Blind Channel Estimation for Multicarrier Systems With Narrowband Interference Suppression

Hongbin Li, *Member, IEEE*

***Abstract*—We present herein a novel blind channel estimator for multicarrier (MC) systems in the presence of *unmodeled* narrowband interference. A generalized multichannel minimum variance principle is invoked to design an equalizing filterbank that preserves desired signal components and suppresses the overall interference. While a channel estimate may be obtained by directly maximizing the filterbank output power through multidimensional nonlinear searches, such an approach is computationally prohibitive and suffers local convergence. To overcome this difficulty, we derive an asymptotically (in SNR) tight lower bound of the filterbank output power and use it for channel estimation, which reduces the problem to a quadratic minimization. Numerical examples show that the proposed scheme compares favorably with a subspace blind channel estimator in the presence of unknown narrowband interference.**

***Index Terms*—Blind channel estimation, equalization, interference suppression, multicarrier (MC) systems.**

I. Introduction

MULTICARRIER (MC) modulation is considered a promising technique for broadband wireless networks. For coherent detection, the channel state information (CSI) of the underlying frequency-selective channel has to be estimated at the receiver. Channel estimates may be obtained by exploiting training symbols [1], [2], or by blind schemes. A popular class of blind channel estimators for MC systems are *subspace schemes* that were considered in several recent studies (e.g., [3]–[5], and references therein.) It is well known that MC systems are sensitive to various interference [2]. When the covariance of the interference can be reliably estimated at the receiver, pre-whitening can be invoked before applying subspace channel estimation. However, when there is insufficient information about the interference so that pre-whitening cannot be performed, subspace channel estimation is in general inaccurate.

In this letter, we present a new blind channel estimator for MC systems that can deal with unknown or unmodeled narrowband interference. The proposed estimator is obtained by utilizing a multichannel generalization of the minimum variance (MV) principle (e.g., [6]) along with a certain bounding technique, which effectively reduces a highly nonlinear estimation problem to a quadratic minimization.

Manuscript received November 20, 2002. The associate editor coordinating the review of this letter and approving it for publication was Prof. S. Batalama. This work was supported in part by the Army Research Office (ARO), the New Jersey Commission on Science and Technology, and by the Center for Wireless Network Security at Stevens Institute of Technology.

The author is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: hli@stevens-tech.edu).

Digital Object Identifier 10.1109/LCOMM.2003.814030

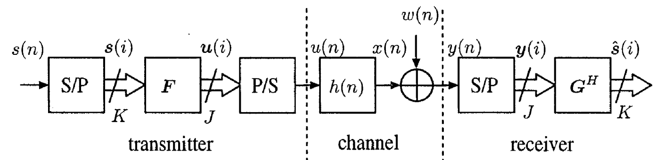

Fig. 1. Baseband discrete-time model of an MC system.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters; superscripts $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, complex conjugate and conjugate transpose, respectively; $\boldsymbol{I}_J$ is the $J\times J$ identity matrix; $\mathbf{0}_{K\times L}$ is an all-zero $K\times L$ matrix; $\mathrm{tr}\{\cdot\}$ denotes the trace; $\mathrm{vec}(\cdot)$ stacks the columns of its matrix argument on top of one another; $\otimes$ denotes the matrix Kronecker product.

II. Problem Statement

Consider a general baseband MC system shown in Fig. 1 [5]. At the transmitter, the information symbols $s(n)$ are first serial-to-parallel (S/P) converted to $K \times 1$ vectors $\boldsymbol{s}(i) \triangleq [s(iK), \ldots, s(iK+K-1)]^T$, which are linearly transformed by a $J\times K$ matrix $\boldsymbol{F}$. The data blocks $\boldsymbol{u}(i) = \boldsymbol{F}s(i)$ are next parallel-to-serial (P/S) converted and sent through a frequency-selective channel with impulse response $h(n)$. To avoid multipath-induced inter-block interference (IBI), transmission redundancy is introduced by choosing $J \geq K + L$, where L is an *upper bound* of the channel order. At the receiver, the received samples $y(n)$, corrupted by channel noise $e(n)$ and possibly interference $w(n)$ with unknown statistics, are S/P converted to $J\times 1$ vectors $\boldsymbol{y}(i) \triangleq [y(iJ), \ldots, y(iJ+J-1)]^T$, which are then equalized by the $J \times K$ matrix $\boldsymbol{G}$ to form $\hat{\boldsymbol{s}}(i) = \boldsymbol{G}^H\boldsymbol{y}(i)$.

