Universality of sequential quantum measurements

Teiko Heinosaari*
Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, Finland

Takayuki Miyadera†
Department of Nuclear Engineering, Kyoto University, 6158540, Kyoto, Japan
(Received 20 February 2014; published 13 February 2015)

We show that any jointly measurable pair of quantum observables can be obtained in a sequential measurement scheme, even if the second observable will be decided after the first measurement. This means that it is possible to perform a measurement of any quantum observable in a way that does not disturb the subsequent measurements more than is dictated by joint measurability. Only measurements with a specific structure have this universality feature. As a supplementing result, we provide a characterization of all possible joint measurements obtained from a sequential measurement lacking universality.

DOI: 10.1103/PhysRevA.91.022110 PACS number(s): 03.65.Ta

I. INTRODUCTION

Sequential quantum measurements have received renewed attention recently and their range of application has become broader. This concept has been investigated, for example, in the context of state discrimination [1], tomography [2], cryptography [3,4], undecidability [5], decoding [6,7], and contextuality [8]. However, the fundamental limitations of the sequential method of performing quantum measurements have not been addressed systematically. Understanding the limitations should be important from the perspective of quantum information science.

On the conceptual side, a sequential measurement can be seen as a special type of joint measurement. The connections of joint measurability with Bell nonlocality [9] and steering [10,11] have been recently clarified. In this respect, one can ask if sequential measurements form a strictly smaller class of joint measurements, or if any jointly measurable set of observables has a sequential realization.

The main result of this work is that any jointly measurable pair of observables can be obtained via a sequential measurement scheme, even if the second observable is decided *after* the first measurement. In particular, suppose an observable A is pairwisely jointly measurable with two observables B and C, but the triplet $(\mathsf{A},\mathsf{B},\mathsf{C})$ is not jointly measurable. One could think that we have to choose beforehand which pair, (A,B) or (A,C), we will measure. However, this intuition is not true. One can, in fact, first measure A and only later decide whether to implement a measurement of B or C.

Our result means that it is possible to perform a measurement of any quantum observable in a way that does not disturb the subsequent measurements more than is dictated by joint measurability. In practice, this means that for two jointly measurable observables A and B, we can find an observable B', depending on both A and B, such that a sequential measurement of A and B' implements a joint measurement of A and B. This striking feature, which we call *universality*, holds only for certain measurement schemes. As a supplementary result, we derive a characterization of all possible joint measurements obtained from any nonuniversal measurement scheme.

II. SEQUENTIAL AND JOINT MEASUREMENTS

The mathematical framework for sequential quantum measurements was provided a long time ago [12]. For our purposes it suffices to use a rudimentary formulation. We will describe a quantum measurement as a pair (A,Λ) consisting of an observable A and a channel Λ, where A assigns a probability distribution of measurement outcomes to each input state and Λ maps each input state ϱ into an output state $\Lambda(\varrho)$; see Fig. 1. One can obviously give more detailed descriptions of a quantum measurement, but this kind of simple description is enough for our present purposes. Two common levels of descriptions are those given by instruments and measurement models. An overview can be found in Ref. [13].

Mathematically, an observable, in its most general form, is presented as a positive-operator valued measure (POVM). We will assume that there are a finite number of outcomes, so an observable is a function $\mathsf{A}: x \mapsto \mathsf{A}(x)$ from a finite set of measurement outcomes $\Omega_{\mathsf{A}} \subset \mathbb{Z}$ to the set of positive of operators on an input Hilbert space $\mathcal{H}_{\text{in}}$, and this function must satisfy the normalization constrain $\sum_{x\in\Omega_{\mathsf{A}}} \mathsf{A}(x) = \mathbb{1}$, where $\mathbb{1}$ is the identity operator on $\mathcal{H}_{\text{in}}$.[1] The probability of obtaining an outcome x for an input state ϱ is $\text{tr}[\varrho \mathsf{A}(x)]$.

A channel Λ is presented as a completely positive and trace preserving linear map on Hilbert space operators. It transforms an input state ϱ on $\mathcal{H}_{\text{in}}$ into an output state $\Lambda(\varrho)$ on $\mathcal{H}_{\text{out}}$. We assume that the outcome of the measurement is not used for selection, so the output state $\Lambda(\varrho)$ describes the unconditional state change. We allow $\mathcal{H}_{\text{out}}$ to be different from $\mathcal{H}_{\text{in}}$; physically this amounts to either including an environment in the description of the output system, or to discarding some part of the input system.

*teiko.heinosaari@utu.fi
†miyadera@nucleng.kyoto-u.ac.jp

[1] In the literature POVMs are also called generalized observables or generalized measurements. A POVM can have also infinite number of outcomes, in which case it assigns a positive operator to sets; see e.g. [13]

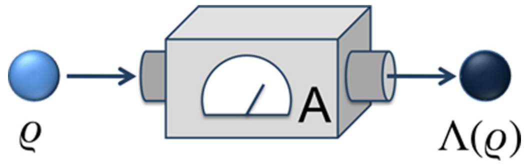

FIG. 1. (Color online) A quantum measurement leads to a classical output (measurement outcome) and quantum output (transformed quantum state). The probabilities for measurement outcomes are given by an observable A, while the averaged output state is given by a channel Λ.

