CHAPTER

Signal Parameter Estimation for the Localization Problem

2

The vast majority of existing positioning systems involve the use of different distance or direction-dependent measurements. For instance, received signal strength (RSS) is dependent, for a given transmitted power, on the distance between a receiver and the transmitting source. Signal propagation time, usually referred to as time of arrival (TOA), is also dependent on distance, which may be directly inferred if the medium propagation speed is known. Angle of arrival (AOA) observations also provide location information about an emitter, and in some cases, angular and range measurements are combined to improve performance. Clearly, the accuracy of a position location scheme will strongly rely on the accuracy of the corresponding range or angular measurements, and for this reason, correct understanding of their measurement and estimation techniques and of their limitations becomes of utmost importance.

This chapter is devoted to presenting some popular AOA, TOA, and RSS measurement/estimation techniques, and will provide the reader with an overview of the different phenomena that limit the accuracy of these measurements. Theoretical accuracy bounds will also be presented, for some specific scenarios, in order to provide more insight on how the different transmitted signal, channel, and receiver parameters affect range and direction measurements.

2.1 AOA MEASUREMENTS

The angle at which a signal arrives at a sensor, such as a hydrophone or an antenna, yields important information about the location of the signal source. As will be explained in Chapter 3, two error-free AOA measurements obtained at two different sensors located at different points in space will be enough to estimate the exact position of the signal-generating source in a 2D scenario. In this section, we discuss several techniques to measure the AOA of a signal, or a

group of signals, in the presence of several impairments such as sensor noise, unknown sensor response, limited sensor observations, and signal propagation distortion such as attenuation and multipath.

Most of the existing AOA estimation techniques rely on observations of signals at the output of multiple sensors spaced by fractions or a few wavelengths, and deployed with a specific geometrical arrangement. These sensor schemes are usually referred to as sensor arrays. A signal $s(t)$ impinging on the array with an angle different than 90 degrees (with respect to the array axis) will reach individual sensors at slightly different times; thus, sensor outputs will contain time- and phase-delayed versions of the original signal. Array processing techniques take advantage of the time and phase delay that exists at the sensor outputs and attempt to extract the corresponding embedded angle of arrival information.

In this chapter we will focus on array processing techniques for AOA estimation. We will present the AOA estimation algorithms and then discuss their strengths and weaknesses and, specifically, their sensitivity to noise and channel impairments. We will also present theoretical bounds on performance for the case of arrays with additive white Gaussian noise (AWGN).

2.1.1 The Uniform Linear Array Model

Consider the problem of solving the directions of arrival of multiple signals produced by narrow band radiating uncorrelated sources in the presence of background noise. We assume D signals incident on a uniform linear array of M sensors from angles $\{\theta_i,\ i = 1, 2, \ldots, D\}, \theta_i \in [-90^\circ, 90^\circ]$. It is also assumed that the array is in the far field of any of the sources such that the impinging signals are planar. Let us first concentrate on the signal generated by the i-th source, as shown in Figure 2.1, and its corresponding output at the m-th sensor. Note that a wave arriving at sensor m travels an extra distance of $d \sin\theta_i$ meters when compared to the distance traveled to reach sensor $m-1$. This means that signals at contiguous sensors have a time delay of $d \sin\theta_i / u$ seconds, or a phase delay of $k_i = \omega_c d \sin\theta_i / u$, where u is the propagation speed in meters per second, and ω_c is the carrier frequency of the narrow band signal. To avoid ambiguities in the spatial sampling of the array, the phase delay should be limited to $|k_i| < \pi,\ i = 1, 2, \ldots, D$. This condition is achieved for any θ_i, $i = 1, 2, \ldots, D$, if the separation between the array sensors is limited to $d < \frac{\lambda_c}{2}$, where $\lambda_c = 2\pi u / \omega_c$ is the wavelength of the impinging narrow band signals.

If we reference the sensor output phases with respect to the first sensor (i.e., sensor one is considered to have zero phase delay), then the output at the m-th sensor due to the i-th source may be written as

$$r_m(t) = H_m(\omega_c, \theta_i) e^{-j(m-1)k_i} s_i(t) + \eta_m(t), \tag{2.1}$$

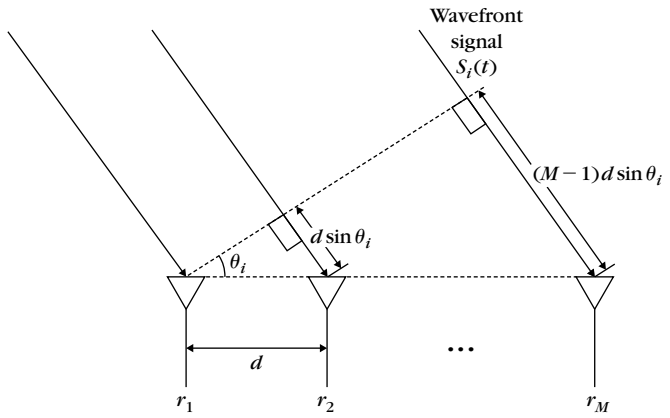

FIGURE 2.1

Uniform linear antenna array.

where $H_m(\omega_c, \theta_i)$ is the frequency response of the m-th sensor evaluated at the center carrier frequency. Note that it has been emphasized that this sensor response is a function of the signal's angle of arrival. This response is mainly determined by the inherent sensor characteristics such as its gain, directivity, sensitivity pattern, and so on. The term $\eta_m(t)$ is additive stationary white Gaussian noise at the m-th sensor with power equal to σ^2, which may correspond to thermal noise generated by the sensor's hardware. This noise is assumed to be uncorrelated with the signals.

Let us now consider all D signals impinging on the array to obtain the field received at the m-th sensor at time t as

$$r_m(t) = \sum_{i=1}^{D} H_m(\omega_c, \theta_i) e^{-j(m-1)k_i} s_i(t) + \eta_m(t); \quad m = 1, \ldots, M, \tag{2.2}$$

For any real number θ, we define the corresponding *steering vector* as

$$\mathbf{a}(\theta) = \left[H_1(\omega_c, \theta)\ H_2(\omega_c, \theta)e^{-jk} H_3(\omega_c, \theta)e^{-2jk} \ldots H_M(\omega_c, \theta)e^{-(M-1)jk}\right]^T,$$
$$k = \frac{\omega_c d}{u} \sin\theta. \tag{2.3}$$

Clearly this vector depends on θ through the phase delay k, and the sensor responses $H_m(\omega_c, \theta)$. The signals received by the M sensors can then be written as

$$\begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_M(t) \end{bmatrix} = [\mathbf{a}(\theta_1)\, \mathbf{a}(\theta_2) \cdots \mathbf{a}(\theta_D)] \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_D(t) \end{bmatrix} + \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \\ \vdots \\ \eta_M(t) \end{bmatrix}, \tag{2.4}$$

or more compactly,

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \boldsymbol{\eta}(t). \tag{2.5}$$

Matrix $\mathbf{A} = [\mathbf{a}(\theta_1)\,\mathbf{a}(\theta_2)\,\cdots\,\mathbf{a}(\theta_D)]$ is usually called the *array manifold*, and its columns are referred to as *steering vectors* since, as will be shown, they may be used to "steer" the array response in any desired direction. It is assumed that for any set of distinct parameters θ_i, $i = 1, 2, \ldots, D, M > D$, the columns of $\mathbf{A}$ are linearly independent. The array manifold may be expressed in a closed analytical form or measured through field calibration procedures. Actually, a major source of error in AOA estimation methods consists precisely of the incorrect calibration of the array sensors, which basically translates into the incorrect modeling of the sensor's transfer functions [44].

In many scenarios, it is possible to assume that all M sensors are omnidirectional and identical; then $H_m(\omega_c, \theta) = constant\ \forall\, m,\ \theta$. In this case, without loss of generality, the constant gain may be embedded in the amplitudes of signals $s_i(t)$ to obtain an array response as in Equation (2.4), but with simplified steering vectors

$$\mathbf{a}(\theta) = \left[1\, e^{-jk}\, e^{-2jk}\, \ldots\, e^{-(M-1)jk}\right]^T. \tag{2.6}$$

Returning to the general case, from Equation (2.5) the spatial array correlation matrix can be written as

$$\mathbf{R} = E\{\mathbf{r}(t)\mathbf{r}(t)^H\} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \mathbf{R_n}, \tag{2.7}$$

where $\mathbf{S}$ and $\mathbf{R_n}$ are the correlation matrices of $\mathbf{s}(t)$, and $\boldsymbol{\eta}(t)$, respectively. The noise is assumed spatially and temporally uncorrelated so that $\mathbf{R_n} = \sigma^2\mathbf{I}$, and the signal correlation matrix is assumed nonsingular and not necessarily diagonal, implying that signals may be partially correlated. Note that if the signals are fully correlated then they are said to be coherent and indeed, in this scenario, $\mathbf{S}$ will become singular.

In many situations it will be convenient to weight the sensor outputs and add the resulting products to obtain an overall array response $r_B(t)$. This is achieved with a so-called beamformer whose structure is shown in Figure 2.2. The resultant beamformer output signal $r_B(t)$ may be written as

$$r_B(t) = \mathbf{w}^H\mathbf{r}(t), \tag{2.8}$$

where $\mathbf{w} = [w_1 \ldots w_M]^T$ corresponds to the vector of weights of the beamformer, and $\mathbf{r}(t)$ is the array output vector described in Equation (2.5). The output average power of the beamformer may then be calculated as

$$E\{|r_B(t)|^2\} = \mathbf{w}^H\mathbf{R}\mathbf{w}. \tag{2.9}$$

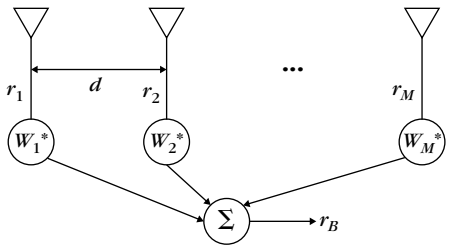

FIGURE 2.2
Beamformer.

Although the array output vector $\mathbf{r}(t)$ presented in Equation (2.4) was derived for a uniform linear array, this equation may apply to an arbitrary array geometry as long as the adequate array manifold matrix $\mathbf{A}$, obtained analytically or by calibration, is considered in the model. The array or beamformer outputs may be sampled at different time instants in order to obtain an ensemble of array snapshots that may be used in the estimation of the angles of arrival θ_i, $i = 1, 2, \ldots, D$. In the following sections we will discuss AOA estimation techniques commonly applied by position estimation schemes. In what follows, we will assume that the number of signal sources D is known or has been successfully estimated. Methods to estimate the number of sources may be found in [47, 48].

2.1.2 Cramer Rao Bound for Array Observations

The Cramer Rao bound (CRB) [38] obtains expressions for the best possible error covariance that may be obtained by any unbiased parameter estimator (independent of the particular estimator's form). The CRB for the observation model given in Equation (2.5) was found in [40] for the case of signals $\{s_i(t),\ i = 1, \ldots, D\}$ with known correlation matrix and unknown noise power σ^2. More general performance bounds may be found in [25, 46]. Given all the assumptions made in the previous section ($M > D$, nonsingular $\mathbf{S}$, linear independence between columns of $\mathbf{A}$, white Gaussian noise), the CRB for any unbiased estimator of the vector of AOAs $\boldsymbol{\theta} = [\theta_1 \ldots \theta_D]^T$, which is based on N independent snapshot observations of M sensor outputs at times $\{t_1 \ldots t_N\}$ is given by

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{\sigma}{2}\left\{\sum_{l=1}^{N} Re\left[\boldsymbol{\Upsilon}(t_l)^H \mathbf{D}^H [\mathbf{I} - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H]\mathbf{D}\boldsymbol{\Upsilon}(t_l)\right]\right\}^{-1}, \tag{2.10}$$

$$VAR_{CRB}(\sigma) = \frac{\sigma^2}{MN},$$

where $\Upsilon(t) = \text{diag}\{s_1(t), \ldots, s_D(t)\}$, and $\mathbf{D} = \left[\frac{d\mathbf{a}_1(\theta)}{d\theta} \ldots \frac{d\mathbf{a}_D(\theta)}{d\theta}\right]$. This bound on the error covariance may be used to compare the performance of the specific parameter estimators to be presented in the following discussion.

2.2 NONPARAMETRIC METHODS FOR AOA ESTIMATION

Nonparametric methods make no assumptions about the correlation structure of the data and assume that the array is calibrated—that is, that the form of the steering vector $\mathbf{a}(\theta)$ is known.

2.2.1 Beamscan AOA Estimator

Analogous to power spectral density estimators, a spatial periodogram may be defined [42] as

$$|\mathbf{a}(\theta)^H \mathbf{r}(t)|^2 = \left| \sum_{m=1}^{M} r_m(t) e^{j(m-1)k} \right|^2, \tag{2.11}$$

which may be computed using a fast Fourier transform (FFT). When the steering vector is scanned over θ, this periodogram yields large variance estimates of the spatial spectral density of the array output vector $\mathbf{r}(t)$. To reduce this variance, several periodograms obtained with N array output snapshots may be averaged to yield

$$P_{\text{BS}}(\theta) = \frac{1}{N} \sum_{t=1}^{N} \left|\mathbf{a}(\theta)^H \mathbf{r}(t)\right|^2 = \mathbf{a}(\theta)^H \hat{\mathbf{R}} \mathbf{a}(\theta), \tag{2.12}$$

where

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{t=1}^{N} \mathbf{r}(t)\mathbf{r}(t)^H, \tag{2.13}$$

which corresponds to the N-snapshot estimate of the array correlation matrix. The estimates of $\theta_i, i = 1, 2, \ldots, D$ may be chosen as the D largest peaks of the beamscan estimator function $P_{\text{BS}}(\theta)$.

The resolution of the beamscan estimator is inversely proportional to the array aperture and, for large array lengths and $d = \frac{\lambda}{2}$, it may be approximated as $\frac{1}{M}$. The beamscan estimator is consistent for the case of a single source. Multiple sources, however, will cause inconsistent AOA estimates, and their bias will grow when the sources are strongly correlated or closely spaced. Note, however, that this estimator will obtain accurate AOA estimates when the sources are uncorrelated and separated by more than the array beam width. Closely

spaced and highly correlated sources may arise in dense multipath scenarios where multiple delayed copies of a signal may reach the array from different directions. These multiple signal copies may be caused by signal reflections and scattering from objects such as buildings, cars, trees, people, and so on. If the reflecting objects are close to each other, so are the AOAs of the different reflections.

