Quantification of Entanglement by Means of Convergent Iterations

Jaroslav Řeháček and Zdeněk Hradil
Department of Optics, Palacký University, 772 00 Olomouc, Czech Republic
(Received 12 June 2002; published 26 March 2003)

The relative entropy of entanglement of a given bipartite quantum state is calculated by means of a convergent iterative algorithm. When this state turns out to be nonseparable, the algorithm also provides the corresponding optimal entanglement-witness measurement.

DOI: 10.1103/PhysRevLett.90.127904 PACS numbers: 03.67.Mn, 03.65.Ud

Entanglement is an important resource of quantum information processing. Although there are quantum protocols based on other features of quantum mechanics such as quantum superposition principle in Deutsch's celebrated searching algorithm [1] or protocols where the use of entanglement can be advantageous but is not essential, such as signaling through depolarizing channels with memory [2], most quantum protocols rely on the existence of nonseparable states. For practical purposes, it is very important to quantify entanglement generated by realistic laboratory sources and thus evaluate the potential usefulness for quantum processing/communication purposes.

One of the measures of entanglement thoroughly studied over the past decade is the relative entropy of entanglement defined as [3]

$$E(\sigma) = \inf_{\rho \mathrm{sep.}} S(\sigma \parallel \rho), \tag{1}$$

where S is the quantum relative entropy,

$$S(\sigma \parallel \rho) = \mathrm{Tr}(\sigma \ln \sigma - \sigma \ln \rho), \tag{2}$$

between states σ and ρ; the infimum in Eq. (1) is taken over the set of separable states. This functional is one possible generalization of the classical relative entropy between two probability distributions [4] to quantum theory. Let us mention that unlike in the case of entropy this generalization is by no means unique. It can be interpreted geometrically as a quasidistance between the state whose entanglement we are interested in and the convex set of separable states. E fulfills most of the requirements usually imposed on a good entanglement measure and has other good properties. Most notably, on the set of pure states it reproduces the Von Neumann reduced entropy [5,6] and is closely related to some other measures of entanglement [7]. Relative entropy also makes a good entanglement measure for multipartite [8] and infinite-dimensional [9] quantum systems.

The analytical form of E is known only for some special sets of states of high symmetry [10–13]. Generally, one has to resort to numerical calculation. In a sense, the problem resembles the reconstruction of quantum states using the maximum likelihood principle [14,15]. Here the given input state σ plays the role of experimental data; once this state is known, the statistics of any possible measurement performed on it is available. The solution can be obtained by means of several numerical methods. The formulation given in [5] is just an example corresponding to the implementation of the steepest descent method. Its efficiency strongly depends on the dimensionality of the problem.

In the procedure proposed here a more analytical approach will be adopted. We derive a set of extremal equations for E and show how to solve them by means of repeated convergent iterations. But this is not the only goal. The extremal equations indicate that there is a structure of quantum measurement associated with the extremal solution. The separable measurement obtained in this way specifies the extremal separable state and, significantly, it provides the optimal entanglement-witness operator, revealing the possible entanglement of the input state σ.

Let us denote by ρ^* the separable state having the smallest quantum relative entropy with respect to σ. Let $f(x, \rho^*, \rho) = S(\sigma \parallel (1 - x)\rho^* + x\rho)$ be the relative entropy of a state obtained by moving from ρ^* towards some ρ. We are looking for the global maximum of a convex functional on the convex set of separable states. Two cases may arise. When σ is separable, the necessary and sufficient condition for the maximum of S is that its variations along the paths lying in the set of separable states vanish,

$$\frac{\partial f}{\partial x}(0, \rho^*, \rho) = 0, \quad \forall \rho \quad \text{separable.} \tag{3}$$

When σ is entangled, S attains its true maximum outside the set of separable states and we must carry on the maximization on the boundary. In that case, Eq. (3) holds only for certain variations along the boundary. It is well known that any separable state from the Hilbert space of dimension $p = d \otimes d$ can be expressed as a convex sum of (at most) p^2 projectors on disentangled pure states (Caratheodory's theorem, see also [5]),

$$\rho = \sum_{k=1}^{p^2} |\varphi_k^1\rangle\langle\varphi_k^1| \otimes |\varphi_k^2\rangle\langle\varphi_k^2|. \tag{4}$$

Here $|\varphi_k^1\rangle$ and $|\varphi_k^2\rangle$ are pure states (not normalized) of the systems 1 and 2, respectively. This could be also seen as the expansion of a separable state in a separable operator basis given by linearly independent but not necessarily orthogonal disentangled projectors.

