Report On Generalized Quantum Measurement

Abhinav Sapra

1 Introduction

This is a report on generalized quantum measurements written for the course Physics of Quantum Information, being offered in CMI during Jan-May 2024, and closely follows the essay "Generalized Quantum Measurements - Imperfect meters and POVMs" by Nicholas Wheeler. Our main goal in this report is to highlight POVMs and their construction, Neumark's dilation theorem, and the theory of qubit discrimination.

2 Basics Of Quantum Measurement

2.1 Postulates Of Quantum Measurement:

The states of quantum system $\mathcal{S}$ are represented by a complex unit vectors $|\psi\rangle$ in a Hilbert Space $\mathcal{H}$.

- Quantum measurements are described by a collection $\{M_m\}$ of measurement operators. The index m refers to the measurement outcomes that may occur in the experiment.
- If the state of the quantum system is $|\psi\rangle$ immediately before the measurement then the probability of the outcome being m is $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$, and the state of the system after measurement becomes $\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}$.
- Moreover, measurement operators satisfy the completeness equation,
  $\sum_m M_m^\dagger M_m = \mathbb{I}$
  This in turn implies that $\sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle = 1$

2.2 Projective Measurement

John von Neumann theorized that the action of quantum measurement devices("perfect meters") can be simulated by the action of self-adjoint linear operator $\mathcal{A}$.

In non-degenerate case (when all the eigenvlaues are distinct):
we can write, $\mathcal{A} = \sum_{k=1}^n a_k\mathbb{P}_k$ where $\mathbb{P}_k = |a_k\rangle\langle a_k|$.

In degenerate case:
we can write, $\mathcal{A} = \sum_k a_k\mathbb{P}_k$ where a_k are distinct and $\mathbb{P}_k$ projects over ν_k-dimensional k-th-eigenspace and $\sum \nu_k = dim(\mathcal{H})$.

In either case the $\mathbb{P}$-matrices are:

- Hermitian ($\mathbb{P}_m^\dagger = \mathbb{P}_m$)
- Positive ($\langle\psi|\mathbb{P}_m|\psi\rangle \geq 0 \ \ \forall \ \ |\psi\rangle$)
- Complete ($\sum_m \mathbb{P}_m = \mathbb{I}$)

- Orthogonal ($\mathbb{P}_i \cdot \mathbb{P}_j = \delta_{ij}\mathbb{P}_i$)

This can be thought of as a special case of Generalized Quantum Measurement, as if we impose the orthogonality condition on our measurement operators $\{M_m\}$, then we get $\mathbb{P}_m = M_m = M_m^\dagger M_m$

2.3 Density Matrix

A density matrix (or density operator) is a representation of the state of a quantum system. It provides a way to describe both pure and mixed states of a system, accounting for statistical mixtures of quantum states as well as quantum entanglement.
Consider an ensemble $\mathcal{E}$ of pure states given by $\{p_i||\psi_i\rangle\}_{i=1}^N$, which provides state $|\psi_i\rangle$ with probability p_i and the associated density matrix with this ensemble is

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i|$$

where $\sum_{i=1}^N p_i = 1$ and $0 \leq p_i \leq 1$. Therefore, for measurement operators $\{M_m\}$ the probability $p(m) = \sum_{i=1}^N p_i\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle$.

$$\begin{aligned}
p(m) &= \sum_{i=1}^{N} p_i\langle\psi_i|M_m^\dagger M_m|\psi_i\rangle \\
&= \sum_i \sum_j p_i\langle\psi_i|M_m^\dagger M_m|e_j\rangle\langle e_j|\psi_i\rangle && [\because \sum_j |e_j\rangle\langle e_j| = \mathbb{I}] \\
&= \sum_i \sum_j p_i\langle e_j|\psi_i\rangle\langle\psi_i|M_m^\dagger M_m|e_j\rangle \\
&= \sum_j \langle e_j|(\sum_i |\psi_i\rangle p_i\langle\psi_i|)M_m^\dagger M_m|e_j\rangle \\
&= tr(\rho M_m^\dagger M_m)
\end{aligned}$$

where $|e_j\rangle$ is the standard basis for $\mathcal{H}$.

Properties Of Densities Matrices:

- Hermitian
- Positive
- Unit Trace

From spectral decomposition of hermitian matrices, we get $\rho = \sum_k r_k|r_k\rangle\langle r_k| = \sum_k r_k\mathbb{P}_k$, where r_k are distinct eigenvalues of ρ and $\mathbb{P}_k$ projects onto the eigenspaces associated to these eigenvalues.