An important point is that not every pair of an observable and channel gives a valid description of a quantum measurement. Specifically, a channel Λ and an observable A can describe the same measurement if and only if there exist completely positive maps Φ_x such that

$$\sum_x \Phi_x = \Lambda \quad \text{and} \quad \text{tr}[\Phi_x(\varrho)] = \text{tr}[\varrho \mathsf{A}(x)] \tag{1}$$

for all outcomes x and input states ϱ [14]; in this case we say that Λ is A-channel. Perhaps the most commonly used A-channel is the Lüders channel $\mathcal{L}_\mathsf{A}$ of A, which is defined as $\mathcal{L}_\mathsf{A}(\varrho) = \sum_x \sqrt{\mathsf{A}(x)}\varrho\sqrt{\mathsf{A}(x)}$.

By a *sequential measurement* we mean a setting where two or more measurements are performed on the same system, one after the other. We will concentrate on the case of two measurements. The first measurement must be described as an observable-channel pair (A,Λ), while for the second measurement it is enough to specify just the observable B since we do not examine the output state after the second measurement. If the initial state of the system is ϱ, then the obtained measurement outcome distributions are $\text{tr}[\varrho\mathsf{A}(x)]$ and $\text{tr}[\Lambda(\varrho)\mathsf{B}(y)]$. We are typically interested in the properties of the input state ϱ rather than the output state $\Lambda(\varrho)$; hence it is convenient to use the Heisenberg picture and write the second measurement outcome distribution as $\text{tr}[\varrho\Lambda^*(\mathsf{B}(y))]$, where Λ^* is the adjoint action of Λ on the set of observables.[2] In essence, a sequential measurement of A and B gives measurement outcomes of A and $\Lambda^*(\mathsf{B})$ on the input state ϱ.

A concept related to sequential measurements is that of a joint measurement. Two observables A and B are *jointly measurable* if there exists an observable M having the product set $\Omega_\mathsf{A} \times \Omega_\mathsf{B}$ as the set of measurement outcomes and satisfying

$$\mathsf{A}(x) = \sum_y \mathsf{M}(x,y)\,, \quad \mathsf{B}(y) = \sum_x \mathsf{M}(x,y) \tag{2}$$

for all $x \in \Omega_\mathsf{A}$ and $y \in \Omega_\mathsf{B}$. Any observable M satisfying (2) is called a *joint observable* of A and B. This definition easily extends to any finite number of observables; $\mathsf{A},\mathsf{B},\mathsf{C},\ldots$, are jointly measurable if there is a single observable M whose marginals coincide with $\mathsf{A},\mathsf{B},\mathsf{C},\ldots$.

[2]If Λ is a channel written in the Schrödinger picture, then Λ^* is defined by the formula $\text{tr}[\Lambda(\varrho)T] = \text{tr}[\varrho\Lambda^*(T)]$, required to hold for all states ϱ and operators T.

A sequential measurement of two quantum observables can be seen as a special type of joint measurement. Formally, if we have maps Φ_x satisfying (1), then we can define $\mathsf{M}(x,y) = \Phi_x^*(\mathsf{B}(y))$. This is a joint observable of A and the perturbed version $\Lambda^*(\mathsf{B})$ of B. At first sight, joint measurement is a broader concept than sequential measurement; a joint measurement is any type of measurement from which one can extract the desired probability distributions of measurement outcomes, whereas in a sequential measurement one has to measure two observables, one after the other. Hence, an immediate question arises: Does the sequential method of measuring two quantum observables suffer from limitations specific to it, or can one perform all possible joint measurements in this way? We show that there are no additional limitations, and there is even a surprising advantage in certain kinds sequential measurements, which we will call *universality*.

III. SHARP OBSERVABLES

As a warm up, let us assume that A is a *sharp observable*, i.e., A is a POVM and each operator $\mathsf{A}(x)$ is a projection. If we perform a standard von Neumann measurement of A, then the state transformation is described by the Lüders channel of A. Supposing that the subsequently measured observable is B, then the actually implemented perturbed version is given by $y \mapsto \sum_x \mathsf{A}(x)\mathsf{B}(y)\mathsf{A}(x)$. We see that if A and B commute [i.e., $\mathsf{A}(x)\mathsf{B}(y) = \mathsf{B}(y)\mathsf{A}(x)$ for all x,y], then $\sum_x \mathsf{A}(x)\mathsf{B}(y)\mathsf{A}(x) = \mathsf{B}(y)$ and this sequential measurement is a joint measurement of A and B. On the other hand, it is well known that a sharp observable is jointly measurable with another observable if and only if they commute; see, e.g., [15]. We conclude that a joint measurement of a sharp observable A and some other observable B can always be implemented as a sequential measurement of A and B.

