OPTIMAL ESTIMATION OF ENSEMBLE AVERAGES FROM A QUANTUM MEASUREMENT

Paolo Perinotti[1,2] and Giacomo M. D'Ariano[1]

[1]Dipartimento di Fisica "A. Volta",
via Bassi 6, 27100 Pavia, Italy
[2]perinotti@fisicavolta.unipv.it

We consider the general measurement scenario in which the ensemble average of an operator is determined via suitable data-processing of the outcomes of a quantum measurement described by a POVM. After reviewing the optimization of data processing that minimizes the statistical error of the estimation, we provide a compact formula for the evaluation of the estimation error.

1. INTRODUCTION

A measurement that can be performed in the lab is described by a POVM (acronym for Positive Operator-Valued Measure), namely a set of (generally nonorthogonal) positive operators $P_i \geq 0$, $1 \leq i \leq N$ which resolve the identity $\sum_{i=1}^N P_i = I$ similarly to the orthogonal projectors of an observable [1]. The probability distribution of the ith outcome is given by the Born rule

$$p(i|\rho) = \mathrm{Tr}[P_i \rho] \tag{1}$$

ρ being the density operator of the state. By such a measurement one can experimentally determine the ensemble averages of (generally complex) operators X. Clearly this is possible if X can be expanded over the POVM elements (mathematically we denote this condition as $X \in \mathrm{Span}\{P_i\}_{i=1,N}$). This means that there exists a set of coefficients $f_i[X]$ such that

$$X = \sum_{i=1}^N f_i[X] P_i, \quad X \in \mathcal{S} := \mathrm{Span}\{P_i\}_{i=1,N} \tag{2}$$

When $\mathcal{S} \equiv \mathcal{B}(\mathcal{H})$ (i. e. when all operators can be expanded over the POVM), then the measurement is informationally complete. Obviously, once the expansion (2) is established one can obtain the ensemble average of X by the following averaging

$$\langle X \rangle = \sum_{i=1}^N f_i[X] p(i|\rho), \tag{3}$$

where the probability distribution is given in Eq. (1).

The above general measurement procedure opens the problem of finding the coefficients $f_i[X]$ in Eq. (2), namely the data-processing of the measurement outcomes needed to determine the ensemble average of X. In general the coefficients $f_i[X]$ are not unique (if $N > \dim(\mathcal{S})$), and one then wants to optimize the data-processing according to a practical criterion, typically minimizing the statistical error. This problem has been solved in the general case in [2], and its solution will be reviewed in this paper. Here in addition We will present a simple formula for the minimum estimation error for arbitrary operator X.

Notice that although the processing functions are intrinsically linear in the definition (2), there is no guarantee that the optimal ones are linear in X. Remarkably, however, the optimal processing function is indeed linear in X, and depends only on the POVM and, in a *Bayesian scheme*, on the ensemble of possible input states. The derivation of the optimal data-processing function requires elementary notions of frame theory [3, 4] and linear algebra [5], which will be introduced in the first part of the paper.

2. MATHEMATICAL TOOLS

A frame in a Hilbert space $\mathcal{K}$ is a set of vectors $\{\phi_n\} \subseteq \mathcal{K}$ satisfying the property

$$a\|\psi\|^2 \leq \sum_n |\langle\phi_n|\psi\rangle|^2 \leq b\|\psi\|^2, \tag{4}$$

for all $\psi \in \mathcal{K}$, with fixed $0 < a \leq b < \infty$. The starting theorem in frame theory states that the set $\{\phi_n\}$ is a frame iff the positive operator, called *frame operator*

$$F = \sum_n |\phi_n\rangle\langle\phi_n|, \tag{5}$$

is bounded and invertible. In this case we can define the *canonical dual frame* $\{\chi_n\}$ by the following formula

$$|\chi_n\rangle = F^{-1}|\phi_n\rangle, \tag{6}$$

and all the vectors $\psi \in \mathcal{K}$ can be written as a linear combination of the vectors $\{\phi_n\}$ as follows

$$|\psi\rangle = \sum_n |\phi_n\rangle\langle\chi_n|\psi\rangle. \tag{7}$$

When the frame is made of linearly dependent vectors, the choice of the coefficients in the expansion Eq. (7) is not unique, and all alternate choices are provided by *alternate dual frames* $\{\eta_n\}$, classified by the relation [6]

$$|\eta_n\rangle = |\chi_n\rangle + |\delta_n\rangle - \sum_m |\delta_m\rangle\langle\phi_m|\chi_n\rangle, \tag{8}$$

where $\{\delta_n\} \subseteq \mathcal{K}$ is an arbitrary set of vectors. This theorem is useful in our case because we can consider the POVM elements P_i as vectors in the space of Hilbert-Schmidt operators—which for finite dimensional systems are all possible operators X—and they provide a frame in the space $\mathcal{S}$. Frame theory then solves the problem of finding all possible sets of coefficients $f_i[X]$ in Eq. (2), which are simply given by the scalar products $\langle D_i|X\rangle := \mathrm{Tr}[D_i^\dagger X]$, $\{D_i\}$ being an alternate dual for the frame $\{P_i\} \subseteq \mathcal{S}$.