Thus, we can think of density matrices in two ways:

$$\rho = \begin{cases} \sum p_i|\psi_i\rangle\langle\psi_i|, \text{as } p_i \text{ weighted mixture of } |\psi_i\rangle\text{-states} \\ \sum r_k|r_k\rangle\langle r_k|, \text{as } r_k \text{ weighted ensemble of } |r_k\rangle\text{-eigenstates} \end{cases}$$

3 Ideal And Imperfect Meters

In classical physics, measurement devices are typically considered to be passive observers that simply reveal the properties of the system being measured. However, in quantum mechanics, the act of measurement is inherently interactive and can have a significant effect on the quantum system being measured.

Von Neumann recognized that measurement devices in quantum mechanics are not passive observers but are instead active components of the measurement process. These devices interact with the quantum system and can potentially alter its state. Moreover, the outcomes of quantum measurements are inherently probabilistic, meaning that we can only predict the probabilities of different measurement outcomes rather than predicting the outcomes themselves with certainty.

$$|\psi\rangle \xrightarrow[\textbf{A-measurement}]{} \begin{cases} |a_1\rangle \text{ with probability } |\langle a_1|\psi\rangle|^2 = \langle\psi|\mathbb{P}_1|\psi\rangle \\ \vdots \\ |a_k\rangle \text{ with probability } |\langle a_k|\psi\rangle|^2 = \langle\psi|\mathbb{P}_k|\psi\rangle \\ \vdots \\ |a_n\rangle \text{ with probability } |\langle a_n|\psi\rangle|^2 = \langle\psi|\mathbb{P}_n|\psi\rangle \end{cases}$$

If the meter is read (and a_k is non-degenerate), then

$$|\psi\rangle \xrightarrow[\textbf{A}\text{-meter reads } a_k]{} |a_k\rangle$$

Here, we get state a_i with probability $\langle\psi|\mathbb{P}_i|\psi\rangle \implies \sum probabilities = \langle\psi|\mathbb{I}|\psi\rangle = 1$

But if a_k is degenerate, then we again have a probabilistic mixture of states

$$|\psi\rangle \xrightarrow[\textbf{A}\text{-meter reads} a_k]{} \begin{cases} |a_{k,1}\rangle \text{ with probability } |\langle a_{k,1}|\psi\rangle|^2 = \langle\psi|\mathbb{P}_{k,1}|\psi\rangle \\ |a_{k,2}\rangle \text{ with probability } |\langle a_{k,2}|\psi\rangle|^2 = \langle\psi|\mathbb{P}_{k,2}|\psi\rangle \\ \vdots \\ |a_{k,\nu}\rangle \text{ with probability } |\langle a_{k,\nu}|\psi\rangle|^2 = \langle\psi|\mathbb{P}_{k,\nu}|\psi\rangle \end{cases}$$

Here, the meter reads $a_{k,i}$ with probability $\langle\psi|\mathbb{P}_{k,i}|\psi\rangle \implies \sum probabilties = \langle\psi|\sum_i \mathbb{P}_{k,i}|\psi\rangle = \langle\psi|\mathbb{P}_k|\psi\rangle$

But regardless of the fact that the spectrum is degenerate or non-degenerate, the expected value of the measurement is given by $\langle A\rangle_\psi = \sum_i a_i \cdot \langle\psi_i|\mathbb{P}_i|\psi_i\rangle = \langle\psi|A|\psi\rangle = tr(\rho A)$

But to perform many such measurements we must possess an ensemble $\mathcal{E}(\mathcal{S}_\psi)$, such that each system is in state $|\psi\rangle$.This leads us to study of Imperfect Meters.

