



OPEN

Maximum information measurement for qubit states

Árpád Varga¹, Peter Adam^{1,2}✉ & János A. Bergou^{1,3}

We determine the optimal measurement that maximizes the average information gain about the state of a qubit system. The qubit is prepared in one of two known states with known prior probabilities. To treat the problem analytically we employ the formalism developed for the maximum confidence quantum state discrimination strategy and obtain the POVM which optimizes the information gain for the entire parameter space of the system. We show that the optimal measurement coincides exactly with the minimum-error quantum measurement only for two pure states, or when the two states have the same Bloch radius or they are on the same diagonal of the Bloch disk.

In quantum information the carriers of information are quantum systems and information is encoded in their states. Extracting this information is a central problem in quantum information processing and it can be done by determining the state via measurements¹.

In many quantum communication schemes, information is the state itself. In these schemes a sender, Alice, prepares an ensemble of quantum systems, each in a state from a set of n known states, $\{\rho_j | j = 1, \dots, n\}$, the letter states. The weight of state ρ_j in this initial ensemble is η_j , called the *a priori* probability or simply prior. Alice then randomly draws a system from this initial ensemble and sends it to the receiver, Bob. The set of possible states as well as their priors are also known to the receiver whose task is to identify the state of the system he received. If the states are mutually orthogonal the task is easy: Bob sets up detectors along these orthogonal directions and a click in one of them will perfectly determine the input. However, if the possible states are not mutually orthogonal, the problem is highly nontrivial. Bob needs to choose a figure of merit and find a measurement which is optimal with this respect. Accordingly, several strategies have been developed with respect to various criteria. Optimization, in general, leads to complex measurement strategies often involving generalized measurements. Some of the frequently employed strategies are discrimination with minimum error (ME)^{2–6}, unambiguous discrimination (UD)^{7–12}, and maximum confidence (MC)^{13–16} discrimination.

The ME strategy was first introduced in Refs.^{2–4} for two states (pure or mixed) with arbitrary priors. In this strategy, every time Bob receives a system he has to make a guess about its state based on the outcome of his measurement. The price to pay is that errors must be allowed. In the optimal strategy the average probability of error is minimized. The ME strategy involving more than two states is known in some special cases only.

The UD strategy was first introduced in Refs.^{7–9} for two pure states with equal priors and was later generalized for arbitrary prior probabilities in Ref.¹⁰. In the UD strategy, no errors are allowed. The price to pay is that Bob must be allowed to return inconclusive answers. In the optimal strategy the average probability of inconclusive answers is minimized. An important result states that UD is possible if the states are linearly independent¹⁷, which is not a requirement for ME. The UD strategy is successively used in sequential state discrimination, which is a strategy for N separate receivers^{18–22}.

We note that each strategy has its own advantages and drawbacks when we try to apply them for a general measurement problem. It is difficult to find measurements realizing unambiguous discrimination for mixed states, but it is relatively easy to generalize this strategy for more than two states, at least in principle²³. The ME strategy handles mixed states and pure states on equal footing but is hard to generalize for more than two states, except for some special, highly symmetric cases^{24,25} (although progress has been made recently in this area^{26,27}).

Another independent strategy, called Maximum Confidence (MC), was introduced in Ref.¹³. The aim of the MC strategy is to construct a measurement which maximizes the confidence C_j : the conditional probability that detector j clicks provided that the state ρ_j was prepared. In the case of linearly independent states this strategy coincides with the UD strategy. However, when the states to be discriminated are not linearly independent, this is an independent strategy²⁸. For further developments in this line of state discrimination studies we refer to the recent reviews^{14–16}.

¹Institute of Physics, University of Pécs, Pécs, Ifjúság útja 6, 7624, Hungary. ²Institute for Solid State Physics and Optics, HUN-REN Wigner Research Centre for Physics, Budapest, P.O. Box 49, 1525, Hungary. ³Department of Physics and Astronomy, Hunter College and the Graduate Center of the City University of New York, 695 Park Avenue, New York, NY 10065, USA. ✉email: adam.peter@wigner.hu

In quantum communication one can look for a measurement strategy maximizing the mutual information between the communicating parties. In this problem the sender sends a sequence of individual quantum systems, each taken from a given set of known states $\mathcal{E} = \{\rho_j | j = 1, \dots, n\}$, and the receiver measures them one by one, possibly by a POVM with the POVM elements Π_m where $m = 1, \dots, N$. Our task is to maximize the Shannon mutual information^{2,3,29},

$$I(m:j) = \sum_{m,j} p(m|j) \eta_j \ln \frac{p(m|j)}{p(m)}, \quad (1)$$

between the measurement outcomes m and the input states j , where η_j is the *a priori* probability of the state j , and $p(m|j) = \text{Tr}(\Pi_m \rho_j)$ is the conditional probability of getting the measurement result m , given that the state ρ_j was prepared.

For a given set of states and their *a priori* probabilities, the problem is to find a measurement which maximizes the mutual information. While the mutual information is known to obey the Holevo bound³⁰, it is important to determine the accessible information, which is the actual maximum over all possible measurements. This is a special problem in state discrimination: we want to maximize the correlation between state preparation and measurement outcomes, i.e., we want to devise a measurement strategy that yields maximum information about which state was prepared. We will refer to such a measurement as maximum information (MI) measurement. Solution to this problem is known only for a few special cases. Even determining the amount of information which can be encoded in a given quantum system is a nontrivial task^{31–36}. We note that the problem we address consists in maximizing mutual information between classical random variables linked via a quantum encoding–decoding scheme. Therefore, it is related to the calculation of channel capacities. It is, however, different from that of quantum channel capacity concepts maximizing quantum (as opposed to classical) mutual information^{37–41}.

The number of POVM elements needed to maximize the information gain is in general unknown. It is known⁴² that for any ensemble in d dimensions there is an optimal strategy with at most N elements, where $d \leq N \leq d^2$. Sasaki et al.⁴³ showed that for the case of real states (that is, states with a real density operator), the upper bound is $d(d+1)/2$. They also gave explicit solutions for the case of real and symmetric states in two dimensions, and showed that at most three POVM elements are necessary. Levitin⁴⁴ conjectured that if the number of the possible states is $N \leq d$, the optimal measurement will always be a von Neumann measurement. This conjecture was proven to hold for two pure states in arbitrary dimensions by Levitin. However, Shor⁴⁵ gave a counterexample involving three real pure states in three dimensions (qutrits). Considering qubits only, Keil⁴⁶ proved that von Neumann measurements are always optimal for two general states.

Fuchs and Peres⁴⁷ studied numerically the trade-off between the information gain and the measurement induced disturbance. Ban et al.⁴⁸ gave analytic results for pure binary signal states, and showed the connection between the ME measurement and the MI measurement for this special case. Řeháček et al.⁴⁹ gave an iterative algorithm to find the optimal POVM for the accessible information and illustrated the method on an example in three dimensions. There are also lower⁵⁰ and upper^{51,52} bounds to the accessible information for simple cases, which depend explicitly only on the message ensemble.

