Colloquium: Incompatible measurements in quantum information science

Otfried Gühne

Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany

Erkka Haapasalo

Centre for Quantum Technologies, National University of Singapore, Science Drive 2, Block S15-03-18, Singapore 117543, Singapore

Tristan Kraft

Institute for Theoretical Physics, University of Innsbruck, Technikerstraße 21A, 6020 Innsbruck, Austria and Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany

Juha-Pekka Pellonpää

Department of Physics and Astronomy, University of Turku, FI-20014 Turun yliopisto, Finland

Roope Uola

Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland

(published 6 February 2023; corrected 7 April 2023)

Some measurements in quantum mechanics disturb each other. This has puzzled physicists since the formulation of the theory, but only in recent decades has the incompatibility of measurements been analyzed in depth and detail, using the notion of joint measurability of generalized measurements. In this Colloquium joint measurability and incompatibility are reviewed from the perspective of quantum information science. The Colloquium starts by discussing the basic definitions and concepts. An overview on applications of incompatibility, such as in measurement uncertainty relations, the characterization of quantum correlations, or information processing tasks like quantum state discrimination, is then presented. Finally, emerging directions of research, such as a resource theory of incompatibility as well as other concepts to grasp the nature of measurements in quantum mechanics, are discussed.

DOI: 10.1103/RevModPhys.95.011003

CONTENTS

I. INTRODUCTION

Measurements in quantum mechanics are different than their classical counterparts. From today's perspective this statement may sometimes seem to be a truism or platitude, but when quantum theory was developed the notion of measurements and their relation to physical quantities was indeed a major roadblock on the way to a better understanding. In 1925, Heisenberg noted that the product of physical quantities in the theory of atoms may depend on their order (Heisenberg, 1925). Directly thereafter, Born and Jordan (1925) pointed out that the fundamental reason for this is that physical quantities in quantum mechanics are described by matrices. Matrix calculus was not common knowledge to physicists in those times, so Born and Jordan found it important to point out directly at the beginning of their paper that for two general matrices A and B

$$AB \neq BA \tag{1}$$

holds. But what is the physical relevance of this noncommutativity?

In the following years, the fact that two observables do not share common eigenstates attracted attention in the form of uncertainty relations (Heisenberg, 1927; Kennard, 1927; Robertson, 1929, 1934). Here the noncommutativity directly plays a role, as in the Robertson relation

$$\Delta(A)\Delta(B) \geq \tfrac{1}{2}|\langle\psi|[A,B]|\psi\rangle|, \tag{2}$$

where $\Delta(A)$ denotes the standard deviation of the observable A and $[A,B] = AB - BA$ is the commutator. Owing to such relations, noncommutativity of observables is sometimes seen as a key phenomenon in quantum mechanics that already contains most of the mysteries of quantum measurements.

It turned out, however, that the notion of observables or Hermitian matrices is much too narrow to describe all measurements in quantum mechanics (Davies, 1976; Helstrom, 1976; Holevo, 1982; Ludwig, 1983; Prugovecki, 1992; Busch *et al.*, 2016). Indeed, the textbook notion of projective measurements can be extended to positive-operator-valued measures (POVMs) [a short historical review was given by Ali *et al.* (2009)]. POVMs can have more outcomes and may be seen as measurements carried out with the help of an additional quantum system. They are at the core of the modern formulation of operational quantum mechanics and provide an advantage in fundamental protocols of quantum physics, such as the discrimination of quantum states. POVMs are the most general description of the outcome statistics of measurements, but if the postmeasurement state is taken into account, one needs to further generalize them and consider so-called quantum instruments, which were introduced by Davies and Lewis (1970).

But what is the extension of the notion of noncommutativity to POVMs? Here several notions have been introduced, but their relation has often not been clear and a direct physical interpretation has been missing. In recent years, however, the situation has changed. The notion of *joint measurability* of POVMs has turned out to be fundamentally related to several other phenomena in quantum mechanics and quantum information theory. Joint measurability is related to measurement uncertainty relations as well as preparation noncontextuality. Moreover, *incompatibility* (i.e., the absence of joint measurability) is essential for the creation and exploitation of quantum correlations, such as in the form of quantum steering.

In this Colloquium, we give an overview on joint measurability from the perspective of quantum information theory. Starting with the basic definitions and properties of joint measurability and related concepts, we discuss their applications, such as in Bell nonlocality or protocols in quantum information processing. Our aim is to present these concepts in simple language in order to provide an introduction for researchers from different backgrounds.

We note that there are several noteworthy works that cover parts of the theory presented in our Colloquium. For instance, incompatibility from an operational point of view was discussed by Heinosaari, Miyadera, and Ziman (2016), and quantum measurement theory from a mathematical perspective was developed in depth by Busch *et al.* (2016). In addition, joint measurability is connected to several topics of quantum information theory, and interested readers can find several detailed overview articles about them. These topics include quantum correlations like Bell nonlocality (Brunner *et al.*, 2014) and quantum steering (Cavalcanti and Skrzypczyk, 2017; Uola, Costa *et al.*, 2020), and phenomena and applications like uncertainty relations (Busch, Lahti, and Werner, 2014a), quantum contextuality (Liang, Spekkens, and Wiseman, 2011; Budroni *et al.*, 2022), and quantum state discrimination (Barnett and Croke, 2009).

This Colloquium is structured as follows. In Sec. II we introduce and explain basic concepts to describe measurements. These include the central notion of POVMs and their joint measurability, as well as concepts like instruments and the disturbance of measurements. Section III discusses important results on joint measurability. We start with analytical results for qubit systems and then discuss measures of incompatibility, including their numerical evaluation via semidefinite programming and constructive methods to obtain joint measurements. Section IV connects joint measurability with different concepts in quantum information processing. We describe intimate connections to various forms of quantum correlations, to foundational effects such as contextuality and macrorealism, and to information processing tasks like state discrimination or random access codes. Finally, Sec. V collects various extensions of the previously discussed concepts, including resource theory aspects, other notions for accessing the nonclassical behavior of quantum measurements, and the incompatibility of quantum channels.

II. CONCEPTS

Throughout this Colloquium, we use the measurement theoretical formulation of quantum measurements; see Busch *et al.* (2016). In this formulation, Hermitian operators are generalized to positive-operator-valued measures, and the state updates caused by measurements are described by quantum instruments. The latter are objects generalizing the projection postulate of the Hermitian formulation. This corresponds to the most general model for quantum measurements and, of relevance to this Colloquium, manages to describe the different operational formulations of measurement incompatibility.

A. Measurements and instruments

A POVM is a collection of positive semidefinite matrices $\{A_a\}$ that normalizes to the identity operator, i.e., $\sum_a A_a = \mathbb{1}$. The positivity and normalization requirements correspond to the requirement on the related measurement outcome statistics to form a probability distribution. In a quantum state ϱ, i.e., a positive unit-trace operator, these probabilities are given by $p(a|\varrho) = \mathrm{tr}[\varrho A_a]$, where a is the outcome. Whenever A_a is a projection for all a, i.e., $A_a^2 = A_a$, the POVM is said to be sharp or a projection-valued measure (PVM). The special case of PVMs is in one-to-one correspondence with the Hermitian formulation by the spectral theorem, i.e., any Hermitian operator is of the form $\sum_a a P_a$ for a unique PVM $\{P_a\}$. Although POVMs are more general than PVMs, any POVM can be seen as a PVM on a larger system through the Naimark dilation; see Sec. III.C.3.

When we describe the entire measurement process, we have to take into account how the state changes are conditioned on registering an outcome. If an outcome a is obtained, the nonnormalized postmeasurement state is σ_a, and we assume that the map $\varrho \mapsto \sigma_a$ is a linear (or rather affine) completely positive map and the sum $\sum_a \sigma_a$ is a quantum state. Thus, a measurement is associated with an instrument $\{\mathcal{I}_a\}$ that is a collection of linear completely positive maps such that the sum $\sum_a \mathcal{I}_a$ is a completely positive trace-preserving (CPTP) map, i.e., a quantum channel. It is evident that the projection postulate $\varrho \mapsto P_a \varrho P_a$ for a PVM $\{P_a\}$ is an instance of a quantum instrument. More generally, any instrument associated with a POVM $\{A_a\}$, i.e., any instrument with the property $\mathrm{tr}[\mathcal{I}_a(\varrho)] = \mathrm{tr}[\varrho A_a]$ holding for all states ϱ, is of the form $\mathcal{I}_a(\varrho) = \Lambda_a(\sqrt{A_a}\varrho\sqrt{A_a})$, where Λ_a is an outcome-dependent quantum channel (from the input system to the output system) (Pellonpää, 2013b). As an important special case, we highlight the von Neumann–Lüders instrument $\mathcal{I}_a^{\mathrm{vN\text{-}L}}(\varrho) = \sqrt{A_a}\varrho\sqrt{A_a}$, which is the most direct generalization of the projection postulate and can be seen as the least disturbing implementation of the POVM $\{A_a\}$; see Sec. V.C.5. Quantum instruments have been analyzed intensively in the literature; see (Davies and Lewis (1970), Davies (1976), Cycon and Hellwig (1977), Ozawa (1984), Holevo (1998), Pellonpää (2013a, 2013b), Busch *et al.* (2016), and Haapasalo and Pellonpää (2017a).

B. Joint measurability

There are three natural distinctions between classical and quantum properties of POVMs. They are given by noncommutativity, inherent measurement disturbance, and the impossibility of a simultaneous readout of the outcomes. Of the three, the last one has found the most profound role in quantum information theory, and consequently is our main focus. We start with the general notion of joint measurabilility and discuss the other two as special cases thereof.

The idea of joint measurability is to simulate the statistics of a set of measurements using only one measurement apparatus. This apparatus is described using a POVM $\{G_\lambda\}$, and its statistics in a state ϱ read $p(\lambda|\varrho) = \mathrm{tr}[G_\lambda \varrho]$. The set of measurements that we aim to simulate is described by a set of POVMs $\{A_{a|x}\}$. In this notation, x labels the choice of the POVM and a denotes the corresponding outcome. The simulation is done classically on the level of statistics and is described by classical postprocessings, i.e., conditional probabilities $\{p(a|x,\lambda)\}$. The simulation is successful if $\mathrm{tr}[A_{a|x}\varrho] = \sum_\lambda p(a|x,\lambda)\mathrm{tr}[G_\lambda \varrho]$ holds for any quantum state ϱ.

This leads to the formal definition of joint measurability: A set of POVMs $\{A_{a|x}\}$ is said to be jointly measurable or compatible if there are a POVM $\{G_\lambda\}$ and classical postprocessings, i.e., a set of conditional probabilities $\{p(a|x,\lambda)\}$, such that

$$A_{a|x} = \sum_\lambda p(a|x,\lambda) G_\lambda. \tag{3}$$

In this case the POVM $\{G_\lambda\}$ is called a joint or parent measurement of the set $\{A_{a|x}\}$. Otherwise, the set $\{A_{a|x}\}$ is called not jointly measurable or incompatible.

We note that the previous definition is equivalent to the existence of a POVM $\{M_{\vec{a}}\}$, where $\vec{a} = (a_1, \ldots, a_n)$ is a vector of the outcomes with the subindex referring to the measurement choice x, from which one gets the original POVMs as margins. More formally, one has

$$A_{a|x} = \sum_{\vec{a} \in E_{a|x}} M_{\vec{a}}, \tag{4}$$

where the set $E_{a|x}$ consists of those outcomes $\vec{a}$ of the joint measurement that include the outcome a of the measurement x. As an example, in the case of two POVMs this reduces to $A_{a_1|1} = \sum_{a_2} M_{a_1,a_2}$ for all outcomes a_1 of the first measurement and $A_{a_2|2} = \sum_{a_1} M_{a_1,a_2}$ for all outcomes a_2 of the second measurement; see Fig. 1. In general, the measurement $\{M_{\vec{a}}\}$ can be viewed as a simultaneous readout of all its

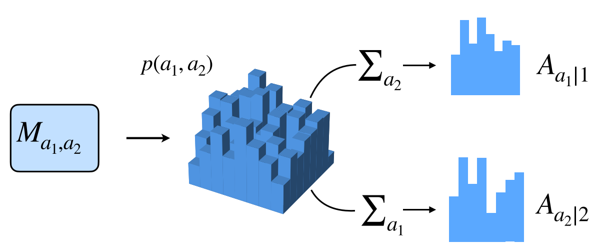

FIG. 1. Marginal form joint measurement of two POVMs. The data of the joint measurement are presented as a grid from which marginalization gives the data of the original pair of measurements.

components in the sense that neglecting the data of all but one component gives the exact statistics of this component POVM. To see the equivalence between Eqs. (3) and (4), one notes that the marginal form is a special case of a general postprocessing. For the other direction, one can set $M_{\vec{a}} := \sum_\lambda [\Pi_x p(a_x|x,\lambda)] G_\lambda$; cf. Ali *et al.* (2009).

The concept of joint measurability is easiest to illustrate with an example. Consider two measurements acting on a qubit system given by the POVM elements $A_{\pm|1}(\mu) = (1/2)(\mathbb{1} \pm \mu\sigma_x)$ and $A_{\pm|2}(\mu) = (1/2)(\mathbb{1} \pm \mu\sigma_z)$. These are noisy versions of the sharp spin measurements along the directions x and z, with the parameter $1-\mu \in [0,1]$ describing the noise. An intuitive way to find a measurement with correct margins is to choose measurement directions that are between x and z (see also Sec. III.C.1), i.e., that define a candidate joint measurement

$$M_{a_1,a_2}(\mu) = (1/4)[\mathbb{1} + \mu(a_1\sigma_x + a_2\sigma_z)]. \tag{5}$$

Equation (5) has the correct margins, and for $\mu \in [0, 1/\sqrt{2}]$ it is a POVM. Hence, the POVMs $A_{\pm|1}(\mu)$ and $A_{\pm|2}(\mu)$ are jointly measurable whenever $\mu \in [0, 1/\sqrt{2}]$. Moreover, whenever $\mu \in (0, 1/\sqrt{2}]$ the POVMs are noncommuting, but a joint measurement nevertheless exists. It can be shown that for $\mu > 1/\sqrt{2}$ the POVMs are not jointly measurable; see Sec. III.A. This is a simple example of a joint measurement, but we note that in general their form can be complex and can require an exponentially increasing number of outcomes (Skrzypczyk *et al.*, 2020).

C. Nondisturbance and commutativity

Joint measurability envelopes another central property of quantum measurements, that is, the possibility of measuring POVMs in a sequence without disturbance. A POVM $\{A_{a_1|1}\}$ is said to be nondisturbing with respect to another POVM $\{A_{a_2|2}\}$ if there is a sequential implementation in which neglecting the outcome of the first measurement $\{A_{a_1|1}\}$ does not affect the statistics of the subsequent measurement $\{A_{a_2|2}\}$. More precisely, one asks for the existence of an instrument $\{\mathcal{I}_{a_1}\}$ associated with $\{A_{a_1|1}\}$ such that $\sum_{a_1} \mathrm{tr}[\mathcal{I}_{a_1}(\varrho)A_{a_2|2}] = \mathrm{tr}[\varrho A_{a_2|2}]$ for all ϱ and a_2. This notion generalizes to more measurements straightforwardly. Nondisturbance implies joint measurability by setting $\mathrm{tr}[M_{a_1,a_2}\varrho] := \mathrm{tr}[\mathcal{I}_{a_1}(\varrho)A_{a_2|2}]$ for all ϱ. There are pairs of jointly measurable POVMs that do not allow for a nondisturbing sequential implementation (Heinosaari and Wolf, 2010). However, jointly measurable pairs can always be measured in a sequence by performing a suitable instrument $\{\mathcal{I}_{a_1}\}$ of $\{A_{a_1|1}\}$ and a retrieving measurement $\{\tilde{A}_{a_2|2}\}$ after it, i.e., $\sum_{a_1} \mathrm{tr}[\mathcal{I}_{a_1}(\varrho)\tilde{A}_{a_2|2}] = \mathrm{tr}[\varrho A_{a_2|2}]$; see Sec. V.C.5 and Heinosaari and Miyadera (2015) and Haapasalo and Pellonpää (2017b). We note that in general the retrieving measurement is different from the original and it can be interpreted in two different ways. Either the measurement is a purely mathematical construction that relies on additional degrees of freedom on a larger Hilbert space or it is a physical one, in which case one uses the output a_1 of the first measurement as an input for the second measurement. These cases are explained in more detail in Sec. V.C.5.

A historically relevant special case of joint measurability is that of commutativity. A set of POVMs $\{A_{a|x}\}$ is said to be commuting if $[A_{a|x}, A_{b|y}] = 0$ for all a, b, and $x \neq y$. Such a set allows a nondisturbing implementation by the use of the von Neumann–Lüders instrument and is jointly measurable with the product POVM $M_{\vec{a}} := A_{a_1|1} \cdots A_{a_n|n}$, where $\vec{a} = (a_1, \ldots, a_n)$. However, the inverse implications do not hold in general: As previously mentioned, there are noncommuting POVMs that allow a joint measurement. Moreover, Heinosaari and Wolf (2010) showed that when the Hilbert space dimension d is equal to 2, nondisturbance reduces to commutativity, but in systems with $d = 3$ this is no longer true.

Although noncommutativity lacks an operational meaning in quantum measurement theory in general, with the exception of two-outcome (also called binary) measurements (Designolle, Uola *et al.*, 2021), it has been central for the development of quantum measurement theory. For example, nondisturbance and joint measurability are equivalent to commutativity in the case of PVMs (Hermitian operators). In addition, a pair of POVMs is jointly measurable if and only if they have a common Naimark dilation in which the projective measurements on the dilation space commute; see Sec. III.C.3.

III. CHARACTERIZING JOINT MEASURABILITY

In this section, we present the basic techniques for characterizing and quantifying incompatibility. First, we discuss analytical criteria for the qubit case and the connection to measurement uncertainty relations. Second, we explain the connections between joint measurability and the optimization method of semidefinite programming. This allows us to introduce quantifiers of incompatibility. Third, we explain general methods to construct parent measurements and discuss connections with the Naimark extension of POVMs, which allows us to formulate some results from a higher perspective. Finally, we discuss algebraic characterizations of specific highly incompatible measurements in arbitrary dimensions.

A. Criteria for joint measurability and measurement uncertainty relations

Given the formal definition of joint measurability, one may ask for analytical criteria to determine whether or not two measurements are jointly measurable. In this section, we focus on analytical criteria for measurements with two outcomes on a single qubit. As it turns out, there is an interesting connection to measurement uncertainty relations.

For the case of qubits, the effects (i.e., positive operators bounded from above by $\mathbb{1}$) of all measurements may, up to normalization, be viewed as vectors on the Bloch sphere. Thus, for a two-outcome measurement $\{A_\pm\}$ we can write

$$A_\pm = \tfrac{1}{2}[(1 \pm \gamma)\mathbb{1} \pm \vec{m} \cdot \vec{\sigma}], \tag{6}$$

with $\vec{m} \cdot \vec{\sigma} = m_x\sigma_x + m_y\sigma_y + m_z\sigma_z$. Here γ is also called the bias of the measurement, while $\|\vec{m}\|$ is called the sharpness (Busch *et al.*, 2016).

The first result on joint measurability of such measurements was obtained by Busch (1986). He considered the case of two dichotomic measurements on a qubit, described by γ_i and $\vec{m}_i$, which are both unbiased ($\gamma_1 = \gamma_2 = 0$). He then showed that these are jointly measurable if and only if

$$\|\vec{m}_1 + \vec{m}_2\| + \|\vec{m}_1 - \vec{m}_2\| \leq 2. \tag{7}$$

Equation (7) has also been used to determine the probability of random measurements to be incompatible (Zhang *et al.*, 2019).

