Chapter 4

Basics on the Theory of Fading Channels and Diversity

Vasileios M. Kapinas, Georgia D. Ntouni, and George K. Karagiannidis

Contents

Abbreviations

2D	two-dimensional
AF	amount of fading
AM	amplitude modulation
AWGN	additive white Gaussian noise
BPSK	binary phase-shift keying
CDMA	code division multiple access
CIR	channel impulse response
CSI	channel state information
DoF	degree of freedom
DS-SS	direct-sequence spread spectrum
FDM	frequency division multiplexing
FH-SS	frequency-hopping spread spectrum
i.i.d.	independent and identically distributed
IQ	in-phase/quadrature
ISI	intersymbol interference
LOS	line-of-sight
LTI	linear time-invariant
LTV	linear time-variant
MIMO	multiple-input multiple-output
MISO	multiple-input single-output
MLD	maximum likelihood detection
MLSE	maximum likelihood sequence estimation
MRC	maximum ratio combining

MUSA	multiple unit steerable antenna
OFDM	orthogonal frequency division multiplexing
PDF	probability density function
PRS	pseudo-random sequence
PSD	power spectral density
QAM	quadrature amplitude modulation
QoS	quality-of-service
QPSK	quadrature phase-shift keying
QSFF	quasi-static flat fading
RF	radio frequency
RV	random variable
Rx	receiver
RxD	receive diversity
SAR	synthetic aperture radar
SER	symbol error rate
SIMO	single-input multiple-output
SISO	single-input single-output
SNR	signal-to-noise ratio
Tx	transmitter
TxD	transmit diversity
WSS	wide-sense stationary

In modern communication systems, the sufficient knowledge of the channel behavior and its induced impairments on the transmitted signal are vital for the sophisticated design of advanced mitigation techniques. Fading is a complicated phenomenon that can severely affect the signal propagation in fixed and mobile wireless communication systems. However, its stochastic and highly varying nature makes it very difficult to model the associated channels precisely. Research efforts over the years have led to various statistical models for fading channels, which depend on the radio propagation environment and the communication scenario under study.

The negative impact of fading on signal transmission can be summarized in two major effects: the distortion of the signal due to intersymbol interference (ISI) and the signal-to-noise ratio (SNR) penalty in the error performance compared to the additive white Gaussian noise (AWGN) channel. These problems need to be efficiently tackled in order to guarantee robust communication with high availability. Diversity schemes can improve the transmission reliability by proper utilization of multiple communication channels with different characteristics. Interestingly, diversity has been proved to be one of the most common and efficient techniques for combating the detrimental effects of fading, interference, and error bursts for over a century now in the history of wireless systems.

The main purpose of this chapter is to introduce the fundamental concepts of fading and diversity, mainly in a qualitative way. A more detailed presentation of these subjects can be found in many textbooks, articles and dissertations in the

literature, such as [1–10] and references therein. This chapter also provides a detailed historical investigation on the roots of the fading phenomenon and the evolution of the various diversity techniques since the beginning of the wireless communications era. To this end, profound evidence of critical works, either not known to the research community or even uncited in the context of wireless communications, are given. Finally, more than 15 different pure concepts of the diversity are recorded here for the reader's ease, serving as a high-level taxonomy of the versatile diversity techniques that can be employed, either alone or in combination, by wireless communication systems to improve their error performance.

The organization of the chapter is as follows. The first half provides the basic theoretical tools for the analysis of fading, including its nature and types as well as the most commonly used statistical models. The impact of fading on signal transmission is also analyzed, focusing on time and frequency selectivity. In the second half, the chapter elaborates on fading mitigation and continues with a complete high-level classification of the diversity techniques proposed in the entire period of the twentieth century, which exploit various dimensions or degrees of freedom (DoF)s, such as space, frequency, time, and polarization, among others.

Throughout this chapter, various notations are used. Other than those described where read for the first time, some common math symbols adopted here include: (1) $j = \sqrt{-1}$ (imaginary unit), (2) $\mathrm{Re}\{z\}$, $\mathrm{Im}\{z\}$ (real and imaginary parts of complex number z), (3) $|z|$, $\angle z$ (amplitude, $|z| = a$, and phase, $\angle z = \theta$, of complex $z = a\exp(j\theta)$), (4) $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Z}$ (sets of complex, real, and integer numbers), (5) $\mathbb{N}$, $\mathbb{N}_0$ (sets of natural and non-negative integer numbers), (6) $\mathbb{R}^+$, $\mathbb{R}_0^+$ (sets of positive real and non-negative real numbers), (7) $\mathrm{E}\{\cdot\}$ (statistical expectation operator), (8) $\mu_z = \mathrm{E}\{z\}$, $\mathrm{var}\{z\} = \mathrm{E}\{|z - \mu_z|^2\}$ (mean and variance of RV $z \in \mathbb{C}$), (9) $\forall$ (for each), and (10) $x \mapsto T(x)$ (mapping from x to $T(x)$, also defining the mapping rule T).

4.1 Nature and Types of Fading

Consider a communication channel between a source S and a destination D, where the former intends to deliver an information message to the latter through the available wireless medium. Typically, not a single radio propagation path will exist between S and D due to the presence of various objects and structures surrounding either or both. These obstacles provide alternative paths for the electromagnetic wave propagation through different mechanisms that can generally be attributed to reflection, diffraction and scattering. As a consequence, radio waves travel along diverse paths and arrive at D from several directions with various delays and after experiencing different attenuations. Thus, the *multipath* nature of the medium introduces time spread in the information-bearing signal transmitted through the radio channel. In addition, *time variations* in the structure of the medium or existence of relative motion between S and D induce extra phase shift on the multipath components due to the Doppler effect.

Taking into account the characteristics of the time-variant multipath channel, the received signal $r_\ell(t)$ at D can be considered as a superposition of a large number of phasors,* let us say L_p, with time-varying amplitudes $a_l(t)$ and phases $\theta_l(t)$ that appear to be unpredictable to the user(s) of the channel

$$r_\ell(t) = \sum_{l=1}^{L_p} a_l(t) \exp\left[j\theta_l(t)\right] \tag{4.1}$$

where:

$\theta_l(t) = -2\pi f_c \tau_l(t)$

$\tau_l(t)$ is the related to the lth path propagation delay

f_c is the carrier frequency

For the sake of clarity, noiseless transmission of an unmodulated carrier is initially considered, while the complete version of the received signal, $r_\ell(t)$, will be treated in Section 4.3.2. In addition, the model in Equation 4.1 describes a discrete multipath channel, whereas in the case of a continuum of multipath components (i.e., *diffused multipath*), the sum has to be replaced by an integral.

4.1.1 Small and Large Scale Fading

The multipath channel model described earlier implies that the received signal $r_\ell(t)$ may be viewed as a complex-valued stochastic process with random amplitude and phase in both time and space domain, the latter defined by the set of all possible locations of D. In a more intuitive approach, the associated multipath components may add constructively or destructively at different time realizations, so that the composite signal received by D may experience distortion, strength fluctuation or both. This is the result of *signal fading* caused by the channel impulse response (CIR), that can be classified into major categories based on the nature of the physical medium, the distance between S and D, and the transmitted symbol rate.†

From the definition of fading, it is clear that when a received signal experiences fading during transmission, both its envelope (i.e., amplitude) and phase fluctuate over time. However, in most communication system scenarios, the receiver (Rx) is able to perform *coherent detection*, that is to reconstruct the carrier with perfect knowledge of the phase and frequency, and proceed further with complex conjugate demodulation. Therefore, it can generally be assumed that the phase variation

* The subscript ℓ in $r_\ell(t)$ indicates the equivalent lowpass signal of the real bandpass signal $r_b(t)$, also called the complex envelope of $r_b(t)$. For more details, see Section 4.3.

† We generally assume, excluding the spread spectrum systems, that the transmitted symbol rate ρ_s is approximately equal to the inverse of the signal bandwidth W, namely $\rho_s \approx 1/W$.

due to fading does not affect the error performance of the system, since the channel phase can be perfectly tracked at the Rx. It follows that performance analysis of digital communication systems mainly requires the knowledge of the fading envelope statistics, such as those given in Section 4.1.2, that widely define the type or distribution of the fading process.

Simply put, the rapid fluctuations of the instantaneous received signal strength over small travel distances or short time intervals, caused by the time-variant multipath channel,* result in the so-called *envelope fading* or more officially termed *small-scale fading* effects. This type of fading is relatively fast and is therefore responsible for the short-term signal variations. It is assumed to be a wide-sense stationary random process in time.† Depending on the nature of the radio propagation environment, there are different models describing the statistical behavior of the multipath fading envelope, the most common being the Rayleigh (1880), Nakagami-*q* or Hoyt (1947), Nakagami-*n* or Rice (1948), and Nakagami-*m* (1960). In the context of mobile communications, small-scale fading can result in signal power variations of up to 30–40 dB for a mobile movement of just a fraction of wavelength, while higher speeds can cause the mobile to pass through several fades in a small period of time.

However, as the Rx moves away from the transmitter (Tx) over much longer distances, the average signal power will gradually decrease as a result of the *large-scale fading* effects. Definitely, the most typical representative of channel impairments that fall within this category is the free-space *path loss*, a factor involved in the well-known Friis equation, for the case where there is a clear and unobstructed line-of-sight (LOS) path between the Tx and Rx. Nevertheless, this is a rather simplistic model since, in a real mobile radio channel, propagation is generally neither free space nor LOS. To this end, generalized measurement-based path loss models exist in the literature, where the path loss exponent can take values up to 4, as opposed to 2 for the free-space model. In addition, several empirical models have been obtained by curve fitting experimental data, two of the most popular being the Okumura-Hata (1968, 1980) and Lee (1985). Finally, in typical mobile radio channels, the mean signal level (or *local mean*) experiences slow variations over distances of several tens of wavelengths due to the presence of large buildings, hilly terrain and foliage. This phenomenon is known as shadow fading or *shadowing* and is usually modeled as a multiplicative and, generally, slowly time-varying random process. Experimental observations have confirmed that the shadow fades follow a Lognormal distribution for various outdoor and indoor environments (see, for instance, [11] and references therein).

* Even the spatial variations in the received signal due to the mobility of D can be observed by the latter as temporal variations while moving through the multipath field.

† It is recalled that a wide-sense stationary random process has the property that its mean value is constant and its autocorrelation function is invariant to any time shift.

4.1.2 Statistics and Modeling of Fading

It has already been discussed that, in a multipath fading channel, delayed versions of the transmitted signal can be received by reflection, diffraction and/or scattering, where each component corresponds to an appropriate channel *fading coefficient*. As will be seen in Section 4.3, fading has a multiplicative effect on the transmitted signal (see for instance Equation 4.53), while, for narrowband systems, the faded signal can be simply modeled by the product of a single fading coefficient and the complex envelope of the transmitted signal (see Equation 4.62). Under these assumptions, the envelope and phase of the noiseless received signal are actually the envelope and phase of the linearly modulated transmitted symbol by the fading coefficient. However, it has been proved that knowing the statistics of just the fading envelope is sufficient for the performance analysis of wireless communication systems [3, Sec. 2.1.1].* Therefore, in the sequel of this section, the envelope statistics of the most popular fading channels will be outlined. Note that, over the years, several models have been adopted to describe the statistical behavior of the fading envelope, each with different flexibility and complexity determined by the nature of the radio propagation environment.

4.1.2.1 Rayleigh Fading

Rayleigh fading is the most commonly used statistical model in wireless communications. Interestingly, Lord Rayleigh (John William Strutt) derived and applied this distribution to the theory of sound [12] long before the understanding of multipath signal reception. In a Rayleigh channel there exists no LOS component and the fading coefficient, h, is modeled as a zero-mean circularly-symmetric complex Gaussian random variable (RV) with variance $\mathrm{var}\{h\} = \mathrm{E}\{|h|^2\}$, that is, $h \sim CN(0, \mathrm{E}\{|h|^2\})$.† The probability density function (PDF) of the fading envelope $|h| = a$ equals

$$f_a(a) = \frac{2a}{\Omega} \exp\left(-\frac{a^2}{\Omega}\right), \quad a \geq 0 \tag{4.2}$$

* Recall also the discussion held in Section 4.1.1 about the coherent detection capability of communication systems.

† The symbol C in $CN(\cdot,\cdot)$ denotes that random samples (here snapshots of the fading channel) are both circularly-symmetric and complex. Additionally, it is known that circularly-symmetric complex jointly-Gaussian random vectors are completely determined by their covariance matrix [13], which for the (complex) scalar case reduces to the variance.

where $\Omega = \mathrm{E}\{a^2\}$ is the *average fading power*. It shall be noted here that, if needed, the phase $\angle h = \theta$, in this case, is typically following a uniform distribution in $[0, 2\pi)$, that is, $\theta \sim \mathcal{U}(0, 2\pi)$ and is considered to be statistically independent of a.

Assuming now that E_s is the symbol energy and N_0 the variance of the AWGN, the instantaneous SNR per symbol, $\gamma = a^2 \left(E_s/N_0\right)$, is exponentially distributed according to

$$f_\gamma(\gamma) = \frac{1}{\mathrm{E}\{\gamma\}} \exp\left(-\frac{\gamma}{\mathrm{E}\{\gamma\}}\right), \quad \gamma \geq 0 \tag{4.3}$$

In order to measure the severity of the experienced fading, Charash [14] introduced a metric called *amount of fading* (AF), which, according to [3, Sec. 1.1.4], can be used as an alternative performance criterion in the more general context of systems with arbitrary combining techniques and channel statistics. The AF metric is quantitatively expressed by

$$\mathrm{AF} = \frac{\mathrm{var}\{\gamma\}}{\mathrm{E}\{\gamma\}^2} = \frac{\mathrm{E}\{\gamma^2\} - \mathrm{E}\{\gamma\}^2}{\mathrm{E}\{\gamma\}^2} \tag{4.4}$$

It is seen from Equation 4.4 that the calculation of AF requires the computation of the first two moments (i.e., mean and variance) of γ, which for the Rayleigh fading case can be determined by [3, eq. (2.9)]

$$\mathrm{E}\{\gamma^k\} = \Gamma(1+k)\mathrm{E}\{\gamma\}^k, \quad k \in \mathbb{N} \tag{4.5}$$

where $\Gamma(z) = \int_0^\infty t^{z-1} \exp(-t)\, dt$, $z \in \mathbb{R}^+$ stands for the Euler gamma function. Hence, $\mathrm{AF}_{\mathrm{Rayleigh}} = 1$.

4.1.2.2 Rician (or Nakagami-n) Fading

If there exists a strong LOS component along with many random weaker ones, the channel fading coefficient has a non-zero mean value. In this case, the fading envelope follows the Rice distribution that was first derived by Rice [15]. Its basic parameter K, called *Rician factor*, denotes the average power ratio of the LOS component to all the other multipath components, where increasing K implies decreasing fading severity conditions. In the telecommunications literature, the Rice distribution is also known as Nakagami-n distribution with PDF given by [16, eq. (50)]

$$f_a(a) = \frac{2(1+n^2)\exp(-n^2)a}{\Omega} \exp\left(-\frac{(1+n^2)a^2}{\Omega}\right) I_0\left(2na\sqrt{\frac{1+n^2}{\Omega}}\right), \quad a \geq 0 \tag{4.6}$$

where $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind defined in [17, eq. (9.6.3)] and $n \geq 0$ is the Nakagami-n fading parameter which is related to

the Rician factor by $K = n^2$. Interestingly, for $n = K = 0$, Equation 4.6 reduces to the Rayleigh PDF in Equation 4.2, while for $n = K \to \infty$ there exists no fading at all but the LOS component.