In this paper we consider the problem of finding the optimal measurement to maximize the mutual information for a general qubit system. Our approach makes use of the method developed for the maximum confidence strategy and leads to analytical insight. In particular, we determine the POVM in parametric form with a single parameter, which maximizes the information gain for the entire parameter space of the system.

Results

Information gain and confidence probabilities: General formalism

We begin with a study of the simplest case of two qubit states in a two-dimensional Hilbert space and present an alternative derivation of Eq. (1) for this case. We cast the result to a form that shows the intrinsic connection of information gain with the confidence probabilities, introduced in Ref.¹³.

Recall that we consider a two-party protocol with a sender, Alice, and a receiver, Bob. Alice prepares an ensemble of qubit systems where each qubit is either in the state ρ_1 or in the state ρ_2 . The first state is prepared a fraction η_1 of the time and the second state is prepared a fraction $\eta_2 = 1 - \eta_1$ of the time. The ensemble is described by the density matrix

$$\rho = \eta_1 \rho_1 + \eta_2 \rho_2 \quad (2)$$

Alice then randomly draws a quantum system from this ensemble and sends it over to Bob.

A more elaborate but equivalent way of describing state preparation is as follows. Alice initially prepares an ensemble of two-qubit states $\rho_{ab} = \eta_1 |0\rangle_a \langle 0| \otimes \rho_{1,b} + \eta_2 |1\rangle_a \langle 1| \otimes \rho_{2,b}$ and sends particle b over to Bob. Then she performs a measurement in the computational basis on the particle in her possession. If she finds the result $|0\rangle$ she knows that Bob's particle is in the state ρ_1 and if she finds the result $|1\rangle$ she knows that Bob's particle is in the state ρ_2 . If this is repeated a large number of times, Bob will receive the state ρ_1 a fraction η_1 of the time and the state ρ_2 a fraction η_2 of the time, on average.

Either way, Bob has no knowledge of the actual state he received, all he knows are the prior probabilities, η_1 and $\eta_2 = 1 - \eta_1$. The initial information uncertainty is given by

$$S_i = - \sum_{j=1}^2 \eta_j \log \eta_j, \quad (3)$$

and the initial information is

$$I_i = 1 - S_i = 1 + \sum_{j=1}^2 \eta_j \log \eta_j \quad (4)$$

The question we are addressing here is: How much information can Bob gain by performing measurement(s) on the system he received? To this end we will consider the following general model of quantum measurement. In accordance with the results described in the Introduction, we assume that Bob has N detectors described by the set of rank-1 operators $\{\Pi_m | m = 1, \dots, N\}$ with $d \leq N \leq d^2$ adding up to the identity operator,

$$\sum_{m=1}^N \Pi_m = \mathbb{1} \quad (5)$$

The latter condition ensures that given the measured system in any state, one of the detectors will click. The conditional probability that Π_m clicks, given a system in the state ρ_j is calculated using Born's rule:

$$P(m|j) = \text{Tr}(\rho_j \Pi_m) \quad (6)$$

In order to have positive probabilities we have to require the positivity (non-negativity) of the detection operators,

$$\Pi_m \geq 0 \quad (7)$$

Equations (5) and (7) define a Positive Operator Valued Measure (POVM), which is simply the decomposition of the identity in terms of positive operators, called the elements of the POVM.

Next, we use Bayesian updating, employing Bayes' theorem, $P(m|j)P(j) = P(j|m)P(m)$, for conditional probabilities. In particular, we apply this formula for the situation when $j = \rho_j$ and $m = \Pi_m$. Then $P(j) = \eta_j$ is the prior probability of state j , $P(m|j)$ is the detection probability given in Eq. (6), i.e., detector m records an event if state j is given, and $P(m) = \eta_1 \text{Tr}(\rho_1 \Pi_m) + \eta_2 \text{Tr}(\rho_2 \Pi_m)$ is the total probability that detector m records an event. Using Eq. (2), the last expression can be written as

$$P(m) = \text{Tr}(\Pi_m \rho) \quad (8)$$

Thus we find that the conditional probability $P(\rho_j|\Pi_m)$, the probability that if detector m records an event it is due to the state j , can be written as

$$P(\rho_j|\Pi_m) \equiv C_{jm} = \frac{\eta_j \text{Tr}(\rho_j \Pi_m)}{\text{Tr}(\Pi_m \rho)} \quad (9)$$

C_{jm} is the confidence (or confidence probability) which is the central quantity in the MC state discrimination strategy.

Equipped with the confidences we next give the information uncertainty for the case when detector m clicks. Clearly, using $C_{2m} = 1 - C_{1m}$,

$$S_m = -C_{1m} \log C_{1m} - (1 - C_{1m}) \log (1 - C_{1m}) \quad (10)$$

The average uncertainty is

$$S_f = \sum_{m=1}^N P(m) S_m, \quad (11)$$

where $P(m)$ is given by Eq. (8). The information after the measurement is given by $I_f = 1 - S_f$.

Finally, the information gain from the measurement can be given as

$$\Delta I = I_f - I_i = S_i - S_f, \quad (12)$$

where S_i is given by Eq. (3) and S_f is given by (11). Substituting the explicit expressions for the various quantities obtained so far into Eq. (12) it can be shown that this equation is identical to Eq. (1). In this formulation the contributions from the prior and posterior information appear clearly separated. Furthermore, S_i is constant for a given set of priors, so it is independent of the measurement we perform. Therefore, optimizing the information gain is equivalent to finding the POVM that minimizes the second term, the information uncertainty S_f , Eq. (11). As noted before, the information gain that is maximized over all possible measurements is also called the accessible information. In the next section we develop a fully analytical theory that provides the accessible information (optimal solution) in parametric form for all values of the parameters.

Accessible information for qubits

To treat the optimization problem effectively, we employ the formalism developed for the maximum confidence strategy in Ref.¹³. Equations (11) and (12) are expressed in terms of the confidence probabilities, so they provide a convenient starting point. The method yields the optimal solution analytically in parametric form, in terms of a single parameter.

As the first step we introduce transformed density and measurement operators by the definitions:

$$\bar{\rho}_j = \eta_j \rho^{-1/2} \rho_j \rho^{-1/2}, \quad (13)$$

$$\bar{\Pi}_m = \rho^{1/2} \Pi_m \rho^{1/2}, \quad (14)$$

and

$$\bar{\bar{\Pi}}_m = \frac{\bar{\Pi}_m}{\text{Tr}(\bar{\Pi}_m)}, \quad (15)$$

where ρ is defined in Eq. (2). The transformed states satisfy

$$\bar{\rho}_1 + \bar{\rho}_2 = \mathbb{1} \quad (16)$$

It follows from this expression that the transformed states $\bar{\rho}_{1,2}$ have the same set of eigenvectors. Using the transformed operators we can write the confidence C_{jm} , Eq. (9), in the more compact form,

$$C_{jm} = \text{Tr}(\bar{\bar{\Pi}}_m \bar{\rho}_j) \quad (17)$$

We wish to maximize the information gain (12) [or the final uncertainty S_f (11)] due to the measurement. S_f is already in terms of the confidences while the outcome probability becomes

