

# Group theoretic structures in the estimation of an unknown unitary transformation

Giulio Chiribella

Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, N2L 2Y5, Canada

**Abstract.** This paper presents a series of general results about the optimal estimation of physical transformations in a given symmetry group. In particular, it is shown how the different symmetries of the problem determine different properties of the optimal estimation strategy. The paper also contains a discussion about the role of entanglement between the representation and multiplicity spaces and about the optimality of square-root measurements.

## 1. Introduction

The estimation of an unknown unitary transformation in a given symmetry group is a general problem related to many stimulating topics, such as high-precision measurements [1], coherent states and uncertainty relations [2], quantum clocks [3], quantum gyroscopes [4] and quantum reference frames [5]. The aim of this paper is to provide a synthetic account of the general theory of optimal estimation for symmetry groups, presenting new proofs and highlighting the underlying group theoretical and algebraic structures into play.

### 1.1. Prologue: dense coding

Let us start with a simple example. Suppose that we have at disposal a quantum system with two-dimensional Hilbert space  $\mathcal{H} = \mathbb{C}^2$  and a black box that performs an unknown transformation  $\mathcal{U}_{ij}, i, j \in \{0, 1\}$ , defined on the set  $\text{St}(\mathcal{H})$  of all density matrices on  $\mathcal{H}$  as  $\mathcal{U}_{ij}(\rho) := U_{ij}\rho U_{ij}^\dagger, \forall \rho \in \text{St}(\mathcal{H})$ , where  $\{U_{ij}\}_{i,j=0}^1$  are the unitary matrices

$$U_{00} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U_{01} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad U_{10} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad U_{11} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The value of the indices  $(i, j)$  is unknown to us and our goal is to find it out using the black box only once. Here the natural figure of merit is the minimization of the probability of error.

Clearly, if we apply the black box to a state  $\rho \in \text{St}(\mathcal{H})$  there will always be an error: the four states  $\{\rho_{ij} := \mathcal{U}_{ij}(\rho)\}_{i,j=0}^1$  will never be perfectly distinguishable (in a two-dimensional Hilbert space there cannot be more than two perfectly distinguishable states!). However, if we introduce a second system  $\mathcal{K} \simeq \mathcal{H}$  and apply the unknown transformation to the projector on the entangled vector  $|\Phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \in \mathcal{H} \otimes \mathcal{K}$ , where  $\{|0\rangle, |1\rangle\}$  is the standard basis for  $\mathbb{C}^2$ , we obtain four states (the projectors on the vectors  $|\Phi_{ij}\rangle := (U_{ij} \otimes I)|\Phi\rangle$ ) that are perfectly distinguishable:  $\langle \Phi_{ij} | \Phi_{kl} \rangle = \text{Tr}[U_{ij}^\dagger U_{kl}]/2 = \delta_{ik}\delta_{jl}$ . This remarkable fact was observed

for the first time by Bennett and Wiesner, who exploited it to construct a protocol known as *dense-coding* [6]. The moral of dense coding is that the use of entanglement with an ancillary system can improve dramatically the discrimination of unknown transformations. Despite the extreme simplicity of the mathematics, this curious example points out structures that are much deeper than it might seem to the first sight. The aim of the present paper is to illustrate these structures in the general context of the estimation of an unknown group transformation. Our analysis will include dense coding, where the group of interest is the Klein group  $\mathbb{Z}_2 \times \mathbb{Z}_2$ .

## 2. General problem: estimation of an unknown group transformation

### 2.1. The problem

Suppose that we have at disposal one use of a black box and that we want to identify the action of the black box. Suppose also that we have some prior knowledge about the black box: in particular, we know that

- (i) the black box acts on a quantum system with finite dimensional Hilbert space  $\mathcal{H} \simeq \mathbb{C}^d$ ,  $d < \infty$ .
- (ii) it performs a deterministic transformation  $\mathcal{U}_g$  belonging to a given representation of a given symmetry group  $\mathbf{G}$ .

Mathematically, a *deterministic transformation* (also known as *quantum channel*) is described by a completely positive trace-preserving linear map  $\mathcal{C}$  acting on the set  $\text{St}(\mathcal{H})$  of quantum states (non-negative matrices with unit trace) on the Hilbert space  $\mathcal{H}$ . The representation of the group  $\mathbf{G}$  is then given a function  $\mathcal{U} : g \mapsto \mathcal{U}_g$  from  $\mathbf{G}$  to the set of quantum channels, with the usual requirements

$$\begin{aligned} \mathcal{U}_e &= \mathcal{I}_{\mathcal{H}} \\ \mathcal{U}_g \mathcal{U}_{g^{-1}} &= \mathcal{I}_{\mathcal{H}} & \forall g \in \mathbf{G} \\ \mathcal{U}_g \mathcal{U}_h &= \mathcal{U}_{gh} & \forall g, h \in \mathbf{G} \end{aligned}$$

where  $e \in \mathbf{G}$  is the identity element of the group and  $\mathcal{I}_{\mathcal{H}}$  is the identity channel, given by  $\mathcal{I}_{\mathcal{H}}(\rho) := \rho, \forall \rho \in \text{St}(\mathcal{H})$ . In this paper the group  $\mathbf{G}$  will be always assumed to be either finite or compact.

### 2.2. Estimation strategies without ancillary systems

Since the group  $\mathbf{G}$  and the representation  $\{\mathcal{U}_g\}_{g \in \mathbf{G}}$  are both known, the problem here is to identify the group element  $g \in \mathbf{G}$  that specifies the action of the black box. How can we accomplish this task? A first idea is to prepare the system in a suitable input state  $\rho \in \text{St}(\mathcal{H})$  and to apply the unknown transformation  $\mathcal{U}_g$  to it, thus obtaining the output state  $\rho_g := \mathcal{U}_g(\rho)$ . The procedure can be represented diagrammatically as

$$\boxed{\rho} \xrightarrow{\mathcal{H}} \boxed{\mathcal{U}_g} \xrightarrow{\mathcal{H}} = \boxed{\rho_g} \xrightarrow{\mathcal{H}}$$

In this way, the output state  $\rho_g$  will carry some information about  $g$  and we can try to extract this information with a quantum measurement and to produce an estimate  $\hat{g} \in \mathbf{G}$ . The combination of the quantum measurement with our classical data processing can be described by a single mathematical object, namely a *positive operator-valued measure* (*POVM*, for short).

Let us denote by  $\text{Lin}(\mathcal{H})$  the set of linear operators on  $\mathcal{H}$  and by  $\text{Lin}_+(\mathcal{H}) \subset \text{Lin}(\mathcal{H})$  the set of non-negative operators. If  $\mathbf{G}$  is a finite group, a POVM with outcomes in  $\mathbf{G}$  is just a function  $P : \mathbf{G} \rightarrow \text{Lin}_+(\mathcal{H})$  sending the element  $\hat{g} \in \mathbf{G}$  to the non-negative operator  $P_{\hat{g}} \in \text{Lin}_+(\mathcal{H})$  and satisfying the requirement  $\sum_{\hat{g} \in \mathbf{G}} P_{\hat{g}} = I_{\mathcal{H}}$ , where  $I_{\mathcal{H}}$  is the identity on  $\mathcal{H}$ . The conditional probability of inferring  $\hat{g}$  when the true value is  $g$  is then given by the Born rule  $p(\hat{g}|g) = \text{Tr}[P_{\hat{g}}\rho_g]$ .

In general, if  $\mathbf{G}$  is a compact group and  $\sigma(\mathbf{G})$  is the collection of its measurable subsets [7], a POVM with outcomes in  $\mathbf{G}$  is an operator-valued measure  $P : \sigma(\mathbf{G}) \rightarrow \text{Lin}_+(\mathcal{H})$  sending the measurable subset  $B \in \sigma(\mathbf{G})$  to the non-negative operator  $P_B \in \text{Lin}_+(\mathcal{H})$  and satisfying the requirements

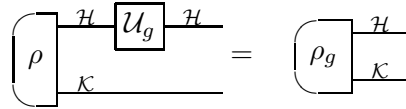
- $P_{\mathbf{G}} = I_{\mathcal{H}}$
- if  $B = \bigcup_{i=1}^{\infty} B_i$  and  $\{B_i\}_{i=1}^{\infty}$  are disjoint, then  $P_B = \sum_{i=1}^{\infty} P_{B_i}$ .

The conditional probability that the estimate  $\hat{g}$  lies in the subset  $B \in \sigma(\mathbf{G})$  is given by the Born rule  $p(B|g) = \text{Tr}[P_B \rho_g]$ .

In the following I will frequently use the notation  $P(d\hat{g})$  to indicate the POVM  $P$ . Accordingly, I will also write  $p(d\hat{g}|g) = \text{Tr}[P(d\hat{g})\rho_g]$  and  $P_B = \int_B P(d\hat{g})$ . In the case of finite groups, the notation  $P(d\hat{g})$  will be understood as synonym of  $P_{\hat{g}}$  and the integral over a subset  $B$  will be understood as a finite sum. The pair  $(\rho, P)$  of an input state and a POVM will be referred to as an *estimation strategy*.

### 2.3. Estimation strategies with ancillary systems

In the previous paragraph we discussed strategies where the unknown transformation was applied to a suitable state  $\rho \in \text{St}(\mathcal{H})$ . However, these strategies are not most general ones: we can introduce an ancillary system with Hilbert space  $\mathcal{K}$ , prepare a *bipartite* input state  $\rho \in \text{St}(\mathcal{H} \otimes \mathcal{K})$ , and then apply the black box on system  $\mathcal{H}$ , thus obtaining the output state  $\rho_g := (\mathcal{U}_g \otimes \mathcal{I}_{\mathcal{K}})(\rho)$ . The schematic of this procedure is



In this case, to obtain an estimate  $\hat{g}$  of the unknown group element  $g$  we will use a POVM  $P$  on the tensor product Hilbert space  $\mathcal{H} \otimes \mathcal{K}$ .

