

## Mixed State Discrimination Using Optimal Control

B. L. Higgins,<sup>1</sup> B. M. Booth,<sup>1</sup> A. C. Doherty,<sup>2</sup> S. D. Bartlett,<sup>3</sup> H. M. Wiseman,<sup>1,\*</sup> and G. J. Pryde<sup>1,†</sup>

<sup>1</sup>Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia

<sup>2</sup>Physics Department, The University of Queensland, Brisbane, 4072, Australia

<sup>3</sup>School of Physics, The University of Sydney, Sydney, 2006, Australia

(Received 7 September 2009; published 24 November 2009)

We present theory and experiment for the task of discriminating two nonorthogonal states, given multiple copies. We implement several local measurement schemes, on both pure states and states mixed by depolarizing noise. We find that schemes which are optimal (or have optimal scaling) without noise perform worse with noise than simply repeating the optimal single-copy measurement. Applying optimal control theory, we derive the globally optimal local measurement strategy, which outperforms all other local schemes, and experimentally implement it for various levels of noise.

DOI: [10.1103/PhysRevLett.103.220503](https://doi.org/10.1103/PhysRevLett.103.220503)

PACS numbers: 03.67.-a, 03.65.Ta, 42.50.Ex

Quantum control—the application of control theory to quantum systems—offers powerful tools to enable quantum technologies to function robustly in the presence of noise and device imperfections [1–5] and to simplify protocols by reducing the need for entangling operations or collective measurements [6,7]. One such tool is adaptive measurement, wherein one adapts future measurements based on the outcomes of previous ones [1]. Quantum control based on adaptive measurements has been used to improve the measurement of an optical phase [4,8,9]. Here, we consider the problem of quantum state discrimination and demonstrate experimentally that adaptive local measurements can discriminate pure states better than nonadaptive ones. Moreover, we show that in the presence of noise, which is unavoidable in practice, the full power of optimal control theory is required to derive the globally optimal adaptive (local) measurement scheme, which we then experimentally implement.

The task of state discrimination is a fundamental primitive in many fields of quantum information science, including quantum communications, cryptography, and computing. If a quantum system is prepared in one of several possible states, this preparation can only be determined with certainty if the possible states are all mutually orthogonal. For nonorthogonal states, two complementary tasks are often considered [1]: minimizing the likelihood of either an incorrect result (an error) [10] or of an inconclusive result with no errors [11–13].

In this Letter, we consider the minimum-error discrimination of two nonorthogonal qubit states, given  $N$  identical copies of the state, using only local measurements, where the cost function  $C_N$  (which is to be minimized) is the probability of error. While continuous measurement schemes for distinguishing two infinite-dimensional pure states from a single copy have been studied elsewhere [5,14], here we consider discrete measurements of each of  $N$  discrete copies of the state. An optimal solution for multiple-copy discrimination of pure states is given by Helstrom [10] (see also [1]) and takes the form of a two-

outcome projective measurement on the joint space of all copies. For  $N > 1$ , this measurement is a nonlocal (collective) measurement on all copies, and schemes in which the same local measurement is performed on each system do not achieve this optimal performance [15]. Remarkably, it has been predicted theoretically that the optimum can be reached using *adaptive* local measurements [15]. In this adaptive scheme each system is measured locally in the basis that minimizes the probability of error immediately after that measurement. We refer to this procedure of  $N$  adaptive measurements as the “locally optimal local measurement” scheme. As shown in [15], for pure states this adaptive measurement performs just as well as the optimal collective measurement on all  $N$  copies of the state. In the asymptotic limit  $N \rightarrow \infty$ , the scaling of  $C_N$  for various state discrimination schemes has been well studied [15–17], with the notable finding that adaptive local measurements do not provide an advantage (in terms of scaling) over fixed strategies, even for mixed states [17].

