Chapter 1

EXPLOITING ANTENNA ARRAYS FOR SYNCHRONIZATION

Gonzalo Seco[1]
Department of Signal Theory & Communications
Universitat Politècnica de Catalunya
Barcelona, Spain

A. Lee Swindlehurst[2]
Department of Electrical & Computer Engineering
Brigham Young University
Provo, Utah

David Astély
Research & Development Department
Nokia Telecommunications
Kista, Sweden

1.1 Introduction

Synchronization is a critical aspect of virtually all communications systems. Accurate frame and symbol synchronization is especially important in time-division multiple access (TDMA) and packet-based systems, or with any protocol that employs training, control, or identification bits interspersed with the raw data. Multiuser detectors for code-division multiple access (CDMA) require reliable code timing information for acceptable performance in near-far environments. In addition, achieving precise synchronization is the key to obtaining location estimates with accuracies of a few meters or better in Global Positioning System (GPS) receivers.

Timing information is also needed for any application where range measure-

[1]Work on this project supported by the Catalan and Spanish Governments under grants 1997FI 00755 APDT, TIC96-0500-C10-01, TIC98-0412, TIC98-0703, TIC99-0849, and CIRIT 1998SGR-00081.

[2]Work on this project supported by the U. S. National Science Foundation under Wireless Initiative Grant CCR 99-79452.

ments are made, as in active radar and sonar systems. In these fields the problem is usually referred to as time-delay estimation, but in mathematical terms there is little difference between this and synchronization. We will use both terms interchangeably. There is a vast literature on both synchronization and time delay estimation, the majority of which focuses on the case where data is measured from a single receiver. However, the performance of single channel timing recovery methods is limited when multipath or co-channel interference (CCI) is present, such as in surveillance systems plagued by jamming and many wireless communications applications. For this reason, attention has recently shifted to the use of antenna arrays for addressing these problems. The spatial selectivity offered by an antenna array can dramatically improve performance in environments with severe interference.

A number of techniques that exploit antenna arrays for synchronization have been developed, each differing from the others on its assumptions regarding multipath, CCI, signal parameterization, and computational load. One of the first such techniques was presented in [1] for DS-CDMA systems, in which a least-squares beamformer is calculated for each possible location of the desired user's codeword over one symbol period, assuming no transition has occurred between consecutive symbols. A similar approach was presented in [2] for TDMA communication systems, using a minimum mean-squared error (MMSE) beamformer calculated for each possible position of a training sequence in a given frame of data. The beamformer that results in an output that is the most strongly correlated with the training sequence is used to cancel CCI and any uncorrelated multipath. The guard intervals present in such systems are also included in order to remove CCI not present during the training interval. More recently, this approach has been revised and extended to better deal with CCI [3], and to handle very short bursts of training data [4].

Other researchers have taken a parameter estimation point of view, attempting to determine the direction of arrival (DOA) and time delay of each arrival of a given signal at the array. While these techniques do not take CCI into account, they exploit the full space-time structure of the multipath. The methods of [5], [6] do not operate directly on the data; instead, they assume that the channel matrix has been estimated in a previous step. The DOAs and time delays are then determined by fitting the model to the estimated channel. While suboptimal, the advantage of this approach is that in certain cases, the channel estimate may be obtained blindly, without the need for training data. The algorithm of [6] requires either a multidimensional (MD) search, or a series of suboptimal one-dimensional (1-D) searches, while that of [5] assumes a uniform linear array (ULA) and achieves closed-form estimates using a 2-D version of the ESPRIT technique [7]. This latter idea is carried one step further in [8], where a uniform rectangular array is used to estimate both azimuth and elevation DOAs along with the time delays using a 3-D ESPRIT implementation. It should be noted that [8] works directly on the received data to estimate the parameters in a single step. A maximum likelihood (ML) approach is taken in [9], [10], [11], in which both the interference and noise are modeled together as a temporally white Gaussian process. The methods of [10],

[11] also assume spatial whiteness. In [12], the joint angle and delay estimation problem is solved via weighted least squares (WLS), where the weighting matrices are designed to account for spatial color, array calibration errors, etc. While offering some claim to optimality, the primary drawback of the ML and WLS approaches is that complicated search procedures are required to estimate the desired parameters.

To obtain DOA estimates of each arrival, the parametric approaches described above must assume the availability of a calibrated antenna array, and a single arrival at each time delay. Errors in the array calibration or deviations of the array from uniformity are inevitable, and can lead to significant performance degradation. Furthermore, in multipath-rich propagation environments, there may be numerous arrivals at each delay due to local scattering near the array. To overcome these difficulties, an unstructured parameterization of the spatial response can be used, as in [13]. While this leads to an increase in the number of parameters to be estimated, the model is linear in the additional parameters, and they can be estimated in closed form. Once the spatial parameters are eliminated from the ML criterion in [13], the time delays are solved for iteratively using either IQML [14] or MODE [15], [16]. A suboptimal delay estimator based on ESPRIT is also presented in [13]. These techniques have recently been extended in [17] to the blind case where no training data is available, although without exploiting additional knowledge about the signal (*e.g.,* known pulse shaping, etc), an absolute time base cannot be established in this case. Note that [13], [17] both assume spatially and temporally white noise, and thus are not suited for situations involving strong CCI. If the desired signal is digitally modulated with a known pulse shape, the iterative method of [18] can be used to account for spatially colored interference via prewhitening. The algorithm presented in [19] also allows for interference with arbitrary unknown spatial color as well as an unstructured array response for the desired signal, although it assumes a slightly different temporal model. Instead of modeling the multipath arrivals using arbitrary delays, the arrivals are assumed to occur on a uniformly spaced time-domain grid with an unknown starting location. While not exact, this model leads to an ML solution requiring only a 1-D search for the starting position of the training sequence.

Other work has focused on the special nature of the synchronization problem in various applications. The use of antenna arrays in code timing recovery for CDMA applications has been addressed in [1] as mentioned above, and more recently in [20], [21], [22], [23], [24]. Each of these approaches estimates the code timing for one user at a time while treating the multiple access CCI as Gaussian interference with unknown spatial color. In [20], [21] the interference is restricted to be temporally white, while the other methods allow for CCI with unknown temporal color. The methods described in [20], [22], [24] assume flat fading and an unstructured spatial response model for the desired user, and use a maximum likelihood approach that leads to a simple 1-D search for the location of the user's codeword. They differ in that [24] assumes a known set of training symbols, while the algorithm in [22] is blind. The algorithm in [20] is general enough for either case, but it operates using only one symbol's worth of data, and the time delay is restricted to be an integer

multiple of the chip period. The approach described in [21] also assumes flat fading and an unstructured array response, but it resorts to an unnecessary asymptotic approximation of the ML criterion to achieve a 1-D parameter search. However, [21] does present a method for transforming the 1-D search for the time delay into rooting a second order polynomial. As an alternative to the above approaches, [23] assumes a uniform linear array and processes the data in the frequency domain to estimate the DOAs and time delays of all of the desired user's multipath arrivals via 2-D ESPRIT.

Exploiting antenna arrays for synchronization in GPS systems is another application that has recently been considered in [25], [26]. Differentiating aspects of this application are that only the time delay of the direct path is of interest, and one can usually assume that the DOA of the direct path is precisely known. Finally, we also mention the work of [27], [28], which focus on exploiting antenna arrays not only for estimation of multipath time delays, but also carrier offsets and Doppler shifts as well.

The goal of this chapter is to present a general mathematical framework for the problem of using antenna arrays for synchronization and time delay estimation, in order to better contrast and compare the techniques described above. We will also present a few specific algorithms that we believe offer the best compromise between model realism and computational complexity.

1.2 Data Model

We assume that an arbitrary m element array receives d scaled and delayed replicas of a known signal $s(t)$. The baseband array output is modeled as the $m \times 1$ complex vector

$$\mathbf{y}[n] = \sum_{k=1}^{d} \mathbf{a}_k s(nT_s - \tau_k) + \mathbf{e}[n] \,, \tag{1.2.1}$$

where T_s is the sampling period, $\mathbf{a}_k \in \mathcal{C}^m$ and τ_k are the spatial signature and time delay of the k^{th} arrival, and $\mathbf{e}[n] \in \mathcal{C}^m$ represents additive noise and interference. We may write (1.2.1) in matrix form as follows:

$$\mathbf{y}\,[n] = \mathbf{A}\,\mathbf{s}\,[n, \boldsymbol{\tau}] + \mathbf{e}\,[n] \tag{1.2.2}$$

where

$$\boldsymbol{\tau} = [\tau_1 \;\; \cdots \;\; \tau_d]^T \qquad d \times 1 \tag{1.2.3}$$

$$\mathbf{A} = [\mathbf{a}_1 \;\; \cdots \;\; \mathbf{a}_d] \qquad m \times d \tag{1.2.4}$$

$$\mathbf{s}[n, \boldsymbol{\tau}] = \left[\; s\,(nT_s - \tau_1) \quad s\,(nT_s - \tau_2) \quad \cdots \quad s\,(nT_s - \tau_d) \;\right]^T \qquad d \times 1 \tag{1.2.5}$$

If N samples are collected from the array, they all may be grouped together into the following $m \times N$ matrix equation:

$$\mathbf{Y} = [\mathbf{y}[1] \;\; \mathbf{y}[2] \;\; \cdots \;\; \mathbf{y}[N]] = \mathbf{A}\,\mathbf{S}\,(\boldsymbol{\tau}) + \mathbf{E} \tag{1.2.6}$$

where $\mathbf{S}(\boldsymbol{\tau}) \in \mathcal{C}^{d \times N}$ and $\mathbf{E} \in \mathcal{C}^{m \times N}$ are formed identically to $\mathbf{Y}$.

It is clear from the above model that we are invoking the standard "narrowband assumption" common to many array signal processing problems; *i.e.,* we assume that the time required for the signal to propagate across the array is much smaller than its inverse bandwidth. Note that, rather than parameterizing the array response $\mathbf{a}_k$ in terms of one or more DOAs, we treat it as an unstructured deterministic vector. Note also that we assume that the received signal can be described by discrete arrivals with distinct delays. This is obviously an approximation, especially in the multipath rich environments often encountered in wireless communications. In situations where temporally and spatially diffuse arrivals are present, the above model is still quite reasonable since arrivals that have nearly the same delay τ (*i.e.,* arrivals whose delays are separated by less than the reciprocal of the signal bandwidth, due for example to a cluster of closely spaced scatterers) can be grouped together into a single term. The temporal spread in a given cluster and the total number of clusters d will then depend on the bandwidth of the transmitted signal; the wider the signal bandwidth, the more clusters that may be necessary. The singular values of the data matrix formed from several snapshots of data can be used to determine d for a given scenario.

This approach is advantageous from a modeling and estimation point of view, as it leads to algorithms with a more reasonable computational cost. In a fully parameterized model, the spatial signature would be decomposed as

$$\mathbf{a}_k = \sum_{i=1}^{d_k} \alpha_{i,k} \mathbf{a}(\theta_{i,k}) \,, \tag{1.2.7}$$

where $\mathbf{a}(\theta)$ represents the far-field array response to a unit amplitude plane wave arriving from DOA θ, and $d_k, \alpha_{i,k}$ and $\theta_{i,k}$ denote the total number of multipaths associated with arrival k, their complex amplitudes, and their DOAs, respectively. While perhaps more concise than assuming an unstructured $\mathbf{a}_k$ (unless d_k is large), such a model requires in general a more complicated estimator due its non-linear dependence on the DOA parameters. In addition, estimation of the DOAs necessitates that the array response $\mathbf{a}(\theta)$ be accurately calibrated, which is a problematic assumption. For these reasons, we feel that the use of an unstructured spatial response model leads to a much more practical approach. If DOA information is needed (*e.g.,* in forming transmit beamformer weights for downlink communication in frequency division duplex communications systems), the directions can be determined from the estimated spatial signatures using a simple least-squares fit, provided that d_k is not too large. See [13], [29] for more information on this approach.

Errors in the above model, together with the effects of background noise and co-channel interference, are all lumped together in the error term $\mathbf{e}[n]$. The CCI contribution to $\mathbf{e}[n]$ could be modeled in the same way as the signal of interest, *e.g.,* as several delayed versions of, say, a finite alphabet sequence. However, taking the CCI structure into account in this way will lead to a search over all finite alphabet sequences transmitted by the interferers. Instead of such a computationally

demanding strategy, we model the CCI contribution, along with additive noise and model errors, as complex Gaussian. This assumption is primarily for modeling purposes, and allows us to develop a metric that takes the spatial covariance of the CCI into account. Thus, $\mathbf{e}[n]$ is modeled as a complex, circularly-symmetric, zero-mean Gaussian process. For simplicity the process is assumed to be temporally white. However, as in, *e.g.,* [9], [19], [20], [21], [30], the CCI is accounted for by modeling the process as *spatially* colored with an arbitrary unknown correlation matrix:

$$\mathcal{E}\left\{\mathbf{e}[n]\mathbf{e}^*[m]\right\} = \mathbf{Q}\,\delta_{n,m}\ , \tag{1.2.8}$$

where $(\cdot)^*$ denotes the complex conjugate transpose operation. While such a model for $\mathbf{e}[n]$ is clearly only approximate, it captures the most significant effects of the noise and interference, and leads to tractable algorithms.

With the above mathematical model in hand, we can succinctly state the problem addressed in this work:

> **(P1)** – Given N snapshots of data in the matrix $\mathbf{Y}$ described by equations (1.2.1)-(1.2.6), estimate the spatial signatures $\mathbf{A}$ and time delays $\boldsymbol{\tau}$ of the arrivals, as well as the spatial covariance $\mathbf{Q}$ of the noise and interference.

