Is an Extremal Grassmannian POVM Attaining the Accessible Information?

N. Karimi¹, 2, * and M. A. Jafarizadeh¹, 2, 3, †

¹Department of Theoretical Physics and Astrophysics, Tabriz University, Tabriz 51664, Iran
²Institute for Studies in Theoretical Physics and Mathematics, Tehran 19395-1795, Iran
³Research Institute for Fundamental Sciences, Tabriz 51664, Iran

(Received October 28, 2006)

With regard to Davies’ theorem for an ensemble of \(d\)-dimensional Hilbert spaces, there exists an optimal positive operator-valued measure (POVM) with \(n\) rank-one operators where \(d \leq n \leq d^2\). Also, with respect to this theorem, for symmetric quantum states with irreducible representation, there is a symmetric POVM maximizing the mutual information. For a given class of POVMs of fixed redundancy, we define the Grassmannian POVM as the one that minimizes the maximal correlation among all POVMs. We show that the (3-2) extremal Grassmannian POVM in \(\mathbb{R}^2\) and the symmetric informationally complete POVM as an extremal Grassmannian POVM in \(\mathbb{C}^d\) are compatible with Davies’ theorem.

PACS numbers: 03.67.-a, 03.67.HK

1. INTRODUCTION

In quantum theory, measurements are represented by positive operator-valued measures (POVMs). In order to be maximally efficient at obtaining information regarding the preparation of a quantum state, such measurements should also be of rank one, i.e., the measurement operators or POVM elements should be positive multiples of projections onto pure states, in which case each POVM element corresponds uniquely (up to a phase) to a subnormalized vector in \(\mathbb{C}^d\). A particularly appealing and potentially useful measurement is one which is Grassmannian, meaning that the cross-correlation among a given class of POVMs is minimal. Such a POVM is called a Grassmannian positive operator-valued measures, or Grassmannian POVM for short. The name Grassmannian POVM is extracted from Grassmannian frame, which itself is motivated by the fact that in finite dimensions Grassmannian frames coincide with optimal packing in certain Grassmannian spaces [1, 2].

The main motivation for the present investigation is twofold: the first one is the remarkable fact that, in quantum information theory, such measurements are relevant to the problem of optimizing the Shannon mutual information, and the other is to provide fundamental studies where they would make for particularly interesting standard quantum measurements [3]. For example, here it is shown that extremal Grassmannian POVMs are useful for the optimization of the classical mutual information for real two-dimensional symmetric quantum sources. It is also shown that the well-known class of symmetric informationally complete POVMs, namely SIC-POVMs [4], is a particular case of extremal Grassmannian POVMs.

The paper is organized as follows: In Section 2 we define a Grassmannian POVM
as an interesting case of a POVM with minimal cross correlation among all POVMs. In Section 3 we briefly review the research question in the language of frame theory. In Section 4, through the use of a theorem due to Benedetto and Kolesar, we give an explicit form of a \((n-2)\)-Grassmannian POVM in real two-dimensional space. We use the \((n-2)\)-Grassmannian POVM to optimize the Shannon mutual information for a real symmetric source \(X\) with equal probability, and show that the maximum Shannon mutual information occurs for an extremal Grassmannian POVM. In Section 5 the relation between SIC-POVMs and Grassmannian POVMs is explained. In Section 6 we explain the group-theoretic formalism for symmetric ensembles, leading to the definition of covariant SIC-POVMs. In Section 7 we show that the covariant SIC-POVM is an optimal strategy for a covariant ensemble.

2. THE \((n,d)\)-GRASSMANNIAN POVM

This section deals with the basic definitions and the main motivation for introducing the \((n,d)\)-Grassmannian POVM. We will deal with the operators acting on a finite dimensional vector space \(\mathbb{C}^d\) or \(\mathbb{R}^d\). Let \(\{P_i\}_{i=1}^d\) be a complete set of orthogonal projections, that is,

\[
P_iP_j = P_i \delta_{ij}, \quad \forall \ i, j \in \{1, \ldots, d\}. \tag{2-1}
\]

