

## The Inaccuracy Principle

Hans Martens<sup>1</sup> and Willem M. de Muynck<sup>1</sup>

Received May 30, 1989; revised July 21, 1989

---

*The problem of joint measurement of incompatible observables is investigated. Measurements are represented by positive operator-valued measures. A quantitative notion of inaccuracy is defined. It is shown that within this framework joint inaccurate measurements are possible for arbitrary maximal projection-valued measures on finite-dimensional spaces. The accuracy of such measurements is limited, as is shown by an inaccuracy inequality we derive. This new type of uncertainty relation can be unambiguously interpreted as referring to measurement precision rather than preparative quality. Several recent experiments are seen to be realizations of such joint measurements.*

---

### 1. INTRODUCTION

In quantum theory, the issue of joint measurability of incompatible observables has been much debated. Discussion centers around two questions: are incompatible observables jointly measurable at all and, if so, is there a relation (typographically) similar to Robertson's<sup>(47)</sup> (i.e.,  $\langle \Delta^2 \mathbf{A} \rangle \langle \Delta^2 \mathbf{B} \rangle \geq \frac{1}{4} | \langle [\mathbf{A}, \mathbf{B}]_- \rangle |^2$ ) restricting the accuracy of such joint measurements? Such a relation would represent an *inaccuracy principle*.

An answer to this question is of some interest because the extension of "the" uncertainty principle to accuracy of joint measurement is dubious<sup>(3,27,43)</sup> but nevertheless frequently suggested (Ref. 5, p. 446; Ref. 16, p. I-II; Ref. 20, p. 81; Ref. 21, p. 162; Ref. 39, Sec. III.4), and also because in such measurements (if possible at all) one might hope to find typically quantum mechanical inaccuracies limiting the possibilities of (future) high-precision measurement devices.

---

<sup>1</sup> Department of Theoretical Physics, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

One's first thought might be that Robertson's relation itself expresses an "inaccuracy principle." The first to attempt to prove this was von Neumann (Ref. 39, Sec. V.4). He used a **C** measurement of the first kind to effect a joint **A-B**-measurement. The **A**-c.q. **B**-scatter in the state-after-measurement  $|c\rangle\langle c|$  is associated with measurement inaccuracy. Von Neumann gave measurement a strange meaning in this argument. Usually this term is applied to a procedure that yields information about quantities associated with the premeasurement state of the object that was not available before. In other words, the crucial aspects of measurement is its *determinative* character. Von Neumann concentrated exclusively on the preparative side (in general we cannot even extract the expectation values of **A** and **B** from a **C**-measurement), and consequently his proposal does not show any quantum limitations of determinative possibilities (i.e., an inaccuracy principle) to exist. More generally, it can be seen that proposals intended to establish an inaccuracy principle based on the scatter in the state-after-measurement are doomed, because there is no fundamental reason<sup>(23)</sup> why this scatter should be related to the measurement accuracy. In particular, destructive measurements would not seem to be affected by an inaccuracy principle derived along these lines.

Others (Ref. 10; Ref. 20, p. 81; Ref. 35, p. 121) have claimed that the Robertson relation can be given such an inaccuracy interpretation if one gives a suitable meaning to  $\langle \Delta^2 \mathbf{A} \rangle$  and  $\langle \Delta^2 \mathbf{B} \rangle$ . The only meaning Born's probabilistic interpretation gives these symbols, however, is scatter in *independent, accurate* measurements (Ref. 3, p. 23; Ref. 27, p. 197; Ref. 43, p. 225). Robertson's relation limits *preparative* possibilities. Measurement results (of **A** and **B** in the Robertson case) only play a part insofar as they are used to characterize the preparative quality, and the quality of the measurements themselves is not at stake. The symbols  $\langle \Delta^2 \mathbf{A} \rangle$  and  $\langle \Delta^2 \mathbf{B} \rangle$  are properties of the preparation device, and consequently Robertson's relation is in general irrelevant for the joint measurement problem (but see Sec. 4). Hence we suggest<sup>(31)</sup> the interpretation of the uncertainty principle in terms of two independent subprinciples: a scatter principle and an inaccuracy principle.

Although the von Neumann-Dirac formalism,<sup>(15,39)</sup> in which scatter relations like the Robertson relation are the only available representatives of the uncertainty principle, still has some authority among physicists, its approach to measurement is losing ground in the light of more modern developments, which generalize it. One argument for such a generalization is that it is unsatisfactory to limit oneself *by postulate* to a class of observables (in the von Neumann-Dirac case those represented by self-adjoint operators) that are, in some "self-evident" sense, "optimal." A more suitable approach would select a class of "optimal" measurements from the

class of all measurements by using *operational* arguments. Only in this way can we be sure that the limitation to “optimal” observables is possible without loss of generality. The generalized frame in which the joint measurement problem can (and has been) discussed, and which we will use, is that of the *positive operator-valued measures* (POVM’s).<sup>(12, 17, 19, 22, 27, 46)</sup> A POVM is (for a discrete outcome set  $K$ ) a family of bounded linear operators  $\{\mathbf{M}_k\}_{k \in K}$  satisfying the relations

$$\forall_{k \in K} \mathbf{M}_k \geq \mathbf{0} \tag{1}$$

$$\sum_{k \in K} \mathbf{M}_k = \mathbf{1} \tag{2}$$

(Operators are boldfaced.) If the object system is in a state represented by the density operator  $\rho$ , a  $\{\mathbf{M}_k\}_{k \in K}$ -measurement will yield outcome  $k$  with probability  $\text{Tr}(\rho \mathbf{M}_k)$ .

Textbook observables (*simple observables*) are represented by *projection-valued measures* (PVM’s ~ orthogonal spectral resolutions of self-adjoint operators). PVM’s obey, in addition to (1) and (2),

$$\forall_{k \in K} \mathbf{M}_k^2 = \mathbf{M}_k \tag{3}$$

Define further:

**Definition 1.** The POVM’s  $\{\mathbf{M}_k\}_{k \in K}$  and  $\{\mathbf{N}_l\}_{l \in L}$  are *coexistent*

:=

$$\exists_{\text{POM}\{\mathbf{O}_{kl}\}_{(k,l) \in K \times L}} \begin{cases} \sum_{k \in K} \mathbf{O}_{kl} = \mathbf{N}_l \\ \sum_{l \in L} \mathbf{O}_{kl} = \mathbf{M}_k \end{cases}$$

The POM  $\{\mathbf{O}_{kl}\}_{(k,l) \in K \times L}$  in Definition 1 is said to represent a *joint (ideal) measurement* of  $\{\mathbf{M}_k\}_{k \in K}$  and  $\{\mathbf{N}_l\}_{l \in L}$ . Two observables are termed *compatible* whenever their POVM’s are coexistent (Ref. 19, Sec. II.6). As is well known, in quantum mechanics not all observables are compatible. Our definition of compatibility coincides with the usual one for simple observables, as the following theorem shows:

**Theorem 1.** If two POVM’s  $\{\mathbf{M}_k\}_{k \in K}$  and  $\{\mathbf{N}_l\}_{l \in L}$  satisfy  $\forall_{k \in K, l \in L} [\mathbf{M}_k, \mathbf{N}_l]_- = 0$ , they are coexistent. If they are PVM’s the converse is also true.

The second part of Theorem 1 is well known (Ref. 19, Proposition II.6.1; Ref. 26). The first part can be proved by the explicit construction of the POM  $\{\mathbf{O}_{kl}\}_{(k,l) \in K \times L}$ ;  $\mathbf{O}_{kl} := \mathbf{M}_k \mathbf{N}_l$ .

In the following we shall not go into the preparative aspects of measurement (i.e., into the question of what happens to the object system after the measurement) since these are irrelevant to the definition of measurement as such.<sup>2</sup> Hence issues concerning, for example, “measurements of the first kind” or quantum nondemolition measurements will not be treated here. In the last section we will compare the quantum estimation theoretic approach to ours, and we shall also mention possibilities of realization.

But first we shall study the joint measurement of incompatible simple observables on finite-dimensional spaces. In de Muynck *et al.*<sup>(24,37,38)</sup> it was required that the joint measurement procedure should yield marginals unbiased with respect to the original outcome sets. Analogous requirements were urged by Busch<sup>(6)</sup> and by Schroeck.<sup>(48)</sup> It was demonstrated<sup>(24,38)</sup> that, on this definition, the spreading  $\langle \Delta^2 \mathbf{A}_{\text{joint}} \rangle$  in the joint measurement procedure must necessarily exceed the value  $\langle \Delta^2 \mathbf{A} \rangle$  obtained in an ideal measurement, thus accounting for an excess spreading obtained already by Arthurs and Kelly<sup>(2,50)</sup> in a joint measurement of position and momentum.

