Semidefinite Relaxation of Nonconvex Quadratic Optimization: A Key Technique in Signal Processing and Communications

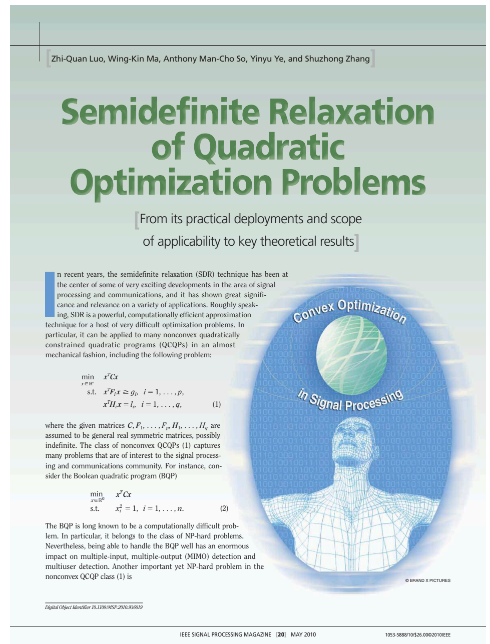
Wing-Kin (Ken) Ma
Dept. Electronic Engineering
The Chinese University of Hong Kong
Hong Kong

Anthony Man-Cho So
Dept. Systems Engg. & Engg. Management
The Chinese University of Hong Kong
Hong Kong

EUSIPCO 2011 Tutorial, Aug 29, 2011

- A recent tutorial article:

> Z.-Q. Luo, W.-K. Ma, Anthony M.-C. So, Y. Ye, & S. Zhang, "Semidefinite relaxation of quadratic optimization problems," in *IEEE SP Magazine*, Special Issue on *Convex Optimization for Signal Process.*, May 2010.


[Zhi-Quan Luo, Wing-Kin Ma, Anthony Man-Cho So, Yinyu Ye, and Shuzhong Zhang]

Semidefinite Relaxation of Quadratic Optimization Problems

[From its practical deployments and scope of applicability to key theoretical results]

- Acknowledgment: Tom Luo, Yinyu Ye, & Shuzhong Zhang for co-authoring the magazine article; Qiang Li & Jiaxian Pan for helping prepare this tutorial note.

- Many thanks also go to everyone who has contributed to this powerful tool.

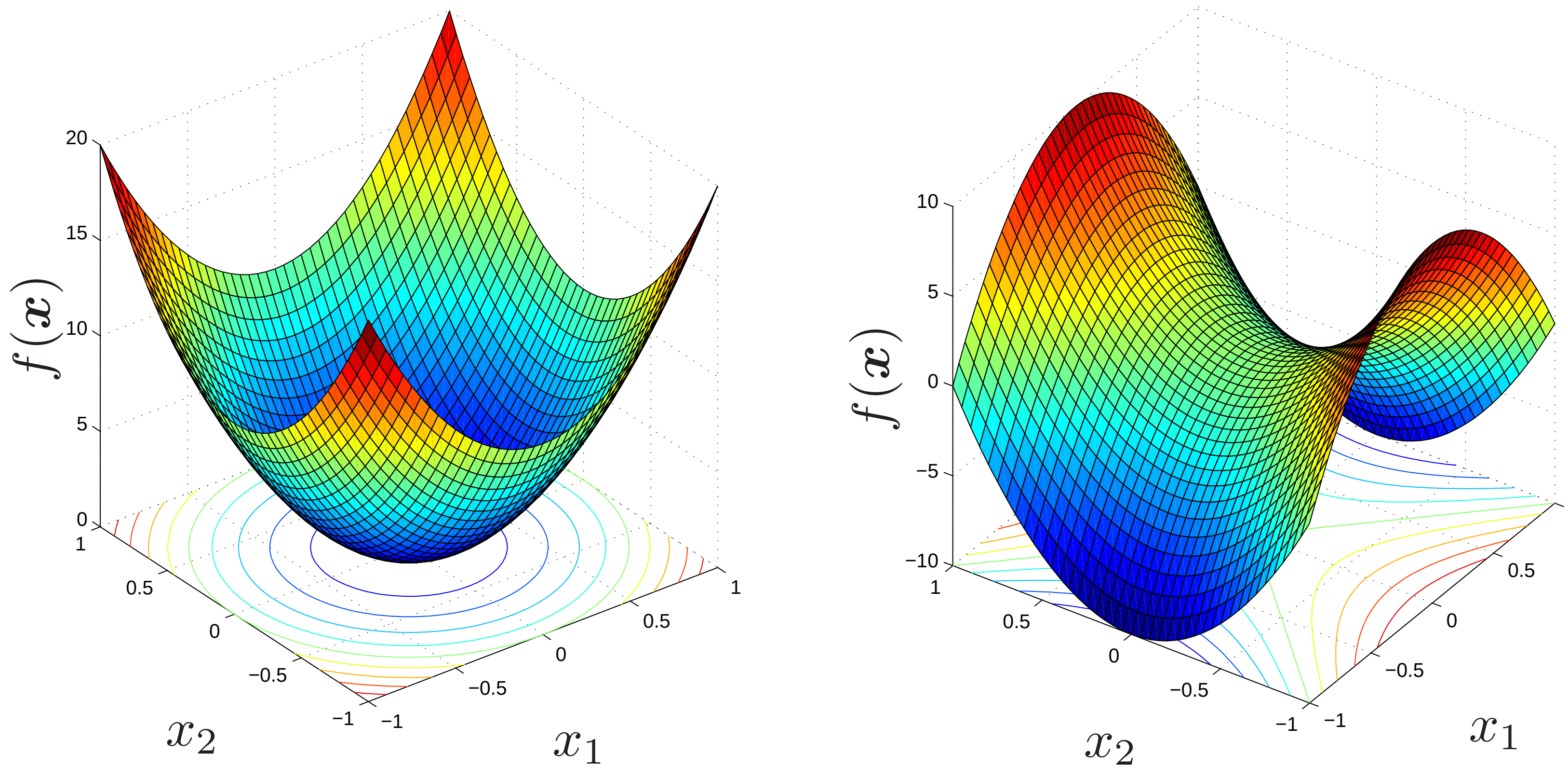
Outline

- Part I: Basic concepts and overview of semidefinite relaxation (SDR)
- Part II: Theory, and implications in practice
- Part III: Applications
  - A. MIMO detection
  - B. Sensor network localization
  - C. Transmit beamforming

Part I: Basic Concepts and Overview

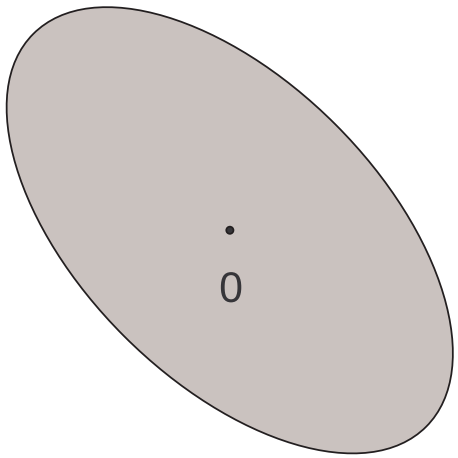
A quick reminder of what convex quadratic functions & constraints are:

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

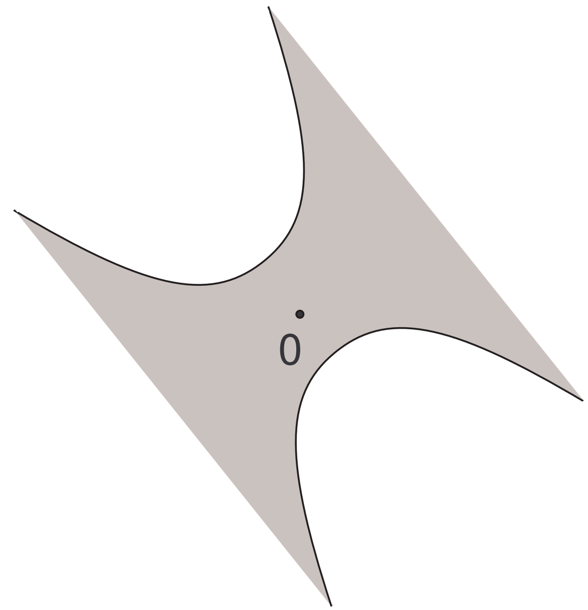


(a) $\boldsymbol{C} \succeq \mathbf{0}$. **(b)** $\boldsymbol{C} \not\succeq \mathbf{0}$.

- A constraint set $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}^T \boldsymbol{F} \boldsymbol{x} \leq 1\}$ is convex if and only if $\boldsymbol{F} \succeq \boldsymbol{0}$.



(a) $\boldsymbol{F} \succeq \boldsymbol{0}$.



(b) $\boldsymbol{F} \not\succeq \boldsymbol{0}$.

- A constraint set $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}^T \boldsymbol{F} \boldsymbol{x} = 1\}$ is nonconvex.

Quadratically Constrained Quadratic Program

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

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T\boldsymbol{F}_i\boldsymbol{x} \geq g_i, \quad i = 1,\ldots,p, \\ & \boldsymbol{x}^T\boldsymbol{H}_i\boldsymbol{x} = l_i, \quad i = 1,\ldots,q, \end{aligned}$$

where $\boldsymbol{C}, \boldsymbol{F}_1, \ldots, \boldsymbol{F}_p, \boldsymbol{H}_1, \ldots, \boldsymbol{H}_q \in \mathbb{S}^n$; $\mathbb{S}^n$ is the set of all $n \times n$ real symmetric matrices.

- We do not consider convex cases, and $\boldsymbol{C}, \boldsymbol{F}_i, \boldsymbol{H}_i$ may be arbitrary.

- Nonconvex QCQP is a very difficult problem in general.

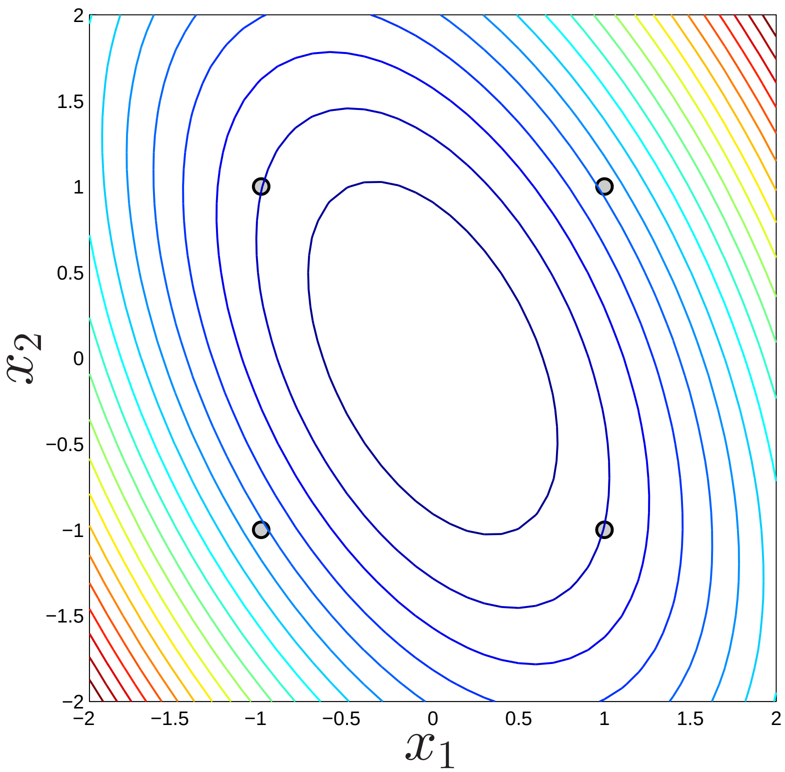
Nonconvex QCQP: How Hard Could it Be?

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

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & x_i^2 = 1, \quad i = 1, \dots, n, \end{aligned}$$

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 $\mathcal{O}(2^n)$, 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 $\mathcal{O}(n^p)$ for any $p > 0$.



Nonconvex QCQP: How Hard Could it Be?

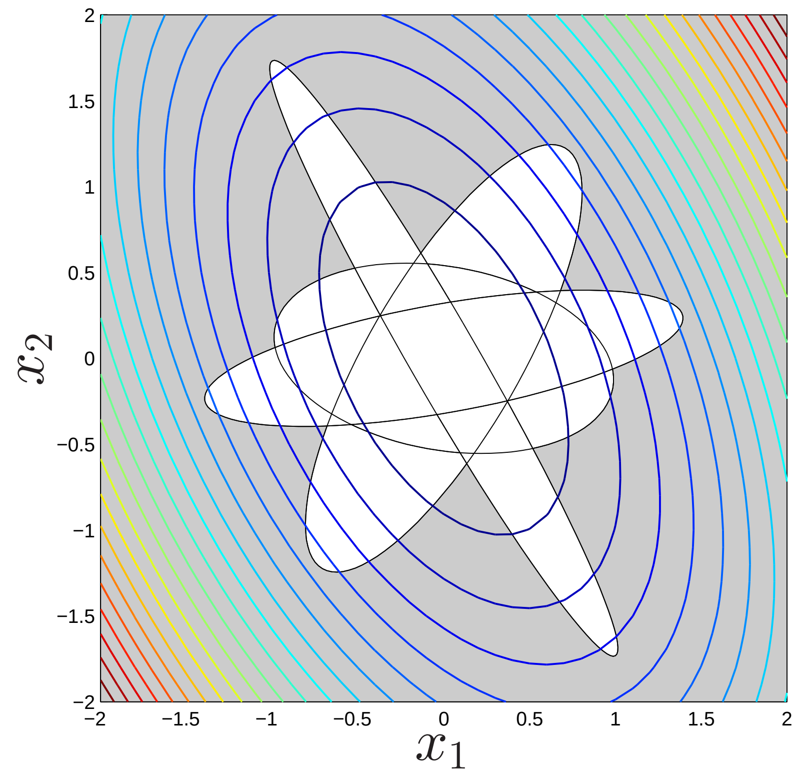
Consider the following problem

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T \boldsymbol{F}_i \boldsymbol{x} \geq 1, \quad i = 1, \ldots, m, \end{aligned}$$

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

- 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

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} \trianglerighteq_i b_i, \quad i=1,\ldots,m. \end{aligned} \tag{QCQP}$$

  Here, '$\trianglerighteq_i$' can represent either '$\geq$', '$=$', or '$\leq$' for each i; $\boldsymbol{C},\boldsymbol{A}_1,\ldots,\boldsymbol{A}_m \in \mathbb{S}^n$; and $b_1,\ldots,b_m \in \mathbb{R}$.

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

$$\boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} = \mathrm{Tr}(\boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}) = \mathrm{Tr}(\boldsymbol{C}\boldsymbol{x}\boldsymbol{x}^T), \quad \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} = \mathrm{Tr}(\boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x}) = \mathrm{Tr}(\boldsymbol{A}_i\boldsymbol{x}\boldsymbol{x}^T),$$

  or, if we let $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$,

$$\boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} = \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}), \quad \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} = \mathrm{Tr}(\boldsymbol{A}_i\boldsymbol{X}).$$

- The objective and constraint functions are **linear** in $\boldsymbol{X}$.

The Concept of SDR

- The condition $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$ is equivalent to $\boldsymbol{X} \succeq \boldsymbol{0}$, $\mathrm{rank}(\boldsymbol{X}) = 1$, thus (QCQP) is the same as

$$\begin{aligned} \min_{\boldsymbol{X} \in \mathbb{S}^n} \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1, \ldots, m \\ & \boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{rank}(\boldsymbol{X}) = 1. \end{aligned} \tag{QCQP}$$

- The constraints $\mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i$ are easy, but $\mathrm{rank}(\boldsymbol{X}) = 1$ is hard.

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

$$\begin{aligned} \min_{\boldsymbol{X} \in \mathbb{S}^n} \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1, \ldots, m, \\ & \boldsymbol{X} \succeq \boldsymbol{0}. \end{aligned} \tag{SDR}$$

(SDR) is a convex problem.

Some Merits We Can Immediately Say

- The SDR

$$\begin{aligned} \min_{\boldsymbol{X}\in\mathbb{S}^n} \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{Tr}(\boldsymbol{A}_i\boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1,\ldots,m \end{aligned} \tag{SDR}$$

  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 $\mathcal{O}(\max\{m,n\}^4 n^{1/2} \log(1/\epsilon))$, ϵ being soln. accuracy).

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

```
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
```

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 $\boldsymbol{X}^\star$, is of rank one; or, equivalently,

$$\boldsymbol{X}^\star = \boldsymbol{x}^\star \boldsymbol{x}^{\star T},$$

  then $\boldsymbol{x}^\star$ 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 $\boldsymbol{X}^\star$, for instances where $\mathrm{rank}(\boldsymbol{X}^\star) > 1$.

QCQP Solution Approximation in SDR: An Example

- Consider again the BQP

$$\begin{array}{ll} \min & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} & x_i^2 = 1, \quad i = 1, \ldots, n. \end{array} \tag{BQP}$$

  The SDR of (BQP) is

$$\begin{array}{ll} \min & \text{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} & \boldsymbol{X} \succeq \boldsymbol{0}, \quad X_{ii} = 1, \; i = 1, \ldots, n. \end{array} \tag{SDR}$$

- An intuitively reasonable idea (true even for engineers) is to apply a rank-1 approximation to the SDR solution $\boldsymbol{X}^\star$:

  1) Carry out the eigen-decomposition

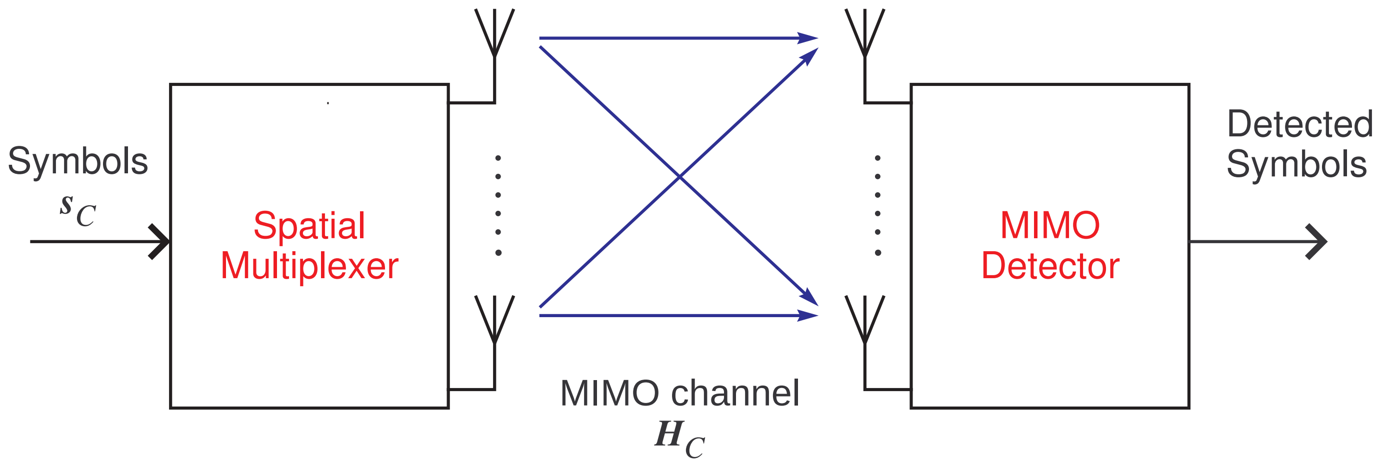
$$\boldsymbol{X}^\star = \textstyle\sum_{i=1}^r \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^T,$$

     where $r = \text{rank}(\boldsymbol{X}^\star)$, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ are the eigenvalues and $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_r \in \mathbb{R}^n$ the respective eigenvectors.

  2) Approximate the BQP by $\hat{\boldsymbol{x}} = \text{sgn}(\sqrt{\lambda_1}\boldsymbol{q}_1)$.

Application: MIMO Detection

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



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

- Received signal model:

$$\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C$$

where $\boldsymbol{H}_C \in \mathbb{C}^{M_r \times M_t}$ is the MIMO channel, $\boldsymbol{s}_C \in \mathbb{C}^{M_t}$ is the transmitted symbol vector, & $\boldsymbol{v}_C \in \mathbb{C}^{M_r}$ is complex circular Gaussian noise.

- Assume QPSK constellations, $\boldsymbol{s}_C \in \{\pm 1 \pm j\}^{M_t}$.

- Problem: maximum-likelihood (ML) detection (NP-hard)

$$\hat{\boldsymbol{s}}_{C,ML} = \arg \min_{\boldsymbol{s}_C \in \{\pm 1 \pm j\}^{M_t}} \|\boldsymbol{y}_C - \boldsymbol{H}_C \boldsymbol{s}_C\|^2.$$

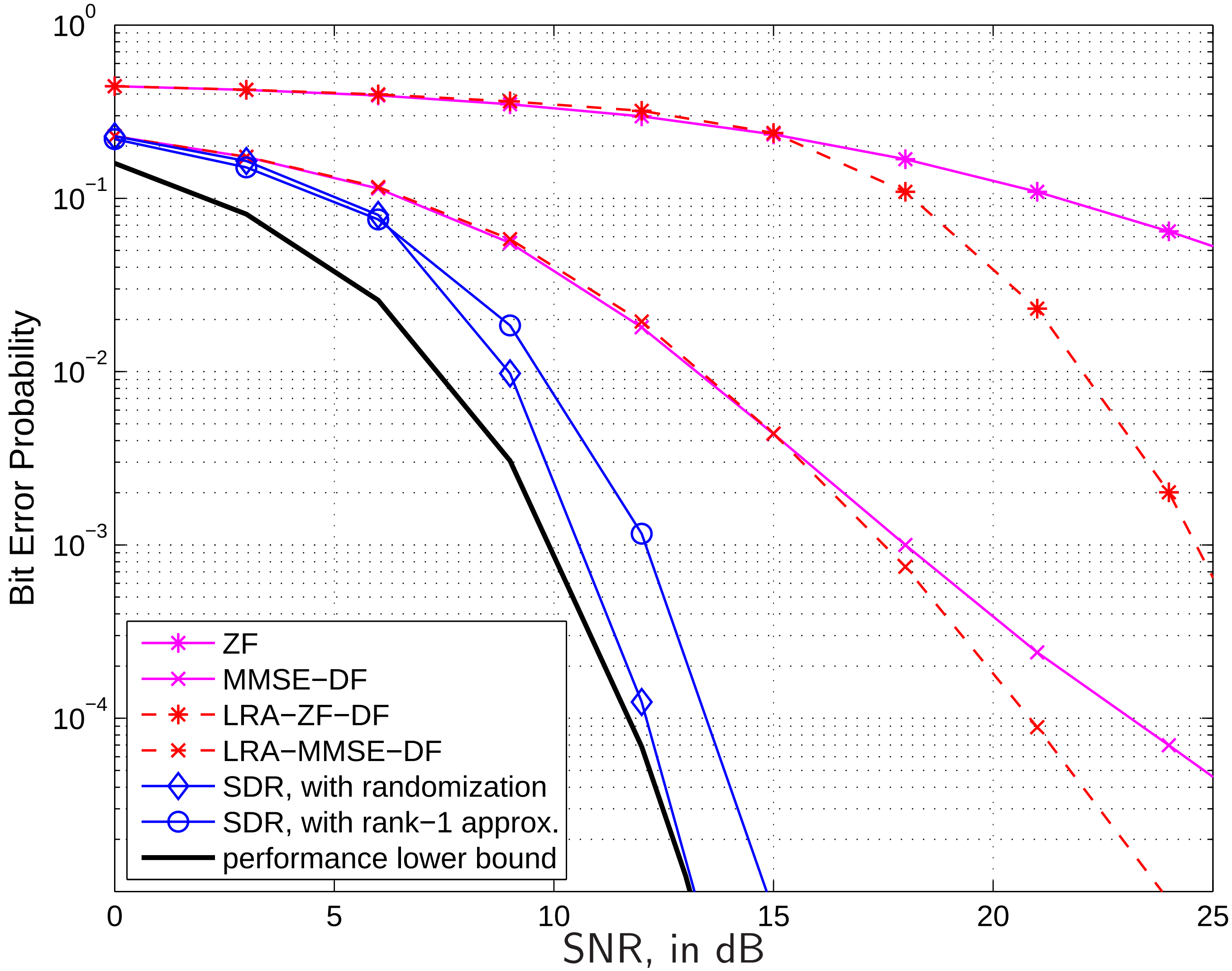
- The received signal model can be converted to a real form

$$\underbrace{\begin{bmatrix} \mathrm{Re}\{\boldsymbol{y}_C\} \\ \mathrm{Im}\{\boldsymbol{y}_C\} \end{bmatrix}}_{\boldsymbol{y}} = \underbrace{\begin{bmatrix} \mathrm{Re}\{\boldsymbol{H}_C\} & -\mathrm{Im}\{\boldsymbol{H}_C\} \\ \mathrm{Im}\{\boldsymbol{H}_C\} & \mathrm{Re}\{\boldsymbol{H}_C\} \end{bmatrix}}_{\boldsymbol{H}} \underbrace{\begin{bmatrix} \mathrm{Re}\{\boldsymbol{s}_C\} \\ \mathrm{Im}\{\boldsymbol{s}_C\} \end{bmatrix}}_{\boldsymbol{s} \in \{\pm 1\}^{2M_t}} + \underbrace{\begin{bmatrix} \mathrm{Re}\{\boldsymbol{v}_C\} \\ \mathrm{Im}\{\boldsymbol{v}_C\} \end{bmatrix}}_{\boldsymbol{v}},$$

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

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

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.

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

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} \trianglerighteq_i b_i, \quad i=1,\ldots,m. \end{aligned} \tag{QCQP}$$

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

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

  where we manipulate the statistics of $\boldsymbol{\xi}$ so that the objective function is minimized & constraints are satisfied *in expectation*.

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

$$\begin{aligned} \min \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{X}\succeq\mathbf{0}, \quad \mathrm{Tr}(\boldsymbol{A}_i\boldsymbol{X}) \trianglerighteq_i b_i, \quad i=1,\ldots,m. \end{aligned} \tag{SDR}$$

- The stochastic QCQP interpretation of SDR

$$\begin{aligned} \min_{\boldsymbol{X} \in \mathbb{S}^n} \quad & \mathrm{E}_{\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{X})} \{ \boldsymbol{\xi}^T \boldsymbol{C} \boldsymbol{\xi} \} \\ \text{s.t.} \quad & \mathrm{E}_{\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{X})} \{ \boldsymbol{\xi}^T \boldsymbol{A}_i \boldsymbol{\xi} \} \trianglerighteq_i b_i, \quad i = 1, \ldots, m \end{aligned} \qquad \text{(E-QCQP)}$$

essentially sheds lights into a different way of approximating QCQP.

- What we could do is the following: generate a random vector $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{X}^\star)$ ($\boldsymbol{X}^\star$ is an SDR soln.), and modify $\boldsymbol{\xi}$ 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]**. (to be covered in Part II)

Example: Randomization in BQP or MIMO Detection

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

$$\begin{aligned} \min \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & x_i^2 = 1, \quad i = 1, \ldots, n. \end{aligned} \tag{BQP}$$

Box 1. ***Gaussian Randomization Procedure for BQP***

given an SDR solution $\boldsymbol{X}^\star$, and a number of randomizations L.

for $\ell = 1, \ldots, L$

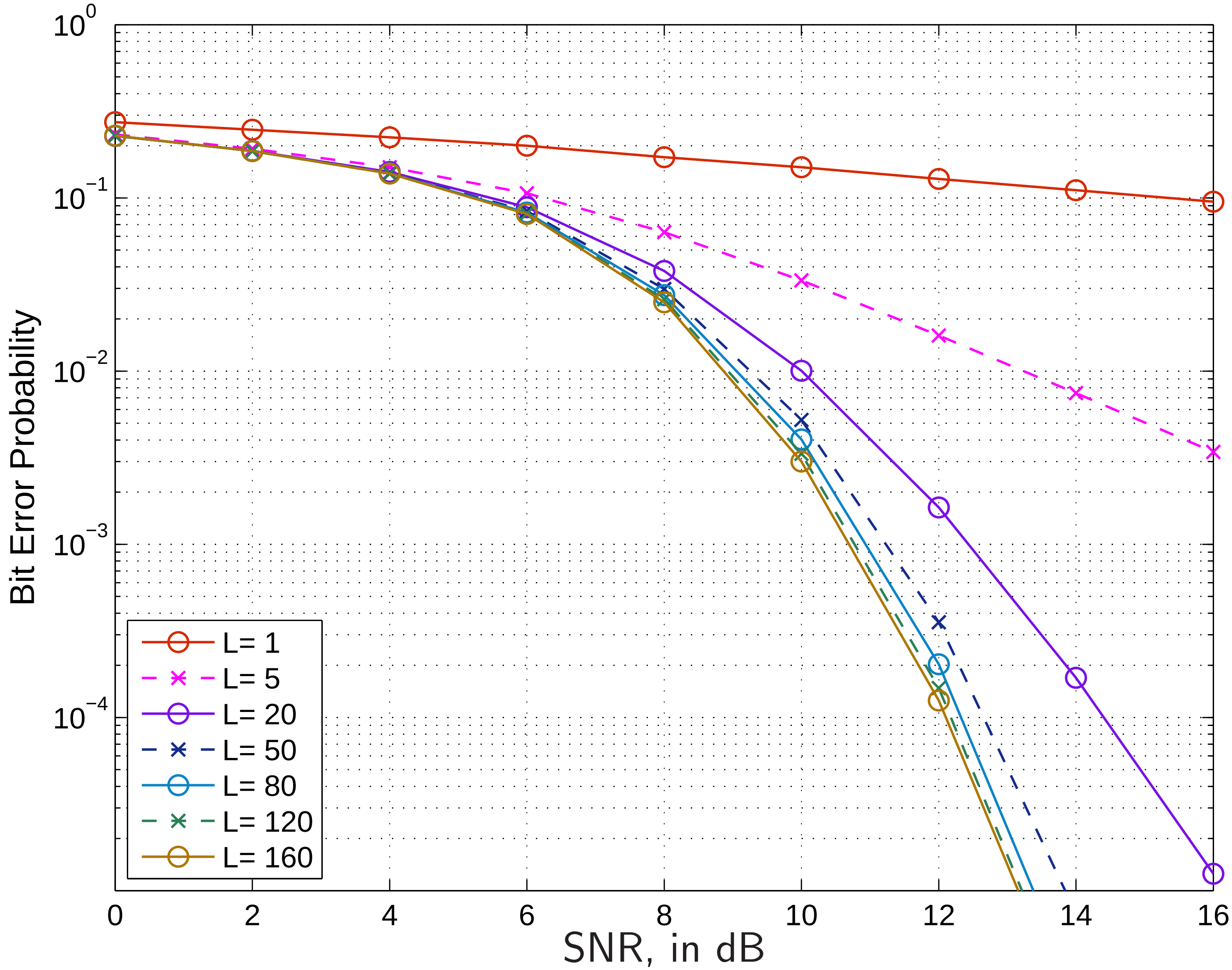
generate $\boldsymbol{\xi}_\ell \sim \mathcal{N}(\mathbf{0}, \boldsymbol{X}^\star)$, and construct a feasible point

$$\tilde{\boldsymbol{x}}_\ell = \operatorname{sgn}(\boldsymbol{\xi}_\ell).$$

end

determine $\ell^\star = \arg \min_{\ell=1,\ldots,L} \tilde{\boldsymbol{x}}_\ell^T \boldsymbol{C} \tilde{\boldsymbol{x}}_\ell$.

output $\hat{\boldsymbol{x}} = \tilde{\boldsymbol{x}}_{\ell^\star}$ as an approximate solution to (BQP).



Performance of various no. of randomizations in MIMO detection. $M_t = M_r = 40$.

Complex-valued QCQP and SDR

- Consider a general complex-valued QCQP

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{C}^n} \quad & \boldsymbol{x}^H \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^H \boldsymbol{A}_i \boldsymbol{x} \trianglerighteq_i b_i, \quad i = 1, \ldots, m, \end{aligned} \tag{1}$$

  where $\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathbb{H}^n$; $\mathbb{H}^n$ denotes the set of $n \times n$ Hermitian matrices.

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

$$\begin{aligned} \min_{\boldsymbol{X} \in \mathbb{H}^n} \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1, \ldots, m. \end{aligned}$$

  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 (will elaborate upon later).

Application: Multicast Transmit Beamforming

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

- The transmit signal:

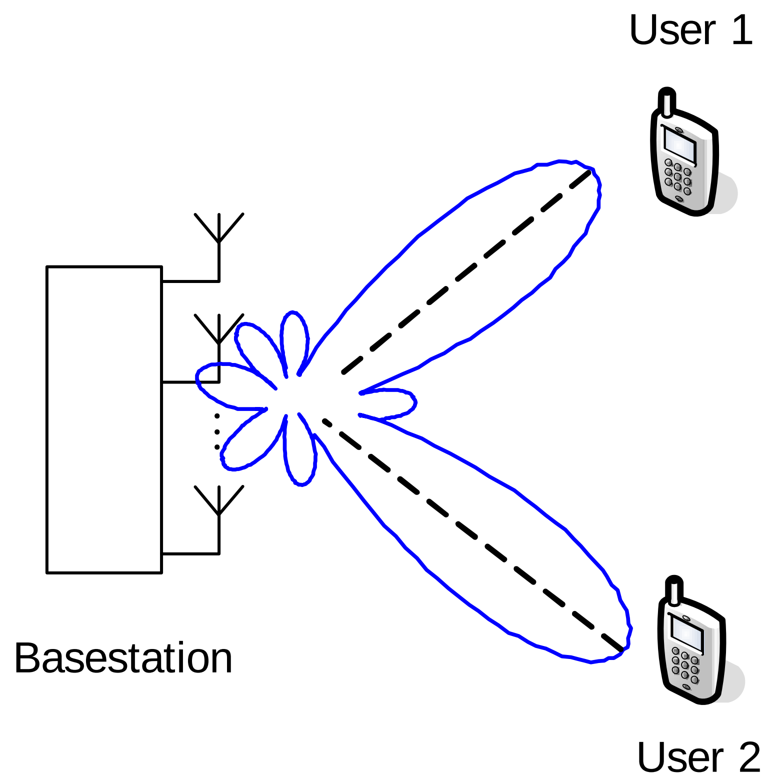
$$\boldsymbol{x}(t) = \boldsymbol{w}s(t),$$

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

- Received signal for user i:

$$y_i(t) = \boldsymbol{h}_i^H \boldsymbol{x}(t) + v_i(t),$$

where $\boldsymbol{h}_i \in \mathbb{C}^{N_t}$ is the channel of user i, & $v_i(t)$ is noise with variance σ_i^2.



- Consider a QoS-assured design:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{C}^{N_t}} \quad & \|\boldsymbol{w}\|^2 \\ \text{s.t.} \quad & \mathsf{SNR}_i \geq \gamma_i, \quad i=1,\ldots,K, \end{aligned}$$

where each γ_i is a prescribed SNR requirement for user i, and

$$\mathsf{SNR}_i = \mathrm{E}\{|\boldsymbol{h}_i^H \boldsymbol{w} s(t)|^2\}/\sigma_i^2 = \boldsymbol{w}^H \boldsymbol{R}_i \boldsymbol{w}/\sigma_i^2,$$

$$\boldsymbol{R}_i = \begin{cases} \boldsymbol{h}_i\boldsymbol{h}_i^H, & \boldsymbol{h}_i \text{ is available (instant CSIT)}, \\ \mathrm{E}\{\boldsymbol{h}_i\boldsymbol{h}_i^H\}, & \boldsymbol{h}_i \text{ is random with known 2nd order stat. (stat. CSIT)}. \end{cases}$$

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

$$\begin{aligned} \min \quad & \|\boldsymbol{w}\|^2 \\ \text{s.t.} \quad & \boldsymbol{w}^H \boldsymbol{A}_i \boldsymbol{w} \geq 1, \quad i=1,\ldots,K, \end{aligned}$$

where $\boldsymbol{A}_i = \boldsymbol{R}_i/\gamma_i\sigma_i^2$.

- This multicast problem is NP-hard in general, but can be approximated by SDR **[Sidiropoulos-Davidson-Luo'06]**.

A Randomization Example Relevant to Multicast Beamforming

Consider the problem

$$\begin{aligned} \min \quad & \boldsymbol{x}^H \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^H \boldsymbol{A}_i \boldsymbol{x} \geq 1, \quad i = 1, \ldots, m, \end{aligned} \qquad (\dagger)$$

where $\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \succeq \boldsymbol{0}$.

Box 2. ***Gaussian Randomization Procedure for (†)***

given an SDR solution $\boldsymbol{X}^\star$, and a number of randomizations L.

for $\ell = 1, \ldots, L$

generate $\boldsymbol{\xi}_\ell \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{X}^\star)$, and construct a feasible point

$$\tilde{\boldsymbol{x}}_\ell = \frac{\boldsymbol{\xi}_\ell}{\sqrt{\min_{i=1,\ldots,m} \boldsymbol{\xi}_\ell^H \boldsymbol{A}_i \boldsymbol{\xi}_\ell}}$$

end

determine $\ell^\star = \arg \min_{\ell=1,\ldots,L} \tilde{\boldsymbol{x}}_\ell^H \boldsymbol{C} \tilde{\boldsymbol{x}}_\ell$.

output $\hat{\boldsymbol{x}} = \tilde{\boldsymbol{x}}_{\ell^\star}$ as an approximate solution to (†).

