Primal Decomposition for Multicarrier Linear MIMO Transceivers

Daniel P. Palomar

Hong Kong University of Science and Technology (HKUST)

ELEC692Q - Convex Optimization
Fall 2007-08, HKUST, Hong Kong
November 19, 2007
Outline of Lecture

• Signal Model: MIMO Channels and Linear MIMO Transceivers
• A Review of Design of the Single-Carrier Linear MIMO Transceiver
• Design of Multicarrier Linear MIMO Transceivers
• Numerical Results
• Conclusions
Signal Model: Single MIMO Channel

- Signal model:

\[ y = Hs + n \]
Signal Model: Multicarrier MIMO Channel

- Signal model:

\[ y_k = H_k s_k + n_k \quad 1 \leq k \leq N \]
Signal Model: Examples of MIMO Channels

• Multiantenna wireless channel

\[ s(t) \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_{n_T}(t) \end{bmatrix} \xrightarrow{\text{Scattering medium}} y(t) \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n_R}(t) \end{bmatrix} \]

\[ s(t) = \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_{n_T}(t) \end{bmatrix} \sim \text{Scattering medium} \to \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n_R}(t) \end{bmatrix} \]

• Matrix convolution signal model

\[
y(n) = \sum_{k=0}^{L} H(k) s(n-k) + n(n)
\]
Signal Model: Examples of MIMO Channels

- DSL channel (bundle of twisted-pair copper wires)

- Matrix convolution signal model

\[ y(n) = \sum_{k=0}^{L} H(k) s(n - k) + n(n) \]
Rewriting the matrix convolution

\[
y(n) = \sum_{k=0}^{L} H(k) s(n-k) + n(n)
\]

Convolution matrix approach

\[
\begin{bmatrix}
    \vdots \\
y(-2) \\
y(-1) \\
y(0) \\
y(1) \\
y(2) \\
    \vdots
\end{bmatrix}
= 
\begin{bmatrix}
    \vdots \\
    H(L) \cdots H(0) \\
    H(L) \cdots H(0) \\
    H(L) \cdots H(0) \\
    H(L) \cdots H(0) \\
    \vdots
\end{bmatrix}
\cdot
\begin{bmatrix}
    \vdots \\
s(-2) \\
s(-1) \\
s(0) \\
s(1) \\
s(2) \\
    \vdots
\end{bmatrix}
\]

Multicarrier Approach

\[
y_k = H_k s_k + n_k \quad 1 \leq k \leq N
\]
**Signal Model: Linear MIMO Transceiver (I)**

- Scheme of a single linear MIMO transceiver:

  ![Diagram of a single linear MIMO transceiver]

  - Interpretation of the beam-matrix as multiple beamvectors:

    ![Diagram showing multiple beamvectors]

Daniel P. Palomar
• Transmitted signal using a beam-matrix:

\[ s = Bx \]

with total average power \( P_T = \text{Tr}(BB^H) \).

• Received signal post-processed with a beam-matrix:

\[ \hat{x} = A^H y. \]

• Global signal model (several substreams established):

\[ \hat{x}_i = a_i^H (Hb_i x_i + n_i) \quad 1 \leq i \leq L. \]
Signal Model: Multicarrier Linear MIMO Transceiver (I)

• Scheme of a linear MIMO transceiver:

\[ \mathbf{x}_1 \rightarrow \mathbf{B}_k \rightarrow \mathbf{H}_k \rightarrow \mathbf{y}_k \rightarrow \mathbf{A}_k^H \rightarrow \hat{\mathbf{x}}_k \]

\[ \vdots \]

\[ \mathbf{x}_k \rightarrow \mathbf{B}_k \rightarrow \mathbf{H}_k \rightarrow \mathbf{y}_k \rightarrow \mathbf{A}_k^H \rightarrow \hat{\mathbf{x}}_k \]

\[ \vdots \]

\[ \mathbf{x}_N \rightarrow \mathbf{B}_k \rightarrow \mathbf{H}_k \rightarrow \mathbf{y}_k \rightarrow \mathbf{A}_k^H \rightarrow \hat{\mathbf{x}}_N \]
Signal Model: Multicarrier Linear MIMO Transceiver (II)

- Transmitted signal using a beam-matrices:
  \[ s_k = B_k x_k \quad 1 \leq k \leq N \]
  with total average power \[ P_T = \sum_{k=1}^{N} \text{Tr} (B_k B_k^H) \].

