X2TE2109: Convex Optimization and Its Applications in Signal Processing

Handout 6: Quadratically Constrained Quadratic Optimization with Complex Variables and Rank Reduction in Semidefinite Programming

Instructor: Anthony Man–Cho So April 14, 2022

1 Introduction

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

$$\begin{array}{ll} \text{minimize} & z^H Q z \\ \text{subject to} & z^H A_i z \geq b_i \quad \text{for } i = 1, \ldots, m, \end{array} \tag{QCQP}$$

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

$$z^H Q z = \text{tr}(z^H Q z) = Q \bullet zz^H.$$

Hence, (QCQP) is equivalent to

$$\begin{array}{ll} \text{minimize} & Q \bullet zz^H \\ \text{subject to} & A_i \bullet zz^H \geq b_i \quad \text{for } i = 1, \ldots, m. \end{array}$$

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

$$Z = zz^H \quad \Longleftrightarrow \quad Z \succeq \mathbf{0}, \ \text{rank}(Z) \leq 1.$$

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

$$\begin{array}{lll} \text{minimize} & Q \bullet Z & \\ \text{subject to} & A_i \bullet Z \geq b_i & \text{for } i = 1, \ldots, m, \\ & Z \succeq \mathbf{0}, \ \text{rank}(Z) \leq 1. & \end{array} \tag{RCSDP}$$

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

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

Problem (QPSDR) is a (complex) SDP, which can be efficiently solved. However, an optimal solution Z^* to (QPSDR) may not be feasible for (QCQP), since we need not have $\text{rank}(Z^*) \leq 1$. This motivates us to consider two fundamental questions:

1. Under what conditions would the relaxation (QPSDR) be *tight*? In particular, when would it be possible to convert an optimal solution Z^* to (QPSDR) into an optimal solution z^* to (QCQP)?
2. In the case where the relaxation (QPSDR) is not tight, how can we extract a feasible solution to (QCQP) from an optimal solution to (QPSDR)? More ambitiously, can we establish any theoretical guarantees on the approximation quality of the extracted solution?

In the sequel, we shall develop some machinery to address the above questions. As we shall see, the CLP duality theory plays an important role in the development of those machinery.

2 Rank of SDP Solutions

2.1 Bounds via Constraint Counting

Consider the following standard form complex SDP:

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

where $C, A_1, \ldots, A_m \in \mathcal{H}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$ are given. The following theorem 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 Z^* to (SDP) satisfying* $\operatorname{rank}(Z^*) \le \lfloor \sqrt{m} \rfloor$. *Moreover, such an optimal solution can be computed efficiently.*

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

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

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

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

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 $Z(W) = LWL^H$ 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):

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

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 that

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

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:

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

Since $W \in \mathcal{H}^r$, it is completely determined by the entries on and above the diagonal. Note that the diagonal entries of W must be real, while the entries above the diagonal can be complex. It follows that (1) is a system of m equations in $r + 2(r(r-1)/2) = r^2$ real variables. Now, if $r^2 > m$, then there exists a non–zero $\bar{W} \in \mathcal{H}^r$ satisfying (1). 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 (1). Consider the matrix $\bar{W}^+ = I + (1/\lambda_{\min}(\bar{W}))\bar{W} \in \mathcal{H}^r$, where $\lambda_{\min}(\bar{W}) < 0$ is the smallest eigenvalue of $\bar{W}$. It can be verified that $\bar{W}^+ \succeq \mathbf{0}$ and $\text{rank}\left(\bar{W}^+\right) < r$. Moreover, 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 $Z\left(\bar{W}^+\right) = L\bar{W}^+L^H$ is optimal for (SDP) and satisfies $\text{rank}\left(Z\left(\bar{W}^+\right)\right) < r$.

By repeating the above procedure until $W = \mathbf{0}$ is the only solution to (1), we obtain an optimal solution Z^* with $\text{rank}(Z^*)^2 \le m$. Since $\text{rank}(Z^*)$ is an integer, we must have $\text{rank}(Z^*) \le \lfloor \sqrt{m} \rfloor$. Lastly, note that an optimal solution to (SDP) and a non–zero solution to (1), if exists, can be found efficiently. This completes the proof of Theorem 1. □

The following is an easy corollary of Theorem 1:

Corollary 1 *Consider the following SDP:*

$$\begin{array}{lll} \inf & C \bullet Z & \\ \text{subject to} & A_i \bullet Z = b_i & \text{for } i = 1, \ldots, m', \\ & A_i \bullet Z \ge b_i & \text{for } i = m'+1, \ldots, m, \\ & Z \succeq \mathbf{0}, & \end{array} \tag{2}$$

where $C, A_1, \ldots, A_m \in \mathcal{H}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$ are given. If problem (2) has an optimal solution, then it has an optimal solution Z^* satisfying $\text{rank}(Z^*) \le \lfloor \sqrt{m} \rfloor$*. Moreover, such an optimal solution can be found efficiently.*

Proof Let $\bar{Z}^*$ be an optimal solution to (2). Then, $\bar{Z}^*$ is an optimal solution to the following SDP:

$$\begin{array}{rll} \inf & C \bullet Z & \\ \text{subject to} & A_i \bullet Z = A_i \bullet \bar{Z}^* & \text{for } i = 1, \ldots, m, \\ & Z \succeq \mathbf{0}. & \end{array} \tag{3}$$

By Theorem 1, an optimal solution Z^* to (3) satisfying $\text{rank}(Z^*) \leq \lfloor \sqrt{m} \rfloor$ can be found efficiently. To complete the proof, it suffices to note that Z^* is also optimal for (2). □

Using Corollary 1, we obtain our first tightness result concerning the relaxation (QPSDR).

Corollary 2 *The semidefinite relaxation (QPSDR) is tight for (QCQP) when $m \leq 3$.*

REMARK: Theorem 1 extends the corresponding result for real SDPs in [14, 1, 12] to complex SDPs. A slightly different formulation of Theorem 1 can be found in [8, Theorem 5.1].