A set of POVMs \(\{\hat{\Pi}_k\}_{k=1}^n\) with \(n \geq d\) is a generalization of an orthogonal basis whose members are hermitian positive operators forming a resolution of the identity, i.e.,

\[
\text{non-negativity} : \quad \hat{\Pi}_k = \hat{\Pi}_k^\dagger \geq 0, \quad k = 1, \ldots, n,
\]

completeness : \(\sum_{k=1}^n \hat{\Pi}_k = \hat{I}\). \tag{2-2}

If we consider the Hilbert-Schmidt product \((\hat{A}, \hat{B}) = \text{Tr}(\hat{A}^\dagger \hat{B})\) between two operators, condition (2-1) essentially states that the elements of orthogonal projections are completely uncorrelated. However, we know that the operators \(\hat{\Pi}_i\) do not need to be orthogonal to each other, that is, \(\text{Tr}(\hat{\Pi}_i\hat{\Pi}_j)\) does not need to be equal to zero when \(i \neq j\). This suggests searching for POVMs \(\{\hat{\Pi}_i\}_{i=1}^n\) with \(n \geq d\), such that the maximal correlation \((\hat{\Pi}_i\hat{\Pi}_j) = \text{Tr}(\hat{\Pi}_i^\dagger \hat{\Pi}_j)\) for all \(i, j \in \{1, \ldots, n\}\) with \(i \neq j\) is as small as possible. This idea will lead us to the so-called Grassmannian POVM, which is characterized by the property that the POVM elements have minimal cross-correlation among a given class of POVMs.

**Definition 1.** For a POVM \(\{\hat{\Pi}_k = \frac{1}{n} |\psi_k\rangle\langle\psi_k|\}_{k=1}^n\) made up of normalized vectors \(\chi_d^n = \{|\psi_k\rangle \in \mathbb{C}^d\}_{k=1}^n\) the quantity

\[
\mathcal{M}(\chi_d^n) = \max_{k \neq l} |\langle\psi_k|\psi_l\rangle|^2 \tag{2-3}
\]

is called the maximum correlation of \(\chi_d^n\).
Definition 2. Let \( n \geq d \), and suppose that
\[
\mu(n, d) := \min \mathcal{M}(\chi_{n,d}),
\]
(2-4)
where the minimum is taken over all POVMs. A set POVM \( \chi_{n,d} = \{ \Pi_k = \frac{d}{n}|\psi_k\rangle\langle \psi_k| \} \) is called an \((n,d)\)-Grassmannian POVM if it satisfies
\[
\max_{k \neq l} |\langle \psi_k | \psi_l \rangle|^2 = \mu(n, d).
\]
(2-5)

In other words, a Grassmannian POVM minimizes the maximal correlation between POVM elements among all POVMs (acting on \( \mathbb{E}^d = \mathbb{C}^d \) or \( \mathbb{R}^d \)) which have the same redundancy.

Two questions arise naturally when studying finite Grassmannian POVMs:

**Question 1:** Can we derive bounds on Grassmannian POVMs for given \( n \) and \( d \)?

**Question 2:** How can we construct Grassmannian POVMs?

Theorem 3 as a fundamental result provides an exhaustive answer to Question 1 and Theorem 5 rigourously answers Question 2 in real two-dimensional vector space.

3. FRAMES AS A USEFUL TOOL FOR THE CONSTRUCTION OF POVMS

The concept of frame theory provides a simple and robust means of putting our problem in a general setting, since a Grassmannian POVM is a particular kind of tight frame. The set of vectors comprising a Grassmannian POVM was first studied by Strohmer and Heath [5] and subsequently by Benedetto and Kolesar [6]. Indeed, attempts have been made to construct a particular POVM by means of frame theory [4, 7–10]. Frames are a generalization of basis sets, with the requirements of orthogonality and normalization relaxed. For a finite-dimensional vector space \( \mathcal{H} \), a collection of vectors \( |\tilde{f}_k\rangle \in \mathcal{H} \) is a frame if there exist constants \( 0 < A \leq B < \infty \) such that
\[
A |\langle f | f \rangle|^2 \leq \sum_k |\langle f | \tilde{f}_k \rangle|^2 \leq B |\langle f | f \rangle|^2
\]
(3-6)
for all \( |f \rangle \in \mathcal{H} \). Any collection of vectors is a frame in the subspace spanned by the vectors. The constants \( A \) and \( B \) are called frame bounds, and if \( A = B \) the frame is said to be tight (TF), and if \( \langle f_i | f_i \rangle = \langle f_j | f_j \rangle \) for all \( i,j \) it is called a uniform tight frame (UTF). We may work with the frames more easily by formulating them in terms of operators. The analysis operator, or frame transform, \( T : \mathcal{H} \to \ell_2 \), decomposes any vector into the sequence of overlaps with the frame elements: \( T|\psi\rangle = \{ \langle \tilde{f} | \psi \rangle \} \). The adjoint is a synthesis operator, or a preframe operator, \( T^\dagger : \ell_2 \to \mathcal{H} \), which creates a vector from a sequence: \( T^\dagger |a_k\rangle = \sum_k a_k |\tilde{f}_k\rangle \). The frame operator is the positive operator \( S : \mathcal{H} \to \mathcal{H} \) such that \( S = T^\dagger T \), i.e.,
\[
S = \sum_k |\tilde{f}_k\rangle \langle \tilde{f}_k|.
\]
(3-7)
Putting the analysis and synthesis operators in the other order $TT^\dagger$ yields the Gram matrix $G_{j,k} = \langle \phi_j | \phi_k \rangle$ of the frame elements. In terms of the frame operator, Equation (3-6) now reads $A I \leq S \leq B I$ where $I$ is the identity matrix. It is evident that for a tight frame $S = A I$. This tight frame condition is equivalent to the completeness condition for the corresponding POVM elements $| f_k \rangle \langle f_k | / A$, and thus rank-one POVMs and tight frames are the same mathematical objects.

