Convex Optimization Techniques for Signal Processing and Communication

Zhi-Quan (Tom) Luo

Department of Electrical and Computer Engineering
University of Minnesota
Minneapolis, MN 55455

luozq@ece.umn.edu

Content

1. Overview
2. Basic concepts, theory and algorithms
3. Applications
   - Digital signal processing
   - communication system design
4. Softwares
5. Bibliography

Part I: Overview

The Role of Optimization in Engineering

- A link/bridge between “science” and “engineering”, with many applications
  - signal processing, digital communication, fibre optic communication
  - financial engineering
  - control systems, mechanical engineering
  - transportation
- What fraction of “real” problems are convex?
  - By no means all
  - Many go unrecognized
  - Convex optimization plays important roles in nonconvex optimization
- Exploiting convexity in engineering context

What is Convex Optimization?

- Convex sets:

$$S = \{(x, y) | x^2 + y^2 \leq 1\}, ...$$

- Convex functions:

$$\begin{aligned} f(x) &= |x|,\ e^x,\ x^2,\ -\ln x, ... \\ f(x, y) &= x^2 + y^2,\ -x + y^4, ... \end{aligned}$$

- Convex optimization problems:

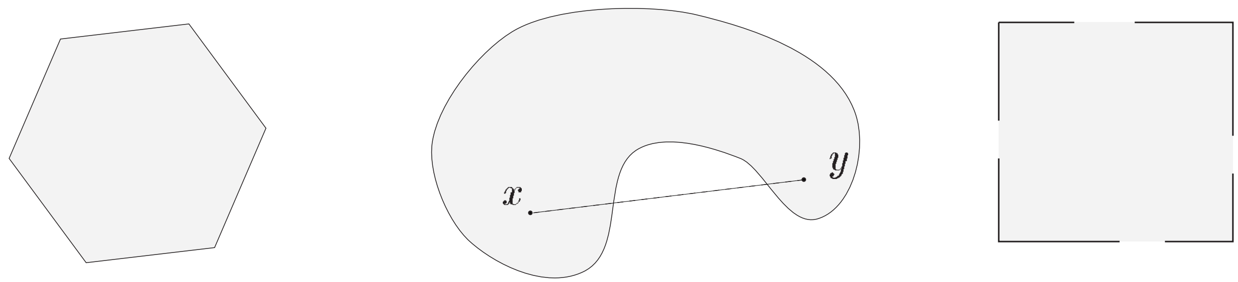
$$\boxed{\text{minimize}\quad f_0(x), \quad \text{subject to } x \in X, \quad f_i(x) \leq 0, \quad i = 1, 2, ..., m,}$$

where X is a closed convex set, each f_i is convex function ($0 \leq i \leq m$).

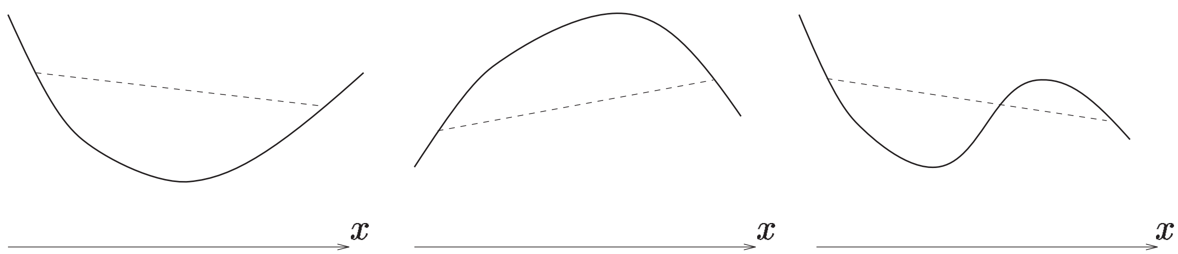
Examples of Convex Sets/Functions

Sets:

convex not convex



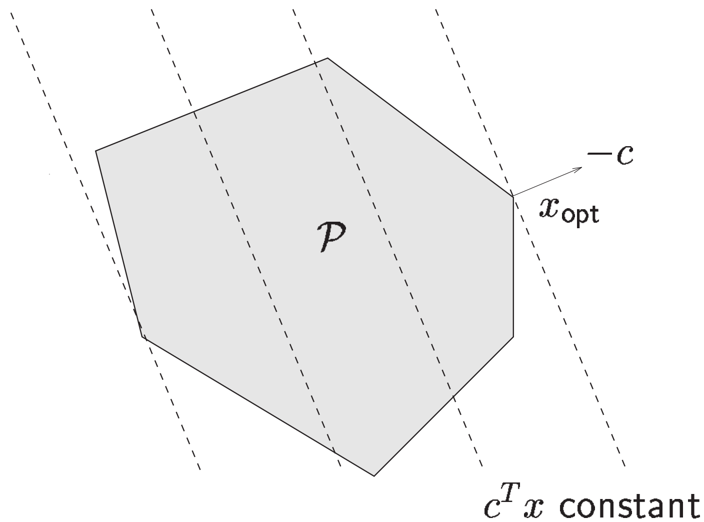
and functions:



convex concave neither

Examples of Convex Optimization Problems

$$\boxed{\text{minimize} \quad f_0(x), \quad \text{subject to} \quad f_i(x) \le 0, \quad i = 1, 2, \ldots, m.}$$



- each f_i is a linear function, $\Rightarrow$ linear programming (LP)
- f_0 is quadratic and f_i $(i \ge 1)$ is linear, $\Rightarrow$ quadratic programming (QP)
- all f_i are convex quadratic (QCQP)

A Well Known Example: Minimum Least Squares

$$\boxed{\text{minimize} \quad \|Ax - b\|^2, \quad \text{subject to} \quad Cx = d}$$

- Well known and widely (wildly?) used in engineering and statistics
- The objective function $f_0(x) = \|Ax - b\|^2$ is convex; the constraint set $Cx = d$ is an affine space, $\Rightarrow$ convex.
- Much of the current practice in engineering design is based on Least Squares.
- It is time to use some new and more powerful optimization tools and models.
- This allows efficient solution of some previously considered intractable engineering problems.

Variations:

- Note that we can take extra bound constraints $\ell_i \leq x_i \leq u_i$. Easy!
- However, constraints of the form $|x_i - a_i| \geq 1$ are difficult to handle (nonconvex). Difficult

Easy problems can appear very similar to difficult ones!

Semidefinite Programming (SDP)

- SDP is a broad class of conic convex optimization problems:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} c_i x_i \\ \text{subject to} & A_0 + x_1 A_1 + x_2 A_2 + \cdots + x_n A_n \succeq 0, \end{array}$$

  where each A_i is a $m \times m$ symmetric matrix. The constraint is called linear matrix inequality constraint (LMI).

- The set of symmetric positive semidefinite matrices forms a convex cone. LMI is a nonlinear constraint on x.

- SDP includes as special cases: LP, convex QP, QCQP, etc.

- Very efficient algorithms and softwares: primal/dual interior point algorithms and efficient matlab implementation (e.g., **SeDuMi**).

An Example of SDP

Covariance Matrix Approximation: For a zero mean stationary stochastic process, the covariance matrix is positive semidefinite and Toeplitz. However, the so called finite sample covariance matrix

$$R_{xx}(n) = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T, \quad x_i \in \Re^m$$

is in general not Toeplitz.

Method 1: minimize$_r$ $\|R_{xx}(n) - \sum_{i=0}^{m-1} r_i E_i\|$, subject to $r \in \Re^m$, where E_i is a Toeplitz matrix whose i and $n-i$ diagonals are equal to 1 and equal to zero elsewhere. However, the resulting Toeplitz approximation matrix $\sum_{i=0}^{m-1} r_i E_i$ may not be positive semidefinite.

Method 2: minimize$_r$ $\|R_{xx}(n) - \sum_{i=0}^{m-1} r_i E_i\|$, subject to $\sum_{i=0}^{m-1} r_i E_i \geq 0$. This is a SDP.

What's the big deal about convex optimization?

Convex optimization formulation

- always achieves global minimum, no local traps
- certificate for infeasibility
- can be solved by polynomial time complexity algorithms (e.g., interior point algorithms)
- highly efficient software available
- the dividing line between "easy" and "difficult" problems (compare with solving linear equations)

$\Rightarrow$ Whenever possible, always go for convex formulation.

Part II: Basic Concepts, Theory & Algorithms

- Convex sets

- Convex functions

- Convex optimization problems: LP, QP, SOCP, SDP

- Interior Point Algorithms

Convex Set

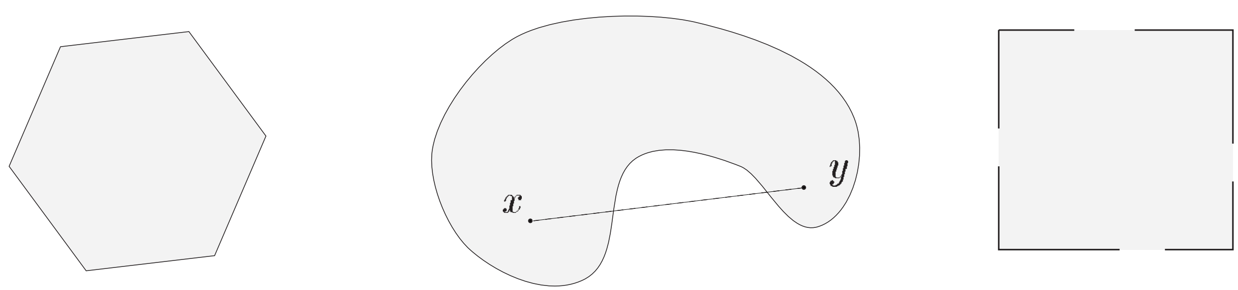
$S \subseteq \Re^n$ is *convex* if

$$x, y \in S,\ \ \lambda, \mu \geq 0,\ \ \lambda + \mu = 1 \ \Rightarrow \lambda x + \mu y \in S.$$

Geometrically, $x, y \in S$ implies the line segment $[x, y] \subseteq S$. S is called a *convex cone* if

$$x, y \in S,\ \ \lambda, \mu \geq 0,\ \Rightarrow \lambda x + \mu y \in S.$$

convex not convex



Examples of Convex Set

- Linear subspace: $S = \{x \mid Ax = 0\}$ is a convex cone
- Affine subspace: $S = \{x \mid Ax = b\}$ is a convex set
- Polyhedral set: $S = \{x \mid Ax \le b\}$ is a convex set
- PSD matrix cone: $S = \{A \mid A \text{ is symmetric, } A \ge 0\}$ is convex
- Second order cone: $S = \{(t, x) \mid t \ge \|x\|\}$ is convex

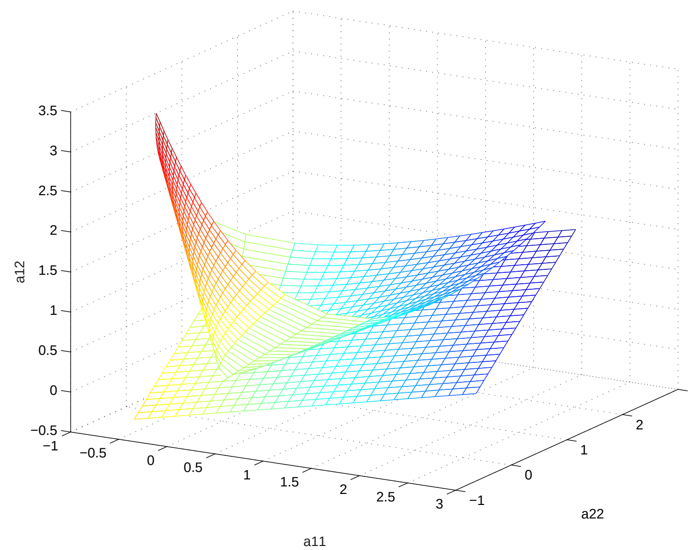
Important property:

$$\text{Intersection of} \begin{pmatrix} \text{linear subspaces} \\ \text{affine subspaces} \\ \text{convex cones} \\ \text{convex sets} \end{pmatrix} \text{ is also a } \begin{pmatrix} \text{linear subspace} \\ \text{affine subspace} \\ \text{convex cone} \\ \text{convex set} \end{pmatrix}$$

Recognizing Convex Sets

PSD matrix cone: $S = \{A \mid A$ is symmetric, positive semidefinite$\}$ is convex since $S = \left(\bigcap_{i \neq j} S_{ij}\right) \bigcap \left(\bigcap_{x \in \Re^n} S_x\right)$, where S_{ij} and S_x are linear subspaces and half spaces respectively

$$S_{ij} = \{A \mid A_{ij} = A_{ji}\}, \quad S_x = \{A \mid x^T A x \geq 0\}.$$

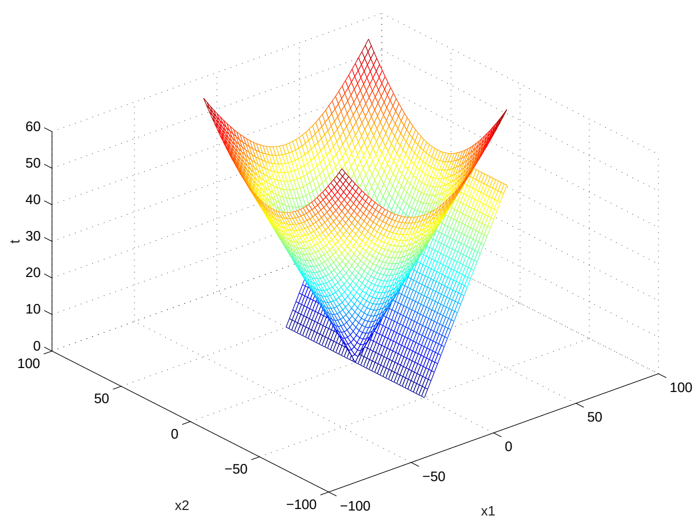


Recognizing Convex Sets

Second order cone: $S = \{(x,t) \mid \|x\| \le t\}$ is a convex cone since

$$S = \bigcap_{\|a\|=1} S_a, \quad S_a = \{(x,t) \mid a^T x \le t\}.$$

Here we note that $\|x\| \le t$ iff $a^T x \le t$ for all $\|a\| = 1$.

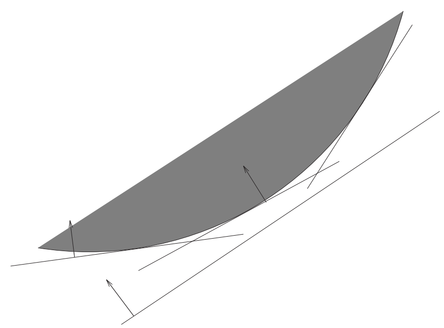


Robust Linear Constraint

- **Linear constraint:** $S = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} \geq b\}$ represents a half space.
- **Robust linear constraint:**

$$\bar{S} = \{\mathbf{x} \mid (\mathbf{a} + \Delta\mathbf{a})^T\mathbf{x} \geq b + \Delta b, \quad \forall \; \|(\Delta\mathbf{a}, \Delta b)\| \leq \epsilon\}$$

is seen as the intersection of infinitely many half spaces



- In fact, robust feasible region can be characterized as

$$\bar{S} = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} - b - \epsilon\sqrt{\|\mathbf{x}\|^2 + 1} \geq 0\}$$

MVDR Beamforming

$$\boxed{\text{mininmize } w^H R_{\mathrm{i+n}} w, \quad \text{subject to } w^H a(\theta_0) = 1}$$ $\Leftarrow$ Least squares!

- uniform linear array; steering vector: $a(\theta) = (1, e^{j\alpha\theta}, e^{j2\alpha\theta}, \ldots, e^{j(M-1)\alpha\theta})^H$
- can handle sidelope bound constraints $|w^H a(\theta)| \leq \epsilon,\ \forall\theta \in \Theta_s$.
- sidelope constraints can also be represented via an LMI (more later):

$$A_0 + w_1 A_1 + w_2 A_2 + \cdots + w_M A_M \geq 0$$

- robust MVDR (more later) via robust linear constraint (more later)

$$\boxed{\text{mininmize } w^H R_{\mathrm{i+n}} w, \quad \text{subject to } |w^H(a(\theta_0) + \Delta a)| \geq 1, \quad \forall \|\Delta a\| \leq \delta}$$

- However, constraints like "only half of w_i can be nonzero" are nonconvex.

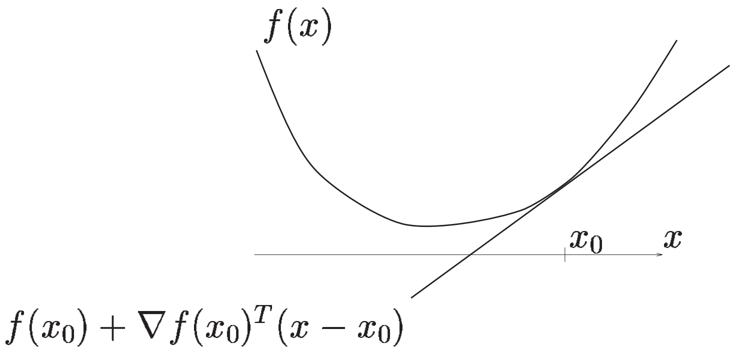
Convex Functions

$f(x) : \Re^n \longrightarrow \Re$ is **convex** if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall\, x, y \in \Re^n.$$

When $f(x)$ is differentiable, $f(x)$ is convex iff

$$f(x) \geq f(x_0) + \Delta f(x_0)^T (x - x_0), \text{ for all } x,\, x_0.$$



Recognizing Convex Functions

If $f(x)$ is twice continuously differentiable, then

$$f \text{ is convex} \quad \Leftrightarrow \quad \nabla^2 f(x) \geq 0 \text{ for all } x \in \Re^n.$$

Examples:

Non-convex	Non-differentiable convex	Differentiable convex
$1 - \|x\|$	$\|x\|$	$x^T A x, \; A \geq 0$
$\cos(e^T x)$	$\|x\|_1$	$e^{-\|x\|^2}$
$x^T A x, \; A \not\geq 0$	$\max\{\|x\|^2, e^T x\}$	$-\log(t^2 - \|x\|^2)$
$\|x\|^3$	$\|x\|_1^2$	$-\log\det(X)$
$\vdots$	$\vdots$	$\vdots$

Properties of Convex Functions

- **Convexity over all lines:**

$$f(x) \text{ is convex} \quad \Leftrightarrow \quad f(x_0 + th) \text{ is convex in } t \text{ for all } x_0 \text{ and } h$$

- **Positive multiple:**

$$f(x) \text{ is convex} \quad \Leftrightarrow \quad \alpha f(x) \text{ is convex, for all } \alpha \geq 0$$

- **Sum of convex functions**:

$$f_1(x),\ f_2(x) \text{ convex} \quad \Rightarrow \quad f_1(x) + f_2(x) \text{ is convex}$$

- **Pointwise maximum:**

$$f_1(x),\ f_2(x) \text{ convex} \quad \Rightarrow \quad \max\{f_1(x), f_2(x)\} \text{ is convex}$$

- **Affine transformation of domain:**

$$f(x) \text{ is convex} \quad \Rightarrow \quad f(Ax + b) \text{ is convex}$$

Some Commonly Used Convex Functions

- **Piecewise-linear functions:** $\max_i\{a_i^T x + b_i\}$ is convex in x
- **Quadratic functions:** $f(x) = x^T Qx + 2q^T x + c$ is convex iff $Q \geq 0$
- **Piecewise-quadratic functions:** $\max_i\{x^T Q_i x + q_i^T x + c_i\}$ is convex in x if $Q_i \geq 0$
- **Norm functions:** $\|x\|_k = \left(\sum_{i=1}^{n} |x_i|^k\right)^{1/k}$, where $k \in [1, \infty)$
- **Convex functions over matrices:** $\mathrm{Tr}(X)$, $\lambda_{\max}(X)$ are convex on $X = X^T$; and $-\log\det(X)$ is convex on the set $\{X \mid X = X^T,\ X \geq 0\}$
- **Logarithmic barrier functions:** $f(x) = \sum_{i=1}^{m} \log(b_i - a_i^T x)^{-1}$ is convex over $\mathcal{P} = \{x \mid a_i^T x \leq b_i,\ 1 \leq i \leq m\}$

Convex Optimization Problems

- **Abstract form:** problem

$$\text{minimize } f(x), \quad \text{subject to } \; x \in C$$

  is convex if C and f are **convex** (set, function)
  Note: **maximizing concave** f over convex C is a convex optimization problem

- **Standard form:** problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & g_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

  is convex if $f_0, \; f_1, \ldots, f_m$ are **convex**, $g_1, \ldots, g_p$ **affine** often written as

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0; \quad i = 1, \ldots, m, \\ & Ax = b. \end{array}$$

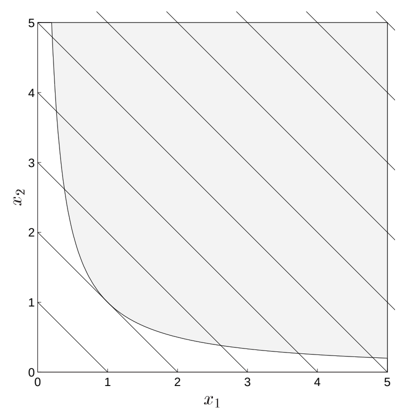
Recognizing Convex Optimization Problems

- Consider

$$\begin{array}{ll} \text{minimize} & x_1 + x_2 \\ \text{subject to} & -x_1 \le 0, \ -x_2 \le 0 \\ & 1 - x_1 x_2 \le 0 \end{array}$$

  It is convex in the abstract form since the feasible set is convex, but nonconvex in standard form.