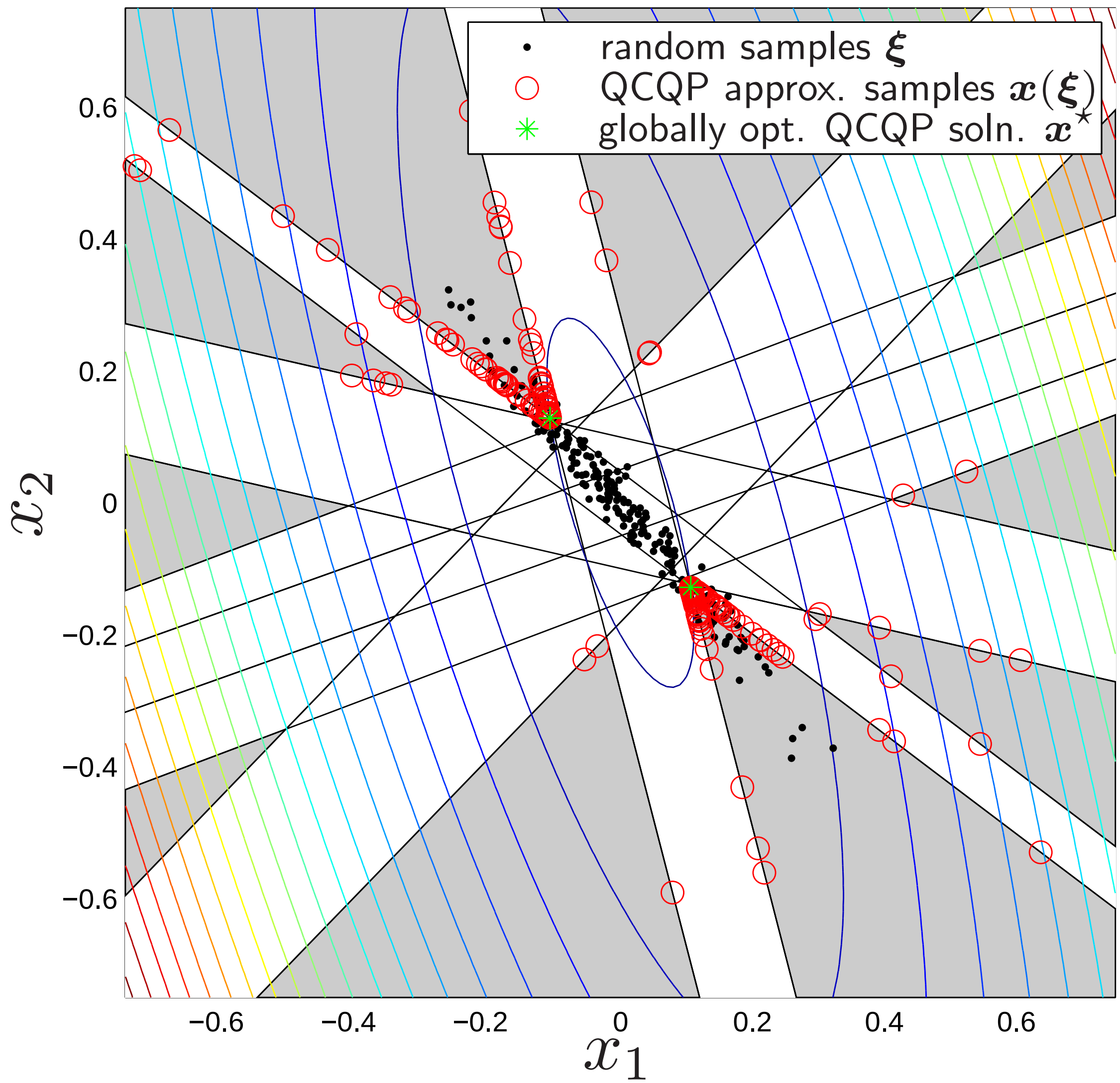
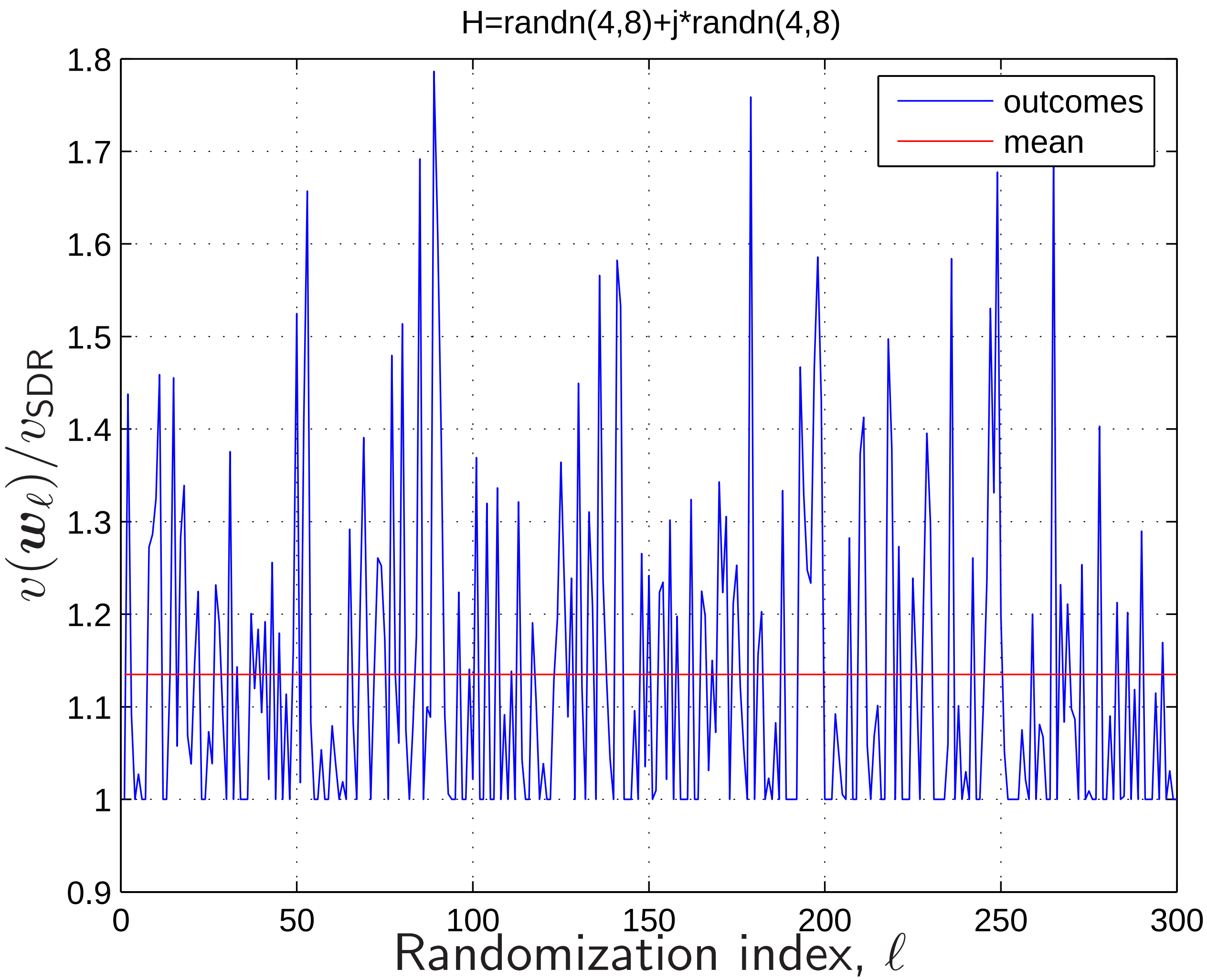


Illustration of randomizations in $\mathbb{R}^2$, for Problem (†). The gray area is the feasible set and colored lines the contour of the objective.



Approximation accuracy of Gaussian randomization in multicast beamforming. $N_t = 4$, $K = 8$, $v(\boldsymbol{w}) = \|\boldsymbol{w}\|^2$ is the objective value, v_{SDR} is the optimal value of SDR. Note that for any feasible $\boldsymbol{w}$, $v(\boldsymbol{w})/v_{\mathsf{SDR}} \geq v_{\mathsf{QP}}/v_{\mathsf{SDR}}$ where v_{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:

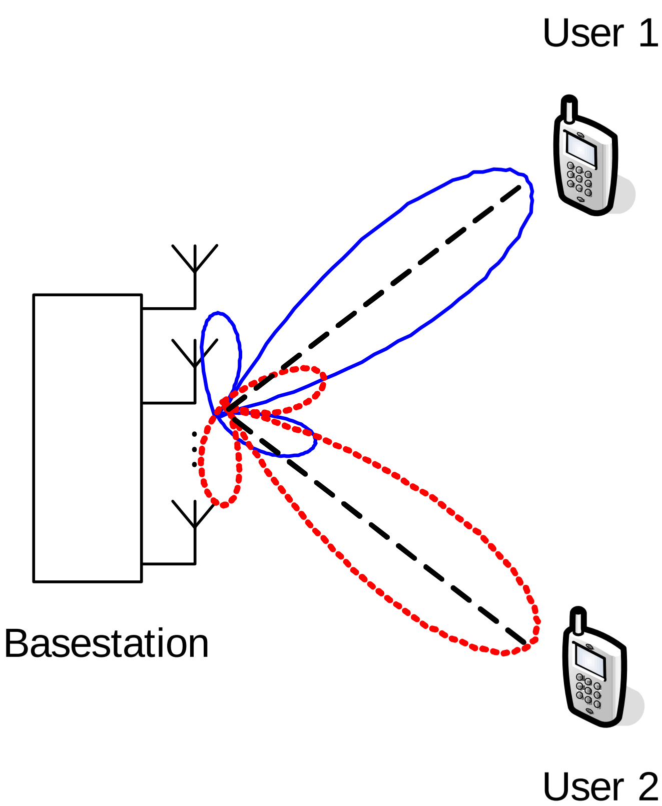
$$
\begin{aligned}
\min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k \in \mathbb{C}^n} \quad & \textstyle\sum_{i=1}^k \boldsymbol{x}_i^H \boldsymbol{C}_i \boldsymbol{x}_i \\
\text{s.t.} \quad & \textstyle\sum_{l=1}^k \boldsymbol{x}_l^H \boldsymbol{A}_{i,l} \boldsymbol{x}_l \trianglerighteq_i b_i, \quad i = 1, \ldots, m.
\end{aligned}
$$

- By writing $\boldsymbol{X}_i = \boldsymbol{x}_i \boldsymbol{x}_i^H$ for all i, and then "semidefinite-relaxing" them, we obtain an SDR

$$
\begin{aligned}
\min_{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k \in \mathbb{H}^n} \quad & \textstyle\sum_{i=1}^k \mathrm{Tr}(\boldsymbol{C}_i \boldsymbol{X}_i) \\
\text{s.t.} \quad & \textstyle\sum_{l=1}^k \mathrm{Tr}(\boldsymbol{A}_{i,l} \boldsymbol{X}_i) \trianglerighteq_i b_i, \quad i = 1, \ldots, m, \\
& \boldsymbol{X}_1 \succeq \boldsymbol{0}, \ldots, \boldsymbol{X}_k \succeq \boldsymbol{0}.
\end{aligned}
$$

Application: Unicast Transmit Downlink Beamforming

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



- Transmit signal:

$$\boldsymbol{x}(t) = \sum_{i=1}^{K} \boldsymbol{w}_i s_i(t),$$

where $s_i(t) \in \mathbb{C}$ is the data stream for user i, & $\boldsymbol{w}_i \in \mathbb{C}^{N_t}$ its tx. beamvector.

- Received signal of user i:

$$\begin{aligned} y_i(t) &= \boldsymbol{h}_i^H \boldsymbol{x}(t) + v_i(t) \\ &= \boldsymbol{h}_i^H \boldsymbol{w}_i s_i(t) + \underbrace{\sum_{l \neq i} \boldsymbol{h}_i^H \boldsymbol{w}_l s_l(t)}_{\text{interference}} + v_i(t). \end{aligned}$$

- The signal-to-interference-and-noise ratio (SINR) of user i:

$$\mathsf{SINR}_i = \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2},$$

where $\boldsymbol{R}_i = \boldsymbol{h}_i \boldsymbol{h}_i^H$ for instant. CSIT, and $\boldsymbol{R}_i = \mathrm{E}\{\boldsymbol{h}_i \boldsymbol{h}_i^H\}$ for stat. CSIT.

- Consider the QoS-assured design:

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

and its SDR

$$\begin{aligned} \min_{\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K \in \mathbb{H}^{N_t}} \quad & \textstyle\sum_{i=1}^K \mathrm{Tr}(\boldsymbol{W}_i) \\ \text{s.t.} \quad & \mathrm{Tr}(\boldsymbol{R}_i \boldsymbol{W}_i) \geq \gamma_i (\textstyle\sum_{l \neq i} \mathrm{Tr}(\boldsymbol{R}_i \boldsymbol{W}_l) + \sigma_i^2), \ i = 1, \ldots, K, \\ & \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \succeq \mathbf{0}. \end{aligned} \qquad (\ddagger)$$

- ($\ddagger$) is shown to have a rank-one solution for $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_K \succeq \mathbf{0}$, via uplink-downlink duality **[Bengtsson-Ottersten'01]**; SDR is optimal to ($\dagger$), so to speak!

- Part II will introduce an "easy" way to identify rank-one SDR instances.

Additional Remarks about the Transmit Beamforming Application

- 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.

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).

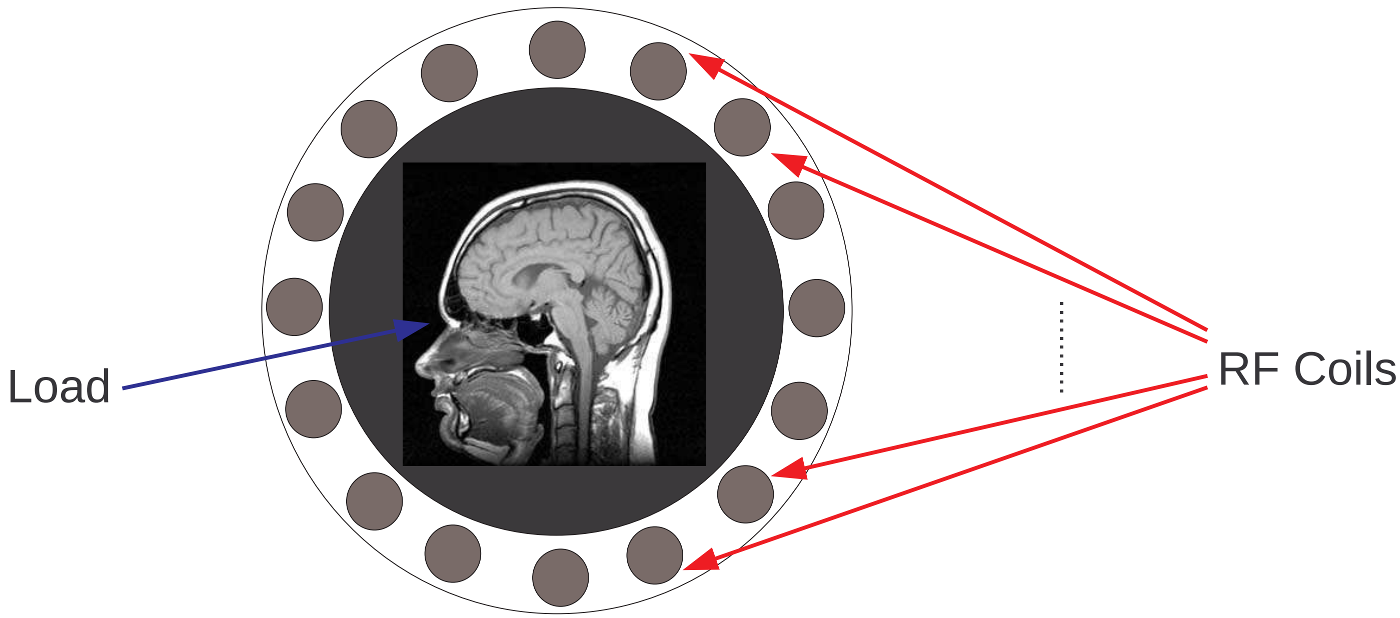
- So should we compare SDR and NPM?

- 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.

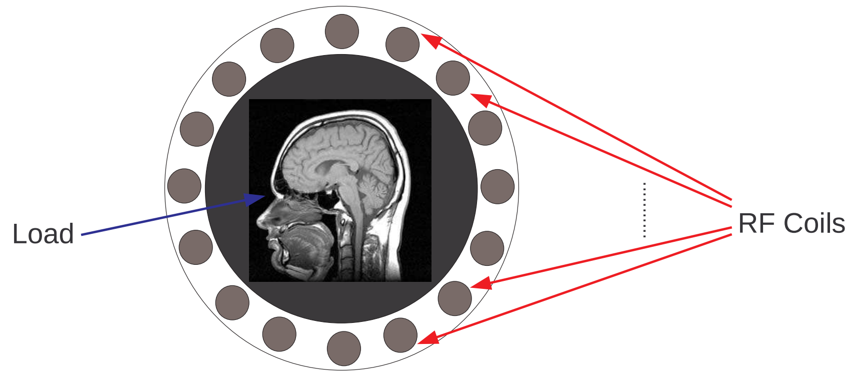
Application: Transmit B_1 Shim in MRI

- In MRI, a transmit RF coil array is used to generate a B_1 field.



- An undesirable effect is that the B_1 field exhibits strong inhomogeneity (spatial non-uniformity) across the load, due to complex interactions between the magnetic field and the loaded tissues.

- The goal is to design the transmit amplitudes and phases of the RF coils such that the resultant B_1 map is as uniform as possible.



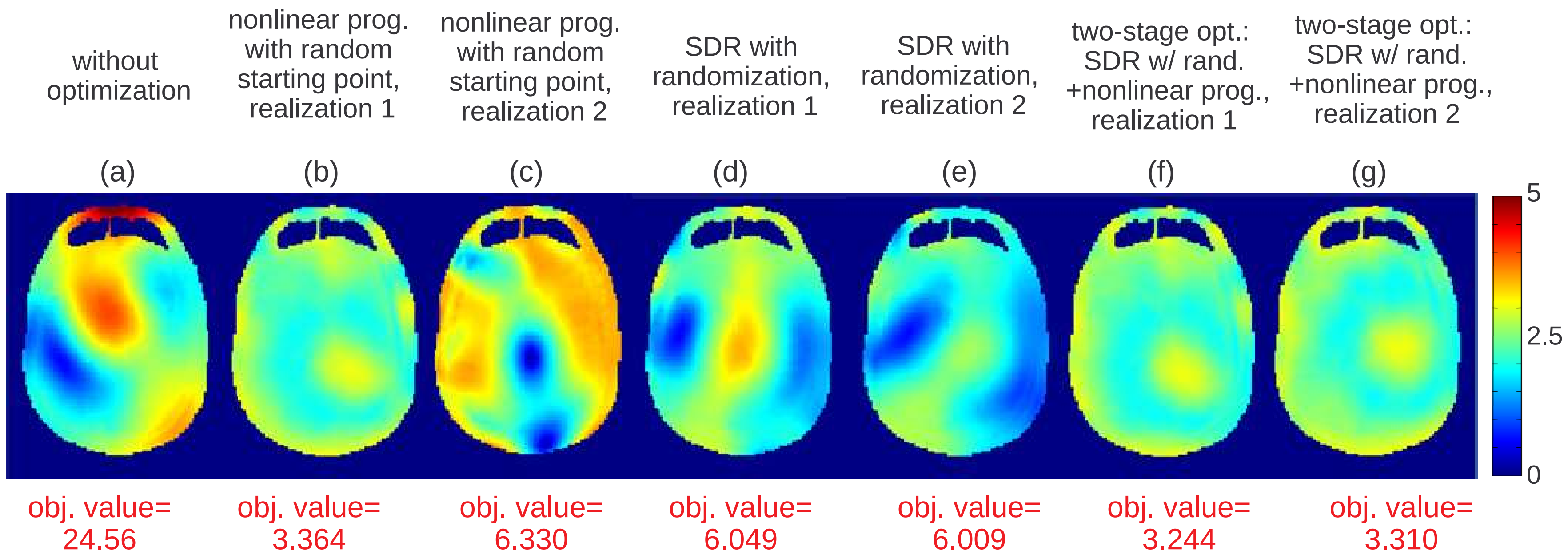
- Let $\boldsymbol{a}_i \in \mathbb{C}^n$, $i = 1, \ldots, m$, be the field response from the array to the ith pixel (MISO); i.e., the ith pixel receives a B_1 field of magnitude $|\boldsymbol{a}_i^T \boldsymbol{x}|$.

- The design problem is to minimize the worst-case field magnitude difference

$$\begin{array}{ll} \min_{\boldsymbol{x} \in \mathbb{C}^n} & \max_{i=1,\ldots,m} \left| |\boldsymbol{a}_i^T \boldsymbol{x}|^2 - b^2 \right| \\ \text{s.t.} & \boldsymbol{x}^H \boldsymbol{G} \boldsymbol{x} \leq \rho. \end{array}$$

Here, $\boldsymbol{x} \in \mathbb{C}^n$ is the transmit vector of the RF coil array, m is the total no. of pixels, $b > 0$ is the desired pixel value (uniform over all pixels), $\boldsymbol{x}^H \boldsymbol{G} \boldsymbol{x}$ is an average specific absorption rate (SAR), and ρ is a pre-specified SAR limit.

- It can be approximated by SDR **[Chang-Luo-Wu *et al.*'08]**.



B_1 maps of various optimization methods. You can see that the two-stage, SDR+NPM method shows better solution fidelity.

Part II: Theory

Provable Approximation Accuracies: Motivation

- So far we have introduced several procedures for generating an approximate QCQP solution from an SDR solution.

- A natural question arises: How good are these procedures?
  - Of course, their performance can be observed empirically. However, can we prove something about their approximation accuracy?
  - Such theoretical results can provide strong justification for the use of SDR in various problem settings.

- To measure the performance of a particular procedure, one intuitive approach is to quantify the gap between the objective value of the QCQP solution generated by the procedure and the optimal value of the QCQP.

Provable Approximation Accuracies: Setup

- Let $v(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x}$, and denote the optimal values of (QCQP) and (SDR) by

$$\begin{aligned} v_{\mathsf{QP}} = \min \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \trianglerighteq_i b_i, \quad i = 1, \ldots, m; \\ v_{\mathsf{SDR}} = \min \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1, \ldots, m. \end{aligned}$$

  Moreover, let $\hat{\boldsymbol{x}}$ be an approximate solution to (QCQP), obtained using one of the solution generation procedures (e.g., randomization). Note that

$$v(\hat{\boldsymbol{x}}) \geq v_{\mathsf{QP}}.$$

- We are interested to know if there exists a finite number $\gamma \geq 1$ (called the approximation ratio) such that

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

  either in expectation, or with high probability, or almost surely (since $\hat{\boldsymbol{x}}$ can be random). In general, the smaller γ, the better the solution generation procedure.

Provable Approximation Accuracies: Remarks

- In the definition of approximation ratio, we are implicitly assuming that $v_{\mathsf{QP}}, v_{\mathsf{SDR}} > 0$.
  - The notion of approximation ratio can be defined for problems where $v_{\mathsf{QP}} \leq 0$. However, we shall not go there in this tutorial.
- Given a solution generation procedure, we are usually interested in its performance on arbitrary instances of (QCQP). Thus, the approximation ratio γ should not depend on the problem data $\{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m, \boldsymbol{b}, \boldsymbol{C}\}$. However, it could depend on the problem dimensions m, n.
- For quadratic maximization problems, the notion of approximation ratio can be defined similarly.
- The problem of proving approximation accuracies has been of great interest to optimization theorists, and it has enormous implications in practice.

The Seminal Approx. Accuracy Result by Goemans & Williamson

- Consider

$$\begin{aligned} v_{\mathsf{QP}} = \max_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & x_i^2 = 1, \quad i = 1, \ldots, n \end{aligned}$$

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

- In **[Goemans-Williamson'95]**, it was shown that if the randomization procedure in Box 1 is used, then

$$\gamma v_{\mathsf{QP}} \leq \mathrm{E}\{v(\hat{\boldsymbol{x}})\} \leq v_{\mathsf{QP}},$$

  where $\gamma \approx 0.87856$.

- In particular, the approximation ratio is independent of the problem dimension n. In the context of MAXCUT, this means that the approximation accuracy is independent of the number of vertices in the graph.

Complex k-ary Quadratic Maximization

- Consider the problem

$$\begin{aligned} v_{\mathsf{QP}} = \max_{\boldsymbol{x}\in\mathbb{C}^n} \quad & \boldsymbol{x}^H \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & x_i \in \{1, \omega, \ldots, \omega^{k-1}\}, \quad i = 1, \ldots, n, \end{aligned} \tag{CQP-k}$$

  where $\boldsymbol{C} \succeq \boldsymbol{0}$ and $\omega = \exp(j2\pi/k)$ is the kth root of unity, for some given integer $k \geq 2$.

  – This is a generalization of the problem considered by Goemans and Williamson.

- Since $|x_i|^2 = 1$ for all i, (CQP-k) can be handled by SDR. Specifically,

$$\begin{aligned} v_{\mathsf{SDR}} = \max_{\boldsymbol{X}\in\mathbb{H}^n} \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{X} \succeq \boldsymbol{0}, \quad X_{ii} = 1, \quad i = 1, \ldots, n. \end{aligned}$$

Randomization Procedure for Complex k-ary Quad. Max.

- Again, a Gaussian randomization procedure can be used to generate a feasible solution to (CQP-k) from an SDR solution.

Box 3. $\boldsymbol{\textit{Gaussian Randomization Procedure for CQP-k}}$

given an SDR solution $\boldsymbol{X}^\star$, and a number of randomizations L.

for $\ell = 1, \ldots, L$

generate $\boldsymbol{\xi}_\ell \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{X}^\star)$, and construct the feasible point $\tilde{\boldsymbol{x}}_\ell \in \mathbb{C}^n$, where $[\tilde{\boldsymbol{x}}_\ell]_i = f([\boldsymbol{\xi}_\ell]_i)$ and

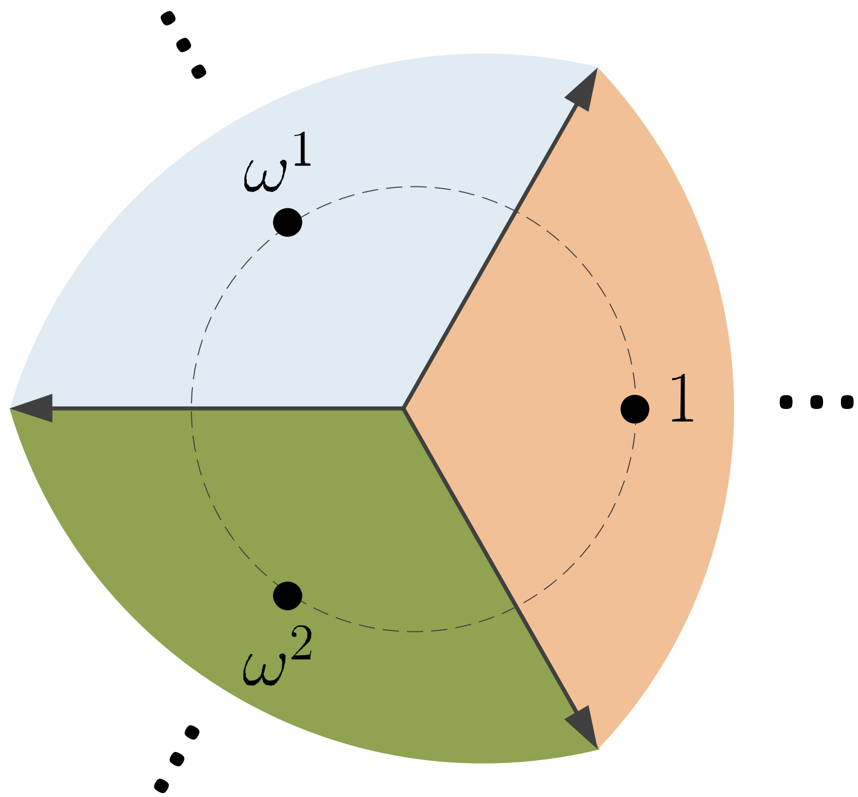
$$f(z) = \begin{cases} 1, & \arg(z) \in [-\pi/k, \pi/k), \\ \omega, & \arg(z) \in [\pi/k, 3\pi/k), \\ \vdots & \vdots \\ \omega^{k-1}, & \arg(z) \in [(2k-3)\pi/k, (2k-1)\pi/k). \end{cases}$$

end

determine $\ell^\star = \arg \max_{\ell=1,\ldots,L} \tilde{\boldsymbol{x}}_\ell^H \boldsymbol{C} \tilde{\boldsymbol{x}}_\ell$.

output $\hat{\boldsymbol{x}} = \tilde{\boldsymbol{x}}_{\ell^\star}$ as the approximate solution to (CQP-k).

Pictorial Illustration of the Randomization Procedure, for $k = 3$



Approx. Accuracy Result for Complex k-ary Quad. Max.

- In **[So-Zhang-Ye'07]**, it is shown that if the randomization procedure in Box 3 is used, then

$$\gamma v_{\mathsf{QP}} \leq \mathrm{E}\{\hat{\boldsymbol{x}}^H \boldsymbol{C} \hat{\boldsymbol{x}}\} \leq v_{\mathsf{QP}},$$

  where $\gamma = \dfrac{(k \sin(\pi/k))^2}{4\pi}$.

- If we take $k = \infty$, then the k-ary constraints in (CQP-k) become

$$|x_i| = 1, \quad i = 1, \ldots, n. \qquad (\dagger)$$

  In **[So-Zhang-Ye'07]** it is shown that by letting the function f in Box 3 to be

$$f(z) = \begin{cases} z/|z|, & |z| > 0, \\ 0, & |z| = 0, \end{cases}$$

  the randomization procedure would yield $\gamma = \pi/4$ for the unit-modulus constraints ($\dagger$). It is interesting (and comforting) to note that

$$\lim_{k \to \infty} \frac{(k \sin(\pi/k))^2}{4\pi} = \frac{\pi}{4}.$$

Applications of Complex k-ary Quadratic Maximization

- (CQP-k) has many applications in signal processing, e.g.:
  - blind orthogonal space-time block code detection **[Zhang-Ma'09]**
  - radar code waveform design **[De Maio *et al.*'09]**

problem	approx. accuracy γ; see (21)-(22) for def.	references
Boolean QP $$\begin{array}{ll}\max_{\boldsymbol{x}\in\mathbb{R}^n} & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}\\ \text{s.t.} & x_i^2=1,\ i=1,\ldots,n\end{array}$$	$$\gamma=\begin{cases}0.87856, & \boldsymbol{C}\succeq\boldsymbol{0},\ C_{ij}\leq 0\ \forall i\neq j\\ 2/\pi\simeq 0.63661, & \boldsymbol{C}\succeq\boldsymbol{0}\\ 1\text{ (opt.)}, & C_{ij}\geq 0, \forall i\neq j\end{cases}$$	Goemans-Williamson [2], Nesterov [3], Zhang [6]. Relevant applications: [24]–[26]
Complex k-ary QP $$\begin{array}{ll}\max_{\boldsymbol{x}\in\mathbb{C}^n} & \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x}\\ \text{s.t.} & x_i\in\{1,\omega,\ldots,\omega^{k-1}\},\\ & i=1,\ldots,n\end{array}$$ where $\omega=e^{j2\pi/k}$, and $k>1$ is an integer.	For $\boldsymbol{C}\succeq\boldsymbol{0}$, $$\gamma=\frac{(k\sin(\pi/k))^2}{4\pi}.$$ e.g., $\gamma=0.7458$ for $k=8$, $\gamma=0.7754$ for $k=16$.	Zhang-Huang [7], So-Zhang-Ye [8]. Relevant applications: [27], [37]
Complex constant-modulus QP $$\begin{array}{ll}\max_{\boldsymbol{x}\in\mathbb{C}^n} & \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x}\\ \text{s.t.} & \lvert x_i\rvert^2=1,\ i=1,\ldots,n\end{array}$$	For $\boldsymbol{C}\succeq\boldsymbol{0}$, $$\gamma=\pi/4=0.7854.$$ Remark: coincide with complex k-ary QP as $k\to\infty$.	Zhang-Huang [7], So-Zhang-Ye [8].
$$\begin{array}{ll}\max_{\boldsymbol{x}\in\mathbb{C}^n} & \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x}\\ \text{s.t.} & (\lvert x_1\rvert^2,\ldots,\lvert x_n\rvert^2)\in\mathcal{F}\end{array}$$ where $\mathcal{F}\subset\mathbb{R}^n$ is a closed convex set.	The same approx. ratio as in complex constant-modulus QP; i.e., $\gamma=\pi/4$ for $\boldsymbol{C}\succeq\boldsymbol{0}$. If the problem is reduced to the real-valued case, then the approx. ratio results are the same as that in Boolean QP.	Ye [4], Zhang [6].
$$\begin{array}{ll}\max_{\boldsymbol{x}\in\mathbb{R}^n} & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x}\\ \text{s.t.} & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x}\leq 1,\ i=1,\ldots,m\end{array}$$ where $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_m\succeq\boldsymbol{0}$.	For any $\boldsymbol{C}\in\mathbb{S}^n$, $$\gamma=\frac{1}{2\ln(2m\mu)}$$ where $\mu=\min\{m,\max_i\operatorname{rank}(\boldsymbol{A}_i)\}$.	Nemirovski-Roos-Terlaky [5]. Extensions: Ye [72], Luo-Sidiropoulos-Tseng-Zhang [9] and So-Ye-Zhang [71].

Known approximation accuracies for quadratic maximization problems. The reference numbers refer to those in our Signal Processing Magazine article.

Approx. Accuracy Result for Quadratic Minimization

- Consider now the problem

$$
\begin{aligned}
v_{\mathsf{QP}} = \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 1, \quad i = 1, \ldots, m
\end{aligned}
\tag{$\dagger$}
$$

  for $\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \succeq \mathbf{0}$, which arises in the study of 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),

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

  where $\gamma = 27m^2/\pi$.

  – For the complex version of (†), one has a better approximation ratio: $\gamma = 8m$.

- Notice that this ratio accommodates the worst possible problem instance $\{\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m\}$. In practice, the approximation accuracies are usually much better— a phenomenon that deserves further investigation.

Interpretation in Multicast Transmit Beamforming

- Recall that in the context of multicast transmit beamforming, we encounter the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{w} \in \mathbb{C}^N} \quad & \|\boldsymbol{w}\|^2 \\ \text{s.t.} \quad & \mathsf{SNR}_i = \frac{1}{\gamma_i \sigma_i^2} \boldsymbol{w}^H \boldsymbol{R}_i \boldsymbol{w} \geq 1, \quad i = 1, \ldots, K, \end{aligned}$$

  which is an instance of the quadratic minimization problem considered in **[Luo-Sidiropoulos-Tseng-Zhang'07]**.

- The aforementioned approximation accuracy result thus says that SDR together with the randomization procedure can produce a transmit beamforming vector that satisfies all the prescribed SNR requirements and whose power is at most $8m$ times the optimal.

- Again, this is just a worst-case guarantee. In practice, the performance is usually much better.

problem	approx. accuracy γ; see (18)-(19) for def.	references
$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{C}^n} \quad & \boldsymbol{x}^H \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^H \boldsymbol{A}_i \boldsymbol{x} \geq 1,\ i=1,\ldots,m \end{aligned}$ where $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_m \succeq \boldsymbol{0}$.	$\gamma = 8m.$ If the problem is reduced to the real-valued case, then $\gamma = \frac{27m^2}{\pi}.$	Luo-Sidiropoulos-Tseng-Zhang [9]; see also So-Ye-Zhang [71]. Relevant applications: [29]
MIMO Detection $\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|_2^2 \\ \text{s.t.} \quad & x_i^2 = 1,\ i=1,\ldots,n \end{aligned}$ where $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s}+\boldsymbol{v}$; $\boldsymbol{H} \in \mathbb{C}^{n\times n}$ has i.i.d. standard complex Gaussian entries; $s_i^2 = 1$ for $i=1,\ldots,n$; and $\boldsymbol{v} \in \mathbb{C}^n$ has i.i.d. complex mean zero Gaussian entries with variance σ^2.	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.	Kisialiou-Luo [67], So [69]. Extensions: So [68], [69]. Related: Jaldén-Ottersten [66]. Relevant applications: [17]–[20], [22], [23]

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, even equal to one.