Although the asymptotic performance of state discrimination schemes is of considerable academic interest, practical applications will require results for finite  $N$ , and moreover must consider the effect of noise (i.e., mixed states). Here, we adjust the local measurement strategies presented in Ref. [15] to function in the presence of noise, and analyze their performance theoretically and experimentally. Importantly, we discover that, with the exception of states that are almost pure, simple nonadaptive “unbiased measurements” (see below) outperform the locally optimal strategy defined above, for a sufficiently large number of copies. However, the globally optimal local measurement strategy, determined using optimal control theory, does outperform unbiased measurements, even though it does not achieve the optimum achievable using nonlocal measurements. For  $N$  up to 10, we theoretically predict and experimentally demonstrate the performance of each scheme with various levels of noise.

All measurements we consider are projective, in a basis  $\{|\phi\rangle, |\phi - \pi/2\rangle\}$ , where  $\phi \in [0, \pi/2)$  and

$|\phi\rangle \equiv \cos\phi|x\rangle + \sin\phi|y\rangle$ , for some orthonormal basis  $\{|x\rangle, |y\rangle\}$ . Initially, we restrict our study to the problem of distinguishing between two nonorthogonal pure states, defined without loss of generality by  $|\psi_{\pm}\rangle = \cos\theta|x\rangle \pm \sin\theta|y\rangle$ . Their overlap is  $c = \langle\psi_+|\psi_-\rangle = \cos 2\theta$ , and they are prepared with probability  $q_{\pm}$  ( $q_+ \geq q_-$ ). The single-copy Helstrom measurement is the projective measurement with  $\phi^{\text{Hel}}(q_+) = \frac{1}{2}\text{arccot}[(q_+ - q_-)\cot 2\theta]$ . From a measurement on a single copy, the most likely state given outcome  $+$  ( $-$ ) is  $|\psi_+\rangle$  ( $|\psi_-\rangle$ ), and the probability of error resulting from this best guess is  $C_1^{\text{Hel}} = (1 - \sqrt{1 - 4q_+q_-c^2})/2$ .

For multiple copies, we first build upon the three local measurement schemes presented in Ref. [15]. We treat these schemes as a prescription for what measurements to make, but unlike [15] we employ Bayesian processing of all results. This analysis allows us to determine the performance of these schemes for distinguishing mixed states. For pure states, however, such analysis is equivalent to the protocols as presented in Ref. [15].

1. *Unbiased measurements.*—Independently perform the single-copy Helstrom measurement on each copy, and decide in favor of the state with the highest posterior probability. For pure states with  $q_+ = q_-$ , this decision reduces to choosing the state with the most favorable outcomes—a “majority vote” as in Ref. [15]. When  $N$  is even there is the potential of a “split vote,” in which case a random guess is made. This scheme performs for general states and odd  $N$  as  $C_N^{\text{un}} = \sum_{m>N/2}^N \binom{N}{m} (C_1^{\text{Hel}})^m \times (1 - C_1^{\text{Hel}})^{N-m}$ , and  $C_N^{\text{un}} = C_{N-1}^{\text{un}}$  for even  $N$ . For pure states, the large  $N$  scaling is  $C_N^{\text{un}} \sim \eta c^N$ , where  $\eta$  is a constant.

2. *Fully biased measurements.*—Independently perform a projective measurement on each copy with  $\phi = \theta$ , that is, in the basis  $\{|\psi_+\rangle, |\psi_-\rangle\}$ . For pure states, the scheme can only guess the  $|\psi_+\rangle$  state if all measurement results are  $|\psi_+\rangle$ , otherwise it must guess  $|\psi_-\rangle$ ; this is the “unanimity vote” scheme of Ref. [15]. Mixed states, however, cannot reliably fulfill unanimity; in general, the best guess must be made via Bayesian analysis. For pure states, the error probability is  $C_N^{\text{fb}} = q_+ c^{2N}$ . Asymptotically, this scheme thus scales quadratically better than unbiased measurements; however, when  $N$  is sufficiently small, unbiased measurements have better performance.

