Operations, disturbance, and simultaneous measurability

Masanao Ozawa
CREST, Japan Science and Technology, School of Informatics and Sciences, Nagoya University, Chikusa-ku, Nagoya 464-8601, Japan

(Received 12 May 2000; published 15 February 2001)

Quantum mechanics predicts the joint probability distributions of the outcomes of simultaneous measurements of commuting observables, but the current formulation lacks the operational definition of simultaneous measurements. In order to provide foundations of joint statistics of local general measurements on entangled systems in a general theoretical framework, the question is answered as to under what condition the outputs of two measuring apparatuses satisfy the joint probability formula for simultaneous measurements of their observables. For this purpose, all the possible state changes caused by measurements of an observable are characterized, and the notion of disturbance in measurement is formalized in terms of operations derived by the measuring interaction.

DOI: 10.1103/PhysRevA.63.032109 PACS number(s): 03.65.Ta, 03.67.−a

I. INTRODUCTION

The probability distribution of the outcome of a measurement is determined by the observable to be measured and the state at the time of the measurement, but the joint probability distribution of the outcomes of two successive measurements on the same object depends on how the first measurement disturbs the object. The disturbance depends not only on the observable and the state, but also on the apparatus to be used. Thus the joint probability distribution of successive measurements will be closely related to how the apparatus disturbs the object. It would be an interesting and significant problem to investigate the relation between the disturbance and the joint probability distribution, although to our knowledge there has been no systematic approach to the problem. This paper investigates, in particular, the relation between the disturbance and the joint probability formula for simultaneous measurements.

In quantum mechanics observables are represented by linear operators, for which the product operation is not necessarily commutative. Two observables are represented by commuting operators if and only if they are simultaneously measurable, and then quantum mechanics predicts the joint probability distribution of the outcomes of their simultaneous measurement (see Ref. [1] p. 228). But, it has not been answered fully what measurement can be considered as a simultaneous measurement of those observables.

In the current formulation we have two arguments to show how commuting observables can be measured simultaneously. The first argument is based on the fact that any commuting observables A and B have a third observable C for which A and B are functions of C (see Ref. [1], p. 173). In this case, the measurement of C gives also the outcomes of the A and the B measurements simultaneously (Ref. [1], p. 228). This argument gives one special instance of simultaneous measurement, but it is quite open from this argument how a pair of measuring apparatuses for A and B makes a simultaneous measurement of A and B.

The second argument assumes the projection postulate proposed by Lüders [2]. The projection postulate uniquely determines the state after the measurement conditional upon the outcome of the measurement, so that for successive measurements of any pair of discrete observables A and B, the joint probability distribution of their outcomes is determined. According to this probability distribution, if A and B commute, we have the standard joint probability formula for the simultaneous measurement of A and B. Thus, under the projection postulate, the successive measurements of A and B are considered effectively as their simultaneous measurement.

If we would restrict the class of measurements to those satisfying the projection postulate, any successive measurements of commuting observables could be considered as a simultaneous measurement. However, this approach has the following limitations: (i) Some of the most familiar measuring apparatuses such as photon counters do not satisfy the projection postulate. (ii) When the observable has a continuous spectrum, no measurements satisfy the repeatability hypothesis [3,4], so that the projection postulate cannot be formulated properly for measurements of continuous observables. (iii) The measurement of a function of an observable C such as $A=f(C)$ using the apparatus measuring C does not satisfy the projection postulate in general [2].

In fact, Einstein, Podolsky, and Rosen (EPR) [5] derived the joint probability formula for the outcomes of measurements of two observables pertaining to entangled subsystems *based on* the projection postulate. This EPR correlation is indeed one instance of the joint probability formula for the simultaneous measurement. The EPR correlation was experimentally tested by optical experiments [6,7]. However, those optical experiments use photon counting, and violate EPR's original assumption that the measurements satisfy the projection postulate. The recent realizations of quantum teleportation [8,9] were also optical realizations of the EPR correlation that violate EPR's assumption of the projection postulate. Thus, if we should be restricted to measurements satisfying the projection postulate, the scope of measurement theory would exclude most of the recent results in quantum information processing using entanglement [10].

The modern measurement theory [3,4,11–16] extends the scope from the measurements satisfying the projection postulate to more general measurements described by operations and effects, or more generally by operation-valued measures introduced from an axiomatic motivation [17–19]. Thus it is natural to expect to have a well-defined way to calculate probabilities in any combinations of measuring apparatuses, once we have identified the operation-valued measure in the general theory with the given model of measuring apparatus. However, determinations of the operation-valued measure so far have relied on the projection postulate [12] or the joint probability formula [3,20]. Thus the foundations of the modern approach in the present status involve the same difficulty as establishing the joint probability formula without assuming the projection postulate.

In this paper we shall abandon the projection postulate as a universal quantum rule and consider the following problem: *under what condition can a successive measurement of two or more observables be considered as a simultaneous measurement of those observables?* The prospective solution could be stated in the intuitive language that the preceding measurement does not disturb the observable to be measured later. However, in the current quantum mechanics very little is known about the disturbance caused by general measurement beyond the projection postulate. In order to answer the question in rigorous language, this paper will attempt to develop a theory of disturbance in general measurements with determining the possible state changes caused by measurements of observables. The justification of the joint probability formula and the EPR correlation will then follow without assuming the projection postulate.

Section II defines the simultaneous measurement for a pair of measuring apparatus. Sections III and IV discuss simultaneous measurements under the repeatability hypothesis and the projection postulate, indicating their limitations. The following three sections develop the theory of general measurement. Section V introduces the nonselective operations and their duals. Section VI discusses the Davies-Lewis postulate for the existence of operation-valued measures corresponding to apparatuses, and shows that the two justifications of their postulate known so far involve the same difficulty encountered the one in establishing the joint probability formula without assuming the projection postulate. Section VII gives a justification of the Davies-Lewis postulate without assuming the projection postulate or the joint probability formula, and proves the factoring property of operation-valued measures. Section VIII formulates the disturbance in the measurement, and establishes in rigorous language the relation between the disturbance and the joint probability formula. Sections IX and X apply the above result to the EPR correlation and the minimum disturbing measurement. Section XI concludes the paper with some remarks on the uncertainty principle.

II. STATISTICAL FORMULA FOR SIMULTANEOUS MEASUREMENTS

A. Born statistical formula

To formulate the problem precisely, let $\mathbf{S}$ be a quantum system with the Hilbert space $\mathcal{H}$ of state vectors. We shall distinguish measuring apparatuses by their own output variables [21], denoting by $\mathbf{A}(\mathbf{x})$ the apparatus measuring the system $\mathbf{S}$ with the output variable $\mathbf{x}$, which, we assume, takes values in the real line $\mathbf{R}$. We shall denote by ‘‘$\mathbf{x}(t)\in\Delta$’’ the probabilistic event that the outcome of the measurement using apparatus $\mathbf{A}(\mathbf{x})$ at time t is in a Borel set Δ in the real line $\mathbf{R}$. (Throughout this paper, ‘‘Borel set’’ can be replaced by ‘‘interval’’ for simplifying the presentation without any loss of generality.)

Let A be an observable of $\mathbf{S}$. The spectral projection of A corresponding to a Borel set Δ is denoted by $E^A(\Delta)$. According to the Born statistical formula, apparatus $\mathbf{A}(\mathbf{a})$, with an output variable $\mathbf{a}$, is said to *measure* an observable A at the time t if the relation

$$\Pr\{\mathbf{a}(t)\in\Delta\}=\mathrm{Tr}[E^A(\Delta)\rho(t)] \tag{1}$$

holds for the state $\rho(t)$ of the system $\mathbf{S}$ at the time t. The state $\rho(t)$ is called the *input state* to apparatus $\mathbf{A}(\mathbf{a})$, and is taken to be an arbitrary density operator.

The relation between the present formulation based on spectral projections due to von Neumann [1] and Dirac's formulation [22] is as follows. If the observable A has the Dirac-type spectral decomposition

$$A=\sum_\nu \sum_\mu \mu|\mu,\nu\rangle\langle\mu,\nu|+\sum_\nu \int \lambda|\lambda,\nu\rangle\langle\lambda,\nu|d\lambda,$$

where μ varies over the discrete eigenvalues, λ varies over the continuous eigenvalues, and ν is the degeneracy parameter, then we have

$$E^A(\Delta)=\sum_\nu \sum_{\mu\in\Delta} |\mu,\nu\rangle\langle\mu,\nu|+\sum_\nu \int_\Delta|\lambda,\nu\rangle\langle\lambda,\nu|d\lambda.$$

In this case, we have

$$\mathrm{Tr}[E^A(\Delta)\rho(t)]=\sum_\nu \sum_{\mu\in\Delta} \langle\mu,\nu|\rho(t)|\mu,\nu\rangle +\sum_\nu \int_\Delta\langle\lambda,\nu|\rho(t)|\lambda,\nu\rangle d\lambda.$$

B. Simultaneous measurements using one apparatus

Any commuting observables A and B are simultaneously measurable, and the joint probability distribution of the outcomes of their simultaneous measurement is given by

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t)\in\Delta'\}=\mathrm{Tr}[E^A(\Delta)E^B(\Delta')\rho(t)], \tag{2}$$

where Δ and Δ' are arbitrary Borel sets, and $\mathbf{a}$ and $\mathbf{b}$ denote the output variables of the apparatuses measuring A and B at time t, respectively.

A well-known proof of this formula from Eq. (1) runs as follows (Ref. [1], p. 228). Since A and B are commutable, there exist an observable C and real-valued functions f and g such that $A=f(C)$ and $B=g(C)$ (Ref. [1], p. 173). Their spectral projections satisfy the relations

$$E^A(\Delta)=E^C(f^{-1}(\Delta)),$$

$$E^B(\Delta')=E^C(g^{-1}(\Delta')).$$

For the outcome c of the C measurement, one defines the outcome of the A measurement to be $f(c)$ and the outcome of the B measurement to be $g(c)$. Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be the output variables of the measurements of A, B, and C, respectively. Then we have

$$\begin{aligned}\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t)\in\Delta'\}\\ &=\Pr\{\mathbf{c}(t)\in f^{-1}(\Delta),\mathbf{c}(t)\in g^{-1}(\Delta')\}\\ &=\Pr\{\mathbf{c}(t)\in f^{-1}(\Delta)\cap g^{-1}(\Delta')\}\\ &=\mathrm{Tr}[E^{C}(f^{-1}(\Delta)\cap g^{-1}(\Delta'))\rho(t)]\\ &=\mathrm{Tr}[E^{C}(f^{-1}(\Delta))E^{C}(g^{-1}(\Delta'))\rho(t)]\\ &=\mathrm{Tr}[E^{A}(\Delta)E^{B}(\Delta')\rho(t)].\end{aligned}$$

Thus their outcomes satisfy Eq. (2), so that the measurement of C at the time t gives a simultaneous measurement of A and B.

C. Simultaneous measurements using two apparatuses

The above proof gives one special instance of simultaneous measurement which uses one measuring apparatus with two output variables, but it is rather open when a pair of measuring apparatuses for A and B makes a simultaneous measurement of A and B.

In order to formulate this problem precisely, suppose that the observer measures A at the time t using the apparatus $\mathbf{A}(\mathbf{a})$. Let $t+\Delta t$ be the time just after the $\mathbf{A}(\mathbf{a})$ measurement. This means precisely that $t+\Delta t$ is the instant of the time just after the interaction is turned off between $\mathbf{A}(\mathbf{a})$ and $\mathbf{S}$, and that after $t+\Delta t$ the object $\mathbf{S}$ is free from the apparatus $\mathbf{A}(\mathbf{a})$. (Note that the last condition precludes the recoupling of the system with the apparatus.)

