PHYSICAL REVIEW
LETTERS

VOLUME 80 23 FEBRUARY 1998 NUMBER 8

Universal Algorithm for Optimal Estimation of Quantum States from Finite Ensembles via Realizable Generalized Measurement

R. Derka,[1] V. Bužek,[2,3] and A. K. Ekert[1]
[1]*Department of Physics, Oxford University, Parks Road, OX1 3PU Oxford, United Kingdom*
[2]*Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 842 28 Bratislava, Slovakia*
[3]*Optics Section, The Blackett Laboratory, Imperial College, London SW7 2BZ, United Kingdom*
(Received 14 July 1997)

We present a universal algorithm for the optimal quantum state estimation of an arbitrary finite dimensional system. The algorithm specifies a *physically realizable* (i.e., finite) positive operator valued measurement on a finite number of identically prepared systems. We illustrate the general formalism by applying it to different scenarios of the state estimation of N independent and identically prepared two-level systems (qubits). [S0031-9007(98)05400-3]

PACS numbers: 03.65.Bz

Suppose we have N quantum objects, each prepared in an unknown *pure* quantum state described by a density operator $\hat{\rho} = |\psi\rangle\langle\psi|$. The question is: *What kind of measurement provides the best possible estimation of* $\hat{\rho}$*?* Clearly, if we have an unlimited supply of particles in the state $\hat{\rho}$, i.e., when N approaches infinity, we can estimate $\hat{\rho}$ with an arbitrary precision. In practice, however, only finite and usually small ensembles of identically prepared quantum systems are available. This leads to an important problem of the optimal state estimation with limited physical resources. It is a generic problem, common to many areas of quantum physics ranging from the ultraprecise quantum metrology to eavesdropping in quantum cryptography.

Within a framework of an elementary group theory the problem of the state estimation can be reformulated as a more general problem of estimating an unknown unitary operation from a group of transformations acting on a given quantum system (i.e., the state estimation follows as a special case). Holevo [1] has shown that this problem can be solved via the *covariant measurement* (CM) approach. Unfortunately, the covariant measurement corresponds to an *infinite* (i.e., consisting of an infinite continuous set of operators) and therefore physically nonrealizable positive operator valued measurement (POVM). We note that from the logic of the CM it follows that if any optimal measurement (finite or infinite) does exist then using a simple formal construction one can generate from the original optimal measurement another measurement which is covariant and which, at the same time, conserves optimality of the original solution. In the present Letter we address the question of how to find *finite* optimal generalized measurements if they exist. This is a fundamental question because only *finite* POVM schemes are experimentally realizable. We propose a universal algorithm about how to look for these POVM schemes and we apply it explicitly in two physically interesting cases of the state estimation of N identically prepared two-level systems (qubits).

In order to set up the scene, let us assume that state $\hat{\rho}$ is generated from a reference state $\hat{\rho}_0 = |\psi_0\rangle\langle\psi_0|$ by a unitary operation $U(\mathbf{x})$ which is an element of a particular unitary finite dimensional representation of a compact Lie group G. Different $\mathbf{x}$ denote different points of the group [e.g., different angles of rotation in the case of the SU(2)] and we assume that all values of $\mathbf{x}$ are equally probable.

Our task is to design the most general POVM, mathematically described as a set $\{\hat{O}_r\}_{r=1}^{R}$ of positive Hermitian operators such that $\sum_r \hat{O}_r = \hat{1}$ [2,3], which when applied to the *combined* system of *all* N copies provides us with the best possible estimation of $\hat{\rho}$ [and therefore also of $U(\mathbf{x})$]. We quantify the quality of the state estimation in terms of the *mean* fidelity

$$\overline{f} = \sum_r \int_G d\mathbf{x}\, \mathrm{Tr}[\hat{O}_r \overbrace{U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x}) \otimes \cdots \otimes U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x})}^{N \text{ times}}] \times \mathrm{Tr}[U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x}) U_r \hat{\rho}_0 U_r^\dagger], \quad (1)$$

which corresponds to a particular choice of a cost function [3] used in a context of detection and estimation theory. The mean fidelity (1) can be understood as follows: In order to assess how good a chosen measurement is we apply it many times *simultaneously* on *all* N particles each in state $U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x})$. The parameter $\mathbf{x}$ varies randomly and isotropically [4] over all points of the group G during many runs of the measurement.