Classically, we would not expect any improvement from the use of an ancillary system because this system will not carry any information about the unknown transformation. However, here the state  $\rho$  can be an entangled state, that is, a state that cannot be written as a convex combination of tensor product states. As the example of dense coding teaches, the use of entanglement can improve the estimation dramatically. Note, however, that mathematically every estimation strategy with ancillary system can be reduced to the form of an estimation strategy without ancillary system by choosing  $\mathcal{H}' := \mathcal{H} \otimes \mathcal{K}$  and  $\mathcal{U}'_g := \mathcal{U}_g \otimes \mathcal{I}_{\mathcal{K}}$ .

### 2.4. Cost of an estimation strategy

To quantify how far is the guess  $\hat{g}$  from the true value one typically introduces a *cost function*  $c : \mathbf{G} \times \mathbf{G} \rightarrow \mathbb{R}$ ,  $(\hat{g}, g) \mapsto c(\hat{g}, g)$  which takes its minimum value when  $\hat{g} = g$ . If the true value is  $g$ , the expected cost of the estimation strategy  $(\rho, P)$  is

$$c(\rho, P|g) := \int_{\mathbf{G}} c(\hat{g}, g) p(d\hat{g}|g) = \int_{\mathbf{G}} c(\hat{g}, g) \text{Tr}[P(d\hat{g})\rho_g].$$

Ideally, we would like to minimize the cost simultaneously for every  $g$ . However, in general this is not possible because the number of states in the orbit  $\{\rho_g\}_{g \in \mathbf{G}}$  is larger than the dimension of the Hilbert space and therefore there is no way in which the states  $\{\rho_g\}_{g \in \mathbf{G}}$  can be perfectly distinguishable.

Since the true value  $g \in \mathbf{G}$  is unknown, one approach to optimization is to minimize the *worst-case cost* of the strategy, defined as

$$c_{wc}(\rho, P) = \max_{g \in \mathbf{G}} c(\rho, P|g). \quad (1)$$

An alternative (and less conservative) approach consists in assuming a prior distribution  $\pi(dg)$  for the true value  $g \in \mathbf{G}$  and in minimizing the average cost. Given that  $g$  is completely unknown, a natural choice of prior is the Haar measure  $dg$ , normalized as  $\int_{\mathbf{G}} dg = 1$ . With this choice, the average cost is given by

$$c_{ave}(\rho, P) := \int_{\mathbf{G}} dg c(\rho, P|g). \quad (2)$$

For finite groups we can understand the integral as a finite sum and  $dg$  as synonymous of  $1/|\mathbf{G}|$ , where  $|\mathbf{G}|$  is the cardinality of the group  $\mathbf{G}$ .

In general, the worst-case and the average approach are very different: for example, the functional  $c_{ave}(\rho, P)$  is linear in both arguments while the functional  $c_{wc}(\rho, P)$  is not. The two approaches may lead to very different optimal strategies: for example, since  $c_{ave}(\rho, P)$  is linear in  $\rho$  the optimal input state can be searched without loss of generality in the set of pure states  $\rho = |\varphi\rangle\langle\varphi|$ , while this reduction is generally not possible for the minimization of  $c_{wc}(\rho, P)$ . However, when the cost function enjoys suitable group symmetries it is possible to show that there is a strategy  $(\rho, P)$  that is optimal for both the worst-case and the average approach. This point will be illustrated in Section 4.

### 3. Basic notions and notations

#### 3.1. Bipartite states

Let  $\{|\phi_i\rangle\}_{i=1}^{d_{\mathcal{H}}}$  (resp.  $\{|\psi_j\rangle\}_{j=1}^{d_{\mathcal{K}}}$ ) be a fixed orthonormal bases for the Hilbert space  $\mathcal{H}$  (resp.  $\mathcal{K}$ ). The choice of orthonormal bases induces a bijective correspondence between linear operators from  $\mathcal{K}$  to  $\mathcal{H}$  and bipartite vectors in  $\mathcal{H} \otimes \mathcal{K}$ . Following the “double-ket notation” of Ref. [8], if  $A \in \text{Lin}(\mathcal{K}, \mathcal{H})$  is a linear operator from  $\mathcal{K}$  to  $\mathcal{H}$  we define the bipartite vector  $|A\rangle\rangle \in \mathcal{H} \otimes \mathcal{K}$  as

$$|A\rangle\rangle := \sum_{i=1}^{d_{\mathcal{H}}} \sum_{j=1}^{d_{\mathcal{K}}} \langle\phi_i|A|\psi_j\rangle |\phi_i\rangle|\psi_j\rangle. \quad (3)$$

Two useful properties of this correspondence are given by the equations

$$\langle\langle A|B\rangle\rangle = \text{Tr}[A^\dagger B] \quad \forall A, B \in \text{Lin}(\mathcal{K}, \mathcal{H}) \quad (4)$$

$$\begin{aligned} |A\rangle\rangle &= (A \otimes I_{\mathcal{K}})|I_{\mathcal{K}}\rangle\rangle \\ &= (I_{\mathcal{H}} \otimes A^T)|I_{\mathcal{H}}\rangle\rangle \quad \forall A \in \text{Lin}(\mathcal{K}, \mathcal{H}), \end{aligned} \quad (5)$$

where  $A^T$  denotes the transpose of  $A$  with respect to the fixed bases. Using the singular value decomposition of  $A$ , it is easy to see that the state  $|A\rangle\rangle$  is of the product form  $|A\rangle\rangle = |\phi\rangle|\psi\rangle$  if and only if  $A$  is rank-one. Moreover, if the dimension of  $\mathcal{K}$  is larger than  $d_{\mathcal{H}}$ , then for every state  $|A\rangle\rangle$  there exists a  $d_{\mathcal{H}}$ -dimensional subspace  $\mathcal{K}_A$  such that  $|A\rangle\rangle \in \mathcal{H} \otimes \mathcal{K}_A$ . This is clear from Eq. (5): the dimension of the image of  $A^T$  cannot exceed the dimension of its domain  $\mathcal{H}$ .

#### 3.2. Group representations

The unknown transformation in our estimation problem belongs to a group representation  $\{\mathcal{U}_g\}_{g \in \mathbf{G}}$ , where each map  $\mathcal{U}_g$  is a completely positive trace preserving map sending operators in  $\text{Lin}(\mathcal{H})$  to operators in  $\text{Lin}(\mathcal{H})$ . Now, the condition  $\mathcal{U}_{g^{-1}}\mathcal{U}_g = \mathcal{I}_{\mathcal{H}}$  implies that the map  $\mathcal{U}_g$  must have the form  $\mathcal{U}_g(\rho) = U_g \rho U_g^\dagger$ , where  $U_g \in \text{Lin}(\mathcal{H})$  is a unitary matrix. In particular, for  $g = e$  we can choose  $U_e = I_{\mathcal{H}}$ . Moreover, the condition  $\mathcal{U}_g\mathcal{U}_h = \mathcal{U}_{gh}, \forall g, h \in \mathbf{G}$  implies that the unitaries  $\{U_g\}_{g \in \mathbf{G}}$  define a projective unitary representation, that is, a function  $U : g \mapsto U_g$  sending elements of the group to unitary operators and satisfying the relation

$U_g U_h = \omega(g, h) U_{gh}, \forall g, h \in \mathbf{G}$ , where  $\omega(g, h) \in \mathbb{C}$  is a multiplier, satisfying the properties

$$\begin{aligned} |\omega(g, h)| &= 1 & \forall g, h \in \mathbf{G} \\ \omega(g, h)\omega(gh, k) &= \omega(g, hk)\omega(h, k) & \forall g, h, k \in \mathbf{G} \\ \omega(g, e) = \omega(e, g) &= 1 & \forall g \in \mathbf{G}. \end{aligned}$$

Since the group  $\mathbf{G}$  is compact and the space  $\mathcal{H}$  is finite dimensional, with a suitable choice of basis the Hilbert space can be decomposed as

$$\mathcal{H} = \bigoplus_{\mu \in \text{lrr}(U)} \mathcal{H}_\mu \otimes \mathbb{C}^{m_\mu}, \quad (6)$$

where the sum runs over the set  $\text{lrr}(U)$  of all irreducible representations contained in the isotypic decomposition of  $U$ ,  $\mathcal{H}_\mu$  is a *representation space* of dimension  $d_\mu$ , carrying the irreducible representation  $U^\mu$ , and  $\mathbb{C}^{m_\mu}$  is a *multiplicity space*, where  $m_\mu$  is the multiplicity of the irreducible representation  $U^\mu$  in the decomposition of  $U$ . Accordingly, the representation  $U$  can be written in the block diagonal form

$$U = \bigoplus_{\mu \in \text{lrr}(U)} U^\mu \otimes I_{m_\mu}. \quad (7)$$

Note that all irreducible representations  $U^\mu$  must be projective unitary representations and must have the same multiplier  $\omega$ .

In the optimization of the estimation strategy  $(\rho, P)$  we can take the multiplicities  $m_\mu$  as large as we want. Indeed, introducing an ancillary system  $\mathcal{K}$  means replacing the Hilbert space  $\mathcal{H}$  with the tensor product  $\mathcal{H}' := \mathcal{H} \otimes \mathcal{K}$  and replacing the unitary  $U_g$  with the unitary  $U'_g := U_g \otimes I_{\mathcal{K}}$  for every  $g \in \mathbf{G}$ . In Eqs. (6) and (7) this means replacing the multiplicity space  $\mathbb{C}^{m_\mu}$  with  $\mathbb{C}^{m_\mu} \otimes \mathcal{K}$  and the multiplicity  $m_\mu$  with  $m'_\mu := m_\mu d_{\mathcal{K}}$ . In particular, we are free to choose the ancillary system  $\mathcal{K}$  so that the condition  $m_\mu \geq d_\mu$  is met for every  $\mu \in \text{lrr}(U)$ .

