Relative-entropy conservation law in hypothesis testing based on quantum measurement

Yui Kuramochi\textsuperscript{1} and Masahito Ueda\textsuperscript{1}

\textsuperscript{1}Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

(Dated: June 10, 2014)

Abstract

By quantifying the distinguishability of two quantum states in terms of a relative entropy of the measured observable of a system, we establish a general condition under which a decrease in the relative entropy of the system equals the relative entropy of the measurement outcome, i.e., the information gain due to measurement. We also show that if the measured observable is an ordinary projection-valued measure the general condition leads to the relative entropy conservation law for an arbitrary pre-measurement state. We obtain three general theorems on the relative entropy conservation law and apply them to quantum non-demolition measurements and three destructive measurements, namely photon counting, homodyne and heterodyne measurements. The applications to the destructive measurements indicate that the relative-entropy conservation in quantum measurement is a wider concept than the quantum non-demolition principle which ensures the conservation of measured quantum numbers.

PACS numbers: 03.67.-a, 03.65.Ta, 42.50.Lc, 42.50.Ar
I. INTRODUCTION

Ever since the early days of quantum theory, the information obtained by quantum measurement and its relation to measurement back-action on the measured system have attracted interest of numerous researchers. A classic study on this subject is Heisenberg’s uncertainty relation [1]. However, early works on the quantum measurement faced two major difficulties.

For one thing, there was no general mathematical formulation of quantum measurement which can describe both statistics of measurement outcomes and measurement back-action for the measurement that cannot be described by the von Neumann’s projection postulate [2]. In the 70’s and 80’s, such a general formulation of quantum measurement was developed [3–6] in which an open system approach to quantum measurement is essential.

Another difficulty was the lack of the quantitative theory of information. Since the seminal work by Shannon [7], the quantitative and operational aspects of the classical information have extensively been investigated and many entropic information contents have been proposed. One of the most important information contents is the relative entropy, or the Kullback-Leibler divergence [8], which is defined by

\[ D(p||q) = \sum_x p(x) \ln \left( \frac{p(x)}{q(x)} \right), \]

if \( p \) and \( q \) are probability distributions on a discrete sample space, or

\[ D(p||q) := \int dx p(x) \ln \left( \frac{p(x)}{q(x)} \right), \]

if \( p \) and \( q \) are probability density functions on a continuous sample space. Statistically the relative entropy can be interpreted as the information about the distinguishability of the two probability distributions as shown in the Chernoff-Stein lemma [9], which states that in the setup of the hypothesis testing the probability of the error of the second kind behaves as \( \sim \exp[-ND(p||q)] \) in the large sample-size limit \( N \to \infty \). The relative entropy is also related to other fundamental information contents such as the mutual information [7] or the Fisher information [10] in parameter estimation.

The conservation of information in quantum measurement was discussed by Ban [11–14] in terms of the mutual information \( I(x : y) \) between a system’s observable \( x \) and the measurement outcome \( y \):

\[ I(x : y) = H_x(\hat{\rho}) - E_{\hat{\rho}}[H_x(\hat{\rho}_y)], \quad (1) \]
where $\hat{\rho}$ is the pre-measurement state, $\hat{\rho}_y$ is the post-measurement state conditioned on the measurement outcome $y$, $E_{\hat{\rho}}[\cdot]$ denotes the ensemble average over the measurement outcome $y$, and $H_x(\hat{\rho})$ 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’s 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 measurement back-action. The crucial assumption in this work is that we know the density operator $\hat{\rho}$ of the pre-measurement state; otherwise it would be impossible to calculate the right-hand side of Eq. (1). What can be stated about the information gain and the state reduction if the pre-measurement state is less known, for example, if we only know that the pre-measurement state is in one of the several candidate states $\hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_n$? The purpose of this paper is to address this question for the case of $n = 2$.

Our strategy is to consider the hypothesis testing problem in which the pre-measurement state is assumed to be prepared in one of two hypothesis states, $\hat{\rho}$ or $\hat{\sigma}$, and the observer infers from the measurement outcome which state is actually prepared. The information content obtained from the measurement is then the relative entropy of the probability distributions of the measurement outcome for the hypothesis states. The primary finding of this paper is the relative-entropy conservation law shown in Eq. (7) 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 observable $x$. The former gives the information gain due to measurement, while the latter is caused by the measurement back-action; thus the relative-entropy conservation law quantitatively characterizes the trade-off relation between the information gain and the measurement back-action.

This paper is organized as follows. In Sec. II, we show the relative-entropy conservation law as Theorem 1 when the system’s observable $x$ is a general positive operator-valued measure. A special case in which $x$ is a usual projection-valued measure is formulated in Theorem 2. By further assuming the discreteness of both the observed quantity $x$ and the measurement outcome $y$, we establish in Theorem 3 the equivalence of the relative-entropy conservation law and other conditions assumed in Theorem 2. In Sec. III, we apply the general theorems on the relative-entropy conservation law to typical quantum measurements, namely quantum non-demolition measurement, photon-counting measurement, balanced homodyne measurement, and heterodyne measurement. In Sec. IV, we summarize the main
results of this paper.

