Available online at www.sciencedirect.com
SCIENCE DIRECT®

Annals of Physics 321 (2006) 744–769

ANNALS of PHYSICS

www.elsevier.com/locate/aop

Quantum perfect correlations ☆

Masanao Ozawa *

Graduate School of Information Sciences, Tôhoku University, Aoba-ku, Sendai 980-8579, Japan

Received 5 May 2005; accepted 21 August 2005
Available online 2 November 2005

Abstract

The notion of perfect correlations between arbitrary observables, or more generally arbitrary POVMs, is introduced in the standard formulation of quantum mechanics, and characterized by several well-established statistical conditions. The transitivity of perfect correlations is proved to generally hold, and applied to a simple articulation for the failure of Hardy's nonlocality proof for maximally entangled states. The notion of perfect correlations between observables and POVMs is used for defining the notion of a precise measurement of a given observable in a given state. A longstanding misconception on the correlation made by the measuring interaction is resolved in the light of the new theory of quantum perfect correlations.

1. Introduction

It is often stressed that quantum mechanics does not speak of the value of an observable in a single event, but only speaks of the average value over a large number of events. In fact, quantum states are characterized as what determine the expectation values of all the observables. However, quantum correlations definitely describe relations of values of observables in a single event as typically in the EPR correlation [1]. In the early days of quantum mechanics, the quantum correlation played a central role in measurement theory since von Neumann [2] generally described a process of making a perfect correlation between two systems. In the recent investigations on quantum information [3], the notion

☆ A preliminary account on this subject has appeared in M. Ozawa, Phys. Lett. A 335, (2005) 11–19.

* Fax: +81227954628.
E-mail address: ozawa@math.is.tohoku.ac.jp.

0003-4916/$ - see front matter
doi:10.1016/j.aop.2005.08.007

of quantum correlations naturally plays a key role, as in classical information theory the amount of information is defined as a measure of statistical correlations for pairs of random variables. Nevertheless, we have not had a general notion of quantum correlation; in those investigations the quantum correlation has rather replaced by the notion of entanglement, which is regarded as quantum correlations restricted to those between commuting observables from different subsystems.

The main aim of this paper is to establish the general notion of quantum perfect correlations. It should be stressed that statistical correlation is a state dependent notion, and it is required to address the problem as to when a pair of observables are considered to be perfectly correlated in a given state. The operational meaning of this condition is that those two observables can be jointly measured in that state and that each joint measurement gives the same value, although the value may distribute randomly. In classical probability theory, it is well accepted that two random variables (observables) are perfectly correlated if and only if the joint probability of any pair of their different values vanishes. Thus, we can immediately generalize this notion to pairs of commuting observables based on the well-defined joint probability distribution of commuting observables. It is well-known that every entangled (pure) state of a bipartite system has the Schmidt decomposition that determines naturally a pair of perfectly correlated observables in respective subsystems. The perfect correlation relevant to the study of entanglement is as such always those for commuting observables. Nevertheless, we have several problems that strongly demand the generalization of the notion of perfect correlations to noncommuting observables.

One of them is the transitivity problem of quantum perfect correlations. Suppose that commuting observables X and Y are perfectly correlated as well as commuting observables Y and Z. If X and Z commute, we can easily say that X and Z are perfectly correlated. However, there are cases where X and Z do not commute, and no existing theory determines whether X and Z are considered to be perfectly correlated.

There has been a longstanding misconception on statistical correlation in measurement. In the conventional model of measurement found by von Neumann [2], the measuring interaction is required to establish two different kinds of perfect correlations: one is between measured observable *before* the interaction and the meter observable *after* the interaction, and the other is between the meter observable *after* the interaction and the measured observable *after* the interaction. The first one ensures that the observation of the meter observable suffices to know the value of the measured observable before the interaction, and the second one ensures that the measurement leaves the measured system in the eigenstate pertaining to the measurement result. However, we have been able to treat only the second correlation, since the Heisenberg operator of the measured observable before the interaction and the Heisenberg operator of the meter observable after the interaction do not commute in general. Moreover, there has been a confusion between the meaning of those two different correlations. Even in the modern approach to measurement theory, the lack of the general theory of quantum perfect correlations has left the fundamental question unanswered as to when the given observable is precisely measured in a given state.

This paper introduces the notion of perfect correlations between arbitrary two observables, and characterizes it by various statistical notions in quantum mechanics. As a result, the above problems are shown to be answered by simple and well-founded conditions in the standard formalism of quantum mechanics.

In Section 2, we introduce the definition of the perfect correlation between two observables. In Section 3, the condition that two observables are perfectly correlated in a given state is characterized in terms of well-formulated statistical notions in the standard quantum mechanics. It is immediate from the definition that two perfectly correlated observables are identically distributed, i.e., having the same probability distribution, but the converse is not true as seen from the case of two independent observables with identical distribution. This section considers the question as to what additional condition ensures that two identically distributed observables are perfectly correlated.

In Section 4, we prove that the perfect correlation between observables in a given state is transitive and consequently is an equivalence relation between observables. In Section 5, we consider the joint probability distribution of perfectly correlated observables, and show that two observables are perfectly correlated if and only if they have joint probability distribution concentrated on the diagonal. We show that our definition of perfectly correlated observables in a given state ensures that they are jointly measurable in that state. We also characterize the quasi-joint probability distribution of perfectly correlated observables. In Section 6, we consider the perfect correlation between observables in bipartite systems, and characterize pairs of perfectly correlated observables from two subsystems. We also apply the transitivity of perfect correlations to a simple explanation for the failure of Hardy's nonlocality proof for the class of maximally entangled states [4].

In Section 7, we consider the perfect correlations between probability operator valued measures (POVMs). We show that any pair of POVMs has a joint dilation to a pair of observables in an extended system in such a way that the given POVMs are perfectly correlated if and only if the corresponding observables are perfectly correlated. In this way, the problem of perfect correlations between POVMs can be reduced to the problem of perfect correlations between observables, and we extend the characterization of perfectly correlations between two observables to those between a POVM and an observable.

In Section 8, we consider perfect correlations in measurements, and gives the definition for precise measurements of an observable in a given state, using the notion of perfect correlations between observables and POVMs. A longstanding misconception on the correlation made by the measuring interaction is resolved in the light of the new theory of quantum perfect correlations. Section 9 concludes the present paper with summary and some remarks.

2. Basic formulations

Let $\mathcal{H}$ be a separable Hilbert space. An *observable* is a self-adjoint operator densely defined in $\mathcal{H}$ and a *state* is a density operator ρ on $\mathcal{H}$, or equivalently a positive operator ρ on $\mathcal{H}$ with unit trace [2]. A unit vector ψ in $\mathcal{H}$ is called a *state vector* or a *vector state* defining the state $\rho = |\psi\rangle\langle\psi|$ that is an extreme point (*pure state*) in the convex set $\mathcal{S}(\mathcal{H})$ of states on $\mathcal{H}$. Denote by $\mathcal{B}(\mathbf{R}^n)$ the Borel σ-field of the Euclidean space $\mathbf{R}^n$ and by $B(\mathbf{R}^n)$ the algebra of (complex-valued) bounded Borel functions on $\mathbf{R}^n$. Denote by $\mathcal{L}(\mathcal{H})$ the algebra of bounded operators on $\mathcal{H}$ and by $\mathcal{L}(\mathcal{H})_+$ the cone of positive operators on $\mathcal{H}$. A *positive operator valued measure* [5] is a mapping Π from $\mathcal{B}(\mathbf{R})$ to $\mathcal{L}(\mathcal{H})_+$ such that $\Pi(\bigcup_j \Delta_j) = \sum_{j=1}^{\infty} \Pi(\Delta_j)$ in the weak operator topology for any disjoint sequence of Borel sets $\Delta_1, \Delta_2, \ldots$. A *probability operator valued measure* (*POVM*) [6,7] is a positive operator valued measure Π such that $\Pi(\mathbf{R}) = I$.

We say that two POVMs Π_1 and Π_2 are *perfectly correlated* in a state ρ iff

$$\mathrm{Tr}[\Pi_1(\Delta)\Pi_2(\Gamma)\rho] = 0 \tag{1}$$

for any disjoint Borel sets Δ, Γ. For any vector state ψ, Eq. (1) is equivalent to

$$\langle \Pi_1(\Delta)\psi, \Pi_2(\Gamma)\psi\rangle = 0. \tag{2}$$

The following proposition generalizes Eq. (1) to arbitrary pairs of Borel sets Δ, Γ.

Proposition 2.1. *For any POVMs Π_1, Π_2, and any state ρ, the following conditions are equivalent.*

(i) *Π_1 and Π_2 are perfectly correlated in ρ.*
(ii) $\mathrm{Tr}[\Pi_1(\Delta)\Pi_2(\Gamma)\rho] = \mathrm{Tr}[\Pi_1(\Delta\cap\Gamma)\rho]$ *for any $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$.*
(iii) $\mathrm{Tr}[\Pi_1(\Delta)\Pi_2(\Gamma)\rho] = \mathrm{Tr}[\Pi_2(\Delta\cap\Gamma)\rho]$ *for any $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$.*

Proof. If Π_1, Π_2 are perfectly correlated, we have

$$\begin{aligned}\mathrm{Tr}[\Pi_1(\Delta)\Pi_2(\Gamma)\rho] &= \mathrm{Tr}[\Pi_1(\Delta\cap\Gamma)\Pi_2(\Gamma)\rho] + \mathrm{Tr}[\Pi_1(\Delta\setminus\Gamma)\Pi_2(\Gamma)\rho]\\ &= \mathrm{Tr}[\Pi_1(\Delta\cap\Gamma)\Pi_2(\Gamma)\rho]\\ &= \mathrm{Tr}[\Pi_1(\Delta\cap\Gamma)\Pi_2(\mathbf{R}\setminus\Gamma)\rho] + \mathrm{Tr}[\Pi_1(\Delta\cap\Gamma)\Pi_2(\Gamma)\rho]\\ &= \mathrm{Tr}[\Pi_1(\Delta\cap\Gamma)\rho]\end{aligned}$$

for any $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$. This proves (i) $\Rightarrow$ (ii). The converse part (ii) $\Rightarrow$ (i) is obvious, and the equivalence (i) $\iff$ (iii) can be proved analogously. □

Let Π be a positive operator valued measure. For any Borel function f on $\mathbf{R}$ the operator $\Pi(f)$ is defined by

$$\mathrm{dom}(\Pi(f)) = \left\{\psi \in \mathcal{H} \,\middle|\, \int_{\mathbf{R}} |f(x)|^2 \langle\psi, \mathrm{d}\Pi(x)\psi\rangle < \infty\right\},$$
$$\langle\psi', \Pi(f)\psi\rangle = \int_{\mathbf{R}} f(x)\langle\psi', \mathrm{d}\Pi(x)\psi\rangle$$

for all $\psi \in \mathrm{dom}(\Pi(f))$ and $\psi' \in \mathcal{H}$; see [8] for comparison with other approaches. For the identity function id on $\mathbf{R}$, i.e., $\mathrm{id}(x) = x$ for all $x \in \mathbf{R}$, the operator $\Pi(\mathrm{id}^n)$ is called the nth *moment operator* of Π. For any real-valued Borel function f on $\mathbf{R}$, the relation

$$\Pi^f(\Delta) = \Pi(f^{-1}(\Delta)), \tag{3}$$

where $\Delta \in \mathcal{B}(\mathbf{R})$, defines a unique positive operator valued measure Π^f. For any real-valued Borel functions f, g, it is easy to see that $\Pi(f\circ g) = \Pi^g(f) = \Pi^{f\circ g}(\mathrm{id})$, where $f\circ g$ is the composition of f and g, i.e., $f\circ g(x) = f(g(x))$ for all $x \in \mathbf{R}$. For any bounded operator A on $\mathcal{H}$, the relation

$$\Pi^A(\Delta) = A^\dagger\Pi(\Delta)A, \tag{4}$$

where $\Delta \in \mathcal{B}(\mathbf{R})$, defines a unique positive operator valued measure Π^A. For any bounded operator A, B, we have $\Pi^{AB} = (\Pi^A)^B$. If Π is a POVM, so are Π^f and Π^U whenever U is isometry.

Now we have the following.

Theorem 2.2. *For any POVMs Π_1,Π_2, state ρ, and unitary operator U on $\mathcal{H}$, the following conditions are equivalent.*

(i) *Π_1 and Π_2 are perfectly correlated in ρ.*
(ii) *Π_1^f and Π_2^f are perfectly correlated in ρ for any real-valued Borel function f.*
(iii) *Π_1^f and Π_2^f are perfectly correlated in ρ for any bounded real-valued Borel function f.*
(iv) *Π_1^f and Π_2^f are perfectly correlated in ρ for a bijective Borel function f from $\mathbf{R}$ to a Borel set $\Omega \in \mathcal{B}(\mathbf{R})$.*
(v) *Π_1^U and Π_2^U are perfectly correlated in $U^\dagger \rho U$.*

Proof. Suppose that Π_1 and Π_2 are perfectly correlated in ρ. Let f be a real-valued Borel function. We have

$$\begin{aligned}\mathrm{Tr}[\Pi_1^f(\Delta)\Pi_2^f(\Gamma)\rho] &= \mathrm{Tr}[\Pi_1(f^{-1}(\Delta))\Pi_2(f^{-1}(\Gamma))\rho] = \mathrm{Tr}[\Pi_1(f^{-1}(\Delta)\cap f^{-1}(\Gamma))\rho]\\ &= \mathrm{Tr}[\Pi_1(f^{-1}(\Delta\cap\Gamma))\rho] = \mathrm{Tr}[\Pi_1^f(\Delta\cap\Gamma)\rho].\end{aligned}$$

Thus, Π_1^f and Π_2^f are perfectly correlated in ρ. This proves (i)⇒(ii). The implications (ii)⇒(iii)⇒(iv) are obvious. Suppose (iv). Then, there is a Borel function g such that $g[f(x)] = x$ for all $x \in \mathbf{R}$. By the implication (i) ⇒ (ii), two POVMs $\Pi_1 = \Pi_1^{g\circ f}$ and $\Pi_2 = \Pi_2^{g\circ f}$ are perfectly correlated in ρ. This proves (iv) ⇒ (i). The equivalence (i) $\iff$ (v) is straightforward from the property of trace and the proof is completed. □

Let X be an observable on $\mathcal{H}$. The *spectral measure* of X is the projection-valued POVM E^X such that $E^X(p) = p(X)$ for any polynomial p. For any Borel function f on $\mathbf{R}$, the operator $f(X)$ is defined by $f(X) = E^X(f)$.

