Multiple-copy state discrimination: Thinking globally, acting locally

B. L. Higgins,[1, *] A. C. Doherty,[2] S. D. Bartlett,[2] G. J. Pryde,[1] and H. M. Wiseman[1, †]

[1]Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia
[2]School of Physics, The University of Sydney, Sydney, 2006, Australia

We theoretically investigate schemes to discriminate between two nonorthogonal quantum states given multiple copies. We consider a number of state discrimination schemes as applied to nonorthogonal, mixed states of a qubit. In particular, we examine the difference that local and global optimization of local measurements makes to the probability of obtaining an erroneous result, in the regime of finite numbers of copies N, and in the asymptotic limit as $N \to \infty$. Five schemes are considered: optimal collective measurements over all copies, locally optimal local measurements in a fixed single-qubit measurement basis, globally optimal fixed local measurements, locally optimal adaptive local measurements, and globally optimal adaptive local measurements. Here an adaptive measurement is one in which the measurement basis can depend on prior measurement results. For each of these measurement schemes we determine the probability of error (for finite N) and scaling of this error in the asymptotic limit. In the asymptotic limit, it is known analytically (and we verify numerically) that adaptive schemes have no advantage over the optimal fixed local scheme. Here we show moreover that, in this limit, the most naive scheme (locally optimal xed local measurements) is as good as any noncollective scheme except for states with less than 2% mixture. For finite N, however, the most sophisticated local scheme (globally optimal adaptive local measurements) is better than any other noncollective scheme, for any degree of mixture.

I. INTRODUCTION

Understanding the principles governing the process of measurement has the potential to illuminate fundamental questions and practical notions of the behaviour of the physical world. Yet despite the exceptional experimental success of quantum theory, many questions regarding the fundamental principles of quantum measurement remain. A key aspect of this is the restriction imposed by quantum measurement that a measurement of an unknown quantum system cannot usually reveal complete information about that system. The consequences of this become evident in the problem of discriminating between two nonorthogonal quantum states [1]. Consider a quantum system that was prepared in one of two known states, but we do not know which one. The task is to determine which of the two preparations took place. If we consider only pure states, we may represent the two possibilities as $|\psi_+\rangle$ and $|\psi_-\rangle$. Except for orthogonal states (where $|\langle\psi_+|\psi_-\rangle| = 0$) there is no measurement that could be applied that will deterministically find the state of the system without error [1].

Despite this, it is possible to construct a measurement that always determines the correct state, but doing so produces a nonzero probability of obtaining an inconclusive result [1–4]. Alternatively, one can construct a measurement in which all results are conclusive, but necessarily possessing some probability of error, which is to be

———

*Present address: Institute for Quantum Computing, University of Waterloo, Waterloo, ON N2L 3G1, Canada
†Electronic address: H.Wiseman@griffith.edu.au

minimized [1, 5]. This minimum-error-probability measurement is known as the Helstrom measurement, possessing an error probability C dependent upon the overlap of the states, $|\langle\psi_+|\psi_-\rangle|$.

Should additional copies of the unknown state be available [6], one can devise measurement schemes that exploit these additional copies to achieve a lowered error probability. An example is the simple "majority vote" scheme, where the Helstrom measurement is applied to each of N copies individually, and the overall result chosen according to which state corresponds to the majority of individual results. The probability of error in such a scheme scales as $C^{\mathrm{maj}} \propto |\langle\psi_+|\psi_-\rangle|^N$.

In comparison, the collective Helstrom measurement, performed at once on the collective state of all copies, $|\psi_\pm\rangle^{\otimes N}$, achieves a probability of error $C^{\mathrm{col}} = (1 - \sqrt{1 - |\langle\psi_+|\psi_-\rangle|^{2N}})/2$. While this strategy is strictly optimal, with C^{col} scaling as $|\langle\psi_+|\psi_-\rangle|^{2N}$ for $N \to \infty$, it is not in general obvious whether the Helstrom measurement on $N > 1$ copies can be performed using only local measurements, making it difficult to practically realize using current experimental techniques.

For this special case of discriminating pure nonorthogonal quantum states, $|\psi_+\rangle$ and $|\psi_-\rangle$, it is possible to achieve the same scaling of error probability as the collective Helstrom measurement by using a fixed local measurement (as in the majority vote scheme but using a different fixed measurement) applied to all N copies [7]. Furthermore, by performing local measurements on each copy adaptively, choosing the measurement bases to minimize the error probability at each stage (thus being "locally optimal" adaptive measurements), it is possible to precisely achieve the optimal performance as defined by the collective Helstrom measurement for any N [7, 8].

The above picture becomes much more complicated for mixed states. In this general case, the locally optimal adaptive local measurement scheme can exhibit worse performance than the simple majority vote scheme when the number of copies exceeds some threshold [9]. Clearly, neither scheme is the optimal local measurement scheme in the general case. To derive a local measurement scheme that is *globally* optimal, with the least probability of error for all N, requires an analysis based on dynamic programming [10]. This was done in Ref. [9], and tested experimentally, for $N \leq 10$.

Here we analyze the performance of a number of multiple-copy two-state discrimination schemes, focussing on the distinction between local optimization and global optimization as it applies to the resulting probability of error of these schemes. We consider two classes of local measurement schemes: those using identical measurement bases for each copy (fixed measurements) and those in which the measurement bases are changed as each measurement result adds new knowledge (adaptive measurements). This gives us four schemes for consideration: locally optimal fixed (LOF) local measurements, globally optimal fixed (GOF) local measurements, locally optimal adaptive (LOA) local measurements, and globally optimal adaptive (GOA) local measurements. For comparison, we also consider the optimal collective measurement (OCM). For each of these schemes we calculate the error probability for states with various levels of depolarizing mixture, for finite N. We then determine the large N scaling of the schemes by directly calculating the corresponding Chernoff bound in each case. Doing so, we verify recent theoretical bounds [11, 12] placed on the asymptotic scaling of local measurement schemes. In addition, we analyze the relationship between the LOF and GOF local measurement schemes as mixture is varied, and find that the latter has no advantage in the asymptotic limit for all qubit states with mixture greater than 2%.

