Capacity of Multi-antenna Gaussian Channel

Wei Zhang

Department of Electrical Engineering

University of Notre Dame

Notre Dame, IN 46556

Multiuser Communication Class Presentation, April 3rd, 2003

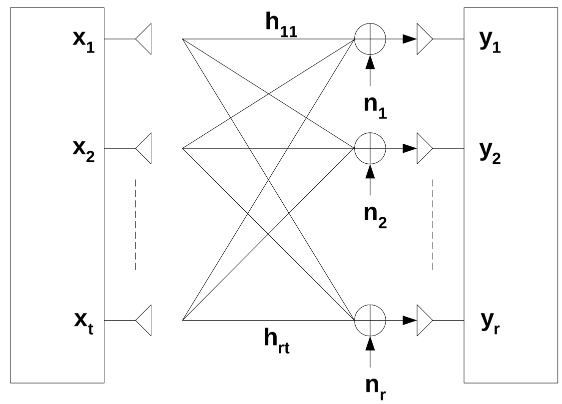
OUTLINE OF THE PRESENTATION

- Introduction of MIMO channel
- Capacity for fixed channel transfer function
- Capacity of the Rayleigh fading channel
- Capacity of the non-ergodic channel

INTRODUCTION OF MIMO CHANNEL

- MIMO is multi-input and multi-output system.
- MIMO vs. SISO
  - Additional dimension to carry information
  - Capacity gain
  - Hot research spot
- Multi-antenna Gaussian channel for a single user
  - Joint processing at the transmitter/receiver

Model of Multi-antenna Gaussian channel



- $\mathbf{y} = H\mathbf{x} + \mathbf{n}$
- $\mathbf{y} \in \mathbb{C}^r$,$H \in \mathcal{C}^{r\times t}$, $\mathbf{x} \in \mathbb{C}^t$
- $\mathbf{n}$ is the noise vector
  - Zero-mean complex Gaussian noise
  - Independent real and imaginary part
  - Covariance $\mathcal{E}[\mathbf{n}\mathbf{n}^\dagger] = \mathbf{I}_r$

Capacity of Multi-antenna Gaussian channel

- Total power constraint P to the transmitter antennas

$$\mathcal{E}[\mathbf{x}^\dagger \mathbf{x}] \leq P$$

- What's the capacity of this channel?
  - H is a deterministic matrix;
  - H is a ergodic random matrix;
  - H is random, but is fixed once it is chosen(Non-ergodic).

Channel with fixed transfer function

- H is deterministic. (Fixed wireless communication, such as WLL etc.)
- Usual sense Shannon capacity

$$C(H,P) = \max_{p(\mathbf{x}):\, \mathcal{E}[\mathbf{x}^\dagger \mathbf{x}] \le P} I(\mathbf{x}, \mathbf{y})$$

- Correlated parallel Gaussian channel
  - Crosstalk from every transmitter $y_i = \sum_j H_{ij} x_j + n_i$
  - Linear method to suppress(cancel) the crosstalk:

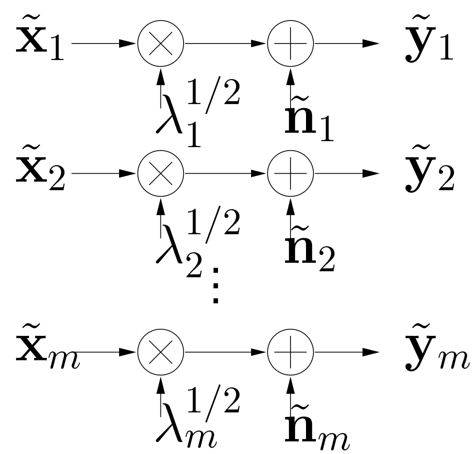
    $H = UDV^\dagger$ (Singular Value Decomposition)

    $\mathbf{y} = UDV^\dagger \mathbf{x} + \mathbf{n} \Rightarrow \tilde{\mathbf{y}} = D\tilde{\mathbf{x}} + \tilde{\mathbf{n}}$

Independent Parallel Gaussian Channel Capacity

- Since D is diagonal, the parallel channels are decorrelated.

  $\tilde{y}_i = \lambda_i^{1/2}\tilde{x}_i + \tilde{n}_i, \quad 1 \leq i \leq m = \min(r,t)$



- Water-filling solution: choose $\tilde{x}_i$'s independent Gaussian with independent zero-mean real and imaginary part.

