

# OPTIMAL COVARIANT MEASUREMENTS: THE CASE OF A COMPACT SYMMETRY GROUP AND PHASE OBSERVABLES

CLAUDIO CARMELI, TEIKO HEINOSAARI, JUHA-PEKKA PELLONPÄÄ, AND ALESSANDRO TOIGO

**ABSTRACT.** We study various optimality criteria for quantum observables. Observables are represented as covariant positive operator valued measures and we consider the case when the symmetry group is compact. Phase observables are examined as an example.

## 1. INTRODUCTION

In this work we study four optimality criteria applied on covariant quantum observables. Quantum observables are described as normalized positive operator valued measures. Covariance means that observables transform in a consistent way under a group operation. For example, phase observables are defined as the normalized positive operator measures on a circle which are covariant under the phase shifts.

A quantum observable can be optimal in various ways. There are at least four different ways how optimality can be defined. Namely, an observable can be

- (a) (approximately) sharp
- (b) extremal
- (c) postprocessing clean
- (d) preprocessing clean

If an observable is (approximately) sharp then for any neighborhood of a point from the set of possible measurement outcomes, one can prepare a state such that the probability of getting a result from the neighborhood is (approximately) 1. This reflects the fact that an approximately sharp observable does not have intrinsic unsharpness – spread in the measurement outcome distribution depends solely on the measured state.

A covariant observable is extremal in the convex set of all covariant observables if it cannot be represented as a nontrivial convex combination of other covariant observables. A convex mixture of observables corresponds to a random choice between measurement apparatuses, and thus an extremal observable is unaffected by this kind of randomness.

After a measurement of an observable, one may try to process the obtained measurement outcome data to get the measurement outcome data of another observable. An observable is called postprocessing clean if it cannot be obtained by manipulating the measurement outcome data of another observable.

Analogously, a preprocessing clean observable cannot be obtained by manipulating the state before the measurement and then measuring some other observable. State manipulation is described by a quantum channel, which typically loses some information. Hence, a preprocessing clean observable is not irreversibly connected to another observable by a channel.

In Section 2 we formulate the optimality criteria and derive some general results in the case of a compact symmetry group. In Section 3 we focus on the phase observables. The phase observables form an interesting class as there is no sharp phase observable (i.e. projection valued measure). This class is also a rich example for the general investigation on optimal observables, showing various connections and differences between the four optimality criteria.

## 2. OPTIMAL COVARIANT OBSERVABLES

**2.1. Basic definitions.** In quantum mechanics, observables are represented as normalized positive operator valued measures. We shortly recall some relevant basic concepts. For more details, we refer to [2], [6], [13].

Let  $\mathcal{H}$  be a complex separable Hilbert space. We denote by  $\mathcal{L}(\mathcal{H})$  the set of bounded operators and  $\mathcal{T}(\mathcal{H})$  the set of trace-class operators on  $\mathcal{H}$ . We denote by  $O$  and  $I$  be the zero and identity operators on  $\mathcal{H}$ , respectively. Let  $\Omega$  be a topological space. We denote by  $\mathcal{B}(\Omega)$  the Borel  $\sigma$ -algebra on  $\Omega$ .

**Definition 1.** A set function  $E : \mathcal{B}(\Omega) \rightarrow \mathcal{L}(\mathcal{H})$  is an *observable* if it satisfies the following conditions:

- $O \leq E(X) \leq I$  for any  $X \in \mathcal{B}(\Omega)$ ;
- $E(\Omega) = I$ ;
- $E(\cup_i X_i) = \sum_i E(X_i)$  for any sequence  $\{X_i\}$  of disjoint Borel sets, the sum converging in the weak operator topology.

The set  $\Omega$  is called the *outcome set* of  $E$ .

Two observables  $E$  and  $F$  are called *equivalent*, and denoted  $E \sim F$ , if there exists a unitary operator  $W \in \mathcal{L}(\mathcal{H})$  such that  $F(X) = WE(X)W^*$  for all  $X \in \mathcal{B}(\Omega)$ .

States of a quantum system are described by (and identified with) positive trace class operators of trace 1. We denote by  $\mathcal{S}(\mathcal{H})$  the convex set of all states. If  $E$  is an observable and  $\varrho$  is a state, then the trace formula

$$p_\varrho^E(X) := \text{tr}[\varrho E(X)], \quad X \in \mathcal{B}(\Omega),$$

defines a probability measure  $p_\varrho^E$  on  $\mathcal{B}(\Omega)$ . The number  $p_\varrho^E(X)$  is interpreted as the probability of getting a measurement outcome  $x$  belonging to  $X$ , when the system is in the state  $\varrho$  and the observable  $E$  is measured.

There is also an equivalent description of observables which we need later. Namely, an observable  $E$  determines a mapping  $\Theta_E$  from the set of states  $\mathcal{S}(\mathcal{H})$  into the set of probability measures  $M_1^+(\Omega)$ ,

$$\varrho \mapsto \Theta_E(\varrho) := p_\varrho^E.$$

The mapping  $\Theta_E$  is affine, i.e., it maps convex combinations of states to convex combinations of corresponding probability measures.

In this paper we study observables which have a specific symmetry property, called covariance. To formulate this concept, let  $G$  be a topological group. For simplicity, here we consider only the case where  $G$  is the symmetry group and also the outcome space. (This corresponds to the situation when the symmetry group acts transitively on the outcome space  $\Omega$  and the stability subgroup is the trivial one element group consisting of the identity element  $e$  only.) By a unitary representation of  $G$  we mean a strongly continuous group homomorphism from  $G$  to the group  $\mathcal{U}(\mathcal{H})$  of unitary operators in a Hilbert space  $\mathcal{H}$ .

**Definition 2.** Let  $U$  be a unitary representation of  $G$  in a Hilbert space  $\mathcal{H}$ . An observable  $E : \mathcal{B}(G) \rightarrow \mathcal{L}(\mathcal{H})$  is *covariant* with respect to  $U$  (or  *$U$ -covariant*, for short) if

$$U(g)E(X)U(g)^* = E(gX)$$

for all  $g \in G$  and  $X \in \mathcal{B}(G)$ .

Two  $U$ -covariant observables  $E$  and  $F$  are called  *$U$ -equivalent*, and denoted  $E \sim_U F$ , if there exists a unitary operator  $W \in \mathcal{L}(\mathcal{H})$  such that  $U(g)W = WU(g)$  for all  $g \in G$  and  $F(X) = WE(X)W^*$  for all  $X \in \mathcal{B}(G)$ .

Covariance of observables arises in many different situations; we refer to [2], [11] for many examples. Trivially, for two  $U$ -covariant observables  $E$  and  $F$ , the equivalence  $E \sim_U F$  implies that  $E \sim F$  but, as we will show later (see Proposition 10),  $E \sim F$  does *not* necessarily imply that  $E \sim_U F$ .

In the rest of this paper,  $G$  is a compact topological group which is Hausdorff and satisfies the second axiom of countability. We denote by  $\mu$  the Haar measure of  $G$ , normalized so that  $\mu(G) = 1$ . We fix a unitary representation  $U$  of  $G$ , and the set of all  $U$ -covariant observables is denoted by  $\mathcal{O}$ .

**2.2. Approximately sharp observables.** Suppose for a moment that  $E : \mathcal{B}(G) \rightarrow \mathcal{L}(\mathcal{H})$  is a projection valued observable, also called a *sharp observable*. Then for any  $X \in \mathcal{B}(G)$  with  $E(X) \neq O$ , we can choose a (pure) state  $\varrho \in \mathcal{S}(\mathcal{H})$  such that  $\text{tr}[\varrho E(X)] = 1$ . This means that we can prepare the system in a state  $\varrho$  such that the probability of getting an outcome  $x \in X$  is 1.

This property of sharp observables is a simple consequence of the fact that each non-zero operator  $E(X)$  has eigenvalue 1. Especially,  $\|E(X)\| = 1$ . This gives a useful generalization of sharp observables [9],[13].

**Definition 3.** An observable  $E : \mathcal{B}(G) \rightarrow \mathcal{L}(\mathcal{H})$  is *approximately sharp* if  $\|E(X)\| = 1$  for every open set  $X \subseteq G$  such that  $E(X) \neq O$ .

Let us first note that the condition  $\|E(X)\| = 1$  does not imply that  $\text{tr}[\varrho E(X)] = 1$  for some state  $\varrho$ . Indeed,  $\|E(X)\| = 1$  means that 1 is in the spectrum of  $E(X)$  but 1 need not be an eigenvalue. The condition  $\|E(X)\| = 1$  is equivalent to the fact that there exists a sequence

$\{\varrho_n\}_{n \in \mathbb{N}}$  of (pure) states such that

$$\lim_{n \rightarrow \infty} \text{tr} [\varrho_n \mathbf{E}(X)] = 1.$$

Therefore, approximate sharpness means that we can prepare the system in a state  $\varrho_n$  such that the probability of getting an outcome from  $X$  is arbitrarily close to 1.

Now, assume that  $\mathbf{E}$  is a  $U$ -covariant observable. Then  $\mathbf{E}(X) = 0$  if and only if  $\mu(X) = 0$ ; see [9]. On the other hand,  $\mu(X) > 0$  for any (non-empty) open set  $X \subseteq G$ . Thus,  $\mathbf{E}$  is approximately sharp if  $\|\mathbf{E}(X)\| = 1$  for every open set  $X \subseteq G$ ,  $X \neq \emptyset$ .

**Proposition 1.** *Let  $\mathbf{E}, \mathbf{F} \in \mathcal{O}$  and  $\mathbf{E} \sim \mathbf{F}$ . If  $\mathbf{E}$  is approximately sharp, then also  $\mathbf{F}$  is such.*

*Proof.* Since  $\mathbf{E} \sim \mathbf{F}$ , there is a unitary operator  $W$  such that  $\mathbf{F}(X) = W\mathbf{E}(X)W^*$  for every  $X \in \mathcal{B}(G)$ . This implies that the operator norms of  $\mathbf{E}(X)$  and  $\mathbf{F}(X)$  are the same, i.e.,  $\|\mathbf{F}(X)\| = \|W\mathbf{E}(X)W^*\| = \|\mathbf{E}(X)\|$ .  $\square$

Until now, it may seem that the concept of an approximately sharp observable is quite artificial. The usefulness and importance of this concept becomes evident in the situations when there are no sharp observables in the set  $\mathcal{O}$ , but there exist approximately sharp observables in  $\mathcal{O}$ . This is the case for phase observables, which we will study in Section 3.