$$P(m) = \text{Tr}(\bar{\bar{\Pi}}_m) \quad (18)$$

in terms of the transformed operators. Using Eqs. (14) and (15) it is easy to show from Eq. (5) that the transformed measurement operators $\bar{\bar{\Pi}}_m$ satisfy the equation:

$$\sum_m P(m) \bar{\bar{\Pi}}_m = \rho \quad (19)$$

We note that the transformation $\Pi_m \rightarrow \bar{\bar{\Pi}}_m$ is rank preserving. Thus they become rank 1 projectors, not necessarily orthogonal. All we can say is that, as seen from Eq. (19), they correspond to a pure state decomposition of ρ , not necessarily in terms of orthogonal pure states. Furthermore, a pair of qubit states are always unitarily equivalent to a pair of real states, their Bloch vectors can be chosen to span the x, z plane of the Bloch sphere. Therefore, we can assume that $\bar{\bar{\Pi}}_m$ is also real. The general expression of a real rank 1 matrix can be written as

$$\bar{\bar{\Pi}}_m = \begin{pmatrix} \cos^2 \alpha_m & \cos \alpha_m \sin \alpha_m \\ \sin \alpha_m \cos \alpha_m & \sin^2 \alpha_m \end{pmatrix} \quad (20)$$

For the calculation that follows it is convenient to use the common eigenvectors of $\bar{\rho}_1$ and $\bar{\rho}_2$ as basis. Let the eigenvectors be $|1\rangle$ and $|2\rangle$, and the eigenvalues of $\bar{\rho}_1$ be λ_1 and λ_2 . Eq. (16) immediately gives that the eigenvalues of $\bar{\rho}_2$ are $1 - \lambda_1$ and $1 - \lambda_2$. Substituting (20) first into (17), we obtain

$$\cos^2 \alpha_m = \frac{C_{1m} - \lambda_2}{\lambda_1 - \lambda_2} \quad (21)$$

Without loss of generality we can assume $\lambda_1 \geq \lambda_2$ from where $\lambda_1 \geq C_{1m} \geq \lambda_2$ and $1 - \lambda_2 \geq C_{2m} \geq 1 - \lambda_1$ follow. Substituting Eq. (20) next into Eq. (19), we find

$$\sum_m P(m) \cos^2 \alpha_m = \rho_{11}, \quad (22)$$

$$\sum_m P(m) \sin^2 \alpha_m = \rho_{22}, \quad (23)$$

$$\sum_m P(m) \cos \alpha_m \sin \alpha_m = \rho_{12} \quad (24)$$

Here, ρ_{ij} are the matrix elements of ρ in the basis formed by the eigenstates of $\bar{\rho}_1$.

Up to this point our consideration is general as we have not imposed any restriction on the number N of POVM elements. However, it has been proven that for a pair of qubit states the optimal measurement is projective^{44,46}, that is, $N = 2$. Then $\Pi_m = P_m$, where $\{P_m | m = 1, 2\}$ are rank 1 orthogonal projectors, $P_m P_{m'} = P_m \delta_{mm'} (m, m' = 1, 2)$. Therefore, from now on, we deal with the case of two orthogonal detectors and use the notation

$$P(j = 1 | m = 1) \equiv C_1, \quad (25)$$

$$P(j = 2 | m = 2) \equiv C_2, \quad (26)$$

since in this case we want to identify a click in detector $m = j$ with ρ_j . Hence C_j is the probability of “good” events for the corresponding detector.

Using this notation in (21) and then the resulting expression in (22) and (23), some lengthy but straightforward algebra yields

$$P(1) = \frac{C_2 - \eta_2}{C_1 + C_2 - 1}, \quad P(2) = \frac{C_1 - \eta_1}{C_1 + C_2 - 1} \quad (27)$$

Substituting (27) into Eq. (11) and then using the resulting expression in (12) we obtain

$$\Delta I = \frac{C_2 - \eta_2}{C_1 + C_2 - 1} [C_1 \log(C_1) + (1 - C_1) \log(1 - C_1)] + \frac{C_1 - \eta_1}{C_1 + C_2 - 1} [C_2 \log(C_2) + (1 - C_2) \log(1 - C_2)] - \eta_1 \log(\eta_1) - \eta_2 \log(\eta_2), \quad (28)$$

which is one of our central results. It expresses the information gain ΔI entirely in terms of the confidences C_1 and C_2 , and the prior probabilities η_1 and η_2 . Remarkably, this expression is independent of the structure of the states to be discriminated.

Equations (22) and (23) together with Eq. (21) allowed us to express the information gain entirely in terms of the confidence probabilities. The remaining Eq. (24) represents the main constraint under which (28) should be optimized. From Eqs. (24) and (27) it is easy to obtain the relation

$$\frac{\sin \alpha_1 \cos \alpha_1 - \rho_{12}}{C_1 - \eta_1} + \frac{\sin \alpha_2 \cos \alpha_2 - \rho_{12}}{C_2 - \eta_2} = 0 \quad (29)$$

The first term on the left-hand side is a function of quantities related to state 1 alone, while the second term is a function of quantities related to state 2 alone. Therefore, they separately must be equal to a universal function of the parameters, which we denote by a . The function a still depends on the parameters of the problem but not on C_1 and C_2 and, in order to satisfy (29), it must be antisymmetric under the exchange $1 \leftrightarrow 2$. In terms of a we can write

$$\frac{\sin \alpha_1 \cos \alpha_1 - \rho_{12}}{C_1 - \eta_1} = a, \quad \frac{\sin \alpha_2 \cos \alpha_2 - \rho_{12}}{C_2 - \eta_2} = -a \quad (30)$$

This provides a straightforward analytical solution to the entire problem. Substituting α_1 and α_2 from Eqs. (21), we obtain the constraint in parametric form. It expresses C_1 and C_2 and hence the information gain in terms of the single parameter a .

More importantly, however, it leads to a visual geometric solution which is the central result of this paper. It can serve as guide to find the exact solution for any values of the parameters specifying the problem. We introduce the geometric approach in the next subsection and illustrate its power on several examples.

Geometric optimization

First, we introduce a convenient parametrization of the problem. Recall that two qubit states are always unitarily equivalent to two real states: the corresponding two Bloch vectors span a plane in the Bloch sphere, and this plane can always be unitarily rotated to the $x - z$ plane. We can thus restrict our discussion to this plane, also termed as the Bloch disk. With a further rotation, the Bloch vector of one of the states, say ρ_1 can be aligned with the z axis. So we assume, without loss of generality, that our states are real from the beginning and the Bloch vector r_1 of the first state is along the z axis, that is, we use the following parametrization of the states:

$$\rho_i(r_i, \theta_i) = \frac{1}{2} (\mathbb{I} + r_i \sin(\theta_i) \sigma_x + r_i \cos(\theta_i) \sigma_z), \quad i = 1, 2. \quad (31)$$

Here, \mathbb{I} is the two-dimensional identity operator, σ_x and σ_z are Pauli matrices, r_i is the Bloch radius, and θ_i is the polar angle of state ρ_i , measured from the z axis. The parameters are shown in Fig. 1 where $0 < r_1, r_2 \leq 1$, $\theta_1 = 0$, $0 \leq \theta_2 \leq \pi$. In the following, we use these parameters to present our results.

