POVM Design for Quantum State Discrimination

Qi Ding*, Catherine Medlock†, Alan Oppenheim‡
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139–4301
Email: *qding@mit.edu, †catherine.lockton@ll.mit.edu, ‡avo@mit.edu

***Abstract*—A qubit pure state can be specified by a point on the Bloch sphere, and similarly certain quantum measurements can be specified by M points on a sphere which we refer to as an Etro sphere. A key objective of this paper is to compare the performance for discriminating between two qubit pure states using POVM designs based on the distribution of M points on the Etro sphere. We specifically address the case where the alignment of the M points relative to the coordinates of the sphere is unknown. Of particular interest is the insensitivity of the POVM designs as measured by the difference between the maximum and minimum probabilities of error over all alignments. We consider distributions of the M points corresponding to Platonic solids as well as optimal distributions with respect to several criteria including, for example, maximum nearest neighbor distance and minimum Riesz s-energy. We provide evidence through simulation of various performance tradeoffs such as the tradeoff between stability and best case performance.**

***Index Terms*—Quantum state discrimination, POVM design, quantum measurement, Platonic solids, Bloch sphere, Bloch vectors, qubit**

I. Introduction

We consider measurement design for discriminating between one of two possible quantum states of a qubit as arises for example in quantum optical communications [1]–[3] and in scenarios of quantum device testing. We consider that an ensemble of L qubits has been prepared in one of two pure states by one of two apparatuses. Our focus relates to deciding which of the two apparatuses was used. Each pure state of a qubit can be represented by a three-dimensional unit vector or equivalently by a point on the Bloch sphere [4]. We consider the case where the relative angle between the two states is known, but their alignment relative to the coordinates of the Bloch sphere is not. It is desirable to choose a quantum measurement whose discrimination performance is stable over all possible rotations of the Bloch sphere coordinates relative to the possible qubit states.

A quantum measurement can be described through an associated positive operator-valued measure (POVM). The POVMs that we consider can be fully specified by a set of M three-dimensional vectors of equal norm $\sqrt{2}/M$, with each POVM element corresponding to a unique point on a sphere of radius $\sqrt{2}/M$, defined in Section III as an Etro sphere. Considerable work by others has focused on POVM design based on inscribing within an Etro sphere one of the five Platonic solids [5]–[9]. In this paper we propose and compare the performance of alternative POVM designs based on several different strategies for arranging M points on an Etro sphere. This and related work is also discussed in [11], [12].

II. Quantum Mechanics Preliminaries

The postulates of quantum mechanics give rise to such problems where the system operate fundamentally differently than a classical system. The two key postulates in [4] which are essential to the formulation of our POVM design strategy are briefly paraphrased below.

A. Quantum State Postulate

The state of an isolated quantum system can be represented by a positive semi-definite unit trace Hermitian operator, the density operator, denoted by ρ, that acts on a complex Hilbert space $\mathcal{H}$ of dimension N. ρ can be written in terms of its eigenbases as

$$\rho = \sum_{i=0}^{N-1} p_i \left|\psi_i\right\rangle \left\langle\psi_i\right| \tag{1}$$

where $\{p_i\}$ sum to 1 and $\{|\psi_i\rangle\}$ are orthonormal. The density operator generally describes a mixed state, i.e. an ensemble of pure states. In this paper, we specifically deal only with pure states in which case the density operator has the form

$$\rho = \left|\psi\right\rangle \left\langle\psi\right|. \tag{2}$$

The vector $|\psi\rangle$ is often referred to as a pure state vector. In this paper, we restrict the dimension of the Hilbert space $\mathcal{H}$ to be $N = 2$.

B. Quantum Measurement Postulate

A quantum measurement with M possible outcomes is represented by a collection of M operators $\{A_k, k = 1, 2, ..., M\}$, with each value of the index k corresponding to a possible measurement outcome. The probability that the kth measurement outcome occurs is

$$p(k) = \mathrm{Tr}(A_k^\dagger A_k \rho) \tag{3}$$

where † denotes Hermitian adjoint and Tr denotes the trace operator.

