

Diversity and Maximum-Ratio Combining (MRC)

Brennan's Paper

石政修

NCKU

February 19, 2026

1. Background
2. Assumption
3. Diversity Method
4. Combining Method
5. Maximal-Ratio Combining
6. Appendix

Multipath Channel

- Usually modeled by the Rayleigh fading model.
- The amplitude of the channel response follows a Rayleigh distribution.
- The phase follows a Uniform distribution.

Problem: Deep Fading

- Occurs when signal paths add up destructively (destructive interference).
- Causes a severe drop in the channel gain.
- The weak received signal can occur at a specific time, frequency, or spatial location.

Solution: Diversity

- Exploits multiple independent signal paths (channels).
- If the probability of deep fading is p , with m independent channels, the joint probability of deep fading drops to p^m .

Assumption

(A) The noise in each channel is independent of the signal, and additive: received signal $\underline{r_i(t) = s_i(t) + n_i(t)}$, where $s_i(t)$ and $n_i(t)$ are the signal and noise components in the i th channel.

(B) Flat fading (so the channel gain $h_i(t)$ is multiplicative): $s_i(t) = h_i(t)m(t)$, where $m(t)$ is the transmitted signal. And **slow fading**: $h_i(t) = h_i$ over the symbol duration. Hence, $\underline{s_i(t) = h_i m(t)}$.

(C) The noise components are uncorrelated and have zero means.

(D) Channel $h_i(t)$ are independent.

Note. In (B), flat fading means the multipath delay spread is much smaller than the symbol duration, i.e., we can ignore the effect of delay. Hence, the original convolution becomes multiplication.

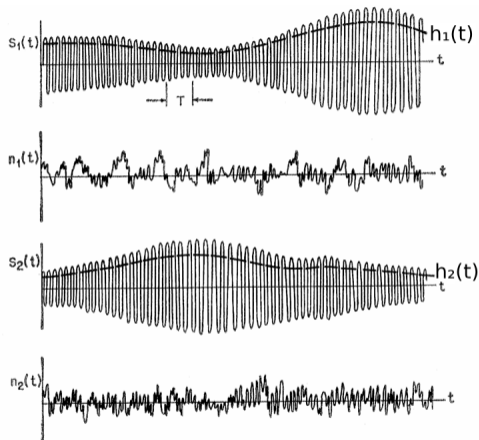
Assumption

Ergodic: Time averages equal ensemble averages.

Normalized average power: $\overline{m^2(t)} = 1$. Hence, the average power of signal component is

$$\begin{aligned}\overline{s_i^2(t)} &= \frac{1}{T} \int_{t-T}^t s_i^2(\tau) d\tau \\ &= \frac{1}{T} \int_{t-T}^t (h_i m(\tau))^2 d\tau \\ &= h_i^2 \frac{1}{T} \int_{t-T}^t m^2(\tau) d\tau \\ &= h_i^2\end{aligned}$$

Therefore, $h_i = h_i(t)$ is the rms value of $s_i(t)$. As shown in the figure, $h_i(t)$ is proportional to the signal's envelope (peak value).



Signals and noise in two diversity system. Assumption (A) to (C).

Diversity System

A system in which one has available two or more closely similar copies of desired signal (i.e., $s_1(t), \dots, s_N(t)$).

Diversity Method

- Methods to obtain two or more signal copies $s_i(t)$.

Linear Diversity Combining Method

- Methods to combine the received copies to obtain an improved signal $y(t)$.
- For example, $y(t) = r_1(t) + r_2(t) = [s_1(t) + s_2(t)] + [n_1(t) + n_2(t)]$. Signal components add up coherently, while noise components add up incoherently. Thus, the Signal-to-Noise Ratio (SNR) is improved.
- General form: $y(t) = \sum_{i=1}^N w_i r_i(t)$, where w_i are the combining weights we want to determine.

Note. The example above assumes coherent addition, which requires phase alignment. This is known as predetection combining. In contrast, postdetection combining signals after the detector (envelope or power) where phase information is lost.

Concept:

Exploits the time-varying nature of the fading channel. For example, send the same symbol three times.

Sacrifice:

Data rate and latency of repetition (or interleaving).

Constraint:

The transmission interval must be larger than the coherence time. Otherwise, channels cannot be viewed as independent, and diversity is no longer exploitable. That is, *fails in slow fading and static channels*.

Concept:

Exploits the frequency-selective nature of the fading channel. For example, send the same information on three different carrier frequencies.

Sacrifice:

Spectral Efficiency.

Constraint:

The frequency separation must be larger than the coherence bandwidth. *Fails in flat fading* because all frequencies fade together.

Concept:

Exploits the spatial independence of the fading channel when the separation between multiple antennas is sufficiently large.