The previously described case is a particular instance of the class of measurement schemes where the first measurement does not disturb the second one at all; the *nondisturbance condition* requires that

$$\text{tr}[\Lambda(\varrho)\mathsf{B}(y)] = \text{tr}[\varrho\mathsf{B}(y)] \tag{3}$$

for all input states ϱ and outcomes y. This condition means that the measurement outcome probabilities for B are the same for all pairs of an input state ϱ and the corresponding output state $\Lambda(\varrho)$. Obviously, if the nondisturbance condition holds, then a sequential measurement described by (A,Λ) and B implements a joint measurement of A and B even if A is not sharp. The condition (3) holds, for instance, if A and B commute and we choose Λ to be the Lüders channel of A. However, for observables that are not sharp, the condition (3) may be fulfilled for some A-channel even if A and B do not commute, and joint measurability may hold even if there is no nondisturbing measurement at all [16]. A special feature of sharp observables is the equivalence of commutativity, nondisturbance, and joint measurability.

IV. CHANNELS HAVING THE UNIVERSAL PROPERTY

In the general case, the first measurement disturbs the second one, but a joint measurement may still be possible. In practice, this means that to obtain a joint measurement of

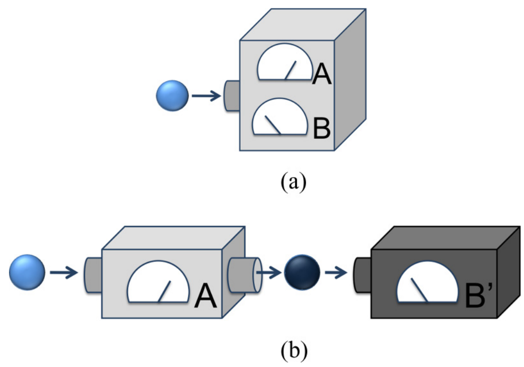

FIG. 2. (Color online) (a) A joint measurement of observables A and B gives measurement outcomes for both A and B. (b) Sequential measurement of A and a modified observable B' is equivalent to the joint measurement if the modification properly compensates for the disturbance that the first measurement causes.

some observables A and B, we may need to measure first A and then B', which is a modified version of B. The modification is aimed to compensate the disturbance that the first measurement causes; see Fig. 2. We are thus looking for an A-channel Λ and an observable B' such that

$$\mathrm{tr}[\Lambda(\varrho)\mathsf{B}'(y)] = \mathrm{tr}[\varrho\mathsf{B}(y)] \tag{4}$$

for all input states ϱ and outcomes y, or equivalently in the Heisenberg picture, $\Lambda^*(\mathsf{B}'(y)) = \mathsf{B}(y)$ for all outcomes y.

This equation can be interpreted in two ways. First, we may measure some observable B' after A, typically as sharp, informative or good as possible, and then (4) determines the actually implemented observable B. Second, we may want to obtain exactly B as the second observable, in which case we can try to tailor Λ and B' in such a way that B is acquired.

As a preliminary step towards our main result, we recall a simple construction which shows that the required objects Λ and B' exist whenever A and B are jointly measurable [16]. Namely, suppose that A and B are jointly measurable; hence there exists M such that $\mathsf{A}(x) = \sum_y \mathsf{M}(x,y)$ and $\mathsf{B}(y) = \sum_x \mathsf{M}(x,y)$. We define a channel Λ^{B} with an output Hilbert space $\mathcal{H}_{\mathrm{out}} = \mathbb{C}^{|\Omega_{\mathsf{B}}|}$ as

$$\Lambda^{\mathsf{B}}(\varrho) = \sum_y \mathrm{tr}[\varrho\mathsf{B}(y)]|y\rangle\langle y|, \tag{5}$$

where $\{|y\rangle\}_{y\in\Omega_{\mathsf{B}}}$ is an orthonormal basis of $\mathcal{H}_{\mathrm{out}}$. Since

$$\Lambda^{\mathsf{B}}(\varrho) = \sum_x \left(\sum_y \mathrm{tr}[\varrho\mathsf{M}(x,y)]|y\rangle\langle y|\right) \tag{6}$$

and

$$\mathrm{tr}\left[\sum_y \mathrm{tr}[\varrho\mathsf{M}(x,y)]|y\rangle\langle y|\right] = \mathrm{tr}[\varrho\mathsf{A}(x)] \tag{7}$$

we conclude that Λ^{B} is an A-channel. Moreover,

$$\mathrm{tr}[\Lambda^{\mathsf{B}}(\varrho)|y\rangle\langle y|] = \mathrm{tr}[\varrho\mathsf{B}(y)] \tag{8}$$

so that (4) holds for the choices $\mathsf{B}'(y) = |y\rangle\langle y|$ and $\Lambda = \Lambda^{\mathsf{B}}$. In conclusion, this sequential measurement scheme implements a joint measurement of A and B.