For given fixed values of the parameters, that is, the prior probabilities η_1 and η_2 and the eigenvalues of the transformed states, λ_1 and λ_2 , the constraint (29) (or its parametric version, (30)) can be easily plotted in the C_1 - C_2 plane. This gives us a unique 8-shaped curve on which we have to find the optimal values of C_1 and C_2 . To this end, we notice that the information gain expression, (28), for a fixed value of ΔI is also a curve in the same plane. If we choose the fixed value ΔI too large, the two curves do not intersect. Lowering the value of ΔI , for a certain threshold value the two curves become tangent. This value is the maximal information gain ΔI_{\max} available by the measurement, that is, the accessible information. The procedure is illustrated in Fig. 2. We should also mention that the values $\Delta I < \Delta I_{\max}$ correspond to feasible (suboptimal) measurements, all the way to $C_1 = C_2 = 0.5$, which corresponds to pure guessing.

Figure 3 shows two examples for the geometric optimization, that is, the constraint (30) and the information gain (28) plotted together in the C_1 , C_2 plane, for two sets of parameters of the input states. The figure shows that increasing the prior probability η_1 of the state ρ_1 increases the optimal confidence probability C_1 of the state while reducing the confidence C_2 of the other state.

We find numerically that the optimal values are $C_1 = 0.4879$ and $C_2 = 0.9469$ for the left panel, while $C_1 = 0.7867$ and $C_2 = 0.8338$ for the right panel. It is interesting to note that in the first case both detectors identify ρ_2 with larger confidence.

In order to interpret these results, we point out that Figs. 2 and 3 are symmetric under reflection about the $C_1 + C_2 = 1$ line. This property follows from the fact that the information gain, Eq. (28), and the constraint, Eq. (29), are invariant under the substitution $C_1 \leftrightarrow 1 - C_2$. In particular, the constraint which is represented by the 8-shaped dashed line in these plots has this symmetry and the point where it intersects itself has coordinates $C_1 = \eta_1$ and $C_2 = \eta_2$. These values correspond to pure guessing with no actual measurement performed and

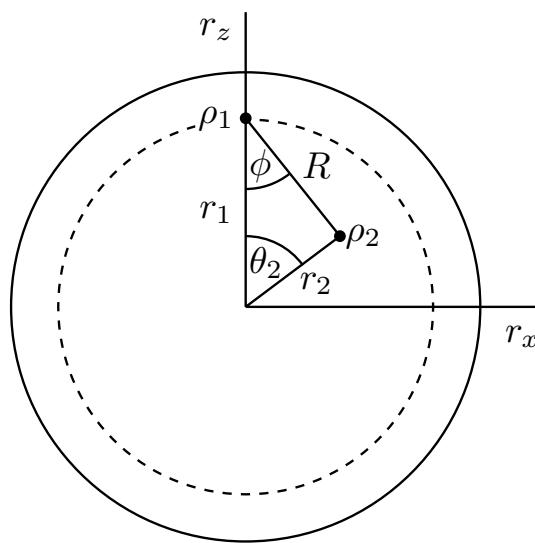


Figure 1. Parametrization of real states: r_1 and r_2 are the Bloch radii measured from the origin, θ_2 is the polar angle relative to the z axis ($\theta_1 = 0$). R is the Euclidean distance of the states ρ_1 and ρ_2 , and ϕ is the angle between r_1 and R .

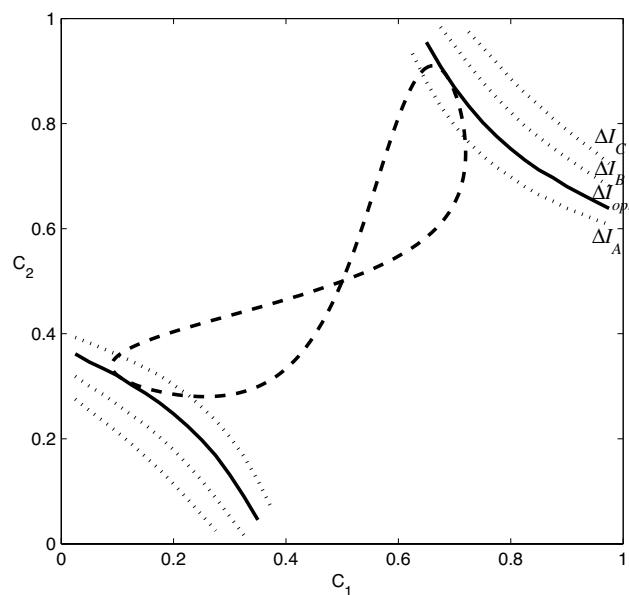


Figure 2. The constraint Eq. (30) (dashed line) and the information gain Eq. (28) (dotted lines) plotted together in the C_1, C_2 plane, for various fixed suboptimal values of the information gain ($\Delta I_A = 0.18$, $\Delta I_B = 0.3$, $\Delta I_C = 0.37$). The solid line corresponds to the maximal information gain, $\Delta I_{\max} = 0.23129$. Optimal values of C_1 and C_2 are the coordinates of the point where the solid line is tangent to the dashed curve (note that there are two sets of optimal solutions). The values of the parameters are $\eta_1 = \eta_2 = 1/2$, $r_1 = 0.9$, $\theta_1 = 0$, and $r_2 = 0.5$, $\theta_2 = \pi/4$.

using them in Eq. (28) leads to $\Delta I_{\min} = 0$. As noted before, Eq. (28) also gives a relation between C_1 and C_2 for a fixed value of the information gain ΔI . When plotted in the $C_1 - C_2$ plane, it exhibits two disjoint segments that are related by the reflection symmetry about the $C_1 + C_2 = 1$ line. The optimal measurement corresponds to the points where the solid line, Eq. (28), is tangent to the dashed line, Eq. (30). It can be seen that there are two sets of solutions, related by the same symmetry. Feasible measurements are in the region bounded by the two solid lines, yielding a value ΔI in the range $\Delta I_{\max} > \Delta I > \Delta I_{\min} = 0$ as we approach, from either of the boundaries, the point $C_1 = \eta_1$ and $C_2 = \eta_2$ where $\Delta I_{\min} = 0$. It should also be noted that, as a consequence of Eq. (21), knowledge of λ_1 and λ_2 , the eigenvalues of the transformed states in Eq. (16), is sufficient to find the optimum measurement.

