Quantum Bayesian Inference in Quasiprobability Representations

Clive Cenxin Aw[1,*] Kelvin Onggadinata,[1] Dagomir Kaszlikowski,[1,2] and Valerio Scarani[1,2]

[1]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore*

[2]*Department of Physics, National University of Singapore, 2 Science Drive 3, 117542, Singapore*

Bayes' rule plays a crucial role in logical inference in information and physical sciences alike. Its extension into the quantum regime has been the subject of several recent works. These quantum versions of Bayes' rule have been expressed in the language of Hilbert spaces. In this paper, we derive the expression for the Petz recovery map within any quasiprobability representation, with explicit formulas for the two canonical choices of "normal quasiprobability representations" (which include discrete-Wigner representations) and of representations based on symmetric informationally complete positive operator-valued measures (SIC-POVMs). By using the same mathematical syntax of (quasi)stochastic matrices acting on (quasi)stochastic vectors, the core difference in logical inference between classical and quantum theory is found in the manipulation of the reference prior rather than in the representation of the channel.

I. INTRODUCTION

Inference is a logical necessity in every science. In information theory and physics, the fundamentality of inference is particularly apparent in notions of process reversibility and state recovery. Here, the most empirically applied and canonical approach is Bayes' rule:

$$\tilde{\mathcal{E}}_\gamma(a|a') = \mathcal{E}(a'|a)\frac{\gamma(a)}{\tilde{\gamma}(a')}. \tag{1}$$

This relation gives us a recipe for obtaining various probability-theoretic objects [1–4]. Of particular note, we may use it to obtain the "reverse" transition $\tilde{\mathcal{E}}_\gamma$ for (i) any given forward process or *transformation* $\mathcal{E}$ and (ii) the reference *prior* γ on the input of said process. The *posterior*, $\tilde{\gamma}(a') = \sum_a \mathcal{E}(a'|a)\gamma(a)$, emerges from these two objects.

While the typical form of Bayes' rule works naturally for classical information theory, an extension to quantum theory requires some work. As one possible reason for this, note that in a classical process $a \to a'$, the joint probability distribution $P(a, a')$ is routinely defined; and from this, one can compute marginal and conditional probabilities. By contrast, for a quantum process $\alpha \to \alpha' = \mathcal{E}(\alpha)$, where $\mathcal{E}$ is a completely positive trace-preserving (CPTP) map, there is no elementary way to construct a state that takes both the input and the output into account. Various proposals have been presented over the years and we refer to a very recent consolidating framework for all the references [5]. In this context of finding a quantum theoretic Bayes' rule, a special role is played by the *Petz recovery map* [6–8]:

$$\hat{\mathcal{E}}_\gamma[\bullet] = \sqrt{\gamma}\,\mathcal{E}^\dagger\left[\frac{1}{\sqrt{\mathcal{E}[\gamma]}} \bullet \frac{1}{\sqrt{\mathcal{E}[\gamma]}}\right]\sqrt{\gamma}. \tag{2}$$

This recovery channel is defined for any CPTP map $\mathcal{E}$ and a reference density operator γ. Notably, when reference priors, input states, and the channel share the same eigenbases, the Petz map reduces to the classical Bayes' rule [5,8–10]. This and other properties pertaining to what may be called the "conservation of divergences" (which is what led to its conception) have built up the reputation of this recovery map as the "quantum Bayes' rule" [11]—a reputation recently vindicated in an axiomatic approach [12]. The Petz-map construction also appears naturally in the definition of fluctuation theorems in thermodynamics [13–15].

Now, that said, it seems that what exactly makes the Petz map similar (or different) to the classical Bayesian update has not been as formalized as it could be. From an information-theoretical perspective, there are correspondences between the action of these recipes. Yet, we know that there are key regime differences in the woodwork. This lack of formal comparison across these regimes is at least partially because the Petz map has thus far only been understood in terms of CPTP maps and density operators, living in a Hilbert space. Meanwhile, the classical

---

*e0006371@u.nus.edu

Bayes' rule exists as a stochastic matrix mapping stochastic vectors, living in a real vector space. While, as we note, it is known that some CPTP maps correspond to the classical case, the reverse of this—writing Petz maps in terms of matrices and vectors, without Hilbert-space formalism—has yet to be done.

In this paper, we attempt to close this gap by investigating the Petz map in *quasiprobability representation* (QPR) [16,17]. This formalism provides a complete description of quantum theories while sharing the familiar mathematical equipment found in classical probability theory. The distinction is that *quasiprobabilities* (or "negative probabilities") are generally necessary in the quantum case [18]. This negativity has been attributed as a resource for advantage in quantum computation [19–21]. As such, we seek to put the Petz maps in the same formal habitat as that of classical Bayesian inversion and in an expression that is comparable to it. From there, we may discuss the similarities, differences, and interpretations wherever they are appropriate. We believe that this work contributes a formal step forward in understanding the essential distinctions between classical and quantum inference.

This paper is set out as follows. In Sec. II, we review features of Bayesian inference for classical and quantum transformations. In Sec. III, we review the formalisms of QPR in quantum theory. Readers familiar with the formal content here may skim through these sections. In Sec. IV, we work toward the key expression of the Petz map in QPR, stating relevant theorems along the way. In Sec. V, we discuss consequential theoretical observations, contrasting notable formal features of the expression to the classical Bayesian update. In Sec. VI, we introduce quasiprobabilistic "transition graphs" that can help visualize the implications of our results. Finally, in Sec. VII, we summarize our findings and state some open lines of inquiry.

II. CLASSICAL AND QUANTUM BAYESIAN INFERENCE

In the context of classical mechanics and probability theory, a physical transformation can be expressed by conditional probabilities $\mathcal{E}(a'|a)$ mapping probability distributions of inputs $p(a)$ to distributions of outputs $\tilde{p}(a') = \sum_a \mathcal{E}(a'|a)p(a)$ residing in some given state space A [22]. This can be captured compactly by a stochastic matrix $S^{\mathcal{E}} = \{\mathcal{E}(a'|a)\}$, mapping $v^p = \{p(a)\}$ to $v^{\tilde{p}} = \{\tilde{p}(a')\}$.

As already discussed, if we want to acquire a stochastically valid and logically sound "reverse" of this transformation $\mathcal{E}$, we must invoke not only the channel in question but also a *reference prior* γ on the input. This is essentially a preexisting best guess of the inputs for which the Bayesian inverse is constructed. This process of acquiring $\tilde{\mathcal{E}}_\gamma$ from $\mathcal{E}$ and γ can be referred to as performing "retrodiction" (inference about the past, in contrast to prediction,

which is inference about the future) for $\mathcal{E}$ on the prior γ. Meanwhile, $S^{\tilde{\mathcal{E}}_\gamma} v^{\tilde{p}}$ gives the "retrodicted input" given an observation $\tilde{p}$. It may also be referred to as the "Bayesian update on γ given $\tilde{p}$."

For every each *individual* transition, we may consult Eq. (1) for the corresponding retrodiction $a' \rightarrow a$. For the mapping of *distributions*, it is more instructive to write the retrodiction map as a stochastic matrix:

$$S_{\text{CL}}^{\tilde{\mathcal{E}}_\gamma} = D_\gamma (S^{\mathcal{E}})^T D_{\mathcal{E}[\gamma]}^{-1}. \tag{3}$$

Here, D_p is a diagonal matrix with entries corresponding to some distribution p.

As introduced in Sec. I, the counterpart to the Bayes' rule in quantum theory is the Petz map given in Eq. (2). It is well defined and CPTP for any full-rank $\mathcal{E}[\gamma]$ [23]. It may also be expressed as

$$\hat{\mathcal{E}}_\gamma = \mathcal{M}_{\gamma^{1/2}} \circ \mathcal{E}^\dagger \circ \mathcal{M}_{\mathcal{E}[\gamma]^{-1/2}}, \tag{4}$$

where $\mathcal{M}_{\alpha^r}[\bullet] = \alpha^r \bullet \alpha^r$ for any density operator α and $r \in \mathbb{R}$, and $\mathcal{E}^\dagger$ is the adjoint of $\mathcal{E}$. This is the unique map for which

$$\text{Tr}(\mathcal{E}[\rho]\sigma) = \text{Tr}(\mathcal{E}^\dagger[\sigma]\rho), \tag{5}$$

for all self-adjoint ρ and σ.

Before continuing, it is important to stress that *Bayesian inference is generically not inversion*. Inference is possible for any map, while inversion is only possible for invertible maps (information-preserving)—and even then, the two operations are generally not the same, since the inverse of a map is generically not a valid map. In fact, it can be proved that inference and inversion coincide if and only if $S^{\mathcal{E}}$ is a permutation (for the classical case) or $\mathcal{E}$ is a unitary channel (in the quantum case) [15,24]. In general, therefore, $S_{\text{CL}}^{\tilde{\mathcal{E}}_\gamma} S^{\mathcal{E}} v^\rho \neq v^\rho$ and $\hat{\mathcal{E}}_\gamma \circ \mathcal{E}[\rho] \neq \rho$; although the reference state is recovered: $S_{\text{CL}}^{\tilde{\mathcal{E}}_\gamma} S^{\mathcal{E}} v^\gamma = v^\gamma$ and $\hat{\mathcal{E}}_\gamma \circ \mathcal{E}[\gamma] = \gamma$ for all γ.

III. QUASIPROBABILITY REPRESENTATIONS

A. Generalities

We now move on to provide a brief review of the essential elements of QPRs for quantum theory. To map quantum theoretic objects acting on a d-dimensional Hilbert space to a QPR, the core is the choice of a *frame* $\{F_j\}_{j \in \Lambda}$, i.e., a set of $d \times d$ Hermitian operators spanning the Hermitian space equipped with Hilbert-Schmidt scalar product. The set of indices Λ may be continuous; its minimal cardinality is d^2 and we assume such minimal frames for the remainder of this paper. Given a frame, one can always

find a *dual frame* $\{G_j\}_{j\in\Lambda}$ such that

$$\forall\,\alpha,\beta : \sum_j \text{Tr}[F_j\alpha]\,\text{Tr}[G_j\beta] = \text{Tr}[\alpha\beta]. \qquad (6)$$

In general, the dual is not unique given a frame. However, for a minimal basis, the frame and dual always satisfy the orthogonality relation $\text{Tr}[F_j G_k] = \delta_{jk}$.

Once a frame and its dual are set, all Hilbert-space objects are in one-to-one correspondence with an object in the QPR. A dictionary of recipes between the two frameworks is summarized in Table I. These recipes can be understood as category-theoretic relationships called "functors," mapping objects across the two formalisms, each of which lives in its own separate category [25,26]. From the normalization of the state quasiprobability, $\sum_a v_a^\rho = 1$, it follows that the frame operators must satisfy $\sum_a F_a = \mathbb{1}$. Similarly, from the fact that each POVM $\{E_m\}$ must satisfy $\sum_m E_m = \mathbb{1}$, it follows that $\text{Tr}[G_j] = 1$ for all $j \in \Lambda$ [27]. As such, the QPR of any CPTP map $\mathcal{E}$ is a quasistochastic matrix $S^{\mathcal{E}}$ as defined in Table I, with entries $S_{a'a}^{\mathcal{E}} \in \mathbb{R}$, $a', a \in \Lambda$, such that, $\forall a$ $\sum_{a'} S_{a'a}^{\mathcal{E}} = 1$. The evolution of a state through a channel is then described by simple matrix multiplication: $\rho' = \mathcal{E}[\rho] \rightarrow v_{a'}^{\rho'} = \sum_a S_{a'a}^{\mathcal{E}} v_a^\rho$ [28]. With a slight abuse of notation, for ease of correspondence with the classical formalism, we also denote the elements of the quasistochastic matrix as $S_{a'a}^{\mathcal{E}} \equiv \mathcal{E}(a'|a)$.

