ETH zürich

Taiwan National
Yang Ming Chiao
Tung University

Energy-Optimal Signaling using the Example of Optical Communication

Stefan M. Moser

26 January 2023

Information Theory and Tapas Workshop — Universidad Carlos III de Madrid

Joint Work with

Longguang Li

East China
Normal University
China

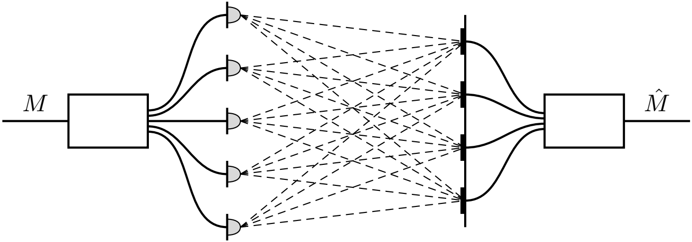
Ligong Wang

ETIS Laboratory
Cergy-Pontoise
France

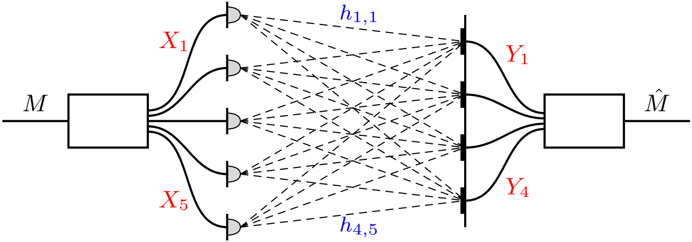
Michèle Wigger

Télécom Paris
France

Model for Optical Communication



Model for Optical Communication



Optical MIMO Channel (1/3)

$$
\begin{aligned}
Y_1 &= h_{1,1}x_1 + h_{1,2}x_2 + \cdots + h_{1,n_\mathsf{T}}x_{n_\mathsf{T}} + Z_1 \\
Y_2 &= h_{2,1}x_1 + h_{2,2}x_2 + \cdots + h_{2,n_\mathsf{T}}x_{n_\mathsf{T}} + Z_2 \\
&\vdots \\
Y_{n_\mathsf{R}} &= h_{n_\mathsf{R},1}x_1 + h_{n_\mathsf{R},2}x_2 + \cdots + h_{n_\mathsf{R},n_\mathsf{T}}x_{n_\mathsf{T}} + Z_{n_\mathsf{R}}
\end{aligned}
$$

where

- $x_1, \ldots, x_{n_\mathsf{T}}$ describe LEDs (light intensity $\Longrightarrow$ nonnegative)
- $Y_1, \ldots, Y_{n_\mathsf{R}}$ describe photo detectors (electrical signal)
- $h_{i,j}$ describe constant channel coefficients (nonnegative)
- $Z_1, \ldots, Z_{n_\mathsf{R}}$ is additive, signal-independent Gaussian noise (due to thermal noise and ambient light; shot noise or relative intensity noise is neglected)

Optical MIMO Channel (2/3)

In vector-notation:

$$\mathbf{Y} = \mathsf{H}\mathbf{x} + \mathbf{Z}$$

- $n_\mathsf{R} \times n_\mathsf{T}$ matrix H with nonnegative entries
- n_T-vector $\mathbf{x}$ with nonnegative entries (intensities!)

Optical MIMO Channel (2/3)

In vector-notation:

$$\mathbf{Y} = \mathsf{H}\mathbf{x} + \mathbf{Z}$$

- $n_\mathsf{R} \times n_\mathsf{T}$ matrix H with nonnegative entries
- n_T-vector $\mathbf{x}$ with nonnegative entries (intensities!)
- **power constraints:**
  - limited total average power $\mathcal{E}$: $$\sum_{i=1}^{n_\mathsf{T}} \mathsf{E}[X_i] \leq \mathcal{E}$$
  - limited per-antenna peak power A: $\Pr[X_i > A] = 0$ $(i = 1, \ldots, n_\mathsf{T})$

Optical MIMO Channel (2/3)

In vector-notation:

$$\mathbf{Y} = \mathsf{H}\mathbf{x} + \mathbf{Z}$$

- $n_\mathsf{R} \times n_\mathsf{T}$ matrix H with nonnegative entries
- n_T-vector $\mathbf{x}$ with nonnegative entries (intensities!)
- **power constraints:**
  - limited total average power E: $\quad \displaystyle\sum_{i=1}^{n_\mathsf{T}} \mathsf{E}[X_i] \leq \mathsf{E}$
  - limited per-antenna peak power A: $\quad \Pr[X_i > \mathrm{A}] = 0$ $(i = 1, \ldots, n_\mathsf{T})$

$\Longrightarrow$ **constraint on first moment**, not on second moment!

Optical MIMO Channel (3/3)

We define average-to-peak power ratio

$$\alpha \triangleq \frac{\mathsf{E}}{A}$$

Optical MIMO Channel (3/3)

We define average-to-peak power ratio

$$\alpha \triangleq \frac{\mathsf{E}}{A}$$

- α is a channel parameter: $0 < \alpha \leq n_{\mathsf{T}}$
- we play with A and have $\mathsf{E} = \alpha A$

Optical MIMO Channel (3/3)

We define average-to-peak power ratio

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}$$

- α is a channel parameter: $0 < \alpha \leq n_{\mathsf{T}}$
- we play with A and have $\mathsf{E} = \alpha\mathsf{A}$

3 cases:

- If "$\alpha = 0$": only average-power constraint

Optical MIMO Channel (3/3)

We define average-to-peak power ratio

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}$$

- α is a channel parameter: $0 < \alpha \leq n_{\mathsf{T}}$
- we play with A and have $\mathsf{E} = \alpha \mathsf{A}$

3 cases:

- If "$\alpha = 0$": only average-power constraint
- If $\alpha = n_{\mathsf{T}}$: only peak-power constraint

Optical MIMO Channel (3/3)

We define average-to-peak power ratio

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}$$

- α is a channel parameter: $0 < \alpha \leq n_{\mathsf{T}}$
- we play with A and have $\mathsf{E} = \alpha \mathsf{A}$

3 cases:

- If "$\alpha = 0$": only average-power constraint
- If $\frac{n_{\mathsf{T}}}{2} \leq \alpha \leq n_{\mathsf{T}}$: average-power constraint inactive $\Longrightarrow$ only peak-power constraint

Optical MIMO Channel (3/3)

We define average-to-peak power ratio

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}$$

- α is a channel parameter: $0 < \alpha \leq n_{\mathsf{T}}$
- we play with A and have $\mathsf{E} = \alpha \mathsf{A}$

3 cases:

- If "$\alpha = 0$": only average-power constraint
- If $\frac{n_{\mathsf{T}}}{2} \leq \alpha \leq n_{\mathsf{T}}$: average-power constraint inactive $\Longrightarrow$ only peak-power constraint
- If $0 < \alpha < \frac{n_{\mathsf{T}}}{2}$: both peak- and average-power constraint

Channel Capacity

- is the maximum rate at which information can be transmitted over the channel reliably
- can be computed as

$$\max_{P_{\mathbf{X}}} \mathrm{I}(\mathbf{X};\mathbf{Y})$$

where $P_{\mathbf{X}}$ must satisfy power constraints

Channel Capacity

- is the maximum rate at which information can be transmitted over the channel reliably
- can be computed as

$$\max_{P_{\mathbf{X}}} \mathrm{I}(\mathbf{X};\mathbf{Y})$$

  where $P_{\mathbf{X}}$ must satisfy power constraints

- is a function of the available average and peak power:

$$\mathsf{C}(\mathsf{A},\mathsf{E}) = \mathsf{C}(\mathsf{A},\alpha\mathsf{A})$$

Channel Capacity

- is the maximum rate at which information can be transmitted over the channel reliably
- can be computed as

$$\max_{P_{\mathbf{X}}} \mathrm{I}(\mathbf{X};\mathbf{Y})$$

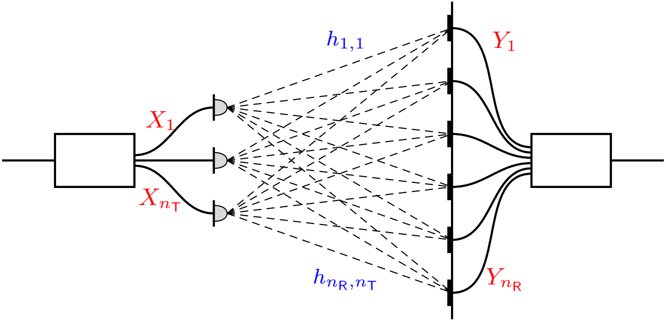
  where $P_{\mathbf{X}}$ must satisfy power constraints
- is a function of the available average and peak power:

$$\mathsf{C}(A,\mathsf{E}) = \mathsf{C}(A,\alpha A)$$

- exact capacity expression is not known
- in SISO case tight bounds are known
- in SISO asymptotic ($A \to \infty$ and $A \to 0$) behavior of capacity is known

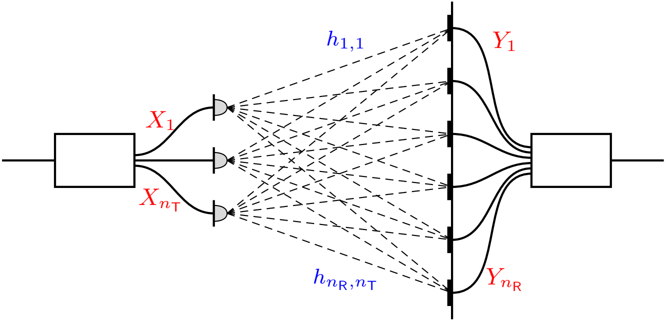
Full-Rank MIMO Channel: $n_\mathsf{T} \leq n_\mathsf{R}$ (1/2)

