von Neumann measurement is optimal for detecting linearly independent mixed quantum states

Yonina C. Eldar*
Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel

(Received 3 June 2003; published 3 November 2003)

We consider the problem of designing a measurement to minimize the probability of a detection error when distinguishing between a collection of possibly nonorthogonal mixed quantum states. We show that if the quantum state ensemble consists of linearly independent density operators, then the optimal measurement is an orthogonal von Neumann measurement consisting of mutually orthogonal projection operators, and not a more general positive operator-valued measure.

DOI: 10.1103/PhysRevA.68.052303 PACS number(s): 03.67.Hk, 03.65.Ta

I. INTRODUCTION

One of the important features of quantum mechanics is that nonorthogonal quantum states cannot be perfectly distinguished [1]. Therefore, a fundamental problem in quantum mechanics is to design measurements optimized to distinguish between a collection of nonorthogonal quantum states.

We consider a quantum state ensemble consisting of m density operators $\{\rho_i, 1 \leqslant i \leqslant m\}$ on an n-dimensional complex Hilbert space $\mathcal{H}$, with prior probabilities $\{p_i > 0, 1 \leqslant i \leqslant m\}$. A density operator ρ is a positive semidefinite (PSD) Hermitian operator with $\mathrm{Tr}(\rho)=1$; we write $\rho \geqslant 0$ to indicate ρ is PSD. A mixed state ensemble is one in which at least one of the density operators ρ_i has rank larger than one. A pure-state ensemble is one in which each density operator ρ_i is rank-one projector $|\phi_i\rangle\langle\phi_i|$, where the vectors $|\phi_i\rangle$, though evidently normalized to unit length, are not necessarily orthogonal.

In a quantum detection problem a transmitter conveys classical information to a receiver using a quantum-mechanical channel. Each message is represented by preparing the quantum channel in one of the ensemble states ρ_i. At the receiver, the information is detected by subjecting the channel to a quantum measurement in order to determine the state prepared. If the quantum states are mutually orthogonal, then the state can be determined correctly with probability one by performing an optimal *von Neumann measurement* [1]. A von Neumann measurement consists of m mutually orthogonal projection operators $\{\Pi_i, 1 \leqslant i \leqslant m\}$ that form a resolution of the identity on $\mathcal{H}$, so that

$$\Pi_i \Pi_j = \delta_{ij} \Pi_i, \quad 1 \leqslant i,j \leqslant m, \tag{1}$$

$$\sum_{i=1}^{m} \Pi_i = I_n, \tag{2}$$

where I_n is the identity operator on $\mathcal{H}$.

If the given states are not orthogonal, then no measurement will distinguish perfectly between them. Our problem is therefore to construct a measurement that minimizes the probability of a detection error. It is well known that the most efficient way of obtaining information about the state of a quantum system is not always by performing orthogonal projections [2,3], but rather by performing more general *positive operator-valued measures* (POVMs). A POVM consists of m PSD Hermitian operators $\{\Pi_i, 1 \leqslant i \leqslant m\}$ that form a resolution of the identity on $\mathcal{H}$ but are not constrained to be projection operators, so that

$$\Pi_i \geqslant 0, \quad 1 \leqslant i \leqslant m,$$

$$\sum_{i=1}^{m} \Pi_i = I_n. \tag{3}$$

Necessary and sufficient conditions for an optimum measurement maximizing the probability of correct detection have been developed [4–6]. However, in general obtaining a closed form expression for the optimal measurement directly from these conditions is a difficult and unsolved problem. Closed-form analytical expressions for the optimal measurement have been derived for several special cases [7–12].

In order to physically realize a POVM, we can rely on Neumark's theorem [2], which states that the space $\mathcal{H}$ can be extended to a larger space, such that the POVM operators $\Pi_i = P A_i P$ are projections of a set of von Neumann measurement operators A_i onto $\mathcal{H}$, where P is the orthogonal projection onto $\mathcal{H}$. The physical interpretation given to Neumark's theorem is that the larger space is obtained by combining an ancilla system with the given known system. Thus, in practice, to realize a POVM, the original system must be enlarged, which in some cases, may be difficult to do in practice.

Kennedy [13] showed that for a pure state ensemble consisting of rank-one density operators $\rho_i = |\phi_i\rangle\langle\phi_i|$ with linearly independent vectors $|\phi_i\rangle$, the optimal measurement maximizing the probability of correct detection is a von Neumann measurement consisting of mutually orthogonal rank-one projection operators. Thus, in this case, an enlargement of the original system is not necessary. However, this implication has not been proven for the more general case of *mixed* state ensembles.