Some of the subsequent derivations apply generally to all representations; others are specific to one of the two canonical choices of QPR that we describe next.

B. Normal quasiprobability representation

The first class of representations are those for which the frame and dual-frame operators are proportional to each other up to some scaling factor c, i.e.,

$$G_j = cF_j, \qquad (7)$$

for all j. For minimal bases, the constant c is equal to the Hilbert-space dimension d. The class of representations satisfying this is known as *normal quasiprobability representation* (NQPR) [29].

An example of NQPR, and perhaps the most widely used representation, is the *discrete-Wigner* (DW) representation [30–33], which is well defined for prime dimension d and composites of them. For odd primes, the frame operators are defined as

$$F_k = F_{r,s} = \frac{1}{d^2}\sum_{x,z=0}^{d-1}\omega^{sx-rz+\frac{xz}{2}}X^xZ^z, \qquad (8)$$

where $k = (r,s) \in \mathbb{Z}_d \times \mathbb{Z}_d$, $\omega = e^{2\pi i/d}$ is the dth root of unity and Z and X are generalized Pauli operators defined as $Z|j\rangle = \omega^j|j\rangle$ and $X|j\rangle = |j+1 \mod d\rangle$, with $\{|j\rangle\}_{j=0}^{d-1}$ as standard orthonormal basis. For a qubit system ($d = 2$), the frame has a simple expression given by

$$F_k = F_{r,s} = \frac{1}{4}\Big[\mathbb{1} + (-1)^r\sigma_x + (-1)^s\sigma_z + (-1)^{r+s}\sigma_y\Big], \qquad (9)$$

where $k = (r,s) \in \mathbb{Z}_2 \times \mathbb{Z}_2$ and σ_x, σ_y, and σ_z are the familiar Pauli operators. For composite $d = d_1 \times d_2 \times \cdots \times d_L$, where $d_1, d_2, \ldots, d_L$ are primes, a tensor structure applies for the frame. That is, the frame operators decompose as

$$F_k = F_{k_1} \otimes F_{k_2} \otimes \cdots \otimes F_{k_L},$$

where $k \rightarrow (k_1, k_2, \ldots, k_L)$ with each $k_l = (r_l, s_l) \in \mathbb{Z}_{d_l} \times \mathbb{Z}_{d_l}$. This tensor structure is enjoyed by any NQPR and thus affords them an esthetic benefit when dealing with composite states and purifications.

C. SIC-POVM representation

Under NQPR, negativity can be found in states, POVM elements, and transformations alike. Symmetric informationally complete positive operator-valued measure (SIC-POVM) representations seek to avoid this by ensuring that all state vectors are positive [34,35]. Negativity features are thus consolidated into the transformations and POVMs.

For d-dimensional Hilbert space, a SIC-POVM is defined as a set of subnormalized rank-1 projectors

TABLE I. The dictionary of relationships between the Hilbert-space and quasiprobability formalisms. $v_a^p = p(a)$ indicates the ath entry in a p distribution. $S_{a'a}^{\mathcal{E}} = \mathcal{E}(a'|a)$ indicates the entry on the a'th column and ath row of a matrix $S^{\mathcal{E}}$. The frame and dual operators F_i and G_j are defined within choices of representations as discussed in Secs. III B and III C, with the most direct expressions found in Eqs. (7), (8), (10), and (11).

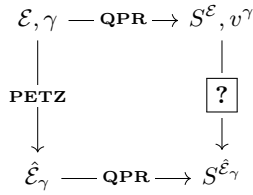| Object	Hilbert-space formalism	Quasiprobability formalism
State	$\rho = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$	$v^\rho \; : \; v_a^\rho = \text{Tr}[\rho\,F_a]$
POVM	$\{E_m \mid E_m \geq 0, \sum_m E_m = \mathbb{1}\}$	$\bar{v}^m \; : \; \bar{v}_{a'}^m = \text{Tr}[E_m\,G_{a'}]$
Unitary	$\mathcal{U}[\bullet] = U \bullet U^\dagger, \; UU^\dagger = \mathbb{1}$	$S^{\mathcal{U}}: S_{a'a}^{\mathcal{U}} = \text{Tr}[F_{a'} U G_a U^\dagger]$
CP maps	$\mathcal{E}[\bullet] = \sum_l \kappa_l \bullet \kappa_l^\dagger$	$S^{\mathcal{E}} \; : \; S_{a'a}^{\mathcal{E}} = \text{Tr}[F_{a'}\mathcal{E}[G_a]]$
Born rule	$\text{Tr}[\rho E_m]$	$v^\rho \cdot \bar{v}^m \in [0,1]$
Dimensionality	$\dim[\mathbb{C}^d] = d$	$\dim[\mathbb{R}^d \otimes \mathbb{R}^d] = d^2$

$$\mathcal{E}, \gamma \xrightarrow{\ \mathbf{QPR}\ } S^{\mathcal{E}}, v^\gamma$$

$$\mathbf{PETZ} \qquad\qquad \boxed{?}$$

$$\hat{\mathcal{E}}_\gamma \xrightarrow{\ \mathbf{QPR}\ } S^{\hat{\mathcal{E}}_\gamma}$$

FIG. 1. A commutativity diagram illustrating the main task of this work: the protocol "**?**" that is to be executable solely within the QPR framework.

$\{\frac{1}{d}\Pi_j\}_{j=1}^{d^2}$, $\Pi_j = |\psi_j\rangle\langle\psi_j|$, such that the elements have equal pairwise Hilbert-Schmidt inner product:

$$\mathrm{Tr}\Big[\Pi_j^\dagger \Pi_k\Big] = |\langle\psi_j|\psi_k\rangle|^2 = \frac{d\delta_{jk}+1}{d+1}. \qquad (10)$$

The solutions to the vectors of SIC-POVM have been found for vast number of dimensions (for the list, see Ref. [36]) and is believed to exist for all [37]. Since the set is informationally complete (i.e., it forms a basis), we can use it as the definition of the frame of the SIC-POVM representation $\{F_j = \frac{1}{d}\Pi_j\}$. From the orthogonality relation, it can be easily deduced that the dual frame is given by

$$G_j = d(d+1)F_j - \mathbb{1} = (d+1)\Pi_j - \mathbb{1}. \qquad (11)$$

For one qubit, the canonical choice is the tetrahedron

$$F_j = \frac{1}{4}\Big[\mathbb{1} + \vec{v}_j \cdot \vec{\sigma}\Big], \qquad (12)$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\vec{v}_0 = \frac{1}{\sqrt{3}}(1,-1,1)$, $\vec{v}_1 = \frac{1}{\sqrt{3}}(-1,1,1)$, $\vec{v}_2 = \frac{1}{\sqrt{3}}(1,1,-1)$, and $\vec{v}_3 = \frac{1}{\sqrt{3}}(-1,-1,-1)$.

IV. THE PETZ MAP IN QUASIPROBABILITY FORMALISMS

Now, our task is to express the Petz recovery map in its QPR, which we denote as $S^{\hat{\mathcal{E}}_\gamma}$. This can obviously be done by invoking the relationship between maps in Table I and then connecting it with Eq. (2). This gives

$$S^{\hat{\mathcal{E}}_\gamma}_{aa'} = \mathrm{Tr}\left[F_a \sqrt{\gamma}\mathcal{E}^\dagger\left[\frac{1}{\sqrt{\mathcal{E}[\gamma]}}G_{a'}\frac{1}{\sqrt{\mathcal{E}[\gamma]}}\right]\sqrt{\gamma}\right]. \qquad (13)$$

But, of course, this affords us no new insight. We are still relying entirely on the Hilbert-space formalism. Nothing novel can be said in comparison to classical Bayesian inference as found in Eq. (3). Our specific task is as illustrated in Fig. 1: write the Petz in a way that *only quasiprobability-theoretic objects* (quasistochastic vectors, matrices, and frames) *are required*.

The naive guess that $S^{\hat{\mathcal{E}}_\gamma}$ could be obtained by grafting the quasiprobabilistic formalism onto the classical Bayesian inverse given in Eq. (3) is easily dismissed: the $S^{\tilde{\mathcal{E}}_\gamma}_{\mathbf{CL}}$ obtained by such a recipe is in general not a valid map in QPR (for explicit counterexamples, see Appendix F), as it results in measurement outcome statistics that are out of bounds. That is,

$$\exists(\mathcal{E}, \gamma, \rho, m) : (S^{\tilde{\mathcal{E}}_\gamma}_{\mathbf{CL}} v^\rho) \cdot \vec{v}^m \notin [0,1].$$

Rather, let us start by noting that the recipe for channels in Table I and the condition given in Eq. (6) imply the concatenation

$$S^{\mathcal{E}} S^{\mathcal{F}} = S^{\mathcal{E} \circ \mathcal{F}} \qquad (14)$$

for two channels $\mathcal{E}$ and $\mathcal{F}$. Thus the Petz map given in Eq. (4) is represented by

$$S^{\hat{\mathcal{E}}_\gamma} = M_{\gamma^{1/2}}\left(S^{\mathcal{E}^\dagger}\right)M_{\mathcal{E}[\gamma]^{-1/2}}, \qquad (15)$$

with

$$M_{\alpha^r} := S^{\alpha^r \bullet \alpha^r}; \qquad (16)$$

i.e.,

$$(M_{\alpha^r})_{a'a} = \mathrm{Tr}[F_{a'}\alpha^r G_a \alpha^r]. \qquad (17)$$

Now, it is crucial for our goals that all objects entering Eq. (15) can be constructed within the quasiprobability formalism.

As a first check, we note that all the entries of the matrices M_{α^r} are real. Indeed, one can rewrite $(M_{\alpha^r})_{a'a} = \mathrm{Tr}[\mathcal{F}^r_{a'}\mathcal{G}^r_a]$ with $\mathcal{F}^r_a = \alpha^{r/2}F_a\alpha^{r/2}$ and $\mathcal{G}^r_a = \alpha^{r/2}G_a\alpha^{r/2}$. These are Hermitian operators, and so $\mathrm{Tr}[\mathcal{F}^r_{a'}\mathcal{G}^r_a] = \frac{1}{2}\mathrm{Tr}[\{\mathcal{F}^r_{a'}, \mathcal{G}^r_a\}]$ is real.

Next, we show that M_α can be expressed using the quasiprobability representation of the state $\alpha = \sum_x v^\alpha_x G_x$. Indeed,

$$(M_\alpha)_{a'a} = \mathrm{Tr}[F_{a'}\alpha G_a \alpha]$$
$$= \sum_{xy} v^\alpha_x v^\alpha_y \mathrm{Tr}[F_{a'}G_x G_a G_y] \qquad (18)$$
$$:= \sum_{xy} v^\alpha_x v^\alpha_y \xi_{a'xay}, \qquad (19)$$

where the

$$\xi_{pqrs} = \mathrm{Tr}[F_p G_q G_r G_s] \qquad (20)$$

are referred to as *structure coefficients*. These same coefficients have appeared in a recent work on the DW representation of maps [38] (see Appendix A). The M_α are real

positive-semidefinite matrices with a unit trace: $M_\alpha \geq 0$ and $\text{Tr}[M_\alpha] = 1$. For NQPRs, they are Hermitian, while for SIC-POVMs they are generically not symmetric and thus not Hermitian (see Appendix E).

