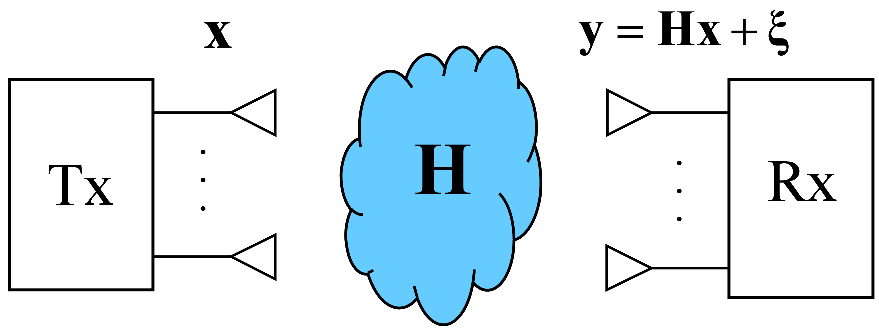
MIMO Systems and Channel Capacity

Consider a MIMO system with m Tx and n Rx antennas.



The power constraint: the total Tx power is $\left\langle |\mathbf{x}|^2 \right\rangle = P_t$.

Component-wise representation of the system model,

$$y_i = \sum_{j=1}^{m} g_{ij} x_j + \xi_i \rightarrow \mathbf{y} = \mathbf{G}\mathbf{x} + \boldsymbol{\xi} \quad (13.1)$$

<u>Theorem</u> (Foschini, Telatar): Under isotropic signaling $\mathbf{R}_x = \left\langle \mathbf{x}\mathbf{x}^+ \right\rangle = (P_t / n)\mathbf{I}$ (e.g. no CSI at the Tx), the capacity of the channel in (13.1) is

$$C = \log\det\left(\mathbf{I}_m + \frac{\gamma}{m}\mathbf{H}\mathbf{H}^+\right) \text{ [bit/s/Hz]} \tag{13.2}$$

where

$\mathbf{H} = \sqrt{P_t / P_r}\,\mathbf{G}$ is the normalized channel matrix,

$\gamma = P_r / \sigma_0^2$ is the aggregate SNR at the Rx end.

$P_r = P_t \cdot tr\left(\mathbf{G}\mathbf{G}^+\right) / m$ is the total received power ,

so that reliable transmission is possible at any rate $R < C$.

Proof: When $\mathbf{x}$ is a complex Gaussian vector, $CN(\mathbf{0}, \mathbf{K}_x)$, its entropy is

$$H(\mathbf{x}) = \log_2\left[(\pi e)^m \det\mathbf{K}_x\right] \tag{13.3}$$

where $\mathbf{K}_x$ is the correlation matrix,

$$\mathbf{K}_x = \langle \mathbf{x}\mathbf{x}^+ \rangle \rightarrow [\mathbf{K}_x]_{ij} = \langle x_i x_j^{\ +} \rangle \tag{13.4}$$

Similarly for the noise and the output vector:

$$H(\xi) = \log_2\left[(\pi e)^n \det \mathbf{K}_\xi\right], \; H(\mathbf{y}) = \log_2\left[(\pi e)^n \det \mathbf{K}_y\right] \tag{13.5}$$

where $\mathbf{K}_\xi = \langle \xi\xi^+ \rangle$, $\mathbf{K}_y = \langle \mathbf{y}\mathbf{y}^+ \rangle$.

Note: this follows from the pdf of a complex Gaussian vector,

$$\rho(\mathbf{x}) = \frac{1}{\pi^m |\mathbf{K}_x|} \exp\left[-\mathbf{x}^+ \mathbf{K}_x^{-1} \mathbf{x}\right] \tag{13.6}$$

and we have assumed that

$$\langle \mathbf{x}_R \mathbf{x}_R^+ \rangle = \langle \mathbf{x}_I \mathbf{x}_I^+ \rangle, \; \langle \mathbf{x}_R \mathbf{x}_I^+ \rangle = -\langle \mathbf{x}_I \mathbf{x}_R^+ \rangle, \; \mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I \tag{13.7}$$

where $\mathbf{x}_R$ and $\mathbf{x}_I$ are real and imaginary parts of $\mathbf{x}$.