- Received signal post-processed with a beam-matrix
  \[ \hat{x}_k = A_k^H y_k \quad 1 \leq k \leq N. \]

- Signal model for the \( k \)th MIMO channel and \( i \)th substream:
  \[ \hat{x}_{k,i} = a_{k,i}^H (H_k b_{k,i} x_{k,i} + n_{k,i}) \).
The performance of a link is commonly measured in terms of MSE, SINR, or BER:

\begin{align*}
\text{MSE}_{k,i} & \triangleq \mathbb{E}[|\hat{x}_{k,i} - x_{k,i}|^2] = |a_{k,i}^H H_k b_{k,i} - 1|^2 + a_{k,i}^H R_{n_{k,i}} a_{k,i} \\
\text{SINR}_{k,i} & \triangleq \frac{\text{desired component}}{\text{undesired component}} = \frac{|a_{k,i}^H H_k b_{k,i}|^2}{a_{k,i}^H R_{n_{k,i}} a_{k,i}} \\
\text{BER}_{k,i} & \triangleq \frac{\# \text{ bits in error}}{\# \text{ transmitted bits}} \approx \tilde{g}_{k,i} (\text{SINR}_{k,i})
\end{align*}
It is notationally convenient to define the $k$th MSE matrix as (natural extension of the scalar MSE):

$$
E_k \triangleq \mathbb{E}\left[ (\hat{x}_k - x_k)(\hat{x}_k - x_k)^H \right]
= (A_k^H H_k B_k - I) (B_k^H H_k^H A_k - I) + A_k^H R_{n_k} A_k
$$

where

$$
\text{MSE}_{k,i} = [E_k]_{ii}.
$$
Design of Single-Carrier MIMO Transceiver

- **General problem formulation**: design system by optimizing some arbitrary measure of quality of the system performance:
  - In terms of MSE:
    \[
    \begin{align*}
    \text{minimize} \quad & f_0 \left( \{ \text{MSE}_i \} \right) \\
    \text{subject to} \quad & \text{Tr} \left( BB^H \right) \leq P_T
    \end{align*}
    \]
    where \( f_0 \) is an arbitrary cost function and
    \[
    \text{MSE}_i = \left[ (A^H H B - I) \left( B^H H^H A - I \right) + A^H R_n A \right]_{ii}.
    \]
  - In terms of SINR:
    \[
    f_0 \left( \{ \text{SINR}_i \} \right)
    \]
  - In terms of BER:
    \[
    f_0 \left( \{ \text{BER}_i \} \right)
    \]
Joint minimization equals nested minimization (even for nonconvex functions):

$$\min_{x,y} f(x,y) = \min_x \left\{ \min_y f(x,y) \right\}.$$ 

In other words,

$$\min_{x,y} f(x,y) = \min_x \tilde{f}(x)$$

where

$$\tilde{f}(x) = \min_y f(x,y).$$
Single Linear Transceiver: Optimal Receiver

- The optimum receiver is the Wiener filter

\[ A = \left( HBB^H H^H + R_n \right)^{-1} HB \]

regardless of \( f_0 \) because the design of each column of \( A, a_i \), decouples (recall that \( \hat{x}_i = a_i^H (Hb_i x_i + n_i) \)).

- All the MSEs, SINRs, and BERs are simultaneously optimized.

- The MSE matrix reduces to

\[ E = \left( I + B^H R_H B \right)^{-1} \]

where \( R_H \triangleq H^H R_n^{-1} H \).
Single Linear Transceiver: Optimal Transmitter (I)

- Since the SINRs and BERs are directly related to the MSEs, we can consider w.l.o.g. the problem formulation as a function of the MSEs:

\[
\begin{align*}
\text{minimize} & \quad f_0 \left( \left[ \left( \mathbf{I} + \mathbf{B}^H \mathbf{R}_H \mathbf{B} \right)^{-1} \right]_{ii} \right) \\
\text{subject to} & \quad \text{Tr} \left( \mathbf{B} \mathbf{B}^H \right) \leq P_T.
\end{align*}
\]

- This is still a terrible nonconvex problem !!!

