NON-DISTURBING QUANTUM MEASUREMENTS

TEIKO HEINOSAARI AND MICHAEL M. WOLF

ABSTRACT. We consider pairs of discrete quantum observables (POVMs) and analyze the relation between the notions of non-disturbance, joint measurability and commutativity. We specify conditions under which these properties coincide or differ—depending for instance on the interplay between the number of outcomes and the Hilbert space dimension or on algebraic properties of the effect operators. We also show that (non-)disturbance is in general not a symmetric relation and that it can be decided and quantified by means of a semidefinite program.

1. INTRODUCTION

One of the main features of quantum mechanics is that measurements of different observables usually disturb each other. This property often comes along with non-commutativity or the impossibility of measuring observables simultaneously. Strictly speaking, however, *non-disturbance, joint measurability* and *commutativity* are different concepts and it is the aim of the present paper to clarify their precise relation.

In general all these notions turn out to be different, but we will see that some coincide with others under certain conditions involving for instance the interplay between the number of measurement outcomes and the Hilbert space dimension or algebraic properties of the effect operators.

In this work we concentrate on discrete observables. Our investigation is organized as follows. A brief overview of the relevant concepts is given in Section 2. The disturbance of one measurement w.r.t. another is then studied in Section 3, where it is shown that (i) non-disturbance is not a symmetric relation, (ii) it is equivalent to commutativity if the second measurement has sufficiently many independent outcomes but (iii) inequivalent to joint measurability and commutativity in general. In Section 4 we discuss measurements which do not disturb themselves, i.e., measurements of the first kind, and their relation to repeatability and commutativity. Finally in Section 5 we

argue that (non-)disturbance can be decided and quantified efficiently by means of a semidefinite program.

2. Preliminaries

This section will fix some notations and introduce the basic concepts. Let $\mathcal{H}$ be a complex Hilbert space, either finite or countably infinite dimensional. We denote by $\mathcal{L}(\mathcal{H})$ the set of bounded linear operators and by $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{L}(\mathcal{H})$ the set of trace class operators on $\mathcal{H}$. A positive operator $\varrho \in \mathcal{T}(\mathcal{H})$ having trace one is a *density operator*, also referred to as *state*, and we denote by $\mathcal{S}(\mathcal{H})$ the set of all states.

2.1. Observables.
Observables are generally described by *positive operator valued measures* (POVMs). In this work we only consider discrete observables. Therefore, an observable is characterized by a finite or countably infinite collection of positive operators $\mathsf{A} \equiv \{\mathsf{A}_x\}$ satisfying $\sum_x \mathsf{A}_x = \mathbb{1}$. Here the sum runs over all $x \in \Omega_{\mathsf{A}}$, where the set Ω_{A} is the collection of all possible measurement outcomes. Whenever convenient we take $\Omega_{\mathsf{A}} \subseteq \mathbb{N}$ as the labeling of the outcomes is irrelevant in our investigation. If a system is prepared in a state ϱ, then a measurement of an observable A will lead to an outcome x with probability $\mathrm{tr}\,[\varrho \mathsf{A}_x]$.

A selfadjoint operator $E \in \mathcal{L}(\mathcal{H})$ satisfying $0 \leq E \leq \mathbb{1}$ is called an *effect* and we denote the set of all effects by $\mathcal{E}(\mathcal{H})$. Note that $\mathsf{A}_x \in \mathcal{E}(\mathcal{H})$, i.e., the elements of a POVM are effects.

2.2. Joint measurability and commutativity.
Given two observables A and B, we say that they are *jointly measurable* if there exists a third observable G with $\Omega_{\mathsf{G}} = \Omega_{\mathsf{A}} \times \Omega_{\mathsf{B}}$ and satisfying $\sum_x \mathsf{G}_{x,y} = \mathsf{B}_y$ for all y and $\sum_y \mathsf{G}_{x,y} = \mathsf{A}_x$ for all x. In other words, A and B correspond to the 'marginals' of G.

A particular case of jointly measurable pairs of observables A and B are those which commute, i.e.,

$$[\mathsf{A}_x, \mathsf{B}_y] = 0$$

for all $x \in \Omega_{\mathsf{A}}$, $y \in \Omega_{\mathsf{B}}$. In this case we can set $\mathsf{G}_{x,y} = \mathsf{A}_x \mathsf{B}_y$ which defines an observable G on the product set $\Omega_{\mathsf{A}} \times \Omega_{\mathsf{B}}$ since the commutativity of A and B guarantees that the operators $\mathsf{G}_{x,y}$ are positive. Note that when talking about commuting pairs of observables we do not necessarily require that the effects within each observable are commuting, i.e., $[\mathsf{A}_x, \mathsf{A}_{x'}]$ may be non-zero for $x \neq x'$.

Two observables A and B can be jointly measurable even if they do not commute. The relation of being jointly measurable is also qualitatively different from commutativity. For instance, even if all partitionings of A and B into two outcome observables are jointly measurable, it may happen that A and B are not [10]. We also recall that not being jointly measurable is closely related and sometimes provably equivalent to the possibility of detecting 'quantum non-locality', i.e., the ability of violating a Bell inequality [14].

2.3. **Instruments.** An observable describes the statistics of the outcomes of a measurement but leaves open how the measurement alters the quantum state. In order to discuss this we need the concept of an *instrument* [9]. An instrument which implements an observable A is a collection of completely positive linear maps $\mathcal{I} \equiv \{\mathcal{I}_x\}$ on $\mathcal{T}(\mathcal{H})$ which satisfy $\mathcal{I}_x^*(\mathbb{1}) = \mathsf{A}_x$ for every x. Here the *adjoint map* $\mathcal{I}_x^*$ is defined via the usual trace duality $\mathrm{tr}\,[X\mathcal{I}_x(Y)] = \mathrm{tr}\,[\mathcal{I}_x^*(X)Y]$ for all $X \in \mathcal{L}(\mathcal{H}), Y \in \mathcal{T}(\mathcal{H})$. In other words, $\mathcal{I}_x^*$ and $\mathcal{I}_x$ correspond to the Heisenberg and Schrödinger pictures, respectively.

If an observable A is implemented via an instrument $\mathcal{I}$, then $\mathcal{I}_x(\varrho)$ is the unnormalized state after having obtained the measurement outcome x upon an input state ϱ. Note that $\mathrm{tr}\,[\mathcal{I}_x(\varrho)] = \mathrm{tr}\,[\varrho\mathsf{A}_x]$ is the probability for this to happen. We always assume that the output system has the same dimension as the input system.

If we ignore the measurement outcome, an instrument $\mathcal{I}$ transforms an input state ϱ to an (unconditional) output state

$$\mathcal{I}_\Omega(\varrho) := \sum_x \mathcal{I}_x(\varrho)\,.$$

The map $\mathcal{I}_\Omega$ is a *quantum channel*, i.e., a completely positive, trace-preserving linear map on $\mathcal{T}(\mathcal{H})$. The dual map $\mathcal{I}_\Omega^*$ is completely positive, identity preserving linear map on $\mathcal{L}(\mathcal{H})$. It is also σ-weakly continuous.

Evidently, many different instruments correspond to the same observable. A particular implementation of an observable A is given by its *Lüders instrument* $\mathcal{I}^L$, defined as

$$\mathcal{I}_x^L(\varrho) = \mathsf{A}_x^{1/2}\varrho\mathsf{A}_x^{1/2}\,.$$

It is also easy to give examples of other instruments. For instance, fix a state ξ_x for each outcome $x \in \Omega_\mathsf{A}$. Then the formula

$$\mathcal{I}_x(\varrho) = \mathrm{tr}\,[\varrho\mathsf{A}(x)]\,\xi_x \tag{1}$$

defines an instrument implementing A.

