Distinguishability measures and ensemble orderings

Anthony Chefles
Department of Physical Sciences, University of Hertfordshire, Hatfield AL10 9AB, Hertfordshire, United Kingdom
(Received 2 October 2001; published 30 October 2002)

It is shown that different distinguishability measures impose different orderings on ensembles of N pure quantum states. This is demonstrated using ensembles of equally probable, linearly independent, symmetrical pure states, with the maximum probabilities of correct hypothesis testing and unambiguous state discrimination being the distinguishability measures. This finding implies that there is no absolute scale for comparing the distinguishability of any two ensembles of N quantum states, and that distinguishability comparison is necessarily relative to a particular discrimination strategy.

DOI: 10.1103/PhysRevA.66.042325 PACS number(s): 03.67.Hk, 03.65.Ta

I. INTRODUCTION

The use of quantum channels to send classical information has many advantages over the use of classical channels. One of these is the fact that it enables one to establish cryptographic keys in a way that is provably secure, a feat which has never been achieved classically. Developments such as this have led to renewed attention being given to the problem of distinguishing between quantum states [1]. In state discrimination, we are faced with the following situation: a quantum system is prepared in one of N states. For the sake of simplicity, we will take these to be pure states $|\psi_j\rangle$, where $j=1, \ldots ,N$. The *a priori* probability of the state of the system being $|\psi_j\rangle$ is p_j. We would like to determine which state has been prepared. Unless the states are orthogonal we cannot determine the state perfectly. We are then faced with the problem of devising a strategy that discriminates between the N potential states as well as possible. This will involve some, possibly generalized, quantum measurement, which should be optimized. The resulting figure of merit, which is typically a probability, can be regarded as a measure of the distinguishability of the states with these *a priori* probabilities. A set of quantum states $|\psi_j\rangle$ considered together with their *a priori* probabilities p_j forms an ensemble $\mathcal{E}=\mathcal{E}(p_j,|\psi_j\rangle)$. So, distinguishability measures will refer to ensembles.

We shall denote a generic distinguishability measure by $D[\mathcal{E}]$. Several distinct measures are in common use for quantifying the distinguishability of states [1]. The question we address in this paper is the following: do all distinguishability measures impose the same ordering on ensembles of N quantum states? That is, suppose that we have two ensembles $\mathcal{E}_1$ and $\mathcal{E}_2$ and two distinguishability measures D_1 and D_2. Then, if $D_1[\mathcal{E}_1]>D_1[\mathcal{E}_2]$, is it the case that $D_2[\mathcal{E}_1]\geqslant D_2[\mathcal{E}_2]$? We shall see that this is not necessarily so.

The distinguishability measures we choose are the maximum probabilities of correct hypothesis testing and unambiguous state discrimination. For the sake of simplicity, we restrict our attention to ensembles of states that have equal *a priori* probabilities. We show that, for two equally probable pure states, this effect cannot be demonstrated for these distinguishability measures, which leads us to consider ensembles of $N>2$ states. We restrict our attention to linearly independent ensembles, since this is a requirement for unambiguous state discrimination [2]. The states in the chosen ensembles are taken to form symmetrical sets. This is done to take advantage of the fact that, for such ensembles, our distinguishability measures can be calculated analytically for equal *a priori* probabilities. The effect we have described is shown to occur for $N=3$. We conclude with a discussion of the connection between this effect and a related one recently discovered by Jozsa and Schlienz [3].

II. DISTINGUISHABILITY MEASURES IMPOSING DIFFERENT ENSEMBLE ORDERINGS

The distinguishability measures we choose are the maximum probabilities of correct hypothesis testing and unambiguous state discrimination. In hypothesis testing among N states in the ensemble $\mathcal{E}(p_j,|\psi_j\rangle)$, we consider an N-outcome generalized measurement, in which the jth outcome is associated with the positive operator valued measure (POVM) element E_j. Our hypothesis is that the outcome of the measurement corresponds exactly to the state. If the state is $|\psi_j\rangle$ and outcome j is obtained, then the hypothesis is correct. If, however, outcome $j'\neq j$ is obtained, then the hypothesis will be incorrect. The maximum probability P_{HYP} that our hypothesis is correct is [4]

$$P_{HYP}(\mathcal{E})=\max_{\{E_j\}}\sum_j p_j\langle\psi_j|E_j|\psi_j\rangle, \tag{2.1}$$

where the maximization is carried out over all sets of N positive operators E_j such that $\Sigma_j E_j=1$. We shall use the maximum probability of correct hypothesis testing, $P_{HYP}(\mathcal{E})$, as a measure of the distinguishability of the ensemble $\mathcal{E}$.

