Quantum Information and Computation, Vol. 8, No. 8&9 (2008) 0797–0818
© Rinton Press

APPROXIMATE JOINT MEASUREMENTS OF QUBIT OBSERVABLES

PAUL BUSCH
Department of Mathematics, University of York
York, UK

TEIKO HEINOSAARI
Research Center for Quantum Information, Slovak Academy of Sciences
Bratislava, Slovakia
and
Department of Physics, University of Turku
Turku, Finland

Received August 10, 2007
Revised April 4, 2008

Joint measurements of qubit observables have recently been studied in conjunction with quantum information processing tasks such as cloning. Considerations of such joint measurements have until now been restricted to a certain class of observables that can be characterized by a form of covariance. Here we investigate conditions for the joint measurability of arbitrary pairs of qubit observables. For pairs of noncommuting sharp qubit observables, a notion of approximate joint measurement is introduced. Optimal approximate joint measurements are shown to lie in the class of covariant joint measurements. The marginal observables found to be optimal approximators are generally not among the coarse-grainings of the observables to be approximated. This yields scope for the improvement of existing joint measurement schemes. Both the quality of the approximations and the intrinsic unsharpness of the approximators are shown to be subject to Heisenberg-type uncertainty relations.

Keywords: joint measurements, unsharp observables, approximate measurements, inaccuracy relation
Communicated by: R Cleve & R Jozsa

1 Introduction

In recent years there has been an increasing interest in the question of joint measurability of noncommuting quantum observables, both from a foundational [1, 2, 3, 4, 5, 6] and quantum information theoretical [7, 8, 9, 10] perspective. The connection of this issue with certain impossible tasks in quantum mechanics, such as universal copier and Bell's telephone, is lucidly explained in [11]. Since two observables represented as selfadjoint operators do not have a joint observable if they do not commute, it is necessary in such cases to understand joint measurability in a wider sense.

As intuitively understood by Heisenberg already in 1927 [12], one has to allow for a degree of imprecision in order to make room for a notion of joint measurement of noncommuting observables. This idea can be appropriately investigated if the wider class of observables represented as positive operator measures (POMs) is taken into consideration. Projection

valued measures among the POMs correspond to the standard observables represented as selfadjoint operators; they are called *sharp observables*.

In the class of POMs, there are pairs of noncommuting observables that possess a joint observable, which thus has these two observables as its marginals. Commutativity is necessary for joint measurability if at least one of the POMs is a sharp observable, but generally commutativity is not required. A joint measurement of two POMs $\mathcal{E}^1$ and $\mathcal{E}^2$ can be regarded as an *approximate joint measurement* of two noncommuting observables $\mathcal{A}$ and $\mathcal{B}$ if $\mathcal{E}^1, \mathcal{E}^2$ are close (in some suitable sense) to $\mathcal{A}, \mathcal{B}$, respectively.

The problem of approximate joint measurements of position and momentum has been treated comprehensively in related publications [13, 14, 15]. The case of observables with discrete spectra requires somewhat different concepts and will be treated in the present paper for the case of joint measurements of qubit observables. We will introduce an appropriate measure of the quality of the approximation of one observable by another observable. It will then be shown that the quality of approximations in an approximate joint measurement of two sharp observables is limited if these observables do not commute. This limitation can indeed be formulated rigorously as a form of Heisenberg uncertainty relation.

One factor limiting the accuracy in an approximate joint measurement of noncommuting sharp observables is the fact that the approximating marginal observables must have a sufficient degree of *intrinsic unsharpness* as a consequence of their joint measurability. Hence there is yet another form of Heisenberg uncertainty relation for appropriately defined degrees of unsharpness in joint measurements. The distinction between the relational feature of inaccuracy (distance between two POMs) and the intrinsic property of unsharpness (of an individual POM) was until now blurred due to the fact that joint measurements were considered in which the marginals were coarse-grained versions of the sharp observables to be approximated; in such cases the intrinsic unsharpness and the inaccuracy are interconnected.

A theory and first models of approximate joint measurements of qubit observables were presented for special cases in [16, 17]. In those works and all subsequent developments, only a restricted class of joint measurements was used to approximate two sharp spin components. Here this restriction will be lifted, thereby allowing one to determine optimal joint measurements and to formulate uncertainty relations that can be used to characterize the optimal cases.

The paper is organized as follows. In Section 2 we review the condition of joint measurability of two simple observables (i.e. observables representing yes-no measurements), and give a precise definition of the approximate joint measurability of two observables. These conditions and concepts are investigated in Sections 3-5 in the case of qubit observables. Section 6 gives our conclusions and an outlook.

2 Simple observables and their (approximate) joint measurability

2.1 Effects and observables

The general definition of an observable $\mathcal{A}$ as a positive operator measure (POM) reduces, in the case of measurements with finitely many outcomes ω_i, to the specification of a map $\omega_i \mapsto \mathcal{A}_i$, where the $\mathcal{A}_i$ are *effects*, that is, positive operators satisfying the ordering relation[a] $\mathbb{O} \leq \mathcal{A}_i \leq \mathbb{1}$. (Here $\mathbb{O}, \mathbb{1}$ are the null and unit operators, respectively.) Together with

[a]Relation $A \leq B$ for two selfadjoint operators A and B means that $\langle \psi \,|\, A\psi \rangle \leq \langle \psi \,|\, B\psi \rangle$ for every vector ψ.

any state (density operator) T, $\mathcal{A}$ determines a probability distribution over the outcomes of $\mathcal{A}$ via the trace formula, $\omega_i \mapsto \mathrm{tr}[T\mathcal{A}_i]$. The additivity and normalization of probability distributions is ensured by the condition $\sum_i \mathcal{A}_i = \mathbb{1}$.

A *simple observable* is one that represents a measurement with two possible outcomes; it is given as a POM with two values and associated effects,

$$\mathcal{A}: \quad \omega_+ \mapsto \mathcal{A}_+, \quad \omega_- \mapsto \mathcal{A}_-. \tag{1}$$

Normalization entails that $\mathcal{A}_+ + \mathcal{A}_- = \mathbb{1}$, so that a simple observable is commutative. We note that any effect A together with its complement effect $A' := \mathbb{1} - A$ defines a class of simple observables, distinguished only by their outcome sets $\{\omega_+, \omega_-\}$.[b]

As noted in the introduction, an observable (POM) is *sharp* if its effects are all projections. Otherwise an observable is called *unsharp*. A measure of the intrinsic unsharpness of an effect A and thus of the associated simple observable $\mathcal{A}$ that is independent of the outcomes of $\mathcal{A}$ is obtained as follows. Let σ_A denote the spectrum of an effect A, then the *spectral width* of A is the length of the smallest closed interval containing σ_A, that is: $\mathfrak{W}(\sigma_A) := \max \sigma_A - \min \sigma_A = \|A\| + \|A'\| - 1$. The *sharpness* of A is defined as

$$\mathfrak{S}(A) := \mathfrak{W}(\sigma_A) - \mathfrak{W}(\sigma_{AA'}). \tag{2}$$

The operator AA' can be written as $AA' = A^{\frac{1}{2}} A' A^{\frac{1}{2}}$. This shows that AA' arises from the sequential Lüders measurement of the complement effects A and A', and this is a motivation for the formula (2); see [18] for further discussion.

Since $\mathfrak{W}(\sigma_A) = \mathfrak{W}(\sigma_{A'})$, it follows that $\mathfrak{S}(A') = \mathfrak{S}(A)$. On can prove the following facts about the sharpness:[c] $\mathfrak{S}(A) \in [0,1]$; $\mathfrak{S}(A) = 0$ exactly when A is a trivial effect (i.e. $A = k\mathbb{1}$ for some $0 \leq k \leq 1$); $\mathfrak{S}(A) = 1$ exactly when A is a nontrivial projection. These are properties one would expect any measure of sharpness to possess: the measure should single out the perfectly sharp effects and the trivial effects.

The *sharpness* and *unsharpness* of a simple observable $\mathcal{A}$ may now be defined as

$$\mathfrak{S}(\mathcal{A}) := \mathfrak{S}(\mathcal{A}_+) = \mathfrak{S}(\mathcal{A}_-), \tag{3}$$

$$\mathfrak{U}(\mathcal{A}) := 1 - \mathfrak{S}(\mathcal{A})^2. \tag{4}$$

2.2 Joint measurability

Two observables are jointly measurable if there is a measurement scheme that allows the determination of the values of both observables. This means that the POM representing that joint measurement contains the two observables as marginals. In this way it is ensured that there is a joint probability distribution for each state. We spell out this definition for the case of a pair of simple observables.[d]

Two simple observables $\mathcal{E}^1$ and $\mathcal{E}^2$ are *jointly measurable* if there is an observable $\mathcal{G}$: $\omega_{ij} \mapsto \mathcal{G}_{ij}$, $i, j = \pm$, such that

$$\begin{aligned} \mathcal{E}^1_+ &= \mathcal{G}_{++} + \mathcal{G}_{+-}, \qquad & \mathcal{E}^1_- &= \mathcal{G}_{-+} + \mathcal{G}_{--}, \\ \mathcal{E}^2_+ &= \mathcal{G}_{++} + \mathcal{G}_{-+}, \qquad & \mathcal{E}^2_- &= \mathcal{G}_{+-} + \mathcal{G}_{--}. \end{aligned} \tag{5}$$

[b] For clarity we denote observables with script capital letters, while effects are denoted with italic letters.
[c] See [18]. These statements are proved for qubit observables in Subsection 3.3.
[d] For the general definition of joint measurability and a detailed discussion on this topic, see e.g. [19] and references given therein.

In this case the observables $\mathcal{E}^1$ and $\mathcal{E}^2$ are the *marginals* of $\mathcal{G}$, and we also say that $\mathcal{G}$ is a *joint observable* for $\mathcal{E}^1$ and $\mathcal{E}^2$. The outcomes ω_{ij} of $\mathcal{G}$ could be taken to be (or replaced by) the pairs (ω_i, ω_j).

The joint measurability of two simple observables $\mathcal{E}^1$ and $\mathcal{E}^2$ is equivalent to the statement [20] that there exists an operator $\mathcal{G}_{++}$ satisfying the following set of operator inequalities:

$$\begin{gathered} \mathbb{O} \leq \mathcal{G}_{++}, \quad \mathcal{G}_{++} \leq \mathcal{E}^1_+, \quad \mathcal{G}_{++} \leq \mathcal{E}^2_+, \\ \mathcal{E}^1_+ + \mathcal{E}^2_+ - \mathbb{1} \leq \mathcal{G}_{++}. \end{gathered} \tag{6}$$

In fact, these inequalities ensure that the following four operators are effects:

$$\begin{gathered} \mathcal{G}_{++}, \quad \mathcal{G}_{+-} \equiv \mathcal{E}^1_+ - \mathcal{G}_{++}, \quad \mathcal{G}_{-+} \equiv \mathcal{E}^2_+ - \mathcal{G}_{++}, \\ \mathcal{G}_{--} \equiv \mathbb{1} - \mathcal{G}_{++} - \mathcal{G}_{+-} - \mathcal{G}_{-+} = \mathbb{1} - \mathcal{E}^1_+ - \mathcal{E}^2_+ + \mathcal{G}_{++}. \end{gathered} \tag{7}$$

It is straightforward to verify that equations (5) hold and hence, these effects define a joint observable $\mathcal{G}$ for $\mathcal{E}^1$ and $\mathcal{E}^2$.

