Semidefinite Relaxation of Quadratic Optimization Problems and Applications

Wing-Kin (Ken) Ma

Department of Electronic Engineering,
The Chinese University of Hong Kong, Hong Kong

Lesson 9, ELEG5481
Reference:


Acknowledgment: Anthony So, Tom Luo, Yinyu Ye, & Shuzhong Zhang.
Outline

• Part I: Basic concepts and overview of semidefinite relaxation (SDR)

• Part II: Theory, and implications in practice

• Part III: Frontier Developments
  – Outage-based Transmit Beamforming Optimization
Part I: Basic Concepts and Overview
A quick reminder of what convex quadratic functions & constraints are:

- A function $f(x) = x^T C x = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j C_{ij}$ is convex if and only if $C \succeq 0$ ($C \succeq 0$ means that $C$ is positive semidefinite (PSD)).
Quadratically Constrained Quadratic Program

Consider the class of real-valued quadratically constrained quadratic programs (QCQPs):

$$\min_{x \in \mathbb{R}^n} x^T C x$$

s.t.  
$$x^T F_i x \geq g_i, \quad i = 1, \ldots, p,$$
$$x^T H_i x = l_i, \quad i = 1, \ldots, q,$$

where $C, F_1, \ldots, F_p, H_1, \ldots, H_q \in \mathbb{S}^n$; $\mathbb{S}^n$ is the set of all $n \times n$ real symmetric matrices.

- We do not consider convex cases, and $C, F_i, H_i$ may be arbitrary.

- Nonconvex QCQP is a very difficult problem in general.
Nonconvex QCQP: How Hard Could it Be?

Consider the **Boolean quadratic program (BQP)**

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad x^T C x \\
\text{s.t.} & \quad x_i^2 = 1, \quad i = 1, \ldots, n,
\end{align*}
\]

a long-known difficult problem falling in the nonconvex QCQP class.

- You could solve it by evaluating all possible combinations; i.e., brute-force search.
- The complexity of a brute-force search is \(O(2^n)\), not okay at all for large \(n\)!
- The BQP is **NP-hard** in general— we still can’t find an algorithm that can solve a general BQP in \(O(n^p)\) for any \(p > 0\).
Nonconvex QCQP: How Hard Could it Be?

Consider the following problem

\[
\begin{align*}
\min_{\mathbf{x} \in \mathbb{R}^n} & \quad \mathbf{x}^T \mathbf{C} \mathbf{x} \\
\text{s.t.} & \quad \mathbf{x}^T \mathbf{F}_i \mathbf{x} \geq 1, \quad i = 1, \ldots, m,
\end{align*}
\]

where \( \mathbf{C}, \mathbf{F}_1, \ldots, \mathbf{F}_m \) are all positive semidefinite, or \( \mathbf{C}, \mathbf{F}_1, \ldots, \mathbf{F}_m \succeq 0 \).

- Difficulty: feasible set is the intersection of the exteriors of ellipsoids.
- This problem is also NP-hard.
Semidefinite Relaxation for QCQP

Semidefinite relaxation (SDR) is a computationally efficient approximation approach to QCQP.

- Approximate QCQPs by a semidefinite program (SDP), a class of convex optimization problems where reliable, efficient algorithms are readily available.

- The idea can be found in an early paper of Lovász in 1979 [Lovász’79].

- It is arguably the work by Goemans & Williamson [Goemans-Williamson’95] that sparked the significant interest in SDR.

- A key notion introduced by Goemans & Williamson is randomization; we will go through that.

- SDR has received much interest in the optimization field; now we have seen a number of theoretically elegant analysis results.

- (This may concern us more) In many applications, SDR works well empirically.
Impacts of SDR in SP and Commun.

• The introduction of SDR in SP and commun. since the early 2000’s has reshaped the way we see many topics today.

• Applications identified include
  – multiuser/MIMO detection [Tan-Rasmussen’01], [Ma-Davidson-Wong-Luo-Ching’02]
  – multiuser downlink tx beamforming: unicast [Bengtsson-Ottersten’01], multicast [Sidiropoulos-Davidson-Luo’06], & more recently, multicell downlinks, relaying (incl. analog network coding), cognitive radio, secrecy...
  – sensor network localization [Biswas-Liang-Wang-Ye’06]
  – robust blind receive beamforming [Ma-Ching-Vo’04]
  – code waveform design in radar [De Maio et al.’08]
  – transmit $B_1$ shim in MRI [Chang-Luo-Wu et al.’08]
  – fusion for distributed detection [Quan-Ma-Cui-Sayed’10]
  – binary image restoration, phase unwrapping
  – large-margin parameter estimation in speech recognition [Li-Jiang’07]
  – ...

and the scope of applications is still expanding.
The Concept of SDR

• For notational conciseness, we write the QCQP as

\[
\min_{x \in \mathbb{R}^n} \quad x^T C x \\
\text{s.t.} \quad x^T A_i x \geq_i b_i, \quad i = 1, \ldots, m.
\]

(QCQP)

Here, ‘\(\geq_i\)’ can represent either ‘\(\geq\)’, ‘\(=\)’, or ‘\(\leq\)’ for each \(i\); \(C, A_1, \ldots, A_m \in \mathbb{S}^n\); and \(b_1, \ldots, b_m \in \mathbb{R}\).

• A crucial first step of understanding SDR is to see that

\[
x^T C x = \text{Tr}(x^T C x) = \text{Tr}(C xx^T), \quad x^T A_i x = \text{Tr}(x^T A_i x) = \text{Tr}(A_i xx^T),
\]

or, if we let \(X = xx^T\),

\[
x^T C x = \text{Tr}(C X), \quad x^T A_i x = \text{Tr}(A_i X)
\]

• The objective and constraint functions are linear in \(X\).
The Concept of SDR

• The condition $X = xx^T$ is equivalent to $X \succeq 0$, $\text{rank}(X) = 1$, thus (QCQP) is the same as

$$\min_{X \in \mathbb{S}^n} \quad \text{Tr}(CX)$$

s.t. \quad \text{Tr}(A_iX) \succeq b_i, \quad i = 1, \ldots, m \quad (\text{QCQP})$$

$X \succeq 0, \quad \text{rank}(X) = 1.$

• The constraints $\text{Tr}(A_iX) \succeq b_i$ are easy, but $\text{rank}(X) = 1$ is hard.