We will also consider a slight variation of the above problem that results when a small but significant modification to the signal model is used. As an alternative to (1.2.5), consider the following definition for $\mathbf{s}[n,\boldsymbol{\tau}]$:

$$\mathbf{s}[n,\tau] = \left[\ s\left(nT_s-\tau\right)\quad s\left((n-1)T_s-\tau\right)\quad\cdots\quad s\left((n-d+1)T_s-\tau\right)\ \right]^T\ . \tag{1.2.9}$$

In this approach, the effect of the multipath channel is modeled as a finite impulse response (FIR) filter, with the k^{th} row of $\mathbf{A}$ containing the filter coefficients for the channel separating the source and the k^{th} antenna. The signal vector is assumed to be known to within the scalar time delay parameter τ, which we will refer to as the *frame* delay to differentiate it from the *propagation* delays in the model of (1.2.5). For the model of (1.2.9), we word the problem statement as follows:

> **(P2)** – Given N snapshots of data in the matrix $\mathbf{Y}$ described by equations (1.2.1)-(1.2.4), (1.2.6), and (1.2.9), estimate the FIR channel matrix $\mathbf{A}$ and frame delay τ of the signal, as well as the spatial covariance $\mathbf{Q}$ of the noise and interference.

The mathematical approach we take to solve (P1) and (P2) is identical, although the complexity of the resulting algorithms will be somewhat different.

Although in the FIR case the effects of temporal oversampling or a pulse shaping filter (for digitally modulated signals) could be factored into the channel matrix $\mathbf{A}$, we will assume that the elements of $\mathbf{s}[n,\tau]$ are samples of the continuous modulated waveform $s(t)$ rather than discrete symbols (this applies to the model in (1.2.5) as well). As such, the matrix $\mathbf{A}$ only describes the propagation effects of the channel,

and τ is a continuous-valued variable. This assumption is somewhat different than that made in [19] and in other work on blind equalization of FIR channels (*e.g.,* [31], [32], [33], [34], [35]). For both (1.2.5) and (1.2.9), we need an additional technical assumption that will be used later on, namely that $s(t)$ is a band-limited finite-average-power signal. Therefore, its analog autocorrelation function

$$c_{ss}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_T s(t+\tau)\, s^*(t)\, dt \tag{1.2.10}$$

is assumed to be continuous with continuous derivatives. Assuming also that the sampling period T_s satisfies the Nyquist criterion, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_N s(nT_s - \tau_l)\, s^*(nT_s - \tau_k) = c_{ss}(\tau_k - \tau_l) \tag{1.2.11}$$

As a final modeling issue, note that in the above discussion we have implicitly assumed that d, the number of multipath rays (clusters) or the length of the FIR channels, is known. Determining d is a non-trivial problem that is beyond the scope of this chapter. A number of possibilities exist, including simple rank tests on $\mathbf{Y}$, use of the Minimum Description Length (MDL) [36] or Akaike's criterion (AIC) [37], sequential tests based on the asymptotic distribution of a given criterion function [38], [39], [40], robust bootstrap techniques [41], among many others.

1.3 Maximum Likelihood Estimator

Let $p(\cdot|\mathbf{Q})$ denote the probability density function (pdf) of a complex Gaussian vector with zero mean and covariance $\mathbf{Q}$. Under the assumption that $\mathbf{e}[n]$ is temporally white, the negative log-likelihood function for N observations of $\mathbf{y}[n]$ is given by

$$f_N(\boldsymbol{\tau}, \mathbf{A}, \mathbf{Q}) = -\sum_{n=1}^{N} \log p\left(\mathbf{y}[n] - \mathbf{A}\mathbf{s}[n, \boldsymbol{\tau}]\,|\mathbf{Q}\right) . \tag{1.3.1}$$

The subscript N is used to explicitly denote the number of data samples used to form the criterion. Note that this equation and all others derived in this section apply to both problems (P1) and (P2) described above; the only difference is that for (P2), $\boldsymbol{\tau}$ is a scalar variable representing the frame delay. Making use of the expression for the complex Gaussian pdf [42] and neglecting irrelevant additive and multiplicative constants, we obtain

$$f_N(\boldsymbol{\tau}, \mathbf{A}, \mathbf{Q}) = \log|\mathbf{Q}| + \mathrm{trace}\left\{\mathbf{C}(\boldsymbol{\tau}, \mathbf{A})\mathbf{Q}^{-1}\right\} , \tag{1.3.2}$$

where

$$\mathbf{C}(\boldsymbol{\tau}, \mathbf{A}) = \hat{\mathbf{R}}_{yy} - \mathbf{A}\hat{\mathbf{R}}^*_{ys}(\boldsymbol{\tau}) - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\mathbf{A}^* + \mathbf{A}\hat{\mathbf{R}}_{ss}(\boldsymbol{\tau})\mathbf{A}^* \tag{1.3.3}$$

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^* \tag{1.3.4}$$

$$\hat{\mathbf{R}}_{ys}(\boldsymbol{\tau}) = \frac{1}{N}\mathbf{Y}\mathbf{S}^*(\boldsymbol{\tau}) \tag{1.3.5}$$

$$\hat{\mathbf{R}}_{ss}(\boldsymbol{\tau}) = \frac{1}{N}\mathbf{S}(\boldsymbol{\tau})\mathbf{S}^*(\boldsymbol{\tau}) \ . \tag{1.3.6}$$

The minimization of (1.3.2) with respect to $\mathbf{Q}$ and $\mathbf{A}$ may be performed explicitly. Using standard matrix calculus results (see *e.g.*, [43]), the gradient of the criterion with respect to $\mathbf{Q}$ is easily shown to be

$$\frac{\partial f_N(\boldsymbol{\tau}, \mathbf{A}, \mathbf{Q})}{\partial \mathbf{Q}} = \mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{C}(\boldsymbol{\tau}, \mathbf{A})\mathbf{Q}^{-1} \ ,$$

from which it is clear that the ML estimate of $\mathbf{Q}$ is given by[3]

$$\hat{\mathbf{Q}}_{ML}(\boldsymbol{\tau}, \mathbf{A}) = \mathbf{C}(\boldsymbol{\tau}, \mathbf{A}) \ . \tag{1.3.7}$$

Replacing $\mathbf{Q}$ in (1.3.2) with (1.3.7) and neglecting the resulting constant term yields

$$f_N(\boldsymbol{\tau}, \mathbf{A}) = \log \left| \hat{\mathbf{R}}_{yy} - \mathbf{A}\hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau}) - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\mathbf{A}^* + \mathbf{A}\hat{\mathbf{R}}_{ss}(\boldsymbol{\tau})\mathbf{A}^* \right| \tag{1.3.8}$$

$$= \log \left| \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau}) \right.$$
$$\left. + \left(\mathbf{A} - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\right)\hat{\mathbf{R}}_{ss}\left(\mathbf{A} - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\right)^* \right| \tag{1.3.9}$$

$$\geq \log \left| \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau}) \right| \ , \tag{1.3.10}$$

where in the second equation we have added and subtracted the term

$$\hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau}) \ .$$

For every $\boldsymbol{\tau}$, the lower bound of (1.3.10) is clearly achieved if $\mathbf{A} = \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})$, so the ML estimates of $\mathbf{A}$ and $\mathbf{Q}$ may be expressed as

$$\hat{\mathbf{A}}_{ML}(\boldsymbol{\tau}) = \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau}) \tag{1.3.11}$$

$$\hat{\mathbf{Q}}_{ML}(\boldsymbol{\tau}) = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau}) \ . \tag{1.3.12}$$

The resulting criterion for $\boldsymbol{\tau}$ is then

$$f_N(\boldsymbol{\tau}) = \log \left| \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau}) \right| = \log \left| \hat{\mathbf{Q}}_{ML}(\boldsymbol{\tau}) \right| \ . \tag{1.3.13}$$

Thus, the maximum likelihood estimate of the spatial signatures (FIR channels), $\hat{\mathbf{A}}_{ML}(\boldsymbol{\tau})$, is given by a least squares fit to the data, and the ML estimate of the spatial noise covariance, $\hat{\mathbf{Q}}_{ML}(\boldsymbol{\tau})$, is simply the sample covariance of the residuals. The delay(s) for which the determinant of the sample covariance of the residuals is

[3] We assume that $N \geq m + d$ so that the matrix $\mathbf{C}(\boldsymbol{\tau}, \mathbf{A})$ is invertible with probability one.

minimized correspond to the propagation delays in the case of (P1), and the frame delay for (P2).

Using the following standard properties of the matrix determinant:

$$|\mathbf{XZ}| = |\mathbf{X}| \cdot |\mathbf{Z}|$$
$$|\mathbf{I} - \mathbf{XZ}| = |\mathbf{I} - \mathbf{ZX}| \ ,$$

where $\mathbf{X}$ and $\mathbf{Z}$ are appropriately dimensioned matrices, it is straightforward to show that

$$f_N(\boldsymbol{\tau}) = \log\left|\hat{\mathbf{R}}_{yy}\right| + \log\left|\mathbf{I} - \hat{\mathbf{R}}_{yy}^{-1}\hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^{*}(\boldsymbol{\tau})\right| \tag{1.3.14}$$
$$= \log\left|\hat{\mathbf{R}}_{yy}\right| + \log\left|\mathbf{I} - \hat{\mathbf{R}}_{yy}^{-\frac{1}{2}}\hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^{*}(\boldsymbol{\tau})\hat{\mathbf{R}}_{yy}^{-\frac{1}{2}}\right| \tag{1.3.15}$$
$$\stackrel{\text{def}}{=} \log\left|\hat{\mathbf{R}}_{yy}\right| + \log|\mathbf{I} - \mathbf{B}_N(\boldsymbol{\tau})| \ , \tag{1.3.16}$$

where we have defined $\mathbf{B}_N(\boldsymbol{\tau})$ in an obvious way. Note that the first term involving the determinant of $\hat{\mathbf{R}}_{yy}$ can be ignored when minimizing with respect to $\boldsymbol{\tau}$ since it is parameter independent. We will define $V_N(\boldsymbol{\tau})$ as the criterion obtained by ignoring the first term:

$$V_N(\boldsymbol{\tau}) = \log|\mathbf{I} - \mathbf{B}_N(\boldsymbol{\tau})| \ .$$

Consequently, the proposed maximum likelihood synchronization approach can be summarized as in Table 1.1. Some properties of the algorithm are discussed below.

$$\hat{\boldsymbol{\tau}}_{ML} = \arg\min_{\boldsymbol{\tau}} V_N(\boldsymbol{\tau}) \tag{1.3.17}$$
$$= \arg\min_{\boldsymbol{\tau}} \log|\mathbf{I} - \mathbf{B}_N(\boldsymbol{\tau})| \tag{1.3.18}$$
$$\mathbf{B}_N(\boldsymbol{\tau}) = \hat{\mathbf{R}}_{yy}^{-\frac{1}{2}}\hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\hat{\mathbf{R}}_{ys}^{*}(\boldsymbol{\tau})\hat{\mathbf{R}}_{yy}^{-\frac{1}{2}} \tag{1.3.19}$$
$$\hat{\mathbf{A}}_{ML}(\boldsymbol{\tau}) = \hat{\mathbf{R}}_{ys}(\hat{\boldsymbol{\tau}}_{ML})\hat{\mathbf{R}}_{ss}^{-1}(\hat{\boldsymbol{\tau}}_{ML}) \tag{1.3.20}$$
$$\hat{\mathbf{Q}}_{ML}(\boldsymbol{\tau}) = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}(\hat{\boldsymbol{\tau}}_{ML})\hat{\mathbf{R}}_{ss}^{-1}(\hat{\boldsymbol{\tau}}_{ML})\hat{\mathbf{R}}_{ys}^{*}(\hat{\boldsymbol{\tau}}_{ML}) \tag{1.3.21}$$

Table 1.1. Summary of the Maximum Likelihood Synchronization Algorithm for Spatially Colored Noise and Interference

