Physics Letters A 373 (2008) 61–64

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Quantum binary channels with mixed states

Norberto Tomassoni [a,b], Matteo G.A. Paris [a,c,d,*]

[a] Dipartimento di Fisica dell'Università di Milano, I-20133, Italy
[b] I.N.RI.M., Strada delle Cacce 91, I-10135 Torino, Italy
[c] CNISM, Udr Milano Università, I-20133 Milano, Italy
[d] Institute for Scientific Interchange Foundation, I-10133 Torino, Italy

A R T I C L E   I N F O

Article history:
Received 31 July 2008
Received in revised form 16 October 2008
Accepted 3 November 2008
Available online 7 November 2008
Communicated by P.R. Holland

A B S T R A C T

We address binary communication channels with symbols encoded in two states of a finite dimensional Hilbert spaces. For pure states we confirm that the optimal decoding stage, maximizing the mutual information, coincides with the projective measure that minimizes the error probability (Bayes criterion). On the other hand, we prove that for communication schemes based on mixed states the optimal decoding, still being a projective measurement, is generally different from the Bayes' one, unless the two density operators commute.

In a classical noise free communication channel we make the basic assumption that different letters of input alphabet correspond to mutually exclusive properties of the information carriers. On the other hand, when the carriers have to be described quantum mechanically, exact discrimination is no longer possible for nonorthogonal quantum states. An optimization problem thus naturally arises, which is aimed to find the best measurement to be performed at the end of the channel, according to the extremization of a suitable figure of merit [1]. Upon looking at the decoding stage of a channel as a decision problem optimization corresponds to the minimization of the error probability or, more generally, of the Bayes' mean cost of the inference strategy [2,3]. On the other hand, from a communication perspective, the quantity to be maximized is the mutual information between the sender and the receiver [4]. Optimal measurements according to Bayes' strategy have been studied for pure states and also for some classes of mixed signals [5,6]. In this Letter, we consider both the approaches for the case of binary channels and prove that they lead to the same measurement for pure states, a result which has been obtained with a different method also in [7], whereas optimization for mixed signals leads to different measurements, unless the two density operators commute. The second result is the original contribution of this Letter.

A binary communication channel consists in a sender (Alice) who encodes symbols 0, 1 in two quantum states, ρ_0 and ρ_1, and a receiver (Bob) who decodes the symbols by measuring a two value probability operator-valued measure (POVM). In general ρ_0 and ρ_1 are not orthogonal, and thus there are no POVMs that allow Bob to discriminate exactly the state sent by Alice. The goal of Bob is thus to optimize the decoding stage in order to retrieve the information sent by Alice as carefully as possible. To this aim Bob may follow different strategies, among which we focus on those minimizing the error probability (Bayes strategy) or maximizing the mutual information. As we will see, the two strategies are equivalent for pure-state encoding, whereas they lead to different optimal POVM for mixed signals.

The POVM that minimizes the error probability may be found as follows. Upon denoting by p_{jk} the conditional probability to infer the symbol j when k was sent by Alice, and by z the a priori probability for the symbol 0, the error probability is given by $P_e(p_{01}, p_{10}, z) = zp_{10} + (1-z)p_{01}$. Assuming that Alice uses the states ρ_0, ρ_1 and Bob implements the POVM $\Pi_0 + \Pi_1 = I$ we have $P_e = z\,\mathrm{Tr}[\Pi_1\rho_0] + (1-z)\,\mathrm{Tr}[\Pi_0\rho_1]$, i.e., $P_e = z + \mathrm{Tr}[\Lambda\Pi_0] = (1-z) - \mathrm{Tr}[\Lambda\Pi_1]$, where $\Lambda = (1-z)\rho_1 - z\rho_0$. The optimal POVM, minimizing the error probability, thus corresponds to a projective POVM with Π_0 being the projector on the subspace spanned by the eigenvectors associated to the negative eigenvalues of Λ and Π_1 the corresponding projector on the positive subspace of Λ.

