Classicality condition on a system observable in a quantum measurement and a relative-entropy conservation law

Yui Kuramochi and Masahito Ueda
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
(Received 9 June 2014; revised manuscript received 10 January 2015; published 16 March 2015)

We consider the information flow on a system observable X corresponding to a positive-operator-valued measure under a quantum measurement process Y described by a completely positive instrument from the viewpoint of the relative entropy. We establish a sufficient condition for the relative-entropy conservation law which states that the average decrease in the relative entropy of the system observable X equals the relative entropy of the measurement outcome of Y, i.e., the information gain due to measurement. This sufficient condition is interpreted as an assumption of classicality in the sense that there exists a sufficient statistic in a joint successive measurement of Y followed by X such that the probability distribution of the statistic coincides with that of a single measurement of X for the premeasurement state. We show that in the case when X is a discrete projection-valued measure and Y is discrete, the classicality condition is equivalent to the relative-entropy conservation for arbitrary states. The general theory on the relative-entropy conservation is applied to typical quantum measurement models, namely, quantum nondemolition measurement, destructive sharp measurements on two-level systems, a photon counting, a quantum counting, homodyne and heterodyne measurements. These examples except for the nondemolition and photon-counting measurements do not satisfy the known Shannon-entropy conservation law proposed by Ban [M. Ban, J. Phys. A: Math. Gen. **32**, 1643 (1999)], implying that our approach based on the relative entropy is applicable to a wider class of quantum measurements.

DOI: 10.1103/PhysRevA.91.032110 PACS number(s): 03.65.Ta, 03.67.−a, 42.50.Lc, 42.50.Ar

I. INTRODUCTION

In spite of the inevitable state change by a quantum measurement process, some quantum measurement models are known to conserve the information about a system observable. Examples of such measurements in optical systems include the quantum nondemolition (QND) measurement [1] and the destructive photon-counting measurement [2–4] on a single-mode photon number. In the QND measurement, the number of photons is not destructed and the classical Bayes rule holds for the photon-number distributions of premeasurement and postmeasurement states. On the other hand, the photon-counting measurement is a destructive measurement on the system photon number, but we can still construct the photon-number distribution of the premeasurement state from the number of counts and the photon number of the postmeasurement state.

This kind of information-conserving quantum measurement was discussed by Ban [5–8] quantitatively in terms of the mutual information $I_{\hat{\rho}}(X:Y)$ between a system observable X described by a positive-operator-valued measure (POVM) and the measurement outcome of a completely positive (CP) instrument Y [9–12]. Ban established a condition for X and Y under which the following Shannon-entropy [13] conservation law holds:

$$I_{\hat{\rho}}(X:Y) = H_{\hat{\rho}}(X) - E_{\hat{\rho}}[H_{\hat{\rho}_y}(X)], \quad (1)$$

where $\hat{\rho}$ is the premeasurement state, $\hat{\rho}_y$ is the postmeasurement state conditioned on the measurement outcome y, $E_{\hat{\rho}}[\ldots]$ denotes the ensemble average over the measurement outcome y for given $\hat{\rho}$, and $H_{\hat{\rho}}(X)$ is the Shannon entropy computed from the distribution of X for state $\hat{\rho}$. The left-hand side of Eq. (1) is the information gain about the system observable X which is obtained from the measurement outcome Y, while the right-hand side is a decrease in the uncertainty about the distribution of X due to the state change of the measurement. The physical meaning of the condition for the Shannon-entropy conservation (1) due to Ban is, however, not clear. There are also measurement models with continuous outcomes in which information about a system observable is conserved, but the Shannon-entropy conservation (1) does not hold due to a strong dependence of the continuous Shannon entropy, or differential entropy, on a reference measure of the probability measure. In this sense, it is difficult to regard Eq. (1) as the quantitative expression of the information conservation about X.

In this paper, we investigate the information flows of the measured observable based on the relative entropies [14] of the measurement process Y and the observable X. Operationally, the consideration of the relative entropies corresponds to the situation when the premeasurement state is assumed to be prepared in one of the two candidate states $\hat{\rho}$ or $\hat{\sigma}$, and the observer infers from the measurement outcome Y which state is actually prepared. This kind of information is quantified as relative entropy of Y between $\hat{\rho}$ and $\hat{\sigma}$. The same consideration applies to X and we can define the relative entropy of X for candidate states $\hat{\rho}$ and $\hat{\sigma}$ in a similar manner. Thus, we can compare these relative entropies as Ban did to the Shannon entropy and mutual information [6,7].

The primary finding of this paper is Theorem 1 which states that a kind of classicality condition for X and Y implies the relative-entropy conservation law which states that the relative entropy of the measurement outcome Y is equal to the ensemble-averaged decrease in the relative entropy of the system with respect to the POVM X. The classicality condition for X and Y assumed in Theorem 1 can be interpreted as the existence of a sufficient statistic [14,15] in a joint successive measurement of Y followed by X such that the distribution of the statistic coincides with that of X for the premeasurement state. This condition permits a classical interpretation of the measurement process Y in the sense that there exists a classical model that simulates the conditional change of the probability distribution of X in the measurement process Y computed from the system's density operator. It is also shown that the conservation of the relative entropy (8) holds in a wider range of quantum measurements than the Shannon-entropy conservation law (1) since the relative entropy is free from the dependence on the reference measure as in the Shannon entropy.

This paper is organized as follows. In Sec. II, we show the relative-entropy conservation law as Theorem 1 under a classicality condition for a system POVM X and a measurement process Y. A special case in which X is projection valued is formulated in Theorem 2. By further assuming the discreteness of both the projection-valued measure X and the measurement outcome of Y, we establish in Theorem 3 the equivalence between the relative-entropy conservation law for arbitrary candidate states and the classicality condition assumed in Theorem 2, i.e., the classicality condition is a necessary and sufficient condition for the relative-entropy conservation law in this case. In Sec. III, we show that typical quantum measurements satisfy the classicality condition, which are quantum nondemolition measurements, destructive sharp measurements on two-level systems, photon-counting measurement, quantum-counter measurement, balanced homodyne measurement, and heterodyne measurement. In these examples, except for the quantum nondemolition and photon-counting measurements, we show that the Shannon-entropy conservation law (1) does not hold. In Sec. IV, we summarize the main results of this paper.

II. RELATIVE-ENTROPY CONSERVATION LAW

In this section, we consider a quantum system described by a Hilbert space $\mathcal{H}$, a system's POVM X, and measurement process Y described by a CP instrument. Here, we assume that X is described by a density $\{\hat{E}^X_x\}_{x\in\Omega_X}$ of POVM with respect to a reference measure $\nu_0(dx)$ and that Y is described by a density of CP instrument $\{\mathcal{E}^Y_y\}_{y\in\Omega_Y}$ with respect to a reference measure $\mu_0(dy)$. The probability densities for the measurement outcomes for X and Y for a given density operator $\hat{\rho}$ are given by

$$p^X_{\hat{\rho}}(x) = \mathrm{tr}\left[\hat{\rho}\hat{E}^X_x\right]$$

and

$$p^Y_{\hat{\rho}}(y) = \mathrm{tr}\left[\mathcal{E}^Y_y(\hat{\rho})\right] = \mathrm{tr}\left[\hat{\rho}\hat{E}^Y_y\right], \tag{2}$$

respectively, where $\hat{E}^Y_y = \mathcal{E}^{Y\dagger}_y(\hat{I})$ is the density of the POVM for the measurement outcome y, $\hat{I}$ is the identity operator, and the adjoint $\mathcal{E}^\dagger$ of a superoperator $\mathcal{E}$ is defined by $\mathrm{tr}[\hat{\rho}\mathcal{E}^\dagger(\hat{A})] := \mathrm{tr}[\mathcal{E}(\hat{\rho})\hat{A}]$ for arbitrary $\hat{\rho}$ and $\hat{A}$. The postmeasurement state for a given measurement outcome y of Y is given by

$$\hat{\rho}_y = \frac{\mathcal{E}^Y_y(\hat{\rho})}{P^Y_{\hat{\rho}}(y)}. \tag{3}$$

The densities of POVMs $\hat{E}^X_x$ and $\hat{E}^Y_y$ satisfy the following completeness conditions:

$$\int \mu_0(dy)\hat{E}^Y_y = \hat{I}, \tag{4}$$

$$\int \nu_0(dx)\hat{E}^X_x = \hat{I}. \tag{5}$$

As the information content of the measurement outcome, we consider the relative entropies of the measurement outcomes for X and Y given by

$$D_X(\hat{\rho}||\hat{\sigma}) := D\big(p^X_{\hat{\rho}}\big|\big|p^X_{\hat{\sigma}}\big) = \int \nu_0(dx)p^X_{\hat{\rho}}(x)\ln\left[\frac{p^X_{\hat{\rho}}(x)}{p^X_{\hat{\sigma}}(x)}\right] \tag{6}$$

and

$$D\big(p^Y_{\hat{\rho}}\big|\big|p^Y_{\hat{\sigma}}\big) = \int \mu_0(dy)p^Y_{\hat{\rho}}(y)\ln\left[\frac{p^Y_{\hat{\rho}}(y)}{p^Y_{\hat{\sigma}}(y)}\right], \tag{7}$$

respectively. The relative entropies in Eqs. (6) and (7) are information contents obtained from the measurement outcomes as to which state $\hat{\rho}$ or $\hat{\sigma}$ is initially prepared.

The main goal of this work is to establish a condition for X and Y such that the relative-entropy conservation law

$$D\big(p^Y_{\hat{\rho}}\big|\big|p^Y_{\hat{\sigma}}\big) = D\big(p^X_{\hat{\rho}}\big|\big|p^X_{\hat{\sigma}}\big) - E_{\hat{\rho}}\big[D\big(p^X_{\hat{\rho}_y}\big|\big|p^X_{\hat{\sigma}_y}\big)\big] \tag{8}$$

holds. Before discussing the condition for X and Y we rewrite Eq. (8) in a more tractable form as in the following lemma.

Lemma 1. Let $\{\hat{E}^X_x\}_{x\in\Omega_X}$ be a density of POVM with respect to a reference measure $\nu_0(dx)$ and let $\{\mathcal{E}^Y_y\}_{y\in\Omega_Y}$ be a density of CP instrument with respect to a reference measure $\mu_0(dy)$. Then, the relative-entropy conservation law (8) is equivalent to

$$D\big(\tilde{p}^{XY}_{\hat{\rho}}\big|\big|\tilde{p}^{XY}_{\hat{\sigma}}\big) = D\big(p^X_{\hat{\rho}}\big|\big|p^X_{\hat{\sigma}}\big), \tag{9}$$

where $\tilde{p}^{XY}(x,y)$ is the probability distribution for a successive joint measurement of Y followed by X.

