The Structure of the Worst Noise in Gaussian Vector Broadcast Channels

Wei Yu

University of Toronto

March 19, 2003

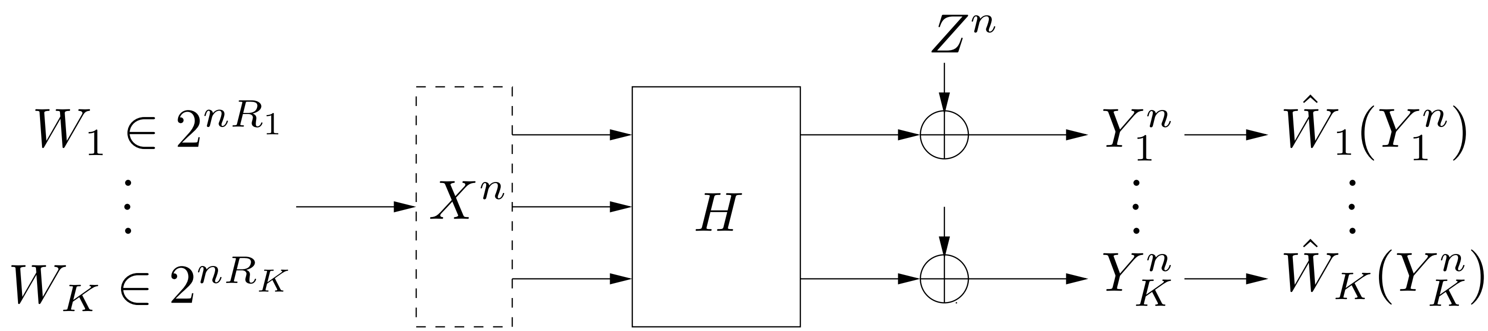
DIMACS Workshop on Network Information Theory

Outline

- Sum capacity of Gaussian vector broadcast channels.
- Complete characterization of the worst-noise.
- Efficient numerical solution for the dual channel.
- Does duality extend beyond the power constrained channels?

Gaussian Vector Broadcast Channel

- Non-degraded broadcast channel:



- Capacity region is still unknown.
  - Sum capacity $C = \max\{R_1 + \cdots + R_K\}$ is recently solved.

Marton's Achievability Region

- For a broadcast channel $p(y_1, y_2|x)$:

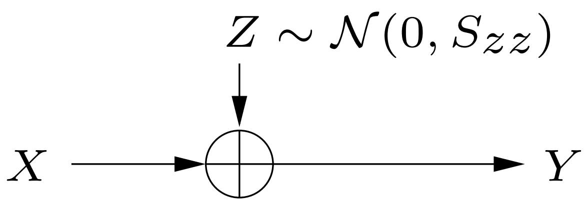
$$\begin{aligned} R_1 &\leq I(U_1; Y_1) \\ R_2 &\leq I(U_2; Y_2) \\ R_1 + R_2 &\leq I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2) \end{aligned}$$

  for some auxiliary random variables $p(u_1, u_2)p(x|u_1, u_2)$.

- For the Gaussian broadcast channel:
  $I(U_2; Y_2) - I(U_1; U_2)$ is achieved with precoding.

