delta_{\bar{j},j}\delta_{\bar{j}',j'}, \quad (C7)$$

for all indices. We expand the expression in the first line:

$$\Gamma(\bar{j}, \bar{j}'; \bar{i}, \bar{i}'|i, i'; j, j', \mathcal{S})$$
$$= \text{Tr}[F_{\bar{j},\bar{j}';\bar{i},\bar{i}'}\mathcal{S}(D_{i,i';j,j'})]$$
$$= \text{Tr}[F_{\bar{j},\bar{j}'} \otimes F_{\bar{i},\bar{i}'}\mathcal{S}(D_{i,i'} \otimes D_{j,j'})]$$
$$= \text{Tr}[(F_{\bar{j},\bar{j}'} \otimes F_{\bar{i},\bar{i}'})(D_{j,j'} \otimes D_{i,i'})]$$
$$= \text{Tr}[F_{\bar{j},\bar{j}'}D_{j,j'}]\text{Tr}[F_{\bar{i},\bar{i}'}D_{i,i'}]$$
$$= \delta_{\bar{i},i}\delta_{\bar{i}',i'}\delta_{\bar{j},j}\delta_{\bar{j}',j'}, \quad (C8)$$

where the second equality follows from the definition of frame and dual frame for two systems [namely, Eqs. (17) and (18)].

These facts together imply [29] that every KD representation of quantum theory (as we have defined it) respects composition of quantum processes. Formally, such representations are symmetric strict monoidal functors—or in other words, diagram-preserving maps.

---

[1] C. Ferrie and J. Emerson, Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations, J. Phys. A **41**, 352001 (2008).

[2] R. W. Spekkens, Negativity and contextuality are equivalent notions of nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[3] E. P. Wigner, On the quantum correction for thermodynamic equilibrium, Phys. Rev. **40**, 749 (1932).

[4] D. Gross, Hudson's theorem for finite-dimensional quantum systems, J. Math. Phys. **47**, 122107 (2006).

[5] J. G. Kirkwood, Quantum statistics of almost classical assemblies, Phys. Rev. **44**, 31 (1933).

[6] P. A. M. Dirac, On the analogy between classical and quantum mechanics, Rev. Mod. Phys. **17**, 195 (1945).

[7] D. R. M. Arvidsson-Shukur, W. F. Braasch Jr., S. D. Bievre, J. Dressel, A. N. Jordan, C. Langrenez, M. Lostaglio, J. S. Lundeen, and N. Y. Halpern, Properties and applications of the Kirkwood-Dirac distribution, arXiv:2403.18899.

[8] D. R. M. Arvidsson-Shukur, N. Y. Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, Quantum advantage in postselected metrology, Nat. Commun. **11**, 3775 (2020).

[9] N. Lupu-Gladstein, Y. B. Yilmaz, D. R. M. Arvidsson-Shukur, A. Brodutch, A. O. T. Pang, A. M. Steinberg, and N. Y. Halpern, Negative quasiprobabilities enhance phase estimation in quantum-optics experiment, Phys. Rev. Lett. **128**, 220504 (2022).

[10] J. H. Jenne and D. R. M. Arvidsson-Shukur, Unbounded and lossless compression of multiparameter quantum information, Phys. Rev. A **106**, 042404 (2022).

[11] M. Lostaglio, Quantum fluctuation theorems, contextuality, and work quasiprobabilities, Phys. Rev. Lett. **120**, 040602 (2018).

[12] N. Y. Halpern, Jarzynski-like equality for the out-of-time-ordered correlator, Phys. Rev. A **95**, 012120 (2017).

[13] T. Upadhyaya, W. F. Braasch Jr., G. T. Landi, and N. Y. Halpern, Non-Abelian transport distinguishes three usually equivalent notions of entropy production, PRX Quantum **5**, 030355 (2024).

[14] E. Bäumer, M. Lostaglio, M. Perarnau-Llobet, and R. Sampaio, Fluctuating work in coherent quantum systems: Proposals and limitations, in *Thermodynamics in the Quantum Regime* (Springer, Berlin, 2018), p. 275–300.

[15] M. Lostaglio, A. Belenchia, A. Levy, S. Hernández-Gómez, N. Fabbri, and S. Gherardini, Kirkwood-Dirac quasiprobability approach to the statistics of incompatible observables, Quantum **7**, 1128 (2023).

[16] S. Hernández-Gómez, S. Gherardini, A. Belenchia, M. Lostaglio, A. Levy, and N. Fabbri, Projective measurements can probe nonclassical work extraction and timecorrelations, Phys. Rev. Res. **6**, 023280 (2024).