Defining $E_k = A_k^\dagger A_k$, it follows from the postulates of quantum mechanics that the collection of operators $\{E_k\}$ satisfy the properties required of a POVM, i.e. E_k is a positive semi-definite operator such that $\sum_{k=1}^{M} E_k = I$ and $p(m) = \langle\psi| E_m |\psi\rangle$. These constraints on the set of operators $\{E_k\}$ are sufficient to determine the probabilities of different

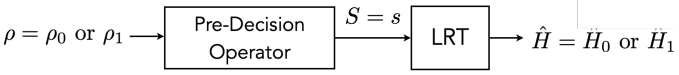

Fig. 1: Framework for qubit state discrimination.

measurement outcomes. The set $\{E_k\}$ is called a POVM and each E_k is referred to as a POVM element.

III. Quantum Binary State Discrimination

In this section, we first present the basic framework and notation for the qubit state discrimination problem, and then summarize the Bloch sphere and Etro spheres, which are geometric representations of density operators and POVM operators.

A. Discrimination Strategy

The two stages of the discrimination process are shown in Fig. 1. The input to the system consists of L identical qubits all prepared in the same quantum state. The corresponding density operator ρ is equal to ρ_0 or ρ_1 with some known prior probabilities. Our objective is to distinguish between these two possibilities. To this end the L qubits are first processed by a pre-decision operator which performs the same M-outcome quantum measurement on each qubit. The M-element POVM associated with the measurement will be denoted by $\{E_k, k = 1, 2, \ldots, M\}$. The measurement outcomes are assembled into a relative frequency vector

$$\vec{s} = \begin{bmatrix} n_1/L & n_2/L & \cdots & n_M/L \end{bmatrix}^T, \tag{4}$$

where $n_k, k = 1, 2, ..., M$ denotes the number of occurrences of the kth measurement outcome. In the second stage of the strategy, a decision of $\hat{H} = H_0$ (indicating that $\rho = \rho_0$) or $\hat{H} = H_1$ (indicating that $\rho = \rho_1$) is made based on the relative frequency vector obtained. We assume that the objective is to minimize probability of error, in which case it is well-known that the optimal decision rule is a likelihood ratio test (LRT) with a threshold equal to the ratio of the prior probabilities [10].

The use of an LRT to minimize probability of error is based on the viewpoint that the observed relative frequency vector $\vec{s}$ can be thought of as one realization of a vector-valued random variable $\vec{S}$. The conditional distributions of $\vec{S}$ given that $\rho = \rho_0$ or ρ_1 are multinomial distributions parameterized by the value of L and by the probabilities given by Equation (3) with $\rho = \rho_0$ or ρ_1. Exact expressions for the conditional distributions can be found in Appendix A of [11].

B. Bloch Sphere and Etro Spheres

Our focus is on single-qubit systems for which ρ is an operator on a two-dimensional complex Hilbert space. Typically, the Bloch sphere is used to geometrically represent the quantum state of a qubit [4]. The density operator ρ that represents a pure state can be written as

$$\rho = \frac{1}{\sqrt{2}}\frac{I}{\sqrt{2}} + r_1\frac{\sigma_1}{\sqrt{2}} + r_2\frac{\sigma_2}{\sqrt{2}} + r_3\frac{\sigma_3}{\sqrt{2}} \tag{5}$$

where I is the identity matrix and $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{6}$$

When ρ represents a pure state, then $r_1^2 + r_2^2 + r_3^2 = 1/\sqrt{2}$ [4]. The normalized identity and Pauli matrices $\{I/\sqrt{2}, \sigma_1/\sqrt{2}, \sigma_2/\sqrt{2}, \sigma_3/\sqrt{2}\}$ form a common orthonormal basis for a two-dimensional complex Hilbert space $\mathcal{H} = \mathbb{C}^2$.