In Rician fading, the instantaneous SNR per symbol follows noncentral chi-square distribution

$$f_\gamma(\gamma) = \frac{(1+n^2)\exp\left(-n^2\right)}{\mathrm{E}\{\gamma\}} \exp\left(-\frac{(1+n^2)\gamma}{\mathrm{E}\{\gamma\}}\right) I_0\left(2n\sqrt{\frac{(1+n^2)\gamma}{\mathrm{E}\{\gamma\}}}\right), \quad \gamma \geq 0 \quad (4.7)$$

while the moments of γ are given by [3, eq. (2.18)]

$$\mathrm{E}\{\gamma^k\} = \frac{\Gamma(1+k)}{(1+n^2)^k} M(-k, 1, -n^2)\mathrm{E}\{\gamma\}^k, \quad k \in \mathbb{N} \quad (4.8)$$

where $M(\cdot,\cdot,\cdot)$ stands for the Kummer's confluent hypergeometric function [17, eq. (13.1.2)].

Therefore, with the aid of Equation 4.8, the AF in Rician fading can be computed from Equation 4.4 as

$$\mathrm{AF}_{\mathrm{Rician}} = \frac{1+2n^2}{(1+n^2)^2} \quad (4.9)$$

From Equation 4.9, it can be validated that as $n \to \infty$ then $\mathrm{AF}_{\mathrm{Rician}} \to 0$, which corresponds to no-fading channel conditions.

4.1.2.3 *Nakagami*-m *Fading*

A more general fading distribution that can be well-fitted to experimental data is the Nakagami-*m* distribution suggested in 1960 by Nakagami. In this case, the PDF of the fading envelope is given by [16]

$$f_a(a) = \frac{2m^m a^{2m-1}}{\Omega^m \Gamma(m)} \exp\left(-\frac{ma^2}{\Omega}\right), \quad a \geq 0 \quad (4.10)$$

where $m \geq 0.5$ is the Nakagami-*m* fading parameter. Nakagami-*m* reduces to Rayleigh when $m = 1$, approximates Rician fading when $m = \left[(K+1)^2/(2K+1)\right]$ and results to no fading when $m \to \infty$. The PDF of the instantaneous SNR per symbol is a Gamma distribution given by

$$f_\gamma(\gamma) = \frac{m^m \gamma^{m-1}}{\mathrm{E}\{\gamma\}^m \Gamma(m)} \exp\left(-\frac{m\gamma}{\mathrm{E}\{\gamma\}}\right), \quad \gamma \geq 0 \quad (4.11)$$

The moments of γ are given by [3, eq. (2.23)]

$$\mathrm{E}\{\gamma^k\} = \frac{\Gamma(m+k)}{m^k\Gamma(m)}\mathrm{E}\{\gamma\}^k, \quad k \in \mathbb{N} \tag{4.12}$$

Again, with the aid of Equation 4.12, the AF in Nakagami-*m* fading can be computed from Equation 4.4 as

$$\mathrm{AF}_{\mathrm{Nakagami}} = \frac{1}{m} \tag{4.13}$$

Given that $\mathrm{AF} \in [0,2]$, Nakagami-*m* can model worse fading conditions than Rayleigh. Particularly, the channel with the most severe fading conditions, corresponding to $\mathrm{AF} = 2$, is referred to as one-sided Gaussian fading and occurs when either $\mathrm{var}\{\mathrm{Re}\{h\}\} = 0$ or $\mathrm{var}\{\mathrm{Im}\{h\}\} = 0$.

4.1.2.4 Hoyt (or Nakagami-q) Fading

The Hoyt model was originally proposed by Hoyt [18] and was further investigated by Nakagami (1960) in his effort to approximate Nakagami-*m* for certain values of *m*. Apart from cellular mobile radio systems, this type of fading channel is common in satellite links affected by strong ionospheric scintillation. The PDF of the fading envelope is given by [16, eq. (52)]

$$f_a(a) = \frac{(1+q^2)a}{q\Omega}\exp\left(-\frac{(1+q^2)^2a^2}{4q^2\Omega}\right)I_0\left(\frac{(1-q^4)a^2}{4q^2\Omega}\right), \quad a \geq 0 \tag{4.14}$$

where $0 \leq q \leq 1$ is the Nakagami-*q* fading parameter. This distribution can be approximated by the Nakagami-*m* after substituting $m = (1+q^2)^2/\left[2(1+2q^4)\right]$.

Regarding the instantaneous SNR per symbol, its PDF reads

$$f_\gamma(\gamma) = \frac{1+q^2}{2q\mathrm{E}\{\gamma\}}\exp\left(-\frac{(1+q^2)^2\gamma}{4q^2\mathrm{E}\{\gamma\}}\right)I_0\left(\frac{(1-q^4)\gamma}{4q^2\mathrm{E}\{\gamma\}}\right), \quad \gamma \geq 0 \tag{4.15}$$

In this case, the moments of γ are given by [3, eq. (2.13)]

$$\mathrm{E}\{\gamma^k\} = \Gamma(1+k)F\left(-\frac{k-1}{2}, -\frac{k}{2}; 1; \left(\frac{1-q^2}{1+q^2}\right)^2\right)\mathrm{E}\{\gamma\}^k, \quad k \in \mathbb{N} \tag{4.16}$$

where $F(\cdot,\cdot;\cdot;\cdot)$ is the Gauss hypergeometric function [17, eq. (15.1.1)], [19, eq. (9.100)].

Finally, the AF in Hoyt fading can be computed with the aid of Equations 4.4 and 4.16 as

$$\mathrm{AF}_{\mathrm{Hyot}} = \frac{2(1+q^4)}{(1+q^2)^2} \tag{4.17}$$

taking values within the interval [1,2].

4.1.2.5 Generalized Gamma (or α − μ) Fading

The generalized Gamma distribution was originally proposed by Stacy [20] as a purely mathematical model in which some statistical properties of a generalized version of the Gamma distribution were investigated. In this sense, it was connected neither with any specific application nor with any physical modeling of any given phenomenon. Generalized Gamma was recently revisited by Yacoub (2007), under the name $\alpha - \mu$ distribution, with its parameters being directly associated with the physical properties of the propagation medium of a fading channel. The generalized Gamma or $\alpha - \mu$ distribution includes several other distributions as special cases, such as the Gamma, Nakagami-*m*, Exponential, Weibull, one-sided Gaussian, and Rayleigh [21], while fading conditions for values of $m < 0.5$ can also be modeled. The PDF of the fading envelope is [22, eq. (1)]

$$f_a(a) = \frac{\beta m^m a^{m\beta-1}}{\Omega_p^m \Gamma(m)} \exp\left(-\frac{m a^\beta}{\Omega_p}\right), \quad a \geq 0 \tag{4.18}$$

where:

$\beta > 0$ is a parameter that yields the best fit to empirical measurements
$m > 0$ is the fading parameter
Ω_p is a power-scaling parameter given by $\Omega_p = \mathrm{E}\{a^\beta\}$

Denoting by $\tau = \left[\Gamma(m)/\Gamma(m + 2/\beta)\right]$, the PDF of the instantaneous SNR per symbol is given by [23, eq. (1)]

$$f_\gamma(\gamma) = \frac{\beta \gamma^{\frac{m\beta}{2}-1}}{2\Gamma(m)(\tau \mathrm{E}\{\gamma\})^{\frac{m\beta}{2}}} \exp\left[-\left(\frac{\gamma}{\tau \mathrm{E}\{\gamma\}}\right)^{\frac{\beta}{2}}\right], \quad \gamma \geq 0 \tag{4.19}$$

The moments of γ in this case are given by [23, eq. (3)]

$$\mathrm{E}\{\gamma^k\} = \frac{\Gamma[m + (2k/\beta)]}{\Gamma(m)} (\tau \mathrm{E}\{\gamma\})^k, \quad k \in \mathbb{N} \tag{4.20}$$

while the AF in generalized Gamma fading can be calculated from Equations 4.4 and 4.20 as

$$\mathrm{AF}_{\text{gen. Gamma}} = \frac{\Gamma[m + (4/\beta)]\Gamma(m)}{\Gamma^2[m + (2/\beta)]} - 1 \tag{4.21}$$

As mentioned earlier, generalized Gamma can describe a great variety of small-scale fading and shadowing conditions by adopting different values for the two parameters β and *m*. Indicatively, for $\beta = 2$ and $m = 1$ the Rayleigh fading channel model is obtained, while the Nakagami-*m* can be derived for $\beta = 2$ and $m > 0.5$.

4.1.2.6 Lognormal Shadowing

In the case of large-scale fading and assuming that the multipath effects are somehow eliminated, shadowing is the only phenomenon affecting the performance of the communication system. As mentioned in Section 4.1.1, empirical measurements have confirmed that the shadow fades can be modeled by a Lognormal distribution. In this case, the PDF of the instantaneous SNR per symbol, γ, is given by [3, eq. (2.53)]

$$f_\gamma(\gamma) = \frac{\xi}{\sigma\sqrt{2\pi\gamma}} \exp\left(-\frac{(10\log\gamma - \mu)^2}{2\sigma^2}\right) \tag{4.22}$$

where $\xi = (10/\ln 10)$ and μ (in dB), σ (in dB) are the mean and standard deviation of RV $10\log\gamma$, respectively.
The moments of γ in this case are given by [3, eq. (2.55)]

$$\mathrm{E}\{\gamma^k\} = \exp\left(\frac{k\mu}{\xi} + \frac{k^2\sigma^2}{2\xi^2}\right), \quad k \in \mathbb{N} \tag{4.23}$$

while the AF yields the expression

$$\mathrm{AF}_{\mathrm{Lognormal}} = \exp\left(\frac{\sigma^2}{\xi^2}\right) - 1 \tag{4.24}$$

In practical situations, the AF in Lognormal shadowing is bounded by a number that exceeds the maximal AF exhibited by the multipath PDF studied in the previous sections by several orders of magnitude [3, Sec. 2.2.2].

4.1.2.7 Generalized-K Composite Fading/Shadowing

Shadowing results in random rather than deterministic average received power levels, while in many communication systems it has to be encountered along with small-scale fading. The combined effect of small and large-scale fading can be modeled by various distributions, which are derived by the combination of one distribution for the fading envelope and another one for the local mean.

The Generalized-K distribution was introduced by Shankar [24] as a generalized fading/shadowing channel model, combining the Nakagami-m fading and Gamma shadowing models. It is general enough to include as special cases (or approximate well) other distributions. For this composite channel, the PDF of the fading envelope is given by [25, eq. (1)]

$$f_a(a) = \frac{4m^{\frac{k+m}{2}}}{\Gamma(m)\Gamma(k)\Omega_p^{\frac{k+m}{2}}} a^{k+m-1} K_{k-m}\left(2a\sqrt{\frac{m}{\Omega_p}}\right), \quad a \geq 0 \tag{4.25}$$

where k, m are the two shaping parameters, $\Omega_p = \mathrm{E}\{a^2\}/k$, and $K_{k-m}(\cdot)$ is the $(k-m)$th order modified Bessel function of the second kind [19, eq. (8.407.1)]. One can see that, for $m=1$, Equation 4.26 reduces to the K distribution, thus approaching the Rayleigh-Lognormal composite fading/shadowing channel model. In addition, as $k \to \infty$ it approximates Nakagami-m, while, for $m, k \to \infty$, the Generalized-K fading channel approaches AWGN (i.e., no-fading case).

The PDF of the instantaneous SNR per symbol is given by [25, eq. (2)]

$$f_\gamma(\gamma) = 2\left(\frac{km}{\mathrm{E}\{\gamma\}}\right)^{\frac{k+m}{2}} \frac{\gamma^{\frac{k+m-2}{2}}}{\Gamma(k)\Gamma(m)} K_{k-m}\left(2\sqrt{\frac{km}{\mathrm{E}\{\gamma\}}\gamma}\right), \quad \gamma \geq 0 \tag{4.26}$$

In addition, the moments of γ for the Generalized-K distribution are given by [25, eq. (5)]

$$\mathrm{E}\{\gamma^n\} = \frac{\Gamma(k+n)\Gamma(m+n)}{\Gamma(k)\Gamma(m)}\left(\frac{\mathrm{E}\{\gamma\}}{km}\right)^n, \quad n \in \mathbb{N} \tag{4.27}$$

Finally, the AF in Generalized-K fading can be directly computed from [24, eq. (8)]

$$\mathrm{AF}_{\text{gen.-K}} = \frac{1}{k} + \frac{1}{m} + \frac{1}{km} \tag{4.28}$$

where it is easy to see that (1) low values of k and m correspond to severe fading, (2) values of $k \to \infty$ render Equation 4.28 identical to 4.13, that is, shadowing vanishes giving rise to Nakagami-m fading only, and (3) values of $k, m \to \infty$ give $\mathrm{AF}_{\text{gen.-}K} = 0$, which corresponds to the ideal case of no fading at all or equivalently to the AWGN channel.

4.1.3 Historical Roots of Fading

The fluctuations in the received signal strength over time and distance were noticed very soon after the first transatlantic transmission, accomplished by Marconi in 1901. Long ago in 1902, Marconi drew attention to the remarkable difference in the strength of signals received by day and by night, and his observations have been repeated and recorded by many other observers [26]. It is very interesting to reproduce below some selected excerpts from the weekly published issues of the "Electrical World and Engineer" periodical of that period, which reveal that the signal intensity variation in long distance transmissions was very early reported by experts and not, though any kind of explanation for this paradox was unable to be provided by that time:

> Atmospheric conditions assuredly play an important part here, and it is a common experience of transatlantic travel to find the wireless not

> working satisfactorily, while at times remarkable results are achieved. At moderate range most of the difficulties seem to disappear.
>
> **—*Elec. World Eng.*, vol. XL, no. 20, Nov. 15, 1902; under the column "Long distance wireless telegraphy"**

> These long-distance feats are tantalizing in their uncertainty. Messages might be received on one day and the next day, for no apparent reason, might fail to have any accurate effect on the receiving apparatus.... It is thought that long distance experiments are useless for practical purposes.
>
> **—*Elec. World Eng.*, vol. XL, no. 20, Nov. 15, 1902; under the column "Telegraphy, telephony and signals"**

> In the wireless service on the transatlantic steamers there has been beautiful success in picking up passing neighbors at long range, but there have also been lugubrious failures in reaching the land stations at very moderate distances.
>
> **—*Elec. World Eng.*, vol. XL, no. 23, Dec. 6, 1902; under the column "Round the world by wireless"**

However, it seems that it was not before 1910s that first in-depth discussions on this phenomenon and possible explanations appear in the literature. In one of these very early works published in 1911, Taylor uses the term *freak communications* (or *freak signals*) to refer to variations in the range of communication (or signal strength) occasionally occurred by day, which are much more pronounced during the night hours [27]. His broader explanation attributes this phenomenon to "variations in the transmitting efficiency of the atmosphere" adding that "perturbances in this respect exhibit almost the same characteristics of irregularity as atmospheric impulses." In order to highlight the importance of this phenomenon, Taylor remarks that "the sudden veiling or obscuring of distant signals together with the equally sudden 'opening out' of signals observed whilst listening on the receiver at a wireless station during the prevalence of this phenomenon is very impressive." Nevertheless, Commander Loring had some doubts about the explanation given previously [28]:

> The prevalence, or otherwise, of atmospherics does not appear to bear any relation to the appearance of freak signals. At any rate, atmospherics have been reported by reliable observers to be least noticeable at periods of freak ranges.
>
> **—Commander Loring (1911); comment on [27]. Reproduced by permission of the Institution of Engineering & Technology.**

In the following years, systematic observations and recordings took place with regard to variations of the radio range or signal strength in several wireless links. Indicatively, Marriott [29] presented a record of the most interesting range variations observed at the Manhattan Beach station of the United Wireless Telegraph Company, located at Coney Island in New York City. His measurements indicated that there were more *freaks* in the month of August than in the winter months, namely the ranges showed a greater percentage variation in the summer than in the winter. Marchant (1915) summarized measurements made mainly between stations in Liverpool and Paris regarding fluctuations in the strength of wireless signals during sunset and night hours, taking into account the weather conditions over the period of one year. Among others, Marchant concluded that (1) the *sunset effect*, (i.e., the strengthening of the signal just after sunset) occurs about 3/4 hr after the actual time of sunset and varies with the weather conditions, whereas under rainy conditions it is much less marked and (2) the freak signals are relatively great during night and occur within the space of a few minutes, while greater increases in signal strength can be observed after the cessation of rain either at the transmitting or receiving station [26]. In the same context, Austin (1915) studied the seasonal variation in strength of signals sent from radio stations in the Philadelphia and Norfolk navy yards via observations made for over two years. His findings, among others, experimentally verified that winter signals in general are stronger than those of summer, especially when the transmission takes place overland, while, "contrary to the ideas previously held, there seems to be no very marked connection between rainfall and the transmission of the signals" [30].

