# 8 State discrimination

## 8.1 Minimum error state discrimination

Suppose we know that a system $\mathsf{Q}$ is in the state $\rho(X)$ where $X$ is a random variable taking values in $\mathcal{A}_X$. That is, we have a density operator $\rho(x)$ for each $x \in \mathcal{A}_X$ and we know that, with probability $P_X(x)$, the state of $\mathsf{Q}$ is $\rho(x)$. (For example, perhaps $\mathcal{A}_X = \{0,1,2\}$, $P_X(0) = 2/3, P_X(1) = 1/6, P_X(2) = 1/6$, and the states of $\mathsf{Q}$ are $\rho(0) = |+\rangle\langle+|, \rho(1) = |0\rangle\langle 0|, \rho(2) = |-\rangle\langle-|$, where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$.

We are interested in how well we can determine $X$ by measuring the system. If $\hat{X}$ is the result of measuring some POVM $E : \mathcal{A}_X \to \mathcal{L}(\mathcal{H}_\mathsf{Q})$ then

$$\Pr(\hat{X} = X) = \sum_{x \in \mathcal{A}_X} \Pr(\hat{X} = x, X = x) = \sum_{x \in \mathcal{A}_X} \Pr(X = x) \Pr(\hat{X} = x | X = x) \tag{8.1}$$

$$= \sum_{x \in \mathcal{A}_X} P_X(x) \mathrm{Tr} E(x) \rho(x). \tag{8.2}$$

Maximising this success probability over all POVMs $E$, (i.e. $E$ such that $E(x) \geq 0$ for all $x$ and $\sum_x E(x) = \mathbb{1}$) is a type of optimisation called a *semidefinite program*, which can be solved efficiently on a computer (in time polynomial in $k$ and $d_\mathsf{Q}$) but which does not, in general, have a closed form solution. However, for the case $k = 2$, the **Holevo-Helstrom theorem** gives us a closed form solution for the maximum probability and a POVM which achieves it.

## 8.2 The Holevo-Helstrom theorem

### 8.2.1 Mathematical preliminaries

As usual, here we are assuming that we are dealing with operators on finite dimensional, complex Hilbert spaces.

1. Any operator of the form $L^\dagger L$ is positive. Any positive operator $A$ has a unique positive square root $A^{1/2}$ such that $(A^{1/2})^\dagger A^{1/2} = A$ and $A \geq 0$. If $A = \sum_j \lambda_j |\alpha_j\rangle\langle\alpha_j|$ is an eigendecomposition, then $A^{1/2} = \sum_j \lambda_j^{1/2} |\alpha_j\rangle\langle\alpha_j|$.

2. If operators $A$ and $M$ satisfy $A \geq 0$ and $M \geq 0$ then $\mathrm{Tr} AM = \mathrm{Tr} A^{1/2} M A^{1/2} \geq 0$ because $A^{1/2} M A^{1/2} \geq 0$. But note that $AM$ is not necessarily positive or even hermitian.

3. For any operator $J$, $|J| := (J^\dagger J)^{1/2}$ and the **trace norm** of $J$ is $\|J\|_1 := \mathrm{Tr}|J|$.

4. The support of $J$ is the orthogonal complement of $\ker(J)$.

5. If $A \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$ is hermitian then $A_+$ and $A_-$ are the unique operators such that $A_+ \geq 0$, $A_- \geq 0$, $A_+ - A_- = A$, and $A_+ A_- = 0$.

6. If $A \in \mathcal{L}(\mathcal{H}_\mathsf{Q})$ is hermitian, with eigendecomposition $A = \sum_{j=1}^{d_\mathsf{Q}} \lambda_j |\alpha_j\rangle\langle\alpha_j|$, then
   (a) $A_+ = \sum_{j:\lambda_j>0} \lambda_j |\alpha_j\rangle\langle\alpha_j|$.
   (b) $A_- = -\sum_{j:\lambda_j<0} \lambda_j |\alpha_j\rangle\langle\alpha_j|$ (note the minus sign).
   (c) $|A| = A_+ + A_-$.
   (d) $\|A\|_1 = \sum_{j=1}^{d_\mathsf{Q}} |\lambda_j|$.
   (e) The projector onto $\mathrm{supp}(A_+)$ is $\sum_{j:\lambda_j>0} |\alpha_j\rangle\langle\alpha_j|$.
   (f) The projector onto $\mathrm{supp}(A_-)$ is $\sum_{j:\lambda_j<0} |\alpha_j\rangle\langle\alpha_j|$.
   (g) The projector onto $\ker(A)$ is $\sum_{j:\lambda_j=0} |\alpha_j\rangle\langle\alpha_j|$.
   (h) The projector onto $\mathrm{supp}(A)$ is $\sum_{j:\lambda_j\neq 0} |\alpha_j\rangle\langle\alpha_j|$.