**2.3. Extremal observables.** The set of all  $U$ -covariant observables  $\mathcal{O}$  is convex, and we denote by  $\text{Ext } \mathcal{O}$  the set of extremal elements of  $\mathcal{O}$ . A convex mixture of two observables corresponds to a random choice or fluctuation between two measurement apparatuses. An extremal observable thus describes an observable which is unaffected by this kind of randomness.

In our earlier article [3] we have characterized the extremal observables in  $\mathcal{O}$ . We do not reproduce this characterization here but we use it in the special case of phase observables in Section 3.5. Extremality and approximate sharpness have the following simple connection.

**Proposition 2.** *Let  $\mathbf{E}, \mathbf{E}_1, \mathbf{E}_2$  be three observables and assume that  $\mathbf{E}$  is a convex combination of  $\mathbf{E}_1$  and  $\mathbf{E}_2$ . If  $\mathbf{E}$  is approximately sharp, then also  $\mathbf{E}_1$  and  $\mathbf{E}_2$  are approximately sharp.*

*Proof.* Let  $X \subseteq G$  be an open set. As  $\mathbf{E}$  is a convex combination of  $\mathbf{E}_1$  and  $\mathbf{E}_2$ , we have  $\mathbf{E}(X) = \lambda \mathbf{E}_1(X) + (1 - \lambda) \mathbf{E}_2(X)$  for some  $0 < \lambda < 1$ . By the triangle inequality, we get

$$\|\mathbf{E}(X)\| \leq \lambda \|\mathbf{E}_1(X)\| + (1 - \lambda) \|\mathbf{E}_2(X)\| \leq 1.$$

Thus, if  $\|\mathbf{E}(X)\| = 1$ , then  $\|\mathbf{E}_1(X)\| = \|\mathbf{E}_2(X)\| = 1$ .  $\square$

**2.4. Postprocessing.** Let  $M_1^+(G)$  be the convex set of probability measures on  $\mathcal{B}(G)$ . As we earlier discussed, each observable  $\mathbf{E}$  defined on  $\mathcal{B}(G)$  determines an affine mapping  $\Theta_{\mathbf{E}}$  from  $\mathcal{S}(\mathcal{H})$  to  $M_1^+(G)$ .

**Definition 4.** Let  $\mathbf{E}, \mathbf{F} \in \mathcal{O}$ . We say that  $\mathbf{F}$  is a *postprocessing* of  $\mathbf{E}$ , and denote  $\mathbf{F} \preceq_{\text{POST}} \mathbf{E}$ , if there exists an affine mapping  $\Psi : M_1^+(G) \rightarrow M_1^+(G)$  such that

$$(1) \quad \Theta_{\mathbf{F}} = \Psi \circ \Theta_{\mathbf{E}}.$$

The relation  $\preceq_{\text{POST}}$  on  $\mathcal{O}$  is reflexive and transitive, thus a preorder. We denote by  $\overset{\text{POST}}{\sim}$  the induced equivalence relation, i.e.,  $E \overset{\text{POST}}{\sim} F$  if and only if  $F \preceq_{\text{POST}} E \preceq_{\text{POST}} F$ .

**Definition 5.** An observable  $E \in \mathcal{O}$  is *postprocessing clean* in  $\mathcal{O}$  if for every  $F \in \mathcal{O}$ , the following implication holds:

$$(2) \quad E \preceq_{\text{POST}} F \Rightarrow E \overset{\text{POST}}{\sim} F.$$

The postprocessing relation and postprocessing cleanliness have been studied in [1], [6], [8], [15], [21]. Some connections with the convex structure of the set of observables and the postprocessing relation were proved in [14]. In the following we derive some results for the postprocessing relation in the case of covariant observables.

For each probability measure  $\nu$  on  $G$ , we define a mapping  $\Psi_\nu : M_1^+(G) \rightarrow M_1^+(G)$  by

$$(3) \quad \Psi_\nu(p) = \nu * p.$$

Here  $\nu * p$  denotes the convolution of these measures, i.e.,

$$(\nu * p)(X) = \int \nu(g^{-1}X) dp(g), \quad X \in \mathcal{B}(G).$$

For  $E \in \mathcal{O}$  and  $\nu \in M_1^+(G)$ , we denote by  $E_\nu$  the observable corresponding to the affine mapping  $\Psi_\nu \circ \Theta_E$ . In other words,  $E_\nu$  is the observable defined by formula

$$(4) \quad E_\nu(X) = \int \nu(g^{-1}X) dE(g), \quad X \in \mathcal{B}(G).$$

By its definition,  $E_\nu$  is a postprocessing of  $E$ . It is straightforward to check that the  $U$ -covariance of  $E$  implies that also  $E_\nu$  is  $U$ -covariant.

**Proposition 3.** *Let  $E, F \in \mathcal{O}$ . The following conditions are equivalent:*

- (i)  $F \preceq_{\text{POST}} E$ .
- (ii) *There is a probability measure  $\nu$  such that  $F = E_\nu$ .*

*Proof.* As explained previously, (ii) implies (i).

Conversely, suppose  $\Psi : M_1^+(G) \rightarrow M_1^+(G)$  is an affine mapping such that Eq. (1) holds. The mapping  $\Psi$  extends uniquely to a bounded linear mapping  $\Psi : M(G) \rightarrow M(G)$ , where  $M(G)$  is the Banach space of Borel complex measures on  $G$ , by setting

$$\begin{aligned} \Psi(m) &= \|m_{1+}\| \Psi\left(\frac{m_{1+}}{\|m_{1+}\|}\right) - \|m_{1-}\| \Psi\left(\frac{m_{1-}}{\|m_{1-}\|}\right) \\ &\quad + i \|m_{2+}\| \Psi\left(\frac{m_{2+}}{\|m_{2+}\|}\right) - i \|m_{2-}\| \Psi\left(\frac{m_{2-}}{\|m_{2-}\|}\right). \end{aligned}$$

Here  $m = m_{1+} - m_{1-} + im_{2+} - im_{2-}$  is the Lebesgue decomposition of a measure  $m$  and  $\|m\|$  is the total variation norm of  $m \in M(G)$  (in the above formula,  $0/0 = 0$  is assumed).

For each  $g \in G$ , we define the following linear isometric isomorphism  $\Lambda_g : M(G) \rightarrow M(G)$

$$\int_G f(h) d(\Lambda_g m)(h) := \int f(gh) dm(h) \quad \forall f \in C_0(G).$$

The mapping  $g \mapsto \Lambda_g m$  is weak\*-continuous, and it is norm continuous if  $m \in L^1(G)$ . In this case, the mapping  $g \mapsto \Lambda_g \Psi \Lambda_{g^{-1}} m$  is weak\*-continuous. In fact, if  $f \in C(G)$ , then

$$\begin{aligned} & \left| \int f(x) d(\Lambda_g \Psi \Lambda_{g^{-1}} m)(x) - \int f(x) d(\Lambda_h \Psi \Lambda_{h^{-1}} m)(x) \right| \\ & \leq \left| \int f(gx) d(\Psi \Lambda_{g^{-1}} m - \Psi \Lambda_{h^{-1}} m)(x) \right| + \left| \int [f(gx) - f(hx)] d(\Psi \Lambda_{h^{-1}} m)(x) \right| \\ & \leq \|f\|_\infty \|\Psi \Lambda_{g^{-1}} m - \Psi \Lambda_{h^{-1}} m\|_M + \|f(g\cdot) - f(h\cdot)\|_\infty \|\Psi \Lambda_{h^{-1}} m\|_M \\ & \leq \|f\|_\infty \|\Psi\| \|\Lambda_{g^{-1}} m - \Lambda_{h^{-1}} m\|_M + \|f(g\cdot) - f(h\cdot)\|_\infty \|\Psi\| \|m\|_M, \end{aligned}$$

and the mappings  $g \mapsto \Lambda_{g^{-1}} m$  and  $g \mapsto f(g\cdot)$  are both norm continuous.

If  $m \in L^1(G)$ , then for all  $f \in C(G)$  the mapping  $g \mapsto \int f(x) d(\Lambda_g \Psi \Lambda_{g^{-1}} m)(x)$  is  $\mu$ -integrable, and we have

$$\left| \int \left[ \int f(x) d(\Lambda_g \Psi \Lambda_{g^{-1}} m)(x) \right] dg \right| \leq \|f\|_\infty \|\Psi\| \|m\|_M.$$

This shows that there exists a measure  $\check{\Psi}m \in M(G)$  such that

$$\int \left[ \int f(x) d(\Lambda_g \Psi \Lambda_{g^{-1}} m)(x) \right] dg = \int f(x) d(\check{\Psi}m)(x) \quad \forall f \in C(G),$$

and the linear mapping  $m \mapsto \check{\Psi}m$  is bounded from  $L^1(G)$  into  $M(G)$ .

Suppose  $m \in L^1(G) \cap M_1^+(G)$ . If  $\{f_n\}_{n \in \mathbb{N}} \subset C(G)$  is such that  $f_n \geq 0$  and  $f_n \uparrow 1$ , then  $\int f_n(x) d(\Lambda_g \Psi \Lambda_{g^{-1}} m)(x) \uparrow 1$  since  $\Lambda_g \Psi \Lambda_{g^{-1}} m \in M_1^+(G)$ . Therefore,  $\int f_n(x) d(\check{\Psi}m)(x) \uparrow 1$  by dominated convergence theorem, thus showing that  $\check{\Psi}m \in M_1^+(G)$ .

Clearly,  $\Lambda_g \check{\Psi}m = \check{\Psi} \Lambda_g m$  for all  $m \in L^1(G)$ . In particular, the map  $g \mapsto \Lambda_g \check{\Psi}m$  is continuous, hence  $\check{\Psi}m \in L^1(G)$  by Theorem 1.6 in [24] (which is unaltered even if  $G$  is not Abelian). By a result of Wendel [26], there exists  $\nu \in M_1^+(G)$  such that  $\check{\Psi}(\phi) = \nu * \phi$  for all  $\phi \in L^1(G)$ .