The joint measurability condition for two simple observables can be interpreted as the requirement that the intersection of four cones in the set of effects is nonempty. The order relation $A \leq B$ for two selfadjoint operators is equivalent to either of $B - A \geq \mathbb{O}$ and $A - B \leq \mathbb{O}$. The condition $A \geq \mathbb{O}$ defines a convex cone in the real vector space of selfadjoint operators.[e] Thus we can define the upward and downward cones of a selfadjoint operator A as $\mathcal{C}^\vee(A) := \{B : A \leq B\}$ and $\mathcal{C}^\wedge(A) := \{B : A \geq B\}$. The joint measurability condition for observables $\mathcal{E}^1$ and $\mathcal{E}^2$ now reads:

$$\mathcal{C}^\vee(\mathbb{O}) \cap \mathcal{C}^\wedge(\mathcal{E}^1_+) \cap \mathcal{C}^\wedge(\mathcal{E}^2_+) \cap \mathcal{C}^\vee(\mathcal{E}^1_+ + \mathcal{E}^2_+ - \mathbb{1}) \neq \emptyset. \tag{8}$$

Example 1 [Trivial cases of joint measurability]: In the following four cases, joint measurability falls out trivially:
(a) $\mathcal{E}^1_+ \geq \mathcal{E}^2_+$: put $\mathcal{G}_{++} = \mathcal{E}^2_+$, $\mathcal{G}_{+-} = \mathcal{E}^1_+ - \mathcal{E}^2_+$, $\mathcal{G}_{-+} = \mathbb{O}$, $\mathcal{G}_{--} = \mathbb{1} - \mathcal{E}^1_+$.
(b) $\mathcal{E}^1_+ \leq \mathcal{E}^2_+$: put $\mathcal{G}_{++} = \mathcal{E}^1_+$, $\mathcal{G}_{+-} = \mathbb{O}$, $\mathcal{G}_{-+} = \mathcal{E}^2_+ - \mathcal{E}^1_+$, $\mathcal{G}_{--} = \mathbb{1} - \mathcal{E}^2_+$.
(c) $\mathcal{E}^1_+ \geq \mathcal{E}^2_-$: put $\mathcal{G}_{++} = \mathcal{E}^1_+ + \mathcal{E}^2_+ - \mathbb{1}$, $\mathcal{G}_{+-} = \mathcal{E}^2_-$, $\mathcal{G}_{-+} = \mathcal{E}^1_-$, $\mathcal{G}_{--} = \mathbb{O}$.
(d) $\mathcal{E}^1_+ \leq \mathcal{E}^2_-$: put $\mathcal{G}_{++} = \mathbb{O}$, $\mathcal{G}_{+-} = \mathcal{E}^1_+$, $\mathcal{G}_{-+} = \mathcal{E}^2_+$, $\mathcal{G}_{--} = \mathbb{1} - \mathcal{E}^1_+ - \mathcal{E}^2_+$.

We conclude that interesting (i.e. nontrivial) cases arise when $\mathcal{E}^1_+ - \mathcal{E}^2_+$ and $\mathcal{E}^1_+ + \mathcal{E}^2_+ - \mathbb{1}$ are neither $\leq \mathbb{O}$ nor $\geq \mathbb{O}$. In terms of a joint observable $\mathcal{G}$, nontrivial cases are exactly those in which $\mathcal{G}_{ij} \neq \mathbb{O}$ for every $i, j = \pm$.

Another simple instance of joint measurability arises from commutativity.

Example 2 [Mutually commuting observables]: If $\mathcal{E}^1$ and $\mathcal{E}^2$ commute mutually in the sense that $\mathcal{E}^1_i \mathcal{E}^2_j = \mathcal{E}^2_j \mathcal{E}^1_i$ for every $i, j = \pm$, then they are jointly measurable. In this case the formula $\mathcal{G}_{ij} = \mathcal{E}^1_i \mathcal{E}^2_j$ defines a joint observable.

It is instructive to prove the following well-known proposition, which supplements Example 2. (Its statement is valid also for observables that are not simple, and then the proof

[e] This means that whenever $A, B \geq \mathbb{O}$, then $tA + (1-t)B \geq \mathbb{O}$ for any $t \in [0, 1]$.

requires only minor changes.)

Proposition 1: *Let $\mathcal{E}^1$ and $\mathcal{E}^2$ be simple observables which are jointly measurable. If one of them is a sharp observable, then they commute and the unique joint observable $\mathcal{G}$ is of the product form $\mathcal{G}_{ij} = \mathcal{E}^1_i \mathcal{E}^2_j$.*

Proof: Let, for instance, $\mathcal{E}^1$ be a sharp observable and suppose that $\mathcal{G}$ is a joint observable for $\mathcal{E}^1$ and $\mathcal{E}^2$. Since $\mathcal{G}_{ij} \leq \mathcal{E}^1_i$, one obtains $\mathcal{E}^1_i \mathcal{G}_{ij} = \mathcal{G}_{ij}\mathcal{E}^1_i = \mathcal{G}_{ij}$. This shows also that $\mathcal{E}^1_i \mathcal{G}_{-ij} = (\mathbb{1} - \mathcal{E}^1_{-i})\mathcal{G}_{-ij} = \mathbb{O}$ and similarly $\mathcal{G}_{-ij}\mathcal{E}^1_i = \mathbb{O}$. It follows that

$$\mathcal{E}^1_i \mathcal{E}^2_j = \mathcal{E}^1_i (\mathcal{G}_{+j} + \mathcal{G}_{-j}) = \mathcal{G}_{ij} \tag{9}$$

and

$$\mathcal{E}^2_j \mathcal{E}^1_i = (\mathcal{G}_{+j} + \mathcal{G}_{-j})\mathcal{E}^1_i = \mathcal{G}_{ij}. \tag{10}$$

A comparison of these equations proves the claim. □

We note that in general two observables may have many different joint observables; this fact will be demonstrated in Section 4.3.

For later use we recall the following general fact. The set of observables on a fixed outcome space is convex: for two observables $\mathcal{E}, \mathcal{F}$ and any $t \in [0,1]$, a new observable $t\mathcal{E} + (1-t)\mathcal{F}$ is defined as $\omega_i \mapsto t\mathcal{E}_i + (1-t)\mathcal{F}_i$.

Proposition 2: *Let $(\mathcal{E}^1, \mathcal{E}^2)$ and $(\mathcal{F}^1, \mathcal{F}^2)$ be two pairs of jointly measurable observables. Then for any $t \in [0,1]$, the observables $t\mathcal{E}^1 + (1-t)\mathcal{F}^1$ and $t\mathcal{E}^2 + (1-t)\mathcal{F}^2$ are jointly measurable.*

Proof: Let $\mathcal{G}$ be a joint observable of $\mathcal{E}^1, \mathcal{E}^2$ and $\mathcal{H}$ of $\mathcal{F}^1, \mathcal{F}^2$. Then $t\mathcal{G} + (1-t)\mathcal{H}$ is a joint observable of $t\mathcal{E}^1 + (1-t)\mathcal{F}^1$ and $t\mathcal{E}^2 + (1-t)\mathcal{F}^2$. □

2.3 Approximate joint measurability

Assume that two observables $\mathcal{A}$ and $\mathcal{B}$ do not have a joint measurement. We may still ask if it could be possible to obtain some information on both observables in a single measurement scheme. One way of approaching this task is to consider whether there are two jointly measurable observables $\mathcal{E}^1, \mathcal{E}^2$ that are close to $\mathcal{A}, \mathcal{B}$, respectively, in a sense to be determined. Any joint measurement of $\mathcal{E}^1$ and $\mathcal{E}^2$ can then be regarded as an approximate joint measurement of $\mathcal{A}$ and $\mathcal{B}$.

A natural characterization of the closeness between two observables (assumed to have the same outcome space) is based on the degree of similarity of their associated probability distributions for all states. Hence we define the distance between observables $\mathcal{A}$ and $\mathcal{B}$ in the following way:

$$\mathfrak{D}(\mathcal{A}, \mathcal{B}) := \max_j \sup_T \left| \mathrm{tr}[T\mathcal{A}_j] - \mathrm{tr}[T\mathcal{B}_j] \right| = \max_j \left\| \mathcal{A}_j - \mathcal{B}_j \right\|. \tag{11}$$

Clearly, $0 \leq \mathfrak{D}(\mathcal{A}, \mathcal{B}) \leq 1$, and $\mathfrak{D}(\mathcal{A}, \mathcal{B}) = 0$ if and only if $\mathcal{A} = \mathcal{B}$. Moreover, the triangle inequality holds for a triple of observables, so that $\mathfrak{D}$ is indeed a metric. If $\mathcal{A}$ and $\mathcal{B}$ are simple

observables, then $\mathcal{A}_- - \mathcal{B}_- = \mathcal{B}_+ - \mathcal{A}_+$ and therefore

$$\mathfrak{D}(\mathcal{A}, \mathcal{B}) = \|\mathcal{A}_+ - \mathcal{B}_+\| = \|\mathcal{A}_- - \mathcal{B}_-\| \,. \tag{12}$$

A conventional approach to realizing approximate joint measurements consists of replacing the observables $\mathcal{A}, \mathcal{B}$ to be approximated with some coarse-grained versions $\mathcal{E}^1, \mathcal{E}^2$. Here we briefly illustrate this approach in the case of simple observables from the perspective of the general framework. We refer to [2, Chapter 7] for a review and examples on this topic.

If $\mathcal{A}$ and $\mathcal{B}$ are simple observables, one defines, using 2×2 stochastic matrices[f] $(\lambda_{ik}), (\mu_{j\ell})$, the coarse-grainings $\mathcal{E}^1$ and $\mathcal{E}^2$ by

$$\mathcal{E}^1_i = \lambda_{i+}\mathcal{A}_+ + \lambda_{i-}\mathcal{A}_-, \quad \mathcal{E}^2_j = \mu_{j+}\mathcal{B}_+ + \mu_{j-}\mathcal{B}_-. \tag{13}$$

We expect that $\mathcal{E}^1$ is a good approximation of $\mathcal{A}$ if λ_{++} is close to 1 and λ_{+-} is close to 0. Indeed, if, for instance, $\mathcal{A}$ is a sharp observable then

$$\mathfrak{D}(\mathcal{E}^1, \mathcal{A}) = \max\{1 - \lambda_{++}, \lambda_{+-}\}. \tag{14}$$

To give an example of jointly measurable observables, assume that $\mathcal{E}^1$ and $\mathcal{E}^2$ are defined as in (13). Since we want to approximate $\mathcal{A}$ by $\mathcal{E}^1$ and $\mathcal{B}$ by $\mathcal{E}^2$, it is natural to require that

$$\lambda_{++} \geq \lambda_{+-}, \quad \mu_{++} \geq \mu_{+-}. \tag{15}$$

In fact, if these inequalities do not hold, one can choose $\lambda'_{ik} = \lambda_{-ik}$, $\mu'_{jl} = \mu_{-jl}$ to obtain new coarse-grainings which do satisfy the inequalities.

