Appendix A
Signal Characteristics

A.1 Bandpass Signals

A bandpass signal has its power spectrum in a spectral band surrounding a carrier frequency, which is usually at the center of the band. The Hilbert transform provides the basis for signal representations that facilitate the analysis of bandpass signals and systems. The Hilbert transform of a real-valued function $g(t)$ is

$$H[g(t)] = \hat{g}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(u)}{t-u} \, du.$$  (A.1)

Since its integrand has a singularity, the integral is defined as its Cauchy principal value:

$$\int_{-\infty}^{\infty} \frac{g(u)}{t-u} \, du = \lim_{\epsilon \to 0} \left[ \int_{-\infty}^{t-\epsilon} \frac{g(u)}{t-u} \, du + \int_{t+\epsilon}^{\infty} \frac{g(u)}{t-u} \, du \right]$$  (A.2)

provided that the limit exists. Since (A.1) has the form of the convolution of $g(t)$ with $1/\pi t$, $\hat{g}(t)$ results from passing $g(t)$ through a linear filter with an impulse response equal to $1/\pi t$. The transfer function of the filter is given by the Fourier transform

$$F \left\{ \frac{1}{\pi t} \right\} = \int_{-\infty}^{\infty} \frac{\exp(-j 2 \pi ft)}{\pi t} \, dt$$  (A.3)

where $j = \sqrt{-1}$. This integral can be rigorously evaluated by using contour integration. Alternatively, we observe that since $1/t$ is an odd function,

$$F \left\{ \frac{1}{\pi t} \right\} = -2j \int_{0}^{\infty} \frac{\sin 2 \pi ft}{\pi t} \, dt$$

$$= -j \text{ sgn}(f)$$  (A.4)
where \( \text{sgn}(f) \) is the *signum function* defined by

\[
\text{sgn}(f) = \begin{cases} 
1, & f > 0 \\
0, & f = 0 \\
-1, & f < 0.
\end{cases}
\] (A.5)

Let \( G(f) = \mathcal{F}\{g(t)\} \), and let \( \hat{G}(f) = \mathcal{F}\{\hat{g}(t)\} \). Equations (A.1) and (A.4) and the convolution theorem imply that

\[
\hat{G}(f) = -j \text{sgn}(f)G(f).
\] (A.6)

Because \( H[\hat{g}(t)] \) results from passing \( g(t) \) through two successive filters, each with transfer function \(-j \text{sgn}(f)\),

\[
H[\hat{g}(t)] = -g(t)
\] (A.7)

provided that \( G(0) = 0 \).

Equation (A.6) indicates that taking the Hilbert transform corresponds to introducing a phase shift of \(-\pi\) radians for all positive frequencies and \(+\pi\) radians for all negative frequencies. Consequently,

\[
H[\cos 2\pi f_c t] = \sin 2\pi f_c t \\
H[\sin 2\pi f_c t] = -\cos 2\pi f_c t.
\] (A.8, A.9)

These relations can be formally verified by taking the Fourier transform of the left-hand side of (A.8) or (A.9), applying (A.6), and then taking the inverse Fourier transform of the result. If \( G(f) = 0 \) for \(|f| > W \) and \( f_c > W \), the same method yields

\[
H[g(t) \cos 2\pi f_c t] = g(t) \sin 2\pi f_c t \\
H[g(t) \sin 2\pi f_c t] = -g(t) \cos 2\pi f_c t.
\] (A.10, A.11)

A *bandpass signal* is one with a Fourier transform that is negligible except for \( f_c - W/2 \leq |f| \leq f_c + W/2 \), where \( 0 \leq W < 2f_c \) and \( f_c \) is the center frequency. If \( W \ll f_c \), the bandpass signal is often called a *narrowband signal*. A complex-valued signal with a Fourier transform that is nonzero only for \( f > 0 \) is called an *analytic signal*.

Consider a bandpass signal \( g(t) \) with Fourier transform \( G(f) \). The analytic signal \( g_a(t) \) associated with \( g(t) \) is defined to be the signal with Fourier transform

\[
G_a(f) = [1 + \text{sgn}(f)]G(f)
\] (A.12)

which is zero for \( f \leq 0 \) and is confined to the band \(|f - f_c| \leq W/2 \) when \( f > 0 \). The inverse Fourier transform of (A.12) and (A.6) imply that

\[
g_a(t) = g(t) + j \hat{g}(t).
\] (A.13)
The complex envelope of \( g(t) \) is defined by

\[
g_l(t) = g_a(t) \exp[-j2\pi f_c t]
\]  

(A.14)

where \( f_c \) is the center frequency if \( g(t) \) is a bandpass signal. Since the Fourier transform of \( g_l(t) \) is \( \hat{G}_a(f + f_c) \), which occupies the band \( |f| \leq W/2 \), the complex envelope is a baseband signal that may be regarded as an equivalent lowpass representation of \( g(t) \). Equations (A.13) and (A.14) imply that \( g(t) \) may be expressed in terms of its complex envelope as

\[
g(t) = \text{Re}[g_l(t) \exp(j2\pi f_c t)].
\]  

(A.15)

The complex envelope can be decomposed as

\[
g_l(t) = g_c(t) + jg_s(t)
\]  

(A.16)

where \( g_c(t) \) and \( g_s(t) \) are real-valued functions. Therefore, (A.15) yields

\[
g(t) = g_c(t) \cos(2\pi f_c t) - g_s(t) \sin(2\pi f_c t).
\]  

(A.17)

Since the two sinusoidal carriers are in phase quadrature, \( g_c(t) \) and \( g_s(t) \) are called the in-phase and quadrature components of \( g(t) \), respectively. These components are lowpass signals confined to \( |f| \leq W/2 \).

Applying Parseval’s identity from Fourier analysis and then (A.6), we obtain

\[
\int_{-\infty}^{\infty} \hat{g}^2(t) \, dt = \int_{-\infty}^{\infty} |\hat{G}(f)|^2 \, df = \int_{-\infty}^{\infty} |G(f)|^2 \, df = \int_{-\infty}^{\infty} g^2(t) \, dt. \tag{A.18}
\]

Therefore,

\[
\int_{-\infty}^{\infty} |g_l(t)|^2 \, dt = \int_{-\infty}^{\infty} |g_a(t)|^2 \, dt = \int_{-\infty}^{\infty} g^2(t) \, dt + \int_{-\infty}^{\infty} \hat{g}^2(t) \, dt
\]

\[= 2 \int_{-\infty}^{\infty} g^2(t) \, dt = 2\mathcal{E} \tag{A.19}
\]

where \( \mathcal{E} \) denotes the energy of the bandpass signal \( g(t) \).

**A.2 Stationary Stochastic Processes**

Consider a stochastic process \( n(t) \) that is a zero-mean, wide-sense stationary process with autocorrelation

\[
R_n(\tau) = E[n(t)n(t + \tau)] \tag{A.20}
\]
where $E[x]$ denotes the expected value of $x$. The Hilbert transform of this process is the stochastic process defined by

$$\hat{n}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{n(u)}{t-u} du \quad (A.21)$$

where it is assumed that the Cauchy principal value of the integral exists for almost every sample function of $n(t)$. This equation indicates that $\hat{n}(t)$ is a zero-mean stochastic process. The zero-mean processes $n(t)$ and $\hat{n}(t)$ are jointly wide-sense stationary if their correlation and cross-correlation functions are not functions of $t$. A straightforward calculation using (A.21) and (A.20) gives the cross-correlation

$$R_{n\hat{n}}(\tau) = E[n(t)\hat{n}(t+\tau)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R_n(u)}{\tau-u} du = \hat{R}_n(\tau). \quad (A.22)$$

A similar derivation using (A.7) yields the autocorrelation

$$R_{\hat{n}}(\tau) = E[\hat{n}(t)\hat{n}(t+\tau)] = R_n(\tau). \quad (A.23)$$

Equations (A.20), (A.22), and (A.23) indicate that $n(t)$ and $\hat{n}(t)$ are jointly wide-sense stationary.

The analytic signal associated with $n(t)$ is the zero-mean process defined by

$$n_a(t) = n(t) + j\hat{n}(t). \quad (A.24)$$

The autocorrelation of the analytic signal is defined as

$$R_a(\tau) = E[n_a^*(t)n_a(t+\tau)] \quad (A.25)$$

where the asterisk denotes the complex conjugate. Using (A.20) and (A.22) to (A.25), we obtain

$$R_a(\tau) = 2R_n(\tau) + 2j\hat{R}_n(\tau) \quad (A.26)$$

which establishes the wide-sense stationarity of the analytic signal.

Since (A.20) indicates that $R_n(\tau)$ is an even function, (A.22) yields

$$R_{n\hat{n}}(0) = \hat{R}_n(0) = 0 \quad (A.27)$$

which indicates that $n(t)$ and $\hat{n}(t)$ are uncorrelated. Equations (A.23), (A.26), and (A.27) yield

$$R_{\hat{n}}(0) = R_n(0) = 1/2R_a(0). \quad (A.28)$$

The complex envelope of $n(t)$ or the equivalent lowpass representation of $n(t)$ is the zero-mean stochastic process defined by

$$n_l(t) = n_a(t) \exp(-j2\pi f_c t) \quad (A.29)$$
where \( f_c \) is an arbitrary frequency usually chosen as the center or carrier frequency of \( n(t) \). The complex envelope can be decomposed as

\[
n(t) = n_c(t) + j n_s(t)
\]  
\[
(A.30)
\]

where \( n_c(t) \) and \( n_s(t) \) are real-valued, zero-mean stochastic processes.