We say that two observables X and Y are *perfectly correlated* in a state ρ iff E^X and E^Y are perfectly correlated in ρ. From Proposition 2.1, X and Y are perfectly correlated in ρ if and only if one of the following equivalent conditions holds:

(i) $\mathrm{Tr}[E^X(\Delta)E^Y(\Gamma)\rho] = 0$ for any disjoint Borel sets $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$.
(ii) $\mathrm{Tr}[E^X(\Delta)E^Y(\Gamma)\rho] = \mathrm{Tr}[E^X(\Delta\cap\Gamma)\rho]$ for any $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$.
(iii) $\mathrm{Tr}[E^X(\Delta)E^Y(\Gamma)\rho] = \mathrm{Tr}[E^Y(\Delta\cap\Gamma)\rho]$ for any $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$.

The following theorem restates Theorem 2.2 for observables.

Theorem 2.3. *For any observables X, Y, state ρ, and unitary operator U on $\mathcal{H}$, the following conditions are all equivalent.*

(i) *X and Y are perfectly correlated in ρ.*
(ii) *$f(X)$ and $f(Y)$ are perfectly correlated in ρ for any real-valued Borel function f.*
(iii) *$f(X)$ and $f(Y)$ are perfectly correlated in ρ for any bounded real-valued Borel function f.*
(iv) *$f(X)$ and $f(Y)$ are perfectly correlated in ρ for a bijective Borel function f from $\mathbf{R}$ to a Borel set $\Omega \in \mathcal{B}(\mathbf{R})$.*
(v) *$U^\dagger XU$ and $U^\dagger YU$ are perfectly correlated in $U^\dagger \rho U$.*

From the above theorem, the perfect correlation between two not necessarily bounded observables X and Y can be reduced to the perfect correlation of a pair of bounded observables, say, $\tan^{-1}X$ and $\tan^{-1}Y$.

3. Characterizations of perfectly correlated observables

The *cyclic subspace* of $\mathcal{H}$ spanned by an observable X and a state vector $\psi \in \mathcal{H}$ is the closed subspace $\mathcal{C}(X, \psi)$ defined by

$$\mathcal{C}(X, \psi) = \text{the closure of } \{f(X)\psi \in \mathcal{H} | f \in B(\mathbf{R})\}.$$

Denote by $\mathcal{C}_1(X, \psi)$ the unit sphere of $\mathcal{C}(X, \psi)$ and by $P_{X,\psi}$ the projection of $\mathcal{H}$ onto $\mathcal{C}(X, \psi)$. A closed subspace of $\mathcal{H}$ is said to be *invariant* under X iff it is invariant under all projections $E^X(\Delta)$ for $\Delta \in \mathcal{B}(\mathbf{R})$. Since $\mathcal{C}(X, \psi)$ is invariant under X, the projection $P_{X,\psi}$ commutes with $E^X(\Delta)$ for all $\Delta \in \mathcal{B}(\mathbf{R})$. Then we obtain the following theorem.

Theorem 3.1. *For any two observables X and Y on $\mathcal{H}$ and any state vector $\psi \in \mathcal{H}$, the following conditions are equivalent.*

(i) *X and Y are perfectly correlated in ψ.*
(ii) *X and Y are perfectly correlated in any $\phi \in \mathcal{C}_1(X, \psi)$.*
(iii) *$E^X(\Delta)\psi = E^Y(\Delta)\psi$ for any $\Delta \in \mathcal{B}(\mathbf{R})$.*
(iv) *$f(X)\psi = f(Y)\psi$ for any $f \in B(\mathbf{R})$.*
(v) *$f(X)P_{X,\psi} = f(Y)P_{X,\psi}$ for any $f \in B(\mathbf{R})$.*
(vi) *$P_{X,\psi} = P_{Y,\psi}$ and $XP_{X,\psi} = YP_{Y,\psi}$.*

Proof. Suppose (i) holds. Let $\Delta \in \mathcal{B}(\mathbf{R})$. Then, we have

$$\begin{aligned}&\|E^X(\Delta)\psi - E^Y(\Delta)\psi\|^2 \\ &\quad = \|E^X(\Delta)\psi\|^2 - \langle E^X(\Delta)\psi, E^Y(\Delta)\psi\rangle - \langle E^Y(\Delta)\psi, E^X(\Delta)\psi\rangle + \|E^Y(\Delta)\psi\|^2 = 0.\end{aligned}$$

Thus, we have $E^X(\Delta)\psi = E^Y(\Delta)\psi$ for every $\Delta \in \mathcal{B}(\mathbf{R})$, and the implication (i) $\Rightarrow$ (iii) follows. Suppose (iii) holds. The set of Borel functions $f \in B(\mathbf{R})$ satisfying $f(X)\psi = f(Y)\psi$ is closed under the linear combination, the uniform convergence, and includes all characteristic functions χ_Δ for $\Delta \in \mathcal{B}(\mathbf{R})$, so that $f(X)\psi = f(Y)\psi$ holds for every $f \in B(\mathbf{R})$. Thus, the implication (iii) $\Rightarrow$ (iv) follows. Suppose that condition (iv) holds. Then, we have $f(X)g(X)\psi = f(Y)g(Y)\psi = f(Y)g(X)\psi$ for any $f, g \in B(\mathbf{R})$. Since every $\phi \in \mathcal{C}(X, \psi)$ is a limit of vectors of the form $\phi = g(X)\psi$ for some $g \in B(\mathbf{R})$, we have $f(X)P_{X,\psi} = g(Y)P_{X,\psi}$. Thus, the implication (iv) $\Rightarrow$ (v) follows. Suppose that condition (v) holds. The implication (v) $\Rightarrow$ (iv) trivially holds, and hence we have $\mathcal{C}(X, \psi) = \mathcal{C}(Y, \psi)$ and $P_{X,\psi} = P_{Y,\psi}$. Letting $f = \chi_\Delta$ in condition (v), we have $E^X(\Delta)P_{X,\psi} = E^Y(\Delta)P_{Y,\psi}$, and hence the spectral measures of the self-adjoint operators $XP_{X,\psi}$ and $YP_{Y,\psi}$ are the same, so that they are identical. Thus, the implication (v) $\Rightarrow$ (vi) follows. Suppose that condition (vi) holds. Let $\phi \in \mathcal{C}(X, \psi)$ and $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$. By the assumption we have $E^X(\Gamma)P_{X,\psi} = E^Y(\Gamma)P_{Y,\psi}$, so that we have $E^X(\Gamma)\phi = E^Y(\Gamma)\phi$, and hence

$$\langle E^X(\Delta)\phi, E^Y(\Gamma)\phi\rangle = \langle E^X(\Delta)\phi, E^X(\Gamma)\phi\rangle = \langle \phi, E^X(\Delta \cap \Gamma)\phi\rangle.$$

It follows that X and Y are perfectly correlated in ϕ, and hence the implication (vi) $\Rightarrow$ (ii) follows. Since the implication (ii) $\Rightarrow$ (i) is obvious, the proof is completed. □

It should be noticed that condition (vi) above does not imply the relation $X\psi = Y\psi$, since ψ may not be in the domain of X or Y. However, for any rapidly decreasing f, i.e., $f \in \mathcal{S}(\mathbf{R})$, we have $f(X)\psi$ is in the domains of X and Y, and that the self-adjoint extension of $XP_{X,\psi} - YP_{Y,\psi}$ coincides with the zero operator.

For bounded X and Y, condition (vi) above is equivalent to that $X\phi = Y\phi$ for all $\phi \in \mathcal{C}_1(X, \psi)$, and the later condition means that the observable $X - Y$ has the definite value zero in state ϕ, so that it is an interesting question to ask whether the relation

$$X\psi = Y\psi \tag{5}$$

ensures that X and Y are perfectly correlated in ψ. If bounded observables X and Y commute, by multiplying $f(X)$ to the both sides we have $Xf(X)\psi = Yf(X)\psi$ for all $f \in B(\mathbf{R})$ so that we have $XP_{X,\psi} = YP_{X,\psi}$, and hence X and Y are perfectly correlated in ψ. Busch et al. [9] pointed out that Eq. (5) does not ensure that X and Y are identically distributed in ψ. Here, we shall show that even unitarily equivalent X and Y satisfying Eq. (5) may fail to be perfectly correlated. Let X, Y, and ψ be two 4×4 matrices and a four-dimensional column vector such that

$$X = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \psi = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Then, it is easy to see that X and Y are unitarily equivalent and satisfy Eq. (5). However, we have $\langle\psi|X^3|\psi\rangle = 4$ but $\langle\psi|Y^3|\psi\rangle = 3$. Thus, the third moments of X and Y are different, so that the observables X and Y have different probability distributions in ψ, and hence from Proposition 2.1 they cannot be perfectly correlated.

Let X be an observable on $\mathcal{H}$ and ρ a state on $\mathcal{H}$. The cyclic subspace of $\mathcal{H}$ spanned by observable X and state ρ is the closed subspace $\mathcal{C}(X, \rho)$ defined by

$$\mathcal{C}(X, \rho) = \text{the closure of } \{f(X)\psi \in \mathcal{H} | f \in B(\mathbf{R}),\ \psi \in \mathrm{ran}(\rho)\}. \tag{6}$$

Then, it is easy to see the following relation

$$\mathcal{C}(X, \rho) = \text{the closure of } \bigcup_{\psi \in \mathrm{ran}(\rho)} \mathcal{C}(X, \psi). \tag{7}$$

In particular, we have $\mathcal{C}(X, |\psi\rangle\langle\psi|) = \mathcal{C}(X, \psi)$ for any state vector $\psi \in \mathcal{H}$. Denote by $\mathcal{C}_1(X, \rho)$ the unit sphere of $\mathcal{C}(X, \rho)$ and by $P_{X,\rho}$ the projection of $\mathcal{H}$ onto $\mathcal{C}(X, \rho)$. Since $P_{X,\rho} = \bigvee_{\psi \in \mathrm{ran}(\rho)} P_{X,\psi}$, we have $[P_{X,\rho}, E^X(\varDelta)] = 0$ for all $\varDelta \in \mathcal{B}(\mathbf{R})$. Denote by $\mathcal{S}(X, \rho)$ the space of states supported in $\mathcal{C}(X, \rho)$, i.e.,

$$\mathcal{S}(X, \rho) = \{\sigma \in \mathcal{S}(\mathcal{H}) |\ \mathrm{ran}(\sigma) \subseteq \mathcal{C}(X, \rho)\}. \tag{8}$$

It is easy to see that the following conditions are equivalent: (i) $\sigma \in \mathcal{S}(X, \rho)$. (ii) $P_{X,\rho}\sigma = \sigma$. (iii) $\sigma P_{X,\rho} = \sigma$. (iv) $P_{X,\rho}\sigma P_{X,\rho} = \sigma$.

Then, we obtain the following characterization of perfect correlation in a mixed state.

Theorem 3.2. *For any two observables X and Y on $\mathcal{H}$ and any state ρ on $\mathcal{H}$, the following conditions are equivalent.*

(i) *X and Y are perfectly correlated in ρ.*
(ii) *X and Y are perfectly correlated in any $\psi \in \mathrm{ran}(\rho)$.*
(iii) *X and Y are perfectly correlated in any $\psi \in \mathcal{C}_1(X, \rho)$.*
(iv) *X and Y are perfectly correlated in any $\sigma \in \mathcal{S}(X, \rho)$.*
(v) *$E^X(\varDelta)\rho = E^Y(\varDelta)\rho$ for any $\varDelta \in \mathcal{B}(\mathbf{R})$.*

(vi) $f(X)P_{X,\rho} = f(Y)P_{X,\rho}$ *for any* $f \in B(\mathbf{R})$.
(vii) $P_{X,\rho} = P_{Y,\rho}$ *and* $XP_{X,\rho} = YP_{Y,\rho}$.

Proof. Suppose (i) holds. Let $\Delta \in \mathcal{B}(\mathbf{R})$. From Proposition 2.1, we have

$$\|E^X(\Delta)\sqrt{\rho} - E^Y(\Delta)\sqrt{\rho}\|_{HS}^2 \\ = \mathrm{Tr}[E^X(\Delta)\rho] - \mathrm{Tr}[E^X(\Delta)E^Y(\Delta)\rho] - \mathrm{Tr}[E^Y(\Delta)E^X(\Delta)\rho] + \mathrm{Tr}[E^Y(\Delta)\rho] = 0,$$

where $\|\cdot\|_{HS}$ stands for the Hilbert-Schmidt norm. It follows that we have $E^X(\Delta)\rho = E^Y(\Delta)\rho$, and hence the implication (i) $\Rightarrow$ (v) follows. The implications (v) $\Rightarrow$ (ii), (ii) $\Rightarrow$ (iii), and (iii) $\Rightarrow$(vi) follow easily from the implications (iii) $\Rightarrow$ (i), (i) $\Rightarrow$ (ii), and (ii) $\Rightarrow$ (iv) in Theorem 3.1, respectively. Assume condition (vi). It follows immediately that $f(X)\psi = f(Y)\psi$ for all $\psi \in \mathrm{ran}(\rho)$, so that from Eq. (6) we have $\mathcal{C}(X,\rho) = \mathcal{C}(Y,\rho)$ and $P_{X,\rho} = P_{Y,\rho}$. Then, by assumption we have $E^X(\Delta)P_{X,\rho} = E^Y(\Delta)P_{Y,\rho}$ for all $\Delta \in \mathcal{B}(\mathbf{R})$, and hence we conclude that $XP_{X,\rho}$ equals $YP_{Y,\rho}$ since their spectral measures coincides. Thus, the implication (vi) $\Rightarrow$ (vii) follows. Assume condition (vii). Suppose $\sigma \in \mathcal{S}(X,\rho)$. Then $P_{X,\rho}\sigma = \sigma$, so that $E^X(\Delta)\sigma = E^Y(\Delta)\sigma$ and it is easy to see that X and Y are perfectly correlated in σ, and the implication (vii) $\Rightarrow$ (iv) follows. The implication (iv) $\Rightarrow$ (i) trivially holds and the proof is completed. □

For observables with a complete orthonormal family of eigenvectors (discrete observables), we have the following important characterization of perfectly correlating states.