The above system is quite versatile and encompasses the popular OFDM (orthogonal frequency-division multiplexing) as a special case [5]. Indeed, let $\boldsymbol{F} \triangleq [\bar{\boldsymbol{F}}_1^T, \bar{\boldsymbol{F}}^T]^T$ and $\boldsymbol{G} \triangleq [\mathbf{0}^T_{(J-K)\times K}, \bar{\boldsymbol{F}}^*]^T$, where $\bar{\boldsymbol{F}}$ denotes the $K \times K$ IDFT matrix with $[\bar{\boldsymbol{F}}]_{k,l} \triangleq K^{-1/2}\exp(j2\pi(k-1)(l-1)/K)$, and $\bar{\boldsymbol{F}}_1 \in \mathbb{C}^{(J-K)\times K}$ is formed from the last $J-K$ rows of $\bar{\boldsymbol{F}}$. Effectively, $\boldsymbol{F}$ performs the IDFT and CP (cyclic prefix) insertion, whereas $\boldsymbol{G}$ performs the CP removal and DFT. Then, the system reduces to OFDM. An alternative technique for IBI avoidance is to use zero-padding (ZP) [5], which amounts to choosing $\boldsymbol{F} \triangleq [\bar{\boldsymbol{F}}^T, \mathbf{0}^T_{(J-K)\times K}]^T$, where $\bar{\boldsymbol{F}}$ is usually a $K\times K$ *unitary* matrix. Typical choices for $\bar{\boldsymbol{F}}$ include the (I)DFT and Walsh-Hadamard matrices. It has been shown that under mild conditions, ZP guarantees symbol recovery irrespective of the channel zero locations [5]. On the other hand, the CP-based OFDM is sensitive to frequency-selective fading.

In this work, we consider ZP based MC systems. This is motivated by not only the advantage of guaranteed symbol recovery, but also the fact that there has been a rich literature on channel estimation for CP based OFDM systems (e.g.,[1], [3], [7], [8], and references therein). With ZP transmission, $\boldsymbol{y}(i)$ can be expressed as [4], [5]

$$\boldsymbol{y}(i) = \boldsymbol{H}\bar{\boldsymbol{F}}\boldsymbol{s}(i) + \boldsymbol{w}(i) + \boldsymbol{e}(i) \tag{1}$$

where the $J \times K$ matrix $\boldsymbol{H}$ is Toeplitz with the first column $[h(0), \ldots, h(L), \mathbf{0}_{1\times(J-L-1)}]^T$ and first row $[h(0), \mathbf{0}_{1\times(K-1)}]$, while $\boldsymbol{w}(i)$ and $\boldsymbol{e}(i)$ denote $J\times 1$ interference and noise vectors, respectively. The problem of interest is to estimate the channel coefficients $\{h(n)\}_{n=0}^{L}$ from the measurements $\{\boldsymbol{y}(i)\}$ without knowing the transmitted symbols.

III. Proposed Scheme

Equation (1) represents effectively a multiple-input multiple-output (MIMO) system with K inputs and J outputs. The mixing matrix $\boldsymbol{H}\bar{\boldsymbol{F}}$ is partially known since $\boldsymbol{H}$ has a known Toeplitz structure and $\bar{\boldsymbol{F}}$ is also known to the receiver. We can exploit this knowledge to design a bank of K FIR filters $\boldsymbol{G} \in \mathbb{C}^{J\times K}$, each passing one symbol with unit-gain, completely annihilating the other $K-1$ interfering symbols, meanwhile suppressing interference $\boldsymbol{w}(i)$ as much as possible. In particular, we design the filterbank by the following *multichannel minimum variance principle*:

$$\boldsymbol{G} = \arg\min_{\boldsymbol{G}\in\mathbb{C}^{J\times K}} \mathrm{tr}\{\boldsymbol{G}^H\boldsymbol{R}_y\boldsymbol{G}\}, \quad \text{subject to} \quad \boldsymbol{G}^H\boldsymbol{H}\bar{\boldsymbol{F}} = \boldsymbol{I}_K \tag{2}$$

