# Semidefinite Relaxation: From Theory to Applications to Latest Advances

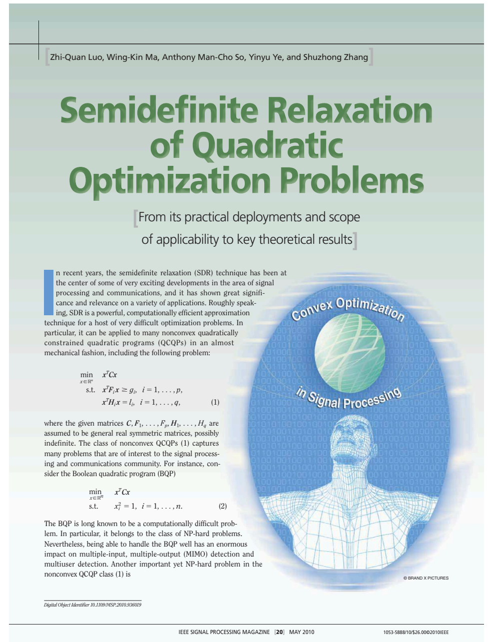
**Wing-Kin (Ken) Ma**
Dept. Electronic Engineering
The Chinese University of Hong Kong
Hong Kong

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

**ICASSP 2014 Tutorial, May 5, 2014**

- The main reference of this tutorial:

Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, & S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, 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]

In recent years, the semidefinite relaxation (SDR) technique has been at the center of some of very exciting developments in the area of signal processing and communications, and it has shown great significance and relevance on a variety of applications. Roughly speaking, SDR is a powerful, computationally efficient approximation technique for a host of very difficult optimization problems. In particular, it can be applied to many nonconvex quadratically constrained quadratic programs (QCQPs) in an almost mechanical fashion, including the following problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T C x \\ \text{s.t.} \quad & x^T F_i x \geq g_i, \quad i = 1, \ldots, p, \\ & x^T H_i x = l_i, \quad i = 1, \ldots, q, \end{aligned} \qquad (1)$$

where the given matrices $C, F_1, \ldots, F_p, H_1, \ldots, H_q$ are assumed to be general real symmetric matrices, possibly indefinite. The class of nonconvex QCQPs (1) captures many problems that are of interest to the signal processing and communications community. For instance, consider the Boolean quadratic program (BQP)

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

The BQP is long known to be a computationally difficult problem. In particular, it belongs to the class of NP-hard problems. Nevertheless, being able to handle the BQP well has an enormous impact on multiple-input, multiple-output (MIMO) detection and multiuser detection. Another important yet NP-hard problem in the nonconvex QCQP class (1) is



Digital Object Identifier 10.1109/MSP.2010.936019

- Acknowledgment: Tom Luo, Yinyu Ye, & Shuzhong Zhang for co-authoring the above article; Qiang Li, Jiaxian Pan, Xiaoxiao Wu & Xiao Fu for helping prepare this slides.
- Thanks also go to everyone who has contributed to this powerful tool.

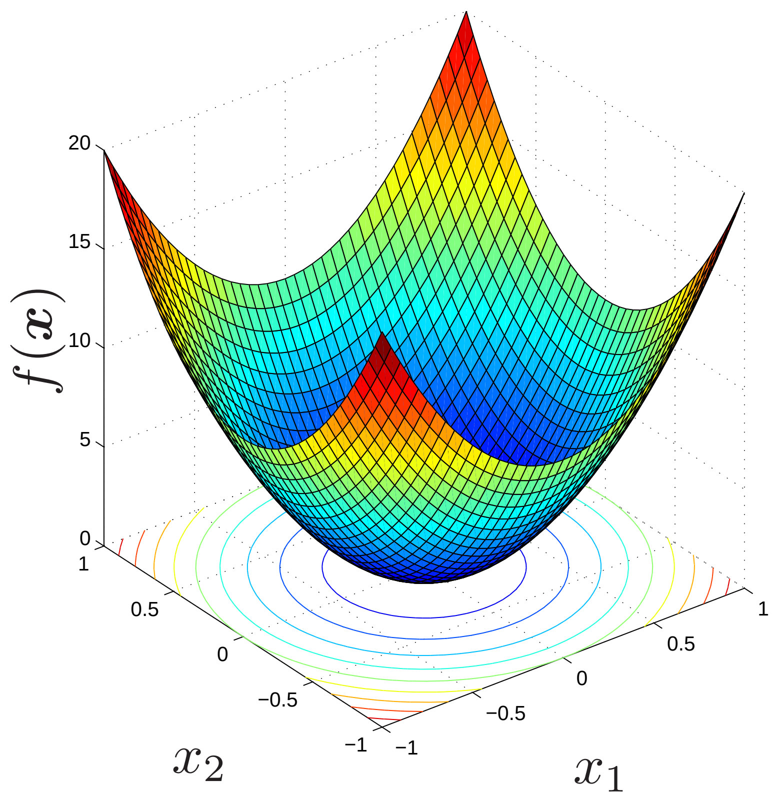
# Outline

- Part I: Basic concepts and overview of semidefinite relaxation (SDR)
- Part II: Theory, and implications in practice
- Part III: Applications and Latest Advances
  - A. transmit beamforming
  - B. advanced topics in transmit beamforming
  - C. sensor network localization
- Conclusion
- Bonus Material: MIMO detection

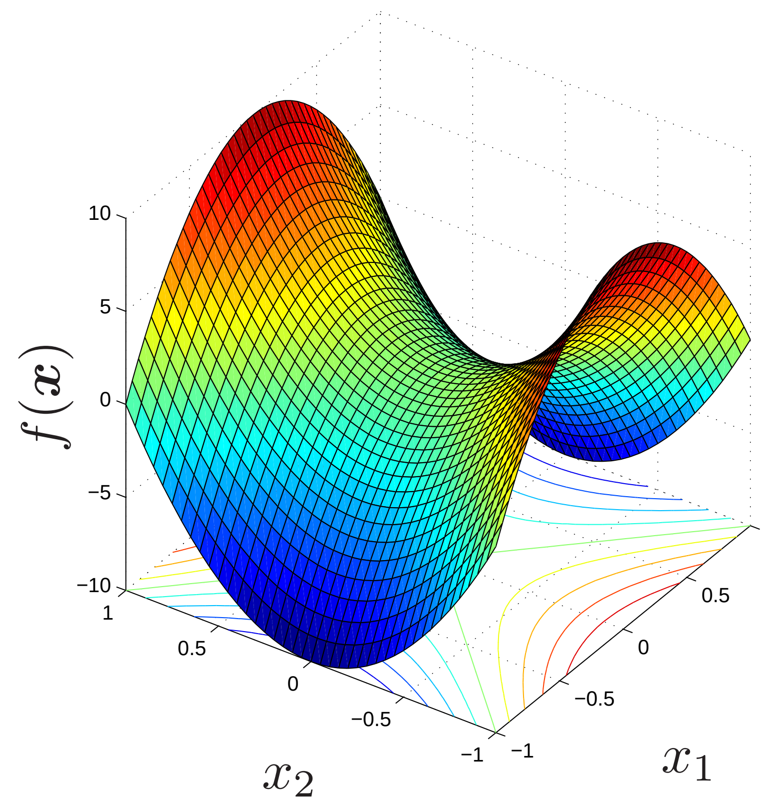
# Part I: Basic Concepts and Overview

**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 $\boldsymbol{C} \succeq \boldsymbol{0}$ ($\boldsymbol{C} \succeq \boldsymbol{0}$ means that $\boldsymbol{C}$ is positive semidefinite (PSD)).

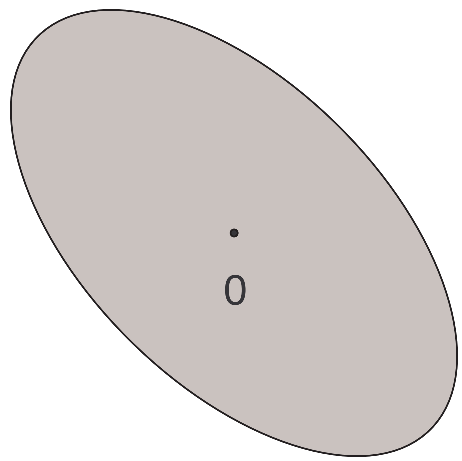



**(a)** $\boldsymbol{C} \succeq \boldsymbol{0}$**.**

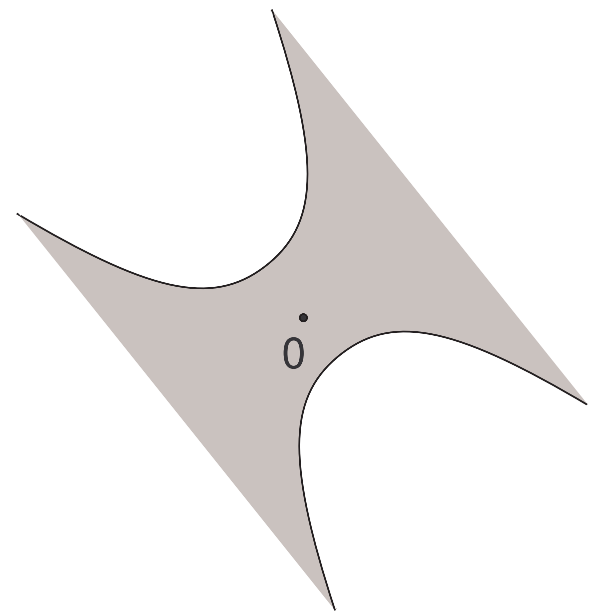



**(b)** $\boldsymbol{C} \not\succeq \boldsymbol{0}$**.**

- 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 $\boldsymbol{F} \succeq \mathbf{0}$.




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




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

- 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 assume convexity here. In particular, $\boldsymbol{C}, \boldsymbol{F}_i, \boldsymbol{H}_i$ can be arbitrary.
- Nonconvex QCQP is a very difficult problem in general.

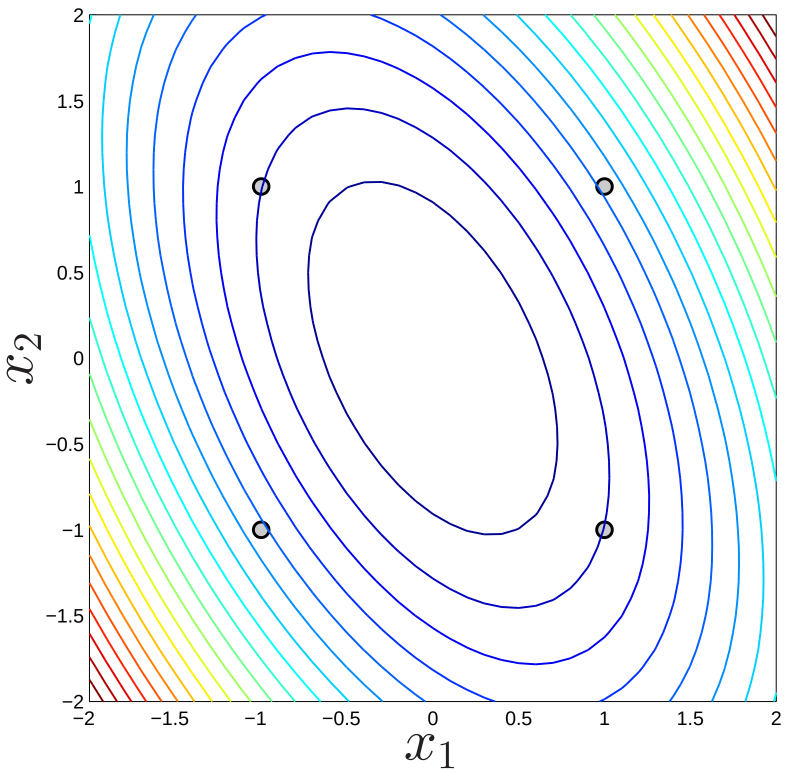
# Nonconvex QCQP: How Hard Could it Be?

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

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

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

- One could solve it by evaluating all possible combinations; i.e., brute-force search.
- The time 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 time $\mathcal{O}(n^p)$ for any fixed $p > 0$.



# Nonconvex QCQP: How Hard Could it Be?

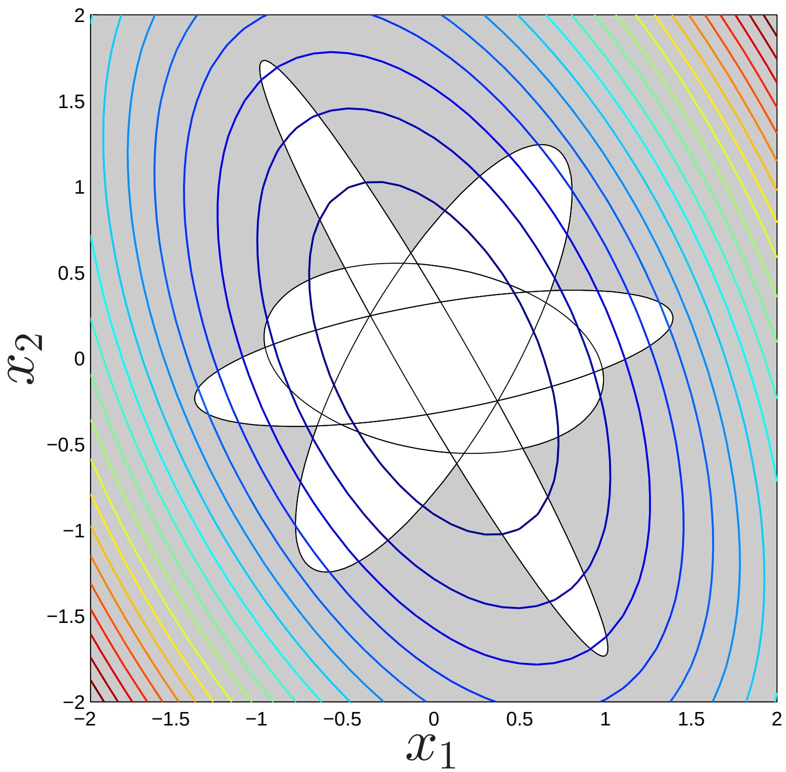
Consider another QCQP:

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

where $\boldsymbol{C}, \boldsymbol{F}_1, \ldots, \boldsymbol{F}_m \succeq \boldsymbol{0}$ (to make the dimension explicit, we will also use the notation $\boldsymbol{C}, \boldsymbol{F}_1, \ldots, \boldsymbol{F}_m \in \mathbb{S}^n_+$).

- 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.
- Existing applications include
  - multiuser/MIMO detection **[Tan-Rasmussen'01]**, **[Ma-Davidson-Wong-Luo-Ching'02]**
  - transmit beamforming: unicast beamforming **[Bengtsson-Ottersten'01]**, multicast beamforming **[Sidiropoulos-Davidson-Luo'06]**, & many others...
  - source localization and sensor network localization **[Cheung-Ma-So'04]**, **[Biswas-Liang-Wang-Ye'06]**
  - code waveform design in radar **[De Maio *et al.*'08]**
  - large-margin parameter estimation in speech recognition **[Li-Jiang'07]**
  - optimal power flow in electrical grids **[Low'14]** (also **[Bienstock'14]**)
  - (related) phase retrieval **[Candès-Eldar-Strohmer-Voroninski'13]**
  - others: robust blind receive beamforming, transmit $B_1$ shim in MRI, distributed detection, phase unwrapping...
- We believe that more applications are on the way.

# 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

$$\boldsymbol{X} \succeq \boldsymbol{0}, \quad \mathrm{rank}(\boldsymbol{X}) \leq 1.$$

Hence, (QCQP) can be reformulated as

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

- The constraints $\mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}) \trianglerighteq_i b_i$ are easy, but $\mathrm{rank}(\boldsymbol{X}) \leq 1$ 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 convex and is an instance of semidefinite program (SDP).

# Some Merits We Can Immediately Say

- A globally optimal solution to an SDP can be found by available numerical algorithms in polynomial time (often by interior-point methods, in $\mathcal{O}(\max\{m,n\}^4 n^{1/2} \log(1/\epsilon))$, $\epsilon$ being soln. accuracy).

- For instance, using the software toolbox CVX, we can solve (SDR) in MATLAB with the following lines: (for simplicity we assume '$\trianglerighteq_i$' $=$ '$\geq$' 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 a solution to (SDR) into 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).

- However, we cannot guarantee that $\boldsymbol{X}^\star$ is always of rank-one. (Otherwise we would have solved an NP-hard problem in polynomial time!)

- There are many ways to produce an approximate QCQP solution from $\boldsymbol{X}^\star$ when $\mathrm{rank}(\boldsymbol{X}^\star) > 1$.

# QCQP Solution Approximation in SDR: An Example

- Consider again the BQP

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

The SDR of (BQP) is

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

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

1) Carry out the eigen-decomposition

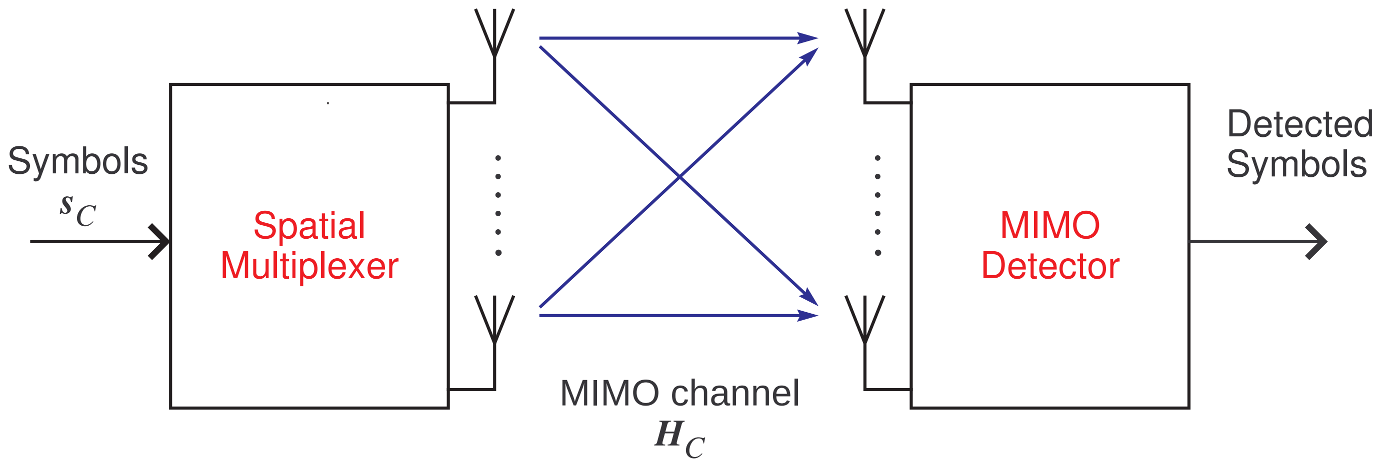
$$\boldsymbol{X}^\star = \sum_{i=1}^{r} \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^T,$$

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

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

# 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)

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

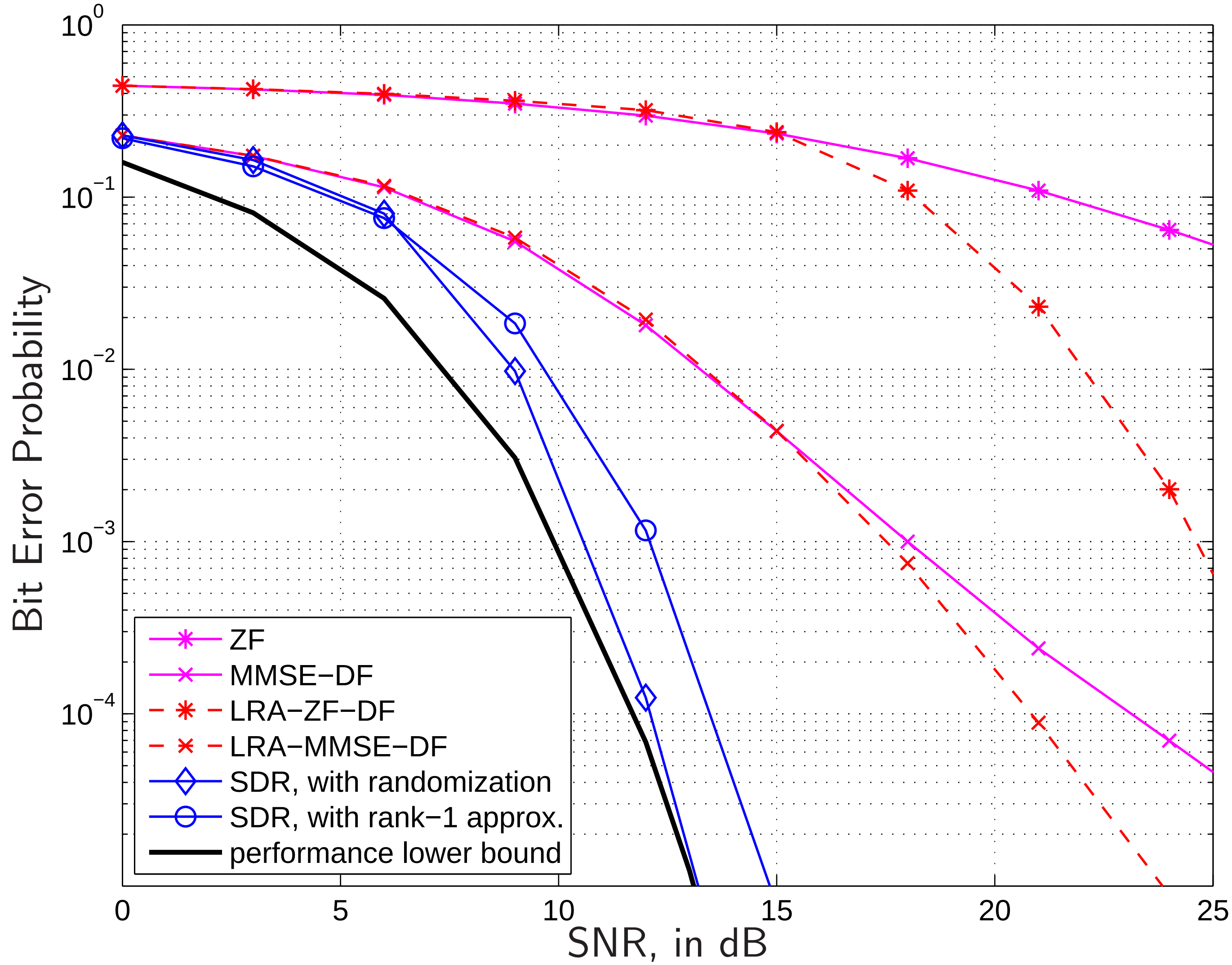
- The received signal model can be converted to a real form

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

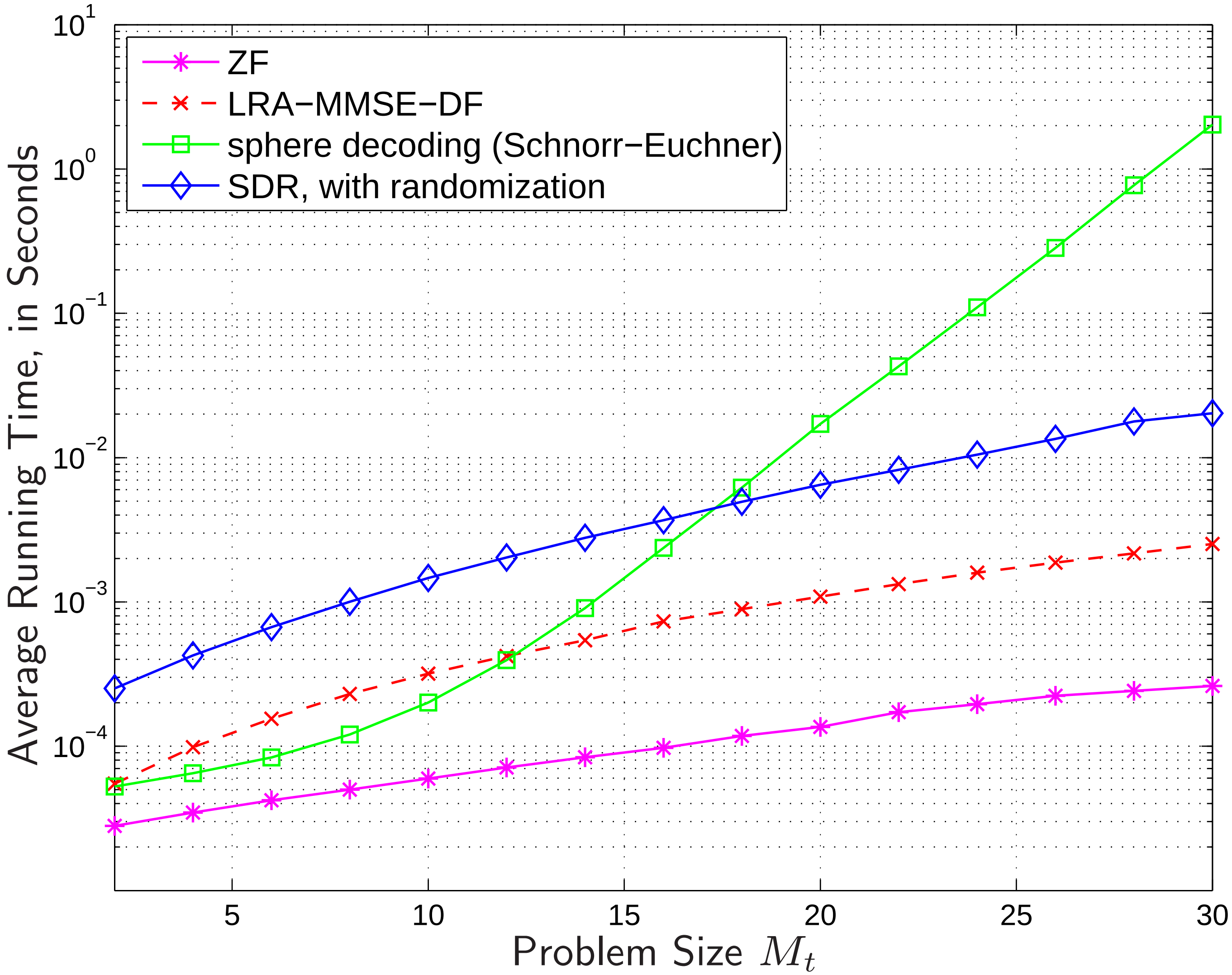
and hence the ML problem can be rewritten (homogenized) as

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

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.




Complexity comparison of various MIMO detectors. SNR= 12dB. Sphere decoding is an exact ML method.

## Additional Remarks about the MIMO Detection Application

- The idea is not restricted to spatial multiplexing! It can also be used in multiuser CDMA, space-time/freq./time-freq. coding, multiuser MIMO, massive 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

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

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

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

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

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

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

- The stochastic QCQP interpretation of SDR

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

motivates another approach to approximating QCQPs, namely

> generate a random vector $\boldsymbol{\xi}\sim\mathcal{N}(\mathbf{0},\boldsymbol{X}^\star)$ ($\boldsymbol{X}^\star$ is an SDR soln.), then 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.

- The stochastic QCQP interpretation allows one to establish many important theoretical SDR approx. accuracy results. This will be explained in Part II.

# Example: Randomization in BQP or MIMO Detection

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

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

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

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

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

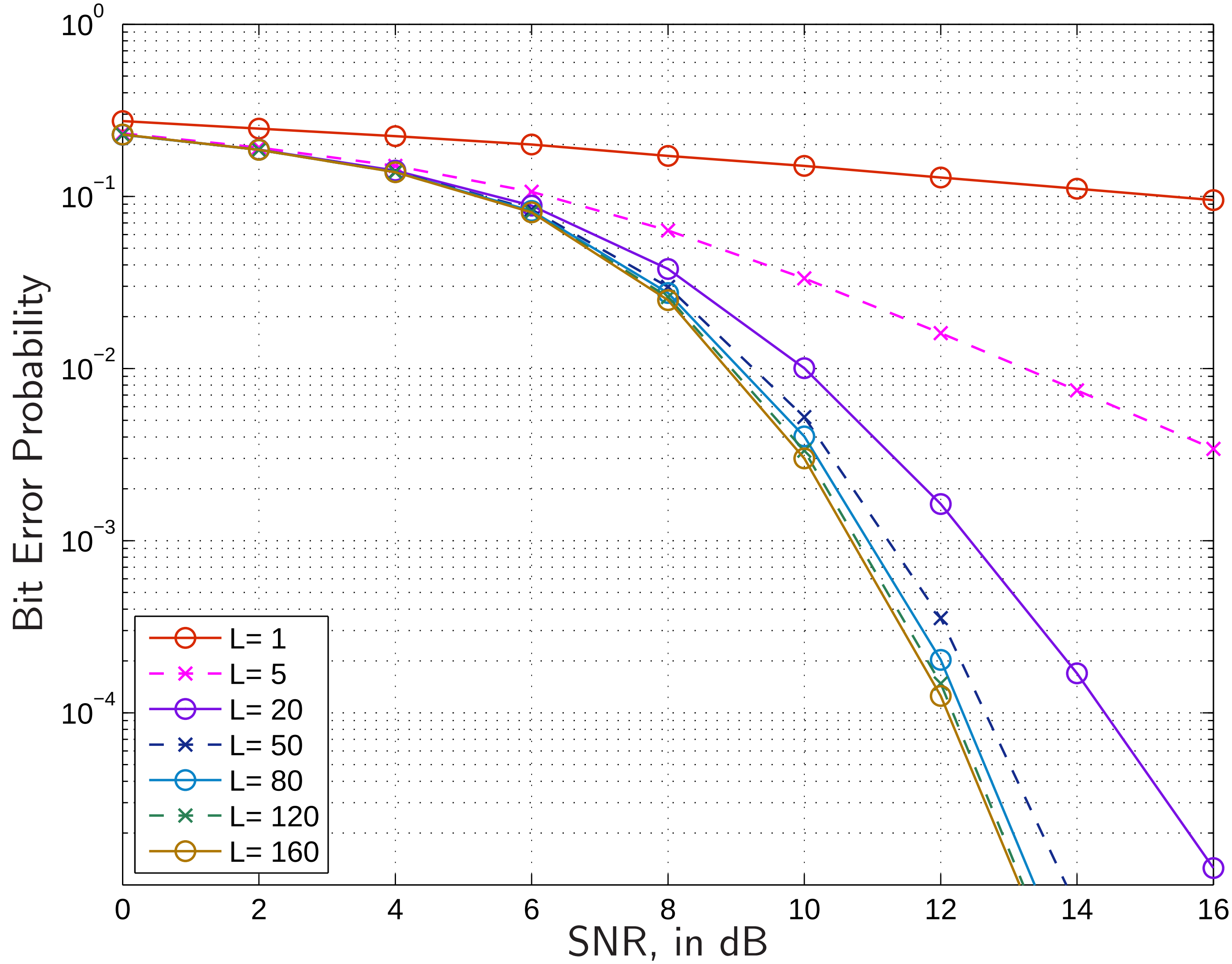
generate $\boldsymbol{\xi}_\ell \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{X}^\star)$, and construct a feasible point

$$\tilde{\boldsymbol{x}}_\ell = \mathrm{sgn}(\boldsymbol{\xi}_\ell);$$

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

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

  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 as before, one can derive an SDR for the complex-valued QCQP:

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

  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 can be different. This will be explained in Part II.

# Application: Multicast Transmit Beamforming

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

- The transmit signal:

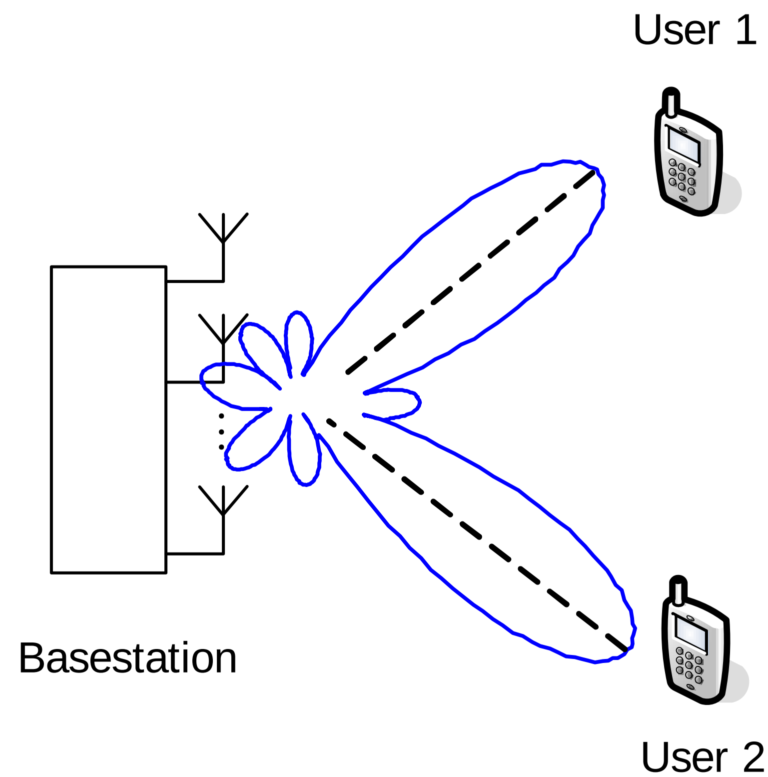
$$\boldsymbol{x}(t) = \boldsymbol{w}s(t),$$

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 $\sigma_i^2$.



- **Problem:** Optimize $\boldsymbol{w}$ by a QoS-assured design:

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

where $\gamma$ is a prescribed SNR requirement for all users, 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

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

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

- 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

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

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 (†).

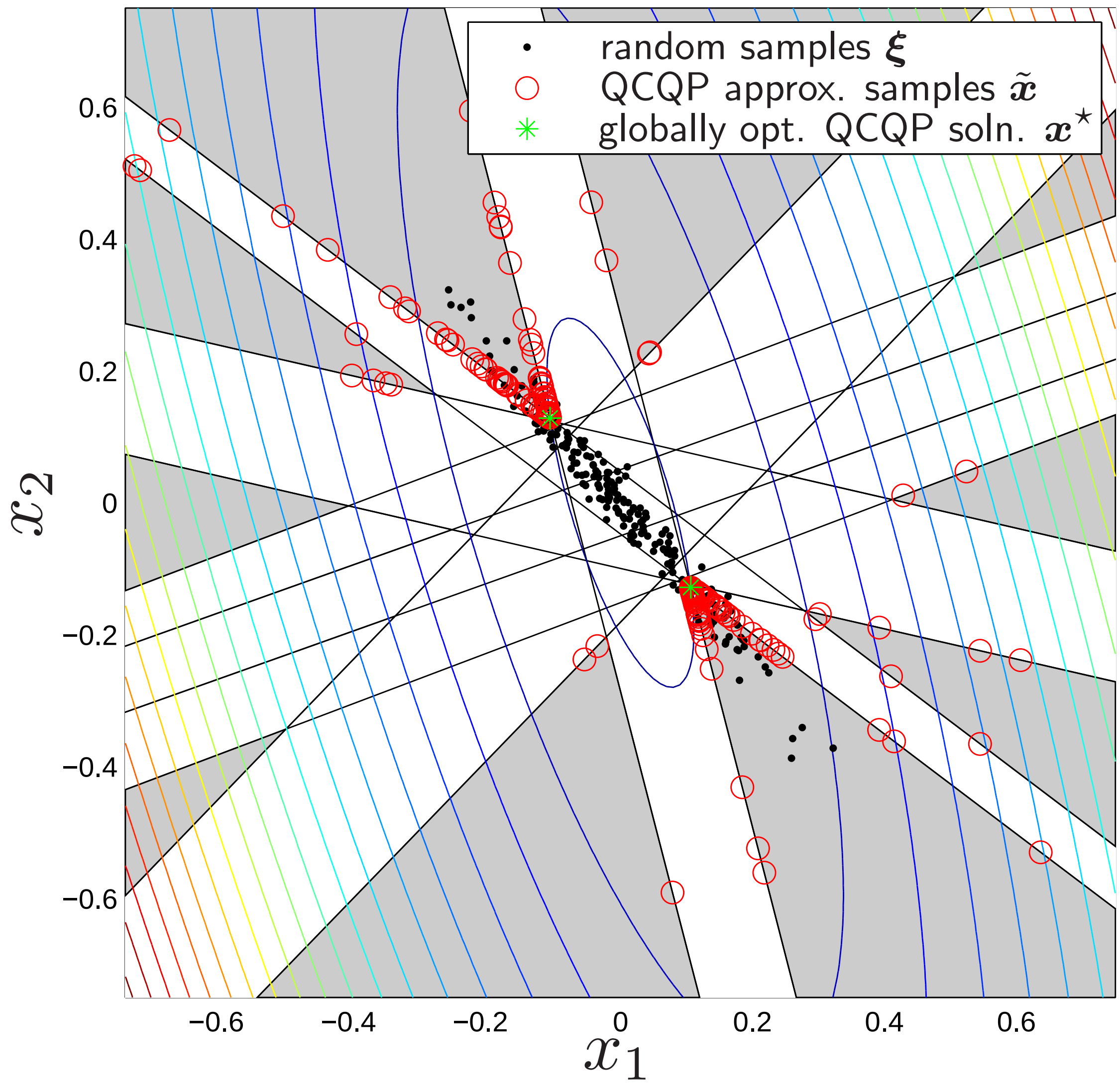
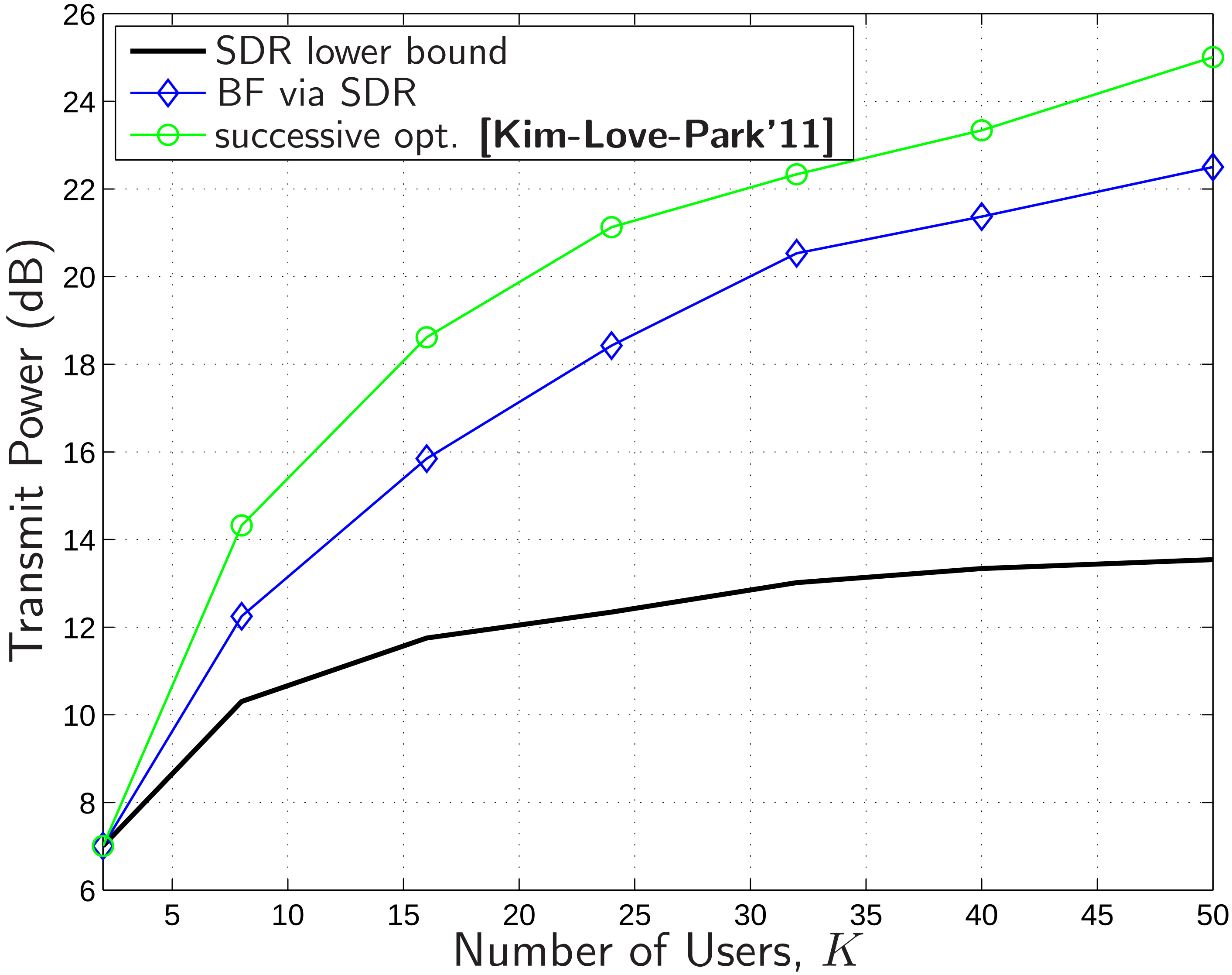



Illustration of randomizations in $\mathbb{R}^2$, for Problem (†). The gray area is the feasible set and colored lines the contour of the objective.