A drawback of this latter definition is that it depends on the labeling of the outcome set (see Sec. 2). For this reason in the present paper another criterion has been used (Definition 2), based on a nonideality notion that is discussed in more detail elsewhere.<sup>(29)</sup> It enables the derivation (Sec. 3) of an inequality limiting the accuracy achievable in the joint measurement of arbitrary simple observables on finite-dimensional spaces. (Some of these results were previewed in Ref. 30.) In Sec. 4 a particular example is studied in more detail that will allow us to contrast our approach to others more closely (Sec. 5).

## 2. JOINT NONIDEAL MEASUREMENTS

The quantum mechanical representatives of measurements are POVM's. Since the class of POVM's is so general, it contains many elements corresponding to uninteresting measurements. Hence it should be profitable to structure this set so that we can discern “better” measurements from “worse” ones. We shall use the notation  $\{\mathbf{N}_l\}_{l \in L} \rightarrow \{\mathbf{M}_k\}_{k \in M}$  for: “a measurement of the POVM  $\{\mathbf{M}_k\}_{k \in K}$  is a nonideal measurement of the POVM  $\{\mathbf{N}_l\}_{l \in L}$ .” This is defined more quantitatively as<sup>(29)</sup>

<sup>2</sup> If we take the preparative aspects into consideration, a measurement device is represented by an *instrument* (Ref. 12).

**Definition 2.**  $\{\mathbf{N}_l\}_{l \in L} \rightarrow \{\mathbf{M}_k\}_{k \in K}$

:=

$$\exists_{\{\lambda_{kl}\}} \left( \lambda_{kl} \geq 0 \wedge \sum_{k \in K} \lambda_{kl} = 1 \wedge \mathbf{M}_k = \sum_{l \in L} \lambda_{kl} \mathbf{N}_l \right)$$

A somewhat more restrictive version of the relation  $\rightarrow$  is defined by

**Definition 3.**  $\{\mathbf{N}_l\}_{l \in L} \rightarrow^i \{\mathbf{M}_k\}_{k \in K}$

:=

$$\begin{cases} \exists_{\{\lambda_{kl}\}} (\lambda_{kl} \geq 0 \wedge \sum_{k \in K} \lambda_{kl} = 1 \wedge \mathbf{M}_k = \sum_{l \in L} \lambda_{kl} \mathbf{N}_l) \\ \exists_{\{\mu_{lk}\}} \mathbf{N}_l = \sum_{k \in K} \mu_{lk} \mathbf{M}_k \end{cases}$$

In this case we speak of an “invertibly nonideal” measurement.

Similar relations have been used by Davies [also by Prugovečki *et al.* (Ref. 46 and references therein) and by Busch<sup>(6-10)</sup>] for the representation of joint measurements of position and momentum.<sup>(11,12)</sup> These earlier approaches differ from ours in that they require  $L = K$  in Definition 2. Such an extra demand is unwanted because it gives the scale an unphysical importance: the scale of a measurement device is merely a matter of convenience; the physics of the measurement device is not altered by changing its scale (Ref. 27, p. 97; Ref. 29). It would for instance exclude the possibility of a (nonideal) spin- $\frac{1}{2}$  measurement with an outcome set equal to the whole of  $\mathbb{R}$  (this would actually be a suitable description of an instrument consisting of a screen placed after a Stern–Gerlach device). Such an extra restriction on Definition 2 would also jeopardize the generality of an inaccuracy relation derived with its help.

We use the term “nonideal” here as a technical term for a *relational* notion which can informally be denoted as “course-grained,” “approximate,” or “inaccurate.” Such a use of the word “nonideal” should be justified on operational grounds, and that can be done as follows. We can use the nonideal measurement to estimate parameters (linear functionals, e.g., the moment or the probability for some interval) of the outcome probability distribution of the observable we wanted to measure. It is, after all, this last distribution in which we are interested. In the case of  $\rightarrow^i$  this is even possible for all parameters, i.e., one can calculate the entire probability distribution of the desired observable, given the distribution of the nonideal measurement.

In both cases the nonideality shows itself in the fact that the number of repetitions needed to estimate such a parameter with a given reliability using the nonideal measurement will always be larger than if one had been

able to measure the ideal.<sup>(29)</sup> A nonideal measurement provides us with less reliable information than the one we wanted to perform.

On a more conceptual level we can, following Prugovečki,<sup>(45,46)</sup> interpret  $\lambda_{kl}$  as a measure for the confidence in the “real” result  $l$ , given that our nonideal measurement gave outcome  $k$ .

It is not difficult to verify that both  $\rightarrow$  and  $\rightarrow^i$  define partial order relations (hence we speak of a nonideality *structure*) between equivalence classes generated by an equivalence relation naturally defined by

$$\begin{aligned} \text{Definition 4. } \quad & \{\mathbf{M}_k\}_{k \in K} \leftrightarrow \{\mathbf{N}_l\}_{l \in L} \\ & := \\ & \{\mathbf{N}_l\}_{l \in L} \rightarrow \{\mathbf{M}_k\}_{k \in K} \wedge \{\mathbf{M}_k\}_{k \in K} \rightarrow \{\mathbf{N}_l\}_{l \in L} \end{aligned}$$

Equivalent POVM’s are associated with procedures that measure physically equivalent “quantities.” Of special interest are the observables that are “optimal,” that cannot be seen as nonideal versions of other, nonequivalent observables. In terms of this partial ordering, these correspond to *maximal* POVM’s:

$$\begin{aligned} \text{Definition 5. } \quad & \{\mathbf{M}_k\}_{k \in K} \text{ is maximal} \\ & := \\ & \{\mathbf{N}_l\}_{l \in L} \rightarrow \{\mathbf{M}_k\}_{k \in K} \Rightarrow \{\mathbf{N}_l\}_{l \in L} \leftrightarrow \{\mathbf{M}_k\}_{k \in K} \end{aligned}$$

It can be shown that a POVM is maximal iff all its elements are up to a scalar factor one-dimensional projectors.<sup>2</sup> This means that our notion of maximality is a generalization of the usual one (= nondegeneracy) that is only applicable to PVM’s. It also means that in quantum mechanics there is no *unique* equivalence class of maximal POVM’s, whereas in classical theories the equivalence class of maximal “POVM”’s is unique.<sup>(29)</sup> This fact immediately implies that there are incompatible observables in quantum theory (and none in classical theories), as the following theorem shows (in conjunction with Definition 5):

$$\begin{aligned} \text{Theorem 2. } \quad & \{\mathbf{M}_k\}_{k \in K} \text{ and } \{\mathbf{N}_l\}_{l \in L} \text{ coexistent} \\ & \Updownarrow \\ & \exists_{\text{POM}\{\mathbf{O}_m\}_{m \in M}} \{\mathbf{O}_m\}_{m \in M} \rightarrow \{\mathbf{M}_k\}_{k \in K} \wedge \{\mathbf{O}_m\}_{m \in M} \rightarrow \{\mathbf{N}_l\}_{l \in L} \end{aligned}$$

Such a maximal POVM need of course *not* be realized in an instrument (Ref. 12) corresponding to (an analog of) a measurement of the first kind.

Proof of the theorem follows straightforwardly from Definition 1 and Definition 2.

The importance of the theorem is that it shows that the typical feature of quantum measurements, viz. incompatibility, can be defined completely in terms of the relation  $\rightarrow$ . This suggests that the properties of the nonideality structure on the class of observables for a given statistical theory can be used to characterize the theory *in abstracto*, much as such an abstract characterization is the aim of quantum logic. Given the above, the natural definition of joint (nonideal) measurement is:

**Definition 6.** The POVM  $\{\mathbf{O}_{mm'}\}_{(m,m') \in M \times M'}$  is associated with a *joint nonideal measurement* of the POVM's  $\{\mathbf{M}_k\}_{k \in K}$  and  $\{\mathbf{N}_l\}_{l \in L}$

$$:= \{\mathbf{N}_l\}_{l \in L} \rightarrow \left\{ \sum_{m \in M} \mathbf{O}_{mm'} \right\}_{m' \in M'} \wedge \{\mathbf{M}_k\}_{k \in K} \rightarrow \left\{ \sum_{m' \in M'} \mathbf{O}_{mm'} \right\}_{m \in M}$$

The term *joint invertibly nonideal measurement* is defined in a similar way, using  $\rightarrow^i$ .