1.3.1 Consistency

The consistency of the ML time delay estimator follows from the fact that as $N \to \infty$, $V_N(\boldsymbol{\tau})$ converges with probability one to its limiting value $V_\infty(\boldsymbol{\tau})$, which is minimized by the true values of the time delays, denoted by the vector $\boldsymbol{\tau}_0$. The

convergence is uniform thanks to the differentiability of $c_{ss}(\tau)$. By (1.2.11), the limiting value of the cost function is

$$\begin{aligned} V_\infty(\boldsymbol{\tau}) &= \log\left|\mathbf{I}-\mathbf{R}_{yy}^{-\frac{1}{2}}\mathbf{A}\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)\mathbf{C}_{ss}^{-1}\left(\boldsymbol{\tau},\boldsymbol{\tau}\right)\mathbf{C}_{ss}^{*}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)\mathbf{A}^{*}\mathbf{R}_{yy}^{-\frac{1}{2}}\right| \\ &= \log\left|\mathbf{I}-\mathbf{R}_{yy}^{-\frac{1}{2}}\mathbf{A}\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}_0\right)\mathbf{A}^{*}\mathbf{R}_{yy}^{-\frac{1}{2}}\right. \\ &\left.+\,\mathbf{R}_{yy}^{-\frac{1}{2}}\mathbf{A}\left(\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}_0\right)-\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)\mathbf{C}_{ss}^{-1}\left(\boldsymbol{\tau},\boldsymbol{\tau}\right)\mathbf{C}_{ss}^{*}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)\right)\mathbf{A}^{*}\mathbf{R}_{yy}^{-\frac{1}{2}}\right| \end{aligned} \tag{1.3.22}$$

where $\mathbf{R}_{yy}$ is the limiting value of $\hat{\mathbf{R}}_{yy}$ and the k,l-th element of the matrix $\mathbf{C}_{ss}(\boldsymbol{\tau},\boldsymbol{\lambda})$ is $c_{ss}(\lambda_l-\tau_k)$. At this point we have to use the following results. First, the determinant is a nondecreasing function. This means that for any positive definite matrix $\mathbf{G}$ and any non-negative definite matrix $\boldsymbol{\Delta}\mathbf{G}$, the determinant satisfies

$$\begin{aligned} |\mathbf{G}+\boldsymbol{\Delta}\mathbf{G}| &= \left|\mathbf{G}(\mathbf{I}+\mathbf{G}^{-1}\boldsymbol{\Delta}\mathbf{G})\right| \\ &= |\mathbf{G}|\left|\mathbf{I}+\mathbf{G}^{-1}\boldsymbol{\Delta}\mathbf{G}\right| \\ &\geq |\mathbf{G}|\ , \end{aligned}$$

since the eigenvalues of $\mathbf{I}+\mathbf{G}^{-1}\boldsymbol{\Delta}\mathbf{G}$ are ≥ 1. Note that the equality only holds for $\boldsymbol{\Delta}\mathbf{G}=\mathbf{0}$. Second, the matrix

$$\mathbf{M}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)=\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}_0\right)-\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)\mathbf{C}_{ss}^{-1}\left(\boldsymbol{\tau},\boldsymbol{\tau}\right)\mathbf{C}_{ss}^{*}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right)$$

is non-negative definite because it is the Schur complement of $\mathbf{C}_{ss}(\boldsymbol{\tau},\boldsymbol{\tau})$ in

$$\begin{bmatrix} \mathbf{C}_{ss}\left(\boldsymbol{\tau},\boldsymbol{\tau}\right) & \mathbf{C}_{ss}^{*}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right) \\ \mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}\right) & \mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}_0\right) \end{bmatrix} = \lim_{N\to\infty}\frac{1}{N}\sum_{N}\begin{bmatrix} \mathbf{s}\left[n,\boldsymbol{\tau}\right] \\ \mathbf{s}\left[n,\boldsymbol{\tau}_0\right] \end{bmatrix}\begin{bmatrix} \mathbf{s}^{*}\left[n,\boldsymbol{\tau}\right] & \mathbf{s}^{*}\left[n,\boldsymbol{\tau}_0\right] \end{bmatrix}, \tag{1.3.23}$$

which is clearly non-negative definite. Therefore, the limiting cost function satisfies

$$V_\infty(\boldsymbol{\tau})\geq\log\left|\mathbf{I}-\mathbf{R}_{yy}^{-\frac{1}{2}}\mathbf{A}\mathbf{C}_{ss}\left(\boldsymbol{\tau}_0,\boldsymbol{\tau}_0\right)\mathbf{A}^{*}\mathbf{R}_{yy}^{-\frac{1}{2}}\right|=V_\infty(\boldsymbol{\tau}_0) \tag{1.3.24}$$

The equality in (1.3.24) holds if and only if the Schur complement $\mathbf{M}(\boldsymbol{\tau}_0,\boldsymbol{\tau})$ is zero. This is only possible for $\boldsymbol{\tau}=\boldsymbol{\tau}_0$ if the following *non-ambiguity condition* is fulfilled: The matrix $\mathbf{C}_{ss}(\widetilde{\boldsymbol{\tau}},\widetilde{\boldsymbol{\tau}})$ is positive definite for any vector $\widetilde{\boldsymbol{\tau}}$ of length $2d$ whose elements are all distinct. This condition is equivalent to the one presented in [44] for the estimation of directions of arrival with large arrays. The consistency of $\hat{\mathbf{A}}_{ML}$ and $\hat{\mathbf{Q}}_{ML}$ follows immediately from (1.3.20)-(1.3.21) and the consistency of $\hat{\boldsymbol{\tau}}_{ML}$.

1.3.2 Cramér-Rao Bound

Since the ML estimates of all of the parameters are consistent, they will also be asymptotically (large N) efficient (*i.e.,* their asymptotic covariance coincides with

the Cramér-Rao bound (CRB)). Using a straightforward modification of the analysis in [45], [46], it can be shown that the CRB for the time delays has the following form:

$$\mathbf{CRB}^{-1}(\boldsymbol{\tau}) = 2\,\mathrm{Re}\left\{ \left(\mathbf{D}\,\mathbf{P}^{\perp}_{\mathbf{S}^*}\,\mathbf{D}^*\right) \odot \left(\mathbf{A}^*\,\mathbf{Q}^{-1}\,\mathbf{A}\right)^T \right\} \tag{1.3.25}$$

where

$$\mathbf{D}(\boldsymbol{\tau}) = \left[\begin{array}{ccc} \mathbf{d}\,(\tau_1) & \cdots & \mathbf{d}\,(\tau_d) \end{array} \right]^T \tag{1.3.26}$$

$$\mathbf{d}\,(\tau_i) = -\left[\begin{array}{ccc} \frac{d\,s(t)}{d\,t}\Big|_{T_s-\tau_i} & \cdots & \frac{d\,s(t)}{d\,t}\Big|_{NT_s-\tau_i} \end{array} \right]^T \tag{1.3.27}$$

$$\mathbf{P}^{\perp}_{\mathbf{S}^*(\boldsymbol{\tau})} = \mathbf{I} - \mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})} = \mathbf{I} - \mathbf{S}^*(\boldsymbol{\tau})\,(\mathbf{S}(\boldsymbol{\tau})\mathbf{S}^*(\boldsymbol{\tau}))^{-1}\,\mathbf{S}(\boldsymbol{\tau})\,. \tag{1.3.28}$$

1.3.3 Computation of the Estimates

For the frame delay estimation problem described in (P2), minimizing (1.3.18) requires only a one-dimensional search, and thus is not overly burdensome. This is in fact the solution considered in [19]. For (P1), however, the complicated non-linear dependence of V_N on $\boldsymbol{\tau}$, especially due to the presence of the determinant, implies that a multidimensional search is the only method that can be used to find the estimates. Although the search can be implemented in more sophisticated ways than brute force evaluation of the cost function on a multidimensional grid (using, for example, a gradient search, expectation-maximization, or alternating projections), a more computationally efficient solution is still desirable.

A simpler solution is possible if the interference and noise are assumed to be spatially white (*e.g.,* no CCI). Under this assumption, the criterion reduces to a trace rather than determinant operation. By transforming the data to the frequency domain, an iterative solution based on the so-called IQML (Iterative Quadratic Maximum Likelihood) [14] or MODE [15], [47] can be used. The computational advantage of these techniques results because the cost function depends linearly on the signal projection matrix $\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}$. By reparameterizing the matrix $\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}$ according to the coefficients of a certain polynomial, and assuming a previous estimate of these coefficients is available, the dependence of $\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}$ and hence the cost function on some trial coefficients becomes quadratic. The quadratic optimization problem is then solved in closed form (see [13] for details).

The ML cost function for both spatially and temporally white noise is

$$f^w_N(\boldsymbol{\tau}) = -\mathrm{Tr}\left\{ \widehat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\,\widehat{\mathbf{R}}^{-1}_{ss}(\boldsymbol{\tau})\,\widehat{\mathbf{R}}^*_{ys}(\boldsymbol{\tau}) \right\} = -\frac{1}{N}\mathrm{Tr}\left\{ \mathbf{Y}\,\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}\,\mathbf{Y}^* \right\} \tag{1.3.29}$$

which satisfies the condition of linear dependence on $\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}$ stated above. This condition is not fulfilled by the ML cost function for unknown correlated noise in (1.3.18) due to the log-determinant operation. Consequently, an IQML-like algorithm can not be directly applied to (1.3.18). The main goal of this chapter is to present and analyze a cost function that is asymptotically equivalent to the original

ML criterion (1.3.18), but that is linear in the signal projection matrix and, therefore makes possible the computation of the estimates using an IQML or MODE approach.

1.4 An Asymptotically Equivalent Estimator

We propose computing the delay estimates as the minimizing arguments of the following criterion function:

$$g_N\left(\boldsymbol{\tau}, \mathbf{W}_0\right) = -\mathrm{Tr}\left\{\mathbf{W}_0\, \mathbf{B}_N(\boldsymbol{\tau})\right\} \tag{1.4.1}$$

where

$$\mathbf{W}_0 \triangleq \left(\mathbf{I} - \mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1} \tag{1.4.2}$$

and $\mathbf{B}_N(\boldsymbol{\tau}_0)$ is defined as in (1.3.19). A proof that this criterion yields asymptotically efficient delay estimates is given below.

1.4.1 Proof of the Asymptotic Equivalence

To begin, we note that it can be shown that the estimates obtained with (1.4.1) are consistent. The proof is similar to that in Section 1.3.1 and will be omitted. We will now establish the asymptotic equivalence between

$$\hat{\boldsymbol{\tau}}_1 = \arg\min_{\boldsymbol{\tau}}\ V_N(\boldsymbol{\tau}) \tag{1.4.3}$$

$$\hat{\boldsymbol{\tau}}_2 = \arg\min_{\boldsymbol{\tau}}\ g_N\left(\boldsymbol{\tau}, \mathbf{W}_0\right)\ , \tag{1.4.4}$$

which means that

$$\hat{\boldsymbol{\tau}}_2 = \hat{\boldsymbol{\tau}}_1 + o_p\left(N^{-1/2}\right)\ , \tag{1.4.5}$$

where $x = o_p(\gamma)$ means that x/γ converges to zero in probability. A sufficient condition for (1.4.5) to hold is that [39]

$$g_N^i\left(\boldsymbol{\tau}_0, \mathbf{W}_0\right) = V_N^i(\boldsymbol{\tau}_0) + o_p\left(N^{-1/2}\right) \tag{1.4.6}$$

$$g_N^{ij}\left(\boldsymbol{\tau}_0, \mathbf{W}_0\right) = V_N^{ij}(\boldsymbol{\tau}_0) + o_p\left(1\right) \tag{1.4.7}$$

where the superscript $(\cdot)^i$ denotes the derivative with respect to τ_i. A double superscript denotes the corresponding second derivatives.

The proof of (1.4.6) is immediate since

$$V_N^i(\boldsymbol{\tau}_0) = -\mathrm{Tr}\left\{\left(\mathbf{I} - \mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1}\, \mathbf{B}_N^i(\boldsymbol{\tau}_0)\right\} = g_N^i\left(\boldsymbol{\tau}_0, \mathbf{W}_0\right)\ . \tag{1.4.8}$$

The second derivatives also satisfy the equivalence condition (1.4.7) because

$$V_N^{ij}(\boldsymbol{\tau}_0) = -\mathrm{Tr}\left\{\left(\mathbf{I} - \mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1}\, \mathbf{B}_N^{ij}(\boldsymbol{\tau}_0)\right\}$$

$$
\begin{aligned}
&+\mathrm{Tr}\left\{ \left(\mathbf{I}-\mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1} \mathbf{B}_N^i(\boldsymbol{\tau}_0) \left(\mathbf{I}-\mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1} \mathbf{B}_N^j(\boldsymbol{\tau}_0) \right\} \\
&= g_N^{ij}\left(\boldsymbol{\tau}_0, \mathbf{W}_0\right) \\
&\quad +\mathrm{Tr}\left\{ \left(\mathbf{I}-\mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1} \mathbf{B}_N^i(\boldsymbol{\tau}_0) \left(\mathbf{I}-\mathbf{B}_N(\boldsymbol{\tau}_0)\right)^{-1} \mathbf{B}_N^j(\boldsymbol{\tau}_0) \right\} \quad (1.4.9)
\end{aligned}
$$

and the last term in (1.4.9) is (at least) $o_p(1)$ since, as shown in Appendix .1, $\mathbf{B}_N^i(\boldsymbol{\tau}_0) = O_p\left(N^{-1/2}\right)$, where $x = O_p(\gamma)$ means that x/γ is bounded in probability.

1.4.2 Calculation of the Weighting Matrix

The weighting matrix $\mathbf{W}_0$ appearing in the proposed cost function depends on the true value of the delays, and hence is unknown. However, it is well known that we can replace it with a consistent estimate $\hat{\mathbf{W}}$ without affecting the asymptotic properties of the estimates. If $\hat{\boldsymbol{\tau}}$ is a consistent estimate of $\boldsymbol{\tau}_0$, then we can construct the *practical* weighting matrix as

$$
\hat{\mathbf{W}} = \left(\mathbf{I}-\mathbf{B}_N\left(\hat{\boldsymbol{\tau}}\right)\right)^{-1} . \quad (1.4.10)
$$

The estimates obtained from $g_N(\boldsymbol{\tau}, \mathbf{W}_0)$ and $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ are asymptotically equivalent because the derivatives of these cost functions satisfy conditions similar to those stated in (1.4.6) and (1.4.7). The proof is immediate since $\hat{\mathbf{W}} = \mathbf{W}_0 + o_p(1)$ (by the definition of consistency), $g_N^i\left(\boldsymbol{\tau}_0, \mathbf{W}_0\right) = O_p\left(N^{-1/2}\right)$ and $g_N^{ij}\left(\boldsymbol{\tau}_0, \mathbf{W}_0\right) = O_p(1)$. It is worth remarking that the *practical* cost function admits the following expression:

$$
g_N(\boldsymbol{\tau}, \hat{\mathbf{W}}) = -\mathrm{Tr}\left\{ \hat{\mathbf{Q}}^{-1/2}\, \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\, \hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\, \hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau})\, \hat{\mathbf{Q}}^{-1/2} \right\} \quad (1.4.11)
$$

where $\hat{\mathbf{Q}} = \hat{\mathbf{R}}_{yy} - \hat{\mathbf{R}}_{ys}(\hat{\boldsymbol{\tau}})\, \hat{\mathbf{R}}_{ss}^{-1}(\hat{\boldsymbol{\tau}})\, \hat{\mathbf{R}}_{ys}^*(\hat{\boldsymbol{\tau}})$ is a consistent estimate of the correlation matrix of the noise. The criterion above resembles the one in the white-noise case (1.3.29); the difference is that now the signals are prewhitened using an estimate of the noise correlation. While the function in (1.4.11) could have been derived using purely heuristic reasoning, the development followed herein has allowed us to prove the equivalence between (1.4.11) and the original criterion (1.3.18), which would have been difficult to do from a simple inspection of those cost functions.

