CS 766/QIC 820 Theory of Quantum Information (Fall 2011)

Lecture 8: Semidefinite programs for fidelity and optimal measurements

This lecture is devoted to two examples of semidefinite programs: one is for the fidelity between two positive semidefinite operators, and the other is for optimal measurements for distinguishing ensembles of states. The primary goal in studying these examples at this point in the course is to gain familiarity with the concept of semidefinite programming and how it may be applied to problems of interest. The examples themselves are interesting, but they should not necessarily be viewed as primary reasons for studying semidefinite programming—they are simply examples making use of concepts we have discussed thus far in the course. We will see further applications of semidefinite programming to quantum information theory later in the course, and there are many more applications that we will not discuss.

8.1 A semidefinite program for the fidelity function

We begin with a semidefinite program whose optimal value equals the fidelity between to given positive semidefinite operators. As it represents the first application of semidefinite programming to quantum information theory that we are studying in the course, we will go through it in some detail.

8.1.1 Specification of the semidefinite program

Suppose $P, Q \in \operatorname{Pos}(\mathcal{X})$, where $\mathcal{X}$ is a complex Euclidean space, and consider the following optimization problem:

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\operatorname{Tr}(X) + \frac{1}{2}\operatorname{Tr}(X^*) \\ \text{subject to:} \quad & \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \geq 0 \\ & X \in \mathrm{L}(\mathcal{X}). \end{aligned}$$

Although it is not phrased in the precise form of a semidefinite program as we formally defined them in the previous lecture, it can be converted to one, as we will now see.

Let us begin by noting that the matrix

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}$$

is a block matrix that describes an operator in the space $\mathrm{L}(\mathcal{X} \oplus \mathcal{X})$. To phrase the optimization problem above as a semidefinite program, we will effectively optimize over all positive semidefinite operators in $\operatorname{Pos}(\mathcal{X} \oplus \mathcal{X})$, using linear constraints to force the diagonal blocks to be P and Q.

With this idea in mind, we define a linear mapping $\Phi : \mathrm{L}(\mathcal{X} \oplus \mathcal{X}) \rightarrow \mathrm{L}(\mathcal{X} \oplus \mathcal{X})$ as follows:

$$\Phi \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} = \begin{pmatrix} X_{1,1} & 0 \\ 0 & X_{2,2} \end{pmatrix}$$

for all choices of $X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2} \in \mathrm{L}(\mathcal{X})$, and we define $A, B \in \mathrm{Herm}(\mathcal{X} \oplus \mathcal{X})$ as

$$A = \frac{1}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}.$$

Now consider the semidefinite program (Φ, A, B), as defined in the previous lecture. The primal objective function takes the form

$$\left\langle \frac{1}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right\rangle = \frac{1}{2} \operatorname{Tr}(X_{1,2}) + \frac{1}{2} \operatorname{Tr}(X_{2,1}).$$

The constraint

$$\Phi \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$$

is equivalent to the conditions $X_{1,1} = P$ and $X_{2,2} = Q$. Of course, the condition

$$\begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X})$$

forces $X_{2,1} = X_{1,2}^*$, as this follows from the Hermiticity of the operator. So, by writing X in place of $X_{1,2}$, we see that the optimization problem stated at the beginning of the section is equivalent to the primal problem associated with (Φ, A, B).

Now let us examine the dual problem. It is as follows:

$$\begin{aligned} \text{minimize:} \quad & \left\langle \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}, \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix} \right\rangle \\ \text{subject to:} \quad & \Phi^* \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix} \geq \frac{1}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \\ & \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix} \in \mathrm{Herm}(\mathcal{X} \oplus \mathcal{X}). \end{aligned}$$

As is typical when trying to understand the relationship between the primal and dual problems of a semidefinite program, we must find an expression for Φ^*. This happens to be easy in the present case, for we have

$$\left\langle \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix}, \Phi \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix}, \begin{pmatrix} X_{1,1} & 0 \\ 0 & X_{2,2} \end{pmatrix} \right\rangle = \langle Y_{1,1}, X_{1,1} \rangle + \langle Y_{2,2}, X_{2,2} \rangle$$

and

$$\left\langle \Phi \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix}, \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} Y_{1,1} & 0 \\ 0 & Y_{2,2} \end{pmatrix}, \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} \right\rangle = \langle Y_{1,1}, X_{1,1} \rangle + \langle Y_{2,2}, X_{2,2} \rangle,$$

so it must hold that $\Phi^* = \Phi$. Simplifying the above problem accordingly yields

$$\begin{aligned} \text{minimize:} \quad & \langle P, Y_{1,1} \rangle + \langle Q, Y_{2,2} \rangle \\ \text{subject to:} \quad & \begin{pmatrix} Y_{1,1} & 0 \\ 0 & Y_{2,2} \end{pmatrix} \geq \frac{1}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \\ & Y_{1,1}, Y_{2,2} \in \operatorname{Herm}(\mathcal{X}). \end{aligned}$$

