From Time-Reversal Symmetry to Quantum Bayes' Rules

Arthur J. Parzygnat [1,*] and James Fullwood [2,†]

[1] *Graduate School of Informatics, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan*
[2] *School of Mathematical Sciences, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, People's Republic of China*

(Received 5 January 2023; revised 7 March 2023; accepted 21 March 2023; published 2 June 2023)

Bayes' rule, $\mathbb{P}(B|A)\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B)$, is one of the simplest yet most profound, ubiquitous, and far-reaching results of classical probability theory, with applications in any field utilizing statistical inference. Many attempts have been made to extend this rule to quantum systems, the significance of which we are only beginning to understand. In this work, we develop a systematic framework for defining Bayes' rule in the quantum setting, and we show that a vast majority of the proposed quantum Bayes' rules appearing in the literature are all instances of our definition. Moreover, our Bayes' rule is based upon a simple relationship between the notions of *state over time* and a time-reversal-symmetry map, both of which are introduced here.

DOI: 10.1103/PRXQuantum.4.020334

I. INTRODUCTION

Bayes' rule is a cornerstone of inference, prediction, retrodiction, and decision-making that is used throughout the natural sciences [1–10]. Due to its ubiquitous stature, many have proposed extensions beyond classical probability theory into the setting of quantum mechanics [11–34], with applications to cosmology [23], entanglement wedge reconstruction in the anti–de Sitter−conformal field theory (AdS-CFT) correspondence [26], and quantum foundations [17,25,35].

A common approach to formulating quantum generalizations of the classical Bayes' rule, namely,

$$\mathbb{P}(y|x)\mathbb{P}(x) = \mathbb{P}(x|y)\mathbb{P}(y), \tag{1}$$

is by "quantizing" Eq. (1), i.e., by defining operator analogs of $\mathbb{P}(y|x), \ldots, \mathbb{P}(y)$ in such a way that the substitution of such analogs into Eq. (1) also yields a valid equation. But since there are various approaches to formulating operator analogs of $\mathbb{P}(y|x), \ldots, \mathbb{P}(y)$ and moreover, since multiplication of such operators is not necessarily commutative, different formulations of quantum Bayes' rules have appeared in the literature, each having its own advantages over the others [11–34].

*parzygnat@nagoya-u.jp; parzygnat@ihes.fr
†fullwood@sjtu.edu.cn

The approach taken here, however, is to view Eq. (1) as a reflection of a certain time-reversal symmetry for classical systems and that the associated symmetry transformation relating both sides of Eq. (1) is more fundamental than the particular form of the equation. Moreover, our emphasis on transformations, as opposed to equations, reveals that various formulations of quantum Bayes' rules are all manifestations of what we refer to as a Bayes' rule *with respect to a state over time*, which is a quantum analog of a joint distribution $\mathbb{P}(x, y)$ associated with a physical system on two timelike separated regions.

To explain how such a perspective may be taken in the classical case, let $(\Omega, \mathbb{P})$ be a finite probability space, where Ω is a finite set corresponding to all possible outcomes of an experiment or data-generating process. Given random variables $\mathcal{X} : \Omega \to X$ and $\mathcal{Y} : \Omega \to Y$ on Ω, let $p = \mathcal{X}_*\mathbb{P}$ and $q = \mathcal{Y}_*\mathbb{P}$ be the associated probability mass functions, so that for all $x \in X$ and $y \in Y$,

$$p(x) = \mathbb{P}\left(\mathcal{X}^{-1}(x)\right) \quad \text{and} \quad q(y) = \mathbb{P}\left(\mathcal{Y}^{-1}(y)\right). \tag{2}$$

Since we are considering not one but two random variables, $\mathcal{X}$ and $\mathcal{Y}$, there are two associated joint distributions or, rather, *states over time*, $\vartheta : X \times Y \to [0, 1]$ and $\vartheta^* : Y \times X \to [0, 1]$, depending on whether the observation of $\mathcal{X}$ precedes $\mathcal{Y}$ or vice versa. In particular, ϑ is the state over time corresponding to first observing $\mathcal{X}$ and then observing $\mathcal{Y}$, which is given by

$$\vartheta(x, y) = p(y|x)p(x), \tag{3}$$

where $p(y|x)$ represents the conditional probability of observing y *given* that x has been observed first. Similarly,

ϑ^* corresponds to the time reversal of this procedure, since it describes the probability of first observing $\mathcal{Y}$ and then observing $\mathcal{X}$, which is given by

$$\vartheta^*(y,x) = q(x|y)q(y). \tag{4}$$

Bayes' rule (1) in this context may then be reformulated as

$$\vartheta = \gamma(\vartheta^*), \tag{5}$$

where γ is the canonical map sending distributions on $Y \times X$ to distributions on $X \times Y$, which we view as a reversal of time for joint distributions representing timelike separated variables. Equation (5) then says that the states over time ϑ and ϑ^* are related by the time-reversal transformation γ and it is this formulation of Bayes' rule that we take as a guide into the quantum realm.

However, while states over time in the classical case are essentially unique, there are various approaches to defining states over time in the quantum setting. This stems from the fact that while the postulates of quantum mechanics make clear that a joint state supported on *spacelike* separated regions of a system is represented by a density matrix on the tensor product of the Hilbert spaces for each region, the postulates are silent regarding what mathematical entity should faithfully represent *timelike* separated states of a system. In particular, if two timelike separated states are causally related, then a suitable notion of state over time for the system should be an operator on the tensor product of the Hilbert spaces at the two different times, encoding not only the two timelike separated states but also causal correlations between the states as well. Moreover, it has been emphasized in Refs. [36,37] that such a state over time should admit negative eigenvalues if it is to encode temporal correlations and, as such, states over time should not be positive in general. It is for these reasons that defining states over time is not so straightforward.

In Ref. [38], a minimal list of axioms has been proposed that any general state-over-time construction should satisfy and a no-go theorem has been proved, stating that there is no such construction satisfying their list of axioms. However, while the aforementioned no-go theorem is mathematically correct, in Ref. [37] the present authors have found a loophole in this theorem by slightly weakening the hypotheses in a way which does not alter their physical significance and interpretation, thus resulting in an explicit state-over-time construction satisfying the axioms put forth in Ref. [38]. While it is still not known whether or not our state-over-time construction is *characterized* by such axioms, we are nevertheless starting to better understand states over time from both a mathematical and physical perspective, as is further supported in this work.

In particular, we use Eq. (5) to formulate a general quantum Bayes' rule that takes into account the choice of a state-over-time construction. By doing so, we show that the various formulations of quantum Bayes' rules appearing in Refs. [11–13,16,17,23,25,28,29,32] may all be obtained from our Bayes' rule once an appropriate notion of state over time is specified in each case. We also establish a list of axioms for state-over-time constructions similar to those in Ref. [38] and we prove general results for arbitrary state-over-time constructions satisfying certain subsets of these axioms. For example, we show that our Bayes' rule with respect to any state-over-time construction satisfying what we refer to as the *classical-limit axiom* yields the state-update rule associated with quantum measurement [12,39,40]. Moreover, with such axioms, we are able to identify the key differences between these various approaches toward quantum Bayes' rules, while at the same time incorporating them all into a single framework.

Another application of our quantum Bayes' rule is in regard to the notion of time reversal in quantum theory. In particular, our Bayes' rule yields a novel notion of a *Bayesian inverse* of a quantum channel with respect to a prior state that is to dynamically evolve according to the channel. Moreover, we show that the Bayesian inverse of a completely positive trace-preserving (CPTP) map generally differs from that of its Hilbert-Schmidt adjoint. Thus, our Bayesian inverses provide a more robust notion of time reversal in quantum theory, in parallel with other results on retrodictability [8,9,41–44]. And while we find that the Hilbert-Schmidt adjoint does not in general provide an appropriate notion of time reversal, in the case of bistochastic channels, we show that the Bayesian inverse with respect to the uniform prior is indeed the Hilbert-Schmidt adjoint of the channel. This result is, in fact, independent of a state-over-time construction provided that it satisfies the classical-limit axiom. This clarifies why bistochastic channels are often viewed as the only channels exhibiting a canonical time-reversal map [45–47].

Throughout our work, we illustrate our definitions through a multitude of examples, including the two-state vector formalism [41,42,48], the time-dependent two-point correlator [49,50], the symmetric bloom of the present authors [37], the noncommutative Bayes' theorem of the first author [29], the quantum Bayes' rule of Fuchs [17], the causal state formalism of Leifer and Spekkens with the Petz recovery map [25,51], the compound states of Ohya [52,53], generalized conditional expectations [32], and the state-update rule associated with quantum measurements, the latter of which has often been called the quantum analog of Bayes' rule by Bub, Ozawa, Tegmark, and others [11,13,23,54].

II. STATES OVER TIME, BAYES' RULES, AND BAYESIAN INVERSES

A. Notation and conventions

In this work, we represent finite-dimensional hybrid classical-quantum systems by *multimatrix algebras* (as in

Refs. [55] and [56, Chapter 2]), which are direct sums of matrix algebras and which are denoted by $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$. Any such multimatrix algebra $\mathcal{A}$ is therefore of the form $\mathcal{A} = \bigoplus_{x\in X} \mathbb{M}_{m_x}$, with X some finite index set, $\{m_x\}$ positive integers, and $\mathbb{M}_d$ denoting $d \times d$ complex matrices. By using multimatrix algebras instead of just operators on some Hilbert space, we have a single framework where all concepts such as conditional probabilities, density matrices, positive operator-valued measures (POVMs), ensemble preparations, instruments, and quantum channels are all instances of CPTP maps between such multimatrix algebras. By working with CPTP maps, we are working in the Schrödinger picture of quantum theory, in contrast to our previous work [37], where our results have been formulated in the Heisenberg picture. As such, some notation and terminology differ. We now provide the basic definitions and notation which are used throughout [55,57].

Every multimatrix algebra $\mathcal{A} = \bigoplus_{x\in X} \mathbb{M}_{m_x}$ has a trace tr the value of which on $A = \bigoplus_{x\in X} A_x$ is given by $\mathrm{tr}(A) = \sum_{x\in X} \mathrm{tr}(A_x)$, where the latter trace is the standard (unnormalized) trace on the matrix algebra $\mathbb{M}_{m_x}$. A *density matrix* (or *state*) in $\mathcal{A}$ is an element ρ of $\mathcal{A}$ such that $\mathrm{tr}(\rho) = 1$ and ρ is positive, which means there exists some element $A \in \mathcal{A}$ such that $A^\dagger A = \rho$. Here, the $\dagger$ denotes the component-wise conjugate transpose, namely $A^\dagger = \bigoplus_{x\in X} A_x^\dagger$, and the multiplication is also component-wise: $AA' = \bigoplus_{x\in X} A_x A'_x$. The set of all states in $\mathcal{A}$ is denoted by $\mathcal{S}(\mathcal{A})$.

We immediately give some examples. First, if $m_x = 1$ for all $x \in X$, then each matrix algebra is one-dimensional, so that ρ corresponds to a collection of non-negative numbers $\{\rho_x\}$ the sum of which satisfies $\sum_{x\in X} \rho_x = 1$. In other words, a density matrix can be viewed as a probability distribution on the index set X. At the other extreme, if X is an index set with only a single element, then ρ is an ordinary density matrix in a matrix algebra. Note that in the general case of a multimatrix algebra, the x component of ρ, namely ρ_x, need not be a density matrix, since its trace can be less than 1. However, if the trace is nonzero, then $\rho_x/\mathrm{tr}(\rho_x)$ is a density matrix on $\mathbb{M}_{m_x}$ in the usual sense. Hence, a density matrix ρ on a multimatrix algebra, also called a *classical-quantum state*, can be viewed as a collection of density matrices on possibly different matrix algebras weighted by some probability distribution, namely $\rho = \bigoplus_{x\in X} \mathrm{tr}(\rho_x)\rho_x/\mathrm{tr}(\rho_x)$.

For more examples, let $\mathcal{B} = \bigoplus_{y\in Y} \mathbb{M}_{n_y}$ be another multimatrix algebra and let $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ be a CPTP map. This specializes to many cases of interest in quantum information theory (for a brief summary, see Table I, and for more details, see the introduction of Ref. [47]):

(1) When $m_x = 1$ and $n_y = 1$ for all $x \in X, y \in Y$, the map $\mathcal{E}$ corresponds to a classical channel, i.e., a collection of conditional probabilities $\mathbb{P}(y|x)$. More explicitly, if δ_x denotes the unit vector with 1 in the x component and 0 otherwise, then $\mathbb{P}(y|x)$ is the y component of $\mathcal{E}(\delta_x)$.
(2) When X and Y have only one element each, $\mathcal{E}$ corresponds to a quantum channel between matrix algebras.
(3) When $n_y = 1$ for all $y \in Y$ and X has only a single element, $\mathcal{E}$ corresponds to a POVM. Indeed, each y component of $\mathcal{E}$ defines a positive functional $\mathcal{E}_y : \mathbb{M}_m \to \mathbb{C}$, which equals $\mathrm{tr}(M_y \cdot\,)$ for some unique positive matrix $M_y \in \mathbb{M}_m$. The trace-preserving condition guarantees that $\sum_{y\in Y} M_y = \mathbb{1}_m$.
(4) Dually, when $m_x = 1$ for all $x \in X$ and Y has only a single element, $\mathcal{E}$ corresponds to an ensemble preparation. Indeed, $\mathcal{E}$ sends each unit vector δ_x to some positive matrix ρ_x in $\mathbb{M}_n$. The trace-preserving condition guarantees that $\mathrm{tr}(\rho_x) = 1$ so that ρ_x is a density matrix for each $x \in X$.
(5) When $n_y = n$ for all $y \in Y$ for some positive integer n and X has only a single element, $\mathcal{E} : \mathbb{M}_m \to \bigoplus_{y\in Y} \mathbb{M}_n$ corresponds to a quantum instrument. Indeed, the projection of $\mathcal{E}$ onto the y component defines a CP map $\mathcal{E}_y : \mathbb{M}_m \to \mathbb{M}_n$. By the trace-preserving condition on $\mathcal{E}$, the sum $\sum_{y\in Y} \mathcal{E}_y : \mathbb{M}_m \to \mathbb{M}_n$ is CPTP. This is the usual definition for an instrument with a finite outcome space Y [58].

TABLE I. A summary of some of the classical and quantum information-theoretic realizations of a CPTP map between multimatrix algebras. The notation m (n), as opposed to m_x (n_y), is used for the size of the matrices in the algebra when $|X| = 1$ ($|Y| = 1$) or when all values coincide.

$\bigoplus_{x\in X} \mathbb{M}_{m_x} \xrightarrow{\mathcal{E}} \bigoplus_{y\in Y} \mathbb{M}_{n_y}$	Concept in quantum information
$\lvert X\rvert = 1, m = 1 = n_y\ \forall\, y$	Probability distribution on Y
$\lvert X\rvert = 1 = \lvert Y\rvert, m = 1$	Density matrix on $\mathbb{M}_n$
$m_x = 1 = n_y\ \forall\, x, y$	Classical channel $X \to Y$
$\lvert X\rvert = 1 = \lvert Y\rvert$	Quantum channel $\mathbb{M}_m \to \mathbb{M}_n$
$\lvert X\rvert = 1, n_y = 1\ \forall\, y$	POVM on $\mathbb{M}_m$
$\lvert Y\rvert = 1, m_x = 1\ \forall\, x$	Ensemble preparation on $\mathbb{M}_n$
$\lvert X\rvert = 1, n_y = n\ \forall\, y$	Quantum instrument

More generally, arbitrary direct sums of matrix algebras can be used to describe certain superselection sectors, as discussed in Refs. [59] and [60, Section 2.3]. Thus, we find that a number of fundamental concepts in quantum information theory may be formulated in terms of completely positive maps between multimatrix algebras. Moreover, as every finite-dimensional C^*-algebra is isomorphic to a multimatrix algebra [55], our formulation provides a stepping stone toward generalizations to infinite-dimensional quantum systems, further justifying our use of the multimatrix algebra formalism.

Henceforth, if $\mathcal{A}$ and $\mathcal{B}$ are multimatrix algebras, the collection of all CPTP maps from $\mathcal{A}$ to $\mathcal{B}$ is denoted by

$\mathrm{CPTP}(\mathcal{A},\mathcal{B})$. If $\mathcal{E}:\mathcal{A}\to\mathcal{B}$ is a linear map, its Hilbert-Schmidt adjoint $\mathcal{E}^*:\mathcal{B}\to\mathcal{A}$ is the unique linear map satisfying

$$\mathrm{tr}\big(\mathcal{E}(A)^\dagger B\big)=\mathrm{tr}\big(A^\dagger\mathcal{E}^*(B)\big) \tag{6}$$

for all $A\in\mathcal{A}$ and $B\in\mathcal{B}$. Let $\mu_{\mathcal{A}}:\mathcal{A}\otimes\mathcal{A}\to\mathcal{A}$ denote the multiplication map, which is uniquely determined by its assignment on tensors via $A_1\otimes A_2\mapsto A_1A_2$. Given $\mathcal{E}$ as above, its associated *channel state* is the element of $\mathcal{A}\otimes\mathcal{B}$ given by

$$\mathscr{D}_{\mathcal{A},\mathcal{B}}[\mathcal{E}]:=(\mathrm{id}_{\mathcal{A}}\otimes\mathcal{E})\big(\mu^*_{\mathcal{A}}(1_{\mathcal{A}})\big), \tag{7}$$

where $1_{\mathcal{A}}$ is the unit element in $\mathcal{A}$ and $\mu^*_{\mathcal{A}}$ is the Hilbert-Schmidt adjoint of the multiplication map. When the algebras are clear from context, the shorthand $\mathscr{D}[\mathcal{E}]$ is used. Although this definition may look unfamiliar at first, it reduces to two familiar cases for certain multimatrix algebras. First, when $\mathcal{A}=\mathbb{M}_m$ and $\mathcal{B}=\mathbb{M}_n$, the channel state reduces to the associated *Jamiołkowski state* [61]

$$\mathscr{D}[\mathcal{E}]=\sum_{i,j}E^{(m)}_{ij}\otimes\mathcal{E}(E^{(m)}_{ji}), \tag{8}$$

where $\{E^{(m)}_{ij}\}$ are the standard matrix units in $\mathbb{M}_m$. In the special case $\mathcal{E}=\mathrm{id}_{\mathcal{A}}$, the Jamiołkowski state becomes the SWAP operator, which satisfies $\mathscr{D}[\mathrm{id}_{\mathcal{A}}]\big(|i\rangle\otimes|j\rangle\big)=|j\rangle\otimes|i\rangle$ for all $i,j\in\{1,\ldots,m\}$ (Dirac notation is implemented). Second, when $\mathcal{A}=\bigoplus_{x\in X}\mathbb{C}\equiv\mathbb{C}^X$ and $\mathcal{B}=\bigoplus_{y\in Y}\mathbb{C}\equiv\mathbb{C}^Y$, so that a positive trace-preserving map $\mathcal{E}$ corresponds to conditional probabilities $\mathbb{P}(y|x)$, the channel state is an element of $\bigoplus_{(x,y)\in X\times Y}\mathbb{C}\equiv\mathbb{C}^{X\times Y}$, the (x,y) component of which is given by $\mathbb{P}(y|x)$.

B. Main definitions

Definition 1.—A *state-over-time function* associates every pair $(\mathcal{A},\mathcal{B})$ of multimatrix algebras with a map $\star_{\mathcal{A}\mathcal{B}}:\mathrm{CPTP}(\mathcal{A},\mathcal{B})\times\mathcal{S}(\mathcal{A})\to\mathcal{A}\otimes\mathcal{B}$, the value of which on $(\mathcal{E},\rho)$ is denoted by $\mathcal{E}\star_{\mathcal{A}\mathcal{B}}\rho$ (or just $\mathcal{E}\star\rho$ when $\mathcal{A}$ and $\mathcal{B}$ are clear) [62], such that $\star_{\mathcal{A}\mathcal{B}}$ preserves marginal states in the sense that

$$\mathrm{tr}_{\mathcal{B}}\left(\mathcal{E}\star_{\mathcal{A}\mathcal{B}}\rho\right)=\rho\quad\text{and}\quad\mathrm{tr}_{\mathcal{A}}\left(\mathcal{E}\star_{\mathcal{A}\mathcal{B}}\rho\right)=\mathcal{E}(\rho), \tag{9}$$

where $\mathrm{tr}_{\mathcal{B}}:\mathcal{A}\otimes\mathcal{B}\to\mathcal{A}$ and $\mathrm{tr}_{\mathcal{A}}:\mathcal{A}\otimes\mathcal{B}\to\mathcal{B}$ denote the partial traces. In such a case, the element $\mathcal{E}\star_{\mathcal{A}\mathcal{B}}\rho\in\mathcal{A}\otimes\mathcal{B}$ is referred to as the *state over time* associated with $\star_{\mathcal{A}\mathcal{B}}$ and the input $(\mathcal{E},\rho)$.

To incorporate more examples and also various approaches to states over time, the definition given here is less restrictive than the definition of a state-over-time function given in Refs. [37,38]. In particular, this definition includes only the bare minimum of what one would expect from a state-over-time function, namely, that the output element $\mathcal{E}\star\rho$ has the expected marginals. Note that unitality, as defined in Refs. [37,38], holds automatically because if ρ is a density matrix and $\mathcal{E}$ is trace preserving, then $\mathrm{tr}(\mathcal{E}\star\rho)=1$ because of the marginal preservation property. Of course, however, one would expect a physically meaningful state-over-time function to satisfy additional properties.