• **Key Insight:** Drop the rank-one constraint to obtain a relaxed QCQP

$$\min_{X \in \mathbb{S}^n} \quad \text{Tr}(CX)$$

s.t. \quad $\text{Tr}(A_iX) \succeq b_i, \quad i = 1, \ldots, m, \quad (\text{SDR})$

$X \succeq 0.$

(SDR) is a convex problem.
Some Merits We Can Immediately Say

- The SDR

\[
\begin{align*}
\min_{X \in \mathbb{S}^n} \quad & \text{Tr}(CX) \\
\text{s.t.} \quad & X \succeq 0, \quad \text{Tr}(A_i X) \geq_i b_i, \quad i = 1, \ldots, m
\end{align*}
\]

(SDR)

is a semidefinite program (SDP), whose globally optimal solution may be found by available numerical algorithms in polynomial time (often by interior-point methods, in \(O(\max\{m, n\}^4 n^{1/2} \log(1/\epsilon))\), \(\epsilon\) being soln. accuracy).

- For instance, using the software toolbox CVX, we can solve (SDR) in MATLAB with the following lines: (for simplicity we assume ‘\(\geq_i\)’ = ‘\(\geq\)’ for all \(i\) here)

```plaintext
cvx_begin
    variable X(n,n) symmetric
    minimize(trace(C*X));
    subject to
        for i=1:m
            trace(A(:,:,i)*X) >= b(i);
        end
    X == semidefinite(n)

cvx_end
```

W.-K. Ma, Dept. EE, The Chinese University of Hong Kong
### Issues with the Use of SDR

- There is no free lunch in turning the NP-hard (QCQP) to the convex, polynomial-time solvable (SDR).

- The issue is how to convert an SDR solution to an approximate QCQP solution.

- If an SDR solution, say, denoted by $X^*$, is of rank one; or, equivalently,

  \[ X^* = x^* x^{*T}, \]

  then $x^*$ is feasible—and in fact optimal—to (QCQP).

- But the case of rank-one SDR solutions does not always hold (otherwise we would have solved an NP-hard problem in polynomial time!)

- There are many ways to produce an approximate QCQP solution from $X^*$, for instances where $\text{rank}(X^*) > 1$. 
QCQP Solution Approximation in SDR: An Example

• Consider again the BQP

\[
\begin{align*}
\min & \quad x^T C x \\
\text{s.t.} & \quad x_i^2 = 1, \quad i = 1, \ldots, n.
\end{align*}
\]  

\text{(BQP)}

The SDR of (BQP) is

\[
\begin{align*}
\min & \quad \text{Tr}(C X) \\
\text{s.t.} & \quad X \succeq 0, \quad X_{ii} = 1, \quad i = 1, \ldots, n.
\end{align*}
\]  

\text{(SDR)}

• An intuitively reasonable idea (true even for engineers) is to apply a rank-1 approximation to the SDR solution \(X^*\):

1) Carry out the eigen-decomposition

\[
X^* = \sum_{i=1}^{\text{rank}(X^*)} \lambda_i q_i q_i^T,
\]

where \(r = \text{rank}(X^*), \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0\) are the eigenvalues and \(q_1, \ldots, q_r \in \mathbb{R}^n\) the respective eigenvectors.

2) Approximate the BQP by \(\hat{x} = \text{sgn}(\sqrt{\lambda_1} q_1)\).
**Application: MIMO Detection**

**Scenario:** A spatial multiplexing system with $M_t$ transmit & $M_r$ receive antennae.

![Diagram of MIMO Detection System]

**Objective:** detect symbols from the received signals, given channel information.

- Received signal model:
  \[
  y_C = H_C s_C + \nu_C
  \]
  where $H_C \in \mathbb{C}^{M_r \times M_t}$ is the MIMO channel, $s_C \in \mathbb{C}^{M_t}$ is the transmitted symbol vector, & $\nu_C \in \mathbb{C}^{M_r}$ is complex circular Gaussian noise.

- Assume QPSK constellations, $s_C \in \{\pm 1 \pm j\}^{M_t}$. 
• Problem: maximum-likelihood (ML) detection (NP-hard)

$$\hat{s}_{C, ML} = \arg \min_{s_C \in \{\pm 1 \pm j\}^{M_t}} ||y_C - H_C s_C||^2$$

• The received signal model can be converted to a real form

$$\begin{bmatrix} \text{Re}\{y_C\} \\ \text{Im}\{y_C\} \end{bmatrix} = \begin{bmatrix} \text{Re}\{H_C\} & -\text{Im}\{H_C\} \\ \text{Im}\{H_C\} & \text{Re}\{H_C\} \end{bmatrix} \begin{bmatrix} \text{Re}\{s_C\} \\ \text{Im}\{s_C\} \end{bmatrix} + \begin{bmatrix} \text{Re}\{v_C\} \\ \text{Im}\{v_C\} \end{bmatrix},$$

and hence the ML problem can be rewritten (homogenized) as

$$\min_{s \in \{\pm 1\}^{2M_t}} ||y - H s||^2 = \min_{s \in \{\pm 1\}^{2M_t}, t \in \{\pm 1\}} ||t y - H s||^2$$

$$= \min_{s \in \{\pm 1\}^{2M_t}, t \in \{\pm 1\}} \begin{bmatrix} s^T & t \end{bmatrix} \begin{bmatrix} H^T H & -H^T y \\ -y^T H & ||y||^2 \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix},$$

which is a BQP. Subsequently, SDR can be applied [Tan-Rasmussen’01], [Ma-Davidson-Wong-Luo-Ching’02].
Bit error rate performance under \((M_r, M_t) = (40, 40)\). ‘ZF’— zero forcing; ‘MMSE-DF’— min. mean square error with decision feedback; ‘LRA’— lattice reduction aided. ‘Randomization’ will be explained shortly.
Complexity comparison of various MIMO detectors. SNR= 12dB. Sphere decoding is an exact ML method.
Additional Remarks about the MIMO Detection Application

• The idea is not restricted to spatial multiplexing! It can also be used in multiuser CDMA, space-time/freq./time-freq. coding, multiuser MIMO, and even blind MIMO [Li-Bai-Ding’03], [Ma-Vo-Davidson-Ching’06],...

