Optimal Linear Precoding for Multiple-Input Multiple-Output Gaussian Channels with Arbitrary Inputs

Fernando Pérez-Cruz[†*], Miguel Rodrigues[‡] and Sergio Verdú[†]

[†]Princeton University
[*]University Carlos III in Madrid
[‡]University of Porto

Centre Tecnològic de Telecomunicacions de Catalunya

May 25th, 2009

Overview

- Problem Set Up.
- Literature Review.
- Optimal Linear Precoding.
- High snr Regime.
- Experimental Results.
- Dependent Inputs.
- Optimal Power Allocation.
- Conclusions.

Set Up

$$\mathbf{y} = \sqrt{\mathsf{snr}}\,\mathbf{H}\,\mathbf{P}\,\mathbf{x} + \mathbf{w}$$

- $\mathbf{y}$: Received vector.
- $\mathbf{x}$: Independent zero-mean unit-variance information vector.
- $\mathbf{w}$: Independent zero-mean unit-variance complex Gaussian noise.
- snr: Scaling factor (total power).
- $\mathbf{H}$: Deterministic channel gains (known encoder and decoder).
- $\mathbf{P}$: Full complex-value linear precoder.

Set Up

- Optimal Power Allocation (Lozano-Tulino-Verdú'06):

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sqrt{\mathsf{snr}} \begin{pmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n \end{pmatrix} \begin{pmatrix} \sqrt{p_1} & 0 & \cdots & 0 \\ 0 & \sqrt{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

- Optimal Precoding:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sqrt{\mathsf{snr}} \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nm} \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1q} \\ p_{21} & p_{22} & \cdots & p_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mq} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Particular Cases

- Optimal Power Allocation (Lozano-Tulino-Verdú'06):

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sqrt{\mathsf{snr}} \begin{pmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n \end{pmatrix} \begin{pmatrix} \sqrt{p_1} & 0 & \cdots & 0 \\ 0 & \sqrt{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

- Optimal Precoding for noninterfering channels:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sqrt{\mathsf{snr}} \begin{pmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1q} \\ p_{21} & p_{22} & \cdots & p_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nq} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Particular Cases

- Optimal Power Allocation (Lozano-Tulino-Verdú'06):

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sqrt{\mathsf{snr}} \begin{pmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_n \end{pmatrix} \begin{pmatrix} \sqrt{p_1} & 0 & \cdots & 0 \\ 0 & \sqrt{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

- Optimal power allocation for MIMO channels:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sqrt{\mathsf{snr}} \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1m} \\ h_{21} & h_{22} & \cdots & h_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nm} \end{pmatrix} \begin{pmatrix} \sqrt{p_1} & 0 & \cdots & 0 \\ 0 & \sqrt{p_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{p_m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Related literature

- Diagonalization + waterfilling + power gap (Bingham et al. (1995), Forney and Ungerboek (1998), Star et al. (1999), ...).

- Relation between mutual information and MMSE (Guo et al. (2005), and Palomar and Verdú (2006)):

$$\frac{\mathrm{d}I(\mathbf{x};\mathbf{y})}{\mathsf{dsnr}} = \mathsf{mmse}(\mathsf{snr})$$

- Power allocation with arbitrary inputs for parallel channels (Lozano et al. (2006)):

Mercury/Waterfilling for parallel Gaussian Channels

Maximum Transmission Rate

- The channel capacity is given by:

$$C = \max_{p(\mathbf{x}),\mathbf{P}} I(\mathbf{x};\mathbf{y}) \qquad \text{s.t.} \qquad \mathrm{Tr}\{\mathbf{P}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{P}^{\dagger}\} \leq 1$$

We achieve capacity for:

- $p^*(\mathbf{x}) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- $\mathbf{P}^*$ diagonalizes the channel and waterfills each input.

- Maximum Transmission Rate:

$$\max_{\mathbf{P}} I(\mathbf{x};\mathbf{y}) \qquad \text{s.t.} \qquad \mathrm{Tr}\{\mathbf{P}\mathbf{P}^{\dagger}\} \leq 1$$

- $p(\mathbf{x})$ is fixed a priori (e.g. QPSK or QAM).

Karush-Kuhn-Tucker Conditions

- We build the Lagrangian (for a nonnegative λ):

$$\mathfrak{L}(\mathbf{P}, \lambda) = -I(\mathbf{x}; \mathbf{y}) - \lambda \left(1 - \mathrm{Tr}\{\mathbf{P}\mathbf{P}^\dagger\}\right)$$

- The first order conditions are given by:

$$\begin{aligned} \nabla_{\mathbf{P}} \mathfrak{L}(\mathbf{P}, \lambda) = -\nabla_{\mathbf{P}} I(\mathbf{x}; \mathbf{y}) + \lambda \mathbf{P} = 0 \\ \lambda \left(1 - \mathrm{Tr}\{\mathbf{P}\mathbf{P}^\dagger\}\right) = 0 \end{aligned}$$

- The precoder can be computed as a fixed-point equation:

$$\mathbf{P}^* = \lambda^{-1} \nabla_{\mathbf{P}} I(\mathbf{x}; \mathbf{y}) = \frac{\mathbf{H}^\dagger \mathbf{H} \mathbf{P}^* \mathbf{E}}{\|\mathbf{H}^\dagger \mathbf{H} \mathbf{P}^* \mathbf{E}\|}$$

where $(\mathbf{E})_{ij} = \mathbb{E}\left[(x_i - \mathbb{E}[x_i|\mathbf{y}])^\dagger (x_j - \mathbb{E}[x_j|\mathbf{y}])\right]$ (MMSE matrix).

