State reduction, information and entropy in quantum measurement processes

1999 J. Phys. A: Math. Gen. 32 1643

(http://iopscience.iop.org/0305-4470/32/9/012)

J. Phys. A: Math. Gen. **32** (1999) 1643–1665. Printed in the UK PII: S0305-4470(99)97600-X

State reduction, information and entropy in quantum measurement processes

Masashi Ban
Advanced Research Laboratory, Hitachi Ltd, Akanuma 2520, Hatoyama, Saitama 350-0395, Japan

Received 11 September 1998, in final form 11 November 1998

Abstract. Information obtained by a quantum measurement process performed on a physical system and the entropy change of the measured physical system are considered in detail. It is shown that the condition for the amount of information obtained by the quantum measurement process to be represented by the Shannon mutual entropy is that the intrinsic observable of the measured physical system commutes with the operational observable defined by the quantum measurement process. When some measurement outcome is obtained, the decrease of the Shannon entropy of the measured system is compared with that of the von Neumann entropy. Furthermore, a condition is established under which the amount of information that can be established by the quantum measurement process becomes equal to the decrease of the Shannon entropy of the measured physical system.

1. Introduction

Entropy is one of the most important quantities not only in thermodynamics and statistical mechanics [1, 2] but also in information theory and statistics [3, 4]. In statistical physics, the second law of thermodynamics and the irreversibility of dynamical processes are characterized by the increase of entropy [5]. In thermal equilibrium, the entropy of a physical system is given by the celebrated Boltzmann formula which clarifies the relation between thermodynamic entropy and probability. Letting p_j be the probability that the physical system is in the jth microscopic state and letting W be the total number of all possible microscopic states of the physical system under some constraint; then the entropy of the physical system is given by the formula, $H(p) = -\sum_{j=1}^{W} p_j \log p_j$, where we set the Boltzmann constant $k_B = 1$. In this paper, we use the logarithm with arbitrary base. When we use the natural logarithm, the entropy is measured in *nats* and when the logarithm with base two is applied, the entropy is measured in *bits*. The maximum value of the entropy becomes $H_{max} = \max_{\{p_j\}} H(p) = \log W$ which is the Boltzmann formula. The thermal equilibrium state of the physical system is derived, according to the entropy-maximum principle [6, 7]. The maximum value of the entropy is attained by $p_j = 1/W$ for all j. This is equivalent to the principle of equal *a priori* probabilities for a microcanonical ensemble in thermal equilibrium [1].

The relation between entropy and information was considered first by Szilard [8] to investigate the Maxwell demon in a thermodynamic system. He showed that the information gain by the measurement process performed on the thermodynamic system decreases the entropy of the measured physical system. Of course, the total entropy of the measured physical system and the measurement apparatus increases due to the second law of thermodynamics. The significance of this work was pointed out by Brillouin [9]. The most important work that showed the clear relation between entropy and information was done by Shannon [10–12] who introduced entropy, conditional entropy and mutual entropy called the Shannon entropies, into communication theory. He showed that the average length of a code word representing a symbol generated from a message source is lower bounded by the Shannon entropy of the message source. Furthermore, he found that the information can be reliably transmitted through a noisy communication channel if the information rate is less than the channel capacity which is the maximum value of the mutual entropy of the communication channel. The former is referred to as the source-coding theorem and the latter is called the channel-coding theorem. The information provided by experiments was investigated based on Shannon information theory by Lindley [13]. Other interesting works that consider the relations between entropy, information and randomness (algorithmic complexity or Kolmogorov complexity) of physical systems have been done by Zurek [14, 15] and Caves [16]. The quantum Maxwell demon has also been investigated by Lloyd [17].

Quantum mechanical entropy was introduced by von Neumann [18] in the quantum theory of measurement processes. The quantum mechanical entropy $S(\hat{\rho}_S)$, called von Neumann entropy, is given by the formula, $S(\hat{\rho}_S) = -\mathrm{Tr}_S(\hat{\rho}_S \log \hat{\rho}_S)$, where $\hat{\rho}_S$ represents a statistical operator which describes a quantum state of a physical system and Tr_S stands for the trace operation over the Hilbert space $\mathcal{H}_S$ of a physical system. When the physical system is prepared in a quantum state $\hat{\rho}_S$, an observable $\hat{X}_S$ of the physical system, which has the eigenstate $|\psi_S(x)\rangle$ with eigenvalue x, takes the value x with probability $P_X(x) = \langle\psi_S(x)|\hat{\rho}_S|\psi_S(x)\rangle$. Then the Shannon entropy of the observable $\hat{X}_S$ in the quantum state $\hat{\rho}_S$ is given by $H(\hat{X}_S) = -\sum_{x\in\Omega_X} P_X(x) \log P_X(x)$, where Ω_X is the spectral set of the observable $\hat{X}_S$. In some cases, this entropy is called the measurement entropy [19]. The Shannon entropy is no less than the von Neumann entropy, that is, $S(\hat{\rho}_S) \leqslant H(\hat{X}_S)$, which is derived from the concavity of the function $f(x) = -x \log x$ [20]. It has recently been found that the von Neumann entropy in quantum information theory [21, 22] plays the same role as Shannon entropy does in classical information theory [3, 4]. Quantum-source-coding theorem has proved that the average number of quantum bits (two-level quantum systems) representing a pure quantum state generated from a quantum message source is lower bounded by the von Neumann entropy of the message source [23, 24]. Furthermore, quantum-channel-coding theorem has found that the information can be reliably transmitted through a noisy quantum channel if the information rate is less than the quantum channel capacity, sometimes called the Holevo bound [25–27].

In this paper, the information gain and the entropy change by a quantum measurement process performed on a physical system will be considered in order to understand the information-theoretical properties of quantum measurement processes. When some quantum measurement is performed on a physical system, the quantum state of the measured physical system inevitably changes due to the effects of the quantum measurement process. Any quantum measurement process that does not disturb the quantum state of the measured physical system gives us no information about the physical system. The state change of the measured physical system induces the changes of the Shannon entropy and the von Neumann entropy. Therefore, it is considered that there is a relation between the amount of information on the physical system that can be obtained by the quantum measurement process and the entropy change of the measured physical system that is caused by the quantum measurement process. In particular, the condition for quantum measurement processes is investigated, under which the amount of information about the physical system is equal to the decrease of the Shannon entropy of the measured physical system. The information gain by the quantum measurement process and the entropy change of the measured physical system are very important in quantum information theory [21, 22]. Although optimization of quantum measurement processes such that the information gain is maximized or the error probability is minimized has been investigated in detail [28–33], the relation between information gain and entropy change has rarely been considered in quantum information theory. The relation between information gain and quantum state disturbance, however, plays an important role in quantum cryptography [34–36] and in error corrections of quantum computation [37–41]. Thus, the results obtained in this paper may have some relevance to quantum information theory.

This paper is organized as follows. In section 2, we briefly summarize the basic formulation of quantum measurement processes [42–46] in a convenient way for our purpose. The state-reduction formula for the measured physical system and the probability distributions of the observable of the physical system and the measurement outcome are given. In section 3, we consider information about the physical system obtained by the quantum measurement process and we obtain the condition under which the information gain can be represented by the Shannon mutual entropy. In section 4, when we obtain the measurement outcome, we compare the decrease of the Shannon entropy of the measured physical system with that of von Neumann entropy in the quantum measurement process. In section 5, we investigate the relation between information gain and entropy decrease and we obtain the condition under which the amount of information obtained by the quantum measurement process becomes equal to the decrease of the Shannon entropy of the measured physical system. In section 6, to examine the general results obtained in sections 3–5, we consider the information gain and the entropy change in the standard position measurement of the physical system and the photon number measurement by means of a lossless beam splitter. In section 7, we summarize the results obtained in this paper.

2. Quantum measurement processes and state reductions

In this section, we briefly review the state-reduction formula for a measured physical system and the probability distribution of measurement outcomes [42–46] and then we introduce the Shannon and von Neumann entropies in quantum measurement processes. An observable $\hat{X}_S$ of a physical system $\mathcal{S}$ on which we perform a quantum measurement can be characterized by a projection-valued measure (PVM) or a spectral measure $\hat{\mathcal{X}}_S(E_X)$ [31, 47] which is expressed in the following form

$$\hat{\mathcal{X}}_S(E_X) = \int_{x\in E_X} \mathrm{d}\mu(x)|\psi_S(x)\rangle\langle\psi_S(x)| \equiv \int_{x\in E_X} \mathrm{d}\mu(x)\hat{\mathcal{X}}_S(x) \tag{1}$$

where $|\psi_S(x)\rangle$ is an orthogonal eigenstate of the observable $\hat{X}_S$ and E_X is an arbitrary subset of the spectral set Ω_X which represents the set of all possible values of the observable $\hat{X}_S$. We can formally write $\hat{\mathcal{X}}_S(x) = \delta\hat{\mathcal{X}}_S(\Omega_X)/\delta\mu(x)$. The PVM $\hat{\mathcal{X}}_S(E_X)$ satisfies the relations

$$\hat{\mathcal{X}}_S(E_X) \geqslant 0 \qquad \hat{\mathcal{X}}_S(\Omega_X) = \hat{I}_S \tag{2}$$

for any subset $E_X \subseteq \Omega_X$. Here $\hat{I}_S$ is an identity operator defined on the Hilbert space $\mathcal{H}_S$ of the physical system. Furthermore the PVM satisfies the equalities

$$\hat{\mathcal{X}}_S(E_X)\hat{\mathcal{X}}_S(F_X) = \hat{\mathcal{X}}_S(E_X \cap F_X) \tag{3}$$

for any disjoined subsets $E_X, F_X \subseteq \Omega_X$. We will consider quantum measurement processes for both continuous and discrete observables in a systematic way. When we investigate a quantum measurement process for a discrete observable, we set $\mathrm{d}\mu(x) = \sum_{x_j\in\Omega_X} \delta(x - x_j)\,\mathrm{d}x$ in equation (1) and we obtain the PVM of the discrete observable

$$\hat{\mathcal{X}}_S(E_X) = \sum_{x_j\in E_X} |\psi_S(x_j)\rangle\langle\psi_S(x_j)|. \tag{4}$$

A function $f(\hat{X}_S)$ of the observable is a Hermitian operator defined by

$$f(\hat{X}_S) = \int_{x\in\Omega_X} \mathrm{d}\mu(x) f(x) |\psi_S(x)\rangle\langle\psi_S(x)| = \int_{x\in\Omega_X} \mathrm{d}\mu(x) f(x) \hat{\mathcal{X}}_S(x) \qquad (5)$$

which satisfies the eigenvalue equation $f(\hat{X}_S)|\psi_S(x)\rangle = f(x)|\psi_S(x)\rangle$. If the function $f(x)$ is expanded as $f(x) = \sum_n f_n x^n$, we obtain the equality $f(\hat{X}_S) = \sum_n f_n \hat{X}_S^n$. The mathematically rigorous treatment of quantum measurement processes of continuous observables was formulated by Ozawa [44]. The information and entropy for generalized observables which cannot be represented by PVMs will be considered in the appendix.

Suppose that before we perform the quantum measurement, the physical system to be measured is in a quantum state described by a statistical operator $\hat{\rho}_{\mathrm{in}}^S$ which is a non-negative and trace one operator defined on the Hilbert space $\mathcal{H}_S$. Then, the probability $\mathcal{P}_{\mathrm{in}}^S(E_X)$ that the observable $\hat{X}_S$ in the quantum state $\hat{\rho}_{\mathrm{in}}^S$ takes some value belonging to the set E_X is given by

$$\mathcal{P}_{\mathrm{in}}^S(E_X) = \mathrm{Tr}_S[\hat{\mathcal{X}}_S(E_X)\hat{\rho}_{\mathrm{in}}^S] = \int_{x\in E_X} \mathrm{d}\mu(x) \langle\psi_S(x)|\hat{\rho}_{\mathrm{in}}^S|\psi_S(x)\rangle \qquad (6)$$

where Tr_S stands for the trace operation over the Hilbert space $\mathcal{H}_S$ of the physical system. Here we introduce a probability function

$$P_{\mathrm{in}}^S(x) = \langle\psi_S(x)|\hat{\rho}_{\mathrm{in}}^S|\psi_S(x)\rangle = \mathrm{Tr}_S[\hat{\mathcal{X}}_S(x)\hat{\rho}_{\mathrm{in}}^S] \qquad (7)$$

which is formally rewritten as

$$P_{\mathrm{in}}^S(x) = \frac{\delta\mathcal{P}_{\mathrm{in}}^S(\Omega_X)}{\delta\mu(x)} = \mathrm{Tr}_S\left[\frac{\delta\hat{\mathcal{X}}_S(\Omega_X)}{\delta\mu(x)}\hat{\rho}_{\mathrm{in}}^S\right]. \qquad (8)$$