The mutual information is

$$I(\mathbf{y},\mathbf{x}) = H(\mathbf{x}) + H(\mathbf{y}) - H(\mathbf{x},\mathbf{y}) = \log \frac{|\mathbf{K}_x| \cdot |\mathbf{K}_y|}{|\mathbf{K}_{xy}|} \tag{13.8}$$

where

$$\mathbf{K}_{xy} = \left\langle \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \begin{bmatrix} \mathbf{x}^+ & \mathbf{y}^+ \end{bmatrix} \right\rangle = \begin{bmatrix} \mathbf{K}_x & \mathbf{K}'_{xy} \\ \mathbf{K}'_{yx} & \mathbf{K}_y \end{bmatrix} \tag{13.9}$$

The capacity is

$$C = \max_{p(\mathbf{x})} \left[I(\mathbf{y},\mathbf{x}) \right] \tag{13.10}$$

under the assumptions that the channel is fixed (static). We further assume that the CSI is available at the Rx only (but not at the Tx).

For Gaussian noise, the maximum is achieved when $\mathbf{x}$ is zero-mean Gaussian.

Under certain circumstances (e.g. no Tx CSI), the covariance is a scaled identity, subject to the power constraint $tr\mathbf{K}_x \le P_t$. Hence,

$$\mathbf{K}_x = \frac{P_t}{m} \mathbf{I}_m \tag{13.11}$$

Here, we also assume that the noise is uncorrelated from one Rx antenna to another,

$$\mathbf{K}_{\xi} = \sigma_0^2 \mathbf{I}_n, \tag{13.12}$$

and that the signal and noise are uncorrelated as well:

$$\left\langle x_i \xi_j^* \right\rangle = 0 \tag{13.13}$$

Under these assumptions,

$$\mathbf{K}_y = \langle \mathbf{y}\mathbf{y}^+ \rangle = \mathbf{G}\mathbf{G}^+ \frac{P_t}{n} + \sigma_0^2 \mathbf{I}_n$$
$$\mathbf{K}'_{xy} = \frac{P_t}{m}\mathbf{G}^+, \mathbf{K}'_{yx} = \frac{P_t}{m}\mathbf{G} \tag{13.14}$$

Using the following identity,

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A}| \left| \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \right| \tag{13.15}$$

we find

$$\left|\mathbf{K}_{xy}\right| = \left|\mathbf{K}_x\right| \left|\mathbf{K}_y - \mathbf{K}'_{xy}\mathbf{K}_x^{-1}\mathbf{K}'_{yx}\right| = \sigma_0^{2n} \left|\mathbf{K}_x\right| \tag{13.16}$$

Finally,

$$C = \log \frac{\left|\mathbf{K}_x\right| \cdot \left|\mathbf{K}_y\right|}{\left|\mathbf{K}_{xy}\right|} = \log \det \left[\mathbf{I}_m + \frac{P_t}{\sigma_0^2 m} \mathbf{G}\mathbf{G}^+ \right] \tag{13.17}$$

Let us normalize **G** as follows,

$$\mathbf{H} = \sqrt{P_t / P_r}\,\mathbf{G} \tag{13.18}$$

where P_r is the total received power. Then,

$$C = \log\det\left(\mathbf{I}_n + \frac{\gamma}{m}\mathbf{H}\mathbf{H}^+\right)\ [\text{bit/s/Hz}] \tag{13.19}$$

where $\gamma = P_r / \sigma_0^2$ is the average SNR at the Rx. This gives the MIMO channel capacity in bit/s/Hz. This is the celebrated Foschini-Telatar formula. Q.E.D.

Important special cases

n parallel independent sub-channels,

$$\mathbf{H} = \mathbf{I} \ (n = m) \tag{13.20}$$

The capacity is

$$C = m \log\left(1 + \frac{\gamma}{m}\right) \tag{13.21}$$

Interpretation: the incoming bit stream is split into m independent sub-streams and transmitted independently over m channels. Since the total Tx power is fixed to P_t, per-channel SNR is γ/m.