Equations (A.29) and (A.30) imply that

\[
n(t) = Re\{n(t) \exp(j2\pi f_c t)\}
\]

\[
= n_c(t) \cos(2\pi f_c t) - n_s(t) \sin(2\pi f_c t).
\]  
\[
(A.31)
\]

Substituting (A.24) and (A.30) into (A.29) we find that

\[
n_c(t) = n(t) \cos(2\pi f_c t) + \hat{n}(t) \sin(2\pi f_c t)
\]  
\[
(A.32)
\]

\[
n_s(t) = \hat{n}(t) \cos(2\pi f_c t) - n(t) \sin(2\pi f_c t).
\]  
\[
(A.33)
\]

The autocorrelations of \( n_c(t) \) and \( n_s(t) \) are defined by

\[
R_c(\tau) = E[n_c(t)n_c(t + \tau)]
\]  
\[
(A.34)
\]

and

\[
R_s(\tau) = E[n_s(t)n_s(t + \tau)].
\]  
\[
(A.35)
\]

Using (A.32) and (A.33) and then (A.20), (A.22), and (A.23) and trigonometric identities, we obtain

\[
R_c(\tau) = R_s(\tau) = R_n(\tau) \cos(2\pi f_c \tau) + \hat{R}_n(\tau) \sin(2\pi f_c \tau)
\]  
\[
(A.36)
\]

which shows explicitly that if \( n(t) \) is wide-sense stationary, then \( n_c(t) \) and \( n_s(t) \) are wide-sense stationary with the same autocorrelation function. The variances of \( n(t) \), \( n_c(t) \), and \( n_s(t) \) are all equal because

\[
R_c(0) = R_s(0) = R_n(0).
\]  
\[
(A.37)
\]

A derivation similar to that of (A.36) gives the cross-correlation

\[
R_{cs}(\tau) = E[n_c(t)n_s(t + \tau)] = \hat{R}_n(\tau) \cos(2\pi f_c \tau) - R_n(\tau) \sin(2\pi f_c \tau).
\]  
\[
(A.38)
\]

Equations (A.36) and (A.38) indicate that \( n_c(t) \) and \( n_s(t) \) are jointly wide-sense stationary, which then implies that

\[
R_{sc}(\tau) = E[n_s(t)n_c(t + \tau)] = R_{cs}(-\tau).
\]  
\[
(A.39)
\]
Equations (A.27) and (A.38) give

\[ R_{cs}(0) = 0 \]  \hspace{1cm} (A.40)

which implies that \( n_c(t) \) and \( n_s(t) \) are uncorrelated.

Since \( n(t) \) is wide-sense stationary, \( R_n(-\tau) \equiv R_n(\tau) \). It then follows from (A.17) and a change of the integration variable that \( \hat{R}_n(-\tau) = -\hat{R}_n(\tau) \). Combining these equations with (A.38) yields \( R_{cs}(-\tau) = -R_{cs}(\tau) \). This equation and (A.39) indicate that

\[ R_{cs}(\tau) = -R_{sc}(\tau). \]  \hspace{1cm} (A.41)

Equations (A.30), (A.37), and (A.41) imply that

\[ E[n_l(t)n_l(t + \tau)] = 0. \]  \hspace{1cm} (A.42)

A complex-valued, zero-mean stochastic process that satisfies this equation is called a circularly symmetric process. Thus, the complex envelope of a zero-mean, wide-sense stationary process is a circularly symmetric process.

Equation (A.21) indicates that \( \hat{n}(t) \) is generated by a linear operation on \( n(t) \). Therefore, if \( n(t) \) is a zero-mean Gaussian process, \( \hat{n}(t) \) and \( n(t) \) are zero-mean jointly Gaussian processes. Equations (A.32) and (A.33) then imply that \( n_c(t) \) and \( n_s(t) \) are zero-mean jointly Gaussian processes. Since they are uncorrelated, \( n_c(t) \) and \( n_s(t) \) are statistically independent, zero-mean Gaussian processes.

The power spectral density of a signal is the Fourier transform of its autocorrelation. Let \( S(f), S_c(f), \) and \( S_s(f) \) denote the power spectral densities of \( n(t), n_c(t), \) and \( n_s(t) \), respectively. We assume that \( S_n(f) \) occupies the band \( f_c - W/2 \leq |f| \leq f_c + W/2 \) and that \( f_c > W/2 \geq 0 \). Taking the Fourier transform of (A.36), using (A.6), and simplifying, we obtain

\[ S_c(f) = S_s(f) = \begin{cases} S_n(f - f_c) + S_n(f + f_c), & |f| \leq W/2 \\ 0, & |f| > W/2. \end{cases} \]  \hspace{1cm} (A.43)

Thus, if \( n(t) \) is a passband process with bandwidth \( W \) of the positive frequencies, then \( n_c(t) \) and \( n_s(t) \) are baseband processes with bandwidths \( W/2 \). This property and the statistical independence of \( n_c(t) \) and \( n_s(t) \) when \( n(t) \) is Gaussian make (A.31) a very useful representation of \( n(t) \).

Similarly, the cross-spectral density of \( n_c(t) \) and \( n_s(t) \) can be derived by taking the Fourier transform of (A.38) and using (A.6). After simplification, the result is

\[ S_{cs}(f) = \begin{cases} j[S_n(f - f_c) - S_n(f + f_c)], & |f| \leq W/2 \\ 0, & |f| > W/2. \end{cases} \]  \hspace{1cm} (A.44)

If \( S_n(f) \) is locally symmetric about \( f_c \), then

\[ S_n(f_c + f) = S_n(f_c - f), \quad |f| \leq W/2. \]  \hspace{1cm} (A.45)
Since a power spectral density is a real-valued, even function, $S_n(f_c - f) = S_n(f - f_c)$. Equation (A.45) then yields $S_n(f + f_c) = S_n(f - f_c)$ for $|f| \leq W/2$. Therefore, (A.44) gives $S_{cs}(f) = 0$, which implies that

$$R_{cs}(\tau) = 0$$

(A.46)

for all $\tau$. Thus, $n_c(t)$ and $n_s(t + \tau)$ are uncorrelated for all $\tau$, and if $n(t)$ is a zero-mean Gaussian process, then $n_c(t)$ and $n_s(t + \tau)$ are statistically independent for all $\tau$.

The autocorrelation of the complex envelope is defined by

$$R_l(\tau) = E[n^*_l(t)n_l(t + \tau)].$$

(A.47)

Equations (A.29) and (A.26) imply that the complex envelope of a zero-mean, wide-sense stationary process is wide-sense stationary. Equations (A.28) and (A.29) yield

$$R_l(0) = 2R_n(0).$$

(A.48)

Substituting (A.30) into (A.47) and using (A.36) and (A.38), we obtain

$$R_l(\tau) = 2R_c(\tau) + j2R_{cs}(\tau).$$

(A.49)

The power spectral density of $n_l(t)$, which we denote by $S_l(f)$, can be derived from (A.49), (A.44), and (A.43). If $S_n(f)$ occupies the band $f_c - W/2 \leq |f| \leq f_c + W/2$ and $f_c > W/2 \geq 0$, then

$$S_l(f) = \begin{cases} 4S_n(f + f_c), & |f| \leq W/2 \\ 0, & |f| > W/2. \end{cases}$$

(A.50)

Equations (A.36) and (A.38) yield

$$R_n(\tau) = 2R_c(\tau) \cos(2\pi f_c \tau) - 2R_{cs}(\tau) \sin(2\pi f_c \tau).$$

(A.51)

Equations (A.51) and (A.49) imply that

$$R_n(\tau) = Re \left[ R_l(\tau) \exp(j2\pi f_c \tau) \right].$$

(A.52)

We expand the right-hand side of this equation by using the fact that $Re[z] = (z + z^*)/2$. Taking the Fourier transform and observing that $S_l(f)$ is a real-valued function, we obtain

$$S_n(f) = \frac{1}{4}S_l(f - f_c) + \frac{1}{4}S_l(-f - f_c).$$

(A.53)
If $S_n(f)$ is locally symmetric about $f_c$, then (A.50) and (A.45) imply that $S_l(-f) = S_l(f)$, and (A.53) becomes

$$S_n(f) = \frac{1}{4}S_l(f - f_c) + \frac{1}{4}S_l(f + f_c). \quad \text{(A.54)}$$

Many communication signals are modeled as bandpass signals having the form

$$s(t) = A \text{Re}[s_l(t) \exp(j2\pi f_c t + \theta)] \quad \text{(A.55)}$$

where $A$ is the amplitude and $\theta$ is an independent random variable that is uniformly distributed over $0 \leq \theta < 2\pi$. Equation (A.15) indicates that the complex envelope of $s(t)$ is $As_l(t) \exp(j\theta)$. The power spectral density of the complex envelope is equal to $AS_l(f)$, where $S_l(f)$ is the power spectral density of $s_l(t)$. The power spectral density of $s(t)$ is calculated by applying (A.54) or (A.53).