- We can replace $1 - x_1 x_2 \le 0$ by a convex constraint $-\log x_1 - \log x_2 \le 0.$



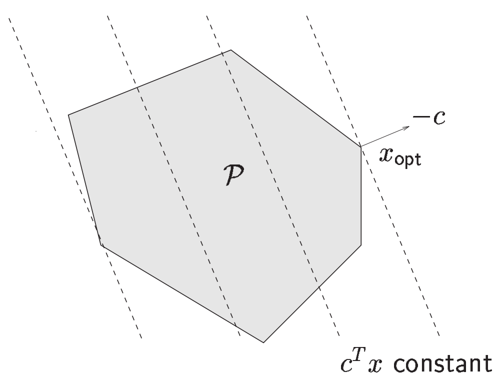
Recognizing Convex Optimization Problems

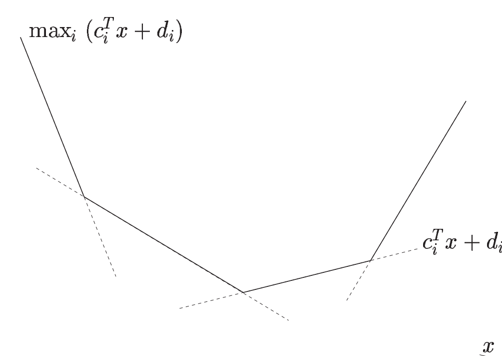
- **Linear program:** f_i all affine yields linear program

$$\begin{array}{ll} \text{minimize} & c_0^T x + d_0 \\ \text{subject to} & c_i^T x + d_i \leq 0, i = 1, \dots, m \\ & Ax = b \end{array}$$

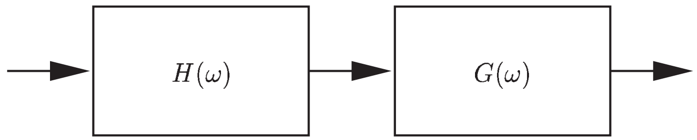
- **Constrained minimax problem:**

$$\begin{array}{ll} \text{minimize} & \max_i \{c_i^T x + d_i\} \\ \text{subject to} & Ax \leq b \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \text{minimize} & t \\ \text{subject to} & Ax \leq b, \; c_i^T x + d_i \leq t, \forall i \end{array}$$





Blind Channel Equalization via Linear Programming



- LTI channel; impulse response h: complex, unknown; QAM modulation
- **Goal:** design equalizer $G(\omega)$ such that $H(\omega)G(\omega) = e^{-jD\omega}$, where D is delay.
- Let x_i be the channel output, and y_k be the equalizer output. It is known that

$$\boxed{\begin{array}{ll} \text{minimize} & \max_k |\text{Re}(y_k)| \\ \text{subject to} & \text{Re}(h_1) + \text{Im}(h_1) = 1 \end{array}} \quad \Rightarrow \text{channel equalization.}$$

where $\text{Re}(h_1) + \text{Im}(h_1) = 1$ is a normalizing constraint.

- Note that $\text{Re}(y_k) = \sum_{i=-N}^{N} \text{Re}(x_{k-i})\text{Re}(h_i) - \text{Im}(x_{k-i})\text{Im}(h_i)$ is linear in h.

Blind Channel Equalization via Linear Programming

- Note that $\max_k\{|\text{Re}(y_k)|\}$ is **nonsmooth** in h.
- **Smoothing approach:** approximating the ∞-norm by p-norm, with p large,

$$\max_k\{|\text{Re}(y_k)|\} \approx \left(\sum_k |\text{Re}(y_k)|^p\right)^{1/p}$$

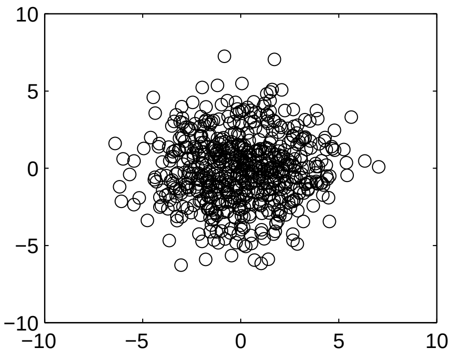
and then remove the exponent $\frac{1}{p}$ to obtain a smooth objective. $\Rightarrow$ Gradient descent.

- **Linear programming formulation:**

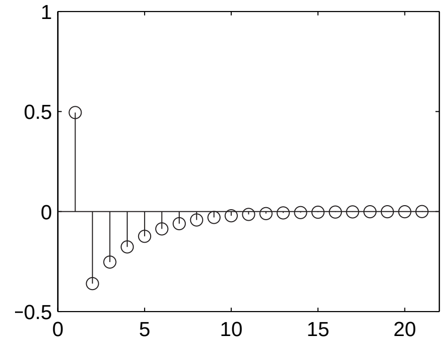
$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \text{Re}(h_1) + \text{Im}(h_1) = 1 \\ & -t \leq \displaystyle\sum_{i=-N}^{N} \text{Re}(x_{k-i})\text{Re}(h_i) - \text{Im}(x_{k-i})\text{Im}(h_i) \leq t, \quad \forall\ k \end{array}$$

- The above formulation requires **fewer samples**, and enjoys **faster convergence**.

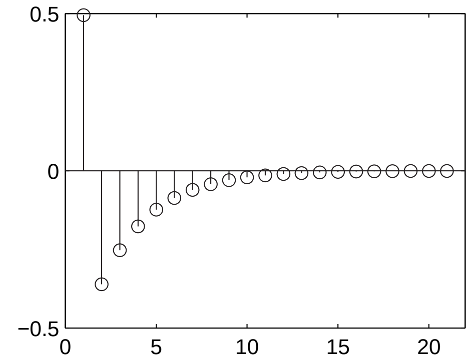
Blind Channel Equalization via Linear Programming



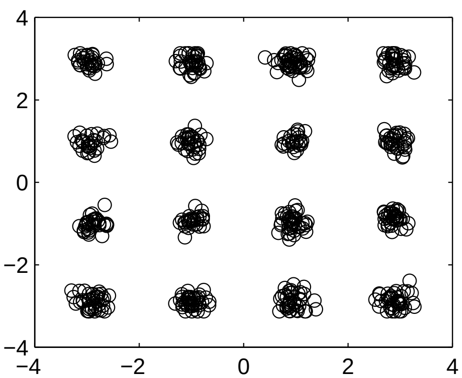
(a) Channel output



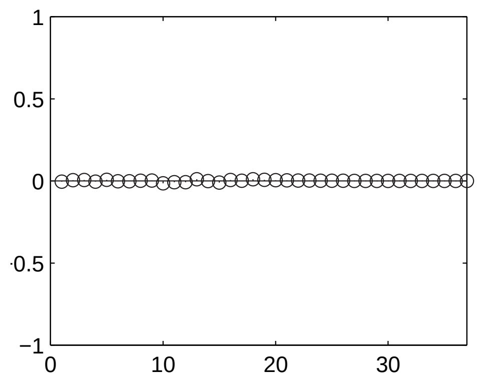
(b) Real part



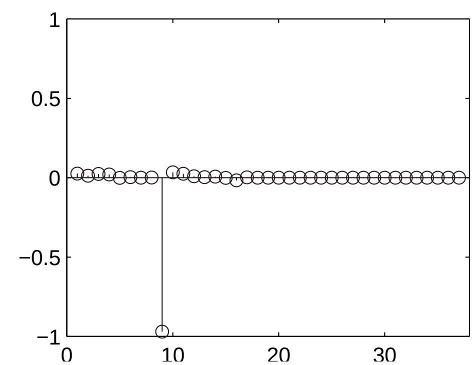
(c) Imaginary part



(d) Equalized output



(e) Real part



(f) Imaginary part

Quadratic Optimization Problems

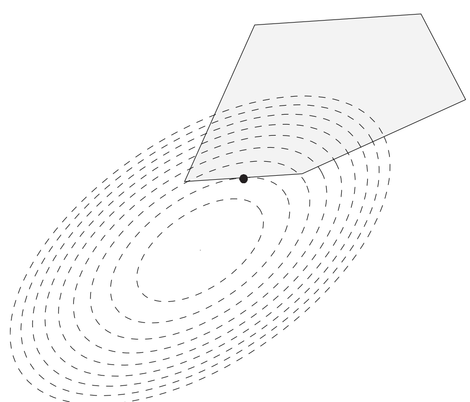
- **Convex quadratically constrained quadratic program (QCQP):** each $Q_i \succeq 0$

$$\begin{array}{ll} \text{minimize} & x^T Q_0 x + p_0^T x + c_0 \\ \text{subject to} & x^T Q_i x + p_i^T x + c_i \le 0, i = 1, \ldots, m \\ & Ax = b. \end{array}$$

  If $Q_i = 0$, $i \ge 1$, then it is a **convex QP**. If some $Q_i \not\succeq 0$, $\Rightarrow$ **nonconvex and hard**!

- **Optimal norm approximation with bound constraints:**

$$\begin{array}{ll} \text{minimize} & \|Ax - b\| \\ \text{subject to} & \ell_i \le x_i \le u_i, \; i = 1, \ldots, n. \end{array}$$



Convex Conic Optimization Problems

- **Second order cone program (SOCP):**

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|A_i x + b_i\| \le d_i^T x + e_i \end{array}$$

  includes LP, convex QP, and convex quadratically constrained QP as special cases.

- **Semi-definite program (SDP):**

$$\begin{array}{ll} \text{minimize} & \mathrm{Tr}(CX) \\ \text{subject to} & \mathrm{Tr}(A_i X) = b_i, \quad i = 1, .., m, \\ & X \succeq 0. \end{array}$$

  includes LP (if C, A_i are diagonal) and SOCP as special cases.

- **Model generality:** LP $<$ QP $<$ QCQP $<$ SOCP $<$ SDP.
- **Solution efficiency:** LP $>$ QP $>$ QCQP $>$ SOCP $>$ SDP. For example, it is much easier to exploit sparsity in LP than in SDP.

Robust Linear Programming via SOCP

$$\begin{array}{ll} \text{minimize} & \max_{\|\Delta c\| \le \epsilon} (c + \Delta c)^T x \\ \text{subject to} & (a_i + \Delta a_i)^T x \le b_i + \Delta b_i, \quad \forall \, \|(\Delta a_i, \Delta b_i)\| \le \epsilon_i \end{array}$$

Robust linear constraint is equivalent to a SOC constraint

$$(a + \Delta a)^T x \le b + \Delta b, \quad \forall \, \|(\Delta a, \Delta b)\| \le \epsilon \quad \Leftrightarrow \quad a^T x + b + \epsilon \sqrt{\|x\|^2 + 1} \le 0$$

So the robust linear program is equivalent to a SOCP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i + \epsilon_i \sqrt{\|x\|^2 + 1} \le 0 \\ & \epsilon \|x\| + c^T x \le t \end{array}$$

Schur Complement

Let $A > 0$ and

$$X = X^T = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, \quad S = C - B^T A^{-1} B.$$

Then S is called **Schur complement** of A in X, and

$$\boxed{X \geq 0 \quad \Leftrightarrow \quad S \geq 0.}$$

Schur complement is useful to represent nonlinear convex constraints as LMIs, e.g.,

$$\boxed{(Ax+b)^T(Ax+b) - c^T x - d \leq 0} \quad \Leftrightarrow \quad \begin{bmatrix} I & (Ax+b) \\ (Ax+b)^T & c^T x + d \end{bmatrix} \geq 0.$$

and

$$\boxed{x \geq 0, \ y \geq 0, \ 1 - xy \leq 0} \quad \Leftrightarrow \quad \begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix} \geq 0$$

Some Nonconvex Problems

'Slight' modification of convex problems can be very hard.

- Convex maximization, concave minimization, e.g.,

$$\begin{array}{ll} \text{maximize} & \|Ax - b\|^2 \\ \text{subject to} & \|x\| \leq 1. \end{array}$$

- Nonlinear equality constraints, e.g.,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x^T Q_i x + q_i^T x + c_i = 0, \;\; 1 \leq i \leq K, \end{array}$$

  where $Q_i \succeq 0$.
- Minimizing over nonconvex sets, e.g., integer constraints

$$\text{find } \; x \; \text{ such that } \; Ax \leq b, \;\; x_i \text{ is integer.}$$

Lagrangian Duality Theory

- **Primal problem:** Let p^* be the optimal value of convex optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_1(x) \le 0, \cdots, f_m(x) \le 0. \end{array}$$

- **(Lagrangian) dual function:**

$$g(\lambda) = \inf_x L(x, \lambda) = \inf_x \{f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)\}$$

  where $L(x, \lambda) = f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)$ is the **Lagrangian function** and λ_i's are Lagrangian multipliers

- **Weak duality:** if $\lambda \ge 0$ (**dual feasible**) and x is **primal feasible** then $f(x) \ge g(\lambda)$.
- **Dual problem:** Let d^* be the optimal value of

$$\begin{array}{ll} \text{maximize} & g(\lambda) \\ \text{subject to} & \lambda \ge 0. \end{array}$$

- **Duality gap:** $p^* \ge d^*$. For convex problems and under mild conditions, $p^* = d^*$.

KKT Optimality Condition

- Suppose each f_i is differentiable. Then (x^*, λ^*) is primal and dual optimal iff

$$\begin{aligned} f_i(x^*) &\leq 0 \\ \lambda^* &\geq 0 \\ \nabla f_0(x^*) + \textstyle\sum_i \lambda_i^* \nabla f_i(x^*) &= 0 \\ \lambda_i^* f_i(x^*) &= 0 \end{aligned}$$

  Here λ^* serves as a **certificate of optimality.**

- **Theorem of alternatives:**
  - there exists x such that $f_i(x) < 0$, for all i
  - there exists a $\lambda \geq 0$ and $\lambda \neq 0$ such that $g(\lambda) \geq 0$.

  Exactly one of the above is true. In practice, λ serves as a **certificate of infeasibility.**

Interior Point Algorithms for SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & A_0 + x_1 A_1 + \ldots + x_n A_n \geq 0 \end{array}$$

- **Logarithmic barrier function:**

$$\phi(x) = -\log\det(A_0 + x_1 A_1 + \ldots + x_n A_n)$$

- follow the **Central path:**

$$x^*(t) = \text{argmin}\left(tc^T x + \phi(x)\right), \quad t \in [0, \infty)$$

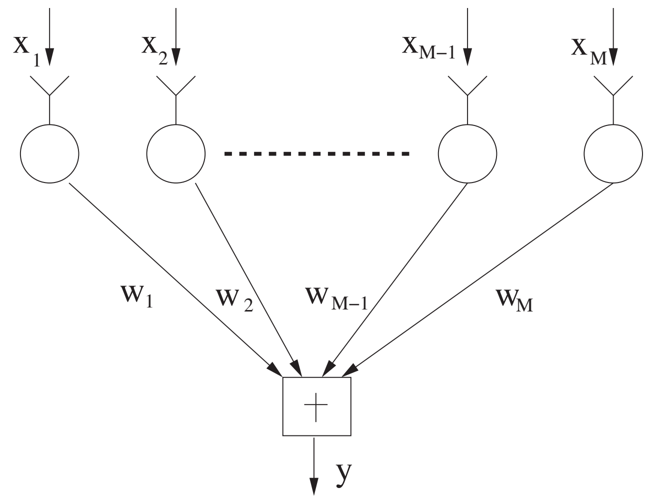
- for any $t \geq 0$, it is known $f_0(x^*(t)) \geq f^* \geq f_0(x^*(t)) - \frac{m}{t}$
- works very well in practice; theoretical worst case polynomial complexity
- other polynomial interior point algorithms: primal-dual, predictor-corrector, etc

Part III: Applications

1. Robust beamforming
2. Optimal linear decentralized estimation
3. Pulse shaping filter design
4. MMSE precoder design for multi-access communication
5. Quasi-maximum likelihood detection via SDP relaxation
6. Transmit beamforming

Robust beamforming

Robust Beamforming Application



- Widely used in wireless communications, microphone array speech processing, radar, sonar, medical imaging, radio astronomy.
- The output of a narrowband beamformer is given by

$$y(k) = \mathbf{w}^H \mathbf{x}(k), \quad \text{where } k \text{ is the sample index.} \tag{1}$$

Robust Beamforming

- The observation vector is given by

$$\mathbf{x}(k) = \mathbf{s}(k) + \mathbf{i}(k) + \mathbf{n}(k) = s(k)\mathbf{a} + \mathbf{i}(k) + \mathbf{n}(k) \tag{2}$$

  where $\mathbf{s}(k)$, $\mathbf{i}(k)$, and $\mathbf{n}(k)$ are the desired signal, interference, and noise components, respectively. Here, $s(k)$ is the signal waveform, and $\mathbf{a}$ is the signal steering vector.

- The robustness of a beamformer to a mismatch between the nominal (presumed) and real signal steering vectors becomes the main issue.

- Such mismatches can occur in practical situations as a consequence of look direction and signal pointing errors, imperfect array calibration and distorted antenna shape, array manifold mismodeling due to source wavefront distortions caused by environmental inhomogeneities, near-far problem, source spreading and local scattering.

Robust Beamforming

- The optimal weight vector can be obtained through the maximization of the Signal-to-Interference-plus-Noise Ratio (SINR)

$$\mathrm{SINR} = \frac{\sigma_{\mathrm{s}}^2 \, |\mathbf{w}^H \mathbf{a}|^2}{\mathbf{w}^H \mathbf{R}_{\mathrm{i+n}} \mathbf{w}} \tag{3}$$

where $\mathbf{R}_{\mathrm{i+n}} = \mathrm{E}\left\{ (\mathbf{i}(t) + \mathbf{n}(t)) \, (\mathbf{i}(t) + \mathbf{n}(t))^H \right\}$.

