Lecture 1

Conic Programming

The first topic we will discuss in the course is *conic programming*, which is a valuable tool for the study of quantum information. In particular, *semidefinite programs*, which are a specific type of conic program, have proven to be particularly useful in the theory of quantum information and computation—and you may already be familiar with some of their applications. There is, however, much to be gained in considering conic programs in greater generality. A few points in support of this claim follow.

1. Conic programming offers a formulation through which fundamental concepts in convex analysis may be conveniently expressed and analyzed. These fundamental concepts can offer valuable insights into semidefinite programs that might otherwise be easily obscured by technical details.

2. Certain key properties of semidefinite programs are, in fact, possessed by conic programs defined over a wide variety of convex cones. A noteworthy example is that *Slater's theorem*, which provides a simple-to-check condition for the critically important property of *strong duality* for many semidefinite programs, generalizes to conic programs.

3. While not all properties of semidefinite programs generalize to conic programs, an understanding of conic programming serves to illuminate the specific attributes of the cone of positive semidefinite operators that have allowed for these properties to hold in the semidefinite programming case.

4. The generality offered by conic programming will be of use in this course.

One must appreciate that semidefinite programs do have a very special property that contributes enormously to their utility, which is that one can generally solve a given semidefinite program with reasonable efficiency and precision using a (classical!) computer. Conic programs, in contrast, are in general hard to

solve with a computer. For example, maximizing a linear function over the set $\mathrm{SepD}(\mathbb{C}^n : \mathbb{C}^n)$ of *bipartite separable density operators* with local dimension n, which is easily expressed as a conic program, is an NP-hard optimization problem, even to approximate with a modest degree of precision.

Having sufficient motivation (I presume) for a study of conic programming, we will now proceed to such a study.

1.1 Preliminaries

This preliminary subsection defines various notions and discuss a few known facts (without proofs) that will be needed for a proper treatment of conic programming, mostly relating to convex analysis.

Let $\mathcal{V}$ be a finite-dimensional real inner product space, with the inner product of any two vectors $u, v \in \mathcal{V}$ being denoted $\langle u, v \rangle$. Note that this inner product is necessarily symmetric in its arguments, given that $\mathcal{V}$ is a vector space over the *real* numbers $\mathbb{R}$, which will be our ground field throughout this entire discussion of conic programming.

A subset $\mathcal{C} \subseteq \mathcal{V}$ is *convex* if, for all $u, v \in \mathcal{C}$ and $\lambda \in [0, 1]$, one has

$$\lambda u + (1 - \lambda) v \in \mathcal{C}. \tag{1.1}$$

A subset $\mathcal{K} \subseteq \mathcal{V}$ is a *cone* if, for all $u \in \mathcal{K}$ and $\lambda \geq 0$, one has $\lambda u \in \mathcal{K}$. We will be principally concerned with subsets having both properties simultaneously, which are aptly named *convex cones*. The letters $\mathcal{K}$ and $\mathcal{L}$ are typical names for convex cones, and we will often make the additional assumption that these cones are *closed* when we are discussing conic programs.

The Cartesian product of any two convex sets is convex. Explicitly, if $\mathcal{C}$ and $\mathcal{D}$ are convex, and $(v_0, w_0), (v_1, w_1) \in \mathcal{C} \times \mathcal{D}$ and $\lambda \in [0, 1]$, then

$$\lambda(v_0, w_0) + (1 - \lambda)(v_1, w_1) = (\lambda v_0 + (1 - \lambda) v_1, \lambda w_0 + (1 - \lambda) w_1) \in \mathcal{C} \times \mathcal{D}. \tag{1.2}$$

Similarly, the Cartesian product of any two cones is a cone: if $\mathcal{K}$ and $\mathcal{L}$ are cones, and $(u, v) \in \mathcal{K} \times \mathcal{L}$ and $\lambda \geq 0$, then

$$\lambda(u, v) = (\lambda u, \lambda v) \in \mathcal{K} \times \mathcal{L}. \tag{1.3}$$

The notion of *separating hyperplane* is fundamental within convex analysis. Here is one form of the *separating hyperplane theorem*, which establishes that for any two disjoint, nonempty, convex sets, there is a hyperplane that separates the two convex sets, with one lying within one of the two closed half-spaces defined by the hyperplane and the second set lying within the opposite closed half-space.

