Transitions in the Communication Capacity of Dissipative Qubit Channels

D. Daems*

QuIC, Ecole Polytechnique, Université Libre de Bruxelles, 1050 Brussels, Belgium

(Received 19 February 2008; revised manuscript received 16 December 2008; published 5 May 2009)

The information transmission is studied for quantum channels in which the noise includes dissipative effects, more specifically, nonunitality. Noise is usually a nuisance but can sometimes be helpful. For these channels, the communication capacity is shown to increase with the dissipative component of the noise and may exhibit transitions beyond which it increases faster. The optimal states are constructed analytically as well as the pertaining "phase" diagram.

DOI: 10.1103/PhysRevLett.102.180503 PACS numbers: 03.67.Hk

A fundamental issue of quantum information theory concerns the capacities of a quantum channel. They specify the largest amount of information that can be reliably transmitted per use of the noisy channel. The coding and decoding required to achieve these maximal transmission rates are such that the probability of errors goes to zero for increasing number of parallel and independent uses of the channel. Early works on this topics were mainly devoted to memoryless unital channels for which consecutive signal transmissions through the channel are not correlated, and which preserve the identity [1–4]. Recently, much attention was given to unital quantum channels with memory in order to increase the rate of classical or quantum information by entangling multiple uses of the channel [5–8].

The focus is here on the transmission of classical information through memoryless channels with a general type of noise (namely nonunital channels, to be defined below), which, in particular, include dissipative effects. Nonunital channels have been considered in the pioneering work by Fuchs [9] who showed that the states maximizing the classical information capacity can be nonorthogonal, which is somehow counterintuitive since nonorthogonal states cannot be distinguished with perfect reliability. Schumacher and Westmoreland [10] have shown that the optimal states of the amplitude damping channel, a paradigm for the description of energy dissipation, also share that property. Finally, examples have been given of nonunital qubit channels that require three [11] or four [12] input states to achieve the Holevo capacity.

In this Letter, we first evaluate analytically the Holevo capacity and the states which are optimal for the coding of classical information in a class of qubit nonunital channels which is sufficiently large to include most channels considered in these early works. Second, we show that for some memoryless nonunital channels, there are transitions between different types of optimal states for the coding when the parameter controlling the nonunitality is increased while all other channel parameters are kept fixed. Moreover, we obtain that the capacity increases with the nonunitality of the channel, and even faster after each transitions. The pertaining "phase" diagram is constructed analytically. The dissipative component of the noise therefore effectively reduces the noise for some states that are then optimal for the information transmission. This can be viewed as one instance of fighting noise with (dissipative) noise, which has been used in quantum cryptography [13].

The maximal amount of classical information which can be extracted from an ensemble of states ρ_i occurring with probabilities p_i is given by [1]

$$\chi(\{p_i, \rho_i\}) = H\Big(\sum_i p_i\rho_i\Big) - \sum_i p_i H(\rho_i), \tag{1}$$

where $H(\rho) = -\mathrm{Tr}(\rho\log_2\rho)$ is the von Neumann entropy. For a quantum communication channel $\mathcal{E}$, one can define the Holevo capacity as the largest value of χ over all ensembles of input states ρ_i and probabilities p_i

$$\chi_{\mathcal{E}} = \sup_{p_i,\rho_i}\chi(\{p_i, \mathcal{E}(\rho_i)\}). \tag{2}$$

It represents the maximal amount of classical information which can be transmitted over a quantum channel when the codewords used for data transmission are composed of tensor products of the optimal signal states and probabilities, i.e., when they are not entangled over multiple uses of the channel. Clearly, the χ quantity would be the largest if for some optimal set, the term $H[\sum_i p_i\mathcal{E}(\rho_i)]$ can be equal to its maximum $\log_2 d$ (for qdits) while the term $\sum_i p_i H[\mathcal{E}(\rho_i)]$ is as small as possible. There exist channels for which the relation $\chi_{\mathcal{E}} = \log_2 d - \inf_{\rho_i} H[\mathcal{E}(\rho_i)]$ is known to hold. These are channels which preserve the identity (unital or doubly stochastic channels). For qubits, all unital channels have that property [2], whereas for qdits, it holds for some unital channels [3] and is known not to hold for some others [4]. For channels which do not preserve the identity (nonunital channels), this relation need not hold. The optimal states are thus a compromise between the two terms as was first observed by Fuchs [9]. Here, we shall restrict ourselves to the qubit case.