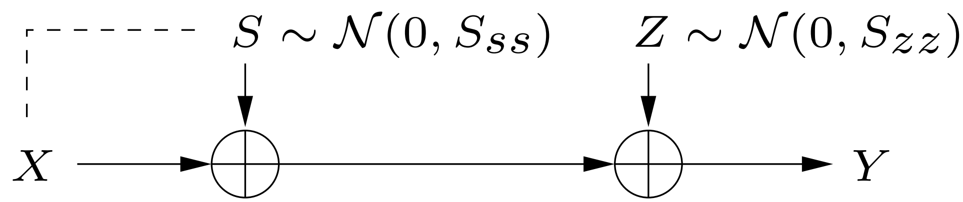
Writing on Dirty Paper

Gaussian Channel



$$C = \frac{1}{2} \log \frac{|S_{xx} + S_{zz}|}{|S_{zz}|}$$

... with Transmitter Side Information

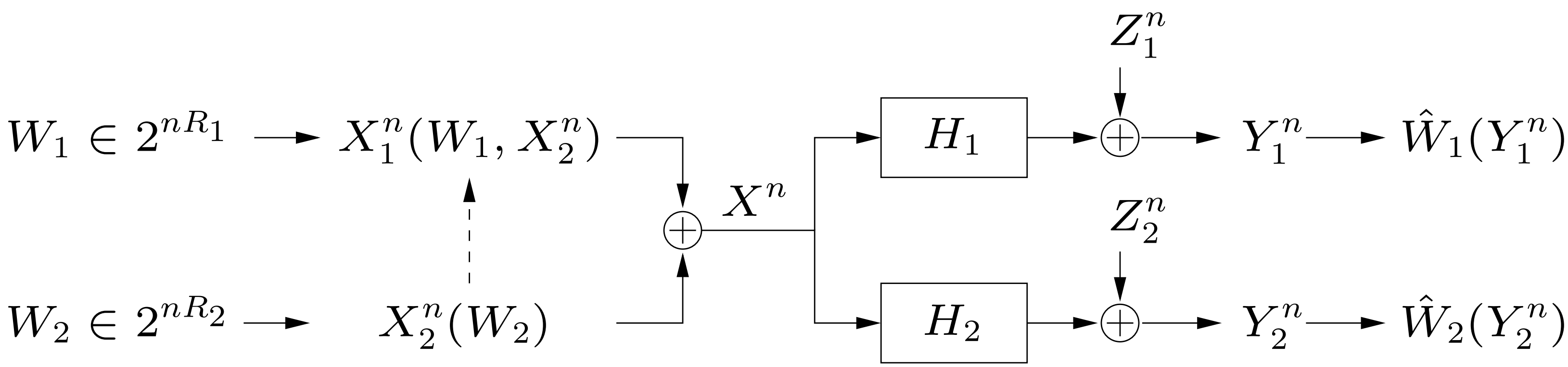


$$C = \frac{1}{2} \log \frac{|S_{xx} + S_{zz}|}{|S_{zz}|}$$

- Capacities are the same if S is known *non-causally* at the transmitter.

$$C = \max_{p(u,x|s)} I(U;Y) - I(U;S) = \max_{p(x)} I(X;Y|S)$$

Precoding for the Broadcast Channel

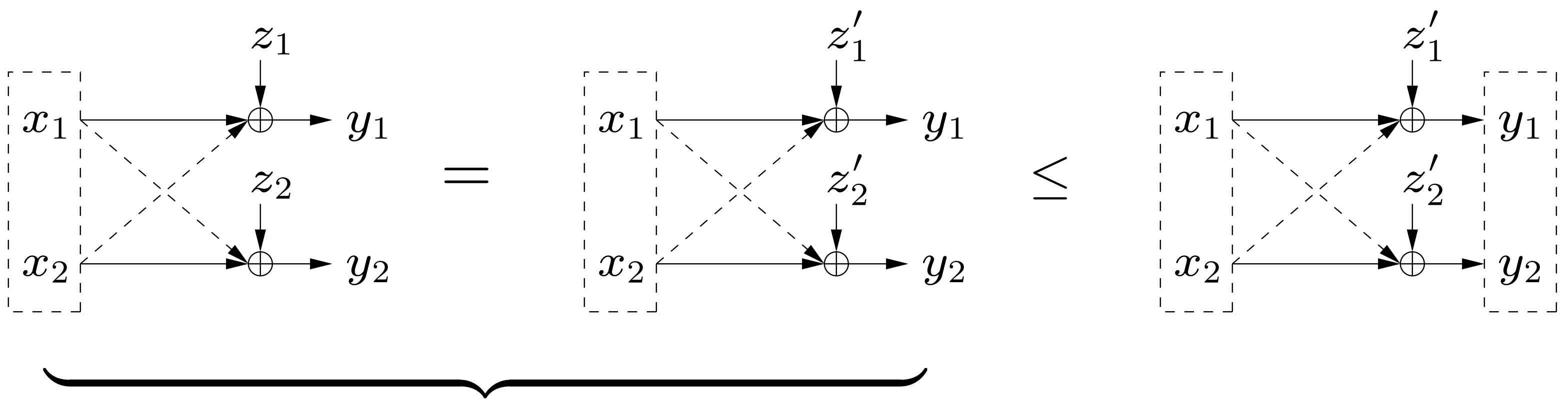


$$R_1 = I(\mathbf{X_1}; \mathbf{Y_1} | \mathbf{X_2}) = \frac{1}{2} \log \frac{|H_1 S_1 H_1^T + S_{z_1 z_1}|}{|S_{z_1 z_1}|}$$

$$R_2 = I(\mathbf{X_2}; \mathbf{Y_2}) = \frac{1}{2} \log \frac{|H_2 S_2 H_2^T + H_2 S_1 H_2^T + S_{z_2 z_2}|}{|H_2 S_1 H_2^T + S_{z_2 z_2}|}$$

Converse: Sato's Outer Bound

- Broadcast capacity does not depend on noise correlation: Sato ('78).



if $\begin{cases} p(z_1) = p(z_1') \\ p(z_2) = p(z_2') \end{cases}$, not necessarily $p(z_1, z_2) = p(z_1', z_2')$.

- So, sum capacity $C \leq \min\limits_{S_{zz}} \max\limits_{S_{xx}} I(\mathbf{X}; \mathbf{Y})$.