The consistent estimate of the time-delays needed to construct $\hat{\mathbf{W}}$ can be obtained as the minimizing argument of

$$
\begin{aligned}
h_N(\boldsymbol{\tau}) \triangleq g_N\left(\boldsymbol{\tau}, \mathbf{I}\right) &= -\mathrm{Tr}\left\{\mathbf{B}_N(\boldsymbol{\tau})\right\} \quad (1.4.12) \\
&= -\mathrm{Tr}\left\{ \hat{\mathbf{R}}_{yy}^{-1/2}\, \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau})\, \hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau})\, \hat{\mathbf{R}}_{ys}^*(\boldsymbol{\tau})\, \hat{\mathbf{R}}_{yy}^{-1/2} \right\} , \quad (1.4.13)
\end{aligned}
$$

in which the unknown weighting matrix is replaced by the identity, which amounts to prewhitening the signals according to the total correlation matrix $\hat{\mathbf{R}}_{yy}$ instead of the correlation of the noise, as done in (1.4.11). Again, the proof of the consistency of $h_N(\boldsymbol{\tau})$ is similar to that in Section 1.3.1 and will be omitted. Note that the criterion $f_N^w(\boldsymbol{\tau})$ obtained for spatially white noise also provides consistent estimates.

Nevertheless, as illustrated in the simulations of Section 1.7, its variance will be much larger than that of the estimates obtained with (1.4.13) when the noise is not spatially white since $f_N^w(\boldsymbol{\tau})$ makes no attempt to prewhiten the signals. Therefore, there is no advantage in employing $f_N^w(\boldsymbol{\tau})$ instead of $h_N(\boldsymbol{\tau})$ (apart from a slightly reduced computational complexity), and the latter is preferred.

1.5 Heuristic Derivations

In the previous section we have directly put forward a new cost function $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ and determined its statistical properties. In this section, three different heuristic ways of deriving the new cost function are presented.

1.5.1 Series Expansion of the Logarithm

Note that

$$\begin{aligned}\mathbf{B}_N(\boldsymbol{\tau}) &= \hat{\mathbf{R}}_{yy}^{-\frac{1}{2}} \hat{\mathbf{R}}_{ys}(\boldsymbol{\tau}) \hat{\mathbf{R}}_{ss}^{-1}(\boldsymbol{\tau}) \hat{\mathbf{R}}_{ys}^{*}(\boldsymbol{\tau}) \hat{\mathbf{R}}_{yy}^{-\frac{1}{2}} \\ &= \underbrace{(\mathbf{Y}\mathbf{Y}^*)^{-1/2}\mathbf{Y}\mathbf{S}^*(\boldsymbol{\tau}) \left(\mathbf{S}(\boldsymbol{\tau})\mathbf{S}^*(\boldsymbol{\tau})\right)^{-1} \mathbf{S}(\boldsymbol{\tau})}_{\mathbf{X}} \cdot \underbrace{\mathbf{Y}^*(\mathbf{Y}\mathbf{Y}^*)^{-1/2}}_{\mathbf{Z}} \, .\end{aligned}$$

If we compare the above equation with

$$\begin{aligned}\mathbf{P}_{\mathbf{Y}^*} \mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})} &= \mathbf{Y}^* \left(\mathbf{Y}\,\mathbf{Y}^*\right)^{-1} \mathbf{Y}\mathbf{S}^*(\boldsymbol{\tau}) \left(\mathbf{S}(\boldsymbol{\tau})\,\mathbf{S}^*(\boldsymbol{\tau})\right)^{-1} \mathbf{S}(\boldsymbol{\tau}) \\ &= \underbrace{\mathbf{Y}^*(\mathbf{Y}\mathbf{Y}^*)^{-1/2}}_{\mathbf{Z}} \cdot \underbrace{(\mathbf{Y}\mathbf{Y}^*)^{-1/2}\mathbf{Y}\mathbf{S}^*(\boldsymbol{\tau}) \left(\mathbf{S}(\boldsymbol{\tau})\mathbf{S}^*(\boldsymbol{\tau})\right)^{-1} \mathbf{S}(\boldsymbol{\tau})}_{\mathbf{X}} \, ,\end{aligned}$$

we see that $\mathbf{B}_N(\boldsymbol{\tau})$ and the product of the two projection matrices above must share the same non-zero eigenvalues. Thus, the eigenvalues of $\mathbf{B}_N(\boldsymbol{\tau})$ satisfy $0 \leq \lambda_i \leq 1$, or in the presence of noise $\lambda_i < 1$. Therefore, we can use the series expansion of the logarithm to express the original ML criterion as follows:

$$V_N(\boldsymbol{\tau}) = -\mathrm{Tr}\left\{\mathbf{B}_N(\boldsymbol{\tau})\right\} - \frac{1}{2}\mathrm{Tr}\left\{\mathbf{B}_N^2(\boldsymbol{\tau})\right\} - \frac{1}{3}\mathrm{Tr}\left\{\mathbf{B}_N^3(\boldsymbol{\tau})\right\} + \cdots \, . \tag{1.5.1}$$

The function $h_N(\boldsymbol{\tau})$, which we have proposed to use in obtaining the initial consistent estimates, is the first term of this expansion. Unlike many other estimation problems (see *e.g.*, [48], [49], [50]), the first-order term is not asymptotically equivalent to the original function because $\lim_{N\to\infty} \mathbf{B}_N(\boldsymbol{\tau}_0) = \mathbf{I} - \mathbf{R}_{yy}^{-1/2}\mathbf{Q}\mathbf{R}_{yy}^{-1/2}$ is not equal to zero. The error in this first order approximation is "small" when all the eigenvalues of $\mathbf{Q}$ are large with respect to the power of the signals. However, this situation will seldom be encountered in practice, where usually only some of eigenvalues of $\mathbf{Q}$ are large, due to the reception of directional interferers. In order to maintain not only consistency but also asymptotic efficiency, all of the terms in the expansion (1.5.1) must be kept. Since the second and higher-order terms are the ones that introduce the undesirable nonlinear dependence on the matrix $\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}$, we

decide to approximate them. The approximation is not done directly over $V_N(\boldsymbol{\tau})$, but over its derivative. If we differentiate (1.5.1) and replace $\mathbf{B}_N(\boldsymbol{\tau})$ by $\mathbf{B}_N(\hat{\boldsymbol{\tau}})$ in all the second and higher-order terms (this is justified since $\mathbf{B}_N^i(\boldsymbol{\tau}_0) = O_p\left(N^{-1/2}\right)$), it results that

$$V_N^i(\boldsymbol{\tau}) \simeq -\mathrm{Tr}\left\{\mathbf{B}_N^i(\boldsymbol{\tau})\left(\mathbf{I} + \mathbf{B}_N(\hat{\boldsymbol{\tau}}) + \mathbf{B}_N^2(\hat{\boldsymbol{\tau}}) + \cdots\right)\right\} \tag{1.5.2}$$

$$= -\mathrm{Tr}\left\{\left(\mathbf{I} - \mathbf{B}_N(\hat{\boldsymbol{\tau}})\right)^{-1}\mathbf{B}_N^i(\boldsymbol{\tau})\right\} . \tag{1.5.3}$$

Thus we retrieve the new criterion presented in Section 1.4 since the value of $\boldsymbol{\tau}$ that minimizes $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ also nulls (1.5.3).

1.5.2 Eigenvalue Weighting

The derivative of the original ML criterion can be written as

$$V_N^i(\boldsymbol{\tau}) = -\sum_{k=1}^{m} \frac{\lambda_k^i(\boldsymbol{\tau})}{1 - \lambda_k(\boldsymbol{\tau})} \tag{1.5.4}$$

while the derivative of the cost function that provides only consistent estimates is

$$h_N^i(\boldsymbol{\tau}) = -\sum_{k=1}^{m} \lambda_k^i(\boldsymbol{\tau}) . \tag{1.5.5}$$

We notice that the difference between the estimator that is asymptotically efficient and the one that is not lies in an appropriate weighting of the eigenvalues. The second criterion (1.5.5) approaches the original one (1.5.4) when all the nonzero eigenvalues are much smaller than one, or when all of them have similar values. Again, this only happens if all the eigenvalues of $\mathbf{Q}$ are much larger than the power of the desired signals. A reasonable approach to approximating the optimal weighting would be to replace the eigenvalues $\lambda_k(\boldsymbol{\tau})$ in (1.5.4) by the eigenvalues of $\mathbf{B}_N(\hat{\boldsymbol{\tau}})$. Using (1.5.4) and the eigendecomposition $\mathbf{B}_N(\boldsymbol{\tau}) = \mathbf{U}(\boldsymbol{\tau})\,\boldsymbol{\Lambda}(\boldsymbol{\tau})\,\mathbf{U}^*(\boldsymbol{\tau})$, this approach results in

$$V_N(\boldsymbol{\tau}) \simeq -\mathrm{Tr}\left\{\left(\mathbf{I} - \boldsymbol{\Lambda}(\hat{\boldsymbol{\tau}})\right)^{-1}\,\boldsymbol{\Lambda}(\boldsymbol{\tau})\right\} \tag{1.5.6}$$

$$= -\mathrm{Tr}\left\{\mathbf{U}(\boldsymbol{\tau})\left(\mathbf{I} - \boldsymbol{\Lambda}(\hat{\boldsymbol{\tau}})\right)^{-1}\mathbf{U}^*(\boldsymbol{\tau})\,\mathbf{U}(\boldsymbol{\tau})\boldsymbol{\Lambda}(\boldsymbol{\tau})\mathbf{U}^*(\boldsymbol{\tau})\right\} \tag{1.5.7}$$

$$\simeq -\mathrm{Tr}\left\{\mathbf{U}(\hat{\boldsymbol{\tau}})\left(\mathbf{I} - \boldsymbol{\Lambda}(\hat{\boldsymbol{\tau}})\right)^{-1}\mathbf{U}^*(\hat{\boldsymbol{\tau}})\,\mathbf{B}_N(\boldsymbol{\tau})\right\} \tag{1.5.8}$$

$$= -\mathrm{Tr}\left\{\left(\mathbf{I} - \mathbf{B}_N(\hat{\boldsymbol{\tau}})\right)^{-1}\mathbf{B}_N(\boldsymbol{\tau})\right\} = g_N(\boldsymbol{\tau}, \hat{\mathbf{W}}) \tag{1.5.9}$$

which is the cost function proposed in Section 1.4.

Note that in the two approaches above the approximations are always carried out on the derivative of the ML cost function, and next the function $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ results by integration. If the approximations had been performed directly on the ML cost function, the resulting criterion would not have been asymptotically efficient.

1.5.3 First Order Approximation

As stated above, a direct first-order approximation of the ML cost function does not yield an asymptotically efficient estimator. However, we can write the ML cost function as

$$\begin{aligned} V_N(\boldsymbol{\tau}) &= \log|\mathbf{I} - \mathbf{B}_N(\hat{\boldsymbol{\tau}}) + \mathbf{B}_N(\hat{\boldsymbol{\tau}}) - \mathbf{B}_N(\boldsymbol{\tau})| \\ &= \log|\mathbf{I} - \mathbf{B}_N(\hat{\boldsymbol{\tau}})| + \log\left|\mathbf{I} + (\mathbf{I} - \mathbf{B}_N(\hat{\boldsymbol{\tau}}))^{-1}(\mathbf{B}_N(\hat{\boldsymbol{\tau}}) - \mathbf{B}_N(\boldsymbol{\tau}))\right| \end{aligned} \quad (1.5.10)$$

Since $\lim_{N\to\infty}(\mathbf{B}_N(\hat{\boldsymbol{\tau}}) - \mathbf{B}_N(\boldsymbol{\tau}_0)) = 0$ and $\mathbf{B}_N^i(\boldsymbol{\tau}_0) = O_p\left(N^{-1/2}\right)$, it is possible to maintain the asymptotic efficiency by only keeping the first-order term in the series expansion of (1.5.10), which is

$$V_N(\boldsymbol{\tau}) \underset{\substack{\text{first order}\\ \text{term}}}{\simeq} \log|\mathbf{I} - \mathbf{B}_N(\hat{\boldsymbol{\tau}})| + \mathrm{Tr}\left\{\hat{\mathbf{W}}\,\mathbf{B}_N(\hat{\boldsymbol{\tau}})\right\} - \mathrm{Tr}\left\{\hat{\mathbf{W}}\,\mathbf{B}_N(\boldsymbol{\tau})\right\} = c + g_N(\boldsymbol{\tau}, \hat{\mathbf{W}}) \;, \quad (1.5.11)$$

where c is a constant. Once more, this coincides with the alternative function we have proposed.

1.6 Calculating the Estimates with IQML and ESPRIT

In this section we outline how the IQML and ESPRIT algorithms can be applied to the cost functions that have appeared in the previous sections. Since the objective in the preceding sections has been to reduce the complexity involved in a direct minimization of a multidimensional criterion, it does not make sense to optimize the new cost functions using a search; instead, the use of computationally efficient algorithms is more appropriate. We focus on applications to problem (P1), where multiple time delays must be estimated. Only a one-dimensional search is required to implement the exact ML criterion for problem (P2), so the techniques described below likely do not offer a significant computational savings for this case.