Performance of SDR-based multicast beamforming with respect to the number of users. $\gamma = 10\text{dB}$. $N_t = 4$.

# Extension to Complex-Valued Separable QCQP

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

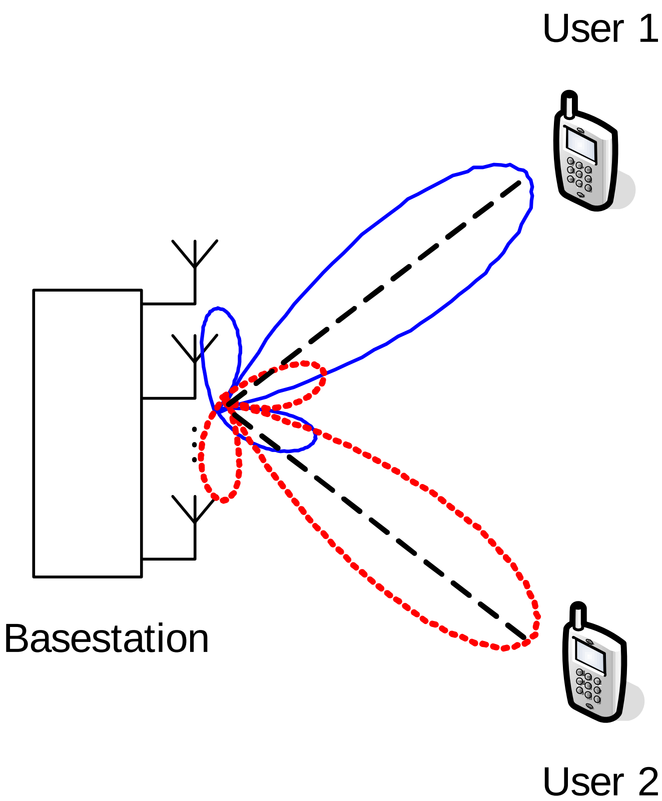
$$
\begin{aligned}
\min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_k \in \mathbb{C}^n} \quad & \sum_{i=1}^{k} \boldsymbol{x}_i^H \boldsymbol{C}_i \boldsymbol{x}_i \\
\text{s.t.} \quad & \sum_{l=1}^{k} \boldsymbol{x}_l^H \boldsymbol{A}_{i,l} \boldsymbol{x}_l \trianglerighteq_i b_i, \quad i = 1, \ldots, m.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

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

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

# 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 $i$ is given by

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

  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.

- **Problem:** Given users' SINR requirements $\gamma_1, \ldots, \gamma_K$, solve

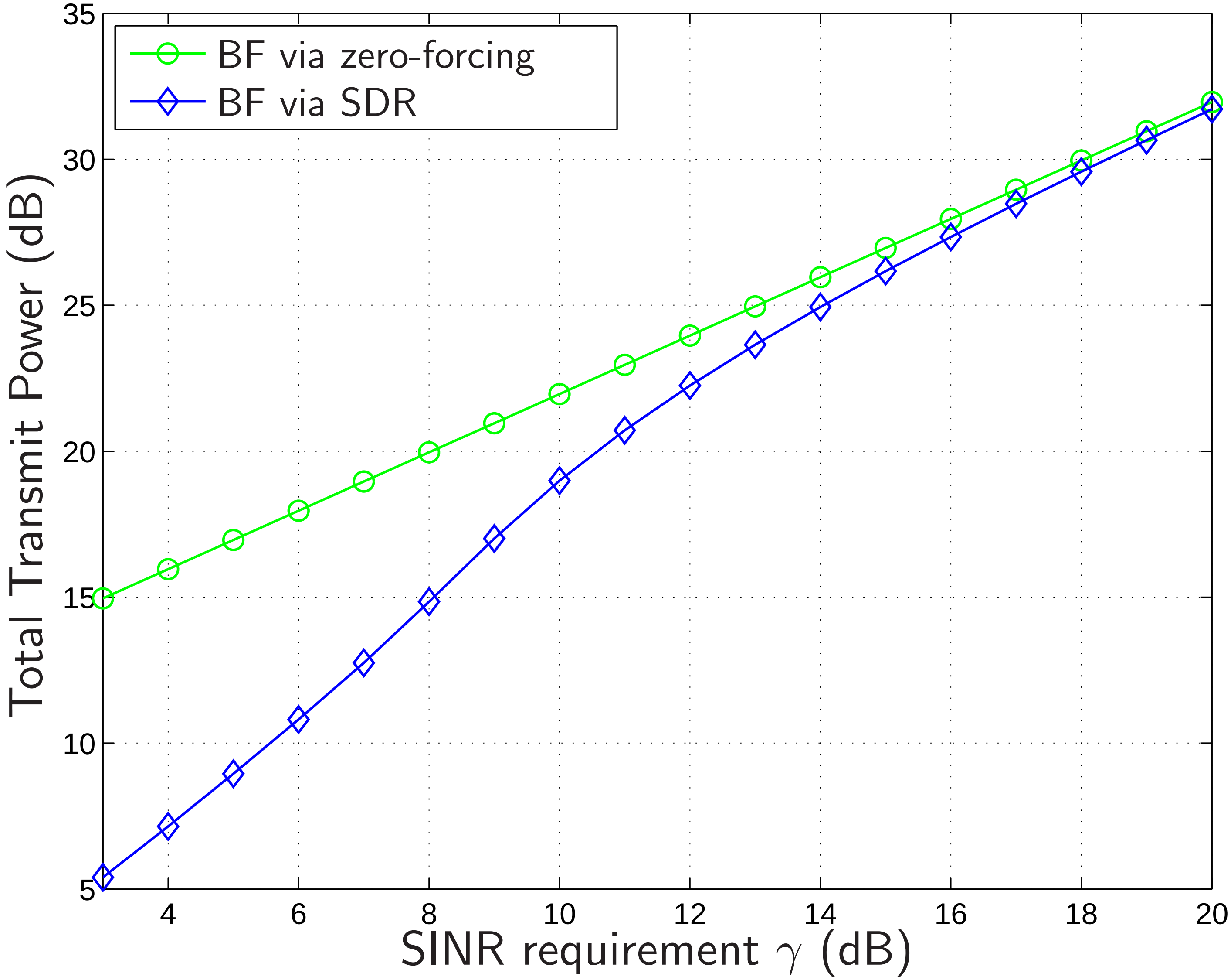
$$\begin{aligned} \min_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K \in \mathbb{C}^{N_t}} \quad & \textstyle\sum_{i=1}^K \|\boldsymbol{w}_i\|^2 \\ \text{s.t.} \quad & \mathsf{SINR}_i \geq \gamma_i, \quad i = 1, \ldots, K. \end{aligned} \tag{$\dagger$}$$

- Write $\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H$. The SDR of ($\dagger$) is

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

- ($\ddagger$) is shown to have a rank-one solution for $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_K \succeq \boldsymbol{0}$, via uplink-downlink duality **[Bengtsson-Ottersten'01]**. Thus, the SDR is tight!

- In Part II, we will introduce an "easier" way to establish the tightness of SDRs.




Performance of unicast beamforming with respect to the SINR requirement. $N_t = K = 8$, $\gamma = \gamma_1 = \cdots = \gamma_K$.

## Additional Remarks about the Transmit Beamforming Application

- Transmit beamforming is now a key topic. For review articles, see **[Gershman-Sidiropoulos-Shahbazpanahi-Bengtsson-Ottersten'10]**, **[Luo-Chang'10]**.

- From the original unicast and multicast beamforming problems, numerous extensions are emerging— e.g., multicell coordinated beamforming, cognitive radio beamforming, relay beamforming, secrecy beamforming, and energy harvesting beamforming— and they will be described in Part III.A.

- All these beamforming problems turn out to be, or be closely related to, nonconvex QCQPs, and hence SDR plays a key role.

- In the transmit beamforming context, SDR is not just a direct application. There are new developments that were not previously seen even in optimization; they will be described in Part III.B.

# 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) (e.g., SQP in the MATLAB Optimization Toolbox).

- So should we compare SDR and NPM?

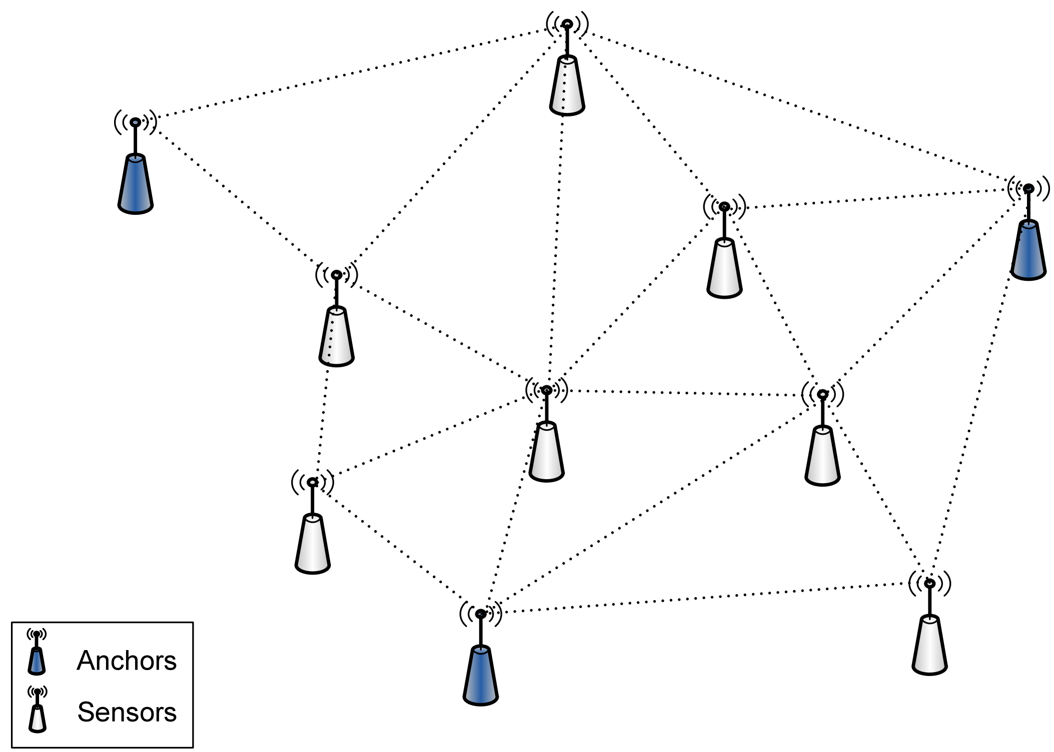
- The interesting argument is that they complement each other, instead of competing:
  - An NPM depends much on a 'good' starting point, and that's usually the missing piece.
  - To SDR, NPMs may serve as a local refinement of the solution.

- One may consider a two-stage approach where SDR is used as a starting point for NPMs.

# Application: Sensor Network Localization (SNL)

**Scenario:** A network of sensors deployed in an area.

- A few sensors, called anchors, have self-localization capability (e.g., by GPS).
- The others (& the majority in the network) do not.




- A pair of sensors that are within comm. range can measure their relative distance, e.g., by measuring the time-of-arrival info., or by ping-pong.
- **Goal:** Estimate the unknown sensor locations using the distance measurements & info. of anchor locations.

- Model:
  - Let
$$\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}, \quad \boldsymbol{a}_i \in \mathbb{R}^2,$$
    be the collection of all (known) anchor coordinates.
  - Let
$$\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}, \quad \boldsymbol{x}_i \in \mathbb{R}^2,$$
    be the collection of all (unknown) sensor coordinates.
  - Let $d_{ij}$ (resp. $\bar{d}_{ij}$) be the distance measurement between sensor $i$ and sensor $j$ (resp. sensor $i$ and anchor $j$). They are modeled as

$$\begin{aligned} d_{ij} &= \|\boldsymbol{x}_i - \boldsymbol{x}_j\| + \eta_{ij}, \\ \bar{d}_{ij} &= \|\boldsymbol{x}_i - \boldsymbol{a}_j\| + \bar{\eta}_{ij}, \end{aligned}$$

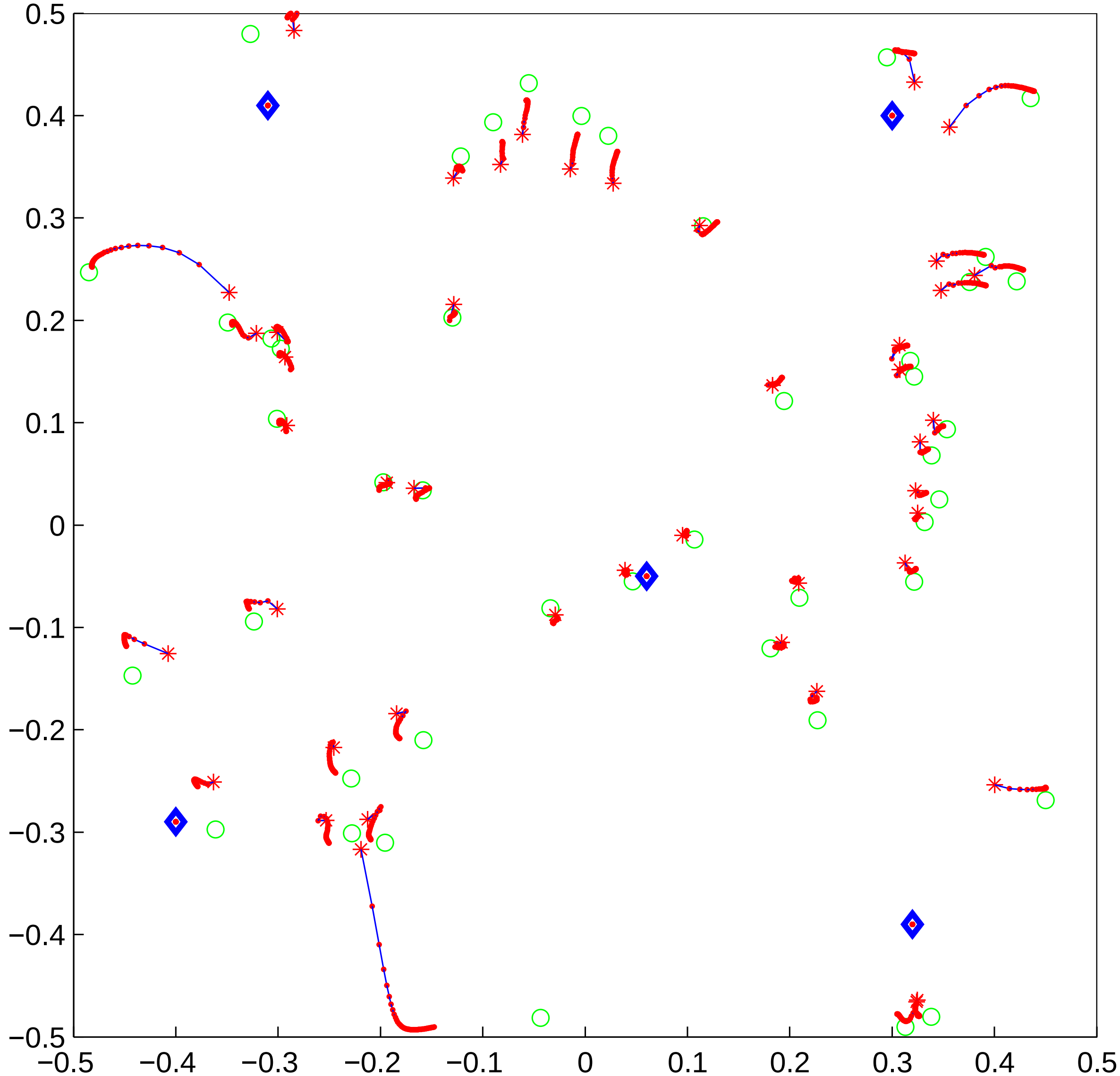
    resp., where $\eta_{ij}, \bar{\eta}_{ij}$ are measurement noise.

- **Problem:** Maximum-likelihood (ML) SNL formulation under noisy distance measurements **[Biswas-Liang-Wang-Ye'06]**:

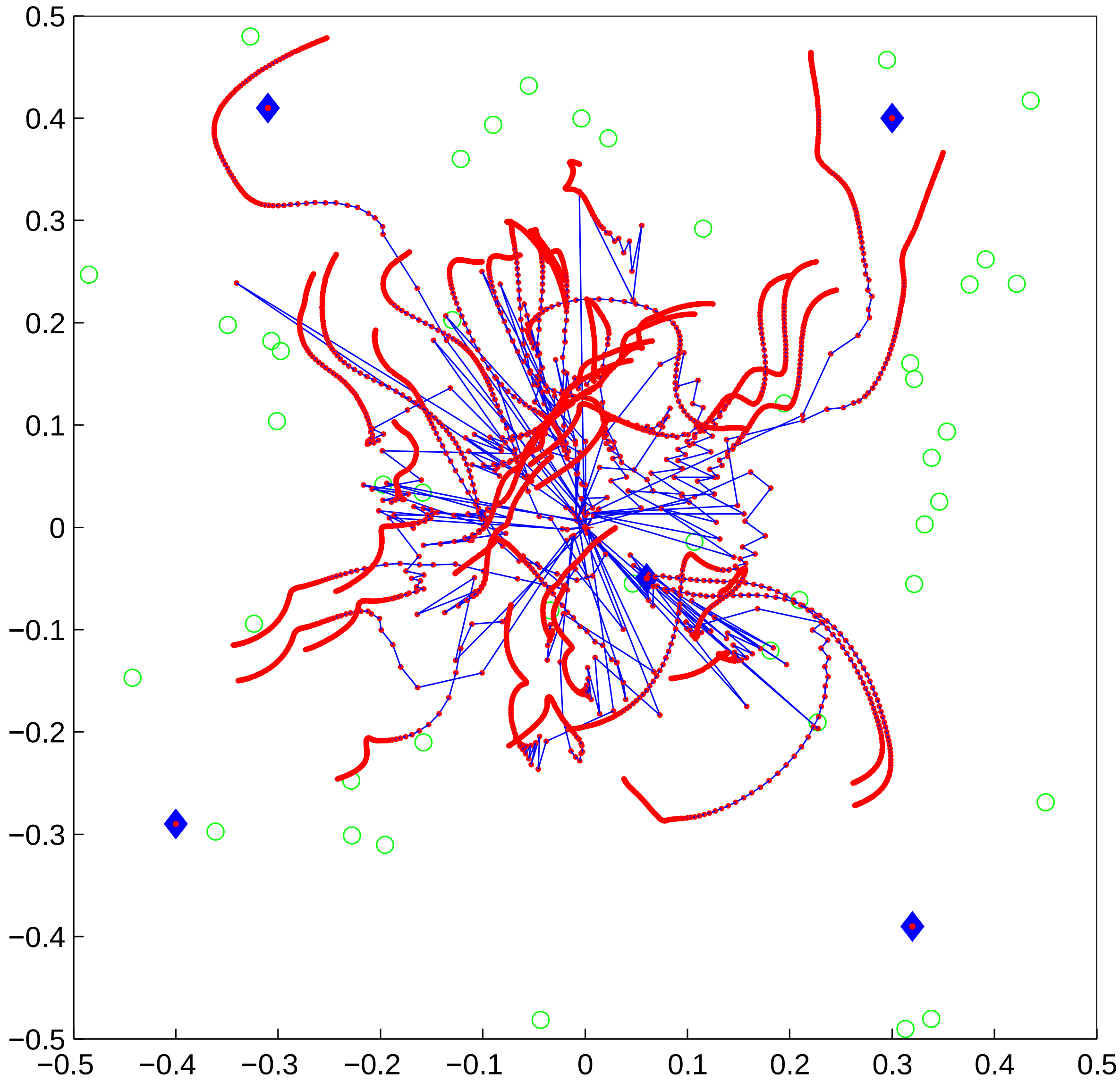
$$\min_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \mathbb{R}^2} \quad \sum_{(i,j)\in E_{ss}} \frac{1}{\sigma_{ij}^2}\left(\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - d_{ij}\right)^2 + \sum_{(i,j)\in E_{sa}} \frac{1}{\bar{\sigma}_{ij}^2}\left(\|\boldsymbol{x}_i - \boldsymbol{a}_j\| - \bar{d}_{ij}\right)^2,$$

where $E_{ss}$ & $E_{sa}$ are the sensor-to-sensor & sensor-to-anchor edge sets, resp.; $\sigma_{ij}^2$ & $\bar{\sigma}_{ij}^2$ are noise variances w.r.t. $d_{ij}$ & $\bar{d}_{ij}$, resp.

  - The ML-SNL problem is an unconstrained nonconvex problem.
  - One may handle the ML-SNL problem by applying gradient descent directly.
  - Alternatively, one can tackle the ML-SNL problem by SDR, which is achieved through a careful reformulation. This will be described in Part III.C.



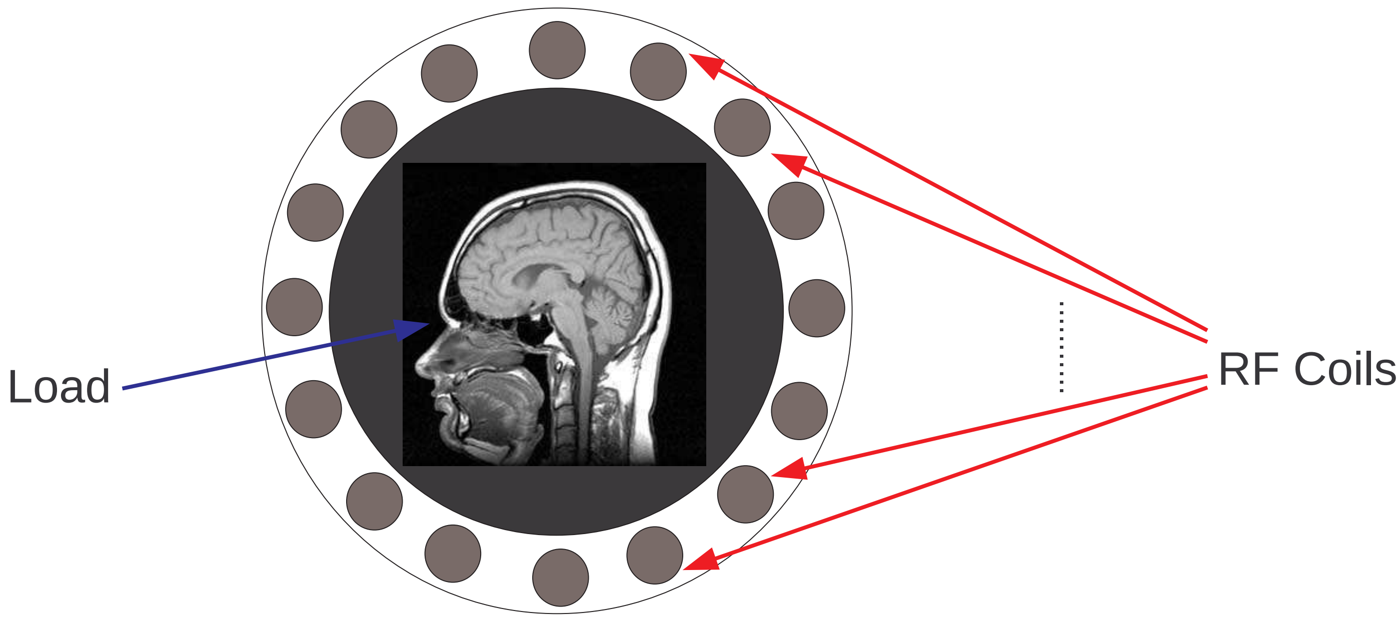

SDR (ML-SNL formulation), plus a 2nd-stage solution refinement by gradient descent. The distance measurements are noisy. ○: true sensor locations; ◊: anchor locations; ∗: SDR solution; — : gradient descent trajectory (50 iterations).




Gradient descent ML-SNL with a random starting point. ○: true sensor locations; ◊: anchor locations; — : gradient descent trajectory (50 iterations).

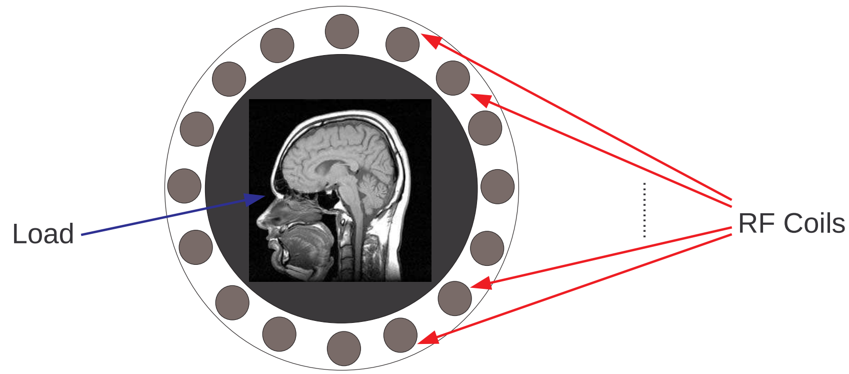
# Application: Transmit $B_1$ Shim in MRI

**Scenario:** 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.



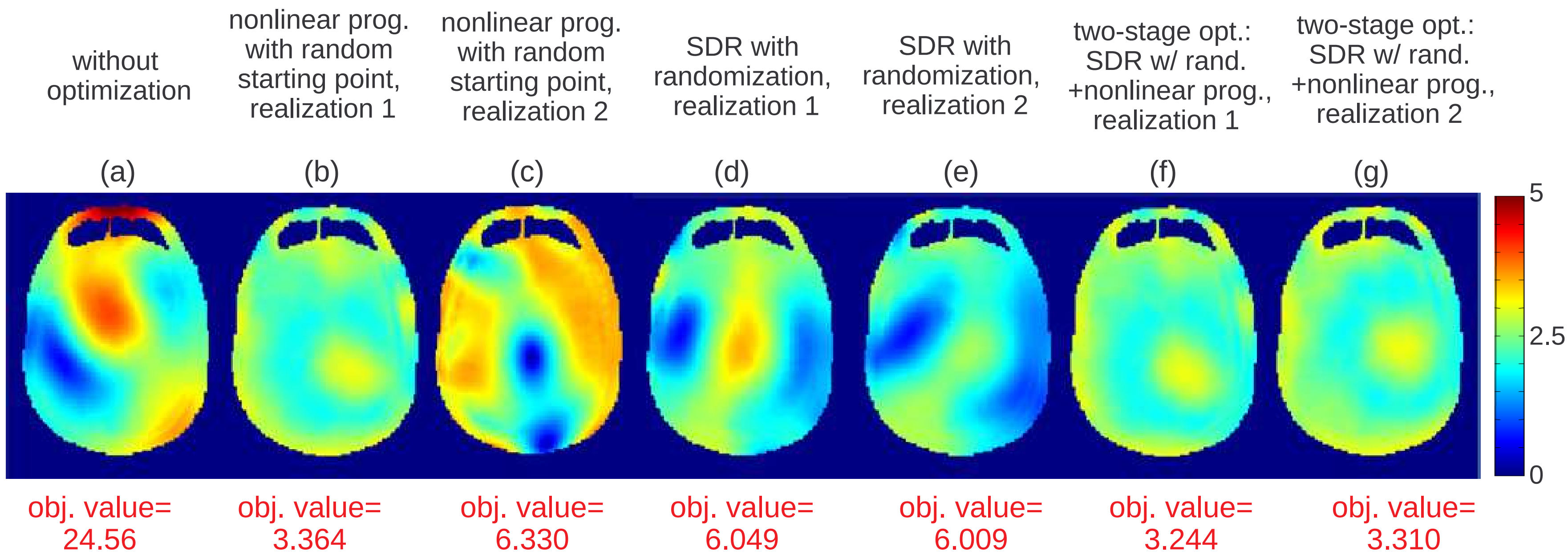

- Let $\boldsymbol{a}_i \in \mathbb{C}^n$, $i = 1, \ldots, m$, be the field response from the array to the $i$th pixel (MISO); i.e., the $i$th pixel receives a $B_1$ field of magnitude $|\boldsymbol{a}_i^T \boldsymbol{x}|$.

- **Problem:** Minimize the worst-case field magnitude difference

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

  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 $\rho$ 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 $\gamma$, 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 through it in this tutorial.

- Given a solution generation procedure, we are usually interested in its performance on arbitrary instances of (QCQP). Thus, the approximation ratio $\gamma$ 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.

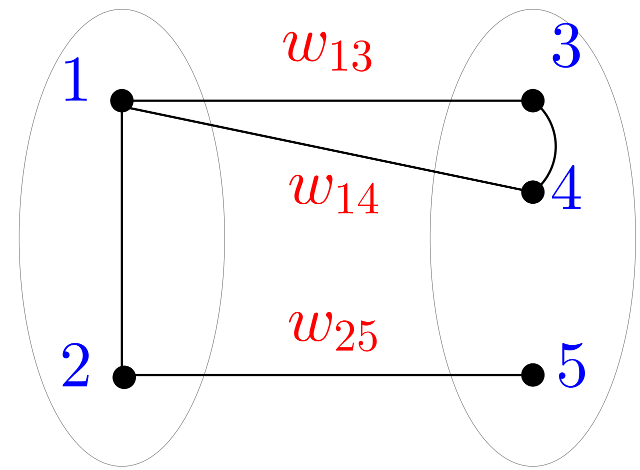
- The problem of proving approximation accuracies has been of great interest to optimization theorists, and it has enormous implications in practice.

# The Seminal Approx. Accuracy Result by Goemans & Williamson

- Consider

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

  with $\boldsymbol{C} \succeq \boldsymbol{0}$, $C_{ij} \leq 0$ for all $i \neq j$.




- Such a problem arises in 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

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

  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 $k$th 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.** ***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 $\boldsymbol{\xi}_\ell \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{X}^\star)$, 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

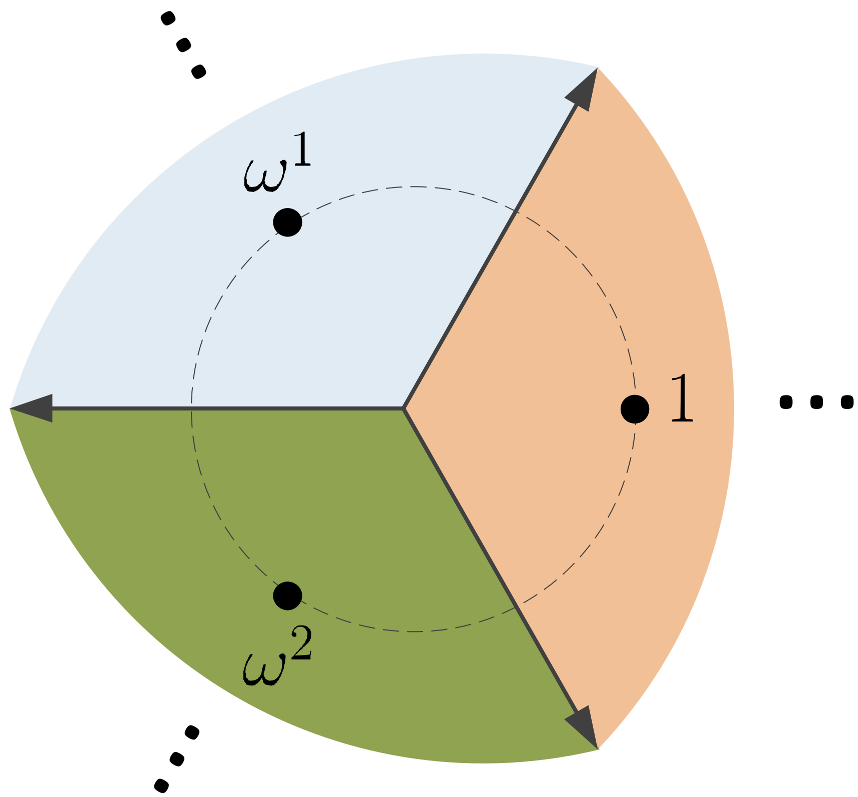
$$f(z) = \begin{cases} 1, & \arg(z) \in [-\pi/k, \pi/k), \\ \omega, & \arg(z) \in [\pi/k, 3\pi/k), \\ \vdots & \vdots \\ \omega^{k-1}, & \arg(z) \in [(2k-3)\pi/k, (2k-1)\pi/k); \end{cases}$$

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

$$|x_i| = 1, \quad i = 1, \ldots, n. \tag{$\dagger$}$$

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]**
  - distributed detection over multiple-access channels **[Banavar-Smith-Tepedelenlioğlu-Spanias'12]**

| problem | approx. accuracy $\gamma$; see (21)-(22) for def. | references |
|---|---|---|
| Boolean QP<br>$$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & x_i^2 = 1,\ i=1,\ldots,n \end{aligned}$$ | $$\gamma = \begin{cases} 0.87856, & \boldsymbol{C}\succeq\boldsymbol{0},\ C_{ij}\le 0\ \forall i\neq j \\ 2/\pi \simeq 0.63661, & \boldsymbol{C}\succeq\boldsymbol{0} \\ 1 \text{ (opt.)}, & C_{ij}\ge 0, \forall i\neq j \end{cases}$$ | Goemans-Williamson [2],<br>Nesterov [3], Zhang [6].<br>Relevant applications: [24]–[26] |
| Complex $k$-ary QP<br>$$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{C}^n} \quad & \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & x_i \in \{1,\omega,\ldots,\omega^{k-1}\}, \\ & i=1,\ldots,n \end{aligned}$$<br>where $\omega = e^{j2\pi/k}$, and $k>1$ is an integer. | For $\boldsymbol{C}\succeq\boldsymbol{0}$,<br>$$\gamma = \frac{(k\sin(\pi/k))^2}{4\pi}.$$<br>e.g., $\gamma=0.7458$ for $k=8$, $\gamma=0.7754$ for $k=16$. | Zhang-Huang [7],<br>So-Zhang-Ye [8].<br>Relevant applications: [27], [37] |
| Complex constant-modulus QP<br>$$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{C}^n} \quad & \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & \lvert x_i\rvert^2 = 1,\ i=1,\ldots,n \end{aligned}$$ | For $\boldsymbol{C}\succeq\boldsymbol{0}$,<br>$$\gamma = \pi/4 = 0.7854.$$<br>Remark: coincide with complex $k$-ary QP as $k\to\infty$. | Zhang-Huang [7],<br>So-Zhang-Ye [8]. |
| $$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{C}^n} \quad & \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & (\lvert x_1\rvert^2,\ldots,\lvert x_n\rvert^2)\in\mathcal{F} \end{aligned}$$<br>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 $\boldsymbol{C}\succeq\boldsymbol{0}$.<br><br>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]. |
| $$\begin{aligned} \max_{\boldsymbol{x}\in\mathbb{R}^n} \quad & \boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{x}^T\boldsymbol{A}_i\boldsymbol{x}\le 1,\ i=1,\ldots,m \end{aligned}$$<br>where $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_m\succeq\boldsymbol{0}$. | For any $\boldsymbol{C}\in\mathbb{S}^n$,<br>$$\gamma = \frac{1}{2\ln(2m\mu)}$$<br>where $\mu=\min\{m,\max_i \operatorname{rank}(\boldsymbol{A}_i)\}$. | Nemirovski-Roos-Terlaky [5].<br>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

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

  where $\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \succeq \boldsymbol{0}$. This arises in the study of multicast 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 ($\dagger$), 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:

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

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

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

| problem | approx. accuracy $\gamma$; see (18)-(19) for def. | references |
|---|---|---|
| $\min_{\boldsymbol{x}\in\mathbb{C}^n} \; \boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x}$ <br> s.t. $\boldsymbol{x}^H\boldsymbol{A}_i\boldsymbol{x} \geq 1,\; i=1,\ldots,m$ <br> where $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_m \succeq \boldsymbol{0}$. | $\gamma = 8m.$ <br> If the problem is reduced to the real-valued case, then <br> $\gamma = \dfrac{27m^2}{\pi}.$ | Luo-Sidiropoulos-Tseng-Zhang [9]; see also So-Ye-Zhang [71]. <br> Relevant applications: [29] |
| MIMO Detection <br> $\min_{\boldsymbol{x}\in\mathbb{R}^n} \; \lVert\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\rVert_2^2$ <br> s.t. $x_i^2 = 1,\; i=1,\ldots,n$ <br> 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 $\sigma^2$. | For $\sigma^2 \geq 60n$ (which corresponds to the low signal-to-noise ratio (SNR) region), with probability at least $1-3\exp(-n/6)$, <br> $\gamma \leq \dfrac{11}{2}.$ <br> For $\sigma^2 = \mathcal{O}(1)$ (which corresponds to the high SNR region), with probability at least $1-\exp(-\mathcal{O}(n))$, <br> $\gamma = 1,$ <br> i.e. the SDR is tight. | Kisialiou-Luo [67], So [69]. <br> Extensions: So [68], [69]. <br> Related: Jaldén-Ottersten [66]. <br> 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

- The SDR methodology introduced so far can be summarized as follows:

  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.