2.4. **Non-disturbing measurements.** Given two observables A and B we say that A *can be measured without disturbing* B if there exists an instrument $\mathcal{I}$ which implements A and for which

$$\operatorname{tr}\left[\mathcal{I}_\Omega(\varrho)\mathsf{B}_y\right] = \operatorname{tr}\left[\varrho\mathsf{B}_y\right] \quad \forall\varrho \in \mathcal{S}(\mathcal{H}), y \in \Omega_\mathsf{B}\,. \tag{2}$$

This means that the measurement statistics of B are the same for all pairs of an input state ϱ and an output state $\mathcal{I}_\Omega(\varrho)$.

We can write the non-disturbance condition (2) in an equivalent form

$$\mathcal{I}_\Omega^*(\mathsf{B}_y) = \mathsf{B}_y \quad \forall y \in \Omega_\mathsf{B}\,. \tag{3}$$

Hence, this is to say all the effects B_y are *fixed points* of $\mathcal{I}_\Omega^*$. We denote by $F(\mathcal{I}_\Omega^*) \subseteq \mathcal{L}(\mathcal{H})$ the set of fixed points of $\mathcal{I}_\Omega^*$. This is clearly a linear subspace of $\mathcal{L}(\mathcal{H})$. Moreover, since $\mathcal{I}_\Omega^*$ is σ-weakly continuous, $F(\mathcal{I}_\Omega^*)$ is σ-weakly closed.

If A and B commute, then a non-disturbing measurement can be achieved by the Lüders instrument implementing A since

$$\begin{aligned}
\operatorname{tr}\left[\mathcal{I}_\Omega^L(\varrho)\mathsf{B}_y\right] &= \sum_x \operatorname{tr}\left[\mathsf{A}_x^{1/2}\varrho\mathsf{A}_x^{1/2}\mathsf{B}_y\right] = \sum_x \operatorname{tr}\left[\mathsf{A}_x\varrho\mathsf{B}_y\right] \\
&= \operatorname{tr}\left[\varrho\mathsf{B}_y\right]\,.
\end{aligned}$$

We recall that if the Hilbert space $\mathcal{H}$ is finite dimensional, then the Lüders instrument implementing A does not disturb B if and *only if* they commute [8]. In an infinite dimensional Hilbert space this is not generally true; there exist non-commuting observables A and B such that the Lüders measurement of A does not disturb B [1], [12].

To give a class of examples of pairs in which a non-disturbing measurement is not possible, suppose that B is an informationally complete observable. This means that the probabilities $\operatorname{tr}\left[\varrho\mathsf{B}_y\right]$ uniquely determine every state ϱ. It is then clear from the non-disturbance condition (2) that $\mathcal{I}_\Omega(\varrho) = \varrho$ for every state ϱ. However, every non-trivial observable A necessarily perturbs at least some state and therefore also disturbs any informationally complete observable B.

If A can be implemented via an instrument $\mathcal{I}$ which does not disturb B, then A and B are jointly measurable. This is quite obvious since their sequential measurement gives the measurement statistics of both A and B without any perturbation. Formally, we can set $\mathsf{G}_{x,y} = \mathcal{I}_x^*(\mathsf{B}_y)$ which defines a joint observable G for A and B.

In Subsection 3.1 we will see an explicit example of two observables A and B which are jointly measurable despite the fact that one necessarily disturbs the other. In that case, there is no instrument $\mathcal{I}$ implementing A and satisfying $\mathcal{I}_\Omega^*(\mathsf{B}_y) = \mathsf{B}_y$ for all y. Hence, the joint measurement cannot be implemented sequentially by first measuring A and then B.

Let us remark that if two observables are jointly measurable, there is always a sequential implementation which encodes the outcomes of the B-measurement into a quantum system and then recovers them again by a final measurement. The latter is then, however, generally different from B and one may have to increase the dimension of the Hilbert space for the encoding. An instrument for this kind of sequential implementation can be chosen to be

$$\mathcal{I}_x(\varrho) = \sum_y \operatorname{tr}\left[\varrho \mathsf{G}_{x,y}\right] |y\rangle\langle y|,$$

where G is a joint observable of A and B and $\{|y\rangle\}$ is an orthonormal basis. The instrument $\mathcal{I}$ implements A, while a subsequent measurement in the basis $|y\rangle$ yields B.

2.5. **Sharp observables.** An observable A is called *sharp* if each effect A_x is a projection. It is well known that if we have two observables A, B and (at least) one of them is sharp, then the three relations - commutativity, joint measurability and non-disturbance - are equivalent. In the following we note two slightly stronger results.

Proposition 1. *Let A and B be two jointly measurable observables. Suppose that B_y is a projection. Then $[\mathsf{A}_x, \mathsf{B}_y] = 0$ for all $x \in \Omega_\mathsf{A}$.*

Proof. Let G be a joint observable for A and B. Since $\sum_x \mathsf{G}_{xy} = \mathsf{B}_y$, we conclude that $\mathsf{G}_{xy} \le \mathsf{B}_y$ for every x. But as B_y is a projection, this implies that $\mathsf{G}_{xy}\mathsf{B}_y = \mathsf{B}_y\mathsf{G}_{xy} = \mathsf{G}_{xy}$ for every x. Also $\mathbb{1} - \mathsf{B}_y$ is a projection and $\sum_{y' \ne y}\sum_x \mathsf{G}_{xy'} = \mathbb{1} - \mathsf{B}_y$. We thus get $\mathsf{G}_{xy'}(\mathbb{1} - \mathsf{B}_y) = (\mathbb{1} - \mathsf{B}_y)\mathsf{G}_{xy'} = \mathsf{G}_{xy'}$ for every x and every $y' \ne y$, implying that $\mathsf{G}_{xy'}\mathsf{B}_y = \mathsf{B}_y\mathsf{G}_{xy'} = 0$. We conclude that B_y commutes with all effects $\mathsf{G}_{xy'}$, therefore also with $\mathsf{A}_x = \sum_{y'} \mathsf{G}_{xy'}$. $\qquad\square$

Proposition 2. *Let A and B be two observables. Suppose that for some y, the effect B_y is proportional to a projection. If an instrument $\mathcal{I}$ implements A and $\mathcal{I}_\Omega^*(\mathsf{B}_y) = \mathsf{B}_y$, then $[\mathsf{A}_x, \mathsf{B}_y] = 0$ for all $x \in \Omega_\mathsf{A}$.*

Proof. The proof follows Lemma 3.3 in [4]. Denote by P the projection onto the support of B_y and $P^\perp = \mathbb{1} - P$. The assumption means that $\mathsf{B}_y = cP$ for some $c > 0$, and $\mathcal{I}_\Omega^*(\mathsf{B}_y) = \mathsf{B}_y$ thus implies that $\mathcal{I}_\Omega^*(P) = P$. Moreover, since $\mathcal{I}_\Omega^*(\mathbb{1}) = \mathbb{1}$ we also have $\mathcal{I}_\Omega^*(P^\perp) = P^\perp$.