Definition 2.—A state-over-time function $\star$

(P1) is *Hermitian* if and only if $\mathcal{E}\star\rho$ is self-adjoint for all $\rho\in\mathcal{S}(\mathcal{A})$ and $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$

(P2) is *locally positive* (or *block positive*) if and only if

$$\mathrm{tr}\big((\mathcal{E}\star\rho)^\dagger(A\otimes B)\big)\geq 0 \tag{10}$$

for all $\rho\in\mathcal{S}(\mathcal{A})$, $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$, and for all positive $A\in\mathcal{A}$ and $B\in\mathcal{B}$

(P3) is *positive* if and only if $\mathcal{E}\star\rho$ is positive for all $\rho\in\mathcal{S}(\mathcal{A})$ and $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$

(P4) is *state linear* if and only if

$$\mathcal{E}\star\big(\lambda\rho+(1-\lambda)\sigma\big)=\lambda\mathcal{E}\star\rho+(1-\lambda)\mathcal{E}\star\sigma \tag{11}$$

for all $\lambda\in[0,1]$, $\rho,\sigma\in\mathcal{S}(\mathcal{A})$, and $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$

(P5) is *process linear* if and only if

$$\big(\lambda\mathcal{E}+(1-\lambda)\mathcal{F}\big)\star\rho=\lambda\mathcal{E}\star\rho+(1-\lambda)\mathcal{F}\star\rho \tag{12}$$

for all $\lambda\in[0,1]$, $\rho\in\mathcal{S}(\mathcal{A})$, and $\mathcal{E},\mathcal{F}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$

(P6) is *bilinear* if and only if $\star$ is state linear and process linear

(P7) satisfies the *classical-limit axiom* if and only if given any $\rho\in\mathcal{S}(\mathcal{A})$ and $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$ satisfying $[\mathscr{D}[\mathcal{E}],\rho\otimes 1_{\mathcal{B}}]=0$ implies [63]

$$\mathcal{E}\star\rho=\mathscr{D}[\mathcal{E}](\rho\otimes 1_{\mathcal{B}}) \tag{13}$$

There is also an associativity axiom guaranteeing that a state-over-time function yields a consistent notion of a *tripartite* state over time on $\mathcal{A}\otimes\mathcal{B}\otimes\mathcal{C}$ associated with every composable pair $\mathcal{A}\xrightarrow{\mathcal{E}}\mathcal{B}\xrightarrow{\mathcal{F}}\mathcal{C}$ of CPTP maps. We mention in our examples when this associativity axiom holds, but otherwise we do not make use of it in this work [64].

Note that if a state-over-time function $\star$ satisfies the classical-limit axiom, then $\mathcal{E}\star(1_{\mathcal{A}}/\mathrm{tr}(1_{\mathcal{A}}))=(1/\mathrm{tr}(1_{\mathcal{A}}))\mathscr{D}[\mathcal{E}]$. This shows that such a $\star$ can be viewed as an extension of the Jamiołkowski isomorphism to include states besides the maximally mixed state (though it need not be bijective). We also note that a state over time that is positive appears in the marginal state problem and is sometimes called a *compound state* [34,52,53,65,66].

Since we do not require positivity, our considerations are more general.

We provide many examples of state-over-time functions in the remaining sections. But first, we introduce the definition of a Bayesian inverse with respect to a state-over-time function in terms of a quantum Bayes' rule.

Definition 3.—Let $\star$ be a state-over-time function. Given a density matrix (a *prior*) $\rho \in \mathcal{S}(\mathcal{A})$ and a CPTP map $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ (a process), a *Bayesian inverse* associated with $(\mathcal{E}, \rho)$ is a CPTP map $\mathcal{E}^{\star}_{\rho} : \mathcal{B} \to \mathcal{A}$ such that

$$\mathcal{E} \star \rho = \tau \left(\mathcal{E}^{\star}_{\rho} \star \mathcal{E}(\rho) \right), \tag{14}$$

where $\tau : \mathcal{B} \otimes \mathcal{A} \to \mathcal{A} \otimes \mathcal{B}$ is the *quantum time-reversal map* for states over time, defined as the unique conjugate-linear extension of the assignment

$$\tau(B \otimes A) = A^{\dagger} \otimes B^{\dagger}. \tag{15}$$

The equation $\mathcal{E} \star \rho = \tau \left(\mathcal{E}^{\star}_{\rho} \star \mathcal{E}(\rho) \right)$ is then referred to as *Bayes' rule* associated with $\star$ and the input $(\mathcal{E}, \rho)$.

By applying the partial trace $\mathrm{tr}_{\mathcal{B}}$ to both sides of Bayes' rule, it follows that

$$\rho = \mathcal{E}^{\star}_{\rho}\big(\mathcal{E}(\rho)\big). \tag{16}$$

In other words, if ρ is thought of as a prior, with $\mathcal{E}(\rho)$ as the associated prediction via $\mathcal{E}$ and $\mathcal{E}^{\star}_{\rho}$ as the associated retrodiction map, then the retrodiction applied to the prediction gives back the prior. If $\mathcal{E}^{\star}_{\rho}$ exists and is unique for all $\mathcal{E}$ and ρ (such that ρ and $\mathcal{E}(\rho)$ are faithful), this assignment defines a (universal) *recovery map* in the language of Refs. [67–70], a *state-retrieval map* in the language of Ref. [71], and a *retrodiction family* in the language of Ref. [44].

While the linear swap map $\gamma : \mathcal{B} \otimes \mathcal{A} \to \mathcal{A} \otimes \mathcal{B}$ provides a suitable notion of time reversal for states over time in the classical setting (as described in Sec. I) [72], we find that composing the swap map with the dagger provides a more robust notion of time reversal for states over time in the quantum setting (see Fig. 1). Indeed, the swap map γ on its own only guarantees that the right-hand side of Bayes' rule is an element of $\mathcal{A} \otimes \mathcal{B}$. However, if a quantum channel $\mathcal{E}$ is invertible with its inverse also a channel, then the usage of τ ensures that $\mathcal{E}^{-1}$ is a Bayesian inverse for $\mathcal{E}$, which would not necessarily be the case if we had simply used γ (for more details, see Remark 2). Remarks 1, 3, and 5 provide further justifications for supplementing the swap map with the dagger in the quantum setting. Moreover, τ reduces to γ for classical systems and the Bayes' rule from Definition 3 coincides with the standard Bayes' rule [Eq. (1)] on commutative algebras when we take the state-over-time function to be the standard one [37, Section 1]. We see this explicitly in Sec. III, along with many other examples from the literature, some of which have also been called quantum Bayes' rules.

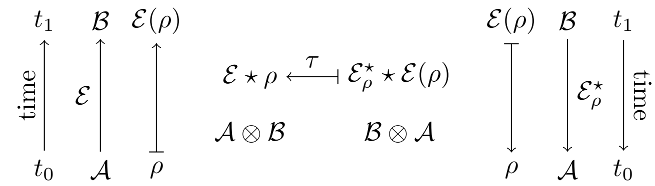

FIG. 1. If $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ is viewed as a CPTP map describing some dynamics from initial time t_0 to final time t_1, then this figure depicts the two states over time associated with $(\mathcal{E}, \rho)$ and a Bayesian inverse. If $\mathcal{E} \star \rho$ is interpreted as having a time orientation $t_0 \to t_1$, then $\mathcal{E}^{\star}_{\rho} \star \mathcal{E}(\rho)$ has time orientation $t_1 \to t_0$. The quantum time-reversal map τ simultaneously reverses the orientation of time and switches the two factors so that the resulting elements can be compared on an equal footing. Bayes' rule says that these two elements are the same. In this way, τ embodies a fundamental time-reversal symmetry for states over time associated with any input process and state $(\mathcal{E}, \rho)$ that admits a Bayesian inverse (see Ref. [44] for a closely related inferential form of time-reversal symmetry).

Furthermore, our formulation of Bayes' rule using our quantum time-reversal map τ solves several open questions in the literature. First, it resolves a puzzle posed by Leifer and Spekkens at the end of Ref. [25, Section VII.B.1], where they observe that using γ alone is not sufficient to provide enough symmetry to relate states over time in the forward and backward time directions. By adding a dagger, we solve this problem and restore the symmetry. Second, we show how Tsang's Bayes' rules, obtained from certain inner products [32], are derived from a certain class of state-over-time functions and our Bayes' rule. This answers Tsang's open remark or question (posed at the end of Ref. [32, Section III.B]) in regard to the relationship between generalized conditional expectations and states over time. Moreover, combining these two previous points yields that our Bayes' rule provides a possible answer to the question of Baez on the relationship between time reversal and the inner product in quantum theory [73]. In particular, our present work combined with Ref. [44] suggests that retrodiction based upon our quantum Bayes' rule may provide a mathematically precise relationship between time reversal and the inner product that is more robust than the ordinary adjoint operation on quantum channels [45–47,73–75].

III. FIRST EXAMPLES

The usefulness of a definition is illustrated through its examples. In what follows, we first justify the significance of the classical-limit axiom by considering not only the case of commutative algebras but also bistochastic channels (between possibly noncommutative algebras).

We then provide an example of a state-over-time function where the classical-limit axiom fails by removing all correlations. Subsequently, we consider the Leifer-Spekkens state over time [25]. From there, we describe two measurement scenarios. First, we show how prepare-evolve-measure scenarios have an inferential time-reversal symmetry involving our Bayes' rule. Second, we show how the state-update rule due to measurement is an instantiation of our Bayes' rule.

A. The classical case

The following example provides some motivation and justification for the importance of the classical-limit axiom. Consider the special case where $\mathcal{A} = \mathbb{C}^X \equiv \bigoplus_{x\in X}\mathbb{C}$ and $\mathcal{B} = \mathbb{C}^Y \equiv \bigoplus_{y\in Y}\mathbb{C}$, with X and Y finite sets. Then, a state ρ on $\mathcal{A}$ corresponds to a probability distribution $\{p_x\}$ on X, and a CPTP map $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ corresponds to a stochastic map from X to Y with conditional probabilities denoted by $\mathcal{E}_{yx}$.

In such a case, we have $[\mathscr{D}[\mathcal{E}], \rho \otimes 1_{\mathcal{B}}] = 0$, since the algebras $\mathcal{A}$ and $\mathcal{B}$ are both commutative. Because the associated channel state $\mathscr{D}[\mathcal{E}]$ is the element of $\mathcal{A}\otimes\mathcal{B} \cong \bigoplus_{x,y}\mathbb{C}$, the (x, y) component of which is $\mathcal{E}_{yx}$, it follows that $(\rho\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}]$ has an (x, y) component given by $p_x\mathcal{E}_{yx}$, which we refer to as the *classical state over time*. Write $q_y := \sum_{x\in X}\mathcal{E}_{yx}p_x$ as the probability distribution on Y corresponding to $\mathcal{E}(\rho)$. A Bayesian inverse of $(\mathcal{E}, \rho)$ is therefore a CPTP map $\mathcal{E}^\star_\rho : \mathcal{B} \to \mathcal{A}$, with corresponding conditional probabilities written as $(\mathcal{E}^\star_\rho)_{xy}$, such that

$$\mathcal{E}_{yx}p_x = (\mathcal{E}^\star_\rho)_{xy}q_y, \tag{17}$$

which is the classical Bayes' rule [cf. Eq. (1)].

If $\mathcal{E}$ is viewed as a genuine stochastic process, the physical intuition behind the term $\mathcal{E}_{yx}p_x$ is that it describes the predictive probability of first measuring x and then measuring y after the system has undergone the evolution described by $\mathcal{E}$ (note that this probability does *not*, in general, equal p_xq_y, which would require the random variables associated with X and Y to be independent and/or uncorrelated). Conversely, $(\mathcal{E}^\star_\rho)_{xy}q_y$ describes the retrodictive probability of measuring y and deducing that x has preceded it in the course of evolution through the inference map $\mathcal{E}^\star_\rho$ [5].

Of course, the classical-limit axiom covers far more cases than this, such as when the algebras $\mathcal{A}$ and $\mathcal{B}$ are not necessarily commutative and yet the commutativity condition still holds. One example is the case of *unital* CPTP maps, which includes bistochastic matrices and which is described in Sec. III B.

B. Bistochastic channels and time-reversal symmetry

Let $\star$ be any state-over-time function that satisfies the classical-limit axiom. Take $\mathcal{A} = \mathbb{M}_m$, $\mathcal{B} = \mathbb{M}_n$ and let $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ be a *unital* quantum channel (sometimes called a *bistochastic channel*), i.e., $\mathcal{E}(\mathbb{1}_m) = \mathbb{1}_n$. If $\rho = \mathbb{1}_m/m$ then, the associated state over time is always given by $\mathcal{E}\star\rho = (1/m)\mathscr{D}[\mathcal{E}]$. Physically, $\mathcal{E}$ describes a stochastic evolution that leaves the infinite-temperature-limit Gibbs state invariant (it is a consequence of unitality and the trace-preserving condition of $\mathcal{E}$ that $m = n$). If $\mathcal{A}$ and $\mathcal{B}$ were commutative algebras instead of matrix algebras, then $\mathcal{E}$ would be a doubly stochastic matrix (as the matrix is necessarily square, since $\mathcal{A} \cong \mathcal{B}$).

Hence, when $\mathcal{E}$ is a unital quantum channel and the prior is the uniform density matrix $\rho = \mathbb{1}_m/m$, then a Bayesian inverse $\mathcal{E}^\star_\rho$ must satisfy the equation

$$\mathscr{D}[\mathcal{E}^*] = \gamma\big(\mathscr{D}[\mathcal{E}]\big) = \gamma\big(\mathscr{D}[\mathcal{E}]^\dagger\big) = \mathscr{D}[\mathcal{E}^\star_\rho], \tag{18}$$

where the first equality holds by Lemma 2 in Appendix A, the second equality holds because $\mathscr{D}[\mathcal{E}]$ is self-adjoint, and the third equality holds by our definition of Bayes' rule. In other words, $\mathcal{E}^\star_\rho = \mathcal{E}^*$, which provides some justification for the notation that we use for a Bayesian inverse.

More importantly, this result has significant implications toward our understanding of time-reversal symmetry in quantum theory. Indeed, it has often been argued that unital quantum channels are the only channels for which a canonical notion of time reversal is possible [45,46] and, in such a case, the canonical time reversal is provided by the Hilbert-Schmidt adjoint. Such a claim, however, is at odds with classical probability theory, as not all classical channels are bistochastic and yet they have a well-defined notion of time-reversal symmetry furnished by the classical Bayes' rule. It then seems plausible that there may exist a more general notion of time reversal for quantum channels, which not only agrees with the Bayesian inverse for classical channels but also reduces to the Hilbert-Schmidt adjoint for bistochastic channels. And while no-go theorems have recently been proved in Refs. [45,46] stating that there is no such notion of time reversal for arbitrary quantum channels, it has recently been shown in Ref. [44] via an explicit construction that time reversal *is* possible for *all* quantum channels once a prior state is incorporated into the data of the theory (the explicit construction is discussed in more detail later in this work). The perspective gained from Ref. [44] is that when considering time reversal in quantum theory, one should be working in the category of *states* on multimatrix algebras and state-preserving channels, as opposed to the category of channels on their own.

We can therefore provide a possible explanation as to why bistochastic channels are often viewed as the only reversible operations in quantum theory. Indeed, our results show that there is a *unique* Bayesian inverse for any bistochastic channel with the uniform prior and this result is *independent* of the channel and a choice of a state-over-time function (as long as the state-over-time function satisfies the classical-limit axiom). Moreover, this *canonical* Bayesian inverse is, in fact, the Hilbert-Schmidt adjoint of the original channel, which agrees with the standard time-reversal map when the channel is unitary. However, the Hilbert-Schmidt adjoint is *rarely* a Bayesian inverse with respect to a state-over-time function when one has a *nonuniform* prior and an *arbitrary* quantum channel [8] and it is precisely this more general situation in which one needs additional input to specify time-reversal symmetry and Bayesian inverses and that input is the (noncanonical) choice of a state-over-time function.

As such, one might hope to find a state-over-time function whose associated Bayesian inverses simultaneously extend both classical Bayesian inversion for arbitrary stochastic channels and the Hilbert-Schmidt adjoint for all unital quantum channels. We consider such examples soon, but first, we give some examples of states over time that forget correlations and entanglement.

C. The uncorrelated state over time

The assignment

$$(\mathcal{E},\rho)\overset{\star}{\mapsto}\rho\otimes\mathcal{E}(\rho) \tag{19}$$

is a state-over-time function, called the *uncorrelated state over time*, that is positive (and hence also Hermitian) and process linear. However, it is not state linear, it is not associative [76], and, most importantly, it does not satisfy the classical-limit axiom.

A Bayesian inverse $\mathcal{E}^{\star}_{\rho}$ must satisfy the equation

$$\rho\otimes\mathcal{E}(\rho)=\mathcal{E}^{\star}_{\rho}\big(\mathcal{E}(\rho)\big)\otimes\mathcal{E}(\rho). \tag{20}$$

Thus, *any* CPTP map $\mathcal{E}^{\star}_{\rho}:\mathcal{B}\to\mathcal{A}$ such that $\mathcal{E}^{\star}_{\rho}\big(\mathcal{E}(\rho)\big)=\rho$ is a Bayesian inverse of $(\mathcal{E},\rho)$. There are many CPTP maps that satisfy this condition, one of which is simply the map that sends B to $\mathrm{tr}(B)\rho$. This, however, is not a very good candidate for Bayesian inversion, since it loses the information of most other states and essentially ignores the evolution $\mathcal{E}$ in its description. Such a state over time therefore treats ρ and $\mathcal{E}(\rho)$ as independent and, as such, it does not encode any correlations between ρ and $\mathcal{E}(\rho)$.

Furthermore, although there might be a better choice among the many state-preserving CPTP maps that satisfy Bayes' rule in this example, this state over time does not offer us any guide as to *which* of those CPTP maps to choose unless we impose further constraints. As such, the uncorrelated state-over-time function does not provide a robust formulation of Bayesian inversion.

D. The separable compound state and the quantum state marginal problem

There is an improvement on the uncorrelated state over time due to Ohya [52], which is henceforth called *Ohya's compound state over time*, following similar terminology in Refs. [52,53]. For matrix algebras $\mathcal{A}=\mathbb{M}_m$ and $\mathcal{B}=\mathbb{M}_n$, it is defined as follows [77].

Given any state $\rho\in\mathcal{A}$, let $\rho=\sum_\alpha\lambda_\alpha P_\alpha$ be the spectral decomposition, where each P_α is the projection onto the λ_α eigenspace. Then, for any $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$, set

$$\mathcal{E}\star\rho=\sum_\alpha\lambda_\alpha P_\alpha\otimes\mathcal{E}\left(\frac{P_\alpha}{\mathrm{tr}(P_\alpha)}\right). \tag{21}$$

This then defines a state over time that is positive and process linear and therefore gives a solution to the marginal state problem. It is neither state linear nor does it satisfy our classical-limit axiom. However, we note that if the nonzero eigenvalues of ρ have multiplicity 1, then the classical-limit axiom *does* hold for $(\mathcal{E},\rho)$ for arbitrary $\mathcal{E}\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})$ such that $\big[\rho\otimes 1_{\mathcal{B}},\mathscr{D}[\mathcal{E}]\big]=0$. Although the compound state over time is not completely uncorrelated, it is still *separable* [78] and hence seems unlikely to keep track of quantum *entanglement* over time.

E. The causal states of Leifer and Spekkens and the Petz recovery map

The assignment sending $(\mathcal{E},\rho)\in\mathrm{CPTP}(\mathcal{A},\mathcal{B})\times\mathcal{S}(\mathcal{A})$ to

$$\big(\sqrt{\rho}\otimes 1_{\mathcal{B}}\big)\mathscr{D}[\mathcal{E}]\big(\sqrt{\rho}\otimes 1_{\mathcal{B}}\big) \tag{22}$$

is called the *Leifer-Spekkens state over time* [20,21,25,37,38]. The Leifer-Spekkens state-over-time function is process linear, Hermitian, locally positive, and satisfies the classical-limit axiom, but it is not in general positive or associative [25,38]. One might argue that the reason for the lack of positivity is because we choose to use the Jamiołkowski channel state $\mathscr{D}[\mathcal{E}]$ as opposed to the *Choi matrix* for $\mathcal{E}$ [79]. However, the Choi matrix suffers from a few problems in our context. One is that it depends on a choice of basis and, hence, it would require additional data to specify a general state over time. In addition, if we had used the Choi matrix, the marginal density matrix $\mathrm{tr}_{\mathcal{A}}(\mathcal{E}\star\rho)$ would not be $\mathcal{E}(\rho)$, as it would be $\mathcal{E}(\rho^{\mathrm{T}})$, where $^{\mathrm{T}}$ denotes the transpose with respect to that chosen basis [31] (repeated eigenvalues prevent one from choosing a canonical basis of eigenvectors for ρ as is done in Ref. [31] for their construction of a positive compound state).

A Bayesian inverse $\mathcal{E}^{\star}_{\rho}$ to the Leifer-Spekkens state-over-time function must satisfy the equation

$$\begin{aligned}\gamma\Big(\big(\sqrt{\rho}\otimes 1_{\mathcal{B}}\big)\mathscr{D}[\mathcal{E}]\big(\sqrt{\rho}\otimes 1_{\mathcal{B}}\big)\Big)\\ =\left(\sqrt{\mathcal{E}(\rho)}\otimes 1_{\mathcal{A}}\right)\mathscr{D}\big[\mathcal{E}^{\star}_{\rho}\big]\left(\sqrt{\mathcal{E}(\rho)}\otimes 1_{\mathcal{A}}\right),\end{aligned} \tag{23}$$

which has the unique solution

$$\mathscr{D}[\mathcal{E}_\rho^\star] = \left(\sqrt{\mathcal{E}(\rho)^{-1}} \otimes \sqrt{\rho}\right) \mathscr{D}[\mathcal{E}^*] \left(\sqrt{\mathcal{E}(\rho)^{-1}} \otimes \sqrt{\rho}\right) \tag{24}$$

whenever $\mathcal{E}(\rho)$ is nonsingular. Solving this using Lemma 2 in Appendix A yields

$$\mathcal{E}_\rho^\star = \mathrm{Ad}_{\rho^{1/2}} \circ \mathcal{E}^* \circ \mathrm{Ad}_{\mathcal{E}(\rho)^{-1/2}}, \tag{25}$$

which is the *Petz recovery map* (sometimes called the *transpose channel*) [51,57,80].

