学位論文

Relative-entropy conservation law in quantum measurement and its applications to continuous measurements
（量子測定における相対エントロピーの保存則とその連続測定への応用）

平成２６年１２月博士（理学）申請

東京大学大学院理学系研究科
物理学専攻
倉持　結

Abstract

In a general quantum measurement, some amount of information on the measured observable corresponding to a system's positive-operator valued measure (POVM) is lost due to inevitable state change of the measurement. However, there is a special class of quantum measurements in which the information about the measured observable is conserved. Quantum non-demolition measurement and photon-counting measurement on a single-mode field are examples of quantum measurements in which the information about system's photon number conserves. Ban discussed the information flows in quantum measurement processes based on the Shannon entropy (M. Ban, Int. Jour. Theor. Phys. **37**, 2491 (1998)). He quantified the obtained information as the mutual information between the measurement outcome and the measured observable, and established a condition for a Shannon entropy conservation which states that the mutual information is equal to the average decrease in the system's Shannon entropy of the measured observable. However, since the Shannon entropy for a continuous variable cannot be interpreted as an information content, there exist several continuous quantum measurements that do not satisfy the Shannon entropy conservation. Furthermore the physical meaning of the condition for the Shannon entropy conservation derived by Ban is not clear. In this thesis, we consider the information flow quantitatively by using the relative entropy and establish the condition for the relative-entropy conservation.

First, we quantify the information carried by the measurement outcome in terms of the relative entropy between the probability measures of two candidate states and establish a sufficient condition for the relative-entropy conservation law which states that the relative entropy of the measurement outcome is equal to a decrease in the relative entropy of the measured observable in the system. The statistical meaning of the condition is clarified by considering a successive joint measurement process of the measurement process followed by a sharp measurement of the observable. In this joint measurement process, the condition can be interpreted as the existence of a sufficient statistic whose probability distribution coincides with that of the measured observable. The condition for the relative-entropy conservation law is less restrictive than that for the Shannon entropy conservation and we compare these conditions in the case in which both the measurement outcome and the measured observable are discrete.

Second we apply the general theory on the relative-entropy conservation law to typical optical continuous measurements, namely photon-counting, quantum-counting, homodyne measurement, and heterodyne measurement. We show that

the Shannon entropy conservation does not hold except for the case of the photon-counting measurement, while the relative-entropy conservation does hold for all of these measurements. The breakdown of the Shannon entropy conservation is shown to be due to the non-unit Jacobian of the sufficient statistic and the strong dependence of the continuous Shannon entropy on the reference measure.

Finally, we consider a problem of whether or not there exists a relative-entropy-conserving POVM of the system for a given measurement process. Assuming that the sample space of the measurement process is a standard Borel space, we construct a relative-entropy-conserving POVM of the system. Physically, the constructed POVM corresponds to an infinite successive joint measurements of the given measurement process.

Contents

Chapter 1

Introduction

The inevitable back-action on the system in a quantum measurement process is one of the key concepts in quantum theory and much has been discussed about the relation between the obtained information and the measurement back-action. Though a quantum measurement has a back-action on the system, some of the measurement process is used to measure a system's observable X corresponding to a system's positive-operator valued measure (POVM) and known to bring us an information about X by a sequence of the measurements. Examples of such measurements include a quantum non-demolition (QND) measurement [1, 2, 3, 4] and a photon-counting measurement [5, 6, 7].

The QND measurement is a measurement in which there is not a measurement back-action on a measured observable corresponding to a system's projection-valued measure (PVM) X. Note that this does not imply that there is no measurement back-action on the system since an observable conjugate to X is actually disturbed. By performing the same QND measurement on the system many times, we can obtain the information about X as much as we can. On the other hand, the photon-counting measurement is a destructive measurement such that the number of photons in the system decreases in a counting event. Still we can obtain the information about the photon number $\hat{n}$ by performing the measurement continuously.

What is common in these measurement processes is a conservation of the information about the measured observable, in the above examples the projection-valued measure of the photon number operator. Ban [8, 9, 10] considered the information conservation of the measured POVM X in a quantum measurement process Y described by a completely positive (CP) instrument quantitatively by using the Shannon entropy and mutual information. In Ref. [9, 10], he showed that under some conditions on the measurement process the Shannon entropy conservation law

$$I(X:Y) = H_{\hat{\rho}}(X) - E[H_{\hat{\rho}_y}(X)] \tag{1.1}$$

holds. Here the $I(X:Y)$ is the mutual information between a system's POVM X and the measurement outcome Y, $H_{\hat{\rho}}(X)$ is the Shannon entropy of X for the pre-measurement state $\hat{\rho}$, $E[\cdot]$ denotes the ensemble average over the measurement outcome y and $H_{\hat{\rho}_y}(X)$ is the Shannon entropy of X for a given post-measurement state $\hat{\rho}_y$ when the measurement outcome is y. The left-hand side of Eq. (1.1) is the amount of the obtained information about X from the measurement outcome Y and the right-hand side is the decrease in the Shannon entropy of X by the measurement back-action. It can be shown that both QND and photon-counting measurements satisfy the established condition and the Shannon entropy conservation (1.1) in these measurements.

The Shannon entropy for continuous variable X is defined by

$$H(X) = \int_{\Omega_X} \mu_0(dx) p^X(x) \ln p^X(x), \tag{1.2}$$

where Ω_X is the sample space of X, μ_0 is a measure on Ω_X and $p^X(x)$ is the probability density of X with respect to μ_0. The continuous Shannon entropy (1.2) depends strongly on the reference measure μ_0 and is known to change its value by a one-to-one transformation of the stochastic variable. Due to this fact we cannot interpret the continuous Shannon entropy (1.2) as an information content. Such a difficulty in some measurement processes with continuous outcome suggests that the left-hand side of Eq. (1.1) does not correspond to a decrease in the system's information. Furthermore, the physical meaning of the condition for the Shannon entropy conservation is not so clear.

Another important information content is the relative entropy [11], or the Kullback-Leibler divergence, defined by

$$D(p^X||q^X) = \int_{\Omega_X} \mu_0(dx) p^X(x) \ln \left(\frac{p^X(x)}{q^X(x)} \right), \tag{1.3}$$

where p^X and q^X are probability density functions with respect to a reference measure μ_0. The relative entropy is shown to be positive and the expression (1.3) does not depend on the choice of the reference measure μ_0. Statistically the relative entropy is regarded as the amount of information obtained from the measurement outcome x about which of the probability measures p and q is actually prepared. The relative entropy is also known to characterize the sufficiency of a system's stochastic variable $T(x)$ on the parameter estimation problem. Here in the parameter estimation problem we consider a set of probability measures

$\{p_\theta^X\}_{\theta\in\Theta}$ parametrized by θ and estimate from the measurement outcome x which θ is actually chosen. The sufficiency of a stochastic variable $T(x)$ is defined by the condition that the conditional probability $p_\theta(x|T(x)=t)$ does not depend on θ for any t. This implies that it is sufficient to know $T(x)$ to obtain the information about θ. The sufficiency is characterized by the conservation of the relative entropies between X and $T(x)$.

In this thesis we address the problem of the information conservation in quantum measurements from the different standpoint, that is, the relative entropy. In our approach, the obtained information is quantified as the relative entropy $D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}})$, where $p^Y_{\hat{\rho}}$ is the probability measure of the measurement outcome when the state is prepared in $\hat{\rho}$ and σ is another candidate state. By assuming some conditions on the measurement process Y and a system observable X, we will establish the following relative-entropy conservation law

$$D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}}) = D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}) - E[D(p^X_{\hat{\rho}_y}||p^X_{\hat{\sigma}_y})], \tag{1.4}$$

where $p^X_{\hat{\rho}}$ is the probability distribution of X for a state $\hat{\rho}$. The condition for the relative-entropy conservation law we found is well understood if we consider a joint successive measurement process in which Y is first measured and then a sharp measurement on X is done. Then the relative-entropy conservation law (1.4) is shown to be equivalent to

$$D(\tilde{p}^{XY}_{\hat{\rho}}||\tilde{p}^{XY}_{\hat{\sigma}}) = D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}), \tag{1.5}$$

where $\tilde{p}^{XY}_{\hat{\rho}}$ is the probability distribution of the joint measurement process for a pre-measurement state $\hat{\rho}$. In this joint measurement process, the condition for the relative entropy conservation (1.4) or (1.5) can be interpreted as a condition that there exists a sufficient statistic $\tilde{x}(x;y)$ such that the probability distribution of $\tilde{x}(x;y)$ coincides with that of X for the pre-measurement state. This condition is logically less restrictive than the condition for the Shannon entropy conservation, i.e. the relative-entropy conservation law applies to a wider class of quantum measurements. In examples of homodyne and heterodyne measurements in which the system's observable X is continuous, it is shown that the relative-entropy conservation law holds, while the Shannon entropy conservation does not. This is due to the difficulties in the definition of the Shannon entropy of a continuous variable. We also show that for a given measurement process Y, a system's observable X corresponding to the infinite joint measurement of Y satisfies the relative-entropy conservation.

This thesis is organized as follows. In Chapter 2 we review quantum measurement theory with continuous sample space. In Chapter 3, we review the classical entropic information contents especially for the continuous case. In Chapter 4, we review the Shannon entropy conservation by Ban. In Chapter 5 we prove the relative entropy conservation law for a measurement of system's observable described by a general positive operator-valued measure (Theorem 5.1.1) and projection-valued measure (Theorem 5.2.1). For the case in which Y is a pure discrete measurement, it is shown that the relative-entropy conservation and the condition for Theorem 5.2.1 is logically equivalent. Furthermore, the condition for the Shannon entropy conservation is compared with that for the relative entropy when X and Y are discrete and X is projection-valued. In Chapter 6, we apply the general discussion about the relative-entropy conservation law to optical destructive measurements, namely photon counting, quantum counting, homodyne measurement, and heterodyne measurement. In these examples except the photon-counting measurement, we show the Shannon entropy conservation law does not hold due to the difficulties in the Shannon entropy for continuous variable, while the relative entropy conservation law does hold. In Chapter 7 we construct a relative-entropy-conserving observable X for a given measurement process Y. In Chapter 8, we summarize this thesis.

The results in Chapter 5 and Chapter 6 are based on Ref. [12] collaborating with M. Ueda.

Chapter 2

Review on Quantum Measurements

In this chapter we will review general theory of quantum measurements which is needed in the main part of this thesis.

2.1 Positive Operator-valued Measure

A positive operator-valued measure (POVM) describes the statistics of the measurement outcome of a quantum measurement. To treat quantum measurements with continuous and discrete sample spaces on a equal footing, we introduce POVM on a general σ-algebra [13].

Definition 2.1.1 (σ-algebra)
Let Ω be a set. A σ-algebra $\mathscr{B}$ on Ω is a family of subsets of Ω such that

1. $\Omega \in \mathscr{B}$;
2. if $A \in \mathscr{B}$ then $\Omega \setminus A \in \mathscr{B}$, where $X \setminus Y := \{x \in X; x \notin Y\}$ is the difference set;
3. for a countable sequence of sets $\{A_n\}_{n=1}^{\infty} \subset \mathscr{B}$, $\bigcup_{n=1}^{\infty} A_n \in \mathscr{B}$.

A pair $(\Omega, \mathscr{B})$ is said to be a **measurable space**, or a **sample space**, iff $\mathscr{B}$ is a σ-algebra on Ω.

Definition 2.1.2 (positive operator-valued measure)
Let $\mathcal{H}$ be a Hilber space, $\mathcal{L}(\mathcal{H})$ be the set of bouded operators on $\mathcal{H}$, Ω be a set and $\mathscr{B}$ be a σ-algebra on Ω. A mapping $\hat{E} : \mathscr{B} \to \mathcal{L}(\mathcal{H})$ is called as a **positive operator-valued measure** (**POVM**) iff

1. $\hat{E}(A) \geq 0$ for all $A \in \mathscr{B}$

2. $\hat{E}(\emptyset) = 0$;

3. $\hat{E}(\Omega) = \hat{I}$;

4. $\hat{E}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \hat{E}(A_n)$ (in the weak sense) for any disjoint $\{A_n\} \subset \mathscr{B}$, where a generalized sequence $\hat{A}_\alpha$ in $\mathcal{L}(\mathcal{H})$ converges weakly to $\hat{A} \in \mathcal{L}(\mathcal{H})$ iff $\langle\psi| \hat{A}_\alpha |\phi\rangle \to \langle\psi| \hat{A} |\phi\rangle$ for all $|\psi\rangle, |\phi\rangle \in \mathcal{H}$.

For the POVM $\hat{E}$, Ω is said to be a sample space of $\hat{E}$ and the measurable space $(\Omega, \mathscr{B})$ is said to be an outcome space of $\hat{E}$. We also refer the triad $\mathsf{E} = (\hat{E}, \Omega, \mathscr{B})$ as the POVM or the observable on $\mathcal{H}$.

Let $\hat{E} : \mathscr{B} \to \mathcal{L}(\mathcal{H})$ be a POVM on a sample space $(\Omega, \mathscr{B})$ and $\hat{\rho}$ be a density operator on $\mathcal{H}$, i.e. a positive operator with unit trace. Then the mapping

$$P_{\hat{\rho}} : \mathscr{B} \ni A \mapsto \mathrm{tr}[\hat{\rho}\hat{E}(A)] \in [0, 1]$$

is a probability measure on a measurable space $(\Omega, \mathscr{B})$. $P_{\hat{\rho}}(A)$ can be interpreted as the probability for the event that the measurement outcome ω is in A. Each element A of $\mathscr{B}$ is assumed to be an event such that the probability for A can be defined.

Next we show some examples of the POVM.

Example 1
Let Ω be a countable set and $\hat{E}$ be a POVM on $(\Omega, 2^\Omega)$, where 2^Ω is the power set of Ω. Then for any $A \in 2^\Omega$

$$\hat{E}(A) = \sum_{\omega \in A} \hat{E}_\omega, \tag{2.1}$$

where $\hat{E}_\omega := \hat{E}(\{\omega\})$. From $\hat{E}(\Omega) = \hat{I}$, the set of positive operators $\{\hat{E}_\omega\}_{\omega\in\Omega}$ satisfies the following completeness condition:

$$\sum_{\omega\in\Omega} \hat{E}_\omega = \hat{I}. \tag{2.2}$$

On the other hand, if a set of positive operators $\{\hat{E}_\omega\}_{\omega\in\Omega}$ satisfies the completeness condition (2.2), a POVM on a sample space $(\Omega, 2^\Omega)$ can be defined by Eq. (2.1). For this reason, the set of operators $\{\hat{E}_\omega\}_{\omega\in\Omega}$ is identified with the POVM itself. This kind of POVM is called discrete.

Example 2
A POVM $\hat{E}$ on a sample space $(\Omega, \mathscr{B})$ is a **projection valued measure** (**PVM**) iff $\hat{E}(A)$ is a projection operator for any $A \in \mathscr{B}$. For a PVM $\hat{E}$, we can show that

$$\hat{E}(A)\hat{E}(B) = \hat{E}(A \cap B) \tag{2.3}$$

for any $A, B \in \mathscr{B}$. To prove Eq. (2.3), we first note that if $X, Y \in \mathscr{B}$ and $X \subset Y$, then

$$\hat{E}(X) \leq \hat{E}(X) + \hat{E}(Y \setminus X) = \hat{E}(Y). \tag{2.4}$$

Since $\hat{E}(X)$ and $\hat{E}(Y)$ are projection operators, Eq. (2.4) implies

$$\hat{E}(Y)\hat{E}(X) = \hat{E}(X). \tag{2.5}$$

Therefore we obtain

$$\begin{aligned} \hat{E}(A) + \hat{E}(A \cap B) &= \hat{E}(A)\hat{E}(A \cup B) + \hat{E}(A)\hat{E}(A \cap B) \\ &= \hat{E}(A)\left(\hat{E}(A) + \hat{E}(B)\right) \\ &= \hat{E}(A) + \hat{E}(A)\hat{E}(B), \end{aligned}$$

and thus Eq. (2.3) is proved.

If $\{|i\rangle\}_{i \in I}$ is a complete orthonormal basis of the Hilbert space $\mathcal{H}$, then $\{|i\rangle \langle i|\}_{i \in I}$ is a discrete PVM on $\mathcal{H}$.

Another important example is the spectral decomposition $\hat{E}$ of a position operator $\hat{x}$ of a one-dimensional particle. In this case, the Hilber space $\mathcal{H}$ is given by $L^2(\mathbb{R})$ which is the set of all the square-integrable complex measurable functions on $\mathbb{R}$. The sample space for $\hat{E}$ is given by $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, where $\mathscr{B}(\mathbb{R})$, called as the Borel σ-algebra of $\mathbb{R}$, is the smallest σ-algebra which contains the family of open sets. The PVM $\hat{E}$ is then given by

$$\hat{E}(A)\psi(x) := 1_A(x)\psi(x)$$

for all $A \in \mathscr{B}(\mathbb{R})$, $\psi \in \mathcal{H}$, and $x \in \mathbb{R}$, where

$$1_A(x) = \begin{cases} 1 \text{ if } x \in A; \\ 0 \text{ if } x \notin A \end{cases}$$

is the indicator function. By using the Dirac ket $|x\rangle$ such that

$$\langle x|x'\rangle = \delta(x - x'), \quad \int_{\mathbb{R}} dx \, |x\rangle \langle x| = \hat{I}, \tag{2.6}$$

the PVM $\hat{E}$ can be written as

$$\hat{E}(A) = \int_A dx \, |x\rangle \langle x| \, .$$

Example 3 (POVM density)
Let $(\hat{E}, \Omega_Y, \mathscr{B}_Y)$ be a POVM on $\mathcal{H}$. For many cases we can take a positive measure μ_0 on a measurable space $(\Omega_Y, \mathscr{B}_Y)$ and a positive operator-valued fuction $\hat{E}_y^Y$ $(y \in \Omega_Y)$ such that the POVM $\hat{E}$ can be written as

$$\hat{E}(A) = \int_A \mu_0(dy) \hat{E}_y^Y \tag{2.7}$$

for any $A \in \mathscr{B}_Y$. $\hat{E}_y^Y$ is said to be a POVM density with respect to a reference measure μ_0. We will encounter many examples of the POVM density in the following chapters.

2.2 Completely Positive (CP) Instruments and Measurement Models

In the quantum measurement, the state change due to the measurement back-action is essential. To describe the measurement back-action, we need a mathematical framework which can derive both the measurement outcome and the state change. Such a description is given by a completely positive instrument discussed in this section. We can also consider a more detailed description of the measurement which includes an interaction involved in the measurement and a pointer observable of the probe system. The relation of these descriptions are also discussed in this section.

2.2.1 CP maps and Kraus representations

Definition 2.2.1 (completely positive super-operator)
Let $\mathcal{E} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}\ (\mathcal{K})$ be a linear map. (A linear map which acts on a space of operators is often called as a superoperator). $\mathcal{E}$ is **positive** iff $\mathcal{E}(A) \geq 0$ for any positive $A \in \mathcal{L}(\mathcal{H})$. $\mathcal{E}$ is **completely positive** (**CP**) iff $\mathcal{E} \otimes \mathcal{I}_n$ is positive for all $n \geq 1$, where $\mathcal{I}_n$ is the identity superoperator on $\mathcal{L}(\mathbb{C}^n)$ and the tensor product $\mathcal{E} \otimes \mathcal{F}$ of superoperators $\mathcal{E}$ and $\mathcal{F}$ is defined by $\mathcal{E} \otimes \mathcal{F}(\hat{A} \otimes \hat{B}) := \mathcal{E}(\hat{A}) \otimes \mathcal{F}(\hat{B})$ for all $\hat{A}$ and $\hat{B}$ in the domains of $\mathcal{E}$ and $\mathcal{F}$, respectively.

The complete positivity states that the state change is positive if another ancilla system $\mathbb{C}^n$ is present. It is known that there exists a positive but not completely positive superoperator [14].