2.2.2 MVDR AOA Estimator

MVDR stands for *minimum variance distortionless response*, and the estimator may be obtained by minimizing the average output power of a beamformer $E\{|y(t)|^2\}$ as given in Equation (2.9) subject to the linear constraint $\mathbf{w}^H\mathbf{a}(\theta) = 1$. It is easy to verify that the minimization is achieved by the weight vector [18],

$$\mathbf{w} = \frac{\mathbf{R}^{-1}\mathbf{a}(\theta)}{\mathbf{a}(\theta)^H\mathbf{R}^{-1}\mathbf{a}(\theta)}. \tag{2.14}$$

The idea is to minimize the array response to every AOA, except a prescribed angle of interest θ. Inserting this optimal weight into Equation (2.9), and replacing the spatial correlation matrix $\mathbf{R}$ by its sample estimate counterpart $\hat{\mathbf{R}}$ yields the MVDR AOA estimator,

$$P_{\text{MVDR}}(\theta) = \frac{1}{\mathbf{a}(\theta)^H\hat{\mathbf{R}}^{-1}\mathbf{a}(\theta)}. \tag{2.15}$$

The estimates of θ_i, $i = 1, 2, \ldots, D$ may be chosen as the D largest peaks of $P_{\text{MVDR}}(\theta)$ when "steered" over all possible values of θ. Note that the sample correlation matrix estimate $\hat{\mathbf{R}}$ will be nonsingular if the noise correlation matrix is nonsingular (which is clearly the case in Equation (2.7) where $\mathbf{R_n} = \sigma^2\mathbf{I}$). The MVDR has been shown, empirically, to have better performance than the beamscan estimator [42].

2.3 PARAMETRIC METHODS FOR AOA ESTIMATION

Performance of the nonparametric AOA estimation techniques may be improved if prior information about the array model or the signal statistics is available.

2.3.1 Maximum Likelihood AOA Estimator

The maximum likelihood (ML) principle is based on the maximization of the likelihood function of the array observations, which follows from the joint density of N snapshots of the array outputs. Under the assumption of Gaussian

uncorrelated noise, the joint probability density function for N array snapshots at times $\{t_l,\ l = 1 \ldots N\}$ is given by

$$f_{\mathbf{r}(t_1)\ldots\mathbf{r}(t_N)} = [(2\pi)^{-MN}(\sigma/2)^{MN}] \exp\left\{-\frac{1}{\sigma^2}\sum_{l=1}^{N}[\mathbf{r}(t_l) - \mathbf{A}\mathbf{s}(t_l)]^H[\mathbf{r}(t_l) - \mathbf{A}\mathbf{s}(t_l)]\right\}. \tag{2.16}$$

Taking the logarithm of this density and ignoring terms that do not depend on the AOA parameters $\theta_i,\ i = 1 \ldots M$, we obtain the log-likelihood function (the negative sign will also be neglected with the effect of converting the maximization problem into a minimization one),

$$L = \frac{1}{\sigma^2}\sum_{l=1}^{N}[\mathbf{r}(t_l) - \mathbf{A}\mathbf{s}(t)]^H[\mathbf{r}(t_l) - \mathbf{A}\mathbf{s}(t)]. \tag{2.17}$$

Clearly, for the case of non-Gaussian noise, the above function corresponds to the sum of squared errors of the model whose minimization leads to a least-squares (LS) estimation scheme.

First we use the well-known LS solution [38, 42] to estimate the unknown signal amplitude vectors $\mathbf{s}(t_l)$ to obtain

$$\hat{\mathbf{s}}(t_l) = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{r}(t_l), \quad l = 1 \ldots N. \tag{2.18}$$

Inserting this result into Equation (2.18), we obtain

$$L = \frac{N}{\sigma^2} tr\left[\mathbf{I} - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\hat{\mathbf{R}}\right], \tag{2.19}$$

where $\hat{\mathbf{R}}$ is the sample array correlation matrix given in Equation (2.13). It follows from this last equation that the ML estimate for $\theta_i,\ i = 1, 2, \ldots, D$ is obtained by solving

$$\{\hat{\theta}_i\}_{i=1}^{D} = \arg\max_{\{\theta_i\}_{i=1}^{D}} tr\left[\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\hat{\mathbf{R}}\right], \tag{2.20}$$

which corresponds to a nonlinear multidimensional maximization problem. The solution to this type of problem will usually be computationally intensive and will suffer from local minima and convergence issues. Implementations of the AOA ML estimator have been discussed in several works (see references 10–12, 33–36, 42 in [40]).

According to Stoica and Nehorai [40], the ML estimator is inefficient in the sense that an infinite number of snapshots will not be enough for the estimator to reach the CRB as long as the number of sensors in the array M is finite. The CRB will only be reached by letting $M \to \infty$, which could only be achieved by an infinite aperture array.

An interesting observation in Stoica and Moses [42] states that the beamscan AOA estimator provides an approximate solution to the ML problem whenever

the AOAs are well separated. In this scenario, it can be seen that $\mathbf{A}^H\mathbf{A} \approx M\mathbf{I}$. Then the objective function in Equation (2.20) may be approximated as

$$tr[\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\hat{\mathbf{R}}] \approx \frac{1}{M}\sum_{i=1}^{D}\mathbf{a}(\theta_i)^H\hat{\mathbf{R}}\mathbf{a}(\theta_i), \tag{2.21}$$

where the right side of the equation will be maximized at the true AOAs as was previously explained in the beamscan estimator discussion.

Due to the complexity of the AOA ML estimator, several other suboptimal estimators are usually preferred, and actually many of these suboptimal schemes will reach an optimal performance under specific scenarios [40, 41].

2.3.2 MUSIC Algorithm for AOA Estimation

The multiple signal classification (MUSIC) algorithm [39] falls into the category of the so-called subspace, or super-resolution, estimators. These algorithms profit from the eigenstructure properties of the array correlation matrix to obtain very-high-resolution estimators with lower computational complexity when compared to ML estimation schemes.

Let us analyze the properties of the spatial correlation matrix $\mathbf{R}$ described in Equation (2.7). It is clear that if the number of array sensors is larger than the number of signal sources (i.e., $M > D$), when $\mathbf{S}$ is positive definite (i.e., the signals $\mathbf{s}_i(t)$ are not fully correlated), the matrix $\mathbf{R} - \mathbf{R_n}$ will have rank D and a null space of dimension $M - D$. Then matrix $\mathbf{R}$ will have D eigenvalues greater than σ^2 and $M - D$ eigenvalues equal to σ^2; these eigenvalues may be ordered from largest to smallest such that $\lambda_1 > \lambda_2 > \cdots > \lambda_D > \lambda_{D+1} = \lambda_{D+2} = \ldots = \lambda_M = \sigma^2$. The eigenvectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_D\}$ corresponding to the largest eigenvalues span the D-dimensional *signal subspace*. We can group these signal eigenvectors in the columns of matrix $\mathbf{E}_s$. The eigenvectors $\{\mathbf{e}_{D+1}, \mathbf{e}_{D+2}, \ldots, \mathbf{e}_M\}$ corresponding to the smallest eigenvalues span the $(M - D)$-dimensional *noise subspace*. These signal eigenvectors can be grouped in the columns of matrix $\mathbf{E}_n$. If we write the eigenvalue problem,

$$\mathbf{R}\mathbf{e}_i = \lambda_i\mathbf{e}_i; \quad i = D+1, D+2, \ldots, M, \tag{2.22}$$

we may substitute $\mathbf{R}$ by the leftmost side of Equation (2.7), and realize that the eigenvalue λ_i for every noise subspace eigenvector is equal to σ^2 to obtain

$$\mathbf{A}\mathbf{S}\mathbf{A}^H\mathbf{e}_i + \sigma^2\mathbf{e}_i - \sigma^2\mathbf{e}_i = \mathbf{0}; \quad i = D+1, D+2, \ldots, M, \tag{2.23}$$

which implies that

$$\mathbf{A}^H\mathbf{e}_i = \mathbf{0}; \quad i = D+1, D+2, \ldots, M. \tag{2.24}$$

Then it is clear that the columns of matrix $\mathbf{A}$ are orthogonal to the eigenvectors that span the noise subspace.

For an exactly known $\mathbf{R}$, we can find the desired angles $\theta_i, i = 1, 2, \ldots, D$, by the following steps that comprise the MUSIC algorithm.

1. Compute the eigenvalues and eigenvectors of the $(M \times M)$ matrix $\mathbf{R}$.
2. D and σ^2 are determined by the facts that the minimum eigenvalue is equal to σ^2 and has multiplicity $M - D$.
3. Group the spanning eigenvectors for the noise subspace in the set $\{\mathbf{e}_{D+1}, \mathbf{e}_{D+2}, \ldots, \mathbf{e}_M\}$.
4. Perform an orthogonality test between different candidate steering vectors and the noise subspace to obtain the MUSIC pseudospectrum,

$$P_{\text{MUSIC}}(\theta) = \frac{\mathbf{a}(\theta)^H \mathbf{a}(\theta)}{\sum_{j=D+1}^{M} [\mathbf{a}(\theta)^H \mathbf{e}_j]^2}. \tag{2.25}$$

Ideally $P_{\text{MUSIC}}(\theta)$ will peak to infinity each time a true $\theta_i, i = 1, 2, \ldots, D$ angle is tested. When neither D nor $\mathbf{R}$ is known, they may be estimated using *model-order identification* techniques presented in *MDL* or *AIC* [47, 48], and an N-snapshot sample covariance matrix, respectively.

We have seen that MUSIC works only when the rank of matrix $\mathbf{R} - \mathbf{R_n}$ is equal to D, that is, when the signals are uncorrelated. Unfortunately, in many multipath scenarios of interest, should the signals incident on the array be strongly or completely correlated, the rank of this matrix would reduce to unity. In this case we will not be able to find the correct noise subspace and the MUSIC algorithm will fail. In order to restore the rank of the matrix $\mathbf{R} - \mathbf{R_n}$, a well-known array processing technique known as *spatial smoothing* may be used [33, 34, 37].

A closed-form expression for the MUSIC AOA asymptotic estimation variance (large N) is given by

$$\begin{aligned} VAR_{\text{MUSIC}}(\hat{\theta}_i) &= \frac{\sigma}{2N} \left\{ [\mathbf{S}^{-1}]_{ii} + \sigma [\mathbf{S}^{-1} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{S}^{-1}]_{ii} \right\} / h(\theta_i), \\ h(\theta) &= \mathbf{a}'(\theta)^H [\mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H] \mathbf{a}'(\theta), \end{aligned} \tag{2.26}$$

where $\mathbf{a}'(\theta) = \frac{d\mathbf{a}(\theta)}{d\theta}$ [40]. Apart from the obvious result that variance increases with the noise power, this equation also shows that variance will increase if the signals are highly correlated (since that would imply that $\mathbf{S}$ is nearly singular), or if their AOAs are closely spaced (which would imply that $\mathbf{A}^H \mathbf{A}$ is nearly singular). Also, Xu et. al. [8, 49] presented results that show that MUSIC may suffer from large bias in scenarios with low SNR, closely spaced sources, or highly correlated sources.

For uncorrelated signals, MUSIC reaches the CRB for large N, M, and SNR scenarios. On the other hand, correlated signals prevent MUSIC from achieving the CRB and may actually cause it to be inefficient even in the presence of large N, M, and SNR cases. Finally, MUSIC is very sensitive to array calibration errors, as has been observed in several works (e.g., [15, 44]).

2.3.3 ESPRIT Algorithm for AOA Estimation

ESPRIT (estimation of signal parameters via rotational invariance techniques) [36] eliminates the array calibration required by MUSIC and offers some computational advantages. It exploits the rotational invariance in the signal subspace that is created by two arrays with a translational invariant structure.

Consider an array with $M+1$ sensors and two M-dimensional array data vectors $\mathbf{r}_L(t) = [r_1(t) \dots r_M(t)]^T$ and $\mathbf{r}_U(t) = [r_2(t) \dots r_{M+1}(t)]^T$, where we recall that $r_m(t)$ is the m-th sensor signal output as given in Equation (2.2). Each of the array data vectors may be expressed as

$$\begin{aligned} \mathbf{r}_L(t) &= \mathbf{A}\mathbf{s}(t) + \boldsymbol{\eta}_L(t), \\ \mathbf{r}_U(t) &= \mathbf{A}\boldsymbol{\Phi}\mathbf{s}(t) + \boldsymbol{\eta}_U(t), \end{aligned} \tag{2.27}$$

where $\boldsymbol{\Phi} = \text{diag}\{e^{jk_1}, e^{jk_2}, \dots, e^{jk_D}\}$ (recall that $k_i = \frac{\omega_c d}{u} \sin\theta_i$ as was defined in Equation (2.3)). Collecting both array responses in a single vector, we obtain

$$\begin{aligned} \bar{\mathbf{r}}(t) &= \begin{bmatrix} \mathbf{r}_L(t) \\ \mathbf{r}_U(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}\boldsymbol{\Phi} \end{bmatrix} \mathbf{s}(t) + \begin{bmatrix} \boldsymbol{\eta}_L(t) \\ \boldsymbol{\eta}_U(t) \end{bmatrix}, \\ \bar{\mathbf{r}}(t) &= \bar{\mathbf{A}}\mathbf{s}(t) + \bar{\boldsymbol{\eta}}(t). \end{aligned} \tag{2.28}$$

The objective is to estimate the elements of $\boldsymbol{\Phi}$ that contain the AOA information without the need to know the array manifold $\mathbf{A}$.

