

# Informationally complete joint measurements on finite quantum systems

Claudio Carmeli\*

*D.I.M.E., Università di Genova, Via Cadorna 2, I-17100 Savona, Italy*

Teiko Heinosaari†

*Department of Physics and Astronomy, Turku Centre for Quantum Physics, University of Turku, FI-20014 Turku, Finland*

Alessandro Toigo‡

*Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133 Milano, Italy and*

*I.N.F.N., Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*

(Received 22 November 2011; published 24 January 2012)

We show that there are informationally complete joint measurements of two conjugated observables on a finite quantum system, meaning that they enable the identification of all quantum states from their measurement outcome statistics. We further demonstrate that it is possible to implement a joint observable as a sequential measurement. If we require minimal noise in the joint measurement, then the joint observable is unique. If  $d$  is odd, then this observable is informationally complete. But if  $d$  is even, then the joint observable is not informationally complete, and one has to allow more noise in order to obtain informational completeness.

DOI: [10.1103/PhysRevA.85.012109](https://doi.org/10.1103/PhysRevA.85.012109)

PACS number(s): 03.65.Ta, 03.65.Aa, 03.67.—a

## I. INTRODUCTION

The general aim in quantum tomography is to identify quantum states from measurement outcome statistics. A collection of observables with this property is called *informationally complete* [1]. Even a single observable can be informationally complete, but then, it must be a noncommutative positive-operator-valued measure (POVM) [2,3]. We will study a class of informationally complete POVMs in dimension  $d$  with the minimal number of  $d^2$  outcomes, and we will explain how they can be implemented as sequential measurements of two  $d$ -outcome measurements. One can interpret the generated joint observable as a phase-space measurement in the discrete phase space  $\mathbb{Z}_d \times \mathbb{Z}_d$  [4,5].

There are some particularly interesting approaches to finite-dimensional quantum tomography, and one of them is based on complete collections of mutually unbiased bases (MUBs) [6–8]. In a  $d$ -dimensional Hilbert space, one needs  $d + 1$  MUBs in order to be able to identify all quantum states, but it is not known if a complete set of MUBs exists in all dimensions. In fact, there is evidence that, for  $d = 6$ , there is no complete set of MUBs [9,10].

In our scheme, we start from two mutually unbiased bases connected by the finite Fourier transform. They define a pair of complementary observables, which cannot be measured jointly. However, it is possible to realize their joint measurement if some additional noise is allowed. We show that their joint measurement can be chosen to be informationally complete and that this can be realized as a sequential measurement where we first perform a weak measurement in one basis and then another successive measurement in the other basis. Compared to the fact that one would need  $d + 1$  complementary observables in order to reach informational

completeness in separate measurements, it is remarkable that, in the sequential scheme, only two observables suffice.

The price to pay for a joint measurement is that the marginal observables are not the original complementary observables but their unsharp versions. We will analyze the required additional noise and will characterize the optimal joint observable from this point of view. The qubit case has been studied in Ref. [11], and our paper generalizes those results for an arbitrary finite dimension.

The covariant phase-space observables, i.e., POVMs covariant under the finite Weyl-Heisenberg group, play a special role in our investigation. We prove that, if a pair of conjugate observables has a joint measurement, then it also has a joint measurement, which is a covariant phase-space observable. Since every covariant phase-space observable arises from a sequential measurement of two conjugate observables [12], the covariant phase-space observables are an outstanding choice for finite-dimensional quantum tomography.

The Weyl-Heisenberg group also plays a pivotal role in the investigations of symmetric informationally complete (SIC) observables [13,14]. It generally is believed that a Weyl-Heisenberg covariant SIC observable exists in every finite dimension, and its existence is tested numerically in all dimensions up to 67 [15]. Our results show that any such observable has a neat sequential realization scheme.

There is an interesting difference between the even- and odd-dimensional Hilbert spaces. If we require minimal noise in both marginal observables, then their joint observable is unique. If  $d$  is odd, then this observable is informationally complete. But if  $d$  is even, then the joint observable is not informationally complete. This result gives an additional aspect to the common observation that quantum tomography is different in even and odd dimensions [16].

## II. PRELIMINARIES

In this section, we fix some notations and introduce the basic concepts.

\*claudio.carmeli@gmail.com

†teiko.heinosaari@utu.fi

‡alessandro.toigo@polimi.it

### A. States and observables

Let  $\mathcal{H}$  be a finite-dimensional Hilbert space with  $\dim \mathcal{H} = d \geq 2$ . We denote the vector space of all linear operators on  $\mathcal{H}$  by  $\mathcal{L}(\mathcal{H})$ . A positive operator  $\varrho \in \mathcal{L}(\mathcal{H})$  having trace 1 is a *state*, and we denote the set of all states by  $\mathcal{S}(\mathcal{H})$ .

Observables generally are described by POVMs [17,18]. In this paper, we only consider observables with a finite number of outcomes. Therefore, an observable can be defined as a function  $\mathbf{A}: x \mapsto \mathbf{A}(x)$ , where each  $\mathbf{A}(x)$  is a positive operator and  $\sum_x \mathbf{A}(x) = \mathbb{1}$ . Here, the sum runs over all  $x \in \Omega_{\mathbf{A}}$ , where set  $\Omega_{\mathbf{A}}$  is the collection of all possible measurement outcomes.

If a system is prepared in state  $\varrho$ , then a measurement of observable  $\mathbf{A}$  leads to an outcome  $x$  with the probability  $\text{tr}[\varrho \mathbf{A}(x)]$ .

### B. Informational completeness

Observable  $\mathbf{A}$  is *informationally complete* if its measurement outcome probability distribution is sufficient to identify a unique state [1]. In other words, two different states must give rise to different probability distributions: for all pairs of states  $\varrho_1, \varrho_2$ ,

$$\text{tr}[\varrho_1 \mathbf{A}(j)] = \text{tr}[\varrho_2 \mathbf{A}(j)] \quad \forall j \in \Omega_{\mathbf{A}} \quad \Rightarrow \quad \varrho_1 = \varrho_2.$$

The informational completeness of observable  $\mathbf{A}$  is equivalent to the property that the linear span of set  $\{\mathbf{A}(j): j \in \Omega_{\mathbf{A}}\}$  is  $\mathcal{L}(\mathcal{H})$  [3,19].

### C. Joint measurability

Given two observables  $\mathbf{A}$  and  $\mathbf{B}$ , we say that they are *jointly measurable* if there exists a third observable  $\mathbf{C}$  with  $\Omega_{\mathbf{C}} = \Omega_{\mathbf{A}} \times \Omega_{\mathbf{B}}$  and satisfying

$$\sum_{x \in \Omega_{\mathbf{A}}} \mathbf{C}(x, y) = \mathbf{B}(y) \quad \forall y, \quad \sum_{y \in \Omega_{\mathbf{B}}} \mathbf{C}(x, y) = \mathbf{A}(x) \quad \forall x.$$

In other words,  $\mathbf{A}$  and  $\mathbf{B}$  correspond to the marginals of  $\mathbf{C}$ . Any observable having  $\mathbf{A}$  and  $\mathbf{B}$  as its marginals is called a *joint observable* of  $\mathbf{A}$  and  $\mathbf{B}$  [20].

We recall that joint measurability is equivalent to the following [21]: There exists an observable  $\mathbf{G}$  and stochastic matrices  $[M_{xz}]$ ,  $[M'_{yz}]$  such that

$$\sum_z M_{xz} \mathbf{G}(z) = \mathbf{A}(x) \quad \forall x, \quad \sum_z M'_{yz} \mathbf{G}(z) = \mathbf{B}(y) \quad \forall y.$$

Hence, two observables are jointly measurable iff they can be postprocessed from a single observable.

We use the following simple fact several times: If  $\mathbf{C}$  and  $\mathbf{C}'$  are joint observables of  $\mathbf{A}$  and  $\mathbf{B}$ , then all their convex combinations  $t\mathbf{C} + (1-t)\mathbf{C}'$  (with  $0 < t < 1$ ) also are joint observables of  $\mathbf{A}$  and  $\mathbf{B}$ . It follows that two jointly measurable observables have either a unique joint observable or an infinite number of them.

Another useful fact is related to unitary transformations. Let  $U$  be a unitary operator on  $\mathcal{H}$ . Two observables  $\mathbf{A}$  and  $\mathbf{B}$  are jointly measurable if and only if the observables  $U\mathbf{A}U^*$  and  $U\mathbf{B}U^*$  are jointly measurable. Indeed, it is easy to see that  $\mathbf{C}$  is a joint observable of  $\mathbf{A}$  and  $\mathbf{B}$  if and only if  $UCU^*$  is a joint observable of  $U\mathbf{A}U^*$  and  $U\mathbf{B}U^*$ .

### D. Instruments

An observable describes the statistics of the outcomes of a measurement but leaves open how the measurement disturbs the input state. In order to discuss this, we need the concept of an *instrument* [22]. An instrument with finitely many outcomes is a mapping  $\mathcal{I}: x \mapsto \mathcal{I}_x$  such that each  $\mathcal{I}_x$  is a completely positive linear map on  $\mathcal{L}(\mathcal{H})$  and  $\sum_x \text{tr}[\mathcal{I}_x(\varrho)] = 1$  for all states  $\varrho$ .

The *adjoint map*  $\mathcal{I}_x^*$  of  $\mathcal{I}_x$  is defined via the usual trace duality,

$$\text{tr}[S\mathcal{I}_x(T)] = \text{tr}[\mathcal{I}_x^*(S)T] \quad \forall S, T \in \mathcal{L}(\mathcal{H}).$$

In other words,  $\mathcal{I}_x^*$  and  $\mathcal{I}_x$  correspond to the Heisenberg and Schrödinger pictures, respectively.

Suppose that  $\mathbf{A}$  is an observable. Then, we say that an instrument  $\mathcal{I}$  is  *$\mathbf{A}$ -compatible* if  $\mathcal{I}_x^*(\mathbb{1}) = \mathbf{A}(x)$  for every  $x$ . Every  $\mathbf{A}$ -compatible instrument describes some particular kind of measurement of  $\mathbf{A}$  [23].

An example of an  $\mathbf{A}$ -compatible instrument is the *Lüders instrument*  $\mathcal{I}^L$ , defined by

$$\mathcal{I}_x^L(\varrho) = \sqrt{\mathbf{A}(x)}\varrho\sqrt{\mathbf{A}(x)}.$$

Any other  $\mathbf{A}$ -compatible instrument  $\mathcal{I}$  is of the form

$$\mathcal{I}_x(\varrho) = \mathcal{E}_x(\mathcal{I}_x^L(\varrho))$$

for some collection  $\{\mathcal{E}_x\}$  of completely positive trace-preserving maps on  $\mathcal{L}(\mathcal{H})$  [24].

### E. Sequential measurements

By a *sequential measurement*, we mean a setting where two measurements are combined into a third measurement by performing them one after the other [25]. Generally, the order in which the measurements are performed is crucial [26].

Suppose we have two  $N$ -outcome observables  $\mathbf{A}, \mathbf{B}$ , and we measure them subsequently; first  $\mathbf{A}$  and then  $\mathbf{B}$ . As a result, we have, in total,  $N^2$  possible measurement outcomes. Generally, we do not obtain a joint measurement of  $\mathbf{A}$  and  $\mathbf{B}$  since the first measurement disturbs the input state. In fact, the overall measurement depends on the way we measure  $\mathbf{A}$ . If the first measurement is described by an  $\mathbf{A}$ -compatible instrument  $\mathcal{I}$ , then the overall observable  $\mathbf{C}$  is given by

$$\text{tr}[\varrho \mathbf{C}(j, k)] = \text{tr}[\mathbf{B}(k)\mathcal{I}_j(\varrho)]$$

for all input states  $\varrho$ , or equivalently,

$$\mathbf{C}(j, k) = \mathcal{I}_j^*(\mathbf{B}(k)).$$

Let us notice that the first marginal of  $\mathbf{C}$  is always  $\mathbf{A}$ , while the second marginal is a perturbed version of  $\mathbf{B}$  and depends on the instrument  $\mathcal{I}$ .

## III. EXAMPLE: SEQUENTIAL MEASUREMENTS OF $\sigma_x$ AND $\sigma_y$

We start with a preliminary example, which is mainly a collection of well-known facts. It hints about the forthcoming developments and clarifies the aims of the later sections. We refer to Ref. [27] for more details and further references.

Fix  $\mathcal{H} = \mathbb{C}^2$ , and let  $\mathbf{A}$  and  $\mathbf{B}$  be the two observables corresponding to the measurements of spin- $\frac{1}{2}$  components in directions  $x$  and  $y$ , respectively. Thus,

$$\mathbf{A}(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \sigma_x), \quad \mathbf{B}(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \sigma_y),$$

where  $\sigma_x, \sigma_y$  are the Pauli spin matrices.

Since  $\mathbf{A}$  and  $\mathbf{B}$  consist of projections and they do not commute mutually, it is not possible to measure them jointly. Moreover, if we measure them separately on two similarly prepared ensembles, we still cannot infer the unknown state.