Let $\mathbf{A}(\mathbf{b})$ be another apparatus measuring an observable B of $\mathbf{S}$ with output variable $\mathbf{b}$. If the measurement using $\mathbf{A}(\mathbf{b})$ is turned on at the time $t+\Delta t$, the two measurements are called the *successive measurement* using $\mathbf{A}(\mathbf{a})$ and $\mathbf{A}(\mathbf{b})$ (in this order). Then the *successive measurement using* $\mathbf{A}(\mathbf{a})$ *and* $\mathbf{A}(\mathbf{b})$ *is defined to be a simultaneous measurement of A and B if and only if the joint probability distribution of their output variables* $\mathbf{a}$ *and* $\mathbf{b}$ *satisfies the standard joint probability formula*

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\}=\mathrm{Tr}[E^{A}(\Delta)E^{B}(\Delta')\rho(t)]. \tag{3}$$

It should be noted that the validity of the above relation depends on how the apparatus $\mathbf{A}(\mathbf{a})$ disturbs $\mathbf{S}$ during the first measurement, but does not depend on the property of the apparatus $\mathbf{A}(\mathbf{b})$ as long as $\mathbf{A}(\mathbf{b})$ measures the observable B in any input state. Therefore, the problem to be considered is to find a necessary and sufficient condition on the apparatus $\mathbf{A}(\mathbf{a})$, measuring A, in order for successive measurements using $\mathbf{A}(\mathbf{a})$ and an arbitrary apparatus $\mathbf{A}(\mathbf{b})$ measuring B to satisfy Eq. (3).

In the conventional approach, von Neumann (Ref. [1], p. 224) proved that if two observables are simultaneously measurable, then they are represented by commuting operators under the repeatability hypothesis. (The repeatability hypothesis will be discussed in detail in Sec. III.) The following theorem, though an easy consequence of the definition, shows that the simultaneous measurability extended by the above definition is still consistent with the old one.

Theorem 1. If the successive measurement of A and B using apparatuses $\mathbf{A}(\mathbf{a})$ *and* $\mathbf{A}(\mathbf{b})$, *respectively, is a simultaneous measurement of observables A and B, then A and B commute.*

Proof. Suppose that Eq. (3) holds. By the positivity of probability, both sides are non-negative. Since $\rho(t)$ is arbitrary, the product $E^{A}(\Delta)E^{B}(\Delta')$ is a positive self-adjoint operator so that $E^{A}(\Delta)$ and $E^{B}(\Delta')$ commute. Since Δ and Δ' are arbitrary, it follows that A and B commute. ■

In the following two sections, we shall re-examine the conventional approach from the operational point of view before starting with the general considerations.

III. SIMULTANEOUS MEASUREMENTS UNDER THE REPEATABILITY HYPOTHESIS

A. von Neumann's formulation

The conventional approach to measurement theory supposes that the measurement leaves the measured system in the eigenstate corresponding to the outcome of the measurement [1,23]. This assumption is equivalent to the *repeatability hypothesis* formulated by von Neumann (Ref. [1], p. 335) as follows: *If a physical quantity is measured twice in succession in a system, then we get the same value each time.*

Even though the repeatability hypothesis was posed as a universal law by von Neumann (Ref. [1], p. 213) based on the experiment of Compton and Simmons, in the modern approach it merely characterizes a class of measuring apparatuses. Thus, in what follows, *by saying that the apparatus* $\mathbf{A}(\mathbf{a})$ *satisfies the repeatability hypothesis, it is meant precisely that the repeatability hypothesis holds for the repeated measurement of A using the apparatus* $\mathbf{A}(\mathbf{a})$ *for the first A measurement.*

B. Repeatability hypothesis and joint probability

Suppose that the system $\mathbf{S}$ is measured at time t by an apparatus $\mathbf{A}(\mathbf{a})$. Let $t+\Delta t$ be the time just after the measurement. For any Borel set Δ, let $\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$ be the state at $t+\Delta t$ of $\mathbf{S}$ conditional upon $\mathbf{a}(t)\in\Delta$. Thus if the system $\mathbf{S}$ is sampled randomly from the subensemble of the similar systems that yield the outcome of the $\mathbf{A}(\mathbf{a})$ measurement in the Borel set Δ, then $\mathbf{S}$ is in the state $\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$ at time $t+\Delta t$. When $\Pr\{\mathbf{a}(t)\in\Delta\}=0$, the state $\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$ is indefinite, and we let $\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$ be an arbitrarily chosen density operator for mathematical convenience.

Suppose that the apparatus $\mathbf{A}(\mathbf{a})$ measures a discrete observable A with eigenvalues $a_1,a_2,\ldots$, and that the $\mathbf{A}(\mathbf{a})$ measurement at time t is followed immediately by an $\mathbf{A}(\mathbf{b})$ measurement measuring B. The conditional probability of $\mathbf{b}(t+\Delta t)\in\Delta'$, conditional upon the outcome $\mathbf{a}(t)=a_n$, is the probability of obtaining the outcome $\mathbf{b}(t+\Delta t)\in\Delta'$ in the state $\rho(t+\Delta|\mathbf{a}(t)=a_n)$, so that we have

$$\begin{aligned}\Pr\{\mathbf{b}(t+\Delta t)\in\Delta'|\mathbf{a}(t)=a_n\}\\ &=\mathrm{Tr}[E^{B}(\Delta')\rho(t+\Delta t|\mathbf{a}(t)=a_n)].\end{aligned} \tag{4}$$

The joint probability distribution of the outcomes of the successive measurement using $\mathbf{A}(\mathbf{a})$ and $\mathbf{A}(\mathbf{b})$ is given by the well-known relation

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} = \sum_{a_n\in\Delta}\Pr\{\mathbf{b}(t+\Delta t)\in\Delta'|\mathbf{a}(t)=a_n\}\Pr\{\mathbf{a}(t)=a_n\}. \tag{5}$$

Now, suppose that A is nondegenerate, and that apparatus $\mathbf{A}(\mathbf{a})$ satisfies the repeatability hypothesis. Then the state of the system just after the $\mathbf{A}(\mathbf{a})$ measurement, conditional upon the outcome $\mathbf{a}(t)=a_n$, is determined uniquely as the normalized eigenstate

$$\rho(t+\Delta t|\mathbf{a}(t)=a_n)=|\phi_n\rangle\langle\phi_n| \tag{6}$$

corresponding to the eigenstate a_n, provided $\Pr\{\mathbf{a}(t)=a_n\}>0$ (Ref. [1], pp. 215–217).

From Eqs. (4) and (6), we have

$$\begin{aligned}\sum_{a_n\in\Delta}&\Pr\{\mathbf{b}(t+\Delta t)\in\Delta'|\mathbf{a}(t)=a_n\}\Pr\{\mathbf{a}(t)=a_n\}\\ &=\sum_{a_n\in\Delta}\langle\phi_n|E^B(\Delta')|\phi_n\rangle\langle\phi_n|\rho(t)|\phi_n\rangle\\ &=\sum_{a_n\in\Delta}\mathrm{Tr}[|\phi_n\rangle\langle\phi_n|E^B(\Delta')|\phi_n\rangle\\ &\quad\times\langle\phi_n|\rho(t)].\end{aligned}$$

Thus, from Eq. (5) we have

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} = \sum_{a_n\in\Delta}\mathrm{Tr}[|\phi_n\rangle\langle\phi_n|E^B(\Delta')|\phi_n\rangle\langle\phi_n|\rho(t)]. \tag{7}$$

If A and B commute, we have

$$\begin{aligned}\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} &= \sum_{a_n\in\Delta}\mathrm{Tr}[|\phi_n\rangle\langle\phi_n|E^B(\Delta')\rho(t)]\\ &=\mathrm{Tr}[E^A(\Delta)E^B(\Delta')\rho(t)],\end{aligned}$$

and hence we obtain the joint probability formula [Eq. (3)].

By the above, we conclude that *if the apparatus* $\mathbf{A}(\mathbf{a})$, *measuring a nondegenerate discrete observable A, satisfies the repeatability hypothesis, the successive measurement using* $\mathbf{A}(\mathbf{a})$ *and* $\mathbf{A}(\mathbf{b})$ *measuring an arbitrary observable B commuting with A is a simultaneous measurement of A and B.*

C. Repeatability hypothesis and measuring processes

Consider a model of a measuring process in which an observable $A=\Sigma_n a_n|\phi_n\rangle\langle\phi_n|$ with nondegenerate eigenvalues a_n is measured by an apparatus $\mathbf{A}(\mathbf{a})$ having the nondegenerate probe observable $B=\Sigma_n a_n|\xi_n\rangle\langle\xi_n|$. The *probe observable* is generally defined to be the quantum-mechanical observable in the apparatus that is to be correlated with the measured observable by the measuring interaction, and amplified in a later stage of the apparatus to the directly sensible variable eventually read out by the observer [24]. In the conventional approach, the notion of the "pointer position" is used ambiguously instead, which sometimes means the "probe observable" and sometimes means the "directly sensible variable." Let $\mathbf{P}$ be the subsystem of the apparatus $\mathbf{A}(\mathbf{a})$ that includes the probe observable and actually interacts with the measured system $\mathbf{S}$. Suppose that $\mathbf{P}$ is prepared in the state $|\xi\rangle$ just before measurement. Let U be the unitary operator representing the time evolution of the composite system $\mathbf{S}+\mathbf{P}$ during the measuring interaction. The apparatus $\mathbf{A}(\mathbf{a})$ satisfies the repeatability hypothesis if and only if U satisfies

$$U:|\phi_n\rangle\otimes|\xi\rangle\mapsto e^{i\theta_n}|\phi_n\rangle\otimes|\xi_n\rangle, \tag{8}$$

where $e^{i\theta_n}$ is an arbitrary phase factor.

D. Measurements violating the repeatability hypothesis

A typical model which does not satisfy Eq. (8) is the photon-counting measurement in which, if the measurement takes place in the number state $|\phi_n\rangle=|n\rangle$, then the apparatus absorbs all the photons and outputs the amplified classical energy proportional to the number of absorbed photons. In this case, U satisfies

$$U:|n\rangle\otimes|\xi\rangle\mapsto|0\rangle\otimes|\xi_n\rangle, \tag{9}$$

and hence U does not satisfy Eq. (8).

For less idealized models of photon-counting measurement, we refer to Refs. [25,26]. For the case of measurements of continuous observables, models of exact position measurements that do not satisfy the repeatability hypothesis even approximately have also been constructed [15,27,28]. They were applied to the position monitoring that breaks the standard quantum limit [27,29].

E. Significance of the repeatability hypothesis

Given A, B, and $|\xi\rangle$, generally the apparatus $\mathbf{A}(\mathbf{a})$ measures the observable A, or equivalently satisfies Eq. (1), if and only if U satisfies

$$U:|\phi_n\rangle\otimes|\xi\rangle\mapsto|\phi_n'\rangle\otimes|\xi_n\rangle, \tag{10}$$

where $\{|\phi_n'\rangle\}$ is an arbitrary family of normalized vectors, not necessarily orthogonal. Thus, if we do not assume the repeatability hypothesis, the measurement correlates causally the input state $|\phi_n\rangle$ of the object before measurement with the output state $|\xi_n\rangle$ of the probe after measurement for some orthonormal basis $\{|\xi_n\rangle\}$. On the other hand, the repeatability hypothesis requires not only that the input state $|\phi_n\rangle$ is correlated to the output state $|\xi_n\rangle$ causally, but also that in the composite system after the measurement the input state $|\phi_n\rangle$ and the output state $|\xi_n\rangle$ are entangled to have a complete statistical correlation.

It is stated quite often that to measure an observable A is to change the input state $|\psi\rangle$ to an eigenstate $|\phi_n\rangle$ with the probability $|\langle\phi_n|\psi\rangle|^2$. This does not follow from the Born statistical formula [Eq. (1)], but assumes the repeatability hypothesis. Thus only when the measurement is assumed to satisfy the repeatability hypothesis, can we say that the measurement changes the state of the object probabilistically to one of the eigenstates of the measured observable.

Unless the repeatability hypothesis is assumed, it is, therefore, not a correct description that the measurement is to make a one-to-one correspondence (or to make an entanglement, in the modern language) between the state of the object before the measurement and the state of the probe after the measurement as described by Eq. (8), by which the problem of measuring the object is transferred to the problem of measuring the probe [30,31]. In what follows, a measurement is called *repeatable* if it is carried out by an apparatus satisfying the repeatability hypothesis.