Asymptotically,

$$m \to \infty \Rightarrow C \approx \frac{\gamma}{\ln 2} \tag{13.22}$$

Note that capacity growth linearly with $\gamma \to$ much faster!

While $C(m)$ is monotonically increasing with m, the increase slows down for large n.

Since increase in n is related to huge increase in system complexity, there is some maximum m, which is approximately $m_{\max} \approx \gamma$.

In practice, one would keep $m \le m_{\max}$.

Compare with the case when all the bits are transmitted over one sub-channel only,

$$C_1 = \log(1+\gamma) \tag{13.23}$$

Clearly,

$$C \ge C_1 \text{ for } \gamma \gg 1 \tag{13.24}$$

Another approach to this problem is based on the cost model,

$$S = \alpha m \tag{13.25}$$

where S = cost, α = cost/stream. Hence, the capacity/cost ratio (capacity per unit cost) is

$$\frac{C}{S} = \frac{1}{\alpha} \log\left(1 + \frac{\gamma}{m}\right) \tag{13.26}$$

This is capacity per unit cost and it is always decreasing function of n. However, the decrease is slow for moderate values of m, and it becomes large only when $m \sim \gamma$.

Next consider diversity combining,

$$\mathbf{H} = [h_1 \quad h_2 \quad \dots \quad h_n]^T,\ m = 1 \tag{13.27}$$

(one Tx and n Rx's). The capacity is

$$C = \log\det\left(\mathbf{I} + \frac{\gamma}{m}\mathbf{H}\mathbf{H}^+\right) = \log\left(1 + \gamma\sum_{i=1}^{n}|h_i|^2\right) \tag{13.28}$$

Clearly, this is the MRC. It proves once again that the MRC is optimum (this time, from capacity (information theoretic) viewpoint).

Selection combining can be represented as,

$$C = \max_i\left\{\log\left(1 + \gamma|h_i|^2\right)\right\} = \log\left(1 + \gamma\max_i\left\{|h_i|^2\right\}\right) \tag{13.29}$$

Hence, "capacity-wise" selection combining is the same as "power-wise" one.

Tx diversity combining, for fixed P_t , is not the same as Rx combining.

Q: prove it!

However, when per-branch Tx power is fixed, it is the same as Rx combining.

Q: prove it!

SVD Decomposition and Channel Capacity

Introduce a (instantaneous) channel covariance matrix:

$$\mathbf{W} = \mathbf{H}^{+}\mathbf{H} \tag{13.30}$$

Eigenvalue decomposition of $\mathbf{W}$ is

$$\mathbf{W} = \mathbf{Q}\Lambda\mathbf{Q}^{+} \tag{13.31}$$

where $\mathbf{Q}$ is a mxm unitary matrix of eigenvectors, and

$$\Lambda = diag\begin{bmatrix}\lambda_1 & \lambda_2 & \dots & \lambda_m\end{bmatrix}$$

diagonal matrix of eigenvalues.

Consider the SVD of $\mathbf{H}$:

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{+} \tag{13.32}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices (of left and right singular vectors of $\mathbf{H}$), and $\boldsymbol{\Sigma}$ is diagonal (mxn) matrix of (non-negative) singular values.

Note that

$$\mathbf{W} = \mathbf{H}^{+}\mathbf{H} = \mathbf{V}\boldsymbol{\Sigma}^{+}\boldsymbol{\Sigma}\mathbf{V}^{+} \tag{13.33}$$

Hence,

$$\mathbf{Q} = \mathbf{V},\ \mathbf{\Lambda} = \mathbf{\Sigma}^{+}\mathbf{\Sigma} \tag{13.33}$$

i.e, eigenvalues of $\mathbf{W}$ are squared singular values of $\mathbf{H}$, $\lambda_i(\mathbf{W}) = \sigma_i^2(\mathbf{H})$.