### A.3 Direct-Conversion Receiver

Receivers often extract the complex envelope of the desired signal before applying it to a matched filter. The main components in a direct-conversion receiver are shown in Fig. A.1a. The spectra of the received signal $g(t)$, the input to the baseband filter $g(t) = g(t) \exp(-j2\pi f_c t)$, and the complex envelope $g_l(t)$ are depicted in Fig. A.1b. Let $h(t)$ denote the impulse response of the filter. The output of the filter is

$$y(t) = \int_{-\infty}^{\infty} 2g(\tau) \exp(-j2\pi f_c \tau) h(t - \tau) d\tau. \quad \text{(A.56)}$$

Using (A.15) and the fact that $\text{Re}(x) = (x + x^*)/2$, where $x^*$ denotes the complex conjugate of $x$, we obtain

$$y(t) = \int_{-\infty}^{\infty} g_l(\tau) h(t - \tau) d\tau + \int_{-\infty}^{\infty} g_l(\tau) h(t - \tau) \exp(-j4\pi f_c \tau) d\tau. \quad \text{(A.57)}$$

The second term is the Fourier transform of $g_l(\tau)h(t - \tau)$ evaluated at frequency $-2f_c$. Assuming that $g_l(\tau)$ and $h(t - \tau)$ have transforms confined to $|f| < f_c$, their product has a transform confined to $|f| < 2f_c$, and the second term in (A.57) vanishes. If the Fourier transform of $h(t)$ is a constant over the passband of $g_l(t)$, then (A.57) implies that $y(t)$ is proportional to $g_l(t)$, as desired. Figure A.1c shows the direct-conversion receiver for real-valued signals.

The direct-conversion receiver alters the character of the noise $n(t)$ entering it. Suppose that $n(t)$ is a zero-mean, white Gaussian noise process with autocorrelation

$$R_n(\tau) = E[n(t)n(t + \tau)] = \frac{N_0}{2} \delta(\tau) \quad \text{(A.58)}$$
where $\delta(\tau)$ denotes the Dirac delta function, and $N_0/2$ is the two-sided noise-power spectral density. The complex-valued noise at the output of Fig. A.1a is

$$z(t) = \int_{-\infty}^{\infty} 2n(u)e^{-j2\pi f_c u}h(t-u)du$$  \hspace{1cm} (A.59)

Since it is a linear function of $n(t)$, $z(t)$ is zero-mean and its real and imaginary parts are jointly Gaussian. The autocorrelation of a wide-sense stationary, complex-valued process $z(t)$ is defined as

$$R_z(\tau) = \frac{1}{2}E[z^*(t)z(t+\tau)].$$  \hspace{1cm} (A.60)
Substituting (A.59), interchanging the expectation and integration operations, using (A.58) to evaluate one of the integrals, and then changing variables, we obtain

\[ R_z(\tau) = N_0 \int_{-\infty}^{\infty} h(u)h^*(u + \tau) du. \]  

(A.61)

If the filter is an ideal bandpass filter with Fourier transform

\[ H(f) = \begin{cases} 1, & |f| \leq W \\ 0, & \text{otherwise} \end{cases} \]  

(A.62)

then evaluating the Fourier transform of both sides of (A.61) gives the power spectral density

\[ S_z(f) = \begin{cases} N_0, & |f| \leq W \\ 0, & \text{otherwise} \end{cases}. \]  

(A.63)

Thus, if the subsequent filters have narrower bandwidths than \( W \) or if \( W \to \infty \), then the autocorrelation of \( z(t) \) may be approximated by

\[ R_z(\tau) = N_0 \delta(\tau). \]  

(A.64)

This approximation permits major analytical simplifications. Equations (A.59) and (A.58) imply that

\[ E[z(t)z(t + \tau)] = 2N_0 e^{-j4\pi f_c t} \int_{-\infty}^{\infty} e^{j4\pi f_c u} h(u + \tau) h(u) du. \]  

(A.65)

If \( W < f_c \), then reasoning similar to that following (A.57) leads to

\[ E[z(t)z(t + \tau)] = 0 \]  

(A.66)

which indicates that the complex-valued stochastic process \( z(t) \) is a \textit{circularly symmetric} process. Let \( z^R(t) \) and \( z^I(t) \) denote the real and imaginary parts of \( z(t) \), respectively. Since \( z(t) \) is zero-mean, \( z^R(t) \) and \( z^I(t) \) are zero-mean. Setting \( \tau = 0 \) in (A.66) and (A.60), and then using (A.61), Parseval’s identity, and (A.62), we obtain

\[ E[(z^R(t))^2] = E[(z^I(t))^2] = 2N_0W \]  

(A.67)

\[ E[z^R(t)z^I(t)] = 0. \]  

(A.68)

Thus, \( z^R(t) \) and \( z^I(t) \) are zero-mean, independent Gaussian processes with the same variance.
Appendix B
Probability Distributions

B.1 Chi-Square Distribution

Consider the random variable

$$Z = \sum_{i=1}^{N} A_i^2$$  \hspace{1cm} (B.1)

where the \( \{A_i\} \) are independent Gaussian random variables with means \( \{m_i\} \) and common variance \( \sigma^2 \). The random variable \( Z \) is said to have a noncentral chi-square \( (\chi^2) \) distribution with \( N \) degrees of freedom and a noncentral parameter

$$\lambda = \sum_{i=1}^{N} m_i^2.$$  \hspace{1cm} (B.2)

To derive the probability density function of \( Z \), we first note that each \( A_i \) has the density function

$$f_{A_i}(x) = \frac{1}{\sqrt{2\pi \sigma}} \exp \left[ -\frac{(x-m_i)^2}{2\sigma^2} \right].$$  \hspace{1cm} (B.3)

From elementary probability, the density of \( Y_i = A_i^2 \) is

$$f_{Y_i}(x) = \frac{1}{2\sqrt{x}} [f_{A_i}(\sqrt{x}) + f_{A_i}(-\sqrt{x})] u(x)$$  \hspace{1cm} (B.4)

where \( u(x) = 1, x \geq 0 \), and \( u(x) = 0, x < 0 \). Substituting (B.3) into (B.4), expanding the exponentials, and simplifying, we obtain the density

$$f_{Y_i}(x) = \frac{1}{\sqrt{2\pi x\sigma}} \exp \left( -\frac{x+m_i^2}{2\sigma^2} \right) \cosh \left( \frac{m_i\sqrt{x}}{\sigma^2} \right) u(x).$$  \hspace{1cm} (B.5)
The characteristic function of a random variable $X$ is defined as

$$C_X(jv) = E[e^{jvX}] = \int_{-\infty}^{\infty} f_X(x) \exp(jv x) \, dx$$  \hspace{1cm} (B.6)$$

where $j = \sqrt{-1}$, and $f_X(x)$ is the density of $X$. Since $C_X(jv)$ is the conjugate Fourier transform of $f_X(x)$,

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C_X(jv) \exp(-jv x) \, dv$$  \hspace{1cm} (B.7)$$

From Laplace or Fourier transform tables, it is found that the characteristic function of $f_{Y_1}(x)$ is

$$C_{Y_1}(jv) = \frac{\exp[jm_1^2v/(1 - j2\sigma^2v)]}{(1 - j\sigma^2v)^{1/2}}.$$  \hspace{1cm} (B.8)$$

The characteristic function of a sum of independent random variables is equal to the product of the individual characteristic functions. Because $Z$ is the sum of the $Y_i$, the characteristic function of $Z$ is

$$C_Z(jv) = \frac{\exp[j\lambda v/(1 - j2\sigma^2v)]}{(1 - j\sigma^2v)^{N/2}}$$  \hspace{1cm} (B.9)$$

where we have used (B.2). From (B.9), (B.7), and Laplace or Fourier transform tables, we obtain the probability density function of noncentral $\chi^2$ random variable with $N$ degrees of freedom and a noncentral parameter $\lambda$:

$$f_Z(x) = \frac{1}{2\sigma^2} \left(\frac{x}{\lambda}\right)^{(N-2)/4} \exp\left[-\frac{x + \lambda}{2\sigma^2}\right] I_{N/2-1} \left(\frac{\sqrt{x\lambda}}{\sigma^2}\right) u(x)$$  \hspace{1cm} (B.10)$$

where $I_n(\ )$ is the modified Bessel function of the first kind and order $n$. This function may be represented by

$$I_n(x) = \sum_{i=0}^{\infty} \frac{(x/2)^{n+2i}}{i! \Gamma(n + i + 1)}$$  \hspace{1cm} (B.11)$$

where the gamma function is defined as

$$\Gamma(x) = \int_0^{\infty} y^{x-1} \exp(-y) \, dy, \quad \text{Re}(x) > 0.$$  \hspace{1cm} (B.12)$$

The probability distribution function of a noncentral $\chi^2$ random variable is

$$F_Z(x) = \int_0^x \frac{1}{2\sigma^2} \left(\frac{y}{\lambda}\right)^{(N-2)/4} \exp\left(-\frac{y + \lambda}{2\sigma^2}\right) I_{N/2-1} \left(\frac{\sqrt{y\lambda}}{\sigma^2}\right) \, dy, \quad x \geq 0.$$  \hspace{1cm} (B.13)$$
If $N$ is even so that $N/2$ is an integer, then $F_Z(\infty) = 1$ and a change of variables in (B.13) yield

$$F_Z(x) = 1 - Q_{N/2} \left( \frac{\sqrt{\lambda}}{\sigma}, \frac{\sqrt{x}}{\sigma} \right), \quad x \geq 0 \quad (B.14)$$

where the generalized Marcum $Q$-function is defined as

$$Q_m(\alpha, \beta) = \int_{\beta}^{\infty} \left( \frac{x}{\alpha} \right)^{m-1} \exp \left( -\frac{x^2 + \alpha^2}{2} \right) I_{m-1}(\alpha x) \, dx \quad (B.15)$$

and $m$ is an integer. Since $Q_m(\alpha, 0) = 1$, it follows that $1 - Q_m(\alpha, \beta)$ is an integral with finite limits that can be numerically integrated. The mean, variance, and moments of $Z$ can be easily obtained by using (B.1) and the properties of independent Gaussian random variables. The mean and variance of $Z$ are

$$E[Z] = N\sigma^2 + \lambda \quad (B.16)$$
$$\sigma_z^2 = 2N\sigma^4 + 4\lambda\sigma^2 \quad (B.17)$$

where $\sigma^2$ is the common variance of the $\{A_i\}$.