If $p_X(x)$ is the a priori distribution of Alice's alphabet, then every letter sent by Alice contains on the average an information given by the Shannon information $H(X) = -\sum_x p_X(x)\log p_X(x)$. Given $p_Y(y)$, the distribution of Bob's outcomes, $H(Y)$ is the corresponding average information per letter. Upon defining the Shannon information $H(X,Y)$ for the joint probability distribution $p_{X,Y}(x,y)$ we have that the average information shared by Alice and Bob is measured by the mutual information $H(X:Y) = H(X) + H(Y) - H(X,Y)$, which is the quantity Bob wants to maximize. For a binary channel the mutual information may be written as a function of the conditional and a priori probabilities as follows

* Corresponding author at: Dipartimento di Fisica dell'Università di Milano, I-20133, Italy.
E-mail address: matteo.paris@fisica.unimi.it (M.G.A. Paris).

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2008.11.002

$$H = z(1-p_{10})\log\frac{1-p_{10}}{q_0} + zp_{10}\log\frac{p_{10}}{q_1} + (1-z)p_{01}\log\frac{p_{01}}{q_0} + (1-z)(1-p_{01})\log\frac{1-p_{01}}{q_1} \tag{1}$$

where $q_0 = z(1-p_{10}) + (1-z)p_{01}$ and $q_1 = zp_{10} + (1-z)(1-p_{01})$ are the overall (unconditional) probabilities of Bob's outcomes 0, 1.

Let us now assume that Alice encodes symbols onto two pure states, $|\psi_0\rangle$ and $|\psi_1\rangle$ of a finite dimensional Hilbert space. If P is the projector on the subspace spanned by the two vectors then $p_{jk} = \mathrm{Tr}[\Pi_j|\psi_k\rangle\langle\psi_k|] = \mathrm{Tr}[P\Pi_j P|\psi_k\rangle\langle\psi_k|]$ where $\{P\Pi_0 P, P\Pi_1 P\}$ is a POVM if $\{\Pi_0, \Pi_1\}$ is a POVM. It follows that in maximizing mutual information, or minimizing error probability, we may focus to the bidimensional Hilbert space spanned by Alice's states. Assuming, without loss of generality, that $|\psi_0\rangle$ and $|\psi_1\rangle$ belong to a bidimensional Hilbert space, we can always find a base $\{|0\rangle, |1\rangle\}$ to write (see Fig. 1)

$$\rho_j = |\psi_j\rangle\langle\psi_j| = \frac{1}{2}\big(I + \cos\phi\sigma_1 + (-)^j\sin\phi\sigma_2\big)$$

where $\phi \in [0, \pi/2]$. Analogously a generic two-value POVM may be written as $\Pi_1 = k_0\sigma_0 + k_x\sigma_1 + k_y\sigma_2 + k_z\sigma_3$, $\Pi_0 = I - \Pi_1$, where σ_k, $k = 1, 3$, are Pauli matrices. The positivity conditions for Π_0 and Π_1 correspond to $0 \leqslant k_0 \leqslant 1$ and $|\vec{k}| \leqslant \min(k_0, 1-k_0)$ with the conditions $k_0 = 1/2$, $|\vec{k}| = 1/2$, $|\vec{k}| = (k_x^2 + k_y^2 + k_z^2)^{1/2}$ corresponding to a POVM made by orthogonal projectors (PVM). The conditional probabilities p_{01} and p_{10} may be thus written as

$$p_{01} = k_0 + k_x\cos\phi - k_y\sin\phi,$$
$$p_{10} = 1 - k_0 - k_x\cos\phi - k_y\sin\phi. \tag{2}$$

The mutual information is a function of the five variables: k_0, k_x, k_y, ϕ, and z. Upon maximizing H over k_0, k_x, k_y (within the domain imposed by the positivity conditions) we found that the optimal POVM is projective, i.e., that the mutual information is maximal for $k_0 = 1/2$ and $k_x^2 + k_y^2 + k_z^2 = 1/4$. In order to continue optimization we parametrize the general PVM as $\Pi_0 = \frac{1}{2}(I - \vec{n}\cdot\sigma)$ where $|\vec{n}| = 1$, i.e., in the polar representation

$$\Pi_0 = \frac{1}{2}(I - \sin\theta\cos\beta\sigma_1 - \sin\theta\sin\beta\sigma_2 - \cos\theta\sigma_3) \tag{3}$$

where $\theta \in [0, \pi]$ and $\beta \in [0, 2\pi]$. The conditional probabilities in term of the three parameters θ, β, ϕ rewrites as

$$p_{01} = \frac{1}{2}\big(1 - \sin\theta\cos(\beta+\phi)\big),$$
$$p_{10} = \frac{1}{2}\big(1 + \sin\theta\cos(\beta-\phi)\big) \tag{4}$$