- Even in the scalar case, \( \frac{1}{1+|b|^2 r} \) is nonconvex in \( b \).

- The problem needs to be simplified somehow ... *majorization theory + convex optimization theory.*
Theorem 1 [PalCioLag’03]: Consider the optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0 \left( \left\{ \left[ (I + B^H R H B)^{-1} \right]_{ii} \right\} \right) \\
\text{subject to} & \quad \text{Tr} \left( BB^H \right) \leq P_T.
\end{align*}
\]

- If \( f_0 \) is Schur-concave, then the optimal solution diagonalizes the channel and the MSE matrix:

\[
B = U_H \Sigma_B.
\]

- If \( f_0 \) is Schur-convex, then the optimal solution diagonalizes the channel up to a rotation and the MSE matrix has equal diagonal elements (non-diagonal):

\[
B = U_H \Sigma_B Q^H.
\]
• Scalarized MSE ($B = U_H \Sigma_B$):

$$\text{MSE}_i = \left[ (I + B^H R_H B)^{-1} \right]_{ii} = \frac{1}{1 + p_i \lambda_{H,i}} \quad 1 \leq i \leq L.$$  

• Problem formulation

$$\begin{align*}
\underset{p}{\text{minimize}} & \quad f_0 \left( \left\{ \frac{1}{1 + p_i \lambda_{H,i}} \right\} \right) \\
\text{subject to} & \quad \sum_{j=1}^{L} p_j \leq P_T \\
& \quad p_i \geq 0 \quad 1 \leq i \leq L.
\end{align*}$$

• The solution depends on the specific choice of $f_0$. 
• Minimization of the sum of the MSEs or, equivalently, of $\text{Tr} \left( \mathbf{E} \right)$ with solution $p_i = \left( \mu \lambda_{H,i}^{-1/2} - \lambda_{H,i}^{-1} \right)^+$.  

• Minimization of the weighted sum of the MSEs or, equivalently, of $\text{Tr} \left( \mathbf{W} \mathbf{E} \right)$, where $\mathbf{W} = \text{diag} \left( \{ w_i \} \right)$ is a diagonal weighting matrix, with solution $p_i = \left( \mu w_i^{1/2} \lambda_{H,i}^{-1/2} - \lambda_{H,i}^{-1} \right)^+$.  

• Minimization of the (exponentially weighted) product of the MSEs with solution $p_i = \left( \mu w_i - \lambda_{H,i}^{-1} \right)^+$.  

• Minimization of $|\mathbf{E}|$ with solution $p_i = \left( \mu - \lambda_{H,i}^{-1} \right)^+$.  

Daniel P. Palomar
List of Schur-Concave Functions (II)

- Maximization of the mutual information, with solution $p_i = (\mu - \lambda_{H,i}^{-1})^+$. 

- Maximization of the (weighted) sum of the SINRs with solution given by allocating all the power on the channel eigenmode with highest weighted gain $w_i \lambda_{H,i}$. 

- Maximization of the (exponentially weighted) product of the SINRs with solution $p_i = P_0 w_i / \sum_j w_j$ (for the unweighted case, it results in a uniform power allocation).
Single Linear Transceiver: Schur-Convex Functions

- Scalarized MSE ($\mathbf{B} = \mathbf{U}_H \mathbf{\Sigma}_B \mathbf{Q}^H$):

\[
\text{MSE}_i = \frac{1}{L} \text{Tr} \left( (\mathbf{I} + \mathbf{B}^H \mathbf{R}_H \mathbf{B})^{-1} \right)
= \frac{1}{L} \sum_{j=1}^{L} \frac{1}{1 + p_j \lambda_{H,j}} \quad 1 \leq i \leq L.
\]