Proof. The joint distribution $\tilde{p}^{XY}(x,y)$ and the conditional probability distribution $\tilde{p}^{X|Y}_{\hat{\rho}}(x|y)$ of X under given measurement outcome y are given by

$$\tilde{p}^{XY}_{\hat{\rho}}(x,y) = \mathrm{tr}\left[\mathcal{E}^Y_y(\hat{\rho})\hat{E}^X_x\right] = \mathrm{tr}\left[\hat{\rho}\mathcal{E}^{Y\dagger}_y(\hat{E}^X_x)\right]$$

and

$$\tilde{p}^{X|Y}_{\hat{\rho}}(x|y) := \frac{\tilde{p}^{XY}_{\hat{\rho}}(x,y)}{p^Y_{\hat{\rho}}(y)} = p^X_{\hat{\rho}_y}(x), \tag{10}$$

respectively. In deriving Eq. (10), we used the fact that the marginal distribution of Y is given by Eq. (2) and the definition of the postmeasurement state in Eq. (3). From the chain rule for the classical relative entropy (e.g., Chap. 2 of Ref. [16]), we have

$$\begin{aligned} D\big(\tilde{p}^{XY}_{\hat{\rho}}\big|\big|\tilde{p}^{XY}_{\hat{\sigma}}\big) &= D\big(p^Y_{\hat{\rho}}\big|\big|p^Y_{\hat{\sigma}}\big) + E_{\hat{\rho}}\big[D(\tilde{p}^{X|Y}_{\hat{\rho}}(\cdot|y)\big|\big|\tilde{p}^{X|Y}_{\hat{\sigma}}(\cdot|y))\big] \\ &= D\big(p^Y_{\hat{\rho}}\big|\big|p^Y_{\hat{\sigma}}\big) + E_{\hat{\rho}}\big[D\big(p^X_{\hat{\rho}_y}\big|\big|p^X_{\hat{\sigma}_y}\big)\big], \end{aligned} \tag{11}$$

where we used Eq. (10) in deriving the second equality. Here, $D(\tilde{p}^{X|Y}_{\hat{\rho}}(\cdot|y)||\tilde{p}^{X|Y}_{\hat{\sigma}}(\cdot|y))$ denotes the relative entropy between the conditional probabilities (10) under a given measurement outcome y. The equivalence between Eqs. (8) and (9) is now evident from Eq. (11). ■

Equation (9) indicates that the information about X contained in the original states $\hat{\rho}$ and $\hat{\sigma}$ is equal to the information obtained from the joint successive measurement of Y followed by X.

Now, our first main result is the following theorem on the relative-entropy conservation law:

Theorem 1. Let X be a density of POVM $\{\hat{E}_x^X\}_{x\in\Omega_X}$ with respect to a reference measure $\nu_0(dx)$ and let Y be a density of an instrument $\{\mathcal{E}_y^Y\}_{y\in\Omega_Y}$ with respect to a reference measure $\mu_0(dy)$. Suppose that X and Y satisfy the following conditions.

(1) POVM of Y is the coarse graining of X, i.e., there exists a conditional probability $p(y|x) \geqslant 0$ such that

$$\hat{E}_y^Y = \int \nu_0(dx)p(y|x)\hat{E}_x^X \tag{12}$$

with the normalization condition

$$\int \mu_0(dy)p(y|x) = 1. \tag{13}$$

(2) There exist functions $\tilde{x}(x;y)$ and $q(x;y) \geqslant 0$ such that

$$\mathcal{E}_y^{Y\dagger}(\hat{E}_x^X) = q(x;y)\hat{E}_{\tilde{x}(x;y)}^X \tag{14}$$

for any x and y.

(3) For any y and any smooth function $F(x)$,

$$\int \nu_0(dx)q(x;y)F(\tilde{x}(x;y)) = \int \nu_0(dx)p(y|x)F(x). \tag{15}$$

Then, the relative-entropy conservation law (8) or (9) holds.

Proof. We prove Eq. (9). By taking a quantum expectation of Eq. (14) with respect to $\hat{\rho}$, we obtain

$$\tilde{p}_{\hat{\rho}}^{XY}(x,y) = q(x;y)p_{\hat{\rho}}^X(\tilde{x}(x;y)). \tag{16}$$

Equation (16) implies that, from the factorization theorem for the sufficient statistic [15], the stochastic variable $\tilde{x}(x;y)$ is a sufficient statistic of the joint successive measurement of Y followed by X. Let us denote the probability distribution function of $\tilde{x}(x;y)$ with respect to the reference measure ν_0 as $p_{\hat{\rho}}^{\tilde{X}}(x)$. From the definition of $p_{\hat{\rho}}^{\tilde{X}}(x)$ and the condition (15), for any function $F(x)$ we have

$$\begin{aligned}
&\int \nu_0(dx)p_{\hat{\rho}}^{\tilde{X}}(x)F(x)\\
&= \int \nu_0(dx)\int \mu_0(dy)\tilde{p}_{\hat{\rho}}^{XY}(x,y)F(\tilde{x}(x;y))\\
&= \int \mu_0(dy)\int \nu_0(dx)p(y|x)p_{\hat{\rho}}^X(x)F(x)\\
&= \int \nu_0(dx)p_{\hat{\rho}}^X(x)F(x),
\end{aligned}$$

which implies that the probability distribution of $\tilde{x}(x;y)$ coincides with that of the single measurement of X. Thus, the condition (15) ensures

$$p_{\hat{\rho}}^{\tilde{X}}(x) = p_{\hat{\rho}}^X(x). \tag{17}$$

From Eqs. (16) and (17), we have

$$D(\tilde{p}_{\hat{\rho}}^{XY}||\tilde{p}_{\hat{\sigma}}^{XY}) = D(p_{\hat{\rho}}^{\tilde{X}}||p_{\hat{\sigma}}^{\tilde{X}}) = D(p_{\hat{\rho}}^X||p_{\hat{\sigma}}^X),$$

where in deriving the first equality, we used the relative entropy conservation for the sufficient statistic due to Kullback and Leibler [14]. ■

The physical meaning of the conditions (14) and (15) is clear from Eqs. (16) and (17); the condition (14) implies that $\tilde{x}(x;y)$ is a sufficient statistic for the joint successive measurement of Y followed by X and the condition (15) ensures that the distribution of $\tilde{x}(x;y)$ is equivalent to that of X for the premeasurement state.

The assumptions 1, 2, 3 in Theorem 1 are interpreted as a kind of classicality condition as the proof uses only the classical probabilities. In fact, a statistical model

$$\tilde{p}(x_{\rm in},y,x_{\rm out}) = \delta_{x_{\rm in},\tilde{x}(x_{\rm out};y)}q(x_{\rm out};y)p_{\hat{\rho}}^X(x_{\rm in})$$

with its sample space $\Omega_X \times \Omega_Y \times \Omega_X$ reproduces all the probabilities that appear in the proof, where $x_{\rm in}$ and $x_{\rm out}$ are the system's values of X before and after the measurement of Y, respectively, and y is the outcome of Y. Here, we assumed the discreteness of Ω_X for simplicity, but the same construction still applies to the continuous case.

In Ref. [7], Ban proves the conservation for the Shannon entropy (1) by assuming Eqs. (12), (13), (15), and

$$\mathcal{E}_y^{Y\dagger}(\hat{E}_x^X) = p(x|\tilde{x}(x;y))\hat{E}_{\tilde{x}(x;y)}^X \tag{18}$$

for all x and y. The condition (18) is stronger than our condition (14) since $q(x;y)$ is, in general, different from $p(x|\tilde{x}(x;y))$. In some examples discussed in the next section, we will show that condition (18) together with the Shannon-entropy conservation law (1) does not hold, whereas our condition for the relative-entropy conservation law (8) does. This implies that our condition can be applicable to a wider range of quantum measurements. Furthermore, for the case in which X is a projection-valued measure and labels x and y are both discrete, we can show that condition (18) is equivalent to the condition that the postmeasurement state is one of eigenstates of X if the premeasurement state is also one of them (see Appendix A for detail).

Now, we consider the case in which the reference POVM is a projection-valued measure (PVM) $\hat{E}_x^X$ which satisfies the following orthonormal completeness condition:

$$\hat{E}_x^X\hat{E}_{x'}^X = \delta_{x,x'}\hat{E}_x^X, \quad \sum_{x\in\Omega_X}\hat{E}_x^X = \hat{I} \quad \text{for discrete } x; \tag{19}$$

$$\hat{E}_x^X\hat{E}_{x'}^X = \delta(x-x')\hat{E}_x^X, \quad \int_{\mathbb{R}} dx\hat{E}_x^X = \hat{I} \quad \text{for continuous } x, \tag{20}$$

where $\delta_{x,x'}$ is the Kronecker delta and $\delta(x-x')$ is the Dirac delta function. If $\hat{E}_x^X$ is written as $|x\rangle\langle x|$, the X-relative entropy

$$D_{\rm diag}(\hat{\rho}||\hat{\sigma}) := \begin{cases}\sum_{x\in\Omega_X}\langle x|\hat{\rho}|x\rangle \ln\left(\frac{\langle x|\hat{\rho}|x\rangle}{\langle x|\hat{\sigma}|x\rangle}\right),\\ \int dx\langle x|\hat{\rho}|x\rangle \ln\left(\frac{\langle x|\hat{\rho}|x\rangle}{\langle x|\hat{\sigma}|x\rangle}\right)\end{cases}$$

is called the diagonal-relative entropy. For this reference PVM, the condition for the relative-entropy conservation law is relaxed as shown in the following theorem.

Theorem 2. Let $\{\mathcal{E}_y^Y\}_{y\in\Omega_Y}$ be a density of an instrument with respect to a reference measure $\mu_0(dy)$ and $\hat{E}_x^X$ be a PVM with the completeness condition (19) or (20). Suppose that X and Y satisfy the condition (14) in Theorem 1. Then there exists a unique positive function $p(y|x)$ satisfying Eqs. (12) and (13). Furthermore, the relative-entropy conservation law in Eq. (8) holds.

Proof. For simplicity, we only consider the case in which the label x for the PVM is discrete. The following proof can easily be generalized to continuous X by replacing the sum $\sum_x \ldots$ with the integral $\int dx \ldots$ and the Kronecker delta $\delta_{x,x'}$ with the Dirac delta function $\delta(x-x')$.

The summation of Eq. (14) with respect to x gives

$$\hat{E}_y^Y = \sum_{x\in\Omega_X} q(x;y)\hat{E}^X_{\tilde{x}(x;y)}$$
$$= \sum_{x'\in\Omega_X}\left[\sum_{x\in\Omega_X}\delta_{x',\tilde{x}(x;y)}q(x;y)\right]\hat{E}^X_{x'}. \tag{21}$$

Therefore,

$$p(y|x) = \sum_{x'\in\Omega_X}\delta_{x,\tilde{x}(x';y)}q(x';y) \tag{22}$$

satisfies Eq. (12). The uniqueness and the normalization condition (13) for $p(y|x)$ follow from Eq. (21) and the completeness condition (4) for $\hat{E}_y^Y$ noting that $\{\hat{E}_x^X\}_{x\in\Omega_X}$ is linearly independent.

Next, we show the relative-entropy conservation law (8). From Theorem 1, it is sufficient to show the condition (15). For an arbitrary function $F(x)$ we have

$$\sum_{x\in\Omega_X} q(x;y)F(\tilde{x}(x;y)) = \sum_{x'\in\Omega_X}\left[\sum_{x\in\Omega_X}\delta_{x',\tilde{x}(x;y)}q(x;y)\right]F(x')$$
$$= \sum_{x\in\Omega_X} p(y|x)F(x),$$

where we used Eq. (22) in the second equality. Then, the condition (15) holds. ■

Next, we consider the case in which X is a discrete PVM $\{\hat{E}_x^X\}_{x\in\Omega_X}$ with the discrete complete orthonormal condition (19) and Y is a discrete measurement on a sample space Ω_Y described by a set of CP maps $\{\mathcal{E}_y^X\}_{y\in\Omega_Y}$ with the completeness condition

$$\sum_{y\in\Omega_Y}\mathcal{E}_y^{Y\dagger}(\hat{I}) = \hat{I}. \tag{23}$$

In this case, we can show the equivalence between the established condition (14) in Theorem 2 and the relative-entropy conservation law (8).