The previously defined sequential measurement is not very useful since the applied A-channel Λ^{B} is designed specifically for B. This weakness becomes clear when we consider a collection of observables that are pairwisely jointly measurable without being jointly measurable as a whole [15,17]: Suppose A, B, and C are such. Hence, A is pairwisely jointly measurable with both B and C, but the triplet $(\mathsf{A},\mathsf{B},\mathsf{C})$ is not jointly measurable. (A simple example of this kind of triplet is formed when the usual x,y,z-spin-component observables of a spin-$\frac{1}{2}$ particle are mixed with uniform noise with a mixing parameter t chosen from the interval $1/\sqrt{3} < t \leqslant 1/\sqrt{2}$ [15]. Interestingly, one can also find a different triplet that violates a Bell inequality [11].) Suppose further that our task is to measure either the pair (A,B) or (A,C), but we will be told the desired pair only after we have performed the measurement of A. Since the triplet $(\mathsf{A},\mathsf{B},\mathsf{C})$ is not jointly measurable, it is not clear how to succeed in this task. In particular, the sequential measurement scheme related to Λ^{B} does not work since in that case the first measurement has to be chosen according to the second one, and one can see that for any observable C' on $\mathcal{H}_{\mathrm{out}}$, we get $(\Lambda^{\mathsf{B}})^*(\mathsf{C}'(z)) = \sum_y \langle y|\mathsf{C}'(z)|y\rangle\mathsf{B}(y)$, which is just a smearing of B.

To be able to overcome this drawback, we need an A-channel Λ that satisfies the criterion (4) for both B and C with some modified versions B' and C' on the left-hand side, respectively. In the best case we would have an A-channel Λ that satisfies the sequential measurement criterion (4) for all observables that are jointly measurable with A. This motivates the following definition.

Definition 1. An A-channel Λ has the *universal property* (relative to A) if for each observable B jointly measurable with A, there exists an observable B' such that

$$\mathrm{tr}[\Lambda(\varrho)\mathsf{B}'(y)] = \mathrm{tr}[\varrho\mathsf{B}(y)] \tag{9}$$

for all input states ϱ and outcomes $y \in \Omega_{\mathsf{B}}$.

The universal property means that the measurement of an observable A limits the future measurements no more than is necessary, thus putting no additional limitations (i.e., other than joint measurability) on the measurements that can be implemented later. Our main result states that these kinds of measurements exist.

Theorem 1. For every observable A, there exists an A-channel Λ_{A} having the universal property.

Before the proof of Theorem 1, we recall that each observable A has a *Naimark dilation*, i.e., a triplet $(\mathcal{K},\hat{\mathsf{A}},V)$ where $\mathcal{K}$ is a Hilbert space, $V:\mathcal{H}_{\mathrm{in}} \to \mathcal{K}$ is an isometry, and $\hat{\mathsf{A}}$ is a sharp observable on $\mathcal{K}$ satisfying $V^*\hat{\mathsf{A}}(x)V = \mathsf{A}(x)$ for each $x \in \Omega_{\mathsf{A}}$. Moreover, there exists a *minimal* Naimark dilation, meaning that the set $\{\sum_x c_x\hat{\mathsf{A}}(x)V\psi : c_x \in \mathbb{C}, \psi \in \mathcal{H}\}$ is dense in $\mathcal{K}$. The minimal Naimark dilation is essentially unique in the sense that if $(\mathcal{K}_1,\hat{\mathsf{A}}_1,V_1)$ is a minimal Naimark dilation and $(\mathcal{K}_2,\hat{\mathsf{A}}_2,V_2)$ is any other Naimark dilation of A, then there exists an isometry $J:\mathcal{K}_1 \to \mathcal{K}_2$ satisfying $J\hat{\mathsf{A}}_1(x) = \hat{\mathsf{A}}_2(x)J$ and $JV_1 = V_2$. In particular, J is a unitary operator if both Naimark dilations are minimal.

Proof. Fix a Naimark dilation $(\mathcal{K},\hat{\mathsf{A}},V)$ of A, i.e., $\mathcal{K}$ is a Hilbert space, $V:\mathcal{H}_{\mathrm{in}} \to \mathcal{K}$ is an isometry, and $\hat{\mathsf{A}}$ is a sharp observable on $\mathcal{K}$ satisfying $V^*\hat{\mathsf{A}}(x)V = \mathsf{A}(x)$ for each $x \in \Omega_{\mathsf{A}}$. Assume that this dilation is minimal. We define a channel Λ_{A} with input and output Hilbert spaces $\mathcal{H}_{\rm in}$ and $\mathcal{K}$, respectively, by

$$\Lambda_{\mathsf{A}}(\varrho) = \sum_x \hat{\mathsf{A}}(x) V \varrho V^* \hat{\mathsf{A}}(x). \tag{10}$$

This is an A-channel since

$$\mathrm{tr}[\hat{\mathsf{A}}(x) V \varrho V^* \hat{\mathsf{A}}(x)] = \mathrm{tr}[\varrho V^* \hat{\mathsf{A}}(x) V] = \mathrm{tr}[\varrho \mathsf{A}(x)].$$

We claim that Λ_{A} has the universal property.