For the case of two potentially biased measurements, this problem was considered at the same time by Stano, Reitzner, and Heinosaari (2008), Busch and Schmidt (2010), and Yu *et al.* (2010). The resulting conditions are mathematically equivalent, but the most compact form was derived by Yu *et al.* (2010). For that, one defines the auxiliary quantities

$$F_i = \tfrac{1}{2}\left[\sqrt{(1+\gamma_i)^2 - \|\vec{m}_i\|^2} + \sqrt{(1-\gamma_i)^2 - \|\vec{m}_i\|^2}\right] \tag{8}$$

for $i = 1, 2$. Thus, the measurements $\{A_{\pm|1}\}$ and $\{A_{\pm|2}\}$ are jointly measurable if and only if

$$(1 - F_1^2 - F_2^2)\left(1 - \frac{\gamma_1^2}{F_1^2} - \frac{\gamma_2^2}{F_2^2}\right) \leq (\vec{m}_1 \cdot \vec{m}_2 - \gamma_1\gamma_2)^2. \tag{9}$$

Finally, several works extended the condition in Eq. (7) to three unbiased measurements. This was first done for unbiased measurements in orthogonal directions (Busch, 1986; Brougham and Andersson, 2007) and for three measurement directions in angles of $2\pi/3$ in a plane (Liang, Spekkens, and Wiseman, 2011). For three general measurements a necessary condition was found by Pal and Ghosh (2011), which was then shown to be sufficient for unbiased measurements by Yu and Oh (2013). It reads as follows: For a set of vectors $\{\vec{v}_k\}$ one defines the Fermat-Torricelli vector $\vec{v}_{\mathrm{FT}}$ as the vector minimizing the sum of the distances $\sum_k \|\vec{v} - \vec{v}_k\|$. Therefore, three unbiased measurements on a qubit are jointly measurable if and only if

$$\sum_{k=0}^{3} \|\vec{T}_k - \vec{T}_{\mathrm{FT}}\| \leq 4, \tag{10}$$

where $\vec{T}_{\mathrm{FT}}$ is the Fermat-Torricelli vector of the four vectors $\vec{T}_0 = \vec{m}_1 + \vec{m}_2 + \vec{m}_3$ and $\vec{T}_k = 2\vec{m}_k - \vec{T}_0$ for $k = 1, 2, 3$.

Thus far we have discussed criteria of joint measurability for pairs or triples of measurements. This leads to the following question: Which joint-measurability structures in a set of POVMs are possible? For instance, one may ask for a triplet of measurements where each pair is jointly measurable but all three are not jointly measurable. Such an example was constructed by Heinosaari, Reitzner, and Stano (2008). In fact, for large sets of measurements, arbitrary joint-measurability structures can be realized (Kunjwal, Heunen, and Fritz, 2014; Andrejic and Kunjwal, 2020).

The previous exact solutions of the joint-measurability problem for certain instances not only are of mathematical interest but also are relevant for deriving measurement uncertainty relations. We saw in Eq. (2) that the commutator of two projective measurements occurs naturally in the formulation of the Robertson uncertainty relation. The Robertson uncertainty relation is a preparation uncertainty relation in the sense that it constrains the ability to prepare states that are close to common eigenstates of the observables, but this is not directly related to the measurement process of the observables.

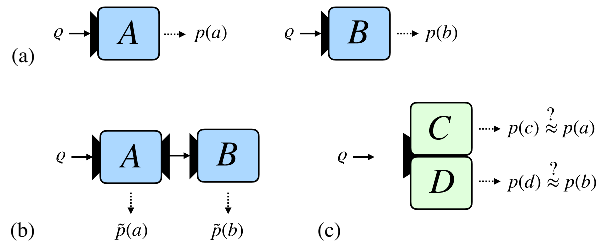

FIG. 2. Schematic view of measurement uncertainty relations. (a) For any quantum state ρ the measurement A yields results with a probability distribution $p(a)$; similarly, B delivers a distribution $p(b)$. (b) A possible way to study the disturbance of B through A is to measure in sequence and compare the distributions $\tilde{p}(a)$ and $\tilde{p}(b)$ to $p(a)$ and $p(b)$. The reverse order of measurements can also be considered. (c) Since the measurements in (b) give probability distributions other than in (a), they can be described by different POVMs C and D. But C and D are then jointly measurable by construction. The question arises: What is the best approximation of A and B by two jointly measurable POVMs C and D?

In recent years, the notion of measurement uncertainty relations has been used to quantify potential constraints and disturbances during the measurement process (Werner, 2004; Busch, Lahti, and Werner, 2014a). We can use Eq. (7) to explain this concept in a simple setting (Bullock and Busch, 2018). Assume that two projective measurements A and B on a qubit shall be implemented simultaneously. Since they may not be jointly measurable, one has to implement two POVMs C and D as approximations of A and B, respectively, where C and D are jointly measurable; see also Fig. 2. This introduces an error that may be quantified by the difference of the probabilities of one result $D^2(A, C) = 4|p(A = +) - p(C = +)|$, which is a simple case of the so-called Wasserstein distance between two probability distributions. If one considers the worst case and maximizes this error over all quantum states, one finds

$$D^2(A, C) + D^2(B, D) \geq \sqrt{2}(\|\vec{a} + \vec{b}\| + \|\vec{a} - \vec{b}\| - 2), \tag{11}$$

where $\vec{a}$ and $\vec{b}$ are the Bloch vectors of the measurements, such as $\vec{m}$ in Eq. (6). This shows that Eq. (7) can result in uncertainty relations similar to the commutator in Eq. (2).

B. Quantification of incompatibility

1. Joint measurability as a semidefinite program

Given a set of measurements, the existence of a joint measurement can be decided through convex optimization techniques that we explain in the following. It is instructive to first consider the case of two arbitrary two-outcome (or binary) POVMs (Wolf, Perez-Garcia, and Fernandez, 2009). Let $A_1 = \{A_{+|1}, A_{-|1}\}$ and $A_2 = \{A_{+|2}, A_{-|2}\}$ be two POVMs in dimension d. They are jointly measurable if there is a four-outcome measurement M_{a_1,a_2} with $a_1, a_2 = \pm$ such that $M_{++} + M_{+-} = A_{+|1}$ and $M_{++} + M_{-+} = A_{+|2}$. By writing the elements of the parent POVM in terms of A_1 and A_2, one realizes that their compatibility is equivalent to the existence of a positive semidefinite operator M_{++} for which

$$A_{+|1} + A_{+|2} - \mathbb{1} \leq M_{++} \leq A_{+|x} \tag{12}$$

for $x = 1, 2$. The problem of deciding if such an operator exists can be cast as a semidefinite program (SDP) by minimizing a real number γ subject to the constraints that $A_{+|1} + A_{+|2} \leq \gamma\mathbb{1} + M_{++}$ and $0 \leq M_{++} \leq A_{+|x}$ for $x = 1, 2$. If $\gamma = 1$ can be reached, then the measurements A_1 and A_2 are jointly measurable.

In its most general form, a SDP can be written as

$$\max\ \mathrm{tr}[AX] \tag{13}$$

$$\text{subject to } \Phi(X) = B, \tag{14}$$

$$X \geq 0, \tag{15}$$

where A and B are Hermitian operators and Φ is a Hermiticity-preserving linear map. Note that in the literature SDPs are also frequently written as a maximization of the function $\sum_i c_i x_i$ over real variables x_i, subject to the constraint that $F_0 + \sum_i x_i F_i \geq 0$ is a positive semidefinite matrix and the F_i are Hermitian matrices.

The theory of convex optimization and, in particular, of semidefinite programming is well developed (Boyd and Vandenberghe, 2004; Gärtner and Matoušek, 2012) and is a frequently used tool in quantum information theory (Watrous, 2018). In fact, the previously mentioned optimization problem can easily be solved using available software such as CVX (Grant and Boyd, 2020) or MOSEK (ApS, 2021). SDPs enjoy many properties that make them useful as a mathematical tool. For instance, each SDP can be associated to its so-called dual program, which reads

$$\min\ \mathrm{tr}[BY] \tag{16}$$

$$\text{subject to } \Phi^\dagger(Y) \geq A, \tag{17}$$

$$Y = Y^\dagger, \tag{18}$$

where $\Phi^\dagger$ denotes the adjoint map of Φ defined by $\mathrm{tr}[T\Phi(X)] = \mathrm{tr}[\Phi^\dagger(T)X]$. An important property of many SDPs is known as *strong duality*, which refers to the fact that under certain conditions known as *Slater's conditions* the optimal values of the primal and the dual problem coincide.

Concerning general POVMs, one can formulate the SDP based on the following considerations. One first observes that the classical postprocessing in Eq. (3) can be chosen as deterministic; i.e., $p(a|x,\lambda) = D(a|x,\lambda)$ takes only the values 0 and 1 (Ali *et al.*, 2009). There is only a finite number of such postprocessings. Thus, for a set $\{A_{a|x}\}$ of POVMs the following SDP decides whether or not they are jointly measurable:

$$\text{given } \{A_{a|x}\}_{a,x}, \{D(a|x,\lambda)\}_\lambda \tag{19}$$

$$\max_{\{G_\lambda\}} \mu \tag{20}$$

$$\text{subject to } \sum_\lambda D(a|x,\lambda) G_\lambda = A_{a|x} \quad \forall\ a, x, \tag{21}$$

$$G_\lambda \geq \mu\mathbb{1}, \qquad \sum_\lambda G_\lambda = \mathbb{1}. \tag{22}$$

This optimization is performed for each fixed deterministic postprocessing $\{D(a|x,\lambda)\}_\lambda$. If this optimization results in a value of μ strictly less than zero, the positivity constraint on the joint observable cannot be fulfilled, which proves incompatibility. Otherwise, a joint observable that proves joint measurability is found.

2. Various quantifiers of incompatibility

Typically, one is interested not only in answering the question as to whether a set of measurements is incompatible but also in quantifying to what extent the measurements are incompatible. Similarly, in entanglement theory not only does one ask whether a state ϱ is entangled or separable, one is also interested in how close a state is to being separable. This can be done by adding a certain amount of noise until an entangled state becomes separable. Different types of noise lead to different quantifiers including the generalized entanglement robustness (Vidal and Tarrach, 1999; Steiner, 2003; Brandão, 2005; Gühne and Tóth, 2009) or the best separable approximation introduced by Lewenstein and Sanpera (1998).

Similarly, for measurement incompatibility one can ask how close a set of POVMs is to the set of compatible sets of POVMs by means of adding a certain type of noise. Consider a set of incompatible POVMs to which we add an amount $p \in [0,1]$ of classical noise. The resulting POVMs are thus given by $A^{(p)}_{a|x} = (1-p)A_{a|x} + p\mathbb{1}/|a|$, where $|a|$ denotes the number of outcomes and the quantity

$$R^{\mathrm{n}}_{\mathrm{inc}}(A_{a|x}) = \inf\{p | A^{(p)}_{a|x} \text{ are compatible}\} \tag{23}$$

is called the incompatibility noise robustness [cf. Heinosaari, Kiukas, and Reitzner (2015)], which is one minus the incompatibility random robustness given by Designolle, Farkas, and Kaniewski (2019). Heinosaari, Kiukas, and Reitzner (2015) showed that this quantity is an incompatibility monotone since it fulfills the following properties: (i) it vanishes on compatible sets, (ii) it is symmetric under exchange of measurements, and (iii) it does not increase under preprocessing by a quantum channel.

When other types of noise are considered, one arrives at similar quantities that all share similar properties; for instance, they act as monotones under certain transformations. One example is the incompatibility weight (Pusey, 2015), which is analogous to the steering weight defined by Gallego and Aolita (2015). The incompatibility weight of a set $\{A_{a|x}\}$ of

POVMs is the smallest value of ν for which the decomposition $A_{a|x} = \nu N_{a|x} + (1-\nu) O_{a|x}$ exists, where $N_{a|x}$ is an arbitrary "noise" POVM and $\{O_{a|x}\}$ are jointly measurable. More precisely,

$$W_{\mathrm{inc}}(A_{a|x}) = \inf\left\{\nu \geq 0 \,\middle|\, \frac{A_{a|x} - \nu N_{a|x}}{1-\nu} = O_{a|x}\right\}. \tag{24}$$

In entanglement theory this quantity is known as the best separable approximation (Lewenstein and Sanpera, 1998). It was furthermore shown that the incompatibility weight is a monotone when transformations more general than quantum preprocessing are allowed (Pusey, 2015). More precisely, the incompatibility weight is a monotone under compatibility nondecreasing operations (CNDOs) that consist of preprocessing by a quantum instrument and conditional classical postprocessing.

A similar construction is the incompatibility robustness (Haapasalo, 2015; Uola *et al.*, 2015) defined by

$$R_{\mathrm{inc}}(A_{a|x}) = \inf\left\{t \geq 0 \,\middle|\, \frac{A_{a|x} + t N_{a|x}}{1+t} = O_{a|x}\right\}, \tag{25}$$

where $\{N_{a|x}\}$ is any set of POVMs and $\{O_{a|x}\}$ are jointly measurable; see also Fig. 3. Like the incompatibility weight, the incompatibility robustness is a monotone under CNDOs.

These quantifiers are all based on the convex distance of incompatible POVMs to the set of jointly measurable ones under the addition of different types of noise. In particular, all these distances can be evaluated numerically, as they fall under the framework of the previously explained SDPs. For instance, it was shown following the construction of the steering robustness (Piani and Watrous, 2015) that the incompatibility robustness can be cast as the following optimization problem (Uola *et al.*, 2015):

$$\min \ \frac{1}{d}\sum_\lambda \mathrm{tr}[G_\lambda] \tag{26}$$

$$\text{subject to} \sum_\lambda D(a|x,\lambda) G_\lambda \geq A_{a|x} \quad \forall \ a, x, \tag{27}$$

$$G_\lambda \geq 0, \tag{28}$$

$$\sum_\lambda G_\lambda = \frac{\mathbb{1}}{d}\left(\sum_\lambda \mathrm{tr}[G_\lambda]\right). \tag{29}$$

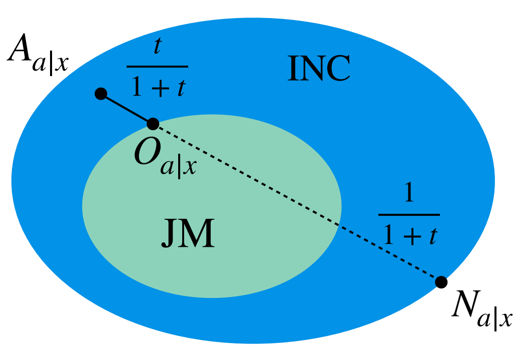

FIG. 3. Illustration of the incompatibility robustness. JM, compatible sets; INC, incompatible sets. The POVMs $\{A_{a|x}\}$ are mixed with arbitrary noise POVMs $\{N_{a|x}\}$ such that their mixture is compatible. The incompatibility robustness is the smallest t that realizes such a decomposition; cf. Eq. (25).

A recent and more detailed review on the numerical evaluation of robustness based incompatibility measures was given by Cavalcanti and Skrzypczyk (2017).

C. Constructing joint measurements

1. Adaptive strategy

An intuitive way to build joint measurements for given POVMs was presented by Uola *et al.* (2016). The idea is to exploit classical randomness between measurements that are in some sense similar to the original ones. In the simplest case of two measurements $\{A_{a|x}\}$ with $x = 1, 2$, one can flip a coin to decide which measurement to perform and assign an outcome to the other measurement based on the gained information, i.e., classically postprocess the outcome. However, in many scenarios it is better to flip a coin between some measurements other than the original pair.

To illustrate the technique, we define a pair of noisy spin measurements as $A_{\pm|1}(\mu) = (1/2)(\mathbb{1} \pm \mu\sigma_x)$ and $A_{\pm|2}(\lambda) = (1/2)(\mathbb{1} \pm \lambda\sigma_z)$. The auxiliary measurements are defined as $B_{\pm|1}(\mu,\lambda) = (1/2)[\mathbb{1} \pm (\mu\sigma_x + \lambda\sigma_z)/N]$ and $B_{\pm|2}(\mu,\lambda) = (1/2)[\mathbb{1} \pm (\mu\sigma_x - \lambda\sigma_z)/N]$; see also Fig. 4 for the case $\mu = \lambda$. Here $N = \sqrt{\mu^2 + \lambda^2}$ is the norm of the Bloch vector, which guarantees the positivity of the effects of the auxiliary measurements. Flipping a coin between these measurement and making the obvious assignments of values leads to a marginal form joint POVM with the effects

$$M_{++}(\mu,\lambda) = \tfrac{1}{2}B_{+|1}(\mu,\lambda), \quad M_{--}(\mu,\lambda) = \tfrac{1}{2}B_{-|1}(\mu,\lambda), \tag{30}$$

$$M_{-+}(\mu,\lambda) = \tfrac{1}{2}B_{-|2}(\mu,\lambda), \quad M_{+-}(\mu,\lambda) = \tfrac{1}{2}B_{+|2}(\mu,\lambda). \tag{31}$$

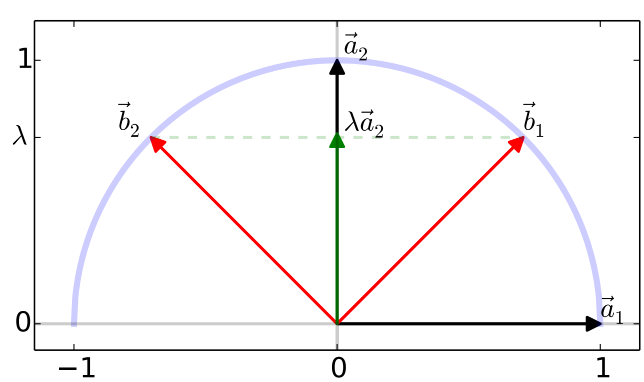

FIG. 4. An equally weighted coin toss between the auxiliary elements $B_{+|1}(\lambda,\lambda)$ and $B_{+|2}(\lambda,\lambda)$ with $\lambda = 1/\sqrt{2}$ results in an effect with the Bloch vector $\lambda\vec{a}_2 = \lambda(0,0,1) = (0,0,1/\sqrt{2})$. Here $\vec{b}_1$ and $\vec{b}_2$ are the Bloch vectors of $B_{+|1}(\lambda,\lambda)$ and $B_{+|2}(\lambda,\lambda)$. The vectors $\vec{a}_1$ and $\vec{a}_2$ represent the directions of the original measurements.

More compactly, we have $M_{i,j} = (1/4)[\mathbb{1} + (i\mu\sigma_x + j\lambda\sigma_z)/N]$. This joint measurement gives the noisy versions $A_{\pm|1}(\mu/N)$ and $A_{\pm|2}(\lambda/N)$ as marginals. For example, in the case of $\mu = \lambda = 1/\sqrt{2}$ we get $N = 1$, which corresponds to the optimal threshold in Eq. (7); cf. Fig. 4.

The previously mentioned approach for finding joint measurements entitles an adaptive strategy (Uola et al., 2016), as it uses the gained information to assign values to other measurements. In principle, one can use unbiased coins and a different number of auxiliary measurements, and every such scenario will lead to a joint measurement of some POVMs. Whereas random or uneducated guesses for the auxiliary measurements are not guaranteed to give a good joint measurement, optimal auxiliary measurements are rather straightforward to find in scenarios with symmetry, such as in the case of pairs of mutually unbiased bases (MUBs) (Uola et al., 2016); see also Sec. III.D for explicit results on symmetric measurement sets.