- It is natural to expect that the lower the rank of the SDP solution, the better the approximation.

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

- However, we can identify special cases where the SDP solution rank is low or even equal to one.

# Shapiro-Barvinok-Pataki (SBP) Result

- Consider the real-valued SDP

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

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

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

- 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 now consider the complex-valued SDP

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

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

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

As a direct corollary, we have

> For a complex-valued QCQP with $m \leq 3$, SDR is tight.

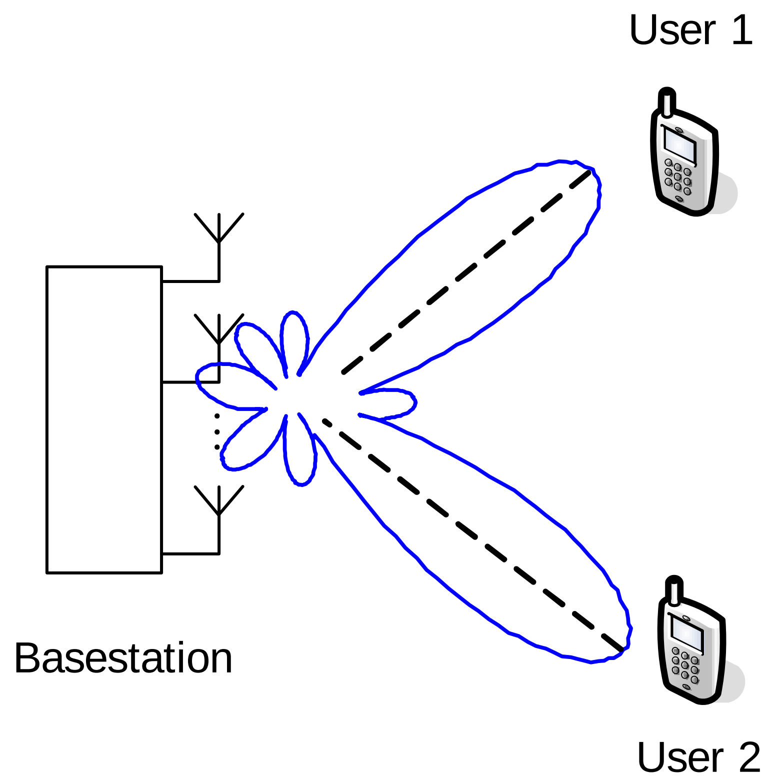
- A complex rank-1 decomposition algorithm for $m \leq 3$ is available **[Huang-Zhang'07]**.

# Application Revisited: Multicast Beamforming

- Recall the multicast beamforming problem:

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

$K$ 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

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

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

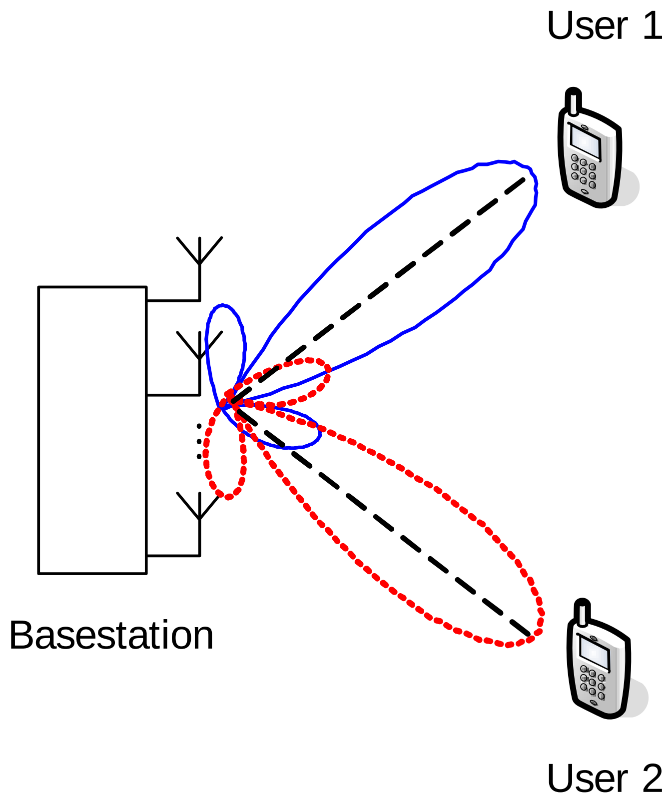
- We have the following generalization of the SBP result **[Huang-Palomar'09]**:

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

  Consequently,

  Suppose that an SDR solution $\{\boldsymbol{X}_i^\star\}_i$ satisfies $\boldsymbol{X}_i^\star \neq \boldsymbol{0}$ for all $i$. Then, the 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 aforementioned result, SDR solves ($\dagger$) optimally for any $\boldsymbol{R}_1, \ldots, \boldsymbol{R}_K \succeq \mathbf{0}$.

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

# Part III.A: Transmit Beamforming

# SDR in Transmit Beamforming

- Many forefront advances of SDR we see recently lie in transmit beamforming (BF) optimization.

From http://money.cnn.com

From http://www.telepresenceoptions.com

From http://www.slashgear.com

# Current Development of SDR in Transmit Beamforming

- Apart from unicast and multicast BF, we have seen numerous extensions:
  - multigroup multicast **[Karipidis-Sidiropoulos-Luo'08]**
  - cognitive radio BF **[Phan-Vorobyov-Sidiropoulos-Tellambura'09]**
  - relay beamforming
    - 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]**
    - interference neutralization **[Ho-Jorswieck'12]**
  - multicell coordinated beamforming **[Bengtsson-Ottersten'01]**, **[Dahrouj-Yu'10]**, **[Shen-Chang-Wang-Qiu-Chi'12]**
  - secrecy beamforming **[Liao-Chang-Ma-Chi'11]**
  - energy harvesting **[Xu-Liu-Zhang'13]**, **[Chalise-Ma-Zhang-Suraweera-Amin'13]**

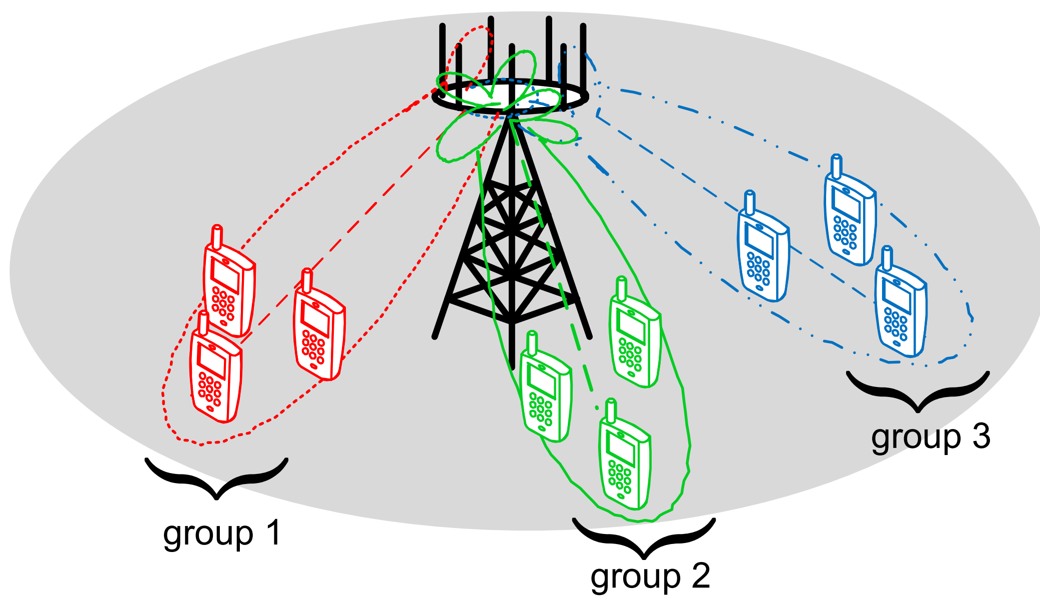
# Overview

- Our focus:
  - QCQP-SDR perspective on various transmit BF problems.
  - A glimpse of some nice formulations.

- What we will not go through:
  - alternative solution approaches and comparison
    * 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]**

# 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 $m$th 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 $m$th 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:**

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

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

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

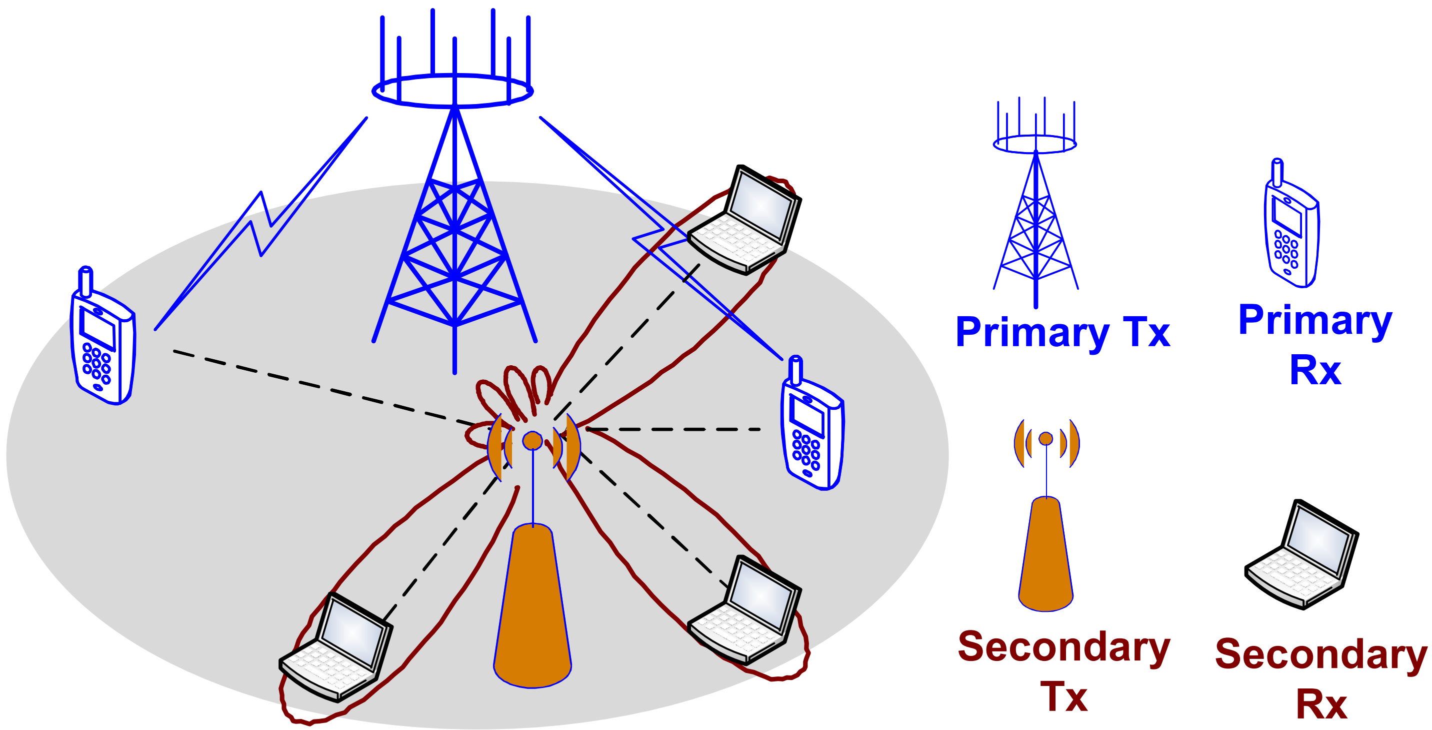
- A separable QCQP with $M$ variables, $\boldsymbol{w}_1,\ldots,\boldsymbol{w}_M$, and $\sum_{m=1}^{M} K_m$ constraints.

- By the SBP rank reduction result in Part II, SDR has rank-$1$ solution (and solves the BF problem optimally) when
  - $K_1 \leq 3,\ K_m = 1\ \forall m \neq 1$ (one group serving $\leq 3$ users, the others $1$ user);
  - $K_1 \leq 2,\ K_2 \leq 2,\ K_m = 1\ \forall m \neq 1,2$ (two groups serving $\leq 2$ users, the others $1$ user).

- For non-rank-$1$ instances, solution approx. can be done by a (more sophisticated) Gaussian randomization procedure **[Karipidis-Sidiropoulos-Luo'08]**.

# Cognitive Radio (CR) Beamforming

- **Goal:** Access the channel owned by primary users through spectrum sharing.
- **Scenario:** MISO downlink with the CR (or secondary) system.




- **Idea:** Avoid excessive interference to the primary users through tx. opt.

- CR spectrum-sharing model:
  - $K$ secondary users (SUs), $L$ single-antenna primary users (PUs)
  - tx. and rx. model for SUs: same as the previous multicast or unicast model
  - interference to the $l$th PU given by

$$|\boldsymbol{g}_l^H \boldsymbol{w}|^2,$$

    where $\boldsymbol{g}_l$ is the channel from the secondary transmitter to the $l$th PU
  - known CSIT from the secondary transmitter to the PUs
- Design for the multicast case **[Phan-Vorobyov-Sidiropoulos-Tellambura'09]**:

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

  where $\boldsymbol{G}_l$ is the CSIT of $l$th PU (defined in the same way as $\boldsymbol{R}_k$); $\delta_l$ is the tolerable interference level to the $l$th PU; $\gamma$ is SUs' SNR requirement.
  - By the SBP result, SDR is tight when $K \leq 2$, $L = 1$ ($\leq 2$ SUs, $1$ PU).

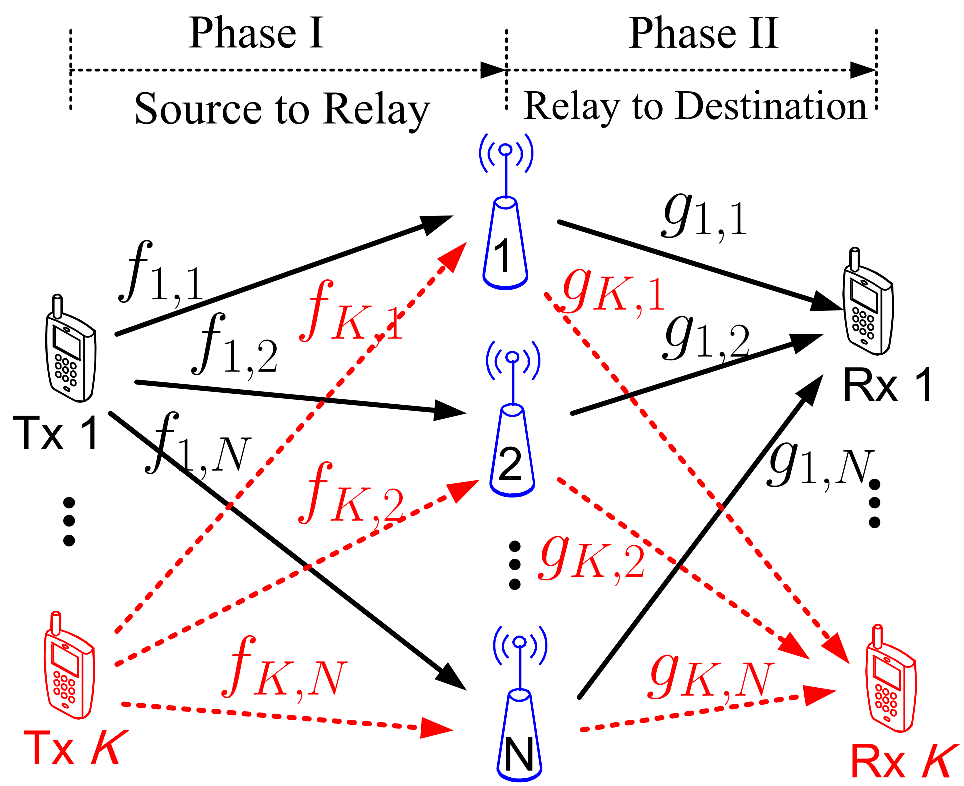
- Design for the unicast case (see, e.g., **[Zhang-Liang-Cui'10]**):

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

  - A separable QCQP with $K$ variables and $K + L$ constraints.
  - By the SBP result, SDR is tight if $L \leq 2$ (two PUs or less).
  - Remark: For instant. CSIT with SUs, SDR can be shown to be rank-$1$ optimal for any $L$. Alternatively, it can be reformulated as an 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 $\sigma_{v,i}^2$.

- Assuming instant. CSIT (for ease of illustration), we have

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

- **Problem:** Let $\boldsymbol{w} = [\ w_1, \ldots, w_N\ ]^T \in \mathbb{C}^N$. Solve

$$\min_{\boldsymbol{w}} \ \mathrm{E}\{\|\boldsymbol{x}(t)\|^2\} = \boldsymbol{w}^H \boldsymbol{C} \boldsymbol{w}$$

$$\text{s.t. } \text{SINR}_i = \frac{\boldsymbol{w}^H \boldsymbol{A}_i \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{B}_i \boldsymbol{w} + \sigma_{v,i}^2} \geq \gamma_i, \ \ i = 1, \ldots, K,$$

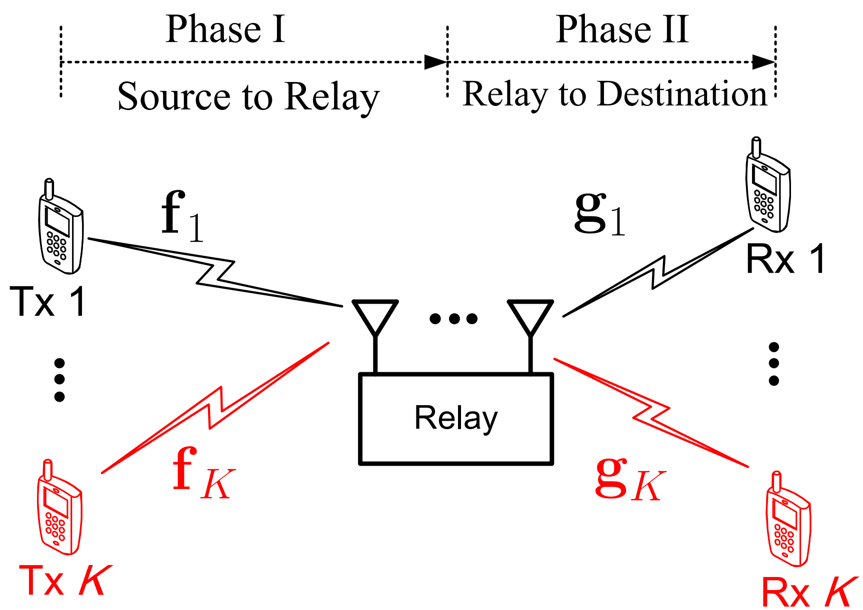
  where $\boldsymbol{A}_i = (\boldsymbol{f}_i^* \odot \boldsymbol{g}_i)(\boldsymbol{f}_i^* \odot \boldsymbol{g}_i)^H$, $\boldsymbol{B}_i = \sum_{k\neq i} (\boldsymbol{f}_k^* \odot \boldsymbol{g}_i)(\boldsymbol{f}_k^* \odot \boldsymbol{g}_i)^H + \mathrm{Diag}(|g_{i,1}|^2 \sigma_{n,1}^2, \ldots, |g_{i,N}|^2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

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

  – Remark: While this relay application has some unicast flavor— i.e., one data stream for one user— it does not imply that the same SDR tightness result in standard unicast BF holds for the relay BF problem.

# 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 $\boldsymbol{w} = \mathrm{vec}(\boldsymbol{W}) \in \mathbb{C}^{N^2}$.

- The design problem (after some careful derivations):

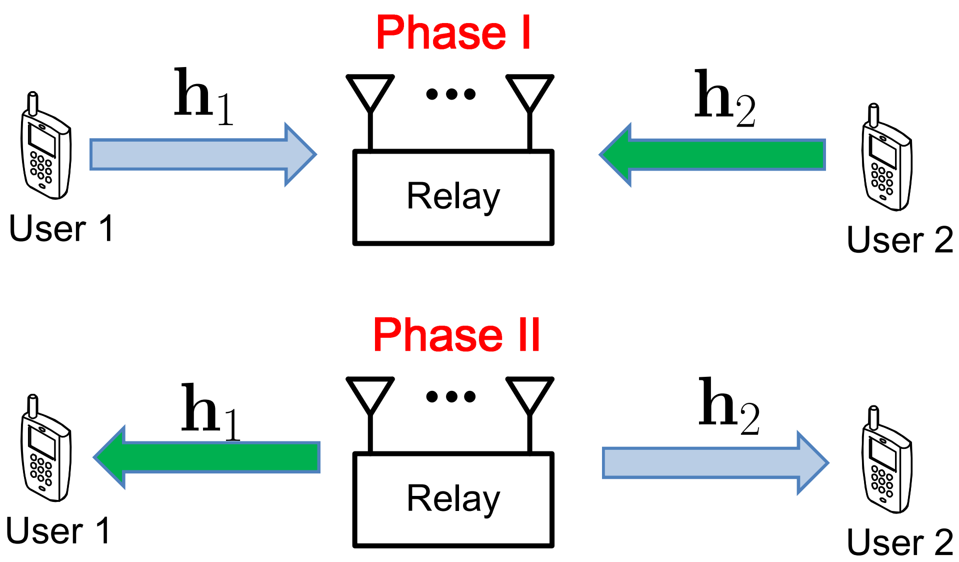
$$\begin{aligned} \min_{\boldsymbol{w}} \; & \mathrm{E}\{\|\boldsymbol{x}(t)\|^2\} = \boldsymbol{w}^H \boldsymbol{C} \boldsymbol{w} \\ \text{s.t. } & \mathsf{SINR}_i = \frac{\boldsymbol{w}^H \boldsymbol{A}_i \boldsymbol{w}}{\boldsymbol{w}^H \boldsymbol{B}_i \boldsymbol{w} + \sigma_{v,i}^2} \geq \gamma_i,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

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, SDR is tight 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.

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

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

- The design problem

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

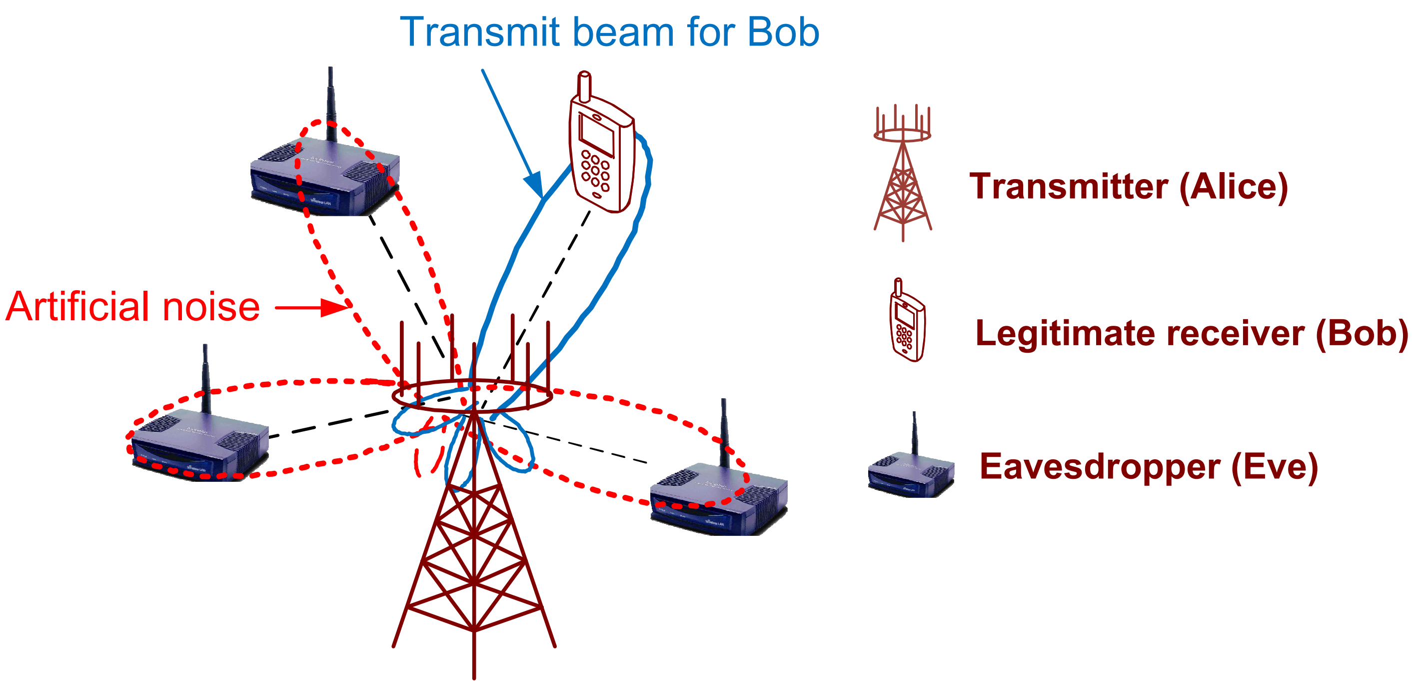
  can be converted to a 2-constraint QCQP, by applying $\boldsymbol{w} = \mathrm{vec}(\boldsymbol{W}) \in \mathbb{C}^{N^2}$ (the same way as in the last example).

- Hence, SDR solves the two-relaying BF problem optimally.

# Physical-Layer Security

- **Scenario:** One intended user (Bob) and multiple illegitimate users (Eves).

- **Goal:** Block eavesdropping by degrading illegitimate users' "QoS".




- **Idea:** Use spatially selective artificial noise (AN) to jam illegitimate users **[Swindlehurst'09]**, **[Liao-Chang-Ma-Chi'11]**.

- tx. signal: $\boldsymbol{x}(t) = \boldsymbol{w}s(t) + \boldsymbol{z}(t)$, where $\boldsymbol{z}(t) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$ is AN. SINRs:

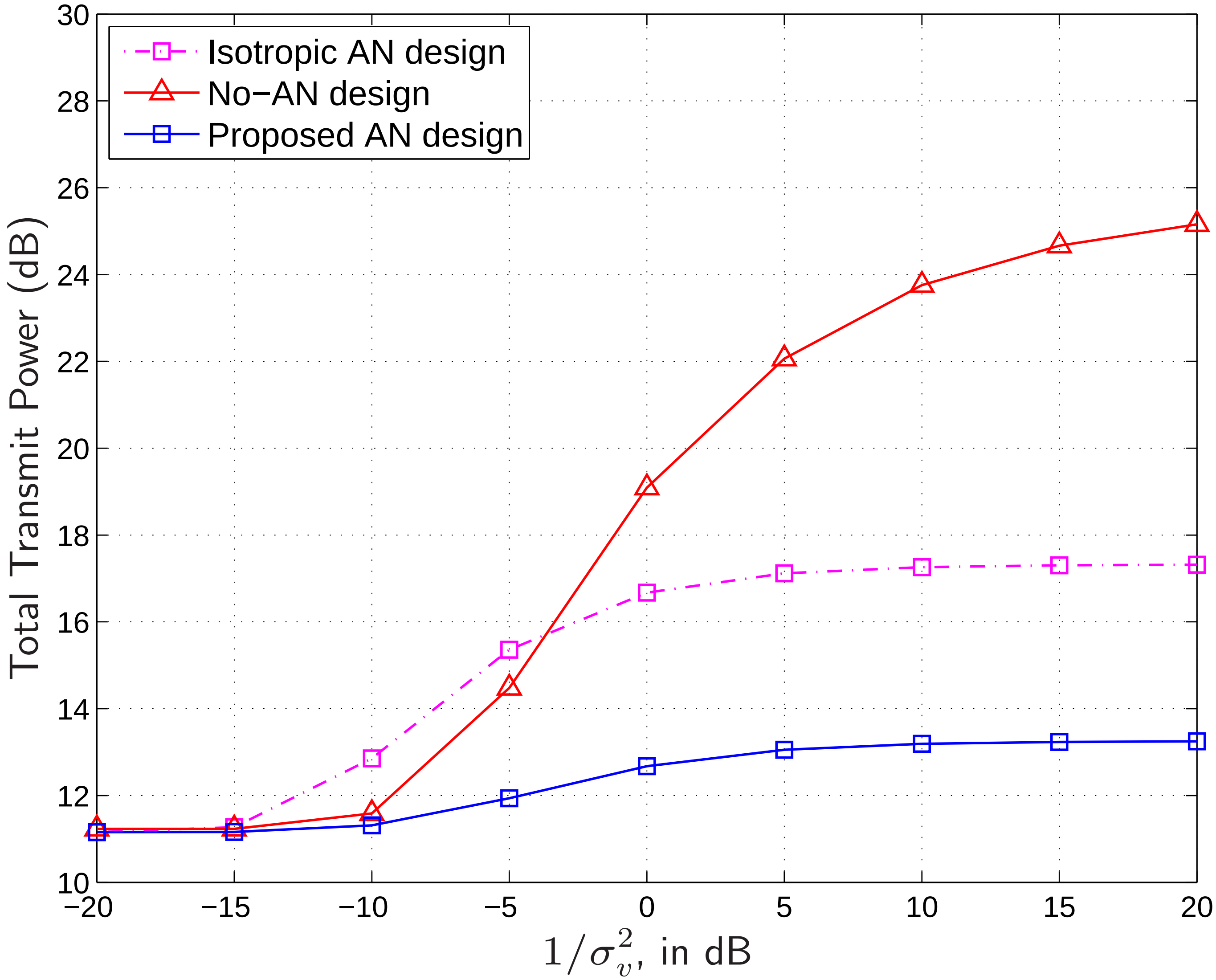
$$\mathsf{SINR}_{\mathsf{Bob}} = \frac{|\boldsymbol{h}^H \boldsymbol{w}|^2}{\mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{h}\boldsymbol{h}^H) + \sigma_n^2}, \quad \mathsf{SINR}_{\mathsf{Eve},i} = \frac{|\boldsymbol{g}_i^H \boldsymbol{w}|^2}{\mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{g}_i\boldsymbol{g}_i^H)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

where $\boldsymbol{h}$ and $\boldsymbol{g}_i$ are the channels of Bob and Eve $i$, resp.

- **Problem:** Given an SINR specification $(\gamma, \beta)$, solve

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

  - $(\dagger)$ provides a secrecy rate guarantee $\log(1+\gamma) - \log(1+\beta)$.
  - $(\dagger)$ can be handled by SDR, by replacing $\boldsymbol{W} = \boldsymbol{w}\boldsymbol{w}^H$ with $\boldsymbol{W} \succeq \mathbf{0}$.
  - By the SBP result, you can (immediately!) declare that SDR is tight for $L \leq 2$.
  - By exploiting specific problem structures of $(\dagger)$, it is proven that SDR always gives rank-1 solution with $\boldsymbol{W}$ for any $L$ **[Liao-Chang-Ma-Chi'11]**.

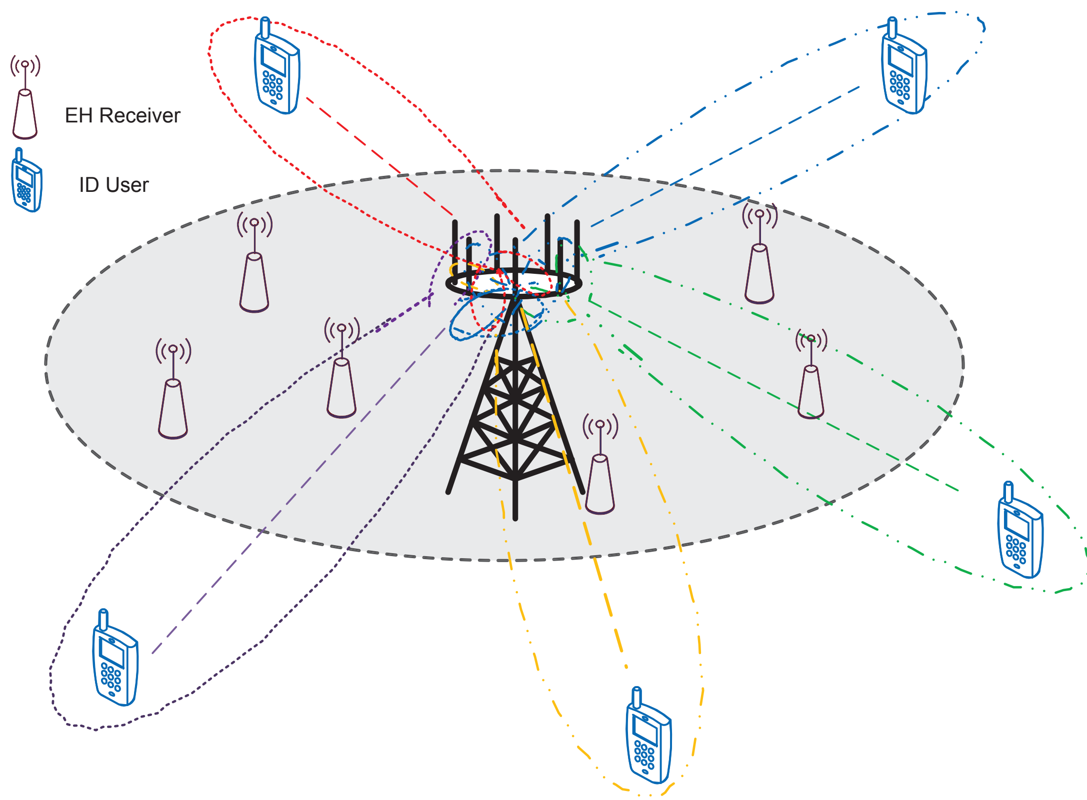



Transmit power performance of various secret BF designs. $N_t = 4$; $L = 3$; $\sigma_n^2 = 0$dB; $\gamma = 10$dB; $\beta = 0$dB. No-AN design refers to (†) without AN. Isotropic AN design refers to a closed-form design in which $\boldsymbol{w} = \sqrt{\alpha P_{\sf max}}\boldsymbol{h}/\|\boldsymbol{h}\|$, $\boldsymbol{\Sigma} = (1-\alpha)P_{\sf max}(\boldsymbol{I} - \boldsymbol{h}\boldsymbol{h}^H/\|\boldsymbol{h}\|^2)$, with $P_{\sf max}$ being the total tx power and $0 < \alpha \leq 1$ being a power allocation factor.

# Energy Harvesting

- **Scenario:** Unicast multiuser MISO downlink, with energy harvesting (EH) receivers that can harvest energy from radio signals.

- **Goal:** Simultaneous information transmission and wireless power transfer via BF **[Xu-Liu-Zhang'13]**, **[Chalise-Ma-Zhang-Suraweera-Amin'13]**.



- System model:
  - $K$ information decoding (ID) users, $L$ EH receivers.
  - tx. signal model:

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

    where
    * $s_i(t)$ & $\boldsymbol{w}_i$ are ID user $i$'s data steam & beamvector, resp.;
    * every $u_i(t)$ is an energy-carrying (& no-info.) signal for transfering energy to EH receivers; $\boldsymbol{v}_i$ is the corresponding energy beamvector.
  - Power harvested by EH receiver $i$:

$$\mathsf{P}_i^{\mathsf{EH}} = \zeta_i \cdot \mathrm{E}\{|y_i^{\mathsf{EH}}(t)|^2\} = \zeta_i \cdot \left( \sum_{i=1}^{M} \boldsymbol{v}_i^H \boldsymbol{G}_i \boldsymbol{v}_i + \sum_{i=1}^{K} \boldsymbol{w}_i^H \boldsymbol{G}_i \boldsymbol{w}_i \right),$$

    where $y_i^{\mathsf{EH}}(t)$ is EH receiver $i$'s rx. signal; $\zeta_i$ the EH efficiency; $\boldsymbol{G}_i$ the CSIT.
  - ID users' SINRs: Every ID user is assumed to know $u_i(t)$ and can cancel them from the rx. signal. By letting $\boldsymbol{R}_i$ to be the CSIT of ID user $i$, the SINRs are

$$\mathsf{SINR}_i^{\mathsf{ID}} = \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2}, \quad i = 1, \ldots, K.$$

- **Problem:** Given ID users' SINR requirements and EH receivers' requirements, denoted by $\gamma_1, \ldots, \gamma_K$ and $\beta_1, \ldots, \beta_L$, resp., solve

$$
\begin{aligned}
\min_{\{\boldsymbol{w}_i\},\{\boldsymbol{v}_i\}} \ & \sum_{i=1}^{K} \|\boldsymbol{w}_i\|^2 + \sum_{i=1}^{M} \|\boldsymbol{v}_i\|^2 \\
\text{s.t. } & \mathsf{SINR}_i^{\mathsf{ID}} \geq \gamma_i, \ i = 1, \ldots, K, \\
& \mathsf{P}_i^{\mathsf{EH}} \geq \beta_i, \ i = 1, \ldots, L.
\end{aligned}
$$

- At first sight, one may do the following SDR formulation: Let

$$
\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H, i = 1, \ldots, K, \quad \boldsymbol{V}_i = \boldsymbol{v}_i \boldsymbol{v}_i^H, \quad i = 1, \ldots, L,
$$

  and then "SDR" them.

- Specific problem structures can be exploited to formulate a simpler SDR.