• Extensions that have been considered:
  – MPSK constellations [Ma-Ching-Ding’04];
  – higher-order QAM constellations [Ma-Su-Jaldén-Chang-Chi’09] (and refs. therein);
  – soft-in-soft-out MIMO detection (a.k.a. BICM-MIMO) [Steingrimsson-Luo-Wong’03];
  – fast implementations [Kisialiou-Luo-Luo’09], [Wai-Ma-So’11]

• Performance analysis for SDR MIMO detection:
  – diversity analysis [Jaldén-Ottersten’08]
  – probabilistic approximation accuracy analysis [Kisialiou-Luo’10], [So’10].
Alternative Interpretation of SDR: Solving QCQP in Expectation

- We return to the SDR solution approximation issue. Recall

\[
\min_{\mathbf{x} \in \mathbb{R}^n} \quad \mathbf{x}^T \mathbf{C} \mathbf{x} \\
\text{s.t.} \quad \mathbf{x}^T \mathbf{A}_i \mathbf{x} \succeq b_i, \quad i = 1, \ldots, m. \tag{QCQP}
\]

- Let \( \xi \sim \mathcal{N}(0, \mathbf{X}) \) where \( \mathbf{X} \) is the covariance. Consider a stochastic QCQP:

\[
\min_{\mathbf{X} \in \mathbb{S}^n, \quad \mathbf{X} \succeq 0} \quad \mathbb{E}_{\xi \sim \mathcal{N}(0, \mathbf{X})} \{\xi^T \mathbf{C} \xi\} \\
\text{s.t.} \quad \mathbb{E}_{\xi \sim \mathcal{N}(0, \mathbf{X})} \{\xi^T \mathbf{A}_i \xi\} \succeq b_i, \quad i = 1, \ldots, m, \tag{E-QCQP}
\]

where we manipulate the statistics of \( \xi \) so that the objective function is minimized & constraints are satisfied in expectation.

- One can show that (E-QCQP) is the same as the SDR

\[
\min \quad \text{Tr}(\mathbf{C} \mathbf{X}) \\
\text{s.t.} \quad \mathbf{X} \succeq 0, \quad \text{Tr}(\mathbf{A}_i \mathbf{X}) \succeq b_i, \quad i = 1, \ldots, m. \tag{SDR}
\]
The stochastic QCQP interpretation of SDR

\[
\min_{X \in S^n} \mathbb{E}_{\xi \sim \mathcal{N}(0, X)} \{\xi^T C \xi\}
\]

s.t. \( \mathbb{E}_{\xi \sim \mathcal{N}(0, X)} \{\xi^T A_i \xi\} \geq b_i, \quad i = 1, \ldots, m \)

(E-QCQP)

essentially sheds lights into a different way of approximating QCQP.

- What we could do is the following: generate a random vector \( \xi \sim \mathcal{N}(0, X^*) \) \((X^* \text{ is an SDR soln.})\), and modify \( \xi \) so that it is QCQP-feasible.

- Such a randomized QCQP soln. approx. may be performed multiple times, to get a better approx.

- (Believe it or not) The stochastic QCQP interpretation is the intuition behind many important theoretical SDR approx. accuracy results, including the famous Goemans-Williamson result [Goemans-Williamson’95].
Example: Randomization in BQP or MIMO Detection

A simple (and very important) example for illustrating randomizations is BQP:

\[
\begin{align*}
\min & \quad x^T C x \\
\text{s.t.} & \quad x_i^2 = 1, \quad i = 1, \ldots, n.
\end{align*}
\]  

(BQP)

Box 1. Gaussian Randomization Procedure for BQP

Given an SDR solution \( X^* \), and a number of randomizations \( L \),
for \( \ell = 1, \ldots, L \)

generate \( \xi_\ell \sim \mathcal{N}(0, X^*) \), and construct a feasible point

\[
\tilde{x}_\ell = \text{sgn}(\xi_\ell).
\]

End

Determine \( \ell^* = \arg \min_{\ell=1,\ldots,L} \tilde{x}_\ell^T C \tilde{x}_\ell \).

Output \( \hat{x} = \tilde{x}_{\ell^*} \) as an approximate solution to (BQP).
Performance of various no. of randomizations in MIMO detection. $M_t = M_r = 40$. 

W.-K. Ma, Dept. EE, The Chinese University of Hong Kong
Complex-valued QCQP and SDR

• Consider a general complex-valued QCQP

\[
\begin{align*}
\min_{x \in \mathbb{C}^n} & \quad x^H C x \\
\text{s.t.} & \quad x^H A_i x \succeq_i b_i, \quad i = 1, \ldots, m,
\end{align*}
\]

(1)

where \( C, A_1, \ldots, A_m \in \mathbb{H}^n \); \( \mathbb{H}^n \) denotes the set of \( n \times n \) Hermitian matrices.

• Using the same idea, SDR can be derived for complex-valued QCQP:

\[
\begin{align*}
\min_{X \in \mathbb{H}^n} & \quad \text{Tr}(C X) \\
\text{s.t.} & \quad X \succeq 0, \quad \text{Tr}(A_i X) \succeq_i b_i, \quad i = 1, \ldots, m.
\end{align*}
\]

The only difference is that the problem domain now is \( \mathbb{H}^n \) (change ‘symmetric’ to ‘hermitian’ in your CVX code).

• Note that while the ideas leading to real and complex SDRs are the same, their performance may be different (we will see this later).
**Application: Multicast Transmit Beamforming**

**Scenario:** Common information broadcast in multiuser MISO downlink, assuming channel state information at the transmitter (CSIT).

- The transmit signal:

  \[ x(t) = ws(t), \]

  where \( s(t) \in \mathbb{C} \) is the tx. data stream, & \( w \in \mathbb{C}^{N_t} \) is the tx. beamvector.

- Received signal for user \( i \):

  \[ y_i(t) = h_i^H x(t) + v_i(t), \]

  where \( h_i \in \mathbb{C}^{N_t} \) is the channel of user \( i \), & \( v_i(t) \) is noise with variance \( \sigma_i^2 \).
• Consider a QoS-assured design:

\[
\min_{\boldsymbol{w} \in \mathbb{C}^{N_t}} \|\boldsymbol{w}\|^2 \\
\text{s.t. } \text{SNR}_i \geq \gamma_i, \quad i = 1, \ldots, K,
\]

where each \(\gamma_i\) is a prescribed SNR requirement for user \(i\), and

\[
\text{SNR}_i = \mathbb{E}\{ |h_i^H \mathbf{s}(t)|^2 \}/\sigma_i^2 = \boldsymbol{w}^H \mathbf{R}_i \boldsymbol{w}/\sigma_i^2,
\]

\[
\mathbf{R}_i = \left\{ \begin{array}{ll}
\mathbf{h}_i \mathbf{h}_i^H, & \text{\(\mathbf{h}_i\) is available (instant CSIT)}, \\
\mathbb{E}\{\mathbf{h}_i \mathbf{h}_i^H\}, & \text{\(\mathbf{h}_i\) is random with known 2nd order stat. (stat. CSIT)}. 
\end{array} \right.
\]

• The design problem can be rewritten as a complex-valued QCQP

\[
\min \|\boldsymbol{w}\|^2 \\
\text{s.t. } \boldsymbol{w}^H \mathbf{A}_i \boldsymbol{w} \geq 1, \quad i = 1, \ldots, K,
\]

where \(\mathbf{A}_i = \mathbf{R}_i/\gamma_i \sigma_i^2\).