In unambiguous state discrimination, there are only two possible outcomes for the state $|\psi_j\rangle$: outcome j and a further inconclusive result "?." There are no errors. Unlike hypothesis testing, unambiguous discrimination is only possible for linearly independent sets [2] and it is to such sets that we will restrict our attention. If $P_{USD}(|\psi_j\rangle)$ is the probability that, given the initial state was $|\psi_j\rangle$, we obtain a conclusive identification rather than an inconclusive result, then the maximum probability of unambiguous state discrimination is

$$P_{USD}(\mathcal{E}) = \max_{\{P_{USD}(|\psi_j\rangle)\}} \sum_j p_j P_{USD}(|\psi_j\rangle), \tag{2.2}$$

where the extremization with respect to the $P_{USD}(|\psi_j\rangle)$ is discussed by Duan and Guo [5]. The distinguishability measure we will use for this strategy will be the maximum probability $P_{USD}(\mathcal{E})$ of unambiguous determination of the state.

Here, we will show that there exist ensemble pairs for which

$$P_{USD}(\mathcal{E}_2) < P_{USD}(\mathcal{E}_1), \tag{2.3}$$

$$P_{HYP}(\mathcal{E}_2) > P_{HYP}(\mathcal{E}_1). \tag{2.4}$$

The effect we wish to demonstrate does not occur for an ensemble of just two equally probable pure states. This can be seen by examining the values of P_{HYP} and P_{USD} for a pair of pure states $|\psi_1\rangle$ and $|\psi_2\rangle$. For later convenience, we will give expressions for these with arbitrary *a priori* probabilities p_1 and p_2. The former is given by Helstrom's bound [4]:

$$P_{HYP} = \frac{1}{2}(1 + \sqrt{1-(1-\Delta^2)|\langle\psi_1|\psi_2\rangle|^2}), \tag{2.5}$$

where $\Delta = |p_1 - p_2|$. Also, the maximum value of P_{USD} for a pair of pure states is given by the Jaeger-Shimony bound [6]

$$P_{USD} = \begin{cases} 1-\sqrt{1-\Delta^2}|\langle\psi_1|\psi_2\rangle|, & \sqrt{\dfrac{1-\Delta}{1+\Delta}} \geqslant |\langle\psi_1|\psi_2\rangle| \\ \frac{1}{2}(1+\Delta)(1-|\langle\psi_1|\psi_2\rangle|^2), & \sqrt{\dfrac{1-\Delta}{1+\Delta}} \leqslant |\langle\psi_1|\psi_2\rangle|. \end{cases} \tag{2.6}$$

When the *a priori* probabilities are equal, $\Delta = 0$ and $P_{HYP} = (1/2)(1+\sqrt{1-[1-(1-P_{USD})^2]})$. From this, one can show that P_{HYP} is an increasing function of P_{USD}, implying that inequalities (2.3) and (2.4) can never be simultaneously satisfied.

To find ensembles of equally probable pure states for which both (2.3) and (2.4) are true, we have to consider at least three states. We focus on ensembles of equally probable, linearly independent, symmetrical states, since the maximum probabilities for unambiguous discrimination and correct hypothesis testing can be calculated explicitly for these. We will demonstrate the existence of ensemble pairs $\mathcal{E}_1$ and $\mathcal{E}_2$ that satisfy inequalities (2.3) and (2.4), in the following way. First, we will consider *all* sets of N linearly independent, symmetric states with equal *a priori* probabilities that have the same, arbitrary but fixed, value of P_{USD}. Over this set, we will find the extremal values of P_{HYP}. Using this information, we choose ensemble pairs that satisfy inequality (2.3), but where $\mathcal{E}_1$ and $\mathcal{E}_2$ have, respectively, the minimum and maximum values of P_{HYP} for their corresponding values of P_{USD}. We will find that, for $N=3$, inequality (2.4) is satisfied for a large range of parameters.