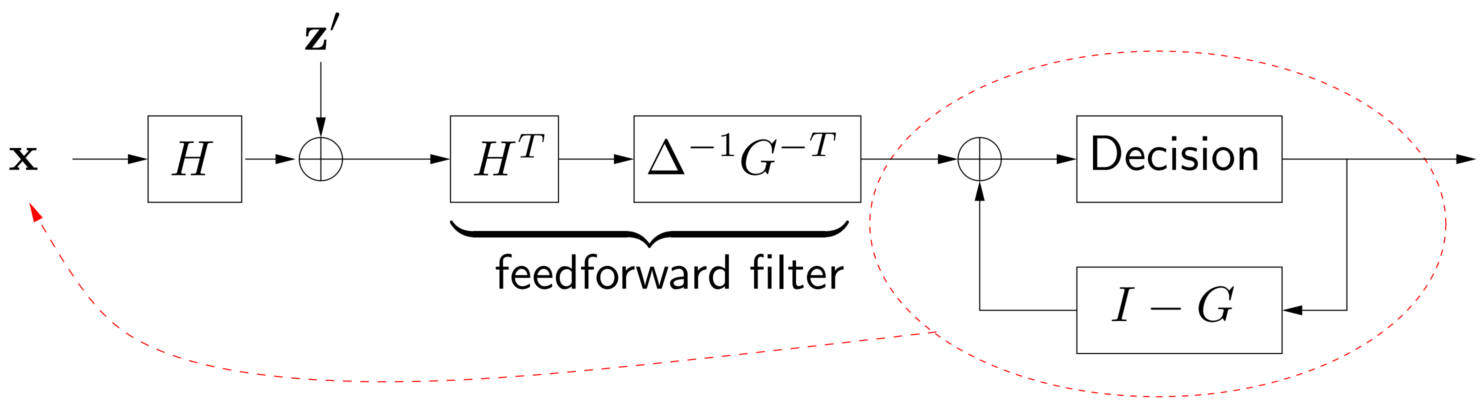
Three Proofs of the Sum Capacity Result

1. Decision-Feedback Equalization approach (Yu, Cioffi)

2. Uplink-Downlink duality approach (Viswanath, Tse)

3. Convex duality approach (Jindal, Vishwanath, Goldsmith)

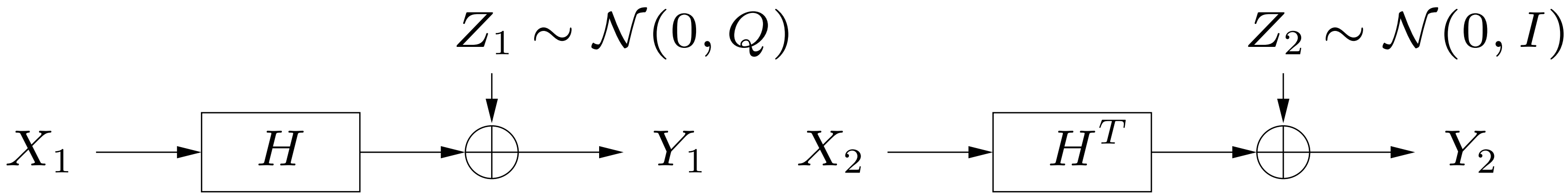
DFE Approach



- Decision-feedback at the receiver is equivalent to transmitter precoding.
- (Non-Singular) Worst Noise $\Longleftrightarrow$ Diagonal feedforward filter

$$\text{Fix } S_{xx}, \quad \min_{S_{zz}} I(\mathbf{X};\mathbf{Y}) \text{ is achievable.}$$

Uplink-Downlink Duality Approach

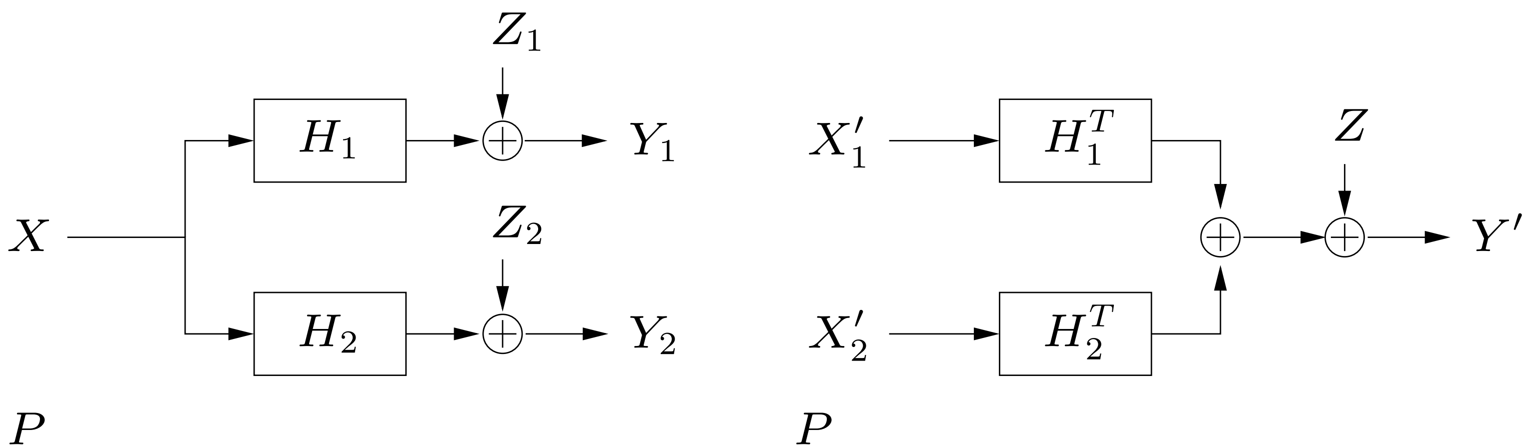


$\mathbf{E}[X_1^T X_1] \leq P$ $\qquad$ $\mathbf{E}[X_2^T Q X_2] \leq P$

- Uplink and downlink channels are duals.
- The noise covariance and input constraint are duals.
- Worst-noise gives an input constraint that decouples the inputs.

$$C = \max_{S_{xx}} \min_{S_{zz}} I(\mathbf{X}; \mathbf{Y})$$

Convex Duality Approach



- Sato's bound: $C \leq \min_{S_{zz}} \max_{S_{xx}} I(\mathbf{X};\mathbf{Y})$.
- Broadcast/Multiple-Access duality: $C \geq \max_{S_{x'x'}} I(\mathbf{X}';\mathbf{Y}')$.
- Convex duality: $\max_{S_{xx}} \min_{S_{zz}} I(\mathbf{X};\mathbf{Y}) = \max_{S_{x'x'}} I(\mathbf{X}';\mathbf{Y}')$.