With minimal additional effort, one can extend Theorem 1 to cover SDPs with certain block structure. Specifically, consider the following problem:

$$\begin{array}{rrll} v^*_{\mathsf{BSDP}} = & \inf & \displaystyle\sum_{k=1}^{K} C_k \bullet Z_k & \\ & \text{subject to} & \displaystyle\sum_{k=1}^{K} A_{ik} \bullet Z_k = b_i & \text{for } i = 1, \ldots, m, \\ & & Z_k \succeq \mathbf{0} & \text{for } k = 1, \ldots, K, \end{array} \tag{BSDP}$$

where $C_k, A_{1k}, \ldots, A_{mk} \in \mathcal{H}^{n_k}$ for $k = 1, \ldots, K$ and $b_1, \ldots, b_m \in \mathbb{R}$ are given. We then have the following theorem:

Theorem 2 *Suppose that (BSDP) has an optimal solution. Then, there exists an optimal solution $(Z_1^*, \ldots, Z_K^*)$ to (BSDP) satisfying $\sum_{k=1}^{K} \text{rank}\,(Z_k^*)^2 \leq m$. Moreover, such an optimal solution can be computed efficiently.*

Proof The proof essentially follows that of Theorem 1. Let $(\bar{Z}_1^*, \ldots, \bar{Z}_K^*)$ be an optimal solution to (BSDP) with $\text{rank}\left(\bar{Z}_k^*\right) = r_k$ for $k = 1, \ldots, K$. Consider the Cholesky factorization $\bar{Z}_k^* = L_k L_k^H$ of $\bar{Z}_k^*$, where $L_k \in \mathbb{C}^{n_k \times r_k}$ and $k = 1, \ldots, K$. Then, using similar argument as in the proof of Theorem 1, it can be shown that every feasible solution $(W_1, \ldots, W_K)$ to the auxiliary SDP

$$\begin{array}{rrll} v^*_{\mathsf{ABSDP}} = & \inf & \displaystyle\sum_{k=1}^{K} \bar{C}_k \bullet W_k & \\ & \text{subject to} & \displaystyle\sum_{k=1}^{K} \bar{A}_{ik} \bullet W_k = b_i & \text{for } i = 1, \ldots, m, \\ & & W_k \succeq \mathbf{0} & \text{for } k = 1, \ldots, K, \end{array} \tag{ABSDP}$$

where $\bar{C}_k = L_k^H C_k L_k \in \mathcal{H}^{r_k}$ and $\bar{A}_{ik} = L_k^H A_{ik} L_k \in \mathcal{H}^{r_k}$ for $k = 1, \ldots, K$, $i = 1, \ldots, m$, is optimal for (ABSDP) and corresponds to an optimal solution $(L_1 W_1 L_1^H, \ldots, L_K W_K L_K^H)$ to (BSDP).

Now, consider the following system of homogeneous linear equations:

$$\sum_{k=1}^{K} \bar{A}_{ik} \bullet W_k = 0 \quad \text{for } i = 1, \ldots, m, \quad W_k \in \mathcal{H}^{r_k} \quad \text{for } k = 1, \ldots, K. \tag{4}$$

Note that the number of real variables in (4) is precisely $r = \sum_{k=1}^{K} r_k^2$. Thus, if $r > m$, then there exist matrices $\bar{W}_k \in \mathcal{H}^{r_k}$ for $k = 1, \ldots, K$, not all zero, such that $(\bar{W}_1, \ldots, \bar{W}_K)$ satisfies (4). Let

$$\mathcal{K} = \left\{ k : \bar{W}_k \neq \mathbf{0} \right\}.$$

Without loss of generality, we may assume that for $k \in \mathcal{K}$, the matrix $\bar{W}_k$ has at least one negative eigenvalue. Set $\lambda = \min\left\{\lambda_{\min}(\bar{W}_k) : k \in \mathcal{K}\right\} < 0$ and

$$\bar{W}_k^+ = \begin{cases} I_{r_k} + (1/\lambda)\bar{W}_k & \text{for } k \in \mathcal{K}, \\ I_{r_k} & \text{for } k \notin \mathcal{K}, \end{cases}$$

where I_{r_k} is the $r_k \times r_k$ identity matrix and $\lambda_{\min}(\bar{W}_k)$ is the smallest eigenvalue of $\bar{W}_k$ for $k = 1, \ldots, K$. It can then be verified that $\bar{W}_k^+ \succeq \mathbf{0}$ for $k = 1, \ldots, K$ and $\operatorname{rank}\left(\bar{W}_k^+\right) < r_k$ for some $k \in \mathcal{K}$. Moreover, since $\bar{W}_k = \mathbf{0}$ for $k \notin \mathcal{K}$ and $(\bar{W}_1, \ldots, \bar{W}_K)$ satisfies (4), we have

$$\sum_{k=1}^{K} \bar{A}_{ik} \bullet \bar{W}_k^+ = \sum_{k=1}^{K} \bar{A}_{ik} \bullet I_{r_k} + \frac{1}{\lambda} \sum_{k=1}^{K} \bar{A}_{ik} \bullet \bar{W}_k = b_i \quad \text{for } i = 1, \ldots, m.$$

It follows that $\left(\bar{W}_1^+, \ldots, \bar{W}_K^+\right)$ is feasible and hence optimal for (ABSDP). By setting $Z_k^* = L_k \bar{W}_k^+ L_k^H$ for $k = 1, \ldots, K$, we conclude that $(Z_1^*, \ldots, Z_K^*)$ is optimal for (BSDP) and satisfies $\sum_{k=1}^{K} \operatorname{rank}(Z_k^*)^2 < r$. To complete the proof, it suffices to repeat the above procedure and follow the argument in the proof of Theorem 1. □

Similar to Corollary 1, we have the following immediate corollary of Theorem 2:

Corollary 3 *Consider the following SDP:*

$$\begin{array}{lll} \inf & \displaystyle\sum_{k=1}^{K} C_k \bullet Z_k & \\ \text{subject to} & \displaystyle\sum_{k=1}^{K} A_{ik} \bullet Z_k = b_i & \text{for } i = 1, \ldots, m', \\ & \displaystyle\sum_{k=1}^{K} A_{ik} \bullet Z_k \geq b_i & \text{for } i = m'+1, \ldots, m, \\ & Z_1, \ldots, Z_K \succeq \mathbf{0} & \text{for } k = 1, \ldots, K, \end{array} \tag{5}$$

where $C_k, A_{1k}, \ldots, A_{mk} \in \mathcal{H}^{n_k}$ and $b_1, \ldots, b_m \in \mathbb{R}$ are given. If problem (5) has an optimal solution, then it has an optimal solution $(Z_1^*, \ldots, Z_K^*)$ satisfying $\sum_{k=1}^{K} \operatorname{rank}\left(Z_k^*\right)^2 \leq m$. Moreover, such an optimal solution can be found efficiently.