which can be used to write the mutual information $H(\theta, \beta, \phi, z)$ and the error probability $P_e(\theta, \beta, \phi, z)$ as functions of the four parameters θ, β, ϕ and z, ϕ and z describing Alice' encoding and θ and β defining the projective measurement performed by Bob. The optimization of the measurement scheme thus reduces to the maximization of H, or the minimization of P_e, with respect to the variables θ and β. Notice that for $\phi = 0$ the two states coincide and no information may be sent. We also notice that $H(\theta, \beta, \phi, z)$=$H(\theta, \beta + \pi, \phi, z)$ and use this symmetry to restrict the domain to $\theta, \beta \in [0, \pi] \times [0, \pi]$. The maximum of $H(\theta, \beta, \phi, z)$ is located at $\theta(\phi, z) = \frac{\pi}{2}$ and

$$\beta(\phi, z) \equiv \beta_H = \arccos\left(\frac{(2z-1)\cos\phi}{\gamma}\right) \tag{5}$$

with value

$$H_M(\phi, z) = \frac{1}{2}\left[\log\cos^2\phi + (1-2z)\log\frac{z}{1-z} + \gamma\log\frac{1+\gamma}{1-\gamma}\right] \tag{6}$$

where $\gamma = \sqrt{\sin^2\phi + (2z-1)^2\cos^2\phi}$ (see Fig. 2).

For the error probability $P_e(\theta, \beta, \phi, z)$ the minimum value is given by the Helstrom bound

$$P_{e,\min}(\phi, z) = \frac{1}{2}[1 - \gamma] \tag{7}$$

and is achieved for $\theta = \pi/2$ and $\beta = \beta_H$, i.e., the measurement maximizing the mutual information coincide with that minimizing the error probability; both of them are projective measurements (see Fig. 2).

Let us suppose now that Alice employs two mixed states belonging to a bidimensional Hilbert space. We can parametrize two general mixed states as

$$\rho_j = \frac{1}{2}\big(I + (1-2p_j)\cos\phi\sigma_1 + (-)^j(1-2p_j)\sin\phi\sigma_2\big), \tag{8}$$

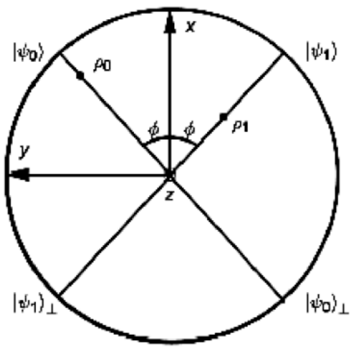

Fig. 1. Representation of a couple of signals ρ_0 and ρ_1 on the (x, y) plane of the Bloch sphere.

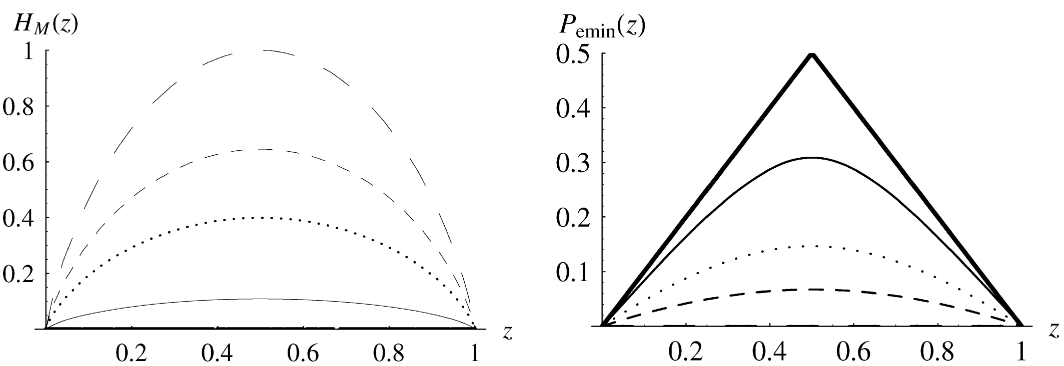

Fig. 2. Maximum mutual information (left) and minimum error probability (right) as functions of z for some values of ϕ: $\phi = 0$ (thick line, coinciding with z-axis in the left graph), $\phi = \pi/8$ (thin line), $\phi = \pi/4$ (dot line), $\phi = \pi/3$ (dashed line), $\phi = \pi/2$ (large dashed line, coinciding with the z-axis in the right graph).

