Covariant quantum measurements that maximize the likelihood

Giulio Chiribella,* Giacomo Mauro D'Ariano,† Paolo Perinotti,‡ and Massimiliano F. Sacchi§
QUIT (Quantum Information Theory) Group,‖ Dipartimento di Fisica "A. Volta" and INFM, via A. Bassi 6, I-27100 Pavia, Italy
(Received 10 March 2004; published 9 December 2004)

We derive the class of covariant measurements that are optimal according to the maximum likelihood criterion. The optimization problem is fully resolved in the case of pure input states, under the physically meaningful hypotheses of unimodularity of the covariance group and measurability of the stability subgroup. The general result is applied to the case of covariant state estimation for finite dimension, and to the Weyl-Heisenberg displacement estimation in infinite dimension. We also consider estimation with multiple copies, and analyze the behavior of the likelihood versus the number of copies. A "continuous-variable" analog of the measurement of direction of the angular momentum with two antiparallel spins by Gisin and Popescu is given.

DOI: 10.1103/PhysRevA.70.062105 PACS number(s): 03.65.Ta, 03.67.−a, 02.20.−a

I. INTRODUCTION

State estimation is a unique kind of quantum measurement in the quality of information that it provides. In fact, the knowledge of the state of a quantum system enables the evaluation of any ensemble average, which is equivalent to the possibility of performing any desired experiment on the system. For its intrinsic versatility such unconventional type of quantum measurement is of interest for the new technology of quantum information [1] in the estimation of parameters that do not correspond to observables [2]—such as the phase of an electromagnetic field—but also as a method to achieve quantum cloning [3,4], whence in designing eavesdropping strategies for quantum cryptography [5].

An exact state estimation without any prior knowledge of the form of the state is impossible [6] due to the no-cloning theorem [7,8]. This also reflects the fact that an optimal approximate state estimation would not be achievable as an orthogonal measurement, since the state estimation is a kind of "informationally complete" measurement [9]. More generally, one can have some prior knowledge of the form of the state, i.e., by parametrizing it with a restricted set of variables. This is the typical situation of the quantum estimation theory of Helstrom [2], where the goal is to determine a multidimensional parameter of a state transformation. When the set of states to be discriminated are orthonormal the parameter corresponds to an "observable" whose eigenstates are the set itself, and the estimation is exact. However, in practice it happens very often that the multidimensional parameter cannot be described by an observable (e.g., it is a phase of a field, or it corresponds to a set of noncompatible observables), whence a measurement represented by a so-called *positive operator valued measure* (POVM) needs to be performed.

For a state estimation that is not equivalent to the measurement of an observable we have a choice of infinitely many POVM's achieving the same task with different strategies. Indeed, there is no universal criterion which is optimal for all situations, and one needs to define the appropriate figure of merit pertaining to the particular problem. Once the optimization problem is solved in terms of an optimal POVM, one can then address the problem of the feasibility of the measurement apparatus by classification of orthogonal dilations of the POVM [2,10,11], or else compare the performance of actual devices to the ultimate theoretical limit.

A statistically meaningful optimization strategy is the maximization of the likelihood that the true value of the estimated parameter coincides with the outcome of the measurement. Such a strategy is actually very general, since for measurements that are *group covariant*, optimization of a generic *goal function* corresponds to optimization of the likelihood for a different input state. Physically, "group covariance" means that there is a group of transformations on the probability space which maps events into events, in such a way that when the quantum system is transformed according to one element of the group, the probability of the given event becomes the probability of the transformed event. This situation is very natural, and occurs in most practical applications. For example, the heterodyne measurement [12,13] is covariant under the group of displacements of the complex field, which means that if we displace the state of radiation by an additional complex averaged field, then the output photocurrent will be displaced by the same complex quantity. Other examples of covariant measurements are the quantum estimation of a "spin orientation" [14–17], or of the phase shift of an electromagnetic field [2,18,19].

The statistics of the measurement can be improved by using many copies of the same quantum system. In this scenario, it is relevant for experiments to distinguish the measurements achievable by local operations and classical communication (LOCC) from more general schemes that require entanglement. Unfortunately, a useful classification of LOCC schemes is still missing. Alternatively, one can give just a mathematical categorization in terms of the POVM of the measurement: (i) "independent" measurements, correspond-

*Electronic address: chiribella@fisicavolta.unipv.it
†Also at Center for Photonic Communication and Computing, Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208, USA. Electronic address: dariano@unipv.it
‡Electronic address: perinotti@fisicavolta.unipv.it
§Electronic address: msacchi@unipv.it
‖URL: http://www.qubit.it

ing to tensor products of independent POVM's; (ii) "separable" measurements, corresponding to POVM's where each element is separable; (iii) "nonseparable" or "entangled" measurements, corresponding to POVM's where some element is entangled. In the first category measurements are performed independently on each copy. In the separable class, on the other hand, the measurement can be performed by means of separable operations; hence all LOCC schemes are included in this category. Notice, however, that not all separable operations can be implemented locally (see, e.g., the case of nonlocality without entanglement of Ref. [20]). Finally, the class of entangled POVM's represents the most general scheme of measurement, and opens the exponential growth of the Hilbert space dimension versus the number of copies N, with the possibility of largely surpassing the statistical efficiency of the independent measurement schemes [21–24]. However, as already noticed in Ref. [25], in the case of pure states and for the maximum likelihood strategy, the optimal schemes can be surprisingly achieved by separable measurements, and here we address this issue for covariant measurements. Under the general assumption of square-summable representation we derive a general "canonical form" for the optimal measurements for pure input states, corresponding to a POVM which is separable or entangled, depending on the group representation.

After introducing in Sec. II the precise formulation of the covariant state estimation problem, in Sec. III we derive some useful mathematical identities for group integrals which are then used to algebraically characterize covariant measurements. This also helps us in deriving a simple upper bound for the maximum likelihood in Sec. IV, along with the canonical form of the optimal measurement given in terms of the group representation. Examples of the canonical form are given in Sec. V in dimension $d<\infty$ for the group SU(d)—corresponding to the estimation of an unknown pure state—and in infinite dimensions for the estimation of displacements on the phase space. The case of multiple copies is then analyzed, discussing the occurrence of entangled versus separable POVM's. For the estimation of displacements on the phase space, the case of two copies experiencing opposite shifts in momentum is also analyzed—the continuous-variable analog of the measurement of direction of the angular momentum with two antiparallel spins by Gisin and Popescu [26]. For coherent states it is shown that such a scheme provides a better estimation of the displacement as compared to the conventional case of identical displacements.