2. Operator measure with correct marginals

Jae et al. (2019) proposed building joint measurements for pairs of measurements using a specific ansatz. For a pair of POVMs $\{A_a\}$ and $\{B_b\}$ that both have n values, they defined a so-called W measure as

$$W_{ab} = C_{ab} + \frac{1}{n}\left(A_a - \sum_j C_{aj}\right) + \frac{1}{n}\left(B_b - \sum_i C_{ib}\right), \quad (32)$$

where $\{C_{ab}\}$ is an arbitrary POVM. $\{W_{ab}\}$ has the original pair of POVMs as its marginals, but its elements are not required to be positive semidefinite. Still, this ansatz already allows a numerical treatment. If one takes $\{C_{ab}\}$ to be the parent POVM $\{M_{ab}\}$, then also $\{W_{ab}\} = \{M_{ab}\}$. Thus, Eq. (32) can be used to formulate an iteration with the desired parent POVM as one fixed point.

Moreover, it is straightforward to show that instead of using the POVM $\{C_{ab}\}$ one can parametrize W measures as

$$W_{ab} = \frac{1}{n}(A_a + B_b) - \Omega_{ab}, \quad (33)$$

where $\{\Omega_{ab}\}$ are Hermitian operators with the property $\sum_a \Omega_{ab} = \sum_b \Omega_{ab} = \mathbb{1}/n$. Jae et al. (2019) noted that the original pair is jointly measurable if and only if there is a collection $\{\Omega_{ab}\}$ such that the corresponding collection $\{W_{ab}\}$ is positive semidefinite.

This allows one to introduce the negativity of a W measure as $\mathcal{N} := (1/d)\sum_{a,b,k}(|\lambda^k_{ab}| - \lambda^k_{ab})$, where $\{\lambda^k_{ab}\}_k$ are the eigenvalues of W_{ab}. $\mathcal{N} = 0$ for some $\{\Omega_{ab}\}$ if and only if the POVMs are jointly measurable. Note that this enables one to decide joint measurability using a direct minimization, which may be solved analytically for special cases without using SDPs.

Indeed, Jae et al. (2019) minimized $\mathcal{N}$ over all collections $\{\Omega_{ab}\}$ for two important cases, namely, general unbiased qubit POVMs with two outcomes and special qubit so-called trinary POVMs with three outcomes. The former results in the known Busch criterion in Eq. (7), and the latter results in

$$\mathcal{N}_{\min} = \max\left(\frac{1}{9}\sum_{a,b}\|\vec{m}_a + \vec{n}_b - \vec{\theta}_{ab}\| - 1, 0\right) \quad (34)$$

for two trinary POVMs with the Bloch vectors $\{\vec{m}_a\}$ of the POVM $\{A_a\}$ and $\{\vec{n}_b\}$ of the POVM $\{B_b\}$, where the three effects for each of the measurements all lie in the same plane of the Bloch sphere. Here $\vec{\theta}_{ab} = \vec{m}_{2(a+b)} + \vec{n}_{2(a+b)}$. This can be translated into a condition on joint measurability reading

$$\sum_{a,b}\|\vec{m}_a + \vec{n}_b - \vec{\theta}_{ab}\| \leq 9. \quad (35)$$

3. Naimark strategy

Here we review an analytical technique for characterizing all possible parent POVMs related to a given measurement (Pellonpää, 2014b; Haapasalo and Pellonpää, 2017b). This relies on extending the POVMs to PVMs on a larger system using the so-called Naimark extension. The method shows a tight connection between joint measurability and commutativity in the extended Hilbert space picture. As a direct consequence, the technique gives an important structural result: All POVMs jointly measurable with a given rank-1 POVM are its postprocessings.

We first introduce the desired Naimark extension (Naimark, 1940; Peres, 1995). Let $\{A_a\}$ be a POVM with n outcomes in a d-dimensional system described by the Hilbert space $\mathcal{H}$. We can dilate $\{A_a\}$ to a larger Hilbert space as follows: Each effect of the POVM can be written as $A_a = \sum_{k=1}^{m_a} |d_{ak}\rangle\langle d_{ak}|$, where the vectors $|d_{ak}\rangle = \sqrt{\lambda_{ak}}|\varphi_{ak}\rangle$ for $k = 1, \ldots, m_a$ are the unnormalized eigenvectors associated with the nonzero eigenvalues λ_{ak} and m_a is the rank of A_a. Consequently, $\{|\varphi_{ak}\rangle\}_{k=1}^d$ is an orthonormal eigenbasis of A_a.

To write the extension, let $\mathcal{H}_\oplus$ be a $(\sum_a m_a)$-dimensional Hilbert space with an orthonormal basis $\{|e_{ak}\rangle\}_{a,k}$ and consider the PVM $P_a = \sum_{k=1}^{m_a} |e_{ak}\rangle\langle e_{ak}|$ with n outcomes a. We can then define the map $J = \sum_{a=1}^n \sum_{k=1}^{m_a} |e_{ak}\rangle\langle d_{ak}|$, which is an isometry since $J^\dagger J = \mathbb{1}$ on $\mathcal{H}$. Now $A_a = J^\dagger P_a J$, so the triplet $(\mathcal{H}_\oplus, J, \{P_a\})$ is a Naimark dilation of $\{A_a\}$. This dilation is minimal, meaning that the set of vectors $\{P_a J|\psi\rangle | a, |\psi\rangle\}$ spans $\mathcal{H}_\oplus$ (Pellonpää, 2014b; Haapasalo and Pellonpää, 2017b). Note that $\{A_a\}$ is sharp exactly when J is unitary. In this case, we may identify $\{A_a\}$ with $\{P_a\}$.

Based on this Naimark dilation, one can obtain several structural insights. If $\{B_b\}$ is a POVM that is jointly measurable with $\{A_a\}$, then any joint measurement $\{M_{ab}\}$ is of the form $M_{ab} = J^\dagger P_a \tilde{B}_b J$, where $\{\tilde{B}_b\}$ is a unique POVM of $\mathcal{H}_\oplus$ that commutes with P_a, i.e., $P_a \tilde{B}_b = \tilde{B}_b P_a$ (Pellonpää, 2014b; Haapasalo and Pellonpää, 2017b). The uniqueness follows since the dilation is minimal. This gives an effective method for constructing all POVMs to be jointly measurable with $\{A_a\}$. They are of the form $B_b = J^\dagger \tilde{B}_b J$.

We have two immediate special cases: If $\{A_a\}$ is sharp (i.e., J is unitary), then $M_{ab} = J^\dagger P_a J J^\dagger \tilde{B}_b J = A_a B_b = B_b A_a$, so the POVMs must commute. If $\{A_a\}$ is of rank 1 [i.e., any $m_a = 1$, $A_a = |d_a\rangle\langle d_a|$, and $P_a = |e_a\rangle\langle e_a|$ (Pellonpää, 2014a)], then each $\tilde{B}_b$ is diagonal in the basis $\{|e_a\rangle\}$, so one can write $\tilde{B}_b = \sum_a p(b|a)|e_a\rangle\langle e_a|$. From this one obtains $M_{ab} = p(b|a)A_a$, where $p(b|a)$ is a conditional probability. Thus, any POVM $\{B_b\}$ that is jointly measurable with $\{A_a\}$ is a classical postprocessing $B_b = \sum_a p(b|a)A_a$. Finally, we note that if $M_{ab} = J^\dagger P_{ab} J$ is a Naimark dilation of a joint POVM of $\{A_a\}$ and $\{B_b\}$, then we get a common Naimark dilation for $A_a = J^\dagger(\sum_b P_{ab})J$ and $B_b = J^\dagger(\sum_a P_{ab})J$ where the marginal PVMs commute.

D. High-dimensional measurements and symmetry

Thus far we have presented methods for the characterization of joint measurability that are applicable mainly to low-dimensional systems. The explicit criteria in Sec. III.A were formulated for qubits, and the computational approaches in Secs. III.B and III.C.2 are naturally restricted due to numerical limitations. As we now explain, for high-dimensional systems symmetries and algebraic relations can often be used to characterize joint measurability.

To start, we consider the case of measurements in MUBs. Two bases $|\psi_i\rangle$ and $|\phi_j\rangle$ of a d-dimensional space are called mutually unbiased if they obey

$$|\langle\psi_i|\phi_j\rangle|^2 = \frac{1}{d} \tag{36}$$

for all i and j. As an example, the eigenstates of the three Pauli matrices form a triplet of MUBs. In fact, MUBs can be seen as a generalization of the Pauli matrices to higher dimensions, and as such one may expect MUBs to correspond to highly incompatible measurements. Regardless, MUBs are relevant for various quantum information processing tasks like quantum tomography or quantum key distribution. It is known that for a given d maximally $d+1$ MUBs can exist, but whether this bound can be reached is an open problem and, indeed, one of the hard problems in quantum information theory; see also Bengtsson (2007) and Horodecki, Rudnicki, and Życzkowski (2020) for an overview.

To quantify the incompatibility of general MUBs, Designolle *et al.* (2019) proceeded as follows. As a quantifier, they used a variant of incompatibility noise robustness as introduced in Eq. (23) by considering the noisy POVM $A^\eta_{a|x} = \eta A_{a|x} + (1-\eta)\mathrm{tr}(A_{a|x})\mathbb{1}/d$ and asking for the maximal η^* such that the POVMs $A^\eta_{a|x}$ were compatible. Such robustness is also called the depolarizing or white noise robustness. The optimal η^* can be computed by a SDP. For deriving upper bounds on η^*, one can consider the dual optimization problem, which is a minimization problem. Inserting a specific instance of the dual variables results in an analytical upper bound on η^*. For instance, for k projective measurements where the projectors are of rank 1, one finds that

$$\eta^* \leq \eta_{\mathrm{up}} = \frac{\lambda - k/d}{k - k/d}, \tag{37}$$

where λ is the largest eigenvalue of an operator X that can be obtained by selecting one outcome per measurement, that is, $X = \sum_{x=1}^k A_{j_x|x}$ for some $\vec{j}$. A set of k MUBs can also be viewed as measurements, and for prime power dimensions there is an explicit construction of $d+1$ MUBs due to Wootters and Fields (1989). It turns out that when taking $k=2$, $k=d$, or $k=d+1$ of these MUBs, one finds that $\eta^* = \eta_{\mathrm{up}}$; as a result, in this case these MUBs are maximally incompatible. We stress, however, that in general two sets of MUBs can be inequivalent (in the sense that they are not connected by a unitary transformation or permutation), and in general MUBs do not reach the bound in Eq. (37). Still, one can prove for general MUBs a lower bound on η^* (Designolle *et al.*, 2019), as well as upper and lower bounds for general measurements (Designolle, Farkas, and Kaniewski, 2019).

This result begs the questionas to which measurements are the most incompatible for a given quantifier of measurement incompatibility. This was studied in detail by Bavaresco *et al.* (2017) and Designolle, Farkas, and Kaniewski (2019). The most incompatible pair of measurements depends on the chosen quantifier, and sometimes other measurements besides MUBs are the most incompatible ones.

Finally, one may ask how interesting sets of measurements, such as those with a high incompatibility or other desirable properties, can be identified from abstract principles. This problem was addressed by Nguyen *et al.* (2020), who studied sets of measurements from group theoretic perspectives. In fact, starting with a complex reflection group G and with a given representation, one can construct a measurement assemblage (i.e., a set of measurements) with certain symmetries. These assemblages often have interesting physical properties such as high incompatibility.

IV. INCOMPATIBILITY AND QUANTUM INFORMATION PROCESSING

Measurement incompatibility is inherently linked to the nonclassical character of quantum correlations. Indeed, for many scenarios it is evident that incompatibility of the performed measurements is necessary for displaying nonclassical correlations. Since such correlations are required for tasks like quantum key distribution or quantum metrology, these connections highlight the resource aspect of measurement incompatibility. In this section, we describe different forms of nonclassical correlations as well as other phenomena, for which the incompatibility of measurements is essential.

A. Bell nonlocality

To start, we discuss the relation between joint measurability and Bell nonlocality (Bell, 1964; Brunner *et al.*, 2014). In this scenario, one considers two parties Alice and Bob, and each of them performs some measurements $\{A_{a|x}\}$ and $\{B_{b|y}\}$, respectively; see also Fig. 5. The question then arises as to whether or not the observed probabilities $p(a,b|x,y)$ of the results a and b for the given inputs x and y can be explained by a local hidden variable model. This means that they can be written as

$$p(a,b|x,y) = \int d\lambda p(\lambda)\chi^A(a|x,\lambda)\chi^B(b|y,\lambda), \tag{38}$$

where λ is the hidden variable occurring with probability $p(\lambda)$ and $\chi^A(a|x,\lambda)$ and $\chi^B(b|y,\lambda)$ are the response functions of

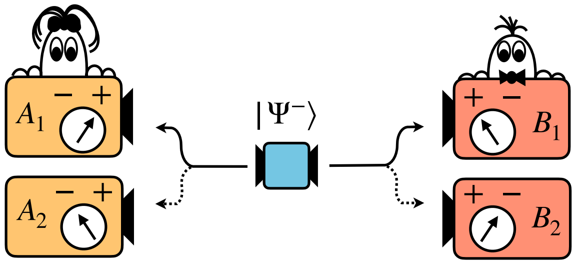

FIG. 5. Schematic of the Bell scenario. A source distributes two particles to two parties named Alice and Bob. They perform measurements on it, and the questions arises as to whether or not the observed probability distribution can be described by a local hidden variable model.

Alice and Bob, respectively. Note that here no reference to quantum mechanics is made and no knowledge about the measurements on each side is assumed.

To start, we explain why jointly measurable observables on Alice's side can never lead to Bell nonlocality (Fine, 1982; Wolf, Perez-Garcia, and Fernandez, 2009). For simplicity, we consider two measurements for Alice, $x = 1, 2$ with outcomes a_1 and a_2, and analogously for Bob. Since Alice's measurements are jointly measurable, Alice may perform the parent POVM and directly obtain the probability distribution $p(a_1, a_2)$. If Bob performs one measurement B_y simultaneously, one will observe the probability distribution $p(a_1, a_2, b_y|B_y)$. This has to obey $p(a_1, a_2) = \sum_{b_y} p(a_1, a_2, b_y|B_y)$ regardless of y since the correlations obey the nonsignaling condition; i.e., Bob cannot send information to Alice by choosing his measurement. However, one may define a global probability distribution as

$$p(a_1, a_2, b_1, b_2) = \frac{p(a_1, a_2, b_1|B_1)p(a_1, a_2, b_2|B_2)}{p(a_1, a_2)}. \qquad (39)$$

Indeed, one can directly verify that this obeys all the properties of a probability distribution. The existence of such a distribution, however, already implies that the conditional distributions obey Eq. (38) since the global distribution can always be expressed as a probabilistic mixture of deterministic assignments (Fine, 1982).

The question remains as to whether any set of incompatible observables on Alice's side can lead to Bell nonlocality if the underlying quantum state and Bob's measurements are properly chosen. For that, one needs to show that the correlations violate some suitable Bell inequality, and connections between violations of Bell inequalities and the degree of incompatibility were soon observed (Andersson, Barnett, and Aspect, 2005; Son *et al.*, 2005).

For the simplest scenario in which Alice and Bob have two measurements each with two possible outcomes ± 1, there is a direct connection between incompatibility and Bell nonlocality. In this case, the only relevant Bell inequality is the one by Clauser, Horne, Shimony, and Holt (CHSH) that reads (Clauser *et al.*, 1969, 1970)

$$S = \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \leq 2. \qquad (40)$$

In Eq. (40) $\langle A_x B_y \rangle = p(+,+|x,y) - p(+,-|x,y) - p(-,+|x,y) + p(-,-|x,y)$ denotes the expectation value of a correlation measurement.

The connection established by Wolf, Perez-Garcia, and Fernandez (2009) uses the SDP formulation of joint measurability. As shown in Sec. III.B, for two measurements $\{A_{a|x}\}$ with $x = 1, 2$ the existence of a parent POVM $\{M_{\pm\pm}\}$ with four outcomes can be rephrased as a search for an effect M_{++} obeying $A_{+|1} + A_{+|2} - \mathbb{1} \leq M_{++} \leq A_{+|x}$ for $x = 1, 2$; see also Eq. (12). As mentioned, the search for such an effect M_{++} can be formulated as a simple SDP considered as a feasibility problem.

Given the SDP formulation in Eq. (12), one can consider the dual SDP. As it turns out, this is directly linked to the CHSH inequality: The additional variables of the dual problem can be viewed as a quantum state, and measurements on Bob's side and the CHSH inequality can be violated if and only if the SDP defined by Eq. (12) is unfeasible or, in other words, the measurements $\{A_{a|x}\}$ are incompatible.

For general scenarios the connection is, however, not as strict anymore. More explicitly, Bene and Vértesi (2018) presented a set of three measurements with two outcomes on a qubit that are pairwise jointly measurable, but there is no common parent POVM for the entire set, so the triplet is incompatible. It is then shown for all quantum states and measurements on Bob's side that the resulting correlations are local in the sense of Eq. (38). This result holds for an arbitrary number of POVMs on Bob's side. Note that for the special case of dichotomic measurements on Bob's side an analogous result was previously shown by Quintino *et al.* (2016), and the case of an infinite number of measurements on Alice's side was considered by Hirsch, Quintino, and Brunner (2018).

B. Quantum steering

As mentioned in Sec. IV.A, there are nonjointly measurable sets that cannot break any Bell inequalities for any quantum state. Here we review the results of Quintino, Vértesi, and Brunner (2014), Uola, Moroder, and Gühne (2014), Uola *et al.* (2015), and Kiukas *et al.* (2017) showing that for a slightly weaker form of correlations the so-called quantum steering incompatibility exactly characterizes the sets of measurements that allow for the relevant nonlocal effect, which may be considered a spooky action at a distance (Uola, Costa *et al.*, 2020).

The modern formulation of quantum steering came from Wiseman, Jones, and Doherty (2007). In this formulation one party, say, Alice, performs measurements $\{A_{a|x}\}$ in her local laboratory on a bipartite quantum state ϱ_{AB}. When asked to perform measurement x, she announces an output a, and the postmeasurement state in Bob's laboratory is given by

$$\sigma_{a|x} = \mathrm{tr}_A[(A_{a|x} \otimes \mathbb{1})\varrho_{AB}]. \qquad (41)$$

Assuming that the experiment is repeated many times and that Bob has access to a tomographically complete set of measurements, he can reconstruct these unnormalized states, also called an assemblage $\{\sigma_{a|x}\}$. It is straightforward to verify that the assemblage fulfills the condition of nonsignaling;

i.e., the operator $\sum_a \sigma_{a|x} \coloneqq \sigma_B = \mathrm{tr}_A[\varrho_{AB}]$ is independent of the input x.

The unsteerability of such an assemblage is associated with the existence of a local hidden state model. This is a model consisting of a local ensemble of states $\{p(\lambda)\varrho_\lambda^B\}$ on Bob's side whose priors $p(\lambda)$ are updated upon learning the classical information (a,x). In other words, an assemblage $\sigma_{a|x}$ is unsteerable if

$$\sigma_{a|x} = p(a|x)\sum_\lambda p(\lambda|a,x)\varrho_\lambda^B = \sum_\lambda p(\lambda)p(a|x,\lambda)\varrho_\lambda^B, \qquad (42)$$

and steerable otherwise. In Eq. (42) $p(a|x) = \mathrm{tr}[\sigma_{a|x}]$, and the last equality follows from the fact that the hidden states are independent of the measurement choice, i.e., $p(x,\lambda) = p(x)p(\lambda)$; see Uola, Costa *et al.* (2020) for a detailed interpretation of such models.

From Eq. (42) we see that separable states $\varrho_{AB}^{\mathrm{sep}} = \sum_\lambda p(\lambda)\varrho_\lambda^A \otimes \varrho_\lambda^B$ cannot lead to steerable assemblages. It is well known, however, that entanglement is not sufficient for quantum steering (Wiseman, Jones, and Doherty, 2007; Quintino *et al.*, 2015). For example, the Werner states (Werner, 1989) and the isotropic states within a certain parameter regime provide examples of entangled states that have a local hidden state model for all measurements on Alice's side (Werner, 1989; Barrett, 2002; Almeida *et al.*, 2007; Wiseman, Jones, and Doherty, 2007; Nguyen and Gühne, 2020a, 2020b).