3. *Locally optimal local measurements.*—Perform an optimal single-copy Helstrom measurement  $\phi^{\text{Hel}}(q_+)$  on the first copy. Via Bayes’ theorem, use the result to update the prior probability  $P_1 = q_+$  to posterior probability  $P_2$ . Using this, apply a new single-copy Helstrom measurement  $\phi^{\text{Hel}}(P_2)$  on the next copy. Repeat this adaptive process with updated probabilities  $P_n$  for all remaining copies. The best guess is the state with the higher final posterior probability. For pure states, this scheme is globally optimal for all  $N$  [15], yielding the same probability of error as the collective  $N$ -copy Helstrom measurement:  $C_N^{\text{loc}} = (1 - \sqrt{1 - 4q_+q_-c^{2N}})/2 \sim q_+q_-c^{2N}$ .

We experimentally demonstrate these schemes with  $\theta = 15^\circ$  and  $q_+ = q_- = 1/2$ , using single-photon polarization to encode the two pure states we wish to discriminate; see Fig. 1. Within the experiment, horizontal photon polarization implements the  $|x\rangle$  and vertical polarization implements the  $|y\rangle$  basis states. A half wave plate (HWP) determines the measurement basis. The measurement outcomes are entirely dependent on the relative angle between the state and the measurement axes, and not on any global orientation of the state or measurement axes. Therefore, we do not separately prepare the two states  $|\psi_+\rangle$  and  $|\psi_-\rangle$ , but rather always prepare  $|\psi_+\rangle$  and offset the measurement axes by an angle  $2\theta$  for experiments on  $|\psi_-\rangle$ . A high-contrast-ratio polarizing beam displacer and single-photon counting modules implement the orthogonal measurement outcomes. The polarization contrast ratio achievable with the apparatus was measured to be better than 0.9999 (the Bayesian processing assumes perfect visibility). The results of running each of the three algorithms in the experiment, and their theoretical predictions, are shown in Fig. 2—the experimental results correspond well with the theoretical predictions.

We now turn to the performance of these schemes in the presence of noise, i.e., for mixed states. This situation describes the addition of noise due to, for example, transmission over a noisy channel, as well as imperfect measurements. In particular, we consider uniform depolarizing noise on qubits [18] of strength  $0 \leq \nu \leq 1$ , so that the two states are now

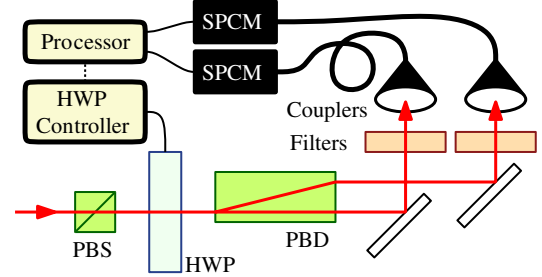


FIG. 1 (color online). Layout of the experiment. A polarizing beam splitter (PBS) acts as a filter to ensure high fidelity horizontally polarized photons. A half wave plate (HWP) in a motorized rotation stage determines the measurement basis. A polarizing beam displacer (PBD) and single-photon counting modules (SPCMs) discriminate between horizontally and vertically polarized photons with high contrast. The result of the measurement is fed to a processor which, depending on the protocol being tested, adjusts the operation of the HWP controller. Single photon inputs are obtained through type-I spontaneous parametric down-conversion—a 410 nm diode laser pumps a BiBO (bismuth borate) crystal, producing pairs of 820 nm single photons in the state  $|HH\rangle$ , with photons in separate spatial modes. One of the photon pair is guided to the input of the experiment through a single-mode optical fiber. The other photon is guided directly to a single-photon counting module. Detection in coincidence ensures high fidelity single photons are measured in the experiment.

