OFDM Channel Estimation in the Presence of Interference

Aleksandar Jeremić, Timothy A. Thomas, *Member, IEEE*, and Arye Nehorai, *Fellow, IEEE*

***Abstract*—We develop a frequency-domain channel estimation algorithm for single-user multiantenna orthogonal frequency division multiplexing (OFDM) wireless systems in the presence of synchronous interference. In this case, the synchronous interferer's signal on each OFDM subcarrier is correlated in space with a rank one spatial covariance matrix. In addition, the interferer's spatial covariance matrix is correlated in frequency based on the delay spread of the interferer's channel. To reduce the number of unknown parameters we develop a structured covariance model that accounts for the structure resulting from the synchronous interference. To further reduce the number of unknown parameters, we model the covariance matrix using *a priori* known set of frequency-dependent functions of joint (global) parameters. We estimate the interference covariance parameters using a residual method of moments (RMM) estimator and the channel parameters by maximum likelihood (ML) estimation. Since our RMM estimates are invariant to the mean, this approach yields simple noniterative estimates of the covariance parameters while having optimal statistical efficiency. Hence, our algorithm outperforms existing channel estimators that do not account for the interference, and at the same time requires smaller number of pilots than the MANOVA method and thus has smaller overhead. Numerical results illustrate the applicability of the proposed algorithm.**

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM) has recently received considerable interest for its advantages in high-bit-rate transmissions over frequency selective fading channels [1]. In OFDM systems, the input high-rate data stream is divided into many low-rate streams that are transmitted in parallel [2], thereby increasing the symbol duration and reducing the intersymbol interference (ISI). These features have motivated the adoption of OFDM as a standard for digital audio broadcasting [3], digital video broadcasting [4], and broadband local area networks [5].

Coherent OFDM transmission invariably requires an estimation of the channel frequency response (i.e., the gains of the OFDM tones); see [6]. Currently, there are two different types of channel parameter estimators: i) blind and ii) pilot-aided. Blind channel estimation techniques (see [7] and [8]) try to estimate the channel without any knowledge of the transmitted data. They are attractive because of the possible savings in training overhead, however they are effective only when a large amount of data can be collected (so that stochastic estimation can be made reliably). This is clearly a disadvantage in the case of mobile wireless systems because of the time-varying nature of the channel. Pilot-aided channel estimation is the other approach in which training sequence consisting of known data symbols (pilots) is transmitted at the beginning of a session (or multiplexed into the user data stream at a later stage) and the initial estimation of the channel parameters is performed using the received pilot signal. It has been shown that the pilot-aided channel estimation is the optimum way to estimate the channel when signal-to-noise ratio (SNR) is sufficiently high [9].

In this paper (see also [10]), we present a frequency domain channel estimation for a single-user and multiple receiving antennas system in the presence of synchronous interference using a pilot-aided algorithm. Interference suppression is of utmost importance in high-rate (high capacity) wireless systems. In the presence of interference, the received signal, in addition to multipath fading, may have frequency-dependent covariance, which means that channel estimation algorithms that treat the interference as white noise [11] are suboptimal. Sufficiently accurate estimation of the covariance matrices, which is necessary for efficient interference suppression in such a model, would impose significant training overhead due to the required large number of pilot symbols.

To reduce the number of unknown parameters, we first observe that in the synchronous case the interference can be modeled as a convolution of the interferer's channel and data sequence. As a result, the interference covariance matrix will have one-dimensional kernel. Using this *a priori* structural information, we develop a structured model with fewer parameters. To further reduce the number of unknown parameters we propose to model the frequency dependence of the covariance matrices using *a priori* known set of frequency dependent functions of joint (global) parameters, which results in a simplified structure of the covariance matrix. In this way the number of required pilot symbols will be dependent on the number of functions needed to achieve sufficiently accurate approximation and not on the number of subcarriers. We also assume that the channel vector is deterministic and completely unknown. In this case the maximum likelihood estimator (MLE) is the best possible estimator as long as the channel impulse response is viewed as a deterministic unknown vector and the estimator is unbiased. Therefore, we propose estimating the unknown parame-

Manuscript received November 21, 2002; revised October 17, 2003. This work was supported by the Air Force Office of Scientific Research under Grants F49620-00-1-0083 and F49620-02-1-0339, the National Science Foundation under Grant CCR-0105334, and the Office of Naval Research under Grant N00014-01-1-0681. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Behrouz Farhang-Boroujeny.

A. Jeremić was with the Department of Electrical and Computer Engineering, The University of Illinois at Chicago, Chicago, IL 60607-7053 USA. He is now with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, L8S 4K1 Canada (e-mail: jeremic@ece.mcmaster.ca)

T. A. Thomas is with the Motorola Labs, Schaumburg, IL 60196 USA (e-mail: T.Thomas@motorola.com).

A. Nehorai is with the Department of Electrical and Computer Engineering, The University of Illinois at Chicago, Chicago, IL 60607-7053 USA (e-mail: nehorai@ece.uic.edu).

Digital Object Identifier 10.1109/TSP.2004.837440

ters using two estimators: i) MLE and ii) a simplified, asymptotic MLE in which the nonlinear covariance parameters are estimated using a method-of-moments (MM) estimator [12].

In Section II, we briefly describe an OFDM system with pilot-symbol-aided channel estimation. In Section III, we present the statistical model. In Section IV, we develop the MLE and discuss numerical implementations of the proposed nonlinear algorithms. Section V demonstrates the applicability of our results through numerical examples. In Section VI, we discuss possible extensions to multiple-input-multiple-output (MIMO) OFDM systems. Concluding remarks are given in Section VII.

II. OFDM Model

In this section we briefly introduce the channel model for an OFDM system with pilot-symbol-aided channel estimator. Our goal is to develop a model that will include unknown random effects due to a presence of unknown interference and enable statistically efficient channel estimation using sufficiently small number of pilots.

A. Channel Model

Consider an OFDM system [2] that consists of $n_{\rm s}$ subcarriers of which $n_{\rm u}+1$ subcarriers at the central spectrum are used for transmission and the other subcarriers at both edges form the guard bands. Each transmission subcarrier is modulated by a data symbol x_{jk}, where j represents the subcarrier index number and k the time slot number (OFDM symbol number). OFDM transmitters usually employ an inverse fast Fourier transform (IFFT) of size $n_{\rm s}$ for the modulation. In order to limit the transmit signal to a bandwidth smaller than $1/T$, where T is the sampling time interval of the OFDM signal, the subcarriers in the guard band are not used. A guard interval is also added for every OFDM symbol to avoid ISI caused by multipath fading channels. As a result, the output complex baseband representation of the transmitted signal is given by

$$\begin{aligned} s(t) &= \sum_{k=-\infty}^{\infty} \sum_{j=-n_{\rm u}/2}^{n_{\rm u}/2} x_{jk}\phi_j(t-k(T+T_{\rm g})) \\ \phi_j(t) &= \begin{cases} \frac{1}{\sqrt{T-T_{\rm g}}} e^{\iota 2\pi(B/n_{\rm s})j(t-T_{\rm g})}, & t\in[0, T-T_{\rm g}] \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{2.1}$$

where ι is the imaginary unit, B the total bandwidth (corresponding to $n_{\rm s}$ subcarriers), and $T_{\rm g}$ the guard interval. Since the guard interval contains a repetition of a preceding part of the signal only, it is not of interest in this discussion, and therefore, we assume without loss of generality that $T_{\rm g}=0$.

The signal $s(t)$ is then transmitted through a multipath wireless channel characterized by

$$\boldsymbol{h}(t,\tau) = \sum_{i=1}^{n_{\rm mu}} \boldsymbol{\alpha}_i(t) c(\tau-\tau_i) \tag{2.2}$$

where $\boldsymbol{h}(t,\tau)$ is an $n_{\rm r}$-dimensional user channel response vector, $n_{\rm r}$ the number of antennas on the receiver side, $n_{\rm mu}$ the number of user multipaths, $\boldsymbol{\alpha}_i(t)$, τ_i the complex amplitude and delay of the ith path, respectively, and $c(t)$ is a shaping pulse. To simplify the derivations we will assume $c(t)=\delta(t)$, where

$$\delta(t) = \begin{cases} 1, & t=0 \\ 0, & \text{elsewhere.} \end{cases}$$

We will show later that the only difference between $\boldsymbol{h}(t,\tau)$ for an arbitrary shaping pulse and $\boldsymbol{h}(t,\tau)$ for $c(t)=\delta(t)$ is in a scaling factor that does not affect our algortihm. In the remainder of the paper, unless otherwise noted, we will assume that the channel is deterministic and time independent, i.e., $\boldsymbol{\alpha}(t)\equiv\boldsymbol{\alpha}_i$, which yields

$$\boldsymbol{h}(t,\tau) \equiv \boldsymbol{h}(\tau) = \sum_{i=1}^{n_{\rm mu}} \boldsymbol{\alpha}_i c(\tau-\tau_i). \tag{2.3}$$

In Section V, we will briefly discuss possible extensions to time-varying channels using time interpolation based on appropriately selected basis functions.

For a time-independent channel, the received signal $\boldsymbol{r}(t)$ in the time domain is given by

$$\boldsymbol{r}(t) = s(t) * \boldsymbol{h}(t) + \boldsymbol{u}(t) + \boldsymbol{e}(t) \tag{2.4}$$

where "$*$" denotes the convolution, $\boldsymbol{u}(t)$ the interference, and $\boldsymbol{e}(t)$ is additive noise.