• This multicast problem is NP-hard in general, but can be approximated by SDR \[\text{[Sidiropoulos-Davidson-Luo'06]}.\]
A Randomization Example Relevant to Multicast Beamforming

Consider the problem

\[
\begin{align*}
\min & \quad x^H C x \\
\text{s.t.} & \quad x^H A_i x \geq 1, \quad i = 1, \ldots, m, \\
\end{align*}
\]  

(\dagger)

where \( C, A_1, \ldots, A_m \succeq 0 \).

Box 2. Gaussian Randomization Procedure for (\dagger)

given an SDR solution \( X^* \), and a number of randomizations \( L \).

for \( \ell = 1, \ldots, L \)

generate \( \xi_\ell \sim \mathcal{CN}(0, X^*) \), and construct a feasible point

\[
\tilde{x}_\ell = \frac{\xi_\ell}{\sqrt{\min_{i=1,\ldots,m} \xi_\ell^H A_i \xi_\ell}}
\]

end

determine \( \ell^* = \arg \min_{\ell=1,\ldots,L} \tilde{x}_\ell^H C \tilde{x}_\ell \).

output \( \hat{x} = \tilde{x}_{\ell^*} \) as an approximate solution to (\dagger).
Illustration of randomizations in $\mathbb{R}^2$, for Problem (†). The gray area is the feasible set and colored lines the contour of the objective.
Approximation accuracy of Gaussian randomization in multicast beamforming. $N_t = 4$, $K = 8$, $v(\mathbf{w}) = \|\mathbf{w}\|^2$ is the objective value, $v_{\text{SDR}}$ is the optimal value of SDR. Note that for any feasible $\mathbf{w}$, $v(\mathbf{w})/v_{\text{SDR}} \geq v_{\text{QP}}/v_{\text{SDR}}$ where $v_{\text{QP}}$ is the optimal value of QCQP. Courtesy to T.-H. Chang and Z.-Q. Luo.
Extension to Complex-Valued Separable QCQP

• Consider a further extension, called complex-valued separable QCQP:

\[
\begin{align*}
& \min_{x_1, \ldots, x_k \in \mathbb{C}^n} \sum_{i=1}^{k} x_i^H C_i x_i \\
& \text{s.t.} \quad \sum_{l=1}^{k} x_l^H A_{i,l} x_l \succeq b_i, \quad i = 1, \ldots, m.
\end{align*}
\]

• By writing \( X_i = x_i x_i^H \) for all \( i \), and then “semidefinite-relaxing” them, we obtain an SDR

\[
\begin{align*}
& \min_{X_1, \ldots, X_k \in \mathbb{H}^n} \sum_{i=1}^{k} \text{Tr}(C_i X_i) \\
& \text{s.t.} \quad \sum_{l=1}^{k} \text{Tr}(A_{i,l} X_l) \succeq b_i, \quad i = 1, \ldots, m, \\
& \quad X_1 \succeq 0, \ldots, X_k \succeq 0.
\end{align*}
\]
**Application: Unicast Transmit Downlink Beamforming**

**Scenario:** multiuser MISO downlink; each user receives an individual data stream.

- **Transmit signal:**
  \[ x(t) = \sum_{i=1}^{K} w_i s_i(t), \]
  where \( s_i(t) \in \mathbb{C} \) is the data stream for user \( i \), & \( w_i \in \mathbb{C}^{N_t} \) its tx. beamvector.

- **Received signal of user \( i \):**
  \[ y_i(t) = h_i^H x(t) + v_i(t) \]
  \[ = h_i^H w_i s_i(t) + \sum_{l \neq i} h_i^H w_i s_l(t) + v_i(t). \]
• The signal-to-interference-and-noise ratio (SINR) of user $i$:

$$\text{SINR}_i = \frac{w_i^H R_i w_i}{\sum_{l \neq i} w_l^H R_i w_l + \sigma_i^2},$$

where $R_i = h_i h_i^H$ for instant. CSIT, and $R_i = E\{h_i h_i^H\}$ for stat. CSIT.

• Consider the QoS-assured design:

$$\begin{align*}
\min_{w_1, \ldots, w_K \in \mathbb{C}^{N_t}} \quad & \sum_{i=1}^{K} \|w_i\|^2 \\
\text{s.t.} \quad & \frac{w_i^H R_i w_i}{\sum_{l \neq i} w_l^H R_i w_l + \sigma_i^2} \geq \gamma_i, \quad i = 1, \ldots, K
\end{align*}$$

(†)

and its SDR

$$\begin{align*}
\min_{W_1, \ldots, W_K \in \mathbb{H}^{N_t}} \quad & \sum_{i=1}^{K} \text{Tr}(W_i) \\
\text{s.t.} \quad & \text{Tr}(R_i W_i) \geq \gamma_i \left(\sum_{l \neq i} \text{Tr}(R_i W_l) + \sigma_i^2\right), \quad i = 1, \ldots, K, \quad (‡)
\end{align*}$$

• ($‡$) is shown to have a rank-one solution for $R_1, \ldots, R_K \succeq 0$, via uplink-downlink duality [Bengtsson-Ottersten'01]; SDR is optimal to (†), so to speak!

• We will introduce an “easy” way to identify rank-one SDR instances.
**SDR Versus Nonlinear Programming: They complement, not compete**

- Since SDR is an approximation method, as an alternative one may choose to approximate (QCQP) by a nonlinear programming method (NPM) (like, SQP in the MATLAB Optimization Toolbox).

- So should we compare SDR and NPM?

- The interesting argument is that they complement each other, instead of competing:
  - An NPM depends much on a ‘good’ starting point, and that’s usually the missing piece.
  - To SDR, NPMs may serve as a local refinement of the solution.

- One may consider a **two-stage approach** where SDR is used as a starting point for NPMs.
Application: Sensor Network Localization

The sensor network localization (SNL) problem is to determine the \((x, y)\) coordinates of the sensors, given distance information between sensors.

- In ad-hoc sensor networks, the sensor locations may not be known.

- A sensor may acquire its location by equipping it with GPS, but this may be too expensive.

- We may have several anchor sensors that have self-localization capability, though.

- Since sensors can communicate with each other, each sensor pair can work out their distance (e.g., by measuring the time-of-arrival info., or by ping-pong).