- Problem formulation

\[
\begin{align*}
\text{minimize} \quad & f_0 \left( \left\{ \frac{1}{L} \sum_{j=1}^{L} \frac{1}{1 + p_j \lambda_{H,j}} \right\}_{i=1}^{L} \right) \\
\text{subject to} \quad & \sum_{j=1}^{L} p_j \leq P_T \\
& p_i \geq 0 \quad 1 \leq i \leq L.
\end{align*}
\]
• Simplified problem formulation (independent of $f_0$)

$$\begin{align*}
\text{minimize} & \quad \frac{1}{L} \sum_{j=1}^{L} \frac{1}{1 + p_j \lambda_{H,j}} \\
\text{subject to} & \quad \sum_{j=1}^{L} p_j \leq P_T \\
& \quad p_i \geq 0 \quad 1 \leq i \leq L.
\end{align*}$$

• Solution given by the following waterfilling (plus the rotation):

$$p_i = \left( \mu \lambda_{H,i}^{-1/2} - \lambda_{H,i}^{-1} \right)^+ \quad 1 \leq i \leq L.$$
List of Schur-Convex Functions

- Minimization of the maximum of the MSEs.
- Maximization of the minimum of the SINRs.
- Maximization of the harmonic mean of the SINRs.
- Minimization of the average BER (with equal constellations).
- Minimization of the maximum of the BERs.

⇒ Solution always given by the following waterfilling (plus the rotation):

\[ p_i = \left( \mu \lambda_{H,i}^{-1/2} - \lambda_{H,i}^{-1} \right)^+ . \]
• General problem formulation: design system by optimizing some arbitrary measure of quality of the system performance:

\[
\begin{align*}
\text{minimize} & \quad f_0(\alpha_1, \cdots, \alpha_N) \\
\text{subject to} & \quad \alpha_k = f_k\left(\{\text{MSE}_{k,i}\}_{i=1}^{L_k}\right) \quad 1 \leq k \leq N \\
& \quad \sum_{k=1}^{N} \text{Tr}\left(B_k B_k^H\right) \leq P_T,
\end{align*}
\]

where \(f_0\) and the \(f_k\)'s are arbitrary cost functions and

\[
\text{MSE}_{k,i} = \left[\left( A_k^H H_k B_k - I \right) \left( B_k^H H_k^H A_k - I \right) + A_k^H R_{n_k} A_k\right]_{ii}.
\]
The optimum receiver is again the Wiener filter

\[
A_k = (H_k B_k B_k^H H_k^H + R_{n,k})^{-1} H_k B_k
\]

regardless of \( f_0 \) and the \( f_k \)'s.

All the MSEs, SINRs, and BERs are simultaneously optimized.

The MSE matrices reduce to

\[
E_k = (I + B_k^H R_{H,k} B_k)^{-1}
\]

where \( R_{H,k} \triangleq H_k^H R_{n,k}^{-1} H_k \).
Multicarrier MIMO Transceivers: Optimal Transmitter (I)

- Using the expression for the optimum receiver (Wiener filter) the problem becomes:

\[
\begin{align*}
\text{minimize} & \quad f_0 (\alpha_1, \ldots, \alpha_N) \\
\text{subject to} & \quad \alpha_k = f_k \left( \left\{ (I + B_k^H R_{H_k} B_k)^{-1} \right\}_{ii}^{L_k} \right), \quad \forall k \\
& \quad \sum_{k=1}^{N} \text{Tr} \left( B_k B_k^H \right) \leq P_T.
\end{align*}
\]

- This problem is even more complicated than the one for the single-carrier system and, of course, it is nonconvex.

- The problem needs to be simplified somehow ...
Multicarrier MIMO Transceivers: Optimal Transmitter (II)

- Brute-force approach [Pal-PhD’03]:
  1) use majorization theory to simplify the problem as in the single-carrier case
  2) try to solve the simplified problem using convex optimization theory (KKT conditions) ... can be nontrivial ...