### 3.3. Bipartite states and group representations

Here we combine the facts observed in the paragraphs 3.1 and 3.2. Let us consider a vector  $|\Psi\rangle \in \mathcal{H}$  in a Hilbert space where the representation  $U$  acts. Using the isotypic decomposition of Eq. (6) and the correspondence of Eq. (3) for each tensor product  $\mathcal{H}_\mu \otimes \mathbb{C}^{m_\mu}$ , we can write

$$|\Psi\rangle = \bigoplus_{\mu \in \text{lrr}(U)} |\Psi_\mu\rangle \in \bigoplus_{\mu \in \text{lrr}(U)} \mathcal{H}_\mu \otimes \mathbb{C}^{m_\mu},$$

where  $\Psi_\mu$  is a linear operator in  $\text{Lin}(\mathbb{C}^{m_\mu}, \mathcal{H}_\mu)$ . In other words, the vector  $|\Psi\rangle$  can be written as a linear superposition of bipartite vectors. As we will see in the next sections, the fact that we can have entanglement between each representation space  $\mathcal{H}_\mu$  and the corresponding multiplicity space  $\mathbb{C}^{m_\mu}$  is the key ingredient to the construction of the optimal estimation strategies.

Suppose that  $U$  has been chosen so that  $m_\mu \geq d_\mu$  for every  $\mu \in \text{lrr}(U)$ . From the last observation of paragraph 3.1 we know that each vector  $|\Psi_\mu\rangle$  will be contained in a subspace  $\mathcal{H}_\mu \otimes \mathcal{K}_\mu$ , where  $\mathcal{K}_\mu \simeq \mathbb{C}^{d_\mu} \simeq \mathcal{H}_\mu$  is a suitable subspace of  $\mathbb{C}^{m_\mu}$ . Therefore, every vector  $|\Psi\rangle$  belongs to a suitable subspace  $\tilde{\mathcal{H}} \simeq \bigoplus_{\mu \in \text{lrr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$ . Note that the whole orbit  $\{|\Psi_g\rangle := U_g |\Psi\rangle\}_{g \in \mathbf{G}}$  belongs to the subspace  $\tilde{\mathcal{H}}$ . In conclusion, as long as we are interested in *pure* input states we can effectively replace  $\mathcal{H}$  with the space  $\tilde{\mathcal{H}} := \bigoplus_{\mu \in \text{lrr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$  where  $m_\mu = d_\mu$  for every  $\mu \in \text{lrr}(U)$  and we can replace the representation  $U$  with the representation  $\tilde{U}$  defined by  $\tilde{U}_g := \bigoplus_{\mu \in \text{lrr}(U)} U_g^\mu \otimes I_\mu$ , where  $I_\mu$  denotes the identity on  $\mathcal{H}_\mu$ .

### 3.4. The class states

An important family of states in  $\tilde{\mathcal{H}} = \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$  is the family of states of the form

$$|\Phi\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \frac{c_\mu}{\sqrt{d_\mu}} |I_\mu\rangle, \quad c_\mu \in \mathbb{C}, \quad \sum_{\mu \in \text{Irr}(U)} |c_\mu|^2 = 1 \quad (8)$$

These states are linear superpositions of maximally entangled states in the sectors  $\mathcal{H}_\mu \otimes \mathcal{H}_\mu$ . In the following I will call these states *class states*. The reason for the name comes from the fact that these states correspond to *class functions* (see definition below) through the Fourier-Plancherel theory (see e.g. [9]). The remaining part of this paragraph is aimed at making this correspondence clear.

Let us start from Fourier-Plancherel theory. Denote by  $\text{Irr}(\mathbf{G}, \omega)$  the set of all irreducible representations of  $\mathbf{G}$  with multiplier  $\omega$  and denote by  $u_{ij}^\mu(g) := \langle \mu, i | U_g^\mu | \mu, j \rangle$  the matrix element of the operator  $U_g^\mu$  with respect to a fixed orthonormal basis  $\{|\mu, i\rangle\}_{i=1}^{d_\mu}$  for the representation space  $\mathcal{H}_\mu$ . Recall that the functions  $\{\tilde{u}_{ij}^{\mu*}(g) := \sqrt{d_\mu} u_{ij}^{\mu*}(g)\}_{\mu \in \text{Irr}(\mathbf{G}, \omega); i, j=1, \dots, d_\mu}$  are an orthonormal basis for the Hilbert space  $L_2(\mathbf{G}, dg)$  endowed with the scalar product  $\langle f_1, f_2 \rangle := \int_{\mathbf{G}} dg f_1^*(g) f_2(g)$ . Using this fact, we can expand every function  $f \in L_2(\mathbf{G}, dg)$  as

$$|f\rangle = \bigoplus_{\mu \in \text{Irr}(\mathbf{G}, \omega)} \sum_{i, j=1}^{d_\mu} f_{ij}^\mu |\tilde{u}_{ij}^{\mu*}\rangle. \quad (9)$$

On the Hilbert space  $L_2(\mathbf{G}, dg)$  we consider the action of  $\mathbf{G}$  given by the left-regular representation  $T$  with multiplier  $\omega$ , defined by  $(T_g f)(h) := \omega(g, g^{-1}h) f(g^{-1}h)$ ,  $\forall f \in L_2(\mathbf{G}, dg), \forall g, h \in \mathbf{G}$ . [10]. In particular, the transformation of the functions  $\tilde{u}_{ij}^{\mu*}(g)$  is given by

$$T_g |\tilde{u}_{ij}^{\mu*}\rangle = \sum_{k=1}^{d_\mu} u_{ki}^\mu(g) |\tilde{u}_{kj}^{\mu*}\rangle, \quad (10)$$

that is, for every fixed value of  $j$  the vectors  $\{|\tilde{u}_{ij}^{\mu*}\rangle\}_{i=1}^{d_\mu}$  span an invariant subspace carrying the irreducible representation  $U^\mu$ . It is then easy to see that the unitary operator  $V : L_2(\mathbf{G}, dg) \rightarrow \bigoplus_{\mu \in \text{Irr}(\mathbf{G}, \omega)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$  defined by

$$V |u_{ij}^{\mu*}\rangle := |\mu, i\rangle |\mu, j\rangle \quad \forall \mu \in \text{Irr}(\mathbf{G}, \omega), \forall i, j = 1, \dots, d_\mu \quad (11)$$

intertwines the two representations  $T_g$  and  $W_g = \bigoplus_{\mu \in \text{Irr}(\mathbf{G}, \omega)} U_g^\mu \otimes I_\mu$ : indeed, using Eqs. (10) and (11) one has  $VT_g |u_{ij}^{\mu*}\rangle = U_g^\mu |\mu, i\rangle |\mu, j\rangle = (U_g^\mu \otimes I_\mu) V |u_{ij}^{\mu*}\rangle$  for every  $\mu \in \text{Irr}(\mathbf{G}, \omega), \forall i, j = 1, \dots, d_\mu$ , and therefore  $VT_g = W_g V$ . Applying the unitary  $V$  to the expansion of Eq. (9) we then obtain

$$V |f\rangle = \bigoplus_{\mu \in \text{Irr}(\mathbf{G}, \omega)} \sum_{i, j=1}^{d_\mu} f_{ij}^\mu |\mu, i\rangle |\mu, j\rangle.$$

Let us come now to class functions. A class function is a function that is constant on conjugacy classes, that is,  $f(hgh^{-1}) = f(g)$  for every  $g, h \in \mathbf{G}$ . Any class function can be written as a linear combination of irreducible characters, namely  $f(g) = \sum_{\mu \in \text{Irr}(\mathbf{G})} f_\mu \chi_\mu^*(g)$ , where  $\chi_\mu^*(g) = \text{Tr}[U_g^{\mu*}]$ . Exploiting the correspondence given by the unitary  $V$  we obtain

$$V |f\rangle = \bigoplus_{\mu \in \text{Irr}(\mathbf{G}, \omega)} \frac{f_\mu}{\sqrt{d_\mu}} |I_\mu\rangle.$$

It is now clear that the class states of Eq. (8) are nothing but the projection of the class functions to the finite dimensional subspace  $\tilde{\mathcal{H}} \subset \bigoplus_{\mu \in \text{lrr}(\mathbb{G}, \omega)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$ .

Besides providing a justification for the choice of the name *class states*, the Fourier-Plancherel theory also provides a hint of the fact that the class states are optimal for estimation. Indeed, if we take the scalar product of two states in the orbit  $\{|\Phi_g\rangle := \tilde{U}_g|\Phi\rangle\}$  we obtain  $\langle \Phi_g | \Phi_h \rangle = \sum_{\mu \in \text{lrr}(U)} \frac{|c_\mu|^2}{d_\mu} \chi_\mu(g^{-1}h)$ . If  $c_\mu$  is chosen so that  $c_\mu = \lambda d_\mu$  for some constant  $\lambda \in \mathbb{C}$  we then obtain  $\langle \Phi_g | \Phi_h \rangle \propto \sum_{\mu \in \text{lrr}(U)} d_\mu \chi_\mu(g^{-1}h)$ , which is nothing but the truncated version of the Dirac delta  $\delta(g, h) = \sum_{\mu \in \text{lrr}(\mathbb{G}, \omega)} d_\mu \chi_\mu(g^{-1}h)$ . This provides an heuristic argument for the optimality of the class states, although the choice of the coefficients  $\{c_\mu\}$  to provide the “best” finite dimensional approximation of the Dirac delta will depend on the choice of the cost function  $c(\hat{g}, g)$ .