Despite recent progress (Nguyen, Nguyen, and Gühne, 2019), the problem of deciding steerability of generic quantum states remains open. Still, a complete characterization of the measurements that lead to steering is known (Quintino, Vértesi, and Brunner, 2014; Uola, Moroder, and Gühne, 2014). The first observation is that the joint measurability of Alice's measurements leads to an unsteerable assemblage for any shared state. To see this, one can simply plug Eq. (3) into Eq. (41). Conversely, use of the maximally entangled state $|\psi^+\rangle = (1/\sqrt{d})\sum_{n=1}^d |n\rangle \otimes |n\rangle$ yields $\sigma_{a|x} = (1/d)A_{a|x}^T$, where T denotes the transpose in the computational basis $\{|n\rangle\}$. Comparing Eq. (42) to Eq. (3) shows that a local hidden state model for $\{\sigma_{a|x}\}$ can be converted into a joint measurement of $\{A_{a|x}\}$ by denoting $G_\lambda = p(\lambda)d(\varrho_\lambda^B)^T$. We arrive at the following result.

Observation.—Joint measurability of Alice's measurements leads to unsteerable assemblages for any shared quantum state. Conversely, for any set of incompatible measurements there is a shared state for which these measurements lead to steering.

The previous connection can be used to translate results on joint measurability to steering, and vice versa (Uola, Moroder, and Gühne, 2014; Cavalcanti and Skrzypczyk, 2016; Chen *et al.*, 2016). For example, the incompatibility robustness of Alice's measurements is known to be lower bounded by the so-called steering robustness of the corresponding assemblage (Cavalcanti and Skrzypczyk, 2016; Chen *et al.*, 2016). As steering verification does not assume Alice's measurements to be trusted, such lower bounds constitute semidevice independent bounds on measurement incompatibility; see also Chen *et al.* (2021) for further quantification techniques in the device-independent setting. Moreover, Uola, Moroder, and Gühne (2014) showed that the characterization of steerability of the isotropic state (Wiseman, Jones, and Doherty, 2007) translates into a characterization of the white noise robustness of all projective measurements in a given dimension.

There are also two broader connections between the concepts of joint measurability and steering. The first connection is that the previous result relies on the use of a pure maximally entangled state. To relax this, Uola *et al.* (2015) showed that the unsteerability of an assemblage $\{\sigma_{a|x}\}$ is equivalent to the joint measurability of the corresponding square root or "pretty good measurements." These are measurements that are known to give a good, but not always optimal, performance in discriminating among the corresponding sets of states (Hausladen and Wootters, 1994), and they are known to be related to the information capacity of quantum measurements (Dall'Arno, D'Ariano, and Sacchi, 2011; Holevo, 2012). They are given by $\sigma_B^{-1/2}\sigma_{a|x}\sigma_B^{-1/2}$, where $\sigma_B = \sum_a \sigma_{a|x}$ and a pseudoinverse is used when necessary. To see the more general connection between steering and measurement incompatibility, one can simply sandwich a local hidden state model in Eq. (42) with $\sigma_B^{-1/2}$, or sandwich a joint measurement in Eq. (3) with $\sigma_B^{1/2}$. We summarize this in the following.

Observation.—A state assemblage $\{\sigma_{a|x}\}$ is steerable if and only if the corresponding pretty good measurements

$$B_{a|x} \coloneqq \sigma_B^{-1/2}\sigma_{a|x}\sigma_B^{-1/2} \qquad (43)$$

are incompatible.

This result gives a direct link between steering criteria and incompatibility conditions. For example, the incompatibility criteria of Sec. III.A were used to fully characterize the steerability of two-input two-output qubit assemblages (Uola *et al.*, 2015) [cf. Chen, Ye, and Fei (2017) for a detailed analysis], and it was shown that the incompatibility robustness of $\{B_{a|x}\}$ can be used to witness the entanglement dimensionality of the underlying bipartite state (Designolle, Srivastav *et al.*, 2021). Another application of the connection was demonstrated by Uola *et al.* (2021), who used steerable states with a positive partial transpose (Moroder *et al.*, 2014) to construct incompatible qutrit measurements that are compatible in every qubit subspace. We note that similar results on incompatibility in subspaces were obtained by Loulidi and Nechita (2021) using different techniques. Furthermore, the connection can be generalized for characterizing so-called channel steering (Piani, 2015) via measurement incompatibility (Uola *et al.*, 2018).

Going one step further, Kiukas *et al.* (2017) showed that one can reformulate the steering problem of a state ϱ_{AB} in the Choi picture. This gives a map between Alice's POVMs and the pretty good measurements on Bob's side. The channel associated with a shared state ϱ_{AB} is given in the Heisenberg picture as

$$\Lambda_{\varrho_{AB}}^\dagger(A_{a|x}) = \sigma_B^{-1/2}\mathrm{tr}_A[(A_{a|x} \otimes \mathbb{1})\varrho_{AB}]^T\sigma_B^{-1/2} = B_{a|x}^T, \qquad (44)$$

where the transpose is in the eigenbasis of σ_B. We note that a generalization of the textbook Choi isomorphism is used here in which the fixed reduced state is σ_B instead of the canonical maximally mixed state (Kiukas *et al.*, 2017). It is straightforward to see that the state ϱ_{AB} is steerable if and only if the corresponding channel $\Lambda^{\dagger}_{\varrho_{AB}}$ does not break the incompatibility of some set of measurements.

The advantage of the channel approach is that it extends the connection between steering and nonjoint measurability of Alice's measurements to the infinite-dimensional case and to POVMs with nondiscrete outcome sets, and that it can unify seemingly different steering problems, such as the steerability of NOON states subjected to photon loss and steerability in systems that have amplitude damping dynamics (Kiukas *et al.*, 2017). Furthermore, a map similar to that in Eq. (44) was used in the solution of the steering problem for two qubits (Nguyen, Nguyen, and Gühne, 2019).

C. Quantum contextuality

Quantum contextuality refers to the fact that the predictions of quantum mechanics cannot be explained by hidden variable models, which are noncontextual for compatible measurements (Kochen and Specker, 1967). In simple terms, measurements are compatible if they can be measured simultaneously or in sequence without disturbance, and noncontextuality means that the model assigns values to a measurement regardless of the context; see also Budroni *et al.* (2022) for a recent review on the topic.

A noteworthy fact is that contextuality can be proven regardless of the preparation of the quantum system. Therefore, the connection between properties of quantum measurements and nonclassical behavior can be extremely strong in various scenarios. This is in stark contrast to distributed scenarios, where a properly chosen entangled state is required as a catalyst to harness the nonclassical behavior of the measurements. We review here the formal connection between the concept of joint measurability and contextuality for two different notions of contextuality.

1. Kochen-Specker contextuality

In the Kochen-Specker setup, the context is defined as a set of projective measurements that can be performed simultaneously. One then asks whether a hidden variable model could explain the outcome statistics of all measurements while assuming that the hidden variable assigns values regardless of the context. This assumption leads to various contradictions to quantum mechanics, such as in the so-called Peres-Mermin square (Mermin, 1990; Peres, 1990).

Kochen-Specker contextuality can be proven in a state-independent manner (Cabello, 2008; Yu and Oh, 2012). Hence, contextuality is a statement about measurements. The phenomenon of Kochen-Specker contextuality is based on the properties of the Hilbert space projections and it is indeed formulated for PVMs. On this level, the notions of joint measurability and nondisturbance reduce to commutativity. It is hence expected that noncommutativity will be essential for the violations of the relevant classical models. However, as one also needs context for measurements, which requires compatibility, one needs to find the exact interplay between compatibility and incompatibility in order to reveal violations of Kochen-Specker noncontextuality. This structure was characterized by Xu and Cabello (2019) using graphs to represent the possible contextuality scenarios. In their graph representation, adjacent vertices represent compatible PVMs. The main result reads as follows.

Observation.—For a given graph, there is a quantum realization with PVMs producing contextuality if and only if the graph is not chordal.

Here chordality means that the graph does not contain induced cycles with a size larger than 3. Induced cycles are subgraphs with a set of vertices S and edges E such that the vertices S are connected in a closed chain, and furthermore every edge of the original graph that has both ends in S is part of the subgraph. This specifically implies that the simplest contextuality scenario requires four measurements.

2. Spekkens contextuality

The notion of operational noncontextuality asks whether one's measurement statistics can be reproduced by means of an ontological model. Such a model assigns a distribution of ontological states λ onto each preparation procedure. This distribution is then classically postprocessed. In short, for a preparation P and a measurement M with outcomes $\{a\}$, an ontological model reads

$$p(a|P,M) = \sum_{\lambda} p(\lambda|P)p(a|M,\lambda), \tag{45}$$

where p represents a probability distribution.

For quantum theory, where preparations are presented as density matrices and measurements as POVMs, such models can be constructed in many ways. For example, (a) one can identify the space of ontological states as that of all mixed quantum states and define $p(\lambda|P)$ as the point measure concentrated on P. Similarly, (b) one can define a point measure on pure states and extend it to mixed states in a nonunique manner (Beltrametti and Bugajski, 1995). In addition, (c) one can identify quantum states with their eigendecompositions and choose $p(\lambda|P)$ to be eigenvalues and λ to be the corresponding eigenprojectors.

To find contradictions with quantum theory, one needs to seek for meaningful restrictions of the ontological model (Spekkens, 2005). One possible restriction is to demand that operationally indistinguishable preparations be represented by the same distribution of ontological states. That is, if P_1 cannot be distinguished from P_2, then $p(\lambda|P_1) = p(\lambda|P_2)$. This assumption is called preparation noncontextuality (Spekkens, 2005). An additional feature of these models is convex linearity: $p(\lambda|\mu P_1 + (1-\mu)P_2) = \mu p(\lambda|P_1) + (1-\mu)p(\lambda|P_2)$. Setting a similar restriction on indistinguishable measurements leads to the notion of measurement noncontextuality. Note, however, that these assumptions are not obeyed by several known hidden variable models (Belinfante, 1973), and their physical relevance has been debated (Ballentine, 2014).

The previous model examples fit to these restrictions as follows. The models (a) and (c) are measurement noncontextual, but they break preparation noncontextuality in the sense that they are not convex linear. The model (b) is measurement noncontextual but breaks preparation noncontextuality in the sense that mixed states do not have a unique decomposition into pure states. Hence, measurement noncontextuality is not sufficient for contradicting quantum theory (Spekkens, 2005). Later we review the results of Tavakoli and Uola (2019), showing that when all quantum state preparations are allowed, the notion of preparation noncontextuality together with its convex linearity are equivalent to joint measurability.

In the case of quantum theory, indistinguishable preparations are presented by the set of density matrices. Hence, the assumption of preparation noncontextuality restricts one to ontological models that depend only on the density matrix ϱ and not on the way that P is prepared, i.e., $p(\lambda|P) = p(\lambda|\varrho)$. The assumption of convex linearity implies that for each λ the map $p_\lambda(\varrho) \coloneqq p(\lambda|\varrho)$ extends to a linear map from trace-class operators to complex numbers. As the dual of the trace class is the set of bounded operators (Busch *et al.*, 2016), one gets $p_\lambda(\varrho) = \mathrm{tr}[\varrho G_\lambda]$ for all states ϱ and for some positive operator G_λ. Noting that $\sum_\lambda p(\lambda|\varrho) = 1$ for each state, $\{G_\lambda\}$ forms a POVM. We summarize this in the following (Tavakoli and Uola, 2019).

Observation.—Any jointly measurable set of POVMs leads to preparation noncontextual correlations for all input states. Conversely, the existence of a preparation noncontextual model for all quantum states implies joint measurability of the involved measurements.

As a direct application, one sees that bounds on preparation contextuality translate to incompatibility criteria. For the explicit form of such witnesses, see Tavakoli and Uola (2019).

We stress that in the previous result all quantum states are considered. In the experimental setting one does not have access to all possible states, and hence the set of considered preparations is finite. In such a scenario with a fixed set of states, there is no guarantee that an incompatible set of measurements would lead to preparation contextual correlations (Selby *et al.*, 2021). It was further noted by Selby *et al.* (2021) that setting the additional restriction of measurement noncontextuality corresponds to a class of models that can be violated even when using compatible measurements.

D. Macrorealism

The notion of macrorealism challenges classical intuition by asking the following question: Can one perform measurements in a way that does not disturb the subsequent evolution of the system? To formalize the concept, Leggett and Garg (1985) suggested probing hidden variable theories that fulfill the assumptions of *macroscopic realism* and *noninvasive measurability*. In short, the first assumption amounts to the existence of a hidden variable λ that carries the information about all measurements (whether or not they were performed), and the second assumption states that one can measure the system without disturbing the distribution of hidden variables. Note that the second assumption is problematic, as it is not verifiable in the experimental setting. This is sometimes referred to as the clumsiness loophole of macrorealism, and there are various proposals for getting around it (Leggett and Garg, 1985; Knee *et al.*, 2012, 2016; Li *et al.*, 2012; Wilde and Mizel, 2012; George *et al.*, 2013; Emary, Lambert, and Nori, 2014; Budroni *et al.*, 2015; Robens *et al.*, 2015; Emary, 2017; Huffman and Mizel, 2017; Ku *et al.*, 2020). For this Colloquium, a central take on the problem is given by the measurement theoretical approach of Uola, Vitagliano, and Budroni (2019), which we explain in the following.

In a typical setting, one has n time steps. At each step, one chooses either to perform or not to perform a measurement designated for that time step. Here we concentrate exclusively on such scenarios. In this case, the assumptions of Leggett and Garg are equivalent to the fact that the resulting probability distributions are no signaling in both directions (Clemente and Kofler, 2016). In other words, all sequences of measuring and not measuring are compatible with one another under the act of marginalizing. As an example, consider the case $n = 2$. Labeling the probability distributions as p_i for a single measurement at time step $i = 1, 2$ and p_{12} for the sequence, no signaling in both directions requires

$$\sum_a p_{12}(a, b) = p_2(b), \tag{46}$$

$$\sum_b p_{12}(a, b) = p_1(a) \tag{47}$$

for all outcomes a and b.

The second condition is satisfied by any physical distribution. However, requiring the first condition for all input states implies measurement compatibility in the form of nondisturbance. The notion of nondisturbance involves the optimization over all possible ways of performing the first measurement, i.e., optimization over all instruments. This raises the following observation (Uola, Vitagliano, and Budroni, 2019); see also Clemente and Kofler (2015), who discussed a connection between measurement compatibility and no-signaling conditions was discussed.

Observation.—When all clumsiness caused by the lack of capability of the observer is removed, measurement incompatibility in the form of inherent measurement disturbance is the property that allows one to distinguish quantum theory from macrorealistic ones.

For longer sequences of measurements, requiring the no-signaling constraints for all input states generates a more involved structure of nondisturbance relations. For example, in the case of three time steps one requires nondisturbance in all pairs (in time order) and from the first measurement to the rest of the sequence. Note that one has to use the same instrument for a given time step in all conditions (Uola, Vitagliano, and Budroni, 2019). In other words, the first measurement has to be nondisturbing with respect to the second, the third, and the nondisturbing sequence of the second and the third all with the same instrument. The structure arising from this generalized notion of nondisturbance was analyzed in more detail by Uola, Vitagliano, and Budroni (2019), who also discussed a resource theoretical take on the topic.

E. Prepare-and-measure scenarios

In this section we discuss the relevance of incompatible measurements in prepare-and-measure scenarios. More precisely, we first review the advantage that incompatible measurements provide over compatible ones in state-discrimination and state exclusion tasks. We then discuss the necessity of performing incompatible measurements in quantum random access codes, which is required to gain an advantage over classical random access codes. Last, we review the role of incompatibility in distributed sampling scenarios.

1. State discrimination and exclusion

A task that is important to quantum information theory, and particularly to quantum communication (Helstrom, 1969; Holevo, 1982) and quantum cryptography (Gisin *et al.*, 2002; Bennett and Brassard, 2014), is *minimum-error state discrimination* (Helstrom, 1976; Barnett and Croke, 2009); see Fig. 6(a). There one aims to correctly guess the label of a state ϱ_a that is randomly drawn from an ensemble $\mathcal{E} = \{p_a, \varrho_a\}_{a\in I}$ with the known probability p_a. To be more precise, upon receiving a state we perform a measurement of a POVM $\{A_a\}$ and guess the state to be ϱ_a whenever we observe the outcome a. The success in correctly guessing the label a can be quantified by the probability of success $p_{\text{guess}}(\mathcal{E}, \{A_a\}) = \sum_a p_a \text{tr}[A_a \varrho_a]$. The maximum probability of success is obtained by simply maximizing over all measurements $p_{\text{guess}}(\mathcal{E}) = \max_{\{A_a\}} p_{\text{guess}}(\mathcal{E}, \{A_a\})$. We emphasize that this task differs from *unambigous state discrimination*, where one is not allowed to make a wrong guess but is allowed to pass and not give an answer at all (Helstrom, 1976; Barnett and Croke, 2009).

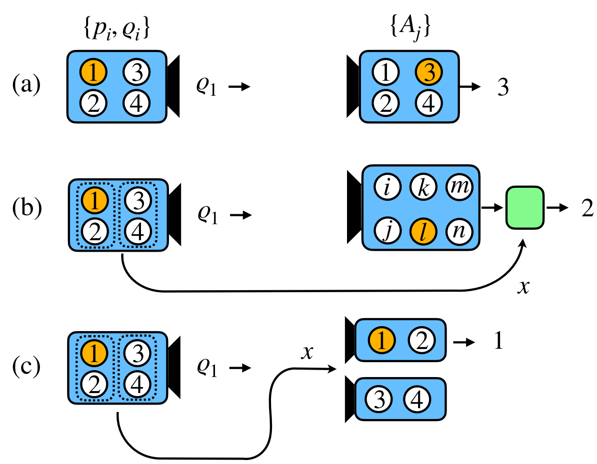

FIG. 6. Quantum state discrimination in different scenarios. (a) In minimum-error state discrimination the task is to guess the correct input label of the state with high probability. (b) In state discrimination with postinformation, the label x is reviled after the measurement. Here the additional information can be used only to postprocess the classical measurement result. (c) In state discrimination with prior information, the partition x of the label is received before the state, and hence a different measurement can be tailored to each subset x of the labels. Adapted from Carmeli, Heinosaari, and Toigo, 2018.

A similar but slightly different task is called minimum-error state discrimination with postmeasurement information (Ballester, Wehner, and Winter, 2008; Gopal and Wehner, 2010); see Fig. 6(b). Suppose that the index set I of the ensemble $\mathcal{E}$ is partitioned into nonempty disjoint sets I_x such that $\cup_x I_x = I$ and that the label x is revealed after the measurement of $\{A_a\}$ has been performed. This information cannot decrease the probability of correctly guessing the label a. However, the probability of success can increase if the label x is revealed prior to the measurement since one can tailor a separate measurement to each label x individually; see Fig. 6(c). Thus, one arrives at the conclusion that $p_{\text{guess}}(\mathcal{E}) \le p^{\text{post}}_{\text{guess}}(\mathcal{E}) \le p^{\text{prior}}_{\text{guess}}(\mathcal{E})$. Carmeli, Heinosaari, and Toigo (2018) proved that $p_{\text{guess}}(\mathcal{E}) = p^{\text{post}}_{\text{guess}}(\mathcal{E}) = p^{\text{prior}}_{\text{guess}}(\mathcal{E})$ if and only if the measurements that maximize the probability of success $p_{\text{guess}}(\mathcal{E})$ for each x are jointly measurable. This also shows that joint measurability can be understood in terms of state-discrimination games; namely, if the two scenarios in Figs. 6(b) and 6(c) are indistinguishable, measurements are compatible.