- More elegant approach [Pal’05]:
  1) use a decomposition technique to simplify the problem
  2) borrow directly the solution from the single-carrier problem + solve a master problem.
Example: minimization of maximum of the MSEs

- Problem formulation:

\[
\begin{align*}
\text{minimize} & \quad t \\
\text{subject to} & \quad t \geq \frac{1}{L_k} \sum_{i=1}^{L_k} \frac{1}{1+\lambda_k,i p_{k,i}} \quad 1 \leq k \leq N, \\
& \quad \sum_{k,i} p_{k,i} \leq P_T, \\
& \quad p_{k,i} \geq 0 \quad 1 \leq k \leq N, 1 \leq i \leq L_k.
\end{align*}
\]

- Waterfilling solution with multiple waterlevels:

\[
p_{k,i} = \left( \mu_k \lambda_k^{-1/2} - \lambda_k^{-1} \right)^+
\]

where the waterlevels need to satisfy the conditions:

\[
\begin{align*}
\text{minimize} & \quad t \\
\text{subject to} & \quad t = \frac{1}{L_k} \sum_{i=1}^{L_k} \frac{1}{1+\lambda_k,i p_{k,i}} \quad 1 \leq k \leq N, \\
& \quad \sum_{k,i} p_{k,i} = P_T.
\end{align*}
\]
Example: minimization of harmonic mean of SINRs

• Problem formulation:

\[
\begin{align*}
\text{minimize} & \quad \sum_{k=1}^{N} \frac{t_k}{L_k-t_k} \\
\text{subject to} & \quad L_k > t_k \geq \sum_{i=1}^{L_k} \frac{1}{1+\lambda_{k,i}p_{k,i}} \quad 1 \leq k \leq N, \\
& \quad \sum_{k,i} p_{k,i} \leq P_T, \\
& \quad p_{k,i} \geq 0 \quad 1 \leq k \leq N, 1 \leq i \leq L_k.
\end{align*}
\]

• Waterfilling solution with multiple waterlevels:

\[
p_{k,i} = \left( \mu_k \lambda_{k,i}^{-1/2} - \lambda_{k,i}^{-1} \right)^+ 
\]

where the waterlevels need to satisfy the conditions:

\[
\begin{align*}
t_k & = \sum_{i=1}^{L_k} \frac{1}{1+\lambda_{k,i}p_{k,i}} \quad 1 \leq k \leq N, \\
\mu_k & = \omega \frac{L_k^{1/2}}{L_k-t_k}, \\
\sum_{k,i} p_{k,i} & = P_T.
\end{align*}
\]
Interlude: Evaluation of Waterfilling Solutions (I)

- Waterfillings with a single waterlevel $\mu$:

$$ p_i = (\mu - \lambda_i^{-1})^+ \quad 1 \leq i \leq L $$

can be easily evaluated numerically by bisection (approximate sol.) or hypothesis testing (exact sol.).

- Hypothesis testing forms all possible hypotheses and chooses the right one:

  i) the initial worst-case complexity corresponds to evaluating $2^L$ hypotheses (active/inactive substreams)

  ii) assuming eigenvalues in decreasing order, the complexity reduces to $L$. 

Daniel P. Palomar
Interlude: Evaluation of Waterfilling Solutions (II)

• However, for waterfillings with multiple waterlevels

\[ p_{k,i} = \left( \mu_k - \lambda_{k,i}^{-1} \right)^+ , \quad 1 \leq k \leq N, \ 1 \leq i \leq L \]

finding the waterlevels is in principle a formidable problem:

i) the initial worst-case complexity corresponds to evaluating \(2^{NL}\) hypotheses

ii) assuming eigenvalues in decreasing order at each carrier, the complexity reduces to \(L^N\)

iii) neat (nontrivial) simplification [Pal-PhD’03] [PalFon’05]: the complexity is actually \(NL\), which is polynomial!!
Example of algorithm for maximization of harmonic mean of SINRs:

1. Set $\omega_{\text{max-ub}} = \max_{1 \leq k \leq N} \left\{ \frac{1}{L_k^{1/2}} \left( \widetilde{L}_k \lambda_{k,1}^{-1/2} - \sum_{i=1}^{\widetilde{L}_k} \lambda_{k,i}^{-1/2} \right) \right\}$.

2. If $\omega_{\text{max-lb}} < \frac{PT + \sum_{k=1}^{N} \left( \sum_{i=1}^{\widetilde{L}_k} \lambda_{k,i}^{-1} - \frac{1}{\widetilde{L}_k} \left( \sum_{i=1}^{\widetilde{L}_k} \lambda_{k,i}^{-1/2} \right)^2 \right)}{\sum_{k=1}^{N} \frac{L_k^{1/2}}{L_k} \left( \sum_{i=1}^{\widetilde{L}_k} \lambda_{k,i}^{-1/2} \right)}$, then accept the hypothesis $\{ \widetilde{L}_k \}$ and go to step 3. Otherwise reject the hypothesis, set $\widetilde{L}_{k_{\text{max}}} = \widetilde{L}_{k_{\text{max}}} - 1$, and go to step 1.