From the aforementioned published works, it can be easily observed that during the first two decades of the wireless telegraphy era,* there was considerable confusion in understanding the actual effects of fading on signal transmission and even more difficulties in providing accurate explanations for its genesis. With this in mind, the terms *freak communications/signals/ranges* were actually related to the long ranges that could be covered by low-power shortwave radio transmitters at certain times and places, a phenomenon which is well-known today as *skywave* or *skip* propagation. However, in these early works, the so-called *freaks* were sometimes unintentionally confused with the large-scale fading effects.† Very interestingly, it was in 1913 that the radio pioneer (and engineer of the Federal Telegraph Co.) De Forest officially coined the term *fading* for the first time to describe the fact that during a station-to-station transmission from Los Angeles to San Francisco, being 350 miles apart, two waves differing only 5% in length (3260 and 3100 m, respectively) were received with great variations in their intensity over certain time intervals, sometimes observing even total extinction of the one or the other but very rarely of both simultaneously [34].‡ In order to give some

* Inventor and entrepreneur Marconi had actually established radio communication links spanning as far as the English Channel (la Manche) since the early 1890s [31].

† Note that, the first consolidated results on "shadowing" appear in the literature several years later, in [32,33].

‡ This first occurrence of the term "fading" in the literature was also reported by Burrows [35].

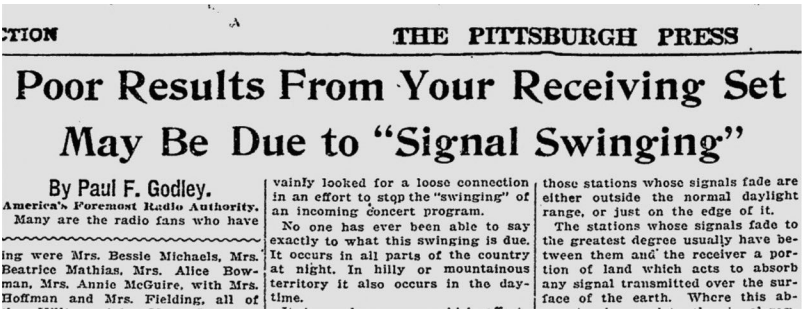
explanation for this paradox, De Forest adopted an existing speculative theory about the possible interference of the direct wave with the reflected ones from the upper conducting layer of the atmosphere, thus giving the birth to multipath fading reasoning. The next parts from this seminal work are very illuminative.

> The duration of this fading effect is often several hours after nightfall; then it suddenly vanishes and thereafter both waves have their normal intensity. This alteration of intensity is sometimes for one wave, and sometimes for the other, and rarely for both. Under such conditions there are acting at the receiving stations two trains of waves, which have travelled over paths of unequal lengths or which have travelled with unequal velocities. Consequently, there will be a phase displacement between them and interference at certain localities… it would account for the fact that the Marconi transatlantic stations can operate sometimes with a few kilowatts and sometimes require 125 to 600 kilowatts.
>
> **—De Forest. © 1913 IEEE. Reprinted, with permission, from [34].**

In 1916, Taylor and Blatterman published a very interesting paper related to nocturnal transmission experiments carried out by radio stations in University of North Dakota and Washington University in St. Louis as an attempt to test the interference theory of fading effects. In this work, the term *swinging* appears for the first time to denote slow variations in signal strength. Particularly, they stated that there seem to be two kinds of fluctuations in nocturnal overland transmission; a rapid fading and a slow swinging in signal strength [36]. In the discussion section of the same paper, some plausible explanations for these two channel oddities are also provided.

> The first (refers to *fading*) may be due to changes, in the nature of interference effects. These could be local at the sender or at the receiver, or they might be caused by rather sharp surfaces of discontinuity almost anywhere between the stations... The second (refers to *swinging*) or slower effect may be due to refracting masses of moving ionized air in the path of transmission, producing at times a lens-like concentration and at other times a dispersive effect.
>
> **—Taylor and Blatterman. © 1916 IEEE. Reprinted, with permission, from [36].**

In the coming years, the understanding of the cause and effects of fading or swinging on radio signal transmission proved itself to be one of the most elusive problems faced by both radio engineers and amateurs. As the radio art was



CTION THE PITTSBURGH PRESS

Poor Results From Your Receiving Set May Be Due to "Signal Swinging"

By Paul F. Godley.
America's Foremost Radio Authority.

Many are the radio fans who have vainly looked for a loose connection in an effort to stop the "swinging" of an incoming concert program.

No one has ever been able to say exactly to what this swinging is due. It occurs in all parts of the country at night. In hilly or mountainous territory it also occurs in the daytime.

The stations whose signals fade are either outside the normal daylight range, or just on the edge of it.

The stations whose signals fade to the greatest degree usually have between them and the receiver a portion of land which acts to absorb any signal transmitted over the surface of the earth. Where this ab-

ing were Mrs. Bessie Michaels, Mrs. Beatrice Mathias, Mrs. Alice Bowman, Mrs. Annie McGuire, with Mrs. Hoffman and Mrs. Fielding, all of

Figure 4.1 Part from a news article on *signal swinging*. *The Pittsburgh Press*, June 25, 1922.

progressively moving from spark telegraphy into continuous wave telegraphy and radio telephone broadcasting, any empirical or scientific contribution on the peculiarities of the wireless transmission medium and the possible ways to improve the message quality at the receiving end became matters of common interest. As a consequence, these issues occupied much space in the academic and public press, an example for the latter given in Figure 4.1 that depicts the headline and part of the body of a 1922 article on signal swinging, published in a major daily afternoon newspaper. During that period, another term, that is, *soaring*, appears in the literature to describe the radio signal variations in the same context. Particularly, Pickard [37] mentions that the terms fading, swinging, and soaring were commonly used interchangeably to refer to the large amplitude short period variations of radio signals.

In the meantime, radio amateurs were among the first to perform large-scale experiments on fading and long-distance shortwave communication even from the early 1910s [31], thus having already obtained valuable experience. However, due to the plethora of variables involved in radio transmission, any conclusions regarding fading could be considered safe only after being corroborated by a large number of observers. To this end, Whittemore and Kruse from the U.S. National Bureau of Standards suggested a giant initiative involving extensive experiments by radio amateurs. The so-called *Fading Tests* took place during 1920–1921 with 243 receiving stations from North America producing thousands of records [31]. The results, after being analyzed by the Bureau of Standards, were finally published in [38], mainly validating the long-standing belief that fading at short wavelengths is more serious than at long ones. However, historical investigations, carried out by Yeang (2004) revealed that "the claims associated with atmospheric conditions did not lead to a theory of fading. Nor was the conclusion very credible to the experiments at the bureau" [31].

4.2 Impact of Fading on Signal Transmission

4.2.1 Multipath Spread and Frequency Selectivity

It has been mentioned that multipath causes small-scale fading due to interference of two or more versions of the transmitted signal that arrive at the Rx at slightly different times. The detrimental effects of multipath can be better realized through the *delay power spectrum* or *multipath intensity profile* of the channel, which gives the average power output of the channel as a function of the time delay τ. Particularly, the range of τ values over which the delay power spectrum is essentially nonzero is defined as the *multipath spread* T_m of the channel. Therefore, if the signaling interval T is smaller than T_m, then the channel introduces ISI since the replicas of the neighboring transmitted signals interfere with the one received over the current interval. Equivalently, the Fourier transform of the multipath intensity profile defines the *spaced-frequency correlation function*, which provides a measure of the correlation between two samples of the channel response taken at different frequencies. A critical parameter of this function is the *coherence bandwidth* B_c of the channel that indicates the frequency range over which the fading process is highly correlated.*

Therefore, if the signal bandwidth W is greater than B_c, then the spectral components of the transmitted signal are affected by different amplitude gains and phase shifts across the band. Under these conditions, that is, $W > B_c$ or equivalently $T < T_m$, the channel is said to be *frequency-selective* and the signal is severely distorted. However, if the signaling interval is selected to be greater than T_m, then the channel introduces a negligible amount of ISI. Likewise, if the bandwidth is smaller than B_c, all the frequency components of the signal undergo the same attenuation and phase shift in transmission through the channel. In this case, that is, $W < B_c$ or equivalently $T > T_m$, the fading is said to be *frequency-nonselective* or *flat* and the multipath components of the channel are not resolvable, thus having a time-varying multiplicative effect on the transmitted signal.

4.2.2 Doppler Spread and Time Selectivity

The rapidity with which the CIR is changing with time is related to the *Doppler power spectrum* that gives the power profile as a function of the Doppler frequency λ. Particularly, the range of λ values over which the Doppler power spectrum is essentially nonzero is defined as the *Doppler spread* B_d of the channel.† Since the time variations in the CIR are evidenced as a Doppler broadening, it follows that a large value of B_d characterizes a rapidly changing channel. Equivalently, the inverse Fourier transform of the Doppler power spectrum defines the *spaced-time correlation*

* The coherence bandwidth of the channel is related to the multipath spread by $B_c \approx 1/T_m$.

† The Doppler spread of the channel is sometimes called *fading rate*.

function, which quantifies the correlation between two samples of the channel response taken at different time instants. A critical parameter of this function is the *coherence time* T_c of the channel that indicates the period of time over which the fading process is highly correlated. Since the reciprocal of B_d is a measure of the channel coherence time, the T_c parameter gives strong evidence of the rate at which the channel attenuation and phase shift change over time due to the nonstatic nature of the environment.

Therefore, a channel with a large Doppler spread compared to the signal bandwidth is also characterized by a small coherence time with respect to the signaling interval. In this case, that is, $W < B_d$ or equivalently $T > T_c$, the transmitted signal undergoes *time-selective* or *fast* fading. In the dual case, if T is selected to be smaller than the coherence time, that is, $W > B_d$ or equivalently $T < T_c$, the channel attenuation and phase shift can be considered fixed during at least one signaling interval, thus giving rise to *time-nonselective* or *slow* fading.

4.2.3 Characterization of Wireless Channels

The adverse effects of signal fading, if not mitigated, can significantly degrade the system's performance, resulting in nonreliable or even no communication during deep fades, a phenomenon called *outage*. Before referring to the versatile ways to tackle the problem of fading, it would be very useful to first classify the fading channels that arise in practice and highlight the different impact they have on signal transmission. Following a top–down approach, small-scale and large-scale fading are the main phenomena involved in the wireless channel propagation. Although they are both experienced by the user(s) of the channel as temporal and spatial fluctuations in the received power, the degradation mechanisms behind them are different, as explained in Section 4.1. From now on, we will focus on small-scale fading and refer to it simply as fading.

In Sections 4.2.1 and 4.2.2, it has been mentioned that fading channels can be characterized by their level of selectivity in frequency and time through the normalized parameters T_m/T and B_d/W, respectively. While the former indicates the dispersion degree of the signal in time due to multipath delays, the latter provides a measure of the signal dispersion in frequency due to Doppler shifts. Interestingly, a simple yet meaningful factor, which encompasses both spread parameters T_m, B_d to characterize a single fading channel, is the *spread factor* given by $T_m B_d \approx (B_c T_c)^{-1}$, where intuitively the $B_c T_c$ quantity reveals the overall coherence of the channel. The value of this factor is critical for the design of a digital communication system as it reflects the ease at which channel estimation and coherent demodulation can be performed, thus affecting the detection performance especially under adaptive transmission. By definition, a channel is said to be *underspread* if $T_m B_d < 1$ and *overspread* if $T_m B_d > 1$. This implies that, in the first case, the CIR can be easily measured, while in the second one, the same process becomes extremely difficult and unreliable, if not impossible.

Taking all these into account, a fading channel can be generally characterized as:

1. Slow flat (or *nondispersive*)
   a. If and only if $T_m < T < T_c$ and $B_d < W < B_c$
   b. If $W \approx 1/T$, then $T_m B_d < 1 \rightarrow$ underspread channel
2. Slow frequency-selective (or *time-dispersive*)
   a. If and only if $T_m > T < T_c$ and $B_d < W > B_c$
   b. If $\rho_s \downarrow$ and B_d is small enough $\rightarrow$ more flat channel
3. Time-selective flat (or *frequency-dispersive*)
   a. If and only if $T_m < T > T_c$ and $B_d > W < B_c$
   b. If $\rho_s \uparrow$ and T_m is small enough $\rightarrow$ more slow channel
4. Doubly-selective (or *doubly-dispersive*)
   a. If and only if $T_m > T > T_c$ and $B_d > W > B_c$
   b. If $W \approx 1/T$, then $T_m B_d > 1 \rightarrow$ overspread channel

The necessary and sufficient conditions per channel case illustrated earlier give evidence that, from a system-level point of view, the transmission characteristics can be chosen in such a way so that the dispersive nature of the channel can be moderated. To this end, in a frequency-selective fading channel, if the symbol rate ρ_s is decreased, the channel becomes more flat.* Correspondingly, in a time-selective fading channel, if ρ_s is increased, the channel varies more slowly. However, the design of more realistic fading mitigation techniques comes up as a requirement in radio applications and their consideration should definitely take into account the different degradation mechanisms involved in each channel type.

For instance, in the presence of slow flat fading with no LOS path, the error probability is asymptotically inversely proportional to the SNR. For a given quality of service (QoS) threshold, this implies a significant loss in terms of SNR compared to the AWGN channel, where the error probability decreases exponentially with the SNR. For visualization purposes, the error performance degradation of binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) systems due to the presence of nondispersive fading is shown in Figures 4.2 and 4.3. All systems are single-input single-output (SISO) and employ maximum likelihood detection (MLD) at the Rx. In many cases, the comparison with respect to the ideal AWGN (i.e., no-fading) channel reveals huge SNR penalties, which for typical target error rates span several tens of dB. The standard methodology and related techniques used to increase the received SNR in wireless communication systems suffering from severe fading are provided in Sections 4.4.2 and 4.4.4.

* In this direction, *narrowband* signal transmission, that is, $W << f_c$, is almost synonymous with flat fading conditions.

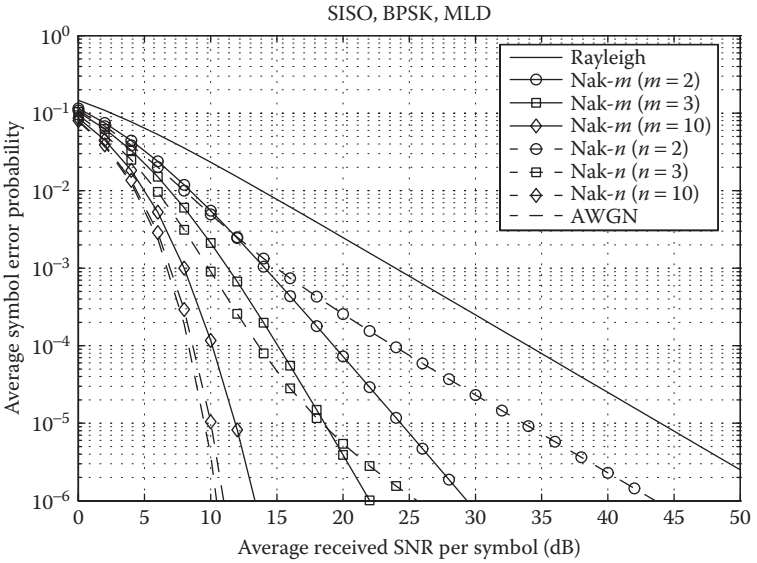


Figure 4.2 Error performance of a BPSK communication system operating over various slow flat fading channels (i.e., Nakagami-*m*, Nakagami-*n*) and comparison with the ideal AWGN case.