### 3.5. Notations

If  $V \in \text{Lin}(\mathcal{H})$  is a unitary matrix, the symbol  $\mathcal{V}$  will always denote the completely positive map  $\mathcal{V}(\rho) = V\rho V^\dagger$ . Let  $\mathcal{C} : \text{Lin}(\mathcal{H}) \rightarrow \text{Lin}(\mathcal{K})$  be a completely positive map and let  $\mathcal{C}(\rho) = \sum_{i=1}^r C_i \rho C_i^\dagger$ ,  $C_i \in \text{Lin}(\mathcal{H}, \mathcal{K})$  be a Kraus form for  $\mathcal{C}$ . The notation  $\mathcal{C}^\dagger, \mathcal{C}^*, \mathcal{C}^T$  will be used for the following maps

$$\begin{aligned} \mathcal{C}^\dagger(\rho) &:= \sum_{i=1}^r C_i^\dagger \rho C_i \\ \mathcal{C}^*(\rho) &:= \sum_{i=1}^r C_i^* \rho C_i^T \\ \mathcal{C}^T(\rho) &:= \sum_{i=1}^r C_i^T \rho C_i^* = (\mathcal{C}^*)^\dagger(\rho), \end{aligned}$$

where the complex conjugation  $*$  and the transpose  $T$  are defined with respect to the fixed bases for  $\mathcal{H}$  and  $\mathcal{K}$ . Note that the map  $\mathcal{C}^\dagger$  is the adjoint of  $\mathcal{C}$  with respect to the Hilbert-Schmidt scalar product  $\langle\langle A|B \rangle\rangle = \text{Tr}[A^\dagger B]$ ,  $A, B \in \text{Lin}(\mathcal{H})$ : indeed, we have  $\langle\langle A|\mathcal{C}(B) \rangle\rangle = \langle\langle \mathcal{C}^\dagger(A)|B \rangle\rangle$ . Note also that the definition of the maps  $\mathcal{C}^\dagger, \mathcal{C}^*, \mathcal{C}^T$  does not depend on the choice of a particular Kraus form.

Consider the Hilbert space  $\tilde{\mathcal{H}} = \bigoplus_{\mu \in \text{lrr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$ , the representation  $\tilde{U} = \bigoplus_{\mu \in \text{lrr}(U)} U^\mu \otimes I_\mu$ , and let

$$\begin{aligned} \tilde{U}' &= \left\{ \bigoplus_{\mu \in \text{lrr}(U)} I_\mu \otimes B_\mu \mid B_\mu \in \text{Lin}(\mathcal{H}_\mu) \right\} \\ \tilde{U}'' &= \left\{ \bigoplus_{\mu \in \text{lrr}(U)} A_\mu \otimes I_\mu \mid A_\mu \in \text{Lin}(\mathcal{H}_\mu) \right\} \end{aligned}$$

be the commutant and the bicommutant of  $\tilde{U}$ , respectively. Every class state in Eq. (8) then defines a modular conjugation: for every operator  $A = \bigoplus_{\mu \in \text{lrr}(U)} A_\mu \otimes I_\mu \in \tilde{U}''$  I will denote by  $A^R \in \tilde{U}'$  its modular conjugate, given by

$$A^R := \bigoplus_{\mu \in \text{lrr}(U)} I_\mu \otimes A_\mu^*. \quad (12)$$

Every class state  $|\Phi\rangle \in \tilde{\mathcal{H}}$  then enjoys the symmetry

$$\tilde{U}_g \tilde{U}_g^R |\Phi\rangle = |\Phi\rangle \quad \forall g \in \mathbf{G}, \quad (13)$$

as it can be easily verified using Eq. (5) for every  $\mu \in \text{Irr}(U)$ . The modular conjugate representation will be denoted by  $\tilde{U}^R$ . For a completely positive map  $\mathcal{C}(\rho) = \sum_{i=1}^r C_i \rho C_i^\dagger$  with  $C_i \in \tilde{U}''$  for every  $i = 1, \dots, r$  we will denote by  $\mathcal{C}^R$  the completely positive map  $\mathcal{C}^R(\rho) := \sum_{i=1}^r C_i^R \rho C_i^{R\dagger}$ . Note that in general we have

$$\mathcal{C}(|\Phi\rangle\langle\Phi|) = \mathcal{C}^{R\dagger}(|\Phi\rangle\langle\Phi|). \quad (14)$$

Introducing the *swap map*  $\mathcal{S}$ , defined by  $\mathcal{S}\left(\bigoplus_{\mu \in \text{Irr}(U)} A_\mu \otimes B_\mu\right) := \bigoplus_{\mu \in \text{Irr}(U)} B_\mu \otimes A_\mu$ , we obtain the relations

$$A = \mathcal{S}(A^{R*}) \quad \forall A \in \tilde{U}'' \quad (15)$$

$$\mathcal{C} = \mathcal{S}\mathcal{C}^{R*}\mathcal{S} \quad \forall \mathcal{C} : \mathcal{C}(\rho) = \sum_{i=1}^r C_i \rho C_i^\dagger, \quad C_i \in \tilde{U}'' \quad \forall i = 1, \dots, r. \quad (16)$$

## 4. Optimal estimation

### 4.1. Left-invariant cost functions: optimality of covariant measurements

In this paragraph I will shortly review two classic results about the structure of covariant measurements and about their optimality. The proofs of these results can be found in the monographs [11, 12, 13].

Let us start from the definition:

**Definition 1** A POVM  $P : \sigma(\mathbf{G}) \rightarrow \text{Lin}_+(\mathcal{H})$  is covariant with respect to the representation  $\{U_g\}_{g \in \mathbf{G}}$  if  $P_{gB} = \mathcal{U}_g(P_B) \quad \forall g \in \mathbf{G}, \forall B \in \sigma(\mathbf{G})$ .

Covariant POVMs have a very simple structure:

**Theorem 1 (Structure of covariant POVMs)** Let  $\mathbf{G}$  be a compact group,  $U$  be a projective unitary representation of  $\mathbf{G}$  on the Hilbert space  $\mathcal{H}$ . A POVM  $P : \sigma(\mathbf{G}) \rightarrow \text{Lin}_+(\mathcal{H})$  is covariant with respect to  $U$  if and only if it has the form

$$P(d\hat{g}) = \mathcal{U}_{\hat{g}}(\xi) d\hat{g} \quad (17)$$

where  $d\hat{g}$  is the normalized Haar measure and  $\xi \in \text{Lin}_+(\mathcal{H})$  is a suitable operator, called the seed of the covariant POVM.

The special form of covariant POVMs provides a great simplification in optimization problems, because it reduces the optimization of the whole POVM to the optimization of a single non-negative operator  $\xi \in \text{Lin}_+(\mathcal{H})$ . Luckily, this simplification can be done without loss of generality in most situations. Precisely, covariant measurement are optimal whenever the cost function  $c(\hat{g}, g)$  is *left-invariant*, that is  $c(h\hat{g}, hg) = c(\hat{g}, g) \forall g, h \in \mathbf{G}$ .

**Theorem 2 (Optimality of covariant POVMs)** Let  $\mathbf{G}$  be a compact group,  $\{U_g\}_{g \in \mathbf{G}}$  be a unitary projective representation of  $\mathbf{G}$  on the Hilbert space  $\mathcal{H}$  and  $c(\hat{g}, g)$  be a left-invariant cost function. Then, for every input state  $\rho \in \text{St}(\mathcal{H})$  the optimal POVM  $P$  for both the worst-case and uniform average approach can be assumed without loss of generality to be covariant. With this choice one has  $c_{ave}(\rho, P) = c_{wc}(\rho, P) = c(\rho, P|g) \forall g \in \mathbf{G}$ .



#### 4.2. Right-invariant cost functions: optimality of class states

In this paragraph I illustrate with a new proof a recent result on the optimality of class states [14]. This result is dual to the classic result on the optimality of covariant POVMs showed in the previous paragraph.

A central feature of quantum theory is the validity of the *purification principle* [15]: for every mixed state  $\rho \in \text{St}(\mathcal{H})$  there exists a Hilbert space  $\mathcal{K}$  and a pure state  $|\Psi\rangle\rangle \in \mathcal{H} \otimes \mathcal{K}$  such that  $\rho = \text{Tr}_{\mathcal{K}}[|\Psi\rangle\rangle\langle\langle\Psi|]$ . The pure state  $\Psi$  is referred to as a *purification* of  $\rho$ . From Eq. (5) we have  $\rho = \Psi\Psi^\dagger$  and  $\text{Supp}(\rho) = \text{Rng}(\Psi)$ , where  $\text{Supp}$  and  $\text{Rng}$  denote the support and the range, respectively. The simplest example of purification of a state  $\rho \in \text{St}(\mathcal{H})$  is the *square-root purification*  $|\rho^{\frac{1}{2}}\rangle\rangle \in \mathcal{H} \otimes \mathcal{H}'$ , where  $\mathcal{H}' \simeq \mathcal{H}$ .