This map has appeared in the context of sufficient quantum statistics and equality conditions for relative entropy [80–83], approximate quantum error correction [84,85], operational time reversal in quantum theory [44,86,87], earlier approaches toward extending Bayesian inversion to the quantum setting [20,21], coarse-grained or observational entropy [88], quantum fluctuation relations [89,90], the renormalization group in quantum field theory [91], entanglement wedge reconstruction and proposals for a resolution of the black-hole information paradox [26,92–94], and many other contexts. The Petz recovery map has recently been shown to define the only presently known quantum retrodiction functor [44], which, in particular, says that it is *compositional* in the sense that

$$(\mathcal{F} \circ \mathcal{E})_\rho^\star = \mathcal{E}_\rho^\star \circ \mathcal{F}_{\mathcal{E}(\rho)}^\star \tag{26}$$

for a composable pair $\mathcal{A} \xrightarrow{\mathcal{E}} \mathcal{B} \xrightarrow{\mathcal{F}} \mathcal{C}$ of CPTP maps and an initial state $\rho \in \mathcal{S}(\mathcal{A})$.

1. The quantum Bayes' rule of Fuchs

In the case of a POVM $\mathcal{E} : \mathbb{M}_n \to \mathbb{C}^X$ given by $\mathcal{E} = \bigoplus_{x\in X} \mathrm{tr}(M_x\,\cdot\,)$ and prior density matrix $\rho \in \mathbb{M}_n$, our quantum Bayes' rule for the Leifer-Spekkens state-over-time function yields (by calculations similar to the above)

$$\bigoplus_{x\in X} p_x\rho_x = \bigoplus_{x\in X} \sqrt{\rho} M_x \sqrt{\rho}, \tag{27}$$

where $\rho_x = \mathcal{E}_\rho^\star(\delta_x)$ and $p_x = \mathrm{tr}(M_x\rho)$ for all $x \in X$. The above equation then yields the quantum Bayes' rule of Fuchs [17, Section 5], namely,

$$\rho_x = \frac{\sqrt{\rho} M_x \sqrt{\rho}}{p_x}, \tag{28}$$

which has also been derived using the formalism of *conditional states* in Ref. [25].

F. The *t*-rotated family and rotated Petz recovery maps

For each $t \in \mathbb{R}$, the assignment sending $(\mathcal{E}, \rho)$ to

$$(\rho^{1/2-it} \otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{1/2+it} \otimes 1_{\mathcal{B}}) \tag{29}$$

defines a state-over-time function, called the *t-rotated state over time*, that is process linear, locally positive, and satisfies the classical-limit axiom for all $t \in \mathbb{R}$. The Bayesian inverse $\mathcal{E}_\rho^\star$ in this case is given by

$$\mathcal{E}_\rho^\star = \mathrm{Ad}_{\rho^{1/2-it}} \circ \mathcal{E}^* \circ \mathrm{Ad}_{\mathcal{E}(\rho)^{-1/2+it}}, \tag{30}$$

which is the *rotated Petz recovery map* [68]. This map, and averaged versions of it, have become especially important in the context of information recovery and the strengthening of data-processing inequalities [67–70,95,96].

More generally, for each $\rho \in \mathcal{S}(\mathcal{A})$, choose a unitary $U_\rho \in \mathcal{A}$ such that $[U_\rho, \rho] = 0$. Then, the assignment sending $(\mathcal{E}, \rho)$ to

$$\left(U_\rho^\dagger \rho^{1/2} \otimes 1_{\mathcal{B}}\right)\mathscr{D}[\mathcal{E}]\left(\rho^{1/2} U_\rho \otimes 1_{\mathcal{B}}\right) \tag{31}$$

defines a state over time that is process linear, locally positive, and satisfies the classical-limit axiom. It is referred to as the *Sutter-Tomamichel-Harrow (STH) state over time* based on its appearances in Refs. [44,96]. The Bayesian inverse $\mathcal{E}_\rho^\star$ in this case is given by

$$\mathcal{E}_\rho^\star = \mathrm{Ad}_{U_\rho^\dagger \rho^{1/2}} \circ \mathcal{E}^* \circ \mathrm{Ad}_{U_{\mathcal{E}(\rho)} \mathcal{E}(\rho)^{-1/2}}, \tag{32}$$

which is a generalization of the rotated Petz recovery map (the special case of the rotated Petz is $U_\rho = \rho^{it}$ for some $t \in \mathbb{R}$).

Although the above rotated Petz recovery maps can be obtained by a judicious choice of state-over-time function, an open question is: what state-over-time function has the universal recovery map of Junge *et al.* [69,95] as the Bayesian inverse? Or, more generally: what state-over-time functions have averaged rotated Petz recovery maps, as defined in Ref. [44], as their Bayesian inverses? Since the universal recovery map of Refs. [69,95] has been claimed to provide a quantum generalization of Bayes' rule in Ref. [26] (based on results from Ref. [97]), we expect that this arises within our framework.

IV. MEASUREMENT

Measurement in quantum mechanics involves hybrid classical-quantum systems. Therefore, our definitions of states over time and Bayesian inverses should specialize to such settings. In this section, we illustrate this with two key examples. The first example involves preparation, evolution, and then measurement. The second example is the state-update rule associated with the measurement of a quantum system in terms of quantum instruments.

A. Prepare-evolve-measure scenarios

Let X be a finite set, let $\mathcal{P} : \mathbb{C}^X \to \mathbb{M}_n$ be a CPTP map, and set $\rho = \mathcal{P}(p)$ for some probability distribution $p \in \mathbb{C}^X$. Such data determine a preparation of the state ρ given by

$$\rho = \sum_{x \in X} p_x \rho_x, \tag{33}$$

where $\rho_x = \mathcal{P}(\delta_x)$. Now suppose that the state ρ is to evolve according to a CPTP map $\mathcal{E} : \mathbb{M}_n \to \mathbb{M}_m$ and let $\mathcal{M} : \mathbb{M}_m \to \mathbb{C}^Y$ be a POVM on the output $\mathbb{M}_m$ of $\mathcal{E}$, so that $\mathcal{M} = \bigoplus_{y\in Y} \operatorname{tr}(M_y\ \cdot\ )$ for a collection of positive operators M_y summing to the identity, indexed by some finite set Y. We then refer to the four-tuple $(p, \mathcal{P}, \mathcal{E}, \mathcal{M})$ as a *prepare-evolve-measure (PEM) scenario*.

Given a PEM scenario $(p, \mathcal{P}, \mathcal{E}, \mathcal{M})$, we can naturally define a classical channel $f : \mathbb{C}^X \to \mathbb{C}^Y$ given by $f = \mathcal{M} \circ \mathcal{E} \circ \mathcal{P}$ (see Fig. 2), which we refer to as the *classical dynamics* of the PEM scenario $(p, \mathcal{P}, \mathcal{E}, \mathcal{M})$. The conditional probabilities f_{yx} associated with f are then interpreted as the predictive probability of measuring outcome y via the POVM $\mathcal{M}$ given that the state ρ_x was prepared and then evolved under the channel $\mathcal{E}$. Conversely, the classical Bayesian inverse $g : \mathbb{C}^Y \to \mathbb{C}^X$ of f is given by

$$g_{xy} = \frac{f_{yx} q_y}{p_x}, \tag{34}$$

where $q_y := \operatorname{tr}\big(M_y \mathcal{E}(\rho)\big)$ gives the probability of obtaining outcome y assuming the probabilistic preparation from p and $\mathcal{P}$. The conditional probability g_{xy} may be interpreted as the retrodictive probability that ρ_x was prepared given a measurement outcome of y. Furthermore, if we set $\sigma = \mathcal{E}(\rho)$ and let $\star$ be the Leifer-Spekkens state-over-time function, then $(q, \mathcal{M}^\star_\sigma, \mathcal{E}^\star_\rho, \mathcal{P}^\star_p)$ is a PEM scenario and compositionality of the Petz recovery map implies that g is the classical dynamics associated with the PEM scenario $(q, \mathcal{M}^\star_\sigma, \mathcal{E}^\star_\rho, \mathcal{P}^\star_p)$, which may be viewed as an inferential time-reverse of the PEM scenario $(p, \mathcal{P}, \mathcal{E}, \mathcal{M})$ [98]. This makes mathematically precise a sense in which measurement and preparation are time reversals of one another, even when there is nonunitary evolution between the prepared state and the measured state.

In the special case where $X = \{1, \ldots, n\}$, $Y = \{1, \ldots, m\}$ and the sets $\{|i\rangle\langle i| \equiv P_i := \mathcal{P}(\delta_i)\}_{i\in X}$ and $\{|k\rangle\langle k| \equiv M_k\}_{k\in Y}$ are chosen to be the orthogonal rank-1 projections associated with spectral decompositions of ρ and σ, respectively, then the Bayesian inverses $\mathcal{P}^\star_p$ and $\mathcal{M}^\star_\sigma$ are *independent* of the state-over-time function provided that it satisfies the classical-limit axiom. Indeed, the Bayesian inverses are simply given by the Hilbert-Schmidt adjoints in this case, so that $\mathcal{M}^\star_\sigma = \mathcal{M}^*$ and $\mathcal{P}^\star_p = \mathcal{P}^*$. However, since $\mathcal{E}$ is not assumed to have a particularly simple structure with respect to these eigenstates, its Hilbert-Schmidt adjoint is not a suitable "reverse" operation in general. By using the Leifer-Spekkens state-over-time function, we find such a suitable "reverse" operation $\mathcal{E}^\star_\rho$ for which the conditional probabilities $\langle i|\mathcal{E}^\star_\rho(M_k)|i\rangle$ indeed satisfy Bayes' rule (see Fig. 2), namely,

$$\langle k|\mathcal{E}(P_i)|k\rangle p_i = \langle i|\mathcal{E}^\star_\rho(M_k)|i\rangle q_k. \tag{35}$$

We note that this resolves an issue brought up by Barandes and Kagan in Ref. [100], where it is stated that the classical Bayes' theorem in this context does not hold in general due to the "generic irreversibility" of $\mathcal{E}$. So not only does our notion of Bayesian inversion restore the classical Bayes' theorem in such a context by providing a suitable time reversal of $\mathcal{E}$ but it does so for any PEM scenario, i.e., for arbitrary preparations, evolutions, and measurements.

An interesting result in this context, which is a reformulation in the language of states over time of a result first proved by Leifer [20], is that the classical state over time $f \star p$ associated with the classical dynamics f of the PEM scenario $(p, \mathcal{P}, \mathcal{E}, \mathcal{M})$ may be obtained from the *quantum* state over time $\mathcal{E} \star \rho$. Indeed, one may show that for all $(x, y) \in X \times Y$,

$$(f \star p)(x, y) = \operatorname{tr}\big((N_x \otimes M_y)(\mathcal{E} \star \rho)\big), \tag{36}$$

where N_x and M_y are the positive operators corresponding to the POVMs $\mathcal{N} := \mathcal{P}^\star_p$ and $\mathcal{M}$, respectively.

$$\begin{tikzcd}
\mathbb{M}_n \arrow[r, "\mathcal{E}"] & \mathbb{M}_m \arrow[d, "\mathcal{M}"] \\
\mathbb{C}^X \arrow[u, "\mathcal{P}"] \arrow[r, "f"'] & \mathbb{C}^Y
\end{tikzcd}
\qquad
\begin{tikzcd}
\rho \arrow[r, mapsto, "\mathcal{E}"] & \sigma \arrow[d, mapsto, "\mathcal{M}"] \\
p \arrow[u, mapsto, "\mathcal{P}"] & q
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbb{M}_n \arrow[d, "\mathcal{P}^\star_p"'] & \mathbb{M}_m \arrow[l, "\mathcal{E}^\star_\rho"'] \\
\mathbb{C}^X & \mathbb{C}^Y \arrow[l, "g"] \arrow[u, "\mathcal{M}^\star_\sigma"']
\end{tikzcd}$$

FIG. 2. Given a PEM scenario $(p, \mathcal{P}, \mathcal{E}, \mathcal{M})$, we can use the latter three maps to define a classical channel $f : \mathbb{C}^X \to \mathbb{C}^Y$ via $f = \mathcal{M} \circ \mathcal{E} \circ \mathcal{P}$. This equality is expressed by saying that the diagram on the left *commutes* (this terminology of a *commuting diagram* should not be confused with commutativity of a set of operators). The diagram in the middle depicts the evolution of the state p along preparation to ρ, evolution to σ, and measurement to q. One can then use the probability p and the classical channel f to define the (classical) Bayesian inverse $g : \mathbb{C}^Y \to \mathbb{C}^X$. However, one can compute *another* map $\mathbb{C}^Y \to \mathbb{C}^X$ via $\mathcal{P}^\star_p \circ \mathcal{E}^\star_\rho \circ \mathcal{M}^\star_\sigma$ by using the Leifer-Spekkens state over time to Bayesian invert each of the pairs $(\mathcal{M}, \sigma)$, $(\mathcal{E}, \rho)$, and $(\mathcal{P}, p)$, where $\sigma = \mathcal{E}(\rho)$ and $\rho = \mathcal{P}(p)$, to arrive at the CPTP maps $\mathcal{M}^\star_\sigma$, $\mathcal{E}^\star_\rho$, and $\mathcal{P}^\star_p$ (note that the Bayesian inverse of a preparation is a measurement and vice versa, so that $(q, \mathcal{M}^\star_\sigma, \mathcal{E}^\star_\rho, \mathcal{P}^\star_p)$ defines another PEM scenario). The two stochastic channels g and $\mathcal{P}^\star_p \circ \mathcal{E}^\star_\rho \circ \mathcal{M}^\star_\sigma$, are equal, i.e., the diagram on the right commutes, because the Petz recovery map is compositional. This compositionality can also be viewed a quantum generalization of Jeffrey's probability kinematics [44,88,99].

Finally, we briefly mention that the quantum Bayes' rule of Schack, Brun, and Caves [16] can be viewed as a PEM scenario, when one appropriately extends our definitions to C^*-algebras. The need for this extension to C^*-algebras is because the probability p is replaced by a prior (Radon) probability measure on $X = \mathcal{S}(\mathcal{A})$, the compact Hausdorff space of states in a matrix algebra $\mathcal{A}$ [101,102]. Namely, $\mathbb{C}^X$ is replaced with $C(X)$, the C^*-algebra of continuous $\mathbb{C}$-valued functions on X, while the preparation $\mathcal{P} : C(X) \to \mathcal{A}$ is replaced by a canonical classical-to-quantum generalization of the *sampling map* from classical statistics [28,103]. Once a measurement outcome is obtained, the prior can then be updated. Such a scenario appears in the context of adaptive strategies for optimal quantum state determination [104,105]. The details of this will be expounded upon elsewhere.

B. Instruments, measurement, and the state-update rule

If $\sigma \in \mathcal{A}$ is an initial quantum state that undergoes measurement associated with some instrument $\mathcal{F} : \mathcal{A} \to \mathcal{B} \otimes \mathbb{C}^X$ and outcome $x \in X$ is measured, then the *state-update rule* dictates that the quantum state of the system becomes [13,17,40,54,58,106–108]

$$\sigma \mapsto \frac{\mathcal{F}_x(\sigma)}{\operatorname{tr}\big(\mathcal{F}_x(\sigma)\big)} \tag{37}$$

after the measurement has been performed. Here, $\mathcal{F}_x : \mathcal{A} \to \mathcal{B}$ is the CP map of the instrument associated with outcome x, so that the sum $\sum_{x\in X} \mathcal{F}_x$ defines a CPTP map from $\mathcal{A}$ to $\mathcal{B}$. It is often the case that $\mathcal{F}_x$ is assumed to be a Kraus rank-1 CP map, i.e., it can be written in the form $\mathcal{F}_x = \mathrm{Ad}_{V_x}$. The special case where the V_x form projection operators gives the Lüders–von Neumann measurement [12,39,109].

This state-update rule has often been called a quantum generalization of Bayes' rule [11,13–15,23,110,111]. Some have also argued that the state-update rule is *not* a quantum generalization Bayesian conditioning and that it is more a combination of belief propagation and conditioning on a measurement outcome [25, Sections V.A.2. and V.B.]. In the present section, we illustrate how one is inevitably led to the state-update rule from Bayesian inverses associated with *any* state-over-time function that satisfies the classical-limit axiom. In other words, our work supports the idea that the state-update rule is indeed a quantum generalization of Bayes' rule. Nevertheless, we still agree with Ref. [25] in the statement that this does not produce a *retrodictive* state, since the state-update rule does not tell us what the state of the quantum system might have been *before* the act of measurement. Furthermore, we find that the state-update rule is indeed conditioning on a measurement, where the concept of conditioning is made precise in terms of operator-algebraic conditional expectations [108,112,113].

To state our result precisely, we introduce additional notation besides what is specified above. First, set $\mathcal{E} : \mathcal{B} \otimes \mathbb{C}^X \to \mathbb{C}^X$ to be the partial trace map that traces out the $\mathcal{B}$ system, i.e., $\mathcal{E} = \operatorname{tr}_{\mathcal{B}} \equiv \operatorname{tr} \otimes \mathrm{id}_{\mathbb{C}^X}$. Let $\rho \in \mathcal{B} \otimes \mathbb{C}^X$ be any state, which we decompose as

$$\rho = \sum_{x\in X} \rho_x \otimes \delta_x, \tag{38}$$

where $\rho_x \in \mathcal{B}$ is some positive element for all $x \in X$ and δ_x is the unit vector the x component of which is 1 while other components are 0.

Proposition 1.—Let $\star$ be any state-over-time function satisfying the classical-limit axiom and assume that $\rho_x \neq 0$ for all $x \in X$. Then, using the notation of the previous paragraph, a Bayesian inverse of $(\mathcal{E}, \rho)$ with respect to $\star$ exists, is unique, and is given by

$$\mathcal{E}^\star_\rho(\delta_x) = \frac{\rho_x}{\operatorname{tr}(\rho_x)} \otimes \delta_x \tag{39}$$

for all $x \in X$.

We first prove this result before discussing the relevance to the state-change associated with a measurement.

Proof.—Explicitly computing $\mathscr{D}[\mathcal{E}] \in (\mathcal{B} \otimes \mathbb{C}^X) \otimes \mathbb{C}^X$ (the parentheses are used to distinguish the input and output algebras) gives

$$\mathscr{D}[\mathcal{E}] = \sum_{x\in X} 1_{\mathcal{B}} \otimes \delta_x \otimes \delta_x. \tag{40}$$

Therefore, $[\rho \otimes 1_{\mathbb{C}^X}, \mathscr{D}[\mathcal{E}]] = 0$, so that

$$\mathcal{E} \star \rho = (\rho \otimes 1_{\mathbb{C}^X})\mathscr{D}[\mathcal{E}] = \sum_{x\in X} \rho_x \otimes \delta_x \otimes \delta_x \tag{41}$$

and

$$\tau(\mathcal{E} \star \rho) = \sum_{x\in X} \delta_x \otimes \rho_x \otimes \delta_x. \tag{42}$$

Furthermore, since $\mathcal{E}(\rho) \in \mathbb{C}^X$ lives in a commutative C^*-algebra, $\mathcal{E}(\rho) \otimes 1_{\mathcal{B}\otimes\mathbb{C}^X}$ necessarily commutes with every element of $\mathbb{C}^X \otimes \mathcal{B} \otimes \mathbb{C}^X$. In particular, $\big[\mathcal{E}(\rho) \otimes 1_{\mathcal{B}\otimes\mathbb{C}^X}, \mathscr{D}[\mathcal{E}^\star_\rho]\big] = 0$, so that

$$\begin{aligned}\mathcal{E}^\star_\rho \star \mathcal{E}(\rho) &= \big(\mathcal{E}(\rho) \otimes 1_{\mathcal{B}} \otimes 1_{\mathbb{C}^X}\big)\mathscr{D}[\mathcal{E}^\star_\rho] \\ &= \left(\sum_{x\in X} \operatorname{tr}(\rho_x)\delta_x \otimes 1_{\mathcal{B}} \otimes 1_{\mathbb{C}^X}\right) \mathscr{D}[\mathcal{E}^\star_\rho].\end{aligned} \tag{43}$$

Bayes' rule $\tau(\mathcal{E} \star \rho) = \mathcal{E}^\star_\rho \star \mathcal{E}(\rho)$ then yields

$$\mathscr{D}[\mathcal{E}^\star_\rho] = \sum_{x\in X} \delta_x \otimes \frac{\rho_x}{\operatorname{tr}(\rho_x)} \otimes \delta_x, \tag{44}$$

so that the Bayesian inverse $\mathcal{E}^{\star}_{\rho}$ is given by the map that sends δ_x, representing the outcome x, to

$$\mathcal{E}^{\star}_{\rho}(\delta_x) = \frac{\rho_x}{\mathrm{tr}(\rho_x)} \otimes \delta_x. \tag{45}$$

■

Using this result, we can relate the state-update rule to state-preserving conditional expectations and Bayesian inverses associated with a state-over-time function satisfying the classical-limit axiom.

Theorem 1.—Let $\sigma \in \mathcal{A}$ be a state, $\mathcal{F} : \mathcal{A} \to \mathcal{B} \otimes \mathbb{C}^X$ an instrument, and $\mathcal{E} : \mathcal{B} \otimes \mathbb{C}^X \to \mathbb{C}^X$ the partial trace, and set $\rho = \mathcal{F}(\sigma)$. Suppose that $\mathrm{tr}(\mathcal{F}_x(\sigma)) \neq 0$ for all $x \in X$. Furthermore, define the following maps:

(1) Let $\Psi : \mathbb{C}^X \to \mathcal{B} \otimes \mathbb{C}^X$ be the state-update map defined as the unique linear extension of

$$\Psi(\delta_x) = \frac{\mathcal{F}_x(\sigma)}{\mathrm{tr}\big(\mathcal{F}_x(\sigma)\big)} \otimes \delta_x. \tag{46}$$

(2) Let $\mathcal{E}^{\star}_{\rho} : \mathbb{C}^X \to \mathcal{B} \otimes \mathbb{C}^X$ be the Bayesian inverse of $(\mathcal{E}, \rho)$ associated with *any* state-over-time function $\star$ satisfying the classical-limit axiom (cf. Proposition 1).

(3) Let $\Omega : \mathbb{C}^X \to \mathcal{B} \otimes \mathbb{C}^X$ be the Hilbert-Schmidt adjoint of a state-preserving conditional expectation associated with the map $\mathcal{E}$ and state ρ, i.e., $\Omega(\mathcal{E}(\rho)) = \rho$ and $\mathcal{E} \circ \Omega = \mathrm{id}_{\mathbb{C}^X}$ (cf. Ref. [114]).