First, we recall the preliminary result demonstrated earlier: Each B that is jointly measurable with A can be written as $\mathsf{B}(y) = (\Lambda^{\mathsf{B}})^*(|y\rangle\langle y|)$, where $\{|y\rangle\}_{y\in\Omega_{\mathsf{B}}}$ is an orthonormal basis and Λ^{B} is the channel defined in Eq. (5). We show that there exists a channel Γ^{B} such that

$$\Lambda^{\mathsf{B}} = \Gamma^{\mathsf{B}} \circ \Lambda_{\mathsf{A}}\,, \tag{11}$$

where $\circ$ denotes the composition of two functions. Then, using the Heisenberg form $(\Gamma^{\mathsf{B}})^*$ of Γ^{B}, we define an observable B' as

$$\mathsf{B}'(y) = (\Gamma^{\mathsf{B}})^*(|y\rangle\langle y|)\,. \tag{12}$$

This observable satisfies (9) since

$$\mathrm{tr}[\Lambda_{\mathsf{A}}(\varrho)\mathsf{B}'(y)] \overset{(12)}{=} \mathrm{tr}[\Gamma^{\mathsf{B}}(\Lambda_{\mathsf{A}}(\varrho))|y\rangle\langle y|] \tag{13}$$

$$\overset{(11)}{=} \mathrm{tr}[\Lambda^{\mathsf{B}}(\varrho)|y\rangle\langle y|] \overset{(8)}{=} \mathrm{tr}[\varrho \mathsf{B}(y)]. \tag{14}$$

Hence, to complete the proof we need to show that for each B that is jointly measurable with A, there exists a channel Γ^{B} such that (11) holds. Let M be a joint observable of A and B, and let $(\mathcal{K}',\hat{\mathsf{M}},V')$ be a Naimark dilation of M. We define a sharp observable $\hat{\mathsf{A}}'$ as $\hat{\mathsf{A}}'(x) = \sum_y \hat{\mathsf{M}}(x,y)$. Since $\mathsf{A}(x) = \sum_y \mathsf{M}(x,y)$, we observe that $(\mathcal{K}',\hat{\mathsf{A}}',V')$ is a Naimark dilation of A. In fact $V'^*\hat{\mathsf{A}}'(x)V' = V'^*\sum_y \hat{\mathsf{M}}(x,y)V' = \sum_y \mathsf{M}(x,y) = \mathsf{A}(x)$ holds. The initially fixed Naimark dilation $(\mathcal{K},\hat{\mathsf{A}},V)$ was chosen to be minimal, so there exists an isometry $J:\mathcal{K}\to\mathcal{K}'$ satisfying $\hat{\mathsf{A}}'(x)J = J\hat{\mathsf{A}}(x)$ and $V' = JV$ due to the uniqueness of the minimal dilation. Taking also into account that $\hat{\mathsf{M}}(x,y)\hat{\mathsf{M}}(x',y) = \delta_{xx'}\hat{\mathsf{M}}(x,y)$, we obtain

$$\hat{\mathsf{M}}(x,y)J\hat{\mathsf{A}}(x') = \delta_{xx'}\hat{\mathsf{M}}(x,y)J\,. \tag{15}$$

Using (15) we can write Λ^{B} in the form

$$\Lambda^{\mathsf{B}}(\varrho) = \sum_{x,y} \mathrm{tr}[\hat{\mathsf{A}}(x)V\varrho V^*\hat{\mathsf{A}}(x)J^*\hat{\mathsf{M}}(x,y)J]|y\rangle\langle y|.$$

We define a channel Γ^{B} as

$$\Gamma^{\mathsf{B}}(\varrho) = \sum_{x,y} \mathrm{tr}[\varrho J^*\hat{\mathsf{M}}(x,y)J]|y\rangle\langle y|\,.$$

Using (15) again one can confirm that the required equation $\Lambda^{\mathsf{B}} = \Gamma^{\mathsf{B}}\circ\Lambda_{\mathsf{A}}$ holds. ∎

The channel Λ_{A} was introduced as the least disturbing A-channel in Ref. [18]. In the context of sequential measurements, the channel Λ_{A} is the appropriate generalization of the Lüders channel of sharp observables; Λ_{A} has the universal property and it reduces to the Lüders channel whenever A is a sharp observable. The price we have to pay is the larger output space $\mathcal{H}_{\rm out}$ compared to the input space $\mathcal{H}_{\rm in}$.

V. CHANNELS WITHOUT THE UNIVERSAL PROPERTY

Suppose two jointly measurable observables A and B are given. We may have limited resources, so perhaps we cannot realize any A-channel having the universal property. Hence, suppose we first perform a measurement described by a pair (A,Λ), where Λ is an A-channel *not* having the universal property. We want to know if we can still implement a measurement of B, i.e., whether there exists an observable B' such that

$$\mathrm{tr}[\Lambda(\varrho)\mathsf{B}'(y)] = \mathrm{tr}[\varrho\mathsf{B}(y)] \tag{16}$$

for all input states ϱ and outcomes $y\in\Omega_{\mathsf{B}}$.