Having expressed M_α in the QPR formalism, one can finally prove that

$$M_{\alpha^r} = M_\alpha^r \qquad (21)$$

holds for any $r \in \mathbb{Q}$ (i.e., for any rational number r; for the proof, see Appendix B). In particular, the M_{α^r} for $r = \pm\frac{1}{2}$ that are needed in Eq. (15) can be constructed from the quasiprobability representation of states by first computing M_α Eq. (19) and then taking the suitable roots.

In summary, we obtain our main result:

Result. The Petz map in any QPR reads

$$S_{\text{QM}}^{\hat{\mathcal{E}}_\gamma} = M_\gamma^{1/2}\left(S^{\mathcal{E}^\dagger}\right) M_{\mathcal{E}[\gamma]}^{-1/2}, \qquad (22)$$

where

$$(M_\gamma)_{a'a} = \sum_{xy} v_x^\gamma v_y^\gamma \, \xi_{a'xay},$$

$$\left(M_{\mathcal{E}[\gamma]}\right)_{a'a} = \sum_{xy} (S^{\mathcal{E}} v^\gamma)_x (S^{\mathcal{E}} v^\gamma)_y \, \xi_{a'xay},$$

and $\xi_{pqrs} = \text{Tr}[F_p G_q G_r G_s]$ are structure coefficients determined by the specific QPR. Everything is expressed exclusively in the quasiprobabilistic formalism: no knowledge of Hilbert-space renditions of the quantum channel or reference state is required.

For the two canonical choices of QPR introduced above, we prove in Appendix C that

$$\text{NQPR} : \ S_{\text{NQ}}^{\mathcal{E}^\dagger} = (S^{\mathcal{E}})^T, \qquad (23)$$

$$\text{SIC-POVM} : \ S_{\text{SP}}^{\mathcal{E}^\dagger} = (S^{\mathcal{E}})^T + J_{\mathcal{E}}, \qquad (24)$$

where $(J_{\mathcal{E}})_{ij} = \frac{1}{d}(\sum_a \mathcal{E}(j\,|a) - 1)$; whence explicitly

$$S_{\text{NQ}}^{\hat{\mathcal{E}}_\gamma} = M_\gamma^{1/2}(S^{\mathcal{E}})^T M_{\mathcal{E}[\gamma]}^{-1/2}, \qquad (25)$$

$$S_{\text{SP}}^{\hat{\mathcal{E}}_\gamma} = M_\gamma^{1/2}\left[(S^{\mathcal{E}})^T + J_{\mathcal{E}}\right] M_{\mathcal{E}[\gamma]}^{-1/2}. \qquad (26)$$

Since the QPRs of unital maps (i.e., $\mathcal{E}[\mathbb{1}] = \mathbb{1}$) are quasi-bistochastic matrices (i.e., $\sum_a \mathcal{E}(j\,|a) = 1$ for all j), for such maps $J_{\mathcal{E}}$ vanishes and the expressions for the NQPR and SIC-POVM representations are formally identical.

TABLE II. Retrodiction maps for classical probabilities $[S_{\mathcal{T}}^{\bar{\mathcal{E}}_\gamma} \to S_{\text{CL}}^{\bar{\mathcal{E}}_\gamma}$, Eqs. (3) or (27)] and quantum quasiprobabilities $[S_{\mathcal{T}}^{\bar{\mathcal{E}}_\gamma} \to S_{\text{QM}}^{\hat{\mathcal{E}}_\gamma}$, Eq. (22)]. The law of transformation of the central object is the same in both classical and quantum theory, while one needs the structure coefficients to obtain the X_γ from the v_γ.

Bayesian inference in theory $\mathcal{T}$		
$S_{\mathcal{T}}^{\bar{\mathcal{E}}_\gamma} = X_\gamma^{1/2}\left(S^{\mathcal{E}^\dagger}\right) X_{\mathcal{E}[\gamma]}^{-1/2}$		
$(X_\gamma)_j = \sum_{xy} v_x^\gamma v_y^\gamma \xi_{ixjy}$		
Object	$\mathcal{T}$: quantum	$\mathcal{T}$: classical
$S^{\mathcal{E}^\dagger}$	NQ: $(S^{\mathcal{E}})^T$	$(S^{\mathcal{E}})^T$
	SP: $(S^{\mathcal{E}})^T + J_{\mathcal{E}}$	
ξ_{ixjy}	$\text{Tr}[F_i G_x G_j G_y]$	$\delta_{ix}\delta_{jy}\delta_{ij}$

V. COMPARING CLASSICAL AND QUANTUM RETRODICTION

A. Main comparison

We are finally in a position to compare classical and quantum Bayesian inference. Having found Eq. (22), it is easy to note that the classical Bayes' rule given in Eq. (3) can be rewritten in the same form,

$$S_{\text{CL}}^{\bar{\mathcal{E}}_\gamma} = (D_\gamma^2)^{1/2}\left(S^{\mathcal{E}^\dagger}\right)(D_{\mathcal{E}[\gamma]}^2)^{-1/2}, \qquad (27)$$

because $(S^{\mathcal{E}})^T = S^{\mathcal{E}^\dagger}$ for classical channels (see Appendix D) and $(D_\gamma)_{ij} = v_i^\gamma \delta_{ij}$. In other words, classical Bayesian inference hides the fact *that the central matrix is an adjoint* and that *the left and right matrices should be seen as square roots of more fundamental matrices X_γ and $X_{\mathcal{E}[\gamma]}$*. This is the *common form* of classical and quantum Bayesian inference that emerges from using QPRs.

Let us now study the *differences* between the two theories. There is, of course, the starting point: for a given system dimension d, in classical theory the v^γ are d-component probability vectors, while in a minimal QPR of quantum theory they are d^2-component quasiprobability vectors. If we leave this aside and focus only on Bayesian inference, the *formal difference* appears in the matrices X_γ (D_γ^2 for classical, M_γ for quantum). Both can be written

$$(X_\gamma)_{ij} = \sum_{x,y=1}^{d^2} v_x^\gamma v_y^\gamma \xi_{ixjy}; \qquad (28)$$

but while the M_γ of quantum theory has $\xi_{ixjy} = \text{Tr}[F_i G_x G_j G_y]$, the D_γ^2 of classical theory is such that $(D_\gamma^2)_{ij} = (v_i^\gamma)^2 \delta_{ij}$, i.e.,

$$\xi_{ixjy} = \delta_{ix}\delta_{jy}\delta_{ij} \qquad [\text{Classical}]. \qquad (29)$$

This comparison is summarized in Table II.

B. Related remarks

Previously, we have mentioned that $S_{\mathrm{QM}}^{\hat{\mathcal{E}}_\gamma}$ is not naively equal to $S_{\mathrm{CL}}^{\tilde{\mathcal{E}}_\gamma}$. We can now be more precise about this difference. First, we note that no frame satisfies $\mathrm{Tr}[F_i G_x G_j G_y] = \delta_{ix}\delta_{jy}\delta_{ij}$. If this were the case, M_γ would be diagonal for all γ; in fact, all conceivable quantum channels would become trivial. Here is the proof: by extending on Ref. [38] (see Appendix A), in any QPR every entry of $S^{\mathcal{E}}$ can be expressed as

$$S_{ij}^{\mathcal{E}} = \sum_{xyl} \mathrm{Tr}[F_x \kappa_l]\, \mathrm{Tr}\left[F_y \kappa_l^\dagger\right] \mathrm{Tr}[F_i G_x G_j G_y], \qquad (30)$$

where $\{\kappa_l\}_l$ defines a Kraus representation of $\mathcal{E}$ [39]. Thus, a hypothetical frame satisfying Eq. (29) would have

$$
\begin{aligned}
S_{ij}^{\mathcal{E}} &= \sum_{xyl} \mathrm{Tr}[F_x \kappa_l]\, \mathrm{Tr}\left[F_y \kappa_l^\dagger\right] \delta_{ix}\delta_{jy}\delta_{ij} \\
&= \delta_{ij} \sum_l \mathrm{Tr}[F_i \kappa_l]\, \mathrm{Tr}\left[F_j \kappa_l^\dagger\right] = \delta_{ij},
\end{aligned}
$$

where the last equality comes from the fact that $S^{\mathcal{E}}$ must be quasistochastic. In other words, we would have $S^{\mathcal{E}} = \mathbb{1}_{d^2}$ for every $\mathcal{E}$, which is absurd [40]. We conjecture that the only way to obtain $\mathrm{Tr}[F_i G_x G_j G_y] = \delta_{ix}\delta_{jy}\delta_{ij}$ with $d \times d$ matrices is to set $F_i = G_i = |i\rangle\langle i|$ for $1 \le i \le d$ and $\langle i|j\rangle = \delta_{ij}$ but this defines a projective measurement; it has only d elements and is certainly not a frame.

Next, one might conjecture that $(S_{\mathrm{QM}}^{\hat{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m = (S_{\mathrm{CL}}^{\tilde{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m$ holds for classical processes: i.e., processes in which ρ, γ, $\mathcal{E}[\gamma]$, and E_m all have the same eigenbasis and no coherence appears. While our numerical exploration suggests that this is the case for some specific scenarios, it is certainly false in general, even for simple examples (see Appendix G).

A last question is whether the quantum Bayes' rule can be written as

$$\hat{\mathcal{E}}_\gamma(a|a') = f\left(\{\mathcal{E}(a'|a)\},\ \{\gamma(a)\}\right) \qquad (31)$$

in analogy to Eq. (1). The answer is: *in principle, yes*, because the square roots of M_γ and $M_{\mathcal{E}[\gamma]}$ are certainly functions of the $\mathcal{E}(a'|a)$ and the $\gamma(a)$. However, writing down this expression *in practice* requires the explicit expressions. Even for the simplest quantum case (the qubit), in general one would have to find the roots of a quartic characteristic equation.

VI. VISUALIZING QUANTUM INFERENCE VIA QPR

A. Introducing transition graphs

A notable advantage of stochastic maps is their ease of visualization. One can draw what might be called "transition graphs," where transitions between a_i and a'_j are depicted by arrows going from the former to the latter. The probability weights on these transitions may be then depicted by a number or by a color function. These kinds of graphs are not straightforward to write for the standard Hilbert-space formalism. This is simply due to the use of complex terms, probability amplitudes, and the plurality of possible basis choices. With QPR, we can illustrate transformations and their quantum Bayesian inverses with transition graphs just as we would for classical stochastic channels, albeit with the added task of depicting negativity in these transitions.

In Appendix H and this section, we consider some choices of $\mathcal{E}$ that give rise to $S^{\mathcal{E}}$ and their retrodictions $S_{\mathrm{DW}}^{\hat{\mathcal{E}}_\gamma}$ and $S_{\mathrm{SP}}^{\hat{\mathcal{E}}_\gamma}$. These are then depicted as transition graphs. We choose to include, in particular, a half-SWAP channel with a $|1\rangle\langle 1|$ ancilla to visually illustrate and explore the properties of quantum retrodiction. Other transformations are also noted in passing with their graphs and expressions consolidated in Appendix H. Before these, we note some illustrative elements of these figures.

First, with transition arrows, we depict negative (positive) quasiprobabilities with cooler (warmer) shades. Furthermore, these negative (positive) arrows are drawn using dashed (solid) lines. A color legend is included in Fig. 2(a).