It is easy to see that the probability function $P_{\mathrm{in}}^S(x)$ is equal to the probability itself for a discrete observable and to the probability density for a continuous observable. In terms of the probability function $P_{\mathrm{in}}^S(x)$, we define the Shannon entropy of the observable $\hat{X}_S$ of the physical system in the quantum state $\hat{\rho}_{\mathrm{in}}^S$ by

$$H(X_{\mathrm{in}}^S) = -\int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{\mathrm{in}}^S(x) \log P_{\mathrm{in}}^S(x) \qquad (9)$$

which becomes the differential entropy for a continuous observable

$$H(X_{\mathrm{in}}^S) = -\int_{x\in\Omega_X} \mathrm{d}x \langle\psi_S(x)|\hat{\rho}_{\mathrm{in}}^S|\psi_S(x)\rangle \log\langle\psi_S(x)|\hat{\rho}_{\mathrm{in}}^S|\psi_S(x)\rangle \qquad (10)$$

and the usual entropy for a discrete observable

$$H(X_{\mathrm{in}}^S) = -\sum_{x_j\in\Omega_X} \langle\psi_S(x_j)|\hat{\rho}_{\mathrm{in}}^S|\psi_S(x_j)\rangle \log\langle\psi_S(x_j)|\hat{\rho}_{\mathrm{in}}^S|\psi_S(x_j)\rangle. \qquad (11)$$

It should be noted that the differential entropy can take negative values [4]. On the other hand, the von Neumann entropy of the physical system in the quantum state $\hat{\rho}_{\mathrm{in}}^S$ is given by

$$S(\hat{\rho}_{\mathrm{in}}^S) = -\mathrm{Tr}_S[\hat{\rho}_{\mathrm{in}}^S \log \hat{\rho}_{\mathrm{in}}^S]. \qquad (12)$$

To measure the observable $\hat{X}_S$ of the physical system, we must first prepare a measurement apparatus $\mathcal{A}$, the initial state of which is given by a statistical operator $\hat{\rho}_{\mathrm{in}}^A$ defined on the Hilbert space $\mathcal{H}_A$ of the measurement apparatus. We then interact the measurement apparatus $\mathcal{A}$ with the physical system $\mathcal{S}$ to make some quantum correlation between them. Here let $\hat{\mathcal{U}}_{SA}$ be a unitary operator that describes the state of change of the physical system and the measurement apparatus, which is caused by the system–apparatus interaction

$$\hat{\mathcal{U}}_{SA} = \mathrm{T}\exp\left\{-\frac{\mathrm{i}}{\hbar}\int_0^\tau \mathrm{d}t[\hat{H}_0^S(t)\otimes\hat{I}_A + \hat{I}_S\otimes\hat{H}_0^A(t) + \hat{H}_{\mathrm{int}}^{SA}(t)]\right\} \qquad (13)$$

where $\hat{H}_0^S(t)$ and $\hat{H}_0^A(t)$ are the individual Hamiltonians of the physical system and the measurement apparatus, which may depend on time and $\hat{H}_{\rm int}^{SA}(t)$ is the system–apparatus interaction Hamiltonian which acts only during the interaction time $\tau_{\rm int}$ ($0 < \tau_{\rm int} < \tau$). In this equation, the symbol 'T' stands for taking the chronological ordering of operators. Just before the readout of the measurement outcome, the compound quantum state of the physical system and the measurement apparatus is given by a statistical operator

$$\hat{\rho}_{\rm out}^{SA} = \hat{\mathcal{U}}_{SA}(\hat{\rho}_{\rm in}^S \otimes \hat{\rho}_{\rm in}^A)\hat{\mathcal{U}}_{SA}^\dagger. \tag{14}$$

Although we have considered the case where the state change can be described by the unitary operator, the results obtained in this paper are still valid even if the state change is given by a completely positive and trace-preserving map [42–46] which includes a unitary transformation as a special case. When the state change $\hat{\rho}_{\rm in}^S \otimes \hat{\rho}_{\rm in}^A \to \hat{\rho}_{\rm out}^{SA}$ is a complete positive and trace-preserving map, we can express the statistical operator $\hat{\rho}_{\rm out}^{SA}$ by introducing an appropriate environmental system [43]

$$\hat{\rho}_{\rm out}^{SA} = {\rm Tr}_E[\hat{\mathcal{U}}_{SA{\rm E}}(\hat{\rho}_{\rm in}^S \otimes \hat{\rho}_{\rm in}^A \otimes \hat{\rho}_{\rm in}^{\rm E})\hat{\mathcal{U}}_{SA{\rm E}}^\dagger] \tag{15}$$

where $\hat{\mathcal{U}}_{SA{\rm E}}$ is a unitary operator, $\hat{\rho}_{\rm in}^{\rm E}$ is the initial state of the environmental system and ${\rm Tr}_{\rm E}$ is the trace operation of the environmental system. In such a case, we substitute $\hat{\mathcal{U}}_{SA{\rm E}}$, $\hat{\rho}_{\rm in}^A \otimes \hat{\rho}_{\rm in}^{\rm E}$, $\hat{\mathcal{Y}}_A \otimes \hat{I}_{\rm E}$ and ${\rm Tr}_{A{\rm E}} = {\rm Tr}_A{\rm Tr}_{\rm E}$ for $\hat{\mathcal{U}}_{SA}$, $\hat{\rho}_{\rm in}^A$, $\hat{\mathcal{Y}}_A$ and ${\rm Tr}_A$ in the results obtained in this paper.

We finally perform the readout of the result of the quantum measurement process. The readout of the measurement outcome is described by a positive operator-valued measure (POVM) defined on the Hilbert space $\mathcal{H}_A$ of the measurement apparatus. Here let $\hat{\mathcal{Y}}_A(E_Y)$ be the POVM that describes the readout process, the outcome of which belongs to a set E_Y ($\subseteq \Omega_Y$), where Ω_Y represents the set of all possible outcomes of the quantum measurement process [30, 31]. We assume that the POVM $\hat{\mathcal{Y}}_A(E_Y)$ can be expressed in the following form

$$\hat{\mathcal{Y}}_A(E_Y) = \int_{y\in E_Y} {\rm d}\nu(y)\hat{\mathcal{Y}}_A(y) \tag{16}$$

which satisfies the relations

$$\hat{\mathcal{Y}}_A(E_Y) \geqslant 0 \qquad \hat{\mathcal{Y}}_A(\Omega_Y) = \hat{I}_A \tag{17}$$

for any subset $E_Y \subseteq \Omega_Y$. Since the operator $\hat{\mathcal{Y}}_A(E_Y)$ is a POVM but not a PVM in general, the relation $\hat{\mathcal{Y}}_A(E_Y)\hat{\mathcal{Y}}_A(F_Y) = \hat{\mathcal{Y}}_A(E_Y \cap F_Y)$ does not necessarily hold even if $E_Y \cap F_Y = \emptyset$. When we consider the quantum measurement process for a discrete observable, we set ${\rm d}\nu(y) = \sum_{y_j\in\Omega_Y} \delta(y - y_j)\,{\rm d}y$ in equation (16). If the POVM $\hat{\mathcal{Y}}_A(E_Y)$ becomes a PVM, the readout of the measurement outcome is equivalent to measuring the pointer observable $g(\hat{Y}_A)$ of the measurement apparatus

$$g(\hat{Y}_A) = \int_{y\in\Omega_Y} {\rm d}\nu(y)\, g(y)\frac{\delta\hat{\mathcal{Y}}_A(\Omega_Y)}{\delta\nu(y)} = \int_{y\in\Omega_Y} {\rm d}\nu(y)\, g(y)\hat{\mathcal{Y}}_A(y). \tag{18}$$

Here, it is important to note that there are quantum measurement processes in which we cannot describe the readout process by any PVM. Typical examples are the photon counting process which is a continuous quantum measurement of photon number [48–51] and the operational phase measurement [52–54]. Therefore, we do not assume the existence of the pointer observable of the measurement apparatus and we use the POVM to describe the readout of the measurement outcome.

We now perform the readout of the result exhibited by the measurement apparatus just after interaction with the physical system. Then the probability that the measurement outcome y belongs to the set $E_Y(\subseteq \Omega_Y)$ is calculated by the formula [42–46]

$$\mathcal{P}_{\rm out}^A(E_Y) = {\rm Tr}_{SA}[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(E_Y))\hat{\rho}_{\rm out}^{SA}] \tag{19}$$

where Tr_{SA} stands for the trace operation over the tensor product Hilbert space $\mathcal{H}_S \otimes \mathcal{H}_A$ and the statistical operator $\hat{\rho}^{SA}_{\mathrm{out}}$ is given by equation (14). Here we introduce the probability function $P^A_{\mathrm{out}}(y)$ by

$$P^A_{\mathrm{out}}(y) = \mathrm{Tr}_{SA}[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\rho}^{SA}_{\mathrm{out}}] = \frac{\delta \mathcal{P}^A_{\mathrm{out}}(\Omega_Y)}{\delta \nu(y)} \tag{20}$$

in terms of which we obtain the output probability $\mathcal{P}^A_{\mathrm{out}}(E_Y) = \int_{y\in E_Y} \mathrm{d}\nu(y)\, P^A_{\mathrm{out}}(y)$. When we have obtained the measurement outcome which belongs to the set E_Y, the post-measurement state of the physical system is calculated by means of the state-reduction formula [42–46]

$$\hat{\rho}^S_{\mathrm{out}}(E_Y) = \frac{1}{\mathcal{P}^A_{\mathrm{out}}(E_Y)} \int_{y\in E_Y} \mathrm{d}\nu(y) P^A_{\mathrm{out}}(y)\hat{\rho}^S_{\mathrm{out}}(y) \tag{21}$$

where the conditional statistical operator $\hat{\rho}^S_{\mathrm{out}}(y)$ of the physical system, called the posterior state, is given by

$$\hat{\rho}^S_{\mathrm{out}}(y) = \frac{\mathrm{Tr}_A[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\rho}^{SA}_{\mathrm{out}}]}{\mathrm{Tr}_{SA}[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\rho}^{SA}_{\mathrm{out}}]} \tag{22}$$

where Tr_A is the trace operation over the Hilbert space $\mathcal{H}_A$ of the measurement apparatus. In the post-measurement state $\hat{\rho}^S_{\mathrm{out}}(E_Y)$ of the physical system, the observable $\hat{X}_S$ takes the value x belonging to the set E_X with conditional probability

$$\mathcal{P}^S_{\mathrm{out}}(E_X|E_Y) = \mathrm{Tr}_S[\hat{\mathcal{X}}_S(E_X)\hat{\rho}^S_{\mathrm{out}}(E_Y)] = \frac{\mathcal{P}^{SA}_{\mathrm{out}}(E_X, E_Y)}{\mathcal{P}^A_{\mathrm{out}}(E_Y)} \tag{23}$$

with

$$\mathcal{P}^{SA}_{\mathrm{out}}(E_X, E_Y) = \mathrm{Tr}_{SA}[(\hat{\mathcal{X}}_S(E_X) \otimes \hat{\mathcal{Y}}_A(E_Y))\hat{\rho}^{SA}_{\mathrm{out}}] \tag{24}$$

which represents the joint probability that in the compound quantum state $\hat{\rho}^{SA}_{\mathrm{out}}$, the observable $\hat{X}_S$ takes the value x in the set E_X and the measurement outcome y belongs to the set E_Y. As we have done in equations (8) and (20), we introduce the probability functions $P^{SA}_{\mathrm{out}}(x, y)$ and $P^S_{\mathrm{out}}(x|y)$ by

$$\begin{aligned} P^{SA}_{\mathrm{out}}(x, y) &= \frac{\delta^2 \mathcal{P}^{SA}_{\mathrm{out}}(\Omega_X, \Omega_Y)}{\delta\mu(x)\delta\nu(y)} \\ &= \mathrm{Tr}_A\langle\psi_S(x)|[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\rho}^{SA}_{\mathrm{out}}]|\psi_S(x)\rangle \qquad (25) \\ P^S_{\mathrm{out}}(x|y) &= \langle\psi_S(x)|\hat{\rho}^S_{\mathrm{out}}(y)|\psi_S(x)\rangle = \frac{P^{SA}_{\mathrm{out}}(x, y)}{P^A_{\mathrm{out}}(y)}. \qquad (26) \end{aligned}$$

Furthermore, according to Bayes theorem [55], we obtain the posterior probability function of the measurement apparatus

$$P^A_{\mathrm{out}}(y|x) = \frac{P^{SA}_{\mathrm{out}}(x, y)}{P^S_{\mathrm{out}}(x)} = \frac{P^S_{\mathrm{out}}(x|y)P^A_{\mathrm{out}}(y)}{P^S_{\mathrm{out}}(x)} \tag{27}$$

where the probability function $P^S_{\mathrm{out}}(x)$ is given by $P^S_{\mathrm{out}}(x) = \mathrm{Tr}_{SA}[(\hat{\mathcal{X}}_S(x) \otimes \hat{I}_A)\hat{\rho}^{SA}_{\mathrm{out}}]$.