As independent variables of our problem, we take $2p^2$ vectors $|\varphi_k^{1,2}\rangle$ rather than $2p^2$ projectors $|\varphi_k^{1,2}\rangle\langle\varphi_k^{1,2}|$ as this has two main advantages: The set of parameters with respect to which we minimize the relative entropy becomes convex, and it has no boundary since any set of $2p^2$ vectors defines a separable state through Eq. (4).

The derivation in Eq. (3) can easily be calculated using the integral representation of the logarithm of a positive operator [5]. It reads

$$\frac{\partial f}{\partial x}(0,\rho^*,\rho) = \int_0^\infty \mathrm{Tr}((\rho^*+t)^{-1}\sigma(\rho^*+t)^{-1}\delta\rho)dt = \mathrm{Tr}A\delta\rho, \tag{5}$$

where we denoted $(1-x)\rho^* + x\rho = \rho^* + \delta\rho$, and operator A has the following matrix elements in the eigenbasis $\{|\lambda_n\rangle\}$ of ρ^* [16]:

$$\langle\lambda_m|A|\lambda_n\rangle = \frac{\log\lambda_n - \log\lambda_m}{\lambda_n - \lambda_m}\langle\lambda_m|\sigma|\lambda_n\rangle. \tag{6}$$

Its meaning will be discussed later. The variation $\delta\rho$ in Eq. (5) is generated by the $2p^2$ variations of the independent variables $|\varphi_k^{1,2}\rangle \to |\varphi_k^{1,2}\rangle + |\delta\varphi_k^{1,2}\rangle$. Keeping only terms to the first order in the small quantities, it reads

$$\delta\rho = \sum_k (|\varphi_k^1\rangle\langle\varphi_k^1| \otimes |\varphi_k^2\rangle\langle\delta\varphi_k^2| + |\varphi_k^1\rangle\langle\delta\varphi_k^1| \otimes |\varphi_k^2\rangle\langle\varphi_k^2|) + \mathrm{h.c.} \tag{7}$$

The necessary condition for the maximum of the relative entropy is that the right-hand side of Eq. (5) vanishes for all $|\delta\varphi_k^{1,2}\rangle$. To make sure that the trace constraint $\mathrm{Tr}(\rho^* + \delta\rho) = 1$ is obeyed, we use a Lagrange multiplier λ. Using Eq. (7), the extremal equation becomes

$$0 = \mathrm{Tr}(A-\lambda)\delta\rho = \sum_k \langle\delta\varphi_k^1|(R_k^1 - \lambda)|\varphi_k^1\rangle\langle\varphi_k^2 \mid \varphi_k^2\rangle + \sum_k \langle\delta\varphi_k^2|(R_k^2-\lambda)|\varphi_k^2\rangle\langle\varphi_k^1 \mid \varphi_k^1\rangle + \mathrm{h.c.}, \tag{8}$$

where we defined $R_k^1 = \mathrm{Tr}_2(A\overline{|\varphi_k^2\rangle\langle\varphi_k^2|})$, $R_k^2 = \mathrm{Tr}_1(A\overline{|\varphi_k^1\rangle\langle\varphi_k^1|})$; the bars denote the normalization to unity. Now since Eq. (8) holds for all $|\delta\varphi_k^{1,2}\rangle$, the $2p^2$ vectors $(R_k^{1,2} - \lambda)|\varphi_k^{1,2}\rangle$ are seen to be zero vectors [17]. This gives us the following set of extremal equations:

$$\begin{aligned} R_k^1|\varphi_k^1\rangle\langle\varphi_k^1| &= \lambda|\varphi_k^1\rangle\langle\varphi_k^1|, \\ R_k^2|\varphi_k^2\rangle\langle\varphi_k^2| &= \lambda|\varphi_k^2\rangle\langle\varphi_k^2|, \\ k &= 1,\ldots,p^2. \end{aligned} \tag{9}$$