A qubit transmission channel is described by a completely positive linear map which converts an input state ρ to an output state $\mathcal{E}(\rho)$ that will be denoted ρ' for brevity. An arbitrary density matrix ρ for a mixed state qubit may be expressed in terms of its Bloch vector r as $\rho = \frac{1}{2}\times(I + r.\sigma)$ where $r = (x, y, z)$ is real and such that $|r| \leq 1$, with equality for pure states, and $\sigma = \text{col}(\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices. The channel can be written explicitly in terms of the Bloch vectors as

$$\mathcal{E}: r \to r' = rM + \eta, \tag{3}$$

where the matrix M and vector η fully characterize the noise properties of the channel in an alternate but equivalent way as the Kraus representation. Here, we consider a diagonal matrix $M = \text{diag}(a, a, b)$ which contracts Bloch vectors equally with a factor $0 \leq a \leq 1$ along the x and y directions and possibly with a different factor $0 \leq b \leq 1$ along the z direction. If $\eta = (0, 0, 0)$, the channel maps the identity to the identity (unital channel). We shall investigate the case of nonunital channels for which $\eta = (0, 0, c)$. Without loss of generality, we shall take $c \geq 0$ so that the effect of the channel is to contract the Bloch sphere and shift it toward the north pole $(0, 0, 1)$ by a distance c. Given a and b, there is a maximum value of c such that the unit sphere is mapped by the channel into itself,

$$\begin{aligned} M &= \text{diag}(a, a, b) \qquad 0 \leq a, b \leq 1 \\ \eta &= (0, 0, c) \qquad 0 \leq c \leq c_{\max}, \end{aligned} \tag{4}$$

where $c_{\max} = 1 - b$ if $a \leq \sqrt{b}$ and $\{1 + b^2 - a^2 - \frac{b^2}{a^2}\}^{1/2}$ otherwise.

The Holevo capacity is determined by the states and probabilities which maximize the difference $H(\bar{\rho}') - \sum_i p_i H(\rho_i')$, where $\bar{\rho} \equiv \sum_i p_i \rho_i$ and the prime indicates the image under $\mathcal{E}$. It is well known that the von Neuman entropy of a qubit state ρ' only depends on the length $|r'| \equiv \ell$ of its Bloch vector: $H(\rho') = -\frac{1+\ell}{2}\log_2\frac{1+\ell}{2} - \frac{1-\ell}{2}\log_2\frac{1-\ell}{2} = S_{\text{bin}}(\frac{1+\ell}{2}) \equiv S(\ell)$. In addition, since $0 \leq \ell \leq 1$, $S(\ell)$ is decreasing with L.

The rotation symmetry of the problem is such that to any state, one can associate another one such that the entropy $H(\bar{\rho}')$ is higher than if a single element of the pair were considered while the value of $\sum_i p_i H(\rho_i')$ is unchanged. Indeed, since $\bar{\ell} = \{a^2\bar{x}^2 + a^2\bar{y}^2 + (b\bar{z} + c)^2\}^{1/2}$, the first two terms of the square root vanish if for every state there is another one occurring with the same probability but with opposite values for the x and y components. Hence, $\bar{\ell}$ reduces to $|b\bar{z} + c|$. On the other hand, taking the same value for the z components yields output Bloch vectors which have the same length and therefore the same individual output entropy. It follows that the optimal states will necessarily come in the form of such pairs.

Specifically, we consider a first pair of states ρ_1 and ρ_2 occurring with probabilities $p_1 = p_2 = \frac{p_+}{2}$ and whose Bloch vectors are $r_1 = (x, y, \beta_+)$ and $r_2 = (-x, -y, \beta_+)$ with $\beta_+ \geq 0$. The second pair of states is ρ_3, ρ_4 with probabilities $p_3 = p_4 = \frac{1-p_+}{2}$ and Bloch vectors $r_3 = (x, y, -\beta_-)$ and $r_4 = (-x, -y, -\beta_-)$ with $\beta_- \geq 0$. The output Bloch vectors have lengths $\ell_1 = \ell_2 = \ell(\beta_+) \equiv \ell_+$ and $\ell_3 = \ell_4 = \ell(-\beta_-) \equiv \ell_-$, respectively, with $\ell_\pm = \{a^2 + c^2 \pm 2bc\beta_\pm + (b^2 - a^2)\beta_\pm^2\}^{1/2}$. Note that the particular values of the x and y components are irrelevant since they do not enter $\bar{\ell}$ or $\ell_\pm$. The optimization is thus carried over p_+, β_+, and β_- for fixed channel parameters a, b, c,