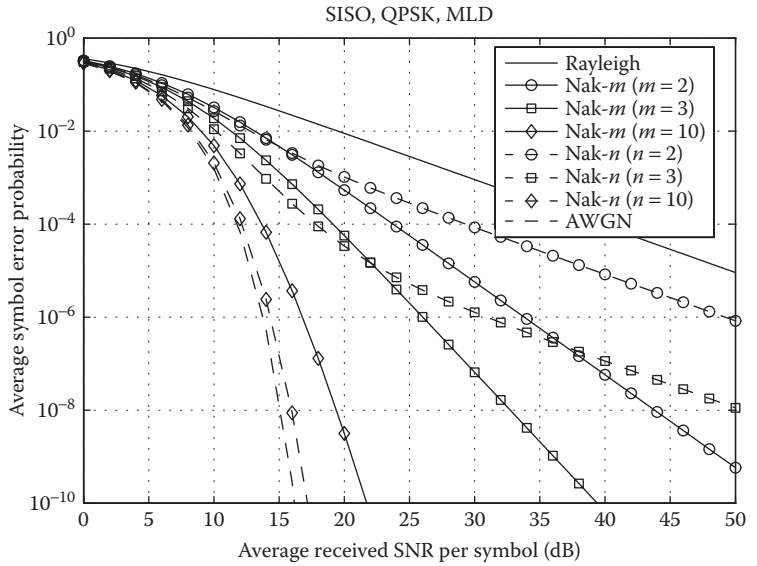


Figure 4.3 Error performance of a QPSK communication system operating over various slow flat fading channels (i.e., Nakagami-*m*, Nakagami-*n*) and comparison with the ideal AWGN case.

Even worse, in the case of transmission through a frequency-selective and/or fast fading channel, the system performance can exhibit an irreducible error-rate level, called *error floor*, due to the significant degradation induced by the ISI and/or Doppler spread.* In such a scenario, no amount of SNR can help achieve the desired level of performance, unless some forms of mitigation are first employed to reduce or even eliminate the signal distortion. In Section 4.4.1, a list of robust transmission techniques is given for the elimination of ISI and the associated performance error floor.

4.3 Formulation of Signal Transmission over Fading Channels

As our analysis does not depend on the carrier frequency f_c, it is mathematically convenient to make use of the *complex baseband representation* of bandpass signals and channels. To begin with, if $s_b(t)$ is the continuous-time linearly modulated narrowband bandpass signal to be transmitted over the wireless medium, then its general form can be written as [39, Sec. 2.5.2]

$$s_b(t) = \alpha(t)\cos(2\pi f_c t + \phi(t)) \tag{4.29}$$

with $\alpha(t)$ and $\phi(t)$ being the envelope and phase of $s_b(t)$, whereas its canonical form follows as

$$s_b(t) = s_I(t)\cos(2\pi f_c t) - s_Q(t)\sin(2\pi f_c t) \tag{4.30}$$

In Equation 4.30, $s_I(t) = \alpha(t)\cos\phi(t)$ and $s_Q(t) = \alpha(t)\sin\phi(t)$ denote the in-phase/quadrature (IQ) components of the real-valued bandpass signal $s_b(t)$, defining the equivalent lowpass (or baseband) signal $s_\ell(t)$ as

$$s_\ell(t) = s_I(t) + js_Q(t) = \alpha(t)\exp\big(j\phi(t)\big) \tag{4.31}$$

For M-ary digital modulation, both $s_I(t)$ and $s_Q(t)$ take values from predefined finite discrete sets with the additional restriction that eligible $(s_I(t), s_Q(t))$-tuples must belong to a two-dimensional (2D) constellation of size M that satisfies particular design criteria. Then, each constellation point is associated with a distinct $\log_2(M)$-bit length sequence according to a lookup table. In this sense, modulation of the binary information message can be envisaged either through $s_b(t)$ or its complex envelope $s_\ell(t)$ due to their interconnection through the relation

$$s_b(t) = \mathrm{Re}\{s_\ell(t)\exp(j2\pi f_c t)\} \tag{4.32}$$

* The varying Doppler shifts on different multipath components result in random frequency modulation.

The corresponding spectrum relation is derived by taking the Fourier transform of $s_b(t)$ in Equation 4.32

$$\begin{aligned} S_b(f) = \mathcal{F}_t\{s_b(t)\} &= \frac{1}{2}\int_{-\infty}^{\infty} s_\ell(t)\exp\left(j2\pi f_c t\right)\exp\left(-j2\pi ft\right)dt \\ &\quad + \frac{1}{2}\int_{-\infty}^{\infty} s_\ell^*(t)\exp\left(-j2\pi f_c t\right)\exp\left(-j2\pi ft\right)dt \\ &= \frac{1}{2}\left[S_\ell(f-f_c) + S_\ell^*(-f-f_c)\right] \end{aligned} \tag{4.33}$$

where $S_b(f)$ and $S_\ell(f)$ are the bandpass and lowpass signal spectra, respectively.

Let us now assume that $s_b(t)$ passes through a continuous-time linear time-invariant (LTI) bandpass filter with impulse response $g_b(t)$ that yields the following representation with respect to its complex envelope $g_\ell(t)$ [40]*

$$g_b(t) = 2\text{Re}\{g_\ell(t)\exp\left(j2\pi f_c t\right)\} \tag{4.34}$$

If $G_b(f)$ and $G_\ell(f)$ are the bandpass and equivalent lowpass frequency responses, respectively, then

$$G_b(f) = \mathcal{F}_t\{g_b(t)\} = G_\ell(f-f_c) + G_\ell^*(-f-f_c) \tag{4.35}$$

Thus, the output of the bandpass system in the time and frequency domains is given by

$$x_b(t) = s_b(t) * g_b(t) \Rightarrow X_b(f) = S_b(f)G_b(f) \tag{4.36}$$

By substituting Equations 4.33 and 4.35 into the right-hand side of Equation 4.36, we obtain

$$\begin{aligned} X_b(f) = \frac{1}{2}\Big[&S_\ell(f-f_c)G_\ell(f-f_c) + S_\ell^*(-f-f_c)G_\ell^*(-f-f_c) \\ &+ S_\ell(f-f_c)G_\ell^*(-f-f_c) + S_\ell^*(-f-f_c)G_\ell(f-f_c)\Big] \end{aligned} \tag{4.37}$$

However, for narrowband signal transmission, where the frequency content of $s_b(t)$ is concentrated in a narrow band around f_c, that is, $W << f_c$, it is easy to see that [1, Sec. 4.1.3]

$$S_\ell(f-f_c)G_\ell^*(-f-f_c) = 0 \quad \text{and} \quad S_\ell^*(-f-f_c)G_\ell(f-f_c) = 0 \tag{4.38}$$

* According to Tranter et al. [40, Sec. 4.2.1], the factor 2 in Equation 4.34 preserves the filter passband gain.

Therefore, by taking the inverse Fourier transform of $X_b(f)$ in Equation 4.37, it is straightforward to show that

$$x_b(t) = \mathcal{F}_f^{-1}\{X_b(f)\} = \mathrm{Re}\{\big(s_\ell(t) * g_\ell(t)\big)\exp\big(j2\pi f_c t\big)\} = \mathrm{Re}\{x_\ell(t)\exp\big(j2\pi f_c t\big)\} \quad (4.39)$$

which finally defines the input-output relation of the equivalent lowpass system in the time domain as

$$x_\ell(t) = s_\ell(t) * g_\ell(t) \quad (4.40)$$

4.3.1 Time-Variant Multipath Channel Model

Consider next that the bandpass signal $x_b(t)$ is transmitted through a time-variant multipath channel with L_p distinct paths. Normally, the number of propagation paths will be a function of time, that is, $t \mapsto L_p(t)$. However, the assumption $L_p(t) \approx \mathrm{E}\{L_p(t)\} \approx L_p, \forall t$, which implies a constant ensemble and time average number of paths over all signaling intervals, does not affect the generality of our analysis, while it simplifies the derivation of the final system model.

In this case, the faded signal $y_b(t)$ at the channel output consists of the superposition of L_p randomly attenuated and delayed copies of the original signal $x_b(t)$, mathematically formulated as

$$y_b(t) = \sum_{l=1}^{L_p} a_l(t) x_b[t - \tau_l(t)] \quad (4.41)$$

with $a_l(t)$ and $\tau_l(t)$ being the time-varying attenuation and propagation delay induced by the lth path, respectively. By substituting the right-hand side of Equation 4.39 into 4.41, we obtain successively

$$\begin{aligned} y_b(t) &= \sum_{l=1}^{L_p} a_l(t)\mathrm{Re}\{x_\ell(t - \tau_l(t))\exp[j2\pi f_c(t - \tau_l(t))]\} \\ &= \mathrm{Re}\{\sum_{l=1}^{L_p} a_l(t)\exp(j\theta_l(t))x_\ell(t - \tau_l(t))\exp(j2\pi f_c t)\} \end{aligned} \quad (4.42)$$

where $\theta_l(t) = -2\pi f_c \tau_l(t)$ is the phase shift associated with the $\tau_l(t)$ delay. By simple inspection, we get the complex envelope $y_\ell(t)$ of the faded signal expressed in terms of the baseband transmitted signal*

$$y_\ell(t) = \sum_{l=1}^{L_p} a_l(t)\exp\big(j\theta_l(t)\big)x_\ell(t - \tau_l(t)) \quad (4.43)$$

* Equation 4.43 reduces to 4.1 for an unmodulated carrier with no filtering, that is, $s_\ell(t) = 1$, $g_\ell(t) = \delta(t), \forall t$.

In order to proceed with the interpretation of Equation 4.43, it is necessary to invoke some elements from the theory of linear time-variant (LTV) filters. Particularly, if $x_\ell(t)$ is the input to an LTV filter with impulse response $h_\ell(t,\theta)$, where t and θ are the observation and impulse input time instants, respectively, then the input–output relation is defined by the general superposition (or convolution) integral [41, Sec. 6.2]

$$x_\ell(t) * h_\ell(t,\theta) \triangleq \int_{-\infty}^{\infty} h_\ell(t,\theta)x_\ell(\theta)d\theta = \int_{-\infty}^{\infty} h_\ell(t,t-\tau)x_\ell(t-\tau)d\tau \tag{4.44}$$

where the second integral form is simply derived from the first one by introducing the variable $\tau = t - \theta$ that physically corresponds to the elapsed time of the filter (or age of the input). For the case of an LTI system it holds that $h_\ell(t,\theta) = h_\ell(t+a,\theta+a), \forall a \in \mathbb{R}$. Hence, by setting $a = -\theta$ we get $h_\ell(t,\theta) = h_\ell(t-\theta,0) \triangleq h_\ell(t-\theta) = h_\ell(\tau)$, which reduces Equation 4.44 to the standard convolution integral $\int_{-\infty}^{\infty} h_\ell(\tau)x_\ell(t-\tau)d\tau$ [42, Sec. 1.3].

Although expressions in Equation 4.44 are general enough to handle both LTI and LTV systems, the impulse response $h_\ell(t,\theta)$ does not involve directly the parameter τ, which becomes a drawback when one wants to define the frequency response of the LTV system through the Fourier transform of $h_\ell(t,\theta)$. For this reason, Kailath [43] introduced the modified impulse response $h_\ell(\tau;t)$ of an LTV system, which satisfies the following relations with respect to $h_\ell(t,\theta)$ [43], [41, Sec. 6.2.1]

$$h_\ell(\tau;t) = h_\ell(t,t-\tau) \quad \text{and} \quad h_\ell(t,\theta) = h_\ell(t-\theta;t) \tag{4.45}$$

Based on these relations, the time-varying convolution integral of Equation 4.44 may be alternatively defined as

$$x_\ell(t) * h_\ell(\tau;t) \triangleq \int_{-\infty}^{\infty} h_\ell(\tau;t)x_\ell(t-\tau)d\tau = \int_{-\infty}^{\infty} h_\ell(t-\theta;t)x_\ell(\theta)d\theta \tag{4.46}$$

The LTV nature of the filter implies that its impulse response generally changes as a function of both the time t at which the response is observed and the time $t-\tau$ at which the impulse is applied. For an LTI filter the impulse response depends only upon the time difference $t-(t-\tau) = \tau$, regardless of when the impulse is applied or the response is observed, since by adopting the modified version of the impulse response for an LTI system, we take $h_\ell(\tau;t) \triangleq h_\ell(\tau) = h_\ell(t-\theta)$, which for a reference input time $\theta = 0$ gives also $h(t)$ [42, Sec. 1.3]. Thus, considering the faded signal $y_\ell(t)$ in Equation 4.43 as the output of the LTV filter in Equation 4.46, the impulse response of the multipath radio channel $h_\ell(\tau;t)$ may be expressed in terms of $a_l(t)$, $\theta_l(t)$, and $\tau_l(t)$ as [1, Sec. 1.3], [40, Sec. 14.4]*

$$h_\ell(\tau;t) = \sum_{l=1}^{L_p} h_l(t)\delta[\tau - \tau_l(t)] \tag{4.47}$$

* There is a typo in [40, eq. (14.21)]; the quantity $\delta[t-\tau_n(t)]$ should be changed to $\delta[\tau-\tau_n(t)]$.

where $h_l(t) = a_l(t)\exp\left[j\theta_l(t)\right]$ represents the generally time-varying complex coefficient of the lth path. Equation 4.47 is well known in the literature as the *tapped delay line model* for the representation of time-varying frequency-selective channels. In this context, $h_l(t)$, $l = 1, 2, \ldots, L_p$, are the associated complex tap coefficients (or weights). It is interesting to note that, by substituting Equation 4.47 into 4.46 and taking into account Equation 4.43, we validate the input-output relation

$$y_\ell(t) = x_\ell(t) * h_\ell(\tau; t) \tag{4.48}$$

Before proceeding, we need to revisit Equation 4.40 to discuss about the lowpass impulse response $g_\ell(t)$ of the LTI filter. In the system under study, the wireless channel is considered to be band-limited to W, namely its time-varying lowpass frequency response, which is defined now as $H_\ell(f; t) = \mathcal{F}_\tau\{h_\ell(\tau; t)\}$, should be zero for $|f| > W/2, \forall t$.* This implies that the transmitted signal $x_\ell(t)$ shall initially be limited to $W/2$, which is accomplished through the transmit pulse-shaping filter with impulse response $p_T(t)$. Hence, the output of the filter is given by replacing $g_\ell(t) = p_T(t)$ into Equation 4.40, that is, $x_\ell(t) = s_\ell(t) * p_T(t)$, which permits the faded signal in Equation 4.40 to be written in the following form

$$y_\ell(t) = s_\ell(t) * p_T(t) * h_\ell(\tau; t) \tag{4.49}$$

In Equation 4.49, the associative property of convolution has been invoked. It would be useful to mention here that the associative and distributive properties still hold for LTV systems, in contrast to the commutativity that is generally valid only for LTI systems [44]. Also, LTV systems satisfy the property of homogeneity of degree one [41, Sec. 6.3.1].

4.3.2 Received Signal in the Presence of Noise

In addition to the fading channel impairments, the Rx front end introduces another source of signal degradation. This is the noise generated by electronic components and devices due to the thermal agitation of electrons, which can be closely modeled by a stationary additive white Gaussian stochastic process with zero mean and flat power spectral density (PSD), a.k.a. power density spectrum, equal to $N_0/4$ W/Hz for all frequencies f.† We need to mention here that, other types of additive noise include atmospheric noise, cosmic noise, man-made noise, and interference coming from other communication systems [1, Ch. 1]. Nevertheless, in this chapter, only the thermal noise is taken into account, which according to the central limit

* Note that the bandwidth of the bandpass signal is twice the bandwidth of the equivalent baseband.