II. RELATIVE-ENTROPY CONSERVATION LAW

We consider a quantum measurement which is described by a set of completely positive (CP) maps $\{\mathcal{E}_y^Y(\cdot)\}$. We assume that the pre-measurement state is prepared in a quantum state described by a density operator $\hat{\rho}$ and that an observer knows the system is prepared in either $\hat{\rho}$ or $\hat{\sigma}$. The probability density of the outcome $y$ and post-measurement density operator for the pre-measurement state $\hat{\rho}$ are given by

$$P^Y_\hat{\rho}(y) = \text{tr}[\mathcal{E}_y^Y(\hat{\rho})] = \text{tr}[\hat{\rho}\hat{E}_y^Y],$$

$$\hat{P}_y = \frac{\mathcal{E}_y^Y(\hat{\rho})}{P^Y_\hat{\rho}(y)},$$

where $\hat{E}_y^Y = \mathcal{E}_y^{Y\dagger}(\hat{I})$ is the positive operator-valued measure (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 $\text{tr}[\mathcal{E}(\hat{\rho})\hat{A}] = \text{tr}[\hat{\rho}\mathcal{E}^\dagger(\hat{A})]$ for arbitrary $\hat{\rho}$ and $\hat{A}$. $P^Y_\hat{\sigma}(y)$ and $\hat{P}_y$ for the pre-measurement state $\hat{\sigma}$ are given in a similar manner. The completeness relation for the POVM $\hat{E}_y^Y$ is

$$\int \mu_0(dy)\hat{E}_y^Y = \hat{I},$$

where $\mu_0$ is a reference measure for the outcome labeled by $y$ and $\hat{I}$ is the identity operator.

We assume that the measurement outcome $y$ conveys the information of a system observable $x$ represented by a POVM $\{\hat{E}_x^X\}$ through the relation

$$\hat{E}_y^Y = \int \nu_0(dx)p(y|x)\hat{E}_x^X,$$  \hspace{1cm} (2)

where $\nu_0$ is a measure for the label $x$, and $p(y|x) \geq 0$ is the conditional probability with the normalization condition

$$\int \mu_0(dy)p(y|x) = 1.$$  \hspace{1cm} (3)

We call $\hat{E}_x^X$ a reference POVM which satisfies the completeness relation

$$\int \nu_0(dx)\hat{E}_x^X = \hat{I},$$  \hspace{1cm} (4)

and gives the probability distribution of the reference POVM for a quantum state $\hat{\rho}$ as

$$P^X_\hat{\rho}(x) = \text{tr}[\hat{\rho}\hat{E}_x^X].$$
The information for the hypothesis testing acquired from the measurement outcome \( y \) is quantitatively characterized by the relative entropy

\[
D(\rho_Y || \sigma_Y) = \int \mu_0(dy) \rho_Y^Y(y) \ln \left( \frac{\rho_Y^Y(y)}{\sigma_Y^Y(y)} \right).
\]

We also define the \( x \)-relative entropy of two quantum states \( \rho \) and \( \sigma \) as

\[
D_x(\rho || \sigma) := D(\rho_X || \sigma_X) = \int \nu_0(dx) \rho_X^X(x) \ln \left( \frac{\rho_X^X(x)}{\sigma_X^X(x)} \right),
\]

which is the information content on the distinguishability of \( \rho \) and \( \sigma \) when the complete information about \( x \) is accessible.

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

**Theorem 1.** Suppose that there exist a positive function \( q(x; y) \) and a function \( \tilde{x}(x; y) \) such that

\[
\mathcal{E}_{y}^X(\hat{E}_x^X) = q(x; y)\hat{E}_{\tilde{x}(x; y)}^X \quad \text{for all } x \text{ and } y, \tag{5}
\]

and

\[
\int \nu_0(dx)q(x; y)F(\tilde{x}(x; y)) = \int \nu_0(dx)p(y|x)F(x) \tag{6}
\]

for arbitrary \( y \) and a function \( F(x) \). Then we have the following relative-entropy conservation law:

\[
D(\rho_Y || \sigma_Y) = D_x(\hat{\rho} || \hat{\sigma}) - E_{\hat{\rho}}[D_x(\hat{\rho}_y || \hat{\sigma}_y)], \tag{7}
\]

where \( E_{\hat{\rho}}[\cdot] \) is the ensemble average over the measurement outcome \( y \) for pre-measurement state \( \hat{\rho} \).

**Proof.** From Eq. (5) the distribution of \( x \) for the post-measurement state \( \hat{\rho}_y \) is given by

\[
P_{\hat{\rho}_y}^X(x) = \frac{q(x; y)\rho_{\tilde{x}(x; y)}^X}{\rho_Y^Y(y)}, \tag{8}
\]

and the \( x \)-relative entropy of the post-measurement states is given by

\[
D_x(\hat{\rho}_y || \hat{\sigma}_y) = -\ln \left( \frac{\rho_Y^Y(y)}{\rho_Y^\sigma(y)} \right) + \int \nu_0(dx)p(y|x)\rho_{\tilde{x}(x; y)}^X \ln \left( \frac{\rho_{\tilde{x}(x; y)}^X}{\rho_{\hat{\sigma}}^X(x)} \right),
\]
where we have used Eq. (6) in deriving the last term. Then the ensemble average of the $x$-relative entropy of the post-measurement state is evaluated to be
\[
E_{\hat{\rho}}[D_x(\hat{\rho}_y||\hat{\sigma}_y)] = \int \mu_0(dy) P^Y_\rho(y) D_x(\hat{\rho}_y||\hat{\sigma}_y)
\]
\[
= -D(\hat{P}^Y_\rho||\hat{P}^Y_\sigma) + \int \nu_0(dx) \mu_0(dy) p(y|x) P^X_\rho(x) \ln \left( \frac{P^X_\rho(x)}{P^X_\sigma(x)} \right)
\]
\[
= -D(\hat{P}^Y_\rho||\hat{P}^Y_\sigma) + D_x(\hat{\rho}||\hat{\sigma}),
\]
where we used Eq. (3) in deriving the last equality, and we obtain Eq. (7).

We remark that $\tilde{x}(x;y)$ in Eq. (5) can be interpreted as the inferred value of $x$ for the pre-measurement state conditioned on the value of $x$ for the post-measurement state and the measurement outcome $y$. Thus the condition in Eq. (5) implies the conservation of the information about $x$ in the sense that initial $x$ is inferred uniquely from the post-measurement $x$ and the measurement outcome $y$.

