Channel Fading Models

We have now seen two of the most famous wireless path loss models. More generally, in cellular wireless systems, we tend to model the received signal power P_R through the following equation:

$$P_R = \alpha^2 \cdot 10^{x/10} \cdot g(d) \cdot P_T \cdot G_T \cdot G_R \tag{1}$$

where α^2 and $10^{x/10}$ represent, respectively, the effects of small scale and large scale fading. The $g(d)P_T G_T G_R$ component is what we have already seen: with $g(d)$ representing the inverse variation of power with distance, this gives the **average received power** at the receiver. These fading factors, then, aim to model the statistical variations experienced around the area-mean-power, with x and α each representing random variables.

Shadow Fading (large-scale)

Shadow fading is in general caused by variations in signal power due to signal attenuation from terrain obstructions, such as hills, buildings, or even leaves. This type of fading manifests over relatively long distances (many wavelengths, on the order of meters), and thus is relatively slowly varying. Over these distances, the measured signal power may differ substantially at different locations, even when we are at the same distance from the transmitter.

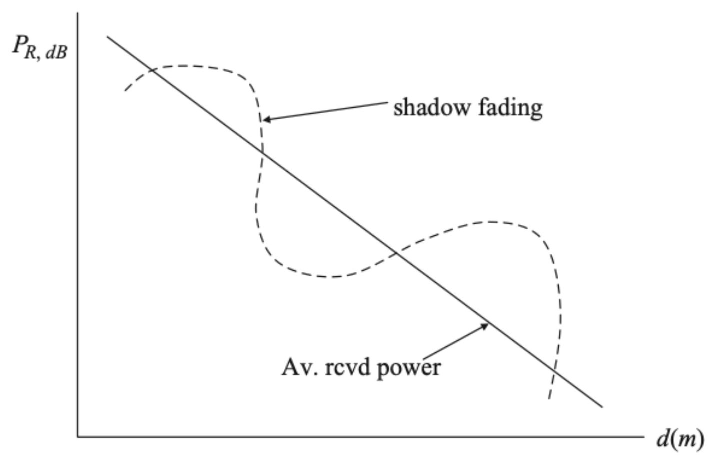
The shadow-fading random variable x, when measured in dB, is often assumed to follow a Gaussian distribution with $\mu = 0$ and variance σ^2. In other words, $X \sim \mathcal{N}(0, \sigma^2)$, where

$$f_X(x) = \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \tag{2}$$

The factor $10^{x/10}$, then, converts from dB to Watts.

In practice, σ^2 is found to typically range between 6 and 10 dB. The log-normal distribution based on these values is found to be a good approximation in practice. The rationale for this is that there are often many obstacles between the transmitter and receiver. The existence of an obstacle s (e.g., a building) with width d_s attenuates a signal by 10^{-ad_s}, where $a > 0$ is some constant. With N potential obstacles, the attenuation is $10^{-a\sum_{s=1}^{N} d_s}$. When N is large, by the central limit theorem, $\sum_{s=1}^{N} d_s$ becomes Gaussian (assuming the obstacles are i.i.d.).

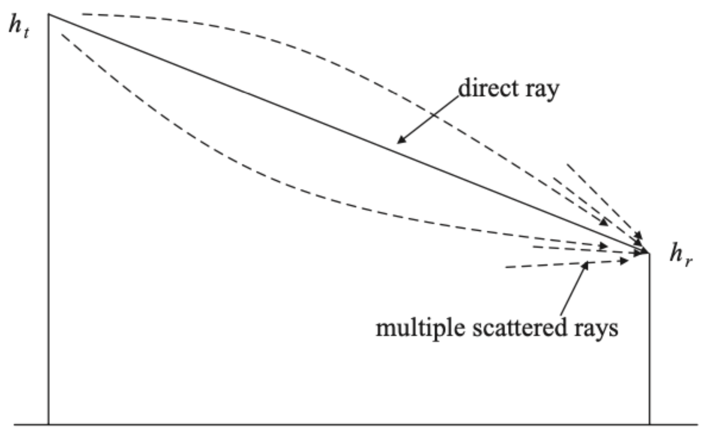
Shadow fading occurs over distances of ten/hundreds of wavelengths, i.e., $d \gg \lambda$. The figure below shows an example of how we would expect the observed received power (in dB) to fluctuate around the average as the distance increases.