Since  $\Theta_E(\mathcal{S}(\mathcal{H})), \Theta_F(\mathcal{S}(\mathcal{H})) \subseteq L^1(G)$  (see e.g. [9]), and  $\Lambda_g \Theta_F = \Psi \Lambda_g \Theta_E$  by covariance of  $E$  and  $F$ , we have  $\Theta_F = \check{\Psi} \circ \Theta_E$ . We conclude that  $F = E_\nu$ , and thus, (i) implies (ii).  $\square$

Let the probability measure  $\nu$  in (4) be the Dirac measure  $\delta_x$  in some point  $x \in G$ . In this case, the corresponding observable  $E_{\delta_x}$  has the form

$$(5) \quad E_{\delta_x}(X) = E(Xx^{-1}) \equiv E_x(X).$$

This observable, which we denote by  $E_x$ , is therefore just a translated version of  $E$ . It is clear that  $E \stackrel{\text{POST}}{\sim} E_x$ . If  $G$  is an Abelian group, then  $E_x(X) = E(x^{-1}X) = U(x^{-1})E(X)U(x)$  and therefore  $E_x \sim_U E$ .

**Proposition 4.** *Let  $E \in \mathcal{O}$  and  $\nu \in M_1^+(G)$ . The following conditions are equivalent:*

- (i)  $E_\nu$  is approximately sharp.
- (ii)  $E$  is approximately sharp and  $\nu = \delta_x$  for some  $x \in G$  (i.e.  $E_\nu = E_x$ ).

*Proof.* Let us first note that since  $G$  is Hausdorff and second countable, it is metrizable and we can choose a left invariant metric  $d$  for  $G$ , i.e.,  $d(gx, gy) = d(x, y)$  for every  $g, x, y \in G$  (see, for instance, Theorem 8.3 in [10] for this fact).

Suppose that (i) holds. We make a counter assumption that  $\nu$  is not a Dirac measure, i.e., there are two different points  $x, y$  in the support of  $\nu$ . Denote  $r := \frac{1}{5}d(x, y)$ . Then the open balls  $B(x; r)$  and  $B(y; r)$  are disjoint and have positive  $\nu$ -measure. Thus, if we set  $\alpha = \min\{\nu(B(x; r)), \nu(B(y; r))\}$ , then  $\alpha > 0$ . As the metric  $d$  is left invariant, we have  $gB(x; r) = B(gx; r)$  for any  $g \in G$ . Moreover, for any  $g \in G$ , we have  $B(gx; r) \cap B(x; r) = \emptyset$  or  $B(gx; r) \cap B(y; r) = \emptyset$ . This implies that  $\nu(gB(x; r)) \leq 1 - \alpha$  for all  $g \in G$ . It then follows that

$$\|\mathbf{E}_\nu(B(x; r))\| = \left\| \int_G \nu(g^{-1}B(x; r)) d\mathbf{E}(g) \right\| \leq \sup_{g \in G} \nu(g^{-1}B(x; r)) \cdot \|\mathbf{E}(G)\| \leq 1 - \alpha < 1.$$

This is in contradiction with our assumption that (i) holds. Hence,  $\nu$  is the Dirac measure  $\delta_x$  for some point  $x \in G$ . For every  $x \in G$  and  $X \in \mathcal{B}(G)$ , we have

$$(6) \quad \|\mathbf{E}_{\delta_x}(Xx)\| = \|\mathbf{E}(X)\|.$$

Therefore,  $\|\mathbf{E}(X)\| = 1$  for every open set  $X \subseteq G$  since the right multiplication is a homeomorphism. This shows that (i)  $\Rightarrow$  (ii).

The fact that (ii)  $\Rightarrow$  (i) is clear from (6).  $\square$

As a consequence of Propositions 3 and 4, we get the following result.

**Proposition 5.** *Let  $\mathbf{E} \in \mathcal{O}$  be an approximately sharp observable. Then  $\mathbf{E}$  is postprocessing clean in  $\mathcal{O}$ . Another observable  $\mathbf{F} \in \mathcal{O}$  is postprocessing equivalent with  $\mathbf{E}$  if and only if  $\mathbf{F} = \mathbf{E}_x$  for some  $x \in G$ .*

*Proof.* Let  $\mathbf{E} \in \mathcal{O}$  be an approximately sharp observable. If  $\mathbf{F} \in \mathcal{O}$  is such that  $\mathbf{E} \preceq_{\text{POST}} \mathbf{F}$ , then by Prop. 3 we have  $\mathbf{E} = \mathbf{F}_\nu$  for some  $\nu \in M_1^+(G)$ . But Prop. 4 now implies that  $\nu = \delta_x$  for some  $x \in G$ , which means, in particular, that  $\mathbf{E} \overset{\text{POST}}{\sim} \mathbf{F}$ . Therefore,  $\mathbf{E}$  is postprocessing clean. This reasoning also proves the second claim.  $\square$

**2.5. Preprocessing.** Let  $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$  be a linear mapping such that its adjoint  $\Phi^* : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$  is completely positive and  $\Phi^*(I) = I$ . We say that  $\Phi$  is a *channel* and denote the set of all channels by  $\text{Ch}(\mathcal{H})$ . Let  $\text{Ch}(\mathcal{H})^*$  be the set of the adjoints of channels, that is,  $\text{Ch}(\mathcal{H})^*$  consists of normal completely positive maps  $\Xi : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H})$  such that  $\Xi(I) = I$ .

**Definition 6.** Let  $\mathbf{E}, \mathbf{F} \in \mathcal{O}$ . We say that  $\mathbf{F}$  is a *preprocessing* of  $\mathbf{E}$ , and denote  $\mathbf{F} \preceq_{\text{PRE}} \mathbf{E}$ , if there exists a channel  $\Phi : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$  such that

$$(7) \quad \Theta_{\mathbf{F}} = \Theta_{\mathbf{E}} \circ \Phi.$$

Written in terms of the adjoint channel  $\Phi^*$ , Eq. (7) amounts to say that  $\mathbf{F}(X) = \Phi^*(\mathbf{E}(X))$  for all  $X \in \mathcal{B}(G)$ .

The relation  $\preceq_{\text{PRE}}$  on  $\mathcal{O}$  is clearly a preorder and we denote by  $\overset{\text{PRE}}{\sim}$  the induced equivalence relation.



**Definition 7.** An observable  $E \in \mathcal{O}$  is *preprocessing clean in  $\mathcal{O}$*  if for every  $F \in \mathcal{O}$ , the following implication holds:

$$(8) \quad E \preceq_{\text{PRE}} F \Rightarrow E \stackrel{\text{PRE}}{\sim} F.$$

We have some simple connections of the preprocessing relation to the other relations.

**Proposition 6.** If  $E, F \in \mathcal{O}$  and  $E \sim F$ , then  $E \stackrel{\text{PRE}}{\sim} F$ .

*Proof.* The relation  $E \sim F$  means that  $F(\cdot) = WE(\cdot)W^*$  for some unitary operator  $W$ . Define unitary channels  $\Phi_1(T) = W^*TW$  and  $\Phi_2(T) = WTW^*$ . Then  $\Theta_F = \Theta_E \circ \Phi_1$  and  $\Theta_E = \Theta_F \circ \Phi_2$ . Hence,  $E \stackrel{\text{PRE}}{\sim} F$ .  $\square$

**Proposition 7.** Let  $E, F \in \mathcal{O}$  and  $F \preceq_{\text{PRE}} E$ . If  $F$  is approximately sharp, then also  $E$  is approximately sharp.

*Proof.* For any  $X \in \mathcal{B}(\Omega)$ , we get

$$\|F(X)\| = \sup_{\varrho \in \mathcal{S}(\mathcal{H})} \text{tr} [\varrho F(X)] = \sup_{\varrho \in \mathcal{S}(\mathcal{H})} \text{tr} [\Phi(\varrho)E(X)] \leq \sup_{\varrho \in \mathcal{S}(\mathcal{H})} \text{tr} [\varrho E(X)] = \|E(X)\|.$$

Thus,  $\|F(X)\| = 1$  implies that  $\|E(X)\| = 1$ .  $\square$

A channel  $\Phi$  is called *U-covariant* if

$$(9) \quad \Phi(U(g)\varrho U(g)^*) = U(g)\Phi(\varrho)U(g)^*$$

for every  $\varrho \in \mathcal{S}(\mathcal{H})$  and  $g \in G$ .

**Proposition 8.** Let  $E, F \in \mathcal{O}$  and  $F \preceq_{\text{PRE}} E$ . Then there is a *U-covariant* channel  $\Phi$  such that (7) holds.

*Proof.* Let  $\Phi$  be a channel such that (7) holds. Since  $E$  and  $F$  are covariant, we have

$$(10) \quad \Theta_F(U(g)^*\varrho U(g)) = \Theta_E(U(g)^*\Phi(\varrho)U(g))$$

for every  $\varrho \in \mathcal{S}(\mathcal{H})$  and  $g \in G$ . We define a linear mapping  $\check{\Phi} : \mathcal{T}(\mathcal{H}) \rightarrow \mathcal{T}(\mathcal{H})$  by setting

$$(11) \quad \check{\Phi}(\varrho) = \int U(g)\Phi(U(g)^*\varrho U(g))U(g)^* dg.$$

Let  $\Phi^*$  and  $\check{\Phi}^*$  be the adjoints of  $\Phi$  and  $\check{\Phi}$ . Hence,

$$(12) \quad \check{\Phi}^*(A) = \int U(g)\Phi^*(U(g)^*AU(g))U(g)^* dg, \quad A \in \mathcal{L}(\mathcal{H}).$$

The fact that  $\Phi^*$  is completely positive equals with

$$(13) \quad \sum_{i,j=1}^n \langle \psi_i | \Phi^*(A_i^* A_j) \psi_j \rangle \geq 0$$



for all  $n = 1, 2, \dots$ ,  $A_1, \dots, A_n \in \mathcal{L}(\mathcal{H})$ , and  $\psi_1, \dots, \psi_n \in \mathcal{H}$ . Since

$$\sum_{i,j=1}^n \langle \psi_i | \check{\Phi}^*(A_i^* A_j) \psi_j \rangle = \int \left[ \sum_{i,j=1}^n \langle U(g)^* \psi_i | \Phi^*((A_i U(g))^* (A_j U(g))) U(g)^* \psi_j \rangle \right] dg \geq 0$$

it follows that  $\check{\Phi}^*$  is completely positive. But  $\check{\Phi}^*(I) = I$  so that  $\check{\Phi}$  is a channel. Using the invariance of the Haar integral, it is straightforward to verify that  $\check{\Phi}$  is  $U$ -covariant. By Eq. (10),  $\Theta_F = \Theta_E \circ \check{\Phi}$ .  $\square$

Proposition 8 shows that the investigation of the preprocessing relation reduces to the study of covariant channels. We will need the following general result in order to characterise the set of  $U$ -covariant elements in  $\text{Ch}(\mathcal{H})^*$ . For a proof, see e.g. [4].