IV. SIMULTANEOUS MEASUREMENTS UNDER THE PROJECTION POSTULATE

A. von Neumann's measurements of degenerate observables

For the observables with a nondegenerate purely discrete spectrum, the repeatability hypothesis determines the state after measurement uniquely; however, when the observable has degenerate eigenvalues, the state after measurement is not determined uniquely but depends on the ''actual measuring arrangement'' (Ref. [1], p. 348). von Neumann (Ref. [1], p. 348) considered the following means of measurement satisfying the repeatability hypothesis: Let $\{|\phi_{n,m}\rangle\}$ be an orthonormal basis, and let A be an observable represented by

$$A=\sum_{n,m} a_n|\phi_{n,m}\rangle\langle\phi_{n,m}|. \tag{11}$$

Suppose that the observer performs a repeatable measurement of a nondegenerate observable A' given by

$$A'=\sum_{n,m} a_{n,m}|\phi_{n,m}\rangle\langle\phi_{n,m}|, \tag{12}$$

where all $a_{n,m}$ are different, and that if the outcome of the A' measurement is $a_{n,m}$ then the outcome of the A measurement is taken to be a_n. Then we have a repeatable measurement of A.

Suppose that the observable A is measured at time t in the above way in a state (vector) $|\psi\rangle$ using the apparatus $\mathbf{A}(\mathbf{a})$. Then, at time $t+\Delta t$, just after the measurement, the object is left in the state (density operator)

$$\rho(t+\Delta t|\mathbf{a}(t)=a_n)=\frac{1}{\sum_m |c_{n,m}|^2}\sum_m |c_{n,m}|^2|\phi_{n,m}\rangle\langle\phi_{n,m}|, \tag{13}$$

where $c_{n,m}=\langle\phi_{n,m}|\psi\rangle$. This state depends not only on the observable A and the outcome a_n but also on the choice of the orthonormal basis $\{|\phi_{n,m}\rangle\}$ that satisfies Eq. (11). Since there are infinitely many essentially different choices of $\{|\phi_{n,m}\rangle\}$, the state change depends on the method of measurement even if the repeatability hypothesis holds.

In this case, the joint probability distribution of the outcomes of the $\mathbf{A}(\mathbf{a})$ measurement and the immediately following B measurement, using the apparatus $\mathbf{A}(\mathbf{b})$, is given by

$$\begin{aligned}\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\}\\ &=\sum_{a_n\in\Delta}\sum_m \mathrm{Tr}[|\phi_{n,m}\rangle\langle\phi_{n,m}|E^B(\Delta')|\phi_{n,m}\rangle\\ &\quad\times\langle\phi_{n,m}|\rho(t)].\end{aligned} \tag{14}$$

From the relation

$$E^A(\Delta)=\sum_{a_n\in\Delta}\sum_m |\phi_{n,m}\rangle\langle\phi_{n,m}|,$$

the joint probability formula, Eq. (3), holds for the arbitrary input state $\rho(t)$ if and only if A' and B commute. Although A and B commute, there are many choices of $\{|\phi_{n,m}\rangle\}$ such that A' and B do not commute. Thus, even if A and B commute and $\mathbf{A}(\mathbf{a})$ satisfies the repeatability hypothesis, the successive measurement using $\mathbf{A}(\mathbf{a})$ and $\mathbf{A}(\mathbf{b})$ cannot be a simultaneous measurement of A and B in general.

B. Lüders's formulation

The previous argument shows the existence of infinitely many different ways of measuring the same observable, which satisfies the repeatability hypothesis but does not satisfy the joint probability formula for the simultaneous measurement. Moreover, Lüders [2] pointed out that the observable corresponding to the identity operator I is considered to be measured without changing the input state, but any one of the above measurement for the identity changes the state unreasonably. Lüders suggested that the above measurement for a degenerate observable is always more disturbing than the desirable one. In order to determine the canonical way of measuring even the degenerate observables, Lüders proposed the following hypothesis: *If an observable A is measured in a state ρ, then at the time just after measurement the object is left in the state*

$$\frac{E^A\{a\}\rho E^A\{a\}}{\mathrm{Tr}[E^A\{a\}\rho]},$$

provided that the object leads to the outcome a with $\mathrm{Tr}[E^A\{a\}\rho]>0$.

In particular, if the object is measured in the vector state $|\psi\rangle$, then the state after measurement is the vector state $E^A\{a\}|\psi\rangle$ up to normalization. Thus the eigenstate corresponding to the outcome a is uniquely chosen as the projection, and hence the above hypothesis is called the *projection postulate*.

C. Projection postulate and joint probability

Suppose that a discrete observable A of $\mathbf{S}$ with eigenvalues $a_1, a_2, \ldots$ is measured at the time t by the apparatus $\mathbf{A}(\mathbf{a})$, measuring A satisfying the projection postulate. Then the state of the system at time $t+\Delta t$ just after the $\mathbf{A}(\mathbf{a})$ measurement conditional upon the outcome $\mathbf{a}(t)=a_n$ is

$$\rho(t+\Delta t|\mathbf{a}(t)=a_n)=\frac{E^A\{a_n\}\rho(t)E^A\{a_n\}}{\mathrm{Tr}[E^A\{a_n\}\rho(t)]}, \tag{15}$$

provided $\Pr\{\mathbf{a}(t)=a_n\}>0$.

From Eqs. (5) and (15), we have

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} = \sum_{a_n\in\Delta}\mathrm{Tr}[E^B(\Delta')E^A\{a_n\}\rho(t)E^A\{a_n\}]. \tag{16}$$

Thus the joint probability formula, Eq. (3), holds for the arbitrary input state $\rho(t)$ if and only if A and B commute. Therefore, we have seen the following theorem [2].

Theorem 2. The successive measurement of commuting observables A and B using apparatuses $\mathbf{A}(\mathbf{a})$ *and* $\mathbf{A}(\mathbf{b})$ *is a simultaneous measurement of A and B, if* $\mathbf{A}(\mathbf{a})$ *satisfies the projection postulate.*

We have seen that in order for a successive measurement of A and B to be a simultaneous measurement, the projection postulate for the apparatus measuring A is a sufficient condition. If we restrict our attention to the measuring apparatuses satisfying the projection postulate, any successive measurement of commuting observables is a simultaneous measurement, but, we have also seen that there are many kinds of measuring apparatuses which do not satisfy the projection postulate. Now we shall turn to the problem of under what condition a successive measurement of two or more observables can be considered as a simultaneous measurement.

V. NONSELECTIVE OPERATIONS OF MEASURING APPARATUSES

Every measuring process is considered to include an interaction, called the *measuring interaction*, between the measured system and the measuring apparatus. Let us consider the following description of the measuring interaction arising in the measurement using the apparatus $\mathbf{A}(\mathbf{a})$. In the following, the *probe* $\mathbf{P}$ is a microscopic subsystem of apparatus $\mathbf{A}(\mathbf{a})$ that actually interacts with $\mathbf{S}$. More precisely, we define the probe $\mathbf{P}$ to be the smallest subsystem of $\mathbf{A}(\mathbf{a})$ such that the composite system $\mathbf{S}+\mathbf{P}$ is isolated during the measuring interaction. Since we assume naturally that $\mathbf{S}+\mathbf{A}(\mathbf{a})$ is isolated during the measuring interaction, the smallest subsystem exists. The measurement is carried out by the interaction between the system $\mathbf{S}$ and the probe $\mathbf{P}$ and by the subsequent measurement on the probe $\mathbf{P}$. We assume that the probe system is a quantum mechanical system described by the Hilbert space $\mathcal{K}$ of state vectors. At the time of measurement t, the probe $\mathbf{P}$ is in a fixed state σ, so that the composite system is in the state

$$\rho_{\mathbf{S}+\mathbf{P}}(t)=\rho(t)\otimes\sigma. \tag{17}$$

The time evolution of the composite system $\mathbf{S}+\mathbf{P}$ during the interaction is described by a unitary operator U on $\mathcal{H}\otimes\mathcal{K}$. Hence, at the time just after the interaction, $t+\Delta t$, the composite system $\mathbf{S}+\mathbf{P}$ is in the state

$$\rho_{\mathbf{S}+\mathbf{P}}(t+\Delta t)=U[\rho(t)\otimes\sigma]U^\dagger. \tag{18}$$

The outcome of this measurement is obtained by a measurement of an observable M of $\mathbf{P}$ at time $t+\Delta t$. The observable M is called the *probe observable*. Thus, the output probability distribution of the apparatus $\mathbf{A}(\mathbf{a})$ is

$$\Pr\{\mathbf{a}(t)\in\Delta\}=\mathrm{Tr}\{[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger\}. \tag{19}$$

We shall call the above description of the measuring process the *indirect measurement model* determined by $(\mathcal{K},\sigma,U,M)$. In this model, from Eq. (18) the system $\mathbf{S}$ is in the state

$$\rho(t+\Delta t)=\mathrm{Tr}_{\mathcal{K}}\{U(\rho(t)\otimes\sigma)U^\dagger\} \tag{20}$$

at time $t+\Delta t$, where $\mathrm{Tr}_{\mathcal{K}}$ is the partial trace over $\mathcal{K}$. The state change $\rho(t)\mapsto\rho(t+\Delta t)$ is determined independently of the outcome of the measurement, and is called the *nonselective state change*.

As introduced previously, we have another type of state change $\rho(t)\mapsto\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$, where $\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$ is the state at $t+\Delta t$ of $\mathbf{S}$ conditional upon $\mathbf{a}(t)\in\Delta$. Since the condition $\mathbf{a}(t)\in\mathbf{R}$ makes no selection, we have

$$\rho(t+\Delta t|\mathbf{a}(t)\in\mathbf{R})=\rho(t+\Delta t). \tag{21}$$

For $\Delta\neq\mathbf{R}$, the state change $\rho(t)\mapsto\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)$ is called the *selective state change*.

We define the transformation $\rho\mapsto\mathbf{T}\rho$ for any trace class operator ρ on the state space $\mathcal{H}$ of $\mathbf{S}$ by the relation

$$\mathbf{T}\rho=\mathrm{Tr}_{\mathcal{K}}[U(\rho\otimes\sigma)U^\dagger]. \tag{22}$$

Then $\mathbf{T}$ is a trace preserving completely positive linear transformation on the space $\tau c(\mathcal{H})$ of trace class operators on $\mathcal{H}$ [11,12,21]. The transformation $\mathbf{T}$ is determined by the apparatus preparation σ and the measuring interaction U, and is called the *nonselective operation* of the apparatus $\mathbf{A}(\mathbf{a})$. The nonselective operation $\mathbf{T}$ represents the open system dynamics of the system $\mathbf{S}$ from t to $t+\Delta t$, and we have

$$\mathbf{T}\rho(t)=\rho(t+\Delta t). \tag{23}$$

As the converse of definition (22), it is well known that for every trace preserving completely positive linear transformation on $\tau c(\mathcal{H})$, there is an indirect measurement model such that $\mathbf{T}$ is the nonselective operation of that model [20,3,12].

For any bounded linear transformation $\mathbf{L}$ on $\tau c(\mathcal{H})$, its *dual* $\mathbf{L}^*$ is defined by

$$\mathrm{Tr}[(\mathbf{L}^*X)\rho]=\mathrm{Tr}[X(\mathbf{L}\rho)]$$

for all $\rho\in\tau c(\mathcal{H})$ and $X\in\mathcal{L}(\mathcal{H})$, where $\mathcal{L}(\mathcal{H})$ stands for the space of bounded operators on $\mathcal{H}$. Let $\mathbf{T}^*$ be the dual of the nonselective operation $\mathbf{T}$. Then $\mathbf{T}^*$ is the normal unit preserving completely positive linear transformation on $\mathcal{L}(\mathcal{H})$, such that

$$\mathrm{Tr}[(\mathbf{T}^*X)\rho]=\mathrm{Tr}[X(\mathbf{T}\rho)] \tag{24}$$

for all $\rho\in\tau c(\mathcal{H})$ and $X\in\mathcal{L}(\mathcal{H})$ (Ref. [11], p. 18). We call $\mathbf{T}^*$ the *dual nonselective operation*. Let $X\in\mathcal{L}(\mathcal{H})$ and $\rho\in\tau c(\mathcal{H})$. From Eq. (22) and a property of the partial trace, we have

$$\begin{aligned}\mathrm{Tr}[X\mathbf{T}\rho]&=\mathrm{Tr}[X\mathrm{Tr}_{\mathcal{K}}[U(\rho\otimes\sigma)U^\dagger]]\\&=\mathrm{Tr}[(X\otimes I)U(\rho\otimes\sigma)U^\dagger]\\&=\mathrm{Tr}[U^\dagger(X\otimes I)U(I\otimes\sigma)(\rho\otimes I)]\\&=\mathrm{Tr}[\mathrm{Tr}_{\mathcal{K}}[U^\dagger(X\otimes I)U(I\otimes\sigma)]\rho].\end{aligned}$$

Hence, from Eq. (24), we have

$$\mathrm{Tr}[(\mathbf{T}^*X)\rho]=\mathrm{Tr}[\mathrm{Tr}_{\mathcal{K}}[U^\dagger(X\otimes I)U(I\otimes\sigma)]\rho].$$

Since ρ is arbitrary, we have

$$\mathbf{T}^*X=\mathrm{Tr}_{\mathcal{K}}[U^\dagger(X\otimes I)U(I\otimes\sigma)] \tag{25}$$

for all $X\in\mathcal{L}(\mathcal{H})$. This characterizes the dual nonselective operation.