The relation between incompatible measurements and state-discrimination tasks can be made more precise by showing that whenever a set of measurements is incompatible, there is an instance of a state-discrimination task with prior information in which this set of measurements performs strictly better than any compatible one (Carmeli, Heinosaari, and Toigo, 2019; Oszmaniec and Biswas, 2019; Skrzypczyk, Šupić, and Cavalcanti, 2019; Uola *et al.*, 2019; Buscemi, Chitambar, and Zhou, 2020). More precisely, for any set of incompatible POVMs $\{A_{a|x}\}$, there is a state-discrimination task in which this set strictly outperforms any set of compatible measurements. The outperformance can be quantified by the incompatibility robustness $R_{\text{inc}}(A_{a|x})$, and we have

$$\sup_{\mathcal{E}} \frac{p_{\text{succ}}(A_{a|x}, \mathcal{E})}{\max_{O_{a|x}\in \text{JM}} p_{\text{succ}}(O_{a|x}, \mathcal{E})} = 1 + R_{\text{inc}}(A_{a|x}). \tag{48}$$

The state-discrimination task can be derived from the optimal solution of the incompatibility robustness SDP in Eq. (26). The connection between state-discrimination games and the incompatibility robustness has been used to experimentally verify the incompatibility of two-qubit measurements given by Smirne *et al.* (2022). We note that a similar result was already shown in the case of steering (Piani and Watrous, 2015). Here it was known that any steerable assemblage leads to a better performance in a suitably chosen *subchannel discrimination task* than in any unsteerable assemblage.

Comparing Eqs. (24) and (25), one sees that the incompatibility robustness and the incompatibility weight are similar by their definition. Therefore, it is natural to ask whether the incompatibility weight has a similar interpretation in terms of state-discrimination tasks. To that end, one first needs to define minimum-error state exclusion tasks with prior information, also known as *antidistinguishability* (Heinosaari and Kerppo, 2018). Such tasks were first formalized by Bandyopadhyay *et al.* (2014) and in the context of the Pusey-Barrett-Rudolph argument against a naive statistical interpretation of the wave function (Pusey, Barrett, and Rudolph, 2012).

The scenario in this task is similar to that in minimum-error state discrimination, with the difference being that one aims to maximize the probability of guessing a state that was not sent, that is, one minimizes the probability $p_{\rm succ}(A_{a|x},\mathcal{E})$ of guessing the state correctly. One then finds that (Uola, Bullock *et al.*, 2020) [see also Ducuara and Skrzypczyk (2020)]

$$\inf_{\mathcal{E}} \frac{p_{\rm succ}(A_{a|x},\mathcal{E})}{\min_{O_{a|x}\in {\rm JM}} p_{\rm succ}(O_{a|x},\mathcal{E})} = 1 - W_{\rm inc}(A_{a|x}), \qquad (49)$$

where the optimization is performed over those sets $\{O_{a|x}\}$ for which the left-hand side is finite. Finally, we note that the role of minimum-error state discrimination as a resource monotone and its connection to the robustness measure extends to the resource theory of single measurements; cf. Skrzypczyk and Linden (2019) and Guff *et al.* (2021).

2. Quantum random access codes

Random access codes (RACs) are an important class of classical communication tasks in which one party encodes a string of n classical bits $\mathbf{x} = (x_1, \ldots, x_n)$ into $m < n$ bits using some encoding strategy. Subsequently, the m bits are communicated to a receiver. The task of the receiver is then to recover, with a high probability of success, a randomly chosen bit x_j of the original string $\mathbf{x}$ using some decoding strategy. This is strongly related to the concept of information causality (Pawłowski *et al.*, 2009), which plays an important role in the foundations of quantum theory and in the problem of singling out quantum correlations from more general nonsignaling correlations; cf. Gallego *et al.* (2011).

The idea of sending quantum states instead of classical information goes back to the work of Wiesner (1983), who discussed it under the name of *conjugate coding*, and was later rediscovered by Ambainis *et al.* (2002) in the field of quantum finite automata. In quantum random access codes (QRACs) the sender encodes the string of n classical bits into a single d-level system using a CPTP map $\mathcal{E}(\mathbf{x})$, which is then called an (n,d) QRAC. The decoding is done by performing a measurement $\{A_{x_j|j}\}$ that depends on which bit x_j was chosen to be recovered; see also Fig. 7. The decoding was successful if the outcome is equal to the value of x_j. The average success probability is then given by

$$P_{\rm QRAC}(A_1,\ldots,A_n) = \frac{1}{nd^n}\sum_{\mathbf{x}} {\rm tr}[\mathcal{E}(\mathbf{x})(A_{x_1|1} + \cdots + A_{x_n|n})], \qquad (50)$$

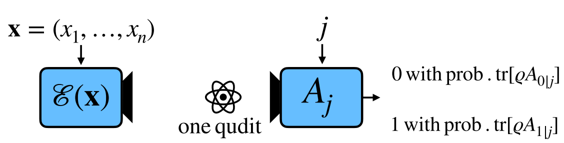

FIG. 7. Schematic of quantum random access codes. A sender holds n bits of information $\mathbf{x}$ and encodes it into a single qudit via the quantum channel $\mathcal{E}(\mathbf{x})$. The receiver wants to recover a random bit x_j from $\mathbf{x}$ by performing a measurement $\{A_{x_j|j}\}$ on the qudit.

where $\mathcal{E}(\mathbf{x})$ is the encoding map and $A_{x_j|j}$ are the measurement effects. When optimized over states, the optimal average success probability is given by $\bar{P}_{\rm QRAC}(A_1,\ldots,A_n) = (1/nd^n)\sum_{\mathbf{x}}\|(A_{x_1|1} + \cdots + A_{x_n|n})\|_\infty$, where $\|\cdot\|_\infty$ denotes the operator norm and the measurements are useful if the average success probability exceeds the classical bound, i.e., if $\bar{P}_{\rm QRAC}(A_1,\ldots,A_n) > P^{n,d}_{\rm RAC}$.

Carmeli, Heinosaari, and Toigo (2020) showed that a QRAC performs better than its classical counterpart only when incompatible measurements are used in the decoding step; see also Frenkel and Weiner (2015). For $n = 2$ the following results are known (Carmeli, Heinosaari, and Toigo, 2020).

(1) For any compatible pair of d-outcome measurements A_1 and A_2, it holds that $\bar{P}_{\rm QRAC}(A_1,A_2) \le P^{2,d}_{\rm RAC}$. The upper bound is tight.

This raises the question as to whether all incompatible measurements provide an advantage in QRACs. While this turns out not to be true in general, it is true in the following two instances.

(2) Let A_1 and A_2 be two sharp d-outcome measurements. $\bar{P}_{\rm QRAC}(A_1,A_2) \ge P^{2,d}_{\rm RAC}$, with equality attained if and only if A_1 and A_2 are compatible.

(3) Two unbiased qubit measurements A_1 and A_2 are incompatible if and only if they are useful for $(2,2)$ QRAC.

In the third case, the result follows directly from the fact that the average success probability is a function of the Busch criterion in Eq. (7). Furthermore, it was shown that there are pairs of biased qubit observables that are incompatible but nevertheless have $\bar{P}_{\rm QRAC}(A_1,A_2) < P^{2,2}_{\rm RAC}$, and thus do not provide an advantage over classical RACs.

Anwer *et al.* (2020) experimentally demonstrated quantum RACs to quantify the *degree of incompatibility*; see Busch, Lahti, and Werner (2014b). Another experimental implementation was reported by Foletto *et al.* (2020).

3. Distributed sampling

Going beyond the state-discrimination scenario in the Sec. IV.E.1, other scenarios have been identified where incompatible measurements play an important role. *Distributed sampling* refers to the task of Alice and Bob being able to sample from the probability distributions ${\rm tr}(\varrho_x B_{b|y})_{b,x,y}$, where ϱ_x is a quantum input of Alice and y is a classical input of Bob (Guerini, Quintino, and Aolita, 2019). If they share a perfect quantum communication channel, Alice could send her input state to Bob, who can then perfectly sample from the desired probability distribution. When Alice's communication to Bob is restricted to classical information, the most general strategy is that Alice performs a measurement $\{A_a\}$ on her input state ϱ_x and sends her result to Bob. Bob then outputs a classical variable b according to some response function $f(b|y,a)$. More precisely, the distributions that Alice and Bob can sample from are of the form $P(b|\varrho_x, y) = \sum {\rm tr}[\varrho_x A_a] f(b|y,a)$. One of the results of Guerini, Quintino, and Aolita (2019) was that a set of measurements is compatible if and only if the behavior $\{{\rm tr}[\varrho_x B_{b|y}]\}$ admits a distributed sampling realization. Moreover, such a sampling task can be used to obtain lower bounds to the incompatibility robustness, and thus quantify the degree of incompatibility of the implemented measurements.

V. FURTHER TOPICS AND APPLICATIONS

In this section we describe various topics related to quantum incompatibility. We start with the potential resource theory of incompatibility. We then review various other notions of incompatibility, such as the incompatibility of channels or other notions for measurements, such as complementarity or coexistence. Finally, we comment on the problem of joint measurability for the infinite-dimensional case, which is important for understanding the incompatibility of position and momentum.

A. Resource theory of incompatibility

Resource theories formalize the idea that certain operations or preparations require fewer resources than others. For instance, the preparation of separable states does not require any nonlocal operations, as local operations and classical communication (LOCC) are sufficient for their preparation. This is in stark contrast to entangled states, which require global operations for their preparation (Gühne and Tóth, 2009; Horodecki *et al.*, 2009). Parts of entanglement theory may be seen as an example of a resource theory in which separable states are deemed free, whereas entangled states are resourceful. A resource theory would then ask for sets of physically motivated monotones that decide whether or not a transformation between two resourceful states is possible, such as the transformation of resources by free operations like LOCC or the distillation of highly resourceful states. The resourceful states can be used to accomplish some task; for instance, they can be used for teleportation or quantum key distribution.

In the case of measurement incompatibility, the compatible sets of POVMs are deemed resourceless and the resourceful measurements are the incompatible sets. To define meaningful resource monotones one first needs to establish a notion of free operations. Heinosaari, Kiukas, and Reitzner (2015) considered as free operations preprocessing by quantum channels; i.e., every incompatibility monotone I needs to satisfy $I[\Lambda(A_1),\Lambda(A_2)] \leq I[A_1,A_2]$ for all unital CPTP maps Λ. It was shown that in such a scenario a scaled violation of the CHSH inequality (Wolf, Perez-Garcia, and Fernandez, 2009) and the noise robustness are such monotones. Guerini *et al.* (2017) and Skrzypczyk, Šupić, and Cavalcanti (2019) considered classical postprocessing as a free operation. In that case, the relevant monotones need to be nonincreasing under classical postprocessing. Skrzypczyk, Šupić, and Cavalcanti (2019) showed that the incompatibility robustness fulfills this property. Moreover, it was shown that state-discrimination games with postmeasurement information forms a complete set of operationally meaningful monotones, in the sense that a set of measurements $\{A_{a|x}\}$ can be transformed into a set $\{\tilde{A}_{a|x}\}$ by classical postprocessing if and only if $P_{\text{guess}}(A_{a|x},\mathcal{E}) \geq P_{\text{guess}}(\tilde{A}_{a|x},\mathcal{E})$ for all state-discrimination games $\mathcal{E}$.

From a physical point of view one is not necessarily constrained to choose between either preprocessing or postprocessing. Pusey (2015) considered CNDO operations as free operations, and it was shown that for two binary projective measurements $\{B_{a|x}\}$ and $\{\tilde{B}_{a|x}\}$ the first one can be converted to the second one by CNDO if and only if they are unitary equivalent. Buscemi, Chitambar, and Zhou (2020) considered both quantum preprocessing and conditional classical postprocessing. It was shown that $\{A_{a|x}\}$ can be transformed to $\{\tilde{A}_{a|x}\}$ via these operations if and only if $\{A_{a|x}\}$ performs at least as well as $\{\tilde{A}_{a|x}\}$ in all discrimination games with postinformation.

Finally, Styliaris and Zanardi (2019) studied the resource theory of measurement incompatibility relative to a basis. Here probabilities arise from measurements on states that are diagonal in a fixed basis, and one asks whether the resulting probability distributions can be converted by classical postprocessing. A connection between resource monotones and multivariate majorization conditions is shown.

B. Channel incompatibility

Incompatibility can also be formulated for quantum objects other than measurements, especially for channels, i.e., completely positive trace-preserving maps. We say that a set of quantum measurement devices (i.e., a set of POVMs and a set of channels) is compatible if there is a single joint measurement process (represented by an instrument) that simultaneously realizes all the POVMs as well as all the channels in the set. These definitions follow the model setup given by Heinosaari and Miyadera (2017); note that there the compatibility of a mixed set of POVMs and channels is seen as the compatibility of channels where the POVMs are replaced with appropriate measure-and-prepare channels. This generalizes the definition of joint measurability, as we later see. In the following, we first give a precise definition of channel incompatibility, then discuss two applications, the quantum marginal problem and the information-disturbance trade-off.

1. Formal definition of channel incompatibility

To formalize the previous idea, consider a set $\{A_{a|x}\}$ of POVMs on a system A and a set of channels Λ_y with the shared input system A and the possibly varying output systems B_y. Besides this, there are no further constraints on the relation between the set of POVMs and the set of channels.

Imagine an instrument $\mathcal{I} = \{\mathcal{I}_{\vec{a}}\}$ whose input system is A and output system is the composition of the systems B_y. Here $\vec{a}$ denotes the vector $(a_x)_x$ that can encode all the outcomes of the POVMs $\{A_{a|x}\}$. Furthermore, we can denote by $E_{a|x}$ the set of all such vectors where the xth component is fixed at a. This instrument can then be related to the previously discussed set of POVMs and set of channels in the following manner.

First, it may reproduce the POVMs if $\mathrm{tr}[\varrho A_{a|x}] = \sum_{\vec{a}\in E_{a|x}} \mathrm{tr}[\mathcal{I}_{\vec{a}}(\varrho)]$ holds for all input states ϱ, all indices x, and all values a of the xth POVM. Second, it may reproduce the channels in the sense that $\Lambda_y(\varrho) = \sum_{\vec{a}} \mathrm{tr}_{B_y^c}[\mathcal{I}_{\vec{a}}(\varrho)]$ for all input states ϱ and output labels y. Here B_y^c is the composition of all systems $B_{y'}$, where $y' \neq y$. If such an instrument $\mathcal{I}$

exists, then we can say that the POVMs $\{A_{a|x}\}$ and the channels Λ_y are contained in a single measurement setting. One can then also say that the POVMs and channels are *compatible*; otherwise, they are *incompatible*. Similar definitions were introduced by Heinosaari, Miyadera, and Reitzner (2014).

One can immediately see that if the set of the output labels y is empty (i.e., no channels are considered), then this definition corresponds to the joint measurability of $\{A_{a|x}\}$ with the joint POVM $\{M_{\vec{a}}\}$ defined through $\mathrm{tr}[\varrho M_{\vec{a}}] = \mathrm{tr}[\mathcal{I}_{\vec{a}}(\varrho)]$ for all ϱ. On the other hand, if no POVMs are considered, i.e., the set of labels x is empty, the compatibility of channels Λ_y is equivalent to the existence of a broadcasting channel (or joint channel) Γ with the input system A, and whose output system is the composition of the systems B_y such that $\Lambda_y(\varrho) = \mathrm{tr}_{B_y^c}[\Gamma(\varrho)]$ for all ϱ and y. This definition coincides then with those given by Heinosaari and Miyadera (2017) and Haapasalo (2019). This notion of channel compatibility can also be generalized to channels that may share the input system only partly and also have overlapping output systems (Hsieh, Lostaglio, and Acin, 2021).

2. Quantum marginal problem

Marginal problems are compatibility problems of quantum states: Starting at given states in different subsystems, one has to determine whether there is a global state from which all the subsystem states can be obtained as reduced states. This problem is also known as the N-representability problem (Coleman, 1963; Ruskai, 1969), and it remains a major problem in quantum chemistry (National Research Council, 1995).

In mathematical terms, one has a collection $\mathcal{A} \coloneqq \{A_j\}_{j\in J}$ of quantum systems where some subsets $X_i \subseteq \mathcal{A}$ with $i \in I$ are considered as subsystems. For each $i \in I$, there is a state ϱ_i given on the system X_i. The marginal problem associated with this setting asks whether there is a global state ϱ of the entire system $\mathcal{A}$ such that ϱ_i is the reduced state of ϱ for each $i \in I$, i.e., $\varrho_i = \mathrm{tr}_{X_i^c}[\varrho]$. In principle, this problem can be formulated as a SDP, but one often has additional constraints; for instance, the global state $\varrho = |\psi\rangle\langle\psi|$ is required to be pure or bosonic or fermionic symmetries must be respected. In this case, systematic approaches using algebraic geometry (Klyachko, 2004, 2006), generalized Pauli constraints (Castillo *et al.*, 2021), or hierarchies of SDPs (Yu *et al.*, 2021) have been developed; nevertheless, the problem remains difficult.

The central result of Haapasalo *et al.* (2021) was that marginal problems and compatibility questions can be identified with each other through the generalized channel-state dualism defined by the fixed margin ϱ_A; see Sec. IV.B and Eq. (44) for the exact form of the map. See also Proposition 12 of Plávala (2017) for the case of the canonical Choi map; i.e., the one where ϱ_A is in the maximally mixed state. In other words, one can devise the following formulation: A collection of $A \to B_i$ channels Λ_i is compatible if and only if the marginal problem involving the corresponding Choi states has a solution.

More explicitly, a tuple $\vec{\Lambda} = (\Lambda_i)_{i=1}^n$ of channels is compatible if and only if, for a full-rank state ϱ_A that can be freely chosen, the marginal problem involving the Choi states $S_{\varrho_A}(\Lambda_i)$ of the channels has a solution. On the other hand, the marginal problem involving a given tuple $\vec{\varrho} = (\varrho_i)_{i=1}^n$ of states on systems A and B_i with the fixed A margins $\mathrm{tr}_B[\varrho_i] = (\varrho_i)_A = \varrho_A$ has a solution if and only if the channels Λ_i from A to B_i such that $\varrho_i = S_{\varrho_A}(\Lambda_i)$ are compatible. Technically, for this direction we need ϱ_A to be invertible but, as pointed out by Haapasalo *et al.* (2021), we are free to suitably restrict system A to make ϱ_A invertible simultaneously, thereby not effectively altering the original marginal problem, so this is not a real restriction. See also Girard, Plávala, and Sikora (2021) for details on similar results.

The previous result enables the translation of results between the fields of compatibility and marginal problems. This was demonstrated by Haapasalo *et al.* (2021), who translated entropic conditions for the solvability of the marginal problem (Carlen, Lebowitz, and Lieb, 2013) to necessary conditions for the compatibility of channels. Moreover, known conditions for the compatibility of channels (Haapasalo, 2019) were used to characterize the solvability of marginal problems involving Bell-diagonal states. Solvability conditions for problems involving higher-dimensional qudit states with depolarizing noise were also obtained, and the quantitative perspective of the connection was discussed.