To give a general answer to this question, we recall that by the *Stinespring dilation theorem* (see, e.g., [19]) any channel Λ can be written in the form

$$\Lambda(\varrho) = \mathrm{tr}_{\mathcal{K}}[V\varrho V^*]\,, \tag{17}$$

where $\mathcal{K}$ is a Hilbert space attached to an environment system, $V:\mathcal{H}_{\rm in}\to\mathcal{H}_{\rm out}\otimes\mathcal{K}$ is an isometry and $\mathrm{tr}_{\mathcal{K}}[\cdot]$ is the partial trace over $\mathcal{K}$. Tracing over $\mathcal{H}_{\rm out}$ rather than $\mathcal{K}$ we obtain the corresponding *conjugate channel* (also called the complementary channel)

$$\bar{\Lambda}(\varrho) = \mathrm{tr}_{\mathcal{H}_{\rm out}}[V\varrho V^*]\,. \tag{18}$$

The answer to the previous question can now be stated in a concise form.

Theorem 2. For a channel Λ and observable B, there exists an observable B' satisfying (16) if and only if the conjugate channel $\bar{\Lambda}$ of Λ is a B-channel.

Proof. The Stinespring dilation of a channel Λ, written in the Heisenberg picture, reads

$$\Lambda^*(T) = V^*(T\otimes\mathbb{1})V\,,$$

and the corresponding conjugate channel is then

$$\bar{\Lambda}^*(S) = V^*(\mathbb{1}\otimes S)V.$$

As explained in Ref. [20], it follows from the Radon-Nikodym theorem for quantum operations [21,22] that $\bar{\Lambda}^*$ is a B-channel if and only if there exists an observable B' such that $V^*(\mathsf{B}'(y)\otimes\mathbb{1})V = \mathsf{B}(y)$. Inserting this into the first equation gives $\Lambda^*(\mathsf{B}'(y)) = \mathsf{B}(y)$. ∎

The formulation of Theorem 2 is slightly loose since the existence of B' may seem to depend on the choice of the conjugate channel. However, all conjugate channels of a given channel Λ are equivalent in the sense that each of them can be obtained from any other by concatenating with some other channel. This implies that if one conjugate channel of Λ is a B-channel, then all conjugate channels of Λ are B-channels. Therefore, it does not matter which conjugate channel we use in Theorem 2. For completeness we provide an argument that all conjugate channels are equivalent.

Let Λ^* be a channel, written in the Heisenberg picture. It has a minimal Stinespring dilation

$$\Lambda^*(T) = V^*(T\otimes\mathbb{1})V\,,$$

where $V:\mathcal{H}_{\rm in}\to\mathcal{H}_{\rm out}\otimes\mathcal{K}$ is an isometry. Suppose we have another Stinespring dilation

$$\Lambda^*(T) = V_1^*(T\otimes\mathbb{1})V_1\,,$$

where $V_1 : \mathcal{H}_{\text{in}} \to \mathcal{H}_{\text{out}} \otimes \mathcal{K}_1$ is an isometry. As the first Stinespring dilation is minimal, there is an isometry $W : \mathcal{K} \to \mathcal{K}_1$ such that $V_1 = (\mathbb{1} \otimes W)V$. We consider the conjugate channels that are determined by these two dilations, i.e.,

$$\bar{\Lambda}^*(S) = V^*(\mathbb{1} \otimes S)V\,, \quad \bar{\Lambda}_1^*(S) = V_1^*(\mathbb{1} \otimes S)V_1.$$

Inserting $V_1 = (\mathbb{1} \otimes W)V$ we obtain $\bar{\Lambda}_1^*(S) = \bar{\Lambda}^*(W^*SW)$, and the map $S \mapsto W^*SW$ is a quantum channel. In a similar way, we obtain $\bar{\Lambda}^*(S) = \bar{\Lambda}_1^*(WSW^*)$, but the map $S \mapsto WSW^*$ need not be unital and hence not a quantum channel. We define a channel $\mathcal{E}^*$ as

$$\mathcal{E}^*(S) = WSW^* + (\mathbb{1} - WW^*)\text{tr}[\rho_0 S], \tag{19}$$

where ρ_0 is a fixed state. Then $\bar{\Lambda}^*(S) = \bar{\Lambda}_1^*(\mathcal{E}(S))$. As a conclusion, $\bar{\Lambda}$ and $\bar{\Lambda}_1$ are equivalent. Therefore, every channel has a unique conjugate channel up to equivalence.