The N pure states $|\psi_j\rangle$, where $j=0,\ldots,N-1$, are linearly independent and symmetric if and only if they can be written as

$$|\psi_j\rangle = \sum_{r=0}^{N-1} c_r e^{2\pi i j r/N}|x_r\rangle, \tag{2.7}$$

for some N orthonormal states $|x_r\rangle$ and nonzero complex coefficients c_r satisfying $\Sigma_r|c_r|^2=1$. Notice that the phase of c_r may be absorbed by $|x_r\rangle$, which implies that we may, without loss of generality, take c_r to be real and nonnegative, which we shall. We will take all states to have equal *a priori* probabilities $p_j = 1/N$. The maximum unambiguous discrimination probability for these states is [7]

$$P_{USD} = N \times \min_r c_r^2. \tag{2.8}$$

The optimum hypothesis testing strategy uses the so-called "square-root" measurement [4]. Define the operator

$$\Phi = \sum_j |\psi_j\rangle\langle\psi_j| \tag{2.9}$$

and the states

$$|\omega_j\rangle = \Phi^{-1/2}|\psi_j\rangle. \tag{2.10}$$

One can quite easily show that the operators $E_r = |\omega_r\rangle\langle\omega_r|$ form a POVM (i.e., that $E_r \geqslant 0$ and $\Sigma_r E_r = 1$.) This POVM is the optimal hypothesis testing strategy, and we find that

$$P_{HYP} = \frac{1}{N}\left(\sum_r c_r\right)^2. \tag{2.11}$$

The square-root measurement optimally discriminates between any set of equally probable, symmetric states, even if they are not linearly independent. This measurement has recently been carried out [8,9] for symmetrical optical polarization states. Applications of this measurement to quantum key distribution are discussed in [10].

We now calculate the global extrema of P_{HYP} for a fixed value of P_{USD}. Our aim is to fix the smallest of the c_r, which is equivalent to fixing P_{USD}, and vary the remaining coefficients to obtain the extremal values of P_{HYP}. We may let $\min_r c_r = c_0$. This allows us to write

$$c_r = c_0 + \cos^2\theta_r, \tag{2.12}$$

for $r=1,\ldots,N-1$ and some angles θ_j. We will now extremize P_{HYP} with respect to these angles using the method of Lagrange multipliers, in order to take into account the normalization constraint. This method will yield the local extrema, over which we will subsequently optimize to find the global extrema. Let $G = (\Sigma_{r=0}^{N-1} c_r^2) - 1$. The constrained, local extrema of P_{HYP} occur where

$$\frac{\partial P_{HYP}}{\partial \theta_r} = \lambda \frac{\partial G}{\partial \theta_r} \tag{2.13}$$

and λ is our Lagrange multiplier. Inserting Eq. (2.12) into the definitions of P_{HYP} and G, we find that this becomes

$$\sin\theta_r\cos\theta_r\left[c_0+\frac{1}{N}\sum_{r'=1}^{N-1}\cos^2\theta_{r'}-\lambda(c_0+\cos^2\theta_r)\right]=0. \tag{2.14}$$

It is a simple matter to show that the normalization requirement will be violated if $\sin\theta_r=0$ for any r. So, for each r, either $\cos\theta_r=0$, in which case $c_r=c_0$, or the term in brackets is zero. To proceed, we will partition the c_r into two sets, where each set corresponds to one of these two possibilities. Let there be N_0 coefficients, including c_0, which are equal to c_0. For the remaining $N-N_0$ coefficients, the term in brackets is zero. All of the coefficients in this latter set must also be equal, as can be seen from the fact that the vanishing of the bracket term implies that the corresponding $\cos^2\theta_r$ are equal. The value that they have is easily deduced from the fact that the other N_0 coefficients are equal to c_0, and normalization. We find they have the value

$$c_r=\sqrt{\frac{1-N_0c_0^2}{N-N_0}}. \tag{2.15}$$

We can now calculate the local extrema of P_{HYP}. Evaluating Eq. (2.11) where N_0 coefficients are equal to c_0, with the remaining $N-N_0$ coefficients being given by Eq. (2.15), and making use of the fact that $c_0=\sqrt{P_{USD}/N}$, we find that the local extrema of P_{HYP} for fixed P_{USD} are

$$P_{HYP}=\frac{1}{N^2}[N_0\sqrt{P_{USD}}+\sqrt{(N-N_0)(N-N_0P_{USD})}]^2. \tag{2.16}$$