Using (2.4) and the discrete Fourier transform, the received signal in the frequency domain can be written as

$$\boldsymbol{y}_{jk} = \boldsymbol{h}_j x_{jk} + \boldsymbol{u}_{jk} + \boldsymbol{n}_{jk} \tag{2.5}$$

where $\boldsymbol{y}_{jk}$ is the $n_{\rm r}$-dimensional received signal vector at the jth subcarrier and kth time slot, $\boldsymbol{h}_j$ the $n_{\rm r}$-dimensional user channel vector at the jth subcarrier, $\boldsymbol{u}_{jk}$ the unknown interference vector at the jth subcarrier and kth time slot, and $\boldsymbol{n}_{jk}$ is additive noise. Observe that we use the subscript j to denote frequency dependence and k to denote time dependence. In the frequency domain, the user channel vector is the discrete Fourier transform of the user channel impulse response $\boldsymbol{h}(t)$. In particular

$$\boldsymbol{h}_j = \sum_{i=1}^{n_{\rm mu}} \boldsymbol{\alpha}_i e^{-2\iota\pi B(j-1-n_{\rm s}/2)\tau_i/n_{\rm s}}. \tag{2.6}$$

There are two types of interference with respect to the cyclic prefix position: synchronous, when the interferer's cyclic prefix is aligned in time with the user's cyclic prefix, and asynchronous, when these prefixes are not aligned. In the case of single synchronous interferer, $\boldsymbol{u}(t)$ can be represented as a convolution of the interferer's channel $\boldsymbol{g}(t)$ and the interferer's data symbol sequence $\psi(t)$, i.e., in the frequency domain

$$\boldsymbol{u}_{jk} = \boldsymbol{g}_j \psi_{jk} \tag{2.7}$$

where ψ_{jk} is an interferer data symbol at the jth subcarrier and kth time slot, and $\boldsymbol{g}_j$ is the $n_{\rm r}$-dimensional interferer channel vector at the jth subcarrier

$$\boldsymbol{g}_j = \sum_{i=1}^{n_{\rm mi}} \boldsymbol{\beta}_i e^{-j2\pi B(j-1-n_{\rm s}/2)\nu_i/n_{\rm s}} \tag{2.8}$$

where $\{\boldsymbol{\beta}_i\}$ and $\{\nu_i\}$ are the complex amplitude and delay of the ith path. Observe that the expression for the user (2.6) and synchronous interferer (2.8) channels are almost the same, i.e., they differ only in the number of multipaths and multipath delays and attenuations. If multiple synchronous interferers are present the interference can be represented as a summation of individual interferers, i.e., in the frequency domain

$$\boldsymbol{u}_{jk} = \sum_{q=1}^{n_i} \boldsymbol{g}_j^q \psi_{jk}^q \tag{2.9}$$

where n_i is the number of interferers, $\boldsymbol{g}_j^q$ the qth interferer channel at the jth subcarrier, and ψ_{jk}^q the qth interferer data symbol at the jth subcarrier and kth time slot. When the interferer is asynchronous such a representation is not possible. However, it may be possible to approximate the asynchronous interferer by a reduced rank model, not necessarily of rank one and obtain reasonable estimates using slightly modified algorithm [13].

III. Statistical Model

In order to derive the MLE for the above OFDM model, we first develop a statistical model of the received signal as a function of the channel and interference parameters. Our goal is to develop a model that will include the unknown random effect due to the presence of unknown interference and enable statistically efficient channel estimation using a small number of pilots.

We start by assuming that both the interference and ambient noise are zero-mean Gaussian wide-sense stationary (WSS) random processes. In addition, we assume that the ambient noise is uncorrelated in space, time, and frequency according to

$$\mathrm{E}[\boldsymbol{n}_{jk}\boldsymbol{n}_{j'k'}^H] = \sigma^2\delta(j-j')\delta(k-k')I_{n_\mathrm{r}} \tag{3.1}$$

where the superscript "^{H}" denotes the Hermitian transpose, I_n the $n \times n$ identity matrix, and the interference is uncorrelated in time according to

$$\mathrm{E}[\boldsymbol{u}_{jk}\boldsymbol{u}_{j'k'}^H] = \Sigma_{j,j'}\delta(k-k') \tag{3.2}$$

where $\Sigma_{j,j'}$ is the $n_\mathrm{r} \times n_\mathrm{r}$ interferer's spatial covariance matrix. In the remainder of the paper, we will omit the dimension subscripts of the identity matrices when they are obvious.

Then, the statistical properties of the received signal are completely determined by its mean

$$\boldsymbol{h}_j x_{jk} \tag{3.3}$$

and covariance of $\boldsymbol{\epsilon}_{jk} = \boldsymbol{u}_{jk} + \boldsymbol{n}_{jk}$

$$\mathrm{E}[\boldsymbol{\epsilon}_{jk}\boldsymbol{\epsilon}_{j'k'}{}^H] = \left(\Sigma_{j,j'} + \sigma^2\delta(j-j')I\right)\delta(k-k'). \tag{3.4}$$

In the single synchronous interferer case, the covariance can be further simplified to

$$\mathrm{E}[\boldsymbol{\epsilon}_{jk}\boldsymbol{\epsilon}_{j'k'}{}^H] = \boldsymbol{g}_j\boldsymbol{g}_{j'}^H \mathrm{E}[\psi_{jk}\psi_{j'k'}^*]\delta(k-k') + \sigma^2 I\delta(j-j')\delta(k-k') \tag{3.5}$$

where the superscript "*" denotes the complex conjugate. Furthermore, we model the interference data symbols ψ_{jk} as a zero-mean Gaussian random process uncorrelated in frequency and time with unit variance. Therefore, the covariance in this case reduces to

$$\mathrm{E}[\boldsymbol{\epsilon}_{jk}\boldsymbol{\epsilon}_{j'k'}{}^H] = \left(\boldsymbol{g}_j\boldsymbol{g}_{j'}^H + \sigma^2 I\right)\delta(j-j')\delta(k-k'). \tag{3.6}$$

Observe that the number of unknown covariance parameters in the above model is $n_\mathrm{s}n_\mathrm{r} + 1$ and that the number of measurements is $n_\mathrm{r}n_\mathrm{s}n_\mathrm{p}$, where n_p is the number of pilots (size of the training sequence). Recall the n_p is relatively small, usually from three to five.

When multiple synchronous interferers are present the model (3.6) becomes

$$\mathrm{E}[\boldsymbol{\epsilon}_{jk}\boldsymbol{\epsilon}_{j'k'}{}^H] = \left(\sum_{q=1}^{n_i} \boldsymbol{g}_j^q \boldsymbol{g}_{j'}^{q\,H} + \sigma^2 I\right)\delta(j-j')\delta(k-k'). \tag{3.7}$$

Note that the Gaussian model for the interferer's data symbols is just an approximation since these symbols come from a finite alphabet and thus typically have uniform distribution. However, the interferer signal constellation may not be known, and then, the Gaussian model is the most reasonable choice. In addition, it provides a model with tractable analysis. In the multiple interferer scenario, due to a possibly large number of interferers, the Gaussian assumptions becomes valid due to the superposition of many independent uniform random variables.

In order to further decrease the number of unknown parameters we propose to exploit the structure of the channel vectors $\boldsymbol{h}_j$ and $\boldsymbol{g}_j$. To choose an adequate approximation we observe that the main difficulty with estimating the unknown parameters in the models (2.6) and (2.8) is the unknown, possibly large, number of multipaths and corresponding multipath delays. In addition, note that if several multipaths have the same delay, they will be represented with only one term in the summation (2.6). Therefore, we propose to substitute, in a manner similar to [15], the *real* wireless channel with an *approximate* channel in which the number of multipaths and discretized delays are *a priori* known, i.e.,

$$\begin{aligned}\boldsymbol{h}_j &\approx \sum_{l=1}^{n_\mathrm{b}} \tilde{\boldsymbol{\alpha}}_l e^{-2\iota\pi B(j-1-n_j/2)\tilde{\tau}_l/n_\mathrm{s}} \\ \boldsymbol{\alpha}_l &= [\tilde{\alpha}_l^1, \ldots, \tilde{\alpha}_l^{n_\mathrm{r}}]^T\end{aligned} \tag{3.8}$$

where n_b is the number of basis functions, $\{\tilde{\boldsymbol{\alpha}}_l\}$ are the unknown multipath attenuations, and $\tilde{\tau}_l$ are known multipath delays. For a given maximum delay spread $\tau_{\max}$, the most reasonable choice for discretized delays (when there is no other *a priori* information available) is linearly spaced temporal grid, i.e., $\tau_l = l\tau_{\max}/n_\mathrm{b}$.