Then, $\Omega = \Psi = \mathcal{E}^{\star}_{\rho}$.

Proof.—The equivalence between items 1 and 2, and hence the equality $\Psi = \mathcal{E}^{\star}_{\rho}$, follows from Proposition 1 by setting $\rho_x = \mathcal{F}_x(\sigma)$. As for the relationship to state-preserving conditional expectations, first note that

$$\begin{aligned}\Psi\big(\mathcal{E}(\rho)\big) &= \Psi\left(\sum_x \mathrm{tr}(\rho_x)\delta_x\right) = \sum_{x\in X} \mathrm{tr}(\rho_x)\Psi(\delta_x) \\ &= \sum_{x\in X} \mathrm{tr}(\rho_x)\left(\frac{\rho_x}{\mathrm{tr}(\rho_x)} \otimes \delta_x\right) = \sum_x \rho_x \otimes \delta_x \\ &= \rho.\end{aligned} \tag{47}$$

Second, we have

$$(\mathcal{E} \circ \Psi)(\delta_x) = \mathcal{E}\left(\frac{\rho_x}{\mathrm{tr}(\rho_x)} \otimes \delta_x\right) = \frac{\mathrm{tr}(\rho_x)}{\mathrm{tr}(\rho_x)}\delta_x = \delta_x \tag{48}$$

for each $x \in X$. By linear extension, this shows that $\mathcal{E} \circ \Psi = \mathrm{id}_{\mathbb{C}^X}$. Thus, Ψ satisfies the definition of a state-preserving conditional expectation. By the uniqueness of state-preserving conditional expectations for faithful states [27,57,113], $\Omega = \Psi = \mathcal{E}^{\star}_{\rho}$. ■

Note that by definition of Ψ, if one acquires soft evidence from the measurement in the form of a probability distribution $r \in \mathbb{C}^X$, then

$$\Psi(r) = \sum_{x\in X} \frac{r_x \mathcal{F}_x(\sigma)}{\mathrm{tr}\big(\mathcal{F}_x(\sigma)\big)} \otimes \delta_x. \tag{49}$$

Note that the $\otimes\delta_x$ term is merely used as a book-keeping device to separate the possible state-updates depending on each outcome. Namely, by tracing out over $\mathbb{C}^X$, one obtains the barycenter

$$(\mathrm{id}_{\mathcal{B}} \otimes \mathrm{tr}_{\mathbb{C}^X})\big(\Psi(r)\big) = \sum_{x\in X} \frac{r_x \mathcal{F}_x(\sigma)}{\mathrm{tr}\big(\mathcal{F}_x(\sigma)\big)} \tag{50}$$

as the updated density matrix in $\mathcal{B}$ after the measurement has been performed and the outcome is only given by soft evidence r. This is a quantum generalization of Jeffrey's rule [7,99]. Nevertheless, this is a special case of Bayesian inverses from our perspective due to the hybrid classical-quantum nature of the channels involved. The full quantum generalization of Bayes' rule is the main definition provided in this work. The point here is that the Bayesian inverse for *every* state-over-time function satisfying the classical-limit axiom *must* reproduce this special version of quantum Bayes' rule.

V. NONPOSITIVE BAYESIAN INVERSES

To incorporate even more instances of Bayes' rule, we extend state-over-time functions to a larger domain. One reason to do this is because there may exist solutions $\mathcal{E}^{\star}_{\rho}$ to Bayes' rule $\mathcal{E} \star \rho = \tau(\mathcal{E}^{\star}_{\rho} \star \mathcal{E}(\rho))$ that are not necessarily completely positive and these solutions may nevertheless have useful applications and/or interpretations. A second reason is to incorporate a wider variety of examples in the literature as instances of states over time and Bayes' rule. These include the two-state formalism [41,42,48], time-dependent correlators [49,50], generalized conditional expectations [32], and many more.

We first consider a solution $\mathcal{E}^{\star}_{\rho}$ to Bayes' rule for a state-over-time function where $\mathcal{E}^{\star}_{\rho}$ is not necessarily CP, but is †-*preserving*, in the sense that $\mathcal{E}^{\star}_{\rho}(B^\dagger) = (\mathcal{E}^{\star}_{\rho}(B))^\dagger$ for all inputs B. In this case, we allow more flexibility in our earlier definition of a state-over-time function by requesting that it is a family of functions each of the form $\star : \mathrm{HPTP}(\mathcal{A},\mathcal{B}) \times \mathcal{S}(\mathcal{A}) \to \mathcal{A} \otimes \mathcal{B}$, where $\mathrm{HPTP}(\mathcal{A},\mathcal{B})$ denotes the set of †-preserving and trace-preserving (HPTP) maps from $\mathcal{A}$ to $\mathcal{B}$. An HPTP solution $\mathcal{E}^{\star}_{\rho}$ to Bayes' rule is called a *Bayes map*. Our primary example is the Jordan-product state-over-time function.

A. The Jordan-product state over time

The *symmetric bloom* (or *Jordan product*) state-over-time function is defined on $(\mathcal{E}, \rho) \in \mathrm{HPTP}(\mathcal{A},\mathcal{B}) \times \mathcal{S}(\mathcal{A})$

by

$$\mathcal{E} \star \rho = \frac{1}{2}\{\rho \otimes 1_{\mathcal{B}}, \mathscr{D}[\mathcal{E}]\}, \tag{51}$$

where $\{X, Y\} = XY + YX$ denotes the Jordan product (anticommutator). The symmetric bloom state-over-time function is Hermitian, bilinear, associative, and satisfies the classical-limit axiom (these statements are all proved in Ref. [37]). The symmetric bloom has recently been shown to define a quantum state-over-time function that bypasses the no-go theorem of Ref. [38], which claims the nonexistence of a state-over-time function satisfying these properties. While the no-go theorem of Ref. [38] is mathematically correct, it assumes a larger domain for state-over-time functions than is physically necessary. The symmetric bloom then bypasses the statement in Ref. [38] by being defined on a smaller, more physically relevant, domain.

A Bayes map $\mathcal{E}^{\star}_{\rho}$ for the symmetric bloom must satisfy the equation

$$\tau\Big(\{\rho \otimes 1_{\mathcal{B}}, \mathscr{D}[\mathcal{E}]\}\Big) = \{\mathcal{E}(\rho) \otimes 1_{\mathcal{A}}, \mathscr{D}[\mathcal{E}^{\star}_{\rho}]\}, \tag{52}$$

which is equivalent to

$$\{1_{\mathcal{B}} \otimes \rho, \mathscr{D}[\mathcal{E}^{*}]\} = \{\mathcal{E}(\rho) \otimes 1_{\mathcal{A}}, \mathscr{D}[\mathcal{E}^{\star}_{\rho}]\} \tag{53}$$

because ρ is self-adjoint and $\mathcal{E}$ is CP. This is a linear-algebra problem of the form $B = \{A, X\}$, where A and B are known and X is desired.

To avoid cumbersome indices, we now restrict to the case where $\mathcal{A} = \mathbb{M}_m$ and $\mathcal{B} = \mathbb{M}_n$ are matrix algebras. In addition, to avoid a discussion of measure zero and noncommutative almost-everywhere equivalence [27], we restrict our attention to the cases where $\mathcal{E}(\rho)$ is nonsingular (has only nonzero eigenvalues). Let

$$\mathcal{E}(\rho) = \sum_{k=1}^{n} q_k |w_k\rangle\langle w_k| \tag{54}$$

be a spectral decomposition into one-dimensional projections (so that some of the q_k may repeat).

Using this and the above equation in terms of the Jordan product, one obtains

$$\mathcal{E}^{\star}_{\rho}\big(|w_k\rangle\langle w_l|\big) = (q_k + q_l)^{-1}\Big\{\rho, \mathcal{E}^{*}\big(|w_k\rangle\langle w_l|\big)\Big\}. \tag{55}$$

Since the $\{|w_k\rangle\}$ form a basis, this determines $\mathcal{E}^{\star}_{\rho}$ as a linear operator. In fact, $\mathcal{E}^{\star}_{\rho}$ is †-preserving because ρ is self-adjoint, the $\{q_k\}$ are real, the anticommutator is †-preserving, and $\mathcal{E}$ is †-preserving. It is not presently known to us what necessary and sufficient conditions guarantee complete positivity of $\mathcal{E}^{\star}_{\rho}$ in the case of the symmetric bloom.

1. The symmetric bloom approximates Leifer-Spekkens

Interestingly, the symmetric bloom state over time provides the linear approximation to the Leifer-Spekkens state over time near the maximally mixed state [37]. Let $\rho_0 := 1_{\mathcal{A}}/\mathrm{tr}(1_{\mathcal{A}})$ denote the uniform state in a multimatrix algebra $\mathcal{A}$, let $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ be a CPTP map, and let $\star_{\mathrm{LS}}$ and $\star_{\mathrm{J}}$ denote the Leifer-Spekkens and symmetric bloom state-over-time functions, respectively. Then for any $A \in \mathcal{A}^{\mathrm{sa}}$ such that $\mathrm{tr}(A) = 0$, one has

$$\mathcal{E} \star_{\mathrm{LS}} (\rho_0 + \epsilon A) - \mathcal{E} \star_{\mathrm{LS}} \rho_0 = \epsilon(\mathcal{E} \star_{\mathrm{J}} A) + \mathcal{O}(\epsilon^2) \tag{56}$$

for sufficiently small ϵ. Note that we linearly extend $\star_{\mathrm{J}}$ from $\mathcal{S}(\mathcal{A})$ to $\mathcal{A}^{\mathrm{sa}}$ in the second argument to make sense of the right-hand side of this identity. Equivalently,

$$\lim_{\epsilon \to 0}\left(\frac{\mathcal{E} \star_{\mathrm{LS}} (\rho_0 + \epsilon A) - \mathcal{E} \star_{\mathrm{LS}} \rho_0}{\epsilon}\right) = \mathcal{E} \star_{\mathrm{J}} A. \tag{57}$$

This, combined with the work of Ref. [37], shows that the symmetric bloom state over time maintains some of the physically relevant features of the Leifer-Spekkens state over time but, in addition, satisfies a larger number of convenient properties, including associativity. At present, we do not know what this says about the relationship between the Petz recovery map and the Bayes map for the symmetric bloom.

2. An operational interpretation of the symmetric bloom

In Refs. [49,115], the symmetric bloom state over time, in the special case where the channel is the identity, appears as the real part of the ideal two-point quantum correlator (for more details, see Sec. V D). The symmetric bloom state over time, when viewed as a function of its second argument, defines a map $\mathrm{id}_{\mathcal{A}} \star \,\cdot\, : \mathcal{S}(\mathcal{A}) \to \mathcal{A} \otimes \mathcal{A}$. This map is given an operational interpretation by decomposing it via a generalized Kraus decomposition into a quantum instrument, which in turn allows one to compute expectation values associated with the input state and instrument [49, Proposition 1]. A unique such decomposition into a quantum instrument can be chosen by specifying that the sum of the absolute values of the associated statistical errors is minimized [49, Proposition 2]. In this special case of the symmetric bloom, this unique decomposition is expressed in terms of symmetric and antisymmetric optimal cloners [49, Proposition 3]. In fact, Ref. [49] has proposed an optical experiment with polarized photons to operationally determine the symmetric bloom.

B. Bayes' rule and linear Bayes maps

We may even go beyond †-preserving maps by finding *linear* solutions to Bayes' rule. To make sense of this, we therefore need to extend state-over-time functions to be definable on arbitrary trace-preserving linear maps. We also find that the quantum time-reversal symmetry τ for states over time, given by the swap map γ and adjoint $\dagger$, is not enough for a robust Bayes' rule in this more general setting. In all that follows, let $\mathrm{TP}(\mathcal{A},\mathcal{B})$ denote the space of trace-preserving linear maps from $\mathcal{A}$ to $\mathcal{B}$.

Definition 4.—An *extended state-over-time function* associates every pair $(\mathcal{A},\mathcal{B})$ of multimatrix algebras with a map $\star : \mathrm{TP}(\mathcal{A},\mathcal{B}) \times \mathcal{S}(\mathcal{A}) \to \mathcal{A}\otimes\mathcal{B}$, the value of which on $(\mathcal{E},\rho)$ is denoted by $\mathcal{E}\star\rho$, such that $\star$ preserves marginals in the sense that

$$\mathrm{tr}_{\mathcal{B}}\left(\mathcal{E}\star\rho\right) = \rho \quad \text{and} \quad \mathrm{tr}_{\mathcal{A}}\left(\mathcal{E}\star\rho\right) = \mathcal{E}(\rho). \tag{58}$$

In such a case, the element $\mathcal{E}\star\rho \in \mathcal{A}\otimes\mathcal{B}$ is referred to as the *state over time* associated with $\star$ and the input $(\mathcal{E},\rho)$.

As before, the word "state" is abusive, since the element $\mathcal{E}\star\rho$ need not be a state. One can extend the properties from Definition 2 to extended state-over-time functions as follows: property (1) allows $\mathcal{E}\in\mathrm{HPTP}(\mathcal{A},\mathcal{B})$, properties (4) and (7) allow $\mathcal{E}\in\mathrm{TP}(\mathcal{A},\mathcal{B})$, property 5 allows $\lambda\in\mathbb{C}$ and $\mathcal{E},\mathcal{F}\in\mathrm{TP}(\mathcal{A},\mathcal{B})$, and properties (2), (3), and (6) are the same as originally stated. For brevity, we also henceforth drop the word "extended" from our phrasing unless emphasis is needed.

Definition 5.—Let $\star$ be a state-over-time function. Given a density matrix $\rho\in\mathcal{S}(\mathcal{A})$ and a CPTP map $\mathcal{E}:\mathcal{A}\to\mathcal{B}$, a *Bayes map* associated with $(\mathcal{E},\rho)$ is a trace-preserving linear map $\mathcal{E}^{\star}_{\rho}:\mathcal{B}\to\mathcal{A}$ such that

$$\mathcal{E}\star\rho = \tau\left(\widetilde{\mathcal{E}}^{\star}_{\rho}\star\mathcal{E}(\rho)\right), \tag{59}$$

where $\widetilde{\mathcal{E}}^{\star}_{\rho} := \dagger\circ\mathcal{E}^{\star}_{\rho}\circ\dagger$ is defined by [116]

$$\widetilde{\mathcal{E}}^{\star}_{\rho}(B) := \left(\mathcal{E}^{\star}_{\rho}(B^{\dagger})\right)^{\dagger}. \tag{60}$$

Note that when $\mathcal{E}^{\star}_{\rho}$ is $\dagger$-preserving, then $\widetilde{\mathcal{E}}^{\star}_{\rho} = \mathcal{E}^{\star}_{\rho}$. We have already seen a special case of this when discussing the symmetric bloom.

Remark 1.—Note that by taking the partial trace $\mathrm{tr}_{\mathcal{B}}$ of both sides of our generalized Bayes' rule, we find that

$$\begin{aligned}\rho &= \mathrm{tr}_{\mathcal{B}}\big(\tau(\widetilde{\mathcal{E}}^{\star}_{\rho}\star\mathcal{E}(\rho)\big) = \Big(\widetilde{\mathcal{E}}^{\star}_{\rho}\big(\mathcal{E}(\rho)\big)\Big)^{\dagger}\\ &= \mathcal{E}^{\star}_{\rho}\big(\mathcal{E}(\rho)^{\dagger}\big) = \mathcal{E}^{\star}_{\rho}\big(\mathcal{E}(\rho)\big),\end{aligned} \tag{61}$$

so that $\mathcal{E}^{\star}_{\rho}$ takes the prediction $\mathcal{E}(\rho)$ back to the prior ρ. This would not necessarily be the case had we used $\mathcal{E}^{\star}_{\rho}$ instead of $\widetilde{\mathcal{E}}^{\star}_{\rho}$ in our definition of Bayes' rule. More reasons for the importance of using $\widetilde{\mathcal{E}}^{\star}_{\rho}$ as opposed to just $\mathcal{E}^{\star}_{\rho}$ in Bayes' rule are provided in the upcoming examples.

C. The (r,s)-parametrized family

We now simultaneously generalize both the Leifer-Spekkens and Jordan state-over-time functions by viewing them as special cases inside a family. For each $r,s\in[0,1]$, the assignment sending $(\mathcal{E},\rho)\in\mathrm{TP}(\mathcal{A},\mathcal{B})\times\mathcal{S}(\mathcal{A})$ to

$$\begin{aligned}&s(\rho^{r}\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{1-r}\otimes 1_{\mathcal{B}})\\ &\quad+(1-s)(\rho^{1-r}\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{r}\otimes 1_{\mathcal{B}})\end{aligned} \tag{62}$$

defines a state-over-time function that is process linear and that satisfies the classical-limit axiom for all $r,s\in[0,1]$. Process linearity follows from the linearity of the channel state in its argument. As for the classical-limit axiom, note that by the functional calculus for matrices, if $[\rho\otimes 1_{\mathcal{B}},\mathscr{D}[\mathcal{E}]]=0$, then $[\rho^{r}\otimes 1_{\mathcal{B}},\mathscr{D}[\mathcal{E}]]=0$. Temporarily introducing the notation $s^{\perp}:=1-s$ and $r^{\perp}:=1-r$, this implies that

$$\begin{aligned}&s(\rho^{r}\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{r^{\perp}}\otimes 1_{\mathcal{B}})\\ &\quad+s^{\perp}(\rho^{r^{\perp}}\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{r}\otimes 1_{\mathcal{B}})\\ &\quad= s(\rho\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}]+(1-s)(\rho\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}]\\ &\quad=(\rho\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}],\end{aligned} \tag{63}$$

so that the classical-limit axiom holds.

This parametrized family specializes to the Leifer-Spekkens state over time for $r=1/2$ and s arbitrary, the symmetric bloom for $r\in\{0,1\}$ and $s=1/2$, and many other cases that have appeared in the literature. These include the *left* and *right blooms*, which specialize to the two-state formalism and time-dependent correlators, as is described next.

D. The right bloom

The cases $(r,s)=(1,1)$ or $(r,s)=(0,0)$ yield $\mathcal{E}\star\rho=(\rho\otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}]$, the *bloom* from Ref. [37], which is referred to here as the *right bloom* [117]. The right bloom is also state linear, so that it is in fact bilinear, and associative.

A Bayes map $\mathcal{E}^{\star}_{\rho}$ for the right bloom must satisfy the equation

$$\mathcal{E}\star\rho = \gamma\Big(\big((\dagger\circ\mathcal{E}^{\star}_{\rho}\circ\dagger)\star\mathcal{E}(\rho)\big)^{\dagger}\Big), \tag{64}$$

which, by applying $\dagger$ to both sides and using the fact that γ is $\dagger$ preserving, gives

$$\begin{aligned}\mathscr{D}[\mathcal{E}](\rho\otimes 1_{\mathcal{B}}) &= \gamma\Big(\big(\mathcal{E}(\rho)\otimes 1_{\mathcal{A}}\big)\mathscr{D}[\dagger\circ\mathcal{E}^{\star}_{\rho}\circ\dagger]\Big)\\ &= \big(1_{\mathcal{A}}\otimes\mathcal{E}(\rho)\big)\mathscr{D}\big[(\mathcal{E}^{\star}_{\rho})^{*}\big]\end{aligned} \tag{65}$$

by Lemma 3 in Appendix A. If we assume that $\mathcal{E}(\rho)$ is strictly positive, this gives the unique solution

$$\begin{aligned}\mathscr{D}\big[\mathcal{E}^{\star*}_{\rho}\big] &= \big(1_{\mathcal{A}}\otimes\mathcal{E}(\rho)^{-1}\big)\mathscr{D}[\mathcal{E}](\rho\otimes 1_{\mathcal{B}})\\ &= \mathscr{D}\big[\mathscr{L}_{\mathcal{E}(\rho)^{-1}}\circ\mathcal{E}\circ\mathscr{L}_{\rho}\big]\end{aligned} \tag{66}$$

by Lemma 3 in Appendix A (as in Lemma 2, $\mathscr{L}$ denotes left multiplication).

Since the channel state assignment $\mathscr{D}$ is a linear isomorphism, the inputs are equal, i.e., $\mathcal{E}_\rho^{\star *} = \mathscr{L}_{\mathcal{E}(\rho)^{-1}} \circ \mathcal{E} \circ \mathscr{L}_\rho$. Taking the Hilbert-Schmidt adjoint therefore gives the solution

$$\mathcal{E}_\rho^\star = \mathscr{L}_\rho^* \circ \mathcal{E}^* \circ \mathscr{L}_{\mathcal{E}(\rho)^{-1}}^* = \mathscr{L}_\rho \circ \mathcal{E}^* \circ \mathscr{L}_{\mathcal{E}(\rho)^{-1}}, \tag{67}$$

where we use the general fact that $\mathscr{L}_A^* = \mathscr{L}_{A^\dagger}$. Explicitly, this means that $\mathcal{E}_\rho^\star$ is given by the formula

$$\mathcal{B} \ni B \overset{\mathcal{E}_\rho^\star}{\mapsto} \rho\mathcal{E}^*\big(\mathcal{E}(\rho)^{-1}B\big), \tag{68}$$

which agrees with the Bayes map of Refs. [28–30,118].

We make two important remarks regarding the right-bloom state-over-time function and the associated Bayes map.

Remark 2.—If $\mathcal{E}$ is a $*$-isomorphism (equivalently, $\mathcal{E}$ is an invertible quantum channel, in the sense that $\mathcal{E}^{-1}$ exists and is also a quantum channel), then the Bayes map satisfies $\mathcal{E}_\rho^\star = \mathcal{E}^{-1}$. This property would generally *fail* if we had instead defined Bayes' rule naively as $\mathcal{E} \star \rho = \gamma\left(\mathcal{E}_\rho^\star \star \mathcal{E}(\rho)\right)$ without using τ on the right-hand side.

Indeed, if $\mathcal{E} = \mathrm{Ad}_U$ is invertible and represented by a unitary U, then the Bayes map given by Eq. (68) yields

$$\begin{aligned}\mathcal{E}_\rho^\star(B) &= \rho\mathrm{Ad}_{U^\dagger}\Big(\big(\mathrm{Ad}_U(\rho)\big)^{-1}B\Big) = \rho U^\dagger\big(U\rho U^\dagger\big)^{-1}BU \\ &= \rho U^\dagger(U^\dagger)^{-1}\rho^{-1}U^{-1}BU = \rho\rho^{-1}U^\dagger BU \\ &= U^\dagger BU,\end{aligned} \tag{69}$$

since $U^{-1} = U^\dagger$. Thus, $\mathcal{E}_\rho^\star = \mathcal{E}^{-1}$.