Theorem 1.1 (Separating hyperplane theorem). *Let $\mathcal{V}$ be a finite-dimensional real inner-product space and let $\mathcal{C}$ and $\mathcal{D}$ be nonempty, disjoint, convex subsets of $\mathcal{V}$. There exists a nonzero vector $w \in \mathcal{V}$ and a real number $\gamma \in \mathbb{R}$ such that*

$$\langle w, u \rangle \leq \gamma \leq \langle w, v \rangle \tag{1.4}$$

for every $u \in \mathcal{C}$ and $v \in \mathcal{D}$. If either of $\mathcal{C}$ or $\mathcal{D}$ is a cone, then there must exist a nonzero vector $w \in \mathcal{V}$ as above for which the inequality (1.4) is true, for all $u \in \mathcal{C}$ and $v \in \mathcal{D}$, when $\gamma = 0$.

Given any set $\mathcal{A} \subseteq \mathcal{V}$, one defines the *dual cone* to $\mathcal{A}$ as

$$\mathcal{A}^* = \big\{ v \in \mathcal{V} \,:\, \langle u, v \rangle \geq 0 \text{ for all } u \in \mathcal{A} \big\}. \tag{1.5}$$

The fact that $\mathcal{A}^*$ is indeed a cone, and is also closed and convex, irrespective of the choice of the set $\mathcal{A}$, can be verified. For any two cones $\mathcal{K}, \mathcal{L} \subseteq \mathcal{V}$, it is the case that

$$(\mathcal{K} \times \mathcal{L})^* = \mathcal{K}^* \times \mathcal{L}^*. \tag{1.6}$$

Finally, if $\mathcal{K}$ is a closed, convex cone, then $\mathcal{K}^{**} = \mathcal{K}$.

1.2 Definitions

Now suppose that a finite-dimensional real inner-product space $\mathcal{V}$ and a closed, convex cone $\mathcal{K} \subseteq \mathcal{V}$ have been fixed. In addition, let $\mathcal{W}$ be a finite-dimensional real inner product space, let $\phi : \mathcal{V} \to \mathcal{W}$ be a linear map, and let $a \in \mathcal{V}$ and $b \in \mathcal{W}$ be vectors. These choices of objects define a *conic program*, with which the following optimization problem is associated.

Problem 1.1 (Standard Conic Program)

Primal problem		Dual problem	
maximize:	$\langle a, x \rangle$	*minimize:*	$\langle b, y \rangle$
subject to:	$\phi(x) = b$	*subject to:*	$\phi^*(y) - a \in \mathcal{K}^*$
	$x \in \mathcal{K}$		$y \in \mathcal{W}$

In the dual problem statement, $\phi^* : \mathcal{W} \to \mathcal{V}$ denotes the uniquely determined linear map, known as the *adjoint map* to ϕ, that satisfies $\langle w, \phi(v) \rangle = \langle \phi^*(w), v \rangle$ for all $v \in \mathcal{V}$ and $w \in \mathcal{W}$.

1.3 Feasible solutions, optimal values, and weak duality

It is convenient to associate two sets of vectors with Optimization Problem 1.1:

$$\mathcal{A} = \big\{x \in \mathcal{K} : \Phi(x) = b\big\} \quad \text{and} \quad \mathcal{B} = \big\{y \in \mathcal{W} : \Phi^*(y) - a \in \mathcal{K}^*\big\} \tag{1.7}$$

are the sets of *primal feasible* and *dual feasible* vectors for that conic program. We also define the *primal optimal* and *dual optimal* values of this conic program as

$$\alpha = \sup_{x \in \mathcal{A}} \langle a, x \rangle \quad \text{and} \quad \beta = \inf_{y \in \mathcal{B}} \langle b, y \rangle, \tag{1.8}$$

respectively. These values may be finite or infinite, and by convention we define $\alpha = -\infty$ in case $\mathcal{A} = \varnothing$ and $\beta = \infty$ in case $\mathcal{B} = \varnothing$.