$$\chi_{\mathcal{E}} = \sup_{p_+,\beta_+,\beta_-}\{S(\bar{\ell}) - p_+S(\ell_+) - (1 - p_+)S(\ell_-)\}. \tag{5}$$

We first consider the following quantity,

$$\frac{\partial\chi}{\partial p_+} = -\frac{b(\beta_+ + \beta_-)\text{arctanh}\bar{\ell}}{\ln 2} + S(\ell_-) - S(\ell_+), \tag{6}$$

which, when set to 0, can be regarded as an equation for p_+ since this latter only enters $\bar{\ell}$. The optimal probability can be given in terms of the function $g(\beta_+, \beta_-) \equiv \tanh\{[S(\ell_-) - S(\ell_+)]\ln 2/b(\beta_+ + \beta_-)\}$ according to the following expression which specifies when the zero of (6) yields a well-defined probability.

$$p_+(\beta_+, \beta_-) = \begin{cases} 0 & g - c < -b\beta_- \\ \frac{g-c+b\beta_-}{b(\beta_+ + \beta_-)} & -b\beta_- \leq g - c < b\beta_+ \\ 1 & g - c \geq b\beta_+ \end{cases}. \tag{7}$$

To determine the optimal values of the z components $\beta_\pm$ of the optimal states, we consider the following quantity and cast it into the form

$$\frac{\partial\chi}{\partial\beta_\pm} = \mp\frac{bp_\pm}{\ln 2}\{\text{arctanh}\bar{\ell} - \text{arctanh}f(\pm\beta_\pm)\}, \tag{8}$$

where we introduce the function $f(\beta) \equiv \tanh\{[c + (b^2 - a^2)\beta/b]\text{arctanh}\ell(\beta)/\ell(\beta)\}$. The right-hand side of (8) vanishes for $p_\pm = 0$ and possibly for values of $\beta_\pm$ such that $\bar{\ell} = f$. One also has to consider the case $\beta_+ = \beta_- = \beta$ for which the Eqs. (8) are replaced by $\frac{\partial\chi}{\partial\beta} = \frac{\partial\chi}{\partial\beta_+} + \frac{\partial\chi}{\partial\beta_-}$.

Depending on the channel parameters a and b and the nonunital parameter c, the optimal states may combine nontrivial values of the occurrence probability p_+ and/or nontrivial values of the z components $\beta_\pm$ as we now describe. For $a > b$, there is one pair of optimal states (labeled A), the northern states ρ_1 and ρ_2, each occurring with a probability 1/2 and characterized by a nontrivial positive z component

$$A: p_+ = 1, \qquad \beta_+ = \beta_A \simeq \frac{f_0 - c}{b - f_1}. \tag{9}$$

In the unital case, the characterization (9) of the optimal states reduces to $p_+ = 1$ and $\beta_A = 0$ which indeed specifies the optimal set obtained for $a > b$ and $c = 0$. It corresponds to a pair of states, diametrically opposed, on the equator of the Bloch sphere. The generalization of theses states to the nonunital case leads to $p_+ = 1$ and a nontrivial z component $\beta_A > 0$. The optimal states ρ_1 and ρ_2 are therefore no longer on a single diameter of the Bloch sphere (see Fig. 1), and are thus nonorthogonal (their scalar product is precisely equal to β_A^2). This observation was already made in Ref. [9] on the basis of a particular example. In the notation of (3) and (4), this example corresponds to $M = \text{diag}(1/3, 1/\sqrt{3}, 0)$ and $\eta = (1/3, 0, 0)$. Actually, as far as the optimal states are concerned, it is strictly equivalent to the case $M = \text{diag}(1/\sqrt{3}, 1/\sqrt{3}, 1/3)$

