JOURNAL OF MATHEMATICAL PHYSICS VOLUME 45, NUMBER 12 DECEMBER 2004

# Extremal covariant positive operator valued measures

Giulio Chiribella[a)]
*QUIT Group, Istituto Nazionale di Fisica della Materia, Unità di Pavia, Dipartimento di Fisica "A. Volta," via Bassi 6, I-27100 Pavia, Italy*

Giacomo Mauro D'Ariano[b)]
*QUIT Group, Istituto Nazionale di Fisica della Materia, Unità di Pavia, Dipartimento di Fisica "A. Volta," via Bassi 6, I-27100 Pavia, Italy, and Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois 60208*

(Received 30 June 2004; accepted 27 July 2004; published 4 November 2004)

We consider the convex set of positive operator valued measures (POVM) which are covariant under a finite dimensional unitary projective representation of a group. We derive a general characterization for the extremal points, and provide bounds for the ranks of the corresponding POVM densities, also relating extremality to uniqueness and stability of optimized measurements. Examples of applications are given. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1806262]

## I. INTRODUCTION

An essential step in the design of the new quantum information technology[1] is to assess the ultimate precision limits achievable by quantum measurements in extracting information from physical systems. For example, the security analysis of a quantum cryptographic protocol[2] is based on the evaluation of the limits posed in principle by the quantum laws to any possible eavesdropping strategy. A general method to establish such limits is to optimize a quantum measurement according to a suitable criterion, and this is the general objective of the so-called *quantum estimation theory*.[3,4] Different criteria can be adopted for optimizing the measurement, the choice of a particular one depending on the particular problem at hand. Moreover, many different optimization problems often share the same form, e.g., they resort to the maximization of a concave function on the set of the possible measurements. We remind that measurements form a convex set, the convex combination corresponding to the random choice between two different apparatuses. Since a concave function attains its maximum in an extremal point, it is clear that the optimization problem is strictly connected to the problem of characterizing the extremal points of the convex set.

The quantum measurements interesting in most applications are *covariant*[4] with respect to a group of physical transformations. In a purely statistical description of a quantum measurement in terms of the outcome probability only—i.e., without considering the state-reduction—the measurement is completely described by a positive operator valued measure (POVM) on its probability space. In terms of POVM's, "group covariant" means that there is an action of the transformation group on the probability space which maps events into events, in such a way that when the measured system is transformed according to a group transformation, the probability of a given event becomes the probability of the transformed event. Such a scenario naturally occurs in the estimation of an unknown group transformation performed on a known input state, e.g., in the estimation of an unknown unitary transformation,[5,6] in the measurement of a phase shift in the radiation field,[4,7] or in the estimation of rotations on a system of spins.[8] A first technique for characterizing extremal covariant POVM's and quantum operations has been presented in Ref. 9

[a)]Electronic mail: chiribella@unipv.it; http://www.qubit.it
[b)]Electronic mail: dariano@unipv.it; http://www.qubit.it

inspired by the method for characterizing extremal correlation matrices of Ref. 10, in particular, classification of extremal POVM's has been presented for the case of trivial stability group, i.e., when the only transformation which leaves the input state unchanged is the identity. Here we solve the characterization problem for extremal covariant POVM's in the general case of nontrivial stability group, providing a simple criterion for extremality in Theorem 1 in terms of minimality of the support of the *seed* of the POVM, presenting iff conditions for extremality in Theorem 3, and providing bounds for the rank of extremal POVM's [in the following we will define the rank of a POVM as the rank of its respective density: see Eq. (6) for its definition]. We show that, contrarily to the usual credo, the optimal covariant POVM can have rank larger than one. Indeed, there are group representations for which covariant POVM cannot have unit rank, since this would violate a general bound for the rank of the POVM in relation to dimensions and multiplicity of the invariant subspaces of the group. In the present paper we adopt the maximum likelihood optimality criterion, which, however, as we will show, is formally equivalent to the solution of the optimization problem in a very large class of optimality criteria. Other issues of practical interest that we address are the uniqueness and the stability of the optimal covariant POVM. The whole derivation is given for finite dimensional Hilbert spaces: as we will show in a simple example, it can be generalized to infinite dimensions, however, at the price of making the theory much more technical.

The paper is organized as follows. After introducing covariant POVM's and their convex structure in Sec. II, the main group theoretical tools that will be used for the characterization of covariant POVM's are presented in Sec. III. In Sec. IV we give a characterization of extremal covariant POVM's in finite dimension with a general stability group, deriving an algebraic extremality criterion, along with a general bound for the rank of the extremal POVM's in terms of the dimensions of the invariant subspaces of the group and of the stability subgroup. Properties of extremal POVM's in relation with optimization problems are analyzed in Sec. V, where also the issues of uniqueness and stability of the optimal covariant POVM's are addressed. Finally, examples of application of the theory to estimation of rotation, state, phase shift, etc., are given in Sec. VI, providing extremal POVM's with a nontrivial stability group and giving examples of optimization problems with solution consisting of extremal POVM with rank greater than one.

## II. CONVEX STRUCTURE OF COVARIANT POVM'S

The general description of the statistics of a measurement is given in terms of a probability space $\mathfrak{X}$—the set of all possible measurement *outcomes*—equipped with a $\sigma$-algebra $\sigma(\mathfrak{X})$ of subsets $\mathsf{B}\subseteq\mathfrak{X}$ and with a probability measure $p$ on $\sigma(\mathfrak{X})$. Each subset $\mathsf{B}\in\sigma(\mathfrak{X})$ describes the event "the outcome $x$ belongs to $\mathsf{B}$" and the statistics of the measurement is fully specified by the probability measure $p$, which associates to any event $\mathsf{B}$ its probability $p(\mathsf{B})$.

In quantum mechanics the probability $p(\mathsf{B})$ is given by the Born rule,

$$p(\mathsf{B}) \doteq \mathrm{Tr}[\rho P(\mathsf{B})], \tag{1}$$

where $\rho$ is a density operator (i.e., a positive semidefinite operator with unit trace) on the Hilbert space $\mathcal{H}$ of the measured system, representing its state, whereas $P$ is the POVM of the apparatus, giving the probability measure $p$ for every given state $\rho$ of the quantum system. Mathematically a POVM $P:\sigma(\mathfrak{X})\to\mathcal{B}(\mathcal{H})$ is a *positive operator valued measure* on $\sigma(\mathfrak{X})$, namely it satisfies the following defining properties:

$$0 \leqslant P(\mathsf{B}) \leqslant I \quad \forall\, \mathsf{B} \in \sigma(\mathfrak{X}), \tag{2}$$

$$P(\cup_{i=1}^{\infty}\mathsf{B}_i) = \sum_{i=1}^{\infty} P(\mathsf{B}_i) \quad \forall\,\{\mathsf{B}_i\}\ \text{disjoint}, \tag{3}$$

$$P(\mathfrak{X}) = I. \tag{4}$$

Notice that the set of POVM's for $\sigma(\mathfrak{X})$ is a convex set, namely, if $P_1$ and $P_2$ are POVM's for $\sigma(\mathfrak{X})$, then also $\lambda P_1+(1-\lambda)P_2$ is a POVM for $\sigma(\mathfrak{X})$ for any $0\leqslant\lambda\leqslant 1$. The measurement described by the POVM $\lambda P_1+(1-\lambda)P_2$ corresponds to randomly choosing between two different measuring apparatuses described by the POVM's $P_1$ and $P_2$, respectively. The extremal points of such convex set of POVM's—the so-called *extremal POVM's*—correspond to measurements that cannot result from a random choice between different measuring apparatuses.