† The PSD of thermal noise is denoted here as $N_0/4$, instead of $N_0/2$, for power normalization purposes.

theorem is Gaussian distributed. In addition, thermal noise is simply characterized as *stationary*, as a wide-sense stationary (WSS) Gaussian stochastic process is strictly stationary as well.

As a white stochastic process cannot be expressed in terms of IQ components due to its wideband character, it is convenient to consider the noise process at the output of an ideal bandpass filter with a completely flat passband that includes the frequencies satisfying the inequalities $|f - f_c| \leq W/2$ and $|f + f_c| \leq W/2$. Particularly, the PSD of the bandpass white noise $\eta_b(t)$ can be written in closed form as $\mathcal{S}_{\eta_b}(f) = N_0/4\left[\Pi\left((f - f_c)/W\right) + \Pi\left((f + f_c)/W\right)\right]$. Thus, the signal $z_\ell(t)$ at the Rx input can be modeled as the faded signal $y_\ell(t)$ subject to the additive complex random signal $\eta_\ell(t)$

$$z_\ell(t) = y_\ell(t) + \eta_\ell(t) \tag{4.50}$$

In Equation 4.50, the sample function $\eta_\ell(t)$ of the equivalent lowpass noise process can now be expressed in terms of its IQ components as $\eta_\ell(t) = \eta_I(t) + j\eta_Q(t)$. Following the analysis in [1], it is straightforward to show that both $\eta_I(t)$ and $\eta_Q(t)$ are sample functions of a zero-mean stationary white Gaussian stochastic process with PSD functions given by $\mathcal{S}_{\eta_I}(f) = \mathcal{S}_{\eta_Q}(f) = N_0/2$ for $|f| \leq W/2$; $= 0$ for $|f| > W/2$, or more compact by $\mathcal{S}_{\eta_I}(f) = \mathcal{S}_{\eta_Q}(f) = (N_0/2)\Pi\left(f/W\right)$. Furthermore, $\eta_I(t), \eta_Q(t)$ are jointly-stationary random signals. As a Gaussian process is completely characterized by its first two moments, it would also be useful to employ the Wiener-Khinchin theorem,* in order to calculate the associated autocorrelation functions as $\mathcal{R}_{\eta_I}(\tau) = \mathcal{F}_f^{-1}\{\mathcal{S}_{\eta_I}(f)\} = \mathcal{F}_f^{-1}\{\mathcal{S}_{\eta_Q}(f)\} = \mathcal{R}_{\eta_Q}(\tau) = (N_0/2)W\text{sinc}(W\tau)$. Finally, given that the cross-correlation function $\mathcal{R}_{\eta_Q\eta_I}(\tau)$ is zero, the zero-mean IQ components of $\eta_\ell(t)$ are uncorrelated, and since they are Gaussian, they are also statistically independent [45, Sec. 7.1].

The next operation involves the convolution of $z_\ell(t)$ with the LTI impulse response $p_R(t)$ of the receive filter. Taking into account Equations 4.49 and 4.50, and by invoking successively the distributive and associative properties of convolution, the filtered version of $z_\ell(t)$ can easily be derived as

$$r_\ell(t) = z_\ell(t) * p_R(t) = s_\ell(t) * \hbar_\ell(\tau; t) + n_\ell(t) \tag{4.51}$$

where:

$\hbar_\ell(\tau; t) = p_T(t) * h_\ell(\tau; t) * p_R(t)$ is the compound channel impulse response
$n_\ell(t) = \eta_\ell(t) * p_R(t)$ is the lowpass noise signal at the output of the receive filter

Thus, given that the output of the baseband IQ modulator can be expressed in the form [46, Sec. 1.2]

* According to this theorem, the PSD and autocorrelation functions of a WSS process form a Fourier transform pair.

$$s_\ell(t) = \sum_{k=-\infty}^{\infty} s_I(kT)\delta(t-kT) + j\sum_{k=-\infty}^{\infty} s_Q(kT)\delta(t-kT) = \sum_{k=-\infty}^{\infty} s_\ell(kT)\delta(t-kT) \quad (4.52)$$

the output of the receive filter is derived by substituting Equation 4.52 into 4.51, recalling the distributive property of convolution, and applying the integral formulation of Equation 4.46 for its calculation, finally yielding

$$r_\ell(t) = \sum_{k=-\infty}^{\infty} s_\ell(kT)\hbar_\ell(t-kT;t) + n_\ell(t) \quad (4.53)$$

It turns out that analysis can proceed with the establishment of the complex baseband system model instead of the original bandpass one by employing the equivalent lowpass signals and channels. In the sequel, the subscript ℓ will be dropped from all complex envelopes for notational simplicity.

4.3.3 Waveform and Discrete Channel Models

We can elaborate more on Equation 4.53 by adopting the quasi-static flat fading (QSFF) channel that was firstly adopted by Foschini and Gans (1998) in their revolutionary work [47] and is very popular in the context of multiple-input multiple-output (MIMO) and especially space-time block coding systems. This channel model corresponds to a special case of a slow fading environment, where the fading coefficients $h_l(t)$, $l = 1,2,\ldots,L_p$ are assumed to be constant over a frame of $N \in \mathbb{N}$ successive time intervals changing independently from frame to frame, implying that the channel coherence time T_c spans N signaling intervals, that is, $T_c = NT$. In this case, inequality $T_c \geq T$ is fulfilled, thus rendering the fading channel time-nonselective.

Under these channel conditions, it is convenient to consider the associated with the mth transmitted block impulse response of the multipath radio channel, which enables Equation 4.47 to be written in terms of τ only as

$$h^{(m)}(\tau;t) \triangleq h^{(m)}(\tau) \triangleq \sum_{l=1}^{L_p} h_l^{(m)}\delta(\tau-\tau_l) \quad (4.54)$$

where the fading coefficient $h_l^{(m)}$ is constant within the time interval $[(m-1)NT, mNT]$, $m \in \mathbb{Z}$, that is the duration (or length) of the mth block. Following the same notation, the associated with the mth transmitted block compound channel impulse response is reduced now to $\hbar^{(m)}(\tau;t) \triangleq \hbar^{(m)}(\tau)$, which obviously corresponds to the impulse response of an equivalent filter consisting of three LTI filters in cascade. Having this in mind, we can now proceed by considering the mth block-related version of Equation 4.53

$$\begin{aligned} r^{(m)}(t) &= \sum_{k=-\infty}^{\infty} s(kT)\hbar^{(m)}(t-kT)+n^{(m)}(t) \\ &= \sum_{k=-\infty}^{\infty} s(kT)\Big(h^{(m)}(t-kT)*p(t)\Big)+n^{(m)}(t) \end{aligned} \tag{4.55}$$

where $p(t)=p_T(t)*p_R(t)$ is the nonrelated to block index m overall impulse response of the combined transmit-receive filter.* Taking into account Equation 4.54, Equation 4.55 can be further simplified to

$$r^{(m)}(t)=\sum_{l=1}^{L_p} h_l^{(m)} \sum_{k=-\infty}^{\infty} s(kT)\,p(t-kT-\tau_l)+n^{(m)}(t), \quad (m-1)NT \le t \le mNT \tag{4.56}$$

This last formulation of the signal at the output of the receive filter is very useful for the analysis of MIMO systems, since it incorporates the analog portion of the system in a single chain, comprising the filter/modulator at the Tx, the noisy fading channel, and the demodulator/filter at the Rx. Since it represents the signal waveform relation from the output of the baseband IQ modulator to the input of the Rx decoder, it is also called *waveform channel model*. As described in [40, Sec. 11.2], an equalizer could also be part of the waveform channel. Nevertheless, the equalization procedure has been intentionally omitted from Equation 4.56, since mitigation of ISI caused by the time-dispersive nature of the channel is not our target in this section.

Consequently, analysis will proceed with the replacement of the waveform-based channel model by a discrete-time one, assuming either perfect equalization or flat fading channel conditions. We highlight that, flat fading is a valid consideration under the narrowband signal transmission scenario that has been adopted here (see Section 4.2), whereas for the case of a frequency-selective channel, orthogonal frequency division multiplexing (OFDM) can be used to convert a broadband channel into a set of parallel narrowband (i.e., flat) subchannels (see Section 4.4). Under this consideration, Equation 4.56 may be further simplified by setting $L_p=1$ and defining $h_1^{(m)} \triangleq h^{(m)}$ and $\tau_1 \triangleq \tau$, which finally takes the form

$$r^{(m)}(t)=h^{(m)}\sum_{k=-\infty}^{\infty} s(kT)\,p(t-kT-\tau)+n^{(m)}(t), \quad (m-1)NT \le t \le mNT \tag{4.57}$$

The *discrete channel model* can be easily derived by sampling the filter output at the symbol rate, namely at time instants $t_\nu=\nu T+\hat{\tau}, \nu \in \mathbb{N}$, where $\hat{\tau}$ is the propagation delay estimate provided by the timing recovery loop of the Rx. Obviously, assuming

* In Equation 4.55, the commutative property of convolution has been utilized, which holds for LTI systems.

perfect symbol synchronization, the timing error $\varepsilon = \hat{\tau} - \tau$ becomes zero, and under causality constraints, that is, $k, \nu \in \mathbb{N}$, the sampled version of Equation 4.57 is

$$r^{(m)}(\nu) = h^{(m)} \sum_{k=1}^{\nu} s(k)\, p(\nu - k) + n^{(m)}(\nu) \tag{4.58}$$

where in Equation 4.58 we have introduced the notations $r^{(m)}(t_\nu) \triangleq r^{(m)}(\nu)$ and $n^{(m)}(t_\nu) \triangleq n^{(m)}(\nu)$ as well as $s(kT) \triangleq s(k)$ and $p(qT) \triangleq p(q)$, with $q \in \mathbb{N}_0$. We may also write Equation 4.58 in the more meaningful form

$$r^{(m)}(\nu) = p(0) h^{(m)} s(\nu) + \upsilon^{(m)}(\nu) + n^{(m)}(\nu) \tag{4.59}$$

with $s(\nu)$ being the desired symbol at the νth sampling instant and $\upsilon^{(m)}(\nu) = h^{(m)} \sum_{k=1}^{\nu-1} s(k)\, p(\nu - k)$ representing the ISI produced from symbols transmitted prior to $s(\nu)$. At this point, a few words have to be spent on the discrete-time impulse response $p(q)$ of the combined transmit-receive filter. From communication theory, it is known that the optimal choice of $p(q)$ is the one that satisfies the Nyquist condition for zero ISI, which is simply expressed as $p(q) = 1$ for $q = 0; = 0$ for $q \neq 0$ [1]. In this case, $\upsilon^{(m)}(\nu) = 0, \forall \nu$, and Equation 4.59 reduces to the more convenient to handle expression

$$r^{(m)}(\nu) = h^{(m)} s(\nu) + n^{(m)}(\nu) \tag{4.60}$$

Considering now the transmission of the mth block, which takes place within the time interval $[(m-1)NT, mNT]$, the associated global sampling index ν is restricted to the set $\{(m-1)N + \kappa : \kappa \in \mathbb{N}; \kappa \leq N\}$. Evidently, the next set of relations hold an equivalent to Equation 4.60 meaning within a block

$$\left.\begin{aligned} r(\kappa) &= h s(\kappa) + n(\kappa) \\ r_\kappa &= h s_\kappa + n_\kappa \end{aligned}\right\}, \quad \kappa = 1, 2, \ldots, N \tag{4.61}$$

where the new block-based sampling index κ takes now the same values for all m. Interestingly, for $N = 1$, which is a special case of the QSFF channel, and after dropping subscript κ for notational convenience, the following scalar input–output relationship is obtained

$$r = hs + n \tag{4.62}$$

The communication system model in Equation 4.62 describes the well-known complex baseband representation of the open-loop SISO scheme for narrowband digital signal transmission over slow fading channels with coherent reception and no transceiver impairments. The multiplicative fading coefficient $h = a \exp(j\theta)$ represents the independent and identically distributed (i.i.d.) snapshots of the flat fading channel, where the statistics of its amplitude (or fading envelope) $|h| = a$ and

phase $\angle h = \theta$ both depend on the nature of the radio propagation environment (for details on the statistics of various flat fading channels, see Section 4.1.2).

4.4 Fading Mitigation Techniques and Diversity

In the system model described by Equation 4.62, N has been interpreted as the number of successive time slots for the single-antenna transmission of N consecutive digital symbols, all experiencing the same fading channel conditions due to the QSFF model assumed. However, this transmission strategy is very vulnerable to fading, as each information symbol realizes just a single instance of the channel. To make it more clear, during a block transmission under unfavorable channel conditions (e.g., deep fading), a whole set of N information symbols is very likely to be erroneously detected. Severe fading also causes failure with large probability in the symbol detection of SISO systems defined by Equation 4.62, although in this case only a single symbol is corrupted by the associated fading coefficient. However, these open-loop configurations can bring benefits in terms of reliability if extra dimensions are introduced and ingeniously utilized along with the spatial one.

In order to give evidence of such a capability, suppose that every individual information symbol utilizes L independent channel snapshots for its transmission through the wireless channel. In this case, there are apparently L new degrees of freedom available in the system. However, a critical question that arises is how these can be efficiently exploited toward more reliable communication. Frankly speaking, there is not a unique methodology for accomplishing this goal. Instead, the happy medium is rather a compromise among the available resources (e.g., bandwidth, power, space) and the practical limitations imposed by the intended application.

In Section 4.2, we have already discussed the two major effects of fading on signal transmission, namely (a) signal distortion mainly due to ISI and (b) SNR penalty in the error performance with respect to the ideal AWGN case. In the sequel of this chapter, efficient ways to tackle these problems will be discussed, starting from the latter.

4.4.1 Compensation of Intersymbol Interference

The ISI caused by a time-dispersive channel can be efficiently compensated by equalization, OFDM, and spread spectrum techniques. In more detail, the process of equalization provides a means to gather the dispersed symbol energy back into its original time interval. To do so, the equalizer adjusts the balance between frequency components, thus approximating an inverse filter of the channel. The target for the combination of the channel and equalizer filter is to provide a flat composite received frequency response and linear phase.

One of the most effective nonlinear equalization techniques is the maximum likelihood sequence estimation (MLSE), first proposed by Forney [48], that is typically implemented using the Viterbi (1967) decoding algorithm. However, in practical scenarios (e.g., mobile radio channels), optimal equalization techniques cannot always be used due to constraints on processing power. In these cases, OFDM is an attractive alternative that achieves the same goal by demultiplexing a data sequence into several streams, transmitted in parallel on different subcarriers and with sufficiently small symbol rates with respect to B_c, so as to consider the subchannels as flat. In addition, the residual ISI can be completely eliminated by the insertion of a cyclic prefix or guard interval, which is always larger than T_m, between successive OFDM symbols. It is interesting to note here that Cimini (1985) proposed and described the use of OFDM in mobile communications. However, the fundamental principle of OFDM is credited to Chang [49] who studied the simultaneous transmission of several messages through a linear bandlimited medium at a maximum data rate without interference.

Spread spectrum techniques can also compensate for the ISI by a different mechanism which is based on bandwidth expansion, namely the transmitted signal is spread over a much larger bandwidth than the minimum required (i.e., the bit data rate). This can be achieved by the introduction of a pseudo-random pattern in the transmitted signal that is later removed at the demodulator to recover the original data sequence without ISI. There are two basic implementations of this technology; the direct-sequence spread spectrum (DS-SS) and the frequency-hopping spread spectrum (FH-SS). According to Scholtz (1982), frequency hopping was a recognized anti-jamming concept during the early 1940s, while the direct sequence followed several years later.