**Proposition 1.** For any hermitian operator $A$, $\|A\|_1 = \max\{\mathrm{Tr}AT : -\mathbb{1} \leq T \leq \mathbb{1}\}$ and the maximum is attained when $T = \Pi_+ + \Pi_0 - \Pi_-$, where $\Pi_+$, $\Pi_-$ and $\Pi_0$ are the projectors onto $\mathrm{supp}(A_+)$, $\mathrm{supp}(A_-)$ and $\ker(A)$ respectively.

*Proof.* The constraints on $T$ are equivalent to $\mathbb{1} - T \geq 0$ and $\mathbb{1} + T \geq 0$. Because $A_+ \geq 0$, $A_- \geq 0$, by item 2 above, for any $T$ which satisfies the constraints we have

$$\mathrm{Tr}AT = \mathrm{Tr}(A_+ + A_-)T \tag{8.3}$$
$$= \mathrm{Tr}A_+ + \mathrm{Tr}A_- - \mathrm{Tr}A_+(\mathbb{1} - T) - \mathrm{Tr}A_-(\mathbb{1} + T) \leq \mathrm{Tr}A_+ + \mathrm{Tr}A_- = \|A\|_1. \tag{8.4}$$

Now let $T' = \Pi_+ + \Pi_0 - \Pi_-$. Since $\mathbb{1} = \Pi_+ + \Pi_0 + \Pi_-$, we have $\mathbb{1} - T' = 2\Pi_- \geq 0$ and $\mathbb{1} + T' = 2\Pi_+ + 2\Pi_0 \geq 0$, so $T'$ satisfies the constraints in the maximisation and has

$$\mathrm{Tr}(A_+ - A_-)T' = \mathrm{Tr}A_+(\Pi_+ + \Pi_0) + \mathrm{Tr}A_-\Pi_- - \mathrm{Tr}A_-(\Pi_+ + \Pi_0) - \mathrm{Tr}A_+\Pi_- \tag{8.5}$$
$$= \mathrm{Tr}A_+ + \mathrm{Tr}A_-. \tag{8.6}$$

□

With this fact in hand, it is relatively easy to give a simple formula for the maximum success probability for discriminating between two states of a system.

### 8.2.2 The Holevo-Helstrom theorem

**Theorem 2** (The Holevo-Helstrom theorem)**.** Suppose we know that a system $\mathsf{Q}$ is in one of two states $\rho(0)$ or $\rho(1)$. If the state is $\rho(X)$ where $X$ is a random variable with $\mathcal{A}_X = \{0, 1\}$, and $\hat{X}$ is the result of measuring some POVM $E : \{0, 1\} \to \mathcal{L}(\mathcal{H}_\mathsf{Q})$ then

$$\max_{\text{POVMs } E} \Pr(\hat{X} = X) = \frac{1}{2}\left(1 + \|\Delta\|_1\right), \text{ where } \Delta = P_X(1)\rho(1) - P_X(0)\rho(0), \tag{8.7}$$

and this is attained by the POVM with $E(1) = \Pi_+ + \Pi_0$, $E(0) = \Pi_-$ where $\Pi_+, \Pi_-$ and $\Pi_0$ are the projectors onto $\mathrm{supp}(\Delta_+)$, $\mathrm{supp}(\Delta_-)$ and $\ker(\Delta)$, respectively.