However, if the alternative proposal $\mathcal{E} \star \rho = \gamma\left(\mathcal{E}_\rho^\star \star \mathcal{E}(\rho)\right)$ for Bayes' rule had been used, then the unique linear solution would be given by $\mathcal{E}_\rho^\star(B) = \rho\mathcal{E}^*\big(B\mathcal{E}(\rho)^{-1}\big)$. Plugging in $\mathcal{E} = \mathrm{Ad}_U$ into *this* expression would instead yield

$$\mathcal{E}_\rho^\star(B) = \rho\mathrm{Ad}_{U^\dagger}\Big(B\big(\mathrm{Ad}_U(\rho)\big)^{-1}\Big) = \rho U^\dagger BU\rho^{-1}, \tag{70}$$

which is not the inverse of $\mathcal{E}$ in general. We find many other reasons supporting the usage of τ (and $\widetilde{\mathcal{E}_\rho^\star}$), as opposed to just γ, in our definition of Bayes' rule in later examples.

Remark 3.—Although $\mathcal{E}_\rho^\star$ is not in general CP, Ref. [30, Proposition 3.2] (see also Ref. [29]) shows that the following are equivalent:

(1) $\mathcal{E}_\rho^\star$ is $\dagger$-preserving
(2) $\rho\mathcal{E}^*\big(\mathcal{E}(\rho)^{-1}B\big) = \mathcal{E}^*\big(B\mathcal{E}(\rho)^{-1}\big)\rho$ for all $B \in \mathcal{B}$
(3) $\mathcal{E}_\rho^\star$ is CP
(4) $\mathrm{Ad}_{\mathcal{E}(\rho)^{it}} \circ \mathcal{E} = \mathcal{E} \circ \mathrm{Ad}_{\rho^{it}}$ for all $t \in \mathbb{R}$

In this case, one can rewrite the formula for $\mathcal{E}_\rho^\star$ as

$$\mathcal{E}_\rho^\star = \mathrm{Ad}_{\rho^{1/2}} \circ \mathcal{E}^* \circ \mathrm{Ad}_{\mathcal{E}(\rho)^{-1/2}}, \tag{71}$$

which agrees with the Petz recovery map.

Note that the last equivalent condition originally appeared in the work of Accardi and Cecchini [119], specifically in the context of GNS-symmetric dynamics and Tomita-Takesaki theory [30,120,121]. However, it also appears in the physics literature as time-symmetric covariant quantum channels [122,123] (see also Ref. [44]), where the time evolution symmetry is generated by the modular Hamiltonians associated with the initial and final states.

Interestingly, if we had ignored the dagger in our formulation of Bayes' rule, the covariance condition would have instead been $\mathrm{Ad}_{\mathcal{E}(\rho)^{it}} \circ \mathcal{E} = \mathcal{E} \circ \mathrm{Ad}_{\rho^{-it}}$ for all $t \in \mathbb{R}$. This would suggest that the modular flow goes forward in time for one state but backward in time for the other state, which seems to be at odds with the natural directionality of time suggested by the modular flow [124].

1. Weak values and the two-state formalism

In the special case where $\mathcal{A} = \mathbb{M}_m$, $\mathcal{B} = \mathbb{C}^X$, and $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ describes a POVM with x component $\mathcal{E}_x = \mathrm{tr}(M_x \,\cdot\,)$ for some positive operator M_x, the state over time associated with the right bloom generalizes the *two-state* from Refs. [41,42,48], which has also appeared recently in the context of holography [125]. In this case, $\mathcal{A} \otimes \mathcal{B} \cong \bigoplus_{x\in X} \mathbb{M}_m$ so that

$$\begin{aligned}\mathcal{E} \star \rho &= (\rho \otimes 1_{\mathbb{C}^X})\left(\sum_{i,j} E_{ij}^{(m)} \otimes \left(\sum_{x\in X}\mathrm{tr}\big(M_x E_{ji}^{(m)}\big)\delta_x\right)\right) \\ &= \sum_{x\in X}(\rho \otimes 1_{\mathbb{C}^X})\left(\sum_{i,j}\big((M_x)_{ij}E_{ij}^{(m)}\big) \otimes \delta_x\right) \\ &= \sum_{x\in X}(\rho M_x) \otimes \delta_x \cong \bigoplus_{x\in X}\rho M_x.\end{aligned} \tag{72}$$

In the special case where $\rho = |\psi\rangle\langle\psi|$ is the one-dimensional projection corresponding to a pure state $|\psi\rangle$ and similarly for $M_x = |\phi_x\rangle\langle\phi_x|$, this state over time becomes

$$\mathcal{E} \star \rho \cong \bigoplus_{x\in X}\langle\psi|\phi_x\rangle|\psi\rangle\langle\phi_x|, \tag{73}$$

which is the two-state appearing in Ref. [42] (combining all possibilities indexed by x together with the overlap coefficient $\langle\psi|\phi_x\rangle$) when the Hamiltonian evolution is trivial. This setup is similar to a PEM scenario in that a single state is prepared and the evolution is trivial. The main difference, however, is the choice of state-over-time function, which in this case yields the two-state and describes pre- and postselection [48].

In the case where the Hamiltonian is not trivial and there are three times $t_0 < t_1 < t_2$ with $U_{t_1 \leftarrow t_0}$, $U_{t_2 \leftarrow t_1}$, and $U_{t_2 \leftarrow t_0}$ describing the unitary evolution from t_0 to t_1, t_1 to t_2, and t_0 to t_2, respectively, (cf. Fig. 3) then, upon setting $\mathcal{E}' := \mathcal{E} \circ \mathrm{Ad}_{U_{t_2 \leftarrow t_0}}$, $|\psi'\rangle := U_{t_1 \leftarrow t_0}|\psi\rangle$, and $|\phi'_x\rangle := U^{\dagger}_{t_2 \leftarrow t_1}|\phi_x\rangle$, we obtain

$$
\begin{aligned}
\mathcal{E}' \star \rho &= (\rho \otimes 1_{\mathbb{C}^X})\mathscr{D}[\mathcal{E} \circ \mathrm{Ad}_{U_{t_2 \leftarrow t_0}}] \\
&= (\rho U^{\dagger}_{t_2 \leftarrow t_0} \otimes 1_{\mathbb{C}^X})\mathscr{D}[\mathcal{E}](U_{t_2 \leftarrow t_0} \otimes 1_{\mathbb{C}^X}) \\
&\cong \bigoplus_{x \in X} \rho U^{\dagger}_{t_2 \leftarrow t_0} M_x U_{t_2 \leftarrow t_0} \\
&= \bigoplus_{x \in X} |\psi\rangle\langle\psi| U^{\dagger}_{t_1 \leftarrow t_0} U^{\dagger}_{t_2 \leftarrow t_1} |\phi_x\rangle\langle\phi_x| U_{t_2 \leftarrow t_1} U_{t_1 \leftarrow t_0} \\
&= \bigoplus_{x \in X} |\psi\rangle\langle\psi'|\phi'_x\rangle\langle\phi'_x| U_{t_1 \leftarrow t_0},
\end{aligned} \tag{74}
$$

where we use Lemma 3 from Appendix A in the second equality. Therefore,

$$
\left(\mathrm{Ad}_{U_{t_1 \leftarrow t_0}} \otimes \mathrm{id}_{\mathbb{C}^X}\right)(\mathcal{E}' \star \rho) \cong \bigoplus_{x \in X} \langle\psi'|\phi'_x\rangle|\psi'\rangle\langle\phi'_x|. \tag{75}
$$

The intuitive reason for the $\mathrm{Ad}_{U_{t_1 \leftarrow t_0}}$ in front of $\mathcal{E}' \star \rho$ is because the two-state of Ref. [42] is viewed at time t_1 rather than the initial time t_0 or the final time t_2 (cf. Fig. 3). As such, it is necessary to propagate our state over time from time t_0 to t_1 in order to obtain the two-state of Refs. [42,126]. The $\mathrm{id}_{\mathbb{C}^X}$ is needed to incorporate all the possible outcomes due to the measurement.

We now examine what Bayes maps look like for the previous setup when the unitary evolution is trivial. Since a Bayes map $\mathcal{E}^{\star}_{\rho} : \mathcal{B} \to \mathcal{A}$ must be trace preserving, it must define an $m \times m$ matrix $\rho_x := \mathcal{E}^{\star}_{\rho}(\delta_x)$ such that $\mathrm{tr}(\rho_x) = 1$ for every $x \in X$. Write $p := \mathcal{E}(\rho) \equiv \bigoplus_x p_x \equiv \sum_x p_x \delta_x$ as the associated probability distribution on X, which is given by $p_x = \mathrm{tr}(M_x \rho)$. Then,

$$
\begin{aligned}
\rho_x &= \rho\mathcal{E}^*\big(\mathcal{E}(\rho)^{-1}\delta_x\big) = \rho\mathcal{E}^*\big(\mathcal{E}_x(\rho)^{-1}\delta_x\big) \\
&= \rho\left(\frac{1}{\mathrm{tr}(M_x\rho)}\right)\mathcal{E}^*(\delta_x) = \frac{\rho M_x}{p_x}
\end{aligned} \tag{76}
$$

whenever $p_x \neq 0$. Note that $\mathcal{E}^{\star}_{\rho}$ is not †-preserving and that it is necessary to use the version of Bayes' rule from Definition 5 to derive this result. Also, since $\mathcal{E}^{\star}_{\rho} : \mathbb{C}^X \to \mathbb{M}_m$ is a linear trace-preserving map from a classical algebra to a matrix algebra, it can be viewed as a sort of ensemble (though not technically, since each ρ_x need not be a density matrix).

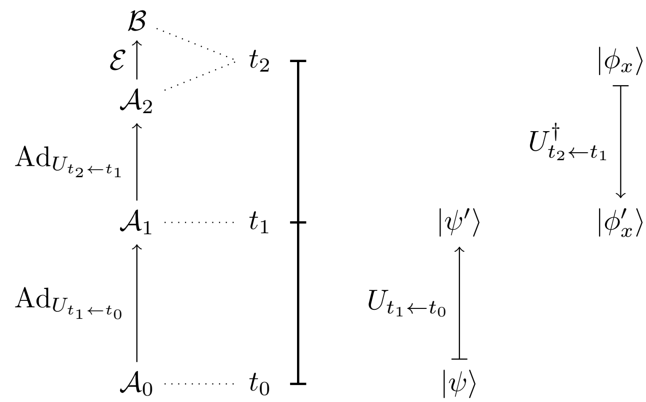

FIG. 3. In this figure, the algebras $\mathcal{A}_0$, $\mathcal{A}_1$, and $\mathcal{A}_2$ are all equal to some fiducial $\mathcal{A}$ and the subscript is meant to label the time. A state $|\psi\rangle$ is initially prepared at time t_0. Then, it evolves from t_0 to t_2 via $U_{t_2 \leftarrow t_0}$. Finally, a state $|\phi_x\rangle$ is measured at time t_2 via the POVM $\mathcal{E}$. The two-state of Refs. [42,126] at some intermediate time t_1 is obtained by forward-propagating $|\psi\rangle$ to t_1 via $U_{t_1 \leftarrow t_0}$ and back-propagating $|\phi_x\rangle$ to t_1 via $U^{\dagger}_{t_2 \leftarrow t_1}$. Taking the outer product of these two defines the two-state $|\psi'\rangle\langle\phi'_x|$. This two-state is precisely the state over time associated with our right bloom $(\mathcal{E} \circ \mathrm{Ad}_{U_{t_2 \leftarrow t_0}}) \star \rho \in \mathcal{A}_0 \otimes \mathcal{B}$ after forward-propagating the latter via $U_{t_1 \leftarrow t_0}$ to get an element of $\mathcal{A}_1 \otimes \mathcal{B}$.

In the special case where $\rho = |\psi\rangle\langle\psi|$ corresponds to a pure state $|\psi\rangle$ and, similarly, $M_x = |\phi_x\rangle\langle\phi_x|$ is a one-dimensional projection, this becomes

$$
\rho_x = \frac{|\psi\rangle\langle\psi|\phi_x\rangle\langle\phi_x|}{|\langle\psi|\phi_x\rangle|^2} = \frac{|\psi\rangle\langle\phi_x|}{\langle\phi_x|\psi\rangle}, \tag{77}
$$

which agrees with the *normalized two-state* from Ref. [42, Eq. (5)] and the *transition matrix* from Ref. [125, Eq. (1.3)]. The *expectation value*

$$
\mathrm{tr}(\rho_x^{\dagger} A) = \frac{\langle\psi|A|\phi_x\rangle}{\langle\psi|\phi_x\rangle} \tag{78}
$$

of this normalized two-state on an observable $A \in \mathbb{M}_m$ then agrees with the *weak value* of Aharonov, Albert, and Vaidman [Ref. [127, Eq. (6)]; see also Ref. [128]]. To contrast this with the earlier PEM scenario where the Bayesian inverse (using the Leifer-Spekkens state over time) of a measurement was a preparation, the current choice of state-over-time function provides a Bayes map that in general lacks positivity and leads to a different interpretation in terms of weak values.

E. The left bloom

The cases $(r,s) = (1,0)$ and $(r,s) = (0,1)$ yield the bloom $\mathcal{E} \star \rho = \mathscr{D}[\mathcal{E}](\rho \otimes 1_{\mathcal{B}})$ from Refs. [37,129], which is referred to here as the *left bloom*. Besides satisfying the classical-limit axiom, the left bloom is also bilinear and associative.

The Bayes maps associated with the left bloom are similar to those for the right bloom, so only the solution for the Bayes maps are provided. The unique linear map solving Bayes's rule is given by

$$\mathcal{B} \ni B \overset{\mathcal{E}^{\star}_{\rho}}{\mapsto} \mathcal{E}^{*}\big(B\mathcal{E}(\rho)^{-1}\big)\rho. \tag{79}$$

This map is CPTP under exactly the same conditions as in the right-bloom case.

1. The two-time correlator

In the special case where $\mathcal{E}$ is the identity channel, the left bloom appears in the context of two-point quantum correlation functions [49], where it is equal to the ideal two-point quantum correlator applied to ρ [130]. Although not a positive (or †-preserving) map, the two-point correlator has been given an operational interpretation in Refs. [49,115] by decomposing it into the sum of two †-preserving operations, one of which is the symmetric bloom state over time and the other of which is proportional to a commutator (for more details, see Sec. V A 2).

In the slightly more general case where $\mathcal{E}$ describes unitary evolution after some time t, i.e., $\mathcal{E} = \mathrm{Ad}_{e^{-iHt}}$ for some Hamiltonian $H \in \mathcal{A}^{\mathrm{sa}}$, then the left bloom defines two-time correlators [50]. Indeed, the *time-dependent correlation* between two observables $A, B \in \mathcal{A}^{\mathrm{sa}}$, is

$$\big\langle B(t)A(0)\big\rangle_{\rho} := \mathrm{tr}\big(e^{iHt}Be^{-iHt}A\rho\big), \tag{80}$$

where $A(s) := \mathrm{Ad}_{e^{isH}}(A)$ for $s \in \mathbb{R}$ and similarly for B. If $\mathcal{E} \star \rho$ denotes the left-bloom state over time, then

$$\big\langle B(t)A(0)\big\rangle_{\rho} = \mathrm{tr}\big((\mathcal{E} \star \rho)^{\dagger}(A \otimes B)\big) \tag{81}$$

for all $A, B \in \mathcal{A}^{\mathrm{sa}}$. Indeed,

$$\begin{aligned}\mathrm{tr}\big((\mathcal{E} \star \rho)^{\dagger}(A\otimes B)\big) &= \mathrm{tr}\big((\rho \otimes 1_{\mathcal{A}})\mathscr{D}[\mathrm{Ad}_{U^{\dagger}}](A \otimes B)\big)\\ &= \mathrm{tr}\Big(\big((\mathrm{id} \otimes \mathrm{Ad}_{U^{\dagger}})\mu^{*}_{\mathcal{A}}(1_{\mathcal{A}})\big)(A\rho \otimes B)\Big)\\ &= \mathrm{tr}\big(\mu^{*}_{\mathcal{A}}(1_{\mathcal{A}})(A\rho \otimes UBU^{\dagger})\big)\\ &= \mathrm{tr}(A\rho UBU^{\dagger})\\ &= \mathrm{tr}(UBU^{\dagger}A\rho),\end{aligned} \tag{82}$$

where we temporarily set $U = e^{itH}$. The third equality follows from self-adjointness of $\mathscr{D}[\mathrm{Ad}_{U^{\dagger}}]$ and $\mu^{*}_{\mathcal{A}}(1_{\mathcal{A}})$, while the fourth equality follows from self-adjointness of $\mu^{*}_{\mathcal{A}}(1_{\mathcal{A}})$ and $1_{\mathcal{A}}$.

2. Restoring the symmetry between left and right blooms

Note that if we denote the left- and right-bloom state-over-time functions by $\star_{\mathrm{L}}$ and $\star_{\mathrm{R}}$, respectively, then

$$\mathcal{E} \star_{\mathrm{L}} \rho = \tau\big(\widetilde{\mathcal{E}}^{\star}_{\rho} \star_{\mathrm{L}} \mathcal{E}(\rho)\big) = \gamma\big(\mathcal{E}^{\star}_{\rho} \star_{\mathrm{R}} \mathcal{E}(\rho)\big), \tag{83}$$

where the first equality is Bayes' rule and the second equality is the relationship between the left- and right-bloom states over time [131]. This relationship between left and right bloom and the connection to Bayes' rule resolves the open question of Leifer and Spekkens with regard to the apparent asymmetry between these two states over time (cf. the last paragraph in Ref. [25, Section VII.B.1]). In fact, it *is* a symmetry.

F. Bayes maps for the (r, s) family

After going through several examples of the (r, s) family, here we derive the general formula for the Bayes map. Since a Bayes map $\mathcal{E}^{\star}_{\rho}$ must satisfy $\gamma(\mathcal{E} \star \rho) = \big((\dagger \circ \mathcal{E}^{\star}_{\rho} \circ \dagger) \star \mathcal{E}(\rho)\big)^{\dagger}$, one can show (by similar calculations to the above) that Bayes' rule is equivalent to

$$\begin{aligned}&s(1_{\mathcal{B}} \otimes \rho^{r})\mathscr{D}[\mathcal{E}^{*}](1_{\mathcal{B}} \otimes \rho^{r^{\perp}})\\ &\quad + s^{\perp}(1_{\mathcal{B}} \otimes \rho^{r^{\perp}})\mathscr{D}[\mathcal{E}^{*}](1_{\mathcal{B}} \otimes \rho^{r})\\ &\quad = s\big(\mathcal{E}(\rho)^{r^{\perp}} \otimes 1_{\mathcal{A}}\big)\mathscr{D}[\mathcal{E}^{\star}_{\rho}]\big(\mathcal{E}(\rho)^{r} \otimes 1_{\mathcal{A}}\big)\\ &\qquad + s^{\perp}\big(\mathcal{E}(\rho)^{r} \otimes 1_{\mathcal{A}}\big)\mathscr{D}[\mathcal{E}^{\star}_{\rho}]\big(\mathcal{E}(\rho)^{r^{\perp}} \otimes 1_{\mathcal{A}}\big)\end{aligned} \tag{84}$$

by Eq. (A10) of Lemma 3. Introducing the same notation q_k and $|w_k\rangle$ as in the case of the symmetric bloom, Bayes' rule is equivalent to

$$\begin{aligned}&\sum_{k,l} e_{kl} \otimes \Big(s\rho^{r}\mathcal{E}^{*}(e_{lk})\rho^{r^{\perp}} + s^{\perp}\rho^{r^{\perp}}\mathcal{E}^{*}(e_{lk})\rho^{r}\Big)\\ &\quad = \sum_{k,l}\Big(\big(sq_k^{r^{\perp}}q_l^{r} + s^{\perp}q_k^{r}q_l^{r^{\perp}}\big)e_{kl}\Big) \otimes \mathcal{E}^{\star}_{\rho}(e_{lk}),\end{aligned} \tag{85}$$

where $e_{kl} := |w_k\rangle\langle w_l|$. Since the e_{kl} are linearly independent, this gives the solution

$$\mathcal{E}^{\star}_{\rho}(e_{kl}) = \frac{s\rho^{r}\mathcal{E}^{*}(e_{kl})\rho^{1-r} + (1-s)\rho^{1-r}\mathcal{E}^{*}(e_{kl})\rho^{r}}{sq_k^{r}q_l^{1-r} + (1-s)q_k^{1-r}q_l^{r}}, \tag{86}$$

which can be linearly extended to define a Bayes map for the (r, s) family. Note that $\mathcal{E}^{\star}_{\rho}$ is †-preserving when $s = 1/2$. It is unclear to us if there is a manifestly basis-independent expression for this Bayes map.

G. Generalized conditional expectations

In Ref. [32], Tsang has argued for the interpretation of generalized conditional expectations as maps defining retrodiction and hence as quantum analogs of Bayes' theorem. Generalized conditional expectations include the Petz recovery maps, the Bayes maps for the entire (r,s) family, and many others (see also Ref. [132, Section 6.1] and the references therein). In the present section, we show that *all* of the generalized conditional expectations described by Tsang are indeed included in our framework of states over time and Bayes maps. Namely, from the data used in Refs. [32,132] to define generalized conditional expectations, we construct state-over-time functions and then we show that (the Hilbert-Schmidt adjoint of) these generalized conditional expectation are Bayes maps for those states over time.

Beyond what we have already mentioned above, our work also achieves several new insights:

(1) To derive our Bayes map, we do not need to minimize or extremize any distance measures as is done in Ref. [32, Section III.B]. Namely, we only require a few consistency conditions and our notion of time-reversal symmetry.
(2) We complete Tsang's open remark or question (posed at the end of Ref. [32, Section III.B]) of relating his formalism to that of Horsman *et al.* [38] and Leifer-Spekkens [25]. We are able to do this precisely because of our proposed relationship between states over time and Bayes' rules (the usages of both τ and $\widetilde{\mathcal{E}}^{\star}_{\rho}$ from Definition 5 are crucial here).
(3) We provide an explicit formula for the Bayes map associated with the (r,s) family.