In Sec. III we show that the optimal measurement for distinguishing between a set of linearly independent mixed quantum states is a von Neumann measurement and not a general POVM. Therefore, in this case, when seeking the optimal measurement, we may restrict our attention to the

*URL: http://www.ee.technion.ac.il/Sites/People/YoninaEldar/; electronic address: yonina@ee.technion.ac.il

class of von Neumann measurements. We also show that the rank of each projection operator is equal to the rank of the corresponding density operator.

In Sec. IV we consider the least-squares measurement (LSM) [11], also known as the square-root measurement [14,15], which is a detection measurement that has many desirable properties and has been employed in many settings. We show that for linearly independent mixed state ensembles the LSM reduces to a von Neumann measurement.

In the next section we present our detection problem and summarize results from [6] pertaining to the conditions on the optimal measurement operators.

II. OPTIMAL DETECTION OF QUANTUM STATES

Assume that a quantum channel is prepared in a quantum state drawn from a collection of given states represented by density operators $\{\rho_i, 1\leqslant i\leqslant m\}$ in an n-dimensional complex Hilbert space $\mathcal{H}$. We assume without loss of generality that the eigenvectors of ρ_i, $1\leqslant i\leqslant m$, collectively span [20] $\mathcal{H}$.

Since each density operator ρ_i is Hermitian and PSD, it can be expressed via the eigendecomposition as $\rho_i=\phi_i\phi_i^*$ where ϕ_i is an $n\times r_i$ matrix of orthogonal eigenvectors $\{|\phi_{ik}\rangle, 1\leqslant k\leqslant r_i\}$ and $r_i=\mathrm{rank}(\rho_i)$. The density operators ρ_i, $1\leqslant i\leqslant m$ are *linearly independent* if the eigenvectors $\{|\phi_{ik}\rangle, 1\leqslant k\leqslant r_i, 1\leqslant i\leqslant m\}$ form a linearly independent set of vectors. Since the eigenvectors of ρ_i, $1\leqslant i\leqslant m$ collectively span the n-dimensional space $\mathcal{H}$, it follows that for linearly independent state sets

$$\sum_{i=1}^{m} r_i=n. \tag{4}$$

At the receiver, the constructed measurement comprises m measurement operators $\{\Pi_i, 1\leqslant i\leqslant m\}$ on $\mathcal{H}$ that satisfy (3). We seek the measurement operators $\{\Pi_i, 1\leqslant i\leqslant m\}$ satisfying (3) that maximize the probability of correct detection, which is given by

$$P_d=\sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i\Pi_i), \tag{5}$$

where $p_i>0$ is the prior probability of ρ_i, with $\Sigma_i p_i=1$.

It was shown in [5,6] that a set of measurement operators $\{\hat{\Pi}_i, 1\leqslant i\leqslant m\}$ maximize the probability of correct detection for a state set $\{\rho_i, 1\leqslant i\leqslant m\}$ with prior probabilities $\{p_i, 1\leqslant i\leqslant m\}$ if and only if there exists an Hermitian $\hat{X}$ satisfying

$$\hat{X}\geqslant p_i\rho_i, \quad 1\leqslant i\leqslant m, \tag{6}$$

such that

$$(\hat{X}-p_i\rho_i)\hat{\Pi}_i=0, \quad 1\leqslant i\leqslant m. \tag{7}$$

The matrix $\hat{X}$ can be determined as the solution to the problem

$$\min_{X\in\mathcal{B}} \mathrm{Tr}(X), \tag{8}$$

where $\mathcal{B}$ is the set of Hermitian operators on $\mathcal{H}$, subject to

$$X\geqslant p_i\rho_i, \quad 1\leqslant i\leqslant m. \tag{9}$$

As shown in [6], the conditions (6) and (7) together imply that

$$t_i\leqslant r_i, \tag{10}$$

where $t_i=\mathrm{rank}(\hat{\Pi}_i)$.

Kennedy [13] showed that for pure state ensembles $\rho_i=|\phi_i\rangle\langle\phi_i|$ with linearly independent vectors $|\phi_i\rangle$ the optimal measurement is a rank-one measurement $\hat{\Pi}_i=|\mu_i\rangle\langle\mu_i|$ with orthonormal vectors $|\mu_i\rangle$, i.e., a von Neumann measurement. However, this implication has not been proven for mixed states. In the following section we use the conditions for optimality to prove that the optimal measurement for linearly independent mixed states is a von Neumann measurement and not a more general POVM.