[17] S. Gherardini and G. D. Chiara, Quasiprobabilities in quantum thermodynamics and many-body systems, PRX Quantum **5**, 030201 (2024).

[18] A. Levy and M. Lostaglio, Quasiprobability distribution for heat fluctuations in the quantum regime, PRX Quantum **1**, 010309 (2020).

[19] R. Wagner and E. F. Galvão, Simple proof that anomalous weak values require coherence, Phys. Rev. A **108**, L040202 (2023).

[20] A. Budiyono and H. K. Dipojono, Quantifying quantum coherence via Kirkwood-Dirac quasiprobability, Phys. Rev. A **107**, 022408 (2023).

[21] M. Ban, On sequential measurements with indefinite causal order, Phys. Lett. A **403**, 127383 (2021).

[22] N. Gao, D. Li, A. Mishra, J. Yan, K. Simonov, and G. Chiribella, Measuring incompatibility and clustering quantum observables with a quantum switch, Phys. Rev. Lett. **130**, 170201 (2023).

[23] D. R. M. Arvidsson-Shukur, J. C. Drori, and N. Y. Halpern, Conditions tighter than noncommutation needed for nonclassicality, J. Phys. A **54**, 284001 (2021).

[24] S. De Bièvre, Complete incompatibility, support uncertainty, and Kirkwood-Dirac nonclassicality, Phys. Rev. Lett. **127**, 190404 (2021).

[25] N. Y. Halpern, B. Swingle, and J. Dressel, Quasiprobability behind the out-of-time-ordered correlator, Phys. Rev. A **97**, 042105 (2018).

[26] J. R. G. Alonso, N. Y. Halpern, and J. Dressel, Out-of-time-ordered-correlator quasiprobabilities robustly witness scrambling, Phys. Rev. Lett. **122**, 040404 (2019).

[27] J. R. G. Alonso, N. Shammah, S. Ahmed, F. Nori, and J. Dressel, Diagnosing quantum chaos with out-of-time-ordered-correlator quasiprobability in the kicked-top model, arXiv:2201.08175.

[28] R. Wagner, Z. Schwartzman-Nowik, I. L. Paiva, A. Te'eni, A. Ruiz-Molero, R. S. Barbosa, E. Cohen, and E. F. Galvão, Quantum circuits for measuring weak values, Kirkwood–Dirac quasiprobability distributions, and state spectra, Quantum Sci. Technol. **9**, 015030 (2024).

[29] S. M. Lane, *Categories for the Working Mathematician* (Springer, Berlin, 2013), Vol. 5.

[30] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, A structure theorem for generalized-noncontextual ontological models, Quantum **8**, 1283 (2024).

[31] C. Ferrie and J. Emerson, Framed Hilbert space: hanging the quasi-probability pictures of quantum theory, New J. Phys. **11**, 063040 (2009).

[32] C. Ferrie, Quasi-probability representations of quantum theory with applications to quantum information science, Rep. Prog. Phys. **74**, 116001 (2011).

[33] O. Christensen, *An Introduction to Frames and Riesz Bases*, Applied and Numerical Harmonic Analysis (Springer, Berlin, 2016).

[34] 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).

[35] B. Tamir and E. Cohen, Introduction to weak measurements and weak values, Quanta **2**, 7 (2013).

[36] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Colloquium: Understanding quantum weak values: Basics and applications, Rev. Mod. Phys. **86**, 307 (2014).

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

[38] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens, Characterization of noncontextuality in the framework of generalized probabilistic theories, PRX Quantum **2**, 010331 (2021).

[39] R. W. Spekkens, The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein, arXiv:1909.04628.

[40] D. Schmid and R. W. Spekkens, Contextual advantage for state discrimination, Phys. Rev. X **8**, 011015 (2018).

[41] M. F. Pusey, Anomalous weak values are proofs of contextuality, Phys. Rev. Lett. **113**, 200401 (2014).

[42] R. Kunjwal, M. Lostaglio, and M. F. Pusey, Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts, Phys. Rev. A **100**, 042116 (2019).

[43] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation contextuality powers parity-oblivious multiplexing, Phys. Rev. Lett. **102**, 010401 (2009).

[44] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, Optimal bounds for parity-oblivious random access codes, New J. Phys. **18**, 045003 (2016).

[45] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, Parity oblivious *d*-level random access codes and class of noncontextuality inequalities, Quantum Inf. Process. **18**, 111 (2019).