Using (3.8) we can model the user channel as

$$\boldsymbol{h}_j = H_j\boldsymbol{\theta} \tag{3.9}$$

where $\boldsymbol{\theta}$ is the unknown coefficients vector

$$\boldsymbol{\theta} = [\tilde{\alpha}_1^1, \ldots, \tilde{\alpha}_{n_\mathrm{b}}^1, \ldots, \tilde{\alpha}_1^l, \ldots, \tilde{\alpha}_{n_\mathrm{b}}^l, \ldots, \tilde{\alpha}_1^{n_\mathrm{r}}, \ldots, \tilde{\alpha}_{n_\mathrm{b}}^{n_\mathrm{r}}]^T \tag{3.10}$$

and

$$\begin{aligned}H_j &= \boldsymbol{f}_j^H \otimes I \\ \boldsymbol{f}_j &= [e^{i2\pi\tau_1 jBW/n_\mathrm{s}}, e^{i2\pi\tau_2 jBW/n_\mathrm{s}}, \ldots, e^{i2\pi\tau_l jBW/n_\mathrm{s}}]^H\end{aligned}$$

where $\otimes$ is the Kronecker product.

Similarly, for the single synchronous interferer, we propose to model the interferer's channel using

$$\boldsymbol{g}_j = U_j \boldsymbol{v} \tag{3.11}$$

where U_j is the interferer channel interpolation matrix computed at the jth subcarrier and $\boldsymbol{v}$ is the vector of interferer channel parameters. A natural choice for the matrix U_j is the same as for the user channel interpolation matrix H_j i.e., $U_j = \boldsymbol{f}_j^H \otimes I$. However, we are not limited to this choice, and thus, for generality, in the remainder of the paper, we will assume that these two matrices are different.

In the multiple interferer scenario we model the qth interferer's channel using

$$\boldsymbol{g}_j^q = U_j \boldsymbol{v}^q \tag{3.12}$$

which yields

$$\mathrm{E}[\boldsymbol{u}_{jk}\boldsymbol{u}_{jk}^H] = U_j \left(\sum_{q=1}^{n_i} \boldsymbol{v}^q \boldsymbol{v}^{qH} \right) U_j^H \tag{3.13}$$

$$= U_j^H \Psi U_j. \tag{3.14}$$

In the single interferer case, i.e., $n_i = 1$, the matrix of covariance parameters Ψ simplifies to $\boldsymbol{\eta}\boldsymbol{\eta}^H$. In the presence of multiple interferers, it may be more efficient to estimate covariance parameters using (3.14) than (3.13) since the number of unknown parameters in (3.13) is proportional to the number of interferers. Therefore, in the remainder of the paper, we will disregard structure $\sum_{q=1}^{n_i} \boldsymbol{v}^q \boldsymbol{v}^{qH}$ and use model (3.14) in the multiple interferers case.

IV. Frequency Domain Channel Estimation

In this section, we derive estimation algorithms for the models discussed in the previous section. We start by deriving MLE for both single and multiple interferer scenarios. Then, we derive a simplified, asymptotic MLE that estimates the unknown covariance parameters using the MM estimator.

A. Single Interferer Case

Recall that the received signal is given by

$$\boldsymbol{y}_{jk} = \boldsymbol{h}_j x_{jk} + \boldsymbol{g}_j \psi_{j,k} + \boldsymbol{n}_{jk} \tag{4.1}$$

with the following statistical description:

$$\begin{aligned} \boldsymbol{y}_{jk} &\sim \mathcal{N}(x_{jk}\boldsymbol{h}_j, \Sigma_j) \\ \Sigma_j &= \sigma^2 \left(I + \frac{1}{\sigma^2} \boldsymbol{g}_j \boldsymbol{g}_j^H \right). \end{aligned} \tag{4.2}$$

The distribution of the received signal is thus given by

$$\boldsymbol{y}_{jk} \sim \frac{1}{\pi^{n_\mathrm{r}} |\Sigma_j|} \exp\left[-\frac{1}{\sigma^2} \boldsymbol{\epsilon}_{jk}{}^H \Sigma_j^{-1} \boldsymbol{\epsilon}_{jk} \right].$$

It is obvious that x_{jk} and $\boldsymbol{h}_j$ cannot be estimated simultaneously. In the pilot-aided channel estimation method, we first send a training sequence of known symbols pilots x_{jk}. Using the received pilot symbols, the log-likelihood function becomes the equation shown at the bottom of the page. Next, using the parametric model from Section III and Woodbury's identity [17], the likelihood functions becomes

$$\begin{aligned} &L(\boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{\eta}, \sigma) \\ &= -n_\mathrm{p} n_\mathrm{s} n_\mathrm{r} \log\sigma - n_\mathrm{p} \sum_{j=1}^{n_\mathrm{s}} \log(1 + \boldsymbol{\eta}^H U_j^H U_j \boldsymbol{\eta}) \\ &\quad + \frac{1}{\sigma^2} \sum_{j=1}^{n_\mathrm{s}} \frac{1}{(1 + \boldsymbol{\eta}^H U_j^H U_j \boldsymbol{\eta})} \sum_{k=1}^{n_\mathrm{p}} \boldsymbol{e}_{jk}{}^H U_j \boldsymbol{\eta}\boldsymbol{\eta}^H U_j^H \boldsymbol{e}_{jk} \\ &\quad - \frac{1}{\sigma^2} \sum_{j=1}^{n_\mathrm{s}} \sum_{k=1}^{n_\mathrm{p}} \boldsymbol{e}_{jk}{}^H \boldsymbol{e}_{jk} \end{aligned} \tag{4.3}$$

where $\boldsymbol{\eta} = (1/\sigma)\boldsymbol{v}$, and

$$\boldsymbol{e}_{jk} = \boldsymbol{y}_{jk} - x_{jk} H_j \boldsymbol{\theta}. \tag{4.4}$$

For a given MLE $\hat{\boldsymbol{\eta}}$, the covariance matrix $\Sigma_j = \sigma^2 D(\hat{\boldsymbol{\eta}})$ becomes known up to the unknown parameter σ^2. Therefore, the

---

$$\begin{aligned} L(\boldsymbol{y}|\boldsymbol{h}, \boldsymbol{u}, \sigma) = &- n_\mathrm{p} n_\mathrm{s} n_\mathrm{r} \log\sigma - n_\mathrm{p} \sum_{j=1}^{n_\mathrm{s}} \log\left(1 + \frac{1}{\sigma^2} \boldsymbol{g}_j^H \boldsymbol{g}_j\right) \\ &- \frac{1}{\sigma^2} \sum_{j=1}^{n_\mathrm{s}} \sum_{k=1}^{n_\mathrm{p}} (\boldsymbol{y}_{jk} - x_{jk}\boldsymbol{h}_j)^H \left(I + \frac{1}{\sigma^2} \boldsymbol{g}_j \boldsymbol{g}_j^H \right)^{-1} (\boldsymbol{y}_{jk} - x_{jk}\boldsymbol{h}_j) \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{y} &= [\boldsymbol{y}_{11}^T, \ldots, \boldsymbol{y}_{n_\mathrm{s}1}^T, \ldots, \boldsymbol{y}_{1n_\mathrm{p}}^T, \ldots, \boldsymbol{y}_{n_\mathrm{s}n_\mathrm{p}}^T, \ldots, \boldsymbol{y}_{1n_k}^T, \ldots, \boldsymbol{y}_{n_\mathrm{s}n_\mathrm{p}}^T]^T \\ \boldsymbol{h} &= [\boldsymbol{h}_1^T, \ldots, \boldsymbol{h}_{n_\mathrm{s}}^T]^T \\ \boldsymbol{g}_j &= [\boldsymbol{g}_1^T, \ldots, \boldsymbol{g}_{n_\mathrm{s}}^T]^T. \end{aligned}$$

ML estimates $\hat{\boldsymbol{\theta}}$ and $\hat{\sigma}^2$ are generalized least squares (GLS) estimates

$$\hat{\boldsymbol{\theta}} = Q^{-1}(\hat{\boldsymbol{\eta}}) \sum_{j=1}^{n_s} H_j^H D_j^{-1}(\hat{\boldsymbol{\eta}}) \boldsymbol{s}_j^{xy}$$
$$\hat{\sigma}^2 = \frac{1}{n_p n_s n_r} \sum_{j=1}^{n_s} \hat{\boldsymbol{e}}_{j-}^H D_j^{-1}(\hat{\boldsymbol{\eta}}) \hat{\boldsymbol{e}}_{j-} \tag{4.5}$$

where

$$\boldsymbol{s}_j^{xy} = \sum_{k=1}^{n_p} x_{jk}^* \boldsymbol{y}_{jk}$$
$$\boldsymbol{e}_{j-} = \sum_{k=1}^{n_p} (\boldsymbol{y}_{jk} - x_{jk} H_j \hat{\boldsymbol{\theta}})$$
$$Q(\hat{\boldsymbol{\eta}}) = \sum_{j=1}^{n_s} s_j^x H_j^H D_j^{-1}(\hat{\boldsymbol{\eta}}) H_j$$
$$s_j^x = \sum_{k=1}^{n_p} x_{jk}^*$$
$$D_j^{-1}(\hat{\boldsymbol{\eta}}) = I - \frac{1}{1 + \hat{\boldsymbol{\eta}}^H U_j^H U_j \hat{\boldsymbol{\eta}}} U_J \hat{\boldsymbol{\eta}} \hat{\boldsymbol{\eta}}^H U_j^H.$$

Using $\hat{\boldsymbol{\theta}}$ and $\hat{\sigma}^2$, we compute the MLE $\hat{\boldsymbol{\eta}}$ by maximizing the concentrated likelihood function $L(\boldsymbol{\eta}|\hat{\boldsymbol{\theta}}, \hat{\sigma})$, i.e., we compute $\hat{\boldsymbol{\eta}}$ using the following equation:

$$\frac{\partial L(\boldsymbol{\eta}|\hat{\boldsymbol{\theta}}, \hat{\sigma})}{\partial \boldsymbol{\eta}} = 0.$$

