Diversity Techniques

Markku Renfors
Department of Electronics and Communications Engineering
Tampere University of Technology, Finland

Contents:

- Forms of diversity
- Diversity combining methods
- Performance analysis of basic diversity schemes
- Outage probability, Outage capacity
- Further discussion about diversity concepts
  - Connections with ARQ schemes
  - Resource usage

Main references:
Forms of diversity

The basic idea of diversity reception is to use several replicas of the same information obtained through a number of independently fading channels.

If the probability that the channel gain is below a critical level is $p$, then with $L$ independent channels the probability is $p^L$.

Diversity can be reached in different ways:

- **Frequency diversity** with $L$ different channels with frequency separation exceeding the coherence bandwidth.
- **Time diversity** with $L$ different time slots with time separation exceeding the coherence time.
- **Space diversity** with $L$ antennas experiencing independently fading channels.

In multiantenna systems, diversity can be applied in various different ways, which will be discussed in later lectures of the course.
System model

Here we consider in simplified scenario with

- frequency flat channels
- slow Rayleigh fading, i.e., the channel is constant during each symbol/frame
- independent fading of the diversity branches.

Receiver antenna diversity in rich scattering environment with narrowband signals is one example of such a scenario.

- In this case, different antennas capture energy originating from a single transmission. This leads to additional array gain, i.e., for fixed transmission power, the total average SNR is $L$ times the average SNR in single antenna reception.
- In many other forms of diversity transmission, such array gain is not experienced.
- The array gain effect is not included in the following developments.
Notations

$\gamma_i$  \hspace{1cm} instantaneous SNR of branch $i$

$\gamma_{MRC, SD}$  \hspace{1cm} instantaneous SNR of combiner output

$\gamma_{av}$  \hspace{1cm} average SNR, assumed to be the same for all branches

$\gamma_{min}$  \hspace{1cm} SNR threshold in outage analysis

$g_i = \alpha_i e^{j\phi_i}$  \hspace{1cm} channel gain of branch $i$

$w_i$  \hspace{1cm} complex weight for branch $i$ in combining
Diversity Schemes

1. Maximum ratio combining, MRC
   - Optimum branch weighting by \( w_i = \alpha_i e^{-j\phi_i} \) to maximize SNR

2. Equal gain combining, EGC
   - Branch weighting by \( w_i = e^{-j\phi_i} \)
   - Coherent combining with phase alignment but constant magnitudes

3. Selection diversity, SD
   - Using one branch signal only, the one with highest instantaneous SNR

4. Switched diversity
   - Simplified implementation of SD, with one receiver branch only
Exercise: Brute-force analysis of MRC

Assume that we have two noisy observations of a signal $x[k]$:

$$y_1[k] = g_1 x[k] + n_1[k]$$
$$y_2[k] = g_2 x[k] + n_2[k]$$

Here $g_1$ and $g_2$ are known complex constants and $n_1$ and $n_2$ are independent white Gaussian noise processes with equal variances. The linear combination of the two signals is used for estimating the original signal:

$$\hat{x}[k] = w_1 y_1[k] + w_2 y_2[k]$$

Determine the optimum values of the gains such that the SNR is maximized at the combiner output. What is the resulting SNR?
Solution:

Let’s choose scaling such that signal gain = 1, i.e., the variance of the estimated signal is equal to the variance of \( x \). This also means that the estimate is unbiased. Then

\[
wg_1 + wg_2 = 1
\]

\[
\Rightarrow wg_2 = \frac{1 - wg_1}{g_2}
\]

The noise power is given by

\[
\sigma_n^2 = \left( |wg_1|^2 + |wg_2|^2 \right) \sigma_n^2 = \left( |wg_1|^2 + \left| \frac{1 - wg_1}{g_2} \right|^2 \right) \sigma_n^2
\]

\[
= \frac{|wg_1|^2 |g_2|^2 + (1 - wg_1)(1 - wg_1^*)}{|g_2|^2} \sigma_n^2
\]

\[
= \frac{|wg_1|^2 \left( |g_1|^2 + |g_2|^2 \right) + 1 - wg_1 - wg_1^*}{|g_2|^2} \sigma_n^2
\]

