# Quantum operations, state transformations and probabilities

Anthony Chefles
*Department of Physical Sciences, University of Hertfordshire, Hatfield AL10 9AB, Hertfordshire, United Kingdom*
(Received 13 September 2001; published 3 May 2002)

In quantum operations, probabilities characterize both the degree of the success of a state transformation and, as density operator eigenvalues, the degree of mixedness of the final state. We give a unified treatment of pure→pure state transformations, covering both probabilistic and deterministic cases. We then discuss the role of majorization in describing the dynamics of mixing in quantum operations. The conditions for mixing enhancement for all initial states are derived. We show that mixing is monotonically decreasing for deterministic pure→pure transformations, and discuss the relationship between these transformations and deterministic local operations with classical communication entanglement transformations.

DOI: 10.1103/PhysRevA.65.052314 PACS number(s): 03.67.Hk, 03.65.Ta

## I. INTRODUCTION

Information is carried by physical systems and encoded in their states. It follows that the ways in which information can be manipulated are determined by the scope of the set of possible operations on the states of the signal carriers. It is for this reason that the recent widespread fascination with the information-theoretic properties of quantum systems [1] has been accompanied by a renaissance in the study of the quantum operations formalism, which determines what we can and cannot do with the state of a quantum system.

In quantum, as in classical-information theory, the systems considered may be in one of many possible states. However, quantum states can have attributes that have no exact classical analogue, such as nonorthogonality and entanglement. These features of quantum states, together with the numerous ways in which quantum states can be manipulated, have given rise to some intriguing discoveries in quantum-information theory, such as teleportation, classical capacity superadditivity, and quantum error correction. Certain limitations on the way in which quantum states can be manipulated, such as the no-cloning theorem, also carry significant benefits, such as the security of quantum key distribution and consistency with special relativity.

The many successes in determining optimal transformations for carrying out specific important tasks, such as state discrimination/estimation, approximate cloning, and entanglement manipulation, have led to some more general questions being asked about the constraints imposed by the quantum formalism on state manipulation. In this respect, Hardy and Song [2] have considered optimal universal manipulation of a qubit, while Alber, Delgado, and Jex [3] have described universal bipartite entanglement processes. Even more recently, Fiurášek [4] has discussed the properties of quantum operations which optimally approximate a given transformation of one set of pure states into another with unit probability. The conditions under which such a transformation can be carried out exactly, at least when the initial states are linearly independent, have been derived in [5].

In this paper we continue to explore the properties of general quantum operations and how they transform quantum states. Section II is devoted to giving a unified treatment of probabilistic and deterministic transformations between sets of pure states. We consider the following scenario: a quantum system is prepared in one of the $N$ pure states $|\psi_j^1\rangle$, where $j=1,\ldots,N$. We wish to implement the transformation $|\psi_j^1\rangle\rightarrow|\psi_j^2\rangle$, for some other set of $N$ pure states $|\psi_j^2\rangle$. In general, the transformation will not be deterministic, and will only succeed with some probability for each state. We obtain necessary and, for linearly independent initial states, sufficient conditions for the existence of a quantum operation which carries out this transformation for a fixed set of success probabilities. We then examine some consequences of these conditions, and show how they lead to simple derivations of established conditions for deterministic state transformations and optimal unambiguous state discrimination.

For a general quantum operation, when the initial state is pure, the final state will often be mixed. This effect is common and occurs under many circumstances where we wish to preserve the information content of a quantum state, such as in quantum communications and quantum computation. To understand this mixing it helps to have an appreciation of its quantitative features. A suitable framework for the discussion of mixing is provided by the concept of majorization. This concept was introduced to quantum mechanics by Uhlmann [6–8] as a means of comparing mixing in density operators. Numerous useful theorems relating to majorization have been discovered [9].

The subject of majorization has recently received renewed attention in quantum-information theory, mainly as a result Nielsen's discovery that it provides a suitable framework for the discussion of pure, bipartite entanglement transformations [10]. More recently, Nielsen [11] has derived several interesting majorization relations for static and dynamic mixing of quantum states, latterly in association with generalized measurements (see also the related analysis by Fuchs and Jacobs [12].) Nielsen has also showed that a density operator can represent some ensemble of pure states with fixed probabilities if and only if a certain majorization relation is satisfied [13].

In Sec. III, we describe and employ the concept of majorization as a tool to help us understand the dynamics of mixing in quantum operations. Many nonunitary quantum operations transform at least some pure states into mixed states. This begs the question: under what conditions does a quantum operation never decrease the extent to which any initial

 

state is mixed? Majorization is a suitable tool for comparing the degree of disorder in the initial and final states, and a sufficient condition for this monotonic mixing was derived, in purely algebraic context, by Bapat and Sunder [14]. We give a simple derivation of their condition within the context of quantum operations, and show that this condition is also in fact a necessary condition. We then examine majorization in relation to deterministic pure state transformations, and derive an intuitive and information-theoretically satisfying majorization relation for such operations.

## II. TRANSFORMATIONS BETWEEN SETS OF PURE STATES

### A. Transformation conditions for fixed probabilities

Consider the following situation: we have in our possession a quantum system with a finite, $D$-dimensional Hilbert space $\mathcal{H}$. The initial state of the system is pure, and is an element of the set $\{|\psi_j^1\rangle\}$, where $j=1,\ldots,N$ for some finite $N$. Our aim is to implement a probabilistic transformation $\mathcal{P}$ which transforms the state $|\psi_j^1\rangle$ into some other pure state $|\psi_j^2\rangle$ for each $j$.

It is well known, from studies of particular transformations such as unambiguous state discrimination [16] and probabilistic cloning [17] , that we cannot in general expect the probability of success to be equal to one. Let $p_j$ be the probability of successful transformation of $|\psi_j^1\rangle$ into $|\psi_j^2\rangle$. These probabilities may be represented as the components of a vector $\mathbf{p}=\{p_j\}$.