Sacrifice:

Increases hardware complexity.

Constraint:

The required separation distance depends on the details of the propagation environment, and can range from as little as one half wavelength (for rich scattering environments) to many wavelengths (for the case of LOS).

Advantage:

Higher data rate than time diversity and higher spectral efficiency than frequency diversity. And no latency penalty (unlike time diversity). Moreover, it *works effectively in flat and slow fading channels*.

Combining Method

The most common receive diversity combining techniques are listed below, where combining is performed at the receiver.

Switched Combining

- The receiver switches to one channel i if its SNR (γ_i) is over the threshold. Then $w_i = 1$ and $w_{j \neq i} = 0$. When γ_i is below the threshold, switch to the next channel.

Selection Combining

- The receiver selects the channel i if $\gamma_i \geq \gamma_j \forall j$. Then $w_i = 1$ and $w_{j \neq i} = 0$.

Equal-Gain Combining (EGC)

- All the weights are equal: $w_i = 1 \forall i$.

Maximal-Ratio Combining (MRC)

- The weights are determined to maximize the SNR: $w_i = h_i / \overline{n_i^2}$ (RMS value of the signal divided by the average noise power). The maximal SNR is $\gamma = \sum_{i=1}^N \gamma_i$.

Average SNR of Selection Combining

As usual, assume h_i follows a Rayleigh distribution. For convenience, we let $E\{h_i^2(t)\} = 1$ and $E\{n_i^2(t)\} = 1$. Then $f(h_i) = 2h_i e^{-h_i^2}$ and $\gamma_i = \overline{s_i^2(t)}/\overline{n_i^2(t)} = h_i^2$. Compute the PDF of γ_i :

$$g(\gamma_i) = f(h_i) \left| \frac{dh_i}{d\gamma_i} \right| = 2h_i e^{-h_i^2} \frac{1}{2h_i} = e^{-\gamma_i},$$

that is, an exponential decay distribution, and the CDF is

$$G(\gamma_i) = 1 - e^{-\gamma_i}.$$

The CDF of the SNR for selection combining from N channels is

$$S_N(\gamma) = Pr(\gamma_i < \gamma \forall i) = (G(\gamma))^N = (1 - e^{-\gamma})^N \approx \boxed{\gamma^N \text{ (for small } \gamma)}.$$

Hence, the average SNR (pf. in Appendix 2) is

$$E\{\gamma\} = \int_0^\infty \gamma S_N(\gamma) d\gamma = \int_0^\infty \gamma dS_N(\gamma) = \boxed{\sum_{k=1}^N \frac{1}{k} \approx \ln N \text{ (for large } N)}.$$

Weights of MRC

The general form of combining is $y(t) = \sum_{i=1}^N w_i r_i(t)$ and with assumption (A) we get:

$$y(t) = \sum w_i r_i(t) = \sum w_i (s_i(t) + n_i(t)) = s(t) + n(t),$$

where $s(t) = \sum w_i s_i(t)$ and $n(t) = \sum w_i n_i(t)$.

The SNR we want to maximize is defined as

$$\gamma = \frac{E\{s^2(t)\}}{E\{n^2(t)\}} = \frac{E\left\{\left(\sum w_i s_i(t)\right)^2\right\}}{E\left\{\left(\sum w_i n_i(t)\right)^2\right\}}.$$

By assumption (B), the numerator becomes

$$E\left\{\left(\sum w_i s_i(t)\right)^2\right\} = E\left\{m^2(t) \left(\sum w_i h_i\right)^2\right\} = E\{m^2(t)\} \left(\sum w_i h_i\right)^2 = \left(\sum w_i h_i\right)^2.$$

Weights of MRC

By assumption (C), the denominator becomes

$$\begin{aligned} E \left\{ \left(\sum_i w_i n_i(t) \right)^2 \right\} &= E \left\{ \sum_i w_i^2 n_i^2(t) + \sum_{i \neq j} w_i w_j n_i(t) n_j(t) \right\} = E \left\{ \sum_i w_i^2 n_i^2(t) \right\} \\ &= \sum w_i^2 E \{ n_i^2(t) \} = \sum w_i^2 \overline{n_i^2(t)}. \end{aligned}$$

Thus, with Cauchy—Schwarz inequality, the SNR is

$$\gamma = \frac{(\sum w_i h_i)^2}{\sum w_i^2 \overline{n_i^2(t)}} = \frac{(\sum w_i h_i)^2}{\sum w_i^2 \left(\sqrt{\overline{n_i^2(t)}} \right)^2} \leq \sum \frac{h_i^2}{\left(\sqrt{\overline{n_i^2(t)}} \right)^2} = \sum \frac{h_i^2}{\overline{n_i^2(t)}} = \sum \gamma_i,$$

where the “=” occurs iff for any $k \in \mathbb{R}$,

$$w_i \sqrt{\overline{n_i^2(t)}} = k \frac{h_i}{\sqrt{\overline{n_i^2(t)}}} \quad \text{or} \quad \boxed{w_i = k \frac{h_i}{\overline{n_i^2(t)}}}.$$

