Eur. Phys. J. D (2012)
DOI: 10.1140/epjd/e2012-20618-3

THE EUROPEAN PHYSICAL JOURNAL D

Regular Article

# Programmed discrimination of multiple sets of qbits with added classical information

A.J.T. Colin[a]

Department of Physics, Scottish Universities Physics Alliance, Strathclyde University, John Anderson Building, 107 Rottenrow, G4 0NG Glasgow, UK

Received 25 October 2011 / Received in final form 31 January 2012
Published online (Inserted Later) – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2012

**Abstract.** This paper offers some new results in the area of programmed quantum discrimination, which determines whether a data qbit is the same as the first or second of a given pair of program qbits. In a recent paper [A.J.T. Colin, S.M. Barnett, J. Jeffers, Eur. Phys. J. D **63**, 463 (2011)] we examined the effect of additional classical information on this process. We now extend our work to consider the effects of replicating some or all of the qbits in the group. As the parameters to the calculations can take many values in several dimensions, we offer only a sample of results. A suite of programs [A.J.T. Colin, Multiple qbit discrimination programs, `ftp://www.singing-baboon.com/discrimination.zip`] allows the reader to explore the problem space in detail. Turning to a possible application of the method, we investigate the efficacy of programmed discrimination in the context of data communication. The technique offers a key advantage over other methods of quantum-based data transmission; namely, that it is insensitive to any unitary transformation that may occur in transit, provided only that the same transformation applies equally to all the qbits in the group. However, the technique is costly in its use of resources. Using the best configuration we could find, with unambiguous discrimination, orthogonal program qbits, and a duplicated data qbit, the transmission of one binary digit reliably still needs 8 qbits.

## 1 Introduction

Much work has been published on a technique called *programmed discrimination* [1,3–15]. Here each bit of data is represented by a *triad* which consists of three qbits:

- a first program qbit $|\psi_1\rangle$;
- a data qbit $|\psi_2\rangle$;
- a second program qbit $|\psi_3\rangle$.

The program qbits are pure states that correspond to different points of the surface of the Bloch sphere. The data qbit is guaranteed to be identical to one or other of the program qbits.

The discrimination task consists of deciding which of the two possibilities, $|\psi_2\rangle \equiv |\psi_1\rangle$ or $|\psi_2\rangle \equiv |\psi_3\rangle$ is true. In this context, as in others, there are two main methods of discrimination, neither of which is guaranteed always to give the correct result.

The methods are:

- Optimal or minimum error, where the discrimination process always returns a result, but there is a finite probability, which we aim to make as small as possible, that this result is wrong. This probability is called the *error rate* [3,16].
- Unambiguous, where the discrimination *may* give a result guaranteed to be correct, but there will always be a certain probability, known as the *failure rate*, that the process will return the answer 'don't know' [3–5,16].

Both types of process disturb the qbits in the triad, so that any further measurement will not return a meaningful result.

As all three qbits in a triad are unknown, the success rate of any discrimination process is quoted as an average of all possible configurations.

Sometimes the problem is not symmetrical between the two program qbits. It may be known in advance that the data qbit is more likely to be identical to one of the program qbits than the other. We can write expressions for the probabilities $\eta_1$ and $\eta_2$:

$$\eta_1 = P(|\psi_2\rangle \equiv |\psi_1\rangle) \quad (1)$$

$$\eta_2 = P(|\psi_2\rangle \equiv |\psi_3\rangle) \quad (2)$$

$$\eta_1 + \eta_2 = 1. \quad (3)$$

This topic has been extensively analysed by several authors. In a key paper Bergou et al. [3] have published analytic expressions that give the expected error rates as functions of $\eta_1$ for both the optimal and unambiguous discrimination methods.

The states of the two program qbits have mostly been taken as random, with no correlation between them.

[a] e-mail: `andrew@crm.scotnet.co.uk`

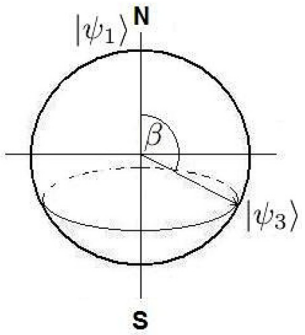


**Fig. 1.** Showing a fixed overlap of $\cos^2(\beta/2)$.

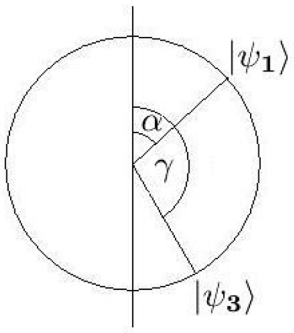


**Fig. 2.** Program qbits on the same great circle.

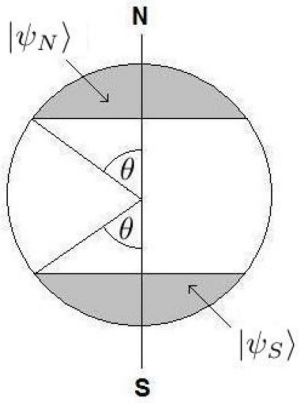


**Fig. 3.** Program qbits confined to polar caps.

However, a recent paper by Colin et al. [1] investigates discrimination rates for triads where the program qbits have certain known relationships. In particular we considered three specific configurations:

(1) The two program qbits have a *known* overlap. In this paper we represent this overlap by an angle $\beta$, such that $\langle\psi_1|\psi_3\rangle = \cos^2(\beta/2)$. This is illustrated in Figure 1.
(2) The program qbits are both located on the same *known* great circle. See Figure 2.
(3) The program qbits are confined to caps, of *known* equal size, centered on opposite poles of the Bloch sphere. The size of each cap is determined by an angle $\theta$, as shown in Figure 3.

The present paper extends this research by considering configurations where the data qbit, and either or both of the program qbits, are supplied in multiple copies. This possibility has also been analysed by other authors [3,7–9] but in all these cases the states of the program qbits have been taken to be randomly distributed, an approach that lends itself to formal mathematical analysis. In contrast to other papers in the area, our approach is necessarily numerical, as it involves finding the eigenvalues and eigenvectors of large matrices derived from a range of parameters. We obtain practical results for many configurations which have so far resisted analytical treatment. For certain particular values of these parameters, analytical results have already been derived [3], and here our computed results agree with the published expressions. In particular, two qbits at random locations on the Bloch sphere have an average overlap of $(\frac{1}{\sqrt{2}})$, and our computation with a (fixed) overlap of $(\frac{1}{\sqrt{2}})$ gives, as expected, precisely the same results.

A final section of the paper assesses the technique of programmed discrimination as a method of data communication. We show that imposing a known relationship between the program qbits can markedly increase the efficiency of communication.

A key aspect of any scientific paper is that the results should be independently verifiable and repeatable. To this end we describe our algorithms in some detail.

## 2 The problem

Our basic aim is to extend the analysis in [1] to configurations where each of the qbits in the triad may be replicated two or more times. We note that analytic solutions to this problem have already been published [3,9,13], but only for symmetrical configurations where the program qbits are uncorrelated.

It is convenient to refer to a configuration which has $x$ copies of one program qbit, $y$ copies of the data qbit, and $z$ copies of the other program qbit as $\{x, y, z\}$ (for example, {1, 3, 2}).

In previous papers the state of a system might have been described as:

$$|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle. \quad (4)$$

From this point we drop the position indicators $_1$, $_2$ and $_3$ because these qbits may be repeated. The notation $|\psi\rangle^{\otimes n}$ signifies the $n$-fold tensor product $\underbrace{|\psi\rangle\otimes|\psi\rangle\otimes\ldots\otimes|\psi\rangle}_{\longleftarrow n\longrightarrow}$.

This problem can be solved in a parameter space that spans many dimensions. Two primary dimensions are:

(a) the type of relationship between the program qbits (fixed overlap, same given great circle, and confined to caps);
(b) the type of discrimination (optimal or unambiguous)

Continuous dimensions include:

(c) the a priori probabilities of identity $\eta_1$ and $\eta_2$;
(d) the degree of overlap (for a known overlap);
(e) the size of the caps (for confined qbits).