The spatial correlation matrices for $\mathbf{r}_L(t)$, $\mathbf{r}_U(t)$, and $\bar{\mathbf{r}}$ are given, respectively, by

$$\begin{aligned} \mathbf{R}_L &= E\{\mathbf{r}_L(t)\mathbf{r}_L(t)^H\} = \mathbf{A}\mathbf{S}\mathbf{A}^H + \sigma^2\mathbf{I}, \\ \mathbf{R}_U &= E\{\mathbf{r}_U(t)\mathbf{r}_U(t)^H\} = \mathbf{A}\boldsymbol{\Phi}\mathbf{S}\boldsymbol{\Phi}^H\mathbf{A}^H + \sigma^2\mathbf{I}, \\ \bar{\mathbf{R}} &= E\{\bar{\mathbf{r}}(t)\bar{\mathbf{r}}^H(t)\} = \bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^H + \sigma^2\boldsymbol{\Sigma}, \end{aligned} \tag{2.29}$$

where the normalized noise covariance matrix $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{I} & \mathbf{Q}_{-1} \\ \mathbf{Q}_{+1} & \mathbf{I} \end{bmatrix}, \tag{2.30}$$

and $\mathbf{Q}_{-1}$ and $\mathbf{Q}_{+1}$ are $M \times M$ matrices with ones immediately below the main diagonal and zeros elsewhere, and ones immediately above the main diagonal

and zeros elsewhere, respectively. The identity matrix $\mathbf{I}$ is also an $M \times M$ matrix. The D largest eigenvectors of matrices $\mathbf{R}_L$ and $\mathbf{R}_U$ span the signal subspace for each subarray and may be grouped in the columns of matrices $\mathbf{E}_{s,L}$ and $\mathbf{E}_{s,U}$, respectively. Further, solving the generalized eigenvalue problem for matrix $\bar{\mathbf{R}}$ given by

$$\bar{\mathbf{R}}\bar{\mathbf{e}}_i = \bar{\lambda}_i \bar{\mathbf{\Sigma}}\bar{\mathbf{e}}_i \tag{2.31}$$

will yield $2M - D$ smallest generalized eigenvalues equal to σ^2 and D generalized eigenvalues greater than σ^2. The corresponding set of largest generalized eigenvectors may be grouped in the columns of matrix $\bar{\mathbf{E}}_s$ that span the signal subspace for the entire array. Due to the invariance structure of the array, $\bar{\mathbf{E}}_s$ may be decomposed into the signal subspaces $\mathbf{E}_{s,L}$ and $\mathbf{E}_{s,U}$. Since the arrays are translationally related, there should exist a unique nonsingular transformation matrix $\mathbf{\Psi}$ such that

$$\mathbf{E}_{s,L}\mathbf{\Psi} = \mathbf{E}_{s,U}. \tag{2.32}$$

Similarly, a nonsingular transformation matrix $\mathbf{T}$ exists such that

$$\bar{\mathbf{E}}_s = \bar{\mathbf{A}}\mathbf{T} = \begin{bmatrix} \mathbf{E}_{s,L} \\ \mathbf{E}_{s,U} \end{bmatrix} = \begin{bmatrix} \mathbf{AT} \\ \mathbf{A\Phi T} \end{bmatrix}. \tag{2.33}$$

It follows from Equations (2.32) and (2.33) that

$$\mathbf{T\Psi T}^{-1} = \mathbf{\Phi}. \tag{2.34}$$

Clearly, Equation (2.34) has the form of an eigenvalue problem where the columns of matrix $\mathbf{T}$ collect the eigenvectors of matrix $\mathbf{\Psi}$, and the diagonal terms of matrix $\mathbf{\Phi}$ collect the corresponding eigenvalues. Hence, estimation of the transformation matrix $\mathbf{\Psi}$ and its corresponding eigenvalues will yield the AOA estimates embedded in the diagonal terms of $\mathbf{\Phi}$.

A total least-squares (TLS) criterion is usually applied for the estimation of matrix $\mathbf{\Psi}$ [16, 36]. The steps of the TLS-based ESPRIT follow:

1. Obtain an estimate of the total array output correlation matrix $\hat{\bar{\mathbf{R}}}$ from N snapshot observations of the array outputs.
2. Compute the generalized eigenvalue problem,

$$\hat{\bar{\mathbf{R}}}\bar{\mathbf{E}} = \Sigma\bar{\mathbf{E}}\mathbf{\Lambda}, \tag{2.35}$$

   where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \ldots, \lambda_{2M}\}, \lambda_1 \geqslant \ldots \geqslant \lambda_{2M}$, and $\bar{\mathbf{E}} = [\mathbf{e}_1 \ldots \mathbf{e}_{2M}]$ (i.e., the set of all generalized eigenvectors ordered from the largest to the smallest).
3. Estimate the number of sources as the D largest generalized eigenvalues in $\mathbf{\Lambda}$ or with the model order estimation techniques that have been mentioned in previous sections.

4. Obtain the signal subspace matrix $\bar{\mathbf{E}}_s$ (from the D largest generalized eigenvectors) and decompose it to obtain $\mathbf{E}_{s,L}$ and $\mathbf{E}_{s,U}$ as in Equation (2.33).
5. Form a new matrix,

$$\mathbf{C} = \begin{bmatrix} \mathbf{E}_{s,L}^H \\ \mathbf{E}_{s,U}^H \end{bmatrix} [\mathbf{E}_{s,L}\, \mathbf{E}_{s,U}], \tag{2.36}$$

   and perform its eigendecomposition to obtain a set of eigenvalues $\lambda_{c,1} \geqslant \ldots \geqslant \lambda_{c,2D}$ and a corresponding eigenvector set collected in columns of matrix $\mathbf{E}_c$.
6. Partition $\mathbf{E}_c$ such that

$$\mathbf{E}_c = \begin{bmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{bmatrix}. \tag{2.37}$$

7. Estimate the rotation operator $\boldsymbol{\Psi}$ as

$$\boldsymbol{\Psi} = -\mathbf{E}_{12}\mathbf{E}_{22}^{-1}, \tag{2.38}$$

   and calculate its eigenvalues $\{\lambda_{\Psi,1} \ldots \lambda_{\Psi,D}\}$.
8. Finally, estimate the AOAs, noting that $\lambda_{\Psi,i} = |\lambda_{\Psi,i}| e^{j \arg(\lambda_{\Psi,i})}$, and then

$$\theta_i = \sin^{-1}\left[\frac{u}{\omega_c d} \arg(\lambda_{\Psi,i})\right]. \tag{2.39}$$

As discussed in Ottersten et al. [25], the ESPRIT and MUSIC algorithms have comparable performance under most scenarios of interest with the advantage that the former reaches its asymptotic behavior with far fewer snapshots, is robust to array calibration errors, and is computationally simpler since it does not require a search over θ to find AOA estimates.

2.4 TOA AND TDOA MEASUREMENTS

First time of arrival may be directly related to the "time of flight" of a transmitted signal. If the propagation speed of the transmission medium is known, and if a direct line of sight (LOS) exists, then the distance between the transmitter and receiver may be directly related to this time of flight.

The basic problem of TOA-based localization techniques is to estimate the propagation delay of a signal arriving from the direct LOS propagation path.

TDOA measurements may be estimated directly from observations at two separate sensors, or indirectly through two TOA estimates obtained at each sensor. The former requires synchronization between the receivers while the latter requires synchronization between the receivers and the transmitter.

An important property of TOA/TDOA-based ranging techniques is that their variance does not increase with distance as is the case in AOA and RSS-based schemes.

2.4.1 The Time of Arrival Problem

Consider a known signal $s(t)$ transmitted through a single path channel. In this scenario, the received signal may be modeled as

$$r(t) = m \cdot s(t - \tau) + \eta(t), \tag{2.40}$$

where m is a random complex-valued gain, τ is the signal propagation delay, and $\eta(t)$ is a stationary zero mean noise process. The ML estimate of τ can be obtained by finding the value of ζ that maximizes the cross-correlation function between the received and transmitted signals when estimated over an observation window of T seconds [30],

$$\hat{p}_{rs}(\zeta) = \frac{1}{T}\int_0^T r(t)s(t-\zeta)dt = m\hat{p}_{ss}(\zeta - \tau) + \nu(\zeta), \tag{2.41}$$

where ζ is the lag time, $\hat{p}_{ss}(\zeta)$ is the finite sample estimate of the autocorrelation function of the transmitted signal, and $\nu(\zeta)$ is the noise as seen at the output of the correlator. Note that we have used the notation $\hat{x}$ to denote the time average estimate of the corresponding exact statistical moment x. This notation will prove useful in Section 2.4.5 when we introduce the generalized correlator. Then, even though $s(t)$ and its exact autocorrelation function are known, the form of this autocorrelation may differ from the exact one at the output of the correlator in Equation (2.41) when T is not large enough, and for this reason we refer to this finite sample estimate as $\hat{p}_{ss}(\zeta)$.

Now consider a known signal $s(t)$ that is transmitted through a multipath channel with impulse response given by

$$h(t) = \sum_{i=1}^{D} m_i \delta(t - \tau_i), \tag{2.42}$$

where D is the number of existing paths, $\{\tau_i\,;\ i = 1, \ldots, D\}$ are the time-delay parameters or the times of arrival of the echoes that include the first time of arrival τ_1, which we want to estimate, and $\{m_i\,;\ i = 1, \ldots, D\}$ are random complex-valued gains that include the scattering and propagation fading effects. In a period of time starting at the moment the signal is transmitted and ending at a time when it is assumed that no more echoes will arrive, a sensor receives a superposition of delayed and scaled versions of $s(t)$ and the received-signal baseband model can be written as

$$r(t) = \sum_{i=1}^{D} m_i s(t - \tau_i) + \eta(t), \tag{2.43}$$

where $\eta(t)$ is an additive white-noise process with a spectral level of σ^2.

The previously described ML cross-correlation scheme may be applied to the received signal to obtain

$$\hat{p}_{rs}(\zeta) = \sum_{i=1}^{D} m_i \hat{p}_{ss}(\zeta - \tau_i) + \nu(\zeta), \tag{2.44}$$

which corresponds to a weighted sum of multiple shifted-signal autocorrelation functions. The correct estimation of the first time of arrival in the multipath scenario will strongly depend on the instantaneous channel response at the time of the measurements, and the signal bandwidth. The performance of the ML estimator presented in Equation (2.44) will degrade considerably whenever the first TOA has a strong fade (due to shadowing or small-scale fading) or when the delay differences between adjacent paths and the first TOA are smaller than the duration of the main signal's autocorrelation peak. The solution to these problems may be achieved by increasing the signal bandwidth (which in turn will decrease the autocorrelation peak duration), or by applying other high-resolution signal processing techniques that will be described shortly.

2.4.2 The Time Difference of Arrival Problem

In the TDOA problem, a signal $s(t)$ transmitted at a remote location is received by two spatially separated receivers. For the case of a flat fading channel, the signals at the two receivers may be described as

$$\begin{aligned} r_a(t) &= s(t) + \eta_a(t), \\ r_b(t) &= ms(t+\tau) + \eta_b(t), \end{aligned} \tag{2.45}$$

where again m is a random complex-valued gain, and now τ corresponds to the time difference with which the signal arrived at each of the two receivers. It is assumed that $\eta_a(t)$ and $\eta_b(t)$ are stationary noise processes. A correlation procedure similar to the one presented for the TOA case may be implemented to estimate τ [23] as follows:

$$\hat{p}_{r_a r_b}(\zeta) = \frac{1}{T}\int_0^T r_a(t) r_b(t-\tau) dt = m\hat{p}_{ss}(\zeta - \tau) + \nu(\zeta), \tag{2.46}$$

where again $\hat{p}_{ss}(\zeta)$ is the finite sample estimate of the autocorrelation of $s(t)$, and $\nu(\zeta)$ includes the noise terms as seen at the output of the correlator. Then the TDOA may be estimated by finding the value of τ that maximizes Equation (2.46).

In the multipath channel case, the signals received at each sensor may be expressed as

$$\begin{aligned} r_a(t) &= \sum_{i=1}^{D} m_{a,i}\, s(t-\tau_{a,i}) + \eta_a(t), \\ r_b(t) &= \sum_{i=1}^{D} m_{b,i}\, s(t-\tau_{b,i}) + \eta_b(t), \end{aligned} \tag{2.47}$$

where subindexes a and b are used to distinguish the parameters corresponding to each of the multipath channels (corresponding to the signal travel from the transmitter to each of the a and b receivers). The estimated cross-correlation function of the received signals is calculated as

$$\hat{p}_{r_a r_b}(\zeta) = \sum_{i=0}^{D}\sum_{j=0}^{D} m_{a,i}\, m_{b,j}\, \hat{p}_{ss}(\zeta - (\tau_{a,i} - \tau_{b,i})) + \nu(\zeta). \tag{2.48}$$

The TDOA estimate will consist of the difference $\tau_{a,1} - \tau_{b,1}$. As stated in Li et al. [23], it is clear in this scenario that the TDOA cannot be detected by finding the delay value that corresponds to the first peak of the cross-correlation function since it will not necessarily correspond to the delay value $\tau_{a,1} - \tau_{b,1}$. Further, the strongest peak of the cross-correlation will not necessarily occur at $\tau_{a,1} - \tau_{b,1}$. Then estimation of the TDOA becomes quite difficult in the presence of multipath since the correlation peak at the true TDOA is not guaranteed to be the first one or the strongest. For this reason, in many applications where the transmitted signal is known a priori, TDOA is measured based on two separate TOA measurements obtained at each receiver. This scheme avoids measurement ambiguities but requires synchronization between the transmitter and both of the receivers. Direct TDOA measurements, on the other hand, only require synchronization between the receivers.

2.4.3 Performance Bound for TOA and TDOA Problems

Lower bounds for the estimation variance of delay in the neighborhood of the true time delay were described in Quazi [30] (and references therein) for the single-path propagation channel model described in Equations (2.40) and (2.45). The derivations assume that the signal and noise spectral densities are constant over the bandwidth $W = f_2 - f_1$ Hz with power of P_0 and σ^2, respectively. The bounds are given by

$$\begin{aligned} \sigma_\tau &\geqslant \frac{1}{\sqrt{8\pi^2}}\frac{1}{SNR}\frac{1}{TW}\frac{1}{f_c}\frac{1}{\sqrt{1+\frac{W^2}{12f_c^2}}}, \quad SNR \ll 1, \\ \sigma_\tau &\geqslant \frac{1}{\sqrt{2\pi^2}}\frac{1}{\sqrt{SNR}}\frac{1}{TW}\frac{1}{f_c}\frac{1}{\sqrt{1+\frac{W^2}{12f_c^2}}}, \quad SNR \gg 1. \end{aligned} \tag{2.49}$$

The bounds in Equation (2.49) show that TOA estimation variance is inversely proportional to $SNR = P_0/\sigma^2$, bandwidth time (BT) product TW, and center carrier frequency f_c.