One important property of the CP map is that it has a Kraus representation shown in the following theorem due to Kraus [15]:

Theorem 2.2.2
Let $\mathcal{H}$ and $\mathcal{K}$ be separable, i.e. countable dimesional, Hilbert spaces and $\mathcal{E}$: $\mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ be a CP map. Then there exists a countable set of bounded linear operators $\hat{M}_i : \mathcal{H} \to \mathcal{K}$ $(i = 1, 2, \cdots)$ such that

$$\mathcal{E}(\hat{\rho}) = \sum_i \hat{M}_i \hat{\rho} \hat{M}_i^\dagger, \tag{2.8}$$

for any $\hat{\rho} \in \mathcal{L}(\mathcal{H})$. Furthermore if both $\mathcal{H}$ and $\mathcal{K}$ are finite dimensional, then the number of operators $\hat{M}_i$ can be taken to be finite.

The representation of $\mathcal{E}$ in Eq. (2.8) is said to be a Kraus representation of $\mathcal{E}$ and the operator $\hat{M}_i$ is said to be a Kraus operator.

For a given CP map $\mathcal{E}$, there is a arbitrariness in the choice of the Kraus representation in Eq. (2.8). In the context of the measurement model, this corresponds to the choice of the measured observable of the probe system with discrete outcomes.

2.2.2 CP instruments

Now we introduce a CP instrument:

Definition 2.2.3 (CP instrument)
Let $(\Omega, \mathscr{B})$ be a sample space, $\mathcal{H}$ be a Hilber space. A mapping

$$\mathscr{B} \times \mathcal{T}(\mathcal{H}) \ni (A, \hat{\rho}) \longmapsto \mathcal{I}_A(\hat{\rho}) \in \mathcal{T}(\mathcal{H}),$$

where $\mathcal{T}(\mathcal{H})$ is the set of trace class operators on $\mathcal{H}$, is called a **CP instrument** iff

1. for any $A \in \mathscr{B}$, a mapping $\mathcal{T}(\mathcal{H}) \ni \hat{\rho} \longmapsto \mathcal{I}_A(\hat{\rho}) \in \mathcal{T}(\mathcal{H})$ is CP and trace-decreasing superoperator;

2. for any state $\hat{\rho}$, $\mathrm{tr}[\mathcal{I}_\Omega(\hat{\rho})] = 1$;

3. if $\hat{\rho}$ is a state and $\{A_n\}_{n=1}^\infty \subset \mathscr{B}$ is disjoint, then

$$\mathrm{tr}\left[\mathcal{I}_{\bigcup_{n=1}^\infty A_n}(\hat{\rho})\hat{B}\right] = \sum_{n=1}^\infty \mathrm{tr}\left[\mathcal{I}_{A_n}(\hat{\rho})\hat{B}\right]$$

   for any operator $\hat{B} \in \mathcal{L}(\mathcal{H})$.

A CP instrument $\mathcal{I}_\cdot(\cdot)$ on a sample space $(\Omega, \mathscr{B})$ determines the statistics of the measurement outcome by

$$P_{\hat{\rho}}(A) = \mathrm{tr}[\mathcal{I}_A(\hat{\rho})] = \mathrm{tr}[\hat{\rho}\hat{E}(A)],$$

where $\hat{E}(A) := \mathcal{I}_A^\dagger(\hat{I})$ 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 any $\hat{\rho} \in \mathcal{T}(\mathcal{H})$ and any $\hat{A} \in \mathcal{L}(\mathcal{H})$. From Def. 2.2.3 $\hat{E}(\cdot)$ is the POVM in Def. 2.1.2. Furthermore, for a pre-measurement state $\hat{\rho}$ and an event $A \in \mathscr{B}$ such that $P_{\hat{\rho}}(A) \neq 0$, the post-measurement state $\hat{\rho}_A$ when the outcome ω is in A is given by

$$\hat{\rho}_A = \frac{\mathcal{I}_A(\hat{\rho})}{P_{\hat{\rho}}(A)}. \tag{2.9}$$

Especially $\hat{\rho}_\Omega = \mathcal{I}_\Omega(\hat{\rho})$ is the non-selective post-measurement state in which all the information of the measurement outcome is discarded. The non-selective state change is described by the CP map $\mathcal{I}_\Omega(\cdot)$ which is trace-preserving (TP). We remark that for a given POVM $\hat{E}$ there exist many CP instruments whose POVM's give the same $\hat{E}$. The difference between these CP instruments corresponds to non-equivalent measurement back-actions.

Next we show some examples of CP instruments.

Example 4 (discrete CP instrument)
Let Ω be a countable set and $\mathcal{I}_\cdot(\cdot)$ be a CP instrument on a sample space $(\Omega, 2^\Omega)$. A CP instrument on a discrete sample space is said to be discrete. We adopt a notation $\mathcal{I}_\omega := \mathcal{I}_{\{\omega\}}$ and the measurement process is completely determined by a set of CP maps $\{\mathcal{I}_\omega\}_{\omega\in\Omega}$ by

$$\mathcal{I}_A = \sum_{\omega\in A} \mathcal{I}_\omega$$

for any $A \in 2^\Omega$.

If for each $\omega \in \Omega$ there exists a bounded operator $\hat{M}_\omega$ such that

$$\mathcal{I}_\omega(\hat{\rho}) = \hat{M}_\omega \hat{\rho} \hat{M}_\omega^\dagger,$$

the measurement is said to be pure. In this case, if the pre-measurement state is a pure state, the post-measurement state is also pure.

Example 5
Let $\mathcal{I}_\cdot(\cdot)$ be a CP instrument on a sample space $(\Omega_Y, \mathscr{B}_Y)$ and μ_0 be a measure on $(\Omega_Y, \mathscr{B}_Y)$. As in the case of the POVM density, for some cases we can take a density of the CP instrument $\mathcal{E}_\omega$ for each $\omega \in \Omega$ such that

$$\mathcal{I}_A(\hat{\rho}) = \int_A \mu_0(d\omega)\mathcal{E}_\omega(\hat{\rho})$$

for any set $A \in \mathscr{B}_Y$ and any state $\hat{\rho}$. In this case, the POVM $\hat{E}(A) = \mathcal{I}_A^\dagger(\hat{I})$ has the density of POVM given by $\hat{E}_\omega = \mathcal{E}_\omega^\dagger(\hat{I})$. $\hat{\rho}_\omega := \mathcal{E}_\omega(\hat{\rho}) / \mathrm{tr}[\mathcal{E}_\omega(\hat{\rho})]$ can be interpreted as the post-measurement state when the measurement outcome is ω.

In the theorem of the relative entropy conservation relation, we assume the existence of such reference measure μ_0 and the density of CP instrument $\mathcal{E}_\omega$.

2.2.3 Indirect measurement model

In this subsection we consider a measurement model which describes the probe system and its interaction with the system together with the measurement outcome and the state change. Formally the indirect measurement model is defined as follows.

Definition 2.2.4
Let $\mathcal{H}$ be a Hilbert space, and $\mathsf{E} = (\hat{E}, \Omega, \mathscr{B})$ be a POVM on $\mathcal{H}$. Suppose that

1. $\mathcal{K}$ is a Hilbert space called a probe system;
2. $\hat{\sigma}$ is a state on $\mathcal{K}$;
3. $\hat{U}$ is a unitary operator on a composite system $\mathcal{H} \otimes \mathcal{K}$;
4. $\mathsf{F} = (\hat{F}, \Omega, \mathscr{B})$ is a POVM on $\mathcal{K}$.

A quadruple $\langle \mathcal{K}, \hat{\sigma}, \hat{U}, \mathsf{F} \rangle$ is called an indirect measurement model, or a measurement model, of E iff

$$\mathrm{tr}[\hat{\rho}\hat{E}(A)] = \mathrm{tr}[\hat{U}(\hat{\rho} \otimes \hat{\sigma})\hat{U}^\dagger \hat{F}(A)]$$

for any state $\hat{\rho}$ and $A \in \mathscr{B}$.

An indirect measurement model $\langle \mathcal{K}, \hat{\sigma}, \hat{U}, \mathsf{F} \rangle$ of a POVM $\mathsf{E} = (\hat{E}, \Omega, \mathscr{B})$ determines the CP instrument on $\mathcal{H}$ by

$$\mathcal{I}_A^{\mathcal{M}}(\hat{\rho}) = \mathrm{tr}_{\mathcal{K}}[\hat{U}(\hat{\rho} \otimes \hat{\sigma})\hat{U}^\dagger (\hat{I} \otimes \hat{F}(A))]$$

for any state $\hat{\rho}$ and $A \in \mathscr{B}$, where $\mathrm{tr}_{\mathcal{K}}$ is the partial trace over the probe Hilbert space $\mathcal{K}$.

Ozawa [16] has shown that for an arbitrary instrument $\mathcal{I}_{\cdot}(\cdot)$ there exists a measurement model $\mathcal{M} = \langle \mathcal{K}, \hat{\sigma}, \hat{U}, \mathsf{F} \rangle$ such that $\mathcal{I} = \mathcal{I}^{\mathcal{M}}$. In this measurement model $\mathcal{M}$, σ can be taken to be a pure state and F be a PVM on $\mathcal{K}$. However, this measurement model $\mathcal{M}$ is constructed in an abstract manner based solely on the mathmatical structure of the instrument $\mathcal{I}_{\cdot}(\cdot)$ and it is hard to relate the constructed $\mathcal{M}$ to any real measurement apparatus. Therefore it is still an interasting question whether or not we can implement a quantum measurement corresponding to a given CP instrument in a realistic situation to which, for example, the probe system $\mathcal{K}$ and the interaction Hamiltonian $\hat{H}_I$ are restricted.

Chapter 3

Classical Entropies

In this chapter we review the properties of information contents needed for the main part of this thesis. We especially investigate the case in which the stochastic variable is continuous. It is pointed out that the Shannon entropy has an arbitrariness in the choice of a reference measure, wheres the relative entropy and the mutual information do not. In the last section, we discuss about the sufficiency of a statistical variable and its relation to the relative entropy.

3.1 Shannon entropy

We first introduce the Shannon entropy [17, 18].

Definition 3.1.1 (discrete Shannon entropy)
Let Ω_X be a discrete sample space and $p^X(x)$ $(x \in \Omega_X)$ be a probability on Ω_X. The Shannon entropy $H(p)$ is defined by

$$H(p^X) := -\sum_{x\in\Omega_X} p^X(x) \ln p^X(x) \tag{3.1}$$

In this thesis, the natural logarithm is adopted in the definitions of entropic information contents and these entropies are expressed in nats. The term $0 \ln 0$, if exists in Eq. (3.1), is understood to be 0 which is obtained from $\lim_{p\to+0} p \ln p$.

The Shannon entropy is interpreted as the randomness of the variable X. For the general reference of the classical information theory, we refer the reader to Ref. [19].

Let us now consider the generalization of the Shannon entropy to a continuous stochastic variable X. For simplicity, we assume that the sample space of X is the real line $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, where $\mathscr{B}(\mathbb{R})$ is the Borel σ-algebra of $\mathbb{R}$ defined in Example 2, and the probability measure of X can be written as $p^X(x)dx$, where $p^X(x)$

is a continuous positive function and dx is the Lebesgue measure. We divide the sample space $\mathbb{R}$ by intervals $I_n(\delta) = (\delta n, \delta(n+1)]$, where $\delta > 0$ and n is an integer. From the continuity of the density function $p^X(x)$, there exists $x_n \in I_n(\delta)$ for each n such that

$$p^X(x_n)\delta = \int_{I_n(\delta)} p^X(x)dx,$$

which can be interpreted as the probability $p_n(\delta)$ of the event that X is in $I_n(\delta)$. As the discretized version of the continuous X, we introduce a stochastic variable

$$X_\delta = x_n \text{ if } X \in I_n(\delta).$$

Then the Shannon entropy for X_δ is given by

$$\begin{aligned} H(p^{X_\delta}) &= -\sum_n p_n(\delta) \ln p_n(\delta) \\ &= -\sum_n \delta p^X(x_n) \ln p^X(x_n) - \ln \delta. \end{aligned} \tag{3.2}$$

In the limit $\delta \to +0$, the first term in Eq. (3.2) converges to a finite value

$$-\int_{\mathbb{R}} dx p^X(x) \ln p^X(x) \tag{3.3}$$

and the second term $-\ln \delta$ diverges to ∞. Thus the simple generalization of the discrete Shannon entropy can be divided to a finite part converging to Eq. (3.2) and the divergent part. As the generalization of Eq. (3.2), we define the following Shannon entropy for continuous observable:

Definition 3.1.2 (differential entropy)
Let $(\Omega_X, \mathscr{B}_X, P^X)$ be a probability space of a stochastic variable X and μ_0 be a measure on $(\Omega_X, \mathscr{B}_X)$. Suppose that P^X has a density function $f(x)$ with respect to μ_0, i.e. $P^X(dx) = p^X(x)\mu_0(dx)$. The differential entropy of X with respect to μ_0 is defined by

$$h(p^X) := -\int_{\Omega_X} \mu_0(dx) p^X(x) \ln p^X(x). \tag{3.4}$$

The differential entropy in Eq. (3.4) is not necessarily positive and strongly depends on the choice of the reference measure μ_0. Furthermore, the value of the differential entropy changes by the trasformation of the variable as the following discussion shows. To be specific, let us again consider the case when X is a stochastic variable on the real line and the probability density of X is $p^X(x)$. We define another stochastic variable Y by

$$Y = y(X),$$

where $y(x)$ is a differentiable one-to-one mapping from $\mathbb{R}$ to $\mathbb{R}$. Then the probability distribution of Y and the differential entropy for Y are given by

$$p^Y(y(x)) = p^X(x)\left(\frac{dy(x)}{dx}\right)^{-1}$$

and

$$\begin{aligned} h(p^Y) &= -\int_{\mathbb{R}} dy p^Y(y) \ln p^Y(y) \\ &= -\int_{\mathbb{R}} dx p^X(x) \ln \left[p^X(x) \left(\frac{dy(x)}{dx}\right)^{-1}\right] \\ &= h(p^X) + \int_{\mathbb{R}} dx p^X(x) \ln\left(\frac{dy(x)}{dx}\right), \end{aligned} \tag{3.5}$$

respectively. The differential entropy for Y does not coincide with that of X in Eq. (3.3) unless the last term in Eq. (3.5) vanishes.

Furthermore a difference between differential entropies for two probability densities p^X and q^X also changes its value by the transformation $x \to y(x)$. In fact, by using Eq. (3.5), we have

$$\begin{aligned} &h(p^Y) - h(q^Y) \\ &= h(p^X) - h(q^X) + \int_{\mathbb{R}} dx p^X(x) \ln\left(\frac{dy(x)}{dx}\right) - \int_{\mathbb{R}} dx q^X(x) \ln\left(\frac{dy(x)}{dx}\right), \end{aligned}$$

which does not in general coincide with $h(p^X) - h(q^X)$.

Due to this fact, it is hard to interpret the differential entropy as an information content of X since the transformation from X to Y is just the change of the label and we neither lose nor obtain any information by this transformation.

3.2 Relative entropy and mutual information

In this section we introduce the relative entropy and show its relation with the mutual information. The relative entropy was first introduced by Kullback and Leibler [11] for general probability measures.

Definition 3.2.1 (relative entropy)
Let P^X and Q^X be probability measures on a sample space $(\Omega_X, \mathscr{B}_X)$. Suppose that P^X and Q^X have the density functions $p^X(x)$ and $q^X(x)$ with respect to a reference measure μ_0 on $(\Omega_X, \mathscr{B}_X)$. Then the relative entropy, or the Kullback-Leibler divergence, of P^X and Q^X is defined by

$$D(P^X||Q^X) := \int_{\Omega_X} \mu_0(dx) p^X(x) \ln \left(\frac{p^X(x)}{q^X(x)} \right). \tag{3.6}$$

We also write $D(P^X||Q^X)$ as a function $D(p^X||q^X)$ of the two density functions p^X and q^X if the reference measure is clear from the context.

We remark that we can take $P^X + Q^X$ as the reference measure μ_0 with respct to which the density functions of P^X and Q^X exist. We also remark that the definition of the relative entropy in Eq. (3.6) does not depend on the choice of the reference measure μ_0. To see the independence, let ν_0 be another reference measure and $p'^X(x)$, $q'^X(x)$ be the density functions of P^X and Q^X with respect to ν_0, respectively. Then we have

$$\begin{aligned} P^X(dx) &= p^X(x)\mu_0(dx) = p'^X(x)\nu_0(dx) \\ Q^X(dx) &= q^X(x)\mu_0(dx) = q'^X(x)\nu_0(dx) \end{aligned} \tag{3.7}$$

and therefore

$$\frac{p^X(x)}{q^X(x)} = \frac{p'^X(x)}{q'^X(x)}. \tag{3.8}$$

From Eqs. (3.7) and (3.8) the right-hand side of Eq. (3.6) can be written as

$$\int_{\Omega_X} \nu_0(dx) p'^X(x) \ln \left(\frac{p'^X(x)}{q'^X(x)} \right),$$

which shows that the relative entropy in Eq. (3.6) is independent of the choice of the reference measure.

Next, we will show some basic properties of the relative entropy.

Proposition 3.2.2 (information inequality)
Let p^X and q^X be probability density functions on a measure space $(\Omega_X, \mathscr{B}_X, \mu_0)$. Then,

$$D(p^X||q^X) \geq 0, \tag{3.9}$$

with the equality if and only if $p^X(x) = q^X(x)$ for almost all x with respect to μ_0.

Proof. Let A be a suppor of p^X, i.e. $A = \{x \in \Omega_X; p^X(x) \neq 0\}$. Then we have

$$\begin{aligned} -D(p^X||q^X) &= \int_A \mu_0(dx) p^X(x) \ln \frac{q^X(x)}{p^X(x)} \\ &\leq \ln \int_A \mu_0(dx) p^X(x) \frac{q^X(x)}{p^X(x)} \\ &= \ln \int_A q^X(x) \\ &\leq \ln 1 = 0. \end{aligned} \tag{3.10}$$

In the inequality (3.10), we have used the Jensen's inequality [20]

$$E_P[f(X)] \leq f\left(E_P[X]\right), \tag{3.11}$$

where f is a concave function, X is a stochastic variable and $E_P(\cdot)$ is the ensemble average over the measure P. The equality holds iff the equality in (3.10) holds, which occurs iff $p^X(x) = q^X(x)$ for almost all x. □

Let us next consider the two stochastic variables X and Y, the sample spaces of which are $(\Omega_X, \mathscr{B}_X)$ and $(\Omega_Y, \mathscr{B}_Y)$, respectively. Let $p^{XY}(x,y)$ and $q^{XY}(x,y)$ be the joint probability density functions of X and Y with respect to a product measure $\mu_0(dx)\nu_0(dy)$. The marginal distribution functions and the conditional distribution functions are defined as

$$\begin{aligned} p^X(x) &= \int_{\Omega_Y} \nu_0(dy) p^{XY}(x,y), \\ p^Y(y) &= \int_{\Omega_X} \mu_0(dx) p^{XY}(x,y), \\ p^{Y|X}(y|x) &= \frac{p^{XY}(x,y)}{p^X(x)}. \end{aligned}$$

q^X, q^Y and $q^{Y|X}$ are defined in a similar manner. Then we have the following chain rule for the relative entropy:

Proposition 3.2.3 (chain rule for the relative entropy)

$$D(p^{XY}||q^{XY}) = D(p^X||q^X) + E_{p^X}[D(p^{Y|X}(\cdot|x)||q^{Y|X}(\cdot|x))] \tag{3.12}$$

Proof.

$$\begin{aligned}
D(p^{XY}||q^{XY}) &= \int_{\Omega_X\times\Omega_Y} \mu_0(dx)\nu_0(dy)p^{XY}(x,y)\ln\left(\frac{p^{XY}(x,y)}{q^{XY}(x,y)}\right) \\
&= \int_{\Omega_X\times\Omega_Y} \mu_0(dx)\nu_0(dy)p^{XY}(x,y)\ln\left(\frac{p^{X}(x)p^{Y|X}(y|x)}{q^{X}(x)q^{Y|X}(y|x)}\right) \\
&= \int_{\Omega_X} \mu_0(dx)p^{X}(x)\ln\left(\frac{p^{X}(x)}{q^{X}(x)}\right) \\
&\quad + \int_{\Omega_X\times\Omega_Y} \mu_0(dx)\nu_0(dy)p^{XY}(x,y)\ln\left(\frac{p^{Y|X}(y|x)}{q^{Y|X}(y|x)}\right) \\
&= D(p^X||q^X) + E_{p^X}[D(p^{Y|X}(\cdot|x)||q^{Y|X}(\cdot|x))].
\end{aligned}$$

□

Equation (3.12) indicates that the joint information of X and Y is divided into the information of X and that of Y conditioned on X. The conditioning by X is due to the statistical correlation between X and Y. Such correlation can be quantified by the following mutual information.