MRC in Complex Vector

Let $\mathbf{r}(t)$, $\mathbf{s}(t)$, $\mathbf{n}(t)$ and weight \mathbf{w} be $N \times 1$ complex vectors. The general form of combining is $y(t) = \mathbf{w}^H \mathbf{r}(t) = \sum_{i=1}^N w_i^* r_i(t)$. With assumption (A) we get:

$$y(t) = \mathbf{w}^H \mathbf{r}(t) = \mathbf{w}^H (\mathbf{s}(t) + \mathbf{n}(t)) = s(t) + n(t),$$

where $s(t) = \mathbf{w}^H \mathbf{s}(t)$ and $n(t) = \mathbf{w}^H \mathbf{n}(t)$.

For complex signals, the SNR is:

$$\gamma = \frac{E\{|s(t)|^2\}}{E\{|n(t)|^2\}} = \frac{E\{|\mathbf{w}^H \mathbf{s}(t)|^2\}}{E\{|\mathbf{w}^H \mathbf{n}(t)|^2\}}.$$

By assumption (B), let $\mathbf{s}(t) = \mathbf{h}m(t)$ where $\mathbf{h} = [h_1 \ h_2 \ \dots \ h_N]^T$ is the complex channel vector, and $h_i = |h_i|e^{j\angle h_i}$ is complex-valued. The numerator becomes:

$$E\{|\mathbf{w}^H \mathbf{s}(t)|^2\} = E\{|m(t)|^2 |\mathbf{w}^H \mathbf{h}|^2\} = E\{|m(t)|^2\} |\mathbf{w}^H \mathbf{h}|^2 = |\mathbf{w}^H \mathbf{h}|^2.$$

MRC in Complex Vector

By assumption (C), complex noise branches are uncorrelated ($E\{n_i(t)n_j^*(t)\} = 0$ for $i \neq j$). Then the noise covariance matrix \mathbf{R}_n becomes diagonal:

$$\mathbf{R}_n = E\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \text{diag}(E\{|n_1|^2\}, \dots, E\{|n_N|^2\}).$$

Thus, the denominator (noise power) simplifies:

$$E\{|\mathbf{w}^H \mathbf{n}(t)|^2\} = E\{\mathbf{w}^H \mathbf{n}(t)\mathbf{n}^H(t)\mathbf{w}\} = \mathbf{w}^H \mathbf{R}_n \mathbf{w}.$$

To maximize SNR, we define $\tilde{\mathbf{w}} = \mathbf{R}_n^{1/2} \mathbf{w}$ and $\tilde{\mathbf{h}} = \mathbf{R}_n^{-1/2} \mathbf{h}$. Applying the Cauchy–Schwarz inequality ($|\tilde{\mathbf{w}}^H \tilde{\mathbf{h}}|^2 \leq \|\tilde{\mathbf{w}}\|^2 \|\tilde{\mathbf{h}}\|^2$):

$$\gamma = \frac{|\mathbf{w}^H \mathbf{h}|^2}{\mathbf{w}^H \mathbf{R}_n \mathbf{w}} = \frac{|\tilde{\mathbf{w}}^H \tilde{\mathbf{h}}|^2}{\|\tilde{\mathbf{w}}\|^2} \leq \|\tilde{\mathbf{h}}\|^2 = \mathbf{h}^H \mathbf{R}_n^{-1} \mathbf{h} = \sum_{i=1}^N \gamma_i,$$

where the “=” occurs iff $\tilde{\mathbf{w}} = k\tilde{\mathbf{h}}$ for any complex k . Reverting to \mathbf{w} :

$$\mathbf{R}_n^{1/2} \mathbf{w} = k\mathbf{R}_n^{-1/2} \mathbf{h} \quad \text{or} \quad \boxed{\mathbf{w} = k\mathbf{R}_n^{-1} \mathbf{h}}.$$

MRC in Complex Vector

Assume all receiver branches have identical and independent noise power N_0 . The covariance matrix becomes $\mathbf{R}_n = N_0\mathbf{I}$. The optimal weight vector simplifies to:

$$\mathbf{w} = k(N_0\mathbf{I})^{-1}\mathbf{h} = \frac{k}{N_0}\mathbf{h}.$$

Since scaling all weights by a constant does not change the SNR ratio, we choose: $\mathbf{w} = \mathbf{h}$. Substituting back into the general combining formula $y(t) = \mathbf{w}^H\mathbf{r}(t)$, we obtain:

$$\boxed{y(t) = \mathbf{h}^H\mathbf{r}(t)}.$$

Note.