We can interpret the relative-entropy conservation law in Eq. (7) as the balance between the information gain on the left-hand side and the decrease in the entropy of the system on the right-hand side. Another expression for the relative entropy-conservation law and its interpretation can be obtained as follows. We consider a situation in which a measurement of $y$ is first performed on $\rho$, and then a measurement of $x$ is carried out. The joint probability distribution of $x$ and $y$ is given by
\[
\tilde{P}^{XY}_{\hat{\rho}}(x,y) = \text{tr}[\mathcal{E}^Y_y(\hat{\rho}) \hat{E}^X_x] = \text{tr}[\hat{\rho} \mathcal{E}^{Y\dagger}_y(\hat{E}^X_x)].
\]
The marginal $y$-distribution of $\tilde{P}^{XY}_{\hat{\rho}}(x,y)$ is $P^Y_\rho(y)$ as is evident from the completeness condition Eq. (4), while the marginal $x$-distribution $\tilde{P}^X_\rho(x)$ does not necessarily coincide with $P^X_\rho(x)$ because of the back-action of the $y$-measurement. A special class of measurements in which two $x$-distributions before and after the measurement coincide are called quantum non-demolition measurements which will be discussed in the next section. From the joint probability $\tilde{P}^{XY}_{\hat{\rho}}(x,y)$, the formal definition of the conditional probability $\tilde{P}^{XY}_{\hat{\rho}}(x|y)$ gives
\[
\tilde{P}^{XY}_{\hat{\rho}}(x|y) := \frac{\tilde{P}^{XY}_{\hat{\rho}}(x,y)}{P^Y_\rho(y)} = P^X_{\rho_y}(x).
\]
From the chain rule for the classical relative entropy (e.g. Chap. 2 of Ref. [9]), we have
\[
D(\tilde{P}^X_{\hat{\rho}}||\tilde{P}^X_{\hat{\sigma}}) = D(P^X_{\hat{\rho}}||P^X_{\sigma}) + E_{\hat{\rho}}[D(\tilde{P}^{XY}_{\hat{\rho}}(\cdot|y)||\tilde{P}^{XY}_{\hat{\sigma}}(\cdot|y))] = D(P^Y_{\hat{\rho}}||P^Y_{\sigma}) + E_{\hat{\rho}}[D(P^X_{\hat{\rho}}||P^X_{\sigma})],
\]
(10)
where we used Eq. (9) in deriving the second equality. Note that in the derivations of Eqs. (9) and (10) we do not assume any additional condition on the measurement such as (2) or (5). From Eq. (10), the relative-entropy conservation law in Eq. (7) is equivalent to

\[ D(\tilde{P}_{\tilde{\rho}}^{XY}||\tilde{P}_{\tilde{\sigma}}^{XY}) = D(\tilde{P}_{\tilde{\rho}}^{X}||\tilde{P}_{\tilde{\sigma}}^{X}). \]  

Equation (11) indicates that the information about \( x \) contained in the original states \( \tilde{\rho} \) and \( \tilde{\sigma} \) is equal to the information obtained from the joint measurement of \( x \) followed by \( y \). Equation (11) provides an alternative expression of the conservation of information about \( x \).

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

\[ \langle x|x' \rangle = \delta_{x,x'}, \quad \sum_x |x\rangle\langle x| = \hat{I} \quad \text{for discrete} \ x; \tag{12} \]

\[ \langle x|x' \rangle = \delta(x - x'), \quad \int dx|x\rangle\langle x| = \hat{I} \quad \text{for continuous} \ x, \tag{13} \]

where \( \delta_{x,x'} \) is the Kronecker delta and \( \delta(x - x') \) is the Dirac delta function. For partly discrete and partly continuous label \( x \), the \( x \)-relative entropy for reference PVM \( |x\rangle\langle x| \) is the diagonal-relative entropy which is defined by

\[ D_{\text{diag}}(\tilde{\rho}||\tilde{\sigma}) := \begin{cases} 
\sum_x \langle x|\tilde{\rho}|x\rangle \ln \left( \frac{\langle x|\tilde{\rho}|x\rangle}{\langle x|\tilde{\sigma}|x\rangle} \right), \\
\int dx \langle x|\tilde{\rho}|x\rangle \ln \left( \frac{\langle x|\tilde{\rho}|x\rangle}{\langle x|\tilde{\sigma}|x\rangle} \right), 
\end{cases} \tag{14} \]

for discrete and continuous \( x \), respectively. For this reference PVM, the condition for the relative-entropy conservation law is relaxed as shown in the following theorem:

**Theorem 2.** Provided that there exist a positive function \( q(y;x) \) and an \( x \)-label valued function \( \tilde{x}(x;y) \) such that

\[ \mathcal{E}^Y_y(|x\rangle\langle x|) = q(y;x)|\tilde{x}(x;y)\rangle\langle \tilde{x}(x;y)| \]  

for all \( x \) and \( y \), the following the relative-entropy conservation law holds:

\[ D(P_{\tilde{\rho}}^{Y}||P_{\tilde{\sigma}}^{Y}) = D_{\text{diag}}(\tilde{\rho}||\tilde{\sigma}) - E_{\tilde{\rho}}[D_{\text{diag}}(\tilde{\rho}_y||\tilde{\sigma}_y)]. \]  

(16)
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 \( \sum x \cdots \) with the integral \( \int dx \cdots \) and the Kronecker delta \( \delta_{x,x'} \) with the Dirac delta function \( \delta(x-x') \).

From Theorem 1, it is sufficient to show that the condition (6) is satisfied. The summation of Eq. (15) with respect to \( x \) gives

\[
\hat{E}_y^Y = \sum_x q(x; y) |\tilde{x}(x; y)\rangle \langle \tilde{x}(x; y)| \\
= \sum_{x'} \left( \sum_x \delta_{x', \tilde{x}(x; y)} q(x; y) \right) |x'\rangle \langle x'|.
\]

(17)

From the complete orthonormality (12) and Eqs. (2) and (17), we obtain

\[
p(y|x) = \sum_{x'} \delta_{x, \tilde{x}(x'; y)} q(x'; y).
\]

Thus for an arbitrary function \( F(x) \) we have

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

which implies the condition in Eq. (6).