Since the signal variance is fixed through scaling, the SNR is maximized when the noise power is minimized, or the numerator in the above term is minimized. For obtaining the derivative, we can use the page from the book: A.H. Sayed, *Adaptive Filters*, summarizing the calculus with derivatives of complex functions, which is available on the course www-page.
\[
\frac{\partial}{\partial w_i} \left( |w_i|^2 \left( |g_1|^2 + |g_2|^2 \right) + 1 - w_i g_1 - w_i^* g_1^* \right) = \left( |g_1|^2 + |g_2|^2 \right) w_i^* - g_1 = 0
\]

\[
\Rightarrow w_1 = \frac{g_1^*}{\left( |g_1|^2 + |g_2|^2 \right)}
\]

\[
\Rightarrow w_2 = \frac{g_2^*}{\left( |g_1|^2 + |g_2|^2 \right)}
\]

The maximum SNR and the corresponding MSE are:

\[
\text{SNR}_{\text{MAX}} = \left( |g_1|^2 + |g_2|^2 \right) \frac{\sigma_x^2}{\sigma_n^2} \quad J_{\text{MIN, unbiased}} = \frac{\sigma_n^2}{\left( |g_1|^2 + |g_2|^2 \right)}
\]

This solution combines the energies of the signal observations. This is basically a special case of the maximum ratio combining solution for two observations. It has the following characteristics:

- The signal phase rotation is compensated and the two observations add up constructively.
- The observations are weighted proportional to the signal levels of the observations: The less noisy observation gets a bigger weight.
**Further discussion:** The above solution maximizes the SNR and it also minimizes the MSE in case of an unbiased estimate (in which case the error is only due to noise). But is the MSE minimized if we allow also biased solutions? We observe that scaling doesn’t effect on SNR, but it has an effect on MSE. Let us assume now that the signal gain takes an arbitrary complex value $w$. The MSE can be expressed as:

$$
|1-w|^2 \sigma_x^2 + \left( |w_1|^2 + |w_2|^2 \right) \sigma_n^2
$$

where

$$
w_1g_1 + w_2g_2 = w
$$

$$
\Rightarrow w_2 = \frac{w - w_1g_1}{g_2}
$$

Modifying the earlier derivation, the noise power becomes now:

$$
\sigma_n^2 = \left( |w_1|^2 + |w_2|^2 \right) \sigma_n^2 = \left( |w_1|^2 + \frac{|w-w_1g_1|^2}{g_2^2} \right) \sigma_n^2
$$

$$
= \left| w_1 \right|^2 \left| g_2 \right|^2 + (w-w_1g_1)(w^*-w_1^*g_1^*) \sigma_n^2
$$

$$
= \left| w_1 \right|^2 \left( \left| g_1 \right|^2 + \left| g_2 \right|^2 \right) + \left| w \right|^2 - w^*w_1g_1 - w_1^*g_1^* \sigma_n^2
$$

$$
= \left| w_1 \right|^2 \left( \left| g_1 \right|^2 + \left| g_2 \right|^2 \right) + \left| w \right|^2 - w^*w_1g_1 - w_1^*g_1^* \sigma_n^2
$$
For fixed \( w \), the noise power is minimized:

\[
\frac{\partial}{\partial g_1} \left( |w_1|^2 \left( |g_1|^2 + |g_2|^2 \right) + |w|^2 - w^* w_1 g_1 - w w_1^* g_1^* \right) = \left( |g_1|^2 + |g_2|^2 \right) w_1^* - w^* g_1 = 0
\]

\[
\Rightarrow w_1 = \frac{w g_1^*}{\left( |g_1|^2 + |g_2|^2 \right)} \quad \text{and} \quad w_2 = \frac{w g_2^*}{\left( |g_1|^2 + |g_2|^2 \right)}
\]

\[
\sigma_{n,\text{min}}^2 = \frac{|w|^2}{\left( |g_1|^2 + |g_2|^2 \right)} \sigma_n^2
\]

The MSE is now:

\[
J = |1 - w|^2 \sigma_x^2 + \frac{|w|^2}{\left( |g_1|^2 + |g_2|^2 \right)} \sigma_n^2
\]

\[
\frac{\partial J}{\partial g} = (w^* - 1) \sigma_x^2 + \frac{w^*}{\left( |g_1|^2 + |g_2|^2 \right)} \sigma_n^2 = 0
\]

\[
\Rightarrow w = \frac{1}{1 + \frac{\sigma_n^2}{\left( |g_1|^2 + |g_2|^2 \right)} \sigma_x^2}
\]

\[
\Rightarrow J_{\text{MIN}} = \frac{\sigma_n^2}{\left( |g_1|^2 + |g_2|^2 \right) + \frac{\sigma_n^2}{\sigma_x^2}}
\]

We see that the MSE solution reduces the signal gain in the low SNR region. It depends on the application, but in many cases the SNR-maximizing non-biased solution is preferred. There are cases were the minimum MSE solution is easier to find, and then it might be appropriate to compensate for the gain effect in order to remove the bias.
MRC analysis #1

The baseband signals of different diversity branches can be expressed as:

\[ r_i(t) = g_i x(t) + n_i(t) \]

where \( n_i(t) \) are independent complex white Gaussian noise processes with equal variances.