- The maximization of (3) is equivalent to

$$\min_{\mathbf{w}} \ \mathbf{w}^H \mathbf{R}_{\mathrm{i+n}} \mathbf{w} \qquad \text{subject to} \qquad \mathbf{w}^H \mathbf{a} = 1$$

- The optimal weight vector is $\mathbf{w}_{\mathrm{opt}} = \alpha \, \mathbf{R}_{\mathrm{i+n}}^{-1} \, \mathbf{a}$, where $\alpha = \left(\mathbf{a}^H \mathbf{R}_{\mathrm{i+n}}^{-1} \mathbf{a} \right)^{-1}$ is the normalization constant (to be omitted for brevity).
- In practice, $\mathbf{R}_{\mathrm{i+n}}$ is replaced by the sample convariance matrix.

Robust Beamforming

- In practical applications, the steering vector distortions $\mathbf{e}$ can be bounded:

$$\|\mathbf{e}\| \leq \epsilon.$$

- Then, the actual signal steering vector belongs to the set

$$\mathcal{A}(\epsilon) = \{\mathbf{c} \mid \mathbf{c} = \mathbf{a} + \mathbf{e}\,,\ \|\mathbf{e}\| \leq \epsilon\}$$

- We impose a constraint that for all vectors in $\mathcal{A}(\epsilon)$, the array response should not be smaller than one, i.e.

$$|\mathbf{w}^H\mathbf{c}| \geq 1 \qquad \text{for all} \qquad \mathbf{c} \in \mathcal{A}(\epsilon)$$

Formulation

- The robust formulation of adaptive beamformer is

$$\boxed{\begin{array}{lll} \text{minimize}_{\mathbf{w}} & \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} & \\ \text{subject to} & |\mathbf{w}^H(\mathbf{a}+\mathbf{e})| \geq 1, & \text{for all} \quad \|\mathbf{e}\| \leq \epsilon. \end{array}}$$

- For each choice of $\mathbf{e}$, the condition $|\mathbf{w}^H(\mathbf{a}+\mathbf{e})| \geq 1$ represents a nonlinear and nonconvex constraint on $\mathbf{w}$.
- Since there are an infinite number of vectors $\mathbf{e}$ with $\|\mathbf{e}\| \leq \epsilon$, the robust beamforming problem is a semi-infinite nonconvex quadratic program.
- It is well known that the general nonconvex quadratically constrained quadratic programming problem is NP-hard and thus intractable.

Convex Reformulation

- However, due to the special structure of the objective function and the constraints, the robust beamforming problem can be reformulated, surprisingly, as a convex second order cone program:

$$\begin{array}{ll} \text{minimize}_{\mathbf{w}} & \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} \\ \text{subject to} & \mathbf{w}^H \mathbf{a} \geq \epsilon \|\mathbf{w}\| + 1 \,, \ \mathrm{Im}\left\{ \mathbf{w}^H \mathbf{a} \right\} = 0. \end{array}$$

[Vorobyov-Gershman-Luo, 2001]

- The reformulation is based on simple $\triangle$- and Cauchy-Schwartz inequalities and the homogeneous nature of the objective function.
- The above SOCP can be efficiently and easily solved via interior point method.

Mathematical Reformulation

$$\boxed{|\mathbf{w}^H(\mathbf{a}+\mathbf{e})| \geq 1, \qquad \text{for all} \quad \|\mathbf{e}\| \leq \epsilon}$$

$\Updownarrow$ ($\triangle$-inequality)

$$\boxed{|\mathbf{w}^H\mathbf{a}| - |\mathbf{w}^H\mathbf{e}| \geq 1, \qquad \text{for all} \quad \|\mathbf{e}\| \leq \epsilon}$$

$\Updownarrow$ (Cauchy-Schwartz)

$$\boxed{|\mathbf{w}^H\mathbf{a}| - \epsilon\|\mathbf{w}\| \geq 1}$$

Therefore, the original robust beamforming formulation is equivalent to

$$\boxed{\begin{array}{ll} \text{minimize}_{\mathbf{w}} & \mathbf{w}^H\hat{\mathbf{R}}\mathbf{w} \\ \text{subject to} & |\mathbf{w}^H\mathbf{a}| \geq \epsilon\|\mathbf{w}\| + 1. \end{array}}$$

$\Leftarrow$ still non-convex!

Mathematical Reformulation

$$\boxed{\begin{array}{ll} \text{minimize}_{\mathbf{w}} & \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} \\ \text{subject to} & |\mathbf{w}^H \mathbf{a}| \geq \epsilon \|\mathbf{w}\| + 1. \end{array}}$$

$\Updownarrow$ Phase rotation

$$\boxed{\begin{array}{ll} \text{minimize}_{\mathbf{w}} & \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} \\ \text{subject to} & \mathbf{w}^H \mathbf{a} \geq \epsilon \|\mathbf{w}\| + 1\,, \ \mathrm{Im}\left\{\mathbf{w}^H \mathbf{a}\right\} = 0. \end{array}}$$

Simulation Examples

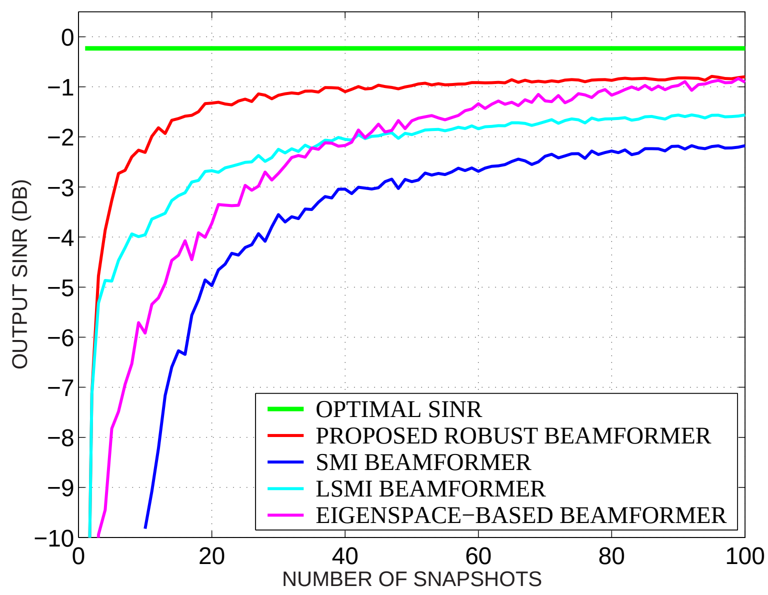
Example 1: Steering vector mismatch due to local scattering

- In this example, the presumed signal steering vector is a plane wave impinging on the array from 3°.
- The real steering vector is formed by five signal paths and is given by

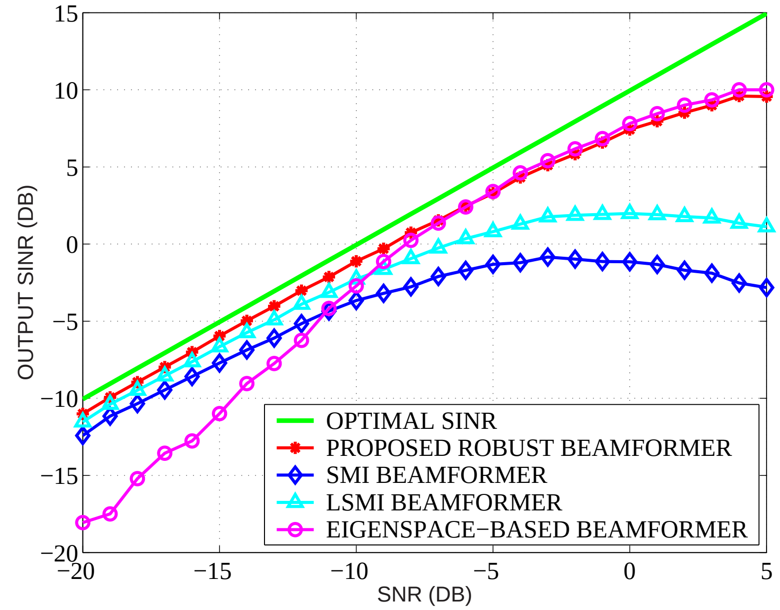
$$\tilde{\boldsymbol{a}} = \boldsymbol{a} + \sum_{i=1}^{4} e^{j\psi_i} \boldsymbol{b}(\theta_i) \tag{4}$$

  where $\boldsymbol{a}$ corresponds to the direct path, whereas $\boldsymbol{b}(\theta_i)$ $(i = 1, 2, 3, 4)$ correspond to the coherently scattered paths, with θ_i, $i = 1, 2, 3, 4$ independently drawn.
- The phases ψ_i, $i = 1, 2, 3, 4$ are independently and uniformly drawn from the interval $[0, 2\pi]$.

Performance Comparison



(a) output SINR v.s. sample size N; 1st example



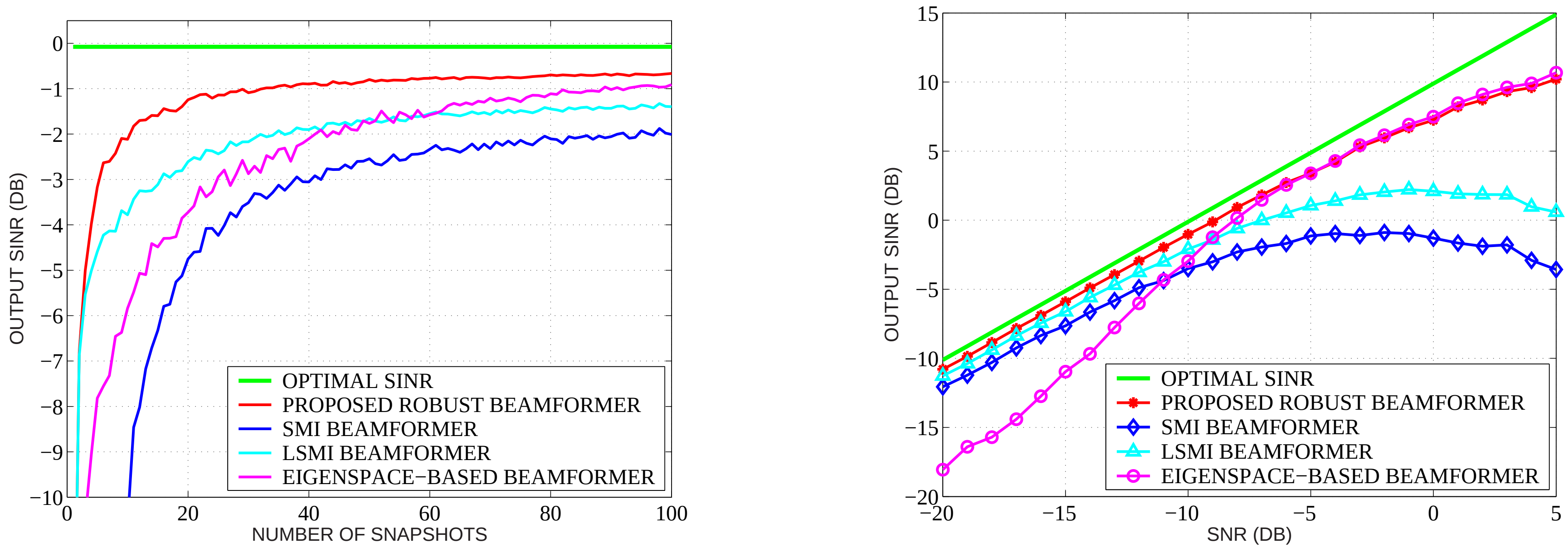
(b) output SINR versus SNR; 1st example

Simulation Examples

Example 2: Near-far steering vector mismatch

- In this example, we model the so-called near-far steering vector mismatch of the desired signal, whereby the presume steering vector of the signal is a plane wave impinging on the array from the normal direction 0°, whereas the real steering vector corresponds to the source located in the near field of the antenna at the distance $D^2/\lambda = (M-1)^2\lambda/4$ from the geometrical center of the array, where $D = (M-1)\lambda/2$ is the length of array aperture.
- The performance of the methods tested versus the number of training snapshots N for the fixed $\mathrm{SNR} = -10$ dB is shown in Fig. (c). Fig. (d) shows the performance of these techniques versus SNR for the fixed training data size $N = 30$.

Performance Comparison



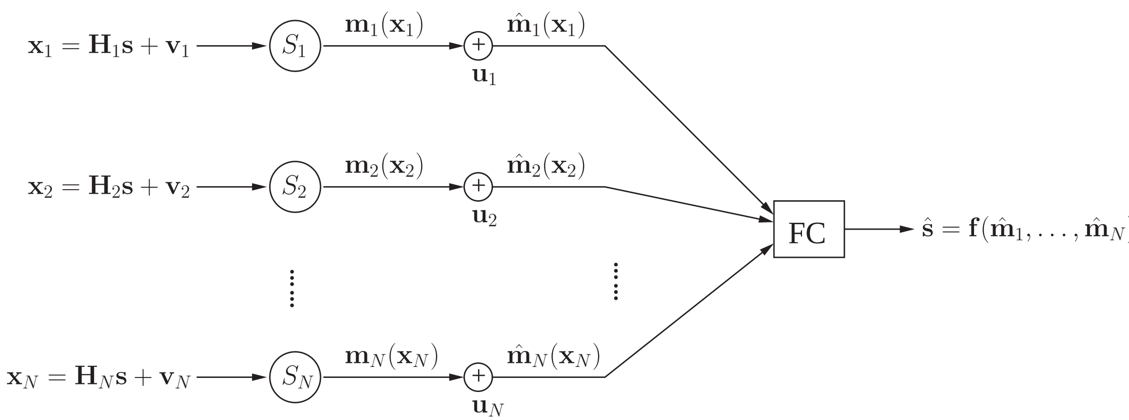
(a) output SINR v.s. sample size N; 2nd example

(b) output SINR versus SNR; 2nd example

Optimal linear decentralized estimation scheme

A Generic Decentralized Estimation Problem

- Our goal is to design a DES that minimizes the mean squared error (MSE) $\mathrm{E}(\mathrm{s} - \hat{\mathrm{s}})^2$.
- **Linear message functions:** $\mathbf{m}_n(\mathbf{x}_n) = \mathbf{Q}_n^T \mathbf{x}_n$, where $\mathbf{Q}_n$ is a **tall** matrix with fixed dimension.
- **Orthogonal AWGN channels:** $\hat{\mathbf{m}}_n(\mathbf{x}_n) = \mathbf{m}_n(\mathbf{x}_n) + \mathbf{u}_n = \mathbf{Q}_n^T \mathbf{x}_n + \mathbf{u}_n$,where $\mathbf{u}_n$ is additive Gaussian noise with a known variance $\mathbf{T}_n$.
- **Linear fusion function:** $\mathbf{f}\left(\hat{\mathbf{m}}_1(\mathbf{x}_1),\ \hat{\mathbf{m}}_2(\mathbf{x}_2), \ldots, \hat{\mathbf{m}}_N(\mathbf{x}_N)\right) = \mathbf{P}_1 \hat{\mathbf{m}}_1 + \cdots + \mathbf{P}_N \hat{\mathbf{m}}_N$, where $\mathbf{P}_n$ are matrices to be designed.



Single Sensor Case

- The mean squared error at the fusion center can be easily calculated as follows:

$$\begin{aligned}
\mathrm{E}(\mathbf{s}-\hat{\mathbf{s}})^2 &= \mathrm{E}\left(\mathbf{s}-\mathbf{P}\left(\mathbf{Q}^T\mathbf{x}+\mathbf{u}\right)\right)^2\\
&= \mathrm{E}\left(\mathbf{s}-\mathbf{P}\left(\mathbf{Q}^T(\mathbf{H}\mathbf{s}+\mathbf{v})+\mathbf{u}\right)\right)^2\\
&= \mathrm{E}\left(\left(\mathbf{I}-\mathbf{P}\mathbf{Q}^T\mathbf{H}\right)\mathbf{s}-\mathbf{P}\left(\mathbf{Q}^T\mathbf{v}+\mathbf{u}\right)\right)^2\\
&= \left\|\mathbf{I}-\mathbf{P}\mathbf{Q}^T\mathbf{H}\right\|_F^2+\mathrm{Tr}\left(\mathbf{P}\left(\mathbf{Q}^T\mathbf{Q}+\mathbf{T}\right)\mathbf{P}^T\right),
\end{aligned}$$

- This leads to the following formulation:

$$\begin{array}{ll}
\text{minimize}_{\mathbf{P},\mathbf{Q}} & \left\|\mathbf{I}-\mathbf{P}\mathbf{Q}^T\mathbf{H}\right\|_F^2+\mathrm{Tr}\left(\mathbf{P}\left(\mathbf{Q}^T\mathbf{Q}+\mathbf{T}\right)\mathbf{P}^T\right)\\
\text{subject to} & \mathrm{Tr}\left(\mathbf{Q}^T\mathbf{Q}\right)\le p,\quad \mathbf{P}\in\mathbb{R}^{\ell\times k},\ \mathbf{Q}\in\mathbb{R}^{\ell\times k},
\end{array}$$

where $p > 0$ is the power budget.

- As a fourth order polynomial, the objective function is nonconvex.

Single Sensor Case

- The objective function

$$g(\mathbf{Q},\mathbf{P}) = \left\|\mathbf{I} - \mathbf{P}\mathbf{Q}^T\mathbf{H}\right\|_F^2 + \mathrm{Tr}\left(\mathbf{P}\left(\mathbf{Q}^T\mathbf{Q} + \mathbf{T}\right)\mathbf{P}^T\right)$$

  is a fourth order nonconvex polynomial in $\mathbf{Q}$, $\mathbf{P}$.

- $g(\mathbf{Q},\mathbf{P})$ is convex in $\mathbf{P}$, and $\mathbf{P}$ is unconstrained $\Rightarrow$ $g(\mathbf{Q},\mathbf{P})$ can be minimized first with respect to $\mathbf{P}$ to obtain $\mathbf{P} = \mathbf{H}^T\mathbf{Q}\left[\mathbf{Q}^T(\mathbf{I}+\mathbf{H}\mathbf{H}^T)\mathbf{Q}+\mathbf{T}\right]^{-1}$ implying

$$g(\mathbf{P}) = n - \mathrm{Tr}\left(\mathbf{H}^T\mathbf{Q}\left[\mathbf{Q}^T(\mathbf{I}+\mathbf{H}\mathbf{H}^T)\mathbf{Q}+\mathbf{T}\right]^{-1}\mathbf{Q}^T\mathbf{H}\right)$$

- Thus, the original formulation becomes

$$\begin{array}{ll} \text{minimize}_{\mathbf{P}} & \mathrm{Tr}\left(\left(\mathbf{I}+\mathbf{H}^T\mathbf{Q}\left(\mathbf{Q}^T\mathbf{Q}+\mathbf{T}\right)^{-1}\mathbf{Q}^T\mathbf{H}\right)^{-1}\right) \\ \text{subject to} & \mathrm{Tr}\left(\mathbf{Q}^T\mathbf{Q}\right) \le p, \quad \mathbf{Q}\in\mathbb{R}^{\ell\times k}, \end{array}$$

Optimal DES Design

- If $\mathbf{T} = 0$, then the objective further simplifies to

$$\hat{g}(\mathbf{Q}) = \mathrm{Tr}\left(\left(\mathbf{I} + \mathbf{H}^T \mathbf{Q}\left(\mathbf{Q}^T\mathbf{Q}\right)^{-1}\mathbf{Q}^T\mathbf{H}\right)^{-1}\right)$$

- The blue matrix is a project matrix of rank k.
- The minimum is then attained by choosing the columns of $\mathbf{Q}$ as the k largest left singular vectors of $\mathbf{H}$.
- This gives the optimal design for the single sensor case $\mathbf{T} = 0$.
- Questions:
  - What if $\mathbf{T} \neq 0$? $\Leftarrow$ still open.
  - What about the multiple sensor case? $\Leftarrow$ partial answers.