Next, let us consider the case in which the labels \( x \) for the reference PVM and \( y \) for the measurement outcome are both purely discrete. We also assume that the \( y \)-measurement is pure and the CP map \( \mathcal{E}_y^Y \) takes the following expression:

\[
\mathcal{E}_y^Y(\hat{\rho}) = \hat{M}_y \hat{\rho} \hat{M}_y^\dagger.
\]

The completeness condition for the POVM \( \hat{E}_y^Y = \hat{M}_y^\dagger \hat{M}_y \) then becomes

\[
\sum_y \hat{E}_y^Y = \hat{I}.
\]

(18)

In this situation, we have the following theorem on the equivalence of conditions:

**Theorem 3.** Let \( \hat{E}_x = |x\rangle \langle x| \) be a reference PVM with discrete label \( x \) and \( \hat{M}_y \) be a measurement operator with the completeness condition (18). Then the following three conditions are equivalent:

\[
\hat{E}_x = |x\rangle \langle x| \text{ is a PVM with discrete label } x, \\
\hat{E}_y^Y = \hat{M}_y^\dagger \hat{M}_y \text{ is a POVM for the } y \text{-measurement, and} \\
\sum_y \hat{E}_y^Y = \hat{I}.
\]
(i) The diagonal elements \( \{\langle x|\hat{M}_y\hat{\rho}\hat{M}_y^\dagger|x\rangle\}_x \) of the post-measurement state depend only on the diagonal elements \( \{\langle x'|\hat{\rho}|x'\rangle\}_x \) of the pre-measurement state for all \( y \).

(ii) The condition (15) holds for all \( x \) and \( y \).

(iii) The relative-entropy conservation law (16) holds for arbitrary states \( \hat{\rho} \) and \( \hat{\sigma} \).

**Proof.** We first show (i) \( \Rightarrow \) (ii). Since the diagonal element of the post-measurement state is given by

\[
\langle x|\hat{M}_y\hat{\rho}\hat{M}_y^\dagger|x\rangle = \sum_{x',x''} \langle x''|\hat{M}_y^\dagger|x\rangle \langle x|\hat{M}_y|x''\rangle \langle x'|\hat{\rho}|x''\rangle,
\]

the condition (i) is equivalent to

\[
\forall y, \forall x, \forall x', \forall x'' \neq x', \quad \langle x''|\hat{M}_y^\dagger|x\rangle \langle x|\hat{M}_y|x''\rangle = 0,
\]

and therefore \( \hat{M}_y^\dagger|x\rangle \langle x|\hat{M}_y \) has no off-diagonal element in the \( |x\rangle \) basis. Since \( \hat{M}_y^\dagger|x\rangle \langle x|\hat{M}_y \) is at most rank 1, we can rewrite it as

\[
\hat{M}_y^\dagger|x\rangle \langle x|\hat{M}_y = q(x; y)|\tilde{x}(x; y)\rangle \langle \tilde{x}(x; y)|
\]

with some positive scalar \( q(x; y) \) and a label \( \tilde{x}(x; y) \). Therefore the condition (15) is satisfied for all \( x \) and \( y \).

(ii) \( \Rightarrow \) (iii) is evident from Theorem 2.

Finally, we show (iii) \( \Rightarrow \) (i) by contraposition, that is, we show that the negation of (i) implies the negation of (iii). From the negation of (i), which is equivalent to the condition (19), there exist \( y_0, x_0, x_1 \) and \( x_2 \neq x_1 \) such that

\[
\langle x_0|\hat{M}_{y_0}|x_1\rangle (\langle x_0|\hat{M}_{y_0}|x_2\rangle)^* \neq 0.
\]

Thus we can take two non-zero complex scalars \( c_1 \) and \( c_2 \) such that

\[
|c_1|^2 + |c_2|^2 = 1
\]

and

\[
\langle x_0|\hat{M}_{y_0}|\psi\rangle = 0,
\]

where \( |\psi\rangle := c_1|x_1\rangle + c_2|x_2\rangle \). If we define quantum states \( \hat{\rho} \) and \( \hat{\sigma} \) by

\[
\hat{\rho} := |c_1|^2|x_1\rangle \langle x_1| + |c_2|^2|x_2\rangle \langle x_2|,
\]

\[
\hat{\sigma} := |\psi\rangle \langle \psi|,
\]

9
we obtain

$$D(P^Y_\hat{\rho} \| P^Y_\hat{\sigma}) = D_{\text{diag}}(\hat{\rho} \| \hat{\sigma}) = 0$$ \hspace{1cm} (22)$$

because $\hat{\rho}$ and $\hat{\sigma}$ have the same diagonal elements. On the other hand, since we have

$$\langle x_0 | \hat{M}_{y_0} \hat{\rho} \hat{M}_{y_0}^\dagger | x_0 \rangle = \sum_{k=1,2} |c_k|^2 |\langle x_0 | M_{j_0} | x_0 \rangle|^2 > 0,$$

$$\langle x_0 | \hat{M}_{y_0} \hat{\sigma} \hat{M}_{y_0}^\dagger | x_0 \rangle = |\langle x_0 | M_{j_0} | \psi \rangle|^2 = 0,$$

we obtain

$$P^Y_{\hat{\rho}} (y_0) = P^Y_{\hat{\sigma}} (y_0) \geq \langle x_0 | \hat{M}_{y_0} \hat{\rho} \hat{M}_{y_0}^\dagger | x_0 \rangle > 0,$$

$$\ln \left( \frac{\langle x_0 | \hat{\rho}_{y_0} | x_0 \rangle}{\langle x_0 | \hat{\sigma}_{y_0} | x_0 \rangle} \right) = \infty,$$

$$D_{\text{diag}}(\hat{\rho}_{y_0} \| \hat{\sigma}_{y_0}) = \sum_x \langle x | \hat{\rho}_{y_0} | x \rangle \ln \left( \frac{\langle x | \hat{\rho}_{y_0} | x \rangle}{\langle x | \hat{\sigma}_{y_0} | x \rangle} \right) = \infty.$$

Thus the ensemble average of the diagonal relative entropy for the post-measurement states satisfies

$$E_{\hat{\rho}}[D_{\text{diag}}(\hat{\rho} \| \hat{\sigma})] = P^Y_{\hat{\rho}} (y_0) D_{\text{diag}}(\hat{\rho}_{y_0} \| \hat{\sigma}_{y_0}) = \infty.$$ \hspace{1cm} (23)

Equations (22) and (23) imply the violation of the relative-entropy conservation law in Eq. (16). Thus, by contraposition, (iii) $\Rightarrow$ (i) is proved.

\[\square\]

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 quantum non-demolition measurements and three optical destructive measurements.