Now let $S^d \subset \mathbb{C}^d$ be the subset consisting of vectors that have unit norm. Any frame can be rewritten in terms of the corresponding normalized vectors, but tightness is not preserved under this transformation.

Since every rank-one POVM is a tight frame, a Grassmanian POVM is clearly something more. Obviously, the minimum in (2-4) just depends on the parameters $n$ and $d$. Thus, we state a lower bound for the maximal correlation between the set of vectors comprising $n$-element POVMs for $E^d = \mathbb{C}^d$ or $\mathbb{R}^d$ [5, 6]. Such lower bounds are useful in coding theory. The following theorem is new in fundamental quantum mechanics, but actually it only unifies and summarizes results from various research areas.

**Theorem 3. (Welch Bound)** Let $n \geq d$, let $\chi^n_d = \{ \hat{\Pi}_k = d/n | \psi_k \rangle \langle \psi_k | \}_{k=1}^n$ be a set of POVMs corresponding to a UTF in $E^d = \mathbb{R}^d$ or $\mathbb{C}^d$. Then

$$M(\chi^n_d) \geq \sqrt{\frac{n-d}{d(n-1)}},$$

where equality holds iff $\chi^n_d$ is equiangular.

Furthermore,

if $E = \mathbb{R}$ equality in (3-8) can only hold if $n \leq \frac{d(d+1)}{2}$,

if $E = \mathbb{C}$ equality in (3-8) can only hold if $n \leq d^2$.

**Proof:** A proof of the bound (3-8) can be found in [11]. It also follows from Lemma 6.1 in [12]. One way to derive (3-8) is to consider the nonzero eigenvalues $\lambda_1, ..., \lambda_d$ of the Gram matrix $[G_{k,l}] = \langle \psi_k | \psi_l \rangle$. These eigenvalues $\lambda_1, ..., \lambda_d$ satisfy $\sum_{k=1}^d \lambda_k = n$ and also

$$\sum_{k=1}^d \lambda_k^2 = \sum_{k=1}^n \sum_{l=1}^n | \langle \psi_k | \psi_l \rangle |^2 \geq \frac{n^2}{d}. \quad (3-11)$$

The bound follows now by taking the maximum over all $| \langle \psi_k | \psi_l \rangle |$ in (3-11) and observing that there are $n(n-1)/2$ different pairs $\langle \psi_k | \psi_l \rangle$ for $k \neq l$.

Equality in (3-8) implies $\lambda_1, ..., \lambda_d = \frac{n}{d}$, which in turn implies $| \langle \psi_k | \psi_l \rangle |^2 = \frac{n-d}{d(n-1)}$ for all $k, l$ with $k \neq l$ which yields the equiangularity. Finally, the bounds on $n$ in (3-9) and (3-10) follow from the bounds in Table II of [13]. Q.E.D.
Definition 4. Let $n \geq d$. A set of POVMs $\chi^n_d = \{\hat{\Pi}_k = \frac{1}{n} |\psi_k\rangle\langle\psi_k|\}_{k=1}^n$ corresponding to a UTF is called an extremal Grassmannian POVM if $\chi^n_d$ satisfies Eq. (3-8) with equality, i.e.,

$$M(\chi^n_d) = \sqrt{\frac{n - d}{d(n - 1)}}.$$ 

4. A $(n, 2)$-GRASSMANNIAN POVM ON $\mathbb{R}^2$ AND AN OPTIMAL STRATEGY FOR OBTAINING CLASSIFICATION INFORMATION ON A (SYMMETRIC) ENSEMBLE

The classification of Grassmannian POVMs in two real dimensions is easy. A useful theorem due to Benedetto and Kolesar, which we state without proof, classifies all $(n, 2)$-Grassmannian POVMs in two real dimensions [6].