A particularly interesting joint measurement problem arises when  $\{\mathbf{M}_a\}_{a \in A}$  and  $\{\mathbf{N}_b\}_{b \in B}$  and PVM's of the self-adjoint operators **A** and **B**, respectively. In that case we can use:

**Theorem 3.** Suppose  $\{\mathbf{M}_a\}_{a \in A}$  is given as the PVM of  $\mathbf{A} = \sum_{a \in A} a \mathbf{M}_a$  on  $\mathcal{H}$ . If  $\{\mathbf{E}_r\}_{r \in R}$  on  $\mathcal{H} \otimes \mathcal{H}'$  is the PVM of  $\mathbf{R} = \sum_{r \in R} r \mathbf{E}_r$ , and  $\mathbf{R}$  can be written as a function  $r(\mathbf{A}, \mathbf{T})$  of  $\mathbf{A}$  and a self-adjoint operator  $\mathbf{T}$  on  $\mathcal{H}'$ , then the POVM  $\{\mathbf{O}_r\}_{r \in R}$  defined by

$$\mathbf{O}_r = \text{Tr}_{\mathcal{H}'}(\rho' \mathbf{E}_r)$$

has the property  $\{\mathbf{M}_a\}_{a \in A} \rightarrow \{\mathbf{O}_r\}_{r \in R}$  for every density operator  $\rho'$  on  $\mathcal{H}'$ .

Proof of this theorem is straightforward, using the spectral theorem on  $\mathcal{H}$ .

This theorem can be a powerful help (and indeed it has been used implicitly<sup>(17,19,42)</sup>) in trying to find a joint measurement of **A** and **B** by looking for a second Hilbert space<sup>4</sup>  $\mathcal{H}'$ , two Hermitian operators **C** and **D** on  $\mathcal{H}'$ , two functions  $f$  and  $g$ , such that the operators  $f(\mathbf{A}, \mathbf{C})$  and  $g(\mathbf{B}, \mathbf{D})$  on  $\mathcal{H} \otimes \mathcal{H}'$  commute. If these exist, they give rise to a joint PVM on  $\mathcal{H} \otimes \mathcal{H}'$ , which gives, after partial tracing over  $\mathcal{H}'$  with some density

<sup>4</sup> In some cases the extra Hilbert space has a physical interpretation in terms of an "ancillary" system (see, e.g., Ref. 17).

operator  $\rho'$  on  $\mathcal{H}'$ , a POVM on  $\mathcal{H}$  that meets the demands of Definition 6. We shall indeed see an example of this use of Theorem 3 in Sec. 4. But already at this point one might think of a very simple solution to the problem:

On  $\mathcal{H}'$  we always have projectors  $\mathbf{E}$  and  $\mathbf{E}^\perp = 1 - \mathbf{E}$ . We can define the following operators on  $\mathcal{H} \otimes \mathcal{H}'$ :

$$\mathbf{F} := \mathbf{E} \otimes \mathbf{A}, \quad \mathbf{G} := \mathbf{E}^\perp \otimes \mathbf{B}$$

The operators  $\mathbf{F}$  and  $\mathbf{G}$  commute, and therefore have a joint PVM. For an arbitrary  $\rho'$  on  $\mathcal{H}'$  this can give a POVM on  $\mathcal{H}$  in the following way: If 0 is not an eigenvalue of  $\mathbf{A}$  (i.e.,  $0 \notin A$ ), introduce  $\mathbf{M}_0 = \mathbf{0}$  into the PVM  $\{\mathbf{M}_a\}_{a \in A}$  and the value 0 into  $A$ . Then the PVM corresponding to  $\mathbf{F}$  always has the outcome set  $A = \{0, a_2, \dots, a_p\}$ .

We treat the PVM  $\{\mathbf{N}_b\}_{b \in B}$ ,  $\mathbf{B}$ , and  $\mathbf{G}$  analogously. This leads to a spectrum  $B := \{0, b_2, \dots, b_q\}$  for  $\mathbf{G}$ . Then the joint PVM  $\{\mathbf{P}_{ab}\}_{(a,b) \in A \times B}$  on  $\mathcal{H} \otimes \mathcal{H}'$  is given by

$$\begin{aligned} \mathbf{P}_{ab} &= \delta_{a0} \delta_{b0} (\mathbf{E} \otimes \mathbf{N}_0 + \mathbf{E}^\perp \otimes \mathbf{M}_0) + (1 - \delta_{a0}) \delta_{b0} \mathbf{E}^\perp \otimes \mathbf{M}_a \\ &\quad + (1 - \delta_{b0}) \delta_{a0} \mathbf{E} \otimes \mathbf{N}_b \\ &= \delta_{b0} \mathbf{E}^\perp \otimes \mathbf{M}_a + \delta_{a0} \mathbf{E} \otimes \mathbf{N}_b \end{aligned}$$

Partial trace gives the POVM  $\{\mathbf{O}_{ab}\}_{(a,b) \in A \times B}$ :

$$\mathbf{O}_{ab} = \text{Tr}_{\mathcal{H}'}(\rho' \mathbf{P}_{ab}) = \lambda \delta_{b0} \mathbf{M}_a + (1 - \lambda) \delta_{a0} \mathbf{N}_b \quad (4)$$

with

$$\lambda = \text{Tr}(\rho' \mathbf{E}')$$

Indeed  $\{\mathbf{O}_{ab}\}$  can be seen to satisfy definition (6) for  $\{\mathbf{M}_a\}$  and  $\{\mathbf{N}_b\}$ . A POVM essentially the same as (4) was proposed by Abu-Zeid.<sup>(1)</sup> Relation (4) can evidently be extended to satisfy Definition 6 for two arbitrary incompatible POVM's. Unfortunately, it represents a *trivial* joint measurement. A joint measurement of two POVM's as in Definition 6 is called trivial iff

$$\forall_{(m,m') \in M \times M'} \mathbf{O}_{mm'} \in \text{span}(\{\mathbf{M}_k\}_{k \in K} \cup \{\mathbf{N}_l\}_{l \in L})$$

One of the most important reasons to do a joint measurement of two observables is the possibility of determining some kind of correlation between the values these two observables "assume." In a trivial joint measurement the joint probability distribution is fixed if its marginals are given. Hence the correlation in the joint probability distribution depends on the particular procedure used, but it does not depend on the state of the

object system at all. Trivial joint measurements do not provide information about  $\rho$  beyond the information obtainable with separate measurements of the observables involved. In this case triviality results from the fact that (4) is a convex sum of two POVM's, each of which has one uninformative marginal. Equation (4) can be realized using a procedure where sometimes (probability  $\lambda$ )  $\{\mathbf{M}_a\}_{a \in A}$  is measured, and sometimes (probability  $1 - \lambda$ )  $\{\mathbf{N}_b\}_{b \in B}$ . The observable that is not measured is attributed the value 0. This means that we shall have to search a little harder to find a meaningful joint nonideal measurement of two incompatible observables.

In order to be able to derive a nonideality inequality we shall have to introduce a nonideality measure that expresses the amount of nonideality inherent in a given relation  $\rightarrow$  as a number. We shall use a measure based on *mutual information*<sup>(32,29)</sup>:

**Definition 7.**

$$I_{\{\mathbf{N}_l\} \rightarrow \{\mathbf{M}_k\}} := \inf_{\{\lambda_{kl}\} \in \Lambda} \left\{ \sum_{k \in K} \sum_{l \in L} q_{kl} \log(q_{kl}/(r_k p_l)) \right\}$$

with

$$\{\mathbf{M}_k\}_{k \in K}, \{\mathbf{N}_l\}_{l \in L} \text{ given POM's with } \{\mathbf{N}_l\}_{l \in L} \rightarrow \{\mathbf{M}_k\}_{k \in K}$$

$$L = \{l_1, \dots, l_m\}$$

$$\Lambda := \left\{ \{\lambda_{kl}\} \mid \lambda_{kl} \geq 0 \wedge \sum_{k \in K} \lambda_{kl} = 1 \wedge \sum_{l \in L} \lambda_{kl} \mathbf{N}_l = \mathbf{M}_k \right\}$$

$$p_l := \text{Tr}(\mathbf{N}_l)/\text{dim}(\mathcal{H}), \quad q_{kl} := \lambda_{kl} p_l, \quad r_k := \sum_{l \in L} q_{kl}$$

The measure of Definition 7 is zero when the nonideal measurement is worthless, and it is maximal when it is perfect. This maximal value is never larger than  $\log(m)$ . As indicated by this definition, we shall restrict ourselves to finite-dimensional spaces. These are technically simple, while preserving the essential characteristics of infinite-dimensional spaces, as far as this problem is concerned.