The problem has no dependence whatsoever on $Y_{1,2}$ and $Y_{2,1}$, so we can ignore them. Let us write $Y = 2Y_{1,1}$ and $Z = 2Y_{2,2}$, so that the problem becomes

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2} \langle P, Y \rangle + \frac{1}{2} \langle Q, Z \rangle \\ \text{subject to:} \quad & \begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \geq 0 \\ & Y, Z \in \operatorname{Herm}(\mathcal{X}). \end{aligned}$$

There is no obvious reason for including the factor of 2 in the specification of Y and Z; it is simply a change of variables that is designed to put the problem into a nicer form for the analysis to come later. The inclusion of the factor of 2 does not, of course, change the fact that Y and Z are free to range over all Hermitian operators.

In summary, we have this pair of problems:

Primal problem

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2} \operatorname{Tr}(X) + \frac{1}{2} \operatorname{Tr}(X^*) \\ \text{subject to:} \quad & \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \geq 0 \\ & X \in \mathrm{L}(\mathcal{X}). \end{aligned}$$

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2} \langle P, Y \rangle + \frac{1}{2} \langle Q, Z \rangle \\ \text{subject to:} \quad & \begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \geq 0 \\ & Y, Z \in \operatorname{Herm}(\mathcal{X}). \end{aligned}$$

We will make some further simplifications to the dual problem a bit later in the lecture, but let us leave it as it is for the time being.

The statement of the primal and dual problems just given is representative of a typical style for specifying semidefinite programs: generally one does not explicitly refer to Φ, A, and B, or operators and mappings coming from other specific forms of semidefinite programs, in applications of the concept in papers or talks. It would not be unusual to see a pair of primal and dual problems presented like this without any indication of how the dual problem was obtained from the primal problem (or vice-versa). This is because the process is more or less routine, once you know how it is done. (Until you've had some practise doing it, however, it may not seem that way.)

8.1.2 Optimal value

Let us observe that strong duality holds for the semidefinite program above. This is easily established by first observing that the primal problem is feasible and the dual problem is strictly feasible, then applying Slater's theorem. To do this formally, we must refer to the triple (Φ, A, B) discussed above. Setting

$$\begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$$

gives a primal feasible operator, so that $\mathcal{A} \neq \varnothing$. Setting

$$\begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}$$

gives

$$\Phi^* \begin{pmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} > \frac{1}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix},$$

owing to the fact that

$$\begin{pmatrix} \mathbb{1} & -\frac{1}{2}\mathbb{1} \\ -\frac{1}{2}\mathbb{1} & \mathbb{1} \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} \otimes \mathbb{1}$$

is positive definite. By Slater's theorem, we have strong duality, and moreover the optimal primal value is achieved by some choice of X.

It so happens that strict primal feasibility may fail to hold: if either of P or Q is not positive definite, it cannot hold that

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} > 0.$$

Note, however, that we cannot conclude from this fact that the optimal dual value will not be achieved—but indeed this is the case for some choices of P and Q. If P and Q are positive definite, strict primal feasibility does hold, and the optimal dual value will be achieved, as follows from Slater's theorem.