Theorem 5. Let $\chi^n_2 = \{\hat{\Pi}_k = \frac{2}{n} |\psi_k\rangle\langle\psi_k|\}_{k=1}^n$ be the set of POVMs corresponding to a UTF in $\mathbb{R}^2$. Then we have the lower bound

$$\cos(\pi/n) \leq M(\chi^n_2).$$

Furthermore, $\chi^n_2$ is an $(n, 2)$-Grassmannian POVM iff there is a $P \in SO(2)$ and a sequence $\{\varepsilon_k\}_{k=1}^n \subset \{\pm 1\}^n$ such that

$$P(\varepsilon \chi^n_2) := \{P(\varepsilon |\psi_k\rangle) : |\psi_k\rangle \in \chi^n_2\} = \left\{\sqrt{\frac{2}{n}} \left(\begin{array}{c}
\cos(k\pi/n) \\
\sin(k\pi/n)
\end{array}\right) : k = 1, \ldots, n\right\}. \quad (4-12)$$

<table>
<thead>
<tr>
<th>$n$</th>
<th>Extremal bound $= \sqrt{\frac{n-2}{2(n-1)}}$</th>
<th>Bound from Theorem 5 $\cos(\pi/n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.5000</td>
<td>0.5000</td>
</tr>
<tr>
<td>4</td>
<td>0.5774</td>
<td>0.7071</td>
</tr>
<tr>
<td>5</td>
<td>0.6124</td>
<td>0.8090</td>
</tr>
<tr>
<td>6</td>
<td>0.6325</td>
<td>0.8660</td>
</tr>
<tr>
<td>7</td>
<td>0.6455</td>
<td>0.9010</td>
</tr>
<tr>
<td>8</td>
<td>0.6547</td>
<td>0.9239</td>
</tr>
<tr>
<td>9</td>
<td>0.6614</td>
<td>0.9397</td>
</tr>
<tr>
<td>10</td>
<td>0.6667</td>
<td>0.9511</td>
</tr>
</tbody>
</table>

TABLE 1: Improvement of the extremal bound derived in Theorem 5 for the case of a $(n, 2)$-Grassmannian POVM.

The above $(n, 2)$-Grassmannian POVM coincides with the covariant POVM introduced by Sasaki et al. [14]. They proposed an optical scheme to implement some $(n, 2)$-Grassmannian
POVMs. Using these \((n,2)\)-Grassmannian POVMs for the optimization of the Shannon mutual information for real symmetric sources,

\[
|\phi_k\rangle = \begin{pmatrix}
\cos \frac{k\pi}{n} \\
\sin \frac{k\pi}{n}
\end{pmatrix},
\]  
(4-13)

with equal prior probability, they investigated and extended Davies’ theorem for such ensembles [15]. Davies’ theorem states that for symmetric quantum states, there exists a symmetric POVM maximizing the mutual information.

In a general communication setting, let \(x_i \in X\) be input letters (ensemble) and let \(P(x_i)\) be their prior probabilities. Let us denote the output letters by \(y_i \in Y\). The Shannon mutual information is defined in terms of the conditional probability \(P(y_j|x_i)\) for obtaining output \(y_j\) provided that the letter sent is \(x_i\); it is defined as

\[
I(X,Y) = -\sum_{y_i} P(y_i) \log P(y_i) + \sum_{x_i,y_j} P(x_i,y_j) \log P(y_j | x_i),
\]  
(4-14)

where \(P(x_i,y_j) = P(y_j|x_i)P(x_i)\) is the joint probability and \(P(y_i) = \sum_{x_j} P(x_j,y_i)\) is the marginal probability of the output letters \(y_i\).

The mutual information \(I(X,Y)\) is a candidate for measuring how much the entropy of one random variable is reduced by knowledge of the other. The maximum value of \(I(X,Y)\) is called the accessible information of the ensemble \(X\). Use the \((n,2)\)-Grassmannian POVM

\[
\chi_n^2 = \left\{ \hat{\Pi}_k = \frac{2}{n} |\psi_k\rangle \langle \psi_k| : |\psi_k\rangle = \begin{pmatrix}
\cos(\theta + \frac{k\pi}{n}) \\
\sin(\theta + \frac{k\pi}{n})
\end{pmatrix}\right\}
\]  
(4-15)

in the optimization of the ensemble (4-13), and note that \(P(y_j|x_i) = \langle \phi_i|\Pi_j|\phi_i\rangle\) yields

\[
I(\theta) = \frac{1}{n} \sum_{k=0}^{n-1} \left[ 1 + \cos(2\theta - \frac{2k\pi}{n}) \right] \times \log\left[ 1 + \cos(2\theta - \frac{2k\pi}{n}) \right].
\]  
(4-16)

For each \(n\), \(I(\theta)\) has a global maximum at \(\theta = \frac{\pi}{2}\). [14]. We see that the maximum value of the mutual information occurs for the \((2,2)\) and \((3,2)\)-Grassmannian POVMs (see Figure 1 in Ref. [14]). This means that an extremal Grassmannian POVM achieves accessible information for equiprobable real symmetric sources. That the amount of maximum mutual information occurs for an extremal Grassmannian POVM is, incidentally, by itself an interesting finding. No wonder \((3,2)\)-Grassmannian POVMs are widely used in the quantum information literature.