and $\eta = (0, 0, 1/3)$ which is of the form of (4). The analytical formula (9) given here yields $\arccos\beta_A \simeq 1.5211$ in excellent agreement with the numerical value 1.5218 provided in Ref. [9]. In this formula, $f_0 \equiv \tanh(c\,\mathrm{arctanh}\ell_0/\ell_0)$ and $f_1 \equiv (1 - f_0^2)\{a^2(b^2 - l_0^2)\mathrm{arctanh}\ell_0/b\ell_0 + bc^2/(1-\ell_0^2)\}/\ell_0^2$ with $\ell_0 = \sqrt{a^2 + c^2}$. It is worth mentioning that expression (9) is nonperturbative in c. Indeed, it is not obtained by expanding $f(\beta_A)$ around $c = 0$ but through an expansion around $\beta_A = 0$. This latter expansion is *a priori* expected to be good since in the expression for the output length $\ell(\beta_A)$, the term resulting from the nonunital shift toward the north pole of the Bloch sphere (linear in β_A), partially compensates the term which arises because of the difference in contraction rates along the z axis and in the x-y plane (quadratic in β_A). The Holevo capacity (5) is then

$$\chi_{\mathcal{E}}^{A} = S(b\beta_A + c) - S[\ell(\beta_A)]. \qquad (10)$$

For $a > b$, there may be another type of optimal states (labeled T) depending on the value of the nonunital parameter c (to be specified below). This set consists of three states: the state strictly pointing north ($\rho_1 \equiv \rho_2$) which occurs with a probability p_+ and a pair of states pointing south (ρ_3 and ρ_4) with a nontrivial z component and occurring each with a probability $(1 - p_+)/2$,

$$T: \quad p_+ = p_T \equiv \frac{1}{2} + \frac{g_T - c + b\beta_T}{b(1+\beta_T)} \qquad \beta_- = \beta_T \qquad \beta_+ = 1, \qquad (11)$$

with $g_T = g(1, \beta_T)$ as given explicitly above (7) and $0 \leq \beta_T \leq 1$. A nontrivial value of β_T corresponds to a zero of (8), i.e., satisfies $g_T = f(-\beta_T)$. Examples of such states have already been given in Ref. [11] through numerical calculations in the case where c is kept fixed to its maximal value $c_{\max} = 1 - b$. If c is too low, then there is no such zero of (8) and the optimal set corresponds to the state pointing north and a pair of opposite state on the equator of the Bloch sphere ($\beta_T = 0$). On the other hand, when c increases, the optimal value β_T increases until it reaches unity, in which case the three states become two as we shall see below. See Fig. 1 for an illustration. The Holevo capacity (5) pertaining to the three states (11) is

$$\chi_{\mathcal{E}}^{T} = S(g_T) - p_T S(b + c) - (1 - p_T)S[\ell(-\beta_T)]. \qquad (12)$$

Finally, for $a < b$, the optimal states (labeled B) are the state strictly pointing north ($\rho_1 \equiv \rho_2$) which occurs with a probability p_+ and the state strictly pointing south ($\rho_3 \equiv \rho_4$) occurring with a probability $1 - p_+$,

$$B: \; p_+ = p_B \equiv \frac{1}{2} + \frac{g_B - c}{2b}, \qquad \beta_\pm = 1, \qquad (13)$$

with $g_B = \tanh\{[S(b - c) - S(b + c)]/2b\log_2 e\}$. Indeed, there is no extremal value of $\beta_\pm$ since $\frac{\partial\chi}{\partial\beta_\pm} > 0$ for any value of c. It follows that $\beta_\pm = 1$ is the supremum of χ while the optimal probability p_+ is given by the second line of (7). Note that the average output length $\bar{\ell} = g_B$. The Holevo capacity (5) is then

$$\chi_{\mathcal{E}}^{B} = 1 - \log_2\sqrt{1 - g_B^2} - \frac{b-c}{2b}S(b+c) - \frac{b+c}{2b}S(b-c). \qquad (14)$$

The states B provide a nontrivial generalization of the optimal states occurring in the unital case when b is the largest channel parameter (hence, the label B). Indeed, for $c = 0$, one recovers $p_+ = 1/2$ and $\beta_\pm = 1$ whereas $\chi_{\mathcal{E}}^{B}$ reduces to $1 - S(b)$. For $c > 0$, we deduce from (13) that the occurrence probability of the northern state in the optimal coding is the largest one ($p_+ > 1/2$) and such that the length of the output average Bloch vector is larger than the nonunital shift ($\bar{\ell}_B > c$). This is illustrated in Fig. 1.