A large number of quantum features, such as teleportation and no cloning, are simple consequences of the purification principle [15]. The structure of the optimal states for the estimation of group parameters is no exception to that. In the following I will give a simple proof of the optimality of the states in Eq. (8) based on the purification principle. To this purpose I will use a remarkable consequence of purification, namely the fact, first noticed by Schrödinger [16], that every ensemble decomposition of a state can be induced from the purification via a quantum measurement on the purifying system:

**Lemma 1** *Let  $(X, \sigma(X))$  be a measurable space and let  $\rho(dx) : \sigma(X) \rightarrow \text{Lin}_+(\mathcal{H}), B \mapsto \rho_B$  be an ensemble, that is, an operator-valued measure such that  $\rho_X \in \text{St}(\mathcal{H})$ . If  $|\Psi\rangle\rangle \in \mathcal{H} \otimes \mathcal{K}$  is a purification of  $\rho_X$  then there exists a POVM  $P : \sigma(X) \rightarrow \text{Lin}_+(\mathcal{K}), B \in \sigma(X) \mapsto P_B$  such that*

$$\rho_B = \text{Tr}_{\mathcal{K}}[(I_{\mathcal{H}} \otimes P_B)|\Psi\rangle\rangle\langle\langle\Psi|] \quad \forall B \in \sigma(X). \quad (18)$$

**Proof.** Let  $\Psi^{-1}$  be the inverse of  $\Psi$  on its support and define the POVM  $P$  via  $P_B := [\Psi^{-1}\rho_B(\Psi^\dagger)^{-1}]^T$ . The POVM  $P$  provides a resolution of the projector on the range of  $\Psi^T$  and can be easily extended to a resolution of the identity on  $\mathcal{K}$ . Using the identity  $\Psi\Psi^{-1} = P_{\text{Supp}(\rho_X)}$  where  $\Pi_{\text{Supp}(\rho_X)}$  is the projector on the support of  $\rho_X$  we have

$$\begin{aligned} \text{Tr}_{\mathcal{K}}[(I_{\mathcal{H}} \otimes P_B)|\Psi\rangle\rangle\langle\langle\Psi|] &= \text{Tr}_{\mathcal{K}}[(I_{\mathcal{H}} \otimes (\Psi^{-1}\rho_B)^T)|\Psi\rangle\rangle\langle\langle\Psi|(I_{\mathcal{H}} \otimes (\Psi^*)^{-1})] \\ &= \text{Tr}_{\mathcal{H}'}[(|\Psi\Psi^{-1}\rho_B\rangle\rangle\langle\langle\Psi\Psi^{-1}|] \\ &= \text{Tr}_{\mathcal{H}'}[(\Pi_{\text{Supp}(\rho_X)}\rho_B)\langle\langle\Pi_{\text{Supp}(\rho_X)}] \\ &= \rho_B. \end{aligned}$$

■

Suppose that the cost function  $c(\hat{g}, g)$  is *right-invariant*, that is  $c(\hat{g}h, gh) = c(\hat{g}, g), \forall \hat{g}, g, h \in \mathbf{G}$ . Then we have the following

**Theorem 3 (Optimality of the purification of invariant states)** *Let  $\mathbf{G}$  be a compact group,  $U$  be a projective unitary representation of  $\mathbf{G}$  on the Hilbert space  $\mathcal{H}$  and  $c(\hat{g}, g)$  be a right-invariant cost function. Then, the optimal input state for both the worst-case and uniform average approach can be assumed without loss of generality to be the purification of an invariant state  $\rho \in U'$ . Denoting by  $|\Psi\rangle\rangle \in \mathcal{H} \otimes \mathcal{L}$  such a purification, there exists an optimal POVM  $P$  on  $\mathcal{H} \otimes \mathcal{L}$  such that  $c_{\text{ave}}(\Psi, P) = c_{\text{wc}}(\Psi, P) = c(\Psi, P|g), \forall g \in \mathbf{G}$ .*

**Proof.** Suppose that  $\sigma \in \text{St}(\mathcal{H} \otimes \mathcal{K})$  is an optimal input state and that  $Q(d\hat{g})$  is an optimal POVM for the average approach (resp. for the worst case approach). We now prove that there is a Hilbert space  $\mathcal{L}$  and another state  $|\Psi\rangle\rangle \in \mathcal{H} \otimes \mathcal{L}$  that is the purification of an invariant state  $\rho \in U'$  and is optimal as well. Consider the orbit  $\{\sigma_h := (U_h \otimes I_{\mathcal{K}})(\sigma)\}_{h \in \mathbf{G}}$  and the operator-valued measure  $\sigma(dh) = \sigma_h dh$ . Clearly the state  $\sigma_{\mathbf{G}} = \int_{\mathbf{G}} dh \sigma_h$  is invariant under  $U \otimes I_{\mathcal{K}}$ . Let  $|\Psi\rangle\rangle \in \mathcal{H} \otimes \mathcal{K} \otimes \mathcal{H}' \otimes \mathcal{K}'$ , with  $\mathcal{H}' \simeq \mathcal{H}, \mathcal{K}' \simeq \mathcal{K}$ , be the square-root purification

of  $\sigma_{\mathbb{G}}$  and define  $\mathcal{L} := \mathcal{K} \otimes \mathcal{H}' \otimes \mathcal{K}'$ . Note that  $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{L}$  is also a purification of the invariant state  $\rho := \text{Tr}_{\mathcal{K}}[\sigma_{\mathbb{G}}] \in U'$ . We now have to show that  $|\Psi\rangle$  is optimal for the uniform average (resp. worst-case) approach. By lemma 1, we know that there is a POVM  $M_h dh$  on  $\mathcal{H}' \otimes \mathcal{K}'$  such that  $\rho_h = \text{Tr}_{\mathcal{H}' \otimes \mathcal{K}'}[(I_{\mathcal{H}} \otimes I_{\mathcal{K}} \otimes M_h)|\Psi\rangle\langle\Psi|] \quad \forall h \in \mathbb{G}$ . Define now the POVM  $P(d\hat{g}) := \int_{\mathbb{G}} dh Q(d(\hat{g}h)) \otimes M_h$ . It is easy to see that the average cost of the strategy  $(\Psi, P)$  is equal to the average cost of the strategy  $(\sigma, Q)$ :

$$\begin{aligned}
c_{ave}(\Psi, P) &= \int_{\mathbb{G}} dg \int_{\mathbb{G}} c(g, \hat{g}) \langle\langle \Psi_g | P(d\hat{g}) | \Psi_g \rangle\rangle \\
&= \int_{\mathbb{G}} dg \int_{\mathbb{G}} dh \int_{\mathbb{G}} c(g, \hat{g}) \langle\langle \Psi | (U_g^\dagger Q(d(\hat{g}h)) U_g \otimes M_h) | \Psi \rangle\rangle \\
&= \int_{\mathbb{G}} dh \int_{\mathbb{G}} dg \int_{\mathbb{G}} c(g, \hat{g}) \text{Tr}[U_g^\dagger Q(d(\hat{g}h)) U_g \sigma_h] \\
&= \int_{\mathbb{G}} dh \int_{\mathbb{G}} d(gh) \int_{\mathbb{G}} c(gh, \hat{g}h) \text{Tr}[Q(d(\hat{g}h)) \sigma_{gh}] \\
&= c_{ave}(\sigma, Q),
\end{aligned}$$

having used the right-invariance of the Haar measure and of the cost function. On the other hand, the worst-case cost of the strategy  $(\Psi, P)$  cannot be larger than the worst-case cost of the strategy  $(\sigma, Q)$ :

$$\begin{aligned}
c_{wc}(\Psi, P) &= \max_{g \in \mathbb{G}} \int_{\mathbb{G}} c(\hat{g}, g) \langle \Psi_g | P(d\hat{g}) | \Psi_g \rangle \\
&= \max_{g \in \mathbb{G}} \int_{\mathbb{G}} dh \int_{\mathbb{G}} c(\hat{g}, g) \langle \Psi | (U_g^\dagger Q(d(\hat{g}h)) U_g \otimes M_h) | \Psi \rangle \\
&= \max_{g \in \mathbb{G}} \int_{\mathbb{G}} dh \int_{\mathbb{G}} c(g, \hat{g}) \text{Tr}[U_g^\dagger Q(d(\hat{g}h)) U_g \sigma_h] \\
&= \max_{g \in \mathbb{G}} \int_{\mathbb{G}} dh \int_{\mathbb{G}} c(\hat{g}h, gh) \text{Tr}[Q(d(\hat{g}h)) \sigma_{gh}] \\
&= \max_{g \in \mathbb{G}} \int_h dh c(\sigma, Q|gh) \\
&\leq \int_h dh \max_{g \in \mathbb{G}} c(\sigma, Q|gh) \\
&= c_{wc}(\sigma, Q).
\end{aligned}$$

Hence, we proved that, both for the average and for the worst case approach, whichever the optimal strategy  $(\sigma, Q)$  is, the strategy  $(\Psi, P)$  cannot be worse. Finally, we observe that for the strategy  $(\Psi, P)$  we have  $c(\Psi, P|g) = \int_{\mathbb{G}} dh \int_{\mathbb{G}} c(\hat{g}h, gh) \text{Tr}[Q(d(\hat{g}h)) \sigma_{gh}] = \int_h dh c(\sigma, Q|h)$ . This proves that  $c(\Psi, P|g)$  is independent of  $g \in \mathbb{G}$ , whence  $c(\Psi, P|g) = c_{ave}(\Psi, P) = c_{wc}(\Psi, P)$ . ■

**Theorem 4 (Optimality of the class states)** *Let  $\mathbb{G}$  be a compact group,  $U$  be a unitary projective representation of  $\mathbb{G}$  on the Hilbert space  $\mathcal{H}$  satisfying the property  $d_\mu \leq m_\mu \forall \mu \in \text{Irr}(U)$  and  $c(\hat{g}, g)$  be a right-invariant cost function. Then, the optimal input state for both the worst-case and uniform average approach can be assumed without loss of generality to be a class state of the form*

$$|\Phi\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{\frac{p_\mu}{d_\mu}} |I_\mu\rangle \in \tilde{\mathcal{H}} = \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu \subseteq \mathcal{H}, \quad (19)$$

where  $p_\mu \geq 0$  and  $\sum_\mu p_\mu = 1$ .