Multiplying the first row of Eq. (9) by $|\varphi_k^2\rangle\langle\varphi_k^2|$, the second by $|\varphi_k^1\rangle\langle\varphi_k^1|$, summing them separately over k, tracing out, and using Eqs. (4) and (6), we find that $\lambda = \mathrm{Tr}A\rho^* = 1$. Finally, we rewrite the left-hand sides of Eq. (9) into an explicitly positive semidefinite form: If $R|\varphi\rangle\langle\varphi| = |\varphi\rangle\langle\varphi|$ holds for an operator R and a vector $|\varphi\rangle$, so does its Hermitian conjugate $|\varphi\rangle\langle\varphi|R^\dagger = |\varphi\rangle\langle\varphi|$. Then if R is Hermitian it follows that $|\varphi\rangle\langle\varphi| = R|\varphi\rangle\langle\varphi| = R(|\varphi\rangle\langle\varphi|R^\dagger) = R|\varphi\rangle\langle\varphi|R$. In the same spirit, we get from Eq. (9) the main formal result of this Letter: The state having the smallest quantum relative entropy with respect to a given state σ satisfies the following $2p^2$ equations:

$$\begin{aligned} R_k^1|\varphi_k^1\rangle\langle\varphi_k^1|R_k^1 &= |\varphi_k^1\rangle\langle\varphi_k^1|, \\ R_k^2|\varphi_k^2\rangle\langle\varphi_k^2|R_k^2 &= |\varphi_k^2\rangle\langle\varphi_k^2|, \qquad k = 1,\ldots,p^2. \end{aligned} \tag{10}$$

Unfortunately, solving such highly nonlinear operator equations by analytical means for anything but the most trivial states seems to be out of the question. One has to turn to numerics. We suggest solving Eqs. (10) by repeated iterations starting from $2p^2$ randomly chosen projectors. Let us note that the iterative procedure based on Eqs. (10) belongs to the family of gradient-type algorithms of the form $x_k^{i+1} = (\partial S(\mathbf{x}^i)/\partial x_k)x_k^i(\partial S(\mathbf{x}^i)/\partial x_k)$. Algorithms of this type are known to behave well; some of them were even proven to converge monotonically [18]. They have found important applications in various optimizations and inverse problems. In our case, we observed that the step generated by operators $R_k^{1,2}$ in Eqs. (10) was often too large—rather than converging to the stationary point, the algorithm would oscillate or diverge. If this happens, the length of the step can be made smaller by mixing the operators $R_k^{1,2}$ with the unity operator:

$$R_k^{1,2} \to (\mathbb{1} + \tfrac{1}{2}\alpha R_k^{1,2})/(1+\tfrac{1}{2}\alpha). \tag{11}$$

Indeed, when α is sufficiently small the algorithm converges monotonically. This can be seen by considering an infinitesimal step with $\alpha \ll 1$. It is convenient to split one iteration of Eqs. (10) into two subsequent steps corresponding to the two rows of Eqs. (10) (projectors of only one of the subsystems are updated at a time). The two steps are completely symmetrical, so we consider an infinitesimal iteration on, say, the projectors $|\varphi_k^{1i}\rangle\langle\varphi_k^{1i}|$ of the first system obtained after the ith iteration. We want to show that after one such step the quantum relative entropy is never increased, $S(\sigma \parallel \rho^{i+1}) \le S(\sigma \parallel \rho^i)$. Using Eq. (11) in Eqs. (10) we get to the first order in α,

$$\rho^{i+1} = (1-\alpha)\rho^i + \alpha\tilde{\rho}, \tag{12}$$

where $\tilde{\rho} = \frac{1}{2}\sum_k(R_k^{1i}|\varphi_k^{1i}\rangle\langle\varphi_k^{1i}| + |\varphi_k^{1i}\rangle\langle\varphi_k^{1i}|R_k^{1i}) \otimes |\varphi_k^{2i}\rangle\langle\varphi_k^{2i}|$, and thus

$$S(\sigma \parallel \rho^{i+1}) - S(\sigma \parallel \rho^i) \propto \frac{\partial f(0,\rho^i,\tilde{\rho})}{\partial\alpha} = 1 - \mathrm{Tr}A^i\tilde{\rho}. \tag{13}$$