Since $\mathcal{I}_x$ is completely positive we can make use of its Kraus decomposition $\mathcal{I}_x(\cdot) = \sum_\alpha K_{\alpha,x} \cdot K_{\alpha,x}^*$. If we insert this into the equation $\mathcal{I}_\Omega^*(P) = P$ and multiply with $P^\perp$ from the left and the right, we get

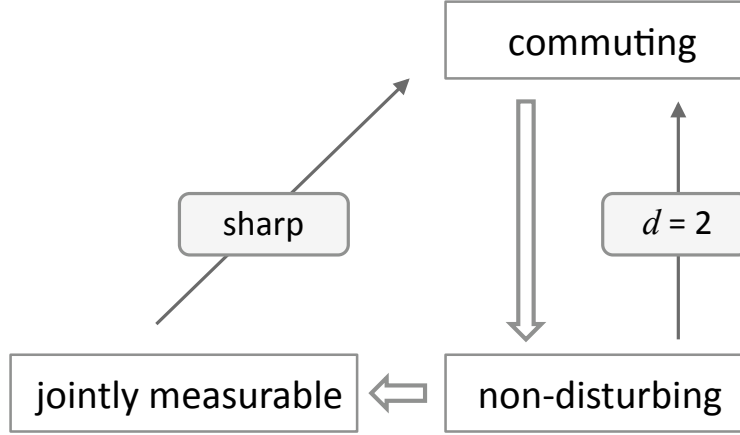$$\sum_{\alpha,x} P^\perp K_{\alpha,x}^* P^2 K_{\alpha,x} P^\perp = 0. \tag{4}$$

FIGURE 1. Implications between the different concepts. The gray boxes indicate the additional conditions under which the implications holds.

Since all the summands are positive this implies $PK_{\alpha,x}P^{\perp} = 0$ and therefore $PK_{\alpha,x}P = PK_{\alpha,x}$. Interchanging P and $P^{\perp}$ in the argument gives $PK_{\alpha,x}P = K_{\alpha,x}P$ and thus $[P, K_{\alpha,x}] = 0$. Taking the adjoint of this equation shows that P also commutes with $K^{*}_{\alpha,x}$, so that $[\mathsf{A}_x, P] = 0$ and finally $[\mathsf{A}_x, \mathsf{B}_y] = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

3. PAIRS OF QUANTUM OBSERVABLES

In this section we investigate the non-disturbance relation for pairs of observables. We start by demonstrating that the non-disturbance criterion differs from the joint measurability and commutativity criteria even for pairs of two-outcome observables (Subsec. 3.1). However, with some additional requirements non-disturbance reduces to commutativity. This happens, for instance, in the case of qubit observables (Subsec. 3.2). The overall picture is schematically summarized in Fig. 1.

We show that even if an observable A can be measured without disturbing another observable B, the converse need not hold (Subsec. 3.3). This means that non-disturbance is not a symmetric relation, unlike commutativity and joint measurability. Finally, we discuss the disturbance caused by a rank-1 observable (Subsec. 3.4).

3.1. **Two-outcome measurements.** In the simplest case an observable A has only two outcomes. It is then determined by a single effect

A_1, since the normalization requires that $A_2 = \mathbb{1} - A_1$. Clearly, a two-outcome observable is commutative.

Let A and B be sharp two-outcome observables. We assume that A and B do not commute, hence they are not jointly measurable since these concepts coincide for sharp observables.

For all $\mu, \nu \in (\frac{1}{2}, 1]$, we define *coarse-grainings* of A and B by

$$A_1^\mu = \mu A_1 + (1 - \mu)A_2, \quad A_2^\mu = (1 - \mu)A_1 + \mu A_2$$

and

$$B_1^\nu = \nu B_1 + (1 - \nu)B_2, \quad B_2^\nu = (1 - \nu)B_1 + \nu B_2.$$

In this way we get observables A^μ and B^ν, which we regard as approximate versions of A and B, respectively. The numbers μ and ν quantify the levels of approximation.

If μ and ν are small enough, then A^μ and B^ν are jointly measurable. For instance, suppose that $\mu = \nu$ and define four operators by

$$G_{1,1} = (1 - \mu)\,\mathbb{1}$$
$$G_{1,2} = (2\mu - 1)A_1$$
$$G_{2,1} = (2\mu - 1)B_1$$
$$G_{2,2} = \mu\,\mathbb{1} - (2\mu - 1)(A_1 + B_1).$$

The first three operators are clearly positive, and taking into account that $0 \le A_1 + B_1 \le 2\mathbb{1}$ we see that also the fourth operator is positive if $\frac{1}{2} < \mu \le \frac{2}{3}$. Under this condition G is a joint observable for A^μ and B^μ. In conclusion, if $\frac{1}{2} < \mu \le \frac{2}{3}$ then A^μ and B^μ are jointly measurable.

We claim that A^μ and B^ν cannot be measured without disturbing each other, no matter how the numbers μ and ν are chosen from the interval $(\frac{1}{2}, 1]$. To prove this, let us first notice that

$$A_1^\mu = (2\mu - 1)A_1 + (1 - \mu)\mathbb{1}, \quad B_1^\nu = (2\nu - 1)B_1 + (1 - \nu)\mathbb{1}.$$

Hence,

$$[A_1^\mu, B_1^\nu] = (2\mu - 1)(2\nu - 1)[A_1, B_1],$$

showing that A^μ and B^ν do not commute for any $\mu, \nu \in (\frac{1}{2}, 1]$. In particular, A^μ and the sharp observable B are not jointly measurable according to Prop.2.

Let us make a counter assumption that there exists an instrument $\mathcal{I}$ implementing A^μ and not disturbing B^ν, i.e.,

$$\mathcal{I}_\Omega^*(B_y^\nu) = B_y^\nu \qquad y = 1, 2. \tag{5}$$

But since $\mathcal{I}_\Omega^*(\mathbb{1}) = \mathbb{1}$ and B_y is a linear combination of $\mathbb{1}$ and B_y^ν, it follows that

$$\mathcal{I}_\Omega^*(B_y) = B_y \qquad y = 1, 2. \tag{6}$$

However, this non-disturbance condition cannot hold as A^μ and B are not jointly measurable. So joint measurability of the above observables does never imply non-disturbance.

Let us proceed by giving an example of non-commuting observables A and B such that there exists an A-measurement not disturbing B. This example originates from Remark 2 in [2].

Let $\mathcal{H} = \mathbb{C}^3$. We define observables A and B as

$$A_1 = \frac{1}{4} \begin{pmatrix} 2 & 0 & -\sqrt{2} \\ 0 & 4 & 0 \\ -\sqrt{2} & 0 & 3 \end{pmatrix}, \quad A_2 = \frac{1}{4} \begin{pmatrix} 2 & 0 & \sqrt{2} \\ 0 & 0 & 0 \\ \sqrt{2} & 0 & 1 \end{pmatrix}$$

and

$$B_1 = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B_2 = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7}$$

Then A and B do not commute.

However, there is an instrument implementing A which does not disturb B. We set

$$K_1 = \frac{1}{2} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad K_2 = \frac{1}{10} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sqrt{10} & 0 \\ 0 & 2\sqrt{10} & 0 \end{pmatrix},$$

$$K_3 = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad K_4 = \frac{1}{20} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4\sqrt{10} & 0 \\ 0 & 2\sqrt{10} & 0 \end{pmatrix},$$

$$K_5 = \frac{1}{2} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

and define $\mathcal{I}_1^*(\cdot) = \sum_{j=1}^4 K_j \cdot K_j^*$ and $\mathcal{I}_2^*(\cdot) = K_5 \cdot K_5^*$. It is straighforward to check that $\mathcal{I}$ satisfies the non-disturbance condition (3).