Multi-path fading (small-scale)

Multi-path fading is used to describe the rapid fluctuation in the amplitude of a radio signal over a short period of time or over a short travel distance (on the order of wavelengths). Variations in path loss and large-scale fading, by contrast, are negligible over such short distances and are not thought of with a temporal component.

Multi-path fading is caused by the interference between two or more versions of the transmitted signal, which arrive at the receiver at slightly different times. Previously, we saw the two-ray pathloss model. Each ray is, in reality, made up of the superposition of multiple rays, due to scattering by buildings and other obstacles along a given path. Thus, the direct ray appearing at the receiver is a superposition of many rays, each differing randomly in amplitude and phase because of the scattering. This is depicted below:



Multi-path fading is modeled by the α variable in (1). When the number of multipath signals is large (> 6), and none of them dominate, α follows a **Raleigh distribution**, which has the following PDF:

$$f_\alpha(a) = \frac{a}{\sigma_R^2} e^{-\left(\frac{a^2}{2\sigma_R^2}\right)}. \tag{3}$$

This tends to hold in macrocellular systems. On the other hand, when only one component dominates (e.g., when there is a strong direct path), α follows a **Ricean distribution** (sometimes called Rician):

$$f_\alpha(a) = \frac{a}{\sigma_R^2} e^{-\left(\frac{a^2+A^2}{2\sigma_R^2}\right)} I_o\left(\frac{aA}{\sigma_R^2}\right), \tag{4}$$

where A is the amplitude of the dominant signal, and $I_o(x)$ is the modified Bessel function of the first kind and 0-order, defined as:

$$I_o(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{x\cos\theta} d\theta \tag{5}$$

This is a more valid model in microcellular systems. We see that as $A \to 0$, Ricean approaches Raleigh, i.e., the case where no component is dominating.

To see how these distributions arise, let the transmitted signal be $s(t) = E_T e^{j\omega_c t}$. The received signal through the kth path can then be written as $a_k E_T e^{j(\omega_c t - \phi_k)}$, where a_k is the attenuation and ϕ_k is the phase shift. The total received signal is then

$$a(t) = \sum_k a_k E_T e^{j(\omega_c t - \phi_k)} = E_T e^{j\omega_c t} \left[\sum_k a_k e^{-j\phi_k} \right], \tag{6}$$

whose magnitude is determined by

$$\left| \sum_k a_k e^{-j\phi_k} \right| = \Big| \underbrace{\sum_k a_k \cos\phi_k}_{X} + j \underbrace{\Big(-\sum_k a_k \sin\phi_k \Big)}_{Y} \Big|. \tag{7}$$

When none of the multipath components dominate, by the central limit theorem, X and Y can be approximated by independent Gaussian random variables with zero mean. In this case,

$$\left| X + jY \right| = \sqrt{X^2 + Y^2} \tag{8}$$

can be shown to follow a Raleigh distribution.

On the other hand, if there is a dominant term, without loss of generality, assume the dominant term is $a_0 e^{j\phi_0} = A$. Additionally, since all phase is relative, we can reference them all to the phase of this dominant term, and take $\phi_0 = 0$ for simplicity. In this case,

$$\left| (X + A) + jY \right| = \sqrt{(X + A)^2 + Y^2} \tag{9}$$

can be shown to follow a Ricean distribution. You will prove both of these facts in a homework problem.

Stochastic Channel Characterization

Our focus so far has been on relating the total power P_T and P_R transmitted and received, respectively. More generally, the received signal power is a stochastic process, both in the time domain and in the frequency domain.