An alternative way is to perform a sequential measurement. The first measurement has to be a weak measurement, meaning that we do not measure  $\mathbf{A}$  but its unsharp version. We define an unsharp version  $\mathbf{A}_\lambda$  of  $\mathbf{A}$  by

$$\mathbf{A}_\lambda(j) := \lambda \mathbf{A}(j) + (1 - \lambda) \frac{1}{2} \mathbb{1}, \quad j = \pm 1.$$

Here,  $\lambda \in [0, 1]$  is a parameter quantifying the noise or imprecision. We can write  $\mathbf{A}_\lambda$  in the form

$$\mathbf{A}_\lambda(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \lambda \sigma_x).$$

In a similar way, we define an unsharp version  $\mathbf{B}_\gamma$  of  $\mathbf{B}$  by

$$\mathbf{B}_\gamma(\pm 1) = \frac{1}{2}(\mathbb{1} \pm \gamma \sigma_y).$$

We want to study the disturbance of the first measurement on the system, and for this reason, we define an instrument related to  $\mathbf{A}_\lambda$ . A class of  $\mathbf{A}_\lambda$ -compatible instruments can be defined by

$$\mathcal{J}_{\pm 1}(\varrho) = L_{\pm 1} \sqrt{\mathbf{A}_\lambda(\pm 1) \varrho \sqrt{\mathbf{A}_\lambda(\pm 1)}} L_{\pm 1}^*,$$

where  $L_1, L_{-1}$  are arbitrary unitary operators. If the subsequent measurement is a measurement of  $\mathbf{B}$ , then the overall statistics of the sequential measurement is given by the observable

$$\mathbf{C}(j, k) = \mathcal{J}_j^*(\mathbf{B}(k)) = \sqrt{\mathbf{A}_\lambda(j)} L_j^* \mathbf{B}(k) L_j \sqrt{\mathbf{A}_\lambda(j)},$$

where  $j, k = \pm 1$ . The properties of  $\mathbf{C}$  obviously depend on  $L_1$  and  $L_{-1}$ . In the following, we consider two different choices of  $L_{\pm 1}$ .

#### A. Optimal joint measurement

If we choose  $L_{\pm 1} = \mathbb{1}$ , then we obtain

$$\mathbf{C}(j, k) = \frac{1}{4}(\mathbb{1} + j\lambda\sigma_x + k\sqrt{1 - \lambda^2}\sigma_y), \quad j, k = \pm 1.$$

In particular, the marginals are

$$\begin{aligned} \mathbf{C}(j, +1) + \mathbf{C}(j, -1) &= \mathbf{A}_\lambda(j), \\ \mathbf{C}(+1, k) + \mathbf{C}(-1, k) &= \mathbf{B}_{\sqrt{1 - \lambda^2}}(k). \end{aligned}$$

The joint observable  $\mathbf{C}$  is an optimal approximate joint measurement of  $\sigma_x$  and  $\sigma_y$ . This means that the unsharp parameters  $\lambda$  and  $\gamma = \sqrt{1 - \lambda^2}$  saturate the inequality

$$\lambda^2 + \gamma^2 \leq 1. \quad (1)$$

Indeed, it is known that this inequality is a necessary and sufficient criterion for two observables  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  to be jointly measurable [11]. Let us also notice that the joint observable of  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  is unique if  $\lambda^2 + \gamma^2 = 1$  [11].

#### B. Informationally complete joint measurement

Another interesting option is to choose

$$L_{\pm 1} = \cos \frac{\theta}{2} \mathbb{1} \mp i \sin \frac{\theta}{2} \sigma_x$$

for some fixed angle  $0 < \theta < \pi/2$ . In this case, we obtain

$$\begin{aligned} \mathbf{C}(j, k) &= \frac{1}{4}(\mathbb{1} + j\lambda\sigma_x + k\sqrt{1 - \lambda^2} \cos \theta \sigma_y \\ &\quad + jk\sqrt{1 - \lambda^2} \sin \theta \sigma_z), \quad j, k = \pm 1 \end{aligned}$$

and the marginals are

$$\begin{aligned} \mathbf{C}(j, +1) + \mathbf{C}(j, -1) &= \mathbf{A}_\lambda(j), \\ \mathbf{C}(+1, k) + \mathbf{C}(-1, k) &= \mathbf{B}_{\sqrt{1 - \lambda^2} \cos \theta}(k). \end{aligned}$$

It is easy to see that the linear span of the four operators  $\mathbf{C}(j, k)$ ,  $j, k = \pm 1$ , is the set of all  $2 \times 2$ -complex matrices. It follows that joint observable  $\mathbf{C}$  is informationally complete.

The unsharpness parameters  $\lambda$  and  $\gamma = \sqrt{1 - \lambda^2} \cos \theta$  do not saturate the inequality (1). By altering the parameter  $\theta$ , we can make the sum  $\lambda^2 + \gamma^2$  as close to 1 as we want, hence, we conclude that  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  admit an informationally complete joint observable if and only if

$$\lambda^2 + \gamma^2 < 1.$$

Finally, we remark that, with choices  $\lambda = 1/\sqrt{3}$  and  $\theta = \pi/4$ , the joint observable  $\mathbf{C}$  is a SIC observable.

### IV. CONJUGATE OBSERVABLES

#### A. Mutually unbiased bases and complementary observables

We start by recalling the usual definition of complementary observables in a finite  $d$ -dimensional Hilbert space and some related basic facts [28,29]. We denote  $\mathbb{Z}_d \equiv \{0, \dots, d-1\}$ . Let  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$  be MUBs, i.e., they are orthonormal bases in  $\mathcal{H}$  and

$$|\langle \varphi_j | \psi_k \rangle|^2 = 1/d \quad \forall j, k \in \mathbb{Z}_d. \quad (2)$$

We define two  $d$ -outcome observables  $\mathbf{A}$  and  $\mathbf{B}$  corresponding to  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$ , respectively. Hence,

$$\mathbf{A}(j) = |\varphi_j\rangle\langle\varphi_j|, \quad \mathbf{B}(k) = |\psi_k\rangle\langle\psi_k|.$$

Obviously, two orthonormal bases  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\varphi'_j\}_{j \in \mathbb{Z}_d}$  define the same observable  $\mathbf{A}$  iff  $\varphi'_j = \alpha_j \varphi_j$  for some complex numbers  $\alpha_j$  with  $|\alpha_j| = 1$ .

The mutual unbiasedness condition (2) can be rephrased by saying that  $\mathbf{A}$  and  $\mathbf{B}$  are *complementary* observables, meaning that, in any state  $\varrho$  where the outcome of  $\mathbf{A}$  is predictable, the  $\mathbf{B}$  distribution is uniform (and vice versa). This entails that the following implications are valid for any state  $\varrho$  and all outcomes  $j, k \in \mathbb{Z}_d$ :

$$\begin{aligned} \text{tr}[\varrho \mathbf{A}(j)] = 1 &\Rightarrow \text{tr}[\varrho \mathbf{B}(k)] = 1/d, \\ \text{tr}[\varrho \mathbf{B}(k)] = 1 &\Rightarrow \text{tr}[\varrho \mathbf{A}(j)] = 1/d. \end{aligned}$$

Since  $\text{tr}[\varrho \mathbf{A}(j)] = 1$  iff  $\varrho = |\varphi_j\rangle\langle\varphi_j|$ , it is easy to see that the complementarity of  $\mathbf{A}$  and  $\mathbf{B}$  indeed is equivalent to the mutual unbiasedness of the bases  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$ .

There is a canonical way to produce two mutually unbiased bases. In the following, suppose an orthonormal basis  $\{\varphi_k\}_{k \in \mathbb{Z}_d}$

of  $\mathcal{H}$  is fixed. By denoting  $\omega \equiv e^{2\pi i/d}$ , we define the following unitary representations  $U$  and  $V$  of the cyclic group  $\mathbb{Z}_d$  in  $\mathcal{H}$ :

$$U_x \varphi_k := \varphi_{k+x}, \quad V_y \varphi_k := \omega^{yk} \varphi_k$$

for all  $x, y, k \in \mathbb{Z}_d$ . In the above formulas and in the rest of the paper, addition and multiplication of elements in  $\mathbb{Z}_d$  are understood mod  $d$ . (For instance, we will often use  $-j = d - j$ .) It is easy to verify that

$$V_y U_x = \omega^{xy} U_x V_y \quad \forall x, y \in \mathbb{Z}_d. \quad (3)$$

The Fourier transform (with respect to the basis  $\{\varphi_k\}_{k \in \mathbb{Z}_d}$ ) is the unitary operator  $\mathcal{F}: \mathcal{H} \rightarrow \mathcal{H}$  defined by

$$\mathcal{F} \varphi_k := \frac{1}{\sqrt{d}} \sum_{h \in \mathbb{Z}_d} \omega^{-hk} \varphi_h. \quad (4)$$

The adjoint operator  $\mathcal{F}^*$  of  $\mathcal{F}$  is given by

$$\mathcal{F}^* \varphi_k = \frac{1}{\sqrt{d}} \sum_{h \in \mathbb{Z}_d} \omega^{hk} \varphi_h = \mathcal{F} \varphi_{-k},$$

and we have  $\mathcal{F}^2 \varphi_k = \mathcal{F}^{*2} \varphi_k = \varphi_{-k}$ . We denote

$$\psi_k := \mathcal{F}^* \varphi_k = \mathcal{F} \varphi_{-k},$$

and it is immediate to check that  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$  are MUBs with  $\langle \varphi_j | \psi_k \rangle = (1/\sqrt{d}) \omega^{jk}$ .

The Fourier transform has the intertwining properties

$$\mathcal{F} U_x = V_x^* \mathcal{F}, \quad \mathcal{F} V_y = U_y \mathcal{F},$$

from which it follows that

$$U_x \psi_k = \omega^{-xk} \psi_k, \quad V_y \psi_k = \psi_{k+y}.$$

The observables  $\mathbf{A}$  and  $\mathbf{B}$  related to  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$ , respectively, satisfy the following conditions for all  $j, k, x, y \in \mathbb{Z}_d$ :

$$U_x \mathbf{A}(j) U_x^* = \mathbf{A}(j+x), \quad V_y \mathbf{A}(j) V_y^* = \mathbf{A}(j), \quad (5)$$

and

$$U_x \mathbf{B}(k) U_x^* = \mathbf{B}(k), \quad V_y \mathbf{B}(k) V_y^* = \mathbf{B}(k+y). \quad (6)$$

In other words,  $\mathbf{A}$  is  $U$  covariant and  $V$  invariant, while  $\mathbf{B}$  is  $U$  invariant and  $V$  covariant. We also note that  $\mathbf{A}$  and  $\mathbf{B}$  are conjugated by  $\mathcal{F}$ , i.e.,

$$\mathbf{B}(k) = \mathcal{F}^* \mathbf{A}(k) \mathcal{F}$$

for all  $k \in \mathbb{Z}_d$ . It is customary to say that  $\mathbf{A}$  and  $\mathbf{B}$  are *canonically conjugated* observables.

The conditions (5) and (6) are analogous to the symmetry properties of the usual position and momentum observables on the real line  $\mathbb{R}$  (see, e.g., Ref. [30]). In some situations, the covariance properties may have some physical meaning or motivation. However, for our purposes, they are just useful features that can be utilized later in our calculations.

### B. Unsharp observables

Measurements of two complementary observables are incompatible and, therefore, have to be performed separately. This means that their measurements require different settings. However, it is possible to perform a simultaneous measurement

of two complementary observables if we allow some additional imprecision or noise. In other words, we can measure jointly *unsharp versions* of  $\mathbf{A}$  and  $\mathbf{B}$ .

We define an unsharp version  $\mathbf{A}_\lambda$  of  $\mathbf{A}$  by

$$\mathbf{A}_\lambda(j) := \lambda \mathbf{A}(j) + (1-\lambda) \frac{1}{d} \mathbb{1}.$$

Here,  $\lambda \in [0, 1]$  is a parameter quantifying the noise. This type of noise is equivalent to the situation where an input state  $\rho$  is first depolarized into a state  $\lambda \rho + ((1-\lambda)/d) \mathbb{1}$  and then, a measurement of  $\mathbf{A}$  is performed.