Generally speaking, the transformation $\mathcal{P}$ will be represented by a completely positive, linear map. We would like to be able to determine unambiguously whether or not the desired transformation has succeeded. This requirement implies that the procedure will have two possible outcomes: success or failure. It will be described by the transformation operators $\{A_{kr}\}$, where $r=S,F$, corresponding to success and failure, respectively, and $k=1,\ldots,M$, for some $M$. If the system is prepared in a state represented by an initial density operator $\rho$, then the probability of the $r$th outcome is determined by the positive quantum detection operator, or positive operator-valued measure (POVM) element

$$E_r=\sum_k A_{kr}^\dagger A_{kr}. \tag{2.1}$$

Throughout this paper, when we speak of a positive operator or matrix, we will, unless otherwise indicated, mean one which is positive semidefinite. The probability of the $r$th outcome is given by

$$p_r(\rho)=\mathrm{Tr}(\rho E_r), \tag{2.2}$$

where

$$\sum_r E_r=1. \tag{2.3}$$

The postmeasurement state corresponding to the $r$th outcome is

$$\rho_r=\sum_k A_{kr}\rho A_{kr}^\dagger/p_r(\rho). \tag{2.4}$$

It is clear from Eq. (2.1) that $E_r$ is positive. From the resolution of the identity in Eq. (2.3) we see that

$$0\leqslant E_r\leqslant 1. \tag{2.5}$$

Let us denote by $\Sigma_{\mathbf{p}}(\mathcal{P})$ the set of admissible probability vectors for the transformation $\mathcal{P}$. We would like to determine the conditions under which a particular probability vector is an element of $\Sigma_{\mathbf{p}}(\mathcal{P})$. The necessary and sufficient conditions for the existence of a transformation which succeeds with probability vector $\mathbf{p}\in\Sigma_{\mathbf{p}}(\mathcal{P})$ are that it can be realized by a set of linear transformation operators as in Eq. (2.4) and that the corresponding POVM element satisfies Eq. (2.1). When this is the case, the operation can always be realized by introducing an ancillary system, with which our original system interacts unitarily. Following this, a projective measurement is carried out on the ancilla, and this has two outcomes. The two transformations of the state of the original system conditioned on these outcomes are the transformation $\mathcal{P}$ and the failure transformation [15].

As a consequence of the necessity and sufficiency of these conditions, it is worth noting that, if we are not particularly concerned about the form of the final states when the attempt to implement $\mathcal{P}$ fails, we may, without loss of generality, assume that there is only one nonzero failure operator, which can be taken to be $A_F=\sqrt{1-E_S}$.

These criteria, while correct, may not always be the most helpful, due to the large number of parameters describing the transformation operators. The following theorem gives simpler necessary, and, for linearly independent initial states, sufficient conditions for the existence of such a transformation.

*Theorem 1*. Let $\{|\psi_j^1\rangle\}$ be a set of $N$ pure quantum states spanning a $D$-dimensional Hilbert space $\mathcal{H}$. Let $\{|\psi_j^2\rangle\}$ be another set of $N$ pure states lying in $\mathcal{H}$. Let the Gram matrices of the initial and final sets be denoted by $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$, respectively. If there exists a probabilistic transformation $\mathcal{P}$: $\{|\psi_j^1\rangle\}\rightarrow\{|\psi_j^2\rangle\}$ with probability vector $\mathbf{p}$, then there exists an $N\times N$ matrix $\mathbf{\Pi}$ which satisfies the following conditions:

(i) $\mathbf{\Pi}\geqslant 0$,
(ii) $\mathrm{diag}(\mathbf{\Pi})=\mathbf{p}$,
(iii) $\mathbf{\Gamma}_1-\mathbf{\Pi}\circ\mathbf{\Gamma}_2\geqslant 0$,

where "$\circ$" denotes the Hadamard (or Schur) matrix product. These conditions are also sufficient if the set $\{|\psi_j^1\rangle\}$ is linearly independent.

Prior to giving a proof of this theorem, we recall that the $N\times N$ Gram matrix $\mathbf{\Gamma}=\{\gamma_{j'j}\}$ corresponding to a set of $N$ pure states $|\psi_j\rangle$ has elements $\gamma_{j'j}=\langle\psi_{j'}|\psi_j\rangle$. Also, the Hadamard product $\mathbf{A}\circ\mathbf{B}$ of two matrices $\mathbf{A}=\{a_{j'j}\}$ and $\mathbf{B}=\{b_{j'j}\}$ has $j'j$ element $a_{j'j}b_{j'j}$.

*Proof*. We begin by proving the necessity part of this theorem. To do this, we note that there must exist complex coefficients $c_{kj}$ such that

$$A_{kS}|\psi_j^1\rangle=c_{kj}|\psi_j^2\rangle. \tag{2.6}$$

We can consider these coefficients to be the elements of an $M\times N$ matrix $\mathbf{C}=\{c_{kj}\}$. Let us now introduce the $N\times N$ matrix $\mathbf{\Pi}=\{\pi_{j'j}\}$ defined by

$$\mathbf{\Pi}=\mathbf{C}^\dagger\mathbf{C}. \tag{2.7}$$

This matrix is clearly positive and thus satisfies condition (i). To see that it also satisfies condition (ii), we make use of the fact that $p_j=\langle\psi_j^1|\Sigma_k A_{kS}^\dagger A_{kS}|\psi_j^1\rangle=\Sigma_k|c_{kj}|^2$. This is easily shown to be equal to $\pi_{jj}$ using Eq. (2.7), which implies that $\mathbf{\Pi}$ satisfies condition (ii). Finally, condition (iii) can be verified by imposing Eq. (2.5), which requires the expectation value of $\Sigma_k A_{kS}^\dagger A_{kS}$ to be no greater than one for any state. Consider an arbitrary pure state $|\phi\rangle$ in the subspace spanned by the $\{|\psi_j^1\rangle\}$. We may write it as $|\phi\rangle=\Sigma_j v_j|\psi_j^1\rangle$ and calculate

$$\langle\phi|\left[\sum_k A_{kS}^\dagger A_{kS}\right]|\phi\rangle=\sum_{jj'} v_{j'}^* v_j \pi_{j'j}\gamma_{j'j}^2,$$

$$\leqslant\langle\phi|\phi\rangle=\sum_{jj'} v_{j'}^* v_j \gamma_{j'j}^1. \tag{2.8}$$

The requirement that $\Sigma_k A_{kS}^\dagger A_{kS}\leqslant 1$ is then seen to be equivalent to the inequality

$$\sum_{jj'} v_{j'}^* v_j(\pi_{j'j}\gamma_{j'j}^2-\gamma_{j'j}^1)\leqslant 0, \tag{2.9}$$

which holds for every vector $\mathbf{v}=\{v_j\}$. From this it follows that the $N\times N$ matrix with elements $\{\gamma_{j'j}^1-\pi_{j'j}\gamma_{j'j}^2\}$ is positive, which is exactly what is expressed, more concisely, by condition (iii).

To prove the converse for linearly independent initial states, we assume the existence of a matrix $\mathbf{\Pi}=\{\pi_{j'j}\}$ which satisfies conditions (i)–(iii). Positivity enables us to factorize $\mathbf{\Pi}$ as $\mathbf{C}^\dagger\mathbf{C}$, for some $M\times N$ matrix $\mathbf{C}=\{c_{kj}\}$, where the integer $M$ may take any value not less than $N$. Let us now define the transformation operators

$$A_{kS}=\sum_j \frac{c_{kj}}{\langle\tilde{\psi}_j^1|\psi_j^1\rangle}|\psi_j^2\rangle\langle\tilde{\psi}_j^1|. \tag{2.10}$$

The $|\tilde{\psi}_j^1\rangle$ are the reciprocal vectors corresponding to the states $|\psi_j^1\rangle$. These have been found, in studies of operations of unambiguous state discrimination [16] and deterministic transformations [5], to be extremely useful in dealing with transformations of sets of linearly independent states. The state $|\tilde{\psi}_j^1\rangle$ is defined as that in $\mathcal{H}$ which is orthogonal to all $|\psi_{j'}^1\rangle$ for $j\neq j'$ and is, up to a phase, unique.