Optimal DES Design: Multiple Sensor Case

- The mean squared error at the fusion center can be easily calculated as follows:

$$\begin{aligned} \mathrm{E}(\mathbf{s}-\hat{\mathbf{s}})^2 &= \mathrm{E}\left(\mathbf{s}-\sum_{n=1}^{N}\mathbf{P}_n\left(\mathbf{Q}_n^T\mathbf{x}_n+\mathbf{u}_n\right)\right)^2 \\ &= \left\|\mathbf{I}-\sum_{n=1}^{N}\mathbf{P}_n\mathbf{Q}_n^T\mathbf{H}_n\right\|_F^2+\sum_{n=1}^{N}\mathrm{Tr}\left(\mathbf{P}_n\left(\mathbf{Q}_n^T\mathbf{Q}_n+\mathbf{T}_n\right)\mathbf{P}_n^T\right), \end{aligned}$$

- The optimal linear DES design problem under individual sensor power constraint:

$$\begin{array}{ll} \text{minimize}_{\mathbf{P}_n,\mathbf{Q}_n} & \left\|\mathbf{I}-\sum_{n=1}^{N}\mathbf{P}_n\mathbf{Q}_n^T\mathbf{H}_n\right\|_F^2+\sum_{n=1}^{N}\mathrm{Tr}\left(\mathbf{P}_n\left(\mathbf{Q}_n^T\mathbf{Q}_n+\mathbf{T}_n\right)\mathbf{P}_n^T\right) \\ \text{subject to} & \mathrm{Tr}\left(\mathbf{Q}_n^T\mathbf{Q}_n\right)\le p_n,\quad \mathbf{P}_n\in\mathbb{R}^{\ell\times k_n},\ \mathbf{Q}_n\in\mathbb{R}^{\ell_n\times k_n},\ \ 1\le n\le N, \end{array}$$

where $p_n > 0$ is the power budget for sensor n.

Optimal DES Design: Multiple Sensor Case

- Minimizing $\mathbf{P}_n$ first yields the equivalent formulation

$$\begin{array}{ll} \text{minimize} & \hat{g}(\mathbf{Q}) = \mathrm{Tr}\left(\left(\mathbf{I} + \sum_{n=1}^{N} \mathbf{H}_n^T \mathbf{Q}_n \left(\mathbf{Q}_n^T \mathbf{Q}_n + \mathbf{T}_n\right)^{-1} \mathbf{Q}_n^T \mathbf{H}_n\right)^{-1}\right) \\ \text{subject to} & \mathrm{Tr}\left(\mathbf{Q}_n \mathbf{Q}_n^T\right) \le p_n, \quad \mathbf{Q}_n \in \mathbb{R}^{\ell_n \times k_n}, \quad n = 1, 2, \ldots, N, \end{array}$$

- The formulation is still nonconvex in the design variables $\{\mathbf{Q}_n\}$, and thus remains difficult to handle numerically.
- **Theorem:** The computational complexity of designing an optimal linear DES is NP-hard, even in the case where $\mathbf{T}_n = \mathbf{0}$ and $\ell_n = 1$ for all n.
- In other words, if each sensor is required to send exactly one linear message to the FC, then the problem of optimally designing these linear message functions (in the MMSE sense) is NP-hard.

Two Special Polynomial Time Solvable Cases

Analytic solution for the case of distortionless channels:

Theorem 1. *If each sensor sends at least* $\min\{\ell_n, \ell\}$ *real-valued messages to the fusion center* $($*i.e.,* $k_n \geq \min\{\ell_n, \ell\}$*, for all* $n)$ *and if sensor channels are distortionless* $($*i.e.,* $\mathbf{T}_n = \mathbf{0}$*, for all* $n)$*, then the optimal linear message functions are given by*

$$\mathbf{m}(\mathbf{x}_n) = \mathbf{Q}_n^T \mathbf{x}_n, \quad \textit{with} \;\; \mathbf{Q}_n = \begin{cases} \sqrt{p_n}\hat{\mathbf{P}}_n^T \begin{bmatrix} \mathbf{I}_\ell \\ \mathbf{0}_{\ell_n - \ell} \end{bmatrix}, & n \in \mathcal{N}_>, \\ \sqrt{p_n} \begin{bmatrix} \mathbf{I}_{\ell_n} & \mathbf{0}_{\ell - \ell_n} \end{bmatrix}, & n \in \mathcal{N}_\leq, \end{cases} \tag{5}$$

where $\hat{\mathbf{P}}_n \in \mathbb{R}^{\ell_n \times \ell_n}$ *is the Q-factor in the QR factorization of* $\mathbf{H}_n$*. Moreover, in this case, the minimum achievable MSE is given by*

$$\mathrm{Tr}\left(\left(\mathbf{I} + \sum_{n=1}^{N} \mathbf{H}_n^T \mathbf{H}_n\right)^{-1}\right). \tag{6}$$

Two Special Polynomial Time Solvable Cases

Semi-definite programming reformulation for the case $k_n \geq \min\{\ell_n, \ell\}$ and $\mathbf{T}_n = t_n \mathbf{I}$

Theorem 2. *Assume sensor channel noises are white with covariance matrices* $\mathbf{T}_n = t_n \mathbf{I}$ *for some* $t_n > 0$*. If the number of transmitted messages from sensor* n *is at least* $k_n \geq \min\{\ell_n, \ell\}$ *for all* n*, then the optimal linear DES design can be obtained by first solving the following SDP* $($*in the matrix variables* $\mathbf{W}$, $\{\mathbf{U}_n, \mathbf{X}_n : n \in \mathcal{N}_{\leq}\}$, $\{\hat{\mathbf{U}}_n, \hat{\mathbf{X}}_n : n \in \mathcal{N}_{>}\})$*:*

$$
\begin{array}{ll}
\textit{minimize} & \mathrm{Tr}\,(\mathbf{W}) \\
\textit{subject to} & \begin{bmatrix} \mathbf{W} & \mathbf{I} \\ \mathbf{I} & \mathbf{S} - \sum\limits_{n \in \mathcal{N}_{\leq}} \mathbf{U}_n - \sum\limits_{n \in \mathcal{N}_{>}} \hat{\mathbf{U}}_n \end{bmatrix} \succeq \mathbf{0}, \\
& \begin{bmatrix} \mathbf{U}_n & \mathbf{H}_n^T \\ \mathbf{H}_n & \mathbf{I} + \mathbf{X}_n \end{bmatrix} \succeq \mathbf{0}, \quad \mathrm{Tr}\,(\mathbf{X}_n) \leq t_n^{-1} p_n, \quad \mathbf{X}_n \succeq \mathbf{0}, \quad \forall\, n \in \mathcal{N}_{\leq}, \\
& \begin{bmatrix} \hat{\mathbf{U}}_n & \hat{\mathbf{R}}_n^T \\ \hat{\mathbf{R}}_n & \mathbf{I} + \hat{\mathbf{X}}_n \end{bmatrix} \succeq \mathbf{0}, \quad \mathrm{Tr}\,(\hat{\mathbf{X}}_n) \leq t_n^{-1} p_n, \quad \hat{\mathbf{X}}_n \succeq \mathbf{0}, \quad \forall\, n \in \mathcal{N}_{>}.
\end{array}
$$

and then setting

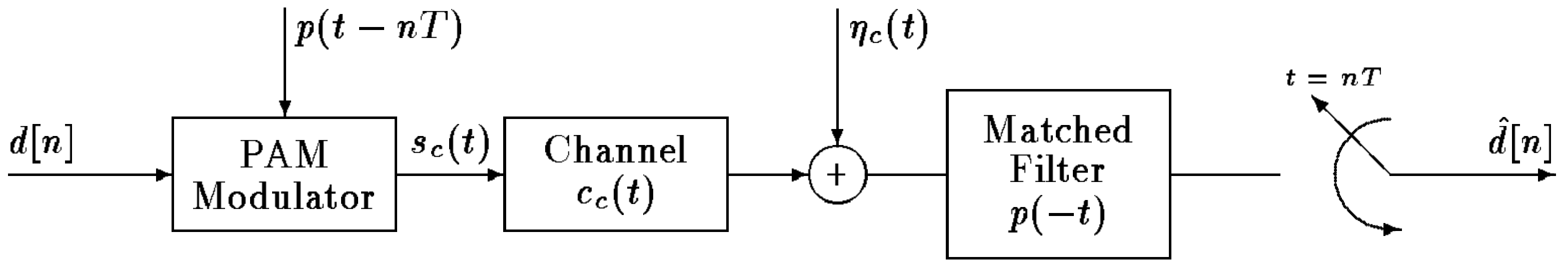
$$\begin{cases} \mathbf{Q}_n = \sqrt{t_n}\begin{bmatrix} \mathbf{X}_n^{1/2} & \mathbf{0}_{k_n - \ell_n} \end{bmatrix}, & n \in \mathcal{N}_{\leq}, \\ \hat{\mathbf{Q}}_n = \sqrt{t_n}\begin{bmatrix} \hat{\mathbf{X}}_n^{1/2} & \mathbf{0}_{k_n - \ell} \end{bmatrix}, & n \in \mathcal{N}_{>}. \end{cases} \tag{7}$$

Pulse Shaping Filter Design

Pulse Shaping Filter Design

- Nonconvex constraints can be converted to semi-infinite convex constraints.
- Some semi-infinite linear constraints can be compactly represented using LMI.
- Design of orthogonal signalling waveforms (pulse shapes) for digital communications can be formulated as an SDP
- Previously awkward design problems can now be efficiently solved, including:
  - minimum transmission bandwidth for a given filter length
  - minimum filter length for a given bandwidth or spectral mask
- Designs are effective—they produce waveforms with superior performance to those chosen in recent standards for digital mobile telephony.

Analogue Baseband PAM



$$s_c(t) = \sum_n d[n]p(t - nT)$$

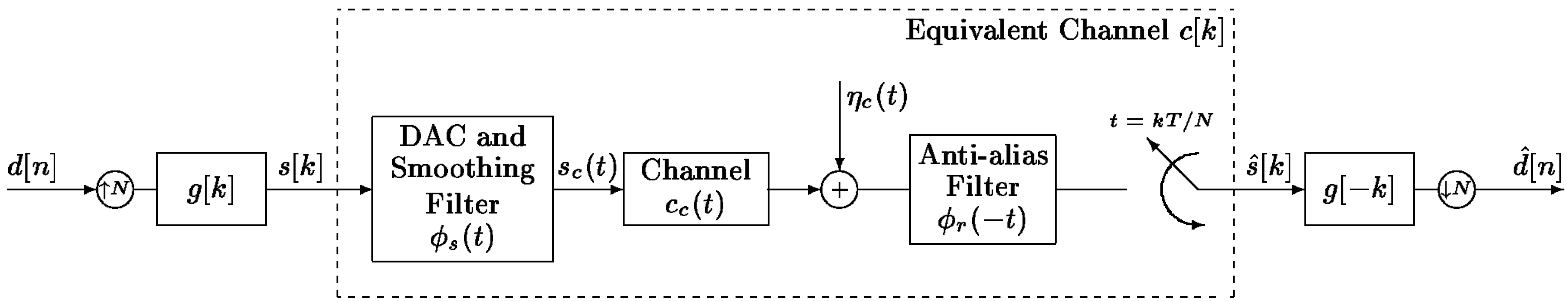
The ‘channel’ generally includes modulators and demodulators

Common Design Goal: Find a waveform $p(t)$ which minimizes the ‘spectral occupation’ subject to the ‘orthogonality’ constraint:

$$\int p(t)p(t - nT)\,dt = \delta[n],$$

where $\delta[n]$ is the Kronecker delta.

DSP-Based Baseband PAM



Now $s_c(t)$ has the same form, with

$$p(t) = \sum_{k=0}^{L-1} g[k]\phi_s(t - kT/N),$$

- waveform design reduces to the design of a finite impulse response (FIR) filter, $g[k]$
- DSP implementation removes many of the physical constraints on analogue waveform design

DSP-Based PAM II

- For most "reasonable" converters, the **design goal** reduces to finding a filter $g[k]$ which minimizes the spectral occupation subject to the orthogonality constraint:

$$\sum_k g[k]g[k-Nn] = \delta[n].$$

- The spectral occupation is specified by a spectral mask:

$$L(e^{j\omega}) \leq |G(e^{j\omega})| \leq U(e^{j\omega}), \qquad \text{for all } \omega \in [0, \pi]$$

for some specified $M_\ell(f)$ and $M_u(f)$.

We now demonstrate the principles of our design technique by studying several formulations of a simple feasibility problem for the filter $g[k]$

A Simple Feasibility Problem

For a given N, L and spectral mask $L(e^{j\omega})$, $U(e^{j\omega})$, either find an orthogonal filter $g[k]$ of length at most L satisfying the spectral mask, or show that none exists.

Formulation 1. *Given* N *and* L*, either find a filter* $g[k]$*,* $0 \leq k \leq L-1$ *such that*

$$\boxed{\sum_{k=\ell N}^{L-1} g[k]g[k-N\ell] = \delta[\ell], \quad \ell = 0, 1, \ldots, \lfloor (L-1)/N \rfloor,} \tag{8a}$$

$$\boxed{L(e^{j\omega}) \leq |G(e^{j\omega})| \leq U(e^{j\omega}), \quad \forall \omega \in [0, \pi]} \tag{8b}$$

or show that none exists.

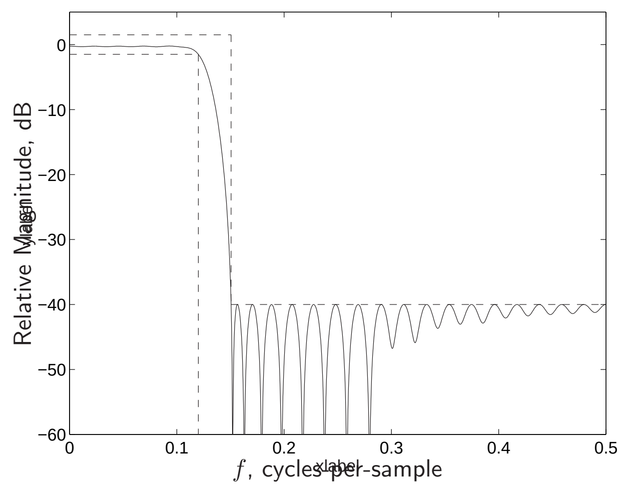
Unfortunately, both constraints are non-convex in $g[k]$.

Spectral Mask Constraints

Spectral mask on polynomial $G(\cdot)$:

$$L(e^{j\omega}) \le |G(e^{j\omega})| \le U(e^{j\omega}) \quad \forall \omega \in [0, 2\pi).$$

- $L(\cdot)$ and $U(\cdot)$ are given. Typically piece-wise constant.
- $G(e^{j\omega}) = \sum_{k=0}^{n} g[k]e^{-jk\omega}$ is a complex trigonometric polynomial
- Coefficients $g[0], \ldots, g[n]$ to be designed



Non-convexity

Each $g[k]$ satisfies

$$L(e^{j\omega}) \leq |G(e^{j\omega})| \leq U(e^{j\omega}) \quad \forall \omega \in [0, 2\pi).$$

(Due to the first '$\leq$' part.)

Observe: if g is real then

$$|G(e^{j\omega})|^2 = \left(\sum_{k=0}^{n} g[k] \cos k\omega \right)^2 + \left(\sum_{k=0}^{n} g[k] \sin k\omega \right)^2.$$

Convex Reformulation

Define autocorrelation coefficients

$$r_g[k] := \sum_i g[k]\overline{g}[k-i].$$

Note in the frequency domain (i.e., under Fourier transform)

$$|G(e^{j\omega})|^2 \quad = \quad \sum_{k=-n}^{n} r_g[k]e^{jk\omega} = r_0 + 2\sum_{k=1}^{n} e^{-jk\omega} r_g[k] =: R_g(e^{j\omega}).$$

The two are equivalent.

Convex Reformulation II

Recall the two nonconvex constraints:

orthogonality constraint: $\boxed{\sum_{k=\ell N}^{L-1} g[k]g[k-N\ell] = \delta[\ell],}$

spectral mask constraint: $\boxed{L(e^{j\omega}) \leq |G(e^{j\omega})| \leq U(e^{j\omega}), \quad \forall \omega \in [0, \pi]}$

Under the new variables $r_g[k]$:

orthogonality constraint $\Rightarrow$ $\boxed{r_g[\ell N] = \delta[\ell],}$

spectral mask constraint $\Rightarrow$ $\boxed{L(e^{j\omega})^2 \leq R_g(e^{j\omega}) \leq U(e^{j\omega})^2, \quad \forall \omega \in [0, 2\pi).}$

A Hidden Condition

- **Question:** how to ensure the existence of $\{g[k]\}$ such that

$$r_g[k] = \sum_i g[k]\overline{g}[k-i].$$

- **Answer:** spectral factorization **(Riesz-Fejer)**

$$\boxed{\exists G(\cdot): \quad R_g(e^{j\omega}) = |G(e^{j\omega})|^2 \quad \forall \omega \in [0, 2\pi)}$$

$$\Updownarrow$$

$$\boxed{R_g(e^{j\omega}) \geq 0 \quad \forall \omega \in [0, 2\pi).}$$

A Semi-infinite LP formulation

For the autocorrelation of the filter $r_g[m] = \sum_k g[k]g[k+m]$, we have $R_g(e^{j\omega}) = |G(e^{j\omega})|^2$. Hence an equivalent formulation is:

Formulation 2. *Given N, L and the spectral mask $L(e^{j\omega})$, $U(e^{j\omega})$, either find an autocorrelation sequence $r_g[m]$, $1-L \leq m \leq L-1$, with $r_g[-m] = r_g[m]$, such that*

$$r_g[\ell N] = \delta[\ell], \qquad \text{for } \ell = 0, 1, \ldots, \lfloor (L-1)/N \rfloor, \tag{9a}$$

$$L(e^{j\omega})^2 \leq R_g(e^{j\omega}) \leq U(e^{j\omega})^2, \quad \forall \omega \in [0, 2\pi). \tag{9b}$$

$$R_g(e^{j\omega}) \geq 0, \qquad \text{for all } \omega \in [0, \pi], \tag{9c}$$

or show that none exists.

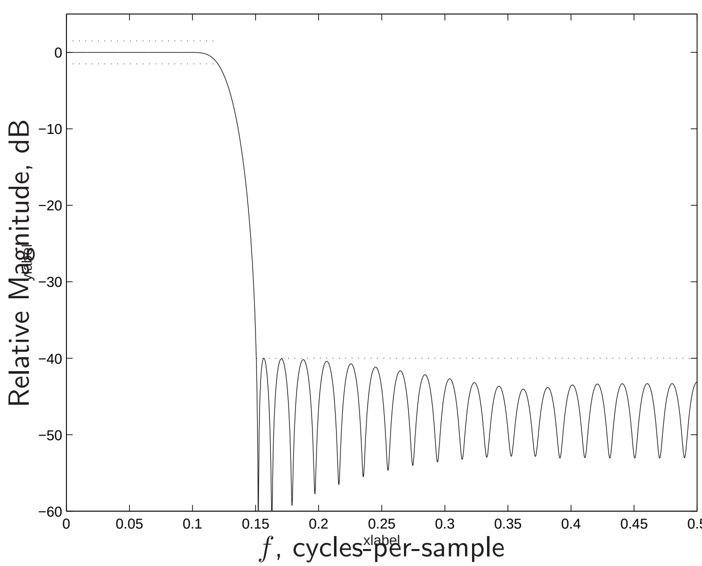
Note: $R_g(e^{j\omega}) = r_g[0] + 2\sum_{m=1}^{L-1} r_g[m]\cos(m\omega)$.