In what follows, our notation differs from Refs. [32,132] to avoid any conflicting notation. For every multimatrix algebra $\mathcal{A}$, let $\Theta:\mathcal{S}(\mathcal{A})\to\mathrm{Map}(\mathcal{A},\mathcal{A})$ send a state ρ to a linear map $\Theta_\rho:\mathcal{A}\to\mathcal{A}$ that satisfies the following axioms (see Refs. [32, Section III.A] and [132, Section 6.1]) [133]:

(T1) $\Theta_\rho(A)=\rho A$ whenever $A\in\mathcal{A}$ satisfies $[\rho,A]=0$.
(T2) $\Theta_\rho\circ\mathcal{E}^{-1}=\mathcal{E}^{-1}\circ\Theta_{\mathcal{E}(\rho)}$ whenever $\mathcal{E}$ is a $*$-isomorphism (e.g., $\Theta_\rho\circ\mathrm{Ad}_{U^\dagger}=\mathrm{Ad}_{U^\dagger}\circ\Theta_{U\rho U^\dagger}$ for all unitaries U).
(T3) If $\mathcal{A}'$ is another multimatrix algebra, then $\Theta_{\rho\otimes\rho'}=\Theta_\rho\otimes\Theta_{\rho'}$ for all states $\rho\in\mathcal{S}(\mathcal{A})$ and $\rho'\in\mathcal{S}(\mathcal{A}')$, provided that at least one of ρ or ρ' are in the center of $\mathcal{A}$ or $\mathcal{A}'$, respectively.
(T4) Θ_ρ is self-adjoint and positive semidefinite with respect to the Hilbert-Schmidt inner product.

We call such a map Θ a *state-rendering map*. Note that in our notation, we suppress the dependence of Θ on the algebra $\mathcal{A}$, as it should be evident from the context.

Remark 4.—The fact that a map $\Phi:\mathcal{A}\to\mathcal{A}$ is self-adjoint with respect to the Hilbert-Schmidt inner product *does not* imply that Φ is $\dagger$-preserving. An example is the map $\Phi(A)=\rho A$ for a state $\rho\in\mathcal{S}(\mathcal{A})$ that is not in the center of $\mathcal{A}$ (when $\mathcal{A}=\mathbb{M}_m$, this means that ρ is not the maximally mixed state). Conversely, if Φ is $\dagger$-preserving, this does not imply that Φ is self-adjoint. An example is the map $\Phi=\mathrm{Ad}_U$ for a unitary $U\in\mathcal{A}$ not proportional to the identity. Similarly, the fact that Φ is positive semidefinite with respect to the Hilbert-Schmidt inner product *does not* mean that Φ is a positive map in the sense that it takes positive elements to positive elements. An example is $\Phi(A)=\rho A$ for a state ρ that is not maximally mixed. Conversely, if Φ is positive, this does not imply Φ is positive semidefinite. An example is $\Phi(A)=A^{\mathrm{T}}$, the transpose map on (complex) matrix algebras.

Theorem 2.—Given a state-rendering map Θ, the assignment $\star$ sending a trace-preserving linear map $\mathcal{E}:\mathcal{A}\to\mathcal{B}$ and $\rho\in\mathcal{S}(\mathcal{A})$ to

$$\mathcal{E}\star\rho:=(\Theta_\rho\otimes\mathrm{id}_{\mathcal{B}})\big(\mathscr{D}[\mathcal{E}]\big) \tag{87}$$

defines a state-over-time function that is process linear and satisfies the classical-limit axiom.

We include the proof of this theorem here to illustrate how the axioms of state-rendering maps are related to those of state-over-time functions.

Proof.—We first check that $\mathcal{E}\star\rho$ has the correct marginals and then we prove the latter two claims.

The marginal on $\mathcal{A}$ is

$$\begin{aligned}\mathrm{tr}_{\mathcal{B}}(\mathcal{E}\star\rho)&=\big(\Theta_\rho\otimes(\mathrm{tr}\circ\mathcal{E})\big)\big(\mu^*_{\mathcal{A}}(1_{\mathcal{A}})\big)\\&=\big(\Theta_\rho\otimes\mathrm{tr}\big)\big(\mu^*_{\mathcal{A}}(1_{\mathcal{A}})\big)\\&=\Theta_\rho(1_{\mathcal{A}})\\&=\rho.\end{aligned} \tag{88}$$

The first equality follows from the definition of the channel state and the definition of the partial trace $\mathrm{tr}_{\mathcal{B}}$. The second equality follows from the fact that $\mathcal{E}$ is trace preserving. The third equality is an identity due to the fact that $\mu^*_{\mathcal{A}}$ is a (nonpositive) broadcasting map (see the axioms of a quantum Markov category in Refs. [28,118]). The last equality follows from axiom (T1).

In the case where $\mathcal{A}=\mathbb{M}_m$, the marginal on $\mathcal{B}$ is

$$\begin{aligned}\mathrm{tr}_{\mathcal{A}}(\mathcal{E}\star\rho)&=\sum_{i,j}\mathrm{tr}\big(\Theta_\rho(E^{(m)}_{ij})\big)\mathcal{E}(E^{(m)}_{ji})\\&=\sum_{i,j}\mathrm{tr}\big(\Theta^*_\rho(1_{\mathcal{A}})^\dagger E^{(m)}_{ij}\big)\mathcal{E}(E^{(m)}_{ji})\\&=\sum_{i,j}\mathrm{tr}\big(\Theta_\rho(1_{\mathcal{A}})^\dagger E^{(m)}_{ij}\big)\mathcal{E}(E^{(m)}_{ji})\\&=\sum_{i,j}\mathrm{tr}\big(\rho E^{(m)}_{ij}\big)\mathcal{E}(E^{(m)}_{ji})\\&=\mathcal{E}(\rho).\end{aligned} \tag{89}$$

The second equality follows from the definition of the Hilbert-Schmidt inner product. The third equality follows from the self-adjointness of Θ_ρ, which is part of axiom (T4). The fourth equality follows from axiom (T1) and the fact that ρ is Hermitian. The last equality follows from the fact that $\mathrm{tr}\big(\rho E_{ij}^{(m)}\big)$ is precisely the ji-matrix entry of ρ. The proof when $\mathcal{A}$ is an arbitrary multimatrix algebra is left as an exercise.

Process linearity of $\star$ follows from linearity of $\mathscr{D}$ and Θ_ρ.

Finally, to see that the classical-limit axiom for $\star$ holds, suppose that $[\rho \otimes 1_{\mathcal{B}}, \mathscr{D}[\mathcal{E}]] = 0$. Then,

$$\left[\rho \otimes \frac{1_{\mathcal{B}}}{\mathrm{tr}(1_{\mathcal{B}})}, \mathscr{D}[\mathcal{E}]\right] = 0 \tag{90}$$

and

$$\begin{aligned}\mathcal{E} \star \rho &= \left(\Theta_\rho \otimes \left(\mathrm{tr}(1_{\mathcal{B}})\Theta_{\frac{1_{\mathcal{B}}}{\mathrm{tr}(1_{\mathcal{B}})}}\right)\right)(\mathscr{D}[\mathcal{E}]) \\ &= \mathrm{tr}(1_{\mathcal{B}})\Theta_{\rho\otimes\frac{1_{\mathcal{B}}}{\mathrm{tr}(1_{\mathcal{B}})}}(\mathscr{D}[\mathcal{E}]) \\ &= (\rho \otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}].\end{aligned} \tag{91}$$

The first equality follows from axiom (T1). The second equality follows from linearity of the tensor product and axiom (T3). The third equality follows from Eq. (90) and axiom (T1). ■

Many of the examples of state-over-time functions that we have given earlier are special cases of this construction [134]:

(1) The right and left blooms are obtained from $\Theta_\rho^{\mathrm{R}}(A) := \rho A$ and $\Theta_\rho^{\mathrm{L}}(A) := A\rho$, respectively.
(2) The Leifer-Spekkens state over time is obtained from $\Theta_\rho^{\mathrm{LS}}(A) := \sqrt{\rho}A\sqrt{\rho}$.
(3) The symmetric bloom state over time is obtained from $\Theta_\rho^{\mathrm{J}}(A) := \frac{1}{2}\{\rho, A\} \equiv \frac{1}{2}(\rho A + A\rho)$.
(4) More generally, the (r,s) family is obtained from $\Theta_\rho^{(r,s)}(A) := s\rho^r A\rho^{1-r} + (1-s)\rho^{1-r}A\rho^r$.

We now introduce the generalized conditional expectation from Ref. [32, Eq. (3.22)] (see also Refs. [132, Eq. (6.21)] and [135]), which has been derived in Ref. [32] by minimizing a certain inner product associated with a state-rendering map Θ.

Given a channel $\mathcal{E} : \mathcal{A} \to \mathcal{B}$, a state $\rho \in \mathcal{S}(\mathcal{A})$, and a state-rendering map Θ, a *generalized conditional expectation* is a linear map $\mathcal{E}_{\Theta,\rho} : \mathcal{A} \to \mathcal{B}$ such that

$$\mathcal{E} \circ \Theta_\rho = \Theta_{\mathcal{E}(\rho)} \circ \mathcal{E}_{\Theta,\rho}. \tag{92}$$

Theorem 3.—Let $\mathcal{E}_{\Theta,\rho}$ be a generalized conditional expectation as in Eq. (92). Then, $\mathcal{E}_{\Theta,\rho}^*$ is a Bayes map for $(\mathcal{E}, \rho)$ associated with the state over time $\star$ defined in Theorem 2.

The proof of Theorem 3 is provided in Appendix B. Interestingly, we have made no use of axiom (T2) in any of our results thus far. The importance of this axiom is due to the next fact, the proof of which is deferred to Appendix B. It shows that if a quantum channel is a $*$-isomorphism, then the inverse is a Bayesian inverse.

Proposition 2.—Let Θ be a state-rendering map and $\star$ its associated state-over-time function. If $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ is a $*$-isomorphism, then $\mathcal{E}^{-1}$ is a Bayesian inverse of $(\mathcal{E}, \rho)$ for every $\rho \in \mathcal{S}(\mathcal{A})$.

Remark 5.—Proposition 2 and Theorem 3 provide additional illustrations of the importance of using the dagger and $\widetilde{\mathcal{E}}_\rho^\star$ in the definition of Bayes' rule. This can be seen more clearly in their proofs.

VI. DISCUSSION AND CONCLUSIONS

In this work, we have provided a rigorous definition for a Bayesian inverse using the notion of a state over time [37,38] in terms of a universal time-reversal-symmetry map that is independent of the input channel and state. In particular, this answers an open question of Leifer and Spekkens [25] on the question about the relationship between Bayes' rules and time-reversal symmetry. It also answers a question posed by Tsang regarding the connection between states over time and generalized conditional expectations [32]. Tsang has also attempted to unify many notions of generalized conditional expectations and our work includes all of the special cases considered in Ref. [32] along with several new ones. We have shown how our definition of a state-over-time function reproduces those of Leifer and Spekkens [25], the two-state vector formalism [41,42,48], the symmetric bloom of the present authors [37], and others, many of which are summarized in Table II.

In addition, using our proposed definitions of Bayesian inverses and, more generally, Bayes maps, associated with a state-over-time function, we have obtained the normalized two-states of Reznik and Aharonov [42], all the rotated Petz recovery maps, two-point quantum correlators, the quantum Bayes' rule of Fuchs [17], and many more concepts arising in various scenarios. Furthermore, we have explained how one would unambiguously arrive at the quantum state-update rule associated with instruments [107] for *any* state over time satisfying the classical-limit axiom. This shows how the state-update rule can be viewed as a quantum Bayes' rule, which has been advocated by Bub [11], Ozawa [13], Tegmark [23], and many others. However, we remark that this version of the quantum Bayes' rule is not as general as our Bayes' rule, the latter of which applies to *arbitrary* maps, not necessarily special kinds of instruments.

Several open questions should be addressed. Although some of these are mentioned earlier, we summarize these questions for convenience and add new ones for further thought:

TABLE II. The many state-over-time functions appearing in this work, along with their formulas, properties satisfied, and associated Bayes maps. The axioms are Hermiticity (P1), block-positivity (P2), positivity (P3), state linearity (P4), process linearity (P5), the classical limit (P7), and associativity A (bilinearity has been removed from the table to avoid redundancy). The $*$ for Ohya's compound state over time is because the classical limit is satisfied for density matrices with no repeating eigenvalues. Note that we do not fully define Ohya's compound state over time for arbitrary CPTP maps between multimatrix algebras (this will be addressed in future work, along with additional examples of state-over-time functions). The question mark represents the fact that we have not yet determined whether the given axiom is satisfied.

Name (page ref.)	State over time $\mathcal{E} \star \rho$	P1	P2	P3	P4	P5	P7	A	Bayes map $\mathcal{E}^{\star}_{\rho}$
Uncorrelated (7)	$\rho \otimes \mathcal{E}(\rho)$	✓	✓	✓	✗	✓	✗	✗	Any CPTP such that $\mathcal{E}^{\star}_{\rho}(\mathcal{E}(\rho)) = \rho$
Ohya compound (7)	$\sum_{\alpha} \lambda_{\alpha} P_{\alpha} \otimes \mathcal{E}\left(\frac{P_{\alpha}}{\mathrm{tr}(P_{\alpha})}\right)$	✓	✓	✓	✗	✓	$*$	?	Not computed here
Leifer-Spekkens (7)	$(\sqrt{\rho} \otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\sqrt{\rho} \otimes 1_{\mathcal{B}})$	✓	✓	✗	✗	✓	✓	✗	Petz map $\mathscr{R}_{\rho,\mathcal{E}} := \mathrm{Ad}_{\rho^{1/2}} \circ \mathcal{E}^{*} \circ \mathrm{Ad}_{\mathcal{E}(\rho)^{-1/2}}$
t-rotated (8)	$(\rho^{1/2-it} \otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{1/2+it} \otimes 1_{\mathcal{B}})$	✓	✓	✗	✗	✓	✓	✗	Rotated Petz map $\mathrm{Ad}_{\rho^{-it}} \circ \mathscr{R}_{\rho,\mathcal{E}} \circ \mathrm{Ad}_{\mathcal{E}(\rho)^{it}}$
STH (8)	$(U_{\rho}^{\dagger}\rho^{1/2} \otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}](\rho^{1/2}U_{\rho} \otimes 1_{\mathcal{B}})$	✓	✓	✗	✗	✓	✓	✗	$\mathrm{Ad}_{U_{\rho}^{\dagger}} \circ \mathscr{R}_{\rho,\mathcal{E}} \circ \mathrm{Ad}_{U_{\mathcal{E}(\rho)}}$
Symmetric bloom (11)	$\frac{1}{2}\{\rho \otimes 1_{\mathcal{B}}, \mathscr{D}[\mathcal{E}]\}$	✓	✗	✗	✓	✓	✓	✓	$\lvert w_k\rangle\langle w_l\rvert \mapsto (q_k + q_l)^{-1}\{\rho, \mathcal{E}^{*}(\lvert w_k\rangle\langle w_l\rvert)\}$
Right bloom (13)	$(\rho \otimes 1_{\mathcal{B}})\mathscr{D}[\mathcal{E}]$ (e.g., two-state)	✗	✗	✗	✓	✓	✓	✓	$B \mapsto \rho\mathcal{E}^{*}(\mathcal{E}(\rho)^{-1}B)$ (e.g., weak values)
Left bloom (15)	$\mathscr{D}[\mathcal{E}](\rho \otimes 1_{\mathcal{B}})$ (e.g., correlator)	✗	✗	✗	✓	✓	✓	✓	$B \mapsto \mathcal{E}^{*}(B\mathcal{E}(\rho)^{-1})\rho$

(1) Are there any physically relevant examples of state-over-time functions that are not process linear?

(2) In regard to Ohya's compound state over time [52], is there a canonical way of isolating an extremal decomposition of every state ρ given a pair $(\mathcal{E}, \rho)$, perhaps by extremizing some information measure? If so, can the compound state over time be improved to define a state-over-time function with more desirable properties?

(3) Is there a state-over-time function the Bayesian inverses of which are the universal recovery maps of Junge *et al.* [69,95]? What about the more general averaged rotated Petz recovery maps [44] or the optimal state retrieval maps of Surace and Scandi [71]?

(4) Which state-over-time functions give Bayesian inverses that define retrodiction functors in the sense of Ref. [44]? Namely, which properties of state-over-time functions correspond to which axioms for retrodiction families? For example, what properties of a state-over-time function imply that the Bayesian inverses are compositional? This is an important question to address so that one can choose a state-over-time function appropriately to achieve the desired properties of retrodiction.

(5) Just as the two-point correlator can be decomposed into Hermitian and anti-Hermitian parts, with a minimal Kraus decompositions the coefficients of which can be given an operational meaning [49], can arbitrary nonpositive Bayes maps also be decomposed in a similar way to provide them with operational meanings? Some progress has recently been made in this direction [136].

(6) Another approach toward generalizing quantum states associated with acausally related regions to causally related ones is the superdensity formalism [137,138]. One major difference between this line of development and ours is that we demand our objects to be defined on the tensor product of the algebras of the associated regions, whereas in Refs. [137,138], the algebra is "doubled." For example, for the case of two causally related regions with algebras $\mathcal{A}$ and $\mathcal{B}$, our states over time are elements of $\mathcal{A} \otimes \mathcal{B}$, whereas superdensity operators are elements of $\mathcal{A}^{*} \otimes \mathcal{B}^{*} \otimes \mathcal{A} \otimes \mathcal{B}$, where the duals are with respect to the Hilbert-Schmidt inner products. As such, we have not yet been able to meaningfully compare our constructions.

(7) Our present work focuses primarily on the setting of two times and hence two algebras mediated by some channel. The associativity axiom, which we alluded to briefly, is closely related to consistently defining *multitime* states over time where one might have a sequence of composable evolutions, $\mathcal{A} \xrightarrow{\mathcal{E}} \mathcal{B} \xrightarrow{\mathcal{F}}$

$\mathcal{C} \to \cdots$, together with an initial state $\rho \in \mathcal{S}(\mathcal{A})$. It is expected that these ideas are closely related to multitime correlators and multitime measurements [139,140], though the details have not yet been worked out.

(8) Just as genuine states in the tensor product $\mathcal{A} \otimes \mathcal{B}$ can be given important information measures, such as entanglement entropy (for a review with applications to quantum field theory, see Ref. [141]), what information measures can be assigned to states over time for causally related regions? Some attempts to extend entanglement entropy to such scenarios have been made in Euclidean field theories for the special case of the right bloom under the name "pseudoentropy" [125,142–144] or in the context of superdensity operators [137,138]. Much work needs to be done to better understand such dynamical information measures. What can these teach us about temporal correlations [31,36,140,145,146] or information extraction from black-hole evaporation [147–149]?

We hope to address these questions, particularly some from the last item, in subsequent work. We hope that by answering such questions, we may eventually provide new methods for detecting and computing temporal correlations that distinguish between classical and quantum systems. More provocatively, we suspect that these differences may provide some insight toward our understanding of quantum information theory and the structure of space and time, and hence quantum gravity.

ACKNOWLEDGMENTS

We thank Clive Cenxin Aw, Francesco Buscemi, John DeBrota, Markus Frembs, Marie Fries, Chris Fuchs, Xiao-Kan Guo, Sachin Gupta, Shintaro Minagawa, Jacques Pienaar, David Schmid, Blake Stacey, Jacopo Surace, Tadashi Takayanagi, Mankei Tsang, Mark Wilde, Edward Witten, Yilè Yīng, and Beata Zjawin for discussions. We thank Carlo Rovelli for bringing Ref. [50] to our attention. We thank two anonymous reviewers for their helpful feedback. A.J.P. thanks the Institut des Hautes Études Scientifiques and the Yukawa Institute for Theoretical Physics at Kyoto University for their hospitality. This work is supported by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT)–Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Transformative Research Areas (A) "Extreme Universe," Grant No. 21H05183. The authors also acknowledge support from the Blaumann Foundation.

APPENDIX A: USEFUL FACTS ABOUT THE CHANNEL STATE

In this appendix, we state some important lemmas that are used in several calculations throughout this work.

Lemma 1.—Let B be an $m \times n$ matrix and let C be an $n \times m$ matrix. Then,

$$\sum_{i,j}^{m} E_{ij}^{(m)} \otimes \left(CE_{ji}^{(m)}B\right) = \sum_{k,l}^{n} \left(BE_{kl}^{(n)}C\right) \otimes E_{lk}^{(n)}. \tag{A1}$$

Proof.—Writing $E_{ij}^{(m)} = |i\rangle\langle j|$ and using the completeness relations $\mathbb{1}_n = \sum_{k=1}^{n} |k\rangle\langle k|$ and $\mathbb{1}_m = \sum_{i=1}^{m} |i\rangle\langle i|$, we obtain

$$\begin{aligned}
&\sum_{i,j}^{m} E_{ij}^{(m)} \otimes \left(CE_{ji}^{(m)}B\right) \\
&= \sum_{i,j}^{m}\sum_{k,l}^{n} |i\rangle\langle j| \otimes \Big(|l\rangle\langle l|C|j\rangle\langle i|B|k\rangle\langle k|\Big) \\
&= \sum_{i,j}^{m}\sum_{k,l}^{n} \Big(|i\rangle\langle i|B|k\rangle\langle l|C|j\rangle\langle j|\Big) \otimes |l\rangle\langle k|,
\end{aligned} \tag{A2}$$

which equals the required expression. ■

Lemma 2.—If $\mathcal{E} : \mathcal{A} \to \mathcal{B}$ is a CP map, then

$$(\mathrm{id}_{\mathcal{A}} \otimes \mathcal{E})\big(\mu_{\mathcal{A}}^{*}(1_{\mathcal{A}})\big) = (\mathcal{E}^{*} \otimes \mathrm{id}_{\mathcal{B}})\big(\mu_{\mathcal{B}}^{*}(1_{\mathcal{B}})\big), \tag{A3}$$

$$\gamma\big(\mathscr{D}[\mathcal{E}]\big) = \mathscr{D}[\mathcal{E}^{*}], \tag{A4}$$

and

$$(A \otimes B)\mathscr{D}[\mathcal{E}](A' \otimes B') = \mathscr{D}[\mathscr{L}_{B} \circ \mathscr{R}_{B'} \circ \mathcal{E} \circ \mathscr{L}_{A'} \circ \mathscr{R}_{A}] \tag{A5}$$

for all $A, A' \in \mathcal{A}$ and $B, B' \in \mathcal{B}$. Here, $\mathscr{L}_{A'}$ and $\mathscr{R}_{A}$ are left and right multiplication by A, namely, $\mathscr{L}_{A'}(X) = A'X$ and $\mathscr{R}_{A}(X) = XA$ for all $X \in \mathcal{A}$ (and similarly for $\mathcal{B}$).

