EE 4TM4: Digital Communications II
Vector Gaussian Channel

I. Water-filling

Theorem 1: Let the random vector $\mathbf{X} \in \mathcal{R}^n$ have zero mean and covariance $K = \mathbb{E}\mathbf{X}\mathbf{X}^T$. Then $h(\mathbf{X}) \leq \frac{1}{2}\log(2\pi e)^n \det(K)$, with equality iff $\mathbf{X} \sim \mathcal{N}(0, K)$.

Theorem 2 (Hadamard's inequality): Let K be a non-negative definite symmetric $n \times n$ matrix. We have $\det(K) \leq \prod_{i=1}^{n} K_{ii}$, with equality iff $K_{ij} = 0$, $i \neq j$.

$\mathbf{Y} = H\mathbf{X} + \mathbf{N}$, where $\mathbf{Y}, \mathbf{N} \in \mathcal{R}^n$, $\mathbf{X} \in \mathcal{R}^m$, $H \in \mathcal{R}^{n\times m}$, and $\mathbf{N} \sim \mathcal{N}(0, K_{\mathbf{N}})$. Let $K_{\mathbf{N}} = U\Sigma U^T$ be the eigenvalue decomposition of $K_{\mathbf{N}}$, where $\Sigma = \text{diag}\{\sigma_1^2, \cdots, \sigma_n^2\}$. Let $\mathbf{N}' = \Sigma^{-1/2}U^T\mathbf{N}$, $H' = \Sigma^{-1/2}\mathbf{U}^T H$, $\mathbf{Y}' = \Sigma^{-1/2}U^T\mathbf{Y}$, and $\mathbf{X}' = \mathbf{X}$. We define an equivalent channel model $\mathbf{Y}' = H'\mathbf{X}' + \mathbf{N}'$, where $\mathbf{N}' \sim \mathcal{N}(0, I_n)$. Let $H' = V\Lambda W^T$ be the singular value decomposition of H', where $\Lambda = \text{diag}\{\lambda_1, \cdots, \lambda_r, 0, \cdots, 0\} \in \mathcal{R}^{n\times m}$ ($r = \text{rank}H'$), and $V \in \mathcal{R}^{n\times n}$, $W \in \mathcal{R}^{m\times m}$ are unitary matrices. Let $\mathbf{X}'' = W^T\mathbf{X}'$, $\mathbf{Y}'' = V^T\mathbf{Y}'$, and $\mathbf{N}'' = V^T\mathbf{N}'$. Now we get another equivalent channel model $\mathbf{Y}'' = \Lambda\mathbf{X}'' + \mathbf{N}''$, where $\mathbf{N}'' \sim \mathcal{N}(0, I_n)$. We have

$$\max_{\text{trace}(\mathbb{E}\mathbf{X}'\mathbf{X}'^T)\leq P} I(\mathbf{X}';\mathbf{Y}') = \max_{\text{trace}(\mathbb{E}\mathbf{X}''\mathbf{X}''^T)\leq P} I(\mathbf{X}'';\mathbf{Y}'').$$

Let the components of $\mathbf{X}''$ be independent[1] and jointly Gaussian with $\mathbb{E}\mathbf{X}''\mathbf{X}''^T = \text{diag}\{\mu_1, \cdots, \mu_r, 0, \cdots, 0\}$. We have

$$\begin{aligned}
\max_{\text{trace}(\mathbb{E}\mathbf{X}''\mathbf{X}''^T)\leq P} I(\mathbf{X}'';\mathbf{Y}'') &= \max_{\sum_{i=1}^r \mu_i \leq P} \frac{1}{2}\sum_{i=1}^{r} \log\left(1 + \mu_i\lambda_i^2\right) \\
&= \sum_{i=1}^{r} \max\left(\frac{1}{2}\log\left(\nu\lambda_i^2\right), 0\right)
\end{aligned}$$

where the maximizer μ_i^* is

$$\mu_i^* = \max\left(\nu - \frac{1}{\lambda_i^2}\right)$$

and ν is given by the equation

$$\sum_{i=1}^{r} \max\left(\nu - \frac{1}{\lambda_i^2}, 0\right) = P.$$

[1]For the independent parallel channels, the input should also be independent over the parallel channels in order to maximize the sum channel capacity.

II. MIMO: Ergodic Capacity

$\mathbf{Y}(t) = \mathbf{H}(t)X(t) + \mathbf{N}(t)$, where $\{\mathbf{H}(t) \in \mathcal{R}^{n\times m}\}_t$ is stationary and ergodic, and $\{\mathbf{N}(t)\}_t$ is i.i.d. with $\mathbf{N}(t) \sim \mathcal{N}(0, \sigma_N^2 I_n)$ for all t.