Remark: Theorem 2 and Corollary 3 are essentially taken from [7].

2.2 Bound via Complementarity

In the previous sub–section, we saw that a bound on the rank of an SDP solution can be obtained by simply counting the number of constraints. In this sub–section, we describe an alternative approach, which exploits the complementarity property of the primal and dual optimal solutions.

To begin, recall that the dual of (SDP) is given by

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

Suppose that (SDP) and (SDD) have optimal solutions Z^* and y^*, respectively. Then, by the CLP Strong Duality Theorem, we have $Z^* \bullet S^* = 0$, where $S^* = C - \sum_{i=1}^m y_i^* A_i \succeq \mathbf{0}$. As the following result shows, this implies a bound on the ranks of Z^* and S^*:

Proposition 1 *Let $A, B \in \mathcal{H}_+^n$ be such that $A \bullet B = 0$. Then,* $\text{rank}(A) + \text{rank}(B) \leq n$.

Proof Let $A = U\Lambda U^H$ be the spectral decomposition of A. Then, we have

$$0 = A \bullet B = \Lambda \bullet U^H BU = \sum_{i=1}^n \Lambda_{ii} (U^H BU)_{ii}.$$

Since $B \succeq \mathbf{0}$, we have $U^H BU \succeq \mathbf{0}$. In particular, we have $(U^H BU)_{ii} = 0$ whenever $\Lambda_{ii} > 0$, or equivalently, the i–th row and i–th column of $U^H BU$ have all zero entries whenever $\Lambda_{ii} > 0$. This implies that $\text{rank}(U^H BU) \leq n - \text{rank}(\Lambda) = n - \text{rank}(A)$. Since B and $U^H BU$ have the same rank, we conclude that $\text{rank}(A) + \text{rank}(B) \leq n$, as desired. □

3 Connection to the $\mathcal{S}$–Procedure

The rank bounds in the previous section can be used to develop the $\mathcal{S}$–procedure, which can be viewed as a theorem of alternatives for quadratic systems. The $\mathcal{S}$–procedure plays a fundamental role in the development of duality theory for non–convex quadratic optimization (see, e.g., [2, 9, 17]) and has applications in many areas of science and engineering (see e.g., [3]). For a historical perspective of the $\mathcal{S}$–procedure, we refer the reader to [6, 13]. In this section, let us use Theorem 1 to prove the following version of the $\mathcal{S}$–procedure:

Theorem 3 *Let $A_1, A_2, Q \in \mathcal{H}^n$ be given. Suppose there exists a $z_0 \in \mathbb{C}^n$ such that $z_0^H A_1 z_0 > 0$ and $z_0^H A_2 z_0 > 0$. Then, the following statements are equivalent:*

(I) *For all $z \in \mathbb{C}^n$, $z^H Q z \geq 0$ whenever $z^H A_1 z \geq 0$ and $z^H A_2 z \geq 0$.*

(II) *There exist $\lambda_1, \lambda_2 \geq 0$ such that $Q - \lambda_1 A_1 - \lambda_2 A_2 \succeq \mathbf{0}$.*

Proof The implication (II)⇒(I) is quite straightforward, so it remains to establish the reverse implication. Towards that end, consider the SDP

$$\begin{array}{rcll} v^* & = & \inf & Q \bullet Z \\ & & \text{subject to} & A_i \bullet Z \geq 0 \quad \text{for } i = 1, 2, \\ & & & I \bullet Z = 1, \\ & & & Z \succeq \mathbf{0}, \end{array} \tag{6}$$

which is a relaxation of the following QCQP:

$$\begin{array}{ll} \inf & z^H Q z \\ \text{subject to} & z^H A_i z \geq 0 \quad \text{for } i = 1, 2, \\ & \|z\|_2^2 = 1. \end{array}$$

Since the feasible region of Problem (6) is compact, there exists an optimal solution to (6). Hence, by Theorem 1, there exists a rank–one optimal solution $Z^* = z^*(z^*)^H$ to (6). Together with (I), this implies that $v^* \geq 0$. Now, consider the dual of (6), which is given by

$$\begin{array}{ll} \sup & \mu \\ \text{subject to} & Q - \lambda_1 A_1 - \lambda_2 A_2 - \mu I \succeq \mathbf{0}, \\ & \lambda_1, \lambda_2 \geq 0. \end{array} \tag{7}$$

Since $z_0^H A_1 z_0 > 0$ and $z_0^H A_2 z_0 > 0$, it is easy to verify that Problem (6) is strictly feasible. Hence, the CLP Strong Duality Theorem implies that Problem (7) has an optimal solution $(\mu^*, \lambda_1^*, \lambda_2^*)$, and that $\mu^* = v^* \geq 0$. In particular, we have $Q - \lambda_1^* A_1 - \lambda_2^* A_2 \succeq \mu^* I \succeq \mathbf{0}$ with $\lambda_1^*, \lambda_2^* \geq 0$. This establishes (II). □

With some additional, rather elementary arguments, one can derive various extensions of Theorem 3. As an illustration, consider the following inhomogeneous version of Theorem 3:

Corollary 4 *Let $A_1, A_2, Q \in \mathcal{H}^n$, $b_1, b_2, q \in \mathbb{C}^n$, and $c_1, c_2, d \in \mathbb{R}$ be given. Define the functions $f_1, f_2, g : \mathbb{C}^n \to \mathbb{R}$ as follows:*

$$\begin{aligned} f_1(z) &= z^H A_1 z + 2\,\mathrm{Re}(b_1^H z) + c_1, \\ f_2(z) &= z^H A_2 z + 2\,\mathrm{Re}(b_2^H z) + c_2, \\ g(z) &= z^H Q z + 2\,\mathrm{Re}(q^H z) + d. \end{aligned}$$

Suppose there exists a $z_0 \in \mathbb{C}^n$ such that $f_1(z_0) > 0$ and $f_2(z_0) > 0$. Then, the following statements are equivalent:

(I) *For all $z \in \mathbb{C}^n$, $g(z) \geq 0$ whenever $f_1(z) \geq 0$ and $f_2(z) \geq 0$.*