Then

$$\sum_{j=1}^{n} U_j^H G_j(\hat{\boldsymbol{\eta}}) U_j \hat{\boldsymbol{\eta}} = 0 \tag{4.6}$$

where

$$G_j(\hat{\boldsymbol{\eta}}) = \frac{1}{\hat{\sigma}^2 (1 + \hat{\boldsymbol{\eta}} U_j^H U_j \hat{\boldsymbol{\eta}})^2} \{[\hat{\sigma}^2 (1 + \hat{\boldsymbol{\eta}} U_j^H U_j \hat{\boldsymbol{\eta}}) - \hat{\boldsymbol{\eta}}^H U_j^H E_j U_j \hat{\boldsymbol{\eta}}] I + (1 + \hat{\boldsymbol{\eta}} U_j^H U_j \hat{\boldsymbol{\eta}}) E_j\}$$
$$E_j = \frac{1}{n_p} \sum_{k=1}^{n_k} \hat{\boldsymbol{e}}_{jk} \hat{\boldsymbol{e}}_{jk}^H \tag{4.7}$$
$$\hat{\boldsymbol{e}}_{jk} = \boldsymbol{y}_{jk} - x_{jk} H_j \hat{\boldsymbol{\theta}}. \tag{4.8}$$

To solve the above equations, we use the following iterative algorithm:

• At the ith step compute $\hat{\boldsymbol{\eta}}_{i+1}$ using $\hat{\sigma}_i$ and $\hat{\boldsymbol{\theta}}_i$. Computing $\hat{\boldsymbol{\eta}}_{i+1}$ requires an additional iterative algorithm since (4.6) is non-linear with respect to $\boldsymbol{\eta}$. Thus, we propose the following algorithm:

$$\hat{\boldsymbol{\eta}}_{i+1}^{m+1} = \frac{1}{\hat{\sigma}_i^2} \left(\sum_{j=1}^{n_s} \frac{U_j^H U_j}{1 + \hat{\boldsymbol{\eta}}_{i+1}^m U_j^H U_j \hat{\boldsymbol{\eta}}_{i+1}^m} \right)^{-1} \times \sum_{j=1}^{n_s} \frac{(1 + \hat{\boldsymbol{\eta}}_{i+1}^m U_j^H U_j \hat{\boldsymbol{\eta}}_{i+1}^m) E_j - \hat{\boldsymbol{\eta}}_{i+1}^m E_j \hat{\boldsymbol{\eta}}_{i+1}^m U_j^H U_j}{(1 + \hat{\boldsymbol{\eta}}_{i+1}^m U_j^H U_j \hat{\boldsymbol{\eta}}_{i+1}^m)} \hat{\boldsymbol{\eta}}_{i+1}^m$$

with initial value $\hat{\boldsymbol{\eta}}_{i+1}^0 = \hat{\boldsymbol{\eta}}_i$

• Using $\hat{\boldsymbol{\eta}}_{i+1}$, compute $\hat{\sigma}_{i+1}$ and $\hat{\boldsymbol{\theta}}_{i+1}$.

• Stop the algorithm if $\|\hat{\boldsymbol{\eta}}_i - \hat{\boldsymbol{\eta}}_{i+1}\|^2 + (\hat{\sigma}_i - \hat{\sigma}_{i+1})^2 \leq \epsilon$. Otherwise, increase i and repeat the process.

We initialize the above algorithm using

$$\hat{\boldsymbol{\theta}}_0 = Q^{-1}(0) \sum_{j=1}^{n_s} H_j^H \boldsymbol{s}_j^{xy}$$
$$\hat{\sigma}_0^2 = \frac{1}{n_p n_s n_r} \sum_{j=1}^{n_s} \hat{\boldsymbol{e}}_{j-}^H \hat{\boldsymbol{e}}_{j-} \tag{4.9}$$
$$\hat{\boldsymbol{\eta}}_0 \in \text{null} \left\{ \sum_{j=1}^{n} U_j^H G_j(0) U_j \right\} \tag{4.10}$$

where null denotes the null space of matrix $\sum_{j=1}^{n} U_j^H G_j(0) U_j$.

B. Multiple Interferers

The likelihood function in this case becomes

$$L(\boldsymbol{y}|\boldsymbol{\theta}, \Psi) = -\sum_{j=1}^{n_s} \sum_{k=1}^{n_p} (\boldsymbol{y}_{jk} - x_{jk} H_j \boldsymbol{\theta})^H \Sigma_j^{-1}(\Psi, \sigma^2) (\boldsymbol{y}_{jk} - x_{jk} H_j \boldsymbol{\theta})$$
$$\Sigma_j(\Psi, \sigma^2) = U_j \Psi U_j^H + \sigma^2 I. \tag{4.11}$$

The estimate of $\boldsymbol{\theta}$ that maximizes (4.11) for any given Ψ is given by

$$\hat{\boldsymbol{\theta}}(\Psi) = \left(\sum_{j=1}^{n_s} s_j \left(U_j \Psi U_j^H + \sigma^2 I \right)^{-1} H_j^H \right)^{-1} \times \sum_{j=1}^{n_s} H_j \left(U_j \Psi U_j^H + \sigma^2 I \right)^{-1} H_j^H \boldsymbol{s}_j. \tag{4.12}$$

Substituting (4.12) in (4.11), we obtain the concentrated log-likelihood function

$$L(\Psi|\hat{\boldsymbol{\theta}}(\Psi)) == -\frac{1}{2} \times \sum_{j=1}^{n_s} \sum_{k=1}^{n_p} (\boldsymbol{y}_{jk} - x_{jk} H_j \hat{\boldsymbol{\theta}})^H \Sigma_j^{-1}(\Psi, \sigma^2) (\boldsymbol{y}_{jk} - x_{jk} H_j \hat{\boldsymbol{\theta}}). \tag{4.13}$$

Note that (4.13) is a highly nonlinear function of the variance components Ψ. Once the ML estimate of $\hat{\Psi}$ computed by maximizing (4.13), the ML estimate of $\boldsymbol{\theta}$ is obtained by substituting $\hat{\Psi}$ into (4.11).

We compute $\hat{\Psi}$ using an expectation-conditional maximization either (ECME) algorithm which results in an iteration between the following two steps:

$$\Sigma_j^{(i)}(\Psi^{(i)}) = U_j\Psi^{(i)}U_j^H + (\sigma^2)^{(i)}I$$

$$\boldsymbol{\theta}^{(i)} = \left(\sum_{j=1}^{n_s} s_j \Sigma_j^{(i)}(\Psi^{(i)})H_j^H\right)^{-1} \times \sum_{j=1}^{n_s} H_j\Sigma_j^{(i)}(\Psi^{(i)})H_j^H \boldsymbol{s}_j$$

$$\boldsymbol{u}_{jk}^{(i)} = \Psi^{(i)}U_j^H U_j \Sigma_j^{(i)^{-1}}\left(\boldsymbol{y}_{jk} - H_j\boldsymbol{\theta}^{(i)}x_{jk}\right) \quad (4.14)$$

for $j = 1, 2, \ldots, n_s$, and (4.15), shown at the bottom of the page. The above iteration can be initialized with $\Psi^{(-1)} = 0$, implying that the initial estimate $\boldsymbol{\theta}^{(-1)}$ is simply the ordinary least squares estimate. After computing $\boldsymbol{\theta}^{(-1)}$, a good initial estimate of $\Psi^{(0)}$ is its MM estimator (see Section IV-C), shown at the bottom of the page.

C. Asymptotic Maximum Likelihood

Since our interest lies primarily in drawing inference about $\boldsymbol{\theta}$ (and subsequently x_{jk}), we propose a simplified estimation scheme: estimate Ψ (or $\boldsymbol{\eta}$) and σ^2 using any consistent estimator, for example, MM. This method provide simple noniterative estimates of the variance-covariance parameters without sacrificing any efficiency in the estimation of $\boldsymbol{\theta}$ and, consequently, symbol detection [12].

The objective function is the same as for the MLE except that $\Sigma_j(\Psi)$ is replaced by $\Sigma_j(\hat{\Psi})$, where $\hat{\Psi}$ is any consistent estimator of Ψ. Thus, $\hat{\boldsymbol{\theta}}(\Psi)$ is given by

$$\hat{\boldsymbol{\theta}}(\Psi) = \left(\sum_{j=1}^{n_s} s_j(U_j\Psi U_j^H + \sigma^2 I)^{-1}H_j^H\right)^{-1} \times \sum_{j=1}^{n_s} H_j(U_j\Psi U_j^H + \sigma^2 I)^{-1}H_j^H\boldsymbol{s}_j \quad (4.16)$$

which is equivalent to the MLE whenever $\hat{\Psi}$ is the MLE. Furthermore, since $\hat{\Psi} \to \Psi$ as $n_p \to \infty$, the above estimator can be shown to be asymptotically equivalent to the MLE [12].

