Multiple antenna transmission

Vahid Meghdadi

References : 1-Goldsmith
2-Tse

June 2007
update Janvier 2009

Chapter 1

Diversity

1.1 Receiver diversity

1.1.1 system model

1.1.2 Selection combining

1.1.3 Threshold combining

1.1.4 Maximum ratio combining[1]

Suppose a MIMO system with M transmit and N receive antennas. Suppose also that at each channel use just one symbol is to be transmitted. The transmitter will multiply the symbol x by a weighting vector $\mathbf{w}$ of size N. So the received signal at the receiver will be a vector of size N as:

$$\mathbf{y} = \sqrt{\gamma}\mathbf{H}\mathbf{w}x + \mathbf{n}$$

where γ is the signal to noise ratio, $\mathbf{H}$ is the $N \times M$ channel vector with normalized circularly symmetric Gaussian random entries and $\mathbf{n}$ is the noise vector with the same distribution as the entries of $\mathbf{H}$. The receiver multiplies the received vector by $\tilde{\mathbf{w}}$ in order to maximize the SNR. So the received metric is:

$$y = \sqrt{\gamma}\tilde{\mathbf{w}}^H\mathbf{H}\mathbf{w}x + \tilde{\mathbf{w}}\mathbf{n}$$

The optimum receiver is that who maximizes the SNR, which is the same as maximum ratio combining. To show that, the SNR of the above equation is

$$SNR = \gamma\frac{\mathbf{w}^H\mathbf{H}^H\tilde{\mathbf{w}}\tilde{\mathbf{w}}^H\mathbf{H}\mathbf{w}}{\tilde{\mathbf{w}}^H\tilde{\mathbf{w}}}$$

To maximize SNR, one could see from this equation that the vectors $\tilde{\mathbf{w}}$ and $\mathbf{H}\mathbf{w}$ should be in the same direction and hence, on should select $\tilde{\mathbf{w}} = \mathbf{H}\mathbf{w}$. The SNR that is obtained is:

$$SNR_{MRC} = \gamma\mathbf{w}^H\mathbf{H}^H\mathbf{H}\mathbf{w}$$

1.1.5 Equal gain combining

[1]Hamid Reza Bahrami and Tho Le-Ngoc, "Maximum Ratio Combining Precoding for Multi-Antenna Relay Systems", ICC 2008

Chapter 2

Multiple Antennas and Space-Time Communication

2.1 Narrowband MIMO model

Consider a transmission system of M_t transmit and M_r receive antennas. This system can be represented by the following equation:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{M_r} \end{bmatrix} = \begin{bmatrix} h_1 1 & \dots & h1M_t \\ \vdots & \ddots & \vdots \\ h_{M-r1} & \dots & h_{M_rM_t} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{M_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_{M_t} \end{bmatrix} \tag{2.1}$$

This is $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$ and h_{ij} represents the gain from TX antenna j to RX antenna i. We assume that the transmit power is P, channel bandwidth is B and the noise is zero mean complex Gaussian with covariance matrix $\sigma^2 \mathbf{I}_{M_r}$ where $\sigma^2 = \mathrm{E}[n_i^2] = N_0/2$. We can assume that the noise variance is one and the transmit power is $P/\sigma^2 = \rho$ and ρ is the average SNR per receive antenna under unity channel gain.

Here we have a MIMO channel. The first assumption is that the channel is zero mean spatially white (ZMSW), so the entries of the channel matrix are i.i.d. zero mean unit-variance, complex circularly symmetric Gaussian random variables. Optimal decoding scheme is ML whose complexity is exponential in M_t. That is because we should test $|\mathcal{X}|^{M_t}$ signals where $|\mathcal{X}|$ is the size of the alphabet set of each TX antenna. If the receiver knows the channel, it is possible to reduce this complexity and transform the MIMO channel into several parallel independent channels. Because in the case of static channel, the receiver can easily estimate the CSI (Channel State Information) through the use of a training sequence, we assume that the channel is known at the receiver.

2.2 Parallel decomposition of the MIMO channel

For any matrix H we can obtain its singular value decomposition (SVD) as

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{2.2}$$

In SVD decomposition, $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\mathbf{\Sigma}$ is a matrix whose any non-diagonal entry is zero. Its diagonal entry is the singular value of the matrix $\mathbf{H}$, it means that $\sigma_i = \sqrt{\lambda_i}$ where λ_i is the eigenvalue of $\mathbf{H}\mathbf{H}^H$. If H is full rank, which is referred to a rich scattering environement, then $R_{\mathbf{H}} = \min(M_t, M_r)$.