Now let us prove that the optimal value is equal to $\mathrm{F}(P, Q)$, beginning with the inequality $\alpha \geq \mathrm{F}(P, Q)$. To prove this inequality, it suffices to exhibit a primal feasible X for which

$$\frac{1}{2} \operatorname{Tr}(X) + \frac{1}{2} \operatorname{Tr}(X^*) = \mathrm{F}(P, Q).$$

We have

$$\mathrm{F}(P, Q) = \mathrm{F}(Q, P) = \left\| \sqrt{Q}\sqrt{P} \right\|_1 = \max \left\{ \left| \operatorname{Tr}\left(U\sqrt{Q}\sqrt{P} \right) \right| \,:\, U \in \mathrm{U}(\mathcal{X}) \right\},$$

and so we may choose a unitary operator $U \in \mathrm{U}(\mathcal{X})$ for which

$$\mathrm{F}(P, Q) = \operatorname{Tr}\left(U\sqrt{Q}\sqrt{P} \right) = \operatorname{Tr}\left(\sqrt{P}U\sqrt{Q} \right).$$

(The absolute value can safely be omitted: we are free to multiply any U maximizing the absolute value with a scalar on the unit circle, obtaining a nonnegative real number for the trace.) Now define

$$X = \sqrt{P}U\sqrt{Q}.$$

It holds that

$$0 \leq \begin{pmatrix} \sqrt{P} & U\sqrt{Q} \end{pmatrix}^* \begin{pmatrix} \sqrt{P} & U\sqrt{Q} \end{pmatrix} = \begin{pmatrix} \sqrt{P} \\ \sqrt{Q}U^* \end{pmatrix} \begin{pmatrix} \sqrt{P} & U\sqrt{Q} \end{pmatrix} = \begin{pmatrix} P & \sqrt{P}U\sqrt{Q} \\ \sqrt{Q}U^*\sqrt{P} & Q \end{pmatrix},$$

so X is primal feasible, and we have

$$\frac{1}{2} \operatorname{Tr}(X) + \frac{1}{2} \operatorname{Tr}(X^*) = \mathrm{F}(P, Q)$$

as claimed.

Now let us prove the reverse inequality: $\alpha \leq \mathrm{F}(P,Q)$. Suppose that $X \in \mathrm{L}(\mathcal{X})$ is primal feasible, meaning that

$$R = \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}$$

is positive semidefinite. We may view that $R \in \mathrm{Pos}(\mathcal{Z} \otimes \mathcal{X})$ for $\mathcal{Z} = \mathbb{C}^2$. (More generally, the m-fold direct sum $\mathbb{C}^\Sigma \oplus \cdots \oplus \mathbb{C}^\Sigma$ may be viewed as being equivalent to the tensor product $\mathbb{C}^m \otimes \mathbb{C}^\Sigma$ by identifying the standard basis element $e_{(j,a)}$ of $\mathbb{C}^\Sigma \oplus \cdots \oplus \mathbb{C}^\Sigma$ with the standard basis element $e_j \otimes e_a$ of $\mathbb{C}^m \otimes \mathbb{C}^\Sigma$, for each $j \in \{1,\ldots,m\}$ and $a \in \Sigma$.) Let $\mathcal{Y}$ be a complex Euclidean space whose dimension is at least $\operatorname{rank}(R)$, and let $u \in \mathcal{Z} \otimes \mathcal{X} \otimes \mathcal{Y}$ be a purification of R:

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = R = E_{1,1} \otimes P + E_{1,2} \otimes X + E_{2,1} \otimes X^* + E_{2,2} \otimes Q.$$

Write

$$u = e_1 \otimes u_1 + e_2 \otimes u_2$$

for $u_1, u_2 \in \mathcal{X}$, and observe that

$$\mathrm{Tr}_{\mathcal{Y}}(u_1 u_1^*) = P, \qquad \mathrm{Tr}_{\mathcal{Y}}(u_2 u_2^*) = Q, \qquad \mathrm{Tr}_{\mathcal{Y}}(u_1 u_2^*) = X, \qquad \text{and} \qquad \mathrm{Tr}_{\mathcal{Y}}(u_2 u_1^*) = X^*.$$

We have

$$\frac{1}{2}\mathrm{Tr}(X) + \frac{1}{2}\mathrm{Tr}(X^*) = \frac{1}{2}\langle u_2, u_1\rangle + \frac{1}{2}\langle u_1, u_2\rangle = \Re(\langle u_1, u_2\rangle) \leq |\langle u_1, u_2\rangle| \leq \mathrm{F}(P,Q),$$

where the last inequality follows from Uhlmann's theorem, along with the fact that u_1 and u_2 purify P and Q, respectively.

8.1.3 Alternate proof of Alberti's theorem

The notes from Lecture 4 include a proof of Alberti's theorem, which states that

$$(\mathrm{F}(P,Q))^2 = \inf_{Y \in \mathrm{Pd}(\mathcal{X})} \langle P, Y\rangle \langle Q, Y^{-1}\rangle,$$

for every choice of positive semidefinite operators $P, Q \in \mathrm{Pos}(\mathcal{X})$. We may use our semidefinite program to obtain an alternate proof of this characterization.

