

Parametric separation of symmetric pure quantum states

M. A. Solís-Prosser* and A. Delgado

*Center for Optics and Photonics, Universidad de Concepción, Casilla 4016, Concepción, Chile,
MSI-Nucleus on Advanced Optics, Universidad de Concepción, Casilla 160-C, Concepción, Chile
and Departamento de Física, Universidad de Concepción, Casilla 160-C, Concepción, Chile*

O. Jiménez

Departamento de Física, Facultad de Ciencias Básicas, Universidad de Antofagasta, Casilla 170, Antofagasta, Chile

L. Neves

Departamento de Física, Universidade Federal de Minas Gerais, Belo Horizonte, MG 30123-970, Brazil
(Received 27 October 2015; published 22 January 2016)

Quantum state separation is a probabilistic map that transforms a given set of pure states into another set of more distinguishable ones. Here we investigate such a map acting onto uniparametric families of symmetric linearly dependent or independent quantum states. We obtained analytical solutions for the success probability of the maps—which is shown to be optimal—as well as explicit constructions in terms of positive operator valued measures. Our results can be used for state discrimination strategies interpolating continuously between minimum-error and unambiguous (or maximum-confidence) discrimination, which, in turn, have many applications in quantum information protocols. As an example, we show that quantum teleportation through a nonmaximally entangled quantum channel can be accomplished with higher probability than the one provided by unambiguous (or maximum-confidence) discrimination and with higher fidelity than the one achievable by minimum-error discrimination. Finally, an optical network is proposed for implementing parametric state separation.

DOI: [10.1103/PhysRevA.93.012337](https://doi.org/10.1103/PhysRevA.93.012337)

I. INTRODUCTION

The physical maps on quantum states are trace-preserving completely positive maps [1,2]. These can be explicitly designed to accomplish a predefined task while simultaneously obeying a set of constraints. For instance, some maps are designed in such a way that they transform a given set of input states onto a fixed set of output states. Two classic examples of this are probabilistic cloning and state discrimination. In the former, a probabilistic map transforms linearly independent states $|\psi_i\rangle$ into states $|\psi_i\rangle|\psi_i\rangle$ [3], which contains perfect clones. The mapping is constructed to achieve the highest cloning probability. The discrimination of quantum states is realized by mapping nonorthogonal states $|\psi_i\rangle$ to be discriminated onto a set of more distinguishable states, possibly orthogonal ones such as $\{|i\rangle\}$. In this process two different quantities can be optimized: the success probability of the identification and the error in the identification of the states (or confidence). This leads to several discrimination protocols, such as unambiguous state discrimination [4–9], maximum confidence [10], and minimum error [11,12], among others.

Here, we investigate probabilistic maps between two sets of pure quantum states. In particular, we focus on maps acting onto uniparametric families of symmetric linearly dependent or independent quantum states. These maps are conclusive: It is possible to know with certainty whether the mapping was successful or not. We obtain analytical solutions for the success probability of the maps as well as explicit constructions in terms of positive operator valued measures (POVMs).

We also show that the success probability attained by these mappings is optimal. It has been shown recently [13] that any discrimination strategy interpolating between minimum-error discrimination and unambiguous (or maximum-confidence, in the case of linearly dependent states) discrimination can be interpreted as a minimum-error implementation preceded by a map between sets of quantum states. The uniparametric families of symmetric states studied here are such that an increase in the value of the parameter leads to an increase in the distinguishability of the states. Consequently, these families of states can be used to interpolate continuously between minimum-error discrimination and unambiguous (or maximum-confidence) discrimination. For this reason we term these mappings as *parametric state separation*. We show that this class of maps finds application in quantum teleportation [14] when it is sufficient to teleport through a poorly entangled quantum channel with a fidelity of transmission higher than a given threshold. By using the map proposed here it is possible to achieve the threshold with a higher probability than the one provided by unambiguous state discrimination (or maximum confidence) [15–17] and with a higher fidelity than the achievable by minimum-error discrimination. Application to other protocols such as entanglement swapping and dense coding are also feasible. Finally, we propose a simple optical network to implement parametric state separation.

This paper is organized as follows: Section II proposes a parametrization for symmetric states intended to interpolate between minimum-error and maximum-confidence (or unambiguous) discrimination, that is, the process of parametric state separation. Section III analyzes the proposed map and shows a closed-form solution for the optimization of the success probability. Section IV displays applications of parametric

*msolisp@udec.cl

state separation, such as assisting quantum state teleportation and entanglement swapping. Section V proposes an optical network able to perform parametric state separation and Sec. VI concludes this paper.

II. PARAMETRIC STATE SEPARATION

In this work, we address the topic of theory and applications of maps that allow one to interpolate between a probabilistic discrimination strategy, as unambiguous or maximum-confidence discrimination, and a deterministic one (e.g., minimum-error discrimination). Since closed-form solutions do not exist for the vast majority of cases, as it also occurs in quantum state discrimination, we shall inspect the context of equally likely symmetric pure states, which has closed-form solutions for several discrimination strategies [18–22]. For this purpose, let us consider a set of N pure quantum states $\{|\psi_j\rangle, j = 0, \dots, N-1\}$ spanning a D -dimensional Hilbert space, with $N \geq D$. These states are *symmetric* under the action of a unitary operator \hat{W} if they satisfy

$$|\psi_j\rangle = \hat{W}|\psi_{j-1}\rangle = \hat{W}^j|\psi_0\rangle, \quad (1)$$

$$|\psi_0\rangle = \hat{W}|\psi_{N-1}\rangle. \quad (2)$$

For every unitary operator \hat{W} , there exists a set of phases and eigenvectors $\{(\zeta_k, |\gamma_k\rangle), k = 0, \dots, N-1\}$ such that

$$\hat{W} = \sum_{k=0}^{N-1} e^{i\zeta_k} |\gamma_k\rangle\langle\gamma_k|. \quad (3)$$

The state $|\psi_0\rangle$ is sometimes known as the *fiducial* state of the set. The condition $\hat{W}|\psi_{N-1}\rangle = |\psi_0\rangle$ implies that $\hat{W}^N = \hat{I}$. For this to be fulfilled, it is necessary that $N\zeta_k = 2\pi n_k, n_k \in \mathbb{Z}$. Then, every unitary operator \hat{W} able to generate a set of N symmetric D -dimensional states can be written as

$$\hat{W} = \sum_{k=0}^{N-1} e^{2\pi i n_k / N} |\gamma_k\rangle\langle\gamma_k|. \quad (4)$$

As remarkable examples, the eigenstates of the phase gates $\hat{Z}|l\rangle = e^{2\pi i l / N} |l\rangle$ correspond to $|\gamma_k\rangle = |k\rangle$, and the ones of the shift gates $\hat{X}|l\rangle = |l \oplus 1\rangle$ correspond to the inverse Fourier transform of the computational basis, i.e., $|\gamma_k\rangle = \hat{F}^{-1}|k\rangle$.

Now, let us define two sets of N symmetric states spanning a D -dimensional Hilbert space, defined by

$$|\alpha_j\rangle = \sum_{k=0}^{N-1} a_k e^{i\phi_k} \omega^{jk} |\mathbf{f}_k\rangle, \quad (5)$$

$$|\beta_j\rangle = \sum_{k=0}^{N-1} b_k e^{i\varphi_k} \omega^{jk} |\mathbf{g}_k\rangle, \quad (6)$$

where $\omega = \exp(2\pi i / N)$ is an N th complex root of the unity, a_k and b_k are real non-negative coefficients where $N - D$ of them adopt null values, ϕ_k and φ_k are real numbers representing phases, $a_k e^{i\phi_k}$ are the coefficients of the fiducial state of the first set, $b_k e^{i\varphi_k}$ are those ones of the second set, $|\mathbf{f}_k\rangle$ and $|\mathbf{g}_k\rangle$ are orthonormal eigenbases for the operators that generate each set. The first set is symmetric under $\hat{Z}_\alpha = \sum_{k=0}^{N-1} \omega^k |\mathbf{f}_k\rangle\langle\mathbf{f}_k|$, and the second one is symmetric

under $\hat{Z}_\beta = \sum_{k=0}^{N-1} \omega^k |\mathbf{g}_k\rangle\langle\mathbf{g}_k|$. Although the operators \hat{Z}_α and \hat{Z}_β were written as phase operators, this representation holds only when they are written in their respective eigenbases.

The general task of finding an operation Λ such that $\Lambda(|\alpha_j\rangle) \propto |\beta_j\rangle$ has been studied. Currently, results about existence and feasibility of these physical operations are known [1,2,23–25]. However, closed-form solutions are known for sets of two states only [3,26,27]. Additionally, it was proven that optimal maps between sets of symmetric states can be found by solving a linear programming problem [28]. This optimization problem can always be numerically solved. Nevertheless, closed-form solutions and details about its implementation, i.e., the corresponding Kraus operators, must be studied for every particular case.