Lower bounds for the variance of estimates of delay in the presence of multipath have been presented in [4, 5], where it has been shown that first TOA estimation accuracy is strongly dependent on delay separation and relative phase between the first path and its closer clutter of arrivals.

2.4.4 Received Signal Model and Its Analogy to the Array Processing Problem

Consider repeating the transmission of a signal $s(t)$ every T seconds through a D-path channel, such as the one described by the impulse response in Equation (2.42), to obtain N responses (referred to as *snapshots*). We assume that after T seconds, no more echoes arrive so that each snapshot has no interference from previous transmissions. It is also assumed that the path delay locations $\{\tau_i;\ i = 1, \ldots, D\}$ do not change during the N-snapshot acquisition period. Then, from Equation (2.43), each snapshot may be expressed as

$$r_j(t) = \sum_{i=1}^{D} m_{ij} s(t - \tau_i) + \eta_j(t) \quad j = 1, 2, \ldots, N, \tag{2.50}$$

where m_{ij} stands for the complex gain of the i-th path as observed at the time when the j-th snapshot was captured.

The TOA estimation problem can be stated as follows: Given a set of snapshots as above, and assuming that we know the number of echoes D (they can be estimated using the MDL or AIC methods [47, 48]), find the time delays corresponding to each return. Note that contrary to the AOA problem where we had multiple sensors, observations for the TOA problem consist of a sequence of N snapshots from a *single sensor*. The analogy to the M sensors is introduced if we sample the received waveforms $r_j(t)$ at M instants and collect the successive samples in a vector of length M. Then for each snapshot we can write in vectorial notation [7]:

$$\begin{bmatrix} r_j(t_1) \\ r_j(t_2) \\ \vdots \\ r_j(t_M) \end{bmatrix} = [\mathbf{a}(\tau_1)\,\mathbf{a}(\tau_2) \cdots \mathbf{a}(\tau_D)] \begin{bmatrix} m_{1j} \\ m_{2j} \\ \vdots \\ m_{Dj} \end{bmatrix} + \begin{bmatrix} \eta_j(t_1) \\ \eta_j(t_2) \\ \vdots \\ \eta_j(t_M) \end{bmatrix}, \tag{2.51}$$

$$\mathbf{r}_j = \mathbf{A}\mathbf{g}_j + \eta_j, \tag{2.52}$$

where now the "steering vectors" are defined as

$$\mathbf{a}(\tau) = [s(t_1 - \tau)\, s(t_2 - \tau) \ \ldots \ s(t_M - \tau)]^T, \tag{2.53}$$

and the correlation matrix is given by

$$\mathbf{R} = E\{\mathbf{r}_j \mathbf{r}_j^H\} = \mathbf{AGA}^H + \mathbf{R_n}, \tag{2.54}$$

where $\mathbf{A} = [\mathbf{a}(\tau_1)\,\mathbf{a}(\tau_2)\ldots\mathbf{a}(\tau_D)]$ is the "array manifold" matrix, and $\mathbf{G}$ and $\mathbf{R_n} = \sigma^2\mathbf{I}$ are the gain and noise correlation matrices, respectively. For each τ, the steering vector can be calculated in a simple way since we know the signal form. In this formulation it is clear that the rank of the matrix $\mathbf{R} - \mathbf{R_n}$ will be equal to D, provided that $M > D$ and that the signal echoes are uncorrelated. For instance, if the channel path gains $\{m_i;\ i = 1, \ldots, D\}$ are random and change rapidly and independently with time, the path decorrelation assumption will hold. It is clear that Equations (2.5) and (2.52) and (2.7) and (2.54) have the same structure.

So the AOA estimation algorithms presented in Section 2.1 may be readily applied to the *one-sensor time delay estimation problem*, provided that the received signal correlation matrix and "steering vectors" are replaced by those given in Equations (2.54) and (2.53), respectively. Apart from the techniques presented in Section 2.1, there are other algorithms that profit from the time delay information obtained with the correlation function between a signal and its delayed counterpart. In the following we will discuss several of these algorithms; for instance, we will present the conventional pseudo-noise (PN) correlation method and its super-resolution counterpart implemented with MUSIC. We will also discuss TOA estimation techniques based on generalized cross-correlation and successive cancellation.

2.4.5 Generalized Cross-Correlation Method for TOA or TDOA Estimation

Let us return to the received signal model presented in Equation (2.45). As mentioned previously, an optimum delay estimator consists of a cross-correlator given by Equation (2.46). Note, however, that if the observation window length is limited (i.e., T may not be arbitrarily large), then the cross-correlation function estimator may become inaccurate, especially in low SNR scenarios. The length of the observation window may be limited for various reasons, and usually the most significant relates to the coherence time of the channel in which the signal and noise may be assumed to be stationary. Noise and finite observation times are then a major source of error in the cross-correlation estimator.

Another important limitation of the estimator may be related to signal bandwidth. A small bandwidth will yield an autocorrelation function that is broad around its peak at the zero lag. These broad autocorrelation peaks will overlap in multipath scenarios to yield new peaks at incorrect delay locations. To improve performance under the previously described scenarios, a *generalized*

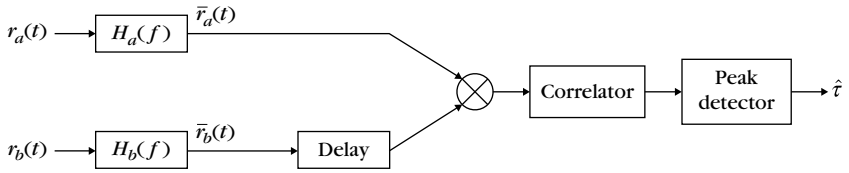

FIGURE 2.3

Generalized correlator.

correlation (GC) method may be implemented instead of the simple correlator. The results presented in this section were obtained from the works of Knapp and Carter [19] and Hassab and Boucher [17]. The GC may be implemented as shown in Figure 2.3 [19]. Prior to the cross-correlation operation, received signals $r_a(t)$ and $r_b(t)$ may be filtered through $H_a(f)$ and $H_b(f)$, respectively. The resulting signals $\bar{r}_a, \bar{r}_b$ are multiplied, integrated, and squared for a range of time shifts ζ, until a peak is obtained at a specific delay and is chosen as the estimate of τ. Note that when the filter's transfer functions are $H_a(f) = H_b(f) = 1$, we return to the simple cross-correlation form in Equation (2.46). The cross-correlation between the received signals may be related to the cross-power spectral density $G_{r_a r_b}(f)$ by [29],

$$p_{r_a r_b}(\zeta) = \int_{-\infty}^{\infty} G_{r_a r_b}(f) e^{j2\pi f\zeta} df. \tag{2.55}$$

After the signals have been filtered, the cross-power spectrum becomes

$$\bar{G}_{r_a r_b}(f) = H_a(f) H_b(f)^* G_{r_a r_b}(f). \tag{2.56}$$

Then the GC between $r_a(t)$ and $r_b(t)$ is given by

$$\bar{p}_{\bar{r}_a \bar{r}_b}(\zeta) = \int_{-\infty}^{\infty} \psi(f) G_{r_a r_b}(f) e^{j2\pi f\zeta} df, \tag{2.57}$$

where $\psi(f) = H_a(f) H_b(f)^*$. Note that in practice, the cross-spectral density between signals $r_a(t)$ and $r_b(t)$ is not known and needs to be estimated from a finite number of observations.

Let us explain the motivation for the use of a GC for estimation of delay. Without making any assumptions about the signal and noise statistics, the statistical cross-correlation between signals $r_a(t)$ and $r_b(t)$ is calculated as

$$\begin{aligned} p_{r_a r_b}(\zeta) &= E\{r_a(t) r_b(t-\zeta)\} = m \cdot p_{ss}(\zeta - \tau) + E\{\eta_a(t)\eta_b(t-\zeta)\} + \phi_s(\zeta), \\ p_{r_a r_b}(\zeta) &= [m \cdot p_{ss}(\zeta) \otimes \delta(\zeta - \tau)] + p_{\eta_a \eta_b}(\zeta) + \phi_s(\zeta), \end{aligned} \tag{2.58}$$

where $\phi_s(\zeta) = E\{s(t)\eta_b(t-\zeta) + m \cdot s(t+\tau-\zeta)\eta_b(t)\}$. Let us analyze the terms on the right side of the second row of Equation (2.58). The first term corresponds to an impulse at the true delay τ, which is being spread by the signal's auto-correlation function. Due to the fact that $p_{ss}(\zeta) < p_{ss}(0)$ [29], this term will in fact peak at the true delay but this peak will be broadened according to the form of the signal auto-correlation function. This broadening will become a serious problem in dense multipath scenarios (where paths will be closely spaced in time) because multiple broad peaks will overlap and cause peaks at incorrect delay locations. Finally, note that if the signal is white, then it will not cause any spreading. The second and third terms are due to the noise processes $\eta_a(t)$ and $\eta_b(t)$. If these terms become large, they may cause large peaks (larger than the signal peak) in $p_{r_a r_b}(\zeta)$, which will yield erroneous estimates of τ.

Note that in the ideal case where the noise processes are uncorrelated and independent from signal $s(t)$, these terms will become zero. Note, however, that even in this ideal scenario, the time smearing of the impulse at delay τ due to the form of the signal's autocorrelation function will not disappear. Thus, one motivation for the use of a GC is that one may chose a frequency weighting factor $\psi(f)$ to maximize the amplitude and "sharpness" of the generalized cross-correlation peaks while maintaining low estimation variances. The other motivation arises when the noise terms in Equation (2.58) are not equal to zero. Non-zero noise terms may arise, even when the uncorrelated noise processes assumption holds, when the exact statistical averaging of Equation (2.58) is replaced by finite sample time averaging (which will always be the case in "real-world" applications). Obviously, noise terms will also appear when $\eta_a(t)$ and $\eta_b(t)$ are correlated. In the presence of noise terms, the frequency weighting factor $\psi(f)$ may be chosen to filter out their spectral components from the cross-spectral density $G_{r_a r_b}(f)$ and thus minimize the occurrence of false peaks.

Several weighting functions have been presented in the literature [1, 2, 17, 19], and some are listed in Table 2.1. To give insight about the rationale of the GC estimator, we will discuss two methods: Roth [35] and phase transform (PHAT) [19]. The reader may consult [19] and references therein for more detail on other methods. In practice, the cross-spectral density $G_{r_a r_b}(f)$ will need to be replaced by its corresponding estimate.

The Roth Generalized Cross-Correlator

As shown in Table 2.1, Roth's GC sets the frequency-weighting function to $\psi(f) = \frac{1}{G_{r_a r_a}(f)}$. Substitution of this function in Equation (2.57) yields

$$\hat{\hat{p}}_{\bar{r}_a \bar{r}_b}(\zeta) = \int_{-\infty}^{\infty} \frac{\hat{G}_{r_a r_b}(f)}{G_{r_a r_a}(f)} e^{j2\pi f \zeta} df. \tag{2.59}$$

Table 2.1 Frequency Weighting for $\psi(f)$ for Generalized Cross-Correlation

Method	$\psi(f)$
Cross-Correlation	1
Roth	$\frac{1}{G_{r_a r_a}(f)}$
SCOT	$\frac{1}{[G_{r_a r_a}(f) G_{r_b r_b}(f)]^{1/2}}$
PHAT	$\frac{1}{\lvert G_{r_a r_b}(f)\rvert}$
Eckart	$\frac{G_{ss}(f)}{G_{\eta_a \eta_a}(f) G_{\eta_b \eta_b}(f)}$
ML	$\frac{\lvert\gamma_{ab}(f)\rvert^2}{\lvert G_{r_a r_b}(f)[1-\gamma_{ab}(f)]\rvert^2}$
	$\gamma_{ab}(f) = \frac{G_{r_a r_b}(f)}{[G_{r_a r_a}(f) G_{r_b r_b}(f)]^{1/2}}$

This last equation is an estimator of the optimum Wiener-Hopf filter, which best approximates the mapping of $r_a(t)$ to $r_b(t)$. In the presence of noise

$$G_{r_a r_a}(f) = G_{ss}(f) + G_{\eta_a \eta_a}(f). \tag{2.60}$$

Substituting this last equation, and Equation (2.58) into (2.59), and assuming uncorrelated noise processes η_a and η_b, and $\phi_s(\zeta) = 0$, we obtain

$$\bar{p}_{\bar{r}_a \bar{r}_b}(\zeta) = \delta(\zeta - \tau) \otimes \int_{-\infty}^{\infty} \frac{m G_{ss}(f)}{G_{ss}(f) + G_{\eta_a \eta_a}(f)} e^{j2\pi f \zeta} df. \tag{2.61}$$

Note that the delta function will be broadened except for the case when $G_{\eta_a \eta_a}(f) = \text{const}\, G_{ss}(f)$. The Roth GC has the effect of suppressing frequency regions where $G_{\eta_a \eta_a}(f)$ is large, and where the cross-spectral density estimate $\hat{G}_{r_a r_b}(f)$ is more likely to be in error.

The Phase Transform Generalized Cross-Correlator

The PHAT frequency weighting function is given by $\psi(f) = \frac{1}{|G_{r_a r_b}(f)|}$. This yields the following GC:

$$\hat{\bar{p}}_{\bar{r}_a \bar{r}_b}(\zeta) = \int_{-\infty}^{\infty} \frac{\hat{G}_{r_a r_b}(f)}{|G_{r_a r_b}(f)|} e^{j2\pi f \zeta} df. \tag{2.62}$$

Then, again assuming uncorrelated noise processes η_a and η_b, and $\phi_s(\zeta) = 0$, we find that

$$|G_{r_a r_b}(f)| = \alpha G_{ss}(f), \tag{2.63}$$

and for the case of perfect estimation of the cross-spectral density, $G_{r_a r_b}(f)$,

$$\frac{G_{r_a r_b}(f)}{|G_{r_a r_b}(f)|} = e^{j2\pi f \tau} \tag{2.64}$$

and

$$\bar{p}_{\bar{r}_a\bar{r}_b}(\zeta) = \delta(\zeta - \tau), \tag{2.65}$$

which is independent of the form of the signal autocorrelation function. Note, however, that when perfect estimation of $G_{r_a r_b}(f)$ is not available, the PHAT GC estimation peak will still be broadened by an amount proportional to the deviation from the ideal cross-spectral density.