More generally, if  $\Lambda$  is a probability distribution on  $\mathbb{Z}_d$ , then, we define an unsharp version  $\mathbf{A}_\Lambda$  of  $\mathbf{A}$  by

$$\mathbf{A}_\Lambda(j) := \sum_{i \in \mathbb{Z}_d} \Lambda(j-i) \mathbf{A}(i).$$

The special case  $\mathbf{A}_\Lambda = \mathbf{A}_\lambda$  corresponds to the probability distribution  $\Lambda$  defined as

$$\Lambda(0) = \lambda + (1-\lambda)/d, \quad \Lambda(j) = (1-\lambda)/d, \quad \text{if } j \neq 0.$$

We can also write this probability distribution in the form

$$\Lambda(j) = \lambda \delta(j) + (1-\lambda) \mu(j),$$

where  $\delta$  is the point distribution at 0 and  $\mu$  is the uniform distribution on  $\mathbb{Z}_d$ , i.e.,

$$\delta(j) = \begin{cases} 1, & \text{if } j = 0 \\ 0, & \text{if } j \neq 0 \end{cases}, \quad \mu(j) = \frac{1}{d} \quad \forall j. \quad (7)$$

In a similar way, a probability distribution  $\Gamma$  on  $\mathbb{Z}_d$  defines an unsharp version  $\mathbf{B}_\Gamma$  of  $\mathbf{B}$  by

$$\mathbf{B}_\Gamma(k) := \sum_{i \in \mathbb{Z}_d} \Gamma(k-i) \mathbf{B}(i).$$

Again, a special class is characterized by noise parameters  $\gamma \in [0, 1]$ , and we denote

$$\mathbf{B}_\gamma(k) := \gamma \mathbf{B}(k) + (1-\gamma) \frac{1}{d} \mathbb{1}.$$

Naturally, there also are other types of approximations of  $\mathbf{A}$  and  $\mathbf{B}$  than the previously defined  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$ . The usefulness of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  is that they satisfy the same covariance and invariance relations as  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. Namely, the observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  satisfy the following conditions:

$$U_x \mathbf{A}_\Lambda(j) U_x^* = \mathbf{A}_\Lambda(j+x), \quad V_y \mathbf{A}_\Lambda(j) V_y^* = \mathbf{A}_\Lambda(j), \quad (8)$$

and

$$U_x \mathbf{B}_\Gamma(k) U_x^* = \mathbf{B}_\Gamma(k), \quad V_y \mathbf{B}_\Gamma(k) V_y^* = \mathbf{B}_\Gamma(k+y). \quad (9)$$

Thus,  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are *conjugated observables*, although they need not be complementary anymore [12]. As we will see, two observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  can have a joint observable even if they do not commute.

*Remark 1.* Suppose that  $\tilde{\mathbf{A}}$  is a  $d$ -outcome observable satisfying

$$U_x \tilde{\mathbf{A}}(j) U_x^* = \tilde{\mathbf{A}}(j+x), \quad V_y \tilde{\mathbf{A}}(j) V_y^* = \tilde{\mathbf{A}}(j) \quad (10)$$

for all  $j, x, y \in \mathbb{Z}_d$ . Then,  $\tilde{\mathbf{A}} = \mathbf{A}_\Lambda$  for some probability distribution  $\Lambda$ . Namely, it follows from the second condition in Eq. (10) that  $\tilde{\mathbf{A}}$  commutes with  $\mathbf{A}$  [since  $\mathbf{A}(j) =$

$(1/d) \sum_y \omega^{-jy} V_y]$  and hence,  $\tilde{\mathbf{A}}(j) = \sum_k p_{j,k} \mathbf{A}(k)$  for some real numbers  $0 \leq p_{j,k} \leq 1$ . The first condition in Eq. (10) then implies that  $p_{j,k} = p_{0,k-j}$ .

## V. JOINT MEASUREMENTS

### A. Covariant observables

We recall that two observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable if they have a joint observable, i.e., an observable  $\mathbf{C}$  on  $\mathbb{Z}_d \times \mathbb{Z}_d$  such that

$$\sum_{k \in \mathbb{Z}_d} \mathbf{C}(j,k) = \mathbf{A}_\Lambda(j), \quad \sum_{j \in \mathbb{Z}_d} \mathbf{C}(j,k) = \mathbf{B}_\Gamma(k)$$

for all  $j, k \in \mathbb{Z}_d$ . A special class of joint observables turns out to be crucial for our developments. We say that an observable  $\mathbf{C}$  on  $\mathbb{Z}_d \times \mathbb{Z}_d$  is a *covariant phase-space observable* if

$$U_x V_y \mathbf{C}(j,k) V_y^* U_x^* = \mathbf{C}(j+x, k+y)$$

for all  $j, k, x, y \in \mathbb{Z}_d$ . The covariant phase-space observables have a simple form [22]. Namely, if  $\mathbf{C}$  is a covariant phase-space observable, then there is a unique operator  $T \in \mathcal{S}(\mathcal{H})$  such that  $\mathbf{C} = \mathbf{C}_T$ , where we have denoted

$$\mathbf{C}_T(j,k) := \frac{1}{d} U_j V_k T V_k^* U_j^*, \quad j, k \in \mathbb{Z}_d.$$

Also, each  $T \in \mathcal{S}(\mathcal{H})$  defines a covariant phase-space observable by this formula. Therefore, the correspondence  $T \leftrightarrow \mathbf{C}_T$  is one to one, and the elements in  $\mathcal{S}(\mathcal{H})$  parametrize the covariant phase-space observables.

The marginals of a covariant phase-space observable  $\mathbf{C}_T$  are conjugated observables on  $\mathbb{Z}_d$ . Indeed, a direct calculation shows that the marginals  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  of  $\mathbf{C}_T$  have probability densities

$$\Lambda(j) = \text{tr}[\mathbf{A}(-j)T], \quad \Gamma(k) = \text{tr}[\mathbf{B}(-k)T]. \quad (11)$$

(This calculation can be found in Ref. [12].)

The essential role of covariant phase-space observables in our discussion becomes clear in the following observation.

*Proposition 1.* If  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable, then they have a joint observable that is a covariant phase-space observable.

*Proof.* Suppose that  $\mathbf{C}$  is a joint observable of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$ . For each  $x, y \in \mathbb{Z}_d$ , we define observable  $\mathbf{C}_{x,y}$  by

$$\mathbf{C}_{x,y}(j,k) := U_x^* V_y^* \mathbf{C}(j+x, k+y) V_y U_x. \quad (12)$$

Using the covariance and invariance properties (8) and (9), it is straightforward to verify that  $\mathbf{C}_{x,y}$  is a joint observable of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$ .

We then define  $\tilde{\mathbf{C}}$  to be the uniform mixture of all  $\mathbf{C}_{x,y}$ , i.e.,

$$\tilde{\mathbf{C}}(j,k) := \frac{1}{d^2} \sum_{x,y \in \mathbb{Z}_d} \mathbf{C}_{x,y}(j,k). \quad (13)$$

Since every  $\mathbf{C}_{x,y}$  is a joint observable of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$ ,  $\tilde{\mathbf{C}}$  also is their joint observable. A direct calculation, using Eq. (3), shows that  $\tilde{\mathbf{C}}$  is a covariant phase-space observable. ■

From Proposition 1, we conclude that two observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable iff their related probability distributions  $\Lambda$  and  $\Gamma$  are of the form (11) for some  $T \in \mathcal{S}(\mathcal{H})$ . Equations (11) can be rewritten in a slightly different form.

Namely, observe that, if  $T \in \mathcal{S}(\mathcal{H})$ , then there exists a unit vector  $\phi \in \mathcal{H} \otimes \mathcal{H}$  such that  $T = \text{tr}_2[|\phi\rangle\langle\phi|]$ , where  $\text{tr}_2$  is the partial trace with respect to the second factor. (A vector state giving  $T$  via the partial trace often is called a *purification* of  $T$ .)

Conversely, if  $\phi \in \mathcal{H} \otimes \mathcal{H}$  is a unit vector, then  $T = \text{tr}_2[|\phi\rangle\langle\phi|]$  is a state. Inserting this form into Eq. (11), we obtain

$$\begin{aligned} \Lambda(j) &= \langle\phi|[\mathbf{A}(-j) \otimes \mathbb{1}]\phi\rangle, \\ \Gamma(k) &= \langle\phi|[\mathbf{B}(-k) \otimes \mathbb{1}]\phi\rangle. \end{aligned} \quad (14)$$

Note that, if a vector  $\phi \in \mathcal{H} \otimes \mathcal{H}$  satisfies the above two equations for some probability densities  $\Lambda$  and  $\Gamma$ , then the normalization  $\|\phi\| = 1$  is automatic. We, thus, have the following characterization of jointly measurable observables.

*Proposition 2.* Let  $\Lambda, \Gamma$  be probability densities on  $\mathbb{Z}_d$ . The following facts are equivalent:

- (i) The observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable.
- (ii) There exists a state  $T \in \mathcal{S}(\mathcal{H})$  such that the probability densities  $\Lambda$  and  $\Gamma$  satisfy Eq. (11).

(iii) There exists a vector  $\phi \in \mathcal{H} \otimes \mathcal{H}$  such that the probability densities  $\Lambda$  and  $\Gamma$  satisfy Eq. (14).

Let us note that any  $\mathbf{A}_\Lambda$  is jointly measurable with some  $\mathbf{B}_\Gamma$ . Namely, for each  $\Lambda$ , we can define a state  $T_\Lambda$  as

$$T_\Lambda := \sum_{j \in \mathbb{Z}_d} \Lambda(-j) |\varphi_j\rangle\langle\varphi_j|.$$

Then,  $\Lambda(j) = \text{tr}[\mathbf{A}(-j)T_\Lambda]$ , and  $\mathbf{A}_\Lambda$ , thus, is jointly measurable with  $\mathbf{B}_\Gamma$ , where  $\Gamma$  is defined as  $\Gamma(k) = \text{tr}[\mathbf{B}(-k)T_\Lambda]$ .

Proposition 2 can be seen as a trade-off relation between the probability distributions  $\Lambda$  and  $\Gamma$  that describe the deviations of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  from  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. For instance, if  $\Lambda = \delta$ , then necessarily  $\Gamma = \mu$ . Hence, we recover the fact that  $\mathbf{A}$  is jointly measurable only with the trivial observable and no other  $\mathbf{B}_\Gamma$ .

We end this subsection with some additional observations.

*Remark 2.* If two observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable, they can have several different covariant phase-space observables as their joint observables.

For instance, let  $\{\zeta_i\}_{i \in \mathbb{Z}_d}$  be an orthonormal basis that is mutually unbiased with respect to both orthonormal bases  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$ . Then, for each  $i \in \mathbb{Z}_d$ , we have

$$\langle\zeta_i|\mathbf{A}(-j)\zeta_i\rangle = \langle\zeta_i|\mathbf{B}(-k)\zeta_i\rangle = 1/d.$$

Therefore, the marginals of the covariant phase-space observables  $\mathbf{C}_{|\zeta_i\rangle\langle\zeta_i|}$  are the same, although  $\mathbf{C}_{|\zeta_i\rangle\langle\zeta_i|} \neq \mathbf{C}_{|\zeta_{i'}\rangle\langle\zeta_{i'}|}$  whenever  $i \neq i'$ .

*Remark 3.* If two observables  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable, they can have a joint observable that is not a covariant phase-space observable.

For instance, let  $p: \mathbb{Z}_d \times \mathbb{Z}_d \rightarrow [0, 1]$  be a bivariate probability distribution with uniform marginals. Then, the observable  $\mathbf{C}(j,k) := p(j,k)\mathbb{1}$  is a joint observable of  $\mathbf{A}_0$  and  $\mathbf{B}_0$ . It is clear that  $\mathbf{C}$  is a covariant phase-space observable only if  $p$  is a uniform distribution. However, a bivariate probability distribution with uniform marginals need not be uniform. For instance, if we set

$$p(i,j) = \frac{1}{d^2} \left[ 1 - \sin\left(2\pi \frac{ij}{d}\right) \right], \quad i, j \in \mathbb{Z}_d,$$

then  $\sum_i p(i,j) = \sum_j p(i,j) = 1/d$ , but  $p$  is not uniform.

The existence of noncovariant joint observables is not limited to the trivial observables  $A_0$  and  $B_0$ . Namely, suppose that  $C$  is a joint observable of  $A_0$  and  $B_0$  and  $C'$  is a joint observable of  $A_\lambda$  and  $B_\gamma$ . Then, the convex combination  $tC' + (1-t)C$ ,  $0 < t < 1$ , is a joint observable of  $A_{t\lambda}$  and  $B_{t\gamma}$ . It is easy to see that, if  $C'$  is a covariant phase-space observable but  $C$  is not, then their convex combination  $tC' + (1-t)C$  cannot be a covariant phase-space observable.

*Remark 4.* Suppose that  $A_\lambda$  and  $B_\gamma$  have a unique joint observable  $C_T$  among the covariant phase-space observables and that  $T^2 = T$ . Then,  $C_T$  is a unique joint observable of  $A_\lambda$  and  $B_\gamma$ .