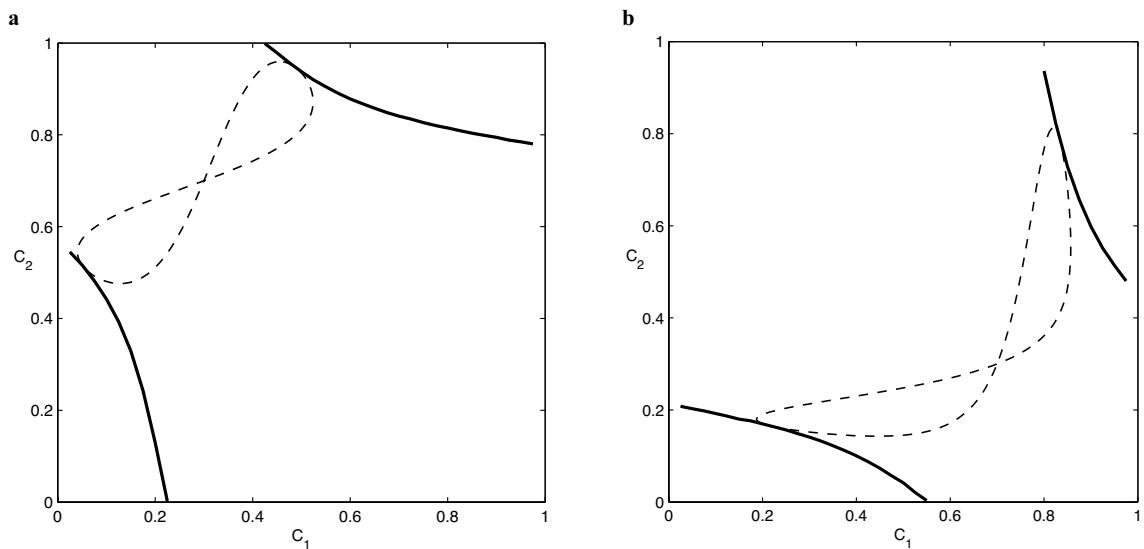


Figure 3. The constraint Eq. (30) (dashed line) and the information gain Eq. (28) (solid line) plotted together in the C_1, C_2 plane. The optimal values of the confidences at the upper points of tangency of the curves are (a) $C_1 = 0.4879$ and $C_2 = 0.9469$, (b) $C_1 = 0.7867$ and $C_2 = 0.8338$. The corresponding maximal information gains are (a) $\Delta I_{\max} = 0.2111$ and (b) $\Delta I_{\max} = 0.1842$. The parameters of the states are (a) $\eta_1 = 0.3, \eta_2 = 0.7$, and (b) $\eta_1 = 0.7, \eta_2 = 0.3$. Other parameters for both figures are $r_1 = 0.9, \theta_1 = 0$, and $r_2 = 0.5, \theta_2 = \pi/4$.

In summary, there are two key points of the geometric approach to optimization. First, the constraint (30), linking C_1 and C_2 , restricts us to a curve in the $C_1 - C_2$ plane, and the maximum of the information gain has to be found along this curve. Second, for a fixed value of ΔI (such that $0 < \Delta I < 1$), the expression for the information gain, Eq. (28), also corresponds to a curve in the same plane. If we choose ΔI too large the two curves may not have any common points. For intermediate values the two curves may have more than one common point. The maximal value ΔI_{\max} of the information gain is the one for which the two curves become tangent. Geometrically, it corresponds to the unique value of ΔI for which the $\Delta I(C_1, C_2)$ curve becomes tangent to the constraint. This can still happen for more than one point and the coordinates of these points, C_1 and C_2 , all correspond to optimal measurements, however, the value of ΔI_{\max} is unique. The actual value can be found numerically. Then we substitute the C_i values corresponding to the tangent points back to (25) and (26) and, using Eqs. (20), (15) and (14), we arrive at the POVM(s) which yield(s) the maximum information about the system (MI POVM).

Comparison of the MI and ME strategies

Although initially it has been assumed that the minimum error and the maximum information measurements coincide, a careful numerical study revealed that, in general, they are different⁴⁹. Therefore, in the following we present a systematic study of how these two measurements compare. Anticipating the results, we find that the ME and MI measurements coincide for the case of two pure states and, generally, for the case when the two states have the same Bloch radius ($r_1 = r_2$) or they are on the same diagonal of the Bloch disk.

In the ME strategy one is looking for measurement operators Π_i that maximize the expression

$$P_S = \eta_1 \text{Tr}(\rho_1 \Pi_1) + \eta_2 \text{Tr}(\rho_2 \Pi_2). \quad (32)$$

P_S is the average probability of correctly guessing the input state, aided with the measurement. Introducing

$$\Lambda = \eta_1 \rho_1 - \eta_2 \rho_2, \quad (33)$$

and using $\Pi_1 + \Pi_2 = \hat{\mathbb{I}}$, we can write Eq. (32) in a more compact form,

$$P_S = \eta_2 + \text{Tr}(\Pi_1 \Lambda). \quad (34)$$

This expression is clearly maximal if Π_1 is the projector to the subspace of Λ belonging to positive eigenvalues. Hence, Π_2 will be the projector to the subspace belonging to negative eigenvalues.

In order to show the relationship between the ME and MI measurements, we first write Eq. (32) in the form

$$P_S = P_1 C_1 + P_2 C_2, \quad (35)$$

which can be obtained if we divide and multiply the first term on the right-hand side of Eq. (32) by P_1 and the second by P_2 and use the definition Eq. (8) for the click probabilities and Eq. (9) for the confidences. It has been observed earlier²⁸ that the minimum error measurement is also the one that maximizes the average confidence. Furthermore, the relation (27) between the click probabilities and the confidences still holds, so finally, we can write P_S as

$$P_S = \frac{C_2 - \eta_2}{C_1 + C_2 - 1} C_1 + \frac{C_1 - \eta_1}{C_1 + C_2 - 1} C_2 \quad (36)$$

This is the cost function to be maximized in the ME measurement. When we compare this to the cost function, Eq. (28), of the MI measurement, we see that there are similarities, e.g., the click probabilities are the same and, in addition, the constraint (29) is also the same for both measurements. Apart from these similarities, the two cost functions are rather different. So, there is no *a priori* reason for the ME and MI measurements to be the same. We will show that they are generally different indeed except for special cases when the input state has some intrinsic symmetry.

What we know is that they are both projective measurements in the xz plane. After determining their respective orientations, the difference between the ME and MI strategies can be characterized by the angle δ between their projectors

$$\delta = \arccos \sqrt{\text{Tr} \Pi_1^{\text{ME}} \Pi_1^{\text{MI}}}. \quad (37)$$

Next, we study the dependence of δ on the structure of the input states. The MI and ME POVMs are determined by using the methods presented previously. Without loss of generality, we will choose the parameters of ρ_1 as r_1 and $\theta_1 = 0$ and study the dependence of δ on the parameters of ρ_2 , that is, on the polar angle θ_2 and the Bloch radius r_2 in the Bloch disk, introduced in Fig. 1. In this paper, we focus on the case of equal priors, $\eta_1 = \eta_2 = 1/2$; the case of arbitrary priors will be addressed in a separate publication. Note that, in the following figures, δ is measured in degrees, while the polar angles are measured in radians.