Second, in order to get a sense of how irreversible a *forward* map is and which states it tends to erase toward, we add colored "bubbles" around the *output* side (denoted $\{a'_j\}$) of every graph for a given $S^{\mathcal{E}}$. The intensity and color of the bubbles are weighted according to the quasiprobability distribution of the state $\mathcal{E}[\mathbb{1}/d]$. Hence, one should expect that these bubbles are colored uniformly for all unital maps (and, thus, for all reversible maps too). Visually speaking, the most irreversible maps would be those for which these colored bubbles correspond exactly to the color of the transition arrows that are drawn toward them (for an example of this, see Fig. 6) [41].

Third, a similar feature is added for the *retrodictive* transition graphs, drawn for $S^{\hat{\mathcal{E}}_\gamma}$ matrices. Crucial for understanding the Bayesian inverse is the reference prior. Hence, for Bayesian-inverting transition graphs, we add colored bubbles on the *input* (denoted $\{a_j\}$, i.e., the input of the *forward* map) side of the graph, weighted according to the distribution of γ. For simplicity, we describe channels acting on qubits and use the most canonical choices of frames: Eq. (9) for DW (with r, s starting from 0) and Eq. (12) for
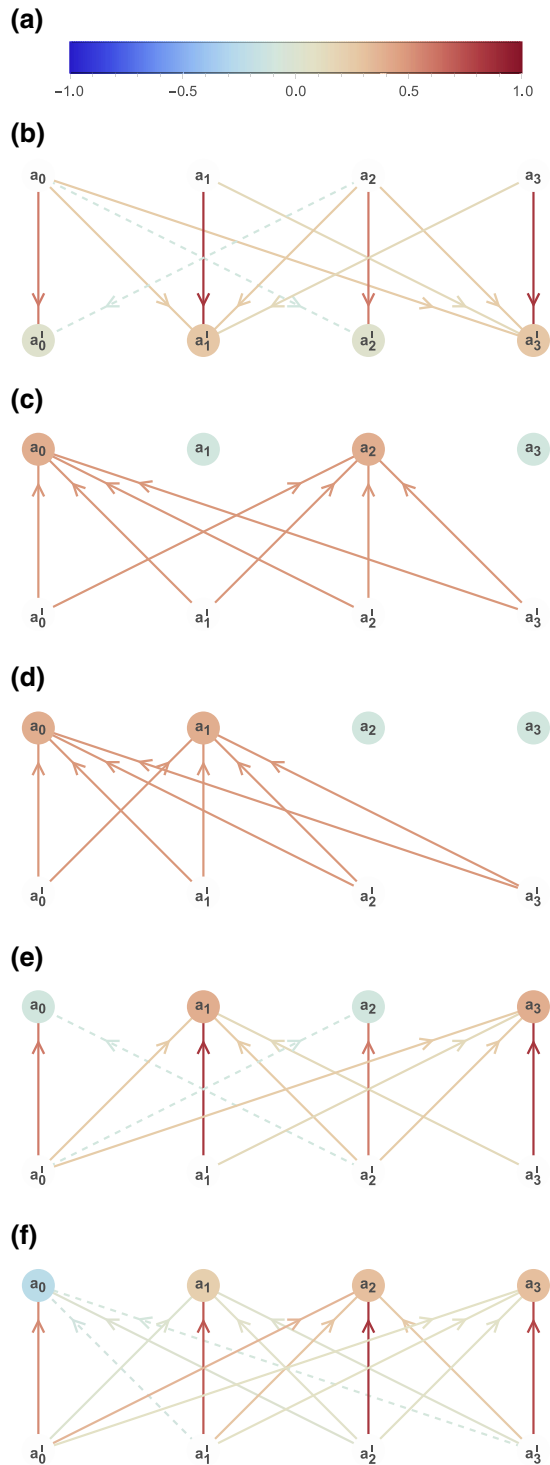
(a)

(b)

(c)

(d)

(e)

(f)

FIG. 2. Transition graphs for a "half-SWAP" with $|1\rangle\langle 1|$ and various retrodictions with a range of reference priors. (a) Color legend for quasiprobabilities. (b) $S_{\mathrm{DW}}^{\mathcal{E}}$ for $\mathcal{E}[\bullet] = \mathrm{Tr}_B[U_{\lambda} \bullet \otimes |1\rangle\langle 1| U_{\lambda}^{\dagger}]$. (c) $S_{\mathrm{DW}}^{\hat{\mathcal{E}}_{\gamma}}$ for $U_{\lambda}, \beta = |1\rangle\langle 1|, \gamma = |0\rangle\langle 0|$. (d) $S_{\mathrm{DW}}^{\hat{\mathcal{E}}_{\gamma}}$ for $U_{\lambda}, \beta = |1\rangle\langle 1|, \gamma = |+\rangle\langle +|$. (e) $S_{\mathrm{DW}}^{\hat{\mathcal{E}}_{\gamma}}$ for $U_{\lambda}, \beta = |1\rangle\langle 1|, \gamma = |1\rangle\langle 1|$. (f) $S_{\mathrm{DW}}^{\hat{\mathcal{E}}_{\gamma}}$ for $U_{\lambda}, \beta = |1\rangle\langle 1|, \gamma(\frac{\pi}{16}, \frac{\pi}{5}, \frac{\pi}{3})$ as per (33).

$$
\begin{array}{ccc}
v^{\rho} \cdot \bar{v}^{\sigma} & \longleftrightarrow & \mathrm{Tr}[\rho\,\sigma] \\
\downarrow & & \downarrow \\
(S^{\mathcal{E}}v^{\rho}) \cdot \bar{v}^{\sigma} & \longleftrightarrow & \mathrm{Tr}\left[\mathcal{E}[\rho]\sigma\right] \\
\downarrow_{\forall\rho\sigma} & & \downarrow_{\forall\rho\sigma} \\
(S^{\mathcal{E}^{\dagger}}v^{\sigma}) \cdot \bar{v}^{\rho} & \longleftrightarrow & \mathrm{Tr}\left[\mathcal{E}^{\dagger}[\sigma]\rho\right]
\end{array}
$$

FIG. 3. The relations within and functors between formalisms pertaining to the adjoint map, illustrated commutatively.

SIC-POVM representations. We employ these frames for all the numerics found in this paper.

B. Fully reversible and fully irreversible

As depicted in Figs. 4 and 5 (found in Appendix H), we observe the provable property that $S^{\hat{\mathcal{U}}_{\gamma}} = S^{\hat{\mathcal{U}}} = (S^{\mathcal{U}})^{T}$, for unitary channels $\mathcal{U}$. The Bayesian inverses simply reflect the transition trajectories back, doing so with equal probability and negativity and regardless of what reference prior is chosen. More interesting features occur for nonunitary channels. We may write any CPTP map as a dilation defined by a global unitary U acting on an extended state space $\mathcal{H}_A \otimes \mathcal{H}_B$ for which the input system $\bullet_A$ and an environment or ancilla β_B is defined:

$$\mathcal{E}[\bullet] = \mathrm{Tr}_B[U \bullet \otimes \beta\, U^{\dagger}]. \tag{32}$$

We stick to the case where both the target and the ancilla are qubits. Arbitrary qubits may be written as

$$\beta(\omega, \theta, \phi) = \sin^2(\omega)\,|\psi\rangle\,\langle\psi| + \cos^2(\omega)\,|\psi^{\perp}\rangle\,\langle\psi^{\perp}|, \tag{33}$$

where $|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\phi}\sin(\theta/2)\,|1\rangle$ and $|\psi^{\perp}\rangle = e^{-i\phi}\sin(\theta/2)\,|0\rangle + \cos(\theta/2)\,|1\rangle$. In maximal contrast to unitary channels, one may consider a quantum total-erasure channel. This is simply a kind of replacement map where a full-SWAP Eq. (H1) acts on a qubit and an ancilla and we trace out the environment. The Bayesian inverses of such quantum channels follow their classical counterparts: they erase back to the reference prior [15]. Since the channel is totally irreversible, the quantum Bayes' rule simply reverts our inference to our best guess about the initial state (illustrated by Fig. 6).

C. Liminally (ir)reversible

For a more conceptually involved and instructive scenario, we consider the half-SWAP U_{λ}, which may be represented in the computational basis as

$$U_{\lambda} \,\hat{=}\, \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{pmatrix}. \tag{34}$$
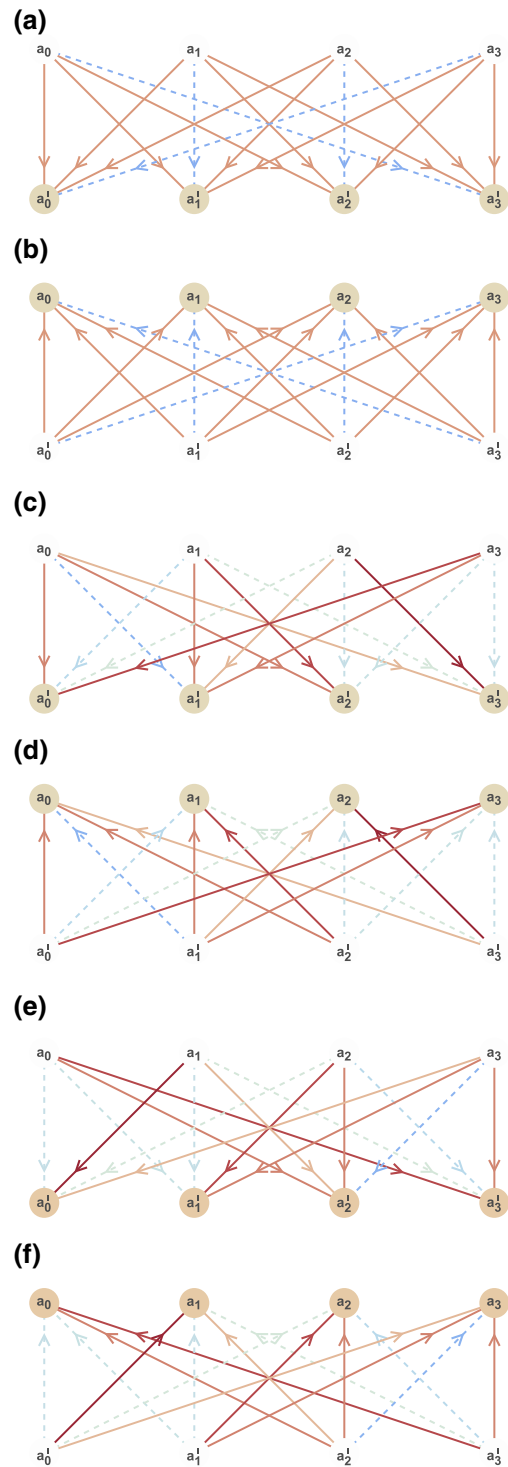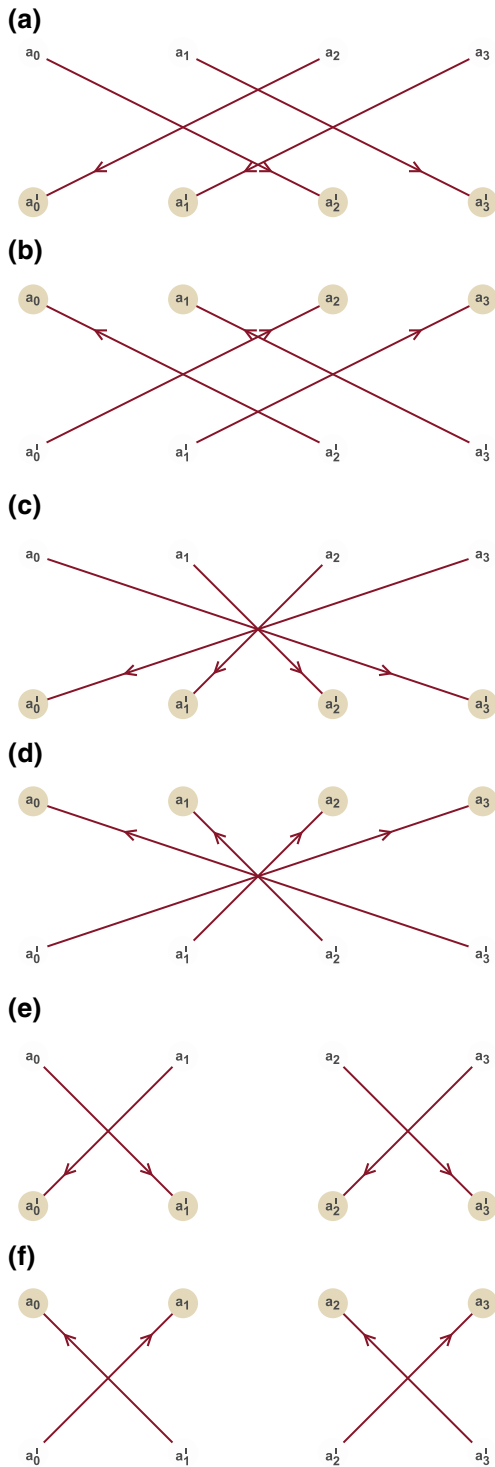
FIG. 4.   Transition graphs for $\sigma_z, \sigma_x,$ and σ_y and their respective retrodictions. (a) $S^{\mathcal{E}}$ for $\mathcal{E}[\bullet] = \sigma_z \bullet \sigma_z$. (b) $S_{\mathbf{SP}}^{\hat{\mathcal{E}}_\gamma}$ and $S_{\mathbf{NQ}}^{\hat{\mathcal{E}}_\gamma}$ for $\mathcal{E}[\bullet] = \sigma_z \bullet \sigma_z$. (c) $S^{\mathcal{E}}$ for $\mathcal{E}[\bullet] = \sigma_y \bullet \sigma_y$. (d) $S_{\mathbf{SP}}^{\hat{\mathcal{E}}_\gamma}$ and $S_{\mathbf{NQ}}^{\hat{\mathcal{E}}_\gamma}$ for $\mathcal{E}[\bullet] = \sigma_y \bullet \sigma_y$. (e) $S^{\mathcal{E}}$ for $\mathcal{E}[\bullet] = \sigma_x \bullet \sigma_x$. (f) $S_{\mathbf{SP}}^{\hat{\mathcal{E}}_\gamma}$ and $S_{\mathbf{NQ}}^{\hat{\mathcal{E}}_\gamma}$ for $\mathcal{E}[\bullet] = \sigma_x \bullet \sigma_x$.