A. Quantum non-demolition measurements

We first consider a quantum non-demolition (QND) measurement [15–17] 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 back-action. This condition is mathematically expressed as

$$P^X_{\hat{\epsilon}}(x) = P^X_{\hat{\rho}}(x)$$ \hspace{1cm} (24)
for all $\hat{\rho}$, where

$$E^y = \int \mu_0(dy)E^y_y$$

is the 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. (24) is also expressed in the Heisenberg representation as

$$E^{y\dagger}(|x\rangle\langle x|) = |x\rangle\langle x|.$$  \hspace{1cm} (25)

Let $\hat{M}_{yz}$ be the Kraus operator [4] of the CP map $E^y_y$ such that

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

Then Eq. (25) becomes

$$\int \mu_0(dy) \sum_z \hat{M}_{yz}^\dagger|x\rangle\langle x|\hat{M}_{yz} = |x\rangle\langle x|.$$  \hspace{1cm} (26)

Taking the diagonal element of Eq. (26) 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. (2), it can be written as

$$\hat{M}_{yz} = \left\{ \begin{array}{l}
\sum_x e^{i\theta(x;y,z)}\sqrt{p(y,z|x)}|x\rangle\langle x|,
\int dx e^{i\theta(x;y,z)}\sqrt{p(y,z|x)}|x\rangle\langle x|, 
\end{array} \right.$$  \hspace{1cm} (27)

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. (27) we have

$$E^{y\dagger}_y(|x\rangle\langle x|) = \sum_z \hat{M}_{yz}^\dagger|x\rangle\langle x|\hat{M}_{yz} = p(y|x)|x\rangle\langle x|,$$  \hspace{1cm} (28)

which ensures the condition (15) with

$$\tilde{x}(x;y) = x,$$
$$q(x;y) = p(x|y).$$
Thus from Theorem 2 the relative-entropy conservation law (16) holds.

The relative entropy conservation relation in Eq. (16) in the QND measurement can be understood in a classical manner as follows. Let us consider a change in the \( x \)-distribution function from \( P^X_{\tilde{\rho}}(x) \) to \( P^X_{\tilde{\rho}_y}(x) \). In the QND measurement, by using Eq. (28), Eq. (8) becomes

\[
P^X_{\tilde{\rho}_y}(x) = \frac{p(y|x)P^X_{\tilde{\rho}}(x)}{P^Y_{\tilde{\rho}}(y)}.
\] (29)

Note that the commutativity of \(|x\rangle\langle x|\) and \(\hat{M}_{yz} \) is essential in deriving Eq. (29). Then Eq. (29) can be interpreted as Bayes’ rule for the conditional probability of \( x \) under measurement outcome \( y \). Since the QND measurement does not disturb the system’s 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 \( y \) based on Bayes’ rule in Eq. (29). 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. (16) from Bayes’ rule in Eq. (29), 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 back-action, yet the relative-entropy conservation law still holds.

### B. Photon-counting measurement

The photon-counting measurement described in Refs. [18–20] 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:

\[
\hat{M}_0(dt) = \hat{I} - \left( i\omega + \frac{\gamma}{2} \right) \hat{n} dt,
\] (30)

\[
\hat{M}_1(dt) = \sqrt{\gamma dt} \hat{a},
\] (31)

where \( \omega \) 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. (30) is called the no-count process in which there is no photocount, while the event corresponding to Eq. (31) is called the one-count process in which a photocount is registered. In the one-count process, the post-measurement 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 the infinitesimal time interval in Eqs. (30) and (31), we can derive an effective measurement operator for a finite time interval $[0, t]$ as follows (cf. Eq. (29) in Ref. [19]):

$$\hat{M}_m(t) = \sqrt{(1 - e^{-\gamma t})^m e^{-(i\omega + \frac{\gamma}{2})t \hat{n}}} \hat{a}^m,$$  \hspace{1cm} (32)

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. (32) can be written as

$$\hat{M}_m^\dagger(t) \hat{M}_m(t) = p(m|\hat{n}; t),$$  \hspace{1cm} (33)

where

$$p(m|n; t) = \binom{n}{m} (1 - e^{-\gamma t})^m e^{-\gamma t(n-m)}.$$  \hspace{1cm} (34)

Equation (33) 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. (34) 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. (32), we obtain

$$\hat{M}_m^\dagger(t)|n\rangle\langle n|\hat{M}_m(t) = q(n; m; t)|\tilde{n}(n; m)\rangle\langle \tilde{n}(n; m)|,$$  \hspace{1cm} (35)

$$\tilde{n}(n; m) = n + m,$$  \hspace{1cm} (36)

$$q(n; m; t) = p(m|m + n; t).$$  \hspace{1cm} (37)

Equation (36) can be interpreted as the photon number of the pre-measurement state when the number of photocounts is $m$ and the photon number remaining in the post-measurement
state is \( n \). From Eqs. (35)-(37), the condition (15) for Theorem 2 is satisfied and we have the relative entropy conservation relation for the photon-counting measurement as

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

where \( p_\rho(m; t) = \text{tr}[\hat{\rho}\hat{M}_m^\dagger(t)\hat{M}_m(t)] \) is the probability distribution of the number of photo-counts \( m \).