- Recall the design problem

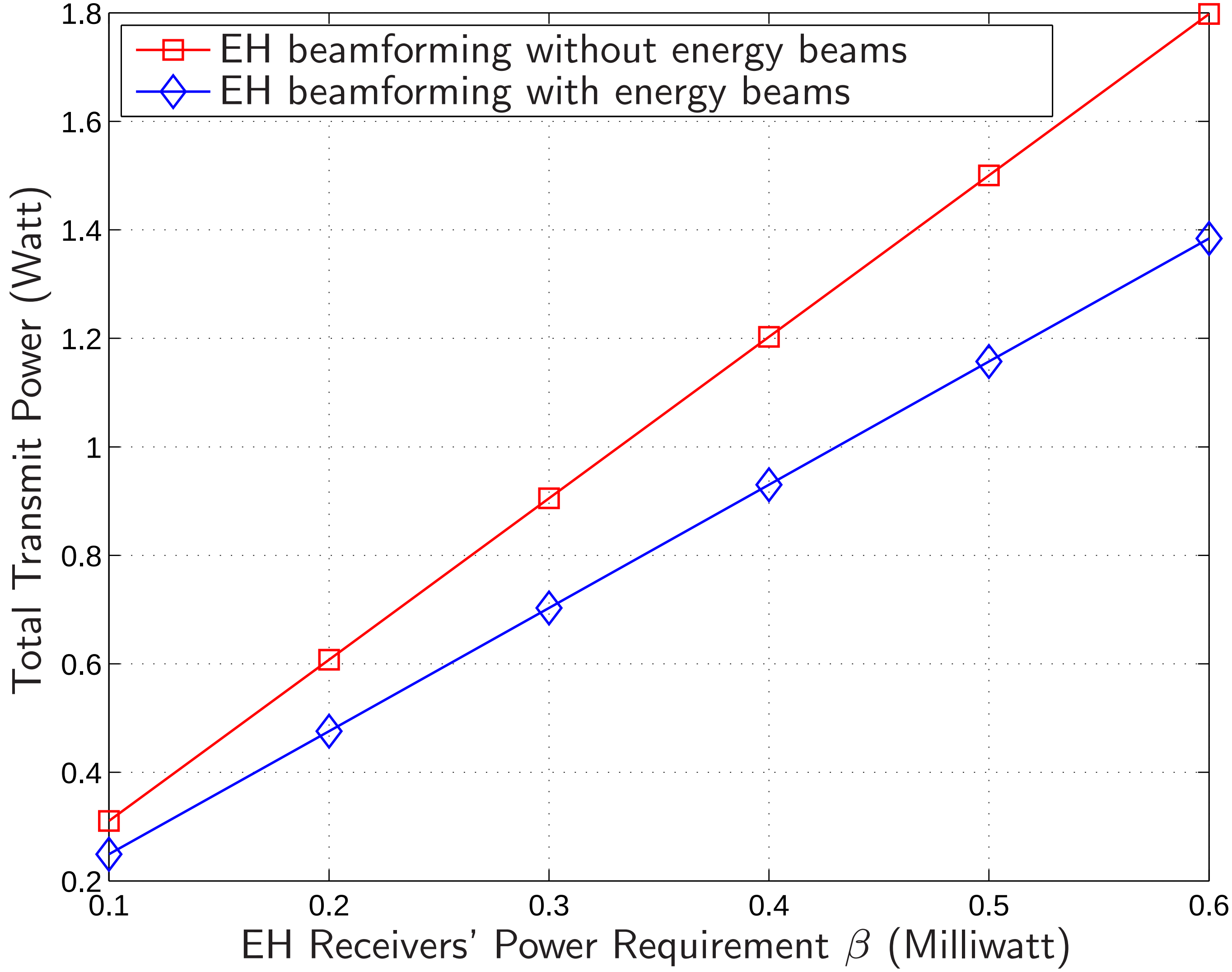
$$\begin{aligned} \min_{\{\boldsymbol{w}_i\},\{\boldsymbol{v}_i\}} \quad & \textstyle\sum_{i=1}^K \|\boldsymbol{w}_i\|^2 + \mathrm{Tr}((\sum_{i=1}^M \boldsymbol{v}_i \boldsymbol{v}_i^H)) \\ \text{s.t.} \quad & \mathsf{SINR}_i^{\mathsf{ID}} = \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l \neq i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2} \geq \gamma_i, \ i = 1, \ldots, K, \\ & \mathsf{P}_i^{\mathsf{EH}} = \zeta_i \cdot \left( \mathrm{Tr}\left( \boldsymbol{G}_i (\textstyle\sum_{i=1}^M \boldsymbol{v}_i \boldsymbol{v}_i^H) \right) + \mathrm{Tr}\left( \boldsymbol{G}_i (\sum_{i=1}^K \boldsymbol{w}_i \boldsymbol{w}_i^H) \right) \right) \geq \beta_i, \\ & \qquad i = 1, \ldots, L. \end{aligned}$$

- Let

$$\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H, i = 1, \ldots, K, \quad \boldsymbol{V} = \textstyle\sum_{i=1}^M \boldsymbol{v}_i \boldsymbol{v}_i^H,$$

where one should note that $\boldsymbol{V} \succeq \boldsymbol{0}, \mathrm{rank}(\boldsymbol{V}) \leq M$. Then, "SDR" $\{\boldsymbol{W}_i\}, \boldsymbol{V}$.

- The resulting SDR has less design variables than the previously mentioned SDR. Also, by the SBP result,
  - SDR has rank-$1$ solution w.r.t. $\{\boldsymbol{W}_i\}$ for $L \leq 2$ (two EH receivers or less);
  - SDR has rank $r \leq 1$ solution w.r.t. $\boldsymbol{V}$ for $L \leq 2$; i.e., when there are two EH receivers or less, it suffices to use $M = 1$, or one energy beam.

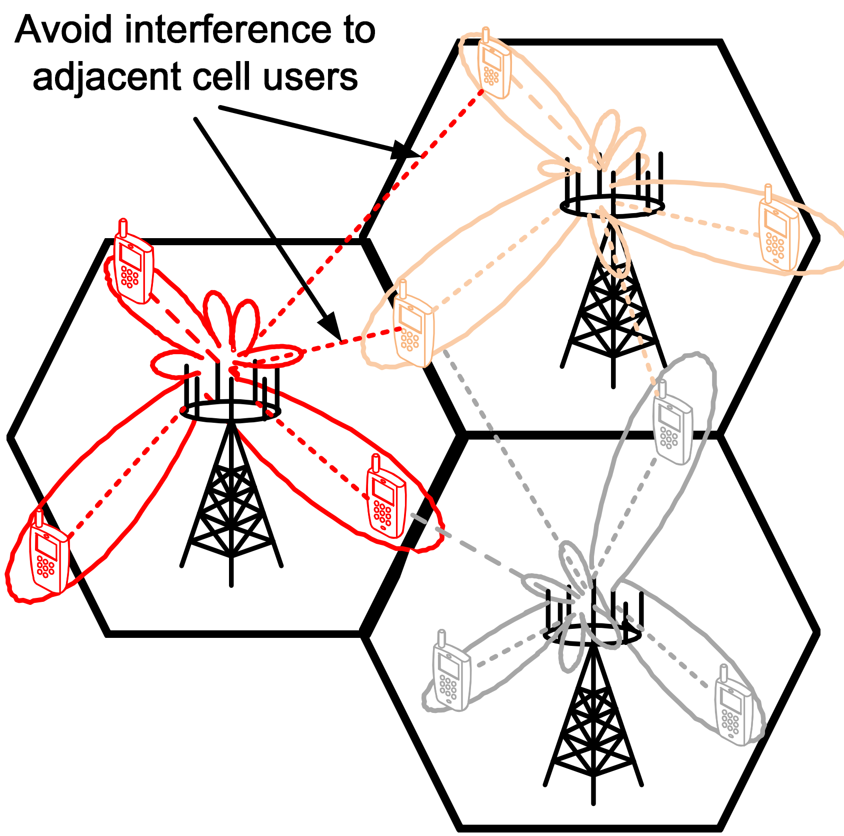



Transmit power performance of EH beamforming with respect to the EH receivers' power requirement. "EH beamforming with energy beams" refers to the EH formulation in the previous slide, while "EH beamforming without energy beams" refers to the same formulation without $\boldsymbol{v}_i$'s. $N_t = 8$, $K = 6$, $L = 2$, $\gamma_1 = \cdots = \gamma_K = 5$dB, $\beta = \beta_1 = \beta_2$, $\zeta_1 = \zeta_2 = 0.5$.

# Multicell Coordinated Beamforming

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

- **Scenario:** Unicast MISO downlink in a multicell scale **[Dahrouj-Yu'10]**, **[Bengtsson-Ottersten'01]**.



- tx. signal of $i$th cell:

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

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

- rx. signal of user $j$ in the $i$th 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 $m$th cell to user $j$ in the $i$th cell.

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

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

- **Problem:**

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

- While ($\dagger$) looks complicated, one can observe that
  - ($\dagger$) is a QCQP with $\sum_{i=1}^{N} K_i$ variables & $\sum_{i=1}^{N} K_i$ constraints;
  - by the SBP result, ($\dagger$) can be optimally solved by SDR.

- A recent direction: **Distributed multicell coordinated BF**
  - practically desirable, free from centralized opt. (requires a central station)
  - in the SDR context, the challenge is the same as solving SDP distributively
  - can be achieved by application of distributed opt. methods, e.g., alternating direction method of multipliers (ADMM) **[Shen-Chang-Wang-Qiu-Chi'12]**

- In ADMM, the idea is to reformulate the SDR of (†) in a consensus opt. form:

$$
\min_{\substack{\{t_{i,j},u_{m,i,j}\}, \\ \{u^{\text{public}}_{m,i,j}\}}} \sum_{i=1}^{N} \left[ \begin{array}{ll} \displaystyle\min_{\{\boldsymbol{W}_{i,j}\}_j} & \displaystyle\sum_j \mathrm{Tr}(\boldsymbol{W}_{i,j}) \\ \text{s.t.} & \dfrac{\mathrm{Tr}(\boldsymbol{R}_{i,i,j}\boldsymbol{W}_{i,j})}{\sum_{l\neq j}\mathrm{Tr}(\boldsymbol{R}_{i,i,j}\boldsymbol{W}_{i,l}) + t_{i,j} + \sigma^2_{i,j}} \geq \gamma_{i,j}, \ \forall i,j, \\ & \boldsymbol{W}_{i,1},\ldots,\boldsymbol{W}_{i,K_i} \succeq \boldsymbol{0}, \\ & u_{i,m,j} = \sum_n \mathrm{Tr}(\boldsymbol{R}_{i,m,j}\boldsymbol{W}_{i,n}), \ \forall m \neq i, j \end{array} \right]
$$

$$
\begin{aligned}
\text{s.t. } & u_{m,i,j} = u^{\text{public}}_{m,i,j}, \ \forall m,i,j, m\neq i, \\
& t_{i,j} = \textstyle\sum_{m\neq i} u^{\text{public}}_{m,i,j}, \ \forall i,j,
\end{aligned}
$$

where $t_{i,j}$ is the sum intercell interference (ICI) from other cells to user $j$ in cell $i$; $u_{m,i,j}$ (resp. $u^{\text{public}}_{m,i,j}$) is local (resp. public) ICI from cell $m$ to user $j$ in cell $i$.

- Then, one can apply ADMM to solve the above problem distributively (details skipped). In layman terms, at each iteration ADMM does the following:
  - Each cell solves a single-cell problem, but with ICI awareness.
  - Cells exchange their local ICI info. and update public ICI info.

# Other Design Formulations

- For tutorial purposes, we so far considered only the QoS-constrained power minimization design formulation.

- SDR can also handle design formulations such as
  - max-min fairness;
  - rate profile problems;
  - joint BF and user selection **[Matskani-Sidiropoulos-Luo-Tassiulas'08]**, **[Wai-Ma'12]**.

- Can SDR be employed to tackle even more challenging design formulations, particularly, the (NP-hard) sum rate maximization (SRM) problem?
  - Tricky, if not impossible...
  - Doable when we consider a harder version of SRM, namely, discrete SRM (DSRM) **[Wai-Li-Ma'13]**.

# Discrete Sum Rate Maximization

- **Scenario:** Unicast multiuser MISO downlink.
- **Problem**: Maximize the discretized sum rate under the total power constraint:

$$\begin{aligned} \max_{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_K} \quad & \textstyle\sum_{i=1}^K \varphi\left(\log_2(1+\mathsf{SINR}_i)\right) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^K \|\boldsymbol{w}\|^2 \le P, \end{aligned} \qquad \text{(DSRM)}$$

where $\mathsf{SINR}_i = \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l\neq i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2}$ with $\boldsymbol{R}_i = \boldsymbol{h}_i \boldsymbol{h}_i^H$ (instant. CSIT);

$$\varphi(r) = \begin{cases} R^M & \text{if } R^M \le r, \\ \vdots & \\ R^1 & \text{if } R^1 \le r < R^2, \\ 0 & \text{if } 0 \le r < R^1, \end{cases}$$

with $0 < R^1 < \cdots < R^M$ being the supported rate values; $P$ is the power limit.

- DSRM is motivated by finite rate constraints in practical modulation and coding schemes **[Cheng-Philipp-Pesavento'12]**.

- DSRM subsumes joint BF and user selection, wherein $M = 1$.

- Consider the simple case of $M = 1$, where $\varphi(r) = 0$ if $0 \le r < R^1$ & $\varphi(r) = R^1$ if $R^1 \le r$. Since

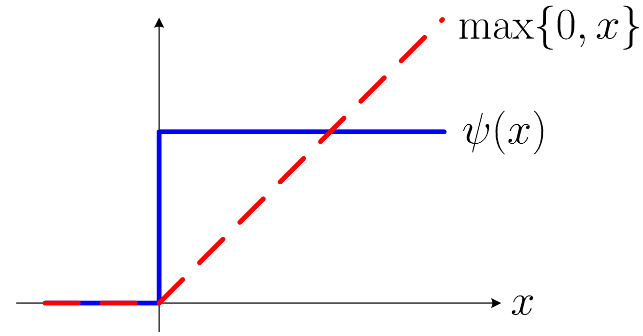
$$R^1 \le \log_2(1 + \mathsf{SINR}_i) \Longleftrightarrow 2^{R^1} \le 1 + \frac{\boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i}{\sum_{l \ne i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l + \sigma_i^2}$$

$$\Longleftrightarrow f_i(\{\boldsymbol{w}_l\}) \triangleq \sum_{l \ne i} \boldsymbol{w}_l^H \boldsymbol{R}_i \boldsymbol{w}_l - \frac{1}{2^{R^1} - 1} \boldsymbol{w}_i^H \boldsymbol{R}_i \boldsymbol{w}_i + \sigma_i^2 \le 0,$$

we can write

$$\varphi(\log_2(1 + \mathsf{SINR}_i)) = R^1 - R^1 \cdot \psi(f_i(\{\boldsymbol{w}_l\})),$$

where $\psi(x)$ is the unit step function.




- **Idea:** Apply the approx. (commonly seen in compressive sensing)

$$\varphi(\log_2(1 + \mathsf{SINR}_i)) \approx R^1 - R^1 \cdot \max\{0, f_i(\{\boldsymbol{w}_l\})\}.$$

Such an idea can be extended to $M > 1$.

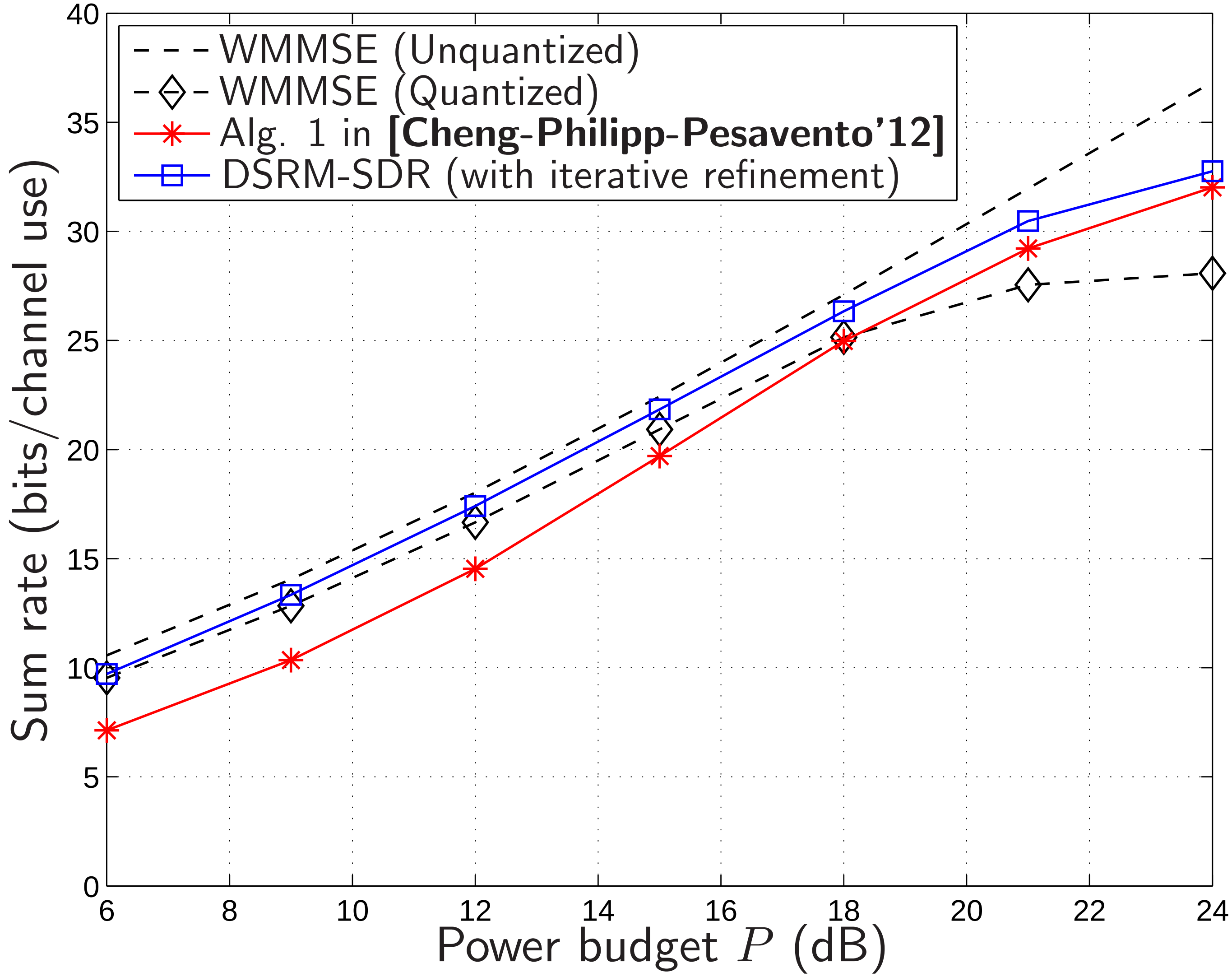
- Using the aforementioned step-to-hinge approx. and then SDR, (DSRM) can be approximated by

$$\max_{\{\boldsymbol{W}_l\}} \sum_{i=1}^{K}\sum_{m=1}^{M}(R^{i-1}-R^i)\max\left\{0, R^i + \frac{R^i}{\sigma_i^2}\mathrm{Tr}\left(\boldsymbol{R}_i\left(\sum_{l\neq i}\boldsymbol{W}_l - \frac{1}{2^{R^i-1}}\boldsymbol{W}_i\right)\right)\right\}$$
$$\text{s.t. } \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K \succeq \mathbf{0},\;\; \textstyle\sum_{i=1}^{K}\mathrm{Tr}(\boldsymbol{W}_i)\leq P, \qquad \text{(DSRM-SDR)}$$

where $R^0 = 0$.

- (DSRM-SDR) is convex.

- Technical remarks:
  - (DSRM-SDR) guarantees rank-one optimal solution with $\{\boldsymbol{W}_l\}$ if we add a small regularization term $-\epsilon\sum_{i=1}^{K}\mathrm{Tr}(\boldsymbol{W}_i)$, $\epsilon > 0$, in the objective.
  - However, (DSRM-SDR) is *not* tight with its discrete rate approx. An iterative refinement procedure is employed to generate an approx. solution; see **[Wai-Li-Ma'13]** for details.




Sum rate performance with respect to the total power limit $P$. $N_t = 6$, $K = 8$, & the discrete rate set follows that in 3GPP LTE standard ($M = 15$, from $R^1 = 0.15$bits to $R^M = 5.55$bits). WMMSE refers to the SRM algorithm in **[Shi-Razaviyayn-Luo-He'11]**, which does not take into account finite rate constraints.

# Part III.B: Advanced Topics in Transmit Beamforming

## Topic 1. Rank-Two Beamforming

# Motivation

- Recall that the rationale of SDR is to use a rank-unconstrained SDP to find a rank-$1$ solution $\boldsymbol{W}$, which can be physically realized by BF.

- A natural question arises: **Can we do rank-$2$ SDR?**

- Suppose that we want to do this. Then, there are two challenges to tackle:
  - From a communication viewpoint, we never said BF is the only way to go. The question is how to design an alternative transmit scheme.
  - From an optimization viewpoint, how to proceed with solution generation, and what is its theoretical performance?

- **Solution:**
  - A combination of BF and the Alamouti space-time block code.
  - The SDP rank reduction theory in **[So-Ye-Zhang'08]**.

# The Alamouti Space-Time Block Code

- Let $\boldsymbol{s} = [\ s_1, s_2\ ]^T$. The Alamouti code is

$$\boldsymbol{C}(\boldsymbol{s}) = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix},$$

  which is arguably the most famous among space-time codes.

- Features:
  - orthogonal: $\boldsymbol{C}(\boldsymbol{s})\boldsymbol{C}^H(\boldsymbol{s}) = \|\boldsymbol{s}\|^2\boldsymbol{I}$
  - easy to detect $\boldsymbol{s}$
  - simple performance characterization
  - ideal choice for isotropic transmission in $2 \times 1$ MISO channels

- Extensions: Beamformed Alamouti coding, often for point-to-point MIMO **[Jöngren-Skoglund-Ottersten'02]**, **[Pascual-Iserte-Palomar-*et al.*'06]**, ...

- Recent development: Beamformed Alamouti for multicasting **[Wu-Ma-So'13]** (also **[Wu-So-Ma'12]**), **[Wen-Law-Alabed-Pesavento'12]**.

# BF Alamouti System Model

- **Scenario:** Multicast MISO downlink with instant. CSIT.

- Parse $\boldsymbol{x}(t)$ into blocks via $\boldsymbol{X}(n) = [\ \boldsymbol{x}(2n)\ \boldsymbol{x}(2n+1)\ ]$. In block $n$, we transmit $\boldsymbol{s}(n) = [\ s(2n)\ s(2n+1)\ ]^T$ by

$$\boldsymbol{X}(n) = \boldsymbol{B}\boldsymbol{C}(\boldsymbol{s}(n)),$$

  where $\boldsymbol{B} \in \mathbb{C}^{N \times 2}$ is a transmit beamforming matrix, and

$$\boldsymbol{C}(\boldsymbol{s}) = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix}$$

  is the Alamouti space-time code.

- By utilizing the special structure of the Alamouti code, rx signals can be equivalently turned to SISO models with SNRs

$$\mathsf{SNR}_i = \frac{\boldsymbol{h}_i^H \boldsymbol{B}\boldsymbol{B}^H \boldsymbol{h}_i}{\sigma_i^2}.$$

# Rank-2 Beamforming Problem

- Consider the power minimization design

$$\begin{aligned} v_{\mathsf{QP2}} = \min_{\boldsymbol{B} \in \mathbb{C}^{N \times 2}} \; & \mathrm{Tr}(\boldsymbol{B}\boldsymbol{B}^H) \\ \text{s.t.} \; & \boldsymbol{h}_i^H \boldsymbol{B}\boldsymbol{B}^H \boldsymbol{h}_i / \sigma_i^2 \geq \gamma, \quad i = 1, \ldots, K. \end{aligned} \tag{PM-ALAM}$$

- Observe that

$$\boldsymbol{W} = \boldsymbol{B}\boldsymbol{B}^H \quad \Longleftrightarrow \quad \boldsymbol{W} \succeq \boldsymbol{0} \text{ and } \mathrm{rank}(\boldsymbol{W}) \leq 2.$$

Hence, upon letting $\boldsymbol{A}_i = \boldsymbol{h}_i \boldsymbol{h}_i^H / (\sigma_i^2 \gamma)$, (PM-ALAM) can be reformulated as

$$\begin{aligned} \max_{\boldsymbol{W} \in \mathbb{H}^N} \quad & \mathrm{Tr}(\boldsymbol{W}) \\ \text{s.t.} \quad & \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{W}) \geq 1, \quad i = 1, \ldots, K, \quad \boldsymbol{W} \succeq \boldsymbol{0}, \quad \mathrm{rank}(\boldsymbol{W}) \leq 2. \end{aligned}$$

- Let us do the same trick— dropping the rank constraint. Then, we get

$$\begin{aligned} \max_{\boldsymbol{W} \in \mathbb{H}^N} \quad & \mathrm{Tr}(\boldsymbol{W}) \\ \text{s.t.} \quad & \mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{W}) \geq 1, \quad i = 1, \ldots, K, \quad \boldsymbol{W} \succeq \boldsymbol{0}, \end{aligned}$$

which is the same SDR as that for the previous beamforming problem!

# Gaussian Randomization

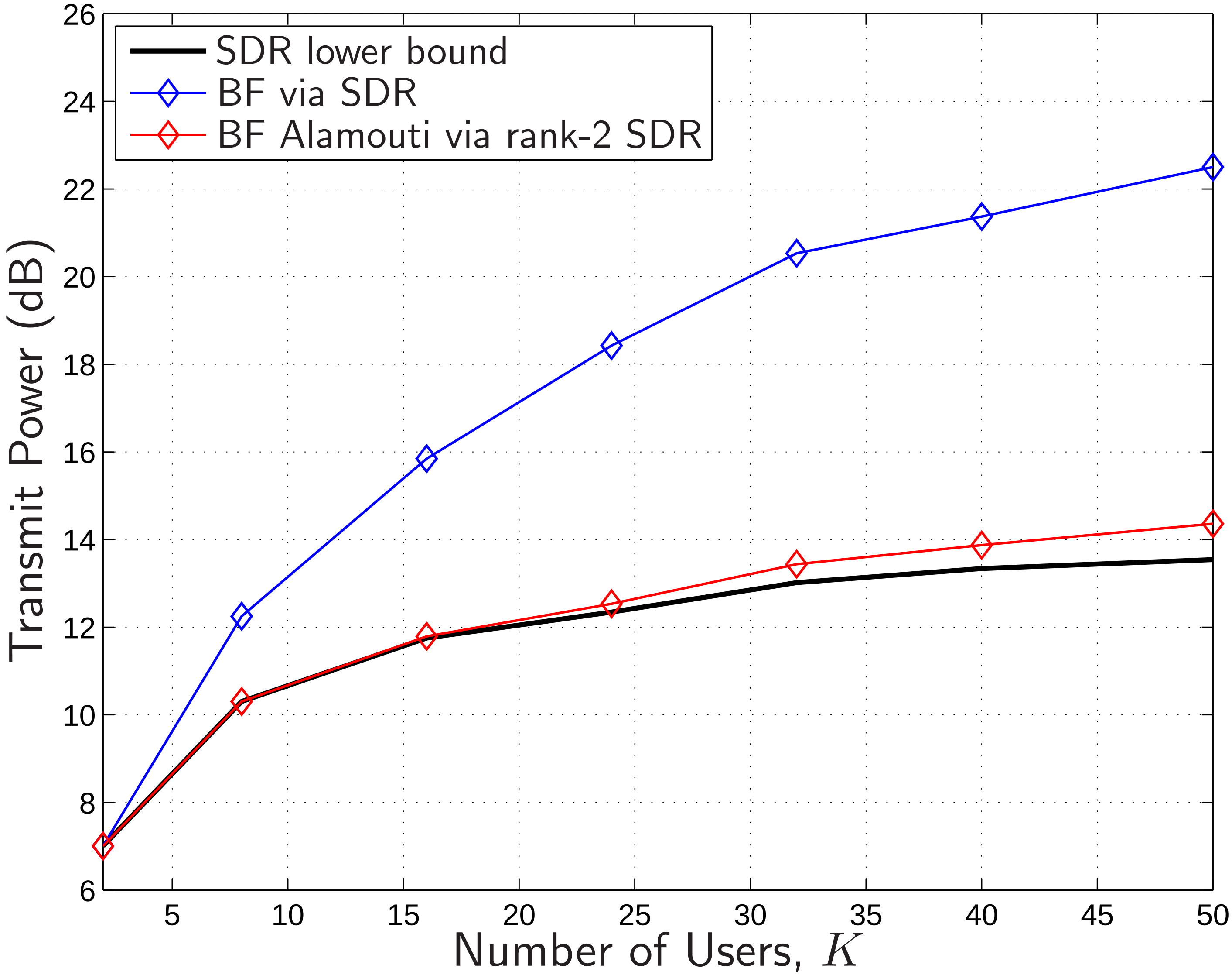
- Let $\boldsymbol{W}^\star$ be a solution to (SDR).

- If $\boldsymbol{W}^\star$ satisfies $\boldsymbol{W}^\star = \hat{\boldsymbol{B}}\hat{\boldsymbol{B}}^H$, or $\mathrm{rank}(\boldsymbol{W}^\star) \leq 2$, then we are done— $\hat{\boldsymbol{B}}$ is optimal to (PM-ALAM).

- Now, consider instances for which $\mathrm{rank}(\boldsymbol{W}^\star) > 2$. Can we do Gaussian randomization for rank-2 $\boldsymbol{W}$?

- The answer is yes.

**Box 4.** ***Gaussian Randomization Procedure for (PM-ALAM)***
**given** an SDR solution $\boldsymbol{W}^\star$, and a number of randomizations $L$.
for $j = 1, \ldots, L$
  generate $\boldsymbol{\xi}_{1,j}, \boldsymbol{\xi}_{2,j} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{W}^\star)$, and define $\tilde{\boldsymbol{B}}_j = [\ \boldsymbol{\xi}_{1,j}\ \boldsymbol{\xi}_{2,j}\ ]$;
  let $\hat{\boldsymbol{B}}_j = \dfrac{\tilde{\boldsymbol{B}}_j}{\sqrt{\min_{i=1,\ldots,K} \mathrm{Tr}(\tilde{\boldsymbol{B}}_j \tilde{\boldsymbol{B}}_j^H \boldsymbol{A}_i)}}$;
end
**output** $\hat{\boldsymbol{B}} = \hat{\boldsymbol{B}}_{j^\star}$, where $j^\star = \arg \min_{j=1,\ldots,L} \mathrm{Tr}(\hat{\boldsymbol{B}}_j \hat{\boldsymbol{B}}_j^H)$.




Performance of the rank-two SDR beamformed Alamouti scheme with respect to the number of users. $\gamma = 10$dB. $N_t = 4$.

## Theoretical Performance of SDR-based BF Alamouti

- By the Shapiro-Barvinok-Pataki (SBP) rank reduction result (see Part II), we have

> When $K \leq 8$, there exists a solution $\boldsymbol{W}^\star$ of (SDR) whose rank satisfies $\mathrm{rank}(\boldsymbol{W}^\star) \leq 2$.

- **Implication:** SDR exactly solves the rank-two beamforming problem for eight users or less (recall that for BF, we have three users or less).

- Let $v(\boldsymbol{B}) = \mathrm{Tr}(\boldsymbol{B}\boldsymbol{B}^H)$. For $K \geq 8$, the following result is established in **[Wu-So-Ma'12]**, **[Wu-Ma-So'13]**:

> With high probability, the solution $\hat{\boldsymbol{B}}$ generated by the rank-2 Gaussian randomization procedure in Box 4 satisfies
>
> $$v_{\mathsf{QP2}} \leq v(\hat{\boldsymbol{B}}) \leq 12.22\sqrt{K} v_{\mathsf{QP2}}.$$

- This provable worst-case performance gain is better than that of BF, which is $8K$.

# Remarks

- The ideas used in the rank-two SDR beamformed Alamouti scheme can also be applied to other scenarios, such as
  - multicast relaying **[Schad-Law-Pesavento'12]**;
  - multi-group multicast **[Ji-Wu-So-Ma'13]**, **[Law-Wen-Pesavento'13]**;
  - energy harvesting **[Chalise-Ma-Zhang-Suraweera-Amin'13]**.

# Extension: Multi-Group Multicast CR Networks

- **Scenario:** Multi-group multicast in a CR network.

- **Problem:** Max-min-fair design formulation under the rank-two beamformed Alamouti scheme:

$$
\begin{aligned}
v_{\mathsf{MMF}} = \max_{\boldsymbol{B}_1,\ldots,\boldsymbol{B}_M \in \mathbb{C}^{N\times 2}} \quad & \min_{\substack{k=1,\ldots,K_m \\ m=1,\ldots,M}} \frac{\mathrm{Tr}(\boldsymbol{R}_{m,k}\boldsymbol{B}_m\boldsymbol{B}_m^H)}{\sum_{l\neq m}\mathrm{Tr}(\boldsymbol{R}_{m,k}\boldsymbol{B}_l\boldsymbol{B}_l^H) + \sigma_{m,k}^2} \\
\text{s.t.} \quad & \sum_{m=1}^{M} \mathrm{Tr}(\boldsymbol{B}_m\boldsymbol{B}_m^H) \le P, \\
& \sum_{m=1}^{M} \mathrm{Tr}(\boldsymbol{G}_l\boldsymbol{B}_m\boldsymbol{B}_m^H) \le \delta_l, \quad l = 1,\ldots,L, \quad \text{(IT constraints)}
\end{aligned}
\qquad \text{(MMF)}
$$

where $P$ is the transmit power limit.

# Theoretical Results for Multi-Group Multicast CR Networks

- Let $\boldsymbol{W}_m = \boldsymbol{B}_m \boldsymbol{B}_m^H$. Using the Huang-Palomar extension of the SBP rank reduction result, we obtain the following:

> When $\sum_{m=1}^M K_m \leq M + 7$, there exists an SDR solution $\{\boldsymbol{W}_m^\star\}$ with $\mathrm{rank}(\boldsymbol{W}_m^\star) \leq 2$ for $m = 1, \ldots, M$.

- Suppose now that $\sum_{m=1}^M K_m > M + 7$. Let

$$v(\{\boldsymbol{B}_m\}) = \min_{\substack{k=1,\ldots,K_m \\ m=1,\ldots,M}} \frac{\mathrm{Tr}(\boldsymbol{R}_{m,k}\boldsymbol{B}_m\boldsymbol{B}_m^H)}{\sum_{l \neq m} \mathrm{Tr}(\boldsymbol{R}_{m,k}\boldsymbol{B}_l\boldsymbol{B}_l^H) + \sigma_{m,k}^2}.$$

In **[Ji-Wu-So-Ma'13]**, the following result is established:

> A Gaussian randomization procedure can generate a feasible solution $\{\hat{\boldsymbol{B}}_m\}$ to (MMF) with $\mathrm{rank}(\hat{\boldsymbol{B}}_m) \leq 2$ for all $m$. Moreover, with high probability,
>
> $$v_{\mathsf{MMF}} \geq v(\{\hat{\boldsymbol{B}}_m\}) \geq \frac{v_{\mathsf{MMF}}}{8\sqrt{\sum_{m=1}^M K_m}(3\log 8(L+1))}.$$

# Theoretical Results for Multi-Group Multicast CR Networks

- The above results generalize those in **[Chang-Luo-Chi'08]**, which concern rank-1 beamforming in the multi-group multicast scenario with no primary user.

## Further Discussion: Can We Do Rank-$r$ Beamforming, $r \geq 3$?

- From an optimization viewpoint, yes.
  - For example, we can consider rank-$3$ SDR, wherein an effective power loss of $\mathcal{O}(K^{1/3})$ can be proven.

- From a realizable physical layer viewpoint, not straightforward.
  - A generalization of the Alamouti code is the class of orthogonal space-time block codes (OSTBCs).
  - Full-rate OSTBCs do not exist for $r > 2$ **[Liang-Xia'03]**.
  - For example, for $r = 3$, the maximal rate is $3/4$, and the code is

$$\boldsymbol{C}(\boldsymbol{s}) = \begin{bmatrix} s_1 & -s_2^* & -s_3^* & 0 \\ s_2 & s_1^* & 0 & -s_3^* \\ s_3 & 0 & s_1^* & s_2^* \end{bmatrix}.$$

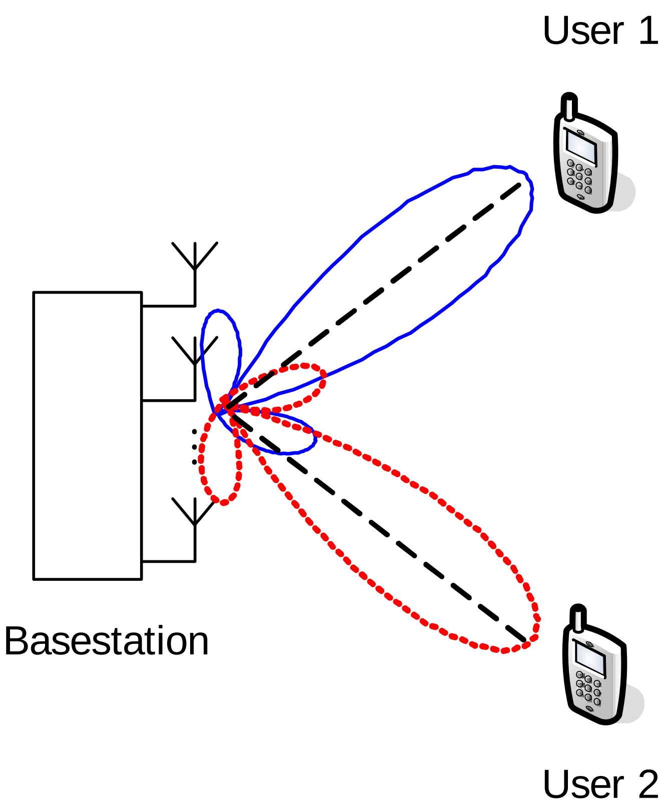
- The question of rank-$r$ beamforming has led to new studies that are no longer about SDR— e.g., *stochastic beamforming* **[Wu-Ma-So'13]**, which can perform virtual rank-$r$ beamforming for any $r \geq 1$.

# Part III.B: Advanced Topics in Transmit Beamforming

## Topic 2. Worst-Case Robust Beamforming

# Motivation

**Scenario:** Unicast multiuser MISO downlink, with instant. CSIT $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K$.