Definition 3.2.4 (mutual information)
Let X, Y, p^{XY} be the same as in the Proposition 3.2.3. Then the mutual information of X and Y with respect to the joint probability distribution p^{XY} is defined by

$$I(X:Y) := D(p^{XY}||p^Xp^Y) = \int_{\Omega_X\times\Omega_Y} \mu_0(dx)\nu_0(dy)p^{XY}(x,y)\ln\left(\frac{p^{XY}(x,y)}{p^X(x)p^Y(y)}\right). \tag{3.13}$$

We note that the mutual information in Eq. (3.13) is independent of the choice of the reference measures μ_0 and ν_0 since it is defined through the relative entropy. From Proposition 3.2.2, we immediately obtain the following proposition:

Proposition 3.2.5

$$I(X:Y) \geq 0 \tag{3.14}$$

with the equality iff X and Y are statistically independent.

3.3 Sufficient statistics and the relative entropy

Definition 3.3.1
Let $(\Omega_X, \mathscr{B}_X)$ and $(\Omega_Y, \mathscr{B}_Y)$ be sample spaces and $T : \Omega_X \to \Omega_Y$ be a measurable map, i.e. $T^{-1}(A) \in \mathscr{B}_X$ for any $A \in \mathscr{B}_Y$. A set of probability measures $\mathcal{P} = \{p_\theta\}_{\theta\in\Theta}$ on $(\Omega_X, \mathscr{B}_X)$ is called a statistical model on $(\Omega_X, \mathscr{B}_X)$. For each $p_\theta \in \mathcal{P}$, we can define a probability measure $p_\theta T^{-1}$ on $(\Omega_Y, \mathscr{B}_Y)$ by

$$p_\theta T^{-1}(A) := p_\theta(T^{-1}(A))$$

for each $A \in \mathscr{B}_Y$. Then, T is said to be **sufficient** for a statistical model $\mathcal{P}$ iff for each $A \in \mathscr{B}_X$, there exists a function $p(E|y)$ on $(\Omega_Y, \mathscr{B}_Y)$ such that for any $p_\theta \in \mathcal{P}$

$$p_\theta(E|y) = p(E|y) \quad p_\theta T^{-1}\text{-a.e. } y \tag{3.15}$$

where $p_\theta(E|y)$ is the conditional probability measure of p_θ when $T(x) = y$. Here a condition depending on y holds μ-almost every y, μ-a.e. y in short, when the set of y which does not satisfy the condition is a μ-null set.

The condition (3.15) means that the information about the probability model $\mathcal{P}$ is completely determined by $T(x)$.

Let μ and ν be finite measures on a sample space $(\Omega, \mathscr{B})$. ν is said to be absolutely continuous with respect to μ, denoted as $\nu \ll \mu$, iff for any $A \in \mathscr{B}$ such that $\mu(A) = 0$, $\nu(A) = 0$. From the Radon-Nikodym theorem [13], if $\nu \ll \mu$ then there exists a nonnegative measurable function f such that

$$\nu(A) = \int_A f(\omega)\mu(d\omega)$$

for all $A \in \mathscr{B}$. The function f is unique up to μ-a.e. in the sense that if g is another μ-integrable function such that $\nu(A) = \int_A g d\mu$ for each $A \in \mathscr{B}$ then $f(\omega) = g(\omega)$ for μ-a.e. $\omega \in \Omega$. The function f is called Radon-Nikodym derivative and denoted as $d\nu/d\mu$.

A probability model $\mathcal{P} = \{p_\theta\}_{\theta\in\Theta}$ on a sample space $(\Omega_X, \mathscr{B}_X)$ is said to be dominated iff there exists a finite measure λ on $(\Omega_X, \mathscr{B}_X)$ such that for any $p_\theta \in \mathcal{P}$, $p_\theta \ll \lambda$. The last statement is shortly denoted as $\mathcal{P} \ll \lambda$.

The following theorem is due to Halmos and Savage [21].

Theorem 3.3.2
Let T be a measurable map from $(\Omega_X, \mathscr{B}_X)$ onto $(\Omega_Y, \mathscr{B}_Y)$ and let $\mathcal{P} = \{p_\theta\}_{\theta\in\Theta}$ be a probability model on $(\Omega_X, \mathscr{B}_X)$ dominated by a finite measure λ. Then the

necessary and sufficient condition that T be sufficient for $\mathcal{P}$ is that, for each $p_\theta \in \mathcal{P}$, $f_\theta := dp_\theta/d\lambda$ can be factorized as

$$f_\theta(x) = g_\theta(T(x))t(x) \quad p_\theta\text{-a.e. } x, \tag{3.16}$$

where g_θ is a measurable positive function on Ω_Y, t is a measurable positive function on Ω_X, $g_\theta \cdot t$ is λ-integrable, and $t = 0$ λ-a.e. on an arbitrary p_θ-null set.

Kullback and Leibler [11] showed that the sufficiency of T can be characterized by the conservation of the relative entropy as shown by the following theorem.

Theorem 3.3.3
Let T be statistic from $(\Omega_X, \mathscr{B}_X)$ to $(\Omega_Y, \mathscr{B}_Y)$, and $\mathcal{P} = \{p_\theta\}_{\theta\in\Theta} \ll \lambda$ be a dominated statistical model. Then

$$D(p_\theta || p_{\theta'}) \geq D(p_\theta T^{-1} || p_{\theta'} T^{-1}) \tag{3.17}$$

for each $p_\theta, p_{\theta'} \in \mathcal{P}$. The equality in (3.17) holds for each $p_\theta, p_{\theta'} \in \mathcal{P}$ iff T is sufficient for $\mathcal{P}$.

The above concepts of the statistical model and the sufficient statistic can be applied to the quantum measurement as follows. For given POVM $(\Omega_X, \mathscr{B}_X, \hat{E}^X)$ on a Hilber space $\mathcal{H}$ and a set of quantum states $\mathcal{M} \subset \mathcal{S}(\mathcal{H})$, where $\mathcal{S}(\mathcal{H}) := \{\hat{\rho} \in \mathcal{L}(\mathcal{H}); \hat{\rho} \geq 0, \operatorname{tr}\hat{\rho} = 1\}$ is the state space on $\mathcal{H}$, a set $\{P^X_{\hat{\rho}}\}_{\hat{\rho}\in\mathcal{M}}$ is a statistical model on $(\Omega_X, \mathscr{B}_X)$, where $P^X_{\hat{\rho}}(A) := \operatorname{tr}[\hat{\rho}\hat{E}^X(A)]$. If the Hilbert space $\mathcal{H}$ is separable, we can show that $\{P^X_{\hat{\rho}}\}_{\hat{\rho}\in\mathcal{M}}$ is dominated by a probability measure and Thereom 3.3.2 and Theorem 3.3.3 are applicable in this measurement theoretical setup.

Proposition 3.3.4
Let $\mathcal{H}$ be a separable Hilbert space and let $(\Omega_X, \mathscr{B}_X, \hat{E}^X)$ be a POVM on $\mathcal{L}(\mathcal{H})$. Then there exists a quantum state $\hat{\rho}_0 \in \mathcal{S}(\mathcal{H})$ such that $\{P^X_{\hat{\rho}}\}_{\hat{\rho}\in\mathcal{S}(\mathcal{H})} \ll P^X_{\hat{\rho}_0}$.

The proof is given in Appendix A.

Chapter 4

Shannon Entropy Conservation Law in Quantum Measurement

In this chapter, we review the Shannon entropy conservation law in quantum measurement shown by Ban [8, 9, 10]. Under some conditions on a quantum measurement, Ban showed [9, 10] the Shannon entropy conservation law which states that a decrease in the Shannon entropy of a system's observable is equal to the mutual information established between the observable and the measurement outcome. Here, based basically on Ref. [10], we reconstruct Ban's discussion. As an example of the conservation law, we consider a quantum non-demolition measurement.

4.1 Shannon entropy conservation for POVM

Let us consider an observable X on a Hilbert space $\mathcal{H}$ and a quantum measurement Y. X is assumed to be represented by a density $\hat{E}_x^X$ of a POVM on a sample space $(\Omega_X, \mathscr{B}_X)$ with respect to a reference measure $\nu_0(dx)$ (cf. Example 3). The probability density of X when the system is prepared in a state $\hat{\rho}$ is

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

The quantum measurement Y corresponds to a density $\mathcal{E}_y^Y(\cdot)$ of a CP instrument on a sample space $(\Omega_Y, \mathscr{B}_Y)$ with respect to a reference measure $\mu_0(dy)$ (cf. Example 5). The densities of the POVM and the probability for the measurement outcome are given by

$$\hat{E}_y^Y = \mathcal{E}_y^{Y\dagger}(\hat{I}),$$
$$p_{\hat{\rho}}^Y(y) = \mathrm{tr}[\hat{\rho}\hat{E}_y^Y],$$

and the post-measurement state for a given measurement outcome y is

$$\hat{\rho}_y = \frac{\mathcal{E}_y^Y(\hat{\rho})}{p_{\hat{\rho}}^Y(y)}.$$

To consider a relation between X and Y we assume that the density of the POVM for Y can be written as

$$\hat{E}_y^Y = \int_{\Omega_X} \nu_0(dx) p(y|x) \hat{E}_x^X, \tag{4.1}$$

where $p(y|x)$ is a conditional probability function such that it is positive and satisfies the normalization condition

$$\int_{\Omega_Y} \mu_0(dy) p(y|x) = 1.$$

By taking the quantum expectation of Eq. (4.1) with respect to a state $\hat{\rho}$, we obtain

$$p_{\hat{\rho}}^Y(y) = \int_{\Omega_X} \nu_0(dx) p(y|x) p_{\hat{\rho}}^X(x). \tag{4.2}$$

The condition (4.1), or equivalently (4.2), implies that the measurement outcome of Y is the coarse-graining of that of X. From Eq. (4.2), we can define the mutual information between X and Y by

$$\begin{aligned} I_{\hat{\rho}}(X:Y) &= \int_{\Omega_X} \nu_0(dx) \int_{\Omega_Y} \mu_0(dy) p(y|x) p_{\hat{\rho}}^X(x) \ln\left(\frac{p(y|x) p_{\hat{\rho}}^X(x)}{p_{\hat{\rho}}^Y(y) p_{\hat{\rho}}^X(x)}\right) \\ &= \int_{\Omega_X} \nu_0(dx) \int_{\Omega_Y} \mu_0(dy) p(y|x) p_{\hat{\rho}}^X(x) \ln\left(\frac{p(y|x)}{p_{\hat{\rho}}^Y(y)}\right), \end{aligned} \tag{4.3}$$

which is the information obtained from the measurement outcome of Y about the distribution of X.

To establish the Shannon entropy conservation law, we further impose the following two conditions on the measurement:

1. There exists a function $\tilde{x}(x;y)$ such that for any x and y

$$\mathcal{E}_y^{Y\dagger}(\hat{E}_x^X) = p(y|\tilde{x}(x;y)) \hat{E}_{\tilde{x}(x;y)}^X; \tag{4.4}$$

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

$$\int_{\Omega_X} \nu_0(dx) p(y|\tilde{x}(x;y)) F(\tilde{x}(x;y)) = \int_{\Omega_X} \nu_0(dx) p(y|x) F(x). \tag{4.5}$$

Based on the assumptions (4.1), (4.4) and (4.5), we evaluate the Shannon entropy of X for the post-measurement state. The probability distribution function of X for the post-measurement state $\hat{\rho}_y$ is evaluated to be

$$\begin{aligned} p^X_{\hat{\rho}_y}(x) &= \frac{\mathrm{tr}[\mathcal{E}^Y_y(\hat{\rho})\hat{E}^X_x]}{p^Y_{\hat{\rho}}(y)} \\ &= \frac{\mathrm{tr}[\hat{\rho}\mathcal{E}^{Y\dagger}_y(\hat{E}^X_x)]}{p^Y_{\hat{\rho}}(y)} \\ &= \frac{p(y|\tilde{x}(x;y))p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)}, \end{aligned} \tag{4.6}$$

where we used the condition (4.4) to derive Eq. (4.6). From Eq. (4.6), the Shannon entropy of X for the post-measurement state $\hat{\rho}_y$ is given by

$$\begin{aligned} h(p^X_{\hat{\rho}_y}) &= -\int_{\Omega_X} \nu_0(dx) p^X_{\hat{\rho}_y}(x) \ln p^X_{\hat{\rho}_y}(x) \\ &= -\int_{\Omega_X} \nu_0(dx) \frac{p(y|\tilde{x}(x;y))p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)} \ln \frac{p(y|\tilde{x}(x;y))p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)} \\ &= -\int_{\Omega_X} \nu_0(dx) \frac{p(y|x)p^X_{\hat{\rho}}(x)}{p^Y_{\hat{\rho}}(y)} \ln \frac{p(y|x)p^X_{\hat{\rho}}(x)}{p^Y_{\hat{\rho}}(y)}, \end{aligned} \tag{4.7}$$

where we used the condition (4.5) in deriving the last equality. By taking the ensemble average of Eq. (4.7) over the measurement outcome y, the average Shannon entropy for the post-measurement state is given by

$$\begin{aligned} E_{\hat{\rho}}[h(p^X_{\hat{\rho}_y})] &= \int_{\Omega_Y} \mu_0(dy) p^Y_{\hat{\rho}}(y) h(p^X_{\hat{\rho}_y}) \\ &= -\int_{\Omega_Y} \mu_0(dy) \int_{\Omega_X} \nu_0(dx) p(y|x) p^X_{\hat{\rho}}(x) \ln \frac{p(y|x)p^X_{\hat{\rho}}(x)}{p^Y_{\hat{\rho}}(y)} \\ &= -\int_{\Omega_Y} \mu_0(dy) \int_{\Omega_X} \nu_0(dx) p(y|x) p^X_{\hat{\rho}}(x) \ln \frac{p(y|x)}{p^Y_{\hat{\rho}}(y)} - \int_{\Omega_X} \nu_0(dx) p^X_{\hat{\rho}}(x) \ln p^X_{\hat{\rho}}(x) \\ &= -I_{\hat{\rho}}(X:Y) + h(p^X_{\hat{\rho}}), \end{aligned}$$

where $E_{\hat{\rho}}[\cdot]$ denotes the ensemble average over the measurement outcome y for a given pre-measurement state $\hat{\rho}$. Therefore we obtain the Shannon entropy conservation law

$$I_{\hat{\rho}}(X:Y) = h(p^X_{\hat{\rho}}) - E_{\hat{\rho}}[h(p^X_{\hat{\rho}_y})]. \tag{4.8}$$

The left-hand side of Eq. (4.8) is the amount of the information about X obtained from Y, while the right-hand side is the decrease in the Shannon entropy of X. The above discussion is summarized as the following theorem.

Theorem 4.1.1
Let X be an observable on a Hilber space $\mathcal{H}$ represented by a density $\hat{E}_x^X$ of a POVM on a sample space $(\Omega_X, \mathscr{B}_X)$ with respect to a reference measure ν_0 and Y be a quantum measurement process on $\mathcal{H}$ which is represented by a density $\mathcal{E}_y^Y$ of a CP instrument on a sample space $(\Omega_Y, \mathscr{B}_Y)$ with respect to a reference measure μ_0. Suppose that X and Y satisfy the conditions (4.1), (4.4) and (4.5). Then the Shannon entropy conservation law (4.8) holds.

4.2 Shannon entropy conservation for PVM

The important class of system's observables is that of PVM's. In this section we consider the case when the system's observable X is a PVM. For simplicity, the PVM is assumed to be of the form $|x\rangle \langle x| \, dx$, where $x \in \mathbb{R}$ and $|x\rangle$ is the Dirac ket which satisfies the complete orthonormal condition (2.6). We note that the following discussion is also applicable to a discrete PVM $|x\rangle \langle x|$ with a discrete complete orthonormal condition

$$\langle x|x'\rangle = \delta_{x,x'}, \quad \sum_x |x\rangle \langle x| = \hat{I}, \tag{4.9}$$

by the following formal correspondences given by

$$\int dx \cdots \longleftrightarrow \sum_x \cdots \tag{4.10}$$

and

$$\delta(x - x') \longleftrightarrow \delta_{x,x'}. \tag{4.11}$$

For the PVM $|x\rangle \langle x|$ and a quantum measurement described by a density $\mathcal{E}_y^Y$ of a CP instrument with respect to a reference measure μ_0, we assume the following two conditions corresponding to (4.1) and (4.4):

1. The density of the POVM for the measurement outcome of Y can be written as

$$\hat{E}_y^Y = \int dx p(y|x) \, |x\rangle \langle x| \, ; \tag{4.12}$$

2. There exists a function $\tilde{x}(x; y)$ such that for any y

$$\mathcal{E}_y^{Y\dagger}(|x\rangle \langle x|) = p(y|\tilde{x}(x; y)) \, |\tilde{x}(x; y)\rangle \langle \tilde{x}(x; y)| \, . \tag{4.13}$$

Note that $p(y|x)$ satisfies the normalization condition

$$\int_{\Omega_Y} \mu_0(dy) p(y|x) = 1, \tag{4.14}$$

because of the completeness condition

$$\int \mu_0(dy) \hat{E}_y^Y = \hat{I} \tag{4.15}$$

and the uniqueness of the spectral decomposition with respect to $|x\rangle\langle x|$. By integrating Eq. (4.13) with respect to x, we obtain

$$\begin{aligned} \hat{E}_y^Y &= \int dx' p(y|\tilde{x}(x';y)) \, |\tilde{x}(x';y)\rangle \langle \tilde{x}(x';y)| \\ &= \int dx' \int dx \delta(x - \tilde{x}(x';y)) p(y|x) \, |x\rangle \langle x| \\ &= \int dx \left(\int dx' \delta(x - \tilde{x}(x';y)) \right) p(y|x) \, |x\rangle \langle x| \, . \end{aligned} \tag{4.16}$$

By comparing Eq. (4.16) with Eq. (4.12), we have

$$\int dx' \delta(x - \tilde{x}(x';y)) = 1 \tag{4.17}$$

for any x and y such that $p(y|x) \neq 0$. Then, for any y and a smooth function $F(x)$

$$\begin{aligned} \int dx' p(y|\tilde{x}(x';y)) F(\tilde{x}(x';y)) &= \int dx' \int dx \delta(x - \tilde{x}(x';y)) p(y|\tilde{x}(x';y)) F(\tilde{x}(x';y)) \\ &= \int dx \left(\int dx' \delta(x - \tilde{x}(x';y)) \right) p(y|x) F(x) \\ &= \int dx p(y|x) F(x), \end{aligned} \tag{4.18}$$

where we used Eq. (4.17) in deriving the last equality. Equation (4.18) implies that the condition (4.5) for X and Y is satisfied and therefore, from Theorem 4.1.1, we obtain the Shannon entropy conservation law (4.8). To summarize, we obtain the following theorem.

Theorem 4.2.1
Let X be a PVM of the form $|x\rangle\langle x| \, dx$ on the real line with a complete orthonormal condition or a discrete rank-1 PVM $|x\rangle\langle x|$ with a discrete complete orthonormal condition and let Y be a quantum measurement process corresponding to a density of the CP instrument $\mathcal{E}_y^Y$ on a sample space $(\Omega_Y, \mathscr{B}_Y)$ with respect to a reference measure $\mu_0(dy)$. Suppose that X and Y satisfy the conditions (4.12) and (4.13). Then the Shannon entropy conservation law (4.8) holds.