Objective

- Completely characterize the worst-noise.
  - Duality through minimax.
  - Worst-noise through duality.
- Efficient numerical solution for the dual channel.
- Does duality extend beyond the power constrained channel?

Minimax Capacity

- Gaussian vector broadcast channel sum capacity is the solution of

$$
\begin{aligned}
\max_{S_{xx}} \min_{S_{zz}} \quad & \frac{1}{2} \log \frac{|H S_{xx} H^T + S_{zz}|}{|S_{zz}|} \\
\text{subject to} \quad & \mathrm{tr}(S_{xx}) \leq P \\
& S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix} \\
& S_{xx}, S_{zz} \geq 0
\end{aligned}
$$

- The minimax problem is **convex** in S_{zz}, **concave** in S_{xx}.
  - How to solve this minimax problem?

Duality through Minimax

- Two KKT conditions must be satisfied simultaneously:

$$H^T(HS_{xx}H^T + S_{zz})^{-1}H = \lambda I$$

$$S_{zz}^{-1} - (HS_{xx}H^T + S_{zz})^{-1} = \begin{bmatrix} \Psi_1 & 0 \\ 0 & \Psi_2 \end{bmatrix}$$

- For the moment, assume that H is invertible.

$$\Rightarrow \quad H^T S_{zz}^{-1} H - \lambda I = H^T \Psi H$$

$$\Rightarrow \quad H(H^T \Psi H + \lambda I)^{-1} H^T = S_{zz}$$

This is a "water-filling" condition for the dual channel.

Power Constraint in the Dual Channel

- Interpretation of dual variable: $\lambda = \frac{\partial C}{\partial P}, \Psi_i = -\frac{\partial C}{\partial S_{z_i z_i}}$.
  - Thus, capacity is preserved if $\lambda \Delta P = \left(\sum_i \Psi_i \right) \Delta S_{z_i z_i}$
- Capacity $C = \min \max \frac{1}{2} \log \frac{|H S_{xx} H^T + S_{zz}|}{|S_{zz}|}$.
  - Thus, capacity is preserved if $\frac{\Delta P}{P} = \frac{\Delta S_{z_i z_i}}{1}$.

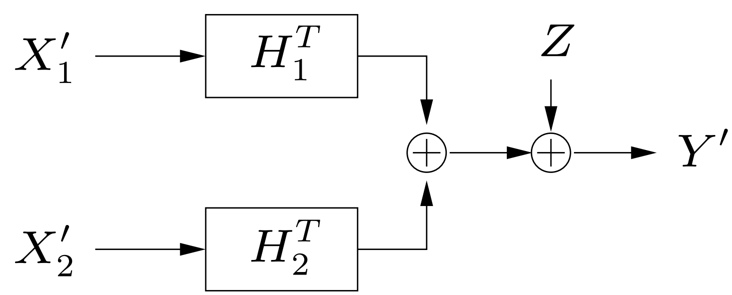
Therefore, $\frac{\sum_i \Psi_i}{\lambda} = P$.

Construct the Dual Channel

KKT condition: $H(H^T D H + I)^{-1} H^T = \frac{1}{\lambda} S_{zz}$

- where $D = \Psi/\lambda$ is diagonal, $\mathrm{trace}(D) = \sum_i \Psi_i/\lambda = P$.
- $S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix}$. Thus, constraint on D: $\mathrm{trace}(D_1) + \mathrm{trace}(D_2) \le P$.

$$\mathbf{E}[X_1' X_1'^T] = D_1$$
$$\mathbf{E}[X_2' X_2'^T] = D_2$$
$$\mathrm{trace}(D_1) + \mathrm{trace}(D_2) \le P$$



Yet Another Derivation for Duality

The duality between broadcast channel and multiple-access channel:

$$\max_{S_{xx}} \min_{S_{zz}} \quad \frac{1}{2} \log \frac{|H S_{xx} H^T + S_{zz}|}{|S_{zz}|}$$

$$\text{s.t.} \quad \mathrm{tr}(S_{xx}) \leq P$$

$$S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix}$$

$$S_{xx}, S_{zz} \geq 0$$

$$\max_{D} \quad \frac{1}{2} \log \frac{|H^T D H + I|}{|I|}$$

$$\text{s.t.} \quad \mathrm{tr}(D) \leq P$$

$$D \;\text{ is diagonal}$$

$$D \geq 0$$

KKT conditions for minimax $\Longrightarrow$ KKT condition for max.