VI. OPERATION VALUED MEASURES

A. Davies-Lewis postulates

Davies and Lewis [19] postulated: *Given an apparatus* $\mathbf{A}(\mathbf{a})$ *for* $\mathbf{S}$, *there is a mapping* $\Delta\mapsto\mathbf{X}(\Delta)$ *from the Borel sets to the positive linear transformations on* $\tau c(\mathcal{H})$ *satisfying the following conditions.*

(a) *For any disjoint sequence of Borel sets* Δ_n *and for any* $\rho\in\tau c(\mathcal{H})$,

$$\mathbf{X}(\cup_n\Delta_n)\rho=\sum_n \mathbf{X}(\Delta_n)\rho. \tag{26a}$$

(b) *For any* $\rho\in\tau c(\mathcal{H})$,

$$\mathrm{Tr}[\mathbf{X}(\mathbf{R})\rho]=\mathrm{Tr}[\rho]. \tag{26b}$$

(c) *For any Borel set* Δ,

$$\Pr\{\mathbf{a}(t)\in\Delta\}=\mathrm{Tr}[\mathbf{X}(\Delta)\rho(t)]. \tag{26c}$$

(d) *For any Borel set* Δ *with* $\Pr\{\mathbf{a}(t)\in\Delta\}>0$,

$$\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)=\frac{\mathbf{X}(\Delta)\rho(t)}{\mathrm{Tr}[\mathbf{X}(\Delta)\rho(t)]}. \tag{26d}$$

We call the above mapping $\mathbf{X}:\Delta\mapsto\mathbf{X}(\Delta)$ the *operation valued measure* or the *operational distribution* of apparatus $\mathbf{A}(\mathbf{a})$. In general, we call any bounded linear transformation on $\tau c(\mathcal{H})$ a *superoperator* for $\mathcal{H}$. Any mapping $\Delta\mapsto\mathbf{X}(\Delta)$ from the Borel sets to the positive superoperators for $\mathcal{H}$ is called a *positive superoperator valued (PSV) measure* if it satisfies condition (a). Moreover, it is called *normalized* if it satisfies condition (b). Accordingly, the operation valued measure of $\mathbf{A}(\mathbf{a})$ is the normalized PSV measure satisfying conditions (c) and (d).

The validity of the Davies-Lewis postulate for the apparatuses with indirect measurement models was previously demonstrated *based on* the joint probability formula in Refs. [20,3], where it was also shown that any normalized PSV measures which are realizable by indirect measurement models are completely positive, and vice versa.

B. Determination of operation valued measures based on the projection postulate

In order to determine the operation valued measure corresponding to the given measuring apparatus, we need to describe the measuring process by an indirect measurement model. Then the measurement is divided into two processes: the measuring interaction in the object-probe composite system, and the probe measurement. Given the indirect measurement model, the current formulation has two arguments to determine the operation valued measure: one relies on the projection postulate (cf. Ref. [12] for yes-no measurements), and the other relies on the joint probability formula [20,3,15].

In the first approach, the probe measurement is explicitly assumed to satisfy the projection postulate. Consequently, the operation valued measure is determined by the unitary operator of the measuring interaction and the projection operator derived by the projection postulate with partial trace over the probe.

The argument runs as follows. Assume that the apparatus $\mathbf{A}(\mathbf{a})$ has the indirect measurement model $(\mathcal{K},\sigma,U,M)$, where the probe observable M is purely discrete with eigenvalues a_n. Let us suppose that the measuring interaction between the system $\mathbf{S}$ and the apparatus $\mathbf{A}(\mathbf{a})$ is turned on from t to $t+\Delta t$, and that the observer measures the probe observable M at the time $t+\Delta t$ using the apparatus $\mathbf{A}(\mathbf{m})$. Let $t+\Delta t+\tau$ be the time just after the measuring interaction is turned off between the probe $\mathbf{P}$ and the apparatus $\mathbf{A}(\mathbf{m})$. The system $\mathbf{A}(\mathbf{m})$ is considered as a subsystem of $\mathbf{A}(\mathbf{a})$, including the later stages after the probe $\mathbf{P}$. Assume that the apparatus $\mathbf{A}(\mathbf{m})$ satisfies the projection postulate and that the outcome is $\mathbf{m}(t+\Delta t)=a_n$. Then, at time $t+\Delta t+\tau$ the composite system $\mathbf{S}+\mathbf{P}$ is in the state

$$\rho_{\mathbf{S}+\mathbf{P}}(t+\Delta t+\tau|\mathbf{m}(t+\Delta t)=a_n)=\frac{(I\otimes E^M\{a_n\})\rho_{\mathbf{S}+\mathbf{P}}(t+\Delta t)(I\otimes E^M\{a_n\})}{\mathrm{Tr}[(I\otimes E^M\{a_n\})\rho_{\mathbf{S}+\mathbf{P}}(t+\Delta t)]}. \tag{27}$$

Since the outcome of the $\mathbf{A}(\mathbf{m})$ measurement at time $t+\Delta t$ is interpreted as the outcome of the $\mathbf{A}(\mathbf{a})$ measurement at time t, the condition $\mathbf{m}(t+\Delta t)=a_n$ is equivalent to the condition $\mathbf{a}(t)=a_n$. It follows that at time $t+\Delta t+\tau$ the system $\mathbf{S}$ is in the state

$$\rho(t+\Delta t+\tau|\mathbf{a}(t)=a_n)$$
$$=\mathrm{Tr}_{\mathcal{K}}[\rho_{\mathbf{S}+\mathbf{P}}(t+\Delta t+\tau|\mathbf{m}(t+\Delta t)=a_n)]. \quad (28)$$

From Eqs. (18), (27), and (28), we have

$$\rho(t+\Delta t+\tau|\mathbf{a}(t)=a_n)$$
$$=\frac{\mathrm{Tr}_{\mathcal{K}}[(I\otimes E^M\{a_n\})U[\rho(t)\otimes\sigma]U^\dagger(I\otimes E^M\{a_n\})]}{\mathrm{Tr}[(I\otimes E^M\{a_n\})U[\rho(t)\otimes\sigma]U^\dagger]}. \quad (29)$$

By the well-known relation

$$\mathrm{Tr}_{\mathcal{K}}[(I\otimes X)Y]=\mathrm{Tr}_{\mathcal{K}}[Y(I\otimes X)] \quad (30)$$

for all $X\in\mathcal{L}(\mathcal{K})$ and $Y\in\mathcal{L}(\mathcal{H}\otimes\mathcal{K})$, we have

$$\mathrm{Tr}_{\mathcal{K}}[(I\otimes E^M\{a_n\})U[\rho(t)\otimes\sigma]U^\dagger(I\otimes E^M\{a_n\})]$$
$$=\mathrm{Tr}_{\mathcal{K}}[(I\otimes E^M\{a_n\})U[\rho(t)\otimes\sigma]U^\dagger]. \quad (31)$$

Hence we have an important relation

$$\rho(t+\Delta t+\tau|\mathbf{a}(t)=a_n)=\frac{\mathrm{Tr}_{\mathcal{K}}[(I\otimes E^M\{a_n\})U[\rho(t)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[(I\otimes E^M\{a_n\})U[\rho(t)\otimes\sigma]U^\dagger]}. \quad (32)$$

Let Δ be an arbitrary Borel set such that $\Pr\{\mathbf{a}(t)\in\Delta\}>0$. Then we have, naturally,

$$\rho(t+\Delta t+\tau|\mathbf{a}(t)\in\Delta)$$
$$=\frac{\sum_{a_n\in\Delta}\Pr\{\mathbf{a}(t)=a_n\}\rho(t+\Delta t+\tau|\mathbf{a}(t)=a_n)}{\Pr\{\mathbf{a}(t)\in\Delta\}}. \quad (33)$$

From Eqs. (19), (32), and (33), we have

$$\rho(t+\Delta t+\tau|\mathbf{a}(t)\in\Delta)=\frac{\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}. \quad (34)$$

To obtain the final result, suppose that the $\mathbf{A}(\mathbf{m})$ measurement is instantaneous, i.e., $\tau\approx 0$, and that there is no interaction between $\mathbf{S}$ and the outside of $\mathbf{A}(\mathbf{a})$ from t to $t+\Delta t+\tau$. Then, in this time interval, the state changes of $\mathbf{S}$ are negligible due to the $\mathbf{A}(\mathbf{m})$ measurement, the time evolution of $\mathbf{S}$, and the decoherence from the environment. Consequently, we have

$$\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)=\rho(t+\Delta t+\tau|\mathbf{a}(t)\in\Delta). \quad (35)$$

Therefore, we have reached the final form

$$\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)=\frac{\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}. \quad (36)$$

From the above, the operation valued measure of $\mathbf{A}(\mathbf{a})$ is determined by

$$\mathbf{X}(\Delta)\rho=\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U(\rho\otimes\sigma)U^\dagger] \quad (37)$$

for all Borel sets Δ and all trace class operators ρ. Furthermore, it follows easily from properties of partial trace that $\mathbf{X}$ satisfies conditions (a)–(d).

Obviously, the above approach explicitly excludes the possibility of measuring the probe by an apparatus not satisfying the projection postulate such as a photon counter. Accordingly, this approach cannot apply correctly to any measurements with continuous probe observables such as the position of the probe pointer, since no apparatuses measuring continuous observables satisfy the repeatability hypothesis [3,4]. Moreover, the argument assumes that the probe measurement should be instantaneous.

C. Determination of operation valued measures based on the joint probability formula

In the second approach, generalizing von Neumann's argument on repeated measurements of the same observable (Ref. [1], pp. 211–223), it is assumed that the observer again measures an arbitrary observable of the object system at the time just after the measuring interaction. Then one can consider the joint probability distribution of the outcomes of the probe measurement and the second object measurement [20,3,15]. By assuming that the above joint probability distribution satisfies the joint probability formula for the simultaneous measurement, we can determine the operation valued measure.

Since the joint probability formula is well formulated even in the case where the probe observable has a continuous spectrum, the second approach can be applied to measurements of continuous observables. Moreover, in the case of the discrete probe observable, the second approach leads to the same operation-valued measure as the first approach, so that the second approach is consistent with the first.