2.4.6 Conventional PN-Correlation Method

Recall that the variance lower bounds in Equation (2.49) show that large BT products are required for better performance. Large BT products are desirable when estimating the delay TOAs of a transmitted signal in a multipath channel in order to cover a large interval of frequencies over a large observation window. The large bandwidth is desired so that all frequency components of the channel are perturbed. Large bandwidth usually implies the ability to resolve multipath arrivals separated by $|\tau_2 - \tau_1| \geqslant \frac{1}{B}$. The large duration is desired because longer signals usually yield higher processing gains. However, transmitting a long duration signal implies that the observation window time T and the number of samples M will be large and will yield large covariance matrix dimensions together with larger computational complexity. This limitation can be solved using the autocorrelation properties of a PN-sequence. A PN-sequence is a long duration broadband signal with a very large BT product but with a short duration (peaky) autocorrelation function [28]. Figure 2.4 shows the *maximal-length shift register* ($m = 7$) PN-sequence autocorrelation function. The chip rate of this sequence was set to 10 MHz. It is observed that the duration of the main autocorrelation function peak is about two times the chip interval T_c, which is the inverse of the chip rate, in this case $T_c = 100$ ns.

Figure 2.5 shows a conventional time-delay estimation method that involves the transmission of a wide-band PN-sequence $c(t)$ and subsequent correlation of the sampled, filtered, and received signal with this sequence [9, 12, 31, 32]. A subset of the local extrema of the resultant magnitude of the correlator output is selected as the multipath delays, and the first delay or the strongest delay may be chosen as the first TOA estimate. Due to the triangular spread of the PN-sequence autocorrelation function, arrivals separated by less than T_c seconds will not be resolved. Thus, the resolution of this method is limited by the chip rate of the signal $c(t)$.

Figure 2.6 shows the power delay profiles, defined as $|z(\zeta)|^2$, of two different two-path channels ($D = 2$). The first channel has delays at $D = [25\,225]$ ns,

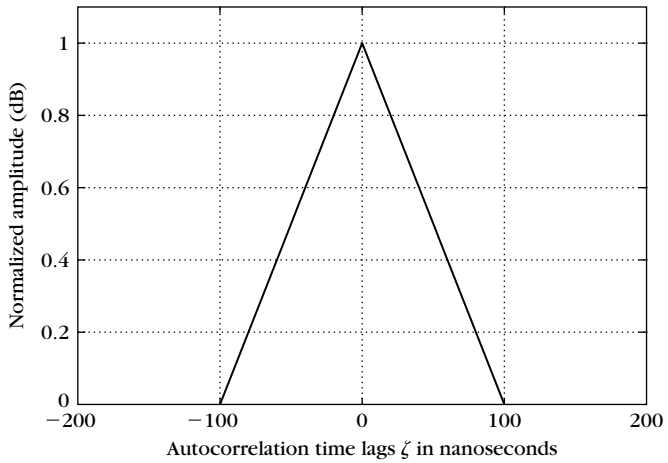

FIGURE 2.4

An $m = 7$ maximal PN-sequence autocorrelation function. $T_C = 100$ ns.

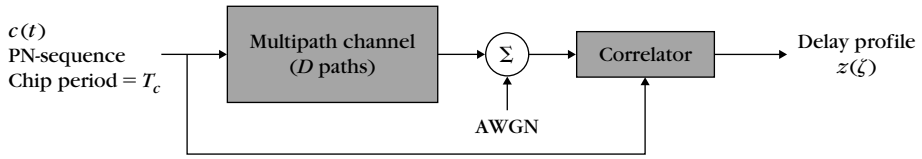

FIGURE 2.5

Conventional measurement of a baseband, multipath channel-delay profile.

a separation of $2\,T_c$. The second channel has delays at $D = [25\,75]$ ns, a separation of $0.5\,T_c$. For this example, $T_c = 100$ ns. Clearly, the conventional PN-correlation method fails to resolve the delays of the second channel. In this scenario, the method yields a first TOA estimate of 50 ns (first peak of the correlation function), which is 25 ns away from the true value. An obvious and easy solution to this limited resolution problem is to increase the chip rate of $c(t)$ to reduce the base width of the autocorrelation function; however, this is not always feasible since we may have bandwidth limitations or hardware complexity constraints.

Another problem faced by the conventional delay profile measurement is that of mutual path cancellations due to differences between their relative phases. These cancellations cause the peaks of the delay profile to move or to completely disappear from the true delay locations.

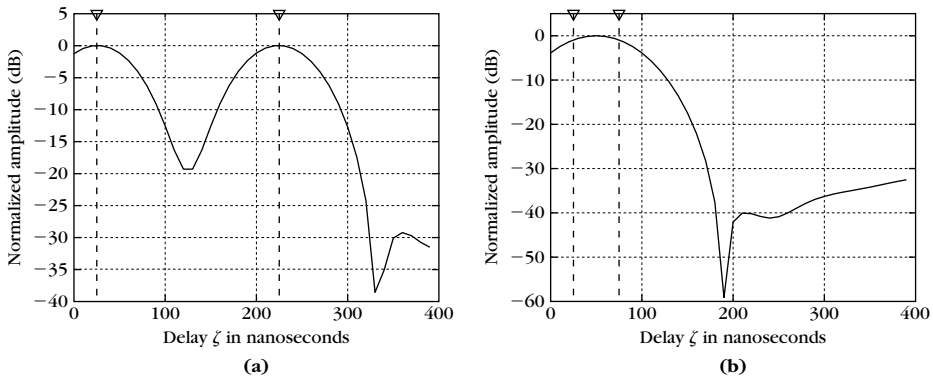

FIGURE 2.6

Power delay profiles of two different two-path channels. The first channel has delays at $D = [25\ 225]$ ns, a separation of $2\ T_c$. The second channel has delays at $D = [25\ 75]$ ns, a separation of $0.5\ T_c$. For this example, $T_c = 100$ ns.

A solution that allows us to keep the signal chip rate untouched and that avoids the path cancellations due to their relative phase difference consists of applying a super-resolution method to the received delay profile $z(\zeta)$.

2.4.7 A Super-Resolution PN-Correlation Method: The SPM Algorithm

This method was presented in [24, 26], and its performance under real transmission conditions was analyzed in [3]. Consider transmitting a signal through a multipath channel consisting of D discrete paths with various time delays $\{\tau_i;\ i = 1, \ldots, D\}$ that we want to estimate, and with an impulse response as given in Equation (2.42). If the baseband modulation signal is a PN-sequence $c(t)$ and ω_c is the carrier frequency, then the baseband received signal is given by

$$r(t) = \sum_{i=1}^{D} m_i c(t - \tau_i) + \eta(t), \tag{2.66}$$

where it is assumed that $\eta(t)$ is a stationary, zero mean, additive white-noise process with variance equal to σ^2. The cross-correlation of $r(t)$ with $c(t)$ is

$$z(\zeta) = \frac{1}{T} \int_0^T c(t - \zeta) r(t)^* dt, \tag{2.67}$$

where T was previously described as the snapshot measurement window that starts at the beginning of each PN-sequence transmission and ends

when no more echoes arrive at the receiver. Substituting Equation (2.66) into Equation (2.67) yields

$$z(\zeta) = \sum_{i=1}^{D} m_i p(\zeta - \tau_i) + \nu(\zeta), \tag{2.68}$$

where

$$p(\zeta) = \frac{1}{T} \int_0^T c(t-\zeta) c(t)^* dt, \tag{2.69}$$

$$\nu(\zeta) = \frac{1}{T} \int_0^T c(t-\zeta) \eta(t)^* dt. \tag{2.70}$$

Thus, we have a new set of functions (autocorrelation and cross-correlation functions) that are defined in the lag domain ζ. It is clear that the delay information we are looking for is inherent in Equation (2.68). So it is very reasonable to consider the signal at the output of the correlator (i.e., the delay profile $z(\zeta)$ signal) as our new received signal and apply the MUSIC algorithm to this new type of data.

Let us sample the delay profile signal at M lags $\{\zeta_1, \zeta_2, \ldots, \zeta_M\}$ to obtain a delay profile vector $\mathbf{z} = [z(\zeta_1)\, z(\zeta_2) \ldots z(\zeta_M)]^T$. It is assumed that the sampling interval is chosen to be a small fraction of the chip interval T_c. From Equation (2.68),

$$\mathbf{z} = \sum_{i=1}^{D} m_i \mathbf{p}(\tau_i) + \boldsymbol{\nu}, \tag{2.71}$$

where our new "steering vector," $\mathbf{p}(\tau)$, is defined as

$$\mathbf{p}(\tau) = [p(\zeta_1 - \tau)\, p(\zeta_2 - \tau) \ldots p(\zeta_M - \tau)]^T. \tag{2.72}$$

This steering vector is easy to calculate for any τ since the autocorrelation function $p(\zeta)$ is known. The new noise vector is given by $\boldsymbol{\nu} = [\nu(\zeta_1)\ \nu(\zeta_2) \ldots \nu(\zeta_M)]^T$. If we define the array manifold matrix and the path gain vector, respectively, as

$$\boldsymbol{\Gamma} = [\mathbf{p}(\tau_1)\ \mathbf{p}(\tau_2)\ \ldots\ \mathbf{p}(\tau_D)], \tag{2.73}$$

and

$$\mathbf{g} = [m_1\ m_2 \ldots m_D]^T, \tag{2.74}$$

then we can write the delay profile vector in matrix notation as

$$\mathbf{z} = \boldsymbol{\Gamma}\mathbf{g} + \boldsymbol{\nu}. \tag{2.75}$$

Finally, the correlation matrix of this data vector $\mathbf{z}$ is expressed as

$$\mathbf{R} = E\{\mathbf{z}\mathbf{z}^H\} = \sum_{i,j} E\{g_i g_j^*\} \mathbf{p}(\tau_i) \mathbf{p}(\tau_j)^H + E\{\boldsymbol{\nu}\boldsymbol{\nu}^H\}. \tag{2.76}$$

In matrix notation,

$$\mathbf{R} = \mathbf{\Gamma} G \mathbf{\Gamma}^H + \mathbf{R_n}, \tag{2.77}$$

where G is the path gain covariance matrix. The noise correlation matrix can be expressed as

$$E\{\boldsymbol{\nu}\boldsymbol{\nu}^H\} = \mathbf{R_n} = \sigma^2 \mathbf{R_0}. \tag{2.78}$$

Note that

$$E\{\nu(\zeta_k)\nu(\zeta_l)^*\} = \left\{ \frac{1}{T} \int_0^T E\{\eta(t)\eta(t)^*\} dt \right\} \left\{ \frac{1}{T} \int_0^T c(t-\zeta_k)c(t-\zeta_l)^* dt \right\} = \sigma^2 p(\zeta_k - \zeta_l), \tag{2.79}$$

so $\mathbf{R_0}$ is a Hermitian matrix whose kl-th element is equal to $p(\zeta_k - \zeta_l)$, and

$$[\mathbf{R_n}]_{k,l} = \sigma^2 p(\zeta_k - \zeta_l). \tag{2.80}$$

From the noise covariance matrix, we observe that now we deal with a colored noise process, so a whitening procedure should be implemented. This may be achieved by applying the linear transformation $\bar{\mathbf{z}} = \mathbf{R_0}^{-\frac{1}{2}} \mathbf{z}$ so that $\bar{\mathbf{R}} = E\{\bar{\mathbf{z}}\bar{\mathbf{z}}^H\} = \mathbf{R_0}^{-\frac{1}{2}} \mathbf{R} \mathbf{R_0}^{-\frac{1}{2}} = [\mathbf{R_0}^{-\frac{1}{2}} \mathbf{\Gamma}] \mathbf{G} [\mathbf{\Gamma}^H \mathbf{R_0}^{-\frac{1}{2}}] + \sigma^2 \mathbf{I}$. Then solving the generalized eigenvalue problem for the whitened observations yields

$$\bar{\mathbf{R}} \bar{\mathbf{e}}_i = [\mathbf{R_0}^{-\frac{1}{2}} \mathbf{R} \mathbf{R_0}^{-\frac{1}{2}}] \bar{\mathbf{e}}_i = \lambda_i \bar{\mathbf{e}}_i, \tag{2.81}$$

which, clearly, leads to the generalized eigenvalue problem of the form

$$\mathbf{R} \mathbf{e_i} = \lambda_i \mathbf{R_0} \mathbf{e}_i. \tag{2.82}$$

It may be concluded that solving the generalized eigenvalue problem in Equation (2.82) is equivalent to whitening the noise [45].

Again we see that the structure of Equations (2.75) and (2.77) is identical to Equations (2.5) and (2.7), respectively. Table 2.2 shows the analogies of the three estimation cases we have discussed in this section and the previous one.