The general expression of the cost function we consider is that given in (1.4.1), which can be written as

$$g_N(\boldsymbol{\tau}, \mathbf{W}) = -\frac{1}{N}\mathrm{Tr}\left\{\mathbf{W}^{1/2}\,\hat{\mathbf{R}}_{yy}^{-1/2}\,\mathbf{Y}\,\mathbf{P}_{\mathbf{S}^*(\boldsymbol{\tau})}\,\mathbf{Y}^*\,\hat{\mathbf{R}}_{yy}^{-1/2}\,\mathbf{W}^{1/2}\right\} \;. \quad (1.6.1)$$

Different criteria are obtained from different choices of the matrix $\mathbf{W}$. That is, if $\mathbf{W}$ is a consistent estimate of $\mathbf{W}_0$ in (1.4.2) then the asymptotically efficient estimator is obtained; if $\mathbf{W}$ is equal to the identity matrix or equal to $\hat{\mathbf{R}}_{yy}$ then the consistent estimator $h_N(\boldsymbol{\tau})$ or the white-noise estimator $f_N^w(\boldsymbol{\tau})$ result, respectively.

If the N temporal samples are transformed into the frequency domain using the DFT, the signals approximately satisfy the following relationship:

$$\mathbf{S}^*(\boldsymbol{\tau}) = \mathbf{S}_\omega^*\,\mathbf{V}(\boldsymbol{\tau}) \quad (1.6.2)$$

where $\mathbf{S}_\omega$ is a diagonal matrix whose entries are the DFT of the vector samples $[s(T_s), \cdots, s(NT_s)]$, and

$$\mathbf{V}(\boldsymbol{\tau}) = \left[\begin{array}{ccc} \mathbf{v}(\tau_1) & \cdots & \mathbf{v}(\tau_N) \end{array} \right] \tag{1.6.3}$$

$$\mathbf{v}(\tau_k) = \left[\begin{array}{ccc} \exp(j\omega_1\tau_k) & \cdots & \exp(j\omega_N\tau_k) \end{array} \right]^T \tag{1.6.4}$$

$$\omega_i = \frac{2\pi}{NT_s}\left(i - 1 - \text{floor}\left(\frac{N}{2}\right)\right) . \tag{1.6.5}$$

Note that the same notation $\mathbf{S}(\boldsymbol{\tau})$ is used for both the time and frequency domain. It should be clear from the context which one is being referred to. Since the noise is assumed to be white in time, it is also white in frequency, so all of the estimators above can be applied in an identical fashion to the frequency rather than time samples.

1.6.1 IQML Algorithm

Let the elements of the vector $\mathbf{g} = [g_0 \ \cdots \ g_d]^T$ be taken from the coefficients of the polynomial

$$g(z) = g_0 z^d + g_1 z^{d-1} + \cdots + g_d \tag{1.6.6}$$

whose roots are $\exp(j2\pi\tau_1/NT_s), \cdots, \exp(j2\pi\tau_d/NT_s)$. Since the roots lie on the unit circle, the coefficients satisfy the so-called conjugate symmetry constraint:

$$g_k = g_{d-k}^*, \quad k = 0, 1, \cdots, d . \tag{1.6.7}$$

It is straightforward to prove the following equality of projection matrices [13]:

$$\mathbf{P}^{\perp}_{\mathbf{S}^*(\boldsymbol{\tau})} = \mathbf{P}_{\mathbf{S}_\omega^{-1}\mathbf{G}} = \mathbf{S}_\omega^{-1}\mathbf{G}\left(\mathbf{G}^*\mathbf{S}_\omega^{-*}\mathbf{S}_\omega^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^*\mathbf{S}_\omega^{-*} \tag{1.6.8}$$

where the $N \times N - d$ Sylvester matrix $\mathbf{G}$ is given by

$$\mathbf{G} = \left[\begin{array}{ccccccc} g_0 & g_1 & \cdots & g_d & 0 & & \\ 0 & g_0 & g_1 & \cdots & g_d & 0 & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ & & 0 & g_0 & g_1 & \cdots & g_d \end{array} \right]^T \tag{1.6.9}$$

Therefore, minimizing the cost function in (1.6.1) is equivalent to minimizing

$$\tilde{g}_N(\mathbf{g}, \mathbf{W}) = \frac{1}{N}\text{Tr}\left\{\mathbf{W}^{\frac{1}{2}}\hat{\mathbf{R}}_{yy}^{-\frac{1}{2}}\mathbf{Y}\mathbf{S}_\omega^{-1}\mathbf{G}\left(\mathbf{G}^*\mathbf{S}_\omega^{-*}\mathbf{S}_\omega^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^*\mathbf{S}_\omega^{-*}\mathbf{Y}^*\hat{\mathbf{R}}_{yy}^{-\frac{1}{2}}\mathbf{W}^{\frac{1}{2}}\right\} \tag{1.6.10}$$

In the IQML algorithm the minimization of (1.6.10) is done iteratively. That is, the matrix $\left(\mathbf{G}_k^*\mathbf{S}_\omega^{-*}\mathbf{S}_\omega^{-1}\mathbf{G}_k\right)^{-1}$ is computed using a given estimate $\mathbf{g}_k$ and held fixed. Then, the resulting criterion is quadratic in $\mathbf{g}_k$ and can be solved in closed-form. Details on the implementation of this step subject to the conjugate symmetry

constraint (1.6.7) and to a certain constraint that avoids the trivial solution (*e.g.* $\|\mathbf{g}_k\| = 1$, or $\mathrm{Re}\{g_0\} = 1$) can be found in [13], [14], [51]. The resulting vector $\mathbf{g}_{k+1}$ is used to fix $\left(\mathbf{G}_{k+1}^{*}\,\mathbf{S}_{\omega}^{-*}\,\mathbf{S}_{\omega}^{-1}\,\mathbf{G}_{k+1}\right)^{-1}$, and the process is repeated until a certain convergence or failure criterion is satisfied (see Section 1.7.1).

If we want the matrix $\mathbf{W}$ to be a consistent estimate of $\mathbf{W}_0$, then we have two iterative processes: the IQML algorithm itself and the computation of the matrix $\mathbf{W}$. Since these two processes can be coupled or not, we can choose two different approaches to implement the complete estimation procedure:

A) Coupled iterations

1. Initialize $k = 0$, $\mathbf{W}_0 = \mathbf{I}$ and $\mathbf{g}_0$ (see Section 1.7.1).
2. Do *only one* iteration of the IQML algorithm and obtain $\mathbf{g}_{k+1}$.
3. Compute $\boldsymbol{\tau}_{k+1}$ from $\mathbf{g}_{k+1}$ using (1.6.6).
4. Compute the weighting matrix $\mathbf{W}_{k+1}$ using $\boldsymbol{\tau}_{k+1}$ and expression (1.4.10). Substitute $\mathbf{W}_{k+1}$ in the cost function that is being minimized (1.6.10).
5. If the convergence / failure condition is satisfied, take the estimate of the delays as $\hat{\boldsymbol{\tau}} = \boldsymbol{\tau}_{k+1}$. If not, set $k = k + 1$ and return to step 2.

B) Decoupled iterations

1. Initialize $\mathbf{W}_0 = \mathbf{I}$ and $\mathbf{g}_0$ (see Section 1.7.1).
2. Perform *all* iterations of the IQML algorithm until it converges or fails (not only one iteration as in the procedure **A**). The result is the vector $\mathbf{g}_1$, and the corresponding time-delay estimate $\boldsymbol{\tau}_1$ is obtained using (1.6.6).
3. Compute the weighting matrix $\mathbf{W}_1$ using $\boldsymbol{\tau}_1$ and expression (1.4.10). Introduce $\mathbf{W}_1$ in the cost function (1.6.10).
4. Again perform *all* iterations of the IQML algorithm until convergence or failure occurs. The result is the vector $\mathbf{g}_2$, and the corresponding time-delay estimate is the final estimate: $\hat{\boldsymbol{\tau}} = \boldsymbol{\tau}_2$.

As we will show in the simulation results, the method **A** presents a lower estimation error and requires less iterations of the IQML algorithm to converge than the method **B**. On the other hand, their computational loads are similar, since in the method **B** the weighting matrix is computed only once. Note that the decoupled iterations method is the one that stems directly from the theoretical results. That is, the delay estimation is divided into two stages: *i)* obtaining the consistent estimate, *ii)* obtaining the asymptotically efficient estimate. However, the coupled iteration method is a logical *ad hoc* modification of the decoupled one given the two iterative processes needed by the proposed estimator. The coupled iterations method does not stem directly from the theoretical study, but it happens to have certain advantages and is preferred in front of the decoupled one.

1.6.2 ESPRIT Algorithm

If $d < m$, $d < N$ and $\mathbf{A}$ is full rank, it is possible to exploit the Vandermonde structure of $\mathbf{V}(\boldsymbol{\tau})$ using the ESPRIT algorithm [7]. The application of this method to time-delay estimation in the white-noise case is detailed in [13]. Given the similarity between the cost functions $f_N^w(\boldsymbol{\tau})$ in (1.3.29) and $g_N(\boldsymbol{\tau}, \mathbf{W})$ in (1.6.1), the extension of ESPRIT to the new criterion is immediate. In our case, the matrix $\mathbf{F} = \mathbf{Y}^* \hat{\mathbf{R}}_{yy}^{-1/2} \mathbf{W}^{1/2}$ plays exactly the same role as $\mathbf{Y}^*$ in the white-noise case.

Let $\mathbf{E}$ denote the d singular vectors of $\mathbf{F}$ associated with the largest singular values. Next, $\mathbf{E}_1$ (resp. $\mathbf{E}_2$) is constructed by taking the first (resp. last) $N-\delta$ rows of $\mathbf{E}$. The matrices $\mathbf{S}_{\omega 1}$ and $\mathbf{S}_{\omega 2}$ are defined similarly. Let the $d \times d$ matrix $\boldsymbol{\Psi}$ be the solution of the following overdetermined system of linear equations:

$$\mathbf{S}_{\omega 1}^* \mathbf{E}_2 = \mathbf{S}_{\omega 2}^* \mathbf{E}_1 \boldsymbol{\Psi} \tag{1.6.11}$$

The time-delay estimates can be determined from the phase ($\angle$) of the eigenvalues λ_i of $\boldsymbol{\Psi}$ (see [7], [13]):

$$\hat{\tau}_i = \frac{N T_s \angle \lambda_i}{2\pi \delta} \quad i = 1, \cdots, d \; . \tag{1.6.12}$$

The shift parameter δ must satisfy two conditions:

$$\delta \leq N - d \quad , \quad \delta < \frac{N T_s}{2 \max_i \tau_i} \tag{1.6.13}$$

The first one prevents the system (1.6.11) from being indefinite, while the second guarantees that the association between eigenvalues and delays is unique. It is possible to solve the system of equations (1.6.11) using either Least Squares (LS) or Total Least Squares (TLS) [52].

1.7 Simulation Results

In this section we analyze the performance of the cost functions proposed above by means of numerical simulations. Our performance metric is the Root Mean Squared Error (RMSE) of the delay estimates produced by each algorithm. We calculate the RMSE for a wide variety of scenarios as a function of the number of samples N, the number of sensors m, the signal to interference ratio, and the relative delay and DOA of the multipath reflections.

1.7.1 Simulation Parameters

All simulations are conducted assuming that $d = 2$ delayed versions of a known signal are received by a uniform linear array with antennas spaced 0.5λ apart. This known signal is a concatenation of K truncated and sampled Nyquist square-root raised cosine pulses. Each pulse has a bandwidth equal to $(1+\alpha)/2T_c$, is truncated to the interval $[-3T_c, 3T_c]$, and the sampling period is $T_c/2$, so there are 13 samples in each pulse (see Figure 1.1). Note that T_c is simply a normalization constant, and

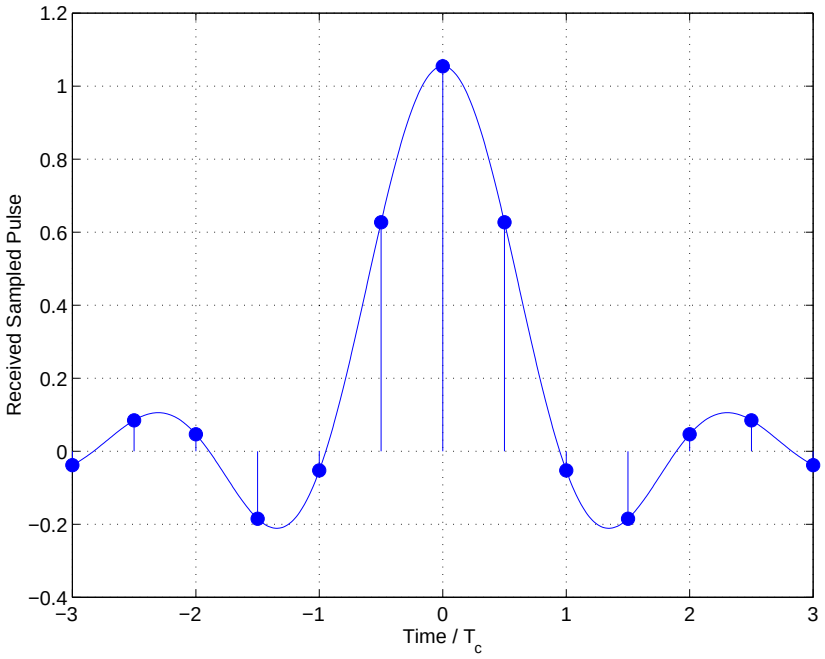

Figure 1.1. Samples of a single pulse and the underlying analog signal.