III. LINEARLY INDEPENDENT STATE ENSEMBLES

Suppose now that the density operators ρ_i are linearly independent, and let $\hat{\Pi}_i$ be the optimal measurement operators that maximize (5) subject to (3). Denoting $\Pi=\Sigma_{i=1}^m\hat{\Pi}_i$ we have that

$$\mathrm{rank}(\Pi)\leqslant\sum_{i=1}^{m}\mathrm{rank}(\hat{\Pi}_i)=\sum_{i=1}^{m} t_i. \tag{11}$$

Since $\Pi=I_n$ we also have

$$\mathrm{rank}(\Pi)=n, \tag{12}$$

from which we conclude that

$$\sum_{i=1}^{m} t_i\geqslant n. \tag{13}$$

Combining (13) with (10) and (4) we conclude that

$$t_i=r_i. \tag{14}$$

Therefore, via the eigendecomposition we can express each measurement operator $\hat{\Pi}_i$ as $\hat{\Pi}_i=\mu_i\mu_i^*$, where μ_i is an $n\times r_i$ matrix of orthogonal eigenvectors $\{|\mu_{ik}\rangle, 1\leqslant k\leqslant r_i\}$. Since $\Sigma_{i=1}^m r_i=n$ we have n vectors $|\mu_{ik}\rangle$. In addition, from Eq. (3),

$$\sum_{ik} |\mu_{ik}\rangle\langle\mu_{ik}|=I_n \tag{15}$$

from which we conclude that the vectors $\{|\mu_{ik}\rangle, 1\leqslant k\leqslant r_i, 1\leqslant i\leqslant m\}$ are linearly independent.

We now show that the vectors $\{|\mu_{ik}\rangle, 1\leqslant k\leqslant r_i, 1\leqslant i\leqslant m\}$ are mutually orthonormal. From Eq. (3) we have that for any $1\leqslant l\leqslant r_i$, $1\leqslant j\leqslant m$,

$$|\mu_{jl}\rangle=\sum_{ik}\langle\mu_{ik}|\mu_{jl}\rangle|\mu_{ik}\rangle. \tag{16}$$

Since the vectors $|\mu_{ik}\rangle$ are linearly independent, we must have that $\langle\mu_{ik}|\mu_{jl}\rangle=\delta_{ij,kl}$.

We conclude that

$$\hat{\Pi}_i=\sum_{k=1}^{r_i}|\mu_{ik}\rangle\langle\mu_{ik}|=P_{\mathcal{S}_i}, \tag{17}$$

where $P_{\mathcal{S}_i}$ is an orthogonal projection onto a subspace $\mathcal{S}_i$ of $\mathcal{H}$ with dimension r_i and

$$P_{\mathcal{S}_i}P_{\mathcal{S}_j}=\delta_{ij}P_{\mathcal{S}_i}, \tag{18}$$

so that $\mathcal{H}=\mathcal{S}_1\oplus\cdots\oplus\mathcal{S}_m$ is the direct sum of the subspaces $\mathcal{S}_i$. Thus, the optimal measurement is a von Neumann measurement. We summarize our results in the following theorem:

Theorem 1. Let $\{\rho_i, 1\leqslant i\leqslant m\}$ be a quantum state ensemble consisting of linearly independent density operators ρ_i with prior probabilities $p_i>0$. Then the optimal measurement is a von Neumann measurement with measurement operators $\{\hat{\Pi}_i=P_{\mathcal{S}_i}, 1\leqslant i\leqslant m\}$ where $P_{\mathcal{S}_i}$ is an orthogonal projection onto an r_i-dimensional subspace $\mathcal{S}_i$ of $\mathcal{H}$ with $r_i=\mathrm{rank}(\rho_i)$ and $P_{\mathcal{S}_i}P_{\mathcal{S}_j}=\delta_{ij}P_{\mathcal{S}_i}$.

Examples of linearly independent states

We now consider several examples of linearly independent quantum states. A state set that is often used in the context of optical communications is the m-array phase-shift-keyed coherent state set [4,8]. The state set consists of the coherent states $|\phi_k\rangle$, where

$$\phi_i=\mu u^i, \quad 0\leqslant i\leqslant m-1, \tag{19}$$

with

$$u=e^{-j(2\pi/m)}, \quad 0\leqslant i\leqslant m-1. \tag{20}$$

The amplitude $\mu>0$, and $s=\mu^2$ is the mean number of transmitted photons, or the intensity of each signal.