Worst-Noise Through Minimax

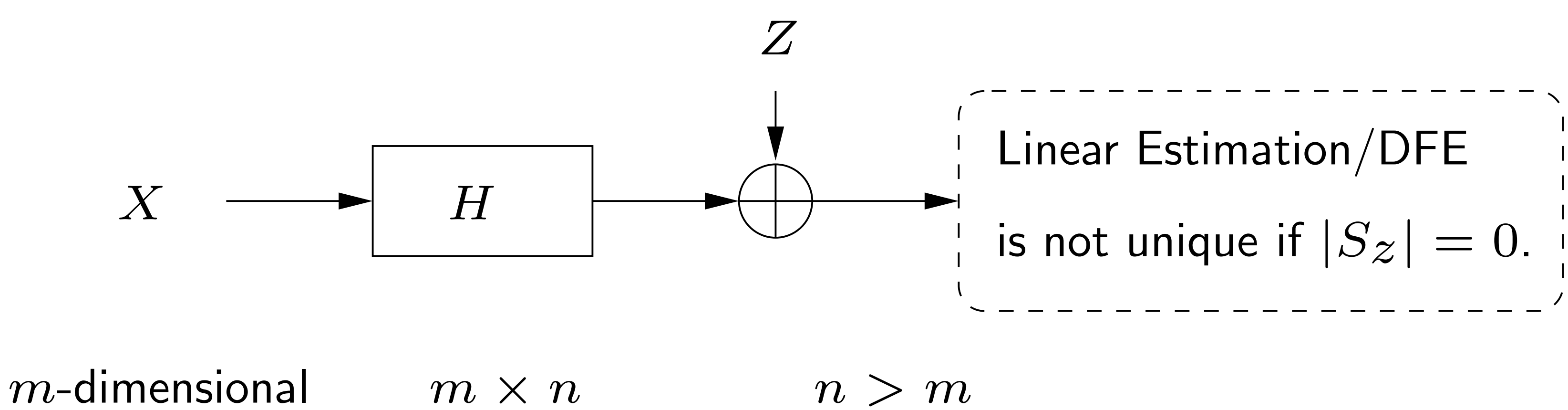
- Solve the dual multiple access channel problem with power constraint P. Obtain (Ψ, λ). Then:

$$\begin{aligned} S_{zz} &= H(H^T \Psi H + \lambda I)^{-1} H^T \\ S_{xx} &= (\lambda I)^{-1} - (H^T \Psi H + \lambda I)^{-1} \end{aligned}$$

- What if H is not invertible, or S_{zz} is singular?

Decision-Feedback Equalization with Singular Noise

- With non-singular noise: $S_{zz}^{-1} - (HS_{xx}H^T + S_{zz})^{-1} = \begin{bmatrix} \Psi_1 & 0 \\ 0 & \Psi_2 \end{bmatrix}$.
- If H is low-rank, S_{zz} can be singular.



Necessary and Sufficient Condition for Diagonalization

- Suppose that the worst-noise $|S_{zz}| = 0$, let

$$S_{zz} = U S_{\tilde{z}\tilde{z}} U^T,$$

where S_{zz} is $n \times n$, $S_{\tilde{z}\tilde{z}}$ is $m \times m$, $m < n$.

- It is always possible to write $H = U\tilde{H}$.

- There exists a DFE with diagonal feedforward filter if and only if

$$S_{\tilde{z}\tilde{z}}^{-1} - (\tilde{H} S_{xx} \tilde{H}^T + S_{\tilde{z}\tilde{z}})^{-1} = U^T \begin{bmatrix} \Psi_1 & 0 \\ 0 & \Psi_2 \end{bmatrix} U$$

Singular Worst-Noise

- It can be verified that the diagonalization condition is satisfied by:

$$\begin{aligned} S_{zz}^{(0)} &= H(H^T \Psi H + \lambda I)^{-1} H^T \\ S_{xx} &= (\lambda I)^{-1} - (H^T \Psi H + \lambda I)^{-1} \end{aligned}$$

- However: $S_{zz}^{(0)}$ does not necessarily have 1's on the diagonal.

$$S_{zz}^{(0)} = \begin{bmatrix} I & \star & \star \\ \star & I & \star \\ \star & \star & \star \end{bmatrix}.$$

Characterization of the Worst-Noise

Theorem 1. *The following steps solve the worst noise in $y = Hx + z$:*

1. *Find the optimal (Ψ, λ) in the dual multiple access channel.*
2. *Form $S_{zz}^{(0)} = H(H^T \Psi H + \lambda I)^{-1} H^T$,*
   $S_{xx} = (\lambda I)^{-1} - (H^T \Psi H + \lambda I)^{-1}$.
3. *If S_{xx} is not full rank, reduce the rank of H, and repeat 1-2.*
4. *The class of worst-noise is precisely $S_{zz}^{(0)} + S'_{zz}$.*

$$\begin{bmatrix} I & \star & \star \\ \star & I & \star \\ \star & \star & \star \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \star \end{bmatrix} = \begin{bmatrix} I & \star & \star \\ \star & I & \star \\ \star & \star & I \end{bmatrix}.$$