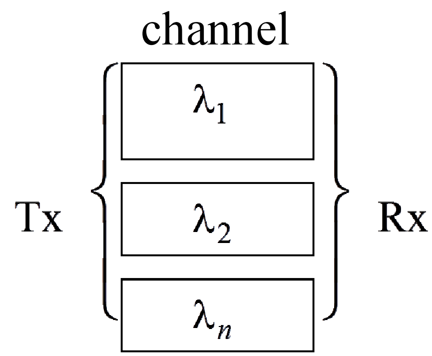
Using the equations above, the capacity can be presented as

$$\boxed{C = \sum_{i=1}^{m} \log\left(1 + \frac{\gamma}{m}\lambda_i\right)} \tag{13.34}$$

Hence, $\lambda_i = \lambda_i(\mathbf{W}) = \sigma_i^2(\mathbf{H})$ are the channel eigenmode power gains.

Q: prove it!

This is alternative representation of the Foschini-Telatar formula.



Note: capacity doesn't change under transformation $\mathbf{H} \to \mathbf{H}^+$

Q: prove it!

Example 1

Consider the all-1 channel $H_{ij} = 1$, find its capacity.

Solution:

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^+$$

where

$$\mathbf{U} = \frac{1}{\sqrt{n}}\begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix}^T, \ \mathbf{\Sigma} = \sqrt{mn}, \ \mathbf{V}^+ = \frac{1}{\sqrt{m}}\begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix} \qquad (13.35)$$

Hence,

$$\lambda_1 = mn, \ \lambda_2, \dots, \lambda_m = 0 \qquad (13.36)$$

$$C = \log(1 + \gamma n) \qquad (13.37)$$

i.e, it increases only logarithmically with n –> this is an example of a correlated (rank-deficient) channel.

Note : the effect of m is hidden in γ since $\gamma = \frac{P_r}{\sigma_0^2}$, where P_r is the total Rx power, $P_r = mP_{r1}$.

Example 2

Consider a multipath channel of the form

$$\mathbf{H} = \frac{1}{\sqrt{mn}} \sum_{i=1}^{M} \mathbf{w}_i \mathbf{v}_i^+ g_i \tag{13.38}$$

where $\mathbf{v}_i$ = Tx array manifold vector, $\mathbf{w}_i$ = Rx array manifold vector , and g_i - channel gains, all for i–th multipath component. Assume that

$$\mathbf{w}_i^+ \mathbf{w}_j = n\delta_{ij}, \mathbf{v}_i^+ \mathbf{v}_j = m\delta_{ij} \tag{13.39}$$

(orthonormal). By inspection, we conclude that

$$\lambda_i = |g_i|^2, \quad C = \sum_{i=1}^{M} \log\left(1 + \frac{\gamma}{m}\lambda_i\right), \text{ if } M < \min(m, n) \tag{13.40}$$

Hence the number of multipath components limits MIMO capacity!

In general case, the number of non-zero eigenvalues is

$$M_0 = \min(m,n,M) \rightarrow C = \sum_{i=1}^{\min(m,n,M)} \log\left(1+\frac{\gamma}{m}\lambda_i\right) \tag{13.41}$$

Assuming that all the eigenvalues are equal,

$$C = \min(m,n,M)\log\left(1+\frac{\gamma}{m}\lambda\right) \tag{13.42}$$

<u>Number of degrees of freedom</u>: a factor in front of log is called a number of degrees of freedom, or capacity slope.

Q.: What is the meaning of it?

Example 3

MISO channel (Tx diverisity), $n=1$, $\mathbf{H}=\begin{bmatrix}1 & 1 & \ldots & 1\end{bmatrix}$:

The capacity is

$$C=\log\left(1+\frac{\gamma}{m}\cdot m\right)=\log(1+\gamma) \tag{13.43}$$

i. e., the same , as SISO with the same total Tx power (the effect of coherent combining is already included in ρ, which is the ratio of total Rx power, in one Rx from all Tx's, to the noise power).

Compare it with SIMO (Rx diversity), $m=1$, $\mathbf{H}=\begin{bmatrix}1 & 1 & 1\end{bmatrix}^T$,

$$C=\log(1+\gamma\cdot n)\neq\log(1+\gamma) \text{ if } n\neq 1 \tag{13.44}$$

Q.: explain the difference!