From (B.9), it follows that the sum of two independent noncentral $\chi^2$ random variables with $N_1$ and $N_2$ degrees of freedom, noncentral parameters $\lambda_1$ and $\lambda_2$, respectively, and the same parameter $\sigma^2$ is a noncentral $\chi^2$ random variable with $N_1 + N_2$ degrees of freedom and noncentral parameter $\lambda_1 + \lambda_2$.

### B.2 Central Chi-Square Distribution

To determine the probability density function of $Z$ when the $\{A_i\}$ have zero means, we substitute (B.11) into (B.10) and then take the limit as $\lambda \to 0$. We obtain

$$f_Z(x) = \frac{1}{(2\sigma^2)^{N/2}\Gamma(N/2)} x^{N/2-1} \exp \left( -\frac{x^2}{2\sigma^2} \right) u(x). \quad (B.18)$$

Alternatively, this equation results if we substitute $\lambda = 0$ into the characteristic function (B.9) and then use (B.7). Equation (B.18) is the probability density function of a central $\chi^2$ random variable with $N$ degrees of freedom. The probability distribution function is

$$F_Z(x) = \int_{0}^{x} \frac{1}{(2\sigma^2)^{N/2}\Gamma(N/2)} y^{N/2-1} \exp \left( -\frac{y^2}{2\sigma^2} \right) dy, \quad x \geq 0. \quad (B.19)$$

If $N$ is even so that $N/2$ is an integer, then integrating this equation by parts $N/2 - 1$ times yields
\[ F_Z(x) = 1 - \exp \left( -\frac{x}{2\sigma^2} \right) \sum_{i=0}^{N/2-1} \frac{1}{i!} \left( \frac{x}{2\sigma^2} \right)^i, \quad x \geq 0. \]  

(B.20)

By direct integration using (B.18) and (B.12) or from (B.16) and (B.17), it is found that the mean and variance of \( Z \) are

\[ E[Z] = N\sigma^2 \]  

(B.21)

\[ \sigma_z^2 = 2N\sigma^4. \]  

(B.22)

### B.3 Rice Distribution

Consider the random variable

\[ R = \sqrt{A_1^2 + A_2^2} \]  

(B.23)

where \( A_1 \) and \( A_2 \) are independent Gaussian random variables with means \( m_1 \) and \( m_2 \), respectively, and a common variance \( \sigma^2 \). The probability distribution function of \( R \) must satisfy \( F_R(r) = F_Z(r^2) \), where \( Z = A_1^2 + A_2^2 \) is a \( \chi^2 \) random variable with two degrees of freedom. Therefore, (B.14) with \( N = 2 \) implies that

\[ F_R(r) = 1 - Q_1 \left( \frac{\sqrt{\lambda}}{\sigma}, \frac{r}{\sigma} \right), \quad r \geq 0 \]  

(B.24)

where \( \lambda = m_1^2 + m_2^2 \). This function is called the Rice probability distribution function. The Rice probability density function, which may be obtained by differentiation of (B.24), is

\[ f_R(r) = \frac{r}{\sigma^2} \exp \left( -\frac{r^2 + \lambda}{2\sigma^2} \right) I_0 \left( \frac{r\sqrt{\lambda}}{\sigma^2} \right) u(r). \]  

(B.25)

The moments of even order can be derived from (B.23) and the moments of the independent Gaussian random variables. The second moment is

\[ E[R^2] = 2\sigma^2 + \lambda. \]  

(B.26)

In general, moments of the Rice distribution are given by an integration over the density in (B.25). Substituting (B.11) into the integrand, interchanging the summation and integration, changing the integration variable, and using (B.12), we obtain a series that is recognized as a special case of the confluent hypergeometric function. Thus,

\[ E[R^n] = (2\sigma^2)^{n/2} \exp \left( -\frac{\lambda}{2\sigma^2} \right) \Gamma \left( 1 + \frac{n}{2} \right) \, {}_1F_1 \left( 1 + \frac{n}{2}, 1; \frac{\lambda}{2\sigma^2} \right), \quad n \geq 0 \]  

(B.27)
where the confluent hypergeometric function is defined as

$$
{1 \choose 0} F_1(\alpha, \beta; x) = \sum_{i=0}^{\infty} \frac{\Gamma(\alpha + i) \Gamma(\beta) x^i}{\Gamma(\alpha) \Gamma(\beta + i) i!}, \quad \beta \neq 0, -1, -2, \ldots.
$$ (B.28)

The Rice density function often arises in the context of a transformation of variables. Let $A_1$ and $A_2$ represent independent Gaussian random variables with common variance $\sigma^2$ and means $\lambda$ and zero, respectively. Let $R$ and $\Theta$ be implicitly defined by $A_1 = R \cos \Theta$ and $A_2 = R \sin \Theta$. Then (B.23) and $\Theta = \tan^{-1}(A_2/A_2)$ describes a transformation of variables. A straightforward calculation yields the joint density function of $R$ and $\Theta$:

$$
f_{R,\Theta}(r, \theta) = \frac{r}{2\pi \sigma^2} \exp \left( -\frac{r^2 - 2r \lambda \cos \theta + \lambda^2}{2\sigma^2} \right), \quad r \geq 0, \ |\theta| \leq \pi.
$$ (B.29)

The density function of the envelope $R$ is obtained by integration over $\Theta$. Since the modified Bessel function of the first kind and order zero satisfies

$$
I_0(x) = \frac{1}{2\pi} \int_{0}^{2\pi} \exp(x \cos u) \, du
$$ (B.30)

this density function reduces to the Rice density function (B.25). The density function of the angle $\Theta$ is obtained by integrating (B.29) over $r$. Completing the square of the argument in (B.29), changing variables, and defining the Q-function

$$
Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left( -\frac{y^2}{2} \right) dy = \frac{1}{2} \text{erfc} \left( \frac{x}{\sqrt{2}} \right)
$$ (B.31)

where erfc( ) is the complementary error function, we obtain

$$
f_{\Theta}(\theta) = \frac{1}{2\pi} \exp \left( -\frac{\lambda^2}{2\sigma^2} \right) + \frac{\lambda \cos \theta}{\sqrt{2\pi} \sigma} \exp \left( -\frac{\lambda^2 \sin^2 \theta}{2\sigma^2} \right) \left[ 1 - Q \left( \frac{\lambda \cos \theta}{\sigma} \right) \right],
$$ (B.32)

$|\Theta| \leq \pi$.

Since (B.29) cannot be written as the product of (B.25) and (B.32), the random variables $R$ and $\Theta$ are not independent.

Since the density function of (B.25) must integrate to unity, we find that

$$
\int_{0}^{\infty} r \exp \left( -\frac{r^2}{2b} \right) I_0 \left( r \sqrt{\frac{\lambda}{b}} \right) \, dr = b \exp \left( \frac{\lambda}{2b} \right)
$$ (B.33)

where $\lambda$ and $b$ are positive constants. This equation is useful in calculations involving the Rice density function.
B.4 Rayleigh Distribution

A Rayleigh-distributed random variable is defined by (B.23) when $A_1$ and $A_2$ are independent Gaussian random variables with zero means and a common variance $\sigma^2$. Since $F_R(r) = F_Z(r^2)$, where $Z$ is a central $\chi^2$ random variable with two degrees of freedom, (B.20) with $N = 2$ implies that the Rayleigh probability distribution function is

$$F_R(r) = 1 - \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0. \quad \text{(B.34)}$$

The Rayleigh probability density function, which may be obtained by differentiation of (B.34), is

$$f_R(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) u(r). \quad \text{(B.35)}$$

By a change of variables in the defining integral, any moment of $R$ can be expressed in terms of the gamma function defined in (B.12). Therefore,

$$E[R^n] = (2\sigma^2)^{n/2}\Gamma\left(1 + \frac{n}{2}\right). \quad \text{(B.36)}$$

Certain properties of the gamma function are needed to simplify (B.36). An integration by parts of (B.12) indicates that $\Gamma(1 + x) = x\Gamma(x)$. A direct integration yields $\Gamma(1) = 1$. Therefore, when $n$ is an integer, $\Gamma(n) = (n - 1)!$. Changing the integration variable by substituting $y = z^2$ in (B.12), it is found that $\Gamma(1/2) = \sqrt{\pi}$.

