Comparison of error probability bounds in quantum state discrimination

Roberto Corvaja[*]
Department of Information Engineering, University of Padova, Via G. Gradenigo 6/B - 35131 Padova, Italy
(Received 2 January 2013; revised manuscript received 19 February 2013; published 23 April 2013)

In quantum discrimination, the value of the minimum error probability and the set of measurement operators which achieve this minimum are often difficult to derive. Here we present a comparison of the performance obtained by the optimal solution and by the available bounds, namely the square root measurement (SRM) and the Chernoff bound. Applied to some Gaussian states, namely to coherent states with thermal noise, it is shown that the SRM provides a much tighter bound with respect to the Chernoff bound, with a comparable numerical complexity.

DOI: 10.1103/PhysRevA.87.042329 PACS number(s): 03.67.Hk

I. INTRODUCTION

In the discrimination of quantum states, not only is it difficult to derive the set of measurement operators achieving the minimum error probability, but in many cases it is difficult to derive also the actual value of the error probability. Closed-form expressions can be obtained only in the cases of pure states with high-symmetry properties. Also, Helstrom's bound for the binary case, in the general case of mixed quantum states, requires a singular value decomposition for the determination of the positive eigenvalues of the difference matrix (Helstrom matrix) needed for the evaluation of the error probability. In many other cases, the optimal measurement set is not known and only numerical solutions are available, requiring one to resort to a heavy convex optimization problem [1] to determine the measurement operators.

However, suboptimal bounds can be derived, in particular the square root measurement (SRM) [2,3], which obtains the set of measurement operators from the Gram operator or from the Gram matrix, and the quantum Chernoff bound, which recently received a great deal of attention, especially for Gaussian quantum states [4], as a simple way to estimate the performance of quantum discrimination [5–7]. Applied to Gaussian states, other bounds are derived in [8] for binary hypothesis testing by fixing one of the conditional error probabilities and minimizing the other conditional error probability.

The Chernoff quantum is limited in that it can only be applied to binary quantum systems, and in this work we compare the SRM and the Chernoff bounds with the optimum error probability in terms of both performance gap and complexity. In fact, we notice that several studies in the literature considered the bounds *separately* and when possible evaluated their relation to the optimum value. In this paper, we will make a systematic comparison of the two bounds. It can be seen that, applied to coherent states with thermal noise, the SRM provides a tighter bound to the error probability than the Chernoff bound with comparable computational complexity. Also, the SRM is available easily also for the M-ary case and has the additional advantage of providing the optimal solution when the states are pure and exhibit geometrical uniform symmetry [9].

[*]corvaja@dei.unipd.it

II. QUANTUM DISCRIMINATION

A general M-ary quantum system with mixed states is described by M density operators $\{\rho_0,\ldots,\rho_{M-1}\}$ in an N-dimensional Hilbert space $\mathcal{H}$, where N may possibly be infinite. The eigenvalues of the operators span a subspace $\mathcal{U}\subseteq\mathcal{H}$.

Quantum discrimination is the operation of choosing among the possible *density operators* ρ_i, $i=0,1,\ldots,M-1$, performed by applying a positive operator valued measurement (POVM) set, that is, a set of M positive semidefinite operators $\Pi_0,\ldots,\Pi_{M-1}$ with the condition

$$\sum_{i=0}^{M-1}\Pi_i = P_{\mathcal{U}}, \tag{1}$$

where $P_{\mathcal{U}}$ is the projector operator onto $\mathcal{U}$. In other words, Π_i must give a resolution of the identity in the subspace $\mathcal{U}$. The probability that the detection outcome is j, provided that the density operator is ρ_i, i.e., the transition probability $p(j|i)$, is given by

$$p(j|i)=\mathrm{Tr}(\rho_i\Pi_j),\quad i,j=0,1,\ldots,M-1, \tag{2}$$

and the corresponding error probability in the detection becomes

$$P_e = 1-\sum_{i=0}^{M-1} q_i\, p(i|i) = 1-\sum_{i=0}^{M-1} q_i\,\mathrm{Tr}(\rho_i\Pi_i), \tag{3}$$

where q_i, $i=0,\ldots,M-1$, are the *a priori* probabilities.