Theorem 3. Let X be a discrete PVM $\{\hat{E}_x^X\}_{x\in\Omega_X}$ with a discrete complete orthonormal condition (19) and let Y be a quantum measurement corresponding to a CP instrument on a discrete sample space Ω_Y described by a set of CP maps $\{\mathcal{E}_y^X\}_{y\in\Omega_Y}$ with the completeness condition (23). Then, the following two conditions are equivalent:

(i) The condition (14) holds for all x and y.

(ii) The relative-entropy conservation law (8) or (9) holds for arbitrary states $\hat{\rho}$ and $\hat{\sigma}$.

To show the theorem, we need the following lemma.

Lemma 2. Let $\{\hat{E}^X\}_{x\in\Omega_X}$ be a PVM with a discrete complete orthonormal condition (19) and let $\{\hat{E}_z^Z\}_{z\in\Omega_Z}$ be a discrete POVM. Suppose that

$$D\big(p_{\hat{\rho}}^X||p_{\hat{\sigma}}^X\big) = D\big(p_{\hat{\rho}}^Z||p_{\hat{\sigma}}^Z\big) \tag{24}$$

holds for any states $\hat{\rho}$ and $\hat{\sigma}$, where $p_{\hat{\rho}}^X(x) = \mathrm{tr}[\hat{\rho}\hat{E}_x^X]$ and $p_{\hat{\rho}}^Z(z) = \mathrm{tr}[\hat{\rho}\hat{E}_z^Z]$. Then, for each $z\in\Omega_Z$ there exist a scalar $q(z)\geqslant 0$ and $\tilde{x}(z)\in\Omega_X$ such that

$$\hat{E}_z^Z = q(z)\hat{E}^X_{\tilde{x}(z)}. \tag{25}$$

Proof of Lemma 2. Let $\hat{U}_x$ be an arbitrary operator such that $\hat{U}_x^\dagger\hat{U}_x = \hat{U}_x\hat{U}_x^\dagger = \hat{E}_x^X$, i.e., $\hat{U}_x$ is an arbitrary unitary operator on a closed subspace $\hat{E}_x^X\mathcal{H}$, where $\mathcal{H}$ is the system's Hilbert space. Define a CP and trace-preserving map $\mathcal{F}$ by

$$\mathcal{F}(\hat{\rho}) := \sum_{x\in\Omega_X}\hat{U}_x\hat{\rho}\hat{U}_x^\dagger.$$

Since $\hat{E}_x\hat{U}_{x'} = \hat{E}_x\hat{U}_{x'}\hat{U}_{x'}^\dagger\hat{U}_{x'} = \hat{E}_x\hat{E}_{x'}\hat{U}_{x'} = \delta_{x,x'}\hat{U}_{x'}$, we have $p_{\hat{\rho}}^X(x) = p^X_{\mathcal{F}(\hat{\rho})}(x)$ for any state $\hat{\rho}$. Therefore, from the assumption (24) we have

$$D\big(p_{\hat{\rho}}^Z||p^Z_{\mathcal{F}(\hat{\rho})}\big) = D\big(p_{\hat{\rho}}^X||p^X_{\mathcal{F}(\hat{\rho})}\big) = 0,$$

and hence we obtain

$$p_{\hat{\rho}}^Z(z) = p^Z_{\mathcal{F}(\hat{\rho})}(z)$$

for any $\hat{\rho}$ and any $z\in\Omega_Z$, which is in the Heisenberg picture represented as

$$\hat{E}_z^Z = \mathcal{F}^\dagger\big(\hat{E}_z^Z\big) = \sum_{x\in\Omega_X}\hat{U}_x^\dagger\hat{E}_z^Z\hat{U}_x. \tag{26}$$

By taking $\hat{U}_x$ as $\hat{E}_x^X$, we have

$$\hat{E}_z^Z = \sum_{x\in\Omega_X}\hat{E}_x^X\hat{E}_z^Z\hat{E}_x^X. \tag{27}$$

From Eqs. (26) and (27), an operator $\hat{E}_x^X\hat{E}_z^Z\hat{E}_x^X$ on $\hat{E}_x^X\mathcal{H}$ commutes with an arbitrary unitary $\hat{U}_x$ on $\hat{E}_x^X\mathcal{H}$, and therefore $\hat{E}_x^X\hat{E}_z^Z\hat{E}_x^X$ is proportional to the projection $\hat{E}_x^X$. Thus, we can rewrite Eq. (27) as

$$\hat{E}_z^Z = \sum_{x\in\Omega_X}\kappa(z|x)\hat{E}_x^X,$$

where $\kappa(z|x)$ is a non-negative scalar that satisfies the normalization condition $\sum_{z\in\Omega_Z}\kappa(z|x) = 1$. Let us define a POVM $\{\hat{E}_{xz}^{XZ}\}_{(x,z)\in\Omega_X\times\Omega_Z}$ by

$$\hat{E}_{xz}^{XZ} := \kappa(z|x)\hat{E}_x^X,$$

whose marginal POVMs are given by $\hat{E}_x^X$ and $\hat{E}_z^Z$, respectively. Since the probability distribution for $\hat{E}_{xz}^{XZ}$ is given by

$$p_{\hat{\rho}}^{XZ}(x,z) := \mathrm{tr}\big[\hat{\rho}\hat{E}_{xz}^{XZ}\big] = \kappa(z|x)p_{\hat{\rho}}^X(x), \tag{28}$$

X is a sufficient statistic for a statistical model $\{p_{\hat{\rho}}^{XZ}(x,z)\}_{\hat{\rho}\in\mathcal{S}(\mathcal{H})}$, where $\mathcal{S}(\mathcal{H})$ is the set of all the density operators on $\mathcal{H}$. Thus, from the sufficiency of X and the assumption (24), we have

$$D\big(p_{\hat{\rho}}^{XZ}||p_{\hat{\sigma}}^{XZ}\big) = D\big(p_{\hat{\rho}}^X||p_{\hat{\sigma}}^X\big) = D\big(p_{\hat{\rho}}^Z||p_{\hat{\sigma}}^Z\big).$$

Since a statistic that does not decrease the relative entropy is a sufficient statistic [14], Z is a sufficient statistic for $\{p_{\hat{\rho}}^{XZ}(x,z)\}_{\hat{\rho}\in\mathcal{S}(\mathcal{H})}$. Therefore, there is a non-negative scalar $r(x|z)$ such that

$$p_{\hat{\rho}}^{XZ}(x,z) = r(x|z)p_{\hat{\rho}}^{Z}(z),$$

or equivalently in the Heisenberg picture

$$\kappa(z|x)\hat{E}_x^X = r(x|z)\hat{E}_z^Z. \tag{29}$$

To prove (25), we have only to consider the case of $\hat{E}_z^Z \neq 0$. For such $z \in \Omega_Z$, there exists $x \in \Omega_X$ such that $\kappa(z|x)\hat{E}_x^X \neq 0$. Thus, from Eq. (29) we have $\hat{E}_z^Z = \frac{\kappa(z|x)}{r(x|z)}\hat{E}_x^X$ and the condition (25) holds. ■

Proof of Theorem 3. (i) ⇒ (ii) is evident from Theorem 2. Conversely, (i) readily follows from (ii) and Lemma 2 by identifying $\hat{E}_z^Z$ with $\mathcal{E}_y^{Y\dagger}(\hat{E}_x^X)$. ■

III. EXAMPLES OF RELATIVE-ENTROPY CONSERVATION LAW

In this section, we apply the general theorem obtained in the previous section to some typical quantum measurements, namely, a quantum nondemolition measurement, a measurement on two-level sytems, a photon-counting measurement, a quantum-counter model, homodyne and heterodyne measurements.

A. Quantum nondemolition measurement

We first consider a quantum nondemolition (QND) measurement [17–19] of a system's PVM $|x\rangle\langle x|$. In the QND measurement, the X distribution of the system is not disturbed by the measurement backaction. This condition is mathematically expressed as

$$p_{\mathcal{E}^Y(\hat{\rho})}^X(x) = p_{\hat{\rho}}^X(x) \tag{30}$$

for all $\hat{\rho}$, where

$$\mathcal{E}^Y = \int \mu_0(dy)\mathcal{E}_y^Y$$

is the completely positive (CP) and trace-preserving map which describes the state change of the system in the measurement of Y in which the measurement outcome is completely discarded. The QND condition in Eq. (30) is also expressed in the Heisenberg representation as

$$\mathcal{E}^{Y\dagger}(|x\rangle\langle x|) = |x\rangle\langle x|. \tag{31}$$

Let $\hat{M}_{yz}$ be the Kraus operator [10] of the CP map $\mathcal{E}_y^Y$ such that

$$\mathcal{E}_y^Y(\hat{\rho}) = \sum_z \hat{M}_{yz}\hat{\rho}\hat{M}_{yz}^\dagger.$$

Then, Eq. (31) becomes

$$\int \mu_0(dy)\sum_z \hat{M}_{yz}^\dagger|x\rangle\langle x|\hat{M}_{yz} = |x\rangle\langle x|. \tag{32}$$

Taking the diagonal element of Eq. (32) over the state $|x'\rangle$ with $x \neq x'$, we have

$$\int \mu_0(dy)\sum_z |\langle x|\hat{M}_{yz}|x'\rangle|^2 = 0.$$

Therefore, the Kraus operator $\hat{M}_{yz}$ is diagonal in the x basis and, from Eq. (12), it can be written as

$$\hat{M}_{yz} = \begin{cases} \sum_x e^{i\theta(x;y,z)}\sqrt{p(y,z|x)}|x\rangle\langle x|, \\ \int dx\, e^{i\theta(x;y,z)}\sqrt{p(y,z|x)}|x\rangle\langle x|, \end{cases} \tag{33}$$

where $p(y,z|x)$ satisfies

$$p(y|x) = \sum_z p(y,z|x).$$

We take the reference PVM $|x\rangle\langle x|$, and from Eq. (33) we have

$$\mathcal{E}_y^{Y\dagger}(|x\rangle\langle x|) = \sum_z \hat{M}_{yz}^\dagger|x\rangle\langle x|\hat{M}_{yz} = p(y|x)|x\rangle\langle x|, \tag{34}$$

which ensures the condition (14) with

$$\tilde{x}(x;y) = x,$$
$$q(x;y) = p(x|y).$$

Thus, from Theorem 2 the relative-entropy conservation law (8) holds. In this case, Ban's condition (18) and Shannon-entropy conservation law in Eq. (1) also hold [7].

