ENGG 5501: Foundations of Optimization **2024–25 First Term**

Handout 6: Some Applications of Conic Linear Programming

Instructor: Anthony Man–Cho So November 11, 2024

1 Introduction

Conic linear programming (CLP), and in particular, semidefinite programming (SDP), has received a lot of attention in recent years. Such a popularity can partly be attributed to the wide applicability of CLP, as well as recent advances in the design of provably efficient interior–point algorithms. In this lecture we will consider some applications that arise in engineering and show how CLP techniques can be used to model them.

2 Robust Linear Programming

To motivate the problem considered in this section, consider the LP

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x \\ \text{subject to} & \hat{A}x \le \hat{b}, \end{array} \tag{1}$$

where $\hat{A} \in \mathbb{R}^{m \times n}$, $\hat{b} \in \mathbb{R}^m$, and $\bar{c} \in \mathbb{R}^n$ are given. As it often occurs in practice, the data of the above LP, namely $\hat{A}$ and $\hat{b}$, are uncertain. In an attempt to tackle such uncertainties, Ben–Tal and Nemirovski [4] introduced the idea of **robust optimization**. In the robust optimization setting, we assume that the uncertain data lie in some given uncertainty set $\mathcal{U}$. Our goal is then to find a solution that is feasible *for all* possible realizations of the uncertain data and optimizes some given objective.

To derive the robust counterpart of (1), let us first rewrite it as

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & Az \le \mathbf{0}, \\ & z_{n+1} = -1, \end{array} \tag{2}$$

where $A = [\hat{A}\ \hat{b}] \in \mathbb{R}^{m \times (n+1)}$, $c = (\bar{c}, 0) \in \mathbb{R}^{n+1}$, and $z \in \mathbb{R}^{n+1}$. Now, suppose that each row $a_i \in \mathbb{R}^{n+1}$ of the matrix A lies in an ellipsoidal region $\mathcal{U}_i$ whose center $u_i \in \mathbb{R}^{n+1}$ is given (here, $i = 1, \ldots, m$). Specifically, we impose the constraint that

$$a_i \in \mathcal{U}_i = \left\{ x \in \mathbb{R}^{n+1} : x = u_i + B_i v,\ \|v\|_2 \le 1 \right\} \qquad \text{for } i = 1, \ldots, m,$$

where a_i is the i–th row of A, $u_i \in \mathbb{R}^{n+1}$ is the center of the ellipsoid $\mathcal{U}_i$, and B_i is some $(n+1)\times(n+1)$ positive semidefinite matrix. Then, the robust counterpart of (2) is the following optimization problem:

$$\begin{array}{lll} \text{minimize} & c^T z & \\ \text{subject to} & Az \le \mathbf{0} & \text{for all } a_i \in \mathcal{U}_i,\ \ i = 1, \ldots, m, \\ & z_{n+1} = -1. & \end{array} \tag{3}$$

Note that in general there are uncountably many constraints in (3). However, we claim that (3) is equivalent to an SOCP problem. To see this, observe that $a_i^T z \le 0$ for all $a_i \in \mathcal{U}_i$ iff

$$0 \ge \max_{v \in \mathbb{R}^{n+1}: \|v\|_2 \le 1} \left\{ (u_i + B_i v)^T z \right\} = u_i^T z + \|B_i z\|_2,$$

where $i = 1, \ldots, m$. Hence, we conclude that (3) is equivalent to

$$\begin{array}{ll} \text{minimize} & c^T z \\ \text{subject to} & \|B_i z\|_2 \le -u_i^T z \quad \text{for } i = 1, \ldots, m, \\ & z_{n+1} = -1, \end{array}$$

which is an SOCP problem.

For further readings on robust optimization, we refer the readers to [14, 1, 5, 3].