We consider the general case of two mixed states. Without loss of generality, we assume that $r_2 \leq r_1$. Figure 4a shows the difference δ between the two measurement strategies as a function of the Bloch radius r_2 and the polar angle θ_2 characterizing the state ρ_2 . Figure 4b shows the same quantity as a function of the polar angle θ_2 , for representative values of the Bloch radius r_2 . In these figures ρ_1 is a fixed mixed state with $r_1 = 0.8$, $\theta_1 = 0$. From these figures one can deduce that the two strategies coincide (that is, $\delta = 0$) only in the case when the two mixed states have the same Bloch radius ($r_1 = r_2$) or they are along the same diagonal of the Bloch disk ($\theta_2 = 0, \pi$). These rules are valid for any value of r_1 . Accordingly, the ME and MI strategies coincide for two pure states ($r_1 = r_2 = 1$). For a given r_2 , δ exhibits maximum, δ_{max} , for a certain value of the polar angle $\theta_2 = \theta_2^{\text{max}}$. Figure 5 displays the polar angle θ_2^{max} , corresponding to the maximum difference $\delta_{\text{max}}(r_2)$ between the two strategies, as a function of the Bloch radius r_2 , for the pure state $\rho_1 = |0\rangle\langle 0|$ ($r_1 = 1, \theta_1 = 0$). In this figure, the value of the polar angle θ_2^{max} decreases linearly, except for values of r_2 close to those of the other Bloch radius r_1 . Note that the value $r_2 = 1$ for which the two strategies coincide is excluded from the domain of the function $\theta_2^{\text{max}}(r_2)$. Figure 4b shows that by increasing the Bloch radius r_2 toward the other radius r_1 the value of the maximal difference $\delta_{\text{max}}(r_2)$ between the ME and MI detection strategies also increases. We have found that the function $\delta_{\text{max}}(r_2)$ practically reaches its saturated value δ_{max} within a precision 0.01 when $r_1 - r_2 \lesssim 0.001$. Figure 6a presents the maximum difference δ_{max} between the ME and MI strategies as a function of the purity r_1^2 of the state ρ_1 . The state ρ_2 is always close to ρ_1 ($r_1 - r_2 = 0.001$), that is, $R \rightarrow 0$ in Fig. 1. The figure shows that increasing the purity r_1^2 of state ρ_1 , δ_{max} grows nearly exponentially, reaching its maximum when ρ_1 is pure. Figure 6b shows the POVMs for this case in the computational basis. The ME POVM is aligned symmetrically around the states. For the MI strategy, however, one of the POVM elements virtually coincides with the pure state, and the other one is perpendicular to rule it out. Although the information provided by the measurements

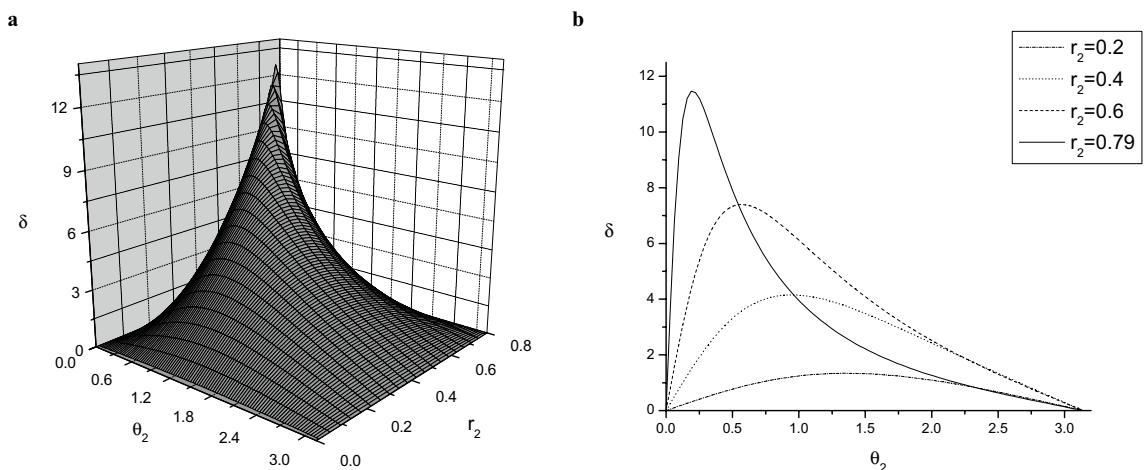


Figure 4. The difference δ between the ME and MI measurement strategies for the case of two mixed states as a function of (a) the Bloch radius r_2 and the polar angle θ_2 characterizing the state ρ_2 , and (b) the polar angle θ_2 , for representative values of r_2 . ρ_1 is a fixed mixed state at $r_1 = 0.8, \theta_1 = 0$. The difference δ is measured in degrees, while the polar angle θ_2 is measured in radians.

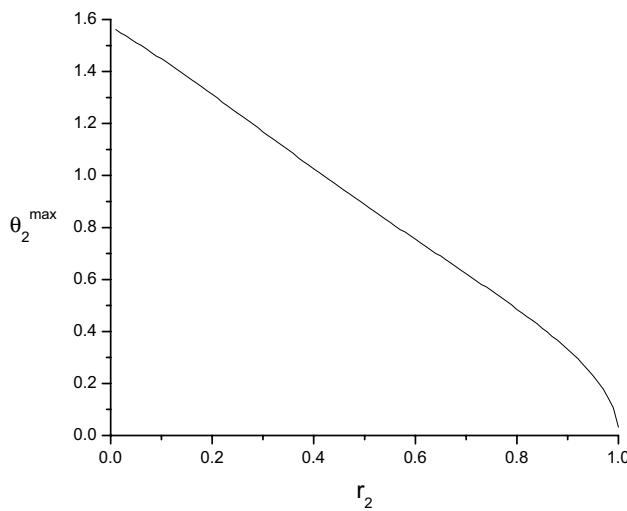


Figure 5. The polar angle θ_2^{\max} for which the difference δ between the ME and MI strategies is maximal, as a function of the Bloch radius r_2 for pure state $\rho_1 = |0\rangle\langle 0|$ ($r_1 = 1, \theta_1 = 0$). The polar angle θ_2^{\max} is measured in radians.