We should make it clear that to apply MUSIC to this new model, the steering vectors $\{\mathbf{p}(\tau_1); \ldots; \mathbf{p}(\tau_D)\}$ should form a linearly independent set so that $\mathbf{\Gamma}$ is full rank. Since the assumption of linear independence between steering vectors is vital for SPM, let us illustrate the degree of dependence now by creating a two-path channel with delays at τ_1 and $\tau_1 + \Delta\tau$ seconds. Using a 10-MHz, chip-rate m-sequence ($T_c = 100$ ns), we calculate a set of array manifold matrices $\mathbf{\Gamma}$ and plot the condition number of $\mathbf{\Gamma}^H \mathbf{\Gamma}$ as a function of $\Delta\tau$. Figure 2.7 shows this plot. As expected, for $\Delta\tau = 0$ the smallest eigenvalue is zero since the rank of $\mathbf{\Gamma}$ becomes one. However, for values of $\Delta\tau > 0$, the smallest eigenvalue has

Table 2.2 Analogies between Methods

Method	AOA	Time Delay	SPM
Signal	$r_m(t) = \sum_{i=1}^{D} s_i(t) e^{-j(m-1)k_i} + \eta$	$r(t) = \sum_{i=1}^{D} m_i s(t - \tau_i) + \eta$	$z(\zeta) = \sum_{i=1}^{D} m_i p(\zeta - \tau_i) + \nu$
M	Number of sensors	Number of time samples	Number of lag samples
D	Number of impinging waves	Number of channel paths	Number of channel paths
Steering vector	$\mathbf{a}_k = [1\, e^{jk} \ \ldots \ e^{(M-1)jk}]^T$	$\mathbf{a}(\tau) = [s(t_1 - \tau) \ \ldots \ s(t_M - \tau)]^T$	$\mathbf{p}(\tau) = [p(\zeta_1 - \tau) \ \ldots \ p(\zeta_M - \tau)]^T$
Model	$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \boldsymbol{\eta}(t)$	$\mathbf{r} = \mathbf{A}\mathbf{g} + \boldsymbol{\eta}$	$\mathbf{z} = \boldsymbol{\Gamma}\mathbf{g} + \boldsymbol{\nu}$
Correlation matrix	$\mathbf{A}\mathbf{S}\mathbf{A}^H + \mathbf{R_n}$	$\mathbf{A}\mathbf{G}\mathbf{A}^H + \mathbf{R_n}$	$\boldsymbol{\Gamma} G \boldsymbol{\Gamma}^H + \mathbf{R_n}$
Noise-correlation matrix	$\mathbf{R_n} = \sigma^2 \mathbf{I}$	$\mathbf{R_n} = \sigma^2 \mathbf{I}$	$\mathbf{R_n} = \sigma^2 \mathbf{R_0}$
Parameter of interest	$\{\theta_i;\ i = 1, \ldots, D\}$	$\{\tau_i;\ i = 1, \ldots, D\}$	$\{\tau_i;\ i = 1, \ldots, D\}$

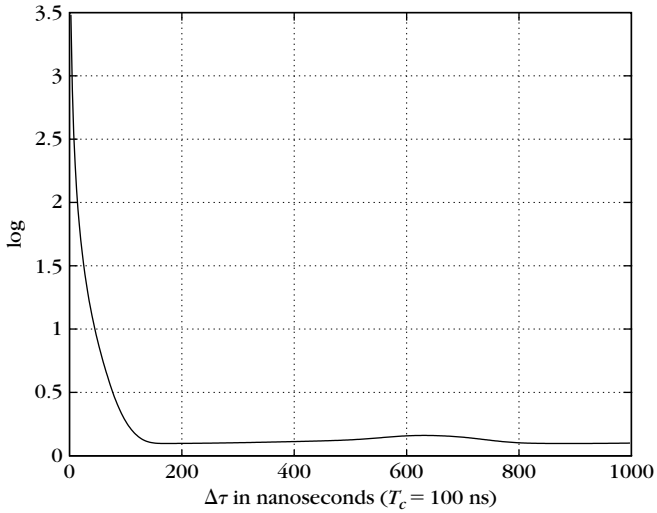

FIGURE 2.7

Condition number of the matrix $\mathbf{\Gamma}^H\mathbf{\Gamma}$ as a function of the delay difference $\Delta\tau$ seconds of a two-path channel.

non-zero values that grow very fast with respect to $\Delta\tau$. This indicates that, at least in certain conditions, $\mathbf{\Gamma}$ is a full-rank matrix when the delay difference is a small fraction of the chip period [3].

In all our discussions thus far, once the correlation matrix was obtained, we could directly apply MUSIC since we assumed that the multipath return signals were uncorrelated—that is, that the path-gain correlation matrix **G** was a diagonal matrix. This is usually not true, especially in slow time-varying channels where the signals will be correlated since the channel path magnitudes and relative phases will remain constant during the acquisition of the N snapshots. This leads to a rank-deficient matrix $\mathbf{R} - \mathbf{R_n}$. To solve this problem, a decorrelating technique similar to spatial smoothing, called *frequency smoothing*, may be used to restore the rank of $\mathbf{R} - \mathbf{R_n}$. This technique consists of transmitting the PN-sequence using distinct carrier frequencies [3, 24].

Assuming decorrelated multipath arrivals, MUSIC may be applied as follows to obtain the SPM algorithm [24]:

1. Estimate the signal correlation matrix **R** using N snapshots.
2. Find the generalized eigenvalues and eigenvectors of $\hat{\mathbf{R}}$.
3. Find the noise subspace spanned by $\{\mathbf{e}_i;\ i = D+1, D+2, \ldots, M\}$, which are the generalized eigenvectors corresponding to the $M-D$ smallest eigenvalues.

4. Calculate the MUSIC spectrum, that is, the *super-resolution delay profile* (SDP):

$$\text{SDP}(\tau) = \frac{\mathbf{p}(\tau)^H \mathbf{R_0}^{-1} \mathbf{p}(\tau)}{\sum_{i=D+1}^{M} |\mathbf{p}(\tau)^H \mathbf{e_i}|^2}. \tag{2.83}$$

The SDP will ideally peak to infinity when the true delay values $\{\tau_i\,;\ i = 1, \ldots, D\}$ of the channel are input into this function. The first TOA estimate may be chosen as the first peak in the SDP.

In an analogy to the array calibration errors in the MUSIC-based AOA estimation problem, MUSIC-based TOA estimation will suffer from unknown filter transfer functions at the transmitter and receiver, which may distort the form of the correlation function and change the form of the "steering vectors" [24]. A way to overcome this problem is to include the filtering distortions in the calculation of the steering vectors $\mathbf{p}(\tau)$. If $h_F(t)$ is the cascade impulse response of all the filtering operations on the signal prior to the SPM processing, then we can modify Equation (2.69) as

$$\tilde{p}(\zeta) = \frac{1}{T} \int_0^T c(t-\zeta)[h_F(t) \otimes c(t)]dt = h_F(\zeta) \otimes p(\zeta), \tag{2.84}$$

and calculate the steering vector in Equation (2.72) using $\tilde{p}(\zeta)$ instead of $p(\zeta)$.

The SDP of four different uncorrelated two-path channels with $|\tau_2 - \tau_1| = \{0.25\,T_c,\ 0.2\,T_c,\ 0.1\,T_c,\ 0.05\,T_c\}$ were obtained at different SNR scenarios. The paths were created with equal amplitudes and zero degrees of relative phase between them, and 10-MHz PN-sequences were used for transmission, which means that the resolution of the conventional correlation method will be limited to 100 ns. The sampling interval at the output of the correlator was set to $0.1T_c$. Note that we will try to measure a delay separation of $0.05\,T_c$, which is smaller than this sampling interval. The noise covariance matrix was calculated theoretically using Equation (2.80), and $N = 150$ snapshots were used to estimate the signal covariance matrix. Figures 2.8, 2.9, and 2.10 show the results for SNRs of 40 dB, 30 dB, and 10 dB, respectively, where SNRs are measured at the correlator output. The true locations of the path delays are marked with arrows. The solid plots correspond to the super-resolution delay profiles of the channels, and the dotted plots correspond to the channel power delay profiles $|z(\zeta)|^2$ obtained with the conventional PN-correlation method.

From these plots it can be seen that for high SNRs we can correctly estimate the true delay locations of the channels, even for path separations smaller than the sampling interval at the output of the correlator. For low SNRs we can still resolve paths separated by $0.2\ T_c$ seconds or more but, because of the large noise, now the estimates have errors and the SDPs do not peak at the exact true

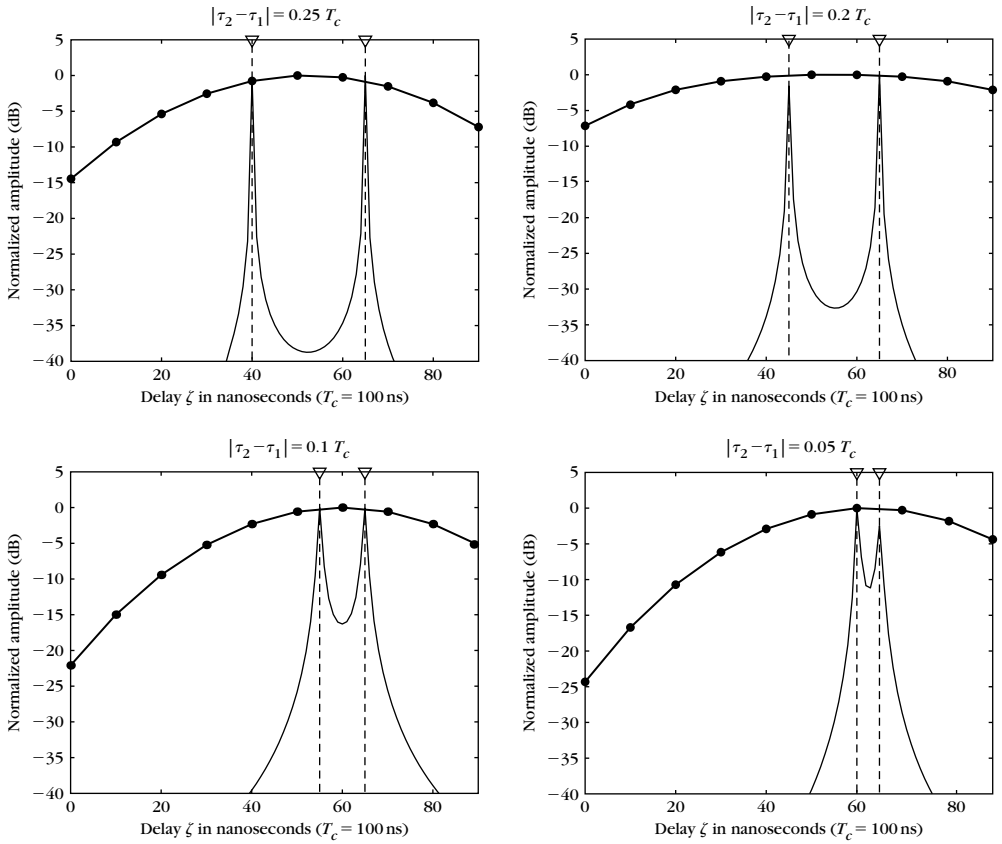

FIGURE 2.8

Super-resolution delay profiles for two-path channels with SNR = 40 dB measured at the correlator output. The dotted plots are the power delay profiles measured at the correlator output; the true path delay locations are marked by arrows.

delay locations. Additive noise lowers the resolution capabilities and causes the SDP estimates to have bias errors. Note how the dotted-line plots corresponding to the conventional correlation method are far from resolving two paths in any of the examples.

The PN-correlation–based methods presented in this section and the previous one choose the first TOA as the first peak on the estimation function (either the correlator output in the conventional case or the SDP in its super-resolution counterpart). These methods commonly suffer from two major sources of error. The first is LOS attenuation, which may be caused by shadowing. When this happens, a contiguous, subsequent strongest non-LOS arrival may be chosen as the first TOA, and this in turn will translate into range estimation errors. The

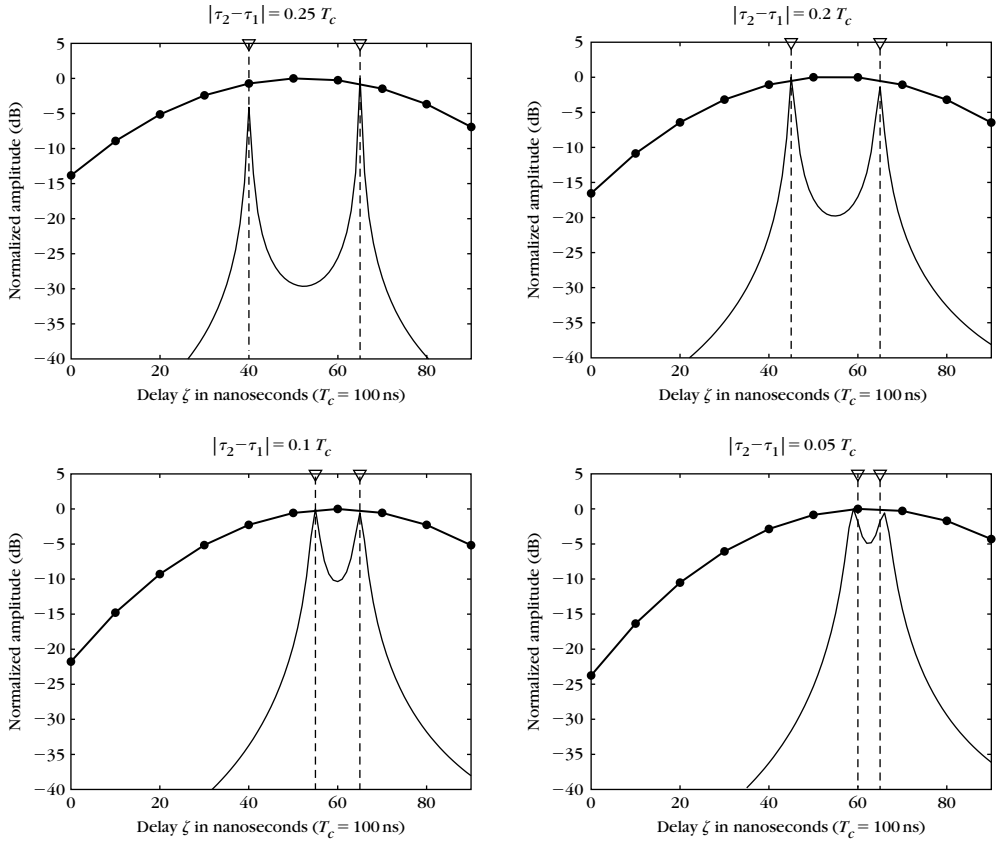

FIGURE 2.9

Super-resolution delay profiles for two-path channels with SNR = 30 dB measured at the correlator output. The dotted plots are the power delay profiles measured at the correlator output; the true path delay locations are marked by arrows.

second one is limited resolution, which may cause multipaths that arrive less than T_c seconds after the LOS arrival time to add up destructively and attenuate the magnitude of the LOS path. Unresolvable multipath components usually shift the estimator peaks to values away from the true first TOA, causing estimation errors.

TOA-based systems require high clock synchronization between all the system components in order for a receiver to know the exact time when a signal was transmitted by a source. Poor synchronization (clock drifts) is another major source of error in TOA-based systems.