  $\mathcal{E}[\mathcal{R}e(\tilde{x}_i)^2] = \mathcal{E}[\mathcal{I}m(\tilde{x}_i)^2] = \frac{1}{2}(\mu - \lambda_i^{-1})^+;$

  $P(\mu) = \sum_i(\mu - \lambda_i^{-1})^+;$

  $C(\mu) = \sum_i[\ln(\mu\lambda_i)]^+.$

Circularly Symmetric Gaussian Random Vector

- Definition: A Gaussian random vector $\mathbf{x}$ is circularly symmetric, if for $\hat{\mathbf{x}} = [\mathcal{R}e[\mathbf{x}]\ \mathcal{I}m[\mathbf{x}]]^{\mathcal{T}}$,

$$cov[\hat{\mathbf{x}}] = \frac{1}{2}\begin{bmatrix} \mathcal{R}e[Q] & -\mathcal{I}m[Q] \\ \mathcal{I}m[Q] & \mathcal{R}e[Q] \end{bmatrix}, \quad \text{where } Q = cov[\mathbf{x}].$$

- Given covariance matrix Q, circularly symmetric Gaussian random vector is entropy maximizer.

$$H(\gamma_Q) = \log\det(\pi e Q)$$

CORRELATED PARALLEL GAUSSIAN CAPACITY

- $I(\mathbf{x};\mathbf{y}) = h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x}) = h(\mathbf{y}) - h(\mathbf{n})$
  - Covariance $\mathcal{E}[\mathbf{y}\mathbf{y}^\dagger] = I_r + HQH^\dagger$
  - Choose circularly symmetric Gaussian to maximize $I(\mathbf{x};\mathbf{y})$
- $I(\mathbf{x};\mathbf{y}) = \log\det[I_r + HQH^\dagger] = \log\det[I_r + QH^\dagger H]$

$$H^\dagger H = U^\dagger \Lambda U \text{ (eigenvalue decomposition)}$$

$$\begin{aligned} I(\mathbf{x};\mathbf{y}) &= \log\det[I_t + \Lambda^{1/2} U Q U^\dagger \Lambda^{1/2}] \\ &= \log\det[I_t + \Lambda^{1/2}\tilde{Q}\Lambda^{1/2}] \\ &\leq \log \prod_i (1 + \tilde{Q}_{ii}\lambda_i) \end{aligned}$$

The equality holds when $\tilde{Q}$ is diagonal.

- Water-filling to the diagonal elements $\tilde{Q}_{ii}$

Gaussian channel with Rayleigh Fading

- Rayleigh fading model:
  - In rich scattering environment, no direct path.
  - H_{ij}'s are i.i.d. Gaussian complex random variables. (Central Limit Theorem)
  - Independent real and imaginary part, each with variance $1/2$.
- Perfect CSI at the receiver, no CSI at the transmitter.
- Best strategy for the transmitter: intuitively, equal power allocation.

Capacity under Rayleigh Fading

- Maximize $I(\mathbf{x};\mathbf{y},\mathbf{H}) = \mathcal{E}_{\mathbf{H}}\{I(\mathbf{x};\mathbf{y}|\mathbf{H}=H)\}$ under power constraint.
- **Theorem 1:** The capacity of multiple antenna Gaussian channel with fading is $\mathcal{E}[\log\det[I_r + (P/t)HH^\dagger]]$. The capacity is achieved when $\mathbf{x}$ is circularly symmetric complex Gaussian with zero-mean and covariance $(P/t)I_t$.

  Sketch of the proof:

  1. $\max \mathcal{E}_{\mathbf{H}}[I(\mathbf{x};\mathbf{y}|\mathbf{H}=H)] \Rightarrow \max \mathcal{E}[\log\det(I_r + \mathbf{H}Q\mathbf{H}^\dagger)]$
  2. Constraint the choice of Q as diagonal matrix.
  3. Use random permutation to construct upper bounding matrix $\tilde{Q} = \frac{1}{t!}\sum_{\Pi} \Pi Q \Pi^\dagger$.

Capacity Evaluation

- **Theorem 2:** The capacity of the channel with t transmitters and r receivers under power constraint P equals

$$\int_0^{\infty} \log(1 + P\lambda/t) \sum_{k=0}^{m-1} \frac{k!}{(k+n-m)!} [L_k^{n-m}(\lambda)]^2 \lambda^{n-m} e^{-\lambda} d\lambda$$

where $m = \min(r, t)$ and $n = \min(r, t)$, and L_j^i are the associated Laguerre polynomials.

Capacity Evaluation

Sketch of the proof:

1. $\mathcal{E}[\log\det[I_r + (P/t)HH^\dagger]] = \mathcal{E}[\log\det[I_t + (P/t)H^\dagger H]]$

2. Define

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger & r > t \\ \mathbf{H}^\dagger\mathbf{H} & r \leq t \end{cases}$$

   The distribution of eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$ of $\mathbf{W}$ is called *Wishart Distribution*.

3. $$\mathcal{E}\left[\sum_{i=1}^{m} \log(1 + (P/t)\lambda_i)\right]$$

   Then find expectation over $p_\lambda(\lambda_i)$.

4. $p_\lambda(\lambda_1, \lambda_2, \cdots, \lambda_m) = K(m,n)\prod_i \lambda_i^{n-m} e^{-\lambda_i} \prod_{i<j}(\lambda_i - \lambda_j)^2 \Rightarrow p_\lambda(\lambda_i)$

5. Recall determinant of Vandermonde matrix

$$D = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_m \\ \vdots & & \vdots \\ \lambda_1^{m-1} & \cdots & \lambda_m^{m-1} \end{bmatrix},$$

then $p_\lambda(\lambda_1, \lambda_2, \cdots, \lambda_m) = K(m,n) \det^2 D \prod_i \lambda_i^{n-m} e^{-\lambda_i}$.