First let us return to the dual problem from above:

$$\underline{\text{Dual problem}}$$

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\langle P, Y\rangle + \frac{1}{2}\langle Q, Z\rangle \\ \text{subject to:} \quad & \begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \geq 0 \\ & Y, Z \in \mathrm{Herm}(\mathcal{X}). \end{aligned}$$

To simplify the problem further, let us prove the following claim.

Claim 8.1. Let $Y, Z \in \mathrm{Herm}(\mathcal{X})$. It holds that

$$\begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{X})$$

if and only if $Y, Z \in \mathrm{Pd}(\mathcal{X})$ and $Z \geq Y^{-1}$.

Proof. Suppose $Y, Z \in \operatorname{Pd}(\mathcal{X})$ and $Z \geq Y^{-1}$. It holds that

$$\begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ -Y^{-1} & \mathbb{1} \end{pmatrix} \begin{pmatrix} Y & 0 \\ 0 & Z - Y^{-1} \end{pmatrix} \begin{pmatrix} \mathbb{1} & -Y^{-1} \\ 0 & \mathbb{1} \end{pmatrix}$$

and therefore

$$\begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{X}).$$

Conversely, suppose that

$$\begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{X}).$$

It holds that

$$0 \leq \begin{pmatrix} u \\ v \end{pmatrix}^{*} \begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = u^* Y u - u^* v - v^* u + v^* Z v$$

for all $u, v \in \mathcal{X}$. If Y were not positive definite, there would exist a unit vector v for which $v^* Y v = 0$, and one could then set

$$u = \frac{1}{2}(\|Z\| + 1)v$$

to obtain

$$\|Z\| \geq v^* Z v \geq \langle u, v\rangle + \langle v, u\rangle = \|Z\| + 1,$$

which is absurd. Thus, $Y \in \operatorname{Pd}(\mathcal{X})$. Finally, by inverting the expression above, we have

$$\begin{pmatrix} Y & 0 \\ 0 & Z - Y^{-1} \end{pmatrix} = \begin{pmatrix} \mathbb{1} & 0 \\ Y^{-1} & \mathbb{1} \end{pmatrix} \begin{pmatrix} Y & -\mathbb{1} \\ -\mathbb{1} & Z \end{pmatrix} \begin{pmatrix} \mathbb{1} & Y^{-1} \\ 0 & \mathbb{1} \end{pmatrix} \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{X}),$$

which implies $Z \geq Y^{-1}$ (and therefore $Z \in \operatorname{Pd}(\mathcal{X})$) as required. □

Now, given that Q is positive semidefinite, it holds that $\langle Q, Z\rangle \geq \langle Q, Y^{-1}\rangle$ whenever $Z \geq Y^{-1}$, so there would be no point in choosing any Z other than Y^{-1} when aiming to minimize the dual objective function subject to that constraint. The dual problem above can therefore be phrased as follows:

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\langle P, Y\rangle + \frac{1}{2}\langle Q, Y^{-1}\rangle \\ \text{subject to:} \quad & Y \in \operatorname{Pd}(\mathcal{X}). \end{aligned}$$

Given that strong duality holds for our semidefinite program, and that we know the optimal value to be $\mathrm{F}(P, Q)$, we have the following theorem.

Theorem 8.2. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \operatorname{Pos}(\mathcal{X})$. It holds that*

$$\mathrm{F}(P, Q) = \inf \left\{ \frac{1}{2}\langle P, Y\rangle + \frac{1}{2}\langle Q, Y^{-1}\rangle \,:\, Y \in \operatorname{Pd}(\mathcal{X}) \right\}.$$

To see that this is equivalent to Alberti's theorem, note that for every $Y \in \text{Pd}(\mathcal{X})$ it holds that

$$\frac{1}{2}\langle P, Y\rangle + \frac{1}{2}\langle Q, Y^{-1}\rangle \geq \sqrt{\langle P, Y\rangle\langle Q, Y^{-1}\rangle},$$

with equality if and only if $\langle P, Y\rangle = \langle Q, Y^{-1}\rangle$ (by the arithmetic-geometric mean inequality). It follows that

$$\inf_{Y\in\text{Pd}(\mathcal{X})} \langle P, Y\rangle \langle Q, Y^{-1}\rangle \leq (\text{F}(P,Q))^2.$$