As an example of the Shannon entropy conservation law, we consider a quantum non-demolition measurement [1, 2, 3, 4].

Example 6 (quantum non-demolition measurement)
Let X be a PVM $|x\rangle\langle x|$ which is either discrete or continuous as in Theorem 4.2.1 and Y be a quantum measurement corresponding to a CP instrument $\mathcal{I}_{dy}^{Y}(\cdot) = \mathcal{E}_y^Y(\cdot)\mu_0(dy)$ on a sample space $(\Omega_Y, \mathscr{B}_Y)$. Y is called quantum non-demolition (QND) measurement of X iff for any $\hat{\rho}$

$$p_{\hat{\rho}}^X(x) = p_{\mathcal{I}_{\Omega_Y}^Y(\hat{\rho})}^X(x), \tag{4.19}$$

where $\mathcal{I}_{\Omega_Y}^Y(\hat{\rho})$ is the non-selective post-measurement state. The QND condition (4.19) states that the probability distribution of X is not disturbed by the measurement back-action of Y. Since Eq. (4.19) can be written as

$$\mathrm{tr}[\hat{\rho}\,|x\rangle\langle x|] = \mathrm{tr}[\hat{\rho}\mathcal{I}_{\Omega_Y}^{Y\,\dagger}(|x\rangle\langle x|)],$$

the QND condition is also expressed as

$$\mathcal{I}_{\Omega_Y}^{Y\,\dagger}(|x\rangle\langle x|) = |x\rangle\langle x|\,. \tag{4.20}$$

Shimizu and Fujita [4] pointed out that the QND condition adopted in some literature, e.g. in Ref. [1], is too strong and they proposed less restrictive QND conditions. Our QND condition (4.19) corresponds to a 'moderate condition' in Ref. [4].

To be definite, we consider the continuous X while the following discussion is still valid for discrete X by the formal correspondence stated in the first paragraph of this section. If we write $\mathcal{E}_y^{Y\,\dagger}$ in the operator-sum form as

$$\mathcal{E}_y^{Y\,\dagger}(\hat{A}) = \sum_k \hat{M}_{yk}^\dagger \hat{A}\hat{M}_{yk},$$

Eq. (4.20) becomes

$$\int_{\Omega_Y} \mu_0(dy) \sum_k \hat{M}_{yk}^\dagger\,|x\rangle\langle x|\,\hat{M}_{yk} = |x\rangle\langle x|\,. \tag{4.21}$$

By taking the expectation value of Eq. (4.21) with respect to an eigenstate $|x'\rangle$ with $x' \neq x$, we obtain

$$\int_{\Omega_Y} \mu_0(dy) \sum_k |\,\langle x|\,\hat{M}_{yk}\,|x'\rangle\,|^2 = 0,$$

which implies that

$$\langle x| \hat{M}_{yk} |x'\rangle = 0 \tag{4.22}$$

for any x and $x' \neq x$. Thus $\hat{M}_{yk}$ is diagonalized in the $|x\rangle$ basis. From the completeness condition (4.15), $\hat{M}_{yk}$, ${\mathcal{E}_y^Y}^\dagger(|x\rangle\langle x|)$ and $\hat{E}_y^Y$ can be written as

$$\hat{M}_{yk} = \sum_x e^{i\theta(x;y,k)} \sqrt{p(y,k|x)}\, |x\rangle\langle x| \,,$$

$${\mathcal{E}_y^Y}^\dagger(|x\rangle\langle x|) = p(y|x)\, |x\rangle\langle x| \,, \tag{4.23}$$

$$\hat{E}_y^Y = \int dx p(y|x)\, |x\rangle\langle x| \,, \tag{4.24}$$

where

$$p(y|x) = \sum_k p(y,k|x) \tag{4.25}$$

and $p(y|x)$ satisfies the normalization condition (4.14). Equations (4.23) and (4.24) ensure the conditions (4.12) and (4.13) for Theorem 4.2.1 with

$$\tilde{x}(x;y) = x \tag{4.26}$$

and we obtain the Shannon entropy conservation law (4.8).

Examples of destructive measurements that satisfy the Shannon entropy conservation law include a photon counting model and a quantum counting model, both of which will be discussed in Chap. 6 where we also consider the relative entropy conservation law derived in the next chapter.

Chapter 5

Relative Entropy Conservation Law in Quantum Measurement

In this chapter, we establish a sufficient condition for the relative entropy conservation law for a quantum measurement process which states that the relative entropy of the measurement outcome between two candidate states is equal to a decrease in the relative entropy of a system's observable. Furthermore, for the case in which the measurement is pure and the observable is a discrete PVM, we show that the sufficient condition amounts to the relative entropy conservation for arbitrary candidate states, i.e. the condition is a necessary and sufficient condition. We also compare our condition with that for the Shannon entropy conservation in Chap. 4 and it is found that our condition is less restrictive. To demonstrate the generality of our condition, we consider a destructive sharp measurement of an observable in which the relative entropy conservation law holds, whereas the Shannon entropy conservation law does not.

5.1 Relative entropy conservation for POVM

In this section, we prove the relative entropy conservation law under some conditions on the measurement process. As in the previous chapter, we consider an observable X on a Hilbert space $\mathcal{H}$ and a quantum measurement Y corresponding to a density of POVM $\hat{E}_x^X$ on a sample space $(\Omega_X, \mathscr{B}_X)$ with respect to a reference measure ν_0 and a density of a CP instrument $\mathcal{E}_y^Y$ on a sample space $(\Omega_Y, \mathscr{B}_Y)$ with respect to a reference measure μ_0.

Here, as the information content of the measurement outcome, we consider the relative entropy of the measurement outcome of Y given by

$$D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}}) = \int_{\Omega_Y} \mu_0(dy) p^Y_{\hat{\rho}}(y) \ln\left(\frac{p^Y_{\hat{\rho}}(y)}{p^Y_{\hat{\sigma}}(y)}\right), \tag{5.1}$$

where $\hat{\rho}$ and $\hat{\sigma}$ are candidate states. Considering the relative entropy as an information content corresponds to an experimental situation in which the pre-measurement state is prepared in $\hat{\rho}$ and the observer knows a priori that the pre-measurement state is either $\hat{\rho}$ or $\hat{\sigma}$. From the measurement outcome y, the observer infers which the true pre-measurement state is. Although it seems that this formalism is applicable only to the hypothesis testing problem with just two candidate quantum states, it is more general since $\hat{\rho}$ and $\hat{\sigma}$ can be arbitrary states. In this sense, the relative entropy (5.1) is regarded as the amount of the information about how we can distinguish possible quantum states if we consider arbitrary $\hat{\rho}$ and $\hat{\sigma}$.

To establish the relative entropy conservation law, we impose the following conditions on X and Y.

1. The condition (4.1) holds, i.e. the measurement outcome of Y is the coarse-graining of X;

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

$$\mathcal{E}^{Y\dagger}_y(\hat{E}^X_x) = q(x;y)\hat{E}^X_{\tilde{x}(x;y)} \tag{5.2}$$

   for any x and y;

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

$$\int_{\Omega_X} \nu_0(dx) q(x;y) F(\tilde{x}(x;y)) = \int_{\Omega_X} \nu_0(dx) p(y|x) F(x). \tag{5.3}$$

Based on these assumptions, let us show the relative entropy conservation law. The probability distribution function of X for the post-measurement state $\hat{\rho}_y$ is given by

$$p^X_{\hat{\rho}_y}(x) = \frac{q(x;y)p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)}, \tag{5.4}$$

where we used Eq. (5.2). Then, the relative entropy of X for the post-measurement states $\hat{\rho}_y$ and $\hat{\sigma}_y$ is given by

$$\begin{aligned} D(p^X_{\hat{\rho}_y}||p^X_{\hat{\sigma}_y}) &= \int_{\Omega_X} \nu_0(dx) \frac{q(x;y)p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)} \ln\left(\frac{p^Y_{\hat{\sigma}}(y)p^X_{\hat{\rho}}(\tilde{x}(x;y))}{p^Y_{\hat{\rho}}(y)p^X_{\hat{\sigma}}(\tilde{x}(x;y))}\right) \\ &= -\ln\left(\frac{p^Y_{\hat{\rho}}(y)}{p^Y_{\hat{\sigma}}(y)}\right) + \int_{\Omega_X} \nu_0(dx) \frac{p(y|x)p^X_{\hat{\rho}}(x)}{p^Y_{\hat{\rho}}(y)} \ln\left(\frac{p^X_{\hat{\rho}}(x)}{p^X_{\hat{\sigma}}(x)}\right), \end{aligned} \tag{5.5}$$

where we used the condition (5.3) in deriving the last equality. Thus the ensemble average over the measurement outcome y is evaluated to be

$$
\begin{aligned}
E_{\hat{\rho}}[D(p^X_{\hat{\rho}_y}||p^X_{\hat{\sigma}_y})] &= \int_{\Omega_X} \mu_0(dy) p^Y_{\hat{\rho}}(y) D(p^X_{\hat{\rho}_y}||p^X_{\hat{\sigma}_y}) \\
&= -D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}}) + \int_{\Omega_X} \nu_0(dx) \int_{\Omega_Y} \mu_0(dy) p(y|x) p^X_{\hat{\rho}}(x) \ln\left(\frac{p^X_{\hat{\rho}}(x)}{p^X_{\hat{\sigma}}(x)}\right) \\
&= -D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}}) + D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}).
\end{aligned}
$$

Therefore we obtain the relative-entropy conservation law

$$
D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}}) = D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}) - E_{\hat{\rho}}[D(p^X_{\hat{\rho}_y}||p^X_{\hat{\sigma}_y})]. \tag{5.6}
$$

The left-hand side of this equation is the information concerning which state is actually prepared. The right-hand side represents a decrease in the relative entropy of X with respect to the candidate states. We thus obtain the following theorem, which is the main result of this thesis.

Theorem 5.1.1 (relative-entropy conservation law)
Let X be an observable on a Hilbert space $\mathcal{H}$ represented by a density $\hat{E}^X_x$ of a POVM on a sample space $(\Omega_X, \mathscr{B}_X)$ with respect to a reference measure ν_0 and Y be a quantum measurement process on $\mathcal{H}$ which is represented by a density $\mathcal{E}^Y_y$ of a CP instrument on a sample space $(\Omega_Y, \mathscr{B}_Y)$ with respect to a reference measure μ_0. Suppose that X and Y satisfy the conditions (4.1), (5.2) and (5.3). Then the relative-entropy conservation law (5.6) holds.

The conditions for Theorem 5.1.1 is *weaker* than that for Theorem 4.1.1 since $q(x;y)$ in the condition (5.2) does not in general coincide with $p(y|\tilde{x}(x;y))$. While this difference might appear to be a minor modification of the condition at this stage, it will be shown that our relative-entropy conservation law is applicable to a much wider range of quantum measurements in the next chapter.

To further understand the meaning of the conditions (5.2) and (5.3), let us consider a joint measurement process in which measurement of Y is performed and then X is performed on the post-measurement process. We remark that since the concept of a quantum state is equivalent to all the probability distributions for the possible measurements, considering the state change due to Y measurement is equivalent to considering the probability distributions of joint measurements following Y. By taking a quantum expectation of Eq. (5.2) with respect to $\hat{\rho}$, we obtain

$$
\tilde{p}^{XY}_{\hat{\rho}}(x, y) = q(x;y) p^X_{\hat{\rho}}(\tilde{x}(x;y)), \tag{5.7}
$$

where

$$\tilde{p}_{\hat{\rho}}^{XY}(x,y) = \mathrm{tr}[\mathcal{E}_y^Y(\hat{\rho})\hat{E}_x^X] = \mathrm{tr}[\hat{\rho}\mathcal{E}_y^{Y\dagger}(\hat{E}_x^X)].$$

is the joint probability density function of Y followed by X. Equation (5.7) implies that, from Theorem 3.3.2, the stochastic variable $\tilde{x}(x;y)$ is a sufficient statistics of the joint measurement process 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 (5.3), for any function $F(x)$ we have

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

which implies that the probability distribution of $\tilde{x}(x;y)$ is equivalent to that of the single measurement of X. Thus the condition (5.3) ensures

$$p_{\hat{\rho}}^{\tilde{X}}(x) = p_{\hat{\rho}}^{X}(x). \tag{5.8}$$

From Eqs. (5.7) and (5.8), we have

$$D(\tilde{p}_{\hat{\rho}}^{XY}||\tilde{p}_{\hat{\sigma}}^{XY}) = D(p_{\hat{\rho}}^{\tilde{X}}||p_{\hat{\sigma}}^{\tilde{X}}) = D(p_{\hat{\rho}}^{X}||p_{\hat{\sigma}}^{X}), \tag{5.9}$$

where in deriving the first equality, we used the relative entropy conservation for the sufficient statistic in Theorem 3.3.3. Equation (5.9) indicates that the information obtained in the joint measurement of Y followed by X is equivalent to that of the single measurement of X. By using the chain rule (3.12) of the relative entropy, the left-hand side of Eq. (5.9) can be written as

$$\begin{aligned}D(\tilde{p}_{\hat{\rho}}^{XY}||\tilde{p}_{\hat{\sigma}}^{XY}) &= D(p_{\hat{\rho}}^{Y}||p_{\hat{\sigma}}^{Y}) + E_{\hat{\rho}}[D(\tilde{p}_{\hat{\rho}}^{X|Y}(\cdot|y)||\tilde{p}_{\hat{\sigma}}^{X|Y}(\cdot|y))] \\ &= D(p_{\hat{\rho}}^{Y}||p_{\hat{\sigma}}^{Y}) + E_{\hat{\rho}}[D(p_{\hat{\rho}_y}^{X}||p_{\hat{\sigma}_y}^{X})],\end{aligned} \tag{5.10}$$

where

$$\tilde{p}_{\hat{\rho}}^{X|Y}(x|y) = \frac{\tilde{p}_{\hat{\rho}}^{XY}(x,y)}{p_{\hat{\rho}}^{Y}(y)} = p_{\hat{\rho}_y}^{X}(x)$$

is the conditional probability density of X subject to a given measurement outcome y. Note that we did not assume the conditions (4.1), (5.2) and (5.3) in deriving Eq. (5.10). From Eq. (5.10), the relative-entropy conservation law (5.6) is equivalent to Eq. (5.9), and in this sense Eq. (5.9) is another expression of the relative-entropy conservation law.

5.2 Relative-entropy conservation for PVM

In this section, we consider the case in which the system's observable X is a PVM $\{\hat{E}_x^X\}$ such that

$$\hat{E}_x^X \hat{E}_{x'}^X = \delta_{x,x'} \hat{E}_x^X, \quad \sum_{x \in \Omega_X} \hat{E}_x^X = \hat{I} \quad \text{for the discrete case;} \tag{5.11}$$

$$\hat{E}_x^X \hat{E}_{x'}^X = \delta(x - x') \hat{E}_x^X, \quad \int_{\mathbb{R}} dx \hat{E}_x^X = \hat{I} \quad \text{for the continuous case.} \tag{5.12}$$

For definiteness we again only consider continuous X in this section, but the discussion is also valid for discrete X due to the formal correspondences (4.10) and (4.11).

To establish the relative-entropy conservation law, we assume the following condition corresponding to (5.2): there exists functions $\tilde{x}(x;y)$ and $q(x;y)$ such that for any x and y

$$\mathcal{E}_y^{Y\dagger}(\hat{E}_x^X) = q(x;y) \hat{E}_{\tilde{x}(x;y)}^X. \tag{5.13}$$

From Theorem 5.1.1, it is sufficient to show the conditions (4.12) and (5.3) to prove the relative-entropy conservation law (5.6). By integrating Eq. (5.13) with respect to x, we obtain

$$\begin{aligned} \hat{E}_y^Y &= \int dx' q(x';y) \hat{E}_{\tilde{x}(x';y)}^X \\ &= \int dx \left(\int dx' \delta(x - \tilde{x}(x';y)) q(x';y) \right) \hat{E}_x^X \\ &= \int dx p(y|x) \hat{E}_x^X, \end{aligned} \tag{5.14}$$

where

$$p(y|x) = \int dx' \delta(x - \tilde{x}(x';y)) q(x';y). \tag{5.15}$$

Note that the conditional probability $p(y|x)$ in Eq. (5.14) is unique because of the linear independence of $\hat{E}_x^X$ and that $p(y|x)$ satisfies the normalization condition (4.14) from the completeness condition for $\hat{E}_y^Y$. Then, for any y and any function $F(x)$, we have

$$\begin{aligned} \int dx' q(x';y) F(\tilde{x}(x';y)) &= \int dx \left(\int dx' \delta(x - \tilde{x}(x';y)) q(x';y) \right) F(x) \\ &= \int dx p(y|x) F(x), \end{aligned}$$

where we used Eq. (5.15) in deriving the last equality. Thus, the condition (5.3) is satisfied, and therefore the relative-entropy conservation law (5.6) holds. To summarize, we have obtained the following theorem.

Theorem 5.2.1 (relative-entropy conservation law for PVM)
Let X be a discrete or continuous PVM of the form $\hat{E}_x^X$ that satisfies condition (5.11) or (5.12) and let Y be a quantum measurement process corresponding to a density of the CP instrument $\mathcal{E}_y^Y$ on a sample space $(\Omega_Y, \mathscr{B}_Y)$ with respect to a reference measure $\mu_0(dy)$. Suppose that X and Y satisfy the condition (5.13). Then there exists a unique conditional probability function $p(y|x)$ with a normalization condition (4.14) such that Eq. (5.14) holds. Furthermore, the relative-entropy conservation law (5.6) or (5.9) holds.

5.3 Equivalence between the relative-entropy conservation and the established condition

In this section, we consider the case in which X is a discrete PVM $\{\hat{E}_x^X\}_{x\in\Omega_X}$ with the discrete complete orthonormal conditions (5.11) and (5.12) and Y is a discrete measurement on a sample space $(\Omega_Y, 2^{\Omega_Y})$ described by a set of CP maps $\{\mathcal{E}_y^X\}_{y\in\Omega_Y}$ with the completeness condition

$$\sum_{y\in\Omega_Y} \mathcal{E}_y^{Y\dagger}(\hat{I}) = \hat{I}. \tag{5.16}$$

In this case, we can show the equivalence between the established condition (5.13) in Theorem 5.2.1 and the relative-entropy conservation law (5.6).

Theorem 5.3.1
Let X be a discrete PVM $\{\hat{E}_x^X\}_{x\in\Omega_X}$ with a discrete complete orthonormal condition (4.9) and let Y be a quantum measurement corresponding to a CP instrument on a discrete sample space $(\Omega_Y, 2^{\Omega_Y})$ described by a set of CP maps $\{\mathcal{E}_y^X\}_{y\in\Omega_Y}$ with the completeness condition (5.16). Then the following two conditions are equivalent:

(i) The condition (5.13) holds for all x and y.

(ii) The relative-entropy conservation law (5.6) or (5.9) holds for arbitrary states $\hat{\rho}$ and $\hat{\sigma}$.

To show the theorem, we need the following lemma.