Now, define $\mathcal{G}_{++} = \min\{\lambda_{+-}, \mu_{+-}\}\,\mathbb{1}$. Condition (15) implies that $\mathbb{O} \leq \mathcal{G}_{++} \leq \mathcal{E}^1_+$ and $\mathcal{G}_{++} \leq \mathcal{E}^2_+$. The remaining inequality required for joint measurability, $\mathcal{E}^1_+ + \mathcal{E}^2_+ - \mathbb{1} \leq \mathcal{G}_{++}$, depends on the specific structure of the observables $\mathcal{A}$ and $\mathcal{B}$; it is ensured to hold independently of $\mathcal{A}$ and $\mathcal{B}$ if

$$\lambda_{++} + \mu_{++} \leq 1 + \min\{\lambda_{+-}, \mu_{+-}\}. \tag{16}$$

This shows that one can always construct (nontrivial) jointly measurable coarse-grainings; a possible choice is, for instance, $\lambda_{++} = \mu_{++} = \frac{2}{3}$ and $\lambda_{+-} = \mu_{+-} = \frac{1}{3}$.

3 Qubit observables

3.1 Effects and observables

In the 2-dimensional Hilbert space of a qubit one can take the unit operator $\mathbb{1}$ together with the Pauli operators $\sigma_1, \sigma_2, \sigma_3$ as a basis of the real vector space of selfadjoint linear operators. The latter can be defined with respect to any fixed basis of orthogonal unit vectors φ_+, φ_- so that the usual relations are satisfied: $\sigma_3\varphi_\pm = \pm\varphi_\pm$, $\sigma_1\varphi_\pm = \varphi_\mp$, $\sigma_2\varphi_\pm = \pm i\varphi_\mp$. We will write $\boldsymbol{\sigma}$ for the operator triple $(\sigma_1, \sigma_2, \sigma_3)$. States of a qubit can be written in the form $T_{\mathbf{r}} = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\sigma})$, where $\mathbf{r} \in \mathbb{R}^3$ and $\|\mathbf{r}\| \leq 1$. The pure states are characterized by the condition $\|\mathbf{r}\| = 1$.

[f] A stochastic matrix is a square matrix whose entries are non-negative real numbers and each column sums to 1.

For each $(\alpha, \mathbf{a}) \in \mathbb{R}^4$, we denote

$$A(\alpha, \mathbf{a}) := \frac{1}{2}\left(\alpha \mathbb{1} + \mathbf{a} \cdot \boldsymbol{\sigma}\right). \tag{17}$$

The eigenvalues of the operator $A(\alpha, \mathbf{a})$ are $\frac{1}{2}(\alpha \pm \|\mathbf{a}\|)$. Hence, $A(\alpha, \mathbf{a})$ is an effect if

$$\|\mathbf{a}\| \leq \alpha \leq 2 - \|\mathbf{a}\|, \tag{18}$$

which implies, in particular, that $\|\mathbf{a}\| \leq 1$. The operator $A(\alpha, \mathbf{a})$ is a nontrivial projection if

$$\alpha = \|\mathbf{a}\| = 1. \tag{19}$$

The spectral decomposition of the effect $A(\alpha, \mathbf{a})$, $\mathbf{a} \neq 0$, is (putting $\hat{\mathbf{a}} := \|\mathbf{a}\|^{-1}\mathbf{a}$)

$$A(\alpha, \mathbf{a}) = \frac{1}{2}(\alpha + \|\mathbf{a}\|)A(1, \hat{\mathbf{a}}) + \frac{1}{2}(\alpha - \|\mathbf{a}\|)A(1, -\hat{\mathbf{a}}). \tag{20}$$

For later use we note the commutator of two effects $A(\alpha, \mathbf{a})$ and $A(\beta, \mathbf{b})$:

$$[A(\alpha, \mathbf{a}), A(\beta, \mathbf{b})] = \tfrac{1}{2}\left(\mathbf{a} \times \mathbf{b}\right) \cdot \boldsymbol{\sigma}. \tag{21}$$

Since the Hilbert space of a qubit is 2-dimensional, any sharp qubit observable is (effectively) simple. It is clear that this restriction does not apply for a qubit observable in general; one can write the identity operator $\mathbb{1}$ as a sum of arbitrarily many different effects. It is also known that for some quantum informational tasks, such as unambiguous state discrimination, one needs other qubit observables than the simple ones; see e.g. [21]. Here we shall, however, concentrate on simple qubit observables as our aim is to study approximate joint measurements of sharp qubit observables. To clarify further the nature of sharp qubit observables, we note that again due to the low dimensionality, the projections that constitute such an observable are of rank one, which implies that their repeatable measurements are von Neumann measurements [1].

We denote by $\mathcal{E}^{\alpha,\mathbf{a}}$ the simple qubit observable defined as

$$\begin{aligned} \omega_+ &\mapsto \mathcal{E}^{\alpha,\mathbf{a}}_+ := A(\alpha, \mathbf{a}), \\ \omega_- &\mapsto \mathcal{E}^{\alpha,\mathbf{a}}_- := \mathbb{1} - A(\alpha, \mathbf{a}) = A(2-\alpha, -\mathbf{a}). \end{aligned}$$

A special case is given by the sharp observables $\mathcal{E}^{1,\hat{\mathbf{a}}}$. The spectral decomposition (20) of $A(\alpha, \mathbf{a})$ shows that $\mathcal{E}^{\alpha,\mathbf{a}}$ is a coarse-graining of $\mathcal{E}^{1,\hat{\mathbf{a}}}$.

From the above commutator formula we recover the well known fact that the observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ commute exactly when the vectors $\mathbf{a}$ and $\mathbf{b}$ are collinear. Together with Proposition 1, this shows that an observable $\mathcal{E}^{\beta,\mathbf{b}}$ is jointly measurable with a sharp observable $\mathcal{E}^{1,\hat{\mathbf{a}}}$ if and only if $\mathcal{E}^{\beta,\mathbf{b}}$ is a coarse-graining of $\mathcal{E}^{1,\hat{\mathbf{a}}}$. A joint measurement of that kind is of little value; one can simply measure $\mathcal{E}^{1,\hat{\mathbf{a}}}$ alone to get the same information.

3.2 Covariance

Let U be a unitary operator describing some symmetry transformation of the system. We assume that $U^2 = \mathbb{1}$, so that $\{\mathbb{1}, U\}$ form a two-element group. In other words, U is a

selfadjoint unitary operator. We say that an observable $\mathcal{E}^{\alpha,\mathbf{a}}$ is *covariant with respect to* U, or U-*covariant* for short, if

$$U\mathcal{E}_{+}^{\alpha,\mathbf{a}}U = \mathcal{E}_{-}^{\alpha,\mathbf{a}}. \tag{22}$$

This covariance condition means that the symmetry transformation described by U swaps the outcomes of the observable $\mathcal{E}^{\alpha,\mathbf{a}}$ but has no other effect on its measurement outcome distributions.

Effects $\mathcal{E}_{+}^{\alpha,\mathbf{a}}$ and $\mathcal{E}_{-}^{\alpha,\mathbf{a}}$ can be unitarily equivalent only if they have the same eigenvalues, which is the case exactly when $\alpha = 1$. Hence, $\mathcal{E}^{\alpha,\mathbf{a}}$ can be covariant only if $\alpha = 1$. Assume that $\alpha = 1$ and fix a unit vector $\mathbf{u} \in \mathbb{R}^3$ orthogonal to $\mathbf{a}$. The operator $U = \mathbf{u} \cdot \boldsymbol{\sigma}$ is a selfadjoint unitary operator and

$$U\mathcal{E}_{+}^{1,\mathbf{a}}U = \mathcal{E}_{-}^{1,\mathbf{a}}. \tag{23}$$

Moreover, any selfadjoint unitary operator U satisfying (23) is of the form $U = \mathbf{u} \cdot \boldsymbol{\sigma}$ for some unit vector $\mathbf{u}$ orthogonal to $\mathbf{a}$.

In [3], an observable $\mathcal{E}^{\alpha,\mathbf{a}}$ was selected in relation to a sharp observable $\mathcal{E}^{1,\mathbf{n}}$ by the requirement that the expectation values of $\mathcal{E}^{\alpha,\mathbf{a}}$ are proportional to those of $\mathcal{E}^{1,\mathbf{n}}$. This requirement, called there *unbiasedness*, is equivalent with the fact that $\alpha = 1$ and the vectors $\mathbf{a}$ and $\mathbf{n}$ are parallel. Hence, the unbiasedness requirement means that $\mathcal{E}^{\alpha,\mathbf{a}}$ is covariant with respect to the same unitary operators as $\mathcal{E}^{1,\mathbf{n}}$, i.e., the observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{n}}$ have the same symmetry properties.

3.3 Sharpness and distance

The spectral width of an operator acting on two dimensional Hilbert space is simply the difference of its greater and lower eigenvalues. The sharpness of an observable $\mathcal{E}^{\alpha,\mathbf{a}}$ is thus found to be

$$\mathfrak{S}(\mathcal{E}^{\alpha,\mathbf{a}}) = \|\mathbf{a}\|\,(1 - |1-\alpha|) = \|\mathbf{a}\| \min\{\alpha, 2-\alpha\}. \tag{24}$$

With this expression one can easily confirm the statements of Section 2.1 for simple qubit observables: $\mathfrak{S}(\mathcal{E}^{\alpha,\mathbf{a}}) = 1$ exactly when $\mathcal{E}^{\alpha,\mathbf{a}}$ is a sharp observable and $\mathfrak{S}(\mathcal{E}^{\alpha,\mathbf{a}}) = 0$ exactly when $\mathcal{E}^{\alpha,\mathbf{a}}$ is a trivial observable. We also note the following useful observation:

$$\mathfrak{S}(\mathcal{E}^{\alpha,\mathbf{a}}) \leq \mathfrak{S}(\mathcal{E}^{1,\mathbf{a}}). \tag{25}$$

The distance between two qubit observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ is given by the formula

$$\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{\beta,\mathbf{b}}) = \tfrac{1}{2}\,\|\mathbf{a} - \mathbf{b}\| + \tfrac{1}{2}|\alpha - \beta|. \tag{26}$$

This shows, in particular, that the distance of a given observable $\mathcal{E}^{\alpha,\mathbf{a}}$ from any sharp observable $\mathcal{E}^{1,\mathbf{n}}$ is minimal when $\mathbf{n} = \hat{\mathbf{a}}$, or in other words, when $\mathcal{E}^{\alpha,\mathbf{a}}$ is a coarse-graining of $\mathcal{E}^{1,\mathbf{n}}$. In this case we have

$$\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\hat{\mathbf{a}}}) = \tfrac{1}{2}\,(1 - \|\mathbf{a}\|) + \tfrac{1}{2}|1-\alpha|. \tag{27}$$

We also note the following:

$$\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\hat{\mathbf{a}}}) \geq \mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\hat{\mathbf{a}}}). \tag{28}$$

Finally, from equations (24) and (27) we get the following relations:

$$\begin{aligned} \mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}) + \tfrac{1}{2}\mathfrak{S}(\mathcal{E}^{\alpha,\mathbf{a}}) &\geq \mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\hat{\mathbf{a}}}) + \tfrac{1}{2}\mathfrak{S}(\mathcal{E}^{\alpha,\mathbf{a}}) \\ &\geq \mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\hat{\mathbf{a}}}) + \tfrac{1}{2}\mathfrak{S}(\mathcal{E}^{1,\mathbf{a}}) = \tfrac{1}{2}. \end{aligned} \tag{29}$$

The last equation shows that the distance between $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\hat{\mathbf{a}}}$ is directly related to the sharpness of $\mathcal{E}^{1,\mathbf{a}}$. This is not surprising when we recall that $\mathcal{E}^{1,\mathbf{a}}$ is a coarse-graining of $\mathcal{E}^{1,\hat{\mathbf{a}}}$.