We have more receiver antennas than transmitter antennas:



Full-Rank MIMO Channel: $n_\mathsf{T} \le n_\mathsf{R}$ (1/2)

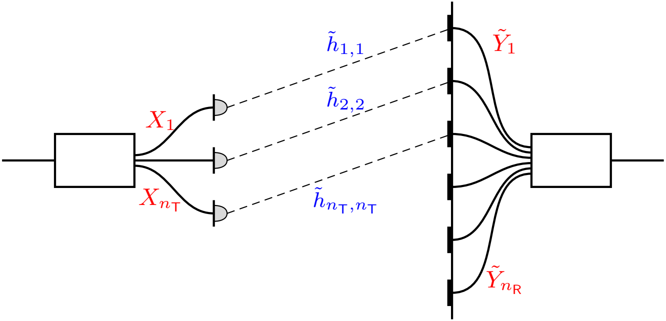
We have more receiver antennas than transmitter antennas:



$\Longrightarrow$ orthogonalize channel matrix H (SVD)

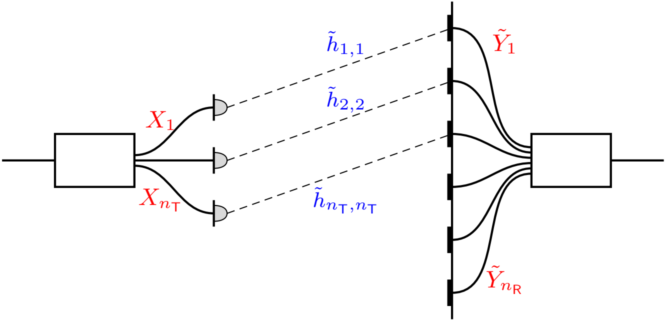
Full-Rank MIMO Channel: $n_\mathsf{T} \leq n_\mathsf{R}$ (2/2)

We obtain equivalent channel with identical capacity:



Full-Rank MIMO Channel: $n_\mathsf{T} \le n_\mathsf{R}$ (2/2)

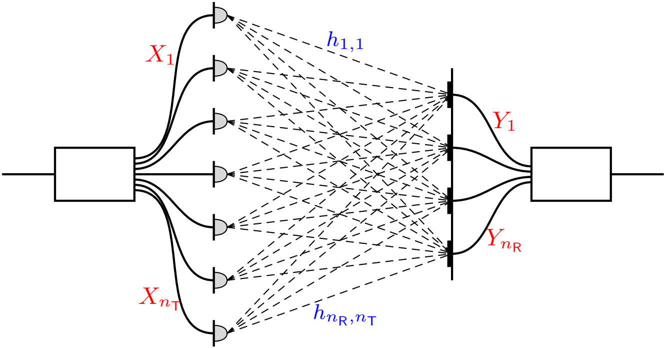
We obtain equivalent channel with identical capacity:



- we ignore redundant photo detectors
- we obtain n_T parallel channels without interference

Rank-Deficient MIMO Channel: $n_\mathsf{T} > n_\mathsf{R}$

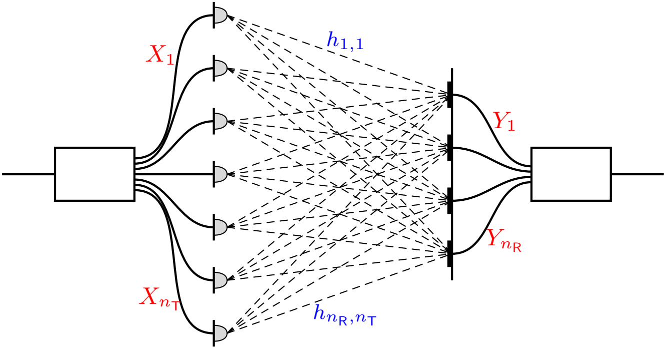
We have more transmitters than receivers:



$\Longrightarrow$ orthogonalization not possible: it is not optimal to ignore inputs

Rank-Deficient MIMO Channel: $n_\mathsf{T} > n_\mathsf{R}$

We have more transmitters than receivers:



$\Longrightarrow$ orthogonalization not possible: it is not optimal to ignore inputs

$\Longrightarrow$ let's start with MISO: $n_\mathsf{T} > 1$, $n_\mathsf{R} = 1$

$$Y = h_1 x_1 + h_2 x_2 + \cdots + h_{n_\mathsf{T}} x_{n_\mathsf{T}} + Z$$

MISO Example: Average-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z \qquad \text{average-power constraint only}$$

MISO Example: Average-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z \qquad \text{average-power constraint only}$$

- X_1 sees a much better channel!

  $\Longrightarrow$ put all energy into first antenna: $X_2 = 0$

  $\Longrightarrow$ like SISO with $h = h_{\mathsf{max}} = h_1 = 9$

An Observation (about SISO Case)

At high power: because of additive noise,

$$\begin{aligned}\mathrm{I}(X;Y) &= \mathrm{h}(Y) - \mathrm{h}(Y|X)\\ &= \mathrm{h}(X+Z) - \mathrm{h}(X+Z|X)\\ &= \mathrm{h}(X+Z) - \mathrm{h}(Z)\\ &\approx \mathrm{h}(X) - \mathrm{h}(Z)\end{aligned}$$

An Observation (about SISO Case)

At high power: because of additive noise,

$$\begin{aligned} \mathrm{I}(X;Y) &= \mathrm{h}(Y) - \mathrm{h}(Y|X) \\ &= \mathrm{h}(X+Z) - \mathrm{h}(X+Z|X) \\ &= \mathrm{h}(X+Z) - \mathrm{h}(Z) \\ &\approx \mathrm{h}(X) - \mathrm{h}(Z) \end{aligned}$$

$\Longrightarrow$ a good input maximizes differential entropy!

An Observation (about SISO Case)

At high power: because of additive noise,

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) \\ &= h(X+Z) - h(X+Z|X) \\ &= h(X+Z) - h(Z) \\ &\approx h(X) - h(Z) \end{aligned}$$

$\Longrightarrow$ a good input maximizes differential entropy!

- if $0 \le X \le \mathsf{A}$: $X \sim \text{Uniform}([0, \mathsf{A}])$
- if $X \ge 0$ and $\mathsf{E}[X] \le \mathrm{E}$: $X \sim \text{Exp}(\mathrm{E})$
- if both $\mathsf{E}[X] \le \mathrm{E}$ and $0 \le X \le \mathsf{A}$: $X \sim \text{truncExp}\big([0, \mathsf{A}], \mathrm{E}\big)$

MISO Example: Average-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z \qquad \text{average-power constraint only}$$

- X_1 sees a much better channel!

  $\Longrightarrow$ put all energy into first antenna: $X_2 = 0$

  $\Longrightarrow$ like SISO with $h = h_{\mathsf{max}} = h_1 = 9$

MISO Example: Average-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z \qquad \text{average-power constraint only}$$

- X_1 sees a much better channel!

  $\Longrightarrow$ put all energy into first antenna: $X_2 = 0$

  $\Longrightarrow$ like SISO with $h = h_{\mathsf{max}} = h_1 = 9$

  $\Longrightarrow$ $X_1 \sim \mathsf{Exp}(\mathsf{E})$ (at high SNR)

Average-power constraint only:

$$\mathsf{C}_{\mathsf{MISO}}(\mathsf{E}) = \mathsf{C}_{\mathsf{SISO}}(h_{\mathsf{max}}\mathsf{E})$$

MISO Example: Peak-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z$$

MISO Example: Peak-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z$$

- make full use of each antenna!

  $\Longrightarrow$ "beamforming": choose $X_1 = X_2 \triangleq X \sim \text{Uniform}([0, A])$

MISO Example: Peak-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z$$

- make full use of each antenna!

  $\Longrightarrow$ "beamforming": choose $X_1 = X_2 \triangleq X \sim \text{Uniform}([0, A])$

  $\Longrightarrow Y = 12X + Z$

  $\Longrightarrow$ like SISO with $h = h_{\text{sum}} = h_1 + h_2 = 12$

MISO Example: Peak-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z$$

- make full use of each antenna!

  $\Longrightarrow$ "beamforming": choose $X_1 = X_2 \triangleq X \sim \text{Uniform}([0, A])$

  $\Longrightarrow Y = 12X + Z$

  $\Longrightarrow$ like SISO with $h = h_{\text{sum}} = h_1 + h_2 = 12$

- Note: $\mathsf{E}[X_1] + \mathsf{E}[X_2] = \frac{A}{2} + \frac{A}{2} = A \Longrightarrow$ any $\alpha \geq 1$ works!

MISO Example: Peak-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 1.2$$

- $\alpha \geq \frac{n_\mathsf{T}}{2} = \frac{2}{2} = 1$

  $\Longrightarrow$ "beamforming": choose $X_1 = X_2 \triangleq X \sim \text{Uniform}([0, \mathsf{A}])$

  $\Longrightarrow Y = 12X + Z$

  $\Longrightarrow$ like SISO with $h = h_{\text{sum}} = h_1 + h_2 = 12$

- Note: $\mathsf{E}[X_1] + \mathsf{E}[X_2] = \frac{\mathsf{A}}{2} + \frac{\mathsf{A}}{2} = \mathsf{A} \Longrightarrow$ any $\alpha \geq 1$ works!