II. THE PROBLEM

Whenever a quantum system $\mathcal{S}$ undergoes a physical transformation belonging to a group $\mathbf{G}$, its state is transformed according to an appropriate representation of $\mathbf{G}$ on the Hilbert space H of the system $\mathcal{S}$. In the following, we will consider the case in which the group $\mathbf{G}$ is a Lie group which acts on H by a (projective) unitary representation $\{U_g\}$, whereas the initial state—also called the *seed* state—is a pure state $|\Psi\rangle$. Notice that the correspondence between transformed states and group elements is generally not injective, since the state $|\Psi\rangle$ may have a nontrivial stability group, say $\mathbf{G}_\Psi$ (we say that a group element h belongs to the stability group $\mathbf{G}_\Psi$ of $|\Psi\rangle$ when $U_h|\Psi\rangle=e^{i\phi_h}|\Psi\rangle$, with ϕ_h a real phase). In this way the transformed states are in one-to-one correspondence with the cosets $g\mathbf{G}_\Psi$: in other words the group-orbit manifold (obviously invariant under the group representation $\{U_g\}$) is identified with the coset space $\mathfrak{X}=\mathbf{G}/\mathbf{G}_\Psi$. We see that in principle from the output state $U_g|\Psi\rangle$ it is possible to estimate the group element g of the transformation U_g only if the stability group $\mathbf{G}_\Psi$ of the input state $|\Psi\rangle$ is trivial. Otherwise, we can estimate the coset $x\in\mathfrak{X}$ which is in one-to-one correspondence with the output state $|\Psi_x\rangle=U_{g(x)}|\Psi\rangle$, $g(x)$ labeling any element of $\mathbf{G}$ in the coset x. In the following we will denote by $x_0\equiv e\mathbf{G}_\Phi$ the coset containing the identity element e, and the seed state is relabeled accordingly as $|\Psi_{x_0}\rangle\equiv|\Psi\rangle$. This notation makes explicit the isomorphism between the coset space $\mathfrak{X}$ and the *homogeneous* manifold of states $|\Psi_x\rangle$ $x\in\mathfrak{X}$, i.e., on which the group acts transitively through its unitary representation as $U_g|\Psi_x\rangle\propto|\Psi_{gx}\rangle$ (apart from a phase factor). In this way, the estimation of the parameter $x\in\mathfrak{X}$ becomes equivalent to a problem of *covariant state estimation*, and it was proved [19] that the optimal probability distribution $p(x|x_0)$ of estimating x for input state $|\Psi_{x_0}\rangle$ satisfies the identity $p(gx|gx_0)=p(x|x_0)$, namely, the probability distribution on the manifold $\mathfrak{X}$ for an input state $U_g|\Psi\rangle$ is equal to the probability distribution for input state $|\Psi\rangle$ but with the manifold shifted by g^{-1}. In the following we will suppose for simplicity that the group $\mathbf{G}$ is unimodular (i.e., the left invariant measure dg on $\mathbf{G}$ is also right invariant) and the stability subgroup is compact. According to a theorem by Holevo [19], for square-integrable representations the covariant estimation is described by a POVM M on the probability space $\mathfrak{X}$ with density of the general form

$$dM(x)=U_{g(x)}\Xi U^\dagger_{g(x)}dx, \tag{1}$$

where dx denotes the invariant measure on $\mathfrak{X}$ induced by invariant measure dg on $\mathbf{G}$ [27], and the positive *kernel* operator Ξ belongs to the commutant $\mathbf{G}'_\Psi$ of the stability group (i.e., $[\Xi,U_h]=0\ \forall\ h\in\mathbf{G}_\Psi$), and satisfies the completeness constraint

$$\int_{\mathfrak{X}}dx\ U_{g(x)}\Xi U^\dagger_{g(x)}\equiv\int_{\mathbf{G}}dg\ U_g\Xi U_g^\dagger=I. \tag{2}$$

The fact that $\Xi\in\mathbf{G}'_\Psi$ guarantees that the POVM does not depend on the particular choice of $g(x)$.

III. GROUP INTEGRALS OF OPERATORS

The completeness constraint in Eq. (2) becomes particularly simple with some abstract considerations on group integrals. Since the group $\mathbf{G}$ is unimodular, its unitary square-summable representations satisfy Schur's lemma for any (generally infinite dimensional) representation space H [28], namely, for any couple $\{U_g^\mu\}$ and $\{U_g^\nu\}$ of irreducible components of the representation with invariant subspaces $\mathsf{H}_\mu,\mathsf{H}_\nu\subseteq\mathsf{H}$, respectively, every operator $O_{\mu\nu}$: $\mathsf{H}_\nu\to\mathsf{H}_\mu$ satisfying the identity $U_g^{\mu} O_{\mu\nu} = O_{\mu\nu} U_g^{\nu}\ \forall\ g \in \mathbf{G}$ must be of the form

$$O_{\mu\nu} = \begin{cases} k I_{\mu\nu} & \text{for } \mu \sim \nu, \\ 0 & \text{otherwise}, \end{cases}$$

where $\sim$ denotes equivalence of irreducible representations, k is a constant, and $I_{\mu\nu}$: $\mathsf{H}_\nu \to \mathsf{H}_\mu$ is the isomorphism mapping the two equivalent components, namely, $U_g^{\mu} = I_{\mu\nu} U_g^{\nu} I_{\mu\nu}^{\dagger}$ $\forall\ g \in \mathbf{G}$ ($I_{\mu\mu}$ is the orthogonal projector onto the invariant irreducible subspace H_μ).

A simple consequence of Schur's lemma is the Wedderburn decomposition of operators O such that $\mathrm{Tr}[I_{\mu\nu} O] < \infty \forall\ \mu, \nu$ [29]

$$\int_G dg\, U_g O U_g^{\dagger} = \sum_\mu \sum_{\nu \sim \mu} a_{\mu\nu} I_{\mu\nu}. \tag{3}$$

Taking the expectation values of both sides of Eq. (3) on an arbitrary element $|e_n^{(\mu)}\rangle$ of an orthonormal basis $\{|e_m^{(\mu)}\rangle\}$ for H_μ one has

$$a_{\mu\mu} = \int_G dg\, \mathrm{Tr}[(U_g^{\dagger} |e_n^{(\mu)}\rangle\langle e_n^{(\mu)}| U_g) O]\ \ \forall\ \ n. \tag{4}$$

Applying now the Wedderburn decomposition to the group average of projectors $|e_n^{(\mu)}\rangle\langle e_n^{(\mu)}|$ and using invariance of the subspace H_μ, one obtains

$$\int_G dg\, U_g^{\dagger} |e_n^{(\mu)}\rangle\langle e_n^{(\mu)}| U_g = b_{\mu\mu} I_{\mu\mu}, \tag{5}$$

where $b_{\mu\mu}$ is a constant to be evaluated. We then have

$$a_{\mu\mu} = b_{\mu\mu} \mathrm{Tr}[I_{\mu\mu} O], \tag{6}$$

where $b_{\mu\mu}$ can be determined by taking the expectation value of both sides of Eq. (5) on any normalized vector in H_μ, in particular on the vector $|e_n^{(\mu)}\rangle$, leading to

$$b_{\mu\mu} = \int_{\mathbf{G}} dg\, |\langle e_n^{(\mu)}| U_g |e_n^{(\mu)}\rangle|^2. \tag{7}$$