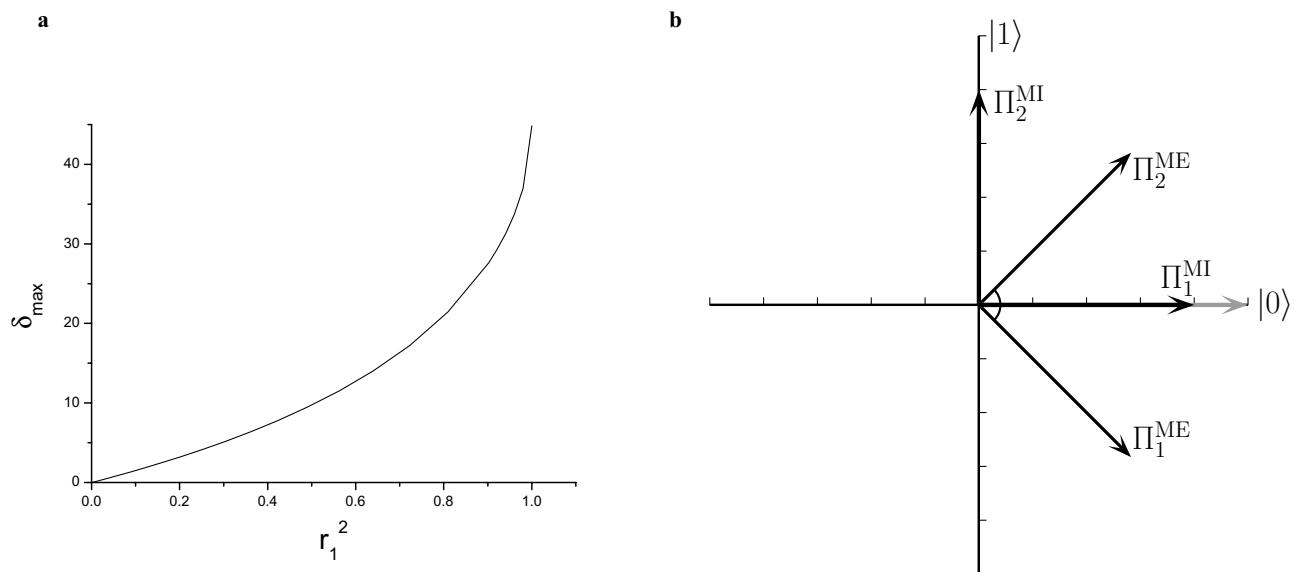


Figure 6. (a) The maximum difference δ_{\max} between the ME and MI strategies as a function of the purity r_1^2 of ρ_1 . The state ρ_2 is always close to ρ_1 ($r_1 - r_2 = 0.001$). (b) The POVMs of the MI (thick vectors) and ME (thin vectors) strategies for two states which are very close to each other on the Bloch disk ($\rho_1 = |0\rangle\langle 0|$, $r_2 = 0.999$, $\theta_2 = 0.001$).

vanishes when the mixed state ρ_2 is in the close vicinity of the pure state ρ_1 , we have found that the MI POVM brings more than twice as much information as the ME POVM for the presented case.

Discussion

We have developed an analytic method, supplemented by a geometric approach to optimization, for finding the measurement that yields the maximum information gain about a qubit system that is prepared in one of two known states with given prior probabilities. We have determined the parameters of the POVM of the maximum information gain measurement for two arbitrary (pure or mixed) states, prepared with equal prior probabilities, building on previous results that the optimal measurement is always a standard von Neumann measurement for this case. We have compared the maximum information measurement to the minimum error one, and showed that the POVMs of the two measurement strategies coincide exactly only when both states have the same Bloch radius or they are along the same diagonal of the Bloch disk. The case of general priors will be addressed in a subsequent publication.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 19 January 2024; Accepted: 16 May 2024

Published online: 24 May 2024

References

1. Nielsen, M. & Chuang, I. *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
2. Holevo, A. Statistical decision theory for quantum systems. *J. Multivar. Anal.* **3**, 337–394. [https://doi.org/10.1016/0047-259X\(73\)90028-6](https://doi.org/10.1016/0047-259X(73)90028-6) (1973).
3. Helstrom, C. *Quantum Detection and Estimation Theory*. Mathematics in Science and Engineering : A series of monographs and textbooks (Academic Press, 1976).
4. Yuen, H., Kennedy, R. & Lax, M. Optimum testing of multiple hypotheses in quantum detection theory. *IEEE Trans. Inf. Theory* **21**, 125–134. <https://doi.org/10.1109/TIT.1975.1055351> (1975).
5. Guerrini, S., Win, M. Z., Chiari, M. & Conti, A. Quantum discrimination of noisy photon-added coherent states. *IEEE J. Sel. Areas Inf. Theory* **1**, 469–479. <https://doi.org/10.1109/JSIT.2020.3012944> (2020).
6. Nakahira, K., Usuda, T. S. & Kato, K. Finding optimal solutions for generalized quantum state discrimination problems. *IEEE Trans. Inf. Theory* **63**, 7845–7856. <https://doi.org/10.1109/TIT.2016.2644666> (2017).
7. Ivanovic, I. How to differentiate between non-orthogonal states. *Phys. Lett. A* **123**, 257–259. [https://doi.org/10.1016/0375-9601\(87\)90222-2](https://doi.org/10.1016/0375-9601(87)90222-2) (1987).
8. Dieks, D. Overlap and distinguishability of quantum states. *Phys. Lett. A* **126**, 303–306. [https://doi.org/10.1016/0375-9601\(88\)90840-7](https://doi.org/10.1016/0375-9601(88)90840-7) (1988).
9. Peres, A. How to differentiate between non-orthogonal states. *Phys. Lett. A* **128**, 19. [https://doi.org/10.1016/0375-9601\(88\)91034-1](https://doi.org/10.1016/0375-9601(88)91034-1) (1988).
10. Jaeger, G. & Shimony, A. Optimal distinction between two non-orthogonal quantum states. *Phys. Lett. A* **197**, 83–87. [https://doi.org/10.1016/0375-9601\(94\)00919-G](https://doi.org/10.1016/0375-9601(94)00919-G) (1995).
11. Takeoka, M., Krovi, H. & Guha, S. Achieving the Holevo capacity of a pure state classical-quantum channel via unambiguous state discrimination. In *IEEE Int. Symp. Inf. Theory*, 166–170, 2013. <https://doi.org/10.1109/ISIT.2013.6620209> (2013).
12. Eldar, Y. A semidefinite programming approach to optimal unambiguous discrimination of quantum states. *IEEE Trans. Inf. Theory* **49**, 446–456. <https://doi.org/10.1109/TIT.2002.807291> (2003).
13. Croke, S., Andersson, E., Barnett, S. M., Gilson, C. R. & Jeffers, J. Maximum confidence quantum measurements. *Phys. Rev. Lett.* **96**, 070401. <https://doi.org/10.1103/PhysRevLett.96.070401> (2006).
14. Barnett, S. M. & Croke, S. Quantum state discrimination. *Adv. Opt. Photon.* **1**, 238–278. <https://doi.org/10.1364/AOP.1.000238> (2009).
15. Bergou, J. A. Discrimination of quantum states. *J. Mod. Opt.* **57**, 160–180. <https://doi.org/10.1080/09500340903477756> (2010).
16. Bae, J. & Kwek, L.-C. Quantum state discrimination and its applications. *J. Phys. A: Math. Theor.* **48**, 083001. <https://doi.org/10.1088/1751-8113/48/8/083001> (2015).
- 17.Chefles, A. Unambiguous discrimination between linearly independent quantum states. *Phys. Lett. A* **239**, 339–347. [https://doi.org/10.1016/S0375-9601\(98\)00064-4](https://doi.org/10.1016/S0375-9601(98)00064-4) (1998).
18. Bergou, J., Feldman, E. & Hillery, M. Extracting information from a qubit by multiple observers: Toward a theory of sequential state discrimination. *Phys. Rev. Lett.* **111**, 100501. <https://doi.org/10.1103/PhysRevLett.111.100501> (2013).
19. Namkung, M. & Kwon, Y. Optimal sequential state discrimination between two mixed quantum states. *Phys. Rev. A* **96**, 022318. <https://doi.org/10.1103/PhysRevA.96.022318> (2017).
20. Namkung, M. & Kwon, Y. Analysis of optimal sequential state discrimination for linearly independent pure quantum states. *Sci. Rep.* **8**, 6515 (2018).
21. Namkung, M. & Kwon, Y. Sequential state discrimination of coherent states. *Sci. Rep.* **8**, 16915 (2018).
22. Namkung, M. & Kwon, Y. Generalized sequential state discrimination for multiparty QKD and its optical implementation. *Sci. Rep.* **10**, 8247 (2020).
23. Bergou, J. A., Futschik, U. & Feldman, E. Optimal unambiguous discrimination of pure quantum states. *Phys. Rev. Lett.* **108**, 250502. <https://doi.org/10.1103/PhysRevLett.108.250502> (2012).
24. Eldar, Y., Megretski, A. & Verghese, G. Optimal detection of symmetric mixed quantum states. *IEEE Trans. Inf. Theory* **50**, 1198–1207. <https://doi.org/10.1109/TIT.2004.828070> (2004).
25. Barnett, S. M. Minimum-error discrimination between multiply symmetric states. *Phys. Rev. A* **64**, 030303. <https://doi.org/10.1103/PhysRevA.64.030303> (2001).
26. Deconinck, M. E. & Terhal, B. M. Qubit state discrimination. *Phys. Rev. A* **81**, 062304. <https://doi.org/10.1103/PhysRevA.81.062304> (2010).
27. Bae, J. Structure of minimum-error quantum state discrimination. *New J. Phys.* **15**, 073037. <https://doi.org/10.1088/1367-2630/15/7/073037> (2013).
28. Herzog, U. Discrimination of two mixed quantum states with maximum confidence and minimum probability of inconclusive results. *Phys. Rev. A* **79**, 032323. <https://doi.org/10.1103/PhysRevA.79.032323> (2009).
29. Peres, A. *Quantum Theory: Concepts and Methods. Fundamental Theories of Physics* (Springer, 1995).
30. Holevo, A. S. Bounds for the quantity of information transmitted by a quantum communication channel. *Probl. Peredachi Inf.* **9**, 3–11 (1973).
31. Yuen, H. P. & Ozawa, M. Ultimate information carrying limit of quantum systems. *Phys. Rev. Lett.* **70**, 363–366. <https://doi.org/10.1103/PhysRevLett.70.363> (1993).
32. Caves, C. M. & Drummond, P. D. Quantum limits on bosonic communication rates. *Rev. Mod. Phys.* **66**, 481–537. <https://doi.org/10.1103/RevModPhys.66.481> (1994).
33. Bagan, E., Baig, M., Brey, A., Muñoz Tapia, R. & Tarrach, R. Optimal strategies for sending information through a quantum channel. *Phys. Rev. Lett.* **85**, 5230–5233. <https://doi.org/10.1103/PhysRevLett.85.5230> (2000).
34. Schumacher, B. Quantum coding. *Phys. Rev. A* **51**, 2738–2747. <https://doi.org/10.1103/PhysRevA.51.2738> (1995).
35. Schumacher, B. & Westmoreland, M. D. Optimal signal ensembles. *Phys. Rev. A* **63**, 022308. <https://doi.org/10.1103/PhysRevA.63.022308> (2001).
36. Massar, S. & Popescu, S. Amount of information obtained by a quantum measurement. *Phys. Rev. A* **61**, 062303. <https://doi.org/10.1103/PhysRevA.61.062303> (2000).
37. Smith, G. Quantum channel capacities. In *2010 IEEE Information Theory Workshop*, 1–5, <https://doi.org/10.1109/CIG.2010.559251> (2010).