6. In the space of real functions in which defines a specific inner product, applying Gram-Schmidt orthogonal transform to the sequence $1, \lambda, \cdots, \lambda^{m-1}$, the D changes into

$$\tilde{D} = \begin{bmatrix} \varphi_1(\lambda_1) & \cdots & \varphi_1(\lambda_m) \\ \vdots & & \vdots \\ \varphi_m(\lambda_1) & \cdots & \varphi_m(\lambda_m) \end{bmatrix}.$$

7. Expand $\det \tilde{D}$, we get
   $p_\lambda = C_{mn} \sum_{\alpha,\beta} (-1)^{per(\alpha)+per(\beta)} \prod_i \varphi_{\alpha_i}(\lambda_i)\varphi_{\beta_i}(\lambda_i)\lambda_i^{n-m} e^{-\lambda_i}$

8. Integrating over $\lambda_2, \cdots, \lambda_m$ using the property of this specific inner product $\int_0^\infty \varphi_i(\lambda)\varphi_j(\lambda)\lambda^{n-m} e^{-\lambda} d\lambda = \delta_{ij}$,

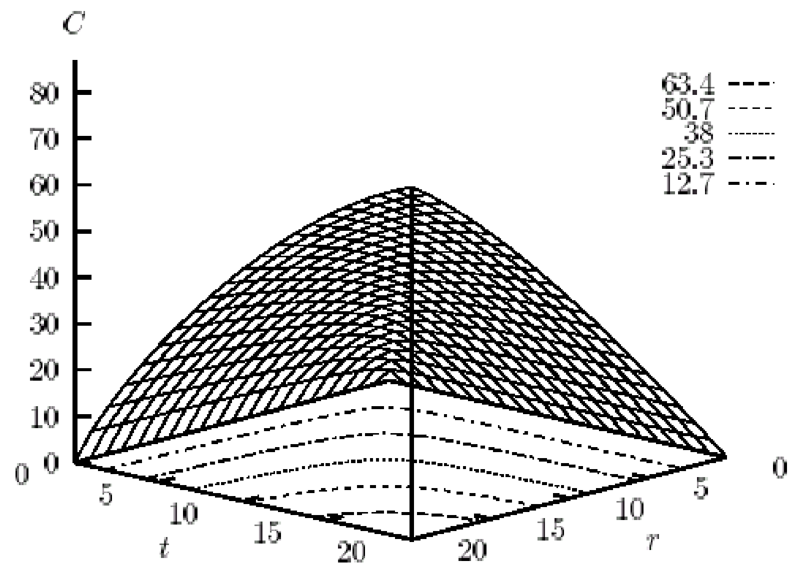
$$p_{\lambda_1}(\lambda_1) = \frac{1}{m} \sum_{i=1}^{m} \varphi_i(\lambda_1)^2 \lambda_1^{n-m} e^{-\lambda_1}.$$

9. From the Gram-Schmidt procedure,

$$\varphi_{k+1}(\lambda) = \left[\frac{k!}{(k+n-m)!}\right]^{1/2} [L_k^{n-m}(\lambda)] \tag{10}$$

USAGE OF CAPACITY EVALUATION

- Calculate the capacity with parameters t, r and SNR



EXAMPLES FOR FIXED CHANNELS

- $H_{ij} = 1$ for all i,j: send the same signal from with equal power all antennas.
  $C = log(1 + rtP)$
- $r = t = n, H = I_n$: send independent signal with equal power from different antennas.
  $C = nlog(1 + P/n)$, when $n \to \infty, C \to P$.

EXAMPLES FOR FADING CHANNELS

- $r < t = n \to \infty$, $C = \mathcal{E}[\log\det[I_r + (P/t)HH^\dagger]]$, by law of large numbers, $1/tHH^\dagger \to I_r$ w.p.1., $C = r\log(1+P)$
- Conversely,$\cdots$.
- So $n = \max(r,t) \to \infty$, $C \propto \min(r,t)$.
- Intuitively, $r = t = n \to \infty$, $C \propto n$.
- Remark: Fades bless the multi-antenna system?!

CAPACITY OF NON-ERGODIC CHANNELS

- The maximum mutual information is in general not equal to the channel capacity because it is not always achievable.
- Average is non sense in this circumstance.
- Another measure of channel capacity that is frequently used is *outage capacity* associated with a outage probability

$$Pr\{C \leq C_{\text{outage}}\} = q.$$

REFERENCE

1. R. G. Gallager, *Information theory and reliable communication*. New York: John Wiley & Sons, 1968.
2. I. Telatar, “Capacity of multi-antenna gaussian channels,” *Bell Labs Technical Memorandum*, June 1995.