II. DISCRIMINATION SCHEMES

Consider a pair of pure qubit states passing through a depolarizing channel [13]. Let $\alpha \leq \pi/2$ be the angle separating the two states in Hilbert space ($\alpha = \pi/2$ thus corresponds to orthogonal states). Without loss of generality we may represent these states as

$$\hat{\rho}_\pm = \tfrac{1}{2}[\hat{I} + (1 - \nu)(\hat{Z}\cos\alpha \pm \hat{X}\sin\alpha)] \qquad (1)$$

where $\hat{X}$ and $\hat{Z}$ are the Pauli operators, and $\nu \in [0, 1]$ quantifies the level of mixture applied by the channel. Given a collection of N identical copies of either the $\hat{\rho}_+$ or the $\hat{\rho}_-$ state (with equal probability), our task is to then determine, conclusively and with minimal probability of error, which state was given.

The four local discrimination schemes are defined by a sequence of local projective measurements applied
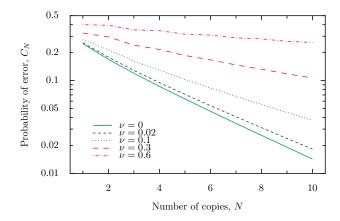

FIG. 1: Probability of error of optimal collective measurement of up to 10 copies with $\alpha = \pi/6$ for various amounts of depolarizing mixture ν, and for equally likely states ($q = 1/2$).

to each of the N copies of the unknown state. A local projective measurement performed in the basis $\{|\phi\rangle\langle\phi|, |\phi^\perp\rangle\langle\phi^\perp|\}$, where

$$|\phi\rangle \equiv \cos\phi|0\rangle + \sin\phi|1\rangle \qquad (2)$$

is parametrized by the angle $\phi \in [0, \pi/2]$. (Due to our choice of Hilbert space basis in defining our states $\hat{\rho}_\pm$, we can restrict measurements to those with real coefficients in this basis.) Each of the local measurement schemes then constitutes a sequence of measurement angles $\{\phi_n\}$, where $1 \leq n \leq N$. Given the state parameters ν and $c \equiv \cos\alpha$, Bayesian inference can be used to optimally analyze the measurement results and make a final determination of that state, with some probability of error, $C = C_N(\nu, c)$.

A. Optimal Single-Copy and Collective Measurement

For any number of copies $N \geq 1$, the minimum possible error probability C can be obtained in principle by measuring all N copies of the state using a *collective* measurement. A collective measurement is one that, in general, cannot be achieved by local measurements, even allowing for the measurement basis to be chosen adaptively (i.e. based on prior results). The optimal collective measurement for any N can be performed by measuring the observable $\hat{\Gamma} = q\hat{\rho}_+^{\otimes N} - (1 - q)\hat{\rho}_-^{\otimes N}$ [1, 5] and guessing the state as $\hat{\rho}_+$ or $\hat{\rho}_-$ corresponding to the sign of the result obtained. Here q is the prior probability of the systems being in the $\hat{\rho}_+$ state. The resulting probability of error, known as the Helstrom lower bound, is found to be $C_N^{\mathrm{OCM}} = 1 - q + \sum_{j:\gamma_j < 0} \gamma_j$, where γ_j are the eigenvalues of $\hat{\Gamma}$.

Given the conditions α and ν, we can compute the eigenvalues of the operator $\hat{\Gamma}$ and determine the probability of error. Some examples for $\alpha = \pi/6$ and various ν are shown in Fig. 1. When $\nu = 0$ (that is, for
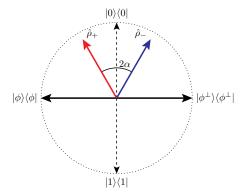
FIG. 2: Single-copy state and measurement on a Bloch sphere cross-section. The states $\hat{\rho}_+$ and $\hat{\rho}_-$, separated by a Bloch-space angle of 2α (corresponding to an angular separation of α in Hilbert space, as used in the text; here $\alpha = \pi/6$), are depicted with $\nu = 0.1$ depolarizing mixture. The measurement basis shown, here with $\phi = \pi/4$, is referred to as an 'unbiased' measurement. In the case of equal prior probability ($q = 1/2$), this is the optimal single-copy measurement.
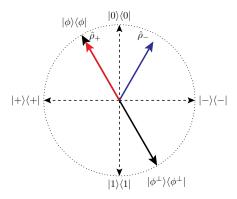

FIG. 3: Single-copy state and measurement on a Bloch sphere cross-section, with states $\hat{\rho}_+$ and $\hat{\rho}_-$ with $\nu = 0.1$ depolarizing mixture. The measurement basis shown is with $\phi = \alpha/2$, optimal in the limit $q \to 1$. This and the opposite case where $\phi = \pi/2 - \alpha/2$ (optimal for $q \to 0$) are referred to as 'fully biased' measurements.

pure states) the eigenvalues of $\hat{\Gamma}$ can be determined analytically, and the optimal collective measurement error probability C_N^{OCM} is given by

$$C_N^{\mathrm{OCM}}\big|_{\nu=0} = \tfrac{1}{2}\left(1 - \sqrt{1 - 4q(1-q)c^{2N}}\right), \quad (3)$$

where c (that is, $\cos\alpha$) in this pure state case can also be written $c = |\langle\psi_+|\psi_-\rangle|$.

The optimal single-copy measurement (OSM) is the special case $N = 1$. This is a projective local measurement, independent of ν due to the invariance of the direction of eigenvectors under depolarization. The measurement angle is given by

$$\phi^{\mathrm{OSM}}(q) = \tfrac{1}{2}\operatorname{arccot}[(2q-1)\cot\alpha]. \quad (4)$$

For the case of equal prior probabilities ($q = 1/2$) the measurement angle ϕ^{OSM} is $\pi/4$, with $|\phi\rangle\langle\phi|$ and
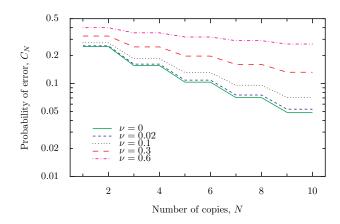

FIG. 4: Probability of error of locally optimal fixed local measurements of up to 10 copies with $\alpha = \pi/6$ for various amounts of depolarizing mixture ν.