Since measurements can be seen as quantum-to-classical channels, joint-measurability questions can be recast as quantum marginal problems too. Indeed, measurements $\{A_{a|x}\}_a$ can be identified with a measure-and-prepare channel Λ_x, $\Lambda_x(\varrho) = \sum_a \mathrm{tr}[\varrho A_{a|x}]|a\rangle\langle a|$, where the output system is a register with the orthonormal basis $\{|a\rangle\}_a$. It is evident that such measurements are jointly measurable if and only if the corresponding channels Λ_x are compatible (Heinosaari and Miyadera, 2017). Moreover, a quick calculation shows that for a full-rank input state in its spectral decomposition $\varrho_A = \sum_m \lambda_m |m\rangle\langle m|$ one finds the Choi states

$$S_{\varrho_A}(\Lambda_x) = \sum_a \varrho_A^{1/2} A_{a|x}^T \varrho_A^{1/2} \otimes |a\rangle\langle a|, \tag{51}$$

where the transpose is taken with respect to the eigenbasis $\{|m\rangle\}_m$ of ϱ_A. Thus, joint-measurability questions can be identified with marginal problems involving block-diagonal states.

3. Information-disturbance trade-off relations

Here we discuss how the incompatibility between a single measurement and a specific quantum channel leads to an information-disturbance trade-off relation. Namely, we review the connection between the information gained in a measurement procedure and the inherent disturbance that it causes to the system (Heinosaari and Miyadera, 2013).

We now study the POVM $\{A_a\}_a$. To describe all the measurement processes describing this POVM, i.e., all the instruments $\{\mathcal{I}_a\}$ such that $\mathrm{tr}[\mathcal{I}_a(\varrho)] = \mathrm{tr}[\varrho A_a]$ for all input states ϱ, we fix a minimal Naimark dilation Δ for this POVM; see also Sec. III.C.3. This consists of an isometry J of the input system to a larger dilation system and a PVM $\{P_a\}_a$ on the larger system, such that $A_a = J^\dagger P_a J$. It can be shown that any instrument $\{\mathcal{I}_a\}$ measuring $\{A_a\}$ is of the form $\mathcal{I}_a(\varrho) = \Phi(J\varrho J^\dagger P_a)$ for some channel Φ from the dilation system to the physical postmeasurement system, where Φ has to obey the additional constraint $\Phi(J\varrho J^\dagger P_a) = \Phi(P_a J\varrho J^\dagger)$ for all input states ϱ and outcomes a (Haapasalo, Heinosaari, and Pellonpää, 2014). Using the condition on the channel Φ, it follows that one can freely replace Φ with $\Phi\circ\mathcal{L}_\Delta$, where $\mathcal{L}_\Delta$ is the Lüders channel $\mathcal{L}_\Delta(\sigma) = \sum_a P_a \sigma P_a$. Recalling the definition of compatibility of POVMs and channels given in Sec. V.B.1, this means that any channel Λ compatible with the single measurement $\{A_a\}$ is of the form $\Lambda = \Phi\circ\Lambda_\Delta$ for some channel Φ from the dilation system to the intended postmeasurement system, where Λ_Δ is the "maximal" channel $\Lambda_\Delta(\varrho) = \mathcal{L}_\Delta(J\varrho J^\dagger)$. This fact should be compared with the characterization given in Sec. III.C.3 for the POVMs compatible with a given POVM that is completely analogical.

Thus, the set $\mathfrak{C}_A$ of channels compatible with a fixed POVM $A = \{A_a\}$ has a simple structure: these channels are all obtained by concatenating any channels to the maximal channel Λ_Δ determined by any Naimark dilation Δ of A. Using the assumption that any dilation can be connected to a minimal one with an isometry (see the construction of Sec. V.C.5 for this well-known fact), it easily follows that, for any other dilation Δ', Λ_Δ and $\Lambda_{\Delta'}$ are equivalent in the sense that they are obtained from each other by channel concatenation. Thus, we may forget about the specific dilation and write $\Lambda_\Delta \eqqcolon \Lambda_A$.

Using this simple structure of channels compatible with a fixed POVM, Heinosaari and Miyadera (2013) proved a qualitative noise-disturbance trade-off relation: the noisier that the POVM $A = \{A_a\}$ is, the larger the set $\mathfrak{C}_A$ of channels compatible with A is. Specifically, given two POVMs $A = \{A_a\}$ and $B = \{B_b\}$ on the same system, there is a postprocessing $p(b|a)$ such that $B_b = \sum_a p(b|a) A_a$ if and only if $\mathfrak{C}_A \subseteq \mathfrak{C}_B$.

Owing to the simple structure of $\mathfrak{C}_A$ and $\mathfrak{C}_B$, the latter condition is equivalent to the existence of a channel Φ such that $\Lambda_A = \Phi\circ\Lambda_B$. Since a channel $\Lambda \in \mathfrak{C}_A$ describes the overall state transformation of the measurement of A, the larger the set $\mathfrak{C}_A$ (i.e., the "higher" the maximal channel Λ_A) is, the less the measurements of A can potentially disturb the system.

Thus, we can interpret the previous noise-disturbance relation in the following form: the more informative (i.e., less noisy) the measurement is, the more the measurement disturbs the system. An extreme example is provided by the trivial POVMs where A_a are all multiples of the identity operator whence $\mathfrak{C}_A$ is the set of all channels with the fixed input system, allowing complete nondisturbance. These POVMs provide absolutely no information on the system being measured. Another example is the case of rank-1 POVMs where $A_a = |\varphi_a\rangle\langle\varphi_a|$. As these POVMs are maximal in the postprocessing order, the sets $\mathfrak{C}_A$ are minimal. In fact, $\mathfrak{C}_A$ consists in this case of the measure-and-prepare channels of the form $\Lambda(\varrho) = \sum_a \mathrm{tr}[\varrho A_a]\sigma_a$ for some postmeasurement states σ_a.

C. Further features of quantum measurements

Joint measurability is the main measurement theoretical notion in this Colloquium due to its various applications in quantum information science. Here we discuss related concepts that have been used to grasp the counterintuive nature of quantum measurements.

1. Simulability of measurements

There are various operationally motivated ways to relax the notion of joint measurability. One possible generalization is to drop the assumption of having only one joint measurement. In other words, one can ask whether there is a simulation scheme that produces the statistics of n measurements from $m < n$ POVMs. For instance, Oszmaniec et al. (2017) defined the notion of measurement simulability as the existence of classical postprocessings $p(a|x, y, \lambda)$ and preprocessings (or classical randomness) $p(y|x)$ such that

$$A_{a|x} = \sum_{\lambda,y} p(y|x) p(a|x, y, \lambda) G_{\lambda|y}, \tag{52}$$

where $x \in \{1, \ldots, n\}$ and $y \in \{1, \ldots, m\}$. Joint measurability corresponds to simulability with one POVM, i.e., $m = 1$. This notion also has other interesting special cases, such as measurements simulable with PVMs (Oszmaniec et al., 2017), measurements simulable with POVMs with a fixed number of outcomes (Kleinmann and Cabello, 2016; Guerini et al., 2017; Shi and Tang, 2020), and sets of measurements simulable with a given number of POVMs (Guerini et al., 2017).

Hirsch et al. (2017) and Oszmaniec et al. (2017) used projective simulability to improve the known noise thresholds for the locality of a two-qubit Werner state. They then showed that any truly nonprojective measurement (i.e., a measurement not simulable with PVMs) provides an advantage in some minimum-error state-discrimination task over all projective simulable ones (Oszmaniec and Biswas, 2019; Uola et al., 2019). Finally, similar results for state discrimination with POVMs that cannot be simulated with measurements having a fixed number of outcomes were reported by Shi and Tang (2020).

2. Joint measurability on many copies

Another extension of joint measurability may be introduced using many copies of the given state (Carmeli et al., 2016). If one has two copies and two measurements, one can reproduce the statistics. In the simplest nontrivial scenario that deviates from joint measurability, one has a set of three POVMs $\{A_{a|x}\}$ and one is asked for a joint measurement on two copies of the original system together with postprocessings for which

$$\mathrm{tr}[A_{a|x}\varrho] = \sum_\lambda p(a|x, \lambda) \mathrm{tr}[(\varrho \otimes \varrho) G_\lambda] \tag{53}$$

holds for all quantum states ϱ. As an example, it was shown by Carmeli et al. (2016) that three orthogonal noisy qubit measurements $A_{\pm|i}(\mu) = (1/2)(\mathbb{1} \pm \mu\sigma_i)$ with $i = x, y, z$ have a joint measurement on two copies if and only if $0 \le \mu \le \sqrt{3}/2$. In other words, there are triplets with a two-copy joint measurement, although each involved pair of measurements is incompatible. The structure of the incompatibility structures arising from compatibility on many copies have also been analyzed in detail (Carmeli et al., 2016).

3. Compatibility of coarse-grained measurements and coexistence

Here we review special instances of incompatibility that arise from coarse graining of measurements and discuss their relation to complementarity. Coarse graining of a POVM corresponds intuitively to combining various POVM elements into a new one. More precisely, for a POVM $\{A_a\}$ we can consider disjoint subsets E_i of the set of outcomes and define the effects $A(E_i) = \sum_{a\in E_i} A_a$ and the coarse-grained POVM $\{A(E_i)\}_{i=1}^s$. Specifically, when $s = 2$, the coarse-grained two-outcome POVM $\{A(E), \mathbb{1} - A(E)\}$ is called a binarization of $\{A_a\}$.

A first notion of compatibility of coarse grainings is that of *coexistence*. This asks whether all yes-no questions, i.e., the set of all possible binarizations $\{A_x(E_x), \mathbb{1} - A_x(E_x)\}$ of given POVMs $\{A_{a|x}\}$, are jointly measurable with a single joint measurement. Here E_x runs over all subsets of outcomes of the input x. Note that for nonbinary measurements, the set of all binarizations consists of more POVMs than the original set. This question can equivalently be reformulated as follows: For any POVM $\{A_a\}$ one can define the range as the set of possible effects $A(E)$ in the previously mentioned notation. For a set of POVMs $\{A_{a|x}\}$ one can then ask whether the union of their ranges is contained in the range of a single POVM $\{G_\lambda\}$ (Ludwig, 1983; Lahti, 2003; Busch *et al.*, 2016); compare to Haapasalo, Pellonpää, and Uola (2015). Indeed, if $A_x(E_x) = \sum_{\lambda\in F} G_\lambda$, then $A_x(E_x) = \sum_\lambda p(\text{yes}|x, E_x, \lambda) G_\lambda$, where $p(\text{yes}|x, E_x, \lambda) = 1$ when $\lambda \in F$ and $p(\text{yes}|x, E_x, \lambda) = 0$ otherwise. Note that joint measurability implies coexistence, but the inverse implication does not hold in general (Reeb, Reitzner, and Wolf, 2013; Pellonpää, 2014b; Uola *et al.*, 2021). Similar notions can also be defined for more general coarse grainings, but they have not been considered in the literature.

A closely related concept was proposed by Saha *et al.* (2020) as the notion of *full complementarity.* This requires an incompatible pair $\{A_a\}$ and $\{B_b\}$ of POVMs to remain incompatible after arbitrary (nontrivial) coarse grainings. They also defined the more general *single-outcome complementarity* by demanding that for each pair (a, b) the POVMs $\{A_a, \mathbb{1} - A_a\}$ and $\{B_b, \mathbb{1} - B_b\}$ are incompatible. Notably here one is interested in the incompatibility of each coarse-grained pair separately. This is in contrast to coexistence.

Note that if one would ask incompatibility of all single-outcome binarizations at once, i.e., with a single joint measurement, one would get a scenario that can be relevant for correlation experiments in which one constructs multi-outcome measurements from two-outcome ones. Albeit important, this last concept has not been analyzed in the literature from a measurement theoretical perspective.

4. Complementarity

Complementarity was defined to be equivalent to nonjoint measurability in the previously mentioned works (Saha *et al.*, 2020), of which the notions of single-outcome complementarity and full complementarity are special cases. We note, however, that nonjoint measurability is not the standard notion of complementarity; see Kiukas *et al.* (2019). Traditionally, a pair of POVMs $\{A_a\}$ and $\{B_b\}$ is called *complementary* if all their measurements are mutually exclusive. Suppose that, for some outcomes a and b, there is a positive operator O such that

$$\mathrm{tr}[\varrho O] \le \mathrm{tr}[\varrho A_a], \qquad \mathrm{tr}[\varrho O] \le \mathrm{tr}[\varrho B_b] \tag{54}$$

for all states ϱ. If $O \neq 0$, then one gets nontrivial information on the measurement probabilities $\mathrm{tr}[\varrho A_a]$ and $\mathrm{tr}[\varrho B_b]$ in each state ϱ for which the probability $\mathrm{tr}[\varrho O]$ of the two-outcome measurement $\{O, \mathbb{1} - O\}$ is not zero. Hence, the minimal requirement for complementarity is for Eq. (54) to yield $O = 0$ for all a and b.

Complementary POVMs are incompatible since, for jointly measurable POVMs $\{A_a\}$ and $\{B_b\}$ with a joint POVM $\{M_{ab}\}$, Eq. (54) holds for $O = M_{ab}$, which is nonzero for some a and b. Furthermore, the PVMs $A_a = |\psi_a\rangle\langle\psi_a|$ and $B_b = |\varphi_b\rangle\langle\varphi_b|$ related to mutually unbiased bases satisfy the minimal requirement. However, it is easy to see in the qubit case that after applying an arbitrarily small amount of white noise the criterion is not fulfilled. Note that such measurements could still be complementary in the sense defined by Saha *et al.* (2020).

One can pose stronger conditions on complementarity by writing the condition (54) for general outcome sets E and F with effects $A(E) = \sum_{a\in E} A_a$ and $B(F) = \sum_{b\in F} B_b$ instead of only singletons $\{a\}$ and $\{b\}$. We must assume that the sets are such that $A(E) \neq \mathbb{1} \neq B(F)$. In this way, one gets many different definitions of complementarity related to different choices of the sets in Eq. (54). This works especially well for continuous POVMs in infinite dimensions. For example, the following pairs of POVMs related to the harmonic oscillator are complementary in the traditional sense: position momentum, position energy, momentum energy, number phase, and energy time (which is essentially the number-phase pair) (Kiukas *et al.*, 2019).

5. Retrieving measurements in the sequential scenario

As discussed in Sec. II.B, joint measurability and nondisturbance are not equivalent notions. However, jointly measurable pairs of POVMs allow a sequential measurement scenario that recovers the data of both POVMs. Namely, instead of measuring a POVM $\{A_a\}$ ($\{B_b\}$) on the first (second) time step, one can measure $\{A_a\}$ on the first step and a retrieving measurement $\{\tilde{B}_b\}$ on the second step; see Fig. 8. In general, the retrieving measurement acts on a larger Hilbert space. It is also possible to remain in the original Hilbert space if one is allowed to choose the second measurement, depending on the outcome of the first; i.e., one uses the output a of the first measurement as an input for the second one.

For a detailed description, recall the notation from Sec. III.C.3 where a POVM $\{A_a\}$ was considered with effects $A_a = \sum_{k=1}^{m_a} |d_{ak}\rangle\langle d_{ak}|$. Instead of the additive Naimark extension in Sec. III.C.3, we now consider a product or auxiliary form Naimark dilation (Preskill, 1998) by choosing Hilbert spaces $\mathcal{H}_n$ and $\mathcal{H}_m$ with bases $\{e_a\}_{a=1}^n$ and $\{f_k\}_{k=1}^m$, $m \ge \max\{m_a\}$, taking the tensor product $\mathcal{H}_n \otimes \mathcal{H}_m$ (instead of $\mathcal{H}_\oplus$), the projection operators $P'_a = |e_a\rangle\langle e_a| \otimes \mathbb{1}$, and the isometry $J' = \sum_{a=1}^n \sum_{k=1}^{m_a} |e_a \otimes f_k\rangle\langle d_{ak}|$ (Haapasalo and

Pellonpää, 2017b). Note that by defining an isometry $W \coloneqq \sum_{a,k} |e_a \otimes f_k\rangle\langle e_{ak}|$ we get $WJ = J'$ and $WP_a = P'_a W$, and the resulting Naimark extension $(\mathcal{H}_n \otimes \mathcal{H}_m, J', \{P'_a\})$ is typically not minimal (in which case, J should be unitary and $nm = \sum_a m_a$, i.e., $m_a \equiv m$). However, for any joint measurement $\{M_{ab}\}$ of $\{A_a\}$ and $\{B_b\}$ there is a possibly nonunique POVM $\{\tilde{B}_b\}$ of $\mathcal{H}_m$ such that $M_{ab} = J'^\dagger(|e_a\rangle\langle e_a| \otimes \tilde{B}_b)J'$ (Haapasalo and Pellonpää, 2017b). Specifically, we obtain the sequential measurement interpretation of the joint measurement. We can write $\mathrm{tr}[M_{ab}\varrho] = \mathrm{tr}[\mathcal{I}_a(\varrho)\tilde{B}_b]$, where we have expressed an instrument $\{\mathcal{I}_a\}$ in the following Stinespring form: $\mathcal{I}_a(\varrho) \coloneqq \mathrm{tr}_{\mathcal{H}_n}[J'\varrho J'^\dagger(|e_a\rangle\langle e_a| \otimes \mathbb{1})]$.

If we choose $m \geq d$ and define the isometries $J_a \coloneqq \sum_{k=1}^d |e_a \otimes f_k\rangle\langle \varphi_{ak}|$, we obtain an alternative form $\mathcal{I}_a(\varrho) = \Lambda_a(\sqrt{A_a}\varrho\sqrt{A_a})$, where $\Lambda_a(\varrho) \coloneqq \mathrm{tr}_{\mathcal{H}_n}[J_a \varrho J_a{}^\dagger]$ is a quantum channel. From here one can find the retrieving measurements on the original system by setting $\tilde{B}_{b|a} \coloneqq \Lambda_a^\dagger(\tilde{B}_b)$. For each a this forms a POVM. One further sees that the Lüders instrument is the least disturbing in the sense that after it the data of any POVM jointly measurable with $\{A_a\}$ can be retrieved. This fact was also found by Heinosaari and Miyadera (2015), who associated the minimal disturbance property, also called the universality, with the channel Λ_A of Sec. V.B.3.

For a further physical interpretation, consider the case $m = d$; see Fig. 9. We can then identify the system's Hilbert space with $\mathcal{H}_m$ and "extend" the isometry J' to a unitary operator U on $\mathcal{H}_n \otimes \mathcal{H}_m$ via $U(|\xi_0\rangle \otimes |\psi\rangle) \coloneqq J'|\psi\rangle$ for $|\psi\rangle \in \mathcal{H}_m$, where $|\xi_0\rangle \in \mathcal{H}_n$ is some fixed ancilla's unit vector (the so-called ready state). Now $\langle \mathcal{H}_n, \{Z_a\}_a, |\xi_0\rangle, U\rangle$ is called a measurement scheme (or measurement model) of $\{A_a\}$. Specifically, $\{Z_a\}$, $Z_a = |e_a\rangle\langle e_a|$, is the pointer PVM; see Busch, Lahti, and Mittelstaedt (1996).

This has a distinct physical meaning: Before the measurement, the initial state of the compound system is $|\xi_0\rangle\langle\xi_0| \otimes \varrho$ since one assumes that the probe (ancilla) and system are dynamically and probabilistically independent of each other. The measurement coupling U transforms the initial state into the final (entangled) state $S_\varrho \coloneqq U(|\xi_0\rangle\langle\xi_0| \otimes \varrho)U^\dagger = J'\varrho J'^\dagger$ that determines the subsystems' final states $\mathrm{tr}_{\mathcal{H}_m}[S_\varrho]$ and $\mathrm{tr}_{\mathcal{H}_n}[S_\varrho]$.