3 Chance Constrained Linear Programming

In the previous section, we showed how one can handle data uncertainties in a linear optimization problem using the robust optimization approach. Such an approach is particularly suitable if there is little information about the nature of the data uncertainty, and/or if the constraints are not to be violated under any circumstance. However, it can produce very conservative solutions. In situations where the distribution of the data uncertainty is (partially) known, one can often obtain a better solution by allowing the constraints to be violated with small probability. To illustrate this possibility, let us consider the following simple LP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a^T x \le b, \\ & x \in P, \end{array} \tag{4}$$

where $a, c \in \mathbb{R}^n$ are given vectors, $b \in \mathbb{R}$ is a given scalar, and $P \subseteq \mathbb{R}^n$ is a given polyhedron. For the sake of simplicity, suppose that P is deterministic, but the data a, b are randomly affinely perturbed; i.e.,

$$a = a^0 + \sum_{i=1}^{l} \epsilon_i a^i, \quad b = b_0 + \sum_{i=1}^{l} \epsilon_i b_i,$$

where $a^0, a^1, \ldots, a^l \in \mathbb{R}^n$ and $b_0, b_1, \ldots, b_l \in \mathbb{R}$ are given, and $\epsilon_1, \ldots, \epsilon_l$ are i.i.d. mean zero random variables supported on $[-1, 1]$. Then, for any given tolerance parameter $\delta \in (0, 1)$, we can formulate the following chance constrained counterpart of (4):

$$\begin{array}{lll} \text{minimize} & c^T x & \\ \text{subject to} & \Pr(a^T x > b) \le \delta, & (\dagger) \\ & x \in P. & \end{array} \tag{5}$$

In other words, a solution $\bar{x} \in P$ is feasible for (5) if it only violates the constraint $a^T x \le b$ with probability at most δ. The constraint $(\dagger)$ is known as a **chance constraint**. Note that when $\delta = 0$, (5) reduces to a robust linear optimization problem. Moreover, if $\bar{x} \in \mathbb{R}^n$ is feasible for (5) at some tolerance level $\bar{\delta} \ge 0$, then it is also feasible for (5) at any $\delta \ge \bar{\delta}$. Thus, by varying δ, we will be able to adjust the conservatism of the solution.

Despite the attractiveness of the chance constraint formulation, it has one major drawback; namely, it is generally computationally intractable. Indeed, even when the distributions of $\epsilon_1, \ldots, \epsilon_l$ are very simple, the feasible set defined by the chance constraint (†) can be non–convex. One way to tackle this problem is to replace the chance constraint by its **safe tractable approximation**; i.e., a system of deterministic constraints $\mathcal{H}$ such that (i) $\bar{x} \in \mathbb{R}^n$ is feasible for (†) whenever it is feasible for $\mathcal{H}$ (safe approximation), and (ii) the constraints in $\mathcal{H}$ are efficiently computable (tractability). To develop a safe tractable approximation of (†), we first observe that (†) is equivalent to the following system of constraints:

$$\Pr\left(y_0 + \sum_{i=1}^{l} \epsilon_i y_i > 0 \right) \leq \delta, \tag{6}$$

$$y_i = (a^i)^T x - b_i \qquad \text{for } i = 0, 1, \ldots, l. \tag{7}$$

Since $\epsilon_1, \ldots, \epsilon_l$ are i.i.d. mean zero random variables supported on $[-1, 1]$, by Hoeffding's inequality [12], we have

$$\Pr\left(\sum_{i=1}^{l} \epsilon_i y_i > t \right) \leq \exp\left(-\frac{t^2}{2\sum_{i=1}^{l} y_i^2} \right)$$

for any $t > 0$. It follows that when

$$-y_0 \geq \sqrt{\left(2 \ln \frac{1}{\delta} \right) \sum_{i=1}^{l} y_i^2}, \tag{8}$$

we have

$$\Pr\left(y_0 + \sum_{i=1}^{l} \epsilon_i y_i > 0 \right) \leq \exp\left(-\frac{y_0^2}{2\sum_{i=1}^{l} y_i^2} \right) \leq \delta.$$

In other words, (8) is a sufficient condition for (6) to hold. The upshot of (8) is that it is a second–order cone constraint. Hence, we conclude that constraints (7) and (8) together serve as a safe tractable approximation of the chance constraint (†).

So far we have only discussed how to handle a single scalar chance constraint. Of course, the case of joint scalar chance constraints; i.e., chance constraints of the form

$$\Pr\left((a^i)^T x > b_i \text{ for } i = 1, 2, \ldots, m \right) \leq \delta$$

are also of great interest. However, they are not as easy to handle as a single scale chance constraint. For some recent results in this direction, see [3, 6, 20, 8].