Using the probability functions, we introduce the Shannon entropies in the quantum measurement process

$$H(X^S_{\mathrm{out}}) = -\int_{x\in\Omega_X} \mathrm{d}\mu(x) P^S_{\mathrm{out}}(x) \log P^S_{\mathrm{out}}(x) \tag{28}$$

$$H(Y^A_{\mathrm{out}}) = -\int_{y\in\Omega_Y} \mathrm{d}\nu(y) P^A_{\mathrm{out}}(y) \log P^A_{\mathrm{out}}(y) \tag{29}$$

$$H(X_{\rm out}^S|Y_{\rm out}^A) = -\int_{x\in\Omega_X} {\rm d}\mu(x) \int_{y\in\Omega_Y} {\rm d}\nu(y) P_{\rm out}^{SA}(x,y) \log P_{\rm out}^S(x|y) \tag{30}$$

$$H(Y_{\rm out}^A|X_{\rm out}^S) = -\int_{x\in\Omega_X} {\rm d}\mu(x) \int_{y\in\Omega_Y} {\rm d}\nu(y) P_{\rm out}^{SA}(x,y) \log P_{\rm out}^A(y|x) \tag{31}$$

$$H(X_{\rm out}^S, Y_{\rm out}^A) = -\int_{x\in\Omega_X} {\rm d}\mu(x) \int_{y\in\Omega_Y} {\rm d}\nu(y) P_{\rm out}^{SA}(x,y) \log P_{\rm out}^{SA}(x,y) \tag{32}$$

which satisfy well known relations among the entropies

$$\begin{aligned} H(X_{\rm out}^S, Y_{\rm out}^A) &= H(X_{\rm out}^S|Y_{\rm out}^A) + H(Y_{\rm out}^A) \\ &= H(Y_{\rm out}^A|X_{\rm out}^S) + H(X_{\rm out}^S). \end{aligned} \tag{33}$$

The mutual entropy between the physical system and the measurement apparatus just after interaction is given by

$$H(Y_{\rm out}^A; X_{\rm out}^S) = H(X_{\rm out}^S) + H(Y_{\rm out}^A) - H(X_{\rm out}^S, Y_{\rm out}^A). \tag{34}$$

Furthermore, we introduce the Shannon entropy $H(Y_{\rm in}^A)$ of the measurement apparatus before interaction with the physical system

$$H(Y_{\rm in}^A) = -\int_{y\in\Omega_Y} {\rm d}\nu(y) P_{\rm in}^A(y) \log P_{\rm in}^A(y) \tag{35}$$

where $P_{\rm in}^A(y) = {\rm Tr}_A[\hat{\mathcal{Y}}_A(y)\hat{\rho}_{\rm in}^A]$. On the other hand, the von Neumann entropy of the post-measurement state of the physical system is given by

$$S(\hat{\rho}_{\rm out}^S|Y_{\rm out}^A) = -\int_{y\in\Omega_Y} {\rm d}\nu(y) P_{\rm out}^A(y) {\rm Tr}_S[\hat{\rho}_{\rm out}^S(y) \log \hat{\rho}_{\rm out}^S(y)] \tag{36}$$

where the output probability function $P_{\rm out}^A(y)$ and the conditional statistical operator $\hat{\rho}_{\rm out}^S(y)$ are given, respectively, by equations (20) and (22). Once we obtain the result of the quantum measurement process, the decreases of the Shannon entropy and the von Neumann entropy of the physical system are calculated by

$$\Delta H(X_{\rm out}^S, X_{\rm in}^S|Y_{\rm out}^A) = H(X_{\rm in}^S) - H(X_{\rm out}^S|Y_{\rm out}^A) \tag{37}$$

$$\Delta S(\hat{\rho}_{\rm out}^S, \hat{\rho}_{\rm in}^S|Y_{\rm out}^A) = S(\hat{\rho}_{\rm in}^S) - S(\hat{\rho}_{\rm out}^S|Y_{\rm out}^A). \tag{38}$$

These quantities will be compared with the amount of information about the observable of the physical system that can be obtained by the quantum measurement process.

3. Information gain in quantum measurement processes

In this section, we consider the information on the observable $\hat{X}_S$ of the physical system in the quantum state $\hat{\rho}_{\rm in}^S$, which can be obtained by the quantum measurement process. The output probability of the quantum measurement process is given by equation (19) or (20). Then, by substituting equation (14) into equations (19) and (20), we obtain

$$\mathcal{P}_{\rm out}^S(E_Y) = {\rm Tr}_S[\hat{\mathcal{Z}}_S(E_Y)\hat{\rho}_{\rm in}^S] \qquad P_{\rm out}^S(y) = {\rm Tr}_S[\hat{\mathcal{Z}}_S(y)\hat{\rho}_{\rm in}^S] \tag{39}$$

where the operators $\hat{\mathcal{Z}}_S(y)$ and $\hat{\mathcal{Z}}_S(E_Y)$ defined on the Hilbert space $\mathcal{H}_S$ of the physical system are given, respectively, by

$$\hat{\mathcal{Z}}_S(y) = {\rm Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)] \tag{40}$$

$$\hat{\mathcal{Z}}_S(E_Y) = {\rm Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(E_Y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)] \tag{41}$$

where the relations $\hat{\mathcal{Z}}_S(E_Y) = \int_{y\in E_Y} \mathrm{d}\nu(y)\hat{\mathcal{Z}}_S(y)$ and $\hat{\mathcal{Z}}_S(y) = \delta\hat{\mathcal{Z}}_S(\Omega_Y)/\delta\nu(y)$ hold. It is found from equation (17) and the unitarity of the operator $\hat{\mathcal{U}}_{SA}$ that the operator $\hat{\mathcal{Z}}_S(E_Y)$ becomes a POVM of the physical system, which satisfies the relations

$$\hat{\mathcal{Z}}_S(E_Y) \geqslant 0 \qquad \hat{\mathcal{Z}}_S(\Omega_Y) = \hat{I}_S. \tag{42}$$

Here, it should be noted that although the operator $\hat{\mathcal{X}}_S(E_X)$ of the physical system is a PVM, the operator $\hat{\mathcal{Z}}_S(E_Y)$ does not becomes a PVM in general. We introduce an operator of the physical system by the following relation

$$g(\hat{Z}_S) = \int_{y\in\Omega_Y} \mathrm{d}\nu(y) g(y)\hat{\mathcal{Z}}_S(y) \tag{43}$$

where $g(y)$ is an analytic function of y. Since $\hat{\mathcal{Z}}_S(E_Y)$ is not a PVM, $g(\hat{Z}_S)$ does not become Hermitian. In fact, it is easy to see that $g(\hat{Z}_S) \neq \sum_n g_n \hat{Z}_S^n$ even if $g(y) = \sum_n g_n y^n$. The operators $\hat{\mathcal{X}}_S(E_X)$ and $f(\hat{X}_S)$ of the physical system do not depend on the measurement apparatus and the system–apparatus interaction while $\hat{\mathcal{Z}}_S(E_Y)$ and $g(\hat{Z}_S)$ defined on the Hilbert space $\mathcal{H}_S$ are determined only by the measurement apparatus and the system–apparatus interaction and are independent of the intrinsic properties of the physical system. For this reason, $\hat{\mathcal{X}}_S(E_X)$ or $f(\hat{X}_S)$ are referred to as the intrinsic observable of the physical system and $\hat{\mathcal{Z}}_S(E_Y)$ or $g(\hat{Z}_S)$ are called the operational observable of the physical system [56–58]. The operational observable is also referred to as the unsharp [59–61] or fuzzy observable [62–66].

To proceed further, we assume that the operational observable $\hat{\mathcal{Z}}_S(y)$ determined by the quantum measurement process commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system, that is, $[\hat{\mathcal{X}}_S(x), \hat{\mathcal{Z}}_S(y)] = 0$ for all $x \in \Omega_X$ and $y \in \Omega_Y$. This commutability is a mathematical condition imposed on quantum measurement processes and at present, the physical implications are not clear. Under this assumption, we can express the operational observable $\hat{\mathcal{Z}}_S(y)$ in the following form

$$\begin{aligned}\hat{\mathcal{Z}}_S(y) &= \int_{x\in\Omega_X} \mathrm{d}\mu(x)|\psi_S(x)\rangle P_{SA}(y|x)\langle\psi_S(x)| \\ &= \int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{SA}(y|x)\hat{\mathcal{X}}_S(x)\end{aligned} \tag{44}$$

which is formally written as $\hat{\mathcal{Z}}_S(y) = P_{SA}(y|\hat{X}_S)$ (see equation (5)). Since we have $\hat{\mathcal{X}}_S(x) \geqslant 0$, $\hat{\mathcal{Z}}_S(y) \geqslant 0$ and $\hat{\mathcal{X}}_S(\Omega_X) = \hat{\mathcal{Z}}_S(\Omega_Y) = \hat{I}_S$, the kernel function $P_{SA}(y|x)$ satisfies the relations

$$P_{SA}(y|x) \geqslant 0 \qquad \int_{y\in\Omega_Y} \mathrm{d}\nu(y) P_{SA}(y|x) = 1. \tag{45}$$

Substituting equation (44) into (39), we find that the kernel function $P_{SA}(y|x)$ gives the relation between the initial probability $P_{\mathrm{in}}^S(x)$ of the physical system and the output probability $P_{\mathrm{out}}^A(y)$ of the quantum measurement process

$$P_{\mathrm{out}}^A(y) = \int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{SA}(y|x) P_{\mathrm{in}}^S(x). \tag{46}$$

Thus, it is seen from equations (45) and (46) that the kernel function $P_{SA}(y|x)$ is the conditional probability (density) that the measurement outcome lies in the infinitesimal range around value y when the intrinsic observable of the physical system in the quantum state $\hat{\rho}_{\mathrm{in}}^S$ takes a value in the infinitesimal range around x. More precisely, the conditional probability that the measurement outcome belongs to the set $E_Y(\subseteq \Omega_Y)$ when the intrinsic observable of the physical system takes a value belonging to the set $E_X(\subseteq \Omega_X)$ is given by

$$P_{SA}(E_Y|E_X) = \int_{x\in E_X} \mathrm{d}\mu(x) \int_{y\in E_Y} \mathrm{d}\nu(y) P_{SA}(y|x). \tag{47}$$

For the quantum measurement processes of discrete observables, the conditional probability $P_{SA}(y_i|x_j)$ and the output probability $P^A_{\text{out}}(y_i)$ become

$$P_{SA}(y_i|x_j) = \langle\psi_S(x_j)|\hat{\mathcal{Z}}_S(y_i)|\psi_S(x_j)\rangle \tag{48}$$

$$P^A_{\text{out}}(y_i) = \sum_{x_j\in\Omega_X} P_{SA}(y_i|x_j)P^S_{\text{in}}(x_j). \tag{49}$$

On the other hand, if the output probability $P^A_{\text{out}}(y)$ is written in the form of equation (46), the operational observable $\hat{\mathcal{Z}}_S(y)$ must satisfy the following relation for any initial statistical operator $\hat{\rho}^S_{\text{in}}$

$$\text{Tr}_S\left\{\left[\hat{\mathcal{Z}}_S(y) - \int_{x\in\Omega_X} \mathrm{d}\mu(x)P_{SA}(y|x)\hat{\mathcal{X}}_S(x)\right]\hat{\rho}^S_{\text{in}}\right\} = 0 \tag{50}$$

which yields equation (44). Therefore, the commutability of the intrinsic and operational observables is necessary and sufficient for the existence of the conditional probability $P_{SA}(y|x)$ in the quantum measurement process.

According to Bayes theorem [55], we obtain the posterior probability $\tilde{P}_{SA}(x|y)$ that the intrinsic observable of the physical system in the initial quantum state $\hat{\rho}^S_{\text{in}}$ takes a value x when the measurement outcome y was given

$$\tilde{P}_{SA}(x|y) = \frac{P_{SA}(y,x)}{P^A_{\text{out}}(y)} = \frac{P_{SA}(y|x)P^S_{\text{in}}(x)}{P^A_{\text{out}}(y)} \tag{51}$$

where $P_{SA}(y,x)$ is the joint probability of the intrinsic observable of the physical system and the outcome of the quantum measurement process. Using these probabilities, the joint and conditional entropies in the quantum measurement process are given, respectively, by

$$H(Y^A_{\text{out}}, X^S_{\text{in}}) = -\int_{x\in\Omega_X} \mathrm{d}\mu(x)\int_{y\in\Omega_Y} \mathrm{d}\nu(y)P_{SA}(y,x)\log P_{SA}(y,x) \tag{52}$$

$$H(Y^A_{\text{out}}|X^S_{\text{in}}) = -\int_{x\in\Omega_X} \mathrm{d}\mu(x)\int_{y\in\Omega_Y} \mathrm{d}\nu(y)P_{SA}(y,x)\log P_{SA}(y|x) \tag{53}$$

$$H(X^S_{\text{in}}|Y^A_{\text{out}}) = -\int_{x\in\Omega_X} \mathrm{d}\mu(x)\int_{y\in\Omega_Y} \mathrm{d}\nu(y)P_{SA}(y,x)\log \tilde{P}_{SA}(x|y) \tag{54}$$

which satisfies the relations

$$\begin{aligned} H(Y^A_{\text{out}}, X^S_{\text{in}}) &= H(Y^A_{\text{out}}|X^S_{\text{in}}) + H(X^S_{\text{in}}) \\ &= H(X^S_{\text{in}}|Y^A_{\text{out}}) + H(Y^A_{\text{out}}) \end{aligned} \tag{55}$$

where $H(X^S_{\text{in}})$ and $H(Y^A_{\text{out}})$ are given, respectively, by equations (9) and (29). The information about the intrinsic observable of the physical system in the initial quantum state $\hat{\rho}^S_{\text{in}}$, which can be obtained by the quantum measurement process, is equivalent to the information transmitted from the physical system in the initial quantum state $\hat{\rho}^S_{\text{in}}$ to the measurement apparatus in the output state $\hat{\rho}^{SA}_{\text{out}}$ by unitary transformation $\hat{\mathcal{U}}_{SA}$. This means that the quantum measurement process can be considered a quantum communication channel between the physical system to be measured and the measurement apparatus. Thus, when the relations given by equations (45) and (46) hold, we can represent the amount of information about the intrinsic observable of the physical system by the Shannon mutual entropy

$$I(Y^A_{\text{out}}; X^S_{\text{in}}) = H(Y^A_{\text{out}}) + H(X^S_{\text{in}}) - H(Y^A_{\text{out}}, X^S_{\text{in}}). \tag{56}$$

Therefore, we can summarize the result in the following theorem.