Moreover, for an arbitrary choice of $Y \in \text{Pd}(\mathcal{X})$, one may choose $\lambda > 0$ so that

$$\langle P, \lambda Y\rangle = \langle Q, (\lambda Y)^{-1}\rangle$$

and therefore

$$\frac{1}{2}\langle P, \lambda Y\rangle + \frac{1}{2}\langle Q, (\lambda Y)^{-1}\rangle = \sqrt{\langle P, \lambda Y\rangle\langle Q, (\lambda Y)^{-1}\rangle} = \sqrt{\langle P, Y\rangle\langle Q, Y^{-1}\rangle}.$$

Thus,

$$\inf_{Y\in\text{Pd}(\mathcal{X})} \langle P, Y\rangle \langle Q, Y^{-1}\rangle \geq (\text{F}(P,Q))^2.$$

We therefore have that Alberti's theorem is a corollary to the theorem above, as claimed.

Theorem 8.3 (Alberti)**.** *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \text{Pos}(\mathcal{X})$. It holds that*

$$(\text{F}(P,Q))^2 = \inf_{Y\in\text{Pd}(\mathcal{X})} \langle P, Y\rangle \langle Q, Y^{-1}\rangle.$$

8.2 Optimal measurements

We will now move on to the second example of the lecture of a semidefinite programming application to quantum information theory. This example concerns the notion of optimal measurements for distinguishing elements of an ensemble of states.

Suppose that $\mathcal{X}$ is a complex Euclidean space, Γ is a finite and nonempty set, $p \in \mathbb{R}^\Gamma$ is a probability vector, and $\{\rho_a : a \in \Gamma\} \subset \text{D}(\mathcal{X})$ is a collection of density operators. Consider the scenario in which Alice randomly selects $a \in \Gamma$ according to the probability distribution described by p, then prepares a register X in the state ρ_a for whichever element $a \in \Gamma$ she selected. She sends X to Bob, whose goal is to identify the element $a \in \Gamma$ selected by Alice with as high a probability as possible. He must do this by means of a measurement $\mu : \Gamma \rightarrow \text{Pos}(\mathcal{X}) : a \mapsto P_a$ on X, without any additional help or input from Alice. Bob's optimal probability is given by the maximum value of

$$\sum_{a\in\Gamma} p(a) \langle P_a, \rho_a\rangle$$

over all measurements $\mu : \Gamma \rightarrow \text{Pos}(\mathcal{X}) : a \mapsto P_a$ on $\mathcal{X}$.

It is natural to associate an *ensemble* of states with the process performed by Alice. This is a collection

$$\mathcal{E} = \{(p(a), \rho_a) : a \in \Gamma\},$$

which can be described more succinctly by a mapping

$$\eta : \Gamma \rightarrow \text{Pos}(\mathcal{X}) : a \mapsto \sigma_a,$$

where $\sigma_a = p(a)\rho_a$ for each $a \in \Gamma$. In general, any mapping η of the above form represents an ensemble if and only if

$$\sum_{a \in \Gamma} \sigma_a \in \mathrm{D}(\mathcal{X}).$$

To recover the description of a collection $\mathcal{E} = \{(p(a), \rho_a) : a \in \Gamma\}$ representing such an ensemble, one may take $p(a) = \operatorname{Tr}(\sigma_a)$ and $\rho_a = \sigma_a / \operatorname{Tr}(\sigma_a)$. Thus, each σ_a is generally not a density operator, but may be viewed as an unnormalized density operator that describes both a density operator and the probability that it is selected.

Now, let us say that a measurement $\mu : \Gamma \to \operatorname{Pos}(\mathcal{X})$ is an *optimal* measurement for a given ensemble $\eta : \Gamma \to \operatorname{Pos}(\mathcal{X})$ if and only if it holds that

$$\sum_{a \in \Gamma} \langle \mu(a), \eta(a) \rangle$$

is maximal among all possible choices of measurements that could be substituted for μ in this expression. We will prove the following theorem, which provides a simple condition (both necessary and sufficient) for a given measurement to be optimal for a given ensemble.