From definition in Eq. (2.10), we see that (i) is automatically satisfied. Also making use of Eq. (2.2), is clear that $p_j$, the transformation probability for the $j$th state, given by $\langle\psi_j|\Sigma_k A_{kS}^\dagger A_{kS}|\psi_j\rangle$, is equal to $\pi_{jj}$. This shows that condition (ii) is satisfied. Finally, the necessary and sufficient condition for the transformation operators in Eq. (2.10) to be physically realizable is that $\Sigma_k A_{kS}^\dagger A_{kS}\leqslant 1$. If condition (iii) is satisfied, then so is inequality (2.5), which is equivalent to $\Sigma_k A_{kS}^\dagger A_{kS}\leqslant 1$. This completes the proof. ■

We have assumed, for the sake of notational convenience, that the final states are in the space spanned by the initial states. This leads to no loss of generality. The linearity of quantum mechanics implies that the dimension of the space spanned by the final states cannot exceed that of the input states (this is why linearly dependent states cannot be unambiguously discriminated; to transform a linearly dependent set into an orthogonal set would necessarily increase the dimension of the space they span.) So, the final states can always be transformed unitarily, and thus reversibly, into a subspace of $\mathcal{H}$. We can then assume that this is done, satisfying the conditions of Theorem 1.

### B. Examples

It is instructive to see how established results relating to specific transformations follow from the general conditions in Theorem 1. The first kind of transformation we shall consider is a deterministic transformation, where all of the $p_j$ are equal to 1. Let us write $\mathbf{G}=\mathbf{\Gamma}_1-\mathbf{\Pi}\circ\mathbf{\Gamma}_2$. As a consequence of (iii), $\mathbf{G}$ must be positive. The diagonal elements of $\mathbf{\Gamma}_1$, $\mathbf{\Gamma}_2$ and, as a consequence of (iii), $\mathbf{\Pi}$ are all equal to 1. It follows that the diagonal elements, and hence the trace, of $\mathbf{G}$ are equal to zero. The only positive matrix with zero trace is the zero matrix. Therefore,

$$\mathbf{\Gamma}_1-\mathbf{\Pi}\circ\mathbf{\Gamma}_2=0. \tag{2.11}$$

One situation which is of particular interest is that which arises when $\mathbf{\Gamma}_2$ has no zero elements, which corresponds to all of the final states being nonorthogonal. When this is so, we can conclude that

$$\mathbf{\Pi}=\mathbf{\Gamma}_1\circ\mathbf{\Gamma}_2^{\circ -1}, \tag{2.12}$$

where $\mathbf{\Gamma}_2^{\circ -1}$ is the Hadamard inverse of $\mathbf{\Gamma}_2$. The Hadamard inverse of a matrix $\mathbf{A}=\{a_{j'j}\}$ has elements $1/a_{j'j}$. Finally, imposing condition (ii) gives

$$\mathbf{\Gamma}_1\circ\mathbf{\Gamma}_2^{\circ -1}\geqslant 0, \tag{2.13}$$

which is identical to condition (ii) in [5] for a deterministic transformation expressed in terms of Gram matrices and Hadamard product notation.

The second case we shall consider is that of unambiguous state discrimination. Here, the final states are orthonormal, so $\mathbf{\Gamma}_2=\mathbf{1}$. Let $\mathbf{\Delta}(\mathbf{p})$ be the matrix with $j'j$ element $p_j\delta_{j'j}$. Then $\mathbf{\Pi}\circ\mathbf{\Gamma}_2=\mathbf{\Delta}(\mathbf{p})$. Inserting this into (iii) gives the inequality

$$\mathbf{\Gamma}_1-\mathbf{\Delta}(\mathbf{p})\geqslant 0. \tag{2.14}$$

This is precisely the inequality obtained by Duan and Guo using a unitary-reduction scheme [17].

For a probability vector $\mathbf{p}$ which satisfies this inequality, the corresponding $\mathbf{\Pi}$ may be assumed to take a particularly simple form. If $\pi_{j'j}=\sqrt{p_{j'}p_j}$, then it can easily be shown

that $\mathbf{\Pi}$ satisfies both conditions (i) and (ii), and that (iii) is equivalent to Eq. (2.14). This $\mathbf{\Pi}$ matrix is clearly proportional to a rank-one projector.

Matrices of this form have an interesting significance in relation to the following question: under what additional conditions can the transformation $\mathcal{P}$ be carried out with probability vector $\mathbf{p}$ when only one of the $A_{kS}$ is nonzero? That is, we are interested in implementing the transformation with just two transformation operators, $A_S$ and $A_F$, respectively implementing and failing to implement the transformation, and satisfying $A_S^\dagger A_S + A_F^\dagger A_F = 1$. Here we shall show that the necessary and, for linearly independent initial states, sufficient condition for the transformation $\mathcal{P}$ be to be implementable this way with probability vector $\mathbf{p}$ is that there exists a matrix $\mathbf{\Pi}$ which, in addition to satisfying conditions (i)–(iii) above, is also proportional to a rank-one projector.

To prove necessity, we observe that for the transformation to meet our specifications, there must exist some operator $A_S$ such that

$$A_S|\psi_j^1\rangle = c_j|\psi_j^2\rangle, \tag{2.15}$$

for some coefficients $c_j$. Let us now define the $N\times N$ matrix $\mathbf{\Pi}=\{\pi_{j'j}\}$ where $\pi_{j'j}=c_{j'}^* c_j$. This matrix is clearly proportional to a rank-one projector. The proof that this matrix must satisfy the three conditions of Theorem 1 proceeds as in the more general case. It is clearly positive, and so satisfies condition (i). Condition (ii) follows from the fact that Eq. (2.15) gives

$$p_j = |c_j|^2, \tag{2.16}$$

and the derivation of condition (iii) is essentially identical to that of the more general case; obtaining it amounts to nothing more than dropping the index $k$. This proves necessary.