Theorem 3.3. *Two discrete observables X and Y are perfectly correlated in a vector state* ψ *if and only if* ψ *is a superposition of common eigenstates of X and Y with common eigenvalues.*

Proof. Suppose that X and Y are perfectly correlated in a state ψ. Then, $\mathcal{C}(X,\psi)$ is generated by eigenstates of $XP_{X,\psi} = YP_{X,\psi}$. Thus, ψ is a superposition of common eigenstates of X and Y with common eigenvalues. Conversely, suppose that ψ is a superposition of common eigenstates of X and Y with common eigenvalues. Then, the subspace $\mathcal{S}$ generated by those eigenstates is invariant under both X and Y and includes ψ. Thus, $\mathcal{C}(X,\psi) \subseteq \mathcal{S}$, and $X = Y$ on $\mathcal{C}(X,\psi)$, and hence from Theorem 3.1, we conclude X and Y are perfectly correlated in ψ. □

We say that two observables X and Y are *identically distributed* in a state ρ iff $\mathrm{Tr}[E^X(\Delta)\rho] = \mathrm{Tr}[E^Y(\Delta)\rho]$ for all $\Delta \in \mathcal{B}(\mathbf{R})$. Then, we have the following.

Theorem 3.4. *For any two observables X and Y on* $\mathcal{H}$ *and any state* $\rho \in \mathcal{S}(\mathcal{H})$, *the following conditions are equivalent.*

(i) *X and Y are perfectly correlated in state* ρ.
(ii) *X and Y are identically distributed in any* $\psi \in \mathcal{C}_1(X,\rho)$.
(iii) *X and Y are identically distributed in any state* $\rho \in \mathcal{S}(X,\rho)$.

Proof. Suppose (i) holds. Let $\Delta \in \mathcal{B}(\mathbf{R})$. From Theorem 3.2, we have $E^X(\Delta)\sigma = E^Y(\Delta)\sigma$ for any $\sigma \in \mathcal{S}(X,\rho)$. Thus, (iii) holds, and (i) $\Rightarrow$ (iii) follows. The implication (iii) $\Rightarrow$ (ii) is obvious. Suppose that (ii) holds. Let $\psi \in \mathcal{C}_1(X,\rho)$. Let Δ, Γ be disjoint Borel sets in $\mathcal{B}(\mathbf{R})$. Then, $E^X(\Delta)\psi \in \mathcal{C}(X,\rho)$, and hence

$$\langle E^X(\Delta)\psi, E^Y(\Gamma)E^X(\Delta)\psi\rangle = \langle E^X(\Delta)\psi, E^X(\Gamma)E^X(\Delta)\psi\rangle = 0.$$

Thus, by the Schwarz inequality we have

$$|\langle E^X(\Delta)\psi, E^Y(\Gamma)\psi\rangle|^2 \leqslant \|E^Y(\Gamma)E^X(\Delta)\psi\|^2 = \langle E^X(\Delta)\psi, E^Y(\Gamma)E^X(\Delta)\psi\rangle = 0.$$

It follows that (i) holds. Thus, the proof is completed. □

It should also be noticed that even two identically distributed commuting observables X and Y may fail to satisfy Eq. (5). To see this, suppose that $\mathcal{H} = \mathcal{K} \otimes \mathcal{K}$ for some Hilbert space $\mathcal{K}$. Let $X = A \otimes I$ and $Y = I \otimes A$ for some bounded operator A on $\mathcal{K}$ and $\psi = \phi \otimes \phi$ for some state vector $\phi \in \mathcal{K}$. Then, we have $\langle\psi|E^X(\Delta)|\psi\rangle = \langle\phi|E^A(\Delta)|\phi\rangle = \langle\psi|E^Y(\Delta)|\psi\rangle$ for all $\Delta \in \mathcal{B}(\mathbf{R})$, and hence they are identically distributed. However, we have $(X - Y)\psi = A\phi \otimes \phi - \phi \otimes A\phi$, and hence Eq. (5) does not hold unless ϕ is an eigenvector of A.

4. Transitivity of perfect correlations

We denote by $\{X = Y\}$ the subspace spanned by all states $\psi \in \mathcal{H}$ such that X and Y are perfectly correlated in ψ, i.e.,

$$\{X = Y\} = \{\psi \in \mathcal{H} | \langle E^X(\Delta)\psi, E^Y(\Gamma)\psi\rangle = 0 \text{ for all disjoint Borel sets } \Delta, \Gamma\}.$$

We shall call $\{X = Y\}$ the *perfectly correlative domain* for X and Y. Then, we have

Theorem 4.1. *The space $\{X = Y\}$ is the largest closed subspace $\mathcal{K}$ of $\mathcal{H}$ satisfying the following conditions.*

(i) *$\mathcal{K}$ is invariant under X and Y for all $\Delta \in \mathcal{B}(\mathbf{R})$.*
(ii) *$E^X(\Delta)\psi = E^Y(\Delta)\psi$ for all $\Delta \in \mathcal{B}(\mathbf{R})$ and $\psi \in \mathcal{K}$.*

Proof. Assume $\psi \in \{X = Y\}$. Then, we have $E^X(\Delta)E^X(\Gamma)\psi = E^X(\Delta \cap \Gamma)\psi = E^Y(\Delta \cap \Gamma)\psi = E^Y(\Delta)E^Y(\Gamma)\psi = E^Y(\Delta)E^X(\Gamma)\psi$. Thus, $\{X = Y\}$ is invariant under X, and similarly under Y. The space $\{X = Y\}$ satisfies condition (ii) obviously from Theorem 3.1. Assume that $\mathcal{K}$ satisfies conditions (i) and (ii). Let $\psi \in \mathcal{K}$. Then, from (ii) we have $\psi \in \{X = Y\}$, and hence $\{X = Y\}$ is the largest. □

From the above theorem, $\psi \in \{X = Y\}$ if and only if $\mathcal{C}(X, \psi) \subseteq \{X = Y\}$. The following theorem shows that the perfect correlation in a given state is an equivalence relation between observables.

Theorem 4.2. *For any observables X, Y, Z, we have $\{X = X\} = \mathcal{H}$, $\{X = Y\} = \{Y = X\}$, and $\{X = Y\} \cap \{Y = Z\} \subseteq \{X = Z\}$.*

Proof. The relations $\{X = X\} = \mathcal{H}$ and $\{X = Y\} = \{Y = X\}$ are obvious. Let $\psi \in \{X = Y\} \cap \{Y = Z\}$ and $f \in B(\mathbf{R})$. Then, we have $f(X)\psi = f(Y)\psi$ and $f(Y)\psi = f(Z)\psi$, so that $f(X)\psi = f(Z)\psi$. Since f is arbitrary, we have $\psi \in \{X = Z\}$. Thus, we conclude $\{X = Y\} \cap \{Y = Z\} \subseteq \{X = Z\}$. □

We denote by $[[X = Y]]$ the projection of $\mathcal{H}$ onto $\{X = Y\}$. From Theorem 4.1 we have

$$\left[E^X(\Delta) - E^Y(\Delta)\right][[X = Y]] = 0 \tag{9}$$

for all $\Delta \in \mathcal{B}(\mathbf{R})$.

Theorem 4.3. *For any two observables X and Y on $\mathcal{H}$ and any state ρ on $\mathcal{H}$, the following conditions are equivalent.*

(i) *X and Y are perfectly correlated in ρ.*
(ii) *ran $(\rho) \subseteq \{X = Y\}$.*
(iii) $[[X = Y]]\rho = \rho$.
(iv) $\rho[[X = Y]] = \rho$.

Proof. Suppose that X and Y are perfectly correlated in a state ρ. Let $\psi \in \mathrm{ran}(\rho) \setminus \{0\}$. Then, $\psi = \sqrt{\rho}\phi$ for some vector $\phi \in \mathcal{H}$. For any disjoint $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$ we have

$$\begin{aligned}\langle E^X(\Delta)\psi, E^Y(\Gamma)\psi\rangle &= \|\phi\|^2\langle\phi/\|\phi\|, \sqrt{\rho}E^X(\Delta)E^Y(\Gamma)\sqrt{\rho}\phi/\|\phi\|\rangle \\ &\leqslant \|\phi\|^2\mathrm{Tr}[\sqrt{\rho}E^X(\Delta)E^Y(\Gamma)\sqrt{\rho}] = 0.\end{aligned}$$

Thus, $\psi \in \{X = Y\}$, so that $\mathrm{ran}(\rho) \subseteq \{X = Y\}$, and the implication (i) $\Rightarrow$ (ii) follows. The implication (ii) $\Rightarrow$ (iii) is obvious. Suppose $[[X = Y]]\rho = \rho$. From Eq. (9), we have $E^X(\Delta)\rho = E^Y(\Delta)\rho$ and hence X and Y are perfectly correlated in ρ, and the implication (iii) $\Rightarrow$ (i) follows. The equivalence (iii) $\iff$ (iv) follows immediately from taking the adjoint of the both sides of relation (iii) or (iv). □

For two observables X, Y, and a state ρ, we denote by $X \equiv_\rho Y$ iff X and Y are perfectly correlated in ρ. The following theorem shows that the relation $\equiv_\rho$ is an equivalence relation between observables and in particular it is transitive.

Theorem 4.4. *For any observables X, Y, Z, and state ρ, we have (i) $X \equiv_\rho X$, (ii) if $X \equiv_\rho Y$ then $Y \equiv_\rho X$, and (iii) if $X \equiv_\rho Y$ and $Y \equiv_\rho Z$ then $X \equiv_\rho Z$.*

Proof. From Theorems 4.2 and 4.3, statements (i) and (ii) follow easily. Suppose $X \equiv_\rho Y$ and $Y \equiv_\rho Z$. Then, from Theorem 4.3 we have $\mathrm{ran}(\rho) \subseteq \{X = Y\}$ and $\mathrm{ran}(\rho) \subseteq \{Y = Z\}$, and hence $\mathrm{ran}(\rho) \subseteq \{X = Y\} \cap \{Y = Z\}$. From Theorem 4.2, we have $\mathrm{ran}(\rho) \subseteq \{X = Z\}$, so that we have shown $X \equiv_\rho Z$, and statement (iii) follows. □

5. Joint distributions

5.1. Perfect correlations and joint probability distributions

Let X and Y be two observables on $\mathcal{H}$. We say that X and Y *commute* on a closed subspace $\mathcal{K} \subseteq \mathcal{H}$ iff $\mathcal{K}$ is invariant under X and Y and $[E^X(\Delta), E^Y(\Gamma)]\psi = 0$ for all $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$ and $\psi \in \mathcal{K}$. The *commutative domain* of X and Y is defined to be the set $\mathrm{com}(X, Y)$ of those vectors $\psi \in \mathcal{H}$ such that $[E^X(\Delta), E^Y(\Gamma)]\psi = 0$ for all $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$. It is clear that if X and Y commute on $\mathcal{K}$ then $\mathcal{K} \subseteq \mathrm{com}(X, Y)$. It can be easily seen that $\mathrm{com}(X, Y)$ is invariant under X and Y; in fact, if $\psi \in \mathrm{com}(X, Y)$, we have $E^X(\Delta_1)E^Y(\Delta_2)E^X(\Delta_3)\psi = E^X(\Delta_1)E^X(\Delta_3)E^Y(\Delta_2)\psi = E^X(\Delta_1 \cap \Delta_3)E^Y(\Delta_2)\psi = E^Y(\Delta_2)E^X(\Delta_1 \cap \Delta_3)\psi = E^Y(\Delta_2)E^X(\Delta_1)E^X(\Delta_3)\psi$, so that $E^X(\Delta_3)\psi \in \mathrm{com}(X, Y)$. Thus, $\mathrm{com}(X, Y)$ is the largest closed subspace on which X and Y commute; see Ylinen [10]. Let $C_{X,Y}$ denote the projection of $\mathcal{H}$ onto $\mathrm{com}(X, Y)$. Then, we have

$$[E^X(\varDelta), E^Y(\varGamma)]C_{X,Y} = 0 \tag{10}$$

for all $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$. The following theorem generalizes Yilnen's theorem [10] on characterization of pure states in $\mathrm{com}(X, Y)$ to mixed states.