$|\phi^\perp\rangle\langle\phi^\perp|$ symmetric about the state vectors $\hat{\rho}_\pm$ (see, e.g., Fig. 2). The probability of error is

$$C^{\mathrm{OSM}} = \tfrac{1}{2}\left[1 - (1-\nu)\sqrt{1-c^2}\right] \quad (5)$$

for all $\nu \in [0,1]$.

We note that in the case of unequal prior probabilities, $q \neq 1/2$, the optimal single-copy measurement angle ϕ^{OSM} 'biases' towards the more likely state. As q increases from $1/2$, the measurement angle decreases such that, in the limit of complete prior determination $q = 1$, $\phi^{\mathrm{OSM}} = \alpha/2$, with $|\phi\rangle\langle\phi|$ lying in the direction of $\hat{\rho}_+$ (see Fig. 3). Similarly, for $q = 0$, $\phi^{\mathrm{OSM}} = \pi/2 - \alpha/2$, placing $|\phi^\perp\rangle\langle\phi^\perp|$ in the direction of $\hat{\rho}_-$. We refer to the measurements in these limits as being 'fully biased', and the case where $\phi = \pi/4$ as being 'unbiased'.

B. Locally Optimal Fixed Local Measurements

For any locally optimal scheme, the optimization is done by minimizing the error probability of the measurement of only a single copy. The measurement angles are therefore defined by ϕ^{OSM} (dependent on q). Here we consider the case where each measurement is also performed independently, i.e. without any information obtained from the measurement of other copies. Given ν, α, and $q = 1/2$, all constants, it follows that the measurement bases of such a scheme are fixed for all copies: $\phi^{\mathrm{LOF}} = \pi/4$.

The LOF scheme reduces to a simple binomial decision problem, with probability of error

$$C_N^{\mathrm{LOF}} = \sum_{n=0}^{\lfloor N/2 \rfloor} \binom{N}{n}\left(1 - C^{\mathrm{OSM}}\right)^n \left(C^{\mathrm{OSM}}\right)^{N-n} \quad (6)$$

for N odd. Because the measurements are fixed and symmetric about the states $\hat{\rho}_\pm$, for N even there is the additional possibility of obtaining measurement results such

that Bayesian analysis equally favors both states. In this case, we may arbitrarily choose the overall discrimination result, because the equal a priori likelihood of the states ensures that any choice will lead to the same probability of error. For example, we can randomly choose the outcome. Equivalently, we may choose the state that would have been chosen if the final measurement had not been made, i.e. by considering only the first $N-1$ measurement results. In doing so, it is evident that the probability of error for the LOF scheme will not differ from any odd $N-1$ copies to an even N copies, regardless of our method for choosing the overall outcome in this situation. This results in a stepwise decrease of the error probability as the number of copies increases, examples of which can be seen in Fig. 4, where various levels of depolarizing mixture ν are applied to the state.

C. Globally Optimal Fixed Local Measurements

To achieve global optimization it is necessary to consider the error probability that arises after all N copies have been measured. One might consider a direct approach, calculating the total error probability by summing over the set of possible outcomes where the Bayesian posterior probability is satisfied (i.e. greater than one half, for each outcome), similar to Eq. (6), and optimizing over the measurement angle.

Instead of this approach, however, we present an alternative approach based on dynamic programming [10]. While more complicated than a direct summation, this approach is considerably more powerful, and will be necessary in following sections. For pedagogical reasons, we introduce it here for the easiest to understand case: a fixed scheme.

The dynamic programming approach allows us to construct a recursive analytic expression for the resulting probability of error. Suppose all N copies are measured in the basis defined by the angle ϕ, and Bayesian analysis is applied to the first $n \leq N$ measurement results. We obtain a value P_{n+1} that quantifies the credulity that the prepared state is $\hat{\rho}_+$ at this stage.

Let us define R_n as the probability of error we expect to obtain after measuring the remaining $N-n$ copies. For $n = N$, this is a simple function of the final credulity, $R_N(P_{N+1}) = \min(P_{N+1}, 1 - P_{N+1})$. For general n, the expected error probability at the $(n-1)$th measurement can be determined from the expected error probability at the nth measurement, as

$$R_{n-1}(P_n) = \sum_{D_n} \Pr[D_n|P_n, \phi] R_n(P_{n+1}), \qquad (7)$$

where D_n represents the possible outcomes of the measurement of the nth copy and $\Pr[D_n|P_n, \phi] = \Pr[D_n|\hat{\rho}_+, \phi] P_n + \Pr[D_n|\hat{\rho}_-, \phi](1 - P_n)$. In order to evaluate R_n, the credulity P_{n+1} can be calculated from the credulity P_n and measurement outcome D_n by applying
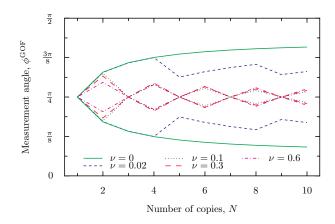

FIG. 5: Angles of globally optimal fixed local measurements ϕ^{GOF} for up to 10 copies with $\alpha = \pi/6$ for various amounts of depolarizing mixture ν.
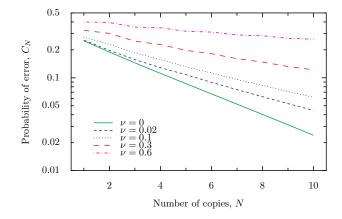

FIG. 6: Probability of error of globally optimal fixed local measurements of up to 10 copies with $\alpha = \pi/6$ for various amounts of depolarizing mixture ν.

Bayes' theorem,

$$P_{n+1} = \frac{\Pr[D_n|\hat{\rho}_+, \phi] P_n}{\Pr[D_n|P_n, \phi]}. \qquad (8)$$

Beginning with $n = N$ and progressing backwards towards $n = 0$, we can thus calculate the expected probability of error for any n by evaluating Eqs. (7) and (8) at each step using

$$\Pr[D_n|\hat{\rho}_\pm, \phi] = \frac{\nu}{2} + (1 - \nu) \cos^2\left(\phi - \frac{\pi}{4}(D_n - 1) \mp \frac{\alpha}{2}\right). \qquad (9)$$

In doing so we traverse a binary tree in which all possible permutations of measurement results are visited. The probability of error for N copies of the state is therefore given by

$$C_N = R_0(P_1), \qquad (10)$$

where $P_1 = q = 1/2$.