- The inter-sensor distances, together with anchor locations, can be used to estimate all the sensor locations in a joint fashion.
• Let \( \{x_1, \ldots, x_n\}, x_i \in \mathbb{R}^2 \) for all \( i \), be the collection of all (unknown) sensor coordinates.

• Let \( \{a_1, \ldots, a_m\}, a_i \in \mathbb{R}^2 \), be the collection of all (known) anchor coordinates.

• The distance between sensor \( i \) and sensor \( j \) is

\[
d_{ij} = \sqrt{(x_{i,1} - x_{j,1})^2 + (x_{i,2} - x_{j,2})^2} = \|x_i - x_j\|
\]

Likewise, the distance between sensor \( i \) and anchor \( j \) is

\[
\bar{d}_{ij} = \|x_i - a_j\|
\]

The obtained \( d_{ij} \) & \( \bar{d}_{ij} \) are assumed noiseless (extension for noisy cases available).

• The SNL problem here is that of finding \( x_1, \ldots, x_n \) such that

\[
\|x_i - x_j\|^2 = d_{ij}^2, \quad (i, j) \in E_{ss}
\]

\[
\|x_i - a_j\|^2 = \bar{d}_{ij}^2, \quad (i, j) \in E_{sa}
\]

where \( E_{ss} \) & \( E_{sa} \) are the sensor-to-sensor & sensor-to-anchor edge sets, resp.
Let $X = [x_1, \ldots, x_n] \in \mathbb{R}^{2 \times n}$. The SNL problem is written as

\[
\begin{align*}
\text{find} & \quad X \\
\text{s.t.} & \quad x_i^T x_i - 2x_i^T x_j + x_j^T x_j = d_{ij}^2, \quad (i,j) \in E_{ss} \\
& \quad x_i^T x_i - 2x_i^T a_j + a_j^T a_j = d_{ij}^2, \quad (i,j) \in E_{sa}
\end{align*}
\]

Let $Y = X^T X \in \mathbb{R}^{n \times n}$. The SNL problem have an equivalent formulation

\[
\begin{align*}
\text{find} & \quad X, Y \\
\text{s.t.} & \quad Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss} \\
& \quad Y_{ii} - 2x_i^T a_j + a_j^T a_j = d_{ij}^2, \quad (i,j) \in E_{sa} \\
& \quad Y = X^T X
\end{align*}
\]

Naturally (and after you have seen how SDR operates), the SNL problem can be approximated by SDR:

\[
\begin{align*}
\text{find} & \quad X, Y \\
\text{s.t.} & \quad Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss} \\
& \quad Y_{ii} - 2x_i^T a_j + a_j^T a_j = d_{ij}^2, \quad (i,j) \in E_{sa} \\
& \quad Y \preceq X^T X
\end{align*}
\]
SDR (ML-SNL formulation), plus a 2nd-stage solution refinement by gradient descent. The distance measurements are noisy. ○: true sensor locations; ♦: anchor locations; *: SDR solution; —: gradient descent trajectory (50 iterations).
Gradient descent ML-SNL with a random starting point. ⭕: true sensor locations; ⭐: anchor locations; — : gradient descent trajectory (50 iterations).
Part II: Theory
Provable Approximation Accuracies

The following problem has been of great interest to optimization theorists, and it has enormous implications in practice.

• Let $v(x) = x^T C x$, and denote the optimal values of (QCQP) and (SDR) by

$$v_{QP} = \min_{x} x^T C x$$
$$\text{s.t. } x^T A_i x \succeq b_i, \quad i = 1, \ldots, m$$

$$v_{SDR} = \min \operatorname{Tr}(C X)$$
$$\text{s.t. } X \succeq 0, \quad \operatorname{Tr}(A_i X) \succeq b_i, \quad i = 1, \ldots, m$$

Moreover, let $\hat{x}$ be an approximate solution of (QCQP), say, using randomization. Note that

$$v_{QP} \leq v(\hat{x}).$$

• The problem is to prove a constant $\gamma$ such that

$$v(\hat{x}) \leq \gamma v_{QP}$$

in a worst case sense, or with high probability. A $\gamma$ close to 1 would mean a near-optimal accuracy.
The Seminal Approx. Accuracy Result by Goemans & Williamson

- Consider

$$v_{QP} = \max_{x \in \mathbb{R}^n} x^T C x$$

s.t. $$x_i^2 = 1, \quad i = 1, \ldots, n$$

with $$C \succeq 0, \quad C_{ij} \leq 0$$ for all $$i \neq j$$ (the so-called MAXCUT in network optimization).

- In [Goemans-Williamson’95], it was shown that if the randomization procedure is used, then

$$\gamma v_{QP} \leq \mathbb{E}\{v(\hat{x})\} \leq v_{QP}$$

where $$\gamma \approx 0.87856$$.

- The work by Goeman and Williamson has triggered much interest, resulting in many more approx. accuracy results being established for a wider class of problems.
**Approx. Accuracy Result for Quadratic Minimization**

- Consider now the problem

\[
v_{QP} = \min_{x \in \mathbb{C}^n} \quad x^H C x \\
\text{s.t.} \quad x^H A_i x \geq 1, \quad i = 1, \ldots, m
\]

for \( C, A_1, \ldots, A_m \succeq 0 \), which arises in multicast downlink beamforming.

- It was shown in [Luo-Sidiropoulos-Tseng-Zhang'07] that if the randomization procedure in Box 2 is used, then with high probability (instead of just in expectation),

\[
v_{QP} \leq v(\hat{x}) \leq \gamma v_{QP},
\]

where \( \gamma = 8m \).

  - In multicast beamforming, this result says that we can produce a transmit beamforming vector that satisfies all the prescribed SNR requirements and whose power is at most \( 8m \) times the optimal.