The probability reproducibility condition $\mathrm{tr}[\varrho A_a] = \mathrm{tr}[S_\varrho(Z_a \otimes \mathbb{1})] = \langle e_a|\mathrm{tr}_{\mathcal{H}_m}[S_\varrho]|e_a\rangle$ guarantees that the measurement outcome probabilities are reproduced in the distribution of the pointer values in the final probe state. The state $S_\varrho^a \coloneqq \mathrm{tr}[\varrho A_a]^{-1}(Z_a \otimes \mathbb{1})S_\varrho(Z_a \otimes \mathbb{1})$ can be interpreted as a conditional state under the condition $|e_a\rangle\langle e_a| \otimes \mathbb{1}$ (Cassinelli and Zanghì, 1983). One obtains the subsystem states $\varrho_a \coloneqq \mathrm{tr}_{\mathcal{H}_n}[S_\varrho^a]$ and $|e_a\rangle\langle e_a|$, where the latter is the state of the probe after the interaction on condition Z_a. The sequential interpretation of the joint probability distribution,

$$\mathrm{tr}[M_{ab}\varrho] = \mathrm{tr}[S_\varrho(Z_a \otimes \tilde{B}_b)] = \mathrm{tr}[\varrho A_a]\mathrm{tr}[\varrho_a \tilde{B}_b], \quad (55)$$

shows that the states ϱ_a give the probabilities for any subsequent measurement on the system. In addition, any joint POVM can be expressed in tensor product form.

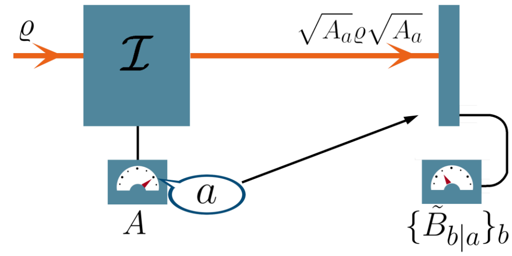

FIG. 8. Illustration for a joint measurement (a parent measurement) for the target POVMs A and B. As long as A and B are jointly measurable, by first performing the Lüders measurement of A and then, conditioned by the outcome a obtained in the A measurement, measuring a conditional POVM $\{\tilde{B}_{b|a}\}_b$ after the A measurement, this sequential scheme realizes a joint measurement of A and B.

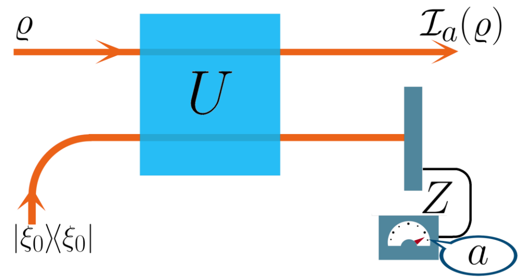

FIG. 9. The "auxiliary" Naimark dilation has a particular form that enables the following physical interpretation of an instrument $\mathcal{I}$ measuring the target POVM A (Ozawa, 1984). The system state is coupled with a probe (the auxiliary system in the dilation) in the initial state $|\xi_0\rangle\langle\xi_0|$. A unitary evolution mediated by the unitary U then takes place followed by a sharp pointer measurement Z of the probe, after which the probe is detached from the system.

D. Compatibility of continuous POVMs

Thus far we have explained the notions of incompatibility for the case of POVMs in finite-dimensional spaces with a finite set of outcomes. But, as mentioned in the Introduction, in the entire research program the case of position and momentum observables played a substantial role (Born and Jordan, 1925; Heisenberg, 1925). We now finally explain some basic facts about joint measurability for the case of "continuous" POVMs to demonstrate that most of the questions studied in this Colloquium are also relevant in the continuous case. However, as the continuous case is not our main focus, we merely outline how compatibility is defined in this setting and mention some generalizations of the previously presented results.

We now study a set $\{A_x\}$ of POVMs labeled by x, where each POVM has a continuous value space Ω_x. Here the allowed events or outcomes are measurable subsets of Ω_x. These POVMs operate in a possibly infinite-dimensional Hilbert space $\mathcal{H}$. We say that the set $\{A_x\}$ is jointly measurable if and only if there is a POVM G and conditional probability measures $p(\cdot|x,\lambda)$ for which

$$A_x(E) = \int p(E|x,\lambda)dG(\lambda) \quad (56)$$

for all measurable $E \subseteq \Omega_x$. Note that in Eq. (57) we present for any A_x a density with respect to a probability measure, giving a concrete way of evaluating the previously mentioned operator integral. Joint measurability can be equivalently defined by requiring there to be a parent POVM M from which A_x can be obtained as margins just as in the discrete case.

In addition, in the continuous case PVMs are jointly measurable only if they commute. From this we immediately see that the canonical position and momentum are not jointly measurable; see also our subsequent discussion on the quadrature observables. Position and momentum are also maximally incompatible in the sense that the addition of the maximum amount of trivial noise is required to make them compatible (Heinosaari et al., 2014). This quantification of incompatibility is similar to the incompatibility random robustness with the addition that the noise can consist of any POVMs whose effects are multiples of the identity operator.

In our previous discussions, such as those in Secs. III.C.3 and V.C.5, the Naimark extension played an important role, so we now also explain this for the continuous case. To begin, a general POVM A of the possibly infinite-dimensional Hilbert space $\mathcal{H}$ with a basis $\{|n\rangle\}$ can be written as

$$A(E) = \int_E \sum_{k=1}^{m_a} |d_{ak}\rangle\langle d_{ak}| d\mu(a), \tag{57}$$

where E is a measurable subset of outcomes, μ is a probability (or positive) distribution on the outcome space, the integral runs over all $a \in E$, and the $|d_{ak}\rangle$'s are generalized vectors (Hytönen, Pellonpää, and Ylinen, 2007). By defining $\mathcal{H}$-valued wave functions $|\psi_n\rangle$ as maps $a \mapsto |\psi_n(a)\rangle = \sum_{k=1}^{m_a} \langle d_{ak}|n\rangle |k\rangle$ and an isometry $J = \sum_n |\psi_n\rangle\langle n|$, one obtains a minimal Naimark dilation

$$\begin{aligned} A(E) &= J^\dagger P(E) J \\ &= \sum_{m,n} \langle \psi_m | P(E) | \psi_n \rangle |m\rangle\langle n| \\ &= \sum_{m,n} \int_E \langle \psi_m(a) | \psi_n(a) \rangle d\mu(a) |m\rangle\langle n|, \end{aligned} \tag{58}$$

where P is the "generalized position PVM" of the wave function space (a direct integral) defined via $[P(E)\psi](a) = \psi(a)$ when $a \in E$, and 0 otherwise.

If a POVM B is jointly measurable with A, then $B(F) = J^\dagger \tilde{B}(F) J$, where the POVM $\tilde{B}$ commutes with P; i.e., it is of the form $[\tilde{B}(F)\psi](a) = \tilde{B}_a(F)\psi(a)$, where $\tilde{B}_a$ is a POVM acting in the m_a-dimensional subspace of $\mathcal{H}$ (Pellonpää, 2014b). Specifically, if each multiplicity is given by $m_a = 1$, meaning that A is of rank 1 (Pellonpää, 2014a), then the operators $\tilde{B}_a(F)$ are conditional probabilities, say, $p(F|a)$, and we have $B(F) = \int p(F|a) dA(a)$, showing that B is a classical postprocessing or smearing of A. This is in line with the finite-dimensional result in Sec. III.C.3. Indeed, in the discrete case μ is the counting measure such that all integrals reduce to sums: $A(E) = \sum_{a \in E} A_a$, where $A_a = \sum_{k=1}^{m_a} |d_{ak}\rangle\langle d_{ak}| = \sum_{m,n} \langle \psi_m | P_a | \psi_n \rangle |m\rangle\langle n| = \sum_{m,n} \langle \psi_m(a) | \psi_n(a) \rangle |m\rangle\langle n|$ and $P_a = |e_a\rangle\langle e_a| \otimes \mathbb{1}$, $\psi_n = \sum_a |e_a \otimes \psi_n(a)\rangle = \sum_{a,k} |e_a\rangle \otimes \langle d_{ak}|n\rangle |k\rangle$. Now the wave function space is simply a direct sum $\mathcal{H}_\oplus = \bigoplus_a \mathcal{H}_a$ where each $\mathcal{H}_a$ is spanned by the vectors $|e_a\rangle \otimes |k\rangle$, $k = 1, 2, \ldots, m_a$. Furthermore, the previous operators $\tilde{B}(F)$ are of the "diagonal block form"; i.e., any $\tilde{B}_a(F)$ is an operator of $\mathcal{H}_a$.

For example, in the case of a covariant phase POVM (Holevo, 1982; Busch et al., 2016)

$$\Phi(E) = \sum_{m,n=0}^{\infty} \langle \eta_m | \eta_n \rangle \int_E e^{i(m-n)\theta} \frac{d\theta}{2\pi} |m\rangle\langle n|, \quad E \subseteq [0, 2\pi), \tag{59}$$

where the $|\eta_n\rangle$'s are unit vectors that span a d-dimensional space, one sees that $m_\theta = d$, $|\psi_n(\theta)\rangle = |\eta_n\rangle e^{-in\theta}$, and any jointly measurable POVM of Φ can be written as

$$B(F) = \sum_{m,n=0}^{\infty} \int_0^{2\pi} \langle \eta_m | \tilde{B}_\theta(F) | \eta_n \rangle e^{i(m-n)\theta} \frac{d\theta}{2\pi} |m\rangle\langle n|. \tag{60}$$

For instance, if $\tilde{B}_\theta = \tilde{B}$ does not depend on θ, we get a smeared number observable $B(F) = \sum_{n=0}^{\infty} \langle \eta_n | \tilde{B}(F) | \eta_n \rangle \times |n\rangle\langle n|$. Or, if Φ is the rank-1 canonical phase (Lahti and Pellonpää, 2000), i.e., any $|\eta_n\rangle = |0\rangle$ and $d = 1$, Eq. (60) reduces to $B(F) = p(F)\mathbb{1}$, with $p(F) = \langle 0 | \tilde{B}(F) | 0 \rangle$, which is a trivial smearing of both the canonical phase and the sharp number. In conclusion, one cannot measure the canonical phase and the number together [actually, they are complementary observables (Lahti, Pellonpää, and Schultz, 2017)], but after suitable smearings they become compatible.

Similarly, the rotated quadratures $Q_\theta = Q\cos\theta + P\sin\theta$ are of rank 1, so they are not jointly measurable [here $(Q\psi)(x) = x\psi(x)$ and $(P\psi)(x) = -i\hbar d\psi(x)/dx$ or, in Dirac's notation, $\langle x|Q|\psi\rangle = x\langle x|\psi\rangle$ and $\langle x|P|\psi\rangle = -i\hbar\partial_x\langle x|\psi\rangle$, are the position and momentum operators]. However, their smeared versions have joint measurements; see Busch et al. (2016).

Finally, many of the results and interpretations of joint measurability presented in this Colloquium generalize to the continuous setting. For example, the results of Sec. IV.E.1 on the advantage that incompatible measurements provide in state-discrimination tasks can be generalized to the continuous variable setting, as done by Kuramochi (2020), who found that the incompatibility robustness still quantifies the advantage. Here we note that the discrete version of the result relies on the SDP formulation of joint measurability, but such a formulation does not exist for the continuous case. This leads to the use of more involved techniques through limit procedures; see Kuramochi (2020) for details. Moreover, the connection between steering and measurement incompatibility detailed in Sec. IV.B is generalized to continuous measurements and infinite-dimensional Hilbert spaces by Kiukas et al. (2017). Furthermore, the W measure of Sec. III.C.2 bears similarity to the continuous variable s parametrized quasiprobability distributions (Cahill and Glauber, 1969). The connection between these distributions and joint measurability was studied by Pellonpää (2001) and Rahimi-Keshari et al. (2021). Pellonpää (2001) showed that between the Wigner and the Q function these distributions relate to operator-valued measures with possibly nonpositive semidefinite elements, whose marginals are the smeared position and momentum

TABLE I. Glossary summarizing different types of measurements that were discussed in this Colloquium.

Term	Definition
Compatible measurements	Can be simultaneously classically postprocessed from a single measurement; see Eq. (3).
Incompatible measurements	Measurements that are not compatible.
Joint measurement	The measurement from which compatible ones can be post-processed via Eq. (3).
Parent (or mother) POVM	A joint measurement that is of the marginal form in Eq. (4).
Complementary measurements	Two POVMs are complementary if sufficiently many pairs of their effects are mutually exclusive.
Unbiased measurement	Produces a uniform probability distribution when measured on the maximally mixed state.
Nondisturbing measurements	A POVM is said to be nondisturbing with respect to another POVM if there exists a sequential implementation in which neglecting the outcome of the first measurement does not affect the statistics of the second measurement.
Pretty good measurement	Performs fairly well (but not optimally) in state-discrimination tasks where states are roughly of equal probability and almost orthogonal.
Retrieving measurement	A measurement that is used to implement compatible POVMs in a sequential order; see Sec. V.C.5.
Simulable measurements	Generalized notion of measurement compatibility where statistics of measurements are simulated by fewer measurements; see Sec. V.C.1.
Coarse-grained measurement	Any measurement that is obtained by binning the outcomes of a measurement.

measurements. In the special case of the Q function, one gets a joint measurement for noisy position and momentum measurements. However, not all of the results obtained for discrete POVMs can be extended to the continuous setting. For example, the fact that for discrete POVMs A and B we find that A is a postprocessing of B if and only if any channel compatible with B is also compatible with A has not yet been established in the continuous case, although it makes operational sense as the core of the information-disturbance tradeoff presented in Sec. V.B.3.

E. Glossary

A glossary summarizing different types of measurements and their definitions can be found in Table I.

VI. CONCLUSION

The puzzling properties of quantum measurements have sparked intense discussions among scientists for nearly 100 years. This has led to many interesting research results, and it has taken some time for key concepts of quantum measurement theory, such as the notion of POVMs and the notion of joint measurability, to emerge and find widespread applications. In this Colloquium, we explained the incompatibility of measurements, highlighting the connections to information processing.

One may desire concepts like POVMs, instruments, and joint measurability to become standard knowledge on quantum mechanics in the physics community in the future. This is motivated by the fact that an operational view on quantum mechanics combined with elements of information theory is becoming more and more standard in physics. For instance, many new textbooks and university courses favor this viewpoint, showing that the method of teaching and understanding quantum mechanics is changing. In addition, novel applications of joint measurability may be found.

There are several interesting open problems connected with the incompatibility and joint measurability of generalized measurements. The following list provides a small selection.

- It has been shown that not all incompatible measurements can lead to Bell nonlocality. Which additional properties of measurements are required for this?
- As we have seen, incompatibility can be quantified by different figures of merit. The following questions arise: Which are the most incompatible measurements, and what are they useful for?
- There are other concepts to grasp the nature of measurement in quantum mechanics, such as coexistence and unbiasedness. What are their applications in quantum information processing?
- Some incompatible measurements do not provide an advantage in QRACs, so is there a stronger form of incompatibility that is necessary and sufficient for QRACs?
- Another open problem is to clarify the role of incompatible measurements in quantum metrology, where measurements are used to characterize one or more parameters of a quantum state with high precision.
- One open direction is to investigate the incompatibility properties of quantum measurements that act on many particles, such as on two or more qubits. In general, it would be interesting to characterize the measurement resources that are needed for generating nonclassical effects in quantum networks.

In addition, with the progress of experimental techniques, complex measurements with interesting incompatibility features may also become an available resource in practical implementations. This will finally lead to further applications of the theory presented in this Colloquium.

ACKNOWLEDGMENTS

We thank Alastair Abbott, Konstantin Beyer, Nicolas Brunner, Costantino Budroni, Tom Bullock, Paul Busch, Sébastien Designolle, Teiko Heinosaari, Benjamin D. M. Jones, Jukka Kiukas, Matthias Kleinmann, Pekka Lahti, Fabiano Lever, Kimmo Luoma, Nikolai Miklin, Takayuki Miyadera, Tobias Moroder, H. Chau Nguyen, Michał Oszmaniec, Marco Piani, Martin Plávala, Jussi Schultz, Michal Sedlák, Jiangwei Shang, Paul Skrzypczyk, Walter T. Strunz, Ivan Šupić, Armin Tavakoli, Giuseppe Vitagliano, Reinhard F. Werner, Zhen-Peng Xu, Kari Ylinen, Xiao-Dong Yu, and Mario Ziman for collaborations and discussions on the topic. We are also thankful for the useful comments given by Sébastien Designolle, Huan-Yu Ku, Peter Morgan, Martin Plávala, Marco Túlio Quintino, and three anonymous referees on an earlier version of the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; Projects No. 447948357 and No. 440958198), the Sino-German Center for Research Promotion (Project No. M-0294), the ERC (Consolidator Grant No. 683107/TempoQ), the DAAD, the Austrian Science Fund (FWF; Grant No. P 32273-N27), and the Swiss National Science Foundation (NCCR, SwissMAP, and Ambizione Grant No. PZ00P2-202179). This research is supported by the National Research Foundation, Singapore and A*STAR under its CQT bridging grant.

REFERENCES

Ali, S. T., C. Carmeli, T. Heinosaari, and A. Toigo, 2009, Found. Phys. **39**, 593.

Almeida, M. L., S. Pionio, J. Barrett, G. Tóth, and A. Acín, 2007, Phys. Rev. Lett. **99**, 040403.

Ambainis, A., A. Nayak, A. Ta-Shma, and U. Vazirani, 2002, J. ACM **49**, 496.

Andersson, E., S. M. Barnett, and A. Aspect, 2005, Phys. Rev. A **72**, 042104.

Andrejic, N., and R. Kunjwal, 2020, Phys. Rev. Res. **2**, 043147.

Anwer, H., S. Muhammad, W. Cherifi, N. Miklin, A. Tavakoli, and M. Bourennane, 2020, Phys. Rev. Lett. **125**, 080403.

ApS, M., 2021, computer code MOSEK, version 9.3.

Ballentine, L., 2014, arXiv:1402.5689.

Ballester, M. A., S. Wehner, and A. Winter, 2008, IEEE Trans. Inf. Theory **54**, 4183.

Bandyopadhyay, S., R. Jain, J. Oppenheim, and C. Perry, 2014, Phys. Rev. A **89**, 022336.

Barnett, S. M., and S. Croke, 2009, Adv. Opt. Photonics **1**, 238.

Barrett, J., 2002, Phys. Rev. A **65**, 042302.

Bavaresco, J., M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, 2017, Phys. Rev. A **96**, 022110.

Belinfante, F. J., 1973, *A Survey of Hidden-Variables Theories*, 1st ed., Monographs in Natural Philosophy (Elsevier, New York).

Bell, J. S., 1964, Physics **1**, 195.

Beltrametti, E. G., and S. Bugajski, 1995, J. Phys. A **28**, 3329.

Bene, E., and T. Vértesi, 2018, New J. Phys. **20**, 013021.

Bengtsson, I., 2007, AIP Conf. Proc. **889**, 40.

Bennett, C. H., and G. Brassard, 2014, Theor. Comput. Sci. **560**, 7.

Born, M., and P. Jordan, 1925, Z. Phys. **34**, 858.

Boyd, S., and L. Vandenberghe, 2004, *Convex Optimization* (Cambridge University Press, Cambridge, England).

Brandão, F. G. S. L., 2005, Phys. Rev. A **72**, 022310.

Brougham, T., and E. Andersson, 2007, Phys. Rev. A **76**, 052313.

Brunner, N., D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, 2014, Rev. Mod. Phys. **86**, 419.

Budroni, C., A. Cabello, O. Gühne, M. Kleinmann, and J.-Å. Larsson, 2022, Rev. Mod. Phys. **94**, 045007.

Budroni, C., G. Vitagliano, G. Colangelo, R. J. Sewell, O. Gühne, G. Tóth, and M. W. Mitchell, 2015, Phys. Rev. Lett. **115**, 200403.

Bullock, T., and P. Busch, 2018, J. Phys. A **51**, 283001.

Buscemi, F., E. Chitambar, and W. Zhou, 2020, Phys. Rev. Lett. **124**, 120401.

Busch, P., 1986, Phys. Rev. D **33**, 2253.