Within all these dimensions we also consider different patterns of replication.

The analysis of any system with a total of $n$ qbits involves matrices of dimension $2^n \times 2^n$. Increasing values of $n$ result in an exponential growth in the size of the problem. For $n > 3$ it is no longer feasible to carry out the analysis by hand except in special cases.

Even with a computer, the exponential nature of the problem sets a practical limit to the size of the configuration that can be analysed. For example, the analysis of $\{1, 2, 4\}$ (7 qbits) is feasible; $\{5, 12, 8\}$ (25 qbits) is not.

The parameter space of the problem is massive; it would be impractical to explore it in full and present the results in this paper. Instead, we display just a few sample results, and give the URL of a website that holds our programs [2]. Readers can use this software to explore other parts of the space in which they have an interest.

## 3 Overview

This section presents a summary of the methods used to establish our results. They are similar to those described in [1], but have been adapted for computer execution. Details of the calculations can be found in the appendix.

The data qbit is always the same as one or other of the program qbits. It follows that for any configuration $\{x, y, z\}$, there are only two possible states of the system:

$$|\Psi_1\rangle = |\psi_1\rangle^{\otimes(x+y)}|\psi_3\rangle^{\otimes z} \tag{5}$$

$$|\Psi_2\rangle = |\psi_1\rangle^{\otimes x}|\psi_3\rangle^{\otimes(y+z)}. \tag{6}$$

The analysis of each configuration takes two main stages:

- The *first stage* sets up density matrices for each of the two possible states, both averaged over those parts of the Bloch sphere in which the qbits might be found:

$$\overline{\hat{\rho}}_1 = \overline{|\Psi_1\rangle\langle\Psi_1|} \tag{7}$$

$$\overline{\hat{\rho}}_2 = \overline{|\Psi_2\rangle\langle\Psi_2|}. \tag{8}$$

- The *second stage* uses these two density matrices to find statistical results.

Helstrom's rule [17] allows the optimal discrimination error rate to be calculated directly from the two density matrices. We substitute various values of $\eta_1$ and $\eta_2$ and take the following steps:

(1) compute a difference operator $\hat{\Lambda} = (\eta_1\overline{\hat{\rho}}_1 - \eta_2\overline{\hat{\rho}}_2)$;
(2) find $s$, the sum of the magnitudes of the eigenvalues of $\hat{\Lambda}$;
(3) calculate the error rate as $(1 - s)/2$.

Finding the unambiguous discrimination failure rate also requires several steps and is an order of magnitude more laborious. Following the procedure described in [5], we establish a POM with three component operators:

- $\hat{\Pi}_1$, which unambiguously recognises state $|\Psi_1\rangle$;
- $\hat{\Pi}_2$, which unambiguously recognises state $|\Psi_2\rangle$;
- $\hat{\Pi}_0$, which returns the result unknown.

Since one of these outcomes is certain to occur, we may write

$$\hat{\Pi}_0 + \hat{\Pi}_1 + \hat{\Pi}_2 = \mathbf{I} \tag{9}$$

$$\text{or } \hat{\Pi}_0 = \mathbf{I} - \hat{\Pi}_1 - \hat{\Pi}_2$$

where $\mathbf{I}$ is the unit matrix.

The steps in the computation are:

(1) Using Jacobi's algorithm, find the eigenvalues and eigenvectors of the two density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$. Discard the eigenvectors with zero eigenvalues, and form the others into two non-square matrices $Z_1$ and $Z_2$.
(2) Construct a projector $\hat{P}_1$ from the vectors in $Z_1$ and $Z_2$. Assume that each vector in $Z_1$ is a linear combination of all the vectors in $Z_2$, plus a residual component that is orthogonal to all the vectors in $Z_2$. Build $\hat{P}_1$ from these residual vectors, discarding the rest. Construct $\hat{P}_2$ in the same way.
Confirm that:

$$\mathrm{Tr}(\overline{\hat{\rho}}_1\hat{P}_2) = \mathrm{Tr}(\overline{\hat{\rho}}_2\hat{P}_1) = 0. \tag{10}$$

The algorithm for this procedure is more fully described in the appendix.
(3) At first sight it would seem that $\hat{P}_1$ could be used as it stands as $\hat{\Pi}_1$, the detector for $|\Psi_1\rangle$, since it cannot possibly be triggered by state $|\Psi_2\rangle$, and *vice versa*. Matters are not so simple, as $\hat{\Pi}_0$ must be positive-definite; that is, its smallest eigenvalue may not be negative. The best unambiguous recognition rate is attained when the smallest eigenvalue of $\hat{\Pi}_0$ is exactly zero. We must set:

$$\hat{\Pi}_1 = c_1\hat{P}_1 \tag{11}$$

$$\hat{\Pi}_2 = c_2\hat{P}_2 \tag{12}$$

where $c_1$ and $c_2$ are positive constants less than 1. Clearly,

$$\hat{\Pi}_0 = \mathbf{I} - c_1\hat{P}_1 - c_2\hat{P}_2. \tag{13}$$

The constants $c_1$ and $c_2$ are not independent, for any given value of one determines the other, but the relationship between them in most cases is not easy to determine analytically. It must be found by an iterative numerical method.
(4) Given that the state is $|\Psi_j\rangle$, $(j = \{1, 2\})$ the probability of detecting is it $\mathrm{Tr}(\hat{\rho}_j\hat{\Pi}_j)$. It follows that the overall probability of successful unambiguous discrimination $Q$ is:

$$\begin{aligned} Q &= \eta_1\mathrm{Tr}(\hat{\rho}_1\hat{\Pi}_1) + \eta_2\mathrm{Tr}(\hat{\rho}_2\hat{\Pi}_2) \\ &= \eta_1\mathrm{Tr}(c_1\hat{\rho}_1\hat{P}_1) + \eta_2\mathrm{Tr}(c_2\hat{\rho}_2\hat{P}_2). \end{aligned} \tag{14}$$

For any given value of $\eta_1$ in the central range this expression can be maximised numerically, using one of $c_1$ or $c_2$ as an independent variable.

**Table 1.** Fixed overlap optimal discrimination {1, 1, 1}.

| $\beta$ | $\pi/6$ | $\pi/3$ | $\pi/2$ | $2\pi/3$ | $5\pi/6$ | $\pi$ |
|---|---|---|---|---|---|---|
| $\eta_1$ | | | | | | |
| 0.25 | 0.241 | 0.217 | 0.183 | 0.150 | 0.125 | 0.116 |
| 0.50 | 0.481 | 0.428 | 0.356 | 0.283 | 0.231 | 0.211 |
| 0.75 | 0.241 | 0.217 | 0.183 | 0.150 | 0.125 | 0.116 |

**Table 2.** Fixed overlap optimal discrimination {1, 6, 1}.

| $\beta$ | $\pi/6$ | $\pi/3$ | $\pi/2$ | $2\pi/3$ | $5\pi/6$ | $\pi$ |
|---|---|---|---|---|---|---|
| $\eta_1$ | | | | | | |
| 0.25 | 0.236 | 0.196 | 0.142 | 0.088 | 0.049 | 0.035 |
| 0.50 | 0.471 | 0.392 | 0.283 | 0.175 | 0.096 | 0.067 |
| 0.75 | 0.236 | 0.196 | 0.142 | 0.088 | 0.049 | 0.035 |

When $\eta_1$ or $\eta_2$ are close to zero, the equation has no solution, and it is sometimes best to discard the POM in favour of one or other of the projectors $\hat{P}_1$ or $\hat{P}_2$. The range over which the full POM is valid depends on the configuration being analysed.

The proportion of measurements, $(1 - Q)$, which return the result 'unknown' is the *failure rate*.