It is instructive to examine the relationship between the robust optimization approach and the safe tractable approximation approach. Upon following the derivations in the previous section, we see that constraint (8) is equivalent to the following robust constraint:

$$d^T y \leq 0 \quad \text{for all } d \in \mathcal{U},$$

where $\mathcal{U}$ is the ellipsoidal uncertainty set given by

$$\mathcal{U} = \left\{ x \in \mathbb{R}^{l+1} : x = e_1 + Bv, \ \|v\|_2 \leq 1 \right\},$$

and $B \in \mathbb{R}^{(l+1)\times(l+1)}$ is given by

$$B = \sqrt{2\ln\frac{1}{\delta}} \cdot \left[\begin{array}{c|c} 0 & \mathbf{0}^T \\ \hline \mathbf{0} & I \end{array}\right].$$

In other words, we are using the following robust optimization problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \displaystyle\sum_{i=0}^{l} d_i \left((a^i)^T x - b_i\right) \le 0 \quad \text{for all } d \in \mathcal{U}, \\ & x \in P \end{array}$$

as a safe tractable approximation of the chance constrained optimization problem (5). For further elaboration on the connection between robust optimization and chance constrained optimization, see [7, 6].

We remark that chance constrained optimization has been employed in many recent applications; see, e.g., [9, 19, 13, 23, 22]. For a more comprehensive treatment and some recent advances of chance constrained optimization, we refer the reader to [18, Chapter 4] and [21].

4 Quadratically Constrained Quadratic Optimization

A class of optimization problems that has frequently arisen in applications is that of quadratically constrained quadratic optimization problems (QCQPs); i.e., problems of the form

$$\begin{array}{ll} \text{minimize} & x^T C x \\ \text{subject to} & x^T A_i x \ge b_i \quad \text{for } i = 1, \dots, m, \end{array} \tag{9}$$

where $C, A_1, \dots, A_m \in \mathcal{S}^n$ are given. In general, due to the non–convexity of the objective function and constraints, problem (9) is intractable. Nevertheless, it can be tackled by the so–called **semidefinite relaxation** technique. To introduce this technique, we first observe that for any $C \in \mathcal{S}^n$,

$$x^T C x = \mathrm{tr}(x^T C x) = \mathrm{tr}(C x x^T) = C \bullet x x^T.$$

Hence, problem (9) is equivalent to

$$\begin{array}{ll} \text{minimize} & C \bullet x x^T \\ \text{subject to} & A_i \bullet x x^T \ge b_i \quad \text{for } i = 1, \dots, m. \end{array}$$

Now, using the spectral theorem for symmetric matrices, one can verify that

$$X = x x^T \quad \Longleftrightarrow \quad X \succeq \mathbf{0},\ \mathrm{rank}(X) \le 1.$$

It follows that problem (9) is equivalent to the following rank–constrained SDP problem:

$$\begin{array}{lll} \text{minimize} & C \bullet X & \\ \text{subject to} & A_i \bullet X \ge b_i & \text{for } i = 1, \dots, m, \\ & X \succeq \mathbf{0},\ \mathrm{rank}(X) \le 1. & \end{array} \tag{10}$$

The advantage of the formulation in (10) over that in (9) is that it reveals where the difficulty of the problem lies; namely, in the non–convex constraint $\mathrm{rank}(X) \leq 1$. By dropping this constraint, we obtain the following semidefinite relaxation of problem (9):

$$\begin{array}{ll} \text{minimize} & C \bullet X \\ \text{subject to} & A_i \bullet X \geq b_i \quad \text{for } i = 1, \ldots, m, \\ & X \succeq \mathbf{0}. \end{array} \tag{11}$$

Problem (11) is an SDP and can be efficiently solved. However, an optimal solution X^* to problem (11) may not be feasible for problem (9), since we need not have $\mathrm{rank}(X^*) \leq 1$. This motivates us to consider two fundamental questions:

1. Under what conditions would the relaxation (11) be *tight*? In particular, when would it be possible to convert an optimal solution X^* to (11) into an optimal solution x^* to (9)?