Having described the three types of optimal sets, we now specify which one occurs for given shrinking factors a and b and nonunital parameter c. The surfaces delimiting the three types of optimal states can be determined analytically and allow us to construct a "phase" diagram. First, note that the three states of the optimal set T merge with the two states of set B when β_T approaches 1. From (8), we may deduce that this transition occurs when a is given by the following exact expression

$$a_{TB}(b, c) = \sqrt{(b - c)\left(b + \frac{S(b-c) - S(b+c)}{2\mathrm{arctanh}(b-c)\log_2 e}\right)}. \qquad (15)$$

This function is equal to b in the unital case and increases monotonically with c. The nonunitality can thus be seen as giving rise to a larger effective value of b since the domain where the type B states are optimal extends with c. Second, in a similar vein, one can give an approximate expression for the transition surface $a_{AT}(b, c)$ between the A and T types of optimal states from the requirement that β_T vanishes.

A cut in the "phase" diagram is displayed in Fig. 2 for $b = 1/2$. For $a < a_{TB}(b, c)$, the optimal states are of type B. They are of type T for $a_{TB}(b, c) < a < a_{AT}(b, c)$. For $a > a_{AT}(b, c)$, the optimal states are of type A. Notice from Fig. 2 that the transition lines increase with c. For $c = 0$,

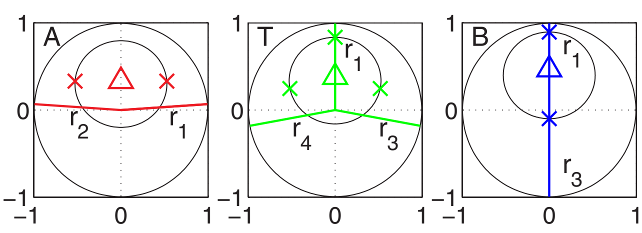

FIG. 1 (color online). Optimal states (solid line) for $a = 0.53$, $b = 0.5$, and, respectively, $c = 0.3$, $c = 0.34$, and $c = 0.4$. The unit circle is the intersection of the Bloch sphere with an arbitrary plane containing the (vertical) z axis. It is mapped by the nonunital channel onto the inner ellipse. The crosses correspond to the image of the optimal states and the triangle to the image of their average.

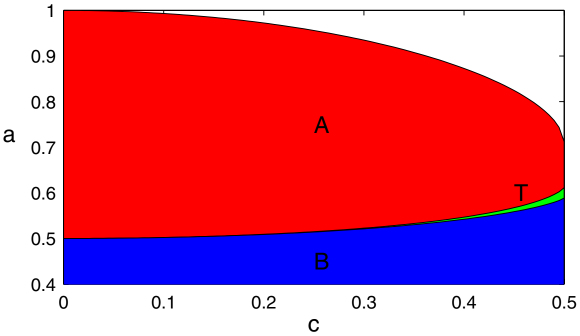

FIG. 2 (color online). Types of optimal states as a function of a and c for $b = 0.5$. The boundaries of this "phase" diagram are determined analytically [cf. below (4) and (5)].

they merge, entailing that in the unital case there is no optimal states of type T and that the transition between types A and B occurs for $a = b$ as is well known for Pauli channels.

When the parameter c controlling the nonunitality is varied, there can be transitions between the different types of optimal states for given contraction rates a and b. This is illustrated in Fig. 2 and detailed in Fig. 3 for $a = 0.53$ and $b = 0.5$. The occurrence probability of each optimal state and their z component are depicted in the lower panels. Notice that the Holevo capacity (upper panel) increases with the nonunitality, and even faster after each transition.