It is thus sufficient to specify the density operator of a pure state by a three-dimensional vector $\vec{r} = [r_1, r_2, r_3]^T$ with $\|\vec{r}\| = 1/\sqrt{2}$. In this way, the density operator of a pure state always corresponds to a three-dimensional vector from the origin to a point on a sphere with radius $1/\sqrt{2}$. $\vec{r}$ is referred to as the Bloch vector of the state.

In our POVM design for state discrimination each POVM element can be described in the same basis as the density operators, i.e. each POVM element E_k can be expressed as

$$E_k = \frac{\text{Tr}(E_k)}{\sqrt{2}}\frac{I}{\sqrt{2}} + c_{k1}\frac{\sigma_1}{\sqrt{2}} + c_{k2}\frac{\sigma_2}{\sqrt{2}} + c_{k3}\frac{\sigma_3}{\sqrt{2}}, \quad k = 1, 2, ..., M \tag{7}$$

We further restrict all of the M POVM elements to be rank one and have equal trace $2/M$. Furthermore with the POVM constraints as described in Section III, we have

$$\sum_{k=1}^{M} \text{Tr}(E_k) = 2 \tag{8a}$$

$$0 \leq \text{Tr}(E_k) \leq 2, \quad k = 1, 2, ..., M \tag{8b}$$

$$\sum_{k=1}^{M} \vec{c}_k = \vec{0} \tag{8c}$$

With these constraints each element of a specific POVM can be represented geometrically as a three dimensional vector $\vec{c}_k = [c_{k1}, c_{k2}, c_{k3}]^T$ with $\|\vec{c}_k\| = \text{Tr}(E_k)/\sqrt{2} = \sqrt{2}/M$ from the origin to a point on a sphere with radius $\sqrt{2}/M$. The sphere associated with any specific POVM will be referred to as an Etro sphere to identify the fact that all POVM elements represented by points on the sphere are equal trace and rank one and correspondingly the design of an M element POVM corresponds to choosing an appropriate set of M points on the sphere.

IV. Problem Statement

We consider the design of POVMs for discriminating between two possible pure states given a collection of L qubits that have all been identically prepared in one of two states. The Bloch vectors of the two pure states are $\vec{r}_i$ for $i \in \{0, 1\}$. The angle α between the $\{\vec{r}_i\}$, illustrated in Fig. 2, takes values between 0 and π. The prior probabilities are denoted by $P(\vec{r}_i)$ for $i \in \{0, 1\}$. As described in Section III, we discriminate between the two hypotheses by performing L identical measurements, one on each of the L qubits and all with the same POVM. A final decision is made by performing an LRT on the vector of relative frequencies of the measurement outcomes. As identified in Section III, the design

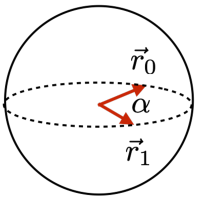

Fig. 2: Bloch vectors of ρ_0 and ρ_1.

of the POVM can be framed as a problem of appropriately distributing M points on the POVM Etro sphere. The objective is to compare the performance of POVMs constructed from different sets of $\{\vec{c}_k\}$ and equivalently different distributions of M points on the POVM Etro sphere, according to their minimum and maximum probabilities of error (P_e) over all possible orientations of the $\{\vec{c}_k\}$.

In the design of the POVM, we consider as an important goal the insensitivity of the performance in discrimination to rotation of the Bloch sphere relative to the qubit Bloch vectors, which for example could result from some error in the apparatuses. In particular, the angle α between the $\{\vec{r}_i\}$ is assumed known but the overall orientation of $\{\vec{r}_i\}$ on the Bloch sphere is unknown. Intuition suggests that designing the POVM to maximally distribute points on the Etro sphere will tend to reduce the variation of performance over all possible orientations.