  - Notice that this ratio accommodates the worst possible problem instance \( \{C, A_1, \ldots, A_m\} \). In practice, the approximation accuracies are usually much better—a phenomenon that deserves further investigation.
**Known approximation accuracies for quadratic maximization problems. The reference numbers refer to those in our Signal Processing Magazine article.**

<table>
<thead>
<tr>
<th>problem</th>
<th>approx. accuracy $\gamma$; see (21)-(22) for def.</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Boolean QP</strong></td>
<td></td>
<td>Goemans-Williamson [2], Nesterov [3], Zhang [6]. Relevant applications: [24]–[26]</td>
</tr>
<tr>
<td>$\max_{x \in \mathbb{R}^n} x^T C x$ s.t. $x_i^2 = 1, \ i = 1, \ldots, n$</td>
<td>$\gamma = \begin{cases} 0.87856, &amp; C \succeq 0, \ C_{ij} \leq 0 \ \forall i \neq j \ 2/\pi \simeq 0.63661, &amp; C \succeq 0 \ 1 \ \text{(opt.)}, &amp; C_{ij} \geq 0, \ \forall i \neq j \end{cases}$</td>
<td></td>
</tr>
<tr>
<td><strong>Complex k-ary QP</strong></td>
<td>For $C \succeq 0$, $\gamma = \frac{(k \sin(\pi/k))^2}{4\pi}$, e.g., $\gamma = 0.7458$ for $k = 8$, $\gamma = 0.7754$ for $k = 16$.</td>
<td>Zhang-Huang [7], So-Zhang-Ye [8]. Relevant applications: [27], [37]</td>
</tr>
<tr>
<td>$\max_{x \in \mathbb{C}^n} x^H C x$ s.t. $x_i \in {1, \omega, \ldots, \omega^{k-1}}, \ i = 1, \ldots, n$</td>
<td>For $C \succeq 0$, $\gamma = \frac{\pi}{4} = 0.7854$. Remark: coincide with complex $k$-ary QP as $k \to \infty$.</td>
<td>Zhang-Huang [7], So-Zhang-Ye [8].</td>
</tr>
<tr>
<td>where $\omega = e^{2\pi i / k}$, and $k &gt; 1$ is an integer.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Complex constant-modulus QP</strong></td>
<td>The same approx. ratio as in complex constant-modulus QP, i.e., $\gamma = \frac{\pi}{4}$ for $C \succeq 0$. If the problem is reduced to the real-valued case, then the approx. ratio results are the same as that in Boolean QP.</td>
<td>Ye [4], Zhang [6].</td>
</tr>
<tr>
<td>$\max_{x \in \mathbb{C}^n} x^H C x$ s.t. $</td>
<td>x_i</td>
<td>^2 = 1, \ i = 1, \ldots, n$</td>
</tr>
<tr>
<td>where $\mathcal{F} \subset \mathbb{R}^n$ is a closed convex set.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>problem</td>
<td>approx. accuracy $\gamma$; see (18)-(19) for def.</td>
<td>references</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>$\min_{x \in \mathbb{C}^n} x^H C x$ s.t. $x^H A_i x \geq 1$, $i = 1, \ldots, m$ where $A_1, \ldots, A_m \succeq 0$.</td>
<td>$\gamma = 8m$. If the problem is reduced to the real-valued case, then $\gamma = \frac{27m^2}{\pi}$.</td>
<td>Luo-Sidiropoulos-Tseng-Zhang [9]; see also So-Ye-Zhang [71]. Relevant applications: [29]</td>
</tr>
<tr>
<td>MIMO Detection $\min_{x \in \mathbb{R}^n} |y - Hx|_2^2$ s.t. $x_i^2 = 1$, $i = 1, \ldots, n$ where $y = Hs + v$; $H \in \mathbb{C}^{n \times n}$ has i.i.d. standard complex Gaussian entries; $s_i^2 = 1$ for $i = 1, \ldots, n$; and $v \in \mathbb{C}^n$ has i.i.d. complex mean zero Gaussian entries with variance $\sigma^2$.</td>
<td>For $\sigma^2 \geq 60n$ (which corresponds to the low signal-to-noise ratio (SNR) region), with probability at least $1 - 3 \exp(-n/6)$, $\gamma \leq \frac{11}{2}$. For $\sigma^2 = \mathcal{O}(1)$ (which corresponds to the high SNR region), with probability at least $1 - \exp(-\mathcal{O}(n))$, $\gamma = 1$, i.e. the SDR is tight.</td>
<td>Kisialiou-Luo [67], So [69]. Extensions: So [68], [69]. Related: Jaldén-Ottersten [66]. Relevant applications: [17]-[20], [22], [23]</td>
</tr>
</tbody>
</table>

Known approximation accuracies for quadratic minimization problems. The reference numbers refer to those in our Signal Processing Magazine article.
Rank Reduction in SDR

• Now you may notice that an SDR methodology basically has the following steps:
  1) formulate a hard problem (nonconvex QCQP) as a rank-one-constrained SDP
  2) remove the rank constraint to obtain an SDP
  3) use some methods, such as randomizations, to produce an approximate solution to the original problem.

• Apparently, the lower the rank of the SDP solution, the better the approximation we would expect.

• Unfortunately, we cannot guarantee a low rank solution for the SDP in general.

• But we can identify special cases where the SDP solution rank is low, and, sometimes, one.
Shapiro-Barvinok-Pataki (SBP) Result

- Consider the real-valued SDP (or SDR)

\[
\min_{X \in S^n} \quad \text{Tr}(CX) \\
\text{s.t.} \quad X \succeq 0, \quad \text{Tr}(A_i X) \geq b_i, \quad i = 1, \ldots, m
\]  

(SDR)

SBP Result [Pataki’98]: there exists an optimal solution \(X^*\) such that

\[
\frac{\text{rank}(X^*)\text{rank}(X^*) + 1}{2} \leq m
\]

- In particular, SBP result implies that for \(m \leq 2\), a rank-1 \(X^*\) exists. Hence,

For a real-valued QCQP with \(m \leq 2\), SDR is tight; i.e., solving the SDR is equivalent to solving the original QCQP.

- Note that a rank reduction algorithm may be required to turn an SDP solution to a rank-one solution [Ye-Zhang’03].
Complex Extension of the Rank Reduction Result

• Let us consider the extension to the complex-valued SDP

\[
\min_{X \in \mathbb{H}^n} \quad \text{Tr}(CX) \\
\text{s.t.} \quad X \succeq 0, \quad \text{Tr}(A_i X) \succeq b_i, \quad i = 1, \ldots, m
\]

• In this case, the SBP result can be generalized to [Huang-Palomar’09]

\[
\text{rank}(X^*)^2 \leq m
\]

and the direct consequence is that

For a complex-valued QCQP with \( m \leq 3 \), SDR is tight; i.e., solving the SDR is equivalent to solving the original QCQP.

• A complex rank-1 decomposition algorithm for \( m \leq 3 \) is available [Huang-Zhang’07].
Application Revisited: Multicast Beamforming

- Recall the multicast beamforming problem:

\[
\min_{w \in \mathbb{C}^{N_t}} \|w\|^2
\]

s.t. \[\text{SNR}_i = \frac{w R_i w}{\sigma_i^2} \geq \gamma_i, \]
\[i = 1, \ldots, K,\]

\[K\] being the number of users.