Using these properties of the gamma function, we obtain the mean and the variance of a Rayleigh-distributed random variable:

$$E[R] = \sqrt{\frac{\pi}{2}} \sigma \quad \text{(B.37)}$$

$$\sigma_R^2 = \left(2 - \frac{\pi}{2}\right) \sigma^2. \quad \text{(B.38)}$$

Since $A_1$ and $A_2$ have zero means, the joint probability density function of the random variables $R = \sqrt{A_1^2 + A_2^2}$ and $\Theta = \tan^{-1}(A_2/A_1)$ is given by (B.29) with $\lambda = 0$. Therefore,

$$f_{R,\Theta}(r, \Theta) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0, \quad |\Theta| \leq \pi. \quad \text{(B.39)}$$

Integration over $\Theta$ yields (A.35), and integration over $r$ yields the uniform probability density function:

$$f_\theta(\theta) = \frac{1}{2\pi}, \quad |\theta| \leq \pi. \quad \text{(B.40)}$$
Since (B.39) equals the product of (B.35) and (B.40), the random variables $R$ and $\Theta$ are independent. In terms of these random variables, $A_1 = R \cos \Theta$ and $A_2 = R \sin \Theta$. A straightforward calculation using the independence and densities of $R$ and $\Theta$ verifies that $A_1$ and $A_2$ are zero-mean, independent, Gaussian random variables with common variance $\sigma^2$. Since the square of a Rayleigh-distributed random variable may be expressed as $R^2 = A_1^2 + A_2^2$, where $A_1$ and $A_2$ are zero-mean, independent, Gaussian random variables with common variance $\sigma^2$, $R^2$ has the distribution of a central chi-square random variable with two degrees of freedom. Therefore, (B.18) with $N = 2$ indicates that the square of a Rayleigh-distributed random variable has an exponential probability density function with mean $2\sigma^2$.

**B.5 Exponentially Distributed Random Variables**

Both the square of a Rayleigh-distributed random variable and a central chi-square random variable with two degrees of freedom have exponential probability distribution and density functions. Consider the random variable

$$Z = \sum_{i=1}^{N} Y_i$$  \hspace{1cm} (B.41)

where the $\{Y_i\}$ are independent, exponentially distributed random variables with unequal positive means $\{m_i\}$. The exponential probability density function of $Y_i$ is

$$f_{Y_i}(x) = \frac{1}{m_i} \exp \left( -\frac{x}{m_i} \right) u(x).$$  \hspace{1cm} (B.42)

A straightforward calculation yields the characteristic function

$$C_{Y_i}(jv) = \frac{1}{1 - jvm_i}.$$  \hspace{1cm} (B.43)

Since $Z$ is the sum of independent random variables, (B.43) implies that its characteristic function is

$$C_Z(jv) = \prod_{i=1}^{N} \frac{1}{1 - jvm_i}.$$  \hspace{1cm} (B.44)

To derive the probability density function of $Z$, (B.7) is applied after first expanding the right-hand side of (B.44) in a partial-fraction expansion. The result is

$$f_Z(x) = \sum_{i=1}^{N} \frac{B_i}{m_i} \exp \left( -\frac{x}{m_i} \right) u(x)$$  \hspace{1cm} (B.45)
where
\[
B_i = \begin{cases} 
\prod_{k=1, k \neq i}^{N} \frac{m_i}{m_i - m_k}, & N \geq 2 \\
1, & N = 1 
\end{cases} \quad (B.46)
\]
and \( m_i \neq m_k, i \neq k \). A direct integration and algebra yields the probability distribution function
\[
F_Z(x) = 1 - \sum_{i=1}^{N} B_i \exp\left(-\frac{x}{m_i}\right), \quad x \geq 0. \quad (B.47)
\]
Equations (B.45) and (B.12) give
\[
E[Z^n] = \Gamma(1 + n) \sum_{i=1}^{N} B_i m_i^n, \quad n \geq 0. \quad (B.48)
\]
When the \( \{m_i\} \) are equal so that \( m_i = m, 1 \leq i \leq N \), then \( C_Z(jv) = (1 - jvm)^{-N} \). Therefore, the probability density function of \( Z \) is
\[
f_Z(x) = \frac{1}{(N-1)!m^N} x^{N-1} \exp\left(-\frac{x}{m}\right) u(x) \quad (B.49)
\]
which is a special case of the gamma density function. Successive integration by parts yields the probability distribution function
\[
F_Z(x) = 1 - \exp\left(-\frac{x}{m}\right) \sum_{i=0}^{N-1} \frac{1}{i!} \left(\frac{x}{m}\right)^i. \quad (B.50)
\]
From (B.49) and (B.12), the mean and variance of \( Z \) are found to be
\[
E[Z] = Nm \quad (B.51)
\]
\[
\sigma_Z^2 = Nm^2. \quad (B.52)
\]
Appendix C
Convergence of Adaptive Algorithms

C.1 LMS Algorithm

C.1.1 Convergence of the Mean

The least-mean-square (LMS) algorithm computes the $N \times 1$ weight vector at iteration $n$ as (Sects. 2.7 and 6.5)

$$W(n + 1) = W(n) + 2\mu \epsilon^*(n) x(n), \ n = 0, 1, \ldots$$ (C.1)

where $x(n)$ is a zero-mean $N \times 1$ input vector,

$$\epsilon(n) = d(n) - y(n) = d(n) - W^H(n)x(n)$$ (C.2)

is the estimation error, $d(n)$ is the desired response, and

$$y(n) = W^H(n)x(n)$$ (C.3)

is the filter output. The adaptation constant $\mu$ controls the rate of convergence of the algorithm. We prove convergence of the mean weight vector under the assumption that the input vectors $\{x(n)\}$ are statistically independent, stationary random vectors. The assumption is valid at least when the sampling times of the components of $x(n + 1)$ are separated in time from those of $x(n)$ by intervals that are large compared to the correlation time of the input process. The Wiener–Hopf equation for the optimal weight vector, which minimizes the mean-square estimation error, is

$$W_0 = R_{xx}^{-1}R_{xd}$$ (C.4)

where

$$R_{xx} = E[x(n)x^H(n)]$$ (C.5)
is the $N \times N$ Hermitian correlation matrix of $x(n)$ and
\[ R_{xd} = E \left[ x(n) d^*(n) \right] \]  
(C.6)
is the $N \times 1$ cross-correlation vector. If we assume that $E[|y(n)|^2] \neq 0$ when $W(n) \neq 0$, then $R_{xx}$ must be positive definite. If $W(n) = W_0$, then (C.2) indicates that the minimum mean-square estimation error is
\[ \epsilon_m^2 = E[|d(n)|^2] - R_{xd}^H R_{xx}^{-1} R_{xd}. \]  
(C.7)

If $x(n + 1)$ is independent of $x(k)$ and $d(k)$, $k \leq n$, (C.1) implies that $W(n)$ is independent of $x(n)$. Thus, the expected value of the weight vector satisfies
\[ E[W(n + 1)] = (1 - 2\mu R_{xx}) E[W(n)] + 2\mu R_{xd}. \]  
(C.8)
This discrete-time equation is linear and time invariant. Its equilibrium point is easily calculated to be $W_0$. From (C.4) and (C.8), it follows that
\[ E[W(n + 1)] - W_0 = (I - 2\mu R_{xx})^{n+1} [W(0) - W_0]. \]  
(C.9)
With an initial weight vector $W(0)$, this equation implies that
\[ E[W(n + 1)] - W_0 = (I - 2\mu R_{xx})^{n+1} [W(0) - W_0] \]  
(C.10)
where $W(0)$ might be the identity matrix or an estimate of $W_0$. Since $R_{xx}$ is Hermitian and positive definite, it can be represented as
\[ R_{xx} = QAQ^{-1} = QAQ^H \]  
(C.11)
where $Q$ is the unitary modal matrix of $R_{xx}$ with eigenvectors as its columns, and $\Lambda$ is the diagonal matrix of eigenvalues of $R_{xx}$. Therefore, (C.10) can be expressed as
\[ E[W(k + 1)] - W_0 = [I - 2\mu QAQ^{-1}]^{n+1} [W(0) - W_0] \]
\[ = Q[I - 2\mu \Lambda]^{n+1} Q^{-1} [W(0) - W_0]. \]  
(C.12)

This equation indicates that
\[ \lim_{n \to \infty} [I - 2\mu \Lambda] = 0 \]  
(C.13)
is necessary and sufficient for the weight vector to converge to its optimal value:
\[ \lim_{n \to \infty} E[W(n)] = W_0 = R_{xx}^{-1} R_{xd}. \]  
(C.14)
A necessary and sufficient condition for (C.13), and hence (C.14), is that the diagonal elements of the diagonal matrix \( [I - 2\mu A] \) have magnitudes less than unity. Since \( R_{xx} \) is Hermitian and positive definite, its eigenvalues, \( \lambda_1, \lambda_2, ..., \lambda_n \), are positive. Therefore, the diagonal elements of \( I - 2\mu \Lambda \), which are \( 1 - 2\mu \lambda_1, 1 - 2\mu \lambda_2, ..., 1 - 2\mu \lambda_N \), have magnitudes less than unity if and only if

\[
|1 - 2\mu \lambda_{\text{max}}| < 1
\]

where \( \lambda_{\text{max}} \) is the maximum eigenvalue of \( R_{xx} \). This equation yields the necessary and sufficient convergence condition:

\[
0 < \mu < \frac{1}{\lambda_{\text{max}}}
\]

Although stronger convergence results can be proved if the inputs are stationary processes and \( \mu \) is allowed to decrease with the iteration number, making \( \mu \) constant gives the adaptive system flexibility in processing nonstationary inputs.