Shapiro-Barvinok-Pataki (SBP) Result

- Consider the real-valued SDP (or SDR)

$$\begin{aligned} \min_{\boldsymbol{X} \in \mathbb{S}^n} \quad & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} \quad & \boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1, \ldots, m. \end{aligned} \qquad \text{(SDR)}$$

SBP Result [Pataki'98]: There exists an optimal solution $\boldsymbol{X}^\star$ such that

$$\frac{\mathrm{rank}(\boldsymbol{X}^\star)(\mathrm{rank}(\boldsymbol{X}^\star)+1)}{2} \leq m.$$

- In particular, SBP result implies that for $m \leq 2$, a rank-1 $\boldsymbol{X}^\star$ 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

$$\begin{array}{ll} \min\limits_{\boldsymbol{X} \in \mathbb{H}^n} & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) \\ \text{s.t.} & \boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i, \quad i = 1, \ldots, m. \end{array}$$

- In this case, the SBP result can be generalized to **[Huang-Zhang'07]**

$$\mathrm{rank}(\boldsymbol{X}^\star)^2 \leq m,$$

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.

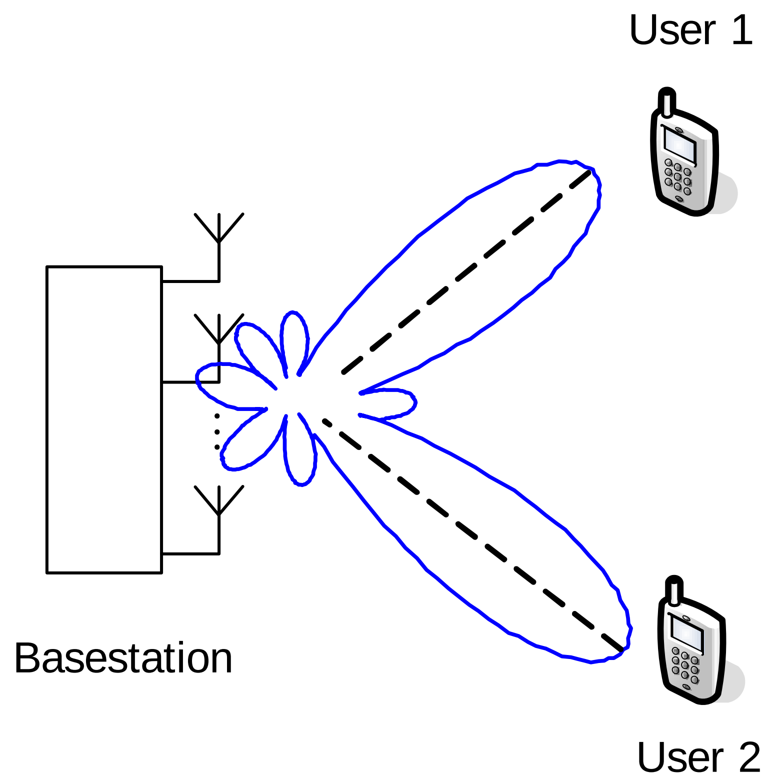
- A complex rank-1 decomposition algorithm for $m \leq 3$ is available **[Huang-Zhang'07]**.

Application Revisited: Multicast Beamforming

- Recall the multicast beamforming problem:

$$
\begin{aligned}
\min_{\boldsymbol{w}\in\mathbb{C}^{N_t}} \quad & \|\boldsymbol{w}\|^2 \\
\text{s.t.} \quad & \mathsf{SNR}_i = \frac{\boldsymbol{w}\boldsymbol{R}_i\boldsymbol{w}}{\sigma_i^2} \geq \gamma_i, \\
& i = 1, \ldots, K,
\end{aligned}
$$

K being the number of users.



- By the SBP result, SDR solves the multicast problem optimally for $K \leq 3$.

Further Extension of the Rank Reduction Result

- Recall the problem

$$\begin{aligned} \min_{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k \in \mathbb{H}^n} \quad & \textstyle\sum_{i=1}^k \mathrm{Tr}(\boldsymbol{C}_i \boldsymbol{X}_i) \\ \text{s.t.} \quad & \textstyle\sum_{l=1}^k \mathrm{Tr}(\boldsymbol{A}_{i,l} \boldsymbol{X}_i) \trianglerighteq_i b_i, \quad i = 1, \ldots, m, \\ & \boldsymbol{X}_1 \succeq \boldsymbol{0}, \ldots, \boldsymbol{X}_k \succeq \boldsymbol{0}, \end{aligned}$$

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

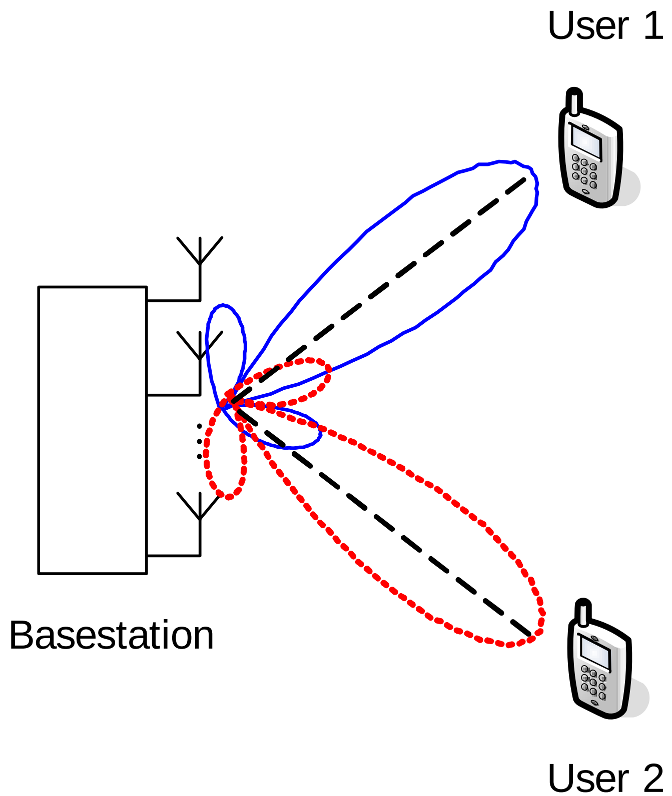
- A generalization of the SBP result **[Huang-Palomar'09]**:

$$\textstyle\sum_{i=1}^k \mathrm{rank}(\boldsymbol{X}_i^\star)^2 \leq m,$$

  and, as a subsequent result:

> Suppose that an SDR solution cannot have $\boldsymbol{X}_i^\star = \boldsymbol{0}$ for any i. Then SDR is tight for $m \leq k + 2$.

Application Revisited: Unicast Beamforming



- Recall the design problem

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

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

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

- And hey, it's still fine if you put two more quadratic constraints in ($\dagger$)!

Further Results in SDR Rank Reduction: Motivation

- As we have seen before, it is not possible in general to find a low-rank solution to a QCQP in polynomial time, unless the number of constraints is small.
  - In fact, most polynomial-time interior-point algorithms will return a solution that has the highest rank.

- In such situations, can we find a low-rank solution that has good approximation accuracy?
  - The randomization procedure introduced before can generate a rank-one solution with good approximation accuracy.
  - How about a rank-r solution, where $r \geq 2$ is fixed and small? This has applications in sensor network localization, which will be covered later.

Further Results in SDR Rank Reduction: Setup

- Consider the following rank-constrained semidefinite feasibility problem (RCSDF):

$$\begin{array}{ll} \text{find} & \boldsymbol{X} \succeq \boldsymbol{0}, \ \text{rank}(\boldsymbol{X}) \leq r \\ \text{s.t.} & \text{Tr}(\boldsymbol{A}_i \boldsymbol{X}) = b_i, \quad i = 1, \ldots, m, \end{array} \qquad \text{(RCSDF)}$$

  where $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \succeq \boldsymbol{0}$, $b_1, \ldots, b_m \geq 0$ and $r \geq 1$ are given.

- What is the relationship between (RCSDF) and the QCQPs we studied before?
  - For instance, consider the QCQP

$$\begin{array}{ll} \min & \boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x}^T \boldsymbol{A}_i \boldsymbol{x} \geq 1, \quad i = 1, \ldots, m, \end{array}$$

    where $\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \succeq \boldsymbol{0}$. Let $\boldsymbol{x}^\star$ be an optimal solution.
  - Then, $\boldsymbol{X}^\star = \boldsymbol{x}^\star \boldsymbol{x}^{\star T}$ is feasible for the following RCSDF problem:

$$\begin{array}{ll} \text{find} & \boldsymbol{X} \succeq \boldsymbol{0}, \ \text{rank}(\boldsymbol{X}) \leq 1 \\ \text{s.t.} & \text{Tr}(\boldsymbol{A}_i \boldsymbol{X}) = \text{Tr}(\boldsymbol{A}_i \boldsymbol{X}^\star), \quad i = 1, \ldots, m, \\ & \text{Tr}(\boldsymbol{C} \boldsymbol{X}) = \text{Tr}(\boldsymbol{C} \boldsymbol{X}^\star). \end{array}$$

Further Results in SDR Rank Reduction: Setup

- There are two difficulties in solving (RCSDF):
  - There is no known polynomial-time algorithm for solving (RCSDF) in general.
  - The problem may be infeasible.

- It is thus natural to study "easier" versions of (RCSDF). For instance, consider

$$\begin{array}{rl} \text{find} & \boldsymbol{X} \succeq \mathbf{0},\ \text{rank}(\boldsymbol{X}) \leq r \\ \text{s.t.} & \alpha \cdot b_i \leq \text{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \leq \beta \cdot b_i, \quad i = 1, \ldots, m, \end{array} \qquad \text{(RCSDF-R)}$$

  where α, β are some parameters.
  - Note that when $\alpha = \beta = 1$, we get back (RCSDF).

- Question: Under what conditions on α and β would (RCSDF-R) be feasible, and a feasible solution can be found efficiently?

SDR Rank Reduction: The So-Ye-Zhang (SYZ) Theorem

- Suppose there exists an $\boldsymbol{X}^\star \succeq \boldsymbol{0}$ such that $\mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}^\star) = b_i$ for $i = 1, \ldots, m$.

- **[So-Ye-Zhang'08]** One can find in randomized polynomial time a rank-r $\hat{\boldsymbol{X}}$ such that

$$\alpha \cdot b_i \le \mathrm{Tr}(\boldsymbol{A}_i \hat{\boldsymbol{X}}) \le \beta \cdot b_i \quad \text{for } i = 1, \ldots, m$$

  with high probability, where

$$\alpha = \begin{cases} \dfrac{1}{e(2m)^{2/r}}, & 1 \le r \le 4\ln(2m), \\ \max\left\{ \dfrac{1}{e(2m)^{2/r}}, 1 - \sqrt{\dfrac{4\ln(2m)}{r}} \right\}, & r > 4\ln(2m), \end{cases}$$

$$\beta = \begin{cases} 1 + \dfrac{12\ln(4mq)}{r}, & 1 \le r \le 12\ln(4mq), \\ 1 + \sqrt{\dfrac{12\ln(4mq)}{r}}, & r > 12\ln(4mq), \end{cases}$$

$$q = \min\{\sqrt{2m}, n\}.$$

- Some quick remarks:
  - The larger the r, the closer α, β are to 1. Intuitively, r dictates how much information about $\boldsymbol{X}^\star$ we can keep.
  - From the definition of q, it is clear that the bounds can be made independent of n, the dimension of the matrices.

Randomization Procedure for SDR Rank Reduction

- How to achieve the claimed bounds? (Surprise) Use Gaussian randomization!

> **Box 4.** ***Gaussian Randomization Procedure for Rank Reduction***
> **given** a solution $\boldsymbol{X}^\star$ that satisfies $\mathrm{Tr}(\boldsymbol{A}_i\boldsymbol{X}^\star) = b_i$ for all i and $\boldsymbol{X}^\star \succeq \mathbf{0}$, and an integer $r \geq 1$.
>
> for $\ell = 1, \ldots, r$
>
>   generate $\boldsymbol{\xi}_\ell \sim \mathcal{N}(\mathbf{0}, \boldsymbol{X}^\star)$;
>
> end
>
> **output** $\hat{\boldsymbol{X}} = \frac{1}{r}\sum_{\ell=1}^{r} \boldsymbol{\xi}_\ell \boldsymbol{\xi}_\ell^T$ as the candidate solution.

- Why this works? Intuitively,

$$\hat{\boldsymbol{X}} \succeq \mathbf{0}, \quad \mathrm{rank}(\hat{\boldsymbol{X}}) \leq r, \quad \mathrm{E}\{\hat{\boldsymbol{X}}\} = \boldsymbol{X}^\star.$$

- This extends our previous rank-1 Gaussian randomization procedures.

An Application of the SYZ Theorem

- Let $G = (V, E)$ be a graph (e.g., the communication graph of a sensor network).

- Assumption: Suppose that we are given a set of nonnegative weights $\{d_{ij} : (i,j) \in E\}$, and that there exist vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbb{R}^\ell$ for some $\ell \geq 1$ with

$$\|\boldsymbol{v}_i - \boldsymbol{v}_j\|^2 = d_{ij}^2, \quad (i,j) \in E.$$

- Question: Can we find vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ in a lower dimensional space, say $\mathbb{R}^2$, such that

$$\|\boldsymbol{u}_i - \boldsymbol{u}_j\|^2 = d_{ij}^2, \quad (i,j) \in E?$$

- In general, this is not possible. However, the SYZ theorem asserts that if we allow some distortion in the distances, then such vectors can be found.
  - The amount of distortion depends on how low the dimension of the space in which we need to embed the points $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$.

Finding the Low Dimensional Vectors

- The assumption implies that the following system is feasible:

$$X_{ii} - 2X_{ij} + X_{jj} = d_{ij}^2, \quad (i,j) \in E; \quad \boldsymbol{X} \succeq \boldsymbol{0}. \tag{$\dagger$}$$

- Indeed, let $\boldsymbol{X}^\star = [\ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\ ]^T [\ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\ ] \succeq \boldsymbol{0}$, so that $X_{ij}^\star = \boldsymbol{v}_i^T \boldsymbol{v}_j$. Then, we have

$$X_{ii}^\star - 2X_{ij}^\star + X_{jj}^\star = \boldsymbol{v}_i^T \boldsymbol{v}_i - 2\boldsymbol{v}_i^T \boldsymbol{v}_j + \boldsymbol{v}_j^T \boldsymbol{v}_j = d_{ij}^2, \quad (i,j) \in E.$$

  - A side observation: $\mathrm{rank}(\boldsymbol{X}^\star) \leq \ell$.

- Note that ($\dagger$) can be written as

$$\mathrm{Tr}(\boldsymbol{E}_{ij}\boldsymbol{X}) = d_{ij}^2, \quad (i,j) \in E; \quad \boldsymbol{X} \succeq \boldsymbol{0},$$

  where $\boldsymbol{E}_{ij} = (\boldsymbol{e}_i - \boldsymbol{e}_j)(\boldsymbol{e}_i - \boldsymbol{e}_j)^T$.

- Hence, our problem becomes

$$\begin{array}{rl} \text{find} & \boldsymbol{X} \succeq \mathbf{0}, \, \mathrm{rank}(\boldsymbol{X}) \leq 2 \\ \text{s.t.} & \mathrm{Tr}(\boldsymbol{E}_{ij}\boldsymbol{X}) = d_{ij}^2, \quad (i,j) \in E. \end{array}$$

- By the SYZ theorem, we can find an $\hat{\boldsymbol{X}} \succeq \mathbf{0}$ with $\mathrm{rank}(\hat{\boldsymbol{X}}) \leq 2$ and

$$\frac{1}{2e|E|} d_{ij}^2 \leq \mathrm{Tr}(\boldsymbol{E}_{ij}\hat{\boldsymbol{X}}) \leq (1 + 6\ln(4q|E|))\, d_{ij}^2, \quad (i,j) \in E, \qquad (\dagger)$$

  where $q = \min\{\sqrt{2|E|}, n\}$.

- Let $\hat{\boldsymbol{X}} = [\ \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\ ]^T [\ \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\ ]$ be the Cholesky factorization of $\hat{\boldsymbol{X}}$, where $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n \in \mathbb{R}^2$. Then ($\dagger$) says that

$$\Omega\left(\frac{1}{|E|}\right) \cdot d_{ij}^2 \leq \|\boldsymbol{u}_i - \boldsymbol{u}_j\|^2 \leq O\left(\ln(|E|)\right) \cdot d_{ij}^2, \quad (i,j) \in E.$$

Part III.A: MIMO Detection

Introduction

- MIMO detection is an important topic with a wide scope of applicability.
- The goal is to achieve good symbol error probability performance, preferably near-optimal, in a computationally efficient manner.
- Note that SDR is not the only efficient high-performance MIMO detection approach. The sphere decoding approach and the lattice reduction aided (LRA) approach are also powerful.
- Our focus:
  - computational or implementation aspects of SDR;
  - alternative interpretations of SDR; connections to other MIMO detectors;
  - SDR for various types of constellations (we went thro' $\{\pm 1\}$ so far);
  - benchmarking SDR and representative MIMO detectors, through extensive simulation results .

Problem Statement

- Consider a generic complex-valued $M_C \times N_C$ MIMO model

$$\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C,$$

where

$\boldsymbol{H}_C \in \mathbb{C}^{M_C \times N_C}$	the MIMO channel;
$\boldsymbol{s}_C \in \mathcal{S}^{N_C}$	the tx symbol vector, with $\mathcal{S} \subset \mathbb{C}$ being the constellation set;
$\boldsymbol{v}_C \in \mathbb{C}^{M_C}$	complex AWGN.

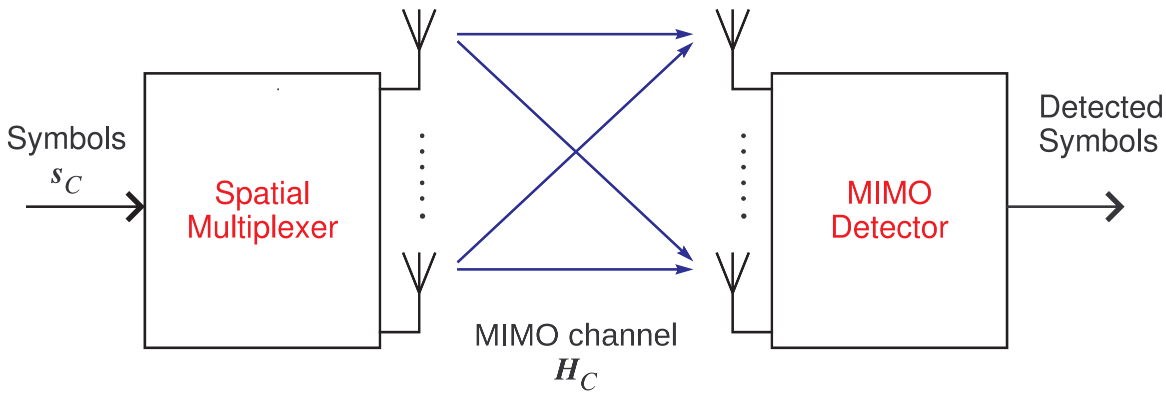
- We will focus on the ML detection problem

$$\hat{\boldsymbol{s}}_{C,\text{ML}} = \arg \min_{\boldsymbol{s}_C \in \mathcal{S}^{N_C}} \|\boldsymbol{y}_C - \boldsymbol{H}_C \boldsymbol{s}_C\|^2.$$

- Constellations:
  - QPSK: $\mathcal{S} = \{ \; s = a + jb \mid a, b \in \{\pm 1\} \; \}$
  - M-ary PSK (MPSK): $\mathcal{S} = \{ \; s = e^{j2\pi k/M} \mid k = 0, 1, \ldots, M-1 \; \}$
  - 4^q-ary QAM: $\mathcal{S} = \{ \; s = a + jb \mid a, b \in \{\pm 1, \pm 3, \ldots, \pm(2^{q-1} - 1)\} \; \}$

Scope of Applicability

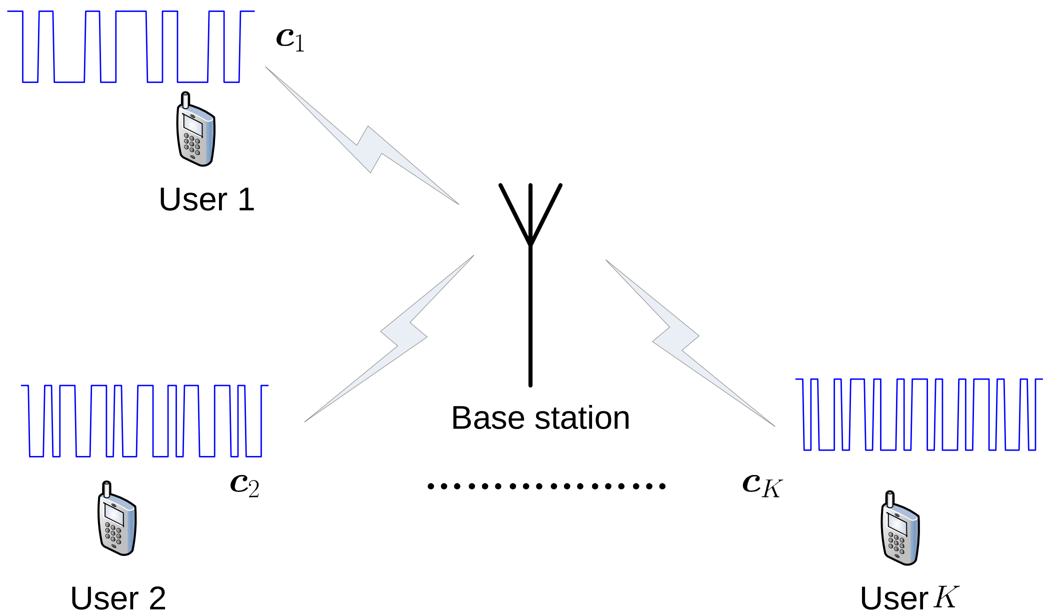
- The simple MIMO model $\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C$ is popularly used in the point-to-point spatial multiplexing scenario.



- Actually, this MIMO model is general enough to cover a wide variety of digital communication scenarios.

- As such, MIMO detection methods developed for the generic MIMO model can be universally applied to many different scenarios.

Example: CDMA Multiuser Detection



- Consider a multiuser CDMA scenario. rx signal model over one symbol interval:

$$\boldsymbol{y} = \sum_{i=1}^{K} \boldsymbol{c}_i \alpha_i s_i + \boldsymbol{v},$$

where $\boldsymbol{y} \in \mathbb{C}^N$ is the rx code vector; $\boldsymbol{c}_i \in \mathbb{C}^N$ spreading code sequence vector of user i; s_i tx symbol of user i; $\alpha_i \in \mathbb{C}$ rx amplitude/phase coefficient of user i.

- can be rewritten as $\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C$ (obviously).

Example: Space-Time Block Coding

- Consider a point-to-point space-time block code (STBC) scenario:

$$\boldsymbol{Y} = \boldsymbol{H}_C \boldsymbol{C}(\boldsymbol{s}_C) + \boldsymbol{V},$$

where $\boldsymbol{H}_C \in \mathbb{C}^{M_r \times M_t}$ the MIMO channel; $\boldsymbol{C}(\boldsymbol{s}_C) \in \mathbb{C}^{M_t \times T}$ is an STBC; $\boldsymbol{Y} \in \mathbb{C}^{M_r \times T}$ is the rx space-time code block, T being the time length.

- Assume a linear dispersion STBC:

$$\boldsymbol{C}(\boldsymbol{s}_C) = \sum_{l=1}^{L} \boldsymbol{A}_l \mathrm{Re}\{s_{C,l}\} + \boldsymbol{B}_l \mathrm{Im}\{s_{C,l}\}.$$

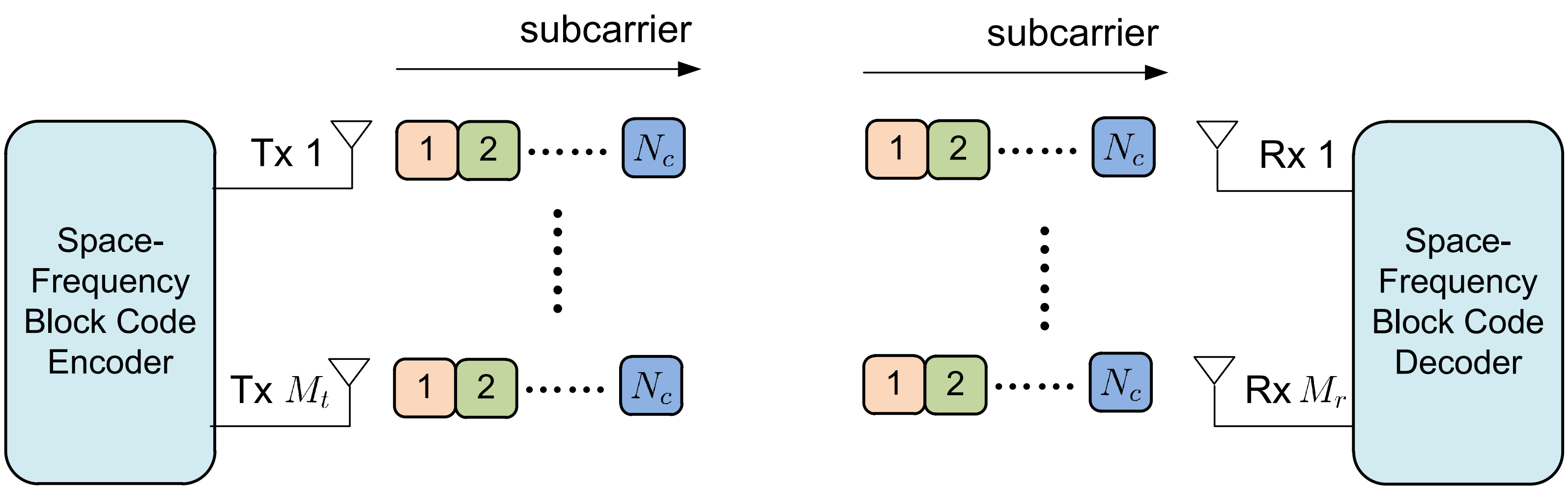
- The rx model can be converted to the generic MIMO form:

$$\mathrm{vec}(\boldsymbol{Y}) = \underbrace{(\boldsymbol{I} \otimes \boldsymbol{H}_C)\boldsymbol{\mathcal{X}}}_{\text{“another } \boldsymbol{H}_C\text{”}} \tilde{\boldsymbol{s}} + \mathrm{vec}(\boldsymbol{V}),$$

where $\boldsymbol{\mathcal{X}} = [\ \mathrm{vec}(\boldsymbol{A}_1), \ldots, \mathrm{vec}(\boldsymbol{A}_L), \mathrm{vec}(\boldsymbol{B}_1), \ldots, \mathrm{vec}(\boldsymbol{B}_L)\ ] \in \mathbb{C}^{M_t T \times 2L}$, $\tilde{\boldsymbol{s}} = [\ \mathrm{Re}\{\boldsymbol{s}_C\}^T, \mathrm{Im}\{\boldsymbol{s}_C\}^T\ ]^T \in \mathbb{R}^{2L}$.

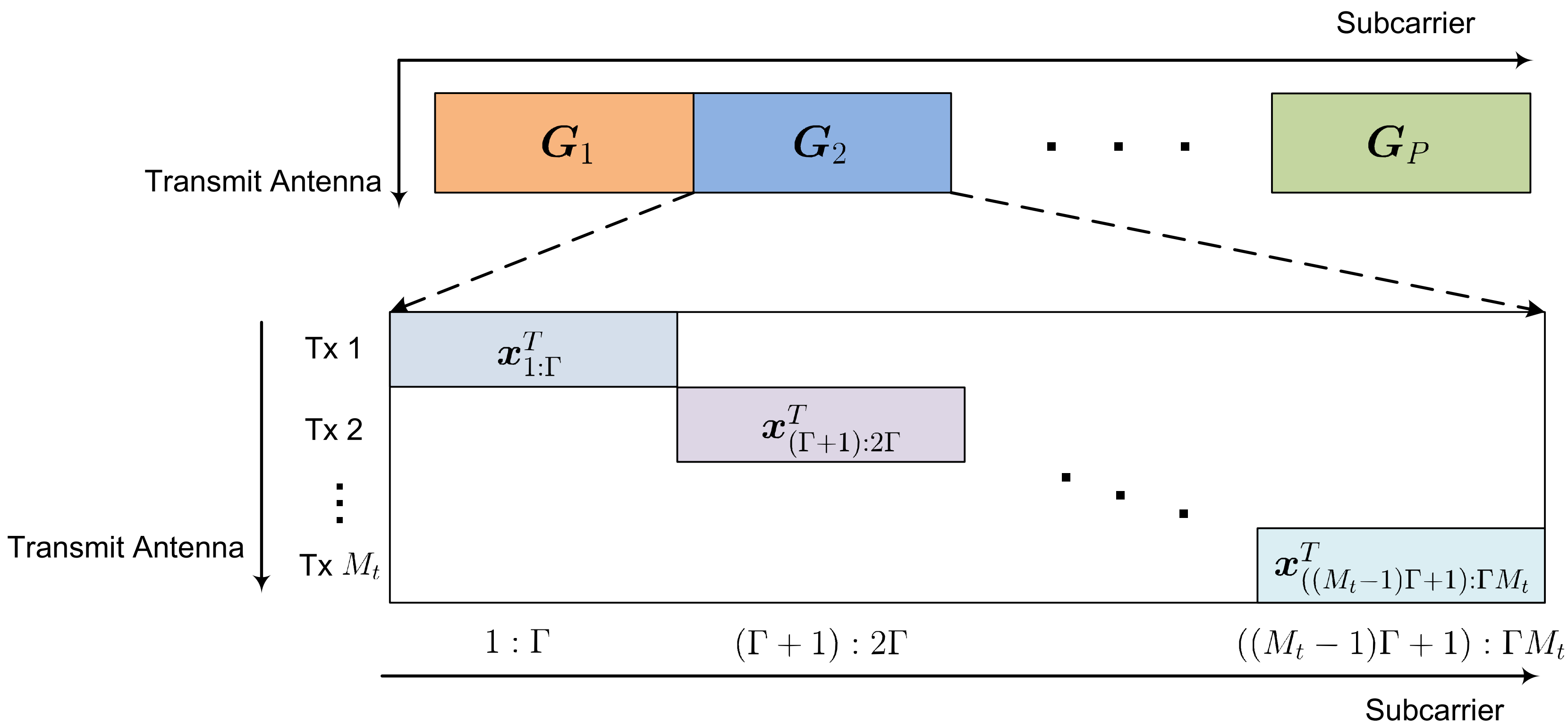
Example: Space-Time Frequency Coding

- **Scenario:** point-to-point MIMO OFDM in the presence of frequency selective multipath channels.



- **Goal:** precode across space and frequency, to harvest space and multipath diversity, esp., full space-multipath diversity.

- Let us have a case study on the algebraic space-frequency code (SFC) scheme **[Su-Safar-Liu'05]**.



- Operations:
  - Subcarriers are partitioned into groups;
  - In each group, symbols are precoded;
  - Precoded symbols ($\boldsymbol{x}$ above) are appropriately interleaved in space and frequency.

- Assume one rx antenna, for ease of illustration.

- The rx signal model in each group can be represented by

$$\boldsymbol{y} = \boldsymbol{D_H}\boldsymbol{x} + \boldsymbol{v},$$
$$\boldsymbol{x} = \boldsymbol{\Theta}\boldsymbol{s},$$

  where

$\boldsymbol{x} \in \mathbb{C}^{\Gamma M_t}$	the precoded symbol vector;
$\boldsymbol{\Theta} \in \mathbb{C}^{\Gamma M_t \times \Gamma M_t}$	the precoder matrix;
$\boldsymbol{s} \in \mathcal{S}^{\Gamma M_t}$	the tx symbol vector;
$\boldsymbol{D_H} \in \mathbb{C}^{\Gamma M_t \times \Gamma M_t}$	a diagonal matrix whose diagonals contain channel freq. responses (dependent on the SFC interleaving pattern).

- A properly designed $\boldsymbol{\Theta}$ can lead to full space-multipath diversity $d = M_t L$, where M_t is the no. of tx antennas & L is the no. of multipaths. To do so, one should choose $\Gamma \geq L$.

- The rx model can again be written as the generic form $\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C$. Note that the problem size in this example, ΓM_t, may be large.

Efficient High-Performance Approaches other than SDR

- **Sphere decoders** **[Mow'92]**, **[Viterbo-Biglieri'93]**, **[Damen-El-Gamal-Caire'03]**:
  - An exact ML solver based on branch and bound, or tree search;
  - Empirical experience with its runtime performance: very fast for high SNRs and small to moderate problem sizes N_C; can be (very) slow otherwise;
  - Exponential complexity w.r.t. the problem size **[Jaldén-Ottersten'05]**.

- **Lattice reduction aided (LRA) detectors** **[Yao-Wornell'02]**, **[Wübben-Seethaler-Jaldén-Matz'11]**:
  - Use lattice reduction to improve the channel "conditioning";
  - Interface well with linear and decision feedback detectors;
  - Exhibit good diversity or diversity multiplexing tradeoff performance **[Taherzadeh-Mobasher-Khandani'07]**, **[Jaldén-Elia'10]**.

Inhomogeneous QCQPs and SDR

- Consider a general inhomogeneous QCQP

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \ & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} + 2\boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \ & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} + 2\boldsymbol{a}_i^T\boldsymbol{x} \trianglerighteq_i b_i, \quad i = 1,\ldots,m. \end{aligned}$$

- An inhomogeneous QCQP can be reformulated as a homogenous QCQP

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}} \ & \begin{bmatrix} \boldsymbol{x}^T & t \end{bmatrix} \begin{bmatrix} \boldsymbol{C} & \boldsymbol{c} \\ \boldsymbol{c}^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ t \end{bmatrix} \\ \text{s.t.} \ & t^2 = 1, \\ & \begin{bmatrix} \boldsymbol{x}^T & t \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_i & \boldsymbol{a}_i \\ \boldsymbol{a}_i^T & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ t \end{bmatrix} \trianglerighteq_i b_i, \quad i = 1,\ldots,m \end{aligned}$$

and then handled by SDR.

An Alternative Way to Derive SDR for Inhomogeneous QCQPs

- Recap of inhomogeneous QCQP:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathbb{R}^n} \ & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} + 2\boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \ & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x} + 2\boldsymbol{a}_i^T\boldsymbol{x} \trianglerighteq_i b_i, \quad i = 1,\ldots,m. \end{aligned}$$

- By letting $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^T$, and then by replacing it with

$$\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T,$$

we can derive an SDR

$$\begin{aligned} \min_{\boldsymbol{X}\in\mathbb{S}^n, \boldsymbol{x}\in\mathbb{R}^n} \ & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) + 2\boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \ & \mathrm{Tr}(\boldsymbol{A}_i\boldsymbol{X}) + 2\boldsymbol{a}_i^T\boldsymbol{x} \trianglerighteq_i b_i, \quad i = 1,\ldots,m, \\ & \boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T. \end{aligned}$$

- This inhomogeneous SDR is equivalent to the SDR from the homogenized QCQP formulation (last page), by Schur complement $\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^T \Longleftrightarrow \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^T & 1 \end{bmatrix} \succeq \boldsymbol{0}$.