To prove this claim, let  $C$  be a joint observable of  $A_\lambda$  and  $B_\gamma$ . We need to show that  $C = C_T$ . We define  $C_{x,y}$  and  $\tilde{C}$  as in Eqs. (12) and (13). Since, by its construction,  $\tilde{C}$  is a covariant phase-space observable, we must have  $\tilde{C} = C_T$  by the assumption on uniqueness. In particular, each operator  $\tilde{C}(j,k) = (1/d)U_j V_k T V_k^* U_j^*$  is rank 1. Since  $0 \leq C_{x,y}(0,0) \leq d^2 \tilde{C}(0,0) = dT$  by Eq. (13), it follows that there exists a real constant  $0 \leq c(x,y) \leq d$  such that  $C_{x,y}(0,0) = c(x,y)T$ , hence,

$$C(x,y) = c(x,y) U_x V_y T V_y^* U_x^*$$

by Eq. (12). Suppose  $\lambda \neq 1$ . Since  $C$  has  $A_\lambda$  as its first marginal, then for all  $x \in \mathbb{Z}_d$ , we obtain

$$\sum_{y \in \mathbb{Z}_d} c(x,y) U_x V_y T V_y^* U_x^* = \lambda A(x) + (1-\lambda) \frac{1}{d} \mathbb{1}.$$

The right-hand side of this equation is a rank- $d$  operator, while on the left-hand side, we have the sum of  $d$  operators with rank 1. It then follows that set  $\{U_x V_y T V_y^* U_x^*\}_{y \in \mathbb{Z}_d}$  is linearly independent in  $\mathcal{L}(\mathcal{H})$ . Since  $C$  and  $\tilde{C}$  have the same marginals,

$$\sum_{y \in \mathbb{Z}_d} c(x,y) U_x V_y T V_y^* U_x^* = \frac{1}{d} \sum_{y \in \mathbb{Z}_d} U_x V_y T V_y^* U_x^*$$

for all  $x$ , hence,  $c(x,y) = 1/d$  for all  $x,y$  by linear independence. The case  $\lambda = 1$  is treated in a similar way by taking the marginal  $B_\gamma$  in the place of  $A_\lambda$  (and now, necessarily,  $\gamma \neq 1$ ). Therefore,  $C_T$  is the unique joint observable of  $A_\lambda$  and  $B_\gamma$ .

### B. Unsharpness inequality

In this subsection, we apply Proposition 2 to the cases where  $A_\lambda = A_\lambda$  and  $B_\gamma = B_\gamma$ . These special types of marginal observables are interesting as we can quantify their unsharpnesses by single numbers  $\lambda$  and  $\gamma$ . In particular, we can ask how small  $\lambda$  and  $\gamma$  must be in order for  $A_\lambda$  and  $B_\gamma$  to become jointly measurable. We first notice that this question is, indeed, meaningful.

*Proposition 3.* Let  $\lambda, \gamma \in (0,1)$ . The following conditions are equivalent:

- (i)  $A_\lambda$  and  $B_\gamma$  are jointly measurable.
- (ii)  $A_{\lambda'}$  and  $B_{\gamma'}$  are jointly measurable for all  $0 \leq \lambda' \leq \lambda$  and  $0 \leq \gamma' \leq \gamma$ .
- (iii)  $A_{\lambda'}$  and  $B_{\gamma'}$  are jointly measurable for all  $0 \leq \lambda' < \lambda$  and  $0 \leq \gamma' < \gamma$ .

*Proof.* Suppose that (i) holds and  $0 \leq \gamma' \leq \gamma$ . We denote  $t := \gamma'/\gamma$  and, hence,  $0 \leq t \leq 1$ . We have

$$B_{\gamma'}(k) = tB_\gamma(k) + (1-t) \frac{1}{d} \mathbb{1},$$

meaning that  $B_{\gamma'}$  is a convex combination of  $B_\gamma$  and the trivial observable  $B_0$ . By the assumption,  $A_\lambda$  is jointly measurable with  $B_\gamma$ , and  $A_\lambda$  is jointly measurable with the trivial observable  $B_0$  (since they commute). If  $C_1$  is a joint observable of  $A_\lambda$  and  $B_\gamma$  and  $C_2$  is a joint observable of  $A_\lambda$  and  $B_0$ , then the convex combination  $tC_1 + (1-t)C_2$  is a joint observable of  $A_\lambda$  and  $B_{\gamma'}$ . Therefore,  $A_\lambda$  and  $B_{\gamma'}$  are jointly measurable. We can interchange the roles of  $A_\lambda$  and  $B_\gamma$  and run the same argument, hence, we obtain (ii).

It is clear that (ii) implies (iii). Hence, to complete the proof, we need to show that (iii) implies (i).

Suppose that (iii) holds. We choose sequences  $(\lambda_n)$  and  $(\gamma_n)$  such that  $0 < \lambda_n < \lambda$ ,  $0 < \gamma_n < \gamma$ , and  $\lim_n \lambda_n = \lambda$ ,  $\lim_n \gamma_n = \gamma$ . For each  $n$ , we fix a state  $T_n$  such that the corresponding covariant phase-space observable  $C_{T_n}$  is a joint observable of  $A_{\lambda_n}$  and  $B_{\gamma_n}$ . The set of states  $\mathcal{S}(\mathcal{H})$  is compact in the operator norm topology, hence, the sequence  $(T_n)$  has a convergent subsequence. By  $T$ , we denote the limit of this convergent subsequence. By passing to the limit in Eq. (11), we see that the covariant phase-space observable  $C_T$  is a joint observable of  $A_\lambda$  and  $B_\gamma$ . Thus, (i) holds. ■

From Proposition 3, we conclude that, for every  $\lambda \in (0,1)$ , there is a number  $\gamma_{\max}(\lambda) \geq 0$  such that  $A_\lambda$  and  $B_\gamma$  are jointly measurable iff  $0 \leq \gamma \leq \gamma_{\max}(\lambda)$ . Similarly, for every  $\gamma \in (0,1)$ , there is a number  $\lambda_{\max}(\gamma) \geq 0$  such that  $A_\lambda$  and  $B_\gamma$  are jointly measurable iff  $0 \leq \lambda \leq \lambda_{\max}(\gamma)$ .

We also know that  $\gamma_{\max}(0) = \lambda_{\max}(0) = 1$  (since a trivial observable is jointly measurable with any other observable) and that  $\gamma_{\max}(1) = \lambda_{\max}(1) = 0$  (see the discussion after Proposition 2).

*Proposition 4.* The equality  $\lambda_{\max}(x) = \gamma_{\max}(x)$  holds for all  $x \in [0,1]$ .

*Proof.* It is enough to show that  $A_\lambda$  and  $B_\gamma$  are jointly measurable if and only if  $A_\gamma$  and  $B_\lambda$  are such. Joint measurability of  $A_\lambda$  and  $B_\gamma$  means that there exists an observable  $C$  on  $\mathbb{Z}_d \times \mathbb{Z}_d$  having marginals  $A_\lambda$  and  $B_\gamma$ , respectively. We set  $\hat{C}(j,k) := \mathcal{F}^* C(k, -j) \mathcal{F}$  for every  $j,k \in \mathbb{Z}_d$ . Then,

$$\begin{aligned} \sum_{k \in \mathbb{Z}_d} \hat{C}(j,k) &= \sum_{k \in \mathbb{Z}_d} \mathcal{F}^* C(k, -j) \mathcal{F} = \mathcal{F}^* B_\gamma(-j) \mathcal{F} \\ &= \mathcal{F}^{*2} A_\gamma(-j) \mathcal{F}^2 = A_\gamma(j), \end{aligned}$$

$$\begin{aligned} \sum_{j \in \mathbb{Z}_d} \hat{C}(j,k) &= \sum_{j \in \mathbb{Z}_d} \mathcal{F}^* C(k, -j) \mathcal{F} = \mathcal{F}^* A_\lambda(k) \mathcal{F} \\ &= B_\lambda(k). \end{aligned}$$

We conclude that  $\hat{C}$  is a joint observable of  $A_\gamma$  and  $B_\lambda$ , hence, the latter two are jointly measurable. ■

We will now find out the function  $\gamma_{\max}(\lambda)$ , or, equivalently,  $\lambda_{\max}(\gamma)$ . Suppose that  $\lambda, \gamma \in [0,1]$  are such that  $A_\lambda$  and  $B_\gamma$  are jointly measurable. By Proposition 2, this is equivalent to the existence of a vector  $\phi \in \mathcal{H} \otimes \mathcal{H}$  satisfying

$$\langle \phi | [A(j) \otimes \mathbb{1}] \phi \rangle = \lambda \delta(j) + (1-\lambda) \mu(j), \quad (15)$$

$$\langle \phi | [B(k) \otimes \mathbb{1}] \phi \rangle = \gamma \delta(k) + (1-\gamma) \mu(k) \quad (16)$$

for all  $j, k \in \mathbb{Z}_d$  [see Eq. (7) for the definition of  $\delta$  and  $\mu$ ]. We now give a condition on the parameters  $\lambda, \gamma$ , which is necessary and sufficient for the existence of a vector  $\phi \in \mathcal{H} \otimes \mathcal{H}$  satisfying the above two equations. Moreover, we show that, for the extreme values of  $\lambda, \gamma$ , the vector  $\phi$  is essentially unique.

*Lemma 1.* Let  $\lambda, \gamma \in [0, 1]$ . Suppose there exists a vector  $\phi \in \mathcal{H} \otimes \mathcal{H}$  satisfying Eqs. (15) and (16). Then,

$$\gamma \leq \frac{(d-2)(1-\lambda) + 2\sqrt{(1-d)\lambda^2 + (d-2)\lambda + 1}}{d}. \quad (17)$$

For any choice of a unit vector  $\eta \in \mathcal{H}$ , the vector

$$\phi = (\alpha_\lambda \varphi_0 + \beta_\lambda \psi_0) \otimes \eta, \quad (18)$$

with

$$\alpha_\lambda = \frac{1}{\sqrt{d}} [\sqrt{(d-1)\lambda + 1} - \sqrt{1-\lambda}], \quad \beta_\lambda = \sqrt{1-\lambda}$$

satisfies Eqs. (15) and (16) with equality in Eq. (17). Hence, the right-hand side in Eq. (17) is equal to  $\gamma_{\max}(\lambda)$ .

If  $\phi'$  is a vector satisfying Eqs. (15) and (16) with  $\gamma = \gamma_{\max}(\lambda)$ , then  $\phi' = (\alpha_\lambda \varphi_0 + \beta_\lambda \psi_0) \otimes \eta'$  for some unit vector  $\eta' \in \mathcal{H}$ .

*Proof.* Suppose  $\phi \in \mathcal{H} \otimes \mathcal{H}$  satisfies Eqs. (15) and (16). We write  $\phi$  in the form

$$\phi = \sum_{i \in \mathbb{Z}_d} \varphi_i \otimes \xi_i,$$

where  $\{\xi_i\}_{i \in \mathbb{Z}_d}$  are vectors in  $\mathcal{H}$ . From Eq. (15), it follows that

$$\|\xi_i\|^2 = \lambda \delta(i) + (1-\lambda)\mu(i),$$

hence, there exist unit vectors  $\{\eta_i\}_{i \in \mathbb{Z}_d}$  such that

$$\xi_0 = \sqrt{\frac{(d-1)\lambda + 1}{d}} \eta_0, \quad \xi_i = \sqrt{\frac{1-\lambda}{d}} \eta_i \quad \forall i \neq 0.$$

On the other hand, we have

$$\begin{aligned} \langle \phi | [\mathbf{B}(k) \otimes \mathbb{1}] \phi \rangle &= \sum_{i, j \in \mathbb{Z}_d} \langle \varphi_j \otimes \xi_j | (|\psi_k\rangle\langle\psi_k| \otimes \mathbb{1}) \varphi_i \otimes \xi_i \rangle \\ &= \frac{1}{d} \sum_{i, j \in \mathbb{Z}_d} \omega^{jk} \omega^{-ik} \langle \xi_j | \xi_i \rangle, \end{aligned}$$

so, by Eq. (16), we must have

$$\frac{1}{d} \sum_{i, j \in \mathbb{Z}_d} \omega^{jk} \omega^{-ik} \langle \xi_j | \xi_i \rangle = \gamma \delta(k) + (1-\gamma)\mu(k).$$

This equation, evaluated at  $k = 0$ , gives

$$\begin{aligned} \left(1 - \frac{1}{d}\right)\gamma + \frac{1}{d} &= \frac{1}{d} \left\| \sum_{i \in \mathbb{Z}_d} \xi_i \right\|^2 \\ &= \frac{1}{d^2} \left\| \sqrt{(d-1)\lambda + 1} \eta_0 + \sum_{i \in \mathbb{Z}_d, i \neq 0} \sqrt{1-\lambda} \eta_i \right\|^2. \end{aligned}$$

The maximum value of  $\gamma$  then is achieved when the right-hand side of this equation is maximal, i.e., when there exists a unit vector  $\eta \in \mathcal{H}$  such that  $\eta_i = \eta \forall i \in \mathbb{Z}_d$ . The corresponding maximum value  $\gamma_{\max}$  of  $\gamma$  is given by

$$d(d-1)\gamma_{\max} + d = [\sqrt{(d-1)\lambda + 1} + (d-1)\sqrt{1-\lambda}]^2,$$

i.e.,

$$\gamma_{\max} = \frac{(d-2)(1-\lambda) + 2\sqrt{(1-d)\lambda^2 + (d-2)\lambda + 1}}{d}.$$

In order to show that, if the sequence  $\{\xi_i\}_{i \in \mathbb{Z}_d}$  is chosen as above, then the corresponding vector

$$\begin{aligned} \phi &= \left( \sqrt{\frac{(d-1)\lambda + 1}{d}} \varphi_0 + \sqrt{\frac{1-\lambda}{d}} \sum_{i \in \mathbb{Z}_d, i \neq 0} \varphi_i \right) \otimes \eta \\ &= (\alpha_\lambda \varphi_0 + \beta_\lambda \psi_0) \otimes \eta, \end{aligned}$$

also satisfies Eq. (16) with  $\gamma = \gamma_{\max}$  (and, thus, the maximum indeed is achieved by  $\phi$ ), we evaluate

$$\begin{aligned} \langle \phi | [\mathbf{B}(k) \otimes \mathbb{1}] \phi \rangle &= \langle \phi | (|\psi_k\rangle\langle\psi_k| \otimes \mathbb{1}) \phi \rangle \\ &= |\langle \psi_k | \alpha_\lambda \varphi_0 + \beta_\lambda \psi_0 \rangle|^2 \\ &= \left( \frac{\alpha_\lambda}{\sqrt{d}} + \beta_\lambda \delta(k) \right)^2 \\ &= \frac{\alpha_\lambda^2}{d} + \left( \beta_\lambda^2 + 2 \frac{\alpha_\lambda \beta_\lambda}{\sqrt{d}} \right) \delta(k) \\ &= (1 - \gamma_{\max})\mu(k) + \gamma_{\max} \delta(k), \end{aligned}$$

which is Eq. (16).  $\blacksquare$

As a consequence of the above discussion, we obtain an inequality for the unsharpnesses of two jointly measurable observables  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$ .