MISO Example: Peak-Power Constraint Only

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 1.2$$

- $\alpha \geq \frac{n_\mathsf{T}}{2} = \frac{2}{2} = 1$

  $\Longrightarrow$ "beamforming": choose $X_1 = X_2 \triangleq X \sim \mathsf{Uniform}([0, \mathsf{A}])$

  $\Longrightarrow Y = 12X + Z$

  $\Longrightarrow$ like SISO with $h = h_{\mathsf{sum}} = h_1 + h_2 = 12$

- Note: $\mathsf{E}[X_1] + \mathsf{E}[X_2] = \frac{\mathsf{A}}{2} + \frac{\mathsf{A}}{2} = \mathsf{A} \Longrightarrow$ any $\alpha \geq 1$ works!

$\alpha \geq \frac{n_\mathsf{T}}{2}$ **(including peak-power constraint only):**

$$\mathsf{C}_{\mathsf{MISO}}(\mathsf{A}, \alpha\mathsf{A}) = \mathsf{C}_{\mathsf{SISO}}\left(h_{\mathsf{sum}}\mathsf{A}, \frac{h_{\mathsf{sum}}}{2}\mathsf{A}\right)$$

MISO Example: Both Constraints: A First Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- $X_1 = X_2 = X \sim \text{Uniform}([0, A])$ is not possible because of average-power constraint

MISO Example: Both Constraints: A First Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- $X_1 = X_2 = X \sim \text{Uniform}([0, A])$ is not possible because of average-power constraint

- Shall we give full average power of 0.5 to X_1 and the rest of 0.4 to X_2?

$$X_1 \sim \text{Uniform}([0, A]) \qquad X_2 = 0.8X_1 \sim \text{Uniform}([0, 0.8A])$$

MISO Example: Both Constraints: A First Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- $X_1 = X_2 = X \sim \text{Uniform}([0, A])$ is not possible because of average-power constraint

- Shall we give full average power of 0.5 to X_1 and the rest of 0.4 to X_2?

$$X_1 \sim \text{Uniform}([0, A]) \qquad X_2 = 0.8X_1 \sim \text{Uniform}([0, 0.8A])$$

But this means that we do not make full use of peak power on X_2!

MISO Example: Both Constraints: A Second Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- Maybe better:

$$X_1 \sim \text{Uniform}([0, A]) \qquad X_2 \sim \text{truncExp}\big([0, A], 0.4A\big)$$

$\Longrightarrow$ now we make full use of peak and average power

MISO Example: Both Constraints: A Second Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- Maybe better:

$$X_1 \sim \mathsf{Uniform}([0, A]) \qquad X_2 \sim \mathsf{truncExp}\big([0, A], 0.4A\big)$$

$\Longrightarrow$ now we make full use of peak and average power

- But: X_1 and X_2 are not correlated, but $X_1 \perp\!\!\!\perp X_2$ (?)

$\Longrightarrow$ $9X_1 + 3X_2$ has a bad mix-distribution

MISO Example: Both Constraints: A Third Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- If we want full correlation:

$$X_1 = X_2 = X \sim \mathsf{truncExp}\big([0, A], 0.45A\big)$$

MISO Example: Both Constraints: A Third Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- If we want full correlation:

$$X_1 = X_2 = X \sim \mathsf{truncExp}\bigl([0, A], 0.45A\bigr)$$

$\Longrightarrow$ now we make full use of peak and average power & we have full correlation

MISO Example: Both Constraints: A Third Attempt

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- If we want full correlation:

$$X_1 = X_2 = X \sim \mathsf{truncExp}\big([0, A], 0.45A\big)$$

$\Longrightarrow$ now we make full use of peak and average power & we have full correlation

- But: we do not favor X_1 over X_2!

How to Optimize Correctly? (1/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

How to Optimize Correctly? (1/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

Clue: think about energy-efficiency!

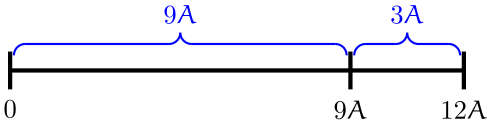
How to Optimize Correctly? (1/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

Clue: think about energy-efficiency!

What is the range that Y can take on (**ignoring noise**)?

$$0 \leq Y \leq (9+3)A = 12A$$



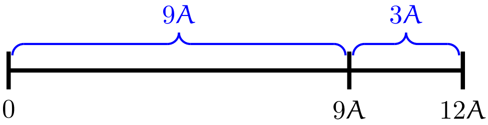
How to Optimize Correctly? (1/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

Clue: think about energy-efficiency!

What is the range that Y can take on (**ignoring noise**)?

$$0 \leq Y \leq (9+3)A = 12A$$



$\Longrightarrow$ maximum $12A$ can only be reached if we make full use of **both** LEDs!

How to Optimize Correctly? (2/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- What is the most energy-efficient way to reach, e.g., $Y = 2A$?

How to Optimize Correctly? (2/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- What is the most energy-efficient way to reach, e.g., $Y = 2A$?

$$9 \cdot \frac{0}{9}A + 3 \cdot \frac{2}{3}A = 2A \qquad \frac{0}{9}A + \frac{2}{3}A = \frac{6}{9}A$$

$$9 \cdot \frac{1}{9}A + 3 \cdot \frac{1}{3}A = 2A \qquad \frac{1}{9}A + \frac{1}{3}A = \frac{4}{9}A$$

$$9 \cdot \frac{2}{9}A + 3 \cdot \frac{0}{3}A = 2A \qquad \frac{2}{9}A + \frac{0}{3}A = \frac{2}{9}A$$

How to Optimize Correctly? (2/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- What is the most energy-efficient way to reach, e.g., $Y = 2A$?

$$\begin{aligned} 9 \cdot \frac{0}{9}A + 3 \cdot \frac{2}{3}A &= 2A & \qquad \frac{0}{9}A + \frac{2}{3}A &= \frac{6}{9}A \\ 9 \cdot \frac{1}{9}A + 3 \cdot \frac{1}{3}A &= 2A & \qquad \frac{1}{9}A + \frac{1}{3}A &= \frac{4}{9}A \\ 9 \cdot \frac{2}{9}A + 3 \cdot \frac{0}{3}A &= 2A & \qquad \frac{2}{9}A + \frac{0}{3}A &= \frac{2}{9}A \end{aligned}$$

$\Longrightarrow$ only use X_1 and set $X_2 = 0$!

How to Optimize Correctly? (2/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- What is the most energy-efficient way to reach, e.g., $Y = 2A$?

$$\begin{aligned} 9 \cdot \frac{0}{9}A + 3 \cdot \frac{2}{3}A &= 2A & \qquad \frac{0}{9}A + \frac{2}{3}A &= \frac{6}{9}A \\ 9 \cdot \frac{1}{9}A + 3 \cdot \frac{1}{3}A &= 2A & \qquad \frac{1}{9}A + \frac{1}{3}A &= \frac{4}{9}A \\ 9 \cdot \frac{2}{9}A + 3 \cdot \frac{0}{3}A &= 2A & \qquad \frac{2}{9}A + \frac{0}{3}A &= \frac{2}{9}A \end{aligned}$$

$\Longrightarrow$ only use X_1 and set $X_2 = 0$!

- What is the most energy-efficient way to reach, e.g., $Y = 10A$?

How to Optimize Correctly? (2/4)

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.9$$

- What is the most energy-efficient way to reach, e.g., $Y = 2A$?

$$\begin{aligned} 9 \cdot \frac{0}{9}A + 3 \cdot \frac{2}{3}A &= 2A & \qquad \frac{0}{9}A + \frac{2}{3}A &= \frac{6}{9}A \\ 9 \cdot \frac{1}{9}A + 3 \cdot \frac{1}{3}A &= 2A & \qquad \frac{1}{9}A + \frac{1}{3}A &= \frac{4}{9}A \\ 9 \cdot \frac{2}{9}A + 3 \cdot \frac{0}{3}A &= 2A & \qquad \frac{2}{9}A + \frac{0}{3}A &= \frac{2}{9}A \end{aligned}$$

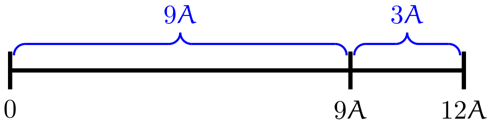
$\Longrightarrow$ only use X_1 and set $X_2 = 0$!

- What is the most energy-efficient way to reach, e.g., $Y = 10A$?

$\Longrightarrow$ set X_1 to full power and get the rest with X_2:

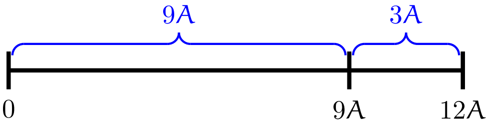
$$9 \cdot 1A + 3 \cdot \frac{1}{3}A = 10A \qquad 1A + \frac{1}{3}A = \frac{4}{3}A$$

How to Optimize Correctly? (3/4)



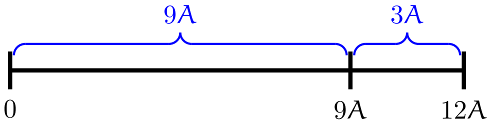
- What is the most efficient way to have $Y \in [0, 9A]$?

How to Optimize Correctly? (3/4)



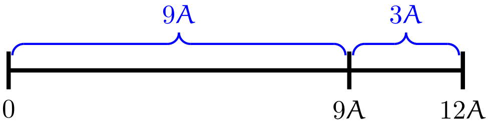
- What is the most efficient way to have $Y \in [0, 9A]$?
  $\Longrightarrow$ only use X_1 for signaling and set $X_2 = 0$!

How to Optimize Correctly? (3/4)



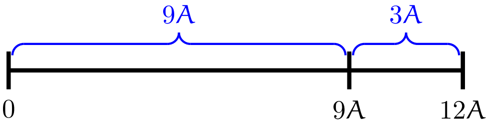
- What is the most efficient way to have $Y \in [0, 9A]$?
  $\Longrightarrow$ only use X_1 for signaling and set $X_2 = 0$!
- What is the most efficient way to have $Y \in [9A, 12A]$?