It remains to be shown that $\mathrm{Tr}A^i\tilde{\rho} \ge 1$. Let us denote $\lambda_k^i = \langle\varphi_k^{1i} \mid \varphi_k^{1i}\rangle\langle\varphi_k^{2i} \mid \varphi_k^{2i}\rangle$. Notice that $\sum_k \lambda_k^i = 1$ by the normalization of ρ^i. Then by using the Schwarz inequality and the concavity of the square function, we obtain

$$\begin{aligned}\mathrm{Tr}A^i\tilde{\rho} &= \sum_k \lambda_k^i \mathrm{Tr}(R_k^{1i}\overline{|\varphi_k^{1i}\rangle\langle\varphi_k^{1i}|}R_k^{1i}) \\ &\geq \sum_k \lambda_k^i [\mathrm{Tr}(R_k^{1i}\overline{|\varphi_k^{1i}\rangle\langle\varphi_k^{1i}|})]^2 \\ &\geq \Big[\sum_k \lambda_k^i \mathrm{Tr}(R_k^{1i}\overline{|\varphi_k^{1i}\rangle\langle\varphi_k^{1i}|})\Big]^2 = (\mathrm{Tr}A^i\rho^i)^2 = 1,\end{aligned} \tag{14}$$

which completes our proof. This means that there exists $\alpha > 0$ such that after the regularization (11) the algorithm (10) will converge monotonically. In practice, the parameter α need not be very small. In $2 \otimes 2$ and $4 \otimes 4$ dimensional problems we tried, a monotonic convergence was observed even when α was of the order of unity.

Now let us go back to the extremal Eqs. (10). The left-hand sides are generated by the operator A, which depends on σ through Eq. (6). From now on, let us assume that σ is an entangled state. Then A represents the gradient of the quantum relative entropy $S(\sigma \parallel \rho)$ at ρ^*—the separable state closest to σ. Loosely speaking, the states giving the same expectation, $\mathrm{Tr}A(\sigma)\rho = \mathrm{const}$, form hyperplanes that are perpendicular to the line connecting σ and ρ^*. Since $\mathrm{TrA}(\sigma)\rho^*=1$ and ρ^* lies at the boundary of the set of separable states, the conjecture is that the operator A is up to a shift of its spectrum a witness operator [19,20] detecting the entanglement of σ. In the following, we prove this conjecture and show that the operator

$$W(\sigma) = \mathbb{1} - A(\sigma) \tag{15}$$

is indeed the optimal witness of the entanglement of σ. The mutual relationship of σ, ρ^*, and the states detected by W is shown in Fig. 1.

First we show that $\mathrm{Tr}A\rho \leq 1$ if ρ is separable. To this end, let us note that

$$\begin{aligned}1 - \mathrm{Tr}A\rho &= \frac{\partial f}{\partial x}(0, \rho^*, \rho) \\ &= \lim_{x\to 0}\frac{S(\sigma \parallel (1-x)\rho^* + x\rho) - S(\sigma \parallel \rho^*)}{x}.\end{aligned} \tag{16}$$

Since both ρ and ρ^* are separable states, so is their convex combination $(1-x)\rho^* + x\rho$. But ρ^* minimizes $S(\sigma \parallel \rho)$ over the set of separable states. Therefore, $S(\sigma \parallel (1-x)\rho^* + x\rho) - S(\sigma \parallel \rho^*) \geq 1$. This holds for all x, so we have

$$\mathrm{Tr}W(\sigma)\rho = 1 - \mathrm{Tr}A(\sigma)\rho \geq 0, \quad \forall\rho \quad \text{separable}. \tag{17}$$

This already means that W is an entanglement-witness operator. To show that W detects σ, we again make use of Eq. (16) with ρ now being substituted by the entangled state σ. Now, because of convexity of S,

$$\frac{S(\sigma \parallel (1-x)\rho^* + x\sigma) - S(\sigma \parallel \rho^*)}{x} \leq -S(\sigma \parallel \rho^*) < 0. \tag{18}$$

The last inequality follows from the assumed nonseparability of σ. Equation (18) also holds for any x, so we obtain

$$\mathrm{Tr}W(\sigma)\sigma = 1 - \mathrm{Tr}A(\sigma)\sigma < 0, \quad \forall\sigma \quad \text{entangled}, \tag{19}$$

which we set out to prove.