To prove sufficiency for linearly independent states, let $\mathbf{\Pi}$ be an $N\times N$ matrix proportional to a rank-one projector. It follows that $\pi_{j'j}=c_{j'}^* c_j$ for some $c_j$. With these coefficients, we construct the operator

$$A_S = \sum_j \frac{c_j}{\langle\tilde{\psi}_j^1|\psi_j^1\rangle}|\psi_j^2\rangle\langle\tilde{\psi}_j^1|. \tag{2.17}$$

The remainder of the proof proceeds as in the more general case. Clearly, $A_S|\psi_j^1\rangle = c_j|\psi_j^2\rangle$, as is required. The success probability for the $j$th state is $p_j = \langle\psi_j^1|A_S^\dagger A_S|\psi_j^1\rangle = \pi_{jj} = |c_j|^2$. We can finally make use of condition (iii) as before to show that $E_S = A_S^\dagger A_S \leqslant 1$.

## III. QUANTUM OPERATIONS AND MAJORIZATION

### A. Majorization relations and mixing

So far we have been considering quantum operations which convert one set of pure states into another, either deterministically or probabilistically. It is well known, however, that quantum operations often convert pure states into mixed states. This effect is often undesirable. For example, one of the principle obstacles currently in the way of realizing quantum computers is the phenomenon of decoherence, which is the mixing of the state of the computer by unwanted, uncontrollable environmental influences.

The mixing of quantum states is intimately connected with entanglement. In this example, decoherence arises due to the entanglement of the computer with the environment. If two systems become entangled, their individual states will be mixed even though the state of the entire system will remain pure.

It follows from this that measures of entanglement and mixedness ought to be intimately related, at least when the entire system is a pure, bipartite state. Indeed, the von Neumann entropy of one of subsystems simultaneously satisfies many of the natural requirements of an entanglement measure and also those of a measure of how mixed a subsystem state is. However, being a single quantity, it is unable to quantify many specific details of entanglement or mixedness, in much the same way that the Shannon entropy of a source in classical-information theory, while being sufficient to describe many important things, like the maximum asymptotically error-free transmission rate, is a less complete description of the source than the source symbols accompanied with their respective *a priori* probabilities.

In the study of pure, bipartite entanglement, the analogous, more complete description is given by the eigenvalues of the subsystem density operators. The prominence of these quantities becomes apparent when the relationship between entanglement and deterministic local operations with classical communication (LOCC) is taken into consideration. Entanglement is nonincreasing under such operations. This implies that, if one state $|\Psi_1\rangle$ can be transformed into another state $|\Psi_2\rangle$ by deterministic LOCC, then $|\Psi_2\rangle$ can be no more entangled than $|\Psi_1\rangle$ with respect to any reasonable entanglement measure. The role of the subsystem density operator eigenvalues in determining the conditions under which such a transformation is possible was made clear by Nielsen [10], who showed that the necessary and sufficient condition for such a transformation to be possible is a simple majorization relation.

In view of this and the connection between entanglement and mixing of subsystem states, we should expect majorization to play a similarly important role in describing mixedness. Indeed, that this is so was understood by Uhlmann [6–8] who pioneered the application of majorization to the study of mixing in quantum states.

It would be helpful, given current concerns about issues such as decoherence, to understand the mixing properties of quantum operations. Majorization provides an eminently suitable framework for the discussion of this issue, and our aim is to use it to help us understand the information loss, which often occurs in quantum operations and manifests itself as mixing. Fortunately, some progress has been made in this direction. Some intriguing theorems in linear algebra due to Bapat and Sunder [14] are particularly useful in this context. Here, we will employ one of their results within the framework of quantum operations, to obtain the necessary and sufficient condition for a quantum operation to increase mixing, in terms of majorization, for every initial state.

We then give an intuitive information-theoretic argument that the density operator for a pure state ensemble should not become more mixed when the pure states undergo a deterministic transformation into another set of pure states, and that the majorization relation we have hitherto considered ought not to apply (except in a certain trivial case) under such circumstances. We then prove that, in fact, it is precisely the reverse majorization relation that is true.

Prior to doing so, we will briefly review the relevant concepts. Consider two $N$ component vectors $\boldsymbol{\lambda}=\{\lambda_r\}$ and $\boldsymbol{\sigma}=\{\sigma_r\}$. The components will be taken to be real. From these vectors, we construct two further vectors $\boldsymbol{\lambda}^{\downarrow}=\{\lambda_r^{\downarrow}\}$ and $\boldsymbol{\sigma}^{\downarrow}=\{\sigma_r^{\downarrow}\}$. The components of $\boldsymbol{\lambda}^{\downarrow}$ and $\boldsymbol{\sigma}^{\downarrow}$ are those of $\boldsymbol{\lambda}$ and $\boldsymbol{\sigma}$ arranged in decreasing order. The vector $\boldsymbol{\lambda}$ is said to *majorize* the vector $\boldsymbol{\sigma}$ if the following conditions hold:

$$\sum_{r=1}^{k}\sigma_r^{\downarrow}\leqslant\sum_{r=1}^{k}\lambda_r^{\downarrow},\quad 1\leqslant k\leqslant N-1, \tag{3.1}$$

$$\sum_{r=1}^{N}\sigma_r^{\downarrow}=\sum_{r=1}^{N}\lambda_r^{\downarrow}. \tag{3.2}$$

This majorization of $\boldsymbol{\sigma}$ by $\boldsymbol{\lambda}$ is written as $\boldsymbol{\sigma}\prec\boldsymbol{\lambda}$.

In the context of probability, the vectors $\boldsymbol{\sigma}$ and $\boldsymbol{\lambda}$ are probability distributions, having positive components which satisfy $\Sigma_r\sigma_r=\Sigma_r\lambda_r=1$. The majorization relation $\boldsymbol{\sigma}\prec\boldsymbol{\lambda}$ says that the distribution $\boldsymbol{\sigma}$ is no less mixed than $\boldsymbol{\lambda}$. Two identities relating to majorization will be of particular importance in what follows. These are:

(a) The vectors $\boldsymbol{\sigma}$ and $\boldsymbol{\lambda}$ satisfy the majorization relation $\boldsymbol{\sigma}\prec\boldsymbol{\lambda}$ if and only if there is a doubly stochastic matrix $\mathbf{S}$ such that $\boldsymbol{\sigma}=\mathbf{S}\boldsymbol{\lambda}$. A doubly stochastic matrix is a matrix whose elements are real, non-negative, and where the sum of the elements in each row and column is equal to 1.

(b) If $\boldsymbol{\sigma}\prec\boldsymbol{\lambda}$, and $\lambda_r=1/N$, then $\sigma_r=1/N$ also. This is effectively a statement of the fact that if a probability distribution $\boldsymbol{\sigma}$ is no less mixed than another probability distribution $\boldsymbol{\lambda}$, and $\boldsymbol{\lambda}$ is the maximally mixed, or uniform distribution, then $\boldsymbol{\sigma}$ must also be the uniform distribution.