SDR MIMO Detection for QPSK Constellations

- Let $N = 2N_C$, $M = 2M_C$,

$$\boldsymbol{y} = \begin{bmatrix} \mathrm{Re}\{\boldsymbol{y}_C\} \\ \mathrm{Im}\{\boldsymbol{y}_C\} \end{bmatrix}, \boldsymbol{s} = \begin{bmatrix} \mathrm{Re}\{\boldsymbol{s}_C\} \\ \mathrm{Im}\{\boldsymbol{s}_C\} \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} \mathrm{Re}\{\boldsymbol{v}_C\} \\ \mathrm{Im}\{\boldsymbol{v}_C\} \end{bmatrix}, \boldsymbol{H} = \begin{bmatrix} \mathrm{Re}\{\boldsymbol{H}_C\} & -\mathrm{Im}\{\boldsymbol{H}_C\} \\ \mathrm{Im}\{\boldsymbol{H}_C\} & \mathrm{Re}\{\boldsymbol{H}_C\} \end{bmatrix}.$$

  The complex-valued model $\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C$ can be turned to a real one

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{v}.$$

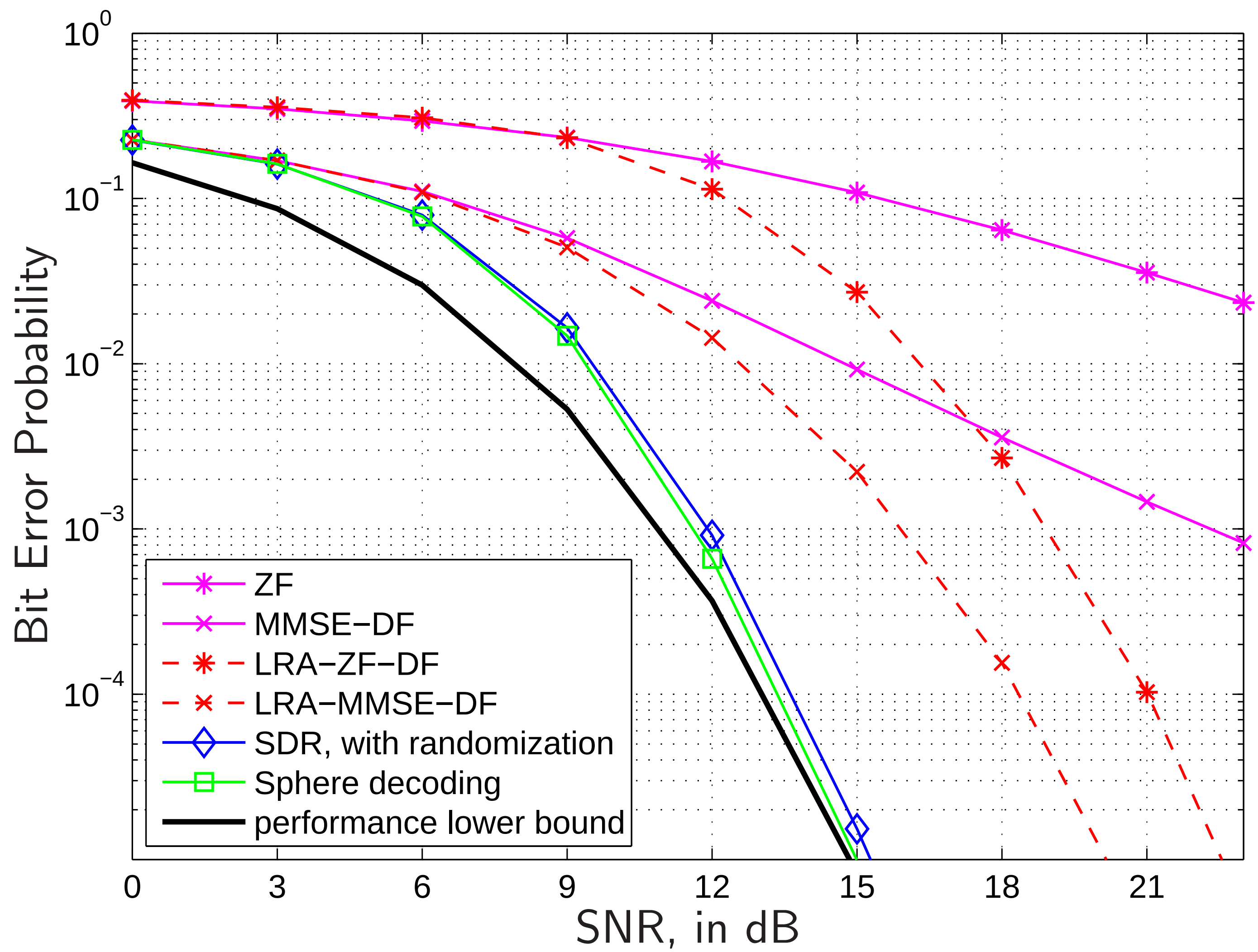
  where $\boldsymbol{s} \in \{\pm 1\}^N$ for QPSK constellations.

- ML detection problem:

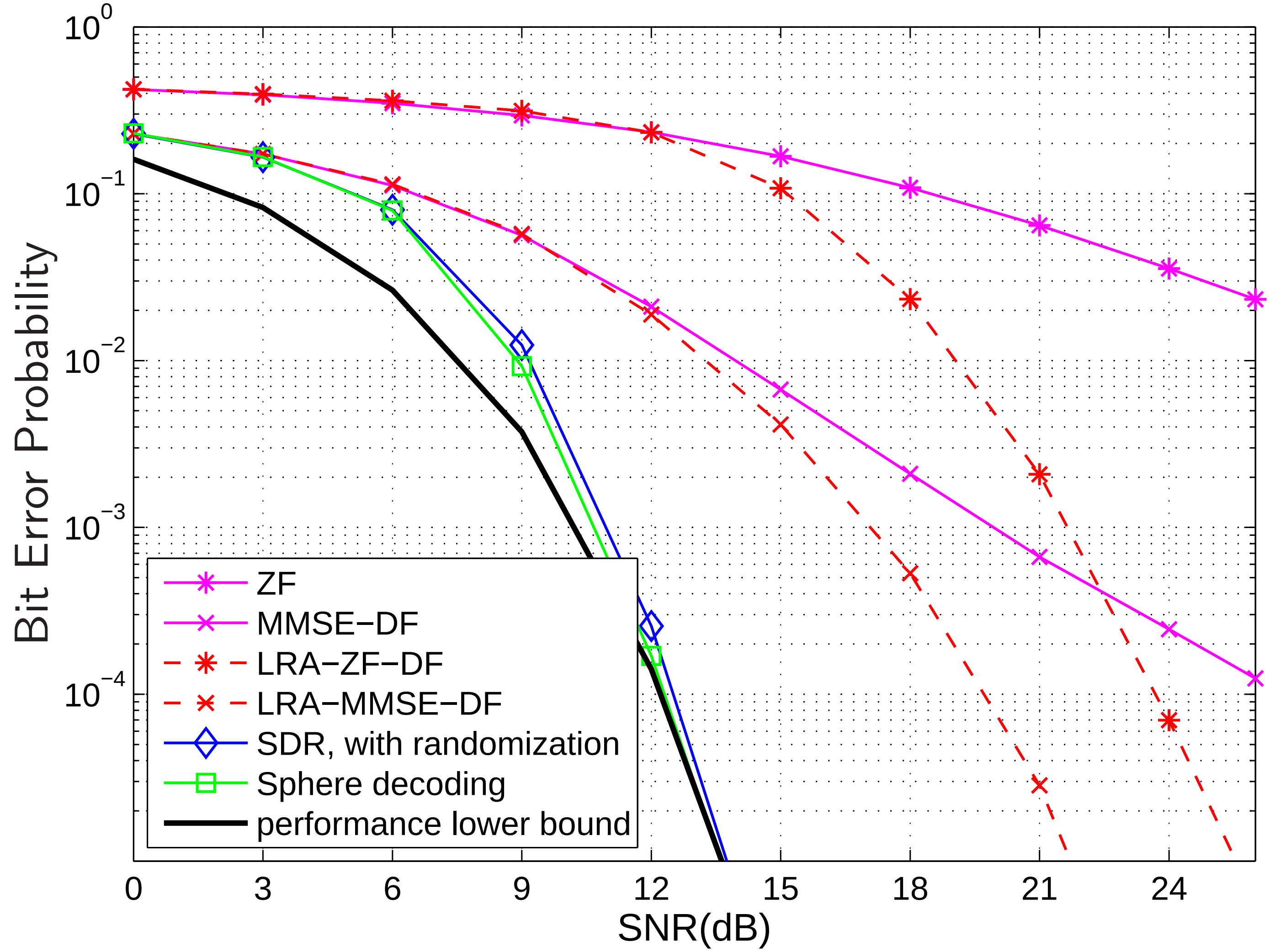
$$\begin{aligned} \min_{\boldsymbol{s} \in \mathbb{R}^N} \ & \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 \\ \text{s.t. } & s_i^2 = 1, \ i = 1, \ldots, N. \end{aligned}$$

- SDR:

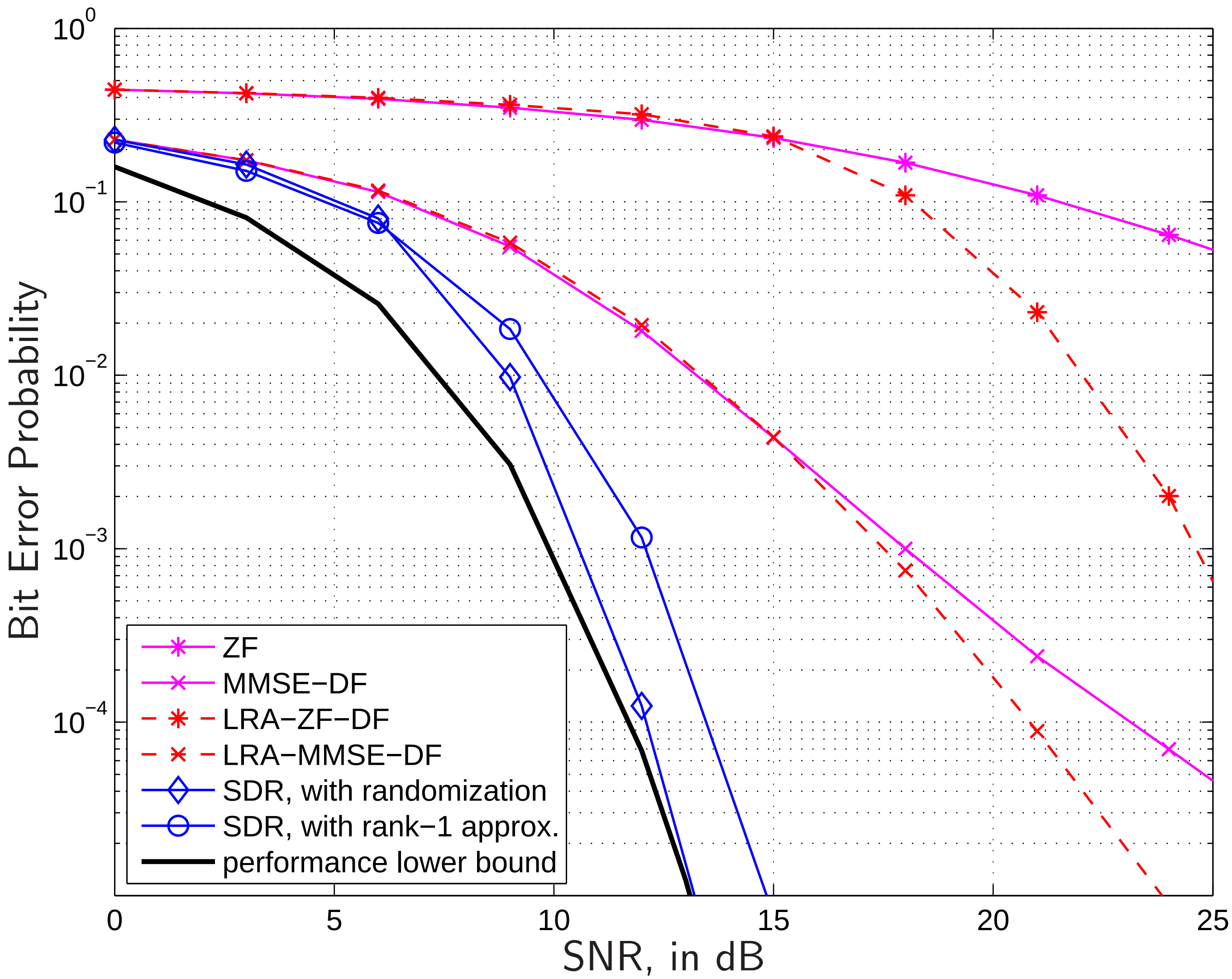
$$\begin{aligned} \min_{\boldsymbol{S} \in \mathbb{S}^N, \boldsymbol{s} \in \mathbb{R}^N} \ & \mathrm{Tr}(\boldsymbol{H}^T \boldsymbol{H} \boldsymbol{S}) - 2\boldsymbol{s}^T \boldsymbol{H}^T \boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t. } & S_{ii} = 1, \ i = 1, \ldots, N, \\ & \boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T. \end{aligned}$$



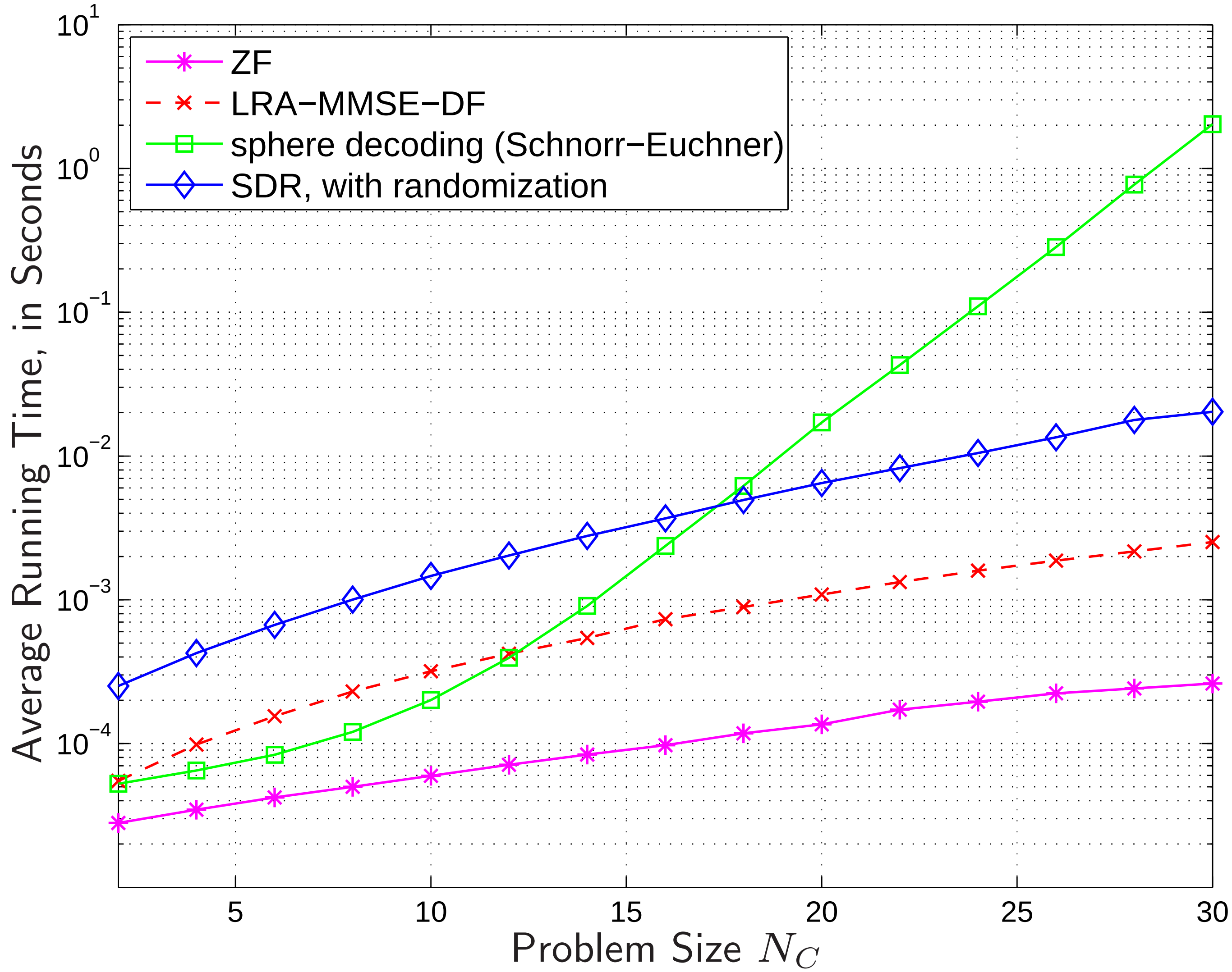
Bit error rate performance under $(M_C, N_C) = (10, 10)$, QPSK constellations. The SNR is defined as $\frac{\mathrm{E}\{\|\boldsymbol{H}_C \boldsymbol{s}_C\|^2\}}{\mathrm{E}\{\|\boldsymbol{v}_C\|^2\}}$. 'ZF'— zero forcing; 'MMSE-DF'— min. mean square error with decision feedback; 'LRA'— lattice reduction aided; the Schnorr-Euchner sphere decoder is used.



Bit error rate performance under $(M_C, N_C) = (20, 20)$, QPSK constellations.



Bit error rate performance under $(M_C, N_C) = (40, 40)$, QPSK constellations. It is too expensive to run sphere decoding in this example.



Complexity comparison of various MIMO detectors. SNR$= 12$dB.

Computational Efficiency of SDR MIMO Detection

- The bulk of complexity lies in solving the SDP.

- A common, arguably dominant, way to solve SDPs is to use the interior point methods (IPMs)— their solution precision is good, & their complexities are provably polynomial-time in the problem size.

- For the SDP in QPSK SDR MIMO detection, an IPM can output a solution with a worst-case complexity of

$$\mathcal{O}((N+1)^{3.5}\log(\epsilon^{-1})) \simeq \mathcal{O}(N^{3.5}),$$

  where $\epsilon > 0$ is the desired solution accuracy.

- A few practical hints:
  - You don't need a very small ϵ in MIMO detection, since you will round the solution anyway.
  - While a general purpose software, such as `CVX`, can be used to solve the SDP conveniently, you'd better off write your own IPM for maximizing the computational efficiency.

Interior-Point Algorithm for SDR MIMO Detection

The SDR problem in homogenous form:

$$\min_{\boldsymbol{X} \in \mathbb{S}^n} \ \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X})$$
$$\text{s.t.} \ \boldsymbol{X} \succeq \boldsymbol{0}, \ X_{ii} = 1, \ i = 1, \ldots, n,$$

where $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{H}^T\boldsymbol{H} & -\boldsymbol{H}^T\boldsymbol{y} \\ -\boldsymbol{y}^T\boldsymbol{H} & \|\boldsymbol{y}\|^2 \end{bmatrix}$, $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{S} & \boldsymbol{s} \\ \boldsymbol{s}^T & 1 \end{bmatrix}$. By exploiting its simple equality constraint structure, a specialized (and fast) IPM can be derived **[Helmberg-*et al.*'96]**.

> **given** $\epsilon > 0$, and strictly feasible $\boldsymbol{X}$, $\boldsymbol{y}$, and $\boldsymbol{Z}$.
>
> **repeat**
>
> 1. update the barrier parameter $\mu := \mathrm{tr}(\boldsymbol{Z}\boldsymbol{X})/2n$.
> 2. compute
>
> $$\Delta\boldsymbol{y} := [(\boldsymbol{Z}^{-1} \circ \boldsymbol{X})]^{-1}(\mu \mathrm{diag}(\boldsymbol{Z}^{-1}) - \boldsymbol{1}),$$
> $$\Delta\boldsymbol{Z} := \mathrm{Diag}(\Delta\boldsymbol{y})$$
> $$\Delta\boldsymbol{X} := \mu\boldsymbol{Z}^{-1} - \boldsymbol{X} - \boldsymbol{Z}^{-1}\Delta\boldsymbol{Z}\boldsymbol{X}, \quad \Delta\boldsymbol{X} := (\Delta\boldsymbol{X} + \Delta\boldsymbol{X}^T)/2$$
>
> 3. find step-sizes $\alpha_p \in (0,1]$ and $\alpha_d \in (0,1]$ such that $\boldsymbol{X} + \alpha_p\Delta\boldsymbol{X} \succ \boldsymbol{0}$ and $\boldsymbol{Z} + \alpha_d\Delta\boldsymbol{Z} \succ \boldsymbol{0}$.
> 4. $\boldsymbol{X} := \boldsymbol{X} + \alpha_p\Delta\boldsymbol{X}$, $\boldsymbol{y} := \boldsymbol{y} + \alpha_d\Delta\boldsymbol{y}$, and $\boldsymbol{Z} := \boldsymbol{Z} + \alpha_d\Delta\boldsymbol{Z}$.
>
> **until** $\mathrm{tr}(\boldsymbol{Z}\boldsymbol{X}) \leq \epsilon$.

Cheap SDR by Row-by-Row Coordinate Descent

- While IPMs have good solution fidelity, they are generally not low complexity options (check out the IPM pseudo code last page).

- Low complexity SDR implementation has received much interest.

- A possible alternative is row-by-row (RBR) coordinate descent **[Wen-Goldfarb-Ma-Scheinberg'09]**, **[Wai-Ma-So'11]**.

- Ready-to-use codes available at http://www.ee.cuhk.edu.hk/~wkma/mimo/.

- To describe RBR, consider a barrier SDR problem

$$\begin{aligned} \min_{\boldsymbol{X} \in \mathbb{S}^n} \ & \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) - \sigma \log\det(\boldsymbol{X}) \\ \text{s.t.} \ & X_{ii} = 1, \ i = 1, \ldots, n, \end{aligned} \qquad \text{(B-SDR)}$$

  where $\sigma > 0$ is the barrier parameter.

- In (B-SDR), the log barrier function is used to enforce $\boldsymbol{X} \succeq \boldsymbol{0}$ (more precisely, $\boldsymbol{X} \succ \boldsymbol{0}$), thereby avoiding to deal with the constraint $\boldsymbol{X} \succeq \boldsymbol{0}$ explicitly.

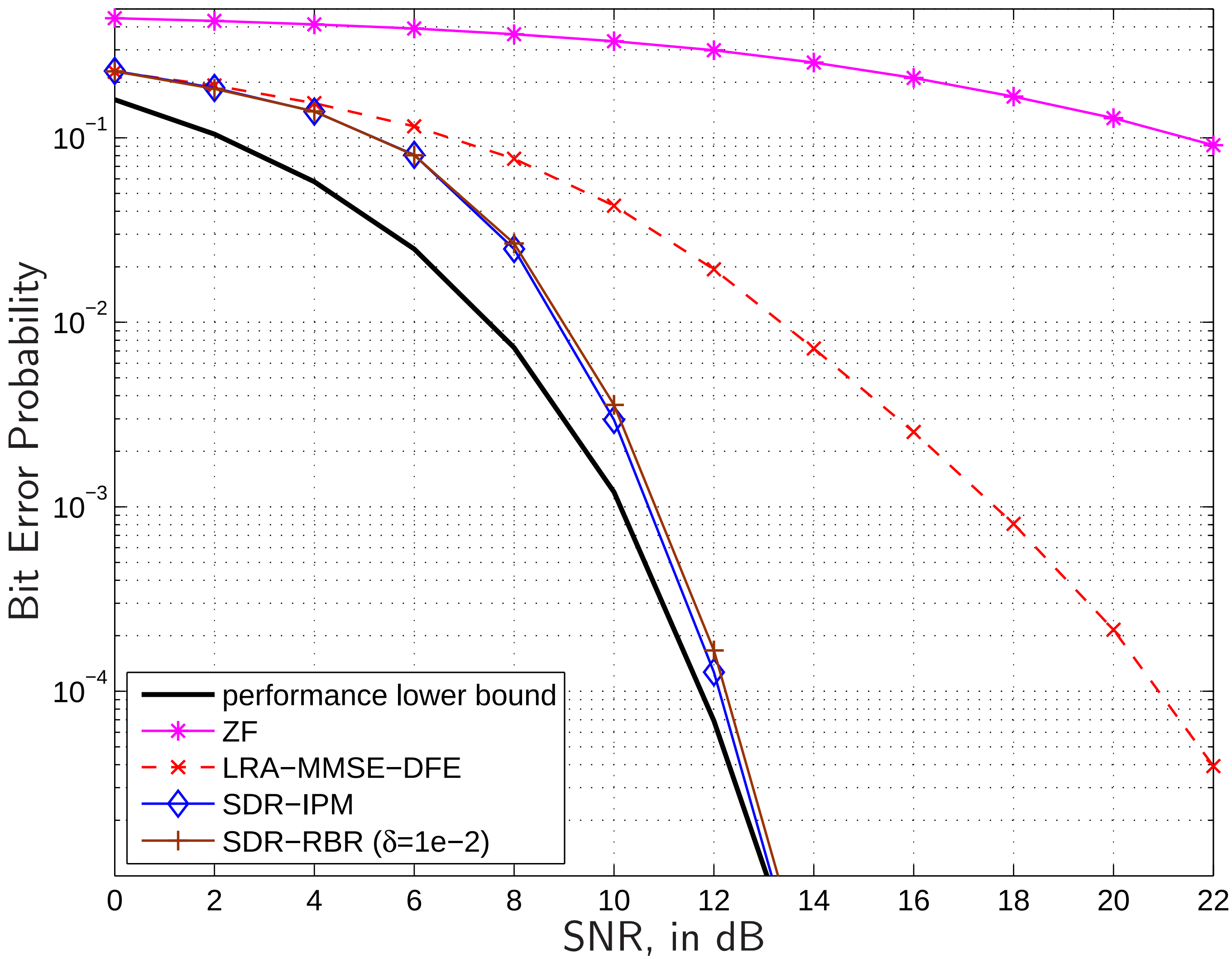
- Let $f(\boldsymbol{X}) = \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) - \sigma \log\det(\boldsymbol{X})$, $\boldsymbol{x}_i$ be the ith row of $\boldsymbol{X}$, & $\boldsymbol{X}_{-i}$ be the collection of all elements of $\boldsymbol{X}$ except for $\boldsymbol{x}_i$.
- **Idea of RBR**: do a block coordinate descent on (B-SDR):

> **given** a starting point $\hat{\boldsymbol{X}}$;
> **repeat**
>   **for** $i = 1, \ldots, n$
> $\hat{\boldsymbol{x}}_i := \arg \min_{\boldsymbol{x}_i,\ X_{ii}=1} f(\boldsymbol{x}_i, \hat{\boldsymbol{X}}_{-i})$;
>   **end;**
> **until** a stopping criterion is satisfied.

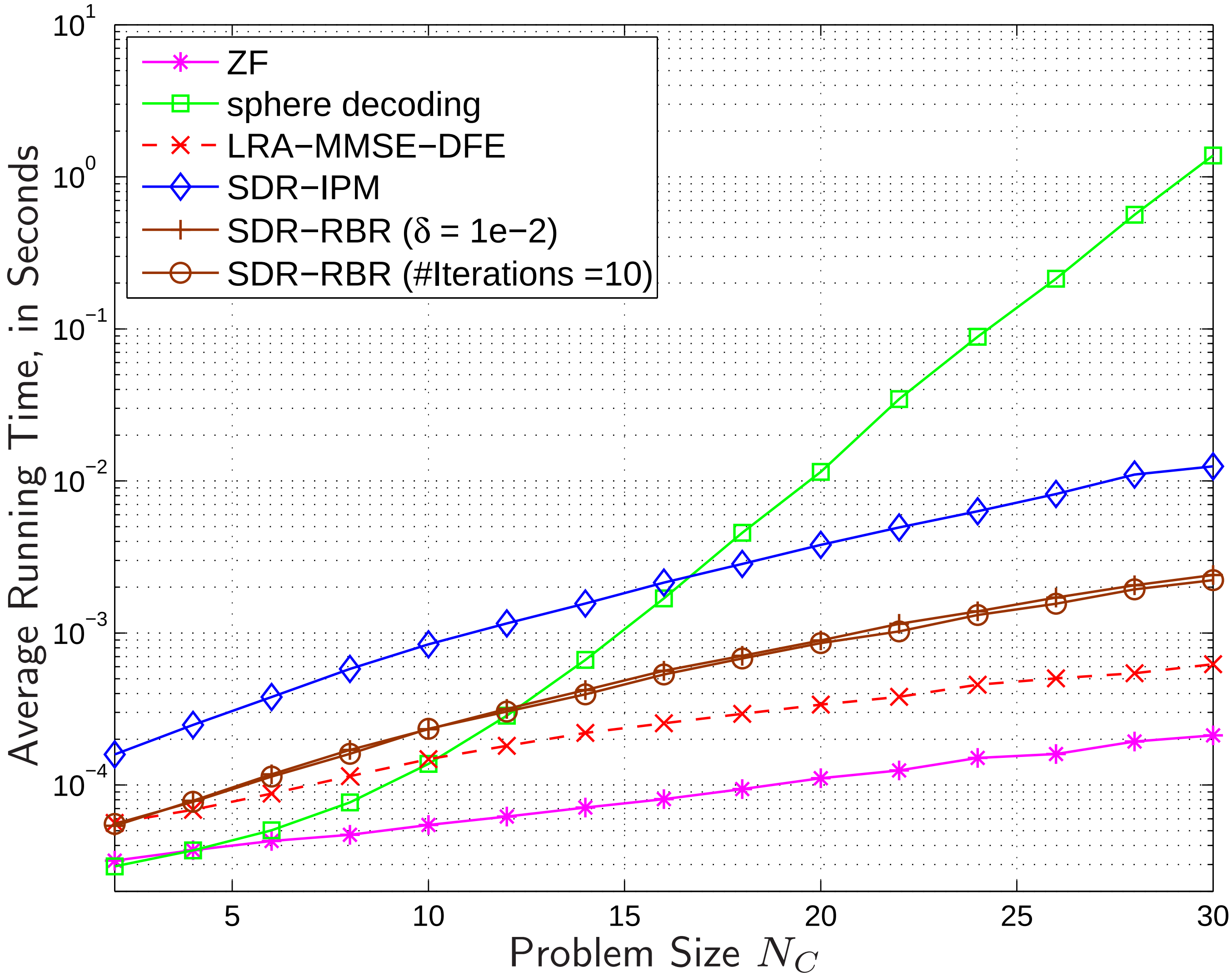
- The iterates are known to converge to the optimal solution of (B-SDR).
- Each per-row update is simple; e.g., the 1st row update can be equiv. written as

$$\min_{\boldsymbol{\xi}_1 \in \mathbb{R}^{n-1}} 2\boldsymbol{c}_1^T \boldsymbol{\xi}_1 - \sigma \log(1 - \boldsymbol{\xi}_1^T \hat{\boldsymbol{X}}^{\dagger}_{2:n,2:n} \boldsymbol{\xi}_1), \qquad (\S)$$

where $\boldsymbol{\xi}_1 = [\boldsymbol{x}_1]_{2:n}$, $\boldsymbol{c}_1 = \boldsymbol{C}_{1,2:n}$. The soln. to (§) is $\boldsymbol{\xi}_1^\star = \kappa \hat{\boldsymbol{X}}_{2:n,2:n} \boldsymbol{c}_1$ for some κ, a simple closed form (matrix multiplication, no inverse)!



Bit error probability performance under $(M_C, N_C) = (40, 40)$. Setting for RBR: $\sigma = 10^{-2}/n$; RBR terminates when $\left| \frac{f^{(k+1)} - f^{(k)}}{f^{(k)}} \right| \leq \delta$, where $f^{(k)}$ is the objective value at iteration k.



Complexity of RBR. A tenfold runtime saving relative to IPM is observed.

Other Relaxations for QPSK ML MIMO Detection

- Generally speaking, relaxation methods work by relaxing the original problem to a tractable problem.
- In that regard, relaxations other than SDR can be considered.
- Unconstrained relaxation (UR):

$$\min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2.$$

  The result is ZF.

- On-Sphere Relaxation (OSR):

$$\min_{\|\boldsymbol{s}\|^2=N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2.$$

  The solution is $\hat{\boldsymbol{s}}_{\sf OSR} = (\boldsymbol{H}^T\boldsymbol{H}+\gamma\boldsymbol{I})^{-1}\boldsymbol{H}^T\boldsymbol{y}$ for some γ; has an MMSE flavor.

- Box Relaxation (BR):

$$\min_{s_i^2\leq 1,\ i=1,\ldots,N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2.$$

Comparison of the Various Relaxations

- In order to compare, let

$$
\begin{aligned}
f^\star_{\mathsf{ML}} &= \min_{\boldsymbol{s}\in\{\pm1\}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2, \\
f^\star_{\mathsf{SDR}} &= \min_{\boldsymbol{S}\succeq\boldsymbol{s}\boldsymbol{s}^T,\ S_{ii}=1\ \forall i} \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2, \\
f^\star_{\mathsf{BR}} &= \min_{s_i^2\leq 1,\ i=1,\ldots,N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2, \\
f^\star_{\mathsf{OSR}} &= \min_{\|\boldsymbol{s}\|^2=N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2, \quad f^\star_{\mathsf{UR}} = \min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2.
\end{aligned}
$$

- It is shown that **[Ma-Davidson-Wong-Luo-Ching'02]**, **[Poljak-Rendl-Wolkowicz'95]**

$$
\max\{f^\star_{\mathsf{UR}}, f^\star_{\mathsf{OSR}}, f^\star_{\mathsf{BR}}\} \leq f^\star_{\mathsf{SDR}} \leq f^\star_{\mathsf{ML}}.
$$

- The result means that SDR provides a relaxation no worse than the other three methods. Or, the other methods may be seen as further relaxations of SDR.

Regularization in LS

- Consider the least squares (LS) problem (for generic applications):

$$\min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2.$$

- Sometimes, in order to make the problem better conditioned, we may turn to a regularized LS:

$$\min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T\boldsymbol{T}\boldsymbol{s},$$

  for some regularizer $\boldsymbol{T} \in \mathbb{S}^N$ (common choice: $\boldsymbol{T} = \rho\boldsymbol{I}$, $\rho > 0$).

- SDR can be interpreted as a regularized LS.

A Regularized LS Perspective on SDR

- Consider a Lagrangian dual of ML, as an approx.:

$$
\begin{aligned}
f^\star_{\sf ML} \geq g^\star_{\sf ML} &= \max_{\boldsymbol{\lambda}\in\mathbb{R}^N} \min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2 + \sum_{i=1}^N \lambda_i (s_i^2 - 1) \\
&= \max_{\boldsymbol{\lambda}\in\mathbb{R}^N} -\boldsymbol{\lambda}^T\mathbf{1} + \underbrace{\min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T D(\boldsymbol{\lambda})\boldsymbol{s}}_{\text{regularized LS}}, \qquad \text{(D)}
\end{aligned}
$$

  where $g^\star_{\sf ML}$ is the dual optimal value, $D(\cdot)$ is a diagonal operator.

- (D) intends to find a 'best' regularization in a $\{\pm 1\}$ LS context.

- SDR is equivalent to (D):

$$f^\star_{\sf SDR} = g^\star_{\sf ML}.$$

  Also, the dual of SDR is (D) (the trick: strong duality of convex problems).

- Recap of SDR in dual form

$$f^\star_{\sf SDR} = \max_{\boldsymbol{\lambda}\in\mathbb{R}^N} \; -\boldsymbol{\lambda}^T\mathbf{1} + \min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T D(\boldsymbol{\lambda})\boldsymbol{s}.$$

- Consider OSR and BR. By strong Lagrangian duality, they can be expressed as

$$f^\star_{\sf OSR} = \max_{\boldsymbol{\lambda}=\gamma\mathbf{1},\ \gamma\in\mathbb{R}} \; -\boldsymbol{\lambda}^T\mathbf{1} + \min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T D(\boldsymbol{\lambda})\boldsymbol{s},$$

$$f^\star_{\sf BR} = \max_{\boldsymbol{\lambda}\succeq\mathbf{0}} \; -\boldsymbol{\lambda}^T\mathbf{1} + \min_{\boldsymbol{s}\in\mathbb{R}^N} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{s}\|^2 + \boldsymbol{s}^T D(\boldsymbol{\lambda})\boldsymbol{s}.$$

- Apart from showing a regularized LS interpretation of OSR and BR, the above eqs. reveal that the feasible set of $\boldsymbol{\lambda}$ in SDR subsumes that in OSR and BR.

- Hence, we can conclude the previous result that

$$f^\star_{\sf SDR} \geq f^\star_{\sf OSR}, \quad f^\star_{\sf SDR} \geq f^\star_{\sf BR}.$$

SDR MIMO Detection for MPSK Constellations

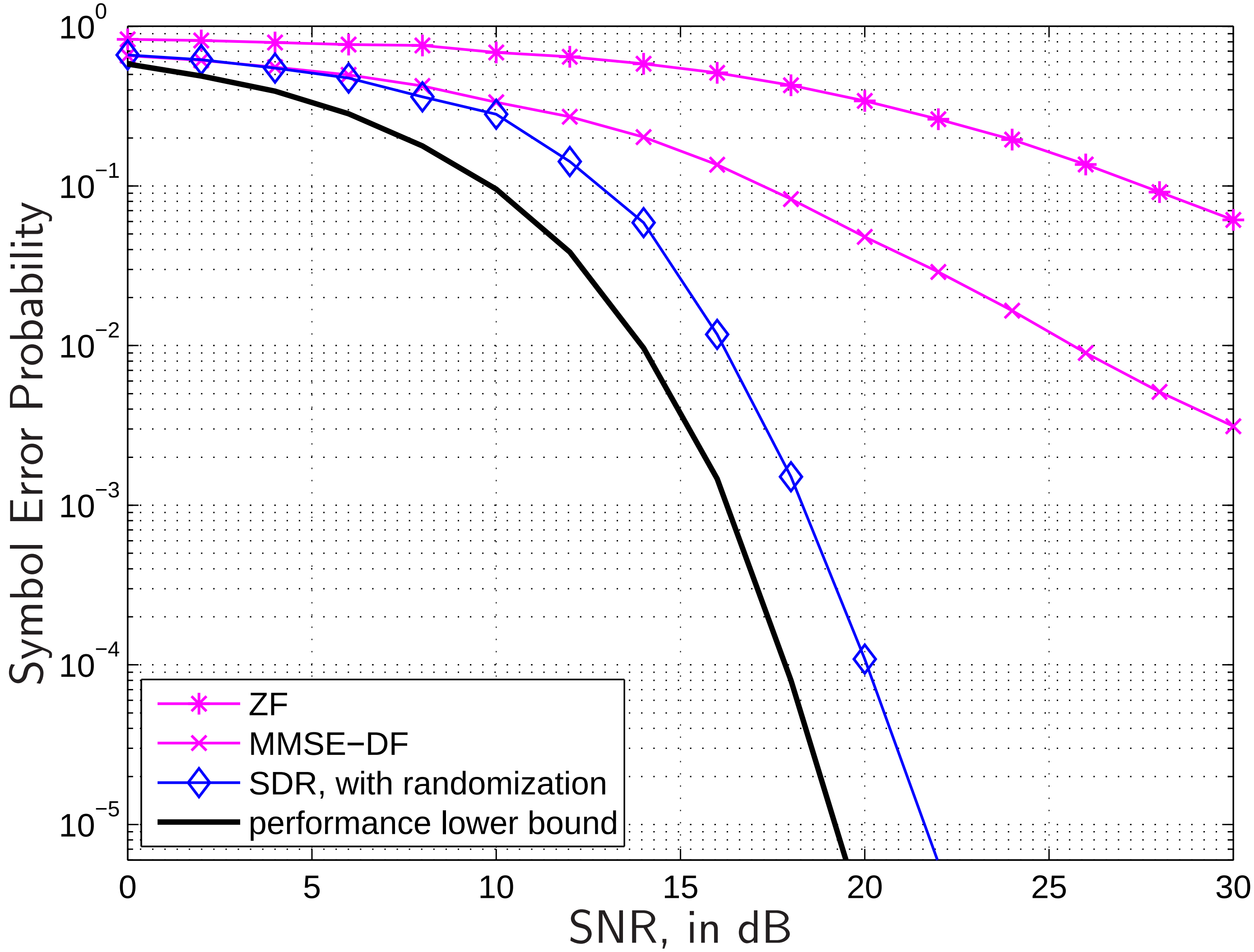
- The ML problem in the MPSK case:

$$\begin{aligned} \min_{\boldsymbol{s}_C \in \mathbb{C}^{N_C}} \quad & \|\boldsymbol{y}_C - \boldsymbol{H}_C \boldsymbol{s}_C\|^2 \\ \text{s.t.} \quad & s_{C,i} \in \{1, e^{j2\pi/M}, \ldots, e^{j2\pi(M-1)/M}\}, \quad i = 1, \ldots, N_C. \end{aligned}$$

- Intuition: relax the constellations constraints to $|s_{C,i}|^2 = 1$, & then apply SDR.