Theorem 1. *If the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system to be measured commutes with the operational observable $\hat{\mathcal{Z}}_S(y)$ determined by the quantum measurement process, that is, $[\hat{\mathcal{X}}_S(x), \hat{\mathcal{Z}}_S(y)] = 0$ for all $x \in \Omega_X$ and $y \in \Omega_Y$, the amount of information about the intrinsic observable of the physical system in the quantum state $\hat{\rho}^S_{\rm in}$, which can be obtained by the quantum measurement process, is given by the Shannon mutual entropy*

$$\begin{aligned} I(Y^A_{\rm out}; X^S_{\rm in}) &= \int_{x\in\Omega_X} {\rm d}\mu(x) \int_{y\in\Omega_Y} {\rm d}\nu(y) P_{SA}(y|x) P^S_{\rm in}(x) \log\left[\frac{P_{SA}(y|x)}{P^A_{\rm out}(y)}\right] \\ &= H(Y^A_{\rm out}) + H(X^S_{\rm in}) - H(Y^A_{\rm out}, X^S_{\rm in}) \\ &= H(Y^A_{\rm out}) - H(Y^A_{\rm out}|X^S_{\rm in}) \\ &= H(X^S_{\rm in}) - H(X^S_{\rm in}|Y^A_{\rm out}) \end{aligned} \tag{57}$$

where the initial probability $P^S_{\rm in}(x)$ of the physical system and the output probability $P^A_{\rm out}(y)$ of the measurement apparatus are given, respectively, by equations (8) and (20) and the conditional probability $P_{SA}(y|x)$ is determined by equation (44).

Thus far we have not imposed any restriction on the quantum measurement process, except for the commutativity of the intrinsic and operational observables of the physical system, $[\hat{\mathcal{X}}_S(x), \hat{\mathcal{Z}}_S(y)] = 0$. In some cases, *the probability reproducibility condition* [61, 67, 68] is introduced for investigating the properties of quantum measurement processes and their interpretation. In our notations, this condition is represented by the following relation

$${\rm Tr}_S[\hat{\mathcal{X}}_S(x)\hat{\rho}^S_{\rm in}] = {\rm Tr}_S[\hat{\mathcal{Z}}_S(g(x))\hat{\rho}^S_{\rm in}] \tag{58}$$

where $g(y)$ is some analytic function which connects the measurement outcome y with the value x of the intrinsic observable of the physical system. Of course, this relation is equivalent to the equality $P^S_{\rm in}(x) = P^A_{\rm out}(g(x))$. Furthermore, the equalities ${\rm d}\mu(g^{-1}(y)) = {\rm d}\nu(y)$ and $g(\Omega_X) = \Omega_Y$ are required, from which we have the consistency for the normalization conditions

$$\begin{aligned} 1 &= \int_{x\in\Omega_X} {\rm d}\mu(x) P^S_{\rm in}(x) = \int_{x\in\Omega_X} {\rm d}\mu(x) P^A_{\rm out}(g(x)) = \int_{y\in g(\Omega_X)} {\rm d}\mu(g^{-1}(y)) P^A_{\rm out}(y) \\ &= \int_{y\in\Omega_Y} {\rm d}\nu(y) P^A_{\rm out}(y). \end{aligned} \tag{59}$$

Since the relation given by equation (58) holds for any statistical operator $\hat{\rho}^S_{\rm in}$ of the physical system, we obtain the equality $\hat{\mathcal{X}}_S(x) = \hat{\mathcal{Z}}_S(g(x))$ of the intrinsic and operational observables, which indicates that the conditional probability in equation (44) is given by $P_{SA}(y|x) = \delta(y - g(x))$ or $P_{SA}(y|x) = \delta_{y,g(x)}$. Thus, it is found that the probability reproducibility condition is stronger than the commutability of the intrinsic and operational observables.

Since the differential entropy does not take a finite value for δ-function probability densities, we consider the quantum measurement process of a discrete observable that satisfies the probability reproducibility condition. In this case, the information gain $I(Y^A_{\rm out}; X^S_{\rm in})$ is calculated to be

$$\begin{aligned} I(Y^A_{\rm out}; X^S_{\rm in}) &= \sum_{x_j\in\Omega_X}\sum_{y_k\in\Omega_Y} P_{SA}(y_k|x_j) P^S_{\rm in}(x_j) \log\left[\frac{P_{SA}(y_k|x_j)}{P^A_{\rm out}(y_k)}\right] \\ &= -\sum_{x_j\in\Omega_X} P^A_{\rm out}(g(x_j)) \log P^A_{\rm out}(g(x_j)) \\ &= -\sum_{y_k\in\Omega_Y} P^A_{\rm out}(y_k) \log P^A_{\rm out}(y_k) = H(Y^A_{\rm out}) \\ &= -\sum_{x_j\in\Omega_X} P^S_{\rm in}(x_j) \log P^S_{\rm in}(x_j) = H(X^S_{\rm in}) \end{aligned} \tag{60}$$

which is equivalent to $H(X_{\rm in}^S|Y_{\rm out}^A) = H(Y_{\rm out}^A|X_{\rm in}^S) = 0$. This result means that complete information can be obtained from the measurement outcomes and the information gain is given by the Shannon entropy of the physical system in the quantum state $\hat{\rho}_{\rm in}^S$. Therefore, it is found that under the probability reproducibility condition, the Shannon entropy of the physical system becomes equal to the amount of information which we can obtain by the quantum measurement process. Furthermore, since the von Neumann entropy is no greater than the Shannon entropy, the probability reproducibility condition yields a relation among the information gain, the Shannon and von Neumann entropies

$$I(Y_{\rm out}^A; X_{\rm in}^S) = H(X_{\rm in}^S) \geqslant S(\hat{\rho}_{\rm in}^S) \tag{61}$$

$$I(Y_{\rm out}^A; X_{\rm in}^S) = H(Y_{\rm out}^A) \geqslant S(\hat{\rho}_{\rm out}^A) \tag{62}$$

where $S(\hat{\rho}_{\rm out}^A)$ is the von Neumann entropy of the measurement apparatus in the quantum state $\hat{\rho}_{\rm out}^A = {\rm Tr}_S \hat{\rho}_{\rm out}^{SA}$.

4. Entropy change in quantum measurement processes

When we perform a quantum measurement on a physical system to obtain information about an intrinsic observable, the quantum state of the measured physical system changes due to the effect of the quantum measurement process. Such a state change induces a decrease of the entropy of the measured physical system since the information about the physical system is obtained and the uncertainty of the physical system is reduced. When we obtain the result of the quantum measurement process, the decreases of the Shannon entropy and the von Neumann entropy of the physical system are given, respectively, by equations (37) and (38). In this section, we compare the decrease of the Shannon entropy with that of the von Neumann entropy. Here, we consider quantum measurement processes of only discrete observables.

We first consider the case where the initial quantum state of the physical system is a statistical mixture of the eigenstates of the intrinsic observable, where the statistical operator $\hat{\rho}_{\rm in}^S$ commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x_j)$ of the physical system, that is $[\hat{\rho}_{\rm in}^S, \hat{\mathcal{X}}_S(x_j)] = 0$ for all $x_j \in \Omega_X$. In this case, we can represent the statistical operator $\hat{\rho}_{\rm in}^S$ in the following form

$$\hat{\rho}_{\rm in}^S = \sum_{x_j \in \Omega_X} P_{\rm in}^S(x_j) |\psi_S(x_j)\rangle\langle\psi_S(x_j)| = P_{\rm in}^S(\hat{X}_S) \tag{63}$$

where $P_{\rm in}^S(x_j) \geqslant 0$ and $\sum_{x_j \in \Omega_X} P_{\rm in}^S(x_j) = 1$. Since $|\psi_S(x_j)\rangle$ is the orthonormal eigenstate state, the Shannon entropy $H(X_{\rm in}^S)$ and the von Neumann entropy $S(\hat{\rho}_{\rm in}^S)$ of the physical system in the initial quantum state $\hat{\rho}_{\rm in}^S$ are equal

$$S(\hat{\rho}_{\rm in}^S) = H(X_{\rm in}^S) = -\sum_{x_j \in \Omega_X} P_{\rm in}^S(x_j) \log P_{\rm in}^S(x_j). \tag{64}$$

On the other hand, when we obtain the measurement outcome, the von Neumann entropy $S(\hat{\rho}_{\rm out}^S|Y_{\rm out}^A)$ of the physical system in the post-measurement state can be evaluated as follows

$$\begin{aligned}
S(\hat{\rho}_{\rm out}^S|Y_{\rm out}^A) &= -\sum_{y_k \in \Omega_Y} P_{\rm out}^A(y_k) {\rm Tr}_S[\hat{\rho}_{\rm out}^S(y_k) \log \hat{\rho}_{\rm out}^S(y_k)] \\
&\leqslant -\sum_{y_k \in \Omega_Y} P_{\rm out}^A(y_k) \sum_{x_j \in \Omega_X} P_{\rm out}^S(x_j|y_k) \log P_{\rm out}^S(x_j|y_k) \\
&= -\sum_{y_k \in \Omega_Y} \sum_{x_j \in \Omega_X} P_{\rm out}^{SA}(x_j, y_k) \log P_{\rm out}^S(x_j|y_k) \\
&= H(X_{\rm out}^S|Y_{\rm out}^A)
\end{aligned} \tag{65}$$

where $P^S_{\rm out}(x_j|y_k) = \langle\psi_S(x_j)|\hat{\rho}^S_{\rm out}(y_k)|\psi_S(x_j)\rangle$ and we have used equations (25) and (26). The inequality on the right-hand side of this equation is ensured by the Jensen inequality [4] or equivalently by the concavity of the entropy function $[f(x) = -x\log x]$ [20]. Using equations (64) and (65), we obtain the inequality

$$\begin{aligned}\Delta S(\hat{\rho}^S_{\rm out}, \hat{\rho}^S_{\rm in}|Y^A_{\rm out}) &= S(\hat{\rho}^S_{\rm in}) - S(\hat{\rho}^S_{\rm out}|Y^A_{\rm out}) \\ &\geqslant H(X^S_{\rm in}) - H(X^S_{\rm out}|Y^A_{\rm out}) \\ &= \Delta H(X^S_{\rm out}, X^S_{\rm in}|Y^A_{\rm out}).\end{aligned} \tag{66}$$

Thus if the commutativity $[\hat{\rho}^S_{\rm in}, \hat{\mathcal{X}}_S(x_j)] = 0$ holds for all $x_j \in \Omega_X$, the decrease of the von Neumann entropy of the physical system is no less than that of the Shannon entropy in the quantum measurement process of the discrete observable.

We next consider the case where the physical system after the measurement is a statistical mixture of the eigenstates of the intrinsic observable, where the conditional statistical operator $\hat{\rho}^S_{\rm out}(y_k)$ of the post-measurement state of the physical system commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x_j)$ of the physical system, that is $[\hat{\rho}^S_{\rm out}(y_k), \hat{\mathcal{X}}_S(x_j)] = 0$ for all $x_j \in \Omega_X$ and $y_k \in \Omega_Y$. Then, the conditional statistical operator $\hat{\rho}^S_{\rm out}(y_k)$ can be expressed as

$$\hat{\rho}^S_{\rm out}(y_k) = \sum_{y_k\in\Omega Y} P^S_{\rm out}(x_j|y_k)|\psi_S(x_j)\rangle\langle\psi_S(x_j)| = P^S_{\rm out}(\hat{X}_S|y_k) \tag{67}$$

where $P^S_{\rm out}(x_j|y_k) \geqslant 0$ and $\sum_{x_j\in\Omega_X} P^S_{\rm out}(x_j|y_k) = 1$. In this case, it is easy to see from equations (30) and (36) that the following equality is established after the quantum measurement process

$$S(\hat{\rho}^S_{\rm out}|Y^A_{\rm out}) = H(X^S_{\rm out}|Y^A_{\rm out}) = -\sum_{y_k\in\Omega_Y}\sum_{x_j\in\Omega_X} P^{SA}_{\rm out}(x_j, y_k)\log P^S_{\rm out}(x_j|y_k). \tag{68}$$

Using the fact that the inequality $S(\hat{\rho}^S_{\rm in}) \leqslant H(X^S_{\rm in})$ holds in general, we find that decreases of the von Neumann entropy of the physical system is no greater than that of the Shannon entropy in the quantum measurement process of the discrete observable

$$\Delta S(\hat{\rho}^S_{\rm out}, \hat{\rho}^S_{\rm in}|Y^A_{\rm out}) \leqslant \Delta H(X^S_{\rm out}, X^S_{\rm in}|Y^A_{\rm out}). \tag{69}$$

Therefore, we can summarize the results in the following theorem.