Equation (1) can be seen as characterizing the marginal distribution of the received signal power, not factoring in statistics of the transmit signal. In reality, a modulated carrier signal would be transmitted, with the modulation carrying the information to be transmitted. In particular, the transmission would be of the form $s(t)e^{j\omega_c t}$, where $s(t)$ is the information-bearing **baseband signal**, also called the modulating signal. We aim to understand how properties of $s(t)$, and of the channel, impact the received signal $a(t)$. To do so, we can model the joint probability density function, or correlation, between different aspects of the received signal.

In general, let $x(t)$ be a random process varying with time. At any selected time t, the instantiation of this process is some random variable, say X_t. A measure of the variation of $x(t)$ with time is given by its **autocorrelation function**:

$$R_x(t_1, t_2) = \mathbb{E}[X_1 X_2]. \tag{10}$$

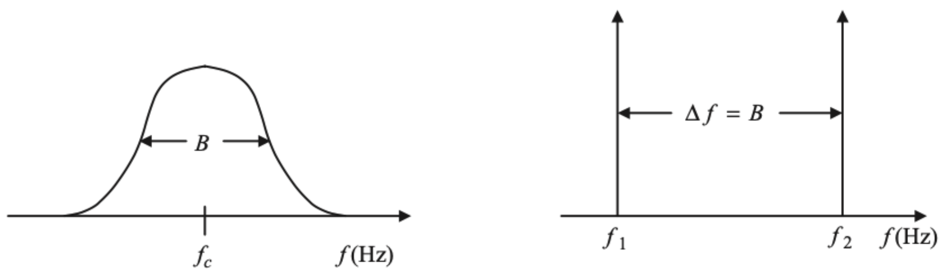
Additionally, we can define the **correlation coefficient**

$$\rho_x = \frac{\mathbb{E}[(X_1 - \mathbb{E}(X_1)) \cdot (X_2 - \mathbb{E}(X_2))]}{\sqrt{\mathrm{Var}(X_1) \cdot \mathrm{Var}(X_2)}}. \tag{11}$$

As $\rho \to 0$, $x(t_1)$ and $x(t_2)$ are uncorrelated, i.e., the dependence is low. On the other hand, as $\rho \to 1$, $x(t_1)$ and $x(t_2)$ are highly correlated, i.e., the dependence is high.

To measure the effect of fading, we want to consider this correlation as we observe the received signal at two different times, t and $t + \Delta t$, and at two different frequencies, f and $f + \Delta f$. The reason for considering two frequencies is visualized below. If the bandwidth of $s(t)$ is $B/2$ (in Hz), then the bandwidth of the transmitted $s(t)e^{j\omega_c t}$ is B (shown in (a)). B will be inversely proportional to the **symbol duration** T_s (in sec), which depends on the digital coding scheme. A wideband signal will have a small T_s and a large B (for example, 1 MHz in IS-95, and 20 MHz in LTE).

Taking the two extreme frequencies $f_c - \frac{B}{2}$ and $f_c + \frac{B}{2}$ (shown in (b)), we can then see whether the fading characteristics change substantially over this range. If they do, the signal can be seen as passing through a filter, and thus can be distorted.



(a) Modulated carrier spectrum (b) Two-carrier model

We may then calculate the correlation coefficient ρ either in time or in frequency:

- *In time*: If it is close to 1, the signal power changes slowly. If it is close to 0, the signal power changes rapidly.
- *In frequency*: If it is close to 1, different frequency components of the signal experience the same level of fading, referred to as **flat fading**. If it is close to 0, then different components experience frequency-selective fading, leading to **distortion**.

Which is better? Is it better to have a channel whose fading levels change slowly (or quickly) in time, or slowly (or quickly) in frequency?

- If the signal power changes slowly in time, the transmitter will have enough time to estimate the channel, and we can design effective control schemes (e.g., power control). However, there may be long periods of **deep fades**, where the received signal drops below a given amplitude.
- If the signal power changes slowly in frequency, then signal distortion will be small, making it easy to do **equalization**. The entire band may run into deep fades, perhaps for only short periods of time.
- On the contrary, for a frequency-selective channel, the chance of having some frequency component with a good signal power is higher. This is useful for having multiple narrow band carriers, e.g., in OFDM.