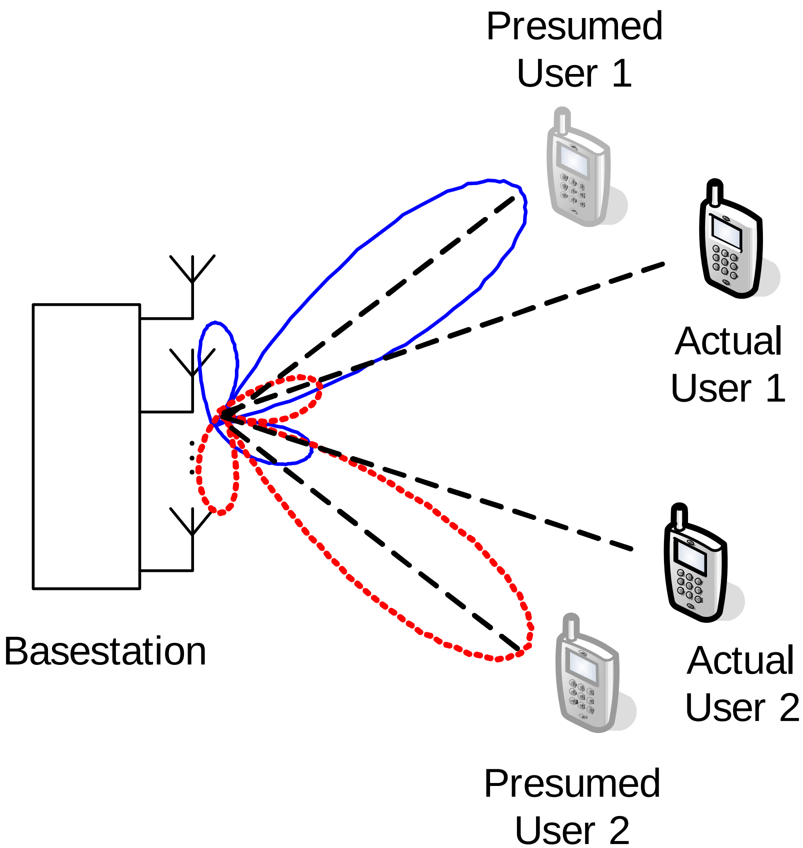

- Previously, we have formulated the following QoS-constrained power minimization problem:

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

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

# Motivation

**Issue:** CSIT is generally imperfectly known in practice.

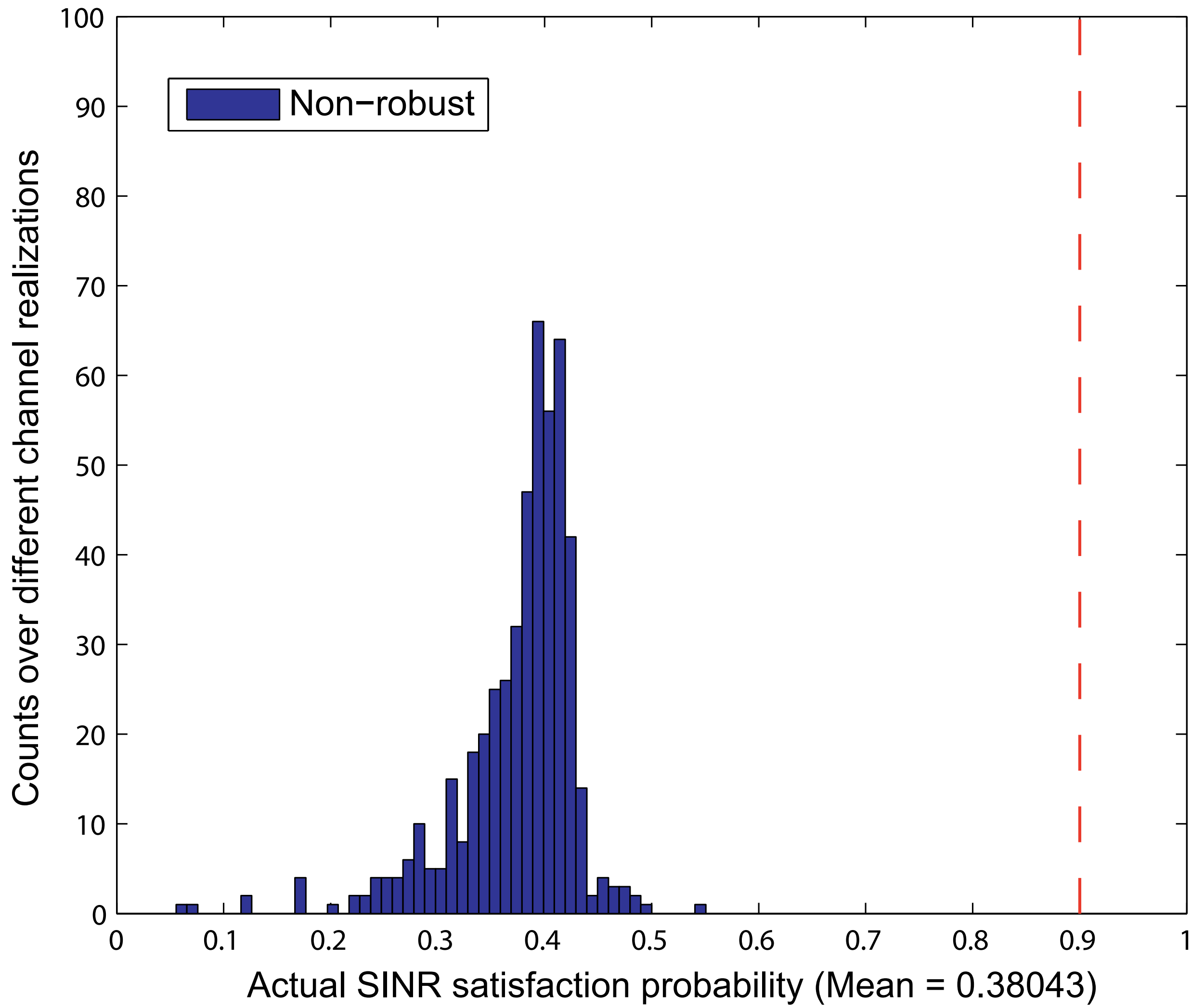



- Suppose that the presumed CSIT, $\{\boldsymbol{h}_i\}$, is inaccurate.

- If we directly substitute the presumed CSIT into the standard power minimization design

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

and run it, then the resultant design may have severe SINR outage.




Histogram of the actual SINR satisfaction probabilities of the non-robust QoS-constrained power minimization 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.

# Worst-Case Robust Beamforming Problem

- **Design goal:** Guarantee the SINR requirements even in the worst-case scenario.

- Let $\boldsymbol{W}_i = \boldsymbol{w}_i \boldsymbol{w}_i^H$, $i = 1, \ldots, K$, and define

$$\mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \boldsymbol{h}_i) = \frac{\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{h}_i \boldsymbol{h}_i^H)}{\sum_{l \neq i} \mathrm{Tr}(\boldsymbol{W}_l \boldsymbol{h}_i \boldsymbol{h}_i^H) + \sigma_i^2}.$$

- We adopt the CSIT model

$$\boldsymbol{h}_i = \bar{\boldsymbol{h}}_i + \boldsymbol{e}_i,$$

where $\bar{\boldsymbol{h}}_i$ is the presumed channel; $\boldsymbol{e}_i$ is a deterministic unknown with $\|\boldsymbol{e}_i\| \leq r_i$ ($r_i$ is known).

- Consider the following worst-case robust BF design problem:

$$\begin{aligned} \min_{\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K} \quad & \textstyle\sum_{i=1}^K \mathrm{Tr}(\boldsymbol{W}_i) \\ \text{s.t.} \quad & \mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \geq \gamma_i, \quad \forall \|\boldsymbol{e}_i\| \leq r_i, \quad i = 1, \ldots, K, \\ & \mathrm{rank}(\boldsymbol{W}_i) \leq 1, \quad \boldsymbol{W}_i \succeq \boldsymbol{0}, \qquad i = 1, \ldots, K. \end{aligned} \tag{RPM}$$

- By noting that

$$\mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \boldsymbol{h}_i) \geq \gamma_i \Longleftrightarrow \boldsymbol{h}_i^H \left( \frac{1}{\gamma_i} \boldsymbol{W}_i - \sum_{l \neq i} \boldsymbol{W}_l \right) \boldsymbol{h}_i \geq \sigma_i^2,$$

(RPM) can be rewritten as

$$\begin{aligned} \min_{\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K} \quad & \textstyle\sum_{i=1}^K \mathrm{Tr}(\boldsymbol{W}_i) \\ \text{s.t.} \quad & (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i)^H \left( \tfrac{1}{\gamma_i} \boldsymbol{W}_i - \textstyle\sum_{l \neq i} \boldsymbol{W}_l \right) (\bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \geq \sigma_i^2, \forall \|\boldsymbol{e}_i\| \leq r_i, \forall i, \\ & \mathrm{rank}(\boldsymbol{W}_i) \leq 1, \quad \boldsymbol{W}_i \succeq \boldsymbol{0}, \qquad i = 1, \ldots, K. \end{aligned} \tag{RPM}$$

  - Again, the idea is to drop the rank constraints in (RPM).
  - (RPM) without rank constraints, or SDR, is convex. However, it does not mean that the problem is easy— there are infinitely many constraints w.r.t. the $\boldsymbol{e}_i$'s.
  - Question: Is the SDR of (RPM) efficiently solvable?
  - The answer is yes, using the so-called $\mathcal{S}$-lemma **[Zheng-Wang-Ng'08]**.

# $\mathcal{S}$-Lemma

- **$\mathcal{S}$-Lemma:** Let

$$f_0(\boldsymbol{x}) = \boldsymbol{x}^H \boldsymbol{A}_0 \boldsymbol{x} + 2\mathrm{Re}\{\boldsymbol{b}_0^H \boldsymbol{x}\} + c_0,$$
$$f_1(\boldsymbol{x}) = \boldsymbol{x}^H \boldsymbol{A}_1 \boldsymbol{x} + 2\mathrm{Re}\{\boldsymbol{b}_1^H \boldsymbol{x}\} + c_1,$$

and suppose that there exists $\hat{\boldsymbol{x}}$ such that $f_1(\hat{\boldsymbol{x}}) < 0$. Then,

$$\begin{array}{c} f_0(\boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \\ \text{satisfying } f_1(\boldsymbol{x}) \leq 0 \end{array} \iff \begin{array}{c} \text{there exists } \lambda \geq 0 \text{ such that} \\ \begin{bmatrix} \boldsymbol{A}_0 & \boldsymbol{b}_0 \\ \boldsymbol{b}_0^H & c_0 \end{bmatrix} + \lambda \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{b}_1 \\ \boldsymbol{b}_1^H & c_1 \end{bmatrix} \succeq \boldsymbol{0}. \end{array} \qquad (\dagger)$$

- In other words, the infinitely many constraints on the LHS of ($\dagger$) is equivalent to the linear matrix inequality on the RHS.

- The $\mathcal{S}$-lemma is widely used in optimization, signal processing, communications and many other areas.

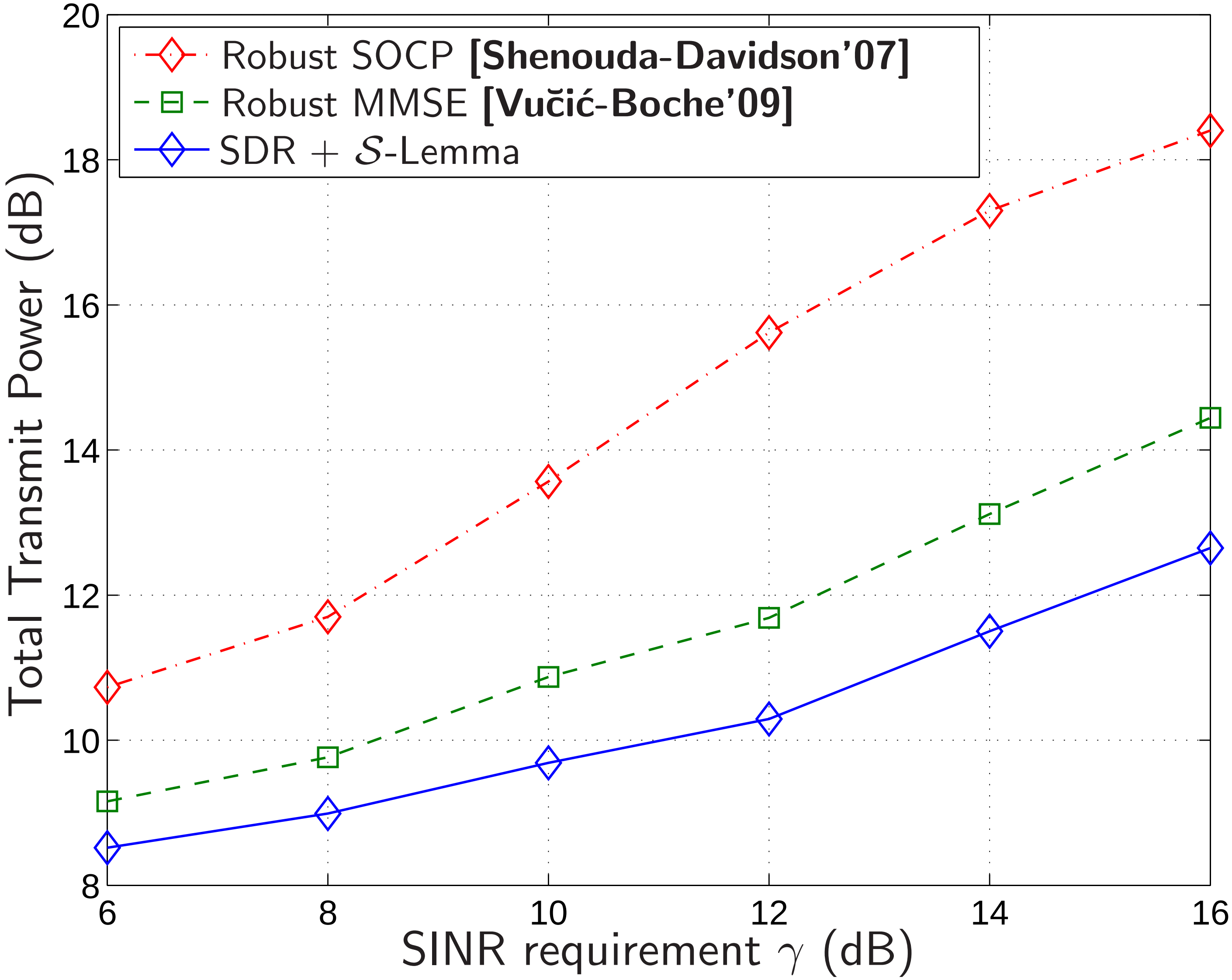
# Worst-Case Robust BF via SDR and $\mathcal{S}$-Lemma

- By applying the $\mathcal{S}$-lemma to (RPM) and dropping the rank constraints, we obtain the following SDR:

$$
\begin{aligned}
\min_{\substack{\lambda_1,\ldots,\lambda_K,\\ \boldsymbol{W}_1,\ldots,\boldsymbol{W}_K}} \; & \sum_{i=1}^{K} \operatorname{Tr}(\boldsymbol{W}_i) \\
\text{s.t.} \; & \begin{bmatrix} \boldsymbol{I} \\ \bar{\boldsymbol{h}}_i^H \end{bmatrix} \left( \frac{1}{\gamma_i} \boldsymbol{W}_i - \sum_{l \neq i} \boldsymbol{W}_l \right) \begin{bmatrix} \boldsymbol{I} & \bar{\boldsymbol{h}}_i \end{bmatrix} + \begin{bmatrix} \lambda_i \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\sigma_i^2 - \lambda_i r_i^2 \end{bmatrix} \succeq \boldsymbol{0}, \; \forall i, \\
& \lambda_1, \ldots, \lambda_K \geq 0, \; \boldsymbol{W}_1, \ldots, \boldsymbol{W}_K \succeq \boldsymbol{0}.
\end{aligned}
\tag{RPM-SDR}
$$

- (RPM-SDR) is an SDP, with a finite number of constraints.

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




Performance of worst-case robust SDR and other solutions. $N_t = K = 4$, $\gamma = \gamma_1 = \cdots = \gamma_K$, $\sigma_1^2 = \cdots = \sigma_K^2 = 0.1$, $r_1 = \cdots = r_K = 0.1$. SDR yielded rank-1 solution in all the trials run.

# Rank-One Optimality for Worst-Case Robust BF?

- **Challenge:** Can we prove (or disprove) the rank-one optimality of SDR for the worst-case robust unicast BF problem

$$
\begin{aligned}
\min_{\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K \in \mathbb{H}^N} \quad & \textstyle\sum_{i=1}^{K} \mathrm{Tr}(\boldsymbol{W}_i) \\
\text{s.t.} \quad & \mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \geq \gamma_i, \quad \forall \|\boldsymbol{e}_i\| \leq r_i, \quad i = 1, \ldots, K, \\
& \boldsymbol{W}_i \succeq \boldsymbol{0}, \qquad i = 1, \ldots, K.
\end{aligned}
$$

- Note: The SBP rank reduction result does not work here!

- Some sufficient conditions are currently available:
  - **[Song-Shi-Sanjabi-Sun-Luo'12]**: An SDR solution $\{\boldsymbol{W}_i^\star\}$ must be of rank one if $r_i$'s are small enough in a problem instance-dependent manner.
  - **[Chang-Ma-Chi'11]**: An SDR solution $\{\boldsymbol{W}_i^\star\}$ must be of rank one if another related problem (what?) has a unique solution.

- **Proving or disproving SDR rank-one optimality in unicast robust BF remains an intriguing open problem.**

# The Result in [Chang-Ma-Chi'11]

- Recall the SDR problem

$$
\begin{aligned}
\min_{\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K \succeq \boldsymbol{0}} \quad & \textstyle\sum_{i=1}^K \mathrm{Tr}(\boldsymbol{W}_i) \\
\text{s.t.} \quad & \min_{\|\boldsymbol{e}_i\| \le r_i} \mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \ge \gamma_i, \quad i = 1,\ldots,K.
\end{aligned} \tag{$\dagger$}
$$

- Consider a different problem

$$
\begin{aligned}
\max_{\substack{\|\boldsymbol{e}_i\| \le r_i, \\ i=1,\ldots,K}} \min_{\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K \succeq \boldsymbol{0}} \quad & \textstyle\sum_{i=1}^K \mathrm{Tr}(\boldsymbol{W}_i) \\
\text{s.t.} \quad & \mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \bar{\boldsymbol{h}}_i + \boldsymbol{e}_i) \ge \gamma_i, \quad i = 1,\ldots,K.
\end{aligned} \tag{$\ddagger$}
$$

- **[Chang-Ma-Chi'11]** reveals the duality between (†) & (‡). Specifically,
  - the SDR of (‡) (w.r.t. $\boldsymbol{e}_i$'s) yields the same optimal value as (†);
  - any solution of (†) is a solution of the SDR of (‡);
  - if the SDR of (‡) has a unique inner solution $\{\hat{\boldsymbol{W}}_i\}$, then a solution of (†) must be of rank one.

# Remarks

- The combination of SDR and $\mathcal{S}$-lemma for robust BF design may also be applied to other scenarios.

- A Gaussian randomization procedure can be developed for the robust BF design case.

# Part III.B: Advanced Topics in Transmit Beamforming

## Topic 3. Outage-based Robust Beamforming

# Motivation

- Recall that to study the effect of imperfect CSIT, we have adopted the model

$$\boldsymbol{h}_i = \bar{\boldsymbol{h}}_i + \boldsymbol{e}_i,$$

where $\bar{\boldsymbol{h}}_i$ is the presumed channel and $\boldsymbol{e}_i$ is the channel error.

- Previously, we assumed that the error $\boldsymbol{e}_i$ lies in a ball centered at the origin with known radius; i.e., $\|\boldsymbol{e}_i\| \leq r_i$, where $r_i$ is known.
  - This gives rise to a worst-case robust BF design problem.

- Alternatively, one can consider the following Gaussian channel error model:

$$\boldsymbol{e}_i \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{C}_i),$$

where $\boldsymbol{C}_i \succeq \mathbf{0}$ is a known covariance matrix.

  - In particular, we have $\boldsymbol{h}_i \sim \mathcal{CN}(\bar{\boldsymbol{h}}_i, \boldsymbol{C}_i)$.

# Motivation

- Under this probabilistic error model, a meaningful, but very difficult, problem is the following outage-based BF design problem:

$$
\begin{aligned}
\min_{\boldsymbol{W}_1,\ldots,\boldsymbol{W}_K} \quad & \textstyle\sum_{i=1}^K \mathrm{Tr}(\boldsymbol{W}_i) \\
\text{s.t.} \quad & \mathrm{Prob}_{\boldsymbol{h}_i \sim \mathcal{CN}(\bar{\boldsymbol{h}}_i, \boldsymbol{C}_i)} \{ \mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \boldsymbol{h}_i) \geq \gamma_i \} \geq 1 - \rho_i, \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1, \ldots, K, \\
& \mathrm{rank}(\boldsymbol{W}_i) \leq 1, \quad \boldsymbol{W}_i \succeq \boldsymbol{0}, \qquad i = 1, \ldots, K.
\end{aligned}
\tag{OPM}
$$

Here, $\rho_i \in (0,1)$ is user $i$'s maximum tolerable outage probability. Recall that

$$
\mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \boldsymbol{h}_i) = \frac{\mathrm{Tr}(\boldsymbol{W}_i \boldsymbol{h}_i \boldsymbol{h}_i^H)}{\sum_{l \neq i} \mathrm{Tr}(\boldsymbol{W}_l \boldsymbol{h}_i \boldsymbol{h}_i^H) + \sigma_i^2}.
$$

- The above outage-based design problem is an instance of the so-called chance-constrained or probabilistically-constrained optimization problem.

# Tackling the Probabilistic Constraints

- To get a more tractable problem, a natural move is to apply SDR and remove the rank constraints. However, the SDRed (OPM) remains hard.

- Indeed, although the outage-based SINR constraints

$$\mathrm{Prob}_{\boldsymbol{h}_i \sim \mathcal{CN}(\bar{\boldsymbol{h}}_i, \boldsymbol{C}_i)} \left\{ \mathsf{SINR}_i(\{\boldsymbol{W}_j\}, \boldsymbol{h}_i) \geq \gamma_i \right\} \geq 1 - \rho_i$$

can be rewritten as

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

the probability on the LHS has no simple closed form expression.

# Tackling 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 $\mathcal{CN}(\mathbf{0}, \boldsymbol{C}_i)$. Here, $L \geq 1$ is the number of independent samples of $\boldsymbol{e}_i$. 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 $\rho_i$), any solution to ($\dagger$) will satisfy the corresponding SDRed probabilistic constraint with high confidence (but not necessarily always— this could be problematic in some applications).

- In addition, this method is extremely time consuming in practice.

# Tackling the Probabilistic Constraint: Convex Restriction

- To circumvent the aforementioned difficulties, let us consider an alternative approach. Let

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

  be the violation probability. Recall that we want

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

- By applying some simple transformations, we can express $V_i$ as

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

  for some $\boldsymbol{Q}$, $\boldsymbol{r}$ and $s$ that depend on $\boldsymbol{C}_i$, $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_K$, and the index $i$. (Here and in the sequel, we drop the index $i$ for notational simplicity.)

- Thus, the crux of the outage-based design problem is how to process the probabilistic constraint

$$\mathrm{Prob}_{\boldsymbol{e}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{I})}\left\{\boldsymbol{e}^H\boldsymbol{Q}\boldsymbol{e}+2\mathrm{Re}\{\boldsymbol{e}^H\boldsymbol{r}\}+s<0\right\}\leq\rho. \tag{PC}$$

- Here is an idea: Suppose that we can find an efficiently computable convex function $f(\boldsymbol{Q},\boldsymbol{r},s)$ such that

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

  Then, by construction, the convex constraint

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

  serves as a sufficient condition for (PC) to hold.

  – We call (CR-PC) a convex restriction of (PC).

# Finding a Convex Restriction

- Can we find such an $f$? Does it even exist?
  - As it turns out, the answer is Yes! (And there are many such functions.)

- The constructions are based on large deviation bounds on the tail probability

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

# Finding a Convex Restriction: Sphere Bounding

- The first construction is motivated by ideas from robust optimization. Consider a set $\mathcal{B}$ such that
$$\text{Prob}_{\boldsymbol{e}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{I})}\{\boldsymbol{e}\in\mathcal{B}\}\geq 1-\rho.$$

- Then, we have the following implication:
$$\begin{aligned} &\boldsymbol{\delta}^H\boldsymbol{Q}\boldsymbol{\delta}+2\text{Re}\{\boldsymbol{\delta}^H\boldsymbol{r}\}+s\geq 0 \quad \text{for all } \boldsymbol{\delta}\in\mathcal{B} \\ \Longrightarrow\quad &\text{Prob}_{\boldsymbol{e}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{I})}\{\boldsymbol{e}^H\boldsymbol{Q}\boldsymbol{e}+2\text{Re}\{\boldsymbol{e}^H\boldsymbol{r}\}+s<0\}\leq\rho. \end{aligned}$$

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

  where
$$f(\boldsymbol{Q},\boldsymbol{r},s)=\begin{cases}\rho & \text{if } \boldsymbol{\delta}^H\boldsymbol{Q}\boldsymbol{\delta}+2\text{Re}\{\boldsymbol{\delta}^H\boldsymbol{r}\}+s\geq 0 \quad \forall\boldsymbol{\delta}\in\mathcal{B}, \\ +\infty & \text{otherwise.}\end{cases}$$

# Finding a Convex Restriction: Sphere Bounding

- Note that $f(\boldsymbol{Q}, \boldsymbol{r}, s) \leq \rho$ if and only if

$$\boldsymbol{\delta}^H \boldsymbol{Q} \boldsymbol{\delta} + 2\mathrm{Re}\{\boldsymbol{\delta}^H \boldsymbol{r}\} + s \geq 0 \quad \forall \boldsymbol{\delta} \in \mathcal{B}. \tag{SB}$$

  Thus, (SB) is a sufficient condition for (PC) to hold.

- Is (SB) an efficiently computable constraint? That depends on what $\mathcal{B}$ is.

- Suppose that we choose

$$\mathcal{B} = \{\boldsymbol{\delta} : \|\boldsymbol{\delta}\| \leq d\}, \quad d = \sqrt{\frac{\Phi^{-1}_{\chi^2_{2n}}(1-\rho)}{2}},$$

  where $\Phi^{-1}_{\chi^2_m}(\cdot)$ is the inverse cumulative distribution function of the (central) Chi-square random variable with $m$ degrees of freedom **[Wang-Chang-Ma-Chi'10]**.

- Then, we have $\mathrm{Prob}_{\boldsymbol{e} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I})}\{\boldsymbol{e} \in \mathcal{B}\} = 1 - \rho$.

## Finding a Convex Restriction: Sphere Bounding

- Moreover, by the $\mathcal{S}$-lemma, the constraint

$$\boldsymbol{\delta}^H \boldsymbol{Q}\boldsymbol{\delta} + 2\mathrm{Re}\{\boldsymbol{\delta}^H \boldsymbol{r}\} + s \geq 0 \quad \forall \boldsymbol{\delta} \in \mathcal{B}$$

  is equivalent to

$$\begin{bmatrix} \boldsymbol{Q} + t\boldsymbol{I} & \boldsymbol{r} \\ \boldsymbol{r}^H & s - td^2 \end{bmatrix} \succeq \mathbf{0}, \quad t \geq 0, \tag{SB-CR}$$

  which is an SDP in the variables $(\boldsymbol{Q}, \boldsymbol{r}, s, t)$.

- Thus, (SB-CR) is a convex restriction of (PC).

# Outage-Based BF Design via Sphere Bounding

- Applying the sphere bounding technique to the outage-based SINR constraint

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

  we obtain the following convex restriction:

$$\begin{bmatrix} \boldsymbol{I} \\ \bar{\boldsymbol{h}}_i^H \end{bmatrix} \left( \tfrac{1}{\gamma_i} \boldsymbol{W}_i - \textstyle\sum_{l \neq i} \boldsymbol{W}_l \right) \begin{bmatrix} \boldsymbol{I} & \bar{\boldsymbol{h}}_i \end{bmatrix} + \begin{bmatrix} t_i \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\sigma_i^2 - \lambda_i d_i^2 \end{bmatrix} \succeq \boldsymbol{0}, \quad t_i \geq 0, \quad (\dagger)$$

  where $d_i^2 = \Phi_{\chi_{2n}^2}^{-1}(1 - \rho_i)/2$.

- It is worth noting that the constraint $(\dagger)$ has exactly the same form as that in the worst-case robust BF design problem (RPM-SDR). The only difference lies in how the parameter $d_i$ is determined.
  - For worst-case robust design, $d_i$ is a pre-specified parameter that determines the radius of the ball in which the error vector $\boldsymbol{e}_i$ lies; i.e., $\|\boldsymbol{e}_i\| \leq d_i$.
  - For outage-based design, $d_i$ is determined by the maximum tolerable outage probability $\rho_i$.

## Finding a Convex Restriction: Bernstein-Type Inequality

- Another construction is based on the so-called Bernstein-type inequality **[Bechar2009]**, **[Wang-Chang-Ma-So-Chi'11]**, which states that

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

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

- Is the constraint

$$f(\boldsymbol{Q},\boldsymbol{r},s)=e^{-T^{-1}(s)}\leq\rho$$

  efficiently computable? Yes! It is equivalent to the SDP

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

  in the variables $(\boldsymbol{Q},\boldsymbol{r},s,t_1,t_2)$.

- Thus, (BI-CR) is a convex restriction of (PC).

# Outage-Based BF Design via Bernstein-Type Inequality

- Applying the Bernstein-type inequality to the outage-based SINR constraints, we obtain another convex restriction of the outage-based BF design problem **[Wang-Chang-Ma-So-Chi'11]**.

- **Yet another mysterious finding in simulations:** For the Bernstein-type inequality approach, rank-one SDR solution is obtained in almost all the problem instances— the same phenomenon as that observed in worst-case robust design (or the sphere bounding approach).

# Comparing the Convex Restrictions

- The above development raises the following natural question:

  What is the approximation quality (w.r.t. the original probabilistic constraint) of the convex restrictions obtained by the sphere bounding and Bernstein-type inequality approaches?

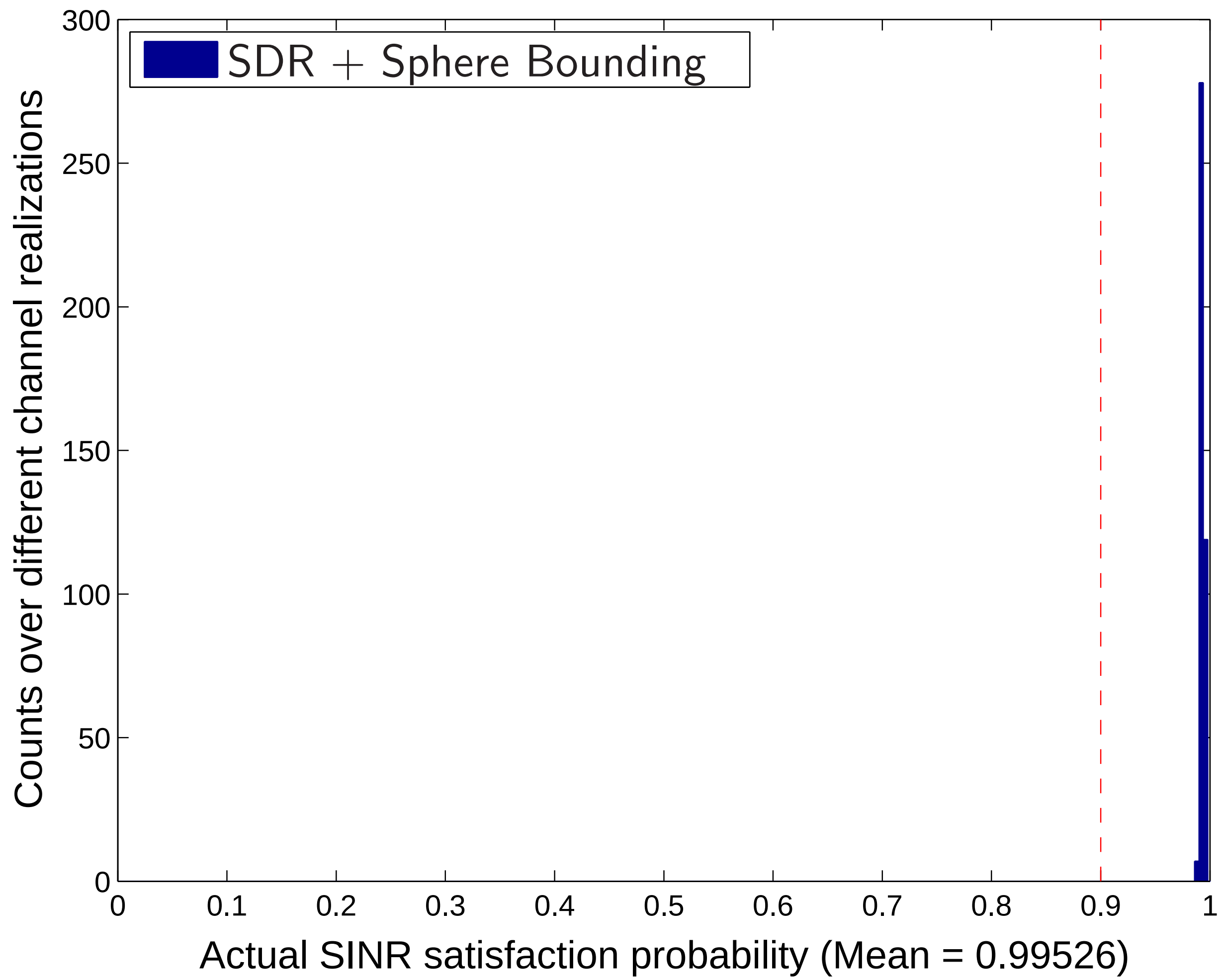
- Unfortunately, this remains an intriguing open question.

- This leads to the next natural question:

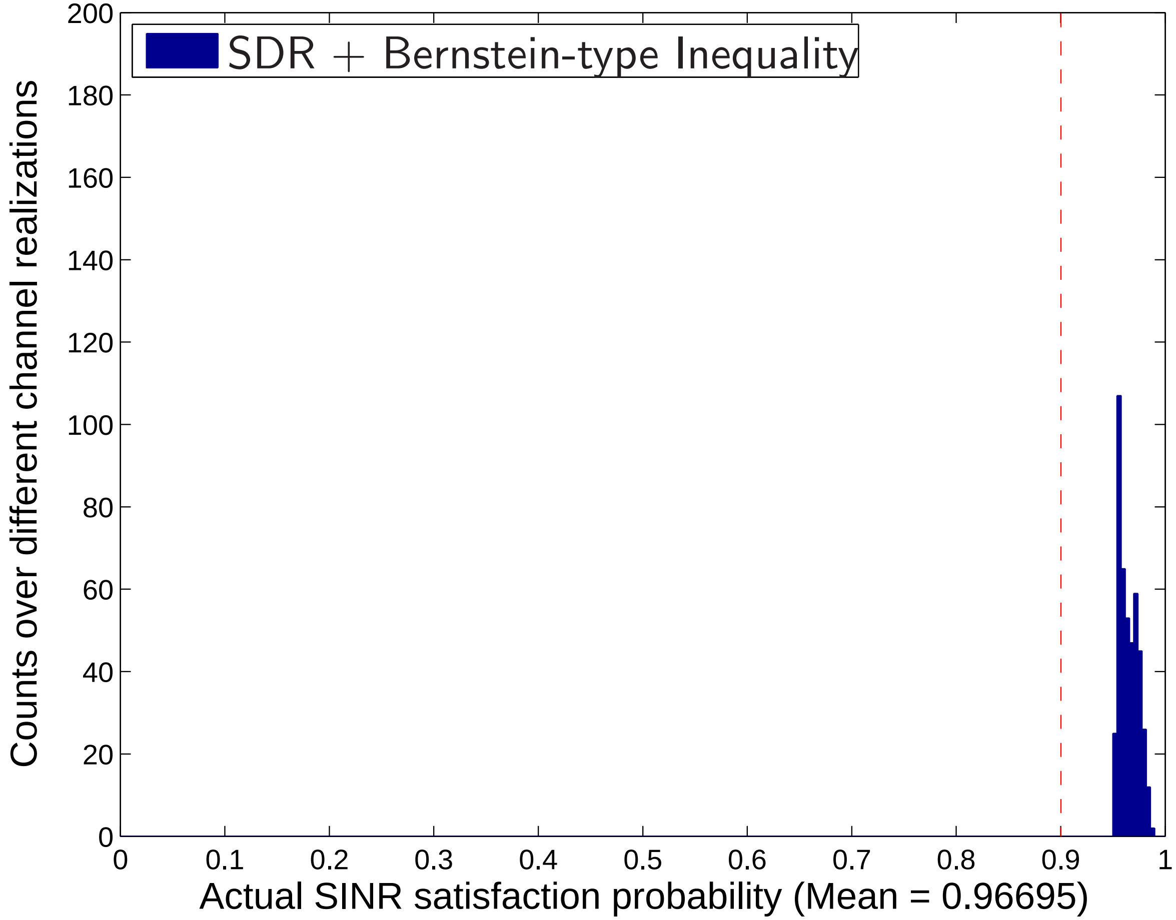
  Which of the two convex restrictions has better approximation quality?