3.1 Imperfect Meters

To make multiple measurements,we need an ensemble of sysytems, $\mathcal{E}(\mathcal{S}_\psi)$.Such an ensemble is called a Pure Ensemble.But to make such an ensemble we need error-free equipments that can make such systems or detect such systems without error.

$$|unknown\ state\rangle \xrightarrow[\text{pre-selection}]{\text{G-meter}} \begin{cases} |\psi\rangle \xrightarrow{A-meter} \left\{ \vdots \right. \\ \\ |\psi\rangle_{undesired} \longrightarrow \text{Discarded} \end{cases}$$

If G-meter is imperfect then we get a mix of states,such that $\mathcal{E}(\mathcal{S}_{\{\psi_1,\psi_2,\dots\}}) = \begin{cases} \vdots \\ |\psi_i\rangle \text{with probability } p_i \\ \vdots \end{cases}$

Now if we measure with perfect A-meter,then we get a_k with probability $= \sum_i p_i \langle\psi_i|\mathbb{P}_k|\psi_i\rangle$.
Thus, sum of all probabilities is

$$\sum_k \sum_i p_i \langle\psi_i|\mathbb{P}_k|\psi_i\rangle = \sum_i p_i \langle\psi_i||\psi_i\rangle = \sum_i p_i = 1$$

Moreover, average or expected value of the measurement is given by

$$\begin{aligned}
\langle A\rangle_{\mathcal{E}} &= \sum_i p_i \langle A\rangle_{\psi_i} \\
&= \sum_i p_i \langle\psi_i|A|\psi_i\rangle \\
&= \sum_j \sum_i p_i \langle\psi_i|A|e_j\rangle\langle e_j|\psi_i\rangle \\
&= \sum_j \sum_i \langle e_j|\psi_i\rangle p_i \langle\psi_i|A|e_j\rangle \\
&= \sum_j \langle e_j(\sum_i |\psi_i\rangle p_i \langle\psi_i|)A|e_j\rangle \\
&= tr(\rho_{\mathcal{E}} A), \quad where \ \ \rho_{\mathcal{E}} = density\ matrix
\end{aligned}$$

4 Neumark's Dilation Theorem : From POVM To PVM

Consider N un-normalized n-vectors $\{|\phi_1\rangle, \dots, |\phi_N\rangle\}$, such that $N \geq n$.
Using this we contruct an N-element POVM $\{\tilde{\mathbb{P}}_1, \dots, \tilde{\mathbb{P}}_N\}$, where $\tilde{\mathbb{P}}_k = |\phi_k\rangle\langle\phi_k|$ operates on $\mathcal{H}_n$.

$$\sum_k \mathbb{P}_k = \mathbb{I} \implies \sum_k (\mathbb{P}_k)_{ij} = \sum_k \phi_{ai}\bar{\phi}_{aj} = \delta_{ij}$$

Now we can interpret ϕ_{ai} as the a-th element of the i-th member $|\Phi_i\rangle$ of a set of N-vectors.Therefore,

$$\sum_{a=1}^{N} \phi_{ai}\bar{\phi}_{aj} = \delta_{ij} \text{ reads } \sum_{a=1}^{N} \Phi_{ia}\Phi_{ja} = \delta_{ij}$$

This means that $|\Phi_i\rangle$ and $|\Phi_j\rangle$ are orthogonal.Thus, using Gram-Schmidt process, we can extend the orthogonal set $\{\Phi_1, \dots, \Phi_n\}$ to an orthogonal basis $\{\Phi_1, \dots, \Phi_n, \Phi_{n+1}, \dots, \Phi_N\}$ of the bigger hilbert space $\mathcal{H}_N$.Now feed the elements of $|\Phi_i\rangle$ in the i-th row of a $N \times N$

$$\mathbb{U} = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \dots & \Phi_{1n} & \dots & \Phi_{1N} \\ \Phi_{21} & \Phi_{22} & \dots & \Phi_{2n} & \dots & \Phi_{2N} \\ \vdots & \vdots & & \vdots & & \vdots \\ \Phi_{n1} & \Phi_{n2} & \dots & \Phi_{nn} & \dots & \Phi_{nN} \\ \Phi_{n+1,1} & \Phi_{n+1,2} & \dots & \Phi_{n+1,n} & \dots & \Phi_{n+1,N} \\ \vdots & \vdots & & \vdots & & \vdots \\ \Phi_{N1} & \Phi_{N2} & \dots & \Phi_{Nn} & \dots & \Phi_{NN} \end{pmatrix}$$

$$= \begin{pmatrix} \phi_{11} & \phi_{21} & \dots & \phi_{n1} & \dots & \phi_{N1} \\ \phi_{12} & \phi_{22} & \dots & \phi_{n2} & \dots & \phi_{N2} \\ \vdots & \vdots & & \vdots & & \vdots \\ \phi_{1n} & \phi_{2n} & \dots & \phi_{nn} & \dots & \phi_{Nn} \\ \Phi_{n+1,1} & \Phi_{n+1,2} & \dots & \Phi_{n+1,n} & \dots & \Phi_{n+1,N} \\ \vdots & \vdots & & \vdots & & \vdots \\ \Phi_{N1} & \Phi_{N2} & \dots & \Phi_{Nn} & \dots & \Phi_{NN} \end{pmatrix}$$