Finding the global extrema of P_{HYP} amounts to extremization of this expression with respect to N_0. One method of doing this is to treat N_0 as a continuous parameter in the interval $[1,N-1]$ and differentiate Eq. (2.16) with respect to it. After some algebra, we find that $\partial P_{HYP}/\partial N_0\leqslant 0$, for P_{HYP} given by Eq. (2.16). This implies that the global maximum and minimum of P_{HYP} occur at the minimum and maximum values of N_0, respectively. Thus, the maximum value of P_{HYP} occurs at $N_0=1$, which gives us the tight upper bound

$$P_{HYP}\leqslant\frac{1}{N^2}[\sqrt{P_{USD}}+\sqrt{(N-1)(N-P_{USD})}]^2. \tag{2.17}$$

For $P_{USD}=1$, we see that the maximum of P_{HYP} is also equal to 1. However, when $P_{USD}=0$, the maximum of P_{HYP} is $(N-1)/N$.

We also see that the minimum value of P_{HYP} occurs at $N_0=N-1$, which gives us the tight lower bound

$$P_{HYP}\geqslant\frac{1}{N^2}[(N-1)\sqrt{P_{USD}}+\sqrt{N-(N-1)P_{USD}}]^2. \tag{2.18}$$

When $P_{USD}=1$, the minimum of P_{HYP} is also 1, as we would expect. However, if $P_{USD}=0$, then the minimum of P_{HYP} is $1/N$, which corresponds to a random guess of which of the N possible states the system has been prepared in. In fact, it is quite easily shown that in this case, all $|\psi_j\rangle$ are, up to a phase, equal to $|x_0\rangle$. Here, the states are linearly dependent and this case is of no interest to us in the present context.

Let us now use the bounds in (2.17) and (2.18) to demonstrate the existence of ensembles for which inequalities (2.3) and (2.4) hold. For the sake of simplicity, we will consider only ensembles for which $N=3$. We will choose $\mathcal{E}_2$ to be an ensemble that saturates the bound in (2.17) and $\mathcal{E}_1$ to be one which saturates the bound in (2.18), yet where $P_{USD}(\mathcal{E}_2)=P_{USD}(\mathcal{E}_1)-\epsilon$, for some positive, nonzero parameter ϵ. This guarantees that inequality (2.3) is satisfied. Figure 1 depicts $P_{HYP}(\mathcal{E}_2)/P_{HYP}(\mathcal{E}_1)$ as a function of $P_{USD}(\mathcal{E}_1)$ and ϵ, where $0\leqslant\epsilon\leqslant P_{USD}(\mathcal{E}_1)$ and $N=3$. Satisfaction of the inequality (2.4) occurs for parameters where $P_{HYP}(\mathcal{E}_2)/P_{HYP}(\mathcal{E}_1)>1$, and a large range of such parameters in clearly visible in the figure.

For example, let $P_{USD}(\mathcal{E}_1)=0.5$ and $P_{USD}(\mathcal{E}_2)=0.4$. This gives $\epsilon=0.1$ and inequality (2.3) is satisfied. Evaluating the maximum correct hypothesis testing probabilities, we find that $P_{HYP}(\mathcal{E}_1)=8/9\sim0.888$ and $P_{HYP}(\mathcal{E}_2)\sim0.943>P_{HYP}(\mathcal{E}_1)$ which satisfies inequality (2.4).

III. DISTINGUISHABILITY AND INFORMATION

We have seen that different distinguishability measures impose different orderings on ensembles of pure quantum states. We considered linearly independent states and demonstrated this effect using the maximum probabilities of correct hypothesis testing and unambiguous state discrimination as distiguishability measures.