- By the SBP result, \textit{SDR solves the multicast problem optimally for } \(K \leq 3\).
Further Extension of the Rank Reduction Result

- Recall the problem

\[
\begin{align*}
\min_{X_1, \ldots, X_k \in \mathbb{H}^n} & \quad \sum_{i=1}^{k} \text{Tr}(C_i X_i) \\
\text{s.t.} & \quad \sum_{l=1}^{k} \text{Tr}(A_{i,l} X_i) \geq b_i, \quad i = 1, \ldots, m, \\
& \quad X_1 \succeq 0, \ldots, X_k \succeq 0,
\end{align*}
\]

which is an SDR of the so-called separable QCQP.

- A generalization of the SBP result [Huang-Palomar'09]:

\[
\sum_{i=1}^{k} \text{rank}(X_i^*)^2 \leq m,
\]

and, as a subsequent result:

Suppose that an SDR solution cannot have \(X_i^* = 0\) for any \(i\). Then SDR is tight for \(m \leq k + 2\).
Application Revisited: Unicast Beamforming

- Recall the design problem

\[
\begin{align*}
\min_{\mathbf{w}_1, \ldots, \mathbf{w}_K \in \mathbb{C}^{N_t}} & \quad \sum_{i=1}^{K} \|\mathbf{w}_i\|^2 \\
\text{s.t.} & \quad \frac{\mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i}{\sum_{l \neq i} \mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_l + \sigma_i^2} \geq \gamma_i, \quad (\dagger)
\end{align*}
\]

which is a separable QCQP with \( K \) variables (beamvectors) and \( K \) constraints (SINR req.).

- By the SBP rank reduction result, SDR solves (\( \dagger \)) optimally for any \( \mathbf{R}_1, \ldots, \mathbf{R}_K \), regardless of \( \mathbf{R}_i \succeq 0 \) or not.

- And hey, it’s still fine if you put two more quadratic constraints in (\( \dagger \))!
Cognitive Radio (CR) Beamforming: A Further Example

- **Goal:** access the channel owned by primary users (PUs) through spectrum sharing.

- **Idea:** the CR system avoids excessive interference to the PUs through tx. opt.

- **Scenario:** MISO downlink with the CR (or secondary) system, either unicast or multicast; \( K \) secondary users (SUs); \( L \) single-antenna PUs

- **Assume known CSIT from the secondary transmitter to the PUs.**
• Consider the multicast case.
  – tx. and rx. model for SUs: same as the previous multicast model.
  – Interference to the $l$th PU:
    \[ |g_l^H w|^2 \]
    where $g_l$ is the channel from the secondary transmitter to the $l$th PU.

• Design problem [Phan-Vorobyov-Sidiropoulos-Tellambura’09]:

\[
\min_w \|w\|^2 \\
\text{s.t. } \text{SNR}_{SU,i} = w^H R_k w / \sigma_k^2 \geq \gamma_k, \ k = 1, \ldots, K, \\
    w^H G_l w \leq \delta_l, \ l = 1, \ldots, L \quad \text{(interference temperature (IT) constraints)}
\]

where $G_l$ is the CSIT of $l$th PU (defined in the same way as $R_k$), $\delta_l$ is the tolerable interference level to $l$ PU, & $\gamma_k$ are SUs’ SNR requirements.

• By the SBP rank result, SDR is optimal when $K \leq 2$, $L = 1$ ($\leq 2$ SUs, 1 PU).
• CR BF design for the unicast case (see, e.g., [Zhang-Liang-Cui’10]):

\[
\min_{\mathbf{w}_1, \ldots, \mathbf{w}_K} \sum_{k=1}^{K} \| \mathbf{w}_k \|^2
\]

s.t. \( \text{SINR}_{SU,i} = \frac{\mathbf{w}_k^H \mathbf{R}_k \mathbf{w}_k}{\sum_{l \neq k} \mathbf{w}_l^H \mathbf{R}_k \mathbf{w}_l + \sigma_k^2} \geq \gamma_k, \ k = 1, \ldots, K, \)

\[
\sum_{k=1}^{K} \mathbf{w}_k^H \mathbf{G}_l \mathbf{w}_k \leq \delta_l, \ l = 1, \ldots, L \quad \text{(IT constraints)}
\]

• A separable QCQP with \( K \) variables and \( K + L \) constraints.

• By the SBP rank result, SDR solves the problem if \( L \leq 2 \).

• Remark: For instance, CSIT with SUs, SDR can be shown to be optimal for any \( L \). Or it can be reformulated, and then solved, by SOCP.
Part III: Frontier Development
Transmit Beamforming

• Transmit beamforming is now a key topic; see [Gershman-Sidiropoulos-Shahbazpanahi-Bengtsson-Ottersten’10], [Luo-Chang’10] for review.

• Apart from standard transmit beamforming, we have seen numerous extensions:
  – one-way relay beamforming [Fazeli-Dehkordy-Shahbazpanahi-Gazor’09], [Chalise-Vandendorpe’09]
  – two-way relay beamforming (a.k.a. analog network coding) [Zhang-Liang-Chai-Cui’09]
  – cognitive radio beamforming [Zhang-Liang-Cui’10]
  – multicell coordinated beamforming [Bengtsson-Ottersten’01], [Dahrouj-Yu’10]
  – secrecy beamforming [Liao-Chang-Ma-Chi’10], [Li-Ma’11],

• Interestingly, all these beamforming problems turn out to be, or be closely related to, nonconvex QCQPs.

• And, as it turns out, SDR plays a key role.
Frontier Problem: Outage-Based Unicast Transmit Beamforming

CSIT is generally imperfectly known in practice.

- Suppose that the presumed CSIT, \( \{ h_i \} \), is inaccurate.

- If we directly substitute the presumed CSIT into the standard QoS-assured design

\[
\min_{w_1, \ldots, w_K \in \mathbb{C}^N} \sum_{i=1}^{K} \| w_i \|^2 \\
\text{s.t.} \quad \frac{| w_i^H h_i |^2}{\sum_{l \neq i} | w_i^H h_i |^2 + \sigma_i^2} \geq \gamma_i,
\]

and run it, then the resultant design may have severe SINR outage.
Histogram of the actual SINR satisfaction probabilities of the non-robust QoS-assured design. $N_t = K = 3$; i.i.d. complex Gaussian CSI errors with zero mean and variance $0.002$; $\gamma = 11\text{dB}$. The design has more than 50% outage most of the time.
Outage-Based Unicast Transmit Beamforming: Formulation

- Let us assume that $h_i \sim \mathcal{CN}(\bar{h}_i, \sigma_e^2 I)$, where $\bar{h}_i$ is the presumed channel, and $\sigma_e^2$ is the CSI uncertainty variance.