The matrix multiplications in (C.12) indicate that during adaptation the weights undergo transients that vary as sums of terms of the form \( (1 - 2\mu \lambda_i)^k \). These transients determine the rate of convergence of the mean vector. The time constants of the convergence are defined so that

\[
|1 - 2\mu \lambda_i|^n = \exp \left( -\frac{n}{\tau_i} \right), \quad i = 1, 2, ..., N
\]

which yields

\[
\tau_i = -\frac{1}{\ln((1 - 2\mu \lambda_i))}, \quad i = 1, 2, ..., N.
\]

As explained subsequently, the convergence of the weight covariances and the mean-square error require, under reasonable assumptions, that \( \mu < 1/2\lambda_{\text{max}} \). With this restriction, it follows from (C.18) that the maximum time constant is

\[
\tau_{\text{max}} = -\frac{1}{\ln(1 - 2\mu \lambda_{\text{min}})}, \quad 0 < \mu < \frac{1}{2\lambda_{\text{max}}}
\]

where \( \lambda_{\text{min}} \) is the smallest eigenvalue of \( R_{xx} \). If \( \mu \) is close to the upper bound in (C.19), then \( \tau_{\text{max}} \) is largely determined by the eigenvalue spread defined as \( \lambda_{\text{max}}/\lambda_{\text{min}} \).

### 1.2 Misadjustment

If the random vectors \( W(n) \) and \( x(n) \) are independent, then (C.2), (C.4), and (C.7) imply that

\[
E[|\epsilon(n)|^2] = \epsilon_m^2 + E[V^H(n)R_{xx}V(n)]
\]
where

\[ V(n) = W(n) - W_0. \]  \hspace{1cm} (C.21)

Even if \( E[W(n)] \to W_0 \), it does not follow that \( E[|\epsilon|^2] \to \epsilon_m^2 \). A measure of the extent to which the LMS algorithm fails to provide the ideal performance is the excess mean-square error, \( E[|\epsilon|^2] - \epsilon_m^2 \). A dimensionless measure of the performance loss, called the misadjustment, is defined as

\[ M = \lim_{k \to \infty} \frac{E[|\epsilon(n)|^2] - \epsilon_m^2}{\epsilon_m^2}. \]  \hspace{1cm} (C.22)

To derive an expression for the misadjustment, we make the following four assumptions:

1. The jointly stationary processes \( x(n + 1) \) and \( d(n + 1) \) are independent of \( x(k) \) and \( d(k) \), \( k \leq n \). It then follows from (C.1) that \( W(n) \) is independent of \( x(n) \) and \( d(n) \).

2. The adaptation constant satisfies

\[ 0 < \mu < \frac{1}{tr(R_{xx})}. \]  \hspace{1cm} (C.23)

3. \( E[\|V(n)\|^2] \) converges as \( n \to \infty \).

4. As \( n \to \infty \), \( |\epsilon(n)|^2 \) and \( \|x(n)\|^2 \) become uncorrelated so that

\[ \lim_{k \to \infty} E[|\epsilon(n)|^2 \|x(n)\|^2] = tr(R_{xx}) \left\{ \lim_{n \to \infty} E[|\epsilon(n)|^2] \right\}. \]  \hspace{1cm} (C.24)

Assumptions 1 and 2 imply convergence of the mean weight vector, which requires (C.16), because the sum of the eigenvalues of a square matrix is equal to its trace, and hence

\[ \lambda_{\text{max}} < \sum_{i=1}^{N} \lambda_i = tr(R_{xx}). \]  \hspace{1cm} (C.25)

The total input power is \( E[\|x(n)\|^2] = tr(R_{xx}) \). For Assumption 3 to be true, a tighter restriction on \( \mu \) than Assumption 2 may be necessary. Assumption 4 is physically plausible, but it is an approximation.

Equations (C.1) and (C.21) imply that

\[ V(n + 1) = V(n) + 2\mu e^*(n)x(n). \]  \hspace{1cm} (C.26)

It follows that

\[ E[\|V(n + 1)\|^2] = E[\|V(n)\|^2] + 4\mu Re\{E[e^*(n)V^H(n)x(n)]\} \]
\[ + 4\mu^2 E[|\epsilon(n)|^2 \|x(n)\|^2]. \]  \hspace{1cm} (C.27)
Assumption 1 and (C.2), (C.5), and (C.6) yield
\[
E[\epsilon^*(n) V^H(n)x(n)] = E[V^H(n)R_{x,d}] - E[V^H(n)R_{xx}W(n)].
\] (C.28)

Substitution of (C.21), (C.4), and (C.20) gives
\[
E[\epsilon^*(n) V^H(n)x(n)] = \epsilon^2_m - E[|\epsilon(n)|^2].
\] (C.29)

Substituting (C.29) into (C.27), taking the limit as \( n \to \infty \), and using Assumptions 3 and 4, we obtain
\[
\lim_{k \to \infty} E[|\epsilon(n)|^2] = \frac{\epsilon^2_m}{1 - \mu tr(R_{xx})}.
\] (C.30)

Assumption 2 ensures that the right-hand side of this equation is positive and finite, which could not be guaranteed if the less restrictive (C.16) were assumed instead. Substituting (C.30) into (C.22), we obtain
\[
M = \frac{\mu tr(R_{xx})}{1 - \mu tr(R_{xx})}.
\] (C.31)

This result applies to both the real and complex discrete-time LMS algorithms. According to (C.31), increasing \( \mu \) to improve the convergence rate has the side effect of increasing the misadjustment. For fixed \( \mu \), the misadjustment increases with the total input power.

C.2 Frost Algorithm

C.2.1 Convergence of the Mean

The Frost or linearly constrained minimum-variance algorithm computes the \( N \times 1 \) weight vector as (Sect. 6.5)
\[
W(0) = \frac{1}{G} p_k
\] (C.32)
\[
W(n + 1) = \left( I - \frac{1}{G} p_k p_k^T \right) \left[ W(n) - 2\mu y(n) \bar{d}_k^*(n) \right] + \frac{1}{G} p_k
\] (C.33)

where \( p_k \) is the \( N \times 1 \) vector of the spreading sequence of user \( k \), \( y(n) \) is the \( N \times 1 \) input vector comprising output samples of a chip-matched filter, and
\[
\bar{d}_k(n) = W^H(n) y(n)
\] (C.34)
is the $N \times 1$ filter output vector. The optimal weight vector is

$$W_0 = \frac{R_y^{-1}p_k}{p_k^T R_y^{-1} p_k} \quad \text{(C.35)}$$

where

$$R_y = E \left[ y(n) y^H(n) \right] \quad \text{(C.36)}$$

is the $N \times N$ Hermitian correlation matrix of $y(n)$. The spreading sequence is normalized so that

$$p_k^T p_k = G \quad \text{(C.37)}$$

If $y(n + 1)$ is independent of $y(k), k \leq n$, then $W(n)$ and $y(n)$ are independent and (C.33) implies that

$$E[W(n + 1)] = A[I - 2\mu R_y]E[W(n)] + \frac{1}{G} p_k, \quad n \geq 0 \quad \text{(C.38)}$$

where

$$A = \left( I - \frac{1}{G} p_k p_k^T \right) \quad \text{(C.39)}$$

Equations (C.38), (C.39), (C.21), (C.35), and (C.37) yield

$$V(n + 1) = AV(n) - 2\mu AR_y V(n), \quad n \geq 0 \quad \text{(C.40)}$$

Direct multiplication verifies that $A^2 = A$. It then follows from (C.40) that $AV(n) = V(n), n \geq 1$. It is easily verified that $AV(0) = V(0)$. Consequently,

$$V(n + 1) = [I - 2\mu AR_y A]V(n) = [I - 2\mu AR_y A]^{n+1}V(0), \quad n \geq 0 \quad \text{(C.41)}$$

A straightforward calculation verifies that the matrix $AR_y A$ is Hermitian, and hence it has a complete set of orthonormal eigenvectors. Direct calculation proves that

$$AR_y A p_k = 0 \quad \text{(C.42)}$$

which indicates that $p_k$ is an eigenvector of $AR_y A$ with eigenvalue equal to zero. Let $e_i, i = 1, 2, \ldots, N - 1,$ denote the $N - 1$ remaining orthonormal eigenvectors. Because the $e_i$ must be orthogonal to $p_k$,

$$p_k^T e_i = 0, \quad i = 1, 2, \ldots, N - 1 \quad \text{(C.43)}$$

From this equation and (C.39), it follows that

$$Ae_i = e_i, \quad i = 1, 2, \ldots, N - 1 \quad \text{(C.44)}$$
Let $\sigma_i$ denote the eigenvalue of $\mathbf{AR}_y\mathbf{A}$ associated with the unit eigenvector $\mathbf{e}_i$. Using (C.44), we obtain

$$\sigma_i = \mathbf{e}_i^H \mathbf{AR}_y \mathbf{A} \mathbf{e}_i = \mathbf{e}_i^H \mathbf{R}_y \mathbf{e}_i, \quad i = 1, 2, ..., N - 1. \quad (C.45)$$

Since $\mathbf{e}_i$ is a unit vector, the Rayleigh quotient of (5.86) implies that

$$\lambda_{\min} \leq \mathbf{e}_i^H \mathbf{R}_y \mathbf{e}_i \leq \lambda_{\max} \quad (C.46)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues, respectively, of the Hermitian matrix $\mathbf{R}_y$. If we assume that $\mathbf{R}_y$ is a positive definite, then $\lambda_{\min} > 0$, and hence $\sigma_i > 0, i = 1, 2, ..., N - 1$. We conclude that the $\{\mathbf{e}_i\}$ correspond to nonzero eigenvalues.