High snr Regime for Discrete Inputs I

- The MMSE for the MIMO channel.

$$\mathsf{mmse}(\mathsf{snr}) = \mathbb{E}\left[||\mathbf{HPx} - \mathbf{HP}\mathbb{E}[\mathbf{x}|\mathbf{y}]||^2\right]$$

- We can bound the MMSE as follows:

$$\frac{1}{2M}\frac{e^{-d_{min}^2\mathsf{snr}/4}}{d_{min}\sqrt{\mathsf{snr}}}\left(\sqrt{\pi} - \frac{4.37}{d_{min}^2\mathsf{snr}}\right) \leq \mathsf{mmse}(\mathsf{snr}) \leq (M-1)d_{max}^2\frac{e^{-d_{min}^2\mathsf{snr}/4}}{d_{min}\sqrt{\mathsf{snr}\pi}}$$

  - Tight for large snr.

- $d_{min} = \min_{\substack{\bar{\mathbf{x}},\mathbf{x} \\ \mathbf{x}\neq\bar{\mathbf{x}}}} ||\mathbf{HPx} - \mathbf{HP}\bar{\mathbf{x}}||$ and $d_{max} = \max_{\bar{\mathbf{x}},\mathbf{x}} ||\mathbf{HPx} - \mathbf{HP}\bar{\mathbf{x}}||$

High snr Regime for Discrete Inputs II

- We can bound the Mutual Information as follows:

$$\log M - \frac{2(M-1)d_{max}^2}{d_{min}^3\sqrt{\mathsf{snr}\pi}} e^{-d_{min}^2\mathsf{snr}/4} \leq I(\mathbf{x};\mathbf{y})$$

$$\leq \log M - \frac{e^{-d_{min}^2\mathsf{snr}/4}}{M d_{min}^3\sqrt{\mathsf{snr}}}\left(\sqrt{\pi} - \frac{4.37 + 2\sqrt{\pi}}{d_{min}^2\mathsf{snr}}\right)$$

  - Tight for large snr.

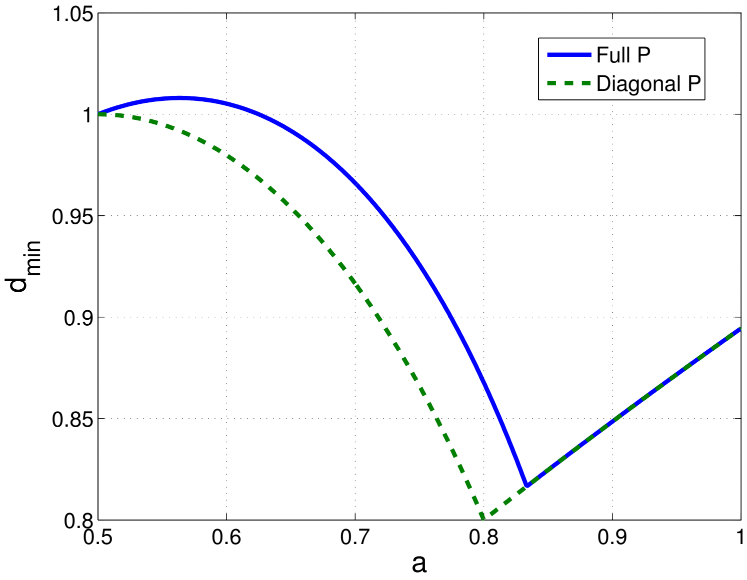
- We have use the previous MMSE bounds and the relation between the MMSE and mutual information:

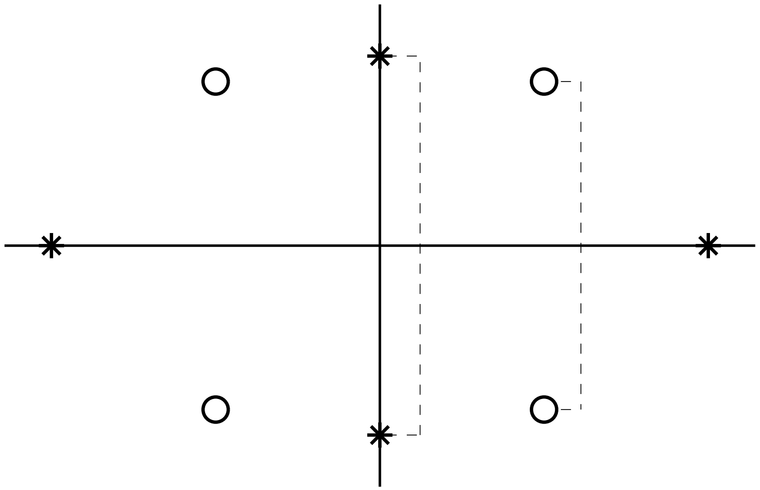
$$I(\mathbf{x};\mathbf{y}) = \int_0^{\mathsf{snr}} \mathsf{mmse}(\xi)d\xi = \log M - \int_{\mathsf{snr}}^{\infty} \mathsf{mmse}(\xi)d\xi$$

High snr Regime for Discrete Inputs III

- Identical $\mathbf{P}^*$ for 3 different criteria:
  - $\max_{\mathbf{P}} I(\mathbf{x};\mathbf{y}) \quad \Leftrightarrow \quad \min_{\mathbf{P}} \mathsf{mmse}(\mathsf{snr}) \quad \Leftrightarrow \quad \max_{\mathbf{P}} d_{min}$
- $\mathbf{P}^*$ does not diagonalize and waterfill the channel.
- For non-interfering channels: $\mathbf{P}^*$ is not diagonal.
- $\mathbf{P}^*$ diagonalizes and waterfills the channel if:
  - Gaussian inputs.
  - Linear equalizer: $\hat{\mathbf{x}} = \mathbf{L}\mathbf{y}$.