Theorem 5.1. *For any state ρ, the following conditions are equivalent.*

(i) $C_{X,Y}\rho = \rho$.
(ii) *There is a spectral measure E on $\mathcal{B}(\mathbf{R}^2)$ such that $E(\varDelta \times \varGamma)\rho = E^X(\varDelta) \wedge E^Y(\varGamma)\rho$ for all $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$.*
(iii) *The function $\varDelta \times \varGamma \mapsto \mathrm{Tr}[E^X(\varDelta) \wedge E^Y(\varGamma)\rho]$ on $\mathcal{B}(\mathbf{R}) \times \mathcal{B}(\mathbf{R})$ extends to a probability measure on $\mathcal{B}(\mathbf{R}^2)$.*
(iv) *$E^X(\varDelta)E^Y(\varGamma)\rho = E^Y(\varGamma)E^X(\varDelta)\rho$ for all $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$.*

Proof. Since $XC_{X,Y}$ and $YC_{X,Y}$ are commuting self-adjoint operators, there is another self-adjoint operator Z and two real-valued Borel functions f, g such that $XC_{X,Y} = f(Z)$ and $YC_{X,Y} = g(Z)$ [2]. Let E be the spectral measure on $\mathcal{B}(\mathbf{R}^2)$ defined by $E(\varDelta \times \varGamma) = E^Z(f^{-1}(\varDelta) \cap g^{-1}(\varGamma))$ for all $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$. Let $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$. We have $E(\varDelta \times \varGamma)C_{X,Y} = E^Z(f^{-1}(\varDelta))E^Z(g^{-1}(\varGamma))C_{X,Y} = E^X(\varDelta)E^Y(\varGamma)C_{X,Y} = E^X(\varDelta) \wedge E^Y(\varGamma)C_{X,Y}$. Thus, it is easy to see that the implication (i) $\Rightarrow$ (ii) follows. The implication (ii) $\Rightarrow$ (iii) follows obviously. Assume condition (iii). Let μ be the probability measure on $\mathcal{B}(\mathbf{R}^2)$ such that $\mu(\varDelta \times \varGamma) = \mathrm{Tr}[E^X(\varDelta) \wedge E^Y(\varGamma)\rho]$. Let $P = E^Y(\varGamma) - E^X(\varDelta) \wedge E^Y(\varGamma) - E^X(\mathbf{R} \setminus \varDelta) \wedge E^Y(\varGamma)$. Then, P is a projection and $E^X(\varDelta)P = E^X(\varDelta)E^Y(\varGamma) - E^X(\varDelta) \wedge E^Y(\varGamma)$. By the countable additivity of μ, we have $\mathrm{Tr}[(P\sqrt{\rho})^\dagger(P\sqrt{\rho})] = \mathrm{Tr}[P\rho] = \mu(\mathbf{R} \times \varGamma) - \mu(\varDelta \times \varGamma) - \mu((\mathbf{R} \setminus \varDelta) \times \varGamma) = 0$. Thus, we have $P\sqrt{\rho} = 0$ so that $E^X(\varDelta)P\rho = 0$, and hence we have $E^X(\varDelta)E^Y(\varGamma)\rho = E^X(\varDelta) \wedge E^Y(\varGamma)\rho$. By symmetry, we also obtain $E^Y(\varGamma)E^X(\varDelta)\rho = E^X(\varDelta) \wedge E^Y(\varGamma)\rho$. Thus, the implication (iii) $\Rightarrow$ (iv) follows. Assume condition (iv). Then, we have $\rho\psi \in \mathrm{com}(X, Y)$ for all $\psi \in \mathcal{H}$. Thus, $C_{X,Y}\rho\psi = \rho\psi$ for all $\psi \in \mathcal{H}$, and hence the implication (iv) $\Rightarrow$ (i) follows. □

Observables X and Y are said to be *compatible* in a state ρ iff $C_{X,Y}\rho = \rho$, and they are said to have the *joint probability distribution* in ρ iff there is a probability measure $\mu_\rho^{X,Y}$ on $\mathcal{B}(\mathbf{R}^2)$ satisfying

$$\mu_\rho^{X,Y}(\varDelta \times \varGamma) = \mathrm{Tr}[E^X(\varDelta) \wedge E^Y(\varGamma)\rho] = \mathrm{Tr}[E^Y(\varGamma)E^X(\varDelta)\rho] = \mathrm{Tr}[E^X(\varDelta)E^Y(\varGamma)\rho] \tag{11}$$

for all $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$. Theorem 5.1 shows that X and Y have the joint probability distribution in ρ if and only if they are compatible in ρ.

Two observables X and Y are called *jointly measurable* in a state ρ iff they have the joint probability distribution $\mu_\rho^{X,Y}$ and satisfy the following relations

$$\mu_\rho^{X,Y}(\varDelta \times \varGamma) = \mathrm{Tr}[E^X(\varDelta)E^Y(\varGamma)E^X(\varDelta)\rho], \tag{12}$$

$$\mu_\rho^{X,Y}(\varDelta \times \varGamma) = \mathrm{Tr}[E^Y(\varGamma)E^X(\varDelta)E^Y(\varGamma)\rho] \tag{13}$$

for any $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$. The above relations ensure that the theoretical joint probability of the event "$X \in \varDelta$ and $Y \in \varGamma$" is obtained as the joint probability of outcomes of the successive projective measurements of projections $E^X(\varDelta)$ and $E^Y(\varGamma)$ irrespective of the order of the measurements [11]. Moreover, for discrete observables X and Y, the above relation ensures that the the joint probability distribution of the outcomes of the successive projective measurements of observables X and Y coincides with the joint probability distribution $\mu_\rho^{X,Y}$ irrespective of the order of the measurements.

Theorem 5.2. *Every pair of observables X and Y compatible in a state ρ is jointly measurable in the state ρ.*

Proof. The assertion follows immediately from Theorem 5.1. □

Denote by $\mathbf{D}$ the diagonal set in $\mathbf{R}^2$, i.e., $\mathbf{D} = \{(x, y) \in \mathbf{R}^2 \mid x = y\}$.

Theorem 5.3. *Two observables X and Y are perfectly correlated in a sate ρ if and only if X and Y are compatible in ρ and the joint probability distribution is concentrated in the diagonal set, i.e., $\mu_\rho^{X,Y}(\mathbf{R} \setminus \mathbf{D}) = 0$.*

Proof. If $x \neq y$, there is a rational number q such that $x < q < y$ or $y < q < x$, and hence it is easy to see that

$$\mathbf{R} \setminus \mathbf{D} = \bigcup_{q \in \mathbf{Q}} (-\infty, q) \times (q, \infty) \cup \bigcup_{q \in \mathbf{Q}} (q, \infty) \times (-\infty, q), \tag{14}$$

where $\mathbf{Q}$ stands for the set of rational numbers. Suppose that X and Y are perfectly correlated in ρ. Then, we have $E^X(\varDelta)E^Y(\varGamma)\rho = E^X(\varDelta \cap \varGamma)\rho = E^Y(\varDelta \cap \varGamma)\rho E^Y(\varGamma)E^Y(\varDelta) = E^Y(\varDelta)E^X(\varDelta)\rho$, and hence X and Y are compatible in ρ. Accordingly, the joint probability distribution satisfies $\mu_\rho^{X,Y}((-\infty, q) \times (q, \infty)) = \mu_\rho^{X,Y}((q, \infty) \times (-\infty, q)) = 0$, so that $\mu_\rho^{X,Y}(\mathbf{R} \setminus \mathbf{D}) = 0$. Conversely, suppose that X and Y are compatible in ρ and $\mu_\rho^{X,Y}(\mathbf{R} \setminus \mathbf{D}) = 0$. Let $\varDelta, \varGamma \in \mathcal{B}(\mathbf{R})$. In general, we have $(\varDelta \times \varGamma) \cap \mathbf{D} = [\mathbf{R} \times (\varDelta \cap \varGamma)] \cap \mathbf{D}$. Thus, if $\varDelta \cap \varGamma = \emptyset$, we have

$$\mathrm{Tr}[E^X(\varDelta)E^Y(\varGamma)\rho] = \mu_\rho^{X,Y}((\varDelta \times \varGamma) \cap \mathbf{D}) = \mu_\rho^{X,Y}([\mathbf{R} \times (\varDelta \cap \varGamma)] \cap \mathbf{D}) = 0, \tag{15}$$

so that X and Y are perfectly correlated in ρ. □

Let $\varepsilon > 0$. Let $\cdots < \mu_{-1} < \mu_0 < \mu_1 < \cdots$ be a partition of the real line $\mathbf{R}$ such that $\mu_{j+1} - \mu_j < \varepsilon$ for all j. Let X_ε and Y_ε be ε approximations of observables X and Y defined by $X_\varepsilon = \sum_{j \in \mathbf{Z}} \lambda_j E^X(\varDelta_j)$ and $Y_\varepsilon = \sum_{j \in \mathbf{Z}} \lambda_j E^Y(\varDelta_j)$, where $\varDelta_j = [\mu_j, \mu_{j+1})$ and $\lambda_j \in \varDelta_j$. If X and Y are discrete observables, there are ε approximations X_ε and Y_ε such that $X = X_\varepsilon$ and $Y = Y_\varepsilon$. From Theorems 5.2 and 5.3 we conclude that two observables X and Y perfectly correlated in a state ρ have the joint probability distribution concentrated in the diagonal set and that each instance of the successive projective measurements of any ε approximations X_ε and Y_ε gives the same output irrespective of the order of the measurements for any $\varepsilon > 0$.

5.2. Perfect correlations and quasi-probability distributions

In [12], Urbanik introduced the following formulation for the quasi-joint probability distribution for any pair of observables, generalizing the quasi-joint probability distribution of the position and the momentum first studied by Wigner [13] and Moyal [14]. Let μ be a probability measure on $\mathcal{B}(\mathbf{R}^2)$. To any pair of real numbers x, y there corresponds the family of lines $S_t^{a,b}$ given by the equation $ax + by = t$, where $t \in \mathbf{R}$. Letting for every Borel subset $\varDelta \subseteq \mathbf{R}$

$$\mu_{a,b}(\Delta) = \mu\left(\bigcup_{t\in\Delta} S_t^{a,b}\right) \tag{16}$$

we obtain a probability measure on $\mathbf{R}$. It is well-know that μ is determined uniquely by the family of probability measures $\mu_{a,b}$. We suppose that for all pair of real numbers a, b the linear combinations $aX + bY$ are self-adjoint operators on $\mathcal{H}$. Consequently, for every pair $a, b \in \mathbf{R}$ and every state vector ψ the probability distribution of $aX + bY$ is defined by

$$\mu_{\psi}^{aX+bY}(\Delta) = \langle \psi, E^{aX+bY}(\Delta)\psi\rangle \tag{17}$$

for all $\Delta \in \mathcal{B}(\mathbf{R})$. Given a state vector ψ, a probability measure μ on $\mathbf{R}^2$ is said to be *the joint probability distribution* of observables X and Y iff $\mu_{a,b}$ is equal to μ_{ψ}^{aX+bY}. The joint probability distribution so defined is uniquely determined, provided it exists. We shall denote by $\nu_{\psi}^{X,Y}$ the joint probability distribution of X, Y in ψ. We also denote by $\Phi_{a,b}^{\psi}$ the characteristic function of the probability distribution μ_{ψ}^{aX+bY}:

$$\Phi_{a,b}^{\psi}(t) = \langle \psi, e^{it(aX+bY)}\psi\rangle$$

for all $t \in \mathbf{R}$. Then, from Bochner's theorem it is easy to see that observables X and Y have the joint probability distribution in a vector state $\psi \in \mathcal{H}$ if and only if the function $\Phi_{t,s}^{\psi}(1)$ of two variable t, s is a continuous positive definite function on $\mathbf{R}^2$.

Theorem 5.4. *For any observables X, Y and any state vector ψ, the following conditions are equivalent.*

(i) *X and Y are perfectly correlated in ψ.*
(ii) *$\Phi_{t,s}^{\phi}(1) = \Phi_{t+s,0}^{\phi}(1)$ for any $t, s \in \mathbf{R}$ and $\phi \in \mathcal{C}_1(X,\psi)$*
(iii) *$\Phi_{t,0}^{\phi}(1) = \Phi_{0,t}^{\phi}(1)$ for any $t \in \mathbf{R}$ and $\phi \in \mathcal{C}_1(X,\psi)$.*

Proof. Assume (i) holds. Then, we have $XP_{X,\psi} = YP_{X,\psi}$, so that $e^{i(tX+sY)}P_{X,\psi} = e^{itX} e^{isY}P_{X,\psi} = e^{itX}e^{isX}P_{X,\psi} = e^{i(t+s)X}P_{X,\psi}$. Let $\phi \in \mathcal{C}_1(X,\psi)$. Then, we have $\Phi_{t,s}^{\phi}(1) = \langle \phi, e^{i(tX+sY)}P_{X,\psi}\phi\rangle = \langle \phi, e^{i(t+s)X}\phi\rangle$, and hence the implication (i) $\Rightarrow$ (ii) follows. The implication (ii) $\Rightarrow$ (iii) is obvious. Assume (iii) holds. Then, we have

$$\langle \phi, e^{itX}\phi\rangle = \langle \phi, e^{itY}\phi\rangle$$

for all $t \in \mathbf{R}$. It follows that X and Y are identically distributed in ϕ. Since $\phi \in \mathcal{C}_1(X,\psi)$ is arbitrary, the implication (iii) $\Rightarrow$ (i) follows from Theorem 3.4. □

Our approach is more coherent with the following definition of "characteristic functions." We define a function $\Psi_{a,b}^{\psi}$ on $\mathbf{R}$ by

$$\Psi_{a,b}^{\psi}(t) = \langle e^{-itaX}\psi, e^{itbY}\psi\rangle$$

for all $t \in \mathbf{R}$. It is easy to see that $\Psi_{a,0}^{\psi}(t) = \Phi_{a,0}^{\psi}(t)$ and $\Psi_{0,a}^{\psi}(t) = \Phi_{0,a}^{\psi}(t)$ for all $a,b,t \in \mathbf{R}$. Then, we have

Theorem 5.5. *For any observables X, Y and any state vector ψ, the following conditions are equivalent.*

(i) *X and Y are perfectly correlated in ψ.*
(ii) *$\Psi_{t,s}^{\psi}(1) = \Psi_{t+s,0}^{\psi}(1)$ for any $t, s \in \mathbf{R}$.*

Proof. Suppose (i) holds. From Theorem 3.1 we have $e^{isY}\psi = e^{isX}\psi$, and hence

$$\Psi^{\psi}_{t,s}(1) = \langle e^{-itX}\psi, e^{isY}\psi\rangle = \langle e^{-itX}\psi, e^{isX}\psi\rangle = \Phi^{\psi}_{t+s,0}(1)$$

for any $t, s \in \mathbf{R}$. Suppose (ii) holds. We have

$$\|e^{itX}\psi - e^{itY}\psi\|^2 = 2 - 2\mathrm{Re}\langle e^{itX}\psi, e^{itY}\psi\rangle = 2 - 2\mathrm{Re}\Psi_{-t,t}(1) = 0.$$

Thus, we have $e^{itX}\psi = e^{itY}\psi$ for any $t \in \mathbf{R}$. Since the von Neumann algebra generated by all e^{itX} with $t \in \mathbf{R}$ coincides with that of all $f(X)$ with $f \in B(\mathbf{R})$, the set of Borel functions f satisfying $f(X)\psi = f(Y)\psi$ includes $B(\mathbf{R})$, and the implication (ii) $\Rightarrow$ (i) follows. $\square$

6. Perfect correlations and entanglement

6.1. Bipartite perfect correlations

The notion of perfect correlation in quantum theory was discussed first by von Neumann [2] to establish a quantum mechanical description of a process of measurement and is closely related to the notion of entanglement recently discussed quite actively in the field of quantum information [3]. In what follows we shall discuss some examples in these fields.

Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be two Hilbert spaces and suppose $\mathcal{H} = \mathcal{K}_1 \otimes \mathcal{K}_2$. Every state vector ψ has two orthonormal sequences $\{\phi_j\}$ and $\{\xi_j\}$ such that

$$\psi = \sum_j \sqrt{p_j}\phi_j \otimes \xi_j, \tag{18}$$

where $p_j > 0$ and $\sum_j p_j = 1$ [2]. The above decomposition is called the *Schmidt decomposition* of ψ. Then, the amount of entanglement [3] of ψ is defined by

$$E(\psi) = -\sum_j p_j \log p_j. \tag{19}$$

Let $\rho_1 = \mathrm{Tr}_2|\psi\rangle\langle\psi|$ and $\rho_2 = \mathrm{Tr}_1|\psi\rangle\langle\psi|$, where Tr_l stands for the partial trace over $\mathcal{K}_l$ for $l = 1, 2$. Then, $E(\psi) = S(\rho_1) = S(\rho_2)$, where S stands for the von Neumann entropy, i.e., $S(\rho_l) = -\mathrm{Tr}[\rho_l \log \rho_l]$ [2]. Let X and Y be observables on $\mathcal{H}$ defined by $X = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|$ and $Y = \sum_j \lambda_j |\xi_j\rangle\langle\xi_j|$ with nondegenerate eigenvalues $\{\lambda_j\}$. Then, we have $\langle E^{X\otimes I}(\{\lambda_j\})\psi, E^{I\otimes Y}(\{\lambda_k\})\psi\rangle = \sqrt{p_j p_k}\langle\phi_j \otimes \xi_j, \phi_k \otimes \xi_k\rangle = \delta_{j,k}p_j$, and hence we can conclude that $X \otimes I$ and $I \otimes Y$ are perfectly correlated in ψ.

Theorem 6.1. *Suppose $\mathcal{H} = \mathcal{K}_1 \otimes \mathcal{K}_2$ with $\dim(\mathcal{H}) < \infty$. Let ψ be a state on $\mathcal{H}$. For any two observables X on $\mathcal{K}_1$ and Y on $\mathcal{K}_2$, the observables $X \otimes I$ and $I \otimes Y$ are perfectly correlated in ψ if and only if there is a pair of orthonormal basis $\{\phi_j\}$ of $\mathcal{K}_1$ and $\{\xi_j\}$ of $\mathcal{K}_2$ and a sequence of nonzero real numbers $\lambda_1, \ldots, \lambda_n$ such that ψ has the Schmidt decomposition $\psi = \sum_{j=1}^n \sqrt{p_j}\phi_j \otimes \xi_j$ with $p_j > 0$ for all $j = 1, \ldots, n$, and that $X\phi_j = \lambda_j\phi_j$ and $Y\xi_j = \lambda_j\xi_j$ for all $j = 1, \ldots, n$.*

Proof. Suppose that $X\otimes I$ and $I\otimes Y$ are perfectly correlated in ψ.Then, by Theorem 3.3 the state ψ is a superposition of common eigenstates with common eigenvalues $\mu_1,\ldots,\mu_m$ of X and Y. Let $\psi_k=[E^X(\{\mu_k\})\otimes E^Y(\{\mu_k\})]\psi/\sqrt{q_k}$, where $\|\Psi_k\|^2=1$ for all $k=1,\ldots,m$.We have $\psi=\sum_{k=1}^m\sqrt{q_k}\psi_k$ with $q_k>0$ and $\sum_{k=1}^m q_k=1$.Let $\psi_k=\sum_{l=1}^{s(k)}\sqrt{r_l^{(k)}}\phi_l^{(k)}\otimes\xi_l^{(k)}$ be a Schmidt decomposition of ψ_k.Then, $(X\otimes I)\psi_k=\sum_{l=1}^{s(k)}\sqrt{r_l^{(k)}}X\phi_l^{(k)}\otimes\xi_l^{(k)}$ and $(X\otimes I)\psi_k=\sum_{l=1}^{s(k)}\sqrt{r_l^{(k)}}\mu_k\phi_l^{(k)}\otimes\xi_l^{(k)}$.Since $\xi_1^{(k)},\ldots,\xi_{s(k)}^{(k)}$ are linearly independent, we have $X\phi_l^{(k)}=\mu_k\phi_l^{(k)}$, and similarly we have $Y\xi_l^{(k)}=\mu_k\xi_l^{(k)}$.Let $\lambda_j=\mu_k$ if $\sum_{l=1}^{k-1}s(l)<j\leqslant\sum_{l=1}^{k}s(l)$, let $\phi_j=\phi_l^{(k)}$, $\xi_j=\xi_l^{(k)}$, $\sqrt{p_j}=\sqrt{q_k r_l^{(k)}}$ if $j=l+\sum_{l=1}^{k-1}s(l)$, and let $n=\sum_{k=1}^m s(k)$. Then, we have a Schmidt decomposition $\psi=\sum_{j=1}^n\sqrt{p_j}\phi_j\otimes\xi_j$ with the desired properties. The converse part is obvious from the discussion preceding the present theorem, and the proof is completed. □

6.2. Nonlocality without inequality

Let us consider the case where $\mathcal{H}=\mathcal{K}_1\otimes\mathcal{K}_2$ and $\mathcal{K}_j\cong\mathbf{C}^2$ for $j=1,2$. Let U, D be two observables on $\mathbf{C}^2$ having eigenvalues 1 and 0. Let $U_1=U\otimes I$, $D_1=D\otimes I$, $U_2=I\otimes U$, and $D_1=I\otimes D$. Hardy [4] showed that any state vector $\psi\in\mathcal{H}$ shows nonlocality if it satisfies

$$P_\psi(U_1=0,U_2=1)=0, \tag{20}$$
$$P_\psi(U_1=1,D_2=0)=0, \tag{21}$$
$$P_\psi(D_1=1,U_2=0)=0, \tag{22}$$
$$P_\psi(D_1=1,D_2=0)>0, \tag{23}$$

where $P_\psi(A=a,\ B=b)=\langle E^A(\{a\})\psi,E^B(\{b\})\psi\rangle$ for $A=U_1,D_1$, and $B=U_2,D_2$, and $a,b=0,1$, and showed that actually we can find such observables U and D for any state ψ unless ψ is a product state *or a maximally entangled state*. This failure of Hardy's nonlocality proof for the class of maximally entangled states has been explained by Cereceda [15] as follows: the perfect correlation for pairs (U_1, U_2), (U_1, D_2), and (D_1, U_2) necessarily entails perfect correlation for the pair (D_1, D_2). Now, we shall show that Cereceda's argument can be considerably simplified by appealing to the general property of the transitivity of perfect correlations.

Let ψ be a general state vector in $\mathcal{H}$. Then, we have a Schmidt decomposition of ψ such that

$$\psi=\sqrt{p_1}\xi_1\otimes\eta_1+\sqrt{p_2}\xi_2\otimes\eta_2. \tag{24}$$

The numbers $0\leqslant p_2\leqslant p_1$ are uniquely determined with $p_1+p_2=1$, and if $p_1\neq 1/2,1$, the vectors $\xi_1\otimes\eta_1$ and $\xi_2\otimes\eta_2$ are uniquely determined up to constant factors. The essential part of Hardy's proof of nonlocality is that if $1/2<p_1<1$, we can always find observables U and D such that $U\neq D$ while they satisfy Eqs. (20)–(23). Now, suppose that ψ is maximally entangled, i.e., $p_1=1/2$. We shall show that Eq. (20) leads to $P_\psi(U_1=1,\ U_2=0)=0$. Let $\{\xi_0,\xi_1\}$ be an orthonormal basis such that $U=|\xi_1\rangle\langle\xi_1|$. Expanding ψ in the basis $\{\xi_j\otimes\xi_k\}_{j,k=0,1}$, we have $\psi=\sum_{j,k}c_{jk}\xi_j\otimes\xi_k$. Then, we have

$P_\psi(U_1 = j, U_2 = k) = |c_{jk}|^2$ for all $j, k = 0, 1$. From Eq. (20), we have $c_{01} = 0$, and hence we have

$$\begin{aligned}\rho_1 =& \mathrm{Tr}_2|\psi\rangle\langle\psi| \\ =& |c_{00}|^2|\xi_0\rangle\langle\xi_0| + c_{00}c_{10}^*|\xi_0\rangle\langle\xi_1| + c_{10}c_{00}^*|\xi_1\rangle\langle\xi_0| + (|c_{10}|^2 + |c_{11}|^2)|\xi_1\rangle\langle\xi_1|.\end{aligned}$$

Since $S(\rho_1) = \log 2$, we have $\rho_1 = I_1/2$, and hence $|c_{00}|^2 = 1/2$ and $c_{10}c_{00}^* = 0$, so that we have $P_\psi(U_1 = 1, U_2 = 0) = |c_{10}|^2 = 0$. Thus, Eq. (20) leads to $U_1 \equiv_\psi U_2$. Similarly, Eq. (21) leads to $U_1 \equiv_\psi D_2$, and Eq. (22) leads to $D_1 \equiv_\psi U_2$. Thus, by the transitivity of perfect correlation, we conclude $D_1 \equiv_\psi D_2$, or $P_\psi(D_1 = j, D_2 = k) = 0$ if $j \neq k$, and this contradicts Eq. (23).

7. Characterizations of perfectly correlated POVMs

7.1. Joint dilations of POVMs

For any POVM Π, there is a triple $(\mathcal{K}, \xi, L)$, called a *Naimark-Holevo dilation* of Π, consisting of a separable Hilbert space $\mathcal{K}$, a state vector $\xi \in \mathcal{K}$, and an observable L on $\mathcal{H} \otimes \mathcal{K}$ satisfying

$$\langle\psi', \Pi(\Delta)\psi\rangle = \langle\psi' \otimes \xi, E^L(\Delta)(\psi \otimes \xi)\rangle \tag{25}$$

for any $\psi, \psi' \in \mathcal{H}$ and $\Delta \in \mathcal{B}(\mathbf{R})$ [16]. We now extends the above notion to any pair of POVMs. A *joint dilation* of POVMs Π_1, Π_2 is a quadruple $(\mathcal{K}, \xi, X, Y)$ consisting of a separable Hilbert space $\mathcal{K}$, a state vector $\xi \in \mathcal{K}$, and observables X, Y on $\mathcal{H} \otimes \mathcal{K}$, satisfying

$$\langle\Pi_1(\Delta)\psi', \Pi_2(\Gamma)\psi\rangle = \langle E^X(\Delta)(\psi' \otimes \xi), E^Y(\Gamma)(\psi \otimes \xi)\rangle \tag{26}$$

for all $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$ and $\psi, \psi' \in \mathcal{H}$. In this case, we have

$$\mathrm{Tr}[\Pi_1(\Delta)\Pi_2(\Gamma)\rho] = \mathrm{Tr}[E^X(\Delta)E^Y(\Gamma)(\rho \otimes |\xi\rangle\langle\xi|)] \tag{27}$$

for all $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$ and $\rho \in \mathcal{H}$. The existence of the joint dilations is given in the following.