α is the roll-off factor which we set equal to 0.2. The use of this type of signal is of interest because each pulse may represent the output of the despreader at every symbol period in a DS-CDMA system. In this type of system, accurate timing estimation is fundamental to achieving satisfactory performance.

The noise plus interference field in which the array operates consists of: *i)* spatially and temporally white Gaussian noise, and *ii)* a temporally white Gaussian interference at DOA -30^o relative to the array broadside, which is responsible for the spatial correlation of the noise plus interference field. Both the noise and the interference are uncorrelated with the desired signal. The remaining scenario parameters, except when one of them is varied, are as follows: $m = 6$ antennas; $K = 3$ pulses constituting each signal; delays of the two signals equal to 0 and $0.4\,T_c$; DOAs of those signals: 0^o, 10^o, respectively; Signal to Noise Ratio (SNR) of the first signal: 16 dB; Signal to Interference Ratio (SIR) of the first signal: -3 dB; the second signal is attenuated 3 dB with respect to the first, and they are in phase at the first sensor. In all cases, only the 7 DFT bins with the strongest signal content are employed by IQML and ESPRIT for estimation of the time delays.

In the figures below, we show the RMSE in the estimation of the time-delay of the first signal (the time-delay estimate of the second signal behaves similarly) obtained from different cost functions and optimization algorithms. The RMSEs are

computed from 500 Monte Carlo realizations of the noise and interference, and they are compared to the CRB given by (1.3.25). Each curve in the figures corresponds to one of the following methods:

1. Asymptotically efficient estimator $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ given by (1.4.11). The minimization is carried out using IQML, with the weight matrix $\hat{\mathbf{W}}$ recomputed at every IQML iteration (*i.e.*, the coupled iteration method **A** of Section 1.6.1).

2. Consistent estimator $g_N(\boldsymbol{\tau}, \mathbf{I})$ given by (1.4.13). The estimates are calculated using LS-ESPRIT, since we have observed that in this case ESPRIT somewhat surprisingly yields a lower RMSE than IQML.

3. ML estimator in spatially white noise $f_N^w(\boldsymbol{\tau})$ given by (1.3.29), with the estimates obtained via IQML as in [13].

4. Asymptotically efficient estimator $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ given by (1.4.11). The minimization is carried out using IQML twice, but the matrix $\hat{\mathbf{W}}$ is computed after IQML has converged (*i.e.* the decoupled iteration method **B** in Section 1.6.1). The performance of this method is shown only in Figure 1.2, and it is referred to as IQML "2 iterations".

Only the RMSEs are plotted because these four methods are essentially unbiased (*i.e.*, their biases are much smaller than their standard deviations). The initial value of $\hat{\mathbf{W}}$ used in methods 1 and 4 above is the identity matrix. In these methods and also in method 3, the matrix $\mathbf{G}^* \mathbf{S}_\omega^{-*} \mathbf{S}_\omega^{-1} \mathbf{G}$ appearing in the IQML algorithm is initialized using estimates obtained with ESPRIT. We have noticed that the alternative of initializing this matrix to the identity results in an increased RMSE. In all cases, LS-ESPRIT is used since the more complicated TLS-ESPRIT does not provide an appreciable performance improvement. The displacement between the two data structures in ESPRIT is $\delta = 2$. IQML is implemented with the quadratic constraint $\|\mathbf{g}\| = 1$, since it seems to give better results than the linear constraint $\mathrm{Re}\{g_0\} = 1$. The IQML iterations are terminated when either of the following is satisfied:

- $\|\mathbf{g}_k - \mathbf{g}_{k-1}\| < \epsilon = 10^{-4}$. (Any value of ϵ between 10^{-2} and 10^{-4} provides essentially the same performance.)

- the number of iterations > 50

1.7.2 Effect of the Number of Samples

The finite-sample and asymptotic performance of the four methods above is illustrated in Figure 1.2. As predicted by the theoretical study, the RMSE of the proposed cost functions (methods 1 and 4) tend to the CRB as the number of samples (equivalently, as the number of pulses) increases. Although both methods are asymptotically efficient, method 1 shows a lower RMSE than method 4 for a small number of samples (*e.g.*, the former attains the CRB for 4 or more pulses, whereas

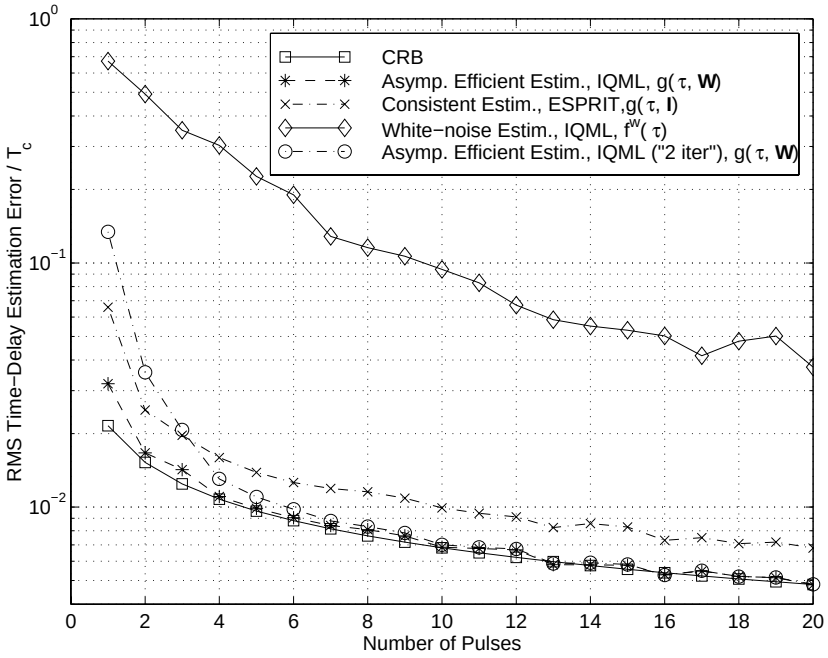

Figure 1.2. RMSE of the proposed estimators as a function of the number of training pulses. Parameters: $\theta_1 = 0^o$, $\theta_2 = 10^o$, $\theta_i = -30^o$, $\tau_1 = 0$, $\tau_2 = 0.4\,T_c$, $m =$ 6 antennas, $\mathrm{SNR}_1 = 16\,\mathrm{dB}$, $\mathrm{SIR}_1 = -3\,\mathrm{dB}$, $\mathrm{SNR}_1/\mathrm{SNR}_2 = 3\,\mathrm{dB}$.

the latter needs at least 7 pulses). Another advantage of method 1 over method 4 is that fewer iterations of IQML are required for convergence. For instance, when the number of received pulses is 3, the number of iterations required by method 1 is 5.59 ± 0.93 (mean $\pm$ standard deviation), while method 4 needs 7.66 ± 1.18 iterations. Though the number of iterations varies depending on the scenario, in the vast majority of the cases we have simulated it is less than 15 and the difference between method 1 and 4 is approximately constant. Thus, the computational time required for the optimization is reduced by several orders of magnitude thanks to the use of IQML and ESPRIT instead of minimizing (1.3.18) or (1.4.11) via a search.

Since method 1 has some advantages over method 4, the former is used for the rest of the simulations. The small difference between its RMSE and the CRB visible in Figures 1.3-1.6 occurs because the method has not achieved its asymptotic behaviour for signals formed by only 3 pulses. Also as predicted by the theoretical study, method 2 does not attain the CRB, but performs much better than the estimator designed for the white-noise case (method 3). This occurs because method 2 (cost function $g_N(\boldsymbol{\tau}, \mathbf{I})$) takes into account the spatial correlation of the noise field, though not in an optimal way. It is worth remarking that, because of its simplicity

(only ESPRIT is applied) and its low RMSE, method 2 is an excellent initialization scheme for the asymptotically efficient estimators based on IQML. The RMSE obtained with the white-noise estimator is plotted for comparison purposes, and the severe degradation that it undergoes when the noise field is spatially correlated is evident in all the simulations.

1.7.3 Effect of the Number of Sensors

Results showing the effect of varying the number of sensors are given in Figure 1.3. In all cases, method 1 outperforms method 2, and both are clearly superior to the white-noise estimator. For the number of sensors shown in the figure, the CRB decreases slightly faster than $1/m$. Now, however, the RMSE of the estimates obtained with $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ do not approach the CRB as the number of sensors increases. This behaviour coincides with two well-known results in sensor array processing: As the data grows in a dimension different from the dimension in which the parameters are estimated, *i)* the deterministic ML estimator is not asymptotically efficient [47], [46], and *ii)* the IQML algorithm is inconsistent [53] and its RMSE does not necessarily decrease [54].

Another effect is that the number of iterations needed by IQML to converge increases with the number of sensors. Although beyond the scope of this paper, it is worth noting that the performance of the methods for a large number of antennas can be improved using a modification of the signal matrix as in the MODE algorithm [15]. That is, the term $\mathbf{Z} = \mathbf{Y}^* \hat{\mathbf{R}}_{yy}^{-1/2} \hat{\mathbf{W}} \hat{\mathbf{R}}_{yy}^{-1/2} \mathbf{Y}$ appearing in $g(\boldsymbol{\tau}, \hat{\mathbf{W}})$ can be replaced by $\hat{\mathbf{E}} \tilde{\boldsymbol{\Lambda}} \hat{\mathbf{E}}^*$, where $\hat{\mathbf{E}}$ represents the eigenvectors of $\mathbf{Z}$ associated with the d largest eigenvalues, and $\tilde{\boldsymbol{\Lambda}}$ is a certain diagonal weighting matrix.

1.7.4 Effect of the SIR

The objective of this simulation, whose results are shown in Figure 1.4, is to show that the RMSE of the consistent and asymptotically efficient estimators proposed herein are robust against arbitrarily strong interferers. Therefore, they are valid approaches for time-delay estimation in interference-limited situations, such as most mobile communication systems. Note that the estimator designed for a white-noise scenario completely fails for SIR < -10 dB.

1.7.5 Closely Spaced Signals

In Figures 1.5 and 1.6, we investigate the ability of the different methods to resolve closely-spaced signals in the temporal and spatial domains. As with all time-delay estimators that do not assume parameterized spatial signatures, the CRB grows without limit as the relative delay of the signals decreases. We observe that the estimator we have proposed is always very close to the CRB, except for the case of relative delays smaller than $0.1T_c$. However, this range of delays lacks practical interest because reliable delay estimates cannot be expected for any algorithm; *i.e.*,

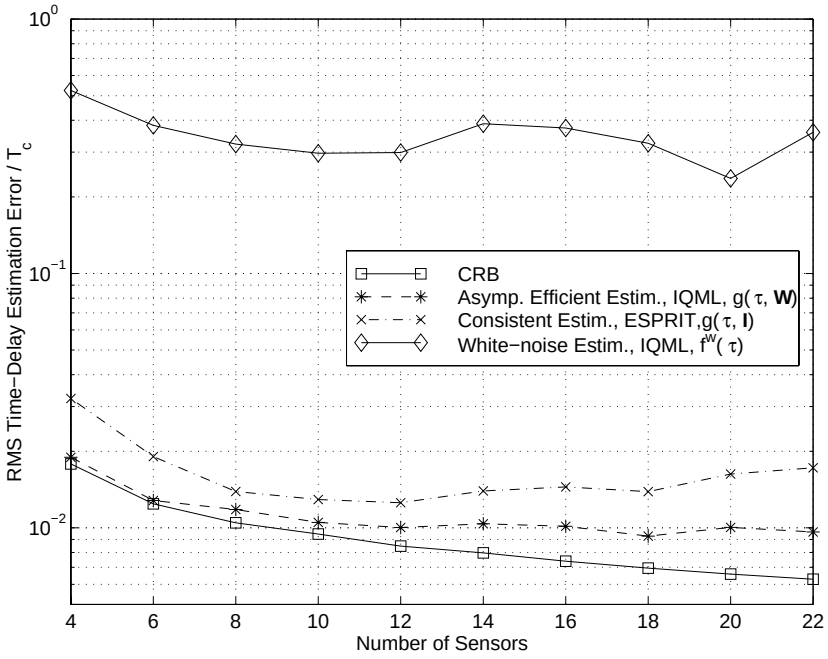

Figure 1.3. RMSE of the proposed estimators as a function of the number of sensors. Parameters: $\theta_1 = 0^{\rm o}$, $\theta_2 = 10^{\rm o}$, $\theta_i = -30^{\rm o}$, $\tau_1 = 0$, $\tau_2 = 0.4\,T_c$, $K = 3$ pulses, $\mathrm{SNR}_1 = 16\,\mathrm{dB}$, $\mathrm{SIR}_1 = -3\,\mathrm{dB}$, $\mathrm{SNR}_1/\mathrm{SNR}_2 = 3\,\mathrm{dB}$.