Suppose we have two sets of N pure states defined by

$$|\alpha_j\rangle = \sum_{k=0}^{N-1} a_k e^{i\phi_k} \omega^{jk} |k\rangle, \quad (7)$$

$$|u_j\rangle = \frac{1}{\sqrt{D}} \sum_{k=0}^{N-1} y_k \omega^{jk} |k\rangle, \quad (8)$$

where y_k is a binary parameter such that

$$y_k = \begin{cases} 1 & \text{if } a_k \neq 0 \\ 0 & \text{if } a_k = 0 \end{cases}, \quad (9)$$

and $D = \sum_{k=0}^{N-1} y_k$, where D is the dimension of the Hilbert space in which states $|\alpha_j\rangle$ lie and is also the number of nonvanishing coefficients a_k . The states $|u_j\rangle$ constitute a set of D orthonormal states when $N = D$. Otherwise, they compose a set with maximal distinguishability within the D -dimensional subspace where they belong. Optimal operations transforming states $|\alpha_j\rangle$ onto states $|u_j\rangle$ have been studied [18–20]. For instance, the problem of discriminating within an ensemble of equally likely symmetric states was studied via unambiguous discrimination [4–8] for linearly independent states [9] and maximum confidence [10] for linearly dependent states.

Here, we consider maps from the set of states $|\alpha_j\rangle$ to the set of states $|\beta_j(\xi)\rangle$ defined as

$$|\beta_j(\xi)\rangle = \sum_{k=0}^{N-1} b_k(\xi) e^{i\varphi_k(\xi)} \omega^{jk} |k\rangle, \quad (10)$$

where

$$\begin{aligned} b_k(\xi)^2 &= \left[a_k^2 + \left(\frac{1}{D} - a_k^2 \right) \xi \right] y_k \\ &= (1 - \xi) a_k^2 + \frac{y_k \xi}{D} \end{aligned} \quad (11)$$

are the coefficients of the fiducial output states. As $|\beta_j(0)\rangle = |\alpha_j\rangle$ and $|\beta_j(1)\rangle = |u_j\rangle$, then ξ is a parameter that indicates a point between the sets $|\alpha_j\rangle$ and $|u_j\rangle$ in which the states $|\beta_j(\xi)\rangle$ are lying on, with $0 \leq \xi \leq 1$. The phase $\varphi(\xi)$ must equal ϕ_k for $\xi = 0$, and 0 for $\xi = 1$. A simple linear

formula, which will be used here,¹ is $\varphi_k(\xi) = (1 - \xi)\phi_k$. Note, from Eq. (11), that we have considered $a_k y_k = a_k$, as can be inferred from the definition of Eq. (9). The way these numbers were defined is based upon the parametric definition of a line segment between two points, in which the endpoints are a_k^2 and $1/D$. This definition for $b_k(\xi)$ preserves the norm of the states and the ordering of the coefficients, i.e., $a_k \geq a_l$ implies that $b_k(\xi) \geq b_l(\xi)$ and $a_m = 0$ implies that $b_m(\xi) = 0$, for every value of ξ in the domain.

A procedure that implements the mapping of states $|\alpha_j\rangle_s$ onto states $|\beta_j(\xi)\rangle_s$ is given by [29]

$$\begin{aligned} \hat{U}_{sa}(\xi)|\alpha_j\rangle_s|\mathcal{A}\rangle_a &= \hat{A}_S(\xi)|\alpha_j\rangle_s|0\rangle_a + \hat{A}_F(\xi)|\alpha_j\rangle_s|1\rangle_a, \\ &= \sqrt{p_S(\xi)}|\beta_j(\xi)\rangle_s|0\rangle_a \\ &\quad + \sqrt{1 - p_S(\xi)}|\tilde{\beta}_j(\xi)\rangle_s|1\rangle_a, \end{aligned} \quad (12)$$

where $|0\rangle_a$ and $|1\rangle_a$ are two orthogonal states from a two-dimensional ancillary system, $\hat{U}_{sa}(\xi)$ is a unitary operator acting on the bipartite Hilbert space of the system (s) and the ancilla (a), $|\mathcal{A}\rangle_a$ is the initial state of the ancilla and $p_S(\xi)$ is the probability of achieving the desired mapping. A measurement on the ancilla performed afterwards can announce whether the map was successful. In this case, it has been assumed the success probability to be the same for all $|\alpha_j\rangle$ states. The *uniformization lemma* of Ref. [28] supports this assumption. The Kraus operators $\hat{A}_S(\xi)$ and $\hat{A}_F(\xi)$ represent the process of mapping successfully and unsuccessfully, respectively. Since $\hat{U}(\xi)$ is a unitary operation, these Kraus operators must satisfy $\hat{A}_S^\dagger(\xi)\hat{A}_S(\xi) + \hat{A}_F^\dagger(\xi)\hat{A}_F(\xi) = \hat{I}$, where \hat{I} is the identity operator acting on the Hilbert space of the symmetric states. The Kraus operator related to successful events can be written as

$$\hat{A}_S(\xi) = \sqrt{p_S} \sum_{k=0}^{N-1} \frac{b_k(\xi)}{a_k} e^{-i\xi\phi_k} |k\rangle\langle k|, \quad (13)$$

where only the terms with nonvanishing coefficients of the fiducial state contribute. The probability of success can be maximized under the constraint that $\hat{A}_F^\dagger(\xi)\hat{A}_F(\xi) = \hat{I} - \hat{A}_S^\dagger(\xi)\hat{A}_S(\xi) \geq 0$. This constraint ensures the non-negativity of the probability of failure given by $p_F = 1 - p_S$. As a result,

$$p_S(\xi) = \min_{\{k\}} \left[\frac{a_k^2}{b_k^2(\xi)} \right] = \frac{1}{(1 - \xi) + \frac{\xi}{Da_{\min}^2}}, \quad (14)$$

where $a_{\min} = \min\{a_j | j = 0, \dots, N-1 \wedge a_j \neq 0\}$. A graph of the success probability is shown in Fig. 1 in terms of the smallest coefficient of the initial fiducial state and number of nonvanishing coefficients (y axis) and the target family of states determined by ξ (x axis). Additionally, it can be shown that

$$b_{\min}^2(\xi) = (1 - \xi)a_{\min}^2 + \xi/D \quad (15)$$

¹Other choices for $\varphi_k(\xi)$ are possible. For instance, $\varphi_k(\xi) = -\phi_k$, which immediately eliminates the phases of the fiducial state no matter the value of ξ .

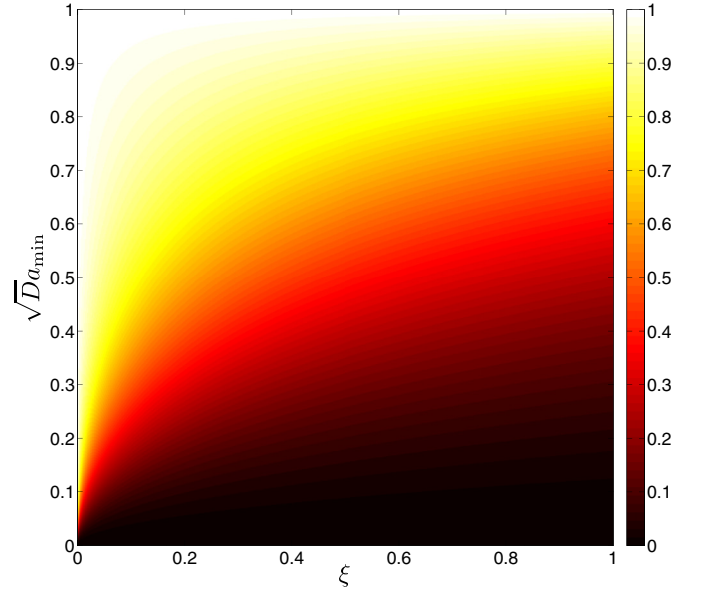


FIG. 1. Success probability $p_S(\xi)$ of Eq. (14) for parametric state separation as a function of the minimum coefficient a_{\min} (scaled by the number of nonvanishing fiducial coefficients D) and the ξ parameter that defines the target set of states.

is the smallest of the nonvanishing $b_k^2(\xi)$ coefficients. The Kraus operators $\hat{A}_S(\xi)$ and $\hat{A}_F(\xi)$ can then be written as

$$\hat{A}_S(\xi) = \sum_{k=0}^{N-1} \sqrt{\frac{1 - \xi + \xi/Da_k^2}{1 - \xi + \xi/Da_{\min}^2}} y_k e^{-i\xi\phi_k} |k\rangle\langle k|, \quad (16)$$

$$\begin{aligned} \hat{A}_F(\xi) &= \sum_{k=0}^{N-1} \left[\sqrt{\frac{\xi}{D} \frac{1/a_{\min}^2 - 1/a_k^2}{1 - \xi + \xi/Da_{\min}^2}} e^{-i\xi\phi_k} y_k \right. \\ &\quad \left. + (1 - y_k) \right] |k\rangle\langle k|. \end{aligned} \quad (17)$$

The action of these operators on the states $|\alpha_j\rangle$ is given by

$$\hat{A}_S(\xi)|\alpha_j\rangle = \sqrt{p_S(\xi)}|\beta_j(\xi)\rangle, \quad (18)$$

$$\hat{A}_F(\xi)|\alpha_j\rangle = \sqrt{1 - p_S(\xi)}|\tilde{\beta}_j(\xi)\rangle, \quad (19)$$

where

$$|\tilde{\beta}_j(\xi)\rangle = \sum_{k=0}^{N-1} \sqrt{\frac{a_k^2 - a_{\min}^2 y_k}{1 - Da_{\min}^2}} \omega^{jk} e^{i(1-\xi)\phi_k} |k\rangle. \quad (20)$$