Theorem 7.1. *Any pair of POVMs has a joint dilation of them.*

Proof. Let Π_1, Π_2 be a pair of POVMs. Let $(\mathcal{K}_j, \xi_j, L_j)$ be a Naimark-Holevo dilation of Π_j for $j = 1, 2$. Let $\phi_1, \phi_2, \ldots$ be an arbitrary orthonormal basis of $\mathcal{H}$. Let $\eta_1^{(j)}, \eta_2^{(j)}, \ldots$ be an orthonormal basis of $\mathcal{K}_j$ such that $\eta_1^{(j)} = \xi_j$ for $j = 1, 2$. Then, by repeated uses of the Parceval identity, for any $\psi, \psi' \in \mathcal{H}$ and $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$ we have

$$\begin{aligned}\langle\Pi_1(\Delta)\psi', \Pi_2(\Gamma)\psi\rangle &= \sum_k \langle\Pi_1(\Delta)\psi', \phi_k\rangle\langle\phi_k, \Pi_2(\Gamma)\psi\rangle \\ &= \sum_k \langle E^{L_1}(\Delta)(\psi' \otimes \xi_1), \phi_k \otimes \xi_1\rangle\langle\phi_k \otimes \xi_2, E^{L_2}(\Gamma)(\psi \otimes \xi_2)\rangle \\ &= \sum_k \langle(E^{L_1}(\Delta) \otimes I_2)(\psi' \otimes \xi_1 \otimes \xi_2), \phi_k \otimes \xi_1 \otimes \xi_2\rangle \\ &\qquad \times \langle\phi_k \otimes \xi_1 \otimes \xi_2, (E^{L_2}(\Gamma) \otimes I_1)(\psi \otimes \xi_1 \otimes \xi_2)\rangle \\ &= \sum_{k,l,m} \langle(E^{L_1}(\Delta) \otimes I_2)(\psi' \otimes \xi_1 \otimes \xi_2), \phi_k \otimes \eta_l^{(1)} \otimes \eta_m^{(2)}\rangle \\ &\qquad \times \langle\phi_k \otimes \eta_l^{(1)} \otimes \eta_m^{(2)}, (E^{L_2}(\Gamma) \otimes I_1)(\psi \otimes \xi_1 \otimes \xi_2)\rangle \\ &= \langle(E^{L_1}(\Delta) \otimes I_2)(\psi' \otimes \xi_1 \otimes \xi_2), (E^{L_2}(\Gamma) \otimes I_1)(\psi \otimes \xi_1 \otimes \xi_2)\rangle,\end{aligned}$$

where I_j is the identity operator on $\mathcal{K}_j$. Thus, we have a joint dilation $(\mathcal{K}_1 \otimes \mathcal{K}_2, \xi_1 \otimes \xi_2, L_1 \otimes I_2, L_2 \otimes I_1)$. □

Using joint dilations, perfect correlations between POVMs are reduced to those between observables.

Theorem 7.2. *For any joint dilation $(\mathcal{K}, \xi, X, Y)$ of a pair of POVMs Π_1, Π_2, the POVMs Π_1 and Π_2 are perfectly correlated in a state $\rho \in \mathcal{S}(\mathcal{H})$ if and only if X and Y are perfectly correlated in $\rho \otimes |\xi\rangle\langle\xi|$. In this case, we have*

$$\Pi_1(f)\rho = \Pi_2(f)\rho \tag{28}$$

for any $f \in B(\mathbf{R})$.

Proof. Let $(\mathcal{K}, \xi, X, Y)$ be a joint dilation of Π_1 and Π_2. Then, from Eq. (27) it is easy to see that Π_1 and Π_2 are perfectly correlated in ρ if and only if so are X and Y in $\rho \otimes |\xi\rangle\langle\xi|$. In this case, from Theorem 3.2 we have

$$E^X(\Delta)\rho \otimes |\xi\rangle\langle\xi| = E^Y(\Delta)\rho \otimes |\xi\rangle\langle\xi|$$

for any $\Delta \in \mathcal{B}(\mathbf{R})$. Note that $\Pi_1(\Delta) = V_\xi^\dagger E^X(\Delta)V_\xi$ and $\Pi_2(\Delta) = V_\xi^\dagger E^Y(\Delta)V_\xi$ for all Δ, where $V_\xi\psi = \psi \otimes \xi$ for any $\psi \in \mathcal{H}$. Let $\psi \in \mathcal{H}$. We have $\Pi_1(\Delta)\rho\psi = V_\xi^\dagger E^X(\Delta)V_\xi\rho\psi = V_\xi^\dagger E^X(\Delta)(\rho\psi \otimes \xi) = V_\xi^\dagger E^Y(\Delta)(\rho\psi \otimes \xi) = V_\xi^\dagger E^X(\Delta)V_\xi\rho\psi = \Pi_2(\Delta)\rho\psi$, and by the standard argument we have $\Pi_1(f)\rho\psi = \Pi_2(f)\rho\psi$ for any $f \in B(\mathbf{R})$. Since ψ is arbitrary, we obtain Eq. (28). □

7.2. Perfect correlations between observables and POVMs

For any observable X and POVM Π, we say that X and Π are *perfectly correlated* in a state ρ iff E^X and Π are perfectly correlated in ρ. Now, we extend Theorem 3.2 to arbitrary pair of an observable and a POVM.

Theorem 7.3. *For any observable X, any POVM Π, and any state $\rho \in \mathcal{S}(\mathcal{H})$, the following conditions are equivalent.*

(i) *X and Π are perfectly correlated in ρ.*
(ii) *X and Π are perfectly correlated in any state $\sigma \in \mathcal{S}(X, \rho)$.*
(iii) *$E^X(\Delta)\rho = \Pi(\Delta)\rho$ for any $\Delta \in \mathcal{B}(\mathbf{R})$.*
(iv) *$f(X)\rho = \Pi(f)\rho$ for any $f \in \mathbf{B}(\mathbf{R})$.*
(v) *$f(X)P_{X,\rho} = \Pi(f)P_{X,\rho}$ for any $f \in \mathbf{B}(\mathbf{R})$.*

Proof. The implication (i) ⇒ (iv) follows from Theorem 7.2. The implication (iv) ⇒ (iii) is obvious. The implication (iii) ⇒ (i) follows from the relations

$$\mathrm{Tr}[E^X(\Delta)\Pi(\Gamma)\rho] = \mathrm{Tr}[E^X(\Delta)E^X(\Gamma)\rho] = \mathrm{Tr}[E^X(\Delta \cap \Gamma)\rho], \tag{29}$$

for any $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$. Now, we shall show the implications (i) ⇒ (v) ⇒ (ii) ⇒ (i). Suppose that condition (i) holds. Let $(\mathcal{K}, \xi, L)$ be a Naimark-Holevo dilation of Π. Then, it is easy to see that $(\mathcal{K}, \xi, X \otimes I, L)$ is a joint dilation of E^X and Π. It follows from the assumption and Theorem 7.2 that $X \otimes I$ and L are perfectly correlated in $\rho \otimes |\xi\rangle\langle\xi|$, and hence

$$f(X)\rho\psi \otimes \xi = f(L)(\rho\psi \otimes \xi) \tag{30}$$

for any $f \in B(\mathbf{R})$ and $\psi \in \mathcal{H}$. Let $\phi \in \mathcal{H}$. Then, we have $f(X)g(X)\rho\psi \otimes \xi = f(L)g\,(L)(\rho\psi \otimes \xi) = f(L)(g(X)\rho\psi \otimes \xi)$ for any $f, g \in B(\mathbf{R})$. Thus, we have

$$\begin{aligned} f(X)g(X)\rho\psi &= V_\xi^\dagger(f(X)g(X)\rho\psi \otimes \xi) = V_\xi^\dagger f(L)(g(X)\rho\psi \otimes \xi) \\ &= V_\xi^\dagger f(L)V_\xi g(X)\rho\psi. \end{aligned}$$

Since the vector of the form $g(X)\rho\psi$ with $g \in B(\mathbf{R})$, $\psi \in \mathcal{H}$ spans $\mathcal{C}(X,\rho)$, we obtain

$$f(X)P_{X,\rho} = V_\xi^\dagger f(L)V_\xi P_{X,\rho},$$

and hence the implication (i) $\Rightarrow$ (v) follows. Suppose that condition (v) holds. Let $\sigma \in \mathcal{S}(X,\rho)$ and $\Delta, \Gamma \in \mathcal{B}(\mathbf{R})$. Then, $P_{X,\rho}\sigma = \sigma$ and we have $E^X(\Gamma)\sigma = \Pi(\Gamma)\sigma$. Thus,

$$\mathrm{Tr}[E^X(\Delta)\Pi(\Gamma)\sigma] = \mathrm{Tr}[E^X(\Delta)E^X(\Gamma)\sigma] = \mathrm{Tr}[E^X(\Delta \cap \Gamma)\sigma].$$

It follows that X and Y are perfectly correlated in σ, and hence the implication (v) $\Rightarrow$ (ii) follows. Since the implication (ii) $\Rightarrow$ (i) is obvious, the proof is completed. □

8. Perfect correlations in measurements

8.1. Quantum instruments and measuring processes

A *measuring process* for $\mathcal{H}$ is defined to be a quadruple $(\mathcal{K}, \xi, U, M)$ consisting of a separable Hilbert space $\mathcal{K}$, a state vector ξ in $\mathcal{K}$, a unitary operator U on $\mathcal{H} \otimes \mathcal{K}$, and an observable M on $\mathcal{K}$ [17]. It is a plausible hypothesis in the theory of measurement that to any measuring apparatus $\mathbf{A}(\mathbf{x})$ with output variable $\mathbf{x}$ for a system $\mathbf{S}$ described by a Hilbert space $\mathcal{H}$, there corresponds a measuring process $(\mathcal{K}, \xi, U, M)$ such that $\mathcal{K}$ describes the probe $\mathbf{P}$ prepared in ξ just before the measurement, U describes the time evolution of the composite system $\mathbf{S} + \mathbf{P}$ during the measuring interaction, and that M describes the meter observable to be actually observed just after the measuring interaction [17–20,11,21]. Then, the probability distribution of the output $\mathbf{x}$ on input state ρ is given by

$$\Pr\{\mathbf{x} \in \Delta \| \rho\} = \mathrm{Tr}[(I \otimes E^M(\Delta))U(\rho \otimes |\xi\rangle\langle\xi|)U^\dagger], \tag{31}$$

and the conditional output state $\rho_{\{\mathbf{x} \in \Delta\}}$ of the apparatus on input state ρ given the outcome $\mathbf{x} \in \Delta$ is described by

$$\rho_{\{\mathbf{x}\in\Delta\}} = \frac{\mathrm{Tr}_{\mathcal{K}}[(I \otimes E^M(\Delta))U(\rho \otimes |\xi\rangle\langle\xi|)U^\dagger]}{\mathrm{Tr}[(I \otimes E^M(\Delta))U(\rho \otimes |\xi\rangle\langle\xi|)U^\dagger]}, \tag{32}$$

where $\mathrm{Tr}_{\mathcal{K}}$ stands for the partial trace over $\mathcal{K}$.

Two measuring apparatuses $\mathbf{A}(\mathbf{x}), \mathbf{A}(\mathbf{y})$, or corresponding measuring processes are called *statistically equivalent* iff they have the same output distributions and the same conditional output states on each input state, i.e., $\Pr\{\mathbf{x} \in \Delta \| \rho\} = \Pr\{\mathbf{y} \in \Delta \| \rho\}$ and $\rho_{\{\mathbf{x}\in\Delta\}} = \rho_{\{\mathbf{y} \in \Delta\}}$ for all $\rho \in \mathcal{S}(\mathcal{H})$ and $\Delta \in \mathcal{B}(\mathbf{R})$. The statistical equivalence classes of all the measuring processes are characterized by completely positive map valued measures as follows.

Denote by $\tau c(\mathcal{H})$ the space of trace class operators on $\mathcal{H}$ and by $\mathcal{L}(\tau c(\mathcal{H}))$ the space of bounded linear transformations on $\tau c(\mathcal{H})$. A linear transformation $T \in \mathcal{L}(\tau c(\mathcal{H}))$ is called *completely positive* iff $T \otimes \mathrm{id}_n \in \mathcal{L}(\tau c(\mathcal{H} \otimes \mathbf{C}^n))$ is a positive transformation for any positive integer n. Denote by $\mathcal{CP}(\tau c(\mathcal{H}))$ the space of completely positive maps on $\tau c(\mathcal{H})$. An *instrument* is a countably additive normalized completely positive map valued measure from $\mathcal{B}(\mathbf{R})$ to $\mathcal{L}(\tau c(\mathcal{H}))$, i.e., a mapping $\mathcal{I} : \mathcal{B}(\mathbf{R}) \to \mathcal{CP}(\tau c(\mathcal{H}))$ satisfying that $\mathcal{I}(\mathbf{R})$ is trace-preserving and $\sum_{j=1}^{\infty} \mathcal{I}(\Delta_j) = \mathcal{I}(\mathbf{R})$ in the strong operator topology for any disjoint Borel sets $\Delta_1, \Delta_2, \ldots$ such that $\cup_j \Delta_j = \mathbf{R}$ [17].

For any instrument $\mathcal{I}$ and state ρ, the relation

$$\mu_\rho^{\mathcal{I}}(\Delta) = \mathrm{Tr}[\mathcal{I}(\Delta)\rho] \tag{33}$$

defines a probability measure on $\mathcal{B}(\mathbf{R})$ called the *output distribution* of $\mathcal{I}$ on input state ρ, and the state

$$\frac{\mathcal{I}(\Delta)\rho}{\mathrm{Tr}[\mathcal{I}(\Delta)\rho]} \tag{34}$$

is called the *output state* of $\mathcal{I}$ on input state ρ given Δ [22]. The dual map of $\mathcal{I}(\Delta)$ is the linear transformation $\mathcal{I}(\Delta)^*$ on $\mathcal{L}(\mathcal{H})$ defined by

$$\mathrm{Tr}[(\mathcal{I}(\Delta)^* A)\rho] = \mathrm{Tr}[A\mathcal{I}(\Delta)\rho] \tag{35}$$

for all $A \in \mathcal{L}(\mathcal{H})$, $\rho \in \tau c(\mathcal{H})$, and $\Delta \in \mathcal{B}(\mathbf{R})$. Then, $\mathcal{I}(\Delta)^*$ is a normal completely positive map on $\mathcal{L}(\mathcal{H})$ [23] and especially $\mathcal{I}(\mathbf{R})^*$ is unit-preserving. The relation

$$\Pi^{\mathcal{I}}(\Delta) = \mathcal{I}(\Delta)^* I, \tag{36}$$

where $\Delta \in \mathcal{B}(\mathbf{R})$ defines a POVM, called the *POVM* of $\mathcal{I}$, which satisfies

$$\mu_\rho^{\mathcal{I}}(\Delta) = \mathrm{Tr}[\Pi^{\mathcal{I}}(\Delta)\rho] \tag{37}$$

for all $\Delta \in \mathcal{B}(\mathbf{R})$ and $\rho \in \mathcal{S}(\mathcal{H})$.