Theorem 2. *In quantum measurement processes of discrete observables, when we obtain the measurement outcome, decreases of the Shannon entropy and the von Neumann entropy of the measured physical system satisfy the inequality*

$$\Delta S(\hat{\rho}^S_{\rm out}, \hat{\rho}^S_{\rm in}|Y^A_{\rm out}) \leqslant \Delta H(X^S_{\rm out}, X^S_{\rm in}|Y^A_{\rm out}) \tag{70}$$

if the conditional statistical operator $\hat{\rho}^S_{\rm out}(y_k)$ of the post-measurement state of the physical system commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x_j)$ and

$$\Delta S(\hat{\rho}^S_{\rm out}, \hat{\rho}^S_{\rm in}|Y^A_{\rm out}) \geqslant \Delta H(X^S_{\rm out}, X^S_{\rm in}|Y^A_{\rm out}) \tag{71}$$

if the initial statistical operator $\hat{\rho}^S_{\rm in}$ of the physical system to be measured commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x_j)$. In equations (70) and (71), the equality holds if the commutativity $[\hat{\rho}^S_{\rm out}(y_k), \hat{\mathcal{X}}_S(x_j)] = [\hat{\rho}^S_{\rm in}, \hat{\mathcal{X}}_S(x_j)] = 0$ is established for all $x_j \in \Omega_X$ and $y_k \in \Omega_Y$.

Finally we remark that the change of von Neumann entropy of the physical system in quantum measurement processes has been investigated by Groenewold, Lindblad and Ozawa [69–72].

5. Information gain and entropy change

In section 3, we investigated the information gain $I(Y_{\rm out}^A; X_{\rm in}^S)$ in the quantum measurement process and in section 4, we considered the entropy decreases $\Delta H(X_{\rm out}^S, X_{\rm in}^S|Y_{\rm out}^A)$ and $\Delta S(\hat{\rho}_{\rm out}^S, \hat{\rho}_{\rm in}^S|Y_{\rm out}^A)$ of the physical system that are caused by the quantum measurement process. In this section, we therefore investigate the relation between the information gain $I(Y_{\rm out}^A; X_{\rm in}^S)$ and the entropy decrease $\Delta H(X_{\rm out}^S, X_{\rm in}^S|Y_{\rm out}^A)$. For this purpose, we first recall that the condition under which the amount of information obtained by the quantum measurement process can be represented by the Shannon mutual entropy is the commutativity of the intrinsic and operational observables of the physical system, that is $[\hat{\mathcal{X}}_S(x), \hat{\mathcal{Z}}_S(y)] = 0$ for all $x \in \Omega_X$ and $y \in \Omega_Y$. Then from equations (40) and (44) we can write this condition as

$$\mathrm{Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)] = \int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{SA}(y|x)\hat{\mathcal{X}}_S(x). \qquad (72)$$

By using the completeness relation $\int_{x\in\Omega_X} \mathrm{d}\mu(x)\hat{\mathcal{X}}_S(x) = \hat{I}_S$, this relation becomes

$$\int_{x\in\Omega_X} \mathrm{d}\mu(x)\mathrm{Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{\mathcal{X}}_S(x) \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)] = \int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{SA}(y|x)\hat{\mathcal{X}}_S(x). \qquad (73)$$

To obtain the relation between information gain and entropy decrease, we further impose a condition on the quantum measurement process. The condition is that the integrand on the left-hand side of equation (73) is equivalent to that on the right-hand side, which can be given in the following form

$$\mathrm{Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{\mathcal{X}}_S(x) \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)] = P_{SA}(y|f(x;y))\hat{\mathcal{X}}_S(f(x;y)) \qquad (74)$$

where $f(x;y)$ is a function of x that in general depends on y and the spectral set Ω_X of the intrinsic observable, the measure $\mathrm{d}\mu(x)$ and the conditional probability $P_{SA}(y|x)$ satisfy the relation

$$\int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{SA}(y|x) F(x) = \int_{x\in\Omega_X} \mathrm{d}\mu(x) P_{SA}(y|f(x;y)) F(f(x;y)). \qquad (75)$$

Here $F(x)$ is an arbitrary non-singular function of x. Of course, the condition given by equations (74) and (75) is stronger than that given by equation (73). In fact, it is easy to see that equation (73) holds if equations (74) and (75) are satisfied. Furthermore, equation (74) means that except for the conditional probability, the intrinsic observable is transformed as $\hat{\mathcal{X}}_S(x) \to \hat{\mathcal{X}}_S(x')$ with $x' = f(x;y)$ by the dual map of the completely positive map $\hat{\rho}_{\rm in}^S \to \hat{\rho}_{\rm out}^S(y)$.

As an example that satisfies the relations given by equations (74) and (75), let us consider a quantum non-demolition measurement of the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system [73–75]. Here we assume that the system–apparatus interaction is sufficiently strong so that $\|\hat{H}_{\rm int}^{SA}(t)\| \gg \|\hat{H}_0^S(t)\|$ and $\|\hat{H}_{\rm int}^{SA}(t)\| \gg \|\hat{H}_0^A(t)\|$, where $\|\hat{X}\|$ is a norm of operator $\hat{X}$. In this case, the PVM $\hat{\mathcal{X}}_S(x)$, the POVM $\hat{\mathcal{Y}}_A(y)$ and the unitary operator $\hat{\mathcal{U}}_{SA}$ satisfy

$$[\hat{\mathcal{X}}_S(x), \hat{\mathcal{U}}_{SA}] = 0 \qquad [\hat{\mathcal{Y}}_A(y), \hat{\mathcal{U}}_{SA}] \neq 0. \qquad (76)$$

It is easy to see from equations (40) and (76) that the commutativity of the intrinsic and operational observables, $[\hat{\mathcal{X}}_S(x), \hat{\mathcal{Z}}_S(y)] = 0$, holds for the quantum non-demolition measurement. Then, we can calculate the left-hand side of equation (74) as follows

$$\begin{aligned}\mathrm{Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{\mathcal{X}}_S(x) \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)] &= \mathrm{Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S \otimes \hat{\rho}_{\rm in}^A)]\hat{\mathcal{X}}_S(x) \\ &= \int_{x'\in\Omega_X} \mathrm{d}\mu(x') P_{SA}(y|x')\hat{\mathcal{X}}_S(x')\hat{\mathcal{X}}_S(x) \\ &= P_{SA}(y|x)\hat{\mathcal{X}}_S(x) \qquad (77)\end{aligned}$$

where we have used equation (72) and the fact that the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system is an orthogonal projector. This result shows that the function $f(x;\, y)$ in equation (74) is given by $f(x;\, y) = x$ for the quantum non-demolition measurement. Therefore, any quantum non-demolition measurement of the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system satisfies the relations given by equations (74) and (75).

To investigate the entropy decrease of the physical system that is caused by the quantum measurement process, we calculate the joint probability function $P^{SA}_{\text{out}}(x, y)$ given by equation (25) under condition (74)

$$\begin{aligned}
P^{SA}_{\text{out}}(x, y) &= \langle\psi_S(x)|\text{Tr}_A[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\rho}^{SA}_{\text{out}}]|\psi_S(x)\rangle \\
&= \text{Tr}_{SA}[(\hat{\mathcal{X}}_S(x) \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{\rho}^S_{\text{in}} \otimes \hat{\rho}^A_{\text{in}})\hat{\mathcal{U}}^\dagger_{SA}] \\
&= \text{Tr}_{SA}[\hat{\mathcal{U}}^\dagger_{SA}(\hat{\mathcal{X}}_S(x) \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{\rho}^S_{\text{in}} \otimes \hat{\rho}^A_{\text{in}})] \\
&= P_{SA}(y|f(x;\, y))\text{Tr}_S[\hat{\mathcal{X}}_S(f(x;\, y))\hat{\rho}^S_{\text{in}}] \\
&= P_{SA}(y|f(x;\, y))\langle\psi_S(x)|\hat{\rho}^S_{\text{in}}|\psi_S(x)\rangle|_{x\to f(x;y)} \\
&= P_{SA}(y|f(x;\, y))P^S_{\text{in}}(f(x;\, y)).
\end{aligned} \tag{78}$$

If the function $f(x;\, y)$ is independent of y, this result yields the equality $P^S_{\text{out}}(x) = P^S_{\text{in}}(f(x))$. This means that when we do not obtain the measurement outcome, the probability of the observable $\hat{X}_S$ in the post-measurement state of the physical system is equal to that of the observable $f(\hat{X}_S)$ in the initial quantum state of the physical system. Using equations (75) and (78), we can calculate the joint entropy $H(X^S_{\text{out}}, Y^A_{\text{out}})$ as follows

$$\begin{aligned}
H(X^S_{\text{out}}, Y^A_{\text{out}}) &= -\int_{x\in\Omega_X} \mathrm{d}\mu(x) \int_{y\in\Omega_Y} \mathrm{d}\nu(y) P^{SA}_{\text{out}}(x, y) \log P^{SA}_{\text{out}}(x, y) \\
&= -\int_{x\in\Omega_X} \mathrm{d}\mu(x) \int_{y\in\Omega_Y} \mathrm{d}\nu(y) P_{SA}(y|f(x;\, y))P^S_{\text{in}}(f(x;\, y)) \\
&\quad \times \log[P_{SA}(y|f(x;\, y))P^S_{\text{in}}(f(x;\, y))] \\
&= -\int_{x\in\Omega_X} \mathrm{d}\mu(x) \int_{y\in\Omega_Y} \mathrm{d}\nu(y) P_{SA}(y|x)P^S_{\text{in}}(x) \log[P_{SA}(y|x)P^S_{\text{in}}(x)] \\
&= H(Y^A_{\text{out}}|X^S_{\text{in}}) + H(X^S_{\text{in}}) \\
&= H(X^S_{\text{in}}) + H(Y^A_{\text{out}}) - I(Y^A_{\text{out}};\, X^S_{\text{in}})
\end{aligned} \tag{79}$$

where we have used equation (57). It should be noted that this relation is different from the well known relation

$$H(X^S_{\text{out}}, Y^A_{\text{out}}) = H(X^S_{\text{out}}) + H(Y^A_{\text{out}}) - I(X^S_{\text{out}};\, Y^A_{\text{out}}). \tag{80}$$

From equations (37) and (79), the entropy decrease of the physical system in the quantum measurement process becomes

$$\begin{aligned}
\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) &= H(X^S_{\text{in}}) - H(X^S_{\text{out}}|Y^A_{\text{out}}) \\
&= H(X^S_{\text{in}}) + H(Y^A_{\text{out}}) - H(X^S_{\text{out}}, Y^A_{\text{out}}) \\
&= I(Y^A_{\text{out}};\, X^S_{\text{in}}).
\end{aligned} \tag{81}$$

Therefore we obtain the following theorem.

Theorem 3. *If the quantum measurement process satisfies the condition given by equations (74) and (75), the amount of information about the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system in the quantum state $\hat{\rho}^S_{\text{in}}$, which can be obtained by the quantum measurement process, is equal to the decrease of the Shannon entropy of the measured physical system*

$$\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) = I(Y^A_{\text{out}};\, X^S_{\text{in}}). \tag{82}$$

In particular, this equality always holds for any quantum non-demolition measurement of the intrinsic observable of the physical system.

It is important to note that the condition of the theorem is sufficient, but not necessary, to hold equality (82). To see this, we consider the case that the intrinsic observable $\hat{\mathcal{X}}_S(x)$ of the physical system has a discrete and non-degenerate spectrum and the quantum measurement process satisfies the probability reproducibility condition (58). It should be noted that the probability reproducibility condition (58) does not guarantee the condition of the theorem. In this case, we obtain the relation from equations (26) and (78)

$$\begin{aligned} P^S_{\text{out}}(x|y) &= \langle\psi_S(x)|\hat{\rho}^S_{\text{out}}(y)|\psi_S(x)\rangle \\ &= \delta_{y,g(f(x;y))}\frac{P^S_{\text{in}}(f(x;y))}{P^A_{\text{out}}(y)} \\ &= \delta_{y,g(f(x;y))} \end{aligned} \tag{83}$$

where we have used $P_{SA}(y|x) = \delta_{y,g(x)}$ and $P^S_{\text{in}}(x) = P^A_{\text{out}}(g(x))$. Since $|\psi_S(x)\rangle$ is the eigenstate of the discrete and non-degenerate observable, equation $y = g(f(x;y))$ has a unique solution $\tilde{x}(y)$ for given y. Thus, we obtain $\hat{\rho}^S_{\text{out}}(y) = |\psi_S(\tilde{x}(y))\rangle\langle\psi_S(\tilde{x}(y))|$ from equation (83). This result indicates that after the measurement outcome was obtained, both the Shannon entropy and von Neumann entropy of the physical system vanish, that is $H(X^S_{\text{out}}|Y^A_{\text{out}}) = S(\hat{\rho}^S_{\text{out}}|Y^A_{\text{out}}) = 0$. This is consistent with the result obtained in section 3 that when the quantum measurement process satisfies the probability reproducibility condition, complete information can be obtained from the measurement outcomes. In this case, since we have $\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) = H(X^S_{\text{in}})$, equality (82) holds.