Simple Illustration





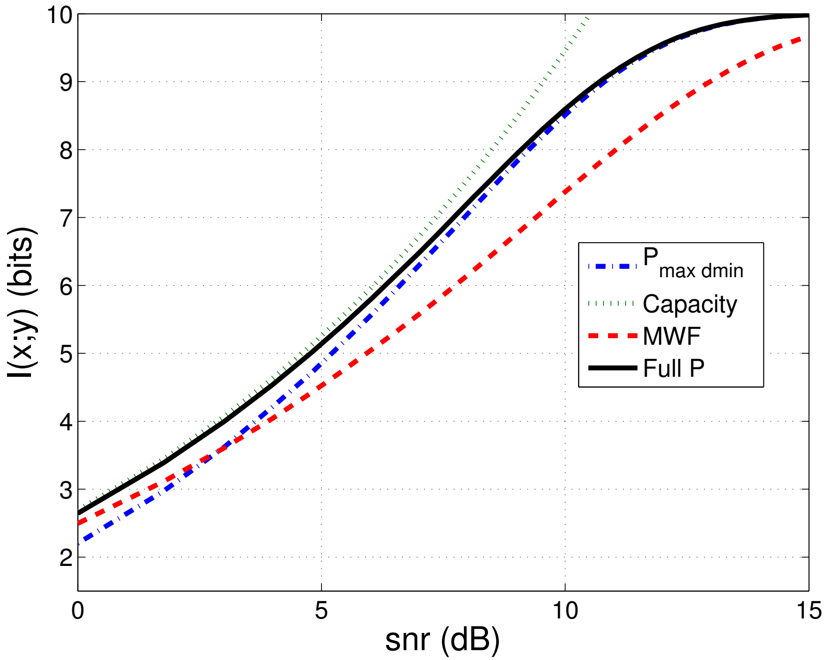
$a = 0.634$

- $\mathbf{H} = \begin{bmatrix} \sqrt{a} & 0 \\ 0 & 1 - \sqrt{a} \end{bmatrix}$
- Both inputs: BPSK (or 4QAM and nothing).

Examples

- Non-interfering channels:
  - Encounter in digital Communications: Digital Subscriber Line.
  - Typical solution: Modified waterfilling.
- Mutually interfering channels:
  - Gigabit DSL: 4 copper pairs (pedestal-drop segment).
  - Typical solution: Diagonalize and modified waterfilling.

Non-Interfering Channel



- QPSK inputs

- $$\mathbf{H} = \begin{bmatrix} 2e^{-j2.5} & 0 & 0 & 0 & 0 \\ 0 & 1.5e^{j1.6} & 0 & 0 & 0 \\ 0 & 0 & e^{-j0.8} & 0 & 0 \\ 0 & 0 & 0 & 0.8e^{-j3.0} & 0 \\ 0 & 0 & 0 & 0 & 0.6e^{-j1.7} \end{bmatrix}$$

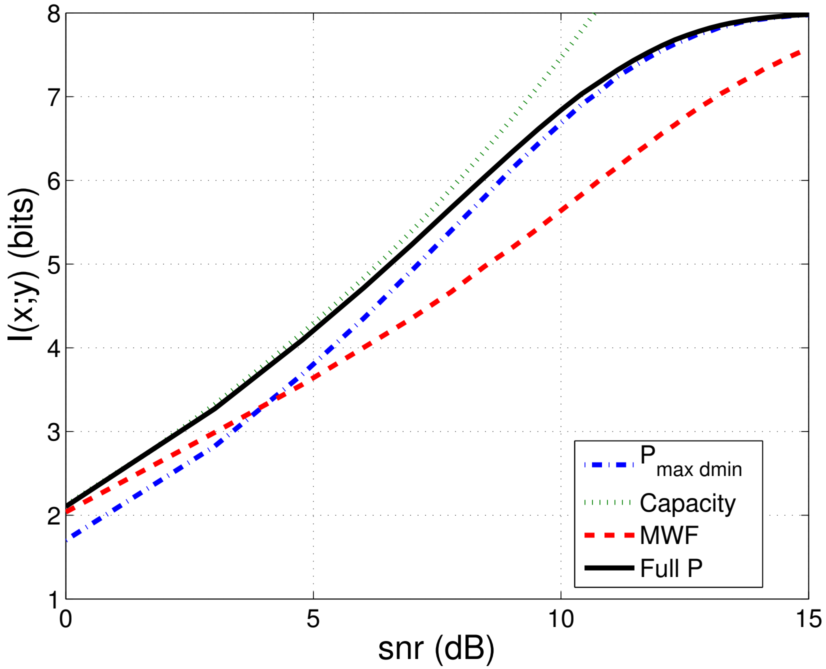
Maximum Minimum Distance Matrices

$$\mathbf{P}_1 = \begin{bmatrix} 0.20 & 0 & 0 & 0 & 0 \\ 0 & 0.27 & 0 & 0 & 0 \\ 0 & 0 & 0.41 & 0 & 0 \\ 0 & 0 & 0 & 0.51 & 0 \\ 0 & 0 & 0 & 0 & 0.68 \end{bmatrix}$$