The argument in the second approach runs as follows. Let $\mathbf{A}(\mathbf{a})$ be an apparatus described by the indirect measurement model $(\mathcal{K},\sigma,U,M)$. Suppose that the system $\mathbf{S}$ is measured at time t by the apparatus $\mathbf{A}(\mathbf{a})$. Suppose that at time $t+\Delta t$, just after the measuring interaction, the observer were to measure an arbitrary observable B of the same object $\mathbf{S}$ by an apparatus $\mathbf{A}(\mathbf{b})$. The conditional probability of $\mathbf{b}(t+\Delta t)\in\Delta'$ given $\mathbf{a}(t)\in\Delta$ is the probability of $\mathbf{b}(t+\Delta t)\in\Delta'$ in the state $\rho(t+\Delta|\mathbf{a}(t)\in\Delta)$, so that the joint probability distribution of $\mathbf{a}(t)$ and $\mathbf{b}(t+\Delta t)$ satisfies

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\}$$
$$=\mathrm{Tr}[E^B(\Delta')\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)]\Pr\{\mathbf{a}(t)\in\Delta\}. \quad (38)$$

For any Borel set Δ, let $\mathbf{X}(\Delta,\rho(t))$ be the trace class operator defined by

$$\mathbf{X}(\Delta,\rho(t))=\Pr\{\mathbf{a}(t)\in\Delta\}\rho(t+\Delta t|\mathbf{a}(t)\in\Delta). \quad (39)$$

From Eq. (38), we have

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\}=\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta,\rho(t))]. \tag{40}$$

On the other hand, by the indirect measurement model the output $\mathbf{a}(t)$ of this measurement is obtained by a measurement of the probe observable M at time $t+\Delta t$. Let $\mathbf{A}(\mathbf{m})$ be the apparatus measuring M at the time $t+\Delta t$. Then, the probabilistic event "$\mathbf{a}(t)\in\Delta$" is equivalent to the probabilistic event "$\mathbf{m}(t+\Delta t)\in\Delta$" and hence we have

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} =\Pr\{\mathbf{m}(t+\Delta t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\}. \tag{41}$$

Since the observable M of $\mathbf{P}$ and the observable B of $\mathbf{S}$ are simultaneously measurable, if the $\mathbf{A}(\mathbf{m})$ and $\mathbf{A}(\mathbf{b})$ measurements can be considered to be simultaneous, we have

$$\begin{aligned}\Pr\{\mathbf{m}(t+\Delta t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} &=\mathrm{Tr}[[E^B(\Delta')\otimes E^M(\Delta)]\rho_{\mathbf{S}+\mathbf{P}}(t+\Delta t)]\\ &=\mathrm{Tr}[[E^B(\Delta')\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger].\end{aligned}$$

By the property of partial trace, we have

$$\Pr\{\mathbf{m}(t+\Delta t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} =\mathrm{Tr}[E^B(\Delta')\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]. \tag{42}$$

Since B and Δ' are arbitrary, from Eqs. (40)–(42) we have

$$\mathbf{X}(\Delta,\rho(t))=\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]. \tag{43}$$

Suppose that $\Pr\{\mathbf{a}(t)\in\Delta\}>0$. From Eq. (39), we have

$$\begin{aligned}\rho(t+\Delta t|\mathbf{a}(t)\in\Delta)&=\frac{\mathbf{X}(\Delta,\rho(t))}{\mathrm{Tr}[\mathbf{X}(\Delta,\rho(t))]}\\ &=\frac{\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}.\end{aligned}$$

Hence we have shown that relation (36) holds for the apparatus $\mathbf{A}(\mathbf{a})$ given in this argument. Let $\mathbf{X}$ be the mapping $\Delta\mapsto\mathbf{X}(\Delta)$ defined by relation (37) for the present apparatus. Then $\mathbf{X}$ satisfies conditions (a) and (b) by the properties of partial trace as before. From Eq. (43), we have

$$\mathbf{X}(\Delta)\rho(t)=\mathbf{X}(\Delta,\rho(t)), \tag{44}$$

and hence $\mathbf{X}$ satisfies conditions (c) and (d). Thus $\mathbf{X}$ satisfies the Davies-Lewis postulate for the apparatus $\mathbf{A}(\mathbf{a})$ given above.

We have shown that the determination (37) of the operation-valued measure holds without assuming the projection postulate for the probe measurement. Nevertheless, in order to justify formula (37) generally, we need to justify the joint probability formula without assuming the projection postulate. This puts a serious constraint on the theoretical device to explore our problem. Indeed, because of the threat of a circular argument, the above arguments do not enable us to take advantage of operation-valued measures for the justification of the joint probability formula. In conventional measurement theory, a similar kind of circular argument has been known as the infinite regress of the von Neumann chain. Despite the above difficulties, in the following sections we shall show an alternative approach without any fear of a circular argument.

VII. STATISTICAL APPROACH TO THE OPERATION VALUED MEASURES

A. Existence of the operation valued measures

In what follows, we shall prove the Davies-Lewis postulate *without* assuming the joint probability formula or the projection postulate. Let us suppose that the system $\mathbf{S}$ is measured at time t by the apparatus $\mathbf{A}(\mathbf{a})$, and at time $t+\Delta t$ immediately after this measurement an observable B of $\mathbf{S}$ is measured using an apparatus $\mathbf{A}(\mathbf{b})$. Then the joint probability distribution of the outcomes of the A and B measurements satisfies Eq. (38). For any Borel set Δ, let $\mathbf{X}(\Delta,\rho(t))$ be the trace class operator defined by Eq. (39). Then, from Eq. (38), $\mathbf{X}(\Delta,\rho(t))$ satisfies Eq. (40). Since the input state $\rho(t)$ is assumed to be an arbitrary density operator, Eq. (39) defines the transformation $\mathbf{X}(\Delta)$ that maps $\rho(t)$ to $\mathbf{X}(\Delta,\rho(t))$. From Eqs. (39) and (40), $\mathbf{X}(\Delta)$ satisfies the relations

$$\mathbf{X}(\Delta)\rho(t)=\Pr\{\mathbf{a}(t)\in\Delta\}\rho(t+\Delta t|\mathbf{a}(t)\in\Delta) \tag{45}$$

and

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\}=\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)\rho(t)]. \tag{46}$$

Suppose that the input state $\rho(t)$ is a mixture of density operators ρ_1 and ρ_2, i.e.,

$$\rho(t)=\alpha\rho_1+(1-\alpha)\rho_2, \tag{47}$$

where $0<\alpha<1$. This means that at time t the measured object $\mathbf{S}$ is sampled randomly from an ensemble of similar systems described by the density operator ρ_1 with probability α, and from another ensemble described by the density operator ρ_2 with probability $1-\alpha$. Thus we have, naturally,

$$\begin{aligned}\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'|\rho(t)=\alpha\rho_1+(1-\alpha)\rho_2\} &=\alpha\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'|\rho(t)=\rho_1\}\\ &\quad+(1-\alpha)\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'|\rho(t)=\rho_2\},\end{aligned} \tag{48}$$

where $\Pr\{E|F\}$ stands for the conditional probability of E given F. From Eqs. (46) and (48), we have

$$\begin{aligned}\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)[\alpha\rho_1+(1-\alpha)\rho_2]] &=\alpha\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)\rho_1]+(1-\alpha)\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)\rho_2]\\ &=\mathrm{Tr}[E^B(\Delta')[\alpha\mathbf{X}(\Delta)\rho_1+(1-\alpha)\mathbf{X}(\Delta)\rho_2]].\end{aligned}$$

Since B and Δ' are arbitrary, we have

$$\mathbf{X}(\Delta)[\alpha\rho_1+(1-\alpha)\rho_2]=\alpha\mathbf{X}(\Delta)\rho_1+(1-\alpha)\mathbf{X}(\Delta)\rho_2. \tag{49}$$

It follows that $\mathbf{X}(\Delta)$ is an affine transformation from the space of density operators to the space of trace class operators, so that it can be extended to a unique positive superoperator [21].

We have proved that for any apparatus $\mathbf{A}(\mathbf{a})$ measuring A there is uniquely a family $\{\mathbf{X}(\Delta) | \Delta\in\mathcal{B}(\mathbf{R})\}$ of positive superoperators such that Eqs. (45) and (46) hold, where $\mathcal{B}(\mathbf{R})$ stands for the collection of all Borel sets.

By the countable additivity of probability, if $\Delta=\cup_n\Delta_n$ for disjoint Borel sets Δ_n, we have

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t+\Delta t)\in\Delta'\} = \sum_n \Pr\{\mathbf{a}(t)\in\Delta_n,\mathbf{b}(t+\Delta t)\in\Delta'\}. \tag{50}$$

From Eqs. (46) and (50), we have

$$\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)\rho(t)] = \sum_n \mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta_n)\rho(t)] = \mathrm{Tr}\left[E^B(\Delta')\sum_n \mathbf{X}(\Delta_n)\rho(t)\right].$$

Since B and Δ' are arbitrary, we have

$$\mathbf{X}(\Delta)\rho(t) = \sum_n \mathbf{X}(\Delta_n)\rho(t).$$

Since $\rho(t)$ is arbitrary, condition (a) holds for arbitrary density operator ρ, and hence by linearity, condition (a) holds for all $\rho\in\tau c(\mathcal{H})$. Conditions (c) and (d) are obvious from Eq. (39). From condition (c), we have

$$\mathrm{Tr}[\mathbf{X}(\mathbf{R})\rho(t)]=1.$$

Since $\rho(t)$ is arbitrary, condition (b) holds for arbitrary density operator ρ, and hence by linearity, condition (b) holds for all $\rho\in\tau c(\mathcal{H})$. Thus the mapping $\mathbf{X}:\Delta\mapsto\mathbf{X}(\Delta)\rho$ satisfies the Davies-Lewis postulate.

It should be noted that the present derivation rely on neither the existence of the indirect measurement model, the joint probability formula, nor the projection postulate. The crucial assumption in the above argument is Eq. (48), which follows from the basic principle underlying the notion of the mixture of states. Thus we can conclude that *every measuring apparatus has the operation valued measure satisfying the Davies-Lewis postulate*.

B. Basic properties of the operation valued measures

Let $\mathbf{A}(\mathbf{a})$ be a measuring apparatus for the system $\mathbf{S}$ with the operation-valued measure $\mathbf{X}$. Let us assume that the apparatus $\mathbf{A}(\mathbf{a})$ measures an observable A. In this case, from Eq. (1) and condition (c), we have

$$\mathrm{Tr}[\mathbf{X}(\Delta)\rho(t)] = \mathrm{Tr}[E^A(\Delta)\rho(t)]. \tag{51}$$

Let $\mathbf{X}(\Delta)^*$ be the dual of $\mathbf{X}(\Delta)$. Then we have

$$\mathrm{Tr}[(\mathbf{X}(\Delta)^*I)\rho(t)] = \mathrm{Tr}[E^A(\Delta)\rho(t)].$$

Since $\rho(t)$ is arbitrary, we conclude

$$\mathbf{X}(\Delta)^*I = E^A(\Delta) \tag{52}$$

for any Borel set Δ.

We say that a PSV measure $\mathbf{X}$ is *A compatible* if $\mathbf{X}$ satisfies relation (52). By the above, the operation-valued measure of the apparatus $\mathbf{A}(\mathbf{a})$ measuring A is an A compatible PSV measure. Now we are ready to state the following important relations for operation valued measures [33].

Theorem 3. Let A be an observable and let $\mathbf{X}$ be an A-compatible PSV measure. Then, for any Borel set Δ and any trace class operator ρ we have

$$\begin{aligned}\mathbf{X}(\Delta)\rho &= \mathbf{X}(\mathbf{R})[E^A(\Delta)\rho] = \mathbf{X}(\mathbf{R})[\rho E^A(\Delta)] \\ &= \mathbf{X}(\mathbf{R})[E^A(\Delta)\rho E^A(\Delta)],\end{aligned} \tag{53}$$

and for any bounded operator B we have

$$\begin{aligned}\mathbf{X}(\Delta)^*B &= [\mathbf{X}(\mathbf{R})^*B]E^A(\Delta) = E^A(\Delta)\mathbf{X}(\mathbf{R})^*B \\ &= E^A(\Delta)[\mathbf{X}(\mathbf{R})^*B]E^A(\Delta).\end{aligned} \tag{54}$$

A proof of the above theorem was given in Ref. [3] for the case where $\mathbf{X}(\Delta)$ is completely positive, and another proof was given in Ref. [32] for the case where A is discrete. The general proof necessary for the above theorem runs as follows.