Combining theorem 3 with theorem 2, we find that if the quantum measurement process of a discrete observable satisfies the relations given by equations (74) and (75), the amount of information $I(Y^A_{\text{out}}; X^S_{\text{in}})$ which can be obtained by the quantum measurement process, decreases of the Shannon entropy and the von Neumann entropy, $\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}})$ and $\Delta S(\hat{\rho}^S_{\text{out}}, \hat{\rho}^S_{\text{in}}|Y^A_{\text{out}})$, of the measured physical system satisfy

$$\Delta S(\hat{\rho}^S_{\text{out}}, \hat{\rho}^S_{\text{in}}|Y^A_{\text{out}}) \leqslant \Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) = I(Y^A_{\text{out}}; X^S_{\text{in}}) \tag{84}$$

if the conditional statistical operator $\hat{\rho}^S_{\text{out}}(y_k)$ of the post-measurement state of the physical system commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x_j)$ and

$$\Delta S(\hat{\rho}^S_{\text{out}}, \hat{\rho}^S_{\text{in}}|Y^A_{\text{out}}) \geqslant \Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) = I(Y^A_{\text{out}}; X^S_{\text{in}}) \tag{85}$$

if the initial statistical operator $\hat{\rho}^S_{\text{in}}$ of the physical system commutes with the intrinsic observable $\hat{\mathcal{X}}_S(x_j)$. Furthermore, if the quantum measurement process satisfies the probability reproducibility condition, the relations given by equations (84) and (85), respectively, become $S(\hat{\rho}^S_{\text{in}}) \leqslant H(X^S_{\text{in}}) = I(Y^A_{\text{out}}; X^S_{\text{in}})$ and $S(\hat{\rho}^S_{\text{in}}) = H(X^S_{\text{in}}) = I(Y^A_{\text{out}}; X^S_{\text{in}})$.

Before closing this section, we consider the case where the physical system before the interaction with the measurement apparatus is in an eigenstate of the intrinsic observable with eigenvalue $\tilde{x}$, that is $\hat{\rho}^S_{\text{in}} = |\psi_S(\tilde{x})\rangle\langle\psi_S(\tilde{x})|$, where we assume a discrete and non-degenerate observable. In this case, since we obtain $H(X^S_{\text{in}}) = 0$ and $P^A_{\text{out}}(y_k) = P_{SA}(y_k|\tilde{x})$ from equations (9) and (46), it is seen from equation (57) that the information gain becomes zero, that is $I(Y^A_{\text{out}}; X^S_{\text{in}}) = 0$. Therefore, when the initial quantum state of the physical system is the eigenstate of the intrinsic observable with discrete and non-degenerate spectrum, the information gain becomes zero. This result is consistent with our intuition that if we have complete knowledge of the intrinsic observable of the physical system, we cannot obtain any further information about it, even though we perform any quantum measurement on the physical system. Furthermore, if the quantum measurement process satisfies the condition given by equations (74) and (75), we obtain the relation $H(X^S_{\rm out}, Y^A_{\rm out}) = H(Y^A_{\rm out})$ and $H(X^S_{\rm out}|Y^A_{\rm out}) = 0$ since $H(X^S_{\rm in}) = I(Y^A_{\rm out}; X^S_{\rm in}) = 0$. Thus, we find that the entropy decrease of the measured physical system becomes zero, that is $\Delta H(X^S_{\rm out}, X^S_{\rm in}|Y^A_{\rm out}) = 0$.

6. Examples of quantum measurement processes

In this section, we consider two simple examples of quantum measurement processes to examine the general results obtained in sections 3–5. One is the standard position measurement [18] and the other is the photon number measurement by means of a lossless beam splitter [76, 77]. We obtain the information gain and the entropy change in these quantum measurement processes.

6.1. Standard position measurement

In position measurement of the physical system, the intrinsic observable of the physical system and the PVM of the measurement apparatus are given, respectively, by $\hat{\mathcal{X}}_S(x) = |x_S\rangle\langle x_S|$ and $\hat{\mathcal{Y}}_A(y) = |y_A\rangle\langle y_A|$, where $|x_S\rangle$ and $|y_A\rangle$ are the eigenstates of the position operators $\hat{x}_S$ and $\hat{x}_A$ of the physical system and the measurement apparatus, satisfying the eigenvalue equations $\hat{x}_S|x_S\rangle = x|x_S\rangle$ and $\hat{x}_A|y_A\rangle = y|y_A\rangle$. The position operator $\hat{x}_A$ corresponds to the pointer observable of the measurement apparatus. In this case, the sets Ω_X and Ω_Y are the set of all real numbers. The unitary operator that describes the state change due to interaction between the physical system and the measurement apparatus in the measurement process is assumed to be

$$\hat{\mathcal{U}}_{SA} = \exp(-{\rm i}\hat{x}_S \otimes \hat{p}_A) \tag{86}$$

where $\hat{p}_A$ is the momentum operator of the measurement apparatus, canonically conjugate to the position operator $\hat{x}_A$ and we set $\lambda\tau_{\rm int}/\hbar = 1$ with the coupling constant λ, for the sake of simplicity. Then the compound quantum state $\hat{\rho}^{SA}_{\rm out}$ of the physical system and the measurement apparatus just after the interaction becomes

$$\begin{aligned}\hat{\rho}^{SA}_{\rm out} = \int_{-\infty}^{\infty} {\rm d}x \int_{-\infty}^{\infty} {\rm d}x' \int_{-\infty}^{\infty} {\rm d}y \int_{-\infty}^{\infty} {\rm d}y' \langle x_S|\hat{\rho}^S_{\rm in}|x'_S\rangle\langle y_A|\hat{\rho}^A_{\rm in}|y'_A\rangle \\ \times |x_S\rangle\langle x'_S| \otimes |x_A + y_A\rangle\langle x'_A + y'_A|.\end{aligned} \tag{87}$$

Hence, we can obtain the conditional statistical operator $\hat{\rho}^S_{\rm out}(r)$ of the physical system and the output probability density $P^A_{\rm out}(r)$ of the measurement apparatus

$$\hat{\rho}^S_{\rm out}(r) = \int_{-\infty}^{\infty} {\rm d}x \int_{-\infty}^{\infty} {\rm d}y |x_S\rangle \left[\frac{\langle x_S|\hat{\rho}^S_{\rm in}|y_S\rangle\langle r_A - x_A|\hat{\rho}^A_{\rm in}|r_A - y_A\rangle}{P^A_{\rm out}(r)}\right]\langle y_S| \tag{88}$$

$$P^A_{\rm out}(r) = \int_{-\infty}^{\infty} {\rm d}x \langle r_A - x_A|\hat{\rho}^A_{\rm in}|r_A - x_A\rangle P^S_{\rm in}(x) \tag{89}$$

where $P^S_{\rm in}(x) = \langle x_S|\hat{\rho}^S_{\rm in}|x_S\rangle$ is the position probability density of the physical system in the initial quantum state $\hat{\rho}^S_{\rm in}$. The operational observable $\hat{\mathcal{Z}}_S(r)$ of the physical system that is determined by the position measurement is obtained from equation (40)

$$\begin{aligned}\hat{\mathcal{Z}}_S(r) &= \int_{-\infty}^{\infty} {\rm d}x |x_S\rangle\langle r_A - x_A|\hat{\rho}^A_{\rm in}|r_A - x_A\rangle\langle x_S| \\ &= \int_{-\infty}^{\infty} {\rm d}x \langle r_A - x_A|\hat{\rho}^A_{\rm in}|r_A - x_A\rangle\hat{\mathcal{X}}_S(x)\end{aligned} \tag{90}$$

which indicates that the conditional probability density $P_{SA}(y|x)$ in the position measurement process is given by

$$P_{SA}(y|x) = \langle y_A - x_A|\hat{\rho}^A_{\rm in}|y_A - x_A\rangle = P^A_{\rm in}(y - x) \tag{91}$$

where $P^A_{\rm in}(y) = \langle y_A|\hat{\rho}^A_{\rm in}|y_A\rangle$ is the initial probability density of the measurement apparatus. It is easy to see from equation (90) that the intrinsic position observable commutes with the operational position observable, namely, $(\hat{\mathcal{X}}_S(x), \hat{\mathcal{Z}}_S(r)) = 0$ for all x and r. Therefore, theorem 1 holds for the standard position measurement of the physical system.

The amount of information $I(Y^A_{\rm out}; X^S_{\rm in})$ about the intrinsic position observable of the physical system is calculated from equations (57) and (91)

$$\begin{aligned} I(Y^A_{\rm out}; X^S_{\rm in}) &= \int_{-\infty}^{\infty} {\rm d}x \int_{-\infty}^{\infty} {\rm d}y\, P_{SA}(x|y) P^S_{\rm in}(y) \log\left[\frac{P_{SA}(x|y)}{P^A_{\rm out}(x)}\right] \\ &= H(Y^A_{\rm out}) + \int_{-\infty}^{\infty} {\rm d}x \int_{-\infty}^{\infty} {\rm d}y\, P^A_{\rm in}(x-y) P^S_{\rm in}(y) \log P^A_{\rm in}(x-y) \\ &= H(Y^A_{\rm out}) + \int_{-\infty}^{\infty} {\rm d}x \int_{-\infty}^{\infty} {\rm d}y\, P^A_{\rm in}(x) P^S_{\rm in}(y) \log P^A_{\rm in}(x) \\ &= H(Y^A_{\rm out}) - H(Y^A_{\rm in}). \end{aligned} \tag{92}$$

This result indicates that the amount of information on the position observable of the physical system is equal to the entropy increase of the measurement apparatus. We also obtain the equality $H(Y^A_{\rm in}) = H(Y^A_{\rm out}|X^S_{\rm in})$ which means that when we have complete knowledge of the position observable of the physical system, the uncertainty of the result of the position measurement is equal to the uncertainty of the initial position of the measurement apparatus.

We next examine the relations given by equations (74) and (75). For the position measurement, it is easy to calculate the left-hand side of equation (74)

$$\begin{aligned} {\rm Tr}_A[\hat{\mathcal{U}}^\dagger_{SA}(\hat{\mathcal{X}}_S(x)\otimes\hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{I}_S\otimes\hat{\rho}^A_{\rm in})] &= \langle y_A|\exp(-{\rm i}x\hat{p}_A)\hat{\rho}^A_{\rm in}\exp({\rm i}x\hat{p})|y_A\rangle|x_S\rangle\langle x_S| \\ &= \langle y_A - x_A|\hat{\rho}^A_{\rm in}|y_A - x_A\rangle|x_S\rangle\langle x_S| \\ &= P_{SA}(y|x)\hat{\mathcal{X}}_S(x) \end{aligned} \tag{93}$$

where we have used equation (91). This result indicates that the position measurement of the physical system satisfies the relations given by equations (74) and (75) with $f(x; y) = x$. Thus, it is found from theorem 3 and equation (92) that the amount of information $I(Y^A_{\rm out}; X^S_{\rm in})$ about the position observable of the physical system is equal to the decrease $\Delta H(X^S_{\rm out}, X^S_{\rm in}|Y^A_{\rm out})$ of the Shannon entropy of the physical system and to the increase $H(Y^A_{\rm out}) - H(Y^A_{\rm in})$ of the Shannon entropy of the measurement apparatus in the position measurement.