Armed with an expression for the probability of error for all N copies, we can determine the globally optimal
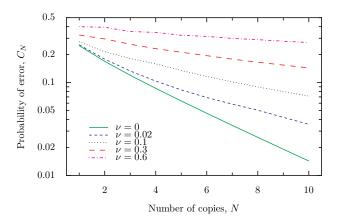
FIG. 7: Probability of error of locally optimal adaptive local measurements of up to 10 copies with $\alpha = \pi/6$ for various amounts of depolarizing mixture ν.
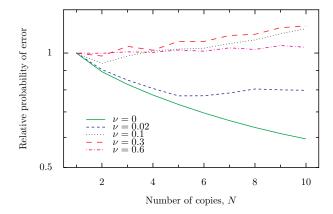

FIG. 8: Ratio of the error probability of locally optimal adaptive local measurements and the error probability of globally optimal fixed local measurements, $C_N^{\text{LOA}}/C_N^{\text{GOF}}$, for various mixtures ν. Values above 1 indicate the LOA scheme exhibiting error probabilities greater (i.e. worse) than can be obtained using the GOF scheme.

fixed measurement angle by simply minimizing over ϕ. In this case, the optimal measurement angle, ϕ^{GOF}, will depend on N and ν, as well as α. Figure 5 shows examples of this globally optimal fixed measurement angle for various amounts of mixture. We see that for $N > 1$ in the pure state case, a pair of measurement angles are equally optimal, becoming increasingly biased towards one or the other states as N increases.

For states with appreciable mixture, the measurements oscillate between being unbiased for odd N, and somewhat biased for even N. In doing so, the scheme avoids the possible condition of obtaining equal credulities for the states following analysis of all measurement results. A mixture of $\nu = 0.02$ appears to be an intermediate regime (but see Sec. III). Figure 6 shows the error probabilities for these measurements.

D. Locally Optimal Adaptive Local Measurements

For the pure state case, $\nu = 0$, it has been shown theoretically [7] and demonstrated experimentally [9] that, with measurements allowed to vary adaptively, one can achieve an error probability exactly equal to that of the collective measurements for a simple, locally optimal, adaptive scheme. That is,

$$C_N^{\text{LOA}}\big|_{\nu=0} = C_N^{\text{OCM}}\big|_{\nu=0} = \tfrac{1}{2}\left(1 - \sqrt{1 - c^{2N}}\right). \quad (11)$$

Here each copy is measured in sequence using the locally optimal single-copy measurement given the Bayesian credulity at each stage, P_n [8]. This measurement basis is defined by the angle $\phi_n^{\text{LOA}} = \phi^{\text{OSM}}(P_n)$. The measurement result is used to determine an updated credulity P_{n+1}, following Eq. (8). This is subsequently used to find the (locally optimal) angle for the next measurement, and so on in turn for each of the N copies.

For the mixed state case, $\nu > 0$, one could consider a direct approach, in a similar fashion as may be considered

for the GOF scheme, by evaluating the error probability over the entire set of measurement outcome strings. Doing so might be feasible for the moderate N considered here, but the complexity of the computation increases exponentially, making such an approach intractable for the large N considered in Sec. III.

The dynamic programming approach described in the previous section can also be used to calculate the error probability of this scheme. Here, however, an expression for the measurement angle is already well-defined, given the credulity P_n at the nth copy (which can be calculated). The overall probability of error for this scheme can therefore be calculated without requiring a numerical optimization. Examples are illustrated in Fig. 7.

The optimality of the LOA scheme (in terms of lowest overall error probability) does not hold for $\nu > 0$. For certain ν and N the LOA scheme results in a higher probability of error than a fixed scheme [9]. Figure 8 illustrates this by comparing the LOA scheme to the GOF scheme. The error probability for each scheme decreases for each newly measured copy (as each measurement provides information) however there appears a complicated relationship in the ratio of error probabilities of each scheme. The ratio is not monotonic in ν for N fixed, nor is it monotonic in N for ν fixed.

E. Globally Optimal Adaptive Local Measurements

To determine the measurements necessary to achieve the lowest error probability possible using local measurements in the general case, we must perform global optimization in a manner similar to that done for the GOF scheme. Relaxing the condition that ϕ is fixed, the measurement angle for the nth measurement becomes a function of n and the credulity P_n, i.e. $\phi_n(P_n)$. To consider the potential values of P_n between 0 and 1, we construct

a table of measurement angles $\phi_n^{\mathrm{GOA}}(P_n)$ with a large number s of linearly spaced samples of $0 \leq P_n \leq 1$. For intermediate values of P_n, the angle ϕ_n^{GOA} is estimated by cubic interpolation of nearest-neighboring samples.

Bellman's principle of optimality [10, 14] requires that the final stages of an optimal scheme, whatever the conditions are when those stages commence, must necessarily themselves be optimal under those conditions. Adhering to this principle, we construct the table of measurement angles by first recognizing that the final measurement angle $\phi_N^{\mathrm{GOA}}(P_N) = \phi^{\mathrm{OSM}}(P_N)$, the optimal single-copy angle, as there is only one copy available (the final copy) to measure at this stage. Thus, entries in the final (Nth) column of the measurements table represent the optimal single-copy measurement angles $\phi^{\mathrm{OSM}}(P_N)$ for all s samples of P_N.

From this we determine the preceding measurement angle, $\phi_{N-1}^{\mathrm{GOA}}(P_{N-1})$, by minimizing $R_{N-1}(P_{N-1}, \phi_{N-1}^{\mathrm{GOA}})$, given P_{N-1}. We do this for each P_{N-1} sample, defining the $(N-1)$th column of the measurements table representing ϕ_n^{GOA}. This then defines the optimal penultimate measurement for whatever credulity P_{N-1} may be found by that stage. The $(N-2)$th measurement can be determined similarly by minimizing $R_{N-2}(P_{N-2}, \phi_{N-2}^{\mathrm{GOA}})$ for all samples of P_{N-2}, and so on for the remaining copies. Satisfying the optimality principle at every stage of this reverse construction guarantees that the entire measurement sequence is, of all possible local measurement sequences, the globally optimal one. After constructing the final column of the measurements table, we find the probability of error $C_N^{\mathrm{GOA}} = R_0^{\mathrm{GOA}}(q)$.