- Following this intuition, we can formulate a complex-valued SDR **[Ma-Ding-Ching'04]**:

$$\begin{aligned} \min_{\boldsymbol{S}_C \in \mathbb{H}^{N_C},\ \boldsymbol{s}_C \in \mathbb{C}^{N_C}} \quad & \mathrm{Tr}({\boldsymbol{H}_C}^H \boldsymbol{H}_C \boldsymbol{S}_C) - 2\mathrm{Re}\{{\boldsymbol{s}_C}^H {\boldsymbol{H}_C}^H \boldsymbol{y}_C\} + \|\boldsymbol{y}_C\|^2 \\ \text{s.t.} \quad & [\boldsymbol{S}_C]_{ii} = 1, \quad i = 1, \ldots, N_C, \\ & \boldsymbol{S}_C \succeq \boldsymbol{s}_C {\boldsymbol{s}_C}^H. \end{aligned}$$



Symbol error rate performance under $(M_C, N_C) = (20, 20)$, 8-PSK constellations. Note that LRA methods are not applicable to MPSK constellations.

SDR MIMO Detection for Higher-Order QAM

- Assume 16-QAM constellations, for ease of illustration.

- The ML problem (under the equivalent real-valued model):

$$\min_{\boldsymbol{s} \in \{\pm 1, \pm 3\}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2.$$

- A number of attempts have been made for SDR of 16-QAM ML detection **[Wiesel-Eldar-Shamai'05]**, **[Sidiropoulos-Luo'06]**, **[Yang-Zhao-Zhou-Wu'07]**, **[Mobasher-Taherzah-Sotirov-Khandani'07]**, **[Mao-Wang-Wang'07]**.

- We consider
  - polynomial inspired SDR (PI-SDR) **[Wiesel-Eldar-Shamai'05]**;
  - bound constrained SDR (BC-SDR) **[Sidiropoulos-Luo'06]**;
  - virtually antipodal SDR (VA-SDR) **[Mao-Wang-Wang'07]**.

Bound Constrained SDR (BC-SDR) **[Sidiropoulos-Luo'06]**:

- The 16-QAM ML problem is equivalent to

$$\begin{aligned} \min_{\boldsymbol{S}\in\mathbb{S}^N, \boldsymbol{s}\in\mathbb{R}^N} \quad & \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\mathbf{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t.} \quad & \boldsymbol{S} = \boldsymbol{s}\boldsymbol{s}^T, \\ & S_{ii} \in \{1,9\}, \ i = 1,\ldots,N. \qquad (\Leftrightarrow s_i^2 \in \{1,9\}) \end{aligned}$$

- Relaxing $\boldsymbol{S} = \boldsymbol{s}\boldsymbol{s}^T$ to $\boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T$ is not enough to yield a convex relaxation.

- BC-SDR also relaxes $\{1,9\}$ to $[1,9]$, leading to

$$\begin{aligned} \min_{\boldsymbol{S}\in\mathbb{S}^N, \boldsymbol{s}\in\mathbb{R}^N} \quad & \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\mathbf{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t.} \quad & \boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T, \\ & 1 \le S_{ii} \le 9, \ i = 1,\ldots,N. \end{aligned} \qquad \text{(BC-SDR)}$$

- BC-SDR is simple to implement, and a specialized IPM is available **[Ma-Su-Jaldén-Chi'08]**.

Polynomial Inspired SDR (PI-SDR) [Wiesel-Eldar-Shamai'05]:

- PI-SDR uses the fact that

$$u \in \{1, 9\} \Longleftrightarrow (u-1)(u-9) = 0 \Longleftrightarrow u^2 - 10u + 9 = 0$$

to reformulate the ML problem as

$$\begin{array}{ll} \min\limits_{\boldsymbol{S},\boldsymbol{s},\boldsymbol{U},\boldsymbol{u}} & \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t.} & \boldsymbol{S} = \boldsymbol{s}\boldsymbol{s}^T, \quad \boldsymbol{U} = \boldsymbol{u}\boldsymbol{u}^T, \\ & d(\boldsymbol{S}) = \boldsymbol{u}, \quad d(\boldsymbol{U}) - 10\boldsymbol{u} + 9\mathbf{1} = \mathbf{0}. \quad (\Leftrightarrow u_i^2 - 10u_i + 9 = 0, \forall i) \end{array}$$

where $d : \mathbb{R}^{N \times N} \to \mathbb{R}^N$ is the diagonal operator.

- PI-SDR is the SDR of the polynomial ML formulation:

$$\begin{array}{ll} \min & \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t.} & \boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T, \quad \boldsymbol{U} \succeq \boldsymbol{u}\boldsymbol{u}^T, \\ & d(\boldsymbol{S}) = \boldsymbol{u}, \quad d(\boldsymbol{U}) - 10\boldsymbol{u} + 9\mathbf{1} = \mathbf{0}. \end{array} \qquad \text{(PI-SDR)}$$

Virtually Antipodal SDR (VA-SDR) [Mao-Wang-Wang'07]:

- VA-SDR uses the fact that

$$s \in \{\pm 1, \pm 3\} \Longleftrightarrow s = b_1 + 2b_2, \quad b_1, b_2 \in \{\pm 1\}.$$

to rewrite the ML problem in a virtually antipodal form

$$\min_{\boldsymbol{b}_1, \boldsymbol{b}_2 \in \{\pm 1\}^N} \|\boldsymbol{y} - \boldsymbol{H}(\boldsymbol{b}_1 + 2\boldsymbol{b}_2)\|^2 = \min_{\boldsymbol{b} \in \{\pm 1\}^{2N}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{W}\boldsymbol{b}\|^2,$$

where $\boldsymbol{W} = [\ \boldsymbol{I}\ 2\boldsymbol{I}\ ], \boldsymbol{b} = [\ \boldsymbol{b}_1^T\ \boldsymbol{b}_2^T\ ]^T$.

- By applying the same SDR as in QPSK constellations, VA-SDR is obtained:

$$\begin{array}{ll} \min & \mathrm{Tr}(\boldsymbol{W}^T\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{W}\boldsymbol{B}) - 2\boldsymbol{b}^T\boldsymbol{W}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t.} & \boldsymbol{B} \succeq \boldsymbol{b}\boldsymbol{b}^T, \quad B_{ii} = 1, \quad i = 1, \ldots, 2N. \end{array} \qquad \text{(VA-SDR)}$$

- Rather unexpectedly, the three SDRs are equivalent **[Ma-Su-Jaldén-Chang-Chi'09]**.

- Consider a unified SDR expression

$$\min_{(\boldsymbol{S},\boldsymbol{s})\in\mathcal{F}} \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2,$$

where $\mathcal{F}$ depends on the SDR employed:

$$\mathcal{F}_{\mathsf{BC-SDR}} = \{ \ (\mathbf{S},\mathbf{s}) \mid \mathbf{S} \succeq \mathbf{s}\mathbf{s}^T, \mathbf{1} \preceq d(\mathbf{S}) \preceq 9\mathbf{1} \ \},$$

$$\mathcal{F}_{\mathsf{PI-SDR}} = \{ \ (\mathbf{S},\mathbf{s}) \mid (\mathbf{U},\mathbf{u},\mathbf{S},\mathbf{s}) \in \mathcal{W}_{\mathsf{PI-SDR}} \ \},$$

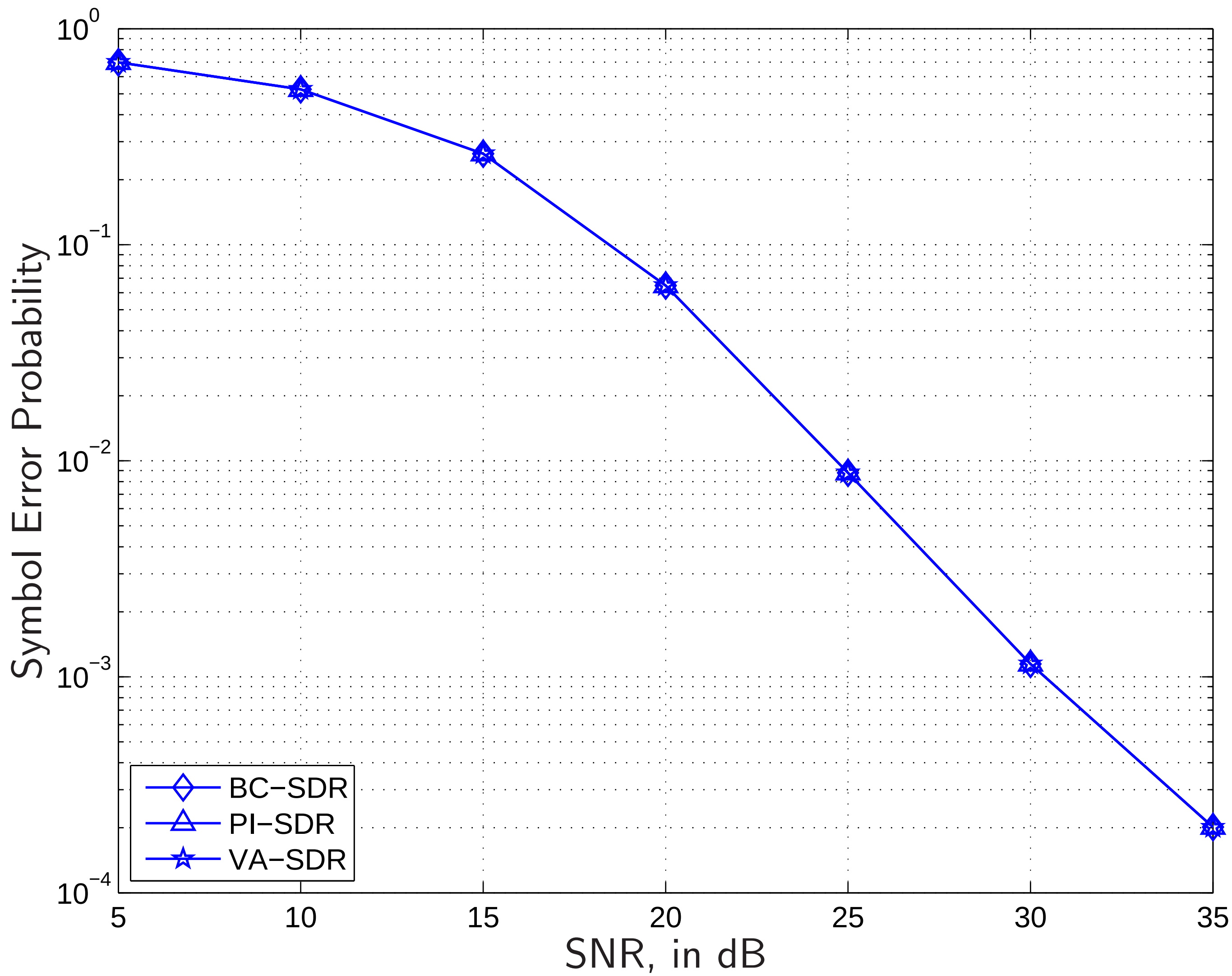
$$\mathcal{W}_{\mathsf{PI-SDR}} = \{ \ (\mathbf{U},\mathbf{u},\mathbf{S},\mathbf{s}) \mid \mathbf{U} \succeq \mathbf{u}\mathbf{u}^T, \mathbf{S} \succeq \mathbf{s}\mathbf{s}^T, d(\mathbf{S}) = \mathbf{u}, d(\mathbf{U}) - 10\mathbf{u} + 9\mathbf{1} = \mathbf{0} \ \},$$

$$\mathcal{F}_{\mathsf{VA-SDR}} = \{ \ (\mathbf{S},\mathbf{s}) = (\mathbf{W}\mathbf{B}\mathbf{W}^T, \mathbf{W}\mathbf{b}) \mid \mathbf{B} \succeq \mathbf{b}\mathbf{b}^T, d(\mathbf{B}) = \mathbf{1} \ \}.$$

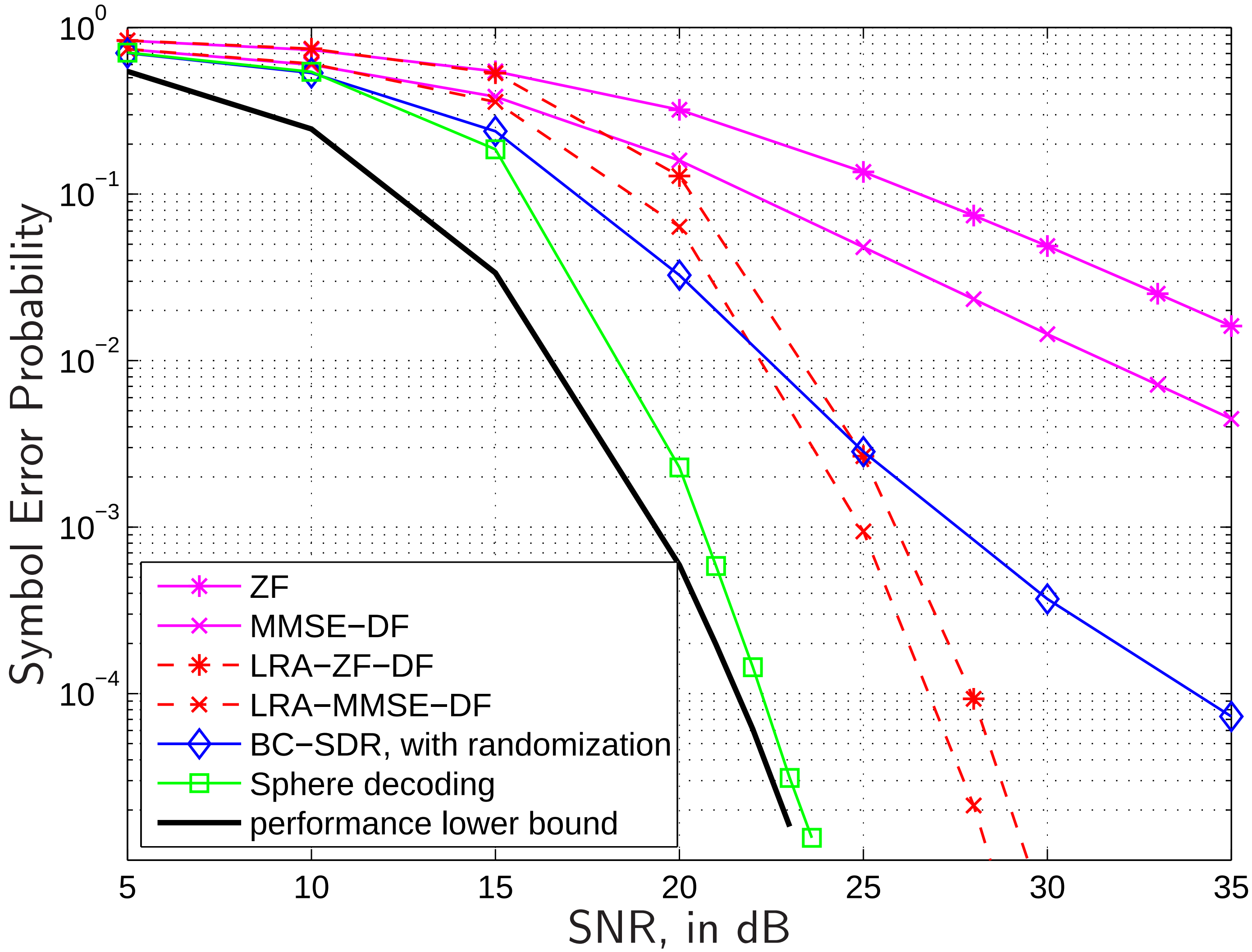
- It is shown by analysis that

$$\mathcal{F}_{\mathsf{BC-SDR}} = \mathcal{F}_{\mathsf{PI-SDR}} = \mathcal{F}_{\mathsf{VA-SDR}}.$$

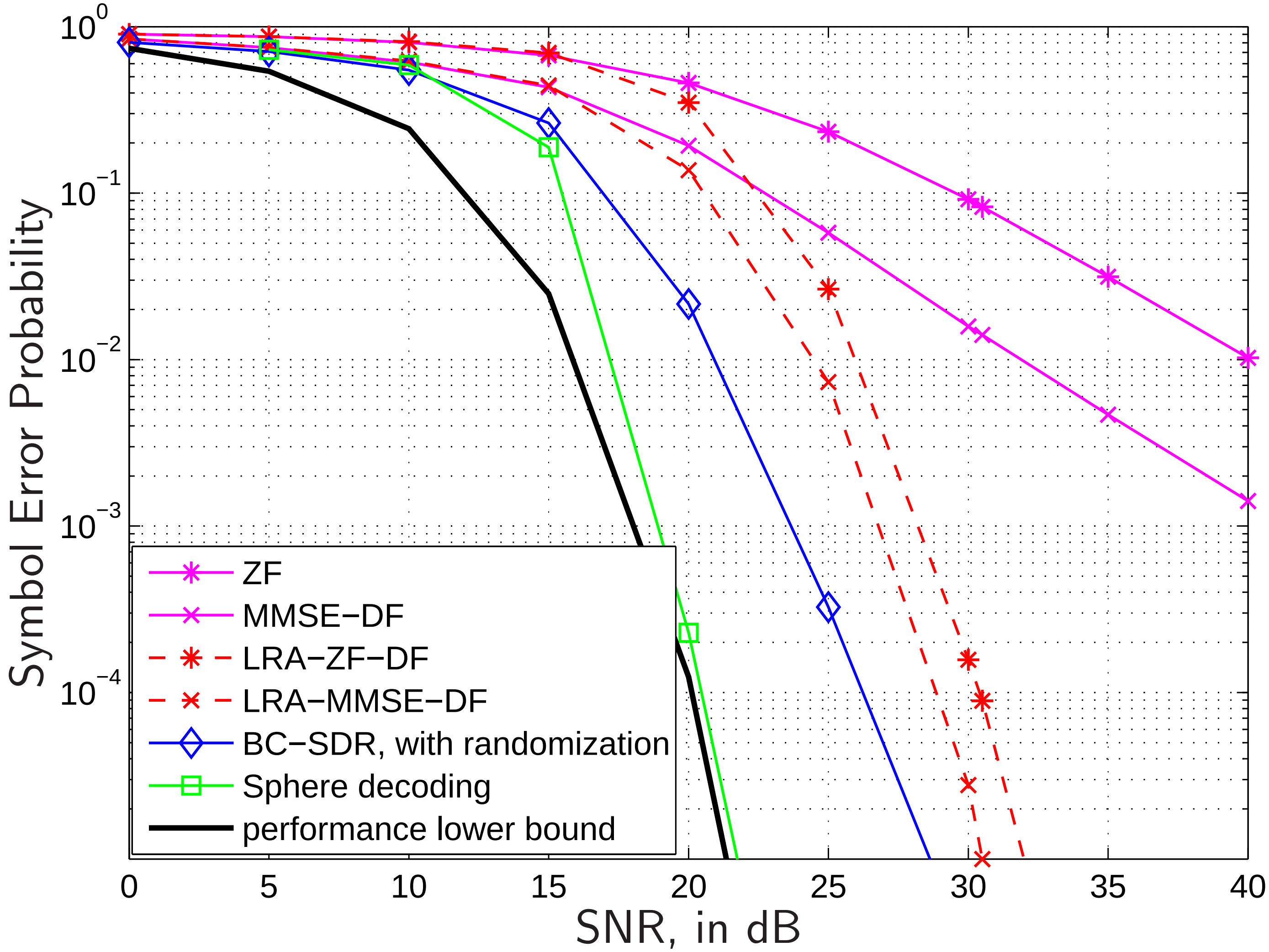
The same equivalence is also proven for 64-QAM PI-SDR, & for any 2^q-QAM VA-SDR.



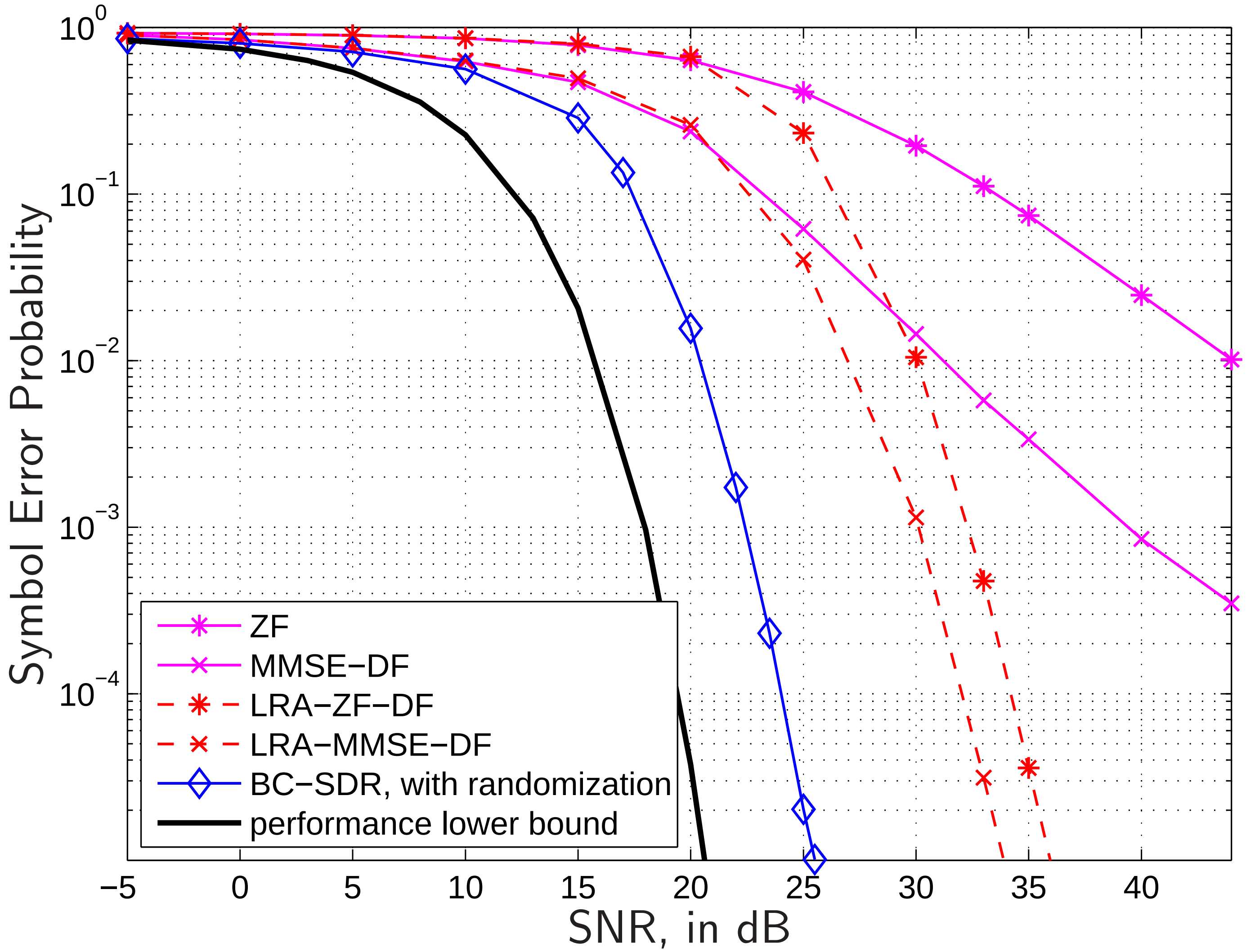
Symbol error rate performance of BC-SDR, PI-SDR, and VA-SDR under $(M_C, N_C) = (8, 8)$, 16-QAM constellations. The three performance plots coincide.



Symbol error rate performance under $(M_C, N_C) = (8, 8)$, 16-QAM constellations.



Symbol error rate performance under $(M_C, N_C) = (16, 16)$, 16-QAM constellations.



Symbol error rate performance under $(M_C, N_C) = (40, 40)$, 16-QAM constellations. Sphere decoding is too expensive to run in this case.

Some Results in Performance Analysis

- Assume QPSK or BPSK constellations. SDR has a high probability of giving a rank-one solution, for high SNRs **[Jaldén-Martin-Ottersten'03]**.

- Assume BPSK constellations, & i.i.d. complex Gaussian $\boldsymbol{H}_C$. SDR is proven to achieve the full rx diversity **[Jaldén-Ottersten'08]**.

- Approximation accuracies: **[So'09], [So'10]** showed that in both the MPSK and 4^q-QAM scenarios, the SDR detector can produce a constant factor approximate solution to the ML detection problem with exponentially high probability if the SNR is sufficiently low. In other words, in the low SNR region, we have

$$\|\boldsymbol{y} - \boldsymbol{H}\hat{\boldsymbol{s}}\|^2 \leq O(1) \cdot \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}^\star\|^2,$$

  with very high probability, where

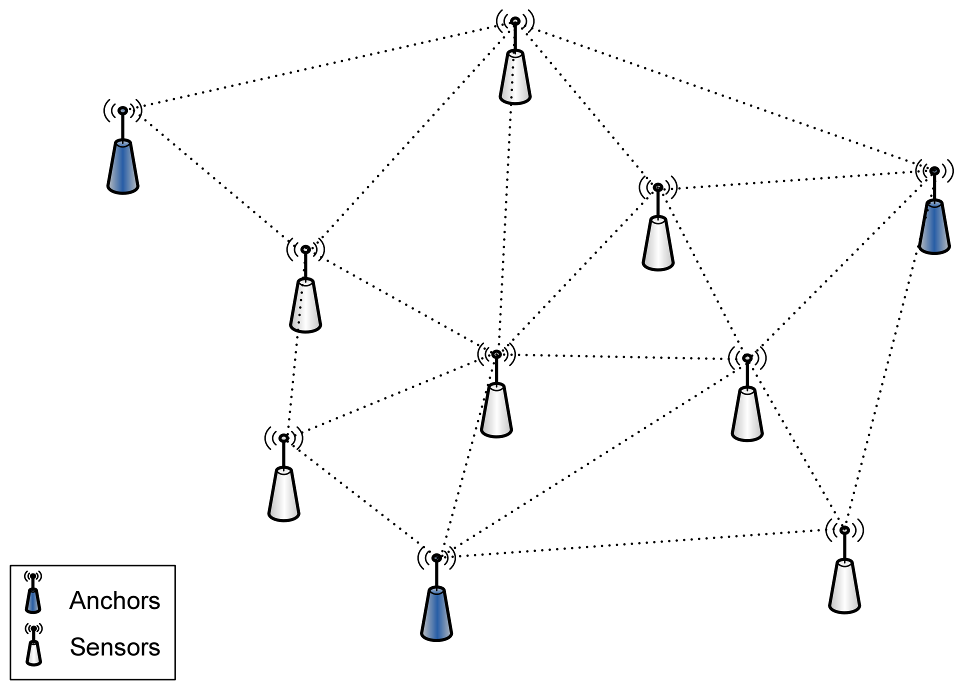
  - $\hat{\boldsymbol{s}}$ is the solution produced by SDR (with a suitable randomization procedure),
  - $\boldsymbol{s}^\star$ is the optimal ML solution.

Part III.B: Sensor Network Localization

Overview

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 $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$, $\boldsymbol{x}_i \in \mathbb{R}^2$, be the collection of all (unknown) sensor coordinates.

- Let $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$, $\boldsymbol{a}_i \in \mathbb{R}^2$, be the collection of all (known) anchor coordinates.

- The distance between sensor i and sensor j is

$$d_{ij} = \sqrt{(x_{i,1} - x_{j,1})^2 + (x_{i,2} - x_{j,2})^2} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|.$$

  Likewise, the distance between sensor i and anchor j is

$$\bar{d}_{ij} = \|\boldsymbol{x}_i - \boldsymbol{a}_j\|.$$

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

- The SNL problem here is that of finding $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ such that

$$\begin{aligned} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 &= d_{ij}^2, \quad (i,j) \in E_{ss}, \\ \|\boldsymbol{x}_i - \boldsymbol{a}_j\|^2 &= \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \end{aligned}$$

  where E_{ss} & E_{sa} are the sensor-sensor & sensor-anchor edge sets, resp.

Deriving an SDR of the SNL Problem: A First Attempt

- Let $\boldsymbol{X} = [\ \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\ ] \in \mathbb{R}^{2\times n}$. The SNL problem can be formulated as

$$\begin{aligned} \text{find} \quad & \boldsymbol{X} \in \mathbb{R}^{2\times n} \\ \text{s.t.} \quad & \boldsymbol{x}_i^T\boldsymbol{x}_i - 2\boldsymbol{x}_i^T\boldsymbol{x}_j + \boldsymbol{x}_j^T\boldsymbol{x}_j = d_{ij}^2, \quad (i,j) \in E_{ss}, \\ & \boldsymbol{x}_i^T\boldsymbol{x}_i - 2\boldsymbol{x}_i^T\boldsymbol{a}_j + \boldsymbol{a}_j^T\boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}. \end{aligned}$$

  This follows since

$$\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 = (\boldsymbol{x}_i - \boldsymbol{x}_j)^T(\boldsymbol{x}_i - \boldsymbol{x}_j) = \boldsymbol{x}_i^T\boldsymbol{x}_i - 2\boldsymbol{x}_i^T\boldsymbol{x}_j + \boldsymbol{x}_j^T\boldsymbol{x}_j$$

  and similarly for $\|\boldsymbol{x}_i - \boldsymbol{a}_j\|^2$.

- By letting $\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X} \in \mathbb{R}^{n\times n}$, we can also formulate the SNL problem as follows:

$$\begin{aligned} \text{find} \quad & \boldsymbol{X} \in \mathbb{R}^{2\times n}, \boldsymbol{Y} \in \mathbb{R}^{n\times n} \\ \text{s.t.} \quad & Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss}, \\ & Y_{ii} - 2\boldsymbol{x}_i^T\boldsymbol{a}_j + \boldsymbol{a}_j^T\boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \\ & \boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}. \end{aligned} \qquad \text{(SNL)}$$

- It is known **[Saxe'79]** that finding a solution to (SNL) is NP-hard.

- Observe that with $\boldsymbol{X} \in \mathbb{R}^{2\times n}$, the constraint $\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}$ is equivalent to

$$\boldsymbol{Y} \succeq \mathbf{0}, \quad \mathrm{rank}(\boldsymbol{Y}) \leq 2.$$

- If we proceed as before and just drop the trouble-causing rank constraint, then we get the following SDR:

$$\begin{array}{rl} \text{find} & \boldsymbol{X} \in \mathbb{R}^{2\times n}, \boldsymbol{Y} \in \mathbb{R}^{n\times n} \\ \text{s.t.} & Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss}, \\ & Y_{ii} - 2\boldsymbol{x}_i^T\boldsymbol{a}_j + \boldsymbol{a}_j^T\boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \\ & \boldsymbol{Y} \succeq \mathbf{0}. \end{array}$$

- In this formulation, there is no connection between $\boldsymbol{X}$ and $\boldsymbol{Y}$. In other words, the information in the original constraint $\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}$ is totally lost. The solution obtained could be quite awful.

Deriving an SDR of the SNL Problem: Another Attempt

- To keep the connection between $\boldsymbol{X}$ and $\boldsymbol{Y}$, instead of relaxing $\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}$ to $\boldsymbol{Y} \succeq \boldsymbol{0}$, we relax it to

$$\boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X}.$$

  – This is an SDP constraint, since by the Schur complement,

$$\boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X} \quad \Longleftrightarrow \quad \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \succeq \boldsymbol{0}.$$

- Then, we have the following SDR of the SNL problem:

$$\begin{array}{rl} \text{find} & \boldsymbol{X} \in \mathbb{R}^{2\times n}, \boldsymbol{Y} \in \mathbb{R}^{n\times n} \\ \text{s.t.} & Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss}, \\ & Y_{ii} - 2\boldsymbol{x}_i^T\boldsymbol{a}_j + \boldsymbol{a}_j^T\boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \\ & \boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X}. \end{array} \qquad \text{(SNL-SDR)}$$

Theoretical Properties of the SDR

- Suppose that we have a solution $(\boldsymbol{X}^\star, \boldsymbol{Y}^\star)$ to (SNL-SDR). Under what conditions will it be a solution to the original problem (SNL)?

- In **[So-Ye'07]**, a complete characterization is obtained.

- **[So-Ye'07]** Suppose that the given SNL instance is connected. Then, the following statements are equivalent:
  - The solution $(\boldsymbol{X}^\star, \boldsymbol{Y}^\star)$ to (SNL-SDR) is feasible for (SNL) (in particular, we have $\boldsymbol{Y}^\star = \boldsymbol{X}^{\star T}\boldsymbol{X}^\star$).
  - The max-rank solution to (SNL-SDR) has rank 2.
  - The given SNL instance is uniquely localizable, i.e., it has a unique solution in all dimensions.

- Since most polynomial-time interior-point algorithms for solving SDPs will return a solution that has the highest rank, we can localize uniquely localizable instances in polynomial time.

- The above result fits the theme of compressed sensing and low-rank optimization, which are two currently very active research areas.

Rank of SDR Solution and Dimension Reduction

- In **[So-Ye'07]** it was also shown that if the solution $\boldsymbol{Y}^\star$ to (SNL-SDR) has rank d, then one can construct d-dimensional coordinates for the sensors so that the distance constraints are satisfied.

- Question: While it is NP-hard to find a rank-2 solution $\boldsymbol{Y}^\star$ to (SNL-SDR), is it possible to find a low rank solution (and hence achieve dimension reduction)?

- One heuristic is to "stretch apart" pairs of non-adjacent nodes. This will tend to flatten the configuration of nodes.

- Mathematically, this corresponds to adding an objective function to (SNL):

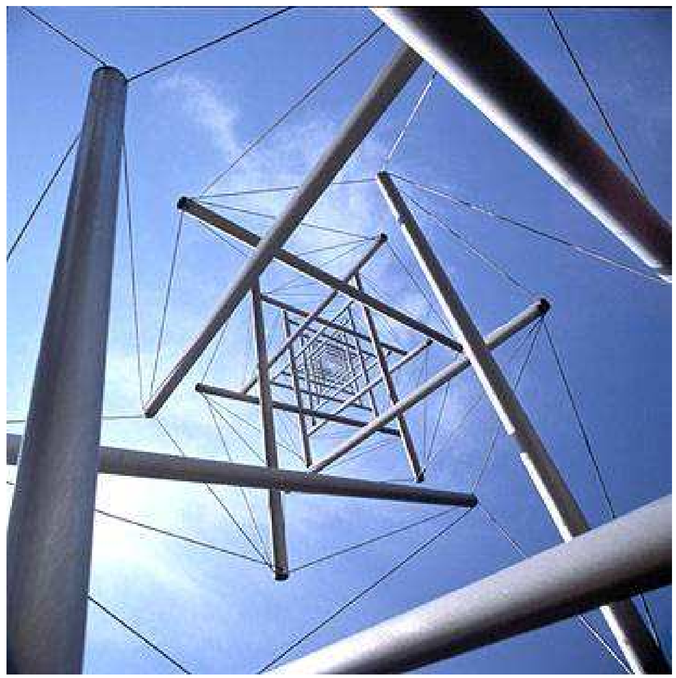
$$
\begin{aligned}
\max_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n} \quad & \sum_{(i,j)\in N_{ss}} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 \\
\text{s.t.} \quad & \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 = d_{ij}^2, \quad (i,j) \in E_{ss}, \\
& \|\boldsymbol{x}_i - \boldsymbol{a}_j\|^2 = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa},
\end{aligned}
\tag{SNL-OBJ}
$$

where $N_{ss} \subset \{(i,j) : (i,j) \notin E_{ss}\}$ is a subset of the non-adjacent pairs.

- Again, we can apply SDR to (SNL-OBJ).

- Interestingly, the solution to the resulting SDR often has low rank.

- In **[So-Ye'06]**, some theoretical justification is given to explain this phenomenon. It is related to the so-called tensegrity theory in discrete geometry.



- If distortion on the distances is allowed, then one can achieve dimension reduction using the SYZ theorem **[So-Ye-Zhang'08]** (see the procedure in Box 4).