where $\boldsymbol{R}_y \triangleq E\{\boldsymbol{y}(i)\boldsymbol{y}^H(i)\}$ denotes the covariance matrix. The solution is (e.g., [9, p. 283])

$$\boldsymbol{G} = \boldsymbol{R}_y^{-1}\boldsymbol{H}\bar{\boldsymbol{F}}\left(\bar{\boldsymbol{F}}^H\boldsymbol{H}\boldsymbol{R}_y^{-1}\boldsymbol{H}\bar{\boldsymbol{F}}\right)^{-1}. \tag{3}$$

Substituting (3) into the cost function in (2), the minimum average filterbank output power is

$$V_1(\boldsymbol{h}) = \mathrm{tr}\left\{\left(\bar{\boldsymbol{F}}^H\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\bar{\boldsymbol{F}}\right)^{-1}\right\} = \mathrm{tr}\left\{\left(\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\right)^{-1}\right\} \tag{4}$$

where $\boldsymbol{h} \triangleq [h(0), \ldots, h(L)]^T$ and in the second equality, we used the assumption that $\bar{\boldsymbol{F}}$ is unitary (see Section II). To find an estimate of $\boldsymbol{h}$, we could maximize $V_1(\boldsymbol{h})$ with respect to $\boldsymbol{h}$, so that $\boldsymbol{G}$ will maximally preserve the signal power. Doing so is computationally involved and suffers local convergence due to the highly nonlinear nature of $V_1(\boldsymbol{h})$. Instead, we propose to maximize an (asymptotically tight) lower bound of $V_1(\boldsymbol{h})$. Specifically, by the Schwartz inequality, we have

$$\begin{aligned} K^2 &= \mathrm{tr}^2(\boldsymbol{I}_K) \\ &= \mathrm{tr}^2\left\{\left(\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\right)^{-\frac{1}{2}}\left(\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\right)^{\frac{1}{2}}\right\} \\ &\le \mathrm{tr}\left\{\left(\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\right)^{-1}\right\}\mathrm{tr}\left\{\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\right\} \end{aligned} \tag{5}$$

with equality holds if and only if $\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}$ is a (scaled) identity. We show in the Appendix that this condition is satisfied asymptotically (for high SNR). Hence, maximizing $V_1(\boldsymbol{h})$ is asymptotically equivalent to

$$\hat{\boldsymbol{h}} = \arg\min_{\boldsymbol{h}\in\mathbb{C}^{K\times 1}} \mathrm{tr}\left\{\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\right\}. \tag{6}$$

To solve for $\hat{\boldsymbol{h}}$, we rewrite the cost function $V_2(\boldsymbol{h}) \triangleq \mathrm{tr}\{\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}\}$ as follows:

$$\begin{aligned} V_2(\boldsymbol{h}) &= \mathrm{vec}^T(\boldsymbol{H}^*)\left[\boldsymbol{I}_K \otimes \boldsymbol{R}_y^{-1}\right]\mathrm{vec}(\boldsymbol{H}) \\ &= \mathrm{vec}^H(\boldsymbol{H})(\boldsymbol{I}_K\otimes\boldsymbol{C}^H)(\boldsymbol{I}_K\otimes\boldsymbol{C})\mathrm{vec}(\boldsymbol{H}) \\ &= \boldsymbol{h}^H\boldsymbol{S}^H(\boldsymbol{I}_K\otimes\boldsymbol{C}^H)(\boldsymbol{I}_K\otimes\boldsymbol{C})\boldsymbol{S}\boldsymbol{h} \triangleq \boldsymbol{h}^H\boldsymbol{\Phi}^H\boldsymbol{\Phi}\boldsymbol{h} \end{aligned} \tag{7}$$