For the binary case, the optimum solution is available by the decomposition of the difference operator $D=q_1\rho_1-q_0\rho_0$ [10], obtaining the Helstrom bound [11] for the minimum error probability achievable,

$$P_e = q_1-\sum_{\eta_k>0}\eta_k, \tag{4}$$

where η_k are the eigenvalues of D and the sum extends over the positive ones.

In the following, for simplicity, we assume that all the states are equiprobable, that is, $q_i=1/M$, $i=0,\ldots,M-1$.

Apart for the binary case, the optimal detection set of POVM can be obtained by convex semidefinite programming (CSP) [9], while suboptimal solutions are achievable by square root measurement (SRM) [3] or by the Chernoff bound, both giving an upper bound to the error probability. The conditions for the optimum POVM set in [12,13] lead to a *convex*

 ©2013 American Physical Society

semidefinite problem with an equivalent "dual problem" as follows [9]. Given the M density operators ρ_i, find the positive semidefinite Hermitian operator Y with the constraint $Y \geqslant \rho_i$ for each i, such that $\mathrm{Tr}(Y)$ is minimum. The minimum trace gives the maximum correct decision probability, and by the equations $(Y - \rho_i)\Pi_i = 0$ one gets the optimal Π_i.

III. SQUARE ROOT MEASUREMENT

A more straightforward, albeit not optimal, approach to the problem is given by the SRM. The approach has been proposed for pure states by Hausladen *et al.* [2] and thoroughly investigated by Eldar and Forney [3]. The generalization to mixed states is due to Eldar *et al.* [9]. Cariolaro and Pierobon [14,15] applied the SRM technique systematically to evaluate the performance of the most popular quantum communication systems. In [3] it is thoroughly shown that the underlying principle of SRM is the least mean square (LMS), in which the set of orthogonal measurement vectors is sought which minimize the quadratic error with respect to the set of state vectors.

The SRM approach is based on the *Gram operator* $\boldsymbol{T} = \boldsymbol{\Gamma}\,\boldsymbol{\Gamma}^*$ and the *Gram matrix* $\boldsymbol{G} = \boldsymbol{\Gamma}^*\,\boldsymbol{\Gamma}$. These matrices usually make reference to pure states $|\gamma_i\rangle$, where the density operators are $\rho_i = |\gamma_i\rangle\langle\gamma_i|$, but they can be extended to the mixed states, the key being the factorization of each density operator in the form $\rho_i = \gamma_i\,\gamma_i^*$, e.g., via the eigendecomposition of ρ_i. If ρ_i has rank k_i, the *state factor* (briefly *state*) γ_i can be chosen with dimensions $N \times k_i$. Hence, the collection of the M state factors as block columns gives the (generalized) state matrix $\boldsymbol{\Gamma} = [\gamma_0, \ldots, \gamma_{m-1}]$. The ith measurement operator with rank not greater than the rank of ρ_i is written in the form $\Pi_i = \mu_i\,\mu_i^*$, where μ_i are $N \times k_i$ matrices. The M *measurement factors* μ_i, collected as block columns, give the (generalized) measurement matrix $\boldsymbol{M} = [\mu_0, \ldots, \mu_{M-1}]$. Both $\boldsymbol{\Gamma}$ and $\boldsymbol{M}$ have dimensions $N \times r$ with $r = k_0 + \cdots + k_{M-1}$.

In SRM, the measurement matrix $\boldsymbol{M}$ is given by the two equivalent expressions [3]

$$\boldsymbol{M} = \boldsymbol{T}^{-1/2}\,\boldsymbol{\Gamma}, \quad \boldsymbol{M} = \boldsymbol{\Gamma}\,\boldsymbol{G}^{-1/2}, \tag{5}$$

where $\boldsymbol{T}^{-1/2}$ and $\boldsymbol{G}^{-1/2}$ are the inverse square roots (in the Moore-Penrose generalized sense) of $\boldsymbol{T}$ and $\boldsymbol{G}$. From the measurement matrix $\boldsymbol{M}$, the operators $\Pi_i = \mu_i\,\mu_i^*$ give the transition probabilities as

$$p(j|i) = \mathrm{Tr}(\rho_i \Pi_j) \tag{6}$$

and the error probability as (3), namely

$$P_e = 1 - \sum_{i=0}^{M-1} q_i \mathrm{Tr}(\gamma_i\,\gamma_i^* \mu_i \mu_i^*). \tag{7}$$