Proof. Let C be a bounded operator such that $0\leqslant C\leqslant I$, and let $\Delta\in\mathcal{B}(\mathbf{R})$. We define

$$\begin{aligned} &A_{11}=\mathbf{X}(\Delta)^*C, && A_{12}=\mathbf{X}(\Delta)^*(I-C),\\ &A_{21}=\mathbf{X}(\mathbf{R}-\Delta)^*C, && A_{22}=\mathbf{X}(\mathbf{R}-\Delta)^*(I-C),\\ &P_1=E^A(\Delta), && P_2=I-E^A(\Delta),\\ &Q_1=\mathbf{X}(\mathbf{R})^*C, && Q_2=I-\mathbf{X}(\mathbf{R})^*C.\end{aligned}$$

Then, for $i,j=1,2$, we have $0\leqslant A_{ij}\leqslant P_i$, so that $[A_{ij},P_i]=[A_{ij},P_j]=0$. It follows that $Q_j=A_{1j}+A_{2j}$ commutes with P_1 and P_2 as well. Thus

$$A_{ij}=P_iA_{ij}\leqslant P_iQ_j.$$

On the other hand, we have $\Sigma_{ij}A_{ij}=I$ and $\Sigma_{ij}P_iQ_j=I$, whence $A_{ij}=P_iQ_j$. It follows that

$$\mathbf{X}(\Delta)^*C = E^A(\Delta)\mathbf{X}(\mathbf{R})^*C.$$

By taking the adjoint, we also have

$$\mathbf{X}(\Delta)^*C = [\mathbf{X}(\mathbf{R})^*C]E^A(\Delta).$$

Since any bounded operator B can be represented by $B=\Sigma_{n=0}^3\lambda_nC_n$ with positive operators $0\leqslant C_n\leqslant I$ and complex numbers λ_n, we have

$$\mathbf{X}(\Delta)^*B = E^A(\Delta)\mathbf{X}(\mathbf{R})^*B = [\mathbf{X}(\mathbf{R})^*B]E^A(\Delta)$$

for any $\Delta\in\mathcal{B}(\mathbf{R})$ and $B\in\mathcal{L}(\mathcal{H})$. By multiplying $E^A(\Delta)$ from both sides, we also have

$$\mathbf{X}(\Delta)^*B = E^A(\Delta)[\mathbf{X}(\mathbf{R})^*B]E^A(\Delta).$$

Hence relations (54) hold. Relations (53) follow easily by taking the duals of $\mathbf{X}(\Delta)^*$ and $\mathbf{X}(\mathbf{R})^*$. ■

By the above theorem, the operation valued measure $\mathbf{X}$ of an arbitrary apparatus $\mathbf{A}(\mathbf{a})$ measuring A is determined uniquely by the nonselective operation $\mathbf{T} = \mathbf{X}(\mathbf{R})$ of $\mathbf{A}(\mathbf{a})$. A mathematical theory of PSV measures was introduced by Davies and Lewis [19] based on conditions (a) and (b) as mathematical axioms; see also Davies [11]. Their relations with measuring processes were established in Refs. [3,4,13,14,20,34], and applied to analyzing various measuring processes in Refs. [15,27,28,35].

C. Operation valued measures of indirect measurement models

Suppose that the apparatus $\mathbf{A}(\mathbf{a})$ measuring A, has an indirect measurement model $(\mathcal{K},\sigma,U,M)$. In this case, we can determine the operation valued measure $\mathbf{X}$ of the apparatus $\mathbf{A}(\mathbf{a})$ without assuming the joint probability distribution or the projection postulate, as follows.

Let $\mathbf{X}$ be the operation valued measure of the apparatus $\mathbf{A}(\mathbf{a})$. Then $\mathbf{X}$ satisfies conditions (a)–(d), and hence $\mathbf{X}$ is an A-compatible PSV measure. It follows from Theorem 3 that $\mathbf{X}$ satisfies

$$\mathbf{X}(\Delta)\rho = \mathbf{X}(\mathbf{R})[E^A(\Delta)\rho], \tag{55}$$

where $\Delta \in \mathcal{B}(\mathbf{R})$ and $\rho \in \tau c(\mathcal{H})$. Since $\mathbf{A}(\mathbf{a})$ has the indirect measurement model $(\mathcal{K},\sigma,U,M)$, relation (20) holds. By condition (d) and Eqs. (20) and (21), we have

$$\mathbf{X}(\mathbf{R})\rho(t) = \mathrm{Tr}_{\mathcal{K}}[U[\rho(t)\otimes\sigma]U^\dagger].$$

Since $\rho(t)$ is arbitrary and $\mathbf{X}(\mathbf{R})$ is linear, the above relation can be extended to trace class operators ρ from density operators $\rho(t)$, so that we have

$$\mathbf{X}(\mathbf{R})\rho = \mathrm{Tr}_{\mathcal{K}}[U(\rho\otimes\sigma)U^\dagger] \tag{56}$$

for all $\rho \in \tau c(\mathcal{H})$. Now, we consider the expression

$$\mathcal{E}(\Delta)\rho = \mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U(\rho\otimes\sigma)U^\dagger], \tag{57}$$

where $\Delta \in \mathcal{B}(\mathbf{R})$ and $\rho \in \tau c(\mathcal{H})$. Then we can show, purely mathematically, that the mapping $\mathcal{E}:\Delta \mapsto \mathcal{E}(\Delta)$ defined above is an A-compatible PSV measure satisfying

$$\mathcal{E}(\mathbf{R})\rho = \mathrm{Tr}_{\mathcal{K}}[U(\rho\otimes\sigma)U^\dagger]. \tag{58}$$

Thus $\mathcal{E}$ satisfies the assumptions of Theorem 3, and hence we have

$$\mathcal{E}(\Delta)\rho = \mathcal{E}(\mathbf{R})[E^A(\Delta)\rho]. \tag{59}$$

From Eqs. (56) and (58), we have

$$\mathcal{E}(\mathbf{R}) = \mathbf{X}(\mathbf{R}) \tag{60}$$

and hence from Eqs. (55) and (59) we have

$$\mathcal{E}(\Delta)\rho = \mathcal{E}(\mathbf{R})[E^A(\Delta)\rho] = \mathbf{X}(\mathbf{R})[E^A(\Delta)\rho] = \mathbf{X}(\Delta)\rho.$$

Therefore, we conclude that $\mathbf{X}$ satisfies Eq. (37).

From Eqs. (37) and (53), we have the following expressions for $\mathbf{X}$:

$$\mathbf{X}(\Delta)\rho = \mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U(\rho\otimes\sigma)U^\dagger] \tag{61a}$$

$$= \mathrm{Tr}_{\mathcal{K}}[U[\rho(t)E^A(\Delta)\otimes\sigma]U^\dagger] \tag{61b}$$

$$= \mathrm{Tr}_{\mathcal{K}}[U[E^A(\Delta)\rho(t)\otimes\sigma]U^\dagger] \tag{61c}$$

$$= \mathrm{Tr}_{\mathcal{K}}[U[E^A(\Delta)\rho(t)E^A(\Delta)\otimes\sigma]U^\dagger]. \tag{61d}$$

Thus, if $\Pr\{\mathbf{a}(t)\in\Delta\} > 0$, we obtain the following relations:

$$\rho(t+\Delta t|\mathbf{a}(t)\in\Delta) = \frac{\mathrm{Tr}_{\mathcal{K}}[[I\otimes E^M(\Delta)]U[\rho(t)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[E^A(\Delta)\rho(t)]} \tag{62a}$$

$$= \frac{\mathrm{Tr}_{\mathcal{K}}[U[\rho(t)E^A(\Delta)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[E^A(\Delta)\rho(t)]} \tag{62b}$$

$$= \frac{\mathrm{Tr}_{\mathcal{K}}[U[E^A(\Delta)\rho(t)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[E^A(\Delta)\rho(t)]} \tag{62c}$$

$$= \frac{\mathrm{Tr}_{\mathcal{K}}[U[E^A(\Delta)\rho(t)E^A(\Delta)\otimes\sigma]U^\dagger]}{\mathrm{Tr}[E^A(\Delta)\rho(t)]}. \tag{62d}$$

VIII. DISTURBANCE IN MEASUREMENT

A. Disturbance and simultaneous measurability

Let B be an arbitrary observable of $\mathbf{S}$. We say that the measurement using an apparatus $\mathbf{A}(\mathbf{a})$ *does not disturb* the observable B if the nonselective state change does not perturb the probability distribution of B; that is, we have

$$\mathrm{Tr}[E^B(\Delta)\rho(t+\Delta t)] = \mathrm{Tr}[E^B(\Delta)e^{-iH\Delta t/\hbar}\rho(t)e^{iH\Delta t/\hbar}] \tag{63}$$

for any Borel set Δ, where H is the Hamiltonian of the system $\mathbf{S}$. The measurement is said to be *instantaneous* if the duration Δt of the measurement is negligible on the time scale of the time evolution of the system $\mathbf{S}$. Thus the instantaneous measurement using the apparatus $\mathbf{A}(\mathbf{a})$ does not disturb B if and only if

$$\mathrm{Tr}[E^B(\Delta)\rho(t+\Delta t)] = \mathrm{Tr}[E^B(\Delta)\rho(t)] \tag{64}$$

for any Borel set Δ.

Let $\mathbf{X}$ be the operation-valued measure of the apparatus $\mathbf{A}(\mathbf{a})$, and $\mathbf{T} = \mathbf{X}(\mathbf{R})$ be the nonselective operation of $\mathbf{A}(\mathbf{a})$. Then, from condition (d), we have

$$\rho(t+\Delta t) = \mathbf{T}\rho(t) \tag{65}$$

and hence Eq. (64) is equivalent to

$$\mathrm{Tr}[E^B(\Delta)\mathbf{T}\rho(t)] = \mathrm{Tr}[E^B(\Delta)\rho(t)]. \tag{66}$$

Let $\mathbf{T}^*$ be the dual nonselective operation of $\mathbf{A}(\mathbf{a})$. It follows from Eq. (66) that Eq. (64) is equivalent to

$$\mathrm{Tr}[[\mathbf{T}^* E^B(\Delta)]\rho(t)] = \mathrm{Tr}[E^B(\Delta)\rho(t)]. \tag{67}$$

Since $\rho(t)$ is arbitrary, Eq. (64) is equivalent to

$$\mathbf{T}^* E^B(\Delta) = E^B(\Delta). \tag{68}$$

Thus we conclude that *the instantaneous measurement using the apparatus* $\mathbf{A}(\mathbf{a})$ *with nonselective operation* $\mathbf{T}$ *does not disturb the observable B if and only if Eq.* (68) *holds for any Borel set* Δ. Now we are ready to state the answer to our problem.

Theorem 4. Let $\mathbf{A}(\mathbf{a})$ *be an apparatus measuring an observable A instantaneously, and let* $\mathbf{A}(\mathbf{b})$ *be an arbitrary apparatus measuring an observable B. Then the successive measurement using* $\mathbf{A}(\mathbf{a})$ *and* $\mathbf{A}(\mathbf{b})$ *is a simultaneous measurement of A and B if and only if* $\mathbf{A}(\mathbf{a})$ *does not disturb B.*

Proof. It suffices to show the equivalence between Eqs. (3) and 68. From Eqs. (46) and (55), we have

$$\begin{aligned}
&\Pr\{\mathbf{a}(t)\in\Delta, \mathbf{b}(t+\Delta t)\in\Delta'\}\\
&\quad = \mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)\rho(t)]\\
&\quad = \mathrm{Tr}[E^B(\Delta')\mathbf{X}(\mathbf{R})[\rho(t)E^A(\Delta)]]\\
&\quad = \mathrm{Tr}[[\mathbf{T}^* E^B(\Delta')]\rho(t)E^A(\Delta)]\\
&\quad = \mathrm{Tr}[E^A(\Delta)[\mathbf{T}^* E^B(\Delta')]\rho(t)]
\end{aligned}$$

Thus the joint probability distribution of A and B is given by

$$\Pr\{\mathbf{a}(t)\in\Delta, \mathbf{b}(t+\Delta t)\in\Delta'\} = \mathrm{Tr}[E^A(\Delta)[\mathbf{T}^* E^B(\Delta')]\rho(t)]. \tag{69}$$

If Eq. (68) holds, Eq. (3) follows immediately from Eq. (69). Conversely, suppose that Eq. (3) holds. By substituting $\Delta = \mathbf{R}$ in Eq. (3), we have

$$\Pr\{\mathbf{a}(t)\in\mathbf{R}, \mathbf{b}(t+\Delta t)\in\Delta'\} = \mathrm{Tr}[E^B(\Delta')\rho(t)]. \tag{70}$$

On the other hand, from Eq. (69) we have

$$\Pr\{\mathbf{a}(t)\in\mathbf{R}, \mathbf{b}(t+\Delta t)\in\Delta'\} = \mathrm{Tr}[[\mathbf{T}^* E^B(\Delta')]\rho(t)]. \tag{71}$$

Since $\rho(t)$ is arbitrary, from Eqs. (70) and (71) we obtain Eq. (68). Therefore, Eqs. (3) and (68) are equivalent. ∎

From Theorems 1 and 4, we can see that if the apparatus $\mathbf{A}(\mathbf{a})$, instantaneously measuring an observable A, does not disturb an observable B, then A and B necessarily commute. Therefore, we can conclude the following statement.