On the other hand, if the representations μ and ν are equivalent, there are two orthonormal basis $\{|e_n^{(\mu)}\rangle\}$ and $\{|e_m^{(\nu)}\rangle\}$ for H_μ and H_ν, respectively, such that $I_{\mu\nu} = \Sigma_n |e_n^{(\mu)}\rangle\langle e_n^{(\nu)}|$. Now, taking the matrix element of both sides of Eq. (3) between vectors $|e_n^{(\mu)}\rangle$ and $|e_n^{(\nu)}\rangle$ one has

$$a_{\mu\nu} = \int_G dg\, \mathrm{Tr}[U_g^{\dagger} |e_n^{(\nu)}\rangle\langle e_n^{(\mu)}| U_g O]. \tag{8}$$

The invariance of both subspaces H_μ and H_ν along with Schur's lemma gives the identity

$$\int_G dg\, U_g^{\dagger} |e_n^{(\nu)}\rangle\langle e_n^{(\mu)}| U_g = b_{\nu\mu} I_{\nu\mu}, \tag{9}$$

for a suitable constant $b_{\nu\mu}$ to be determined. Substituting the last equation into Eq. (8) gives

$$a_{\mu\nu} = b_{\nu\mu} \mathrm{Tr}[I_{\nu\mu} O], \tag{10}$$

and the constant $b_{\nu\mu}$ can be determined by taking the matrix element of Eq. (9) between vectors $|e_n^{(\nu)}\rangle$ and $|e_n^{(\mu)}\rangle$, namely,

$$b_{\nu\mu} = \int_{\mathbf{G}} dg\, \langle e_n^{(\nu)}| U_g^{\dagger} |e_n^{(\nu)}\rangle\langle e_n^{(\mu)}| U_g |e_n^{(\mu)}\rangle. \tag{11}$$

Notice that for equivalent components $\mu \sim \nu$ for our choice of bases one has $\langle e_n^{(\mu)}| U_g |e_n^{(\mu)}\rangle = \langle e_n^{(\nu)}| U_g |e_n^{(\nu)}\rangle$, whence $b_{\nu\mu} = b_{\mu\mu} = b_{\nu\nu} \equiv b_\mu$. Summarizing, we have the decomposition

$$\int_G dg\, U_g O U_g^{\dagger} = \sum_\mu b_\mu \sum_{\nu \sim \mu} \mathrm{Tr}[I_{\nu\mu} O] I_{\mu\nu},$$

$$b_\mu = \int_{\mathbf{G}} dg\, |\langle e_n^{(\mu)}| U_g |e_n^{(\mu)}\rangle|^2. \tag{12}$$

If the group $\mathbf{G}$ is compact and its measure dg is normalized (i.e., $\int_{\mathbf{G}} dg = 1$), then it is easy to show that $b_\mu = 1/d_\mu$, where $d_\mu = \dim(\mathsf{H}_\mu)$ (irreducible representations of compact groups are finite dimensional). In fact, summation over all n in Eqs. (4) and (8) provides in a direct way the values $a_{\mu\mu} = \mathrm{Tr}[I_{\mu\mu} O]/d_\mu$ and $a_{\mu\nu} = \mathrm{Tr}[I_{\mu\nu} O]/d_\mu$ for the coefficients in Eq. (3). On the other hand, the derivation given above holds for unitary square-summable representations, even with a Dirac-orthogonal basis $\{|e_x^{(\mu)}\rangle\}$ for H_μ, namely, $\langle e_x^{\mu} | e_{x'}^{\mu}\rangle = \delta(x - x')$. The coefficients b_μ^{-1} are generally noninteger, are called *formal dimensions*, and carry information about the structure of the irreducible components of the group representation.

IV. MEASUREMENTS WITH MAXIMUM LIKELIHOOD

We will now consider measurements which maximize the likelihood, namely, the conditional probability density $p(x|x)$ of having the outcome equal to the true value for any x. Because of covariance this optimality criterion is equivalent to maximize the *likelihood functional* $\mathcal{L}_\rho[\Xi] = \mathrm{Tr}[\Xi\rho]$ with $\rho = |\Psi\rangle\langle\Psi|$, $|\Psi\rangle$ being the input state.

Notice that the general solution to the maximum likelihood problem, which at first sight may appear of limited value, is actually equivalent to the solution of any quantum estimation problem with positive summable "goal" function $f(\hat{x}, x)$ [the goal function is the opposite of the customary cost function—$f(\hat{x}, x)$ [2]]. This consists in associating with each measurement outcome $\hat{x}$ a "score" $f(\hat{x}, x)$, with the function $f(\hat{x}, x)$ increasing versus $\hat{x}$ for $\hat{x}$ approaching the true value x. Then, the optimal measurement is the one that maximizes the average score. In a covariant estimation problem a meaningful goal function must satisfy the invariance property $f(\hat{x}, x) = f(g\hat{x}, gx) \forall\ g \in \mathbf{G}$, and this allows us to define a function $h(\hat{g}, g)$ on the group via the relation $h(\hat{g}, g) \equiv f(\hat{g}x_0, gx_0)$ for fixed x_0. Then, the function h is positive (bounded from below), summable, and satisfies $h(\hat{g}, g) = h(g^{-1}\hat{g}, e)$, e denoting the identity transformation. Now, thanks to covariance the average score can be written as

$$\overline{s} = \int_{\mathbf{G}} dg\, h(g,e)\mathrm{Tr}[\rho U_g \Xi U_g^\dagger] = \left(\int_{\mathbf{G}} dg\, h(g,e)\right)\mathcal{L}_{\mathcal{M}(\rho)}[\Xi],$$

where

$$\mathcal{M}(\rho) = \frac{\int_{\mathbf{G}} dg\, h(g,e) U_g^\dagger \rho U_g}{\int_{\mathbf{G}} dg\, h(g,e)}$$

is a completely positive trace preserving map. Therefore, the maximization of a goal function can be viewed as a maximum likelihood scheme on the transformed state $\mathcal{M}(\rho)$, and depending on the form of the function h the choice of the input state may be restricted to special states, possibly mixed. Nevertheless, in this paper we will give a complete solution only for pure input states.