In the following we will focus attention to the case of probability space $\mathfrak{X}$ given by the quotient $\mathbf{G}/\mathbf{G}_0$ of a compact Lie group $\mathbf{G}$ with respect to a subgroup $\mathbf{G}_0$. Physically, this situation arises when the POVM is designed to estimate a state in the group-orbit $\{U_g\rho U_g^\dagger | g\in\mathbf{G}\}$ of a given state $\rho$, with the group $\mathbf{G}$ acting on the Hilbert space $\mathcal{H}$ of a quantum system via the unitary projective representation $\mathsf{R}(\mathbf{G})\doteq\{U_g | g\in\mathbf{G}\}$. In such a case, in fact, the probability space of the POVM is exactly $\mathfrak{X}=\mathbf{G}/\mathbf{G}_0$, and $\mathbf{G}_0=\{h\in\mathbf{G} | U_h\rho U_h^\dagger=\rho\}$ is the stability group of $\rho$, whence the points of the orbit are in one-to-one correspondence with the elements of $\mathfrak{X}=\mathbf{G}/\mathbf{G}_0$. Notice that in the following the fact that the representation is projective is inconsequential, whence there will be no need for reminding.

An important class of measurements with $\mathfrak{X}=\mathbf{G}/\mathbf{G}_0$ is described by the *covariant* POVM's,[4] namely those POVM's which enjoy the property

$$P(g\mathsf{B}) = U_g P(\mathsf{B}) U_g^\dagger \quad \forall\, \mathsf{B}\in\sigma(\mathfrak{X}), \quad \forall\, g\in\mathbf{G}, \tag{5}$$

where $g\mathsf{B}\doteq\{gx | x\in\mathsf{B}\}$. Any POVM $P$ in this class is absolutely continuous with respect to the measure $\mathrm{d}x$ induced on $\mathfrak{X}$ by the normalized Haar measure $\mathrm{d}g$ on the group $\mathbf{G}$, and admits an operator density $M$, namely

$$M:\mathfrak{X}\to\mathcal{B}(\mathcal{H}), \quad P(\mathsf{B}) = \int_{\mathsf{B}} \mathrm{d}x\, M(x). \tag{6}$$

For a covariant POVM, the operator density has the form[4]

$$M(x) = U_{g(x)}\Xi U_{g(x)}^\dagger, \tag{7}$$

where $g(x)\in\mathbf{G}$ is any element in the equivalence class $x\in\mathfrak{X}=\mathbf{G}/\mathrm{G}_0$, and $\Xi$ is an Hermitian operator satisfying the constraints

$$\Xi\geqslant 0, \quad \int_{\mathbf{G}} \mathrm{d}g\, U_g\Xi U_g^\dagger = I, \tag{8}$$

$$[\Xi, U_h] = 0 \quad \forall\, h\in\mathbf{G}_0. \tag{9}$$

The operator $\Xi$ is usually referred to as the *seed* of the covariant POVM.[11]

Notice that the constraints (8) are needed for positivity and normalization of the probability density, whereas identity (9) guarantees that $M(x)=U_{g(x)}\Xi U_{g(x)}^\dagger$ does not depend on the particular element $g(x)$ in the equivalence class $x$. It is easy to see that the constraints (8) and (9) still define a convex set $\mathsf{C}$, namely, for any $\Xi_1,\Xi_2\in\mathsf{C}$ and for any $0\leqslant\lambda\leqslant 1$ one has $\lambda\Xi_1+(1-\lambda)\Xi_2\in\mathsf{C}$. Precisely, the convex set $\mathsf{C}$ is the intersection of the cone of positive semidefinite operators with the two affine hyperplanes given by identity (9) and by the normalization condition in Eq. (8). Since a covariant POVM is completely specified by its seed $\Xi$ as in Eq. (7), the classification of the the extremal covariant POVM's resorts to the classification of the extremal points in the convex set $\mathsf{C}$.

## III. GROUP THEORETIC TOOLS

Let $\mathbf{G}$ be a compact Lie group, with invariant Haar measure $\mathrm{d}g$ normalized as $\int_{\mathbf{G}}\mathrm{d}g=1$, and consider a unitary representation $\mathsf{R}(\mathbf{G})=\{U_g | g\in\mathbf{G}\}$ on a finite dimensional Hilbert space $\mathcal{H}$. Then $\mathcal{H}$ is decomposed as direct sum of orthogonal irreducible subspaces as follows:

$$\mathcal{H}=\bigoplus_{\mu\in\mathsf{S}}\bigoplus_{i=1}^{m_\mu}\mathcal{H}_i^{(\mu)}, \tag{10}$$

$\mathsf{S}$ denoting the collection of equivalence classes of irreducible components of the representation, the classes being labeled by the greek index $\mu$, whereas the italic index $i$ numbers equivalent representations in the same class. Let $T_{ij}^{(\mu)}:\mathcal{H}_j^{(\mu)}\to\mathcal{H}_i^{(\mu)}$ denote invariant isomorphisms connecting the irreducible representations of the equivalence class $\mu$ of dimension $d_\mu$, namely for any $i,j=1,\ldots,m_\mu$ $T_{ij}^{(\mu)}:\mathcal{H}_j^{(\mu)}\to\mathcal{H}_i^{(\mu)}$ is an invertible operator satisfying the identity

$$U_gT_{ij}^{(\mu)}U_g^\dagger=T_{ij}^{(\mu)},\quad \forall\, g\in\mathbf{G}. \tag{11}$$

Consistently with this notation $T_{ii}^{(\mu)}$ will denote the projection operator on $\mathcal{H}_i^{(\mu)}$. Since all subspaces $\mathcal{H}_i^{(\mu)}$ are isomorphic, we can equivalently write

$$\bigoplus_{i=1}^{m_\mu}\mathcal{H}_i^{(\mu)}\equiv\mathcal{H}_\mu\otimes\mathcal{M}_\mu, \tag{12}$$

where $\mathcal{H}_\mu$ denotes the *representation space*, i.e., an abstract $d_\mu$-dimensional subspace where a representation of the class $\mu$ acts, while $\mathcal{M}_\mu$ denotes the *multiplicity space*, i.e., a $m_\mu$-dimensional space which is unaffected by the action of the group. In this way, the decomposition (10) can be written in the Wedderburn's form,[12]

$$\mathcal{H}=\bigoplus_{\mu\in\mathsf{S}}\mathcal{H}_\mu\otimes\mathcal{M}_\mu. \tag{13}$$

Due to Schur lemmas, an operator $O$ in the commutant of the representation $\mathsf{R}(\mathbf{G})$ can be decomposed as follows:[13]

$$O=\sum_\mu\sum_{i,j=1}^{m_\mu}\frac{\mathrm{Tr}[T_{ji}^{(\mu)}O]}{d_\mu}T_{ij}^{(\mu)}, \tag{14}$$

whereas, in terms of the decomposition (13) one has

$$O=\oplus_{\mu\in\mathsf{S}}(I_\mu\otimes O_\mu), \tag{15}$$

$I_\mu$ denoting the identity on the representation space $\mathcal{H}_\mu$, and $O_\mu\in\mathcal{B}(\mathcal{M}_\mu)$ being a suitable set of operators on the multiplicity spaces $\mathcal{M}_\mu$.

In this paper we will consider covariant POVM's with $\mathfrak{X}=\mathbf{G}/\mathbf{G}_0$ where both $\mathbf{G}$ and $\mathbf{G}_0$ are compact Lie groups, represented on the Hilbert space $\mathcal{H}$ by the unitary representations $\mathsf{R}(\mathbf{G})=\{U_g|g\in\mathbf{G}\}$ and $\mathsf{R}(\mathbf{G}_0)=\{U_h|h\in\mathbf{G}_0\}$. We will denote with $\mathsf{S}$ and $\mathsf{S}_0$ the equivalence classes of irreducible representations of $\mathsf{R}(\mathbf{G})$ and $\mathsf{R}(\mathbf{G}_0)$, respectively. The constraints (8) and (9) can be rewritten in a remarkably simple form using the decompositions of $\mathcal{H}$ in irreducible subspaces under the action of $\mathbf{G}$ and $\mathbf{G}_0$. In fact, due to the invariance of the Haar measure $\mathrm{d}g$, the integral in (8) belongs to the commutant of $\mathsf{R}(\mathbf{G})$. Rewriting the constraint (8) by using (14), one obtains easily,

$$\mathrm{Tr}[T_{ij}^{(\mu)}\Xi]=d_\mu\delta_{ij},\quad\forall\,\mu\in\mathsf{S},\quad\forall\, i,j=1,\ldots,m_\mu. \tag{16}$$

Moreover, according to (8) and (9), the operator $\Xi$ must be a positive semidefinite operator in the commutant of $R(\mathbf{G}_0)$ (9), then we have

$$\Xi = \oplus_{\nu \in \mathsf{S}_0} (I_\nu \otimes X_\nu^\dagger X_\nu), \tag{17}$$

where $X_\nu$ is an operator on the multiplicity subspace $\mathcal{M}_\nu$.