where $\boldsymbol{C} \in \mathbb{C}^{J\times J}$ denotes the triangular Cholesky factor of $\boldsymbol{R}_y^{-1}$, i.e., $\boldsymbol{R}_y^{-1} = \boldsymbol{C}^H\boldsymbol{C}$, and $\boldsymbol{\Phi} \triangleq (\boldsymbol{I}_K\otimes\boldsymbol{C})\boldsymbol{S}$. The second equality of (7) is due to the identity $\mathrm{tr}\{\boldsymbol{ABCD}\} = \mathrm{vec}^T(\boldsymbol{A}^T)(\boldsymbol{D}^T\otimes\boldsymbol{B})\mathrm{vec}(\boldsymbol{C})$ for any matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ with compatible dimensions; and the third equality is due to a linear dependence of $\boldsymbol{H}$ on $\boldsymbol{h}$: $\mathrm{vec}(\boldsymbol{H}) = \boldsymbol{S}\boldsymbol{h}$. The $JK\times(L+1)$ selection matrix $\boldsymbol{S}$ is given by $\boldsymbol{S} \triangleq [\boldsymbol{S}_1^T, \ldots, \boldsymbol{S}_K^T]^T$, where $\boldsymbol{S}_k^T \triangleq [\mathbf{0}_{(L+1)\times(k-1)}, \boldsymbol{I}_{L+1}, \mathbf{0}_{(L+1)\times(K-k)}]$, $k = 1, \ldots, K$. It should be noted that due to the sparse structure of $\boldsymbol{S}$, $\boldsymbol{\Phi}$ can be readily obtained from $\boldsymbol{C}$ with no additional computations. Specifically, let us decompose $\boldsymbol{\Phi}$ into K equal-size blocks: $\boldsymbol{\Phi} = [\boldsymbol{\Phi}_1^T, \ldots, \boldsymbol{\Phi}_K^T]^T$. It is easy to verify that $\boldsymbol{\Phi}_k = \boldsymbol{C}(:, k : k+L)$, where the Matlab notation $\boldsymbol{C}(:, k : k+L)$ denotes a $J\times(L+1)$ matrix formed from columns $k, k+1, \ldots$, and $k+L$ of $\boldsymbol{C}$.

The solution to the quadratic minimization (6) under constraint $\|\hat{\boldsymbol{h}}\| = 1$ is the eigenvector of $\boldsymbol{\Phi}^H\boldsymbol{\Phi}$ associated with the smallest eigenvalue. Note that for implementation, $\boldsymbol{R}_y$ has to be replaced by some covariance matrix estimate, e.g., the sample covariance matrix $\hat{\boldsymbol{R}}_y = I^{-1}\sum_{i=0}^{I-1}\boldsymbol{y}(i)\boldsymbol{y}^H(i)$ or some adaptive estimate of $\boldsymbol{R}_y$. It can be shown (e.g., [6, Proposition 2]) that $\hat{\boldsymbol{h}}$ converges to the true channel $\boldsymbol{h}$ (up to a scalar factor) as the interference and noise vanish. For finite SNR and in the presence of interference, we evaluate the accuracy of $\hat{\boldsymbol{h}}$ via simulations in Section IV. Finally, like all other blind schemes, the channel estimate $\hat{\boldsymbol{h}}$ (6) has a scalar ambiguity, which can be resolved either by differential coding or by transmitting a *few* pilot symbols.

IV. Simulation Results and Discussions

We compare here the proposed and the subspace blind channel estimators. The system utilizes the IDFT transform [recall that the proposed scheme is invariant to the choice of $\bar{\boldsymbol{F}}$ as long as it is unitary; see (4)] and a BPSK constellation with $K = 48$. The channel is a four-tap $(L = 3)$ FIR channel that is identical to the one in Example 2 of [4]. Two narrowband interfering signals are added with various values of signal-to-interference ratio (SIR). Both estimators use a total of $I = J = 51$ blocks of data for channel estimation. We first plot the cost function $V_1(\boldsymbol{h})$ and the lower bound versus the SNR in Fig. 2. It is seen that the lower bound is indeed asymptotically tight. Fig. 3 depicts the normalized mean-squared errors (MSE) of the channel estimates versus SNR and SIR. In the absence of interference (i.e., $\mathrm{SIR} = \infty$), the subspace estimator outperforms the proposed scheme slightly. However, even with fairly weak interference (i.e., $\mathrm{SIR} = 10\,\mathrm{dB}$), the subspace estimator degrades significantly.