It is important to note that the backwards optimizing construction of the measurements table ensures that, for any n, the final n measurements are globally optimal. It follows that the values of R_n^{GOA} obtained by minimization during this backwards process equal the forwards probabilities of error C_N^{GOA} for each $N = n$. Making use of this fact simplifies the calculation of the measurement angles and error probabilities for multiple (consecutive) values of N.

Figure 9 illustrates the globally optimal adaptive measurement angles in a mixture regime for up to ten copies, calculated using $s = 2501$ samples in the measurement and error probability tables. We found this to be sufficiently accurate as increasing the number of samples made no discernible difference to the outcome for all cases considered here. For large N, the initial measurements are approximately unbiased, that is, $\phi \approx \pi/4$, independent of the credulity P_n. A complex pattern of measurement angles becomes apparent as more copies are measured, with the final measurement being the optimal single-copy measurement by definition. That is, $\phi(P_n) = \phi^{\mathrm{OSM}}(P_n)$ of Eq. 4.

Also illustrated in Fig. 9 are six example trajectories, found by simulating the GOA scheme with measurement results generated randomly according to the outcome probabilities as a function of the state and measurement angles. The actual state is $\hat{\rho}_+$ in each case plotted. As
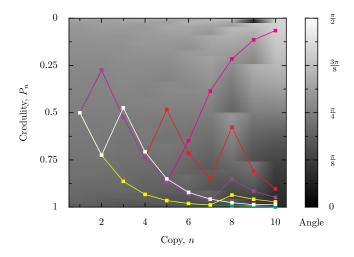

FIG. 9: Measurement angles ϕ for globally optimal adaptive local measurements of up to 10 copies with $\alpha = \pi/6$ and $\nu = 0.1$. As the scheme acquires measurement results it updates the credulity P_n, and chooses the subsequent measurement setting corresponding to the shade at the intersection of P_n and $n+1$. The lines illustrate examples of this measurement sequence.
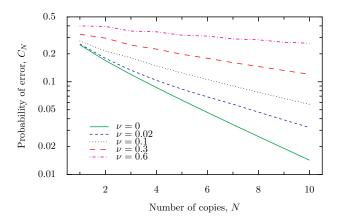

FIG. 10: Probability of error of globally optimal adaptive local measurements of up to 10 copies with $\alpha = \pi/6$ for various amounts of depolarizing mixture ν.

each copy is measured, the credulity moves in a direction corresponding to the measurement result. In some cases the measurement results are such that the direction is away from the true state. In the average case (and in the asymptotic limit) the credulity P_n increasingly corresponds to the true state.

The probability of error for the GOA scheme is plotted, for various levels of mixture, in Fig. 10. For pure states this scheme reduces to the LOA scheme, achieving the same error probability as the optimal collective measurement. For mixed states the GOA scheme does not achieve the error probability of the optimal collective measurement, but no other local measurement scheme exhibits lower probability of error.

III.   IN THE LIMIT OF LARGE N

As the number of copies grows large, i.e. $N \to \infty$, the asymptotic behaviour of the error probability is described by a function decreasing exponentially in N [15],

$$C_N \sim e^{-\xi N}. \tag{12}$$

The Chernoff bound,

$$\xi = -\lim_{N \to \infty} \frac{\log C_N}{N}, \tag{13}$$

corresponding to a given measurement scheme, is a constant quantifying the asymptotic scaling of the error probability, with higher Chernoff bound equating to faster scaling.

Some notable theoretical results already exist in this asymptotic regime. It is known that fixed local measurements are capable of achieving asymptotic scaling equal to that of adaptive local measurements [11]. This result was recently corroborated by an approach in which the error probability of state discrimination is reformed as a semidefinite programming problem, and for which strict bounds on the error probability were calculated [12] using efficient numerical algorithms [16]. Reference [12] used the same method to investigate other schemes similar to those presented here.

In the following sections we directly calculate the large N behaviour for each scheme considered in the previous sections. We consider optimal schemes under the condition of local and collective measurements given states under depolarizing mixture, and thereby provide a direct verification of the asymptotic results of Refs. [11] and [12].

A.   Optimal Collective Measurement

For equally likely states $\hat{\rho}_+$ and $\hat{\rho}_-$, the Chernoff bound for optimal collective measurements (i.e. the *quantum* Chernoff bound) is [6]

$$\xi^{\mathrm{OCM}} = -\log \min_{0 \le a \le 1} \mathrm{Tr}[\hat{\rho}_+^a \hat{\rho}_-^{1-a}]. \tag{14}$$

We can write the states in a diagonal representation, each as a mixture of a pair of orthogonal pure states $\hat{\sigma}_\pm$ and $\hat{\sigma}_\pm^\perp$, with $\hat{\rho}_\pm = (1 - \nu/2)\hat{\sigma}_\pm + (\nu/2)\hat{\sigma}_\pm^\perp$. Substituting, we find

$$\xi^{\mathrm{OCM}} = -\log \min_{0 \le a \le 1} \mathrm{Tr}\Big\{ \big[(1 - \nu/2)^a \hat{\sigma}_+ + (\nu/2)^a \hat{\sigma}_+^\perp \big]$$
$$\times \big[(1 - \nu/2)^{1-a} \hat{\sigma}_- + (\nu/2)^{1-a} \hat{\sigma}_-^\perp \big] \Big\} \tag{15}$$
$$= -\log \min_{0 \le a \le 1} \Big\{ c^2 + (1 - c^2) \big[(1 - \nu/2)^a (\nu/2)^{1-a}$$
$$+ (1 - \nu/2)^{1-a} (\nu/2)^a \big] \Big\}. \tag{16}$$

The minimum, which is unique due to convexity [6], is satisfied for $a = 1/2$. The Chernoff bound for the optimal collective measurement of mixed states $\hat{\rho}_\pm$ is therefore

$$\xi^{\mathrm{OCM}} = -\log \left[1 - (1 - c^2) \left(1 - \sqrt{1 - (1 - \nu)^2} \right) \right]. \tag{17}$$

For pure states ($\nu = 0$) this reduces to $-2\log c$, consistent with the error probability scaling $C_N^{\mathrm{OCM}} \sim c^{2N}$ found by taking Eq. (3) in the limit of large N (with equal prior probabilities).