1) $\{\mathbf{H}(t)\}_t$ is known at the receiver: Let $K_{\mathbf{X}} = \mathbb{E}\mathbf{X}\mathbf{X}^T$. We have

$$\begin{aligned} C &= \max_{\text{trace}(K_{\mathbf{X}})\le P} I(\mathbf{X};\mathbf{Y}|\mathbf{H}) \\ &= \max_{\text{trace}(K_{\mathbf{X}})\le P} \mathbb{E}_{\mathbf{H}} I(\mathbf{X};\mathbf{Y}|\mathbf{H}=H) \\ &= \max_{\text{trace}(K_{\mathbf{X}})\le P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}K_{\mathbf{X}}\mathbf{H}^T\right). \end{aligned}$$

Raleigh fading: the entries of $\mathbf{H}$ are i.i.d. $\mathcal{N}(0,1)$.

Let $K_{\mathbf{X}} = UDU^T$ be the eigenvalue decomposition of $K_{\mathbf{X}}$, where U is a unitary matrix and D is a diagonal with nonnegative entries. Note: $\text{trace}(K_{\mathbf{X}}) = \text{trace}(D)$.

Lemma 1: The distribution of $\mathbf{H}U$ is the same as that of $\mathbf{H}$.

Proof: Let $\mathbf{g}_{i,j}$ be the (i,j) entry of $\mathbf{H}U$. Let $\mathbf{H}_i^T$ be the ith row of $\mathbf{H}$, and U_j be the jth column of U. We have

$$\begin{aligned} \mathbb{E}\mathbf{g}_{i,j}\mathbf{g}_{k,l} &= \mathbb{E}\left(\mathbf{H}_i^T U_j \mathbf{H}_k^T U_l\right) \\ &= \mathbb{E}\left(\mathbf{H}_i^T U_j U_l^T \mathbf{H}_k\right) \\ &= \mathbb{E}\left(U_l^T \mathbf{H}_k \mathbf{H}_i^T U_j\right) \\ &= U_l^T \mathbb{E}\left(\mathbf{H}_k\mathbf{H}_i^T\right) U_j. \end{aligned}$$

If $i \neq k$, we have $\mathbb{E}\left(\mathbf{H}_k\mathbf{H}_i^T\right) = 0$ and thus $\mathbb{E}\mathbf{g}_{i,j}\mathbf{g}_{k,l} = 0$. For $i = k$, we have $\mathbb{E}\left(\mathbf{H}_k\mathbf{H}_i^T\right) = I_m$ and thus $\mathbb{E}\mathbf{g}_{i,j}\mathbf{g}_{k,l} = 0$ if $j \neq l$ and 1 otherwise. ∎

By the above lemma, we have

$$\max_{\text{trace}(K_{\mathbf{X}})\le P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}K_{\mathbf{X}}\mathbf{H}^T\right) = \max_{\text{trace}(D)\le P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}D\mathbf{H}^T\right).$$

By symmetry, if D_π is a diagonal matrix obtained via permutating the diagonal entries of D, then

$$\frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}D_\pi\mathbf{H}^T\right) = \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}D\mathbf{H}^T\right).$$

Hence by the concavity of $\log\det$ and the power constraint, the optimal D must be $\frac{P}{m}I_m$. Therefore, we have

$$C = \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{P}{m\sigma_N^2}\mathbf{H}\mathbf{H}^T\right).$$

Let $\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{W}^T$ be the singular value decomposition of $\mathbf{H}$. Let $\lambda_1, \lambda_2, \cdots, \lambda_{\min(m,n)}$ be the singular values of $\mathbf{H}$. Note: $\lambda_1^2, \lambda_2^2, \cdots, \lambda_{\min(m,n)}^2$ are the non-zero eigenvalues of $\mathbf{H}\mathbf{H}^T$ and $\mathbf{H}^T\mathbf{H}$. We have

$$\begin{aligned} C &= \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{P}{m\sigma_N^2}\mathbf{H}\mathbf{H}^T\right) \\ &= \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{P}{m\sigma_N^2}\mathbf{V}\mathbf{\Lambda}\mathbf{\Lambda}^T\mathbf{V}^T\right) \\ &= \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{P}{m\sigma_N^2}\mathbf{\Lambda}\mathbf{\Lambda}^T\right) \\ &= \frac{1}{2}\mathbb{E}_{\mathbf{H}} \sum_{i=1}^{\min(m,n)} \log\left(1 + \frac{P\lambda_i^2}{m\sigma_N^2}\right). \end{aligned}$$

Note: $\text{trace}AA^T = \text{trace}A^TA = \sum_{i,j} a_{i,j}^2$. Law of energy (/power) conservation:

$$\sum_{i,j} \mathbf{h}_{i,j}^2 = \text{trace}(\mathbf{H}\mathbf{H}^T) = \text{trace}(\mathbf{V}\mathbf{\Lambda}\mathbf{\Lambda}^T\mathbf{V}^T) = \text{trace}(\mathbf{\Lambda}^T\mathbf{\Lambda}) = \sum_{i=1}^{\min(m,n)} \lambda_i^2.$$