TDOA systems avoid the requirement of clock synchronization at the point of interest or tag-end by considering differences in TOA of signals that originate at

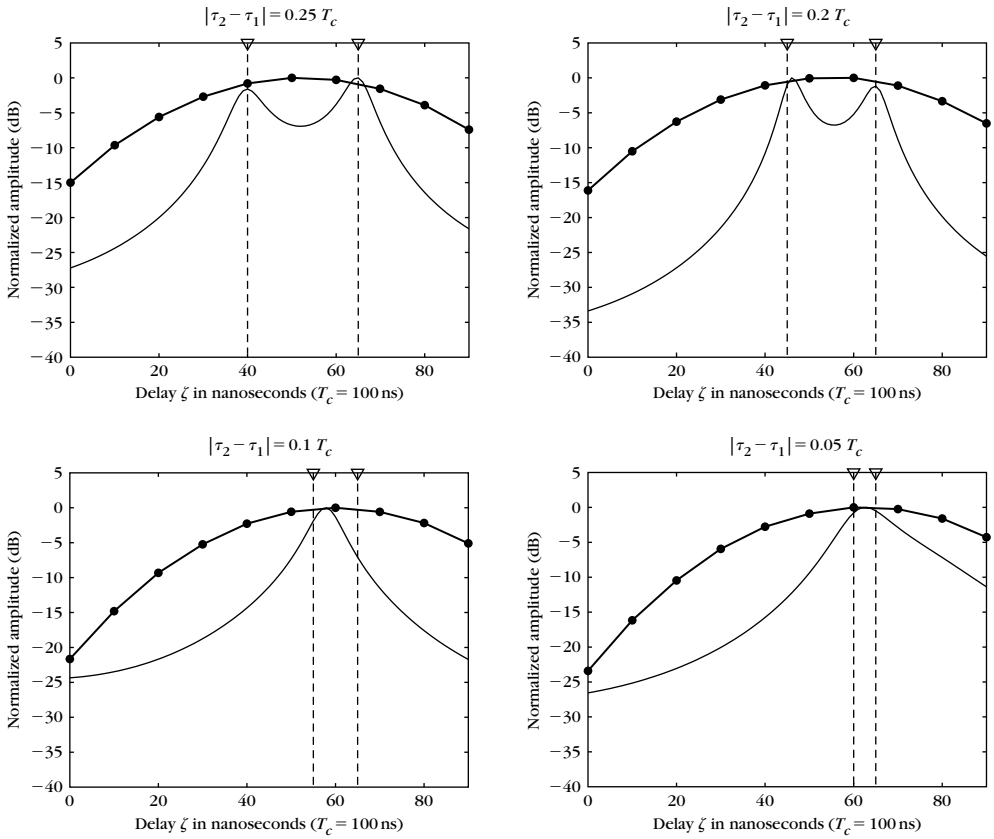

FIGURE 2.10

Super-resolution delay profiles for two-path channels with **SNR = 10 dB** measured at the correlator output. The dotted plots are the power delay profiles measured at the correlator output; the true path delay locations are marked by arrows.

two different reference points. The TDOA calculation will effectively cancel time synchronization errors at the tag. Note that TDOA-based positioning schemes still require clock synchronization among all the receivers in the system.

2.4.8 TOA Estimation by Successive Cancellation

A whitening operation was discussed in the previous section, which led to substitution of the simple eigenvalue problem to a generalized one. The whitening idea has been used effectively to improve SPM under noisy scenarios. For instance, Bouchereau et al. [3] discusses the performance of MUSIC in the presence of narrow-band jammers. In this work, it was concluded that if an estimate of the covariance matrix of the ensemble of additive and jammer noise as seen

at the output of the correlator was available, then it could be included in the generalized eigenvalue problem (as described in Equation (2.82)) to effectively eliminate the jamming effects. This could be stated as

$$\mathbf{Re_i} = \lambda_i[\mathbf{R_0} + \mathbf{R_J}]\mathbf{e}_i, \tag{2.85}$$

where $\mathbf{R_0}$ has already been defined as the normalized correlator output noise covariance matrix, and $\mathbf{R_J}$ is a covariance matrix of any other noise process present in the channel (for the jamming example, $\mathbf{R_J}$ would stand for the jamming process covariance matrix). This same idea was used in Krasny and Koorapaty [20] to obtain an improved MUSIC and conventional correlation estimator for TOA; the technique was named *successive cancellation* by the authors.

Successive cancellation is an iterative algorithm. At the i-th iteration, a channel delay profile is estimated by MUSIC (or SPM if a PN-sequence correlation is used as the received signal) or by the conventional correlation method. The position of the D largest peaks of this profile are then chosen as delay estimates $\{\hat{\tau}_1^i, \ldots, \hat{\tau}_D^i\}$, where the superscript has been used to denote the iteration number. Then at the next iteration a new noise correlation matrix is formed as

$$\mathbf{R_n}^{(i+1)} = \mathbf{R_n}^{(i)} + \hat{\mathbf{A}}^{(i)}\mathbf{G}\hat{\mathbf{A}}^{(i)H}, \tag{2.86}$$

where $\hat{\mathbf{A}}^{(i)}\mathbf{G}\hat{\mathbf{A}}^{(i)H}$ corresponds to the signal correlation term as given in Equation (2.54) which is evaluated at the delays estimated at the i-th iteration such that

$$\hat{\mathbf{A}}^{(i)} = [\mathbf{a}(\hat{\tau}_1^{(i)})\,\mathbf{a}(\hat{\tau}_2^{(i)}) \cdots \mathbf{a}(\hat{\tau}_D^{(i)})], \tag{2.87}$$

where $\mathbf{a}(\tau)$ is a steering vector as defined in Equation (2.53). As it can be seen, at each iteration, the algorithm treats delayed transmitted signals $\{s(t - \hat{\tau}_j^{(i)}),\ i = 1, \ldots, D\}$ as noise and attempts to cancel them out by embedding their correlation matrix in the overall noise matrix. This is equivalent to attempting to eliminate multipath components that may be overlapping in the delay profile function. For the case of MUSIC, this cancellation is achieved by solving, at each iteration, a generalized eigenvalue problem of the form

$$\mathbf{Re}^{(i+1)} = \lambda^{(i+1)}\mathbf{R_n}^{(i+1)}\mathbf{e}^{(i+1)}. \tag{2.88}$$

The delay profile of the channel and the corresponding D delay estimates may then be estimated with the resulting generalized eigenvectors using Equation (2.25).

For the case of the conventional correlation delay estimation method, the generalized correlation function given by

$$p^{(i+1)}(\tau) = \frac{1}{N}\sum_{i=1}^{N} |[\mathbf{R}_\mathbf{n}^{(i+1)}]^{-1}\mathbf{a}(\tau)^H\mathbf{r}_i|^2, \tag{2.89}$$

may be used to suppress delay components at $\{\hat{\tau}_1^{(i)}, \ldots, \hat{\tau}_D^{(i)}\}$ from the correlation function estimate. This last function uses N received signal snapshots $\{\mathbf{r}_i,\ i = 1, \ldots, N\}$ as described in Equation (2.52).

In the first iteration, the noise correlation matrix will be simply $\sigma^2\mathbf{I}$ (or $\sigma^2\mathbf{R_0}$ for SPM). After the last iteration, all the time delays for which a signal was detected are tabulated and the earliest delay is chosen as the first TOA of the received signal. In the tabulation process, all detected delays that lie outside a window $[\tau_p - \delta,\ \tau_p + \delta]$ are ignored. Here, τ_p is the delay at which either the correlation function or MUSIC pseudospectrum presented the maximum peak at the first iteration. The iterations may be stopped when every peak on the delay profile estimate is below a threshold value. It is clear that this algorithm will be strongly influenced by the capacity of the delay profile estimation method to resolve two consecutive paths; this capacity is measured using a probability of resolution metric. An excellent discussion on MUSIC's probability of resolution can be found in Zhang [50].

The successive cancellation algorithm may effectively improve TOA estimation performance in the presence of multipath components, as has been shown in Krasny and Koorapaty [21].

2.5 RANGE ESTIMATION BASED ON RECEIVED SIGNAL STRENGTH

RSS measurements are useful in PL systems since they may be related to distance between a transmitter and a receiver through an adequate propagation model. Although signal strength measurements are quite simple to obtain, obtaining an accurate propagation model might be the opposite. Propagation models are usually based on extensive measurement campaigns and are strongly dependent on the specific scenario (indoors, outdoors, heavy clutter, etc.), frequency band, weather conditions, and sometimes even time of day.

RSS measurements are affected by two major channel impairments known as large-scale fading and small-scale fading (see [32]), which are usually treated as random processes. Ideally, the two fading processes need to be considered when obtaining a propagation model, and this may be done by considering compound fading models [6, 13, 14]. However, due to the mathematical complexity involved in the estimation of the parameters of these schemes, it is common to ignore one of the fading processes or to reduce its impact on the signal strength measurements by means of temporal and spatial averaging as the following explains.

In this section, we will discuss ways to measure received signal strength based on spatial samples or on the channel's power delay profile. We will also

discuss averaging techniques to eliminate small-scale fading effects. Finally, we will introduce the log-normal path-loss model that is widely used in RSS-based PL schemes. Here we will present ML estimators to find the log-normal parameters based on a set of power measurements and, subsequently, an estimator of range based on these parameters.

2.6 RECEIVED SIGNAL STRENGTH MEASUREMENTS

The power of a received signal over a flat-fading wireless channel is composed of a fast-fading superimposed on a large-scale faded signal [6, 13, 14, 22]. Then the received signal power may be modeled as

$$p_r(t) = \gamma(t)p_0(t), \tag{2.90}$$

or

$$p_r(l) = m(l)p_0(l), \tag{2.91}$$

where the two equations represent the signal as measured at time t or at position l, respectively (clearly these two quantities may be related through the receiver's velocity). The objective is to find an estimate of $\gamma(t)$, which refers to the local average power, by eliminating the small-scale fading term $p_0(t)$. It is common to assume that the two fading processes are statistically independent and that $E\{p_0\} = 1$ [6, 14] (this last assumption causes no loss of generality in the model since a nonunity small-scale fading gain will be absorbed by the average of the large-scale fading component). Further, it is usually assumed that $\gamma(t)$ follows a log-normal distribution, whereas $p_0(t)$ may be described by a Rayleigh, Ricean, Suzuki, and Nakagami-m distribution, among others [6].

RSS may be estimated by obtaining the spatial average given by

$$\hat{\gamma}(l) = \frac{1}{2L}\int_{l-L}^{l+L} p_r(\alpha)d\alpha. \tag{2.92}$$

Since it is generally accepted that a log-normal fading component is constant over a spatial dimension that exceeds a few hundred wavelengths [43], it is reasonable to assume that in the PL estimation problem $m(l) = m$, Equation (2.92) becomes

$$\hat{\gamma}(l) = \frac{\gamma}{2L}\int_{l-L}^{l+L} p_0(\alpha)d\alpha. \tag{2.93}$$

Clearly, $\hat{\gamma} \to \gamma$ whenever the integral in Equation (2.93) becomes unity, which is congruent with the fact that $E\{p_0\} = 1$. The variance of the above RSS

estimator was found in [22] for the case of Rayleigh small-scale fading and is given by

$$\mathrm{VAR}(\hat{\gamma}) = \frac{\lambda}{4L}\sigma_R\sqrt{\frac{\pi}{2}}\int_0^{2L\lambda}\left(1-\frac{\lambda\alpha}{2L}\right)J_0^2(2\pi\alpha)d\alpha, \tag{2.94}$$

where $\sigma_R = \frac{1}{2}E\{p_r^2\}$, and $J_0(\cdot)$ is the zero-order Bessel function of the first kind. If L is made too short, the estimate will have large variance since it will still contain small-scale fading information. On the other hand, large values of L risk the smoothing out of large-scale fading information. An averaging over 20 to 40 wavelengths is recommended in [22].

Whenever N discrete, uncorrelated RSS measurements are available at locations $\{l_i,\ i = 1,\ \ldots,\ N\}$, the RSS estimator becomes

$$\hat{\gamma}_{dB} = \frac{1}{N}\sum_{i=1}^{N} p_{r_{dB}}(l_i), \tag{2.95}$$

where subscript *dB* denotes that the quantities are expressed in dBs. Note then that, by the central limit theorem, for large N, $\hat{\gamma}$ will be log-normal regardless of the distribution of p_r. The mean and variance, respectively, of this estimator are found in [22] as

$$\begin{aligned} E\{\hat{\gamma}\} &= \gamma, \\ \mathrm{VAR}(\hat{\gamma}) &= \bigcirc\left(\tfrac{1}{N}\right). \end{aligned} \tag{2.96}$$

According to [22], a value of $N \geqslant 36$ will achieve a $\hat{\gamma}$ that lies within ± 1 dB of its true mean. Finally, the uncorrelated sample assumption will hold whenever the distance between measurements is greater than or equal to $0.8\,\lambda$.

As a final comment, we note that elimination of small-scale fading effects is achieved through spatial averaging, which implies that measurements must be obtained at distinct locations. If this is not possible, then large-scale fading effects will be difficult to eliminate and a range estimator based on the compound fading model in Equation (2.90) may be necessary to increase accuracy. Parameter estimators for compound fading models have been presented in [6, 13, 14, 22], for example.

Now let us consider a multipath channel with D resolvable paths with impulse response $h(t)$ and a received signal $r(t)$, as given in Equations (2.42) and (2.43), respectively. In this scenario, the instantaneous received signal power at the j-th snapshot time t_j may be expressed [32] as

$$p_r(t_j) = \frac{1}{\tau_D}\int_0^{\tau_D}|r(t)|^2dt = k\sum_{i=1}^{D}|m_{ij}|^2, \tag{2.97}$$

where k is a constant that depends on the amplitude and duration (energy) of the transmitted signals. Taking the expectation of this equation and normalizing with respect to k yields the average received signal strength for the multipath channel as

$$\bar{\gamma} = \sum_{i=1}^{D} E\{|m_i|^2\} \approx \sum_{i=1}^{D} \gamma_i. \tag{2.98}$$

Note that if the power delay profile of the channel is obtained at N different uncorrelated locations, γ_i, the local average power of the i-th path may be obtained as explained previously using Equation (2.92) or (2.95). Either equation would have to be applied for each of the D resolvable paths observed on each of the N available power delay profiles.