*Proof.* Since $E(0) + E(1) = \mathbb{1}$, we can write $E(1) = (\mathbb{1} + T)/2$ and $E(0) = (\mathbb{1} - T)/2$ where $T$ is the hermitian operator $E(1) - E(0)$. In terms of $T$,

$$\Pr(\hat{X} = X) = \frac{1}{2}P_X(0)\mathrm{Tr}(\mathbb{1} - T)\rho(0) + \frac{1}{2}P_X(1)\mathrm{Tr}(\mathbb{1} + T)\rho(1) \tag{8.8}$$
$$= \frac{1}{2}\left(P_X(0)\mathrm{Tr}\rho(0) + P_X(1)\mathrm{Tr}\rho(1) + \mathrm{Tr}T\Delta\right) = \frac{1}{2}\left(1 + \mathrm{Tr}T\Delta\right). \tag{8.9}$$

Since $E(1) \geq 0$ iff $T \geq -\mathbb{1}$ and $E(0) \geq 0$ iff $T \leq \mathbb{1}$, maximising over valid POVMs is equivalent to maximising over $T$ such that $-\mathbb{1} \leq T \leq \mathbb{1}$. Therefore, according to Proposition 1, $\max_{\text{POVMs } E} \Pr(\hat{X} = X) = \frac{1}{2}(1 + \|\Delta\|_1)$ which is attained when $T = \Pi_+ + \Pi_0 - \Pi_-$ or, equivalently, when $E(1) = \Pi_+ + \Pi_0$, and $E(0) = \Pi_-$. □

# 8.3 Example

Let $X$ be uniformly distributed bit i.e. $\mathcal{A}_X = \{0, 1\}$, $P_X(0) = P_X(1) = 1/2$. Suppose that system $\mathsf{Q}$ is a qubit, which we know is in the state $\rho(X)$ where

$$\rho(0) = \eta[0] = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \rho(1) = \eta[\pi/2] = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

Here the $\{\eta[\phi] : \phi \in \mathbb{R}\}$ are the pure states on the equator of the Bloch sphere defined in Handout 2. Figure 8.1 shows the position of the two states on the equator of the Bloch sphere.

## 8.3.1 Minimum error state discrimination

First, let us compute the maximum value of $\Pr(\hat{X} = X)$ which can be achieved by measuring a POVM $E : \{0, 1\} \to \mathcal{L}(\mathcal{H}_{\mathsf{Q}})$ to obtain an estimate $\hat{X}$ of $X$. The operator $\Delta$ which appears in the formula (8.7) is

$$\Delta = P_X(1)\rho(1) - P_X(0)\rho(0) = \frac{-1}{4}\begin{pmatrix} 0 & 1+i \\ 1-i & 0 \end{pmatrix}$$

Computing the eigenvectors and eigenvalues of $\Delta$ we find that it has the eigendecomposition

$$\Delta = \frac{1}{2\sqrt{2}}\eta[3\pi/4] - \frac{1}{2\sqrt{2}}\eta[-\pi/4], \quad \text{so } \Delta_- = \frac{1}{2\sqrt{2}}\eta[-\pi/4], \Delta_+ = \frac{1}{2\sqrt{2}}\eta[3\pi/4], \tag{8.10}$$

$$|\Delta| = \frac{1}{2\sqrt{2}}(\eta[-\pi/4] + \eta[3\pi/4]) = \frac{1}{2\sqrt{2}}\mathbb{1}, \text{ and } \|\Delta\|_1 = 1/\sqrt{2}. \tag{8.11}$$

Therefore, the Holevo-Helstrom theorem tells us that the POVM $E'$ with

$$E'(0) = \eta[-\pi/4] \text{ and } E'(1) = \eta[3\pi/4]$$

is optimal and achieves $\Pr(\hat{X} = X) = (1+1/\sqrt{2})/2$. Since the POVM elements are rank-one, trace-one, positive operators we can represent them points on the Bloch sphere in Figure 8.1, along with the states.