5. A SIC-POVM AS AN EXTREMAL GRASSMANNIAN POVM

A special kind of quantum measurement that has received attention lately is what is known as a symmetric informationally complete positive-operator-valued measure (SIC-POVM). One of our results is that a SIC-POVM is a special case of an extremal Grassmannian POVM when we deal with a \(d\)-dimensional complex space \(\mathbb{C}^d\). SIC-POVMs were
discovered by Zauner [16] and independently by Renes et al. [4]. A SIC-POVM $S$ on a $d$-dimensional Hilbert space $\mathbb{C}^d$ is a POVM with $d^2$ elements $\hat{\Pi}_i$ such that each $\hat{\Pi}_i \in S$ is rank one, i.e., $\hat{\Pi}_i = \frac{1}{d} |\psi_i\rangle \langle \psi_i|$ for some $|\psi_i\rangle \in \mathbb{C}^d$, and each pair of normalized vectors satisfies

$$|\langle \psi_i | \psi_j \rangle|^2 = \frac{d \delta_{ij} + 1}{d+1}.$$  \hspace{1cm} (5-17)

Exact solutions to Eq. (5-17) exist in dimensions 2–8 and 19, and numerical examples exist in all dimensions $\leq 45$. It has been conjectured that SIC-POVMs exist in all dimensions [4, 7, 16, 17]. Now we can prove this corollary.

**Corollary 6.** If the set $S = \{\hat{\Pi}_i, i = 1, \ldots, d^2\}$ forms an SIC-POVM on $\mathbb{C}^d$, then it is an extremal Grassmannian POVM.

**Proof:** To prove this, first note that the Welch bound, Eq. (3-8), is saturated for equiangular POVMs. On the other hand, for a SIC-POVM we have

$$\frac{1}{d+1} = \frac{n-d}{d(n-1)}.$$  \hspace{1cm} (5-18)

The only solution to this equation is $n = d^2$. Q.E.D.

Notice that by Definition 3, the inverse of the above corollary is not true. For instance, the $(3,2)$-Grassmannian POVM in $\mathbb{R}^2$ is extremal, but it is not a SIC-POVM.

6. A GROUP-THEORETIC APPROACH AND COVARIANT SIC-POVMS

For two given elements of a group: $g, g' \in G$ with product $gg'$, a unitary representation $\hat{U}_g$ in a Hilbert space is defined by $\hat{U}_g \hat{U}_g^\dagger = \hat{U}_{gg'}$. A more general type of representation, called a ray or projective representation, is relevant for describing the symmetries of quantum mechanical systems. Let $G$ be a group with identity $\hat{I}$ of order $|G|$, $\mathbb{C}$ the field of complex numbers, $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$, and $GL(n, \mathbb{C})$ the group of invertible $n \times n$ matrices over $\mathbb{C}$. A projective unitary representation of degree $n$ of $G$ is a map $\hat{U} : G \rightarrow GL(n, \mathbb{C})$ such that for $g, g' \in G$  

$$\hat{U}_g \hat{U}_{g'} = e^{i \xi_{gg'}} \hat{U}_{gg'},$$  \hspace{1cm} (6-19)

where $e^{i \xi_{gg'}}$ is a phase. A finite ensemble $\mathcal{E}$ of equiprobable (generally mixed) states is said to be symmetric with respect to the group $G$, or $G$-covariant, if there is a projective representation $\Gamma = \{\hat{U}_g\}$ of $G$ such that for all $g, \hat{\rho} \in \mathcal{E}$ whenever $\hat{\rho}$ is in $\mathcal{E}$. We define

$$g\hat{\rho} := \hat{U}_g \hat{\rho} \hat{U}_g^\dagger,$$  \hspace{1cm} (6-20)
for the action of $g$ on the state $\hat{\rho}$. Note that the phase $e^{ix g \hat{\rho}}$ does not appear in Eq. (6-20) and $g_1(g_2(\hat{\rho})) = (g_1g_2)(\hat{\rho})$. A $G$-covariant POVM $\Pi$ for the projective unitary representation $\Gamma$ is a POVM such that $U_g \Pi U_g^\dagger$ is in $\Pi$ whenever $\hat{\Pi}$ is in $\Pi$. We define
\[ g\hat{\Pi} := U_g \hat{\Pi} U_g^\dagger, \] (6-21)
for the action of $g$ on a POVM element $\hat{\Pi}$. From Eqs. (6-20) and (6-21) we see that
\[ Tr(\hat{\Pi} \hat{\rho}) = Tr(g\hat{\Pi} . g\hat{\rho}), \]
so that the set of probabilities of the $G$-shifted outputs $g\hat{\Pi}$ on a fixed input $\hat{\rho}$ is obtained as a permutation of the set of probabilities of the unshifted output $\hat{\Pi}$ acting on suitably shifted inputs. We will use Schur’s First Lemma, the proof of which we shall give below.