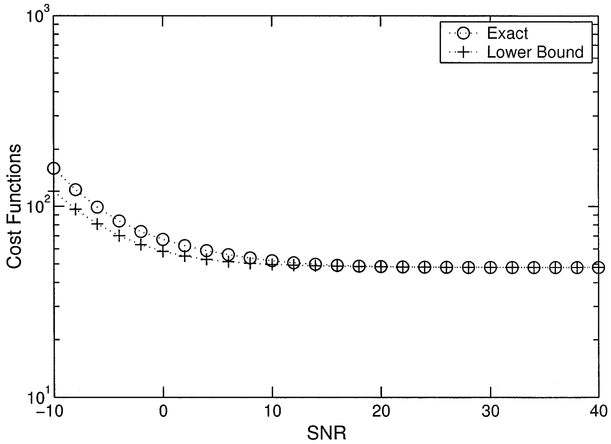

Fig. 2. Cost function $V_1(\boldsymbol{h})$ and its lower bound.

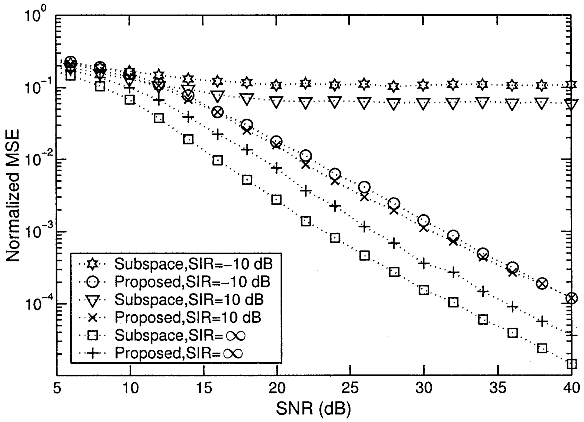

Fig. 3. Normalized MSE of the proposed and subspace blind channel estimates versus SNR and SIR.

The complexity of the proposed scheme is $O(IJ^2)+O(J^3)+O(JK(L+1)^2)+O((L+1)^3)$, which comes from calculating the sample covariance matrix $\hat{\boldsymbol{R}}_y$, the Cholesky decomposition $\boldsymbol{R}_y^{-1} = \boldsymbol{C}^H\boldsymbol{C}$, the matrix product $\boldsymbol{\Phi}^H\boldsymbol{\Phi}$, and the eigendecomposition of $\boldsymbol{\Phi}^H\boldsymbol{\Phi}$, respectively. In general, the last two terms are negligible since L is much smaller than I, J or K in typical applications. Overall, the complexity of the proposed scheme is comparable to that of the subspace estimator, with our scheme being slightly simpler.

Finally, we remark that while only ZP-based transmissions are considered here, the proposed scheme can be extended to CP-based systems for joint channel estimation and narrowband interference mitigation.

APPENDIX
ASYMOTOTIC STRUCTURE OF $\boldsymbol{H}^H\boldsymbol{R}_y^{-1}\boldsymbol{H}$

Proposition 1: Suppose that i) the channel noise is white with $E\{\boldsymbol{e}(i)\boldsymbol{e}^H(i)\} = \sigma^2\boldsymbol{I}_J$, where σ^2 denotes the noise variance; ii) the interference covariance matrix $\boldsymbol{R}_w \triangleq E\{\boldsymbol{w}(i)\boldsymbol{w}^H(i)\}$ has a low rank decomposition $\boldsymbol{R}_w = \boldsymbol{\Gamma}\boldsymbol{\Gamma}^H$, where $\boldsymbol{\Gamma}$ is a $J \times M$ matrix with $M < J - K$; and iii) $\tilde{\boldsymbol{H}} \triangleq [\boldsymbol{H}, \boldsymbol{\Gamma}]$ is full column-rank. Then,

$$\lim_{\sigma^2\to\infty} \boldsymbol{H}^H\boldsymbol{R}_y^{-1}(\sigma^2)\boldsymbol{H} = \boldsymbol{I}_K. \tag{8}$$