*Proposition 5.* Two observables  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  are jointly measurable if and only if

$$\begin{aligned} \gamma &\leq \gamma_{\max}(\lambda) \\ &= \frac{1}{d} [(d-2)(1-\lambda) + 2\sqrt{(1-d)\lambda^2 + (d-2)\lambda + 1}] \end{aligned} \quad (19)$$

(or, equivalently, its modified form under the exchange  $\gamma \leftrightarrow \lambda$ ).

If  $\gamma = \gamma_{\max}(\lambda)$ , then  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  have a unique joint observable. This unique joint observable is the covariant phase-space observable  $\mathbf{C}_T$  defined by state

$$T = |\chi_\lambda\rangle\langle\chi_\lambda|, \quad \chi_\lambda = \alpha_\lambda \varphi_0 + \beta_\lambda \psi_0,$$

with

$$\begin{aligned} \alpha_\lambda &= \frac{1}{\sqrt{d}} [\sqrt{(d-1)\lambda + 1} - \sqrt{1-\lambda}], \\ \beta_\lambda &= \sqrt{1-\lambda}. \end{aligned} \quad (20)$$

*Proof.* If  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  are jointly measurable, then the inequality follows from Proposition 2 and Lemma 1. Conversely, if  $\gamma_{\max}$  is given by Eq. (19), then the pair  $\mathbf{A}_\lambda$  and  $\mathbf{B}_{\gamma_{\max}}$  are jointly measurable again by an application of Proposition 2 and Lemma 1. Then,  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  are jointly measurable by Proposition 3.

Now, suppose  $\lambda$  and  $\gamma$  achieve the bound (19), and let  $\mathbf{C}_T$  be a covariant joint observable of  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$ . Pick  $\phi \in \mathcal{H} \otimes \mathcal{H}$  such that  $T = \text{tr}_2[[\phi]\langle\phi|]$ . As  $\phi$  satisfies Eqs. (15) and (16) with  $\gamma = \gamma_{\max}$ , by Lemma 1, it must be given by Eq. (18) for some choice of a unit vector  $\eta \in \mathcal{H}$ , and  $T = \text{tr}_2[[\phi]\langle\phi|] = |\chi_\lambda\rangle\langle\chi_\lambda|$  with  $\chi_\lambda$  as in the statement of the proposition.

Finally, we need to prove that  $\mathbf{C}_T$  is the unique joint observable (and not only unique among covariant phase-space

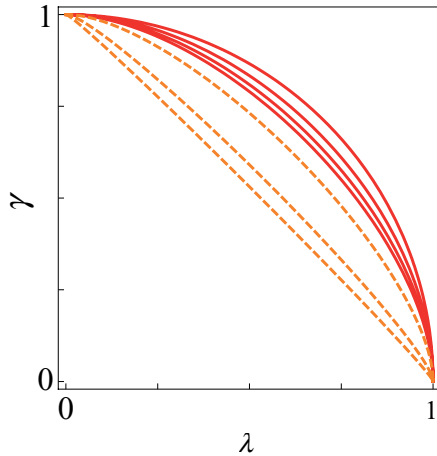


FIG. 1. (Color online) The boundary curve  $\lambda \mapsto \gamma_{\max}(\lambda)$  for  $d = 2, 3, 4, 5$  (red solid curves) and for  $d = 10, 100, 1000$  (orange dashed curves).

observables). We notice that  $T^2 = T$ , and the claim, thus, follows from Remark 4. ■

The graph of the function  $\lambda \mapsto \gamma_{\max}(\lambda)$  is a part of an ellipse. In Fig. 1, we have depicted it for  $d = 2, 3, 4, 5, 10, 100, 1000$ .

*Example 1.* Suppose that two jointly measurable observables  $A_\lambda$  and  $B_\gamma$  are equally unsharp but as close to  $A$  and  $B$  as possible, i.e.,

$$\gamma = \lambda = \lambda_{\max}(\gamma).$$

In this case, Proposition 5 gives

$$\gamma = \lambda = \frac{d + \sqrt{d} - 2}{2(d - 1)} = \frac{1}{2} \left( 1 + \frac{1}{1 + \sqrt{d}} \right).$$

The observables  $A_\lambda$  and  $B_\gamma$  then have a unique joint observable, which is the covariant phase-space observable  $C_T$  associated with state  $T = |\chi\rangle\langle\chi|$ , with

$$\chi = \sqrt{\frac{\sqrt{d}}{2(1 + \sqrt{d})}} (\varphi_0 + \psi_0).$$

Vector state  $\chi$ , hence, is an equal superposition of vector states  $\varphi_0$  and  $\psi_0$ .

By a direct calculation, one can verify that the inequality (19) can be rewritten in the following equivalent form, which is symmetric in  $\lambda$  and  $\gamma$ .

*Proposition 6.* Two observables  $A_\lambda$  and  $B_\gamma$  are jointly measurable iff

$$\gamma + \lambda \leq 1, \tag{21}$$

$$\text{or } \gamma^2 + \lambda^2 + \frac{2(d-2)}{d}(1-\gamma)(1-\lambda) \leq 1. \tag{22}$$

The latter inequality describes a full ellipse. Therefore, the first condition is needed to ignore the lower part of the ellipse, which is not a correct boundary for joint measurability. This is depicted in Fig. 2.

Let us notice that, for  $d = 2$ , the inequality (22) becomes  $\gamma^2 + \lambda^2 \leq 1$ , and then, the first condition is redundant. We, thus, recover the single condition stated in Eq. (1) and proved in Ref. [11]. Also, for  $d = 3$  and  $d = 4$ , a direct calculation

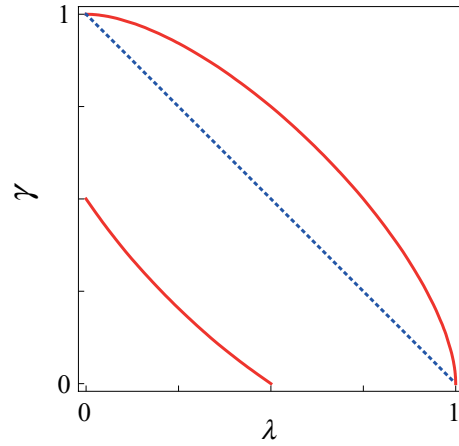


FIG. 2. (Color online) In this picture,  $d = 8$ . The line (dashed blue) and the ellipse (solid red) that define the boundaries of Eqs. (21) and (22), respectively. Only the upper side of the ellipse is relevant for the joint measurability.

shows that the linear inequality (21) is redundant, and the quadratic inequality (22) is necessary and sufficient for the joint measurability. For  $d \geq 5$ , we need both conditions (21) and (22).

By inspecting the function  $\lambda \mapsto \gamma_{\max}(\lambda)$ , we see that, for every  $\epsilon > 0$ , there is a pair of observables  $A_\lambda$  and  $B_\gamma$  such that they are not jointly measurable and  $\lambda + \gamma < 1 + \epsilon$ . Thus, the criterion  $\gamma + \lambda \leq 1$  is the best *sufficient* condition for joint measurability, which is linear and symmetric in  $\lambda$  and  $\gamma$ .

The best linear and symmetric *necessary* condition for joint measurability is achieved by taking the tangent of the boundary curve in the point where it crosses the line  $\gamma = \lambda$ . In this way, we obtain the following conclusion.

*Proposition 7.* If  $A_\lambda$  and  $B_\gamma$  are jointly measurable, then

$$\gamma + \lambda \leq 1 + \frac{\sqrt{d} - 1}{d - 1}.$$

One can also see Proposition 7 in the opposite order; if  $\gamma + \lambda > 1 + \frac{\sqrt{d}-1}{d-1}$ , then  $A_\lambda$  and  $B_\gamma$  are not jointly measurable. In Fig. 3, we have depicted the linear necessary and sufficient conditions in the case  $d = 10$ .

### C. Noncovariant observables

So far, we have concentrated on covariant observables  $A_\Lambda$  and  $B_\Gamma$ . Let us take a quick look at a class of noncovariant observables.

Let  $p$  and  $r$  be two probability distributions on  $\mathbb{Z}_d$ . We define

$$A_{\lambda;p}(j) := \lambda A(j) + (1 - \lambda)p(j)\mathbb{1},$$

and

$$B_{\gamma;r}(k) := \gamma B(k) + (1 - \gamma)r(k)\mathbb{1}.$$

It is straightforward to verify that  $A_{\lambda;p}$  is  $U$  covariant and  $V$  invariant iff  $p$  is the uniform distribution on  $\mathbb{Z}_d$ , in which case,  $A_{\lambda;p} = A_\lambda$ . (The analogous statement holds for  $B_{\gamma;r}$ .) The following result is a generalization of Proposition 5 in Ref. [27].



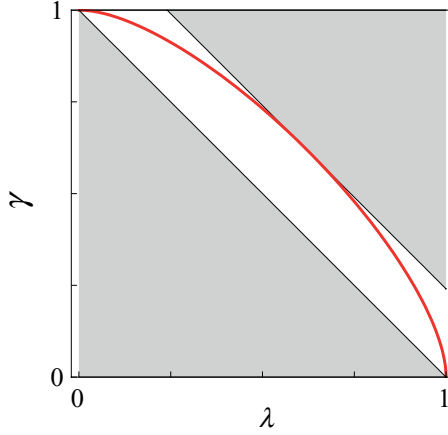


FIG. 3. (Color online) In this picture,  $d = 10$ . The gray regions represent the necessary and sufficient linear conditions. In the white stripe, one has to invoke the quadratic ellipse criterion (red curve), whereas, otherwise, the joint measurability can be deduced from the simple linear criteria.

*Proposition 8.* If  $A_{\lambda;p}$  and  $B_{\gamma;r}$  are jointly measurable, then  $A_\lambda$  and  $B_\gamma$  are jointly measurable.

*Proof.* Suppose that  $A_{\lambda;p}$  and  $B_{\gamma;r}$  are jointly measurable, and let  $C$  be their joint observable. As in the proof of Proposition 1, we define the observable  $\tilde{C}$ , given by

$$\tilde{C}(j,k) := \frac{1}{d^2} \sum_{x,y \in \mathbb{Z}_d} U_x^* V_y^* C(j+x, k+y) V_y U_x.$$

For each  $j$ , we have

$$\begin{aligned} \sum_{k \in \mathbb{Z}_d} \tilde{C}(j,k) &= \frac{1}{d^2} \sum_{x,y \in \mathbb{Z}_d} U_x^* V_y^* \sum_{k \in \mathbb{Z}_d} C(j+x, k+y) V_y U_x \\ &= \frac{1}{d^2} \sum_{x,y \in \mathbb{Z}_d} U_x^* V_y^* A_{\lambda;p}(j+x) V_y U_x \\ &= \lambda A(j) + (1-\lambda) \frac{1}{d^2} \sum_{x,y \in \mathbb{Z}_d} p(j+x) \mathbb{1} \\ &= A_\lambda(j). \end{aligned}$$

In a similar way, we obtain  $\sum_{j \in \mathbb{Z}_d} \tilde{C}(j,k) = B_\gamma(k)$  for every  $k$ . Therefore,  $\tilde{C}$  is a joint observable for  $A_\lambda$  and  $B_\gamma$ . ■

As a consequence of Propositions 5 and 8, we conclude the following necessary criterion for joint measurability.