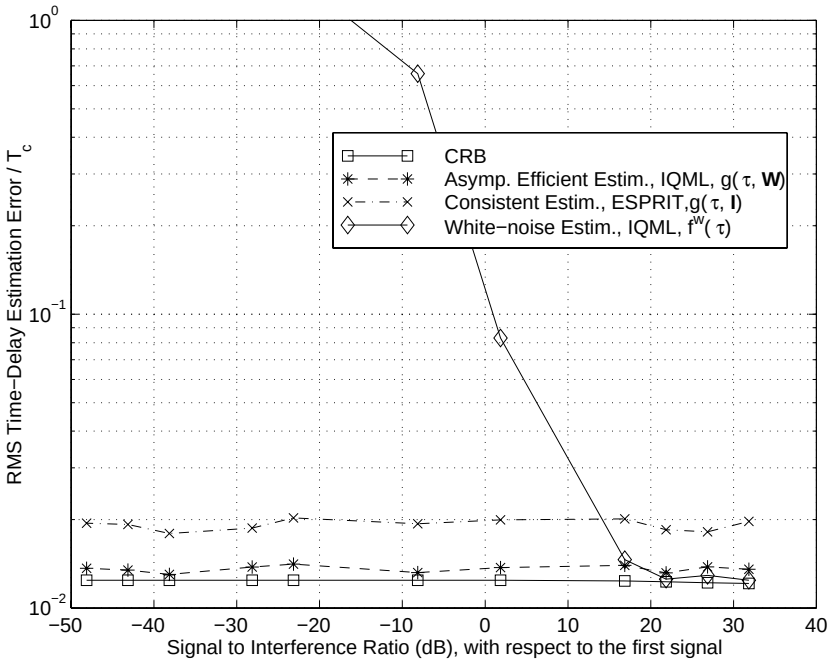

Figure 1.4. RMSE of the proposed estimators as a function of the interference power. Parameters: $\theta_1 = 0^{\circ}$, $\theta_2 = 10^{\circ}$, $\theta_i = -30^{\circ}$, $\tau_1 = 0$, $\tau_2 = 0.4\,T_c$, $m = 6$ antennas, $K = 3$ pulses, $\mathrm{SNR}_1 = 16\,\mathrm{dB}$, $\mathrm{SNR}_1/\mathrm{SNR}_2 = 3\,\mathrm{dB}$.

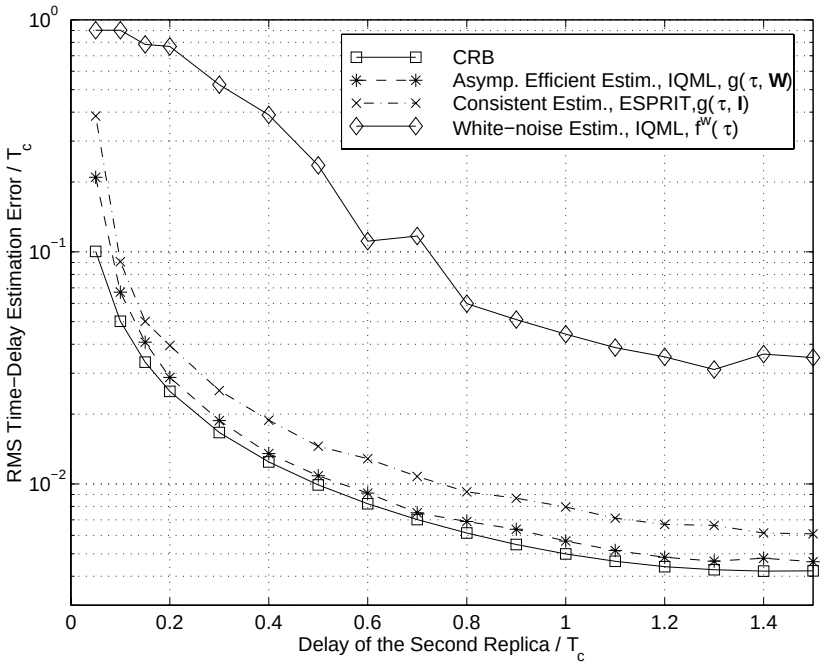

Figure 1.5. RMSE of the proposed estimators as a function of time-delay separation. Parameters: $\theta_1 = 0^o$, $\theta_2 = 10^o$, $\theta_i = -30^o$, $\tau_1 = 0$, $m = 6$ antennas, $K = 3$ pulses, $\mathrm{SNR}_1 = 16\,\mathrm{dB}$, $\mathrm{SIR}_1 = -3\,\mathrm{dB}$, $\mathrm{SNR}_1/\mathrm{SNR}_2 = 3\,\mathrm{dB}$.

in this range the best achievable standard deviation is larger than half the time-delay separation between the two signals [38].

When the DOA separation of the signals is smaller than the beamwidth of the sensor array, the CRB increases as the DOA separation decreases; but it does not tend to infinity as in the case of delay separation. Also when the DOA separation is smaller than the beamwidth, ESPRIT undergoes a severe degradation because the matrix $\mathbf{A}$ tends to be rank deficient. However, the performance of the proposed method is always very close to the CRB, even though it is initialized with ESPRIT.

1.7.6 Performance Using a Search

We have also analyzed the RMSEs obtained when the criteria $V_N(\boldsymbol{\tau})$ in (1.3.18), $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ in (1.4.11), $g_N(\boldsymbol{\tau}, \mathbf{I})$ in (1.4.13), and $f_N^w(\boldsymbol{\tau})$ in (1.3.29) are minimized using a search. We have observed that direct minimization of $f_N(\boldsymbol{\tau})$ and $g_N(\boldsymbol{\tau}, \hat{\mathbf{W}})$ and method 1 all yield nearly the same RMSE even for a small number of samples. Therefore, the new cost function we have proposed, besides being asymptotically efficient, does not entail any degradation in the finite-sample case with respect to

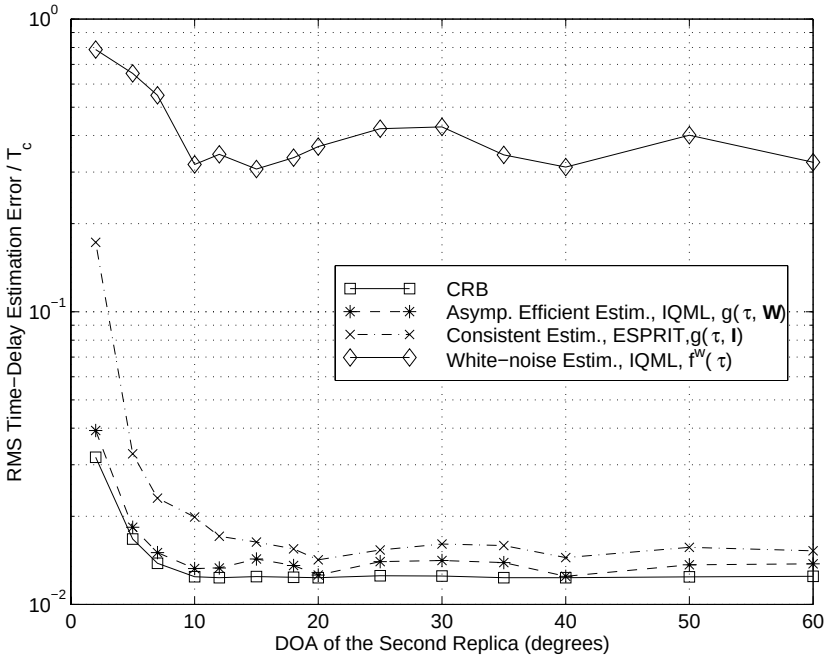

Figure 1.6. RMSE of the proposed estimators as a function of DOA separation. Parameters: $\theta_1 = 0^{\circ}$, $\theta_i = -30^{\circ}$, $\tau_1 = 0$, $\tau_2 = 0.4\,T_c$, $m = 6$ antennas, $K = 3$ pulses, $\mathrm{SNR}_1 = 16\,\mathrm{dB}$, $\mathrm{SIR}_1 = -3\,\mathrm{dB}$, $\mathrm{SNR}_1/\mathrm{SNR}_2 = 3\,\mathrm{dB}$.

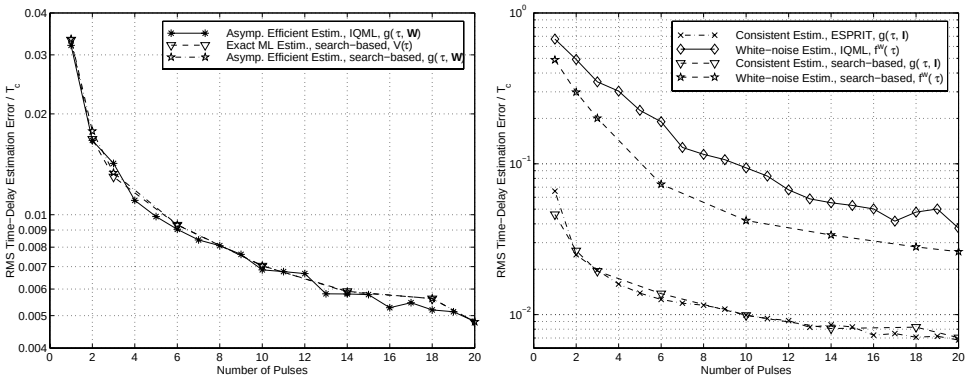

Figure 1.7. Comparison with the search-based estimators. Parameters: $\theta_1 = 0^o, \theta_2 = 10^o, \theta_i = -30^o, \tau_1 = 0, \tau_2 = 0.4\,T_c$, $m = 6$ antennas, $\mathrm{SNR}_1 = 16\,\mathrm{dB}$, $\mathrm{SIR}_1 = -3\,\mathrm{dB}$, $\mathrm{SNR}_1/\mathrm{SNR}_2 = 3\,\mathrm{dB}$.

the exact ML estimator (1.3.18). Moreover, the minimization using IQML does not introduce any impairment with respect to the use of a search. Secondly, the minimization of $g_N(\boldsymbol{\tau}, \mathbf{I})$ by means of a search or using ESPRIT (*i.e.*, method 2) results in approximately the same RMSE. On the contrary, the RMSE obtained from the minimization of $f_N^w(\boldsymbol{\tau})$ using a search is slightly smaller than that obtained using IQML (method 3). Numerical results supporting these claims are provided in Figure 1.7.

1.8 Conclusions

Synchronization and time delay estimation are important components of many signal processing and communications systems. This paper has focused on how multiple receive antennas can be used to improve the accuracy of such systems. An antenna array is particularly useful in situations where multipath and co-channel interference are limiting factors on performance. The spatial selectivity of the array provides an additional dimension in which to differentiate the desired user's signal from the noise and interference. Under the assumption that the noise and CCI are spatially colored but temporally white Gaussian processes, the maximum likelihood solution to the general time delay estimation problem was derived. The resulting concentrated criterion for the delays is highly non-linear, and not conducive to simple minimization procedures. Using various techniques, it was shown how the optimal ML criterion could be approximated by a simpler cost function that was shown to provide asymptotically equivalent (and hence statistically efficient) delay estimates. The form of the new criterion lends itself to minimization by the IQML algorithm, an iterative approach that avoids the need for gradient-based or exhaustive searches. The existence of simple yet accurate initialization schemes based

on ESPRIT or identity weightings makes the approach viable for practical implementation. A number of simulation studies were presented that demonstrate the performance advantage of the proposed technique over competing delay estimators.

.1 Appendix

In this appendix we compute the asymptotic order of the matrix $\mathbf{B}_N^i(\boldsymbol{\tau}_0)$. Using (1.2.6), (1.3.4)-(1.3.6), (1.3.19), and the expression for the derivative of a projection matrix [39]

$$\mathbf{P}_{\mathbf{S}^*}^i = \mathbf{P}_{\mathbf{S}^*}^{\perp} \left(\mathbf{S}^*\right)^i \left(\mathbf{S}^*\right)^{\dagger} + \left(\cdots\right)^* \tag{.1.1}$$

where $(\cdot)^{\dagger}$ denotes the pseudoinverse, we get

$$\mathbf{B}_N^i(\boldsymbol{\tau}_0) = \frac{1}{N}\hat{\mathbf{R}}_{yy}^{-1/2}\,\mathbf{E}\,\mathbf{P}_{\mathbf{S}^*}^{\perp}\left(\mathbf{S}^*\right)^i\mathbf{A}^*\hat{\mathbf{R}}_{yy}^{-1/2} + \left(\cdots\right)^* \tag{.1.2}$$

$$+\frac{1}{N}\hat{\mathbf{R}}_{yy}^{-1/2}\,\mathbf{E}\,\mathbf{P}_{\mathbf{S}^*}^{\perp}\left(\mathbf{S}^*\right)^i\left(\mathbf{S}^*\right)^{\dagger}\mathbf{E}^*\hat{\mathbf{R}}_{yy}^{-1/2} + \left(\cdots\right)^* \tag{.1.3}$$

where all functions of $\boldsymbol{\tau}$ are evaluated at $\boldsymbol{\tau}_0$, the relation $\mathbf{S}\mathbf{P}_{\mathbf{S}^*}^{\perp} = \mathbf{0}$ is used, and the notation $\left(\cdots\right)^*$ means that the same expression appears again transposed and conjugated. In order to obtain an asymptotic expression, the matrix $\hat{\mathbf{R}}_{yy}$ is replaced by $\mathbf{R}_{yy}$ and the terms in (.1.3) are neglected. Thus, we can write

$$\mathbf{B}_N^i(\boldsymbol{\tau}_0) \simeq \frac{1}{N}\mathbf{R}_{yy}^{-1/2}\,\mathbf{E}\,\mathbf{P}_{\mathbf{S}^*}^{\perp}\left(\mathbf{S}^*\right)^i\mathbf{A}^*\mathbf{R}_{yy}^{-1/2} + \left(\cdots\right)^* \ . \tag{.1.4}$$