In more detail, in DS-SS systems, a pseudo-random sequence (PRS) of pulses (or chips) is impressed on the original data sequence expanding the allocated bandwidth by a factor equal to the number of chips per information bit, called *processing gain*.* As the spread signal bandwidth is approximately equal to the chip rate, multipath components delayed by more than the chip interval can be suppressed by correlating the signal with the same synchronized PRS at the Rx. DS-SS is the key technology behind code division multiple access (CDMA), introduced by Qualcomm, Inc. in 1989, that allows several users to share the same band simultaneously by using distinct PRS codes among different pairs.

In a different way, in a FH-SS system, the total spread bandwidth is divided into a large number of contiguous frequency slots. In any signaling interval, one or more frequency slots are pseudo-randomly selected, according to the PRS, to be occupied by the transmitted signal. Obviously, during this interval, the Rx must be set to the same PRS in order to be tuned to the proper frequency band(s) and hop to the next scheduled slot(s) before the arrival of multipath components.†

* Equivalently, a spread spectrum system is characterized by a chip rate that is much larger than the bit data rate.

† The hopping rate of a FH-SS system must be at least equal to the symbol rate.

Nevertheless, in the presence of doubly-selective fading, multipath mitigation strategies alone cannot compensate for the error floor if not assisted by adaptive channel estimation and robust modulation schemes, such as noncoherent, partially coherent, and pilot symbol assisted modulation,* among other techniques.

4.4.2 The Concept of Diversity Gain

The next step toward reliable communication involves examining ways to reduce the fading-induced penalty on the SNR. This can be efficiently achieved through *diversity*, a classic and well-known concept that has been used for several decades. However, the term *diversity* is rather general and comprises the vast majority of fading mitigation techniques. Diversity is based on the notion that the detector is more prone to errors when the channel is in a deep fade. Therefore, if the Rx is supplied with multiple versions of the same information-bearing signal transmitted over independent fading channels, then the probability that all signals have experienced deep fading will be considerably reduced. In other words, if L is the number of independent fading channels utilized for the transmission of the same signal and p is the probability that any signal will fade below some critical value, then p^L is the probability that all replicas will fade below that critical value.

From another standpoint, diversity has the effect of steepening the rate of descent of the error probability $P_e(\overline{\gamma})$ versus the average SNR, $\overline{\gamma}$. Particularly, in the high SNR regime, the term *diversity order* is used to define the negative slope of $P_e(\overline{\gamma})$ curve on a log–log scale, namely

$$d = -\lim_{\overline{\gamma}\to\infty} \frac{\log P_e(\overline{\gamma})}{\log \overline{\gamma}} \tag{4.63}$$

By definition, a system is said to achieve *full diversity* if the diversity order d is equal to the number of independent channel realizations (i.e., $d = L$). In this context, a full-diversity scheme is able to exploit all the available DoFs in the direction of decreasing the error probability for a given average SNR and spectral efficiency, while a system with no diversity is characterized by $d = 1$. In order to reveal the intuition behind Equation 4.63, let assume an $1\times N$ single-input multiple-output (SIMO) maximum ratio combining (MRC) system employing a 16th order quadrature amplitude modulation (QAM) constellation and MLD at the Rx. The fading channel is considered to be slow flat Rayleigh with unity average fading power, that is, $\Omega = 1$ (see Section 4.1.2.1). The average symbol error rate (SER) performance of the system with respect to the average SNR per symbol is shown in Figure 4.4 for various values of N, namely for $N = 1,2,\ldots,10$. We first select a relatively high SNR value, for example 50 dB, and measure the corresponding θ_N angles for all N. Then, the slope of each SER curve at SNR equal to 50 dB is given by definition from $-\theta_N$.

* Analysis of pilot symbol assisted modulation was initially provided by Cavers (1991) for flat Rayleigh fading.

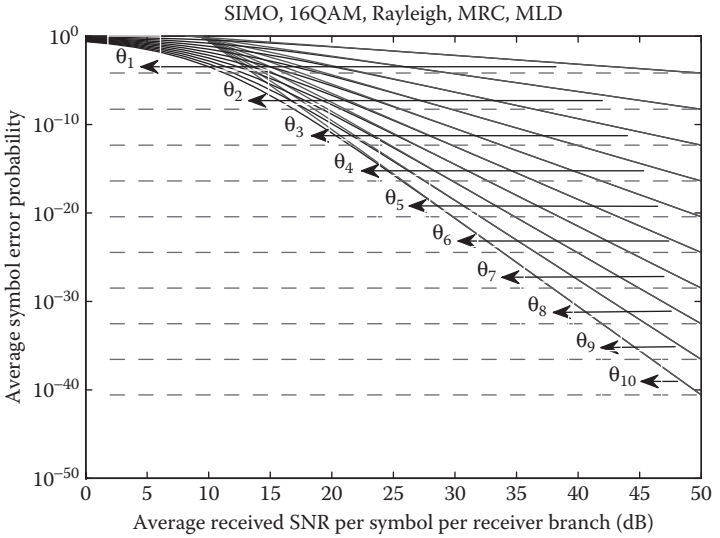


Figure 4.4 Graphical illustration of *d* for a SIMO MRC system with one-ten antennas at the Rx.

The graphically-based values for the diversity order can be derived by making use of Equation 4.63 or equivalently by $d_{gr} = \tan(-\text{slope})$. The calculated values are shown in Table 4.1 and obviously they are very close to the expected ones, that is, $d_{th} = N$, which stem from the theory of MIMO systems (note that SIMO MRC are full-diversity schemes).

Table 4.1 Calculation of *d* with the Aid of Figure 4.4

N	θ_N (*deg*)	*Slope* (*deg*)	d_{gr}	d_{th}
1	44.9970	−44.9970	0.9999	1
2	63.4324	−63.4324	1.9998	2
3	71.5631	−71.5631	2.9997	3
4	75.9622	−75.9622	3.9995	4
5	78.6888	−78.6888	4.9994	5
6	80.5366	−80.5366	5.9993	6
7	81.8689	−81.8689	6.9992	7
8	82.8741	−82.8741	7.9990	8
9	83.6590	−83.6590	8.9989	9
10	84.2887	−84.2887	9.9988	10

The benefit of diversity on the system error performance can also be quantified through the related term *diversity gain*, which determines the improvement in the average received SNR obtained at a given error probability using some form of diversity. Very often, the terms diversity order and diversity gain are used interchangeably in the literature to define the increase in the slope of the error rate. However, the distinct definitions coined previously for the two terms are more common [50]. Interestingly, in the high SNR limit, the error probability can be approximated by $P_e(\overline{\gamma}) \approx (G_c \overline{\gamma})^{-d}$, where G_c is the *coding gain* of the system that determines the average SNR shift (in dB) of the $P_e(\overline{\gamma})$ curve relative to a benchmark transmission scheme [51].

In Figures 4.5 and 4.6, the performance of SIMO MRC communication systems in terms of average SER is illustrated for several values of the diversity order d. In all cases, QAM is employed at the Tx and MLD at the Rx, while the fading channel is assumed to be slow flat Rayleigh. Given that SIMO MRC are full-diversity schemes, the diversity order is always equal to the number of deployed antennas at the Rx. It can be easily observed that, for a giver error probability, increasing the diversity order implies substantial diversity or equivalently SNR gains, which are translated to essential savings in the required transmitted power. This great advantage has been exploited for several decades now in the uplink of wireless networks where the design of the Tx is often subject to fabrication and/or power limitations.

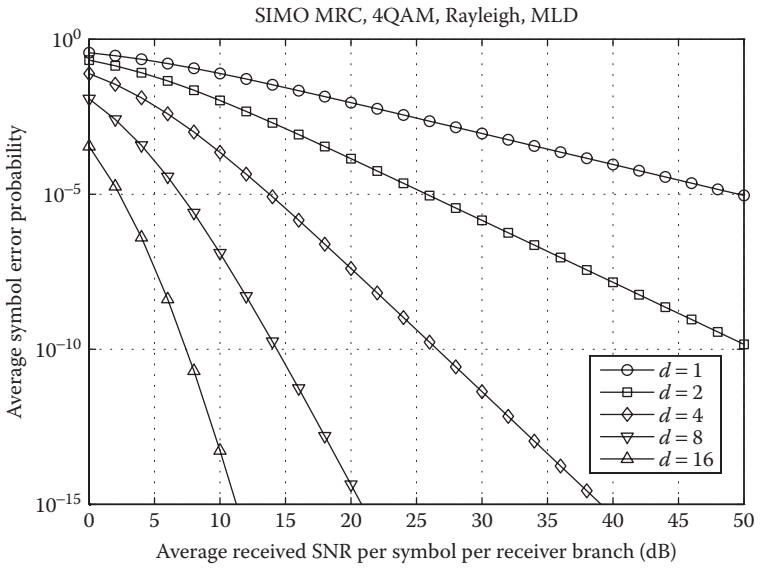


Figure 4.5 Error performance of a 4QAM communication system operating over a slow flat Rayleigh fading channel for different values of the diversity order *d*.

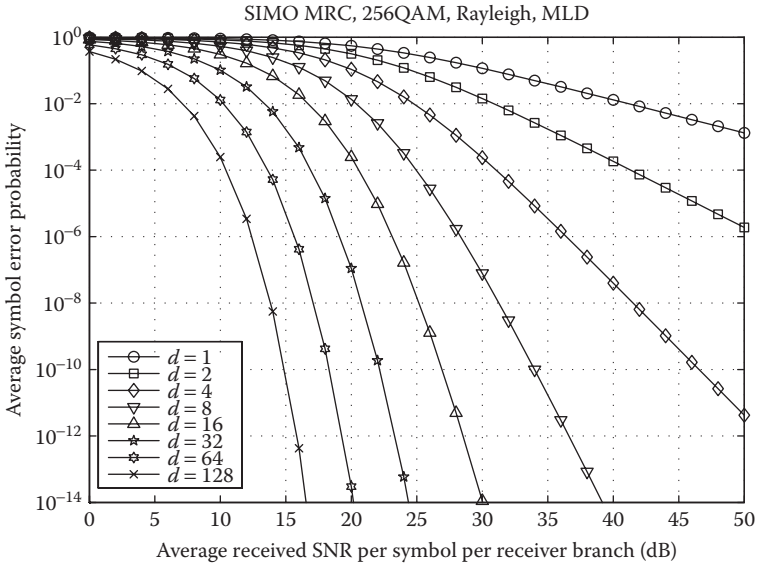


Figure 4.6 Error performance of a 256QAM communication system operating over a slow flat Rayleigh fading channel for different values of the diversity order *d*.

4.4.3 Tracing the Roots of Diversity

It is very common in the literature to trace the roots of diversity back to 1940s or even early 1950s, probably due to the plethora of diversity receivers proposed during that period, whereas just a few works point out related techniques emerged in 1930s in the context of wireless telegraphy; just to mention the seminal work of Beverage and Peterson (1931) that gives implementation details of some commercial transceiver systems employing three different principles of diversity [52]. An early rack installation of diversity equipment is shown in Figure 4.7a, which illustrates a double rack holding diversity apparatus for two circuits. In 1931, the Riverhead receiving station of R.C.A. Communications, Inc., contained twenty such racks, providing facilities for forty circuits in total. These racks were arranged back to back facing aisles with four circuits per aisle. Ordinarily one man was taking care of the circuits in an aisle. Figure 4.7b presents the end view of an aisle.

However, the diversity concept was known several years before. Particularly, Eckersley [53] gives a detailed diagram, reproduced in Figure 4.8, of an Rx apparatus employing one horizontal and six vertical polarized aerials placed in a line at right angles to the station ray about 12 ft above the surface of the earth. These were connected by phasing coils consisting of vertical lecher wires approximately (1/4)λ in length, λ being the wavelength. The Rx was coupled through a cable to the central lecher wire. This special arrangement was used in a series of tests carried out during

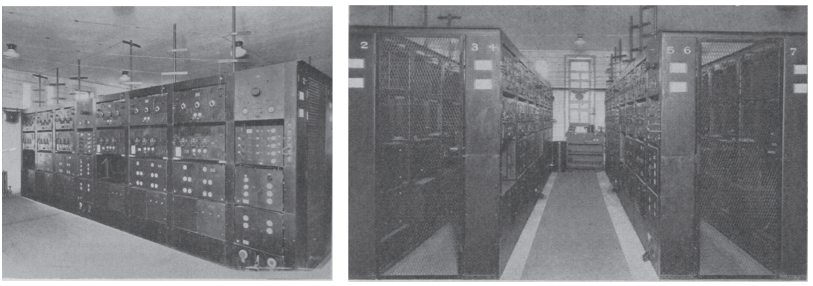


Figure 4.7 Diversity receiving system of R.C.A. Communications, Inc. © 1931 IEEE. (a) Diversity Rx installation. (b) Aisle of four complete diversity sets. (Reprinted with permission from Beverage, H.H. and Peterson, H.O., *Proc. IRE*, 19, 531–561, © 1931 IEEE.)

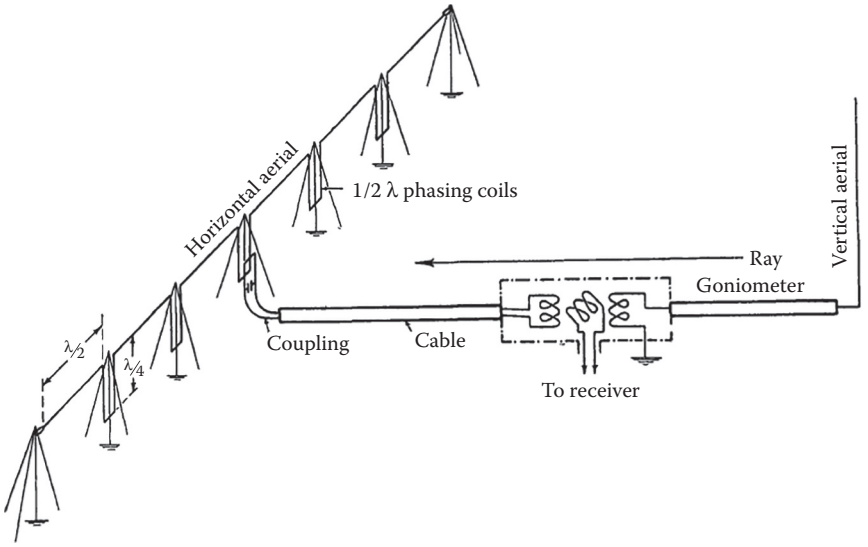


Figure 4.8 The Rx apparatus of Eckersley (1929) for testing polarization diversity. (From Eckersley, T.L., *J. IEEE*, 67, 992–1029, 1929. Reproduced by permission of the Institution of Engineering & Technology.)

the winter of 1927–1928 to compare the relative fading on the two types of aerials, aiming to design a polarization diversity system (see Section 4.4.4.4) through particular combination of the rectified currents coming from the two differently polarized aerials. It is of great interest that, in the same paper, the term *diversity* appears for the first time in the literature to describe one of the very first receive diversity (RxD)

techniques, based on *spaced aerials*. The following excerpt reveals that diversity systems had already been quite common among radio experts by that time:

> The fading on two aerials spaced many wave-lengths apart should be very different, and it is actually found to be so. "Diversity" systems of this type, as they are called in America, are in use both here and in that country, and show considerable gain in leveling the fading.
>
> **—Eckersley (1929) [53].**

Furthermore, in 1928, another form of diversity, that is time diversity (see Section 4.4.4.3), was conceived by Verdan and Loiseau, inventors of the antiparasitic device, which was capable of improving the link performance of the Baudot telegraph system subject to atmospheric disturbances [54]. This apparatus, of the later called Baudot-Verdan system, and the associated Verdan principle that calls for retransmission of the same letter with certain delays, are described in a related paper by Harrison [55]. Also, it is almost unknown in the context of wireless communications that, a directional receiving aerial system, patented by Adcock [56] for the purpose of eliminating the errors on closed-coil direction-finders due to down-coming waves, has probably served as a basis for the design of primitive transmit diversity (TxD) systems, such as the rotating-beacon Tx scheme of Smith-Rose. As shown in the schematic diagram in Figure 4.9, the Adcock system may be regarded simply as a pair of spaced vertical aerials with the currents flowing in opposite phase [57].