*Corollary 1.* If two observables  $A_{\lambda;p}$  and  $B_{\gamma;r}$  are jointly measurable, then

$$\gamma \leq \frac{1}{d} [(d-2)(1-\lambda) + 2\sqrt{(1-d)\lambda^2 + (d-2)\lambda + 1}].$$

A necessary and sufficient inequality for the joint measurability of  $A_{\lambda;p}$  and  $B_{\gamma;r}$  also must contain  $p$  and  $r$  in one form or another. Thus, it is clear that Corollary 1 does not give a sufficient condition. A necessary and sufficient condition in case  $d = 2$  has been obtained in Refs. [31–33].

We remark that a general necessary condition for the joint measurability of two observables on a finite-dimensional system has been presented in Ref. [34]. A comparison to Proposition 5 shows that this condition is not sufficient. We

leave it as an open problem to find a necessary and sufficient condition for the joint measurability of  $A_{\lambda;p}$  and  $B_{\gamma;r}$ .

## VI. INFORMATIONAL COMPLETENESS

We will now study the informational completeness of joint observables of  $A_\lambda$  and  $B_\gamma$ . Let us first recall that the informational completeness of a covariant phase-space observable  $C_T$  is equivalent to the criterion

$$\text{tr}[T U_x V_y] \neq 0 \quad \forall x, y \in \mathbb{Z}_d. \quad (23)$$

This result has been discussed, e.g., in Refs. [35–37]. For completeness, we provide a proof in the Appendix.

*Proposition 9.* Suppose  $A_\lambda$  and  $B_\gamma$  are two observables with  $\lambda \notin \{0,1\}$  and  $\gamma = \gamma_{\max}(\lambda)$ . Then,  $A_\lambda$  and  $B_\gamma$  have a unique joint observable  $C$ . The observable  $C$  is informationally complete if and only if  $d$  is odd.

*Proof.* From Proposition 5, we know  $A_\lambda$  and  $B_\gamma$  have a unique joint observable  $C_T$ , generated by state  $T = |\chi_\lambda\rangle\langle\chi_\lambda|$  with  $\chi_\lambda = \alpha_\lambda \varphi_0 + \beta_\lambda \psi_0$ . The informational completeness of  $C_T$  is equivalent to the condition (23), and a straightforward calculation gives

$$\begin{aligned} \text{tr}[T U_x V_y] &= \langle\chi_\lambda| U_x V_y |\chi_\lambda\rangle \\ &= \alpha_\lambda^2 \delta_{x,0} + \beta_\lambda^2 \delta_{y,0} + \frac{\alpha_\lambda \beta_\lambda}{\sqrt{d}} (\omega^{-xy} + 1). \end{aligned}$$

Let us first notice that  $\alpha_\lambda > 0$  and  $\beta_\lambda > 0$  since both  $\lambda$  and  $\gamma$  are nonzero [see Eq. (20)]. Hence,  $\text{tr}[T U_x V_y] = 0$  exactly when  $\omega^{-xy} = -1$ . The latter condition is equivalent to  $2xy \equiv d \pmod{2d}$ . We conclude that the informational completeness of  $C_T$  is equivalent to the fact that the equation  $2x = d \pmod{2d}$  has no solution  $x \in \mathbb{Z}_d$ , and this holds if and only if  $d$  is odd. ■

In Proposition 9, the crucial assumption is that  $\gamma = \gamma_{\max}(\lambda)$ . This guarantees that  $A_\lambda$  and  $B_\gamma$  have a unique joint observable. If we have  $0 < \gamma < \gamma_{\max}(\lambda)$ , then  $A_\lambda$  and  $B_\gamma$  have infinitely many joint observables. In this case, it is always possible to choose an informationally complete joint observable as we prove in the following.

*Proposition 10.* Suppose  $A_\lambda$  and  $B_\gamma$  are two observables with  $\lambda \notin \{0,1\}$  and  $0 < \gamma < \gamma_{\max}(\lambda)$ . Then, they have an informationally complete covariant joint observable.

*Proof.* Let  $(\gamma_0, \lambda_0)$  be the intersection of the half line  $\mathbb{R}_+(\gamma, \lambda)$  with the boundary of the domain (19) in  $\mathbb{R}_+^2$ , and let  $t_0 > 1$  such that  $t_0(\gamma, \lambda) = (\gamma_0, \lambda_0)$ . Let  $\tau = 1 - 1/t_0 \in (0, 1)$ . We treat the cases of odd and even  $d$  separately.

(1) Suppose that  $d = 2n$  is even. For all  $k \in \mathbb{Z}_d$ , we denote

$$X_k := \frac{i}{2} (|\varphi_{-k}\rangle\langle\varphi_0| - |\varphi_0\rangle\langle\varphi_{-k}| + |\varphi_k\rangle\langle\varphi_0| - |\varphi_0\rangle\langle\varphi_k|).$$

The linear maps  $X_k$  are selfadjoint operators with trace 0 for every  $k$ , and it is easy to check that

$$\sum_{x \in \mathbb{Z}_d} U_x V_y X_k V_y^* U_x^* = 0, \quad \sum_{y \in \mathbb{Z}_d} U_x V_y X_k V_y^* U_x^* = 0.$$

We introduce the self-adjoint operators

$$X = \sum_{k \in \mathbb{Z}_d} X_k,$$

and, for  $\kappa > 0$ ,

$$S_\kappa = \frac{1}{d}\mathbb{1} + \kappa X.$$

If  $\kappa < 1/(d\|X\|)$ , then  $S_\kappa \in \mathcal{S}(\mathcal{H})$ . Moreover, the associated covariant phase-space observable  $\mathbf{C}_{S_\kappa}$  has trivial marginals  $\mathbf{A}_0$  and  $\mathbf{B}_0$ . A straightforward calculation gives

$$\text{tr}[S_\kappa U_x V_y] = \delta_{x,0}\delta_{y,0} + i\kappa(\omega^{-xy} - 1).$$

The covariant phase-space observable  $\mathbf{C}_{T_\kappa}$  associated with state  $T_\kappa = (1 - \tau)|\chi_{\lambda_0}\rangle\langle\chi_{\lambda_0}| + \tau S_\kappa$ , with  $\chi_{\lambda_0} = \alpha_{\lambda_0}\varphi_0 + \beta_{\lambda_0}\psi_0$ , has marginals  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$ . Moreover,

$$\begin{aligned} \text{tr}[T_\kappa U_x V_y] &= (1 - \tau)(\alpha_{\lambda_0}^2 \delta_{x,0} + \beta_{\lambda_0}^2 \delta_{y,0}) + \tau \delta_{x,0}\delta_{y,0} \\ &+ (1 - \tau) \frac{\alpha_{\lambda_0}\beta_{\lambda_0}}{\sqrt{d}}(\omega^{-xy} + 1) + i\kappa\tau(\omega^{-xy} - 1). \end{aligned}$$

Let

$$\varepsilon = \min_{k \in \mathbb{Z}_d, k \neq n} |\omega^k + 1|, \quad \delta = \max_{k \in \mathbb{Z}_d} |\omega^k + 1|.$$

For  $\kappa < \min\{\alpha_{\lambda_0}\beta_{\lambda_0}(1 - \tau)\varepsilon/(\tau\delta\sqrt{d}), 1/(d\|X\|)\}$ , the right-hand side of the last equation is nonzero for all  $x, y \in \mathbb{Z}_d$ , which proves informational completeness of  $\mathbf{C}_{T_\kappa}$  by the criterion (23).

(2) Suppose that  $d$  is odd. Then, for

$$T = (1 - \tau)|\chi_{\lambda_0}\rangle\langle\chi_{\lambda_0}| + \frac{\tau}{d}\mathbb{1},$$

the associated covariant phase-space observable  $\mathbf{C}_T$  has marginals  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$ , and

$$\begin{aligned} \text{tr}[T U_x V_y] &= (1 - \tau)[\alpha_{\lambda_0}^2 \delta_{x,0} + \beta_{\lambda_0}^2 \delta_{y,0}] \\ &+ (1 - \tau) \left[ \frac{\alpha_{\lambda_0}\beta_{\lambda_0}}{\sqrt{d}}(\omega^{-xy} + 1) \right] + \tau \delta_{x,0}\delta_{y,0}, \end{aligned}$$

which is nonzero for all  $x, y \in \mathbb{Z}_d$ . The informational completeness of  $\mathbf{C}_T$  then follows from the criterion (23). ■

The two trivial cases  $\lambda = 0$  or  $\gamma = 0$  are not very interesting, but for completeness, we make the following observation.

*Proposition 11.* Suppose  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  are two observables with  $\lambda = 0$  or  $\gamma = 0$ . Then, they have no informationally complete joint observable.

*Proof.* We consider only the case  $\lambda = 0$ , the case  $\gamma = 0$  being similar. Suppose that  $\mathbf{C}$  is a joint observable of  $\mathbf{A}_0 = \mu\mathbb{1}$  and  $\mathbf{B}_\lambda$ . We have

$$\sum_{k \in \mathbb{Z}_d} \mathbf{C}(j, k) = \mathbf{A}_0(j) = \frac{1}{d}\mathbb{1} \quad \forall j \in \mathbb{Z}_d,$$

hence, the linear span of set  $\{\mathbf{C}(j, k) | j, k \in \mathbb{Z}_d\}$  is the same as the linear span of set  $\{\mathbb{1}, \mathbf{C}(j, k) | j \in \mathbb{Z}_d, k \in \mathbb{Z}_d \setminus \{0\}\}$  and then, for  $d \geq 2$ ,

$$\dim \text{span}\{\mathbf{C}(j, k) | j, k \in \mathbb{Z}_d\} \leq 1 + d(d - 1) < d^2.$$

Thus,  $\mathbf{C}$  is not informationally complete. ■

## VII. SEQUENTIAL IMPLEMENTATION OF JOINT OBSERVABLES

In this section, we discuss the sequential implementation of joint observables of  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\Gamma$  in light of the recent results

obtained in Refs. [12,38]. For illustrative purposes, we point out that two naive methods do not work.

### A. Nondisturbing measurement

Suppose that  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\Gamma$  are jointly measurable, i.e., they satisfy the condition stated in Proposition 2. The most uncomplicated way to realize their joint measurement would be to perform an  $\mathbf{A}_\lambda$  measurement without disturbing the subsequent  $\mathbf{B}_\Gamma$  measurement. In terms of instruments, this would mean that we choose an  $\mathbf{A}_\lambda$ -compatible instrument  $\mathcal{I}$  such that

$$\sum_{j \in \mathbb{Z}_d} \mathcal{I}_j^*(\mathbf{B}_\Gamma(k)) = \mathbf{B}_\Gamma(k)$$

for all  $k$ . However, this type of measurement typically is not possible since a quantum measurement necessarily disturbs the input state.

Let us first notice that  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  commute if and only if  $\lambda\gamma = 0$ , meaning that one of them is a trivial observable. Generally, a nondisturbing measurement can be possible even if two observables do not commute. But by applying Proposition 3 from Ref. [38], we see that this possibility is excluded whenever  $\mathbf{B}_\gamma$  is informationally equivalent with  $\mathbf{B}$  in the sense that the linear spans of sets  $\{\mathbf{B}_\gamma(k) : k \in \mathbb{Z}_d\}$  and  $\{\mathbf{B}(k) : k \in \mathbb{Z}_d\}$  are equal. This property is satisfied by any  $\mathbf{B}_\gamma$  with  $\gamma \neq 0$ . Therefore, whenever both observables are nontrivial, then the  $\mathbf{A}_\lambda$  measurement disturbs the subsequent  $\mathbf{B}_\gamma$  measurement, and the resulting observable is not a joint measurement of  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$ .

### B. Measuring only part of the ensemble

Suppose we have a measurement setup for  $\mathbf{A}$  and that the corresponding instrument is  $\mathcal{I}$ . We can implement an unsharp observable  $\mathbf{A}_\lambda$  by performing the  $\mathbf{A}$  measurement in a randomly chosen  $\lambda$  part of the ensemble and doing nothing for the remaining  $(1 - \lambda)$  part. The corresponding  $\mathbf{A}_\lambda$ -compatible instrument  $\mathcal{I}'$  is then

$$\mathcal{I}'_j(\varrho) = \lambda \mathcal{I}_j(\varrho) + \frac{1 - \lambda}{d} \varrho. \quad (24)$$

This clearly is a very direct way to decrease the disturbance that an  $\mathbf{A}$  measurement would cause. By measuring the observable  $\mathbf{B}$  after the first measurement, one could expect to have a useful joint measurement of  $\mathbf{A}_\lambda$  and some approximate version of  $\mathbf{B}$ . However, this type of method does not yield an informationally complete joint measurement.