This result has important implications for any situation where we wish to transmit nonorthogonal states to send classical information. For example, in quantum key distribution, we may have a choice of sending states prepared in either ensemble $\mathcal{E}_1$ or ensemble $\mathcal{E}_2$. The distinguishability of these ensembles may determine the potential information available to an eavesdropper and also the rate at which key bits can be generated. It would therefore be desirable to be able to compare the distinguishability of both ensembles. What the results in the preceding show is that this cannot be done in any absolute sense; we must know in advance which distinguishability measure is being used. However, we may not know which measure would apply to a potential eavesdropper since we may not know in advance which detection strategy they would employ.

To put the results of the preceding section in context, it is helpful to consider a related finding recently made by Jozsa and Schlienz [3]. They considered two ensembles of pure states, $\mathcal{E}_1$ and $\mathcal{E}_2$, with all states in each ensemble having equal *a priori* probabilities. They used the pairwise overlaps of pairs of states as a means of quantifying the distinguishability of each ensemble, and the von Neumann entropy S_i of the density operator of ensemble $\mathcal{E}_i$ as a means of quantifying information content. They showed that there exist en-

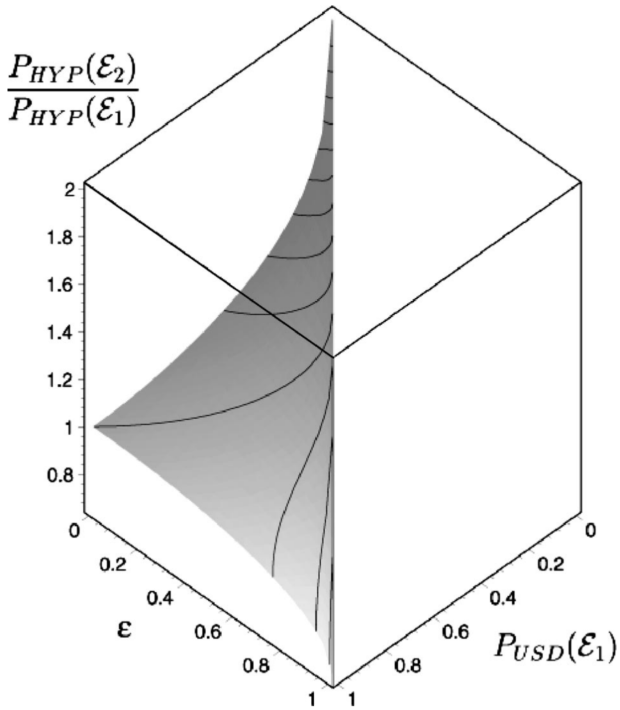

FIG. 1. Contour plot of the ratio of the maximum probabilities of correct hypothesis testing for ensembles $\mathcal{E}_1$ and $\mathcal{E}_2$ of three linearly independent, equally probable, symmetric states, versus the maximum unambiguous discrimination probability $P_{USD}(\mathcal{E}_1)$ of $\mathcal{E}_1$ and the parameter $\epsilon = P_{USD}(\mathcal{E}_1) - P_{USD}(\mathcal{E}_2)$. This parameter is positive, which implies that, except at the origin, inequality (2.3) is satisfied. Inequality (2.4) is satisfied whenever the ratio $P_{HYP}(\mathcal{E}_2)/P_{HYP}(\mathcal{E}_1) > 1$, which, on this plot, corresponds to points lying above the fourth contour from the bottom.

semble pairs in which the pairwise overlaps, and hence the distiguishability of $\mathcal{E}_2$, are less than those of $\mathcal{E}_1$, yet where $S_2 > S_1$, implying that the information content of $\mathcal{E}_2$ is greater than that of $\mathcal{E}_1$. This is somewhat counterintuitive, as information and distinguishability are often regarded as interchangeable concepts.