## 4 Results

The results in this section are generated by a suite of six programs that cover the three types of relationship (fixed overlap, known great circle, and polar caps) and both modes of discrimination (optimal and unambiguous) as applied to each relationship. We present the results of the fixed overlap programs in some detail, but give only brief summaries for the other two types.

### 4.1 Fixed overlap with optimal discrimination

First we examine multiple copies of the data qbit. Some results are given in Table 1 for the configuration {1, 1, 1} and in Table 2 for {1, 6, 1}. In these tables the vertical axis covers different values of $\eta_1$. The horizontal axis gives different values of $\beta$, where the overlap is $\cos^2(\beta/2)$. Figures in the body of the table represent error rates. Results for intermediate cases are not shown but can easily be calculated using the programs in [2]. In all cases the user can adjust the intervals between the values of $\beta$ and $\eta_1$.

Figure 4 is based on the configuration $\{1, n, 1\}(1 \leq n \leq 6)$. and shows how the error rate varies with the number of copies of the data qbit, for various degrees of overlap. The error rate generally falls when the number of copies is increased, but this effect is most marked when the overlap is small, and barely visible when the overlap is large.

Next we investigate the case where there are multiple copies of the program qbits, arranged symmetrically. We compare configurations of the forms $\{1, 2n - 1, 1\}$ with those of $\{n, 1, n\}$, which have the same number of qbits. Results show that the error rates are very similar.

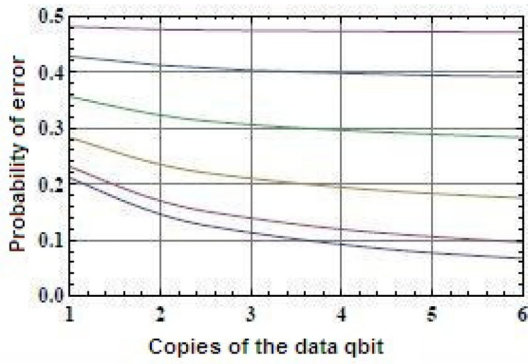


**Fig. 4.** Configuration $\{1, n, 1\}$. Optimal error for fixed overlap given by $\beta = \pi/6$ (top curve) to $\beta = \pi$ (bottom curve) in steps of $\pi/6$.

**Table 3.** Fixed overlap optimal discrimination for symmetrical duplication of program qbits, compared to repetition of the data qbit.

| Configuration | qbits | Error rate ($\beta = \pi/2$) | Error rate ($\beta = \pi$) |
|---|---|---|---|
| {1, 1, 1} | 3 | 0.356 | 0211 |
| {1, 3, 1} | 5 | 0.306 | 0.113 |
| {2, 1, 2} | | 0.308 | 0.115 |
| {1, 5, 1} | 7 | 0.289 | 0.077 |
| {3, 1, 3} | | 0.288 | 0.076 |
| {1, 7, 1} | 9 | 0.280 | 0.059 |
| {4, 1, 4} | | 0.279 | 0.057 |

**Table 4.** Fixed overlap optimal discrimination with asymmetrical replication of program qbits.

| Confiuration | qbits | Error rate ($\beta = \pi/2$) | Error rate ($\beta = \pi$) |
|---|---|---|---|
| {1, 1, 1} | 3 | 0.356 | 0211 |
| {1, 2, 1} | 4 | 0.323 | 0.146 |
| {1, 1, 2} | | 0.319 | 0.138 |
| {1, 3, 1} | 5 | 0.306 | 0.113 |
| {1, 2, 2} | | 0.287 | 0.073 |
| {1, 1, 3} | | 0.302 | 0.105 |
| {1, 4, 1} | 6 | 0.296 | 0.092 |
| {1, 3, 2} | | 0.272 | 0.045 |
| {2, 1, 3} | | 0.294 | 0.089 |
| {1, 2, 3} | | 0.272 | 0.045 |

Table 3, which is based on $\eta_1 = 0.5$ and includes columns for $\beta = \pi/2$ and $\beta = \pi$ makes this clear.

To end this section, we consider various non-symmetrical configurations. For $\beta = \pi/2$ the error rates for al l configurations with the same number of qbits are close to one another, but for $\beta = \pi$ the relationship breaks down. Table 4 gives the figures.

The interesting conclusion is that in this mode (fixed overlap, optimum discrimination) it makes little difference how the qbits are used provided the arrangement is symmetrical: any symmetric configuration with $n$ qbits in total will result in substantially the same probable error rate.

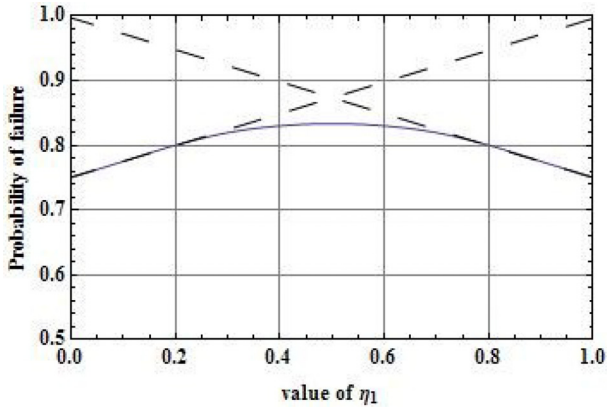


**Fig. 5.** Fixed overlap = $\frac{1}{\sqrt{2}}$, failure rate in unambiguous discrimination for $\{1, 1, 1\}$.

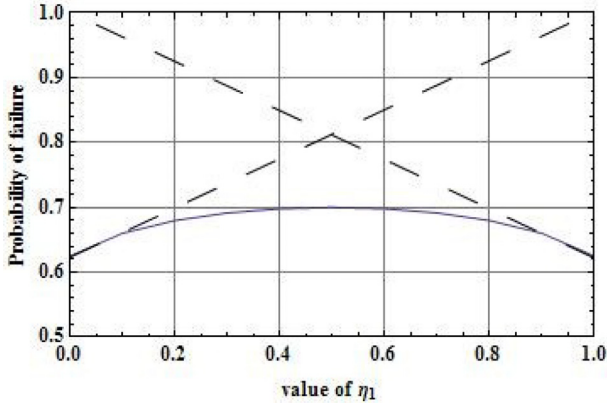


**Fig. 6.** Fixed overlap = $\frac{1}{\sqrt{2}}$, failure rate in unambiguous discrimination for $\{1, 3, 1\}$.

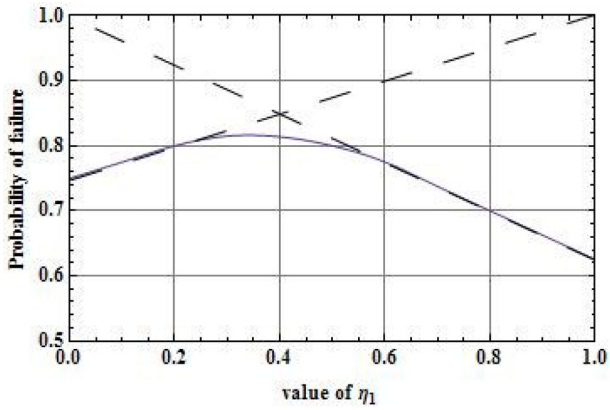


**Fig. 7.** Fixed overlap = $\frac{1}{\sqrt{2}}$, probable error rate in unambiguous discrimination for $\{1, 1, 3\}$.

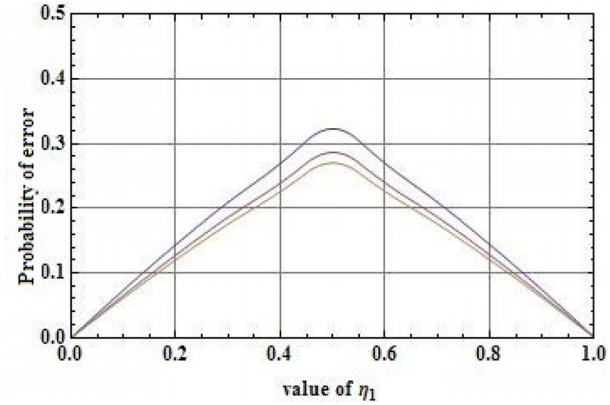


**Fig. 8.** Known great circle, error rate in optimal discrimination for $\{1, 1, 1\}$ (upper curve) to $\{1, 3, 1\}$ (lower curve).