Lemma 5.3.2
Let $\{\hat{E}^X\}_{x\in\Omega_X}$ be a PVM with a discrete complete orthonormal condition (5.11) and let $\{\hat{E}^Z_z\}_{z\in\Omega_Z}$ be a discrete POVM. Suppose that

$$D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}) = D(p^Z_{\hat{\rho}}||p^Z_{\hat{\sigma}}) \tag{5.17}$$

holds for any states $\hat{\rho}$ and $\hat{\sigma}$, where $p^X_{\hat{\rho}}(x) = \mathrm{tr}[\hat{\rho}\hat{E}^X_x]$ and $p^Z_{\hat{\rho}}(z) = \mathrm{tr}[\hat{\rho}\hat{E}^Z_z]$. Then for each $z \in \Omega_Z$ there exist a scalar $q(z) \geq 0$ and $\tilde{x}(z) \in \Omega_X$ such that

$$\hat{E}^Z_z = q(z)\hat{E}^X_{\tilde{x}(z)}. \tag{5.18}$$

Proof of Lemma 5.3.2. Let $\hat{U}_x$ be an arbitrary operator such that $\hat{U}^\dagger_x\hat{U}_x = \hat{U}_x\hat{U}^\dagger_x = \hat{E}^X_x$, i.e. $\hat{U}_x$ is an arbitrary unitary operator on a closed subspace $\hat{E}^X_x\mathcal{H}$. Define a CP and trace-preserving map $\mathcal{F}$ by

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

Since $\hat{E}_x\hat{U}_{x'} = \hat{E}_x\hat{U}_{x'}\hat{U}^\dagger_{x'}\hat{U}_{x'} = \hat{E}_x\hat{E}_{x'}\hat{U}_{x'} = \delta_{x,x'}\hat{U}_{x'}$, we have $p^X_{\hat{\rho}}(x) = p^X_{\mathcal{F}(\hat{\rho})}(x)$ for any state $\hat{\rho}$. Therefore, from the assumption (5.17) we have

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

and hence we obtain

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

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

$$\hat{E}^Z_z = \mathcal{F}^\dagger(\hat{E}^Z_z) = \sum_{x\in\Omega_X} \hat{U}^\dagger_x\hat{E}^Z_z\hat{U}_x. \tag{5.19}$$

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

$$\hat{E}^Z_z = \sum_{x\in\Omega_X} \hat{E}^X_x\hat{E}^Z_z\hat{E}^X_x. \tag{5.20}$$

From Eqs. (5.19) and (5.20), 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. (5.20) as

$$\hat{E}^Z_z = \sum_{x\in\Omega_X} \kappa(z|x)\hat{E}^X_x,$$

where $\kappa(z|x)$ is a nonnegative 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

$$p^{XZ}_{\hat{\rho}}(x,z) := \mathrm{tr}[\hat{\rho}\hat{E}^{XZ}_{xz}] = \kappa(z|x)p^X_{\hat{\rho}}(x), \tag{5.21}$$

X is a sufficient statistic for a statistical model $\{p^{XZ}_{\hat{\rho}}(x,z)\}_{\hat{\rho}\in\mathcal{S}(\mathcal{H})}$. Thus, from Theorem 3.3.3 and the assumption (5.17), we have

$$D(p^{XZ}_{\hat{\rho}}||p^{XZ}_{\hat{\sigma}}) = D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}) = D(p^Z_{\hat{\rho}}||p^Z_{\hat{\sigma}}),$$

and again from Theorem 3.3.3, Z is a sufficient statistic for $\{p^{XZ}_{\hat{\rho}}(x,z)\}_{\hat{\rho}\in\mathcal{S}(\mathcal{H})}$. Therefore there is a nonnegative scalar $r(x|z)$ such that

$$p^{XZ}_{\hat{\rho}}(x,z) = r(x|z)p^Z_{\hat{\rho}}(z),$$

or equivalently in the Heisenberg picture

$$\kappa(z|x)\hat{E}^X_x = r(x|z)\hat{E}^Z_z. \tag{5.22}$$

To prove (5.18), 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. (5.22) we have $\hat{E}^Z_z = \frac{\kappa(z|x)}{r(x|z)}\hat{E}^X_x$ and the condition (5.18) holds. □

Proof of Theorem 5.3.1. (i) ⇒ (ii) is evident from Theorem 5.2.1. Conversely, (i) readily follows from (ii) and Lemma 5.3.2 by identifying $\hat{E}^Z_z$ with $\mathcal{E}^{Y\dagger}_y(\hat{E}^X_x)$. □

5.4 Comparison with the Shannon entropy conservation

In this section, we consider the relation between the conditions for the relative-entropy conservation law and that for the Shannon entropy conservation law when the system's observable X is a PVM. As we have remarked just below Theorem 5.1.1, our condition (5.13) for Theorem 5.2.1 is less restrictive than the condition (4.13) for Theorem 4.2.1. Thus we consider rephrasing the condition (4.13) under the assumption of the condition (5.13). We divide the discussion into two parts corresponding to discrete and continuous X and Y.

5.4.1 Discrete case

Here we consider the case when X is a discrete PVM and the sample space of Y is also discrete. We do not assume Y to be pure. Then, under the condition (5.13) we can show that the condition (4.4) required by Ban is equivalent to a condition that if a pre-measurement state is an eigenstate of X, then the conditional post-measurement state for given measurement outcome y is also an eigenstate of X as the following theorem shows.

Theorem 5.4.1
Let X be a PVM $\{|x\rangle\langle x|\}_{x\in\Omega_X}$ with a discrete complete orthonormal condition (4.9) and Y be a quantum measurement described by a CP instrument $\mathcal{I}^Y_\cdot(\cdot)$ on a discrete sample space $(\Omega_Y, 2^{\Omega_Y})$ and let us define $\mathcal{E}^Y_y := \mathcal{I}^Y_{\{y\}}$ for $y \in \Omega_Y$. Suppose that X and Y satisfy the condition (5.13) for Theorem 5.2.1. Then the following four conditions are equivalent:

1. The condition (4.13) holds, i.e. $q(x; y) = p(y|\tilde{x}(x; y))$.

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

$$\sum_{x'\in\Omega_X} \delta_{x,\tilde{x}(x';y)} = 1. \tag{5.23}$$

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

4. The conditional post-measurement state is an eigenstate of X if the pre-measurement state is an eigenstate. Namely, for any x and y, there exist functions $\bar{x}(x; y)$ and $r(x; y) \geq 0$ such that

$$\mathcal{E}^Y_y(|x\rangle\langle x|) = r(x; y)\,|\bar{x}(x; y)\rangle\langle\bar{x}(x; y)|\,. \tag{5.24}$$

Here, we take the convention that for any x and y such that $q(x; y) = 0$, $\tilde{x}(x; y)$ is defined to be $\emptyset$ which is a symbol out of the sample set Ω_X. We also define $p(y|\emptyset) := 0$ for any $y \in \Omega_Y$.

Proof. First, we remark that from Theorem 5.2.1, there exists a unique conditional probability $p(y|x)$ such that

$$\hat{E}^Y_y = \mathcal{E}^{Y\dagger}_y(\hat{I}) = \sum_{x\in\Omega_X} p(y|x)\,|x\rangle\langle x|\,. \tag{5.25}$$

By taking the summation of Eq. (5.13) with respect to x, we also have

$$\begin{aligned}\hat{E}^Y_y &= \sum_{x'\in\Omega_X} q(x';y)\,|\tilde{x}(x';y)\rangle\,\langle\tilde{x}(x';y)| \\ &= \sum_{x'\in\Omega_X}\left(\sum_{x\in\Omega_X}\delta_{x,\tilde{x}(x';y)}\right) q(x';y)\,|\tilde{x}(x';y)\rangle\,\langle\tilde{x}(x';y)| \\ &= \sum_{x\in\Omega_X}\left(\sum_{x'\in\Omega_X}\delta_{x,\tilde{x}(x';y)}q(x';y)\right)|x\rangle\,\langle x|\,. \end{aligned} \tag{5.26}$$

Since Eq. (5.26) coincides with Eq. (5.25), we obtain

$$p(y|x) = \sum_{x'\in\Omega_X}\delta_{x,\tilde{x}(x';y)}q(x';y). \tag{5.27}$$

$1 \Rightarrow 2$: From the condition $q(x;y) = p(y|\tilde{x}(x;y))$ and Eq. (5.27), we have

$$\begin{aligned}p(y|x) &= \sum_{x'\in\Omega_X}\delta_{x,\tilde{x}(x';y)}p(y|\tilde{x}(x;y)) \\ &= \sum_{x'\in\Omega_X}\delta_{x,\tilde{x}(x';y)}p(y|x),\end{aligned}$$

and therefore, Eq. (5.23) holds for any x and y such that $p(y|x) \neq 0$.

$2 \Rightarrow 3$ is evident from the definition of the Kronecker delta.

$3 \Rightarrow 4$: From Eq. (5.25) we have

$$p(y|x) = \mathrm{tr}\left[|x\rangle\,\langle x|\,\mathcal{E}^{Y\dagger}_y(\hat{I})\right] = \mathrm{tr}\left[\mathcal{E}^Y_y(|x\rangle\,\langle x|)\right]. \tag{5.28}$$

For the case of $p(y|x) = 0$, from Eq. (5.28) and the positivity of the superoperator $\mathcal{E}^Y_y$, we have $\mathcal{E}^Y_y(|x\rangle\,\langle x|) = 0$ and the condition (5.24) holds. Let us assume $p(y|x) \neq 0$. From the complete positivity of the superoperator $\mathcal{E}^{Y\dagger}_y$ and Theorem 2.2.2, there exists a set of bounded operators $\{\hat{M}_{yk}\}_k$ such that

$$\mathcal{E}^{Y\dagger}_y(\hat{A}) = \sum_k \hat{M}^\dagger_{yk}\hat{A}\hat{M}_{yk} \tag{5.29}$$

for any bounded operator $\hat{A}$. From Eqs. (5.13) and (5.29), we have

$$\sum_k \hat{M}^\dagger_{yk}\,|x\rangle\,\langle x|\,\hat{M}_{yk} = q(x;y)\,|\tilde{x}(x;y)\rangle\,\langle\tilde{x}(x;y)|\,,$$

and thus we obtain

$$\hat{M}_{yk}^{\dagger} |x\rangle = a(x;y,k) |\tilde{x}(x;y)\rangle , \tag{5.30}$$

where $a(x;y,k)$ is a complex scalar such that

$$\sum_k |a(x;y,k)|^2 = q(x;y). \tag{5.31}$$

From Eqs. (5.29) and (5.30), we have

$$\begin{aligned} \mathcal{E}_y^{Y\dagger}(|x''\rangle\langle x'|) &= \sum_k \hat{M}_{yk}^{\dagger} |x''\rangle \langle x'| \hat{M}_{yk} \\ &= \left(\sum_k a(x'';y,k) a^*(x';y,k) \right) |\tilde{x}(x'';y)\rangle \langle \tilde{x}(x';y)| . \end{aligned} \tag{5.32}$$

Therefore

$$\begin{aligned} \langle x'| \mathcal{E}_y^Y(|x\rangle\langle x|) |x''\rangle &= \mathrm{tr}\left[\mathcal{E}_y^Y(|x\rangle\langle x|) |x''\rangle\langle x'| \right] \\ &= \mathrm{tr}\left[|x\rangle\langle x| \mathcal{E}_y^{Y\dagger}(|x''\rangle\langle x'|) \right] \\ &= \left(\sum_k a(x'';y,k) a^*(x';y,k) \right) \delta_{x,\tilde{x}(x'';y)} \delta_{x,\tilde{x}(x';y)} , \end{aligned} \tag{5.33}$$

where we used Eq. (5.32) in deriving the last equality. From the condition 3, we can define a function $\bar{x}(x;y)$ such that $x = \tilde{x}(x';y)$ implies $x' = \bar{x}(x;y)$ for any x and y with $p(y|x) \neq 0$. Thus Eq. (5.33) becomes

$$\begin{aligned} \langle x'| \mathcal{E}_y^Y(|x\rangle\langle x|) |x''\rangle &= \left(\sum_k |a(x';y,k)|^2 \right) \delta_{x',\bar{x}(x;y)} \delta_{x'',\bar{x}(x;y)} \\ &= q(x';y) \delta_{x',\bar{x}(x;y)} \delta_{x'',\bar{x}(x;y)} , \end{aligned} \tag{5.34}$$

where we used Eq. (5.31) in deriving the last equality. Equation (5.34) implies

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

and the condition (5.24) holds.

$4 \Rightarrow 1$: From the condition (5.24) and Eq. (5.25), $p(y|x)$ becomes

$$\begin{aligned} p(y|x) &= \mathrm{tr}[|x\rangle\langle x| \mathcal{E}_y^{Y\dagger}(\hat{I})] \\ &= \mathrm{tr}[\mathcal{E}_y^Y(|x\rangle\langle x|)] \\ &= r(x;y). \end{aligned} \tag{5.35}$$

Then $q(x;y)$ is given by

$$\begin{aligned} q(x;y) &= \mathrm{tr}\left[|\tilde{x}(x;y)\rangle \langle \tilde{x}(x;y)| \, \mathcal{E}_y^{Y\dagger}(|x\rangle\langle x|) \right] \\ &= \mathrm{tr}\left[\mathcal{E}_y^{Y}(|\tilde{x}(x;y)\rangle\langle \tilde{x}(x;y)|) \, |x\rangle\langle x| \right] \\ &= p(y|\tilde{x}(x;y)) \delta_{x,\bar{x}(\tilde{x}(x;y);y)}, \end{aligned} \tag{5.36}$$

where we used the condition (5.13) in deriving the first equality and the last equality follows from the Eqs. (5.24) and (5.35). Therefore, if $q(x;y) \neq 0$, Eq. (5.36) implies $q(x;y) = p(y|\tilde{x}(x;y))$ and the condition (4.13) holds. If $q(x;y) = 0$, then $p(y|\tilde{x}(x;y)) = p(y|\emptyset) = 0 = q(x;y)$, and thus the condition (4.13) also holds in this case. □

As an example that does not satisfy Ban's condition (4.13) but does satisfy the relative-entropy conservation law, we will consider a destructive measurement of X as follows.

Example 7 (destructive sharp measurement of X)
Let X be a PVM $\{|x\rangle\langle x|\}_{x\in\Omega_X}$ with a discrete complete orthonormal condition (4.9) and let Y be a quantum measurement corresponding to a CP instrument $\mathcal{I}^Y_\cdot(\cdot)$ on the same discrete sample space $(\Omega_X, 2^{\Omega_X})$. Define $\mathcal{E}_y^Y := \mathcal{I}_{\{y\}}^Y$ for $y \in \Omega_X$ as usual. Suppose that the CP instrument $\mathcal{E}_y^Y$ is given by

$$\mathcal{E}_y^Y(\hat{\rho}) = \langle y| \hat{\rho} |y\rangle \hat{\rho}_y, \tag{5.37}$$

where $\hat{\rho}_y$ is an arbitrary state. Note that the superoperator (5.37) is CP because it has a Kraus representation

$$\langle y| \hat{\rho} |y\rangle \hat{\rho}_y = \sum_k \left(\sqrt{\kappa(k|y)} \, |\phi_{yk}\rangle\langle y| \right) \hat{\rho} \left(\sqrt{\kappa(k|y)} \, |y\rangle\langle \phi_{yk}| \right), \tag{5.38}$$

where

$$\hat{\rho}_y = \sum_k \kappa(k|y) \, |\phi_{yk}\rangle\langle \phi_{yk}|$$

is the spectral decomposition of $\hat{\rho}_y$. From Eq. (5.37), we have

$$\begin{aligned} \mathcal{E}_y^{Y\dagger}(\hat{A}) &= \mathrm{tr}[\hat{\rho}_y \hat{A}] \, |y\rangle\langle y| , \\ \mathcal{E}_y^{Y\dagger}(|x\rangle\langle x|) &= \langle x| \hat{\rho}_y |x\rangle \, |y\rangle\langle y| \end{aligned} \tag{5.39}$$

and

$$\hat{E}_y^Y = |y\rangle\langle y| . \tag{5.40}$$

Thus, from Eq. (5.39), the condition (5.13) is satisfied with

$$q(x;y) = \langle x| \hat{\rho}_y |x\rangle ,$$
$$\tilde{x}(x;y) = y.$$

Then, from Theorem 5.2.1, the relative entropy conservation law (5.6) holds. On the other hand, the condition (4.13) is not necessarily satisfied since the post-measurement state $\hat{\rho}_y$ is in general not an eigenstate of X. To be definite, let the dimension d of the system's Hilbert space $\mathcal{H}$ be finite, and $\hat{\rho}_y$ be a maximally mixed state $\hat{I}/d$ for all $y \in \Omega_X$. In this case, the decrease in the Shannon entropy of X is evaluated to be

$$H_{\hat{\rho}}(X) - H_{\hat{\rho}_y}(X) = H_{\hat{\rho}}(X) - \ln d, \tag{5.41}$$

while the mutual information between X and Y is given by

$$I_{\hat{\rho}}(X : Y) = H_{\hat{\rho}}(X),$$

which differs from Eq. (5.41) by the factor $-\ln d$. The factor $-\ln d$ is the Shannon entropy for the post-measurement state $\hat{\rho}_y$ which in general depends strongly upon the choice of the post-measurement state $\hat{\rho}_y$. On the other hand, our formalism focuses on the information about discriminating the pre-measurement state and the post-measurement state $\hat{\rho}_y$ does not have any information about the pre-measurement state. Reflecting this fact properly, the relative entropy of X for the post-measurement candidate states vanishes and the relative entropy conservation law holds with $D(p^X_{\hat{\rho}}||p^X_{\hat{\sigma}}) = D(p^Y_{\hat{\rho}}||p^Y_{\hat{\sigma}})$.

5.4.2 Continuous case

Next we consider a continuous X. In this case we cannot establish a simple rephrasing of the condition (4.13) as in the discrete case. We can still show Eq. (4.17) from the condition (4.13), which is the continuous analogue of Eq. (5.23). However, the formal correspondences (4.10) and (4.11) do not work in this case, for we may not conclude the condition 3 in Theorem 5.3.1 from Eq. (4.17). For simplicity let us assume that $\tilde{x}(x;y)$ is a differentiable one-to-one function with respect to x for each y. In this case, Eq. (4.17), which is a necessary condition for (4.13), implies

$$\left| \frac{\partial \tilde{x}(x;y)}{\partial x} \right| = 1, \tag{5.42}$$

i.e. the Jacobian of the transformation $x \to \tilde{x}(x;y)$ should be unity. The condition (5.42) reflects the strong dependence of the differential entropy on the reference measure, and we will see in the next chapter the examples in which the unit-Jacobian condition (5.42) breaks down. In this sense, the condition (4.13) for continuous X is even more restrictive than that for discrete X.

Chapter 6

Applications to Continuous Measurements

In this chapter, we apply the general theorems on the relative-entropy conservation obtained in Chapter 5 to four typical continuous measurements on a single mode photon field, namely photon counting, quantum counting, homodyne measurement, and heterodyne measurement. In these examples, we also examine the Shannon entropy conservation and show that the Shannon entropy conservation does not hold except for the photon-counting model.

6.1 Photon-counting model

In the photon-counting model [5, 6, 7], the photon number of a single-mode field is measured in a destructive manner. The model is a measurement continuous in time and the measurement operators of an infinitesimal time interval dt are given by

$$\hat{M}_0^{\mathrm{pc}}(dt) = \hat{I} - \frac{\gamma}{2}\hat{n}dt, \tag{6.1}$$

$$\hat{M}_1^{\mathrm{pc}}(dt) = \sqrt{\gamma dt}\hat{a}, \tag{6.2}$$

where $\gamma > 0$ 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. Here we adopt the interaction picture in which the unitary term $e^{-i\omega t\hat{n}}$ for the free motion is omitted. The event corresponding to the measurement operator (6.1) is called a no-count process in which no photon is detected. On the other hand, the event corresponding to (6.2) is called a one-count process in which a photocount is registered. The photon-counting process is known to be implemented by a measurement model in which resonant two-level atoms initially prepared in the ground state interact with the photon field and the level of the atoms after their interaction with the photon field are then measured [7].