### C. Balanced homodyne measurement

The balanced homodyne measurement [21–23] measures one of the quadrature amplitudes of a photon field \( \hat{a} \) in a destructive manner such that the system’s 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:

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

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

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

where \( \gamma \) 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

\[
(d\xi(t))^2 = dt.
\]

The reference measure \( \mu_0(\xi(\cdot)) \) 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. (38), the ensemble average of the outcome \( d\xi(t) \) for the system’s state \( \hat{\rho}(t) \) at time \( t \) is given by

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

where \( \langle \hat{A} \rangle_{\hat{\rho}} := \text{tr}[\hat{\rho}\hat{A}] \). Equation (40) 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. [24, 25].
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 [22]

$$|\psi(t)\rangle = \hat{M}_y(t)|\psi_0\rangle,$$

where

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

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

Note that $\hat{a}^2$ term should be included in the exponent on the right-hand side of Eq. (42) to be consistent with the Itô rule given in Eq. (39). We also mention that the measurement operator in Eq. (42) 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. (41) 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 $|x\rangle_{11} \langle x|$ of the quadrature amplitude operator $\hat{X}_1$, where $|x\rangle_1$ satisfies

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

Then, the operator $\hat{M}_y(t)|x\rangle_{11} \langle y| \hat{M}_y(t)$ and the POVM for the measurement outcome $y(t)$ are evaluated to be (see Appendix A for derivation)

$$\hat{M}^\dagger_y(t)|x\rangle_{11} \langle y| \hat{M}_y(t) = q(x; y(t); t) |\bar{x}(x; y(t); t)\rangle_{11} \langle \tilde{x}(x; y(t); t)|,$$

$$q(x; y(t); t) = \exp \left( e^{-\frac{\gamma t}{2}} x + \frac{y(t)}{\sqrt{2}} \right)^2 - x^2, \quad \bar{x}(x; y(t); t) = e^{-\frac{\gamma t}{2}} x + \frac{y(t)}{\sqrt{2}},$$

$$\mu_0(dy)\hat{M}^\dagger_y(t)\hat{M}_y(t) = \frac{dy}{\sqrt{2\pi e^{-\gamma t}(1 - e^{-\gamma t})}} \exp \left[ -\frac{(y - \sqrt{2}(1 - e^{-\gamma t})\hat{X}_1)^2}{2e^{-\gamma t}(1 - e^{-\gamma t})} \right],$$

where the arguments of $\bar{x}(x; y)$ in Eq. (45) are 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. (46) 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 (44) indicates that the condition (15) for Theorem 2 is satisfied, and we obtain
the relative-entropy conservation law

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

where

\[
P^Y_\rho(y; t)dy = \text{tr}[\hat{\rho}\hat{M}_y(t)\hat{M}_y(t)\dagger]\mu_0(dy)
\]
is the probability distribution function of the measurement outcome \( y(t) \) which is computed
from the POVM in Eq. (46), \( \hat{\rho}_{y(t)}(t) \) 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 \( \hat{X}_1 \).

D. Heterodyne measurement

The heterodyne measurement simultaneously measures the two non-commuting quadra-
ture amplitudes \( \hat{X}_1 \) and \( \hat{X}_2 \) in a destructive manner as in the homodyne measurement. One
way of the 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 correspond
to the two quadrature amplitudes [23].

The measurement operator for the heterodyne measurement in an infinitesimal time in-
terval \( dt \) is given by

\[
\hat{M}(d\zeta(t); dt) = \hat{I} - \frac{\gamma}{2}\hat{n}dt + \sqrt{\gamma}d\zeta(t), \tag{47}
\]

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

\[
(d\zeta(t))^2 = (d\zeta^*(t))^2 = 0, \quad d\zeta(t)d\zeta^*(t) = dt. \tag{48}
\]

As in the homodyne measurement, we consider the time evolution in the interaction picture.
The reference measure \( \mu_0 \) for the measurement outcome \( \zeta(\cdot) \) is the complex Wiener measure.
in which real and imaginary parts of $d\zeta(\cdot)$ are statistically independent Gaussian variables with 0 mean and second order moments consistent with the complex Itô rules in Eq. (48).

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

$$|\psi(t + dt)\rangle = \hat{M}(dt; d\zeta(t))|\psi(t)\rangle.$$  \hspace{1cm} (49)

The solution of Eq. (49) for the initial condition $|\psi_0\rangle$ at $t = 0$ is given by [22]

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

where

$$\hat{M}_{y(t)}(t) = e^{-\frac{\gamma}{2}a^2}e^{y(t)a^\dagger},$$  \hspace{1cm} (50)$$y(t) = \sqrt{\gamma} \int_0^t e^{-\frac{\gamma}{2}d\zeta(s)}.$$  \hspace{1cm} (51)

Here the measurement operator in Eq. (50) does not involve the $a^2$ term unlike the case of the homodyne measurement in Eq. (42) because $(d\zeta(t))^2$ vanishes in this case.

Let us evaluate the POVM for the measurement outcome $y(t)$ in Eq. (51). From Eq. (50), we have

$$\hat{M}_{y(t)}^\dagger(t)\hat{M}_{y(t)}(t) = \mathcal{A}\{\exp[\gamma t - (e^{\gamma t} - 1)\hat{a}\hat{a}^\dagger + e^{\gamma t}(y(t)\hat{a} + y^*(t)\hat{a}^\dagger) - e^{\gamma t}|y(t)|^2]\},$$  \hspace{1cm} (52)

where $\mathcal{A}\{f(\hat{a},\hat{a}^\dagger)\}$ 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 $\hat{M}_{y(t)}^\dagger(t)\hat{M}_{y(t)}(t)$ 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. (51) 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

$$\mu_0(dy(t)) = \frac{e^{-\frac{|y(t)|^2}{2(1-e^{-\gamma t})}}}{\pi(1-e^{-\gamma t})}d^2y(t),$$  \hspace{1cm} (53)

where $d^2y = d(\text{Re}y)d(\text{Im}y)$. From Eqs. (52) and (53), the POVM for $y(t)$ is given by

$$d^2y(t)\mathcal{A}\{p(y(t)|\hat{a},\hat{a}^\dagger;t)\},$$
where
\[ p(y(t)|\alpha, \alpha^*; t) = \exp \left[ -\frac{|y(t)-(1-e^{-\gamma t})\alpha^*/2|^2}{\pi e^{-\gamma t}(1-e^{-\gamma t})} \right]. \] (54)