Equations (C.21), (C.32), (C.35), and (C.37) indicate that $\mathbf{p}_k^T \mathbf{V}(0) = \mathbf{0}$. Therefore, $\mathbf{V}(0)$ is equal to a linear combination of the $\mathbf{e}_i, i = 1, 2, ..., N - 1$, which are the eigenvectors of $\mathbf{AR}_y\mathbf{A}$ corresponding to the nonzero eigenvalues. If $\mathbf{V}(0)$ is equal to the eigenvector $\mathbf{e}_l$ with eigenvalue $\sigma_l$, then (C.41) indicates that

$$\mathbf{V}(n + 1) = (1 - 2\mu\sigma_l)^n \mathbf{e}_l, \quad n \geq 0. \quad (C.47)$$

Therefore, a necessary and sufficient condition for the convergence of the mean weight vector is that $|1 - 2\mu\sigma_l| < 1$ for $i = 1, 2, ..., N - 1$. Since $\sigma_i > 0$, the necessary and sufficient condition for convergence is

$$0 < \mu < \frac{1}{\sigma_{\max}}. \quad (C.48)$$

Analogously to (C.18), the convergence of the mean weight vector of the Frost algorithm has transients that can be characterized by the time constants

$$\tau_i = -\frac{1}{\ln(|1 - 2\mu\sigma_i|)}, \quad i = 1, 2, ..., N - 1. \quad (C.49)$$

If $0 < \mu < 1/2\sigma_{\max}$, the largest time constant is

$$\tau_{\max} = -\frac{1}{\ln(1 - 2\mu\sigma_{\min})}, \quad 0 < \mu < \frac{1}{2\sigma_{\max}} \quad (C.50)$$

where $\sigma_{\min}$ is the smallest nonzero eigenvalue of $\mathbf{AR}_y\mathbf{A}$. If $\mu$ is close to the upper bound in (C.50), then $\tau_{\max}$ is largely determined by the eigenvalue spread defined as $\sigma_{\max}/\sigma_{\min}$. 
Index

A
Acquisition, 216–237
   consecutive-count strategy, 220, 260
   lock mode, 221
   matched-filter, 217–218, 227, 250–257
   multiple-dwell, 220
   multiple-dwell system, 260
   parallel array, 216
   search control system, 260
   sequential detection, 237
   sequential estimation, 217
   serial-search, 219–237, 257–262
   single-dwell, 220
   up-down strategy, 220, 260
   verification mode, 221
Acquisition correlator, 238–244
Acquisition time, 221
density function, 232–233
Ad hoc network, 398, 423–430
Adaptive ACM filter, 154–156
Adaptive filters, 140–144
Adjacent splatter ratio, 173
Alamouti code, 316–318
Analytic signal, 546
Aperiodic autocorrelation, 127
Approximate conditional mean(ACM) filter, 153–154
Area-mean power, 267
Attenuation power law, 268
Autocorrelation
   of stationary process, 547
   Autoregressive process, 152
   Average autocorrelation, 93
B
BCJR algorithm, 49–53
Beta function, 313
BICM, 68–70, 327
BICM-ID, 69, 327
Bit-interleaved coded modulation, see BICM,
   see BICM
Block code, 1–29
   BCH, 4, 11
   cyclic, 4, 10
   extended, 5
   Golay, 5
   Hamming, 5
   linear, 4
   maximum-distance-separable, 6
   perfect, 4
   Reed–Solomon, 6
   repetition, 4
   systematic, 6
Bluetooth, 207, 208, 210, 420
Burst communications, 217
C
Cauchy–Riemann conditions, 141
Cauchy–Schwarz inequality, 301
CDMA2000, 70, 315, 316, 510
Cell, 219
Cellular network, 396–420, 423–425, 430–440
Channel
   frequency response, 283
   impulse response, 282–284
Channel capacity, 521
   constrained, 522
Channel estimation, 197, 353–354, 360–361,
   504–542
   blind, 504
Channelization code, 397
Channelized radiometer, 488–494
Characteristic function, 556
Characteristic polynomial, 378
Chase algorithm, 19, 60
Chebyshev’s inequality, 223
Chernoff bound, 41–43
Chi-square distribution, 555–558
Chip waveform, 101
   rectangular, 80
   sinusoidal, 105
Circulant matrix, 345
Circular state diagram, 233
Circular symmetry, 16, 550
Clipping function, 448
Code rate, 5
Code tracking, 216, 244–250, 262–264
delay-locked loop, 245–248
early-late gate, 263
tau-dither loop, 248–250
Code-aided methods, 156
Code-division multiple access (CDMA)
definition, 365
Code-shift keying (CSK), 132–135
Coded modulation, 44, 68
Coding gain, 19
Coherence bandwidth, 281
Coherence time, 275
Complementary error function, 13
Complete data vector, 498
Complex envelope, 547
Complex gradient, 140–142
Complex-valued quaternary sequence, 377–381
Concatenated code, 47–49
Confluent hypergeometric function, 559
Constellation labeling, 65
Constraint length, 30
Continuous-phase frequency-shift keying, see CPFSK
Continuous-phase modulation, see CPM
Convex function, 381
Convolutional code, 29–41
catastrophic, 36
constraint length, 30
generating function, 40
generators, 30
linear, 30
minimum free distance, 34
punctured, 37
sequential decoding, 34
state, 32
systematic, 30
trellis diagram, 33
Viterbi decoder, 33
Convolver, 130
CPFSK, 174
CPM, 173–185
Cross-correlation
   aperiodic, 377
   continuous-time partial, 376
   parameter, 180
   periodic, 367
Cyclic prefix, 343
Cyclostationary process, 94

D
Decimation, 372
Decision-directed demodulator, 138
Decoder
   bounded-distance, 2
   complete, 2
   erasing, 7
   incomplete, 2
   reproducing, 7
   sequential, 34
   Viterbi, 33
Decoding
   errors-and-erasures, 14
   hard-decision, 7
   soft-decision, 14
Degree distribution
   check nodes, 63
   variable nodes, 63
Dehopping, 161
Delay spread, 280
Despreading, 82
Deviation ratio, 173
Differential phase-shift keying (DPSK), 135
Direct-conversion receiver, 552–554
Discrete Fourier transform, 343
Diversity, 284
   frequency, 284
   path, 327
   polarization, 284
   selection, 309
   spatial, 284, 430
   time, 284
Diversity order, 296
Divider, 205–207
dual-modulus, 206
Doppler
   factor, 416
   shift, 270
   spectrum, 276, 284
   spread, 275, 415
Double-dwell system, 225–228
Double-mix-divide system, 199
Downlink, 397
Downlink capacity, 437
Index

DS/CDMA, 366–420, 437–440
DS/CDMA network, 456, 457
Duplexing, 398
Duty factor, 423
Dwell interval, 162

E
Early-late gate, 263
Energy detector, see Radiometer
Equal-gain combining, 301–309
Equalization, 347–354, 358–360
Equalizer
linear, 347
maximal-ratio combining, 348, 358
minimum mean-square error (MMSE), 349, 359
zero-forcing, 348, 358
Equivalent noise-power spectral density, 386
Erasure, 14
Error probability
channel-symbol, 7
decoded-symbol, 10
decoding failure, 8
information-bit, 10
information-symbol, 10
undetected, 8
word, 8
Error rate
decoded-symbol, 10
information-symbol, 10
Error-floor region, 69
Euclidean distance, 17
Euler function, 97
Expectation Maximization (EM), 497–502
algorithm, 499

F
Fading, 268–282
block, 505
fast, 276
frequency-selective, 280
Nakagami, 273
Rayleigh, 272
Ricean, 273
slow, 276
Fading rate, 277–278
Fast frequency hopping, 162
Feedback shift register, 85
FH/CDMA, 420–440, 459
asynchronous network, 421
synchronous network, 421
FH/CPFSK
power spectral density, 183
FH/CPM, 181–185
FH/DPSK, 178–180
FH/FSK, 163–172
Fixed-point iteration, 502–503
Fractional power, 176
Frequency channel, 159
Frequency discriminator, 184
Frequency synthesizer, 199–210
digital, 201–204
direct, 199–201
fractional-N, 209–210
indirect, 204–210
multiple-loop, 207–209
Frequency-domain equalization (FDE), 356–362
Frequency-hopping pattern, 159
Frequency-selective fading, 280–282
Frequency-shift keying, see FSK
Frost algorithm, 454, 567
FSK, 24–27