Theorem 5. Every apparatus measuring an observable disturbs all the observables that do not commute with the measured observable.

B. Disturbance in indirect measurements

From Eq. (20), and by the property of the partial trace, we have

$$\begin{aligned}
&\mathrm{Tr}[E^B(\Delta)\rho(t+\Delta t)]\\
&\quad = \mathrm{Tr}[E^B(\Delta)\mathrm{Tr}_{\mathcal{K}}[U[\rho(t)\otimes\sigma]U^\dagger]]\\
&\quad = \mathrm{Tr}[[E^B(\Delta)\otimes I]U[\rho(t)\otimes\sigma]U^\dagger]\\
&\quad = \mathrm{Tr}[U^\dagger[E^B(\Delta)\otimes I]U(I\otimes\sigma)[\rho(t)\otimes I]]\\
&\quad = \mathrm{Tr}[\mathrm{Tr}_{\mathcal{K}}[U^\dagger[E^B(\Delta)\otimes I]U(I\otimes\sigma)]\rho(t)].
\end{aligned}$$

Hence Eq. (64) is equivalent to

$$\mathrm{Tr}(\mathrm{Tr}_{\mathcal{K}}\{U^\dagger[E^B(\Delta)\otimes I]U(I\otimes\sigma)\}\rho(t)) = \mathrm{Tr}[E^B(\Delta)\rho(t)].$$

Since $\rho(t)$ is arbitrary, Eq. (64) is equivalent to

$$\mathrm{Tr}_{\mathcal{K}}\{U^\dagger[E^B(\Delta)\otimes I]U(I\otimes\sigma)\} = E^B(\Delta) \tag{72}$$

for any Borel set Δ.

Obviously, from Eq. (72), if U and $B\otimes I$ commute, i.e.,

$$[U, E^B(\Delta)\otimes I] = 0 \tag{73}$$

for any Borel set Δ, then the A measurement does not disturb the observable B. However, Eq. (73) is not a necessary condition for nondisturbing measurement. In the case where σ is a pure state $\sigma = |\xi\rangle\langle\xi|$, from Eq. (72) we have the following theorem.

Theorem 6. Let $\mathbf{A}(\mathbf{a})$ *be an apparatus measuring an observable A instantaneously with indirect measurement model* $(\mathcal{K}, |\Phi\rangle\langle\Phi|, U, M)$. *The Apparatus* $\mathbf{A}(\mathbf{a})$ *does not disturb an observable B if and only if*

$$[U, E^B(\Delta)\otimes I]|\psi\otimes\Phi\rangle = 0 \tag{74}$$

for any Borel set Δ *and any state vector* ψ *of* $\mathbf{S}$.

Proof. First, we note that in the case where $\sigma = |\Phi\rangle\langle\Phi|$, relation (72) holds if and only if

$$\langle\psi\otimes\Phi|U^\dagger[E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle = \langle\psi|E^B(\Delta)|\psi\rangle \tag{75}$$

holds for any state vector ψ. Suppose that Eq. (74) holds. We have

$$U[E^B(\Delta)\otimes I]|\psi\otimes\Phi\rangle = [E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle.$$

Multiplying $U^\dagger$ from the left, we have

$$[E^B(\Delta)\otimes I]|\psi\otimes\Phi\rangle = U^\dagger[E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle,$$

and, hence, we have Eq. (75). Thus if Eq. (74) holds for any Borel set Δ and any state vector ψ, then $\mathbf{A}(\mathbf{a})$ does not disturb B. Conversely, suppose that $\mathbf{A}(\mathbf{a})$ does not disturb B. Then, from Eq. (72), with $\sigma = |\Phi\rangle\langle\Phi|$, we have

$$\begin{aligned}
&\langle\phi'\otimes\Phi|U^\dagger[E^B(\Delta)\otimes I]U|\phi\otimes\Phi\rangle\\
&\quad = \langle\phi'\otimes\Phi|E^B(\Delta)\otimes I|\phi\otimes\Phi\rangle
\end{aligned}$$

for any vectors $\phi, \phi' \in \mathcal{H}$. Let ψ be a state vector. If $|\phi\rangle = |\psi\rangle$ and $|\phi'\rangle = E^B(\Delta)|\psi\rangle$, we have

$$\langle \psi\otimes\Phi|[E^B(\Delta)\otimes I]U^\dagger[E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle = \langle \psi\otimes\Phi|E^B(\Delta)\otimes I|\psi\otimes\Phi\rangle. \tag{76}$$

By taking a complex conjugate, we have

$$\langle \psi\otimes\Phi|U^\dagger[E^B(\Delta)\otimes I]U[E^B(\Delta)\otimes I]|\psi\otimes\Phi\rangle = \langle \psi\otimes\Phi|E^B(\Delta)\otimes I|\psi\otimes\Phi\rangle. \tag{77}$$

From Eqs. (75)–(77), we have

$$\begin{aligned}
&\|\{E^B(\Delta)\otimes I - U^\dagger[E^B(\Delta)\otimes I]U\}|\psi\otimes\Phi\rangle\|^2 \\
&= \langle \psi\otimes\Phi|E^B(\Delta)\otimes I|\psi\otimes\Phi\rangle \\
&\quad - \langle \psi\otimes\Phi|[E^B(\Delta)\otimes I]U^\dagger[E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle \\
&\quad - \langle \psi\otimes\Phi|U^\dagger[E^B(\Delta)\otimes I]U[E^B(\Delta)\otimes I]|\psi\otimes\Phi\rangle \\
&\quad + \langle \psi\otimes\Phi|U^\dagger[E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle \\
&= 0.
\end{aligned}$$

Thus we have

$$[E^B(\Delta)\otimes I]|\psi\otimes\Phi\rangle = U^\dagger[E^B(\Delta)\otimes I]U|\psi\otimes\Phi\rangle.$$

Multiplying U from the left, we have

$$\{U[E^B(\Delta)\otimes I] - [E^B(\Delta)\otimes I]U\}|\psi\otimes\Phi\rangle = 0,$$

and hence we have Eq. (74). Therefore, we conclude that if $\mathbf{A}(\mathbf{a})$ does not disturb B, then Eq. (74) holds for any Borel set Δ and any state vector ψ of $\mathbf{S}$. ■

IX. LOCAL MEASUREMENTS OF OBSERVABLES OF TWO ENTANGLED SYSTEM

If the two observables to be measured belong to two different subsystems, then they commute each other, and the measurement of one is not considered to disturb the other in general, so that the result obtained in Sec. VIII applies to this situation. The purpose of this section is to state this fact in the rigorous language.

Let C be an observable of an system $\mathbf{S}_1$ with Hilbert space $\mathcal{H}_1$, and D an observable of another system $\mathbf{S}_2$ with Hilbert space $\mathcal{H}_2$. Suppose that the composite system $\mathbf{S}=\mathbf{S}_1+\mathbf{S}_2$ is in a state $\rho(t)$ at time t. Let us suppose that one measures the observable C at time t using an apparatus $\mathbf{A}(\mathbf{a})$ and that at time $t+\Delta t$, just after the C measurement, one measures D using any apparatus $\mathbf{A}(\mathbf{b})$ measuring D. We assume that after time t there is no interaction between $\mathbf{S}_1$ and $\mathbf{S}_2$.

First we shall consider the case where the measurement of C satisfies the projection postulate. In this case, in the composite system $\mathbf{S}_{12}$, the observable $A=C\otimes I_2$ is measured at time t, and the observable $B=I_1\otimes D$ is measured immediately after the A measurement, where I_1 and I_2 are the identity operators on $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. From Theorem 2 the joint probability distribution satisfies

$$\Pr\{\mathbf{a}(t)\in\Delta, \mathbf{b}(t+\Delta t)\in\Delta'\} = \mathrm{Tr}[[E^C(\Delta)\otimes E^D(\Delta')]\rho(t)]. \tag{78}$$

In order to compare this result with the argument given by EPR [5], let us consider the special case where C and D are nondegenerate observables in their own subsystems and the initial state $\rho(t)$ is a pure state. In this case, the state $\rho(t)$ is represented by a state vector $\Psi(t)$ in the Hilbert space $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$ as

$$\rho(t) = |\Psi(t)\rangle\langle\Psi(t)|.$$

Let us suppose that the observables C and D have the spectral decompositions

$$C = \sum_n a_n |\phi_n\rangle\langle\phi_n|,$$

$$D = \sum_m b_m |\xi_m\rangle\langle\xi_m|.$$

EPR expanded $\Psi(t)$ using the basis $\{\phi_n\}$ of $\mathcal{H}_1$ as

$$\Psi(t) = \sum_n |\phi_n\otimes\eta_n\rangle, \tag{79}$$

where η_n are uniquely determined vectors in $\mathcal{H}_2$ not necessarily orthogonal and, according to EPR, are to be regarded merely as coefficients of the expansion of $\Psi(t)$ into a series of orthogonal vectors ϕ_n. Then EPR considered the process of ‘‘reduction of the wave packet’’

$$\sum_n |\phi_n\otimes\eta_n\rangle \mapsto N|\phi_n\otimes\eta_n\rangle, \tag{80}$$

where N is the normalization constant determined up to a phase factor by

$$N = \|\phi_n\otimes\eta_n\|^{-1}, \tag{81}$$

and stated that the state after the measurement conditional upon the outcome $\mathbf{a}(t)=a_n$ is determined as

$$\Psi(t+\Delta t|\mathbf{a}(t)=a_n) = N|\phi_n\otimes\eta_n\rangle, \tag{82}$$

where

$$|\Psi(t+\Delta t|\mathbf{a}(t)=a_n)\rangle\langle\Psi(t+\Delta t|\mathbf{a}(t)=a_n)| = \rho(t+\Delta t|\mathbf{a}(t)=a_n). \tag{83}$$

From this, we have the joint probability formula

$$\Pr\{\mathbf{a}(t)=a_n, \mathbf{b}(t+\Delta t)=b_m\} = |\langle\phi_n\otimes\xi_m|\Psi(t)\rangle|^2, \tag{84}$$

which is a special case of Eq. (78).

Now, let us show that the EPR argument is equivalent to the argument based on the projection postulate for the A measurement. From the projection postulate, if the outcome of the $\mathbf{A}(\mathbf{a})$ measurement is a_n, the state of the composite system at the time just after the measurement is

$$\Psi(t+\Delta t|\mathbf{a}(t)=a_n) = \frac{(|\phi_n\rangle\langle\phi_n|\otimes I_2)\Psi(t)}{\|(|\phi_n\rangle\langle\phi_n|\otimes I_2)\Psi(t)\|}. \tag{85}$$

Then, from Eq. (79), we have

$$(|\phi_n\rangle\langle\phi_n|\otimes I_2)\Psi(t)=|\phi_n\otimes\eta_n\rangle. \tag{86}$$

Thus we have shown that Eq. (82) is the consequence from the projection postulate Eq. (85).

In the following, we shall consider the general case. For instance, consider the case where the A measurement leaves the system $\mathbf{S}_1$ in a fixed state ϕ_1 independent of the outcome such as the vacuum state after photon counting. Does Eq. (78) hold even in this case? The answer to this question might depend on the method of measuring A. However, if the measurement of A is carried out so as not to affect the system $\mathbf{S}_2$, then from the result in Sec. VIII we will be able to conclude relation (78). In order to ensure that the measurement of A does not affect the system $\mathbf{S}_2$, we introduce the following condition.