**Lemma 7.** (Schur) If an operator $A$ commutes with all the operators $M(G)$ of an irreducible representation $\Gamma$, then $A$ is multiple of the unit operator.

**Proof:** By hypothesis
\[ M(G)A = AM(G). \] (6-23)
Let $V_n$ be the representation space of $\Gamma$. In this space, there exists at least one eigenvector of the operator $A$. Let us write this vector as $|\phi\rangle$, which is such that
\[ A|\phi\rangle = \lambda|\phi\rangle, \] (6-24)
where $|\phi\rangle$ also represents the operator of the components of the vector $|\phi\rangle$. By applying one of the transformations of the group to $|\phi\rangle$, we obtain
\[ M(G)|\phi\rangle = |\varphi\rangle. \] (6-25)
The vector $|\varphi\rangle$ is also an eigenvector of $A$ associated with the eigenvalue $\lambda$. Because of Eqs. (6-23), (6-24), and (6-25), we have
\[ A|\varphi\rangle = AM(G)|\phi\rangle = M(G)A|\phi\rangle = M(G)\lambda|\phi\rangle = \lambda|\varphi\rangle. \] (6-26)
The vector subspace generated by the eigenvectors of $A$ associated with the same eigenvalue is thus stable under $G$. However, since the representation $\Gamma$ is assumed to be irreducible, this subspace must coincide with $V_n$ completely. As a consequence, the operator $A$ transforming every vector $|\phi\rangle$ of $V_n$ into a collinear vector $\lambda|\phi\rangle$ is an operator such that $A = \lambda I_n$, where $I_n$ is the unit operator of order $n$. Q.E.D.

The SIC-POVMs which have been constructed to date all have a certain group covariance property. Let $G$ be a finite group having $d^2$ elements, and let $\Gamma$ be projective representation of $G$ acting on $d$-dimensional Hilbert space. Suppose one can find a vector $|\psi\rangle \in \mathbb{C}^d$ such that
\[ |(\langle \psi| \hat{U}_g |\psi\rangle)| = \frac{1}{\sqrt{d+1}}, \quad \langle \psi| \psi\rangle = 1 \quad \forall \ g \neq e, \] (6-27)
where $e$ is the identity of $G$. (The vector $|\psi\rangle$ is said to be a fiducial vector). Then, we can prove the following lemma:
Lemma 8. Suppose $\Gamma$ is the projective representation of $G$ such that $\Gamma$ acts irreducibly on $\mathbb{C}^d$ and $|\psi\rangle$ is a fiducial vector. Then the assignment

$$\hat{E}_g = \frac{d}{|G|} \hat{U}_g |\psi\rangle \langle \psi| \hat{U}_g^\dagger$$

(6-28)

defines a covariant SIC-POVM on $\mathbb{C}^d$.

Proof: To show this, let $A = \sum_{g \in G} \hat{E}_g$. Then $A \hat{U}_g = \hat{U}_g A$ for all $g \in G$. Also $A$ is a hermitian positive operator, so it has a real positive eigenvalue $\lambda \geq 0$. Then, by Lemma 7, $A = \lambda \hat{I}$. Since $\text{Tr} \hat{E}_g = \frac{d}{|G|}$ for all $g \in G$, we get $\text{Tr} A = d = \text{Tr} \hat{I}$, so $\lambda = 1$. Q.E.D.

7. COVARIANT SIC-POVM AS AN OPTIMAL STRATEGY

In this section, we show that a covariant SIC-POVM is an optimal strategy for a covariant ensemble $\mathcal{E}$. To do this we first state a definition.