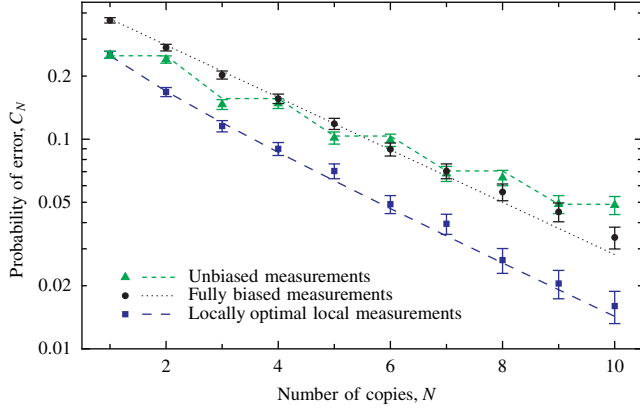


FIG. 2 (color online). Probability of error  $C_N$  in  $N$ -copy state discrimination using various schemes in the absence of noise. Lines represent theoretical predictions, points represent experimental data, each 2000 measurements. Error bars represent 1 standard deviation of the mean of a binomial distribution. The locally optimal local measurement scheme performs best for all  $N$ ; where  $N = 1$  it is equivalent to unbiased measurements, both schemes using one single-copy optimal measurement.

$$\rho_{\pm} = \frac{1}{2} [1 + (1 - \nu)(Z \cos 2\theta \pm X \sin 2\theta)]. \quad (1)$$

Here  $X = |x\rangle\langle y| + |y\rangle\langle x|$  and  $Z = |x\rangle\langle x| - |y\rangle\langle y|$  are Pauli operators. This is simulated experimentally by performing bit, phase, and bit-phase flips in the measurement basis, each with probability  $\nu/4$  [19]. Because the noise is depolarizing, the angles for the fixed measurement schemes are the same in the mixed state case as in the pure case. Even so, noise will clearly have a detrimental effect on the performance of the schemes described above. Indeed, it is now the case that no local scheme can achieve the globally optimal performance achievable with a collective measurement.

We have calculated the respective error probabilities  $C_N$  exactly as a function of noise; see Figs. 3 and 4. Both the fully biased measurement scheme and the locally optimal adaptive scheme lose their superiority over unbiased measurements as  $\nu$  is increased. Our theoretical analysis confirms this behavior for general  $\theta$  and  $q_+$ , with the value of  $\nu$  at which the error probability curves cross depending on  $\theta$ ,  $q_+$ , and  $N$ .

The locally optimal scheme maximizes the discriminating power of each measurement individually, but that is not the same as maximizing the discriminating power of all  $N$  measurements together (even when restricting to local, not collective, measurements). Because the locally optimal local measurement scheme is evidently not the globally optimal local measurement scheme *in general*, we now turn to finding such a scheme.

**4. Globally optimal local measurements.**—To determine the optimal discrimination scheme using local adaptive measurements, we use dynamic programming [20]. This will in general yield an adaptive scheme that depends explicitly on the total number of measurements  $N$  that

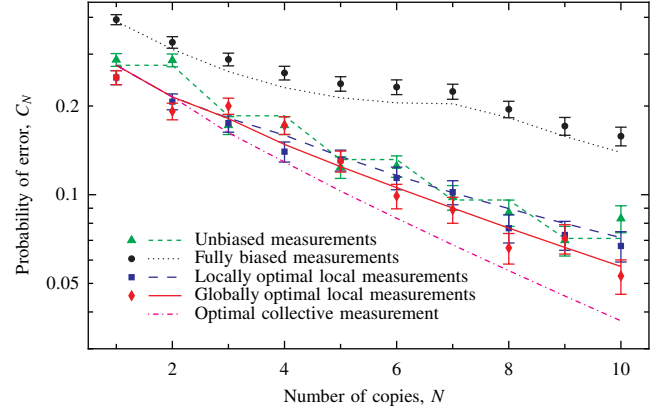


FIG. 3 (color online). Error probability  $C_N$  of discrimination schemes under  $\nu = 10\%$  depolarizing noise. Points represent 1000 experimental discriminations. The addition of noise detrimentally impacts the locally optimal local measurement scheme more than the unbiased scheme. Indeed, theory predicts that the latter outperforms the former for  $N = 5$ ,  $N = 7$ , and  $N \geq 9$ . The globally optimal local measurement scheme performs better than all other local measurement schemes in the presence of noise. The theoretical optimal collective measurement cost is plotted for comparison.