2. In the case where the relaxation (11) is not tight, how can we extract a feasible solution to (9) from an optimal solution to (11)? More ambitiously, can we establish any theoretical guarantees on the approximation quality of the extracted solution?

In the sequel, we shall present some possible approaches to tackling the above questions.

4.1 Rank Bound via Constraint Counting

Consider the following standard form SDP:

$$\begin{array}{rcll} v^*_{\mathsf{SDP}} & = & \inf & C \bullet X \\ & & \text{subject to} & A_i \bullet X = b_i \quad \text{for } i = 1, \ldots, m, \\ & & & X \succeq \mathbf{0}, \end{array} \tag{SDP}$$

where $C, A_1, \ldots, A_m \in \mathcal{S}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$ are given. The following theorem, which is discovered independently by Shapiro [17], Barvinok [2], and Pataki [16], shows that if (SDP) has an optimal solution, then it has an optimal solution whose rank is bounded by a function of m, the number of constraints.

Theorem 1 *Suppose that* (SDP) *has an optimal solution. Then, there exists an optimal solution X^* to* (SDP) *satisfying $r(r+1)/2 \leq m$, where $r = \mathrm{rank}(X^*)$. Moreover, such an optimal solution can be computed efficiently.*

Proof Let $\bar{X}^*$ be an optimal solution to (SDP) and $\bar{r} = \mathrm{rank}\left(\bar{X}^*\right)$. Suppose that $\bar{r}(\bar{r}+1)/2 > m$. Let $\bar{X}^* = LL^T$ be the Cholesky factorization of $\bar{X}^*$, where $L \in \mathbb{R}^{n \times \bar{r}}$. Set $\bar{C} = L^T C L \in \mathcal{S}^{\bar{r}}$ and $\bar{A}_i = L^T A_i L \in \mathcal{S}^{\bar{r}}$ for $i = 1, \ldots, m$, and consider the following auxiliary SDP:

$$\begin{array}{rcll} v^*_{\mathsf{ASDP}} & = & \inf & \bar{C} \bullet W \\ & & \text{subject to} & \bar{A}_i \bullet W = b_i \quad \text{for } i = 1, \ldots, m, \\ & & & W \succeq \mathbf{0}. \end{array} \tag{ASDP}$$

We claim that $v^*_{\mathsf{SDP}} = v^*_{\mathsf{ASDP}}$, and that $W = I \in \mathcal{S}^{\bar{r}}$ is an optimal solution to (ASDP). Indeed, observe that the solution $W = I$ is feasible for (ASDP) and has an objective value

$$\bar{C} \bullet I = C \bullet LL^T = C \bullet X^* = v^*_{\mathsf{SDP}}.$$

This implies that $v^*_{\mathsf{ASDP}} \le v^*_{\mathsf{SDP}}$. On the other hand, every feasible solution W to (ASDP) corresponds to a feasible solution $Z(W) = LWL^T$ to (SDP), which implies that $v^*_{\mathsf{ASDP}} \ge v^*_{\mathsf{SDP}}$. Thus, the claim is established.

Next, we show that every feasible solution to (ASDP) is in fact optimal for (ASDP). Towards that end, consider the dual of (ASDP):

$$\begin{array}{ll} \sup & b^T y \\ \text{subject to} & \bar{C} - \displaystyle\sum_{i=1}^m y_i \bar{A}_i \succeq \mathbf{0}, \\ & y \in \mathbb{R}^m. \end{array} \tag{ASDD}$$

Since (ASDP) is bounded below and strictly feasible, by the CLP Strong Duality Theorem, (ASDD) has an optimal solution y^*. Moreover, since $W = I$ is optimal for (ASDP), the CLP Strong Duality Theorem yields

$$I \bullet \left(\bar{C} - \sum_{i=1}^m y_i^* \bar{A}_i \right) = 0.$$

This, together with the fact that $\bar{C} - \sum_{i=1}^m y_i^* \bar{A}_i \succeq \mathbf{0}$, implies

$$\bar{C} - \sum_{i=1}^m y_i^* \bar{A}_i = \mathbf{0}.$$

It follows that every feasible solution $\bar{W}$ to (ASDP) satisfies the complementarity condition

$$\bar{W} \bullet \left(\bar{C} - \sum_{i=1}^m y_i^* \bar{A}_i \right) = 0.$$

Hence, by the CLP Strong Duality Theorem, we conclude that $\bar{W}$ is optimal for (ASDP).