The relative-entropy conservation law in Eq. (8) in the QND measurement can be understood in a classical manner as follows. Let us consider a change in the x-distribution function from $p_{\hat{\rho}}^X(x)$ to $p_{\hat{\rho}_y}^X(x)$. In the QND measurement, by using Eq. (34), the distribution of X for the conditional postmeasurement state becomes

$$p_{\hat{\rho}_y}^X(x) = \frac{p(y|x)p_{\hat{\rho}}^X(x)}{p_{\hat{\rho}}^Y(y)}. \tag{35}$$

Note that the commutativity of $|x\rangle\langle x|$ and $\hat{M}_{yz}$ is essential in deriving Eq. (35). Then, Eq. (35) can be interpreted as Bayes' rule for the conditional probability of X under measurement outcome of Y. Since the QND measurement does not disturb the system observable X, the change in the X distribution of the system is only the modification of observer's knowledge so as to be consistent with the obtained measurement outcome of Y based on Bayes' rule in Eq. (35). Bayes' rule is also valid in a classical setup in which the information about the system X is conveyed from the classical measurement outcome Y without disturbing X. Since we can derive the relative-entropy conservation law in Eq. (8) from Bayes' rule in Eq. (35), we can conclude that the relative-entropy conservation law in both classical and QND measurements is derived from the same Bayes' rule, or the modification of the observer's knowledge.

The rest of this section is devoted to examples of demolition measurements in which the reference POVM observable X is disturbed by the measurement backaction, yet the relative-entropy conservation law still holds.

B. Measurements on two-level systems

We consider a two-level system corresponding to a two-dimensional Hilbert space spanned by complete orthonormal kets $|0\rangle$ and $|1\rangle$. As the reference PVM of the system, we take

$$\hat{E}_x^X = |x\rangle\langle x| \quad (x = 0,1). \tag{36}$$

We consider a measurement Y described by the following instrument:

$$\mathcal{E}_y^Y(\hat{\rho}) = \hat{\phi}_y \langle y|\hat{\rho}|y\rangle \quad (y = 0,1), \tag{37}$$

where $\hat{\phi}_y$ is an arbitrary state. From Eq. (37) we can show that

$$\hat{E}_y^Y = |y\rangle\langle y|,$$

$$\mathcal{E}_y^{Y\dagger}(|x\rangle\langle x|) = \langle x|\hat{\phi}_y|x\rangle |y\rangle\langle y|$$

or

$$p(y|x) = \delta_{x,y}, \tag{38}$$

$$q(x;y) = \langle x|\hat{\phi}_y|x\rangle, \tag{39}$$

$$\tilde{x}(x;y) = y. \tag{40}$$

Then, the conditions for Theorem 2 are satisfied and the relative-entropy conservation law

$$D\big(p_{\hat{\rho}}^Y \big|\big| p_{\hat{\sigma}}^Y\big) = D_X(\hat{\rho}||\hat{\sigma}) - E_{\hat{\rho}}[D_X(\hat{\rho}_y||\hat{\sigma}_y)]$$
$$= D_X(\hat{\rho}||\hat{\sigma})$$

holds. The second equality follows from $\hat{\rho}_y = \hat{\sigma}_y$. On the other hand, from Eqs. (38)–(40), Ban's condition (18) does not hold if the postmeasurement state $\hat{\phi}_y$ does not coincide with one of eigenstates $|x\rangle\langle x|$.

Let us examine the Shannon-entropy conservation law (1). To make the discussion concrete, we assume $\hat{\phi}_y = \hat{I}/2$. Then, the Shannon entropy of X and the mutual information between X and Y are evaluated to be

$$I_{\hat{\rho}}(X:Y) = H_{\hat{\rho}}(X) = -\sum_{x=0,1} \langle x|\hat{\rho}|x\rangle \ln\langle x|\hat{\rho}|x\rangle,$$

$$H_{\hat{\rho}_y}(X) = H_{\hat{\phi}_y}(X) = \ln 2.$$

Thus,

$$H_{\hat{\rho}}(X) - H_{\hat{\rho}_y}(X) = I_{\hat{\rho}}(X:Y) - \ln 2 \neq I_{\hat{\rho}}(X:Y).$$

Therefore, the Shannon-entropy conservation law (1) does not hold. In this measurement model, the measured information of Y is maximal and any information is not contained in the postmeasurement state. This fact is properly reflected in the fact $D_X(\hat{\rho}_y||\hat{\sigma}_y) = 0$ if we consider the relative entropy, while the Shannon entropy is nonzero if the postmeasurement state is an eigenstate. This is the reason why the Shannon-entropy conservation law (1) does not hold.

C. Photon-counting measurement

The photon-counting measurement described in Refs. [2–4] measures the photon number in a closed cavity in a destructive manner and continuously in time. The measurement process in an infinitesimal time interval dt is described by the following measurement operators:

$$\hat{M}_0(dt) = \hat{I} - \left(i\omega + \frac{\gamma}{2}\right)\hat{n}\,dt, \tag{41}$$

$$\hat{M}_1(dt) = \sqrt{\gamma dt}\,\hat{a}, \tag{42}$$

where ω is the angular frequency of the observed cavity photon mode, $\gamma > 0$ is the coupling constant of the photon field with the detector, $\hat{a}$ is the annihilation operator of the photon field, and $\hat{n} := \hat{a}^\dagger\hat{a}$ is the photon-number operator. The event corresponding to the measurement operator in Eq. (41) is called the no-count process in which there is no photocount, while the event corresponding to Eq. (42) is called the one-count process in which a photocount is registered. In the one-count process, the postmeasurement wave function is multiplied by the annihilation operator $\hat{a}$ which decreases the number of photons in the cavity by one. Thus, this measurement is not a QND measurement.

From the measurement operators for an infinitesimal time interval in Eqs. (41) and (42), we can derive an effective measurement operator for a finite-time interval $[0,t)$ as follows [cf. Eq. (29) in Ref. [3]]:

$$\hat{M}_m(t) = \sqrt{\frac{(1-e^{-\gamma t})^m}{m!}}\, e^{-(i\omega+\frac{\gamma}{2})t\hat{n}}\hat{a}^m, \tag{43}$$

where m is the number of photocounts in the time interval $[0,t)$, which corresponds to the measurement outcome y in Sec. II. The POVM for the measurement operator in Eq. (43) can be written as

$$\hat{M}_m^\dagger(t)\hat{M}_m(t) = p(m|\hat{n};t), \tag{44}$$

where

$$p(m|n;t) = \binom{n}{m}(1-e^{-\gamma t})^m e^{-\gamma t(n-m)}. \tag{45}$$

Equation (44) shows that the measurement outcome m conveys the information about the cavity photon number $\hat{n}$. Especially in the infinite-time limit $t\to\infty$, the conditional probability in Eq. (45) becomes $\delta_{m,n}$, indicating that the number of counts m conveys the complete information about the photon-number distribution of the system. Then, we take the reference PVM as the projection operator into the number state $|n\rangle\langle n|$, with $\hat{n}|n\rangle = n|n\rangle$ and the orthonormal condition $\langle n|n'\rangle = \delta_{n,n'}$. From the measurement operator in Eq. (43), we obtain

$$\hat{M}_m^\dagger(t)|n\rangle\langle n|\hat{M}_m(t) = q(n;m;t)|\tilde{n}(n;m)\rangle\langle\tilde{n}(n;m)|, \tag{46}$$

$$\tilde{n}(n;m) = n + m, \tag{47}$$

$$q(n;m;t) = p(m|m+n;t). \tag{48}$$

Equation (47) can be interpreted as the photon number of the premeasurement state when the number of photocounts is m and the photon number remaining in the postmeasurement state is n. From Eqs. (46)–(48), the condition (14) for Theorem 2, together with Ban's condition (18), is satisfied and we have the relative-entropy conservation relation for the photon-counting measurement as

$$D(p_{\hat{\rho}}(\cdot;t)||p_{\hat{\sigma}}(\cdot;t)) = D_{\text{diag}}(\hat{\rho}||\hat{\sigma}) - E[D_{\text{diag}}(\hat{\rho}_m(t)||\hat{\sigma}_m(t))],$$

where $p_{\hat{\rho}}(m;t) = \mathrm{tr}[\hat{\rho}\hat{M}_m^\dagger(t)\hat{M}_m(t)]$ is the probability distribution of the number of photocounts m. We remark that the Shannon-entropy conservation law in Eq. (1) also holds in this measurement [5].

D. Quantum-counter model

A quantum-counter model [20,21] is a continuous-in-time measurement on a single-mode photon field in which no-count and one-count measurement operators for an infinitesimal time interval dt are given by

$$\hat{M}_0(dt) = \hat{I} - \frac{\gamma}{2}\hat{a}\hat{a}^\dagger dt, \quad \hat{M}_1(dt) = \sqrt{\gamma dt}\hat{a}^\dagger,$$

respectively. The effective measurement operator for a finite-time interval $[0,t]$ is known to be dependent only on the total number m of counting events in the time interval and given by [21]

$$\hat{M}_m^{\mathrm{qc}}(t) = \sqrt{\frac{(e^{\gamma t}-1)^m}{m!}}e^{-\gamma t\hat{a}\hat{a}^\dagger/2}(\hat{a}^\dagger)^m. \tag{49}$$

The POVM for this measurement is then

$$\hat{E}_m^{\mathrm{qc}}(t) = \hat{M}_m^{\mathrm{qc}\dagger}(t)\hat{M}_m^{\mathrm{qc}}(t) = \frac{(e^{\gamma t}-1)^m}{m!}\hat{a}^m e^{-\gamma t\hat{a}\hat{a}^\dagger}(\hat{a}^\dagger)^m$$
$$= p^{\mathrm{qc}}(m|\hat{n};t),$$

where

$$p^{\mathrm{qc}}(m|n;t) = \binom{n+m}{m}(e^{\gamma t}-1)^m e^{-\gamma t(n+m+1)}.$$

In this measurement model, we can show two kinds of relative-entropy conservation laws corresponding to two different system observables. As the first observable, we take the PVM $|n\rangle\langle n|$. Then, from Eq. (49), we have

$$\hat{M}_m^{\mathrm{qc}\dagger}(t)|n\rangle\langle n|\hat{M}_m^{\mathrm{qc}}(t) = p^{\mathrm{qc}}(m|\tilde{n}(n;m);t)|\tilde{n}(n;m)\rangle\langle\tilde{n}(n;m)|, \tag{50}$$

$$\tilde{n}(n;m) = n - m \tag{51}$$

and the conditions for Theorem 2, together with Ban's condition (18), hold. Therefore the relative-entropy conservation law

$$D\big(p_{\hat{\rho}}^{\mathrm{qc}}(\cdot;t)\big|\big|p_{\hat{\sigma}}^{\mathrm{qc}}(\cdot;t)\big) = D\big(p_{\hat{\rho}}^N||p_{\hat{\sigma}}^N\big) - E_{\hat{\rho}}\big[D\big(p^N_{\hat{\rho}_m(t)}||p^N_{\hat{\sigma}_m(t)}\big)\big] \tag{52}$$

holds, where

$$p_{\hat{\rho}}^{\mathrm{qc}}(m;t) = \mathrm{tr}\left[\hat{\rho}\hat{E}_m^{\mathrm{qc}}(t)\right] = \sum_{n=0}^{\infty}p^{\mathrm{qc}}(m|n;t)\langle n|\hat{\rho}|n\rangle,$$
$$p_{\hat{\rho}}^N(n) = \langle n|\hat{\rho}|n\rangle,$$

with $\hat{\rho}_m(t)$ being the postmeasurement state when the measurement outcome is m.