The linear combiner output can be expressed as:

\[ y(t) = \sum_{i=1}^{L} w_i (g_i x(t) + n_i(t)) = x(t) \cdot \sum_{i=1}^{L} w_i g_i + \sum_{i=1}^{L} w_i n_i(t) \]

The instantaneous signal and noise powers at the combiner output are:

\[ \sigma_y^2 = \text{E} \left[ |x(t)|^2 \right] \cdot \left| \sum_{i=1}^{L} w_i g_i \right|^2 \]

\[ \sigma_{nc}^2 = \sigma_n^2 \cdot \sum_{i=1}^{L} |w_i|^2 \]

The instantaneous SNR at the combiner output can be expressed as

\[ \gamma_c = \gamma_{av} \cdot \frac{\left| \sum_{i=1}^{L} w_i g_i \right|^2}{\sum_{i=1}^{L} |w_i|^2} \]
According to the Schwarz inequality for complex parameters, we have
\[
\left| \sum_{i=1}^{L} w_i g_i \right|^2 \leq \sum_{i=1}^{L} |w_i|^2 \cdot \sum_{i=1}^{L} |g_i|^2
\]
which holds with equality for \( w_i = c g_i^* \), where \( c \) is an arbitrary complex constant.

Choosing \( c = 1 \), the instantaneous SNR is maximized with \( w_i = g_i^* \) and it becomes
\[
\gamma_{MRC} = \gamma_{av} \cdot \sum_{i=1}^{L} |g_i|^2 = \sum_{i=1}^{L} \gamma_i
\]
where \( \gamma_i \) is the instantaneous SNR of branch \( i \).

Next we are interested to find the probability distribution of the SNR at the combiner output. The previous expression for \( \gamma_{MRC} \) can be seen as the squared magnitude sum of \( L \) independent Gaussian distributed random variables. It is known to follow the chi-square distribution with \( 2L \) degrees of freedom. It has the following probability density function:
\[
f_{\Gamma} = \frac{1}{(L-1)!} \cdot \frac{\gamma_{MRC}^{L-1}}{\gamma_{av}^L} \cdot e^{-\gamma_{MRC}/\gamma_{av}}.
\]
Below we see the resulting probability density functions (from [1]) for different number of diversity branches $N_r = L$. Here the horizontal axis is $x = \frac{\gamma_{MRC}}{\gamma_{av}}$.

We can see that the probability of obtaining low SNR values is greatly reduced by diversity, i.e., the transmission becomes much more robust.

**Figure 9.22** Normalized probability density function $f_x(x) = \frac{1}{(N_r - 1)^x} N_r^{-1} \exp(-x)$ for a varying number of $N_r$ receive antennas.
The worst kind of channel conditions are encountered with narrowband transmission in slowly fading channel, i.e., when

- signal bandwidth < coherence bandwidth
- transmission burst length (code block length) < coherence time

Then it is inevitable that some of the transmission packets are lost because the SNR over the packet is too small for detection. Such a situation is called outage. One important metric in this context is the **outage probability**, defined as the probability that the instantaneous SNR takes a value lower than a pre-defined critical threshold level for detection. The outage probability can be obtained from the cumulative distribution function as follows:

\[
P(\gamma_{MRC} < \gamma) = \int_{0}^{\gamma} f_\Gamma(\gamma_{MRC}) \, d\gamma_{MRC} = 1 - \int_{\gamma}^{\infty} f_\Gamma(\gamma_{MRC}) \, d\gamma_{MRC}
\]

\[
= 1 - e^{-\gamma_{MRC}/\gamma_{av}} \sum_{i=1}^{L} \left(\frac{\gamma_{MRC}}{\gamma_{av}}\right)^{i-1} \frac{1}{(i-1)!}
\]
The following figure (from [1]) shows the outage probability with MRC for different number of diversity branches ($N_r = L$). The horizontal axis shows the threshold level $\gamma$ in reference to the SNR value which gives 50% outage probability without diversity. We can see that MRC effectively reduces the outage probability.