## IV. EXTREMAL COVARIANT POVM'S WITH A NONTRIVIAL STABILITY GROUP

In this section we will classify the extremal points of the convex set $\mathsf{C}$ of covariant seeds, namely the convex set of operators that satisfy both conditions (8) and (9). For the characterization of the extremal points of a convex set we will use the well-known method of perturbations. We will say that the operator $\Theta \in \mathcal{B}(\mathcal{H})$ is a "perturbation" of a given $\Xi \in \mathsf{C}$ if and only if there exists an $\epsilon > 0$ such that $\Xi + t\Theta \in \mathsf{C}$ for any $t \in [-\epsilon, \epsilon]$. With such definition one has that an operator $\Xi$ is extremal if and only if its unique perturbation is the trivial one, namely if $\Theta$ is a perturbation of $\Xi$ then $\Theta = 0$.

Let us start with a simple lemma which is useful for the characterization of the perturbations of a given seed $\Xi$.

*Lemma 1: Let $\Xi \in \mathcal{B}(\mathcal{H})$ be a positive semidefinite operator. Then, for any Hermitian $\Theta \in \mathcal{B}(\mathcal{H})$ the condition*

$$\exists \epsilon > 0: \quad \forall\, t \in [-\epsilon, \epsilon]\ \Xi + t\Theta \geqslant 0 \tag{18}$$

*is equivalent to*

$$\mathsf{Supp}(\Theta) \subseteq \mathsf{Supp}(\Xi). \tag{19}$$

*Proof:* Suppose that the condition (18) holds. Then for any $|\phi\rangle \in \mathsf{Ker}(\Xi)$ one necessarily has $\langle \phi | \Theta | \phi \rangle = 0$. Therefore, for any vector $|\psi\rangle \in \mathcal{H}$ one has

$$|\langle \psi | \Theta | \phi \rangle| = \frac{1}{t} |\langle \psi | (\Xi + t\Theta) | \phi \rangle| \leqslant \frac{1}{t} \sqrt{\langle \psi | (\Xi + t\Theta) | \psi \rangle \langle \phi | (\Xi + t\Theta) | \phi \rangle} = 0.$$

Hence $\mathsf{Ker}(\Xi) \subseteq \mathsf{Ker}(\Theta)$, implying that $\mathsf{Supp}(\Theta) \subseteq \mathsf{Supp}(\Xi)$. Conversely, suppose that (19) holds. Let us denote by $\lambda$ the smallest nonzero eigenvalue of $\Xi$ and by $\|\Theta\|$ the norm of $\Theta$, then condition (18) holds with $\epsilon = \lambda / \|\Theta\|$. ■

Using the previous lemma we can state that an Hermitian operator $\Theta$ is a perturbation for a given seed $\Xi$ if and only if the following conditions are satisfied:

$$\mathsf{Supp}(\Theta) \subseteq \mathsf{Supp}(\Xi), \tag{20}$$

$$\mathrm{Tr}[\Theta T_{ij}^{(\mu)}] = 0 \quad \forall\, \mu \in \mathsf{S}, \quad \forall\, i,j = 1, \ldots, m_\mu, \tag{21}$$

$$[\Theta, U_h] = 0 \quad \forall\, h \in \mathbf{G}_0 \tag{22}$$

[conditions (21) and (22) follow directly from the normalization constraints (16) and (17)].

This set of conditions leads to an interesting property of extremal seeds.

**Theorem 1:** *$\Xi$ is an extremal point of $\mathsf{C}$ if and only if for any $\zeta \in \mathsf{C}$ one has*

$$\mathsf{Supp}(\zeta) \subseteq \mathsf{Supp}(\Xi) \Rightarrow \zeta = \Xi. \tag{23}$$

*Proof:* To prove necessity it is sufficient to define $\Theta \doteq \Xi - \zeta$ and note that it is a perturbation of $\Xi$. In fact, $\Theta$ is in the commutant of $\mathsf{R}(\mathbf{G}_0)$, $\mathsf{Supp}(\Theta) \subseteq \Xi$, and $\mathrm{Tr}[\Theta T_{ij}^{\mu}] = 0\ \forall \mu \in \mathsf{S}$, $\forall i$, $j = 1, \ldots, m_\mu$. But, since $\Xi$ is extremal, then $\Theta$ must be zero.

Vice versa, assume (23). If $\Theta$ is a perturbation for $\Xi$, then there exists some $t \neq 0$ such that $\zeta \doteq \Xi + t\Theta \in \mathsf{C}$. But a perturbation must satisfy (19), then $\mathsf{Supp}(\zeta) \subseteq \mathsf{Supp}(\Xi)$. Using (23) it is then clear that $\Theta = t^{-1}(\zeta - \Xi) = 0$. ■

The proposition tells us that extremal seeds have "minimal support," in the sense that there is no element $\zeta \in \mathsf{C}$ with $\mathsf{Supp}(\zeta) \subseteq \mathsf{Supp}(\Xi)$ which is different from $\Xi$.

**Theorem 2:** *Let be $\Xi\in\mathsf{C}$. Write $\Xi$ in the form* (17). *Then an operator $\Theta$ is a perturbation of $\Xi$ if and only if*

$$\mathrm{Tr}[\Theta T_{ij}^{(\mu)}]=0 \quad \forall\,\mu\in\mathsf{S},\quad \forall\, i,j=1,\dots,m_\mu \tag{24}$$

*and $\Theta$ can be written as follows:*

$$\Theta=\oplus_{\nu\in\mathsf{S}_0}(I_\nu\otimes X_\nu^\dagger A_\nu X_\nu), \tag{25}$$

*with $X_\nu\in\mathcal{B}(\mathcal{M}_\nu)$ and $A_\nu\in\mathcal{B}(\mathsf{Rng}(X_\nu))$ Hermitian $\forall\,\nu\in\mathsf{S}_0$.*

*Proof:* Suppose $\Theta$ is a perturbation. Condition (21) is the same as (24). Due to the condition (22), $\Theta$ must be an Hermitian operator in the commutant of $\mathsf{R}(\mathbf{G}_0)$, then we can write it in the block form $\Theta=\oplus_{\nu\in\mathsf{S}_0}(I_\nu\otimes O_\nu)$, with each $O_\nu\in\mathcal{B}(\mathcal{M}_\nu)$ Hermitian. Moreover, condition (20) along with (17) imply that each operator $O_\nu$ must have $\mathsf{Supp}(O_\nu)\subseteq\mathsf{Supp}(X_\nu^\dagger X_\nu)=\mathsf{Supp}(X_\nu)$. Using the singular value decomposition $X_\nu=\Sigma_{i=1}^{r_\nu}\lambda_i^{(\nu)}|w_i^{(\nu)}\rangle\langle v_i^\nu|$ [$\{|v_i^\nu\rangle\}$ and $\{|w_i^{(\nu)}\rangle\}$ are orthonormal bases for $\mathsf{Supp}(X_\nu)$ and $\mathsf{Rng}(X_\nu)$, respectively] one can see that any Hermitian operator $O_\nu$ with $\mathsf{Supp}(O_\nu)\subseteq\mathsf{Supp}(X_\nu)$ admit the decomposition $O_\nu=X_\nu^\dagger A_\nu X_\nu$, with $A_\nu$ Hermitian operator in $\mathcal{B}(\mathsf{Rng}(X_\nu))$. Conversely, if both conditions (24) and (25) hold, then conditions (20)–(22) are obviously fulfilled. ■