To complete the proof of Theorem 1, consider the following system of homogeneous linear equations:

$$\bar{A}_i \bullet W = 0 \quad \text{for } i = 1, \ldots, m, \quad W \in \mathcal{S}^{\bar{r}}. \tag{12}$$

Since $W \in \mathcal{S}^{\bar{r}}$, it is completely determined by the entries on and above the diagonal. Thus, we see that (12) is a system of m equations in $\bar{r}(\bar{r}+1)/2$ variables. Now, since we assume that $\bar{r}(\bar{r}+1)/2 > m$, there exists a non–zero $\bar{W} \in \mathcal{S}^{\bar{r}}$ satisfying (12). We may assume without loss that $\bar{W}$ has at least one negative eigenvalue, for otherwise we can simply consider $-\bar{W}$, which also satisfies (12). Consider the matrix $\bar{W}^+ = I - (1/\lambda_{\min}(\bar{W}))\bar{W} \in \mathcal{S}^{\bar{r}}$, where $\lambda_{\min}(\bar{W}) < 0$ is the smallest eigenvalue of $\bar{W}$. Note that $\bar{W}^+ \succeq \mathbf{0}$, since for any $u \in \mathbb{R}^{\bar{r}}$ with $\|u\|_2 = 1$, we have

$$u^T \bar{W}^+ u = 1 - \frac{1}{\lambda_{\min}(\bar{W})} u^T \bar{W} u \ge 1 - \frac{1}{\lambda_{\min}(\bar{W})} \left(\min_{\|u\|_2 = 1} u^T \bar{W} u \right) = 0$$

by the Courant–Fischer theorem. In addition, it can be verified that $\operatorname{rank}\left(\bar{W}^+\right) < \bar{r}$. Lastly, a direct calculation yields

$$\bar{A}_i \bullet \bar{W}^+ = \bar{A}_i \bullet I = b_i \quad \text{for } i = 1, \ldots, m.$$

It follows that $\bar{W}^+$ is feasible and hence optimal for (ASDP). This implies that $X\left(\bar{W}^+\right) = L\bar{W}^+L^T$ is optimal for (SDP) and satisfies $\operatorname{rank}\left(X\left(\bar{W}^+\right)\right) < \bar{r}$.

By repeating the above procedure until $W = \mathbf{0}$ is the only solution to (12), we obtain an optimal solution X^* with $r(r+1)/2 \le m$, where $r = \mathrm{rank}(X^*)$. Moreover, note that an optimal solution to (SDP) and a non–zero solution to (12), if exists, can be found efficiently. This completes the proof of Theorem 1. $\square$

The following is an easy corollary of Theorem 1:

Corollary 1 *Consider the following SDP:*

$$\begin{array}{ll} \inf & C \bullet X \\ \text{subject to} & A_i \bullet X = b_i \quad \text{for } i = 1, \ldots, m', \\ & A_i \bullet X \ge b_i \quad \text{for } i = m'+1, \ldots, m, \\ & X \succeq \mathbf{0}, \end{array} \tag{13}$$

where $C, A_1, \ldots, A_m \in \mathcal{S}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$ are given. If problem (13) has an optimal solution, then it has an optimal solution X^* satisfying $r(r+1)/2 \le m$, where $r = \mathrm{rank}(X^*)$. Moreover, such an optimal solution can be found efficiently.

Proof Let $\bar{X}^*$ be an optimal solution to (13). Then, $\bar{X}^*$ is an optimal solution to the following SDP:

$$\begin{array}{ll} \inf & C \bullet X \\ \text{subject to} & A_i \bullet X = A_i \bullet \bar{X}^* \quad \text{for } i = 1, \ldots, m, \\ & X \succeq \mathbf{0}. \end{array} \tag{14}$$

By Theorem 1, an optimal solution X^* to (14) satisfying $r(r+1)/2$, where $r = \mathrm{rank}(X^*)$, can be found efficiently. To complete the proof, it suffices to note that X^* is also optimal for (13). $\square$

Using Corollary 1, we obtain the following tightness result concerning the relaxation (11).