For each result r of the measurement, i.e., for each operator $\hat{O}_r$, we prescribe the state $|\psi_r\rangle = U_r|\psi_0\rangle$ representing our guess (i.e., estimation) of the original state. The probability of the outcome r is equal to $\mathrm{Tr}[\hat{O}_r U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x}) \otimes \cdots \otimes U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x})]$, while the corresponding fidelity of our estimation is $\mathrm{Tr}[U(\mathbf{x})\hat{\rho}_0 \times U^\dagger(\mathbf{x}) U_r \hat{\rho}_0 U_r^\dagger]$. This fidelity is then averaged over all possible outcomes and over many independent runs of the measurement with randomly and isotropically distributed parameters $\mathbf{x}$. We want to find the generalized measurement which *maximizes* the mean fidelity $\overline{f}$ given by Eq. (1).

The combined system of N identically prepared reference states always remains within the *totally symmetric subspace* of $H^k \otimes H^k \otimes \cdots H^k$, where H^k is k-dimensional Hilbert space of the reference state in which the corresponding unitary representation $U(\mathbf{x})$ acts. Thus the dimensionality d of the space in which we construct the POVM $\{\hat{O}_r\}$ is $d = \binom{N+k-1}{k-1}$. In this case the first trace in Eq. (1) can be rewritten as

$$\overline{f} = \sum_r \int_G \mathrm{Tr}[\hat{O}_r U^N(\mathbf{x})\hat{\Omega}_0 U^{N\dagger}(\mathbf{x})] \times \mathrm{Tr}[U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x}) U_r \hat{\rho}_0 U_r^\dagger]\, d\mathbf{x}\,, \quad (2)$$

where $U^N(\mathbf{x})$ is a new representation of the same group G; it is equivalent to the N-fold symmetrized direct product [5] of the original representation $U(\mathbf{x})$. Here $U^N(\mathbf{x})$ transforms the $\binom{N+k-1}{k-1}$-dimensional reference state denoted as $\hat{\Omega}_0$.

We can insert the identity operator $U_r^N U_r^{N\dagger}$ into the first trace in Eq. (2) and, taking into account that in Eq. (2) we integrate over the whole group G parametrized by $\mathbf{x}$, we can substitute $U^N(\mathbf{x})U_r^{N\dagger} \to U^N(\mathbf{x})$ and $U(\mathbf{x})U_r^\dagger \to U(\mathbf{x})$. Now, using the linearity of the trace operation as well as the linearity of the representation of the group G ($U\hat{\rho}U^\dagger$ is a linear adjoint representation) we rewrite Eq. (2) as

$$\overline{f} = \sum_r \mathrm{Tr}[\hat{O}_r U_r^N \hat{F} U_r^{N\dagger}], \quad (3)$$

where

$$\hat{F} = \int_G U^N(\mathbf{x})\hat{\Omega}_0 U^{N\dagger}(\mathbf{x})\, \mathrm{Tr}[U(\mathbf{x})\hat{\rho}_0 U^\dagger(\mathbf{x})\hat{\rho}_0]\, d\mathbf{x}\,, \quad (4)$$

is a positive Hermitian operator.

Let us now derive an upper bound on the mean fidelity. Taking into account positivity of the operator $\hat{F}$ (i.e., $\hat{F} = \sum_i \lambda_i |\phi_i\rangle\langle\phi_i|$; $\lambda_i \geq 0$) and the completeness condition for POVM (i.e., $\sum_r \hat{O}_r = \hat{1}$) we obtain

$$\overline{f} = \sum_r \mathrm{Tr}[\hat{O}_r U_r^N \hat{F} U_r^{N\dagger}] = \sum_{ir} \lambda_i \mathrm{Tr}[\hat{O}_r U_r^N |\phi_i\rangle\langle\phi_i| U_r^{N\dagger}] \leq \lambda_{\max} \sum_{ir} \mathrm{Tr}[\hat{O}_r U_r^N |\phi_i\rangle\langle\phi_i| U_r^{N\dagger}] = \lambda_{\max} \sum_r \mathrm{Tr}[\hat{O}_r U_r^N \hat{1} U_r^{N\dagger}] = \lambda_{\max} \mathrm{Tr}[\hat{1}] = \lambda_{\max} d\,. \quad (5)$$