$$\mathbf{P}_{\max d_{\min}} = \begin{bmatrix} 0.28e^{j2.5} & 0.26e^{j1.8} & 0.27e^{-j0.1} & 0.16e^{j1.5} & 0.23e^{j1.8} \\ 0.09e^{j1.9} & 0.21e^{j0.9} & 0.16e^{-j3.1} & 0.34e^{-j2.5} & 0.25e^{j2.7} \\ 0.26e^{j0.7} & 0.09e^{-j2.6} & 0.15e^{j1.6} & 0.17e^{j0.1} & 0.27e^{j3.0} \\ 0.07e^{j0.6} & 0.23e^{j1.4} & 0.21e^{j1.8} & 0.02e^{-j0.1} & 0.17e^{-j0.7} \\ 0.18e^{j0.8} & 0.17e^{-j2.9} & 0.15e^{-j0.5} & 0.21e^{-j1.0} & 0.12e^{-j1.0} \end{bmatrix}$$

- $\mathbf{H}\mathbf{P}_1 = c\mathbf{I}$.
- Linear precoder assigns equal power to every input and transmits more power through the strongest and more reliable channel.

Mutually Interfering Channel



- QPSK inputs

- $$\mathbf{H} = \begin{bmatrix} e^{j2.80} & 0.5e^{-j0.97} & 0.01e^{-j1.02} & 10^{-3}e^{-j1.06} \\ 0.5e^{-j0.43} & e^{-j2.98} & 0.1e^{j0.12} & 0.01e^{-j3.12} \\ 0.01e^{j2.93} & 0.1e^{j1.27} & e^{-j2.00} & 0.5e^{-j1.83} \\ 10^{-3}e^{j2.63} & 0.01e^{-j2.06} & 0.5e^{j2.38} & e^{j2.69} \end{bmatrix}$$

Dependent inputs

- We drop the independent constraint over $\mathbf{x}$:

$$\max_{\mathbf{P},p(\mathbf{x})} I(\mathbf{x};\mathbf{y}) \qquad \textit{where} \qquad p(\mathbf{x}) = \prod p(x_i|x_1,\ldots,x_{i-1}).$$

subject to: $\text{Tr}\{\mathbf{P}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{P}^\dagger\} \leq 1$, $\sum_{\mathbf{x}} p(\mathbf{x}) = 1$ and $p(\mathbf{x}) \geq 0 \quad \forall\mathbf{x}$, where $(\boldsymbol{\Sigma}_{\mathbf{x}})_{ij} = \sum_{\mathbf{x}} x_i^\dagger x_j p(\mathbf{x})$.

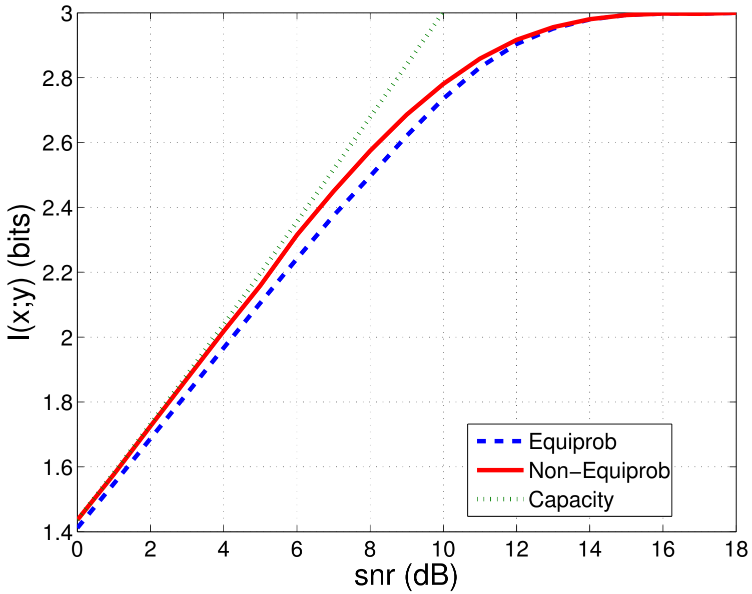
- Exponential family representation of $p(\mathbf{x})$:

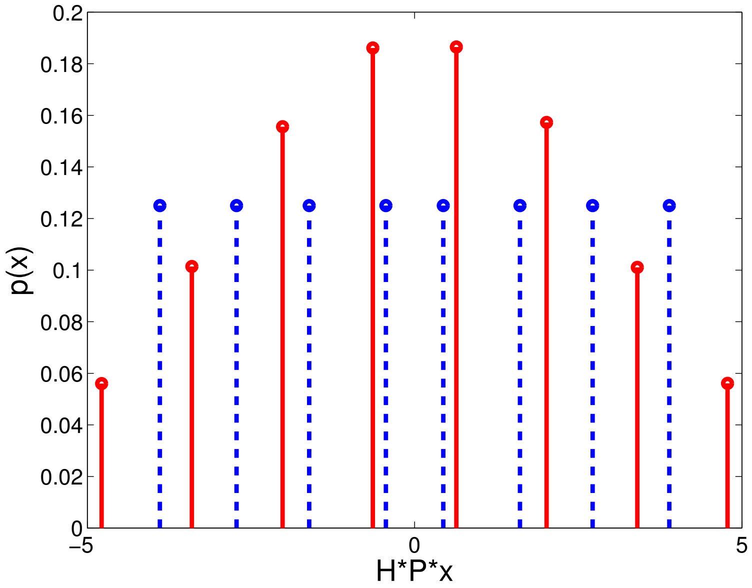
$$p(x_i|x_1,\ldots,x_{i-1}) = \frac{\exp\left(x_i \sum_{j=1}^{i-1} \theta_{ji} x_j\right)}{2\cosh\left(\sum_{j=1}^{i-1} \theta_{ji} x_j\right)}, \qquad x_j \in \{\pm 1\}.$$