[46] D. Saha, P. Horodecki, and M. Pawłowski, State independent contextuality advances one-way communication, New J. Phys. **21**, 093057 (2019).

[47] M. Lostaglio and G. Senno, Contextual advantage for state-dependent cloning, Quantum **4**, 258 (2020).

[48] R. Raussendorf, Contextuality in measurement-based quantum computation, Phys. Rev. A **88**, 022322 (2013).

[49] M. J. Hoban, E. T. Campbell, K. Loukopoulos, and D. E. Browne, Non-adaptive measurement-based quantum computation and multi-party Bell inequalities, New J. Phys. **13**, 023014 (2011).

[50] R. Wagner, A. Camillini, and E. F. Galvão, Coherence and contextuality in a Mach-Zehnder interferometer, Quantum **8**, 1240 (2024).

[51] S. A. Yadavalli and R. Kunjwal, Contextuality in entanglement-assisted one-shot classical communication, Quantum **6**, 839 (2022).

[52] D. Saha and A. Chaturvedi, Preparation contextuality as an essential feature underlying quantum communication advantage, Phys. Rev. A **100**, 022108 (2019).

[53] M. Lostaglio, Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response, Phys. Rev. Lett. **125**, 230603 (2020).

[54] R. Kunjwal and R. W. Spekkens, From the kochen-specker theorem to noncontextuality inequalities without assuming determinism, Phys. Rev. Lett. **115**, 110403 (2015).

[55] 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**, ncomms11780 (2016).

[56] T. Giordani, R. Wagner, C. Esposito, A. Camillini, F. Hoch, G. Carvacho, C. Pentangelo, F. Ceccarelli, S. Piacentini, A. Crespi, N. Spagnolo, R. Osellame, E. F. Galvão, and F. Sciarrino, Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor, Sci. Adv. **9**, eadj4249 (2023).

[57] R. Kunjwal, Contextuality beyond the Kochen-Specker theorem, arXiv:1612.07250.

[58] D. Schmid, R. W. Spekkens, and E. Wolfe, All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences, Phys. Rev. A **97**, 062103 (2018).

[59] M. F. Pusey, L. del Rio, and B. Meyer, Contextuality without access to a tomographically complete set, arXiv:1904.08699.

[60] J. H. Selby, E. Wolfe, D. Schmid, A. B. Sainz, and V. P. Rossi, Linear program for testing nonclassicality and an open-source implementation, Phys. Rev. Lett. **132**, 050202 (2024).

[61] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Accessible fragments of generalized probabilistic theories, cone equivalence, and applications to witnessing nonclassicality, Phys. Rev. A **107**, 062203 (2023).

[62] D. Schmid, J. H. Selby, and R. W. Spekkens, Addressing some common objections to generalized noncontextuality, Phys. Rev. A **109**, 022228 (2024).

[63] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey, Uniqueness of noncontextual models for stabilizer subtheories, Phys. Rev. Lett. **129**, 120403 (2022).

[64] F. Shahandeh, Quantum computational advantage implies contextuality, arXiv:2112.00024.

[65] K. Flatt, H. Lee, C. R. I. Carceller, J. B. Brask, and J. Bae, Contextual advantages and certification for maximum-confidence discrimination, PRX Quantum **3**, 030337 (2022).

[66] S. Mukherjee, S. Naonit, and A. K. Pan, Discriminating three mirror-symmetric states with a restricted contextual advantage, Phys. Rev. A **106**, 012216 (2022).

[67] J. Shin, D. Ha, and Y. Kwon, Quantum contextual advantage depending on nonzero prior probabilities in state discrimination of mixed qubit states, Entropy **23**, 1583 (2021).

[68] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, Why interference phenomena do not capture the essence of quantum theory, Quantum **7**, 1119 (2023).

[69] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, Reply to "Comment on 'Why interference phenomena do not capture the essence of quantum theory' ", arXiv:2207.11791.

[70] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, Aspects of the phenomenology of interference that are genuinely nonclassical, Phys. Rev. A **108**, 022207 (2023).

[71] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Contextuality without incompatibility, Phys. Rev. Lett. **130**, 230201 (2023).

[72] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, What is nonclassical about uncertainty relations? Phys. Rev. Lett. **129**, 240401 (2022).

[73] N. E. Comar, D. Cius, L. F. Santos, R. Wagner, and B. Amaral, Contextuality in anomalous heat flow, arXiv:2406.09715.