From Eq. (5) it clearly follows that the upper bound can be achieved if and only if all operators $\hat{O}_r$ forming the POVM satisfy the following conditions: (i) Each $\hat{O}_r$ is proportional to a suitably rotated (by some U_r^N) projector on the eigenvector of $\hat{F}$ with the highest eigenvalue, i.e., for all $\hat{O}_r$ there exists U_r^N, such that $\hat{O}_r = c_r U_r^N |\phi_{\max}\rangle\langle\phi_{\max}| U_r^{N\dagger}$. This U_r^N, or more precisely $U_r|\psi_0\rangle$, is our guess associated with the result "r." (ii) All c_r are real and positive, to assure that all $\hat{O}_r$ are positive operators. (iii) Finally, the operators $\hat{O}_r$ have to satisfy the completeness criterion $\sum_r c_r U_r^N |\phi_{\max}\rangle\langle\phi_{\max}| U_r^{N\dagger} = \hat{1}$. As shown by Holevo [1] in the case of the *infinite* POVM condition (iii) is fulfilled for covariant measurements, providing the representation U^N of the group G is irreducible (see example B below). This statement follows from the Shur lemma. However, in the general case of reducible representation and specifically for *finite* realizable POVMs this argument cannot be used and we have to proceed differently.

To find the solution of the problem we start with the following observation. Let us assume that we have some POVM $\{\hat{O}_r\}_{r=1}^R$ and the corresponding guesses U_r^N which maximize the mean fidelity $\overline{f}$. We can always construct another POVM with more elements which is also optimal. For example, let us consider a one-parametric subgroup $U(\phi) = \exp(i\hat{X}\phi)$ of our original group G and choose a basis $\{|m\rangle\}_{m=1}^d$ in which the action of this subgroup is equivalent to multiplication by a factor $\mathbf{e}^{i\omega_m \phi}$ [i.e., the operator $U(\phi)$ is diagonal in this basis and ω_m are eigenvalues of the generator $\hat{X}$]. Then we take d points ϕ_s ($s = 1, \ldots, d$) and generate from each original operator $\hat{O}_r$ a set of d operators $\hat{O}_{rs} = \frac{1}{d} U^N(\phi_s)\hat{O}_r U^{N\dagger}(\phi_s)$. In this way we obtain a new set of $(d \cdot R)$ operators such that the mean fidelity for

this new set of operators, $\overline{f} = \sum_{r,s} \mathrm{Tr}[\hat{O}_{r,s} U_{r,s}^N \hat{F} U_{r,s}^{N\dagger}]$, is *equal* to the mean fidelity of the original POVM $\{\hat{O}_r\}$ because we ascribe to each eventual result $[r, s]$ a new guess $U_{r,s} = U(\phi_s)U_r$. However, in order to guarantee that the new set of operators $\hat{O}_{rs}$ is indeed a POVM we have to satisfy the completeness condition

$$\begin{aligned} \hat{1} &= \sum_s \sum_r \hat{O}_{rs} = \sum_s \sum_r \frac{1}{d}\, U^N(\phi_s)\hat{O}_r U^{N\dagger}(\phi_s) \\ &= \sum_s \sum_{m,n} \frac{e^{i\phi_s(\omega_m - \omega_n)}}{d} \sum_r (\hat{O}_r)_{mn} |m\rangle\langle n| \,. \end{aligned} \tag{6}$$

Let us notice that, by the appropriate choice of ϕ_s, the sum $\sum_s \frac{e^{i\phi_s(\omega_m-\omega_n)}}{d}$ can *always* be made equal to $\delta_{m,n}$ providing all eigenvalues are nondegenerate [6] (this is basically a discrete Fourier transform and we illustrate this point in detail in example A). In this case, the conditions (6) for the off-diagonal terms in the basis $|m\rangle$ are *trivially* satisfied whereas the diagonal terms are equal to unity because the original POVM $\{\hat{O}_r\}$ guarantees that $\sum_r (\hat{O}_r)_{mm} = 1$. Moreover, even if the original set of operators $\{\hat{O}_r\}$ does not satisfy the full completeness condition and the conditions for the off-diagonal terms are *not* satisfied (i.e., these operators do not constitute a POVM) we can, using our extension ansatz, always construct a *proper* POVM $\{\hat{O}_{r,s}\}$. This *proves* that when we maximize the mean fidelity (3) it is enough to assume d diagonal conditions rather than the original complete set of d^2 constraints for diagonal *and* off-diagonal elements.