Corollary 2 *The semidefinite relaxation* (11) *is tight for* (9) *when* $m \le 2$.

4.2 An Approximation Algorithm for Maximum Cut in Graphs

Suppose that we are given a simple undirected graph $G = (V, E)$ and a function $w : E \to \mathbb{R}_+$ that assigns to each edge $e \in E$ a non–negative weight w_e. The Maximum Cut Problem (Max–Cut) is that of finding a set $S \subseteq V$ of vertices such that the total weight of the edges in the cut $(S, V \backslash S)$; i.e., sum of the weights of the edges with one endpoint in S and the other in $V \backslash S$, is maximized. By setting $w_{ij} = 0$ if $(i,j) \notin E$, we may denote the weight of a cut $(S, V \backslash S)$ by

$$w(S, V \backslash S) = \sum_{i \in S, j \in V \backslash S} w_{ij}, \tag{15}$$

and our goal is to choose a set $S \subseteq V$ such that the quantity in (15) is maximized. The Max–Cut problem is one of the fundamental computational problems on graphs and has been extensively studied by many researchers. It has been shown that the Max–Cut problem is unlikely to have a polynomial–time algorithm (see, e.g, [10]). On the other hand, in a seminal work, Goemans and Williamson [11] showed how SDP can be used to design a 0.878–approximation algorithm for the Max–Cut problem; i.e., given an instance (G, w) of the Max–Cut problem, the algorithm will find a cut $(S, V \backslash S)$ whose value $w(S, V \backslash S)$ is at least 0.878 times the optimal value. In this section, we will describe the algorithm of Goemans and Williamson and prove its approximation guarantee.

To begin, let (G, w) be a given instance of the MAX–CUT problem, with $n = |V|$. Then, we can formulate the MAX–CUT problem as an integer quadratic program, viz.

$$\begin{array}{rcll} v^* & = & \text{maximize} & \displaystyle\frac{1}{2}\sum_{(i,j)\in E} w_{ij}(1 - x_i x_j) \\ & & \text{subject to} & x_i^2 = 1 \qquad\qquad\qquad \text{for } i = 1, \ldots, n. \end{array} \tag{16}$$

Here, the variable x_i indicates which side of the cut vertex i belongs to. Specifically, the cut $(S, V\backslash S)$ is given by $S = \{i \in \{1, \ldots, n\} : x_i = 1\}$. Note that if vertices i and j belong to the same side of a cut, then $x_i = x_j$, and hence its contribution to the objective function in (16) is zero. Otherwise, we have $x_i \neq x_j$, and its contribution to the objective function is $w_{ij}(1-(-1))/2 = w_{ij}$.

In general, problem (16) is hard to solve. Thus, we consider relaxations of (16). One approach is to observe that both the objective function and the constraints in (16) are *linear* in $x_i x_j$, where $1 \le i, j \le n$. In particular, if we let $X = xx^T \in \mathbb{R}^{n\times n}$, then problem (16) can be written as

$$\begin{array}{rcll} v^* & = & \text{maximize} & \displaystyle\frac{1}{2}\sum_{(i,j)\in E} w_{ij}(1 - X_{ij}) \\ & & \text{subject to} & \operatorname{diag}(X) = e, \\ & & & X = xx^T. \end{array} \tag{17}$$

Since $X = xx^T$ iff $X \succeq \mathbf{0}$ and $\operatorname{rank}(X) \le 1$, from our earlier discussion, we can drop the non–convex rank constraint and arrive at the following relaxation of (17):

$$\begin{array}{rcll} v^*_{sdp} & = & \text{maximize} & \displaystyle\frac{1}{2}\sum_{(i,j)\in E} w_{ij}(1 - X_{ij}) \\ & & \text{subject to} & \operatorname{diag}(X) = e, \\ & & & X \succeq \mathbf{0}. \end{array} \tag{18}$$

Note that (18) is an SDP. Moreover, since (18) is a relaxation of (17), we have $v^*_{sdp} \ge v^*$.