In order to answer the main question of the paper, namely which is the dual frame $\{D_i\}$ providing the minimum statistical error, we will first show that the statistical error can be written in terms of a norm for the vector $\{f_i[X]\}$ of coefficients. Indeed, if we consider the ensemble of possible input states $\{\rho_k, p_k\}$, we can define the statistical error in a fixed state ρ_k, and use its average over the ensemble as a figure of merit. We have

$$\delta_D(X) := \sum_{i=1}^{N} \mathrm{Tr}[\rho_{\mathcal{E}} P_i]|f_i[X]|^2 - \overline{\langle X\rangle^2}_{\mathcal{E}}, \tag{9}$$

where $\rho_{\mathcal{E}} := \sum_k p_k \rho_k$, and $\overline{\langle X\rangle^2}_{\mathcal{E}} := \sum_k p_k |\mathrm{Tr}[\rho_k X]|^2$. The second term in Eq.(9) does not depend on the choice of the dual, then the minimization problem can be stated as the minimization of the norm

$$\|f[X]\|_\pi := \sum_{i=1}^{N} \pi_i |f_i[X]|^2, \tag{10}$$

where $\pi_i = \mathrm{Tr}[P_i X]$. If we now consider the following linear map that takes vectors of coefficients to operators

$$\Lambda : c \rightarrow \sum_{i=1}^{N} c_i P_i, \tag{11}$$

its matrix elements are given by $\Lambda_{mn,i} = (P_i)_{mn}$. One can easily prove that all generalized inverses Γ of Λ, satisfying $\Lambda\Gamma\Lambda = \Lambda$, have matrix elements $\Gamma_{i,mn} = (D_i^*)_{mn}$ where $\{D_i\}$ is an alternate dual frame for $\{P_i\}$. The minimum noise can be obtained through the *minimum norm* generalized inverse Γ that must satisfy the relation [7]

$$\pi\Gamma\Lambda = \Lambda^\dagger\Gamma^\dagger\pi, \tag{12}$$

where π is the positive diagonal matrix with eigenvalues π_i.

3. MINIMIZATION OF ERROR

Since the minimum norm generalized inverse is unique and does not depend on the vector, the optimal dual does not depend on X, and the function $f_i[X] = \langle D_i | X\rangle$ is linear, as anticipated. One can prove that the optimal dual frame $\{D_i\}$ corresponding to such Γ is unique and can be expressed as follows [2]

$$D_i = \Delta_i - \sum_j \{[(I-M)\pi(I-M)]^\ddagger \pi\}_{ij} \Delta_j, \tag{13}$$

where $\{\Delta_i\}$ is the canonical dual and M is the projection matrix with elements $M_{ij} = \mathrm{Tr}[\Delta_i P_j]$. The minimum noise for X can be expressed as

$$\delta_D(X) = \langle X|\Gamma^\dagger\pi\Gamma|X\rangle - \overline{\langle X\rangle^2}_{\mathcal{E}} = \langle X| \left(\sum_{i=1}^{N} \mathrm{Tr}[\rho_{\mathcal{E}} P_i] |D_i\rangle\langle D_i| \right) |X\rangle - \overline{\langle X\rangle^2}_{\mathcal{E}}. \tag{14}$$

On the other hand, one can prove the following identity

$$\Gamma^\dagger\pi\Gamma\Lambda\pi^{-1}\Lambda^\dagger = \Gamma^\dagger\Lambda^\dagger\Gamma^\dagger\Lambda^\dagger = \Gamma^\dagger\Lambda^\dagger = \sum_{i=1}^{N} |D_i\rangle\langle P_i| = I_{\mathcal{S}}. \tag{15}$$

This implies that $\Gamma^\dagger\pi\Gamma = (\Lambda\pi^{-1}\Lambda^\dagger)^{-1}$, and finally, one can express the minimum noise in terms of the POVM and the ensemble only, as follows

$$\delta_D(X) = \langle X|(\Lambda\pi^{-1}\Lambda^\dagger)^{-1}|X\rangle - \overline{\langle X\rangle^2}_{\mathcal{E}}. \tag{16}$$

The optimal dual has been obtained in a completely different framework in [8] in the particular case when $\rho_{\mathcal{E}} = \frac{I}{d}$, and the figure of merit considered therein is the Hilbert-Schmidt distance between the estimated state and the true state.

4. CONCLUSION

In this paper we reviewed the problem of estimation of ensemble averages of operators by indirect measurements, through a fixed measurement whose statistics is described by a POVM $\{P_i\}$. The coefficients for the expansion of of an operator on the POVM elements provide the processing functions, and their calculation is possible in principle by using elementary results in frame theory. The difficult problem is to decide which processing function is the best in order to minimize the statistical error in the estimation of the ensemble average. We restated this problem as the inversion of a rectangular matrix with the constraint of minimum norm. We reviewed the general solution derived in Ref. [2], and we present a synthetic formula for the evaluation of the minimum noise in terms of the POVM elements and the input ensemble.

ACKNOWLEDGMENTS

P. P. thanks A. J. Scott for an interesting exchange of e-mails, which inspired Eq. (16). This work has been supported by Ministero Italiano dell'Università e della Ricerca (MIUR) through PRIN 2005. P. P. acknowledges financial support by EC under project SECOQC (contract n. IST-2003-506813)

REFERENCES

[1] J. Von Neumann, *Mathematical Principles of Quantum Mechanics*, (Princeton University Press, Princeton, 1955).

[2] G. M. D'Ariano and P. Perinotti, Phys. Rev. Lett. **98**, 020403 (2007).

[3] R. J. Duffin and A. C. Schaeffer, Trans. Am. Math. Soc. **72**, 341 (1952); P. G. Casazza, Taiw. J. Math. **4**, 129 (2000).

[4] P. Casazza, D. Han, and D. R. Larson, Contemp. Math. **247**, 149 (1999).

[5] R. B. Bhapat, *Linear Algebra and Linear Models*, (Springer-Verlag, New York, 2000).

[6] S. Li, Numer. Funct. Anal. Optim. **16**, 1181 (1995).

[7] G. M. D'Ariano, P. Perinotti, and M. F. Sacchi, Phys. Rev. A **72**, 042108 (2005).

[8] A. J. Scott, J. Phys. A **39**, 13507 (2006).