We will say that the apparatus $\mathbf{A}(\mathbf{a})$, measuring A, is local in the system $\mathbf{S}_1$ if the measuring interaction is confined in the system $\mathbf{S}_1$ and the apparatus $\mathbf{A}(\mathbf{a})$, as formulated precisely as follows. Let $\mathcal{K}$ be the Hilbert space of the probe $\mathbf{P}$ in apparatus $\mathbf{A}(\mathbf{a})$, and suppose that $\mathbf{P}$ is prepared in state σ at time t of the measurement, and let U be the unitary operator of $\mathcal{K}\otimes\mathcal{H}_1\otimes\mathcal{H}_2$ representing the time evolution of the composite system $\mathbf{S}+\mathbf{P}$. Then the apparatus $\mathbf{A}(\mathbf{a})$ is said to be *local* in the system $\mathbf{S}_1$ if we have

$$[U,I_1\otimes X\otimes I_{\mathcal{K}}]=0 \tag{87}$$

for any bounded operator X on $\mathcal{H}_2$, where $I_{\mathcal{K}}$ is the identity on $\mathcal{K}$.

Theorem 7. Suppose that the composite system $\mathbf{S}=\mathbf{S}_1+\mathbf{S}_2$ *is in state* $\rho(t)$ *at time t of the measurement. Let C and D be observables of* $\mathbf{S}_1$ *and* $\mathbf{S}_2$, *respectively. If the apparatus* $\mathbf{A}(\mathbf{a})$, *measuring* $A=C\otimes I_2$ *instantaneously, is local in the system* $\mathbf{S}_1$ *then Eq.* (78) *holds.*

Proof. Let σ be the state of the probe at t. From Theorem 5 it suffices to show that $\mathbf{A}(\mathbf{a})$ does not disturb the observable $B=I\otimes D$. By assumption, we have

$$[U,E^B(\Delta)\otimes I_{\mathcal{K}}]=[U,I_1\otimes E^D(\Delta)\otimes I_{\mathcal{K}}]=0$$

for any Borel set Δ. Thus relation (73) holds, so that $\mathbf{A}(\mathbf{a})$ does not disturb the observable $B=I_1\otimes D$. Therefore, Eq. (78) follows from Theorem 4. ∎

From the above theorem, we also have the following statement: *Any pair of local instantaneous measuring apparatuses of* $A=C\otimes I_2$ *and* $B=I_1\otimes D$ *satisfies the joint probability formula*

$$\Pr\{\mathbf{a}(t)\in\Delta,\mathbf{b}(t)\in\Delta'\}=\mathrm{Tr}[[E^C(\Delta)\otimes E^D(\Delta')]\rho(t)] \tag{88}$$

regardless of the order of the measurement, where we identify t with $t+\Delta t$.

In the EPR paper [5], the so-called EPR correlation is derived theoretically under the assumption that the pair of measurements satisfies the projection postulate, but the present result concludes that the EPR correlation holds for any pair of local instantaneous measurements, as experiments have already suggested.

X. MINIMUM DISTURBING MEASUREMENTS

Classical measurements are usually considered to disturb no measured systems. This does not mean, however, that no classical measurement disturbs the system, but that among all the possible measurements the minimum disturbing measurement does not disturb the system in principle. In this section, we shall introduce the notion of the minimum disturbing measurement in quantum mechanics, and show that this is equivalent to a measurement satisfying the projection postulate.

For an apparatus $\mathbf{A}(\mathbf{x})$, we denote by $\mathcal{D}(\mathbf{x})$ the set of observables that are disturbed by $\mathbf{A}(\mathbf{x})$, i.e., $\mathcal{D}(\mathbf{x})$ is the set of observables B such that $\mathbf{T}^*E^B(\Delta)\neq E^B(\Delta)$ for some Borel set Δ, where $\mathbf{T}$ is the nonselective operation of $\mathbf{A}(\mathbf{x})$ and $\mathbf{T}^*$ its dual. Let A be an observable of the system $\mathbf{S}$, and let $\mathbf{A}(\mathbf{a})$ be an apparatus measuring A instantaneously. The apparatus $\mathbf{A}(\mathbf{a})$ is called *minimum disturbing* if $\mathcal{D}(\mathbf{a})\subset\mathcal{D}(\mathbf{x})$ for any apparatus $\mathbf{A}(\mathbf{x})$ measuring A instantaneously. Then we have the following statement.

Theorem 8. Let $\mathbf{A}(\mathbf{a})$ *be an apparatus measuring a discrete observable A instantaneously. The apparatus* $\mathbf{A}(\mathbf{a})$ *is minimum disturbing if and only if* $\mathbf{A}(\mathbf{a})$ *satisfies the projection postulate.*

Proof. Let $\mathcal{C}(A)$ be the set of observables that do not commute with A. From Theorem 5, we have

$$\mathcal{C}(A)^c\subset\mathcal{D}(\mathbf{x}) \tag{89}$$

for any apparatus $\mathbf{A}(\mathbf{x})$ measuring A instantaneously, where superscript c stands for the complement in the set of observables. Let $\mathbf{A}(\mathbf{a})$ be an apparatus measuring A instantaneously. Suppose that $\mathbf{A}(\mathbf{a})$ satisfies the projection postulate. Then, from Theorem 2, we have

$$\mathcal{D}(\mathbf{a})\subset\mathcal{C}(A)^c, \tag{90}$$

and hence from Eq. (89) we conclude that $\mathbf{A}(\mathbf{a})$ is minimum disturbing and

$$\mathcal{D}(\mathbf{a})=\mathcal{C}(A)^c. \tag{91}$$

Conversely, suppose that $\mathbf{A}(\mathbf{a})$ is minimum disturbing. We have an indirect measurement model that measures A instantaneously and satisfying the projection postulate [3]. Hence there is an apparatus $\mathbf{A}(\mathbf{x})$ measuring A instantaneously such that $\mathcal{D}(\mathbf{x})=\mathcal{C}(A)^c$. By assumption, $\mathbf{A}(\mathbf{a})$ is minimum disturbing, so that $\mathcal{D}(\mathbf{a})=\mathcal{C}(A)^c$. Then the operation valued measure $\mathbf{X}$ of $\mathbf{A}(\mathbf{a})$ is such that $\mathbf{X}(\mathbf{R})^*E^B(\Delta')=E^B(\Delta')$ for all $B\in\mathcal{C}(A)$ and $\Delta'\in\mathcal{B}(\mathbf{R})$. Thus we have

$$\begin{aligned}
&\mathrm{Tr}[E^B(\Delta')\mathbf{X}(\Delta)\rho(t)]\\
&\quad=\mathrm{Tr}[[\mathbf{X}(\Delta)^*E^B(\Delta')]\rho(t)]\\
&\quad=\sum_{a\in\Delta}\mathrm{Tr}[[\mathbf{X}\{a\}^*E^B(\Delta')]\rho(t)]\\
&\quad=\sum_{a\in\Delta}\mathrm{Tr}[(E^A\{a\}[\mathbf{X}(\mathbf{R})^*E^B(\Delta')]E^A\{a\})\rho(t)]
\end{aligned}$$

$$= \sum_{a \in \Delta} \mathrm{Tr}[E^A\{a\}E^B(\Delta')E^A\{a\}\rho(t)]$$

$$= \mathrm{Tr}\left[E^B(\Delta') \sum_{a \in \Delta} E^A\{a\}\rho(t)E^A\{a\}\right].$$

Since B and Δ' are arbitrary, we have

$$\mathbf{X}(\Delta)\rho(t) = \sum_{a \in \Delta} E^A\{a\}\rho(t)E^A\{a\},$$

and hence

$$\rho(t+\Delta t|\mathbf{a}\in\Delta) = \frac{\sum_{a \in \Delta} E^A\{a\}\rho(t)E^A\{a\}}{\mathrm{Tr}[E^A(\Delta)\rho]}.$$

Thus $\mathbf{A}(\mathbf{a})$ satisfies the projection postulate. ■

We refer to Refs. [3,19] for different approaches to the minimum disturbance condition. The present approach leads to the simplest characterization of the measurements satisfying the projection postulate, which can eventually be called the *minimum disturbing measurements*.

XI. CONCLUDING REMARKS

As anticipated from the ordinary interpretation of the uncertainty principle or the principle of complementarity formulated by the noncommutativity of observables, every measurement of an observable disturbs every observable that does not commute with the measured observable. It should be noted, however, that this does not imply a prevailing interpretation of the Heisenberg uncertainty principle that the measurement of the position with accuracy ϵ must bring about an indeterminacy $\eta = \hbar/2\epsilon$ in the value of the momentum (Ref. [1], p. 239). In fact, we can construct an indirect measurement model of the postion measurement that counters the above statement [36]; this model has complete accuracy, $\epsilon = 0$, but disturbs the momentum arbitrarily small if the input state is arbitarily close to the momentum eigenstate. This example suggests that the relation between the accuracy and the disturbance is more complicated than the relation $\epsilon\eta \geqslant \hbar/2$ suggested by the Robertson uncertainty relation [37]. A detailed investigation will be presented in a forthcoming paper.

[1] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University, Princeton, 1955).
[2] G. Lüders, Ann. Phys. (Leipzig) **8**, 322 (1951).
[3] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
[4] M. Ozawa, Publ. Res. Inst. Math. Sci., Kyoto Univ. **21**, 279 (1985).
[5] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[6] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **49**, 91 (1982).
[7] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
[8] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, Nature (London) **390**, 575 (1997).
[9] D. Boschi, S. Branca, F. De Martini, L. Hardy, and S. Popescu, Phys. Rev. Lett. **80**, 1121 (1998).
[10] C. H. Bennett and D. DiVincenzo, Nature (London) **404**, 247 (2000).
[11] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
[12] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Physics Vol. 190 (Springer, Berlin, 1983).
[13] M. Ozawa, J. Math. Phys. **26**, 1948 (1985).
[14] M. Ozawa, J. Math. Phys. **27**, 759 (1986).
[15] M. Ozawa, in *Squeezed and Nonclassical Light*, edited by P. Tombesi and E. R. Pike (Plenum, New York, 1989), p. 263.
[16] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, Lecture Notes in Physics Vol. 2 (Springer, Berlin, 1991).
[17] R. Haag and D. Kastler, J. Math. Phys. **5**, 848 (1964).
[18] G. Ludwig, Commun. Math. Phys. **4**, 331 (1967).
[19] E. B. Davies and J. T. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[20] M. Ozawa, in *Probability Theory and Mathematical Statistics*, Lecture Notes in Mathematics Vol. 1021, edited by K. Itô and J. V. Prohorov (Springer, Berlin, 1983), p. 518.
[21] M. Ozawa, Phys. Rev. A **62**, 062101 (2000).
[22] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. (Oxford University Press, Oxford, 1958).
[23] E. Schrödinger, Naturwissenschaften **23**, 807 (1935); **23**, 823 (1935); **23**, 844 (1935) [Proc. Am. Philos. Soc. **124**, 323 (1980)].
[24] M. Ozawa, in *Fifth International Conference on Squeezed States and Uncertainty Relations*, edited by D. Han, J. Jansky, Y. S. Kim, and V. I. Man'ko (NASA, Goddard, 1998), p. 517.
[25] M. D. Srinivas and E. B. Davies, Opt. Acta **28**, 981 (1981).
[26] N. Imoto, M. Ueda, and T. Ogawa, Phys. Rev. A **41**, 4127 (1990).
[27] M. Ozawa, Phys. Rev. Lett. **60**, 385 (1988).
[28] M. Ozawa, Phys. Rev. A **41**, 1735 (1990).
[29] J. Maddox, Nature (London) **331**, 559 (1988).
[30] F. London and E. Bauer, *La Théorie de l'Observation en Mécanique Quantique* (Hermann, Paris, 1939) [in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, 1983), p. 217].
[31] E. P. Wigner, Am. J. Phys. **31**, 6 (1963).
[32] M. Ozawa, Ann. Phys. (N.Y.) **259**, 121 (1997).
[33] M. Ozawa, in *Quantum Communication, Computing, and Measurement 3*, edited by P. Tombesi and O. Hirota (Kluwer/Plenum, New York, 2001).
[34] M. Ozawa, in *Quantum Communications and Measurement*, edited by V. P. Belavkin, O. Hirota, and R. L. Hudson (Plenum, New York, 1995), p. 109.
[35] M. Ozawa, J. Math. Phys. **34**, 5596 (1993).
[36] M. Ozawa, e-print, LANL quant-ph/9805033.
[37] H. Robertson, Phys. Rev. **34**, 163 (1929).