We can multiply the information symbols $\tilde{\mathbf{x}}$ by the matrix $\mathbf{V}$ which is called precoding. At the receiver, the received vector is multiplied by $\mathbf{U}^H$. We can write then:

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{U}^H(\mathbf{H}\mathbf{x} + \mathbf{n}) \\ &= \mathbf{U}^H(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{x} + \mathbf{n}) \\ &= \mathbf{\Sigma}\tilde{\mathbf{x}} + \tilde{\mathbf{n}} \end{aligned} \tag{2.3}$$

Since $\mathbf{U}$ is a unitary matrix, the characteristics of $\tilde{\mathbf{n}}$ is the same as $\mathbf{n}$. The interpretation of the equation 2.3 is that there are $R_{\mathbf{H}}$ independent channels which can potentially increase the rate by a factor of $R_{\mathbf{H}}$. This is called the multiplexing gain in a MIMO system. Note that the complexity here is linear with M_r and not exponential as before. Since the gain of each channel is σ_i, the channel with small gain has a small capacity. In order to optimize the power allocated to each channel one may use the water-filling theory if the channel is known at the transmitter.

2.3 MIMO channel capacity

The channel capacity is the maximum of mutual information between the input and the output of the channel.

2.3.1 Static Channel

The capacity of a MIMO channel if the channel is known at the receiver is defined as the mutual information of the input and output of the channel.

$$C = \max_{p(\mathbf{x})} I(\mathbf{X};\mathbf{Y}) = \max_{p(\mathbf{x})}[H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X})] \tag{2.4}$$

Since given $\mathbf{X}$ and knowing the channel $\mathbf{H}$, the uncertainty over $\mathbf{Y}$ is the same as over $\mathbf{n}$, so $H(\mathbf{Y}|\mathbf{X}) = H(\mathbf{n})$. Since the entropy of noise is independent of the channel input, maximizing the mutual information is equivalent to maximizing the entropy in $\mathbf{y}$. We can calculate the covariance matrix of $\mathbf{Y}$ from $\mathbf{R}_{\mathbf{x}}$ as:

$$\mathbf{R}_{\mathbf{y}} = \mathrm{E}[\mathbf{y}\mathbf{y}^H] = \mathbf{H}\mathbf{R}_{\mathbf{x}}\mathbf{H}^H + \mathbf{I}_{M_r} \tag{2.5}$$

It is proved that for a continuous random variable with a given power, *the entropy is maximum if the continuous random variable is a zero-mean circularly symmetric complex Gaussian* (ZMCSCG) random variable. But $\mathbf{y}$ is Gaussian if $\mathbf{x}$ is Gaussian because the channel matrix $\mathbf{H}$ is assumed to be fixed and the noise is Gaussian. So the optimal distribution for $\mathbf{x}$ is ZMCSCG, subject to the power constraint $\mathrm{Tr}(\mathbf{R}_{\mathbf{x}}) = \rho$.

The entropy of a real Gaussian random variable is $H(\mathbf{X}) = \frac{1}{2}\log_2(2\pi e\sigma^2)$ and if the variable is complex Gaussian the entropy will be $H(\mathbf{X}) = log_2(\pi e\sigma^2)$. If we have a real correlated Gaussian vector $\mathbf{X}$ with covariance matrix $\mathbf{R}$, the entropy of $\mathbf{X}$ will be $H(\mathbf{X}) = \frac{1}{2}\log_2[(2\pi e)^n. \det(\mathbf{R})]$ and if we have a complex correlated Gaussian vector $H(\mathbf{X}) = \log_2[(\pi e)^n. \det(\mathbf{R})]$ bit per symbol.