**3. MAIN RESULTS**

We take  $\mathcal{H} = \mathbb{C}^n$ , with the usual inner product. Two complete orthonormal bases  $(|x_k\rangle)_{k \in K}$  and  $(|y_k\rangle)_{k \in K}$  ( $K = \{0, \dots, n-1\}$ ) are given. They define the maximal PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$ :

$$\mathbf{E}_k := |x_k\rangle\langle x_k|, \quad \mathbf{F}_k := |y_k\rangle\langle y_k|$$

Our results will be derived for these PVM's. Since they are completely arbitrary, apart from the finite-dimensionality of the Hilbert space, the results are nevertheless quite general.

The characteristic parameter of this problem is the inner product matrix  $\{f_{kl}\}$ :

$$f_{kl} := \langle x_k | y_l \rangle$$

We shall assume (without loss of generality) that this matrix is in block form:

$$\{f_{kl}\} = \begin{bmatrix} & | & 0 & \dots & 0 \\ 0 & \boxed{\phantom{0}} & & & \vdots \\ \vdots & & & \boxed{\phantom{0}} & \\ 0 & \dots & 0 & & | \end{bmatrix}$$

and that all blocks are irreducible. The PVM's in question are compatible iff the matrix consists entirely of 1's and 0's. Then:

**Theorem 4.** The PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  (defined above) are jointly nonideally measurable. The joint measurement can be both nontrivial and invertible.

**Theorem 5.** Suppose a POVM  $\{\mathbf{M}_{lm}\}_{(l,m) \in L \times M}$  represents a joint nonideal measurement of the PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  (defined above). The quality of this joint measurement is limited by

$$I_{\{\mathbf{E}_k\} \rightarrow \{\sum_{\overline{m}} \mathbf{M}_{lm}\}} + I_{\{\mathbf{F}_k\} \rightarrow \{\sum_{\overline{l}} \mathbf{M}_{lm}\}} \leq 2 \log(n) + \sum_i \frac{n_i}{n} \log(F_i) \quad (5)$$

with

$$F_i := \max_{k,l \in \text{block } \#i} |f_{kl}|^2$$

$$n_i := \dim(\text{block } \#i)$$

Proof can be found in the appendix.

#### Remarks.

Note that in this case for any nonideal measurement of  $\{\mathbf{E}_k\}_{k \in K}$  (or  $\{\mathbf{F}_k\}_{k \in K}$ ) we have  $0 \leq I \leq \log(n)$ , so that Theorem 5 gives a trivial bound

iff  $F_i = 1$  for all  $i$ . This, however, is equivalent to the statement that the PVM's are compatible.

A result comparable to (5) has been obtained for  $\mathbb{C}^2$ , under the assumption  $M = L = K$ , by Busch.<sup>(7,8)</sup>

#### 4. AN EXAMPLE

We shall now proceed with a special case. We have

$$\langle x_k | y_l \rangle = \frac{1}{\sqrt{n}} \exp\left(i \frac{2\pi}{n} kl\right) \quad (k, l \in K = \{0, \dots, n-1\}) \quad (6)$$

The PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  are defined as in the previous section. The  $\mathbb{C}^2$  version of this example corresponds to the case of spin- $\frac{1}{2}$  quantities  $\sigma_x$  and  $\sigma_z$ . More generally, this kind of PVM may be considered a finite-dimensional analog of position and momentum.<sup>(49)</sup> Therefore the results we can obtain should be of some interest outside this finite-dimensional context as well.

Moreover, previous results for the joint measurement problem have almost exclusively been derived for cases similar to this case. To be able to compare these results to ours, it is important to look into this example in more detail.

Define

$$\mathbf{S}_x := \sum_{k \in K} |x_{[[k+1]]}\rangle \langle x_k|, \quad \mathbf{S}_y := \sum_{k \in K} |y_{[[k+1]]}\rangle \langle y_k|$$

with

$$[[a]] := a \bmod n$$

It can be seen that these shift operators satisfy (7a), a relation analogous to the Weyl CCR. The analogy with position-momentum is completed<sup>(49)</sup> by (7b):

$$\mathbf{S}_x^a \mathbf{S}_y^b = \mathbf{S}_y^b \mathbf{S}_x^a \exp\left(-i \frac{2\pi}{n} ab\right) \quad (a, b \in \mathbb{Z}) \quad (7a)$$

$$\mathbf{S}_x = \exp\left(-i \frac{2\pi}{n} \mathbf{Y}\right), \quad \mathbf{S}_y = \exp\left(i \frac{2\pi}{n} \mathbf{X}\right) \quad (7b)$$

with

$$\mathbf{X} := \sum_{k \in K} k \mathbf{E}_k, \quad \mathbf{Y} := \sum_{k \in K} k \mathbf{F}_k$$

In the context of this example we can show<sup>(29)</sup> that for a POVM  $\{\mathbf{O}_l\}_{l \in L}$ :

**Theorem 6.**  $\{\mathbf{E}_k\}_{k \in K} \rightarrow \{\mathbf{O}_l\}_{l \in L} \Leftrightarrow \forall_{l \in L} \mathbf{S}_y \mathbf{O}_l \mathbf{S}_y^\dagger = \mathbf{O}_l$  (invariance)  
[the PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  as defined above, including (6)].

In quantum estimation theory the notion of *covariance* is current:

**Definition 8.** A POVM  $\{\mathbf{M}_k\}_{k \in K}$  is (X-)covariant

:=

$$\mathbf{S}_x \mathbf{M}_k \mathbf{S}_x^\dagger = \mathbf{M}_{[[k+1]]}$$

This leads (Ref. 17; Ref. 19, p. 122) to the following criterion for a joint measurement of two such PVM's:

**Definition 9.** A POVM  $\{\mathbf{M}_{kl}\}_{(k,l) \in K \times K}$  is associated with a covariant joint nonideal measurement of the PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  [defined above, including (6)]

:=

$$\mathbf{S}_x \mathbf{M}_{kl} \mathbf{S}^\dagger = \mathbf{M}_{[[k+1]]l}, \quad \mathbf{S}_y \mathbf{M}_{kl} \mathbf{S}^\dagger = \mathbf{M}_{k[[l+1]]}$$

We can see that this nomenclature is consistent, that the covariance of Definition 9 indeed implies association with a joint nonideal measurement, if we realize that the covariance as in Definition 9 leads to invariance (as in Theorem 6) of the marginals.

Since  $\mathbf{M}_{00} \geq 0$  and  $\mathbf{M}_{kl} = (\mathbf{S}_x^k \mathbf{S}_y^l) \mathbf{M}_{00} (\mathbf{S}_x^k \mathbf{S}_y^l)^\dagger$ , the condition  $\sum_k \sum_l \mathbf{M}_{kl} = \mathbf{1}$  implies that  $\text{Tr}(\mathbf{M}_{00}) = 1/n$ . This means there is a density operator  $\rho_0$  such that (Ref. 19, Sec. III.6)

$$\mathbf{M}_{kl} = \frac{1}{n} (\mathbf{S}_x^k \mathbf{S}_y^l) \rho_0 (\mathbf{S}_x^k \mathbf{S}_y^l)^\dagger$$

The POVM  $\{\mathbf{M}_{kl}\}_{(k,l) \in K \times K}$  is maximal iff  $\rho_0$  is pure.<sup>(29)</sup> This covariant POVM is also of the form (A2). This can be seen by taking

$$C_{k'l'}^{(kl)} = \frac{1}{2} \text{Tr}(\rho_0 \mathbf{F}_{[[l'-l]]} \mathbf{E}_{[[k'-k]]})$$

It is not difficult to verify that the nonideality matrices of the marginals are

$$\begin{aligned} \lambda_{kl}^{(x)} &= \langle x_{[[l-k]]} | \rho_0 | x_{[[l-k]]} \rangle \\ \lambda_{kl}^{(y)} &= \langle y_{[[l-k]]} | \rho_0 | y_{[[l-k]]} \rangle \end{aligned} \quad (8)$$

These nonideality matrices are *symmetric*<sup>(32)</sup>: every row is a permutation of every other row and every column is a permutation of every other column.