Proposition 1.2 (Weak duality for conic programs)**.** *Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional real inner product spaces, let $\mathcal{K} \subseteq \mathcal{V}$ be a closed, convex cone, let $\Phi : \mathcal{V} \to \mathcal{W}$ be a linear map, and let $a \in \mathcal{V}$ and $b \in \mathcal{W}$ be vectors. For $\alpha, \beta \in \mathbb{R} \cup \{-\infty, \infty\}$ as defined in (1.8) above, it is the case that $\alpha \leq \beta$.*

Proof. If either of the sets $\mathcal{A}$ and $\mathcal{B}$ defined in (1.7) are empty, then the proposition is vacuously true: either $-\infty \leq \beta$ or $\alpha \leq \infty$. It therefore suffices to consider the case in which $\mathcal{A}$ and $\mathcal{B}$ are nonempty.

Suppose $x \in \mathcal{A}$ and $y \in \mathcal{B}$ are chosen arbitrarily. The set $\mathcal{A}$ is a subset of $\mathcal{K}$, so $x \in \mathcal{K}$, and because $y \in \mathcal{B}$ it is the case that $\Phi^*(y) - a \in \mathcal{K}^*$, and therefore

$$\langle \Phi^*(y) - a, x \rangle \geq 0. \tag{1.9}$$

We may therefore observe the following inequality and chain of equalities:

$$\langle a, x \rangle \leq \langle \Phi^*(y), x \rangle = \langle y, \Phi(x) \rangle = \langle y, b \rangle = \langle b, y \rangle. \tag{1.10}$$

This inequality is maintained as one takes the supremum over all $x \in \mathcal{A}$ and infimum over $y \in \mathcal{B}$, and therefore $\alpha \leq \beta$, as required. □

1.4 Minimization and maximization

In some situations, it may be convenient or natural to take the primal problem to be a minimization rather than a maximization problem. The dual problem then becomes a maximization problem, as follows.

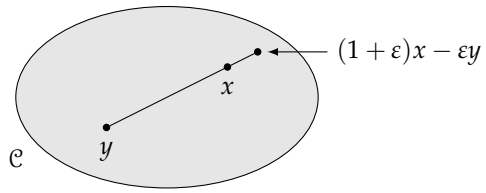

Figure 1.1: An illustration of the point $(1+\varepsilon)x - \varepsilon y$, as it relates to points x and y in a convex set $\mathcal{C}$.

Problem 1.2

Primal problem		Dual problem	
minimize:	$\langle a, x \rangle$	*maximize:*	$\langle b, y \rangle$
subject to:	$\phi(x) = b$	*subject to:*	$a - \phi^*(y) \in \mathcal{K}^*$
	$x \in \mathcal{K}$		$y \in \mathcal{W}$

Notice, in particular, that the dual constraint $a - \phi^*(y) \in \mathcal{K}^*$ has replaced the constraint $\phi^*(y) - a \in \mathcal{K}^*$ in Optimization Problem 1.1. The reason for this substitution is that Optimization Problem 1.2 is equivalent, up to a negation of sign, to an instance of Optimization Problem 1.1 in which a, b, and ϕ have been replaced with $-a$, $-b$, and $-\phi$, respectively.

With this equivalence in mind, we shall feel free to minimize or maximize in the primal problem of any conic program as we see fit—naturally selecting the corresponding dual problem formulation—but at the same time we shall lose no generality by adopting Optimization Problem 1.1 as the standard form for conic programs.