3.2. When does non-disturbance reduce to commutativity? It is a fundamental fact of quantum theory that every measurement perturbs the system. Therefore, we expect that in a sequence of non-disturbing measurements, the second measurement cannot be too informative since otherwise we would detect the perturbation caused by the first measurement. In the following we give some precise conditions for this intuitive idea.

For each observable B, we denote by $\mathrm{span}\mathsf{B}$ the linear subspace in $\mathcal{L}(\mathcal{H})$ generated by the set $\{\mathsf{B}_y : y \in \Omega_\mathsf{B}\}$, i.e.,

$$\mathrm{span}\mathsf{B} = \{\sum_y c_y \mathsf{B}_y \mid c_y \in \mathbb{C}, c_y \neq 0 \text{ for finitely many } y\}\,.$$

By $\overline{\mathrm{span}}\mathsf{B}$ we denote the closure of $\mathrm{span}\mathsf{B}$ in the σ-weak operator topology. (Clearly, if $\dim\mathcal{H} < \infty$, then $\overline{\mathrm{span}}\mathsf{B} = \mathrm{span}\mathsf{B}$.)

Proposition 3. *Suppose that an observable* B *has the following property:*

$$\forall y \in \Omega_\mathsf{B} : \quad \mathsf{B}_y^2 \in \overline{\mathrm{span}}\mathsf{B}\,. \tag{8}$$

Then it is possible to measure an observable A *without disturbing* B *if and only if* A *and* B *commute.*

Proof. Let $\mathcal{I}$ be an instrument which implements A and does not disturb B. The set $F(\mathcal{I}_\Omega^*)$ of the fixed points of $\mathcal{I}_\Omega^*$ is a σ-weakly closed linear subspace of $\mathcal{L}(\mathcal{H})$. Hence, from (8) follows that every B_y^2 is a fixed point of $\mathcal{I}_\Omega^*$.

Let $\mathcal{I}_\Omega(\varrho) = \sum_j K_j \varrho K_j^*$ be a Kraus decomposition for $\mathcal{I}_\Omega$. Then for every $y \in \Omega_\mathsf{B}$, we get

$$\sum_j [K_j, \mathsf{B}_y]^* \, [K_j, \mathsf{B}_y] = \mathcal{I}_\Omega^*(\mathsf{B}_y^2) + \mathsf{B}_y^2 - \mathsf{B}_y \mathcal{I}_\Omega^*(\mathsf{B}_y) - \mathcal{I}_\Omega^*(\mathsf{B}_y)\mathsf{B}_y = 0$$

and therefore $[K_j, \mathsf{B}_y] = 0$ for each j. This implies that A and B commute. $\square$

To give a class of examples where the condition (8) holds, suppose that B is a classical coarse-graining of a sharp observable P in the sense that there is a stochastic square matrix M such that $\mathsf{B}_y = \sum_{y'} M_{yy'} \mathsf{P}_{y'}$. Suppose further that M is invertible (but the inverse need not be a stochastic matrix). Then $\mathsf{P}_{y'} = \sum_y M_{y'y}^{-1} \mathsf{B}_y$, implying that $\mathsf{B}_y^2 \in \mathrm{span}\mathsf{B}$ for every $y \in \Omega_\mathsf{B}$. The condition of M being invertible means that B and P are *informationally equivalent* [6]; B gives different measurement outcome distributions for two states ϱ_1 and ϱ_2 if and only if their measurement outcome distributions are different in a P-measurement.

The remaining results in this subsection rest on the following.

Proposition 4. *Let* A *be an observable and* $\mathcal{I}$ *an instrument implementing* A. *Suppose that the channel* $\mathcal{I}_\Omega$ *has a full rank fixed point* ϱ. *If* $\mathcal{I}$ *does not disturb an observable* B, *then* A *and* B *commute.*

Proof. As shown in Lemma 3.4 in [4], the set of fixed points $F(\mathcal{I}_\Omega^*)$ forms an algebra if $\mathcal{I}_\Omega$ has a full rank fixed point. The claim then follows from Prop. 3. $\square$

In the rest of this subsection we assume that $\mathcal{H}$ is a finite dimensional Hilbert space. We can then identify $\mathcal{T}(\mathcal{H})$ and $\mathcal{L}(\mathcal{H})$ with the set $\mathcal{M}_d(\mathbb{C})$ of $d \times d$ complex matrices, where $d = \dim \mathcal{H}$.

Let $\mathcal{I}$ be an instrument. By fixing a basis in $\mathcal{M}_d(\mathbb{C})$, we can consider the channel $\mathcal{I}_\Omega$ as a matrix acting on the d^2-dimensional vector space $\mathcal{M}_d(\mathbb{C})$. In this way, the fixed points of $\mathcal{I}_\Omega$ are the right eigenvectors with eigenvalue 1, while the fixed points of the dual channel $\mathcal{I}_\Omega^*$ are the left eigenvectors with eigenvalue 1. In particular, the subspaces $F(\mathcal{I}_\Omega)$ and $F(\mathcal{I}_\Omega^*)$ consisting of the fixed points of $\mathcal{I}_\Omega$ and $\mathcal{I}_\Omega^*$ have the same dimension, $\dim F(\mathcal{I}_\Omega) = \dim F(\mathcal{I}_\Omega^*)$.

To formulate the following statement, we denote by $\dim \mathsf{B}$ the dimension of the linear subspace $\mathrm{span}\,\mathsf{B} \subset \mathcal{M}_d(\mathbb{C})$. Roughly speaking, $\dim \mathsf{B}$ is the number of independent measurement outcomes obtained in a B-measurement.

Proposition 5. *Let* B *be an observable such that*

$$\dim \mathsf{B} \geq (d-1)^2 + 1\,. \tag{9}$$

It is possible to measure an observable A *without disturbing* B *if and only if* A *and* B *commute.*

Proof. Let $\mathcal{I}$ be an instrument which implements A and does not disturb B. We will show that the condition (9) implies that $\mathcal{I}_\Omega$ has a full rank fixed point. The claim then follows from Proposition 4.

As proved in [11] (see also [13]), there is a unitary matrix U and a set of states ϱ_k such that

$$F(\mathcal{I}_\Omega) = U \left(0 \oplus \bigoplus_{k=1}^{K} \mathcal{M}_{d_k}(\mathbb{C}) \otimes \varrho_k \right) U^* \tag{10}$$

for an appropriate decomposition of the Hilbert space $\mathbb{C}^d = \mathbb{C}^{d_0} \oplus \bigoplus_k \mathbb{C}^{d_k} \otimes \mathbb{C}^{m_k}$. In particular, there are natural numbers $d_1, \ldots, d_K$ and $m_1, \ldots, m_K$ such that $\dim F(\mathcal{I}_\Omega) = \sum_k d_k^2$ and $\sum_k d_k m_k \leq d$. For convenience, we may assume that $d_k \geq d_{k+1}$ for all $k = 1, \ldots, K-1$.