For any measuring process $\mathbf{M} = (\mathcal{K}, \xi, U, M)$, the relation

$$\mathcal{I}_{\mathbf{M}}(\Delta)\rho = \mathrm{Tr}_{\mathcal{K}}[(I \otimes E^M(\Delta))U(\rho \otimes |\xi\rangle\langle\xi|)U^\dagger], \tag{38}$$

where $\rho \in \mathcal{S}(\mathcal{H})$ and $\Delta \in \mathcal{B}(\mathbf{R})$, defines an instrument $\mathcal{I}_{\mathbf{M}}$, called the *instrument* of $\mathbf{M}$. Then, the POVM of $\mathcal{I}$ is called the *POVM* of $\mathbf{M}$ and denoted by $\Pi_{\mathbf{M}}$. We have

$$\Pi_{\mathbf{M}}(\Delta) = \mathcal{I}_{\mathbf{M}}(\Delta)^* I = \mathrm{Tr}_{\mathcal{K}}[U^\dagger(I \otimes E^M(\Delta))U(I \otimes |\xi\rangle\langle\xi|)] \tag{39}$$

for all $\Delta \in \mathcal{B}(\mathbf{R})$. For all $\rho \in \mathcal{S}(\mathcal{H})$ and $\Delta \in \mathcal{B}(\mathbf{R})$, we have

$$\Pr\{\mathbf{x} \in \Delta \| \rho\} = \mathrm{Tr}[\mathcal{I}_{\mathbf{M}}(\Delta)\rho] = \mathrm{Tr}[\Pi_{\mathbf{M}}(\Delta)\rho] \tag{40}$$

and

$$\rho_{\{\mathbf{x}\in\Delta\}} = \frac{\mathcal{I}(\Delta)\rho}{\mathrm{Tr}[\mathcal{I}(\Delta)\rho]}, \tag{41}$$

provided $\mathrm{Tr}[\mathcal{I}(\Delta)\rho] > 0$. Thus, two measuring processes are statistically equivalent if and only if they have the same instrument.

Conversely, it has been proved in [17] that for any instrument $\mathcal{I}$, there exists a measuring process $\mathbf{M} = (\mathcal{K}, \xi, U, M)$ such that $\mathcal{I} = \mathcal{I}_{\mathbf{M}}$. Thus, every instrument corresponds at least one measuring process, and therefore the instruments are in one-to-one correspondence with the statistical equivalence classes of measuring processes.

8.2. Precise measurements of observables

Once the notion of measurement has been fully generalized by the notion of instruments, a fundamental problem is to recover the conventional notion of measurements of observables in this general formulation. In what follows, we shall give an answer to this problem in the light of the notion of quantum perfect correlations.

According to a fundamental postulate of quantum mechanics, if an apparatus $\mathbf{A}(\mathbf{x})$ measures an observable A in a state ρ, the probability distribution of the output $\mathbf{x}$ on input state ρ should satisfy the *Born statistical formula* (BSF)

$$\Pr\{\mathbf{x} \in \Delta \| \rho\} = \mathrm{Tr}[E^A(\Delta)\rho], \tag{42}$$

where $\Delta \in \mathcal{B}(\mathbf{R})$. From the above it is tempting to say that an apparatus $\mathbf{A}(\mathbf{x})$ measures observable A in state ρ iff it satisfies the BSF Eq. (42). However, to reproduce the probability distribution of observable A in state ρ is a necessary but not sufficient condition for the apparatus $\mathbf{A}(\mathbf{x})$ to measure A in ρ. For example, suppose $\mathcal{H} = \mathcal{K} \otimes \mathcal{K}$, $\rho = \sigma \otimes \sigma$, and $A = X \otimes I$ and $B = I \otimes X$ for some Hilbert space $\mathcal{K}$, a state σ of $\mathcal{K}$, and an observable X of $\mathcal{K}$. In this case, we have $\mathrm{Tr}[E^A(\Delta)\rho] = \mathrm{Tr}[E^B(\Delta)\rho] = \mathrm{Tr}[E^X(\Delta)\sigma]$, so that every apparatus $\mathbf{A}(\mathbf{y})$ measuring B in state ρ also satisfies the BSF for A in ρ. However, we cannot consider that the apparatus $\mathbf{A}(\mathbf{y})$ measures A in ρ as well as B in ρ. Since A and B are independent observables in the separated subsystems, so that another apparatus $\mathbf{A}(\mathbf{x})$ may simultaneously measure A and may obtain a different outcome of the A measurement. In this case, we can say that the apparatus $\mathbf{A}(\mathbf{x})$ measures A but the apparatus $\mathbf{A}(\mathbf{y})$ does not.

To find a satisfactory condition to ensure that a given instrument $\mathcal{I}$ measures A in ρ, let us consider a measuring process $\mathbf{M} = (\mathcal{K}, \xi, U, M)$ of $\mathcal{I}$. Suppose that we measure A at time t at which the system $\mathbf{S}$ described by Hilbert space $\mathcal{H}$ is in state ρ and that the measuring interaction turns on from time t to $t + \Delta t$. In the Heisenberg picture with the original state $\rho \otimes |\xi\rangle\langle\xi|$, we write $A(t) = A \otimes I$, $A(t + \Delta t) = U^\dagger(A \otimes I)U$, $M(t) = I \otimes M$, and $M(t + \Delta t) = U^\dagger(I \otimes M)U$. Then, to measure $A(t)$, this measurement actually measures $M(t + \Delta t)$, so that observables $A(t)$ and $M(t + \Delta t)$ should be perfectly correlated in the original state $\rho \otimes |\xi\rangle\langle\xi|$.

In the previous example, it is concluded that the meter observable of $\mathbf{A}(\mathbf{y})$ after the measuring interaction, $M(t + \Delta t)$, cannot be perfectly correlated with the observable A before the interaction, $A(t)$. In fact, $M(t + \Delta t)$ should be perfectly correlated with the observable B before the interaction, $B(t)$, while $A(t)$ and $B(t)$ are not perfectly correlated before the interaction. It follows from the transitivity of perfect correlations that $A(t)$ and $M(t + \Delta t)$ cannot be perfectly correlated.

It is also clear that given two "meter" observables M_1 and M_2 in the external system described by a Hilbert space $\mathcal{K}$ and given the original state $\rho \otimes |\xi\rangle\langle\xi|$ of $\mathcal{H} \otimes \mathcal{K}$ at time t, if both the pair of $A(t)$ and $M_1(t + \Delta t_1)$ and the pair of $A(t)$ and $M_2(t + \Delta t_2)$ are perfectly correlated in the original state, then we can conclude that both meters give the concordant outcome from the transitivity of perfect correlations.

According to the above consideration, it is natural to say that a measuring process $\mathbf{M} = (\mathcal{K}, \xi, U, M)$ *precisely measures* an observable A on input state ρ iff the observable $A \otimes I$ and $U^\dagger(I \otimes M)U$ are perfectly correlated in the state $\rho \otimes |\xi\rangle\langle\xi|$, and that an instrument $\mathcal{I}$ *precisely measures* an observable A on input state ρ iff every measuring process $\mathbf{M}$ for $\mathcal{I}$ precisely measures A on input state ρ. In the above, the adverb "precisely" is used to distinguish this case from any approximate measurements of the same observable.

The following theorem shows that whether the measuring process $\mathbf{M}$ precisely measures A on ρ is determined solely by the corresponding POVM.

Theorem 8.1. *A measuring process* $\mathbf{M} = (\mathcal{K}, \xi, U, M)$ *precisely measures an observable* A *in a state* ρ *if and only if the POVM of* $\mathbf{M}$ *is perfectly correlated with the observable* A *in the state* ρ.

Proof. The assertion follows immediately from the relations

$$\mathrm{Tr}[(E^A(\Delta) \otimes I)U^\dagger(I \otimes E^M(\Gamma))U(\rho \otimes |\xi\rangle\langle\xi|)] = \mathrm{Tr}[E^A(\Delta)\Pi_{\mathbf{M}}(\Gamma)\rho],$$
$$\mathrm{Tr}[(E^A(\Delta) \otimes I)(\rho \otimes |\xi\rangle\langle\xi|)] = \mathrm{Tr}[E^A(\Delta)\rho]. \qquad \square$$

The following theorem characterizes, up to statistical equivalence, the precise measurements of an observable in a given state.

Theorem 8.2. *For any instrument* $\mathcal{I}$ *with POVM* $\Pi^{\mathcal{I}}$, *any observable* A, *and any state* ρ, *the following conditions are all equivalent.*

(i) $\mathcal{I}$ *precisely measures* A *in* ρ.
(ii) $\Pi^{\mathcal{I}}$ *is perfectly correlated to* A *in* ρ.
(iii) $\Pi^{\mathcal{I}}$ *is perfectly correlated to* A *in any state* $\sigma \in \mathcal{S}(A, \rho)$.
(iv) $\mathcal{I}$ *satisfies the BSF for* A *in any state* $\sigma \in \mathcal{S}(A, \rho)$.
(v) $\Pi^{\mathcal{I}}(\Delta)\sigma = E^A(\Delta)\sigma$ *for any* $\sigma \in \mathcal{S}(A, \rho)$ *and* $\Delta \in \mathcal{B}(\mathbf{R})$.
(vi) $\Pi^{\mathcal{I}}(\Delta)P_{A,\rho} = E^A(\Delta)P_{A,\rho}$ *for any* $\Delta \in \mathcal{B}(\mathbf{R})$.

Proof. The assertion follows easily from Theorems 7.3 and 8.1. $\square$

In the conventional interpretation of instruments proposed by Davies and Lewis [22], an instrument $\mathcal{I}$ is considered to precisely measure A in *every* state ρ iff it satisfies the BSF for A in every state ρ. Since the BSF for A in a given state ρ does not ensure that the instrument $\mathcal{I}$ precisely measures A in ρ, the above hypothesis lacks an immediate justification in the sense that it is not immediately clear whether this hypothesis excludes the ambiguity of the simultaneous meter readings of the same observable. However, this hypothesis has been finally justified by the above theorem, which concludes that $\mathcal{I}$ precisely measures A in every state ρ if and only if $\mathcal{I}$ satisfies the BSF for A in every state.

8.3. von Neumann's model of repeatable measurement

It was shown by von Neumann [2] that a repeatable measurement of an observable

$$A = \sum_n a_n|\phi_n\rangle\langle\phi_n| \tag{43}$$

on $\mathcal{H}$ with eigenvalues $a_1, a_2, \ldots$ and orthonormal basis of eigenvectors $\phi_1, \phi_2, \ldots$ can be realized by a unitary operator U on the tensor product $\mathcal{H} \otimes \mathcal{K}$ with another separable Hilbert space $\mathcal{K}$ with orthonormal basis $\{\xi_n\}$ such that

$$U(\phi_n \otimes \xi) = \alpha_n\phi_n \otimes \xi_n, \tag{44}$$

where ξ is an arbitrary state vector in $\mathcal{K}$, and α_n is an arbitrary phase factor, i.e., $|\alpha_n| = 1$, for all n. Let

$$M = \sum_n a_n |\xi_n\rangle\langle\xi_n| \tag{45}$$

be an observable on $\mathcal{K}$ called the meter. von Neumann's model defines an apparatus $\mathbf{A}(\mathbf{x})$ with measuring process $(\mathcal{K}, \xi, U, M)$.

Let us suppose that the initial state of the system is given by an arbitrary state vector $\psi = \sum_n \sqrt{p_n}\phi_n$. Then, it follows from the linearity of U we have

$$U(\psi \otimes \xi) = \sum_n \sqrt{p_n}\phi_n \otimes \xi_n. \tag{46}$$

The conventional explanation as to why this transformation can be regarded as a measurement is as follows; symbols are adapted to the present context in the quote below. "In the state (46), obtained by the measurement, there is a statistical correlation between the state of the object and that of the apparatus: the simultaneous measurement on the system—object-plus-apparatus—of the two quantities, one of which is the originally measured quantity of the object and the second the position of the pointer of the apparatus, always leads to concordant results. As a result, one of these measurements is unnecessary: the state of the object can be ascertained by an observation on the apparatus. This is a consequence of the special form of the state vector (46), on not containing any $\phi_m \otimes \xi_n$ term with $n \neq m$ [24]." "The equations of motion permit the description of the process whereby the state of the object is mirrored by the state of an apparatus. The problem of a measurement on the object is thereby transformed into the problem of an observation on the apparatus [24]."

The above explanation correctly points out the existence of the statistical correlation between the measured observable A and the meter observable M in the state (46). However, this is not the statistical correlation between the measured observable before the interaction and the meter observable after the interaction, but that between those observables after the interaction. Thus, the above statistical correlation does not even ensure that the probability distribution of the measured observable before the interaction is reproduced by the observation of the meter observable after the interaction.

The role of the measuring interaction described by U should be to make the following two correlations: (i) the correlation between the measured observable A *before* the interaction and the meter M *after* the interaction, and (ii) the correlation between the meter M *after* the interaction and the measured observable A *after* the interaction. The first correlation is required by the *value reproducing requirement* that the interaction transfers the value of the measured observable A before the interaction to the value of the meter M after the interaction. The second correlation is required by the *repeatability hypothesis* that if the meter observable M has the value a_n after the interaction, then the observable A also have the same value a_n after the interaction so that the second measurement of A after the interaction reproduce the same value of the meter of the first measurement of A.