The problem is now to find a positive operator Ξ which maximizes the likelihood functional $\mathcal{L}_\rho[\Xi]=\mathrm{Tr}[\Xi\rho]$, and, at the same time, satisfies the completeness constraints (2). Once an optimal Ξ is found, the presence of a nontrivial stability group $\mathbf{G}_\Psi$ for $|\Psi\rangle$ can be taken into account by replacing Ξ with its group average over $\mathbf{G}_\Psi$

$$\overline{\Xi} = \frac{\int_{G_\Psi} dg\, U_g \Xi U_g^\dagger}{\int_{G_\Psi} dg}. \tag{13}$$

Notice that the value of the likelihood functional remains unchanged after this replacement, and the group average is still optimal [it is easy to show that the same occurs with $\mathcal{M}(\rho)$ in the case of a general goal function]. As a consequence of the Wedderburn decomposition (12), the completeness constraint (2) for Ξ can be written as

$$\mathrm{Tr}[I_{\mu\nu}\Xi] = \delta_{\mu\nu} b_\mu^{-1}\ \forall\ \mu \sim \nu. \tag{14}$$

It is now convenient to decompose the input state $|\Psi\rangle$ over the invariant subspaces H_μ of the representation as $|\Psi\rangle = \Sigma_\mu c_\mu |\Psi_\mu\rangle$. This allows us to simply derive the following chain of inequalities:

$$\mathcal{L}_\Psi[\Xi] = \sum_{\mu,\nu} c_\mu^* c_\nu \langle\Psi_\mu|\Xi|\Psi_\nu\rangle \leqslant \sum_{\mu,\nu} |c_\mu||c_\nu||\xi_{\mu\nu}| \leqslant \sum_{\mu,\nu} |c_\mu| \times |c_\nu|\sqrt{\xi_{\mu\mu}\xi_{\nu\nu}} \leqslant \left(\sum_\mu |c_\mu|\sqrt{b_\mu^{-1}}\right)^2 \leqslant \sum_\mu b_\mu^{-1},$$

where the sums range in the set M_Ψ of all invariant subspaces which are nonorthogonal to $|\Psi\rangle$, $\mathcal{L}_\Psi[\Xi]$ denotes the likelihood functional defined by the pure state $|\Psi\rangle$, and $\xi_{\mu\nu}$ denotes the matrix element $\langle\Psi_\mu|\Xi|\Psi_\nu\rangle$. The first inequality can be saturated by the choice $\xi_{\mu\nu}=e^{i(\vartheta_\mu-\vartheta_\nu)}|\xi_{\mu\nu}|$ where ϑ_μ is the phase of c_μ. The second inequality is a necessary condition for positivity of Ξ, and saturates for $|\xi_{\mu\nu}|=\sqrt{\xi_{\mu\mu}\xi_{\nu\nu}}$ (notice that this inequality is not also a sufficient condition for positivity, whence the positivity of the optimal Ξ must be checked *a posteriori*). The third inequality is due to the fact that $\xi_{\mu\mu} \leqslant \mathrm{Tr}[I_{\mu\mu}\Xi]=b_\mu^{-1}$. Finally, the last Schwartz inequality sets the following general upper bound for the maximum likelihood of covariant measurements:

$$\mathcal{L}_\Psi[\Xi] \leqslant \sum_{\mu\in \mathrm{M}_\Psi} b_\mu^{-1}. \tag{15}$$

In the case of a compact group the inequality (15) implies that the likelihood is always less than the sum of dimensions of invariant subspaces supporting $|\Psi\rangle$. For infinite dimensions, on the other hand, the bound (15) and the likelihood itself may diverge. One can see now that the following choice of the operator Ξ:

$$\Xi = |\eta\rangle\langle\eta|, \quad |\eta\rangle = \sum_{\mu\in \mathrm{M}_\Psi} e^{i\vartheta_\mu}\sqrt{b_\mu^{-1}}|\Psi_\mu\rangle, \tag{16}$$

attains the bound $(\Sigma_{\mu\in \mathsf{M}_\Psi}|c_\mu|\sqrt{b_\mu^{-1}})^2$ for the likelihood functional. Note that, if $|\Psi\rangle$ has no component in some irreducible subspace H_ν, then the operator Ξ must be extended to the whole space H, in order to satisfy the constraints $\mathrm{Tr}[I_{\mu\mu}\Xi]=b_\mu^{-1}$ for all μ. Obviously, such extension is generally not unique, e.g., one can take

$$\Xi = |\eta\rangle\langle\eta| + \sum_{\nu\notin \mathrm{M}_\Psi} b_\nu^{-1}|\Phi_\nu\rangle\langle\Phi_\nu|, \tag{17}$$

where $|\Phi_\nu\rangle$ is any normalized vector in H_ν, which both guarantees $\Xi \geqslant 0$ and satisfies the constraints $\mathrm{Tr}[I_{\mu\mu}\Xi]=b_\mu^{-1}$ for all μ. Notice that the presence of equivalent representations in Eq. (17) generally improves the likelihood (this feature was missed in Refs. [15–17]).

If there are no equivalent representations in the decomposition of $|\Psi\rangle$, then the kernel (17) averaged over the stability subgroup $\mathbf{G}_\Psi$ of $|\Psi\rangle$ is optimal. However, in the presence of equivalent representations, one also wants the off-diagonal constraints $\mathrm{Tr}[I_{\mu\nu}\Xi]=0$ to be satisfied $\forall\ \mu\sim\nu$. One can see that the kernel in Eq. (17) satisfies also the off-diagonal constraints when the decomposition $|\Psi\rangle=\Sigma_\mu c_\mu|\Psi_\mu\rangle$ satisfies

$$\langle\Psi_\mu|I_{\mu\nu}|\Psi_\nu\rangle = 0, \quad \mu\sim\nu. \tag{18}$$

As shown in the Appendix, the subspaces carrying equivalent irreducible components of the representation can always be chosen in such a way as to satisfy Eq. (18). It is worth noticing that the present "canonical" form for maximum likelihood measurements generalizes the case of the optimal covariant phase estimation given by Holevo [19], further generalized in Ref. [30]. Finally, notice that the result derived here also holds for discrete groups, such as the permutation group or $\mathbb{Z}_d\times\mathbb{Z}_d$ by just substituting integrals with sums.

V. EXAMPLES

While it is obvious that averaging the result over a number $N>1$ of equally prepared identical copies always improves the precision of estimation—either classically or not—a legitimate question is whether nonindependent measurements on copies can be exploited to further enhance the precision, compared to this conventional independent measurement scheme. For the maximum likelihood strategy, when measurements are performed independently on each copy, in order to estimate a d-dimensional parameter x, the value of the likelihood is bounded as follows:

$$\mathcal{L}_{\rm av}^{(N)} \equiv p\left(\frac{\sum_{i=1}^{N} x_i}{N} = x \middle| x\right)$$
$$= \int dx_1 \cdots dx_N p(x_1|x) \cdots p(x_N|x) \delta\left(\frac{\sum_{i=1}^{N} x_i}{N} - x\right)$$
$$= N^d \int dx_2 \cdots dx_N p(Nx - x_2 - \cdots - x_N|x)$$
$$\times p(x_2|x) \cdots p(x_N|x) \leq N^d \max_{x'} \{p(x'|x)\}. \qquad (19)$$

Unfortunately, the bound in Eq. (19) is generally not achievable, and for increasing value of the dimensionality d of the parameter x it becomes quite loose. However, the optimal measurement on N copies of the same state can achieve a higher value of the likelihood with respect to the *semiclassical* scheme involving independent measurements. Moreover, the case of preparation in different input states can lead to further improvement in the estimation of the group transformation U_g, since the decomposition of the global state may involve a larger number of invariant subspaces than just those belonging to the symmetric space.