Now we turn back to our original problem of how to construct the POVM which maximizes the mean fidelity. To do so we first express the operators $\hat{O}_r$ in the form $\hat{O}_r = c_r U_r^N |\Psi_r\rangle\langle\Psi_r| U_r^{N\dagger}$, where $|\Psi_r\rangle$ are general normalized states in the d-dimensional space in which the operators $\hat{O}_r$ act, and c_r are positive constants. This substitution is done without any loss of generality [7] and it permits us to rewrite Eq. (3) so that the mean fidelity $\overline{f}$ does not explicitly depend on U_r^N, i.e.,

$$\overline{f} = \sum_r c_r \,\mathrm{Tr}[|\Psi_r\rangle\langle\Psi_r|\hat{F}] \,. \tag{7}$$

Obviously, the completeness condition $\sum_r \hat{O}_r = \hat{1}$ is now modified and it reads

$$\sum_r c_r U_r^N |\Psi_r\rangle\langle\Psi_r| U_r^{N\dagger} = \hat{1} \,. \tag{8}$$

From our discussion above it follows that when maximizing the mean fidelity (7) it is enough to apply only d constraints $\sum_r c_r |\langle m|U_{N,r}|\Psi_r\rangle|^2 = 1$ (here $m = 1, \ldots, d$) out of the d^2 constraints (8). Therefore to accomplish our task we solve a set of Lagrange equations with d Lagrange multipliers L_m. If we express L_m as eigenvalues of the operator $\hat{L} = \sum_m L_m |m\rangle\langle m|$ then we obtain the final very compact set of equations determining the optimal POVM

$$\begin{aligned} [\hat{F} - U_r^{N\dagger}\hat{L}U_r^N]|\Psi_r\rangle &= 0 \,, \\ \sum_r c_r |\langle m|U_r^N|\Psi_r\rangle|^2 &= 1 \,. \end{aligned} \tag{9}$$

From here it follows that $|\Psi_r\rangle$ are determined as zero-eigenvalue eigenstates. More specifically, they are functions of d Lagrange multipliers $\{L_m\}_{m=1}^d$ and R vectors $\{\mathbf{x}_r\}_{r=1}^R$ [where $\mathbf{x}_r$ determine U_r as $U_r = U(\mathbf{x}_r)$]. These free parameters are in turn related via R conditions $\mathrm{Det}[(\hat{F} - U_r^{N\dagger}\hat{L}U_r^N)] = 0$. The mean fidelity now is equal to $\mathrm{Tr}\,\hat{L}$. At this stage we solve a system of d linear equations [see the second formula in Eq. (9)] for R unknown parameters c_r. All solutions for c_r parametrically depend on L_m and $\mathbf{x}_r$ which are specified above. We note that the number of free parameters in our problem depends on R which has not been specified yet. We choose R such that there are enough free parameters so that the mean fidelity is maximized and simultaneously all c_r are positive. This freedom in the choice of the value of R also reflects the fact that there is an infinite number of equivalent (i.e., with the same value of the mean fidelity) optimal POVMs. The whole algorithm is completed by finding ϕ_s from Eq. (6) which explicitly determine the *finite* optimal POVM $\{\hat{O}_{rs}\}$. This is the main result of our Letter.

In the following we will apply this general algorithm into two physically important examples.

Example A.—Suppose we have N identical copies of spin $1/2$ all prepared in the same but unknown pure quantum state. If we choose the group G to be U(2), i.e., the complete unitary group transforming a two-level quantum system, we can straightforwardly apply the optimal estimation scheme as described above. To be more precise, due to the fact that there exist elements of the group U(2) for which the reference state is the fixed point (i.e., it is insensitive to its action) we have to work only with the coset space $^{\mathrm{SU}(n)}|_{\mathrm{U}(n-1)}$ [5]. In the present case this is a subset of the SU(2) group parametrized by two Euler angles θ, ψ (the third Euler angle χ is fixed and equal to zero). This subset is isomorphic to the Poincaré sphere.