Definition 9. Let $\Pi_1 = \{\hat{\Pi}_1^{(1)}, \ldots, \hat{\Pi}_1^{(m_1)}\}$, $\Pi_2 = \{\hat{\Pi}_2^{(1)}, \ldots, \hat{\Pi}_2^{(m_2)}\}, \ldots, \Pi_{n-1} = \{\hat{\Pi}_{n-1}^{(1)}, \ldots, \hat{\Pi}_{n-1}^{(m_{n-1})}\}$ and $\Pi_n = \{\hat{\Pi}_1^{(n)}, \ldots, \hat{\Pi}_n^{(n)}\}$ be POVMs of a system. For $\lambda_i \in [0, 1]$ and $\sum_i \lambda_i = 1$, the convex combination is defined as

$$\Pi = \lambda_1 \Pi_1 + \lambda_2 \Pi_2 + \cdots + \lambda_n \Pi_n := \{\lambda_1 \hat{\Pi}_1^{(1)}, \ldots, \lambda_1 \hat{\Pi}_1^{(m_1)}, \lambda_2 \hat{\Pi}_2^{(1)}, \ldots, \lambda_2 \hat{\Pi}_2^{(m_2)}, \ldots, \lambda_n \hat{\Pi}_n^{(1)}, \ldots, \lambda_n \hat{\Pi}_n^{(n)}\}.$$ 

This convex combination corresponds to a random selection between POVMs. We should bear in mind which POVM we have chosen after the measurement, i.e., we assume that the results of the POVMs are distinct. In the quantum context, the optimization of $I(X : Y)$ is carried out with respect to the choice of POVM $\{\hat{\Pi}_k\}$ for a fixed ensemble $\mathcal{E} = \{\hat{\rho}_i; P(i)\}$ (i.e., with fixed letter states $\hat{\rho}_i$ and fixed prior probabilities $P(i)$). The set $\Pi$ of all POVMs is a convex set, and $I(X : Y)$ for a fixed ensemble $\mathcal{E} = \{\hat{\rho}_i; P(i)\}$ is a convex function on $\Pi$ [18]. Let $I(\mathcal{E} : \Pi)$ denote the mutual information obtained from the POVM $\Pi$ applied to the ensemble $\mathcal{E}$ and let $I(\mathcal{E} : \Pi)$ be a convex combination of POVMs $\Pi_i$, that is,

$$\Pi = \lambda_1 \Pi_1 + \cdots + \lambda_n \Pi_n,$$  

(7-29)

then it follows from convexity that

$$I(\mathcal{E} : \Pi) \leq \sum_i \lambda_i I(\mathcal{E} : \Pi_i) \leq \max_i I(\mathcal{E} : \Pi_i).$$

(7-30)

Theorem 10. Let $\mathcal{E}$ be any ensemble of equiprobable states in dimension $d$. Suppose that $\mathcal{E}$ is $G$-covariant with respect to a projective unitary representation $\Gamma$, such that $\Gamma$ acts irreducibly on $\mathbb{C}^d$. Then the covariant SIC-POVM defined by

$$\hat{E}_g = \frac{d}{|G|} \hat{U}_g |\psi\rangle \langle \psi| \hat{U}_g^\dagger$$

(7-31)

is an optimal strategy for $\mathcal{E}$. 
Proof: Let $B = \{ \hat{B}_1, \ldots, \hat{B}_m \}$ be any rank-one optimal POVM for $\mathcal{E}$. Now look at 

$$\hat{C}_{kg} = \frac{1}{|G|} g \hat{B}_k, \quad \text{for } g \in G \text{ and } k = 1, \ldots, m. \quad (7-32)$$

Note that $\sum_{kg} \hat{C}_{kg} = \hat{I}$, since $\sum \hat{B}_k = \hat{I}$ and $g \hat{I} = \hat{I}$ for all $g$. Let $\mathcal{C} = \{ \hat{C}_{kg} \}$ be the corresponding POVM with $|G|m$ elements. Thus $\mathcal{C}$ is $G$-covariant, but the action of $G$ is not transitive. We aim to cut down $\mathcal{C}$ to a smaller optimal $G$-covariant POVM with elements labeled by $G$. Let $I(\mathcal{E} : \Pi)$ denote the mutual information obtained from any POVM $\Pi$ applied to any ensemble $\mathcal{E}$. First we show that $I(\mathcal{E} : \mathcal{C}) = I(\mathcal{E} : \mathcal{B})$, so that $\mathcal{C}$ remains optimal. Let us label the inputs by $i \in \mathcal{I}$ and denote conditional probabilities for $\mathcal{C}$ by $P(\hat{C}_{kg} | i)$. Denote the conditional probabilities for $\mathcal{B}$ by $P_B(k | i)$, and let $\xi$ be the constant prior input probability. Then

$$P(k | i) = \text{Tr} \hat{C}_{kg} \hat{\rho}_i = \frac{1}{|G|} \text{Tr} g \hat{B}_k \hat{\rho}_i, \quad (7-33)$$

according to Eq. (5-24), for each fixed $g$ and $k$, the resulting set of probabilities labeled by $i \in \mathcal{I}$ will just be a permutation of the set $P_B(k | i)$, rescaled by $\frac{1}{|G|}$. Thus,