(II) *There exist $\lambda_1, \lambda_2 \geq 0$ such that*

$$\begin{bmatrix} Q & q \\ q^H & d \end{bmatrix} - \lambda_1 \begin{bmatrix} A_1 & b_1 \\ b_1^H & c_1 \end{bmatrix} - \lambda_2 \begin{bmatrix} A_2 & b_2 \\ b_2^H & c_2 \end{bmatrix} \succeq \mathbf{0}.$$

Proof Consider the following homogenizations of f_1, f_2, g:

$$\begin{aligned} &\mathbb{C}^n \times \mathbb{C} \ni (z,t) \mapsto f_1^\circ(z,t) = z^H A_1 z + 2\,\mathrm{Re}(t z^H b_1) + c_1|t|^2 = (z,t)^H A_1^\circ (z,t), \\ &\mathbb{C}^n \times \mathbb{C} \ni (z,t) \mapsto f_2^\circ(z,t) = z^H A_2 z + 2\,\mathrm{Re}(t z^H b_2) + c_2|t|^2 = (z,t)^H A_2^\circ (z,t), \\ &\mathbb{C}^n \times \mathbb{C} \ni (z,t) \mapsto g^\circ(z,t) = z^H Q z + 2\,\mathrm{Re}(t z^H q) + d|t|^2 = (z,t)^H Q^\circ (z,t), \end{aligned}$$

where

$$A_1^\circ = \begin{bmatrix} A_1 & b_1 \\ b_1^H & c_1 \end{bmatrix}, \quad A_2^\circ = \begin{bmatrix} A_2 & b_2 \\ b_2^H & c_2 \end{bmatrix}, \quad Q^\circ = \begin{bmatrix} Q & q \\ q^H & d \end{bmatrix}.$$

By construction, we have $f_1(z) = f_1^\circ(z,1)$, $f_2(z) = f_2^\circ(z,1)$, and $g(z) = g^\circ(z,1)$. It follows that (II)⇒(I). To prove the converse, our strategy is to show that (I) is equivalent to

(I') For all $(z,t) \in \mathbb{C}^n \times \mathbb{C}$, $g^\circ(z,t) \ge 0$ whenever $f_1^\circ(z,t) \ge 0$ and $f_2^\circ(z,t) \ge 0$.

The desired result would then follow from Theorem 3. Clearly, we have the implication (I')⇒(I). Conversely, suppose that (I) holds and $(z',t') \in \mathbb{C}^n \times \mathbb{C}$ is such that $f_1^\circ(z',t') \ge 0$ and $f_2^\circ(z',t') \ge 0$. Consider the following two cases:

Case 1: $t' \neq 0$. Then, we have

$$0 \le f_1^\circ\left(z',t'\right) = |t'|^2 f_1^\circ\left(z'/t',1\right) = |t'|^2 f_1\left(z'/t'\right).$$

Similarly, $f_2\left(z'/t'\right) \ge 0$. Since (I) holds, we get $g\left(z'/t'\right) \ge 0$, which implies that $g^\circ(z',t') \ge 0$; i.e., (I') holds.

Case 2: $t' = 0$. Then, we have $f_1^\circ(z',0) = (z')^H A_1 z' \ge 0$ and $f_2^\circ(z',0) = (z')^H A_2 z' \ge 0$. Suppose to the contrary that $g^\circ(z',0) = (z')^H Q z' < 0$. Let $w(t) = z_0 + tz'$ and note that

$$\begin{aligned} f_1(w(t)) &= f_1(z_0) + |t|^2 (z')^H A_1 z' + 2\,\mathrm{Re}\left(t(A_1 z_0 + b_1)^H z'\right), \\ f_2(w(t)) &= f_2(z_0) + |t|^2 (z')^H A_2 z' + 2\,\mathrm{Re}\left(t(A_2 z_0 + b_2)^H z'\right), \\ g(w(t)) &= g(z_0) + |t|^2 (z')^H Q z' + 2\,\mathrm{Re}\left(t(Q z_0 + q)^H z'\right). \end{aligned}$$

If $\mathrm{Re}\left((A_1 z_0 + b_1)^H z'\right)$ and $\mathrm{Re}\left((A_2 z_0 + b_2)^H z'\right)$ have the same sign, then by taking t to be real with the same sign as $\mathrm{Re}\left((A_1 z_0 + b_1)^H z'\right)$ and letting $|t|$ to be sufficiently large, we have $f_1(w(t)) \ge 0$, $f_2(w(t)) \ge 0$, and $g(w(t)) < 0$, which contradicts (I). On the other hand, if $\mathrm{Re}\left((A_1 z_0 + b_1)^H z'\right)$ and $\mathrm{Re}\left((A_2 z_0 + b_2)^H z'\right)$ have different signs, then we can find a $\theta \in [0,2\pi)$ such that $\mathrm{Re}\left(e^{i\theta}(A_1 z_0 + b_1)^H z'\right)$ and $\mathrm{Re}\left(e^{i\theta}(A_2 z_0 + b_2)^H z'\right)$ have the same sign. It follows that by taking $t = re^{i\theta}$ with $r > 0$ sufficiently large, we again have $f_1(w(t)) \ge 0$, $f_2(w(t)) \ge 0$, and $g(w(t)) < 0$, which contradicts (I).

From the above analysis, we see that $g^\circ(z',0) \ge 0$, which implies that (I') holds. This completes the proof. □

As a further illustration of the power of the rank bound established in Theorem 1, let us develop an $\mathcal{S}$–procedure with both equality and inequality constraints:

Corollary 5 *Let $A, Q \in \mathcal{H}^n$, $b, q \in \mathbb{C}^n$, and $c, d \in \mathbb{R}$ be given. Define the functions $f, g : \mathbb{C}^n \to \mathbb{R}$ as follows:*

$$\begin{aligned} f(z) &= z^H A z + 2\,\mathrm{Re}(b^H z) + c, \\ g(z) &= z^H Q z + 2\,\mathrm{Re}(q^H z) + d. \end{aligned}$$

Suppose there exist $z_0, z_1 \in \mathbb{C}^n$ such that $\|z_0\|_2 \le 1$, $\|z_1\|_2 \le 1$, and $f(z_0) < 0 < f(z_1)$. Then, the following statements are equivalent:

(I) *For all $z \in \mathbb{C}^n$, $g(z) \ge 0$ whenever $z^H z \le 1$ and $f(z) = 0$.*

(II) *There exist $\lambda_1 \ge 0$ and $\lambda_2 \in \mathbb{R}$ such that*

$$\begin{bmatrix} Q & q \\ q^H & d \end{bmatrix} - \lambda_1 \begin{bmatrix} -I & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} A & b \\ b^H & c \end{bmatrix} \succeq \mathbf{0}.$$

Proof The implication (II)⇒(I) is again straightforward. To prove the converse, we follow the idea in the proof of Theorem 3 and consider the SDP

$$\begin{array}{rcl} v^* & = & \inf \quad \begin{bmatrix} Q & q \\ q^H & d \end{bmatrix} \bullet Z \\ & & \text{subject to} \quad \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -1 \end{bmatrix} \bullet Z \le 0, \\ & & \begin{bmatrix} A & b \\ b^H & c \end{bmatrix} \bullet Z = 0, \\ & & Z = \begin{bmatrix} Z_{11} & z \\ z^H & 1 \end{bmatrix} \succeq \mathbf{0}. \end{array} \tag{8}$$

Since the feasible region of Problem (8) is compact, there exists an optimal solution to (8). Hence, by Theorem 1, there exists a rank–one optimal solution $Z^* = z^*(z^*)^H$ to (8). Together with (I), this implies that $v^* \ge 0$. Now, let $Z_0 \succ \mathbf{0}$ be such that

$$Z_0 = \begin{bmatrix} Z_{01} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \succ \mathbf{0}, \quad \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -1 \end{bmatrix} \bullet Z_0 < 0.$$

(For instance, one can take $Z_{01} = (n+1)^{-1} I \in \mathcal{H}^n$.) Furthermore, let

$$\theta = \begin{bmatrix} A & b \\ b^H & c \end{bmatrix} \bullet Z_0.$$

If $\theta = 0$, then Z_0 is a strictly feasible solution to Problem (8). Otherwise, if $\theta < 0$, then by the assumption on $z_1 \in \mathbb{C}^n$, there exists an $\alpha \in (0,1)$ such that

$$Z_0^- = \alpha Z_0 + (1-\alpha) \begin{bmatrix} z_1 z_1^H & z_1 \\ z_1^H & 1 \end{bmatrix}$$

is strictly feasible for Problem (8). Similarly, if $\theta > 0$, then the assumption on $z_0 \in \mathbb{C}^n$ implies the existence of an $\alpha \in (0,1)$ such that

$$Z_0^+ = \alpha Z_0 + (1-\alpha) \begin{bmatrix} z_0 z_0^H & z_0 \\ z_0^H & 1 \end{bmatrix}$$

is strictly feasible for Problem (8). Summarizing the above discussion, we see that Problem (8) is strictly feasible. Thus, by the CLP Strong Duality Theorem, the dual of (8), which is given by

$$\begin{array}{ll} \sup & \mu \\ \text{subject to} & \begin{bmatrix} Q & q \\ q^H & d \end{bmatrix} - \lambda_1 \begin{bmatrix} -I & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} A & b \\ b^H & c \end{bmatrix} - \mu \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \succeq \mathbf{0}, \\ & \lambda_1 \ge 0, \end{array}$$

has an optimal solution $(\mu^*, \lambda_1^*, \lambda_2^*)$, and that $v^* = \mu^* \ge 0$. It then follows that (II) holds. □

REMARK: In general, the question of whether an $\mathcal{S}$–procedure exists for systems with quadratic equality constraints is a delicate one. For some recent progress in this direction, see [20].

4 Applications

4.1 Unicast Transmit Downlink Beamforming

Consider the scenario in which a base station equipped with N_t antennae is transmitting individual data streams to M single–antenna users. The signal transmitted by the base station is modeled as

$$x(t) = \sum_{i=1}^{M} s_i(t) w_i \quad \text{for } t = 1, \ldots, T,$$

where $s_i(t)$ and $w_i \in \mathbb{C}^{N_t}$ are the stream of unit–power data symbols and beamforming vector for user i, respectively. The signal received by the i–th user is given by

$$\begin{aligned} y_i(t) &= h_i^H x(t) + n_i(t) \\ &= h_i^H w_i s_i(t) + \sum_{j \neq i} h_i^H w_j s_j(t) + n_i(t) \quad \text{for } t = 1, \ldots, T, \end{aligned} \tag{9}$$

where $h_i \in \mathbb{C}^{N_t}$ is the channel vector of user i and $n_i(t) \sim \mathcal{CN}(0, \sigma_i^2)$ is the additive noise at user i with power σ_i^2. Based on (9), a design problem of interest is that of signal–to–interference–and–noise (SINR) balancing; namely, to minimize the total transmit power while guaranteeing a certain level of received SINR for each user. Assuming that the channels $\{h_i\}_{i=1}^M$ are randomly fading and only the second–order statistics $R_i = \mathbb{E}\left[h_i h_i^H\right]$, where $i = 1, \ldots, M$, are known, the SINR balancing problem can be formulated as follows:

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{i=1}^{M} \|w_i\|_2^2 & \\ \text{subject to} & \mathsf{SINR}_i = \dfrac{w_i^H R_i w_i}{\sum_{j \neq i} w_j^H R_i w_j + \sigma_i^2} \geq \gamma_i & \text{for } i = 1, \ldots, M, \\ & w_i \in \mathbb{C}^{N_t} & \text{for } i = 1, \ldots, M; \end{array} \tag{10}$$

see, e.g., [5]. It is easy to see that Problem (10) is a QCQP. Thus, using the techniques introduced in Section 1, we obtain the following SDR of Problem (10):

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{i=1}^{M} I \bullet W_i & \\ \text{subject to} & R_i \bullet W_i - \gamma_i \displaystyle\sum_{j \neq i} R_i \bullet W_j \geq \gamma_i \sigma_i^2 & \text{for } i = 1, \ldots, M, \\ & W_i \succeq \mathbf{0} & \text{for } i = 1, \ldots, M. \end{array} \tag{11}$$

It can be shown that the dual of (11) is strictly feasible. Hence, if Problem (11) is feasible, then by the CLP Strong Duality Theorem, it has an optimal solution. In this case, Theorem 2 implies that Problem (11) has an optimal solution $(W_1^*, \ldots, W_M^*)$ satisfying $\sum_{i=1}^M \text{rank}(W_i^*)^2 \leq M$. However, this is only possible when $\text{rank}(W_i^*) \leq 1$ for $i = 1, \ldots, M$. It follows that the SDR (11) is tight for the SINR balancing problem (10).