## 8.3.2 Unambiguous state discrimination

Now suppose that instead we measure a POVM $F : \{0, 1, ?\} \to \mathcal{L}(\mathcal{H}_{\mathsf{Q}})$ of the form

$$F(0) = a\eta[-\pi/2], \; F(1) = a\eta[\pi], \; F(?) = \mathbb{1} - F(0) - F(1),$$

where $a$ is some positive real number, obtaining a result $Y$. It is easy to compute

$$P_{Y|X}(0|0) = \mathrm{Tr}F(0)\rho(0) = a/2, \; P_{Y|X}(1|0) = \mathrm{Tr}F(1)\rho(0) = 0, \tag{8.12}$$

$$P_{Y|X}(0|1) = \mathrm{Tr}F(0)\rho(1) = 0, \; P_{Y|X}(1|1) = \mathrm{Tr}F(1)\rho(1) = a/2, \tag{8.13}$$

and it follows that

$$P_Y(0) = P_{YX}(0,0) + P_{YX}(0,1) = P_{Y|X}(0|0)P_X(0) + P_{Y|X}(0|1)P_X(1) = a/4, \tag{8.14}$$

$$P_Y(1) = P_{YX}(1,0) + P_{YX}(1,1) = P_{Y|X}(1|0)P_X(0) + P_{Y|X}(1|1)P_X(1) = a/4, \tag{8.15}$$

$$P_Y(?) = 1 - a/2. \tag{8.16}$$

So, if $Y = 0$ or $Y = 1$ then we know *for sure* that $X = Y$, even though the states we are discriminating are not orthogonal! This comes at the expense of having probability $1 - a/2$ that $Y = ?$. How large can we make $a$? The only constraint is that $F(?) \geq 0$ which is true iff

$$a(\eta[-\pi/2] + \eta[\pi]) \leq \mathbb{1}.$$

This is equivalent to saying that the largest eigenvalue of $a(\eta[-\pi/2] + \eta[\pi])$ must be less than or equal to one. The characteristic polynomial of

$$\eta[-\pi/2] + \eta[\pi] = \begin{pmatrix} 1 & (i-1)/2 \\ -(1+i)/2 & 1 \end{pmatrix}$$

is $(1-\lambda)^2 - (i-1)(i+1)/4 = 0$ so the eigenvalues are $\lambda_1 = 1 - 1/\sqrt{2}$ and $\lambda_2 = 1 + 1/\sqrt{2}$. So, the largest we can make $a$ is $1/(1 + 1/\sqrt{2}) \approx 0.59$ for which $P_Y(?) = 1/\sqrt{2} \approx 0.71$.

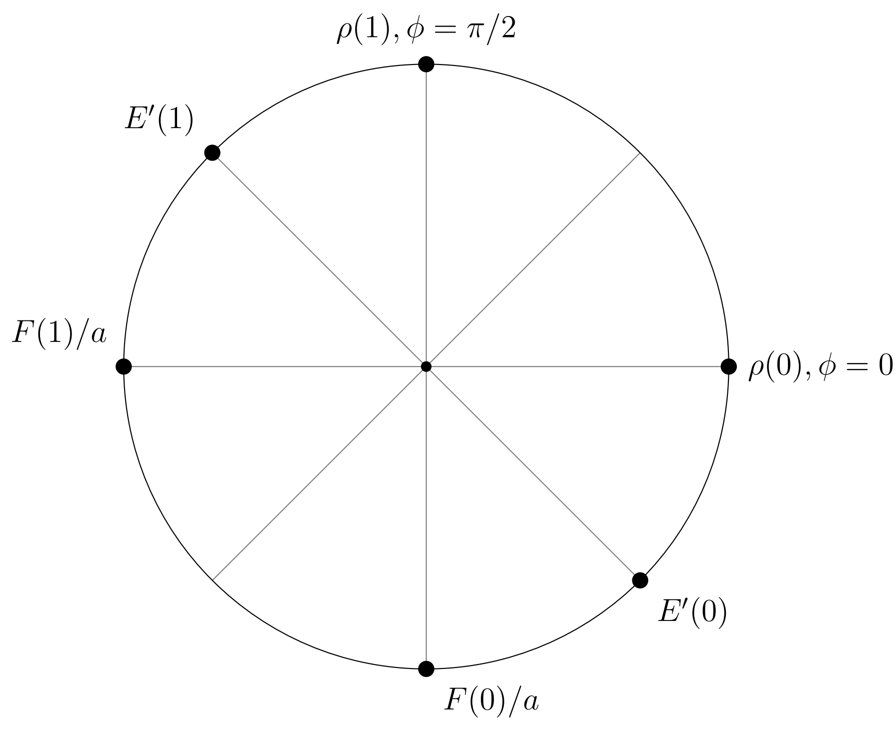


Figure 8.1: States and POVM elements described in the Section 8.3.