$$\sum_i \xi P(k | i) = \frac{1}{|G|} \sum_i \xi P_B(k | i), \quad (7-34)$$

will be independent of $g$ and also

$$\sum_i \xi P(k | i) \log \frac{P(k | i)}{\xi \sum_i P(k | i)} = \frac{1}{|G|} \sum_i P_B(k | i) \log \frac{P_B(k | i)}{\sum_i \xi P_B(k | i)}, \quad (7-35)$$

will be independent of $g$. Thus mutual information $I(\mathcal{E} : \mathcal{C})$ and $I(\mathcal{E} : \mathcal{B})$ are given by

$$I(\mathcal{E} : \mathcal{C}) = \sum_i \xi \sum_{kg} P(k | i) \log \frac{P(k | i)}{\xi \sum_i P(k | i)} , \quad (7-36)$$

$$I(\mathcal{E} : \mathcal{B}) = \sum_i \xi \sum_k P_B(k | i) \log \frac{P_B(k | i)}{\xi \sum_i P_B(k | i)} . \quad (7-37)$$

By substituting the $G$-covariant expressions with $I(\mathcal{E} : \mathcal{C})$, we readily get $I(\mathcal{E} : \mathcal{C}) = I(\mathcal{E} : \mathcal{B})$. Finally, note that for each $i$, $\frac{\hat{B}_i}{\text{Tr} \hat{B}_i}$ is pure state, so by Lemma 8,

$$E_i = \{ \frac{d}{|G|} \frac{g \hat{B}_i}{\text{Tr} \hat{B}_i} : g \in G \} \quad (7-38)$$

is a POVM for each $i$. Let $\text{Tr} \hat{B}_i E_i = \{ \frac{1}{|G|} g \hat{B}_i : g \in G \}$, so $\mathcal{C}$ is a convex combination

$$\mathcal{C} = \sum_{i=1}^m \frac{\text{Tr} \hat{B}_i}{d} E_i . \quad (7-39)$$
Hence by convexity we have
\[ I(\mathcal{E} : \mathcal{C}) \leq \max_i I(\mathcal{E} : E_i). \] (7-40)

Since \( \mathcal{C} \) was optimal, it follows that at least one of the \( E_i \)'s is optimal. This gives an optimal covariant SIC-POVM with rank-one elements, parameterized by \( G \), completing the proof.

Q.E.D.

7-1. Tetrahedral measurement

An example of a rank-one POVM is given by the tetrahedral measurement, which is specified by the four unit vectors
\[
|\psi_1\rangle = (0, 0, 1)^T, \\
|\psi_2\rangle = \left(\frac{\sqrt{3}}{3}, 0, -\frac{1}{3}\right)^T, \\
|\psi_3\rangle = \left(-\frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{3}, -\frac{1}{3}\right)^T, \\
|\psi_4\rangle = \left(-\frac{\sqrt{2}}{3}, -\frac{\sqrt{2}}{3}, -\frac{1}{3}\right)^T.
\] (7-41)

This POVM is a (4-3) Grassmannian POVM [6] and also a SIC-POVM by virtue of the symmetric placement of the Bloch vectors, which connect the origin to the vertices of a tetrahedron whose apex is at the north pole of the Bloch sphere. Therefore by Theorem 10 it is an optimal strategy for a covariant ensemble.

8. CONCLUSION

At first glance, minimizing the maximal correlation among a given class of POVMs seems abstract and complicated, but it is very promising when we encounter some cases which coincide with Grassmannian POVMs, such as SIC-POVMs and covariant POVMs. The SIC-POVM as a special case of an extremal Grassmannian POVM was deduced from the Welch bound. In the case of a \((n,2)\)-Grassmannian POVM, we saw that the extremity had a good physical interpretation, i.e., accessible information for equiprobable symmetric sources occurs for extremal \((2,2)\) and \((3,2)\)-Grassmannian POVMs at \( \theta = \frac{\pi}{2} \). In fact, we showed that \((3-2)\) extremal Grassmannian POVMs and SIC-POVMs are an optimal strategy for covariant ensembles. Also, the findings indicate that the number of POVMs which are necessary for maximizing mutual information are compatible with the upper bound of Davies’ theorem. This compatibility was proved through examining two instances. However, it is yet unknown whether all extremal Grassmannian POVMs behave similarly.

Moreover, based on our definitions, various interesting issues remain unresolved. For example, we hope to classify all Grassmannian POVMs in \( \mathbb{E}^d = \mathbb{C}^d \) or \( \mathbb{R}^d \). Except for \( \mathbb{R}^2 \), intriguingly, the challenge to find explicit forms of Grassmannian POVMs still lies ahead. Hopefully, these lines of research will be pursued further.
References

* Electronic address: na.karimi@tabrizu.ac.ir
† Electronic address: jafarizadeh@tabrizu.ac.ir