- In **[Wang-So-Chang-Ma-Chi'14]**, the following result is shown:

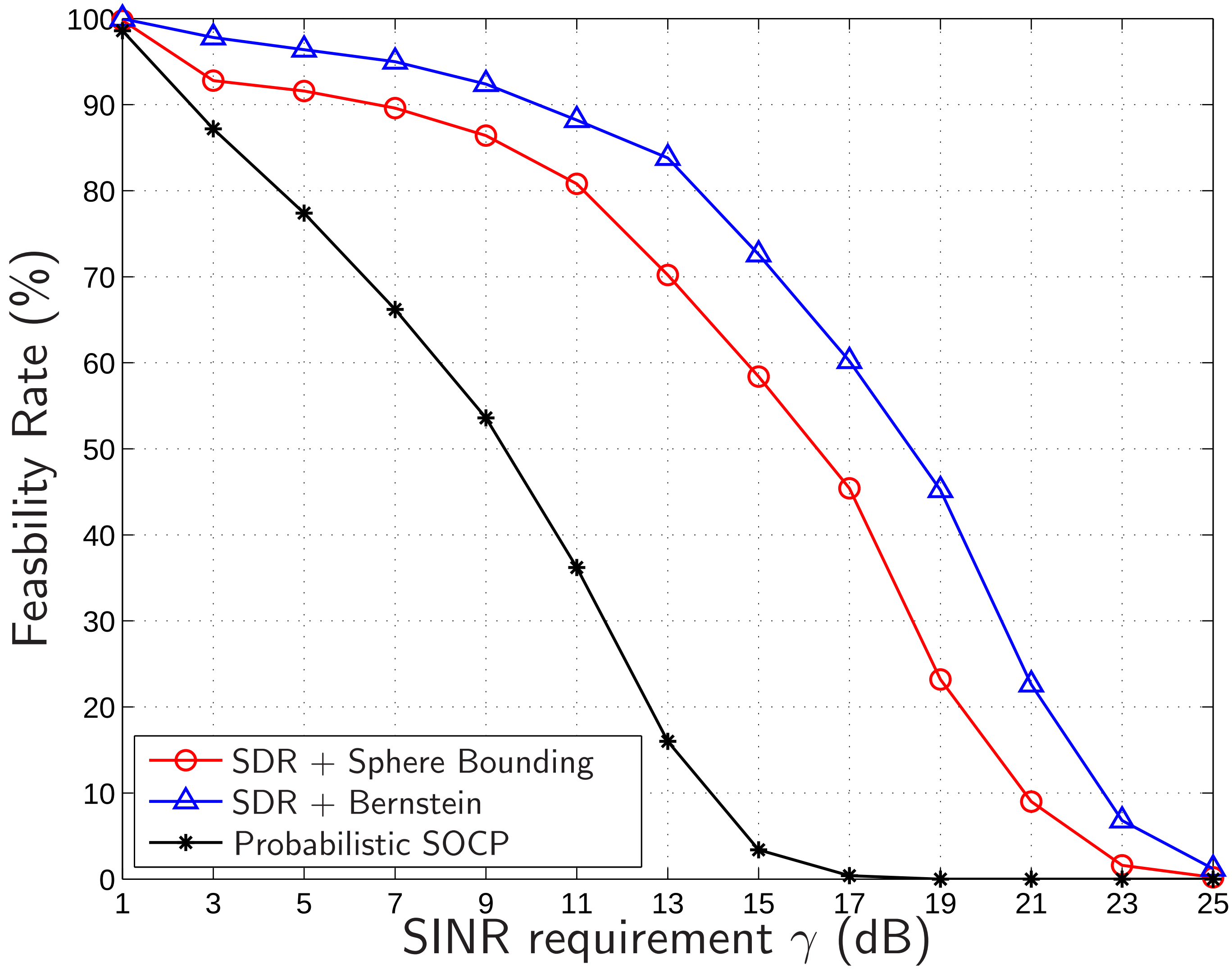
  > Under some mild assumptions, the convex restriction based on the Bernstein-type inequality approach yields a better approximation of the original probabilistic constraint than that based on the sphere bounding approach.



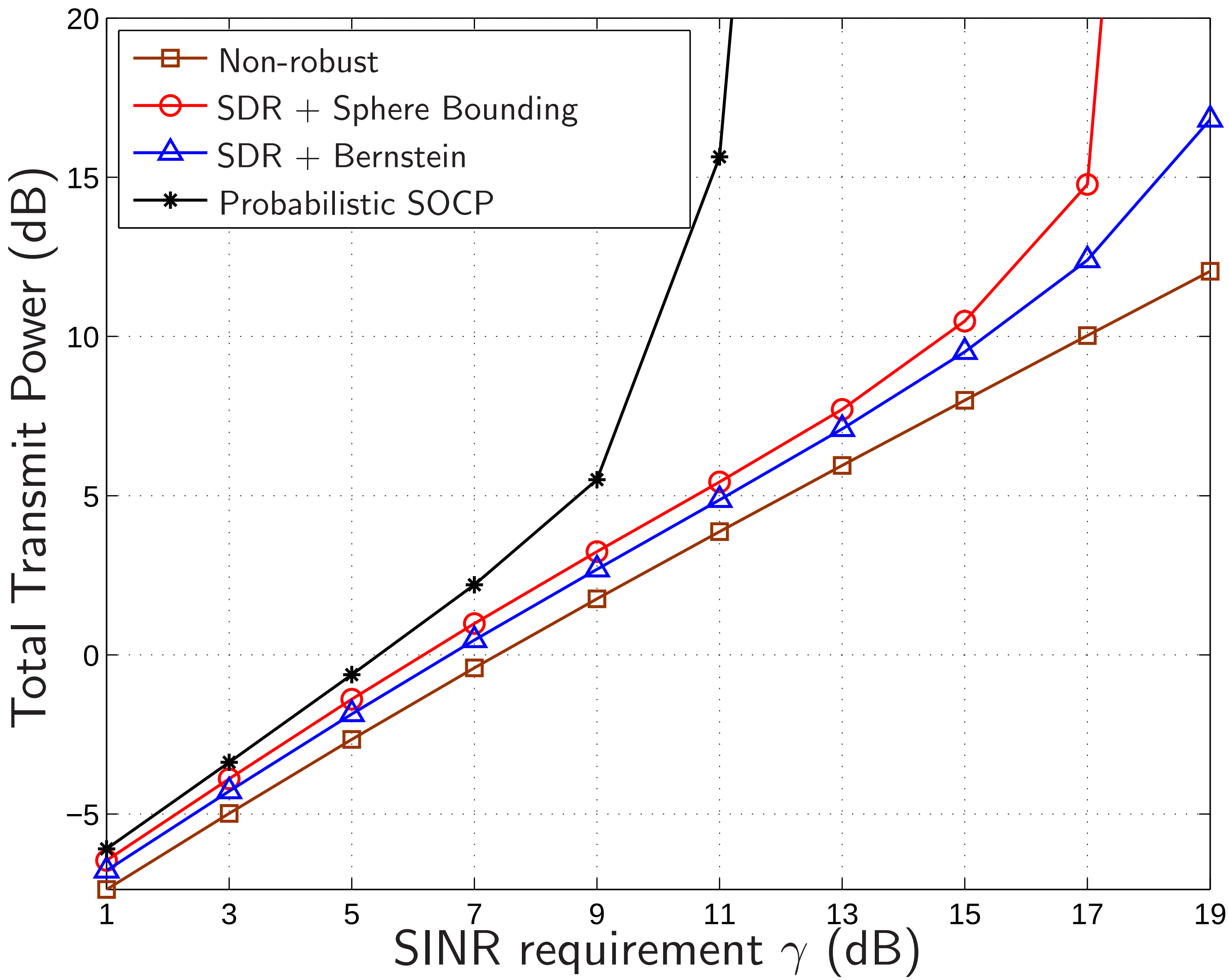

Histogram of the actual SINR satisfaction probabilities of the SDR+sphere bounding method. $N_t = K = 3$; i.i.d. complex Gaussian CSI errors with zero mean and variance $0.002$; $\gamma = 11$dB; $\rho = 0.1$ ($90\%$ SINR satisfaction).




Histogram of the actual SINR satisfaction probabilities of the 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 SDR methods 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).



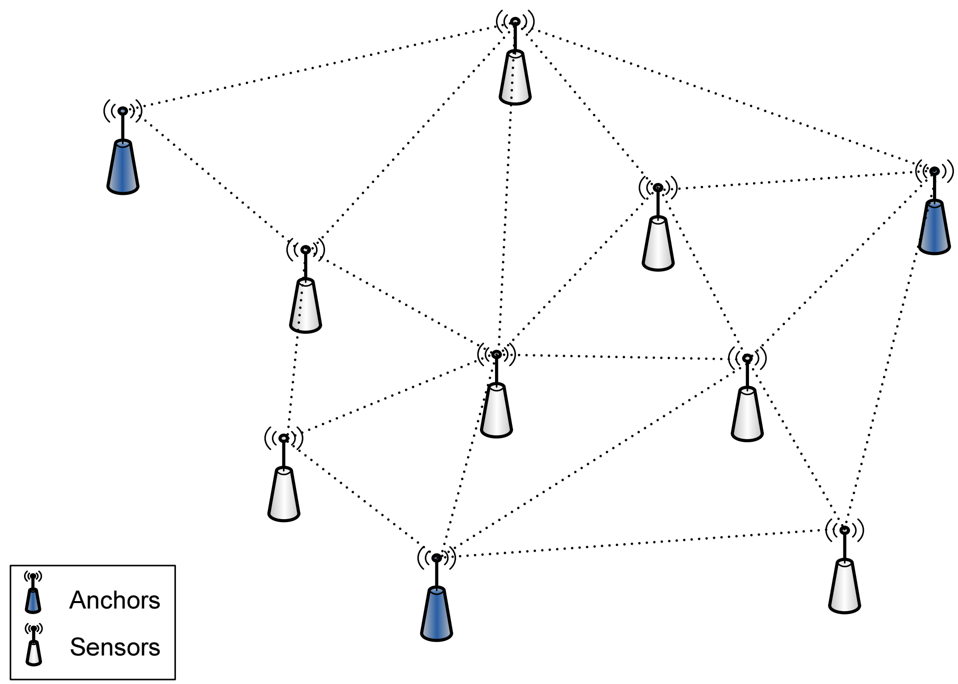

Transmit power performance of the SDR methods and the probabilistic SOCP method. $N_t = K = 3$; $\sigma_e^2 = 0.002$; $\rho = 0.1$ (90% SINR satisfaction). "Non-robust" refers to the perfect CSIT-based design, which is not robust against CSIT errors.

# Part III.C: Sensor Network Localization

# Overview

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

- In ad-hoc sensor networks, sensors' locations are important but may not be known.
- Though one can equip every sensor with GPS, it is too expensive to do so.
- Thus, we may only have several sensors, called *anchors*, that have self-localization capability.




- A pair of sensors that are within communication range of each other can measure the distance between themselves (e.g., via TOA, RSS, etc.).
- The inter-sensor distance measurements, together with anchor locations, can be used to jointly estimate the sensors' locations.

# Problem Formulation

- 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 $j$ (resp. sensor $i$ and anchor $j$) is $\|\boldsymbol{x}_i - \boldsymbol{x}_j\|$ (resp. $\|\boldsymbol{x}_i - \boldsymbol{a}_j\|$).

- Let $E_{ss}$ and $E_{sa}$ denote the set of sensor-sensor and sensor-anchor edges, resp.

- **Problem:** Assuming the distance measurements $\{d_{ij}\}_{(i,j)\in E_{ss}}$ and $\{\bar{d}_{ij}\}_{(i,j)\in E_{sa}}$ are noiseless (extensions for noisy cases will be discussed later), we need to find $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$ such that

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

# 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

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

  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

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

- 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 $\boldsymbol{Y} \succeq \mathbf{0}$, we relax it to
$$\boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X}.$$
  - This is an SDP constraint, since by the Schur complement,
$$\boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X} \quad \Longleftrightarrow \quad \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \succeq \mathbf{0}.$$
- Then, we have the following SDR of the SNL problem:
$$\begin{array}{rl} \text{find} & \boldsymbol{X} \in \mathbb{R}^{2\times n}, \boldsymbol{Y} \in \mathbb{R}^{n\times n} \\ \text{s.t.} & Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss}, \\ & Y_{ii} - 2\boldsymbol{x}_i^T\boldsymbol{a}_j + \boldsymbol{a}_j^T\boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \\ & \boldsymbol{Y} \succeq \boldsymbol{X}^T\boldsymbol{X}. \end{array} \tag{SNL-SDR}$$
- Note that $\operatorname{rank}(\boldsymbol{Z}) \leq 2 \iff \boldsymbol{Y} = \boldsymbol{X}^T\boldsymbol{X} \iff \operatorname{rank}(\boldsymbol{Y}) \leq 2$.
- **Remark:** Although we focus on 2-D localization, our techniques can be easily extended to handle $r$-D localization for any $r \geq 2$.

# 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)?

- The following complete characterization is obtained in **[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 at most $2$.
> – 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

- The following was also established in **[So-Ye'07]**:

  > Every rank $d \geq 2$ solution to (SNL-SDR) corresponds to a set of feasible (w.r.t. the distance constraints) $d$-dimensional coordinates for the sensors.

- Question: While it is NP-hard to find a rank-2 solution to (SNL-SDR), is it possible to find a low rank solution (and hence achieve dimension reduction)?

# Dimension Reduction via Unfolding

- 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):

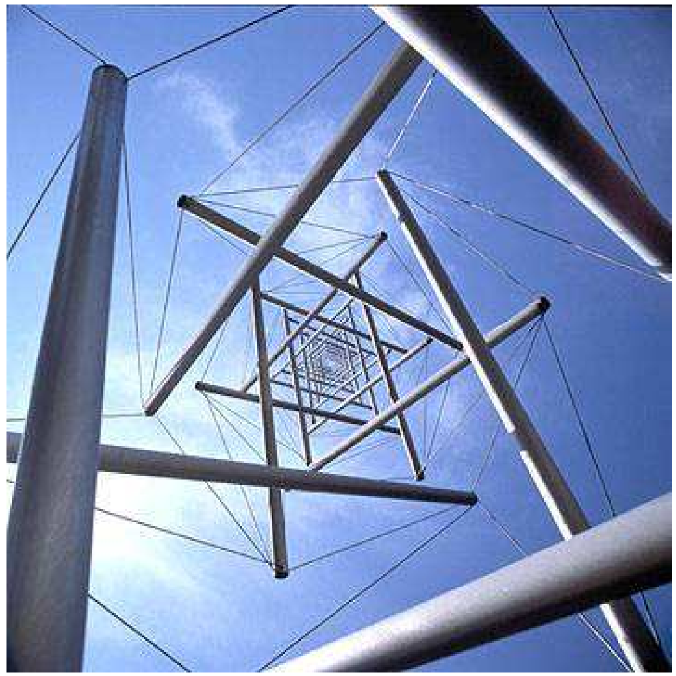
$$
\begin{aligned}
\max_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \mathbb{R}^2} \quad & \sum_{(i,j)\in N_{ss}} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 \\
\text{s.t.} \quad & \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 = d_{ij}^2, \quad (i,j) \in E_{ss}, \\
& \|\boldsymbol{x}_i - \boldsymbol{a}_j\|^2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

  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.



# Dimension Reduction via SDP Rank Reduction Theory

- If distortion of distances is allowed, then one can achieve dimension reduction using the SDP rank reduction theory in **[So-Ye-Zhang'08]**.
- To fix ideas, let us focus only on the sensor-sensor distance constraints:

$$\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2 = d_{ij}^2, \quad (i,j) \in E_{ss}. \tag{$\dagger$}$$

  Using the techniques introduced earlier, we obtain the following SDR of (†):

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

- Note that if $\{\bar{\boldsymbol{x}}_i\}_i$ is a feasible solution to (†), then $\bar{\boldsymbol{X}} = [\bar{X}_{ij}]$, where $\bar{X}_{ij} = \bar{\boldsymbol{x}}_i^T \bar{\boldsymbol{x}}_j$, is a feasible solution to (††).

  Conversely, if $\bar{\boldsymbol{X}}$ is a rank-two matrix satisfying (††), then we can extract from $\bar{\boldsymbol{X}}$ a feasible solution $\{\bar{\boldsymbol{x}}_i\}_i$ to (†) using Cholesky factorization.
- However, there is no guarantee that the solution returned by a polynomial-time algorithm for solving (††) is of rank-two.
  - Question: Can we extract an (approximate) rank-two solution?

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

- The answer is yes! The key lies in the following result from **[So-Ye-Zhang'08]**:

> Let $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \succeq \boldsymbol{0}$ and $b_1, \ldots, b_m \geq 0$ be given. Suppose there exists an $\boldsymbol{X}^\star \in \mathbb{S}^n$ such that
>
> $$\mathrm{Tr}(\boldsymbol{A}_i \boldsymbol{X}^\star) = b_i, \quad i = 1, \ldots, m; \quad \boldsymbol{X}^\star \succeq \boldsymbol{0}.$$
>
> Then, for any given $r \geq 1$, one can find in randomized polynomial time a rank-$r$ matrix $\hat{\boldsymbol{X}} \succeq \boldsymbol{0}$ such that
>
> $$\alpha \cdot b_i \leq \mathrm{Tr}(\boldsymbol{A}_i \hat{\boldsymbol{X}}) \leq \beta \cdot b_i, \quad i = 1, \ldots, m$$
>
> holds with high probability, where
>
> $$\alpha = \Omega\left(m^{-2/r}\right), \quad \beta = O\left(\frac{\ln m}{r}\right) \quad \text{when } r = O(\ln m), \text{ and}$$
>
> $$\alpha = 1 - O\left(\sqrt{\frac{\ln m}{r}}\right), \quad \beta = 1 + O\left(\sqrt{\frac{\ln m}{r}}\right) \quad \text{when } r = \Omega(\ln m).$$

## Application of the SYZ Theorem to the SNL Problem

- Observe that the system

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

is equivalent to

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

where $\boldsymbol{E}_{ij} = (\boldsymbol{e}_i - \boldsymbol{e}_j)(\boldsymbol{e}_i - \boldsymbol{e}_j)^T \succeq \boldsymbol{0}$. Here, $\boldsymbol{e}_i$ is the $i$th basis vector.

- Hence, by the SYZ theorem, we can find an $\hat{\boldsymbol{X}} \succeq \boldsymbol{0}$ with $\mathrm{rank}(\hat{\boldsymbol{X}}) \leq 2$ such that

$$\Omega\left(\frac{1}{|E_{ss}|}\right) d_{ij}^2 \leq \mathrm{Tr}(\boldsymbol{E}_{ij}\hat{\boldsymbol{X}}) \leq O(\ln(|E_{ss}|)) d_{ij}^2, \quad (i,j) \in E_{ss}.$$

- In particular, let $\hat{\boldsymbol{X}} = [\ \hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_n\ ]^T [\ \hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_n\ ]$ be the Cholesky factorization of $\hat{\boldsymbol{X}}$, where $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_n \in \mathbb{R}^2$. Then, we have

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

# Randomization Procedure for SDP Rank Reduction

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

> **Box 5.** ***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 generalizes our previous rank-1 and rank-2 Gaussian randomization procedures for beamforming problems.

## Extensions of the Basic SDR: Noisy Distance Measurements

- 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

$$\begin{aligned} d_{ij} &= \|\boldsymbol{x}_i - \boldsymbol{x}_j\| + \eta_{ij}, \\ \bar{d}_{ij} &= \|\boldsymbol{x}_i - \boldsymbol{a}_j\| + \bar{\eta}_{ij}, \end{aligned}$$

  where $\{\eta_{ij}\}$ (resp. $\{\bar{\eta}_{ij}\}$) are i.i.d. Gaussian random variables with mean $0$ and variance $\sigma_{ij}^2$ (resp. $\bar{\sigma}_{ij}^2$).

- In **[Biswas-Liang-Wang-Ye'06]**, the following maximum-likelihood (ML) SNL formulation is considered:

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

- 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

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

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

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

  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

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

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

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

(ML-SNL-SDR)

# Extensions of the Basic SDR: Uncertain Anchor Positions

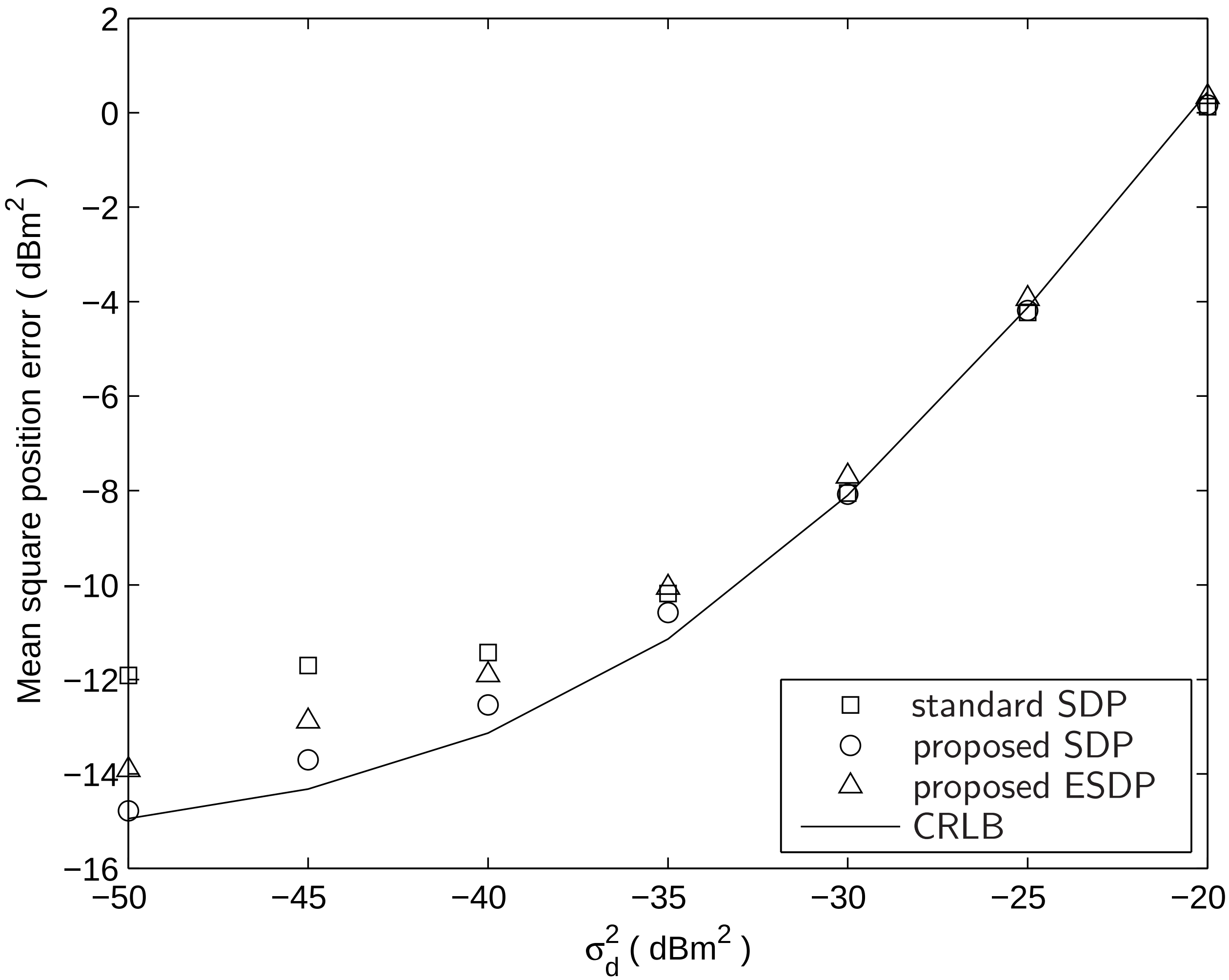
- SDR can also be employed to handle ML-SNL formulations with uncertain anchor locations, and/or with uncertain propagation speed **[Lui-Ma-So-Chan'09]** (the latter happens in underground sensor networks).
- For uncertain anchor locations, one could adopt the following error model:

$$\boldsymbol{a}_i = \bar{\boldsymbol{a}}_i + \boldsymbol{\varepsilon}_i,$$

  where $\bar{\boldsymbol{a}}_i$ is the presumed or nominal coordinate of anchor $i$, and $\boldsymbol{\varepsilon} \in \mathbb{R}^2$ is a Gaussian random vector with mean zero and covariance matrix $\boldsymbol{\Phi}_i$.
- Then, one has the following ML formulation:

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

  - This is a robust formulation where the anchors' uncertainties are accommodated by re-estimating $\boldsymbol{a}_i$.
  - It 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]**.

# 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]**:

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

- Note that (SNL-ESDR) has $|E_{ss}|$ $4 \times 4$ 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.

# Speeding Up the Computation: 2-D Single Source Localization

- Consider 2-D single source localization, a special case of SNL.

- The model reduces to

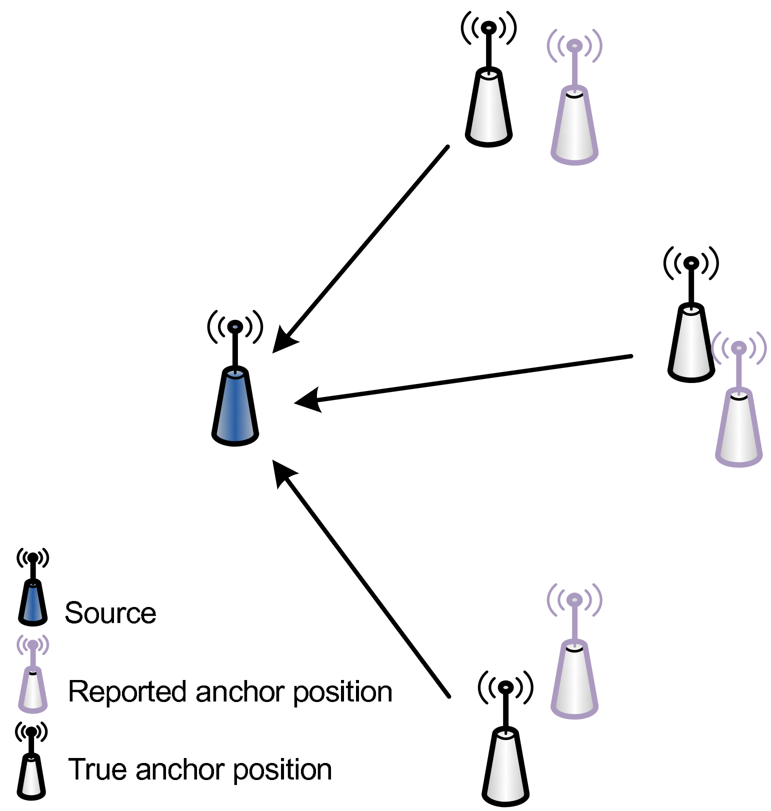
$$d_i = \|\boldsymbol{x} - \boldsymbol{a}_i\| + \eta_i, \quad i = 1, \ldots, m,$$

where $\boldsymbol{x} \in \mathbb{R}^2$ (resp. $\boldsymbol{a}_i \in \mathbb{R}^2$) is the source coordinate (resp. anchor $i$ coordinate); $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$ is noise.

- Uncertain anchor locations are assumed:

$$\boldsymbol{a}_i = \bar{\boldsymbol{a}}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, m,$$

$\bar{\boldsymbol{a}}_i \in \mathbb{R}^2$ is the presumed coordinate of anchor $i$; $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(\boldsymbol{0}, \lambda_i^2 \boldsymbol{I})$ is the uncertainty.



- Consider the ML source localization formulation (same as in SNL):

$$\min_{\substack{\boldsymbol{x}\in\mathbb{R}^2,\ \boldsymbol{a}_i\in\mathbb{R}^2,\\ \text{for } i=1,\ldots,m}} \sum_{i=1}^{m}\left(\frac{(\|\boldsymbol{x}-\boldsymbol{a}_i\|-d_i)^2}{\sigma_i^2}+\frac{\|\boldsymbol{a}_i-\bar{\boldsymbol{a}}_i\|^2}{\lambda_i^2}\right).$$

- Interestingly, the ML problem admits a more compact SDR.

- Indeed, as shown in **[Fu-Chan-Ma-So'12]**, the ML problem can be reformulated as a complex-valued constant-modulus quadratic program (CMQP):

$$\begin{aligned}\min_{\boldsymbol{z}\in\mathbb{C}^m} \quad & \boldsymbol{z}^H\boldsymbol{C}\boldsymbol{z}+2\mathrm{Re}\{\boldsymbol{c}^H\boldsymbol{z}\}\\ \text{s.t.} \quad & |z_i|^2=1,\ i=1,\ldots,m.\end{aligned}$$

  - It is very similar to the ML MIMO detection problem described in Part I.
  - The CMQP, as well as the ML MIMO problem, are very well structured.
  - For such nice problems, don't use `CVX`— we can exploit their problem structures to derive fast SDR solvers, e.g., **[Ma-Ching-Ding'04]**, **[Wai-Ma-So'11]**.

# How to Rewrite the ML Problem as a CMQP?

- We take the insight from **[Oğuz-Ekim-Gomes-Xavier-Oliverira'10]**. Recall

$$\min_{\substack{\boldsymbol{x}\in\mathbb{R}^2,\ \boldsymbol{a}_i\in\mathbb{R}^2,\\ \text{for } i=1,\ldots,m}} \sum_{i=1}^{m} \left( \frac{(\|\boldsymbol{x}-\boldsymbol{a}_i\| - d_i)^2}{\sigma_i^2} + \frac{\|\boldsymbol{a}_i - \bar{\boldsymbol{a}}_i\|^2}{\lambda_i^2} \right).$$

- Let $y = x_1 + jx_2$, $b_i = a_{i,1} + ja_{i,2}$, and $\bar{b}_i = \bar{a}_{i,1} + j\bar{a}_{i,2}$. We have

$$\min_{\substack{y\in\mathbb{C},\ b_i\in\mathbb{C}\\ \text{for } i=1,\ldots,m}} \sum_{i=1}^{m} \left( \frac{(|y-b_i| - d_i)^2}{\sigma_i^2} + \frac{|b_i - \bar{b}_i|^2}{\lambda_i^2} \right).$$

- By introducing additional phase shift variables $z_i \in \mathbb{C}$, the above problem can be equivalently written as

$$\begin{aligned} \min_{\substack{y\in\mathbb{C},\ b_i\in\mathbb{C},\ z_i\in\mathbb{C}\\ \text{for } i=1,\ldots,m}} \quad & \sum_{i=1}^{m} \left( \frac{|y-b_i-z_i d_i|^2}{\sigma_i^2} + \frac{|b_i - \bar{b}_i|^2}{\lambda_i^2} \right) \\ \text{s.t.} \quad & |z_i|^2 = 1,\ i = 1,\ldots,m. \end{aligned}$$

- Let $\boldsymbol{D} = \mathrm{Diag}(d_1, \ldots, d_m)$, $\boldsymbol{b} = [b_1, \ldots, b_m]^T$, $\bar{\boldsymbol{b}} = [\bar{b}_1, \ldots, \bar{b}_m]^T$, $\boldsymbol{z} = [z_1, \ldots, z_m]^T$, $\boldsymbol{\Sigma} = \mathrm{Diag}(\sigma_1, \ldots, \sigma_m)$, and $\boldsymbol{\Lambda} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_m)$. Then, the ML source localization problem becomes

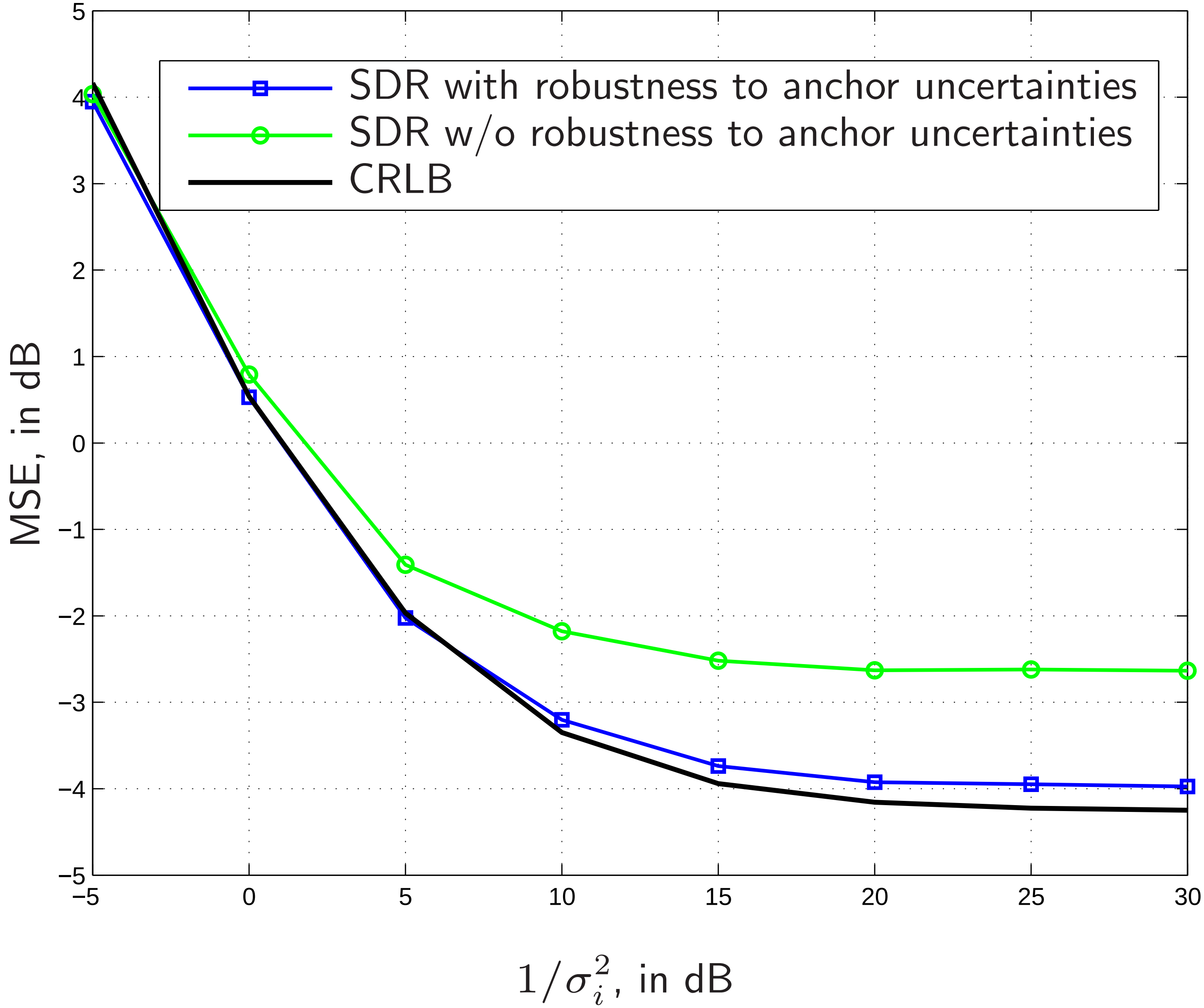
$$\begin{aligned} \min_{\substack{y \in \mathbb{C},\ \boldsymbol{b} \in \mathbb{C}^m \\ \boldsymbol{z} \in \mathbb{C}^m}} \quad & \|\boldsymbol{\Sigma}^{-1}(y\mathbf{1} - \boldsymbol{b} - \boldsymbol{D}\boldsymbol{z})\|^2 + \|\boldsymbol{\Lambda}^{-1}(\boldsymbol{b} - \bar{\boldsymbol{b}})\|^2 \\ \text{s.t.} \quad & |z_i|^2 = 1,\ i = 1, \ldots, m. \end{aligned} \tag{†}$$

- Fixing $\boldsymbol{z}$, the minimization of (†) w.r.t. $\boldsymbol{b}$ & $y$ admits a closed form solution.

- By marginalizing (†) w.r.t. $\boldsymbol{b}$ & $y$, the ML problem can be expressed as the CMQP

$$\begin{aligned} \min_{\boldsymbol{z} \in \mathbb{C}^m} \quad & \boldsymbol{z}^H \boldsymbol{C} \boldsymbol{z} + 2\mathrm{Re}\{\boldsymbol{c}^H \boldsymbol{z}\} \\ \text{s.t.} \quad & |z_i|^2 = 1,\ i = 1, \ldots, m, \end{aligned}$$

where $\boldsymbol{C} = \boldsymbol{D}\boldsymbol{G}\boldsymbol{D} - \boldsymbol{D}\boldsymbol{G}(\boldsymbol{G} + \boldsymbol{\Lambda}^{-2})^{-1}\boldsymbol{G}\boldsymbol{D}$, $\boldsymbol{c} = \boldsymbol{D}\boldsymbol{G}(\boldsymbol{G} + \boldsymbol{\Lambda}^{-2})^{-1}\boldsymbol{\Lambda}^{-2}\bar{\boldsymbol{b}}$, $\boldsymbol{G} = \boldsymbol{\Sigma}^{-1}\boldsymbol{P}^{\perp}\boldsymbol{\Sigma}^{-1}$ and $\boldsymbol{P}^{\perp} = \boldsymbol{I} - (\boldsymbol{\Sigma}^{-1}\mathbf{1})(\boldsymbol{\Sigma}^{-1}\mathbf{1})^{\dagger}$.




MSE Performance of CQMP-based SDR for 2-D single source localization. $m = 6$, $\mathbf{\Lambda} = \mathrm{Diag}(1, 0.8, 0.6, 1.2, 0.5, 1)$; the anchors are uniformly distributed in a circle surrounding the source, with radius $r = 5$.

# Beyond SDR: What Else?

- The hard problems discussed in this tutorial so far, namely, nonconvex QCQPs, can all be reformulated as rank-constrained SDPs; i.e., optimization problems of the form

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

  where $\boldsymbol{C}, \boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathbb{S}^n$, $b_1, \ldots, b_m \in \mathbb{R}$, and $r \ge 1$ are given.

- The SDR methodology entails removing the hard rank constraint in (RCSDP). The resulting problem is then an SDP, which is polynomial-time solvable.

- However, we can no longer guarantee that the SDP solution has the desired rank.

# Beyond SDR: What Else?

- To recover some of the effects of the rank constraint, one natural idea is to use regularizers. Specifically, consider the problem

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

  where

  - $f : \mathbb{S}^n_+ \to \mathbb{R}$ is a function that favors low-rank matrices (i.e., $f(\boldsymbol{X})$ tends to take smaller values when $\boldsymbol{X}$ is of low rank), and
  - $\mu > 0$ is a parameter controlling the effect of the regularizer $f$.

- Question: Which regularizer to use?

# Choosing a Regularizer

- Ideally, we want $f(\boldsymbol{X}) = \mathrm{rank}(\boldsymbol{X})$, but the resulting problem is just as hard as (RCSDP).
  - The function $\boldsymbol{X} \mapsto \mathrm{rank}(\boldsymbol{X})$ is not only nonconvex but also discontinuous.
  - This motivates us to look for tractable surrogates of the rank function.