Speeding Up the Computation

- When the number of sensors/edges is large, solving (SNL-SDR) could take a long time. The bottleneck comes not only from the large number of constraints, but also the large $((n+2) \times (n+2))$ positive semidefinite (PSD) constraint

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \succeq \boldsymbol{0}.$$

- Complexity-reduced implementations, at the cost of some SNL performance, have recently received attention in large-scale sensor network applications.

Speeding Up the Computation: The Edge-Based SDR

- To circumvent the large PSD constraint, one approach is to first observe that each edge $(i,j) \in E_{ss}$ is responsible for the following constraints in (SNL):

$$
\begin{aligned}
& Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \\
& Y_{ii} = \boldsymbol{x}_i^T \boldsymbol{x}_i, \quad Y_{ij} = \boldsymbol{x}_i^T \boldsymbol{x}_j, \quad Y_{jj} = \boldsymbol{x}_j^T \boldsymbol{x}_j.
\end{aligned}
\tag{$\dagger$}
$$

- Now, we can treat the constraints in ($\dagger$) as a group and relax them using our previous technique, i.e.,

$$
\mathbb{R}^{4\times 4} \ni \boldsymbol{Z}_{ij} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{x}_i & \boldsymbol{x}_j \\ \boldsymbol{x}_i^T & Y_{ii} & Y_{ij} \\ \boldsymbol{x}_j^T & Y_{ij} & Y_{jj} \end{bmatrix} \succeq \boldsymbol{0}.
$$

- This approach results in the following so-called edge-based SDR of the SNL problem, which was presented in **[Wang-Zheng-Ye-Boyd'08]**:

$$\begin{array}{ll} \text{find} & \boldsymbol{X} \in \mathbb{R}^{2\times n}, \boldsymbol{Y} \in \mathbb{R}^{n\times n} \\ \text{s.t.} & Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss}, \\ & Y_{ii} - 2\boldsymbol{x}_i^T \boldsymbol{a}_j + \boldsymbol{a}_j^T \boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \\ & \boldsymbol{Z}_{ij} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{x}_i & \boldsymbol{x}_j \\ \boldsymbol{x}_i^T & Y_{ii} & Y_{ij} \\ \boldsymbol{x}_j^T & Y_{ij} & Y_{jj} \end{bmatrix} \succeq \boldsymbol{0}, \quad (i,j) \in E_{ss}. \end{array} \tag{SNL-ESDR}$$

- Note that (SNL-ESDR) has $|E_{ss}|$ 4×4 PSD constraints, instead of one $(n+2) \times (n+2)$ PSD constraint in (SNL-SDR). The smaller dimension (i.e., 4) of the PSD constraints in (SNL-ESDR) is computationally easier to handle, thus allowing a speedup in computation.

- However, it should be noted that (SNL-SDR) is a tighter relaxation than (SNL-ESDR). Indeed, each $\boldsymbol{Z}_{ij}$ is a principal submatrix of $\boldsymbol{Z}$, and every principal submatrix of a PSD matrix must also be PSD.

Extensions of the Basic SDR

- So far we have only considered the noiseless version of the SNL problem.
- In general, the distance measurements $\{d_{ij}\}$ and $\{\bar{d}_{ij}\}$ could be corrupted. A commonly used error model is

$$\begin{aligned} d_{ij} &= d_{ij}^{nom} + g_{ij}, \\ \bar{d}_{ij} &= \bar{d}_{ij}^{nom} + \bar{g}_{ij}, \end{aligned}$$

  where $\{d_{ij}^{nom}\}$ and $\{\bar{d}_{ij}^{nom}\}$ are the nominal measurements, and $\{g_{ij}\}$ (resp. $\{\bar{g}_{ij}\}$) are i.i.d. Gaussian random variables with mean 0 and variance σ_{ij}^2 (resp. $\bar{\sigma}_{ij}^2$).
- In **[Biswas-Liang-Wang-Ye'06]**, the following maximum-likelihood (ML) SNL formulation is considered:

$$\min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n} \quad \sum_{(i,j)\in E_{ss}} \frac{1}{\sigma_{ij}^2}\left(\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - d_{ij}\right)^2 + \sum_{(i,j)\in E_{sa}} \frac{1}{\bar{\sigma}_{ij}^2}\left(\|\boldsymbol{x}_i - \boldsymbol{a}_j\| - \bar{d}_{ij}\right)^2 . \qquad \text{(ML-SNL)}$$

- As shown in **[Biswas-Liang-Wang-Ye'06]**, SDR can be employed to tackle the nonconvex problem (ML-SNL).

- The key lies in constructing suitable linearizations of the expressions

$$(\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - d_{ij})^2 \quad \text{and} \quad \left(\|\boldsymbol{x}_i - \boldsymbol{a}_j\| - \bar{d}_{ij}\right)^2.$$

- Let us focus on the former. The strategy is to proceed "one level at a time". Let

$$\epsilon_{ij} = (\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - d_{ij})^2 = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 - 2d_{ij}\|\boldsymbol{x}_i - \boldsymbol{x}_j\| + d_{ij}^2. \qquad (\dagger)$$

  Upon defining

$$u_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|, \quad v_{ij} = \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2,$$

  we see that ($\dagger$) is equivalent to

$$v_{ij} - 2d_{ij}u_{ij} + d_{ij}^2 = \epsilon_{ij}, \quad v_{ij} = u_{ij}^2, \quad v_{ij} = Y_{ii} - 2Y_{ij} + Y_{jj}, \quad \boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}.$$

- Now, we can relax $v_{ij} = u_{ij}^2$ to $v_{ij} \geq u_{ij}^2$, and $\boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X}$ to $\boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X}$. Using the Schur complement, these are equivalent to the SDP constraints

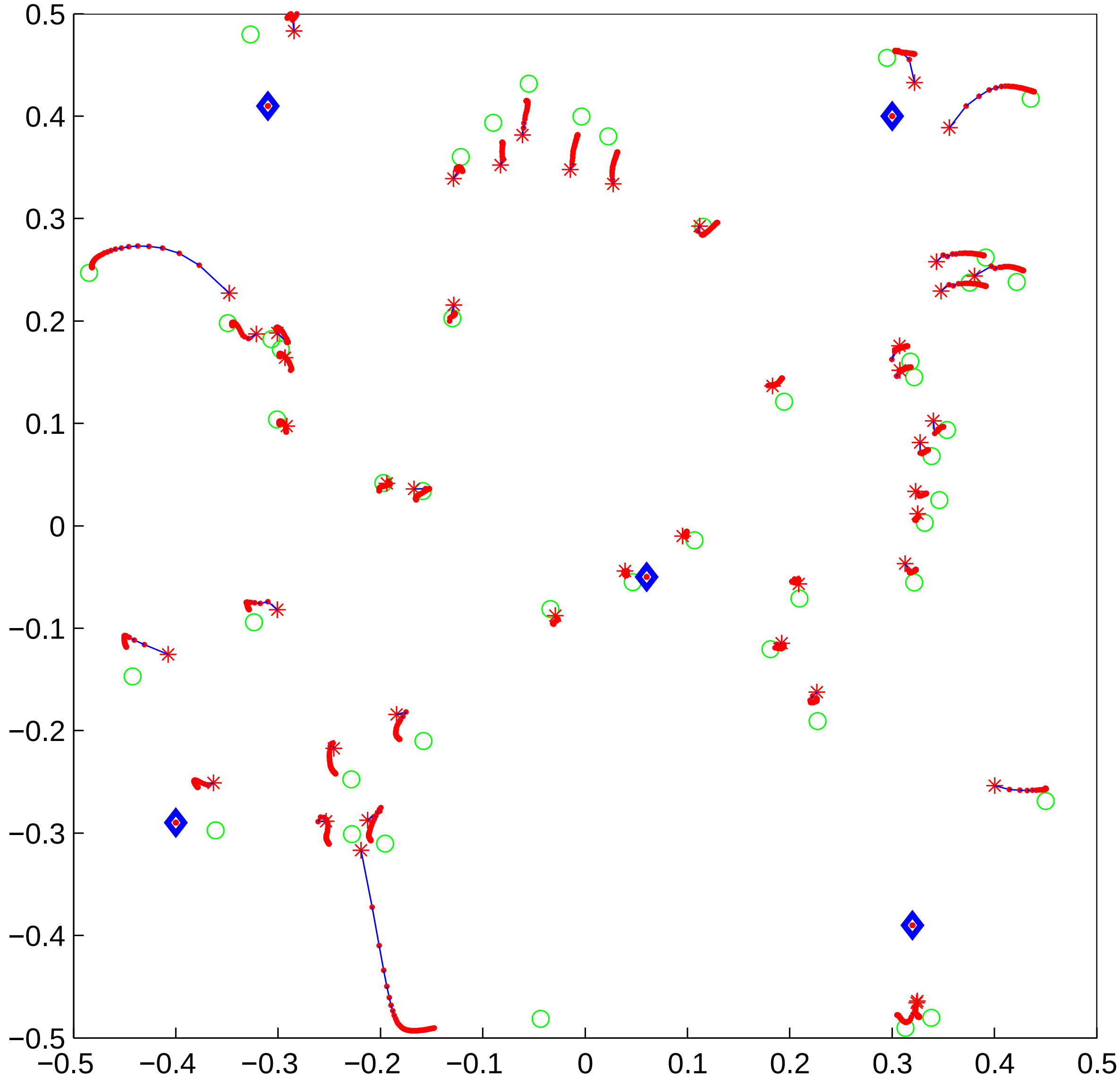
$$\boldsymbol{U}_{ij} = \begin{bmatrix} 1 & u_{ij} \\ u_{ij} & v_{ij} \end{bmatrix} \succeq \boldsymbol{0}, \quad \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \succeq \boldsymbol{0}.$$

- Hence, we obtain the following SDR of (ML-SNL):

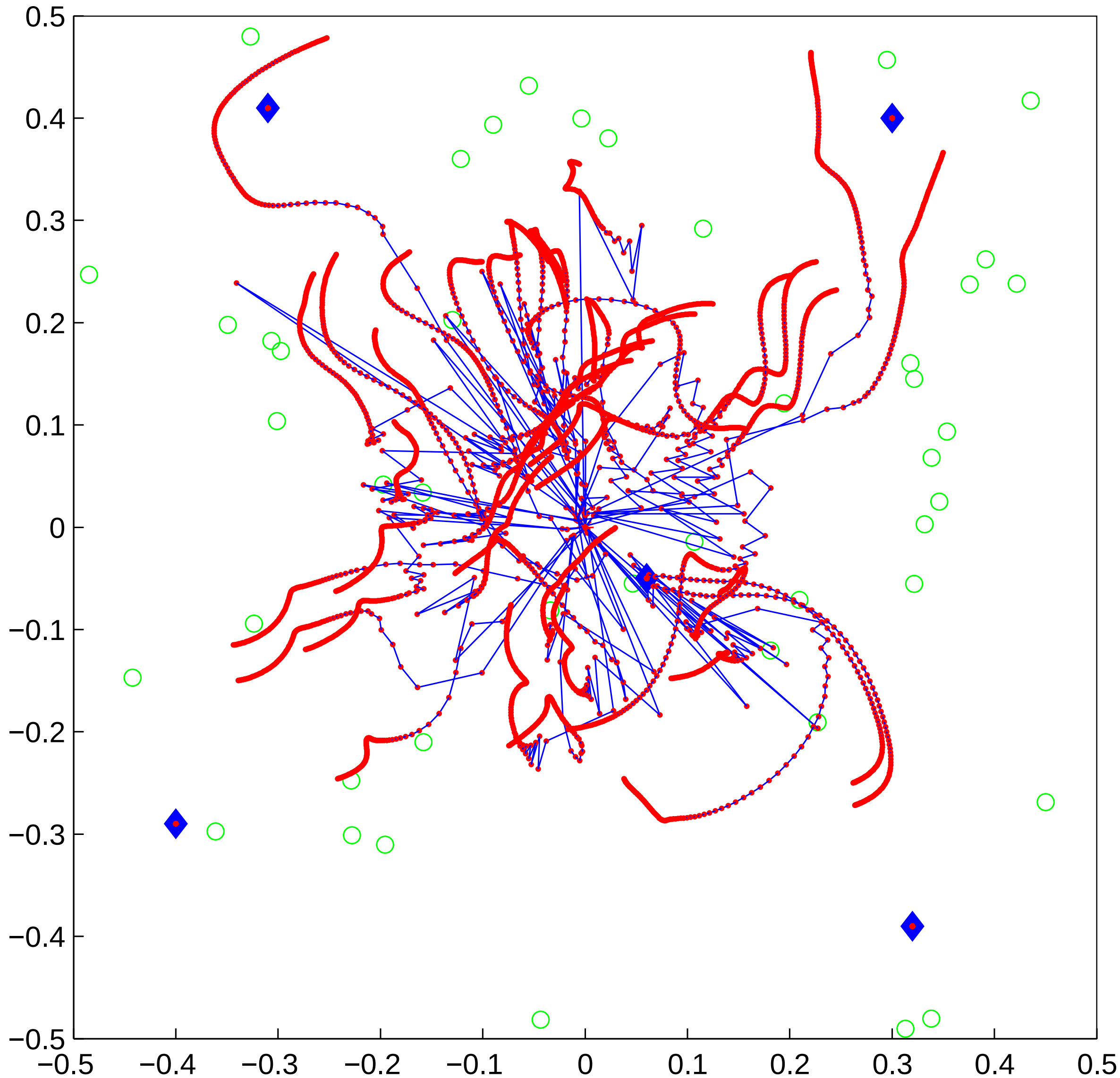
$$\begin{array}{ll} \displaystyle\min_{\boldsymbol{X},\boldsymbol{Y},\{\boldsymbol{U}_{ij}\},\{\bar{\boldsymbol{U}}_{ij}\}} & \displaystyle\sum_{(i,j)\in E_{ss}} \frac{1}{\sigma_{ij}^2}\epsilon_{ij} + \sum_{(i,j)\in E_{sa}} \frac{1}{\bar{\sigma}_{ij}^2}\bar{\epsilon}_{ij} \\ \text{s.t.} & Y_{ii} - 2Y_{ij} + Y_{jj} = v_{ij}^2, \quad (i,j) \in E_{ss}, \\ & Y_{ii} - 2\boldsymbol{x}_i^T\boldsymbol{a}_j + \boldsymbol{a}_j^T\boldsymbol{a}_j = \bar{v}_{ij}^2, \quad (i,j) \in E_{sa}, \\ & v_{ij} - 2d_{ij}u_{ij} + d_{ij}^2 = \epsilon_{ij}, \quad (i,j) \in E_{ss}, \\ & \bar{v}_{ij} - 2\bar{d}_{ij}\bar{u}_{ij} + \bar{d}_{ij}^2 = \bar{\epsilon}_{ij}, \quad (i,j) \in E_{sa}, \\ & \boldsymbol{U}_{ij} \succeq \boldsymbol{0}, \quad (i,j) \in E_{ss}; \quad \bar{\boldsymbol{U}}_{ij} \succeq \boldsymbol{0}, \quad (i,j) \in E_{sa}; \quad \boldsymbol{Z} \succeq \boldsymbol{0}. \end{array}$$

(ML-SNL-SDR)

- Note that the solution to (ML-SNL-SDR) can be used as a starting iterate for a gradient descent procedure, which can further improve the solution quality.



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).

Extensions of the Basic SDR

- SDR can also be employed to handle ML-SNL formulations with uncertain anchor locations, and/or with uncertain propagation speed **[Lui-Ma-So-Chan'09]** (the latter happens in underground sensor networks).

- For uncertain anchor locations, one could adopt the following error model:

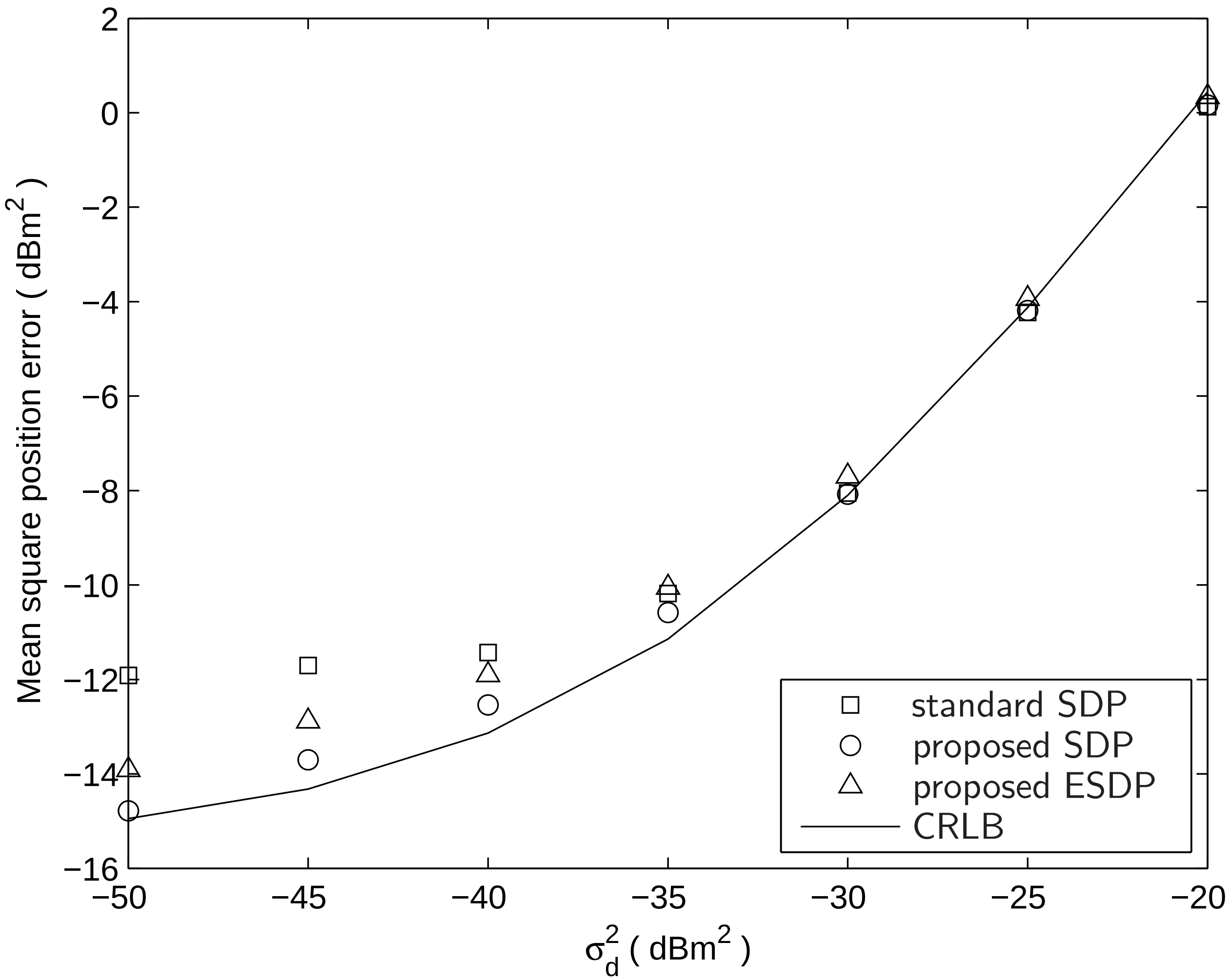
$$\boldsymbol{a}_i = \boldsymbol{a}_i^{nom} + \boldsymbol{z}_i,$$

  where $\boldsymbol{a}_i^{nom} \in \mathbb{R}^2$ is the nominal location of anchor i, and $\boldsymbol{z}_i \in \mathbb{R}^2$ is a Gaussian random vector with mean zero and covariance matrix $\boldsymbol{\Phi}_i$.

- Then, one has the following ML formulation (which is nonconvex):

$$\begin{aligned} \min_{\substack{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \\ \boldsymbol{a}_1,\ldots,\boldsymbol{a}_m}} \quad & \sum_{(i,j)\in E_{ss}} \frac{1}{\sigma_{ij}^2} \left(\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - d_{ij}\right)^2 + \sum_{(i,j)\in E_{sa}} \frac{1}{\bar{\sigma}_{ij}^2} \left(\|\boldsymbol{x}_i - \boldsymbol{a}_j\| - \bar{d}_{ij}\right)^2 \\ & + \sum_i (\boldsymbol{a}_i - \boldsymbol{a}_i^{nom})^T \boldsymbol{\Phi}_i^{-1} (\boldsymbol{a}_i - \boldsymbol{a}_i^{nom}). \end{aligned}$$

- This can be handled using the previously introduced SDR techniques.



Mean square position error performance versus noise power in the presence of anchor position uncertainty. Standard SDP— SDR without anchor position uncertainty; Proposed SDP— SDR with anchor position uncertainty; Proposed ESDP— Edge-based SDR with anchor position uncertainty. Details available in **[Lui-Ma-So-Chan'09]**.

Part III.C: Transmit Beamforming

Overview

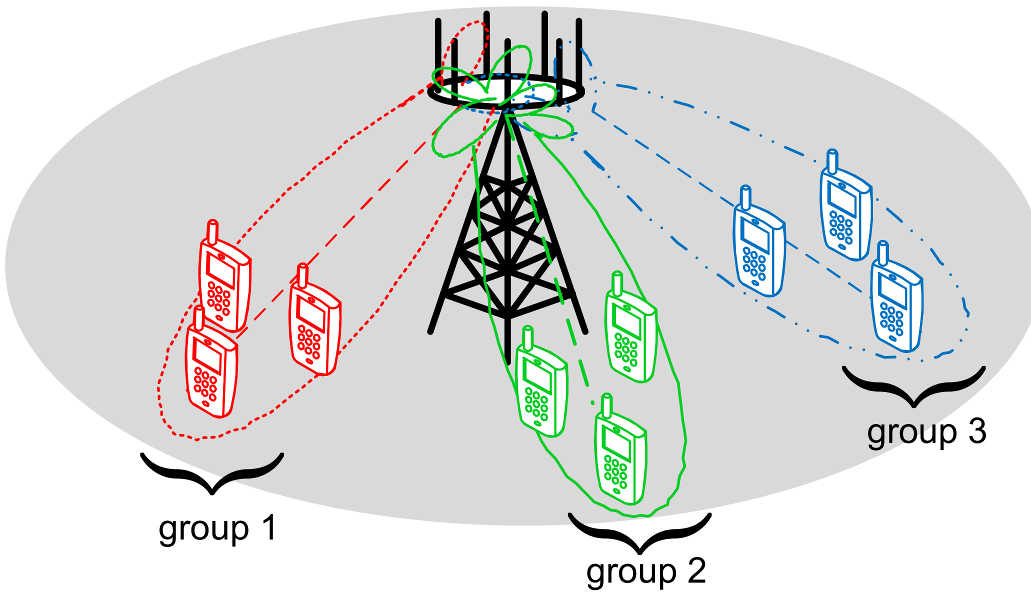
- Tx. beamforming (BF) has received much attention, with numerous formulations, application scenarios, solution approaches, their combinations,...

- Our focus:
  - A QCQP-SDR perspective on various tx. BF problems;
  - robust solutions under imperfect CSIT.

- What we will not go through:
  - alternative solution approaches:
    * second-order cone program (SOCP) (for unicast BF with instant. CSIT only) **[Wiesel-Eldar-Shamai'06]**;
    * uplink-downlink duality (for unicast BF only) **[Schubert-Boche'04]**;
  - alternative design formulations:
    * max-min-fairness (can be handled by SDR);
    * user admission;
    * proportional fairness, sum rate max., ...

Multi-Group Multicast Beamforming

- A natural generalization of unicast and multicast BF.
- **Scenario:** multiuser MISO downlink with M groups of users, & with each group receiving the same info. **[Karipidis-Sidiropoulos-Luo'08]**



- Transmit signal:

$$\boldsymbol{x}(t) = \sum_{m=1}^{M} \boldsymbol{w}_m s_m(t)$$

where $s_m(t) \in \mathbb{C}$ is the data stream for group m, & $\boldsymbol{w}_m \in \mathbb{C}^{N_t}$ its beamvector.

- Received signal of user k in the mth group:

$$\begin{aligned} y_{m,k}(t) &= \boldsymbol{h}_{m,k}^H \boldsymbol{x}(t) + v_{m,k}(t) \\ &= \boldsymbol{h}_{m,k}^H \boldsymbol{w}_m s_m(t) + \underbrace{\sum_{l \neq m} \boldsymbol{h}_{m,k}^H \boldsymbol{w}_l s_l(t)}_{\text{inter-group interference}} + v_{m,k}(t), \end{aligned}$$

where $k = 1, \ldots, K_m$, & K_m is the number of users in the mth group.

- SINR:

$$\mathsf{SINR}_{m,k} = \frac{\boldsymbol{w}_m^H \boldsymbol{R}_{m,k} \boldsymbol{w}_m}{\sum_{l \neq m} \boldsymbol{w}_l^H \boldsymbol{R}_{m,k} \boldsymbol{w}_l + \sigma_{m,k}^2}$$

where $\boldsymbol{R}_{m,k} = \boldsymbol{h}_{m,k} \boldsymbol{h}_{m,k}^H$ for instant. CSIT, & $\boldsymbol{R}_{m,k} = \mathrm{E}\{\boldsymbol{h}_{m,k} \boldsymbol{h}_{m,k}^H\}$ for stat. CSIT.

- Problem: a QoS-assured, total power minimizing, design

$$\min_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_M \in \mathbb{C}^{N_t}} \sum_{m=1}^{M} \|\boldsymbol{w}_m\|^2$$

$$\text{s.t. } \mathsf{SINR}_{m,k} = \frac{\boldsymbol{w}_m^H \boldsymbol{R}_{m,k} \boldsymbol{w}_m}{\sum_{l \neq m} \boldsymbol{w}_l^H \boldsymbol{R}_{m,k} \boldsymbol{w}_l + \sigma_{m,k}^2} \geq \gamma_{m,k}, \quad \begin{array}{l} k = 1, \ldots, K_m, \\ m = 1, \ldots, M, \end{array}$$

where $\gamma_{m,k}$'s are prescribed SINR requirements.

- A separable QCQP with M variables, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_M$, and $\sum_{m=1}^{M} K_m$ constraints.

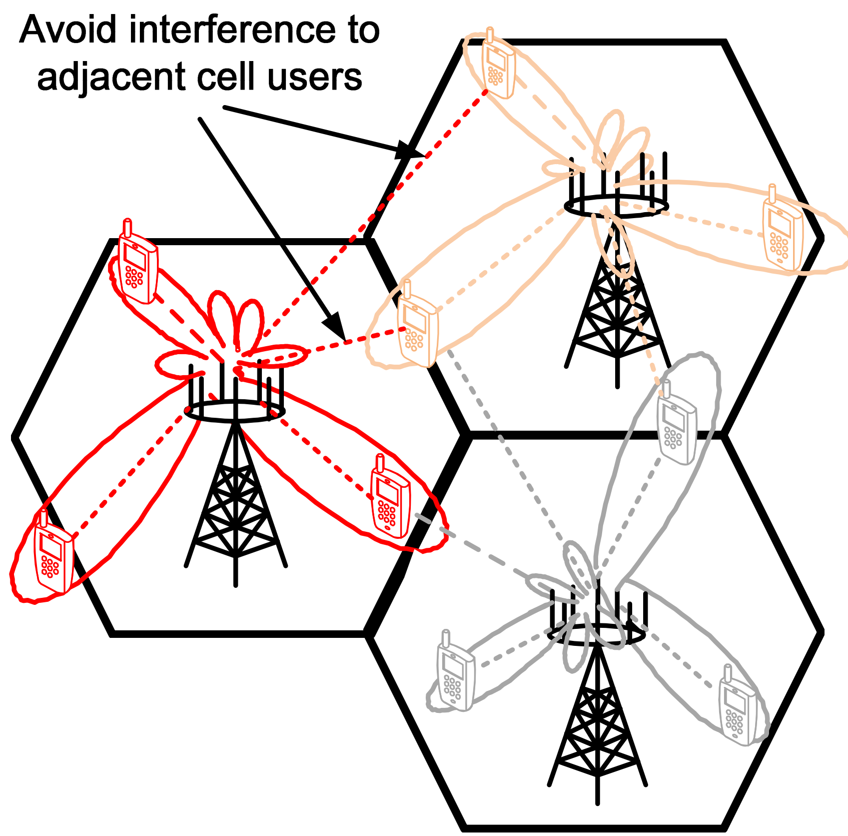
- By the SBP rank result, the rank-one optimality of SDR is assured when:
  - $K_1 \leq 3$, $K_m = 1 \; \forall m \neq 1$ (one group serving ≤ 3 users, the others 1 user);
  - $K_1 \leq 2$, $K_2 \leq 2$, $K_m = 1 \; \forall m \neq 1, 2$ (two groups serving ≤ 2 users, the others 1 user).

- requires a more sophisticated randomization procedure **[Karipidis-Sidiropoulos-Luo'08]**.

Multi-Cell Coordinated Beamforming

- Motivation: provide better interference management by coordinating the transmissions of base stations at different cells.

- **Scenario:** Unicast transmit beamforming in a multi-cell scale **[Dahrouj-Yu'10]**, **[Bengtsson-Ottersten'01]**



- Transmit signal of ith cell:

$$\boldsymbol{x}_i(t) = \sum_{j=1}^{K_i} \boldsymbol{w}_{i,j} s_{i,j}(t), \quad i = 1, \ldots, N$$

where $s_{i,j}(t)$ / $\boldsymbol{w}_{i,j}$ is the tx. stream/ beamvector for user j in ith cell, resp.; K_i is the no. of users in ith cell; N is the no. of cells.

- Received signal of user j in the ith cell:

$$y_{i,j}(t) = \boldsymbol{h}_{i,i,j}^H \boldsymbol{x}_i(t) + \sum_{m \neq i} \boldsymbol{h}_{m,i,j}^H \boldsymbol{x}_m(t) + v_{i,j}(t), \quad j = 1, \ldots, K_i$$

where $\boldsymbol{h}_{m,i,j}$ is the channel from mth cell to user j in the ith cell.

- Define CSIT $\boldsymbol{R}_{m,i,j}$ in the same way as before. SINR:

$$\mathsf{SINR}_{i,j} = \frac{\boldsymbol{w}_{i,j}^H \boldsymbol{R}_{i,i,j} \boldsymbol{w}_{i,j}}{\underbrace{\sum_{l \neq j} \boldsymbol{w}_{i,l}^H \boldsymbol{R}_{i,i,j} \boldsymbol{w}_{i,l}}_{\text{intra-cell interference}} + \underbrace{\sum_{m \neq i} \sum_{n} \boldsymbol{w}_{m,n}^H \boldsymbol{R}_{m,i,j} \boldsymbol{w}_{m,n}}_{\text{inter-cell interference}} + \sigma_{i,j}^2}$$

- Design problem:

$$\begin{aligned}
\min_{\{\boldsymbol{w}_{i,j}\}} \ & \sum_{i,j} \|\boldsymbol{w}_{i,j}\|^2 \\
\text{s.t.} \ & \frac{\boldsymbol{w}_{i,j}^H \boldsymbol{R}_{i,i,j} \boldsymbol{w}_{i,j}}{\sum_{l \neq j} \boldsymbol{w}_{i,l}^H \boldsymbol{R}_{i,i,j} \boldsymbol{w}_{i,l} + \sum_{m \neq i,n} \boldsymbol{w}_{m,n}^H \boldsymbol{R}_{m,i,j} \boldsymbol{w}_{m,n} + \sigma_{i,j}^2} \geq \gamma_{i,j}, \\
& j = 1, \ldots, K_i, \ i = 1, \ldots, N
\end{aligned}$$

- A QCQP with $\sum_{i=1}^N K_i$ variables & $\sum_{i=1}^N K_i$ constraints (despite its complicated appearance).

- SDR solves the multi-cell unicast problem optimally, by the SBP rank result.

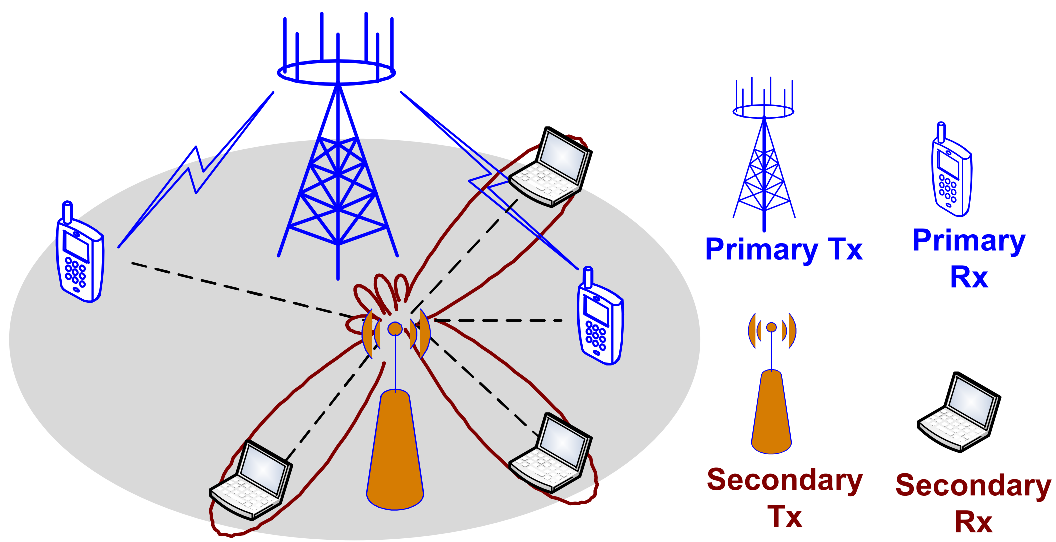
- Practical issue: distributed optimization algorithms.

Cognitive Radio (CR) Beamforming

- 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 lth PU:
    $$|\boldsymbol{g}_l^H \boldsymbol{w}|^2$$
    where $\boldsymbol{g}_l$ is the channel from the secondary transmitter to the lth PU.

- Design problem **[Phan-Vorobyov-Sidiropoulos-Tellambura'09]**:

$$\begin{aligned} \min_{\boldsymbol{w}} \ & \|\boldsymbol{w}\|^2 \\ \text{s.t.} \ & \mathsf{SNR}_{\mathsf{SU},i} = \boldsymbol{w}^H \boldsymbol{R}_k \boldsymbol{w} / \sigma_k^2 \geq \gamma_k, \ k = 1, \ldots, K, \\ & \boldsymbol{w}^H \boldsymbol{G}_l \boldsymbol{w} \leq \delta_l, \ l = 1, \ldots, L \quad \text{(interference temperature (IT) constraints)} \end{aligned}$$

  where $\boldsymbol{G}_l$ is the CSIT of lth PU (defined in the same way as $\boldsymbol{R}_k$), δ_l is the tolerable interference level to l PU, & γ_k are SUs' SNR requirements.

- By the SBP rank result, SDR is optimal when $K \leq 2$, $L = 1$ (≤ 2 SUs, 1 PU).

- CR BF design for the unicast case (see, e.g., **[Zhang-Liang-Cui'10]**):

$$
\begin{aligned}
\min_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K} \quad & \sum_{k=1}^{K} \|\boldsymbol{w}_k\|^2 \\
\text{s.t.} \quad & \mathsf{SINR}_{\mathsf{SU},i} = \frac{\boldsymbol{w}_k^H \boldsymbol{R}_k \boldsymbol{w}_k}{\sum_{l \neq k} \boldsymbol{w}_l^H \boldsymbol{R}_k \boldsymbol{w}_l + \sigma_k^2} \geq \gamma_k, \;\; k = 1, \ldots, K, \\
& \textstyle\sum_{k=1}^{K} \boldsymbol{w}_k^H \boldsymbol{G}_l \boldsymbol{w}_k \leq \delta_l, \;\; l = 1, \ldots, L \quad \text{(IT constraints)}
\end{aligned}
$$

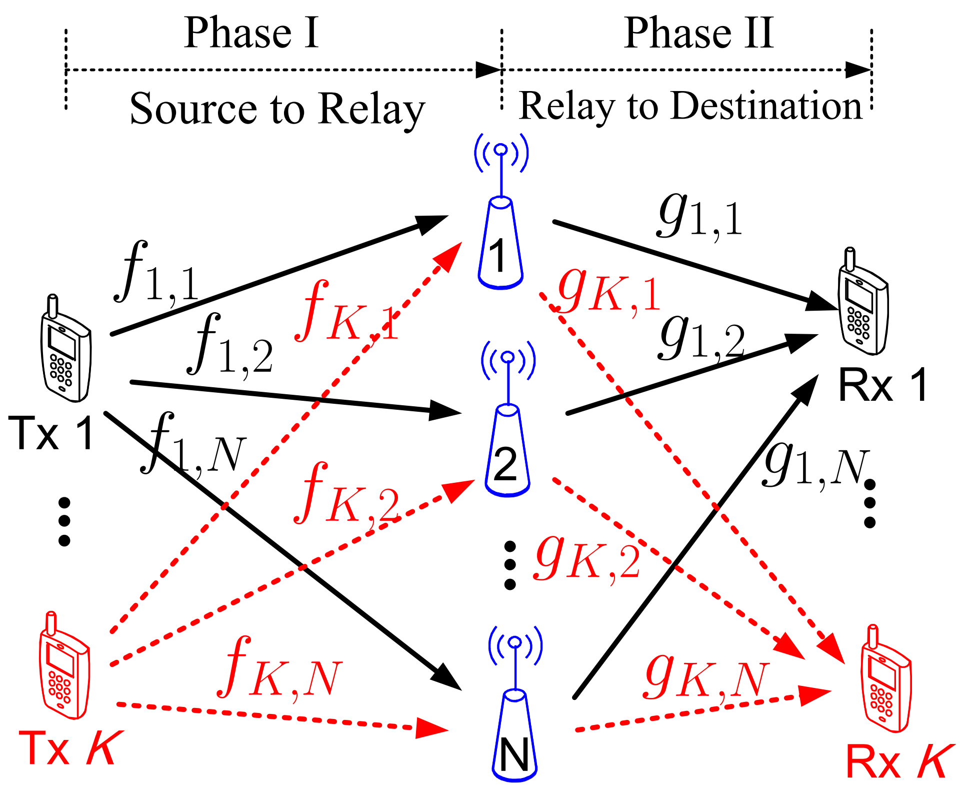
- A separable QCQP with K variables and $K + L$ constraints.