A. Universal state estimation

1. SU(d)-covariant estimation: Pure state estimation

The estimation of a pure state in a finite dimensional Hilbert space H can be regarded as a covariant estimation with respect to the defining representation of the group SU(d), where $d=\dim(\mathsf{H})$. Indeed, the orbit of a given pure state contains all pure states of H. Clearly, the optimal kernel is $\Xi = d|\psi\rangle\langle\psi|$, according to Refs. [2,19], and consequently the value of the maximum likelihood is $\mathcal{L}^{(1)} = d$.

2. Pure state estimation with N>1 copies in the same state

This corresponds to the case of estimation of the group element $g \in \mathrm{SU}(d)$ in the reducible representation $U_g^{\otimes N}$ with initial state $|\Psi\rangle = |\psi\rangle^{\otimes N}$. There are inequivalent components corresponding to the symmetric subspace $(\mathsf{H}^{\otimes N})_+$, along with all other permutation invariant subspaces. Since $|\Psi\rangle$ belongs to the symmetric subspace $(\mathsf{H}^{\otimes N})_+$, the optimal Ξ is not unique, e.g., we can take $\Xi = d_+(|\psi\rangle\langle\psi|)^{\otimes N} + I_{\mathsf{W}}$, where $d_+ = \binom{d+N-1}{d-1}$ is the dimension of $(\mathsf{H}^{\otimes N})_+$ and W is the orthogonal complement of $(\mathsf{H}^{\otimes N})_+$. In any case we have $\mathcal{L}^{(N)} = \dim(\mathsf{H}^{\otimes N})_+$. Notice that the POVM is not separable, due to the presence of the orthogonal projector I_{W}.

3. SU(d) estimation with two copies in different states

In this case $|\Psi\rangle = |\psi\rangle|\phi\rangle$ can be decomposed as $\sqrt{(1+s^2)/2}|\Psi_+\rangle + \sqrt{(1-s^2)/2}|\Psi_-\rangle$, where $s = |\langle\psi|\phi\rangle|$ and $|\Psi_\pm\rangle = [1/\sqrt{2(1\pm s^2)}](|\psi\rangle|\phi\rangle \pm |\phi\rangle|\psi\rangle)$. Then the optimal kernel Ξ is proportional to the projector onto the vector $|\eta\rangle = \sqrt{d_+}|\Psi_+\rangle + \sqrt{d_-}|\Psi_-\rangle$ and the likelihood takes the value $(\sqrt{d_+(1+s^2)/2} + \sqrt{d_-(1-s^2)/2})^2 \leq d^2$ (by the Schwartz inequality), d_- denoting the dimension of the antisymmetric Hilbert space. It is easily seen that this bound can be attained by choosing $s^2 = 1/d$. The optimal POVM is separable (the optimal kernel is actually *factorized* as $\Xi = d^2|\psi\rangle\langle\psi| \otimes |\phi\rangle \times \langle\phi|$). No further improvement can be achieved, since a likelihood greater than d^2 is not compatible with the completeness of the POVM (in fact $\mathcal{L}_\Psi[\Xi] \leq \mathrm{Tr}[\Xi] = d^2$).

B. Weyl-Heisenberg covariant estimation

1. Estimation of displacement on the phase space

This case corresponds to consider the Weyl-Heisenberg irreducible representation $\{D(z)\}$ of the translation group on the complex plane, $D(z)$ denoting the displacement operator $D(z) = e^{za^\dagger - z^* a}$ with $[a, a^\dagger] = 1$. Being noncompact, the representation space H is infinite dimensional. Physically $D(z)$ represents a joint shift of position and momentum of a quantum harmonic oscillator, and the covariant state estimation corresponds to a joint measurement of position and momentum. Here one has $b = \int_{\mathbb{C}} (d^2z/\pi) |\langle n|D(z)|n\rangle|^2$, where $|n\rangle$ denotes an element of any orthonormal basis for H, which we can conveniently take as the set of eigenstates of the number operator $a^\dagger a$. Choosing $n=0$ one obtains $b=1$, whence the optimal kernel for initial state $|\psi\rangle$ is $\Xi = |\psi\rangle\langle\psi|$ and the maximum likelihood is $\mathcal{L}[\Xi] = 1$. Notice that for $|\psi\rangle = |0\rangle$ we get the well known coherent-state POVM describing the heterodyne measurement [12,13].

2. Estimation of displacement with identical shifts on N>1 quantum oscillators

This case corresponds to the tensor representation $\{D^{\otimes N}(z)\}$ of the Weyl-Heisenberg group. The irreducible representations can be easily obtained by the linear change of modes represented by the unitary transformation

$$V = e^{\phi_N [a_1^\dagger (a_2 + \ldots + a_N) - a_1 (a_2^\dagger + \ldots + a_N^\dagger)]},$$

with $\phi = (1/\sqrt{N-1})$ arctan $\sqrt{N-1}$ so that $VD^{\otimes N}(z)V^\dagger = D(\sqrt{N}z) \otimes I^{\otimes(N-1)}$. Then the irreducible subspaces are given by $\mathsf{H}_n = \{V^\dagger|\varphi\rangle \otimes |\Phi_n\rangle, |\varphi\rangle \in \mathsf{H}\}$, where $\{|\Phi_n\rangle\}$ is an orthonormal basis for $\mathsf{H}^{\otimes(N-1)}$. The formal dimension coefficients are easily obtained as follows:

$$b_n = \int_{\mathbb{C}} \frac{d^2z}{\pi} |\langle 0|\langle\Phi_n|VD^{\otimes N}(z)V^\dagger|0\rangle|\Phi_n\rangle|^2$$
$$= \int_{\mathbb{C}} \frac{d^2z}{\pi} |\langle 0|\langle\Phi_n|D(\sqrt{N}z) \otimes I^{\otimes(N-1)}|0\rangle|\Phi_n\rangle|^2$$
$$= \frac{1}{N}\int_{\mathbb{C}} \frac{d^2z}{\pi} |\langle 0|D(z)|0\rangle|^2 = \frac{1}{N}.$$

Since the invariant subspaces carry all equivalent representations—the isomorphism between two of them is $I_{mn} = V^\dagger(I \otimes |\Phi_m\rangle\langle\Phi_n|)V$—the problem of choosing a suitable decomposition of the initial state $|\Psi\rangle$ in irreducible representations arises. In the general case, one should apply the full construction given in the Appendix, while a simpler solution is possible for states of the form $|\Psi\rangle=|i_1\rangle|i_2\rangle\cdots|i_N\rangle$. In this case, one has only to write $V|\Psi\rangle=\Sigma_{i_1,i_2,\ldots,i_N}c_{i_1 i_2\ldots i_N}|i_1\rangle|i_2\rangle\cdots|i_N\rangle$, and to define $|\Phi_n\rangle=C_n^{-1}\Sigma_{i_2\ldots i_N}c_{ni_2\ldots i_N}|i_2\rangle\cdots|i_N\rangle$, where $C_n=\sqrt{\Sigma_{i_2,\ldots,i_N}|c_{ni_2\ldots i_N}|^2}$, obtaining the desired decomposition $|\Psi\rangle=V^\dagger\Sigma_n C_n|n\rangle|\Phi_n\rangle$ (notice that $\langle\Phi_m|\Phi_n\rangle=\delta_{mn}$ since they are eigenstates corresponding to different eigenvalues of the number operator). The value of the likelihood is then $\mathcal{L}[\Xi]=N(\Sigma_n C_n)^2$.