**Theorem 3:** *Let $P_\nu$ be the projection operator onto the subspace $\mathcal{H}_\nu\otimes\mathcal{M}_\nu\subseteq\mathcal{H}$ corresponding to the class $\nu\in\mathsf{S}_0$. An operator $\Xi\in\mathsf{C}$ written in the form $\Xi=\oplus_{\nu\in\mathsf{S}_0}(I_\nu\otimes X_\nu^\dagger X_\nu)$ is extremal if and only if*

$$\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))=\mathsf{Span}\{F_{ij}^{(\mu)}|\mu\in\mathsf{S},\ i,j=1,\dots,m_\mu\}, \tag{26}$$

*where*

$$F_{ij}^{(\mu)}\doteq\oplus_{\nu\in\mathsf{S}_0}X_\nu\,\mathrm{Tr}_{\mathcal{H}_\nu}[P_\nu T_{ij}^{(\mu)}P_\nu]X_\nu^\dagger.$$

*Proof:* Using the characterization of Theorem 2, we know that $\Xi$ is extremal if and only if for any operator $\Theta$ satisfying (24) and (25) one has $\Theta=0$. Let us take $\Theta$ in the form (25), and rewrite the direct sum as an ordinary sum

$$\Theta=\sum_{\nu\in\mathsf{S}_0}P_\nu(I_\nu\otimes X_\nu^\dagger A_\nu X_\nu)P_\nu, \tag{27}$$

using the projectors $P_\nu$ onto $\mathcal{H}_\nu\otimes\mathcal{M}_\nu$. Using invariance of trace under cyclic permutations, we can write

$$\mathrm{Tr}[\Theta T_{ij}^{(\mu)}]=\sum_{\nu\in\mathsf{S}_0}\mathrm{Tr}[(I_\nu\otimes A_\nu)(I_\nu\otimes X_\nu)P_\nu T_{ij}^{(\mu)}P_\nu(I_\nu\otimes X_\nu^\dagger)]=\sum_{\nu\in\mathsf{S}_0}\mathrm{Tr}[A_\nu X_\nu\mathrm{Tr}_{\mathcal{H}_\nu}[P_\nu T_{ij}^{(\mu)}P_\nu]X_\nu^\dagger]. \tag{28}$$

Define the space $\mathcal{R}\doteq\oplus_{\nu\in\mathsf{S}_0}\mathsf{Rng}(X_\nu)$ and denote as $\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))$ the linear space of operators acting on $\mathcal{R}$ which are block diagonal on the subspaces $\mathsf{Rng}(X_\nu)$, $\nu\in\mathsf{S}_0$. Then, the extremality condition for $\Xi$ becomes: for any Hermitian operator $A\in\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))$ one has

$$\mathrm{Tr}[AF_{ij}^{(\mu)}]=0\ \ \forall\,\mu\in\mathsf{S},\quad\forall\, i,j=1,\dots,m_\mu\Rightarrow A=0. \tag{29}$$

In terms of the Hilbert–Schmidt product $(A,B)\doteq\mathrm{Tr}[A^\dagger B]$ this condition says that the unique Hermitian operator $A\in\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))$ which is orthogonal to the whole set of operators $\mathsf{F}\doteq\{F_{ij}^{(\mu)}|\mu\in\mathsf{S},\ i,j=1,\dots,m_\mu\}$ is the null operator. Orthogonality to the set $\mathsf{F}$ is equivalent to orthogonality to the set of Hermitian operators $\mathsf{F}'=\{(F_{ij}^{(\mu)}+F_{ji}^{(\mu)}),i(F_{ij}^{(\mu)}-F_{ji}^{(\mu)})|\mu\in\mathsf{S},\ i,j=1,\dots,m_\mu\}$. Such orthogonality holds if and only if $\mathsf{F}'$ is a spanning set for the real space of Hermitian operators in $\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu)$. Nevertheless, using the Cartesian decomposition we see that any complex block operator $O\in\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))$ can be written as a sum of two Hermitian ones, whence the extremality condition is equivalent to $\mathsf{Span}(\mathsf{F}')=\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))$ . Finally, the observation $\mathsf{Span}(\mathsf{F}')=\mathsf{Span}(\mathrm{F})$ completes the proof. ■

Notice that for trivial stability group $\mathbf{G}_0=\{e\}$ ($e$ denotes the identity element), we recover the characterization of Ref. 9: there, one has indeed a single equivalence class $\overline{\nu}$ in $\mathsf{S}_0$ with one-dimensional representation space $\mathcal{H}_{\overline{\nu}}$, so that the whole Hilbert space $\mathcal{H}$ is isomorphic to the multiplicity space $\mathcal{M}_{\overline{\nu}}$ and the extremality condition (26) reduces to $\mathsf{Span}\{XT_{ij}^{(\mu)}X^\dagger|\mu\in\mathsf{S},\ i,j=1,\ldots,m_\mu\}=\mathcal{B}(\mathsf{Rng}(X))$.

*Corollary 1: Any rank-one seed is extremal.*

*Proof:* Let $\Xi$ be a rank-one seed. In this case there is only one class $\nu_0$ in the decomposition (17) of $\Xi$ (otherwise $\Xi$ could not have unit rank), and the space $\mathcal{B}(\mathsf{Rng}(X_{\nu_0}))$ to be spanned is one dimensional, whence the condition (26) is always satisfied. ■

An alternative proof of Corollary 1 follows by observing that any rank-one element of the cone $\mathsf{D}$ of positive semidefinite operators is necessarily extremal for such cone: since the convex set $\mathsf{C}$ is a subset of $\mathsf{D}$, a rank-one seed $\Xi\in\mathsf{C}$ is necessarily an extreme point of $\mathsf{C}$.

*Corollary 2: Let $\Xi\in\mathsf{C}$ be an extremal seed and write it in the form $\Xi=\oplus_{\nu\in\mathsf{S}_0}(I_\nu\otimes X_\nu^\dagger X_\nu)$. Define $r_\nu\doteq\mathrm{rank}(X_\nu)$. Then*

$$\sum_{\nu\in\mathsf{S}_0} r_\nu^2\leqslant\sum_{\mu\in\mathsf{S}} m_\mu^2. \tag{30}$$

*Proof:* This relation follows directly from the extremality condition by noting that the left-hand side is the dimension of the complex linear space of block operators $\oplus_{\nu\in\mathsf{S}_0}\mathcal{B}(\mathsf{Rng}(X_\nu))$, while the right-hand side is the cardinality of the spanning set $\mathsf{F}=\{F_{ij}^{(\mu)}|\mu\in\mathsf{S},\ i,j=1,\ldots,m_\mu\}$.■

In Sec. VI we will see an explicit example of extremal POVM which achieves this bound.

## V. EXTREMAL POVM'S AND OPTIMIZATION PROBLEMS

A crucial step in a quantum estimation approach is the optimization of the estimation strategy for a given figure of merit. This consists in finding the POVM which maximizes some linear (more generally concave) functional $\mathcal{F}$—e.g., the average fidelity of the estimated state with the true one. Then, the convex structure of the set of POVM's plays a fundamental role in this problem, since, due to concavity of $\mathcal{F}$, one can restrict the optimization procedure to the extremal POVM's only.