Theorem 8.4. *Let $\mathcal{X}$ be a complex Euclidean space, let Γ be a finite and nonempty set, let $\eta : \Gamma \to \operatorname{Pos}(\mathcal{X}) : a \mapsto \sigma_a$ be an ensemble of states, and let $\mu : \Gamma \to \operatorname{Pos}(\mathcal{X}) : a \mapsto P_a$ be a measurement. It holds that μ is optimal for η if and only if the operator*

$$Y = \sum_{a \in \Gamma} \sigma_a P_a$$

is Hermitian and satisfies $Y \geq \sigma_a$ for each $a \in \Gamma$.

The following proposition, which states a property known as *complementary slackness* for semidefinite programs, will be used to prove the theorem.

Proposition 8.5 (Complementary slackness for SDPs)**.** *Suppose (Φ, A, B) is a semidefinite program, and that $X \in \mathcal{A}$ and $Y \in \mathcal{B}$ satisfy $\langle A, X \rangle = \langle B, Y \rangle$. It holds that*

$$\Phi^*(Y)X = AX \qquad \textit{and} \qquad \Phi(X)Y = BY.$$

Remark 8.6. Note that the second equality stated in the proposition is completely trivial, given that $\Phi(X) = B$ for all $X \in \mathcal{A}$. It is stated nevertheless in the interest of illustrating the symmetry between the primal and dual forms of semidefinite programs.

Proof. It holds that

$$\langle A, X \rangle = \langle B, Y \rangle = \langle \Phi(X), Y \rangle = \langle \Phi^*(Y), X \rangle,$$

so

$$\langle \Phi^*(Y) - A, X \rangle = 0.$$

Both $\Phi^*(Y) - A$ and X are positive semidefinite, given that X and Y are feasible. The inner product of two positive semidefinite operators is zero if and only if their product is zero, and so we obtain

$$(\Phi^*(Y) - A)\, X = 0.$$

This implies the first equality in the proposition, as required. □

Next, we will phrase the problem of maximizing the probability of correctly identifying the states in an ensemble as a semidefinite program. We suppose that an ensemble

$$\eta : \Gamma \rightarrow \mathrm{Pos}\left(\mathcal{X}\right) : a \mapsto \sigma_a$$

is given, and define a semidefinite program as follows. Let $\mathcal{Y} = \mathbb{C}^{\Gamma}$, let $A \in \mathrm{Herm}\left(\mathcal{Y} \otimes \mathcal{X}\right)$ be given by

$$A = \sum_{a \in \Gamma} E_{a,a} \otimes \sigma_a,$$

and consider the partial trace $\mathrm{Tr}_{\mathcal{Y}}$ as an element of $\mathrm{T}\left(\mathcal{Y} \otimes \mathcal{X}, \mathcal{X}\right)$. The semidefinite program to be considered is $(\mathrm{Tr}_{\mathcal{Y}}, A, \mathbb{1}_{\mathcal{X}})$, and with it one associates the following problems:

Primal problem		Dual problem	
maximize:	$\langle A, X\rangle$	minimize:	$\mathrm{Tr}(Y)$
subject to:	$\mathrm{Tr}_{\mathcal{Y}}(X) = \mathbb{1}_{\mathcal{X}}$,	subject to:	$\mathbb{1}_{\mathcal{Y}} \otimes Y \geq A$
	$X \in \mathrm{Pos}\left(\mathcal{Y} \otimes \mathcal{X}\right)$.		$Y \in \mathrm{Herm}\left(\mathcal{X}\right)$.

To see that the primal problem represents the optimization problem we are interested in, which is the maximization of

$$\sum_{a \in \Gamma} \langle P_a, \sigma_a \rangle = \sum_{a \in \Gamma} \langle \sigma_a, P_a \rangle$$

over all measurements $\{P_a \,:\, a \in \Gamma\}$, we note that any $X \in \mathrm{L}\left(\mathcal{Y} \otimes \mathcal{X}\right)$ may be written

$$X = \sum_{a,b \in \Gamma} E_{a,b} \otimes X_{a,b}$$

for $\{X_{a,b} \,:\, a, b \in \Gamma\} \subset \mathrm{L}\left(\mathcal{X}\right)$, that the objective function is then given by

$$\langle A, X \rangle = \sum_{a \in \Gamma} \langle \sigma_a, X_{a,a} \rangle$$

and that the constraint $\mathrm{Tr}_{\mathcal{Y}}(X) = \mathbb{1}_{\mathcal{X}}$ is given by

$$\sum_{a \in \Gamma} X_{a,a} = \mathbb{1}_{\mathcal{X}}.$$

As X ranges over all positive semidefinite operators in $\mathrm{Pos}\left(\mathcal{Y} \otimes \mathcal{X}\right)$, the operators $X_{a,a}$ individually and independently range over all possible positive semidefinite operators in $\mathrm{Pos}\left(\mathcal{X}\right)$. The "off-diagonal" operators $X_{a,b}$, for $a \neq b$, have no influence on the problem at all, and can safely be ignored. Writing P_a in place of $X_{a,a}$, we see that the primal problem can alternately be written