Row-Orthogonality $\iff$ $\mathbb{U}$ is Unitary $\iff$ Column-Orthogonality
Let $\mathbb{P}_k = |E_k\rangle\langle E_k|$, where $|E_k\rangle$ is the k-th column of $\mathbb{U}$.
The upshot of using $\{\mathbb{P}_i\}$ set of operators on $\mathcal{H}_N$ is that it is a PVM on $\mathcal{H}_N$ and when we write

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{pmatrix} \quad \text{as} \quad |\Psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{we get, } \langle\psi|\tilde{\mathbb{P}}_k|\psi\rangle = \langle\Psi|\mathbb{P}_k|\Psi\rangle$$

Therefore, probability of a given outcome is unaltered, but we are using **PVM** instead of **POVM**.

4.1 Constructing POVMs

It is well known that we can write a traceless 2×2 hermitian matrix as

$$\mathbb{H} = h_1\sigma_1 + h_2\sigma_2 + h_3\sigma_3 = \begin{pmatrix} h_3 & h_1 - ih_2 \\ h_1 + ih_2 & -h_3 \end{pmatrix}$$

The eigenvalues of this matrix are $\pm\sqrt{h_1^2 + h_2^2 + h_3^2}$ and if eigenvalues are ± 1 then the orthonormal eigenspaces are $h_\pm = \begin{pmatrix} \sqrt{\frac{1 \pm h_3}{2}} \\ \pm\sqrt{\frac{1}{2(1 \pm h_3)}}(h_1 + ih_2) \end{pmatrix}$. Then the associated projectors are given by

$$\mathbb{P}_+(h) = |h_+\rangle\langle h_+| = \begin{pmatrix} (1 + h_3)/2 & (h_1 - ih_2)/2 \\ (h_1 + ih_2)/2 & (1 - h_3)/2 \end{pmatrix}$$

$$\mathbb{P}_-(h) = |h_-\rangle\langle h_-| = \begin{pmatrix} (1-h_3)/2 & -(h_1-ih_2)/2 \\ -(h_1+ih_2)/2 & (1+h_3)/2 \end{pmatrix}$$

For $n \geq 3$,we have vectors $h_1, h_2, \ldots, h_n$ such that $\sum h_i = 0$, then the associated projectors add up to $\sum \mathbb{P}_+(h_i) = \frac{n}{2}\mathbb{I}$ and thus, $\{\frac{2}{n}\mathbb{P}_+(h_1), \ldots, \frac{2}{n}\mathbb{P}_+(h_n)\}$ is POVM if $\sqrt{h_i \cdot h_i} \leq 1$(because eigenvalues of operators are non-negative).

4.2 Example : Neumark's Theorem In Action

let $h_1 = (0,0,1)$,$h_2 = (\frac{\sqrt{3}}{2}, 0, \frac{-1}{2})$ and $h_3 = (\frac{-\sqrt{3}}{2}, 0, \frac{-1}{2})$ then we get a set of vectors $\{|\psi_i\rangle\}_{i=1}^{3}$ in $\mathcal{H}_2$,

$$|h_{1+}\rangle = |\psi_1\rangle = \begin{pmatrix} \sqrt{\frac{2}{3}} \\ 0 \end{pmatrix}, |h_{2+}\rangle = |\psi_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ 1/\sqrt{2} \end{pmatrix}, |h_{3+}\rangle = |\psi_3\rangle = \begin{pmatrix} \frac{-1}{\sqrt{6}} \\ 1/\sqrt{2} \end{pmatrix}$$

$$\tilde{\mathbb{P}}_1 = |\psi_1\rangle\langle\psi_1| = \begin{pmatrix} 2/3 & 0 \\ 0 & 0 \end{pmatrix}, \tilde{\mathbb{P}}_2 = |\psi_2\rangle\langle\psi_2| = \begin{pmatrix} 1/6 & 1/2\sqrt{3} \\ 1/2\sqrt{3} & 1/2 \end{pmatrix}, \tilde{\mathbb{P}}_3 = |\psi_3\rangle\langle\psi_3| = \begin{pmatrix} 1/6 & -1/2\sqrt{3} \\ -1/2\sqrt{3} & 1/2 \end{pmatrix}$$

then $\{\tilde{\mathbb{P}}_1, \tilde{\mathbb{P}}_2, \tilde{\mathbb{P}}_3\}$ is a POVM.