Now, suppose that we have an optimal solution X^* to (18). In general, the matrix X^* need not be in the form xx^T, and hence it does not immediately yield a feasible solution to (17). However, we can extract from X^* a solution $x' \in \{-1, 1\}^n$ to (17) via the following randomized rounding procedure:

1. Compute the *Cholesky factorization* $X^* = U^T U$ of X^*, where $U \in \mathbb{R}^{n\times n}$. Let $u_i \in \mathbb{R}^n$ be the i–th column of U. Note that $\|u_i\|_2^2 = 1$ for $i = 1, \ldots, n$.
2. Let $r \in \mathbb{R}^n$ be a vector uniformly distributed on the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$.
3. Set $x'_i = \operatorname{sgn}\left(u_i^T r\right)$ for $i = 1, \ldots, n$, where

$$\operatorname{sgn}(z) = \begin{cases} 1 & \text{if } z \ge 0, \\ -1 & \text{otherwise.} \end{cases}$$

In other words, we choose a random hyperplane through the origin (with r as its normal) and partition the vertices according to whether their corresponding vectors lie "above" or "below" the hyperplane.

Since the solution $x' \in \{-1,1\}^n$ is produced via a randomized procedure, we are interested in its expected objective value; i.e.,

$$\frac{1}{2}\mathbb{E}\left[\sum_{(i,j)\in E} w_{ij}\left(1-x_i'x_j'\right)\right] = \frac{1}{2}\sum_{(i,j)\in E} w_{ij}\mathbb{E}\left[1-x_i'x_j'\right] = \sum_{(i,j)\in E} w_{ij}\Pr\left[\operatorname{sgn}\left(u_i^T r\right) \neq \operatorname{sgn}\left(u_j^T r\right)\right]. \tag{19}$$

The following theorem provides a lower bound on the expected objective value (19) and allows us to compare it with the optimal value v^*_{sdp} of (17).

Theorem 2 *Let $u, v \in S^{n-1}$, and let r be a vector uniformly distributed on S^{n-1}. Then, we have*

$$\Pr\left[sgn\left(u^T r\right) \neq sgn\left(v^T r\right)\right] = \frac{1}{\pi}\arccos\left(u^T v\right). \tag{20}$$

Furthermore, for any $z \in [-1,1]$, we have

$$\frac{1}{\pi}\arccos(z) \geq \alpha \cdot \frac{1}{2}(1-z) > 0.878 \cdot \frac{1}{2}(1-z), \tag{21}$$

where

$$\alpha = \min_{0\leq\theta\leq\pi} \frac{2\theta}{\pi(1-\cos\theta)}.$$

Proof To establish (20), observe that by symmetry, we have

$$\Pr\left[\operatorname{sgn}\left(u^T r\right) \neq \operatorname{sgn}\left(v^T r\right)\right] = 2\Pr\left(u^T r \geq 0,\, v^T r < 0\right).$$

Now, by projecting r onto the plane containing u and v, we see that $u^T r \geq 0$ and $v^T r < 0$ iff the projection lies in the wedge formed by u and v. Since r is chosen from a spherically symmetric distribution, its projection will be a random direction on the plane containing u and v. Hence, we have

$$\Pr\left(u^T r \geq 0,\, v^T r < 0\right) = \frac{\arccos\left(u^T v\right)}{2\pi},$$

as desired.

To establish the first inequality in (21), consider the change of variable $z = \cos\theta$. Since $z \in [-1,1]$, we have $\theta \in [0,\pi]$. Thus, it follows that

$$\frac{1}{\pi}\arccos(z) = \frac{\theta}{\pi} = \frac{2\theta}{\pi(1-\cos\theta)} \cdot \frac{1}{2}(1-\cos\theta) \geq \alpha \cdot \frac{1}{2}(1-z),$$

as desired. The second inequality in (21) can be established using calculus and we leave the proof to the readers. $\square$

Corollary 3 *Given an instance (G,w) of the* Max–Cut *problem and an optimal solution to (18), the randomized rounding procedure above will produce a cut $(S', V\backslash S')$ whose expected objective value satisfies $w(S', V\backslash S') \geq 0.878v^*$.*

Proof Let x' be the solution obtained from the randomized rounding procedure, and let S' be the corresponding cut. By (19) and Theorem 2, we have

$$\begin{aligned}\mathbb{E}\left[w(S',V\backslash S')\right] &= \frac{1}{\pi}\sum_{(i,j)\in E} w_{ij}\cdot\arccos\left(u_i^Tu_j\right)\\ &\ge 0.878\cdot\frac{1}{2}\sum_{(i,j)\in E} w_{ij}\left(1-u_i^Tu_j\right)\\ &= 0.878v_{sdp}^*\\ &\ge 0.878v^*,\end{aligned}$$

as desired. □

For a survey of the theory and the many applications of the semidefinite relaxation technique, we refer the reader to [15].