Busch, P., P. Lahti, and P. Mittelstaedt, 1996, *The Quantum Theory of Measurement*, Lecture Notes in Physics Monographs Vol. 2 (Springer, New York).

Busch, P., P. Lahti, J.-P. Pellonpää, and K.Ylinen, 2016, *Quantum Measurement*, Theoretical and Mathematical Physics (Springer, New York).

Busch, P., P. Lahti, and R. F. Werner, 2014a, Rev. Mod. Phys. **86**, 1261.

Busch, P., P. Lahti, and R. F. Werner, 2014b, Phys. Rev. A **89**, 012129.

Busch, P., and H.-J. Schmidt, 2010, Quantum Inf. Process. **9**, 143.

Cabello, A., 2008, Phys. Rev. Lett. **101**, 210401.

Cahill, K. E., and R. J. Glauber, 1969, Phys. Rev. **177**, 1857.

Carlen, E. A., J. L. Lebowitz, and E. H. Lieb, 2013, J. Math. Phys. (N.Y.) **54**, 062103.

Carmeli, C., T. Heinosaari, D. Reitzner, J. Schultz, and A. Toigo, 2016, Mathematics **4**, 54.

Carmeli, C., T. Heinosaari, and A. Toigo, 2018, Phys. Rev. A **98**, 012126.

Carmeli, C., T. Heinosaari, and A. Toigo, 2019, Phys. Rev. Lett. **122**, 130402.

Carmeli, C., T. Heinosaari, and A. Toigo, 2020, Europhys. Lett. **130**, 50001.

Cassinelli, G., and N. Zanghì, 1983, Nuovo Cimento Soc. Ital. Fis. **73B**, 237.

Castillo, F., J.-P. Labbé, J. Liebert, A. Padrol, E. Philippe, and C. Schilling, 2021, arXiv:2105.06459.

Cavalcanti, D., and P. Skrzypczyk, 2016, Phys. Rev. A **93**, 052112.

Cavalcanti, D., and P. Skrzypczyk, 2017, Rep. Prog. Phys. **80**, 024001.

Chen, S.-L., C. Budroni, Y.-C. Liang, and Y.-N. Chen, 2016, Phys. Rev. Lett. **116**, 240401.

Chen, S.-L., N. Miklin, C. Budroni, and Y.-N. Chen, 2021, Phys. Rev. Res. **3**, 023143.

Chen, Z., X. Ye, and S. Fei, 2017, Sci. Rep. **7**, 15822.

Clauser, J. F., M. A. Horne, A. Shimony, and R. A. Holt, 1969, Phys. Rev. Lett. **23**, 880.

Clauser, J. F., M. A. Horne, A. Shimony, and R. A. Holt, 1970, Phys. Rev. Lett. **24**, 549(E).

Clemente, L., and J. Kofler, 2015, Phys. Rev. A **91**, 062103.

Clemente, L., and J. Kofler, 2016, Phys. Rev. Lett. **116**, 150401.

Coleman, A. J., 1963, Rev. Mod. Phys. **35**, 668.

Cycon, H., and K. Hellwig, 1977, J. Math. Phys. (N.Y.) **18**, 1154.

Dall'Arno, M., G. M. D'Ariano, and M. F. Sacchi, 2011, Phys. Rev. A **83**, 062304.

Davies, E. B., 1976, *Quantum Theory of Open Systems* (Academic Press, London).

Davies, E. B., and J. T. Lewis, 1970, Commun. Math. Phys. **17**, 239.

Designolle, S., M. Farkas, and J. Kaniewski, 2019, New J. Phys. **21**, 113053.

Designolle, S., P. Skrzypczyk, F. Fröwis, and N. Brunner, 2019, Phys. Rev. Lett. **122**, 050402.

Designolle, S., V. Srivastav, R. Uola, N. H. Valencia, W. McCutcheon, M. Malik, and N. Brunner, 2021, Phys. Rev. Lett. **126**, 200404.

Designolle, S., R. Uola, K. Luoma, and N. Brunner, 2021, Phys. Rev. Lett. **126**, 220404.

Ducuara, A. F., and P. Skrzypczyk, 2020, Phys. Rev. Lett. **125**, 110401.

Emary, C., 2017, Phys. Rev. A **96**, 042102.
Emary, C., N. Lambert, and F. Nori, 2014, Rep. Prog. Phys. **77**, 016001.
Fine, A., 1982, Phys. Rev. Lett. **48**, 291.
Foletto, G., L. Calderaro, G. Vallone, and P. Villoresi, 2020, Phys. Rev. Res. **2**, 033205.
Frenkel, P. E., and M. Weiner, 2015, Commun. Math. Phys. **340**, 563.
Gallego, R., and L. Aolita, 2015, Phys. Rev. X **5**, 041008.
Gallego, R., L. E. Würflinger, A. Acín, and M. Navascués, 2011, Phys. Rev. Lett. **107**, 210403.
Gärtner, B., and J. Matoušek, 2012, *Approximation Algorithms and Semidefinite Programming* (Springer-Verlag, Berlin).
George, R. E., L. M. Robledo, O. J. E. Maroney, M. S. Blok, H. Bernien, M. L. Markham, D. J. Twitchen, J. J. L. Morton, G. A. D. Briggs, and R. Hanson, 2013, Proc. Natl. Acad. Sci. U.S.A. **110**, 3777.
Girard, M., M. Plávala, and J. Sikora, 2021, Nat. Commun. **12**, 2129.
Gisin, N., G. Ribordy, W. Tittel, and H. Zbinden, 2002, Rev. Mod. Phys. **74**, 145.
Gopal, D., and S. Wehner, 2010, Phys. Rev. A **82**, 022326.
Grant, M., and S. Boyd, 2020, computer code CVX, version 2.2, http://cvxr.com/cvx.
Guerini, L., J. Bavaresco, M. Terra Cunha, and A. Acín, 2017, J. Math. Phys. (N.Y.) **58**, 092102.
Guerini, L., M. T. Quintino, and L. Aolita, 2019, Phys. Rev. A **100**, 042308.
Guff, T., N. A. McMahon, Y. R. Sanders, and A. Gilchrist, 2021, J. Phys. A **54**, 225301.
Gühne, O., and G. Tóth, 2009, Phys. Rep. **474**, 1.
Haapasalo, E., 2015, J. Phys. A **48**, 255303.
Haapasalo, E., 2019, Ann. Inst. Henri Poincaré **20**, 3163.
Haapasalo, E., T. Heinosaari, and J.-P. Pellonpää, 2014, Rev. Math. Phys. **26**, 1450002.
Haapasalo, E., T. Kraft, N. Miklin, and R. Uola, 2021, Quantum **5**, 476.
Haapasalo, E., and J.-P. Pellonpää, 2017a, Rev. Math. Phys. **29**, 1750020.
Haapasalo, E., and J.-P. Pellonpää, 2017b, J. Math. Phys. (N.Y.) **58**, 122104.
Haapasalo, E., J.-P. Pellonpää, and R. Uola, 2015, Lett. Math. Phys. **105**, 661.
Hausladen, P., and W. K. Wootters, 1994, J. Mod. Opt. **41**, 2385.
Heinosaari, T., and O. Kerppo, 2018, J. Phys. A **51**, 365303.
Heinosaari, T., J. Kiukas, and D. Reitzner, 2015, Phys. Rev. A **92**, 022115.
Heinosaari, T., and T. Miyadera, 2013, Phys. Rev. A **88**, 042117.
Heinosaari, T., and T. Miyadera, 2015, Phys. Rev. A **91**, 022110.
Heinosaari, T., and T. Miyadera, 2017, J. Phys. A **50**, 135302.
Heinosaari, T., T. Miyadera, and D. Reitzner, 2014, Found. Phys. **44**, 34.
Heinosaari, T., T. Miyadera, and M. Ziman, 2016, J. Phys. A **49**, 123001.
Heinosaari, T., D. Reitzner, and P. Stano, 2008, Found. Phys. **38**, 1133.
Heinosaari, T., J. Schultz, A. Toigo, and M. Ziman, 2014, Phys. Lett. A **378**, 1695.
Heinosaari, T., and M. Wolf, 2010, J. Math. Phys. (N.Y.) **51**, 092201.
Heisenberg, W., 1925, Z. Phys. **33**, 879.
Heisenberg, W., 1927, Z. Phys. **43**, 172.
Helstrom, C. W., 1969, J. Stat. Phys. **1**, 231.
Helstrom, C. W., 1976, *Quantum Detection and Estimation Theory* (Academic Press, New York).
Hirsch, F., M. Quintino, and N. Brunner, 2018, Phys. Rev. A **97**, 012129.
Hirsch, F., M. T. Quintino, T. Vértesi, M. Navascués, and N. Brunner, 2017, Quantum **1**, 3.
Holevo, A. S., 1982, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam).
Holevo, A. S., 1998, J. Math. Phys. (N.Y.) **39**, 1373.
Holevo, A. S., 2012, Probl. Inf. Transm. **48**, 1.
Horodecki, P., Ł. Rudnicki, and K. Życzkowski, 2020, arXiv:2002.03233.
Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki, 2009, Rev. Mod. Phys. **81**, 865.
Hsieh, C.-Y., M. Lostaglio, and A. Acin, 2021, arXiv:2102.10926.
Huffman, E., and A. Mizel, 2017, Phys. Rev. A **95**, 032131.
Hytönen, T., J.-P. Pellonpää, and K. Ylinen, 2007, J. Math. Anal. Appl. **336**, 1287.
Jae, J., K. Baek, J. Ryu, and J. Lee, 2019, Phys. Rev. A **100**, 032113.
Kennard, E. H., 1927, Z. Phys. **44**, 326.
Kiukas, J., C. Budroni, R. Uola, and J.-P. Pellonpää, 2017, Phys. Rev. A **96**, 042331.
Kiukas, J., P. Lahti, J.-P. Pellonpää, and K. Ylinen, 2019, Found. Phys. **49**, 506.
Kleinmann, M., and A. Cabello, 2016, Phys. Rev. Lett. **117**, 150401.
Klyachko, A., 2004, arXiv:quant-ph/0409113.
Klyachko, A. A., 2006, J. Phys. Conf. Ser. **36**, 72.
Knee, G. C., K. Kakuyanagi, M.-C. Yeh, Y. Matsuzaki, H. Toida, H. Yamaguchi, S. Saito, A. J. Leggett, and W. J. Munro, 2016, Nat. Commun. **7**, 13253.
Knee, G. C., *et al.*, 2012, Nat. Commun. **3**, 606.
Kochen, S., and E. P. Specker, 1967, J. Math. Mech. **17**, 59.
Ku, H.-Y., N. Lambert, F.-J. Chan, C. Emary, Y.-N. Chen, and F. Nori, 2020, npj Quantum Inf. **6**, 98.
Kunjwal, R., C. Heunen, and T. Fritz, 2014, Phys. Rev. A **89**, 052126.
Kuramochi, Y., 2020, arXiv:2002.03504.
Lahti, P., 2003, Int. J. Theor. Phys. **42**, 893.
Lahti, P., and J.-P. Pellonpää, 2000, J. Math. Phys. (N.Y.) **41**, 7352.
Lahti, P., J.-P. Pellonpää, and J. Schultz, 2017, J. Phys. A **50**, 375301.
Leggett, A. J., and A. Garg, 1985, Phys. Rev. Lett. **54**, 857.
Lewenstein, M., and A. Sanpera, 1998, Phys. Rev. Lett. **80**, 2261.
Li, C.-M., N. Lambert, Y.-N. Chen, G.-Y. Chen, and F. Nori, 2012, Sci. Rep. **2**, 885.
Liang, Y.-C., R. W. Spekkens, and H. M. Wiseman, 2011, Phys. Rep. **506**, 1.
Loulidi, F., and I. Nechita, 2021, J. Math. Phys. (N.Y.) **62**, 042205.
Ludwig, G., 1983, *Foundations of Quantum Mechanics*, Vol. I (Springer-Verlag, Berlin).
Mermin, N. D., 1990, Phys. Rev. Lett. **65**, 3373.
Moroder, T., O. Gittsovich, M. Huber, and O. Gühne, 2014, Phys. Rev. Lett. **113**, 050404.
Naimark, M. A., 1940, Izv. Akad. Nauk SSSR, Ser. Fiz.-Mat. Tekh. Nauk **4**, 277.
National Research Council, 1995, *Mathematical Challenges from Theoretical/Computational Chemistry* (National Academies Press, Washington, DC).
Nguyen, H. C., S. Designolle, M. Barakat, and O. Gühne, 2020, arXiv:2003.12553.
Nguyen, H. C., and O. Gühne, 2020a, Phys. Rev. A **101**, 042125.
Nguyen, H. C., and O. Gühne, 2020b, Phys. Rev. Lett. **125**, 230402.
Nguyen, H. C., H.-V. Nguyen, and O. Gühne, 2019, Phys. Rev. Lett. **122**, 240401.
Oszmaniec, M., and T. Biswas, 2019, Quantum **3**, 133.
Oszmaniec, M., L. Guerini, P. Wittek, and A. Acín, 2017, Phys. Rev. Lett. **119**, 190501.

Ozawa, M., 1984, J. Math. Phys. (N.Y.) **25**, 79.
Pal, R., and S. Ghosh, 2011, J. Phys. A **44**, 485303.
Pawłowski, M., T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, 2009, Nature (London) **461**, 1101.
Pellonpää, J.-P., 2001, J. Phys. A **34**, 7901.
Pellonpää, J.-P., 2013a, J. Phys. A **46**, 025302.
Pellonpää, J.-P., 2013b, J. Phys. A **46**, 025303.
Pellonpää, J.-P., 2014a, Found. Phys. **44**, 71.
Pellonpää, J.-P., 2014b, J. Phys. A **47**, 052002.
Peres, A., 1990, Phys. Lett. A **151**, 107.
Peres, A., 1995, *Quantum Theory: Concepts and Methods* (Kluwer Academic Publishers, Dordrecht).
Piani, M., 2015, J. Opt. Soc. Am. B **32**, A1.
Piani, M., and J. Watrous, 2015, Phys. Rev. Lett. **114**, 060404.
Plávala, M., 2017, Phys. Rev. A **96**, 052127.
Preskill, J., 1998, lecture notes, Physics 229: Quantum Information and Computation, http://theory.caltech.edu/~preskill/ph229/.
Prugovecki, E., 1992, *Quantum Geometry—A Framework for Quantum General Relativity* (Kluwer Academic Publishers, Dordrecht).
Pusey, M. F., 2015, J. Opt. Soc. Am. B **32**, A56.
Pusey, M. F., J. Barrett, and T. Rudolph, 2012, Nat. Phys. **8**, 475.
Quintino, M. T., J. Bowles, F. Hirsch, and N. Brunner, 2016, Phys. Rev. A **93**, 052115.
Quintino, M. T., T. Vértesi, and N. Brunner, 2014, Phys. Rev. Lett. **113**, 160402.
Quintino, M. T., T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, 2015, Phys. Rev. A **92**, 032107.
Rahimi-Keshari, S., M. Mehboudi, D. De Santis, D. Cavalcanti, and A. Acín, 2021, Phys. Rev. A **104**, 042212.
Reeb, D., D. Reitzner, and M. M. Wolf, 2013, J. Phys. A **46**, 462002.
Robens, C., W. Alt, D. Meschede, C. Emary, and A. Alberti, 2015, Phys. Rev. X **5**, 011003.
Robertson, H. P., 1929, Phys. Rev. **34**, 163.
Robertson, H. P., 1934, Phys. Rev. **46**, 794.
Ruskai, M. B., 1969, Phys. Rev. **183**, 129.
Saha, D., M. Oszmaniec, L. Czekaj, M. Horodecki, and R. Horodecki, 2020, Phys. Rev. A **101**, 052104.
Selby, J. H., D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, 2021, arXiv:2106.09045.
Shi, W., and C. Tang, 2020, arXiv:2005.02722.
Skrzypczyk, P., M. J. Hoban, A. B. Sainz, and N. Linden, 2020, Phys. Rev. Res. **2**, 023292.
Skrzypczyk, P., and N. Linden, 2019, Phys. Rev. Lett. **122**, 140403.
Skrzypczyk, P., I. Šupić, and D. Cavalcanti, 2019, Phys. Rev. Lett. **122**, 130403.
Smirne, A., S. Cialdi, D. Cipriani, C. Carmeli, A. Toigo, and B. Vacchini, 2022, Quantum Sci. Technol. **7**, 025016.
Son, W., E. Andersson, S. M. Barnett, and M. S. Kim, 2005, Phys. Rev. A **72**, 052116.
Spekkens, R. W., 2005, Phys. Rev. A **71**, 052108.
Stano, P., D. Reitzner, and T. Heinosaari, 2008, Phys. Rev. A **78**, 012315.
Steiner, M., 2003, Phys. Rev. A **67**, 054305.
Styliaris, G., and P. Zanardi, 2019, Phys. Rev. Lett. **123**, 070401.
Tavakoli, A., and R. Uola, 2019, arXiv:1905.03614.
Uola, R., C. Budroni, O. Gühne, and J.-P. Pellonpää, 2015, Phys. Rev. Lett. **115**, 230402.
Uola, R., T. Bullock, T. Kraft, J.-P. Pellonpää, and N. Brunner, 2020, Phys. Rev. Lett. **125**, 110402.
Uola, R., A. C. S. Costa, H. C. Nguyen, and O. Gühne, 2020, Rev. Mod. Phys. **92**, 015001.
Uola, R., T. Kraft, S. Designolle, N. Miklin, A. Tavakoli, J.-P. Pellonpää, O. Gühne, and N. Brunner, 2021, Phys. Rev. A **103**, 022203.
Uola, R., T. Kraft, J. Shang, X.-D. Yu, and O. Gühne, 2019, Phys. Rev. Lett. **122**, 130404.
Uola, R., F. Lever, O. Gühne, and J.-P. Pellonpää, 2018, Phys. Rev. A **97**, 032301.
Uola, R., K. Luoma, T. Moroder, and T. Heinosaari, 2016, Phys. Rev. A **94**, 022109.
Uola, R., T. Moroder, and O. Gühne, 2014, Phys. Rev. Lett. **113**, 160403.
Uola, R., G. Vitagliano, and C. Budroni, 2019, Phys. Rev. A **100**, 042117.
Vidal, G., and R. Tarrach, 1999, Phys. Rev. A **59**, 141.
Watrous, J., 2018, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, England).
Werner, R., 2004, Quantum Inf. Comput. **4**, 546.
Werner, R. F., 1989, Phys. Rev. A **40**, 4277.
Wiesner, S., 1983, ACM SIGACT News **15**, 78.
Wilde, M. M., and A. Mizel, 2012, Found. Phys. **42**, 256.
Wiseman, H. M., S. J. Jones, and A. C. Doherty, 2007, Phys. Rev. Lett. **98**, 140402.
Wolf, M., D. Perez-Garcia, and C. Fernandez, 2009, Phys. Rev. Lett. **103**, 230402.
Wootters, W. K., and B. D. Fields, 1989, Ann. Phys. (N.Y.) **191**, 363.
Xu, Z.-P., and A. Cabello, 2019, Phys. Rev. A **99**, 020103.
Yu, S., N. Liu, L. Li, and C. Oh, 2010, Phys. Rev. A **81**, 062116.
Yu, S., and C. H. Oh, 2012, Phys. Rev. Lett. **108**, 030402.
Yu, S., and C. H. Oh, 2013, arXiv:1312.6470.
Yu, X.-D., T. Simnacher, N. Wyderka, H. C. Nguyen, and O. Gühne, 2021, Nat. Commun. **12**, 1012.
Zhang, L., H. Xiang, X. Li-Jost, and S.-M. Fei, 2019, Phys. Rev. E **100**, 062139.

Correction: Equations (8) and (9) contained minor errors and have been remedied. The previously published running head contained a transposition error and has been fixed.