We can relate the amount of nonideality in the matrices (8) to the scatter in  $\rho_0$ :

$$\begin{aligned}
 I_{\{E_k\} \rightarrow \{\sum_l M_{kl}\}} &= \log(n) + \sum_{k \in K} p_k \log(p_k) \\
 I_{\{F_k\} \rightarrow \{\sum_k M_{kl}\}} &= \log(n) + \sum_{k \in K} q_k \log(q_k)
 \end{aligned}
 \tag{9}$$

with

$$p_k := \langle x_k | \rho_0 | x_k \rangle, \quad q_k := \langle y_k | \rho_0 | y_k \rangle \quad (k \in K)$$

In this  $n$ -dimensional case  $0 \leq I \leq \log(n)$ . The maximum is achieved iff the associated distribution  $(p_k)_{k \in K}$  c.q.  $(q_k)_{k \in K}$  is dispersion free. That is not achievable for both distributions at the same time, since the observables are incompatible. This is most conveniently represented by the entropic scatter relation<sup>(14,23,28)</sup> we used in the proof of Theorem 5:

$$-\sum_k p_k \log(p_k) - \sum_k q_k \log(q_k) \geq \log(n)
 \tag{10}$$

Combination of (9) and (10) gives a special case of Theorem 5:

$$I_{\{E_k\} \rightarrow \{\sum_l M_{kl}\}} + I_{\{F_k\} \rightarrow \{\sum_k M_{kl}\}} \leq \log(n)
 \tag{11}$$

**Remarks.**

Such an inaccuracy relation for covariant joint measurements has been derived for the (related) position momentum case by Davies and by Kuryshkin.<sup>(11,12,25)</sup> A similar inequality for the special case of spin  $\frac{1}{2}$  was discovered by Prugovečki.<sup>(45)</sup>

Covariance (Definition 8) explicitly confines the labeling of the non-ideal observable: not only does it require  $L = K$  ( $M = K$ ), but covariance also fixes the order of the elements in  $L(M)$  through the restrictions (i.e.,  $\lambda_{kl}^{(j)} = \tilde{\lambda}_{[[k-l]]}^{(j)}$ ) it places on the nonideality matrices. For the analogous position case these extra demands exclude, for instance, the PVM  $\{\mathbf{M}_n\}_{n \in \mathbb{Z}}$ ,  $\mathbf{M}_n := \int_n^{n+1} |x\rangle \langle x| dx$ , as a nonideal position measurement, whereas it is from a practical point of view perfectly acceptable as such.

Hence the demand of covariance limits the generality of (11) and its analogs severely. It is also undesirable in view of the fact that the labeling, however convenient, should not play a fundamental role (Sec. 2).

Of course the fact that covariance is not to be considered a fundamental factor for the choice of a scale for a measuring instrument, in no way diminishes the importance of group theory in quantum mechanics.<sup>5</sup> In fact, group theory shows why hardly anybody has perceived as unsatisfactory the fact that textbooks use (almost) exclusively the von Neumann–Dirac formalism. Practical quantum problems (for the treatment of which these textbooks train the student), such as scattering, can be treated quite well using self-adjoint operators; POVM's are not at all involved. The reason for this is that in such problems “physical quantities” are involved. All these “quantities” (position, momentum, spin,...) have a group theoretical background (like the correspondence principle<sup>(44)</sup>), and the generators of the quantum representations of the transformations associated with these “quantities” are self-adjoint operators.

But in the treatment of problems like scattering, measurement is not involved, and therefore our conclusion regarding the applicability of covariance to the scale of a measuring instrument is not at all related to the usefulness of (self-adjoint operators or) group theory to other problems.

We can now show how to apply Theorem 3 to this case. Take  $\mathcal{H}' = \mathbb{C}^n$ , and define operators  $X'$  and  $Y'$  on  $\mathcal{H}'$  just as  $X$  and  $Y$  were defined on  $\mathcal{H}$  [including (6)].

**Theorem 7.** The operators  $X'' := [[X \otimes \mathbf{1}' + \mathbf{1} \otimes X']]$  and  $Y'' := [[Y \otimes \mathbf{1}' - \mathbf{1} \otimes Y']]$  on  $\mathcal{H} \otimes \mathcal{H}'$  commute.

Now, using a density operator  $\rho_0$  on  $\mathcal{H}'$ , we get our covariant joint-measurement POVM from the joint PVM of  $X''$  and  $Y''$  through partial trace, as in Theorem 3. The infinite-dimensional analog (in the case of position-momentum) of Theorem 7 is well known and has been applied to the joint measurement problem<sup>(17,18,19,42)</sup> in the same way as Theorem 7.

Theorem 7 also gives us the opportunity to illustrate in what way scatter and inaccuracy principle can be related:

(a) The observables associated with the operators  $X'' := [[X \otimes \mathbf{1}' + \mathbf{1} \otimes X']]$  and  $Y'' := [[Y \otimes \mathbf{1}' - \mathbf{1} \otimes Y']]$  on  $\mathcal{H} \otimes \mathcal{H}'$  are compatible (Theorem 7). Then the accuracy of a  $(X'', Y'')$ -measurement, interpreted as a joint nonideal  $(X, Y)$ -measurement, is directly related to the scatters in  $X'$  and in  $Y'$ .

<sup>5</sup> The work of, for example, Ludwig (Ref. 27) shows that insistence on the unimportance of scales of measurement devices does not mean that group theory should not have a prominent position in quantum mechanics.

(b) There exist density operators on  $\mathcal{H} \otimes \mathcal{H}'$  that are sharp in both  $\mathbf{X}''$  and  $\mathbf{Y}''$ , since these operators commute (Theorem 7). Then a joint non-ideal measurement of  $\mathbf{X}'$  and  $\mathbf{Y}'$  performed on such a state can be used to prepare states on  $\mathcal{H}$ . For states that are thus prepared the scatters in  $\mathbf{X}$  and in  $\mathbf{Y}$  are directly related to the respective nonidealities in the joint  $(\mathbf{X}', \mathbf{Y}')$  measurement on  $\mathcal{H}'$ .

The relation described above can only exist for joint measurement problems solvable through Theorem 3, as is the case here, and with, for example, position-momentum. Note that in case (b) the scatter is associated with the auxiliary Hilbert space rather than with the object system. This fact makes it hardly accurate to speak of a “reinterpretation of (the symbols in) a scatter relation.”

The covariant POVM of Definition 9 has one more interesting property. It can separate the states<sup>(6,8,56)</sup> (*informational completeness*). The PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  do not have this property, not even together.<sup>(4,52)</sup> Trivial joint measurement POVM's therefore cannot achieve informational completeness either. The following theorem gives the demands  $\rho_0$  is to satisfy, in order to get informational completeness:

**Theorem 8.** There is a 1-1 relation between  $\rho$  and the probability distribution  $(p_{kl})$  ( $p_{kl} := \text{Tr}(\rho \mathbf{M}_{kl})$ , the POVM  $\{\mathbf{M}_{kl}\}_{(k,l) \in K \times K}$  as in Definition 9) iff the density operator  $\rho_0 := n \mathbf{M}_{00}$  satisfies

$$\forall_{a,b \in K} \text{Tr}(\rho_0 \mathbf{S}_x^a \mathbf{S}_y^b) \neq 0$$

The possibility of informational completeness has prompted Busch,<sup>(6)</sup> for example, to argue that  $\{\lambda_{kl}\}$  should not be interpreted as nonideality. But we do not interpret the POVM *itself* as nonideal. It is the marginals that are nonideal. The joint measurement POVM can be informationally complete, and it can also be maximal. The POVM as a whole contains “correlations” that give information which cannot be obtained through the marginals alone. For this reason the properties of the marginals should be clearly distinguished from those of the POVM itself. Moreover, an argument like Busch's would hardly affect the operationalization we mentioned in Sec. 2, which ensures the identification of  $\{\lambda_{kl}\}$  with nonideality in a practically meaningful sense.

The state separation condition is, not surprisingly, stronger than the invertibility condition. This follows from the next theorem:

**Theorem 9.** The POVM  $\{\mathbf{M}_{kl}\}_{(k,l) \in K \times K}$  as in Definition 9 is associated with a joint invertibly nonideal measurement of the PVM's  $\{\mathbf{E}_k\}_{k \in K}$  and

$\{\mathbf{F}_k\}_{k \in L}$  [defined above, including (6)] iff the density operator  $\rho_0 := n\mathbf{M}_{00}$  satisfies

$$\forall_{a \in K} \text{Tr}(\rho_0 \mathbf{S}_x^a) \neq 0 \wedge \text{Tr}(\rho_0 \mathbf{S}_y^a) \neq 0$$

Proofs can be found in the appendix.

## 5. CONCLUDING REMARKS

We have derived for the first time an inaccuracy inequality (Theorem 5) that is applicable to a quite general class of pairs of incompatible observables. This inequality shows that quantum mechanics indeed entails an inaccuracy principle. It cannot be interpreted as a scatter relation since it relates measurement accuracies, i.e., properties of the measurement device alone, independent of preparation (density operator). It also shows the precise nature of the bounds quantum mechanics sets to our ability to measure.