Capacity of an Ergodic Rayleigh-fading Channel

We assume that Rx knows the channel (H), but Tx doesn't.

Start with the general MIMO capacity expression for a given (fixed) **H** :

$$C = \log\det\left(\mathbf{I} + \frac{\gamma}{m}\mathbf{H}\mathbf{H}^{+}\right) \tag{13.45}$$

Since **H** is random for a fading channel, C is random as well. For any given realization of **H**, C can be evaluated, but it varies from realization to realization, i.e. random variable.

Define the mean (ergodic) capacity,

$$\overline{C} = \left\langle C \right\rangle_{\mathbf{H}} = \left\langle \log\det\left(\mathbf{I} + \frac{\gamma}{m}\mathbf{H}\mathbf{H}^{+}\right)\right\rangle_{\mathbf{H}} \tag{13.46}$$

Telatar gives a detailed formulation, based on the mutual information $I\left(x,\left(y,H\right)\right)$, and proves that the maximum is achieved when **x** are i.i.d. Gaussian.

Note that the mean capacity makes practical sense for ergodic channels only, i.e., when the expectation over realizations is the same as expectations over time.

<u>Interpretation of ergodic capacity</u>: it can be proved (based on information theory) that there exists a single code that achieves the capacity in (13.46). Alternatively, an adaptive system can be built that achieves the instantaneous capacity in (13.45) and its average capacity is as in (13.46).

For a Rayleigh channel, h_{ij} are i.i.d. complex Gaussians with unit variance, $h_{ij} \sim C\mathbb{N}\left(0,1\right)$. Telatar describes in details the evaluation of capacity in this case (analytical approach).

<u>Consider some special cases.</u>

Example 1

n is fixed, and $m \to \infty$. In this case,

$$\frac{1}{m}\mathbf{HH}^{+} = \frac{1}{m}\sum_{j=1}^{m} h_{ij}h_{kj}^{*} \to \mathbf{I}_n \tag{13.47}$$

in probability due to the central limit theorem. Hence , the capacity is

$$\bar{C} = n\log\left(1+\gamma\right) \tag{13.48}$$

i.e., the same as the capacity of n parallel independent channels, AWGN, no fading. This is the effect of Tx diversity.

Consider $n=1,\ m \to \infty$ then,

$$\bar{C} = \log\left(1+\gamma\right) \tag{13.49}$$

i.e, an infinite-order diversity transforms the Rayleigh-fading channel into a fixed (non-fading) AWGN channel.

Example 2: m is fixed and $n \to \infty$. Homework. References: [1-3, 5, 6].

Summary

- MIMO capacity. Basic concepts (entropy, mutual information) for random vectors.
- Canonical form of the MIMO capacity. Large n limit.
- Comparison to conventional systems. Capacity of diversity combining systems. The impact of multipath, Tx and Rx antenna number.
- MIMO capacity of a Rayleigh channel. The impact of fading.

References

1. J.R. Barry, E.A. Lee, D.G. Messerschmitt, Digital Communication, 2003 (Third Edition). Ch. 10, 11.
2. D. Tse, P. Viswanath, Fundamentals of Wireless Communications, Cambridge, 2005.
3. P.P. Vaidyanathan et al, Signal Processing and Optimization for Transceiver Systems, Cambridge University Press, 2010; Ch. 6, 22, App. B, C.
4. D.W. Bliss, S. Govindasamy, Adaptive Wireless Communications: MIMO Channels and Networks, Cambridge University Press, 2013.
5. A. Paulraj, R. Nabar, D. Gore, Introduction to Space-Time Wireless Communications, Cambridge University Press; 2003.
6. E.G. Larson, P. Stoica, Space-Time Block Coding for Wireless Communications, Cambridge University Press, 2003.
7. I.E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," AT&T Bell Lab. Internal Tech. Memo., June 1995 (European Trans. Telecom., v.10, N.6, Dec.1999).
8. Foschini, G.J., Gans M.J.: ‘On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas’, Wireless Personal Communications, vol. 6, No. 3, pp. 311-335, March 1998.
9. Rayleigh, G.G., Gioffi, J.M.: "Spatio-Temporal Coding for Wireless Communications," IEEE Trans. Commun., v.46, N.3, pp. 357-366, 1998.