1.5 Slater's theorem

Now let us state and prove Slater's theorem. To do this, we must refer to the concept of the *relative interior* of a set $\mathcal{S} \subseteq \mathcal{V}$. This is the set denoted $\operatorname{relint}(\mathcal{S})$ that is obtained by taking the interior of the set $\mathcal{S}$, assuming that we have restricted our attention to the smallest affine subspace of $\mathcal{V}$ that contains $\mathcal{S}$.

The relative interior of a convex set $\mathcal{C}$ may be described in the following simple way:

$$\operatorname{relint}(\mathcal{C}) = \big\{ x \in \mathcal{C} \,:\, (\forall y \in \mathcal{C})(\exists \varepsilon > 0)\big((1+\varepsilon)x - \varepsilon y \in \mathcal{C}\big) \big\}. \tag{1.11}$$

Figure 1.1 illustrates how the point $(1+\varepsilon)x - \varepsilon y$ relates to x and y.

Theorem 1.3 (Slater's theorem for conic programs). *Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional real inner-product spaces, let $\mathcal{K} \subseteq \mathcal{V}$ be a closed, convex cone, let $\phi : \mathcal{V} \to \mathcal{W}$ be a linear map, and let $a \in \mathcal{V}$ and $b \in \mathcal{W}$ be vectors. With respect to the notations $\mathcal{A}$, $\mathcal{B}$, α, and β defined in Subsection 1.3, the following two statements are true.*

1. *If $\mathcal{B}$ is nonempty and there exists $x \in \operatorname{relint}(\mathcal{K})$ such that $\phi(x) = b$, then there must exist $y \in \mathcal{B}$ such that $\langle b, y \rangle = \alpha$.*
2. *If $\mathcal{A}$ is nonempty and there exists $y \in \mathcal{W}$ for which $\phi^*(y) - a \in \operatorname{relint}(\mathcal{K}^*)$, then there must exist $x \in \mathcal{A}$ such that $\langle a, x \rangle = \beta$.*

Both statements imply the equality $\alpha = \beta$.

Proof. We will prove just the first statement—the second statement can be proved through a similar technique, or one may conclude that the second statement is true given the first by formulating the dual problem of Optimization Problem 1.1 as the primal problem of a (different but equivalent) conic program. We will also make the simplifying assumptions that $\mathcal{V} = \operatorname{span}(\mathcal{K})$ and $\mathcal{W} = \operatorname{im}(\phi)$, both of which cause no loss of generality. Note that α and β are both necessarily finite, as the assumptions of the first statement imply that $\mathcal{A}$ and $\mathcal{B}$ are nonempty.

Define two subsets of $\mathcal{W} \oplus \mathcal{V} \oplus \mathbb{R}$ as follows:

$$\begin{aligned} \mathcal{C} &= \big\{(b - \phi(x), z, \langle a, x \rangle) \,:\, x, z \in \mathcal{V},\ x - z \in \mathcal{K}\big\}, \\ \mathcal{D} &= \big\{(0, 0, \eta) \,:\, \eta > \alpha\big\}. \end{aligned} \tag{1.12}$$

Both of these sets are evidently convex, and they are disjoint by the definition of α, so they are separated by a hyperplane. That is, there must exist a nonzero vector $(y, u, \lambda) \in \mathcal{W} \oplus \mathcal{V} \oplus \mathbb{R}$ such that

$$\big\langle (y, u, \lambda), (b - \phi(x), z, \langle a, x \rangle) \big\rangle \leq \big\langle (y, u, \lambda), (0, 0, \eta) \big\rangle, \tag{1.13}$$

or equivalently

$$\langle y, b - \phi(x) \rangle + \langle u, z \rangle + \lambda \langle a, x \rangle \leq \lambda \eta, \tag{1.14}$$

for all $x, z \in \mathcal{V}$ for which $x - z \in \mathcal{K}$ and all $\eta > \alpha$. Let us observe that there is no loss of generality in assuming $\lambda \in \{-1, 0, 1\}$, as the inequality (1.14) remains true when the vector (y, u, λ) is rescaled (i.e., multiplied by any positive real number).