Historically the most interesting pair of incompatible observables is position-momentum. Because of the technically difficult situation (infinite-dimensional Hilbert space, continuous spectrum) a nonideality inequality as general as Theorem 5 has not been derived for this case as yet. That these observables are jointly measurable is nevertheless well known (Refs. 11, 12; Ref. 17, Sec. V.6; Ref. 19, Sec. II.6). This is especially easy to see<sup>(42)</sup> through the use of Theorem 3. For other infinite-dimensional cases, joint nonideal measurement POVM's can probably also be derived, for instance using extensions of the Hilbert space, like in Theorem 3 (see Ref. 30 for an example).

If we try to compare our results to other approaches in the literature, we see that a major problem in the discussion around the joint measurement problem is the divergence of the opponents' opinions about the meaning of the words "joint," "measurement," and "accuracy." It is clear that the answers to the relevant questions depend to a large extent on the context of these words. One such proposal using a measurement concept differing from ours, leading to different conclusions as regards joint measurement possibilities, by Park and Margenau<sup>(40,41)</sup> has already been criticized elsewhere.<sup>(37)</sup> We shall remark here only that all the "joint measurement" procedures they give are trivial. Very strong demands on "measurement" lead other authors<sup>(33,51)</sup> to conclude that joint measurements of incompatible observables are entirely impossible. This conclusion can only be obtained by ignoring or underestimating the possibility of allowing "inaccuracy" in the measurement, as the results in the previous

sections have shown. We shall consider two other characteristic proposals in somewhat more detail.

One proposal is due to Busch and Lahti.<sup>(9)</sup> They argue that the thesis that quantum mechanics admits arbitrarily “sharp” joint measurements of incompatible observables can be substantiated if only “measurement” itself is given a suitable content. The proposal depends on an interpretation of the natural partial ordering on the set of *effects* (Hermitian operators,  $\geq 0$ ,  $\leq 1$ ). If  $A \leq B$ , they claim this means that a procedure realizing the effect  $A$  “gives information” about  $B$ . A joint measurement of two effects  $A$  and  $B$  should therefore realize an effect  $C \neq 0$ ,  $C \leq A$ ,  $C \leq B$ . That this interpretation leads to difficulties<sup>(13)</sup> is clear from the following example:

Let  $A = 1/2$ ;  $B = 1/2 + C/2$ ;  $C$  an arbitrary effect. Then obviously  $A \leq B$ , so that a realization of  $A$  “gives information” about  $B$ . But the probability of a positive result of an  $A$ -realization does not depend on the state of the object system, so that such a procedure gives no information at all, let alone information about  $B$ .

These difficulties surface in a description of joint measurement along the lines of Busch and Lahti. Consider the case of Sec. 4 with  $\mathcal{H} = \mathbb{C}^2$ . We have two effects  $A = aE_1 + (1-a)E_2$  and  $B = bF_1 + (1-b)F_2$  ( $0 \leq a \leq b \leq \frac{1}{2}$ ). Then it can be seen that the greatest lower bound of  $A$  and  $B$  is equal to  $X = a1 + 2[(b-a)(1-b-a)/(1-2a)]E_2$ . Obviously  $X \neq 0$  for arbitrary small  $a, b > 0$ , for arbitrarily small “fuzziness.” But  $X$  tells us no more about the  $B$ - or  $\{F_k\}$ -probabilities than did  $A$  or indeed  $\{E_k\}$  itself. Thus a measurement based on  $X$  can be “bad,” even though the fuzziness in  $A$  and  $B$  is very small. Hence calling an  $A$ -realization a  $B$ -measurement if only  $A \leq B$  seems to imply an unreasonably weak definition of “measurement.”

Inequality (11), a nonideality inequality for covariant joint measurement, is also widely used in quantum estimation theory<sup>(17-19)</sup> (QET). The results of QET in this area rely heavily on *unbiasedness* ( $:=$  correctness of the expectation value) as well as covariance. It can be shown<sup>(29)</sup> that these requirements together imply (a special case of)  $\rightarrow^i$ . Hence Eq. (11) shows that for *the particular case* of Sec. 4 the results of QET are compatible with ours, although weaker.

One can, however, doubt the more general relevance of the methods QET uses for this case. Covariance and unbiasedness are unnecessary for the purpose of generalizing notions like “momentum measurement” as our nonideality concept offers a natural alternative (*viz.* Theorem 6). These demands are also unnecessary for meaningful joint measurements, as Sec. 3 shows.

Moreover, as the consistency of Definition 9 uses the special properties

of the operators  $X$  and  $Y$  of Definition 7 in a highly nontrivial way, it does not seem plausible that the QET approach to joint measurement is applicable outside the range of the observables in this section (and their analogs, such as position-momentum). Indeed the QET literature offers no other examples.

To give covariance and unbiasedness a prominent position is also not acceptable for a number of other reasons. In the first place, as we discussed in Sec. 4, covariance fixes the scale of the measurement device. In the second place it is hardly so that the expectation value is always the only (or even the most) interesting result of a sequence of measurements. Therefore it is a bit odd to give the expectation value such precedence over other parameters of the outcome distribution through the criteria for joint measurement.

In the third place the QET approach gives no clue as to how individual results are to be interpreted as “inaccurate” results of an ideal measurement. This last aspect is precisely the characteristic element of inaccuracy or nonideality in physics. In our approach this interpretation can be achieved by giving  $\{\lambda_{ki}\}$  a likelihood-like meaning.

For these reasons it seems slightly premature to call the QET approach to joint measurement “the” solution to the joint measurement problem, as, for example, Primas maintains (Ref. 44, p. 232).

We end the paper with some remarks on experimental realizations. These can be seen to fit well in our approach, showing that it indeed leads to practicable joint measurements.

Joint measurements of position and momentum have been realized in optical heterodyning<sup>(54)</sup> (see also Ref. 53 and references therein; heterodyning is a destructive measurement). In the case of spin  $\frac{1}{2}$  (a special case of the example in Sec. 4) several experiments have been proposed by Busch<sup>(8)</sup> (these are in general biased). A realization of a simplified version of one of these by Mittelstaedt *et al.*<sup>(34)</sup> to obtain mixed wave-particle behavior shows that Busch’s proposals are indeed achievable.

#### ACKNOWLEDGMENT

One of the author’s (HM) investigations were supported by the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO). The authors would like to thank J. de Graaf for his careful reading of the manuscript.

#### APPENDIX

We use the following lemma in the proof of Theorem 4:

**Lemma 1.**  $\mathbf{X} := \sum_{k'} \sum_{l'} C_{k'l'} \mathbf{E}_{k'} \mathbf{F}_{l'} + \sum_{k'} \sum_{l'} C_{k'l'}^* \mathbf{F}_{l'} \mathbf{E}_{k'} \geq 0$

↑

The are complex  $U_{kl}$ ,  $V_{kl}$ ,  $\alpha_{kl}$  and real  $\beta_k$ ,  $\gamma_l$  such that

$$\begin{aligned} C_{kl} &= \alpha_{kl} + \beta_k + \gamma_l \\ U_{kl} V_{kl}^* &= \alpha_{kl} f_{kl} \\ 2\beta_k &\geq \sum_l |U_{kl}|^2 \\ 2\gamma_l &\geq \sum_k |V_{kl}|^2 \end{aligned} \tag{A1}$$

*Proof of Lemma 1.* Define the following vectors:

$$|c_{kl}\rangle := U_{kl} |x_k\rangle + V_{kl} |y_l\rangle$$

and the projector (up to a scalar factor):

$$\mathbf{N}_{kl} := |c_{kl}\rangle \langle c_{kl}| \geq 0$$

Explicit calculation using (A1) shows that

$$\mathbf{N}_{kl} = |U_{kl}|^2 \mathbf{E}_k + |V_{kl}|^2 \mathbf{F}_l + \alpha_{kl} \mathbf{E}_k \mathbf{F}_l + \alpha_{kl}^* \mathbf{F}_l \mathbf{E}_k$$

Moreover

$$\begin{aligned} \sum_k \sum_l \mathbf{N}_{kl} &\leq 2 \sum_k \beta_k \mathbf{E}_k + 2 \sum_l \gamma_l \mathbf{F}_l \\ &+ \sum_k \sum_l \alpha_{kl} \mathbf{E}_k \mathbf{F}_l + \sum_k \sum_l \alpha_{kl}^* \mathbf{F}_l \mathbf{E}_k = \mathbf{X} \quad \square \end{aligned}$$