**Proof.** By theorem 3, we know that there exists an optimal state that is the purification of an invariant state  $\rho \in U'$ . Now the form of an invariant state is  $\rho = \bigoplus_{\mu \in \text{Irr}(U)} \frac{p_\mu}{d_\mu} I_\mu \otimes \rho_\mu$ , with  $\rho_\mu \in \text{St}(\mathbb{C}^{m_\mu})$ . The square-root purification of  $\rho$  is then given by

$$|\Psi\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{\frac{p_\mu}{d_\mu}} |I_\mu\rangle |\rho_\mu^{\frac{1}{2}}\rangle \in \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu \otimes \mathbb{C}^{m_\mu} \otimes \mathbb{C}^{m_\mu}.$$

Here the first copy of  $\mathcal{H}_\mu$  is the representation space, while the remaining spaces contribute to the multiplicity. The action of the representation  $\{U_g \otimes I_{\mathcal{H}}\}_{g \in \mathbf{G}}$  then generates the orbit  $|\Psi_g\rangle := \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{\frac{p_\mu}{d_\mu}} |U_g^\mu\rangle |\rho_\mu^{\frac{1}{2}}\rangle$ . This orbit can be obtained from the orbit  $|\Phi_g\rangle := \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{\frac{p_\mu}{d_\mu}} |U_g^\mu\rangle \in \tilde{\mathcal{H}}$  by applying the isometry  $V : \bigoplus_{\mu \in \text{Irr}(U_g)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu \rightarrow \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu \otimes \mathbb{C}^{m_\mu} \otimes \mathbb{C}^{m_\mu}$  defined by  $V(\bigoplus_{\mu \in \text{Irr}(U)} |v_\mu\rangle) = \bigoplus_{\mu \in \text{Irr}(U)} |v_\mu\rangle |\rho_\mu^{\frac{1}{2}}\rangle$ . This proves that the state  $|\Phi\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{\frac{p_\mu}{d_\mu}} |I_\mu\rangle$  is an optimal input state. ■

#### 4.3. Invariant cost functions and isotropic states: optimality of isotropic seeds

Suppose that the cost function  $c(\hat{g}, g)$  is *invariant*, that is, both left- and right-invariant. Equivalently, this means that  $c(\hat{g}, g) = f(\hat{g}^{-1}g)$  where  $f$  is a class function. From the discussion of the previous paragraphs we already know that *i)* the optimal input state will be class state and *ii)* the optimal POVM will be covariant. In fact, in this case the optimal POVM has an additional symmetry that reflects the symmetries of class states (see Eq. (13)). Furthermore, this additional property appears not only for class states but also for a more general family of input states that we call *isotropic*.

**Definition 2 (Isotropic operators)** An operator  $A \in \text{Lin}(\tilde{\mathcal{H}})$  with  $\tilde{\mathcal{H}} = \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$  is isotropic if

$$\mathcal{U}_g \mathcal{U}_g^R(A) = A \quad \forall g \in \mathbf{G}.$$

The set of isotropic states (i.e. non-negative isotropic operators with unit trace) is convex. Clearly, every class state is isotropic. As a consequence, every mixture of class states is an isotropic state.

**Theorem 5 (Optimality of isotropic seeds)** Let  $\mathbf{G}$  be a compact group,  $\tilde{U}$  be a projective unitary representation of  $\mathbf{G}$  on the Hilbert space  $\tilde{\mathcal{H}} = \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$ ,  $c(\hat{g}, g)$  be an invariant cost function and  $\rho \in \text{St}(\mathcal{H})$  be an isotropic state. Then, the seed  $\xi \in \text{Lin}_+(\tilde{\mathcal{H}})$  of the optimal covariant measurement for the input state  $\rho$  is isotropic.

**Proof.** Since the cost function is left-invariant, the optimal POVM can be assumed to be covariant (theorem 2) and there is no difference between the average and worst case approaches. Moreover, for any covariant POVM  $P(dg) = \tilde{\mathcal{U}}_g(\xi) dg$  we can define another covariant POVM  $P'(dg) = \tilde{\mathcal{U}}_g(\xi') dg$  with  $\xi$  where  $\xi'$  is isotropic. Indeed, it is enough to choose  $\xi' := \int_{\mathbf{G}} dh \tilde{\mathcal{U}}_h \tilde{\mathcal{U}}_h^R(\xi)$ . Note that  $P'(dg)$  is still a resolution of the identity. The uniform average cost of the strategy  $(\rho, P')$  is the same of the strategy  $(\rho, P)$ :

$$\begin{aligned} c_{ave}(\rho, P') &= \int_{\mathbf{G}} dg \int_{\mathbf{G}} d\hat{g} c(\hat{g}, g) \text{Tr}[\tilde{\mathcal{U}}_{\hat{g}}(\xi') \tilde{\mathcal{U}}_g(\rho)] \\ &= \int_{\mathbf{G}} dg \int_{\mathbf{G}} d\hat{g} \int_{\mathbf{G}} dh c(\hat{g}, g) \text{Tr}[\tilde{\mathcal{U}}_{\hat{g}h}(\xi) \tilde{\mathcal{U}}_g \tilde{\mathcal{U}}_h^{\text{R}\dagger}(\rho)] \\ &= \int_{\mathbf{G}} dh \int_{\mathbf{G}} dg \int_{\mathbf{G}} d\hat{g} c(\hat{g}h, gh) \text{Tr}[\tilde{\mathcal{U}}_{\hat{g}h}(\xi) \tilde{\mathcal{U}}_{gh}(\rho)] \\ &= c_{ave}(\rho, P), \end{aligned}$$

having used the fact that  $\rho$  is isotropic (and hence  $\tilde{\mathcal{U}}_h^{R\dagger}(\rho) = \tilde{\mathcal{U}}_h(\rho)$ ) and the right-invariance of the cost function. ■

An example of isotropic seed is given by the *class seed*  $\xi = |\eta\rangle\langle\eta|$ , where  $|\eta\rangle \in \mathcal{H}$  is the class vector

$$|\eta\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{d_\mu} |I\mu\rangle. \quad (20)$$

We will now see that this seed plays a special role: any isotropic seed can be obtained from it by applying a suitable physical transformation.

**Theorem 6**  $\xi \in \text{Lin}_+(\tilde{\mathcal{H}})$  is an isotropic seed if and only if  $\xi = \mathcal{C}(|\eta\rangle\langle\eta|)$  for some completely positive map  $\mathcal{C} : \text{Lin}(\tilde{\mathcal{H}}) \rightarrow \text{Lin}(\tilde{\mathcal{H}})$  that is bistochastic (i.e. trace-preserving and unit-preserving) and leaves all elements in the algebra  $U'$  invariant.

**Proof.** Let us write  $\xi = \sum_{i=1}^r |\eta_i\rangle\langle\eta_i|$  and  $|\eta_i\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{d_\mu} |\eta_{i\mu}\rangle$ . Defining the operators  $C_i := \bigoplus_{\mu \in \text{Irr}(U)} \eta_{i\mu} \otimes I_\mu \in U''$  and defining the completely positive map  $\mathcal{C}$  as  $\mathcal{C}(\rho) := \sum_{i=1}^r C_i \rho C_i^\dagger$  we then obtain  $\xi = \mathcal{C}(|\eta\rangle\langle\eta|)$ . Clearly, one has  $\mathcal{C}(A) = A$  for every  $A \in U'$ . The requirement that  $\mathcal{C}$  be bistochastic is equivalent to the requirement that  $\xi$  be an isotropic seed. First, we can prove that  $\mathcal{C}$  is unital:

$$\begin{aligned} \mathcal{C}(I) &= \int_{\mathbb{G}} dg \mathcal{C}\tilde{\mathcal{U}}_g(|\eta\rangle\langle\eta|) = \int_{\mathbb{G}} dg \mathcal{C}\tilde{\mathcal{U}}_g^{R\dagger}(|\eta\rangle\langle\eta|) \\ &= \int_{\mathbb{G}} dg \tilde{\mathcal{U}}_g^{R\dagger} \mathcal{C}(|\eta\rangle\langle\eta|) = \int_{\mathbb{G}} dg \tilde{\mathcal{U}}_g^{R\dagger}(\xi) = \int_{\mathbb{G}} dg \tilde{\mathcal{U}}_g(\xi) = I \end{aligned}$$

To prove that  $\mathcal{C}$  is trace-preserving, we equivalently prove that  $\mathcal{C}^\dagger$  is unital. Using Eq. (14) it is immediate to obtain  $\mathcal{C}^{R\dagger}(I) = I$ : indeed, we have

$$\begin{aligned} \mathcal{C}^{R\dagger}(I) &= \int_{\mathbb{G}} dg \mathcal{C}^{R\dagger} \tilde{\mathcal{U}}_g(|\eta\rangle\langle\eta|) = \int_{\mathbb{G}} dg \tilde{\mathcal{U}}_g \mathcal{C}^{R\dagger}(|\eta\rangle\langle\eta|) \\ &= \int_{\mathbb{G}} dg \tilde{\mathcal{U}}_g \mathcal{C}(|\eta\rangle\langle\eta|) = \int_{\mathbb{G}} dg \tilde{\mathcal{U}}_g(\xi) = I \end{aligned}$$

Finally, it is enough to observe that  $\mathcal{C}^\dagger = \mathcal{S}(\mathcal{C}^{R\dagger})^* \mathcal{S}$  (see Eq. (15)) and, therefore,  $\mathcal{C}^\dagger(I) = I$ . ■

**Corollary 1** Let  $\xi \in \text{Lin}(\mathcal{H})$  be an isotropic seed, written as  $\xi = \sum_{i=1}^r |\eta_i\rangle\langle\eta_i|$ ,  $|\eta_i\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{d_\mu} |\eta_{i\mu}\rangle$ . Then we have  $\sum_{i=1}^r \eta_{i\mu} \eta_{i\mu}^\dagger = \sum_{i=1}^r \eta_{i\mu}^T \eta_{i\mu}^* = I_\mu$ .

In the following the covariant POVM  $E(d\hat{g}) = \tilde{\mathcal{U}}_{\hat{g}}(|\eta\rangle\langle\eta|)d\hat{g}$  generated by the class vector  $|\eta\rangle$  in Eq. (20) will be called *class POVM*.

**Corollary 2** Any covariant POVM  $P(d\hat{g}) = \tilde{\mathcal{U}}_{\hat{g}}(\xi)$  with isotropic seed can be obtained from the class POVM  $E(d\hat{g}) = \tilde{\mathcal{U}}_{\hat{g}}(|\eta\rangle\langle\eta|)d\hat{g}$  by application of a bistochastic map  $\mathcal{C}^{R\dagger}$ .