This finding suggests that there are distinctions to be made between information and distinguishability. That they are not entirely interchangeable is perhaps suggested by the fact that they arise in different contexts. In attempting to distinguish between a set of states, we are given only one copy, and must make optimum use of it. In information transmission, however, we perform a collective measurement on a large number of states drawn from the same ensemble. The length of the strings and the subset chosen from the set of all possible strings is such that these strings are highly distinguishable. Indeed, they are chosen in such a way that the probability of failing to distinguish between the states perfectly can be made arbitrarily small. The information content is the number of bits of classical information that can be transmitted per signal in each string, where the signals obey the statistical constraints of the ensemble. In information transmission, we select strings which are sufficiently long so that we can neglect the distinguishability issue in one way. However, it is possible that this issue does return in some form to dictate the information content of a single signal in ways that are not currently well understood.

A conclusive distinction, which does not apply to situations in which all states have the same *a priori* probabilities, is that information content and distinguishability can vary in opposite ways when the *a priori* probabilities of the states are altered. To illustrate this, consider two ensembles $\mathcal{E}_1$ and $\mathcal{E}_2$, where both consist of the pure states $|\psi_1\rangle$ and $|\psi_2\rangle$. In $\mathcal{E}_1$, both states have equal *a priori* probabilities equal to 1/2. However, in $\mathcal{E}_2$, they have unequal *a priori* probabilities p_1 and p_2. In state discrimination, our goal is to determine the state. If one state is more probable than the other, then we can take advantage of this fact to tailor our measurement to weight it in favor of the state that has higher *a priori* probability, so that, on average, we will be able to improve our ability to determine the state. It follows that $\mathcal{E}_2$ ought to have higher distinguishability than $\mathcal{E}_1$. This is made explicit if we look back at either the Helstrom bound (2.5) or the Jaeger-Shimony bound (2.6). These both increase as $\Delta = |p_1 - p_2|$ increases, confirming this expectation.

However, if we wish to use states in ensemble $\mathcal{E}_2$ to send messages, then the constraint that one state must always be more probable than the other implies that, to regain our freedom in the message we might choose to send, we must send more signal states per message as Δ increases. Quantitatively, for $\mathcal{E}_2$, the information content is given by the von Neumann entropy of the ensemble density operator, which is equal to the binary entropy function

$$S_2 = -x\log_2 x - (1-x)\log_2(1-x), \tag{3.1}$$

where

$$x = \frac{1}{2}\left[1 + \sqrt{1 - (1-\Delta^2)(1 - |\langle\psi_1|\psi_2\rangle|^2)}\right]. \tag{3.2}$$

As Δ increases, x also increases, which implies that, as can easily be shown, S_2 decreases. The extreme situation is where one of the *a priori* probabilities, say p_1, is equal to 1, in which case $p_2 = 0$. Here, we know in advance what the state is, and so we can always determine it perfectly. We may say that this ensemble has perfect distinguishability. However, the sender has no freedom in which state to send and the state is always entirely predictable. For this reason, it is impossible to transmit any information using this ensemble, for which S_2 is easily seen to be zero.

As the above argument demonstrates, a clear distinction between information content and distinguishability emerges if variation of the *a priori* probabilities of the states is permitted. However, it is not entirely clear why they should be distinct concepts when the *a priori* probabilities of all states are equal. Consequently, for ensembles with equal *a priori* probabilities, the observation that information and distinguishability impose different orderings remains to be properly understood.

The results in this paper suggest a potential explanation of this phenomenon. Given that it is not clear why information content should not be regarded as a distinguishability measure when we restrict our attention to ensembles of states with equal *a priori* probabilities, let us assume that it is in fact a suitable distinguishability measure under these circumstances. The Jozsa-Schlienz effect could then be explained as a demonstration of the fact that, even for ensembles of equally probable states, different distinguishability measures impose different ensemble orderings, which is what we have shown in this paper, using quantities that are indisputably suitable distinguishability measures.

It remains to be understood why different distinguishability measures impose different orderings. Perhaps, for two distinguishability measures to be nontrivially distinct from one another, they must be sensitive to different aspects of ensembles of states, which can vary, with some degree of independence, from one ensemble to another. This seems reminiscent of the discovery recently made by Virmani and Plenio [11] that different, good measures of the entanglement of mixed states must order these states differently.