- In view of the recent work on low-rank matrix completion/recovery, one may consider using the *trace norm*; i.e., $f(\boldsymbol{X}) = \mathrm{Tr}(\boldsymbol{X})$.
  - The resulting regularized problem (RP) becomes an SDP, which is polynomial-time solvable.
  - Perhaps somewhat surprisingly, such a choice does not always yield good low-rank solutions empirically, especially for the SNL problem.
  - In fact, existing theoretical recovery results concerning the trace norm do not apply to most of the problems discussed in this tutorial.

# Choosing a Regularizer

- To better induce low-rank solutions, let us consider the regularizer $f_p : \mathbb{S}^n_+ \to \mathbb{R}_+$ defined by

$$f_p(\boldsymbol{X}) = \mathrm{Tr}(\boldsymbol{X}^p) = \sum_{i=1}^{n} \lambda_i^p(\boldsymbol{X}),$$

  where $p \in (0,1)$ is fixed and $\lambda_i(\boldsymbol{X})$ is the $i$th largest eigenvalue of $\boldsymbol{X}$.

- The function $f_p$ is known as the Schatten $p$-quasi-norm of $\boldsymbol{X}$.

- Since $f_p(\boldsymbol{X}) \to \mathrm{rank}(\boldsymbol{X})$ as $p \searrow 0$ and $f_p$ is continuous for all $p \in (0,1)$, the Schatten quasi-norm can be viewed as a continuous surrogate of $\mathrm{rank}(\boldsymbol{X})$.

# Schatten Quasi-Norm Regularized SDP

- By fixing a $p \in (0,1)$, we obtain the following Schatten $p$-quasi-norm regularized SDP:

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

  (For concreteness, we replace '$\trianglerighteq_i$' by '$=$' for all $i$.)

- The function $f_p$ is concave on $\mathbb{S}^n_+$, but we are minimizing in (SRP$_p$)...

- In fact, for each $p \in (0,1)$, (SRP$_p$) is NP-hard **[Ge-Jiang-Ye'11]**, so it seems that we are back to square one.

# Schatten Quasi-Norm Regularized SDP

- However, not all is lost, as we have the following result from **[Ji-Sze-Zhou-So-Ye'13]**:

> Let $p \in (0, 1)$ be fixed. Then, a point satisfying the first-order optimality conditions of (SRP$_p$) can be found in polynomial time.

  – Since $f_p$ is not everywhere differentiable, some care is needed when deriving the first-order optimality conditions.

- The above result is established in two steps:
  – Design an interior-point algorithm for (SRP$_p$).
  – Show that the algorithm returns a first-order point in polynomial time.

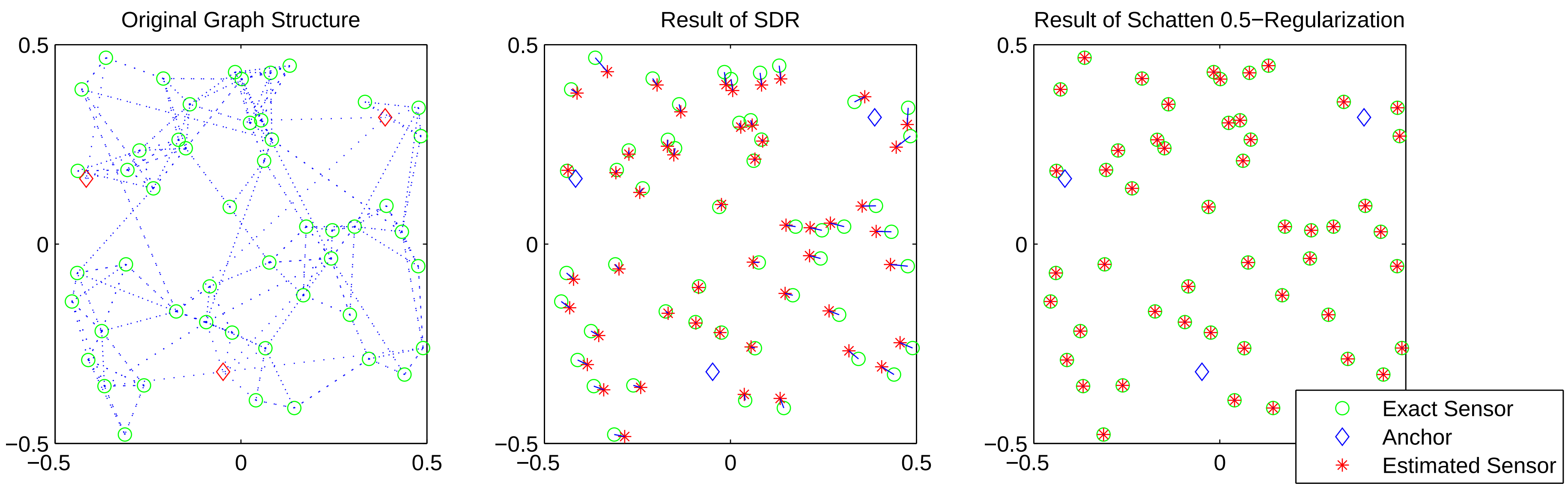
- Question: Since the point returned by the algorithm only satisfies the first-order optimality conditions of (SRP$_p$), is it any good?

## Application: Regularizing the SDR of the SNL Problem

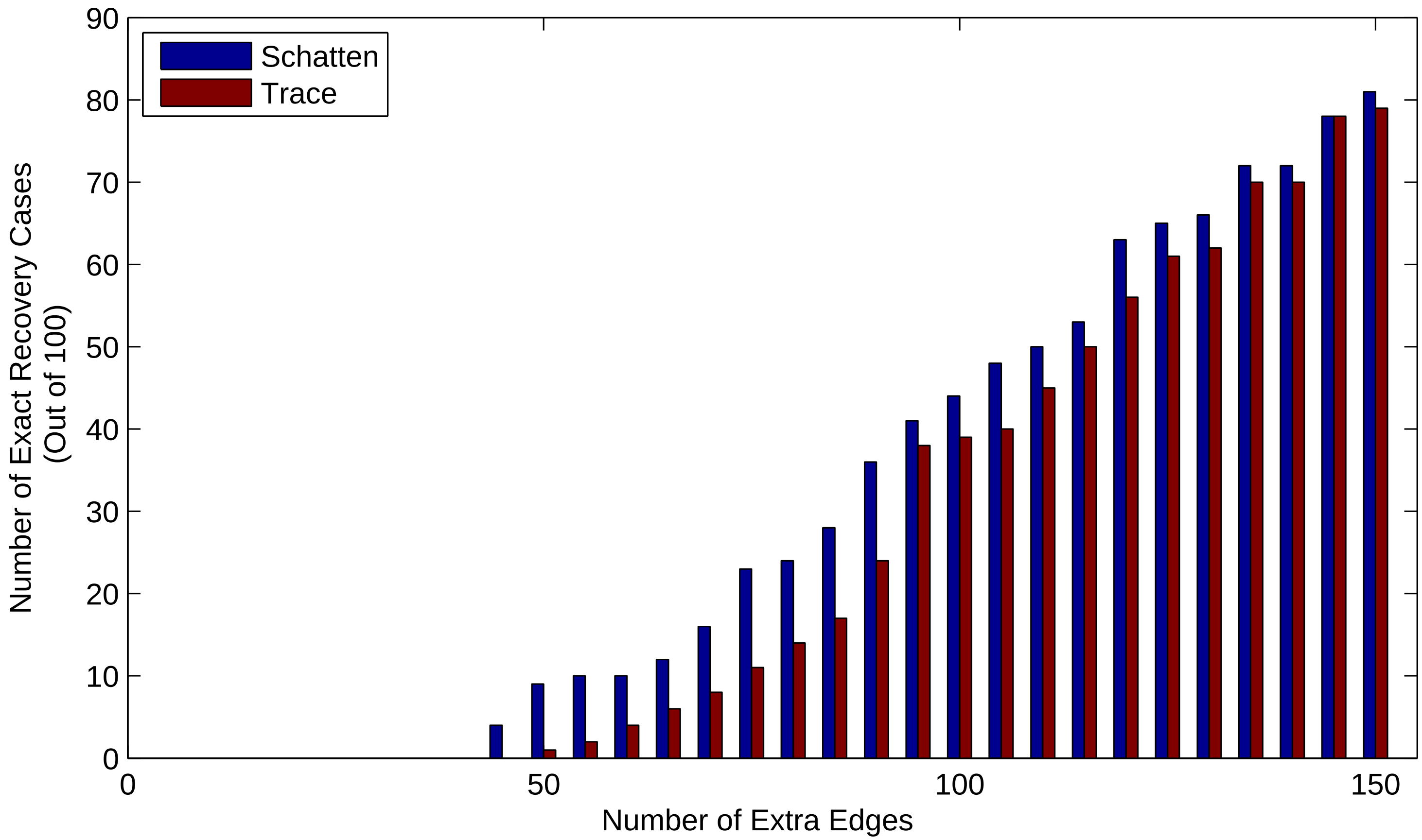
- Consider the following Schatten quasi-norm regularized SDR of the SNL problem:

$$
\begin{array}{rl}
\min & f_p(\boldsymbol{Z}) \\
\text{s.t.} & Y_{ii} - 2Y_{ij} + Y_{jj} = d_{ij}^2, \quad (i,j) \in E_{ss}, \\
& Y_{ii} - 2\boldsymbol{x}_i^T \boldsymbol{a}_j + \boldsymbol{a}_j^T \boldsymbol{a}_j = \bar{d}_{ij}^2, \quad (i,j) \in E_{sa}, \\
& \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{X} \\ \boldsymbol{X}^T & \boldsymbol{Y} \end{bmatrix} \succeq \boldsymbol{0}.
\end{array}
\qquad (\text{SNL-SRP}_p)
$$

- The aforementioned interior-point algorithm can be applied to this problem.




Performance of the plain SDR (SNL-SDR) and the Schatten $0.5$-quasi-norm regularized SDR (SNL-SRP$_{0.5}$). The latter is handled by the interior-point algorithm developed in **[Ji-Sze-Zhou-So-Ye'13]**.




Number of $50$-sensor globally rigid instances (out of a total of 100) for which a rank-$2$ solution is obtained by the trace norm regularized SDR (SNL-SRP$_1$) and the Schatten $0.5$-quasi-norm regularized SDR (SNL-SRP$_{0.5}$), as a function of number of edges added. Details can be found in **[Ji-Sze-Zhou-So-Ye'13]**.

# Conclusion

# Conclusion

- This tutorial has provided an overview of SDR, from practical deployments, applications, theoretical results to latest advances.

- We hope you would be convinced that SDR holds great potential in its wide scope of applicability, and in its powerful approximation accuracies.

- Many researchers have found their SDR applications. We hope you would find yours too in the future.

- We did not cover details of one application— MIMO detection, which is not only important but also elegant— owing to time limitation. We nevertheless append the MIMO detection topic as a "bonus material" in this tutorial slides, and hope you will find them useful.

# Conclusion

- This tutorial has provided an overview of SDR, from practical deployments, applications, theoretical results to latest advances.

- We hope you would be convinced that SDR holds great potential in its wide scope of applicability, and in its powerful approximation accuracies.

- Many researchers have found their SDR applications. We hope you would find yours too in the future.

- We did not cover details of one application— MIMO detection, which is not only important but also elegant— owing to time limitation. We nevertheless append the MIMO detection topic as a "bonus material" in this tutorial slides, and hope you will find them useful.

# Thank you!

# Bonus Material: 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. |

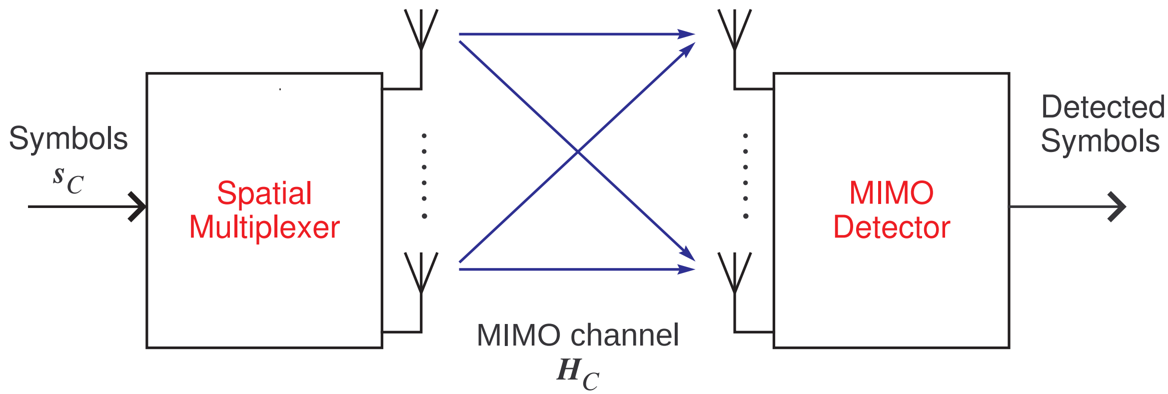
- We will focus on the ML detection problem

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

- 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

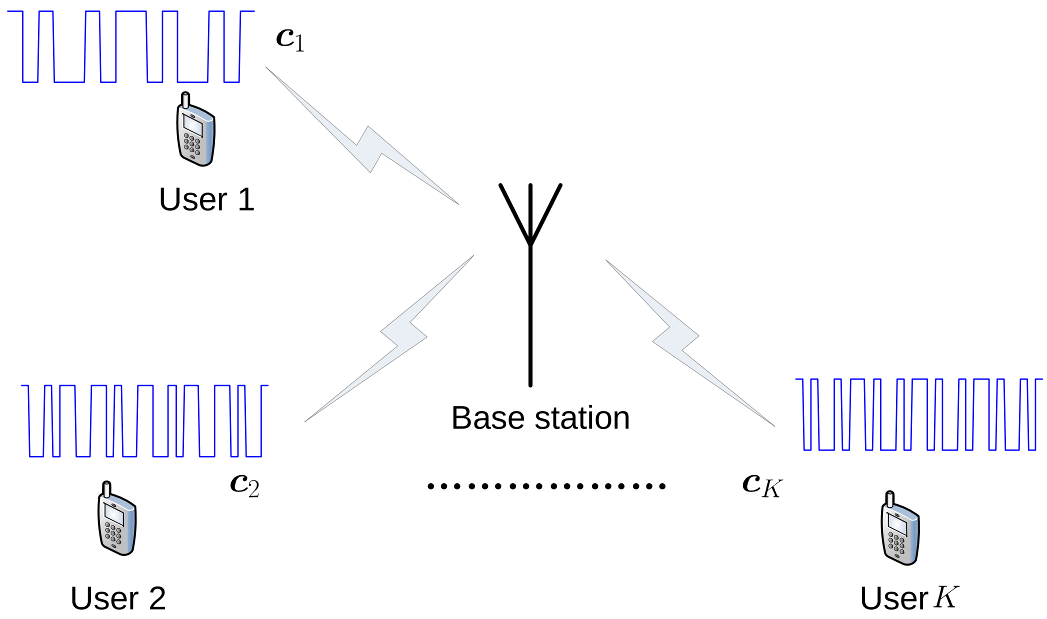
- The simple MIMO model $\boldsymbol{y}_C = \boldsymbol{H}_C \boldsymbol{s}_C + \boldsymbol{v}_C$ is popularly used in the point-to-point spatial multiplexing scenario.




- 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:

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

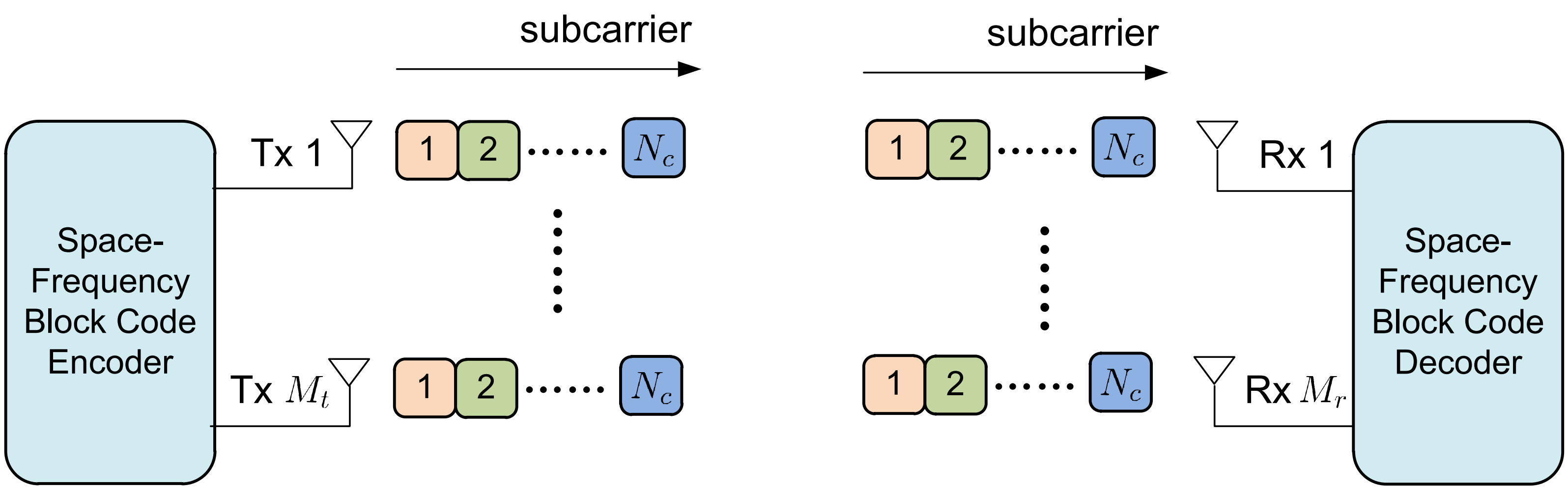
- The rx model can be converted to the generic MIMO form:

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

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

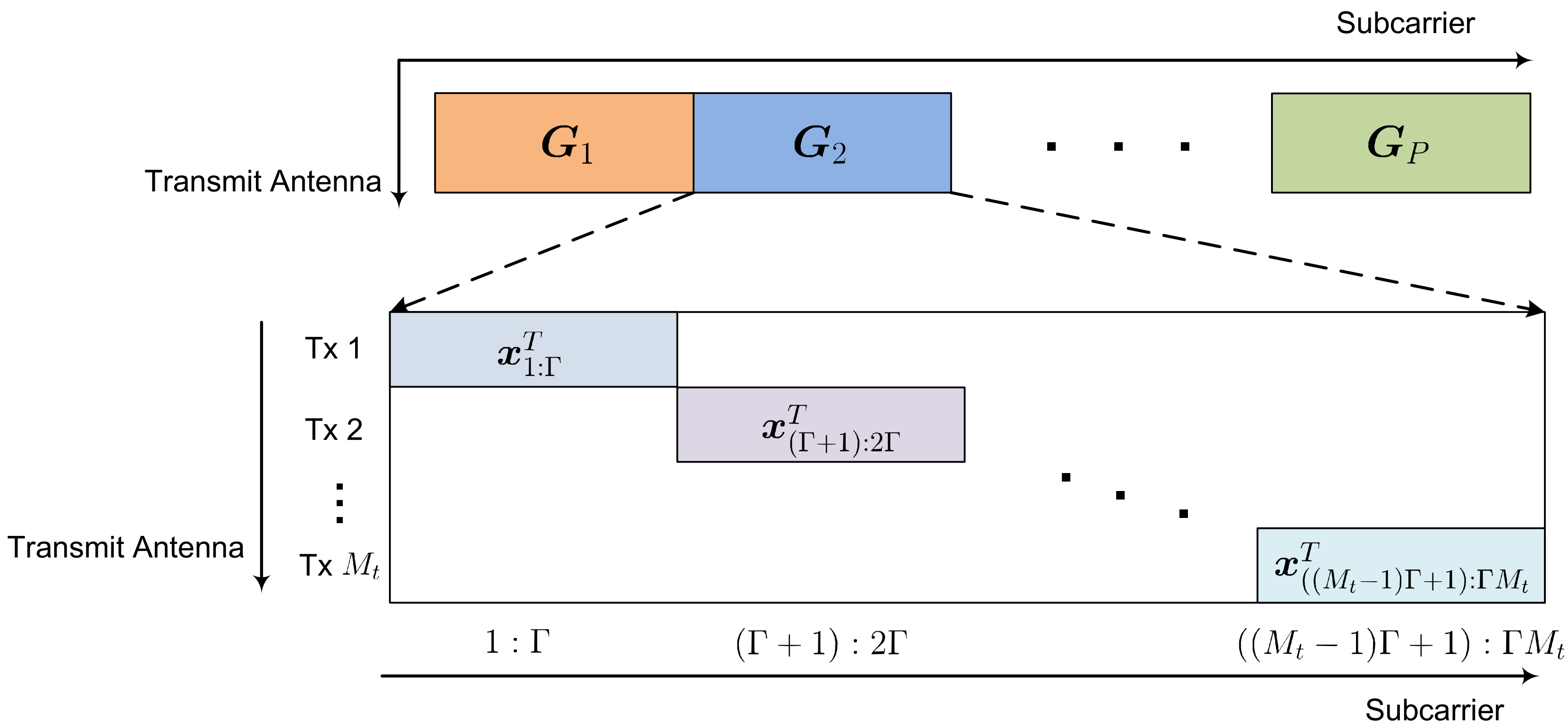
# Example: Space-Time Frequency Coding

- **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

$$
\begin{aligned}
\boldsymbol{y} &= \boldsymbol{D}_{\boldsymbol{H}}\boldsymbol{x} + \boldsymbol{v}, \\
\boldsymbol{x} &= \boldsymbol{\Theta}\boldsymbol{s},
\end{aligned}
$$

  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; |
  | $\boldsymbol{D}_{\boldsymbol{H}} \in \mathbb{C}^{\Gamma M_t \times \Gamma M_t}$ | 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, $\Gamma M_t$, 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

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

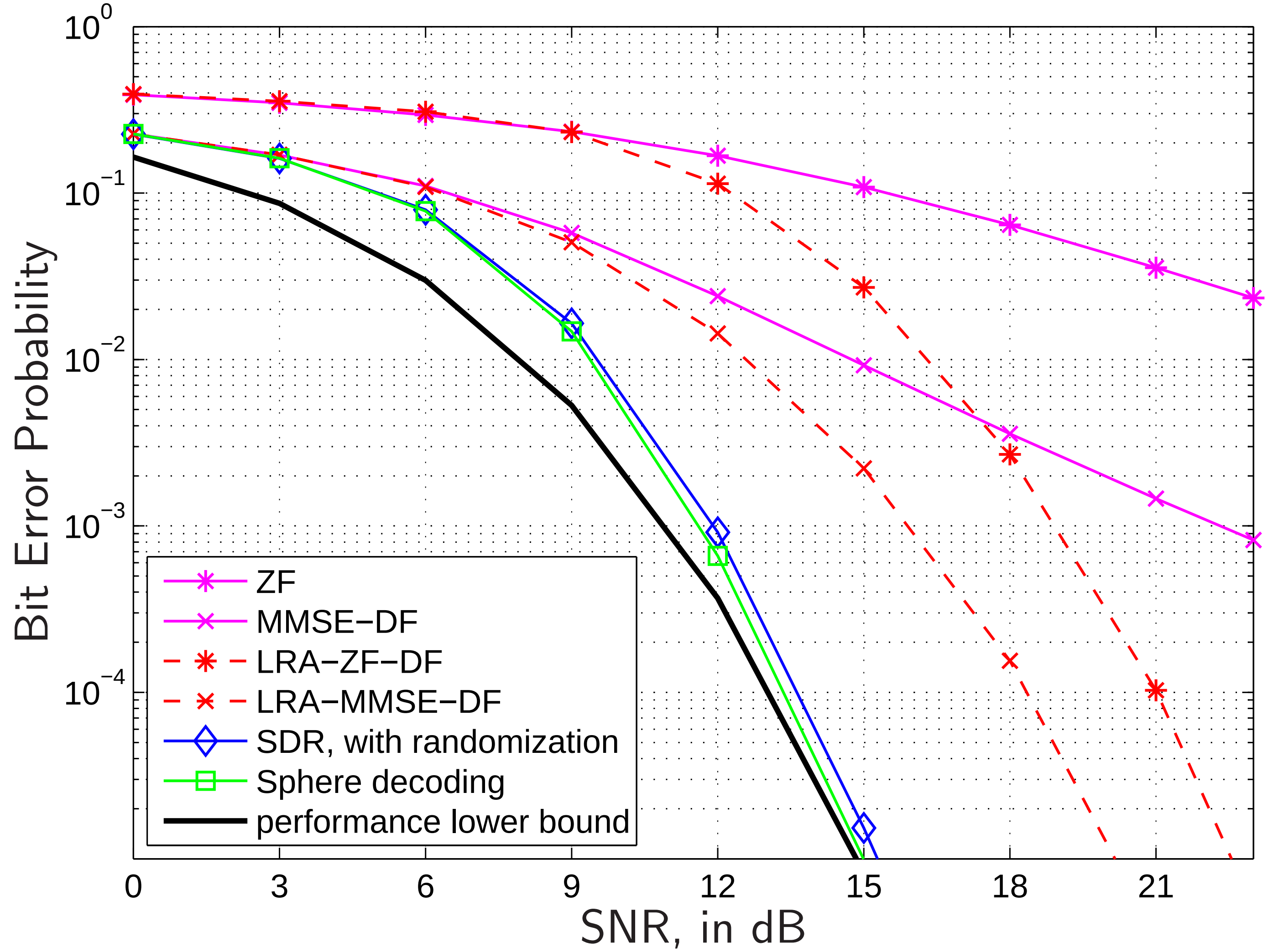
  where $\boldsymbol{s} \in \{\pm 1\}^N$ for QPSK constellations.

- ML detection problem:

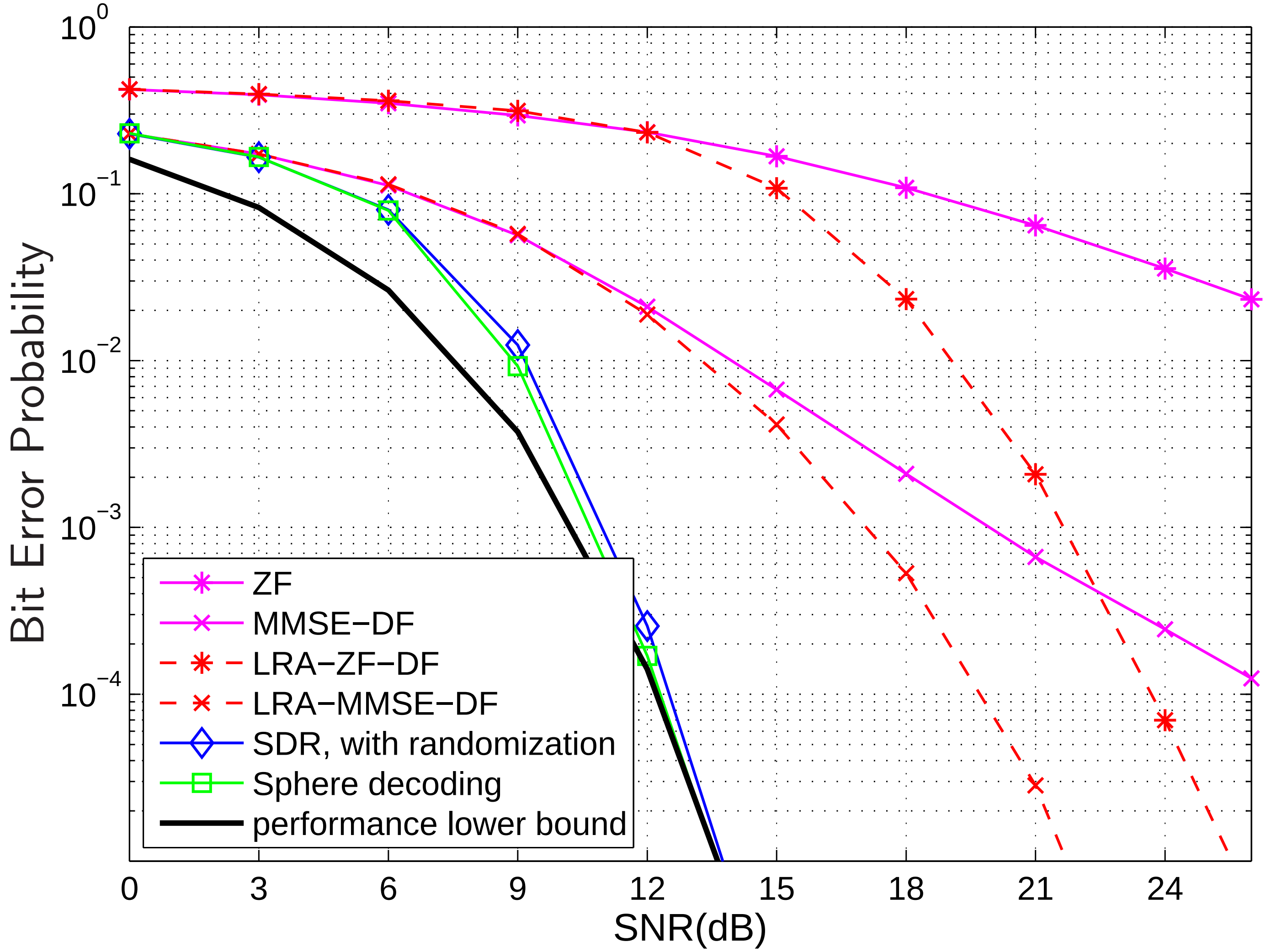
$$\begin{aligned} \min_{\boldsymbol{s} \in \mathbb{R}^N} \ & \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2 \\ \text{s.t.} \ & s_i^2 = 1, \ i = 1, \ldots, N. \end{aligned}$$

- SDR:

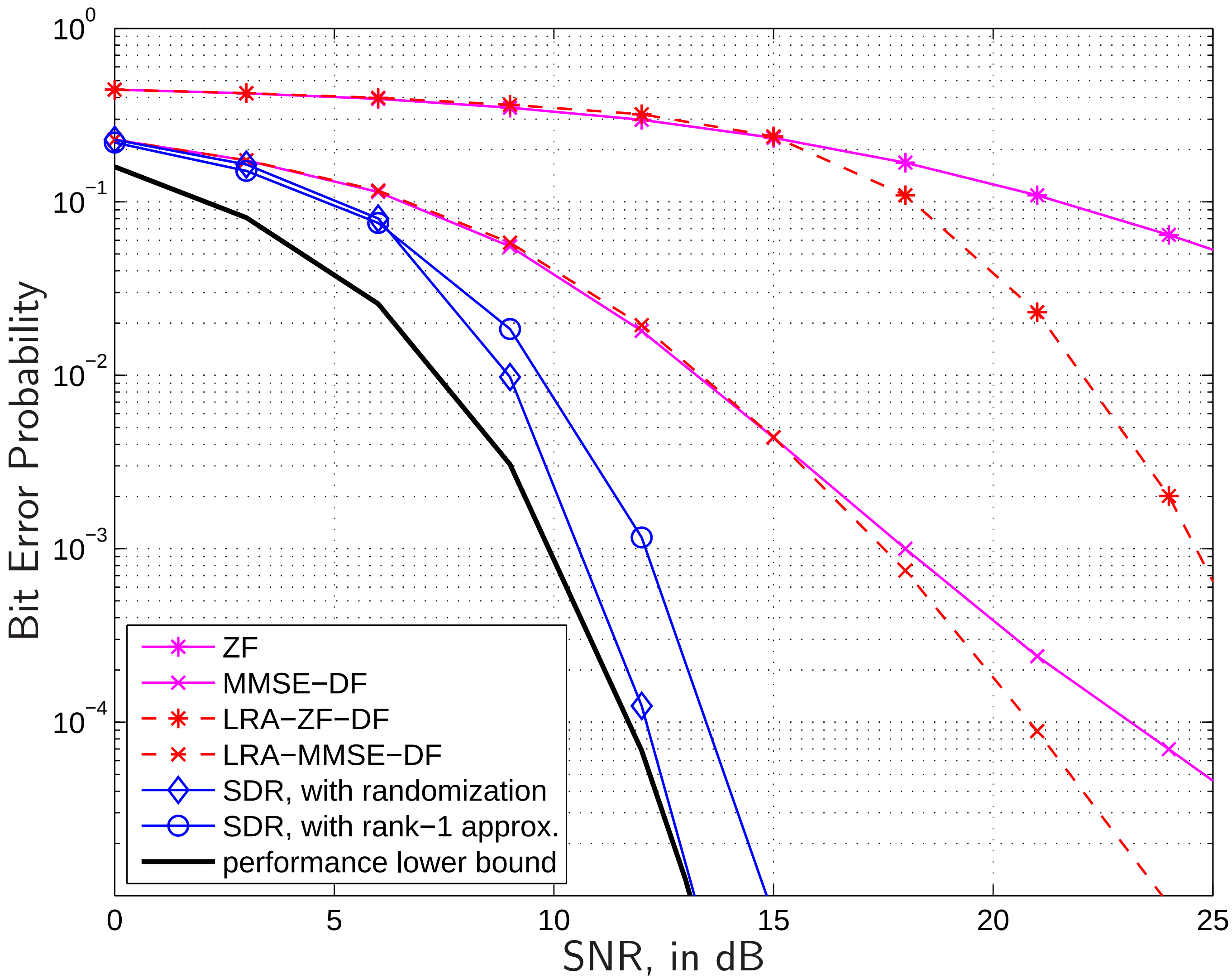
$$\begin{aligned} \min_{\boldsymbol{S} \in \mathbb{S}^N, \boldsymbol{s} \in \mathbb{R}^N} \ & \mathrm{Tr}(\boldsymbol{H}^T \boldsymbol{H} \boldsymbol{S}) - 2\boldsymbol{s}^T \boldsymbol{H}^T \boldsymbol{y} + \|\boldsymbol{y}\|^2 \\ \text{s.t.} \ & S_{ii} = 1, \ i = 1, \ldots, N, \\ & \boldsymbol{S} \succeq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



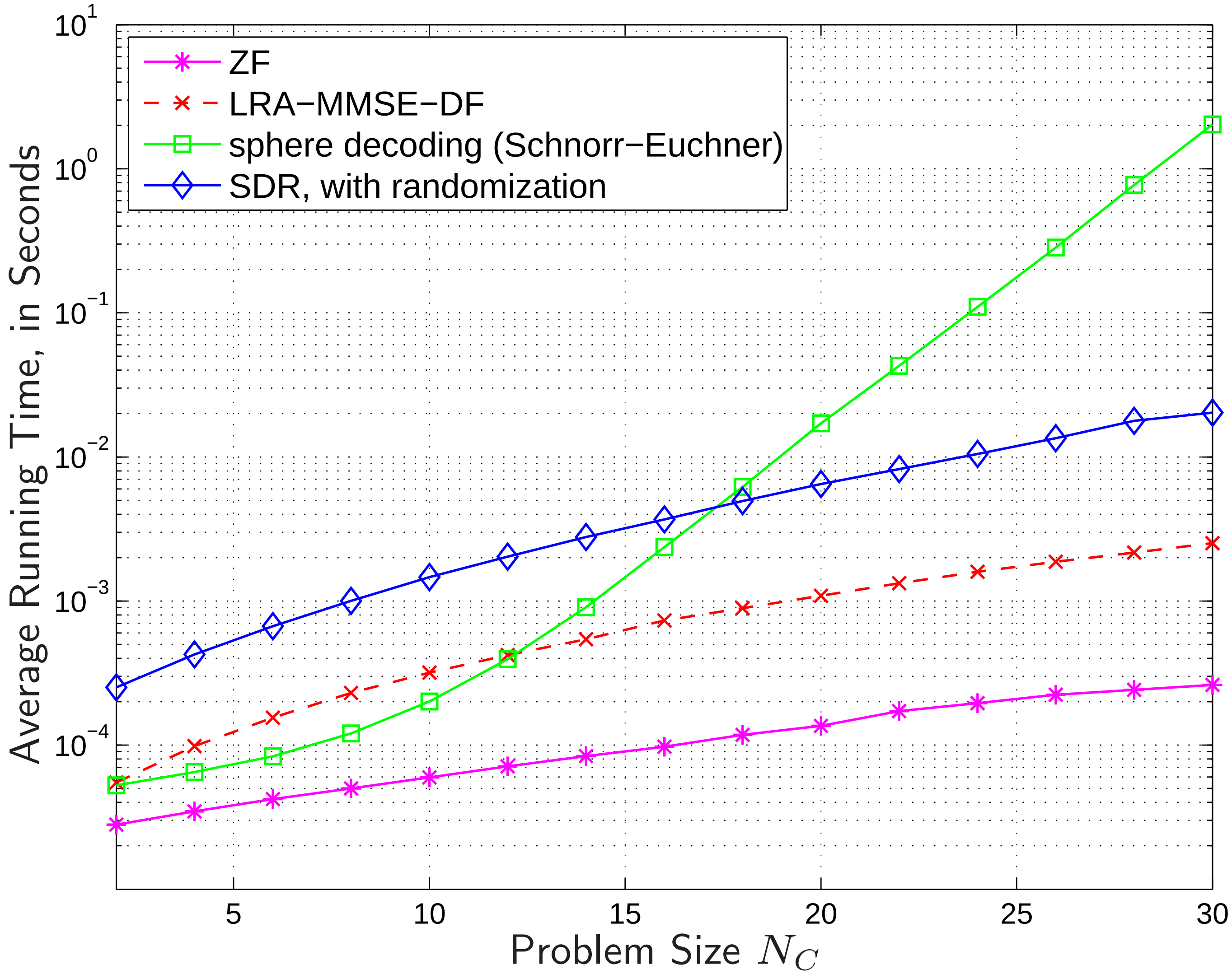

Bit error rate performance under $(M_C, N_C) = (10, 10)$, QPSK constellations. The SNR is defined as $\frac{\mathrm{E}\{\|\boldsymbol{H}_C \boldsymbol{s}_C\|^2\}}{\mathrm{E}\{\|\boldsymbol{v}_C\|^2\}}$. 'ZF'— zero forcing; 'MMSE-DF'— min. mean square error with decision feedback; 'LRA'— lattice reduction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

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

  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.