We now consider two special cases.

(i) N copies of vacuum state $|0\rangle$. This case corresponds to the estimation of the complex shift z on the set $\{|z\rangle^{\otimes N}\}$ of N copies of a coherent state $|z\rangle$. Here, the vacuum state $|\Psi\rangle=|0\rangle^{\otimes N}$ belongs to just one invariant subspace, since $V|\Psi\rangle=|\Psi\rangle$. The optimal kernel is not unique, and is given by any completion of $\Xi=N(|0\rangle\langle 0|)^N$, and the pertaining likelihood value is N. For $N=2$ it can be shown that an optimal POVM corresponds to averaging the outcomes of independent heterodyne measurements on two copies, while another optimal one corresponds to the independent measurement of the position $(1/2)(a_1+a_1^\dagger)$ and the momentum $(1/2i)(a_2-a_2^\dagger)$, taking as the outcome $\alpha=x+iy$, where x and y are the two separate outcomes.

(ii) Two copies of a number state: $|\Psi\rangle=|n\rangle|n\rangle$ with $n>0$. The maximum value of the likelihood is $\mathcal{L}[\Xi]=2\left[\Sigma_{k=0}^n(1/2^n n!)\binom{n}{k}\sqrt{(2k)!(2n-2k)!}\right]^2$, and numerical calculation shows an asymptotic linear behavior versus n. In the case of two copies of a one-photon state $|\Psi\rangle=|1\rangle|1\rangle$. Decomposing the seed state we obtain $|\Psi\rangle=-(\sqrt{2}/2)(V^\dagger|20\rangle+V^\dagger|02\rangle)$: an example of the optimal kernel is then $\Xi=2[2(|1\rangle\langle 1|)^{\otimes 2}+\Sigma_{i\neq 0,2}V^\dagger|0i\rangle\langle 0i|V]$, achieving a likelihood equal to 4, which can be shown to be twice the semiclassical value.

3. Estimation of displacement on two copies, with identical shifts in position and opposite shifts in momentum

This case corresponds to the representation $\{V(z)=D(z)\otimes D(z^*)\}$, which is reducible, but does not possess any irreducible proper component in $\mathsf{H}^{\otimes 2}$, and thus is beyond the hypotheses of our general results. In fact, the irreducible representations are all inequivalent, and make a continuous set, each component being supported by the Dirac-normalized eigenvectors [31] $(1/\sqrt{\pi})|D(w)\rangle\rangle=(1/\sqrt{\pi})\Sigma_{m,n}\langle m|D(w)|n\rangle|m\rangle|n\rangle$ of the normal operator $W=a\otimes I-I\otimes a^\dagger$ (the heterodyne photocurrent [12,13,32]). Upon expanding the operators $V(z)=\exp(zW^\dagger-z^*W)$ over the Dirac-orthonormal basis, one has

$$\int_{\mathbb{C}}\frac{d^2z}{\pi}V(z)OV^\dagger(z) =\int_{\mathbb{C}}\frac{d^2z}{\pi}\int_{\mathbb{C}}\frac{d^2w}{\pi}\int_{\mathbb{C}}\frac{d^2w'}{\pi}e^{z(w-w')^*-z^*(w-w')}|D(w)\rangle\rangle\langle\langle D(w)| \times O|D(w')\rangle\rangle\langle\langle D(w')| =\int_{\mathbb{C}}\frac{d^2w}{\pi}|D(w)\rangle\rangle\langle\langle D(w)|O|D(w)\rangle\rangle\langle\langle D(w)|,$$

namely, a continuous version of the Wedderburn decomposition still holds,

$$\int_{\mathbb{C}}\frac{d^2z}{\pi}V(z)OV^\dagger(z)=\int_{\mathbb{C}}d^2w\ a_wP_w,$$

for any O such that $\mathrm{Tr}[P_wO]<\infty$, with $P_w=|D(w)\rangle\rangle\langle\langle D(w)|$ and $a_w=\pi^{-1}\mathrm{Tr}[P_wO]$ (in a proper mathematical setting the integral over w in the last equation should be interpreted as a direct integral). The maximum likelihood covariant measurement for state estimation among the set generated by the seed $|\Psi\rangle\in\mathsf{H}^{\otimes 2}$ is given by the *entangled* kernel $\Xi=|\eta\rangle\langle\eta|$, where

$$|\eta\rangle=\int_{\mathbb{C}}\frac{d^2w}{\pi}e^{i\theta_w}|D(w)\rangle\rangle,$$

which is the analog of Eq. (16) for a continuous spectrum [as in that previous case, θ_w is the phase of $\langle\langle D(w)|\Psi\rangle$]. It is worth noticing that for $|\Psi\rangle=|0\rangle|0\rangle$ the problem corresponds to estimating the amplitude z of the set of coherent states $\{|z\rangle|z^*\rangle\}$, and the value of the likelihood for the optimal measurement is 4, namely, twice the likelihood for the amplitude estimation for identical states $\{|z\rangle|z\rangle\}$. The probability distributions are indeed Gaussian in both cases, but the variance in this case is half the variance of the Gaussian for the states $|z\rangle|z\rangle$. The fidelity of the estimate is 2/3 for the states $|z\rangle|z\rangle$, while it is 4/5 for $|z\rangle|z^*\rangle$. The last example can be regarded as the "continuous-variable" analog of the measurement of the direction of two antiparallel spins by Gisin and Popescu [26], as previously studied in Ref. [33].

4. Estimation of displacement on one part of a bipartite entangled system

We consider here the representation $\{D(z)\otimes I\}$ acting on two optical modes. In this case the invariant subspaces are $\mathsf{H}_n=\{|\psi\rangle\otimes|\varphi_n\rangle,|\psi\rangle\in\mathsf{H}\}$, where $\{|\varphi_n\rangle\}$ is any orthonormal basis in H, all of them supporting equivalent representations, and the formal dimensions b_n are all equal to 1. If we take a twin beam $|\Psi\rangle=\sqrt{1-x^2}\Sigma_n x^n|n\rangle|n\rangle$ as initial state, its decomposition is trivial, and $|n\rangle|n\rangle$ are precisely the components $|\Psi_n\rangle$ on the irreducible subspaces. Then the optimal POVM is given by $|\eta\rangle=\Sigma_n|n\rangle|n\rangle$, namely, it is the two-mode heterodyne POVM [12,31,32]. Correspondingly, the value of the likelihood is $\mathcal{L}[\Xi]=(1+x)/(1-x)$, showing a strong enhancement by the effect of entanglement in agreement with Ref. [23].