### 4.2 Fixed overlap with unambiguous discrimination

The results in this section are presented for a mid-range overlap ($\beta = \pi/2$). Figures for other degrees of overlap can be found by running the programs in [2].

As we mentioned briefly in the previous section, the graph of the relationship between the error rate and $\eta_1$, for any configuration, has three distinct sectors. In the centre of the range the best unambiguous discrimination is provided by a POM with three components. Outside this range, the 'best' discrimination rate is obtained by using one or other of the projectors $\hat{P}_1$ or $\hat{P}_2$. When $\eta_1$ is small, the error rate is given by:

$$\text{Failure rate} = \eta_1 + (1 - \eta_1)\text{Tr}(\hat{\rho}_2 \hat{P}_2). \qquad (15)$$

A similar equation holds when $\eta_2$ is small.

It is worth noting that when only one projector is used, one of the possible states of the system can sometimes be recognised unambiguously, but the other can never be detected. This suggests that in some circumstances, it might be better to keep using the three-component POM, even though it does not give the 'best' discrimination rates.

Typical curves are shown in Figures 5–7. Figures 5 and 6 use a fixed overlap of $\frac{1}{\sqrt{2}}$. Corresponding figures drawn from analytic studies based on random program bits would look the same. We include our results here to give a more complete overview of the topic. The discrimination rates that correspond to the individual projectors are shown as dashed lines.

These curves illustrate two trends:

(1) Symmetrical configurations, where the data qbit is repeated several times, improve the rate of successful discrimination. They also widen the useful range of the three-component POM.
(2) Non-symmetrical configurations, in which one of the program qbits is duplicated, narrow the range of the POM, and give hardly any improvement in the rate of discrimination.

### 4.3 Known great circle with optimal discrimination

This set of results has one dimension fewer than the previous set, as we do not need to investigate different overlaps.

Figure 8 shows the error rates for three symmetrical configurations – $\{1, 1, 1\}$, $\{1, 2, 1\}$ and $\{1, 3, 1\}$. Increasing the number of data qbits gives a modest improvement of some 20% in the correct discrimination rate.

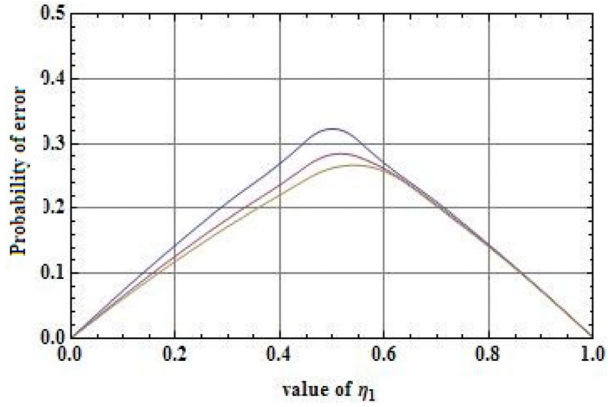


**Fig. 9.** Error rates (great circle, optimal discrimination) for $\{1,1,1\}$ (upper curve) to $\{1,1,3\}$ (lower curve).

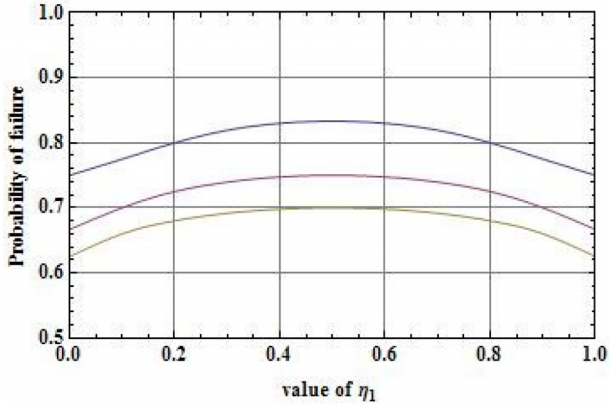


**Fig. 10.** Failure rates (great circle, unambiguous discrimination) for $\{1,1,1\}$ (upper curve) to $\{1,3,1\}$, (lower curve).

Figure 9 illustrates the error rate for non-symmetrical configurations – $\{1,1,1\}$ to $\{1,1,3\}$. As the number of program qbits on the right increases, the discrimination rate improves slightly, and the optimum value of $\eta_1$ shifts away from 0.5.

### 4.4 Known great circle with unambiguous discrimination

Failure rates for $\{1,1,1\}$ to $\{1,3,1\}$ are shown in Figure 10.

Finally, Figure 11 shows the failure rates for non-symmetrical configurations – $\{1,1,1\}$ to $\{1,1,3\}$.

### 4.5 Confinement to polar caps with optimal discrimination

Figure 3 shows how the angle $\theta$ is related to the size of the polar caps that hold the program qbits. Figure 12 plots the variation of the minimum error against $\eta_1$ for different values of $\theta$. As the area of confinement decreases, so does the minimum error.

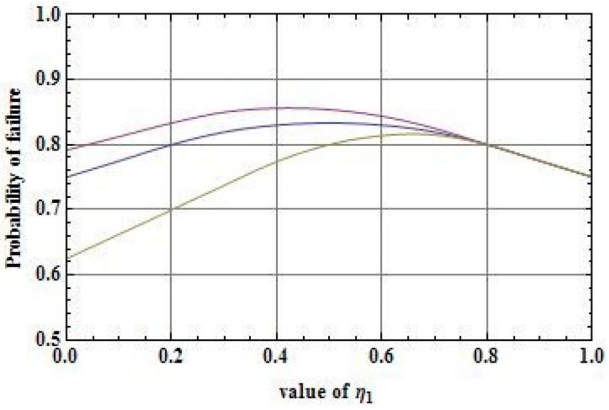


**Fig. 11.** Failure rates (great circle, unambiguous discrimination) for $\{1,1,1\}$ (upper curve), $\{1,1,2\}$ (central curve) and $\{1,1,3\}$ (lower curve).

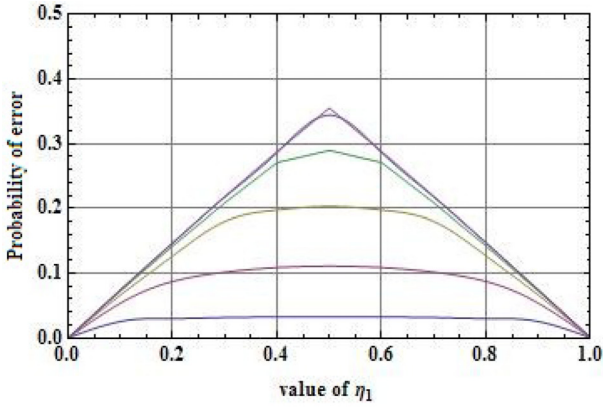


**Fig. 12.** Optimal error rates (polar cap confinement) for $\{1,1,1\}$. The upper curve is for $\theta = \pi$, and the lower one for $\theta = \pi/6$.

Doubling up the data qbit has little overall effect except to decrease the error rate slightly. Doubling up the right program qbit skews the diagram to the right.

### 4.6 Confinement to polar caps with unambiguous discrimination

Figure 13 shows how the unambiguous discrimination rate varies with the degree of confinement for various sizes of polar cap. This diagram is for the asymmetric configuration $\{1,1,3\}$, but apart from a small skew it is similar to other configurations in this context.