4 Joint measurability of qubit observables

4.1 General criterion for joint measurability

The joint measurability conditions (8) applied to two qubit observables $\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{\beta,\mathbf{b}}$ takes the following form: there exists an operator $\mathcal{G}_{++} = \frac{1}{2}(\gamma\mathbb{1} + \mathbf{g}\cdot\boldsymbol{\sigma})$ such that

$$\|\mathbf{g}\| \leq \gamma; \tag{30}$$

$$\|\mathbf{a}-\mathbf{g}\| \leq \alpha-\gamma; \tag{31}$$

$$\|\mathbf{b}-\mathbf{g}\| \leq \beta-\gamma; \tag{32}$$

$$\|\mathbf{a}+\mathbf{b}-\mathbf{g}\| \leq 2+\gamma-\alpha-\beta. \tag{33}$$

Let $\mathbf{B}(\mathbf{x}, r)$ denote the closed ball with center $\mathbf{x}$ and radius r. Then it is seen that the joint measurability of $\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{\beta,\mathbf{b}}$ is equivalent to the statement that there exists a number $\gamma \geq 0$ such that the intersection of four balls is non-empty:

$$\mathbf{B}(\mathbf{0},\gamma) \cap \mathbf{B}(\mathbf{a},\alpha-\gamma) \cap \mathbf{B}(\mathbf{b},\beta-\gamma) \cap \mathbf{B}(\mathbf{a}+\mathbf{b}, 2+\gamma-\alpha-\beta) \neq \emptyset. \tag{34}$$

The criterion (34) immediately gives the following as a necessary condition for joint measurability: the two pairs of balls diagonally opposite to each other must have separations which are no greater than the sum of their radii; thus, there must be a $\gamma \geq 0$ such that

$$\|\mathbf{a}-\mathbf{b}\| \leq \alpha+\beta-2\gamma, \tag{35}$$

$$\|\mathbf{a}+\mathbf{b}\| \leq 2-\alpha-\beta+2\gamma, \tag{36}$$

or equivalently,

$$\gamma_1 := \tfrac{1}{2}\|\mathbf{a}+\mathbf{b}\| + \tfrac{1}{2}[\alpha+\beta-2] \leq \gamma \leq \tfrac{1}{2}[\alpha+\beta] - \tfrac{1}{2}\|\mathbf{a}-\mathbf{b}\| =: \gamma_2. \tag{37}$$

This gives an interval for γ to lie in which has to be nonempty. Therefore the following is a necessary joint measurability condition:

$$\gamma_2 - \gamma_1 = 1 - [\tfrac{1}{2}\|\mathbf{a}+\mathbf{b}\| + \tfrac{1}{2}\|\mathbf{a}-\mathbf{b}\|] \geq 0. \tag{38}$$

Proposition 3: *If observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ are jointly measurable, then*[g]

$$\|\mathbf{a}+\mathbf{b}\| + \|\mathbf{a}-\mathbf{b}\| \leq 2. \tag{39}$$

[g]This condition has the following geometric meaning: for a observable $\mathcal{E}^{1,\mathbf{a}}$, a jointly measurable observable $\mathcal{E}^{1,\mathbf{b}}$ is such that the vector $\mathbf{b}$ is inside a prolate spheroid. The center of the spheroid is in the origin and its major axis is in the direction of $\mathbf{a}$. The polar radius of the spheroid is 1 and the equatorial radius is $(1-\|\mathbf{a}\|^2)^{1/2}$. In fact, in coordinates for which $\mathbf{a}$ is in the z-direction, the inequality becomes $b_x^2 + b_y^2 + (1-a^2)b_z^2 \leq 1-a^2$, to be read as a condition for $\mathbf{b}$.

In the case of covariant qubit observables (for which $\alpha = \beta = 1$) the condition (39) is found to be also sufficient for joint measurability, as was shown in [16]. A new proof of this fact, stated below, will arise as a corollary of our investigation in Subsection 4.3.

Proposition 4: *Observables $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are jointly measurable if and only if inequality* (39) *holds.*

In the following example we demonstrate that (39) is not sufficient in general to guarantee the joint measurability of observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$.

Example 3: Let us consider the case where the vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal and equality holds in (39), or in other words, $\|\mathbf{a}+\mathbf{b}\| = \|\mathbf{a}-\mathbf{b}\| = 1$. Assume that $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ are jointly measurable observables. We have $\gamma = \gamma_1 = \gamma_2$ and therefore, there is only one point $\mathbf{g}$ in the intersection $\mathbf{B}(\mathbf{a}, \alpha-\gamma) \cap \mathbf{B}(\mathbf{b}, \beta-\gamma)$, and similarly in the intersection $\mathbf{B}(\mathbf{0}, \gamma) \cap \mathbf{B}(\mathbf{a}+\mathbf{b}, \gamma+2-\alpha-\beta)$. Thus, $\mathbf{g}$ is in the boundary of $\mathbf{B}(\mathbf{a}, \alpha-\gamma)$ and it must satisfy the equation

$$\mathbf{g} = \mathbf{a} + (\alpha-\gamma)(\mathbf{b}-\mathbf{a}) \tag{40}$$

and three similar equations corresponding to the other balls. These equations taken together imply that $\alpha = \beta = 1$. As condition (39) does not restrict α and β, we conclude that (39) is not sufficient to ensure the joint measurability of $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$. In fact, we could have chosen $\alpha = \beta = \|\mathbf{a}\| = \|\mathbf{b}\| = 1/\sqrt{2}$, in which case the observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ are not jointly measurable although (39) is satisfied.

Propositions 3 and 4 lead to the following observation, which we will need later.

Proposition 5: *If $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ are jointly measurable, then also $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are jointly measurable.*

4.2 Sufficient conditions for joint measurability

The problem of finding necessary *and* sufficient conditions for the joint measurability of a pair of qubit observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ beyond the above case of $\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{b}}$ has only recently been solved by the present authors in different collaborations. In [22], this is achieved by analyzing the sphere intersection condition (34), whereas in [23] the cone intersection condition (8) is elucidated. The sets of inequalities found for $\alpha, \mathbf{a}, \beta, \mathbf{b}$ are rather involved and not easily comparable, hence we refrain from reproducing them here. Instead we give a sufficient condition for the joint measurability of $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ which is an obvious strengthening of (39). The fact that this stronger condition may appear quite natural at first sight but is actually not necessary highlights the intricate nature of the general problem solved in [22] and [23].

First we identify two distinguished effects $A_1 := A(\gamma_1, \mathbf{g}_1)$ and $A_2 := A(\gamma_2, \mathbf{g}_2)$, where

γ_1, γ_2 are the parameters from Eq. (37) and

$$\begin{aligned} \mathbf{g}_1 &= \frac{1}{2}\left[1 - \frac{2-\alpha-\beta}{\|\mathbf{a}+\mathbf{b}\|}\right](\mathbf{a}+\mathbf{b}), \\ \mathbf{g}_2 &= \frac{1}{2}(\mathbf{a}+\mathbf{b}) - \frac{\alpha-\beta}{\|\mathbf{a}-\mathbf{b}\|}\frac{1}{2}(\mathbf{a}-\mathbf{b}). \end{aligned} \tag{41}$$

The effect A_1 is in $\mathcal{C}^\vee(\mathbb{O}) \cap \mathcal{C}^\vee(\mathcal{E}_+^1 + \mathcal{E}_+^2 - \mathbb{1})$ and it is the unique element of all effects $A(\gamma, \mathbf{g})$ in that intersection with the lowest possible γ. The effect A_2 is in $\mathcal{C}^\wedge(\mathcal{E}_+^1) \cap \mathcal{C}^\wedge(\mathcal{E}_+^2)$ and it is the unique element of all effects $A(\gamma, \mathbf{g})$ in that intersection with the greatest possible γ. Now, joint measurability is guaranteed if $A_1 \leq A_2$, which is equivalent to the condition:

$$\|\mathbf{a}+\mathbf{b}\| + \|\mathbf{a}-\mathbf{b}\| + \left\|\frac{2-\alpha-\beta}{\|\mathbf{a}+\mathbf{b}\|}(\mathbf{a}+\mathbf{b}) - \frac{\alpha-\beta}{\|\mathbf{a}-\mathbf{b}\|}(\mathbf{a}-\mathbf{b})\right\| \leq 2. \tag{42}$$

It is not hard to verify that this condition is automatically satisfied in all trivial and commutative cases, where joint measurability is given. Furthermore it follows from the stronger condition

$$\|\mathbf{a}+\mathbf{b}\| + \|\mathbf{a}-\mathbf{b}\| + |2-\alpha-\beta| + |\alpha-\beta| \leq 2, \tag{43}$$

which can also be written in operator terms as

$$\left\|\mathcal{E}_+^{\alpha,\mathbf{a}} - \mathcal{E}_+^{\beta,\mathbf{b}}\right\| + \left\|\mathcal{E}_+^{\alpha,\mathbf{a}} - \mathcal{E}_-^{\beta,\mathbf{b}}\right\| \leq 1. \tag{44}$$

This sufficient condition for joint measurability is satisfied in all cases with $\alpha = \beta = 1$ but is generally not necessary, as can be seen from the example $\mathcal{E}_+^{\alpha,\mathbf{a}} = \mathbb{1}$, $\mathcal{E}_+^{\beta,\mathbf{b}} = A(1, \mathbf{n})$ (where $\mathbf{n}$ is any unit vector).

In nontrivial cases the above sufficient joint measurability conditions can be further strengthened and simplified. For two qubit observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$, the nontriviality requirement in the sense of Example 1 amounts to the following:

$$\begin{aligned} |\alpha-\beta| &< \|\mathbf{a}-\mathbf{b}\| \quad (\text{not (a), (b)}); \\ |2-\alpha-\beta| &< \|\mathbf{a}+\mathbf{b}\| \quad (\text{not (c), (d)}). \end{aligned} \tag{45}$$

Under these nontriviality assumptions, the conditions (42) and (43) are seen to be satisfied if

$$\|\mathbf{a}+\mathbf{b}\| + \|\mathbf{a}-\mathbf{b}\| \leq 1. \tag{46}$$

The next example shows that the sufficient condition (42) is not a necessary condition.