VI. CONCLUSIONS

By group theoretic arguments we have derived the class of measurements of covariant parameters that are optimal according to the maximum likelihood criterion. The optimization problem has been completely resolved for pure states under the simple hypotheses of unimodularity of the group and measurable stability group. The general method has been applied to the case of finite dimensional quantum state estimation with many input copies, and, for infinite dimensions, to the Weyl-Heisenberg covariant estimation, also giving a continuous-variable analog of the estimation of the direction on two antiparallel spins by Gisin and Popescu. The increasing statistical efficiency with the number of copies N is essentially related to two factors: (i) the way in which the dimension of the N-fold tensor product Hilbert space increases versus N; (ii) its decomposition into irreducible subspaces. Moreover, the nonseparability of the optimal measurement—either in its POVM or in the optimal states—is strictly related to the structure of the group representation.

We conclude by mentioning that the optimal covariant estimation for mixed input states is still an open problem, and an explicit analytical optimization seems a very difficult task. For the case of phase estimation the problem can be analytically solved in the special instance of states which are *phase pure* [34]. For a general covariance group representation the concept of a phase-pure state can be generalized by choosing a vector $|\Psi_\mu\rangle$ for each invariant subspace such that for every $\mu\sim\nu$ one has $\langle\Psi_\mu|I_{\mu\nu}|\Psi_\nu\rangle=0$. Then every state ρ satisfying $\mathrm{Supp}\{\rho\}\subset\mathrm{Span}\{|\Psi_\mu\rangle\}$ and $\langle\Psi_\mu|\rho|\Psi_\nu\rangle=e^{i(x_\mu-x_\nu)}|\langle\Psi_\mu|\rho|\Psi_\nu\rangle|$ behaves as a pure state in all respects, so that the upper bound and the canonical form of the optimal POVM still hold. It is likely that a generalization of this approach may extend the validity of the present solution of the covariant estimation problem for a special class of mixed states.

ACKNOWLEDGMENTS

We thank M. Guta for useful and stimulating discussions. This work has been sponsored by INFM through the Project PRA-2002-CLON, and by EC and MIUR through the cosponsored ATESIT Project IST-2000-29681 and Cofinanziamento 2003.

APPENDIX

In this appendix we show how to chose invariant subspaces H_μ of equivalent irreducible representations in order to satisfy Eq. (18), namely, $\langle\Psi_\mu|I_{\mu\nu}|\Psi_\nu\rangle=0$, for the decomposition $|\Psi\rangle=\Sigma_\mu c_\mu|\Psi_\mu\rangle$, in such a way that all invariant subspaces effectively behave as supporting inequivalent irreducible representations. This choice of invariant subspaces guarantees that every operator $\Xi=|\eta\rangle\langle\eta|$, where $|\eta\rangle$ is any linear combination of $|\Psi_\nu\rangle$, will satisfy the constraints $\mathrm{Tr}[\Xi I_{\mu\nu}]=0$ for any $\mu\sim\nu$. This method will allow one to extend the general treatment of the phase estimation problem given in Ref. [30] to any square-summable group representation.

Let us consider an irreducible component of a unitary representation of $\mathbf{G}$ with multiplicity $m\leqslant\infty$, and denote by $\mathsf{H}^{(\omega)}$ the invariant subspace carrying all equivalent irreducible components, and by $\mathsf{H}^{(\omega)}=\oplus_{\mu=1}^m\mathsf{H}_\mu$ a given choice of invariant orthogonal subspaces H_μ, each carrying an equivalent irreducible representation. Since inequivalent irreducible components already satisfy Eq. (18), we can just focus attention on the component $|\Psi_\omega\rangle$ of $|\Psi\rangle$ on $\mathsf{H}^{(\omega)}$.

Let us denote by $I_{\mu\nu}$ the isomorphisms mapping H_ν into H_μ, and satisfying $[I_{\mu\nu},U_g]=0\ \forall\ g\in\mathbf{G}$. As already mentioned, it is always possible to choose an orthonormal basis $\mathsf{B}_\mu=\{|e_n^{(\mu)}\rangle\}$ for each subspace H_μ in such a way that $\mathsf{B}_\mu=I_{\mu\nu}\mathsf{B}_\nu$ for any μ,ν, where equality between bases is defined elementwise, i.e., $|e_n^{(\mu)}\rangle=I_{\mu\nu}|e_n^{(\nu)}\rangle$ for all n. We have now the following simple lemma.

Lemma (choice of the decomposition into equivalent components). For each unitary matrix $\{V_{\mu\nu}\}\in\mathbb{M}_m$ the linear combinations $\mathsf{B}'_\mu=\Sigma_\nu V_{\mu\nu}\mathsf{B}_\nu$ provide a new decomposition $\mathsf{H}^{(\omega)}=\oplus_{\mu=1}^m\mathsf{H}'_\mu$ of $\mathsf{H}^{(\omega)}$ into subspaces supporting equivalent irreducible components, where $\mathsf{H}'_\mu\equiv\mathrm{Span}(\mathsf{B}'_\mu)$.

Proof. The subspaces H'_μ are orthogonal. In fact, upon defining $\mathsf{B}'_\mu=\{|f_l^{(\mu)}\rangle\}$, we obtain

$$\langle f_l^{(\mu)}|f_n^{(\nu)}\rangle=\sum_{\alpha,\beta}\langle e_l^{(\alpha)}|V_{\mu\alpha}^*V_{\nu\beta}|e_n^{(\beta)}\rangle=\delta_{ln}\sum_\alpha(V_{\nu\alpha}V_{\alpha\mu}^\dagger)=\delta_{ln}\delta_{\mu\nu}.$$

Moreover, each H'_μ carries a representation equivalent to that of, say, H_1. In fact, the operator $S_{\mu1}\equiv\Sigma_\nu V_{\mu\nu}I_{\nu1}$ is indeed an isomorphism between the subspaces H_1 and H'_μ, since it defines a one-to-one correspondence between them via $\mathsf{B}'_\mu=S_{\mu1}\mathsf{B}_1$, and commutes with U_g for all $g\in\mathbf{G}$, since each $I_{\mu\nu}$ commutes. This proves that the spaces $\{\mathsf{H}'_\mu\}$ provide a new orthogonal decomposition $\mathsf{H}^{(\omega)}=\oplus_\mu\mathsf{H}'_\mu$ into invariant subspaces carrying equivalent components of the representation.