Define $\mathsf{SNR} = \frac{P}{\sigma_N^2}$. In the low SNR regime, we have

$$\begin{aligned} C &\approx \frac{1}{2m}\mathsf{SNR}\log_2 e \sum_{i=1}^{\min(m,n)} \mathbb{E}_{\mathbf{H}}\lambda_i^2 \\ &= \frac{1}{2m}\mathsf{SNR}\log_2 e \sum_{i=1}^{\min(m,n)} \mathbb{E}_{\mathbf{H}}\text{trace}(\mathbf{H}\mathbf{H}^T) \\ &= \frac{1}{2m}\mathsf{SNR}\log_2 e \sum_{i,j} \mathbb{E}\mathbf{h}_{i,j}^2 \\ &= \frac{n}{2}\mathsf{SNR}\log_2 e \end{aligned}$$

Thus at low SNR, an m by n system yields a power gain of n over a single antenna system. This is due to the fact that the multiple receive antennas can coherently combine their received signals to get a power boost (Note: at low SNR, the capacity scales linearly with the power). Note that increasing the number of transmit antennas does not increase the power gain since, unlike the case when the channel is known at the transmitter, transmit beamforming cannot be done to constructively add signals from the different antennas. Thus, at low SNR and without channel knowledge at the transmitter, multiple transmit antennas are not very useful: the performance of an m by n channel is comparable with that of a 1 by n channel.

At high SNR, we have

$$\begin{aligned} C &\approx \frac{1}{2}\mathbb{E}_{\mathbf{H}} \sum_{i=1}^{\min(m,n)} \log\left(\frac{1}{m}\mathsf{SNR}\lambda_i^2\right) \\ &= \frac{\min(m,n)}{2}\log\mathsf{SNR} + \frac{1}{2}\mathbb{E}_{\mathbf{H}} \sum_{i=1}^{\min(m,n)} \log\frac{\lambda_i^2}{m}. \end{aligned}$$

Hence the capacity gain is $\min(m,n)$ over a single antenna system. $\min(m,n)$ is referred to as the degree of freedom.

2) $\{\mathbf{H}(t)\}_t$ is known at both the transmitter and receiver:

$$\begin{aligned} C &= \max_{\text{trace}(\mathbb{E}K_{\mathbf{X}}(\mathbf{H}))\leq P} I(\mathbf{X};\mathbf{Y}|\mathbf{H}) \\ &= \max_{\text{trace}(\mathbb{E}K_{\mathbf{X}}(\mathbf{H}))\leq P} \mathbb{E}_{\mathbf{H}} I(\mathbf{X};\mathbf{Y}|\mathbf{H}=H) \\ &= \max_{\text{trace}(\mathbb{E}K_{\mathbf{X}}(\mathbf{H}))\leq P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}K_{\mathbf{X}}(\mathbf{H})\mathbf{H}^T\right) \\ &= \max_{\text{trace}(\mathbb{E}K_{\mathbf{X}}(\mathbf{H}))\leq P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}K_{\mathbf{X}}(\mathbf{H})\mathbf{H}^T\right). \end{aligned}$$

Let $\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{W}^T$ be the singular decomposition of $\mathbf{H}$. We have

$$\begin{aligned} & \max_{\text{trace}(\mathbb{E}K_{\mathbf{X}}(\mathbf{H}))\leq P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{H}K_{\mathbf{X}}(\mathbf{H})\mathbf{H}^T\right) \\ =& \max_{\text{trace}(\mathbb{E}K_{\mathbf{X}}(\mathbf{\Lambda}))\leq P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \log\det\left(I_n + \frac{1}{\sigma_N^2}\mathbf{\Lambda}K_{\mathbf{X}}(\mathbf{\Lambda})\mathbf{\Lambda}^T\right) \\ =& \max_{\sum_{i=1}^{\min(m,n)} \mathbb{E}P(\lambda_i)\leq P} \frac{1}{2}\mathbb{E}_{\mathbf{H}} \sum_{i=1}^{\min(m,n)} \log\left(1 + \frac{P(\lambda_i)\lambda_i^2}{\sigma_N^2}\right). \end{aligned}$$

The optimal power control policy $P(\lambda_i)$ is given by

$$P^*(\lambda_i) = \max\left(\mu - \frac{\sigma_N^2}{\lambda_i^2}, 0\right),$$

where μ is determined by the power constraint

$$\sum_{i=1}^{\min(m,n)} \mathbb{E}P^*(\lambda_i) = P.$$

Note that this is waterfilling over time and space (the eigenmodes).

REFERENCES

[1] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: Wiley, 1991.

[2] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, to be published by Cambridge University Press. Online: http://degas.eecs.berkeley.edu/dtse.

[3] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.