Appendix 1: SVD and optimization of MIMO Channel Capacity

When the T_x signal correlation matrix $\mathbf{P} = \mathbf{E}\left\{\mathbf{x}\mathbf{x}^+\right\} \neq \mathbf{I}$, the MIMO capacity is

$$\mathbf{C} = \log\left|1 + \frac{1}{\sigma_0^2}\mathbf{HPH}^+\right| \tag{1}$$

If CSI (channel state information) is available at the T_x, **P** can be chosen to maximize C, subject to the total T_x power constrain:

$$\sum_{i=1}^{m} P_{ii} = tr(\mathbf{P}) \leq P_T \tag{2}$$

where P_{ii} is the i-th T_x power.

Consider the $m \times n$ MIMO channel,

$$\mathbf{y} = \mathbf{Hx} + \boldsymbol{\xi} \tag{3}$$

Using the SVD of **H**,

$$\mathbf{H} = \mathbf{U\Sigma V}^+ \tag{4}$$

where **U**,**V** are $n \times n$ and $m \times m$ unitary matrices, $\mathbf{U}^+\mathbf{U} = \mathbf{V}^+\mathbf{V} = \mathbf{I}$, and

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \tag{5}$$

and $\Sigma_1 = diag\left[\sigma_1, \quad \sigma_2, \quad .. \quad \sigma_k\right]$ are non-zero singular values of **H**.

Using (4) and (3),

$$\mathbf{y} = \mathbf{U\Sigma V}^{+}\mathbf{x} + \boldsymbol{\xi} \quad , \quad \tilde{\mathbf{y}} = \mathbf{\Sigma}\tilde{\mathbf{x}} + \tilde{\boldsymbol{\xi}} \tag{6}$$

$$\widetilde{y_i} = \sigma_i \widetilde{x_i} + \widetilde{\xi_i} \quad , \quad i = 1, 2..k \tag{7}$$

where $\tilde{\mathbf{y}} = \mathbf{U}^{+}\mathbf{y}, \quad \tilde{\mathbf{x}} = \mathbf{V}^{+}\mathbf{x}, \quad \tilde{\boldsymbol{\xi}} = \mathbf{U}^{+}\boldsymbol{\xi}$. Note that multiplication by a unitary matrix does not change statistics of a random vector and, hence, does not affect the mutual information and the capacity.

Hence, the channel in (7) has the same capacity as the original channel (3). But the channel in (7) is as of k independent sub-channels, with per-sub-channel SNR:

$$\tilde{\gamma} = \lambda_i \frac{P_i}{\sigma_0^2} \tag{8}$$

and $\lambda_i = \sigma_i^2$ are the eigenvalues of $\mathbf{HH}^{+}$, and $P_i = P_{ii}$. Its capacity is

$$\mathbf{C} = \sum_{i=1}^{k} \log_2\left(1 + \lambda_i \frac{p_i}{\sigma_0^2}\right) \tag{9}$$

Optimum P_i can be found using water-filling (WF) technique as follows:

$$P_i = \left[\mu - \frac{\sigma_0^2}{\lambda_i} \right]_+ , \quad i = 1, 2, ..k_1 \tag{10}$$

$$\sum_{i=1}^{k_1} P_i = P_T \tag{11}$$

where $\left[x \right]_+ = x$ if x>0 and 0 otherwise; k_1 is the number of active eigenmodes (i.e. with non-zero P_i), and constant μ is found from (10). Note that (10) and (11) also give (implicitly) k_1.

Water-filling technique can be formulated as iterative algorithm as follows [1]:

1) order eigenvalues, set iteration index p=0
2) find μ as follows

$$\mu = \frac{1}{k-p}\left(p_T + \sigma_0^2 \sum_{i=1}^{k-p} \frac{1}{\lambda_i} \right) \tag{12}$$

3) set P_i using (10) with $k_1 = k - p$
4) if there is zero P_i, set $p = p + 1$, eliminate λ_i and go to step 2
5) finish when all P_i (i=1, 2,..k-p) are non-zero.