Important: the conditions (9b), (9c) can be transformed to LMI!

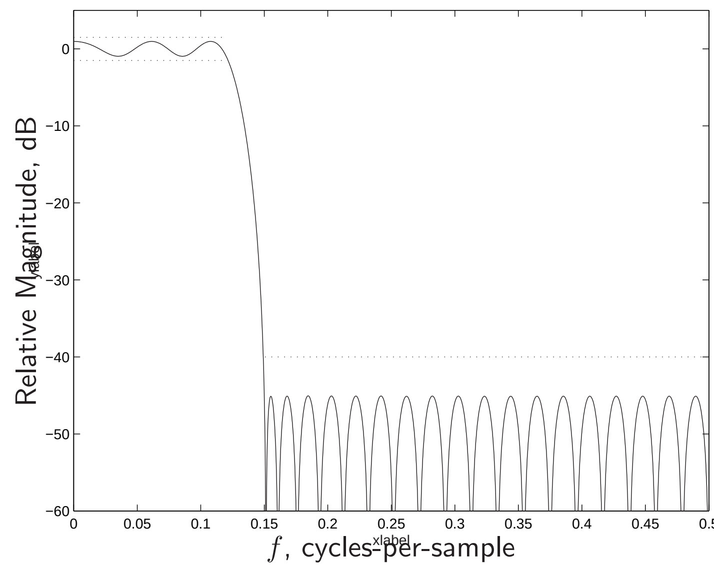
Example: Chip Waveform Design for IS95

- Design a 'chip' waveform to compete with that specified in the IS95 standard for Code Division Multiple Access (CDMA) mobile telephony.
- Design specifications from IS95: relative spectral mask in dB, $N = 4$
- IS95 recommended filter: $L = 48$ with symmetric coefficients (linear phase), but does not satisfy the orthogonality constraints
- Design problem: Find the minimum length filter which satisfies the spectral mask and the orthogonality constraints
- Results in a filter of length $L = 51$
- This provides a substantial performance improvement for a marginal increase in implementation cost

Spectra of the filters



(a) Designed filter, $L = 51$

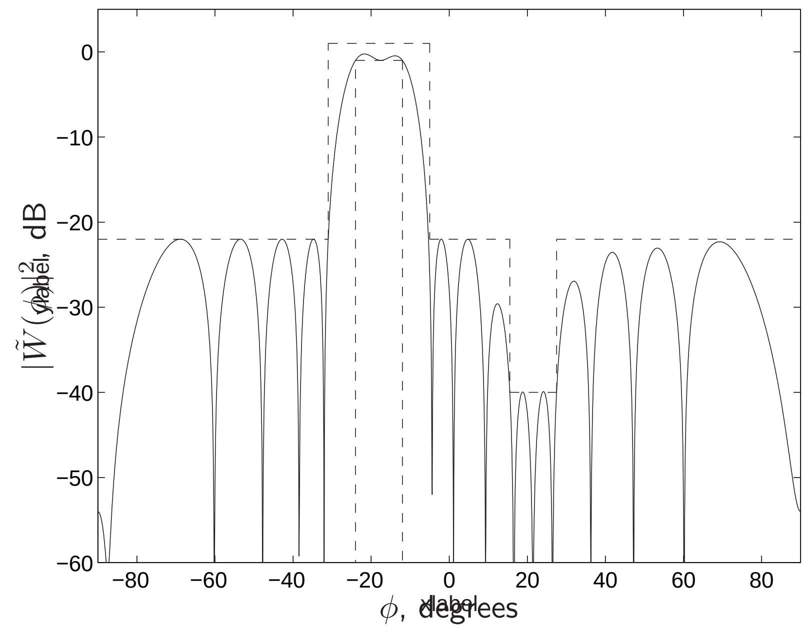


(b) IS95 filter, $L = 48$

Power spectra of the filters in the example with the magnitude bounds from the IS95 standard. Note: $L = 60$ is required to obtain orthogonality and the mask $achieved$ by the IS95 filter

Spectral Mask Constraint: A Beamforming Example

- A linear array of equi-spaced antennas behaves like a spatial filter.
- Spatial frequency mask cosntraint: explicitly suppresses sidelobes



MMSE Multi-Access Transceiver Design

Optimal Transceiver Design: Two-User Multi-Access

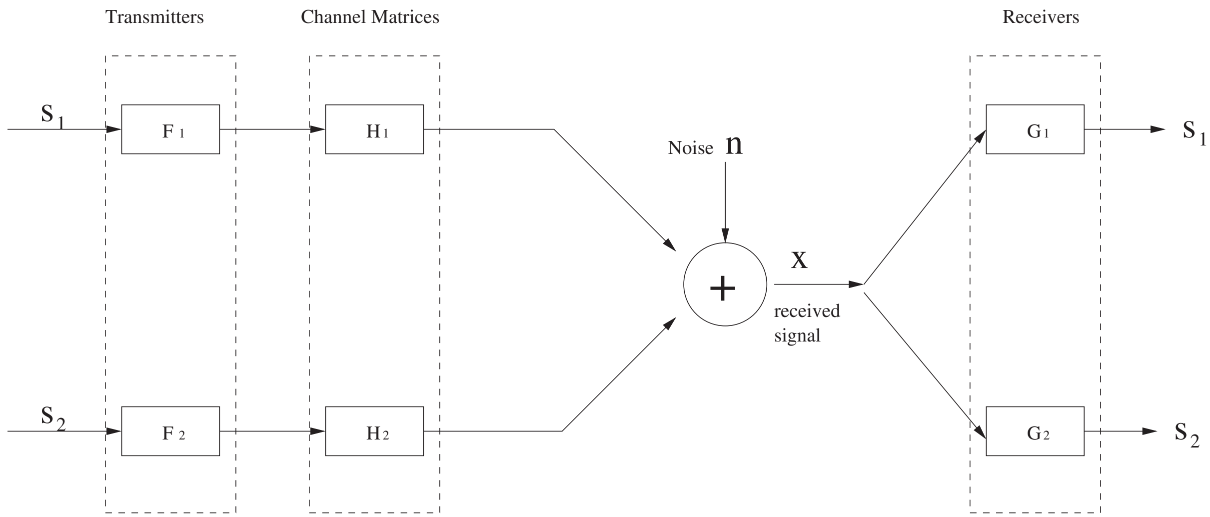


Diagram of Two–user Communication System

Mathematical model: $\mathbf{x} = \mathbf{H}_1\mathbf{F}_1\mathbf{s}_1 + \mathbf{H}_2\mathbf{F}_2\mathbf{s}_2 + \rho\mathbf{n}, \quad \rho > 0.$
Detection: $\mathbf{s}_i = \mathrm{sign}\left(\mathbf{G}_i\mathbf{x}\right),\ i = 1, 2.$

Goal: Given the channel matrices, $\mathbf{H}_1$, $\mathbf{H}_2$, design transceivers $\mathbf{F}_1$, $\mathbf{F}_2$, $\mathbf{G}_1$, $\mathbf{G}_2$.

Mean Square Error

- The detection with receiver (equalizer) $\mathbf{G}_i$: $\hat{\mathbf{s}}_i = \mathrm{sign}\ (\mathbf{G}_i\mathbf{x})$.
- Let $\mathbf{e}_i$ denote the error vector (before making the hard decision) for user i, $i = 1, 2$. Then

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{G}_1\mathbf{x} - \mathbf{s}_1 = \mathbf{G}_1(\mathbf{H}_1\mathbf{F}_1\mathbf{s}_1 + \mathbf{H}_2\mathbf{F}_2\mathbf{s}_2 + \rho\mathbf{n}) - \mathbf{s}_1 \\ &= (\mathbf{G}_1\mathbf{H}_1\mathbf{F}_1 - \mathbf{I})\,\mathbf{s}_1 + \mathbf{G}_1\mathbf{H}_2\mathbf{F}_2\mathbf{s}_2 + \rho\mathbf{G}_1\mathbf{n}. \end{aligned}$$

- This further implies

$$E(\mathbf{e}_1\mathbf{e}_1^\dagger) = (\mathbf{G}_1\mathbf{H}_1\mathbf{F}_1 - \mathbf{I})\,(\mathbf{G}_1\mathbf{H}_1\mathbf{F}_1 - \mathbf{I})^\dagger + (\mathbf{G}_1\mathbf{H}_2\mathbf{F}_2)\,(\mathbf{G}_1\mathbf{H}_2\mathbf{F}_2)^\dagger + \rho^2\mathbf{G}_1\mathbf{G}_1^\dagger$$

- Similarly, we have

$$E(\mathbf{e}_2\mathbf{e}_2^\dagger) = (\mathbf{G}_2\mathbf{H}_2\mathbf{F}_2 - \mathbf{I})\,(\mathbf{G}_2\mathbf{H}_2\mathbf{F}_2 - \mathbf{I})^\dagger + (\mathbf{G}_2\mathbf{H}_1\mathbf{F}_1)\,(\mathbf{G}_2\mathbf{H}_1\mathbf{F}_1)^\dagger + \rho^2\mathbf{G}_2\mathbf{G}_2^\dagger.$$

Formulation: MMSE Transceiver Design

- Our goal is to design a set of transmitting matrix filters $\mathbf{F}_i$ and a set of matrix equalizers $\mathbf{G}_i$ such that the total mean squared error

$$\text{MSE} = \mathrm{Tr}(E(\mathbf{e}_1\mathbf{e}_1^\dagger)) + \mathrm{Tr}(E(\mathbf{e}_2\mathbf{e}_2^\dagger))$$

  is minimized.
- As is always the case in practice, there are power constraints on the transmitting matrix filters:

$$\mathrm{Tr}(\mathbf{F}_1\mathbf{F}_1^\dagger) \le p_1, \quad \mathrm{Tr}(\mathbf{F}_2\mathbf{F}_2^\dagger) \le p_2$$

- The above is nonconvex.
- We first eliminate the variables $\mathbf{G}_1$, $\mathbf{G}_2$: the MMSE equalizers.

Formulation: MMSE Transceiver Design

- By minimizing $E(\mathbf{e}_1\mathbf{e}_1^\dagger)$ with respect to $\mathbf{G}_1$, we obtain the following MMSE equalizer for user 1: $\mathbf{G}_1 = \mathbf{F}_1^\dagger \mathbf{H}_1^\dagger \mathbf{W}$, where

$$\mathbf{W} = \left(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^\dagger\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^\dagger\mathbf{H}_2^\dagger + \rho^2\mathbf{I}\right)^{-1}.$$

- Substituting this into $E(\mathbf{e}_1\mathbf{e}_1^\dagger)$ gives:

$$E(\mathbf{e}_1\mathbf{e}_1^\dagger) = -\mathbf{F}_1^\dagger\mathbf{H}_1^\dagger\mathbf{W}\mathbf{H}_1\mathbf{F}_1 + \mathbf{I}.$$

- Similarly, the MMSE equalizer $\mathbf{G}_2$ for user 2 is given by $\mathbf{G}_2 = \mathbf{F}_2^\dagger\mathbf{H}_2^\dagger\mathbf{W}$ and resulting minimized (with respect to $\mathbf{G}_2$) mean square error for user 2 is given by:

$$E(\mathbf{e}_2\mathbf{e}_2^\dagger) = -\mathbf{F}_2^\dagger\mathbf{H}_2^\dagger\mathbf{W}\mathbf{H}_2\mathbf{F}_2 + \mathbf{I}.$$

Total MSE

Substituting into the above expression gives rise to

$$\begin{aligned}
\text{MSE} &= \mathrm{Tr}(E(\mathbf{e}_1\mathbf{e}_1^\dagger)) + \mathrm{Tr}(E(\mathbf{e}_2\mathbf{e}_2^\dagger)) \\
&= -\mathrm{Tr}\left(\mathbf{F}_1^\dagger\mathbf{H}_1^\dagger\mathbf{W}\mathbf{H}_1\mathbf{F}_1\right) - \mathrm{Tr}\left(\mathbf{F}_2^\dagger\mathbf{H}_2^\dagger\mathbf{W}\mathbf{H}_2\mathbf{F}_2\right) + 2n \\
&= -\mathrm{Tr}\left(\mathbf{W}\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^\dagger\mathbf{H}_1^\dagger\right) - \mathrm{Tr}\left(\mathbf{W}\mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^\dagger\mathbf{H}_2^\dagger\right) + 2n \\
&= -\mathrm{Tr}\left(\mathbf{W}(\mathbf{H}_1\mathbf{F}_1\mathbf{F}_1^\dagger\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{F}_2\mathbf{F}_2^\dagger\mathbf{H}_2^\dagger)\right) + 2n \\
&= \rho^2\mathrm{Tr}\left(\mathbf{W}\right) + n,
\end{aligned}$$

where the last step follows from the definition of $\mathbf{W}$.

Formulation: MMSE Transceiver Design

- Introduce matrix variables: $\mathbf{U}_1 = \mathbf{F}_1\mathbf{F}_1^\dagger, \quad \mathbf{U}_2 = \mathbf{F}_2\mathbf{F}_2^\dagger$.
- Then the MMSE transceiver design problem becomes

$$\begin{array}{ll} \text{minimize}_{\mathbf{U}_1,\mathbf{U}_2} & \mathrm{Tr}\left((\mathbf{H}_1\mathbf{U}_1\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{U}_2\mathbf{H}_2^\dagger + \rho^2\mathbf{I})^{-1}\right) \\ \text{subject to} & \mathrm{Tr}(\mathbf{U}_1) \le p_1, \quad \mathrm{Tr}(\mathbf{U}_2) \le p_2, \\ & \mathbf{U}_1 \ge \mathbf{0}, \quad \mathbf{U}_2 \ge \mathbf{0}. \end{array}$$

- Reformulate using the auxiliary matrix variable $\mathbf{W}$:

$$\begin{array}{ll} \text{minimize}_{\mathbf{W},\mathbf{U}_1,\mathbf{U}_2} & \mathrm{Tr}\left(\mathbf{W}\right) \\ \text{subject to} & \mathrm{Tr}(\mathbf{U}_1) \le p_1, \quad \mathrm{Tr}(\mathbf{U}_2) \le p_2, \\ & \mathbf{W} \ge (\mathbf{H}_1\mathbf{U}_1\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{U}_2\mathbf{H}_2^\dagger + \rho^2\mathbf{I})^{-1} \\ & \mathbf{U}_1 \ge \mathbf{0}, \quad \mathbf{U}_2 \ge \mathbf{0}. \end{array}$$

SDP Formulation

- The constraint $\mathbf{W} \geq (\mathbf{H}_1\mathbf{U}_1\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{U}_2\mathbf{H}_2^\dagger + \rho^2\mathbf{I})^{-1}$ is equivalent to LMI:

$$\begin{bmatrix} \mathbf{W} & \mathbf{I} \\ \mathbf{I} & \mathbf{H}_1\mathbf{U}_1\mathbf{H}_1^\dagger + \mathbf{H}_2\mathbf{U}_2\mathbf{H}_2^\dagger + \rho^2\mathbf{I} \end{bmatrix} \geq 0. \qquad (3)$$

- We obtain an SDP formulation:

$$\boxed{\begin{array}{ll} \text{minimize}_{\mathbf{W},\mathbf{U}_1,\mathbf{U}_2} & \mathrm{Tr}\,(\mathbf{W}) \\ \text{subject to} & \mathrm{Tr}(\mathbf{U}_1) \leq p_1, \quad \mathrm{Tr}(\mathbf{U}_2) \leq p_2, \\ & \mathbf{W} \text{ satisfies (3)}, \\ & \mathbf{U}_1 \geq \mathbf{0}, \quad \mathbf{U}_2 \geq \mathbf{0}. \end{array}}$$

- Interior point method with arithmetic complexity $O(n^{6.5}\log(1/\epsilon))$, $\epsilon > 0$ is the solution accuracy.

OFDM: Diagonal Designs are Optimal!

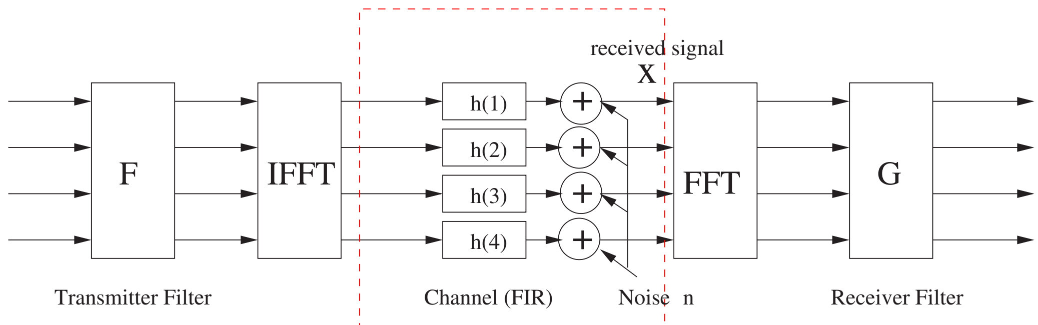
Result

If $\mathbf{H}_1$ and $\mathbf{H}_2$ are diagonal, as in the OFDM systems, then the optimal transmitters are also diagonal.

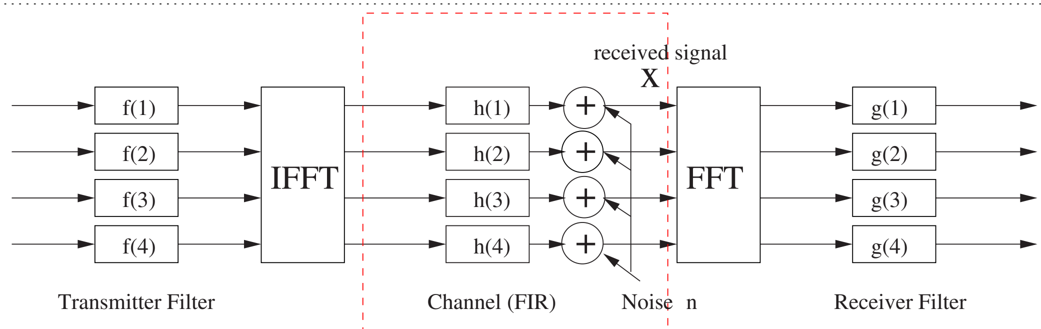
Implication

The MMSE transceivers for an multi-user OFDM system can be implemented by optimally setting the data rates and allocating power to each subcarrier for all the users.

Linearly Precoded/Power Loaded OFDM



General Linearly Precoded OFDM System



Power–Loaded OFDM System

From SDP to SOCP Formulation

- Restricting to diagonal designs, the SDP becomes SOC:

$$
\begin{array}{ll}
\text{minimize}_{\mathbf{w},\mathbf{u}_1,\mathbf{u}_2} & \displaystyle\sum_{i=1}^{n} \mathbf{w}_i \\
\text{subject to} & \displaystyle\sum_{i=1}^{n} \mathbf{u}_1(i) \le p_1, \quad \sum_{i=1}^{n} \mathbf{u}_2(i) \le p_2, \\
& \mathbf{w}_i \left(|\mathbf{h}_1(i)|^2 \mathbf{u}_1(i) + |\mathbf{h}_2(i)|^2 \mathbf{u}_2(i) + \rho^2 \right) \ge 1, \\
& \mathbf{u}_1(i) \ge 0, \quad \mathbf{u}_2(i) \ge 0, \quad i = 1, 2, \ldots, n.
\end{array}
$$

- There exist highly efficient (general purpose) interior point methods to solve the above second order cone program.