How to Optimize Correctly? (3/4)



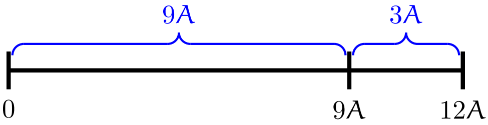
- What is the most efficient way to have $Y \in [0, 9A]$?
  $\Longrightarrow$ only use X_1 for signaling and set $X_2 = 0$!
- What is the most efficient way to have $Y \in [9A, 12A]$?
  $\Longrightarrow$ set $X_1 = A$ and use X_2 for signaling!

How to Optimize Correctly? (3/4)



- What is the most efficient way to have $Y \in [0, 9A]$?
  $\Longrightarrow$ only use X_1 for signaling and set $X_2 = 0$!
- What is the most efficient way to have $Y \in [9A, 12A]$?
  $\Longrightarrow$ set $X_1 = A$ and use X_2 for signaling!
- How to combine?

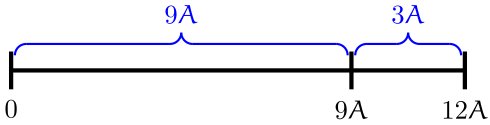
How to Optimize Correctly? (3/4)



- What is the most efficient way to have $Y \in [0, 9A]$?
  $\Longrightarrow$ only use X_1 for signaling and set $X_2 = 0$!
- What is the most efficient way to have $Y \in [9A, 12A]$?
  $\Longrightarrow$ set $X_1 = A$ and use X_2 for signaling!
- How to combine?

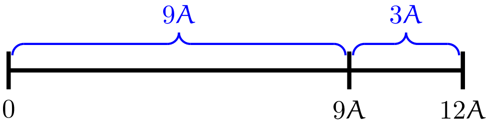
  Note: if possible Y should be uniform!

How to Optimize Correctly? (4/4)



- with probability $\frac{9A}{9A+3A} = \frac{3}{4}$ choose $X_1 \sim \text{Uniform}([0, A])$ and $X_2 = 0$
- with probability $\frac{3A}{9A+3A} = \frac{1}{4}$ choose $X_1 = A$ and $X_2 \sim \text{Uniform}([0, A])$

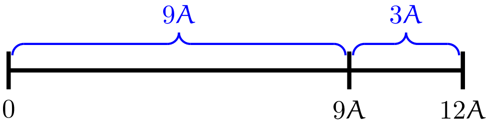
How to Optimize Correctly? (4/4)



- with probability $\frac{9A}{9A+3A} = \frac{3}{4}$ choose $X_1 \sim \text{Uniform}([0, A])$ and $X_2 = 0$
- with probability $\frac{3A}{9A+3A} = \frac{1}{4}$ choose $X_1 = A$ and $X_2 \sim \text{Uniform}([0, A])$

$\Longrightarrow Y \sim \text{Uniform}([0, 12A])$

How to Optimize Correctly? (4/4)



- with probability $\frac{9A}{9A+3A} = \frac{3}{4}$ choose $X_1 \sim \text{Uniform}([0, A])$ and $X_2 = 0$
- with probability $\frac{3A}{9A+3A} = \frac{1}{4}$ choose $X_1 = A$ and $X_2 \sim \text{Uniform}([0, A])$

$\Longrightarrow Y \sim \text{Uniform}([0, 12A])$

- Check:

$$\mathsf{E}[X_1 + X_2] = \frac{3}{4}\left(\frac{A}{2} + 0\right) + \frac{1}{4}\left(A + \frac{A}{2}\right) = \frac{6}{8}A = 0.75A \leq 0.9A$$

$\Longrightarrow$ works for all $\alpha \geq 0.75 \triangleq \alpha_{\text{th}}$

What about $\alpha < \alpha_{\text{th}}$?

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.1$$

- The scheme above does not work: average power too large!

What about $\alpha < \alpha_{\text{th}}$?

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.1$$

- The scheme above does not work: average power too large!
  $\Longrightarrow$ reduce average power in each interval to λA for some λ $(0 < \lambda < \alpha)$

What about $\alpha < \alpha_{\text{th}}$?

$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.1$$

- The scheme above does not work: average power too large!
  $\Longrightarrow$ reduce average power in each interval to λA for some λ $(0 < \lambda < \alpha)$
  $\Longrightarrow$ instead of uniform, use truncExp:

$$\begin{array}{lll} \text{with prob. } 1-(\alpha-\lambda): & X_1 \sim \mathsf{truncExp}\big([0,A],\lambda A\big) & X_2 = 0 \\ \text{with prob. } \alpha-\lambda: & X_1 = A & X_2 \sim \mathsf{truncExp}\big([0,A],\lambda A\big) \end{array}$$

What about $\alpha < \alpha_{\text{th}}$?

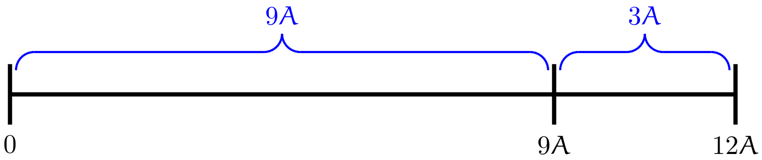
$$Y = 9X_1 + 3X_2 + Z \qquad \text{with } \alpha = 0.1$$

- The scheme above does not work: average power too large!
  $\Longrightarrow$ reduce average power in each interval to λA for some λ $(0 < \lambda < \alpha)$
  $\Longrightarrow$ instead of uniform, use truncExp:

$$\begin{array}{lll} \text{with prob. } 1-(\alpha-\lambda)\text{:} & X_1 \sim \text{truncExp}\big([0,A],\lambda A\big) & X_2 = 0 \\ \text{with prob. } \alpha-\lambda\text{:} & X_1 = A & X_2 \sim \text{truncExp}\big([0,A],\lambda A\big) \end{array}$$

- Note:
  - choice of prob. $\alpha - \lambda$ results in $\mathsf{E}[X_1 + X_2] = \alpha A$
  - we need to numerically optimize over λ

Principle: Find Most Energy-Efficient Signaling (1/2)

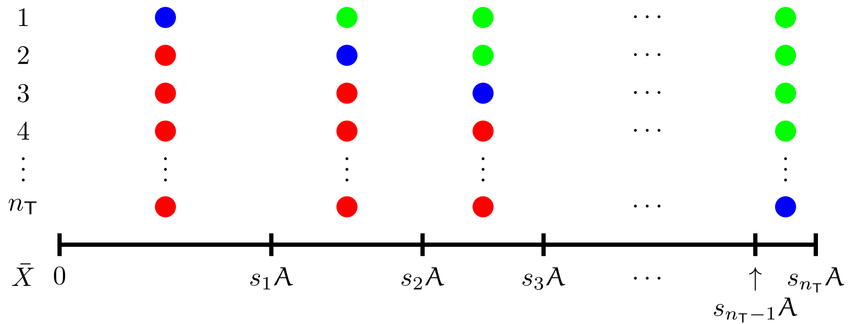


- To have $\bar{X} \triangleq 9X_1 + 3X_2$ operate in the interval $[0, 9A]$, it is most energy-efficient to set $X_2 = 0$
- To have $\bar{X} = 9X_1 + 3X_2$ operate in the interval $[9A, 12A]$, it is most energy-efficient to set $X_1 = A$

Principle: Find Most Energy-Efficient Signaling (2/2)

Let $h_1 \geq h_2 \geq \cdots \geq h_{n_\mathsf{T}}$ be ordered and define

$$s_k \triangleq \sum_{i=1}^{k} h_i \qquad \bar{X} \triangleq \sum_{k=1}^{n_\mathsf{T}} h_k X_k$$



Extension to Rank-Deficient MIMO: $n_\mathsf{T} > n_\mathsf{R}$

$$
\begin{aligned}
\mathbf{Y} &= \mathsf{H}\mathbf{X} + \mathbf{Z} \\
&= \begin{bmatrix} \mathbf{h}_1 \; \mathbf{h}_2 \; \cdots \; \mathbf{h}_{n_\mathsf{T}} \end{bmatrix} \mathbf{X} + \mathbf{Z} \\
&= \mathbf{h}_1 X_1 + \mathbf{h}_2 X_2 + \cdots + \mathbf{h}_{n_\mathsf{T}} X_{n_\mathsf{T}} + \mathbf{Z}
\end{aligned}
$$

- Linear combination of vectors $\mathbf{h}_1, \ldots, \mathbf{h}_{n_\mathsf{T}}$ with factors $0 \leq X_i \leq A$!