This complex-valued channel is called memoryless SIMO channel: $\mathbf{r}(t) = m(t)\mathbf{h}(t) + \mathbf{n}(t)$. And its optimal combining is $y(t) = \mathbf{h}^H\mathbf{r}(t)$.

Average SNR and CDF of MRC

The average SNR of MRC if $E\{\gamma_i\} = 1$:

$$E\{\gamma\} = E\left\{\sum_{i=1}^N \gamma_i\right\} = \sum_{i=1}^N E\{\gamma_i\} = N.$$

From $f(\gamma_i) = e^{-\gamma_i}$, the joint PDF:

$$f(\gamma_1, \gamma_2, \dots, \gamma_N) = f(\gamma_1)f(\gamma_2) \dots f(\gamma_N) = e^{-(\gamma_1 + \gamma_2 + \dots + \gamma_N)} = e^{-\gamma}.$$

Using change of variable, the PDF of SNR for MRC from N channels is

$$g(\gamma) = \frac{\gamma^{N-1} e^{-\gamma}}{(N-1)!}.$$

The correspondent CDF is an Erlang distribution:

$$G(\gamma) = e^{-\gamma} \sum_{k=N}^{\infty} \frac{\gamma^k}{k!} \approx \frac{\gamma^N}{N!} \text{ (for small } \gamma).$$

Conclusion

Assuming independent Rayleigh fading channels with unit average SNR per branch. For a large number of branches N and small threshold γ :

	Selection	EGC	MRC
Average SNR $E\{\gamma\}$	$\ln N$	$[1 + (N - 1)\frac{\pi}{4}]$	N
Outage Prob. (CDF)	γ^N	$\propto \gamma^N$	$\gamma^N/N!$
Hardware Complexity	Low	Medium (Phase only)	High (Phase & Amplitude)

Note.

- **Diversity Gain:** All three methods achieve a diversity order of N , sharply reducing the outage probability from γ to γ^N .
- **Array Gain:** MRC is the optimal combining scheme. Unlike selection's logarithmic growth, MRC's average SNR grows linearly with N , and it further suppresses the outage probability by a factor of $N!$.

Appendix 1

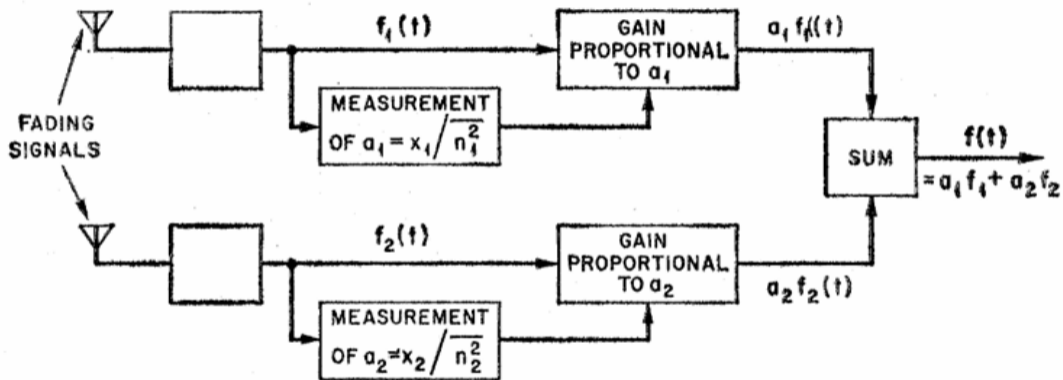


Figure: MRC Scheme

Appendix 2

Consider the following integral, where $S_N(\gamma) = (1 - e^{-\gamma})^N$:

$$\begin{aligned} I &= \int_0^{\infty} \gamma dS_N(\gamma) = \left[\gamma(S_N(\gamma) - 1) \right]_0^{\infty} - \int_0^{\infty} (S_N(\gamma) - 1) d\gamma \\ &= \int_0^{\infty} (1 - S_N(\gamma)) d\gamma = \int_0^{\infty} (1 - (1 - e^{-\gamma})^N) d\gamma. \end{aligned}$$

Let $x = 1 - e^{-\gamma}$, then $d\gamma = \frac{dx}{1-x}$:

$$\begin{aligned} I &= \int_0^1 \frac{1 - x^N}{1 - x} dx = \int_0^1 (1 + x + x^2 + \dots + x^{N-1}) dx \\ &= \left[x + \frac{x^2}{2} + \frac{x^3}{3} + \dots + \frac{x^N}{N} \right]_0^1 = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{N}. \quad \square \end{aligned}$$

The End