### B. Mixing enhancement and trace-preserving maps

Here, we will employ majorization as a tool to help us understand the increase of disorder in the state of a system which occurs in many quantum operations. We will consider a quantum system initially prepared in the state $\rho_1$ which then undergoes the transformation

$$\rho_1\rightarrow\rho_2=\sum_{k=1}^{M}A_k\rho_1A_k^{\dagger}, \tag{3.3}$$

where

$$\sum_{k=1}^{M}A_k^{\dagger}A_k=1. \tag{3.4}$$

The degree of mixedness of a quantum state is completely characterized by the density operator eigenvalues. The vector of eigenvalues of a density operator $\rho$ will be denoted by $\boldsymbol{\lambda}(\rho)$. When is it true that $\boldsymbol{\lambda}(\rho_2)\prec\boldsymbol{\lambda}(\rho_1)$, the final state $\rho_2$ can be characterized as being at least as mixed as the initial state $\rho_1$.

It is not true that for every quantum operation, the final state will always be at least as mixed as the initial state, for every initial state. For example, suppose that we carry out a von Neumann measurement in the orthonormal basis $\{|x_k\rangle\}$, and when we obtain result $k$, carry out a unitary transformation which converts the state $|x_k\rangle$ into some pure state $|x\rangle$. For this procedure, the final state will be the pure state $|x\rangle$, irrespective of the initial state, and how mixed it is. An operation of this kind, which may be viewed as an idealized kind of state preparation procedure, clearly does not increase mixedness.

The following question then arises: under what conditions does a trace-preserving quantum operation always increase mixedness or disorder in the sense of majorization, for every initial state? The answer is given by

*Theorem 2*. Consider a completely positive, linear, trace-preserving map described by Eqs. (3.3) and (3.4). The eigenvalues of the initial density operator majorize those of the final density operator, that is,

$$\boldsymbol{\lambda}(\rho_2)\prec\boldsymbol{\lambda}(\rho_1), \tag{3.5}$$

for every initial density operator $\rho_1$ if, and only if,

$$\sum_{k}A_kA_k^{\dagger}=1. \tag{3.6}$$

*Proof*. The sufficiency part of this theorem comes from a more general result due to Bapat and Sunder [14], and we will establish it by a variation on the relevant parts of their argument. Numerous extensions and consequences of their work are discussed by Visick [18]. Let $\{|\phi_r^1\rangle\}$ and $\{|\phi_{r'}^2\rangle\}$ be complete, orthonormal sets of eigenvectors of $\rho_1$ and $\rho_2$, respectively. If either density operator has zero eigenvalues, then we simply complete the orthonormal basis with an orthonormal set spanning the kernel. From Eq. (3.3), we obtain

$$\boldsymbol{\lambda}(\rho_2)=\mathbf{S}\boldsymbol{\lambda}(\rho_1), \tag{3.7}$$

where we have defined the matrix $\mathbf{S}=\{S_{r'r}\}$ with elements

$$S_{r'r}=\sum_{k}|\langle\phi_{r'}^2|A_k|\phi_r^1\rangle|^2. \tag{3.8}$$

Clearly, $S_{r'r}$ is real and nonnegative. The majorization relation (3.5) will hold for every initial density operator $\rho_1$ if $\mathbf{S}$ is doubly stochastic, which will be the case if the row and column sums of $\mathbf{S}$ are equal to one. For the column sum, we have

$$\sum_{r'}S_{r'r}=\langle\phi_r^1|\left[\sum_{k}A_k^{\dagger}A_k\right]|\phi_r^1\rangle=1, \tag{3.9}$$

as a consequence of the completeness of the $\{|\phi_{r'}^2\rangle\}$ and the resolution of the identity in Eq. (3.4). For the row sum, we see that

$$\sum_r S_{r'r} = \langle \phi^2_{r'} | \left[ \sum_k A_k A_k^\dagger \right] | \phi^2_{r'} \rangle = 1, \tag{3.10}$$

when Eq. (3.6) holds, where we have used the completeness of the $\{|\phi^1_r\rangle\}$. So, when Eq. (3.6) is true, the matrix **S** is doubly stochastic and the majorization relation in Eq. (3.5) holds for every initial density operator $\rho_1$. This proves sufficiency.

To prove necessity, we must show that Eq. (3.6) follows if the majorization relation (3.5) is true for every initial density operator $\rho_1$. Actually, we need only consider the case when $\rho_1$ is the maximally mixed state, that is, $\rho_1 = 1/D$, which implies that $\lambda_r(\rho_1) = 1/D$. This, together with the identity (b), suffices to determine the final density operator $\rho_2$ completely. As a consequence of identity (b), the only possible choice for $\boldsymbol{\lambda}(\rho_2)$ which is consistent with $\lambda_r(\rho_1) = 1/D$ is $\boldsymbol{\lambda}(\rho_2) = \boldsymbol{\lambda}(\rho_1)$. It follows that $\rho_2$ must also be the maximally mixed state. Inserting $\rho_1 = \rho_2 = 1/D$ into Eq. (3.3), and multiplying both sides by $D$ immediately gives Eq. (3.6), completing the proof. ■

Condition (3.6) is always satisfied if the $A_k$ are normal operators, which is a sufficient condition for the sums in Eqs. (3.4) and (3.6) to be identical. It follows that any generalized measurement described by a POVM with elements $E_k$ will satisfy Eq. (3.6) if we choose the transformation operators to be $A_k = \sqrt{E_k}$. This choice of transformation operators for a generalized measurement has been termed the ''rawest'' implementation by Fuchs and Jacobs [12]. Theorem 2 gives this ''rawness'' a concrete meaning. The term ''raw'' has connotations of simplicity and unembellishment. These descriptions fit this implementation of a generalized measurement, reflecting as they do the absence of an attempt to restore or increase the purity of the state following acquisition of the measurement outcome. This is captured by the majorization relation (3.5).

It is instructive to compare and contrast Theorem 2 with a related theorem due to Uhlmann [6–8]. This states that the eigenvalues of two density operators $\rho_1$ and $\rho_2$ obey the majorization relation $\boldsymbol{\lambda}(\rho_2) \prec \boldsymbol{\lambda}(\rho_1)$ if and only if there exists a probability distribution $p_k$ and unitary operators $U_k$ such that

$$\rho_2 = \sum_k p_k U_k \rho_1 U_k^\dagger . \tag{3.11}$$

For a further proof and discussion of this theorem, see Wehrl [19]. Nielsen and Chuang [1] also give a particularly direct proof whose sufficiency part parallels that of the proof we have given of Theorem 2 above. It is obvious that Eq. (3.11) is a valid quantum operation, indeed one which satisfies Eq. (3.4). In fact, the sufficiency part of Uhlmann's theorem is easily seen to follow from the sufficiency part of Theorem 2 in the special case where $A_k = \sqrt{p_k} U_k$.