Now, let us consider the component $|\Psi_\omega\rangle$ of $|\Psi\rangle$ on $\mathsf{H}^{(\omega)}$, and write its decomposition using the set of bases $\{\mathsf{B}_\mu\}$ as follows:

$$|\Psi_\omega\rangle=\sum_{\mu n}\Psi_{\mu n}^\omega|e_n^{(\mu)}\rangle. \tag{A1}$$

We want to construct a new decomposition H'_ν of $\mathsf{H}^{(\omega)}$ such that the components of $|\Psi_\omega\rangle$ on invariant subspaces satisfy Eq. (18), namely, they behave as belonging to inequivalent representations. This can be done as follows. Define $\infty\geqslant d=\dim(\mathsf{H}_\mu)$ and consider the $m\times d$ matrix $\Psi^\omega=\{\Psi_{\mu n}^\omega\}$. According to Eq. (A1), Ψ^ω is a Hilbert-Schmidt operator, and hence we can write the singular value decomposition

$$\Psi^\omega=V^T\Sigma U, \tag{A2}$$

where Σ is a $m\times d$ matrix with all vanishing off-diagonal elements, and V and U are $m\times m$ and $d\times d$ unitaries, respectively. From Eq. (A2) one obtains $\Psi_{\mu n}^\omega=\Sigma_{\nu l}V_{\nu\mu}\sigma_\nu\delta_{\nu l}U_{ln}$, where the sums run from 1 to $r=\mathrm{rank}(\Psi^\omega)\leqslant\min(m,d)$. Equation (A1) is then rewritten

$$|\Psi_\omega\rangle=\sum_{\nu l}\sigma_\nu\delta_{\nu l}|g_l^{(\nu)}\rangle, \tag{A3}$$

where

$$|g_l^{(\nu)}\rangle=\sum_{\mu n}V_{\nu\mu}U_{ln}|e_n^{(\mu)}\rangle. \tag{A4}$$

The new bases $\mathsf{B}'_\nu=\{|g_l^{(\nu)}\rangle\}$ provide a new decomposition into invariant subspaces H'_ν supporting equivalent components. In fact, starting from the set $\{\mathsf{B}_\mu\}$ the first unitary transformation U over each basis preserves the relations $\mathsf{B}_\mu=I_{\mu\nu}\mathsf{B}_\nu$, whereas the second transformation V, according to the previous lemma, gives the new decomposition $\mathsf{H}^{(\omega)}=\oplus_{\mu=1}^m\mathsf{H}'_\mu$, with $\mathsf{H}'_\mu \equiv \mathrm{Span}(\mathsf{B}'_\mu)$. The state $|\Psi_\omega\rangle$ in Eq. (A3) then satisfies

$$\langle \Psi_\mu | I_{\mu\nu} | \Psi_\nu \rangle = 0 \text{ for } \mu \sim \nu$$

since the spaces $\mathsf{H}'_\nu = \mathrm{Span}\{|g_k^{(\nu)}\rangle\}$ have been chosen such that each component of $|\Psi_\omega\rangle$ on H'_ν is just proportional to a single element $|g_l^{(\nu)}\rangle$ of the orthonormal basis with different l for different ν.

Notice that, if the group is noncompact and there is an infinite number of equivalent irreducible subspaces, the spectrum of singular values of Ψ may be continuous, and the sums in the above derivation must be replaced by integrals, with some care in the generalization of definition and theorems.

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

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

[3] D. Bruß, A. Ekert, and C. Macchiavello, Phys. Rev. Lett. **81**, 2598 (1998).

[4] G. M. D'Ariano, C. Macchiavello, and M. F. Sacchi, J. Opt. B: Quantum Semiclassical Opt. **3**, 44 (2001).

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

[6] G. M. D'Ariano and H. P. Yuen, Phys. Rev. Lett. **76**, 2832 (1996).

[7] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).

[8] H. P. Yuen, Phys. Lett. **113A**, 405 (1986).

[9] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Europhys. Lett. **65**, 165 (2004).

[10] M. A. Naimark, Izv. Akad. Nauk SSSR, Ser. Mat. **4**, 277 (1940).

[11] F. Buscemi, G. M. D'Ariano, and M. F. Sacchi, Phys. Rev. A **68**, 042113 (2003).

[12] H. P. Yuen and J. H. Shapiro, IEEE Trans. Inf. Theory **24**, 657 (1978); **25**, 179 (1979); **26**, 78 (1980).

[13] G. M. D'Ariano, in *Concepts and Advances in Quantum Optics and Spectroscopy of Solids*, edited by T. Hakioğlu and A. S. Shumovsky (Kluwer, Amsterdam 1997), pp. 139–174.

[14] G. M. D'Ariano, P. Lo Presti, and M. F. Sacchi, Phys. Lett. A **292**, 233 (2002).

[15] A. Peres and P. F. Scudo, Phys. Rev. Lett. **86**, 4160 (2001).

[16] E. Bagan, M. Baig, and R. Muñoz-Tapia, Phys. Rev. Lett. **87**, 257903 (2001).

[17] E. Bagan, M. Baig, and R. Muñoz-Tapia, Phys. Rev. A **64**, 022305 (2001).

[18] L. Susskind and J. Glogower, Physics (Long Island City, N.Y.) **1**, 49 (1964).

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

[20] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **59**, 1070 (1999).

[21] R. Derka, V. Bužek, and A. K. Ekert, Phys. Rev. Lett. **80**, 1571 (1998).

[22] A. Chefles, Phys. Rev. A **64**, 062305 (2001).

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

[24] A. S. Holevo, e-print quant-ph/0307225.

[25] R. D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).

[26] N. Gisin and S. Popescu, Phys. Rev. Lett. **83**, 432 (1999).

[27] For compact $\mathbf{G}_\Psi$, which is always our case (even for noncompact $\mathbf{G}$), then an invariant measure on $\mathfrak{X}$ can be explicitly constructed using the map $g(x)$ [19]. The measure dx is defined by the requirement that $\int_{\mathfrak{X}} f(x)dx = \int_{\mathbf{G}} f(gx_0)dg$, for any function f integrable on $\mathfrak{X}$, where x_0 is a fixed element in $\mathfrak{X}$. Compactness of $\mathbf{G}_\Psi$ ensures normalizability of dx, and invariance follows from left invariance of dg. Moreover, right invariance of dg implies that dx is independent of the particular choice of x_0.

[28] A. Grossmann, J. Morlet, and T. Pauli, J. Math. Phys. **26**, 2473 (1985).

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

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

[31] G. M. D'Ariano and M. F. Sacchi, Phys. Rev. A **52**, R4309 (1995).

[32] J. H. Shapiro and S. S. Wagner, IEEE J. Quantum Electron. **QE20**, 803 (1984).

[33] N. J. Cerf and S. Iblisdir, Phys. Rev. A **64**, 032307 (2001).

[34] G. M. D'Ariano, C. Macchiavello, P. Perinotti, and M. F. Sacchi, Phys. Lett. A **268**, 241 (2000).