**Proposition 9.** *Suppose  $U$  is a unitary representation of  $G$  in a Hilbert space  $\mathcal{H}$ , and  $\Xi \in \text{Ch}(\mathcal{H})^*$  is a  $U$ -covariant channel. There exists a separable Hilbert space  $\mathcal{K}$ , a unitary representation  $D$  of  $G$  in  $\mathcal{K}$ , and an isometry  $W : \mathcal{H} \rightarrow \mathcal{K} \otimes \mathcal{H}$  such that*

$$(14) \quad \begin{aligned} WU(g) &= (D(g) \otimes U(g))W \quad \forall g \in G, \\ \Xi(A) &= W^*(I \otimes A)W \quad \forall A \in \mathcal{L}(\mathcal{H}). \end{aligned}$$

We apply Proposition 9 in Section 3.7 in the study of phase observables.

### 3. OPTIMAL PHASE OBSERVABLES

**3.1. Structure of phase observables.** Let  $\mathcal{H}$  be a complex Hilbert space with an orthonormal basis  $\{|n\rangle \mid n \in \mathbb{N}\}$ ,  $\mathbb{N} = \{0, 1, 2, \dots\}$ . We define the lowering, raising, and number operators as

$$a := \sum_{n=0}^{\infty} \sqrt{n+1} |n\rangle \langle n+1|, \quad a^* := \sum_{n=0}^{\infty} \sqrt{n+1} |n+1\rangle \langle n|, \quad N := a^* a = \sum_{n=0}^{\infty} n |n\rangle \langle n|,$$

respectively. Physically the Hilbert space  $\mathcal{H}$  and the above operators are associated with a single-mode optical field. The vectors  $|n\rangle$  are called *number states*.

Coherent states  $|z\rangle := e^{-|z|^2/2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} |n\rangle$ , where  $z \in \mathbb{C}$ , describe laser light which is quasi-monochromatic and thus can be approximated as a single-mode system; here  $|z|$  is the energy and  $\arg z$  is the phase parameter of  $|z\rangle$ .

An observable  $\mathbf{E}$  describing a phase parameter measurement should have the interval  $[0, 2\pi)$  as its outcome space. For convenience, we use one dimensional torus  $\mathbb{T} = \{t \in \mathbb{C} : |t| = 1\}$  as an equivalent description. Let  $|z\rangle$  be a coherent state and  $t \in \mathbb{T}$ . Then  $|tz\rangle$  is another coherent state, now having phase parameter  $\arg z + \arg t$  (addition modulo  $2\pi$ ). Therefore, we require that  $\mathbf{E}$  describing a phase parameter measurement satisfies the condition

$$(15) \quad p_{|tz\rangle}^{\mathbf{E}}(X) = p_{|z\rangle}^{\mathbf{E}}(t^{-1}X)$$

for all  $z \in \mathbb{C}$ ,  $t \in \mathbb{T}$  and  $X \in \mathcal{B}(\mathbb{T})$ .

As shown in [20], an observable  $\mathbf{E} : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$  satisfies condition (15) if and only if

$$(16) \quad U(t)\mathbf{E}(X)U(t)^* = \mathbf{E}(tX)$$

for all  $t \in \mathbb{T}$  and  $X \in \mathcal{B}(\mathbb{T})$ , where  $U$  is the *number representation* of  $\mathbb{T}$  in  $\mathcal{H}$ , i.e.,

$$U(t)|n\rangle = t^n|n\rangle \quad \forall n \in \mathbb{N}.$$

Clearly,  $U(t)|z\rangle = |tz\rangle$ . We take this covariance condition as the definition for phase observables.

**Definition 8.** An observable  $E : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$  is a *phase observable* if it satisfies the covariance condition (16).

The following phase theorem characterizing phase observables has been proved in various different methods in [5, 12, 18].

**Theorem 1** (Phase Theorem). *An observable  $E : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$  is a phase observable if and only if*

$$\langle m|E(X)|n\rangle = c_{m,n} \int_X t^{m-n} dt = c_{m,n} \int_{\arg X} e^{i(m-n)\theta} \frac{d\theta}{2\pi} \quad \forall m, n \in \mathbb{N},$$

where the phase matrix  $(c_{m,n})$  is a positive semidefinite complex  $\mathbb{N} \times \mathbb{N}$ -matrix and  $c_{n,n} = 1$  for all  $n \in \mathbb{N}$ .

If  $E_1, E_2$  are phase observables with phase matrices  $(c_{m,n}^1)$  and  $(c_{m,n}^2)$ , then  $E_1 \sim_U E_2$  if and only if there exists a sequence  $\{\lambda_n\}_{n \in \mathbb{N}} \in \mathbb{T}$  such that  $c_{m,n}^1 = \lambda_n \overline{\lambda_m} c_{m,n}^2$ .

Using Theorem 1 we can easily see some general properties of phase observables. For instance, the probability measure  $p_{|n\rangle}^E$  of a phase observable  $E$  in a number state  $|n\rangle$  is uniformly distributed, which is expected as number states do not have specific phase. Also, a phase observable is never sharp (a projection valued measure); see e.g. [18].

We notice that  $U$ -equivalence for two phase observables is, in general, stronger requirement than equivalence. This is demonstrated in the following proposition.

**Proposition 10.** *There exist phase observables  $E_1$  and  $E_2$  such that  $E_1$  is equivalent to  $E_2$ , but  $E_1$  is not  $U$ -equivalent to  $E_2$ .*

*Proof.* Fix  $z \in \mathbb{C}$ ,  $0 < |z| < 1$ , and define phase observables  $E_1$  and  $E_2$  by setting

$$\begin{aligned} E_1(X) &= \int_X dt I + z \int_X t^{-1} dt |0\rangle\langle 1| + \overline{z} \int_X t dt |1\rangle\langle 0|, \\ E_2(X) &= \int_X dt I + z \int_X t^{-1} dt |2\rangle\langle 3| + \overline{z} \int_X t dt |3\rangle\langle 2| \end{aligned}$$

for all  $X \in \mathcal{B}(\mathbb{T})$ . Define a unitary operator  $W$  as

$$W = |2\rangle\langle 0| + |0\rangle\langle 2| + |1\rangle\langle 3| + |3\rangle\langle 1| + \sum_{n=4}^{\infty} |n\rangle\langle n|.$$

Then  $E_2 = WE_1W^*$ . From Theorem 1 one sees easily that  $E_1$  and  $E_2$  cannot be  $U$ -equivalent.  $\square$

**3.2. Canonical phase observable.** The *canonical phase observable*  $E_{\text{can}}$  is the phase observable determined by the phase matrix  $c_{m,n} = 1$  for all  $m, n \in \mathbb{N}$ . If  $E$  is another phase observable with a phase matrix  $(c_{m,n})$ , then  $E \sim_U E_{\text{can}}$  if and only if  $|c_{m,n}| = 1$  for all  $m, n \in \mathbb{N}$ ; see [19].

The canonical phase observable  $E_{\text{can}}$  has some properties which make it special among all phase observables. For example, it is the only phase observable (up to equivalence) which generates the number shifts [19]. A drawback of  $E_{\text{can}}$  is that a realistic measurement scheme for it is not known.

We will see that  $E_{\text{can}}$  is optimal phase observable in all the four different ways we listed in Section 1. However, (perhaps surprisingly) it is not the unique phase observable having this feature.

**3.3. State generated phase observables.** Let  $D(z) := e^{za^* - \bar{z}a}$ ,  $z \in \mathbb{C}$ , be the shift operator, and denote by  $\mathcal{S}(\mathcal{H})_{\text{diag}}$  the convex set of all *diagonal states*, that is, the states of the form  $T = \sum_{n=0}^{\infty} \lambda_n |n\rangle\langle n|$ , where  $\lambda_n \geq 0$  for all  $n \in \mathbb{N}$  and  $\sum_{n=0}^{\infty} \lambda_n = 1$ . Obviously, the extremal elements of  $\mathcal{S}(\mathcal{H})_{\text{diag}}$  are the one-dimensional projections of the form  $|n\rangle\langle n|$ ,  $n \in \mathbb{N}$ .

The *phase observable generated by*  $T \in \mathcal{S}(\mathcal{H})_{\text{diag}}$  (or a *phase space phase observable*), denoted by  $E_T$ , is defined by

$$(17) \quad E_T(X) = \frac{1}{\pi} \int_{\arg X} \int_0^\infty D(re^{i\theta}) T D(re^{i\theta})^* r dr d\theta \quad (\text{weakly})$$

for all  $X \in \mathcal{B}(\mathbb{T})$  [18]. Using the decomposition of  $T$ , one gets

$$E_T(X) = \sum_{n=0}^{\infty} \lambda_n E_{|n\rangle\langle n|}(X) \quad (\text{weakly}) \text{ for all } X \in \mathcal{B}(\mathbb{T}),$$

where we have denoted  $E_{|n\rangle\langle n|}$  briefly by  $E_{|n\rangle}$ . Note that the operator  $T$  in formula (17) must be diagonal in the number basis so that the corresponding observable would be phase shift covariant [18].

State generated phase observables are important since some of them have been measured. Indeed,  $E_{|0\rangle}$  can be seen as an angle margin observable of a phase space observable associated to a  $Q$ -function, and this has been measured by Walker and Carrol for a coherent state input field [25]. Moreover, in principle all state generated phase observables can be measured using an eight-port homodyne detector [16].

We will see that state generated phase observables are approximately sharp and postprocessing clean. They are not extremal nor preprocessing clean.