VI. NECESSITY OF LARGER OUTPUT SPACE

In our investigation we have allowed the output space $\mathcal{H}_{\text{out}}$ to be different than the input space $\mathcal{H}_{\text{in}}$. This means that some part of the input system may be discarded or part of the environment may be included in the description of the output system. This freedom is essential in the definition of Λ_{A}, and therefore also in Theorem 1. However, one may wonder if an observable A can have some compatible channel with the universal property for $\mathcal{H}_{\text{out}} = \mathcal{H}_{\text{in}}$. In fact, as noted earlier, this is true for the Lüders channel of a sharp observable. In the following two examples we demonstrate that, first, the Lüders channel is not universal even for some common observables, and second, it is often necessary to have $\mathcal{H}_{\text{out}}$ larger than $\mathcal{H}_{\text{in}}$ to reach the universal property. Physically, the larger output space makes the subsequent measurement able to take advantage also of the "information leaked to the environment," thereby leading to the universality property.

Example 1. Let us fix $\mathcal{H}_{\text{in}} = \mathbb{C}^2$ and consider two families of binary qubit observables A_s and $\mathsf{B}_{t,\theta}$, where

$$\mathsf{A}_s(\pm 1) = \tfrac{1}{2}(\mathbb{1} \pm s\sigma_z),$$
$$\mathsf{B}_{t,\theta}(\pm 1) = \tfrac{1}{2}[\mathbb{1} \pm t(\sin\theta\sigma_x + \cos\theta\sigma_z)],$$

and the parameters belong to the intervals $s,t \in (0,1]$ and $\theta \in [0,\pi/2]$, respectively. As proved in Ref. [23], A_s and $\mathsf{B}_{t,\theta}$ are jointly measurable if and only if

$$s^2 + t^2 - \cos^2\theta s^2 t^2 \leqslant 1.$$

First, we want to see how the pairs satisfying this inequality can be implemented sequentially.

There exists an A_s-channel satisfying the nondisturbance condition for $\mathsf{B}_{t,\theta}$ if and only if A_s and $\mathsf{B}_{t,\theta}$ commute [16, Prop. 6], which is the case when $\theta \in \{0,\pi/2\}$. Therefore, most of the realizable joint measurements must be performed by first measuring A_s, followed by some modified version B' of $\mathsf{B}_{t,\theta}$. The modified observable B' and a suitable A_s-channel are not difficult to find in this simple case; we can choose Λ to be the Lüders channel of A_s and $\mathsf{B}' = \mathsf{B}_{1,\theta}$. With these choices we have $\mathcal{L}^*_{\mathsf{A}_s}(\mathsf{B}') = \mathsf{B}_{t,\theta}$, hence the Lüders channel can be used to measure all binary observables jointly measurable with A_s.

Let us then consider another pair of observables to demonstrate that the Lüders channel of A_s does not have the universal property. Fix $0 < s < 1$ and let C be the following four outcome observable:

$$\mathsf{C}(1,\pm 1) = \frac{1 \pm s}{4}(1 \pm \sigma_z), \quad \mathsf{C}(-1,1) = \frac{1-s}{4}(1+\sigma_x)$$
$$\mathsf{C}(-1,-1) = \frac{1-s}{4}(1-\sigma_x) + \frac{s}{2}(1-\sigma_z).$$

We have $\mathsf{C}(1,1) + \mathsf{C}(1,-1) = \mathsf{A}_s(1)$ and $\mathsf{C}(-1,1) + \mathsf{C}(-1,-1) = \mathsf{A}_s(-1)$, hence C and A_s are jointly measurable. One can utilize Theorem 2 to see that there is no C' such that

$$\mathcal{L}^*_{\mathsf{A}}(\mathsf{C}'(j,k)) = \mathsf{C}(j,k) \quad \text{for all j,k} = \pm 1.$$

We can also confirm this fact directly; let us make a counter assumption that there exists an observable C' satisfying the previous equation. Since the operator $\mathsf{C}(-1,1)$ is rank 1 and the operators $\mathsf{A}_s(\pm 1)$ are invertible, it follows that $\mathsf{C}'(-1,1)$ is rank 1 as well. As Λ is trace preserving, we conclude that $\Lambda(P) = \frac{1}{2}(1+\sigma_x)$ must hold for some one-dimensional projection P. A direct calculation shows that $\Lambda(P)$ is a projection if and only if $P = \frac{1}{2}(1 \pm \sigma_z)$, in which case we have $\Lambda(P) = P$. Therefore, $\Lambda(P) = \frac{1}{2}(1+\sigma_x)$ cannot be satisfied by any projection P. As a conclusion, the Lüders channel of A_s does not have the universal property.