[74] V. P. Rossi, D. Schmid, J. H. Selby, and A. B. Sainz, Contextuality with vanishing coherence and maximal robustness to dephasing, Phys. Rev. A **108**, 032213 (2023).

[75] R. Wagner, R. S. Barbosa, and E. F. Galvão, Inequalities witnessing coherence, nonlocality, and contextuality, Phys. Rev. A **109**, 032220 (2024).

[76] R. D. Baldijão, R. Wagner, C. Duarte, B. Amaral, and M. T. Cunha, Emergence of noncontextuality under quantum darwinism, PRX Quantum **2**, 030351 (2021).

[77] V. J. Wright and M. Farkas, Invertible map between Bell nonlocal and contextuality scenarios, Phys. Rev. Lett. **131**, 220202 (2023).

[78] D. Schmid, J. H. Selby, and R. W. Spekkens, Unscrambling the omelette of causation and inference: The framework of causal-inferential theories, arXiv:2009.03297.

[79] R. Kunjwal and R. W. Spekkens, From statistical proofs of the Kochen-Specker theorem to noise-robust noncontextuality inequalities, Phys. Rev. A **97**, 052110 (2018).

[80] R. Kunjwal, Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements, Quantum **3**, 184 (2019).

[81] R. Kunjwal, Hypergraph framework for irreducible noncontextuality inequalities from logical proofs of the Kochen-Specker theorem, Quantum **4**, 219 (2020).

[82] R. Kunjwal, Fine's theorem, noncontextuality, and correlations in Specker's scenario, Phys. Rev. A **91**, 022108 (2015).

[83] T. Gonda, R. Kunjwal, D. Schmid, E. Wolfe, and A. B. Sainz, Almost quantum correlations are inconsistent with specker's principle, Quantum **2**, 87 (2018).

[84] C. Fernandes, R. Wagner, L. Novo, and E. F. Galvão, Unitary-invariant witnesses of quantum imaginarity, arXiv:2403.15066.

[85] S. De Bièvre, Relating incompatibility, noncommutativity, uncertainty, and Kirkwood–Dirac nonclassicality, J. Math. Phys. **64**, 022202 (2023).

[86] J. Xu, Kirkwood-dirac classical pure states, Phys. Lett. A **510**, 129529 (2024).

[87] C. Langrenez, D. R. M. Arvidsson-Shukur, and S. D. Bièvre, Characterizing the geometry of the Kirkwood-Dirac positive states, J. Math. Phys. **65**, 072201 (2024).

[88] H. F. Hofmann, Complex joint probabilities as expressions of reversible transformations in quantum mechanics, New J. Phys. **14**, 043031 (2012).

[89] M. Oszmaniec, D. J. Brod, and E. F. Galvão, Measuring relational information between quantum states, and applications, New J. Phys. **26**, 013053 (2024).

[90] L. M. Yosef, S. Ubaru, L. Horesh, and H. Avron, Multivariate trace estimation using quantum state space linear algebra, arXiv:2405.01098.

[91] Y. Quek, E. Kaur, and M. M. Wilde, Multivariate trace estimation in constant quantum depth, Quantum **8**, 1220 (2024).

[92] K. Husimi, Some formal properties of the density matrix, Proc. Phys.-Math. Soc. Jpn. **22**, 264 (1940).

[93] R. J. Glauber, Coherent and incoherent states of the radiation field, Phys. Rev. **131**, 2766 (1963).

[94] E. C. G. Sudarshan, Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams, Phys. Rev. Lett. **10**, 277 (1963).

[95] N. Harrigan and R. W. Spekkens, Einstein, Incompleteness, and the epistemic view of quantum states, Found. Phys. **40**, 125 (2010).

[96] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Probabilistic theories with purification, Phys. Rev. A **81**, 062348 (2010).

[97] R. W. Spekkens, Quasi-quantization: Classical statistical theories with an epistemic restriction, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer Netherlands, Dordrecht, 2016), pp. 83–135.

[98] L. M. Johansen, Quantum theory of successive projective measurements, Phys. Rev. A **76**, 012119 (2007).

[99] V. Bollen, Y. M. Sua, and K. F. Lee, Direct measurement of the Kirkwood-Rihaczek distribution for the spatial properties of a coherent light beam, Phys. Rev. A **81**, 063826 (2010).

[100] M. E. Gorbunov and O. A. Koval, Kirkwood and wigner quantum densities, their properties, and applications in radiophysics, Radiophys. Quantum Electron. **64**, 140 (2021).

[101] L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012.

[102] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).