*Proof of Theorem 4.* Consider the operator-valued measure  $\{\mathbf{M}_{kl}\}_{(k,l) \in L \times M}$ :

$$\mathbf{M}_{kl} := \sum_{k'} \sum_{l'} C_{k'l'}^{(kl)} \mathbf{E}_{k'} \mathbf{F}_{l'} + \sum_{k'} \sum_{l'} C_{k'l'}^{(kl)*} \mathbf{F}_{l'} \mathbf{E}_{k'} \tag{A2}$$

with  $C_{k'l'}^{(kl)}$  complex ( $k' \in K$ ,  $l' \in K$ ),

$$\sum_k \sum_l C_{k'l'}^{(kl)} = \frac{1}{2}, \quad \sum_l C_{k'l'}^{(kl)} = A_{k'}^k, \quad \sum_k C_{k'l'}^{(kl)} = B_l^l.$$

If each element of the operator-valued measure defined by (A2) satisfies the condition of Lemma 1, it is a POVM. It is associated with a joint nonideal

measurement of  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  as follows. Since  $\sum_k \sum_l \mathbf{M}_{kl} = 1$  is evident, from

$$\sum_{k \in L} \mathbf{M}_{kl} = \sum_{m \in K} \lambda_{lm}^{(y)} \mathbf{F}_m, \quad \sum_{l \in M} \mathbf{M}_{kl} = \sum_{m \in K} \lambda_{km}^{(x)} \mathbf{E}_m$$

with

$$\lambda_{lm}^{(y)} = 2 \operatorname{Re}(B_m^l)$$

$$\lambda_{lm}^{(x)} = 2 \operatorname{Re}(A_m^l)$$

Given that  $\{\mathbf{M}_{kl}\}_{(k,l) \in L \times M}$  is a POVM, the matrices  $\lambda_{kl}^{(x)}$  and  $\lambda_{kl}^{(y)}$  are non-ideality matrices because  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$  are PVM's. This leaves us with the task of providing a set of coefficients  $C_{k'l'}^{(kl)}$  that satisfies Lemma 1 and the assertions of the theorem.

The POVM  $\{\mathbf{M}_{kl}\}_{(k,l) \in L \times M}$  is nontrivial if  $C_{k'l'}^{(kl)}$  is not separable into a sum of the form  $\zeta_{k'}^{(kl)} + \xi_{l'}^{(kl)}$ . This requirement, as well as invertibility, is attainable by the following choice:

$$C_{k'l'}^{(kl)} = \varepsilon(2\delta_{[[k-k']]_0} \delta_{[[l-l']]_0} + \sqrt{F}\delta_{[[l-l']]_0} + \sqrt{F}\delta_{[[k-k']]_0}) \quad (\text{A3})$$

with

$$[[a]] := a \bmod n, \quad L = M = K$$

$$F := \max_{k,l} \{|f_{kl}|^2\}, \quad \varepsilon := (4 + 4n\sqrt{F})^{-1}$$

This example is in accord with Lemma 1, as we can see if we take

$$U_{k'l'}^{(kl)} = V_{k'l'}^{(kl)} = \sqrt{2\varepsilon} \delta_{[[k-k']]_0} \delta_{[[k-k']]_0} \delta_{[[l-l']]_0} \sqrt{|f_{k'l'}|} \exp(i\frac{1}{2}\theta_{k'l'})$$

$$\alpha_{k'l'}^{(kl)} = 2\varepsilon \delta_{[[k-k']]_0} \delta_{[[l-l']]_0}$$

$$\beta_{k'}^{(kl)} = \gamma_{k'}^{(kl)} = \varepsilon \sqrt{F} \delta_{[[k-k']]_0}$$

with

$$f_{kl} = |f_{kl}| \exp(i\theta_{kl})$$

and note that

$$\sqrt{F} \geq |f_{kl}|$$

For the nonideality matrices we have in this case

$$\lambda_{kl}^{(x)} = \lambda_{kl}^{(y)} = 2\varepsilon[(2 + n\sqrt{F})\delta_{kl} + \sqrt{F}] \quad (\text{A4})$$

Note that formula (A2) can easily be extended to the joint measurement of arbitrary noncoexistent POVM's. The analogous extension of (A1) might not be so easy. In any case, however, we see that positivity conditions like (A1) will prevent such a joint measurement from having arbitrarily small nonideality.

*Proof of Theorem 5.* Suppose  $\{\mathbf{E}_k\}_{k \in K} \rightarrow \{\mathbf{O}_l\}_{l \in L}$ , i.e.,  $\mathbf{O}_l = \sum_{k \in K} \lambda_{lk} \mathbf{E}_k$  with unique  $\lambda_{lk} \geq 0$  and  $\sum_{l \in L} \lambda_{lk} = 1$ . Then  $I$  reduces to

$$\begin{aligned} I_{\{\mathbf{E}_k\} \rightarrow \{\mathbf{O}_l\}} &= \sum_{l \in L} \sum_{k \in K} \lambda_{lk} \frac{1}{n} \log \left( \frac{\lambda_{lk}(1/n)}{[\sum_k \lambda_{lk}(1/n)](1/n)} \right) \\ &= \log(n) - \sum_{l \in L} \frac{\text{Tr}(\mathbf{O}_l)}{n} H_x(\mathbf{O}_l / \text{Tr}(\mathbf{O}_l)) \end{aligned}$$

with

$$H_x(\rho) := \sum_{k \in K} -\langle x_k | \rho | x_k \rangle \log(\langle x_k | \rho | x_k \rangle)$$

We have used the fact that  $\lambda_{lk} = \langle x_k | \mathbf{O}_l | x_k \rangle$ . Now we take  $\mathbf{O}_l = \sum_{m \in M} \mathbf{M}_{lm}$ ;  $\{\mathbf{M}_{lm}\}_{(l,m) \in L \times M}$  representing a joint nonideal measurement of  $\{\mathbf{E}_k\}_{k \in K}$  and  $\{\mathbf{F}_k\}_{k \in K}$ . The entropy functional  $H_x$  is concave, so that

$$I_{\{\mathbf{E}_k\} \rightarrow \{\sum_m \mathbf{M}_{lm}\}} \leq \log(n) - \sum_{l \in L} \frac{\text{Tr}(\mathbf{O}_l)}{n} \sum_{m \in M} \frac{\text{Tr}(\mathbf{M}_{lm})}{\text{Tr}(\mathbf{O}_l)} H_x(\mathbf{M}_{lm} / \text{Tr}(\mathbf{M}_{lm}))$$

Proceeding analogously for the other marginal, we get

$$\begin{aligned} &I_{\{\mathbf{E}_k\} \rightarrow \{\sum_m \mathbf{M}_{lm}\}} + I_{\{\mathbf{F}_k\} \rightarrow \{\sum_l \mathbf{M}_{lm}\}} \\ &\leq 2 \log(n) - \sum_{l \in L} \sum_{m \in M} (H_x(\mathbf{M}_{lm} / \text{Tr}(\mathbf{M}_{lm})) + H_y(\mathbf{M}_{lm} / \text{Tr}(\mathbf{M}_{lm}))) \frac{\text{Tr}(\mathbf{M}_{lm})}{n} \end{aligned}$$

The functional  $H_y$  is defined similarly to  $H_x$ . At this point we can use a slight generalization of a relation recently derived by Maassen and Uffink<sup>(14,23,28)</sup>:

$$H_x(\rho) + H_y(\rho) \geq \text{Tr}(\rho \mathbf{R}) \quad (\text{A5})$$

with

$$\mathbf{R} := -\sum_i \mathbf{P}_i \log(F_i)$$

$$\mathbf{P}_i := \sum_{k \in \text{block } \#i} |x_k\rangle \langle x_k|$$

Substituting this yields (5), since  $\sum_{l \in L} \sum_{m \in M} \mathbf{M}_{lm} = \mathbf{1}$ .  $\square$

Note that the bound of Theorem 5 is okay if the PVM's have common elements. As an example, we can consider a three-dimensional situation:

$$\{f_{kl}\} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ 0 & \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

In this case we will get the bound  $2 \log(3) - 2 \log(2)/3$ , well below  $2 \log(3)$ . If, however, we have an inner product matrix close to this one, but irreducible, our bound will be quite bad. This suggests that a better operator  $\mathbf{R}$  in (A5) can be found.

Results of numerical work along these lines for three- and four-dimensional spaces inspire the conjecture that a bound as low as

$$2 \log(n) + \frac{1}{n} \sum_{k \in K} \log(\max_{k' \in K} |f_{kk'}|^2)$$

is achievable for (5).