**3.4. Approximately sharp phase observables.** Let  $C(\mathbb{T})$  be a Banach space of continuous complex functions on  $\mathbb{T}$  equipped with the sup norm. By the Riesz representation theorem, the topological dual  $M(\mathbb{T})$  of  $C(\mathbb{T})$  consists of regular complex Borel measures on  $\mathbb{T}$  (or Baire measures). We equip  $M(\mathbb{T})$  with the weak-star topology. In this topology, a net  $\{p_i\}_{i \in \mathcal{I}} \subset M(\mathbb{T})$  converges to a point  $p \in M(\mathbb{T})$  if  $\lim_{i \in \mathcal{I}} \int f dp_i = \int f dp$  for all  $f \in C(\mathbb{T})$ ; we denote  $p = w^*\text{-}\lim_{i \in \mathcal{I}} p_i$ .

Let  $\delta_t : C(\mathbb{T}) \rightarrow \mathbb{C}$ ,  $f \mapsto \delta_t(f) = f(t)$  be the Dirac distribution (or Dirac measure) concentrated on  $t \in \mathbb{T}$ . Let us write the parameter  $z$  of a coherent state  $|z\rangle$  in the form  $z = rt$ , where  $r = |z| \in \mathbb{R}$  and  $t = z/|z| \in \mathbb{T}$ . Let us then recall the following result, proved in [19].

**Proposition 11.** *Let  $E$  be a phase observable with a phase matrix  $(c_{m,n})$ , and let  $u \in \mathbb{T}$ . Then*

$$\text{w}^*\text{-}\lim_{r \rightarrow \infty} p_{|rt\rangle}^E = \delta_{tu}$$

*if and only if*

$$\lim_{m \rightarrow \infty} c_{m,m+k} = u^k$$

*for all  $k = 1, 2, 3, \dots$*

This leads to the following useful conclusion.

**Proposition 12.** *Let  $E$  be a phase observable with a phase matrix  $(c_{m,n})$ . If there exists  $u \in \mathbb{T}$  such that*

$$(18) \quad \lim_{m \rightarrow \infty} c_{m,m+k} = u^k$$

*for all  $k = 1, 2, 3, \dots$ , then  $E$  is approximately sharp.*

*Proof.* If condition (18) holds, then  $\text{w}^*\text{-}\lim_{r \rightarrow \infty} p_{|rt\rangle}^E = \delta_{tu}$  by the previous proposition. Let  $X \in \mathcal{B}(\mathbb{T})$  be a nonempty open set and choose  $t$  such that  $tu \in X$ . Let  $f \in C(\mathbb{T})$  be such that  $f(tu) = 1$ , the support of  $f$  is contained in  $X$  and  $0 \leq f \leq 1$ . For any  $\epsilon > 0$  there exists  $r_\epsilon \geq 0$  such that  $\int f dp_{|rt\rangle}^E > 1 - \epsilon$  for all  $r \geq r_\epsilon$ . Since, for all  $r \geq r_\epsilon$ ,

$$1 = p_{|rt\rangle}^E(X) + p_{|rt\rangle}^E(\mathbb{T} \setminus X) \geq \int f dp_{|rt\rangle}^E + p_{|rt\rangle}^E(\mathbb{T} \setminus X) > 1 - \epsilon + p_{|rt\rangle}^E(\mathbb{T} \setminus X)$$

it follows that  $p_{|rt\rangle}^E(\mathbb{T} \setminus X) < \epsilon$  for all  $r \geq r_\epsilon$ . Thus,  $\lim_{r \rightarrow \infty} p_{|rt\rangle}^E(\mathbb{T} \setminus X) = 0$  and, hence,  $\lim_{r \rightarrow \infty} p_{|rt\rangle}^E(X) = 1$ . This implies that  $\|E(X)\| = 1$ .  $\square$

Note that (in the context of the above proof)  $\lim_{r \rightarrow \infty} p_{|rt\rangle}^E(X) = \delta_{tu}(X)$  *does not* hold for all open  $X$ . Take, e.g.,  $X = \mathbb{T} \setminus \{tu\}$  and note that then  $p_{|rt\rangle}^E(X) = 1$  since  $p_{|rt\rangle}^E$  is absolutely continuous measure with respect to the Haar measure.

Since the canonical phase observable  $E_{\text{can}}$  corresponds to the phase matrix  $c_{m,n} = 1$ , it follows from Proposition 12 that  $E_{\text{can}}$  is approximately sharp (see also [9] for a different proof).

**Proposition 13.** *All state generated phase observables  $E_T$ ,  $T \in \mathcal{S}(\mathcal{H})_{\text{diag}}$ , are approximately sharp.*

*Proof.* Let  $E_T$  be the phase observable generated by  $T = \sum_{s \in \mathbb{N}} \lambda_s |s\rangle \langle s| \in \mathcal{S}(\mathcal{H})_{\text{diag}}$ . Since  $E_T = \sum_{s=0}^{\infty} \lambda_s E_{|s\rangle}$  (weakly), the phase matrix elements of  $E_T$  are  $c_{m,n}^T = \sum_{s=0}^{\infty} \lambda_s c_{m,n}^{(s)}$ . It has been shown in [19] that  $\lim_{m \rightarrow \infty} c_{m,m+k}^{(s)} = 1$  for all  $k \in \mathbb{N}$  and  $s \in \mathbb{N}$ . Since  $|c_{m,m+k}^{(s)}| \leq 1$ , we have  $\lim_{m \rightarrow \infty} c_{m,m+k}^T = 1$  for all  $k \in \mathbb{N}$ . The claim follows then from Prop. 12 with  $u = 1$ .  $\square$

Finally, let us notice that, as seen in the proof of Proposition 12,

$$\lim_{r \rightarrow \infty} \langle rt | E_T(X) | rt \rangle = 1$$

where  $X \subseteq \mathbb{T}$  is open set and  $t \in X$ . This means that in the classical limit when the energy of the state  $|z\rangle$  is large, the phase parameter of  $|z\rangle$  can be determined accurately by measuring  $E_T$ . However, there is no phase observable  $E$  for which  $\langle rt | E(X) | rt \rangle \approx 1$  for finite energies  $r = |z|$ . There exist (phase shift covariant) generalized operator measures satisfying this condition; these measures describe measurements where only a restricted class of state preparations are available [22].

**3.5. Extremal phase observables.** Let  $\mathcal{C}$  be the convex set of all phase matrices and let  $\text{Ext } \mathcal{C}$  be the set of the extremal elements. The correspondence between the convex sets  $\mathcal{C}$  and  $\mathcal{O}$  preserves the convex structures. Especially, a phase observable  $E$  determined by a phase matrix  $(c_{n,m})$  is extremal in  $\mathcal{O}$  if and only if  $(c_{n,m})$  is extremal in  $\mathcal{C}$ .

Let  $(c_{n,m})$  be a phase matrix. It is always possible to choose a Hilbert space  $\mathcal{K}$  and a sequence of unit vectors  $\{\eta_n\}_{n \in \mathbb{N}}$  which is total in  $\mathcal{K}$  such that  $c_{n,m} = \langle \eta_n | \eta_m \rangle$  for all  $n, m \in \mathbb{N}$ ; see [5]. The dimension of  $\mathcal{K}$  is uniquely determined by  $(c_{n,m})$  and we call  $\dim \mathcal{K}$  the *rank* of  $(c_{n,m})$ . Let us denote by  $\mathcal{T}(\mathcal{K})$  the Banach space of trace-class operators on  $\mathcal{K}$ . As proved in [3], [17], the phase matrix  $(c_{n,m})$  is extremal if and only if the trace-class closure of  $\text{lin}_{\mathbb{C}} \{|\eta_n\rangle\langle\eta_n| : n \in \mathbb{N}\} \subseteq \mathcal{T}(\mathcal{K})$  is  $\mathcal{T}(\mathcal{K})$ .

**Example 1.** Let  $\xi \in \mathbb{C}$ ,  $|\xi| \leq 1$ . Define unit vectors  $\eta_{2n}^\xi := |0\rangle$  and  $\eta_{2n+1}^\xi := \xi|0\rangle + \sqrt{1-|\xi|^2}|1\rangle$  for all  $n \in \mathbb{N}$  and a phase matrix  $(c_{n,m}^\xi)$  by  $c_{n,m}^\xi := \langle \eta_n^\xi | \eta_m^\xi \rangle$ . Thus,  $(c_{n,m}^\xi)$  is the phase matrix of the so-called *chess-board phase observable*  $E_\xi$ ; see [19]. Assume that  $|\xi| \neq 1$ . Then the Hilbert space  $\mathcal{K}$  is  $\mathbb{C}|0\rangle + \mathbb{C}|1\rangle$  and the rank of  $(c_{n,m}^\xi)$  is 2. Moreover,

$$|0\rangle\langle 1| \notin \text{lin}_{\mathbb{C}} \{|\eta_n^\xi\rangle\langle\eta_n^\xi| : n \in \mathbb{N}\} = \mathbb{C}|0\rangle\langle 0| + \mathbb{C}(\xi|0\rangle\langle 1| + \bar{\xi}|1\rangle\langle 0| + \sqrt{1-|\xi|^2}|1\rangle\langle 1|)$$

so that  $E_\xi$  is not extremal. If  $|\xi| = 1$ , then the rank of  $(c_{n,m}^\xi)$  is 1 and  $E_\xi$  is automatically extremal; see Proposition 5 in [3]. In addition,  $|c_{n,m}^\xi| = 1$  for all  $n, m \in \mathbb{N}$  so that  $E_\xi \sim E_{\text{can}}$  [19]. For more examples of extremal phase observables, we refer to [17].