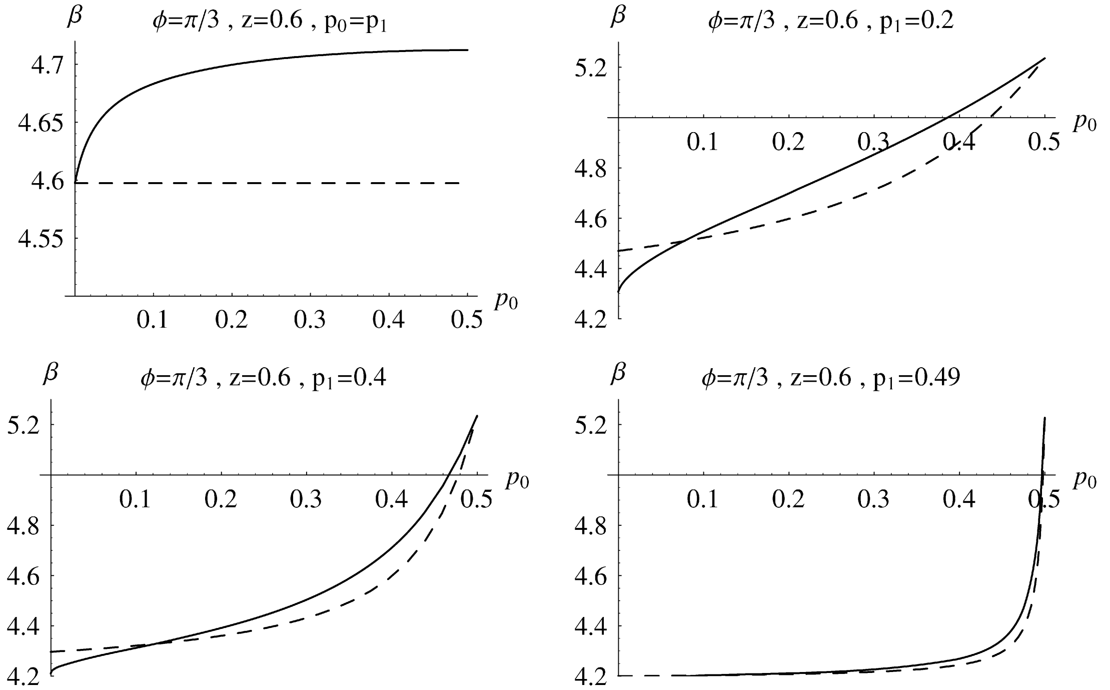

Fig. 3. Angle β optimizing the error probability (dashed line), and angle β optimizing the mutual information (solid line) as functions of p_0. Top left: for $p_1 = p_0$, $\phi = \pi/6$ and $z = 0.6$. Other panels: for $\phi = \pi/3$, $z = 0.6$ and some values of p_1.

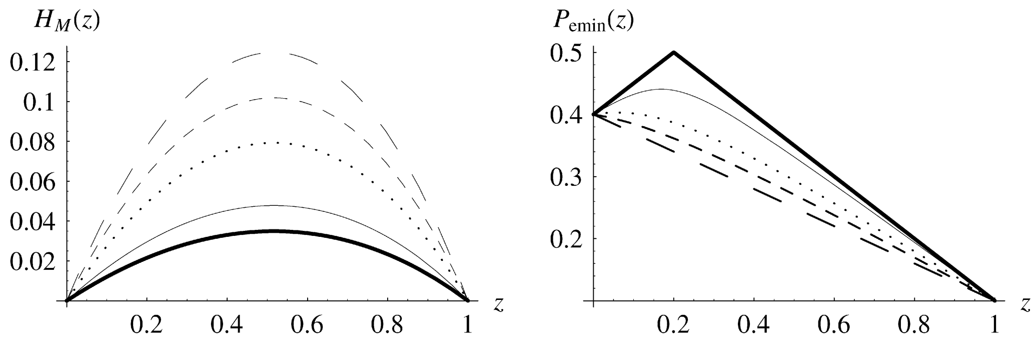

Fig. 4. Maximum mutual information (left) and minimum error probability (right) as functions of z for $\mu_0 = 0.68$ and $\mu_1 = 0.52$ for some values of ϕ: $\phi = 0$ (thick line), $\phi = \pi/8$ (thin line), $\phi = \pi/4$ (dot line), $\phi = \pi/3$ (dashed line), $\phi = \pi/2$ (large dashed line).