Recapitulating, once the bipartite unitary operation has been carried out, the user must perform a measurement on the ancilla. If the outcome is $|0\rangle_a$ ($|1\rangle_a$), the state $|\alpha_j\rangle$ has been mapped on $|\beta_j(\xi)\rangle$ ($|\tilde{\beta}_j(\xi)\rangle$) and the process is considered successful (failed). The states $|\beta_j(\xi)\rangle$ and $|\tilde{\beta}_j(\xi)\rangle$ are given by Eqs. (10) and (20), respectively. It is worth mentioning that, although the components $|\langle k|\beta_j(\xi)\rangle| = b_k(\xi)$ of the successfully transformed states depend on ξ , the coefficients $|\langle k|\tilde{\beta}_j(\xi)\rangle|$ are not functions of ξ . Additionally, there will always be at least one value of k , say k_0 , such that $a_{k_0} = a_{\min}$ and, correspondingly, $b_{k_0}(\xi) = b_{\min}(\xi)$, making the

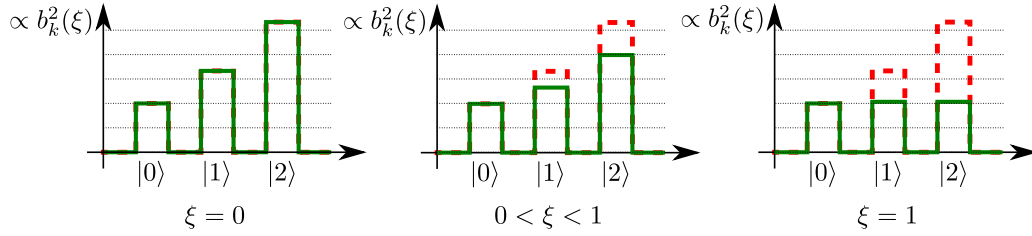


FIG. 2. Simplified depiction of the coefficients during parametric state separation. The dashed line represents the initial coefficients a_k^2 , which are equal to $b_k^2(0)$. The continuous line depicts the values of $b_k^2(\xi)$ for $\xi > 0$, showing the relevant changes on these probabilities.

k_0 th coefficient of $|\tilde{\beta}_j(\xi)\rangle$ vanish in the case of a failed attempt. Then, the smallest coefficient(s) is (are) not transferred to the failure states, which are even less distinguishable among them than the original set $\{|\alpha_j\rangle\}$ since it is the same number of vectors distributed over a smaller subspace.

It is useful to interpret these results in terms of amplitude modulation. Figure 2 sketches this topic through a three-dimensional example. This analysis can be applied for every physical system able to encode a set of symmetric states, where the rectangle functions represent the squared amplitude of, for instance, the energy levels of an atom, the paths of a multiport interferometer, an array of slits, among other examples. It can be seen that the smallest coefficient keeps unaltered and the larger coefficients are attenuated until they reach the value of the smaller one according to the parameter ξ increases. The closer ξ is to 1, the smaller the difference among the coefficients. As can be observed from Fig. 2, the probabilities for each basis component is reduced according to ξ increases. The difference between the original probabilities of the initial states and the final ones, for each $\xi > 0$, will construct the unsuccessful case and, consequently, the coefficients of failure states $|\tilde{\beta}_j(\xi)\rangle$. This is consistent with Eq. (20) as they have null probability in the basis vector corresponding to the smallest coefficient.

A. Distinguishability measures

The map of states $|\alpha_j\rangle$ onto $|\beta_j(\xi)\rangle$ defined by Eqs. (7) and (10), respectively, is called parametric state separation, with the parametrization defined by Eq. (11). A natural question that arises here concerns the distinguishability of the resulting states $|\beta_j(\xi)\rangle$. From the geometrical point of view, two vectors have been *separated* if the angle between them at the end is larger than at the beginning. For more states, we must resort to a distinguishability measure in order to show that the parametrization given by Eq. (11) makes the transformed states more distinguishable than the initial ones and the opposite behavior is not occurring. There exist several quantities to measure distinguishability among the states belonging to a given set. In this work, we will make use of two of them. The first one is the optimal probability of having a successful discrimination via unambiguous discrimination or maximum confidence—according to their linear dependence—among these states, which we shall define as $\mathcal{D}_1(\xi)$. The second one is the optimal probability of correctly discriminating these states by means of optimal minimum-error discrimination, denoted as $\mathcal{D}_2(\xi)$ throughout this section. Since the states we are dealing with are symmetric, these two quantities

are known [18,19,21]. The use of at least two different measures is necessary because there exist sets of states whose ordering, by distinguishability, is different depending on which measure is used [30]. For this reason, there is not an absolute distinguishability measure within an ensemble of quantum states and, consequently, the use of only one measure is not reliable.

Figure 3 shows an example of these measures applied on a given fiducial state. As expected, both increase according to ξ adopting larger values. But, beyond a single example, a more rigorous proof of the increased distinguishability of states $|\beta_j(\xi)\rangle$ with respect to $|\alpha_j\rangle$ is necessary. According to the first measure and taking Eq. (15) into account, we obtain [18–20]

$$\mathcal{D}_1(\alpha) = Da_{\min}^2, \quad (21)$$

$$\begin{aligned} \mathcal{D}_1(\beta(\xi)) &= Db_{\min}^2(\xi) \\ &= \mathcal{D}_1(\alpha) + (1 - Da_{\min}^2)\xi. \end{aligned} \quad (22)$$

Since $a_{\min} \leq 1/\sqrt{D}$, the second term of the right-hand side of Eq. (22) is always positive. Then, for every value of $\xi \in [0, 1]$, it occurs that $\mathcal{D}_1(\beta(\xi)) \geq \mathcal{D}_1(\alpha)$ and the states $|\beta_j(\xi)\rangle$ are more distinguishable than the states $|\alpha_j\rangle$ according to the first

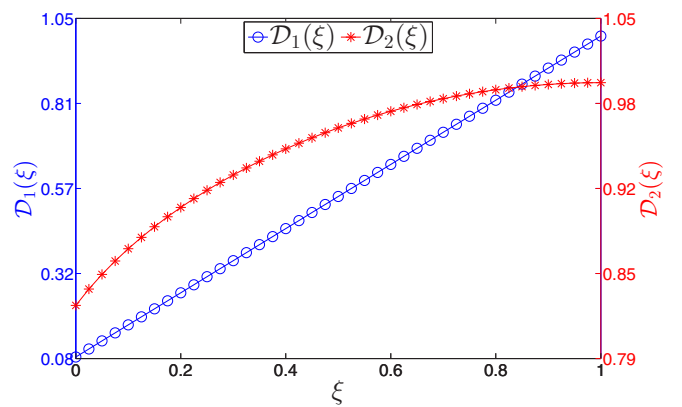


FIG. 3. Example of distinguishability measures behavior according to ξ increases. For this example, $N = D = 5$, $a_0 = 0.6386$, $a_1 = 0.5841$, $a_2 = 0.3817$, $a_3 = 0.1321$, and $a_4 = 0.2964$. Left vertical axis represents the first distinguishability measure given by Eq. (22), which is shown in the graph as circles. Conversely, right vertical axis stands for the second measure defined by Eq. (24), represented in the graph as stars. For this particular example, both measures are increasing functions of ξ . A proof of this behavior for any set of initial states can be found in the main text.

measure function. Moreover, distinguishability increases as ξ grows.

On the other hand, the second measure is the probability of optimal minimum-error discrimination [11,12], given by [21,22]

$$\mathcal{D}_2(\alpha) = \frac{1}{N} \left(\sum_{k=0}^{N-1} a_k \right)^2, \quad (23)$$

$$\mathcal{D}_2(\beta(\xi)) = \frac{1}{N} \left(\sum_{k=0}^{N-1} b_k(\xi) \right)^2. \quad (24)$$

The behavior of the second distinguishability measure can be studied if we calculate its derivatives with respect to ξ . For this purpose, we shall define $\mathcal{E}(\xi) = \sqrt{N\mathcal{D}_2(\beta(\xi))}$, whose behavior is simpler to analyze. So, from Eqs. (9) and (11),

$$\begin{aligned} \mathcal{E}(\xi) &= \sum_{k=0}^{N-1} b_k(\xi) \\ &= \sum_{k=0}^{N-1} y_k \sqrt{(1-\xi)a_k^2 + \xi/D}. \end{aligned} \quad (25)$$

Then,

$$\frac{d}{d\xi} \mathcal{E}(\xi) = \sum_{k=0}^{N-1} y_k \frac{1/D - a_k^2}{2\sqrt{(1-\xi)a_k^2 + \xi/D}}, \quad (26)$$

$$\frac{d^2}{d\xi^2} \mathcal{E}(\xi) = - \sum_{k=0}^{N-1} y_k \frac{(1/D - a_k^2)^2}{4[(1-\xi)a_k^2 + \xi/D]^{3/2}} \leq 0, \quad \forall \xi. \quad (27)$$

We can see that the graph of $\mathcal{E}(\xi)$ is concave downwards since its second derivative is always negative. Although it is not easy to find critical points from its first derivative, we can observe for $\xi = 1$ that

$$\begin{aligned} \frac{d}{d\xi} \mathcal{E}(1) &= \sum_{k=0}^{N-1} y_k \frac{1/D - a_k^2}{2\sqrt{(1-1)a_k^2 + 1/D}} \\ &= \frac{\sqrt{D}}{2} \left(\frac{1}{D} \sum_{k=0}^{N-1} y_k - \sum_{k=0}^{N-1} a_k^2 \right) = 0. \end{aligned} \quad (28)$$

Then, $\xi = 1$ is a critical point. The second derivative in this point is negative, so this is a maximum. This is expected because $\xi = 1$ leads to orthogonal states when $N = D$. Since $\mathcal{E}(\xi)$ is concave downwards and has its maximum in $\xi = 1$, we can ensure that $\mathcal{E}(\xi)$ is monotonically increasing for $0 \leq \xi \leq 1$ and, in consequence, $\mathcal{D}_2 = \mathcal{E}^2(\xi)/N$ is also an increasing function. Therefore, this distinguishability measure increases accordingly as ξ does and $\mathcal{D}_2(\beta(\xi)) \geq \mathcal{D}_2(\beta(0)) = \mathcal{D}_2(\alpha)$.