In the covariant case, the problem resorts to optimize the state estimation in the orbit $\{U_g\rho U_g^\dagger|g\in\mathbf{G}\}\simeq\mathbf{G}/\mathbf{G}_0$ of a given state $\rho$ under the action of a group $\mathbf{G}$, $\mathbf{G}_0$ being the stability group of $\rho$. The optimization typically is the maximization of a linear functional corresponding to the average value of a positive function $f(x,x_*)$, where the average is taken over all the couples $(x,x_*)$ of measured and true values $x$, $x_*\in\mathcal{X}\doteq\mathbf{G}/\mathbf{G}_0$, respectively. The joint probability density $p(x,x_*)$ is connected to the conditional density $p(x|x_*)$ given by the Born rule via Bayes, assuming an *a priori* probability distribution of the true value $x_*$. In the covariant problem the function $f$ enjoys the invariance property $f(gx,gx_*)=f(x,x_*)$ $\forall g\in\mathbf{G}$, and is taken as a decreasing function of the distance $|x-x_*|$ of the measured value $x$ from the true one $x_*$. In the case of compact $\mathbf{G}$ one can assume a uniform *a priori* distribution for $x_*$ values, so that the functional corresponding to the average can be written as follows:

$$\mathcal{F}_\rho[\Xi]=\int_{\mathbf{G}}\mathrm{d}g\int_{\mathbf{G}}\mathrm{d}g_*\,f(gx_0,g_*x_0)\mathrm{Tr}[U_{g_*}\rho U_{g_*}^\dagger U_g\Xi U_g^\dagger] \tag{31}$$

$$=\int_{\mathbf{G}}\mathrm{d}g\,f(x_0,gx_0)\mathrm{Tr}[U_g\rho U_g^\dagger\Xi], \tag{32}$$

where $x_0$ is the equivalence class containing the identity. In the following, we will consider as the prototype optimization problem the maximization of the likelihood functional[3,4]

$$\mathcal{L}_\rho[\Xi] \doteq \mathrm{Tr}[\rho\Xi], \tag{33}$$

corresponding to the choice $f(x,x_*)=\delta(x-x_*)$ in Eq. (31). Maximizing $\mathcal{L}_\rho[\Xi]$ means maximizing the probability density that the measured value $x$ coincides with the true value $x_*$. For such estimation strategy the optimization problem has a remarkably simple form, enabling a general treatment for a large class of group representations.[13] Moreover, the solution of the maximum likelihood is formally equivalent to the solution of any optimization problem with a positive (which, a part from an additive constant, means bounded from below) summable function $f(x,x_*)$. Indeed, we can define the map

$$\mathcal{M}(\rho) = k^{-1}\int_{\mathbf{G}} \mathrm{d}g\, f(x_0,gx_0)U_g\rho U_g^\dagger, \tag{34}$$

where $k=\int_{\mathbf{G}}\mathrm{d}g\, f(x_0,gx_0)$. This map is completely positive, unital and trace preserving, and, in particular, $\mathcal{M}(\rho)$ is a state. With this definition, we have

$$\mathcal{F}_\rho[\Xi] = k\mathcal{L}_{\mathcal{M}(\rho)}[\Xi], \tag{35}$$

whence the maximization of $\mathcal{F}_\rho$ is equivalent to the maximization of the likelihood for the transformed state $\mathcal{M}(\rho)$.

Essentially all optimal covariant measurements known in the literature are represented by rank-one operators. The rank-one assumption often provides a useful instrument for simplifying calculations. Nevertheless, as we will show in the following, the occurrence of POVM's with rank grater than one is unavoidable in some relevant situations.

*Proposition 1: For any* $\Xi \in \mathsf{C}$,

$$\mathrm{rank}[\Xi] \geqslant \max_{\mu\in\mathsf{S}}\left(\frac{m_\mu}{d_\mu}\right). \tag{36}$$

*Proof:* Let us decompose $\mathcal{H}$ into irreducible subspaces for the representation $\mathsf{R}(\mathbf{G})$ of $\mathbf{G}$ as follows:

$$\mathcal{H} = \oplus_{\mu\in\mathsf{S}}\oplus_{i=1}^{m_\mu}\mathcal{H}_i^{(\mu)}. \tag{37}$$

Take an orthonormal basis $\mathsf{B}_i^{(\mu)}=\{|(\mu,i),n\rangle | n=1,\ldots,d_\mu\}$ for each subspace $\mathcal{H}_i^{(\mu)}$ in such a way that $|(\mu,i),n\rangle=T_{ij}^{(\mu)}|(\mu,j),n\rangle$ for any $n$, $T_{ij}^{(\mu)}:\mathcal{H}_j\to\mathcal{H}_i$ being the invariant isomorphism which intertwines the equivalent representations $(\mu,i)$ and $(\mu,j)$. Diagonalize $\Xi$ as

$$\Xi = \sum_{k=1}^{\mathrm{rank}(\Xi)} |\eta_k\rangle\langle\eta_k| \tag{38}$$

and write

$$|\eta_k\rangle = \sum_{\mu\in\mathsf{S}}\sum_{i=1}^{m_\mu}\sum_{n=1}^{d_\mu} c_{(\mu,i),n}^k|(\mu,i),n\rangle. \tag{39}$$

Since $\langle\eta_k|T_{ij}^{(\mu)}|\eta_k\rangle=\Sigma_{n=1}^{d_\mu}c_{(\mu,i),n}^{k*}c_{(\mu,j),n}^k$, the normalization constraints (16) become

$$\sum_{k=1}^{\mathrm{rank}(\Xi)}\sum_{n=1}^{d_\mu} c_{(\mu,i),n}^{k*}c_{(\mu,j),n}^k = d_\mu\delta_{ij}. \tag{40}$$

This relation implies that for any $\mu\in\mathsf{S}$ the vectors $\{\mathbf{c}_{(\mu,i)}|i=1,\ldots,m_\mu\}$ defined by $(\mathbf{c}_{(\mu,i)})_{k,n} \doteq c_{(\mu,i),n}^k$ are orthogonal: since they are $m_\mu$ orthogonal vectors in a linear space whose dimension is $d_\mu\times\mathrm{rank}(\Xi)$, it follows that $m_\mu\leqslant d_\mu\times\mathrm{rank}(\Xi)$, hence $\mathrm{rank}(\Xi)\geqslant m_\mu/d_\mu$ $\forall\,\mu\in\mathsf{S}$. ■

Summarizing, every times $m_\mu > d_\mu$ for some class $\mu \in \mathsf{S}$, a covariant POVM cannot be represented by a rank-one seed, due to the normalization constraints.

The previous proposition exhibits a structural reason for which, in the presence of equivalent representations, the set $\mathsf{C}$ of covariant seeds may contain only elements with rank greater than one. On the other hand, in the following we will discuss the occurrence of covariant POVM's with rank greater than one in explicit optimization problems, independently of the presence of equivalent representations.

*Proposition 2: Let* $\Xi$ *be an extremal point of* $\mathsf{C}$. *Denote by P the projector onto* $\mathsf{Supp}(\Xi)$, *and let* $r \doteq \mathrm{rank}(P)$. *Then* $\Xi$ *is the unique seed which maximizes the likelihood for the state* $\rho = P/r$.

*Proof:* First, we need to prove that $\Xi$ commutes with the representation $\mathsf{R}(\mathbf{H}_0) \doteq \{U_k | k \in \mathbf{H}_0\}$, where $\mathbf{H}_0$ is the stability group of $\rho$. Define the group average

$$\xi \doteq \frac{\displaystyle\int_{\mathbf{H}_0} \mathrm{d}h\, U_h \Xi U_h^\dagger}{\displaystyle\int_{\mathbf{H}_0} \mathrm{d}h}. \tag{41}$$

Since $\mathsf{R}(\mathbf{H}_0)$ is the stability group of the projector onto $\mathsf{Supp}(\Xi)$, clearly $\mathsf{Supp}(\Xi)$ is invariant under $\mathsf{R}(\mathbf{H}_0)$, whence $\xi$ satisfies $\mathsf{Supp}(\xi) \subseteq \mathsf{Supp}(\Xi)$. Moreover, using the invariance of the Haar measure it is easy to see that $\xi$ commutes with $\mathsf{R}(\mathbf{H}_0)$. Finally, $\xi$ is an element of $\mathsf{C}$. In fact, it is positive semidefinite, satisfies (16) and commutes with $\mathsf{R}(\mathbf{G}_0)$, which is by definition a subset of $\mathsf{R}(\mathbf{H}_0)$. Since $\Xi$ is extremal, using Theorem 1 we can conclude that $\Xi = \xi$, whence $\Xi$ commutes with $\mathsf{R}(\mathbf{H}_0)$.