with purity of the two signals given by $\mu_j = \frac{1}{2}(1 - \sqrt{2p_j - 1})$. In a way similar to the one used for pure states it is possible to prove that the general measure that maximizes the mutual information is a projective measure. Parametrizing the general PVM as in the pure states situation we can write the mutual information and the error probability as functions of six variables:

$$H = H(\phi, p_0, p_1, z, \theta, \beta), \qquad P_e = P_e(\phi, p_0, p_1, z, \theta, \beta) \tag{9}$$

where again the variables θ and β define the PVM implemented by Bob. Upon maximizing mutual information one may prove analytically that also for mixed states $\theta = \pi/2$. The optimal value of β may be instead found numerically. Concerning the error probability, the minimization may be solved analytically leading to $\theta = \pi/2$ and

$$\beta = \arccos \frac{[2z(1-p_0) - \delta]\cos\phi}{[\delta - 2z(1-p_0)]^2\cos^2\phi + [\delta - 2zp_0]^2\sin^2\phi} \tag{10}$$

where $\delta = 1 - 2p_1(1 - z)$. Comparing now the values of β that maximize mutual information with those minimizing the error probability it is apparent, see Fig. 3, that they generally differ. In the limiting case $p_0 = p_1$ the POVM that minimizes the error probability does not depend on p_0 whereas the POVM that maximizes mutual information show a clear dependence.

We conclude that the measurement that minimizes error probability differs from the measurement that maximizes mutual information. As expected, the states used by Alice commute, then the two optimizing measurements coincide with the projectors on the elements of the common basis. In Fig. 4 we report the maximum mutual information and the minimum error probability as a function of the a priori probability z and some values of the state parameters ϕ, p_0, p_1. The mutual information no longer vanishes for $\phi = 0$, as in the pure states case, since the two states are not generally coincident unless $p_0 = p_1 = 0$. Similarly, the error probability does not vanish for $\phi = \pi/2$, because the two states are not perfectly distinguishable unless $p_0 = p_1 = 0$.

Of course, also after optimization, the protocol with mixed signals cannot outperform the pure state one. The point here is that maximization of H is in general different from minimizing the error probability, as far as the involved signals are mixed, and that