As shown in this section, there exist at least two distinguishability measures that consider the states $|\beta_j(\xi)\rangle$ as more distinguishable than the $|\alpha_j\rangle$. Then, the parametrization proposed in Eq. (10) is correct in terms of the initial goal of an intermediate map between the original states $|\alpha_j\rangle$ and the states $|u_j\rangle$ of Eq. (8). This opens the possibility of considering alternative strategies in state discrimination equivalent to

perform a map onto a more distinguishable set followed by minimum-error discrimination over the final set [13].

III. THE OPTIMAL TRANSFORM

A. Maps between symmetric pure states as a linear programming problem

In order to map from the symmetric states $|\alpha_j\rangle$ to the symmetric ones $|\beta_j\rangle$, we can also regard a more general transformation for pure states, given by [28]

$$\begin{aligned} \hat{U}_{\text{sea}} |\alpha_j\rangle_s |\mathcal{M}\rangle_e |\mathcal{A}\rangle_a &= \sqrt{p} |\beta_j\rangle_s |\psi_j\rangle_e |0\rangle_a \\ &\quad + \sqrt{1-p} |\Phi_j\rangle_{se} |1\rangle_a, \end{aligned} \quad (29)$$

where p is the success probability, $|\psi_j\rangle$ are states generated as a by-product of the process, and $|\Phi_j\rangle_{se}$ are the resulting states when the process fails. This transformation differs from the one of Eq. (12) since, in addition to the main system (s) and ancilla (a), it includes a third system representing the environment (e) or a second auxiliary system. These three-system transforms may allow higher success probability than two-system transforms [25]. It is noteworthy that transforms like the one of Eq. (12) are particular cases of Eq. (29) in which the environment state evolves independently of the other two systems. Dunjko and Andersson showed that if such transform exists, then there exists an equivalent one in which the $|\psi_j\rangle$ states are also symmetric. This result was shown in the *symmetrization lemma* [28]. Additionally, they proved that the transform mapping $|\alpha_j\rangle$ to $|\beta_j\rangle$ with optimal probability of success can be found by solving a linear problem which, when applied to the states studied throughout this paper, can be summarized as

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} = p, \\ \text{subject to:} \quad & \begin{cases} 0 \leq \mathbf{x} \leq \mathbf{1} \\ \mathbf{M} \mathbf{x} \leq \mathbf{n} \end{cases}, \end{aligned} \quad (30)$$

where

$$\begin{aligned} \mathbf{n} &= \sum_{r=0}^{N-1} a_r^2 \vec{\mathbf{e}}_r, \quad \mathbf{c} = \sum_{k=0}^{N-1} \vec{\mathbf{e}}_k, \\ \mathbf{M} &= \sum_{r=0}^{N-1} b_r^2 \hat{\mathbf{X}}^r, \quad \hat{\mathbf{X}} = \sum_{k=0}^{N-1} \vec{\mathbf{e}}_{k+1} \vec{\mathbf{e}}_k^\dagger \end{aligned} \quad (31)$$

are fixed vectors and matrices, $\{\vec{\mathbf{e}}_j\}_{j=0}^{N-1}$ is the N -dimensional computational Euclidean basis, and

$$\mathbf{x} = \sum_{k=0}^{N-1} p \Psi_k^2 \vec{\mathbf{e}}_k \quad (32)$$

is the unknown N -dimensional vector to be found [28]. The parameters Ψ_s are also the coefficients of the state $|\psi_0\rangle$ and they give information about the Kraus operators involved in the transform, as shown in the appendix. A transform in which all of the $|\psi_j\rangle$ states are equal is referred to as a *leakless* transform [28], whose solution has the form $\mathbf{x} = p \vec{\mathbf{e}}_{k_0}$ for a certain k_0 . Examples of leakless transforms include the ones of Eq. (12). Particularly, Eq. (14) is represented by $\mathbf{x} = p_S(\xi) \vec{\mathbf{e}}_0$.

Although linear programming problems like the one of Eq. (30) can always be numerically solved, the search of closed-form solutions for algebraic coefficients and arbitrary dimensions and number of states is not always an easy task to solve. Nevertheless, we were able to prove that the leakless transform of Eq. (12) actually represents the *optimal* procedure to carry out parametric state separation, since any physically acceptable solution different from $\mathbf{x} = p_S(\xi)\mathbf{e}_0$ yields a smaller success probability. The complete proof has been detailed and commented on in the following subsection (Sec. III B). In addition to showing the optimal success probability, these deductions affirm that the assistance of a third system is not required for the implementation of optimal parametric state separation and a two-dimensional ancilla suffices for this purpose. Consequently, as expected, the use of a third system will not outperform unambiguous or maximum-confidence discrimination when carried out assisted by a two-level ancilla.

B. Proof of optimality of $p_S(\xi)$ for parametric state separation

Thousands of numerical simulations showed that the probability shown in Eq. (14) is also the optimal probability that can be obtained by solving the optimization problem of Dunjko and Andersson [Eq. (30)]. In order to show an analytical proof of this fact, let us study a solution in a neighborhood of the one of Eq. (14), which can be written as $\mathbf{x}_0 = p_S\mathbf{e}_0$ for the optimization algorithm since it was obtained from a leakless transform. Then, considering extra parameters κ_0 and ϵ_k , the \mathbf{x} unknown vector can be written as

$$\mathbf{x} = (p_S - \kappa_0)\mathbf{e}_0 + \sum_{k=1}^{N-1} \epsilon_k \mathbf{e}_k, \quad (33)$$

where the new unknowns are κ_0 and ϵ_k (for $k = 1, \dots, N-1$). Bearing in mind, from Eq. (32), that \mathbf{x} has only non-negative coefficients, we must impose that $\kappa_0 \leq p_S$ and $\epsilon_k \geq 0$. Then, the linear optimization problem is now established as

$$\begin{aligned} \max_{\{\kappa_0, \epsilon_k\}} \quad & p = p_S - \kappa_0 + \sum_{k=1}^{N-1} \epsilon_k \\ \text{subject to} \quad & \Delta_k = a_k^2 - \sum_{s=0}^{N-1} b_s^2 x_{k-s} \geq 0, \end{aligned} \quad (34)$$

where all vector indexes must be taken modulo N . By applying the proposal of Eq. (33), we obtain

$$\Delta_k = a_k^2 - (p_S - \kappa_0)b_k^2 - \sum_{\substack{s=0 \\ s \neq k}}^{N-1} b_s^2 \epsilon_{k-s} \geq 0. \quad (35)$$

The positivity of Δ_k from Eq. (35) must be fulfilled for every value of k . Particularly, when $k = k_{\min}$ such that $a_{k_{\min}} = a_{\min}$, we obtain

$$\kappa_0 \geq \sum_{\substack{s=0 \\ s \neq k_{\min}}}^{N-1} \frac{b_s^2}{b_{\min}^2} \epsilon_{k_{\min}-s} \geq \sum_{\substack{s=0 \\ s \neq k_{\min}}}^{N-1} y_s \epsilon_{k_{\min}-s}. \quad (36)$$

For the case of linearly independent states, every coefficient y_s is equal to one. Since $\epsilon_r \geq 0$, either (i) $\kappa_0 = \epsilon_r = 0$, or (ii) $\kappa_0 > 0$. The success probability for the first case is $p_{(i)} = p_S$, whereas for the second case is $p_{(ii)} = p_S - \kappa_0 + \sum_r \epsilon_r$, considering $0 \neq \kappa_0 \geq \sum_r \epsilon_r$. We can observe that

$$p_{(ii)} - p_{(i)} = -\kappa_0 + \sum_r \epsilon_r \leq 0.$$

Then, $p_{(i)} \geq p_{(ii)}$, showing that $\kappa_0 = \epsilon_r = 0$ is an optimal solution, which also satisfies Eq. (35) for $k \neq k_{\min}$.