B.   Locally Optimal Fixed Local Measurements

We can rewrite the LOF error probability Eq. (6) as

$$C_N^{\mathrm{LOF}} = \left(C^{\mathrm{OSM}} \right)^N \sum_{n=0}^{\lfloor N/2 \rfloor} \binom{N}{n} \left(\frac{1 - C^{\mathrm{OSM}}}{C^{\mathrm{OSM}}} \right)^n. \tag{18}$$

For large N this has the same scaling as the integral

$$C_N^{\mathrm{LOF}} \sim \left(C^{\mathrm{OSM}} \right)^N \int_0^{N/2} \binom{N}{n} \left(\frac{1 - C^{\mathrm{OSM}}}{C^{\mathrm{OSM}}} \right)^n dn. \tag{19}$$

When n is near $N/2$, the binomial coefficient function varies slowly (approximately as a Gaussian of variance $N/4$), but as n approaches $N/2$, $[(1 - C^{\mathrm{OSM}})/C^{\mathrm{OSM}}]^n$ grows exponentially. The scaling is therefore dominated by this term, and the error probability can be approximated as

$$C_N^{\mathrm{LOF}} \sim \left(C^{\mathrm{OSM}} \right)^N \binom{N}{N/2} \int_0^{N/2} \left(\frac{1 - C^{\mathrm{OSM}}}{C^{\mathrm{OSM}}} \right)^n dn \tag{20}$$
$$\sim \left(C^{\mathrm{OSM}} \right)^N \frac{N!}{[(N/2)!]^2} \frac{[(1 - C^{\mathrm{OSM}})/C^{\mathrm{OSM}}]^{N/2}}{\log[(1 - C^{\mathrm{OSM}})/C^{\mathrm{OSM}}]}. \tag{21}$$

To determine the LOF Chernoff bound ξ^{LOF} we take the natural logarithm of Eq. (21),

$$\log C_N^{\mathrm{LOF}} \sim (N/2) \log[(1 - C^{\mathrm{OSM}})/C^{\mathrm{OSM}}] + N \log C^{\mathrm{OSM}}$$
$$+ \log(N!) - 2\log[(N/2)!] + \mathrm{const.}, \tag{22}$$

where terms constant in N have been omitted. Using Stirling's approximation, this becomes

$$\log C_N^{\mathrm{LOF}} \sim N \log \left[2\sqrt{(1 - C^{\mathrm{OSM}})C^{\mathrm{OSM}}} \right]$$
$$- (1/2)\log N + \mathrm{const.} \tag{23}$$

Substituting into Eq. (13) and taking the limit $N \to \infty$ we obtain

$$\xi^{\mathrm{LOF}} = -\log \left[2\sqrt{(1 - C^{\mathrm{OSM}})C^{\mathrm{OSM}}} \right] \tag{24}$$
$$= -(1/2)\log \left[1 - (1 - \nu)^2 (1 - c^2) \right]. \tag{25}$$

For pure states this reduces to $-\log c$ for a probability of error scaling as $C_N^{\mathrm{LOF}} \sim c^N$, quadratically worse than $C_N^{\mathrm{OCM}} \sim c^{2N}$, consistent with the known scaling [7, 9].
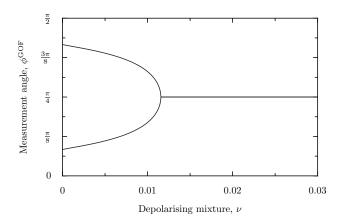
FIG. 11: Angles of globally optimal fixed local measurements ϕ^{GOF} for an unlimited number of copies with $\alpha = \pi/6$ as the amount of depolarizing mixture ν varies. For pure states there are two equally optimal measurements, fully biased such that in each case one projection is in the direction of one of the states $\hat{\rho}_\pm$. As mixture increases, the bias of the measurement decreases. At the critical point ν^{crit} the globally optimal measurement angles equal. For any amount of mixture beyond this point the globally optimal fixed measurement angle is $\pi/4$.

C. Globally Optimal Fixed Local Measurements

It has previously been shown [17] that the scaling of the GOF scheme can be quantified by considering the *classical* Chernoff bound [15] applied to fixed local measurements of each independent copy. Doing so for states $\hat{\rho}_\pm$ gives

$$\xi^{\text{GOF}} = -\log \min_{0 \le a \le 1} \min_{0 \le \phi \le \pi/2} M(a, \phi), \qquad (26)$$

where

$$M(a, \phi) = (\text{Tr}[|\phi\rangle\langle\phi|\hat{\rho}_+])^a \, (\text{Tr}[|\phi\rangle\langle\phi|\hat{\rho}_-])^{1-a}$$
$$+ \left(\text{Tr}[|\phi^\perp\rangle\langle\phi^\perp|\hat{\rho}_+]\right)^a \left(\text{Tr}[|\phi^\perp\rangle\langle\phi^\perp|\hat{\rho}_-]\right)^{1-a}. \qquad (27)$$

We numerically minimize the function $M(a, \phi)$, thus determining ϕ^{GOF}, given the parameters ν and α that define the states $\hat{\rho}_\pm$. An example is shown in Fig. 11 for a fixed α. For pure states we find that the globally optimal fixed measurement is fully biased—that is, $\phi = \alpha/2$ or $\phi^\perp = \phi - \pi/2 = -\alpha/2$, such that one of the measurement vectors lies in the direction of one of the two states. This is consistent with what we might expect to find given the results plotted in Fig. 5 for finite N. It is also consistent with the proof given in Ref. [7] that, for pure states, fully biased measurements exhibit error probability with the same scaling ξ as optimal collective measurements.