FIG. 5.   Transition graphs for the Hadamard gate and an arbitrarily chosen qubit unitary and their respective retrodictions. (a) $S^{\mathcal{U}_{\mathrm{H}}}$ for $\mathcal{U}_{\mathrm{H}}[\bullet] = U_{\mathrm{H}} \bullet U_{\mathrm{H}}^{\dagger}$. (b) $S^{\hat{\mathcal{U}}_{\mathrm{H}}}$ for $\mathcal{U}_{\mathrm{H}}[\bullet] = U_{\mathrm{H}} \bullet U_{\mathrm{H}}^{\dagger}$. (c) $S_{\mathbf{DW}}^{\mathcal{U}_{\mathrm{eg}}}$ for $\mathcal{U}_{\mathrm{eg}}[\bullet] = U_{\mathrm{eg}} \bullet U_{\mathrm{eg}}^{\dagger}$. (d) $S_{\mathbf{DW}}^{\hat{\mathcal{U}}_{\mathrm{eg}}}$ for $\mathcal{U}_{\mathrm{eg}}[\bullet] = U_{\mathrm{eg}} \bullet U_{\mathrm{eg}}^{\dagger}$. (e) $S_{\mathbf{SP}}^{\mathcal{U}_{\mathrm{eg}}}$ for $\mathcal{U}_{\mathrm{eg}}[\bullet] = U_{\mathrm{eg}} \bullet U_{\mathrm{eg}}^{\dagger}$. (f) $S_{\mathbf{SP}}^{\hat{\mathcal{U}}_{\mathrm{eg}}}$ for $\mathcal{U}_{\mathrm{eg}}[\bullet] = U_{\mathrm{eg}} \bullet U_{\mathrm{eg}}^{\dagger}$.

As depicted in Fig. 2, we have the forward and retrodictive transition graphs for a channel given by $\mathcal{E}[\bullet] = \mathrm{Tr}_B[U_{\curlywedge} \bullet \otimes |1\rangle\langle 1| \; U_{\curlywedge}^\dagger]$. To understand the retrodictive action given by the Petz, we can gain some intuitions by writing out these mappings:

$$|01\rangle \overset{U_{\curlywedge}}{\to} \frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big) \overset{\mathrm{Tr}_B}{\to} \boxed{\frac{1}{2}\mathbb{1}},$$

$$|{+}1\rangle \overset{U_{\curlywedge}}{\to} \frac{1}{2}|11\rangle + \frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big),$$

$$\overset{\mathrm{Tr}_B}{\to} \boxed{\frac{1}{4}|0\rangle\langle +| \frac{3}{4}|1\rangle\langle +| \frac{1}{2\sqrt{2}}\big(|1\rangle\langle 0| + |0\rangle\langle 1|\big)},$$

$$|11\rangle \overset{U_{\curlywedge}}{\to} |11\rangle \overset{\mathrm{Tr}_B}{\to} \boxed{|1\rangle\langle 1|}.$$

We see that if the reference state is $\gamma = |0\rangle\langle 0|$ or $|+\rangle\langle +|$, then any state is compatible to its output (they are unambiguously full rank in $\mathbb{C}^2$). Hence, the Petz map erases all (output) states back to the reference, in full consistency with the earlier comments about the quantum total-erasure channel. This is depicted in Figs. 2(c) and 2(d).

A very different situation occurs for $\gamma = |1\rangle\langle 1|$. In this case, only $|1\rangle\langle 1|$ is allowed as an output. Thus, the Petz sends $|1\rangle\langle 1|$ to itself, while all other states are retrodicted in (complicated but logically consistent) ways dependent on that forward transitions of the channel, reflected in Fig. 2(e).

To explain this more symmetrically: in the former two scenarios, all outputs are compatible with the absolute conviction (as enforced by state purity) given to the reference state; hence all outputs are retrodicted to it. Meanwhile, in this latter case, *only one pure output* (which just so happens to be the same as the reference) is compatible with the pure reference state. Hence, all *other states* (beside the expected output) are retrodicted in accordance with the channel without any regard the reference, since the reference already excludes the possibility of such states. These more complicated Bayesian inversions come together and cumulate into a vertical reflection of the forward channel, as depicted in Fig. 2(e). For an arbitrary γ, we obtain a classical mixture of all these key effects together. We depict the case where $\gamma = \gamma(\frac{\pi}{16}, \frac{\pi}{5}, \frac{\pi}{3})$ in Fig. 2(f).

It should be said the interplay of reference and channel dependencies that we review here is fundamental in classical retrodiction scenarios as well. The half-SWAP illustrates that these same Bayesian features hold in the quantum regime via the inferential structure of the Petz map, even when complementarity and entanglement is introduced.

VII. CONCLUSIONS

By expressing the Petz recovery map as a decomposition of matrices given by Eq. (22), we situate quantum Bayesian inference in the same formal language as that of its classical counterpart given by Eq. (27). We also highlight what we find to be the most noteworthy (and interpretation neutral) similarities and differences between these two theories when it comes to logical inference.

Bayesian inference in both theories involves a similar structure (see $S_T^{\tilde{\mathcal{E}}_\gamma}$ in Table II). Given that the transpose of a classical channel is also its adjoint, the key difference between $S_{\mathbf{CL}}^{\tilde{\mathcal{E}}_\gamma}$ and $S_{\mathbf{QM}}^{\hat{\mathcal{E}}_\gamma}$ lies not in the central matrix $S^{\mathcal{E}^\dagger}$ but in the right and left matrices $X_{\mathcal{E}[\gamma]}$ and X_γ that capture the description of the priors. This affirms the fact that what separates quantum theory from classical theory is not so much in its dynamics (which is in many ways conceptually similar) but in the description of states.

Mathematically, the difference is captured by the form of the structure coefficients ξ_{ixjy}. In classical theory, the structure coefficients render the matrices diagonal. By contrast, in a QPR of quantum theory, the structure coefficients introduce weighted products $v_x^\alpha v_y^\alpha$ of every pair of entries of the distribution v^α. This is a consequence of the fact that frames are tomographically complete—ultimately, a signature of complementarity.

After these key results, we illustrate some examples of quantum Bayesian reasoning using transition graphs, which offer some visual intuitions about how the Petz map produces inferences in a quantum regime. Finally, we point to two possible directions to enlarge the perspective open by this work. One may repeat the current study for alternative proposals of quantum Bayes' rule that are not equivalent to the Petz map [5]. Also, having chosen (quasi)stochastic processes as the common language for the comparison, we leave out representations of quantum theory that have a different structure and use complex numbers: notably, the Kirkwood-Dirac representation, which has recently been shown to be related to metrological advantages [42].

APPENDIX A: $S^{\mathcal{E}}$ IN TERMS OF ξ_{pqrs}

We recall in reference to Table I that

$$S_{ij}^{\mathcal{E}} = \mathrm{Tr}\big[F_i \mathcal{E}[G_j]\big]. \tag{A1}$$

Braasch and Wootters have observed [see Ref. [38, Eq. (19)]], in the context of DW representation, that one

can write an output state $\rho' = \mathcal{E}[\rho]$ of a channel $\mathcal{E}[\bullet] = \sum_l \kappa_l \bullet \kappa_l^\dagger$ in the following QPR expression:

$$v_i^{\rho'} = \frac{1}{d} \sum_{xyj} \text{Tr}[G_i G_x G_j G_y] \mathcal{K}_{xy} v_j^\rho, \qquad (A2)$$

where $\mathcal{K}_{xy} = \sum_l K_x^{(l)} K_y^{(l)*}$ and $K_x^{(l)} = \frac{1}{d}\text{Tr}[G_x \kappa_l]$. Hence,

$$v_i^{\rho'} = \frac{1}{d^3} \sum_{xylj} \text{Tr}[G_x \kappa_l] \text{Tr}[G_y \kappa_l^\dagger] \text{Tr}[G_i G_x G_j G_y] v_j^\rho. \qquad (A3)$$

Since $v^{\rho'} = S^{\mathcal{E}} v^\rho$, this implies that

$$(S_{\text{DW}}^{\mathcal{E}})_{ij} = \frac{1}{d^3} \sum_{xyl} \text{Tr}[G_x \kappa_l] \text{Tr}[G_y \kappa_l^\dagger] \text{Tr}[G_i G_x G_j G_y]. \qquad (A4)$$

We build upon this observation by deriving directly from Eq. (A1) an expression akin to Eq. (A4) but valid for any QPR:

$$S_{ij}^{\mathcal{E}} = \text{Tr}[F_i \mathcal{E}[G_j]] \qquad (A5)$$

$$= \sum_l \text{Tr}[F_i \kappa_l G_j \kappa_l^\dagger] \qquad (A6)$$

$$= \sum_{xl} \text{Tr}[F_x \kappa_l] \text{Tr}[G_x G_j \kappa_l^\dagger F_i] \qquad (A7)$$

$$= \sum_{xyl} \text{Tr}[F_x \kappa_l] \text{Tr}[F_y \kappa_l^\dagger] \text{Tr}[F_i G_x G_j G_y] \qquad (A8)$$

$$= \sum_{xyl} \text{Tr}[F_x \kappa_l] \text{Tr}[F_y \kappa_l^\dagger] \xi_{ixjy}. \qquad (A9)$$

Recalling Eq. (7), we recover Eq. (A4) for DW representations. In Eq. (A9), we have an expression much like Eq. (19), now catering for general CP maps. We may compare these expressions more instructively, by writing Eq. (19) as

$$S_{ij}^{\mathcal{M}_\alpha} = (M_\alpha)_{ij} = \sum_{xy} \text{Tr}[F_x \alpha] \text{Tr}[F_y \alpha] \xi_{ixjy}. \qquad (A10)$$

This comparison gives us a sense of what the structure coefficients embed into objects in which they reside. They structurally encode the choice of frame and representation into the entries of these maps, whether they are CPTP or not.