This algorithm gives all non-zero P_i . All the other P_i are zeros (i.e., those eigenmodes are not used).

Proof of the water-filling technique :

using Lagrange multipliers with the following goal function,

$$F=\sum_i \log(1+\frac{\lambda_i P_i}{\sigma_0^2})-\alpha\left(\sum_i P_i - P_T\right) \quad (13)$$

$$\frac{dF}{dP_i}=0, \quad \frac{dF}{d\alpha}=0 \quad (14)$$

where α is a Lagrange multiplier. From (14), one obtains (10) and (11).

Finally, optimum **P** is found using (4)

$$\mathbf{P}=\mathbf{VDV}^+ \quad (15)$$

where $D = diag\left[p_1, p_2, ..p_{k_1}, 0...0\right].$

Effect of T_x CSI on the Capacity [2]

Compare the MIMO channel capacity in 2 cases:

1) no T_x CSI (uninformed T_x-UT)

2) full T_x CSI (informed T_x-IT)

In case 1, the capacity is given by (9) with $P_i = \frac{P_T}{m}$

$$\mathbf{C}_{UT} = \sum_{i=1}^{k} \log(1 + \frac{P_T}{m\sigma_0^2}\lambda_i) \tag{15}$$

In case 2, the capacity is given by (9) with P_i given by (10)

$$C_{IT} = \sum_{i=1}^{k_1} \log(1 + \frac{\lambda_i P_i}{\sigma_0^2}) \tag{18}$$

$$P_i = \left[\mu - \frac{\sigma_0^2}{\lambda_i}\right]_+ \tag{18a}$$

$$\sum_{i=1}^{k_1} P_i = P_T \tag{18b}$$

Consider the ratio

$$\beta = \frac{C_{IT}}{C_{UT}} \tag{19}$$

when $P_T / \sigma_0^2 \to \infty$ i.e. high SNR mode.

Assuming that P_T=const and $\sigma_0^2 \to \infty$, it is clear from (18a) and (18b) that $P_i = P_T/m$ (assuming k=m, i.e. full-rank channel), and

$$\frac{C_{IT}}{C_{UT}} \to 1 \quad as \quad \frac{P_T}{\sigma_0^2} \to \infty \tag{20}$$

Hence, optimum power allocation does not provide advantage in high SNR mode – parallel transmission (spatial multiplexing) with equal powers is optimum.

Consider the case of low SNR, $P_T/\sigma_0^2 \to 0$. Assume that $\sigma_0^2 \to \infty$, then from (18)-(18b) one finds that $P_{i_{\max}} = P_T$ and all the other $P_i = 0^*$, where $i_{\max}$ is the largest eigenmode index.

Hence,

$$C_{IT} = \log\left(1 + \frac{\lambda_{\max} P_T}{\sigma_0^2}\right) \approx \frac{\lambda_{\max} P_T}{\sigma_0^2} \log e \qquad (21)$$

Similarly,

$$C_{UT} \approx \frac{P_T}{m\sigma_0^2} \sum_{i=1}^{m} \lambda_i \log e$$

Hence,

$$\frac{C_{IT}}{C_{UT}} \approx \frac{m\lambda_{\max}}{\sum_{i=1}^{m} \lambda_i} = \frac{m\lambda_{\max}(\mathbf{HH}^+)}{tr(\mathbf{HH}^+)} \qquad (22)$$

Important conclusion: in low SNR case, the best strategy is to use the largest eigenmode only $\to$ this is beamforming!

In high SNR mode, the best strategy is to use spatial multiplexing (parallel transmission on all eigenmodes).

References:

1. A. Paulraj, R. Nabar, D. Gore, Introduction to Space-Time Wireless Communications, Cambridge University Press, 2003.

2. G. Larsson, P. Stoica, Space-Time Block Coding for Wireless Communications, Cambridge University Press, 2003.