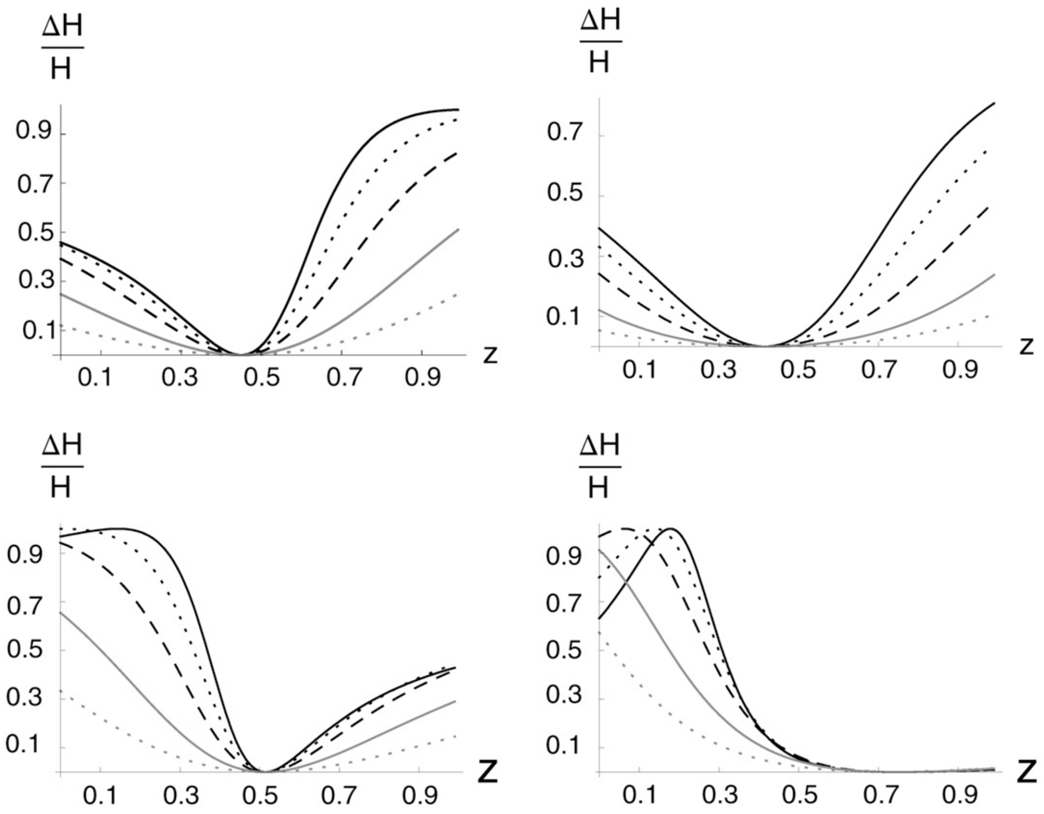

Fig. 5. Relative gain of mutual information (see text) as a function of the a priori probability for different values of ϕ and signals' purities. Top left: $\mu_1 = 0.7$, $\mu_2 = 0.8$. Top right: $\mu_1 = 0.9$, $\mu_2 = 0.95$. Bottom left: $\mu_1 = 0.6$, $\mu_2 = 0.55$. Bottom right: $\mu_1 = 0.95$, $\mu_2 = 0.55$. In all the plots the different curves are for, from top to bottom, $\phi = \pi/10$ (solid black), $\phi = \pi/8$ (dotted black), $\phi = \pi/6$ (dashed black), $\phi = \pi/4$ (solid gray), $\phi = \pi/3$ (doted gray).

optimization is worth to be performed, as it is apparent from Fig. 5, where we compare the mutual information achieved with the optimized measurement with that achievable using Bayes optimal one. We report, for different values of the signals' purities, the relative gain of mutual information $\Delta H = (H_M - H_B)/H_M$, where H_B is the mutual information obtained by performing he measurement which leads to minimum error probability, as a function of the a priori probability z.

The expression (6) represent the maximum mutual information, optimized over two-value POVMs. On the other hand, the maximum mutual information, optimized over all possible POVMs [8], is bounded by the so-called *accessible information*, as expressed by the Holevo bound

$$H_M \leqslant S(\rho) - zS(\rho_0) - (1-z)S(\rho_1) = \chi_H \tag{11}$$

where $\rho = z\rho_0 + (1-z)\rho_1$ and $S(\rho) = -\mathrm{Tr}[\rho \ln \rho]$ denotes the Von Neumann entropy of the state ρ. The Holevo bound is saturated by H_M iff one of the following relation is satisfied: $z = 0, 1$, $\phi = 0\pi/2$, $p_0 = 1/2$, $p_1 = 1/2$. In the first case we have $H_M = \chi_H = 0$. In the other cases the states used by Alice to communicate commute. When the bound is saturated the quantity H_M is of course equal to the accessible information.

In conclusion, in a binary communication channel based on pure states of an arbitrary, finite-dimensional, Hilbert space the measurement that maximizes the mutual information is a projective measurement and coincides with the measurement that minimizes the error probability. If the encoding signals are mixed states belonging to a bidimensional Hilbert space the measurement maximizing the mutual information is still a projective measurement but is generally different from the measurement that minimizes error probability. Only for "classical" commuting signals, the two POVMs coincide, and are given by the projectors on the elements of the common basis. Remarkably, optimization for mixed signals generally leads to a consistent gain of mutual information compared to the Bayes one. The maximum mutual information over all two-value measurements is less or equal to the accessible information and saturates the Holevo bound when the signals commute.

References

[1] C.W. Helstrom, Quantum Detection and Estimation Theory, Academic Press, New York, 1976.
[2] T. Chefles, in: Lecture Notes in Physics, vol. 646, Springer, Berlin, 2004.
[3] J. Bergou, U. Herzog, M. Hillery, in: Lecture Notes in Physics, vol. 646, Springer, Berlin, 2004.
[4] C.M. Caves, P.D. Drummond, Rev. Mod. Phys. 66 (1994) 481.
[5] M. Sasaki, et al., Phys. Rev. A 55 (1997) 3222.
[6] Y. Fujihara, et al., Electron. Commun. Jpn. 86 (2003) 63.
[7] M. Osaki, O. Hirota, M. Ban, J. Mod. Opt. 45 (1998) 269.
[8] M. Sasaki, et al., Phys. Rev. A 59 (1999) 3325.