It follows from this decomposition of $F(\mathcal{I}_\Omega)$ that $\dim F(\mathcal{I}_\Omega)$ can take only some specific values. Clearly, the largest value is $\dim F(\mathcal{I}_\Omega) = d^2$ and then $F(\mathcal{I}_\Omega) = \mathcal{M}_d(\mathbb{C})$, thus $\mathcal{I}_\Omega$ clearly has a full rank invariant state. The second largest value is $\dim F(\mathcal{I}_\Omega) = (d-1)^2 + 1$ and this means that $d_1 = d - 1$ and $d_2 = m_1 = m_2 = 1$. In this case $\mathbb{1} \in F(\mathcal{I}_\Omega)$, hence $\mathcal{I}_\Omega$ has a full rank fixed point. Since $\dim F(\mathcal{I}_\Omega) \geq \dim \mathsf{B}$, the claim follows. $\square$

It is a direct consequence of Proposition 5 that for qubit observables (i.e. $d = 2$) non-disturbance and commutativity are equivalent conditions. We find it useful to give also a simplified proof of this fact.

Proposition 6. *Let* $\dim \mathcal{H} = 2$. *For two observables* A *and* B, *the following conditions are equivalent:*

(i) *It is possible to measure* A *without disturbing* B.
(ii) *It is possible to measure* B *without disturbing* A.
(iii) A *and* B *commute.*

Proof. Let $\mathcal{I}$ be an instrument which implements A without disturbing B. If the channel $\mathcal{I}_\Omega$ has a full rank fixed point, then A and B commute by Proposition 4.

So let us then assume that $\mathcal{I}_\Omega$ does not have a full rank fixed point. This implies that $\dim F(\mathcal{I}_\Omega) = 1$, hence also $\dim F(\mathcal{I}_\Omega^*) = 1$. But $\mathcal{I}_\Omega^*(\mathbb{1}) = \mathbb{1}$, and therefore each B_y is a scalar multiple of the identity operator $\mathbb{1}$. Hence, B commutes with A. $\square$

3.3. **Non-disturbance is not symmetric.** In the following we demonstrate that the non-disturbance relation is not symmetric; there exist observables A and B such that every measurement of B disturbs A while a suitable A-measurement does not disturb B. Again this example originates from Remark 2 in [2]. The construction requires that $\dim \mathcal{H} \geq 3$, and we have seen in Section 3.2 that for $\dim \mathcal{H} = 2$ non-disturbance is equivalent to commutativity, hence a symmetric relation.

Let $\mathcal{H} = \mathbb{C}^3$. We choose B and $K_1, \ldots, K_5$ as in the end of Subsec. 3.1. We take A to be the five outcome observable $A_x = K_x K_x^*$, $x = 1, \ldots, 5$. The effects are thus

$$A_1 = \frac{1}{4} \begin{pmatrix} 2 & 0 & -\sqrt{2} \\ 0 & 0 & 0 \\ -\sqrt{2} & 0 & 1 \end{pmatrix}, \quad A_2 = \frac{1}{10} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & -2 & 4 \end{pmatrix},$$

$$A_3 = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_4 = \frac{1}{10} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 1 \end{pmatrix},$$

$$A_5 = \frac{1}{4} \begin{pmatrix} 2 & 0 & \sqrt{2} \\ 0 & 0 & 0 \\ \sqrt{2} & 0 & 1 \end{pmatrix}.$$

The instrument $\mathcal{I}_x^*(\cdot) = K_x \cdot K_x^*$ implements A and does not disturb B. However, the matrices $A_1, \ldots, A_5$ are linearly independent, implying that $\dim A = 5$. As A and B do not commute, it follows from Proposition 5 that all B-measurements disturb A.

3.4. **Rank-1 observables.** An effect $E \in \mathcal{E}(\mathcal{H})$ is *rank-1* if there is a one-dimensional projection P and a number $0 < e \leq 1$ such that $E = eP$. A discrete observable A is called *rank-1 observable* if each effect A_x is rank-1. Rank-1 observables form an important subset of all observables.

From Proposition 2 we conclude the following.

Proposition 7. *Let* B *be a rank-1 observable. It is possible to measure an observable* A *without disturbing* B *if and only if* A *and* B *commute.*

Suppose then that A is a rank-1 observable and it can be measured without disturbing another observable B. This does *not* imply that A and B commute. Indeed, the example given in Subsec. 3.3 serves as a counterexample.

In spite of this fact, a measurement of a rank-1 observable does make all subsequent measurements useless. In the following we make this statement precise and we characterize all instruments implementing rank-1 observables.

Proposition 8. *Let* E *be an effect. The following conditions are equivalent:*

(i) *E is rank-1.*
(ii) *Every completely positive linear mapping* Φ *on* $\mathcal{T}(\mathcal{H})$ *which satisfies* $\Phi^*(\mathbb{1}) = E$ *is of the form*

$$\Phi(\varrho) = \operatorname{tr}[\varrho E]\, \xi \tag{11}$$

for some state ξ.

Proof. (i)$\Rightarrow$(ii): Let $\{K_j\}$ be the set of Kraus operators for Φ, so that

$$\Phi(\varrho) = \sum_j K_j \varrho K_j^*, \quad \Phi^*(\mathbb{1}) = \sum_j K_j^* K_j = E. \tag{12}$$

The last equation implies that for each j, $K_j^* K_j \leq E$. Since E is rank-1, there is a number $0 < k_j \leq 1$ such that $K_j^* K_j = k_j E$. Clearly, $\sum_j k_j = 1$.

Let $K_j = S_j |K_j|$ be the polar decomposition of K_j. Here S_j is a partial isometry with $\ker S_j = \ker K_j$ and

$$|K_j| = \sqrt{K_j^* K_j} = \sqrt{k_j}\sqrt{E} = \sqrt{ek_j}P.$$

For every state ϱ, we then get

$$K_j \varrho K_j^* = ek_j S_j P \varrho P S_j^* = ek_j \operatorname{tr}[\varrho P]\, S_j P S_j^* = \operatorname{tr}[\varrho E]\, k_j S_j P S_j^*$$

and hence

$$\Phi(\varrho) = \mathrm{tr}\,[\varrho E] \sum_j k_j S_j P S_j^*\,.$$

The remaining thing is to show that the operator $S_j P S_j^*$ is a state for each j, implying that the convex sum $\sum_j k_j S_j P S_j^* =: \xi$ is a state also. The operator $S_j P S_j^*$ is clearly positive. The operator $S_j^* S_j$ is the projection on the closure of $\mathrm{ran}\,|K_j|$, thus $S_j^* S_j = P$. Therefore,

$$\mathrm{tr}\,[S_j P S_j^*] = \mathrm{tr}\,[S_j^* S_j P] = \mathrm{tr}\,[P] = 1\,.$$

(ii)$\Rightarrow$(i): We assume that an effect E is not rank-1 and show that there exists an operation satisfying $\Phi^*(\mathbb{1}) = E$ but not being of the form (11). Let $E = \int_{\sigma(E)} \lambda d\Pi(\lambda)$ be the spectral decomposition of E. We split the spectrum $\sigma(E)$ into two disjoint parts σ_1, σ_2 such that the operators $E_j = \int_{\sigma_j} \lambda d\Pi(\lambda)$, $j = 1, 2$ are nonzero. Clearly, $E_1 + E_2 = E$ and both E_1 and E_2 have eigenvalue 0. We fix two different states ξ_1, ξ_2 and define $\Phi_j(\cdot) = \mathrm{tr}\,[\cdot\, E_j]\, \xi_j$ for $j = 1, 2$. Then the operation $\Phi = \Phi_1 + \Phi_2$ satisfies $\Phi^*(\mathbb{1}) = E$. Suppose that ξ is a state satisfying (11). Choose a unit vector ψ_1 such that $E_1 \psi_1 = 0$ and $E_2 \psi_1 \neq 0$. Then

$$\Phi(|\psi_1\rangle\langle\psi_1|) = \langle\, \psi_1 \,|\, E\psi_1 \,\rangle\, \xi = \langle\, \psi_1 \,|\, E_2\psi_1 \,\rangle\, \xi$$

and, on the other hand,

$$\Phi(|\psi_1\rangle\langle\psi_1|) = \Phi_1(|\psi_1\rangle\langle\psi_1|) + \Phi_2(|\psi_1\rangle\langle\psi_1|) = \langle\, \psi_1 \,|\, E_2\psi_1 \,\rangle\, \xi_2\,.$$