The necessity parts of Uhlmann's theorem and Theorem 2 are, on the other hand, disjoint. In Uhlmann's theorem, the emphasis is on the density operators. It says that if $\boldsymbol{\lambda}(\rho_2) \prec \boldsymbol{\lambda}(\rho_1)$ then there must be a probability distribution $p_k$ and unitary operators $U_k$ *which depend on the initial and final density operators* and satisfy Eq. (3.11). In contrast, the emphasis in Theorem 2 is on the quantum operation, *which is independent of the density operators* and makes a statement about the properties that a particular operation must have if it is never to decrease mixedness for any density operator.

Any quantum operation which satisfies Eqs. (3.4) and (3.6) is said to be doubly stochastic. As it happens, for a two-dimensional quantum system, any doubly stochastic quantum operation may be written in the form (3.11), that is, as a convex combination of unitary operations. This point is discussed by Ruskai *et al.* [20]. However, for higher-dimensional systems, this is not always possible, as was shown by Landau and Streater [21].

### C. Majorization and deterministic transformations

Only operations which satisfy condition (3.6) do not decrease mixedness, in the sense quantified by majorization, for any state. A well-known property of majorization is that if $\boldsymbol{\lambda}(\rho_2) \prec \boldsymbol{\lambda}(\rho_1)$, then $S(\rho_2) \geqslant S(\rho_1)$, where $S(\rho) = -\mathrm{Tr}(\rho \log \rho)$ is the von Neumann entropy. It follows that if $\rho_2$ is at least as mixed as $\rho_1$ in the sense of majorization, then its von Neumann entropy is also at least as high as that of $\rho_1$.

The von Neumann entropy has long been used to quantify mixedness, in the sense of *disorder*, in quantum mechanics. However, with the advent of the noiseless coding theorems for classical and quantum information transmission, it has acquired further significance as a measure of information which is directly analogous to that of the Shannon entropy in classical information theory. In this context, the density operator represents an ensemble of pure states. Consider two ensembles $\mathcal{E}_1 = \{q_j, |\psi^1_j\rangle\}$ and $\mathcal{E}_2 = \{q_j, |\psi^2_j\rangle\}$, where $q_j$ is the *a priori* probability of both $|\psi^1_j\rangle$ and $|\psi^2_j\rangle$. These ensembles have the density operators

$$\rho_1(\mathbf{q}) = \sum_j q_j |\psi^1_j\rangle\langle\psi^1_j| , \tag{3.12}$$

$$\rho_2(\mathbf{q}) = \sum_j q_j |\psi^2_j\rangle\langle\psi^2_j| . \tag{3.13}$$

The noiseless coding theorem for classical [22] (quantum [23]) information with pure quantum states implies that the maximum rate of asymptotically error-free classical (quantum) information transmission using ensemble $\mathcal{E}_i$ is $S(\rho_i(\mathbf{q}))$ bits (qubits) per signal. However, suppose that we could transform $\mathcal{E}_1$ into $\mathcal{E}_2$ with unit probability. If $S(\rho_2(\mathbf{q})) > S(\rho_1(\mathbf{q}))$, then clearly these coding theorems would be violated. Such ensemble transformations, which increase the von Neumann entropy, must be impossible, and lead us to suspect that ensemble transformations giving rise to the majorization relation $\boldsymbol{\lambda}(\rho_2(\mathbf{q})) \prec \boldsymbol{\lambda}(\rho_1(\mathbf{q}))$ will also be impossible [except in the trivial case where all of the equalities in Eq. (3.1) are satisfied].

The transformations we have in mind here are clearly the deterministic transformations described in the preceding section. The question is then: do the eigenvalues of ensemble density operators whose constituent pure states are related by

a deterministic transformation obey *any* majorization relation? The answer, as we will now see, is yes: it is precisely the reverse of that considered in Theorem 2, which is highly satisfactory in view of the above considerations.

*Theorem 3*. Let $\{|\psi_j^1\rangle\}$ and $\{|\psi_j^2\rangle\}$ be sets of $N$ pure states. Consider the mixtures $\rho_1(\mathbf{q})$ and $\rho_2(\mathbf{q})$ defined by Eqs. (3.12) and (3.13). If there is a deterministic transformation $\mathcal{D}:|\psi_j^1\rangle\rightarrow|\psi_j^2\rangle\ \forall\ j$, then

$$\boldsymbol{\lambda}[\rho_1(\mathbf{q})]\prec\boldsymbol{\lambda}[\rho_2(\mathbf{q})], \tag{3.14}$$

for every *a priori* probability vector $\mathbf{q}$.

Prior to proving this, we note that to speak of majorization relations, $\rho_1(\mathbf{q})$ and $\rho_2(\mathbf{q})$ must have the same number of eigenvalues. This condition is easily satisfied by ''padding out'' the spectrum with the lower number of nonzero eigenvalues with zeroes so that the spectra of both density operators are of equal size.

*Proof*. We start with the following observation made by Jozsa and Schlienz [24]. For the *a priori* probability vector $\mathbf{q}$, we define the matrix $\mathbf{Q}=\{\sqrt{q_jq_{j'}}\}$. Then $\rho_1(\mathbf{q})$ has the same nonzero eigenvalues, with the same multiplicities, as $\mathbf{Q}\circ\boldsymbol{\Gamma}_1$, and likewise with $\rho_2(\mathbf{q})$ and $\mathbf{Q}\circ\boldsymbol{\Gamma}_2$. To see why, consider the following entangled state of two systems, $a$ and $b$:

$$|\Phi\rangle=\sum_j\ \sqrt{q_j}|\psi_j\rangle_a\otimes|x_j\rangle_b\,, \tag{3.15}$$

where $\{|\psi_j\rangle\}$ may be either the set $\{|\psi_j^1\rangle\}$ or $\{|\psi_j^2\rangle\}$, and $\{|x_j\rangle\}$ is an orthonormal set. The purity of this state implies that the eigenvalues of the reduced density operators are the same for each subsystem. We find that

$$\rho_a=\sum_j\ q_j|\psi_j\rangle\langle\psi_j|, \tag{3.16}$$

$$\rho_b=\sum_{jj'}\ \sqrt{q_jq_{j'}}\langle\psi_j|\psi_{j'}\rangle|x_{j'}\rangle\langle x_j|,$$

$$=\sum_{jj'}\ \{(\mathbf{Q}\circ\boldsymbol{\Gamma})^T\}_{j'j}|x_{j'}\rangle\langle x_j|, \tag{3.17}$$

where $\boldsymbol{\Gamma}$ is the Gram matrix of the set $\{|\psi_j\rangle\}$. Equation (3.17) tells us that the elements of $\rho_b$ in the $\{|x_j\rangle\}$ basis give the matrix $(\mathbf{Q}\circ\boldsymbol{\Gamma})^T$, where the superscript $T$ denotes the transpose. Any Hermitian matrix has the same nonzero eigenvalues as its transpose (with corresponding eigenvectors being related by complex conjugation in the standard basis). So, we see that $\rho_a$ and $\mathbf{Q}\circ\boldsymbol{\Gamma}$ have the same nonzero eigenvalues. This implies that

$$\boldsymbol{\lambda}[\rho_1(\mathbf{q})]\prec\boldsymbol{\lambda}[\rho_2(\mathbf{q})]\Leftrightarrow\boldsymbol{\lambda}(\mathbf{Q}\circ\boldsymbol{\Gamma}_1)\prec\boldsymbol{\lambda}(\mathbf{Q}\circ\boldsymbol{\Gamma}_2)\forall\,\mathbf{q}. \tag{3.18}$$