IV. DISCUSSION

We have shown in this paper that two of the most common measures of the distinguishability of states, the maximum probabilities of correct hypothesis testing and unambiguous state discrimination, are essentially incompatible with each other. By this, we mean that they do not impose the same ordering on ensembles of pure states. It is possible to have one ensemble $\mathcal{E}_1$ which is more distinguishable than another ensemble $\mathcal{E}_2$ when the states must be distinguished unambiguously, yet where $\mathcal{E}_2$ is more distinguishable than $\mathcal{E}_1$ if we wish to identify the state by minimum error hypothesis testing. In general, we cannot, in any absolute sense, characterize one ensemble of states as being more distinguishable than another. Distinguishability comparison must necessarily refer to a particular discrimination strategy.

As with any interesting phenomenon, it is important to determine the conditions under which this effect can be demonstrated optimally. Here, we are faced with the fact that, for generic ensembles of states, the optimization problems that must be solved to obtain values of distinguishability measures are difficult to tackle analytically. At the time of writing, the only ensembles for $N>2$ for which both of our chosen distinguishability measures can be calculated analytically are ensembles of equally probable, linearly independent, symmetric pure states. For such ensembles, inequalities (2.16) and (2.17) give the extremal values of the correct hypothesis testing probability for a fixed value of the maximum probability of unambiguous state discrimination, and so the optimal conditions for demonstrating this effect for these ensembles are given by these relations. We also explored in some detail the case of $N=3$.

The effect we have demonstrated is somewhat reminiscent of a related one recently discovered by Jozsa and Schlienz. These authors showed that one can construct an ensemble pair $\mathcal{E}_1, \mathcal{E}_2$ where the distinguishability of $\mathcal{E}_1$, as measured by the pairwise overlaps of states, is greater than that of $\mathcal{E}_2$, yet where the information content of $\mathcal{E}_2$, which is quantified by the von Neumann entropy, is greater than that of $\mathcal{E}_1$.

This is an important finding, not least because the concepts of distinguishability and information content are sometimes used interchangeably. We described how essential differences between these concepts do arise when, in moving from one ensemble to another, the *a priori* probabilities of the states are changed. However, if the *a priori* probabilities of all states in both ensembles are the same, it is by no means clear that, despite the contrast between the asymptotic nature of information and the ''one shot'' nature of distinguishability, information measures cannot also serve as measures of distinguishability.

Jozsa and Schlienz demonstrated their effect using such ensembles. What the results in this paper suggest is a possible interpretation of the Jozsa-Schlienz effect. They showed that, for ensembles of equally probable pure states, information and a particular distinguishability measure suffer an ordering incompatibility problem. Does this imply that for such ensembles information content is not a suitable distinguishability measure? The results in this paper suggest that this is not necessarily the case, since different distinguishability measures suffer analogous ordering problems. It could then be the case that, for the ensembles considered by Jozsa and Schlienz, information content is in fact a suitable distinguishability measure and that the ordering incompatibility they discovered is a consequence of the fact that such problems arise with regard to generic distinguishibility measures.

ACKNOWLEDGMENTS

The author would like to thank Masahide Sasaki and Martin B. Plenio for interesting discussions. This work was supported by the U.K. Engineering and Physical Sciences Research Council.

[1] A. Chefles, Contemp. Phys. **41**, 401 (2000).
[2] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[3] R. Jozsa and J. Schlienz, Phys. Rev. A **62**, 012301 (2000).
[4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976); M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[5] L.-M. Duan and G.-C. Guo, Phys. Rev. Lett. **80**, 4999 (1998).
[6] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[7] A. Chefles and S.M. Barnett, Phys. Lett. A **250**, 223 (1998).
[8] R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, E. Riis, and M. Sasaki, Phys. Rev. A **64**, 012303 (2001).
[9] J. Mizuno, M. Fujiwara, M. Akiba, T. Kawanishi, S. M. Barnett, and M. Sasaki, Phys. Rev. A **65**, 012315 (2002).
[10] S. J. D. Phoenix, S. M. Barnett, and A. Chefles, J. Mod. Opt. **47**, 507 (2000).
[11] S. Virmani and M. B. Plenio, Phys. Lett. A **268**, 31 (2000).