2.6.1 Log-Normal Propagation Model

The log-normal path-loss model may be considered as a generalization of the free-space Friis equation [32] where the power is allowed to decrease at a rate of $(1/d)^n$ (where d denotes distance or range), and where a random variable is added in order to account for shadowing (large-scale fading) effects. The model may be expressed as

$$p_r(d)_{dB} = \bar{p}_r(d_0)_{dB} - 10n \log\left(\frac{d}{d_0}\right) + \chi, \quad d > d_0, \tag{2.99}$$

where log stands for base 10 logarithm, $p_r(d)_{dB}$ is the received power at a distance d meters from the transmitter, n is the path-loss exponent that defines the rate of decay of power with respect to distance, and χ is a Gaussian random variable with zero mean and variance σ_χ^2 that is defined in dBs. Further, $\bar{p}_r(d_0)_{dB}$ denotes the ensemble average over all possible received power values for a given reference distance denoted as d_0 meters. This reference power is usually measured a priori or calculated with the free space Friis equation. In all cases, d_0 should be as small as possible while being in the far-field region from the transmitter. In our discussion, we will consider that power measurements exclude small-scale fading since they were obtained by the averaging techniques presented in Section 2.6. The log-normal model basically states that the received power is not uniform when measured at different locations while maintaining the same distance separation between the transmitter and receiver. Note that a Gaussian random variable defined in decibels becomes a log-normal random variable when transformed into the linear domain. The log-normal Equation (2.99) is clearly a line with slope $10n$ when plotted versus distance values given in decibels. Typical values of path-loss exponents range between 1.5 and 5.

2.6.2 ML Estimation of Log-Normal Parameters

Parameters n and σ_χ^2 are usually unknown and must be estimated based on channel measurements. Once these parameters are obtained, distance estimation from the log-normal model becomes relatively straightforward. Let us then obtain expressions for the estimates of n and σ_χ^2.

Consider that a set of N average power observations have been obtained at different distances from a transmitter and at different uncorrelated locations in the area of interest such that N measurement pairs $\{p_r(d_i), d_i\}$ are available. For simplicity we will drop the subscript dB and assume that all power measurements are given in decibels. The log-likelihood function is obtained from the joint probability density of all N power observations as

$$L(n, \sigma_\chi^2) = -N\ln\left(\sqrt{2\pi\sigma_\chi^2}\right) - \frac{1}{2\sigma_\chi^2}\sum_{i=1}^{N}\left[p_r(d_i) - p_r(d_0) - 10n\,\log\left(\frac{d_i}{d_0}\right)\right]^2. \quad \textbf{(2.100)}$$

Differentiating with respect to n and equating to zero leads to the estimate

$$\hat{n} = \frac{\sum_{i=1}^{N}\log\left(\frac{d_i}{d_0}\right)\left[p_r(d_i) - p_r(d_0)\right]}{10\sum_{i=1}^{N}\left(\log\left(\frac{d_i}{d_0}\right)\right)^2}. \quad \textbf{(2.101)}$$

Substituting n by its estimate $\hat{n}$ in Equation (2.100), taking derivatives with respect to σ^2 and equating to zero, we obtain

$$\hat{\sigma}_\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\left[p_r(d_i) - p_r(d_0) - 10\hat{n}\,\log\left(\frac{d_i}{d_0}\right)\right]^2. \quad \textbf{(2.102)}$$

2.6.3 Log-Normal Range Estimator

When the log-normal parameters have been estimated, power measurements may be used to infer the distance between the transmitter and receiver. Substituting estimates $\hat{n}$ and $\hat{\sigma}_\chi^2$ and rewriting Equation (2.100) for a single power observation p_r, we obtain

$$L(\hat{m}, \hat{\sigma}_\chi^2, d) = -\ln\left(\sqrt{2\pi\hat{\sigma}_\chi^2}\right) - \frac{1}{2\hat{\sigma}_\chi^2}\left[p_r - p_r(d_0) - 10\hat{n}\,\log\left(\frac{d}{d_0}\right)\right]^2. \quad \textbf{(2.103)}$$

Taking a derivative with respect to d and equating to zero yields the distance estimator

$$\hat{d} = d_0 10\hat{}\left[\frac{p_r(d_0) - p_r}{10\hat{n}}\right]. \quad \textbf{(2.104)}$$

Let us now calculate the expected value and variance of this distance estimator assuming that $\hat{n}$ is a constant (i.e., not random). First, we note from Equation (2.99) that $p_r(d_0) - p_r = 10\hat{n}\log\left(\frac{d}{d_0}\right) + \chi$, and thus the estimator in (2.104) may be rewritten as

$$\hat{d} = d10^{\frac{\chi}{10\hat{n}}}. \tag{2.105}$$

Note that if $\chi \sim N(0, \sigma_\chi^2)$ and $10\log\xi \sim N(0, \sigma_\chi^2)$, then $\xi = e^{\frac{\chi}{10\log e}}$ is a log-normal distributed random variable. Equation (2.105) may then be redefined in terms of this variable as

$$\hat{d} = d\xi^{1/\hat{n}} = d(e^{\frac{\chi}{10\log e}})^{1/\hat{n}}. \tag{2.106}$$

Finally, it follows from the Gaussian moment-generating function [27] that

$$E\{\hat{d}\} = d\exp\left[\frac{\sigma_\chi^2}{200\hat{n}^2\log^2 e}\right], \tag{2.107}$$

$$\text{VAR}\{\hat{d}\} = d^2\left[\exp\left(\frac{\sigma_\chi^2}{50\hat{n}^2\log^2 e}\right) - \exp\left(\frac{\sigma_\chi^2}{100\hat{n}^2\log^2 e}\right)\right], \tag{2.108}$$

which show that the estimator is biased. An unbiased estimator is readily obtained by letting

$$\tilde{d} = \frac{\hat{d}}{c}, \tag{2.109}$$

where

$$c = \exp\left[\left(\frac{\sigma_\chi^2}{200n^2\log^2 e}\right)\right], \tag{2.110}$$

which yields a variance of

$$\text{VAR}\{\tilde{d}\} = d^2\left[\exp\left(\frac{\sigma_\chi^2}{100\hat{n}^2\log^2 e}\right) - 1\right]. \tag{2.111}$$

Note that the variance of the estimates of range increases with range d. This property is common to RSS-based range estimation schemes.

Distance estimation based on RSS measurements will suffer from non–line-of-sight scenarios (shadowing), where the received power will be reduced by attenuation due to wave propagation through an obstacle.

REFERENCES

[1] M. Azaria, D. Hertz, Time delay estimation by generalized cross correlation methods, IEEE Transaction on Acoustics, Speech, and Signal Processing, 32 (2) (1984) 280–285.

[2] M. Bekara, M. Van Der Baan, A new parametric method for time delay estimation, IEEE International Conference on Acoustic Speech and Signal Processing, 3 (2007) 1033–1036.

[3] F. Bouchereau, D. Brady, C. Lanzl, Multipath delay estimation using a superresolution PN-correlation method, IEEE Transactions on Signal Processing, 49 (5) (2001) 938–949.

[4] F. Bouchereau, D. Brady, Bounds on range-resolution degradation using RSSI measurements, in: Proceedings IEEE International Conference on Communications, June (2004) 3246–3250.

[5] F. Bouchereau, D. Brady, Resolution bounds for statistical location estimation in a hallway, in: Proceedings of the International Workshop on Wireless Communications in Underground and Confined Areas, June 2005.

[6] F. Bouchereau, D. Brady, Method-of-moments parameter estimation for compound fading processes, IEEE Transactions on Communications, 56 (2) (2008) 166–172.

[7] A.M. Bruckstein, T.-J. Shan, T. Kailath, The resolution of overlapping echos, IEEE Transactions on Acoustics, Speech, and Signal Processing, 33 (6) (1985) 1357–1367.

[8] X-L. Xu, K.M. Buckley, Bias analysis of the MUSIC location estimator, IEEE Transactions on Signal Processing, 40 (10) (1992) 2559–2569.

[9] R.J.C. Bultitude, S.A. Mahmoud, W.A. Sullivan, A comparison of indoor radio propagation characteristics at 910 MHz and 1.75 GHz, IEEE Journal on Selected Areas in Communications, 7 (1989) 20–30.

[10] G.C. Carter, A.H. Nutall, P.G. Cable, The smoothed coherence transform, IEEE Proceedings, 61 (1973) 1497–1498.

[11] Y. Chengyou, X. Shanjia, W. Dongjin, On the asymptotic analysis of Cramer-Rao bound for time delay estimation, Proceedings of ICSP (1998) 109–112.

[12] D.M.J. Devasirvatham, Time delay and signal measurements of 850 MHz radio waves in building environments, IEEE Transactions on Antennas and Propagation, AP-34 (1986) 1300–1305.

[13] A. Dogandzic, J. Jin, Maximum likelihood estimation of statistical properties of composite gamma-lognormal fading channels, IEEE Transactions on Signal Processing, 52 (10) (2004) 2940–2945.

[14] A. Dogandzic, B. Zhang, Dynamic shadow-power estimation for wireless communications, IEEE Transactions on Signal Processing, 53 (8) (2005) 2942–2948.

[15] B. Frielander, A sensitivity analysis of the MUSIC algorithm, IEEE Transactions on Acoustics, Speech, and Signal Processing, 38 (10) (1990) 1740–1751.

[16] F. Gross, Smart Antennas for Wireless Communications, McGraw-Hill, 2005.

[17] J. Hassab, R.E. Boucher, Optimum estimation of time delay by a generalized correlator, IEEE Transactions on Acoustics, Speech, and Signal Processing, 27 (4) (1979) 373–380.

[18] S. Haykin, Adaptive Filter Theory, Prentice Hall, 1995.

[19] C. Knapp, C. Carter, The generalized correlation method for estimation of time delay, IEEE Transactions on Acoustics, Speech, and Signal Processing, 24 (4) (1976) 320–327.

[20] L. Krasny, H. Koorapaty, Enhanced time of arrival estimation with successive cancellation, IEEE Vehicular Technology Conference, 6 (2002) 851–855.

[21] L. Krasny, H. Koorapaty, Performance of successive cancellation techniques for time of arrival estimation, IEEE Vehicular Technology Conference, 4 (2002) 2278–2282.

[22] W.C.Y. Lee, Estimate of local average power of a mobile radio signal, IEEE Transactions on Vehicular Technology, 34 (1) (1985).

[23] X. Li, K. Pahlavan, J. Beneat, Performance of TOA estimation techniques in indoor multipath channels, IEEE International Symposium on Personal Indoor and Mobile Radio Communications, 2 (2002) 911–915.

[24] T. Manabe, H. Takai, Superresolution of multipath delay profiles measured by PN correlation method, IEEE Transactions on Antennas and Propagation, 40 (5) (1992) 500–509.

[25] B. Ottersten, M. Viberg, T. Kailath, Performance analysis of the total least squares ESPRIT algorithm, IEEE Transactions on Signal Processing, 39 (5) (1991) 1122–1134.

[26] M-A. Pallas, G. Jourdain, Active high resolution time delay estimation for large BT signals, IEEE Transactions on Signal Processing, 39 (4) (1991) 781–787.

[27] A. Papoulis, S.U. Pillai, Probability, Random Variables and Stochastic Processes, McGraw-Hill, 2001.

[28] R.L. Peterson, R.E. Ziemer, D.E. Borth, Introduction to Spread Spectrum Communications, Prentice Hall, 1995.

[29] J.G. Proakis, D.G. Manolakis, Digital Signal Processing, Principles, Algorithms, and Applications, Prentice Hall, 2006.

[30] A. Quazi, An overview on time delay estimation in active and passive systems for target localization, IEEE Transactions on Acoustics, Speech, and Signal Processing, 29 (3) (1981) 527–533.

[31] T.S. Rappaport, Characterization of UHF multipath radio channels in factory buildings, IEEE Transactions on Antennas and Propagation, AP-37 (1989) 1058–1069.

[32] T.S. Rappaport, Wireless Communications, Principles and Practice, Prentice Hall, 2002.

[33] V.U. Reddy, A. Paulraj, T. Kailath, Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing, IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-35 (7) (1987) 927–935.

[34] K. Maheswara Reddy, V.U. Reddy, Further results in spatial smoothing, Signal Processing, 18 (1965) 217–224.

[35] P.R. Roth, Effective measurements using digital signal analysis, IEEE Spectrum, 8 (1971) 62–70.

[36] R. Roy, T. Kailath, ESPRIT- estimation of signal parameters via rotational invariance techniques, IEEE Transactions on Acoustics, Speech, and Signal Processing, 37 (7) (1989) 984–994.

[37] T-J. Shan, M. Wax, T. Kailath, On spatial smoothing for direction of arrival estimation of coherent signals, IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-33 (4) (1985) 806–811.

[38] L.L Scharf, Statistical Signal Processing—Detection, Estimation, and Time Series Analysis, 2nd edition, Addison-Wesley, 1991.

[39] R.O. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Transactions on Antennas and Propagation, AP-34 (1986) 276–280.

[40] P. Stoica, A. Nehorai, MUSIC, maximum likelihood, and Cramer-Rao bound, IEEE Transactions on Acoustics, Speech, and Signal Processing, 37 (5) (1989) 720–741.

[41] P. Stoica, K.C. Sharman, Maximum likelihood methods for direction-of-arrival estimation, IEEE Transactions on Acoustics, Speech, and Signal Processing, 38 (7) (1990) 1132–1142.

[42] P. Stoica, R. Moses, Spectral Analysis of Signals, Pearson-Prentice Hall, 2005.

[43] H. Suzuki, A statistical model for urban radio propagation, IEEE Transactions on Communications, 25 (7) (1977) 673–680.

[44] A. Swindlehurst, T. Kailath, A performance analysis of subspace-based methods in the presence of model errors: part I—The MUSIC algorithm, IEEE Transactions on Signal Processing, 40 (7) (1992) 1758–1774.

[45] C.W. Therrien, Discrete Random Signals and Statistical Signal Processing, Prentice Hall, 1992.

[46] H.L. Van Trees, Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory, John Wiley and Sons, 2002.

[47] M. Wax, T. Kailath, Detection of signals by information theoretic criteria, IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-33 (2) (1985) 387–392.

[48] M. Wax, I. Ziskind, Detection of the number of coherent signals by the MDL principle, IEEE Transactions on Acoustics, Speech, and Signal Processing, 37 (8) (1989) 1190–1196.

[49] W. Xu, M. Kaveh, Comparative study of the biases of MUSIC-like estimators, Signal Processing, 50 (1996) 39–55.

[50] Q.T. Zhang, Probability of resolution of the MUSIC algorithm, IEEE Transactions on Signal Processing, 43 (4) (1995) 978–987.