Proof.—In this proof, we assume $\mathcal{A} = \mathbb{M}_m$ and $\mathcal{B} = \mathbb{M}_n$ for simplicity. Then, Eq. (A3) reads

$$\mathscr{D}[\mathcal{E}] = \sum_{k,l} \mathcal{E}^{*}\big(E_{kl}^{(n)}\big) \otimes E_{lk}^{(n)}. \tag{A6}$$

To see that this holds, note that since $\mathcal{E}$ is CP, there exist Kraus operators $\{V_\alpha\}$ such that $\mathcal{E} = \sum_\alpha \mathrm{Ad}_{V_\alpha}$ [79]. Hence,

$$\begin{aligned}
\mathscr{D}[\mathcal{E}] &= \sum_{\alpha}\sum_{i,j} E_{ij}^{(m)} \otimes \left(V_\alpha E_{ji}^{(m)} V_\alpha^{\dagger}\right) \\
&= \sum_{\alpha}\sum_{k,l} \left(V_\alpha^{\dagger} E_{kl}^{(n)} V_\alpha\right) \otimes E_{lk}^{(n)} \\
&= \gamma\big(\mathscr{D}[\mathcal{E}^{*}]\big)
\end{aligned} \tag{A7}$$

by Lemma 1. Equation (A4) follows directly from this. For the last identity, and using Lemma 1 as well, we obtain

$$\begin{aligned}
&(A\otimes B)\mathscr{D}[\mathcal{E}](A'\otimes B')\\
&=\sum_{\alpha}(1_{\mathcal{A}}\otimes BV_{\alpha})\left(\sum_{i,j}AE_{ij}^{(m)}A'\otimes E_{ji}^{(m)}\right)(1_{\mathcal{A}}\otimes V_{\alpha}^{\dagger}B')\\
&=\sum_{\alpha}(1_{\mathcal{A}}\otimes BV_{\alpha})\left(\sum_{i,j}E_{ij}^{(m)}\otimes A'E_{ji}^{(m)}A\right)(1_{\mathcal{A}}\otimes V_{\alpha}^{\dagger}B')\\
&=\sum_{i,j}E_{ij}^{(m)}\otimes\big(B\mathcal{E}(A'E_{ji}^{(m)}A)B'\big),
\end{aligned}\tag{A8}$$

which equals (A5). ■

We now generalize Lemma 2 to allow for arbitrary linear maps.

Lemma 3.—If $\mathcal{E}:\mathcal{A}\to\mathcal{B}$ is a linear map, then

$$(\mathrm{id}_{\mathcal{A}}\otimes\mathcal{E})\big(\mu_{\mathcal{A}}^{*}(1_{\mathcal{A}})\big)=\big((\dagger\circ\mathcal{E}^{*}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big)\big(\mu_{\mathcal{B}}^{*}(1_{\mathcal{B}})\big),\tag{A9}$$

$$\gamma\big(\mathscr{D}[\mathcal{E}]\big)=\mathscr{D}[\dagger\circ\mathcal{E}^{*}\circ\dagger]=\mathscr{D}[\mathcal{E}^{*}]^{\dagger},\tag{A10}$$

$$(A\otimes B)\mathscr{D}[\mathcal{E}](A'\otimes B')=\mathscr{D}[\mathscr{L}_{B}\circ\mathscr{R}_{B'}\circ\mathcal{E}\circ\mathscr{L}_{A'}\circ\mathscr{R}_{A}],\tag{A11}$$

and

$$(A^{\dagger}\otimes B)\mathscr{D}[\mathcal{E}](A\otimes B^{\dagger})=\mathscr{D}[\mathrm{Ad}_{B}\circ\mathcal{E}\circ\mathrm{Ad}_{A}]\tag{A12}$$

for all $A,A'\in\mathcal{A}$ and $B,B'\in\mathcal{B}$.

Proof.—The first equality can be seen by first taking the Hilbert-Schmidt adjoints of both sides ($1_{\mathcal{A}}$ is viewed as the unique linear map $\mathbb{C}\to\mathcal{A}$ sending $\lambda\in\mathbb{C}$ to $\lambda 1_{\mathcal{A}}$, and similarly for $\mathcal{B}$). The claim is then equivalent to [150]

$$\mathrm{tr}\circ\mu_{\mathcal{A}}\circ(\mathrm{id}_{\mathcal{A}}\otimes\mathcal{E}^{*})=\mathrm{tr}\circ\mu_{\mathcal{B}}\circ\big((\dagger\circ\mathcal{E}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big).\tag{A13}$$

But this identity follows immediately from applying the left map to $A\otimes B\in\mathcal{A}\otimes\mathcal{B}$, which results in

$$\mathrm{tr}\big(A\mathcal{E}^{*}(B)\big)=\mathrm{tr}\big((A^{\dagger})^{\dagger}\mathcal{E}^{*}(B)\big)=\mathrm{tr}\big(\mathcal{E}(A^{\dagger})^{\dagger}B\big).\tag{A14}$$

The first equality in Eq. (A10) follows from Eq. (A9). The second equality in Eq. (A10) follows from

$$\begin{aligned}
\mathscr{D}[\dagger\circ\mathcal{E}\circ\dagger]^{\dagger}&=\Big(\big(\mathrm{id}_{\mathcal{A}}\otimes(\dagger\circ\mathcal{E}\circ\dagger)\big)\big(\mu_{\mathcal{A}}^{*}(1_{\mathcal{A}})\big)\Big)^{\dagger}\\
&=\big((\mathrm{id}_{\mathcal{A}}\otimes\mathcal{E})\circ(\dagger\otimes\dagger)\big)\big(\mu_{\mathcal{A}}^{*}(1_{\mathcal{A}})\big)\\
&=(\mathrm{id}_{\mathcal{A}}\otimes\mathcal{E})\big(\mu_{\mathcal{A}}^{*}(1_{\mathcal{A}})\big)\\
&=\mathscr{D}[\mathcal{E}]
\end{aligned}\tag{A15}$$

(the string-diagram language of quantum Markov categories from Ref. [28] makes these identities immediately apparent). Finally, Eq. (A11) follows from Eq. (A5) in Lemma 2 together with linearity of the channel state with respect to its input. This is because every linear map $\mathcal{E}$ is a complex combination of at most four CP maps [79,151,152]. Indeed, one can first decompose the Choi matrix $\mathscr{C}[\mathcal{E}]$ of $\mathcal{E}$ into a complex combination of self-adjoint elements

$$\mathscr{C}[\mathcal{E}]=\frac{1}{2}\big(\mathscr{C}[\mathcal{E}]+\mathscr{C}[\mathcal{E}]^{\dagger}\big)+\frac{1}{2i}\big(i\mathscr{C}[\mathcal{E}]-i\mathscr{C}[\mathcal{E}]^{\dagger}\big)\tag{A16}$$

and then further decompose each of these self-adjoint elements into linear combinations of positive elements. The end result is a complex combination of the form

$$\mathscr{C}[\mathcal{E}]=\mathscr{C}[\mathcal{E}]_{\mathfrak{R}}^{+}-\mathscr{C}[\mathcal{E}]_{\mathfrak{R}}^{-}+i\mathscr{C}[\mathcal{E}]_{\mathfrak{I}}^{+}-i\mathscr{C}[\mathcal{E}]_{\mathfrak{I}}^{-},\tag{A17}$$

where $\mathscr{C}[\mathcal{E}]_{\mathfrak{R}}^{+},\mathscr{C}[\mathcal{E}]_{\mathfrak{R}}^{-},\mathscr{C}[\mathcal{E}]_{\mathfrak{I}}^{+},\mathscr{C}[\mathcal{E}]_{\mathfrak{I}}^{-}$ are all positive. Then, by using Choi's theorem, each of these corresponds to a CP map, with an overall Kraus decomposition given by

$$\mathcal{E}=\sum_{\alpha}\mathrm{Ad}_{V_{\alpha}}-\sum_{\beta}\mathrm{Ad}_{W_{\beta}}+i\sum_{\gamma}\mathrm{Ad}_{X_{\gamma}}-i\sum_{\delta}\mathrm{Ad}_{Y_{\delta}}.\tag{A18}$$

Since the channel state is linear in its argument, this decomposition and Eq. (A5) from Lemma 2 imply Eq. (A11). Equation (A12) is a special case of Eq. (A11). ■

APPENDIX B: PROOFS OF SOME MAIN RESULTS

This appendix contains proofs of our main theorems that are not included in the main text. Before proving Theorem 3 about the relationship between Bayes maps and generalized conditional expectations, we compute the reverse orientation state over time [cf. Ref. [116]] in the following lemma.

Lemma 4.—Let Θ be a state-rendering map and let $\star$ denote the associated state-over-time function from Theorem 2. Then,

$$\big(\widetilde{\mathcal{E}}\star\rho\big)^{\dagger}=\big((\dagger\circ\Theta_{\rho}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big)\mathscr{D}[\mathcal{E}]\tag{B1}$$

for all linear maps $\mathcal{E}:\mathcal{A}\to\mathcal{B}$ and states $\rho\in\mathcal{S}(\mathcal{A})$.

Proof.—By definition of the left-hand side, we have

$$\begin{aligned}(\widetilde{\mathcal{E}}\star\rho)^{\dagger}&=\big((\dagger\circ\mathcal{E}\circ\dagger)\star\rho\big)^{\dagger}\\&=\big((\Theta_{\rho}\otimes\mathrm{id}_{\mathcal{B}})\mathscr{D}[\dagger\circ\mathcal{E}\circ\dagger]\big)^{\dagger}\\&=\big((\dagger\circ\Theta_{\rho}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big)\big(\mathscr{D}[\dagger\circ\mathcal{E}\circ\dagger]^{\dagger}\big)\\&=\big((\dagger\circ\Theta_{\rho}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big)\mathscr{D}[\mathcal{E}],\end{aligned}\tag{B2}$$

where the third equality follows from the properties of the involution † [153] and the fourth equality follows from Eq. (A10) of Lemma 3. ■

Proof of Theorem 3.—The goal is to prove $\gamma(\mathcal{E}\star\rho)=\big(\widetilde{\mathcal{E}^{*}_{\Theta,\rho}}\star\mathcal{E}(\rho)\big)^{\dagger}$. Taking the Hilbert-Schmidt adjoint of both sides of Eq. (92) and using axiom (T4) gives the equivalent condition:

$$\Theta_{\rho}\circ\mathcal{E}^{*}=\mathcal{E}^{*}_{\Theta,\rho}\circ\Theta_{\mathcal{E}(\rho)}.\tag{B3}$$

Using this, we obtain

$$\begin{aligned}\gamma(\mathcal{E}\star\rho)&=\gamma\Big(\big(\Theta_{\rho}\otimes\mathrm{id}_{\mathcal{B}}\big)\big(\mathscr{D}[\mathcal{E}]\big)\Big)\\&=\big(\mathrm{id}_{\mathcal{B}}\otimes\Theta_{\rho}\big)\gamma\big(\mathscr{D}[\mathcal{E}]\big)\\&=\big(\mathrm{id}_{\mathcal{B}}\otimes\Theta_{\rho}\big)\big(\mathscr{D}[\mathcal{E}^{*}]\big)\\&=\big(\mathrm{id}_{\mathcal{B}}\otimes(\Theta_{\rho}\circ\mathcal{E}^{*})\big)\big(\mu^{*}_{\mathcal{B}}(1_{\mathcal{B}})\big)\\&=\big(\mathrm{id}_{\mathcal{B}}\otimes(\mathcal{E}^{*}_{\Theta,\rho}\circ\Theta_{\mathcal{E}(\rho)})\big)\big(\mu^{*}_{\mathcal{B}}(1_{\mathcal{B}})\big)\\&=\Big(\big(\mathrm{id}_{\mathcal{B}}\otimes\mathcal{E}^{*}_{\Theta,\rho}\big)\circ\big(\mathrm{id}_{\mathcal{B}}\otimes\Theta_{\mathcal{E}(\rho)}\big)\Big)\big(\mu^{*}_{\mathcal{B}}(1_{\mathcal{B}})\big)\\&=\Big(\big(\mathrm{id}_{\mathcal{B}}\otimes\mathcal{E}^{*}_{\Theta,\rho}\big)\circ\big((\dagger\circ\Theta_{\mathcal{E}(\rho)}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big)\Big)\big(\mu^{*}_{\mathcal{B}}(1_{\mathcal{B}})\big)\\&=\big((\dagger\circ\Theta_{\mathcal{E}(\rho)}\circ\dagger)\otimes\mathrm{id}_{\mathcal{B}}\big)\mathscr{D}[\mathcal{E}^{*}_{\Theta,\rho}]\\&=\big(\widetilde{\mathcal{E}^{*}_{\Theta,\rho}}\star\mathcal{E}(\rho)\big)^{\dagger}.\end{aligned}\tag{B4}$$

The second equality follows from the property of the swap map γ. The third equality follows from Lemma 2. The fourth equality is the definition of the channel state. The fifth equality follows from Eq. (B3). The sixth equality follows from the interchange law relating $\circ$ and $\otimes$. The seventh equality follows from Eq. (A9) of Lemma 3 and axiom (T4). The eighth equality follows from the interchange law and the definition of the channel state. The ninth equality follows from Lemma 4. ■

Proof of Proposition 2.—By axiom (T2), $\mathcal{E}$ is a generalized conditional expectation for every $\rho\in\mathcal{S}(\mathcal{A})$. Hence, $\mathcal{E}^{*}=\mathcal{E}^{-1}$ is a Bayes map for $(\mathcal{E},\rho)$ by Theorem 3. Since $\mathcal{E}^{-1}$ is a channel, $\mathcal{E}^{-1}$ is a Bayesian inverse of $(\mathcal{E},\rho)$.

[1] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Elsevier, San Francisco, 1988).

[2] V. Balasubramanian, Statistical Inference, Occam's Razor, and Statistical Mechanics on the Space of Probability Distributions, Neural Comput. **9**, 349 (1997).

[3] D. A. Berry, Bayesian clinical trials, Nat. Rev. Drug Discov. **5**, 27 (2006).

[4] K. Friston, The free-energy principle: A unified brain theory?, Nat. Rev. Neurosci. **11**, 127 (2010).

[5] S. M. Barnett, in *Quantum Information and Coherence*, edited by E. Andersson and P. Öhberg (Springer International Publishing, Cham, 2014), p. 1.

[6] Z. Ghahramani, Probabilistic machine learning and artificial intelligence, Nature **521**, 452 (2015).

[7] B. Jacobs, The mathematics of changing one's mind, via Jeffrey's or via Pearl's update rule, J. Artificial Intelligence Res. **65**, 783 (2019).

[8] F. Buscemi and V. Scarani, Fluctuation theorems from Bayesian retrodiction, Phys. Rev. E **103**, 052111 (2021).

[9] C. C. Aw, F. Buscemi, and V. Scarani, Fluctuation theorems with retrodiction rather than reverse processes, AVS Quantum Science **3**, 045601 (2021).

[10] T. St. Clere Smithe, Compositional Active Inference I: Bayesian Lenses. Statistical Games, arXiv:2109.04461 [math.ST] (2021).

[11] J. Bub, Von Neumann's projection postulate as a probability conditionalization rule in quantum mechanics, J. Philos. Log. **6**, 381 (1977).

[12] G. Lüders, Concerning the state-change due to the measurement process, Ann. Phys. **15**, 663 (2006).

[13] M. Ozawa, in *Quantum Communication, Computing, and Measurement* (Springer, 1997), p. 233, arXiv:quant-ph/9705030 [quant-ph].

[14] A. N. Korotkov, Continuous quantum measurement of a double dot, Phys. Rev. B **60**, 5737 (1999).

[15] A. N. Korotkov, Selective quantum evolution of a qubit state due to continuous measurement, Phys. Rev. B **63**, 115403 (2001).

[16] R. Schack, T. A. Brun, and C. M. Caves, Quantum Bayes rule, Phys. Rev. A **64**, 014305 (2001).

[17] C. A. Fuchs, in *Decoherence and its implications in quantum computation and information transfer*, Vol. 182, edited by G. Tony and P. E. Turchi (IOS Press, 2001), p. 38, arXiv:quant-ph/0106166 [quant-ph].

[18] M. K. Warmuth, in *Advances in Neural Information Processing Systems*, Vol. 18, edited by Y. Weiss, B. Schölkopf, and J. Platt (MIT Press, Cambridge, 2005).

[19] M. K. Warmuth and D. Kuzmin, Bayesian generalized probability calculus for density matrices, Mach. Learn **78**, 63 (2010).

[20] M. S. Leifer, Quantum dynamics as an analog of conditional probability, Phys. Rev. A **74**, 042310 (2006).

[21] M. S. Leifer, in *Foundations of Probability and Physics - 4*, American Institute of Physics Conference Series, Vol. 889, edited by G. Adenier, C. Fuchs, and A. Y. Khrennikov (2007), p. 172.

[22] B. Coecke and R. W. Spekkens, Picturing classical and quantum Bayesian inference, Synthese **186**, 651 (2012).

[23] M. Tegmark, How unitary cosmology generalizes thermodynamics and solves the inflationary entropy problem, Phys. Rev. D **85**, 123517 (2012).

[24] D. Farenick and M. J. Kozdron, Conditional expectation and Bayes' rule for quantum random variables and positive operator valued measures, J. Math. Phys **53**, 042201 (2012).
[25] M. S. Leifer and R. W. Spekkens, Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference, Phys. Rev. A **88**, 052130 (2013).
[26] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle, and M. Walter, Entanglement Wedge Reconstruction via Universal Recovery Channels, Phys. Rev. X **9**, 031011 (2019).
[27] A. J. Parzygnat and B. P. Russo, Non-commutative disintegrations: existence and uniqueness in finite dimensions,' (2023), to appear in J. Noncommut. Geom., arXiv:1907.09689 [quant-ph].
[28] A. J. Parzygnat, Inverses, disintegrations, and Bayesian inversion in quantum Markov categories, arXiv:2001.08375 [quant-ph] (2020).
[29] A. J. Parzygnat and B. P. Russo, A non-commutative Bayes' theorem, Linear Algebra Its Appl. **644**, 28 (2022).
[30] L. Giorgetti, A. J. Parzygnat, A. Ranallo, and B. P. Russo, Bayesian inversion and the Tomita-Takesaki modular group, Quart. J. Math. **00**, 1 (2023).
[31] D. Chruściński and T. Matsuoka, Quantum Conditional Probability and Measurement Induced Disturbance of a Quantum Channel, Rep. Math. Phys. **86**, 115 (2020).
[32] M. Tsang, Generalized conditional expectations for quantum retrodiction and smoothing, Phys. Rev. A **105**, 042213 (2022).
[33] H. Liu, A Quantum Bayes' Rule and Related Inference, arXiv:2204.08845 [quant-ph] (2022).
[34] T. Matsuoka and D. Chruściński, in *International Conference on Quantum Probability & Related Topics* (Springer International Publishing, Cham, 2022), p. 135.
[35] M. L. Bera and M. N. Bera, Quantum Bayes' rule affirms consistency in measurement inferences in quantum mechanics, arXiv:2207.08623 [quant-ph] (2022).
[36] J. F. Fitzsimons, J. A. Jones, and V. Vedral, Quantum correlations which imply causation, Sci. Rep. **5**, 18281 (2015).
[37] J. Fullwood and A. J. Parzygnat, On quantum states over time, Proc. R. Soc. A **478** (2022).
[38] D. Horsman, C. Heunen, M. F. Pusey, J. Barrett, and R. W. Spekkens, Can a quantum state over time resemble a quantum state at a single time?, Proc. R. Soc. A **473**, 20170395 (2017).
[39] J. von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition* (Princeton University Press, Princeton, 2018).
[40] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Physics Vol. 190 (Springer-Verlag, Berlin, 1983).
[41] S. Watanabe, Symmetry of Physical Laws. Part III. Prediction and Retrodiction, Rev. Mod. Phys. **27**, 179 (1955).
[42] B. Reznik and Y. Aharonov, Time-symmetric formulation of quantum mechanics, Phys. Rev. A **52**, 2538 (1995).
[43] A. Di Biagio, P. Donà, and C. Rovelli, The arrow of time in operational formulations of quantum theory, Quantum **5** (2021).
[44] A. J. Parzygnat and F. Buscemi, Axioms for retrodiction: achieving time-reversal symmetry with a prior, Quantum **7**, 1013 (2023).
[45] B. Coecke, S. Gogioso, and J. H. Selby, The time-reverse of any causal theory is eternal noise, arXiv:1711.05511 [quant-ph] (2017).
[46] G. Chiribella, E. Aurell, and K. Życzkowski, Symmetries of quantum evolutions, Phys. Rev. Res. **3**, 033028 (2021).
[47] J. H. Selby, M. E. Stasinou, S. Gogioso, and B. Coecke, Time symmetry in quantum theories and beyond, arXiv:2209.07867 [quant-ph] (2022).
[48] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Time Symmetry in the Quantum Process of Measurement, Phys. Rev. **134**, B1410 (1964).
[49] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, Direct observation of any two-point quantum correlation function, arXiv:1312.4240 [quant-ph] (2013).
[50] E. Bianchi, H. M. Haggard, and C. Rovelli, The boundary is mixed, Gen. Relativ. Gravit. 49 (2017).
[51] D. Petz, A dual in von Neumann algebras with weights, Q. J. Math. **35**, 475 (1984).
[52] M. Ohya, Note on quantum probability, Lett. Nuovo Cimento **38**, 402 (1983).
[53] M. Ohya, On compound state and mutual information in quantum information theory, IEEE Trans. Inf. Theory **29**, 770 (1983).
[54] M. Ozawa, Quantum State Reduction: An Operational Approach, Fortschr. Phys. **46**, 615 (1998).
[55] D. R. Farenick, *Algebras of Linear Transformations*, Universitext (Springer, New York, 2001), p. xiv+238.
[56] F. M. Goodman, P. de la Harpe, and V. F. R. Jones, *Coxeter Graphs and Towers of Algebras*, Mathematical Sciences Research Institute Publications, Vol. 14 (Springer, New York, 1989), p. x+288.
[57] M. Ohya and D. Petz, *Quantum Entropy and its Use*, Texts and Monographs in Physics (Springer-Verlag, Berlin, 1993), p. viii+335.
[58] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Comm. Math. Phys. **17**, 239 (1970).
[59] G. C. Wick, A. Wightman, and E. Wigner, The Intrinsic Parity of Elementary Particles, Phys. Rev. **88**, 101 (1952).
[60] R. F. Streater, Outline of axiomatic relativistic quantum field theory, Rep. Prog. Phys. **38**, 771 (1975).
[61] The Jamiołkowski state is commonly defined as the partial transpose of the *Choi matrix* [79]. Equation (7) provides a manifestly basis-independent formula for the Jamiołkowski state that is also applicable to arbitrary multimatrix algebras.
[62] The notation $\mathcal{E} \star \rho$ is based on Refs. [25,38].
[63] The condition $[\mathscr{D}[\mathcal{E}], \rho \otimes 1_{\mathcal{B}}] = 0$ is equivalent to $[\rho, \mathcal{E}^*(B)] = 0$ for all $B \in \mathcal{B}$.
[64] For reference, and to provide a formula in the Schrödinger picture, a state-over-time function $\star$ is *associative* if and only if

$$\mathscr{D}^{-1}_{\mathcal{A},\mathcal{B}\otimes\mathcal{C}}\left[\operatorname{tr}(1_{\mathcal{A}})\left((\mathcal{F}\circ\operatorname{tr}_{\mathcal{A}})\star\left(\frac{\mathscr{D}_{\mathcal{A},\mathcal{B}}[\mathcal{E}]}{\operatorname{tr}(1_{\mathcal{A}})}\right)\right)\right]\star\rho$$
$$= (\mathcal{F}\circ\operatorname{tr}_{\mathcal{A}})\star(\mathcal{E}\star\rho)$$

for all $\rho \in \mathcal{S}(\mathcal{A})$ and composable pairs $\mathcal{A} \xrightarrow{\mathcal{E}} \mathcal{B} \xrightarrow{\mathcal{F}} \mathcal{C}$ of CPTP maps. Note that one must be careful about domains to even make sense of this axiom. In particular, we point out that the formula presented here is a more appropriate axiom for associativity than the one considered in Ref. [37] due to the fact that $\star$ need not be state linear (this is why we include the factors of $\mathrm{tr}(1_{\mathcal{A}})$ in the present formulation). Indeed, if $\star$ is state linear, then these factors cancel, which need not happen if $\star$ is not state linear. Note that this change does not alter the main theorem of Ref. [37], since the state-over-time function constructed there is state linear.