Consequently, we will be able to establish the majorization relation (3.14) if we can establish that on the right-hand side of Eq. (3.18). It turns out that the latter relation can be proven rather straightforwardly using the following result obtained by Bapat and Sunder [14]: let $\mathbf{A}$ and $\mathbf{B}$ be $N\times N$ Hermitian matrices. If $\mathbf{A}\geqslant 0$, and the diagonal elements of $\mathbf{A}$ are all equal to 1, then [25]

$$\boldsymbol{\lambda}(\mathbf{A}\circ\mathbf{B})\prec\boldsymbol{\lambda}(\mathbf{B}). \tag{3.19}$$

Let us apply this relation, making the identifications

$$\mathbf{A}=\boldsymbol{\Pi}, \tag{3.20}$$

$$\mathbf{B}=\mathbf{Q}\circ\boldsymbol{\Gamma}_2, \tag{3.21}$$

where $\boldsymbol{\Pi}$ is a positive, $N\times N$ matrix with diagonal elements equal to 1 and $\mathbf{B}$ is easily shown to be Hermitian, indeed positive, as a consequence of the positivity of Gram matrices and projectors ($\mathbf{Q}$ clearly being a projector) and Schur's product theorem [26], which states that the Hadamard product of two positive matrices is also positive. Substituting these definitions into Eq. (3.19) gives

$$\boldsymbol{\lambda}[\boldsymbol{\Pi}\circ(\mathbf{Q}\circ\boldsymbol{\Gamma}_2)]\prec\boldsymbol{\lambda}(\mathbf{Q}\circ\boldsymbol{\Gamma}_2). \tag{3.22}$$

We know from Eq. (2.11) that for a deterministic transformation, there exists a positive matrix $\boldsymbol{\Pi}$ such that $\boldsymbol{\Gamma}_1-\boldsymbol{\Pi}\circ\boldsymbol{\Gamma}_2=0$. We can see from this equation that the diagonal elements of $\boldsymbol{\Pi}$ must be equal to 1. Making use of the commutativity of the Hadamard product, we can see that, for a deterministic transformation

$$\mathbf{Q}\circ\boldsymbol{\Gamma}_1=\boldsymbol{\Pi}\circ(\mathbf{Q}\circ\boldsymbol{\Gamma}_2). \tag{3.23}$$

We can substitute $\mathbf{Q}\circ\boldsymbol{\Gamma}_1$ into the left-hand side of Eq. (3.22), giving the majorization relation on the right-hand side of (3.18). This completes the proof. ■

A question of obvious importance is whether or not the converse of Theorem 3 is true, that is, whether or not satisfaction of the majorization relation (3.14) is a sufficient condition for the existence of a deterministic transformation $\mathcal{D}:|\psi_j^1\rangle\rightarrow|\psi_j^2\rangle\forall\,j$. At the time of writing, this question is open. If it is ever to be answered in the affirmative, then this could suggest an interesting parallel between the theory of deterministic transformations of sets of pure states, and that of deterministic LOCC on pure, bipartite entangled states, which is covered by a theorem due to Nielsen which we mentioned earlier. To be specific, let $|\Psi_1\rangle$ and $|\Psi_2\rangle$ be a pair of pure, bipartite entangled states, and $\rho_1,\rho_2$ be the corresponding reduced density operators for one of the subsystems. Then Nielsen's theorem [10] states the necessary and sufficient condition for the existence of a deterministic LOCC procedure which transforms $|\Psi_1\rangle$ into $|\Psi_2\rangle$ is

$$\boldsymbol{\lambda}(\rho_1)\prec\boldsymbol{\lambda}(\rho_2). \tag{3.24}$$

The similarity between Eqs. (3.14) and (3.24) is striking, especially when we consider the fact that, in both contexts, the mixing, whose nonincrease is expressed by the appropriate majorization relation, is related to a useful quantity or resource, rather than simple disorder. In the context of deterministic transformations of sets of pure states, the degree of mixing can be intuitively understood as expressing the distinguishability of the set of states, at least when they are

linearly independent. We feel that a further open problem, whose solution may require that of the preceding one, is how one can make this intuition quantitatively precise.

In the second context, that of deterministic LOCC entanglement transformations, the degree of mixing relates to how entangled the state is. The fact that useful quantities such as entanglement and distinguishability cannot increase under the appropriate kinds of deterministic transformation, and that this fact can be expressed by simple, similar majorization relations suggests that both scenarios are related, and that this relationship could be understood in terms of some broader, as yet unproposed unifying framework.

## IV. DISCUSSION

In this paper, we have obtained some general results relating to transformations of quantum states, and associated probabilities or density operator eigenvalues, which are closely related to and in some contexts can be interpreted as probabilities. The main emphasis has been on transformations of pure states. Probabilities play an essential role in quantum mechanics in quantifying the likelihood of a particular measurement outcome, given certain information about how the system was prepared, namely, its initial state vector or, more generally, density operator. This has been known since the early days of quantum theory. However, in recent decades, it has become apparent, though a careful analysis of the postulates of quantum mechanics and exploiting the possibilities afforded by interactions between quantum systems, that the quantum formalism permits more general measurements than those whose outcome probabilities are obtained by direct application of Born's rule, and where the resulting post-measurement states are obtained by direct application of the von Neumann–Lüders projection postulate. Such measurements are known as generalized measurements. The formalism of quantum operations, which describes both aspects of this general measurement process, has been of enormous interest recently, especially due to its relevance to the developing field of quantum-information theory.

Since the early days of quantum theory, it was recognized that the measurement process is inextricably bound up with a disturbance of the state of the system. With the development of generalized measurements, it has become recognized that the large disturbance associated with a sharp, von Neumann measurement is an extremal case of a general trade-off between information and disturbance [12,27]. In this context, information is treated as a ''good'' thing, while the associated disturbance is considered to be an undesirable but unavoidable by-product. However, in situations where we aim to tailor the disturbance to produce a particular state, and where we wish to minimize the probability of other transformations being carried out, it is almost as though the conventional ''morality'' of the information/disturbance trade-off is inverted.