Extension to Rank-Deficient MIMO: $n_\mathsf{T} > n_\mathsf{R}$

$$\begin{aligned}
\mathbf{Y} &= \mathsf{H}\mathbf{X} + \mathbf{Z} \\
&= \begin{bmatrix} \mathbf{h}_1 \; \mathbf{h}_2 \; \cdots \; \mathbf{h}_{n_\mathsf{T}} \end{bmatrix} \mathbf{X} + \mathbf{Z} \\
&= \mathbf{h}_1 X_1 + \mathbf{h}_2 X_2 + \cdots + \mathbf{h}_{n_\mathsf{T}} X_{n_\mathsf{T}} + \mathbf{Z}
\end{aligned}$$

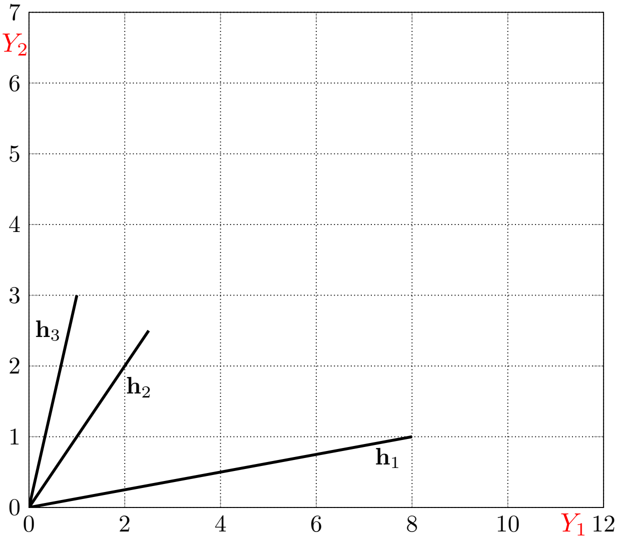
- Linear combination of vectors $\mathbf{h}_1, \ldots, \mathbf{h}_{n_\mathsf{T}}$ with factors $0 \leq X_i \leq A$!
- "Ordering" of antennas depending on their gain not trivial anymore

Extension to Rank-Deficient MIMO: $n_\mathsf{T} > n_\mathsf{R}$

$$\begin{aligned}\mathbf{Y} &= \mathsf{H}\mathbf{X} + \mathbf{Z} \\ &= \begin{bmatrix}\mathbf{h}_1 \ \mathbf{h}_2 \ \cdots \ \mathbf{h}_{n_\mathsf{T}}\end{bmatrix}\mathbf{X} + \mathbf{Z} \\ &= \mathbf{h}_1 X_1 + \mathbf{h}_2 X_2 + \cdots + \mathbf{h}_{n_\mathsf{T}} X_{n_\mathsf{T}} + \mathbf{Z}\end{aligned}$$

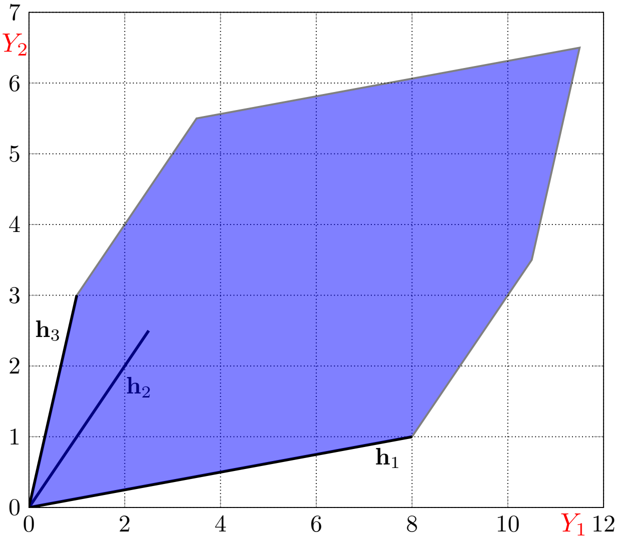
- Linear combination of vectors $\mathbf{h}_1, \ldots, \mathbf{h}_{n_\mathsf{T}}$ with factors $0 \leq X_i \leq A$!
- "Ordering" of antennas depending on their gain not trivial anymore
- What values can $\mathbf{Y}$ achieve (**ignoring noise**)?

MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$): Parallelepiped



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

MIMO Example ($n_{\mathsf{T}} = 3$, $n_{\mathsf{R}} = 2$): Parallelepiped



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

Energy-Optimal Signaling

- Basic idea the same as before: every point in the parallelepiped can be reached in a **most energy-efficient** way:
  - "weak" antennas switched off
  - "strong" antennas switched to full power A
  - n_{R} antennas chosen for signaling

Energy-Optimal Signaling

- Basic idea the same as before: every point in the parallelepiped can be reached in a **most energy-efficient** way:
  - "weak" antennas switched off
  - "strong" antennas switched to full power A
  - n_{R} antennas chosen for signaling

- The big questions is:

Which are the strong and which are the weak antennas?

Energy-Optimal Signaling

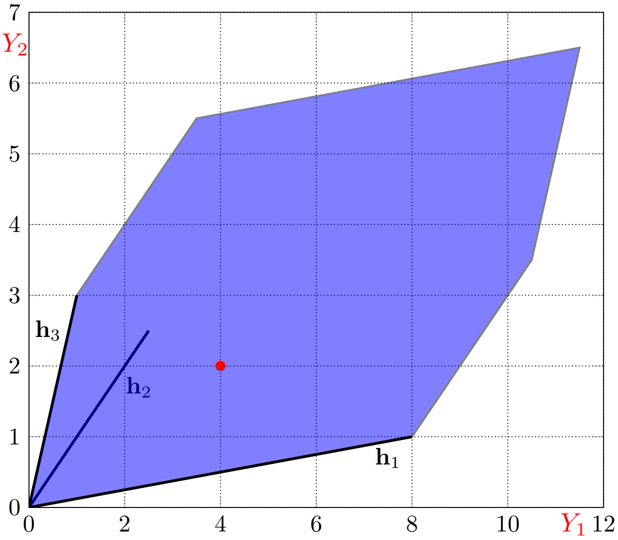
- Basic idea the same as before: every point in the parallelepiped can be reached in a **most energy-efficient** way:
  - "weak" antennas switched off
  - "strong" antennas switched to full power A
  - n_R antennas chosen for signaling

- The big questions is:

Which are the strong and which are the weak antennas?

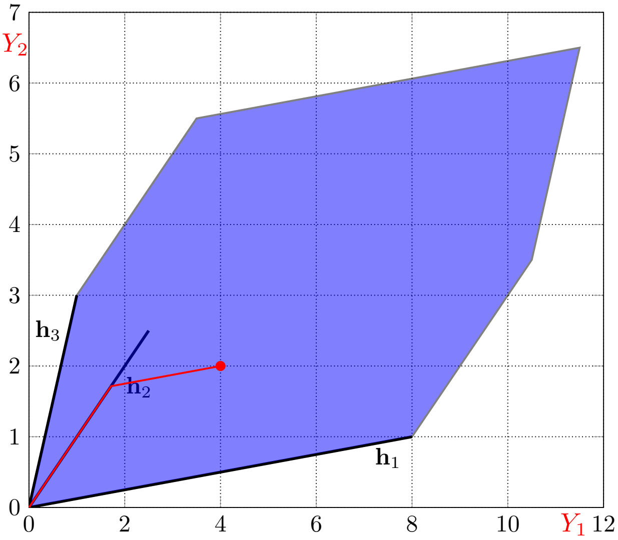
$\Longrightarrow$ depends on position of point!

MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



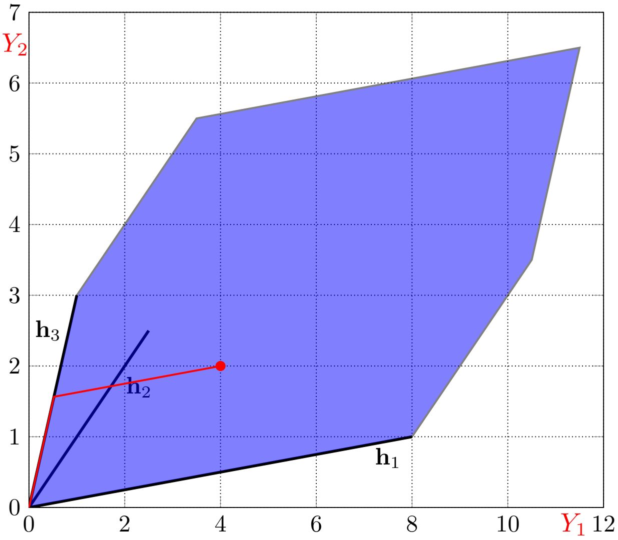
$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

MIMO Example ($n_{\mathsf{T}} = 3$, $n_{\mathsf{R}} = 2$)



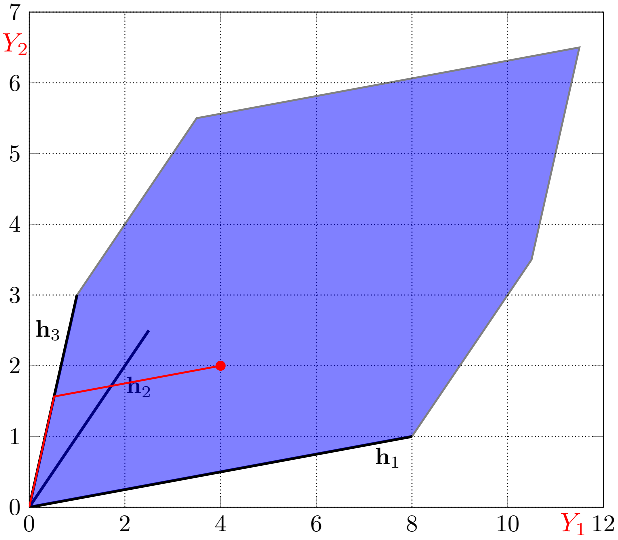
$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

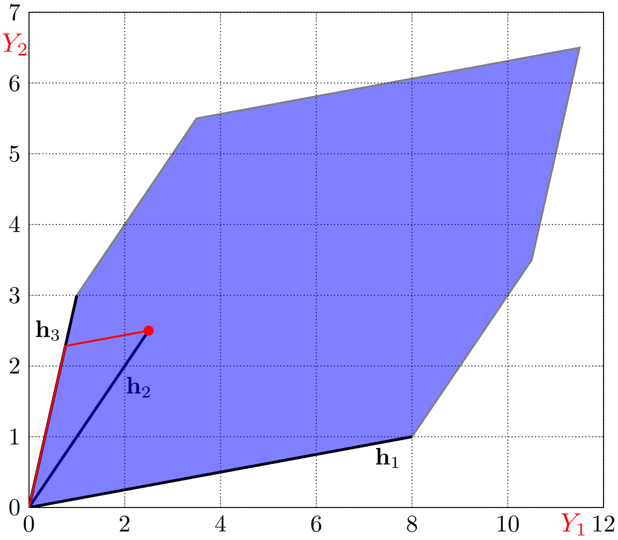
MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

this second way is more efficient!

MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)

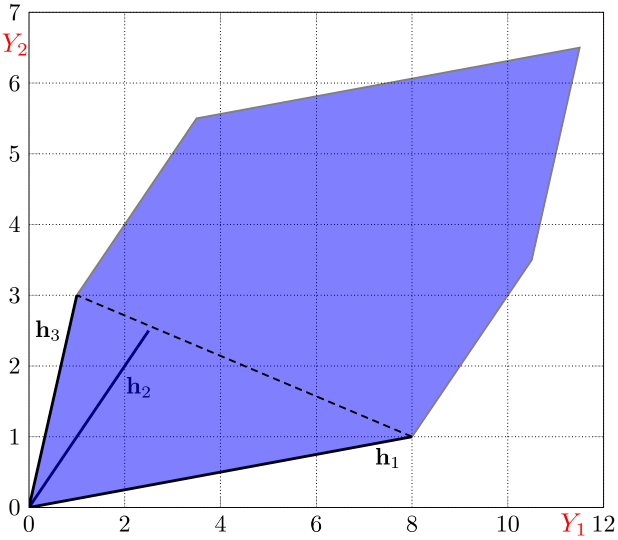


$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

red path uses power $0.22A + 0.76A = 0.98A$

direct path uses power $1A$

MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)

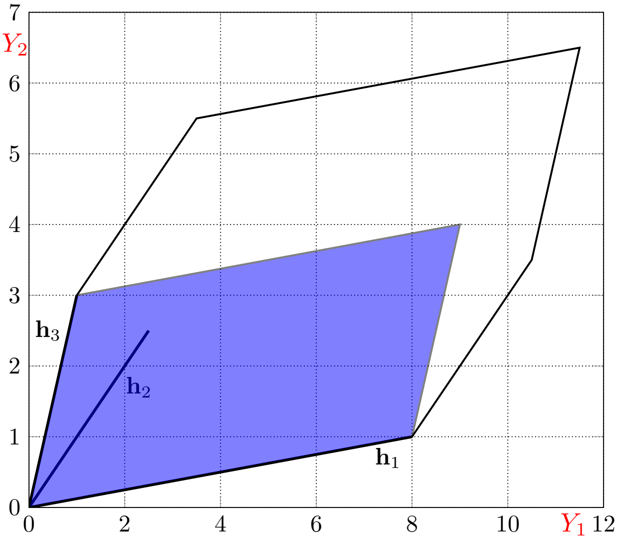


$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

red path uses power $0.22A + 0.76A = 0.98A$
direct path uses power $1A$

$\Longrightarrow$ $\mathbf{h}_2$ is weaker!

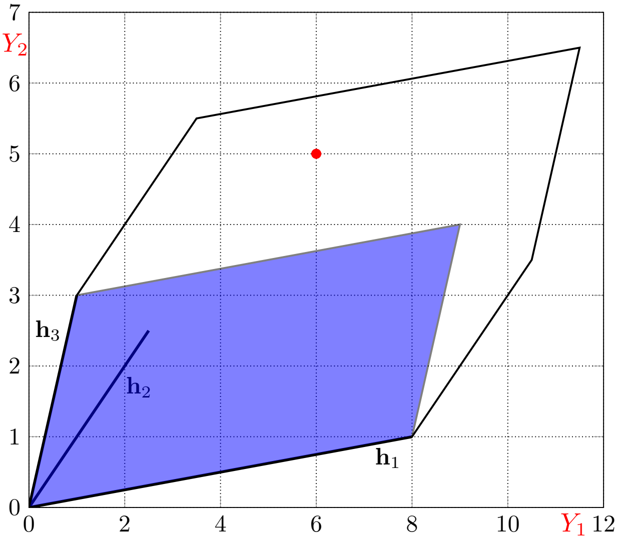
MIMO Example ($n_T = 3$, $n_R = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

$\Longrightarrow$ first switch X_2 off!

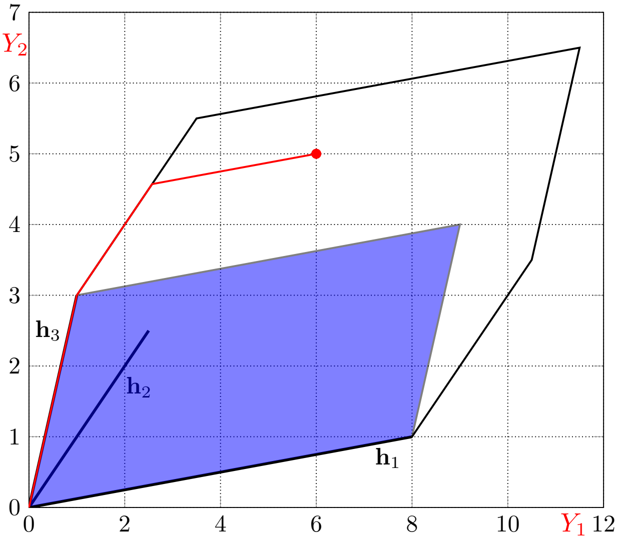
MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

$\Longrightarrow$ first switch X_2 off!

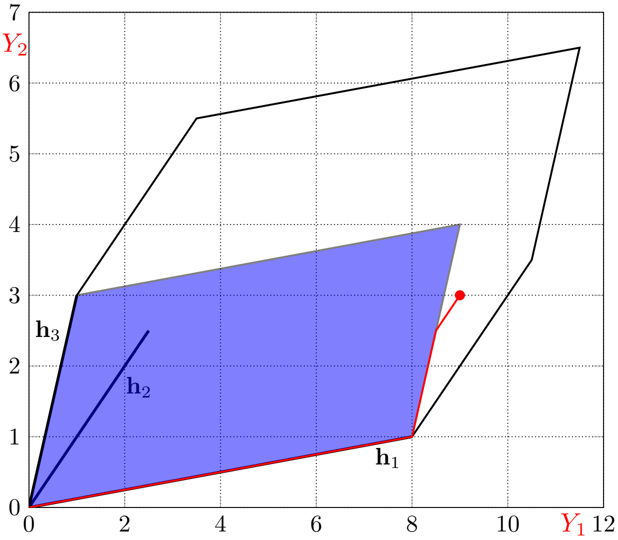
MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

$\Longrightarrow$ set X_3 to A!

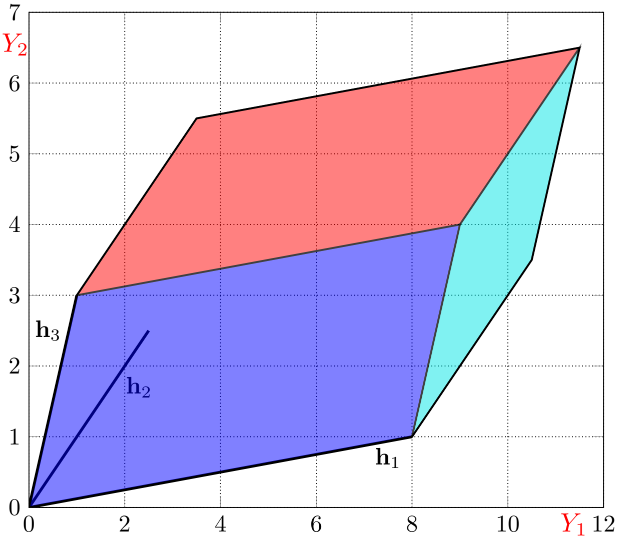
MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

$\Longrightarrow$ set X_1 to A!

MIMO Example ($n_{\mathsf{T}} = 3$, $n_{\mathsf{R}} = 2$)



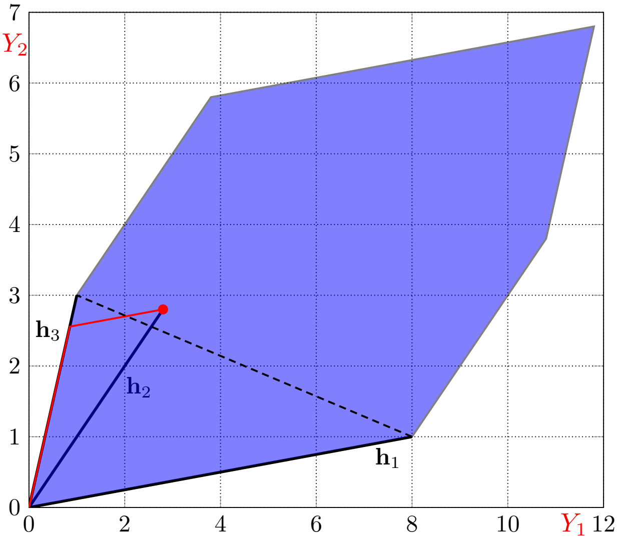
$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

blue area: $X_2 = 0$
use X_1, X_3

red area: $X_3 = A$
use X_1, X_2

cyan area: $X_1 = A$
use X_2, X_3

Another MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



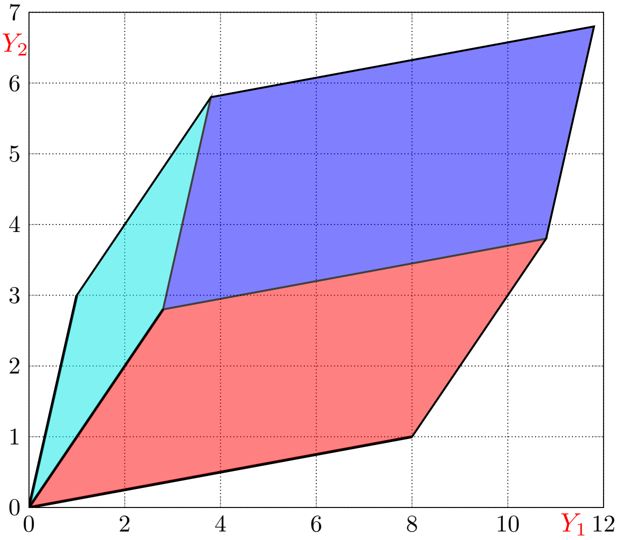
$$\mathsf{H} = \begin{bmatrix} 8 & 2.8 & 1 \\ 1 & 2.8 & 3 \end{bmatrix}$$

red path uses power $0.24A + 0.85A = 1.09A$

direct path uses power $1A$

$\Longrightarrow$ $\mathbf{h}_2$ is stronger!