**Proposition 14.** *If the phase matrix  $(c_{m,n})$  has rank greater than 1 and  $c_{m,n} \in \mathbb{R}$  for all  $m, n \in \mathbb{N}$ , then  $(c_{m,n}) \notin \text{Ext } \mathcal{C}$ , that is, the corresponding phase observable  $E$  is not extremal.*

*Proof.* Let  $\eta_m, \eta_n$  be two linearly independent vectors in  $\mathcal{K}$ , and define a (nonzero) bounded operator  $B = |\eta_m\rangle\langle\eta_n| - |\eta_n\rangle\langle\eta_m|$  on  $\mathcal{K}$ . For every  $j \in \mathbb{N}$ , we have

$$\text{tr}[B|\eta_j\rangle\langle\eta_j|] = c_{n,j}c_{j,m} - c_{m,j}c_{j,n} = c_{n,j}c_{j,m} - \overline{c_{n,j}c_{j,m}} = c_{n,j}c_{j,m} - c_{n,j}c_{j,m} = 0.$$

By the Hahn-Banach theorem this implies that the set  $\text{lin}_{\mathbb{C}} \{|\eta_n\rangle\langle\eta_n| : n \in \mathbb{N}\}$  is not dense in  $\mathcal{T}(\mathcal{K})$ . Therefore, using the above cited criterion the phase matrix  $(c_{m,n})$  is not extremal.  $\square$

Fix  $s \in \mathbb{N}$ . Elements of the phase matrix  $(c_{m,n}^{[s]})$  of the state generated phase observable  $E_{[s]}$  are of the form

$$c_{m,n}^{[s]} = \int_0^\infty f_m^s(x) f_n^s(x) e^{-x} dx,$$

where

$$f_n^s(x) = (-1)^{\max\{0, s-n\}} \sqrt{\frac{(\min\{n, s\})!}{(\max\{n, s\})!}} x^{|s-n|/2} L_{\min\{n, s\}}^{|s-n|}(x)$$

and

$$L_k^\alpha(x) = \sum_{l=0}^k (-1)^l \binom{k+\alpha}{k-l} \frac{x^l}{l!}$$

is the associated Laguerre polynomial; see [18, 19]. Hence, the  $\eta_n$ -vectors of  $(c_{m,n}^{[s]})$  are the unit vectors  $f_n^s$ ,  $n \in \mathbb{N}$ , of  $L^2(\mathbb{R}_+, e^{-x}dx)$ , where  $\mathbb{R}_+$  is the set of positive reals. All the functions  $f_n^s$ ,  $n \in \mathbb{N}$ , are linearly independent, and therefore the rank of  $(c_{m,n}^{[s]})$  is infinite. It is clear that  $c_{m,n}^{[s]} \in \mathbb{R}$  for all  $m, n \in \mathbb{N}$ , and hence the phase observable  $E_{|s\rangle}$  is not extremal by Proposition 14.

Let  $E_T$  be the phase observable generated by a state  $T = \sum_{s=0}^\infty \lambda_s |s\rangle\langle s| \in \mathcal{S}(\mathcal{H})_{\text{diag}}$ . Then  $E_T = \sum_{s=0}^\infty \lambda_s E_{|s\rangle}$  and hence,  $E_T$  is not extremal. We thus have the following conclusion.

**Proposition 15.** *Any  $E_T$ ,  $T \in \mathcal{S}(\mathcal{H})_{\text{diag}}$ , is not extremal in the convex set of all phase observables.*

For completeness, we note that the phase matrix elements of  $E_T$  are  $c_{m,n}^T = \sum_{s=0}^\infty \lambda_s c_{m,n}^{[s]} = \langle f_m^T | f_n^T \rangle$ , where  $f_n^T = \sum_{s=0}^\infty \sqrt{\lambda_s} f_n^s \otimes |s\rangle$  is a unit vector of the Hilbert space  $L^2(\mathbb{R}_+, e^{-x}dx) \otimes \mathcal{H}$ , and it follows that  $\text{rank}(c_{m,n}^T) = \infty$ .

**Remark 1.** Since the physically relevant phase observables  $E_T$  and  $E_{\text{can}}$  all have real-valued phase matrices, it would be interesting to study the extremals of the smaller convex set of real-valued phase matrices  $\mathcal{C}_{\mathbb{R}} = \{(c_{m,n}) \in \mathcal{C} \mid c_{m,n} \in \mathbb{R}, m, n \in \mathbb{N}\}$ . The complete characterization of such extremals is given in [17, Theorem 1]. The canonical phase observable is extremal in  $\mathcal{C}_{\mathbb{R}}$ , but the question of the extremality of the state generated phase observables is open. The method of the proof of Proposition 14 cannot be directly applied in this case as the operator  $|\eta_m\rangle\langle\eta_m| - |\eta_n\rangle\langle\eta_n|$  is trivial on the real Banach space  $\mathcal{T}_s(\mathcal{K})$  of selfadjoint trace class operators on  $\mathcal{K} = \overline{\text{lin}_{\mathbb{R}} \{\eta_n : n \in \mathbb{N}\}}$ .

**3.6. Postprocessing.** As explained in Section 3.4, the observables  $E_{\text{can}}$  and  $E_T$  are approximately sharp. Therefore, it follows directly from Proposition 5 that they are all postprocessing clean in  $\mathcal{O}$ . It still remains to check whether they are postprocessing equivalent. Let us start with the following observation.

**Proposition 16.** *Let  $T, T' \in \mathcal{S}(\mathcal{H})_{\text{diag}}$ . If  $E_T = E_{T'}$ , then  $T = T'$ .*

*Proof.* Let  $T = \sum_{s \in \mathbb{N}} \lambda_s |s\rangle\langle s| \in \mathcal{S}(\mathcal{H})_{\text{diag}}$  and  $T' = \sum_{s \in \mathbb{N}} \lambda'_s |s\rangle\langle s| \in \mathcal{S}(\mathcal{H})_{\text{diag}}$ . Thus,  $c_{m,n}^T = \sum_{s=0}^\infty \lambda_s c_{m,n}^{[s]}$  and  $c_{m,n}^{T'} = \sum_{s=0}^\infty \lambda'_s c_{m,n}^{[s]}$ . Assume  $E_T = E_{T'}$  so that  $c_{m,n}^T = c_{m,n}^{T'}$  for all  $m, n \in \mathbb{N}$ .

For all  $s, k \in \mathbb{N}$ ,

$$\begin{aligned} c_{0,2k}^{[s]} &= (-1)^{s+\max\{0,s-2k\}} \sqrt{\frac{(\min\{2k,s\})!}{s!(\max\{2k,s\})!}} \int_0^\infty x^{(s+|s-2k|)/2} L_{\min\{2k,s\}}^{|s-2k|}(x) e^{-x} dx \\ &= (-1)^{s+\max\{0,s-2k\}} \frac{((|s-2k|+s)/2)! \Gamma(k)}{\sqrt{(2k)!} s! \Gamma((|s-2k|-s)/2)} \end{aligned}$$

where  $\Gamma$  is the Gamma function and we have used the following equation (see formula 7.414(11) in [7]):

$$\int_0^\infty x^{\gamma-1} L_n^\alpha(x) e^{-x} dx = \frac{\Gamma(\gamma) \Gamma(1+\alpha-\gamma+n)}{n! \Gamma(1+\alpha-\gamma)}, \quad \gamma > 0, \alpha, n \in \mathbb{N}.$$

Since  $\lim_{x \rightarrow -n} |\Gamma(x)| = \infty$  for all  $n \in \mathbb{N}$ , it follows that

$$c_{0,2k}^{[s]} = 0$$

if and only if  $0 < k \leq s$ . Thus,

$$c_{0,2k}^T = \sum_{s=0}^{k-1} \lambda_s c_{0,2k}^{[s]} = \sum_{s=0}^{k-1} \lambda'_s c_{0,2k}^{[s]}.$$

By induction  $T = T'$ . Indeed,  $c_{0,2}^T = \lambda_0 c_{0,2}^{[0]} = \lambda'_0 c_{0,2}^{[0]}$  implies that  $\lambda_0 = \lambda'_0$ . If  $\lambda_s = \lambda'_s$  for all  $s = 0, 1, \dots, k-1$  then

$$(\lambda_k - \lambda'_k) c_{0,2(k+1)}^{[k]} = c_{0,2(k+1)}^T - c_{0,2(k+1)}^{T'} = 0.$$

Thus,  $\lambda_k = \lambda'_k$ . □

**Proposition 17.** *Phase observables  $E_{\text{can}}$  and  $E_T$  are postprocessing clean. They are all postprocessing non-equivalent.*

*Proof.* We need to prove the second claim. For two approximately sharp phase observables  $E$  and  $E'$  with phase matrices  $(c_{m,n})$  and  $(c'_{m,n})$ , Prop. 5 implies (taking into account that  $\mathbb{T}$  is Abelian group) that  $E \stackrel{\text{POST}}{\sim} E'$  equals  $E'(\cdot) = U(t)E(\cdot)U(t)^*$  for some  $t \in \mathbb{T}$ . By Theorem 1 this means that  $c'_{m,n} = t^{m-n} c_{m,n}$  for all  $m, n \in \mathbb{N}$ .

Let  $T, T' \in \mathcal{S}(\mathcal{H})_{\text{diag}}$  and assume that, for some  $t \in \mathbb{T}$ ,  $E_{T'}(\cdot) = U(t)E_T(\cdot)U(t)^*$ . We have seen in the proof of Prop. 13 that  $\lim_{m \rightarrow \infty} c_{m,m+k}^T = \lim_{m \rightarrow \infty} c_{m,m+k}^{T'} = 1$  for all  $k \in \mathbb{N}$ . On the other hand,  $c_{m,m+k}^{T'} = t^{-k} c_{m,m+k}^T$ , from which  $t = 1$  follows. Hence,  $E_T \stackrel{\text{POST}}{\sim} E_{T'}$  only if  $E_T = E_{T'}$ , i. e. only if  $T = T'$  by Prop. 16.

Assume then that  $E_{\text{can}} \stackrel{\text{POST}}{\sim} E_T$ . It follows that  $|c_{0,2}^T| = 1$ . This is impossible since  $c_{0,2}^T = \lambda_0 c_{0,2}^{[0]}$  (see the proof of Prop. 16) and

$$c_{0,2}^{[0]} = \int_0^\infty f_0^0(x) f_2^0(x) e^{-x} dx = \frac{1}{\sqrt{2}} \int_0^\infty x e^{-x} dx = \frac{1}{\sqrt{2}} < 1.$$

□



**3.7. Preprocessing.** Let  $\mathbf{E}$  be a phase observable determined by the phase matrix  $(c_{m,n}) \in \mathcal{C}$ . As shown in [23], we can define a  $U$ -covariant channel by formula

$$\mathcal{T}(\mathcal{H}) \ni T \mapsto \Phi_{\mathbf{E}}(T) := \sum_{m,n=0}^{\infty} c_{n,m} \langle m|T|n \rangle |m\rangle \langle n| \in \mathcal{T}(\mathcal{H}).$$

Moreover, this channel satisfies

$$\Theta_{\mathbf{E}} = \Theta_{\mathbf{E}_{\text{can}}} \circ \Phi_{\mathbf{E}}.$$

This shows that every phase observable is a preprocessing of the canonical phase observable  $\mathbf{E}_{\text{can}}$ . In particular,  $\mathbf{E}_{\text{can}}$  is preprocessing clean.