Note that the bistochastic map  $\mathcal{C}^{R\dagger}$  is *covariant*, that is  $\tilde{\mathcal{U}}_g \mathcal{C}^{R\dagger} = \mathcal{C}^{R\dagger} \tilde{\mathcal{U}}_g \forall g \in \mathbb{G}$ .

**Corollary 3** Let  $\rho \in \text{St}(\tilde{\mathcal{H}})$  be an isotropic state and  $P(d\hat{g}) = \tilde{\mathcal{U}}_{\hat{g}}(\xi)$  be a covariant POVM with isotropic seed. Then, the estimation strategy  $(\rho, P)$  has the same cost of the estimation strategy  $(\mathcal{C}^{R\dagger}(\rho), E)$ , where  $E$  is the class POVM  $E(d\hat{g}) = \tilde{\mathcal{U}}_{\hat{g}}(|\eta\rangle\langle\eta|)d\hat{g}$ .

The question now is: why should we perform a physical transformation on the input state  $\rho$  before performing the measurement? Intuitively, one might expect that the application of the transformation  $\mathcal{C}^{R\dagger}$  can only degrade the information contained in the signal states  $\rho_g := \tilde{\mathcal{U}}_g(\rho)$ . In the next paragraph I will show that the intuition is correct in the case where  $\rho$  is a class state and the cost function  $c(\hat{g}, g)$  has negative Fourier transform.

#### 4.4. Invariant cost functions with negative Fourier transform: optimality of square-root measurements

Here we will confine our attention to a special family of invariant cost functions, that is, to cost functions of the form

$$c(\hat{g}, g) = \sum_{\sigma \in \text{Urr}(\mathbf{G})} a_{\sigma} \chi_{\sigma}^*(\hat{g}^{-1}g) \quad a_{\sigma} \leq 0 \quad \forall \sigma \in \text{Urr}(\mathbf{G}). \quad (21)$$

where  $\text{Urr}(\mathbf{G})$  denotes the set of all unitary irreducible representations of  $\mathbf{G}$  (the complex conjugate of the character has been introduced just for later convenience). Optimization problems with cost functions in this family are the maximization of the *entanglement fidelity*, corresponding to the cost function given by  $c(\hat{g}, g) = -|\text{Tr}[V_{\hat{g}^{-1}g}]/n|^2$ , where  $V : \mathbf{G} \rightarrow \text{Lin}(\mathbb{C}^n)$  is a unitary representation of  $\mathbf{G}$ , and the *maximum likelihood* [17], corresponding to the cost function  $c(\hat{g}, g) = -\delta(\hat{g}^{-1}g)$ .

Incidentally, we observe that each cost function of the form in Eq. (21) defines a positive semidefinite product on the vector space of class functions, given by

$$\begin{aligned} \langle f_1, f_2 \rangle_c &:= - \int_{\mathbf{G}} d\hat{g} \int_{\mathbf{G}} dg \, c(\hat{g}, g) f_1^*(\hat{g}) f_2(g) \\ &= - \sum_{\sigma \in \text{Urr}(\mathbf{G})} a_{\sigma} f_{1,\sigma}^* f_{2,\sigma}. \end{aligned}$$

The next theorem summarizes the structures presented so far:

**Theorem 7 Optimal estimation strategy for invariant functions with negative Fourier transform.** *Let  $\mathbf{G}$  be a compact group,  $\tilde{U}$  be a unitary projective representation of  $\mathbf{G}$  on the Hilbert space  $\mathcal{H} = \bigoplus_{\mu \in \text{lrr}(U)} \mathcal{H}_{\mu} \otimes \mathcal{H}_{\mu}$  and  $c(\hat{g}, g)$  be an invariant cost function of the form of Eq. (21). Then, the class POVM  $E(d\hat{g}) = \mathcal{U}_g(|\eta\rangle\langle\eta|)d\hat{g}$  is optimal for every class state  $|\Phi\rangle$  of the form  $|\Phi\rangle = \bigoplus_{\mu \in \text{lrr}(U)} \sqrt{p_{\mu}/d_{\mu}} |I_{\mu}\rangle$ . The cost is given by  $c_{ave} = \sum_{\sigma \in \text{Urr}(\mathbf{G})} a_{\sigma} m_{\sigma}^{\mu\nu*} \sqrt{p_{\mu}p_{\nu}}$ , where  $m_{\sigma}^{\mu\nu*}$  is the multiplicity of the irreducible representation  $U^{\sigma}$  in the tensor product  $U^{\mu} \otimes U^{\nu*}$ .*

**Proof.** By theorem 5 we can assume without loss of generality that the optimal POVM is covariant, i.e.  $P(d\hat{g}) = \mathcal{U}_g(\xi)d\hat{g}$ , and that the seed  $\xi$  is isotropic. Writing the seed as  $\xi = \sum_{i=1}^r |\eta\rangle\langle\eta|$  with  $|\eta\rangle = \bigoplus_{\mu \in \text{lrr}(U)} \sqrt{d_{\mu}} |\eta_{i\mu}\rangle$  we obtain for the cost

$$\begin{aligned} c(P, \Phi|e) &= \int_{\mathbf{G}} d\hat{g} \, c(\hat{g}, e) \langle \Phi | \tilde{\mathcal{U}}_{\hat{g}}(\xi) | \Phi \rangle \\ &= \sum_{\sigma \in \text{Urr}(\mathbf{G})} \sum_{\mu, \nu \in \text{lrr}(U)} a_{\sigma} \sqrt{\frac{p_{\mu}p_{\nu}}{d_{\mu}d_{\nu}}} \int_{\mathbf{G}} d\hat{g} \, \chi_{\sigma}^*(\hat{g}) \langle \langle I_{\mu} | \tilde{\mathcal{U}}_{\hat{g}}(\xi) | I_{\nu} \rangle \rangle \\ &= \sum_{\sigma \in \text{Urr}(\mathbf{G})} \sum_{\mu, \nu \in \text{lrr}(U)} a_{\sigma} \sqrt{p_{\mu}p_{\nu}} \sum_{i=1}^r \int_{\mathbf{G}} d\hat{g} \, \chi_{\sigma}^*(\hat{g}) \langle \langle I_{\mu} | (U_{\hat{g}}^{\mu} \otimes I_{\mu}) | \eta_{i\mu} \rangle \rangle \langle \langle \eta_{i\nu} | (U_{\hat{g}}^{\nu\dagger} \otimes I_{\nu}) | I_{\nu} \rangle \rangle \\ &= \sum_{\sigma \in \text{Urr}(\mathbf{G})} \sum_{\mu, \nu \in \text{lrr}(U)} a_{\sigma} c \sqrt{p_{\mu}p_{\nu}} \sum_{i=1}^r \int_{\mathbf{G}} d\hat{g} \, \chi_{\sigma}^*(\hat{g}) \langle \langle I_{\mu} | U_{\hat{g}}^{\mu} \eta_{i\mu} \rangle \rangle \langle \langle U_{\hat{g}}^{\nu} \eta_{i\nu} | I_{\nu} \rangle \rangle \\ &= \sum_{\sigma \in \text{Urr}(\mathbf{G})} \sum_{\mu, \nu \in \text{lrr}(U)} a_{\sigma} \sqrt{p_{\mu}p_{\nu}} \sum_{i=1}^r \int_{\mathbf{G}} d\hat{g} \, \chi_{\sigma}^*(\hat{g}) \text{Tr}[U_{\hat{g}}^{\mu} \eta_{i\mu} \otimes U_{\hat{g}}^{\nu*} \eta_{i\nu}^*] \\ &= \sum_{\sigma \in \text{Urr}(\mathbf{G})} \sum_{\mu, \nu \in \text{lrr}(U)} \frac{a_{\sigma} \sqrt{p_{\mu}p_{\nu}}}{d_{\sigma}} \sum_{i=1}^r \text{Tr}[\Pi_{\sigma}^{\mu\nu*} (\eta_{i\mu} \otimes \eta_{i\nu}^*)], \end{aligned} \quad (22)$$

where  $P_\sigma^{\mu\nu*}$  is the projector on the direct sum of all irreducible subspaces of  $\mathcal{H}_\mu \otimes \mathcal{H}_\nu$  that carry the irreducible representation  $U^\sigma$  as a subrepresentation of the tensor product  $U^\mu \otimes U^{\nu*}$ . Since  $a_\sigma \leq 0$  for every  $\sigma$ , we have the bound

$$\begin{aligned} c(P, \Phi|e) &\leq \sum_{\sigma \in \text{Uirr}(\mathbb{G})} \sum_{\mu, \nu \in \text{Irr}(U)} \frac{a_\sigma \sqrt{p_\mu p_\nu}}{d_\sigma} \left| \sum_{i=1}^r \text{Tr}[\Pi_\sigma^{\mu\nu*}(\eta_{i\mu} \otimes \eta_{i\nu}^*)] \right|, \\ &= \sum_{\sigma \in \text{Uirr}(\mathbb{G})} \sum_{\mu, \nu \in \text{Irr}(U)} \frac{a_\sigma \sqrt{p_\mu p_\nu}}{d_\sigma} \left| \sum_{i=1}^r \text{Tr}[\Pi_\sigma^{\mu\nu*}(\eta_{i\mu} \otimes I_\nu)(I_\mu \otimes \eta_{i\nu}^*)\Pi_\sigma^{\mu\nu*}] \right|, \\ &\leq \sum_{\sigma \in \text{Uirr}(\mathbb{G})} \sum_{\mu, \nu \in \text{Irr}(U)} \frac{a_\sigma \sqrt{p_\mu p_\nu}}{d_\sigma} \sqrt{\sum_{i=1}^r \text{Tr}[\Pi_\sigma^{\mu\nu*}(\eta_{i\mu} \eta_{i\mu}^\dagger \otimes I_\nu)]} \sqrt{\sum_{j=1}^r \text{Tr}[(I_\nu \otimes \eta_{i\nu}^T \eta_{i\nu}^*)\Pi_\sigma^{\mu\nu*}]}, \end{aligned}$$