IV. QUANTUM CHERNOFF BOUND

The Chernoff bound is usually employed in telecommunications and probability theory to establish an upper bound to the error probability [16] or more in general to bound the probability that a random variable exceeds a certain quantity, based on the knowledge of the characteristic function or of the moments of the random variable. Its application in the case of classical hypothesis testing (H_0 and H_1) is considered in [17], where the maximum *a posteriori* probability decision rule is applied to a sequence of n random variables with conditional probability density functions $f_0(a)$ and $f_1(a)$. The Chernoff bound on the error probability over n attempts $P_e(n)$ is expressed by

$$\frac{1}{n} \ln P_e(n) \leqslant \ln \int f_0^s(a)\, f_1^{1-s}(a)\, da, \tag{8}$$

with $0 \leqslant s \leqslant 1$. Since (8) holds true for any value of s, one can take its minimum over $0 \leqslant s \leqslant 1$. The extension of the Chernoff bound to quantum systems, leading to the *quantum Chernoff bound*, is considered in several works [4–7,18], employing the bound as a tool to estimate the error probability in the discrimination of quantum states, both for single-mode and multimode states. The Chernoff bound can be seen also as a distance measure between operators. The Chernoff distance has been investigated, for example, in [5–7] and related to other distinguishability measures, such as the fidelity.

For the binary case $M = 2$, where the states are described by the density operators ρ_0 and ρ_1, the quantum Chernoff bound states that error probability can be bounded by the expression

$$P_e \leqslant \frac{1}{2} \inf_{0 \leqslant s \leqslant 1} \mathrm{Tr}\left[\rho_0^s \rho_1^{1-s}\right]. \tag{9}$$

The bound requires the minimization with respect to the real value s. Note, however, that when the Gaussian states have the same covariance matrix or the same thermal noise component and no relative displacement, the optimum is attained [7] for $s = 1/2$. Note that in this case the square root of the density operators must be evaluated and the bound becomes

$$P_e \leqslant \frac{1}{2}\mathrm{Tr}[\sqrt{\rho_0}\sqrt{\rho_1}], \tag{10}$$

derived as Lemma 3.2 in [10]. This bound is also called the quantum Bhattacharyya bound [19]. Also in the application example considered in the following of this paper, the conditions leading to a minimum value for $s = 1/2$ hold. In general, a closed-form expression of (9) is not available even in the case of Gaussian states. Only in the case of GUS with pure states can a factorization of the operators be derived by means of the Fourier matrices, as was done in [15].

V. COMPLEXITY COMPARISON

In terms of computational complexity, the evaluation of the optimum POVM set requires a convex constrained optimization procedure, where the optimum operator of size $N \times N$ is searched. The optimization software available to solve this kind of problem is very sophisticated and typically utilizes iterative interior-point methods. Note that convex optimization algorithms convert the matrix constraints into a larger search matrix, so that for the optimization of measurement operators of size $N \times N$, dimensions of the order of $MN \times MN$ are obtained. In any case, the algorithms are typically iterative with a precision parameter determining the stop condition, so that the numerical complexity can vary a lot not only due to the optimization package, but also to the required accuracy. In terms of execution time, we could see that this procedure requires a time which could be orders of magnitude longer than that required by the SRM approach or the Chernoff bound.

The quantum Chernoff bound requires the evaluation of the fractional power of a matrix of size $N \times N$ for all the values of the minimization parameter s. This requires an eigendecomposition of the kind

$$\rho_i = U_i \Lambda_i U_i^* \longrightarrow \rho_i^s = U_i \Lambda_i^s U_i^*. \tag{11}$$

Therefore, apart from the cases in which the value of s corresponding to the minimum is known, each value of s evaluated for minimization requires us to evaluate the fractional power of $N \times N$ matrices, with complexity of the order $O(N^3)$.

The SRM requires an evaluation of a matrix factorization for each of the density operators ρ_i, therefore again M singular value decompositions of matrices of size $N \times N$ as (11), each with complexity $O(N^3)$, followed by the square root of the Gram matrix or the Gram operator, with size r, where $r < N$, and associated complexity $O(r^3)$. Moreover, in some cases such as the case in which the states have geometrically uniform symmetry (GUS) [9], the complexity could be greatly reduced by working in an equivalent compressed space [20].