- By the SBP rank result, SDR solves the problem if $L \leq 2$.

- Remark: For instant. CSIT with SUs, SDR can be shown to be optimal for any L. Or it can be reformulated, and then solved, by SOCP.

One-Way Relay Network Beamforming

- **Scenario:** one-way cooperative communication by a network of N single-antenna amplify-forward (AF) relays, K tx-rx pairs **[Fazeli-Dehkordy-Shahbazpanahi-Gazor'09]**.



- Goal: Design the AF weights so that the SINR requirements are met, and the total relay tx. power is minimized.

- System model:
  - rx. signals for the source-to-relay link:

$$\boldsymbol{r}(t) = \sum_{i=1}^{K} \boldsymbol{f}_i s_i(t) + \boldsymbol{n}(t),$$

where $\boldsymbol{r}(t) = [\ r_1(t), \ldots, r_N(t)\ ]$, $r_i(t)$ being the rx. signal of relay i;
$s_i(t)$ is the data stream from source i to destination i;
$\boldsymbol{f}_i \in \mathbb{C}^N$ the channel from source i to the relays;
$\boldsymbol{n}(t)$ is noise with covariance $\boldsymbol{\Sigma}_n = \mathrm{Diag}(\sigma_{n,1}^2, \ldots, \sigma_{n,N}^2)$.

  - AF process:

$$\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{r}(t),$$

where $\boldsymbol{W} = \mathrm{Diag}(w_1, \ldots, w_N)$, w_i is the AF weight at relay i.

  - rx. signals for the relay-to-destination link:

$$y_i(t) = \boldsymbol{g}_i^H \boldsymbol{x}(t) + v_i(t), \qquad i = 1, \ldots, K,$$

where $\boldsymbol{g}_i$ is the channel from the relays to destination i; $v_i(t)$ is noise with variance $\sigma_{v,i}^2$.

- Assume instant. CSIT (for ease of illustration).

$$\text{SINR}_i = \frac{|\boldsymbol{g}_i^H \boldsymbol{W} \boldsymbol{f}_i|^2}{\underbrace{\sum_{k \neq i} |\boldsymbol{g}_i^H \boldsymbol{W} \boldsymbol{f}_k|^2}_{\text{interference}} + \underbrace{\boldsymbol{g}_i^H \boldsymbol{W} \boldsymbol{\Sigma}_n \boldsymbol{W}^H \boldsymbol{g}_i}_{\text{noise amplification due to AF}} + \sigma_{v,i}^2}, \quad i = 1, \ldots, K.$$

- Let $\boldsymbol{w} = [\ w_1, \ldots, w_N\ ]^T \in \mathbb{C}^N$. The design problem:

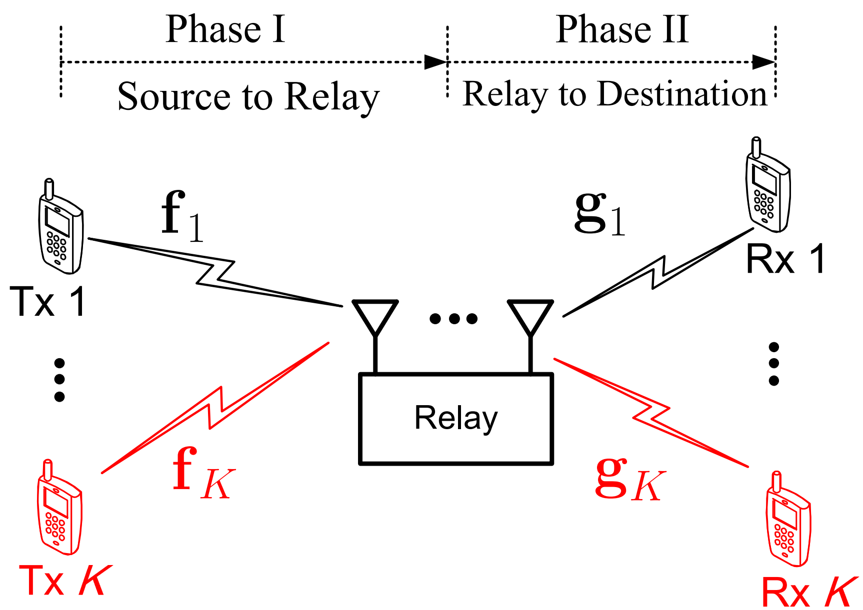
$$\begin{aligned} \min_{\boldsymbol{w}} \ & \mathrm{E}\{\|\boldsymbol{x}(t)\|^2\} = \boldsymbol{w}^H \boldsymbol{C} \boldsymbol{w} \\ \text{s.t. } & \text{SINR}_i = \frac{\boldsymbol{w}^H \boldsymbol{A}_i \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{B}_i \boldsymbol{w} + \sigma_{v,i}^2} \geq \gamma_k, \ k = 1, \ldots, K, \end{aligned}$$

where $\boldsymbol{A}_i = (\boldsymbol{f}_i^* \odot \boldsymbol{g}_i)(\boldsymbol{f}_i^* \odot \boldsymbol{g}_i)^H$, $\boldsymbol{B}_i = \sum_{k \neq i} (\boldsymbol{f}_k^* \odot \boldsymbol{g}_i)(\boldsymbol{f}_k^* \odot \boldsymbol{g}_i)^H + \mathrm{Diag}(|g_{i,1}|^2 \sigma_{n,1}^2, \ldots, |g_{i,N}|^2 \sigma_{n,N}^2)$, $\boldsymbol{C} = \mathrm{Diag}(\|\boldsymbol{f}_1\|^2 + \sigma_{n,1}^2, \ldots, \|\boldsymbol{f}_N\|^2 + \sigma_{n,N}^2)$.

- The problem is a QCQP with K constraints; SDR is optimal for $K \leq 3$.

One-Way MIMO Relay Beamforming

- **Scenario:** one-way relaying by an MIMO AF relay, K tx-rx pairs **[Chalise-Vandendorpe'09]**.



- Everything is the same as that in the last relay example, except that a matrix AF process is considered:

$$\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{r}(t),$$

where $\boldsymbol{W} \in \mathbb{C}^{N \times N}$ is a general $N \times N$ matrix (instead of being diagonal).

- Let $\boldsymbol{w} = \text{vec}(\boldsymbol{W}) \in \mathbb{C}^{N^2}$.

- The design problem (after some careful derivations):

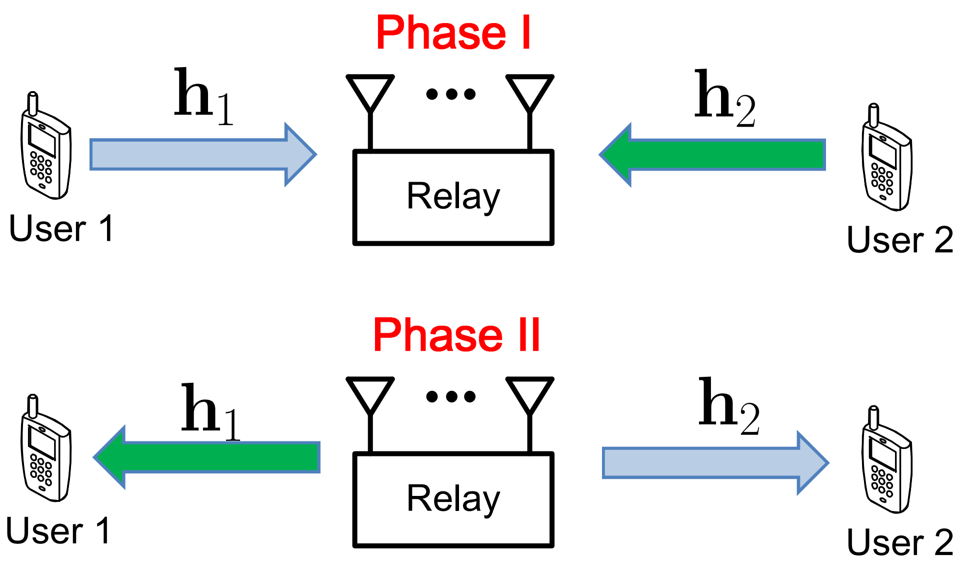
$$\begin{aligned} \min_{\boldsymbol{w}} \ & \text{E}\{\|\boldsymbol{x}(t)\|^2\} = \boldsymbol{w}^H \boldsymbol{C} \boldsymbol{w} \\ \text{s.t.} \ & \mathsf{SINR}_i = \frac{\boldsymbol{w}^H \boldsymbol{A}_i \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{B}_i \boldsymbol{w} + \sigma_{v,i}^2} \geq \gamma_k, \ \ k = 1, \ldots, K, \end{aligned}$$

  where $\boldsymbol{A}_i = (\boldsymbol{f}_i^* \otimes \boldsymbol{g}_i)(\boldsymbol{f}_i^* \otimes \boldsymbol{g}_i)^H$, $\boldsymbol{B}_i = \sum_{k \neq i} (\boldsymbol{f}_k^* \otimes \boldsymbol{g}_i)(\boldsymbol{f}_k^* \otimes \boldsymbol{g}_i)^H + \boldsymbol{\Sigma}_n^T \otimes (\boldsymbol{g}_i \boldsymbol{g}_i^H)$, & $\boldsymbol{C} = (\sum_{i=1}^K \boldsymbol{f}_i^* \boldsymbol{f}_i^T + \boldsymbol{\Sigma}_n^T) \otimes \boldsymbol{I}$.

- Again, optimality of SDR is assured for $K \leq 3$.

Two-Way Relay Beamforming

- **Scenario:** two-way communication between two users, using an MIMO AF relay **[Zhang-Liang-Chai-Cui'09]**



  – Phase I: two users transmit

$$\boldsymbol{r}(t) = \boldsymbol{h}_1 s_1(t) + \boldsymbol{h}_2 s_2(t) + \boldsymbol{n}(t).$$

  – Phase II: matrix AF relaying

$$\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{r}(t).$$

- In addition, the users can self cancel their previously tx. data.

$$y_1(t) = \boldsymbol{h}_1^H \boldsymbol{x}(t) + v_1(t) = \underbrace{\boldsymbol{h}_1^H \boldsymbol{W}\boldsymbol{h}_1 s_1(t)}_{\substack{\text{self interference,}\\ \text{cancelled}}} + \boldsymbol{h}_1^H \boldsymbol{W}\boldsymbol{h}_2 s_2(t) + \underbrace{\boldsymbol{h}_1^H \boldsymbol{W}\boldsymbol{n}(t)}_{\text{noise amp.}} + v_1(t),$$

$$y_2(t) = \boldsymbol{h}_2^H \boldsymbol{x}(t) + v_2(t) = \boldsymbol{h}_2^H \boldsymbol{W}\boldsymbol{h}_1 s_1(t) + \underbrace{\boldsymbol{h}_2^H \boldsymbol{W}\boldsymbol{h}_2 s_2(t)}_{\substack{\text{self interference,}\\ \text{cancelled}}} + \underbrace{\boldsymbol{h}_1^H \boldsymbol{W}\boldsymbol{n}(t)}_{\text{noise amp.}} + v_2(t).$$

- The design problem

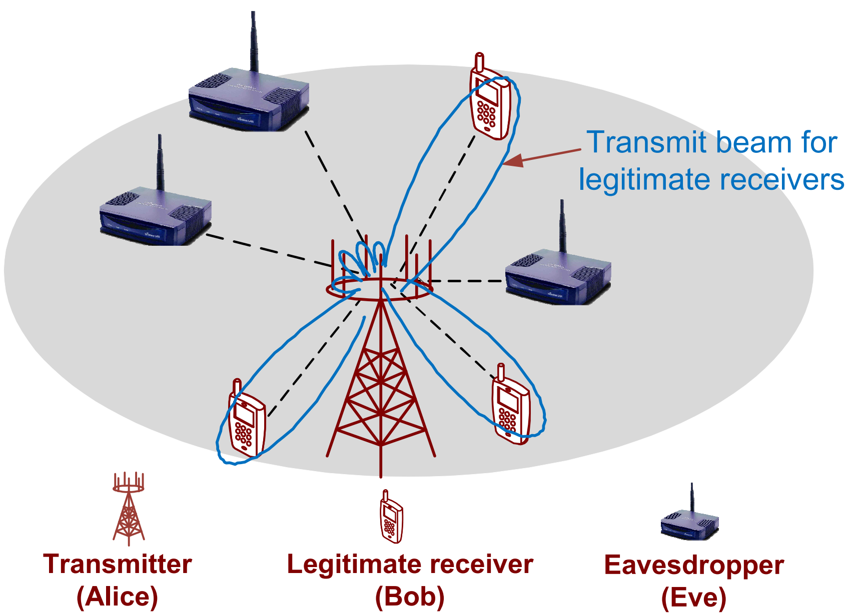
$$
\begin{aligned}
\min_{\boldsymbol{W}} \; & \mathrm{E}\{\|\boldsymbol{x}(t)\|^2\} \\
\text{s.t. } & \mathsf{SNR}_1 = \frac{|\boldsymbol{h}_1^H \boldsymbol{W} \boldsymbol{h}_2|^2}{\boldsymbol{h}_1^H \boldsymbol{W} \boldsymbol{\Sigma}_n \boldsymbol{W}^H \boldsymbol{h}_1 + \sigma_{v,1}^2} \geq \gamma_1, \\
& \mathsf{SNR}_1 = \frac{|\boldsymbol{h}_2^H \boldsymbol{W} \boldsymbol{h}_1|^2}{\boldsymbol{h}_2^H \boldsymbol{W} \boldsymbol{\Sigma}_n \boldsymbol{W}^H \boldsymbol{h}_2 + \sigma_{v,2}^2} \geq \gamma_2
\end{aligned}
$$

can be converted to a 2-constraints QCQP, by applying $\boldsymbol{w} = \mathrm{vec}(\boldsymbol{W}) \in \mathbb{C}^{N^2}$ (the same way as in the last example).

- SDR is optimal for two-way relaying.

Physical-Layer Secrecy

- A topic concerned with physical-layer techniques for delivering secure communication in the presence of eavesdroppers.
- A potentially powerful alternative, or complement, to the network-layer-based cryptographic encryption techniques.
- The idea has been there around the 70's in info. theory; e.g., **[Wyner'75]**.
- Receive growing attention recently, usually with MIMO.



A Simple Illustrating Example of Physical-Layer Secrecy

- We will call the transmitter Alice, the intended receiver Bob, & the eavesdroppers Eves, resp.

- Consider a one-Bob, one-Eve MISO scenario. Signal model:

$$y_{\text{Bob}}(t) = \boldsymbol{h}^H \boldsymbol{x}(t) + n(t), \qquad y_{\text{Eve}}(t) = \boldsymbol{g}^H \boldsymbol{x}(t) + v(t),$$

  where $\boldsymbol{x}(t) = \boldsymbol{w} s(t)$ is the tx signal (BF, as before); $\boldsymbol{h}$ / $\boldsymbol{g}$ is the Alice-to-Bob / Alice-to-Eve channel, resp.; $n(t)$ / $v(t)$ is noise at Bob / Eve, resp. Also, define

$$\text{SNR}_{\text{Bob}} = \frac{|\boldsymbol{h}^H \boldsymbol{w}|^2}{\sigma_n^2}, \qquad \text{SNR}_{\text{Eve}} = \frac{|\boldsymbol{g}^H \boldsymbol{w}|^2}{\sigma_v^2}.$$

- A secrecy rate maximization problem:

$$\begin{aligned} C = \max_{\boldsymbol{w} \in \mathbb{C}^{N_t}} \; & \log(1 + \text{SNR}_{\text{Bob}}) - \log(1 + \text{SNR}_{\text{Eve}}) \\ \text{s.t.} \; & \|\boldsymbol{w}\|^2 \leq P_{\text{max}}. \end{aligned}$$

  This secrecy rate is achievable— there exist codes such that Bob can reliably receive a message at C bps/Hz, while Eve can retrieve almost nothing.

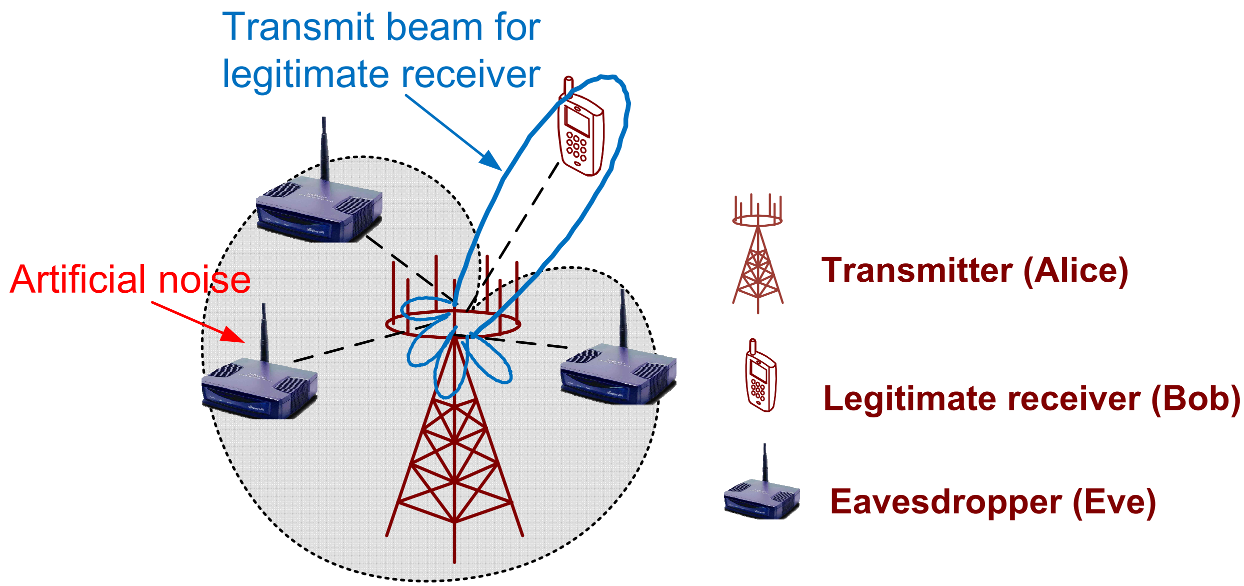
Artificial Noise (AN) for Physical-Layer Secrecy

- An intuitively insightful idea—jamming Eves by transmitting artificially generated noise **[Negi-Goel'05]**.

- Tx model:

$$\boldsymbol{x}(t) = \boldsymbol{w}s(t) + \boldsymbol{z}(t),$$
$$\boldsymbol{z}(t) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{z}(t)$ is AN, and $\boldsymbol{\Sigma}$ its covariance.



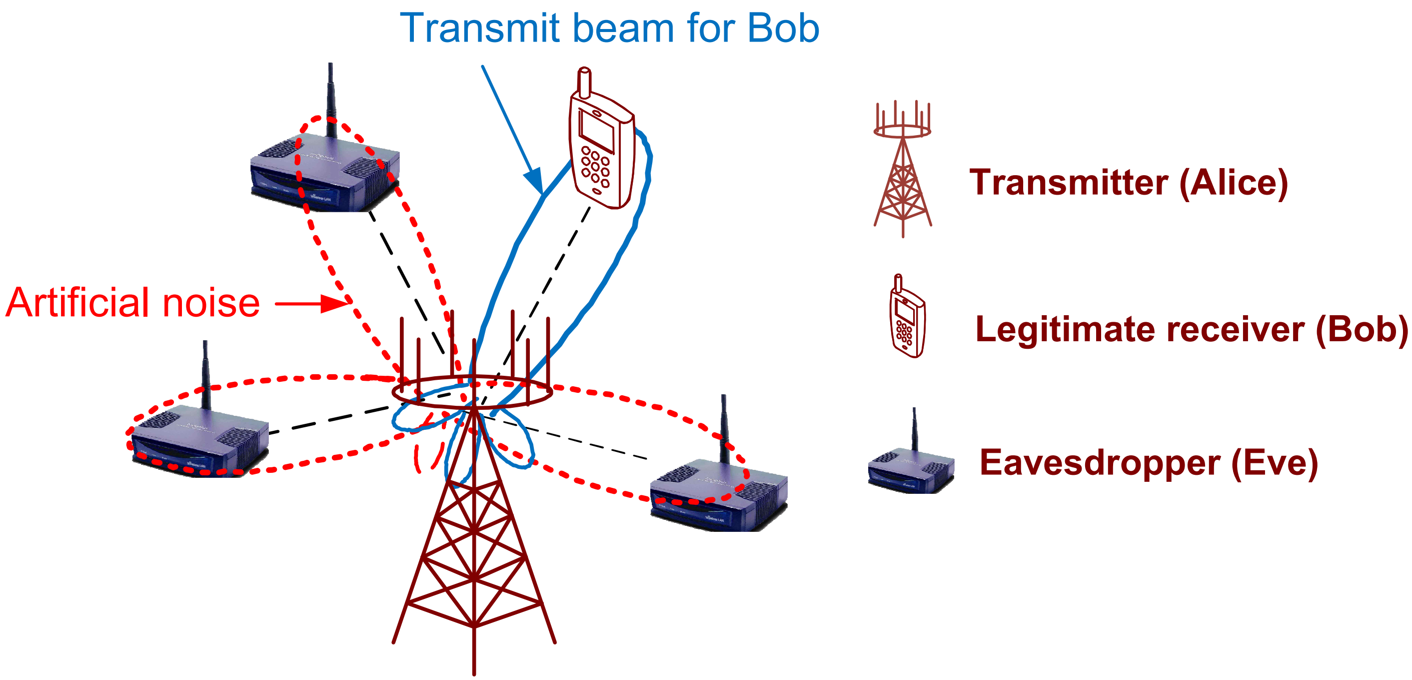
- A popular design is to make AN (almost) isotropic:

$$\boldsymbol{w} = \sqrt{\alpha P_{\mathsf{max}}}\boldsymbol{h}/\|\boldsymbol{h}\|, \qquad \boldsymbol{\Sigma} = (1-\alpha)P_{\mathsf{max}}(\boldsymbol{I} - \boldsymbol{h}\boldsymbol{h}^H/\|\boldsymbol{h}\|^2),$$

where P_{max} is the total tx power, $0 < \alpha \leq 1$ is a power allocation factor.

QoS-based BF Approach to MISO Channel Secrecy

- **Scenario:** One Bob, multiple Eves, all MISO **[Liao-Chang-Ma-Chi'11]**.



- AN can be spatially selective, jamming Eves in a more spatially focused fashion.

- The beamvector and AN covariance are jointly optimized.

- tx-rx model:

$$\boldsymbol{x}(t) = \boldsymbol{w}s(t) + \boldsymbol{z}(t), \quad \boldsymbol{z}(t) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{\Sigma}),$$

$$y_{\text{Bob}}(t) = \boldsymbol{h}^H \boldsymbol{x}(t) + n(t) = \boldsymbol{h}^H \boldsymbol{w} s(t) + \underbrace{\boldsymbol{h}^H \boldsymbol{z}(t)}_{\text{AN interference}} + n(t),$$

$$y_{\text{Eve},i}(t) = \boldsymbol{g}_i^H \boldsymbol{x}(t) + v_i(t) = \boldsymbol{g}_i^H \boldsymbol{w} s(t) + \underbrace{\boldsymbol{g}_i^H \boldsymbol{z}(t)}_{\text{AN interference}} + v_i(t), \; i = 1, \ldots, L.$$

- SINRs:

$$\text{SINR}_{\text{Bob}} = \frac{|\boldsymbol{h}^H \boldsymbol{w}|^2}{\text{Tr}(\boldsymbol{\Sigma} \boldsymbol{h} \boldsymbol{h}^H) + \sigma_n^2}, \quad \text{SINR}_{\text{Eve},i} = \frac{|\boldsymbol{g}_i^H \boldsymbol{w}|^2}{\text{Tr}(\boldsymbol{\Sigma} \boldsymbol{g}_i \boldsymbol{g}_i^H) + \sigma_{v,i}^2}, \; i = 1, \ldots, L.$$

- Design: given a specification (γ, β), solve

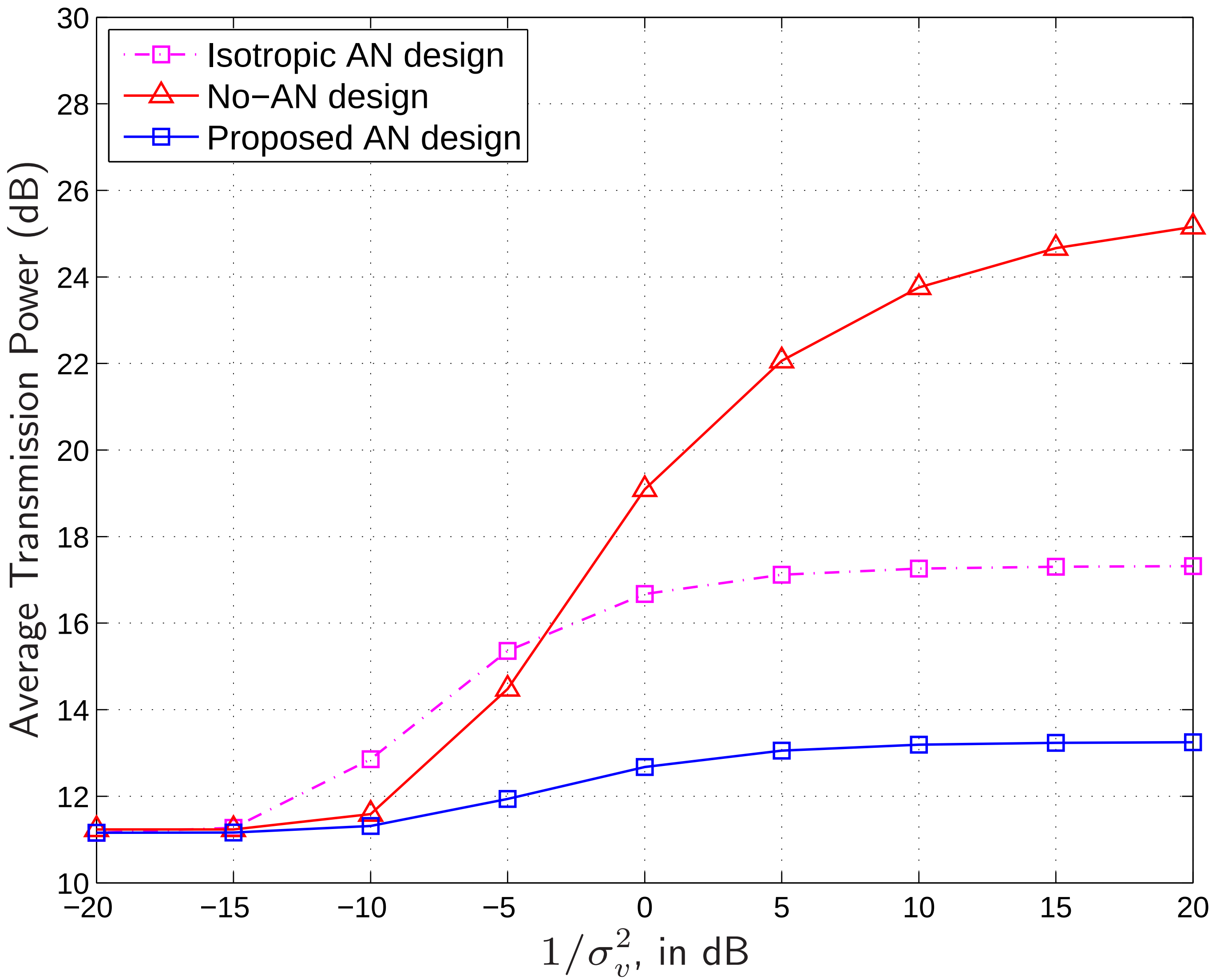
$$\begin{aligned} \min_{\boldsymbol{w}, \boldsymbol{\Sigma} \succeq \mathbf{0}} \; & \|\boldsymbol{w}\|^2 + \text{Tr}(\boldsymbol{\Sigma}) \\ \text{s.t.} \; & \text{SINR}_{\text{Bob}} \geq \gamma, \quad \text{SINR}_{\text{Eve},i} \leq \beta, \quad i = 1, \ldots, L. \end{aligned}$$

  The design achieves a secrecy rate of at least $\log(1 + \gamma) - \log(1 + \beta)$ bps/Hz, with min. power.

Aspects with solving the QoS-based secure tx. BF design:

$$\begin{aligned} \min_{\boldsymbol{w}, \boldsymbol{\Sigma} \succeq \mathbf{0}} \ & \|\boldsymbol{w}\|^2 + \mathrm{Tr}(\boldsymbol{\Sigma}) \\ \text{s.t. } & \mathsf{SINR}_{\mathsf{Bob}} = \frac{|\boldsymbol{h}^H \boldsymbol{w}|^2}{\mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{h}\boldsymbol{h}^H) + \sigma_n^2} \geq \gamma, \\ & \mathsf{SINR}_{\mathsf{Eve},i} = \frac{|\boldsymbol{g}_i^H \boldsymbol{w}|^2}{\mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{g}_i\boldsymbol{g}_i^H) + \sigma_{v,i}^2} \leq \beta, \quad i = 1, \ldots, L. \end{aligned}$$

- cannot be solved by the SOCP approach, except when AN is disabled.
- can be handled by SDR, by replacing $\boldsymbol{W} = \boldsymbol{w}\boldsymbol{w}^H$ with $\boldsymbol{W} \succeq \mathbf{0}$.
- By the now familiarized SBP rank result, rank-one optimality of SDR (w.r.t. $\boldsymbol{W}$ only!) is assured when $L \leq 2$.
- BUT, by examining the KKT condition of this problem, a better result is found— SDR has a rank-one solution with $\boldsymbol{W}$ for any L **[Liao-Chang-Ma-Chi'11]**.
- For stat. CSIT, a similar, but slightly weaker result is established **[Liao-Chang-Ma-Chi'11]**.



Transmit power performance of various secret BF designs. $N_t = 4$; $L = 3$; $\sigma_n^2 = 0$dB; $\gamma = 10$dB; $\beta = 0$dB.

MISO Secrecy Rate Maximization

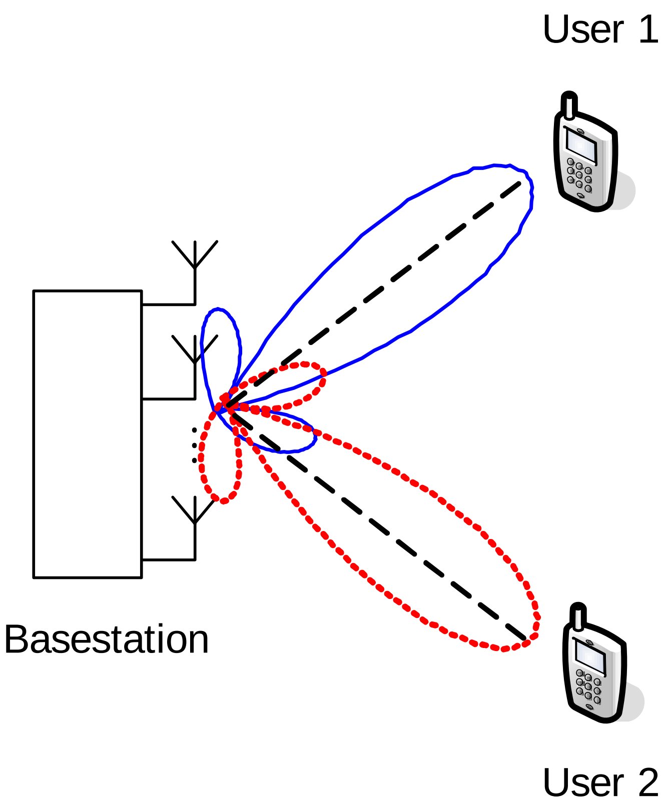
Consider the secrecy rate maximization design:

$$\begin{aligned} \max_{\boldsymbol{w}, \boldsymbol{\Sigma} \succeq \mathbf{0}} \ & \log(1 + \mathsf{SINR}_{\mathsf{Bob}}) - \max_{i=1,\ldots,L} \log(1 + \mathsf{SINR}_{\mathsf{Eve},i}) \\ \text{s.t.} \ & \|\boldsymbol{w}\|^2 + \mathrm{Tr}(\boldsymbol{\Sigma}) \leq P_{\mathsf{max}}. \end{aligned} \tag{$\dagger$}$$

- Problem (†) is nonconvex, even after SDR.
- Despite that, insights from SDR provide vital clues. Recent results **[Li-Ma'11]**, **[Li-Ma'11-B]**:
  - (†) can be solved by a one-variable parameterized SDP; rank-one solution w.r.t. $\boldsymbol{W}$ is proven to exist;
  - without AN, (†) can be solved by an SDP; rank-one solution w.r.t. $\boldsymbol{W}$ is proven to exist uniquely;
  - the above two results can be generalized to handle imperfect CSIT in a worst-case sense (also **[Li-Ma-So'11]** for outage-based robust designs);
  - multi-antenna or colluding Eves can be handled (also **[Liao-Chang-Ma-Chi'11]**).

Frontier Problem: Outage-Based Unicast Transmit Beamforming

Recall the unicast transmit downlink beamforming problem, where we have multiuser MISO downlink, with each user receiving an independent data stream.



- Assume that instantaneous CSIT is perfectly known, $\boldsymbol{R}_i = \boldsymbol{h}_i \boldsymbol{h}_i^H$, $i = 1, \ldots, K$.

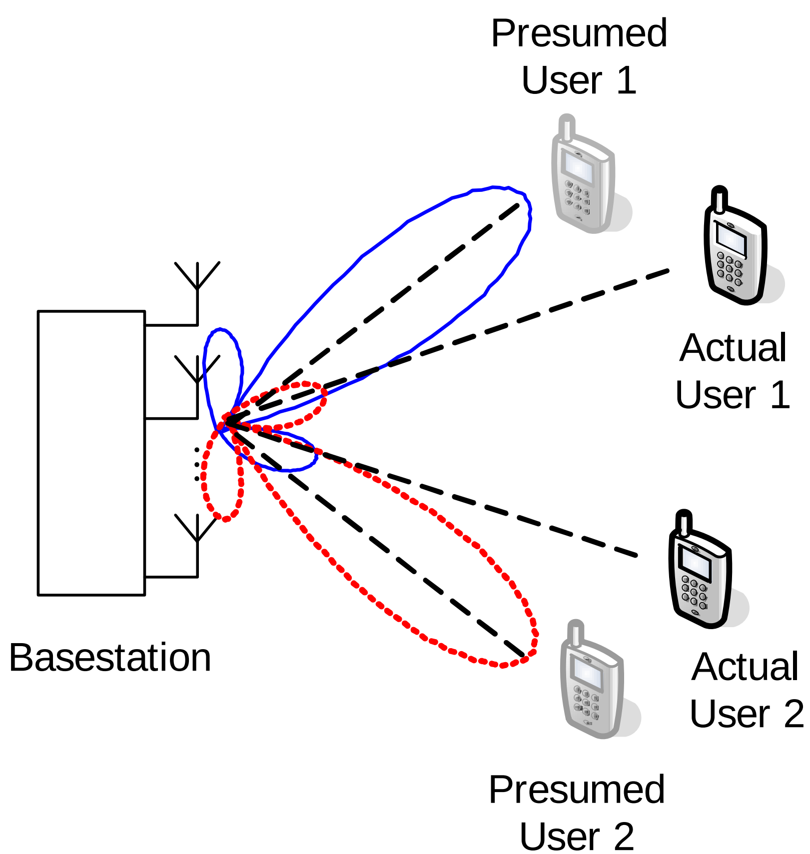
- Previously, we have formulated the following QoS-assured design problem:

$$\begin{aligned} \min_{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K \in \mathbb{C}^N} \quad & \sum_{i=1}^K \|\boldsymbol{w}_i\|^2 \\ \text{s.t.} \quad & \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2} \geq \gamma_i, \\ & \qquad i = 1, \ldots, K. \end{aligned}$$

- As seen before, this can be handled by SDR.

Frontier Problem: Outage-Based Unicast Transmit Beamforming

CSIT is generally imperfectly known in practice.