As we can see, understanding these properties is important for designing appropriate communication and multiple-access schemes. We will next discuss the factors that contribute to time- and frequency-selectivity of the channel, and also typical approaches to deal with it.

In (6), both the received amplitudes a_k and phases ϕ_k are stochastic processes. As we saw in the two-ray model, the phase randomness typically dominates. There are two major contributors to variations in ϕ_k:

(1) Travel Distance

Different paths travel different distances. This leads to a dispersion in the time-shift (or delay) τ_k (in sec) among different paths. The impact on the phase of path k is

$$\theta_k = \omega_c \tau_k = 2\pi f_c \tau_k. \tag{12}$$

We define the **delay spread** as the difference in delay τ_k between the first path and last path to arrive. Typical values for the delay-spread are 0.2 to 0.5 μsec in suburban areas, and 3 to 8 μsec in urban areas.

The effect of this time dispersion will be manifested by frequency-selectivity, since different frequency components see different amounts of phase shift. For digital signals, this distortion manifests as **inter-symbol interference** (ISI), where successive digital signals overlap into adjacent symbol intervals.

As a rough estimate, consider the case when τ_k is deterministic (i.e., non-random). For two frequency components f_c and $f_c + \Delta f$, the phase difference of the two paths will vary by $2\pi \Delta f \tau_k$. If $2\pi \Delta f \tau_k = \pi$, then significant variations of the fading levels will be observed by the two frequency components (i.e., a complete 180 degree phase shift). This occurs when

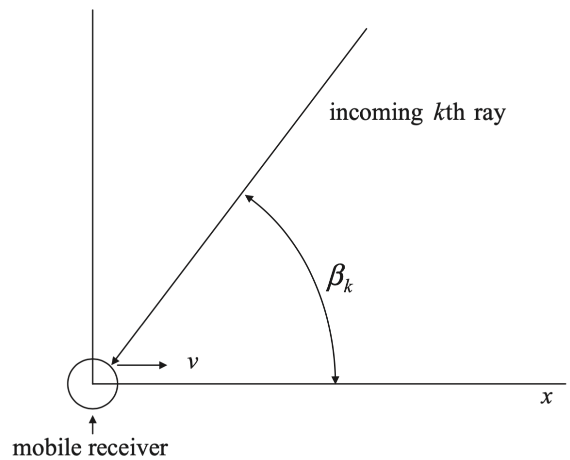
$$\Delta f = \frac{1}{2\tau_k}, \tag{13}$$

which is often referred to as the **coherence bandwidth**.

Δf determines the frequency-selectivity of the channel. If $B > \Delta f$, we will experience a frequency-selective channel. Otherwise, the channel will be "flat" over the range of frequencies.

(2) Doppler Effect

So far, we have ignored the potential mobility of transmitters and receivers. Motion in the receiver terminal can effect the amount of fading experienced.



Consider the diagram above. A mobile receiver is moving at a speed of v (in m/s), and the incoming kth ray forms an angle β_k with the receiver's travel path. The **Doppler effect** leads to a frequency shift of

$$f_k = \frac{v \cos \beta_k}{\lambda}. \tag{14}$$

Different paths can have different f_k due to their angles of arrival. The maximum Dopper shift occurs when $\beta_k = 0$ (i.e., the ray arrives in exactly the opposite direction of the travel path), given by

$$f_m = \frac{v}{\lambda}. \tag{15}$$

For example, suppose $v = 100$ km/hr $= 28$ m/s. A carrier frequency $f_c = 1$ GHz, which corresponds to $\lambda = 0.3$ m, has a maximum shift of $f_m = 90$ Hz, which would imply a phase change of 2π every $1/90$ sec. At a slower speed of $v = 0.1$ m/s, the shift would only be $f_m = 0.33$ Hz. At $f_c = 30$ GHz (i.e., mmWave), we have a smaller $\lambda = 0.01$ m, so even $v = 0.1$ m/s would still experience a maximum shift of $f_m = 10$ Hz.