The probability distribution of the outcome \( y(t) \) when the system is prepared in \( \hat{\rho}_0 \) at \( t = 0 \) is given by
\[ P^Y_{\hat{\rho}_0}(y; t) = \int d^2\alpha p(y(t)|\alpha, \alpha^*; t)Q_{\hat{\rho}_0}(\alpha, \alpha^*), \] (55)
where \( Q_{\hat{\rho}}(\alpha, \alpha^*) := \langle \alpha | \hat{\rho} | \alpha \rangle / \pi \) is the Q-function \([26, 27]\), and \( |\alpha\rangle \) is a coherent state \([28]\) defined by
\[ |\alpha\rangle = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}|0\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle. \] (56)

From Eq. (54), in the infinite-time limit \( t \to \infty \), the probability distribution of outcomes in Eq. (55) reduces to \( Q_{\hat{\rho}_0}(y^*, y) \). Thus the heterodyne measurement actually measures the non-commuting quadrature amplitudes simultaneously in the sense that the probability distribution of outcomes is the Q-function of the initial state \([29]\).

As a reference POVM, we take
\[ d^2\alpha \hat{E}_\alpha = \frac{d^2\alpha}{\pi} |\alpha\rangle \langle \alpha| \] (56)
which generates the Q-function of the density operator. From Eqs. (50) and (56) we have
\[ \mu_0(dy) \hat{M}_y^\dagger(t) \hat{E}_\alpha \hat{M}_y(t) = d^2y(t)q(\alpha, \alpha^*; y)\hat{E}_{\hat{\alpha}(\alpha, y)}, \] (57)
where
\[ \hat{\alpha}(\alpha, y) = e^{-\frac{y^2}{4}}\alpha + y^*, \] (58)
\[ q(\alpha, \alpha^*; y) = \frac{e^{-|\alpha|^2} - |\hat{\alpha}(\alpha, y)|^2 - \frac{|\alpha|^2}{4 e^{-\gamma t}}}{\pi(1 - e^{-\gamma t})}. \] (59)

Note that the inferred quadrature amplitude in Eq. (58) allows a similar interpretation given in the homodyne analysis.

Equation (57) ensures the condition in Eq. (5). From Eqs. (54), (58) and (59), for an arbitrary smooth function \( F(\alpha, \alpha^*) \), we have
\[ \int d^2\alpha q(\alpha, \alpha^*; y)F(\hat{\alpha}(\alpha; y), \hat{\alpha}^*(\alpha; y)) \]
\[ = \int d^2\alpha (e^{\frac{y^2}{4}})q(e^{\frac{y^2}{4}}(\hat{\alpha} + y^*), e^{\frac{y^2}{4}}(\hat{\alpha}^* + y); y)F(\hat{\alpha}, \hat{\alpha}^*) \]
\[ = \int d^2\alpha p(y|\alpha, \alpha^*; t)F(\alpha, \alpha^*). \]
Thus, the condition (6) for Theorem 1 is satisfied and the relative-entropy conservation law

\[ D(P_{\hat{\rho}_0}(\cdot; t) || P_{\hat{\sigma}_0}(\cdot; t)) = D_Q(\hat{\rho}_0 || \hat{\sigma}_0) - E_{\hat{\rho}_0}[D_Q(\hat{\rho}_{y(t)} || \hat{\sigma}_{y(t)})] \]

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

\[ D_Q(\hat{\rho} || \hat{\sigma}) = \int d^2 \alpha \frac{Q_{\hat{\rho}}(\alpha, \alpha^*)}{Q_{\hat{\sigma}}(\alpha, \alpha^*)} \ln \left( \frac{Q_{\hat{\rho}}(\alpha, \alpha^*)}{Q_{\hat{\sigma}}(\alpha, \alpha^*)} \right). \] (60)

Since the Q-function has the complete quantum information about the quantum state, the Q-function relative entropy in Eq. (60) vanishes if and only if \( \hat{\rho} = \hat{\sigma} \), which is not the case in the diagonal relative entropies in the preceding examples. (For example, the diagonal relative entropy of \( \hat{\rho} \) and \( \hat{\sigma} \) in Eqs. (20) and (21) is zero but \( \hat{\rho} \neq \hat{\sigma} \).) Still the Q-function relative entropy is bounded from above by the quantum relative entropy \( S(\hat{\rho} || \hat{\sigma}) := \text{tr} [\hat{\rho} (\ln \hat{\rho} - \ln \hat{\sigma})] \), for the relative entropy of probability distributions on the measurement outcome of a POVM is always smaller than the quantum relative entropy [30].

**IV. SUMMARY**

In this paper we have examined the information flow concerning the hypothesis testing problem of the two quantum states with respect to a reference observable of the system and established the general condition for the relative-entropy conservation law for the general reference POVM (Theorem 1) and PVM (Theorem 2). We have also investigated the case in which the labels of the reference PVM and the measurement outcome are both discrete and shown that the condition for the CP maps of the measurement in Theorem 2 is equivalent to the relative-entropy conservation law for arbitrary states (Theorem 3). We have applied the general theorems to the typical quantum measurements. In the QND measurement, the relative entropy conservation relation can be understood as a result of the classical Bayes’ rule which is the mathematical expression of the modification of our knowledge based on the outcome of the measurement. In the next three examples, namely photon-counting, balanced homodyne and heterodyne measurements, a photon field is measured destructively and the measurement outcomes convey information about the photon number, one and both quadrature amplitude(s), respectively. In spite of the destructive nature of the measurements, the general theorems are still applicable and we have shown that
the relative-entropy conservation laws hold for these measurements. Especially in the heterodyne measurement the reference POVM generates the Q-function and is not an ordinary PVM, reflecting the fact that the non-commuting observables are measured simultaneously. These examples of destructive measurements suggest that the information conservation in a quantum measurement is a much wider concept than the quantum non-demolition principle.