We will now draw several conclusions from the fact that (1.14) holds for all $x, z \in \mathcal{V}$ for which $x - z \in \mathcal{K}$ and all $\eta > \alpha$.

1. The inequality (1.14) must be true when $x = 0$ and $z = 0$, and therefore

   $$\langle y, b \rangle \leq \lambda \eta \tag{1.15}$$

   for all $\eta > \alpha$. This implies that $\lambda \geq 0$, for otherwise the right-hand side of the inequality tends to $-\infty$ as η becomes large while the left-hand side remains fixed. Thus, $\lambda = -1$ is impossible, so $\lambda \in \{0, 1\}$.

2. For any choice of $x \in \mathcal{A}$ and $z \in -\mathcal{K}$, it is the case that $x - z \in \mathcal{K}$. Substituting these vectors into the inequality (1.14) and rearranging yields

$$\langle u, z \rangle \leq \lambda(\eta - \langle a, x \rangle) \tag{1.16}$$

for every $\eta > \alpha$. We conclude that $u \in \mathcal{K}^*$, for otherwise the left-hand side of the above inequality can be made to approach ∞ while the right-hand side remains bounded, for any fixed $\eta > \alpha$, through an appropriate selection of $z \in -\mathcal{K}$.

3. Assume toward contradiction that $\lambda = 0$. Fix any choice of $x \in \mathcal{A} \cap \operatorname{relint}(\mathcal{K})$, which is possible by the assumption of the statement being proved. The inequality (1.14) simplifies to

$$\langle u, z \rangle \leq 0 \tag{1.17}$$

for every choice of $z \in \mathcal{V}$ for which $x - z \in \mathcal{K}$.

Setting $z = x$, we have that $x - z \in \mathcal{K}$, and therefore

$$\langle u, x \rangle \leq 0. \tag{1.18}$$

As $x \in \mathcal{K}$ and $u \in \mathcal{K}^*$, we conclude that $\langle u, x \rangle = 0$.

On the other hand, for an arbitrarily chosen vector $v \in \mathcal{K}$, there must exist $\varepsilon > 0$ such that

$$x - \varepsilon(v - x) = (1 + \varepsilon)x - \varepsilon v \in \mathcal{K}, \tag{1.19}$$

by virtue of the fact that x is in the relative interior of $\mathcal{K}$. Setting $z = \varepsilon v - \varepsilon x$, one therefore has that $x - z \in \mathcal{K}$, and so

$$\varepsilon\langle u, v \rangle = \varepsilon\langle u, v \rangle - \varepsilon\langle u, x \rangle = \langle u, z \rangle \leq 0. \tag{1.20}$$

As it was for x, we find that $\langle u, v \rangle = 0$. Seeing that this is true for all $v \in \mathcal{K}$, and recognizing that $u \in \mathcal{V} = \operatorname{span}(\mathcal{K})$, we conclude that $u = 0$.

But if $\lambda = 0$ and $u = 0$, then we may free the vector x to range over all of $\mathcal{V}$ and set $z = x$ to conclude from (1.14) that

$$\langle y, b - \phi(x) \rangle \leq 0 \tag{1.21}$$

for every $x \in \mathcal{V}$. Bearing in mind the assumption that $\mathcal{W} = \operatorname{im}(\phi)$, we conclude that $y = 0$.

This, however, is a contradiction to the assumption that (y, u, λ) is nonzero. One concludes that $\lambda = 1$.