Example 4: We consider the case where $\mathbf{a} \perp \mathbf{b}$. Furthermore, let $\mathbf{a} = \hat{\mathbf{a}}$ be a unit vector, so that $A := \mathcal{E}_+^{\alpha,\mathbf{a}} = \alpha A(1, \hat{\mathbf{a}})$ is a multiple of a projection. Note that A being an effect entails that $\alpha \leq 1$. Next we denote $B := \mathcal{E}_+^{\beta,\mathbf{b}} = A(1, \mathbf{b})$, where we assume that $b := \|\mathbf{b}\| \neq 0$.

Joint measurability of A, B is given if and only if there is an operator G which is bounded above by A, B and bounded below by $\mathbb{O}$, $A + B - \mathbb{1}$. The inequality $G \leq A$ is satisfied if and only if G is a multiple of the projection $A(1, \hat{\mathbf{a}})$, hence: $G = \gamma A(1, \hat{\mathbf{a}})$, and $\gamma \leq \alpha$. Further, γ must be chosen such that $\gamma A(1, \hat{\mathbf{a}}) \leq B$; thus:

$$1 - \gamma \geq \sqrt{\gamma^2 + b^2}.$$

This is equivalent to $\gamma \leq \gamma_0 := \frac{1}{2}(1-b^2)$. The inequality $\mathbb{1} - A - B + G \geq \mathbb{O}$ is equivalent to

$$1 - \alpha + \gamma \geq \sqrt{(\alpha-\gamma)^2 + b^2}.$$

This is solved by $\gamma \geq \alpha - \frac{1}{2}(1-b^2) = \alpha - \gamma_0$.

To summarize: the given effects A, B are jointly measurable if and only if

$$\alpha - \gamma_0 \leq \min\{\gamma_0, \alpha\}, \quad \gamma_0 \equiv \tfrac{1}{2}(1-b^2).$$

The nontriviality conditions assume here the form

$$|\alpha - \beta| = |2 - \alpha - \beta| = 1 - \alpha < \|\mathbf{a} - \mathbf{b}\| = \|\mathbf{a} + \mathbf{b}\| = \sqrt{\alpha^2 + b^2},$$

which is equivalent to $\gamma_0 < \alpha$. In this case the joint measurability condition reduces to $\alpha/2 \leq \gamma_0$.

We are now ready to show that condition (42) can be violated in nontrivial cases. In the given constellation, this inequality assumes the form

$$\sqrt{\alpha^2 + b^2} + \frac{\alpha(1-\alpha)}{\sqrt{\alpha^2 + b^2}} \leq 1.$$

For the choice $\alpha = \frac{1}{2} = 1 - b^2 = 2\gamma_0$, the left hand side becomes $2/\sqrt{3}$, which is greater than 1. However, this choice fulfills the joint measurability and nontriviality conditions.

4.3 Covariant joint observables

In what follows we will investigate implications of covariance. In this way we establish a far-reaching analogy to similar studies made on approximate joint measurements of position and momentum where covariance (under translations on phase space) was found to be paramount (cf. the review [14]).

Let us first note that there is a unitary operator U such that both $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are covariant with respect to U. Namely, fix a unit vector $\mathbf{u}$ orthogonal to both $\mathbf{a}$ and $\mathbf{b}$ and choose $U = \mathbf{u} \cdot \boldsymbol{\sigma}$.

We say that a joint observable $\mathcal{G}$ of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ is covariant with respect to U, or U-covariant, if

$$\begin{aligned} U\mathcal{G}_{++}U &= \mathcal{G}_{--}, \\ U\mathcal{G}_{+-}U &= \mathcal{G}_{-+}. \end{aligned} \tag{47}$$

Since $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are U-covariant, the two equations in (47) are equivalent and thus, already one of them implies that $\mathcal{G}$ is U-covariant.

Proposition 6: *If $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are jointly measurable, then they have a U-covariant joint observable.*

Proof: Let $\mathcal{G}$ be a joint observable of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$. Define

$$\begin{aligned} \widetilde{\mathcal{G}}_{++} &= \tfrac{1}{2}\left(\mathcal{G}_{++} + U\mathcal{G}_{--}U\right), \\ \widetilde{\mathcal{G}}_{+-} &= \tfrac{1}{2}\left(\mathcal{G}_{+-} + U\mathcal{G}_{-+}U\right), \\ \widetilde{\mathcal{G}}_{-+} &= \tfrac{1}{2}\left(\mathcal{G}_{-+} + U\mathcal{G}_{+-}U\right), \\ \widetilde{\mathcal{G}}_{--} &= \tfrac{1}{2}\left(\mathcal{G}_{--} + U\mathcal{G}_{++}U\right). \end{aligned}$$

Each operator $\widetilde{\mathcal{G}}_{\pm\pm}$ is a convex combination of two effects, hence an effect. Moreover, the sum of these effects is $\mathbb{1}$ and thus, $\widetilde{\mathcal{G}}$ is an observable.

We have

$$\begin{aligned}\widetilde{\mathcal{G}}_{++}+\widetilde{\mathcal{G}}_{+-} &= \mathcal{E}^{1,\mathbf{a}}_{+},\\ \widetilde{\mathcal{G}}_{++}+\widetilde{\mathcal{G}}_{-+} &= \mathcal{E}^{1,\mathbf{b}}_{+},\end{aligned}$$

showing that $\widetilde{\mathcal{G}}$ is a joint observable of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$. Using the fact that $U^2=\mathbb{1}$ we immediately see that $U\widetilde{\mathcal{G}}_{++}U=\widetilde{\mathcal{G}}_{--}$, meaning that $\widetilde{\mathcal{G}}$ is U-covariant. □

We proceed by characterizing all U-covariant joint observables of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$. Denoting $\mathcal{G}_{++}=\frac{1}{2}\left(\gamma\mathbb{1}+\mathbf{g}\cdot\boldsymbol{\sigma}\right)$ the covariance condition (47) can be written in the form

$$\mathbf{g}-(\mathbf{u}\cdot\mathbf{g})\mathbf{u}=\tfrac{1}{2}(\mathbf{a}+\mathbf{b}), \tag{48}$$

which means that $\mathbf{g}=\frac{1}{2}(\mathbf{a}+\mathbf{b})+p\mathbf{u}$ for some $p\in\mathbb{R}$. The joint measurability condition (6) reduces to the requirement that

$$\sqrt{\tfrac{1}{4}\left\|\mathbf{a}+\mathbf{b}\right\|^2+p^2}\leq\gamma\leq 1-\sqrt{\tfrac{1}{4}\left\|\mathbf{a}-\mathbf{b}\right\|^2+p^2}. \tag{49}$$

We conclude that U-covariant joint observables of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are characterized by the pairs (γ,p) satisfying (49). The covariant joint observable $\mathcal{G}$ corresponding to (γ,p) is

$$\begin{aligned}\mathcal{G}_{++} &= \frac{\gamma}{2}\mathbb{1}+\frac{1}{4}(\mathbf{a}+\mathbf{b})\cdot\boldsymbol{\sigma}+\frac{p}{2}\mathbf{u}\cdot\boldsymbol{\sigma},\\ \mathcal{G}_{+-} &= \frac{1-\gamma}{2}\mathbb{1}+\frac{1}{4}(\mathbf{a}-\mathbf{b})\cdot\boldsymbol{\sigma}-\frac{p}{2}\mathbf{u}\cdot\boldsymbol{\sigma},\\ \mathcal{G}_{-+} &= \frac{1-\gamma}{2}\mathbb{1}-\frac{1}{4}(\mathbf{a}-\mathbf{b})\cdot\boldsymbol{\sigma}-\frac{p}{2}\mathbf{u}\cdot\boldsymbol{\sigma},\\ \mathcal{G}_{--} &= \frac{\gamma}{2}\mathbb{1}-\frac{1}{4}(\mathbf{a}+\mathbf{b})\cdot\boldsymbol{\sigma}+\frac{p}{2}\mathbf{u}\cdot\boldsymbol{\sigma}.\end{aligned}$$

If a pair (γ,p) satisfies condition (49), then so does $(\gamma,0)$. Hence, $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ have a U-covariant joint observable if and only if there is a γ such that

$$\tfrac{1}{2}\left\|\mathbf{a}+\mathbf{b}\right\|\leq\gamma\leq 1-\tfrac{1}{2}\left\|\mathbf{a}-\mathbf{b}\right\|, \tag{50}$$

or equivalently, if and only if inequality (39) holds. This together with Proposition 6 gives the result cited in Proposition 4. Inequality (50) implies that

$$\tfrac{1}{2}\left\|\mathbf{a}+\mathbf{b}\right\|\leq\tfrac{1}{2}(1+\mathbf{a}\cdot\mathbf{b})\leq 1-\tfrac{1}{2}\left\|\mathbf{a}-\mathbf{b}\right\|. \tag{51}$$

Thus, if $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are jointly measurable, then they have a joint observable $\mathcal{G}^0$ corresponding to the choice $\gamma=\gamma_0:=\frac{1}{2}(1+\mathbf{a}\cdot\mathbf{b})$ and $p=0$. The effects of $\mathcal{G}^0$ can be written in the form

$$\mathcal{G}^0_{ij}=\tfrac{1}{2}\left(\mathcal{E}^{1,\mathbf{a}}_i\mathcal{E}^{1,\mathbf{b}}_j+\mathcal{E}^{1,\mathbf{b}}_j\mathcal{E}^{1,\mathbf{a}}_i\right),\quad i,j=\pm. \tag{52}$$

If we have the limiting case of condition (50), i.e.

$$\tfrac{1}{2}\left\|\mathbf{a}+\mathbf{b}\right\|=1-\tfrac{1}{2}\left\|\mathbf{a}-\mathbf{b}\right\|, \tag{53}$$

then $(\gamma_0, 0)$ is the only possible pair and hence, in this case $\mathcal{G}^0$ is the unique U-covariant joint observable of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$. In all other situations of covariant joint measurements except this limiting case, there is a continuum of possible pairs (γ, p). The joint observable $\mathcal{G}$ corresponding to (γ, p) is informationally complete if and only if $p \neq 0$; this follows directly from [16, Theorem 4.7].

Finally, we note that the covariance of $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ does not imply that they have only covariant joint observables. To give an example, assume that the vectors $\mathbf{a}$ and $\mathbf{b}$ satisfy $\mathbf{a} \cdot \mathbf{b} \geq 0$ and $\|\mathbf{a} + \mathbf{b}\| < 1$, so that $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are jointly measurable. Fix a number t such that $0 < t \leq \|\mathbf{a} + \mathbf{b}\|^{-1} - 1$, and define

$$\begin{aligned}
\mathcal{G}_{++} &= \frac{1}{4}\mathbb{1} + \frac{1}{4}(1+t)(\mathbf{a}+\mathbf{b}) \cdot \boldsymbol{\sigma}, \\
\mathcal{G}_{+-} &= \frac{1}{4}\mathbb{1} + \frac{1}{4}((1-t)\mathbf{a} - (1+t)\mathbf{b}) \cdot \boldsymbol{\sigma}, \\
\mathcal{G}_{-+} &= \frac{1}{4}\mathbb{1} + \frac{1}{4}((1-t)\mathbf{b} - (1+t)\mathbf{a}) \cdot \boldsymbol{\sigma} \\
\mathcal{G}_{--} &= \frac{1}{4}\mathbb{1} - \frac{1}{4}(1-t)(\mathbf{a}+\mathbf{b}) \cdot \boldsymbol{\sigma}.
\end{aligned}$$

Then $\mathcal{G}$ is a joint observable for $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ but it is not covariant. Indeed, the above condition on t guarantees that inequalities (6) are satisfied. The eigenvalues of $\mathcal{G}_{++}$ and $\mathcal{G}_{--}$ are different and thus, the covariance condition (47) cannot be satisfied with any unitary operator U.