We start by rewriting the received signal as

$$\boldsymbol{y}_{jk} = x_{jk}H_j\boldsymbol{\theta} + U_j\boldsymbol{\xi}_{jk} + \boldsymbol{e}_{jk} \quad (4.17)$$

where $\boldsymbol{\xi}_{jk}$ is a random vector distributed as

$$\boldsymbol{\xi}_{jk} \sim \mathcal{N}(0, \Psi)$$
$$\mathrm{E}[\boldsymbol{\xi}_{jk}\boldsymbol{\xi}_{j'k'}^H] = \Psi\delta(j-j')\delta(k-k'). \quad (4.18)$$

Next, we note that the unobserved residuals can be written as

$$\boldsymbol{\epsilon}_{jk}(\boldsymbol{\theta}) = \boldsymbol{y}_{jk} - H_j\boldsymbol{\theta} = U_j\boldsymbol{\xi}_{jk} + \boldsymbol{e}_{jk}. \quad (4.19)$$

If we estimate $\boldsymbol{\xi}_{jk}$ via the usual least squares estimator

$$\hat{\boldsymbol{\xi}}_{jk} = (U_j^H U_j)^{-1}U_j^H\boldsymbol{\epsilon}_{jk}(\boldsymbol{\theta}) \quad (4.20)$$

and σ^2 via the usual mean square for error

$$\boldsymbol{s}_{jk} = \boldsymbol{e}_{jk}^H(\boldsymbol{\theta})[I - U_j(U_j^H U_j)^{-1}U_j^H]\frac{\boldsymbol{e}_{jk}(\boldsymbol{\theta})}{(n_r - 1)} \quad (4.21)$$

then the MM estimator can be obtained by taking the expectations of $\psi_{jk}\psi_{jk}^*$ and averaging them across the $n_s n_k$ measurements. Since a pooled MM estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n_s n_p(n_r - 1)}\sum_{j=1}^{n_s}\sum_{k=1}^{n_k} s_{jk}^2 \quad (4.22)$$

and since

$$\mathrm{E}[\hat{\boldsymbol{\xi}}_{jk}\hat{\boldsymbol{\xi}}_{jk}^H] = \Psi + \hat{\sigma}^2(U_j^H U_j)^{-1} \quad (4.23)$$

for a fixed $\boldsymbol{\theta}$, it follows that a simple MM estimator of Ψ is

$$\widehat{(\Psi)} = \frac{1}{n_s n_p}\sum_{j=1}^{n_s}\sum_{k=1}^{n_p}\hat{\boldsymbol{\xi}}_{jk}\hat{\boldsymbol{\xi}}_{jk}^H - \frac{\hat{\sigma}^2}{n_s n_p}\sum_{j=1}^{n_s}(U_j^H U_j)^{-1}. \quad (4.24)$$

A nice feature of $\hat{\sigma}^2$ is that it is invariant [12] to the value of $\boldsymbol{\theta}$ used in forming the residual vector $\boldsymbol{e}_{jk}$, i.e., it is computed using only the received signal and the interpolation matrices. Unfortunately, this invariance property does not hold for $\widehat{\Psi}$. By writing $\hat{\Psi}$ as $\hat{\Psi}(\boldsymbol{\theta})$ to reflect the dependence of the MM estimator

$$\Psi^{(i+1)} = \Psi^{(i)} + \frac{1}{n_s}\sum_{j=1}^{n_s}\boldsymbol{u}_{jk}^{(i)}(\boldsymbol{u}_{jk}^i)^H - \Psi^{(i)}U_j^H U_j((\sigma^2)^{(i)}U_j^H U_j + U_j^H U_j\Psi^{(i)}U_j^H U_j)^{-1}U_j^H U_j\Psi^{(i)}$$

$$(\sigma^2)^{(i+1)} = \frac{n_r}{n_p}(\sigma^2)^{(i)} + \frac{1}{n_p n_s n_r} \times \sum_{j=1}^{n_s}\left((\boldsymbol{y}_{jk} - H_j\boldsymbol{\theta}^{(i)}x_{jk} - \boldsymbol{u}_{jk}^{(i)})^H(\boldsymbol{y}_{jk} - H_j\boldsymbol{\theta}^{(i)}x_{jk} - \boldsymbol{u}_{jk}^{(i)}) - [(\sigma^2)^i]^2\mathrm{tr}\{[(\sigma^2)^{(i)}U_j^H U_j + U_j^H U_j\Psi^{(i)}U_j^H U_j]\}\right). \quad (4.15)$$

$$\Psi^{(0)} = \frac{1}{n_s n_p}\sum_{j=1}^{n_s}\sum_{k=1}^{n_p}(U_j^H U_j)^{-1}U_j^H(\boldsymbol{y}_{jk} - H_j\boldsymbol{\theta}^{(-1)}x_{jk})(\boldsymbol{y}_{jk} - H_j\boldsymbol{\theta}^{(-1)}x_{jk})(\boldsymbol{y}_{jk} - H_j\boldsymbol{\theta}^{(-1)}x_{jk})^H U_j(U_j^H U_j)^{-1}.$$

on $\boldsymbol{\theta}$, we can estimate Ψ using some iterative algorithm such as iteratively reweighted least squares [12].

Rather than performing iteratively reweighted least squares, an alternative approach would be to use an MM estimator of Ψ, which, like $\hat{\sigma}^2$ is invariant with respect to $\boldsymbol{\theta}$. In general such an estimator may not be available. However, for a particular choice of interpolation matrices H_j and U_j, we can derive a residual method of moments (RMM) estimate of Ψ, similar to the estimator proposed in [12]. Namely, for interpolation matrices such that

$$H_j = U_j A_j \tag{4.25}$$
$$A_j = I \otimes \boldsymbol{a}_j^H \tag{4.26}$$

the RMM estimate (independent of $\boldsymbol{\theta}$) can be shown to be

$$\widehat{\Psi}_{\rm RMM} = S - \frac{\hat{\sigma}^2}{n_{\rm s} n_{\rm p}} \sum_{j=1}^{n_{\rm s}} [1 - \boldsymbol{a}_j^H (A^H A)^{-1} \boldsymbol{a}_j](U_j^H U_j)^{-1} \tag{4.27}$$

where

$$A = [\boldsymbol{a}_1 \dots \boldsymbol{a}_{n_{\rm s}}]^H \tag{4.28}$$
$$S = \tilde{B}(I - A(A^H A)A^H))\tilde{B} \tag{4.29}$$
$$\tilde{B} = [\tilde{b}_{11}, \dots, \tilde{b}_{1n_{\rm p}}, \tilde{b}_{21}, \dots\dots, \tilde{b}_{n_{\rm s} n_{\rm p}}]^H \tag{4.30}$$
$$\tilde{b}_{jk} = (U_j^H U_J)^{-1} U_j^H \boldsymbol{y}_{jk}. \tag{4.31}$$

The RMM estimate is similar to the residual or restricted maximum likelihood (REML) estimation in that it corrects for a bias due to small number of pilots.

1) Asymptotic Maximum Likelihood-Single Interferer: The asymptotic MLE from the previous section estimates $n_{\rm b}(n_{\rm b}+1)/2$ unknown variance parameters. When only one interferer is present, the proposed algorithm is suboptimal since it does not account for additional structural information $\Psi = \boldsymbol{\eta}\boldsymbol{\eta}^H$. To account for the additional information, an iterative algorithm with respect to $\boldsymbol{\eta}$ is required. It is possible, then, that it will provide little improvement in computing time over the fully iterated MLE presented in Section IV.B. To avoid iterations, we propose to employ a suboptimal approach and fit $\hat{\Psi}$ to $\boldsymbol{\eta}\boldsymbol{\eta}^H$ using the generalized least squares (GLS). Then, the objective function for the GLS can be written as

$$c_{\rm GLS}(\boldsymbol{\eta}) = \sum_{i,j=1}^{n_{\rm b}} w_{ij}(\|[\hat{\Psi}]_{ij} - \eta_i \eta_j^*\|^2) \tag{4.32}$$

where w_{ij} are the covariances of $[\hat{\Psi}^{-1}]_{ij}$. Since $\hat{\Psi}$ has a noncentral Wishart distribution [14] with the noncentrality factor dependent on $\boldsymbol{\eta}$, the weighting coefficients w_{ij} are also dependent on $\boldsymbol{\eta}$. The estimation equations

$$\frac{\partial c_{\rm GLS}(\boldsymbol{\eta})}{\partial \eta_l} = \sum_{i,j=1}^{n_{\rm b}} \frac{\partial w_{ij}}{\eta_l}(\|[\hat{\Psi}]_{ij} - \eta_i \eta_j^*\|^2) - 2\sum_{i,j=1}^{n_{\rm b}} w_{ij}\eta_j^*(\|[\hat{\Psi}]_{ij} - \eta_i \eta_j^*\|) \quad l = 1, \dots, n_{\rm b} \tag{4.33}$$

are highly nonlinear and thus may not significantly reduce the computational complexity. Instead, we propose to use the empirical covariances $\hat{\Lambda}^{-1}$, which yields the following set of $n_{\rm b}$ nonlinear equations:

$$\sum_{j=1}^{n_{\rm b}} [\hat{\Lambda}^{-1}]_{ij} \eta_j^* \left([\hat{\Lambda}]_{ij} - \eta_i^* \eta_j\right) = 0 \tag{4.34}$$

which can be solved easily using any standard method for solving nonlinear equations. It is important to note that the choice of the weighting factors w_{ij} will have a little impact on inference about $\boldsymbol{\theta}$, but they may influence inference related to σ^2 and $\boldsymbol{\eta}$.