4.2 Transmit Design for MISO Channel Secrecy

Consider the scenario in which a base station equipped with N_t antennae is transmitting a data stream to a legitimate single–antenna receiver, but is being eavesdropped by M illegitimate multi–antenna receivers. We assume that the i–th illegitimate receiver has $N_{e,i}$ antennae, where $i = 1, \ldots, M$. In the literature, the base station, the legitimate receiver, and the illegitimate receiver (or eavesdropper) are commonly called Alice, Bob, and Eve, respectively. A fundamental problem in such a scenario is to design a transmit scheme for the base station so that it can reliably communicate with the legitimate receiver while preventing the eavesdroppers from obtaining information from the transmitted signals. To begin, let

$$\begin{aligned} y_b(t) &= h^H x(t) + n(t) \quad \text{for } t = 1, \ldots, T, \\ y_{e,i}(t) &= G_i^H x(t) + v_i(t) \quad \text{for } t = 1, \ldots, T;\ i = 1, \ldots, M \end{aligned}$$

be the received signal of Bob and the i–th Eve, respectively, where $x(t) \in \mathbb{C}^{N_t}$ is the signal transmitted by Alice; $h \in \mathbb{C}^{N_t}$ is the multiple–input single–output (MISO) channel between Alice and Bob; $G_i \in \mathbb{C}^{N_t \times N_{e,i}}$ is the multiple–input multiple–output (MIMO) channel between Alice and the i–th Eve; $n(t) \in \mathbb{C}$ and $v_i(t) \in \mathbb{C}^{N_{e,i}}$ are additive white Gaussian noise at Bob and the i–th Eve, respectively. Without loss of generality, we assume that the $n(t)$ and $v_i(t)$ have unit variance. Furthermore, let $W = \mathbb{E}\left[x(t)x(t)^H\right]$ be the transmit covariance. We can then formulate the aforementioned problem using the notion of physical–layer secrecy (see, e.g., [10]). Specifically, we are interested in minimizing the average transmit power while guaranteeing a certain level of minimum achievable secrecy rate, viz.

$$\begin{array}{ll} \text{minimize} & \operatorname{tr}(W) \\ \text{subject to} & \displaystyle\min_{i=1,\ldots,M} f_i(W) \geq R, \\ & W \succeq \mathbf{0}. \end{array} \tag{12}$$

Here, $f_i(W) = \log(1 + h^H W h) - \log\det(I + G_i^H W G_i)$ is the so–called secrecy rate function associated with the i–th Eve and $R > 0$ is a given minimum secrecy rate threshold.

Problem (12) is not an SDP, as it is non–convex. The main difficulty lies in the $\log\det$ function. To circumvent such difficulty, let us prove the following lemma:

Lemma 1 *Let $A \in \mathcal{H}_+^n$ be given. Then, we have*

$$\det(I + A) \geq 1 + \operatorname{tr}(A)$$

and equality holds iff $\operatorname{rank}(A) \leq 1$.

Proof Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be the eigenvalues of A. Then, we have

$$\det(I + A) = \prod_{i=1}^{n} (1 + \lambda_i) \geq 1 + \sum_{i=1}^{n} \lambda_i = 1 + \operatorname{tr}(A),$$

as desired. □

Armed with the above lemma, we can relax Problem (12) as follows:

$$\begin{array}{ll} \text{minimize} & \operatorname{tr}(W) \\ \text{subject to} & 1 + h^H W h \geq 2^R \left(1 + \operatorname{tr}(G_i^H W G_i)\right) \quad \text{for } i = 1, \ldots, M, \\ & W \succeq \mathbf{0}. \end{array}$$

Note that the above problem is an SDP, as it can be rewritten as

$$\begin{array}{ll} \text{minimize} & I \bullet W \\ \text{subject to} & \left(hh^H - 2^R G_i G_i^H \right) \bullet W \geq 2^R - 1 \quad \text{for } i = 1, \ldots, M, \\ & W \succeq \mathbf{0}. \end{array} \tag{13}$$

The dual of (13) is given by

$$\begin{array}{ll} \text{maximize} & (2^R - 1)e^T y \\ \text{subject to} & S(y) = I - \sum_{i=1}^{M} y_i \left(hh^H - 2^R G_i G_i^H \right) \succeq \mathbf{0}, \\ & y \geq \mathbf{0}. \end{array} \tag{14}$$

Observe that if $\bar{W}$ is feasible for Problem (13), then there exist $\alpha > 1$ and $\beta > 0$ such that $\bar{W}^+ = \alpha \bar{W} + \beta I$ is strictly feasible for (13). Moreover, if we let

$$\lambda = 1 + \max \left\{ \max_{i=1,\ldots,M} \lambda_{\max} \left(hh^H - 2^R G_i G_i^H \right), 0 \right\}$$

and set $\bar{y} = (1/M\lambda)e > \mathbf{0}$, then $(S(\bar{y}), \bar{y})$ is strictly feasible for Problem (14). Hence, by the CLP Strong Duality Theorem and Proposition 1, there exists an optimal solution W^* to (13) and an optimal solution $(S(y^*), y^*)$ to (14) such that $\text{rank}(W^*) + \text{rank}(S(y^*)) \leq N_t$. Now, observe that

$$B = I + 2^R \sum_{i=1}^{M} y_i^* G_i G_i^H \succ \mathbf{0}.$$

This yields

$$\begin{aligned} \text{rank}(S(y^*)) &= \text{rank} \left(B^{-1/2} S(y^*) B^{-1/2} \right) \\ &= \text{rank} \left(I - \left(\sum_{i=1}^{M} y_i^* \right) \left(B^{-1/2} h \right) \left(B^{-1/2} h \right)^H \right) \\ &\geq N_t - 1. \end{aligned}$$

Hence, we conclude that $\text{rank}(W^*) \leq 1$.