References

[1] F. Alizadeh and D. Goldfarb. Second–Order Cone Programming. *Mathematical Programming, Series B*, 95:3–51, 2003.

[2] A. I. Barvinok. Problems of Distance Geometry and Convex Properties of Quadratic Maps. *Discrete and Computational Geometry*, 13:189–202, 1995.

[3] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization.* Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2009.

[4] A. Ben-Tal and A. Nemirovski. Robust Convex Optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[5] A. Ben-Tal and A. Nemirovski. Selected Topics in Robust Convex Optimization. *Mathematical Programming, Series B*, 112(1):125–158, 2008.

[6] W. Chen, M. Sim, J. Sun, and C.-P. Teo. From CVaR to Uncertainty Set: Implications in Joint Chance–Constrained Optimization. *Operations Research*, 58(2):470–485, 2010.

[7] X. Chen, M. Sim, and P. Sun. A Robust Optimization Perspective on Stochastic Programming. *Operations Research*, 55(6):1058–1071, 2007.

[8] S.-S. Cheung, A. M.-C. So, and K. Wang. Linear Matrix Inequalities with Stochastically Dependent Perturbations and Applications to Chance–Constrained Semidefinite Optimization. *SIAM Journal on Optimization*, 22(4):1394–1430, 2012.

[9] L. El Ghaoui, M. Oks, and F. Oustry. Worst–Case Value–at–Risk and Robust Portfolio Optimization: A Conic Programming Approach. *Operations Research*, 51(4):543–556, 2003.

[10] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP–Completeness.* W. H. Freeman and Company, New York, 1979.

[11] M. X. Goemans and D. P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[12] W. Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[13] W. W.-L. Li, Y. J. Zhang, A. M.-C. So, and M. Z. Win. Slow Adaptive OFDMA Systems through Chance Constrained Programming. *IEEE Transactions on Signal Processing*, 58(7):3858–3869, 2010.

[14] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of Second–Order Cone Programming. *Linear Algebra and Its Applications*, 284:193–228, 1998.

[15] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang. Semidefinite Relaxation of Quadratic Optimization Problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.

[16] G. Pataki. On the Rank of Extreme Matrices in Semidefinite Programs and the Multiplicity of Optimal Eigenvalues. *Mathematics of Operations Research*, 23(2):339–358, 1998.

[17] A. Shapiro. Rank–Reducibility of a Symmetric Matrix and Sampling Theory of Minimum Trace Factor Analysis. *Psychometrika*, 47(2):187–199, 1982.

[18] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*, volume 9 of *MPS–SIAM Series on Optimization.* Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 2009.

[19] M. B. Shenouda and T. N. Davidson. Outage–Based Designs for Multi–User Transceivers. In *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009)*, pages 2389–2392, 2009.

[20] A. M.-C. So. Moment Inequalities for Sums of Random Matrices and Their Applications in Optimization. *Mathematical Programming, Series A*, 130(1):125–151, 2011.

[21] A. M.-C. So. Lecture Notes on Chance Constrained Optimization. Available at `http://www.se.cuhk.edu.hk/~manchoso/papers/ccopt.pdf`, 2013.

[22] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi. Outage Constrained Robust Transmit Optimization for Multiuser MISO Downlinks: Tractable Approximations by Conic Optimization. *IEEE Transactions on Signal Processing*, 62(21):5690–5705, 2014.

[23] Y. J. Zhang and A. M.-C. So. Optimal Spectrum Sharing in MIMO Cognitive Radio Networks via Semidefinite Programming. *IEEE Journal on Selected Areas in Communications*, 29(2):362–373, 2011.