Worst-Noise is Not Unique

- The same S_{xx} water-fills the entire class of $S_{zz}^{(0)} + S'_{zz}$.
- $S_{zz}^{(0)} + \begin{bmatrix} 0 & 0 \\ 0 & S'_{zz} \end{bmatrix} = [U|U'] \left(\begin{bmatrix} S_{\tilde{z}\tilde{z}} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} S'_{11} & S'_{12} \\ S'_{21} & S'_{22} \end{bmatrix} \right) [U|U']^T$,
  - where $S'_{11} - S'_{12} S'^{-1}_{22} S'_{21} = 0$.
  - The entire class of worst-noise is related by linear estimation:

$$\mathbf{E}[\tilde{z} + z'_1 | z'_2] = \tilde{z}.$$

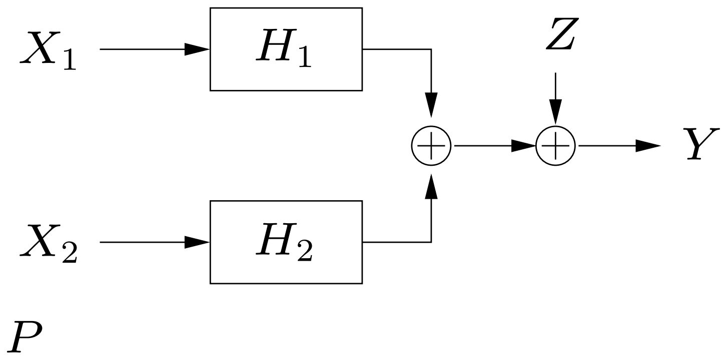
- The class of (S_{xx}, S_{zz}) that satisfies the KKT condition is precisely:

$$(S_{xx}, S_{zz}^{(0)} + S'_{zz})$$

Outline

- Complete characterization of the worst-noise.
  - Duality through minimax.
  - Worst-noise through duality.
- Efficient numerical solution for the dual channel.
- Does duality extend beyond the power constrained channel?

Sum Power Gaussian Vector Multiple Access Channel



$$\max_{S_{xx}} \quad \frac{1}{2}\log|H^T S_{xx} H + I|$$
$$\text{s.t.} \quad \mathrm{tr}(S_{xx}) \le P$$
$$S_{xx} \quad \text{is diagonal}$$
$$S_{xx} \ge 0$$

- An efficient way to find the worst-noise is to solve the dual problem.
  - Previous numerical solution: Jindal, Jafar, Vishwanath, Goldsmith.

Iterative Water-filling

- Iterative water-filling: Optimize each of S_i while fixing all others.

$$\max_{S_i} \quad \frac{1}{2}\log\left|\sum_i H_i S_i H_i^T + I\right| \qquad \text{s.t.} \quad \mathrm{tr}(S_i) \le P_i, \quad S_i \ge 0$$

Individual Constraints

$$\max_{S_i} \quad \frac{1}{2}\log\left|\sum_i H_i S_i H_i^T + I\right| \qquad \text{s.t.} \quad \sum_i \mathrm{tr}(S_i) \le P, \quad S_i \ge 0$$

Coupled Constraint

- Iterative water-filling only works with the individual power constraints.

Dual Decomposition for the Sum-Power Problem

Take Lagrangian dual with respect to the coupled constraint only:

$$\begin{aligned} \max \quad & \frac{1}{2}\log\left|\sum_i H_i S_i H_i^T + I\right| \\ \text{s.t.} \quad & \sum_i P_i \le P \\ & \text{tr}(S_i) \le P_i \\ & S_i \ge 0 \end{aligned}$$

$$\begin{aligned} g(\nu) = \max \quad & \frac{1}{2}\log\left|\sum_i H_i S_i H_i^T + I\right| \\ & - \nu\left(\sum_i P_i - P\right) \\ \text{s.t.} \quad & \text{tr}(S_i) \le P_i \\ & S_i \ge 0 \end{aligned}$$

$$\boxed{\text{Sum Power Capacity} = \min_{\nu > 0} g(\nu)}$$

Iterative Water-filling for the Dual Problem

- By introducing a Lagrange multiplier ν, constraints are decoupled:

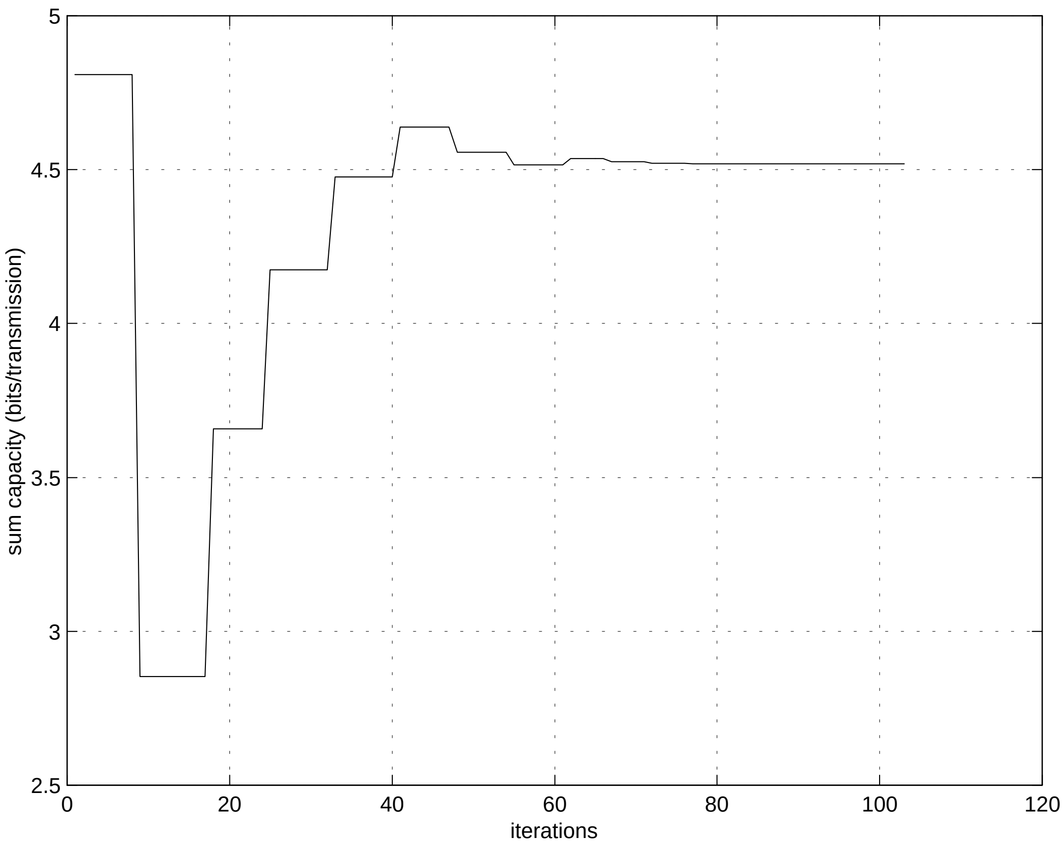
$$
\begin{aligned}
g(\nu) = \quad & \max_{S_i} \quad \frac{1}{2}\log\left|\sum_i H_i S_i H_i^T + I\right| - \nu\left(\sum_i P_i - P\right) \\
& \text{s.t.} \quad \operatorname{tr}(S_i) \le P_i \\
& \qquad\quad S_i \ge 0
\end{aligned}
$$

  - To solve $g(\nu)$: Iteratively optimize each of (S_i, P_i).
  - To find $\min g(\nu)$ over $\nu > 0$:

Decrease ν if $\sum_i P_i < P$. Increase ν if $\sum_i P_i > P$.

Convergence of the Dual Decomposition Algorithm

- 3 transmit antennas
- 50 receivers each with a single antenna
  - typically 3-6 active
- i.i.d. Gaussian channel
- Bisection on ν.



Outline

- Complete characterization of the worst-noise.
  - Duality through minimax.
  - Worst-noise through duality.
- Efficient numerical solution for the dual channel.
- Does duality extend beyond the power constrained channel?

Broadcast Channel under Linear Covariance Constraint

- The DFE achievability result works with any fixed S_{xx}.
- The capacity of the broadcast channel under covariance constraint:

$$
\begin{aligned}
\max_{S_{xx}} \min_{S_{zz}} \quad & \frac{1}{2} \log \frac{|H S_{xx} H^T + S_{zz}|}{|S_{zz}|} \\
\text{subject to} \quad & \operatorname{tr}(Q S_{xx}) \leq P \\
& S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix} \\
& S_{xx}, S_{zz} \geq 0
\end{aligned}
$$

- What is the duality result in this case?

KKT Condition for Minimax

- Two KKT conditions must be satisfied simultaneously:

$$H^T(HS_{xx}H^T + S_{zz})^{-1}H = \lambda Q$$

$$S_{zz}^{-1} - (HS_{xx}H^T + S_{zz})^{-1} = \begin{bmatrix} \Psi_1 & 0 \\ 0 & \Psi_2 \end{bmatrix}$$

- For simplicity, assume invertible H.