3. Obtain the definitive waterlevels and power allocation.
Now, recall the problem we want to solve:

\[
\begin{align*}
\text{minimize} & \quad f_0 (\alpha_1, \cdots, \alpha_N) \\
\text{subject to} & \quad f_k \left( \left\{ \left[ (I + B_k^H R_{H_k} B_k)^{-1} \right]_{ii} \right\}_{i=1}^{L_k} \right) \leq \alpha_k \quad \forall k \\
& \quad \sum_{k=1}^{N} \text{Tr} (B_k B_k^H) \leq P_T \\
& \quad P_k \geq 0.
\end{align*}
\]

Let’s try to solve it in a more elegant and simple way.
Multicarrier MIMO Transceivers: Optimal Transmitter (IV)

- Let’s introduce additional variables $P_1, \cdots, P_N$ to denote the power allocation among the carriers:

$$\begin{align*}
\text{minimize} \quad & f_0(\alpha_1, \cdots, \alpha_N) \\
\text{subject to} \quad & f_k \left( \left\{ \left[ (I + B_k^H R_{H_k} B_k)^{-1} \right]_{ii} \right\}_{i=1}^{L_k} \right) \leq \alpha_k \quad \forall k \\
& \text{Tr} (B_k B_k^H) \leq P_k \\
& \sum_{k=1}^{N} P_k \leq P_T \\
& P_k \geq 0.
\end{align*}$$

- Now it is clear that if we fix the power allocation $P_1, \cdots, P_N$ the problem decouples into a set of $N$ independent problems identical to the previous single-carrier problem.
Multicarrier MIMO Transceivers: Optimal Transmitter (V)

• The idea is to use

\[
\minimize \{B_k, \alpha_k, P_k\} \equiv \minimize \{P_k\} \minimize \{B_k, \alpha_k\}
\]

which is the essence of the primal decomposition.

• The inner minimization decouples into \(N\) independent minimizations (subproblems) with a fixed power allocation which can be solved in parallel.

• The outer minimization (master problem) is in charge of updating the power allocation (it needs the subgradients of the subproblems).
Characterization of the Subproblems

- For a fixed power allocation $P_1, \cdots, P_N$ each subproblem is

  \[
  \begin{align*}
  \text{minimize} & \quad f_k \left( \left\{ \left[ (I + B_k^H R_{H_k} B_k)^{-1} \right]_{ii} \right\}_{i=1}^{L_k} \right) \\
  \text{subject to} & \quad \text{Tr} \left( B_k B_k^H \right) \leq P_k.
  \end{align*}
  \]

- The $N$ problems can be solved in parallel using the previous result for Schur-concave/convex cost functions.

- Let $f_k^*(P_k)$ denote the optimal cost value of the problem for a given power $P_k$. Now we have to properly characterize $f_k^*(P_k)$ for the master problem.
Characterization of the Subproblems for Schur-Concave Functions

(a) The function $f_k^*(P_k)$ is convex (assuming $f_k(\rho_k)$ convex).

(b) A subgradient of $f_k^*(P_k)$ at $P_k$ is $-\mu_k$, where $\mu_k$ is the optimal Lagrange multiplier of the subproblem associated to the power constraint $\text{Tr} \left( B_k B_k^H \right) \leq P_k$.

(c) The function $f_k^*(P_k)$ is differentiable if $f_k(\rho_k)$ is differentiable (and then $\nabla f_k^*(P_k) = -\mu_k$).
Characterization of the Subproblems for Schur-Convex Functions

(a) The function $f_k^*(P_k)$ is convex (assuming $f_k(\rho_k)$ convex).

(b) A subgradient of $f_k^*(P_k)$ at $P_k$ can be readily obtained from the optimal Lagrange multiplier $\mu_k$ of the subproblem associated to the power constraint $\text{Tr} \left( B_k B_k^H \right) \leq P_k$.