Possible applications of our algorithm are twofold: First, it can be used for checking whether a given state is separable or not. Second, it can be used for quantifying the amount of entanglement the state contains. As a test of separability, we tested the algorithm on many randomly generated separable states and entangled states with negative as well as positive partial transposition of dimensions $2 \otimes 2$, $3 \otimes 3$, and $4 \otimes 4$; see Fig. 2 for typical results.

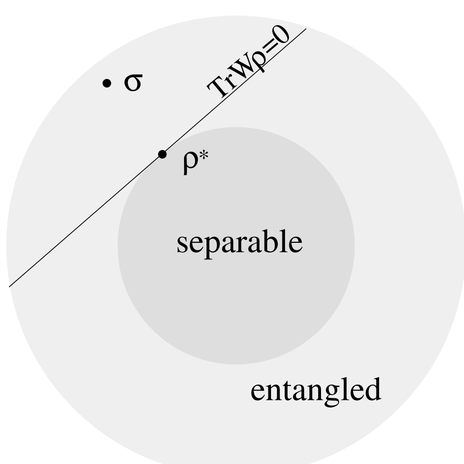

FIG. 1. Mutual relationship of an entangled state σ, the separable state ρ^* closest to it in the sense of quantum relative entropy, and the entangled states detected by the witness operator W.

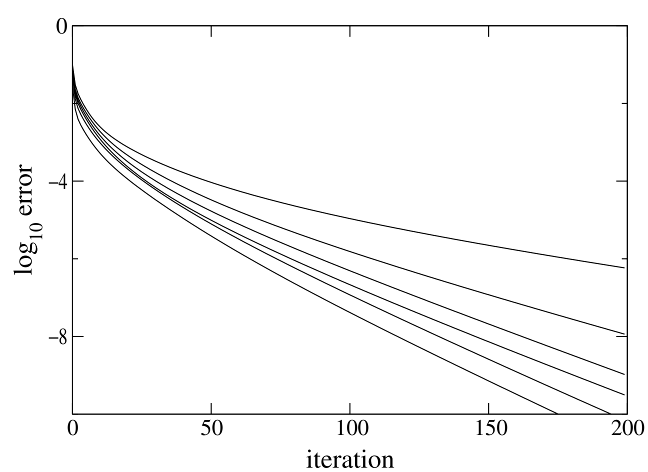

FIG. 2. Test of separability. It is shown how the calculated relative entropy of entanglement of several randomly generated separable states approaches zero in the course of iterating. The ordinate is labeled by the precision in decimal digits.

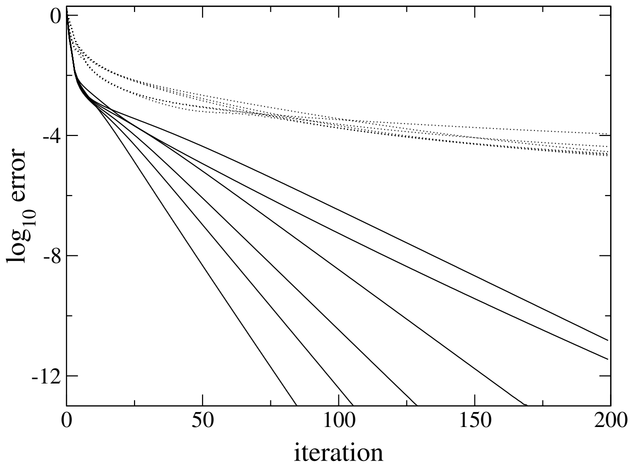

FIG. 3. The accuracy of the calculated relative entropy of entanglement is shown after a given number of iterations for six Werner states with $f \in [-0.05, -0.8]$. The ordinate is labeled by the precision in decimal digits. Solid lines: the proposed algorithm based on Eq. (10); dotted lines: the steepest descent minimization suggested in [5].

Recently, another numerical test of separability has been proposed [21] consisting of a hierarchy of gradually more and more complex separability criteria that can be formulated as separate problems of the linear optimization theory. The algorithm we propose is much more simple. There is just one set of equations to be solved by repeated iterations and after that one finds not only whether the input state is entangled but also how much.