The second system's POVM is given by

$$\hat{E}_x^X dx = p^X(x|\hat{n})dx, \quad p^X(x|n) = \frac{e^{-x}x^n}{n!}, \tag{53}$$

where x is a real positive variable. The probability distribution of X

$$p_{\hat{\rho}}^X(x)dx = \mathrm{tr}\left[\hat{\rho}\hat{E}_x^X\right]dx$$

is known to be the distribution of $\lim_{t\to\infty} m/e^{\gamma t}$, corresponding to the total information obtained during the infinite-time interval [21]. Equation (53) implies that X is obtained by coarse graining $\hat{n}$. It can be shown [21] that the distribution $p_{\hat{\rho}}^X(x)$ determines the photon-number distribution by

$$\langle n|\hat{\rho}|n\rangle = \frac{d^n}{dx^n}\left[e^x p_{\hat{\rho}}^X(x)\right]\bigg|_{x=0}.$$

However, this just implies that the Markov mapping

$$p_{\hat{\rho}}^X(x) = \sum_{n=0}^{\infty}p^X(x|n)p_{\hat{\rho}}^N(n)$$

is injective and we cannot conclude that the information contained in X and $\hat{n}$ are the same as the following discussion shows.

From Eqs. (49) and (53) we obtain

$$\hat{M}_m^{\mathrm{qc}\dagger}(t)p^X(x|\hat{n})\hat{M}_m^{\mathrm{qc}}(t) = q(x;m)p^X(\tilde{x}(x;m)|\hat{n}), \tag{54}$$

$$q(x;m) = e^{-\gamma t}p^{\mathrm{qc}}(m|\tilde{x}(x;m)), \tag{55}$$

$$p^{\mathrm{qc}}(m|x) = \frac{[(e^{\gamma t}-1)x]^m}{m!}\exp[-(e^{\gamma t}-1)x],$$
$$\tilde{x}(x;m) = e^{-\gamma t}x. \tag{56}$$

Here, $p^{\mathrm{qc}}(m|x)$ satisfies $\sum_{m=0}^{\infty}p^{\mathrm{qc}}(m|x) = 1$. Furthermore, for an arbitrary function $F(x)$,

$$\int_0^\infty dx\, q(x;m)F(\tilde{x}(x;m))$$
$$= \int_0^\infty d(e^{-\gamma t}x)p^{\mathrm{qc}}(m|e^{-\gamma t}x)F(e^{-\gamma t}x)$$
$$= \int_0^\infty dx\, p^{\mathrm{qc}}(m|x)F(x). \tag{57}$$

The POVM for the measurement outcome m can be written as

$$\hat{M}_m^{\mathrm{qc}\dagger}(t)\hat{M}_m^{\mathrm{qc}}(t) = \int_0^\infty dx\,\hat{M}_m^{\mathrm{qc}\dagger}(t)p^X(x|\hat{n})\hat{M}_m^{\mathrm{qc}}(t)$$
$$= \int_0^\infty dx\, q(x;m)p^X(\tilde{x}(x;m)|\hat{n})$$
$$= \int_0^\infty dx\, p^{\mathrm{qc}}(m|x)p^X(x|\hat{n}). \tag{58}$$

From Eqs. (54), (57), and (58) and Theorem 1, the relative-entropy conservation law

$$D\big(p_{\hat{\rho}}^{\mathrm{qc}}(\cdot;t)\big|\big|p_{\hat{\sigma}}^{\mathrm{qc}}(\cdot;t)\big) = D\big(p_{\hat{\rho}}^X\big|\big|p_{\hat{\sigma}}^X\big) - E_{\hat{\rho}}\left[D\big(p^X_{\hat{\rho}_m(t)}||p^X_{\hat{\sigma}_m(t)}\big)\right] \tag{59}$$

holds.

Let us consider the asymptotic behaviors of relative entropies in the limit $t\to\infty$. Since $m/e^{\gamma t}$ converges to X in distribution, we have

$$D\big(p_{\hat{\rho}}^{\mathrm{qc}}(\cdot;t)\big|\big|p_{\hat{\sigma}}^{\mathrm{qc}}(\cdot;t)\big) \xrightarrow{t\to\infty} D\big(p_{\hat{\rho}}^X\big|\big|p_{\hat{\sigma}}^X\big). \tag{60}$$

From Eqs. (52), (59) and (60) we obtain

$$E_{\hat{\rho}}\left[D\left(p^N_{\hat{\rho}_m(t)}\middle|\middle|p^N_{\hat{\sigma}_m(t)}\right)\right] \xrightarrow{t\to\infty} D\left(p^N_{\hat{\rho}}\middle|\middle|p^N_{\hat{\sigma}}\right) - D\left(p^X_{\hat{\rho}}\middle|\middle|p^X_{\hat{\sigma}}\right), \tag{61}$$

$$E_{\hat{\rho}}\left[D\left(p^X_{\hat{\rho}_m(t)}\middle|\middle|p^X_{\hat{\sigma}_m(t)}\right)\right] \xrightarrow{t\to\infty} 0. \tag{62}$$

From the chain rule of relative entropy [16], the right-hand side of Eq. (61) is evaluated to be

$$\int_0^\infty dx\, p^X_{\hat{\rho}}(x) D\left(p^N_{\hat{\rho}}(\cdot|x)\middle|\middle|p^N_{\hat{\sigma}}(\cdot|x)\right) \geqslant 0, \tag{63}$$

where

$$p^N_{\hat{\rho}}(n|x) = \frac{p^X(x|n)p^N_{\hat{\rho}}(n)}{p^X_{\hat{\rho}}(x)} \tag{64}$$

is the photon-number distribution conditioned by X. The equality in (63) holds if and only if the photon-number distributions of $\hat{\rho}$ and $\hat{\sigma}$ coincide. This can be shown as follows.

If the equality in Eq. (63) holds, we have $D(p^N_{\hat{\rho}}(\cdot|x)||p^N_{\hat{\sigma}}(\cdot|x)) = 0$ for almost all $x \geqslant 0$. Thus,

$$\forall\, n \geqslant 0, \quad p^N_{\hat{\rho}}(n|x) = p^N_{\hat{\sigma}}(n|x) \tag{65}$$

for almost all $x > 0$, and therefore we can take at least one $x > 0$ satisfying Eq. (65). From Eqs. (64) and (65), we have

$$\forall\, n \geqslant 0, \quad \frac{\langle n|\hat{\rho}|n\rangle}{p^X_{\hat{\rho}}(x)} = \frac{\langle n|\hat{\sigma}|n\rangle}{p^X_{\hat{\sigma}}(x)}. \tag{66}$$

Taking the summation of Eq. (66) over n, we have

$$p^X_{\hat{\rho}}(x) = p^X_{\hat{\sigma}}(x). \tag{67}$$

From Eqs. (66) and (67), we finally obtain $\langle n|\hat{\rho}|n\rangle = \langle n|\hat{\sigma}|n\rangle (\forall\, n \geqslant 0)$.

Since the right-hand side of Eq. (61) is the difference between the information contents of $\hat{n}$ and X, the above discussion shows that the measurement outcome m carries strictly smaller information than that contained in the photon-number distribution. Equation (61) also shows that the difference of these information contents is obtained by a projection measurement on the postmeasurement state.

From Eq. (55), Ban's condition (18) does not hold for X. The difference between the Shannon entropies of premeasurement and postmeasurement states is given by

$$\begin{aligned} &H_{\hat{\rho}}(X) - E_{\hat{\rho}}\left[H_{\hat{\rho}_m(t)}(X)\right] \\ &= H_{\hat{\rho}}(X) + \sum_{m=0}^{\infty} p^{\mathrm{qc}}_{\hat{\rho}}(m) \int_0^\infty dx\, p^X_{\hat{\rho}_m(t)}(x) \ln p^X_{\hat{\rho}_m(t)}(x) \\ &= H_{\hat{\rho}}(X) + \sum_{m=0}^{\infty} \int_0^\infty dx\, e^{-\gamma t} p^{\mathrm{qc}}(m|e^{-\gamma t}x) p^X_{\hat{\rho}}(e^{-\gamma t}x) \\ &\quad \times \ln\left[\frac{e^{-\gamma t} p^{\mathrm{qc}}(m|e^{-\gamma t}x) p^X_{\hat{\rho}}(e^{-\gamma t}x)}{p^{\mathrm{qc}}_{\hat{\rho}}(m)}\right] \\ &= -\gamma t + I_{\hat{\rho}}(X:\mathrm{qc}) \neq I_{\hat{\rho}}(X:\mathrm{qc}), \end{aligned} \tag{68}$$

and the Shannon-entropy conservation law (1) does not hold. The term $-\gamma t$ in Eq. (68) comes from the Jacobian of the variable transformation $x \to \tilde{x}(x;y) = e^{-\gamma t}x$ and the strong dependence of the Shannon entropy for a continuous variable on the reference measure dx. On the other hand, if we take the relative entropy, such dependence on the reference measure is absent and we can analyze both of information conservations of $\hat{n}$ and X in a consistent manner.

E. Balanced homodyne measurement

The balanced homodyne measurement [22–24] measures one of the quadrature amplitudes of a photon field $\hat{a}$ in a destructive manner such that the system photon field relaxes into a vacuum state $|0\rangle$. This measurement process is implemented by mixing the signal photon field with a classical local-oscillator field into two output modes via a 50%-50% beam splitter and taking the difference of the photocurrents of the two output signals. For later convenience, we define the following quadrature amplitude operators:

$$\hat{X}_1 := \frac{\hat{a}+\hat{a}^\dagger}{\sqrt{2}}, \quad \hat{X}_2 := \frac{\hat{a}-\hat{a}^\dagger}{\sqrt{2}i}.$$

The measurement operator in the interaction picture for an infinitesimal time interval dt is given by

$$\hat{M}(d\xi(t);dt) = \hat{I} - \frac{\gamma}{2}\hat{n}\,dt + \sqrt{\gamma}\hat{a}\,d\xi(t), \tag{69}$$

where γ is the strength of the coupling with the detector, $d\xi(t)$ is a real stochastic variable corresponding to the output homodyne current which satisfies the Itô rule

$$[d\xi(t)]^2 = dt. \tag{70}$$

The reference measure $\mu_0[\xi(\ldots)]$ for the measurement outcome is the Wiener measure in which infinitesimal increments $\{d\xi(s)\}_{s\in[0,t)}$ are independent Gaussian stochastic variables with mean 0 and variance dt. From the measurement operator in Eq. (69), the ensemble average of the outcome $d\xi(t)$ for the system's state $\hat{\rho}(t)$ at time t is given by

$$E[d\xi(t)|\hat{\rho}(t)] = \sqrt{2\gamma}\langle\hat{X}_1\rangle_{\hat{\rho}(t)}, \tag{71}$$

where $\langle\hat{A}\rangle_{\hat{\rho}} := \mathrm{tr}[\hat{\rho}\hat{A}]$. Equation (71) indicates that $d\xi(t)$ measures the quadrature amplitude of the system. The general properties of the continuous quantum measurement with such diffusive terms are investigated in Refs. [25,26].