4.4 Joint measurability vs. sharpness

It is instructive to write down the joint measurement condition for two covariant observables $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$, assuming that the vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal. The inequality (39) takes now the form

$$\mathfrak{U}(\mathcal{E}^{1,\mathbf{a}}) + \mathfrak{U}(\mathcal{E}^{1,\mathbf{b}}) \geq 1, \tag{54}$$

showing that the joint measurability is achieved exactly when the observables are made unsharp enough.

In the general case, we can transform (39) by repeated squaring into the equivalent inequality

$$\|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 \leq 1 + (\mathbf{a} \cdot \mathbf{b})^2, \tag{55}$$

which can be written in the form

$$\|\mathbf{a} \times \mathbf{b}\|^2 \leq (1 - \|\mathbf{a}\|^2)(1 - \|\mathbf{b}\|^2). \tag{56}$$

The term on the left hand side is equal to $4\left\|[\mathcal{E}_+^{\alpha,\mathbf{a}}, \mathcal{E}_+^{\beta,\mathbf{b}}]\right\|^2$. The term on the right hand side turns out to give a bound for the degrees of sharpness of $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$. Considering the formulas (4) and (25) we obtain the following.

Proposition 7: *If two qubit observables $\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{\beta,\mathbf{b}}$ are jointly measurable, the degrees of their unsharpness satisfy the inequality*

$$\mathfrak{U}(\mathcal{E}^{\alpha,\mathbf{a}})\,\mathfrak{U}(\mathcal{E}^{\beta,\mathbf{b}}) \;\geq\; \mathfrak{U}(\mathcal{E}^{1,\mathbf{a}})\,\mathfrak{U}(\mathcal{E}^{1,\mathbf{b}}) \;\geq\; 4\left\|[\mathcal{E}_+^{\alpha,\mathbf{a}}, \mathcal{E}_+^{\beta,\mathbf{b}}]\right\|^2. \tag{57}$$

If $\alpha = \beta = 1$, this inequality is in fact equivalent to the joint measurability condition.

This shows that the intrinsic sharpness of two jointly measurable simple qubit observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ is limited by the noncommutativity of the generating effects.

5 Approximate joint measurement for two sharp qubit observables

Two sharp observables $\mathcal{E}^{1,\mathbf{n}}$ and $\mathcal{E}^{1,\mathbf{m}}$ are jointly measurable exactly when they commute, and this happens if and only if $\mathbf{n} = \pm\mathbf{m}$. In this section we consider the case $\mathbf{n} \neq \pm\mathbf{m}$, so that only approximate joint measurements are possible. The idea is to choose a jointly measurable pair $(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{\beta,\mathbf{b}})$ to approximate the sharp pair $(\mathcal{E}^{1,\mathbf{n}}, \mathcal{E}^{1,\mathbf{m}})$. To be specific, and without loss of generality, we assume $\cos\theta := \mathbf{n} \cdot \mathbf{m} > 0$.

We call a point $(\mathfrak{D}_1, \mathfrak{D}_2) \in [0,1] \times [0,1]$ *admissible* if $\mathfrak{D}_1 = \mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}})$ and $\mathfrak{D}_2 = \mathfrak{D}(\mathcal{E}^{\beta,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})$ for some jointly measurable observables $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$. Not all points in the square $[0,1] \times [0,1]$ are admissible; for instance the point $(0,0)$ is not an admissible point since this would mean that $\mathcal{E}^{\alpha,\mathbf{a}} = \mathcal{E}^{1,\mathbf{n}}$ and $\mathcal{E}^{\beta,\mathbf{b}} = \mathcal{E}^{1,\mathbf{m}}$. We show in the following that there are also other points which are not admissible. The set of admissible points gives us a characterization on the quality of possible approximate joint measurements.

The search for admissible points $(\mathfrak{D}_1, \mathfrak{D}_2)$ is narrowed down by the following simple observation:

Example 5: Let $\alpha \in [0,2]$. Then $\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{0}}, \mathcal{E}^{1,\mathbf{n}}) = \frac{1}{2}\max\{\alpha, 2-\alpha\}$ and therefore

$$\left\{\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{0}}, \mathcal{E}^{1,\mathbf{n}}) \,:\, \alpha \in [0,2]\right\} = [\tfrac{1}{2}, 1]. \tag{58}$$

Thus, approximations by means of trivial observables will never give distances below $\frac{1}{2}$. Furthermore, since $\mathcal{E}^{\alpha,\mathbf{0}}$ is jointly measurable with any observable $\mathcal{E}^{\beta,\mathbf{b}}$, and since $\mathfrak{D}(\mathcal{E}^{\beta,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})$ can assume any value in $[0,1]$, it follows that all points in the set $[0,1] \times [0,1] \setminus [0,\frac{1}{2}] \times [0,\frac{1}{2}]$ are trivially admissible. We will therefore concentrate on admissible points $(\mathfrak{D}_1, \mathfrak{D}_2)$ in the region $[0,\frac{1}{2}] \times [0,\frac{1}{2}]$.

The next two results are not complicated but require some preparation and will be proven in the Appendix.

Proposition 8: *Any admissible point $(\mathfrak{D}_1, \mathfrak{D}_2) \in [0,\frac{1}{2}] \times [0,\frac{1}{2}]$ has a realization of the type $\mathfrak{D}_1 = \mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}})$, $\mathfrak{D}_2 = \mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})$, where $\mathbf{a}$ and $\mathbf{b}$ are in the plane spanned by $\mathbf{n}$ and $\mathbf{m}$.*

Proposition 9: *The set of admissible points is a closed convex set which is reflection symmetric with respect to the axis $\mathfrak{D}_1 = \mathfrak{D}_2$; that is, with every admissible point $(\mathfrak{D}_1, \mathfrak{D}_2)$ the point $(\mathfrak{D}_2, \mathfrak{D}_1)$ is also admissible. Thus the segment of the boundary curve defined as the graph of the function*

$$\mathfrak{D}_1 \mapsto \inf\{\mathfrak{D}_2 \,:\, (\mathfrak{D}_1, \mathfrak{D}_2) \text{ is admissible}\} \tag{59}$$

is convex, symmetric and belongs to the set of admissible points.

Example 6 : If $\mathfrak{D}_1 = \mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}) = 0$ (i.e. $\mathbf{a} = \mathbf{n}$), then the joint measurability requirement implies that $\mathbf{a}||\mathbf{b}$ and thus,

$$\mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}}) = \tfrac{1}{2}\,\|\mathbf{b} - \mathbf{m}\| \geq \tfrac{1}{2}\sqrt{1 - (\mathbf{n}\cdot\mathbf{m})^2} = \tfrac{1}{2}\sin\theta.$$

The lower bound is attained when $\mathbf{b} = \cos\theta\ \mathbf{n} = (\mathbf{n}\cdot\mathbf{m})\mathbf{n}$. We conclude that $\left(0, \frac{1}{2}\sin\theta\right)$ and $\left(\frac{1}{2}\sin\theta, 0\right)$ are points in the boundary of the admissible region.

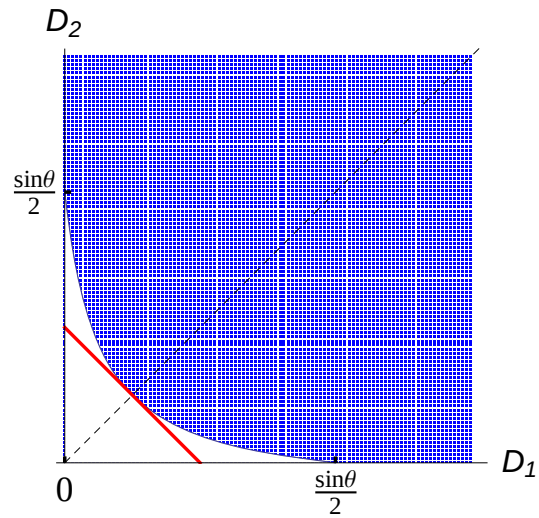

Fig. 1. The admissible region (dotted area) and the line $\mathfrak{D}_1 + \mathfrak{D}_2 = 2\mathfrak{D}_0$ (thick line). The dashed line is the symmetry axis $\mathfrak{D}_1 = \mathfrak{D}_2$.

We next determine the boundary point with $\mathfrak{D}_1 = \mathfrak{D}_2 =: \mathfrak{D}_0$. Due to the convexity of the admissible region and its reflection symmetry with respect to the line $\mathfrak{D}_1 = \mathfrak{D}_2$, it follows immediately that the admissible region is bounded below tightly by the straight line $\mathfrak{D}_1 + \mathfrak{D}_2 = 2\mathfrak{D}_0$. This situation is sketched in Fig. 1. Determination of the value of $\mathfrak{D}_0$ yields the following result.

Proposition 10: *Any admissible point* $(\mathfrak{D}_1, \mathfrak{D}_2) = \left(\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}), \mathfrak{D}(\mathcal{E}^{\beta,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})\right)$ *satisfies the inaccuracy trade-off relation*

$$\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}) + \mathfrak{D}(\mathcal{E}^{\beta,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}}) \geq 2\mathfrak{D}_0, \tag{60}$$

where

$$2\mathfrak{D}_0 = \tfrac{1}{\sqrt{2}}\left[\tfrac{1}{2}\,\|\mathbf{n} + \mathbf{m}\| + \tfrac{1}{2}\,\|\mathbf{n} - \mathbf{m}\| - 1\right] = \tfrac{1}{\sqrt{2}}\left(\cos\tfrac{\theta}{2} + \sin\tfrac{\theta}{2} - 1\right). \tag{61}$$

The point $(\mathfrak{D}_0, \mathfrak{D}_0)$ *is admissible.*

Proof: Consider the set of all jointly measurable covariant observables $\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{b}}$ such that $\mathbf{a}, \mathbf{b}$ have equal fixed distance from $\mathbf{n}, \mathbf{m}$, respectively: $\|\mathbf{a} - \mathbf{n}\| = \|\mathbf{b} - \mathbf{m}\| \equiv d$ (so that