6.2. Photon number measurement

We next consider the photon number measurement of the physical system by means of a lossless beam splitter [76, 77]. In this case, the intrinsic observable of the physical system and the PVM of the measurement apparatus becomes $\hat{\mathcal{X}}_S(n) = |n_S\rangle\langle n_S|$ and $\hat{\mathcal{Y}}_A(n) = |n_A\rangle\langle n_A|$, where $|n_S\rangle$ and $|n_A\rangle$ are the Fock states of the physical system and the measurement apparatus. The pointer observable is the photon number operator of the measurement apparatus. The unitary operator that describes the state change due to the system–apparatus interaction (beam splitting) is given by

$$\hat{\mathcal{U}}_{SA} = \exp[-\theta(\hat{a}^\dagger_S\otimes\hat{a}_A - \hat{a}_S\otimes\hat{a}^\dagger_A)] \tag{94}$$

where $\hat{a}_S$ and $\hat{a}_S^\dagger$ ($\hat{a}_A$ and $\hat{a}_A^\dagger$) are the annihilation and creation operators of the physical system (the measurement apparatus). Furthermore, the initial state of the measurement apparatus is assumed to be the vacuum state $\hat{\rho}_{\rm in}^A = |0_A\rangle\langle 0_A|$. Then the compound quantum state $\hat{\rho}_{\rm out}^{SA}$ of the physical system and the measurement apparatus just after interaction becomes [76, 77]

$$\hat{\rho}_{\rm out}^{SA} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\left[\frac{1}{m!n!}\left(\frac{\mathcal{R}}{\mathcal{T}}\right)^{m+n}\right]^{1/2}\hat{a}_S^m\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{\rho}_{\rm in}^S\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{a}_S^{\dagger n}\otimes|m_A\rangle\langle n_A| \tag{95}$$

where $\mathcal{T} = \cos^2\theta$ and $\mathcal{R} = \sin^2\theta$ are the transmittance and reflectance of the beam splitter. The conditional statistical operator $\hat{\rho}_{\rm out}^S(m)$ of the physical system and the output probability $P_{\rm out}^A(m)$ of the photon number measurement are given, respectively, by

$$\hat{\rho}_{\rm out}^S(m) = \frac{\hat{a}_S^m\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{\rho}_{\rm in}^S\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{a}_S^{\dagger m}}{{\rm Tr}_S[\hat{a}_S^m\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{\rho}_{\rm in}^S\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{a}_S^{\dagger m}]} \tag{96}$$

$$P_{\rm out}^A(m) = \sum_{n=m}^{\infty}\frac{n!}{m!(n-m)!}\mathcal{R}^m\mathcal{T}^{n-m}P_{\rm in}^S(n) \tag{97}$$

where $P_{\rm in}^S(n) = \langle n_S|\hat{\rho}_{\rm in}^S|n_S\rangle$ is the photon number probability of the physical system in the initial quantum state $\hat{\rho}_{\rm in}^S$. The operational observable of the physical system defined by this photon number measurement is obtained from equation (40)

$$\begin{aligned}\hat{\mathcal{Z}}_S(m) &= \sum_{n=m}^{\infty}|n_S\rangle\frac{n!}{m!(n-m)!}\mathcal{R}^m\mathcal{T}^{n-m}\langle n_S| \\ &= \sum_{n=0}^{\infty}\frac{n!}{m!(n-m)!}\mathcal{R}^m\mathcal{T}^{n-m}\hat{\mathcal{X}}_S(n)\end{aligned} \tag{98}$$

which means that the conditional probability $P_{SA}(m|n)$ is given by

$$P_{SA}(m|n) = \frac{n!}{m!(n-m)!}\mathcal{R}^m\mathcal{T}^{n-m}. \tag{99}$$

In the second equality of equation (98), we have used the fact that $n! \to \infty$ ($1/n! \to 0$) if n is a negative integer. It is easy to see from equation (98) that the operational photon number observable $\hat{\mathcal{Z}}_S(m)$ commutes with the intrinsic photon number observable $\hat{\mathcal{X}}_S(n)$. Therefore, theorem 1 holds for the photon number measurement by means of the lossless beam splitter. Here we remark that the operational observable in the homodyne detection was obtained by Banaszek and Wódkiewicz [57] and their result shows the commutability of the intrinsic and operational observables. Thus theorem 1 is established in homodyne detection.

To examine the sufficient condition for theorem 3, we first calculate the left-hand side of equation (74)

$$\begin{aligned}{\rm Tr}_A[\hat{\mathcal{U}}_{SA}^\dagger(\hat{\mathcal{X}}_S(n)\otimes\hat{\mathcal{Y}}_A(m))\hat{\mathcal{U}}_{SA}(\hat{I}_S\otimes|0_A\rangle\langle 0_A|)] &= \frac{1}{m!}\left(\frac{\mathcal{R}}{\mathcal{T}}\right)^m\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S}\hat{a}_S^{\dagger m}|n_S\rangle\langle n_S|\hat{a}_S^m\mathcal{T}^{\frac{1}{2}\hat{a}_S^\dagger\hat{a}_S} \\ &= \frac{(m+n)!}{m!n!}\mathcal{R}^m\mathcal{T}^n|m_S+n_S\rangle\langle m_S+n_S| \\ &= P_{SA}(m|m+n)\hat{\mathcal{X}}_S(m+n)\end{aligned} \tag{100}$$

where we have used equation (99). Thus, we have found that the relation given by equation (74) is satisfied and the function $f(x;y)$ from equation (74) is given by $f(x;y) = x+y$. Furthermore, we can easily verify the relation given by equation (75)

$$\begin{aligned}\sum_{n=0}^{\infty} P_{SA}(m|n+m)F(m+n) &= \sum_{n=0}^{\infty} \frac{(m+n)!}{m!n!}\mathcal{R}^m\mathcal{T}^n F(m+n)\\ &= \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!}\mathcal{R}^m\mathcal{T}^{n-m}F(n)\\ &= \sum_{n=0}^{\infty} \frac{n!}{m!(n-m)!}\mathcal{R}^m\mathcal{T}^{n-m}F(n)\\ &= \sum_{n=0}^{\infty} P_{SA}(m|n)F(n).\end{aligned} \tag{101}$$

This result means that the relation given by equation (75) is satisfied in the photon number measurement. Therefore, since theorem 3 holds, we see that the amount of information $I(Y^A_{\text{out}}; X^S_{\text{in}})$ on the photon number of the physical system is equal to the entropy decrease $\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}})$ of the physical system in the photon number measurement. Although we have to consider the photon number measurement by means of the beam splitter, the same results can also be obtained in the photon number measurement with the non-degenerate parametric amplifier, where the unitary operator $\hat{\mathcal{U}}_{SA}$ is given by $\hat{\mathcal{U}}_{SA} = \exp(-\theta(\hat{a}^\dagger_S \otimes \hat{a}^\dagger_A - \hat{a}_S \otimes \hat{a}_A))$. In this case we have $f(x;y) = x-y$ in equation (74).

We next consider changes of the Shannon entropy and the von Neumann entropy in the photon number measurement by means of the lossless beam splitter. Since the conditional statistical operator $\hat{\rho}^S_{\text{out}}(m)$ of the post-measurement state of the physical system is given by equation (96), we obtain the equality $\Delta S(\hat{\rho}^S_{\text{out}}, \hat{\rho}^S_{\text{in}}|Y^A_{\text{out}}) = \Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}})$ if the initial quantum state $\hat{\rho}^S_{\text{in}}$ of the physical system is diagonal with respect to the photon number eigenstate $|n_S\rangle$, that is $\hat{\rho}^S_{\text{in}} = \sum_{n=0}^{\infty} P^S_{\text{in}}(n)|n_S\rangle\langle n_S|$, since the conditional statistical operator $\hat{\rho}^S_{\text{out}}(m)$ becomes diagonal if $\hat{\rho}^S_{\text{in}}$ is diagonal with respect to the photon number eigenstate $|n_S\rangle$. On the other hand, when the initial quantum state of the physical system is a superposition of vacuum and one-photon states [78]

$$\hat{\rho}^S_{\text{in}} = |\psi^S_{\text{in}}\rangle\langle\psi^S_{\text{in}}| \qquad |\psi^S_{\text{in}}\rangle = a|0_S\rangle + b|1_S\rangle \tag{102}$$

with $|a|^2 + |b|^2 = 1$, the conditional statistical operator of the physical system is given by

$$\hat{\rho}^S_{\text{out}}(0) = \frac{1}{|a|^2 + \mathcal{T}|b|^2}|\psi^S_{\text{out}}\rangle\langle\psi^S_{\text{out}}| \qquad \hat{\rho}^S_{\text{out}}(1) = |0_S\rangle\langle 0_S| \tag{103}$$

where $|\psi^S_{\text{out}}\rangle = a|0_S\rangle + \mathcal{T}^{1/2}b|1_S\rangle$. Furthermore, the output probability of the measurement apparatus becomes

$$P^A_{\text{out}}(0) = |a|^2 + \mathcal{T}|b|^2 \qquad P^A_{\text{out}}(1) = \mathcal{R}|b|^2. \tag{104}$$

Since both the input and conditional output states, $\hat{\rho}^S_{\text{in}}$ and $\hat{\rho}^S_{\text{out}}(m)$, of the physical system are pure, we obtain

$$S(\hat{\rho}^S_{\text{in}}) = S(\hat{\rho}^S_{\text{out}}|Y^A_{\text{out}}) = \Delta S(\hat{\rho}^S_{\text{out}}, \hat{\rho}^S_{\text{in}}|Y^A_{\text{out}}) = 0. \tag{105}$$

The decrease of Shannon entropy of the measured physical system is calculated to be

$$\begin{aligned}\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) &= -(1-\mathcal{R}|a|^2)\log(1-\mathcal{R}|a|^2)\\ &\quad -\mathcal{R}|b|^2\log|b|^2 + (1-\mathcal{R})|b|^2\log(1-\mathcal{R})\end{aligned} \tag{106}$$

which yields the inequality $\Delta H(X^S_{\text{out}}, X^S_{\text{in}}|Y^A_{\text{out}}) \geqslant \Delta S(\hat{\rho}^S_{\text{out}}, \hat{\rho}^S_{\text{in}}|Y^A_{\text{out}}) = 0$, where the equality holds for $\mathcal{T} = 0$ or $\mathcal{T} = 1$.

Finally we remark that the two quantum measurement processes considered here do not satisfy the probability reproducibility condition. In fact, from equations (90) and (98), the probability reproducibility condition (58) is expressed as

$$\langle f_A(x) - y_A|\hat{\rho}^A_{\text{in}}|f_A(x) - y_A\rangle = \delta(x - y) \tag{107}$$

for the position measurement and

$$\frac{m!}{g(n)![m - g(n)]!}\mathcal{R}^{g(n)}\mathcal{T}^{m-g(n)} = \delta_{m,n} \tag{108}$$

for the photon-number measurement, where $f_A(x)$ is some real-valued function and $g(n)$ is a non-negative integer. For any physical state $\hat{\rho}^A_{\text{in}}$ of the measurement apparatus, (107) does not hold, whatever function $f_A(x)$ is. Furthermore, it is easy to see that (108) is not satisfied for any function $g_A(n)$.

7. Summary

In this paper we have considered the amount of information which we can obtain by means of the quantum measurement process of the intrinsic observable of the physical system and we have also investigated the entropy change of the measured physical system that is caused by the quantum measurement process. We first obtained the condition under which the information gain can be represented by Shannon mutual entropy. The condition is that the intrinsic observable of the measured physical system commutes with the operational observable defined by the quantum measurement process. As the example that satisfies this condition, the standard position measurement and the photon number measurement by means of a lossless beam splitter have been considered. Of course, there are many other quantum measurement processes in which the condition holds [79]. We next investigated the entropy decreases of the physical system that are caused by quantum measurement processes of discrete observables. We have found that the decrease of the von Neumann entropy is no greater than that of the Shannon entropy when the conditional statistical operator of the post-measurement state of the physical system commutes with the intrinsic observable. On the other hand, a decrease of the von Neumann entropy is no less than that of the Shannon entropy when the statistical operator of the initial quantum state of the physical system commutes with the intrinsic observable. Then, we have compared the amount of information which we can obtain by the quantum measurement process with the entropy decrease of the measured physical system. We have found that the conditions for information gain and entropy decrease are equal. In particular, any quantum non-demolition measurement satisfies this condition. The general results obtained in this paper are summarized in theorems 1, 2 and 3. In this paper, we have confined ourselves to considering the cases where the intrinsic observable of the physical system is represented by the PVM. When the intrinsic observable is given by the POVM, the results obtained in this paper are slightly modified (see the appendix).