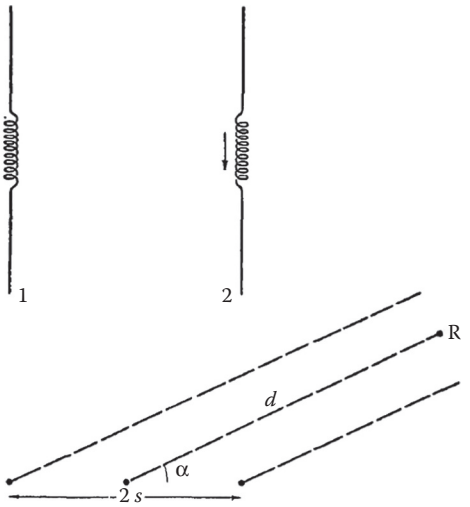


Figure 4.9 Schematic diagram of the Adcock aerial system given by Smith-Rose. (From Smith-Rose, R.L., *J. IEEE*, 66, 270–279, 1928. Reproduced by permission of the Institution of Engineering & Technology.)

Finally, a very interesting spatial TxD system was presented by Hansell (1955), affiliated with R.C.A. Laboratories. In more detail, a wireless system employing two Txs, connected to spaced antennas, and a single Rx antenna was found to be equivalent to a conventional RxD system in terms of overcoming the effects of fading, given that the two Txs are separated in frequency by a very small amount of offset [58]. According to Hansell, the proposed system could be especially suitable for shore to ship transmission because of the impracticability of providing spaced receiving antennas on ships. The paper concludes that it is more cost-effective to improve the signal reception by the addition of a second low-power Tx rather than to increase the power of the single Tx in order to obtain the same performance. Particularly, to do so, the single Tx would require 16–1000 times more power than each of the two Txs. The closing paragraph of [58] is worth being highlighted:

> There is another important advantage that is obtained by transmitter diversity compared to receiver diversity. The receiving equipment is... smaller in size and also easier to operate. This may be quite important, especially for mobile application where quite often space and operating personnel are at a premium.
>
> **—Hansell (1955) [58].**

Surprisingly, the adopted approach and the discussion made in this early work are very consistent with the revolution that took place some decades later in the context of spatial TxD systems, especially through the use of space-time coding and other MIMO techniques.

Before closing this section, it would be a great omission not to mention the Canadian-born inventor Fessenden, whose technological achievements definitely assisted the understanding of fading effects on signal transmission and their mitigation via the development of proper diversity techniques. In this context, if Marconi is considered to be the pioneer of wireless transmission over long distances, then Fessenden is credited as the first to demonstrate the transmission of voice and music by radio [59]. The works in [60,61] refer to Fessenden as a genius and mathematician, being the father of amplitude modulation (AM) radio and a primary pioneer of radio as we know it today. In November 1899, at the meeting of the American Institute of Electrical Engineers, Fessenden presented a paper on the "possibilities of wireless telegraphy" [62], and soon afterwards, in December 1900, he transmitted, for the first time in the wireless history, intelligible words by electromagnetic waves [60]. The transmission took place on Cobb Island, Maryland, over a distance of 1600 m using a mechanical spark interrupter [59]. However, the fact that his voice was badly distorted and accompanied by a loud noise, convinced him that he needed a "continuously acting, proportional indicating receiver" (Fessenden's words) [60]. For this reason, he advocated for

continuous-wave against the spark systems utilized by Marconi [62,63]. In January 1906, the first two-way transatlantic radio telegraphy transmission took place between Fessenden's stations at Brant Rock, Massachusetts and Machrihanish Scotland [61]. Interestingly, Fessenden's concept of continuous-wave radio signals and his development of the heterodyne principle are considered as his greatest contributions to radio technology [61].

4.4.4 Classification of Diversity Techniques

As already discussed, a critical issue in the design of digital communication systems is the provision for high immunity levels to fading by efficient deployment of diversity techniques. As will be shortly described, there is not a single technique that can guarantee this goal. Primarily, diversity systems can be classified according to the nature of fading they are intended to mitigate. Following the discussion in Section 4.2, we focus on *microdiversity* techniques that are used to combat the short-term (i.e., small-scale) fading effects.* Before proceeding with a brief summary of them, it would be important to note that practical diversity systems usually combine more than one diversity techniques for better exploitation of the available DoFs of the system.

4.4.4.1 Spatial Diversity

Spatial diversity, also known as *antenna diversity*,† is the most commonly used form to obtain several copies of the same information-bearing signal by deploying multiple antennas at the Rx or/and Tx side. The redundant signals are then skillfully combined in order to increase the total average SNR. The term *site diversity* is sometimes used as an alternative to spatial diversity. However, site diversity is a macrodiversity technique that fits better in the context of satellite communications systems operating in Ka band and above, where two or more ground stations are linked together to provide more robust downlink signal reception against rain fading.

Traditionally, the most popular schemes achieving diversity over space have been the SIMO systems, which are usually equipped with linear combiners that require partial or full channel state information (CSI) only at the Rx side. Though the first variants of RxD techniques date back to 1920s, they still offer one of the greatest potentials for radio link performance improvement to many of the current and future wireless technologies. The dual case of SIMO are the multiple-input

* In analogy, *macrodiversity* can be employed to mitigate the effects of long-term (i.e., large-scale) fading.

† More generally, antenna diversity can refer to either of the three techniques, spatial, polarization, or pattern diversity.

single-output (MISO) systems that can obtain TxD under some more sophisticated architecture, usually requiring signal processing and some level of CSI awareness at the Tx through a dedicated feedback link between the two sides of the communication link.

In the more general case, a MIMO system can offer diversity and coding gains, with the challenge being to satisfy the complexity constraints and transmitted power limitations put on the design. However, all configurations require sufficient separation between antenna elements on the same side to ensure independent fading conditions over all channels and realize the diversity advantage. Not very often, the minimum distance between antennas that yield uncorrelated fading is called *coherence space*.

4.4.4.2 Frequency Diversity

In *frequency diversity*, the same information-bearing signal is transmitted on L carriers, where the separation between successive carriers must equal or exceed B_c. Obviously, this approach is very well suited to multicarrier systems working in frequency division multiplexing (FDM) mode. One of the very first applications of this technique in wireless telegraphy is described by Beverage and Peterson [52]. However, the major disadvantages of similar systems are generally the non-efficient bandwidth usage and the requirement for receivers with multiple radio frequency (RF) chains. A more refined version of multicarrier operation to achieve the same target is the combination of OFDM with channel coding and frequency interleaving in the presence of time-dispersive channels.*

In a broader sense, frequency diversity is an intrinsic characteristic of wideband channels, that is, channels with a bandwidth greater than B_c. Particularly, wideband transmission can provide the receiver with several independently fading signal replicas, thus achieving frequency diversity of order $d = W/B_c$, given that the induced ISI is efficiently mitigated. A possible approach to deal with this issue is single-carrier transmission with channel equalization. Alternatively, FH-SS technology utilizes bandwidth expansion to exploit the inherent frequency diversity of the channel. Olofsson et al. [64] showed that FH-SS also introduces *interference diversity*, since different signals interfering with the desired one at different times can result in an averaging effect, which improves performance in system level.

Finally, a DS-SS system employing a *RAKE receiver* is capable of resolving multipath components at different time delays, a phenomenon also called *multipath diversity*. The RAKE receiver, proposed by Price and Green [65], uses several correlators or *fingers*, each assigned to a different multipath component, to separate and coherently combine all the delayed replicas of the wideband signal.

* Frequency interleaving is the mapping of the coded bits to the subcarriers.

4.4.4.3 Time Diversity

Time diversity exploits the frequency-dispersive nature of time-varying wireless channels to provide more reliable communication. This can be easily accomplished by using multiple time slots separated by at least the coherence time, T_c, of the channel to transmit the same data symbols. Particularly, every symbol is repeated L times and then interleaved before being transmitted over the wireless channel, so that consecutive symbols experience independent fading conditions.* Obviously, this technique can be viewed as repetition coding of the information sequence, with the length L of the code defining the diversity order obtained. However, the achieved gain comes at the expense of spectral efficiency, since a single symbol requires transmission over several time intervals. Furthermore, a repetition code does not effectively exploit all the DoFs available in a fading channel.

Instead, more sophisticated coding schemes, such as linear block codes or convolutional codes, along with interleaving can be employed to realize additional coding gain.† In this way, the encoded symbols can be efficiently dispersed over different coherence periods, thus taking full advantage of the inherent time diversity in time-selective channels. Nevertheless, as the fading rate (i.e., the B_d) of a fading channel decreases, for example, due to mobile terminals that slow down, it may not be possible to obtain time diversity without introducing unacceptable delays, since in this case the inevitable large T_c limits the performance of a given interleaver.

4.4.4.4 Polarization Diversity

Early measurements on long-distance transmission revealed that propagation characteristics of a wireless medium are not the same for differently polarized waves. In addition, multiple reflections between the Tx and Rx cause depolarization of the radio waves, thus dispersing some energy of the transmitted signal into other polarization directions.‡ Due to that attribute, linearly polarized transmitted waves can come out at the Rx end with an additional nontrivial orthogonal component. These observations motivated many researchers to investigate the fading statistics associated with radio signals received by antennas of various polarization modes.

Among the first, Glaser and Faber [66] demonstrated that the vertical and horizontal polarized components of the same received signal undergo almost statistically independent fading while propagating through certain wireless

* This implies that $(L-1)T \geq T_c \Rightarrow L \geq (T_c/T)+1$, where T is the signaling interval.

† Diversity order provided by channel coding is directly related to the minimum Hamming distance of the codes.

‡ Breit (1927) maybe was the first to elaborate a theory on this phenomenon.

environments. This fact has been utilized to improve radio system performance through the use of *polarization diversity*, where two or more spatially separated unipolarized antennas are replaced by a single antenna structure of almost co-located elements by employing multiple polarizations. The simplest polarization diversity systems utilize dual-polarized antennas with two orthogonal components, that is, copolarized and cross-polarized, rendering them very handy for use in cellular radio networks owing to the minimized antenna installation space and the fact that mobile terminals can experience reliable communication whatever the handset angle of tilt.

4.4.4.5 Pattern Diversity

Antenna *pattern diversity* comprises another form of exploiting the inherent DoFs of multipath propagation environment. The concept behind this approach lies in the relationship that holds between the received signal fluctuation and the direction of the main lobe of the Rx antenna pattern. The first experimental results validating this statement were reported by Bruce [67] and Bruce and Beck [68], who found remarkable reduction in fading by using a horizontal rhombic antenna with an extremely sharp directional pattern. The success of their attempt was based on the existence of stable angular separation among signal components following different paths, thus enabling a steerable antenna with sufficiently sharp directivity to accept only one of them, maybe the best after some calibration, at any time.

Friis et al. [69] clearly demonstrated the qualitative relation between angles of arrival and propagation delays of the received signal components; the greater the delay the greater the angle above the horizontal (or azimuth) plane. In [70], Friis and Feldman (1937) exploited this phenomenon with the aid of a special arrangement whereby individual wave groups arriving at different vertical angles were received separately and, after delay equalization, combined coherently to offer *directional diversity* or *directivity diversity*. They accomplished that with the multiple unit steerable antenna (MUSA) system. According to Gregory and Newsome (2010), the MUSA array was the last major technological development in the short-wave communication era, representing the ultimate receiving system [71]. It was probably the most complex radio Rx ever built and gave valuable service between the 1940s and 1960s. Later, Vogelman et al. (1959) extended the concept of a single beam steerable to a multibeam antenna, where signal components associated with individual beams were particularly combined to offer *angle diversity* or *angle-of-arrival diversity* [72]. However, the effectiveness of this system depends upon the antenna characteristics, which have to be properly chosen so that the angle between adjacent beams is always smaller than the *angular spread* of the channel and that low correlation among different beams is ensured.

Obviously, pattern diversity can also be provided via several antennas with diverse patterns spaced apart from each other. An interesting case arises when

antennas with similar patterns are placed close to each other with respect to the coherence space defined in Section 4.4.4.1. Particularly, mutual coupling effects alter their antenna pattern, thus yielding again a quite effective pattern diversity scheme. In addition, *field component diversity*, first proposed by Gilbert [73], utilizes sophisticated energy density antenna arrangements to receive uncorrelated electric and magnetic field components of the transmitted signal so as to reduce fading in mobile radio. The conception of the energy density antenna as a means to mitigate signal fading in mobile radio channels is credited to Pierce, who communicated this idea to other researchers, such as Gilbert (1965) and Lee (1967).

4.4.4.6 Other Forms of Diversity

It has been mentioned earlier that the inherent diversity of frequency-selective fading channels can be exploited by spread spectrum techniques (e.g., DS-SS) with RAKE reception. However, the RAKE receiver is optimal for time-dispersive channels only, whereas for rapid temporal variations of the channel (i.e., fast fading), the resulting Doppler spread induces performance degradation due to inaccurate channel estimation. This fact motivated Sayeed and Aazhang [74] to suggest the dual version of the conventional RAKE receiver, called *Doppler RAKE receiver*, that is capable of resolving Doppler components at different frequency shifts. By adopting a joint time-frequency representation for the received signal, they first revealed the inherent diversity mechanism of a doubly-dispersive channel and further proved that it can be exploited by appropriate signal processing. This new receiver structure, which actually realizes B_d as another DoF, performs optimally in frequency-dispersive channels and offers *Doppler diversity* of order proportional to TB_d. In the more general case of a doubly-dispersive channel, Sayeed and Aazhang proposed again the *time-frequency RAKE receiver* that can offer in an optimal fashion joint multipath-Doppler diversity of order proportional to the product of the channel spread factor and the processing gain of the spread spectrum signaling technique [74].

Knopp and Humblet [75] studied the optimal power control scheme that maximizes information capacity in the uplink of single-cell multiuser communications corrupted by flat fading. They proved that the best strategy, that is scheduling at any time only the user with the largest SNR, provides *multiuser diversity* of order proportional to the number of users. Since then, the presence of multiple users on the same system has been conceived as a new DoF that was very effectively exploited by Sendonaris et al. [76] and a little later by Laneman et al. [77] through the introduction of the *user cooperation diversity* and *cooperative*

diversity concepts, respectively. These diversity gains can be achieved via the cooperation of intra-cell users that share their antennas and other resources so as to create a virtual array in a distributed manner. For this reason, Laneman and Wornell (2000) adopted also the term *distributed spatial diversity* [78].

A related concept, suggested by Boyer et al. [79], is the *multihop diversity*, where the benefits of spatial diversity are achieved from the concurrent reception of signals that have been transmitted by multiple previous terminals along a single primary route. Particularly, two models of multihop diversity channels were defined in [79]; the *decoded relaying* and the *amplified relaying*. In the former, each intermediate terminal combines, decodes, and reencodes the received signals from all preceding terminals, whereas in the latter, the intermediate terminals simply combine and amplify them, before retransmission. In addition, Boutros and Viterbo [80] introduced the *signal space diversity* or *modulation diversity*, whereby significant coding gains can be realized over fading channels with the aid of specially designed multidimensional QAM constellations. The key point behind this technique lies in the employment of particular lattice rotation, IQ interleaving/deinterleaving and decoding procedures.