It remains to find the equivalence class of  $\mathbf{E}_{\text{can}}$  in the preprocessing relation. The preprocessing clean phase observables are exactly the phase observables belonging to this equivalence class.

**Proposition 18.** *If  $\Xi \in \text{Ch}(\mathcal{H})^*$  is such that*

$$\Xi(U(t)AU(t)^*) = U(t)\Xi(A)U(t)^*$$

*for all  $t \in \mathbb{T}$  and  $A \in \mathcal{L}(\mathcal{H})$ , then there exists a sequence of vectors  $\{\phi_i\}_{i \in \mathbb{N}}$  in  $\mathcal{H} \otimes \mathcal{H}$  with  $\|\phi_i\| = 1$  such that*

$$\langle j|\Xi(A)|i\rangle = \int t^{i-j} \langle \phi_j | (I \otimes U(t)AU(t)^*) \phi_i \rangle dt \quad \forall i, j \in \mathbb{N}.$$

*Proof.* Let  $\rho$  be the right regular representation of  $\mathbb{T}$  in  $L^2(\mathbb{T})$ . Since  $\mathbb{T}$  is compact and  $\dim \mathcal{H} = \infty$ , every separable unitary representation of  $\mathbb{T}$  is contained in the tensor product representation  $\rho \otimes I$  acting in the space  $L^2(\mathbb{T}) \otimes \mathcal{H} = L^2(\mathbb{T}; \mathcal{H})$ . Therefore,  $\mathcal{K} = L^2(\mathbb{T}; \mathcal{H})$  and  $D = \rho \otimes I$  is the most general choice in the context of Prop. 9.

We define the following unitary operator  $T : L^2(\mathbb{T}) \otimes \mathcal{H} \otimes \mathcal{H} \rightarrow L^2(\mathbb{T}) \otimes \mathcal{H} \otimes \mathcal{H}$

$$[T(f \otimes u \otimes v)](t) = f(t) \otimes u \otimes U(t)^*v.$$

Clearly,

$$T(\rho \otimes I \otimes I) = (\rho \otimes I \otimes U)T = (D \otimes U)T.$$

Therefore, every isometry  $W : \mathcal{H} \rightarrow \mathcal{K} \otimes \mathcal{H}$  intertwining  $U$  with  $D \otimes U$  is of the form  $W = T\tilde{W}$ , where  $\tilde{W}$  is an isometry intertwining  $U$  with  $\rho \otimes I \otimes I$ . The most general such  $\tilde{W}$  is given by

$$\tilde{W}|n\rangle = f_n \otimes \phi_n,$$

where  $\{\phi_i\}_{i \in \mathbb{N}}$  is a sequence of unit vectors in  $\mathcal{H} \otimes \mathcal{H}$ . By Prop. 9, every  $U$ -covariant  $\Xi \in \text{Ch}(\mathcal{H})^*$  is thus of the form

$$\begin{aligned} \langle j|\Xi(A)|i\rangle &= \langle j|\tilde{W}^*T^*(I_{\mathcal{K}} \otimes A)T\tilde{W}|i\rangle \\ &= \langle T(f_j \otimes \phi_j) | (I_{L^2(\mathbb{T})} \otimes I_{\mathcal{H}} \otimes A)T(f_i \otimes \phi_i) \rangle \\ &= \int t^{i-j} \langle \phi_j | (I \otimes U(t)AU(t)^*) \phi_i \rangle dt \end{aligned}$$

for all  $A \in \mathcal{L}(\mathcal{H})$  and  $i, j \in \mathbb{N}$ . □

**Proposition 19.** *If  $\mathbf{E}$  is a phase observable with phase matrix  $(c_{m,n})$ , then the following are equivalent:*

- (i)  $\mathbf{E} \stackrel{\text{PRE}}{\sim} \mathbf{E}_{\text{can}}$
- (ii) *there exists  $n_0 \in \mathbb{N}$  and a sequence  $\{\lambda_n\}_{n \geq n_0}$  in  $\mathbb{T}$  such that  $c_{m,n} = \bar{\lambda}_m \lambda_n$  for all  $m, n \geq n_0$ .*

*Proof.* Let  $L^\infty(\mathbb{T})$  be the Banach space of bounded measurable functions on the unit circle  $\mathbb{T}$  with the sup norm. We recall that, if  $\mathbf{E}$  is a (not necessarily covariant) POVM based on  $\mathbb{T}$  with values in  $\mathcal{H}$ , then we can define a norm decreasing linear map  $\mathbf{E} : L^\infty(\mathbb{T}) \rightarrow \mathcal{L}(\mathcal{H})$  in the following way. If  $T \in \mathcal{T}(\mathcal{H})$ , let  $p_T^{\mathbf{E}}$  be the bounded complex measure on  $\mathbb{T}$  given by

$$p_T^{\mathbf{E}}(X) = \text{tr}[T\mathbf{E}(X)] \quad \forall X \in \mathcal{B}(\mathbb{T}).$$

If  $f \in L^\infty(\mathbb{T})$ , the operator  $\mathbf{E}(f)$  is then defined by

$$\text{tr}[T\mathbf{E}(f)] = \int f(t) dp_T^{\mathbf{E}}(t) \quad \forall T \in \mathcal{T}(\mathcal{H}).$$

For all  $j \in \mathbb{Z}$ , let  $f_j(t) = t^j$ . Then

$$\begin{aligned} \mathbf{E}(f_j) &= \sum_{n \geq 0} c_{n,j+n} |n\rangle \langle j+n| \quad \text{if } j \geq 0 \\ \mathbf{E}(f_j) &= \sum_{n \geq 0} c_{n-j,n} |n-j\rangle \langle n| \quad \text{if } j \leq 0, \end{aligned}$$

the sums converging in the weak sense. Clearly,  $\mathbf{E}(f_j)^* = \mathbf{E}(f_{-j})$  for all  $j \in \mathbb{N}$ .

If  $\Xi \in \text{Ch}(\mathcal{H})^*$ , then it is easy to check that  $\Xi(\mathbf{E}(f)) = \mathbf{E}^\Xi(f)$ , where  $\mathbf{E}^\Xi$  is the POVM given by

$$\mathbf{E}^\Xi(X) = \Xi(\mathbf{E}(X)) \quad \forall X \in \mathcal{B}(\mathbb{T}).$$

Suppose  $\Xi$  is a  $U$ -covariant channel, and let  $\{\phi_i\}_{i \in \mathbb{N}}$  be a sequence of unit vectors in  $\mathcal{H} \otimes \mathcal{H}$  associated to  $\Xi$  as in Prop. 18. Then by Fourier analysis  $\mathbf{E}_{\text{can}} = \mathbf{E}^\Xi$  if and only if  $\mathbf{E}_{\text{can}}(f_j) = \mathbf{E}^\Xi(f_j)$  for all  $j \in \mathbb{N}$ .

For  $j \in \mathbb{N}$ , we have

$$\begin{aligned} \langle q | \mathbf{E}^\Xi(f_j) | p \rangle &= \int t^{p-q} \langle \phi_q | (I \otimes U(t) \mathbf{E}(f_j) U(t)^*) \phi_p \rangle dt \\ &= \int t^{p-q} \sum_{n \geq 0} c_{n,j+n} \langle \phi_q | (I \otimes U(t) |n\rangle \langle j+n| U(t)^*) \phi_p \rangle dt \\ &= \sum_{n \geq 0} c_{n,j+n} \int t^{p-q-j} \langle \phi_q | (I \otimes |n\rangle \langle j+n|) \phi_p \rangle dt \\ &= \delta_{j,p-q} \sum_{n \geq 0} c_{n,j+n} \langle \phi_q | (I \otimes |n\rangle \langle j+n|) \phi_p \rangle \end{aligned}$$

for all  $p, q \in \mathbb{N}$ . Setting

$$(19) \quad \phi_k = \sum_{r \geq 0} \phi_k^r \otimes |r\rangle$$

with  $\phi_k^r \in \mathcal{H}$  such that  $\sum_{r \geq 0} \|\phi_k^r\|^2 = 1$ , the last equation becomes

$$\langle q | \mathbf{E}^\Xi(f_j) | p \rangle = \delta_{j,p-q} \sum_{n \geq 0} c_{n,j+n} \langle \phi_q^n | \phi_p^{j+n} \rangle.$$

This must be compared with

$$\langle q | \mathbf{E}_{\text{can}}(f_j) | p \rangle = \delta_{j,p-q}.$$

The two expressions are the same if and only if

$$(20) \quad \sum_{n \geq 0} c_{n,p-q+n} \langle \phi_q^n | \phi_p^{p-q+n} \rangle = 1 \quad \text{for all } p \geq q \geq 0.$$

If  $\{\eta_n\}_{n \in \mathbb{N}}$  are unit vectors in  $\mathcal{H}$  such that  $c_{n,m} = \langle \eta_n | \eta_m \rangle$  (see [5]), eq. (20) amounts to

$$(21) \quad \sum_{n \geq 0} \langle \eta_n \otimes \phi_q^n | \eta_{p-q+n} \otimes \phi_p^{p-q+n} \rangle = 1 \quad \text{for all } p \geq q \geq 0.$$

Since

$$\sum_{n \geq 0} \|\eta_{j+n} \otimes \phi_k^{j+n}\|^2 = \sum_{n \geq 0} \|\phi_k^{j+n}\|^2 \leq 1 \quad \text{for all } j, k \in \mathbb{N},$$

by Cauchy-Schwartz inequality Eq. (21) holds if and only if

$$\eta_n \otimes \phi_q^n = \eta_{p-q+n} \otimes \phi_p^{p-q+n} \quad \forall n \in \mathbb{N} \quad \text{and} \quad \sum_{n \geq 0} \|\phi_q^n\|^2 = \sum_{n \geq 0} \|\phi_p^{p-q+n}\|^2 = 1.$$

for all  $p \geq q \geq 0$ . These two conditions are in turn equivalent to the following

- (1)  $\phi_p^n = 0$  for all  $0 \leq n < p$ , and  $\sum_{n \geq p} \|\phi_p^n\|^2 = 1$ ;
- (2)  $\eta_n \otimes \phi_0^n = \eta_{p+n} \otimes \phi_p^{p+n}$  for all  $n, p \in \mathbb{N}$ .