Therefore, in the binary case, where sometimes also the closed-form optimal solution can be obtained with limited complexity, the derivation of the SRM solution has a computational complexity comparable to or lower than the evaluation of the quantum Chernoff bound, and in the general M-ary cases the quantum Chernoff bound cannot be applied, while the SRM has a limited complexity compared to the optimal solution.

VI. PERFORMANCE

We consider as an application example a quantum optical communication system using coherent states with *phase modulation* [phase shift keying (PSK)]. This modulation consists in preparing the state by a choice among M coherent states, generated according to the constellation of M complex numbers (complex envelopes),

$$\alpha_i = |\alpha| e^{j2\pi \frac{i}{M}}, \quad i = 0, \ldots, M-1. \tag{12}$$

In this modulation, all the states have the same average number of photons per symbol, $|\alpha_i|^2 = |\alpha|^2 = N_\alpha$.

A. Density operators from Glauber's theory

According to the well known theory of Glauber [21], a *coherent state*, representing the monochromatic radiation of a laser in the presence of thermal noise, is an element of an infinite-dimensional Hilbert space. The corresponding density operator $\rho(\alpha)$ has the Fock representation given by

$$\rho_{mn}(\alpha) = \frac{v^n e^{-\frac{|\alpha|^2}{\mathcal{N}+1}}}{\mathcal{N}+1} \sqrt{\frac{m!}{n!}} \left(\frac{\alpha^*}{\mathcal{N}}\right)^{n-m} L_m^{n-m}\left(\frac{(v-1)|\alpha|^2}{\mathcal{N}}\right), \quad m,n = 0,1,2,\ldots, \tag{13}$$

where α is the *signal* complex envelope, $N_\alpha = |\alpha|^2$ gives the average number of *signal photons*, $\mathcal{N}$ is the average number of *thermal noise photons*, $v = \mathcal{N}/(1+\mathcal{N})$, and $L_m^k(x)$ are the generalized Laguerre polynomials. For PSK, the value of the complex number α is chosen in the set (12).

For practical calculations we need a finite approximation of the matrix (13) to N terms, where N can be chosen according to the *quasiunitary trace criterion* [14] as the smallest integer such that $\sum_{m=0}^{N} \rho_{mm}(\alpha) \geqslant 1 - \epsilon$ for a given accuracy ϵ. Once N is established, the EID of $\rho(\alpha)$ gives the factorization

$$\rho(\alpha) = \gamma(\alpha)\gamma^*(\alpha) \tag{14}$$

with $\gamma(\alpha)$ of dimension $N \times r$, where r is the rank. An investigation into these approximations was performed in [14], where the size N is related to the range of error probabilities P_e that one needs to investigate, showing that the rule $\epsilon = P_e/10$ is largely adequate. For the range of error probabilities of interest in quantum communication systems, corresponding to a few signal photons per symbol and $P_e \geqslant 10^{-9}$, we see that an adequate value for the size of the operators is $N \leqslant 50$.

In Fig. 1 we present the case of 2-PSK, that is, $M = 2$, so that the states correspond to the complex values $\pm\alpha$. Note that in this case, since the system is binary, the optimum solution can be derived by means of the Helstrom theory [11]. The error probability is presented for three values of the thermal

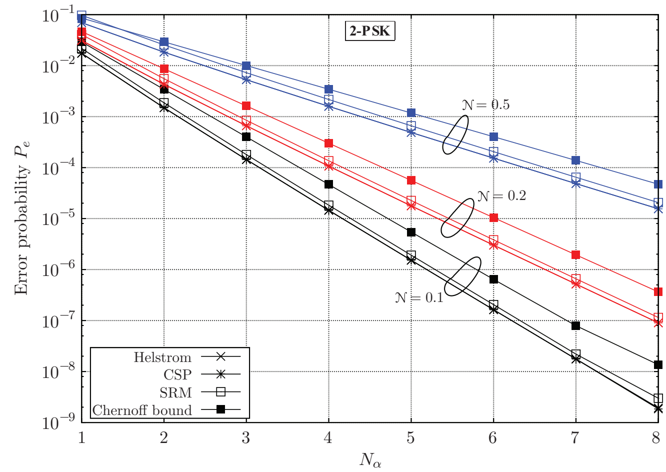

FIG. 1. (Color online) Error probability in 2-PSK vs the average number of photons per symbol N_α for three values of the thermal noise parameter $\mathcal{N}$.