The steps just described have allowed us to conclude that there exist vectors $y \in \mathcal{W}$ and $u \in \mathcal{K}^*$ such that

$$\langle y, b - \phi(x) \rangle + \langle u, z \rangle \leq \eta - \langle a, x \rangle \tag{1.22}$$

for all $x, z \in \mathcal{V}$ for which $x - z \in \mathcal{K}$ and all $\eta > \alpha$. Setting $z = 0$ and flipping sign, we find that

$$\langle \phi^*(y) - a, x \rangle \geq \langle y, b \rangle - \eta \tag{1.23}$$

for all $x \in \mathcal{K}$ and $\eta > \alpha$. This implies

$$\phi^*(y) - a \in \mathcal{K}^*, \tag{1.24}$$

for otherwise the left-hand side of the previous inequality can be made to approach $-\infty$ while the right-hand side remains fixed. The vector y is therefore a dual-feasible point: $y \in \mathcal{B}$. Finally, again considering the possibility that $x = 0$ and $z = 0$, we conclude that $\langle b, y \rangle \leq \eta$ for all $\eta > \alpha$. It is therefore the case that $\langle b, y \rangle \leq \alpha$, and hence $\langle b, y \rangle = \alpha$ by weak duality. This concludes the proof (of the first statement). □

1.6 Example: conic program for optimal measurements

In this section we will discuss an example of a conic program that is relevant to quantum information. We'll begin with a somewhat general example, or perhaps a category of examples, and then discuss a specific, concrete example.

Suppose $\mathcal{X}$ is a complex Euclidean space and $\mathcal{K} \subseteq \mathrm{Herm}(\mathcal{X})$ is a closed, convex cone. Suppose further that $H_1, \ldots, H_n \in \mathrm{Herm}(\mathcal{X})$, and consider this optimization problem.

$$\begin{aligned} \text{maximize:} \quad & \langle H_1, X_1 \rangle + \cdots + \langle H_n, X_n \rangle \\ \text{subject to:} \quad & X_1 + \cdots + X_n = \mathbb{1}_{\mathcal{X}} \\ & X_1, \ldots, X_n \in \mathcal{K} \end{aligned}$$

This is essentially an *optimal measurement* problem, where $X_1, \ldots, X_n$ represent measurement operators; these operators must sum to the identity as usual, but in place of the usual constraint on measurement operators being positive semidefinite, we are instead constraining them to the cone $\mathcal{K}$. By choosing $\mathcal{K}$ to be the cone of positive semidefinite operators $\mathrm{Pos}(\mathcal{X})$, which is closed and convex, we naturally obtain the ordinary optimal measurement problem, but we can consider any closed, convex cone $\mathcal{K}$ we choose. For example, we may take $\mathcal{K}$ to be the cone of *separable operators* when $\mathcal{X} = \mathcal{Y} \otimes \mathcal{Z}$ is a bipartite tensor product space.

We can set this problem up as a conic program as follows. First, observe that $\mathcal{K}^n$ is a closed, convex cone. The problem may therefore be expressed as the primal problem of a conic program:

$$\begin{array}{ll} \text{maximize:} & \big\langle (H_1, \ldots, H_n), (X_1, \ldots, X_n) \big\rangle \\ \text{subject to:} & \phi(X_1, \ldots, X_n) \stackrel{\diamond}{=} X_1 + \cdots + X_n = \mathbb{1}_{\mathcal{X}} \\ & (X_1, \ldots, X_n) \in \mathcal{K}^n \end{array}$$

The symbol $\stackrel{\diamond}{=}$ indicates that this is the definition of the function ϕ, whereas the ordinary equal sign represents a constraint.

Here is the dual problem, as is dictated by Optimization Problem 1.1:

$$\begin{array}{ll} \text{minimize:} & \big\langle \mathbb{1}_{\mathcal{X}}, Y \big\rangle \\ \text{subject to:} & \phi^*(Y) - (H_1, \ldots, H_n) \in \big(\mathcal{K}^n\big)^* \\ & Y \in \mathrm{Herm}(\mathcal{X}) \end{array}$$

We can simplify this by observing that $\phi^*(Y) = (Y, \ldots, Y)$ and $(\mathcal{K}^n)^* = (\mathcal{K}^*)^n$, and naturally replacing the inner-product with the identity operator as the trace in the objective function. The following problem is obtained.

$$\begin{array}{ll} \text{minimize:} & \mathrm{Tr}(Y) \\ \text{subject to:} & Y - H_1 \in \mathcal{K}^* \\ & \qquad \vdots \\ & Y - H_n \in \mathcal{K}^* \\ & Y \in \mathrm{Herm}(\mathcal{X}) \end{array}$$

In summary, the following conic program, expressed in a simplified form, has been obtained.