will be performed, unlike the locally optimal scheme. The scheme is defined by a table of measurement angles, with rows corresponding to  $n$ , the copy to be measured ( $1 \leq n \leq N$ ), and columns corresponding to  $P_n$ , the probability prior to the  $n$ th measurement that the prepared state is  $|\psi_+\rangle$ , conditioned on the measurement results of the first  $n - 1$  copies ( $P_1 = q_+$ ). Thus, at the  $n$ th step, we consult the table to obtain the measurement angle  $\phi_n(P_n)$  to be used. The result of this measurement is then used to calculate a posterior  $P_{n+1}$  via Bayes' theorem, and we proceed to the next step. Linear interpolation resolves the

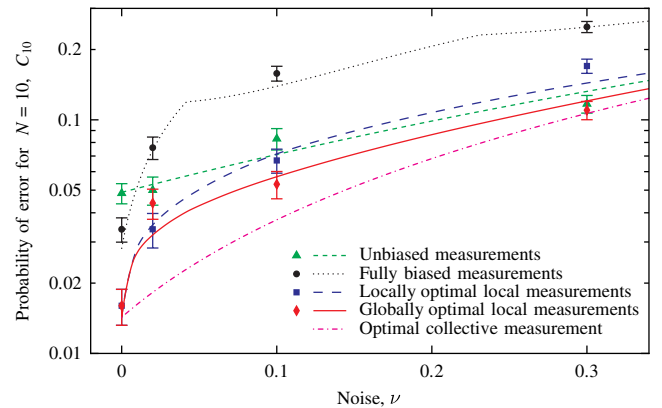


FIG. 4 (color online). Error probability  $C_N$  of discrimination schemes under various levels of noise  $\nu$  for  $N = 10$  measured copies. Points each represent 2000 ( $\nu = 0$ ) or 1000 ( $\nu > 0$ ) experimental discriminations. Here, unbiased measurements outperform the locally optimal local measurement scheme for noise  $\nu \geq 10\%$ .



discreteness in the table's representation of  $P_n$  (here we use 2501 samples).

We construct this table as follows. In all cases, the optimal measurement on the final copy  $n = N$  must be the single-copy Helstrom measurement,  $\phi_N(P_N) = \phi^{\text{Hel}}(P_N)$ , as this measurement will minimize the error probability  $C_N$  regardless of the previous measurement choices. Starting from this fact, the globally optimal local measurement scheme for  $N$  copies is constructed in reverse. Using the recursive relationship between the expected error probabilities after  $n$  and  $n + 1$  measurements (see [19]), the penultimate measurement  $\phi_{N-1}(P_{N-1})$  that minimizes  $C_N$  can be found by a numerical search, given  $P_{N-1}$ . When calculated for samples of the range  $0 \leq P_{N-1} \leq 1$ , this defines row  $N - 1$  of the measurement table.

The optimal measurement that precedes the final two measurements can similarly be obtained by minimizing the expected error probability over the measurement  $\phi_{N-2}(P_{N-2})$  for some  $P_{N-2}$ . This constructs row  $N - 2$  of the measurement table. Continuing this analysis, we construct a table of  $N$  measurement settings, defining  $\phi_n^{\text{glo}}(P_n)$ , which results in the lowest final error probability,  $C_N^{\text{glo}}$ .

To determine the performance of nonglobally optimal measurements when noise is present one can use the same procedure, but with a nonoptimal measurement choice. For example, the locally optimal local measurement  $\phi_n^{\text{loc}}(P_n)$  defines  $C_N^{\text{loc}}$ . As the measurements are Markovian, sampling is unnecessary, and for moderate  $N$  as used in this Letter, the probability of error can be calculated exactly.

The globally optimal local measurement scheme constructed according to the above procedure reduces to the locally optimal scheme in the noiseless case. For high noise, we have found numerically that the measurement setting  $\phi^{\text{glo}}$  for all but the final few copies approaches  $\pi/4$ , as for unbiased measurements when  $q_+ = q_-$ . Its performance also approaches that of unbiased measurements, in this regime where  $\nu$  is not small. Importantly, for all  $\nu > 0$ , we have  $C_N^{\text{glo}} < \min(C_N^{\text{loc}}, C_N^{\text{un}})$  for  $N \geq 3$ , as expected. But we also have  $C_N^{\text{glo}} > C_N^{\text{col}}$ , the probability of error from a collective measurement over all copies, achieved by the  $N$ -copy Helstrom measurement [1,10]. This is illustrated in Figs. 3 and 4.