G
Galois field, 3
Gamma function, 474, 556
Gaussian approximation
improved, 391
standard, 391
Gaussian interference, 109–111
Gaussian MSK (GMSK), 177
Generalized Marcum Q-function, 474, 557
Generating function
acquisition time, 235
convolutional code, 40
Gold sequence, 374
moment, see Moment generating function
polynomial, 94
Global System for Mobile (GSM), 177, 398
Gold sequence, 373–374
Gradient vector, 498
Gray labeling, 69
Guard interval, 343

H
Hadamard matrix, 369
Hamming bound, 3
Hamming correlation, 421
Hamming distance, 2
Hard-decision decoding, 7–14
Hilbert transform, 545
Hop duration, 159
Hop interval, 159
Hop rate, 159
Hopping band, 159
Hopset, 159
Hybrid systems, 185–186

I
Ideal detection, 469, 487
Incomplete beta function, 489
Incomplete data vector, 498
Incomplete gamma function, 473
Interblock interference, 343
Intercell interference, 399
Interchip interference, 82, 101, 103
Interference canceller, 455–459
multistage, 457
parallel, 458
successive, 456
Interleaver, 45–47
block, 46
convolutional, 46
helical, 46
pseudorandom, 47
S-random, 47
Intersymbol interference, 101, 103
Intracell interference, 399
IRA code, 64–65
Irregular repeat-accumulate code, see IRA code
IS-95, 396, 398, 415, 510
Isotropic scattering, 274
Iterative demodulation and decoding, 65–70

J
Jensen’s inequality, 381–383

K
Kalman–Bucy filter, 149, 152
Kasami sequence, 374–375
Key, 101, 161

L
LDPC code, 61–65, 70, 197–198
irregular, 61
regular, 61
Least-mean-square algorithm, see LMS algorithm
Likelihood equation, 498
Likelihood function, 15
Linear span, 161
Linearly constrained minimum-variance algorithm, 454
LMS algorithm, 144, 146, 563
Local-mean power, 268
Lock detector, 229
Log-likelihood function, 15
Log-MAP algorithm, 52–53
Lognormal distribution, 268
Low probability of interception, 465
Low-density parity-check code, see LDPC code

M
MAP algorithm, 49–53
Marcum Q-function, 254
Matched filter, 127–139
bandpass, 127
convolver, 130
SAW transversal filter, 129
Matched-filter acquisition, 217–218, 227
frequency-hopping, 250–257
Max-star function, 52
Maximal sequence, 89–97
preferred pair, 373
Maximal-ratio combining, 289–301
MC-CDMA system, 341–356
Message privacy, 80
Metric, 15, 19–27
AGC, 122
BPSK, 292, 319
coherent FSK, 321
coherent FSK, 299
correlation, 443
maximum-likelihood, 17, 22, 23, 121
Rayleigh, 23, 322
self-normalization, 170
variable-gain, 166
white-noise, 124
Minimum distance, 2
Minimum free distance, 34
Minimum-shift keying, see MSK
Missing data vector, 500
Modified Bessel function, 22
first kind and order n, 473, 556
Moment generating function, 41
Mother code, 369
Moving-window detection, 491
MSK, 174
Multicarrier CDMA system, see MC-CDMA system
Multicarrier direct-sequence system, 339
Index

Multipath, 269
  diffuse components, 272
  intensity profile, 283
  intensity vector, 337
  resolvable components, 281, 328
  specular components, 272
  unresolvable components, 270
Multiple access, 365
Multiuser detector, 440–459
  adaptive, 452
  blind, 453
  decorrelating, 443–448
  for frequency hopping, 459
  interference canceller, 455–459
  minimum mean-square error (MMSE), 448–452
  optimum, 441–443
Mutual Information, 521

N
  Nakagami density, 273
  Nakagami fading, 273
  Narrowband interference, 140–156
  Near-far problem, 397
  Network user capacity, 395
  Noncoherent combining loss, 165
  Noncoherent correlator, 217
  Nonlinear filter, 149–156
  Nonlinear generator, 100–101
  Norm, Euclidean, 287

O
  OFDM, 342
  Optimal array, 285–288
  Orthogonal frequency-division multiplexing, see (OFDM)
  Orthogonal sequences, 368
  Orthogonal variable-spreading-factor codes, 369
  Orthogonality condition, 317
  Outage, 403
    local-mean, 409
  Output threshold test, 192

P
  Packing density, 11
  PAPR, 355–356
  Partial-band interference, 186–195
  Peak-to-average-power ratio, see PAPR
  Peer-to-peer network, see Ad hoc network
  Penalty time, 222, 229
  Periodic autocorrelation, 90–94
  Phase stripper, 290
  Pilot symbols, 190, 504
  Pilot-assisted channel estimation (PACE), 504
  Polynomial, 94–97
    characteristic, 94, 378
    generating function, 94
    irreducible, 96
    primitive, 96
  Power control, 398–399, 406–420
    closed-loop, 398
    open-loop, 398
  Power spectral density, 176, 550
    average, 93
    of direct-sequence signal, 94
    of spreading waveform, 93
  Prefix factor, 347
  Probability densities, see Probability distributions
  Probability distributions, 555–563
    central chi-square, 557
    chi-square, 555
    exponential, 561
    lognormal, 268
    Nakagami, 273
    noncentral chi-square, 555
    Poisson, 405
    Rayleigh, 272, 560
    Rice, 273, 558
  Processing gain, 80, 102
  Product code, 59
  Pseudonoise sequence, 93
  Psi function, 400
  Pulsed interference, 118–126

Q
  Q-ary symmetric channel, 13
  Q-function, 13, 559
  Quaternary system, 112–116
    balanced, 115
dual, 112

R
  Radiometer, 123, 469–494
  Rake demodulator, 327–335
    fingers, 329
    path crosstalk, 331
  Rake receiver, 335
  Random binary sequence, 83–84
    autocorrelation, 84
  Ratio threshold test, 192
Rayleigh distribution, 560–561
Rayleigh fading, 272
Rayleigh metric, 23, 306, 322
Rayleigh quotient, 287
Recirculation loop, 136–138
Reed–Solomon code, 189–195
Repeat-accumulate code, 64
Rewinding time, 222
Rice distribution, 558–559
Rice factor, 273
Ricean fading, 273
Rayleigh quotient, 287
SAW elastic convolver, 130–132
SAW transversal filter, 129
Scrambling code, 397
Search control system, 260
Search strategy
   broken-center Z, 221
   equiexpanding, 231
   expanding-window, 230
   nonuniform alternating, 232
   uniform, 221
   uniform alternating, 231
   Z, 230
Sector antenna, 430
Selection diversity, 309–315
generalized, 332
Self-interference, 240
Separated orthogonal signals, 431
Serial-search acquisition, 219–237
   frequency-hopping, 257–262
Shadowing, 268
   factor, 268
Shift register sequence, 85–101
   linear, 85
   maximal, 89
Side information, 165, 190
Signature sequences, 366
Signum function, 443, 546
Single-dwell system, 226
Singleton bound, 6
SISO algorithm, 52
   log-MAP, 52
   max-log-MAP algorithm, 53
   SOVA algorithm, 53
Slow frequency hopping, 163
Small misalignment technique, 259
Soft-decision decoding, 14–28, 165–170
Soft-in soft-out algorithm, see SISO algorithm
Space-time codes, 316
Spatial diversity, 278–282
Spatial reliability, 427
Spectral notching, 161
Spectral splatter, 172
Spreading factor, 369
Spreading sequence, 80, 83–101
   linear complexity, 100
   long, 98
   short, 98
Spreading waveform, 80, 83–101
Steepest descent
   method, 144
Step size, 219
Sum-product algorithm, 61–63
Switch-and-stay combining, 315
Switching time, 162
Tanner graph, 61
   check nodes, 61
   cycle, 61
   girth, 61
   variable nodes, 61
Third-generation cellular systems, 440, 510
Throughput, 513
Time of day (TOD), 161, 260
Time-domain adaptive filter, 144–146
Tone interference, 105–109
Trace of matrix, 348
Transform-domain processor, 147–149
Transmission security, 161
Transmit diversity, 315–318
Trellis-coded modulation, 44–45, 58, 195–197
Triangular function, 84
Turbo code, 197–199
   BCH, 58
   block, 57
   channel reliability factor, 56
   convolutional, 53
   error floor, 55
   extrinsic information, 56
   product code, 60
   serially concatenated, 59
   system latency, 55
   trellis-coded modulation, 58
UMTS, 534
Uncorrelated scattering, 282
Union bound, 18
Uplink, 397
Uplink capacity, 410, 436
Index

V
Voice-activity detector, 365

W
W-CDMA, 504
Wait technique, 260
Walsh sequence, 369
Waterfall region, 69

Weight
distribution, 6
Hamming, 6
information-weight spectrum, 35
total information, 119
Welch bound, 372
Wiener–Hopf equation, 143, 288, 563
WiMAX, 72