4.3 Robust Unicast Downlink Precoder Design

In this sub–section, let us revisit the scenario considered in Section 4.1. In practice, the channel vector of each user (i.e., $h_i \in \mathbb{C}^{N_t}$, where $i = 1, \ldots, M$) needs to be estimated by the base station and thus is not accurately known. In order to account for the channel estimation errors in our design process, we need to first specify a model of the channel errors. A popular choice is the so–called *norm–bounded* error model (see, e.g., [15, 18]), in which the actual channel vector of user i (where $i = 1, \ldots, M$) is given by

$$h_i = \bar{h}_i + e_i.$$

Here, $\bar{h}_i \in \mathbb{C}^{N_t}$ is the base station's estimate of user i's channel vector and $e_i \in \mathbb{C}^{N_t}$ is the channel error vector satisfying $\|e_i\|_2 \le \epsilon_i$ for some given threshold $\epsilon_i \ge 0$. Then, the robust precoder design problem can be formulated as

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{M} I \bullet W_i \\ \text{subject to} \quad & (\bar{h}_i + e_i)^H W_i (\bar{h}_i + e_i) \\ & -\gamma_i \sum_{j \neq i} (\bar{h}_i + e_i)^H W_j (\bar{h}_i + e_i)^H \ge \gamma_i \sigma_i^2 \qquad \text{for all } \|e_i\|_2 \le \epsilon_i,\ i = 1, \ldots, M, \qquad \text{(15a)} \\ & W_i \succeq \mathbf{0} \qquad \text{for } i = 1, \ldots, M. \qquad \text{(15b)} \end{aligned}$$

Note that Problem (15) is similar to Problem (11), except that it contains the semi–infinite constraints (15a). Thus, Problem (15) is not an SDP. Nevertheless, it can be converted into an SDP using the $\mathcal{S}$–procedure. Indeed, observe that for $i = 1, \ldots, M$,

$$\begin{aligned} & (\bar{h}_i + e_i)^H W_i (\bar{h}_i + e_i) - \gamma_i \sum_{j \neq i} (\bar{h}_i + e_i)^H W_j (\bar{h}_i + e_i)^H - \gamma_i \sigma_i^2 \\ = \quad & e_i^H \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) e_i + 2 \operatorname{Re} \left[\bar{h}_i^H \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) e_i \right] \\ + \quad & \bar{h}_i^H \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) \bar{h}_i - \gamma_i \sigma_i^2. \end{aligned}$$

Hence, by taking

$$\begin{aligned} & A_1 = -I, \quad b_1 = \mathbf{0}, \quad c_1 = \epsilon_i^2, \\ & A_2 = \mathbf{0}, \quad b_2 = \mathbf{0}, \quad c_2 = 1, \\ & Q = W_i - \gamma_i \sum_{j \neq i} W_j, \quad q = \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) \bar{h}_i, \quad d = \bar{h}_i^H \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) \bar{h}_i - \gamma_i \sigma_i^2 \end{aligned}$$

in Corollary 4, we see that constraint (15a) is equivalent to

$$\begin{bmatrix} W_i - \gamma_i \sum_{j \neq i} W_j & \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) \bar{h}_i \\ \bar{h}_i^H \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) & \bar{h}_i^H \left(W_i - \gamma_i \sum_{j \neq i} W_j \right) \bar{h}_i - \gamma_i \sigma_i^2 \end{bmatrix} - \lambda_i \begin{bmatrix} -I & \mathbf{0} \\ \mathbf{0} & \epsilon_i^2 \end{bmatrix} \succeq \mathbf{0}, \qquad \text{(16a)}$$

$$\lambda_i \ge 0. \qquad \text{(16b)}$$

It follows that Problem (15) can be reformulated as

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{M} I \bullet W_i \\ \text{subject to} \quad & \text{(15b)},\ \text{(16a)}, \text{and (16b)}, \end{aligned} \qquad \text{(17)}$$

which is an SDP. An interesting open question here is whether Problem (17) always admits an optimal solution $(W_1^*, \ldots, W_M^*)$ satisfying $\text{rank}(W_i^*) \le 1$ for $i = 1, \ldots, M$. Some partial results can be found in [4, 16, 19].

Recently, Medra et al. [11] have considered a frequency division duplex (FDD) system with structured vector quantization and proposed a channel error model that can more accurately reflect the nature of estimation errors in such system. Specifically, let $\bar{h}_i \in \mathbb{C}^{N_t}$ be the base station's estimate of user i's channel, where $i = 1, \ldots, M$. The base station uses a Grassmannian codebook $\mathcal{C} = \{v_1, \ldots, v_P\}$, where $v_j \in \mathbb{C}^{N_t}$ with $\|v_j\|_2 = 1$ is given for $j = 1, \ldots, P$, to determine the direction of user i's channel via

$$d_i = \arg\min_{v \in \mathcal{C}} \left\{ 1 - \frac{|\bar{h}_i^H v|^2}{\|\bar{h}_i\|_2^2} \right\}.$$

Under suitable conditions, the actual channel vector of user i can then be expressed as

$$h_i = \|\bar{h}_i\|_2 (d_i + e_i),$$

where $e_i \in \mathbb{C}^{N_t}$ is the channel error vector, whose statistics depend both on the statistics of the channel and the codebook used. As is often the case, the statistics of e_i are difficult to characterize. Thus, let us assume for simplicity that e_i lies in a region defined by the following system:

$$\|e_i\|_2 \le \epsilon, \quad \|d_i + e_i\|_2 = 1. \tag{18}$$

Typically, $\epsilon \ge 0$ is regarded as a given parameter, and a robust precoder design problem similar to Problem (15) can then be formulated. However, let us explore another possibility; namely, we treat $\epsilon \ge 0$ as a decision variable and use it to determine the largest region in which the error vectors can reside without compromising the quality–of–service to the users. Specifically, consider the following formulation (see [11]):

$$\begin{aligned} \text{maximize} \quad & \epsilon \\ \text{subject to} \quad & (d_i + e_i)^H W_i (d_i + e_i) \\ & -\gamma_i \sum_{j \ne i} (d_i + e_i)^H W_j (d_i + e_i)^H \ge \frac{\gamma_i \sigma_i^2}{\|\bar{h}_i\|_2} \quad \text{for all } e_i \text{ satisfying (18)},\ i = 1, \ldots, M, \quad \text{(19a)} \\ & \sum_{i=1}^M I \bullet W_i \le P, \quad \text{(19b)} \\ & W_i \succeq \mathbf{0} \qquad \text{for } i = 1, \ldots, M. \quad \text{(19c)} \end{aligned}$$