Primal problem	
maximize:	$\sum_{a \in \Gamma} \langle \sigma_a, P_a \rangle$
subject to:	$\{P_a \,:\, a \in \Gamma\} \subset \mathrm{Pos}\left(\mathcal{X}\right)$
	$\sum_{a \in \Gamma} P_a = \mathbb{1}_{\mathcal{X}}$,

which is the optimization problem of interest.

The dual problem can be simplified by noting that the constraint

$$\mathbb{1}_{\mathcal{Y}} \otimes Y \geq A$$

is equivalent to

$$\sum_{a \in \Gamma} E_{a,a} \otimes (Y - \sigma_a) \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{X}),$$

which in turn is equivalent to $Y \geq \sigma_a$ for each $a \in \Gamma$.

To summarize, we have the following pair of optimization problems:

Primal problem

$$\begin{array}{rl} \text{maximize:} & \displaystyle\sum_{a \in \Gamma} \langle \sigma_a, P_a \rangle \\ \text{subject to:} & \{P_a \,:\, a \in \Gamma\} \subset \operatorname{Pos}(\mathcal{X}) \\ & \displaystyle\sum_{a \in \Gamma} P_a = \mathbb{1}_{\mathcal{X}}, \end{array}$$

Dual problem

$$\begin{array}{rl} \text{minimize:} & \operatorname{Tr}(Y) \\ \text{subject to:} & Y \geq \sigma_a \quad (\text{for all } a \in \Gamma) \\ & Y \in \operatorname{Herm}(\mathcal{X}). \end{array}$$

Strict feasibility is easy to show for this semidefinite program: we may take

$$X = \frac{1}{|\Gamma|} \mathbb{1}_{\mathcal{Y}} \otimes \mathbb{1}_{\mathcal{X}} \qquad \text{and} \qquad Y = 2\mathbb{1}_{\mathcal{X}}$$

to obtain strictly feasible primal and dual solutions. By Slater's theorem, we have strong duality, and moreover that optimal values are always achieved in both problems.

We are now in a position to prove Theorem 8.4. Suppose first that the measurement μ is optimal for η, so that $\{P_a \,:\, a \in \Gamma\}$ is optimal for the semidefinite program above. Somewhat more formally, we have that

$$X = \sum_{a \in \Gamma} E_{a,a} \otimes P_a$$

is an optimal primal solution to the semidefinite program $(\operatorname{Tr}_{\mathcal{Y}}, A, \mathbb{1}_{\mathcal{X}})$. Take Z to be any optimal solution to the dual problem, which we know exists because the optimal solution is always achievable for both the primal and dual problems. By complementary slackness (i.e., Proposition 8.5) it holds that

$$\operatorname{Tr}^*_{\mathcal{Y}}(Z) X = AX,$$

which expands to

$$\sum_{a \in \Gamma} E_{a,a} \otimes Z P_a = \sum_{a \in \Gamma} E_{a,a} \otimes \sigma_a P_a,$$

implying

$$Z P_a = \sigma_a P_a$$

for each $a \in \Gamma$. Summing over $a \in \Gamma$ yields

$$Z = \sum_{a \in \Gamma} \sigma_a P_a = Y.$$

It therefore holds that Y is dual feasible, implying that Y is Hermitian and satisfies $Y \geq \sigma_a$ for each $a \in \Gamma$.

Conversely, suppose that Y is Hermitian and satisfies $Y \geq \sigma_a$ for each $a \in \Gamma$. This means that Y is dual feasible. Given that

$$\operatorname{Tr}(Y) = \sum_{a \in \Gamma} \langle \sigma_a, P_a \rangle ,$$

we find that $\{P_a : a \in \Gamma\}$ must be an optimal primal solution by weak duality, as it equals the value achieved by a dual feasible solution. The measurement μ is therefore optimal for the ensemble η.