having used the Cauchy-Schwarz inequality  $|\text{Tr}[A^\dagger B]| \leq \sqrt{\text{Tr}[A^\dagger A] \text{Tr}[B^\dagger B]}$  for block matrices  $A = \bigoplus_{i=1}^r A_i$  and  $B = \bigoplus_{i=1}^r B_i$ . Finally, using the normalization conditions of corollary 1 we obtain

$$c(P, \Phi, e) \leq \sum_{\sigma \in \text{Uirr}(\mathbb{G})} \sum_{\mu, \nu \in \text{Irr}(U)} \frac{a_\sigma \sqrt{p_\mu p_\nu}}{d_\sigma} \text{Tr}[\Pi_\sigma^{\mu\nu*}] \leq \sum_{\sigma \in \text{Uirr}(\mathbb{G})} \sum_{\mu, \nu \in \text{Irr}(U)} a_\sigma m_\sigma^{\mu\nu*} \sqrt{p_\mu p_\nu},$$

where  $m_\sigma^{\mu\nu*}$  is the multiplicity of  $U^\sigma$  in the tensor product of  $U^\mu \otimes U^{\nu*}$ . On the other hand, by substituting in Eq. (22) it is immediate to see that the class POVM  $E(d\hat{g}) = \mathcal{U}_{\hat{g}}(|\eta\rangle\langle\eta|)$  achieves the bound and hence is the best possible POVM. ■

The optimality of the class POVM can be rephrased as a result about the optimality of *square-root measurements* [18] (see also [13] for a surprising *ante litteram* example). The general definition of square-root measurement that we will consider here is the following:

**Definition 3** *Let  $(X, \sigma(X))$  be a measurable space and let  $\rho(dx) : \sigma(X) \rightarrow \text{Lin}_+(\mathcal{H})$ ,  $B \mapsto \rho_B$  be an ensemble. The square-root measurement associated to  $\rho$  is the POVM  $P : \sigma(B) \mapsto \text{Lin}_+(\text{Supp}(\rho_X))$  given by  $P_B := \rho_X^{-\frac{1}{2}} \rho_B \rho_X^{-\frac{1}{2}}$ , where  $\rho_X^{-\frac{1}{2}}$  is the inverse of  $\rho_X^{\frac{1}{2}}$  on its support.*

It is easy to see that the class POVM  $E(d\hat{g}) = \tilde{\mathcal{U}}_{\hat{g}}(|\eta\rangle\langle\eta|)d\hat{g}$  is the square-root measurement associated to the ensemble  $\rho(dg) := |\Phi_g\rangle\langle\Phi_g|dg$  where  $|\Phi_g\rangle = \tilde{U}_g|\Phi\rangle$  for some class state  $|\Phi\rangle = \bigoplus_{\mu \in \text{Irr}(U)} \sqrt{\frac{p_\mu}{d_\mu}} |I_\mu\rangle$ . Indeed, we have  $\rho_G = \int_G dg |\Phi_g\rangle\langle\Phi_g| = \bigoplus_{\mu \in \text{Irr}(U)} \frac{p_\mu}{d_\mu^2} I_\mu \otimes I_\mu$ , and therefore  $\rho_G^{-\frac{1}{2}} |\Phi_g\rangle = \tilde{U}_g \rho_G^{-\frac{1}{2}} |\Phi\rangle = \tilde{U}_g \left( \sum_{\mu \in \text{Irr}(U)} \sqrt{d_\mu} |I_\mu\rangle \right) = \tilde{U}_g |\eta\rangle$ , which means  $E(d\hat{g}) = \rho_G^{-\frac{1}{2}} |\Phi_g\rangle\langle\Phi_g| \rho_G^{-\frac{1}{2}} d\hat{g}$ .

A slight generalization of Theorem 7 is then given by the following:

**Corollary 4** *Let  $\mathbb{G}$  be a compact group,  $\tilde{U}$  be a projective unitary representation of  $\mathbb{G}$  on the Hilbert space  $\tilde{\mathcal{H}} = \bigoplus_{\mu \in \text{Irr}(U)} \mathcal{H}_\mu \otimes \mathcal{H}_\mu$  and  $c(\hat{g}, g)$  be an invariant cost function of the form of Eq. (21). Then, the square-root measurement is optimal for every state  $|\Psi\rangle \in \tilde{\mathcal{H}}$  of the form  $|\Psi\rangle = V|\Phi\rangle$  where  $|\Phi\rangle \in \tilde{\mathcal{H}}$  is a class state as in Eq. (19) and  $V \in \text{Lin}(\tilde{\mathcal{H}})$  is a unitary matrix in  $U'$ .*

**Proof.** The square-root measurement associated to the ensemble  $|\Psi_g\rangle\langle\Psi_g|dg$  is  $P(d\hat{g}) = \mathcal{U}_g \mathcal{V}(|\eta\rangle\langle\eta|)$ . It is easy to see that the cost of the strategy  $(\Psi, P)$  equal the cost of the strategy

$(\Phi, E)$ : indeed, the conditional probabilities are the same:

$$\begin{aligned} p'(\mathrm{d}\hat{g}|g) &:= \mathrm{Tr}[P(\mathrm{d}\hat{g})\mathcal{U}_g(|\Psi\rangle\langle\Psi|)] \\ &= \mathrm{Tr}[\mathcal{U}_{\hat{g}}\mathcal{V}(|\eta\rangle\langle\eta|)\mathcal{U}_g\mathcal{V}(|\Phi\rangle\langle\Phi|)]\mathrm{d}\hat{g} \\ &= \mathrm{Tr}[\mathcal{U}_{\hat{g}}(|\eta\rangle\langle\eta|)\mathcal{U}(|\Phi\rangle\langle\Phi|)]\mathrm{d}\hat{g} := p(\mathrm{d}\hat{g}|g) \end{aligned}$$

This means that the square-root measurement  $P(\mathrm{d}\hat{g})$  must be optimal for  $|\Psi\rangle$ , otherwise  $E$  would not be optimal for  $|\Phi\rangle$ . ■

## 5. Conclusions and open problems

This paper provided a synthesis of the main group theoretical structures appearing in the estimation of unitary transformations on finite dimensional Hilbert spaces. I deliberately avoided any mention to infinite dimensions, in order to keep the exposition at the simplest possible level and to highlight the structures that underly optimal estimation in the neatest possible way. The development of a general theory of optimal estimation for infinite dimensional systems is mostly an open problem, especially in the case of non-compact groups, for which there are many particular results (especially for abelian groups) but still no unifying framework. The development of a general theory of estimation for non-abelian non-compact groups rests as a stimulating challenge for future research.

## 6. Acknowledgments

This paper contains an advanced elaboration of some ideas present in my PhD thesis “Optimal estimation of quantum signals in the presence of symmetry” (Pavia, October 2006). I would like to take this opportunity to express my gratitude to my PhD advisor and friend G. M. D’Ariano and to my extraordinary collaborators P. Perinotti and M. F. Sacchi. Research at Perimeter Institute for Theoretical Physics is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI.

- [1] D. Leibfried *et al*, Nature **438**, 639 (2005). T. Nagata, R. Okamoto, J. L. O’Brien, K. Sasaki, and S. Takeuchi, Science **316**, 726 (2007). B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, G. J. Pryde, Nature **450**, 393 (2007).
- [2] S. T. Ali, J. P. Antoine and J. P. Gazeau, *Coherent States, Wavelets and Their Generalizations*, Springer-Verlag, 2000.
- [3] V. Bužek, R. Derka, and S. Massar, Phys. Rev. Lett. **82**, 2207, (1999).
- [4] G. Chiribella, G. M. D’Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. Lett. **93**, 180503 (2004).
- [5] S. D. Bartlett, T. Rudolph, R. W. Spekkens Rev. Mod. Phys. **79**, 555 (2007).
- [6] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
- [7] Precisely,  $\sigma(\mathbf{G})$  denotes the Borel  $\sigma$ -algebra generated by the topology of  $\mathbf{G}$ .
- [8] G. M. D’Ariano, P. Lo Presti, and M. F. Sacchi, Phys. Lett. A **272**, 32 (2000).
- [9] G. B. Folland, editor. *A course in abstract harmonic analysis*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995.
- [10] The definition is more intuitive if we use formally the Dirac vectors  $\{|h\rangle\}_{h\in\mathbf{G}}$ , with  $\langle h_1|h_2\rangle = \delta(h_1^{-1}h_2)$  and write  $|f\rangle = \int_{\mathbf{G}} \mathrm{d}h f(h)|h\rangle$ . With this notation we have  $T_g|h\rangle = \omega(g, h)|gh\rangle$ , which shows clearly that  $T$  is a projective representation with multiplier  $\omega$ .
- [11] C. W. Helstrom, *Quantum detection and estimation theory* (Academic Press, New York, 1976).
- [12] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North Holland, Amsterdam, 1982).
- [13] V. P. Belavkin and V. P. Maslov, *Design of Optimal Dynamic Analyzers: Mathematical Aspects of Wave Pattern Recognition*, Mathematical Aspects Of Computer Engineering Advances in Science and Technology in the USSR, Mir Publisher (1988), <http://arxiv.org/abs/quant-ph/041203>.
- [14] G. Chiribella, G. M. D’Ariano, and M. F. Sacchi, Phys. Rev. A **72**, 042338 (2005).
- [15] G. Chiribella, G. M. D’Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).
- [16] E. Schrödinger, Proc. Camb. Phil. Soc. **31**, 555 (1935).
- [17] G. Chiribella, G. M. D’Ariano, P. Perinotti, M. F. Sacchi, Int. J. Quantum Inf. **4**, 453 (2006).
- [18] P. Hausladen and W. K. Wootters, J. Mod. Opt. **41**, 2385 (1994).