Another MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



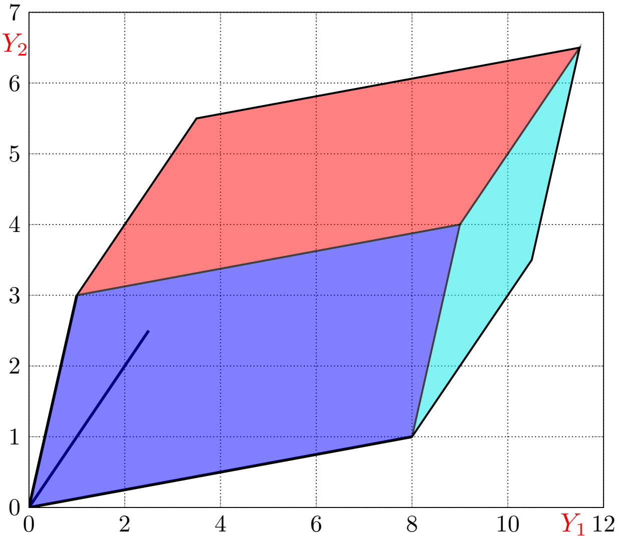
$$\mathsf{H} = \begin{bmatrix} 8 & 2.8 & 1 \\ 1 & 2.8 & 3 \end{bmatrix}$$

red area: $X_3 = 0$
use X_1, X_2

cyan area: $X_1 = 0$
use X_2, X_3

blue area: $X_2 = A$
use X_1, X_3

Another MIMO Example ($n_\mathsf{T} = 3$, $n_\mathsf{R} = 2$)



$$\mathsf{H} = \begin{bmatrix} 8 & 2.5 & 1 \\ 1 & 2.5 & 3 \end{bmatrix}$$

blue area: $X_2 = 0$
use X_1, X_3

red area: $X_3 = A$
use X_1, X_2

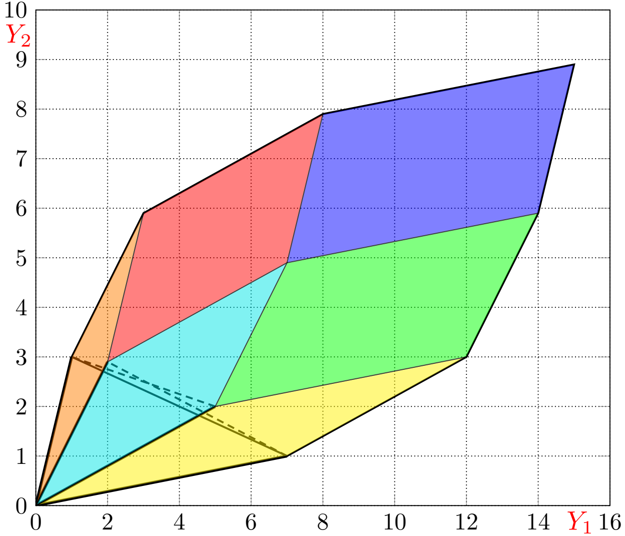
cyan area: $X_1 = A$
use X_2, X_3

Choose Distribution on Energy-Optimal Signaling

For high power:

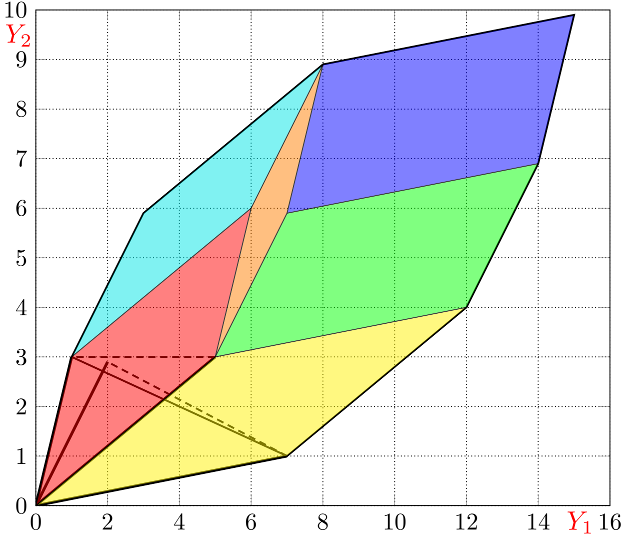
- if $\alpha \geq \alpha_{\text{th}}$: use uniform distribution on each parallelogram
- if $\alpha < \alpha_{\text{th}}$: use generalized truncExp distribution on each parallelogram

MIMO with $n_\mathsf{T} = 4$ and $n_\mathsf{R} = 2$: Example 1



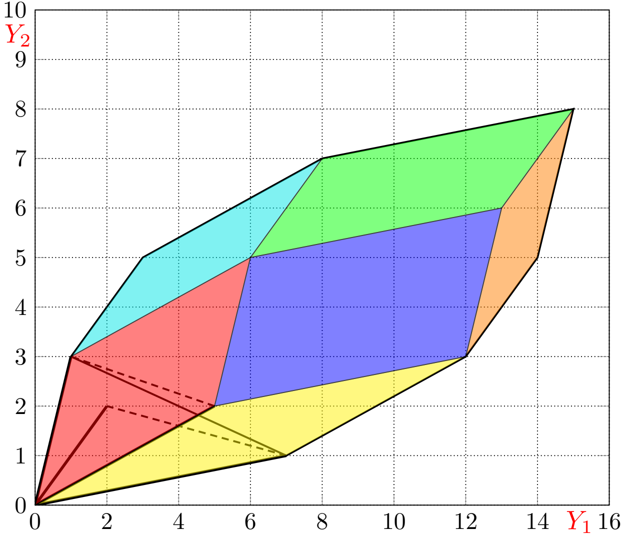
$$\mathsf{H} = \begin{bmatrix} 7 & 5 & 2 & 1 \\ 1 & 2 & 2.9 & 3 \end{bmatrix}$$

MIMO with $n_\mathsf{T} = 4$ and $n_\mathsf{R} = 2$: Example 2



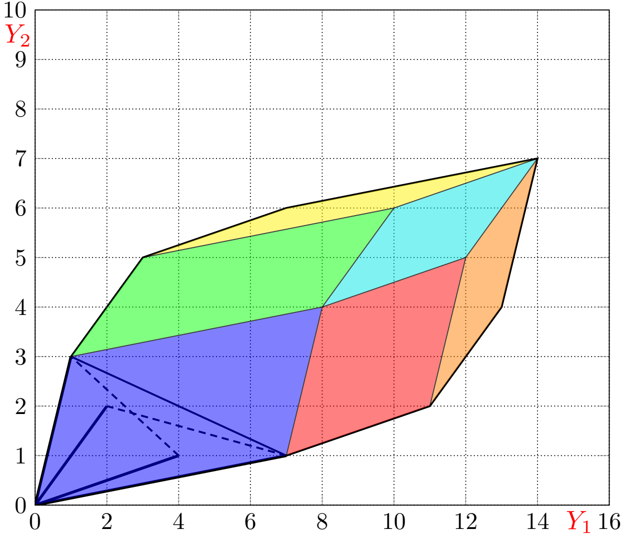
$$\mathsf{H} = \begin{bmatrix} 7 & 5 & 2 & 1 \\ 1 & 3 & 2.9 & 3 \end{bmatrix}$$

MIMO with $n_\mathsf{T} = 4$ and $n_\mathsf{R} = 2$: Example 3



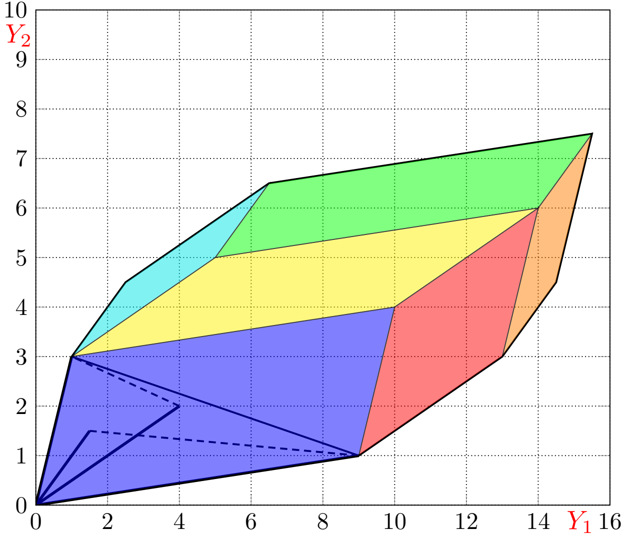
$$\mathsf{H} = \begin{bmatrix} 7 & 5 & 2 & 1 \\ 1 & 2 & 2 & 3 \end{bmatrix}$$