- A meaningful, but very difficult, design problem:

$$\min_{w_1, \ldots, w_K \in \mathbb{C}^N} \sum_{i=1}^{K} \|w_i\|^2$$

s.t. $$\text{Prob}_{h_i \sim \mathcal{CN}(\bar{h}_i, \sigma_e^2 I)} \left\{ \frac{w_i^H h_i h_i^H w_i}{\sum_{l \neq i} w_l^H h_l h_l^H w_l + \sigma_i^2} \geq \gamma_i \right\} \geq 1 - \rho_i, \quad i = 1, \ldots, K,$$

where the $\rho_i$’s are the maximum tolerable outage probabilities.
• The outage-based SINR constraints

\[
\text{Prob}_{h_i \sim \mathcal{CN}(\bar{h}_i, \sigma^2_i I)} \left\{ \frac{w_i^H h_i h_i^H w_i}{\sum_{l \neq i} w_l^H h_i h_i^H w_l + \sigma^2_i} \geq \gamma_i \right\} \geq 1 - \rho_i
\]

can be rewritten as

\[
\text{Prob}_{e_i \sim \mathcal{CN}(0, \sigma^2_i I)} \left\{ (\bar{h}_i + e_i)^H \left( \frac{1}{\gamma_i} w_i w_i^H - \sum_{l \neq i} w_l w_l^H \right) (\bar{h}_i + e_i) \geq \sigma^2_i \right\} \geq 1 - \rho_i.
\]

• Challenges:
  – The probability on the LHS has no simple closed form expression.
  – The quadratic function

\[
(\bar{h}_i + e_i)^H \left( \frac{1}{\gamma_i} w_i w_i^H - \sum_{l \neq i} w_l w_l^H \right) (\bar{h}_i + e_i)
\]

is indefinite (and hence nonconvex) in the design variables \(w_1, \ldots, w_K\).
Let us first do the thing we are good at — SDR.

By SDR, we have

\[
\Pr_{e_i \sim \mathcal{CN}(0, \sigma_e^2 I)} \left\{ (\bar{h}_i + e_i)^H \left( \frac{1}{\gamma_i} W_i - \sum_{l \neq i} W_l \right) (\bar{h}_i + e_i) \geq \sigma_i^2 \right\} \geq 1 - \rho_i.
\]

Now, the function

\[
(\bar{h}_i + e_i)^H \left( \frac{1}{\gamma_i} W_i - \sum_{l \neq i} W_l \right) (\bar{h}_i + e_i)
\]

is linear in the variables \( W_1, \ldots, W_k \), which is good.

However, the probability still does not admit a simple closed form expression.
Processing the Probabilistic Constraint: Convex Restriction

• Let

\[ V_i(\{W_j\}) = \text{Prob}_{e_i \sim \mathcal{C}\mathcal{N}(0, \sigma_i^2 I)} \left\{ (\bar{h}_i + e_i)^H \left( \frac{1}{\gamma_i} W_i - \sum_{l \neq i} W_l \right) (\bar{h}_i + e_i) < \sigma_i^2 \right\} \]

be the violation probability. Recall that we want

\[ V_i(\{W_j\}) \leq \rho_i. \]

• It is not hard to see that \( V_i \) can be expressed as

\[ V_i(\{W_j\}) = \text{Prob}_{e \sim \mathcal{C}\mathcal{N}(0, I)} \left\{ e^H Q e + 2\text{Re}\{e^H r\} + s < 0 \right\} \]

for some \( Q, r \) and \( s \) that depend on \( W_1, \ldots, W_K \) and the index \( i \). (Here and in the sequel, we drop the index \( i \) for notational simplicity.)
• To process the violation probability $V_i$, another idea is to find an efficiently computable convex function $f(Q, r, s, t)$, where $t$ is an additional decision vector, such that

$$V_i(\{W_j\}) = \text{Prob}_{e \sim \mathcal{C} \mathcal{N}(0, I)} \left\{ e^H Q e + 2 \text{Re}\{e^H r\} + s < 0 \right\} \leq f(Q, r, s, t).$$

• Then, by construction, the convex constraint

$$f(Q, r, s, t) \leq \rho$$  \hspace{1cm} (CR-PC)

serves as a sufficient condition for the probabilistic constraint

$$V_i(\{W_j\}) \leq \rho$$  \hspace{1cm} (PC)

to hold. We call (CR-PC) a convex restriction of (PC).
Finding the Convex Restriction

• Can we find such a convex function? Does it even exist? The answer is: Yes! (And there are many such functions.)

• For instance, we can employ a Bernstein-type inequality $[Bechar2009]$, which states that

$$\text{Prob}_{e \sim \mathcal{CN}(0,I)}\{e^H Q e + 2 \text{Re}\{e^H r\} + s < 0\} \leq e^{-T^{-1}(s)},$$

where $T(\eta) = \text{Tr}(Q) - \sqrt{2\eta}\sqrt{\|Q\|_F^2 + \|r\|^2} - \eta \max\{\lambda_{\max}(-Q), 0\}$.

• Is the constraint $e^{-T^{-1}(s)} \leq \rho$ convex? Yes! It is equivalent to

$$\text{Tr}(Q) - \sqrt{-2 \ln(\rho)} \cdot t_1 + \ln(\rho) \cdot t_2 + s \geq 0,$$
$$\sqrt{\|Q\|_F^2 + 2\|r\|^2} \leq t_1,$$
$$t_2 I + Q \succeq 0,$$
$$t_2 \geq 0.$$
Putting Things Together: The Relaxation-Restriction Approach

- Applying the Bernstein-type inequality to the SDR’ed SINR constraints (with some additional work), a convex relaxation-restriction approximation is developed [Wang-Chang-Ma-So-Chi’11].

- **A mysterious finding in simulations**: rank-one SDR solution is obtained in almost all the problem instances!
Histogram of the actual SINR satisfaction probabilities of the proposed SDR+Bernstein method. $N_t = K = 3$; i.i.d. complex Gaussian CSI errors with zero mean and variance 0.002; $\gamma = 11\text{dB}$; $\rho = 0.1$ (90\% SINR satisfaction).
Feasibility performance of the proposed method and the probabilistic SOCP method [Shenouda-Davidson'08]. $N_t = K = 3; \sigma_e^2 = 0.002; \gamma = 11\text{dB}; \rho = 0.1$ (90% SINR satisfaction).
Transmit power performance of the proposed method and the probabilistic SOCP method. $N_t = K = 3; \sigma_e^2 = 0.002; \rho = 0.1$ (90% SINR satisfaction).
References


[Fazeli-Dehkordy-Shahbazpanahi-Gazor’09]  S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor,