Hence, $\xi = \xi_2$. But similarly we get $\xi = \xi_1$ if we repeat the calculation with a vector ψ_2 satisfying $E_2 \psi_2 = 0$ and $E_1 \psi_2 \neq 0$. This leads to the conclusion $\xi_1 = \xi_2$, which is a contradiction. Therefore, ξ does not exist. $\qquad\square$

Corollary 1. *Let* A *be an observable. The following conditions are equivalent:*

(i) A *is a rank-1 observable.*

(ii) *All instruments implementing* A *are of the form*

$$\mathcal{I}_x(\varrho) = \mathrm{tr}\,[\varrho \mathsf{A}_x]\, \xi_x\,, \tag{13}$$

where $\{\xi_x\}$ *is a set of states.*

Suppose we measure first a rank-1 observable A and then some other observable B. The instrument $\mathcal{I}$ describing the A-measurement is of the form (13) and the joint probability distribution is thus given by

$$\mathrm{tr}\,[\mathcal{I}_x(\varrho)\mathsf{B}_y] = \mathrm{tr}\,[\varrho \mathsf{A}_x]\, \mathrm{tr}\,[\xi_x \mathsf{B}_y]\,.$$

Here we see that the joint probability distribution can be calculated already after the first measurement since the numbers $\mathrm{tr}\,[\xi_x \mathsf{B}_y]$ do not

depend on the initial state ϱ at all. Therefore, the B-measurement is completely redundant.

4. Measurements of the first kind

A *first kind measurement* is one which does not disturb itself. More precisely, an instrument $\mathcal{I}$, implementing an observable A, is a *first kind instrument* if

$$\mathcal{I}^*_\Omega(\mathsf{A}_x) = \mathsf{A}_x \quad \text{for all } x \in \Omega_\mathsf{A}. \tag{14}$$

We say that an observable A *admits a first kind measurement* if there exists a first kind instrument which implements A. In this section we study some conditions guaranteeing that an observable A admits or does not admit a first kind measurement. The overall picture is summarized in Fig. 2.

A Lüders instrument $\mathcal{I}^L$ implementing a commutative observable A satisfies the first kind condition (14). Therefore, all commutative observables admit first kind measurements. As we have seen in Section 3, commutativity is also a necessary condition if A is either a qubit observable (Prop. 6) or a rank-1 observable (Prop. 7).

An example of an observable not admitting a first kind measurement is an informationally complete observable. Namely, any non-trivial measurement necessarily perturbs at least some state. But an informationally complete observable gives different measurement outcome distributions for all states, hence a subsequent measurement of the same observable detects any perturbation caused by the first measurement. In a finite dimension $d \geq 3$ this observation is generalized by the condition that if $\dim \mathsf{A} \geq (d-1)^2 + 1$, then A does not admit a first kind measurement. Namely, A cannot be commutative since $(d-1)^2 + 1 > d$, and the conclusion therefore follows from Prop. 5.

A more stringent condition than the first kind condition (14) is *repeatability*; an instrument $\mathcal{I}$, implementing an observable A, is repeatable if

$$\mathcal{I}^*_x(\mathsf{A}_y) = 0 \quad \text{whenever } x \neq y. \tag{15}$$

This condition means that measuring repeatedly gives not only the same statistics but repeated measurement outcomes.

It is clear that a repeatable instrument is of the first kind. The converse is, however, not true. For instance, it is easy to see that a Lüders instrument of a commutative observable is repeatable if and only if the associated observable is sharp. But as we pointed out earlier, the Lüders instrument of any commutative observable is first kind.
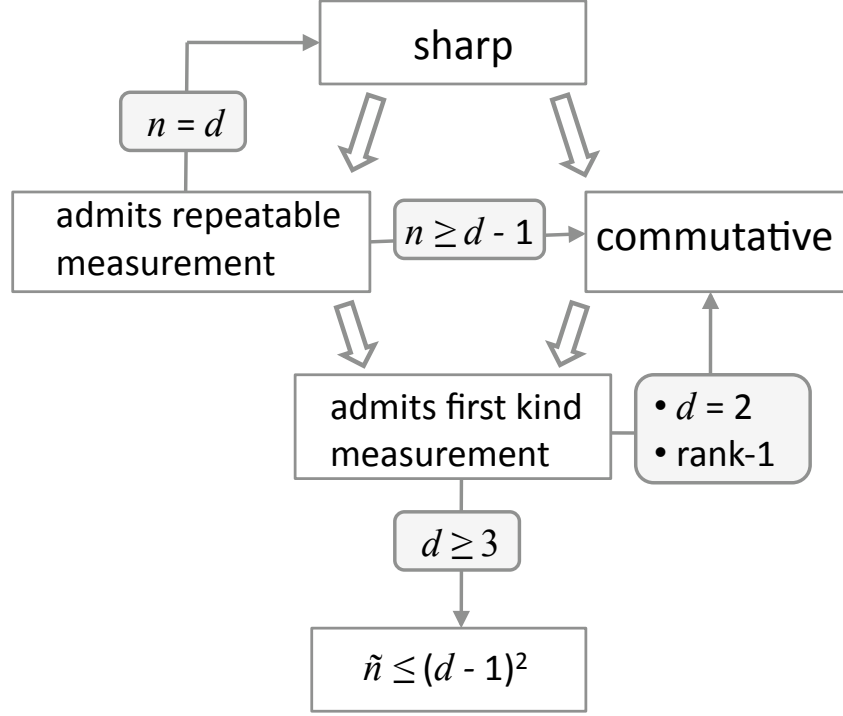
FIGURE 2. Implications between the different concepts. Here $d = \dim \mathcal{H}$, n is the number of outcomes and $\widetilde{n} = \dim \mathsf{A}$. The gray boxes indicate the additional conditions under which the implications holds. Each of the two conditions in the bullet list is sufficient for the implication.

An observable A admits a repeatable measurement if and only if each effect A_x has an eigenvalue 1 [7]. Namely, it follows from (15) that whenever $\mathcal{I}_x(\varrho) \neq 0$, then $\mathcal{I}_x(\varrho)$ is an unnormalized eigenstate of A_x with eigenvalue 1. On the other hand, if A satisfies this eigenvalue condition, we can construct a repeatable instrument by first fixing for each outcome x a unit vector ψ_x satisfying $\mathsf{A}_x \psi_x = \psi_x$, and then defining

$$\mathcal{I}_x(\varrho) = \operatorname{tr}\left[\varrho \mathsf{A}_x\right] |\psi_x\rangle\langle\psi_x| \,.$$

Since $\sum_x \mathsf{A}_x = \mathbb{1}$, we have $\mathsf{A}_y \psi_x = 0$ whenever $y \neq x$, implying that $\mathcal{I}$ is repeatable.

We conclude that there are two sufficient conditions for an observable A to admit a first kind measurement:

- A is commutative.

- Each effect A_x has eigenvalue 1.