APPENDIX B: $M_{\alpha^r} = M_\alpha^r$ FOR ALL $r \in \mathbb{Q}$

From the concatenation of maps $S^{\mathcal{E}} S^{\mathcal{F}} = S^{\mathcal{E} \circ \mathcal{F}}$ [Eq. (14)] and the definition Eq. (16), it is immediate that $M_{\alpha^r} =$

M_α^r holds for $r \in \mathbb{Z}^+$. We now proceed to prove that it is valid for $r \in \mathbb{Q}$.

We recall the entry-wise definition in Eq. (17):

$$(M_{\alpha^r})_{ij} = \text{Tr}[F_i \alpha^r G_j \alpha^r]. \qquad (B1)$$

First, we prove that

$$(M_{\alpha^{1/2}} M_{\alpha^{1/2}})_{ij} = \sum_k (M_{\alpha^{1/2}})_{ik} (M_{\alpha^{1/2}})_{kj}$$

$$= \sum_k \text{Tr}[F_i \sqrt{\alpha} G_k \sqrt{\alpha}] \text{Tr}[F_k \sqrt{\alpha} G_j \sqrt{\alpha}]$$

$$= \sum_k \text{Tr}[G_k \sqrt{\alpha} F_i \sqrt{\alpha}] \text{Tr}[F_k \sqrt{\alpha} G_j \sqrt{\alpha}]$$

$$= \text{Tr}[\sqrt{\alpha} F_i \sqrt{\alpha} \sqrt{\alpha} G_j \sqrt{\alpha}]$$

$$= \text{Tr}[F_i \alpha G_j \alpha] = (M_\alpha)_{ij}.$$

Crucially, in the fourth equality we use the property given in Eq. (6) in QPR. Hence,

$$M_{\alpha^{1/2}} = M_\alpha^{1/2}.$$

By reiterating this (i.e., sending $\alpha \to \sqrt{\alpha}$), we obtain

$$\forall n \in \mathbb{Z}^+ : \ M_{\alpha^{1/2n}} = M_\alpha^{\frac{1}{2n}},$$

Using Eq. (14) for $N \in \mathbb{Z}^+$, we have

$$M_{\alpha^{N/2n}} = S^{\mathcal{M}_{\alpha^{N/2n}}}$$

$$= S^{\overbrace{\mathcal{M}_{\alpha^{1/2n}} \circ \mathcal{M}_{\alpha^{1/2n}} \circ \cdots \circ \mathcal{M}_{\alpha^{1/2n}}}^{N \text{ times}}}$$

$$= \prod_N M_{\alpha^{1/2n}} = \prod_N M_\alpha^{1/2n}.$$

Since any positive rational number can be written as $q = N/2n$, we prove that

$$M_{\alpha^q} = M_\alpha^q, \quad q \in \mathbb{Q}^+. \qquad (B2)$$

Second, we note that

$$(M_\alpha M_{\alpha^{-1}})_{ij} = \sum_k \text{Tr}[F_i \alpha G_k \alpha] \text{Tr}[F_k \alpha^{-1} G_j \alpha^{-1}]$$

$$= \sum_k \text{Tr}[G_k \alpha F_i \alpha] \text{Tr}[F_k \alpha^{-1} G_j \alpha^{-1}]$$

$$= \text{Tr}[\alpha F_i \alpha \alpha^{-1} G_j \alpha^{-1}]$$

$$= \text{Tr}[F_i G_j] = \delta_{ij} = \mathbb{1}_{ij}.$$

Hence, $M_{\alpha^{-1}} = M_\alpha^{-1}$. Repeating this, we can easily see that for any $N \in \mathbb{Z}^+$:

$$M_{\alpha^{-N}} = M_\alpha^{-N}. \qquad (B3)$$

Finally, taking from Eqs. (14), (B2) and (B3), we find that

$$
\begin{aligned}
M_{\alpha^{q-N}} = S^{\mathcal{M}_{\alpha^{q-N}}} &= S^{\mathcal{M}_{\alpha^q} \circ \mathcal{M}_{\alpha^{-N}}} \\
&= M_{\alpha^q} M_{\alpha^{-N}} = M_\alpha^q M_\alpha^{-N} \\
&= M_\alpha^{q-N}.
\end{aligned}
$$

Since any rational number can be written as a positive rational number minus a positive integer, we prove that $M_{\alpha^r} = M_\alpha^r$ for any $r \in \mathbb{Q}$.

APPENDIX C: $S^{\mathcal{E}^\dagger}$ IN TERMS OF $(S^\mathcal{E})^T$ FOR QPRs

We derive the QPR expressions for $S^{\mathcal{E}^\dagger}$ for some CP map $\mathcal{E}[\bullet] = \sum_l \kappa_l \bullet \kappa_l^\dagger$. For NQPRs, we find easily that

$$
\begin{aligned}
(S_{\mathbf{NQ}}^{\mathcal{E}^\dagger})_{ij} = \mathrm{Tr}\big[F_i \mathcal{E}^\dagger[G_j]\big] &= \mathrm{Tr}\bigg[F_i \sum_l \kappa_l^\dagger G_j \kappa_l\bigg] \\
&= \sum_l \mathrm{Tr}\big[G_j \kappa_l F_i \kappa_l^\dagger\big] = \sum_l \mathrm{Tr}\big[F_j \kappa_l G_i \kappa_l^\dagger\big] \\
&= \mathrm{Tr}\bigg[F_j \sum_l \kappa_l G_i \kappa_l^\dagger\bigg] = \mathrm{Tr}\big[F_j \mathcal{E}[G_i]\big] = S_{ji}^\mathcal{E}.
\end{aligned}
$$

Thus, for NQPRs $S_{\mathbf{NQ}}^{\mathcal{E}^\dagger} = (S^\mathcal{E})^T$. For SIC-POVM representations, we have a more complicated expression. We first use Eq. (11):

$$
\begin{aligned}
(S_{\mathbf{SP}}^{\mathcal{E}^\dagger})_{ij} &= \mathrm{Tr}\big[F_i \mathcal{E}^\dagger[G_j]\big] \\
&= \mathrm{Tr}\bigg[\frac{G_i + \mathbb{1}}{d(d+1)} \mathcal{E}^\dagger\big[d(d+1)F_j - \mathbb{1}\big]\bigg].
\end{aligned}
$$

By expanding the terms and noting the unitality of every adjoint map (i.e., $\mathcal{E}^\dagger[\mathbb{1}] = \mathbb{1}$), we arrive at the expression

$$
\begin{aligned}
(S_{\mathbf{SP}}^{\mathcal{E}^\dagger})_{ij} &= \mathrm{Tr}\bigg[F_i \sum_l \kappa_l^\dagger G_j \kappa_l\bigg] + \mathrm{Tr}\big[\mathcal{E}^\dagger[F_j]\big] - \mathrm{Tr}[F_i] \\
&= S_{ji}^\mathcal{E} + \mathrm{Tr}\big[\mathcal{E}^\dagger[F_j]\big] - \frac{1}{d}.
\end{aligned}
$$

By taking note of the relations found in Table I, we may write $\mathrm{Tr}\big[\mathcal{E}^\dagger[F_j]\big] = \sum_l \mathrm{Tr}\big[\kappa_l^\dagger F_j \kappa_l\big] = \sum_l \mathrm{Tr}\big[F_j \kappa_l \mathbb{1} \kappa_l^\dagger\big] = \mathrm{Tr}\big[F_j \mathcal{E}[\mathbb{1}]\big] = \frac{1}{d}(S^\mathcal{E} v^1)_j = \frac{1}{d} \sum_a \mathcal{E}(j|a)$. Hence, we can write the total expression of each entry for SIC-POVM representation as

$$
(S_{\mathbf{SP}}^{\mathcal{E}^\dagger})_{ij} = S_{ji}^\mathcal{E} + \frac{1}{d}\bigg(\sum_a \mathcal{E}(j|a) - 1\bigg). \tag{C1}
$$

This can be written, on the matrix level, as Eq. (24).

APPENDIX D: $(S^\mathcal{E})^T$ AS $S^{\mathcal{E}^\dagger}$ FOR CLASSICAL

In the previous appendix, we prove that for quantum channels (expressed in QPRs), we can express the adjoint channel in terms of the transpose of the channel. Here, we prove the opposite relation for classical channels: that the transpose of a classical channel is the adjoint of that channel. Namely, the transpose map is the map for which Eq. (5) is fulfilled in the case of classical scenarios. Noting first the commutative diagram found in Fig. 3 (which invokes the relationships found in Table I), we see how Eq. (5) is fulfilled by a map for which

$$
(S^\mathcal{E} v^\rho) \cdot \bar{v}^\sigma = (S^{\mathcal{E}^\dagger} v^\sigma) \cdot \bar{v}^\rho, \tag{D1}
$$

for all ρ and σ. With this, we expand the left-hand side of Eq. (D1):

$$
\begin{aligned}
(S^\mathcal{E} v^\rho) \cdot \bar{v}^\sigma &= \sum_y (S^\mathcal{E} v^\rho)_y \bar{v}_y^\sigma \tag{D2} \\
&= \sum_{xy} S_{yx}^\mathcal{E} v_x^\rho \bar{v}_y^\sigma. \tag{D3}
\end{aligned}
$$

Next we expand the following, in order to check if the transpose qualifies as the adjoint:

$$
\begin{aligned}
((S^\mathcal{E})^T v^\sigma) \cdot \bar{v}^\rho &= \sum_{xy} (S^\mathcal{E})_{yx}^T v_x^\sigma \bar{v}_y^\rho \tag{D4} \\
&= \sum_{xy} S_{xy}^\mathcal{E} v_x^\sigma \bar{v}_y^\rho \tag{D5} \\
&= \sum_{xy} S_{yx}^\mathcal{E} \bar{v}_x^\rho v_y^\sigma. \tag{D6}
\end{aligned}
$$

Now, for classical scenarios, the trace of two states, if treated like quantum states in Hilbert space, would simply be the inner product of its density spectra: $v^\rho \cdot v^\sigma = \mathrm{Tr}[\rho\,\sigma]$. This is because the states, being classical distributions, would be diagonalized in the same way. Thus we could have replaced $\bar{v}^\rho$ with v^ρ in all the above calculations and in Fig. 3. The reason why we write $\bar{v}^\rho$ as opposed to v^ρ is to simply highlight that while, indeed, Eq. (D3) is identical to Eq. (D6) for classical scenarios, because $\bar{v}^\rho = v^\rho$ there (and NQPR for that matter, since $\bar{v}^\rho = c\, v^\rho$), the same does *not* hold for SIC-POVM. The transpose qualifies as an adjoint for both NQPR and classical channels but not for SIC-POVM. Hence, the relation proved for classical states and channels does not contradict the ones proved in the previous appendix for QPRs.