MIMO with $n_\mathsf{T} = 4$ and $n_\mathsf{R} = 2$: Example 4



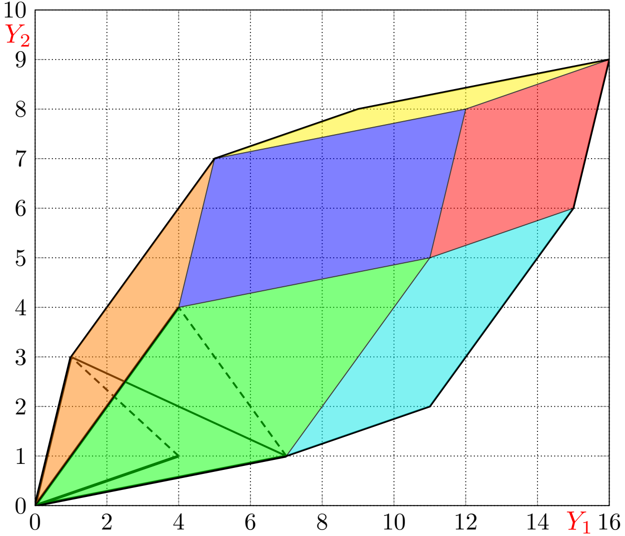
$$\mathsf{H} = \begin{bmatrix} 7 & 4 & 2 & 1 \\ 1 & 1 & 2 & 3 \end{bmatrix}$$

MIMO with $n_\mathsf{T} = 4$ and $n_\mathsf{R} = 2$: Example 5



$$\mathsf{H} = \begin{bmatrix} 9 & 4 & 1.5 & 1 \\ 1 & 2 & 1.5 & 3 \end{bmatrix}$$

MIMO with $n_\mathsf{T} = 4$ and $n_\mathsf{R} = 2$: Example 6



$$\mathsf{H} = \begin{bmatrix} 7 & 4 & 4 & 1 \\ 1 & 1 & 4 & 3 \end{bmatrix}$$

Summary and Comments

- **Optimal signaling:** when thinking about capacity, firstly one needs to figure out the **most energy-efficient signaling**

Summary and Comments

- **Optimal signaling:** when thinking about capacity, firstly one needs to figure out the **most energy-efficient signaling**

- For optical MIMO free-space communication: with certain probability
  - pick n_R antennas for signaling
  - set the other antennas to 0 or A

$\Longrightarrow$ this splits achievable parallelepiped into parallelograms

Summary and Comments

- **Optimal signaling:** when thinking about capacity, firstly one needs to figure out the **most energy-efficient signaling**
- For optical MIMO free-space communication: with certain probability
  - pick n_{R} antennas for signaling
  - set the other antennas to 0 or A

  $\Longrightarrow$ this splits achievable parallelepiped into parallelograms
- This signaling tells also a lot about concrete use of system

Summary and Comments

- **Optimal signaling:** when thinking about capacity, firstly one needs to figure out the **most energy-efficient signaling**
- For optical MIMO free-space communication: with certain probability
  - pick n_{R} antennas for signaling
  - set the other antennas to 0 or A

  $\Longrightarrow$ this splits achievable parallelepiped into parallelograms
- This signaling tells also a lot about concrete use of system
- We have focused on high SNR (ignoring noise), but principle is general

Summary and Comments

- **Optimal signaling:** when thinking about capacity, firstly one needs to figure out the **most energy-efficient signaling**

- For optical MIMO free-space communication: with certain probability
  - pick n_{R} antennas for signaling
  - set the other antennas to 0 or A

  $\Longrightarrow$ this splits achievable parallelepiped into parallelograms

- This signaling tells also a lot about concrete use of system

- We have focused on high SNR (ignoring noise), but principle is general

- The nonnegativity constraint is not crucial for this discussion

Summary and Comments

- **Optimal signaling:** when thinking about capacity, firstly one needs to figure out the **most energy-efficient signaling**

- For optical MIMO free-space communication: with certain probability
  - pick n_{R} antennas for signaling
  - set the other antennas to 0 or A

  $\Longrightarrow$ this splits achievable parallelepiped into parallelograms

- This signaling tells also a lot about concrete use of system

- We have focused on high SNR (ignoring noise), but principle is general

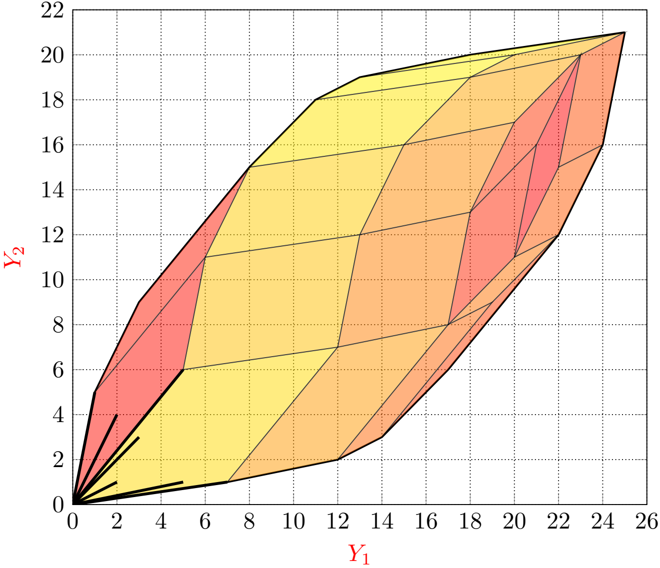
- The nonnegativity constraint is not crucial for this discussion

- We have used the same approach also in the case when one has both first- and second-moment constraints

References

- Stefan M. Moser, Ligong Wang, Michèle Wigger: *Capacity Results on Multiple-Input Single-Output Wireless Optical Channels*, IEEE Transactions on Information Theory, vol. 64, no. 11, pp. 6954–6966, November 2018.
- Longguang Li, Stefan M. Moser, Ligong Wang, Michèle Wigger: "On the Capacity of MIMO Optical Wireless Channels", *IEEE Transactions on Information Theory*, vol. 66, no. 9, pp. 5660–5682, September 2020.
- Shuai Ma, Stefan M. Moser, Ligong Wang, Michèle Wigger: "Signaling for MISO Channels Under First- and Second-Moment Constraints", in Proceedings *2022 IEEE International Symposium on Information Theory (ISIT'22)*, Helsinki, Finland, Jun. 26 – Jul. 1, 2022, pp. 2648–2653.

MIMO Example with $n_\mathsf{T} = 7$ and $n_\mathsf{R} = 2$



Orthogonalization of H ($n_\mathsf{R} \geq n_\mathsf{T}$) with $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathsf{K})$

$$\begin{aligned}
\mathrm{I}(\mathbf{X}; \mathsf{H}\mathbf{X} + \mathbf{Z}) &= \mathrm{I}(\mathbf{X}; \mathsf{S}^{-\mathsf{T}}\mathsf{H}\mathbf{X} + \mathsf{S}^{-\mathsf{T}}\mathbf{Z}) && \mathsf{K} = \mathsf{S}^\mathsf{T}\mathsf{S} \\
&= \mathrm{I}(\mathbf{X}; \mathsf{S}^{-\mathsf{T}}\mathsf{H}\mathbf{X} + \tilde{\mathbf{Z}}) && \tilde{\mathbf{Z}} \sim \text{IID} \\
&= \mathrm{I}(\mathbf{X}; \mathsf{U}\Sigma\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}) && \text{SVD: } \mathsf{S}^{-\mathsf{T}}\mathsf{H} = \mathsf{U}\Sigma\mathsf{V} \\
&= \mathrm{I}(\mathbf{X}; \Sigma\mathsf{V}\mathbf{X} + \mathsf{U}^\mathsf{T}\tilde{\mathbf{Z}}) && \\
&= \mathrm{I}(\mathbf{X}; \Sigma\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}) && \Sigma = \begin{pmatrix} \Sigma_{n_\mathsf{T}} \\ 0 \end{pmatrix} \\
&= \mathrm{I}\big(\mathbf{X}; \Sigma_{n_\mathsf{T}}\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}^{(n_\mathsf{T})}, \tilde{Z}_{n_\mathsf{T}+1}, \ldots, \tilde{Z}_{n_\mathsf{R}}\big) && \tilde{Z}_{n_\mathsf{T}+1}^{n_\mathsf{R}} \text{ indep.} \\
&= \mathrm{I}\big(\mathbf{X}; \Sigma_{n_\mathsf{T}}\mathsf{V}\mathbf{X} + \tilde{\mathbf{Z}}^{(n_\mathsf{T})}\big) && \Sigma_{n_\mathsf{T}}\mathsf{V} \text{ is square} \\
&= \mathrm{I}\big(\mathbf{X}; \tilde{\mathsf{H}}\mathbf{X} + \tilde{\mathbf{Z}}^{(n_\mathsf{T})}\big) && \text{parallel channels} \\
&= \mathrm{I}\big(\mathbf{X}; \mathbf{X} + \mathbf{Z}'^{(n_\mathsf{T})}\big) && \mathbf{Z}'^{(n_\mathsf{T})} \sim \mathcal{N}\big(\mathbf{0}, (\mathsf{H}^\mathsf{T}\mathsf{K}^{-1}\mathsf{H})^{-1}\big)
\end{aligned}$$

$\Longrightarrow$ after transformation redundant receiver antennas are ignored

$\Longrightarrow$ reduced to $n_\mathsf{T} \times n_\mathsf{T}$ square case with parallel channels