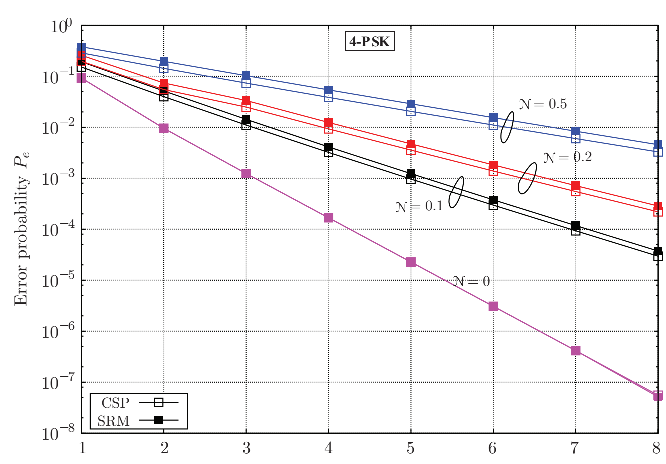

FIG. 2. (Color online) Error probability in 4-PSK vs the average number of photons per symbol N_α for three values of the thermal noise parameter $\mathcal{N}$.

noise parameter $\mathcal{N}$, showing that in general the bound given by the SRM solution is very close to the optimum value, while the Chernoff bound is pessimistic. The optimum obtained by CSP coincides with the Helstrom bound. The Chernoff bound results coincide with the ones derived as in [7], as expected.

In Fig. 2, instead, a multilevel modulation is employed, considering the 4-PSK modulation. In this case, the optimum set of measurement POVM can be obtained only by a lengthy numerical optimization, at least in the presence of thermal noise. Also in this case it is clear that the SRM method outperforms the Chernoff bound, apart from when there is an absence of thermal noise, in which case the bound coincides with the optimum value. Note, however, that in this case, due to the geometric uniform symmetry of the states, the SRM provides the optimum measurement set [9].

VII. CONCLUSIONS

The SRM and quantum Chernoff bound have been considered in the quantum discrimination of Gaussian states and compared with the optimal POVM solution, which in most cases can be derived only by resorting to a heavy numerical optimization procedure. It is shown that for mixed states the SRM solution provides a tighter bound than the Chernoff bound in the binary case and has the advantage that can be applied also to the general M-ary case.

ACKNOWLEDGMENTS

The author would like to thank G. Cariolaro and G. Pierobon for their invaluable help and support. This work has been supported in part by the Project “Q-FUTURE” (prot. STPD08ZXSJ) of the University of Padova.

[1] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **49**, 1007 (2003).
[2] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[3] Y. Eldar and G. D. Forney, IEEE Trans. Inf. Theory **47**, 858 (2001).
[4] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Rev. Mod. Phys. **84**, 621 (2012).
[5] K. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, Commun. Math. Phys. **279**, 251 (2008).
[6] K. M. R. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[7] J. Calsamiglia, R. Munoz-Tapia, L. Masanes, A. Acin, and E. Bagan, Phys. Rev. A **77**, 032311 (2008).
[8] W. Kumagai and M. Hayashi, Commun. Math. Phys. **318**, 535 (2013).
[9] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **50**, 1198 (2004).
[10] M. Hayashi, *Quantum Information: An Introduction* (Springer, Berlin Heidelberg, 2006).
[11] C. Helstrom, J. Liu, and J. Gordon, Proc. IEEE **58**, 1578 (1970).
[12] A. S. Holevo, J. Multivar. Anal. **3**, 337 (1973).
[13] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **21**, 125 (1975).
[14] G. Cariolaro and G. Pierobon, IEEE Trans. Commun. **58**, 623 (2010).
[15] G. Cariolaro and G. Pierobon, IEEE Trans. Commun. **58**, 1213 (2010).
[16] J. Wozencraft and I. Jacobs, *Principles of Communication Engineering* (Wiley, New York, 1965).
[17] T. M. Cover and J. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
[18] M. Nussbaum and A. Szkoła, Ann. Stat. **37**, 1040 (2009).
[19] S. Pirandola and S. Lloyd, Phys. Rev. A **78**, 012331 (2008).
[20] G. Cariolaro, R. Corvaja, and G. L. Pierobon, *Proceedings of Global Telecommunications Conference (GLOBECOM 2011), 5–9 December 2011, Houston, TX* (IEEE, Piscataway, NJ, USA, 2011).
[21] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).