Proof: Assume that the information symbols $s(n)$ are independent with unit average energy. Then, we can write $\boldsymbol{R}_y(\sigma^2) = \boldsymbol{H}\boldsymbol{H}^H + \boldsymbol{R}_w + \sigma^2\boldsymbol{I}_J = \boldsymbol{H}\boldsymbol{H}^H + \boldsymbol{\Gamma}\boldsymbol{\Gamma}^H + \sigma^2\boldsymbol{I}_J = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^H + \sigma^2\boldsymbol{I}_J$. Let the eigenvalue decomposition (EVD) of $\boldsymbol{R}_y$ be expressed as

$$\boldsymbol{R}_y(\sigma^2) = [\boldsymbol{U}_s,\ \boldsymbol{U}_n] \begin{bmatrix} \boldsymbol{\Lambda}_s + \sigma^2\boldsymbol{I}_{K+M} & \\ & \sigma^2\boldsymbol{I}_{J-K-M} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_s^H \\ \boldsymbol{U}_n^H \end{bmatrix}$$

Accordingly, the EVD of $\tilde{\boldsymbol{H}}\tilde{\boldsymbol{H}}^H$ can be written as $\boldsymbol{U}_s\boldsymbol{\Lambda}_s\boldsymbol{U}_s^H$ [9]. It is trivial to show $\boldsymbol{U}_s^H\tilde{\boldsymbol{H}}$ is square and full rank. Hence

$$\tilde{\boldsymbol{H}}^H\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{U}_s^H\tilde{\boldsymbol{H}} = \boldsymbol{I}_{K+M}. \tag{9}$$

A Taylor expansion of $\boldsymbol{R}_y^{-1}(\sigma^2)$ at high SNR is [6, eq. (35)]

$$\boldsymbol{R}_y^{-1}(\sigma^2) = \sigma^{-2}\boldsymbol{U}_n\boldsymbol{U}_n^H + \boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{U}_s^H + O(\sigma^2). \tag{10}$$

Using the above Taylor expansion, along with (9) and the observation $\tilde{\boldsymbol{H}}^H\boldsymbol{U}_n = \boldsymbol{0}$, we have

$$\tilde{\boldsymbol{H}}^H\boldsymbol{R}_y^{-1}\tilde{\boldsymbol{H}} = \boldsymbol{I}_{K+M} + O(\sigma^2) \xrightarrow{\sigma^2\to 0} \boldsymbol{I}_{K+M}. \tag{11}$$

Equation (8) follows immediately from (11). ■

Remark: The low-rank assumption implies that the interference occupies only a portion (i.e., $M < J - K$ out of K subcarriers) of the overall bandwidth. It is interesting to note that the ZP-induced transmission redundancy (recall $J \geq K + L$), initially intended for IBI removal, also helps with narrowband interference cancellation. Proposition 1 indicates that the approximation incurred in Section III, i.e., replacing the highly nonlinear $V_1(\boldsymbol{h})$ by the quadratic $V_2(\boldsymbol{h})$, might not be a good approximation if $M < J - K$ is not satisfied. We have also observed empirically in our simulations (not shown due to space limitation) that when more than $J - K$ tones are corrupted, the performance of the proposed scheme degrades considerably. Interference on more subcarriers can be handled by increasing J, or process several consecutive data blocks $y(i)$ simultaneously.

REFERENCES

[1] Y. G. Li, L. J. Cimini Jr., and N. R. Sollengerger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, July 1998.

[2] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2451–2464, Oct. 2002.

[3] C. Li and S. Roy, "Subspace based blind channel estimation for OFDM by exploiting virtual carriers," in *Proc. IEEE GLOBECOM*, San Antonio, TX, Nov. 2001, pp. 295–299.

[4] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter precoders and equalizers part II: blind channel estimation, synchronization direct equalization," *IEEE Trans. Signal Processing*, vol. 47, pp. 2007–2032, July 1999.

[5] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.

[6] M. K. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receiver," *IEEE Trans. Signal Processing*, vol. 46, pp. 3014–3022, Nov. 1998.

[7] J. Li, G. Liu, and G. B. Giannakis, "Carrier frequency offset estimation for OFDM based WLANs," *IEEE Signal Processing Lett.*, vol. 8, pp. 80–82, Mar. 2001.

[8] S. Zhou and G. B. Giannakis, "Space-time coding with maximum diversity gains over frequency-selective fading channels," *IEEE Signal Processing Lett.*, vol. 8, pp. 269–272, Oct. 2001.

[9] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice Hall, 1997.