In summary, it appears that discrimination rates in both the minimum error and unambiguous regimes depend strongly on the degree to which the program qbits are confined to the poles of the Bloch sphere. The number of qbits available, whether as data or program qbits, has only minor influence.

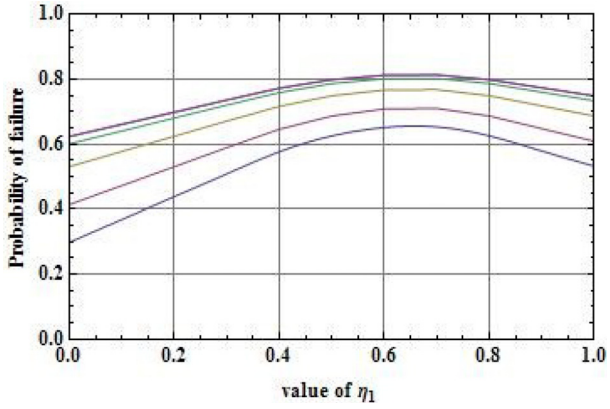


**Fig. 13.** Failure rates (polar cap confinement) for $\{1, 1, 1\}$. The upper curve is for $\theta = \pi$, and the lower one for $\theta = \pi/6$.

## 5 Transmitting data using programmed discrimination

The merits of using qbits to ensure data security are well established, and several appropriate protocols have been defined [18–20].

A potential application of programmed discrimination is in data transmission. The two possibilities $|\psi_2\rangle \equiv |\psi_1\rangle$ and $|\psi_2\rangle \equiv |\psi_3\rangle$ are used to encode the binary digits 0 and 1. In this section we examine the data-carrying capacity of a data transmission system that uses programmed discrimination. We use some of the results presented in previous sections of the paper, as well as those derived from expressions in other published material [3]. A useful measure will be the amount of information, in *bits*, reliably transmitted by a single qbit.

The advantage of programmed discrimination over other methods is that the qbits can be allowed to undergo any unitary transformation during transmission without affecting the data they represent, as long as the same transformation applies to all the qbits in a group. In our study this advantage holds for groups in which the overlap of the program qbits is fixed, but not for the other two modes, as an arbitrary transformation could move the qbits away from a known great circle, or away from given polar caps on the Bloch sphere. Another restriction is that we only consider configurations in which the two possible signals are handled symmetrically.

It is well known that any digital data transmission system will work at its highest efficiency if the expectations of zeros and ones are statistically equal. There exist many algorithms to compress data to this standard. We would therefore expect that in any discrimination system, $\eta_1$ and $\eta_2$ (as previously defined in Eqs. (1) and (2) should both be 0.5. If one used a configuration with unequal expectations where, for example, $\eta_1 = \frac{1}{3}$ and $\eta_2 = \frac{2}{3}$, the data would have to be recoded so that the ratio of zeros and ones was 2 to 1. This is not a realistic proposal.

Another form of asymmetry occurs when the numbers of program qbits of each type are not equal, as in the configuration $\{1, 3, 2\}$. The figures we have derived for these configurations are *averages*, but without symmetry the recognition rates will be different for zeros, as opposed to ones. The analysis of this situation is beyond the scope of this paper.

**Table 5.** Bits per qbit, with random program qbits and optimal discrimination.

| Configuration | $\{1, n, 1\}$ | $\{2, n, 2\}$ | $\{3, n, 3\}$ |
|---|---|---|---|
| Data qbits (n) | | | |
| 1 | 0.02033 | 0.21921 | 0.01908 |
| 2 | 0.02304 | 0.02411 | 0.02064 |
| 3 | 0.02222 | 0.02292 | 0.01959 |
| 4 | 0.02063 | 0.02111 | |
| 5 | 0.01899 | | |
| 6 | 0.01747 | | |
| 7 | 0.01613 | | |

### 5.1 Using optimal recognition

Consider a channel in which the probability of receiving a binary digit incorrectly is $p$. Shannon's Noisy Channel theorem [21] states that provided that $p \neq 0.5$, standard error correction methods allow the channel to be used for error-free transmission, albeit at a reduced data rate. The effective data rate $R$, per bit transmitted, is:

$$R = 1 + p \times \log_2(p) + (1 - p) \times \log_2(1 - p) \text{ bits.} \quad (16)$$

Suppose a group of $n$ qbits is used to transmit a (classical) bit, with an expected error rate of $p$. The amount of information $k$ sent by one qbit is:

$$k = \frac{1 + p \times \log_2(p) + (1 - p) \times \log_2(1 - p)}{n}. \quad (17)$$

Using error rates for optimal discrimination with no correlation between the program qbits [3], the information transmitted per qbit, for various configurations, is shown in Table 5. The efficiency of communication in this context, when compared to a possible standard of one qbit per bit, is barely 2% with the configuration $\{1, 1, 1\}$. This figure improves slightly as the data bit is duplicated, and then falls away again when further copies are brought into play. With multiple program qbits the best configuration, by a small margin, is $\{2, 2, 2\}$.

In previous sections of this paper, we have calculated the expected error rates for various configurations, and for various degrees of fixed overlap. For orthogonal program qbits (where the overlap is 0) the number of bits per qbit is shown in Table 6. Here the presence of classical information gives a four-fold improvement, and further illustrates the small, but real advantage of using multiple qbits.

### 5.2 Using unambiguous recognition

Shannon's theorems are not relevant to the protocol that might be used with unambigous discrimination. Suppose Alice is sending data to Bob [22], using the basic $\{1, 1, 1\}$ configuration with no correlation between the data qbits.

**Table 6.** Bits per qbit, with orthogonal program qbits and optimal discrimination.

| Configuration | {1, $n$, 1} | {2, $n$, 2} | {3, $n$, 3} |
|---|---|---|---|
| Data qbits ($n$) | | | |
| 1 | 0.08533 | 0.09690 | 0.08721 |
| 2 | 0.09978 | 0.11219 | 0.10013 |
| 3 | 0.09840 | 0.10996 | 0.09800 |
| 4 | 0.09295 | 0.10339 | |
| 5 | 0.08671 | | |
| 6 | 0.08068 | | |
| 7 | 0.07513 | | |

**Table 7.** Bits per qbit, with random program qbits and unambiguous discrimination.

| Program qbits | $1+1$ | $2+2$ | $3+3$ |
|---|---|---|---|
| Data qbits | | | |
| 1 | 0.05556 | 0.04000 | 0.03061 |
| 2 | 0.06250 | 0.04630 | 0.03515 |
| 3 | 0.06000 | 0.04463 | |
| 4 | 0.05556 | 0.04375 | |
| 5 | 0.05102 | | |
| 6 | 0.04688 | | |
| 7 | 0.04321 | | |

**Table 8.** Bits per qbit, with orthogonal program qbits and unambiguous discrimination.

| Program qbits | $1+1$ | $2+2$ | $3+3$ |
|---|---|---|---|
| Data qbits | | | |
| 1 | 0.11111 | 0.08000 | 0.06123 |
| 2 | 0.12500 | 0.09259 | 0.07031 |
| 3 | 0.12000 | 0.09184 | |
| 4 | 0.11111 | 0.08750 | |
| 5 | 0.10204 | | |
| 6 | 0.09375 | | |
| 7 | 0.08642 | | |

We know that Bob can only make an identification in one sixth of the triads, but in these cases there is no uncertainty about the outcome. For effective communication Bob must use a conventional data link to tell Alice when he receives a triad that he recognises. To send a bit reliably, Alice must keep re-transmitting the corresponding triad until she learns that Bob has received it. With the basic configuration Alice must send an average of six triads or 18 qbits for every bit, giving a rate of 0.0555 bits per qbit.

The effectiveness of different configurations without correlation is shown in Table 7, and that for zero overlap in Table 8. The best overall rate is achieved with the configuration $\{1, 2, 1\}$, using orthogonal program qbits and unambiguous recognition.

# 6 Conclusion

Our overall conclusion, which applies over all the cases we have considered, is that increasing the qbit count generally gives only modest increases in the rate of successful discrimination.