Unfortunately, explicit formulas for the quantum relative entropy are known only in very few cases. One such exception is the family of Werner states defined as follows:

$$\rho_w = \frac{d-f}{d(d^2-1)}\mathbb{1} + \frac{fd-1}{d(d^2-1)}F, \tag{20}$$

where F is the flip operator and $f = \mathrm{Tr}\rho_w F \in [-1, 1]$ is a parameter. Figure 3 shows the performance of our algorithm for several entangled Werner states of dimension $4 \otimes 4$. It is worth mentioning that the optimal entanglement-witness W for the detection of Werner states of two qubits generated by the operator A in Eq. (6) is simply $W = 1 - 2|\Psi_-\rangle\langle\Psi_-|$, where Ψ_- is the singlet state. The expectation value of W is then just a renormalized singlet fraction.

The curves appearing in Figs. 2 and 3 suggest that the convergence of the proposed algorithm is faster than polynomial but slower than exponential. In some cases such as that shown in Fig. 3, the convergence is nearly exponentially fast (the number of accurate digits grows linearly per iteration). This is just a qualitative statement since we did not attempt to do any optimization of the length of the iteration step. Even so, the performance of our algorithm is good compared to other methods. See Fig. 3 for comparison with the steepest descent method suggested in [5]; the downhill simplex method that was suggested for the maximization of the likelihood functional in quantum tomography [22] is even slower. In all probability, more optimization would result in further speedup compared to our examples given in Figs. 2 and 3.

In conclusion, we derived a convergent iterative algorithm for the calculation of the relative entropy of entanglement. It can be used for checking whether a given input state is entangled. If it is entangled the algorithm calculates its relative entropy of entanglement, finds its best separable approximation, and provides the corresponding optimal entanglement-witness measurement.

We thank Koenraad Audenaert for his pertinent comments. This work was supported by Projects No. J14/98 and No. LN00A015 of the Czech Ministry of Education and by the EU grant under QIPC, Project No. IST-1999-13071 (QUICOV).

[1] D. Deutsch and R. Jozsa, Proc. R. Soc. London, Ser. A **439**, 553 (1992).
[2] C. Macchiavello and G. M. Palma, Phys. Rev. A **65**, 050301 (2002).
[3] V. Vedral, M. B. Plenio, K. Jacobs, and P. L. Knight, Phys. Rev. A **56**, 4452 (1997).
[4] S. L. Braunstein, Phys. Lett. A **219**, 169 (1996).
[5] V. Vedral and M. B. Plenio, Phys. Rev. A **57**, 1619 (1998).
[6] M. B. Plenio, S. Virmani, and P. Papadopoulos, J. Phys. A **33**, 193 (2000).
[7] L. Henderson and V. Vedral, Phys. Rev. Lett. **84**, 2263 (2000).
[8] M. B. Plenio and V. Vedral, J. Phys. A **34**, 6997 (2001).
[9] J. Eisert, C. Simon, and M. B. Plenio, J. Phys. A **35**, 3911 (2002).
[10] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275 (1997).
[11] K. Audenaert, J. Eisert, E. Jané, M. B. Plenio, S. Virmani, and B. De Moor, Phys. Rev. Lett. **87**, 217902 (2001).
[12] K. G. H. Vollbrecht and R. F. Werner, Phys. Rev. A **64**, 062307 (2001).
[13] K. Audenaert, B. De Moor, K. G. H. Vollbrecht, and R. F. Werner, quant-ph/0204143.
[14] Z. Hradil, Phys. Rev. A **55**, 1561(R) (1997).
[15] J. Řeháček, Z. Hradil, and M. Ježek, Phys. Rev. A **63**, 040303(R) (2001).
[16] When $m = n$ or $\lambda_n = \lambda_m$, the corresponding coefficient should be replaced by the limit value of λ_n^{-1}.
[17] Here we assume that all p^2 vectors $|\varphi_k^{1,2}\rangle$ have nonzero norm. With this assumption we lose no generality, because the decomposition (4) is not unique. For example, a disentangled pure state can also be decomposed into p^2 disentangled projectors, all of them being proportional to the pure state itself.
[18] C. L. Byrne, IEEE Trans. Image Process. **2**, 96 (1993).
[19] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[20] B. M. Terhal, Phys. Lett. A **271**, 319 (2000).
[21] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002).
[22] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A **61**, 010304(R) (1999).