APPENDIX E: PROPERTIES OF M_α

Here, we note some interesting properties of M_α, namely, that it is a matrix with all real entries and nonnegative eigenvalues that sum to 1.
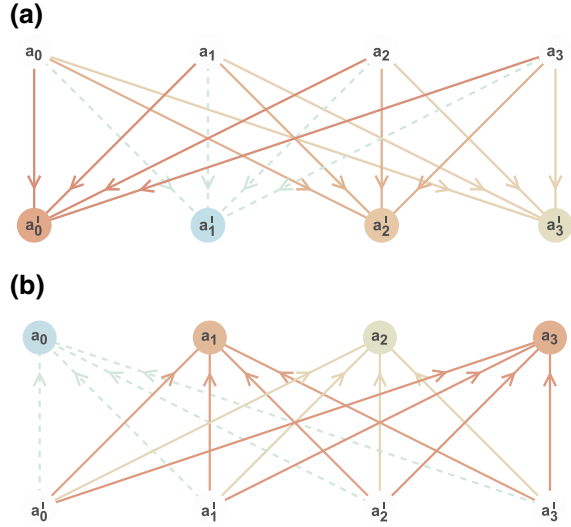
(a)

(b)

FIG. 6. Transition graphs for a quantum total-erasure channel with arbitrary ancilla β and a corresponding retrodiction with reference prior γ. (a) $S^{\mathcal{E}}_{\textbf{DW}}$ for $\mathcal{E}[\bullet] = \text{Tr}_B[U_\leftrightarrow \bullet \otimes \beta(\frac{7\pi}{16}, \frac{3\pi}{5}, \frac{\pi}{6}) U_\leftrightarrow{}^\dagger]$ (b) $SS^{\hat{\mathcal{E}}_\gamma}_{\textbf{DW}}$ for $U_\leftrightarrow, \beta(\frac{7\pi}{16}, \frac{3\pi}{5}, \frac{\pi}{6}), \gamma(\frac{\pi}{16}, \frac{\pi}{5}, \frac{\pi}{8})$.

1. Real entries

It can be shown that all the entries of M_α are real: $(M_\alpha)_{ij} = \text{Tr}[F_i \alpha G_j \alpha] \in \mathbb{R}$. A proof is given in the main text, valid for any M_{α^r}; we repeat it here for completeness.

We first note that the anticommutator for any two Hermitian operators A and B is always also Hermitian—$\{A, B\}^\dagger = \{A, B\}$—while the trace of the commutator of any two operators is always zero (in finite dimension) due to cyclicity: $\text{Tr}[[A, B]] = \text{Tr}[AB] - \text{Tr}[BA] = 0$. Hence

$$\text{Tr}[AB] = \text{Tr}\left[\frac{\{A, B\}}{2} + \frac{[A, B]}{2}\right] = \frac{1}{2}\text{Tr}[\{A, B\}] \in \mathbb{R}. \tag{E1}$$

Noting that $\sqrt{\alpha}F_i\sqrt{\alpha}$ and $\sqrt{\alpha}G_j\sqrt{\alpha}$ are both Hermitian (frame and dual operators are always Hermitian and α is a density operator in Hilbert space), we apply Eq. (E1) to $(M_\alpha)_{ij}$. The entries of M_α are thus proven to be always real.

2. Positive semidefiniteness

For any NQPR, we can always write

$$(M_\alpha)_{ij} = c\,\text{Tr}\left[\sqrt{\alpha}F_i\sqrt{\alpha}\left(\sqrt{\alpha}F_j\sqrt{\alpha}\right)^\dagger\right].$$

Hence, M_α is a Gram matrix with some positive factor c. Thus it is positive semidefinite. For SIC-POVM, we

expand $(M_\alpha)_{ij}$ via Eq. (11), arriving at

$$(M_\alpha)_{ij} = \frac{1}{d(d+1)}\left(\text{Tr}\left[\sqrt{\alpha}G_i\sqrt{\alpha}\left(\sqrt{\alpha}G_j\sqrt{\alpha}\right)^\dagger\right] + \text{Tr}[G_j\alpha^2]\right).$$

The first term, as with the NQPR case, corresponds to a Gram matrix, which is positive semidefinite. One can then note that the second term corresponds to a matrix J_α (i.e., $(J_\alpha)_{ij} = \text{Tr}[G_j\alpha^2]$) with duplicate rows (every jth column filled with identical entries. This simply implies that the only nonzero eigenvalue would be the sum of the entries of any given row, which just means that $\text{eig}[J_\alpha] = \{\sum_j \text{Tr}[G_j\alpha^2], 0\} = \{\text{Tr}[\alpha^2], 0\} \geq 0$. So M_α is the sum of two positive semidefinitive matrices and thus we may conclude $M_\alpha \geq 0$ for SIC-POVM as well.

3. Unit trace

The trace of M_α is given by

$$\text{Tr}[M_\alpha] = \sum_i (M_\alpha)_{ii} = \text{Tr}\left[\underbrace{\sum_i F_i\alpha G_i}_{\mathbb{1}}\,\alpha\right] = 1.$$

To prove the relation invoked for the final equality, we use the previously found result in Ref. [43]. Consider the superoperator

$$\Lambda[\bullet] = \sum_{i=1}^{d^2} \Pi_i \bullet \Pi_i. \tag{E2}$$

it can be shown that

$$\Lambda[\Pi_i] = \frac{d}{d+1}(\Pi_i + \mathbb{1}). \tag{E3}$$

Since the set $\{\Pi_i\}$ forms a basis, we can express the superoperator as

$$\Lambda = \frac{d}{d+1}(\mathcal{I} + \mathbb{1}), \tag{E4}$$

where $\mathcal{I}[A] = \text{Tr}[A]\,\mathbb{1}$. Using this, we can easily show that for SIC-POVM representation, we have

$$\sum_i F_i\alpha G_i = \frac{d+1}{d}\sum_i \Pi_i\alpha\Pi_i - \frac{1}{d}\alpha\sum_i \Pi_i$$

$$= \mathbb{1}. \tag{E5}$$

For DW representation, Zhu [29] has shown that the dual frame can always be expressed as follows:

$$G_i = -\sqrt{d+1}\,\Pi_i + \left(\frac{1+\sqrt{d+1}}{d}\right)\mathbb{1}. \quad (E6)$$

Thus, it can also be easily shown that $\sum_i F_i \alpha G_i = \mathbb{1}$ in this representation.

APPENDIX F: GENERAL EXAMPLES FOR $S_{QM}^{\hat{\mathcal{E}}_\gamma} \neq S_{CL}^{\tilde{\mathcal{E}}_\gamma}$

As discussed in Sec. VI C, it is the case that $\hat{\mathcal{E}}_\gamma[\rho] = \hat{\mathcal{E}}_+[\rho] = |+\rangle\langle+|$ for all ρ when $\mathcal{E}[\bullet] = \mathrm{Tr}_B[U_\lambda \bullet \otimes |1\rangle\langle 1| \, U_\lambda^\dagger]$ and $\gamma = |+\rangle\langle+|$. Yet we can easily find that, for the canonical state representations for DW and SIC-POVM, we have

$$S_{DW}^{\tilde{\mathcal{E}}_+} = \begin{pmatrix} 1 & \frac{1}{7}\left(3-\sqrt{2}\right) & 1 & \frac{1}{7}\left(\sqrt{2}+3\right) \\ 0 & \frac{1}{7}\left(\sqrt{2}+4\right) & 0 & \frac{1}{7}\left(4-\sqrt{2}\right) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\neq \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = S_{DW}^{\hat{\mathcal{E}}_+}.$$

Likewise,

$$S_{SP}^{\tilde{\mathcal{E}}_+} = \begin{pmatrix} 0.925 & 0.183 & -0.264 & 0.353 \\ 0.0744 & 0.744 & 0.275 & 0.168 \\ -0.0191 & 0.0491 & 0.915 & 0.0947 \\ 0.0199 & 0.0233 & 0.0737 & 0.384 \end{pmatrix}$$

$$\neq \frac{1}{12}\begin{pmatrix} \sqrt{3}+3 & \sqrt{3}+3 & \sqrt{3}+3 & \sqrt{3}+3 \\ \sqrt{3}+3 & \sqrt{3}+3 & \sqrt{3}+3 & \sqrt{3}+3 \\ 3-\sqrt{3} & 3-\sqrt{3} & 3-\sqrt{3} & 3-\sqrt{3} \\ 3-\sqrt{3} & 3-\sqrt{3} & 3-\sqrt{3} & 3-\sqrt{3} \end{pmatrix}$$

$$= S_{SP}^{\hat{\mathcal{E}}_+}.$$

Indeed, for some channels, one can find states for which the postmeasurement probabilities violate acceptable bounds. This means that $S_{CL}^{\tilde{\mathcal{E}}_\gamma}$ fails to represent a generally valid quantum transformation. For instance, for a unitary transformation $\mathcal{U}[\bullet] = U \bullet U^\dagger$, where $U = \frac{i}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$ We find that

$$(S_{DW}^{\tilde{\mathcal{U}}_+} v^+) \cdot \bar{v}^0 = \frac{1}{2}(1+\sqrt{3}) > 1$$

$$(S_{SP}^{\tilde{\mathcal{U}}_+} v^0) \cdot \bar{v}^+ = \frac{1}{13}(2-5\sqrt{3}) < 0.$$

$S^{\hat{\mathcal{E}}_\gamma} \neq S_{CL}^{\tilde{\mathcal{E}}_\gamma}$ is thus easily shown.

APPENDIX G: $S_{QM}^{\hat{\mathcal{E}}_\gamma} \neq S_{CL}^{\tilde{\mathcal{E}}_\gamma}$ EVEN WHEN $\mathcal{E}[\gamma], \gamma, \rho, E_m$ ALL COMMUTE

Some may expect that quantum retrodiction, under QPR formalism and using its mathematical equipment, would go to classical retrodiction once all states and transformations share the same eigenbasis. Put differently, there may be an expectation that, for every choice of QPR and if $\mathcal{E}[\gamma]$, γ, ρ, and E_m commute, $(S_{QM}^{\hat{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m = (S_{CL}^{\tilde{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m$.