On the other hand, the case of linearly dependent states must be analyzed separately. Although for this case we have some coefficients a_k equal to zero, Eq. (36) remains valid. Consequently, κ_0 must be a non-negative number. Now, let $Q = \{\mu = 0, \dots, N-1 | y_\mu = 0\}$ be the set of indexes related to vanishing coefficients of the fiducial states. The optimal probability is

$$\begin{aligned} p &= p_S - \kappa_0 + \sum_{\substack{s \notin Q \\ s \neq k_{\min}}} \epsilon_{k_{\min}-s} + \sum_{\mu \in Q} \epsilon_{k_{\min}-\mu} \\ &\leq p_S + \sum_{\mu \in Q} \epsilon_{k_{\min}-\mu}, \end{aligned} \quad (37)$$

where the inequality of Eq. (36) was applied. Let $Q = N - D$ be the number of elements of Q . As it transpires from Eq. (37), the bounds on the optimal probability now depend on Q parameters ϵ_r . For $\mu \in Q$, Eq. (35) becomes

$$\Delta_\mu = - \sum_{\substack{s=0 \\ s \neq q}}^{N-1} b_s^2 \epsilon_{\mu-s} = - \sum_{s \notin Q} b_s^2 \epsilon_{\mu-s} \geq 0, \quad (38)$$

which can be satisfied only if $\forall s \notin Q, \epsilon_{\mu-s} = 0$. In other words, $y_s \neq 0$ implies that $\epsilon_{\mu-s} = 0$, setting the value of D of the total of $N-1$ available ϵ_r as zero. When $N = D+1$, this condition imposes all ϵ_r must be zero, preventing the optimal probability of Eq. (37) from being greater than p_S . Similarly, for $Q = 2$ we have $Q = \{\mu_1, \mu_2\}$. Thus, $\Delta_{\mu_1} \geq 0$ implies that only $\epsilon_{\mu_1-\mu_2}$ *might* differ from zero and $\Delta_{\mu_2} \geq 0$ indicates that only $\epsilon_{\mu_2-\mu_1}$ *might* not vanish. Together, these two statements suggest that every ϵ_r , including $\epsilon_{\mu_2-\mu_1}$ and $\epsilon_{\mu_1-\mu_2}$, are zero *unless* $\mu_1 - \mu_2 \equiv \mu_2 - \mu_1 \pmod{N}$, for which no information is given. Although interesting, this last case does not exert influence on the upper bound of Eq. (37) since it contains only terms of the type $\epsilon_{k_{\min}-\mu}$. Reminding that $k_{\min} \notin Q$, these terms in the upper bound vanish as a consequence of $\Delta_{\mu_1} \geq 0$ and $\Delta_{\mu_2} \geq 0$, as aforementioned. Then, the optimal probability cannot surpass p_S . In the same way, Eq. (38) for $Q > 2$ constructs a set of Q inequalities, each one imposing D of the ϵ_r to adopt null value. For every $\mu_j \in Q$, $\Delta_{\mu_j} \geq 0$ informs that only coefficients of the class $\epsilon_{\mu_j-\mu_k}$ *might* adopt values different from zero, where μ_k are other elements of Q . Otherwise, any other ϵ_r must be zero and the upper bound of Eq. (37) cannot exceed p_S since terms like $\epsilon_{k_{\min}-\mu}$ do not belong to the type of coefficients $\epsilon_{\mu_j-\mu_k}$ that could have been greater than zero. For all of these studied cases, the fact that the optimal probability is p_S means that $\kappa_0 = \epsilon_r = 0$. This result that also satisfies the constraints of Eq. (35) when $k \neq k_{\min}$ and $k \notin Q$. This completes the proof. ■

We have proven that $\mathbf{x} = p_S \vec{\mathbf{e}}_0$ is the optimal solution of the optimization problem for this class of state separation, for both linearly independent and linearly dependent states, since any deviation $\sum_r \epsilon_r \vec{\mathbf{e}}_r - \kappa_0 \vec{\mathbf{e}}_0$ from the solution $p_S \vec{\mathbf{e}}_0$ yields a smaller probability. In principle, this solution is a local maximum. Nevertheless, the local extrema of linear functions are also global extrema. The solution found is, then, the global maximum of the probability of success.

It is noteworthy that the result from this optimization does not depend on the explicit parametrization of Eq. (11) possibly as a consequence of the preserved ordering of the coefficients. As an additional comment, these results also show that the use of a three-system strategy for unambiguous discrimination of symmetric states does not overcome the use of only two systems. The same conclusion holds for maximum-confidence discrimination among symmetric states.

IV. APPLICATION TO QUANTUM TELEPORTATION AND ENTANGLEMENT SWAPPING

In quantum teleportation, an unknown qubit state is deterministically and faithfully transmitted from a sender (Alice) to a receiver (Bob) with the consumption of a shared maximally entangled two-qubit state (the quantum channel) and two bits of forward classical communication [14]. Otherwise, when the quantum channel is not maximally entangled, faithful teleportation is achieved only probabilistically, whereas a deterministic implementation implies in less than unit average fidelity for the protocol. In either case one seeks for optimized schemes. On one hand perfect conclusive teleportation [31,32] achieves unit fidelity with maximum success probability. On the other hand, with the standard protocol [14] one achieves the optimal average teleportation fidelity deterministically [33,34].

Quantum teleportation of pure qudit states (D -level quantum systems) using nonmaximally entangled channels comprises the problem of discriminating among a set of equally likely symmetric states, whose linear dependence is based on the number of nonvanishing coefficients of the entangled state used as a channel. According to this, unambiguous discrimination or maximum-confidence discrimination can be applied to teleport, probabilistically, with the maximum achievable fidelity [15–17]. Otherwise, the standard procedure is equivalent to perform minimum-error discrimination and the teleportation is accomplished deterministically with an average fidelity lower than the probabilistic case [21,22]. In this regard, we can resort to state separation of symmetric states to interpolate between these two cases, i.e., increase the success probability compared to the former and the teleportation fidelity compared to the latter. Briefly, consider a two-qudit state given by

$$|\Psi\rangle_{12} = \sum_{m=0}^{N-1} a_m |m\rangle_1 |m\rangle_2, \quad (39)$$

where a_m are Schmidt coefficients, non-negative real numbers satisfying $\sum_{m=0}^{N-1} a_m^2 = 1$. The number of nonvanishing Schmidt coefficients is $D (\leq N)$. Let $|\phi\rangle_3$ be an N -dimensional state to be teleported. Particle 1 is in Bob's possession, whereas Alice owns particles 2 and 3. Quantum teleportation can be

described by, first, Alice applying a $\hat{G}_{23}^{\text{XOR}}$ operation on the pair of particles in her possession. The state of the three-particle system becomes

$$\hat{G}_{23}^{\text{XOR}} |\Psi\rangle_{12} |\phi\rangle_3 = \frac{1}{N} \sum_{l,k=0}^{N-1} \hat{Z}_1^{N-l} \hat{X}_1^k |\phi\rangle_1 |\alpha_l\rangle_2 |k\rangle_3, \quad (40)$$

where $\hat{G}_{23}^{\text{XOR}}$ is a generalization of a controlled-NOT gate, whose action on the computational basis is defined by $\hat{G}_{23}^{\text{XOR}} |i\rangle_2 |j\rangle_3 = |i\rangle_2 |i \oplus j\rangle_3$, where the addition \oplus and subtraction \ominus are modulo N operations [37,38]. The symmetric states $|\alpha_j\rangle$ are given by Eq. (7) with $\phi_k = 0$, whose fiducial coefficients are the Schmidt coefficients of the entangled state used as the channel. The operators \hat{Z} and \hat{X} are the generalized Pauli operators described in Sec. I. The standard procedure consists now in Alice performing von Neumann measurements on particles 2 and 3 in the inverse Fourier and computational bases, respectively. Alice, then, communicates her results to Bob in order for him to apply the necessary unitary operations to obtain the state. If, instead, Alice decides to separate the states $|\alpha_j\rangle$ before measuring, she will obtain, from Eqs. (12) and (40),

$$\begin{aligned} & \mathcal{U}_{2a}(\xi) \hat{G}_{23}^{\text{XOR}} |\Psi\rangle_{12} |\phi\rangle_3 |\mathcal{A}\rangle_a \\ &= \frac{\sqrt{p_S(\xi)}}{N} \sum_{l,k=0}^{N-1} \hat{Z}_1^{N-l} \hat{X}_1^k |\phi\rangle_1 |\beta_l(\xi)\rangle_2 |k\rangle_3 |0\rangle_a \\ &+ \frac{\sqrt{1-p_S(\xi)}}{N} \sum_{l,k=0}^{N-1} \hat{Z}_1^{N-l} \hat{X}_1^k |\phi\rangle_1 |\tilde{\beta}_l\rangle_2 |k\rangle_3 |1\rangle_a. \end{aligned} \quad (41)$$

Therefore, with success probability $p_S(\xi)$ [Eq. (14)], Alice can teleport the state as they had a bipartite entangled state with Schmidt coefficients $b_k(\xi)$ —which are present in states $|\beta_j(\xi)\rangle$ —instead of the original a_k . The probabilistic average fidelity is [33–36]

$$\mathcal{F}_{\text{ave}}(\xi) = \frac{1}{1+N} \left[1 + \left(\sum_{k=0}^{N-1} b_k(\xi) \right)^2 \right]. \quad (42)$$

We must note that $\xi = 0$ recovers the standard procedure and $\xi = 1$ stands for using unambiguous or maximum-confidence discrimination [16,17]. It can be seen that the number of Schmidt coefficients different from zero is meaningful for $\mathcal{F}_{\text{ave}}(\xi)$, since it can attain $(D+1)/(N+1)$ as its maximum value.