Using the above considerations, $H(\mathbf{Y}) = B\log_2 \det[\pi e\mathbf{R}_y]$ and $H(\mathbf{n}) = B\log_2 \det[\pi e\mathbf{I}_{M_r}]$ bit per second and we can calculate the mutual information of $\mathbf{X}$ and $\mathbf{Y}$ as follows:

$$I(\mathbf{X};\mathbf{Y}) = B\log_2 \frac{\det[\mathbf{R_y}]}{\det[\mathbf{I}_{M_r}]} = B\log_2 \det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H] \quad (2.6)$$

So the capacity of a multi antenna system is obtained by maximizing the equation (2.6) over $\mathbf{R_x}$ subject to $\mathrm{Tr}(\mathbf{R_x}) = \rho$:

$$C = \max_{\mathbf{R_x}:\mathrm{Tr}(\mathbf{R_x})=\rho} B\log_2 \det[\mathbf{I}_{M_r} + \mathbf{H}\mathbf{R_x}\mathbf{H}^H] \quad (2.7)$$

This equation is valid whether or not the channel is known at the transmitter. If the Tx knows the channel, it may use water filling to obtain the optimize $\mathbf{R_x}$ and if not, it may use equal power strategy.

Channel known at the transmitter

If the channel is known at the transmitter the MIMO system is equivalent to $R_\mathbf{H}$ parallel channel according to equation (2.3). Each channel is just a SISO channel with capacity $C_i = B\log_2(1 + \frac{P_{x_i}}{P_n})$. The total capacity is the sum of all this channels subjected to the total power constraint. So the only parameter to optimize is the distribution of total available power among the $R_\mathbf{H}$ SISO channels. Defining $\rho_i = P_{x_i}/P_n$ and σ_i which is the gain of i^{th} channel, the capacity of the MIMO system is as follows:

$$C = \max_{\rho_i:\sum_i \rho_i \le \rho} \sum_{i=1}^{R_\mathbf{H}} B\log_2(1 + \sigma_i^2\rho_i) \quad (2.8)$$

As a recall to water-filling problem, the optimal solution to power allocation among all the Tx antennas can be obtained by the following set of equations:

$$\begin{cases} \frac{P_1}{\sigma^2} = (Cte - \frac{1}{\sigma_1^2})^+ \\ \quad\cdots \\ \frac{P_{R_\mathbf{H}}}{\sigma^2} = (Cte - \frac{1}{\sigma_{R_\mathbf{H}}^2})^+ \end{cases} \quad (2.9)$$

where $(x)^+ = \max(x, 0)$. To solve this equation, we use the constraint that the sum of all the powers is equal to the power constraint. So the *cte* can be calculated and then all the p_i will be known. If however one of them is negative, it means that the corresponding σ_i is too small and the channel should not be used. So we eliminate the corresponding equation and resolve the new set. Note that in this configuration, we obtain a multiplexing gain of $R_\mathbf{H}$ because $R_\mathbf{H}$ channels are used in parallel and we say that the system has $R_\mathbf{H}$ *degrees of freedom.*

Channel unknown at the transmitter

If the channel is not known at the transmitter, the best way is to allocate equal power to each Tx antenna, resulting the covariance matrix of $\mathbf{R_x} = (\rho/M_t)\mathbf{I}_M$.

The mutual information corresponding to the capacity is:

$$I(\mathbf{X};\mathbf{Y}) = B\log_2 \det[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H] \tag{2.10}$$

Using the property $\det(\mathbf{I} + a\mathbf{A}) = \prod(1 + a\lambda_i)$ we obtain:

$$I(\mathbf{X};\mathbf{Y}) = \sum_{i=1}^{R_{\mathbf{H}}} B\log_2\left(1 + \frac{\rho}{M_t}\sigma_i^2\right) \tag{2.11}$$

This is the capacity of the channel for a particular realization of $\mathbf{H}$. Since $\mathbf{H}$ is random, the capacity can be found by averaging the equation (2.11) over the distribution of $\mathbf{H}$. However, because the transmitter does not know the channel realization, it does not know at what rate to transmit. If the capacity is less than the transmitter rate, the outage occurs. In this case, the outage probability is defined where the transmission rate is fixed to R and the probability that the capacity is less than this rate is called the outage probability. This probability is simply:

$$P_{out} = P\left(\mathbf{H} : B\log_2 \det[\mathbf{I}_{M_r} + \frac{\rho}{M_t}\mathbf{H}\mathbf{H}^H] < R\right) \tag{2.12}$$

In order to obtain this probability, one should calculate the statistics of the eigenvalues of $\mathbf{H}\mathbf{H}^H$. This problem is well studied in the recent literature [1]. However, it can be obtained by computer calculation and generation the matrix $\mathbf{H}$ following its distribution function.