Problem 1.3

Primal problem		Dual problem	
maximize:	$\langle H_1, X_1 \rangle + \cdots + \langle H_n, X_n \rangle$	*minimize:*	$\mathrm{Tr}(Y)$
subject to:	$X_1 + \cdots + X_n = \mathbb{1}_{\mathcal{X}}$	*subject to:*	$Y - H_1 \in \mathcal{K}^*$
	$X_1, \ldots, X_n \in \mathcal{K}$		$\vdots$
			$Y - H_n \in \mathcal{K}^*$
			$Y \in \mathrm{Herm}(\mathcal{X})$

Now let us consider a specific instance of this conic program. Define four pure states $|\psi_1\rangle, \ldots, |\psi_4\rangle \in \mathbb{C}^4 \otimes \mathbb{C}^4$ as follows:

$$|\psi_1\rangle = \frac{1}{2}\big(|0\rangle|0\rangle + |1\rangle|1\rangle + |2\rangle|2\rangle + |3\rangle|3\rangle\big) = \frac{1}{2}\operatorname{vec}(\mathbb{1} \otimes \mathbb{1}),$$

$$|\psi_2\rangle = \frac{1}{2}\big(|0\rangle|3\rangle + |1\rangle|2\rangle + |2\rangle|1\rangle + |3\rangle|0\rangle\big) = \frac{1}{2}\operatorname{vec}(\sigma_x \otimes \sigma_x),$$

$$|\psi_3\rangle = \frac{1}{2}\big(|0\rangle|3\rangle + |1\rangle|2\rangle - |2\rangle|1\rangle - |3\rangle|0\rangle\big) = \frac{i}{2}\operatorname{vec}(\sigma_y \otimes \sigma_x),$$

$$|\psi_4\rangle = \frac{1}{2}\big(|0\rangle|1\rangle + |1\rangle|0\rangle - |2\rangle|3\rangle - |3\rangle|2\rangle\big) = \frac{1}{2}\operatorname{vec}(\sigma_z \otimes \sigma_x).$$

These states were identified by Yu, Duan, and Ying (2012), who proved that they cannot be perfectly discriminated by a PPT measurement. Cosentino (2013) subsequently showed that the optimal PPT discrimination probability is $7/8$, by solving the associated semidefinite program.

We will prove that no *separable measurement* can discriminate these states with probability greater than $3/4$, assuming that one of the four states is selected uniformly at random. This is easily achievable by coarse-graining a measurement with respect to the standard basis, so this is in fact the optimal probability to correctly discriminate the states by a separable measurement. Here is a precise description of the corresponding conic program.

Problem 1.4

Primal problem

$$\begin{aligned} \textit{maximize:} \quad & \frac{1}{4}\langle\psi_1|X_1|\psi_1\rangle + \cdots + \frac{1}{4}\langle\psi_4|X_4|\psi_4\rangle \\ \textit{subject to:} \quad & X_1 + \cdots + X_4 = \mathbb{1}_4 \otimes \mathbb{1}_4 \\ & X_1, \ldots, X_4 \in \operatorname{Sep}(\mathbb{C}^4 : \mathbb{C}^4) \end{aligned}$$

Dual problem

$$\begin{aligned} \textit{minimize:} \quad & \operatorname{Tr}(Y) \\ \textit{subject to:} \quad & Y - \frac{1}{4}|\psi_k\rangle\langle\psi_k| \in \operatorname{Sep}(\mathbb{C}^4 : \mathbb{C}^4)^* \quad (1 \leq k \leq 4) \\ & Y \in \operatorname{Herm}(\mathbb{C}^4 \otimes \mathbb{C}^4) \end{aligned}$$

Our goal will be to describe a dual-feasible solution having objective value $3/4$, for this will then be an upper-bound on the probability of a correct discrimination by weak duality.