The correlation between the k,lth element of $\mathbf{B}_N^i(\boldsymbol{\tau}_0)$ and the r,sth element of $\mathbf{B}_N^j(\boldsymbol{\tau}_0)$ for any value $k,l,r,s = 1,\cdots,m$ and $i,j = 1,\cdots,d$ is

$$\Psi_{\mathbf{B}'}\left(k,l,r,s;i,j\right) \triangleq \mathcal{E}\left\{\left[\mathbf{B}_N^i(\boldsymbol{\tau}_0)\right]_{k,l}\left[\mathbf{B}_N^j(\boldsymbol{\tau}_0)\right]_{r,s}^*\right\} = \tag{.1.5}$$

$$= \frac{1}{N^2}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,s}^*\mathbf{A}\,\mathbf{S}^j\,\mathbf{P}_{\mathbf{S}^*}^{\perp}\mathcal{E}\left\{\mathbf{E}^*\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,r}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,k}^*\mathbf{E}\right\}\mathbf{P}_{\mathbf{S}^*}^{\perp}\left(\mathbf{S}^*\right)^i\mathbf{A}^*\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,l}$$

$$+\ \frac{1}{N^2}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,k}^*\mathbf{A}\,\mathbf{S}^i\,\mathbf{P}_{\mathbf{S}^*}^{\perp}\mathcal{E}\left\{\mathbf{E}^*\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,l}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,s}^*\mathbf{E}\right\}\mathbf{P}_{\mathbf{S}^*}^{\perp}\left(\mathbf{S}^*\right)^j\mathbf{A}^*\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,r}$$

where $\left[\cdot\right]_{:,k}$ represents the kth column of a given matrix, and $\mathcal{E}\left\{\cdot\right\}$ is the expectation operator. For the noise model that we have considered,

$$\mathcal{E}\left\{\mathbf{E}^*\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,r}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,k}^*\mathbf{E}\right\} = \mathbf{I}\left(\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,k}^*\mathbf{Q}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{:,r}\right) \tag{.1.6}$$

$$= \mathbf{I}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\,\mathbf{Q}\,\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{k,r} \ . \tag{.1.7}$$

This result and the fact that each row of $\mathbf{S}$ only depends on one variable yield

$$\Psi_{\mathbf{B}'}\left(k,l,r,s;i,j\right) =$$

$$\frac{1}{N^2}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\,\mathbf{Q}\,\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{k,r}\left[\mathbf{R}_{yy}^{-\frac{1}{2}}\mathbf{A}\right]_{s,j}\left[\mathbf{A}^*\,\mathbf{R}_{yy}^{-\frac{1}{2}}\right]_{i,l}\left(\mathbf{h}^T\left(\tau_j\right)\mathbf{P}_{\mathbf{S}^*}^{\perp}\mathbf{h}^c\left(\tau_i\right)\right) + \cdots \tag{.1.8}$$

where the second term has been omitted because it can be readily obtained from the first one and (.1.5) by an appropriate change of indices. Since we have assumed that $s(t)$ is a finite-average-power signal and is sampled above the Nyquist rate, the order of $\mathbf{h}^T(\tau_j)\,\mathbf{P}^{\perp}_{\mathbf{S}^*}\mathbf{h}^c(\tau_i)$ is $O(N)$. Therefore, the correlation (.1.5) of the elements of $\mathbf{B}_N(\boldsymbol{\tau}_0)^i$ is $O\left(N^{-1}\right)$, which completes the proof that $\mathbf{B}^i_N(\boldsymbol{\tau}_0) = O_p\left(N^{-1/2}\right)$.

BIBLIOGRAPHY

[1] L. Brennan and I. Reed, "An Adaptive Signal Processing Algorithm for Communications", *IEEE Trans. Aerosp. Electron. Syst.*, **AES-18**(1):124–130, Jan. 1982.

[2] J. Leary and R. Gooch, "Adaptive Beamforming for TDMA Signals", In *Proc. 29th Asilomar Conf. on Signals, Systems, and Computers*, pages 1378–1382, 1995.

[3] A. Keerthi and J. Shynk, "Separation of Cochannel Signals in TDMA Mobile Radio", *IEEE Trans. Sig. Proc.*, **46**(10):2684–2697, Oct. 1998.

[4] A. Kuzminskiy and D. Hatzinakos, "Semi-Blind Symbol Synchronization for Antenna Array Signal Processing in Short Burst Multiuser SDMA Systems", In *Proc. SPAWC Workshop*, pages 142–145, Annapolis, MD, 1999.

[5] A.-J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint Angle and Delay Estimation Using Shift-Invariance Techniques", *IEEE Trans. SP*, **46**(2):405–418, Feb. 1998.

[6] M. Chenu-Tournier, A. Ferreol, and P. Larzabal, "Low Complexity Blind Space-Time Identification of Propagation Parameters", In *Proc. ICASSP*, volume V, pages 2873–2876, Phoenix, AZ, 1999.

[7] R. Roy and T. Kailath, "ESPRIT – Estimation of Signal Parameters via Rotational Invariance Techniques", *IEEE Trans. on ASSP*, **37**(7):984–995, July 1989.

[8] M. Haardt, C. Brunner, and J. Nossek, "Efficient High-Resolution 3-D Channel Sounding", In *Proc. IEEE Vehicular Technology Conf.*, Ottawa, Canada, 1998.

[9] M. Cedervall and A. Paulraj, "Joint Channel and Space-Time Parameter Estimation", In *Proc. 30th Asilomar Conf. on Signals, Systems, and Computers*, pages 375–379, 1996.

[10] M. Wax and A. Leshem, "Joint Estimation of Time Delays and Directions of Arrival of Multiple Reflections of a Known Signal", *IEEE Trans. SP*, **45**(10):2477–2484, Oct. 1997.

[11] N. Bertaux, P. Larzabal, C. Adnet, and E. Chaumette, "A Parameterized Maximum Likelihood Method for Multipaths Channels Estimation", In *Proc. SPAWC Workshop*, pages 391–394, Annapolis, MD, 1999.

[12] P. Pelin, *Space-Time Algorithms for Mobile Communications*, PhD thesis, Chalmers University of Technology, 1999.

[13] A. Swindlehurst, "Time Delay and Spatial Signature Estimation Using Known Asynchronous Signals", *IEEE Trans. SP*, **46**(2):449–462, Feb. 1998.

[14] Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise", *IEEE Trans. on ASSP*, **34**(5):1081–1089, October 1986.

[15] P. Stoica and K. Sharman, "Maximum Likelihood Methods for Direction-of-Arrival Estimation", *IEEE Trans. on ASSP*, **38**(7):1132–1143, July 1990.

[16] P. Stoica and A. Nehorai, "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation", *IEEE Trans. on ASSP*, **38**(10):1783–1795, October 1990.

[17] A. Swindlehurst and J. Gunther, "Methods for Blind Equalization and Resolution of Overlapping Echoes of Unknown Shape", *IEEE Trans. SP*, **47**(5):1245–1254, May 1999.

[18] P. Pelin, "Iterative Least Squares Techniques for Self-Synchronization and Equalization in Adaptive Antenna Systems", *IEEE Trans. on Vehic. Tech.*, 1998 (submitted).

[19] D. Astély, A. Jakobsson, and A. Swindlehurst, "Burst Synchronization on Unknown Frequency Selective Channels with Co-Channel Interference Using an Antenna Array", In *Proc. IEEE Vehicular Technology Conf.*, Houston, TX, March 1999.

[20] D. Dlugos and R. Scholtz, "Acquisition of Spread Spectrum Signals by an Adaptive Array", *IEEE Trans. on ASSP*, **37**(8):1253–1270, August 1989.

[21] Z.-S. Liu, J. Li, and S.L. Miller, "A Receiver Diversity Based Code-Timing Estimator for Asynchronous DS-CDMA Systems", In *Proc. ICASSP*, volume VI, pages 3245–3248, Seattle, WA, 1998.

[22] A. Jakobsson, A. Swindlehurst, D. Asztély, and C. Tidestav, "A Blind Frequency Domain Method for DS-CDMA Synchronization Using Antenna Arrays", In *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, 1998.

[23] Y.-F. Chen and M. Zoltowski, "Joint Angle and Delay Estimation for DS-CDMA with Application to Reduced Dimension Space-Time RAKE Receivers", In *Proc. ICASSP*, volume V, pages 2933–2936, Phoenix, AZ, 1999.

[24] G. Seco, J. Fernández-Rubio, and A. Swindlehurst, "Code-Timing Synchronization in DS-CDMA Systems Using Space-Time Diversity", In *Proc. 5th Bayona Workshop on Emerging Technologies in Telecommunications – COST 254*, Bayona, Spain, Sept. 1999.

[25] G. Seco and J. Fernández-Rubio, "Maximum Likelihood Propagation-Delay Estimation in Unknown Correlated Noise Using Antenna Arrays", In *Proc. ICASSP*, volume IV, pages 2065–2068, Seattle, WA, 1998.

[26] G. Seco and J. Fernández-Rubio, "Maximum Likelihood Time-of-Arrival Estimation Using Antenna Arrays: Application to Global Navigation Satellite Systems", In *Proc. EUSIPCO*, Rhodes, Greece, Sept. 1998.

[27] A. Jakobbson, A. Swindlehurst, and P. Stoica, "Subspace-Based Estimation of Time Delays and Doppler Shifts", *IEEE Trans. on Sig. Proc.*, **46**(9):2472–2483, Sept 1998.

[28] A. Dogandzic and A. Nehorai, "Estimating Range, Velocity, and Direction with a Radar Array", In *Proc. ICASSP*, volume V, pages 2773–2776, Phoenix, AZ, 1999.

[29] G. Xu and H. Liu, "An Effective Transmission Beamforming Scheme for Frequency Division Duplex Digital Wireless Communication Systems", In *Proc. ICASSP*, pages 1729–1732, Detroit, MI, 1995.

[30] K. Molnar and G. Bottomley, "Adaptive Array Processing MLSE Receivers for TDMA Digital Cellular/PCS Communications", *IEEE J. Select. Areas Commun.*, **16**:1340–1351, Oct. 1998.

[31] L. Tong, G. Xu, and T. Kailath, "Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach", *IEEE Trans. Info. Theory*, **40**(2):340–349, March 1994.

[32] D. Slock, "Blind Fractionally Spaced Equalization, Perfect Reconstruction Filter Banks, and Multichannel Linear Prediction", In *Proc. ICASSP*, pages IV–585 – IV–588, Adelaide, Australia, 1994.

[33] E. Moulines, P. Duhamel, J.F. Cardoso, and S. Mayrargue, "Subspace Methods for the Blind Identification of Multichannel FIR Filters", *IEEE Trans. Sig. Proc.*, **43**(2):516–525, Feb. 1995.

[34] J. K. Tugnait, "On Blind Equalization of Multipath Channels Using Fractional Sampling and Second-Order Cyclostationary Statistics", *IEEE Trans. Info. Theory*, **41**(1):308–311, Jan. 1995.

[35] Z. Ding, "Matrix Outer-Product Decomposition Method for Blind Multiple Channel Equalization", *IEEE Trans. on Sig. Proc.*, **45**(12):3053–3061, Dec. 1997.

[36] J. Rissanen, "Modeling by Shortest Data Description", *Automatica*, **14**:465–471, 1978.

[37] H. Akaike, "A New Look at Statistical Model Identification", *IEEE Trans. on Automatic Control*, **19**:716–723, 1974.

[38] M. Viberg, B. Ottersten, and T. Kailath, "Detection and Estimation in Sensor Arrays Using Weighted Subspace Fitting", *IEEE Trans. on Sig. Proc.*, **39**(11):2436–2449, Nov. 1991.

[39] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing", In *Radar Array Processing*, Haykin, Litva, and Shepherd, editors, pages 99–151. Springer-Verlag, Berlin, 1993.

[40] P. Pelin, "Decoupled Direction Finding: Detection", In *Proc. ICASSP*, Seattle, WA, 1998.

[41] M. Viberg and I. Bogdan, "Using the Bootstrap for Robust Detection in Array Signal Processing", In *Proc. 33rd Asilomar Conf. on Signals, Systems, and Computers*, 1999.

[42] N. R. Goodman, "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)", *Ann. Math. Statist.*, **34**:152–176, 1963.

[43] A. Graham, *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood Limited, Chichester, England, 1981.

[44] M. Viberg, B. Ottersten, and A. Nehorai, "Performance Analysis of Direction Finding with Large Arrays and Finite Data", *IEEE Trans. on Sig. Proc.*, **43**(2):469–477, Feb. 1995.

[45] W. J. Bangs, *Array Processing with Generalized Beamformers*, PhD thesis, Yale University, New Haven, CT, 1971.

[46] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramér-Rao Bound", *IEEE Trans. on ASSP*, **37**(5):720–741, May 1989.

[47] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramér-Rao Bound: Further Results and Comparisons", *IEEE Trans. on ASSP*, **38**(12):2140–2150, December 1990.

[48] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally Efficient Angle Estimation for Signals with Known Waveforms", *IEEE Trans. on Sig. Proc.*, **43**(9):2154–2163, Sept. 1995.

[49] M. Viberg, P. Stoica, and B. Ottersten, "Maximum Likelihood Array Processing in Spatially Colored Noise Fields Using Parameterized Signals", *IEEE Trans. on Sig. Proc.*, **45**(4):996–1004, April 1997.

[50] D. Zheng, J. Li, S.L. Miller, and E.G. Ström, "An Efficient Code-Timing Estimator for DS-CDMA Signals", *IEEE Trans. SP*, **45**(1):82–89, Jan. 1997.

[51] Y. Hua, "The Most Efficient Implementation of the IQML Algorithm", *IEEE Trans. Sig. Proc.*, **43**(8):2203–2204, August 1994.

[52] G. H. Golub and C. F. Van Loan, *Matrix Computations (3rd edition)*, The John Hopkins University Press, 1996.

[53] P. Stoica, J. Li, and T. Sderstrm, "On the Inconsistency of IQML", *Signal Processing*, **56**:185–190, Jan. 1997.

[54] J. Li, P. Stoica, and Z. Liu, "Comparative Study of IQML and MODE Direction-of-Arrival Estimators", *IEEE Trans. SP*, **46**(1):149–160, Jan. 1998.