*Proof of Theorem 7.* This can be seen when one realizes that:

$$\begin{aligned} \exp\left(i \frac{2\pi}{n} [[\mathbf{X} \otimes \mathbf{1}' + \mathbf{1} \otimes \mathbf{X}']]\right) &= \exp\left(i \frac{2\pi}{n} (\mathbf{X} \otimes \mathbf{1}' + \mathbf{1} \otimes \mathbf{X}')\right) \\ &= \exp\left(i \frac{2\pi}{n} \mathbf{X}\right) \otimes \exp\left(i \frac{2\pi}{n} \mathbf{X}'\right) \end{aligned}$$

and that

$$\left[ \exp\left(i \frac{2\pi}{n} \mathbf{X}\right) \otimes \exp\left(i \frac{2\pi}{n} \mathbf{X}'\right), \exp\left(i \frac{2\pi}{n} \mathbf{Y}\right) \otimes \exp\left(i \frac{2\pi}{n} \mathbf{Y}'\right) \right]_- = \mathbf{0}$$

as follows from (7). □

*Proof of Theorem 8.* We can write<sup>(49)</sup>  $\rho_0 = \sum_a \sum_b r_{ab} \mathbf{S}_x^a \mathbf{S}_y^b$ . So

$$\begin{aligned} \text{Tr}(\mathbf{M}_{kl} \mathbf{S}_x^a \mathbf{S}_y^b) &= \frac{1}{n} \sum_{a'} \sum_{b'} r_{a'b'} \text{Tr}(\mathbf{S}_x^k \mathbf{S}_y^l \mathbf{S}_x^{a'} \mathbf{S}_y^{b'} (\mathbf{S}_x^k \mathbf{S}_y^l)^\dagger \mathbf{S}_x^{a'} \mathbf{S}_y^{b'}) \\ &= \exp\left(i \frac{2\pi}{n} (-al + kb - ab)\right) r_{[[ -a ]][[ -b ]]} \end{aligned}$$

Fourier transformation gives

$$\begin{aligned} & \sum_k \sum_l \exp\left(i \frac{2\pi}{n} (ck + dl)\right) \text{Tr}(\mathbf{M}_{kl} \mathbf{S}_x^a \mathbf{S}_y^b) \\ &= n^2 r_{[[ -d ]]c} \exp\left(i \frac{2\pi}{n} cd\right) \delta_{[[c+b]]0} \delta_{[[d-a]]0} \end{aligned}$$

For an arbitrary  $\rho$ , write:  $\rho = \sum_a \sum_b f_{ab} \mathbf{S}_x^a \mathbf{S}_y^b$  to get

$$\sum_k \sum_l \exp\left(i \frac{2\pi}{n} (ck + dl)\right) \text{Tr}(\mathbf{M}_{kl} \rho) = n^2 r_{[[ -d ]]c} f_{d[[ -c ]]} \exp\left(i \frac{2\pi}{n} cd\right)$$

Thus, if the distribution ( $\text{Tr}[\mathbf{M}_{kl} \rho]$ ) is known,  $f_{d[[ -c ]]}$  can be calculated in the Fourier domain iff  $r_{[[ -d ]]c} \neq 0$  for all  $d, c$ .  $\square$

The proof of Theorem 9 is similar.

## REFERENCES

1. O. Abu-Zeid, *Phys. Lett. A* **125**, 167 (1987).
2. E. Arthurs and J. Kelly, *Bell Syst. Tech. J.* **44**, 725 (1965).
3. L. Ballentine, *Rev. Mod. Phys.* **42**, 358 (1970).
4. V. Bargmann, as quoted by H. Reichenbach, *Philosophical Foundations of Quantum Mechanics* (University of California Press, Berkeley, 1944), p. 92.
5. A. Bohm, *Quantum Mechanics* (Springer, Berlin, 1979).
6. P. Busch, *Int. J. Theor. Phys.* **24**, 63 (1985).
7. P. Busch, *Phys. Rev. D* **33**, 2235 (1986).
8. P. Busch, *Found. Phys.* **17**, 905 (1987).
9. P. Busch and P. Lahti, *Phys. Rev. D* **29**, 1634 (1984).
10. P. Busch and P. Lahti, *Philos. Sc.* **32**, 809 (1985).
11. E. Davies, *J. Funct. Anal.* **6**, 318 (1970).
12. E. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).
13. E. Davies and J. Lewis, *Commun. Math. Phys.* **17**, 239 (1969).
14. D. Deutsch, *Phys. Rev. Lett.* **50**, 631 (1983).
15. P. Dirac, *The Principles of Quantum Mechanics* (Clarendon, Oxford, 1930).
16. R. Feynman, R. Leighton, and M. Sands, *Feynman Lectures on Physics*, Vol. 3 (Addison-Wesley, Reading, 1965).
17. C. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
18. C. Helstrom, *Found. Phys.* **4**, 453 (1974).
19. A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
20. M. Jammer, *The Philosophy of Quantum Mechanics* (Wiley, New York, 1974).
21. J. Jauch, *Foundations of Quantum Mechanics* (Addison-Wesley, Reading, 1968).
22. K. Kraus, *States, Effects and Operations* (Lecture Notes in Physics **190**) (Berlin, Springer, 1983).
23. K. Kraus, *Phys. Rev. D* **35**, 3070 (1987).

24. P. Kruszynski and W. de Muynck, *J. Math. Phys.* **28**, 1761 (1987).
25. V. Kuryshkin, *Int. J. Theor. Phys.* **7**, 451 (1973).
26. R. Loudon, *Quantum Theory of Light*, 2nd edn. (Clarendon, Oxford, 1983).
27. G. Ludwig, *Foundations of Quantum Mechanics*, Vol. I (Springer, Berlin, 1983).
28. H. Maassen and J. Uffink, *Phys. Rev. Lett.* **60**, 1103 (1988).
29. H. Martens and W. de Muynck, *Found. Phys.* **20**, 255 (1990).
30. H. Martens and W. de Muynck, "Towards a new type of uncertainty relation," submitted to *Phys. Lett. A*.
31. H. Martens and W. de Muynck, "The uncertainty principle," preprint, 1989, submitted to *Found. Phys.*
32. R. McEliece, *The Theory of Information and Coding* (Addison-Wesley, London, 1977).
33. E. McMullin, *The Principle of Uncertainty*, PhD Thesis, Catholic University of Louvain, Belgium, unpublished, 1954.
34. P. Mittelstädt, A. Prieur, and R. Schieder, *Found. Phys.* **17**, 891 (1987).
35. D. Murdoch, *Niels Bohr's Philosophy of Physics* (Cambridge University Press, Cambridge, 1987).
36. W. de Muynck and J. van den Eijnde, *Found. Phys.* **14**, 111 (1984).
37. W. de Muynck, P. Janssen, and A. Santman, *Found. Phys.* **9**, 71 (1979).
38. W. de Muynck and J. Koelman, *Phys. Lett. A* **98**, 1 (1983).
39. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932, 1982).
40. J. Park, W. Band, and W. Yourgrau, *Ann. Phys.* **37**, 189 (1980).
41. J. Park and H. Margenau, *Int. J. Theor. Phys.* **1**, 211 (1968).
42. S. Personick, *Bell Syst. Tech. J.* **50**, 213 (1971).
43. K. Popper, *Logic of Scientific Discovery*, 6th rev. impr. (Hutchison, London, 1972).
44. H. Primas, *Chemistry, Quantum Mechanics and Reductionism*, 2nd edn. (Springer, Berlin, 1983).
45. E. Prugovečki, *J. Phys. A* **10**, 543 (1977).
46. E. Prugovečki, *Stochastic Quantum Mechanics and Quantum Space-time* (Reidel, Dordrecht, 1984).
47. H. Robertson, *Phys. Rev.* **34**, 163 (1929).
48. F. Schroeck, *Found. Phys.* **12**, 825 (1982).
49. J. Schwinger, *Proc. Natl. Acad. Sci. USA* **46**, 570 (1960).
50. C. She and H. Heffner, *Phys. Rev.* **152**, 1103 (1966).
51. P. Suppes, *Philos. Sci.* **28**, 378 (1961).
52. A. Vogt, in: A. Marlow, ed., *Mathematical Foundations of Quantum Theory* (Academic, New York, 1978), p. 368.
53. Y. Yamamoto and H. Haus, *Rev. Mod. Phys.* **58**, 1001 (1986).
54. H. Yuen and J. Shapiro, *IEEE J. Inform. Theory* **IT-24**, 657 (1978); **IT-25**, 179 (1979); **IT-26**, 78 (1980).