The observable  $\mathbf{A}$  consists of rank-1 operators, and any  $\mathbf{A}$ -compatible instrument  $\mathcal{I}$  is of the form

$$\mathcal{I}_j(\varrho) = \text{tr}[\varrho \mathbf{A}(j)] \xi_j$$

for some set of states  $\{\xi_j : j \in \mathbb{Z}_d\}$  [38]. If we insert this form into Eq. (24), we see that a sequential measurement consisting of  $\mathcal{I}'$  followed by a  $\mathbf{B}$  measurement leads to the joint observable,

$$\mathbf{C}(j, k) = \lambda \text{tr}[\xi_j \mathbf{B}(k)] \mathbf{A}(j) + \frac{1 - \lambda}{d} \mathbf{B}(k), \quad j, k \in \mathbb{Z}_d.$$

The linear span of set  $\{\mathbf{C}(j, k) : j, k \in \mathbb{Z}_d\}$  is contained in the linear span of union  $\{\mathbf{A}(j) : j \in \mathbb{Z}_d\} \cup \{\mathbf{B}(k) : k \in \mathbb{Z}_d\}$ . The

latter is strictly smaller than  $\mathcal{L}(\mathcal{H})$ , hence,  $\mathbf{C}$  is not informationally complete. We also see that this kind of approach cannot give more information than separate measurements of  $\mathbf{A}$  and  $\mathbf{B}$  would give.

### C. General joint observables

From Ref. [38], we recall that every joint observable of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$  can be implemented as a sequential measurement of  $\mathbf{A}_\Lambda$  followed by a measurement of  $\mathbf{B}$ . Namely, suppose that  $\mathbf{C}$  is a joint observable of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}_\Gamma$ . We define an instrument  $\mathcal{I}$  by

$$\mathcal{I}_j(\varrho) = \sum_{k \in \mathbb{Z}_d} \text{tr}[\varrho \mathbf{C}(j, k)] \mathbf{B}(k). \quad (25)$$

This is an  $\mathbf{A}_\Lambda$ -compatible instrument, and from  $\mathbf{B}(k)\mathbf{B}(k') = \delta_{k,k'}\mathbf{B}(k)$ , it follows that

$$\text{tr}[\mathbf{B}(k)\mathcal{I}_j(\varrho)] = \text{tr}[\varrho \mathbf{C}(j, k)].$$

Hence,  $\mathbf{C}(j, k) = \mathcal{I}_j^*(\mathbf{B}(k))$ , and we conclude that  $\mathbf{C}$  is implemented as a sequential measurement of  $\mathbf{A}_\Lambda$  followed by a measurement of  $\mathbf{B}$  as claimed.

### D. Covariant phase-space observables

The instrument defined in Eq. (25) may look quite artificial, and before we know the structure of  $\mathbf{C}$ , the formula does not give us any hint on the structure of  $\mathcal{I}$ . In contrast, every covariant phase-space observable can be implemented as a sequential measurement of  $\mathbf{A}_\Lambda$  and  $\mathbf{B}$  in a very specific form.

As explained in Ref. [12], every covariant  $\mathbf{A}_\Lambda$ -compatible instrument gives rise to a covariant phase-space observable. Here, covariance of an instrument  $\mathcal{I}$  means that

$$U_x V_y \mathcal{I}_j(V_y^* U_x^* \varrho U_x V_y) V_y^* U_x^* = \mathcal{I}_{j+x}(\varrho) \quad (26)$$

for all  $x, y, j \in \mathbb{Z}_d$  and  $\varrho \in \mathcal{S}(\mathcal{H})$ . It is straightforward to verify that the joint observable  $\mathbf{C}(j, k) := \mathcal{I}_j^*(\mathbf{B}(k))$  is a covariant phase-space observable.

We demonstrate this method by choosing the  $\mathbf{A}_\Lambda$ -compatible Lüders instrument  $\mathcal{I}^L$ , defined as

$$\mathcal{I}_j^L(\varrho) = \sqrt{\mathbf{A}_\lambda(j)} \varrho \sqrt{\mathbf{A}_\lambda(j)}.$$

It is straightforward to see that  $\mathcal{I}^L$  satisfies Eq. (26). The covariant joint observable then is

$$\mathbf{C}(j, k) = \sqrt{\mathbf{A}_\lambda(j)} \mathbf{B}(k) \sqrt{\mathbf{A}_\lambda(j)},$$

and its associated state is

$$T = d\mathbf{C}(0, 0) = d\sqrt{\mathbf{A}_\lambda(0)} |\psi_0\rangle \langle \psi_0| \sqrt{\mathbf{A}_\lambda(0)}.$$

Since

$$\sqrt{\mathbf{A}_\lambda(j)} = \frac{\beta_\lambda}{\sqrt{d}} \mathbb{1} + \alpha_\lambda |\varphi_j\rangle \langle \varphi_j|,$$

and

$$\sqrt{\mathbf{A}_\lambda(0)} \psi_0 = \frac{1}{\sqrt{d}} (\beta_\lambda \psi_0 + \alpha_\lambda \varphi_0) = \frac{1}{\sqrt{d}} \chi_\lambda,$$

we see that  $T = |\chi_\lambda\rangle \langle \chi_\lambda|$ , hence, by Proposition 5, the marginal  $\mathbf{B}_\gamma$  is such that  $\gamma = \gamma_{\max}(\lambda)$ .

In conclusion, this type of sequential measurement of  $\mathbf{A}_\lambda$  and  $\mathbf{B}$  effectively is a joint measurement of  $\mathbf{A}_\lambda$  and  $\mathbf{B}_\gamma$  with minimal unsharpnesses.

## VIII. DISCUSSION

In our investigation, we have concentrated on canonically conjugated pairs of observables, i.e., the orthonormal bases  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$  have been assumed to be Fourier connected with respect to the Fourier transform of the cyclic group  $\mathbb{Z}_d$ ; see Eq. (4). Equivalently, we have assumed that the two bases satisfy  $\langle \varphi_j | \psi_k \rangle = (1/\sqrt{d}) \omega^{jk}$  for all  $j, k \in \mathbb{Z}_d$ . As a consequence, the observables  $\mathbf{A}$  and  $\mathbf{B}$ , both defined on  $\Omega_A = \Omega_B = \mathbb{Z}_d$  as  $\mathbf{A}(j) = |\varphi_j\rangle \langle \varphi_j|$  and  $\mathbf{B}(k) = |\psi_k\rangle \langle \psi_k|$ , satisfy the covariance and invariance conditions (5) and (6), which turn out to be very useful in our calculations.

Our approach covers more cases than it may seem at first sight. Namely, we recall that two orthonormal bases  $\{\varphi_j\}_j$  and  $\{\varphi'_j\}_j$  define the same observable iff there are complex numbers  $\alpha_j$  with  $|\alpha_j| = 1$  such that  $\varphi'_j = \alpha_j \varphi_j$ . To illustrate an application of this many-to-one correspondence, suppose that dimension  $d$  is an odd prime number, say  $d = p$  (the generalization for the case  $d = p^r$ , with  $r$  as a positive integer, is straightforward). In this case, it is easy to give a full set of  $p + 1$  MUBs [8]: Fix an orthonormal basis  $\{\varphi_j\}_{j \in \mathbb{Z}_p}$ , and define  $p$  orthonormal bases  $\{\psi_k^a\}_{k \in \mathbb{Z}_p}$ , each one labeled by  $a \in \mathbb{Z}_p$ , as

$$\psi_k^a = \frac{1}{\sqrt{p}} \sum_{x \in \mathbb{Z}_p} \omega^{ax^2 + kx} \varphi_x.$$

The fact that these are MUBs follows from the Gauss summation formula

$$\frac{1}{\sqrt{p}} \sum_{x \in \mathbb{Z}_p} \omega^{ax^2} = \left(\frac{a}{p}\right) \begin{cases} 1, & \text{if } p \in 4\mathbb{N} + 1 \\ i, & \text{if } p \in 4\mathbb{N} - 1 \end{cases} \quad (27)$$

where  $\left(\frac{a}{p}\right)$  is the Legendre symbol (see, e.g., Ref. [39]).

It is immediate to see that the orthonormal basis  $\{\psi_k^a\}_{k \in \mathbb{Z}_p}$  is Fourier connected to the orthonormal basis  $\{\varphi'_j\}_{j \in \mathbb{Z}_p}$  given by

$$\varphi'_j = \omega^{aj^2} \varphi_j \quad \forall j,$$

i.e.,  $\langle \varphi'_j | \psi_k^a \rangle = (1/\sqrt{p}) \omega^{jk}$ . Moreover, for  $a, b \in \mathbb{Z}_p \setminus \{0\}$ , with  $a \neq b$ , define the rescaled orthonormal bases  $\{\psi_j^{a'}\}_{j \in \mathbb{Z}_p}$  and  $\{\psi_k^{b'}\}_{k \in \mathbb{Z}_p}$ , given by

$$\begin{aligned} \psi_j^{a'} &= \omega^{-4^{-1}j^2(b-a)^{-1}} \psi_j^a, \\ \psi_k^{b'} &= \omega^{k^2(b-a)} \left(\frac{b-a}{p}\right) \psi_{2k(b-a)}^b \begin{cases} 1, & \text{if } p \in 4\mathbb{N} + 1 \\ -i, & \text{if } p \in 4\mathbb{N} - 1 \end{cases} \end{aligned}$$

(Here,  $\omega^{x^{-1}}$  means  $\omega$  to the inverse of  $x$  in the field  $\mathbb{Z}_p$  and should not be confused with  $e^{2\pi i/px}$ ). Then, an easy computation using the Gauss formula (27) yields

$$\langle \psi_h^{b'} | \psi_k^{a'} \rangle = \frac{1}{\sqrt{p}} \omega^{-hk},$$

which shows that also  $\{\psi_j^{a'}\}_{j \in \mathbb{Z}_p}$  and  $\{\psi_k^{b'}\}_{k \in \mathbb{Z}_p}$  are Fourier connected.

More generally, one can start from a complementary pair of observables, which means that  $\{\varphi_j\}_j$  and  $\{\psi_k\}_k$  are mutually unbiased but are not necessarily Fourier connected. Obviously, we can still ask similar questions about joint measurements. Especially, it would be interesting to know whether Proposition 6 still is valid under this more general setting. In other words, the question is whether all complementary pairs are essentially similar with respect to joint measurability

Even if we leave this question open in the general case, we can see that our approach can be generalized for a larger domain than we explicitly have used it for. Indeed, all our results are still valid (and with only very slight modifications in some of the proofs) if we consider the Fourier transform with respect to a generic Abelian group  $G$  with order  $d$ , i.e.,  $G = \mathbb{Z}_{d_1} \times \dots \times \mathbb{Z}_{d_n}$  for  $d_1 + \dots + d_n = d$  and  $d_i = p_i^{r_i}$  with  $p_i$  as a prime number and  $r_i$  as a positive integer for all  $i = 1, \dots, n$ . In this case,  $\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_n$  with  $\dim \mathcal{H}_i = d_i$ , a basis  $\{\varphi_j^i\}_{j \in \mathbb{Z}_{d_i}}$  is chosen in each factor Hilbert space  $\mathcal{H}_i$ , and the  $G$ -Fourier transform of  $\mathcal{H}$  is just the tensor product  $\mathcal{F} = \mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n$ , where each  $\mathcal{F}_i$  is the  $\mathbb{Z}_{d_i}$ -Fourier transform in  $\mathcal{H}_i$  with respect to the basis  $\{\varphi_j^i\}_{j \in \mathbb{Z}_{d_i}}$  as defined in Eq. (4). The mutually unbiased bases  $\{\varphi_j\}_{j \in \mathbb{Z}_d}$  and  $\{\psi_k\}_{k \in \mathbb{Z}_d}$  are replaced by the bases  $\{\varphi_{j_1, \dots, j_n}\}_{j_1 \in \mathbb{Z}_{d_1}, \dots, j_n \in \mathbb{Z}_{d_n}}$  and  $\{\psi_{k_1, \dots, k_n}\}_{k_1 \in \mathbb{Z}_{d_1}, \dots, k_n \in \mathbb{Z}_{d_n}}$  of  $\mathcal{H}$ , given by

$$\begin{aligned} \varphi_{j_1, \dots, j_n} &= \varphi_{j_1}^1 \otimes \dots \otimes \varphi_{j_n}^n, \\ \psi_{k_1, \dots, k_n} &= \mathcal{F}^*(\varphi_{j_1, \dots, j_n}) = \psi_{k_1}^1 \otimes \dots \otimes \psi_{k_n}^n. \end{aligned}$$

Now, their associated complementary observables  $\mathbf{A}$  and  $\mathbf{B}$  both are defined on  $G$  and are given by  $\mathbf{A}(j_1, \dots, j_n) = \mathbf{A}(j_1) \otimes \dots \otimes \mathbf{A}(j_n)$  and  $\mathbf{B}(k_1, \dots, k_n) = \mathbf{B}(k_1) \otimes \dots \otimes \mathbf{B}(k_n)$ . They still satisfy the analogs of the covariance and invariance conditions (5) and (6) if representations  $U$  and  $V$  are replaced by suitable tensor products.