**Figure 9.23** Outage probability of maximal-ratio combiner for a varying number $N_r$ of receiver antennas.
Selection diversity analysis  #1

The instantaneous SNR at SD output is readily

\[ \gamma_{SD} = \max \{ \gamma_1, \gamma_2, \ldots, \gamma_L \} \]

The probability density functions of the instantaneous SNRs of all the branches follow the exponential distribution (chi-square with two degrees of freedom):

\[ f_\Gamma (\gamma_i) = \frac{1}{\gamma_{av}} e^{-\frac{\gamma_i}{\gamma_{av}}}, \quad \gamma_i \geq 0. \]

The associated cumulative distribution is

\[ P(\gamma_i \leq \gamma) = \int_{-\infty}^{\gamma} f_\Gamma (\gamma_i) \, d\gamma_i = 1 - e^{-\gamma/\gamma_{av}}, \quad \gamma \geq 0. \]

The cumulative distribution function for the instantaneous SNR after selection combining can be written as:

\[ F_\Gamma (\gamma_{SD}) = P(\gamma_i \leq \gamma_{SD}, i = 1, \ldots, L) = \prod_{i=1}^{L} P(\gamma_i \leq \gamma_{SD}) \]
\[ = \prod_{i=1}^{L} \left(1 - e^{-\gamma_{SD}/\gamma_{av}}\right) = \left(1 - e^{-\gamma_{SD}/\gamma_{av}}\right)^L, \quad \gamma_{SD} \geq 0. \]

This is the outage probability in selection combining.
Selection diversity analysis  #2

The corresponding probability density function can be obtained as the derivative of the previous expression:

\[
f_{\Gamma}(\gamma_{SD}) = \frac{d}{d\gamma_{SD}} F_{\Gamma}(\gamma_{SD}) = \frac{L}{\gamma_{SD}} \cdot e^{-\gamma_{SD}/\gamma_{av}} \cdot \left(1 - e^{-\gamma_{SD}/\gamma_{av}}\right)^{L-1}, \quad \gamma_{SD} \geq 0.
\]

The probability density function with selection diversity for different number of diversity branches can be seen in the following figure (from [1]).

---

**Figure 9.19** Normalized probability density function \(f_x(x) = N_r \exp(-x)\left[1 - \exp(-x)\right]^{N_r-1}\) for a varying number \(N_r\) of receive antennas.
Selection diversity analysis #3

The outage probability with selection diversity for different number of diversity branches can be seen in the following figure.

While SD significantly improves the robustness of transmission link, its performance is clearly lower than that of MRC.
**Numerical example of diversity gains**

As an example, for 10 % outage probability, diversity order 4 with MRC gives about 12 dB reduction in the needed average SNR compared to the SNR needed in the no diversity case.

With selection combining, the gain is about 9 dB.

In receiver antenna diversity, these are also the gains in the total transmission power, assuming that the diversity paths are independently fading.

In other diversity schemes, the transmission power has to be split between the diversity branches, and the corresponding gains become about 6 dB for MRC and 3 dB for SC.
BER analysis

It is interesting to look at the bit error rate performance with and without diversity. Analytical derivations for BER analysis, mostly for binary modulations, can be found in the *Digital Communications* book by J.G. Proakis. However, we skip the derivations and show only the resulting plots.

- With high SNR, the error rate in Rayleigh fading channel is inversely proportional to the SNR.
- This is in contrast with the AWGN channel case, where the decrease is exponential.
- We can see that the diversity order $L$ determines the slope with high SNR.
The Shannon capacity of a slowly flat-fading channel without CSI (channel state information) in the transmitter can be expressed as

\[ C_S = \int_0^\infty B \log_2 (1 + \gamma) p(\gamma) d\gamma \]

where \( B \) is the bandwidth and \( \gamma \) is the instantaneous SNR [2].

- This is the theoretical capacity, and in practice it would be quite difficult to approach it.
- This so called ergodic capacity is smaller than the capacity calculated form the average SNR \( \gamma_{av} \).

For defining the outage capacity, we make the following assumptions:

- No CSI on transmitter side.
- The instantaneous SNR \( \gamma \) is constant over transmission blocks.
- The transmitter encodes the data blocks in a such a way that the specified outage probability is reached, i.e, for target received SNR of \( \gamma_{min} \).