The unitary representation U is now the representation $(\frac{1}{2})$ [we use a standard classification of SU(2) representations, where (j) is the spin number]. Its N-fold symmetrized direct product (we denote this representation as U^N) is the representation classified as $(\frac{N}{2})$ (which transforms a spin-$N/2$ particle). Choosing the standard basis $|j, m\rangle$ with $m = -j, \ldots, j$ in which the coordinate expression for $U(\theta, \psi)$ corresponds to standard rotation matrices $D^j_{m,n}(\theta, \psi, 0) = e^{-im\psi} d^j_{m,n}(\theta)$ [8], we obtain the matrix expression for the operator $\hat{F}$

$$\begin{aligned} F_{m,n} &= \int_0^{2\pi} d\phi \int_0^{\pi} \frac{\sin(\theta)\, d\theta}{8\pi} (1 + \cos\theta) \\ &\quad \times D^{\frac{N}{2}}_{m,\frac{N}{2}}(\theta, \phi) D^{\frac{N}{2}*}_{n,\frac{N}{2}}(\theta, \phi) \\ &= \frac{N/2 + m + 1}{(N+2)(N+1)} \,\delta_{m,n} \,. \end{aligned} \tag{10}$$

When we insert this operator in Eq. (5) we immediately

find the upper bound on the mean fidelity to be equal to $\frac{N+1}{N+2}$.

This is the main result of the paper by Massar and Popescu [9] who noted that this upper bound can be attained using the special POVM which consists of an *infinite* continuous set of operators proportional to isotropically rotated projector $|\frac{N}{2},\frac{N}{2}\rangle\langle\frac{N}{2},\frac{N}{2}|$. This result is closely related to the covariant measurements of Holevo [1].

However, our aim is to construct an optimal and *finite* POVM. To do so, we have to find a finite set of pairs of angles $\{(\theta_r,\psi_r)\}$ such that the completeness conditions (8) which now take the form

$$\sum_r c_r e^{-i\psi_r(m-n)} d^{\frac{N}{2}}_{m,\frac{N}{2}}(\theta_r) d^{\frac{N}{2}}_{n,\frac{N}{2}}(\theta_r) = \delta_{m,n} \qquad (11)$$

are fulfilled. Following our general scheme we first satisfy the completeness conditions (11) for diagonal terms [compare with Eq. (9)]

$$\sum_r c_r d^{\frac{N}{2}}_{m,\frac{N}{2}}(\theta_r)^2 = 1; \quad m = -N/2,\ldots,N/2. \qquad (12)$$

To satisfy these completeness conditions we choose $N+1$ angles θ_r to be equidistantly distributed in the $\langle 0,\pi\rangle$ [obviously, there are many other choices which may suit the purpose—see discussion below Eq. (9)]. Then we solve the system of linear equations for $N+1$ variables c_r. For this choice of θ_r the system (12) has nonnegative solutions. Finally we satisfy the off-diagonal conditions by choosing $N+1$ angles $\psi_s = \frac{2s\pi}{N+1}$ for each θ_r. In this case $\frac{1}{N+1}\sum_{s=0}^{N} e^{i\psi_s y} = \delta_{y,0}$ for all $y = -N/2,\ldots,N/2$ and the off-diagonal conditions are satisfied straightforwardly. This concludes the construction of the *optimal* and *finite* POVM for the spin-1/2 state estimation.

Example B.—Consider a system of N effectively two-level atoms (qubits), all initially prepared in the reference state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ by applying so called $\frac{\pi}{2}$ pulse to initially deexcited atoms. Then the atoms undergo the free evolution effectively described by the U(1) group; i.e., the state of the single qubit evolves as $\frac{1}{\sqrt{2}}(|0\rangle + \exp\{i\psi(t)\}|1\rangle)$. Our task is to find a measurement which provides the optimal estimation of the phase $\psi(t)$ of the U(1) rotation which carries the information about the interaction parameters.

In the standard classification of representations of the U(1) group the single isolated qubit is described by the direct sum of two one-dimensional representations $U = (0) \oplus (1)$. The representation U^N transforming the entire system of N qubits is then equal to the direct sum of representations of the form $(0) \oplus (1) \oplus \cdots (N)$. This acts in the $N+1$ dimensional space spanned by basis vectors $|m\rangle$, $m = 0,1,\ldots,N$. In this basis matrix elements $\hat{F}_{m,n}$ of the operator $\hat{F}$ given by Eq. (4) take the form

$$\hat{F}_{m,n} = \int_0^{2\pi} \frac{d\psi}{2\pi} \frac{\sqrt{\binom{N}{N-m}\binom{N}{N-n}}}{2^{N+1}} e^{i\psi(n-m)}(1+\cos\psi) \qquad (13)$$
$$= \frac{\sqrt{\binom{N}{N-m}\binom{N}{N-n}}}{2^{N+2}} (2\delta_{m,n} + \delta_{m,n+1} + \delta_{m+1,n}).$$