[65] M. Asorey, A. Kossakowski, G. Marmo, and E. C. G. Sudarshan, Relations Between Quantum Maps and Quantum States, Open Syst. Inf. Dyn. **12**, 319 (2005).

[66] M. Girard, M. Plávala, and J. Sikora, Jordan products of quantum channels and their compatibility, Nat. Commun. **12**, 2129 (2021).

[67] O. Fawzi and R. Renner, Quantum Conditional Mutual Information and Approximate Markov Chains, Commun. Math. Phys. **340**, 575 (2015).

[68] M. M. Wilde, Recoverability in quantum information theory, Proc. R. Soc. A **471**, 20150338 (2015).

[69] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, in *2016 IEEE International Symposium on Information Theory (ISIT)* (2016), p. 2494.

[70] E. A. Carlen and A. Vershynina, Recovery map stability for the data processing inequality, J. Phys. A **53**, 035204 (2020).

[71] J. Surace and M. Scandi, State retrieval beyond Bayes' retrodiction, Quantum **7**, 990 (2023).

[72] For general algebras, the swap map $\gamma : \mathcal{A} \otimes \mathcal{B} \to \mathcal{B} \otimes \mathcal{A}$ is not to be confused with the SWAP operator $\mathscr{D}[\mathrm{id}_{\mathcal{A}}]$, the latter of which is an *element* of $\mathcal{A} \otimes \mathcal{A}$.

[73] J. C. Baez, in *Structural Foundations of Quantum Gravity*, edited by S. French, D. Rickles, and J. Saatsi (Oxford University Press, Oxford, 2006), p. 240, arXiv:quant-ph/0404040 [quant-ph].

[74] B. Coecke and A. Kissinger, *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge University Press, Cambridge, 2017).

[75] G. Chiribella and Z. Liu, Quantum operations with indefinite time direction, Commun. Phys. **5**, 1 (2022).

[76] This is because

$$\mathscr{D}^{-1}_{\mathcal{A},\mathcal{B}\otimes\mathcal{C}}\left[\mathrm{tr}(1_{\mathcal{A}})\left((\mathcal{F}\circ\mathrm{tr}_{\mathcal{A}})\star\left(\frac{\mathscr{D}_{\mathcal{A},\mathcal{B}}[\mathcal{E}]}{\mathrm{tr}(1_{\mathcal{A}})}\right)\right)\right]\star\rho$$
$$=\rho\otimes\mathcal{E}(\rho)\otimes\mathcal{F}\left(\mathcal{E}\left(\frac{1_{\mathcal{A}}}{\mathrm{tr}(1_{\mathcal{A}})}\right)\right),$$

while

$$(\mathcal{F}\circ\mathrm{tr}_{\mathcal{A}})\star(\mathcal{E}\star\rho)=\rho\otimes\mathcal{E}(\rho)\otimes\mathcal{F}\big(\mathcal{E}(\rho)\big).$$

Since $\mathcal{F}\left(\mathcal{E}\left(\frac{1_{\mathcal{A}}}{\mathrm{tr}(1_{\mathcal{A}})}\right)\right)$ need not equal $\mathcal{F}\big(\mathcal{E}(\rho)\big)$ for arbitrary $\mathcal{F}, \mathcal{E}$, and ρ, this shows that associativity fails in general.

[77] Technically, Ohya's original definition of a compound state does not give a well-defined state-over-time function because it assumes *additional data* when the input state has repeating eigenvalues. Namely, it requires a *choice* of an eigenvector decomposition. Therefore, we slightly modify Ohya's construction so that no such additional data are needed and so that we obtain a well-defined state-over-time function.

[78] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[79] M. D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).

[80] D. Petz, Sufficiency of channels over von Neumann algebras, Q. J. Math. **39**, 97 (1988).

[81] D. Petz, Sufficient subalgebras and the relative entropy of states of a von Neumann algebra, Commun.Math. Phys. **105**, 123 (1986).

[82] D. Petz, Monotonicity of quantum relative entropy revisited, Rev. Math. Phys. **15**, 79 (2003).

[83] A. Jenčová, Preservation of a quantum Rényi relative entropy implies existence of a recovery map, J. Phys. A: Math. Theor. **50**, 085303 (2017).

[84] H. Barnum and E. Knill, Reversing quantum dynamics with near-optimal quantum and classical fidelity, J. Math. Phys. **43**, 2097 (2002).

[85] H. K. Ng and P. Mandayam, Simple approach to approximate quantum error correction based on the transpose channel, Phys. Rev. A **81**, 062342 (2010).

[86] G. E. Crooks, Quantum operation time reversal, Phys. Rev. A. **77**, 034101 (2008).

[87] M. S. Leifer and M. F. Pusey, Is a time symmetric interpretation of quantum theory possible without retrocausality?, Proc. R. Soc. A **473**, 20160607 (2017).

[88] F. Buscemi, J. Schindler, and D. Šafránek, Observational entropy, coarse-grained states, and the Petz recovery map: information-theoretic properties and bounds, New J. Phys. **25**, 053002 (2023).

[89] H. Kwon and M. S. Kim, Fluctuation Theorems for a Quantum Channel, Phys. Rev. X **9**, 031029 (2019).

[90] Z. Huang, Fluctuation theorems for multitime processes, Phys. Rev. A **105**, 062217 (2022).

[91] K. Furuya, N. Lashkari, and S. Ouseph, Real-space RG, error correction and Petz map, J. High Energ. Phys. **2022** (2022).

[92] C.-F. Chen, G. Penington, and G. Salton, Entanglement wedge reconstruction using the Petz map, J. High Energy Phys. **2020**, 168 (2020).

[93] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, Replica wormholes and the black hole interior, J. High Energ. Phys. **2022**, 1 (2022).

[94] T. Kibe, P. Mandayam, and A. Mukhopadhyay, Holographic spacetime, black holes and quantum error correcting codes: A review, Eur. Phys. J. C **82**, 463 (2022).

[95] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, Universal Recovery Maps and Approximate Sufficiency of Quantum Relative Entropy, Ann. Henri Poincaré **19**, 2955 (2018).

[96] D. Sutter, M. Tomamichel, and A. W. Harrow, Strengthened monotonicity of relative entropy via pinched Petz

recovery map, IEEE Transactions on Information Theory **62**, 2907 (2016).
[97] D. Sutter, M. Berta, and M. Tomamichel, Multivariate Trace Inequalities, Comm. Math. Phys. **352**, 37 (2017).
[98] This is not to be confused with the *operational* time-reverse in the sense defined by Leifer and Pusey [87], because our "control" and "uncontrolled" variables are switched.
[99] R. C. Jeffrey, *The Logic of Decision* (University of Chicago Press, Chicago, 1990), 2nd ed.
[100] J. A. Barandes and D. Kagan, Quantum conditional probabilities and new measures of quantum information, Ann. Phys. (NY), 169192 (2022).
[101] R. Furber and B. Jacobs, From Kleisli Categories to Commutative C*-algebras: Probabilistic Gelfand Duality, Log. Methods Comput. Sci. **11**, 1 (2015).
[102] A. J. Parzygnat, Discrete probabilistic and algebraic dynamics: A stochastic commutative Gelfand-Naimark theorem, arXiv:1708.00091 [math.FA] (2017).
[103] T. Fritz, T. Gonda, P. Perrone, and E. Fjeldgren Rischel, Representable Markov categories and comparison of statistical experiments in categorical probability, Theor. Comput. Sci. **961**, 113896 (2023).
[104] K. R. W. Jones, Principles of quantum inference, Ann. Phys. **207**, 140 (1991).
[105] K. R. W. Jones, *Quantum Inference and the Optimal Determination of Quantum States* (Createspace Independent Publishing Platform, Scotts Valley, 2012).
[106] K.-E. Hellwig and K. Kraus, Pure operations and measurements, Comm. Math. Phys. **11**, 214 (1969).
[107] K. Kraus, General state changes in quantum theory, Ann. Physics **64**, 311 (1971).
[108] M. Ozawa, Conditional probability and a posteriori states in quantum mechanics, Publ. Res. Inst. Math. Sci. **21**, 279 (1985).
[109] G. Lüders, Über die Zustandsänderung durch den Meßprozeß, Ann. Phys. **8**, 322 (1950).
[110] K. Vanslette, The quantum Bayes rule and generalizations from the quantum maximum entropy method, Commun. Phys. **2**, 025017 (2018).
[111] D. Freeman, D. Giannakis, and J. Slawinska, Quantum mechanics for closure of dynamical systems, arXiv:2208.03390 [math.DS] (2022).
[112] H. Umegaki, Conditional expectation in an operator algebra. IV. Entropy and information, Kodai Math. Sem. Rep. **14**, 59 (1962).
[113] M. Takesaki, Conditional expectations in von Neumann algebras, J. Funct. Anal. **9**, 306 (1972).
[114] R. Alicki, Comment on "Reduced Dynamics Need Not Be Completely Positive", Phys. Rev. Lett. **75**, 3020 (1995).
[115] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, Universal optimal quantum correlator, Int. J. Quantum Inf. **12**, 1560002 (2014).
[116] One could equivalently define a new state-over-time function $\star^{\dagger} : \mathrm{TP}(\mathcal{A},\mathcal{B}) \times \mathcal{S}(\mathcal{A}) \to \mathcal{A} \otimes \mathcal{B}$ the value of which on $(\mathcal{E},\rho)$ is given by $\mathcal{E} \star^{\dagger} \rho := \big((\dagger \circ \mathcal{E} \circ \dagger) \star \rho\big)^{\dagger}$. In this case, we could call $\star^{\dagger}$ the *reverse orientation* state-over-time function associated with $\star$. Then, an equivalent formulation of Bayes' rule reads $\mathcal{E} \star \rho = \gamma\big(\mathcal{E}^{\star}_{\rho} \star^{\dagger} \mathcal{E}(\rho)\big)$, where only the swap map γ is used, but with the additional modification of using the reverse orientation state over time on the right.
[117] Using notation from Ref. [37], this is because the functional associated with $\mathcal{E} \star \rho$ sends $A \otimes B \in \mathcal{A} \otimes \mathcal{B}$ to

$$\langle \mathcal{E} \star \rho, A \otimes B \rangle = \mathrm{tr}\big((\mathcal{E} \star \rho)^{\dagger}(A \otimes B)\big) = \mathrm{tr}\big(\rho A \mathcal{E}^{*}(B)\big),$$

which agrees with $\omega\big(AF(B)\big)$ upon setting $F = \mathcal{E}^{*}$ and $\rho = \mathscr{D}[\omega]$.
[118] A. J. Parzygnat, in *Proceedings 18th International Conference on Quantum Physics and Logic*, Electronic Proceedings in Theoretical Computer Science (EPTCS) (2021) arXiv:2102.01529 [quant-ph].
[119] L. Accardi and C. Cecchini, Conditional expectations in von Neumann algebras and a theorem of Takesaki, J. Funct. Anal. **45**, 245 (1982).
[120] J. P. Bannon, J. Cameron, and K. Mukherjee, The modular symmetry of Markov maps, J. Math. Anal. Appl. **439**, 701 (2016).
[121] C. Anantharaman-Delaroche, On ergodic theorems for free group actions on noncommutative spaces, Probab. Theory Relat. Fields **135**, 520 (2006).
[122] I. Marvian, Ph.D. thesis, Department of Physics and Astronomy, University of Waterloo, UWSpace, 2012.
[123] M. Lostaglio, K. Korzekwa, D. Jennings, and T. Rudolph, Quantum Coherence, Time-Translation Symmetry, and Thermodynamics, Phys. Rev. X **5**, 021001 (2015).
[124] A. Connes and C. Rovelli, Von Neumann algebra automorphisms and time-thermodynamics relation in generally covariant quantum theories, Class. Quant. Grav. **11**, 2899 (1994).
[125] Y. Nakata, T. Takayanagi, Y. Taki, K. Tamaoka, and Z. Wei, New holographic generalization of entanglement entropy, Phys. Rev. D **103**, 026005 (2021).
[126] Y. Aharonov and L. Vaidman, in *Time in Quantum Mechanics*, edited by J. Muga, R. S. Mayato, and Í. Egusquiza (Springer-Verlag, Berlin, 2008), p. 399.
[127] 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).
[128] When going from the Schrödinger picture to the Heisenberg picture, the two-state ρ_x gets sent to its weak-value-expectation functional, sending an observable A to $\mathrm{tr}(\rho_x^{\dagger} A)$. This duality, which is obtained by using the Hilbert-Schmidt adjoint, is a conjugate-linear isomorphism with our convention. This is why $\rho_x^{\dagger}$, as opposed to ρ_x, appears in the expression for weak values. Technically, $\mathrm{tr}(\rho_x A)$ reproduces the expression [127, Eq. (6)] exactly. This is due to a matter of convention, since, if we use the right bloom (to be defined in Sec. V E), the two-state would instead be given by $\rho_x = \frac{M\rho_x}{\mathrm{tr}(M\rho_x)} = \frac{|\phi_x\rangle\langle\psi|}{\langle\psi|\phi_x\rangle}$. The associated weak values for this two-state would then equal $\mathrm{tr}(\rho_x^{\dagger} A)$, in complete agreement with Ref. [127, Eq. (6)].
[129] J. Fullwood and A. J. Parzygnat, The Information Loss of a Stochastic Map, Entropy **23** (2021).
[130] Indeed, when $\mathcal{A} = \mathbb{M}_m$, the channel state $\mathscr{D}[\mathrm{id}_{\mathcal{A}}]$ equals the SWAP operator, sending $|i\rangle \otimes |j\rangle$ to $|j\rangle \otimes |i\rangle$ in Ref. [49].

[131] In terms of reverse orientation state-over-time functions, this reads $\star_L^\dagger = \star_R$, and a proof is given by

$$\begin{aligned}\mathcal{E} \star_{\mathrm{L}}^\dagger \rho &= \big((\dagger \circ \mathcal{E} \circ \dagger) \star_{\mathrm{L}} \rho\big)^\dagger = \big((\rho \otimes 1_{\mathcal{B}})\mathscr{D}[\dagger \circ \mathcal{E} \circ \dagger]\big)^\dagger \\ &= \mathscr{D}[\dagger \circ \mathcal{E} \circ \dagger]^\dagger(\rho \otimes 1_{\mathcal{B}}) = \mathscr{D}[\mathcal{E}](\rho \otimes 1_{\mathcal{B}}) \\ &= \mathcal{E} \star_{\mathrm{R}} \rho,\end{aligned}$$

where we use the second identity in Eq. (A10) of Lemma 3 in the second-to-last equality.

[132] M. Hayashi, *Quantum Information Theory* (Springer-Verlag, Heidelberg, 2016), 2nd ed.

[133] Axiom (T3) is weaker than what appears in Refs. [32,132], where the latter demands $\Theta_{\rho\otimes\rho'} = \Theta_\rho \otimes \Theta_{\rho'}$. See Ref. [134] for the significance of this.

[134] As mentioned briefly in Ref. [133], our axiom (T3) is weaker than in Refs. [32,132]. If we had used $\Theta_{\rho\otimes\rho'} = \Theta_\rho \otimes \Theta_{\rho'}$ as in Refs. [32,132], then the symmetric bloom and (r,s) family, with $s \in (0,1)$, would *not* satisfy this axiom (however, it does hold if $s \in \{0,1\}$). It suffices to illustrate this in the case of the symmetric bloom. A counterexample can be obtained for $\mathcal{A} = \mathbb{M}_2 = \mathcal{A}'$ by taking $\rho = \left[\begin{smallmatrix} p & 0 \\ 0 & 1-p \end{smallmatrix}\right]$, $\rho' = \left[\begin{smallmatrix} p' & 0 \\ 0 & 1-p' \end{smallmatrix}\right]$, $A = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] = A'$, and $p, p' \in [0, \frac{1}{2}) \cup (\frac{1}{2}, 1]$. Indeed, $\Theta^{\mathrm{J}}_{\rho\otimes\rho'}(A \otimes A') \neq \Theta^{\mathrm{J}}_\rho(A) \otimes \Theta^{\mathrm{J}}_{\rho'}(A')$.

[135] D. Petz, in *Quantum Probability and Applications III*, edited by L. Accardi and W. von Waldenfels (Springer-Verlag, Berlin, 1988), p. 251.

[136] M. Tsang, Operational meaning of a generalized conditional expectation in quantum metrology, arXiv:2212.13162 [quant-ph] (2022).

[137] J. Cotler, C.-M. Jian, X.-L. Qi, and F. Wilczek, Superdensity operators for spacetime quantum mechanics, J. High Energ. Phys. **2018**, 93 (2018).

[138] J. Cotler, X. Han, X.-L. Qi, and Z. Yang, Quantum causal influence, J. High Energ. Phys. **2019** (2019).

[139] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, Multiple-time states and multiple-time measurements in quantum mechanics, Phys. Rev. A **79**, 052110 (2009).

[140] Z. Huang and X.-K. Guo, Leggett-Garg inequalities for multitime processes, arXiv:2211.13396 [quant-ph] (2022).

[141] E. Witten, APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory, Rev. Mod. Phys. **90**, 045003 (2018).

[142] A. Mollabashi, N. Shiba, T. Takayanagi, K. Tamaoka, and Z. Wei, Pseudo-Entropy in Free Quantum Field Theories, Phys. Rev. Lett. **126**, 081601 (2021).

[143] A. Mollabashi, N. Shiba, T. Takayanagi, K. Tamaoka, and Z. Wei, Aspects of pseudoentropy in field theories, Phys. Rev. Res. **3**, 033254 (2021).

[144] K. Doi, J. Harper, A. Mollabashi, T. Takayanagi, and Y. Taki, Pseudoentropy in dS/CFT and Timelike Entanglement Entropy, Phys. Rev. Lett. **130**, 031601 (2023).

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

[146] B. Schumacher and M. A. Nielsen, Quantum data processing and error correction, Phys. Rev. A **54**, 2629 (1996).

[147] S. W. Hawking, Particle creation by black holes, Comm. Math. Phys. **43**, 199 (1975).

[148] S. W. Hawking, Breakdown of predictability in gravitational collapse, Phys. Rev. D **14**, 2460 (1976).

[149] S. W. Hawking, The unpredictability of quantum gravity, Comm. Math. Phys. **87**, 395 (1982).

[150] The Hilbert-Schmidt adjoint T^* of a conjugate-linear map $T : \mathcal{H} \to \mathcal{K}$ between Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ is defined by $\langle v, T^* w\rangle_{\mathcal{H}} = \langle w, Tv\rangle_{\mathcal{K}}$ for all $v \in \mathcal{H}$ and $w \in \mathcal{K}$. In this case, the conjugate-linear map in question is † itself. One can show that † is self-adjoint with respect to this definition.

[151] J. de Pillis, Linear transformations which preserve Hermitian and positive semidefinite operators, Pac. J. Math. **23**, 129 (1967).

[152] R. D. Hill, Linear transformations which preserve Hermitian matrices, Linear Algebra Appl. **6**, 257 (1973).

[153] As with some earlier proofs, the sequence of calculations from this proof is much more easily visualized using the string diagrams of quantum Markov categories [28].