Various approaches to and criteria for distributing M points on a sphere have been reported in the literature [13]–[15]. We first consider distributions of points that correspond to the vertices of a Platonic solid. In addition, we consider the distribution of the points according to minimizing Riesz s-energy for a given value of M with the constraint that $\{\vec{c}_k\}$ must sum to zero. In three dimensions the s-energy of a set of M points on the unit sphere is defined as follows. Assume that $\{\vec{c}_k\}$ is the set of unit vectors that extend from the origin to each of the M points and let $\|\vec{c}_j - \vec{c}_k\|$ denote the distance between $\vec{c}_j$ and $\vec{c}_k$. In this paper Euclidean distance is used. The s-energy of the points is

$$E(s) = \begin{cases} \sum_{1\le j<k\le M} \log \|\vec{c}_j - \vec{c}_k\|^{-1}, & \text{if } s = 0 \\ \sum_{1\le j<k\le M} \|\vec{c}_j - \vec{c}_k\|^{-s}, & \text{if } s > 0 \end{cases} \tag{9}$$

Minimizing $E(0)$ is equivalent to maximizing the product of distances between points. Minimizing $E(1)$ is equivalent to minimizing the electric potential energy of a system with point charges located at points specified by $\{\vec{c}_k\}$. As $s \to \infty$, only the two closest points contribute to the sum and minimizing $E(s)$ corresponds to maximizing nearest neighbor distance.

In the simulation results presented in Section V, we also consider optimal numerical solutions with respect to maximum nearest neighbor distance, maximum convex hull volume, and minimum covering radius criteria [13], [16]–[19]. The latter two criteria are defined, respectively, as

$$\max_{\vec{c}_1,\vec{c}_2,\ldots,\vec{c}_M} \min_{1\le j<k\le M} \|\vec{c}_j - \vec{c}_k\| \tag{10a}$$

$$\min_{\vec{c}_1,\vec{c}_2,\ldots,\vec{c}_M} \max_{\vec{x}:\|\vec{x}\|=\sqrt{2}/M} \min_{1\le k\le M} \|\vec{x} - \vec{c}_k\| \tag{10b}$$

Note that the optimal numerical solutions are computed without imposing the constraint that $\{\vec{c}_k\}$ must sum to zero, although it is not surprising that many solutions are symmetric and thus naturally sum to zero or to a vector with a very small norm. To ensure that the condition of Equation (8c) is met, we append an extra error vector $\vec{\epsilon} = -\sum_k \vec{c}_k$ to $\{\vec{c}_k\}$ with corresponding POVM element E_ϵ. This additional element does not change the broad trend in our observations as we require $\|\vec{\epsilon}\| \le 10^{-8}$.

V. Simulations and Observations

A sampling of preliminary simulation results is shown in Tables I to III for collection size $L = 5$. The trends shown were observed more broadly than in the examples given here, but notably, there are some important observations that require further study. Note that table entries marked as "N/A" indicate that no Platonic solid is available for that value of M or that the optimal configuration of points do not satisfy the requirements required of a POVM. For fixed α and $P(\vec{r}_1)$, larger values of M often correspond to lower maximum P_e but higher minimum P_e over all possible orientations (see Table I). This tradeoff had been previously demonstrated using only point distributions corresponding to Platonic solids in [9]. For a fixed value of $M \in \{4, 6, 8, 12, 20\}$, the Platonic solid with M vertices is not necessarily the best arrangement of points in terms of its sensitivity to rotation. This can be seen, for example, in Table I for the case where $M = 6$. For fixed $P(\vec{r}_1)$, the increase in stability with M is more pronounced for smaller values of α (see Table II), which makes intuitive sense since smaller values of α correspond to Bloch vectors states that are more collinear and thus more sensitive to small changes in the orientation of the Bloch sphere. For fixed M, larger values of α and values of $P(\vec{r}_1)$ that are further from $1/2$ generally lead to lower minimum and maximum P_e (see Table III).

In addition to considering distributions of points that are optimal with respect to various criteria, we have also conducted experiments with distributions of points that have a range of s-energy values for $s = 0$, including but not limited to the minimum possible value. We require that each point distribution form a tight frame for $\mathbb{R}^3$ to facilitate more direct comparison with the Platonic solids, which all have this property. Our results suggest that for a given value of M, those distributions of points with lower values of $E(0)$ on average achieve lower maximum P_e at the cost of expecting a higher minimum P_e. The monotonic trend between $E(0)$ and sensitivity to rotation was observed for a large range of values of both α and $P(\vec{r}_1)$.