If conditions (1) and (2) hold, let  $n_0$  be such that  $\phi_0^{n_0} \neq 0$ . Condition (2) with  $n = n_0$  then implies  $\eta_{p+n_0} = \lambda_p \eta_{n_0}$  for all  $p \in \mathbb{N}$ , and thus  $c_{n_0+p, n_0+q} = \bar{\lambda}_p \lambda_q$  for all  $p, q \geq 0$ .

If conversely there exists  $n_0 \in \mathbb{N}$  and a sequence  $\{\lambda_n\}_{n \geq n_0}$  such that  $c_{m,n} = \bar{\lambda}_m \lambda_n$  for all  $m, n \geq n_0$ , then by Cauchy-Schwartz inequality  $\eta_n = \bar{\lambda}_{n_0} \lambda_n \eta_{n_0}$  for all  $n \geq n_0$ . Choose  $\phi_k$  in Eq. (19) such that

$$\phi_q^n = \delta_{n_0, n-q} \bar{\lambda}_n \lambda_{n_0} \phi \quad \forall n \in \mathbb{N},$$

with  $\phi \in \mathcal{H}$  such that  $\|\phi\| = 1$  (in the above formula,  $\phi_q^n = 0$  if  $n < n_0$ ). Then the sequence  $\{\phi_q^n\}_{q,n \in \mathbb{N}}$  satisfies conditions (1) and (2) above.  $\square$

A phase observable can be preprocessing clean without being extremal. Indeed, we have the following example.

**Example 2.** Let  $n_0 \geq 1$ . Define  $c_{nm} = 1$  if  $n \geq n_0$  and  $m \geq n_0$ , and  $c_{nm} = \delta_{nm}$  (Kronecker delta) otherwise. Obviously,  $(c_{nm})$  is a phase matrix and the corresponding phase observable  $\mathbf{E}$  is not unitarily equivalent with  $\mathbf{E}_{\text{can}}$ . As a consequence of Proposition 19 we conclude that  $\mathbf{E} \stackrel{\text{PRE}}{\sim} \mathbf{E}_{\text{can}}$ . However, in Proposition 5 we have shown that  $\mathbf{E} \sim_U \mathbf{E}_{\text{can}}$  equals  $\mathbf{E} \stackrel{\text{POST}}{\sim} \mathbf{E}_{\text{can}}$ . Therefore,  $\mathbf{E}$  and  $\mathbf{E}_{\text{can}}$  are not postprocessing equivalent. Note that  $c_{m,n} \in \mathbb{R}$  for all  $m, n$  and the rank of  $(c_{nm})$  is  $n_0 + 1$  (this can be seen easily by noting that the  $\eta$ -sequence of  $(c_{nm})$  can be chosen such that the first  $n_0 + 1$   $\eta_n$ 's are mutually orthonormal and  $\eta_n = \eta_{n_0+1}$  for all  $n > n_0 + 1$ ), hence  $\mathbf{E}$  is not extremal by Prop. 14. Moreover,  $\mathbf{E}$  is approximately sharp by Prop. 12 and, thus, postprocessing clean.

We have seen that  $\mathbf{E}_{\text{can}}$  is optimal in all four ways. However, there are also other phase observables sharing this feature. This is illustrated in the following example.

**Example 3.** Let  $\{f_1, f_2\}$  be the canonical basis of  $\mathbb{C}^2$ , and let  $\eta_0 = f_1$ ,  $\eta_1 = f_2$ ,  $\eta_2 = 2^{-1/2}(f_1 + f_2)$ ,  $\eta_3 = 2^{-1/2}(f_1 + if_2)$ , and  $\eta_n = f_1$  for all  $n \geq 4$ . Define the phase matrix  $c_{m,n} = \langle \eta_m | \eta_n \rangle$ , and let  $\mathbf{E}$  be the associated phase observable.

It has been shown in [17] that  $(c_{m,n})$  is a rank 2 phase matrix which is extremal in the convex set of phase matrices, hence  $\mathbf{E}$  is extremal in the convex set  $\mathcal{O}$ . Moreover,  $\lim_{m \rightarrow \infty} c_{m,m+k} = 1$  for all  $k \in \mathbb{N}$ , and therefore  $\mathbf{E}$  is approximately sharp by Proposition 12. It follows from Proposition 5 that it is also postprocessing clean. By Proposition 19 the observable  $\mathbf{E}$  is preprocessing clean. However,  $\mathbf{E} \not\sim_U \mathbf{E}_{\text{can}}$  since the ranks of the associated phase matrices are different.

**Acknowledgments.** The authors thank Jukka Kiukas and Kari Ylinen for valuable discussions.

## REFERENCES

- [1] F. Buscemi, G.M. D'Ariano, M. Keyl, P. Perinotti, and R.F. Werner. Clean positive operator valued measures. *J. Math. Phys.*, 46:082109, 2005.
- [2] P. Busch, M. Grabowski, and P.J. Lahti. *Operational Quantum Physics*. Springer-Verlag, Berlin, 1997. second corrected printing.
- [3] C. Carmeli, T. Heinosaari, J.-P. Pellonpää, and A. Toigo. Extremal covariant positive operator valued measures: the case of a compact symmetry group. *J. Math. Phys.*, 49:063504, 2008.
- [4] C. Carmeli, T. Heinosaari, and A. Toigo. Covariant quantum instruments. arXiv:0805.3917v1 [math-ph], 2008.
- [5] G. Cassinelli, E. De Vito, P. Lahti, and J.-P. Pellonpää. Covariant localizations in the torus and the phase observables. *J. Math. Phys.*, 43:693–704, 2002.
- [6] W.M. de Muynck. *Foundations of Quantum Mechanics, an Empiricist Approach*. Kluwer Academic Publishers, Dordrecht, 2002.
- [7] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Academic Press, New York, 1980. Corrected and enlarged edition edited by A. Jeffrey.
- [8] T. Heinonen. Optimal measurements in quantum mechanics. *Phys. Lett. A*, 346:77–86, 2005.
- [9] T. Heinonen, P. Lahti, J.-P. Pellonpää, S. Pulmannova, and K. Ylinen. The norm-1-property of a quantum observable. *J. Math. Phys.*, 44:1998–2008, 2003.
- [10] E. Hewitt and K.A. Ross. *Abstract Harmonic Analysis. Vol. I: Structure of Topological Groups. Integration Theory, Group Representations*. Academic Press, New York, 1963.

- [11] A.S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland Publishing Co., Amsterdam, 1982.
- [12] A.S. Holevo. Generalized imprimitivity systems for abelian groups. *Sov. Math. (Iz. VUZ)*, 27:53–80, 1983.
- [13] A.S. Holevo. *Statistical Structure of Quantum Theory*. Springer-Verlag, Berlin, 2001.
- [14] A. Jenčová and S. Pulmannová. How sharp are PV measures? *Rep. Math. Phys.*, 59:257–266, 2007.
- [15] A. Jenčová, S. Pulmannová, and E. Vinceková. Sharp and fuzzy observables on effect algebras. *Int. J. Theor. Phys.*, 47:125–148, 2008.
- [16] J. Kiukas and P. Lahti. On the moment limit of quantum observables, with an application to the balanced homodyne detection. *J. Mod. Opt.*, 55:1175 – 1198, 2008.
- [17] J. Kiukas and J.-P. Pellonpää. A note on infinite extreme correlation matrices. *Linear Algebra Appl.*, 428:2501–2508, 2008.
- [18] P. Lahti and J.-P. Pellonpää. Covariant phase observables in quantum mechanics. *J. Math. Phys.*, 40:4688–4698, 1999.
- [19] P. Lahti and J.-P. Pellonpää. Characterizations of the canonical phase observable. *J. Math. Phys.*, 41:7352–7381, 2000.
- [20] P. Lahti and J.-P. Pellonpää. The Pegg-Barnett formalism and covariant phase observables. *Physica Scripta*, 66:66–70, 2002.
- [21] H. Martens and W.M. de Muynck. Nonideal quantum measurements. *Found. Phys.*, 20:255–281, 1990.
- [22] J.-P. Pellonpää. Phase observables, phase operators and operator orderings. *J. Phys. A*, 34:7901–7916, 2001.
- [23] J.-P. Pellonpää. On the structure of covariant phase observables. *J. Math. Phys.*, 43:1299–1308, 2002.
- [24] W. Rudin. Measure algebras on abelian groups. *Bull. Amer. Math. Soc.*, 65:227–247, 1959.
- [25] N.G. Walker and J.E. Carrol. Multiport homodyne detection near the quantum noise limit. *Opt. Quant. Electron.*, 18:355–363, 1986.
- [26] J.G. Wendel. Left centralizers and isomorphisms of group algebras. *Pacific J. Math.*, 2:251–261, 1952.

CLAUDIO CARMELI, DIPARTIMENTO DI FISICA, UNIVERSITÀ DI GENOVA, AND I.N.F.N., SEZIONE DI GENOVA, VIA DODECANESO 33, 16146 GENOVA, ITALY

*E-mail address:* `carmeli@ge.infn.it`

TEIKO HEINOSAARI, DEPARTMENT OF PHYSICS, UNIVERSITY OF TURKU, FIN-20014 TURKU, FINLAND

*E-mail address:* `heinosaari@gmail.com`

JUHA-PEKKA PELLONPÄÄ, DEPARTMENT OF PHYSICS, UNIVERSITY OF TURKU, FIN-20014 TURKU, FINLAND

*E-mail address:* `juhpeppo@utu.fi`

ALESSANDRO TOIGO, DIPARTIMENTO DI INFORMATICA, UNIVERSITÀ DI GENOVA, VIA DODECANESO 35, AND I.N.F.N., SEZIONE DI GENOVA, VIA DODECANESO 33, 16146 GENOVA, ITALY

*E-mail address:* `toigo@ge.infn.it`