Let us prove now optimality. For any arbitrary seed $\zeta \in \mathsf{C}$, the following bound holds:

$$\mathcal{L}_\rho[\zeta] = \mathrm{Tr}[\rho\zeta] = \frac{\mathrm{Tr}[P\zeta]}{r} \leqslant \frac{\mathrm{Tr}[\zeta]}{r} = \frac{\dim(\mathcal{H})}{r}, \tag{42}$$

where the last equality follows from the normalization constraints (16). Clearly $\Xi$ achieves the bound, whence it is optimal. Notice that the inequality $\mathrm{Tr}[P\zeta] \leqslant \mathrm{Tr}[\zeta]$ becomes equality if and only if $\mathsf{Supp}(\zeta) \subseteq \mathsf{Supp}(\Xi)$, then using Theorem 1 we can see that $\Xi$ represents the unique optimal POVM. ■

Consider now a density matrix $\sigma$ with support in the orthogonal complement of $\mathsf{Supp}(\Xi)$, and consider the randomization

$$\rho = (1-\alpha)\frac{P}{r} + \alpha\sigma, \tag{43}$$

with $0 \leqslant \alpha \leqslant 1$. In the following we prove that, for sufficiently small $\alpha > 0$, $\Xi$ is still optimal for the maximum likelihood strategy. In other words, the extremal POVM represented by $\Xi$ is stable under randomization, and the same measuring apparatus can be used for a larger class of mixed states.

*Proposition 3: Consider the randomized state* $\rho$ *in (43) and denote by* $\overline{q}$ *the maximum eigenvalue of* $\sigma$. *If* $\alpha < 1/(1 + r\overline{q})$, *then* $\Xi$ *is the unique seed which maximizes the likelihood for the state* $\rho$.

*Proof:* First, notice that $\Xi$ commutes with the representation $\mathsf{R}(\mathbf{H}_0)$ of the stability group of $\rho$. This follows from the observation that the condition $\alpha < 1/(1 + r\overline{q})$ implies that $(1-\alpha)/r$ is strictly the largest eigenvalue of $\rho$. Then, $P$ is the projector on the eigenspace with maximum eigenvalue of $\rho$, while, for any $h \in \mathbf{G}$, $P_h \doteq U_h P U_h^\dagger$ is the projector on the eigenspace with maximum eigenvalue of $\rho_h \doteq U_h \rho U_h^\dagger$. If $h \in \mathbf{H}_0$ then it must be $\rho_h = \rho$, and, necessarily, $P_h = P$. Therefore $\mathbf{H}_0$ is a subgroup of the stability group of $P$. But $\Xi$ commutes with the representation of the stability group of $P$, as proven in Proposition 2, then it commutes also with $\mathsf{R}(\mathbf{H}_0)$.

Now we prove optimality of $\Xi$. Let us denote by $Q$ the projection onto $\mathsf{Supp}(\sigma)$. The following bound holds for any $\zeta\in\mathsf{C}$:

$$\mathcal{L}_\rho[\zeta]=\frac{(1-\alpha)}{r}\,\mathrm{Tr}[P\zeta]+\alpha\,\mathrm{Tr}[\sigma\zeta] \tag{44}$$

$$\leqslant\frac{(1-\alpha)}{r}\,\mathrm{Tr}[P\zeta]+\alpha\overline{q}\,\mathrm{Tr}[Q\zeta] \tag{45}$$

$$\leqslant\frac{(1-\alpha)}{r}\,\mathrm{Tr}[(P+Q)\zeta] \tag{46}$$

$$\leqslant\frac{(1-\alpha)}{r}\,\mathrm{Tr}[\zeta]=\frac{(1-\alpha)}{r}\,\dim(\mathcal{H}). \tag{47}$$

This bound is achieved by $\Xi$, proving its optimality. Notice that $\Xi$ is the unique optimal seed. In fact, equality in (46) is attained if and only if $\mathrm{Tr}[Q\zeta]=0$, namely when $\mathsf{Supp}(Q)\subseteq\mathsf{Ker}(\zeta)$, while in (47) equality is attained if and only if $\mathsf{Supp}(\zeta)\subseteq\mathsf{Supp}(P)\oplus\mathsf{Supp}(Q)$. Therefore the bound is achieved if and only if $\mathsf{Supp}(\zeta)\subseteq\mathsf{Supp}(P)=\mathsf{Supp}(\Xi)$, implying $\zeta=\Xi$. ■

## VI. EXAMPLES

### A. Extremal POVM's with a nontrivial stability group

*Example 1:* Consider the group of rotations, represented in a $(2j+1)$-dimensional Hilbert space $\mathcal{H}_j$ by the irreducible representation $R_{\mathbf{n},\varphi}\doteq e^{i\varphi\mathbf{n}\cdot\mathbf{j}}$, where $\varphi$ is an angle, $\mathbf{n}$ is a unit vector, and $\mathbf{j}\doteq(j_x,j_y,j_z)$ is the angular momentum operator. In this case a covariant estimation in the orbit of a pure state $|\psi\rangle$ generally may involve a nontrivial stability group. This is actually the case when $|\psi\rangle\doteq|jm\rangle_{\mathbf{n_0}}$, is an eigenvector of $\mathbf{n_0}\cdot\mathbf{j}$ for some unit vector $\mathbf{n_0}$. Clearly in such case the stability group $\mathbf{G}_0$ consists of rotations around $\mathbf{n_0}$, and the state estimation in the orbit reduces to the estimation of a rotated direction $\mathbf{n}'$. The same situation arises for any state $\rho$ mixture of eigenvectors of $\mathbf{n_0}\cdot\mathbf{j}$. Without loss of generality, let us take $\mathbf{n_0}$ as the direction of the $z$ axis, and write $\rho=\Sigma_{m=-j}^{j}p_m|jm\rangle\langle jm|$ with $p_m\geqslant0\ \forall m$. Let us denote by $P$ the projector onto $\mathsf{Supp}(\rho)$, and take $\overline{m}$ such that $p_{\overline{m}}=\max_m\{p_m\}$. Then, since

$$\mathrm{Tr}[\rho\zeta]\leqslant p_{\overline{m}}\,\mathrm{Tr}[P\Xi]\leqslant p_{\overline{m}}\,\mathrm{Tr}[\Xi]=p_{\overline{m}}(2j+1),$$

one has that $\Xi=(2j+1)|j\overline{m}\rangle\langle j\overline{m}|$ is the optimal POVM. Notice that such POVM commutes with the stability group $\mathsf{R}(\mathbf{G}_0)$ and is extremal, as a consequence of Corollary 1.

*Example 2:* Consider the group $\mathbb{SU}(d)$ of unitary $d\times d$ matrices with unit determinant, acting on the space $\mathcal{H}\doteq\mathbb{C}^d$. It is easy to see that each vector $|\psi\rangle\in\mathcal{H}$ has a nontrivial stability group $\mathbf{G}_0\equiv\mathbb{U}(d-1)$. In fact, by introducing an orthonormal basis $\mathsf{B}_\perp\doteq\{|n\rangle|n=1,\ldots,d-1\}$ for the orthogonal complement $\mathcal{H}^\perp$ of the line $\mathsf{Span}\{|\psi\rangle\}$, and the basis $\mathsf{B}\doteq|\psi\rangle\cup\mathsf{B}_\perp$ for $\mathcal{H}$, the stability group $\mathbf{G}_0$ consists on matrices of the form

$$U_h=\begin{pmatrix}\omega_h & \mathbf{0}\\ \mathbf{0} & V_h\end{pmatrix}, \tag{48}$$

where $\omega_h\in\mathbb{C}$, $|\omega_h|=1$, and $V_h$ is a unitary $(d-1)\times(d-1)$ matrix with $\mathrm{Det}(V_h)=\omega_h^*$. Let us consider now the tensor representation $\mathsf{R}(\mathbf{G})=\{U_g^{\otimes2}|U_g\in\mathbb{SU}(d)\}$ on the space $\mathcal{H}^{\otimes2}$. This representation has two irreducible subspaces, the symmetric and the antisymmetric ones $\mathcal{H}_+$ and $\mathcal{H}_-$, with dimensions $d_+=d(d+1)/2$ and $d_-=d(d-1)/2$, respectively. Denote by $P_+$ and $P_-$ the projectors on $\mathcal{H}_+$ and $\mathcal{H}_-$. Let us apply the representation $\mathsf{R}(\mathbf{G})$ on the state $|\psi\rangle^{\otimes2}\in\mathcal{H}^{\otimes2}$. Clearly the stability group is the same $\mathbf{G}_0$ as before, and it is represented by $\mathsf{R}(\mathbf{G}_0)=\{U_h^{\otimes2}|h\in\mathbf{G}_0\}$. It is easy to see that