A common drawback of MM type estimators is that they may occasionally produce negative-definite estimates. To ensure having a non-negative definite estimate of Ψ, we can apply the correction procedure described in [12]. However, this procedure would require an iterative algorithm, in which case, MLE may be better a choice.

D. MANOVA Model Estimator

The above asymptotic MLE exploits the covariance structure $\Sigma_j(\Psi, \sigma^2)$, which results in smaller training overhead requirements. For comparison purposes, in this section, we derive an MLE algorithm for the unstructured covariance (MANOVA) model. Using these results, we will compare the performances of the asymptotic MLE and MLE for various interference scenarios in Section V.

First, we lump the received waveforms into one vector

$$\boldsymbol{y}_k = \begin{bmatrix} \boldsymbol{y}_{1k} \\ \vdots \\ \boldsymbol{y}_{n_{\rm s}k} \end{bmatrix}. \tag{4.35}$$

Next, rewrite the received signal model (2.5) as

$$\boldsymbol{y}_k = Q_k \boldsymbol{\theta} + \boldsymbol{e}_k \tag{4.36}$$

where

$$Q_k = \begin{bmatrix} x_1 H_1 \\ \vdots \\ x_{n_{\rm s}} H_{n_{\rm s}} \end{bmatrix}. \tag{4.37}$$

In the MANOVA model, $\boldsymbol{e}_k$ are independent and identically distributed (i.i.d.) according to a multivariate normal distribution with zero mean and covariance matrix Σ

$$\boldsymbol{e}_k \sim \mathcal{N}(0, \Sigma). \tag{4.38}$$

The log-likelihood function for this model is

$$L(\boldsymbol{y}|\boldsymbol{\theta}, \Sigma) = -\frac{1}{2}\text{trace}\left\{\Sigma^{-1} \sum_{i=1}^{n_{\rm s}} \boldsymbol{e}_k \boldsymbol{e}_k^H\right\} - \frac{n_{\rm s} n_{\rm p}}{2}|\Sigma| \tag{4.39}$$

$$\boldsymbol{e}_k = \boldsymbol{y}_k - Q_k \boldsymbol{\theta}. \tag{4.40}$$

The MLE of $\boldsymbol{\theta}$ and Σ are, respectively

$$\hat{\boldsymbol{\theta}} = (Q^T Q)^{-1} Q^T \boldsymbol{y} \tag{4.41}$$

$$\hat{\Sigma} = \frac{1}{n_{\mathrm{s}} n_k} \boldsymbol{y}(I - Q(Q^T Q)^{-1} Q^T)\boldsymbol{y}^T \tag{4.42}$$

$$Q = \begin{bmatrix} Q_1 \\ \vdots \\ Q_{n_{\mathrm{p}}} \end{bmatrix} \tag{4.43}$$

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_{\mathrm{p}}} \end{bmatrix}. \tag{4.44}$$

Observe that the above model (4.36) and (4.38) does not assume the structured covariance $\Sigma_j(\theta) = U_j \Psi U_j^H + \sigma^2 I$ and may perform better than when the covariance deviates significantly from this structure. However, since the number of unknown parameters is much larger in the unstructured model, the asymptotic MLE may still give better results than MANOVA MLE for the same number of pilots. However, the smaller number of parameters does not necessarily mean better performance, for example the OLS algorithm always under performs both the asymptotic MLE and MANOVA MLE.

E. Symbol Detection

In the previous section, we pointed out that the symbols x_{jk} and channel/interference parameters $(\boldsymbol{\theta}, \Psi, \sigma^2)$ cannot be estimated jointly. We have also shown that by sending *known* pilot symbols we can compute the channel parameters $\hat{\boldsymbol{\theta}}$ and $\hat{\Psi}$. Then, assuming that the channel and interferer properties do not change in time (or change sufficiently slowly) we can use these estimates to obtain ML estimates of the unknown symbols. Namely, we can rewrite the likelihood function (4.3) as

$$\begin{aligned} L(\boldsymbol{x}) = &-\frac{1}{2} \\ &\times \sum_{j=1}^{n_{\mathrm{s}}} \sum_{k=1}^{n_{\mathrm{p}}} \left(\boldsymbol{y}_{jk} - x_{jk} H_j \hat{\boldsymbol{\theta}}\right)^H \Sigma_j^{-1}(\hat{\Psi}, \hat{\sigma}^2) \left(\boldsymbol{y}_{jk} - x_{jk} H_j \hat{\boldsymbol{\theta}}\right) \\ \Sigma_j(\hat{\Psi}, \hat{\sigma}^2) &\\ = &U_j \hat{\Psi} U_j^H + \hat{\sigma}^2 I. \end{aligned} \tag{4.45}$$

Next, let $\{\xi_1, \ldots, \xi_n\}$ be an OFDM signal constellation. The detected symbol $\hat{x}_{jk}$ is then given by

$$\hat{x}_{jk} = \underset{\xi_l}{\operatorname{argmin}} \left[\left(\boldsymbol{y}_{jk} - \xi_l H_j \hat{\boldsymbol{\theta}}\right)^H \Sigma_j^{-1}(\hat{\Psi}, \hat{\sigma}^2) \left(\boldsymbol{y}_{jk} - \xi_l H_j \hat{\boldsymbol{\theta}}\right) \right]. \tag{4.46}$$

In the single interferer case the above expression simplifies to

$$\hat{x}_{jk}^* = \underset{\xi_l}{\operatorname{argmin}} \left(\frac{\boldsymbol{y}_{jk}^H D_n^{-1}(\hat{\boldsymbol{\eta}}) H_j \hat{\boldsymbol{\theta}}}{\hat{\boldsymbol{\theta}}^H H_j^H D_n^{-1}(\hat{\boldsymbol{\eta}}) H_j \hat{\boldsymbol{\theta}}} - \xi_l^* \right). \tag{4.47}$$

V. Numerical Examples

We demonstrate the performance of the proposed estimators by numerical examples. In all examples, unless otherwise stated, we assume $n_{\mathrm{r}} = 4$ element antenna array and a 16-QAM signal occupying a bandwidth $B = 6.0$ MHz. The number of usable carriers is set to $n_{\mathrm{s}} = 370$. The number of pilot bauds is $n_{\mathrm{p}} = 3$ with pilot symbols on all subcarriers. The number of basis functions for both interpolation matrices is $n_{\mathrm{b}} = 40$ (ten for each antenna). The maximum multipath delay is $\tau_{\max} = 8.0$ μs. We define the signal-to-interference (SIR) as the ratio of signal power to interference power and set it to 15 dB. The number of user and interferer (for each interferer) multipaths is $n_{\mathrm{mu}} = n_{\mathrm{mi}} = 100$ with an exponential power delay profile. The phase shift on each path is uniformly distributed over $[0, 2\pi)$. We generate synchronous interference using QPSK random source and filtering it through the interferer channel.

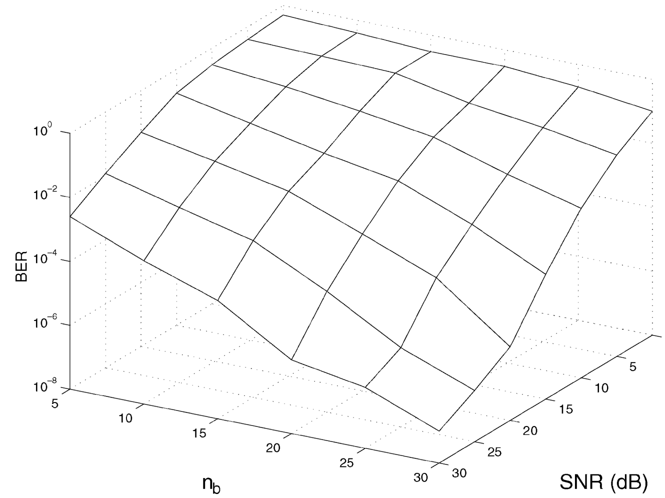


Fig. 1. BER as a function of number of basis functions and SNR—single interferer.

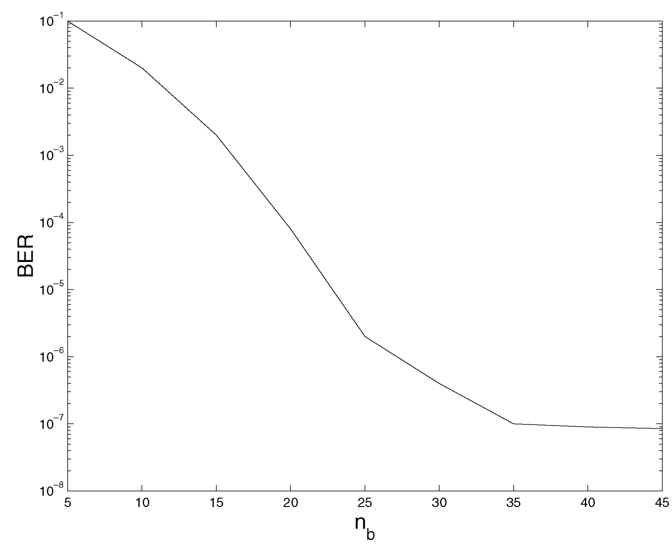


Fig. 2. BER as a function of number of basis functions SNR $= 25$ db.