References

[1] C. W. Helstrom, J. W. S. Liu, and J. P. Gordon, "Quantum-mechanical communication theory," Proc. IEEE, vol. 58, no. 10, pp. 1578–1598, Oct. 1970.

[2] G. Cariolaro, Quantum communications. Switzerland: Springer, 2015, Part II: Quantum Communications.

[3] S. J. Van Enk, Juan I. Cirac, and Peter Zoller. "Ideal quantum communication over noisy channels: a quantum optical implementation." Physical Review Letters 78.22 (1997): 4293.

TABLE I: Minimum, maximum, and maximum minus minimum probabilities of error (P_e) for different distributions of M points on a sphere. The values below were generated using $\alpha = \pi/2$ and $P(\vec{r}_1) = 1/2$.

M	**Description**	**Min** P_e	**Max** P_e	**Difference**
4	Tetrahedron	0.076	0.214	0.138
	Max Convex Hull Vol	0.076	0.213	0.137
	Max N.N. Dist	0.076	0.214	0.138
	Min Covering Radius	0.077	0.211	0.135
	Min Riesz 0-energy	0.076	0.213	0.137
5	(No Platonic Solid)	N/A	N/A	N/A
	Max Convex Hull Vol	0.097	0.224	0.126
	Max N.N. Dist	0.101	0.221	0.120
	Min Covering Radius	0.092	0.223	0.131
	Min Riesz 0-energy	0.091	0.198	0.107
6	Octahedron	0.111	0.174	0.063
	Max Convex Hull Vol	0.111	0.174	0.063
	Max N.N. Dist	0.111	0.174	0.063
	Min Covering Radius	0.111	0.174	0.063
	Min Riesz 0-energy	0.111	0.174	0.063
7	(No Platonic Solid)	N/A	N/A	N/A
	Max Convex Hull Vol	0.121	0.199	0.079
	Max N.N. Dist	N/A	N/A	N/A
	Min Covering Radius	0.120	0.196	0.076
	Min Riesz 0-energy	0.112	0.199	0.087
8	Cube	0.122	0.170	0.048
	Max Convex Hull Vol	0.118	0.168	0.050
	Max N.N. Dist	0.113	0.173	0.060
	Min Covering Radius	0.116	0.160	0.044
	Min Riesz 0-energy	0.124	0.160	0.036

TABLE II: Minimum and maximum probabilities of error for the M points on a sphere that minimize Riesz 0-energy with the Euclidean distance metric. The values below were generated using $P(\vec{r}_1) = 1/2$.

α	M	**Min** P_e	**Max** P_e	**Difference**
$\pi/4$	4	0.223	0.343	0.119
	5	0.235	0.332	0.097
	6	0.255	0.316	0.061
	7	0.252	0.332	0.080
	8	0.261	0.305	0.043
	9	0.271	0.300	0.029
	10	0.268	0.296	0.028
$\pi/2$	4	0.076	0.213	0.137
	5	0.091	0.198	0.107
	6	0.111	0.174	0.063
	7	0.112	0.199	0.087
	8	0.124	0.160	0.036
	9	0.129	0.156	0.027
	10	0.132	0.151	0.019
$3\pi/4$	4	0.029	0.121	0.092
	5	0.044	0.116	0.072
	6	0.063	0.092	0.029
	7	0.059	0.118	0.060
	8	0.066	0.091	0.025
	9	0.068	0.085	0.016
	10	0.070	0.086	0.015

TABLE III: Minimum and maximum probabilities of error for the $M = 6$ points on the sphere that minimize covering radius.