The effect of this Dopper shift will be manifested by time-selectivity, since the value of the phase shift changes in time. As a rough estimate, consider the case when f_k is deterministic. For two time instances t and $t + \Delta t$ (i.e., the symbol interval), the phase difference of the two paths will vary by $2\pi f_k \Delta t$. A phase shift of π occurs when

$$\Delta t = \frac{1}{2 f_k} \tag{16}$$

which is often referred to as the **coherence time**. Δt determines the time-selectivity, or rate of fading, for the channel.

Correlation in Frequency and Time

When f_k and τ_k are random, we must study the how the correlation varies in frequency and time. Specifically, the combined effect of receiver mobility and fading on signal transmission is quantified by the correlation coefficient ρ_a of two received signals.

If we assume Rayleigh fading and an exponentially distributed incremental delay between signals, then after some detailed analysis, it can be shown that

$$\rho_a(\Delta t, \Delta f) \approx \frac{J_0^2(2\pi f_m \Delta t)}{1 + (2\pi \Delta f \tau_{av})^2}, \tag{17}$$

where $J_0(x)$ is the Bessel function of the first kind and zeroth order (which has a slightly different form than (5)). More details are given in the Schwartz textbook, Chapter 2.5. There are two major components here:

(1) τ_{av} is the stochastic version of the delay spread, which is the standard deviation of the path delays τ_k. The formal definition of the **coherence bandwidth** is the value of Δf for which $\rho_a(0, \Delta f) = 0.5$, below which correlation is low. Since $J_0(0) = 1$, the coherence bandwidth is

$$\Delta f = \frac{1}{2\pi \tau_{av}}. \tag{18}$$

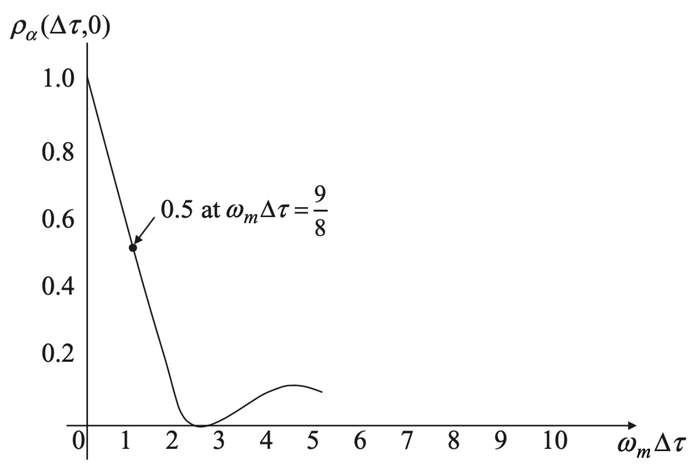
If $B > \Delta f$, we will experience frequency-selective fading, also known as **time dispersion**. In terms of the symbol duration $T_s \approx 1/B$, this gives the following condition:

$$T_s < 2\pi\tau_{av} \approx 6\tau_{av}. \tag{19}$$

We can consider three cases based on the relationship between B and Δf:

- If $B < \Delta f$, we experience flat fading.
- If $B > \Delta f$, ISI will occur.
- If $B \gg \Delta f$, distinct signal echoes may appear, each corresponding to different multipath rays, making them individually resolvable. This phenomenon can be exploited to improve signal detectability in high-bandwidth signals (this approach is taken in CDMA).

(2) f_m is the maximum Doppler shift. The formal definition of the **coherence time** is the time difference Δt for which $\rho_a(\Delta t, 0) = 0.5$. Since $\rho_a(\Delta t, 0) = J_0^2(2\pi f_m \Delta t)$, based on the behavior of $J_0(x)$, we can determine that this happens for $2\pi f_m \Delta t = 9/8$, as shown in the plot below.



Thus, the coherence time is

$$\Delta t = \frac{9}{16\pi f_m} \approx \frac{0.18}{f_m}. \tag{20}$$

If the signal symbol interval $T_s > \Delta t$, distortion within duration of a block occurs. This is known as time-selective fading, or **frequency dispersion**.