$\mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}) = \mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}}) = d/2$). If $(\mathbf{a}, \mathbf{b})$ is not symmetric under reflection with respect to the line parallel to $\mathbf{n} + \mathbf{m}$, denote by $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ the mirror images of $\mathbf{b}$ and $\mathbf{a}$, respectively. Then, if $\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{b}}$ are jointly measurable, so are $\mathcal{E}^{1,\bar{\mathbf{a}}}, \mathcal{E}^{1,\bar{\mathbf{b}}}$ as the condition (39) is invariant under reflections. Due to Proposition 2, the observables $\frac{1}{2}\mathcal{E}^{1,\mathbf{a}} + \frac{1}{2}\mathcal{E}^{1,\bar{\mathbf{a}}} = \mathcal{E}^{1,\frac{1}{2}(\mathbf{a}+\bar{\mathbf{a}})}$ and $\frac{1}{2}\mathcal{E}^{1,\mathbf{b}} + \frac{1}{2}\mathcal{E}^{1,\bar{\mathbf{b}}} = \mathcal{E}^{1,\frac{1}{2}(\mathbf{b}+\bar{\mathbf{b}})}$ are jointly measurable. It is clear from their definitions that the vectors $\frac{1}{2}(\mathbf{a}+\bar{\mathbf{a}})$ and $\frac{1}{2}(\mathbf{b}+\bar{\mathbf{b}})$ are mirror images of each other. As $\mathbf{a}, \mathbf{b}$ have equal distance d from $\mathbf{n}, \mathbf{m}$, respectively, this means that $\mathbf{a}$ and $\bar{\mathbf{a}}$ have equal distance d from $\mathbf{n}$. It follows that the distance from $\mathbf{n}$ to $\frac{1}{2}(\mathbf{a}+\bar{\mathbf{a}})$ is less than d (or d if $\mathbf{a} = \bar{\mathbf{a}}$). We conclude that if $\mathbf{a}, \mathbf{b}$ are not mirror images of each other, there is a pair of jointly measurable covariant observables with smaller (equal) distances from $\mathcal{E}^{1,\mathbf{n}}, \mathcal{E}^{1,\mathbf{m}}$ and mirror symmetric vectors. This shows that the minimal equal distance approximations of $\mathcal{E}^{1,\mathbf{n}}, \mathcal{E}^{1,\mathbf{m}}$ by means of jointly measurable observables occur among the covariant pairs with $\mathbf{a}, \mathbf{b}$ mirror symmetric with respect to $\mathbf{n}+\mathbf{m}$.

If coordinates are chosen such that $\mathbf{n} = (\sin\frac{\theta}{2}, \cos\frac{\theta}{2})$, $\mathbf{m} = (-\sin\frac{\theta}{2}, \cos\frac{\theta}{2})$, then let a symmetric pair $\mathbf{a}, \mathbf{b}$ be given by $\mathbf{a} = (u, v)$ and $\mathbf{b} = (-u, v)$, with $u, v > 0$. For such pairs, the joint measurability condition for $\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{b}}$ assumes the form $u + v \leq 1$. It follows that the shortest (equal) distances d of $\mathbf{a}, \mathbf{b}$ from $\mathbf{n}, \mathbf{m}$ are assumed when $u + v = 1$ and $\mathbf{n} - \mathbf{a}$ is perpendicular to the line $u + v = 1$. But this distance d is equal to the distance of the lines $u + v = 1$ and $u + v = \cos(\frac{\theta}{2}) + \sin(\frac{\theta}{2})$, hence

$$d = \frac{1}{\sqrt{2}}\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2} - 1\right).$$

□

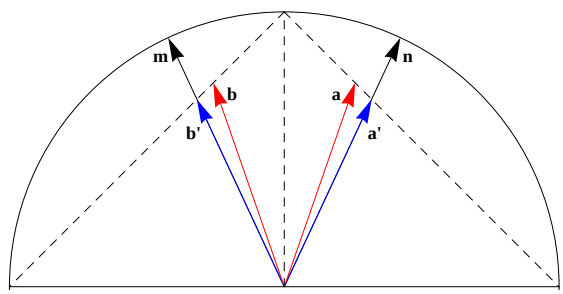

Fig. 2. The vectors corresponding to the optimal approximations $\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{b}}$ and of optimal coarse-grainings $\mathcal{E}^{1,\mathbf{a}'}, \mathcal{E}^{1,\mathbf{b}'}$.

The result of Proposition 10 shows in which way the quality of the approximations is limited by the separation of the sharp observables to be approximated in a simultaneous measurement. This relation becomes perhaps even more transparent when we write the number $\mathfrak{D}_0$ in the form

$$\mathfrak{D}_0 = \frac{1}{2\sqrt{2}}\left[\mathfrak{D}(\mathcal{E}^{1,\mathbf{n}}, \mathcal{E}^{1,\mathbf{m}}) + \mathfrak{D}(\mathcal{E}^{1,\mathbf{n}}, \mathcal{E}^{1,-\mathbf{m}}) - 1\right]. \tag{62}$$

The appearance of $\mathfrak{D}(\mathcal{E}^{1,\mathbf{n}}, \mathcal{E}^{1,-\mathbf{m}})$ in (62) is explained by the fact that the joint measurability criterion is blind to the labeling of outcomes.

Note that $\cos\frac{\theta}{2}+\sin\frac{\theta}{2}=(1+\sin\theta)^{1/2}$ and $\sin\theta=\|\mathbf{n}\times\mathbf{m}\|=2\left\|[\mathcal{E}^{1,\mathbf{n}},\mathcal{E}^{1,\mathbf{m}}]\right\|$. Thus $\mathfrak{D}_0$ is an increasing function of the degree of noncommutativity of the sharp observables to be estimated.

The approximations $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ leading to the boundary point $(\mathfrak{D}_0,\mathfrak{D}_0)$ are generally not among the coarse-grainings of $\mathcal{E}^{1,\mathbf{n}}$ and $\mathcal{E}^{1,\mathbf{m}}$ (in the sense of Section 2.3). Indeed, let us denote by $\mathfrak{D}_0^c$ the smallest number achieved under the assumptions that $\mathfrak{D}_0^c=\mathfrak{D}(\mathcal{E}^{1,\mathbf{a}'},\mathcal{E}^{1,\mathbf{n}})=\mathfrak{D}(\mathcal{E}^{1,\mathbf{b}'},\mathcal{E}^{1,\mathbf{m}})$ and that $\mathcal{E}^{1,\mathbf{a}'},\mathcal{E}^{1,\mathbf{b}'}$ are jointly measurable and coarse-grainings of $\mathcal{E}^{1,\mathbf{n}},\mathcal{E}^{1,\mathbf{m}}$, respectively. If the vectors $\mathbf{n}$ and $\mathbf{m}$ are orthogonal, then $\mathfrak{D}_0^c=\mathfrak{D}_0$. However, if $0<\theta<\frac{\pi}{2}$, then

$$\mathfrak{D}_0^c=\tfrac{1}{2}\left(1-\frac{\sqrt{1-\sin\theta}}{\cos\theta}\right)>\mathfrak{D}_0. \tag{63}$$

The vectors $\mathbf{a},\mathbf{b}$ and $\mathbf{a}',\mathbf{b}'$ are illustrated in Fig. 2. We conclude that to attain the best jointly measurable approximations of two sharp qubit observables, we are forced to seek approximating observables beyond their coarse-grainings.

Finally, we note that it would be interesting to determine the full convex boundary curve of the region of admissible points $(\mathfrak{D}_1,\mathfrak{D}_2)$. Some numerically calculated boundary curves are drawn in [24], but their analytic form is not yet known.

6 Conclusion and outlook

In this paper we have quantified the necessary inaccuracies in approximating noncommuting sharp qubit observables by means of a pair of jointly measurable pair of observables (Eq. (60)). We also exhibited the necessary unsharpness that observables $\mathcal{E}^{\alpha,\mathbf{a}},\mathcal{E}^{\beta,\mathbf{b}}$ must have in order to be jointly measurable (Eq. (57)). If a sharp observable $\mathcal{E}^{1,\hat{\mathbf{a}}}$ is approximated by one of its coarse-grainings $\mathcal{E}^{\alpha,\mathbf{a}}$, the distance is related to the sharpness of $\mathcal{E}^{\alpha,\mathbf{a}}$ via the relation (29).

Trough the case study of qubit observables we have demonstrated the conceptual difference of measurement inaccuracy and intrinsic unsharpness. This sheds some new light on the joint measurement problem raised by Uffink in [25], so we shortly recall his argumentation. Uffink analyzed a definition of "non-ideal" or "unsharp" joint measurement of two noncommuting observables that had previously been sketched out more or less informally by various authors. This definition captures the idea that smearings of two noncommuting sharp observables may have a joint observable. As it was formulated, this definition allowed any smeared or coarse-grained version of an observable to be an approximation of that observable, without further stipulations on the quality of the approximation. This entails that even trivial observables (which are always among the coarse-grainings of any observable) can be taken to represent a sort of non-ideal measurement of a given observable.

Uffink presented an example that makes this definition look absurdly comprehensive and indeed counter-intuitive: he considered two pairs of observables, (σ_x,Q) and (σ_z,P) and took σ_x as a coarse-graining of the first pair and P as a coarse-graining of the second. Then (σ_x,P) is a joint observable for these two, and according to the letter of the definition, it would have to be considered as representing a non-ideal or unsharp joint measurement of the original pairs.

Now, Uffink argued that while the final joint observable had σ_x and P as coarse-grainings (namely, marginals), the original observables were in no way coarse-grainings of it. Hence there was no plausible sense in which (σ_x,P) could be regarded as representing a non-ideal

joint measurement of the original pairs of observables. He thus pointed out rightly that a universal definition or criterion of approximate joint measurability was missing. But then he jumped to the conclusion that POMs do not contribute to solving the joint measurement problem.

We think that the present paper and many preceding it demonstrate that POMs do provide an appropriate language to clarify the definition and quantification of approximate measurements, and to determine any limitations to the accuracy of joint approximations of noncommuting pairs of observables. It is obvious that any measurement can be considered as an "approximate" joint measurement of an arbitrary collection of observables. Even doing nothing and randomly picking outcomes constitutes a trivial "non-ideal" joint measurement of any given set of observables. There is no problem in allowing a definition of non-ideal or approximate joint measurements to include trivial cases; what makes any such definition useful is whether it allows one to give quantifications of how well each of the observables in question is being approximated by a given scheme. As we have shown in this paper and its companion [14], such quantifications can indeed be formulated and yield a nontrivial notion of approximate measurement, leading to the conclusion that there are universal limitations to the accuracies with which noncommuting pairs of observables can be approximately measured together.

If the quality of the approximation is to be optimized, the approximating observables being measured jointly must be unsharp; and the required degree of unsharpness is linked with the quality of the approximations specified. Using the definition of approximation introduced here, and keeping in mind the conceptual difference between the relation of approximation and the property of intrinsic unsharpness, it is clear that the above "absurd" example considered by Uffink is simply not based on good approximations and would therefore not be regarded as a useful joint measurement.

The quantifications of inaccuracy and intrinsic unsharpness presented here for the case of qubit observables complements analogous investigations carried out in the case of continuous observables in [14, 15, 26]. A unified approach and associated trade-off relations for the approximate joint measurements of general pairs of noncommuting quantities is still outstanding.