One exception to the rule, worth noting, is that confining the program qbits to small areas near the poles of the Bloch sphere does lead to a substantial improvement in the rate of correct discrimination. But this restriction violates the freedom to subject all the qbits in the triad to *any* unitary transformation, so the finding is of doubtful value, at least where data transmission is concerned.

Another curious result is that in the minimum error mode, the error rate depends mainly on the number of qbits used to transmit a bit, and much less on how they are distributed between program and data qbits.

For data transmission, the best performance we could find is given by a configuration with orthogonal qbits, unambiguous discrimination and some replication of the data qbit. We would need 8 qbits to transmit a single binary digit reliably. This figure is poor when compared to other communication methods. There would need to be good reasons to justify the use of this method in a practical situation.

This work was supported by the UK Engineering and Physical Research Council. I am grateful to the referees for their comments on the first version of the paper, and to the University of Strathclyde, whose financial suppport allowed this work to be done. I specially wish to thank my supervisors, Prof. S.M. Barnett and Dr. J. Jeffers, for their support and encouragement, and to Andrew Mark Colin and Veronica Colin for a careful reading of the manuscript.

# Appendix A: Details of the algorithms used

This appendix is split into two main sections. The first deals with the calculation of density matrices for each of the three configurations. Once density matrices have been found, they can be used to work out out both the optimal error rate and the unambiguous failure rate in discrimination.

The second section describes the algorithms that derive the actual rates for a given pair of density matrices. These methods are used for all three configurations.

## A.1 Computing density matrices

This section deals with computing density matrices for each of the three types of relationship between the program qbits: fixed overlap, given great circle, and confinement to polar caps. In each case, the method is based closely on that described in [1], but extended to allow for multiple qbits.

### A.1.1 Density matrices for fixed overlap

Following the approach in [1], we provisionally place one of the program qbits at the north pole of the Bloch sphere.

$$|\psi_{1(prov)}\rangle = |0\rangle.$$

The other provisional program qbit can be written as:

$$|\psi_{3(prov)}\rangle = \cos(\beta/2)|0\rangle + e^{i\alpha}\sin(\beta/2)|1\rangle$$

where the known overlap is $\cos^2(\beta/2)$. This is illustrated in Figure 1. This leads to a provisional density matrix:

$$\begin{aligned}\overline{\hat{\rho}_{1(prov)}} &= \cos^2(\beta/2)|0\rangle^{\otimes(x+y+z)}\langle 0|^{\otimes(x+y+z)} \quad (A.1)\\ &+ \sin^2(\beta/2)|0\rangle^{\otimes(x+y)}|1\rangle^{\otimes z}\langle 0|^{\otimes(x+y)}|\langle 1|^{\otimes z}\end{aligned}$$

where $x$, $y$, and $z$ form the pattern of replication.

The provisional density matrix for $|\Psi_{2(prov)}\rangle$ is found by placing the *other* program qbit at the north pole. It is:

$$\begin{aligned}\overline{\hat{\rho}_{2(prov)}} &= \cos^2(\beta/2)|0\rangle^{\otimes(x+y+z)}\langle 0|^{\otimes(x+y+z)}\\ &+ \sin^2(\beta/2)|1\rangle^{\otimes x}|0\rangle^{\otimes(y+z)}\langle 1|^{\otimes x}\langle 0|^{\otimes(y+z)}.\end{aligned} \quad (A.2)$$

Next we apply a generalising transformation that lets both qbits be anywhere on the Bloch sphere. We replace $|0\rangle$ by the general form:

$$|0'\rangle = \cos(\theta/2)|0\rangle + e^{i\psi}\sin(\theta/2)|1\rangle \quad (A.3)$$

and $|1\rangle$ by the state orthogonal to $|0'\rangle$, namely,

$$|1'\rangle = \sin(\theta/2)|0\rangle - e^{i\psi}\cos(\theta/2)|1\rangle. \quad (A.4)$$

The generalised forms of $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are now given by

$$\begin{aligned}|\Psi_1\rangle &= \cos(\beta/2)|0'0'0'\rangle^{\otimes(x+y+z)} \quad (A.5)\\ &+ e^{i\alpha}\sin(\beta/2)|0'0'\rangle^{\otimes(x+y)}|1'\rangle^{\otimes z} \quad (A.6)\\ \Psi_2\rangle &= \cos(\beta/2)|0'0'0'\rangle^{\otimes(x+y+z)} \quad (A.7)\\ &+ e^{i\alpha}\sin(\beta/2)|1'\rangle^{\otimes x}|0'0'\rangle^{\otimes(y+z)}.\end{aligned}$$

We can write $|\Psi_1\rangle$ as the weighted sum of two column vectors:

$$|\Psi_1\rangle = \cos(\beta/2)|p\rangle + \sin(\beta/2)|q\rangle \quad (A.8)$$

where $|p\rangle$ is the expansion of the term $|0'0'0'\rangle^{\otimes(x+y+z)}$, in terms of the computational basis states $|0\rangle$ and $|1\rangle$, and $|q\rangle$ is the expansion of $\left(|0'0'\rangle^{\otimes(x+y)}|1'\rangle^{\otimes z}\right)$.

Similarly,

$$|\Psi_2\rangle = \cos(\beta/2)|p\rangle + \sin(\beta/2)|r\rangle \quad (A.9)$$

where $|p\rangle$ is the same as before and $|r\rangle$ is the expansion of the term $\left(|1'\rangle^{\otimes x}|0'0'\rangle^{\otimes(y+z)}\right)$.

Let $\mathbf{A}$ denote the outer product $|p\rangle\langle p|$ and $\mathbf{B}$ denote $|q\rangle\langle q|$. The density matrix $\overline{\hat{\rho}}_1$ is the average, over the Bloch sphere, of the weighted sum of two matrices $\mathbf{A}$ and $\mathbf{B}$:

$$\overline{\hat{\rho}}_1 = \frac{1}{4\pi}\left(\cos^2(\beta/2)\int_S \mathbf{A}dS + \sin^2(\beta/2)\int_S \mathbf{B}dS\right). \quad (A.10)$$

The integral $\int_S dS$ denotes integration over the surface of the Bloch sphere.

Similarly, using the definintion of $\overline{\hat{\rho}}_2$ we find

$$\overline{\hat{\rho}}_2 = \frac{1}{4\pi}\left(\cos^2(\beta/2)\int_S \mathbf{A}dS + \sin^2(\beta/2)\int_S \mathbf{C}dS\right) \quad (A.11)$$

where $\mathbf{C}$ is the outer product $|r\rangle\langle r|$.

We now turn to the calculation of the individual elements of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. Let $n$ be the total number of qbits in the configuration ($n = x+y+z$). Each of the matrices will be of order $2^n$, and we can label the rows and columns by the binary numbers 0 to $(2^n-1)$, so that – for example – $A[J,K]$ is the element of A at row $J$ and column $K$.

Each element of any of the three matrices is the average, taken over the surface of the Bloch sphere, of an expression that has two factors:

(a) a ket $|X_j\rangle$ that is the same for all the elements in column $J$.
(b) a bra $\langle Y_k|$ that is the same for all the elements in row $K$.

We may put:

$$I[J,K] = \left(\frac{1}{4\pi}\right)\int_S |X_J\rangle\langle Y_K|$$

where the integration is over the whole sphere. $\mathbf{I}$ stands for $\mathbf{A}$, $\mathbf{B}$ or $\mathbf{C}$ as appropriate.