Fig. 1 illustrates the bit-error-rate (BER) performance of the proposed asymptotic MLE as a function of the number of basis functions n_{b} and SNR in the presence of single synchronous interferer. It can be seen that by increasing number of basis functions we can achieve significant improvement in the performance. In Fig. 2, we show BER performance as a function of number of basis functions for a fixed SNR $= 25$ db. It can

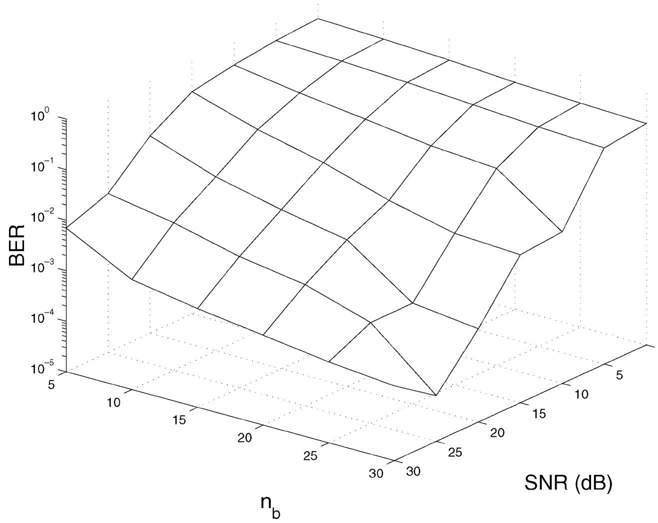


Fig. 3. BER as a function of n_b and SNR—multiple interferers.

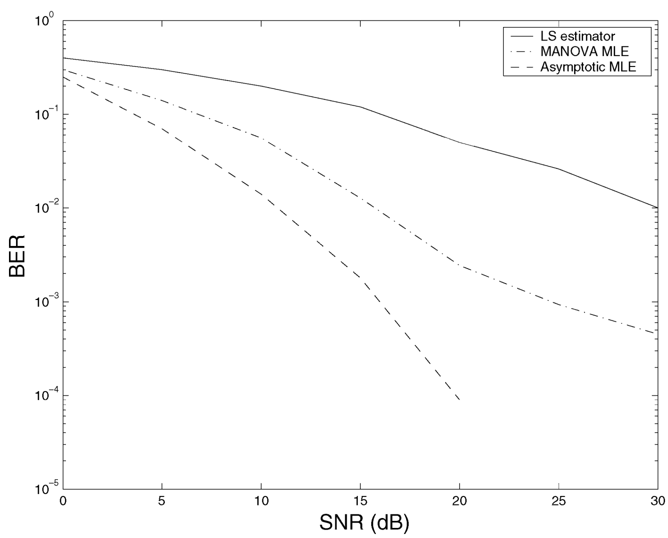


Fig. 4. BER performance comparison—single synchronous interferer.

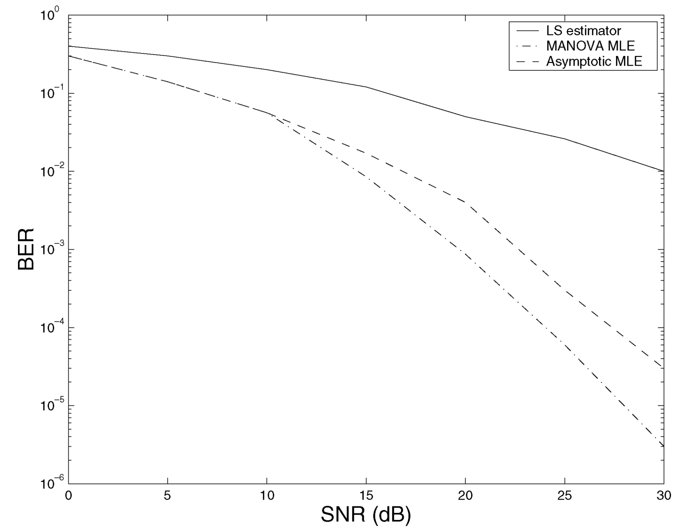


Fig. 5. BER performance comparison—single asynchronous interferer.

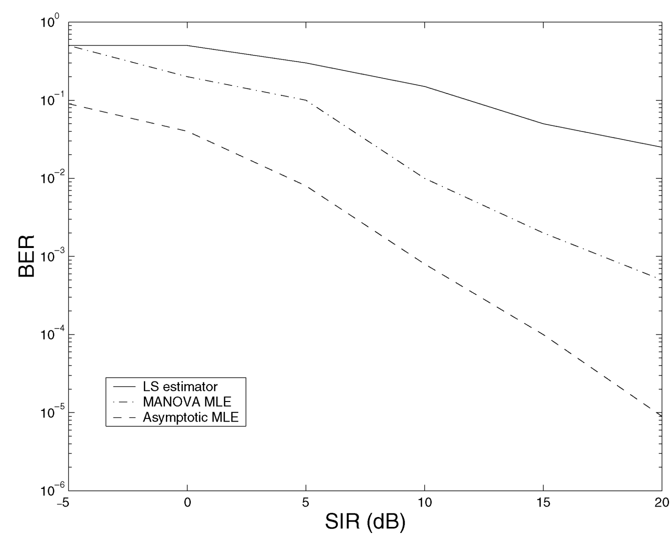


Fig. 6. BER performance comparison as a function of SIR.

be seen that for a particular SNR there exists a threshold above which increasing the number of basis functions will not yield significant improvement in the performance.

In Fig. 3, we illustrate the same results as in Fig. 1 but when three synchronous interferers are present. For comparison purposes, we keep the SNR at the same level. Observe that in this case the performance is improved since, as we stated earlier, the Gaussian assumption is better approximation in the multiple interferers case.

For comparison purposes, in Fig. 4, we present the performance of the OLS, the asymptotic MLE, and the unstructured MLE as a function of SNR for the single interferer scenario. As expected, the asymptotic MLE outperforms the other two estimators. The OLS estimator has smaller number of parameters (one for the covariance) but does not account for heteroskedastic interference properties, namely frequency-dependent spatial covariance and thus under-performs the asymptotic MLE. On the other hand, the unstructured MLE models the covariance using large number of parameters that cannot be estimated efficiently with a small number of pilots. Fig. 5 illustrates the same results for a single asynchronous interferer. Recall that in the asynchronous case the cyclic prefixes do not align in time and thus, the interference can be modeled as a Gaussian process uncorrelated in time and frequency with a frequency-dependent spatial covariance. Therefore, to account for asynchronous interference, we generated 370 random vectors with zero mean Gaussian distribution and different, randomly chosen spatial covariances. In this case, the covariance structure imposed by the asymptotic MLE may not be adequate and thus the unstructured MLE may outperform the asymptotic MLE regardless of the larger number of parameters.

In Fig. 6, we present the performance of OLS, asymptotic MLE, and the unstructured MLE as a function of signal-to-interference (SIR) ratio (single interferer scenario) for a fixed SNR = 20 db. As expected, the structured model shows significant improvement in the performance, especially for larger SIR.

VI. Possible Extensions

A. Multiple Users

It has been shown [18], [19] that antenna arrays can increase the system capacity by allowing multiple users to share the same time-frequency resources, a practice called spatial division multiple access (SDMA). In this section, we discuss possible extensions of our results to the multiple user scenario.

Assume there are n_c SDMA users (or transmit antennas in MIMO/BLAST case) in the system sharing the channel. The received signal in the frequency domain is

$$\boldsymbol{y}_{jk} = \sum_{l=1}^{n_c} \boldsymbol{h}_{j,l} x_{jk,l} + \boldsymbol{u}_{jk} + \boldsymbol{e}_{jk}. \tag{6.1}$$

We propose to model the the lth user channel using

$$\boldsymbol{h}_{j,l} = H_j \boldsymbol{\theta}_l. \tag{6.2}$$

Observe that we model different user channels with the same interpolation matrix H_j but different parameters $\boldsymbol{\theta}_l$. If the interpolation matrices are different the unknown linear parameters in the resulting model will be identifiable only when specific conditions are satisfied [20].