Appendix: Proofs of Propositions 8 and 9

(a) If $(\mathfrak{D}_1, \mathfrak{D}_2)$ is an admissible point, then also $(\mathfrak{D}_2, \mathfrak{D}_1)$ is an admissible point.
Proof: If $(\alpha, \mathbf{a})$ and $(\beta, \mathbf{b})$ realize the distances $\mathfrak{D}_1$ and $\mathfrak{D}_2$, respectively, then choose $(\alpha', \mathbf{a}')$ and $(\beta', \mathbf{b}')$ as follows: $\alpha' = \beta$, $\mathbf{a}'$ has the length of $\mathbf{b}$ and its angle relative to $\mathbf{n}$ is equal to the angle between $\mathbf{b}$ and $\mathbf{m}$; similarly, $\beta' = \alpha$, $\mathbf{b}'$ has the length of $\mathbf{a}$ and its angle relative to $\mathbf{m}$ is the same as the angle between $\mathbf{a}$ and $\mathbf{n}$. This ensures that $(\mathfrak{D}_1', \mathfrak{D}_2') = (\mathfrak{D}_2, \mathfrak{D}_1)$. $\square$

(b) Assume that $(\mathfrak{D}_1, \mathfrak{D}_2) = \big(\mathfrak{D}(\mathcal{E}^{\alpha,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}), \mathfrak{D}(\mathcal{E}^{\beta,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})\big)$ is an admissible point. As shown in Proposition 5, the joint measurability of $\mathcal{E}^{\alpha,\mathbf{a}}$ and $\mathcal{E}^{\beta,\mathbf{b}}$ implies that $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,\mathbf{b}}$ are jointly measurable. Define $\mathbf{a}_0$ and $\mathbf{b}_0$ to be the projections of the vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, onto the plane spanned by $\mathbf{n}$ and $\mathbf{m}$. Then

$$\|\mathbf{a} + \mathbf{b}\| \geq \|\mathbf{a}_0 + \mathbf{b}_0\|\,, \qquad \|\mathbf{a} - \mathbf{b}\| \geq \|\mathbf{a}_0 - \mathbf{b}_0\|\,,$$

and hence, $\mathcal{E}^{1,\mathbf{a}_0}$ and $\mathcal{E}^{1,\mathbf{b}_0}$ are jointly measurable. Using (26) one finds that

$$\begin{aligned} \mathfrak{D}(\mathcal{E}^{1,\mathbf{a}_0}, \mathcal{E}^{1,\mathbf{n}}) &\leq \mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}) \leq \mathfrak{D}_1, \\ \mathfrak{D}(\mathcal{E}^{1,\mathbf{b}_0}, \mathcal{E}^{1,\mathbf{m}}) &\leq \mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}}) \leq \mathfrak{D}_2. \end{aligned} \tag{64}$$

We conclude that the best approximations are to be found from the subset of covariant qubit observables, with vectors $\mathbf{a}$ and $\mathbf{b}$ in the plane spanned by $\mathbf{n}$ and $\mathbf{m}$.

(c) If $(\mathfrak{D}_1, \mathfrak{D}_2)$ is an admissible point, then also $(\mathfrak{D}_1', \mathfrak{D}_2')$ is an admissible point whenever $\mathfrak{D}_i \leq \mathfrak{D}_i' \leq \frac{1}{2}$.
Proof: In view of (b) it is sufficient to show the result for admissible points which have realizations $\big(\mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}), \mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})\big)$. Thus let $\mathcal{E}^{1,\mathbf{a}}$, $\mathcal{E}^{1,\mathbf{b}}$ be two jointly measurable observables. Using Proposition 4, we note that also $\mathcal{E}^{1,\mathbf{a}}$ and $\mathcal{E}^{1,r\mathbf{b}}$ are jointly measurable for any $0 \leq r \leq 1$. Since the function

$$r \mapsto \mathfrak{D}(\mathcal{E}^{1,r\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}}) = \tfrac{1}{2} \|\mathbf{m} - r\mathbf{b}\|$$

is continuous, it takes all values between $\mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})$ and $\frac{1}{2}$. We can similarly realize all values between $\mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}})$ and $\frac{1}{2}$. □

(d) Observations (b) and (c) taken together entail Proposition 8. □

(e) The admissible region is a convex set.
Proof: Let $(\mathfrak{D}_1, \mathfrak{D}_2)$ and $(\mathfrak{D}_1', \mathfrak{D}_2')$ be realized by $(\alpha, \mathbf{a}), (\beta, \mathbf{b})$ and $(\alpha', \mathbf{a}'), (\beta', \mathbf{b}')$ respectively. Let $t \in [0,1]$. Then for $(\alpha_t, \mathbf{a}_t) := (t\alpha + (1-t)\alpha', t\mathbf{a} + (1-t)\mathbf{a}')$ and $(\beta_t, \mathbf{b}_t) := (t\beta + (1-t)\beta', t\mathbf{b} + (1-t)\mathbf{b}')$, we obtain associated distances $\mathfrak{D}_{1,t}$ and $\mathfrak{D}_{2,t}$ which satisfy

$$\mathfrak{D}_{k,t} \leq t\mathfrak{D}_k + (1-t)\mathfrak{D}_k', \quad k = 1, 2.$$

This together with (c) proves the claim. □

(f) The set of admissible points is closed.
Proof: The mapping

$$(\mathbf{a}, \mathbf{b}) \mapsto (\mathfrak{D}(\mathcal{E}^{1,\mathbf{a}}, \mathcal{E}^{1,\mathbf{n}}), \mathfrak{D}(\mathcal{E}^{1,\mathbf{b}}, \mathcal{E}^{1,\mathbf{m}})) = \tfrac{1}{2}(\|\mathbf{a} - \mathbf{n}\|, \|\mathbf{b} - \mathbf{m}\|) \tag{65}$$

from $\mathbb{R}^3 \times \mathbb{R}^3$ to $\mathbb{R} \times \mathbb{R}$ is continuous. The set of admissible points is the image of the compact set

$$\{(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|\mathbf{a}\| \leq 1, \|\mathbf{b}\| \leq 1, \|\mathbf{a} - \mathbf{b}\| + \|\mathbf{a} + \mathbf{b}\| \leq 2\}, \tag{66}$$

hence it is itself closed and contains its boundary. This and (b) shows that for given $\mathfrak{D}_1 \in [0, \frac{1}{2}]$, there is a minimal number $\mathfrak{D}_2^{\min}(\mathfrak{D}_1)$ such that all $(\mathfrak{D}_1, \mathfrak{D}_2)$ with $\mathfrak{D}_2^{\min}(\mathfrak{D}_1) \leq \mathfrak{D}_2 \leq \frac{1}{2}$ are admissible pairs while pairs with $\mathfrak{D}_2 < \mathfrak{D}_2^{\min}(\mathfrak{D}_1)$ are not admissible. □

(g) Since the admissible region is a convex set, the function $\mathfrak{D}_1 \mapsto \mathfrak{D}_2^{\min}(\mathfrak{D}_1)$ is convex and therefore continuous. Due to (a), the curve is symmetric under reflection with respect to the line $\mathfrak{D}_1 = \mathfrak{D}_2$. We conclude that this function gives the lower boundary curve of the set of

admissible points, and that the points on this curve are admissible. Together with (e) and (f), this completes the proof of Proposition 9.

Acknowledgement This work was initiated during T.H.'s visit at Perimeter Institute. Hospitality and support to both authors during this visit and to P.B. during the completion phase are gratefully acknowledged. T.H. acknowledges the support of the European Union project CONQUEST during the final phase of this work.

References

1. P. Busch, M. Grabowski, and P.J. Lahti. *Operational Quantum Physics*. Springer-Verlag, Berlin, 1997. second corrected printing.
2. W.M. de Muynck. *Foundations of Quantum Mechanics, an Empiricist Approach*. Kluwer Academic Publishers, Dordrecht, 2002.
3. E. Andersson, S.M. Barnett, and A. Aspect. Joint measurements of spin, operational locality, and uncertainty. *Phys. Rev. A*, 72:042104/1–8, 2005.
4. W. Son, E. Andersson, S.M. Barnett, and M.S. Kim. Joint measurements and Bell inequalities. *Phys. Rev. A*, 72:052116/1–7, 2005.
5. S.K. Choudhary, G. Kar, S. Kunkri, and R. Rahaman. Joint measurement of spin and Tsirelson's bound. *quant-ph*/0608100, 2006.
6. B. Janssens and L. Bouten. Optimal pointers for joint measurement of σ_x and σ_z via homodyne detection. *J. Phys. A: Math. Gen.*, 39:2773–2790, 2006.
7. G.M. D'Ariano, C. Macchiavello, and M.F. Sacchi. Quantum cloning optimal for joint measurements. *quant-ph*/0009080, 2000.
8. G. M. D'Ariano, C. Macchiavello, and M.F. Sacchi. Joint measurements via quantum cloning. *J. Optics B: Quantum and Semiclassical Optics*, 3:44–50, 2001.
9. T. Brougham, E. Andersson, and S.M. Barnett. Cloning and joint measurements of incompatible components of spin. *Phys. Rev. A*, 73:062319/1–7, 2006.
10. A. Ferraro and M.G.A. Paris. Joint measurements on qubits and cloning of observables. *Open Sys. & Information Dyn.*, 14:149–157, 2007.
11. R.F. Werner. Quantum information theory – an invitation. In *Quantum Information: an Introduction to Basic Theoretical Concepts and Experiments*, chapter 2, pages 14–57. Springer-Verlag, 2001.
12. W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Phys.*, 43:172–198, 1927.
13. R.F. Werner. The uncertainty relation for joint measurement of position and momentum. *Qu. Inf. Comp.*, 4:546–562, 2004.
14. P. Busch, T. Heinonen, and P.J. Lahti. Heisenberg's uncertainty principle. *Physics Reports*, 452:155–176, 2007.
15. P. Busch and D.B. Pearson. Universal joint-measurement uncertainty relation for error bars. *J. Math. Phys.*, 48:082103, 2007.
16. P. Busch. Unsharp reality and joint measurements for spin observables. *Phys. Rev. D*, 33:2253–2261, 1986.
17. P. Busch. Some realizable joint measurements of complementary observables. *Found. Phys.*, 17:905–937, 1987.
18. P. Busch. On the sharpness of an effect. *math-ph*/0706.3532v1, 2007.
19. P. Lahti. Coexistence and joint measurability in quantum mechanics. *Int. J. Theor. Phys.*, 42:893–906, 2003.
20. K. Kraus. *States, Effects, and Operations*. Springer-Verlag, Berlin, 1983.
21. A. Chefles. Quantum state discrimination. *Contemp. Phys.*, 41:401–424, 2000.
22. P. Stano, D. Reitzner and T. Heinosaari. Coexistence of qubit effects. *quant-ph*/0802.4248, 2008.

23. P. Busch and H.-J. Schmidt. Coxistence of qubit effects. *quant-ph*/0802.4167, 2008.
24. T. Heinosaari, D. Reitzner and P. Stano. Approximate joint measurability of spin along two directions. *quant-ph*/0801.2712, 2008.
25. J. Uffink. The joint measurement problem. *Int. J. Theor. Phys.*, 33:199–212, 1994.
26. C. Carmeli, T. Heinonen, and A. Toigo. Intrinsic unsharpness and approximate repeatability of quantum measurements. *J. Phys. A: Math. Theor.*, 40:1303–1323, 2007.