From measurement operators (6.1) and (6.2), for a finite time interval $[0, t]$, it can be shown that both the state change and the statistic of the measurement outcome are dependent on a single integer m, which is the number of photocounts in the time interval, and the measurement operator is given by [6]

$$\hat{M}_m^{\mathrm{pc}}(t) = \sqrt{\frac{(1-e^{-\gamma t})^m}{m!}} e^{-\frac{\gamma t \hat{n}}{2}} \hat{a}^m. \tag{6.3}$$

From Eq. (6.3), the POVM for the measurement outcome m is evaluated to be

$$\hat{M}_m^{\mathrm{pc}\dagger}(t)\hat{M}_m^{\mathrm{pc}}(t) = p^{\mathrm{pc}}(m|\hat{n};t), \tag{6.4}$$

where

$$p^{\mathrm{pc}}(m|n;t) = \binom{n}{m} (1-e^{-\gamma t})^m e^{-\gamma t(n-m)}. \tag{6.5}$$

Equation (6.5) implies that the information of the measurement outcome m is obtained by a coarse-graining of the photon-number distribution. In the infinite-time limit, the conditional probability Eq. (6.5) reduces to $\delta_{m,n}$, indicating that the complete information on the photon-number distribution is obtained by the measurement outcome of photon counting.

Let us examine the relative-entropy conservation. As the system observable X, we take a discrete PVM $|n\rangle\langle n|$, where $|n\rangle$ is the normalized eigenstate of $\hat{n}$. Then from Eq. (6.4), the condition (4.1) holds. From the measurement operator (6.3) we obtain

$$\hat{M}_m^\dagger(t) |n\rangle \langle n| \hat{M}_m(t) = p^{\mathrm{pc}}(m|\tilde{n}(n;m);t) |\tilde{n}(n;m)\rangle \langle \tilde{n}(n;m)|, \tag{6.6}$$

where

$$\tilde{n}(n;m) = n + m. \tag{6.7}$$

Equation (6.7) can be interpreted as the photon number for the pre-measurement state given by the sum of the photon number of the pre-measurement state and the number of photocounts, which is to be contrasted with the QND measurement in which $\tilde{x}(x;y)$ coincides with x for post-measurement state. Equation (6.6) shows that the condition (5.13) is satisfied. Thus from Theorem 5.2.1, we have the relative-entropy conservation law

$$D(p_{\hat{\rho}}^{\mathrm{pc}}(\cdot;t)||p_{\hat{\sigma}}^{\mathrm{pc}}(\cdot;t)) = D(p_{\hat{\rho}}^N||p_{\hat{\sigma}}^N) - E[D(p_{\hat{\rho}_m}^N||p_{\hat{\sigma}_m}^N)],$$

where $p_{\hat{\rho}}^{N} := \langle n| \hat{\rho} |n\rangle$ is the photon number distribution for a state $\hat{\rho}$,

$$p_{\hat{\rho}}^{\mathrm{pc}}(m;t) := \mathrm{tr}[\hat{\rho}\hat{M}_m^{\mathrm{pc}\dagger}(t)\hat{M}_m^{\mathrm{pc}}(t)]$$

is the distribution for the number of the photocounts m, and $\hat{\rho}_m$ is the post-measurement state for the measurement outcome m. Note that in this model Ban's condition (4.13) is also satisfied as seen from Eq. (6.6) and the Shannon entropy conservation law holds [8].

6.2 Quantum-counter model

The quantum counter model [22, 23] also measures the number of photons in a destructive manner continuously in time but differs in that the present model increases the number of photons in a one-count process whereas it decreases for the photon counting model. The measurement operators for the no-count and one-count processes in an infinitesimal time interval dt are given by

$$\hat{M}_0^{\mathrm{qc}}(dt) = \hat{I} - \frac{\gamma}{2}\hat{a}\hat{a}^\dagger dt, \tag{6.8}$$

$$\hat{M}_1^{\mathrm{qc}}(dt) = \sqrt{\gamma dt}\hat{a}^\dagger, \tag{6.9}$$

where we again adopt the interaction picture. The effective measurement operator for a finite time interval $[0,t]$ again depends only on the total number of counts m in the time interval and given by [23]

$$\hat{M}_m^{\mathrm{qc}}(t) = \sqrt{\frac{(e^{\gamma t}-1)^m}{m!}} e^{-\gamma t\hat{a}\hat{a}^\dagger/2} \left(\hat{a}^\dagger\right)^m. \tag{6.10}$$

The POVM for m is evaluated to be

$$\begin{aligned}\hat{E}_m^{\mathrm{qc}}(t) &:= \hat{M}_m^{\mathrm{qc}\dagger}(t)\hat{M}_m^{\mathrm{qc}}(t)\\ &= \frac{(e^{\gamma t}-1)^m}{m!}\hat{a}^m e^{-\gamma t\hat{a}\hat{a}^\dagger}(\hat{a}^\dagger)^m\\ &= p^{\mathrm{qc}}(m|\hat{n};t),\end{aligned} \tag{6.11}$$

where

$$p^{\mathrm{qc}}(m|n;t) = \binom{n+m}{m}(e^{\gamma t}-1)^m e^{-\gamma t(n+m+1)}. \tag{6.12}$$

In this measurement model, two kinds of relative-entropy conservation laws can be shown. The first one is for the photon number distribution. As in the photon counting case, we obtain

$$\hat{M}_m^{\mathrm{qc}\dagger}(t)\,|n\rangle\langle n|\,\hat{M}_m^{\mathrm{qc}}(t) = p^{\mathrm{qc}}(m|\tilde{n}(n;m);t)\,|\tilde{n}(n;m)\rangle\langle\tilde{n}(n;m)|\,, \tag{6.13}$$

$$\tilde{n}(n;m) = n - m \tag{6.14}$$

and the condition (5.13) as well as Ban's condition (4.13) holds. Thus the relative-entropy conservation law

$$D(p_{\hat{\rho}}^{\mathrm{qc}}(\cdot;t)||p_{\hat{\sigma}}^{\mathrm{qc}}(\cdot;t)) = D(p_{\hat{\rho}}^{N}||p_{\hat{\sigma}}^{N}) - E_{\hat{\rho}}[D(p_{\hat{\rho}_m(t)}^{N}||p_{\hat{\sigma}_m(t)}^{N})] \tag{6.15}$$

holds, where

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

and $\hat{\rho}_m(t)$ is the post-measurement state for the given number of counts m.

The second conservation law is for the POVM defined by

$$\hat{E}_x^X dx = p^X(x|\hat{n})dx, \tag{6.16}$$

where

$$p^X(x|n) = \frac{e^{-x}x^n}{n!} \tag{6.17}$$

and $x \in [0,\infty)$. The probability distribution function of X is defined by

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

It is known [23] that the distribution of a stochastic variable

$$\frac{m}{e^{\gamma t}}$$

converges to that of X in the limit $t \to \infty$. In other words, X represents the total information obtained in the quantum-counter measurement. The photon number distribution is determined by that of X as shown in the following equation [23]:

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

Still we can show that X is less informative than the photon number distribution.

From Eqs. (6.10) and (6.16), we obtain

$$\hat{M}_m^{\mathrm{qc}\dagger}(t)\hat{E}_x^X\hat{M}_m^{\mathrm{qc}}(t) = q(x;m)p^X(\tilde{x}(x;m)|\hat{n}), \tag{6.18}$$

where

$$q(x;m) = e^{-\gamma t}p^{\mathrm{qc}}(m|\tilde{x}(x;m)), \tag{6.19}$$

$$p^{\mathrm{qc}}(m|x) = \frac{[(e^{\gamma t}-1)x]^m}{m!}\exp\left[-(e^{\gamma t}-1)x\right], \tag{6.20}$$

$$\tilde{x}(x;m) = e^{-\gamma t}x. \tag{6.21}$$

Here $p^{\mathrm{qc}}(m|x)$ satisfies the normalization condition

$$\sum_{m=0}^{\infty} p^{\mathrm{qc}}(m|x) = 1.$$

Furthermore, for any integrable function $F(x)$, we have

$$\begin{aligned}\int_0^\infty dx q(x;m) F(\tilde{x}(x;m)) &= \int_0^\infty d(e^{-\gamma t}x) p^{\mathrm{qc}}(m|e^{-\gamma t}x) F(e^{-\gamma t}x) \\ &= \int_0^\infty dx p^{\mathrm{qc}}(m|x) F(x). \end{aligned} \tag{6.22}$$

The POVM for m can be written as

$$\begin{aligned}\hat{M}_m^{\mathrm{qc}\dagger}(t)\hat{M}_m^{\mathrm{qc}}(t) &= \int_0^\infty dx \hat{M}_m^{\mathrm{qc}\dagger}(t)\hat{E}_x^X \hat{M}_m^{\mathrm{qc}}(t) \\ &= \int_0^\infty dx q(x;m) p^X(\tilde{x}(x;m)|\hat{n}) \\ &= \int_0^\infty dx p^{\mathrm{qc}}(m|x) p^X(x|\hat{n}), \end{aligned} \tag{6.23}$$

where we used Eqs. (6.18) and (6.22). Equations (6.18), (6.22) and (6.23) ensure the condition for Theorem 5.1.1 and we obtain the relative-entropy conservation law

$$D(p_{\hat{\rho}}^{\mathrm{qc}}(\cdot;t)||p_{\hat{\sigma}}^{\mathrm{qc}}(\cdot;t)) = D(p_{\hat{\rho}}^X||p_{\hat{\sigma}}^X) - E_{\hat{\rho}}\left[D(p_{\hat{\rho}_m(t)}^X||p_{\hat{\sigma}_m(t)}^X)\right]. \tag{6.24}$$

Since X is equivalent to the total information involved in the measurement outcome, $D(p_{\hat{\rho}}^{\mathrm{qc}}(\cdot;t)||p_{\hat{\sigma}}^{\mathrm{qc}}(\cdot;t))$ converges to $D(p_{\hat{\rho}}^X||p_{\hat{\sigma}}^X)$ in the infinite-time limit. Thus from Eqs. (6.15) and (6.24) we have

$$E_{\hat{\rho}}[D(p_{\hat{\rho}_m(t)}^N||p_{\hat{\sigma}_m(t)}^N)] \xrightarrow{t\to\infty} D(p_{\hat{\rho}}^N||p_{\hat{\sigma}}^N) - D(p_{\hat{\rho}}^X||p_{\hat{\sigma}}^X), \tag{6.25}$$

$$E_{\hat{\rho}}\left[D(p_{\hat{\rho}_m(t)}^X||p_{\hat{\sigma}_m(t)}^X)\right] \xrightarrow{t\to\infty} 0. \tag{6.26}$$

This equations show the difference between the asymptotic behaviors of the relative entropies for the post-measurement state. The right-hand side of Eq. (6.25) is the difference of the relative entropies for $\hat{n}$ and X. From the chain rule (3.12) it can be written as

$$\int_0^\infty dx p_{\hat{\rho}}^X(x) D(p_{\hat{\rho}}^N(\cdot|x)||p_{\hat{\sigma}}^N(\cdot|x)) \geq 0, \tag{6.27}$$

where

$$p_{\hat{\rho}}^{N}(n|x) = \frac{p^{X}(x|n)p_{\hat{\rho}}^{N}(n)}{p_{\hat{\rho}}^{X}(x)} \tag{6.28}$$

is the photon-number distribution conditioned by X. Equation (6.28) vanishes if and only if the photon number distributions of $\hat{\rho}$ and $\hat{\sigma}$ coincide. To show this, let the left-hand side of Eq. (6.28) be 0. Then from Proposition 3.2.2 we have

$$\forall n \geq 0, \quad p_{\hat{\rho}}^{N}(n|x) = p_{\hat{\sigma}}^{N}(n|x) \tag{6.29}$$

for almost all $x > 0$. Therefore there exists at least one $x > 0$ which satisfies Eq. (6.29). For such x, from Eqs. (6.28) and (6.29) we have

$$\forall n \geq 0, \quad \frac{\langle n|\,\hat{\rho}\,|n\rangle}{p_{\hat{\rho}}^{X}(x)} = \frac{\langle n|\,\hat{\sigma}\,|n\rangle}{p_{\hat{\sigma}}^{X}(x)}. \tag{6.30}$$

By taking the summation of Eq. (6.30) we have $p_{\hat{\rho}}^{X}(x) = p_{\hat{\sigma}}^{X}(x)$ and again from Eq. (6.30) we obtain $p_{\hat{\rho}}^{N}(n) = p_{\hat{\sigma}}^{N}(n)$ for all $n \geq 0$. Thus we have shown that $\hat{n}$ is more informative than X unless $\hat{\rho}$ and $\hat{\sigma}$ have the same photon-number distribution.

The Shannon entropy conservation for the photon number holds since Ban's condition (4.13) holds from Eq. (6.13). In the case of X, however, the corresponding condition (4.4) does not hold as shown in Eq. (6.18). Furthermore, the amount of the decrease in the Shannon entropies for X is evaluated to be

$$\begin{aligned}
& h(p_{\hat{\rho}}^{X}) - E_{\hat{\rho}}[h(p_{\hat{\rho}_m(t)}^{X})] \\
&= h(p_{\hat{\rho}}^{X}) + \sum_{m=0}^{\infty} p_{\hat{\rho}}^{\mathrm{qc}}(m) \int_0^\infty dx p_{\hat{\rho}_m(t)}^{X}(x) \ln p_{\hat{\rho}_m(t)}^{X}(x) \\
&= h(p_{\hat{\rho}}^{X}) + \sum_{m=0}^{\infty} \int_0^\infty dx e^{-\gamma t} p^{\mathrm{qc}}(m|e^{-\gamma t}x) p_{\hat{\rho}}^{X}(e^{-\gamma t}x) \ln \left(\frac{e^{-\gamma t} p^{\mathrm{qc}}(m|e^{-\gamma t}x) p_{\hat{\rho}}^{X}(e^{-\gamma t}x)}{p_{\hat{\rho}}^{\mathrm{qc}}(m)} \right) \\
&= -\gamma t + I_{\hat{\rho}}(X:\mathrm{qc}) \neq I_{\hat{\rho}}(X:\mathrm{qc}),
\end{aligned} \tag{6.31}$$

where $I_{\hat{\rho}}(X:\mathrm{qc})$ is the mutual information between the measurement outcome m and X. The term $-\gamma t$ comes from the Jacobian of the transformation $x \to \tilde{x}(x;m)$ as explained in Sec. 5.3.2. Note that in this case X is not a PVM.

6.3 Homodyne measurement

In a balanced homodyne measurement [24, 25, 26], one of the quadrature amplitudes of a photon field is measured in a destructive manner such that the post-measurement state relaxes to a vacuum state $|0\rangle$. Here the quadrature-amplitude operators are defined by

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

so that they satisfy the canonical commutation relation $[\hat{X}_1, \hat{X}_2] = i$. One way of implementing this measurement is to mix the signal photon field with a classical local oscillator and divide the mixed field into two output fields via a 50%-50% beam splitter and measuring the difference of the photocurrents of the two output signals.

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

$$\hat{M}(d\xi(t); dt) = \hat{I} - \frac{\gamma}{2}\hat{n}dt + \sqrt{\gamma}\hat{a}\, d\xi(t), \tag{6.32}$$

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

$$(d\xi(t))^2 = dt.$$

The reference measure μ_0 for $d\xi$ is the Wiener measure in which $d\xi(t)$'s at different times obey a mutually independent Gaussian distribution with the first and second moments given by

$$E_{\mu_0}(d\xi) = 0,$$
$$E_{\mu_0}((d\xi)^2) = dt,$$

where $E_{\mu_0}(\cdot)$ denotes the expectation with respect to the Winer measure μ_0.

To derive the effective instrument for a finite time interval $[0, t]$, let us consider the case in which the initial state is a pure state $|\psi_0\rangle$. In this case the state change is described by the following stochastic Shrödinger equation:

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

The solution is given by [25]

$$|\psi(t)\rangle = \hat{M}_{y(t)}(t)\, |\psi_0\rangle\, , \tag{6.33}$$

where

$$\hat{M}_{y(t)}(t) = e^{-\frac{\gamma t}{2}\hat{n}} \exp\left[y(t)\hat{a} - \frac{1}{2}(1-e^{-\gamma t})\hat{a}^2 \right], \tag{6.34}$$

$$y(t) = \sqrt{\gamma}\int_0^t e^{-\frac{\gamma s}{2}} d\xi(s). \tag{6.35}$$

In the Wiener measure, $y(t)$ in Eq. (6.35) is a Gaussian variable with the vanishing first moment and the second moment

$$E_{\mu_0}[y(t)^2] = \gamma \int_0^t e^{-\gamma s} ds = 1 - e^{-\gamma t}.$$

Thus the reference measure $\mu_0(dy)$ for $y(t)$ is given by

$$\mu_0(dy) = \frac{dy}{\sqrt{2\pi(1-e^{-\gamma t})}} \exp\left[-\frac{y^2}{2(1-e^{-\gamma t})}\right]. \tag{6.36}$$

As the system's observable, we take a continuous PVM $|x\rangle_{1\,1}\langle x|$, where $|x\rangle_1$ is the Dirac ket such that

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

Then we can show the following relations (see Appendix B for the derivation):

$$\mu_0(dy)\hat{M}_y^\dagger(t)\hat{M}_y(t) = dyp(y|\hat{X}_1), \tag{6.37}$$

$$p(y|x) = \frac{1}{\sqrt{2\pi e^{-\gamma t}(1-e^{-\gamma t})}} \exp\left[-\frac{\left(y-\sqrt{2}(1-e^{-\gamma t})x\right)^2}{2e^{-\gamma t}(1-e^{-\gamma t})}\right], \tag{6.38}$$

$$\mu_0(dy)\hat{M}_y^\dagger(t)|x\rangle_{1\,1}\langle x|\hat{M}_y(t) = dyq(x;y;t)|\tilde{x}(x;y;t)\rangle_{1\,1}\langle\tilde{x}(x;y;t)|, \tag{6.39}$$

$$q(x;y;t) = e^{-\gamma t/2} p(y|\tilde{x}(x;y;t)), \tag{6.40}$$

$$\tilde{x}(x;y;t) = e^{-\frac{\gamma t}{2}} x + \frac{y}{\sqrt{2}}. \tag{6.41}$$

See Appendix B for the derivation. We can see the destructive nature of the measurement from Eq. (6.41). Equation (6.39) ensures the condition (5.13) and the relative-entropy conservation law

$$D(p_{\hat{\rho}}^Y(\cdot;t)||p_{\hat{\sigma}}^Y(\cdot;t)) = D(p_{\hat{\rho}}^{X_1}||p_{\hat{\sigma}}^{X_1}) - E_{\hat{\rho}}[D(p_{\hat{\rho}_y}^{X_1}||p_{\hat{\sigma}_y}^{X_1})], \tag{6.42}$$

where $p_{\hat{\rho}}^{X_1}(x) = {}_1\langle x|\,\hat{\rho}\,|x\rangle_1$ and

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

is the probability measure of the measurement outcome $y(t)$.

The condition (4.13) for Theorem 4.2.1 is not satisfied as can be seen from Eq. (6.39). Furthermore, we have

$$\begin{aligned} &h(p_{\hat{\rho}}^{X_1}) - E_{\hat{\rho}}[h(p_{\hat{\rho}_y}^{X_1})] \\ &= h(p_{\hat{\rho}}^{X_1}) + \int dxdye^{-\gamma t/2}p(y|\tilde{x}(x;y))p_{\hat{\rho}}^{X_1}(\tilde{x}(x;y)) \ln\left(\frac{e^{-\gamma t/2}p(y|\tilde{x}(x;y))p_{\hat{\rho}}^{X_1}(\tilde{x}(x;y))}{p_{\hat{\rho}}^Y(y)}\right) \\ &= -\frac{\gamma t}{2} + I_{\hat{\rho}}(X_1:Y) \neq I_{\hat{\rho}}(X_1:Y), \end{aligned} \tag{6.43}$$

and the Shannon entropy conservation law does not hold. The term $-\gamma t/2$ comes from the Jacobian of the transformation $x \to \tilde{x}(x; y)$.

6.4 Heterodyne measurement

The final example is the heterodyne measurement in which both of the quadrature amplitudes $\hat{X}_1$ and $\hat{X}_2$ are measured simultaneously in a destructive manner as in the homodyne measurement. This measurement is implemented by detuning the frequency of the local oscillator in the balanced homodyne setting. The sine and cosine components of the output signal correspond to two quadrature amplitudes [26].