To conclude, the states achieving the Holevo capacity and their transitions have been derived analytically for a class of dissipative channels which is sufficiently large to include the amplitude damping channel [10] as well as other nonunital channels considered in the literature [9–11]. Here, we consider a possibly variable dissipative

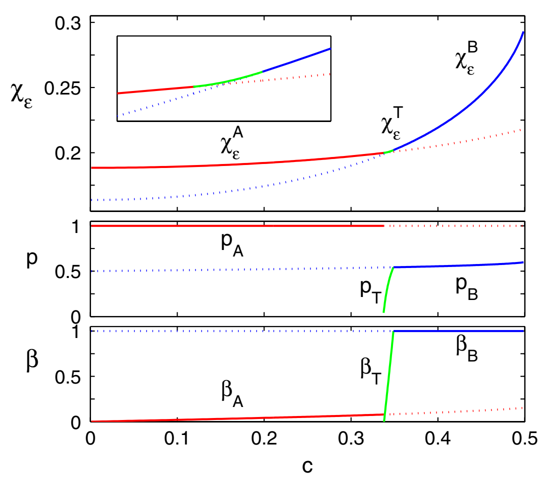

FIG. 3 (color online). Holevo capacity ($\chi_\mathcal{E}$), total probability of occurrence of the pair of northern states (p), and z component of their Bloch vector (β) as a function of c for $a = 0.53$, $b = 0.5$. The corresponding analytical expressions are given by (9)–(14). The dotted lines are displayed to help visualize the cross-over between the optimal states of type A and B.

component of the noise and may regard it as an added noise which allows one to obtain more efficient optimal states for the coding than in its absence. This situation is somehow analogous to a scheme introduced in quantum cryptography where the performance of some protocols is increased if one of the parties adds noise to the measurement data before the error correction [13].

Actual physical systems typically involve both dissipative effects modeled through nonunitality, as studied here, and a memory effect between successive uses of the channels [5–8,14]. Another application is to enhance the fidelity of quantum teleportation. Indeed, it has been shown to increase by subjecting one of the parties to an amplitude damping channel [15,16]. The dissipative channels studied here are more general and will thus also lead to, at least, such a fidelity enhancement.

The author acknowledges stimulating discussions with N. J. Cerf and support from the EU projects QAP and COMPAS.

*ddaems@ulb.ac.be

[1] A. S. Holevo, Probl. Inf. Transm. **9**, 177 (1973); B. Schumacher and M. D. Westmoreland, Phys. Rev. A **56**, 131 (1997); A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).
[2] G. G. Amosov, A. S. Holevo, and R. F. Werner, Probl. Inf. Transm. **36**, 25 (2000); C. King and M. B. Ruskai, IEEE Trans. Inf. Theory **47**, 192 (2001).
[3] A. S. Holevo, arXiv:quant-ph/0212025; J. Cortese, Phys. Rev. A **69**, 022302 (2004).
[4] N. Datta and M. B. Ruskai, J. Phys. A **38**, 9785 (2005).
[5] Ch. Macchiavello and G. M. Palma, Phys. Rev. A **65**, 050301(R) (2002); Ch. Macchiavello, G. M. Palma, S. Virmani, Phys. Rev. A **69**, 010303(R) (2004).
[6] G. Bowen and S. Mancini, Phys. Rev. A **69**, 012306 (2004); D. Kretschmann and R. F. Werner, Phys. Rev. A **72**, 062323 (2005).
[7] E. Karpov, D. Daems, and N. J. Cerf, Phys. Rev. A **74**, 032320 (2006); D. Daems, Phys. Rev. A **76**, 012310 (2007).
[8] M. B. Plenio and S. Virmani, Phys. Rev. Lett. **99**, 120504 (2007).
[9] C. A. Fuchs, Phys. Rev. Lett. **79**, 1162 (1997).
[10] B. Schumacher and M. D. Westmoreland, Phys. Rev. A **63**, 022308 (2001).
[11] C. King, M. Nathanson, and M. B. Ruskai, Phys. Rev. Lett. **88**, 057901 (2002).
[12] M. Hayashi, H. Imai, K. Matsumoto, M. B. Ruskai, and T. Shimono, Quantum Inf. Comput. **5**, 13 (2005).
[13] R. Renner, N. Gisin, and B. Kraus, Phys. Rev. A **72**, 012332 (2005).
[14] F. Caruso, V. Giovannetti, C. Macchiavello, and M. B. Ruskai, Phys. Rev. A **77**, 052323 (2008).
[15] P. Badziag, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **62**, 012311 (2000).
[16] Y. Yeo, Phys. Rev. A **78**, 022334 (2008).