- Maximum transmission rate:

$$\max_{\mathbf{P},\boldsymbol{\theta}} I(\mathbf{x};\mathbf{y}) \qquad s.t. \qquad \text{Tr}\{\mathbf{P}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{P}^\dagger\} \leq 1.$$

Mutually Interfering Channel

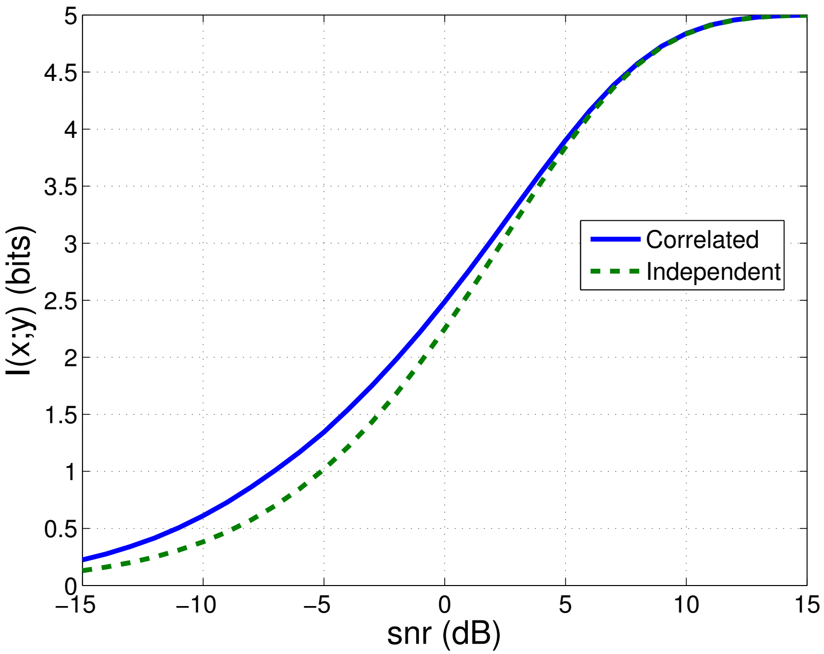




- $\mathbf{H} = \begin{bmatrix} 2 & 1.3 & 0.8 \end{bmatrix}$

- BPSK Inputs

Mutually Interfering Channel & Correlated Inputs



- BPSK inputs

- $$\mathbf{H} = \begin{bmatrix} 0.928 & -0.308 & -0.573 & 0.428 & -0.281 \\ 0.1733 & 0.267 & -0.347 & 1.478 & 0.995 \\ -0.6912 & 1.335 & -0.644 & -0.692 & 0.174 \\ -0.723 & -1.331 & 0.868 & -2.678 & -0.706 \\ -0.574 & 1.399 & 1.639 & 1.068 & -2.087 \end{bmatrix}$$

Optimal Power Allocation

- Diagonal Precoder: $(\mathbf{P})_{ij} = \sqrt{p_j}\delta_{ij}$

$$\max_{p_j} I(\mathbf{x};\mathbf{y})$$

  - subject to: $\sum_j p_j \leq 1$ and $p_j \geq 0$.
  - $p(\mathbf{x})$ is fixed a priori (e.g. QPSK or QAM) and independent.
- Construct the Lagrangian and get the KKT conditions.
- Solution:

$$p_j^* = \frac{\left(\mathbf{P}^*\mathbf{H}^\dagger\mathbf{H}\mathbf{P}^*\mathbf{E}\right)_{jj}}{\mathrm{Tr}\left\{\mathbf{P}^*\mathbf{H}^\dagger\mathbf{H}\mathbf{P}^*\mathbf{E}\right\}}$$

- For diagonal $\mathbf{H}$, we recover the Mercury/Waterfilling solution.

Generalized Mercury/Waterfilling

- Water level:

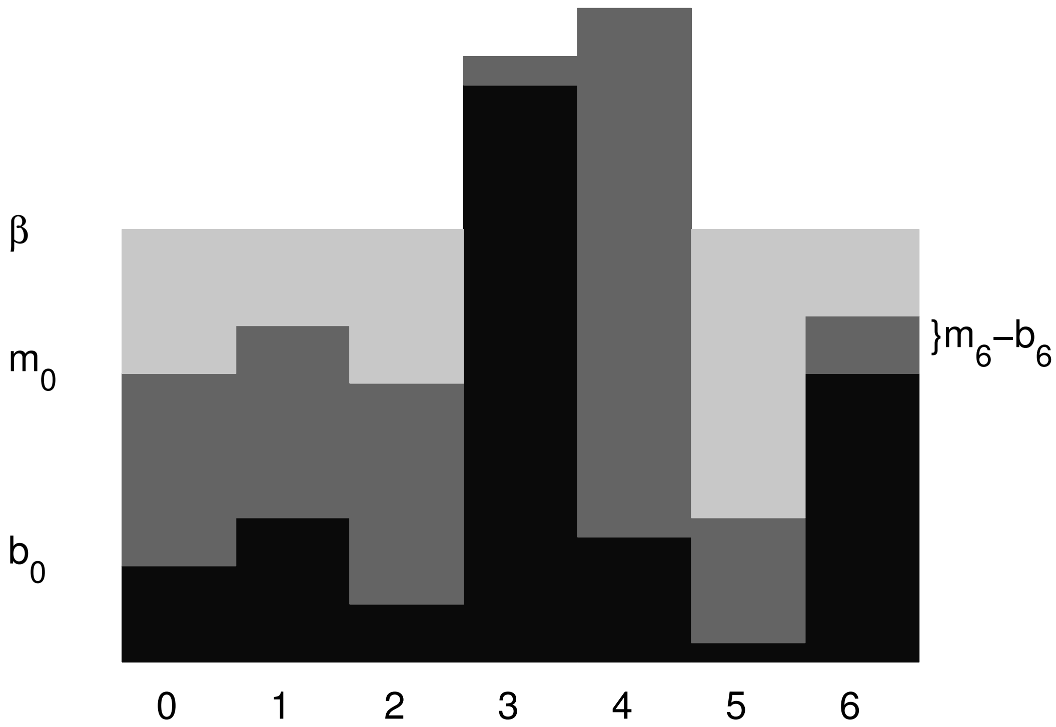
$$\beta = \frac{1}{\mathsf{snr}\,\mathrm{Tr}\left\{\mathbf{P}^*\mathbf{H}^\dagger\mathbf{H}\mathbf{P}^*\mathbf{E}\right\}}$$

- Base Level:

$$b_j = \frac{1}{\mathsf{snr}||\mathbf{h}_j||^2}$$

- Mercury Level:

$$m_j = \frac{1}{\left.\frac{\partial I(\mathbf{x};\mathbf{y})}{\partial p_j}\right|_{\mathbf{P}=\mathbf{P}^*}} - p_j^*$$

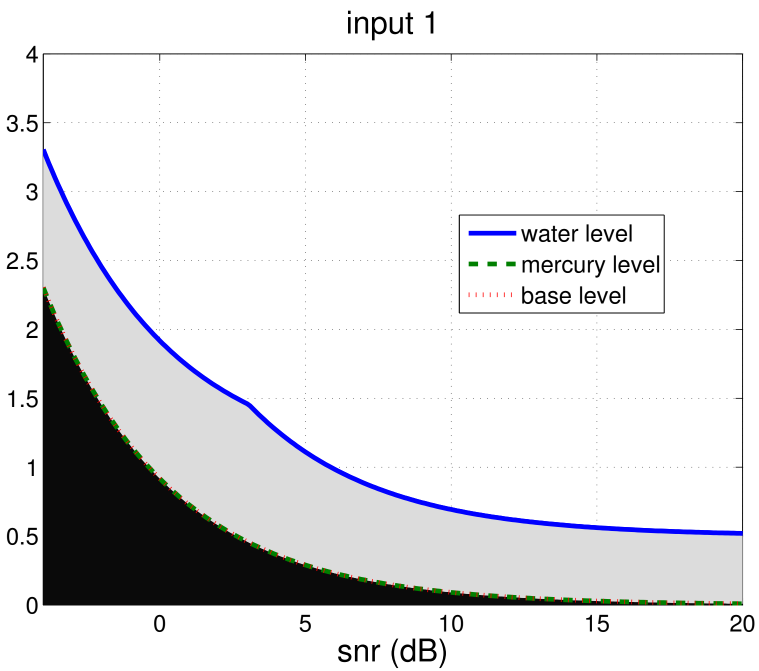


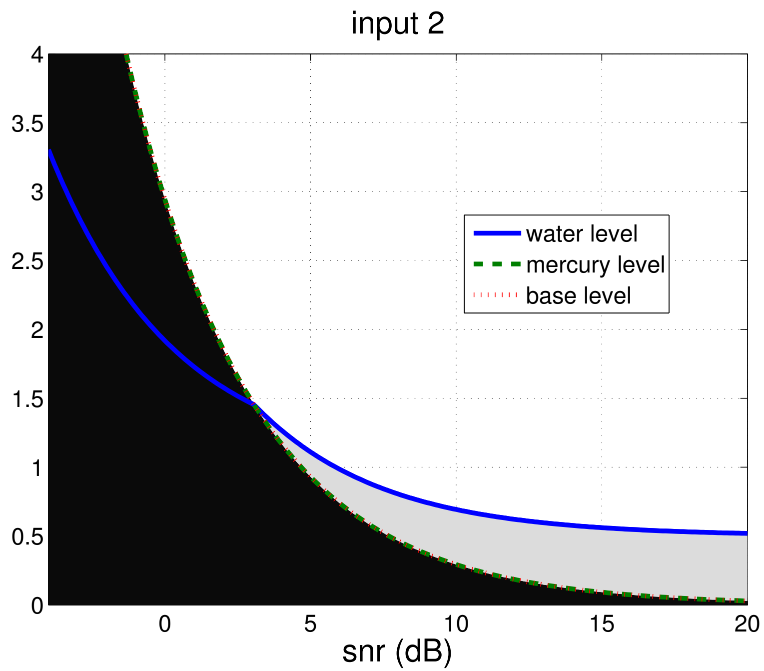
Examples

- 2 Channel matrices:
  - $\mathbf{H} = \begin{bmatrix} \sqrt{1.09} & 0 \\ 0 & \sqrt{0.34} \end{bmatrix}$
  - $\mathbf{H} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 0.5 \end{bmatrix}$
- 2 inputs:
  - Gaussian.
  - BPSK.
- Base, mercury and water levels as a function of the snr.