To demonstrate that we now can handle larger classes of complementary observables, let  $\mathcal{H} = \mathbb{C}^4$ , choose an orthonormal basis  $\{\varphi_j\}_{j \in \{0, \dots, 3\}}$  of  $\mathbb{C}^4$ , and set

$$\begin{aligned} \psi_0 &= \frac{1}{2}(\varphi_0 + \varphi_1 + \varphi_2 + \varphi_3), \\ \psi_1 &= \frac{1}{2}(\varphi_0 - \varphi_1 + \varphi_2 - \varphi_3), \\ \psi_2 &= \frac{1}{2}(\varphi_0 + \varphi_1 - \varphi_2 - \varphi_3), \\ \psi_3 &= \frac{1}{2}(\varphi_0 - \varphi_1 - \varphi_2 + \varphi_3). \end{aligned}$$

Then, the observables  $\mathbf{A}(j) = |\varphi_j\rangle\langle\varphi_j|$  and  $\mathbf{B}(k) = |\psi_k\rangle\langle\psi_k|$  are complementary. They are equally defined by any two orthonormal bases  $\{\alpha_j \varphi_j\}_{j \in \{0, \dots, 3\}}$  and  $\{\beta_k \psi_k\}_{k \in \{0, \dots, 3\}}$ , where  $\alpha_j, \beta_k$  are complex numbers with  $|\alpha_j| = |\beta_k| = 1$ . If some pair of these orthonormal bases were connected by the  $\mathbb{Z}_4$ -Fourier transform, then the matrix of their scalar products  $[\bar{\alpha}_j \beta_k \langle\varphi_j|\psi_k\rangle]$  should be equal to the  $\mathbb{Z}_4$ -Fourier matrix,

$$\frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix},$$

or to a matrix obtained from the above by some permutations of its rows and columns. It is straightforward to verify that the deriving set of equations for  $\alpha_j$  and  $\beta_k$  has no solution. However, the matrix of scalar products  $[\langle\varphi_j|\psi_k\rangle]$  is just the Fourier matrix of  $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ , i.e.,

$$\frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

In other words, the two orthonormal bases are connected by the Fourier transform of  $\mathbb{Z}_2 \times \mathbb{Z}_2$ .

### ACKNOWLEDGMENTS

The authors thank Daniel Reitzner for his help with MATHEMATICA. T.H. is grateful to Cosmo Lupo and Mario Ziman for illuminating discussions on qubit tomography. T.H. acknowledges financial support from the Academy of Finland (Grant No. 138135) and the Magnus Ehrnrooth Foundation.

### APPENDIX: CRITERION FOR INFORMATIONAL COMPLETENESS

*Theorem 1.* Let  $\mathbf{C}_T$  be a covariant phase-space observable. Then,  $\mathbf{C}_T$  is informationally complete if and only if

$$\text{tr}[TU_x V_y] \neq 0 \quad \forall x, y \in \mathbb{Z}_d. \quad (\text{A1})$$

Our proof of the above theorem relies on the following well-known reconstruction formula for the Weyl-Heisenberg group, which is just a special case of orthogonality relations for irreducible representations of compact groups.

*Proposition 12.* The following reconstruction formula holds for every  $A \in \mathcal{L}(\mathcal{H})$ :

$$\frac{1}{d} \sum_{x, y \in \mathbb{Z}_d} \text{tr}[A V_y^* U_x^*] U_x V_y = A. \quad (\text{A2})$$

*Proof.* For all  $h, k \in \mathbb{Z}_d$ , we have  $\langle\varphi_k|U_x V_y \varphi_h\rangle = \omega^{hy} \delta_{h+x, k}$ . Thus, for all  $h, k, m, n \in \mathbb{Z}_d$ , we obtain

$$\begin{aligned} & \left\langle \varphi_k \left| \left[ \sum_{x, y \in \mathbb{Z}_d} \text{tr}[|\varphi_m\rangle\langle\varphi_n| V_y^* U_x^*] U_x V_y \right] \varphi_h \right. \right\rangle \\ &= \sum_{x, y \in \mathbb{Z}_d} \langle U_x V_y \varphi_n | \varphi_m \rangle \langle \varphi_k | U_x V_y \varphi_h \rangle \\ &= \sum_{x, y \in \mathbb{Z}_d} \omega^{(h-n)y} \delta_{h+x, k} \delta_{n+x, m} \\ &= \sum_{y \in \mathbb{Z}_d} \omega^{(h-n)y} \delta_{h-n, k-m} \\ &= d \delta_{h, n} \delta_{h-n, k-m} = d \delta_{h, n} \delta_{k, m} \\ &= d \langle \varphi_k | \varphi_m \rangle \langle \varphi_n | \varphi_h \rangle, \end{aligned}$$

which proves Eq. (A2) for  $A = |\varphi_m\rangle\langle\varphi_n|$ . Since every  $A \in \mathcal{L}(\mathcal{H})$  is a linear combination of operators of this type, the claim follows. ■

*Proof of Theorem 1.* Let  $\ell(\mathbb{Z}_d^2)$  be the linear space of complex functions on  $\mathbb{Z}_d^2 \equiv \mathbb{Z}_d \times \mathbb{Z}_d$ . We recall that, by

Proposition 5.1 in Ref. [40],  $\mathbf{C}_T$  is informationally complete if and only if the linear map

$$V_T: \mathcal{L}(\mathcal{H}) \rightarrow \ell(\mathbb{Z}_d^2), \quad [V_T(A)](x, y) = \text{tr}[\mathbf{C}_T(x, y)A]$$

is injective. Since the dimensions of  $\mathcal{L}(\mathcal{H})$  and  $\ell(\mathbb{Z}_d^2)$  are both  $d^2$ , we conclude that  $\mathbf{C}_T$  is informationally complete if and only if  $V_T$  is an isomorphism.

We define the following three linear maps:

$$\begin{aligned} \Phi: \mathcal{L}(\mathcal{H}) &\rightarrow \mathcal{H} \otimes \mathcal{H}, \\ \Phi(A) &= \frac{1}{d} \sum_{x, y \in \mathbb{Z}_d} \text{tr}[AV_y^* U_x^*] \varphi_x \otimes \varphi_y, \end{aligned}$$

$$\begin{aligned} M_T: \mathcal{H} \otimes \mathcal{H} &\rightarrow \mathcal{H} \otimes \mathcal{H}, \\ M_T(\varphi_x \otimes \varphi_y) &= \text{tr}[TU_x V_y] \varphi_x \otimes \varphi_y, \end{aligned}$$

$$\begin{aligned} R: \mathcal{H} \otimes \mathcal{H} &\rightarrow \ell(\mathbb{Z}_d^2), \\ R\phi(x, y) &= \langle \varphi_y \otimes \varphi_x | \phi \rangle. \end{aligned}$$

Map  $R$  clearly is a linear isomorphism,  $\Phi$  is a linear isomorphism by Proposition 12, and  $M_T$  is a linear isomorphism if and only if Eq. (A1) holds. We now evaluate the composition map  $R(\mathcal{F} \otimes \mathcal{F}^*)M_T\Phi$ . For all  $A \in \mathcal{L}(\mathcal{H})$ , we obtain

$$\begin{aligned} &[R(\mathcal{F} \otimes \mathcal{F}^*)M_T\Phi(A)](h, k) \\ &= \frac{1}{d} \sum_{x, y \in \mathbb{Z}_d} \text{tr}[TU_x V_y] \text{tr}[AV_y^* U_x^*] \\ &\quad \times \langle \varphi_k \otimes \varphi_h | (\mathcal{F} \otimes \mathcal{F}^*)(\varphi_x \otimes \varphi_y) \rangle \\ &= \frac{1}{d^2} \sum_{x, y \in \mathbb{Z}_d} \text{tr}[TU_x V_y] \text{tr}[AV_y^* U_x^*] \omega^{yh-xk} \\ &= \frac{1}{d^2} \sum_{x', y' \in \mathbb{Z}_d} \omega^{y'h-x'k} \text{tr}[TU_h^* U_{x'} V_k^* V_{y'}] \text{tr}[AV_k V_{y'}^* U_h U_{x'}^*] \\ &= \frac{1}{d^2} \sum_{x', y' \in \mathbb{Z}_d} \text{tr}[TU_h^* V_k^* U_{x'} V_{y'}] \text{tr}[AV_k U_h V_{y'}^* U_{x'}^*] \\ &= \frac{1}{d} \text{tr}[TU_h^* V_k^* AV_k U_h] = [V_T(A)](h, k) \end{aligned}$$

[in the third equality, we set  $x = x' - h$  and  $y = y' - k$ , in the fourth equality, we used the commutation relation for  $U$  and  $V$ , and in the fifth equality, we applied the reconstruction formula (A2)]. As map  $R(\mathcal{F} \otimes \mathcal{F}^*)M_T\Phi$  is an isomorphism if and only if Eq. (A1) holds, the same is true for map  $V_T$ , and the theorem is proved. ■

- 
- [1] E. Prugovečki, *Int. J. Theor. Phys.* **16**, 321 (1977).  
[2] P. Busch and P. Lahti, *Found. Phys.* **19**, 633 (1989).  
[3] P. Busch, *Internat. J. Theoret. Phys.* **30**, 1217 (1991).  
[4] T. Opatrný, V. Bužek, J. Bajer, and G. Drobný, *Phys. Rev. A* **52**, 2419 (1995).  
[5] A. Vourdas, *Rep. Math. Phys.* **40**, 367 (1997).  
[6] I. D. Ivanović, *J. Phys. A* **14**, 3241 (1981).  
[7] W. K. Wootters, *Found. Phys.* **16**, 391 (1986).  
[8] W. K. Wootters and B. D. Fields, *Ann. Phys.* **191**, 363 (1989).  
[9] S. Brierley and S. Weigert, *Phys. Rev. A* **78**, 042312 (2008).  
[10] P. Raynal, X. Lü, and B.-G. Englert, *Phys. Rev. A*, **83**, 062303 (2011).  
[11] P. Busch, *Phys. Rev. D* **33**, 2253 (1986).  
[12] C. Carmeli, T. Heinosaari, and A. Toigo, *J. Phys. A: Math. Theor.* **44**, 285304 (2011).  
[13] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, *J. Math. Phys.* **45**, 2171 (2004).  
[14] D. M. Appleby, *J. Math. Phys.* **46**, 052107 (2005).  
[15] A. J. Scott and M. Grassl, *J. Math. Phys.* **51**, 042203 (2010).  
[16] M. Shalaby and A. Vourdas, *J. Phys. A: Math. Theor.* **44**, 345303 (2011).  
[17] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer-Verlag, Berlin, 1997), second corrected printing.  
[18] W. M. de Muynck, *Foundations of Quantum Mechanics, an Empiricist Approach* (Kluwer Academic, Dordrecht, 2002).  
[19] M. Singer and W. Stulpe, *J. Math. Phys.* **33**, 131 (1992).  
[20] P. Lahti and S. Pulmannová, *Rep. Math. Phys.* **39**, 339 (1997).  
[21] S. T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo, *Found. Phys.* **39**, 593 (2009).  
[22] E. B. Davies, *Quantum Theory of Open Systems* (Academic, London, 1976).  
[23] M. Ozawa, *J. Math. Phys.* **25**, 79 (1984).  
[24] M. Hayashi, *Quantum Information* (Springer-Verlag, Berlin, 2006). Translated from the 2003 Japanese original.  
[25] E. B. Davies and J. T. Lewis, *Commun. Math. Phys.* **17**, 239 (1970).  
[26] P. Busch, G. Cassinelli, and P. Lahti, *Found. Phys.* **20**, 757 (1990).  
[27] P. Busch and T. Heinosaari, *Quantum Inf. Comput.* **8**, 0797 (2008).  
[28] J. Schwinger, *Proc. Natl. Acad. Sci. USA* **46**, 570 (1960).  
[29] K. Kraus, *Phys. Rev. D* **35**, 3070 (1987).  
[30] C. Carmeli, T. Heinonen, and A. Toigo, *J. Math. Phys.* **45**, 2526 (2004).  
[31] P. Stano, D. Reitzner, and T. Heinosaari, *Phys. Rev. A* **78**, 012315 (2008).  
[32] P. Busch and H.-J. Schmidt, *Quantum Inf. Process.* **9**, 143 (2010).  
[33] S. Yu, N.-I. Liu, L. Li, and C. H. Oh, *Phys. Rev. A* **81**, 062116 (2010).  
[34] T. Miyadera and H. Imai, *Phys. Rev. A* **78**, 052119 (2008).  
[35] S. T. Ali and E. E. Prugovečki, *J. Math. Phys.* **18**, 219 (1977).  
[36] G. Cassinelli, E. De Vito, P. Lahti, and A. Levrero, *J. Math. Phys.* **41**, 5883 (2000).  
[37] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, *J. Opt. B* **6**, S487 (2004).  
[38] T. Heinosaari and M. M. Wolf, *J. Math. Phys.* **51**, 092201 (2010).  
[39] B. C. Berndt and R. J. Evans, *Bull. Am. Math. Soc.* **2**, 107 (1981).  
[40] P. Busch, G. Cassinelli, and P. J. Lahti, *Rev. Math. Phys.* **7**, 1105 (1995).