- Arithmetic complexity $O(n^{3.5}\log(1/\epsilon))$, $\epsilon > 0$ is the accuracy.

Properties of Optimal MMSE Transceiver

- Let $\mathbf{u}_1^* \geq 0$, $\mathbf{u}_2^* \geq 0$ be the optimal transceivers. Define:

$$\begin{cases} I_1 = \{i \mid \mathbf{u}_1^*(i) > 0,\ \mathbf{u}_2^*(i) = 0\}, & I_2 = \{i \mid \mathbf{u}_1^*(i) = 0,\ \mathbf{u}_2^*(i) > 0\}, \\ I_s = \{i \mid \mathbf{u}_1^*(i) > 0,\ \mathbf{u}_2^*(i) > 0\}, & I_u = \{i \mid \mathbf{u}_1^*(i) = 0,\ \mathbf{u}_2^*(i) = 0\}. \end{cases}$$

- I_1, I_2: subcarriers allocated to user 1 and user 2;
  I_s and I_u: subcarriers *shared* and *unused*;
  data rates: $(|I_1| + |I_s|)/n$, $(|I_2| + |I_s|)/n$

- – For each $i \in I_1$ and $j \in I_2$, we have $\frac{|\mathbf{h}_1(i)|^2}{|\mathbf{h}_2(i)|^2} \geq \frac{|\mathbf{h}_1(j)|^2}{|\mathbf{h}_2(j)|^2}$.

  – For all $i, j \in I_s$, we have $\frac{|\mathbf{h}_1(i)|^2}{|\mathbf{h}_2(i)|^2} = \frac{|\mathbf{h}_1(j)|^2}{|\mathbf{h}_2(j)|^2}$.

  – For any $i \in I_u$ and any $j \in I_1 \cup I_s$, we have $|\mathbf{h}_1(i)|^2 < |\mathbf{h}_1(j)|^2$. Similarly, for any $i \in I_u$ and any $j \in I_2 \cup I_s$, we have $|\mathbf{h}_2(i)|^2 < |\mathbf{h}_2(j)|^2$.

Intuitive Interpretation

- $\mathbf{x} = \mathbf{H}_1\mathbf{F}_1\mathbf{s}_1 + \mathbf{H}_2\mathbf{F}_2\mathbf{s}_2 + \rho\mathbf{n}$, with $\mathbf{H}_i$, $\mathbf{F}_i$ diagonal; $\mathbf{x}(i) = \mathbf{h}_1(i)\mathbf{f}_1(i)\mathbf{s}_1(i) + \mathbf{h}_2(i)\mathbf{f}_2(i)\mathbf{s}_2(i) + \rho^2\mathbf{n}(i)$.
- In a fading environment, the path gains $|\mathbf{h}_1(i)|^2$, $|\mathbf{h}_2(i)|^2$ are random,

  $\Rightarrow$ the probability of having two equal path gains is zero.

  $\Rightarrow$ I_s is singleton: *at most one subcarrier should be shared by the two users.*
- The remaining subcarriers are allocated to the two users according to the path gain ratios: subcarrier i to user 1 and subcarrier j to user 2 only if

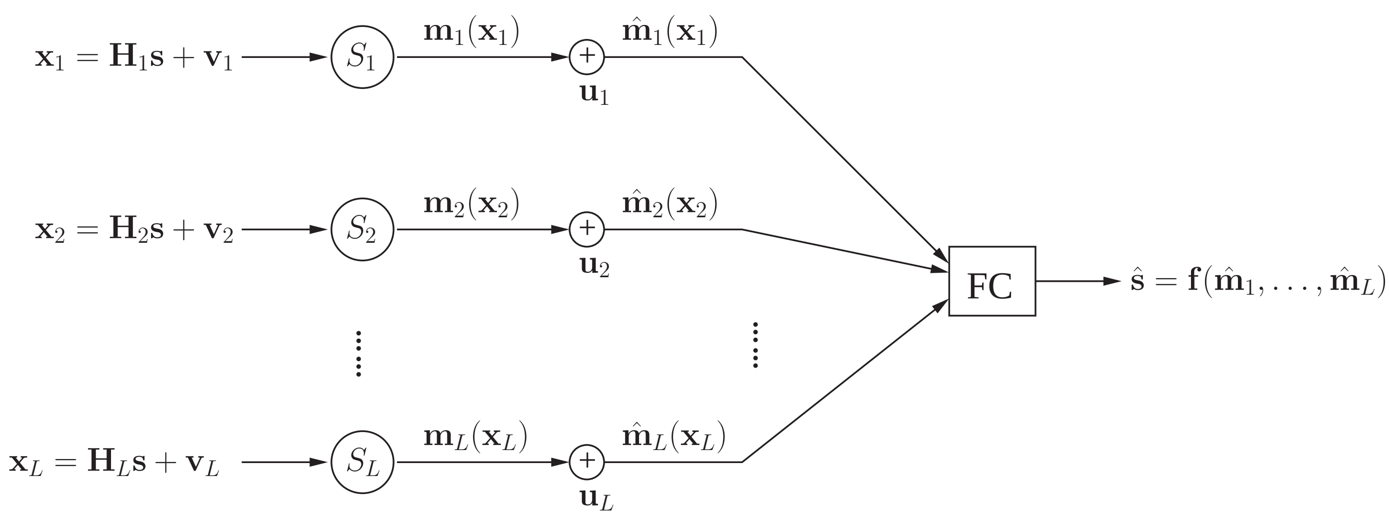
$$\frac{|\mathbf{h}_1(i)|^2}{|\mathbf{h}_2(i)|^2} \geq \frac{|\mathbf{h}_1(j)|^2}{|\mathbf{h}_2(j)|^2}.$$

- The subcarriers in I_u have small path gains for both users (i.e., both $|\mathbf{h}_1(i)|^2$ and $|\mathbf{h}_2(i)|^2$ are small), and they should not be used by either user, i.e., they are useless subcarriers!

Optimal Rate Allocation for Multi-Terminal Source Coding

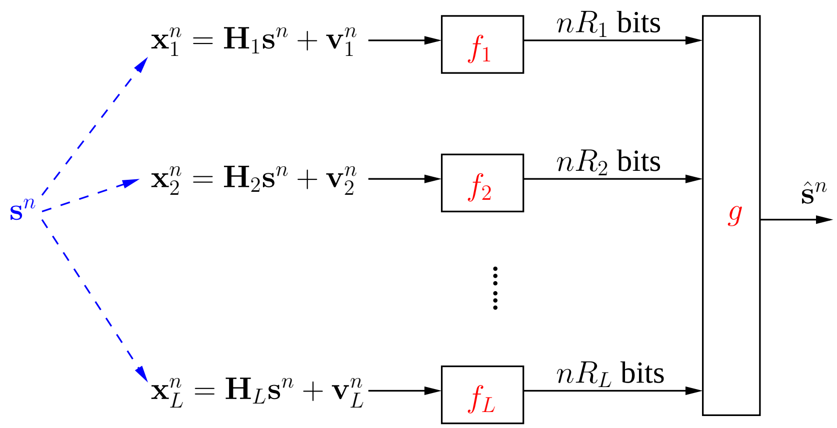
Distributed Vector Source Estimation

- Assume a common vector source $\mathbf{s} \in \mathbb{R}^\ell$
- Sensor observations: $\mathbf{x}_i = \mathbf{H}_i\mathbf{s} + \mathbf{v}_i$, with
- Communication channels: orthogonal AWGN, differing noise variances
- Given a power/bandwidth/MSE requirement, design:
  - ⋆ local message functions $\mathbf{m}_i(\mathbf{x}_i)$,
  - ⋆ final fusion function $\hat{\mathbf{s}} = \mathbf{f}(\hat{\mathbf{m}}_1, \ldots, \hat{\mathbf{m}}_L)$



The Vector Gaussian CEO Problem

- Source $\mathbf{s} \sim \mathcal{N}(0, \mathbf{I}_\ell)$; Linear observation model with (deterministic and known) $\mathbf{H}_i$, and noise $\mathbf{v}_i \sim \mathcal{N}(0, \Omega_{\mathbf{v}_i})$.
- Local observations are quantized to **digital messages with finite rates**, and are transmitted to the FC without error; FC reconstructs $\mathbf{s}$ based on received finite-rate messages.
- Assuming orthogonal links, (the separation theorem holds). **[Xiao-Luo, 2005]**
  $\Rightarrow$ optimal joint design $=$ optimal source coding $+$ optimal channel coding.
- **Rate distortion region** $\mathbb{R}(\mathrm{D})$: the set of all rates that allow the reconstruction of $\mathbf{s}$ at the FC within a given distortion D.



Existing Work

- Berger-Tung **[Berger-Tung'78]** proposed an achievable region for the vector CEO problem.
- The tightness of the Berger-Tung achievable region has not been proved in general.
- For the vector Gaussian CEO problem,
  - **[Gastpar et al'05]** computed the sum-rate in the Berger-Tung achievable region and interpreted it as a distributed Karhunen-Loève transform. They showed that the problem is nonconvex.
  - **[Schizas-Giannakis-Jindal'05]** also analyzed the achievable sum-rate in the Berger-Tung region in a general framework of EC or CE schemes.
  - They both proposed iterative algorithms to compute the optimal rate allocations which minimizes sum-rate in the B-T region.
- **Our work:** the optimal rate allocation problem can be reformulated as a convex **Maxdet** problem.

Berger-Tung Region

- Let us denote $\mathcal{I}_L = \{1, 2, \ldots, L\}$ and define

$$\mathcal{R}(D_0; \Omega_{\mathbf{w}_1}, \Omega_{\mathbf{w}_2}, \ldots, \Omega_{\mathbf{w}_L}) \stackrel{\text{def}}{=} \left\{ (R_1, \ldots, R_L) \;\middle|\; \sum_{i \in \mathcal{A}} R_i \geq I(\mathbf{X}_{\mathcal{A}}; \mathbf{Z}_{\mathcal{A}} | \mathbf{Z}_{\mathcal{A}^c}); \forall\ \mathcal{A} \subseteq \mathcal{I}_L \right\}.$$

- The computation of B-T region requires specification of the conditional probability distribution $P(\mathbf{z}_i \mid \mathbf{x}_i)$.
- Berger-Tung considered the following **Gaussian test channels**

$$\mathbf{z}_i = \mathbf{x}_i + \mathbf{w}_i, \quad i = 1, 2, \ldots,$$

where $\mathbf{w}_i \sim \mathcal{N}(\mathbf{0}, \Omega_{\mathbf{w}_i})$ are independent of $\mathbf{x}_i$. The resulting achievable region is

$$\boxed{\mathcal{R}^{BT}(D_0) = \bigcup_{\{\Omega_{\mathbf{w}_i}\}} \mathcal{R}(D_0; \Omega_{\mathbf{w}_1}, \Omega_{\mathbf{w}_2}, \ldots, \Omega_{\mathbf{w}_L}).}$$

Berger-Tung Region for the Vector Gaussian CEO Problem

- The **sum-rate** achieved through the Gaussian test channels for any given $\{\Omega_{\mathbf{w}_i}\}$ is

$$R_\Sigma \stackrel{\text{def}}{=} \sum_{i=1}^{L} R_i = I(\mathbf{x}_1 \ldots, \mathbf{x}_L; \mathbf{z}_1, \ldots, \mathbf{z}_L)$$

$$= \frac{1}{2}\log\left\{\det\left(\mathbf{I}_\ell + \sum_{i=1}^{L} \mathbf{H}_i^T(\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}\mathbf{H}_i\right)\prod_{i=1}^{L}\det(\mathbf{I}_{\ell_i} + \Omega_{\mathbf{w}_i}^{-1}\Omega_{\mathbf{v}_i})\right\}.$$

- The specification of a Gaussian test channel also determines the final **distortion covariance matrix** of estimating s from $\{\mathbf{z}_i\}$:

$$\mathbf{D} = \left(\mathbf{I}_\ell + \sum_{i=1}^{L} \mathbf{H}_i^T(\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}\mathbf{H}_i\right)^{-1}.$$

Optimal Rate Allocation

- Minimizing **the sum-rate** subject to a distortion constraint leads to

$$\begin{aligned} \min \quad & R_\Sigma = R_1 + \ldots + R_L \\ \text{s.t.} \quad & \mathrm{Tr}(D) \le D_0. \end{aligned}$$

- The optimal rate allocation that is equivalent to **determining the optimal noise covariance matrices** $\Omega_{\mathbf{w}_i}$ for L parallel Gaussian test channels.

- In terms of $\Omega_{\mathbf{w}_i}$, the **optimal rate allocation problem** can be formulated as

$$\boxed{\begin{aligned} \min_{\Omega_{\mathbf{w}_i}} \quad & \frac{1}{2}\log\det\left(\mathbf{I}_\ell + \sum_{i=1}^{L} \mathbf{H}_i^T(\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}\mathbf{H}_i\right) + \frac{1}{2}\sum_{i=1}^{L} \det(\mathbf{I}_{\ell_i} + \Omega_{\mathbf{w}_i}^{-1}\Omega_{\mathbf{v}_i}) \\ \text{s.t.} \quad & \mathrm{Tr}\left(\mathbf{I}_\ell + \sum_{i=1}^{L} \mathbf{H}_i^T(\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}\mathbf{H}_i\right)^{-1} \le D_0, \\ & \Omega_{\mathbf{w}_i} \succeq 0. \end{aligned}}$$

Centralized Case ($L = 1$)

- If $L = 1$, or **full cooperation** is allowed among encoders, then there is analytical solution for the previous problem.
- For the centralized case, the **optimal rate allocation** can be constructed as follows **[Schizas-Giannakis-Jindal'05]**:
  - obtain an MMSE estimator of $\mathbf{s}$ based on $\mathbf{x}$;
  - perform a classical vector Gaussian compression of the MMSE estimator (inverse water-filling).

General Multi-sensor Case: $L > 1$

- When $L > 1$, the optimal rate allocation problem is **not convex** in terms of the matrix variables $\{\Omega_{\mathrm{w}_i} : i = 1, 2, \ldots, L\}$.
- For the general case of $L > 1$, Ω_{w} must have **block-diagonal structure**

$$\Omega_{\mathrm{w}} = \begin{bmatrix} \Omega_{\mathrm{w}_1} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \Omega_{\mathrm{w}_2} & \ldots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \Omega_{\mathrm{w}_L} \end{bmatrix}.$$

  This condition is needed to represent requirement that **encoders must operate independently**.
- **[Gastpar et al'05]** **[Schizas-Giannakis-Jindal'05]** proposed Gauss-Seidel type optimization procedures to choose $\{\Omega_{\mathrm{w}_i}\}$.
  - local traps
  - initialization, stepsize selections, termination

Convex Reformulation

- We can reformulate it as an **equivalent Maxdet problem** that is convex.
- **Step 1:** introduce a distortion covariance matrix $\mathbf{D}$, then we obtain

$$\begin{array}{ll} \min\limits_{\mathbf{D};\Omega_{\mathbf{w}_i}} & -\frac{1}{2}\log\det\mathbf{D} + \sum_{i=1}^{L}\frac{1}{2}\log\det(\mathbf{I}_{\ell_i} + \Omega_{\mathbf{w}_i}^{-1}\Omega_{\mathbf{v}_i}) \\ \text{s.t.} & \mathrm{Tr}(\mathbf{D}) \le D_0, \\ & \Omega_{\mathbf{w}_i} \succeq 0, \ \mathbf{D} \succeq 0, \\ & \left(\mathbf{I}_\ell + \sum_{i=1}^{L}\mathbf{H}_i^T(\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}\mathbf{H}_i\right)^{-1} \preceq \mathbf{D}. \end{array}$$

- **Step 2:** use the Schur complement property to obtain the following equivalent matrix inequality for the third constraint

$$\begin{bmatrix} \mathbf{I}_\ell + \sum_{i=1}^{L}\mathbf{H}_i^T(\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}\mathbf{H}_i & \mathbf{I} \\ \mathbf{I} & \mathbf{D} \end{bmatrix} \succeq 0.$$

Convex Reformulation

- **Step 3:** Introduce new matrix variables $\mathbf{Q}_i = (\Omega_{\mathbf{v}_i} + \Omega_{\mathbf{w}_i})^{-1}$ for each i. Then,

$$\mathbf{I}_{\ell_i} + \Omega_{\mathbf{w}_i}^{-1}\Omega_{\mathbf{v}_i} = (\mathbf{I}_{\ell_i} - \mathbf{Q}_i\Omega_{\mathbf{v}_i})^{-1}.$$

Replacing $\Omega_{\mathbf{w}_i}$ by $\mathbf{Q}_i$, we can get the following equivalent problem:

$$\begin{array}{ll}
\min\limits_{D;\ \mathbf{Q}_i} & -\frac{1}{2}\log\det\mathbf{D} - \sum_{i=1}^{L}\frac{1}{2}\log\det(\mathbf{I}_{\ell_i} - \mathbf{Q}_i\Omega_{\mathbf{v}_i}) \\
\text{s.t.} & \mathrm{Tr}(\mathbf{D}) \le D_0, \\
& \mathbf{D} \succeq 0, \\
& \begin{bmatrix} \mathbf{I}_\ell + \sum_{i=1}^{L}\mathbf{H}_i^T\mathbf{Q}_i\mathbf{H}_i & \mathbf{I} \\ \mathbf{I} & \mathbf{D} \end{bmatrix} \succeq 0.
\end{array}$$

- The above **Maxdet** problem is now convex in the new variables $\{\mathbf{D}, \mathbf{Q}_i : i = 1, 2, \ldots, L\}$.