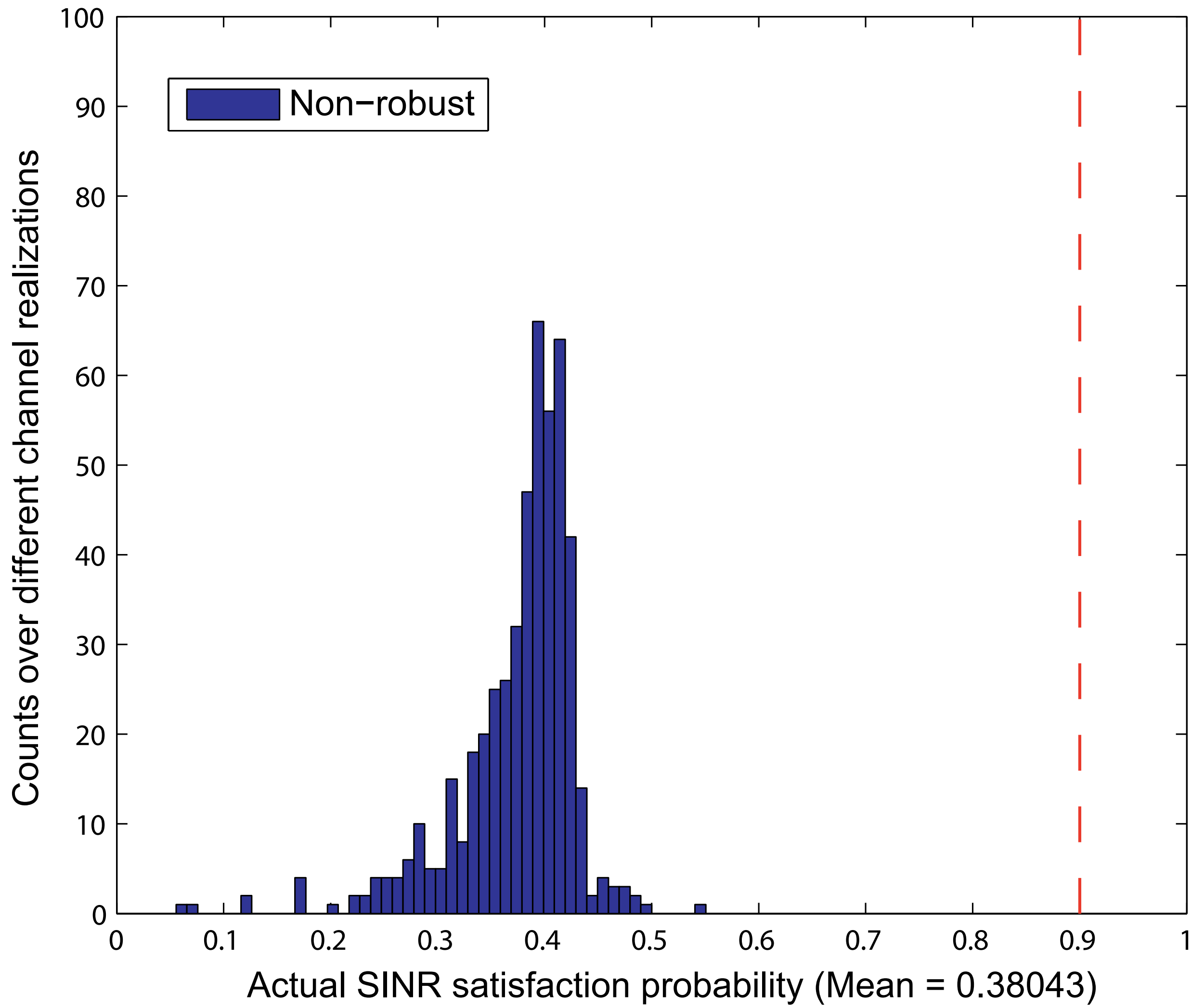


- Suppose that the presumed CSIT is inaccurate.

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

$$
\begin{aligned}
\min_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K \in \mathbb{C}^N} \quad & \sum_{i=1}^K \|\boldsymbol{w}_i\|^2 \\
\text{s.t.} \quad & \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2} \geq \gamma_i, \\
& i = 1, \ldots, K,
\end{aligned}
$$

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; $\gamma = 11$dB. The design has more than 50% outage most of the time.

Outage-Based Unicast Transmit Beamforming: Formulation

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

- A meaningful, but very difficult, design problem:

$$
\begin{aligned}
\min_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K \in \mathbb{C}^N} \quad & \textstyle\sum_{i=1}^K \|\boldsymbol{w}_i\|^2 \\
\text{s.t.} \quad & \text{Prob}_{\boldsymbol{h}_i \sim \mathcal{CN}(\bar{\boldsymbol{h}}_i, \sigma_e^2 \boldsymbol{I})} \left\{ \frac{\boldsymbol{w}_i^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w}_l + \sigma_i^2} \geq \gamma_i \right\} \geq 1 - \rho_i, \\
& \hspace{22em} i = 1, \ldots, K,
\end{aligned}
$$

where the ρ_i's are the maximum tolerable outage probabilities.

- The outage-based SINR constraints

$$\mathrm{Prob}_{\boldsymbol{h}_i \sim \mathcal{CN}(\bar{\boldsymbol{h}}_i, \sigma_e^2 \boldsymbol{I})} \left\{ \frac{\boldsymbol{w}_i^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{h}_i \boldsymbol{h}_i^H \boldsymbol{w}_l + \sigma_i^2} \geq \gamma_i \right\} \geq 1 - \rho_i$$

can be rewritten as

$$\mathrm{Prob}_{\boldsymbol{e}_i \sim \mathcal{CN}(\boldsymbol{0}, \sigma_e^2 \boldsymbol{I})} \left\{ (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)^H \left(\frac{1}{\gamma_i} \boldsymbol{w}_i \boldsymbol{w}_i^H - \sum_{l \neq i} \boldsymbol{w}_l \boldsymbol{w}_l^H \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \geq \sigma_i^2 \right\} \geq 1 - \rho_i.$$

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

$$(\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)^H \left(\frac{1}{\gamma_i} \boldsymbol{w}_i \boldsymbol{w}_i^H - \sum_{l \neq i} \boldsymbol{w}_l \boldsymbol{w}_l^H \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)$$

    is indefinite (and hence nonconvex) in the design variables $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_K$.

Tackling the Nonconvexity: SDR

- Let us first do the thing we are good at — SDR.

- By SDR, we have

$$\text{Prob}_{\boldsymbol{e}_i \sim \mathcal{CN}(\mathbf{0}, \sigma_e^2 \boldsymbol{I})} \left\{ (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)^H \left(\frac{1}{\gamma_i} \boldsymbol{W}_i - \sum_{l \neq i} \boldsymbol{W}_l \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \geq \sigma_i^2 \right\} \geq 1 - \rho_i.$$

- Now, the function

$$(\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)^H \left(\frac{1}{\gamma_i} \boldsymbol{W}_i - \sum_{l \neq i} \boldsymbol{W}_l \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)$$

  is linear in the variables $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_k$, which is good.

- However, the probability still does not admit a simple closed form expression.

Processing the Probabilistic Constraint: Monte Carlo?

- In principle, we can handle the SDRed probabilistic constraint by Monte Carlo methods.

- Specifically, let $\boldsymbol{e}_i^1, \ldots, \boldsymbol{e}_i^L$ be i.i.d. according to $\mathcal{CN}(\mathbf{0}, \sigma_e^2 \boldsymbol{I})$. Here, $L \geq 1$ is some parameter. Consider the SDP constraints

$$
\begin{aligned}
&(\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i^\ell)^H \left(\frac{1}{\gamma_i} \boldsymbol{W}_i - \sum_{l \neq i} \boldsymbol{W}_l \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i^\ell) \geq \sigma_i^2, \quad \ell = 1, \ldots, L, \\
&\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \succeq \mathbf{0}.
\end{aligned} \tag{$\dagger$}
$$

- It can be shown **[Calafiore-Campi'05]** that for sufficiently large L (which depends on the outage tolerance ρ_i), any solution to ($\dagger$) will satisfy the corresponding SDRed probabilistic constraint with high confidence.

- However, this method is extremely time consuming in practice.

Processing the Probabilistic Constraint: Convex Restriction

- Let

$$V_i(\{\boldsymbol{W}_j\}) = \text{Prob}_{\boldsymbol{e}_i \sim \mathcal{CN}(\boldsymbol{0}, \sigma_e^2 \boldsymbol{I})} \left\{ (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)^H \left(\frac{1}{\gamma_i} \boldsymbol{W}_i - \sum_{l \neq i} \boldsymbol{W}_l \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) < \sigma_i^2 \right\}$$

be the violation probability. Recall that we want

$$V_i(\{\boldsymbol{W}_j\}) \leq \rho_i.$$

- It is not hard to see that V_i can be expressed as

$$V_i(\{\boldsymbol{W}_j\}) = \text{Prob}_{\boldsymbol{e} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I})} \left\{ \boldsymbol{e}^H \boldsymbol{Q} \boldsymbol{e} + 2\text{Re}\{\boldsymbol{e}^H \boldsymbol{r}\} + s < 0 \right\}$$

for some $\boldsymbol{Q}$, $\boldsymbol{r}$ and s that depend on $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K$ 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 $f(\boldsymbol{Q}, \boldsymbol{r}, s, \boldsymbol{t})$, where $\boldsymbol{t}$ is an additional decision vector, such that

$$V_i(\{\boldsymbol{W}_j\}) = \mathrm{Prob}_{\boldsymbol{e}\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{I})}\left\{\boldsymbol{e}^H\boldsymbol{Q}\boldsymbol{e} + 2\mathrm{Re}\{\boldsymbol{e}^H\boldsymbol{r}\} + s < 0\right\} \leq f(\boldsymbol{Q}, \boldsymbol{r}, s, \boldsymbol{t}).$$

- Then, by construction, the convex constraint

$$f(\boldsymbol{Q}, \boldsymbol{r}, s, \boldsymbol{t}) \leq \rho \qquad \text{(CR-PC)}$$

serves as a sufficient condition for the probabilistic constraint

$$V_i(\{\boldsymbol{W}_j\}) \leq \rho \qquad \text{(PC)}$$

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

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

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

- Is the constraint

$$e^{-T^{-1}(s)} \leq \rho$$

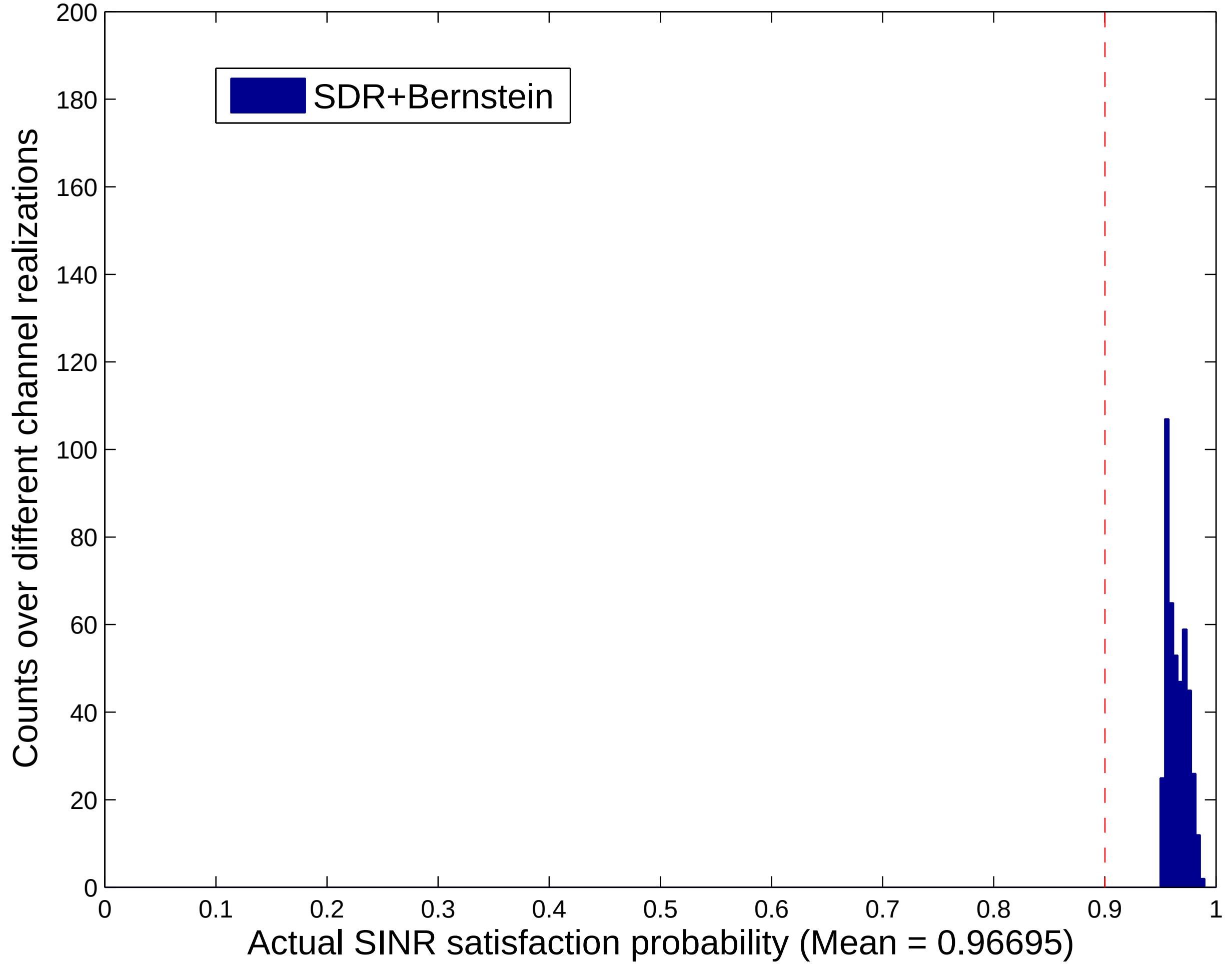
convex? Yes! It is equivalent to

$$\begin{aligned}
&\text{Tr}(\boldsymbol{Q}) - \sqrt{-2\ln(\rho)} \cdot t_1 + \ln(\rho) \cdot t_2 + s \geq 0, \\
&\sqrt{\|\boldsymbol{Q}\|_F^2 + 2\|\boldsymbol{r}\|^2} \leq t_1, \\
&t_2\boldsymbol{I} + \boldsymbol{Q} \succeq \mathbf{0}, \\
&t_2 \geq 0.
\end{aligned}$$

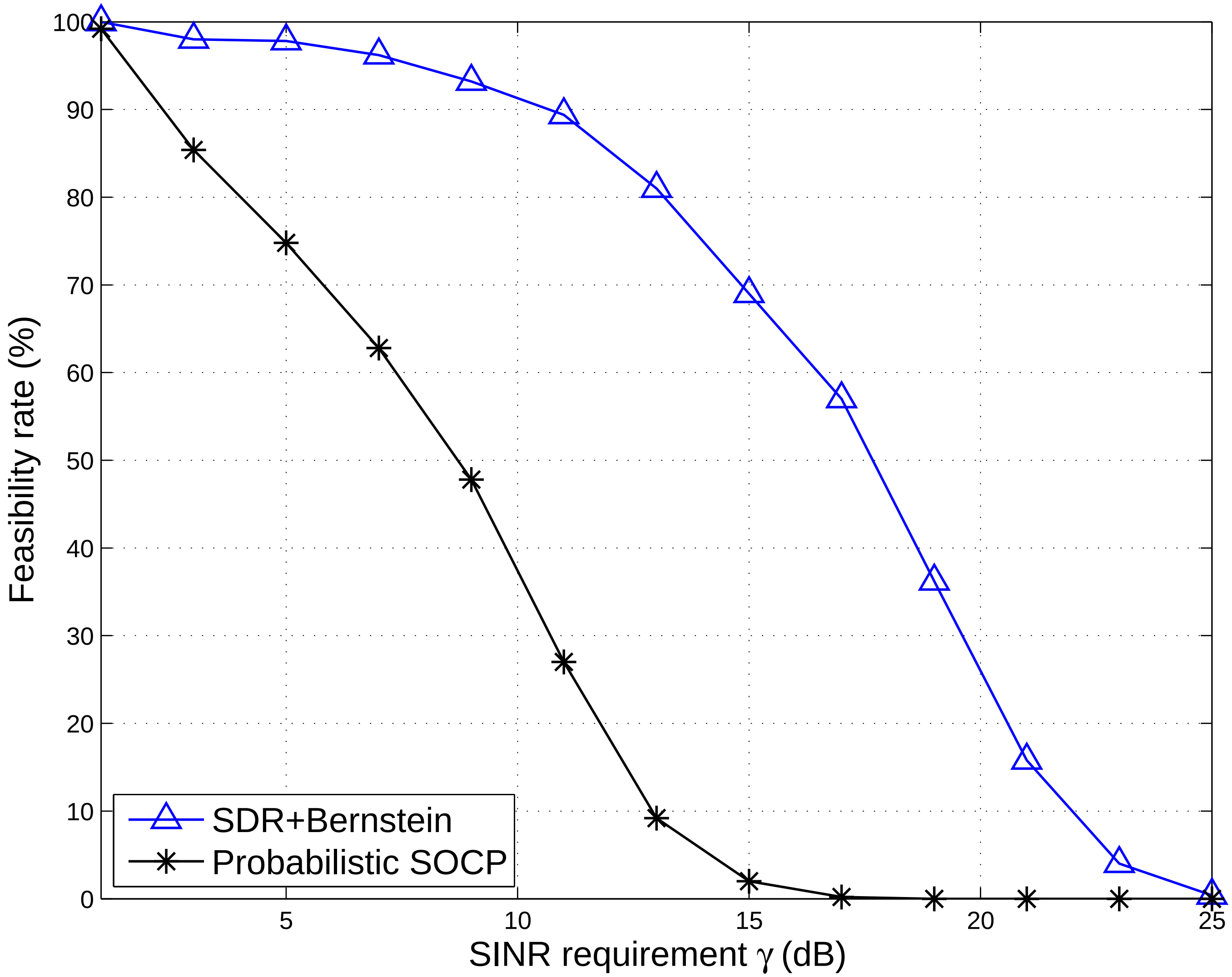
Putting Things Together: The Relaxation-Restriction Approach

- 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]**.

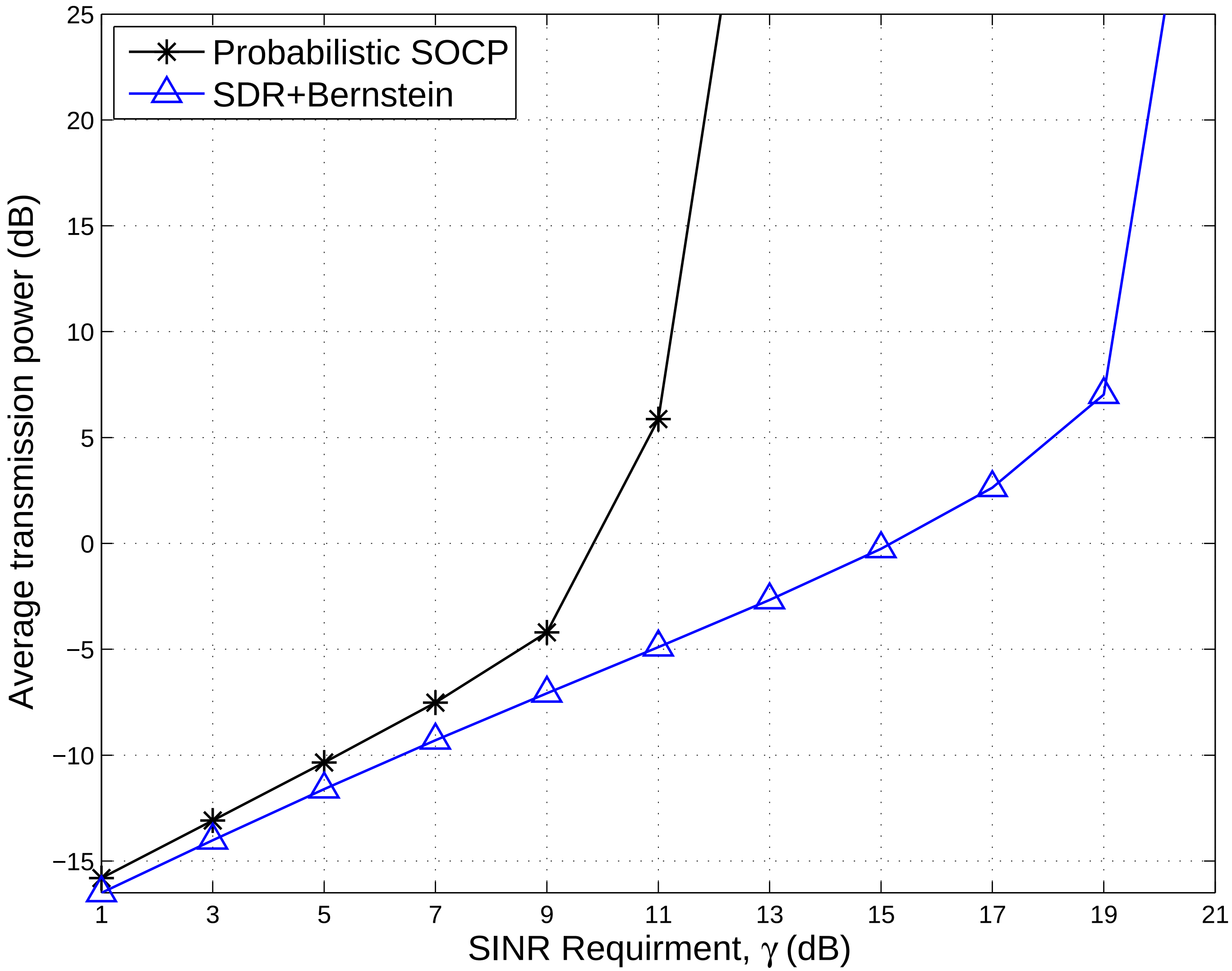
- **A mysterious finding in simulations:** rank-one SDR solution is obtained in almost all the problem instances!



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; $\gamma = 11$dB; $\rho = 0.1$ (90% SINR satisfaction).



Feasibility performance of the proposed method and the probabilistic SOCP method **[Shenouda-Davidson'08]**. $N_t = K = 3$; $\sigma_e^2 = 0.002$; $\gamma = 11$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

[Lovász'79] L. Lovász, "On the Shannon capacity of a graph," *IEEE Trans. Inf. Theory*, 1979.

[Goemans-Williamson'95] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problem using semi-definite programming," *J. ACM*, 1995.

[Tan-Rasmussen'01] P.H. Tan & L.K. Rasmussen, "The application of SDP for detection in CDMA," *IEEE JSAC*, 2001.

[Ma-Davidson-Wong-Luo-Ching'02] W.-K. Ma, T.N. Davidson, K.M. Wong, Z.-Q. Luo, & P.C. Ching, "Quasi-ML multiuser detection using SDR with application to sync. CDMA," *IEEE Trans. Signal Process.*, 2002.

[Bengtsson-Ottersten'01] M. Bengtsson & B. Ottersten, "Optimal & suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Commun.*, L. C. Godara, Ed., CRC Press, 2001.

[Sidiropoulos-Davidson-Luo'06] N.D. Sidiropoulos, T.N. Davidson, & Z.-Q. Luo, "Transmit beamforming for physical layer multicasting," *IEEE Trans. Signal Proc.*, 2006.

[Biswas-Liang-Wang-Ye'06] P. Biswas, T.-C. Liang, T.-C. Wang and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Trans. Sensor Networks*, May 2006.

[Ma-Ching-Vo'04] W.-K. Ma, P.-C. Ching, and B.-N. Vo, "Crosstalk resilient interference cancellation in microphone arrays using capon beamforming," *IEEE Trans. Audio Speech Process.*, 2004.

[De Maio *et al.*'08] A. De Maio, S. De Nicola, Y. Huang, S. Zhang, and A. Farina, "Code design to optimize radar detection performance under accuracy and similarity constraints," *IEEE Trans. Signal Process.*, 2008.

[Chang-Luo-Wu *et al.*'08] T.-H. Chang, Z.-Q. Luo, X. Wu, C. Akgun, J. Vaughan, K. Ugurbil, and P.-F. V. de Moortele, "Transmit B1 shimming at high field with SAR constraints: A two stage optimization independent of the initial set of RF phases and amplitudes," in *Proc. 16th Intl. Society for Magnetic Resonance in Medicine*, May 2008.

[Quan-Ma-Cui-Sayed'10] Z. Quan, W.-K. Ma, S. Cui, and A. Sayed, "Optimal linear fusion for distributed spectrum sensing via semidefinite programming," *IEEE Trans. Signal Process.*, 2010.

[Li-Jiang'07] X. Li and H. Jiang, "Solving large-margin hidden Markov model estimation via semidefinite programming," *IEEE Trans. Audio Speech Lang. Process.*, Nov. 2007.

[Li-Bai-Ding'03] Q. Li, E.-W. Bai & Z. Ding, "Blind source separation of signals with known alphabets using ϵ-approximation algorithms," *IEEE Trans. Signal Process.*, 2003.

[Ma-Vo-Davidson-Ching'06] W.-K. Ma, B.-N. Vo, T.N. Davidson, & P.C. Ching, "Blind ML detection of orthogonal space-time block codes: Efficient high-performance implementations," *IEEE Trans. Signal Process.*, 2006.

[Ma-Ching-Ding'04] W.-K. Ma, P.C. Ching, & Z. Ding, "SDR based multiuser detection for MPSK multiuser systems," *IEEE Trans. Signal Process.*, 2004.

[Ma-Su-Jaldén-Chang-Chi'09] W.-K. Ma, C.-C. Su, J. Jaldén, T.-H. Chang, & C.-Y. Chi, "The equivalence of semidefinite relaxation MIMO Detectors for higher-order QAM," *IEEE J. Sel. Topics Signal Process.*, Dec 2009.

[Steingrimsson-Luo-Wong'03] B. Steingrimsson, Z.-Q. Luo, and K. M. Wong, "Soft quasi-maximumlikelihood detection for multiple-antenna wireless channels," *IEEE Trans. Signal Process.*, Nov. 2003.

[Kisialiou-Luo-Luo'09] M. Kisialiou, X. Luo, and Z.-Q. Luo, "Efficient implementation of quasi-maximum-likelihood detection based on semidefinite relaxation," *IEEE Trans. Signal Process.*, 2009.

[Wai-Ma-So'11] H.-T. Wai, W.-K. Ma, and A. M.-C. So, "Cheap semidefinite relaxation MIMO detection using row-by-row block coordinate descent," in *ICASSP 2011*, 2011.

[Jaldén-Ottersten'08] J. Jaldén and B. Ottersten, "The diversity order of the SDR detector," *IEEE Trans. Inf. Theory*, 2008.

[Kisialiou-Luo'10] M. Kisialiou and Z.-Q. Luo, "Probabilistic analysis of semidefinite relaxation for binary quadratic minimization," *SIAM Journal on Optimization*, 2010.

[So'10] A. M.-C. So, "Non-asymptotic performance analysis of the semidefinite relaxation detector in digital communications," Preprint, 2010.

[Gershman-Sidiropoulos-Shahbazpanahi-Bengtsson-Ottersten'10] 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*, Special Issue on Convex Optimization for Signal Processing, 2010.

[Luo-Chang'10] Z.-Q. Luo and T.-H. Chang, "SDP relaxation of homogeneous quadratic optimization: Approximation bounds and applications," in *Convex Optimization in Signal Processing and Communications*, 2010.

[Fazeli-Dehkordy-Shahbazpanahi-Gazor'09] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Trans. Signal Process.*, Aug. 2009.

[Chalise-Vandendorpe'09] B.K. Chalise and L. Vandendorpe, "MIMO Relay Design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Signal Process.*, July 2009.

[Zhang-Liang-Chai-Cui'09] R. Zhang, Y.-C. Liang, C.-C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Select. Areas Commun.*, June 2009.

[Zhang-Liang-Cui'10] R. Zhang, Y. C. Liang, and S. Cui, "Dynamic resource allocation in cognitive radio networks," *IEEE Signal Processing Magazine*, Special Issue on Convex Optimization for Signal Processing, May 2010.

[Dahrouj-Yu'10] H. Dahrouj and W. Yu, "Coordinated beamforming for multicell multi-antenna wireless system," *IEEE Trans. Wireless Comm.*, May 2010.

[Liao-Chang-Ma-Chi'11] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An artificial-noise-aided approach," *IEEE Tans. Signal Process.*, Mar 2011.

[Li-Ma'11] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, Aug. 2011.

[So-Zhang-Ye'07] A. M.-C. So, J. Zhang, and Y. Ye, "On approximating complex quadratic optimization problems via semidefinite programming relaxations," *Math. Program.*, 2007.

[Zhang-Ma'09] J.-K. Zhang and W.-K. Ma, "Full diversity blind Alamouti space-time block codes for unique identification of flat-fading channels," *IEEE Trans. Signal Proc.*, 2009.

[De Maio *et al.*'09] A. De Maio, S. De Nicola, Z.-Q. Luo, and S. Zhang, "Design of phase codes for radar performance optimization with a similarity constraint," *IEEE Trans. Signal Proc.*, 2009.

[Luo-Sidiropoulos-Tseng-Zhang'07] Z.-Q. Luo, N. D. Sidiropoulos, P. Tseng and S. Zhang, "Approximation bounds for quadratic optimization with homogeneous quadratic constraints," *SIAM Journal on optimization*, Feb. 2007.

[Pataki'98] G. Pataki, "On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues," *Math. Oper. Res.*, 1998.

[Ye-Zhang'03] Y. Ye and S. Zhang, "New results on quadratic minimization," *SIAM J. Optim.*, 2003.

[Huang-Zhang'07] Y. W. Huang and S. Zhang, "Complex matrix decomposition and quadratic programming," *Math. Oper. Res.*, Aug. 2007.

[Huang-Palomar'09] Y. Huang and D. Palomar, "Rank-constrained separable semidefinite program with applications to optimal beamforming," *IEEE Trans. SP*, 2010.

[So-Ye-Zhang'08] A. M.-C. So, Y. Ye, and J. Zhang, "A unified theorem for SDP rank reduction", *Math. Oper. Res.*, 2008.

[Su-Safar-Liu'05] W. Su, Z. Safar, and K. J. R. Liu, "Full-rate full-diversity space-frequency codes with optimum coding advantage," *IEEE Trans. Inf. Theory*, Jan 2005.

[Mow'92] W. H. Mow, "Maximum likelihood sequence estimation from the lattice viewpoint," in *Proc. ICSS/ISITA*, Nov. 1992.

[Viterbo-Biglieri'93] E. Viterbo and E. Biglieri, "A universal lattice decoder," in *Proc.* 14^{eme} *Colloq. GRETSI*, Sept. 1993.

[Damen-El-Gamal-Caire'03] O. Damen, H. El Gamal, $\&$ G. Caire, "On ML detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, 2003.

[Jaldén-Ottersten'05] J. Jaldén & B. Ottersten, "On the complexity of sphere decoding in digital comm.," *IEEE Trans. Signal Process.*, 2005.

[Yao-Wornell'02] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *GLOBECOM 2002*, Nov. 2002.

[Wübben-Seethaler-Jaldén-Matz'11] D. Wübben, D. Seethaler, J. Jaldén, and G. Matz, "Lattice reduction," *IEEE Signal Process. Magazine*, May 2011.

[Taherzadeh-Mobasher-Khandani'07] M. Taherzadeh, A. Mobasher, and A.K. Khandani, "LLL reduction achieves the receive diversity in MIMO decoding," *IEEE Trans. Inf. Theory,* 2007.

[Jaldén-Elia'10] J. Jaldén and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and systems models," *IEEE Trans. Inf. Theory*, Oct. 2010.

[Helmberg-Rendl-Vanderbei-Wolkowicz'96] C. Helmberg, F. Rendl, R.J. Vanderbei, and H. Wolkowicz, "An interior-point method for SDP", *SIAM J. Optim.*, 1996.

[Wen-Goldfarb-Ma-Scheinberg'09] Z.Wen, D. Goldfarb, S. Ma, and K. Scheinberg, "Row by row methods for semidefinite programming," *Tech. Rep.*, Dept of IEOR, Columbia University, Apr. 2009.

[Poljak-Rendl-Wolkowicz'95] S. Poljak, F. Rendl, and H. Wolkowicz, "A recipe for semidefinite relaxation for (0, 1)-quadratic programming," *J. Global Optimiz.*, 1995.

[Wiesel-Eldar-Shamai'05] A. Wiesel, Y.C. Eldar, & S. Shamai, "SDR for detection in 16-QAM signalling in MIMO channels," *IEEE Signal Process. Lett.*, 2005.

[Sidiropoulos-Luo'06] N. D. Sidiropoulos & Z.-Q. Luo, "A SDR approach to MIMO detection for higher-order constellations," *IEEE Signal Process. Lett.*, 2006.

[Mao-Wang-Wang'07] Z. Mao, X. Wang, & X. Wang, "Semidefinite programming relaxation approach for multiuser detection of QAM signals," *IEEE Trans. Wireless Commun.*, 2007.

[Yang-Zhao-Zhou-Wu'07] Y. Yang, C. Zhao, P. Zhou, and W. Xu, "MIMO detection of 16-QAM signaling based on semidefinite relaxation," *IEEE Signal Process. Lett.*, Nov. 2007.

[Mobasher-Taherzadeh-Sotirov-Khandani'07] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, "A near-maximum-likelihood decoding algorithm for MIMO systems based on semi-definite programming," *IEEE Trans. Inf. Theory*, Nov. 2007.

[Ma-Su-Jaldén-Chi'08] W.-K. Ma, C.-C. Su, J. Jaldén, and C.-Y. Chi, "Some results on 16-QAM MIMO detection using semidefinite relaxation," in *Proc. IEEE ICASSP 2008*.

[Jaldén-Martin-Ottersten'03] J. Jaldén, C. Martin, & B. Ottersten, "Semidefinite programming for detection in linear systems— Optimality conditions and space-time decoding," in *Proc. IEEE ICASSP 2003*.

[So'09] A. M.-C. So, "On the performance of semidefinite relaxation MIMO detectors for QAM constellations", in *Proc. IEEE ICASSP 2009*.

[Saxe'79] J. B. Saxe, "Embeddability of weighted graphs in k-space is strongly NP-hard", *Proc. 17th Allerton Conf. Commun. Control Comput.*, 1979.

[So-Ye'07] A. M.-C. So and Y. Ye, "Theory of semidefinite programming for sensor network localization", *Math. Prog.*, 2007.

[So-Ye'06] A. M.-C. So and Y. Ye, "A semidefinite programming approach to tensegrity theory and realizability of graphs", *Proc. 17th SODA*, 2006.

[Wang-Zheng-Ye-Boyd'08] Z. Wang, S. Zheng, Y. Ye, and S. Boyd, "Further relaxations of the semidefinite programming approach to sensor network localization", *SIAM J. Opt.*, 2008.

[Biswas-Liang-Wang-Ye'06] P. Biswas, T.-C. Liang, T.-C. Wang and Y. Ye, "Semidefinite programming based algorithms for sensor network localization", *ACM Trans. Sensor Networks*, May 2006.

[Lui-Ma-So-Chan'09] W. K. Lui, W.-K. Ma, H. C. So, and K. W. Chan, "Semi-definite programming algorithms for sensor network node localization with uncertainties in anchor positions and/or propagation speed", *IEEE Trans. Signal Proc.*, 2009.

[Wiesel-Eldar-Shamai'06] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, 2006.

[Schubert-Boche'04] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, Jan. 2004.

[Karipidis-Sidiropoulos-Luo'08] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, March 2008.

[Phan-Vorobyov-Sidiropoulos-Tellambura'09] K. Phan, S. Vorobyov, N. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, June 2009.

[Wyner'75] A. D. Wyner, "The wiretap channel," *The Bell System Technical Journal*, vol. 54, October 1975, pp. 1355–1387.

[Nego-Goel'05] R. Negi and S. Goel, "Secret communication using artificial noise," in *IEEE Vehicular Technology Conference*, Sept. 2005, pp. 1906-1910.

[Li-Ma'11-B] Q. Li and W.-K. Ma, "A robust artificial noise aided transmit design for MISO secrecy," in *ICASSP 2011*.

[Li-Ma-So'11] Q. Li, W.-K. Ma, and A. M.-C. So, "Safe convex approximation to outage-based MISO secrecy rate Optimization under imperfect CSI and with artificial noise," to appear in *Asilomar 2011*.

[Calafiore-Campi'05] G. Calafiore and M. C. Campi, "Uncertain convex programs: Randomized solutions and confidence levels", *Math. Prog.*, 2005.

[Bechar'09] I. Bechar, "A Bernstein-type inequality for stochastic processes of quadratic forms of Gaussian variables," *Preprint*, 2009.

[Wang-Chang-Ma-So-Chi'11] K.-Y. Wang, T.-H. Chang, W.-K. Ma, A. M.-C. So, and C.-Y. Chi, "Probabilistic SINR constrained robust transmit beamforming: A Bernstein-type inequality based conservative approach," in *ICASSP 2011*, 2011.

[Shenouda-Davidson'08] M. B. Shenouda and T. N. Davidson, "Probabilistically-constrained approaches to the design of the multiple antenna downlink," in *Proc. 42nd Asilomar Conference 2008*.