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 MEXT of Japan. Y. K. acknowledges support by Advanced Leading Graduate Course for Photon Science (ALPS) at the University of Tokyo.

Appendix A: Derivations of Eqs. (44) and (46)

To evaluate the operator \( \hat{M}_{y(t)}(t) |x\rangle_1 \langle x| \hat{M}_{y(t)}(t) \), 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

\[
|\alpha\rangle_1 = \pi^{-1/4} \exp \left[ -\frac{1}{2} (x - \sqrt{2}\alpha)^2 - \frac{1}{2} (\alpha^2 + |\alpha|^2) \right],
\]

we have

\[
\langle \alpha | x \rangle_1 \langle x | \alpha \rangle = \pi^{-1/2} \exp \left[ - \left( x - \frac{\alpha + \alpha^*}{\sqrt{2}} \right)^2 \right],
\]

which implies the following normally ordered expression

\[
|x\rangle_1 \langle x| = \pi^{-1/2} : \exp \left[ - \left( x - \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}} \right)^2 \right] :.
\]

By using Eq. (A1) and a formula

\[
e^{-\lambda \hat{a}} |\alpha\rangle = e^{-|\alpha|^2/2 (1-e^{-2\lambda})} e^{-\lambda \alpha},
\]

\[
e^{-\lambda \hat{a}^\dagger} |\alpha\rangle = e^{-|\alpha|^2/2 (1-e^{2\lambda})} e^{-\lambda \alpha^*},
\]

\[
e^{-\lambda \hat{a}^\dagger} |\alpha\rangle = e^{-|\alpha|^2/2 (1-e^{-2\lambda})} e^{-\lambda \alpha^*},
\]

\[
e^{-\lambda \hat{a}} |\alpha\rangle = e^{-|\alpha|^2/2 (1-e^{2\lambda})} e^{-\lambda \alpha}.
\]
which is valid for real \( \lambda \), 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

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

(A2)

Substituting Eq. (A1) in Eq. (A2), we obtain Eq. (44). By integrating Eq. (44) 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

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

(A3)

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 \( \xi(\cdot) \) obeys a Wiener distribution. Here \( y(t) \) in Eq. (43) under a Wiener measure \( \mu_0 \) 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 \( E_0[\cdot] \) denotes the expectation with respect to the Wiener measure. Thus \( \mu_0(dy(t)) \) is given by

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

(A4)

Multiplying Eq. (A3) by Eq. (A4), we obtain Eq. (46).

[1] W. Heisenberg, Z. Phys. 43, 172 (1927).
[2] J. Von Neumann, Mathematische Grundlagen der Quantenmechanik (Berlin: Springer, 1932).
[3] E. Davies and J. Lewis, Commun. Math. Phys. 17, 239 (1970).
[4] K. Kraus, Ann. Phys. (NY) 64, 311 (1971).
[5] K. Kraus, States, Effects, and Operations (Berlin: Springer, 1983).
[6] M. Ozawa, J. Math. Phys. 25, 79 (1984).

[7] C. E. Shannon, Bell Sys. Tech. J. 27, 379 (1948); 27, 623 (1948).

[8] S. Kullback and R. A. Leibler, Ann. Math. Stat. 22, 79 (1951).

[9] T. M. Cover and J. A. Thomas, Elements of information theory (John Wiley & Sons, 2012).

[10] R. A. Fisher, Philos. T. Roy. Soc. A 222, 309 (1922).

[11] M. Ban, Phys. Lett. A 235, 209 (1997).

[12] M. Ban, Int. Jour. Theor. Phys. 37, 2491 (1998).

[13] M. Ban, J. Phys. A: Math. Gen. 32, 1643 (1999).

[14] M. Ban, J. Phys. A: Math. Gen. 32, 6527 (1999).

[15] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, Rev. Mod. Phys. 52, 341 (1980).

[16] V. B. Braginsky and F. Y. Khalili, Quantum measurement (Cambridge University Press, 1995).

[17] V. B. Braginsky and F. Y. Khalili, Rev. Mod. Phys. 68, 1 (1996).

[18] M. Srinivas and E. Davies, Opt. Acta 28, 981 (1981).

[19] M. Ueda, Quantum Opt.: Journal of the European Optical Society Part B 1, 131 (1989).

[20] N. Imoto, M. Ueda, and T. Ogawa, Phys. Rev. A 41, 4127 (1990).

[21] H. M. Wiseman and G. J. Milburn, Phys. Rev. A 47, 642 (1993).

[22] H. M. Wiseman, Quantum Semiclass. Opt. 8, 205 (1996).

[23] H. M. Wiseman and G. J. Milburn, Quantum Measurement and Control (2010).

[24] H. Wiseman and L. Diósi, Chem. Phys. 268, 91 (2001).

[25] A. Barchielli and M. Gregoratti, Quantum trajectories and measurements in continuous time: the diffusive case, Vol. 782 (Springer, 2009).

[26] K. Husimi, Proc. Phys.-Math. Soc. Japan. 3rd Series 22, 264 (1940).

[27] C. L. Mehta and E. C. G. Sudarshan, Phys. Rev. 138, B274 (1965).

[28] R. J. Glauber, Phys. Rev. 131, 2766 (1963).

[29] H. Carmichael, Statistical Methods in Quantum Optics 2: Non-Classical Fields, Vol. 2 (Springer, 2008) Chap. 18.3.2.

[30] M. Hayashi, Quantum information: An introduction (Springer, 2006).