For states with sufficient mixture ν, we find the optimal fixed measurement is unbiased, equivalent to the LOF scheme. This confirms the results presented in Ref. [9] that showed unbiased measurements performing better
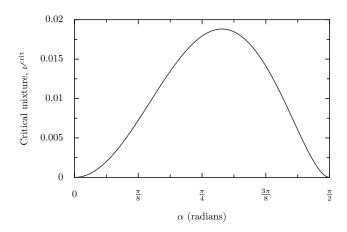

FIG. 12: The minimum amount of depolarizing mixture for which the globally optimal fixed local measurement in the asymptotic limit is $\phi^{\text{GOF}} = \pi/4$ (critical mixture, ν^{crit}) as the separation between the states α varies.

than other schemes, given enough copies of sufficiently mixed states.

As the purity of the states increases, the optimal fixed measurement angle bifurcates at a critical mixture ν^{crit} [12]. For mixture $\nu \ge \nu^{\text{crit}}$, $\phi^{\text{GOF}} = \pi/4$. For mixture $\nu < \nu^{\text{crit}}$ there exists two unique and equally optimal fixed measurement angles related by $\phi_1^{\text{GOF}} + \phi_2^{\text{GOF}} = \pi/2$. In Fig. 11, the bifurcation is found at $\nu^{\text{crit}} \approx 0.012$. As the mixture decreases, the bias of the optimal measurement increases monotonically.

We find that different separations of the states α exhibit different critical mixtures. Figure 12 shows the dependence of ν^{crit} on α, obtained by numeric minimization. The first notable fact is that there is an absolute upper bound on ν^{crit}, of about 0.0188. That is, there is an error threshold of about 2% depolarizing noise, beyond which there is no advantage asymptotically to using any local strategy beyond the simplest (locally optimal fixed) one, in which the final guess is determined by whichever result occurred in the majority of measurements.

The second notable fact is that ν^{crit} attains this upper bound at an intermediate degree of nonorthogonality of the pure states (prior to depolarization). As the two pure states approach orthogonality ($\alpha = \pi/2$) or identity ($\alpha = 0$), the value of ν^{crit} is zero. That is, any degree of mixture added to such states reduces the optimal solution to a repetition of the optimal single-qubit measurement. In the limit $\alpha \to \pi/2$, this result is not surprising, as in this limit for pure states the fully biased measurement (which is the globally optimal fixed measurement) becomes the same as the unbiased measurement, because the states become orthogonal. But in the other limit, $\alpha \to 0$, these two measurements are as different as they can be, so the discontinuous change in the optimal measurement as mixture is added is surprising, and may even be thought paradoxical. What must be borne in mind is that the asymptotic calculation makes no reference to how many copies are required
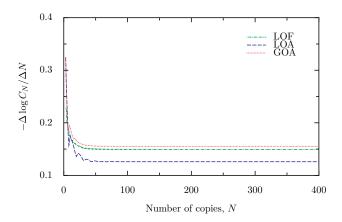
FIG. 13: Difference of nearby log error probabilities as a function of N, approaching ξ as N grows large. Values are calculated for $\alpha = \pi/6$ and $\nu = 0.02$ using the table sampling and interpolating approximation method with $s = 2501$ samples. $\Delta N = 2$ is used to avoid large variations due to the stepwise nature of the LOF error probabilities.

to approach the asymptotic regime. For almost identical states, ξ^{LOF} is very small, reflecting the fact that it is very difficult to distinguish the states. From Eq. (23) we can see that the asymptotic regime will be reached only when $N/\log N \gg 1/\xi^{\mathrm{LOF}}$. That is, as $\alpha \to 0$, the N required to reach the asymptotic regime diverges. Thus for any fixed N, no matter how large, as $\alpha \to 0$ the problem is necessarily non-asymptotic, and the globally optimal fixed measurement for pure states (the fully biased one) will remain close to the globally optimal fixed measurement for states with a small amount of depolarizing noise.

D. Locally Optimal Adaptive Local Measurements

We may find the Chernoff bound ξ for the adaptive schemes by extending the dynamic programming approach (introduced in Sec. II C) into the asymptotic regime. We take the difference of the logarithm of error probability between two nearby points in the regime of large N, thereby obtaining an estimate of

$$\left.\frac{d\log C_N}{dN}\right|_{N\to\infty} = -\xi. \tag{28}$$

For the conditions we consider, we find that this gradient reaches a constant (indicating asymptotic behaviour) before $N = 400$ copies—see, for example, Fig. 13.

For such large N, calculation using the recursive approach as detailed in Sec. II C becomes computationally infeasible. To avoid this, we instead employ the table sampling and interpolating approximation approach used to determine the error probability of the GOA scheme in Sec. II E. For the LOA scheme, the measurement angles $\phi_n^{\mathrm{LOA}}(P_n)$ are already well-defined, equal to the optimal single-copy measurement angle $\phi^{\mathrm{OSM}}(P_n)$ where, as before, P_n is the Bayesian credulity of having the $\hat{\rho}_+$ state

given the results of the previous $n-1$ measurements. We may then use this definition to calculate the error probabilities, making use of the fact that $R_n^{\mathrm{LOA}} = C_{N=n}^{\mathrm{LOA}}$. We do so for up to 400 copies.

We have found that, due to the limitations of machine precision, the interpolating approximation becomes problematic when attempting to correctly determine the error probabilities for such large N. The problem is best understood by considering the results of this approach as applied to the LOF scheme (i.e. with predefined $\phi_n = \phi^{\mathrm{LOF}}$) as compared to the results we expect due to the exact expression for the error probability. The sharp stepwise pattern of error probabilities, evident in the exact result for this scheme and which can be seen in Fig. 4, becomes less distinct and "washed out" for increasing N when calculating using the interpolating approximation. The outcome is a value for the error probability that is significantly lower than the correct value, and a value for the LOF Chernoff bound, ξ^{LOF}, significantly greater than the value as determined by our analytic derivation given in Sec. III B.