Specific probabilistic transformations, such as cloning and unambiguous state discrimination (which is a probabilistic transformation of a nonorthogonal set into an orthogonal set) have been studied in detail. A further kind of transformation which has been examined is a deterministic transformation, which converts one set of pure states into another with unit probability. However, probabilistic transformations, of which deterministic transformations represent a limiting case, have not previously been investigated in full generality. To do so was the objective of Sec. II. For pure state transformations, we derived necessary and, when the initial states are linearly independent, sufficient conditions for such a transformation to be possible with given conditional probabilities for each of the states.

Extending our analysis to cover more general quantum operations, it is easily shown that the purity of states is not preserved in general. For the sake of simplicity, the probabilistic assumption was removed and our emphasis shifted from selective to nonselective operations. This scenario is of considerable practical importance since it applies to a quantum system whose state we wish to control, deterministically, such as that of a quantum computer, but which is subject to uncontrollable influences such as that of the environment.

One of the most basic questions we can ask about such quantum operations is: under what circumstances is the final state always at least as mixed as the initial state, for every possible initial state? Quantifying the extent to which a state is mixed, at least when the Hilbert space dimension exceeds two, is nontrivial. However, under certain circumstances, we can unambiguously compare the degree of mixing of two quantum states for arbitrary quantum systems: specifically, when the eigenvalues of one density operator majorize those of the other. The nontriviality of mixing comparison is quantitatively captured by the fact that majorization enforces only a partial order on equivalence classes of density operators (with respect to unitary symmetry) which allows for incomparable states. We described how a simple, elegant, sufficient condition obtained by Bapat and Sunder is also necessary. We then showed that the eigenvalues of the source density operators for initial and final pure state ensembles related by a deterministic transformation obey the opposite majorization relation. In this context, mixing, rather than characterizing disorder, is related to the information content or distinguishability of the ensemble, and this majorization relation expresses the fact that such aspects of an ensemble cannot be amplified and is perhaps in the same spirit as the no-cloning theorem. Indeed, it is quite simple to show that the strong form of the no-cloning theorem, which states that it is impossible to deterministically copy a set of nonorthogonal states, follows from this majorization relation.

We noted the resemblance between this majorization relation and that obtained by Nielsen in the context of LOCC entanglement transformations. The latter is a necessary and sufficient condition for deterministic LOCC transformation of one pure, bipartite entangled state into another. The former is only known to be a necessary condition for a deterministic transformation of one set of pure states into another pure set. We argued that if it can also be shown to be sufficient, then there is the possibility that deterministic LOCC and pure set transformations could be incorporated within and understood in terms of a broader encompassing framework. This could lead to interesting insights into the relationship between entanglement and distinguishability.

With this possibility in mind, let us consider the fact that the majorization relation (3.14) implies that the quantities

$$\mu_k(\mathbf{q}) = \sum_{r=1}^{k} \lambda_r^{\downarrow}\left[\sum_j q_j |\psi_j^1\rangle\langle\psi_j^1|\right], \quad 1 \leqslant k \leqslant N \qquad (4.1)$$

are nonincreasing under any deterministic transformation $\mathcal{D}$: $|\psi_j^1\rangle \rightarrow |\psi_j^2\rangle \forall j$ for any set of final states $\{|\psi_j^2\rangle\}$. Can we refer to such quantities as "distinguishability monotones," by analogy with the concept of entanglement monotones introduced by Vidal [28]? If so, then how are they related to operations which distinguish between quantum states? How are they related to more general sets of distinguishability monotones? Indeed, what criteria are the necessary and sufficient conditions to qualify a functional as being a distinguishability monotone, or measure? To answer these questions, we would require a greater understanding of the distinguishability of sets of pure quantum states, comparable to that which we have of pure, bipartite entanglement.

## ACKNOWLEDGMENT

This work was supported by the UK Engineering and Physical Sciences Research Council.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[2] L. Hardy and D. D. Song, Phys. Rev. A **63**, 032304 (2001).
[3] G. Alber, A. Delgado, and I. Jex, LANL e-print quant-ph/0006040.
[4] J. Fiurášek, Phys. Rev. A **64**, 062301 (2001).
[5] A. Chefles, Phys. Lett. A **270**, 14 (2000).
[6] A. Uhlmann, Wiss. Z. Karl-Marx-Univ. Leipzig. **20**, 633 (1971).
[7] A. Uhlmann, Wiss. Z. Karl-Marx-Univ. Leipzig. **21**, 427 (1972).
[8] A. Uhlmann, Wiss. Z. Karl-Marx-Univ. Leipzig. **22**, 139 (1973).
[9] R. Bhatia, *Matrix Analysis* (Springer-Verlag, Berlin, 1991).
[10] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[11] M. A. Nielsen, Phys. Rev. A **63**, 022114 (2001).
[12] C. A. Fuchs and K. Jacobs, Phys. Rev. A **63**, 062305 (2001).
[13] M. A. Nielsen, Phys. Rev. A **62**, 052308 (2000).
[14] R. B. Bapat and V. C. Sunder, Linear Algebra Appl. **72**, 107 (1985).
[15] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, Berlin, 1983).
[16] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[17] L.-M. Duan and G.-C. Guo, Phys. Rev. Lett. **80**, 4999 (1998).
[18] G. Visick, Linear Algebra Appl. **304**, 45 (2000).
[19] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[20] M. B. Ruskai, S. Szarek, and E. Werner, LANL e-print quant-ph/0101003.
[21] L. J. Landau and R. F. Streater, Linear Algebra Appl. **193**, 107 (1993).
[22] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[23] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[24] R. Jozsa and J. Schlienz, Phys. Rev. A **62**, 012301 (2000).
[25] The majorization relation (3.19) is actually a special case of the following more general result discussed by Visick [18] : let $\mathbf{A}$ and $\mathbf{B}$ be $N \times N$ Hermitian matrices. If $\mathbf{A} \geqslant 0$ then $\boldsymbol{\lambda}(\mathbf{A} \circ \mathbf{B}) < \boldsymbol{\lambda}[(1 \circ \mathbf{A})\mathbf{B}]$. This reduces to the majorization relation (3.19) in the special case where all of the diagonal elements of $\mathbf{A}$ are equal to 1, in which case $\mathbf{1} \circ \mathbf{A} = \mathbf{1}$.
[26] R. Horn and C. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).
[27] C. A. Fuchs and A. Peres, Phys. Rev. A **53**, 2038 (1996).
[28] G. Vidal, J. Mod. Opt. **47**, 355 (2000).