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

$$\hat{M}(d\zeta(t); dt) = \hat{I} - \frac{\gamma}{2}\hat{n}dt + \sqrt{\gamma}\hat{a}d\zeta(t), \tag{6.44}$$

where $d\zeta(t)$ is a complex stochastic variable with the complex Itô rule

$$(d\zeta(t))^2 = (d\zeta^*(t))^2 = 0, \quad d\zeta(t)d\zeta^*(t) = dt. \tag{6.45}$$

The reference measure μ_0 for $d\zeta$ is the complex Wiener measure in which real and imaginary parts of $d\zeta(t)$ obey independent Wiener measures consistent with the Itô rule (6.45).

The stochastic Schrödinger equation

$$|\psi(t+dt)\rangle = \hat{M}(dt; d\zeta(t))\,|\psi(t)\rangle \tag{6.46}$$

has the solution

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

where $|\psi_0\rangle$ is the initial state at $t = 0$ and

$$\hat{M}_{y(t)}(t) = e^{-\frac{\gamma t}{2}\hat{n}}e^{y(t)\hat{a}}, \tag{6.47}$$

$$y(t) = \sqrt{\gamma}\int_0^t e^{-\frac{\gamma s}{2}}d\zeta(s). \tag{6.48}$$

In the complex Wiener measure, y is a complex Gaussian variable with the vanishing first moment and the second moments

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

Thus the reference measure $\mu_0(dy)$ is given by

$$\mu_0(dy) = \frac{e^{-\frac{|y(t)|^2}{1-e^{-\gamma t}}}}{\pi(1 - e^{-\gamma t})}d^2y, \tag{6.49}$$

where $d^2y = d(\mathrm{Re}y)d(\mathrm{Im}y)$. The density of the POVM for y is given by

$$\hat{M}_y^\dagger(t)\hat{M}_y(t) = \mathscr{A}\left\{\exp\left[\gamma t - (e^{\gamma t}-1)\hat{a}\hat{a}^\dagger + e^{\gamma t}(y\hat{a}+y^*\hat{a}^\dagger) - e^{\gamma t}|y|^2\right]\right\}, \tag{6.50}$$

where $\mathscr{A}$ is the anti-normal ordering in which annihilation operators are placed to the left of the creation operators. By using the overcompleteness condition for the coherent state (B.1)

$$\int d^2\alpha \hat{E}_\alpha^Q = \hat{I}, \tag{6.51}$$

where

$$\hat{E}_\alpha^Q = \frac{|\alpha\rangle\langle\alpha|}{\pi}, \tag{6.52}$$

we have

$$\begin{aligned}\mathscr{A}\left\{\hat{a}^n(\hat{a}^\dagger)^m\right\} &= \int \frac{d^2\alpha}{\pi}\hat{a}^n|\alpha\rangle\langle\alpha|(\hat{a}^\dagger)^m \\ &= \int d^2\alpha\alpha^n(\alpha^*)^m\hat{E}_\alpha^Q,\end{aligned}$$

and therefore, for any function f, we have

$$\mathscr{A}\left\{f(\hat{a},\hat{a}^\dagger)\right\} = \int d^2\alpha f(\alpha,\alpha^*)\hat{E}_\alpha^Q.$$

From Eqs. (6.50) and (6.49), the POVM for $y(t)$ is given by

$$d^2y\mathscr{A}\left\{p(y|\hat{a},\hat{a}^\dagger;t)\right\} = d^2y\int d^2\alpha p(y|\alpha,\alpha^*;t)\hat{E}_\alpha^Q, \tag{6.53}$$

where

$$p(y(t)|\alpha,\alpha^*;t) = \frac{\exp\left[-\frac{\left|y(t)-(1-e^{-\gamma t})\alpha^*\right|^2}{e^{-\gamma t}(1-e^{-\gamma t})}\right]}{\pi e^{-\gamma t}(1-e^{-\gamma t})}. \tag{6.54}$$

The probability density function for y is given by

$$p_{\hat{\rho}_0}^Y(y;t) = \int d^2\alpha p(y(t)|\alpha,\alpha^*;t)Q_{\hat{\rho}_0}(\alpha,\alpha^*), \tag{6.55}$$

where

$$Q_{\hat{\rho}}(\alpha,\alpha^*) = \frac{\langle\alpha|\hat{\rho}|\alpha\rangle}{\pi} = \mathrm{tr}[\hat{\rho}\hat{E}_\alpha^Q]$$

is the Q-function [27, 28]. Since $p(y(t)|\alpha,\alpha^*;t) \to \delta^2(y-\alpha^*)$ in the limit of $t \to \infty$, $p^Y_{\hat{\rho}}(y;t)$ reduces to $Q_{\hat{\rho}}(y^*,y)$ in the infinite-time limit. Thus the measurement outcome of y gives us the information about the Q-function [29].

To show the relative entropy conservation for the system's POVM $\hat{E}^Q_\alpha$, we confirm the conditions for Theorem 5.1.1. From Eq. (6.53) the condition (4.1) is satisfied. From Eq. (6.47) we have

$$\mu_0(dy)\hat{M}^\dagger_y(t)\hat{E}^Q_\alpha\hat{M}_y(t) = d^2y(t)q(\alpha,\alpha^*;y)\hat{E}^Q_{\tilde{\alpha}(\alpha,y)}, \tag{6.56}$$

where

$$\tilde{\alpha}(\alpha,y) = e^{-\frac{\gamma t}{2}}\alpha + y^*, \tag{6.57}$$

$$q(\alpha,\alpha^*;y) = e^{-\gamma t}p(y|\tilde{\alpha}(\alpha;y),\tilde{\alpha}^*(\alpha;y)). \tag{6.58}$$

Thus the condition (5.2) is satisfied. For any smooth function $F(\alpha,\alpha^*)$, we have

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

and the condition (5.3) is satisfied. Thus the assumptions for Theorem 5.1.1 are satisfied and therefore we obtain the relative-entropy conservation law

$$D(p^Y_{\hat{\rho}}(\cdot;t)||p^Y_{\hat{\sigma}}(\cdot;t)) = D(Q_{\hat{\rho}}||Q_{\hat{\sigma}}) - E_{\hat{\rho}_0}[D(Q_{\hat{\rho}_y}||Q_{\hat{\sigma}_y})], \tag{6.59}$$

where $\hat{\rho}_y$ is the post-measurement state for the given measurement outcome y.

In this measurement process, the Shannon entropy conservation does not hold again. In fact, the difference of the Shannon entropies is evaluated to be

$$\begin{aligned}
&h(Q_{\hat{\rho}}) - E_{\hat{\rho}}[h(Q_{\hat{\rho}_y})] \\
&= h(Q_{\hat{\rho}}) + \int d^2\tilde{\alpha}d^2yp(y|\tilde{\alpha},\tilde{\alpha}^*)Q_{\hat{\rho}}(\tilde{\alpha},\tilde{\alpha}^*)\ln\left(\frac{e^{-\gamma t}p(y|\tilde{\alpha},\tilde{\alpha}^*)Q_{\hat{\rho}}(\tilde{\alpha},\tilde{\alpha}^*)}{p^Y_{\hat{\rho}}(y)}\right) \\
&= -\gamma t + I_{\hat{\rho}}(Q:Y) \neq I_{\hat{\rho}}(Q:Y),
\end{aligned} \tag{6.60}$$

where $I_{\hat{\rho}}(Q:Y)$ is the mutual information between the measurement outcome and the system's observable $\hat{E}^Q_\alpha$. The term $-\gamma t$ again comes from the Jacobian of the transformation $\alpha \to \tilde{\alpha}(\alpha;y)$.

Chapter 7

Construction of a Conserved Observable

In the previous chapters, the quantum measurement Y and the system's observable X are first given independently, and then we examined the conditions for the relative-entropy conservation. Then it is natural to ask whether or not there exists an observable which satisfies the relative-entropy conservation law for a given measurement process. The answer is yes, with a relatively weak condition on the sample space of the measurement outcome. In this chapter, we construct such a relative-entropy-conserving observable for a given instrument. The meaning of the constructed observable corresponds to the measurement outcome of the infinite joint measurements of the given quantum measurement.

Before going to the detailed discussion, let us describe the idea of the construction in an informal manner. An observable X corresponding to a POVM outputs a measurement outcome x for a given input state $\hat{\rho}$. This is schematically represented by the following diagram.

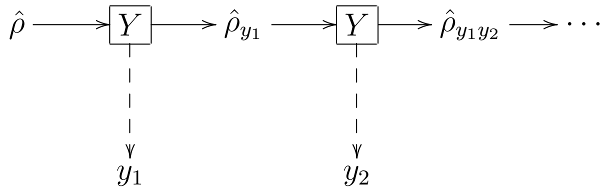
$$\hat{\rho} \longrightarrow \boxed{X} \dashrightarrow x$$

Here dashed line represents the classical outcome. On the other hand, a quantum measurement Y corresponding to a CP instrument outputs a measurement outcome y and the conditional post-measurement state $\hat{\rho}_y$ for a given input state $\hat{\rho}$ as the following diagram shows.

$$\begin{array}{ccccc} \hat{\rho} & \longrightarrow & \boxed{Y} & \longrightarrow & \hat{\rho}_y \\ & & \Big\downarrow & & \\ & & y & & \end{array}$$

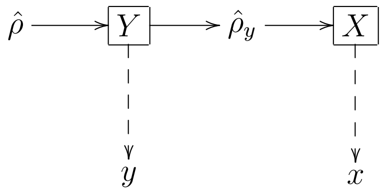
Let X be an observable corresponding to the measurement outcome of the infinite successive measurement of Y. Then X outputs the measurement outcome $x =$

$(y_1, y_2, \cdots)$, where y_k is the measurement outcome of k-th Y-measurement as represented by the following diagram.



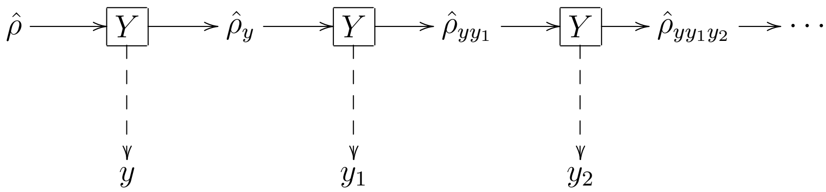
(7.1)

Let us consider the joint successive measurement of X following Y. This is represented by the following diagram.



(7.2)

Since X is represented by the diagram (7.1), the diagram (7.2) is equivalent to the following diagram.



(7.3)

The joint measurement (7.3) is equivalent to the single measurement of X (7.1) with the measurement outcome $(y, y_1, y_2, \cdots)$. Therefore information obtained from these measurement processes coincide and the relative-entropy conservation law (5.9) holds.

The above discussion assumed that X is a well-defined POVM. However, even in a simple case in which the sample space Ω_Y of Y is finite larger than 2, the sample space of X is the infinite product space $\Omega_X = \prod_{k=1}^{\infty} \Omega_k$ ($\Omega_k = \Omega_Y$) and Ω_X has the cardinality of the continuum. Therefore Ω_X is not a discrete space and we must specify the σ-algebra $\mathscr{B}_X$ on Ω_X and show that X is a well-defined POVM on $(\Omega_X, \mathscr{B}_X)$. The most of the following discussion is devoted to such measure theoretic considerations.

7.1 Mathematical preliminaries

7.1.1 Standard Borel space

For the construction of the conserved observable, we assume that the sample space of the measurement is a standard Borel space defined as follows.

Definition 7.1.1 (standard Borel space)

1. Let $(\Omega_1, \mathscr{B}_1)$ and $(\Omega_2, \mathscr{B}_2)$ be measurable spaces (cf. Definition 2.1.1). $(\Omega_1, \mathscr{B}_1)$ and $(\Omega_2, \mathscr{B}_2)$ are said to be **Borel isomorphic** iff there exists a bijection $f : \Omega_1 \to \Omega_2$ which is bimeasurable, i.e. for any $A_1 \in \Omega_1$ and any $A_2 \in \Omega_2$, $f^{-1}(A_2) \in \mathscr{B}_1$ and $f(A_1) \in \mathscr{B}_2$.

2. A topological space X is called a **Polish space** iff X is metrized by a complete metric d and X is separable, i.e. there exists a countable dense subset of X.

3. Let Ω be a topological space and $\mathscr{O}_\Omega$ be the family of open sets. We can define a σ-algebra $\mathscr{B}(\Omega)$ in a natural way by $\mathscr{B}(\Omega) := \sigma(\mathscr{O}_\Omega)$. Here $\sigma(\mathscr{A})$, called the generated σ-algebra of $\mathscr{A}$, denotes the smallest σ-algebra which contains a family $\mathscr{A}$ of subsets of Ω. In this sense, a topological space is considered as a measurable space.

4. A **standard Borel space** $(\Omega, \mathscr{B})$ is a measurable space which is Borel isomorphic to a Polish space.

Two Borel isomorphic measurable spaces are equivalent in the sense it is a relabelling of the measurement outcome.

A discrete space $(\Omega, 2^\Omega)$, where Ω is a countable set, is a standard Borel space. As an continuous example, the Euclidean space $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ is a standard Borel space. In this sense, the concept of the standard Borel space is so general as to include the sample spaces of the measurement outcome encountered in the physical problem.

Next we consider the products of measurable spaces. Let $(\Omega_i, \mathscr{B}_i)$ be a measurable space $(i = 1, 2, \cdots)$. For $n \geq 1$ we can define a σ-algebra $\mathscr{B}_1 \times \mathscr{B}_2 \times \cdots \times \mathscr{B}_n = \prod_{i=1}^n \mathscr{B}_i$ on a product set $\prod_{i=1}^n \Omega_i$ by the σ-algebra generated by a family of sets

$$\{A_1 \times A_2 \times \cdots \times A_n; A_i \in \mathscr{B}_i \, (i = 1, 2, \cdots, n)\},$$

an element of which is called a cylinder set. For an infinite product set $\tilde{\Omega} := \prod_{i=1}^\infty \Omega_i$, the product σ-algebra $\prod_{i=1}^\infty \mathscr{B}_i$ is the σ-algebra generated by a family of sets

$$\bigcup_{i=1}^\infty \{\pi_i^{-1}(A); A \in \mathscr{B}_i\},$$

where $\pi_i : \prod_{i=1}^\infty \Omega_i \to \Omega_i$ is the canonical projection. If $\Omega_i = \Omega$ and $\mathscr{B}_i = \mathscr{B}$ for each $i \geq 1$, the product spaces $(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathscr{B}_i)$ and $(\prod_{i=1}^\infty \Omega_i, \prod_{i=1}^\infty \mathscr{B}_i)$ are denoted by $(\Omega^n, \mathscr{B}^n)$ and $(\Omega^\mathbb{N}, \mathscr{B}^\mathbb{N})$, respectively.

Let $(\Omega_i, \mathscr{B}_i)$ be a standard Borel space for any $i \geq 1$. It is known that the product spaces $(\prod_{i=1}^{n} \Omega_i, \prod_{i=1}^{n} \mathscr{B}_i)$ and $(\prod_{i=1}^{\infty} \Omega_i, \prod_{i=1}^{\infty} \mathscr{B}_i)$ are standard Borel spaces [30].

7.1.2 Composition of instruments

In this section we consider a composition of two instruments $\mathcal{I}^1_\cdot(\cdot)$ and $\mathcal{I}^2_\cdot(\cdot)$ on sample spaces $(\Omega_1, \mathscr{B}_1)$ and $(\Omega_1, \mathscr{B}_1)$, respectively. The composition corresponds to a joint measurement of 1 after 2. If the sample spaces are discrete, this is given by

$$\mathcal{I}^{12}_A(\hat{\rho}) = \sum_{(\omega_1,\omega_2)\in A} \mathcal{I}^1_{\{\omega_1\}} \circ \mathcal{I}^2_{\{\omega_2\}}(\hat{\rho})$$

for each $A \subset 2^{\Omega_1 \times \Omega_2}$ and each density operator $\hat{\rho}$.

The composition of two instruments can be constructed when the sample spaces are both standard Borel spaces. The following theorem is due to Davies and Lewis [31, 32].

Theorem 7.1.2
Let $\mathcal{I}^i_\cdot(\cdot)$ be a CP instrument on a standard Borel space $(\Omega_i, \mathscr{B}_i)$ for $i = 1, 2$. Then there exists a unique CP instrument $\mathcal{I}^{12}_\cdot(\cdot)$ on the product space $(\Omega_1 \times \Omega_2, \mathscr{B}_1 \times \mathscr{B}_2)$ such that

$$\mathcal{I}^{12}_{A_1\times A_2}(\hat{\rho}) = \mathcal{I}^1_{A_1} \circ \mathcal{I}^2_{A_2}(\hat{\rho})$$

for any $A_i \in \mathscr{B}_i$ $(i = 1, 2)$ and any density operator $\hat{\rho}$. The constructed instrument $\mathcal{I}^{12}$ is denoted as $\mathcal{I}^1 * \mathcal{I}^2$.

We remark that the original statement in [31, 32] is for positive (P) instruments, the definition of which is obtained by weakening the complete positivity to a mere positivity in the definition of the CP instrument. The above statement is readily obtained if we note that a P instrument $\mathcal{I}$ is a CP instrument iff $\mathcal{I} \otimes \mathcal{I}_n$ is a P instrument for each $n \geq 1$, where $\mathcal{I}_n$ is the identity superoperator on $\mathcal{L}(\mathbb{C}^n)$.

If we ignore the post-measurement state after $\mathcal{I}^1$ and only consider the measurement outcome, we obtain the following theorem.

Theorem 7.1.3
Let $\hat{E}^1(\cdot)$ be a POVM on a standard Borel space $(\Omega_1, \mathscr{B}_1)$ and $\mathcal{I}^2_\cdot(\cdot)$ be a CP instrument on a standard Borel space $(\Omega_2, \mathscr{B}_2)$. Then there exists a unique POVM $\hat{E}^{12}(\cdot)$ on the product space $(\Omega_1 \times \Omega_2, \mathscr{B}_1 \times \mathscr{B}_2)$ such that

$$\mathrm{tr}[\hat{\rho}\hat{E}^{12}(A_1 \times A_2)] = \mathrm{tr}[\hat{E}^1(A_1)\mathcal{I}^2_{A_2}(\hat{\rho})] \tag{7.4}$$

for any $A_i \in \mathscr{B}_i$ $(i = 1, 2)$ and any density operator $\hat{\rho}$. We call the POVM $\hat{E}^{12}$ as the composed POVM of the POVM $\hat{E}^1$ and the CP instrument $\mathcal{I}^2$.

Proof. Let $\mathcal{I}^1_\cdot(\cdot)$ be a CP instrument on a sample space $(\Omega_1, \mathscr{B}_1)$ such that ${\mathcal{I}^1_A}^\dagger(\hat{I}) = \hat{E}^1(A)$ for any $A \in \mathscr{B}_1$. Such CP instrument is, for example, given by $\mathcal{I}^1_A(\hat{\rho}) = \mathrm{tr}[\hat{E}^1(A)\hat{\rho}]\hat{\rho}_1$ for any $A \in \mathscr{B}_1$ and any density operator $\hat{\rho}$, where $\hat{\rho}_1$ is an arbitrary density operator. Then the POVM $\hat{E}^{12}(A) := (\mathcal{I}^1 * \mathcal{I}^2)^\dagger_A(\hat{I})$ on the product sample space satisfies the condition (7.4). The uniqueness follows from the uniqueness of the measure $\mathrm{tr}[\hat{\rho}\hat{E}^{12}(\cdot)]$ on the product space for each $\hat{\rho}$. $\square$

Let $\mathcal{I}^i_\cdot(\cdot)$ be a CP instrument on a standard Borel space $(\Omega_i, \mathscr{B}_i)$ for $1 \le i \le n$. The composition of n CP instruments on a product space $(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathscr{B}_i)$ is defined by

$$\mathcal{I}^1 * \mathcal{I}^2 * \cdots * \mathcal{I}^n := \mathcal{I}^1 * (\mathcal{I}^2 * (\cdots * \mathcal{I}^n) \cdots). \tag{7.5}$$

This is a unique CP instrument on the product space such that

$$\left(\mathcal{I}^1 * \mathcal{I}^2 * \cdots * \mathcal{I}^n\right)_{\prod_{i=1}^n A_i} = \mathcal{I}^1_{A_1} \circ \mathcal{I}^2_{A_2} \circ \cdots \circ \mathcal{I}^n_{A_n}$$

for any $A_i \in \mathscr{B}_i$ $(1 \le i \le n)$. If $\Omega_i = \Omega$, $\mathscr{B}_i = \mathscr{B}$ and $\mathcal{I}^i = \mathcal{I}$ for all $1 \le i \le n$, the composition of instruments (7.5) is denoted as $\mathcal{I}^{*n}$.