These two conditions overlap (e.g. sharp observables), but they also cover different situations. To give an example of a non-commutative observable A having the property that each effect A_x has eigenvalue 1, suppose that $\Omega_A = \{1, 2, 3\}$ and $d = \dim \mathcal{H} \geq 5$. We split $\mathcal{H} = \mathcal{H}_3 \oplus \mathcal{H}_{d-3}$ into a 3 -dimensional and $d-3$ -dimensional subspaces $\mathcal{H}_3$ and $\mathcal{H}_{d-3}$. In $\mathcal{H}_3$ we fix three orthogonal one-dimensional projections P_1, P_2, P_3, while in $\mathcal{H}_{d-3}$ we fix two non-commuting projections R_1, R_2. Then A, defined as

$$A_1 = P_1 \oplus \tfrac{1}{2}R_1 \quad A_2 = P_2 \oplus \tfrac{1}{2}R_2 \,,$$
$$A_3 = P_3 \oplus (\mathbb{1} - \tfrac{1}{2}R_1 - \tfrac{1}{2}R_2) \,,$$

has the required properties.

The previous example is actually minimal in the sense that a non-commutative observable must have at least 3 outcomes, and, as we show in Proposition 9 below, a non-commutative observable can satisfy the eigenvalue condition only if $|\Omega_A| \leq d - 2$. In particular, for $d \leq 4$ only commutative observables can have repeatable measurements.

Proposition 9. *Suppose that $d = \dim \mathcal{H} < \infty$ and let A be an obsevable admitting a repeatable measurement. Then*

(a) *$|\Omega_A| \leq d$.*
(b) *If $|\Omega_A| = d$, then A is sharp.*
(c) *If $|\Omega_A| \geq d - 1$, then A is commutative.*

Proof. For each x, we write $A_x = P_x + R_x$, where P_x is a one-dimensional projection and part of the spectral projection of A_x associated with eigenvalue 1. Due to normalization two projections P_x, P_y must be orthogonal whenever $x \neq y$. Since there can be only d orthogonal projections in $\mathcal{H}$, we conclude that (a) holds.

For every x, y, we have $P_x + R_y \leq \mathbb{1}$ (since $\sum_x A_x = \mathbb{1}$). Multiplying by P_x on both sides gives $P_x R_y P_x \leq 0$, and thus $P_x R_y P_x = 0$. As $R_y^2 \leq R_y$ (since $R_y \in \mathcal{E}(\mathcal{H})$), we get

$$0 \leq P_x R_y (P_x R_y)^* = P_x R_y^2 P_x \leq P_x R_y P_x = 0 \,.$$

Thus, $P_x R_y = 0$ and $(\sum_x P_x) R_y = 0$. The sum $\sum_x P_x$ is a $|\Omega_A|$-dimensional projection. If $|\Omega_A| = d$, then $R_y = 0$. This proves (b).

If $|\Omega_A| = d-1$, then the previous calculation shows that all the effects R_y are scalar multiples of the 1-dimensional projection $\mathbb{1} - \sum_x P_x$. This proves (c). $\qquad\square$

5. Quantifying and deciding non-disturbance

In this section we assume that $\dim \mathcal{H} < \infty$ and that observables have only a finite number of outcomes. We show how to decide whether an observable A can be measured without disturbing B and, if non-disturbance cannot be achieved, how to quantify the least amount of disturbance induced on B by measuring A.

We first propose a quantification of the non-disturbance relation in a form of a measure of disturbance $D_{\mathsf{A}}(\mathsf{B})$ (Subsec. 5.1). This number satisfies $D_{\mathsf{A}}(\mathsf{B}) = 0$ if and only if A can be measured without disturbing B and it has a simple physical interpretation.

We then demonstrate that calculating the number $D_{\mathsf{A}}(\mathsf{B})$ is a semi-definite program (Subsec. 5.2). That is, the question whether an observable A can be measured without disturbing another observable B can be answered in an efficient and certifiable way. Note that this also implies that the question whether or not an observable admits a first kind measurement can be efficiently decided by setting $\mathsf{A} = \mathsf{B}$.

The problem of deciding whether or not observables are jointly measurable has been phrased in terms of a semidefinite program in [14].

5.1. **Quantification of the non-disturbance relation.** Let $\mathcal{I}$ be an instrument which implements an observable A. A second observable B, measured after A, is then possibly perturbed. Its measurement outcome probabilities, relative to the state prior to the A-measurement, are given by the modified effects $\mathcal{I}_{\Omega}^*(\mathsf{B}_y)$. We want to quantify the difference between B and its perturbed version.

Let us first notice that there is always a number $\lambda \in [0, 1]$ such that

$$-\lambda \mathbb{1} \leq \mathsf{B}_y - \mathcal{I}_{\Omega}^*(\mathsf{B}_y) \leq \lambda \mathbb{1} \qquad \forall y. \tag{16}$$

If A can be measured without disturbing B, then we can choose $\mathcal{I}$ in a way that $\lambda = 0$ in (16). Generally, the condition (16) is equivalent with the requirement that for every state ϱ,

$$|\mathrm{tr}\,[\varrho \mathsf{B}_y] - \mathrm{tr}\,[\varrho \mathcal{I}_{\Omega}^*(\mathsf{B}_y)]| \leq \lambda \qquad \forall y.$$

This inequality is expressing that the measurement outcome probabilities differ at most by λ.

We conclude that the number

$$\max_{y} \sup_{\varrho} |\mathrm{tr}\,[\varrho \mathsf{B}_y] - \mathrm{tr}\,[\varrho \mathcal{I}_{\Omega}^*(\mathsf{B}_y)]| = \max_{y} \|\mathsf{B}_y - \mathcal{I}_{\Omega}^*(\mathsf{B}_y)\|$$

gives a natural quantification of the difference between B and its perturbed version $\mathcal{I}_{\Omega}^*(\mathsf{B})$. In [5] this kind of distance between two observables was used in the study of approximate joint measurability.

We now want to quantify the least amount of disturbance induced on B by an A-measurement. Hence, we denote by $D_\mathsf{A}(\mathsf{B})$ the smallest number λ attained by any implementation of A, i.e.,

$$D_\mathsf{A}(\mathsf{B}) := \inf \left\{ \lambda \in [0,1] \mid \forall y : \ -\lambda \mathbb{1} \leq \mathsf{B}_y - \mathcal{I}_\Omega^*(\mathsf{B}_y) \leq \lambda \mathbb{1} \right\}, \quad (17)$$

where the infimum is taken over all instruments implementing A.

We will see in Subsec. 5.2 that the infimum in (17) is always attained. Therefore, $D_\mathsf{A}(\mathsf{B}) = 0$ if and only if A can be implemented without disturbing B. In general $D_\mathsf{A}(\mathsf{B})$ is, by construction, the maximal possible disturbance of measured probabilities minimized over all instruments implementing A.