As an illustrative example, we can review this procedure applied to $N = D = 2$. A useful decomposition for $|\alpha_j\rangle$ ($j = 0, 1$) is $|\alpha_j\rangle = \cos(\alpha/2)|0\rangle + e^{i\pi j} \sin(\alpha/2)|1\rangle$, with $0 \leq \alpha \leq \pi/2$, so that $\alpha = \cos^{-1}\langle\alpha_0|\alpha_1\rangle$ is the angle between $|\alpha_0\rangle$ and $|\alpha_1\rangle$, as Fig. 4(a) shows. The standard teleportation procedure, also known as optimal deterministic teleportation, can teleport states with average fidelity $\mathcal{F}_{\text{ave}}^{(\alpha)} = [2 + \sin(\alpha)]/3$ [33,34]. In order to attain a fidelity equal to one, in a process known as perfect conclusive teleportation [31,32], Alice discriminates the states $|\alpha_j\rangle$ unambiguously with the optimal success probability [4–6],

$$P_\alpha = 1 - \langle\alpha_0|\alpha_1\rangle = 2 \sin^2(\alpha/2), \quad (43)$$

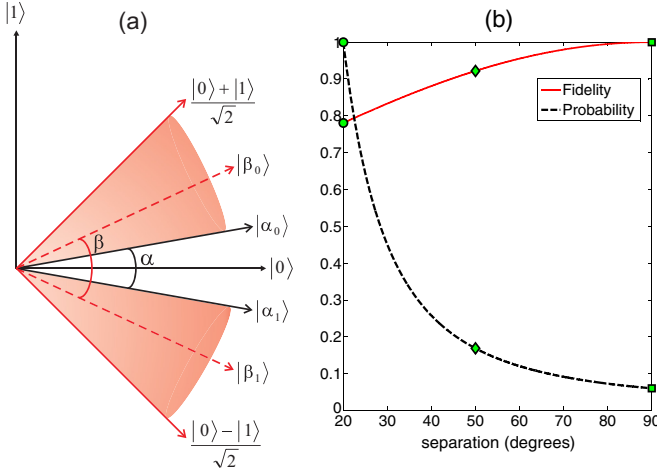


FIG. 4. (a) The initial separation α can be probabilistically increased to any angle β within the shaded region. (b) For $\alpha = 20^\circ$ the curves show the success probability and average teleportation fidelity as a function of the separation angle β . The symbols indicate the following protocols: (circles) optimal deterministic, (squares) perfect conclusive, and (diamonds) imperfect conclusive for $\beta = 50^\circ$.

which is then the probability of a successful teleportation. She tells Bob whether the discrimination succeeded or not by sending him one classical bit. When it succeeds, the protocol is accomplished as in the standard case and the teleported state has unit fidelity. Otherwise, in case of failure they interrupt the process.

Now, we will analyze a third approach that can be recognized as *imperfect conclusive teleportation*, in which an imperfect replica of state $|\phi\rangle$ is obtained after teleportation assisted by a conclusive attempt of state separation. Let $|\beta_0\rangle$ and $|\beta_1\rangle$ be two states separated by angle β , as illustrated in Fig. 4(a). In this process, the states $|\alpha_j\rangle$ are probabilistically mapped onto $|\beta_j\rangle$ through parametric state separation, with $\xi = 1 - \cos(\beta)/\cos(\alpha)$, and a success probability of [3]

$$P_\beta = \frac{1 - \langle \alpha_0 | \alpha_1 \rangle}{1 - \langle \beta_0 | \beta_1 \rangle} = \frac{P_\alpha}{2 \sin^2(\beta/2)}, \quad (44)$$

which, as expected, coincides with the optimal probability of Eq. (14) with $a_{\min} = \sin(\alpha/2)$ and $\xi = 1 - \cos(\beta)/\cos(\alpha)$. This is the probability that the teleportation will be carried out. It is easy to see that $P_\beta \geq P_\alpha$. Alice sends Bob one classical bit communicating whether the mapping $|\alpha_j\rangle \rightarrow |\beta_j\rangle$ succeeded or not. When it fails, they interrupt the protocol; otherwise, the protocol proceeds as in the standard case. The average teleportation fidelity is [33,34]

$$\mathcal{F}_{\text{ave}}^{(\beta)} = [2 + \sin(\beta)]/3, \quad (45)$$

which is the optimal one and satisfies $\mathcal{F}_{\text{ave}}^{(\beta)} \geq \mathcal{F}_{\text{ave}}^{(\alpha)}$.

Now, let us assume that for Alice and Bob purposes it is sufficient to accomplish teleportation with a fidelity of transmission better than a given threshold. In addition, they have scarce resources (e.g., entangled states) and share a quantum channel with a small degree of entanglement (which means small α). In this scenario, standard teleportation may not be enough to overcome that threshold deterministically while perfect conclusive teleportation would not be a clever

choice for them since unit fidelity is not required and the probability of success P_α is small. The best thing to do is for Alice to apply state separation on qubit 2 by setting a separation angle $\beta (< \pi/2)$ that, when successful, is just enough to enable an average teleportation fidelity above the threshold. Doing so, they will increase their transmission efficiency. This situation can be illustrated from Fig. 4(b). Assume a fidelity threshold $\mathcal{F}_{\text{thr}} = 0.91$. For the quantum channel considered there with $\alpha = 20^\circ$, the deterministic protocol delivers a fidelity $\mathcal{F}_\alpha \approx 0.78$ (circle), while the perfect conclusive one succeeds with probability $P_\alpha \approx 0.06$ (square). In this case, an imperfect conclusive protocol that gives an average teleportation fidelity $\mathcal{F}_\beta \approx 0.92$, will succeed with probability $P_\beta \approx 0.17$ (diamonds).

An analogous analysis can be scrutinized for entanglement swapping. Unlike the previous case, the description of this protocol is more complicated and, for the sake of simplicity and ease of comprehension, only basic details are given. In entanglement swapping, two users have, each one, a two-qubit entangled state, say $|\kappa\rangle_{12}$ for Alice and $|\chi\rangle_{34}$ for Bob. Each party sends one particle, say 1 and 4, to a third party who will perform a Bell state measurement on them, preparing an entangled state shared along two parties that never interacted, i.e., an entangled state $|\tau\rangle_{23}$ [14,39]. When at least one of the initial bipartite states ($|\kappa\rangle_{12}$ or $|\chi\rangle_{34}$) is not maximally entangled, the resulting state $|\tau\rangle_{23}$ will not be maximally entangled either. A deterministic implementation is related with an optimal minimum-error strategy, while unambiguous (or maximum-confidence) discrimination enables one to accomplish the protocol with the maximum achievable entanglement for the quantum channels involved [40,41]. With the state separation presented here one can interpolate between these two extreme cases, trading off success probability and the amount of entanglement of the states resulting from swapping.

V. PROPOSAL FOR OPTICAL IMPLEMENTATION

Figure 5 illustrates a simplified optical setup able to perform optimal parametric state separation. This particular case considers $\phi_k = \varphi_k = 0, \forall k$ [see Eqs. (7) and (10)]. Half-wave plates (HWP) are used to assist the state preparation and to implement conditional operations depending on which path the photon followed. An HWP whose fast axis is oriented at an angle ζ with respect to the horizontal axis is represented by a Jones matrix given by [42]

$$\hat{H}(\zeta) = \begin{pmatrix} \cos(2\zeta) & \sin(2\zeta) \\ \sin(2\zeta) & -\cos(2\zeta) \end{pmatrix}, \quad (46)$$

written in the $\{|h\rangle, |v\rangle\}$ basis. Stage I is aimed to prepare the symmetric states encoded in the longitudinal spatial modes of a single photon, whereas its polarization is the degree of freedom used as ancilla. Phase shifters are used to encode the phases corresponding to each symmetric state. The state of the photon after this stage is $|\alpha_j\rangle|v\rangle$, considering that polarizing beam splitters (PBSs) reflect incident vertically polarized photons and transmit horizontally polarized photons. In Stage II, the unitary operation of Eq. (12) is implemented by the HWP at

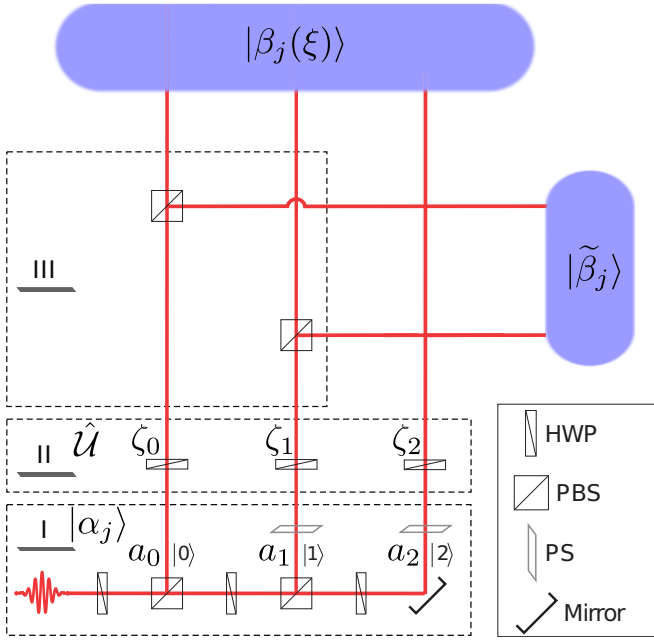


FIG. 5. Experimental proposal for implementation of quantum state separation of N states from a three-dimensional Hilbert space. HWP, half-wave plate; PBS, polarizing beam splitter; PS, phase shifter (see text for details).