Example 2. Let us fix $\mathcal{H}_{\text{in}} = \mathbb{C}^d$ and consider an N-outcome observable A whose elements are all rank 1, that is, each $\mathsf{A}(n)$ is represented as $\mathsf{A}(n) = a_n|\psi_n\rangle\langle\psi_n|$ for some unit vector $\psi_n \in \mathcal{H}_{\text{in}}$ and coefficient $0 < a_n \leqslant 1$. We further assume that no two vectors ψ_n and ψ_m are parallel, so that the operators $\mathsf{A}(n)$ and $\mathsf{A}(m)$ are linearly independent. Let Λ be an A-channel having the universal property. We denote its Stinespring representation in the Heisenberg picture by $\Lambda^*(T) = V^*(T \otimes \mathbb{1})V$. According to the Radon-Nikodym theorem [21,22], there exists an observable $\mathsf{B}(n)$ on an auxiliary Hilbert space satisfying

$$V^*(\mathbb{1} \otimes \mathsf{B}(n))V = \mathsf{A}(n).$$

As A itself is jointly measurable with A, Theorem 2 implies that there exists an observable C satisfying

$$V^*(\mathsf{C}(m) \otimes \mathbb{1})V = \mathsf{A}(m).$$

Since

$$V^*(\mathsf{C}(m) \otimes \mathsf{B}(n))V \leqslant V^*(\mathsf{C}(m) \otimes \mathbb{1})V = \mathsf{A}(m)$$
$$V^*(\mathsf{C}(m) \otimes \mathsf{B}(n))V \leqslant V^*(\mathbb{1} \otimes \mathsf{B}(n))V = \mathsf{B}(n)\,,$$

the rank 1 property of A implies that there exists a family of numbers $\{c_n\}$ satisfying $0 \leqslant c_n \leqslant 1$ and

$$V^*(\mathsf{C}(m) \otimes \mathsf{B}(n))V = \delta_{mn} c_n \mathsf{A}(n)\,.$$

Taking its summation over m and n, we conclude that $\sum_n c_n \mathsf{A}(n) = \mathbb{1}$ and thus $c_n = 1$. Then they satisfy $\langle\psi_n|V^*(\mathsf{C}(m) \otimes \mathsf{B}(n))V|\psi_n\rangle = \delta_{mn}$. This is possible only when $\dim \mathcal{H}_{\text{out}} \geqslant N$.

VII. CONCLUSION

Sequential measurements can be seen as special type of joint measurements. In this sense, they suffer from the same limitations as joint measurements; two incompatible measurements cannot be implemented sequentially as the first measurement disturbs the second one. We have shown that there are no other limitations in the sense that any joint measurement can be replaced by a sequential measurement scheme. In addition to that, considering sequential schemes opens up some additional perspectives. In particular, there is a temporal aspect in sequential measurements and one can therefore make a choice after the first measurement has been performed. An interesting open question is whether the present result can be extended to sequential measurements of more than two observables.

ACKNOWLEDGMENTS

The authors are grateful to David Reeb, Daniel Reitzner, and Leon Loveridge for their comments on an earlier version of this paper. T.H. acknowledges the financial support from the Academy of Finland (Grant No. 138135). T.M. thanks JSPS for the financial support (JSPS KAKENHI Grant No. 22740078).

[1] J. Bergou, E. Feldman, and M. Hillery, Phys. Rev. Lett. **111**, 100501 (2013).
[2] C. Carmeli, T. Heinosaari, and A. Toigo, Phys. Rev. A **85**, 012109 (2012).
[3] A. H. Werner, T. Franz, and R. F. Werner, Phys. Rev. Lett. **103**, 220504 (2009).
[4] T. Miyadera, Int. J. Quantum. Inform. **9**, 1427 (2011).
[5] J. Eisert, M. P. Müller, and C. Gogolin, Phys. Rev. Lett. **108**, 260501 (2012).
[6] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. A **85**, 012302 (2012).
[7] M. Wilde, Proc. R. Soc. A **469**, 20130259 (2013).
[8] J.-Å. Larsson, M. Kleinmann, O. Gühne, and A. Cabello, *Violating noncontextual realism through sequential measurements*, in AIP Conf. Proc. 1327, edited by G. Jaegger, A. Y. Khrennikov, M. Schlosshauer, and G. Weihs (American Institute of Physics, New York, 2011), pp. 401–409.
[9] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).
[10] R. Uola, T. Moroder, and O. Gühne, Phys. Rev. Lett. **113**, 160403 (2014).
[11] M. T. Quintino, T. Vértesi, and N. Brunner, Phys. Rev. Lett. **113**, 160402 (2014).
[12] E. Davies and J. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[13] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory* (Cambridge University Press, Cambridge, England, 2012).
[14] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
[15] T. Heinosaari, D. Reitzner, and P. Stano, Found. Phys. **38**, 1133 (2008).
[16] T. Heinosaari and M. Wolf, J. Math. Phys. **51**, 092201 (2010).
[17] R. Kunjwal, C. Heunen, and T. Fritz, Phys. Rev. A **89**, 052126 (2014).
[18] T. Heinosaari and T. Miyadera, Phys. Rev. A **88**, 042117 (2013).
[19] V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, Cambridge, England, 2003).
[20] T. Heinosaari, T. Miyadera, and D. Reitzner, Found. Phys. **44**, 34 (2014).
[21] W. Arveson, Acta Math. **123**, 141 (1969).
[22] M. Raginsky, J. Math. Phys. **44**, 5003 (2003).
[23] P. Busch, Phys. Rev. D **33**, 2253 (1986).