Again, Problem (19) contains semi–infinite constraints (see (19a)). To tackle them, we first observe that since $\|d_i\|_2 = 1$, we have

$$1 = \|d_i + e_i\|_2 \iff e_i^H e_i + 2\text{Re}(d_i^H e_i) = 0.$$

In particular, for any $\epsilon > 0$, there exist $e_i^0, e_i^1 \in \mathbb{C}^{N_t}$ such that $\|e_i^0\|_2 < \epsilon$, $\|e_i^1\|_2 < \epsilon$, and $(e_i^0)^H e_i^0 + 2\mathrm{Re}(d_i^H e_i^0) < 0 < (e_i^1)^H e_i^1 + 2\mathrm{Re}(d_i^H e_i^1)$. Now, by taking

$$
\begin{aligned}
&A = I, \quad b = d_i, \quad c = 0, \\
&Q = W_i - \gamma_i \sum_{j\neq i} W_j, \quad q = \left(W_i - \gamma_i \sum_{j\neq i} W_j \right) d_i, \quad d = d_i^H \left(W_i - \gamma_i \sum_{j\neq i} W_j \right) d_i - \frac{\gamma_i \sigma_i^2}{\|\bar{h}_i\|_2}
\end{aligned}
$$

in Corollary 5, we see that constraint (19a) is equivalent to

$$
\begin{bmatrix} W_i - \gamma_i \sum_{j\neq i} W_j & \left(W_i - \gamma_i \sum_{j\neq i} W_j \right) d_i \\ d_i^H \left(W_i - \gamma_i \sum_{j\neq i} W_j \right) & d_i^H \left(W_i - \gamma_i \sum_{j\neq i} W_j \right) d_i - \frac{\gamma_i \sigma_i^2}{\|\bar{h}_i\|_2} \end{bmatrix} - \lambda_{1,i} \begin{bmatrix} -I & \mathbf{0} \\ \mathbf{0} & \epsilon^2 \end{bmatrix} - \lambda_{2,i} \begin{bmatrix} I & d_i \\ d_i^H & 0 \end{bmatrix} \succeq \mathbf{0}, \tag{20a}
$$

$$
\lambda_{1,i} \geq 0. \tag{20b}
$$

Hence, Problem (19) can be reformulated as

$$
\begin{array}{ll} \text{maximize} & \epsilon \\ \text{subject to} & (19\text{b}),\ (19\text{c}),\ (20\text{a}), \text{and } (20\text{b}). \end{array} \tag{21}
$$

It should be noted that Problem (21) is still not an SDP, as constraint (20a) involves the nonlinear term $\lambda_{1,i}\epsilon^2$. However, for a fixed $\epsilon \geq 0$, Problem (21) is an SDP. Thus, we can approximate the optimal solution to Problem (21) to arbitrary accuracy efficiently by performing a bisection search on ϵ.

References

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

[2] A. Beck and Y. C. Eldar. Strong Duality in Nonconvex Quadratic Optimization with Two Quadratic Constraints. *SIAM Journal on Optimization*, 17(3):844–860, 2006.

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

[4] T.-H. Chang, W.-K. Ma, and C.-Y. Chi. Worst–Case Robust Multiuser Transmit Beamforming Using Semidefinite Relaxation: Duality and Implications. In *2011 Conference Record of the 45th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 1579–1583, 2011.

[5] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten. Convex Optimization–Based Beamforming: From Receive to Transmit and Network Designs. *IEEE Signal Processing Magazine*, 27(3):62–75, 2010.

[6] S. V. Gusev and A. L. Likhtarnikov. Kalman–Popov–Yakubovich Lemma and the S–Procedure: A Historical Essay. *Automation and Remote Control*, 67(11):1768–1810, 2006.

[7] Y. Huang and D. P. Palomar. Rank–Constrained Separable Semidefinite Programming with Applications to Optimal Beamforming. *IEEE Transactions on Signal Processing*, 58(2):664–678, 2010.

[8] Y. Huang and S. Zhang. Complex Matrix Decomposition and Quadratic Programming. *Mathematics of Operations Research*, 32(3):758–768, 2007.

[9] V. Jeyakumar, G. M. Lee, and G. Y. Li. Alternative Theorems for Quadratic Inequality Systems and Global Quadratic Optimization. *SIAM Journal on Optimization*, 20(2):983–1001, 2009.

[10] Q. Li and W.-K. Ma. Optimal and Robust Transmit Designs for MISO Channel Secrecy by Semidefinite Programming. *IEEE Transactions on Signal Processing*, 59(8):3799–3812, 2011.

[11] M. Medra, W.-K. Ma, and T. N. Davidson. Low–Complexity Robust MISO Downlink Precoder Optimization for the Limited Feedback Case. In *Proceedings of the 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2015)*, pages 3108–3112, 2015.

[12] 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.

[13] I. Pólik and T. Terlaky. A Survey of the S–Lemma. *SIAM Review*, 49(3):371–418, 2007.

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

[15] M. B. Shenouda and T. N. Davidson. Convex Conic Formulations of Robust Downlink Precoder Designs with Quality of Service Constraints. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):714–724, 2007.

[16] E. Song, Q. Shi, M. Sanjabi, R.-Y. Sun, and Z.-Q. Luo. Robust SINR–Constrained MISO Downlink Beamforming: When is Semidefinite Programming Relaxation Tight? *EURASIP Journal on Wireless Communications and Networking*, 2012(243), 2012.

[17] H. Tuy and H. D. Tuan. Generalized S–Lemma and Strong Duality in Nonconvex Quadratic Programming. *Journal on Global Optimization*, 56(3):1045–1072, 2013.

[18] N. Vučić and H. Boche. Robust QoS–Constrained Optimization of Downlink Multiuser MISO Systems. *IEEE Transactions on Signal Processing*, 57(2):714–725, 2009.

[19] Y. Wang and R. Shi. Tightness of Semidefinite Programming Relaxation to Robust Transmit Beamforming with SINR Constraints. *Mathematical Problems in Engineering*, 2013:Article ID 508014, 2013.

[20] Y. Xia, S. Wang, and R.-L. Sheu. S–Lemma with Equality and Its Applications. Accepted for publication in *Mathematical Programming, Series A*, 2015.