We experimentally investigate all four local measurement schemes in the presence of 2%, 10%, 30%, and 60% noise. The results for  $\nu = 10\%$  noise for  $N$  up to 10 are shown in Fig. 3, and for fixed  $N = 10$  under various noise in Fig. 4. Further results may be found in the supplementary material [19]. Theoretical curves are determined numerically using the dynamic programming method described above. The discontinuities in the gradient of  $C_N^{\text{fb}}$  arise due to the discreteness of the number of outcomes required to guess  $|\psi_-\rangle$ . In all cases, experimental data agree with theoretical predictions, within expected statistical variation. The globally optimal scheme has the best performance for all levels of noise and for all  $N$ .

We have shown that local adaptive  $N$ -copy discrimination schemes which are optimal in the noiseless regime are significantly impacted by the addition of noise. The locally optimal local measurement scheme, in particular, performs more poorly than nonadaptive unbiased measurements. Subsequently, by a dynamic programming analysis, we have demonstrated the adaptive local measurement scheme that is globally optimal, having, in all cases, the lowest probability of an incorrect discrimination of any local measurement scheme for any  $N$ . In addition to illuminating part of the fundamentally interesting problem of quantum state discrimination, our work provides an insight into the fragility of idealized models practically applied and demonstrates the usefulness of optimal quantum control techniques in mitigating the real-world issues that face the application of quantum technologies.

We thank Ramon Muñoz-Tapia for helpful discussions. This work was supported by the Australian Research Council.

\*H.Wiseman@griffith.edu.au

†G.Pryde@griffith.edu.au

- [1] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, Cambridge, England, 2009).
- [2] J. M. Geremia, J. K. Stockton, A. C. Doherty, and H. Mabuchi, *Phys. Rev. Lett.* **91**, 250801 (2003).
- [3] A. M. Brańczyk *et al.*, *Phys. Rev. A* **75**, 012329 (2007).
- [4] M. A. Armen *et al.*, *Phys. Rev. Lett.* **89**, 133602 (2002).
- [5] R. L. Cook, P. J. Martin, and J. M. Geremia, *Nature (London)* **446**, 774 (2007).
- [6] R. B. Griffiths and C. S. Niu, *Phys. Rev. Lett.* **76**, 3228 (1996).
- [7] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, *Phys. Rev. Lett.* **85**, 4972 (2000).
- [8] B. L. Higgins *et al.*, *Nature (London)* **450**, 393 (2007).
- [9] B. L. Higgins *et al.*, *New J. Phys.* **11**, 073023 (2009).
- [10] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [11] I. D. Ivanovic, *Phys. Lett. A* **123**, 257 (1987).
- [12] D. Dieks, *Phys. Lett. A* **126**, 303 (1988).
- [13] A. Peres, *Phys. Lett. A* **128**, 19 (1988).
- [14] C. Wittmann *et al.*, *Phys. Rev. Lett.* **101**, 210501 (2008).
- [15] A. Acín *et al.*, *Phys. Rev. A* **71**, 032338 (2005).
- [16] J. Calsamiglia *et al.*, *Phys. Rev. A* **77**, 032311 (2008).
- [17] M. Hayashi, arXiv:0804.0686v1.
- [18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
- [19] See EPAPS Document No. E-PRLTAO-103-040949 for more information about the experimental simulation of depolarizing noise, theoretical details of the globally optimal local measurement scheme derivation, and plots following Fig. 3 of error probability of discrimination schemes under 2%, 30%, and 60% noise. For more information on EPAPS, see <http://www.aip.org/pubservs/epaps.html>.
- [20] G. L. Nemhauser, *Introduction to Dynamic Programming* (John Wiley and Sons, New York, 1966).