38. Winter, A. & Yang, D. Potential capacities of quantum channels. *IEEE Trans. Inf. Theory* **62**, 1415–1424. <https://doi.org/10.1109/TIT.2016.2519920> (2016).
39. Ding, D., Pavlichin, D. S. & Wilde, M. M. Quantum channel capacities per unit cost. *IEEE Trans. Inf. Theory* **65**, 418–435. <https://doi.org/10.1109/TIT.2018.2854747> (2019).
40. Wang, X. Pursuing the fundamental limits for quantum communication. *IEEE Trans. Inf. Theory* **67**, 4524–4532. <https://doi.org/10.1109/TIT.2021.3068818> (2021).
41. Hayashi, M. Analytical calculation formulas for capacities of classical and classical-quantum channels. In *IEEE Int. Symp. Inf. Theory (ISIT)*, 91–96, 2022. <https://doi.org/10.1109/ISIT50566.2022.9834600> (2022).
42. Davies, E. Information and quantum measurement. *IEEE Trans. Inf. Theory* **24**, 596–599. <https://doi.org/10.1109/TIT.1978.1055941> (1978).
43. Sasaki, M., Barnett, S. M., Jozsa, R., Osaki, M. & Hirota, O. Accessible information and optimal strategies for real symmetrical quantum sources. *Phys. Rev. A* **59**, 3325–3335. <https://doi.org/10.1103/PhysRevA.59.3325> (1999).
44. Levitin, L. B. Optimal Quantum Measurements for Two Pure and Mixed States. In *Quantum Communications and Measurement* (eds Belavkin, V. *et al.*) 439 (Springer, 1995).
45. Shor, P. W. On the number of elements needed in a POVM attaining the accessible information (2000). [arXiv:quant-ph/0009077](https://arxiv.org/abs/quant-ph/0009077), quant-ph/0009077.
46. Keil, A. Proof of the orthogonal measurement conjecture for qubit states (2008). [arXiv:0809.0232](https://arxiv.org/abs/0809.0232), 0809.0232.
47. Fuchs, C. A. & Peres, A. Quantum-state disturbance versus information gain: Uncertainty relations for quantum information. *Phys. Rev. A* **53**, 2038–2045. <https://doi.org/10.1103/PhysRevA.53.2038> (1996).
48. Ban, M., Yamazaki, K. & Hirota, O. Accessible information in combined and sequential quantum measurements on a binary-state signal. *Phys. Rev. A* **55**, 22–26. <https://doi.org/10.1103/PhysRevA.55.22> (1997).
49. Režáček, J., Englert, B.-G. & Kaszlikowski, D. Iterative procedure for computing accessible information in quantum communication. *Phys. Rev. A* **71**, 054303. <https://doi.org/10.1103/PhysRevA.71.054303> (2005).
50. Jozsa, R., Robb, D. & Wootters, W. K. Lower bound for accessible information in quantum mechanics. *Phys. Rev. A* **49**, 668–677. <https://doi.org/10.1103/PhysRevA.49.668> (1994).
51. Fuchs, C. A. & Caves, C. M. Ensemble-dependent bounds for accessible information in quantum mechanics. *Phys. Rev. Lett.* **73**, 3047–3050. <https://doi.org/10.1103/PhysRevLett.73.3047> (1994).
52. Jain, R. & Nayak, A. Accessible versus Holevo information for a binary random variable (2007). [arXiv:quant-ph/0603278](https://arxiv.org/abs/quant-ph/0603278), quant-ph/0603278.

Acknowledgements

This research was supported by the National Research, Development and Innovation Office, Hungary (Projects No. TKP 2021-NVA-04, TKP2021-EGA-17). The research of J. A. B. was supported by NSF Grant No. FET-2106447 and by the DEVCOM Army Research Laboratory and was accomplished under Cooperative Agreement No. W911NF-20-2-0097.

Author contributions

All authors contributed to the development of the idea raised by J.A.B., and to the preparation of the manuscript. Á.V. performed the numerical calculations.

Funding

Open access funding provided by HUN-REN Wigner Research Centre for Physics.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024