(c) The function $f_k^*(P_k)$ is differentiable if $f_k(\rho_k)$ is differentiable (and then $\nabla f_k^*(P_k) = -\mu_k \frac{1}{L_k} \sum_{i=1}^{L} \frac{\partial f_k}{\partial \rho_k,i}$).
Characterization of the Subproblems for Schur-Convex Functions: Examples (I)

- Maximum of MSEs: \( f(\rho) = \max_i \{\rho_i\} \) (not differentiable). The function \( f^*(P) = \rho \) is differentiable with gradient

\[
\nabla f^*(P) = -\frac{\mu}{L}.
\]

- Harmonic mean of SINRs: the function to be minimized (inverse of the harmonic mean) is \( f(\rho) = \sum_i (\rho_i^{-1} - 1)^{-1} \). The function \( f^*(P) = \frac{L}{\rho^{-1} - 1} \) is differentiable with gradient

\[
\nabla f^*(P) = -\frac{\mu}{(1 - \rho)^2}.
\]
• Average BER (assuming equal constellations, i.e., $g_i = g$ for all $i$): $f(\rho) = \frac{1}{L} \sum_{i=1}^{L} g(\rho_i)$. The function $f^*(P) = g(\rho)$ is differentiable with gradient

$$
\nabla f^*(P) = -\frac{\mu}{L} g'(\rho).
$$
Primal Decomposition of the Problem

- Recall the problem we want to solve:

\[
\min_{\{P_k\}} \min_{\{B_k, \alpha_k\}} \quad f_0(\alpha_1, \cdots, \alpha_N)
\]

subject to

\[
f_k \left( \left\{ \left[ (I + B_k^H R_{H_k} B_k)^{-1} \right]_{ii} \right\}_{i=1}^{L_k} \right) \leq \alpha_k \quad \forall k
\]

\[
\text{Tr} \left( B_k B_k^H \right) \leq P_k
\]

\[
\sum_{k=1}^{N} P_k \leq P_T
\]

\[
P_k \geq 0.
\]

- The inner minimization is characterized by the convex (maybe differentiable) functions \(f_k^*(P_k)\).
Characterization of the Master Problem

- Using $f^*_k(P_k)$ the problem simplifies to

$$\begin{align*}
\text{minimize} & \quad f_0(P_1, \cdots, P_N) \\
\text{subject to} & \quad \sum_{k=1}^{N} P_k \leq P_0 \\
& \quad P_k \geq 0, \quad 1 \leq k \leq N
\end{align*}$$

where $f_0(P_1, \cdots, P_N) \triangleq f_0(f^*_1(P_1), \cdots, f^*_N(P_N))$.

- The problem is convex as long as $f_0$ is convex (recall that each $f^*_k(P_k)$ is convex).

- Additional constraints can be included, for example, in the form of spectral masks: $P_k \leq P_k^{\max}$.
Primal Decomposition of the Problem

- The original problem has been effectively decomposed into several subproblems controlled by a master problem:

Original Problem \[\xrightarrow{\text{Decomposition}}\]

Subproblem 1
\[\alpha_1 = f_1(P_1)\]

\[\vdots\]

Master Problem
\[f_0(\alpha_1, \ldots, \alpha_N)\]

Subproblem \(N\)
\[\alpha_N = f_N(P_N)\]
Practical Algorithm for the Master Problem

• The master problem is convex and can be solved using any general-purpose optimization method (interior-point method, cutting-plane method, ellipsoid method, etc.).

• We choose a simple subgradient method:

\[ x_{k+1} = \left[ x_k - \alpha_k s_k \right]^+ . \]

• The subgradient \( s_k \) is readily available and the stepsize can be chosen, for example, with the diminishing stepsize rule:

\[ \alpha_k = \alpha_0 \frac{1 + m}{k + m} . \]

• But what about the projection \( [\cdot]^+_\mathcal{X} \) onto the simplex?
The projection of $x_0$ onto a simplex is defined as the problem:

$$\begin{align*}
\text{minimize} & \quad \|x - x_0\|^2 \\
\text{subject to} & \quad x_i \geq 0, \quad 1 \leq i \leq N \\
& \quad \sum_{i=1}^{N} x_i \leq P.
\end{align*}$$

The optimal solution is unique and is given by

$$x_i = (x_{0,i} - \mu)^+$$

where $\mu$ is chosen as the minimum nonnegative value such that $\sum_i x_i \leq P$. 
Projection Onto a Simplex (II): Practical Algorithm

Algorithm 1. 1. First try $\mu = 0$: if $\sum_i (x_0,i)^+ \leq P$, then $x_i = (x_0,i)^+$ and finish.