each path. This operation can be written as

$$\begin{aligned} \hat{U} &= \sum_{k=0}^{D-1} |k\rangle\langle k| \otimes \hat{H}(\zeta_k) \\ &= \left(\sum_{k=0}^{D-1} \cos(2\zeta_k) |k\rangle\langle k| \right) \otimes (|h\rangle\langle h| - |v\rangle\langle v|) \\ &\quad + \left(\sum_{k=0}^{D-1} \sin(2\zeta_k) |k\rangle\langle k| \right) \otimes (|h\rangle\langle v| + |v\rangle\langle h|). \end{aligned} \quad (47)$$

By setting the angles of each HWP such that

$$\sin(2\zeta_k) = \sqrt{\frac{1 - \xi + \xi/Da_k^2}{1 - \xi + \xi/Da_{\min}^2}}, \quad (48)$$

$$\cos(2\zeta_k) = -\sqrt{\frac{\xi}{D} \frac{1/a_{\min}^2 - 1/a_k^2}{1 - \xi + \xi/Da_{\min}^2}}, \quad (49)$$

we obtain

$$\hat{U} = \hat{A}_S \otimes (|h\rangle\langle v| + |v\rangle\langle h|) + \hat{A}_F \otimes (|v\rangle\langle v| - |h\rangle\langle h|), \quad (50)$$

which reduces to Eq. (12), considering $|\mathcal{A}\rangle_a = |v\rangle$, $|0\rangle_a = |h\rangle$, and $|1\rangle_a = |v\rangle$, when applied on $|\alpha_j\rangle|v\rangle$. In this case, failure states $|\beta_j\rangle$ [Eq. (19)] do not depend on ξ since null phases for the fiducial states were chosen. Finally, Stage III performs a measurement on photon polarization which, in the case of a successful event, yields a photon prepared in state $|\beta_j(\xi)\rangle$, shown above Stage III. Otherwise, the state will be $|\tilde{\beta}_j\rangle$ as sketched at the right side of Stage III.

VI. CONCLUSIONS

Besides minimum error, unambiguous discrimination, and maximum confidence as strategies for discrimination of nonorthogonal states, there exist other strategies that interpolate between minimum-error and unambiguous discrimination for linearly independent states, and between minimum error and maximum confidence for linearly dependent ones. These strategies usually impose either a fixed value or an upper bound for the probability of mistaking the retrodiction of the state [43–47] or for the probability of a failed attempt [13,48–50] while optimizing the probability of conclusive and correct results, respectively. Although these strategies comprise different approaches for quantum state discrimination and different optimization problems, it has been shown that they are closely related and the optimization of the probability of correct results under a fixed rate of inconclusive results can be found by optimizing the same probability of correct results under a certain fixed error probability and vice versa [50]. Moreover, it has been shown [13] that a given set of measurements aimed to discriminate quantum states with a particular failure probability attains the optimal probability of correct results if and only if the whole discrimination process can be described by a map from the original set of states onto a specific set of new states and then performing optimal minimum-error probability on the latter (Theorem 1 of Ref. [13]). This finding allows a new approach to the aforementioned optimization problems by treating them all as maps between sets of quantum states followed by minimum-error discrimination. Concerning the topic studied throughout this paper, we can use the probability of success [Eq. (14)] to set a fixed rate of inconclusive results and find the value of ξ , which can be labeled as ξ_{fixed} , that fits with this constraint. Afterwards, the use of Eqs. (16) and (17) allow one to find the Kraus operators needed to map from $|\alpha_j\rangle$ to $|\beta_j(\xi_{\text{fixed}})\rangle$. Since these new states are also equally likely symmetric states, the solution for optimal minimum-error discrimination among them is known in the literature [21,22]. Thus, discrimination strategies interpolating between minimum-error and unambiguous (or maximum-confidence) discrimination can benefit from the map proposed in this work. It is worth mentioning that the states resulting from a failed attempt, which are given by Eq. (20), are also symmetric pure states. This opens the possibility of studying sequential maps analogous to the sequential discrimination proposed in Ref. [19] in which it is still possible to use the inconclusive results to attempt discrimination instead of just discarding these results.

Summarizing, we have proposed a map between sets of uniparametric symmetric pure quantum states that leads to more distinguishable states. The results reported here consider both linearly independent and dependent states. The optimal probability of success was obtained and applications to teleportation of quantum states and entanglement swapping were shown.

As with teleportation, entanglement swapping with non-maximally entangled states can be assisted by state separation of symmetric states. Thus several applications of entanglement swapping can also be improved with the protocol presented here, as, for instance, quantum repeaters [51], entanglement generation between distant users [52–54], some

entanglement concentration schemes [55–58], experimental studies on nonlocality [39], generation of Greenberger-Horne-Zeilinger (GHZ) states via multiparticle entanglement swapping [59,60], quantum secret sharing [61], and other quantum communication protocols [62–66].

ACKNOWLEDGMENTS

This work was supported by CONICyT Grant No. PFB-0824, Millenium Scientific Initiative Grant No. RC130001, and FONDECYT Grants No. 11085057 and No. 1140635. M.A.S.P. acknowledges financial support from CONICyT. L.N. acknowledges financial support from Brazilian agencies CNPq (Grant No. 485401/2013-4), FAPEMIG (Grant No. APQ-00149-13), and PRPq/UFGM (Grant No. ADRC-01/2013). O.J. acknowledges financial support from FONDECYT Grant No. 11121318.

APPENDIX: OPTIMIZATION AND KRAUS OPERATORS

The linear programming problem of Eq. (30) offers a solution to the task of mapping one set of symmetric states onto another. Although the solution is explicit in terms of probability and the leak of the transform, no information about the Kraus operators is provided. Equation (29) shows the Stinespring representation of the linear map, although determining a Kraus operator decomposition is desirable as well. So, let us consider a set of $N + 1$ Kraus operators $\{\hat{t}_{\mathcal{F}}, \hat{t}_{kS}\}_{k=0}^{N-1}$ such that [67]

$$\sum_{k=0}^{N-1} (\hat{t}_{kS}^\dagger \hat{t}_{kS}) + \hat{t}_{\mathcal{F}}^\dagger \hat{t}_{\mathcal{F}} = \hat{I}, \quad (\text{A1})$$

where \hat{t}_{kS} and $\hat{t}_{\mathcal{F}}$ stand for the successful and unsuccessful cases, respectively. The matrix elements of each operator \hat{t}_{kS} can be calculated as [68]

$$\langle y | \hat{t}_{kS} | z \rangle = \langle y \otimes k \otimes 0 | \hat{U} | z \otimes \mathcal{M} \otimes \mathcal{A} \rangle. \quad (\text{A2})$$

From Eq. (29) we have

$$\begin{aligned} & \hat{U} \sum_{k'=0}^{N-1} a_{k'} \omega^{j k'} |k'\rangle_s | \mathcal{M} \rangle_e | \mathcal{A} \rangle_a \\ &= \sqrt{p} \sum_{l,m=0}^{N-1} b_l \Psi_m \omega^{j(l+m)} |l\rangle_s |m\rangle_e |0\rangle_a \\ &+ \sqrt{1-p} |\Phi_j\rangle_{se} |1\rangle_a. \end{aligned}$$

Applying $\langle y \otimes k \otimes 0 |$ from the left, we obtain

$$\begin{aligned} & \sum_{k'=0}^{N-1} a_{k'} \omega^{k' j} \langle y \otimes k \otimes 0 | \hat{U} | k' \otimes \mathcal{M} \otimes \mathcal{A} \rangle \\ &= \sqrt{p} \sum_{l,m=0}^{N-1} b_l \Psi_m \omega^{j(l+m)} \langle y \otimes k \otimes 0 | l \otimes m \otimes 0 \rangle \\ &+ \sqrt{1-p} \langle y \otimes k \otimes 0 | \Phi_j \rangle_{se} |1\rangle_a, \end{aligned}$$

or, equivalently,

$$\sum_{k'=0}^{N-1} a_{k'} \omega^{k' j} \langle y | \hat{t}_{kS} | k' \rangle = \sqrt{p} b_y \Psi_k \omega^{j(y+k)}.$$

The $\langle y | \hat{t}_{kS} | z \rangle$ components can be isolated by applying inverse Fourier transform ($\sum_{j=0}^{N-1} \omega^{-jz} / \sqrt{N}$). This leads us to obtain

$$\langle y | \hat{t}_{kS} | z \rangle = \sqrt{p} \frac{b_y \Psi_k}{a_z} \delta_{y+k,z}.$$

Thus,

$$\hat{t}_{kS} = \sum_{y,z=0}^{N-1} |y\rangle \langle z| \langle y | \hat{t}_{kS} | z \rangle = \sqrt{p} \sum_{y=0}^{N-1} \frac{b_y \Psi_k}{a_{y+k}} |y\rangle \langle y+k|.$$