Waterfilling

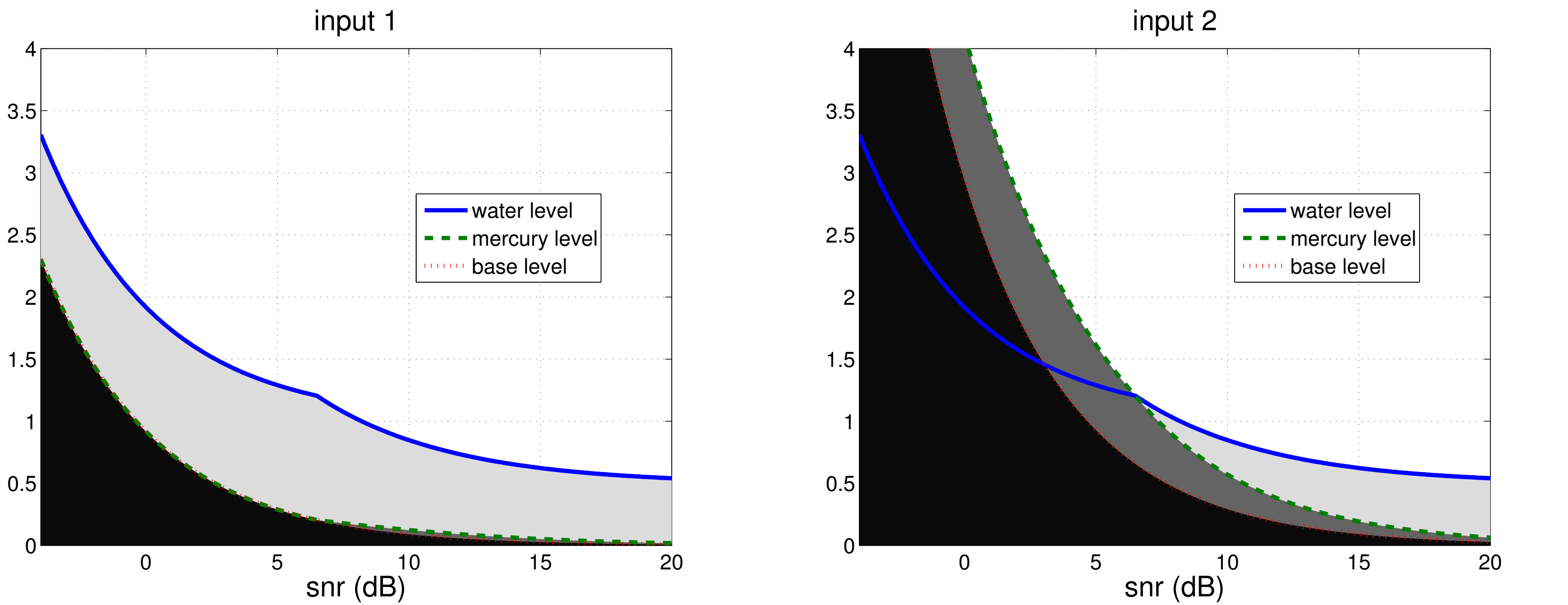
- Diagonal $\mathbf{H}$ and Gaussian inputs:





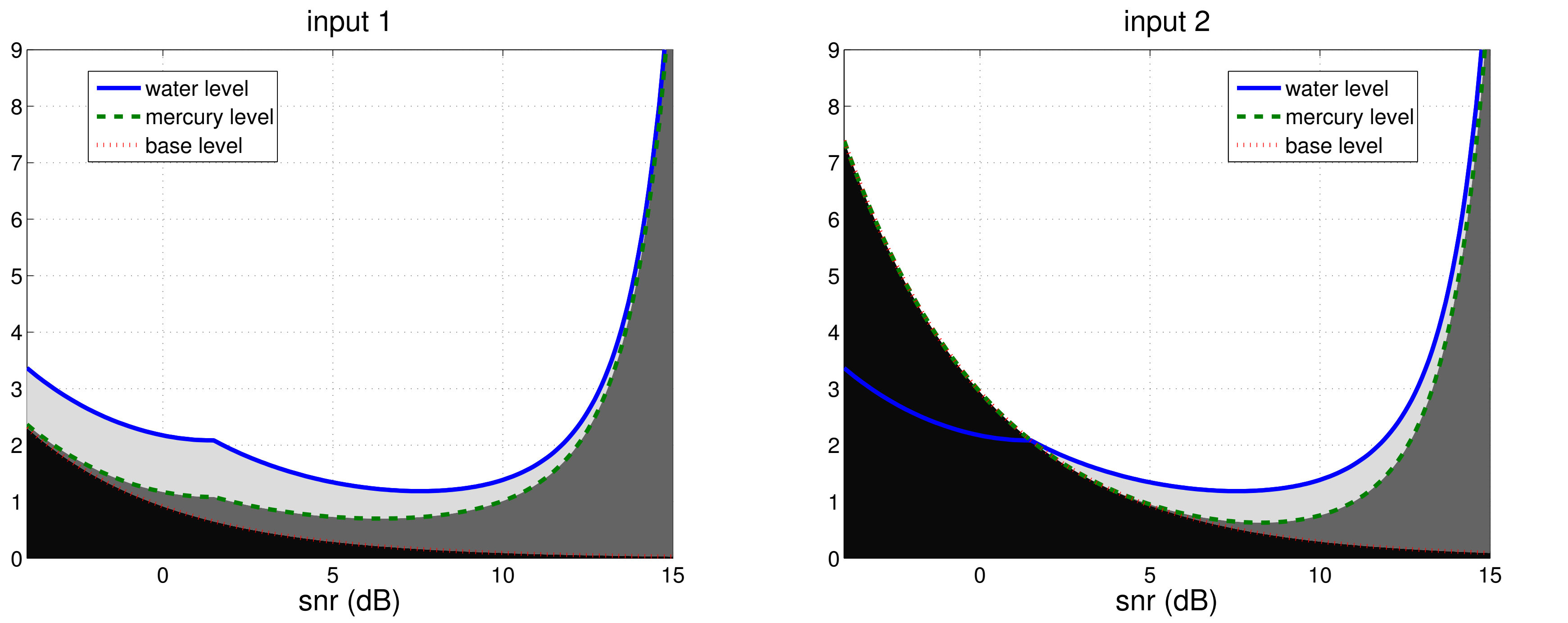
Interfering Gaussians

- Non-diagonal $\mathbf{H}$ and Gaussian inputs:



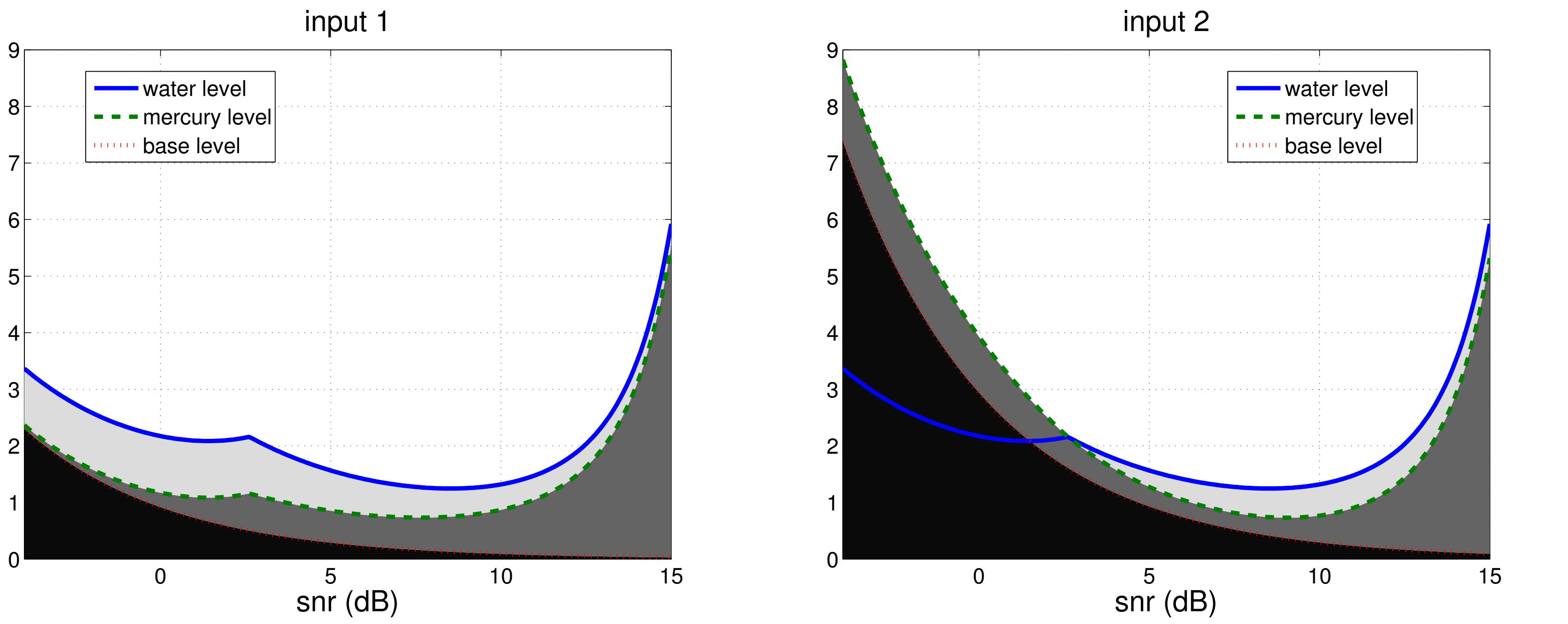
Mercury/Watefilling

- Diagonal $\mathbf{H}$ and BPSK inputs:



Generalized Mercury/Watefilling

- Non-diagonal $\mathbf{H}$ and BPSK inputs:



Conclusions

- Power allocation for MIMO channels with arbitrary inputs.
- Solution based on mutual information and MMSE relation.
- Upper/Lower bounds for mutual information in high snr regime.
- Generalized Mercury/Waterfilling solution.
- Extension for dependent inputs.