7.2 Construction of a relative-entropy-conserving observable

For the construction of a relative-entropy-conserving POVM for a given instrument, we need the following lemma [33].

Lemma 7.2.1 (quantum Kolmogorov extension theorem)
Let $(\Omega_i, \mathscr{B}_i)$ be a standard Borel space $(i = 1, 2, \cdots)$ and $\hat{E}_n(\cdot)$ be a POVM on a product space $(\prod_{i=1}^n \Omega_i, \prod_{i=1}^n \mathscr{B}_i)$ for each n. Suppose that, for each $1 \le n < m$,

$$\hat{E}_n(A) = \hat{E}_m\left(A \times \prod_{n<i\le m} \Omega_i\right) \tag{7.6}$$

holds for any $A \in \prod_{i=1}^n \mathscr{B}_i$. Then there exists a unique POVM $\hat{E}(\cdot)$ on the infinite product space $(\prod_{i=1}^\infty \Omega_i, \prod_{i=1}^\infty \mathscr{B}_i)$ such that

$$\hat{E}_n(A) = \hat{E}\left(A \times \prod_{n<i<\infty} \Omega_i\right)$$

for any $n \ge 1$ and any $A \in \prod_{i=1}^n \mathscr{B}_i$.

The condition (7.6) is called **Kolmogorov consistency condition**. We remark that the classical version of Lemma 7.2.1 is called as a Kolmogorov extension theorem well known in the measure theoretic probability theory (e.g. [20]). The Kolmogorov extension theorem is used, for example, to construct a probability space of an infinite coin-toss process.

Now let us construct a POVM for a given CP instrument $\mathcal{I}^Y_\cdot(\cdot)$ on a standard Borel space $(\Omega_Y, \mathscr{B}_Y)$. Let $\hat{E}^Y_n$ be a POVM on a product space $(\Omega_Y^n, \mathscr{B}_Y^n)$ such that

$$\hat{E}^Y_n\left(\prod_{i=1}^n A_i\right) = (\mathcal{I}^{Y*n})^\dagger_{A_n\times\cdots A_2\times A_1}(\hat{I}) = \mathcal{I}^{Y\ \dagger}_{A_1} \circ \mathcal{I}^{Y\ \dagger}_{A_2} \cdots \circ \mathcal{I}^{Y\ \dagger}_{A_n}(\hat{I}) \tag{7.7}$$

for any $A_i \in \mathscr{B}_Y$ $(i = 1, 2, \cdots, n)$, which corresponds to the measurement outcome of an n-composition of $\mathcal{I}^Y$. Then it satisfies the Kolmogorov consistency condition (7.6) and from the quantum Kolmogorov extension theorem there exists a unique POVM $\hat{E}^X$ on an infinite product space $(\Omega_X, \mathscr{B}_X) = (\Omega_Y^{\mathbb{N}}, \mathscr{B}_Y^{\mathbb{N}})$ such that

$$\hat{E}^X(A \times \Omega_Y^{\mathbb{N}}) = \hat{E}^Y_n(A) \tag{7.8}$$

for any $A \in \mathscr{B}_Y^n$. The POVM $\hat{E}^X$ corresponds to the measurement outcome of the infinite composition of the given measurement process Y. Note that we cannot in general define the post-measurement *state* corresponding to the infinite composition of a CP instrument. An example of such a measurement is the quantum counter measurement described in Sec. 6.2 in which the photon number of the post-measurement state diverges in the infinite composition limit.

Next we prove the relative-entropy conservation law. Here we consider one of the equivalent forms of the conservation law in Eq. (5.9). From Theorem 7.1.3, there exists a POVM $\hat{E}^{XY}$ on a product space $(\Omega_X \times \Omega_Y, \mathscr{B}_X \times \mathscr{B}_Y)$ such that

$$\hat{E}^{XY}(A_X \times A_Y) = \mathcal{I}^{Y\ \dagger}_{A_Y}(\hat{E}^X(A_X)) \tag{7.9}$$

for any $A_X \in \mathscr{B}_X$ and $A_Y \in \mathscr{B}_Y$. The POVM $\hat{E}^{XY}$ corresponds to the outcome of the joint measurement process of X after Y. The sample space $\Omega_X \times \Omega_Y$ is Borel isomorphic to $\Omega_X = \Omega_Y^{\mathbb{N}}$ by a mapping

$$\Omega_X \times \Omega_Y \ni (x, y) \mapsto (y, x(1), x(2), \cdots) \in \Omega_X, \tag{7.10}$$

where $x = \prod_{i=1}^{\infty} x(i)$ and thus $\hat{E}^{XY}$ is identified with a POVM $\hat{E}^{\tilde{X}}(\cdot)$ on $(\Omega_X, \mathscr{B}_X)$ induced by the mapping (7.10). It follows from Eq. (7.9) that

$$\begin{aligned}
\hat{E}^{\tilde{X}}\left(\prod_{i=0}^{n} A_i \times \Omega_Y^{\mathbb{N}}\right) &= \mathcal{I}_{A_0}^{Y\,\dagger}\left(\hat{E}_n^Y\left(\prod_{i=0}^{n} A_i\right)\right) \\
&= \mathcal{I}_{A_0}^{Y\,\dagger} \circ \mathcal{I}_{A_1}^{Y\,\dagger} \circ \cdots \circ \mathcal{I}_{A_n}^{Y\,\dagger}(\hat{I}) \\
&= \hat{E}^X\left(\prod_{i=0}^{n} A_i \times \Omega_Y^{\mathbb{N}}\right)
\end{aligned}$$

for every $A_0, A_1, \cdots, A_n \in \mathscr{B}_Y$ ($n \geq 1$). Thus from the uniqueness of $\hat{E}^X$ we have $\hat{E}^{\tilde{X}} = \hat{E}^X$. Thus we have the relative-entropy conservation law

$$D(p_{\hat{\rho}}^{XY}||p_{\hat{\sigma}}^{XY}) = D(p_{\hat{\rho}}^{X}||p_{\hat{\sigma}}^{X}), \tag{7.11}$$

where $p_{\hat{\rho}}^{XY}$ and $p_{\hat{\rho}}^{X}$ are the probability measures for a quantum state $\hat{\rho}$ defined by

$$\begin{aligned}
p_{\hat{\rho}}^{XY}(A) &= \mathrm{tr}[\hat{\rho}\hat{E}^{XY}(A)], \\
p_{\hat{\rho}}^{X}(B) &= \mathrm{tr}[\hat{\rho}\hat{E}^{X}(B)]
\end{aligned}$$

for each $A \in \mathscr{B}_X \times \mathscr{B}_Y$ and $B \in \mathscr{B}_X$.

The above discussion is summarized as the following theorem.

Theorem 7.2.2

Let Y be a measurement process described by a CP instrument $\mathcal{I}^Y$ on a standard Borel space $(\Omega_Y, \mathscr{B}_Y)$. Then there exists a unique system's observable X described by a POVM $\hat{E}^X$ on a product sample space $(\Omega_X, \mathscr{B}_X) = (\Omega_Y^{\mathbb{N}}, \mathscr{B}_Y^{\mathbb{N}})$ such that the condition (7.8) holds. Furthermore, X and Y satisfy the relative-entropy conservation law (7.11).

We remark that the mapping (7.10) corresponds to $\tilde{x}(x; y)$ in Chapter 5. We also mention that in the example of the quantum-counter model in Sec. 6.2, the POVM in Eq. (6.17) corresponds to the constructed observable (7.8).

Chapter 8

Summary

We have studied information flow in quantum measurement processes based on the relative entropy and identified the conditions for the relative-entropy conservation law. In this chapter we summarize the results obtained in this thesis.

In Chapter 1, we described the motivation and background of this thesis. In Chapter 2, we reviewed the quantum theory of measurement. There we have introduced the POVM, the CP instrument and the measurement model. In the introduction of these concepts, the sample space of the measurement outcome is described by a general measurable space, which enables us to handle discrete and continuous sample spaces in a consistent manner.

In Chapter 3, we have reviewed the classical information theory and introduced classical entropic information contents, namely the Shannon entropy, the mutual information and the relative entropy. We have discussed properties of these entropies for a continuous variable and shown that the Shannon entropy for a continuous variable cannot be interpreted as an information content and depends on the choice of a reference measure of the variable, whereas the other two entropies are independent of the reference measure. Furthermore, we have introduced a concept of a sufficient statistic and seen that the sufficiency of a statistic is characterized by a conservation of the relative entropy.

In Chapter 4, we reviewed a Shannon entropy conservation for quantum measurements established by Ban [10]. For a given system's observable X described by a POVM and a measurement process Y described by a CP instrument, we have proved the Shannon entropy conservation by assuming some conditions on X and Y. We have discussed a special case in which X is projection-valued and shown that the Shannon entropy conservation can be proven under a less restrictive condition than the general case. As an example of Shannon-entropy-conserving measurements, we have discussed a quantum non-demolition measurement.

In Chapter 5, we have derived the condition for a quantum measurement Y and a system's observable X such that the relative-entropy conservation law holds.

The obtained information about which of the two candidate states is actually prepared is quantified in terms of the relative entropy of the measurement outcome. The derived relative-entropy conservation law states that the relative entropy of the measurement outcome is equal to the decrease in the relative entropy of the measured observable X. To clarify the meaning of the established condition for the relative entropy conservation law, we have considered the successive joint measurement of Y followed by X, and derived another equivalent relative-entropy conservation law for this joint measurement process, which states that the relative entropy of the measurement outcome of the joint measurement coincides with the relative entropy of the measured observable X for the initial state. The established condition for the relative-entropy conservation law can be interpreted as the existence of a sufficient statistic $\tilde{x}(x; y)$ in the successive joint measurement such that the probability distribution of $\tilde{x}(x; y)$ coincides with that of X for the initial state. We have also shown that for the case in which Y is discrete and X is a discrete PVM, the relative-entropy conservation law is equivalent to the established condition. The established condition is less restrictive than the condition for the Shannon entropy conservation derived by Ban and we have compared these conditions. For the case in which X and Y are both discrete, Ban's condition is shown to be equivalent to the condition that the post-measurement state is an eigenstate of X if the pre-measurement state is an eigenstate of X. An example in which the Shannon entropy conservation does not hold but the relative-entropy conservation does is given by a destructive sharp measurement of X, in which the measurement outcome is equivalent to that of the projective measurement while the post-measurement state is a maximally mixed state.

In Chapter 6, we have applied the general theorem for the relative-entropy conservation law to typical examples of optical continuous measurements, namely photon-counting, quantum-counting, homodyne and heterodyne measurements, and shown the relative-entropy conservation for each measurement process. We have shown that these measurements except for the photon-counting measurement do not satisfy the Shannon entropy conservation due to the non-unit Jacobian of the transformation $x \to \tilde{x}(x; y)$. Among these measurements, the heterodyne measurement is special in the sense that the probability for the measurement outcome, Q-function, involves all the matrix elements of the system's density matrix as different from other examples in which the diagonal elements are only relevant to the measurement process.

In Chapter 7, we have constructed a relative-entropy-conserving observable for a given measurement process described by a CP instrument on a standard Borel space. The constructed observable is an infinite composition of the given instrument which corresponds to the measurement outcome of an infinite successive joint measurement of the given measurement. For the quantum-counting measurement, there are two relative-entropy-conserving observables, namely the photon number and the infinite composition of the measurement constructed in Chapter 7. In this example the latter one is obtained by the coarse-graining of the photon number. From this observation, one may wonder what is the relation between the constructed observable and other relative-entropy-conserving observables in general. This question remains an outstanding issue.

Appendix A

Proof of Proposition 3.3.4

In this appendix we prove Proposition 3.3.4.

Lemma A.0.3
If $\mathcal{H}$ is separable, then $\mathcal{S}(\mathcal{H})$ is separable with respect to the trace norm.

Proof. Since $\mathcal{H}$ is separable, there exists a countable dense subset $\{|\psi_n\rangle\} \subset \mathcal{H}$. Define $\mathcal{T}_0$ as a set of operators of the form

$$\sum_{k=1}^{m} |\psi_{n_k}\rangle \langle\psi_{n_k}| \quad (m = 1, 2, \cdots)$$

and let $\mathcal{S}_0 := \{\hat{\rho}/\operatorname{tr}[\hat{\rho}]; \hat{\rho} \in \mathcal{T}_0\}$. Then $\mathcal{S}_0$ is countable and it is easy to show that $\mathcal{S}_0$ is dense in $\mathcal{S}(\mathcal{H})$. □

From Lemma A.0.3, there exists a dense countable subset $\{\hat{\rho}_n\}_{n\geq 1} \subset \mathcal{S}(\mathcal{H})$. Define a state $\hat{\rho}_0 \in \mathcal{S}(\mathcal{H})$ by

$$\hat{\rho}_0 := \sum_{n\geq 1} 2^{-n} \hat{\rho}_n. \tag{A.1}$$

To show $P^X_{\hat{\rho}} \ll P^X_{\hat{\rho}_0}$ for each $\hat{\rho} \in \mathcal{S}(\mathcal{H})$, take an arbitrary set $A \in \mathscr{B}_X$ such that $P^X_{\hat{\rho}_0}(A) = 0$. From the definition of $\hat{\rho}_0$ (A.1), we have

$$0 = \sum_{n\geq 1} 2^{-n} \operatorname{tr}[\hat{\rho}_n \hat{E}^X(A)] = \sum_{n\geq 1} 2^{-n} P^X_{\hat{\rho}_n}(A), \tag{A.2}$$

and thus $P^X_{\hat{\rho}_n}(A) = 0$ for all $n \geq 1$. Since $\{\hat{\rho}_n\}_{n\geq 1}$ is dense in $\mathcal{S}(\mathcal{H})$, there exists a subsequence $\{\hat{\rho}_{n_k}\}_{k\geq 1}$ such that $||\hat{\rho} - \hat{\rho}_{n_k}||_1 \to 0$, where $||\hat{A}||_1 := \operatorname{tr}\sqrt{\hat{A}^\dagger \hat{A}}$ is the trace norm. Hence we have

$$P^X_{\hat{\rho}}(A) = \operatorname{tr}[\hat{\rho}\hat{E}^X(A)] = \lim_{k\to\infty} \operatorname{tr}[\hat{\rho}_{n_k} \hat{E}^X(A)] = 0. \tag{A.3}$$

Thus $P^X_{\hat{\rho}} \ll P^X_{\hat{\rho}_0}$ holds and Proposition 3.3.4 is proved.

Appendix B

Derivation of Eqs. (6.37) and (6.39)

We at first evaluate $\hat{M}_y^\dagger(t) \left|x\right\rangle_1 {}_1\left\langle x\right| \hat{M}_y(t)$. In the evaluation of this operator, we use the normal ordering. The normally ordered expression $:O(\hat{a},\hat{a}^\dagger):$ for a scalar function $O(\alpha,\alpha^*)$ of a complex variable α is an operator in which annihilation operators are placed to the right of creation operators. Any operator $\hat{O}$ has a unique normally ordered expression $:O(\hat{a},\hat{a}^\dagger):$ with

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

where

$$\left|\alpha\right\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \left|n\right\rangle \tag{B.1}$$

is a coherent state. Since the coherent state in the $\left|x\right\rangle_1$ basis can be written as

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

we have

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

which implies the following normally ordered expression

$$\left|x\right\rangle_1 {}_1\left\langle x\right| = \pi^{-1/2} : \exp\left[-\left(x-\frac{\hat{a}+\hat{a}^\dagger}{\sqrt{2}}\right)^2\right] : . \tag{B.2}$$

By using Eq. (B.2) and the formula

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

which is valid for real λ, the operator $\hat{M}_y^\dagger(t) |x\rangle_{1\,1} \langle x| \hat{M}_y(t)$ in normal ordering is evaluated to be

$$\begin{aligned}
&\langle\alpha| \hat{M}_y^\dagger(t) |x\rangle_{1\,1} \langle x| \hat{M}_y(t) |\alpha\rangle \\
&= \pi^{-1/2} \exp\left[-\left(e^{-\frac{\gamma t}{2}} x + \frac{y}{\sqrt{2}} - \frac{\alpha+\alpha^*}{\sqrt{2}} \right)^2 + \left(e^{-\frac{\gamma t}{2}} x + \frac{y}{\sqrt{2}} \right)^2 - x^2 \right].
\end{aligned} \tag{B.3}$$

Using again Eq. (B.2) in Eq. (B.3), we obtain Eq. (6.39). By integrating Eq. (6.39) with respect to x, we obtain

$$\hat{M}_y(t)^\dagger \hat{M}_y(t) = \exp\left[\frac{\gamma t}{2} + \hat{X}_1^2 - e^{\gamma t} \left(\hat{X}_1 - \frac{y}{\sqrt{2}} \right)^2 \right]. \tag{B.4}$$

By multiplying Eq. (B.4) with Eq. (6.36), we obtain Eq. (6.37).

Acknowledgements

I would like to express my sincere gratitude to Prof. Masahito Ueda, my supervisor, for his guidance, helpful comments, and continuous encouragements during research. I would also like to thank Tomohiro Shiatara and Prof. Masato Koashi for helpful discussions. I would like to thank Prof. Mio Murao, Prof. Akira Shimizu, Prof. Yuji Tachikawa, Prof. Masaki Sano, and Prof. Izumi Tsutui for refereeing my thesis and for valuable discussions.

This research has been supported by the the Leading Graduates Schools Program “Advanced Leading Graduate Course for Photon Science” and the JSPS Research Fellowships for Young Scientists DC2 (Grant No. 14J09905).

Bibliography

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

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

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

[4] A. Shimizu and K. Fujita, First-kind measurements, non-demolition measurements, and conservation laws, in *Quantum control and measurement: proceedings of the ISQM satellite workshop, ARL, Hitachi, Hatoyama, Saitama, August 28-29, 1992*, pp. 191–196, Amsterdam, 1993, North-Holland.

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

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

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

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

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

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

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

[12] Y. Kuramochi and M. Ueda, Classicality condition on a system's observable in a quantum measurement and relative-entropy conservation law, 2014, arXiv:1406.2130.

[13] P. R. Halmos, *Measure Theory* (Springer, 1974).

[14] W. F. Stinespring, Proc. Amer. Math. Soc. **6**, 211 (1955).

[15] K. Kraus, Ann. Phys. (NY) **64**, 311 (1971).

[16] M. Ozawa, J. Math. Phys. **25**, 79 (1984).

[17] C. E. Shannon, Bell Sys. Tech. J. **27**, 379 (1948).

[18] C. E. Shannon, Bell Sys. Tech. J. **27**, 623 (1948).

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

[20] K. Itô, *An Introduction to Probability Theory* (Cambridge University Press, 1984).

[21] P. R. Halmos and L. J. Savage, Ann. Math. Stat. **20**, 225 (1949).

[22] M. Ueda and M. Kitagawa, Phys. Rev. Lett. **68**, 3424 (1992).

[23] M. Ueda, N. Imoto, and H. Nagaoka, Phys. Rev. A **53**, 3808 (1996).

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

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

[26] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2010).

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

[30] S. M. Srivastava, *A course on Borel sets* (Springer, 1998).

[31] E. B. Davies and J. Lewis, Commun. Math. Phys. **17**, 239 (1970).

[32] E. B. Davies, *Quantum theory of open systems* (IMA, 1976).

[33] R. Tumulka, Lett. Math. Phys. **84**, 41 (2008).