Example 1. Let A and B be two sharp qubit observables. We assume that they do not commute, hence an A-measurement necessarily disturbs B. By Corollary 1, an instrument $\mathcal{I}$ implementing A is determined by two states ξ_1 and ξ_2, and we thus have

$$\mathcal{I}_\Omega^*(\mathsf{B}_y) = \operatorname{tr}\left[\xi_2 \mathsf{B}_y\right] \mathbb{1} + \operatorname{tr}\left[(\xi_1 - \xi_2)\mathsf{B}_y\right] \mathsf{A}_1, \quad y = 1, 2.$$

The two inequalities (16) for $y = 1$ and $y = 2$ are equivalent. To find $D_\mathsf{A}(\mathsf{B})$ we need to choose ξ_1 and ξ_2 in a way that the norm of the operator $\mathsf{B}_1 - \operatorname{tr}\left[\xi_2 \mathsf{B}_1\right]\mathbb{1} - \operatorname{tr}\left[(\xi_1 - \xi_2)\mathsf{B}_1\right]\mathsf{A}_1$ is as small as possible. To calculate the norm, we write $\mathsf{A}_1 = \frac{1}{2}(\mathbb{1} + \vec{a} \cdot \vec{\sigma})$ and $\mathsf{B}_1 = \frac{1}{2}(\mathbb{1} + \vec{b} \cdot \vec{\sigma})$, where $\vec{a}, \vec{b}$ are unit vectors and $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices. A straightforward calculation then shows that the smallest norm is $\frac{1}{2}\sin\theta_{ab}$. We thus conclude that $D_\mathsf{A}(\mathsf{B}) = \frac{1}{2}\sin\theta_{ab} = \|\mathsf{A}_1\mathsf{B}_1 - \mathsf{B}_1\mathsf{A}_1\|$. The disturbance is therefore directly connected with the degree noncommutativity of A and B.

5.2. **Non-disturbance as a semidefinite program.** In the following we show that deciding whether an observable can be measured without disturbing another observable is a semidefinite program.

To cast the quantification problem of Subsec. 5.1 into a semidefinite program, we notice that the task is to find the smallest number λ under the conditions that there exists a collection $\{\mathcal{I}_x^*\}$ of linear maps satisfying

 (a) $\mathcal{I}_x^*(\mathbb{1}) = \mathsf{A}_x$ for all x,
 (b) $\mathcal{I}_x^*$ is completely positive for all x,
 (c) $-\lambda\mathbb{1} \leq \mathsf{B}_y - \mathcal{I}_\Omega^*(\mathsf{B}_y) \leq \lambda\mathbb{1}$ for all y.

The condition for completely positivity can be written as

$$(\mathcal{I}_x \otimes \operatorname{id})(\omega) \geq 0,$$

where ω is the maximally entangled state $\omega = \frac{1}{d}\sum_{j,k}|jj\rangle\langle kk|$.

To proceed, we fix a selfadjoint operator basis $\{E_i\}$ satisfying $\operatorname{tr}[E_i E_j] = \delta_{ij}$ and $E_0 = \frac{1}{\sqrt{d}}\mathbb{1}$. Conditions (a)-(c) can now be written as follows:

(a') $\pm\frac{1}{\sqrt{d}}\mathsf{A}_x \leq \pm\sum_j \operatorname{tr}[E_j \mathcal{I}_x^*(E_0)] E_j \quad \forall x$
(b') $0 \leq \sum_{ij} \operatorname{tr}[E_j \mathcal{I}_x^*(E_i)] E_j \otimes E_i^T \quad \forall x$
(c') $\pm\mathsf{B}_y \leq \lambda\sqrt{d}E_0 \pm \sum_{ijx} \operatorname{tr}[E_j \mathcal{I}_x^*(E_i)] \operatorname{tr}[E_i \mathsf{B}_y] E_j \quad \forall y$

To get (b') we have used the fact that the operator $\sum_i E_i \otimes E_i^T$ is proportional to the maximally entangled state ω.

We hence see that the task of minimizing λ under conditions (a')-(c') can, after combining the constraints using a direct sum, be brought to the form

$$\inf_{c\in\mathbb{R}^n} \left\{ \langle v \,|\, c \rangle \,|\, F_0 \leq \sum_{\ell=1}^{n} c_\ell F_\ell \right\}, \tag{18}$$

where $v \in \mathbb{R}^n$ is a vector and F_0, F_ℓ are Hermitian matrices. It is therefore a semidefinite program, and the dual problem is

$$\sup_{C\geq 0} \left\{ \operatorname{tr}[F_0 C] \,|\, \operatorname{tr}[F_\ell C] = v_\ell \; \forall\ell \right\}. \tag{19}$$

In the case under investigation, the dual takes the form

$$D_\mathsf{A}^*(\mathsf{B}) := \sup\left(\sum_x \operatorname{tr}[H_x \mathsf{A}_x] - \sum_y \operatorname{tr}[K_y \mathsf{B}_y] \right)$$

where the supremum is taken over all selfadjoint operators H_x, K_y satisfying

(d) $H_x \otimes \mathbb{1} \leq \sum_y K_y \otimes \mathsf{B}_y^T$ for all x
(e) $\sum_y \operatorname{tr}[|K_y|] = 1$.

By the general theory of semidefinite programs [3], $D_\mathsf{A}^*(\mathsf{B})$ is automatically a lower bound on $D_\mathsf{A}(\mathsf{B})$. Actually, when written in the standard form (19) the dual program is seen to be strictly feasible. It follows that $D_\mathsf{A}^*(\mathsf{B}) = D_\mathsf{A}(\mathsf{B})$ and the extremum is attained for $D_\mathsf{A}(\mathsf{B})$. This also means that the proposed measure $D_\mathsf{A}(\mathsf{B})$ can be efficiently computed numerically for any given pair A and B, and that the obtained result can be certified by the solution of the dual.

References

[1] A. Arias, A. Gheondea, and S. Gudder. Fixed points of quantum operations. *J. Math. Phys.*, 43:5872–5881, 2002.

[2] W. Arveson. Subalgebras of C^*-algebras. II. *Acta Math.*, 128:271–308, 1972.

[3] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.

[4] O. Bratteli, P.E.T. Jorgensen, A. Kishimoto, and R.F. Werner. Pure states on O_d. *J. Operator Theory*, 43:97–143, 2000.

[5] P. Busch and T. Heinosaari. Approximate joint measurements of qubit observables. *Quant. Inf. Comp.*, 8:0797–0818, 2008.

[6] S.T. Ali and H.D. Doebner. On the equivalence of nonrelativistic quantum mechanics based upon sharp and fuzzy measurements. *J. Math. Phys.*, 17:1105–1111, 1976.

[7] P. Busch, P.J. Lahti, and P. Mittelstaedt. *The Quantum Theory of Measurement*. Springer-Verlag, Berlin, second revised edition, 1996.

[8] P. Busch and J. Singh. Lüders theorem for unsharp quantum measurements. *Phys. Lett. A*, 249:10–12, 1998.

[9] E.B. Davies. *Quantum Theory of Open Systems*. Academic Press, London, 1976.

[10] T. Heinosaari, D. Reitzner, and P. Stano. Notes on joint measurability of quantum observables. *Found. Phys.*, 38:1133–1147, 2008.

[11] G. Lindblad. A general no-cloning theorem. *Lett. Math. Phys.*, 47:189–196, 1999.

[12] L. Weihua and W. Junde. On fixed points of Lüders operation. *J. Math. Phys.*, 50:103531, 2009.

[13] M.M. Wolf. Quantum channels & operations. Lecture notes, available in $www.nbi.dk/ \sim wolf/notes.pdf$, 2010.

[14] M.M. Wolf, D. Perez-Garcia, and C. Fernandez. Measurements incompatible in quantum theory cannot be measured jointly in any other no-signaling theory. *Phys. Rev. Lett.*, 103:230402, 2009.

Teiko Heinosaari, Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark

   E-mail address: heinosaari@nbi.dk

Michael M. Wolf, Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark

   E-mail address: wolf@nbi.dk