The upper bound on the fidelity Eq. (5) is now too conservative to be of any use (greater than unity). We can, however, solve the system of Eqs. (9) which in this particular case of the commutative group reads

$$[\hat{F} - \hat{L}]|\Psi\rangle = 0; \qquad |\langle m|\Psi\rangle|^2 = 1; \quad \forall m. \qquad (14)$$

The condition $\mathrm{Det}(\hat{F} - \hat{L}) = 0$ now determines the eigenvector $|\Psi\rangle$ with the zero eigenvalue as a function of Lagrange multipliers L_m. When we substitute this eigenvector into the second equation in Eq. (14) we obtain a set of equations for L_m from which the state $|\Psi\rangle$ can be determined. The final POVM is then constructed by rotation of $|\Psi\rangle$ by $N+1$ angles ϕ_s in such a way that all off-diagonal elements of $\sum_s (\hat{O}_s)_{m,n}$ become equal to zero. This is done in exactly the same way as in example A. The resulting POVM corresponds to the *von Neumann measurement* performed on the *composite* system of *all* N ions characterized by the set of orthogonal projectors

$$\hat{P}_s = |\Psi_s\rangle\langle\Psi_s|; \qquad |\Psi_s\rangle = \frac{1}{\sqrt{N+1}} \sum_{q=0}^{N} e^{i\frac{2\pi}{N+1}sq}|q\rangle, \qquad (15)$$

and the maximal mean fidelity $\overline{f}$ is given as the sum: $\overline{f} = 1/2 + 1/2^{N+1}\sum_{i=0}^{N-1}\sqrt{\binom{N}{i}\binom{N}{i+1}}$.

Finally, we note that the Hermitian operator $\hat{\Phi}$ constructed from the optimal POVM (15)

$$\hat{\Phi} = \sum_{s=0}^{N} \frac{2\pi}{N+1} s\hat{P}_s, \qquad (16)$$

with the corresponding guesses as eigenvalues, is identical to the Pegg-Barnett Hermitian phase operator [11] originally introduced within a completely different context.

In conclusion, we have presented a general algorithm for the optimal state estimation from finite ensembles. It provides finite POVMs which, following the Neumark theorem [10], can, at least in principle, be implemented as simple quantum computations. We discuss these aspects in detail elsewhere [12].

We thank Serge Massar, Jason Twamley, Susana Huelga, Thomas Pellizzari, and Chiara Macchiavello for helpful discussions. This work was supported by the Open Society Fund and FCO, the United Kingdom EPSRC, European TMR Network ERP-4061PL95-1412, Hewlett-Packard, Elsag-Bailey, and The Royal Society.

[1] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982), p. 163, and references therein.

[2] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic Publishers, Dordrecht, 1993).
[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[4] We note that this isotropy condition is equivalent to a "no *a priori* information" condition and is associated with the specific integration measure in Eq. (1). This measure has to be invariant under the action of all unitary transformations on the state space of pure states.
[5] A. O. Barut and R. Raczka, *Theory of Group Representations and Applications* (World Scientific, Singapore, 1986).
[6] In the case when the spectrum of the generator $\hat{X}$ is degenerate, i.e., for some m and n we have $\omega_m = \omega_n$, then our algorithm is still valid, provided we increase a number of Lagrange multipliers in Eq. (9) to account for off-diagonal elements L_{mn} and L_{nm} in the definition of the operator $\hat{L}$ in Eq. (9).
[7] The most general choice of $\hat{O}_r$ would be $\hat{O}_r = \sum_i c_{r,i}|\Psi_{r,i}\rangle\langle\Psi_{r,i}|$. However, from the point of view of optimality of the POVM these operators are always less effective than operators $\hat{O}_r = c_r U_r^N|\Psi_r\rangle\langle\Psi_r|U_r^{N\dagger}$ which are proportional to one-dimensional projectors.
[8] R. N. Zare, *Angular Momentum* (Wiley, New York, 1988).
[9] S. Massar and S. Popescu, Phys. Rev. Lett. **74**, 1259 (1995).
[10] M. A. Neumark, C.R. Acad. Sci. USSR **41**, 359 (1943).
[11] D. T. Pegg and S. M. Barnett, Europhys. Lett. **6**, 483 (1988).
[12] R. Derka and V. Bužek, "Optimal Estimation of Quantum States from Finite Ensembles: From Pure Theory to Hypothetic Experiments (unpublished).