$$H(H^T\Psi H + \lambda Q)^{-1}H^T = S_{zz}$$

$$\text{with} \quad \frac{\sum_i \operatorname{tr}(\Psi_i)}{\lambda} = P$$

Duality under Linear Covariance Constraint

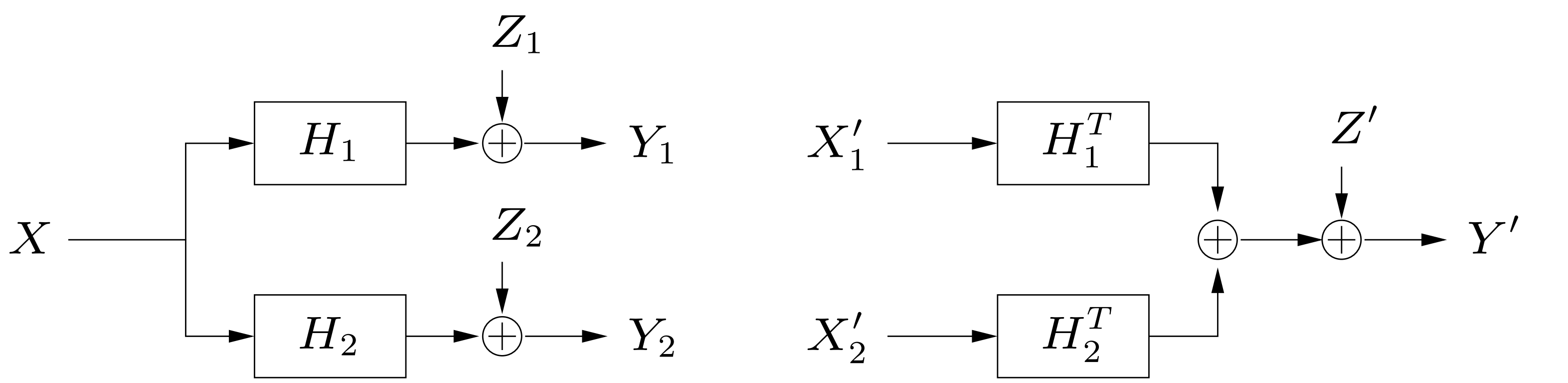
The duality between broadcast channel and multiple-access channel:

$$\begin{aligned} \max_{S_{xx}} \min_{S_{zz}} \quad & \frac{1}{2}\log\frac{|HS_{xx}H^T+S_{zz}|}{|S_{zz}|} \\ \text{s.t.} \quad & \mathrm{tr}(QS_{xx}) \leq P \\ & S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix} \\ & S_{xx}, S_{zz} \geq 0 \end{aligned}$$

$$\begin{aligned} \max_{D} \quad & \frac{1}{2}\log\frac{|H^TDH+Q|}{|Q|} \\ \text{s.t.} \quad & \mathrm{tr}(D) \leq P \\ & D \ \text{ is diagonal} \\ & D \geq 0 \end{aligned}$$

The above two problems have the same KKT conditions.

Generalized Duality



$\text{tr}(S_{xx}Q_1) \leq P \quad S_{zz} \sim \mathcal{N}(0, Q_2) \quad \text{tr}(S_{x'x'}Q_2) \leq P \quad S_{z'z'} \sim \mathcal{N}(0, Q_1)$

Q_1: Input constraint in BC and Noise covariance in MAC.
Q_2: Worst noise covariance in BC and Input constraint in MAC.

Broadcast Channel under Convex Covariance Constraint

- Under arbitrary convex constraint, DFE still works.

$$
\begin{aligned}
\max_{S_{xx}} \min_{S_{zz}} \quad & \frac{1}{2} \log \frac{|H S_{xx} H^T + S_{zz}|}{|S_{zz}|} \\
\text{subject to} \quad & f(S_{xx}) \leq P \\
& S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix} \\
& S_{xx}, S_{zz} \geq 0
\end{aligned}
$$

Does duality exist in this case?

Duality under Convex Covariance Constraint

Duality still exists, but the values of the dual variables are not known:

$$\begin{aligned} \max_{S_{xx}} \min_{S_{zz}} \quad & \frac{1}{2} \log \frac{|H S_{xx} H^T + S_{zz}|}{|S_{zz}|} \\ \text{s.t.} \quad & f(S_{xx}) \leq P \\ & S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix} \\ & S_{xx}, S_{zz} \geq 0 \end{aligned}$$

$$\begin{aligned} \max_{D} \quad & \frac{1}{2} \log \frac{|H^T \Psi H + \lambda Q|}{|\lambda Q|} \\ \text{s.t.} \quad & \operatorname{tr}(\Psi) \leq P' \\ & D \ \text{ is diagonal} \\ & D \geq 0 \end{aligned}$$

$Q = f'(\cdot)$. But if $f(\cdot)$ is non-linear, $\operatorname{tr}(\Psi) \neq \lambda P$.

Peak Power Constrained Broadcast Channel

- Duality exists, but not computationally useful. Need to solve minimax.

$$
\begin{aligned}
\max_{S_{xx}} \min_{S_{zz}} \quad & \frac{1}{2} \log \frac{|HS_{xx}H^T + S_{zz}|}{|S_{zz}|} \\
\text{s.t.} \quad & S_{xx}(i,i) \leq P_i \\
& S_{zz} = \begin{bmatrix} I & \star \\ \star & I \end{bmatrix} \\
& S_{xx}, S_{zz} \geq 0
\end{aligned}
\qquad
\begin{aligned}
\max_{D} \quad & \frac{1}{2} \log \frac{|H^T \Psi H + Q|}{|Q|} \\
\text{s.t.} \quad & \mathrm{tr}(\Psi) \leq P' \\
& D \ \text{ is diagonal} \\
& D \geq 0
\end{aligned}
$$

- Here, $Q = \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{bmatrix}$. But, μ_i, P' are not known.

Concluding Remarks

- Sum capacity of a Gaussian vector broadcast channel is:

$$C = \max_{S_{xx}} \min_{S_{zz}} \frac{1}{2} \log \frac{|HS_{xx}H^T + S_{zz}|}{|S_{zz}|}$$

- If the input constraint is a linear covariance constraint:

$$C = \max_{D} \frac{1}{2} \log \frac{|H^T DH + Q|}{|Q|}$$

- Minimax is a more fundamental expression than duality.

- Duality, when exists, has computational advantage.