It is useful to code the current configuration as three binary numbers $P$, $Q$ and $R$:

$$P = \underbrace{0000000000\ldots0000000000}_{\longleftarrow\ n\ \longrightarrow} \quad (A.12)$$

$$Q = \underbrace{00000\ldots00000}_{\longleftarrow\ x+y\ \longrightarrow}\,\underbrace{11\ldots11}_{\longleftarrow\ z\ \longrightarrow} \quad (A.13)$$

$$R = \underbrace{111\ldots111}_{\longleftarrow\ x\ \longrightarrow}\,\underbrace{00000\ldots00000}_{\longleftarrow\ y+z\ \longrightarrow} \quad (A.14)$$

$|X_J\rangle$ for any of the matrices $\mathbf{A}$, $\mathbf{B}$ or $\mathbf{C}$ can now be worked out by taking the appropriate number ($P$, $Q$ or $R$) and matching it, bit by bit, against the binary version of the column label $J$ . It is clear, from the definition of $|0'\rangle$ and $|1'\rangle$ that $|X_J\rangle$ will be the product of:

- a sign factor $s_1$, which may be $+1$ or $-1$;
- a term $e^{t_1 i\phi}$ where $t_1$ is an integer such that $(t_1 \geq 0)$;
- the cosine of the angle $(\theta/2)$, raised to the power $u_1$;
- the sine of the angle $(\theta/2)$, raised to the power $v_1$. In all cases $(u_1 + v_1) = n$.

The generation of each term in $|X_J\rangle$ starts by initialising four integer variables which represent a partial product:

- $s_1 = +1$ (the sign of the term);
- $u_1 = 0$ (the power of the cosine term);
- $v_1 = 0$ (the power of the sine term);
- $t_1 = 0$ (the coeffiicent of $\psi$).

The final product will be:

$$|X_J\rangle = s_1 e^{t_1 i\phi}\cos^{u_1}(\theta/2)\sin^{v_1}(\theta/2).$$

Each step in generating this product consists of taking a digit from one of $P$, $Q$ or $R$ (say $i$) and the corresponding digit from the binary form of $J$, (say $j$) and applying the following rule. The order in which the bits are used is immaterial:

- $i = 0, j = 0$: increment the cosine count $u_1$;
- $i = 0, j = 1$: increment the sine count $v_1$ and the exponent count $t_1$;
- $i = 1, j = 0$: increment the sine count $v_1$;
- $i = 1, j = 1$: increment the cosine count $u_1$, and the exponent count $t_1$. Change the sign of $s_1$.

The rule works because each of the four possible combinations of $i$ and $j$ identifies one of the terms in the expansion of $|0'\rangle$ and $|0'\rangle$ and multiplies it into the partial product.

The derivation of $\langle Y_K|$ is the same, except that we use a row label $K$ instead of a column label. As we are working on a ket, we decrement the exponent count instead of adding to it. We find:

$$\langle Y_K| = s_2 e^{t_2 i\phi} \cos^{u_2}(\theta/2) \sin^{v_2}(\theta/2)$$

and expect $t_2$ to be *negative*.

The final term to be integrated is the product of $|X_J\rangle$ and $\langle Y_K|$. As the integration is over the surface of a sphere we factor in $\sin(\theta)$, in the form $2\sin(\theta/2)\cos(\theta/2)$.

The final expression for the average is:

$$\frac{s}{4\pi}\left(\int_0^{2\pi} d\phi e^{ti\phi} \int_0^{\pi} d\theta \cos^{u+1}(\theta/2)\cos^{v+1}(\theta/2)\right)$$

where $s = s_1 \times s_2, t = t_1 + t_2, u = u_1 + u_2,$ and$v = v_1 + v_2$.

A useful feature of this expression is that it evaluates to zero for all values of $t$ except 0. As the values of $t_1$ and $t_2$ for any column or row depend only on the number of ones in the binary expansions of $J$ and $K$, it follows that the only non-zero terms in the matrix will be those where the intersecting row and column labels have the same number of ones. For these cells, the average can be worked out by a simple recursive method that uses a standard integral:

Define $I(m, n)$ as $\int_0^{\pi} \cos^n(\theta/2)\sin^m(\theta/2)d\theta$. Then

$$n = 1 \Rightarrow I(m, n) = \frac{1}{m+1}$$

otherwise,

$$I(m, n) = I(m, n-2) - I(m+2, n-2).$$

### A.1.2 Density matrices when both program qbits are on a given great circle

In this section, it is given a priori that the program qbits lie on a pre-defined great circle, as shown in Figure 2. We note that any great circle on the Bloch sphere can be transformed into any other, just by rotating the sphere. The most convenient one to use is the 'Greenwich meridian', where the azimuth angle is always zero. We may put:

$$|\psi_1\rangle = \cos(\alpha/2)|0\rangle + \sin(\alpha/2)|1\rangle \quad \text{(A.15)}$$

$$|\psi_3\rangle = \cos(\gamma/2)|0\rangle + \sin(\gamma/2)|1\rangle. \quad \text{(A.16)}$$

Following the established pattern,

$$|\Psi_1\rangle = |\psi_1\rangle^{\otimes(x+y)}|\psi_3\rangle^{\otimes z}.$$

As $|\psi_1\rangle$ and $|\psi_3\rangle$ are independent, we can define the density matrix of the first state as:

$$\overline{\hat{\rho}_1} = \hat{\lambda}_1 \otimes \hat{\mu}_1$$

where

$$\hat{\lambda}_1 = \psi_1\rangle^{\otimes(x+y)}\langle\psi_1|^{\otimes(x+y)} \quad \text{(A.17)}$$

$$\hat{\mu}_1 = \psi_3\rangle^{\otimes z}\langle\psi_3|^{\otimes z}. \quad \text{(A.18)}$$

We compute $\hat{\lambda}_1$ and $\hat{\mu}_1$ separately, and then form their tensor product to find $\overline{\hat{\rho}_1}$.

The derivation of the individual terms in either of the matrices follows the general lines described in the previous section.

Each term has the form:

$$\left(\frac{1}{2\pi}\int_0^{2\pi} d\omega \cos^v(\omega/2)\sin^u(\omega/2)\right)$$

where $\omega$ stands for $\alpha$ or $\gamma$, and $(v+u)$ is twice the number of bits in each row or column label.

The integration rule is:

- define $J(m, n)$ as $\int_0^{2\pi} d\omega \cos^n(\omega/2)\sin^m(\omega/2)$. Note that $(m + n)$ cannot be odd;
- $m$ and $n$ both odd $\Rightarrow J(m, n) = 0$;
- $n = 0 \Rightarrow J(m, n) = \prod_{h=1}^{m/2}\left(\frac{2h-1}{2h}\right)$;
- otherwise $J(m, n) = J(m, n-2) - J(m+2, n-2)$.

The density matrix of the second possible state $\overline{\hat{\rho}_2}$ is found by multiplying $\hat{\lambda}$ and $\hat{\mu}$ in the opposite order:

$$\overline{\hat{\rho}_2} = \hat{\mu} \otimes \hat{\lambda}. \quad \text{(A.19)}$$

### A.1.3 Density matrices when the program qbits are confined to polar caps

Figure 3 shows the assumed positions of the polar caps that hold the program qbits. Since these qbits are independent, we can use the approach set out in the previous section, and define each density matrix as the tensor product of two smaller matrices, entirely based on one or other of the polar caps.

We can again define the density matrix of the first state as:

$$\overline{\hat{\rho}_1} = \hat{\lambda}_1 \otimes \hat{\mu}_1$$

where

$$\hat{\lambda}_1 = \psi_1\rangle^{\otimes(x+y)}\langle\psi_1|^{\otimes(x+y)} \quad \text{(A.20)}$$

$$\hat{\mu}_1 = \psi_3\rangle^{\otimes z}\langle\psi_3|^{\otimes z}. \quad \text{(A.21)}$$

The general form of $|\psi_1\rangle$ is $\cos(\alpha/2) + e^{i\beta}\sin(\alpha/2)$. Assuming that $|\psi_1\rangle$ is in the northern cap, bounded by the

angle $\theta$, each term in the density matrix $\hat{\lambda}$ will be of the form:

$$\frac{1}{\kappa}\int_0^{2\pi} d\beta e^{ni\beta}\int_0^{\theta} d\alpha \cos^u(\alpha/2)sin^v \sin(\alpha/2)$$

where $\kappa$ is the area of the cap ($\kappa = 2\pi(1-\cos(\theta))$) and $(u+v)$ = the number of bits needed to label a row or column of the matrix.