$\mathsf{R}(\mathbf{G}_0)$ contains five irreducible components, carried by the subspaces $\mathcal{H}_1=\mathsf{Span}\{|\psi\rangle^{\otimes 2}\}$, $\mathcal{H}_2=\mathsf{Span}\{|\psi\rangle\}\otimes\mathcal{H}^{\perp}$, $\mathcal{H}_3=\mathcal{H}^{\perp}\otimes\mathsf{Span}\{|\psi\rangle\}$, $\mathcal{H}_4=P_+(\mathcal{H}^{\perp\otimes 2})$, and $\mathcal{H}_5=P_-(\mathcal{H}^{\perp\otimes 2})$. Notice that $\mathcal{H}_2$ and $\mathcal{H}_3$ carry equivalent representations, corresponding to a two-dimensional multiplicity space. An example of extremal POVM is given by

$$\Xi=\frac{d(d+1)}{2}|\psi\rangle\langle\psi|^{\otimes 2}\oplus\frac{d}{d-2}P_-QP_-,$$

where $Q$ is the projection on $\mathcal{H}^{\perp\otimes 2}$. Since the two summands are proportional to $|\psi\rangle\langle\psi|^{\otimes 2}$ and $P_-QP_-$, which are the projectors on $\mathcal{H}_1$ and $\mathcal{H}_5$, respectively, then $\Xi$ belongs to the commutant of $\mathsf{R}(\mathbf{G}_0)=\{U_h^{\otimes 2}|h\in\mathbf{G}_0\}$. Notice that the subspaces $\mathcal{H}_1$ and $\mathcal{H}_5$ have multiplicities $m_1=m_5=1$, corresponding to one-dimensional multiplicity spaces $\mathcal{M}_1\equiv\mathcal{M}_5\equiv\mathbb{C}$ (whence the partial traces over $\mathcal{H}_{1,5}$ will be $c$ numbers). Moreover, using the fact that $\mathrm{Tr}_{\mathcal{H}_1}[P_+]=1$, $\mathrm{Tr}_{\mathcal{H}_1}[P_-]=0$, $\mathrm{Tr}_{\mathcal{H}_5}[P_+]=0$, $\mathrm{Tr}_{\mathcal{H}_5}[P_-]=(d-1)(d-2)/2$ one can check extremality using the condition (26). Let us observe that in this example we have $r_1=r_5=1$ and $m_+=m_-=1$, where $r_1$ and $r_5$ are defined as in Corollary 2, while $m_+$ and $m_-$ are the multiplicities of the two irreducible representations of $\mathsf{R}(\mathbf{G})$. Then the bound of (30) is saturated. Finally, we remark that this POVM is optimal for discriminating states in the orbit of $|\psi\rangle^{\otimes 2}$,[13] in the orbit of $\rho=(1/r)(|\psi\rangle\langle\psi|^{\otimes 2}+P_-QP_-)$ where $r=1+[(d-1)(d-2)/2]$ because of Proposition 2, and also in the orbit of any randomization $\rho'=(1-\alpha)\rho+\alpha\sigma$ where $\sigma$ is density matrix with $\mathsf{Supp}(\sigma)\subseteq\mathsf{Ker}(P)$, and $\alpha<1/(1+r)$, because of Proposition 3.

### B. Extremal POVM's with rank greater than one

*Example 1:* Consider the Abelian group $\mathbf{G}=\mathbb{U}(1)$ of phase shifts, acting in the space $\mathcal{H}=\mathbb{C}^d$ by the representation $\mathsf{R}(\mathbf{G})=\{U(\varphi)=\exp(i\varphi N)\}|\ \varphi\in[-\pi,\pi]\}$, where the generator $N$ is given by $N=\sum_{n=0}^{d-1}n|n\rangle\langle n|$ for some orthonormal basis $\{|n\rangle|n=0,1,\ldots,d-1\}$. The stability group $\mathbf{G}_0$ may be either the whole $\mathbb{U}(1)$ (for $\rho$ diagonal on the eigenstates of the generator), or a discrete subgroup $\mathbf{G}_0=\mathbb{Z}_k$ for some integer $k$, including the case $k=1$ of trivial stability group. We exclude the degenerate case $\mathbf{G}_0=\mathbb{U}(1)$ of shift invariant states. The parameter space $\mathfrak{X}=\mathbb{U}(1)/\mathbb{Z}_k$ will be a circle, parametrized by an angle $\theta\in[-\pi,\pi]$, and the action of a group element $g(\varphi)\in\mathbf{G}$ on an element $\theta\in\mathfrak{X}$ will be given by $g(\varphi)\theta=\theta+k\varphi$.

Due to constraint (16), a seed $\Xi$ is represented in the eigenbasis of the generator by a correlation matrix, namely by a positive semidefinite matrix with unit diagonal entries. Vice versa, any correlation matrix corresponds to a seed in the case of trivial stability group $\mathbf{G}_0$. In Ref. 10 one can find a constructive method which provides extremal correlation matrices with rank $r>1$: here we show that any of such matrices can be viewed as the optimal seed for the estimation problem in the orbit of a particular state. Let us choose as optimality criterion the maximization of the average value of a positive summable function $f:\mathfrak{X}\times\mathfrak{X}\to\mathbb{R}_+$ depending only on the difference $\theta-\theta_*$ between the measured and the true value. Suppose $\rho$ a state with stability group $\mathbf{G}_0=\mathbb{Z}_k$. As we noted at the beginning of Sec. V, the maximization of $\mathsf{F}_\rho[\Xi]$—the average value of $f(\theta-\theta_*)$—corresponds to the maximization of the likelihood $\mathcal{L}_{\mathcal{M}(\rho)}[\Xi]$ for the transformed state $\mathcal{M}(\rho)=f_0^{-1}\int_{-\pi}^{\pi}(\mathrm{d}\varphi/2\pi)f(-k\varphi)U_\varphi\rho U_\varphi^\dagger$ [from Eq. (34)]. Notice that the map $\mathcal{M}$ is trivially covariant—i.e., $\mathcal{M}(U_\phi\rho U_\phi^\dagger)=U_\phi\mathcal{M}(\rho)U_\phi^\dagger$—since the group is Abelian. For simplicity here we require that the map $\mathcal{M}$ is invertible, whence also $\mathcal{M}^{-1}$ is covariant and trace preserving (but generally not positive). Covariance of $\mathcal{M}$ implies that the stability group of $\mathcal{M}(\rho)$ contains the stability group of $\rho$, and covariance of $\mathcal{M}^{-1}$ implies the reverse inclusion, whence the stability group is not changed by the maps.