Numerical Plots of the Sum-Rate Distortion Function

- In the simulation, we take $L = 2$, and

$$\Omega_{\mathbf{s}} = \mathbf{I}_2;$$

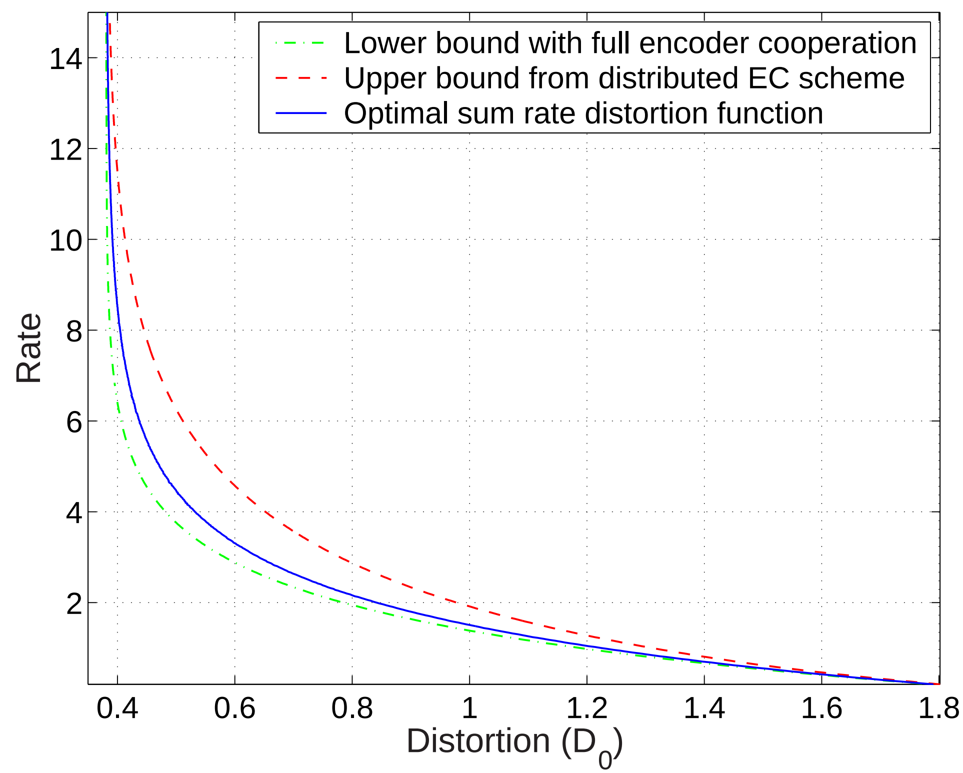
$$\Omega_{\mathbf{v}_1} = \Omega_{\mathbf{v}_2} = \mathbf{I}_2;$$

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

$$\mathbf{H}_2 = \mathbf{U} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \mathbf{U}^T,$$

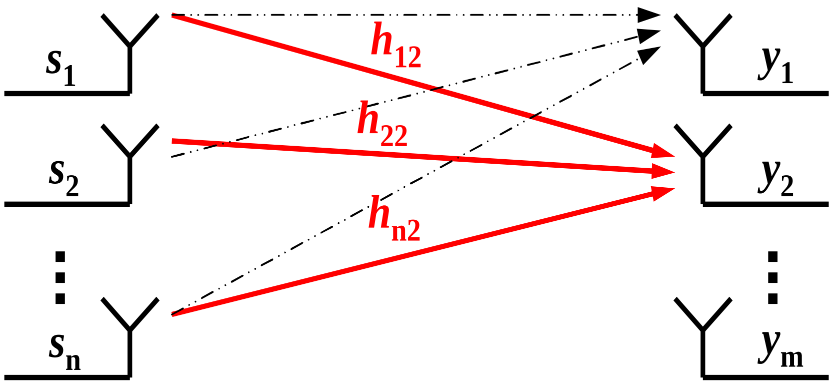
where

$$\mathbf{U} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$



Quasi-ML Detection Via SDP Relaxation

MIMO Channel



$$y_i = \sqrt{\frac{\rho}{n}} \sum_k h_{ik} s_k + v_i$$

Written in the vector form:

$$\boxed{\mathbf{y} = \sqrt{\frac{\rho}{n}}\, \mathbf{H}\ \mathbf{s} + \mathbf{v}}$$

- n - number of transmit antennas;
- m - number of receive antennas;
- $\mathcal{S}^n$ - constellation of signals with dimension n;
- $\mathbf{s} \in \mathcal{S}^n$ - vector of transmitted signals drawn from $\mathcal{C}^n$;
- $\mathbf{y} \in \mathbb{C}^m$ - vector of received signals;
- $\mathbf{H} \in \mathbb{C}^{m \times n}$ - the matrix of fading coefficients, $h_{ik} \sim \mathcal{N}(0,1)$, $\forall i, k$;
- ρ - expected value of SNR at each receiver antenna;
- $\mathbf{v} \in \mathbb{C}^m$ - i.i.d. noise, $v_i \sim \mathcal{N}(0,1)$, $\forall i$.
- Consider $n = m$ in this talk.

Maximum-likelihood detection

- For a memoryless channel with equiprobable input signals (typical in practice), the maximum-likelihood (ML) detection achieves the minimum error probability P_e.
- ML Detector solves the following optimization problem:

$$\mathbf{s}_{ML} = \arg\max_{\mathbf{s}\in\mathcal{C}^n} p(\mathbf{y}|\mathbf{s},\mathbf{H})$$

- For the Gaussian noise ML detection can be written in the form:

$$\mathbf{s}_{ML} = \arg\min_{\mathbf{s}\in\mathcal{C}^n} \left\| \mathbf{y} - \sqrt{\rho/n}\,\mathbf{H}\,\mathbf{s} \right\|^2$$

- The problem is NP-hard in general due to the discrete nature of the constellation set.

Challenge of ML detection

- MIMO systems:
  - Exhaustive search can be applied for small n.
  - For reasonably large n (e.g., $n = 200$), exhaustive search is prohibitively expensive
- Existing approaches:
  - either has polynomial complexity ($\mathcal{O}(n^3)$) but sacrifices performance,
  - or ensures good decoding performance with exponential complexity

Existing suboptimal approaches

- Solve an **unconstrained** problem with **penalty** function:

$$\bar{\mathbf{s}} = \arg\min_{\mathbf{s}\in\mathbb{R}^n} \|\mathbf{y} - \sqrt{\rho/n}\,\mathbf{H}\mathbf{s}\|^2 + \frac{\gamma}{2}\|\mathbf{s}\|^2$$

- Assuming: invertibility of $\mathbf{H}^T\mathbf{H} + \gamma\mathbf{I}$; BPSK signalling, the estimate of ML solution can be expressed as

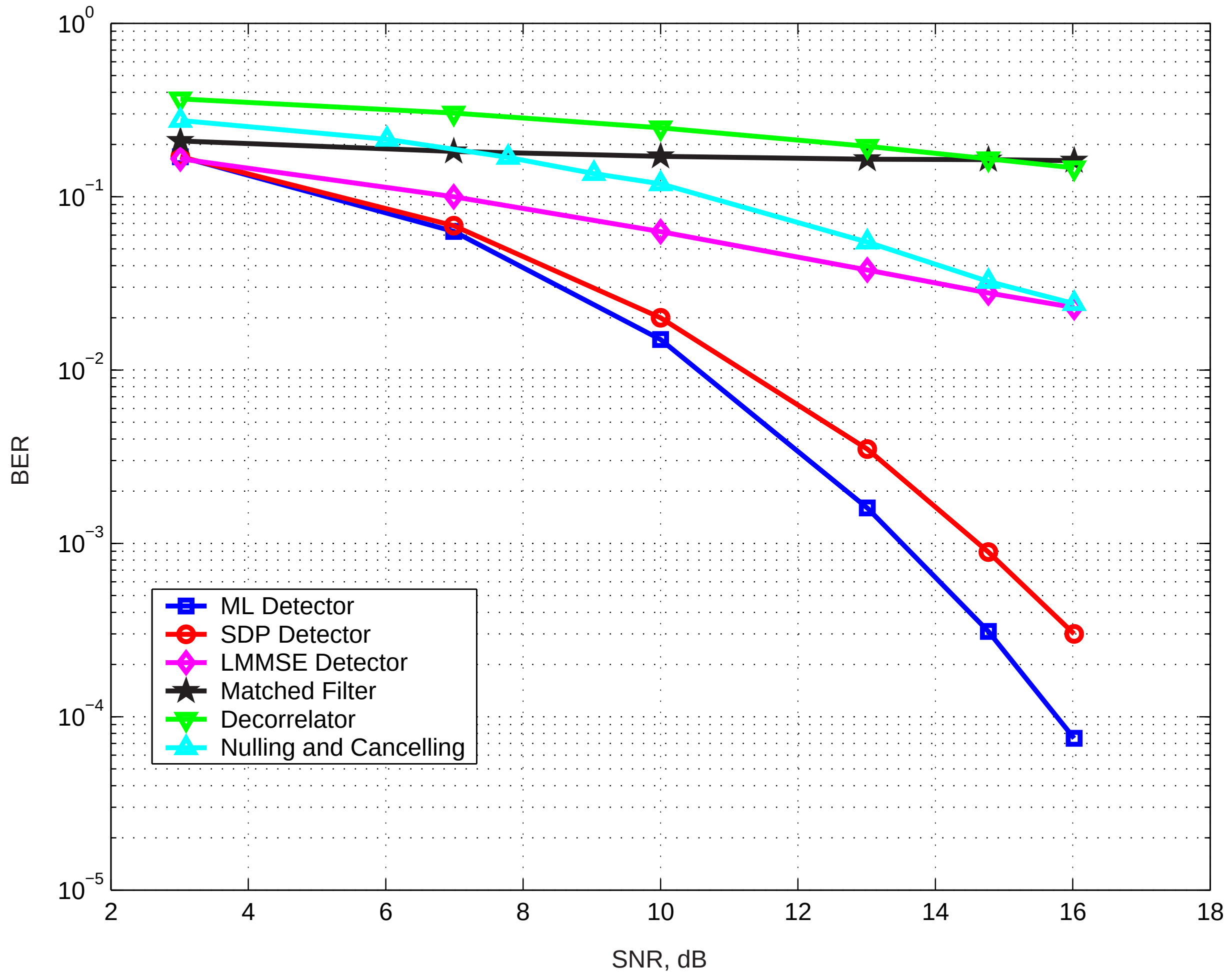
$$\hat{\mathbf{s}} = \operatorname{sign}\left(\sqrt{\frac{\rho}{n}}\left(\frac{\rho}{n}\mathbf{H}^T\mathbf{H} + \gamma\mathbf{I}\right)^{-1}\mathbf{H}^T\mathbf{y}\right)$$

- Choosing different values for penalty factor γ we obtain

$$\begin{aligned} \gamma &= 0, && \text{for } \textbf{decorrelator} \\ \gamma &\to \infty, && \text{for } \textbf{matched filter} \\ \gamma &= 1, && \text{for } \textbf{LMMSE detector}. \end{aligned}$$

- All with $O(n^3)$ complexity, but sacrifices performance.

Performance degradation of suboptimal schemes



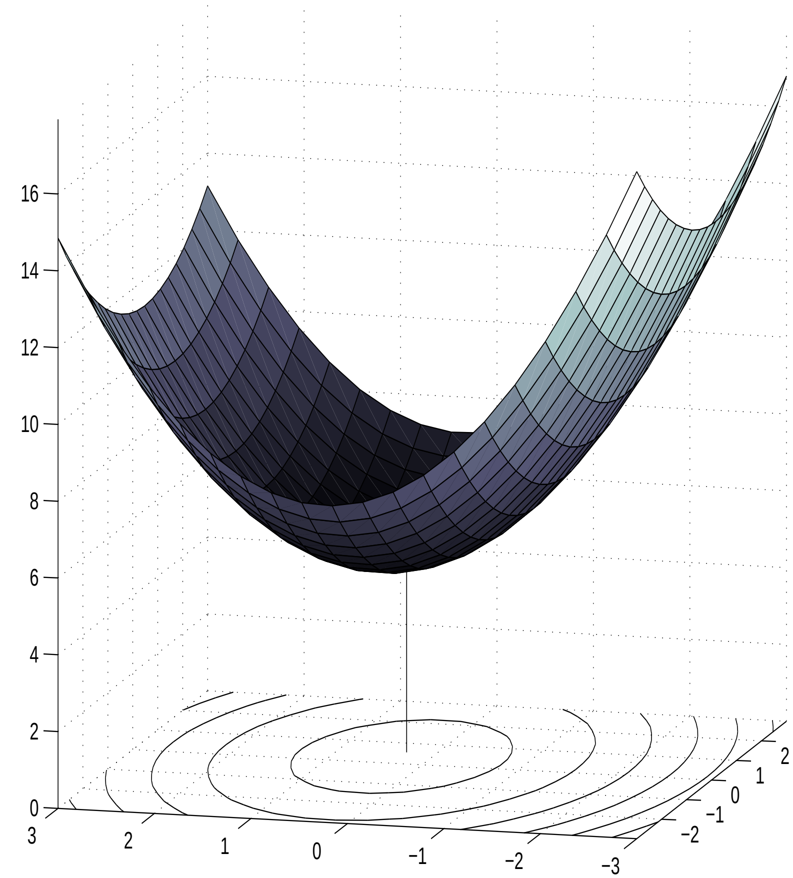
Another suboptimal scheme: Sphere Decoder

Intuition:

- The objective function is convex over $\mathbb{R}^n$.
- Integer solution must be close to the unconstrained minimum:

$$\begin{aligned}\hat{\mathbf{s}} &= \arg\min_{\mathbf{s}\in\mathbb{R}^n} \|\mathbf{y} - \sqrt{\rho/n}\,\mathbf{H}\mathbf{s}\|^2 \\ &= \sqrt{n/\rho}\,(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}\end{aligned}$$

- Idea: localize the exhaustive search to the sphere around zero-forcing solution $\hat{\mathbf{s}}$.
- Two questions:
  - How to choose sphere radius r?
  - How to tell which points are inside the sphere?



Another suboptimal scheme: Sphere Decoder

- Sphere Decoder [Fincke-Pohst, 1985] answers the second question:
  - Let $\sqrt{\rho/n}\,\mathbf{H} = \mathbf{VR}$ (QR-decomposition), then we have

$$\begin{aligned}\|\mathbf{y} - \sqrt{\rho/n}\,\mathbf{Hs}\|^2 &= \frac{\rho}{n}(\mathbf{s}-\hat{\mathbf{s}})^T\mathbf{H}^T\mathbf{H}(\mathbf{s}-\hat{\mathbf{s}}) = (\mathbf{s}-\hat{\mathbf{s}})^T\mathbf{R}^T\mathbf{R}(\mathbf{s}-\hat{\mathbf{s}}) \\ &= \sum_{i=1}^{n} r_{ii}^2(s_i - \hat{s}_i + \sum_{j=i+1}^{n}\frac{r_{ij}}{r_{ii}}(x_j - \hat{x}_j))^2 \le r^2\end{aligned}$$

  - A necessary condition for $\mathbf{s}$ inside the sphere is $r_{nn}^2(s_n - \hat{s}_n)^2 \le r^2$. The remaining coordinates $s_{n-1}, \ldots, s_1$ can be searched iteratively. Back-track if necessary.
- The expected complexity is exponential [Jalden-Ottersten, 2004], despite good empirical performance for small values of n and large ρ.

Sphere Decoder Performance

SD has excellent performance when it works, but can fail in low SNR region or for large systems

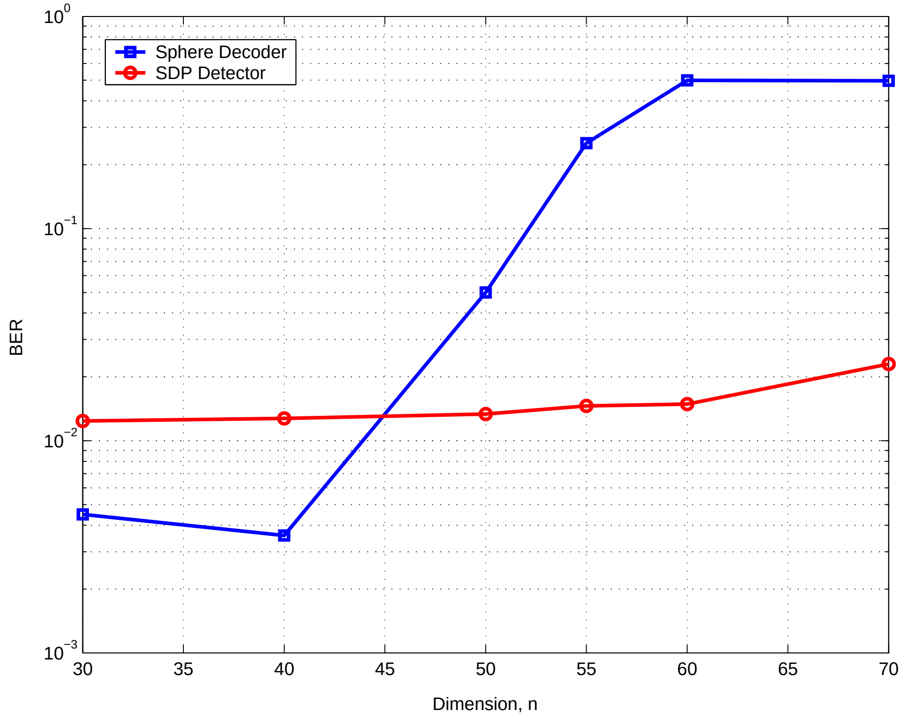


Figure 1: BER degradation due to the limit on detection time. Simulation parameters: BPSK modulation, SNR $= 10$ dB and time limit per bit $= 6.3$ ms.

Dilemma

Existing approaches for MIMO detection

- either has a worst case **polynomial complexity** $\mathcal{O}(n^3)$, at a substantial **performance loss**
- or finds the **ML** solution via heuristic search, but with an **exponential** average complexity
- Questions:
  - is there a polynomial approximation algorithm which offers a constant factor performance guarantee?
  - How about SDP relaxation approach?

A close look at ML detection problem

- Data model: $\mathbf{H} \sim N(\mathbf{0}, \mathbf{I})$; $\mathbf{y} = \sqrt{\frac{\rho}{n}}\mathbf{H}\mathbf{e} + \mathbf{v}$; $\mathbf{v} \sim N(\mathbf{0}, \mathbf{I})$.
- Define matrix $\mathbf{Q}$ and new variable $\mathbf{x}$:

$$\mathbf{Q} = \begin{bmatrix} (\rho/n)\, \mathbf{H}^T\mathbf{H} & -\sqrt{\rho/n}\, \mathbf{H}^T\mathbf{y} \\ -\sqrt{\rho/n}\, \mathbf{y}^T\mathbf{H} & \|\mathbf{y}\|^2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \mathbf{s} \\ 1 \end{bmatrix}$$

- The objective function (negative log likelihood function) can be written as

$$\left\| \mathbf{y} - \sqrt{\rho/n}\, \mathbf{H}\mathbf{s} \right\|^2 = \mathbf{x}^T\mathbf{Q}\mathbf{x}, \quad \Leftarrow \textbf{homogenization}$$

and the ML detection problem becomes

$$\boxed{\begin{array}{rl} f_{ML} = \text{minimize} & \mathbf{x}^H\mathbf{Q}\mathbf{x} \\ \text{subject to} & \mathbf{x} \in \{-1, 1\}^n \end{array}} \quad \Leftarrow \textbf{BQP}\text{: NP-hard}$$

SDP Relaxation of BQP

- Introducing matrix variable $\mathbf{X} = \mathbf{x}\mathbf{x}^H \succeq \mathbf{0}$, then $\mathbf{X}$ is rank 1 and

$$\mathbf{x}^H\mathbf{Q}\mathbf{x} = \mathrm{Tr}(\mathbf{Q}\mathbf{x}\mathbf{x}^H) = \mathrm{Tr}(\mathbf{Q}\mathbf{X}), \quad \text{and} \quad \mathbf{X}_{i,i} = \mathbf{x}_i^2 = 1.$$

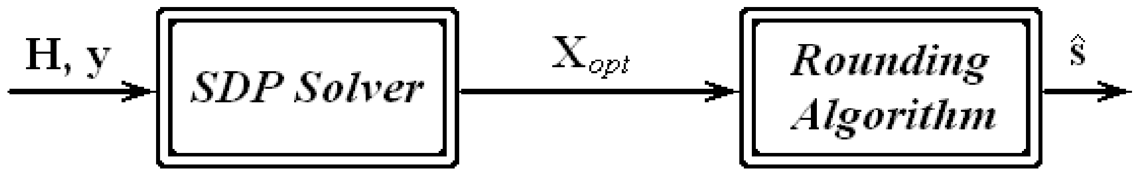
- ML detection problem is equivalent to the problem

$$\boxed{\begin{array}{rl} f_{ML} := \min & \mathrm{Tr}(\mathbf{Q}\mathbf{X}) \\ \text{s.t.} & \mathbf{X} \succeq \mathbf{0}, \ \mathbf{X} \text{ is rank-1} \\ & X_{i,i} = 1, \ i = 1, \ldots, n+1. \end{array}} \quad \Leftarrow \textbf{BQP} \text{ reformulation}$$

$$\Updownarrow$$

$$\boxed{\begin{array}{rl} f_{ML} = \text{minimize} & \mathbf{x}^H\mathbf{Q}\mathbf{x} \\ \text{subject to} & \mathbf{x} \in \{-1, 1\}^n \end{array}} \quad \Leftarrow \textbf{BQP}\text{: NP-hard}$$

SDP Detector, Relaxation

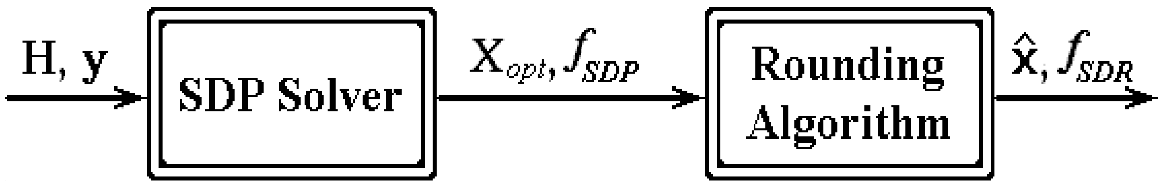


- SDP Detector drops the only non-convex constraint $\mathrm{rank}(\mathbf{X}) = 1$ and solves

$$\boxed{\begin{aligned} f_{SDP} := \min \ & \mathrm{Tr}(\mathbf{QX}) \\ \text{s.t.} \ & \mathbf{X} \succeq 0, \\ & X_{i,i} = 1, \ i = 1, \ldots, n+1. \end{aligned}}$$

- The output $\mathbf{X}_{opt}$ of the SDP problem is no longer rank-1. (But almost rank-1)
- Different randomized rounding procedures can be employed to generate an estimate of transmitted signals $\hat{\mathrm{s}}$, given $\mathbf{X}_{opt}$.
- The complexity of the problem is $\mathcal{O}(n^{3.5})$ (interior point methods).