The time evolution of the system prepared in a pure state $|\psi_0\rangle$ at $t = 0$ is given by the following stochastic Schrödinger equation:

$$|\psi(t+dt)\rangle = \hat{M}(d\xi(t);dt)|\psi(t)\rangle.$$

The solution is given by [23]

$$|\psi(t)\rangle = \hat{M}_{y(t)}(t)|\psi_0\rangle, \tag{72}$$

where

$$\hat{M}_{y(t)}(t) = e^{-\frac{\gamma t}{2}\hat{n}} \exp\left[y(t)\hat{a} - \tfrac{1}{2}(1-e^{-\gamma t})\hat{a}^2\right], \tag{73}$$

$$y(t) = \sqrt{\gamma}\int_0^t e^{-\frac{\gamma s}{2}} d\xi(s). \tag{74}$$

Note that the $\hat{a}^2$ term should be included in the exponent on the right-hand side of Eq. (73) to be consistent with the Itô rule given in Eq. (70). We also mention that the measurement operator in Eq. (73) does not commute with the quadrature amplitude operator $\hat{X}_1$ and therefore this measurement disturbs $\hat{X}_1$. In the infinite-time limit $t \to \infty$, the stochastic wave function in Eq. (72) approaches the vacuum state $|0\rangle$ regardless of the initial state, which also indicates the destructive nature of the measurement.

As the reference PVM, we take the spectral measure $|x\rangle_{1\,1}\langle x|$ of the quadrature amplitude operator $\hat{X}_1$, where $|x\rangle_1$ satisfies

$$\hat{X}_1|x\rangle_1 = x|x\rangle_1, \quad {}_1\langle x|x'\rangle_1 = \delta(x-x').$$

Then, the operator $\hat{M}^\dagger_{y(t)}(t)|x\rangle_{1\,1}\langle x|\hat{M}_{y(t)}(t)$ and the POVM for the measurement outcome $y(t)$ are evaluated to be (see Appendix B for derivation)

$$\hat{M}^\dagger_{y(t)}(t)|x\rangle_{1\,1}\langle x|\hat{M}_{y(t)}(t) = q(x;y(t);t)\,|\tilde{x}(x;y(t);t)\rangle_{1\,1}\langle\tilde{x}(x;y(t);t)|, \tag{75}$$

$$q(x;y(t);t) = e^{-\gamma t/2}p(y|\tilde{x}(x;y(t))), \tag{76}$$

$$p(y|x) = \frac{1}{\sqrt{2\pi e^{-\gamma t}(1-e^{-\gamma t})}} \times \exp\left[-\frac{[y-\sqrt{2}(1-e^{-\gamma t})x]^2}{2e^{-\gamma t}(1-e^{-\gamma t})}\right], \tag{77}$$

$$\tilde{x}(x;y(t);t) = e^{-\frac{\gamma t}{2}}x + \frac{y(t)}{\sqrt{2}}, \tag{78}$$

$$\mu_0(dy)\hat{M}^\dagger_y(t)\hat{M}_y(t) = dy\; p(y|\hat{X}_1), \tag{79}$$

where the arguments of $\tilde{x}(x;y)$ in Eq. (78) are the the measurement outcome $[y(t)/\sqrt{2}$ on the right-hand side] and the remaining signal of the system ($e^{-\frac{\gamma t}{2}}x$ on the right-hand side), in which the exponential decay factor describes the system's relaxation to the vacuum state and the loss of the initial information contained in the system. The POVM in Eq. (79) shows that the measurement outcome $y(t)$ contains unsharp information about the quadrature amplitude $\hat{X}_1$ and that in the infinite-time limit $t \to \infty$ the measurement reduces to the sharp measurement of $\sqrt{2}\hat{X}_1$.

Equation (75) indicates that the condition (14) for Theorem 2 is satisfied, and we obtain the relative-entropy conservation law

$$D\big(p^Y_{\hat{\rho}}(\cdot;t)\big|\big|p^Y_{\hat{\sigma}}(\cdot;t)\big) = D_{X_1}(\hat{\rho}||\hat{\sigma}) - E_{\hat{\rho}}\big[D_{X_1}(\hat{\rho}_{y(t)}(t)||\hat{\sigma}_{y(t)}(t))\big],$$

where

$$p^Y_{\hat{\rho}}(y;t)dy = \mathrm{tr}[\hat{\rho}\hat{M}_y(t)^\dagger\hat{M}_y(t)]\mu_0(dy)$$

is the probability distribution function of the measurement outcome $y(t)$ which is computed from the POVM in Eq. (79), $\hat{\rho}_{y(t)}(t)$ and $\hat{\sigma}_{y(t)}(t)$ are the conditional density operators for given measurement outcome $y(t)$, and $D_{X_1}(\hat{\rho}||\hat{\sigma})$ is the diagonal relative entropy of the quadrature amplitude operator $\hat{X}_1$.

On the other hand, from Eq. (76) Ban's condition (18) does not hold. The difference between the Shannon entropies is evaluated to be

$$\begin{aligned}H_{\hat{\rho}}(X) &- E_{\hat{\rho}}\big[H_{\hat{\rho}_y}(X)\big]\\ &= H_{\hat{\rho}}(X) + \int dx\,dy\, e^{-\gamma t/2}p(y|\tilde{x}(x;y))p^X_{\hat{\rho}}(\tilde{x}(x;y))\\ &\quad\times \ln\left[\frac{e^{-\gamma t/2}p(y|\tilde{x}(x;y))p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)}\right]\\ &= -\frac{\gamma t}{2} + I_{\hat{\rho}}(X:Y) \neq I_{\hat{\rho}}(X:Y),\end{aligned} \tag{80}$$

and Shannon-entropy conservation law does not hold. The term $-\gamma t/2$ in Eq. (80) again arises from the nonunit Jacobian of the transformation $x \to \tilde{x}(x;y)$ as in Eq. (68).

F. Heterodyne measurement

The heterodyne measurement simultaneously measures the two noncommuting quadrature amplitudes $\hat{X}_1$ and $\hat{X}_2$ in a destructive manner as in the homodyne measurement. One way of implementation is to take a large detuning of the local oscillator in the balanced homodyne setup. Then, the cosine and sine components of the homodyne current give the two quadrature amplitudes [24].

The measurement operator for the heterodyne measurement in an infinitesimal time interval dt is given by

$$\hat{M}(d\zeta(t);dt) = \hat{I} - \frac{\gamma}{2}\hat{n}\,dt + \sqrt{\gamma}\hat{a}\,d\zeta(t), \tag{81}$$

where $d\zeta(t)$ is a complex variable obeying the complex Itô rules

$$[d\zeta(t)]^2 = [d\zeta^*(t)]^2 = 0, \quad d\zeta(t)d\zeta^*(t) = dt. \tag{82}$$

As in the homodyne measurement, we consider the time evolution in the interaction picture. The reference measure μ_0 for the measurement outcome $\zeta(\ldots)$ is the complex Wiener measure in which real and imaginary parts of $d\zeta(\ldots)$ are statistically independent Gaussian variables with zero mean and second-order moments consistent with the complex Itô rules in Eq. (82).

The stochastic evolution of the wave function is described by the following stochastic Schrödinger equation:

$$|\psi(t+dt)\rangle = \hat{M}(dt;d\zeta(t))|\psi(t)\rangle. \tag{83}$$

The solution of Eq. (83) for the initial condition $|\psi_0\rangle$ at $t=0$ is given by [23]

$$|\tilde{\psi}(t)\rangle = \hat{M}_{y(t)}(t)|\psi_0\rangle,$$

where

$$\hat{M}_{y(t)}(t) = e^{-\frac{\gamma t}{2}\hat{n}}e^{y(t)\hat{a}}, \tag{84}$$

$$y(t) = \sqrt{\gamma}\int_0^t e^{-\frac{\gamma s}{2}}d\zeta(s). \tag{85}$$

Here, the measurement operator in Eq. (84) does not involve the $\hat{a}^2$ term unlike the case of the homodyne measurement in Eq. (73) because $[d\zeta(t)]^2$ vanishes in this case.

Let us evaluate the POVM for the measurement outcome $y(t)$ in Eq. (85). From Eq. (84), we have

$$\hat{M}^{\dagger}_{y(t)}(t)\hat{M}_{y(t)}(t) = \mathscr{A}\{\exp[\gamma t - (e^{\gamma t}-1)\hat{a}\hat{a}^{\dagger} + e^{\gamma t}[y(t)\hat{a} + y^*(t)\hat{a}^{\dagger}] - e^{\gamma t}|y(t)|^2]\}, \tag{86}$$

where $\mathscr{A}\{f(\hat{a},\hat{a}^{\dagger})\}$ denotes the antinormal ordering in which the annihilation operators are placed to the left of the creation operators. To obtain the proper POVM for the measurement outcome $y(t)$, we have to multiply the operator $\hat{M}^{\dagger}_{y(t)}(t)\hat{M}_{y(t)}(t)$ by the measure $\mu_0[dy(t)]$ which is the measure for the reference complex Wiener measure. In the complex Wiener measure, the variable $y(t)$ in Eq. (85) is a Gaussian variable with zero mean and the second-order moments

$$E_0[y^2(t)] = 0, \quad E_0[|y(t)|^2] = 1 - e^{-\gamma t}.$$

Thus, the reference measure $\mu_0[dy(t)]$ is given by

$$\mu_0[dy(t)] = \frac{e^{-\frac{|y(t)|^2}{1-e^{-\gamma t}}}}{\pi(1-e^{-\gamma t})} d^2y(t), \tag{87}$$

where $d^2y = d(\mathrm{Re}y)d(\mathrm{Im}y)$. From Eqs. (86) and (87), the POVM for $y(t)$ is given by

$$d^2y(t)\mathscr{A}\{p(y(t)|\hat{a},\hat{a}^{\dagger};t)\},$$

where

$$p(y(t)|\alpha,\alpha^*;t) = \frac{\exp\left[-\frac{|y(t)-(1-e^{-\gamma t})\alpha^*|^2}{e^{-\gamma t}(1-e^{-\gamma t})}\right]}{\pi e^{-\gamma t}(1-e^{-\gamma t})}. \tag{88}$$

The probability distribution of the outcome $y(t)$ when the system is prepared in $\hat{\rho}_0$ at $t=0$ is given by

$$p^{Y}_{\hat{\rho}_0}(y;t) = \int d^2\alpha\, p(y(t)|\alpha,\alpha^*;t)Q_{\hat{\rho}_0}(\alpha,\alpha^*), \tag{89}$$

where $Q_{\hat{\rho}}(\alpha,\alpha^*) := \langle\alpha|\hat{\rho}|\alpha\rangle/\pi$ is the Q function [27,28], and $|\alpha\rangle$ is a coherent state [29] defined by

$$|\alpha\rangle = e^{\alpha\hat{a}^{\dagger}-\alpha^*\hat{a}}|0\rangle = e^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle.$$

From Eq. (88), in the infinite-time limit $t\to\infty$, the probability distribution of outcomes in Eq. (89) reduces to $Q_{\hat{\rho}_0}(y^*,y)$. Thus, the heterodyne measurement actually measures the noncommuting quadrature amplitudes simultaneously in the sense that the probability distribution of outcomes is the Q function of the initial state [30].