We now show that the states defined by (19) are linearly independent. To this end it suffices to show that the Gram matrix of inner products with the ikth element given by $\langle\phi_i|\phi_k\rangle$ is invertible, or equivalently, has eigenvalues $\lambda_i>0$. It was shown in [8] that

$$\lambda_i=me^{-s}\sum_{k=0}^{\infty}\frac{s^{i+nm}}{(i+nm)!}, \quad 1\leqslant i\leqslant m, \tag{21}$$

so that clearly $\lambda_i>0$. We therefore conclude that the m-array phase-shift-keyed coherent state set is linearly independent.

Another example of a linearly independent state set is the state set considered by Peres and Wootters in [3]. Specifically, they consider the case of two noninteracting spin-$\frac{1}{2}$ particles, prepared with the same polarization, where there are three possible preparations: Both spins may be directed along the z direction, or both may be in the x-z plane, tilted at 120° or −120° from the z axis. In this problem the states to be distinguished are given by $|\phi_1\rangle=|aa\rangle$, $|\phi_2\rangle=|bb\rangle$, and $|\phi_3\rangle=|cc\rangle$, where $|a\rangle$, $|b\rangle$, and $|c\rangle$ correspond to polarizations of a photon at 0°, 120°, and −120°. For this state set,

$$\langle\phi_1|\phi_2\rangle=\langle\phi_1|\phi_3\rangle=\langle\phi_2|\phi_3\rangle=-\tfrac{1}{2}, \tag{22}$$

so that the Gram matrix of inner products is invertible, and the state set is linearly independent.

We note that both of the state sets considered above are special cases of geometrically uniform states [11] (assuming equal prior probabilities). The optimal measurement for distinguishing these state sets is derived in [11,12] and, as we expect, is a von Neumann measurement.

Both of the examples above are examples of pure state sets. From each pure state set we can generate a mixed state set by choosing one of a subset of the possible state vectors with a certain prior probability. For example, suppose we generate a state set in which the first state is either μu^1 or μu^2 with probability p and $1-p$, respectively, for some $p<1$, and the second state is either μu^3 or μu^4 with probability q and $1-q$, respectively, for some $q<1$. Since the states μu^i are linearly independent, any subset of the states is also linearly independent, which implies that the generated mixed state set is linearly independent.

IV. LEAST-SQUARES MEASUREMENT

A suboptimal measurement that has been employed as a detection measurement in many applications is the least-squares measurement (LSM) [11,16], also known as the square-root measurement [10,14,15,17–19]. The LSM has many desirable properties. Its construction is relatively simple; it can be determined directly from the given collection of states; it minimizes the probability of a detection error for pure and mixed state ensembles that exhibit certain symmetries [6,11]; it is "pretty good" when the states to be distinguished are equally likely and almost orthogonal [14]; and it is asymptotically optimal [15,16].

The LSM corresponding to a set of density operators $\{\rho_i=\phi_i\phi_i^*, 1\leqslant i\leqslant m\}$ with eigenvectors that collectively span $\mathcal{H}$ and prior probabilities $\{p_i, 1\leqslant i\leqslant m\}$ consists of the measurement operators $\{\Sigma_i=\mu_i\mu_i^*, 1\leqslant i\leqslant m\}$ where [11,16]

$$\mu_i=(\Psi\Psi^*)^{-1/2}\psi_i. \tag{23}$$

Here

$$\Psi=[\psi_1 \;\; \psi_2 \;\; \cdots \;\; \psi_m] \tag{24}$$

is the matrix of (block) columns $\psi_i=\sqrt{p_i}\phi_i$, and $(\cdot)^{1/2}$ is the unique Hermitian square root of the corresponding matrix. Note that since the eigenvectors of the $\{\rho_i\}$ collectively span $\mathcal{H}$, the columns of the $\{\psi_i\}$ also together span $\mathcal{H}$, so $\Psi\Psi^*$ is invertible.

We now show that the LSM satisfies the conditions of Theorem 1 so that if the columns of $\{\phi_i\}$ are linearly independent, then $\Sigma_i\Sigma_j=\Sigma_i\delta_{ij}$ and the LSM is a von Neumann measurement.