If the number of Tx antennas goes to infinity, the capacity of MIMO channel given in equations(2.10) and (2.11) can be simplified. The law of large number and for zero-mean spatially white (ZMSW) model:

$$\lim_{M_t\to\infty} \frac{1}{M_t}\mathbf{H}\mathbf{H}^H = \mathbf{I}_{M_r} \tag{2.13}$$

Substituting this in (2.10) we obtain $C = M_r B\log_2(1+\rho)$. Defining $M = \min(M_t, M_r)$, it is shown that the capacity grows linearly with M in the absence of CSI in TX. Note that in this case, M is the rank of $\mathbf{H}\mathbf{H}^H$ and represents the degree of freedom of the system. So, the capacity of the MIMO system increases linearly with the degree of freedom of the system.

2.4 MIMO diversity gain: Beam forming

We saw in the previous section how using multiple antennas the capacity of the channel can be increased. However, it is possible to use the antennas to obtain diversity gain. In this case, the same information symbol, weighted by a complex scale factor, is sent over each Tx antenna. Here, the input covariance matrix has unit rank. This scheme is also referred as *beam forming*.

So there is just one input stream x and one output stream y. From x a vector $\mathbf{x}$ of dimension M_t is created and transmitted through MIMO channel. At the channel output, the vector $\mathbf{y}$ is combined to construct the system output y. So the precoding and shaping matrix are just column vectors $\mathbf{v}$ and $\mathbf{u}$. Both

[1]see the papers of Verdu and Tulino

the transmit and receive weight vectors are normalized so that $||\mathbf{u}|| = ||\mathbf{v}|| = 1$. The output is given by:

$$y = \mathbf{u}^H \mathbf{H} \mathbf{v} x + \mathbf{u}^H \mathbf{n} \tag{2.14}$$

Note that $\mathbf{u}^H \mathbf{n}$ has the statistics as $\mathbf{n}$. Beam forming provides diversity because the same symbol experiences different fading coefficient. It provides also array gain because of coherent combining of multiple signal paths (CSI is assumed known at Rx).

Channel known to the transmitter

When the channel is known at the transmitter the *optimum* precoding vector is obtain by singular decomposition of $\mathbf{H}$. In this case, we select $\sigma_1 = \sigma_{\max}$, the maximum singular value of $\mathbf{H}$, so $\mathbf{u}$ and $\mathbf{v}$ are the first columns of $\mathbf{U}$ and $\mathbf{V}$. The SNR at the output will be $\gamma = \sigma_{\max}^2 \rho$. Therefore, the resulting capacity will be

$$C = B \log_2(1 + \sigma_{\max}^2 \rho) \tag{2.15}$$

This is the capacity of a SISO channel with channel power gain $\sigma_{\max}^2$. It can be shown that for $\mathbf{H}$ a ZMSW matrix, the array gain of beam forming diversity is between $\max(M_t, M_r)$ and $M_t M_r$ and that the diversity gain is $M_t M_r$.

Channel not known to the transmitter

When the channel is not known to the transmitter, space time block coding can be used. For 2 transmit antennas, using Alamouti's scheme, an array gain of M_r and the maximum diversity gain of $2M_r$ can be obtained.

2.5 Diversity-Multiplexing Trade-offs

As we have seen, when we have multiple antennas, we may make use of them to increase the capacity of the channel and send more information. This was done by using SVD decomposition and transforming the MIMO channel into several parallel orthogonal channels. Then independent information is sent through each virtual channel. The capacity gain obtained is also called multiplexing gain. In this case, higher rate are assigned to the channel with higher singular value (which is in fact the gain of corresponding virtual channel).

On the other hand, we can make use of our antennas to increase the reliability and robustness of the link by using beam forming scheme. Here channel gains are coherently combined and high diversity gain can be obtained.

However, we can use the both techniques where some of the space-time dimensions are used for diversity gain and the remaining dimensions used for multiplexing gain. So the important issue of diversity-multiplexing trade-off is to be considered. In a multiple antenna system, the multiplexing gain is defined as follows:

$$r = \lim_{\mathtt{SNR} \to \infty} \frac{R(\mathtt{SNR})}{\log_2 \mathtt{SNR}} \tag{2.16}$$

and the diversity gain is:

$$d = - \lim_{\mathtt{SNR} \to \infty} \frac{\log P_e(\mathtt{SNR})}{\log \mathtt{SNR}} \tag{2.17}$$