However, it turns out that this is not the case. We simply list two examples where this does not obtain. For the canonical choice of DW representation,

$$U = \frac{1}{4}\begin{pmatrix} \sqrt{2}+2 & 1-i & -1+i & i\left(\sqrt{2}-2\right) \\ -1-i & \sqrt{2}+2 & 2-\sqrt{2} & 1-i \\ 1+i & 2-\sqrt{2} & \sqrt{2}+2 & -1+i \\ -i\left(\sqrt{2}-2\right) & -1-i & 1+i & \sqrt{2}+2 \end{pmatrix},$$

$$\beta = \frac{1}{2}\mathbb{1}, \ \gamma = \rho = E_m = |-\alpha\rangle\langle-\alpha|,$$

where U and β define the unitary dilation for $\mathcal{E}$, $|+\alpha\rangle = \cos(\pi/4)|0\rangle + e^{i\pi/8}\sin(\pi/4)|1\rangle$, $|-\alpha\rangle = e^{-i\pi/8}\sin(\pi/4)|0\rangle - \cos(\pi/4)|1\rangle$. Essentially, this gives a simple noise-inducing channel (for qubits) that preserves the coherence of states in the $|+\alpha\rangle$ and $|-\alpha\rangle$ basis. It can be found that

$$(S_{QM}^{\hat{\mathcal{E}}_\gamma} v^\rho)^T = \frac{1}{4}(1-\sqrt{2},\ 1,\ 1+\sqrt{2},\ 1),$$

$$(S_{CL}^{\tilde{\mathcal{E}}_\gamma} v^\rho)^T = \frac{1}{16}(8-5\sqrt{2},\ 2+\sqrt{2},\ 4 + 3\sqrt{2},\ 2+\sqrt{2}),$$

$$(S_{QM}^{\hat{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m = \frac{1}{2},\ \ (S_{CL}^{\tilde{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m = \frac{1}{16}(8-\sqrt{2}).$$

Similarly, for the canonical choice of frames under SIC-POVM,

$$U = U_\lambda, \ \gamma = \rho = \frac{1}{2}\mathbb{1},$$

$$\beta = |0\rangle\langle 0|, \ E_m = |1\rangle\langle 1|.$$

Here, we have an amplitude-damping channel for the $|0\rangle$, $|1\rangle$ basis. It can be found that

$$(S_{\mathbf{QM}}^{\hat{\mathcal{E}}_\gamma} v^\rho)^T = \frac{1}{36}(9 - \sqrt{3},\ 9 + \sqrt{3},\ 9 - \sqrt{3},\ 9 + \sqrt{3}),$$

$$(S_{\mathbf{CL}}^{\tilde{\mathcal{E}}_\gamma} v^\rho)^T = \frac{1}{44}(11 - \sqrt{3},\ 11 + \sqrt{3},\ 11 - \sqrt{3},\ 11 + \sqrt{3}),$$

$$(S_{\mathbf{QM}}^{\hat{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m = \frac{7}{11},\quad (S_{\mathbf{CL}}^{\tilde{\mathcal{E}}_\gamma} v^\rho) \cdot \bar{v}^m = \frac{2}{3},$$

so even when all relevant states and POVMs are diagonal in Hilbert space, the QPR formalism does not make it such that quantum Bayesian inference goes to the classical Bayes' rule. This highlights the categorical difference between quantum and classical theory even when they are using similar mathematical equipment. Once one employs tomographically complete frames in one's QPR, one may not simply reinvoke classical Bayes' rule hoping that it will simulate the quantum retrodiction, even when all quantum states and transformations commute.

Note that this does not contradict the known result, mentioned in Sec. I, that the Petz map given in Eq. (2) reduces to classical Bayes' rule when all relevant quantum objects are diagonal in the same basis. This is because that result holds in the context of the d-dimensional Hilbert-space formalism. Meanwhile, the above conclusion is made under a QPR, when a valid frame has been affixed to define $d \times d$ real space.

APPENDIX H: OTHER TRANSITION GRAPHS

In this appendix, we include illustrative cases of $S^{\mathcal{E}}$, some respective retrodictions, and their transition graphs. In Fig. 4, the transition graphs are depicted for very familiar Pauli rotations. It so happens that these unitaries translate to $S^{\mathcal{E}}$ that give permutations. This is seen in the bold bijective-transition arrows. Like other unitary channels, all retrodictions are reference-prior independent. Transition graphs of such retrodictions are thus always mirror images of the corresponding forward transition graph. That said, most unitaries do not enjoy a permutative structure that exists for these SU(2) rotations. The Hadamard gate, for instance, defined by the following computationally represented operator and gives the respective quasistochastic matrix:

$$U_H \,\hat{=}\, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S^{\mathcal{U}_H} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix},$$

which is consistent across the canonical choices of the DW and SIC-POVM representations.

Likewise, for an arbitrarily chosen unitary U_{eg},

$$U_{\mathrm{eg}} \,\hat{=}\, \frac{1}{4} \begin{pmatrix} i\left(\sqrt{3}+2i\right) & 0 & 3i & 0 \\ 0 & i\left(\sqrt{3}+2i\right) & 0 & 3i \\ -3i & 0 & 2+i\sqrt{3} & 0 \\ 0 & -3i & 0 & 2+i\sqrt{3} \end{pmatrix},$$

one has the following quasiprobability objects:

$$S_{\mathbf{DW}}^{\mathcal{U}_{\mathrm{eg}}} = \frac{1}{16} \begin{pmatrix} 9 & \sqrt{3}-6 & 4-3\sqrt{3} & 2\sqrt{3}+9 \\ -\sqrt{3}-6 & 9 & 9-2\sqrt{3} & 3\sqrt{3}+4 \\ 3\sqrt{3}+4 & 2\sqrt{3}+9 & -3 & 6-5\sqrt{3} \\ 9-2\sqrt{3} & 4-3\sqrt{3} & 5\sqrt{3}+6 & -3 \end{pmatrix},$$

$$S_{\mathbf{SP}}^{\mathcal{U}_{\mathrm{eg}}} = \frac{1}{16} \begin{pmatrix} -3 & 5\sqrt{3}+6 & 4-3\sqrt{3} & 9-2\sqrt{3} \\ 6-5\sqrt{3} & -3 & 2\sqrt{3}+9 & 3\sqrt{3}+4 \\ 3\sqrt{3}+4 & 9-2\sqrt{3} & 9 & -\sqrt{3}-6 \\ 2\sqrt{3}+9 & 4-3\sqrt{3} & \sqrt{3}-6 & 9 \end{pmatrix}.$$

It is clear that these forward channels do not give permutative QPRs. Nevertheless, the property that $S^{\hat{\mathcal{U}}_\gamma} = S^{\hat{\mathcal{U}}} = (S^{\mathcal{U}})^T$ is still reflected clearly in Fig. 5. In contrast to these reversible maps, we can speak of the quantum total-erasure channel mentioned in Sec. VI B. The full swap is expressed as follows:

$$U_{\leftrightarrow} \; \hat{=} \; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{H1}$$

As clearly depicted in Fig. 6, both the forward channel and its retrodiction are erased perfectly to the relevant state (the ancilla for the forward map and the reference state for the retrodiction).

[1] S. Watanabe, Conditional probabilities in physics, Progr. Theor. Phys. Suppl. **E65**, 135 (1965).

[2] S. Watanabe, Symmetry of physical laws. Part III. Prediction and retrodiction, Rev. Mod. Phys. **27**, 179 (1955).

[3] R. Jeffrey, *The Logic of Decision* (McGraw-Hill, New York, 1965).

[4] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, Cambridge, 2003).

[5] A. J. Parzygnat and J. Fullwood, From time-reversal symmetry to quantum Bayes' rules (2022), ArXiv:2212.08088.

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

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

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

[9] M. M. Wilde, From classical to quantum Shannon theory (2011), arXiv preprint ArXiv:1106.1445.

[10] K. Li and A. Winter, Squashed entanglement, **k**-extendibility, quantum Markov chains, and recovery maps, Found. Phys. **48**, 910 (2018).

[11] 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).

[12] A. J. Parzygnat and F. Buscemi, Axioms for retrodiction: Achieving time-reversal symmetry with a prior (2022), arXiv preprint ArXiv:2210.13531.

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

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

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

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

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

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

[19] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, Negative quasi-probability as a resource for quantum computation, New J. Phys. **14**, 113011 (2012).

[20] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Contextuality supplies the "magic" for quantum computation, Nature **510**, 351 (2014).

[21] H. Pashayan, J. J. Wallman, and S. D. Bartlett, Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities, Phys. Rev. Lett. **115**, 070501 (2015).

[22] Of course, one can have it that a' and a are defined in different state spaces A and A' but we can always take $A'' = A \cup A'$ and characterize the channel in this larger alphabet.

[23] This constraint also exists in the classical Bayes update and is likewise of no practical concern, as one can always ensure that that γ is full rank by adding some arbitrarily small weights into its spectrum and adding some arbitrarily small mapping probability in $\mathcal{E}$ as well. These contributions can then be sent to zero on the recovered state.

[24] A. J. Parzygnat and F. Buscemi, Axioms for retrodiction: Achieving time-reversal symmetry with a prior (2022), ArXiv:2210.13531.

[25] S. Mac Lane, *Categories for the Working Mathematician* Vol. 5 (Springer Science & Business Media, New York, 1998), 13.

[26] R. Carnap, *The Logical Syntax of Language* (Open Court Publishing, Chicago and La Salle, Illinois, 2002).

[27] It is the case, for SIC-POVM and DW representation, that $\text{Tr}[F_j] = 1/d$ for all $j \in \Lambda$. But this is not generally the case for all valid QPRs.

[28] M. Ruzzi and D. Galetti, Quantum discrete phase space dynamics and its continuous limit, J. Phys. A: Math. Gen. **33**, 1065 (2000).

[29] H. Zhu, Quasiprobability Representations of Quantum Mechanics with Minimal Negativity, Phys. Rev. Lett. **117**, 120404 (2016).

[30] W. K. Wootters, A Wigner-function formulation of finite-state quantum mechanics, Ann. Phys. (N. Y) **176**, 1 (1987).

[31] A. B. Klimov and C. Muñoz, Discrete Wigner function dynamics, J. Opt. B: Quantum Semiclass. Opt. **7**, S588 (2005).

[32] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, Discrete phase space based on finite fields, Phys. Rev. A **70**, 062101 (2004).

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

[34] M. Appleby, C. A. Fuchs, B. C. Stacey, and H. Zhu, Introducing the Qplex: A novel arena for quantum theory, Eur. Phys. J. D **71**, 1 (2017).

[35] E. O. Kiktenko, A. O. Malyshev, A. S. Mastiukova, V. I. Man'ko, A. K. Fedorov, and D. Chruściński, Probability representation of quantum dynamics using pseudostochastic maps, Phys. Rev. A **101**, 052320 (2020).

[36] J. DeBrota, C. Fuchs, and B. Stacey, Qbism research group, http://www.physics.umb.edu/Research/QBism/.

[37] D. M. Appleby, H. Yadsan-Appleby, and G. Zauner, Galois automorphisms of a symmetric measurement, Quantum Inf. Comput. **13**, 672 (2013).

[38] W. F. Braasch and W. K. Wootters, Transition probabilities and transition rates in discrete phase space, Phys. Rev. A **102**, 052204 (2020).

[39] Our main claim, summarized in Table II, is that, in deriving the retrodiction *inside a QPR*, the structure coefficients are used only to obtain the X_γ from the v_γ. Equation (30) shows that one can use the structure coefficients to obtain the QPR $S^{\mathcal{E}}$ *from the objects of the Hilbert space*: thus, this fact does not contradict our main claim.

[40] For NQPR and SIC-POVMs, the proof is even simpler using the fact that $\sum_x G_x = d\mathbb{1}$: indeed, in this case we have $\sum_{xy} \text{Tr}[F_i G_x G_j G_y] = d^2 \text{Tr}[F_i G_j]$; but $\sum_{xy} \delta_{ix}\delta_{jy}\delta_{ij} = \delta_{ij}$ and therefore Eq. (29) can hold only if $d^2 \text{Tr}[F_i G_j] = \delta_{ij}$, i.e., $d^2 = 1$.

[41] The equivalence in color is a statement of irreversibility, since it implies $\mathcal{E}(a'|a) = \hat{\gamma}(a')$. Thus $\sum_a \mathcal{E}(a'|a)q(a) = \hat{\gamma}(a')$ for all $q(a)$, which is just to say that the channel irreversibily erases all information about the input.

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

[43] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. **45**, 2171 (2004).