From Eq. (23) we have that

$$\Sigma_i\Sigma_j=(\Psi\Psi^*)^{-1/2}\psi_i\psi_i^*(\Psi\Psi^*)^{-1}\psi_j\psi_j^*(\Psi\Psi^*)^{-1/2}. \quad (25)$$

To simplify Eq. (25) we express ψ_i as

$$\psi_i=\Psi E_i. \quad (26)$$

Here E_i is an $n\times r_i$ matrix where the qth column of E_i has one nonzero element equal to 1 in the pth position with $p=\Sigma_{k=1}^{i-1}r_k+q$. We then have that

$$\psi_i^*(\Psi\Psi^*)^{-1}\psi_j=E_i^*\Psi^*(\Psi\Psi^*)^{-1}\Psi E_j. \quad (27)$$

If the density operators $\{\rho_i\}$ are linearly independent, then $\Sigma_i r_i=n$ and the operators $\{\psi_i\}$ are also linearly independent. Since each matrix ψ_i has dimension $n\times r_i$ we conclude that Ψ is an $n\times n$ matrix with linearly independent columns and is therefore invertible. Thus, $\Psi^*(\Psi\Psi^*)^{-1}\Psi=I_n$ and

$$\psi_i^*(\Psi\Psi^*)^{-1}\psi_j=E_i^*E_j=\delta_{ij}I. \quad (28)$$

Substituting Eq. (28) into Eq. (25),

$$\Sigma_i\Sigma_j=\delta_{ij}(\Psi\Psi^*)^{-1/2}\psi_i\psi_i^*(\Psi\Psi^*)^{-1/2}=\delta_{ij}\Sigma_i \quad (29)$$

and the LSM is a von Neumann measurement consisting of mutually orthogonal projection operators. We summarize our results regarding the LSM in the following theorem.

Theorem 2. Let $\{\rho_i, 1\leq i\leq m\}$ be a quantum state ensemble consisting of linearly independent density operators ρ_i with prior probabilities $p_i>0$. Then the least-squares measurement is a von Neumann measurement with measurement operators $\{\Sigma_i=P_{\mathcal{S}_i}, 1\leq i\leq m\}$ where $P_{\mathcal{S}_i}$ is an orthogonal projection onto an r_i-dimensional subspace $\mathcal{S}_i$ of $\mathcal{H}$ with $r_i=\text{rank}(\rho_i)$ and $P_{\mathcal{S}_i}P_{\mathcal{S}_j}=\delta_{ij}P_{\mathcal{S}_i}$.

We note that in many cases the LSM minimizes the probability of a detection error, both for linearly independent states and for linearly dependent states [11,12]. A sufficient condition on the state sets under which the LSM is optimal is given in [12]. It is also shown that state sets that exhibit a wide class of symmetry properties satisfy this condition and therefore, for such state sets, the LSM is optimal. However, there are also cases in which the LSM is not optimal; for an example, see [6]. An interesting direction for future research is to derive *necessary* and sufficient conditions for the LSM to be optimal.

ACKNOWLEDGMENTS

The author acknowledges the Taub Foundation for financial support.

[1] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Boston, 1995).
[2] A. Peres, Found. Phys. **20**, 1441 (1990).
[3] A. Peres and W. K. Wootters, Phys. Rev. Lett. **66**, 1119 (1991).
[4] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[5] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).
[6] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **49**, (2003).
[7] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[8] M. Charbit, C. Bendjaballah, and C. W. Helstrom, IEEE Trans. Inf. Theory **35**, 1131 (1989).
[9] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).
[10] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[11] Y. C. Eldar and G. D. Forney, Jr., IEEE Trans. Inf. Theory **47**, 858 (2001).
[12] Y. C. Eldar, A. Megretski, and G. C. Verghese, e-print quant-ph/0211111.
[13] R. S. Kennedy, MIT Research Laboratory Electronic Quarterly Progress Report, Technical Report No. 110, 1973.
[14] P. Hausladen and W. K. Wootters, J. Mod. Opt. **41**, 2385 (1994).
[15] P. Hausladen, R. Josza, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[16] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[17] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 146 (1998).
[18] M. Sasaki, T. Sasaki-Usuda, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 159 (1998).
[19] K. Kato, M. Osaki, M. Sasaki, and O. Hirota, IEEE Trans. Commun. **47**, 248 (1999).
[20] Otherwise we can transform the problem to a problem equivalent to the one considered here by reformulating the problem on the subspace spanned by the eigenvectors of $\{\rho_i, 1\leq i\leq m\}$.