As a reference POVM, we take

$$d^2\alpha\hat{E}_{\alpha} = \frac{d^2\alpha}{\pi}|\alpha\rangle\langle\alpha| \tag{90}$$

which generates the Q function of the density operator. From Eqs. (84) and (90) we have

$$\mu_0(dy)\hat{M}^{\dagger}_{y}(t)\hat{E}_{\alpha}\hat{M}_{y}(t) = d^2y(t)q(\alpha,\alpha^*;y)\hat{E}_{\tilde{\alpha}(\alpha,y)}, \tag{91}$$

where

$$\tilde{\alpha}(\alpha,y) = e^{-\frac{\gamma t}{2}}\alpha + y^*, \tag{92}$$

$$q(\alpha,\alpha^*;y) = e^{-\gamma t}p(y|\tilde{\alpha}(\alpha;y),\tilde{\alpha}^*(\alpha;y)). \tag{93}$$

Note that the inferred quadrature amplitude in Eq. (92) allows a similar interpretation given in the homodyne analysis.

Equation (91) ensures the condition in Eq. (14). From Eqs. (88), (92), and (93), for an arbitrary smooth function $F(\alpha,\alpha^*)$, we have

$$\begin{aligned}&\int d^2\alpha\, q(\alpha,\alpha^*;y)F(\tilde{\alpha}(\alpha;y),\tilde{\alpha}^*(\alpha;y))\\ &= \int d^2\tilde{\alpha}\left(e^{\frac{\gamma t}{2}}\right)^2 q\left(e^{\frac{\gamma t}{2}}(\tilde{\alpha}+y^*),e^{\frac{\gamma t}{2}}(\tilde{\alpha}^*+y);y\right)F(\tilde{\alpha},\tilde{\alpha}^*)\\ &= \int d^2\alpha p(y|\alpha,\alpha^*;t)F(\alpha,\alpha^*).\end{aligned}$$

Thus, the condition (15) for Theorem 1 is satisfied and the relative-entropy conservation law

$$D\left(P^{Y}_{\hat{\rho}_0}(\cdot;t)\middle|\middle|P^{Y}_{\hat{\sigma}_0}(\cdot;t)\right) = D_{\mathrm{Q}}(\hat{\rho}_0||\hat{\sigma}_0) - E_{\hat{\rho}_0}[D_{\mathrm{Q}}(\hat{\rho}_{y(t)}||\hat{\sigma}_{y(t)})]$$

holds, where $\hat{\rho}_{y(t)}$ and $\hat{\sigma}_{y(t)}$ are the conditional density operators for a given measurement outcome $y(t)$ and $D_{\mathrm{Q}}(\hat{\rho}||\hat{\sigma})$ is the Q-function relative entropy defined as

$$D_{\mathrm{Q}}(\hat{\rho}||\hat{\sigma}) = \int d^2\alpha\, Q_{\hat{\rho}}(\alpha,\alpha^*)\ln\left[\frac{Q_{\hat{\rho}}(\alpha,\alpha^*)}{Q_{\hat{\sigma}}(\alpha,\alpha^*)}\right]. \tag{94}$$

Since the Q function has the complete quantum information about the quantum state, the Q-function relative entropy in Eq. (94) vanishes if and only if $\hat{\rho}=\hat{\sigma}$, which is not the case in the diagonal relative entropies in the preceding examples. Still, the Q-function relative entropy is bounded from above by the quantum relative entropy $S(\hat{\rho}||\hat{\sigma}) := \mathrm{tr}[\hat{\rho}(\ln\hat{\rho}-\ln\hat{\sigma})]$, for the relative entropy of probability distributions on the measurement outcome of a POVM is always smaller than the quantum relative entropy [31].

Equation (93) implies the violation of Ban's condition (18). The difference of the Shannon entropies is given by

$$\begin{aligned}&H_{\hat{\rho}}(Q) - E_{\hat{\rho}}\left[H_{\hat{\rho}_y}(Q)\right]\\ &= H_{\hat{\rho}}(Q) + \int d^2\alpha\, d^2y\, e^{-\gamma t}p(y|\tilde{\alpha}(\alpha;y))Q_{\hat{\rho}}(\tilde{\alpha}(\alpha;y))\\ &\quad\times\ln\left[\frac{e^{-\gamma t}p(y|\tilde{\alpha}(\alpha;y))Q_{\hat{\rho}}(\tilde{\alpha}(\alpha;y))}{p^{Y}_{\hat{\rho}}(y)}\right]\\ &= -\gamma t + I_{\hat{\rho}}(Q:Y) \neq I_{\hat{\rho}}(Q:Y)\end{aligned} \tag{95}$$

and the Shannon-entropy conservation does not hold. Again, the the term $-\gamma t$ in Eq. (95) originates from the nonunite Jacobian of the transformation $x\to\tilde{x}(x;y)$.

IV. SUMMARY

In this paper, we have examined the information flow in a general quantum measurement process Y concerning the relative entropy of the two quantum states with respect to a system's POVM X of the system. By assuming the classicality condition on X and Y, we have proved the relative-entropy conservation law when X is a general POVM (Theorem 1) and when X is a PVM (Theorem 2). The classicality condition can be interpreted as the existence of a sufficient statistic in a joint successive measurement of Y followed by X such that the distribution of the statistic coincides with that of X for the premeasurement state. This condition may be interpreted as a classicality condition because there exists a classical statistical model which generates all the relevant probability distributions of X and Y. We have also investigated the case in which the labels of the PVM X and the measurement outcome of Y are both discrete and we have shown the equivalence between the classicality condition in Theorem 2 and the relative-entropy conservation law for arbitrary states (Theorem 3). We have applied the general theorems to some typical quantum measurements. In the QND measurement, the relative-entropy conservation law can be understood as a result of the classical Bayes' rule which is a mathematical expression of the modification of our knowledge based on the outcome of the measurement. In the destructive sharp measurement of two-level systems, Ban's condition together with the Shannon-entropy conservation law does not hold, while our relative-entropy conservation law does. The next examples, namely, photon counting, quantum counter, balanced homodyne and heterodyne measurements, are non-QND measurements on a single-mode photon field and the measurement outcomes convey information about the photon number, part of the photon number, one and both quadrature amplitude(s), respectively. In spite of the destructive nature of the measurements, the classicality condition is still satisfied and we have shown that the relative-entropy conservation laws hold for these measurements. In the quantum-counter model, we can take two kinds of POVMs of the system satisfying the two relative-entropy conservation laws. In the heterodyne measurement X is the POVM which generates the Q function and is not an ordinary PVM, reflecting the fact that the noncommuting observables are measured simultaneously. In the examples of quantum counter, homodyne and heterodyne measurements, the Shannon-entropy conservation laws do not hold due to the nonunit Jacobian of the transformation $x \to \tilde{x}(x;y)$. These examples of nonconserving Shannon entropies suggest that our approach to the information transfer of the system observable is applicable to a wider range of measurement models than that based on the Shannon entropy.

ACKNOWLEDGMENTS

This work was supported by KAKENHI Grant No. 26287088 from the Japan Society for the Promotion of Science, and a Grant-in-Aid for Scientific Research on Innovation Areas "Topological Quantum Phenomena" (KAKENHI Grant No. 22103005), and the Photon Frontier Network Program from Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. Y.K. acknowledges supports by Advanced Leading Graduate Course for Photon Science (ALPS) at the University of Tokyo and Japan Society for the Promotion of Science (KAKENHI Grant No. 269905).

APPENDIX A: EQUIVALENT CONDITIONS FOR (18) WHEN X AND Y ARE DISCRETE

In this appendix, we characterize the condition (18) required by Ban when the reference POVM X is a discrete PVM and the measurement Y is also discrete. In this case, the condition (18) is equivalent to the condition that if a premeasurement state is an eigenstate of X, then the postmeasurement state is another eigenstate of X as shown in the following theorem.

Theorem 4. Let $\mathcal{E}_y^Y$ be a CP instrument with discrete measurement outcome y and $\hat{E}_x^X = |x\rangle\langle x|$ satisfying the assumption (14) of Theorem 2. Then, the following conditions are equivalent:

(1) Ban's condition (18) holds, i.e., $q(x;y) = p(y|\tilde{x}(x;y))$.

(2) For all x and y such that $p(y|x) \neq 0$,

$$\sum_{x'} \delta_{x,\tilde{x}(x';y)} = 1. \tag{A1}$$

(3) For all x and y such that $p(y|x) \neq 0$, there exists a unique x' such that $x = \tilde{x}(x';y)$.

(4) The postmeasurement state is an eigenstate of X if the premeasurement state is an eigenstate. Namely, for all x and y, there exist functions $\tilde{x}(x;y)$ and $r(x;y) \geqslant 0$ such that

$$\mathcal{E}_y^Y(|x\rangle\langle x|) = r(x;y)|\tilde{x}(x;y)\rangle\langle\tilde{x}(x;y)|. \tag{A2}$$

Before proving this theorem, we make a comment on the arbitrariness of the definition of $\tilde{x}(x;y)$ when $q(x;y) = 0$. In this case, $\tilde{x}(x;y)$ may take any value and we define it as $\emptyset$, which is out of the range of label space of X. We also define $p(y|\emptyset) = 0$ for any y.

Proof. $1 \Rightarrow 2$: We first note that $p(y|x)$ in this case is given by Eq. (22). By substituting $q(x';y) = p(y|\tilde{x}(x';y))$ into Eq. (22), we obtain

$$p(y|x) = \sum_{x'} \delta_{x,\tilde{x}(x';y)} p(y|\tilde{x}(x';y)) = \left(\sum_{x'} \delta_{x,\tilde{x}(x';y)}\right) p(y|x).$$

Therefore, Eq. (A1) holds whenever $p(y|x) \neq 0$.

The condition 3 immediately follows from 2 by noting the definition of the Kronecker's delta.