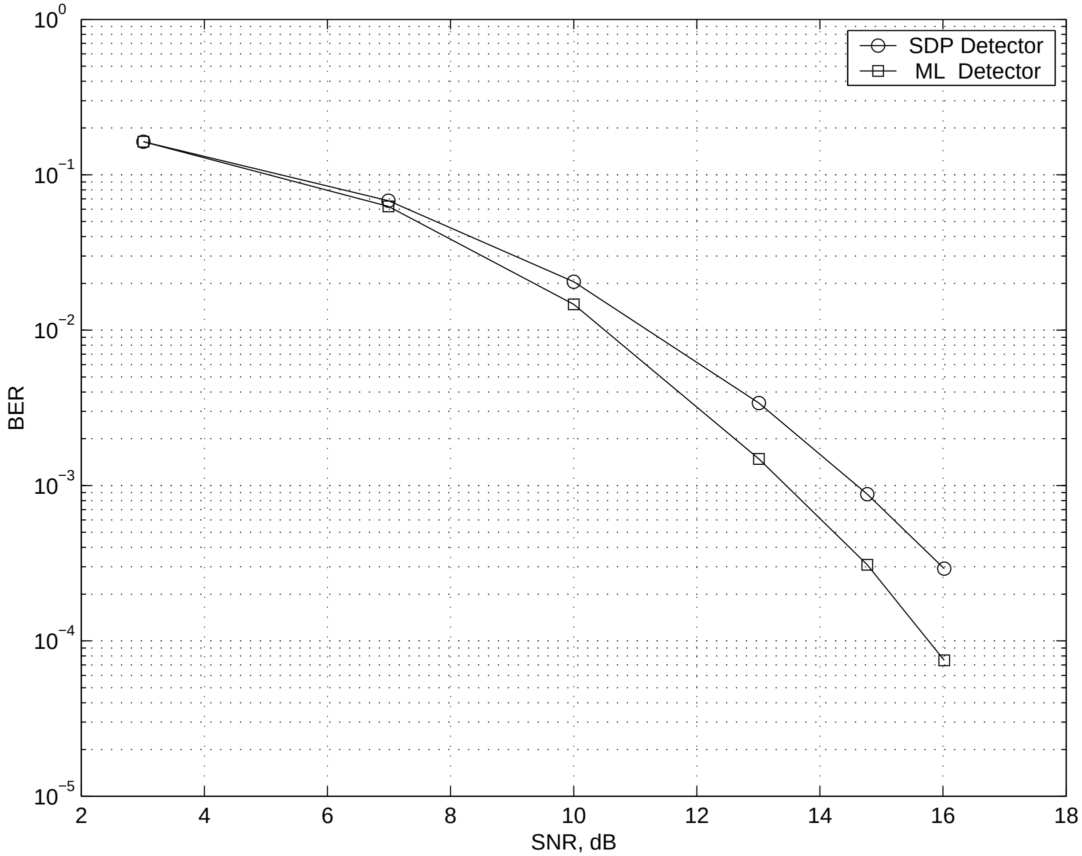
SDP Detector, Randomized Rounding



1. Take a spectral decomposition:
$$\mathbf{X}_{opt} = \sum_{i=1}^{n+1} \lambda_i \mathbf{u}_i \mathbf{u}_i^T.$$

2. Pick eigenvector $\bar{\mathbf{u}}$ that corresponds to the largest eigenvalue $\bar{\lambda}$.
3. Generate a fixed number of vectors $\mathbf{x}$ randomly according to the distribution:
$$\Pr\{x_i = +1\} = \frac{1 + \sqrt{\bar{\lambda}}\ \bar{u}_i}{2}, \quad \Pr\{x_i = -1\} = \frac{1 - \sqrt{\bar{\lambda}}\ \bar{u}_i}{2}.$$
4. Note: $E\{x_i\} = \sqrt{\bar{\lambda}}\ \bar{u}_i$.
5. Output $\hat{\mathbf{x}}$ that achieves the smallest objective value $f_{SDR} = \mathbf{x}^T \mathbf{Q} \mathbf{x}$.

SDP Detector, BER vs SNR, $n = 10$



Key property: for all ρ, performance gap bounded for all and n!

Part VI: Softwares*

* This list is somewhat outdated now. Consult www for the latest versions of software available.

SDPA

- **Authors:** Fujisawa, Kojima, Nakata
- **Version; available:** 5.02, 9/2000; yes (binary only)
- **Key paper:** [43]
  The software manual and the **SDPA** source codes can be downloaded from `http://is-mj.archi.kyoto-u.ac.jp/~fujisawa/research.html`
- **Features:** primal-dual method, users Meschbach library
- **Language, Input format:** C++, SDPA
- **Solves:** SDP
- **Remarks:** numerically robust, good for small size SDPs; parallel versions avaliable.

SeDuMi

- **Authors:** Sturm
- **Version; available:** 1.04, 9/2000; yes
- **Key paper:** [22]
- **Features:** self-dual embedding, dense column handling
- **Language, Input format:** Matlab + C, Matlab, SDPA, SDPpack
- **Solves:** SDP, convex quadratic program, SOCP, linear program
- **Remarks:** benchmark SDP/SOCP solver; now being maintained at McMaster University, Canada; various Matlab interfaces have been built for SeDuMi, including "**convex**" from Stanford University.

CSDP

- **Authors:** Borchers
- **Version; available:** 3.2, 12/15/2000; yes
- **Key paper:** [44]
- **Features:** infesaible predictor-corrector path following interior method
- **Language, Input format:** C; SDPA
- **Solves:** SDP
- **Remarks:** small to medium sized SDP's. No support for SOCP constraints

DSPA

- **Authors:** Benson and Ye
- **Version; available:** 3.2, 11/2000; yes
- **Key paper:** [45]
- **Features:** Dual scaling potential reduction method, Matlab interface, generates primal solution only when requested
- **Language, Input format:** C; SDPA
- **Solves:** SDP
- **Remarks:** good for solving large sparse problems arising from combinatorial optimization

SDPT3

- **Authors:** Toh, Todd, Tutuncu
- **Version; available:** 3.0; yes
- **Key paper:** [46]
- **Features:** infeasible primal-dual and homogeneous self-dual methods, Meschbach library, Lanczos steplength control
- **Language, Input format:** Matlab+C, SDPA
- **Solves:** SDP and SOCP
- **Remarks:** allows Hermitian matrix variables, good for small/medium size problems (with up to around a thousand constraints involving matrices of order up to around a thousand; can also solve some large sparse problems (up to 20,000 constraints and 50,000 variables)

MOSEK

- **Authors:** Andersen
- **Version; available:** 1.4-31; yes (binary)
- **Key paper:** [47]
- **Features:** special SOCP algorithm
- **Language, Input format:** C; QPS, AMPL
- **Solves:** SOCP
- **Remarks:** similar to **SeDuMi** except is 100% C code for speed; callable from Matlab; exploits sparsity and handles fixed and upper bounded variables; does not handle semidefinite matrix cone.

Part V: References*

* Somewhat outdated.

References

[1] Nesterov, Y. and Nemirovskii, A., *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1994.

[2] Ye, Y., *Interior Point Algorithms: Theory and Analysis*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, 1997.

[3] Boyd, S., El Ghaoui, L., Feron, E. and Balakrishnan, V., *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, SIAM, Philadelphia, 1994.

[4] Boyd, S. and Vandenberghe, L., *Convex Optimization,* Cambridge University Press, 2003. [Some of the figures and examples in this tutorial were modelled from this book.]

[5] Mittelmann, H.D., "An Independent Benchmarking of SDP and SOCP Solvers," preprint, Department of Mathematics, Arizona State University, 2001.

[6] Wang, X., Lu, W.S. and Antoniou, A., "Constrained Minimum-BER Multiuser Detection," *IEEE Transactions on Signal Processing,* Vol. 48, No. 10, Oct. 2000, pp. 2903–2909.

[7] Vandenberghe, L. and Balakrishnan, V., "Algorithms And Software Tools For LMI Problems In Control: An Overview," *Proceedings of the 1996 IEEE International Symposium on Computer-Aided Control System Design*, 1996, pp. 229–234.

[8] Vandenberghe, L. and Balakrishnan, V., "Algorithms And Software For LMI Problems In Control," *IEEE Control Systems Magazine*, Vol. 17, No. 5, Oct. 1997, pp. 89–95.

[9] Wu, S.-P., Boyd, S. and Vandenberghe, L., "FIR Filter Design Via Semidefinite Programming And Spectral Factorization," *Proceedings of the 35th IEEE Conference on Decision and Control,* Vol. 1, 1996, pp. 271–276.

[10] Alkire, B. and Vandenberghe, L., "Handling Nonnegative Constraints In Spectral Estimation," *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers,* Vol. 1 , 2000, pp. 202–206.

[11] Balakrishnan, V., Wang, F. and Vandenberghe, L., "Applications Of Semidefinite Programming In Process Control," *Proceedings of the 2000 American Control Conference*, Vol. 5, 2000, pp. 3219–3223.

[12] Vandenberghe, L. and Balakrishnan, V., "Semidefinite Programming Duality And Linear System Theory: Connections And Implications For Computation," *Proceedings of the 38th IEEE Conference on Decision and Control,* Vol. 1, 1999, pp. 989–994.

[13] Xiao, L., Johansson, M., Hindi, H., Boyd, S. and Goldsmith, A., "Joint Optimization Of Communication Rates And Linear Systems", *Proceedings 2001 Conference on Decision and Control,* December 2001,

[14] Kosut, R., Chung, W., Johnson, C. Boyd, S., "On Achieving Reduced Error Propagation Sensitivity In DFE Design Via Convex Optimization" *Proceedings of the 39th IEEE Conference on Decision and Control,* Vol. 5, pp. 4320 - 4323 , December 2000.

[15] Wu, S.-P. and Boyd, S., "SDPSOL: A Parser/Solver For Semidefinite Programs With Matrix Structure", in *Recent Advances in LMI Methods for Control,* L. El Ghaoui and S.-I. Niculescu, editors, SIAM, chapter 4, pp. 79–91, 2000.

[16] Lebret, H. and Boyd, S., "Antenna Array Pattern Synthesis Via Convex Optimization," *IEEE Transactions on Signal Processing,* Vol. 45, No. 3, March 1997, pp. 526–532.

[17] Lobo, M., Vandenberghe, L., Boyd, S. and Lebret, H., “Applications Of Second-Order Cone Programming,” *Linear Algebra and its Applications,* Vol. 284, November 1998, pp. 193–228.

[18] Vandenberghe, L. Boyd, S. and El Gamal, A., “Optimizing Dominant Time Constant In RC Circuits” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* Vol. 17, No. 2, Feb. 1998, pp. 110–125.

[19] Alkire, B. and Vandenberghe, L., “Interior-Point Methods For Magnitude Filter Design,” *Proceedings of 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 6, 2001, pp. 3821–3824.

[20] Dumitrescu, B., Tabus, I. and Stoica, P., “On The Parameterization Of Positive Real Sequences And Ma Parameter Estimation,” *IEEE Transactions on Signal Processing*, Vol. 49, No. 11, Nov. 2001, pp. 2630 -2639.

[21] Stoica, P.; McKelvey, T.; Mari, J., “MA Estimation In Polynomial Time,” *IEEE Transactions on Signal Processing*, Vol. 48, No. 7, July 2000, pp. 1999 -2012.

[22] Sturm, J. F., “Using Sedumi 1.02, A Matlab Toolbox For Optimization Over Symmetric Cones,” *Optim. Meth. Software*, vol. 11–12, pp. 625–653, 1999. See `http://fewcal.kub.nl/sturm/software/sedumi.html` for updates.

[23] Vandenberghe, L. and Boyd, S., “Semidefinite Programming”, *SIAM Review*, vol. 31, pp. 49–95, 1996.

[24] Yu, W., Rhee, W. and Cioffi, J., “Multiuser Transmitter Optimization for Vector Multiple Access Channels”, manuscript, STAR laboratory, Department of Electrical Engineering, Stanford University, CA 94305-9515.

[25] Pesavento, M., Gershman, A. and Luo, Z.-Q., "Robust Array Interpolation Using Second-Order Cone Programming," Accepted for publication in *IEEE Signal Processing Letters*.

[26] Ma, W.-K., Davidson, T.N., Wong, K.M., Luo, Z.-Q. and Ching, P.-C., "Quasi-maximum-likelihood Multiuser Detection Using Semi-definite Relaxation," Accepted for publication in *IEEE Transactions on Signal Processing*.

[27] Luo, Z.-Q., Meng, M., Wong, K.M. and Zhang, J.-K., "A Fractionally Spaced Blind Equalizer Based on Linear Programming," Accepted for publication in *IEEE Transaction on Signal Processing*.

[28] Li, L., Luo, Z.-Q., Wong, K.M. and Bossé, E., "Semidefinite Programming Solutions to the Robust State Estimation Problem", Accepted for publication in *SIAM Journal on Optimization*.

[29] Afkhamie, K.H., Luo, Z.-Q. and Wong, K.M., "MMSE Decision-Feedback Equalization with Short Training Sequences: An Application of Interior Point Least Squares," Accepted for publication in *IEEE Transactions on Signal Processing*.

[30] Li, L., Luo, Z.-Q., Wong, K.M. and Bossé, E. "Convex Optimization Approach to Identity Fusion For Multi-Sensor Target Tracking," *IEEE Transaction on Systems, Man and Cybernetics*, Part A: Systems and Humans, Vol. 31, No. 3, pp. 172–178, 2001.

[31] Fu, M., Souza, C. and Luo, Z.-Q., "Finite Horizon Robust Kalman Filter Design," Accepted for publication in *IEEE Transactions on Signal Processing*.

[32] Ding, Z. and Luo, Z.-Q., "A Fast Linear Programming Algorithm for Blind Equalization," *IEEE Transactions on Communications*, September, 2000.

[33] Afkhamie, K.H. and Luo, Z.-Q., "Blind Identification of FIR Systems Driven by Markov-Like Input Signals," *IEEE Transactions on Signal Processing*, Vol. 48, No. 6, pp. 1726–1736, 2000.

[34] Davidson, T.N., Luo, Z.-Q. and Wong, K.M., "Spectrally-Efficient Orthogonal Pulse Shape Design via Semidefinite Programming," *IEEE Transactions on Signal Processing*, Vol. 48, No. 5, pp. 1433–1445, 2000.

[35] Afkhamie, K., Luo, Z.-Q. and Wong, K.M., "Adaptive Linear Filtering Using Interior Point Optimization Techniques," *IEEE Tran. on Signal Processing*, Vol. 48, No. 6, pp. 1637–1648, 2000.

[36] Li, X.-L., Luo, Z.-Q., Wong, K.M. and Bosse, E., "An Interior Point Linear Programming Approach to Two-Scan Data Association," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 35, No. 2, pp. 474–490, 1999.

[37] Shahbazpanahi, S., Gershman, A., Luo, Z.-Q. and Wong, K.M., "Robust Adaptive Beamforming For General-Rank Signal Models," manuscript, Department of Electrical and Computer Engineering, McMaster University, February 2002.

[38] Maricic, B., Luo, Z.-Q. and Davidson, T.N., "Blind Constant Modulus Equalization via Convex Optimization," manuscript, Department of Electrical and Computer Engineering, McMaster University, April 2001.

[39] Luo, Z.-Q., Davidson, T.N., Giannakis, G.B. and Wong, K.M., "Transceiver Optimization for Multiple Access through ISI Channels", manuscript, Department of Electrical and Computer Engineering, McMaster University, September 2001.

[40] Davidson, T.N., Luo, Z.-Q. and Sturm, J., "Linear Matrix Inequality Formulation of Spectral Mask Constraints," manuscript, Department of Electrical and Computer Engineering, McMaster University, 2001.

[41] Cui, S., Luo, Z.-Q. and Ding, Z., "Robust Blind Multiuser Detection against Signature Waveform

Mismatch," manuscript, Department of Electrical and Computer Engineering, McMaster University, 2001.

[42] Vorobyov, S., Gershman, A. and Luo, Z.-Q., "Robust Adaptive Beamforming Using Worst-Case Performance Optimization," *Proceedings of 2002 ICASSP*, Orlando, Florida, 2002.

[43] Fujisawa, K., Kojima, M. and Nakata, K., "SDPA (Semidefinite Programming Algorithm) – User's Manual –," Technical Report B-308, Tokyo Institute of Technology, Revised on August 2000.

[44] Borchers, B., "CSDP, A C Library for Semidefinite Programming," *Optimization Methods and Software*, Vol. 11, 1999, pp. 613–623.

[45] Benson, S.J. and Ye, Y., "DSDP3: Dual Scaling Algorithm for General Positive Semidefinite Programming," preprint ANL/MCS-P851-1000, Argonne National Labs.

[46] Toh, K.C., Todd, M.J., and Tutuncu, R.H., "SDPT3 – a Matlab Software Packgage for Semidefinite Programming, version 3.0," Technical Report, National University of Singapore, June 2001.

[47] Andersen, E.D., Roos, C. and Terlaky, T., "On Implementing a Primal-Dual Interior-Point Method for Conic Quadratic Optimization," Report W-274, Helsinki School of Economics and Business Administration, Department of Economics and Management Science, Helsinki, Finland, December 2000.

Acknowledgement

- Graduate students: Afkhamie, Cui, Dam, Meng, Li, Liu, Lu, Maricic, Saad, Tan, Yu, Kisialiou, Xiao,...
- Postdocs: Davidson, Gao, Jin, Li, Sturm, Wu, Zhang, Vorobyov...
- Colleagues: Wong, Davidson, Gershman
- Industrial sponsors: Nortel/Mitel/Ontario Hydro/...
- Federal/provincial support: NSF/ARO/DREV/DREO/NSERC/CITO...

Thank You!

`http://www.ece.umn.edu/users/luozq`