Using the above model, (6.1) can be written as

$$\boldsymbol{y}_{jk} = H_j \Theta \boldsymbol{x}_{jk} + \boldsymbol{u}_{jk} + \boldsymbol{e}_{jk} \tag{6.3}$$

where

$$\Theta = [\boldsymbol{\theta}_1 \cdots \boldsymbol{\theta}_{n_c}] \tag{6.4}$$

$$\boldsymbol{x}_{jk} = [x_{jk,1}, \ldots, x_{jk,n_c}]^T. \tag{6.5}$$

Next, let

$$Y_j = [\boldsymbol{y}_{j1} \cdots \boldsymbol{y}_{jn_p}] \tag{6.6}$$

be the collection of all received pilots, and

$$X_j = [\boldsymbol{x}_{j1} \cdots \boldsymbol{x}_{jn_p}] \tag{6.7}$$

be the collection of all pilot symbols at the jth subcarrier. The received signal is described in a matrix form as

$$Y_j = H_j \Theta X_j + \Upsilon_j + E_j \tag{6.8}$$

where Υ_j represents the random interference matrix, and E_j represents the random ambient noise matrix. We assume that the rows of E_j are i.i.d. according to a normal distribution, i.e.,

$$\text{Vec}(E_j) \sim \mathcal{N}(0, \sigma^2 I). \tag{6.9}$$

We continue to assume a synchronous interferer case, i.e.,

$$\Upsilon_j = U_j \boldsymbol{\eta} \otimes \boldsymbol{\psi}_j \tag{6.10}$$

$$\boldsymbol{\psi}_j = [\psi_{j1}, \ldots, \psi_{jn_k}]^T \tag{6.11}$$

resulting in the following distribution:

$$\text{Vec}(\mathrm{U}_j) \sim \mathcal{N}(0, (\mathrm{U}_j \boldsymbol{\eta}\boldsymbol{\eta}^H \mathrm{U}_j^H) \otimes \mathrm{I}). \tag{6.12}$$

Computing the MLE of Θ and $\boldsymbol{\eta}$ in the above model requires an iterative algorithm. However, we can further simplify the model, assuming that the pilot symbols sent on different carriers are the same, i.e., $X_j = X_i$ for all i, j. In this case the distribution of the received signal will be given by

$$\text{Vec}(Y) \sim \mathcal{N}(0, [U(\boldsymbol{\eta}\boldsymbol{\eta}^H \otimes I)] \otimes I) \tag{6.13}$$

where

$$U = \text{bdiag}[U_1, \ldots, U_{n_s}] \tag{6.14}$$

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_{n_s} \end{bmatrix} \tag{6.15}$$

where bdiag is the block-diagonal operator.

Extending the asymptotic MLE presented in Section IV-B to the present model is straightforward.

VII. Conclusions

We presented pilot-aided channel estimation algorithms for estimating OFDM wireless channels in the presence of synchronous interference. We demonstrated that the number of covariance parameters in the synchronous case can be significantly reduced by exploiting *a priori* available information about the interference structure. To further reduce the number of unknown parameters, we modeled the user and interference channel responses using basis functions. As a result of the above approximations we obtained a structured covariance model with smaller number of parameters without significant loss in detection efficiency. We derived MLE and asymptotic MLE algorithms for estimating these parameters. The asymptotic MLE estimates the unknown interference parameters using a residual method of moments estimator invariant to the user channel parameters resulting in a computationally efficient noniterative algorithm. We have also compared the performance of three algorithms: ordinary least squares, asymptotic MLE, and unstructured MLE. We showed that the asymptotic MLE can achieve desired performance with a small number of pilots, unlike the unstructured MLE. We demonstrated that even when asynchronous or multiple synchronous interferers are present, the performance of the asymptotic MLE is comparable to the performance of the MANOVA MLE. An effort will be made to examine the applicability of our approach to time-varying channel and multiple user scenarios.

References

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

[2] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Trans. Commun.*, vol. COM-19, pp. 628–634, Oct. 1971.

[3] "Radio Broadcasting System, Audio Broadcasting (DAB) to Mobile, Portable, and Fixed Receivers," Eur. Telecommun. Stand. Inst., Sophia-Antapolis, Valbonne, France, ETS 300 401, 1995–1997.

[4] "Digital Broadcasting System Television, Sound and Data Services; Framing Structure, Channel Coding, and Modulation Digital Terrestrial Television," Eur. Telecommun. Stand. Inst., Sophia-Antapolis, Valbonne, France, ETS 300 744, 1996.

[5] "Broadband Radio Access Networks (BRAN); High Performance Radio Local Area Networks (HIPERLAN) Type 2; System Overview," Eur. Telecommun. Stand. Inst., Sophia-Antapolis, Valbonne, France, ETR 101 683 114, 1999.

[6] J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Veh. Technol. Conf.*, Chicago, IL, 1995, pp. 815–819.

[7] X. Zhuang, Z. Ding, and A. L. Swindlehurst, "A statistical subspace method for blind channel identification in OFDM communications," in *Proc. IEEE ICASSP*, Istanbul, Turkey, 2000.

[8] C. Tepedelenlioglu and G. B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. Signal Processing*, vol. 48, pp. 2029–2043, July 2000.

[9] B. Hassibi and B. Hochwald, "Optimal training in space-time systems," in *Proc. 34th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 2000.

[10] A. Jeremicć, T. A. Thomas, and A. Nehorai, "OFDM channel estimation in the presence of interference," in *IEEE SAM*, Rosslyn, VA, Aug. 2002.

[11] Y. Li and N. R. Sollenberger, "Adaptive antenna arrays for OFDM systems with co-channel interference," *IEEE Trans. Commun.*, vol. 47, pp. 217–229, Feb. 1999.

[12] E. F. Vonesh and R. L. Carter, "Efficient inference for a random coefficient growth curve model with unbalanced data," *Biometrics*, vol. 43, pp. 617–628, 1987.

[13] A. Jeremić, T. Thomas, and A. Nehorai, "OFDM channel estimation in the presence of interference," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Hong Kong, June 2003.

[14] C. R. Rao, *Linear Statistical Inference and Its Applications*, Second ed. New York: Wiley, 1973.

[15] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1985, Nov. 1998.

[16] T. A. Thomas, F. W. Wook, and K. L. Baum, "Least-squares multi-user frequency-domain channel estimation for broadband wireless systems," in *Proc. 37th Annual Allerton Conf. Commun., Contr. Comput.*, Monticello, IL, Sept. 1999, pp. 1026–1035.

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third ed. Baltimore, MD: John Hopkins Univ. Press, 1996.

[18] R. T. Compton Jr., *Adaptive Antennas: Concepts and Performance*. Englewood Cliff, NJ: Prentice-Hall, 1988.

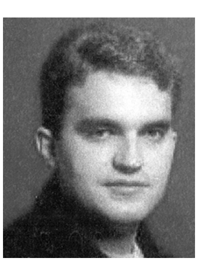
[19] J. H. Winters, "Optimum combining in digital mobile radio with co-channel interference," *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 1200–1204, Aug. 1984.

[20] A. P. Verbyla and W. N. Venables, "An extension of the growth curve model," *Biometrika*, vol. 75, pp. 129–138, 1988.

[21] P. W. Wolniansky, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE*, Pisa, Italy, 1998.

[22] H. Bolsckei, D. Gesbert, and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[23] Y. G. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1207–1216, July 2000.



Aleksandar Jeremić received the B.Sc. degree in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1995 and the M.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Chicago in 1997 and 2003, respectively.

Currently, he is an Assistant Professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. His research interests are in statistical signal processing and biomedicine.

Timothy A. Thomas (S'95–M'98) received the B.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1989, the M.S. degree in electrical engineering from the University of Michigan, Ann Arbor, in 1990, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1997.

From 1991 to 1993, he was a Member of Technical Staff at Bell Communications Research (Bellcore), Red Bank, NJ. Since 1997, he has been with the Communications Systems Research Lab, Motorola Labs, Schaumburg, IL. His research interests are in physical layer communications and, in particular, channel measurements, channel estimation, and adaptive antenna algorithms for mobile broadband communications.

Dr. Thomas is currently serving as associate editor for the IEEE SIGNAL PROCESSING MAGAZINE.

Arye Nehorai (S'80–M'83–SM'90–F'94) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, Haifa, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

After graduation, he worked as a Research Engineer for Systems Control Technology, Inc., Palo Alto, CA. From 1985 to 1989, he was Assistant Professor, and from 1989 to 1995, he was Associate Professor with the Department of Electrical Engineering, Yale University, New Haven, CT. In 1995, he joined the Department of Electrical Engineering and Computer Science, The University of Illinois at Chicago (UIC), as a Full Professor. From 2000 to 2001, he was Chair of the department's Electrical and Computer Engineering (ECE) Division, which is now a new department. He holds a joint professorship with the ECE and Bioengineering Departments at UIC. His research interests are in signal processing, communications, and biomedicine.

Dr. Nehorai is Vice President—Publications and Chair of the Publications Board of the IEEE Signal Processing Society. He is also a member of the Board of Governors and of the Executive Committee of this Society. He was Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from January 2000 to December 2002 and is currently a Member of the Editorial Boards of *Signal Processing*, the IEEE SIGNAL PROCESSING MAGAZINE, and *The Journal of the Franklin Institute*. He is the founder and Guest Editor of the special columns on Leadership Reflections in the the IEEE SIGNAL PROCESSING MAGAZINE. He was an Associate Editor of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, the IEEE SIGNAL PROCESSING LETTERS, the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, the IEEE JOURNAL OF OCEANIC ENGINEERING, and *Circuits, Systems, and Signal Processing*. He served as Chairman of the Connecticut IEEE Signal Processing Chapter from 1986 to 1995 and as a Founding Member, Vice-Chair, and later Chair of the IEEE Signal Processing Society's Technical Committee on Sensor Array and Multichannel (SAM) Processing from 1998 to 2002. He was the co-General Chair of the First and Second IEEE SAM Signal Processing Workshops held in 2000 and 2002. He was co-recipient, with P. Stoica, of the 1989 IEEE Signal Processing Society's Senior Award for Best Paper and co-author of the 2003 Young Author Best Paper Award of this Society, with A. Dogandzic. He received the Faculty Research Award from the UIC College of Engineering in 1999 and was Adviser of the UIC Outstanding Ph.D. Thesis Award in 2001. In addition, in 2001, he was named University Scholar of the University of Illinois. He was elected Distinguished Lecturer of the IEEE Signal Processing Society for the term 2004 to 2005. He has been a Fellow of the Royal Statistical Society since 1996.