Finally, in a different context, Callaghan and Longstaff [81] proposed the *waveform diversity*; a technique employing multiple transmit waveforms for suppressing synthetic aperture radar (SAR) range ambiguities (i.e., ambiguous range target returns). The waveform diversity scheme embeds, let say, N_b bits of information per pulse by selecting the waveform on a pulse-to-pulse basis from a set of 2^{N_b} waveforms, where each waveform represents a distinct N_b-bit symbol. The Rx obtains the embedded information in each pulse by determining the radar waveform that has been transmitted [82]. It is interesting to note that, according to the IEEE Standard Radar Definitions (IEEE Std 686-2008), this term is defined as the adaptivity of the waveform to dynamically optimize radar performance, while exploiting other domains such as time, frequency, coding, antenna radiation pattern, and polarization.

4.4.4.7 High-Level List of Pure Diversity Techniques

In Table 4.2, a quite extensive list of pure diversity techniques proposed so far in the literature is given. With the term *pure* it is implied here that only the diversity techniques which utilize a single basic principle or concept are included. Therefore, in Table 4.2, other very popular diversity strategies, such as space-time coding, space-time-frequency coding, and so on, are not included.

Table 4.2 List of Pure Diversity Techniques in Chronological Order

Diversity	*Details*	*Suggested by*	*Year, Reference*
Spatial	RxD techniques	Several	1920s
	TxD techniques	Several	1990s
Time	Antiparasitic device	Verdan and Loiseau	1928, [54]
	Coding with interleaving	Several	1950s
Polarization	Rx apparatus with aerials	Eckersley	1929, [53]
	Dual-polarized antenna	Glaser and Faber, Jr.	1953, [66]
Antenna pattern	Steerable antenna	Bruce	1931, [67]
	Directional diversity	Friis and Feldman	1937, [70]
	Angle diversity	Vogelman et al.	1959, [72]
	Field component diversity	Gilbert	1965, [73]
Frequency	FDM	Beverage and Peterson	1931, [52]
	FH-SS	Several	1940s
	OFDM	Chang	1966, [49]
Multipath	RAKE receiver	Price and Green	1958, [65]
Multiuser	–	Knopp and Humblet	1995, [75]
Interference	–	Olofsson et al.	1995, [64]
Waveform	–	Callaghan and Longstaff	1997, [81]
User cooperation	–	Sendonaris et al.	1998, [76]
Signal space	a.k.a. modulation diversity	Boutros and Viterbo	1998, [80]
Doppler	–	Sayeed and Aazhang	1999, [74]
Distributed spatial	–	Laneman and Wornell	2000, [78]
Cooperative	–	Laneman et al.	2001, [77]
Multihop	–	Boyer et al.	2001, [79]

References

1. J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.
2. T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice Hall, 1996.
3. M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 2005.
4. G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. New York: Kluwer Academic Publishers, 2002.
5. D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, UK: Cambridge University Press, 2005.
6. B. Sklar, Fading channels, Chapter 4, In *Handbook of Antennas in Wireless Communications*, Godara, L.C., (Ed.). Boca Raton, FL: CRC Press, 2002.
7. A. F. Naguib and A. R. Calderbank, Diversity in wireless systems, Chapter 3, In *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, Bölcskei, H., Gesbert, D., Papadias, C.B., and van der Veen, A.-J., (Eds.). Cambridge, UK: Cambridge University Press, 2006.
8. E. Dahlman, S. Parkvall, and J. Sköld, *4G LTE/LTE-Advanced for Mobile Broadband*. Oxford, UK: Academic Press, 2011.
9. E. Biglieri, J. Proakis, and S. Shamai, Fading channels: Information-theoretic and communications aspects, *IEEE Trans. Inf. Theory*, 44, 6, 2619–2692, 1998.
10. V. M. Kapinas, Optimization and performance evaluation of digital wireless communication systems with multiple transmit and receive antennas, PhD dissertation, Aristotle University of Thessaloniki, Thessaloniki, Greece, January 2014.
11. A. J. Coulson, A. G. Williamson, and R. G. Vaughan, A statistical basis for lognormal shadowing effects in multipath fading channels, *IEEE Trans. Commun.*, 46, 4, 494–502, 1998.
12. L. Rayleigh, On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, *Phil. Mag.*, 10, 60, 73–78, 1880.
13. R. G. Gallager, *Circularly-Symmetric Gaussian Random Vectors, Appendix to Principles of Digital Communication*. Cambridge, UK: Cambridge University Press, 2008.
14. U. Charash, A study of multipath reception with unknown delays, PhD dissertation, University of California, Berkeley, CA, January 1974.
15. S. O. Rice, Statistical properties of a sine wave plus random noise, *Bell Sys. Tech. J.*, 27, 1, 109–157, 1948.
16. M. Nakagami, The *m*-distribution–A general formula of intensity distribution of rapid fading, In *Statistical Method of Radio Propagation*, Hoffman, W.C., (Ed.). Oxford, UK: Pergamon Press, 1960, pp. 3–36.
17. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover Publications, 1972.
18. R. S. Hoyt, Probability functions for the modulus and angle of the normal complex variate, *Bell Sys. Tech. J.*, 26, 2, 318–59, 1947.
19. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic Press, 2007.
20. E. W. Stacy, A generalization of the gamma distribution, *Ann. Math. Statist.*, 33, 3, 1187–1192, 1962.
21. M. D. Yacoub, The $\alpha - \mu$ distribution: A physical fading model for the Stacy distribution, *IEEE Trans. Veh. Technol.*, 56, 1, 27–34, 2007.

22. V. A. Aalo, T. Piboongungon, and C.-D. Iskander, Bit-error rate of binary digital modulation schemes in generalized gamma fading channels, *IEEE Commun. Lett.*, 9, 2, 139–141, 2005.
23. P. S. Bithas, N. C. Sagias, and P. T. Mathiopoulos, GSC diversity receivers over generalized-gamma fading channels, *IEEE Commun. Lett.*, 11, 12, 964–966, 2007.
24. P. M. Shankar, Error rates in generalized shadowed fading channels, *Wirel. Pers. Commun.*, 28, 4, 233–238, 2004.
25. P. S. Bithas, N. C. Sagias, P. T. Mathiopoulos, G. K. Karagiannidis, and A. A. Rontogiannis, On the performance analysis of digital communications over generalized-*K* fading channels, *IEEE Commun. Lett.*, 10, 5, 353–355, 2006.
26. E. W. Marchant, Conditions affecting the variations in strength of wireless signals, *J. IEEE*, 53, 243, 329–340, 1915.
27. J. E. Taylor, Wireless telegraphy in relation to interferences and perturbations, *J. IEEE*, 47, 208, 119–140, 1911.
28. F. G. Loring et al., Discussion on wireless telegraphy in relation to interferences and perturbations, *J. IEEE*, 47, 208, 140–166, 1911.
29. R. H. Marriott, Radio range variation, *Proc. IRE*, 2, 1, 37–53, 1914.
30. L. W. Austin, Seasonal variation in the strength of radiotelegraphic signals, *Proc. IRE*, 3, 2, 103–106, 1915.
31. C.-P. Yeang, Characterizing radio channels: The science and technology of propagation and interference, 1900–1935, PhD dissertation, Program in Science, Technology and Society, M.I.T., Cambridge, MA, September 2004.
32. R. Bown and G. D. Gillett, Distribution of radio waves from broadcasting stations over city districts, *Proc. IRE*, 12, 4, 395–409, 1924.
33. R. Bown, D. K. Martin, and R. K. Potter, Some studies in radio broadcast transmission, *Proc. IRE*, 14, 1, 57–131, 1926.
34. L. De Forest, Recent developments in the work of the Federal Telegraph Company, *Proc. IRE*, 1, 1, 37–51, 1913.
35. C. R. Burrows, The history of radio wave propagation up to the end of World War I, *Proc. IRE*, 50, 5, 682–684, 1962.
36. A. H. Taylor and A. S. Blatterman, Variations in nocturnal transmission, *Proc. IRE*, 4, 2, 131–148, 1916.
37. G. W. Pickard, Short period variations in radio reception, *Proc. IRE*, 12, 2, 119–158, 1924.
38. J. H. Dellinger, L. E. Whittemore, and S. Kruse, A study of radio signal fading, In *Scientific Papers of the Bureau of Standards*, Vol. 19. Washington, DC: U.S. Government Printing Office, 1923, pp. 193–230.
39. G. K. Karagiannidis, *Telecommunication Systems*, (in Greek) 2nd ed. Thessaloniki, Greece: Tziolas Publications, 2011.
40. W. H. Tranter, K. S. Shanmugan, T. S. Rappaport, and K. L. Kosbar, *Principles of Communication Systems Simulation with Wireless Applications*. Upper Saddle River, NJ: Prentice Hall, 2004.
41. Y. Shmaliy, *Continuous-Time Systems*. Dordrecht, the Netherlands: Springer, 2007.
42. T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice Hall, 1980.
43. T. Kailath, Sampling models for linear time-variant filters, M.I.T. Research Laboratory of Electronics, Cambridge, MA, Technical Report 352, May 1959.

44. S. Barbarossa and A. Scaglione, Time-varying fading channels, Chapter 4, In *Signal Processing Advances in Wireless and Mobile Communications, Volume II: Trends in Single- and Multi-User Systems*, Stoica, P., Giannakis, G., Hua, Y., and Tong, L., (Eds.). Upper Saddle River, NJ: Prentice Hall, 2000.
45. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
46. C. Oestges and B. Clerckx, *MIMO Wireless Communications: From Real-World Propagation to Space-Time Code Design*. San Diego, CA: Academic Press, 2007.
47. G. J. Foschini and M. J. Gans, On limits of wireless communications in a fading environment when using multiple antennas, *Wirel. Pers. Commun.*, 6, 3, 311–335, 1998.
48. G. D. Forney, Jr., Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference, *IEEE Trans. Inf. Theory*, 18, 3, 363–378, 1972.
49. R. W. Chang, Synthesis of band-limited orthogonal signals for multichannel data transmission, *Bell Syst. Tech. J.*, 45, 10, 1775–1796, 1966.
50. C. B. Dietrich, Jr., K. Dietze, J. R. Nealy, and W. L. Stutzman, Spatial, polarization, and pattern diversity for wireless handheld terminals, *IEEE Trans. Antennas Propag.*, 49, 9, 1271–1281, 2001.
51. Z. Wang and G. B. Giannakis, A simple and general parameterization quantifying performance in fading channels, *IEEE Trans. Commun.*, 51, 8, 1389–1398, 2003.
52. H. H. Beverage and H. O. Peterson, Diversity receiving system of RCA Communications, Inc., for radiotelegraphy, *Proc. IRE*, 19, 4, 531–561, 1931.
53. T. L. Eckersley, An investigation of short waves, *J. IEEE*, 67, 392, 992–1029, 1929.
54. C. Verdan and L. Loiseau, Signaling system applicable to telegraphy and telemechanical transmission, U.S. Patent 1 677 062, July 1928.
55. H. H. Harrison, Developments in machine telegraph systems and methods of operation, *J. IEEE*, 68, 407, 1369–1453, 1930.
56. F. Adcock, Improvement in means for determining the direction of a distant source of electromagnetic radiation, British Patent 130 490, August 1919.
57. R. L. Smith-Rose, A theoretical discussion of various possible aerial arrangements for rotating-beacon transmitters, *J. IEEE*, 66, 375, 270–279, 1928.
58. G. E. Hansell, Transmitter space diversity as applied to shipboard reception, *IRE Trans. Prof. Commun. Syst.*, 3, 1, 44–46, 1955.
59. I. Brodsky, How Reginald Fessenden put wireless on the right technological footing, In *Proceedings of IEEE Global Telecommunications Conference*, New Orleans, LO, November 2008, pp. 1–5.
60. J. S. Belrose, Reginald Aubrey Fessenden and the birth of wireless telephony, *IEEE Antennas Propag. Mag.*, 44, 2, 38–47, 2002.
61. J. S. Belrose, Fessenden and Marconi: Their differing technologies and transatlantic experiments during the first decade of this century, In *Proceedings of 1995 International Conference on 100 Years of Radio*, London, UK, September 1995, pp. 32–43.
62. J. E. Brittain and A. Reginald, Fessenden and the origins of radio [scanning the past], *Proc. IEEE*, 84, 12, 1852–1853, 1996.
63. J. E. Brittain and A. Reginald, Electrical engineering hall of fame: Reginald A. Fessenden, *Proc. IEEE*, 92, 11, 1866–1869, 2004.

64. H. Olofsson, J. Naslund, and J. Skold, Interference diversity gain in frequency hopping GSM, In *Proceedings on IEEE Vehicular Technology Conference*, Chicago, IL, July 1995.
65. R. Price and E. Green, A communication technique for multipath channels, *Proc. IRE*, 46, 3, 555–570, 1958.
66. J. L. Glaser and L. P. Faber, Jr., Evaluation of polarization diversity performance, *Proc. IRE*, 41, 12, 1774–1778, 1953.
67. E. Bruce, Developments in short-wave directive antennas, *Proc. IRE*, 19, 8, 1406–1433, 1931.
68. E. Bruce and A. C. Beck, Experiments with directivity steering for fading reduction, *Proc. IRE*, 23, 4, 357–371, 1935.
69. H. T. Friis, C. B. Feldman, and W. M. Sharpless, The determination of the direction of arrival of short radio waves, *Proc. IRE*, 22, 1, 47–78, 1934.
70. H. T. Friis and C. B. Feldman, A multiple unit steerable antenna for short-wave reception, *Proc. IRE*, 25, 7, 841–917, 1937.
71. D. Gregory and S. Newsome, Cooling radio station, Hoo Peninsula, Kent: An archaeological investigation of a short-wave receiving station, English Heritage Research Department, Portsmouth, UK, Technical Report 110-2010, 2010.
72. J. H. Vogelman, J. L. Ryerson, and M. H. Bickelhaupt, Tropospheric scatter system using angle diversity, *Proc. IRE*, 47, 5, 688–696, 1959.
73. E. N. Gilbert, Energy reception for mobile radio, *Bell Sys. Tech. J.*, 44, 8, 1779–1803, 1965.
74. A. M. Sayeed and B. Aazhang, Joint multipath-doppler diversity in mobile wireless communications, *IEEE Trans. Commun.*, 47, 1, 123–132, 1999.
75. R. Knopp and P. A. Humblet, Information capacity and power control in single-cell multiuser communications, In *Proceedings on IEEE International Conference on Communications*, Seattle, WA, June 1995.
76. A. Sendonaris, E. Erkip, and B. Aazhang, Increasing uplink capacity via user cooperation diversity, In *Proceedings on IEEE International Symposium on Information Theory*, Cambridge, MA, August 1998.
77. J. N. Laneman, G. W. Wornell, and D. N. C. Tse, An efficient protocol for realizing cooperative diversity in wireless networks, In *Proceedings on IEEE International Symposium on Information Theory*, Washington, DC, June 2001.
78. J. N. Laneman and G. W. Wornell, Exploiting distributed spatial diversity in wireless networks, In *Proceedings on Allerton Conference on Communication, Control, and Computing*, Urbana-Champagne, IL, October 2000.
79. J. Boyer, D. Falconer, and H. Yanikomeroglu, A theoretical characterization of the multihop wireless communications channel with diversity, In *Proceedings on IEEE Global Communications Conference*, San Antonio, TX, November 2001.
80. J. Boutros and E. Viterbo, Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel, *IEEE Trans. Inf. Theory*, 44, 4, 1453–1467, 1998.
81. G. D. Callaghan and I. D. Longstaff, Wide-swath space-borne SAR and range ambiguity, In *Proceedings on Radar 97 (Conference Publication No 449)*, Edinburgh, UK, October 1997.
82. A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad, Signaling strategies for dual-function radar communications: An overview, *IEEE Aerosp. Electron. Syst. Mag.*, 31, 10, 36–45, 2016.
83. The Electrical World and Engineer, vol. XL, no. 20, Nov. 15, 1902.
84. The Electrical World and Engineer, vol. XL, no. 23, Dec. 6, 1902.