Appendix. Information and entropy for generalized observables

In this appendix, we consider the relation between the information gain and the entropy decrease in a quantum measurement process for a generalized intrinsic observable of the physical system, where we assume that the generalized observable can be represented by

$$\hat{\mathcal{X}}_S(E_X) = \int_{x\in E_X} \mathrm{d}\mu(x)\hat{\mathcal{X}}_S(x) \tag{A.1}$$

where $\hat{\mathcal{X}}_S(x)$ is not an orthogonal projector. Note that the generalized observable satisfies equation (2) but not equation (3). The Susskind–Glogower phase observable is the typical example of the generalized observable [80, 81]. To describe the quantum measurement process, we introduce a superoperator $\hat{\mathcal{L}}_S(y)$

$$\hat{\mathcal{L}}_S(y)\hat{O}_S = \mathrm{Tr}_A[(\hat{I}_S \otimes \hat{\mathcal{Y}}_A(y))\hat{\mathcal{U}}_{SA}(\hat{O}_S \otimes \hat{\rho}^A_{\mathrm{in}})\hat{\mathcal{U}}^\dagger_{SA}] \tag{A.2}$$

for any operator $\hat{O}_S$ of the physical system. This superoperator satisfies the relations

$$\hat{\mathcal{L}}_S(y) \geqslant 0 \qquad \int_{y\in\Omega_Y} \mathrm{d}\nu(y)\hat{\mathcal{L}}^\dagger_S(y) = \hat{I}_S \tag{A.3}$$

where the superoperator $\hat{\mathcal{L}}^\dagger_S(y)$ is defined by the relation

$$\mathrm{Tr}_S\{\hat{V}_S[\hat{\mathcal{L}}_S(y)\hat{W}_S]\} = \mathrm{Tr}_S\{[\hat{\mathcal{L}}^\dagger_S(y)\hat{V}_S]\hat{W}_S\} \tag{A.4}$$

for any operators $\hat{V}_S$ and $\hat{W}_S$ of the physical system. Using the superoperator $\hat{\mathcal{L}}_S(y)$, we can express the probability function $P^A_{\mathrm{out}}(y)$ of the measurement outcomes, the conditional statistical operator $\hat{\rho}^S_{\mathrm{out}}(y)$ of the post-measurement state of the physical system and the operational observable $\hat{\mathcal{Z}}_S(y)$ as follows

$$P^A_{\mathrm{out}}(y) = \mathrm{Tr}_S[\hat{\mathcal{L}}_S(y)\hat{\rho}^S_{\mathrm{in}}] \tag{A.5}$$

$$\hat{\rho}^S_{\mathrm{out}} = \frac{\hat{\mathcal{L}}_S(y)\hat{\rho}^S_{\mathrm{in}}}{\mathrm{Tr}_S[\hat{\mathcal{L}}_S(y)\hat{\rho}^S_{\mathrm{in}}]} \tag{A.6}$$

$$\hat{\mathcal{Z}}_S(y) = \hat{\mathcal{L}}^\dagger_S(y)\hat{I}_S. \tag{A.7}$$

To obtain the relation between the information gain and the entropy decrease, we impose the condition on the quantum measurement process that the superoperator $\hat{\mathcal{L}}^\dagger_S(y)$ maps the POVM $\hat{\mathcal{X}}_S(x)$ of the intrinsic observable as follows

$$\hat{\mathcal{L}}^\dagger_S(y)\hat{\mathcal{X}}_S(x) = \mathcal{K}(y|h(x;y))\hat{\mathcal{X}}_S(h(x;y)) \tag{A.8}$$

where the functions $\mathcal{K}(y|x)$ and $h(x;y)$ satisfy the relation

$$\int_{x\in\Omega_X} \mathrm{d}\mu(x)\mathcal{K}(y|h(x;y))F(h(x;y)) = \int_{x\in\Omega_X} \mathrm{d}\mu(x)\mathcal{K}(y|x)F(x) \tag{A.9}$$

for any non-singular function $F(x)$. In this case, because of the linearity of the superoperator $\hat{\mathcal{L}}_S(y)$, the operational observable $\hat{\mathcal{Z}}_S(y)$ is calculated to be

$$\begin{aligned}\hat{\mathcal{Z}}_S(y) &= \hat{\mathcal{L}}^\dagger_S\hat{I}_S = \hat{\mathcal{L}}^\dagger_S\int_{x\in\Omega_X} \mathrm{d}\mu(x)\hat{\mathcal{X}}_S(x)\\ &= \int_{x\in\Omega_X} \mathrm{d}\mu(x)\hat{\mathcal{L}}^\dagger_S\hat{\mathcal{X}}_S(x)\\ &= \int_{x\in\Omega_X} \mathrm{d}\mu(x)\mathcal{K}(y|h(x;y))\hat{\mathcal{X}}_S(h(x;y))\\ &= \int_{x\in\Omega_X} \mathrm{d}\mu(x)\mathcal{K}(y|x)\hat{\mathcal{X}}_S(x).\end{aligned} \tag{A.10}$$

This result indicates that the function $\mathcal{K}(y|x)$ represents the conditional probability $P_{SA}(y|x)$ of the quantum measurement process, that is

$$P^A_{\mathrm{out}}(y) = \int_{x\in\Omega_X} \mathrm{d}\mu(x)\mathcal{K}(y|x)P^S_{\mathrm{in}}(x). \tag{A.11}$$

Therefore, if the relations given by equations (A.8) and (A.9) are satisfied, theorem 1 holds for the quantum measurement process for the generalized observable. It should be noted here that the operational observable $\hat{\mathcal{Z}}_S(y)$ does not commute with the intrinsic observable $\hat{\mathcal{X}}_S(x)$ unless $\hat{\mathcal{X}}_S(x)$ is an orthogonal projector. We next calculate the joint probability $P^{SA}_{\text{out}}(x, y)$ under the conditions given by equations (A.8) and (A.9)

$$\begin{aligned} P^{SA}_{\text{out}}(x, y) &= P^S_{\text{out}}(x|y) P^A_{\text{out}}(y) = \text{Tr}_S[\hat{\mathcal{X}}_S(x)\hat{\rho}^S_{\text{out}}(y)] P^A_{\text{out}}(y) \\ &= \text{Tr}_S[\hat{\mathcal{X}}_S(x)\hat{\mathcal{L}}_S(y)\hat{\rho}^S_{\text{in}}] \\ &= \text{Tr}_S[\hat{\rho}^S_{\text{in}}\hat{\mathcal{L}}^\dagger_S(y)\hat{\mathcal{X}}_S(x)] \\ &= \mathcal{K}(y|h(x; y))\text{Tr}_S[\hat{\mathcal{X}}_S(h(x; y))\hat{\rho}^S_{\text{in}}] \\ &= \mathcal{K}(y|h(x; y)) P^S_{\text{in}}(h(x; y)) \end{aligned} \tag{A.12}$$

which is equivalent to the relation given by equation (78). Therefore, it is easy to see that theorem 3 holds for the quantum measurement process for the generalized observable that satisfies the relations given by equations (A.8) and (A.9). We finally note that although theorem 1 and theorem 3 hold for any quantum non-demolition measurement in the case that the intrinsic observable is represented by PVM, they are, in general, no longer valid for quantum non-demolition measurements for generalized observables.

References

[1] Mayer J E and Mayer M G 1977 *Statistical Mechanics* (New York: Wiley)
[2] Kubo R, Toda M and Hashitsume N 1985 *Statistical Physics* (Berlin: Springer)
[3] Gray R M 1990 *Entropy and Information Theory* (Berlin: Springer)
[4] Cover T M and Thomas J A 1991 *Elements of Information Theory* (New York: Wiley)
[5] Prigogine I 1980 *From Being to Becoming* (San Francisco: Freeman)
[6] Jaynes E T 1957 *Phys. Rev.* **106** 620
[7] Jaynes E T 1957 *Phys. Rev.* **108** 171
[8] Szilard L 1929 *Z. Phys.* **53** 840
[9] Brillouin L 1956 *Science and Information* (New York: Academic)
[10] Shannon C E 1948 *Bell Sys. Tech. J.* **27** 379
[11] Shannon C E 1948 *Bell Sys. Tech. J.* **27** 623
[12] Shannon C E and Weaver W W 1949 *The Mathematical Theory of Communication* (Champaign, IL: University of Illinois Press)
[13] Lindley D V 1956 *Ann. Math. Stat.* **27** 986
[14] Zurek W H 1989 *Nature* **341** 119
[15] Zurek W H 1989 *Phys. Rev.* A **40** 4731
[16] Caves C M 1993 *Phys. Rev.* E **47** 4010
[17] Lloyd S 1996 *Phys. Rev.* A **56** 3374
[18] von Neumann J 1955 *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press)
[19] Ballan R, Vénéroni M and Balazs N 1986 *Europhys. Lett.* **1** 1
[20] Wehrl A 1987 *Rev. Mod. Phys.* **50** 221
[21] Belavkin V P, Hirota O and Hudson R L (eds) 1995 *Quantum Communication and Measurement* (New York: Plenum)
[22] Hirota O, Holevo A S and Caves C M (eds) 1997 *Quantum Communication, Computation and Measurement* (New York: Plenum)
[23] Schumacher B 1995 *Phys. Rev.* A **51** 2738
[24] Jozsa R and Schumacher B 1994 *J. Mod. Opt.* **41** 2343
[25] Hausladen P, Jozsa R, Schumacher B, Westmoreland M and Wootters W 1996 *Phys. Rev.* A **54** 1869
[26] Schumacher B and Westmoreland M 1997 *Phys. Rev.* A **56** 131
[27] Holevo A S 1998 *IEEE Trans. Inf. Theor.* **44** 269
[28] Holevo A S 1973 *Multivar. Anal.* **3** 337
[29] Davies E B 1978 *IEEE Trans. Inf. Theor.* **24** 596
[30] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[31] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[32] Yuen H P, Kennedy R S and Lax M 1975 *IEEE Trans. Inf. Theor.* **21** 125

[33] Belavkin V P and Stratonovich R L 1973 *Radiotekh. Elecktron.* **18** 1839
[34] Ekert A K, Huttner B, Palma C M and Peres A 1997 *Phys. Rev.* A **50** 1047
[35] Fuchs C A and Peres A 1996 *Phys. Rev.* A **53** 2038
[36] Fuchs C A 1996 *LANL e-print* quant-ph/9611010
[37] Ekert A and Jozsa R 1996 *Rev. Mod. Phys.* **68** 733
[38] Steane A 1998 *Rep. Prog. Phys.* **61** 117
[39] DiVincenzo D P and Shor P 1996 *Phys. Rev. Lett.* **77** 3260
[40] DiVincenzo D P 1998 *Proc. R. Soc.* A **454** 261
[41] Preskill J 1998 *Proc. R. Soc.* A **454** 385
[42] Davies E B 1976 *Quantum Theory of Open Systems* (New York: Academic)
[43] Kraus K 1983 *States, Effects, and Operations* (Berlin: Springer)
[44] Ozawa M 1984 *J. Math. Phys.* **25** 79
[45] Ozawa M 1993 *J. Math. Phys.* **34** 5596
[46] Ozawa M 1997 *Ann. Phys., NY* **259** 121
[47] Prugovečki E 1971 *Quantum Mechanics in Hilbert Space* (New York: Academic)
[48] Srinivas M D and Davies E B 1981 *Opt. Acta* **28** 981
[49] Srinivas M D and Davies E B 1982 *Opt. Acta* **29** 235
[50] Srinivas M D 1996 *Pramana J. Phys.* **47** 1
[51] Imoto N, Ueda M and Ogawa T 1990 *Phys. Rev.* A **41** 4127
[52] Noh J W, Fougères A and Mandel L 1991 *Phys. Rev. Lett.* **67** 1426
[53] Noh J W, Fougères A and Mandel L 1992 *Phys. Rev.* A **45** 424
[54] Noh J W, Fougères A and Mandel L 1992 *Phys. Rev.* A **46** 2840
[55] Caves C M and Drummond P 1994 *Rev. Mod. Phys.* **66** 481
[56] Englert B and Wódkiewicz K 1995 *Phys. Rev.* A **51** R2661
[57] Banaszek K and Wódkiewicz K 1997 *Phys. Rev.* A **55** 3117
[58] Ban M 1997 *Int. J. Theor. Phys.* **36** 2583
[59] Busch P 1985 *Int. J. Theor. Phys.* **24** 63
[60] Busch P and Lahti P J 1989 *Found. Phys.* **19** 633
[61] Busch P, Grabowski M and Lahti P J 1995 *Operational Quantum Physics* (Berlin: Springer)
[62] Prugovečki E 1973 *Found. Phys.* **3** 3
[63] Prugovečki E 1976 *J. Math. Phys.* **17** 517
[64] Prugovečki E 1976 *J. Math. Phys.* **17** 1673
[65] Twareque Ali S and Prugovečki E 1997 *J. Math. Phys.* **18** 219
[66] Prugovečki E 1978 *Ann. Phys., NY* **110** 102
[67] Busch P, Lahti P J and Mittelstaedt P 1991 *The Quantum Theory of Measurement* (Berlin: Springer)
[68] Mittelstaedt P 1998 *The Interpretation of Quantum Mechanics and the Measurement Process* (Cambridge: Cambridge University Press)
[69] Groenewold H J 1971 *Int. J. Theor. Phys.* **9** 327
[70] Lindblad G 1972 *Commun. Math. Phys.* **28** 245
[71] Lindblad G 1973 *Commun. Math. Phys.* **33** 305
[72] Ozawa M 1986 *J. Math. Phys.* **27** 759
[73] Caves C M, Thorne K S, Drever R W P, Sandberg V D and Zimmermann M 1980 *Rev. Mod. Phys.* **52** 341
[74] Braginski V B and Khalili F Ya 1992 *Quantum Measurement* (Cambridge: Cambridge University Press)
[75] Braginski V B and Khalili F Ya 1996 *Rev. Mod. Phys.* **68** 1
[76] Ban M 1996 *J. Mod. Opt.* **43** 1281
[77] Ban M 1994 *Phys. Rev.* A **49** 5078
[78] Wódkiewicz K, Knight P L, Buckle S J and Barnett S M 1987 *Phys. Rev.* A **35** 2567
[79] Ban M 1998 *Int. J. Theor. Phys.* **37** 2555
[80] Susskind L and Glogower J 1964 *Physics, NY* **1** 49
[81] Carruthers P and Nieto M M 1968 *Rev. Mod. Phys.* **40** 411