We believe that a similar effect also results in the undervaluing of LOA error probabilities and overestimation of the LOA Chernoff bound. This belief is justified as we note that decreasing the number of samples s results in an exaggeration of this effect. Conversely, increasing s reduces the effect, however the reduction appears to be exponentially decreasing in s. Indeed, for the LOF scheme, the overestimation is still quite apparent even for $s = 10001$ samples. Increasing s quickly becomes computationally infeasible.

To nevertheless extract an accurate estimate of ξ^{LOA} from these results, we calculate representative samples, ξ_s^{LOA}, for various $s \in \{501, 1001, 1501, 2001, 2501, 10001\}$ and fit them to the function

$$\xi_s^{\mathrm{LOA}} = x + \frac{y}{(\log s)^z} \tag{29}$$

using a least-squares method. Here, x and y are allowed to vary. To maintain stability of the solution, z is fixed to the value of 1.22, which is found to work best.

We finally evaluate the fit function in the limit $s \to \infty$; that is, we use $\xi_\infty^{\mathrm{LOA}} = x$. The result of this is a best-guess approximation of the Chernoff bound for the scheme. We have found that this approach works well for all but the highest-purity states, where the projection becomes unstable and the results clearly incorrect (we thus omit those results). The results are shown in Fig. 14 as a function of ν. As this shows, applying the approach described here to the LOF scheme produces ξ^{LOF} that lie only a tiny amount ≈ 0.0024 below the exact analytical results, which validates the approach.

E. Globally Optimal Adaptive Local Measurements

We find the Chernoff bound for the GOA in the same manner as we do for the LOA scheme, by extrapolating
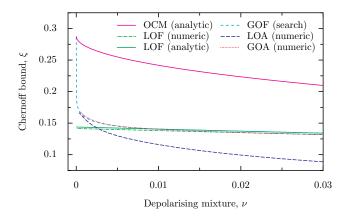
FIG. 14: Chernoff bounds for various state discrimination schemes applied to states with $\alpha = \pi/6$ and various levels of mixture. Results for OCM and LOF schemes are calculated analytically, indicated by solid lines (upper and lower, respectively). Results for the LOA and GOA schemes, as well as the LOF scheme for comparison, are calculated using the extrapolated sampling approximation method. Results for the GOF scheme are found by direct optimizing search. The results confirm the GOF and GOA schemes possessing the same scaling in the asymptotic limit.

fits of the log error probability gradient in the high-N regime. Unlike for the LOA scheme, in this case the measurements are not well-defined in advance, and must be determined by minimizing the error probability. The GOA Chernoff bound is calculated for several numbers of samples s, and an extrapolation for $s \to \infty$ gives us a final estimate of the GOA Chernoff bound, ξ^{GOA}.

The Chernoff bounds for each of the four local schemes are shown in Fig. 14 for states with $\alpha = \pi/6$. The quantum Chernoff bound of optimal collective measurements is also shown. Optimal collective measurements scale better than any other scheme in all cases except for pure states. For pure states, all but the LOF scheme have a Chernoff bound of $\xi = -2\log c$ (this is $-\log c$ for the LOF scheme). The Chernoff bound of all local measurement schemes rapidly degrades as mixture is introduced. For mixed states, the GOA and GOF schemes have equivalent scaling, optimal for local measurement

schemes. For states of high purity, the LOA scheme has better scaling than the LOF scheme, but the converse becomes true as the mixture of the states is increased. For any significant amount of mixture (here greater than approx. $\nu = 0.008$), the LOF, GOF, and GOA schemes all agree. While the numerically calculated ξ^{GOA} is a tiny bit lower than the analytically calculated ξ^{LOF} and ξ^{GOF}, it lies on top of the numerically calculated ξ^{LOF}, as expected. Thus we have confirmed that, for all but very pure states, the LOF, GOF, and GOA are all optimal local schemes in the asymptotic limit, outperforming the LOA scheme.

IV. CONCLUSION

We have investigated the probability of error exhibited by a representative set of local and collective measurement schemes for multiple-copy state discrimination acting on depolarized qubit states. We find that for any nonzero amount of mixture, local schemes fail to exhibit error probabilities approaching that of optimal collective measurements, while the globally optimal adaptive scheme introduced in Ref. [9] outperforms all other local schemes.

We have presented various ways in which we may calculate the error probabilities and measurement settings of the four local measurement schemes for finite N. In addition, we have presented analytical expressions and numerical calculations of the asymptotic behaviour (the Chernoff bound) for each scheme. These results provide a direct verification of the asymptotic scaling behaviour predicted of local measurement schemes in works such as Refs. [11] and [12]. An important conclusion of our work is that there exists a critical degree of mixture, $\nu_{\max}^{\mathrm{crit}} \approx 0.0188$, such that, in the asymptotic regime, for any pair of states with at least this much depolarization, the optimal local measurement scheme to distinguish them is the simplest: the locally optimal fixed measurement scheme.

[1] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control*, (Cambridge Univ. Press, Cambridge, 2010)
[2] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[3] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[4] A. Peres, Phys. Lett. A **128**, 19 (1988).
[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[6] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, Ll. Masanes, A. Acin, F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[7] A. Acín, E. Bagan, M. Baig, Ll. Masanes, and R. Muñoz-

Tapia, Phys. Rev. A **71**, 032338 (2005).
[8] D. Brody and B. Meister, Phys. Rev. Lett. **76**, 1 (1996).
[9] B. L. Higgins *et al.*, Phys. Rev. Lett. **103**, 220503 (2009).
[10] M. Sniedovich, *Dynamic Programming: Foundations and Principles, Second Edition* (CRC Press, New York, 2010).
[11] M. Hayashi, IEEE Trans. Inf. Theory **55**, 3807 (2009).
[12] J. Calsamiglia, J. I. de Vicente, R. Muñoz-Tapia, and E. Bagan, Phys. Rev. Lett. **105**, 080504 (2010).
[13] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).

[14] R. Bellman, *Dynamic Programming* (Princeton Univ. Press, New Jersey, 1957).

[15] H. Chernoff, Ann Math. Stat. **23**, 493 (1952).

[16] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).

[17] J. Calsamiglia, R. Muñoz-Tapia, Ll. Masanes, A. Acín, and E. Bagan, Phys. Rev. A **77**, 032311 (2008).