$3 \Rightarrow 4$: From Eq. (2),

$$p(y|x) = \mathrm{tr}\left[|x\rangle\langle x|\mathcal{E}_y^{Y\dagger}(\hat{I})\right] = \mathrm{tr}\left[\mathcal{E}_y^Y(|x\rangle\langle x|)\right]. \tag{A3}$$

If $p(y|x) = 0$, from Eq. (A3) and the positivity of $\mathcal{E}_y^Y(|x\rangle\langle x|)$, $\mathcal{E}_y^Y(|x\rangle\langle x|) = 0$ and the condition 4 hold. Let us consider the case in which $p(y|x) \neq 0$. Since $\mathcal{E}_y^{Y\dagger}$ is a CP map, it has the following Kraus representation [10]:

$$\mathcal{E}_y^{Y\dagger}(\hat{A}) = \sum_z \hat{M}_{yz}^\dagger \hat{A} \hat{M}_{yz}. \tag{A4}$$

From Eq. (14), we have

$$\sum_z \hat{M}_{yz}^\dagger |x\rangle\langle x|\hat{M}_{yz} = q(x;y)|\tilde{x}(x;y)\rangle\langle\tilde{x}(x;y)|.$$

Therefore, we can put

$$\hat{M}_{yz}^\dagger|x\rangle = a(x;y,z)|\tilde{x}(x;y)\rangle, \tag{A5}$$

where

$$\sum_z |a(x;y,z)|^2 = q(x;y). \tag{A6}$$

From Eqs. (A4) and (A5) we obtain

$$\mathcal{E}_y^{Y\dagger}(|x''\rangle\langle x'|) = \sum_z \hat{M}_{yz}^\dagger |x''\rangle\langle x'|\hat{M}_{yz} = \left[\sum_z a(x'';y,z)a^*(x';y,z)\right] \times |\tilde{x}(x'';y)\rangle\langle\tilde{x}(x';y)|. \quad \text{(A7)}$$

The matrix element of $\mathcal{E}_y^Y(|x\rangle\langle x|)$ is evaluated as

$$\langle x'|\mathcal{E}_y^Y(|x\rangle\langle x|)|x''\rangle = \mathrm{tr}\left[\mathcal{E}_y^Y(|x\rangle\langle x|)|x''\rangle\langle x'|\right] = \mathrm{tr}\left[|x\rangle\langle x|\mathcal{E}_y^{Y\dagger}(|x''\rangle\langle x'|)\right] = \left[\sum_z a(x'';y,z)a^*(x';y,z)\right] \times \delta_{x,\tilde{x}(x'';y)}\delta_{x,\tilde{x}(x';y)}, \quad \text{(A8)}$$

where we used Eq. (A7) in the last equality. From the condition 3, there exists a unique x' such that $x = \tilde{x}(x';y)$ and we write this x' as $\bar{x}(x;y)$. Then, Eq. (A8) becomes

$$\left[\sum_z |a(x';y,z)|^2\right]\delta_{x',\bar{x}(x;y)}\delta_{x'',\bar{x}(x;y)} = q(x';y)\delta_{x',\bar{x}(x;y)}\delta_{x'',\bar{x}(x;y)}, \quad \text{(A9)}$$

where we used Eq. (A6). Equation (A9) implies

$$\mathcal{E}_y^Y(|x\rangle\langle x|) = q(\bar{x}(x;y);y)|\bar{x}(x;y)\rangle\langle\bar{x}(x;y)|,$$

which is nothing but the condition 4.

$4 \Rightarrow 1$: From

$$\hat{E}_y^Y = \mathcal{E}_y^{Y\dagger}(\hat{I}) = \sum_x p(y|x)|x\rangle\langle x|$$

and Eq. (A2), we have

$$p(y|x) = \mathrm{tr}\left[|x\rangle\langle x|\mathcal{E}_y^{Y\dagger}(\hat{I})\right] = \mathrm{tr}\left[\mathcal{E}_y^Y(|x\rangle\langle x|)\right]x = r(x;y). \quad \text{(A10)}$$

From Eqs. (14), (A2), and (A10), we obtain

$$q(x;y) = \mathrm{tr}\left[|\tilde{x}(x;y)\rangle\langle\tilde{x}(x;y)|\mathcal{E}_y^{Y\dagger}(|x\rangle\langle x|)\right] = \mathrm{tr}\left[\mathcal{E}_y^Y[|\tilde{x}(x;y)\rangle\langle\tilde{x}(x;y)|]|x\rangle\langle x|\right] = p(y|\tilde{x}(x;y))\delta_{x,\bar{x}(\tilde{x}(x;y);y)}. \quad \text{(A11)}$$

When $q(x;y) \neq 0$, Eq. (A11) implies $q(x;y) = p(y|\tilde{x}(x;y))$. If $q(x;y) = 0$, $\tilde{x}(x;y) = \emptyset$ and $p(y|\emptyset) = q(x;y) = 0$ from the remark above the present proof. Thus, the condition (18) holds. ■

We briefly remark on the case when the PVM $|x\rangle\langle x|$ is continuous with the complete orthonormal condition (20). Under the same assumptions of Theorem 4, we can show that Ban's condition (18) implies

$$\int dx'\delta[x - \tilde{x}(x';y)] = 1 \quad \text{(A12)}$$

for any x and y such that $p(y|x) \neq 0$. The proof of Eq. (A12) is formally as the same as that of $1 \Rightarrow 2$ in Theorem 4. However, the formal correspondence between continuous and discrete X fails when we consider the other part of the proof of Theorem 4. For example, we cannot conclude from Eq. (A12) the existence and uniqueness of x' such that $\tilde{x}(x';y) = x$. For simplicity, let us assume the uniqueness of x' holds. Still, the condition (A12) is very restrictive since it implies

$$\left|\frac{\partial\tilde{x}(x';y)}{\partial x'}\right| = 1,$$

i.e., the Jacobian of the transformation $x \to \tilde{x}(x;y)$ should be 1. This reflects the strong dependence of the Shannon entropy on the reference measure, which is not the case in the relative entropy.

APPENDIX B: DERIVATIONS OF EQS. (75) AND (79)

To evaluate the operator $\hat{M}_{y(t)}^\dagger(t)|x\rangle_{11}\langle x|\hat{M}_{y(t)}(t)$, we utilize the technique of normal ordering. We first note that the normally ordered expression $:O(\hat{a},\hat{a}^\dagger):$ of an operator $\hat{O}$, in which the annihilation operators are placed to the right of the creation operators, is given by a coherent-state expectation as

$$O(\alpha,\alpha^*) = \langle\alpha|\hat{O}|\alpha\rangle.$$

Since the coherent state $|\alpha\rangle$ in the $|x\rangle_1$ representation is given by

$${}_1\langle x|\alpha\rangle = \pi^{-1/4}\exp\left[-\tfrac{1}{2}(x-\sqrt{2}\alpha)^2 - \tfrac{1}{2}(\alpha^2+|\alpha|^2)\right],$$

we have

$$\langle\alpha|x\rangle_{11}\langle x|\alpha\rangle = \pi^{-1/2}\exp\left[-\left(x - \frac{\alpha+\alpha^*}{\sqrt{2}}\right)^2\right],$$

which implies the following normally ordered expression:

$$|x\rangle_{11}\langle x| = \pi^{-1/2} : \exp\left[-\left(x - \frac{\hat{a}+\hat{a}^\dagger}{\sqrt{2}}\right)^2\right] : . \quad \text{(B1)}$$

By using Eq. (B1) and the formula

$$e^{-\lambda\hat{n}}|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}(1-e^{-2\lambda})}|e^{-\lambda}\alpha\rangle,$$

which is valid for real λ, the expectation of the operator $\hat{M}_y^\dagger(t)|x\rangle_{11}\langle x|\hat{M}_y(t)$ over the coherent state $|\alpha\rangle$ is evaluated to be

$$\langle\alpha|\hat{M}_{y(t)}^\dagger(t)|x\rangle_{11}\langle x|\hat{M}_{y(t)}(t)|\alpha\rangle = \pi^{-1/2}\exp\left\{-\left[e^{-\frac{\gamma t}{2}}x + \frac{y(t)}{\sqrt{2}} - \frac{\alpha+\alpha^*}{\sqrt{2}}\right]^2 + \left[e^{-\frac{\gamma t}{2}}x + \frac{y(t)}{\sqrt{2}}\right]^2 - x^2\right\}. \quad \text{(B2)}$$

Substituting Eq. (B1) in Eq. (B2), we obtain Eq. (75). By integrating Eq. (75) with respect to x and noting a relation

$$f(\hat{X}_1) = \int dx\, f(x)|x\rangle_{11}\langle x|,$$

which is valid for an arbitrary function $f(x)$, we obtain

$$\hat{M}_y(t)^\dagger \hat{M}_y(t) = \exp\left[\frac{\gamma t}{2} + \hat{X}_1^2 - e^{\gamma t}\left(\hat{X}_1 - \frac{y}{\sqrt{2}}\right)^2\right]. \tag{B3}$$

To evaluate the proper POVM for the outcome y, we need to multiply $\hat{M}_y^\dagger(t)\hat{M}_y(t)$ by $\mu_0[dy(t)]$, where $\mu_0[dy(t)]$ is the probability measure of $y(t)$, provided that $\xi(\ldots)$ obeys a Wiener distribution. Here, $y(t)$ in Eq. (74) under a Wiener measure μ_0 is a Gaussian stochastic variable with the first and second moments

$$E_0[y(t)] = 0,$$

$$E_0[y^2(t)] = \gamma \int_0^t e^{-\gamma s} ds = 1 - e^{-\gamma t},$$

where $E_0[\ldots]$ denotes the expectation with respect to the Wiener measure. Thus, $\mu_0[dy(t)]$ is given by

$$\frac{dy}{\sqrt{2\pi(1-e^{-\gamma t})}} \exp\left[-\frac{y^2}{2(1-e^{-\gamma t})}\right]. \tag{B4}$$

Multiplying Eq. (B3) by Eq. (B4), we obtain Eq. (79).

[1] C. Guerlin, J. Bernu, S. Deleglise, C. Sayrin, S. Gleyzes, S. Kuhr, M. Brune, J.-M. Raimond, and S. Haroche, Nature (London) **448**, 889 (2007).
[2] M. Srinivas and E. Davies, Opt. Acta **28**, 981 (1981).
[3] M. Ueda, Quantum Opt. **1**, 131 (1989).
[4] N. Imoto, M. Ueda, and T. Ogawa, Phys. Rev. A **41**, 4127 (1990).
[5] M. Ban, Phys. Lett. A **235**, 209 (1997).
[6] M. Ban, Int. J. Theor. Phys. **37**, 2491 (1998).
[7] M. Ban, J. Phys. A: Math. Gen. **32**, 1643 (1999).
[8] M. Ban, J. Phys. A: Math. Gen. **32**, 6527 (1999).
[9] E. B. Davies and J. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[10] K. Kraus, Ann. Phys. (NY) **64**, 311 (1971).
[11] K. Kraus, *States, Effects, and Operations* (Springer, Berlin, 1983).
[12] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
[13] C. E. Shannon, Bell Syst. Technol. J. **27**, 379 (1948); **27**, 623 (1948).
[14] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[15] P. R. Halmos and L. J. Savage, Ann. Math. Stat. **20**, 225 (1949).
[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, Hoboken, NJ, 2012).
[17] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, Rev. Mod. Phys. **52**, 341 (1980).
[18] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, UK, 1995).
[19] V. B. Braginsky and F. Y. Khalili, Rev. Mod. Phys. **68**, 1 (1996).
[20] M. Ueda and M. Kitagawa, Phys. Rev. Lett. **68**, 3424 (1992).
[21] M. Ueda, N. Imoto, and H. Nagaoka, Phys. Rev. A **53**, 3808 (1996).
[22] H. M. Wiseman and G. J. Milburn, Phys. Rev. A **47**, 642 (1993).
[23] H. M. Wiseman, Quantum Semiclass. Opt. **8**, 205 (1996).
[24] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, UK, 2010).
[25] H. M. Wiseman and L. Diósi, Chem. Phys. **268**, 91 (2001).
[26] A. Barchielli and M. Gregoratti, *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case*, Vol. 782 (Springer, Berlin, 2009).
[27] K. Husimi, Proc. Phys. Math. Soc. Japan. 3rd Series **22**, 264 (1940).
[28] C. L. Mehta and E. C. G. Sudarshan, Phys. Rev. **138**, B274 (1965).
[29] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).
[30] H. Carmichael, *Statistical Methods in Quantum Optics 2: Non-Classical Fields* (Springer, Berlin, 2008).
[31] M. Hayashi, *Quantum Information: An Introduction* (Springer, Berlin, 2006).