Then the capacity can be expressed as:

\[ C_{Out} = B(1 - P_{out}) \log_2 (1 + \gamma_{min}) \]

So the fact that the SNR is most of the time higher than \( \gamma_{min} \) cannot be utilized!
Outage capacity #2

The following figure (from [2]) shows the normalized outage capacity as a function of outage probability for slow Rayleigh fading with 20 dB average SNR.

We can see significant loss of capacity with low outage probability.

The average rate of correctly received bits can be maximized through the choice of the threshold. As an example, for $P_{\text{out}} = 0.1$ this normalized rate is about $0.9 \cdot 3.5 = 3.15$ bits/Hz and for $P_{\text{out}} = 0.01$ the rate is about $0.99 \cdot 1 = 0.99$ bits/Hz.

Figure 4.2: Normalized capacity (C/B) versus outage probability.
About diversity concepts #1

Narrowband transmission in slowly fading channel is essentially the most difficult case, especially regarding the outage probability. There is basically no diversity!

There are various approaches to include diversity in the transmission:

- frequency diversity using coded multicarrier modulation
- CDMA techniques with RAKE receiver
- multicarrier CDMA is basically a simple frequency diversity scheme from a single users point of view, but it allows multiple users to share the same transmission band by using orthogonal codes.
- multiantenna techniques with spatial diversity
  - receiver antenna diversity
  - transmit diversity
  - MIMO
- wideband single-carrier transmission has inherent frequency diversity
  - the notches in frequency response don’t have catastrophic effects on performance
  - the RAKE receiver is another interpretation, which shows the presence of diversity in an intuitive way
  - channel equalization becomes a challenge, but frequency-domain channel equalization is considered to be a solution with realistic complexity
In basic diversity schemes, the same data is transmitted in different diversity branches. This can be seen as repetition coding, i.e., as a very elementary error correction coding technique. However, repetition coding has fairly high overhead in terms of spectral efficiency.

Especially in CDMA schemes, a certain number of users are using the same frequency band at the same time, thus improving the system spectral efficiency, while exploiting frequency diversity. The interferences between users are ideally avoided by using orthogonal codes. However, the orthogonality is usually lost in practice due to various reasons, leading to the need for more complicated detection techniques for optimum system performance.

An alternative approach is to combine diversity in some form with more efficient error control coding methods (e.g., LDPC or turbo-codes), instead of repetition coding. The coding can be done in such a way that the code word bits have time/frequency/spatial diversity. Such approaches lead to spectrally efficient transmission schemes with certain amount of diversity and robustness against fading, and reduced outage probability.
Diversity and ARQ

ARQ (automatic repeat request) schemes can also be considered in the context of diversity, like ‘diversity on demand’:

- If it is not possible to detect the received data frame, then a retransmission is requested.
- Basic ARQ is associated with selection diversity, i.e., the earlier transmissions are ignored and only the latest one is used for detection.
- In HARQ (hybrid ARQ), information about all received data packets is combined to improve detection performance. Among various other schemes, diversity combing methods could be considered as possible ways of combining them. However, in advanced HARQ schemes, the data carried by retransmission is usually incremental, i.e., not the same as the original data packet.

Outage capacity considerations in the ARQ context help us also to understand the dimensioning of modern wireless systems:

- Outage probability with no diversity gives us the probability of (first) retransmission. On the other hand, we have seen that low outage probability results in low capacity. Then it becomes obvious that fairly frequent retransmissions might be acceptable to maximize the capacity.
- In fact the optimization of current wireless systems (which typically include HARQ) has resulted in parametrizations where the coded block error rate is in the order of 5 … 10 %. The uncoded BER is typically in the same order, and coded BER in the order of 0.0001 … 0.001.
In receiver antenna diversity, the diversity gain is achieved without a need for additional transmission resources, in terms of bandwidth/spectrum efficiency or transmission energy. Of course, multiple antennas and multiple receiver chains are needed in the receiver.

In transmit (antenna) diversity, the goal is to reach the diversity gain without losing spectrum efficiency. This is possible in some specific cases (notably using the Alamouti code with two antennas), but not generally. In transmit diversity, the overall transmission energy is split between diversity branches (transmit antennas). Multiple transmitters are needed, however, with reduced power level if the overall transmission power is fixed.

In basic frequency diversity or time diversity schemes, additional spectral resources are needed to reach diversity. Also here, the overall transmission energy is split between diversity branches.

Various coding schemes are aiming at good compromise between time/frequency diversity and spectrum efficiency.

In ARQ schemes, diversity is included in terms of retransmissions, when needed. The effect on spectrum efficiency and energy consumption depends on the retransmission probability.