Now, we shall show that those requirements are actually satisfied. Let $\eta_0, \eta_1, \ldots$ be an orthonormal basis of $\mathcal{H}$ such that $\eta_0 = \xi$, namely an orthonormal basis extending $\{\xi\}$. Let $\Psi_{n,m}$ be a unit vector in $\mathcal{H}$ defined by $\Psi_{n,m} = U^\dagger(\phi_n \otimes \xi_m)$ for any n, m. Then, we have $\Psi_{n,n} = \phi_n \otimes \xi$ and the family $\{\Psi_{n,m}\}$ is an orthonormal basis of $\mathcal{H}$. By simple calculations, we have

$$A \otimes I = A \otimes |\xi\rangle\langle\xi| + \sum_{m \neq 0} A \otimes |\eta_m\rangle\langle\eta_m|, \tag{47}$$

$$U^{\dagger}(A \otimes I)U = A \otimes |\xi\rangle\langle\xi| + \sum_{n \neq m} a_n |\Psi_{n,m}\rangle\langle\Psi_{n,m}|, \tag{48}$$

$$U^{\dagger}(I \otimes M)U = A \otimes |\xi\rangle\langle\xi| + \sum_{n \neq m} a_m |\Psi_{n,m}\rangle\langle\Psi_{n,m}|, \tag{49}$$

where $\sum_{n \neq m}$ stands for the summation over all n, m with $n \neq m$. By the above relations it is now obvious that $A \otimes I = U^{\dagger}(A \otimes I)U = U^{\dagger}(I \otimes M)U$ on their common invariant subspace $\mathcal{H} \otimes [\xi]$, so that those three observables are perfectly correlated in the state $\psi \otimes \xi$ for every state vector ψ in $\mathcal{H}$. Therefore, von Neumann's model $(\mathcal{K}, \xi, U, M)$ satisfies both the the value reproducing requirement and the repeatability hypothesis.

The following theorem characterizes the unitary operators that fulfil the above two requirements.

Theorem 8.3. *Let $\{\phi_n\}$ and $\{\xi_n\}$ be orthonormal bases of $\mathcal{H}$ and $\mathcal{K}$, respectively, and the observables A and B be given by Eq. (43) and Eq. (45), respectively. Then, a unitary operator U on $\mathcal{H} \otimes \mathcal{K}$ and a state vector $\xi \in \mathcal{K}$ satisfy Eq. (44) if and only if (i) $A \otimes I$ and $U^{\dagger}(I \otimes B)U$ are perfectly correlated in $\psi \otimes \xi$ and that (ii) $U^{\dagger}(A \otimes I)U$ and $U^{\dagger}(I \otimes B)U$ are perfectly correlated in $\psi \otimes \xi$ for every state vector $\psi \in \mathcal{H}$.*

Proof. Suppose that U and ξ satisfy Eq. (44). Without any loss of generality we assume $U(\phi_n \otimes \xi) = \phi_n \otimes \xi_n$ for all n; otherwise, we can replace ξ_n by $\alpha_n \xi_n$ without changing B. Let $\psi = \sum_n c_n \phi_n$. By linearity of U we have $U(\psi \otimes \xi) = \sum_n c_n \phi_n \otimes \xi_n$. Thus, it follows from the argument on the entangled state given Eq. (18), $A \otimes I$ and $I \otimes B$ are perfectly correlated in $U(\psi \otimes \xi)$. By Theorem 2.3, $U^{\dagger}(A \otimes I)U$ and $U^{\dagger}(I \otimes B)U$ are perfectly correlated in $\psi \otimes \xi$. Thus, condition (ii) holds. Let $\{\eta_n\}$ be an orthonormal basis of $\mathcal{K}$ such that $\eta_1 = \xi$. Then, we have

$$U(E^A(a_n) \otimes I)(\psi \otimes \xi) = U(|\phi_n\rangle\langle\phi_n| \otimes I) \sum_j c_j \phi_j \otimes \xi = c_n U(\phi_n \otimes \xi) = c_n \phi_n \otimes \xi_n,$$

and

$$(I \otimes E^B(a_m))U(\psi \otimes \xi) = (I \otimes |\xi_m\rangle\langle\xi_m|) \sum_j c_j \phi_j \otimes \xi_j = c_m \phi_m \otimes \xi_m.$$

Thus, we have

$$\langle (E^A(a_n) \otimes I)(\psi \otimes \xi), U^{\dagger}(I \otimes E^B(a_m))U(\psi \otimes \xi)\rangle = c_n^* c_m \delta_{n,m},$$

and this shows that $A \otimes I$ and $U^{\dagger}(I \otimes B)U$ are perfectly correlated in $\psi \otimes \xi$. Thus, we have proved the necessity of conditions (i) and (ii). Conversely, suppose that conditions (i) and (ii) hold. Let $\psi = \phi_n$. Since $A \otimes I$ and $U^{\dagger}(I \otimes B)U$ are perfectly correlated in $\psi \otimes \xi$, we have

$$\langle (I \otimes E^B(a_n))U(\phi_n \otimes \xi), U(\phi_n \otimes \xi)\rangle = \langle (E^A(a_n) \otimes I)(\phi_n \otimes \xi), (\phi_n \otimes \xi)\rangle = 1.$$

Thus, $U(\phi_n \otimes \xi) = \eta_n \otimes \xi_n$ for some state vector η_n. Since $A \otimes I$ and $I \otimes B$ are perfectly correlated in $U(\psi \otimes \xi)$, we have

$$\begin{aligned}\langle E^A(a_n)\eta_n, \eta_n\rangle &= \langle (E^A(a_n) \otimes I)(\eta_n \otimes \xi_n), (\eta_n \otimes \xi_n)\rangle \\ &= \langle (I \otimes E^B(a_n))(\eta_n \otimes \xi_n), (\eta_n \otimes \xi_n)\rangle = 1.\end{aligned}$$

Thus, $|\eta_n\rangle\langle\eta_n| = |\phi_n\rangle\langle\phi_n|$, so that U and ξ satisfy Eq. (44).

Now, we return to von Neumann's measurement model described by Eq. (44). The measurement is said to satisfy the *nondemolition condition*, iff the measured observable is not disturbed by the measuring interaction, so that $A \otimes I$ and $U^{\dagger}(A \otimes I)U$ are perfectly correlated in $\psi \otimes \xi$. As we have shown in Theorem 4.2 perfect correlations are transitive. Thus, the perfect correlation between $A \otimes I$ and $U^{\dagger}(I \otimes B)U$ and that between $U^{\dagger}(A \otimes I)U$ and $U^{\dagger}(I \otimes B)U$ implies the perfect correlation between $A \otimes I$ and $U^{\dagger}(A \otimes I)U$. In the same way, we will be able to explain that two out of three conditions, (i) the valued reproducing condition, (ii) the repeatability hypothesis, and (iii) the nondemolition condition, imply the other one, as straightforward consequence of the transitivity of perfect correlations.

9. Concluding remarks

Let X, Y be a pair of (discrete) observables and ψ a state. Consider the following conditions.

(i) (Equi-valuedness) No joint measurements of X and Y in ψ, if any, give different values, i.e.,

$$\langle E^X(\varDelta)\psi, E^Y(\varGamma)\psi\rangle = 0$$

if $\varDelta \cap \varGamma = \emptyset$.

(ii) (Reproducibility) Successive projective measurements of X and Y in ψ always give the same value irrespective of the order of measurements, i.e.,

$$\sum_{y\in\varGamma} \|E^X(\varDelta)E^Y(\{y\})\psi\|^2 = \sum_{x\in\varDelta} \|E^Y(\varGamma)E^X(\{x\})\psi\|^2 = 0$$

if $\varDelta \cap \varGamma = \emptyset$.

(iii) (Zero difference) The difference $X - Y$ has the definite value zero in ψ. i.e.,

$$(X - Y)\psi = 0.$$

(iv) (Identical distributivity) Independent measurements of X and Y in ψ have the identical output probability distribution, i.e.,

$$\|E^X(\varDelta)\psi\|^2 = \|E^Y(\varDelta)\psi\|^2$$

for any $\varDelta \in \mathcal{B}(\mathbf{R})$.

In this paper, we have shown the following logical relations among the above conditions. The following implications holds: (i) $\iff$ (ii), (i) $\Rightarrow$ (iii), (i) $\Rightarrow$ (iv). However, none of the implications (iii) $\Rightarrow$ (i), (iii) $\Rightarrow$ (iv), (iv) $\Rightarrow$ (i), and (iv) $\Rightarrow$ (iii) hold. If X and Y commute, (i) $\iff$ (iii) and (iii) $\Rightarrow$ (iv) holds, but (iv) $\Rightarrow$ (iii) still does not hold. To clarify the mutual relations, we have considered the notion of the cyclic subspace $\mathcal{C}(X,\psi)$ or $\mathcal{C}(Y,\psi)$ and required conditions (iii) and (iv) to be satisfied by any state ϕ in $\mathcal{C}(X,\psi)$ or $\mathcal{C}(Y,\psi)$, as follows.

(iii)′ $(X - Y)\phi = 0$ for any $\phi \in \mathcal{C}(X,\psi)$.

(iv)′ $\|E^X(\varDelta)\phi\|^2 = \|E^Y(\varDelta)\phi\|^2$ for any $\varDelta \in \mathcal{B}(\mathbf{R})$ and any $\phi \in \mathcal{C}(X,\psi)$.

Then, we have shown that all the conditions (i), (ii), (iii)′, and (iv)′ are mutually equivalent. According to this, we have proposed and justified to say that X and Y are perfectly correlated in ψ iff one of the above equivalent conditions is satisfied.

We have also given an appropriate generalizations of the above considerations to arbitrary observables X, Y and arbitrary state ρ.

We have shown that so defined relation $X \equiv_\rho Y$ meaning X and Y are perfectly correlated in ρ is an equivalence relation on all the observables. In particular, if X and Y are perfectly correlated as well as Y and Z, we can conclude that so are X and Z. This suggests that perfectly correlated observables can be interpreted to have the same value that can be realized by joint measurements of them, even though the quantum state determines it only randomly.

The above interpretation has given a new insight on the state dependent definition of precise measurements of observables. Even though the outcome of a measurement might be used to infer what is the state before or after the measurement as in quantum state estimation or quantum state reduction, this inference cannot be done without appealing to the fact that any measurement measures some observable in the sense of the Born rule; recall that even a POVM measurement corresponds to a measurement of an observable in a larger system and as such a mathematical POVM can be identified with a real experiment. Thus, the most fundamental question in measurement theory is the one as to what observable is (precisely) measured by a given apparatus.

Conventionally, this question has been answered only in a state independent manner as follows: The apparatus measures an observable X if and only if the probability reproducing condition (PRC) is satisfied for *any* input state, where the PRC requires that the output probability distribution reproduces the theoretical probability distribution predicted by the Born rule. However, the justification of the above definition has not been clear, since the probability reproducing condition for *a given* input state does not imply that the measurement is precise in that state. In this respect, our result has successfully justified the conventional definition in that we have given a definition of a precise measurement in a given state and showed that the conventional definition indeed requires the measurement is precise in any input state.

The state dependent definition is not a pedantic justification of the conventional approach. In fact, some measuring apparatus in a laboratory can accept only a small class of states from the whole Hilbert space of the state vectors. For instance, every microscope cannot measure the position of a particle outside of the scope. Thus, the experimenter should have a criterion to judge whether or not the apparatus measures the given observable depending on the input state. Such a criterion was not even discussed in measurement theory before the present investigation.

Acknowledgments

This work was supported by the Strategic Information and Communications R&D Promotion Scheme of the MPHPT of Japan, by the CREST project of the JST, and by the Grant-in-Aid for Scientific Research of the JSPS.

References

[1] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. 47 (1935) 777.
[2] J. von Neumann, Mathematische Grundlagen der Quantenmechanik, Springer, Berlin, 1932.
[3] M.A. Nielsen, I.L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000.
[4] L. Hardy, Phys. Rev. Lett. 71 (1993) 1665.
[5] S.K. Berberian, Notes on Spectral Theory, Van Nostrand, Princeton, NJ, 1966.
[6] C.W. Helstrom, Quantum Detection and Estimation Theory, Academic, New York, 1976.

[7] A.S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory, North-Holland, Amsterdam, 1982.
[8] P. Lahti, J.-P. Pellonpää, K. Ylinen, J. Math. Phys. 40 (1999) 2181.
[9] P. Busch, T. Heinonen, P. Lahti, Phys. Lett. A 320 (2004) 261.
[10] K. Ylinen, On a theorem of Gudder on joint distributions of observables, in: P. Lahti, P. Mittelstaedt (Eds.), Symposium on the Foundations of Modern Physics, World Scientific, Singerpore, 1985, pp. 691–694.
[11] M. Ozawa, Phys. Rev. A 63 (2001) 032109.
[12] K. Urbanik, Studia Math. 21 (1961) 117.
[13] E. Wigner, Phys. Rev. 40 (1932) 749.
[14] J.E. Moyal, Proc. Cambr. Phil. Soc. 45 (1949) 99.
[15] J.L. Cereceda, Found. Phys. Lett. 12 (1999) 211.
[16] A.S. Holevo, J. Multivar. Anal. 3 (1973) 337.
[17] M. Ozawa, J. Math. Phys. 25 (1984) 79.
[18] P. Busch, P.J. Lahti, P. Mittelstaedt, The Quantum Theory of Measurement, Lecture Notes in Physics m2, Springer, Berlin, 1991.
[19] V.B. Braginsky, F.Y. Khalili, Quantum Measurement, Cambridge University Press, Cambridge, 1992.
[20] M. Ozawa, Phys. Rev. A 62 (2000) 062101.
[21] M. Ozawa, Ann. Phys. 311 (2004) 350.
[22] E.B. Davies, J.T. Lewis, Commun. Math. Phys. 17 (1970) 239.
[23] M. Takesaki, Theory of Operator Algebras I, Springer, New York, 1979.
[24] E.P. Wigner, Am. J. Phys. 31 (1963) 6.