2. Obtain $\mu > 0$ such that $\sum_i x_i = P$ as follows.

2.0 Reorder $x_{0,i}$ in decreasing order and set $\tilde{N} = N$.

2.1 If $x_{0,\tilde{N}} > x_{0,\tilde{N}+1}$ and $x_{0,\tilde{N}} > \left(\sum_{i=1}^{\tilde{N}} x_{0,i} - P\right) / \tilde{N}$ then accept hypothesis and go to step 2.2.

Otherwise, reject hypothesis, form a new one by setting $\tilde{N} = \tilde{N} - 1$, and go to step 2.1.

2.2 Set $\mu = \left(\sum_{i=1}^{\tilde{N}} x_{0,i} - P\right) / \tilde{N}$, obtain the optimal solution as $x_i = (x_{0,i} - \mu)^+$, undo the reordering done at step 2.1, and finish.
Example: Minimization of the Maximum MSE (I)

- The problem formulation is *(finite minimax problem)*

\[
\begin{align*}
\text{minimize} & \quad \max \{\alpha_1, \cdots, \alpha_N\} \\
\text{subject to} & \quad \max_i \left\{ \left[ (I + B_k^H R_{H_k} B_k)^{-1} \right]_{ii} \right\} \leq \alpha_k \quad \forall k \\
& \quad \sum_{k=1}^{N} \text{Tr} (B_k B_k^H) \leq P_T,
\end{align*}
\]

- Invoking the decomposition result, the problem reduces to

\[
\begin{align*}
\text{minimize} & \quad \max \{f_1^*(P_1), \cdots, f_N^*(P_N)\} \\
\text{subject to} & \quad \sum_{k=1}^{N} P_k \leq P_T \\
& \quad P_k \geq 0 \quad 1 \leq k \leq N.
\end{align*}
\]
Example: Minimization of the Maximum MSE (II)

- Each $f_k^*(P_k)$ is evaluated by the subproblem with solution
  
  $$p_{k,i} = \left( \mu_k^{-1/2} \lambda_{k,i}^{-1/2} - \lambda_{k,i}^{-1} \right)^+$$

  and has gradient
  
  $$\nabla f_k^*(P_k) = -\mu_k / L_k.$$  

- One possible subgradient of the global cost function $f_0(f_1^*(P_1), \cdots, f_N^*(P_N))$ (which is nondifferentiable) is
  
  $$-\frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} \frac{\mu_k}{L_k} e_k$$

  where $\mathcal{K}$ is the set of subproblems that achieve the maximum value (active subproblems).
Numerical Results: Convergence

- Convergence of the subgradient method for the MAX-MSE design:
Numerical Results: Steady-State Results

- BER vs. the SNR for the methods: SUM-MSE, MAX-MSE, HARM-SINR, and AVE-BER:

![Outage BER (QPSK) for 16 multiple 4x4 MIMO channels with L=3](image-url)
Numerical Results: Additional Constraints $P_k \leq P_k^{\text{max}}$

Power allocation for the AVE–BER without peak constraints

Power allocation for the AVE–BER with peak constraints
Conclusions

• Building on the solutions for Schur-concave/convex cost functions in a single MIMO channel, we have approached the case of a multicarrier MIMO channel.

• Four levels can be identified:
  1. Nonconvex multicarrier problem formulation (no global solution!)
  2. Convex reformulation of multicarrier problem based on majorization theory (global sol. with interior-point method)
  3. Efficient algorithm based on the KKT conditions to solve the waterfilling solutions with multiple waterlevels (no need for IPM)
  4. Simpler approach based on a primal decomposition technique (no need for KKT conditions and computation of coupled multiple waterlevels)
References