Let us take now an extremal correlation matrix $\Xi$ with $\mathrm{rank}(\Xi)=r\geqslant 1$ and denote by $P$ the projector onto $\mathsf{Rng}(\Xi)$. Using Proposition 2, we can see that $\Xi$ commutes with the representation $\mathsf{R}(\mathbf{H}_0)$, where $\mathbf{H}_0$ is the stability group of $P$. Call λ the modulus of the minimum eigenvalue of $\mathcal{M}^{-1}(P/r)$, then

$$\rho = \frac{\lambda}{1 + \mathrm{d}\lambda} I + \frac{1}{1 + \mathrm{d}\lambda} \mathcal{M}^{-1}(P/r)$$

is a density operator. Notice that the stability group $\mathbf{G}_0$ of $\rho$ is the same stability group of $\mathcal{M}^{-1}(P)$, which coincides with $\mathbf{H}_0$, the stability group of $P$. Therefore $\Xi$ commutes with the representation $\mathsf{R}(\mathbf{G}_0)$. It is easy to show that $\Xi$ is the unique seed commuting with $\mathsf{R}(\mathbf{G}_0)$ which is also optimal for the estimation of states in the orbit of $\rho$. In fact, for any $\zeta$ in the convex set $\mathsf{C}$ of the seeds with stability group $\mathbf{G}_0$, we have

$$\mathsf{F}_\rho[\zeta] = f_0 \, \mathrm{Tr}[\zeta \mathcal{M}(\rho)] = f_0 \left( \frac{\lambda}{1 + \mathrm{d}\lambda} \mathrm{Tr}[\zeta] + \frac{1}{r(1 + \mathrm{d}\lambda)} \, \mathrm{Tr}[\zeta P] \right) \leqslant f_0 \left( \frac{d}{r} \right) \left( \frac{1 + r\lambda}{1 + \mathrm{d}\lambda} \right).$$

This bound is achieved choosing $\zeta = \Xi$, moreover, as in Proposition 2, we can observe that the functional $\mathrm{Tr}[\zeta P]$ with $\zeta \in \mathsf{C}$ is maximum if and only if $\zeta = \Xi$, then the maximum is unique.

*Example 2:* We provide now an example with a noncompact group represented in an infinite dimensional Hilbert space. This example is out of the general treatment of the present paper—which considers only finite dimensions—and is given only with the purpose of showing that our results could be generalized to infinite dimensions, however at the price of much more technical proofs.

Take $\mathcal{H}$ as the Fock space, and consider the projective representationon $\mathcal{H}$ of the group of translations on the complex plane $\mathbb{C}$ in terms of the Weyl–Heisenberg operators $\mathsf{R}(\mathbf{G}) = \{D(\alpha) = e^{\alpha a^\dagger - \bar{\alpha} a} | \alpha \in \mathbb{C}\}$, where $[a, a^\dagger] = 1$. Here we will consider the twofold tensor representation $\{D(\alpha)^{\otimes 2} \lfloor \alpha \in \mathbb{C}\}$ on $\mathcal{H}^{\otimes 2}$. Using the unitary operator $V = e^{(\pi/4)(a_1 a_2^\dagger - a_1^\dagger a_2)}$, one can write $D(\alpha)^{\otimes 2} = V(D(\sqrt{2}\alpha) \otimes I)V^\dagger$ and see that the irreducible subspaces of this representation are $\mathcal{H}_n = V(\mathcal{H} \otimes \mathsf{Span}\{|\phi_n\rangle\})$, where $\{|\phi_n\rangle | n = 1, 2, \ldots, \infty\}$ is any orthonormal basis for $\mathcal{H}$. All these subspaces carry equivalent representations, the isomorphism between $\mathcal{H}_m$ and $\mathcal{H}_n$ being

$$T_{mn} = V(I \otimes |\phi_m\rangle\langle\phi_n|)V^\dagger. \tag{49}$$

In terms of these isomorphisms, the normalization constraints (16) for a seed operator become[13]

$$\mathrm{Tr}[T_{mn}\zeta] = 2\,\delta_{mn}. \tag{50}$$

Notice that the number 2 in this formula has nothing to do with the dimension of $\mathcal{H}_n$ which is infinite: in the noncompact case the dimensions are replaced by positive numbers depending only on the equivalence class of representations. In principle, since the space $\mathcal{H}^{\otimes 2}$ is infinite dimensional, there is the possibility of extremal covariant POVM's with an infinite rank. Actually we can provide the remarkable example

$$\Xi = 2V(|0\rangle\langle 0| \otimes I)V^\dagger, \tag{51}$$

where $|0\rangle$ is the vacuum state of the Fock basis $\{|m\rangle | a^\dagger a|m\rangle = m|m\rangle\}$. The corresponding POVM can be realized by averaging the outcomes of two independent measurements with $\Xi_1 = |0\rangle\langle 0| \otimes I$ and $\Xi_2 = I \otimes |0\rangle\langle 0|$,[13] which in quantum optics correspond to two heterodyne measurements.[14]

We can observe that $\Xi$ maximizes the likelihood functional for any state of the form $\rho = V(|0\rangle\langle 0| \otimes \sigma)V^\dagger$, where $\sigma = \Sigma_{n=0}^{\infty} p_n |\phi_n\rangle\langle\phi_n|$, is a mixed state with $p_n > 0\,\forall n$. In fact, for any seed $\zeta$, one has the bound

$$\begin{aligned} \mathrm{Tr}[V(|0\rangle\langle 0| \otimes \sigma)V^\dagger \zeta] &= \sum_{n=0}^{\infty} p_n \, \mathrm{Tr}[V(|0\rangle\langle 0| \otimes |\phi_n\rangle\langle\phi_n|)V^\dagger \zeta] \\ &\leqslant \sum_{n=0}^{\infty} p_n \, \mathrm{Tr}[V(I \otimes |\phi_n\rangle\langle\phi_n|)V^\dagger \zeta] = \sum_{n}^{\infty} p_n \, \mathrm{Tr}[T_{nn}\zeta] = 2, \end{aligned} \tag{52}$$

and since $\Xi$ achieves the bound (52), it is optimal. Moreover $\Xi$ is the unique optimal seed. In fact, the equality in (52) is achieved if and only if $\mathrm{Tr}[V(|0\rangle\langle 0|\otimes|\phi_n\rangle\langle\phi_n|)V^\dagger\zeta]=\mathrm{Tr}[V(I\otimes|\phi_n\rangle \times\langle\phi_n|)V^\dagger\zeta]$ for any $n$: by expanding the identity on the Fock basis, the positivity of $\zeta$ implies $\langle m|\langle\phi_n|V^\dagger\zeta V|m\rangle|\phi_n\rangle=0$ for any $m\neq 0$. Hence the unique nonzero diagonal elements of $\zeta$ are on the vectors $V|0\rangle|\phi_n\rangle$. On the other hand, the positivity of $\zeta$ along with the normalization constraint $\mathrm{Tr}[T_{mn}\zeta]=0\ \forall m\neq n$ imply that all the off diagonal elements of $\zeta$ are zero. Hence $\zeta=2V\Sigma_{n=1}^{\infty}(|0\rangle \times\langle 0|\otimes|\phi_n\rangle\langle\phi_n|)V^\dagger=2V(|0\rangle\langle 0|\otimes I)V^\dagger=\Xi$. The fact that $\Xi$ is the unique optimal seed ensures that it is also extremal, otherwise there would be two different seeds which are equally optimal. Notice that $\Xi$ is extremal also according to our characterization (26).

[1] I. L. Chuang and M. A. Nielsen, *Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, 2000).

[2] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[4] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam 1982).

[5] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).

[6] A. Acin, E. Jane, and G. Vidal, Phys. Rev. A **64**, 050302 (2001).

[7] G. M. D'Ariano, C. Macchiavello, and M. F. Sacchi, Phys. Lett. A **248**, 103 (1998).

[8] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. Lett. (to be published), quant-ph/0405095.

[9] G. M. D'Ariano, J. Math. Phys. (to be published).

[10] C.-K. Li and B.-S. Tam, SIAM J. Matrix Anal. Appl. **15**, 903 (1994).

[11] Notice that in infinite dimensions the POVM density, whence $\Xi$, can become unbounded, or not even anymore an operator, and it can be treated rigorously in the framework of forms.

[12] D. P. Zhelobenko, *Compact Lie Groups and Their Representations* (American Mathematical Society, Providence, RI, 1973).

[13] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. A (to be published), quant-ph/0403083.

[14] G. M. D'Ariano, "Quantum estimation theory and optical detection," in *Quantum Optics and the Spectroscopy of Solids*, edited by T. Hakioğlu and A. S. Shumovsky (Kluwer, Dordrecht, 1997), pp. 139–174.