Using Neumark's Dilation Theorem, we get

$$\mathbb{U} = \begin{pmatrix} \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ x & y & z \end{pmatrix}$$

such that all the rows are orthogonal.First two rows are orthogonal. For the last row to be a unit vector orthogonal to the two rows , x, y, z should satisfy the conditions:

$$\sqrt{\frac{2}{3}}x + \frac{1}{\sqrt{6}}y - \frac{1}{\sqrt{6}}z = 0\ ,\ y/\sqrt{2} + z/\sqrt{2} = 0\ ,\ x^2 + y^2 + z^2 = 1$$

Hence, we get $x = z = \frac{1}{\sqrt{3}}$ *and* $y = \frac{-1}{\sqrt{3}} \implies \mathbb{U} = \begin{pmatrix} \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$

Therefore, we get that the corresponding PVM operators are:

$$\mathbb{P}_1 = \begin{pmatrix} \frac{2}{3} & 0 & \frac{\sqrt{2}}{3} \\ 0 & 0 & 0 \\ \frac{\sqrt{2}}{3} & 0 & \frac{1}{3} \end{pmatrix}, \mathbb{P}_2 = \begin{pmatrix} \frac{1}{6} & \frac{1}{2\sqrt{3}} & \frac{1}{3\sqrt{2}} \\ \frac{1}{2\sqrt{3}} & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}, \mathbb{P}_3 = \begin{pmatrix} \frac{1}{6} & \frac{-1}{2\sqrt{3}} & \frac{-1}{3\sqrt{2}} \\ \frac{-1}{2\sqrt{3}} & \frac{1}{2} & \frac{1}{\sqrt{6}} \\ \frac{-1}{3\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{3} \end{pmatrix}$$

It is easy to see that these operators are Orthogonal,thus $\{\mathbb{P}_i\}$ is a PVM.Therefore, we have successfully made a PVM from a given POVM.

5 Theory Of Qubit Discrimination

So the problem at hand is that, Alice wants to send an encode message to Bob using two pre-determined non-orthogonal states, say $|\alpha\rangle$, $|\beta\rangle$.Bob's assignment is to read the message as best as he can.
clearly, we cannot use a PVM, because of non-orthogonality of the chosen states.
So we construct the operators $E_1 = k(\mathbb{I} - |\alpha\rangle\langle\alpha|)$, $E_2 = k(\mathbb{I} - |\beta\rangle\langle\beta|)$ and $E_3 = \mathbb{I} - E_1 - E_2$, where k is chosen such that E_3 is positive, i.e., $0 < k \leq \frac{1}{1+|\langle\alpha|\beta\rangle|}$.Therefore, every time Bob receives a qubit, his meter flashes with probabilities

State	#1	#2	#3
$\|\alpha\rangle$	$\langle\alpha\|E_1\|\alpha\rangle$	$\langle\alpha\|E_2\|\alpha\rangle$	$\langle\alpha\|E_3\|\alpha\rangle$
$\|\beta\rangle$	$\langle\beta\|E_1\|\beta\rangle$	$\langle\beta\|E_2\|\beta\rangle$	$\langle\beta\|E_3\|\beta\rangle$

State	#1	#2	#3
$\|\alpha\rangle$	0	$k(1-x^2)$	$1-k(1-x^2)$
$\|\beta\rangle$	$k(1-x^2)$	0	$1-k(1-x^2)$

On calculating the probabilities we get

where $x = |\langle\alpha|\beta\rangle|$.

Thus,
#1 flashes $\longrightarrow$ we can say for sure it was $|\beta\rangle$
#2 flashes $\longrightarrow$ we can say for sure it was $|\alpha\rangle$
#3 flashes $\longrightarrow$ we cannot say for sure it was $|\alpha\rangle$ or $|\beta\rangle$
we want to minimize the probability of #3,and we can do that by putting $k = k_max = 1/(1+x)$, so that probability of #3 flashing comes down to $|\langle\alpha|\beta\rangle|$

The key take-away from this is that imperfect(PVOM) meters can be designed to exploit facts to provide information that remains forever beyond the reach of perfect(PVM) meters.