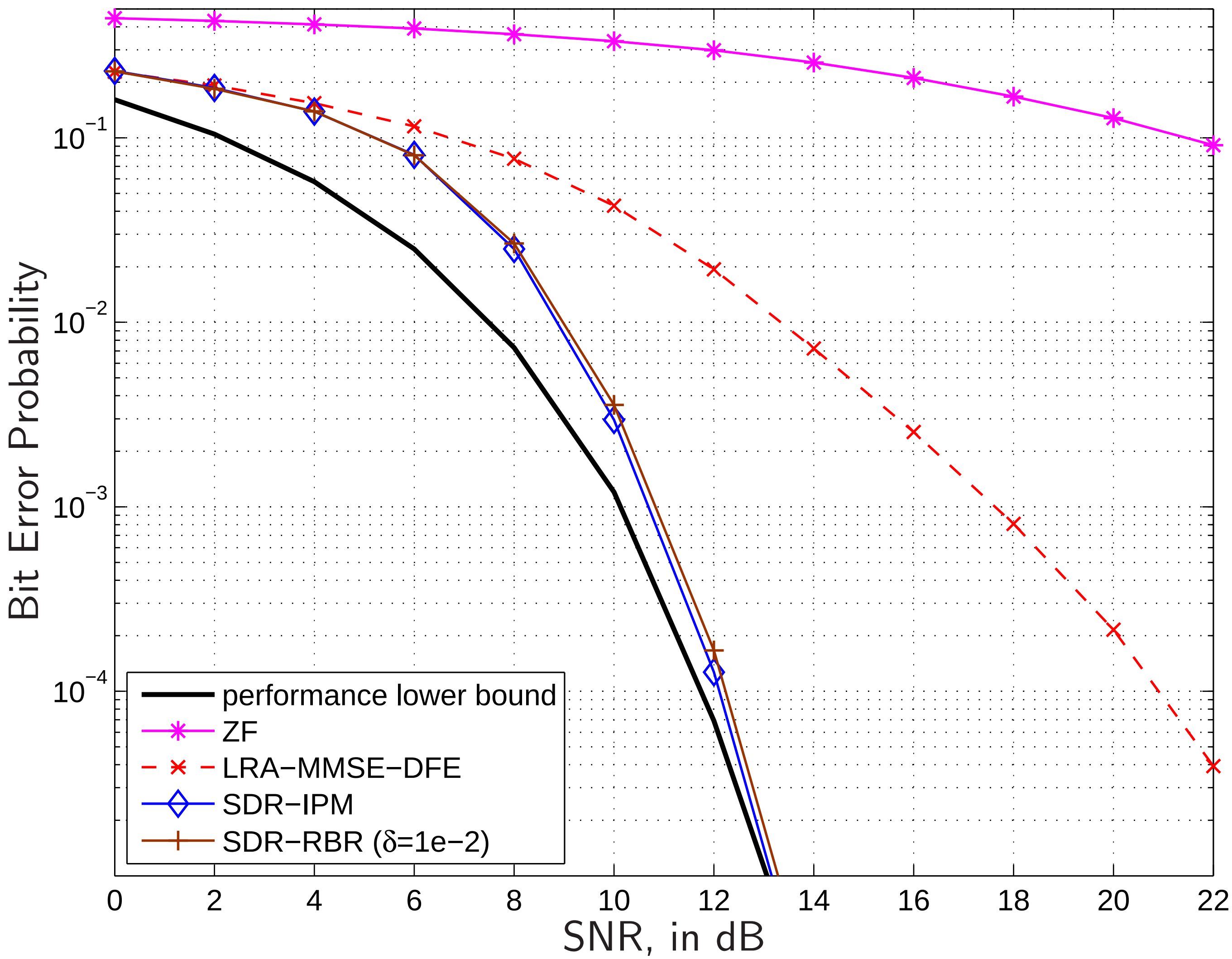
- Let $f(\boldsymbol{X}) = \mathrm{Tr}(\boldsymbol{C}\boldsymbol{X}) - \sigma \log\det(\boldsymbol{X})$, $\boldsymbol{x}_i$ be the $i$th row of $\boldsymbol{X}$, & $\boldsymbol{X}_{-i}$ be the collection of all elements of $\boldsymbol{X}$ except for $\boldsymbol{x}_i$.
- **Idea of RBR**: do a block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

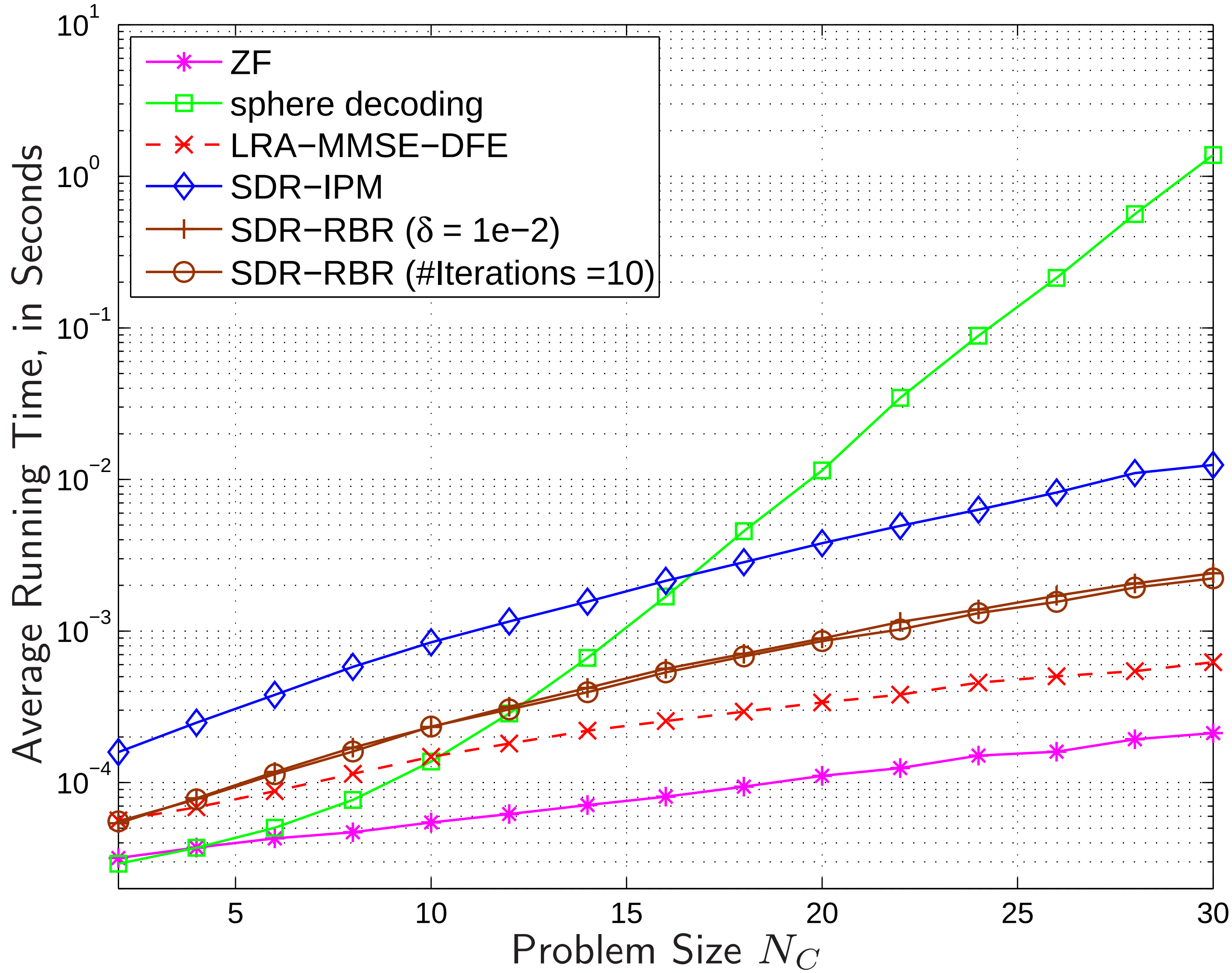
- The iterates are known to converge to the optimal solution of (B-SDR).
- Each per-row update is simple; e.g., the 1st row update can be equiv. written as

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

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 $\kappa$, 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 $\hat{\boldsymbol{s}}_{\mathsf{OSR}} = (\boldsymbol{H}^T\boldsymbol{H} + \gamma\boldsymbol{I})^{-1}\boldsymbol{H}^T\boldsymbol{y}$ for some $\gamma$; 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

$$f^\star_{\sf ML} = \min_{\boldsymbol{s} \in \{\pm 1\}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2,$$

$$f^\star_{\sf SDR} = \min_{\boldsymbol{S} \succeq \boldsymbol{s}\boldsymbol{s}^T,\ S_{ii}=1\ \forall i} \mathrm{Tr}(\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{S}) - 2\boldsymbol{s}^T\boldsymbol{H}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2,$$

$$f^\star_{\sf BR} = \min_{s_i^2 \le 1,\ i=1,\ldots,N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2,$$

$$f^\star_{\sf OSR} = \min_{\|\boldsymbol{s}\|^2 = N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2, \quad f^\star_{\sf UR} = \min_{\boldsymbol{s} \in \mathbb{R}^N} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}\|^2.$$

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

$$\max\{f^\star_{\sf UR}, f^\star_{\sf OSR}, f^\star_{\sf BR}\} \le f^\star_{\sf SDR} \le f^\star_{\sf ML}.$$

- 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.:

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

  where $g_{\sf ML}^\star$ 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):

$$
f_{\sf SDR}^\star = g_{\sf ML}^\star.
$$

  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

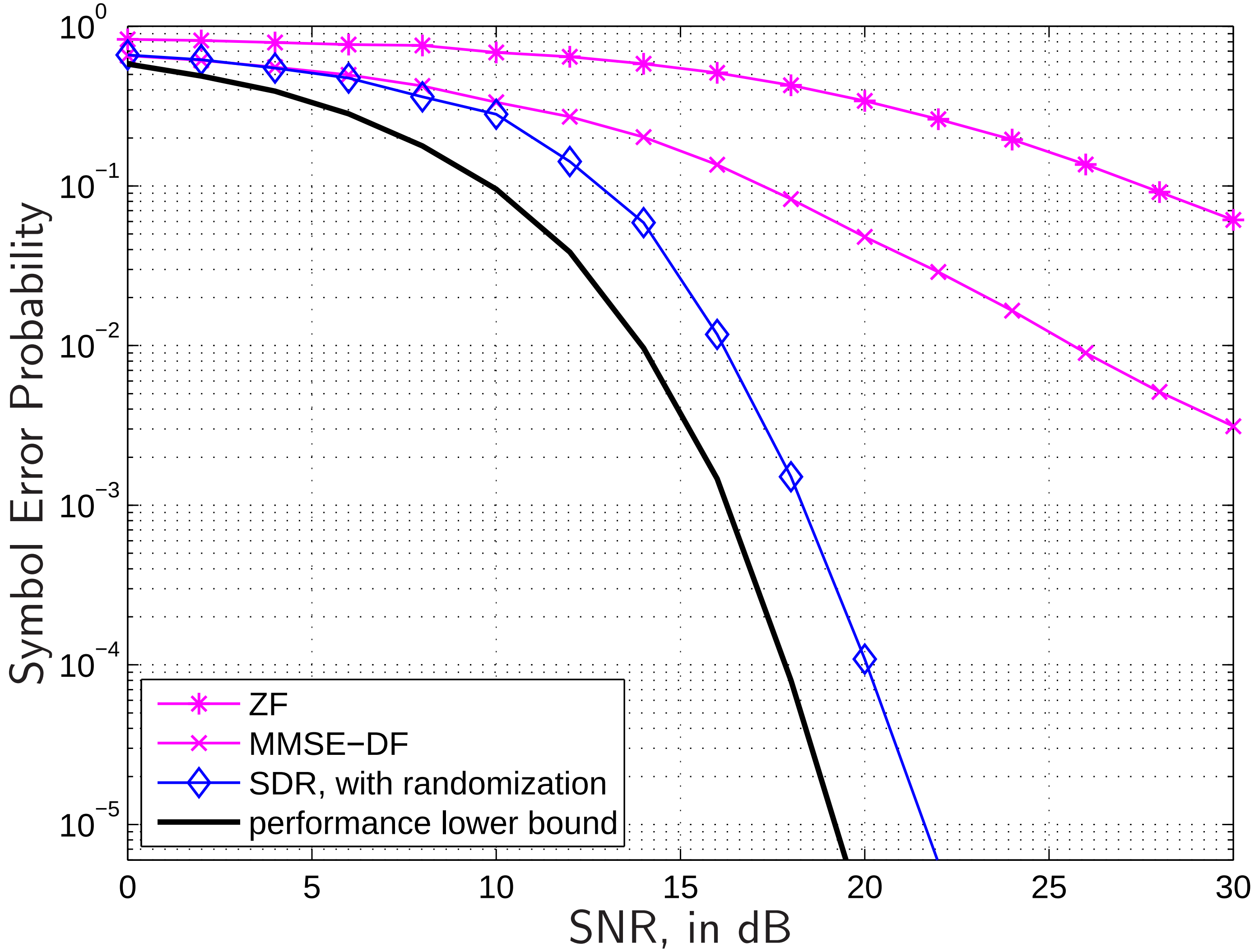
- The ML problem in the MPSK case:

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

- 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

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

to reformulate the ML problem as

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

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:

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

## 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:

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

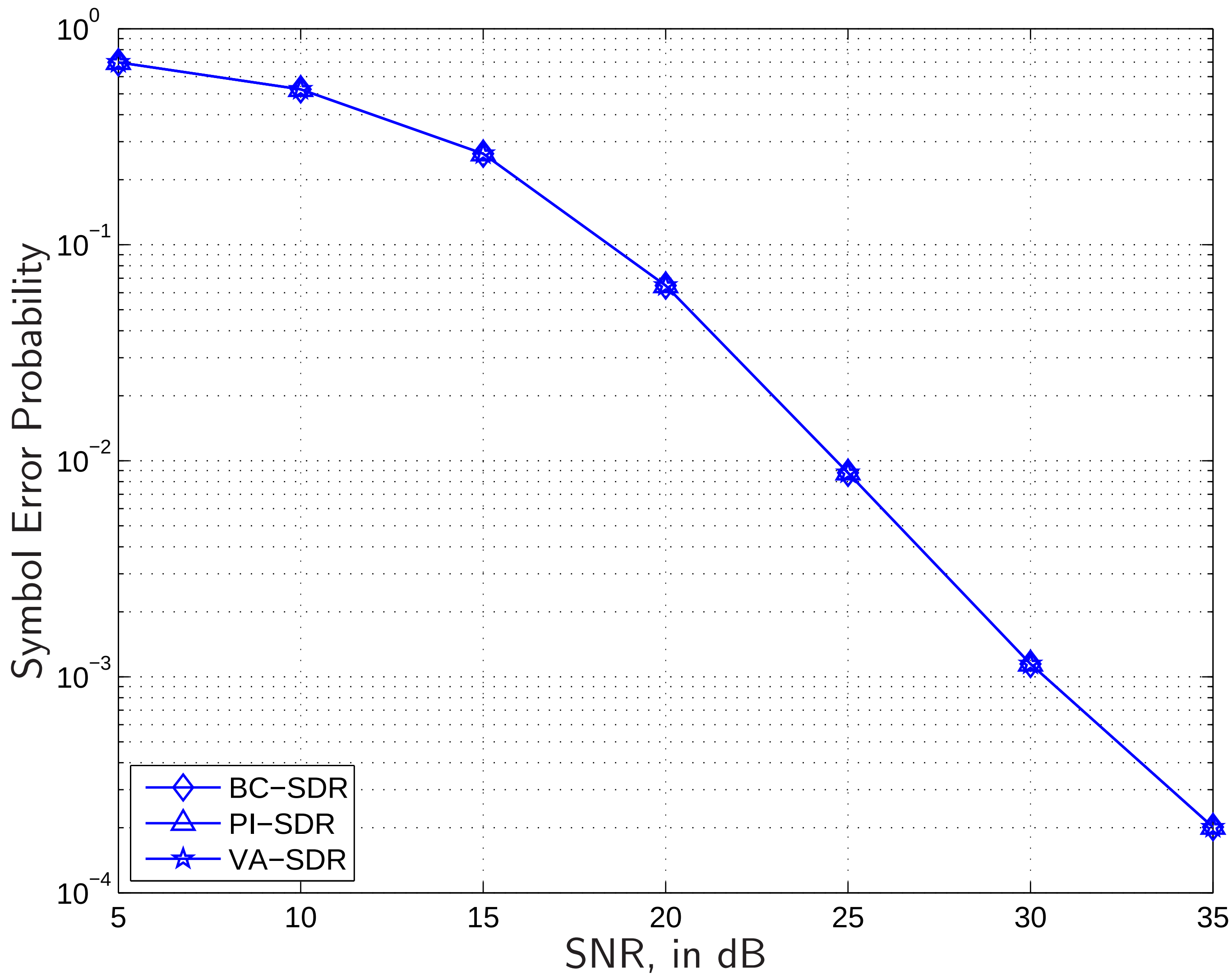
$$\mathcal{F}_{\mathsf{PI-SDR}} = \{\ (\mathbf{S},\mathbf{s}) \mid (\mathbf{U},\mathbf{u},\mathbf{S},\mathbf{s}) \in \mathcal{W}_{\mathsf{PI-SDR}}\ \},$$
$$\mathcal{W}_{\mathsf{PI-SDR}} = \{\ (\mathbf{U},\mathbf{u},\mathbf{S},\mathbf{s}) \mid \mathbf{U} \succeq \mathbf{u}\mathbf{u}^T, \mathbf{S} \succeq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

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

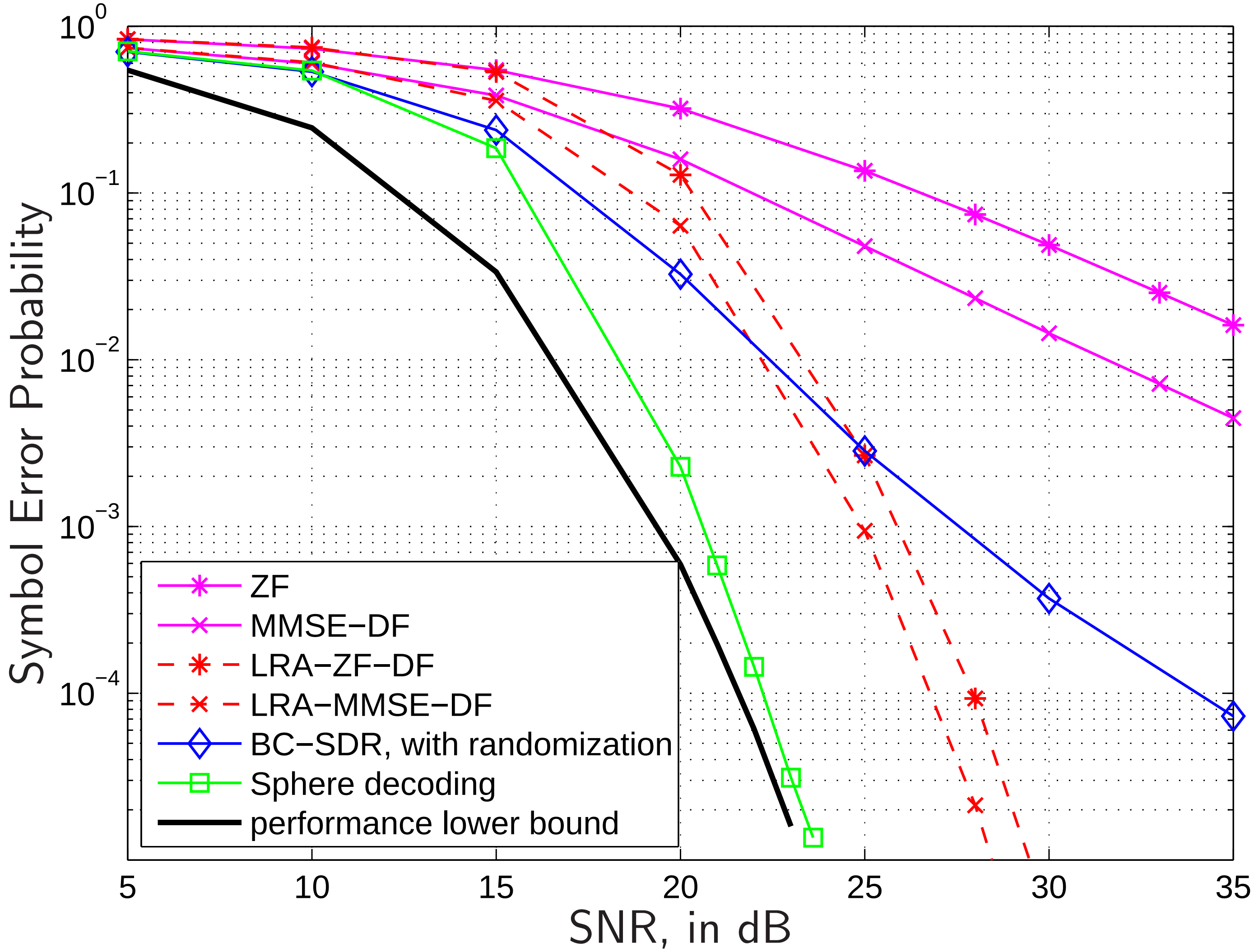
- It is shown by analysis that

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

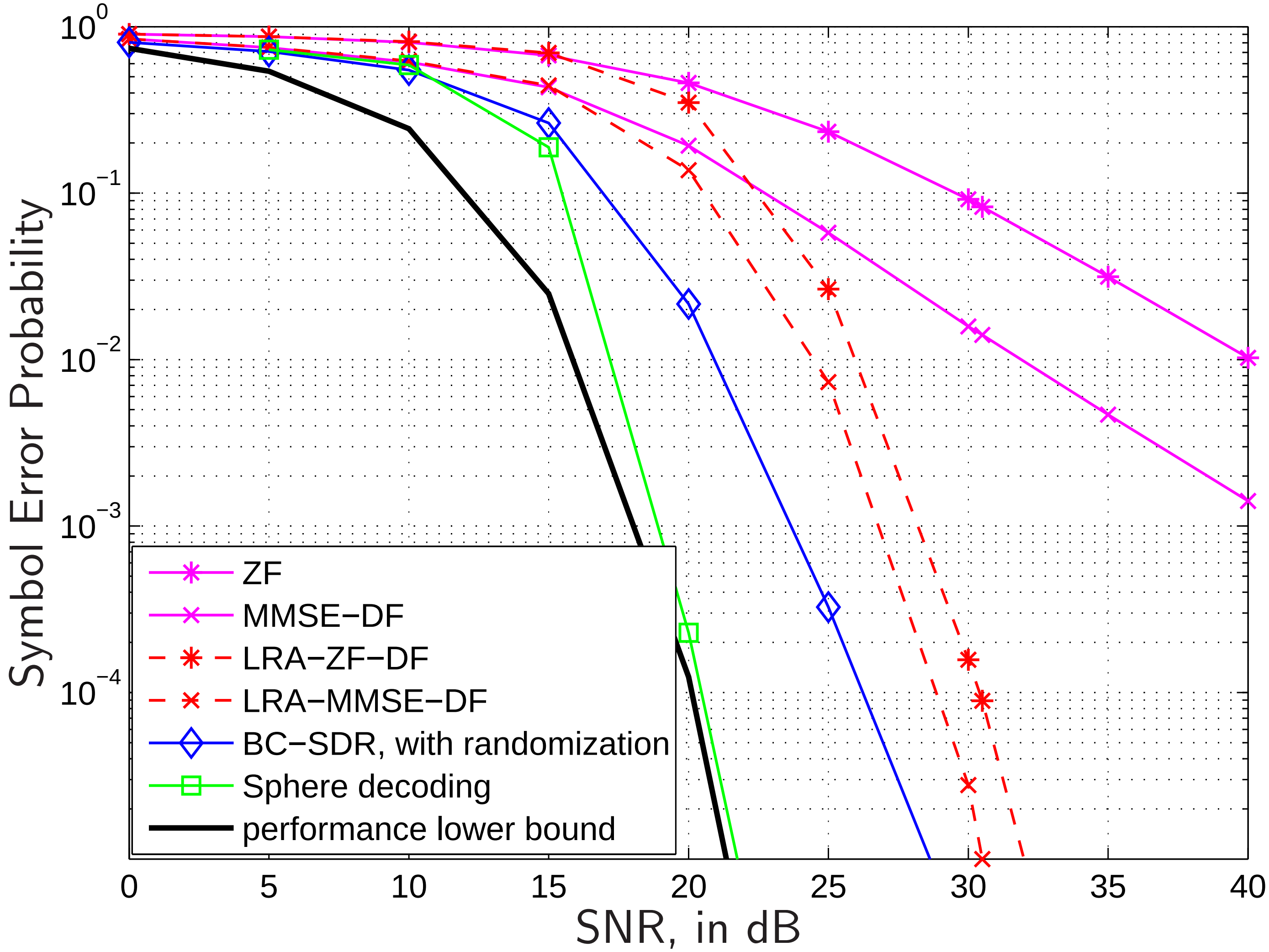
  The same equivalence is also proven for $64$-QAM PI-SDR, & for any $2^q$-QAM VA-SDR.



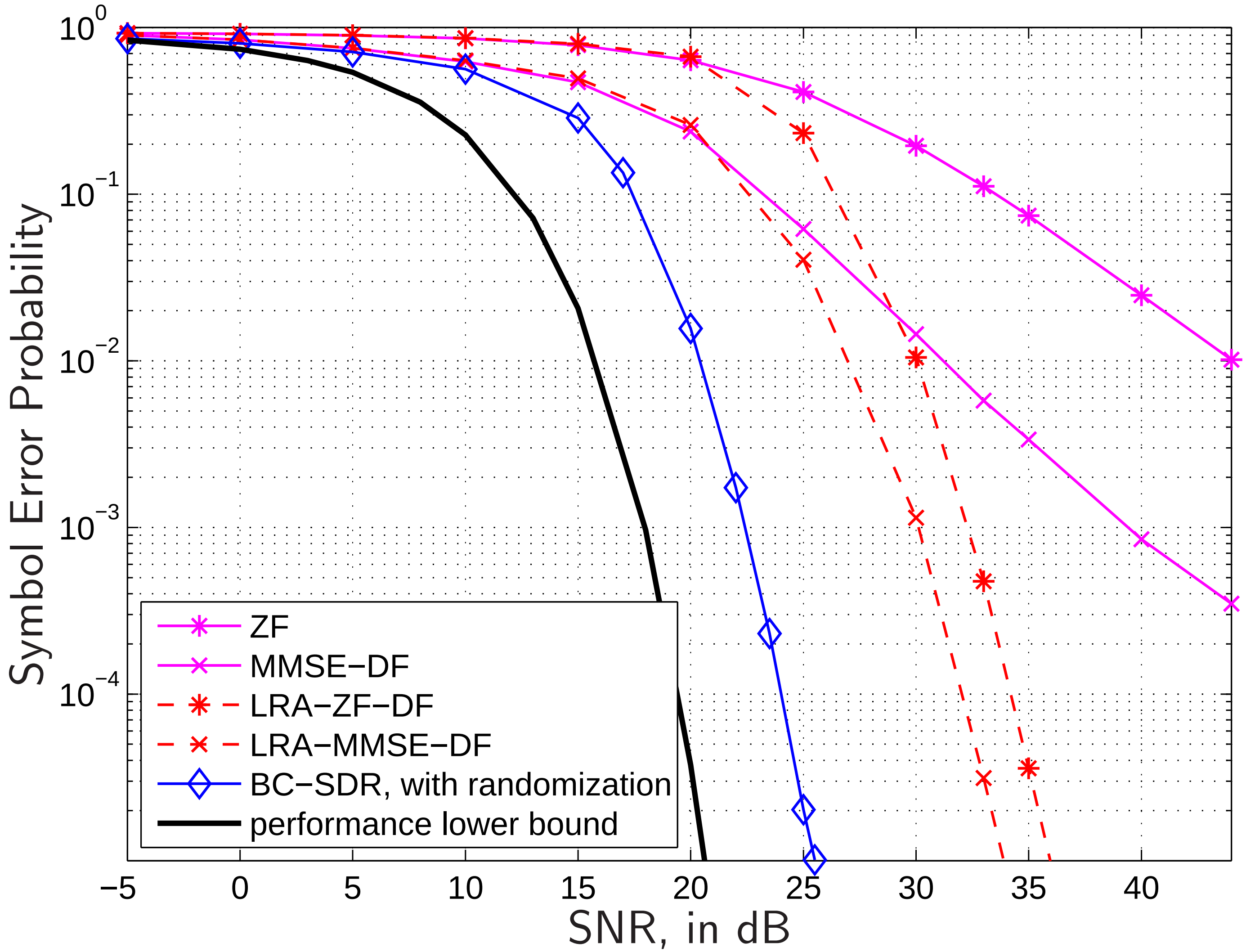

Symbol error rate performance of BC-SDR, PI-SDR, and VA-SDR under $(M_C, N_C) = (8, 8)$, $16$-QAM constellations. The three performance plots coincide.




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.

# 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 J. Select. Areas Commun.*, 2001.

**[Ma-Davidson-Wong-Luo-Ching'02]** W.-K. Ma, T.N. Davidson, K.M. Wong, Z.-Q. Luo, & P.C. Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous 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 Process.*, 2006.

**[Cheung-Ma-So'04]** K. W. Cheung, W.-K. Ma, & H. C. So, "Accurate approximation algorithm for TOA-based maximum likelihood mobile location using semidefinite programming," in *Proc. ICASSP 2004*, 2004.

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

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

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

**[Low'14]** S. H. Low, "Convex relaxation of optimal power flow: Parts I-II," *IEEE Trans. Control of Network Systems*, 2014.

**[Bienstock'14]** D. Bienstock, "progress on solving power flow problems," *OPTIMA*, issue 93, Dec. 2013.

**[Candès-Eldar-Strohmer-Voroninski'13]** E. J. Candes, Y. C. Eldar, T. Strohmer, & V. Voroninski, "Phase retrieval via matrix completion," *SIAM J. Imaging Sci.*, 2013.

**[Li-Bai-Ding'03]** Q. Li, E.-W. Bai & Z. Ding, "Blind source separation of signals with known alphabets using $\epsilon$-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, & K. M. Wong, "Soft quasi-maximum likelihood detection for multiple-antenna wireless channels," *IEEE Trans. Signal Process.*, Nov. 2003.

**[Kisialiou-Luo-Luo'09]** M. Kisialiou, X. Luo, & 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, & A. M.-C. So, "Cheap semidefinite relaxation MIMO detection using row-by-row block coordinate descent," in *Proc. ICASSP 2011*, 2011.

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

**[Kisialiou-Luo'10]** M. Kisialiou & 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.

**[Kim-Love-Park'11]** I. H. Kim, D. J. Love, & S. Y. Park, "Optimal and successive approaches to signal design for multiple antenna physical layer multicasting," *IEEE Trans. Commun.*, 2011.

**[Gershman-Sidiropoulos-Shahbazpanahi-Bengtsson-Ottersten'10]** A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, & B. Ottersten, "Convex optimization-based beamforming: From receive to transmit and network designs," *IEEE Signal Process. Mag.*, May 2010.

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

**[Chang-Luo-Wu *et al.*'08]** T.-H. Chang, Z.-Q. Luo, X. Wu, C. Akgun, J. Vaughan, K. Ugurbil, & 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.

**[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 Process.*, 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 Process.*, 2009.

**[Banavar-Smith-Tepedelenlioğlu-Spanias'12]** M. K. Banavar, A. D. Smith, C. Tepedelenlioğlu, and A. Spanias, "On the effectiveness of multiple antennas in distributed detection over fading MACs," *IEEE Trans. Wireless Commun.*, May 2012.

**[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. Signal Process.*, 2010.

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

*Signal Process.*, Mar. 2008.

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

**[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.

**[Ho-Jorswieck'12]** Z. Ho and E. Jorswieck, "Instantaneous relaying: Optimal strategies and interference neutralization," *IEEE Trans. Signal Process.*, 2012.

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

**[Shen-Chang-Wang-Qiu-Chi'12]** "Distributed robust multi-cell coordinated beamforming with imperfect CSI: An ADMM approach," *IEEE Trans. Signal Process.*, June 2012.

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

**[Xu-Liu-Zhang'13]** J. Xu, L. Liu and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power Transfer," in *Proc. ICASSP 2013*, 2013.

**[Chalise-Ma-Zhang-Suraweera-Amin'13]** B. K. Chalise, W.-K. Ma, Y. D Zhang, H. A Suraweera, & M. G Amin, "Optimum performance boundaries of OSTBC based AF-MIMO relay system with energy harvesting receiver," *IEEE Trans. Signal Process.*, Sep. 2013.

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

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

**[Zhang-Liang-Cui'10]** R. Zhang, Y. C. Liang, & S. Cui, "Dynamic resource allocation in cognitive radio networks," *IEEE Signal Process. Mag.*, May 2010.

**[Swindlehurst'09]** A. L. Swindlehurst, "Fixed SINR solutions for the MIMO wiretap channel," in in *Proc. ICASSP 2009*, 2009.

**[Matskani-Sidiropoulos-Luo-Tassiulas'08]** E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, & L. Tassiulas, "Convex approximation techniques for joint multiuser downlink beamforming and admission control," *IEEE Trans. Wireless Commun.*, 2008.

**[Wai-Ma'12]** H.-T. Wai & W.-K. Ma, "A decentralized method for joint admission control and beamforming in coordinated multicell downlink," in *Proc. Asilomar 2012*, 2012.

**[Wai-Li-Ma'13]** H.-T. Wai, Q. Li & W.-K. Ma, "A convex approximation method for multiuser MISO sum rate maximization under discrete rate constraints," in *Proc. ICASSP 2013*, 2013.

**[Cheng-Philipp-Pesavento'12]** Y. Cheng, A. Philipp, & M. Pesavento, "Dynamic rate adaptation and multiuser downlink beamforming using mixed integer conic programming," in *Proc. EUSIPCO 2012*, 2012.

**[Shi-Razaviyayn-Luo-He'11]** Q. Shi, M. Razaviyayn, Z.-Q. Luo, & C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, Sep. 2011.

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

**[Jöngren-Skoglund-Ottersten'02]** G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, 2002.

**[Pascual-Iserte-Palomar-*et al.*'06]** A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, 2006.

**[Wu-Ma-So'13]** X. Wu, W.-K. Ma, & A. M.-C. So, "Physical-layer multicasting by stochastic beamforming and Alamouti space-time coding," *IEEE Trans. Signal Process.*, Sep. 2013.

**[Wu-So-Ma'12]** X. Wu, A. M.-C. So, & W.-K. Ma, "Rank-two transmit beamformed Alamouti space-time coding for physical-layer multicasting," in *Proc. ICASSP 2012*, 2012.

**[Wen-Law-Alabed-Pesavento'12]** X. Wen, K. L. Law, Samer J. Alabed, and M. Pesavento, "Rank-two beamforming for single-group multicasting networks using OSTBC," in *Proc. SAM 2012*, 2012.

**[Schad-Law-Pesavento'12]** A. Schad, K. Law, and M. Pesavento, "A convex inner approximation technique for rank-two beamforming in multicasting relay networks," in *Proc. EUSIPCO 2012*, 2012.

**[Ji-Wu-So-Ma'13]** S. Ji, X. Wu, A. M.-C. So, and W.-K. Ma, "multi-group multicast beamforming in cognitive radio networks via rank-two transmit beamformed Alamouti space-time coding," in *Proc. ICASSP 2013*, 2013.

**[Law-Wen-Pesavento'13]** K. L. Law, X. Wen, & M. Pesavento, "General-rank transmit beamforming for multi-group nulticasting networks using OSTBC," in *Proc. SPAWC 2013*, 2013.

**[Chang-Luo-Chi'08]** T.-H. Chang, Z.-Q. Luo, and C.-Y. Chi, "Approximation bounds for semidefinite relaxation of max-min-fair multicast transmit beamforming problem," *IEEE Trans. Signal Process.*, Aug. 2008.

**[Liang-Xia'03]** X.-B. Liang and X.-G. Xia, "On the nonexistence of rate-one generalized complex orthogonal designs," *IEEE Trans. Inf. Theory*, 2003.

**[Zheng-Wang-Ng'08]** G. Zheng, K. K. Wong, and T. S. Ng, "Robust MIMO in the downlink: A worse-case optimization with ellipsoidal uncertainty regions," *EURASIP Journal on Advances in Signal Processing,* 2008.

**[Shenouda-Davidson'07]** M. B. Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE J. Sel. Topics in Signal Process.* Dec. 2007.

**[Vučić-Boche'09]** N. Vučić and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems," *IEEE Trans. Signal Process.*, Feb. 2009.

**[Song-Shi-Sanjabi-Sun-Luo'12]** E. Song, Q. Shi, M. Sanjabi, R. Sun, and Z.-Q. Luo, "Robust SINR- constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?" *EURASIP Journal on Wireless Communications and Networking,* 2012.

**[Chang-Ma-Chi'11]** T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Worst-case robust multiuser transmit beamforming using semidefinite relaxation: Duality and implications," in *Proc. Asilomar 2011*, 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-Chi'10]** K.-Y. Wang, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "A semidenite relaxation based conservative approach to robust transmit beamforming with probabilistic SINR constraints," in *Proc. 18th EUSIPCO*, Aug. 2010.

**[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.

**[Wang-So-Chang-Ma-Chi'14]** K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," submitted to *IEEE Trans. Signal Process.*, 2014.

**[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 Conf.*, 2008.

**[Saxe'79]** J. B. Saxe, "Embeddability of weighted graphs in $k$-space is strongly NP-hard," in *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," in *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.

**[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 Process.*, 2009.

**[Fu-Chan-Ma-So'12]** X. Fu, W.-K. Ma, F. W. K. Chan and H. C. So, "Complex-valued semidefinite relaxation approach for two-dimensional source localization using distance measurements and imperfect receiver positions," in *Proc. Int. Conf. Signal Process. 2012*, 2012.

**[Oğuz-Ekim-Gomes-Xavier-Oliverira'10]** P. Oğuz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, "A convex relaxation for approximate maximum-likelihood 2D source localization from range measurements," in *Proc. ICASSP 2010*, 2010.

**[Ge-Jiang-Ye'11]** D. Ge, X. Jiang, and Y. Ye, "A note on the complexity of $L_p$ minimization," *Math. Prog.*, 2011.

**[Ji-Sze-Zhou-So-Ye'13]** S. Ji, K.-F. Sze, Z. Zhou, A. M.-C. So, and Y. Ye, "Beyond convex relaxation: A polynomial-time non-convex optimization approach to network localization", in *Proc. 32nd IEEE INFOCOM*, 2013.

**[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. Mag.*, 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.

**[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.

**[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.

**[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. ICASSP 2008*, 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. ICASSP 2003*, 2003.

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