Note that the dual-feasibility of a given $Y \in \mathrm{Herm}(\mathbb{C}^4 \otimes \mathbb{C}^4)$ is equivalent to

$$Y - \frac{1}{16} \operatorname{vec}(U_k) \operatorname{vec}(U_k)^* \in \mathrm{Sep}(\mathbb{C}^4 : \mathbb{C}^4)^* \tag{1.25}$$

for

$$U_1 = \mathbb{1} \otimes \mathbb{1}, \quad U_2 = \sigma_x \otimes \sigma_x, \quad U_3 = i\sigma_y \otimes \sigma_x, \quad U_4 = \sigma_z \otimes \sigma_x. \tag{1.26}$$

In order to prove the dual-feasibility of a specific choice for Y that will be specified shortly, we will make use of the following lemma.

Lemma 1.4 (Breuer–Hall)**.** *Let $U, V \in \mathrm{U}(\mathbb{C}^n)$ be unitary operators such that $V^{\mathsf{T}}U$ is anti-symmetric: $(V^{\mathsf{T}}U)^{\mathsf{T}} = -V^{\mathsf{T}}U$. The operator*

$$Z = \mathbb{1}_n \otimes \mathbb{1}_n - \operatorname{vec}(U) \operatorname{vec}(U)^* - (\mathrm{T} \otimes \mathbb{1})(\operatorname{vec}(V) \operatorname{vec}(V)^*) \tag{1.27}$$

is contained in $\mathrm{Sep}(\mathbb{C}^n : \mathbb{C}^n)^*$.

Proof. For any unit vector $z \in \mathbb{C}^n$, we find that

$$(\mathbb{1} \otimes z)^* Z (\mathbb{1} \otimes z) = \mathbb{1} - U\bar{z}z^{\mathsf{T}}U^* - \overline{V}zz^*V^{\mathsf{T}} \geq 0. \tag{1.28}$$

This follows from the observation that the vectors $U\bar{z}$ and $\overline{V}z$ must be orthogonal unit vectors:

$$\begin{aligned} \left\langle \overline{V}z, U\bar{z} \right\rangle &= z^*V^{\mathsf{T}}U\bar{z} = \left\langle zz^{\mathsf{T}}, V^{\mathsf{T}}U \right\rangle \\ &= \left\langle (zz^{\mathsf{T}})^{\mathsf{T}}, (V^{\mathsf{T}}U)^{\mathsf{T}} \right\rangle = -\left\langle zz^{\mathsf{T}}, V^{\mathsf{T}}U \right\rangle = 0. \end{aligned} \tag{1.29}$$

It follows that

$$\langle yy^* \otimes zz^*, Z \rangle = (y \otimes z)^* Z (y \otimes z) \geq 0 \tag{1.30}$$

for every $y \in \mathbb{C}^n$. The required containment follows by convexity. □

Now define $V = \sigma_y \otimes \sigma_z$, and observe that $V^{\mathsf{T}}U_1$, $V^{\mathsf{T}}U_2$, $V^{\mathsf{T}}U_3$, and $V^{\mathsf{T}}U_4$ are all anti-symmetric. By the Breuer–Hall lemma, the operator

$$Y = \frac{1}{16}\big(\mathbb{1}_4 \otimes \mathbb{1}_4 - (\mathrm{T} \otimes \mathbb{1})(\operatorname{vec}(V) \operatorname{vec}(V)^*)\big) \tag{1.31}$$

is dual-feasible. Given that $\mathrm{Tr}(Y) = (16 - 4)/16 = 3/4$, we have obtained the claimed upper-bound on correctly discriminating these states by a separable measurement.