The Kraus operators $\hat{t}_{\mathcal{F}}$ can be obtained, although not uniquely, by

$$\hat{t}_{\mathcal{F}} = \left[\hat{I} - \sum_{y=0}^{N-1} \hat{t}_{kS}^\dagger \hat{t}_{kS} \right]^{1/2}, \quad (\text{A3})$$

and, recalling that $\mathbf{x} = p \sum_{k=0}^{N-1} \Psi_k^2 |k\rangle$, we then have

$$\hat{t}_{kS} = \sqrt{x_k} \sum_{y=0}^{N-1} \frac{b_y}{a_{y+k}} |y\rangle \langle y+k|. \quad (\text{A4})$$

In consequence, the results obtained from the linear optimization program reveal not only the optimal probability but also the number of Kraus operators required for the successful case, which is the number of nonvanishing components of \mathbf{x} . So, the $|\psi_y\rangle$ vectors contain information about the physical process in terms of Kraus operators. As a safety check, it can be readily shown that

$$\sum_{k=0}^{N-1} \hat{t}_{kS} |\alpha_j\rangle \langle \alpha_j | \hat{t}_{kS}^\dagger = p |\beta_j\rangle \langle \beta_j|.$$

We have considered only one Kraus operator $\hat{t}_{\mathcal{F}}$ for the inconclusive case. Although there can be several failure Kraus operators $\hat{t}_{r\mathcal{F}}$, we have considered that $\hat{t}_{\mathcal{F}}^\dagger \hat{t}_{\mathcal{F}} = \sum_r \hat{t}_{r\mathcal{F}}^\dagger \hat{t}_{r\mathcal{F}}$ since the main interest remains focused on the Kraus operators associated with successful attempts.

[1] A. Chefles, *Phys. Lett. A* **270**, 14 (2000).

[2] A. Chefles, R. Jozsa, and A. Winter, *Int. J. Quantum. Inf.* **2**, 11 (2004).

[3] A. Chefles and S. M. Barnett, *J. Phys. A* **31**, 10097 (1998).

[4] I. D. Ivanovic, *Phys. Lett. A* **123**, 257 (1987).

[5] D. Dieks, *Phys. Lett. A* **126**, 303 (1988).

[6] A. Peres, *Phys. Lett. A* **128**, 19 (1988).

[7] G. Jaeger and A. Shimony, *Phys. Lett. A* **197**, 83 (1995).

- [8] A. Peres and D. R. Terno, *J. Phys. A* **31**, 7105 (1998).
- [9] A. Chefles, *Phys. Lett. A* **239**, 339 (1998).
- [10] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, *Phys. Rev. Lett.* **96**, 070401 (2006).
- [11] A. S. Holevo, *J. Multivariate Anal.* **3**, 337 (1973).
- [12] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [13] K. Nakahira, T. S. Usuda, and K. Kato, *Phys. Rev. A* **86**, 032316 (2012).
- [14] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [15] L. Roa, A. Delgado, and I. Fuentes-Guridi, *Phys. Rev. A* **68**, 022310 (2003).
- [16] L. Neves, M. A. Solís-Prosser, A. Delgado, and O. Jiménez, *Phys. Rev. A* **85**, 062322 (2012).
- [17] M. A. Solís-Prosser, O. Jiménez, L. Neves, and A. Delgado, *Phys. Scr.*, **T 153**, 014058 (2013).
- [18] A. Chefles and S. M. Barnett, *Phys. Lett. A* **250**, 223 (1998).
- [19] O. Jiménez, M. A. Solís-Prosser, A. Delgado, and L. Neves, *Phys. Rev. A* **84**, 062315 (2011).
- [20] U. Herzog, *Phys. Rev. A* **85**, 032312 (2012).
- [21] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, *Int. J. Theor. Phys.* **36**, 1269 (1997).
- [22] S. M. Barnett, *Phys. Rev. A* **64**, 030303 (2001).
- [23] A. Chefles, *Phys. Rev. A* **65**, 052314 (2002).
- [24] Y. Feng, R. Duan, and Z. Ji, *Phys. Rev. A* **72**, 012313 (2005).
- [25] X.-F. Zhou, Q. Lin, Y.-S. Zhang, and G.-C. Guo, *Phys. Rev. A* **75**, 012321 (2007).
- [26] L. Roa, M. L. Ladrón de Guevara, and A. Delgado, *Phys. Rev. A* **81**, 034101 (2010).
- [27] F. A. Torres-Ruiz, J. Aguirre, A. Delgado, G. Lima, L. Neves, S. Pádua, L. Roa, and C. Saavedra, *Phys. Rev. A* **79**, 052113 (2009).
- [28] V. Dunjko and E. Andersson, *J. Phys. A* **45**, 365304 (2012).
- [29] B. He and J. A. Bergou, *Phys. Lett. A* **356**, 306 (2006).
- [30] A. Chefles, *Phys. Rev. A* **66**, 042325 (2002).
- [31] T. Mor, [arXiv:quant-ph/9608005](https://arxiv.org/abs/quant-ph/9608005).
- [32] Y.-J. Gu, Y.-Z. Zheng, and G.-C. Guo, *Phys. Lett. A* **296**, 157 (2002).
- [33] K. Banaszek, *Phys. Rev. A* **62**, 024301 (2000).
- [34] G. Vidal, D. Jonathan, and M. A. Nielsen, *Phys. Rev. A* **62**, 012304 (2000).
- [35] K. Banaszek, *Phys. Rev. Lett.* **86**, 1366 (2001).
- [36] M. Horodecki, P. Horodecki, and R. Horodecki, *Phys. Rev. A* **60**, 1888 (1999).
- [37] G. Alber, A. Delgado, N. Gisin, and I. Jex, *J. Phys. A* **34**, 8821 (2001).
- [38] I. Jex, G. Alber, S. M. Barnett, and A. Delgado, *Fortschr. Phys.* **51**, 172 (2003).
- [39] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, *Phys. Rev. Lett.* **71**, 4287 (1993).
- [40] A. Delgado, L. Roa, J. C. Retamal, and C. Saavedra, *Phys. Rev. A* **71**, 012303 (2005).
- [41] M. A. Solís-Prosser, A. Delgado, O. Jiménez, and L. Neves, *Phys. Rev. A* **89**, 012337 (2014).
- [42] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, New York, 1991).
- [43] M. A. P. Touzel, R. B. A. Adamson, and A. M. Steinberg, *Phys. Rev. A* **76**, 062314 (2007).
- [44] A. Hayashi, T. Hashimoto, and M. Horibe, *Phys. Rev. A* **78**, 012333 (2008).
- [45] H. Sugimoto, T. Hashimoto, M. Horibe, and A. Hayashi, *Phys. Rev. A* **80**, 052322 (2009).
- [46] H. Sugimoto, Y. Taninaka, and A. Hayashi, *Phys. Rev. A* **86**, 042311 (2012).
- [47] G. Sentís, E. Bagan, J. Calsamiglia, and R. Muñoz-Tapia, *Phys. Rev. A* **88**, 052304 (2013).
- [48] A. Chefles and S. M. Barnett, *J. Mod. Opt.* **45**, 1295 (1998).
- [49] E. Bagan, R. Muñoz-Tapia, G. A. Olivares-Rentería, and J. A. Bergou, *Phys. Rev. A* **86**, 040303(R) (2012).
- [50] U. Herzog, *Phys. Rev. A* **86**, 032314 (2012).
- [51] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, *Rev. Mod. Phys.* **83**, 33 (2011).
- [52] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, *Phys. Rev. Lett.* **81**, 5932 (1998).
- [53] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, *Phys. Rev. A* **59**, 169 (1999).
- [54] E. Waks, A. Zeevi, and Y. Yamamoto, *Phys. Rev. A* **65**, 052310 (2002).
- [55] B.-S. Shi, Y.-K. Jiang, and G.-C. Guo, *Phys. Rev. A* **62**, 054301 (2000).
- [56] L.-Y. Hsu, *Phys. Lett. A* **297**, 126 (2002).
- [57] J. Modlawska and A. Grudka, *Phys. Rev. A* **78**, 032321 (2008).
- [58] M. Yang, A. Delgado, L. Roa, and C. Saavedra, *Opt. Commun.* **282**, 1482 (2009).
- [59] S. Bose, V. Vedral, and P. L. Knight, *Phys. Rev. A* **57**, 822 (1998).
- [60] L. Hardy and D. D. Song, *Phys. Rev. A* **62**, 052315 (2000).
- [61] P. Zhou, F.-G. Deng, and H.-Y. Zhou, *Phys. Scr.* **79**, 035005 (2009).
- [62] Y. Xia, J. Song, and H.-S. Song, *Phys. Scr.* **76**, 363 (2007).
- [63] Y.-B. Zhan, L.-L. Zhang, and Q.-Y. Zhang, *Opt. Commun.* **282**, 4633 (2009).
- [64] S.-J. Qin, F. Gao, Q.-Y. Wen, and F.-C. Zhu, *Opt. Commun.* **283**, 1566 (2010).
- [65] A. Scherer, B. C. Sanders, and W. Tittel, *Opt. Express* **19**, 3004 (2011).
- [66] N. Zhou, L. Wang, L. Gong, X. Zuo, and Y. Liu, *Opt. Commun.* **284**, 4836 (2011).
- [67] K. Kraus, in *Foundations of Quantum Mechanics and Ordered Linear Spaces*, Lecture Notes in Physics Vol 29 (Springer, Berlin/Heidelberg/New York, 1974), pp. 206–229.
- [68] M. Keyl, *Phys. Rep.* **369**, 431 (2002).