$P(\vec{r}_1)$	α	**Min** P_e	**Max** P_e	**Difference**
$1/8$	$\pi/4$	0.095	0.123	0.028
	$\pi/2$	0.046	0.086	0.040
	$3\pi/4$	0.023	0.052	0.029
$1/4$	$\pi/4$	0.177	0.224	0.047
	$\pi/2$	0.081	0.133	0.051
	$3\pi/4$	0.042	0.075	0.033
$3/8$	$\pi/4$	0.234	0.286	0.053
	$\pi/2$	0.103	0.155	0.052
	$3\pi/4$	0.055	0.086	0.031
$1/2$	$\pi/4$	0.255	0.317	0.061
	$\pi/2$	0.111	0.174	0.063
	$3\pi/4$	0.063	0.092	0.029

[4] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information: 10th Anniversary Edition. Cambridge University Press, 2010.

[5] H. Zhu, "Quantum state estimation and symmetric informationally complete POMs," Ph.D. thesis, NUS Graduate School for Integrative Sciences and Engineering Centre for Quantum Technologies, Nat. Univ. of Singapore, Singapore, 2012.

[6] Huangjun Zhu. "Quantum state estimation with informationally overcomplete measurements," Physical Review A, vol. 90, no. 1, pp. 012115-1 – 012115-14, Jul. 2014, doi: *10.1103/PhysRevA.90.012115*.

[7] T. Decker, D. Janzing, and T. Beth. "Quantum circuits for single-qubit measurements corresponding to platonic solids." International Journal of Quantum Information, vol. 2, no. 3, pp. 353-377, Sept. 2004.

[8] W. Słomczyński and A. Szymusiak. "Highly symmetric POVMs and their informational power," Quantum Information Processing, vol. 15, no. 1, pp. 565-606, Jan. 2016.

[9] C. Medlock and A. Oppenheim and P. Boufounos, "Informationally overcomplete POVMs for quantum state estimation and binary detection," 2020, arXiv:2012.05355.

[10] A. V. Oppenheim and G. C. Verghese, Signals, systems and inference. Hoboken, NJ: Pearson, 2016, p. 356-357.

[11] C. A. Medlock and A. V. Oppenheim (2021), "Operating Characteristics for Classical and Quantum Binary Hypothesis Testing," Foundations and Trends(r) in Signal Processing: Vol. 15: No. 1, pp 1-120. http://dx.doi.org/10.1561/2000000106.

[12] C. Medlock, "Quantum Binary State Discrimination with Overcom-

pleteness," Sc.D. thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, USA, 2021.
[13] D. P. Hardin, T. J. Michaels, and E. B. Saff. "A comparison of popular point configurations on $\mathbb{S}^2$," 2016, arXiv:1607.04590.
[14] E. B. Saff and A. B. Kuijlaars. "Distributing many points on a sphere," The Mathematical Intelligencer, vol. 19, no. 1, pp. 5-11, Dec. 1997, doi: *https://doi.org/10.1007/BF03024331*.
[15] P. C. Leopardi, "Distributing points on the sphere: partitions, separation, quadrature and energy," Ph.D. thesis, Univ. of New South Wales, Sydney, Australia, 2007.
[16] N. J. A. Sloane, "Part 1, Dimension 3 through 5, With Up To 130 Points", Spherical codes: nice arrangements of points on a sphere in various dimensions. [Online]. Available: http://neilsloane.com/packings/index.html. Based on joint work with R. H. Hardin, W. D. Smith and others. [Accessed: Aug. 2021].
[17] R. H. Hardin, N. J. A. Sloane, and D. Smith, "Spherical coverings". [Online]. Available: http://neilsloane.com/coverings/index.html. [Accessed: Aug. 2021].
[18] R. H. Hardin, N. J. A. Sloane, and D. Smith, "Maximal volume spherical codes". [Online]. Available: http://neilsloane.com/maxvolumes/index.html. [Accessed: Aug. 2021].
[19] Y. Zhou, "Arrangements of points on the sphere," Ph.D. thesis, Dept. of Mathematics, Univ. of South Florida, FL, USA, 1995.