The powers $v$ and $u$ are obtained in the same way as in the other cases, but the integration process is different:

- define $K(m,n,\theta)$ as $\int_0^{\theta} d\omega \cos^n(\omega/2)\sin^m(\omega/2)$;
- $m$ and $n$ both odd $\Rightarrow K(m,n)=0$;
- $n = 0 \Rightarrow K(m,n) = \left(\frac{4}{m-1}\right)\left(\frac{1}{1-\cos\theta}\right)(1-\cos^{m+1}(\theta/2))$;
- otherwise $K(m,n) = K(m,n-2,\theta) - K(m+2,n-2,\theta)$.

## A.2 Derivation of error rates from the density matrices

This section describes how error rates can be found from any pair of density matrices derived from the two possible states $|\Psi_1\rangle$ and $|\Psi_2\rangle$. To calculate the results presented in Section 4 we applied these methods to density matrices generated in the three different ways: fixed overlap, known great circle and polar caps. The methods would, however, be valid for other pairs of density matrices as well.

In all cases the calculations involve finding the eigenvalues of large matrices. For example, to analyse the configuration $\{x,y,z\}$ we must find the eigenvalues of two matrices, each with $2^{2(x+y+z)}$ elements. This is computationally hard, and the time needed grows exponentially with the number of elements.

We define $s$ as $2^n$ where $n = x+y+z$, total number of qbits in the configuration.

The algorithm for finding the optimum error rate uses Helstrom's rule. It is straightforward and is amply described in the overview Section 3.

Finding the unambiguous discrimination error rate requires two main steps and follows the method described in [5].

First we need to find two projectors $\hat{P}_1$ and $\hat{P}_2$, such that:

(a) $\mathrm{Tr}(\hat{\rho}_1.\hat{P}_2) = \mathrm{Tr}(\hat{\rho}_2.\hat{P}_1) = 0$;
(b) $\hat{P}_1$ and $\hat{P}_2$ are completely orthogonal to one another.

Using Jacobi's method, we calculate the $s$ eigenvectors of $\hat{\rho}_1$ and $\hat{\rho}_2$. and discard those withzero eigenvalues. We are left with a (rectangular) array $Z_1$ with $s$ rows and $x$ columns from $\hat{\rho}_1$, and an array $Z_2$ of dimension $s \times y$ from $\hat{\rho}_2$. Both $x$ and $y$ may be less than $s$.

We suppose that each vector in $Z_2$ is composed of a linear combination of all the vectors in $Z_1$, plus a residual vector, which is what we need. This allows to write down $x$ equations in $s$ unknowns, which would not in general be enough to determine the residual vector, as $x \leq s$. We can find the extra equations by specifying that the residual vector is orthogonal to each of the vectors in $Z_1$.

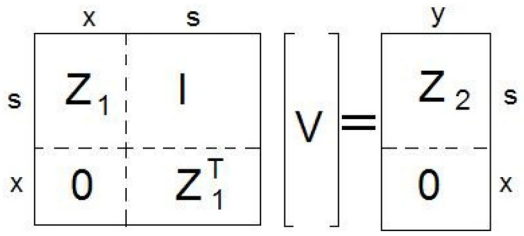


**Fig. A.1.** System of equations to calculate projectors.

The conditions can be coded into a set of simultaneous linear equations, as shown in Figure A.1. Here all the vectors in $Z_2$ have been assembled into a matrix, padded out with zeros.

When this system is solved using Gaussian elimination, the resultant array $V$ holds the orthogonal vectors in its lower $s$ columns. Naming them $|v_1\rangle$ to $|v_y\rangle$, the projector $\hat{P}_1$ can be constructed as:

$$\hat{P}_1 = \sum_{k=1}^{y}(|v_k\rangle\langle v_k|) \qquad (A.22)$$

$\hat{P}_2$ is constructed in exactly the same way.

As we have seen in equation (14), the probability for successful unambiguous discrimination is:

$$Q = \eta_1 \mathrm{Tr}(c_1\hat{\rho}_1\hat{P}_1) + \eta_2 \mathrm{Tr}(c_2\hat{\rho}_2\hat{P}_2) \qquad (A.23)$$

$c_1$ and $c_2$ are constants which must be selected so that $\hat{\Pi}_0$, defined in equation (13), has a smallest eigenvalue of zero, and $Q$ has a minimum value.

It turns out that the relationship between the constants and the smallest eigenvalue of $\hat{\Pi}_0$ is not linear. We use a version of the *golden section search* [23] to satisfy these conditions.

For certain small values of $\eta_1$ and $\eta_2$ the search does not converge, and there are no acceptable values of $c_1$ and $c_2$. This defines the range where the three-component POM is not valid.

The calculation of unambiguous error rates for any configuration takes *much longer* than that for minimum error rates.

## References


1. A.J.T. Colin, S.M. Barnett, J. Jeffers, Eur. Phys. J. D **63**, 463 (2011)
2. A.J.T. Colin, Multiple qbit discrimination programs, `ftp://www.singing-baboon.com/discrimination.zip`
3. J.A. Bergou, V. Bužek, E. Feldman, U. Herzog, M. Hillery, Phys. Rev. A **73**, 062334 (2006)
4. J.A. Bergou, M. Hillery, Phys.: Rev. Lett. **94**, 160501 (2005)
5. U. Herzog, J.A. Bergou, J. Phys.: Conf. Ser. **36**, 49 (2006)
6. D. Hanneke, J.P. Home, J.D. Jost, J.M. Amini, D. Leibfried, D.J. Wineland, Nat. Phys., `DOI:10.1038/nphys1453`

7. A. Hayashi, T. Hashimoto, M. Horibe, Phys. Rev. A **72**, 032325 (2005)
8. A. Hayashi, M. Horibe, T. Hashimoto, Phys. Rev. A **73**, 012328 (2006)
9. B. He, J.A. Bergou, Phys. Rev. A **75**, 032316 (2007)
10. G. Sentis, E. Bagan, J. Calsamiglia, R. Muñoz-Topia, Phys. Rev. A **82**, 042312 (2010)
11. S.M. Barnett, A. Chefles, I. Jex, Phys. Lett. A **307**, 195 (2003)
12. I. Jex, G. Alber, S.M. Barnett, A. Delgado, Fortschr. Phys. **51**, 172, DOI:10.1002/prop.200310021
13. M. Sedlak, M. Ziman, V. Bužek, M. Hillery, Phys. Rev. A **77**, 042304 (2008)
14. M. Kleinmann, H. Kampermann, D. Bruß, Phys. Rev. A **72**, 032308 (2005)
15. S. Croke, *Maximum Confidence Measurements*, Ph.D. thesis, Strathclyde University, 2007
16. S.M. Barnett, *Quantum Information* (Oxford University Press, Oxford, 2009)
17. C.W. Helstrom, *Quantum detection and estimation theory* (Academic Press, New York, 1976)
18. C.H. Bennett, G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing* (Bangalore, 1984), p. 175
19. T. Jennewein, C. Simon, G. Weihs, H. Weinurter, A. Zellinger, Phys. Rev. Lett. **68**, 3121 (1992)
20. A.K. Ekert, J.G. Rarity, P.R. Tapster, G.M. Palma, Phys. Rev. Lett. **84**, 4729 (2000)
21. C.E. Shannon, W. Weaver, *The Mathematical Theory of Communication. Urbana* (University of Illinois Press, IL, 1949, reprinted 1998)
22. B. Schneier, *Applied Cryptography* (John Wiley and Sons, 1994)
23. W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, New York, 1988)
24. N.D. Mermin, *Quantum Computer Science* (Cambridge University Press, 2007)
25. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000)