Semi-Blind Most Significant Tap Detection for Sparse Channel Estimation of OFDM Systems

Feng Wan, *Student Member, IEEE*, W.-P. Zhu, *Senior Member, IEEE*, and M. N. S. Swamy, *Fellow, IEEE*

***Abstract*—In this paper, a very efficient semi-blind approach for the detection of most significant taps (MSTs) in sparse orthogonal frequency-division multiplexing (OFDM) channel estimation is developed. The least square (LS) estimation problem of sparse OFDM channels is first formulated, showing that the key to sparse channel estimation lies in the detection of the MSTs. An in-depth study of the second-order statistics of the signal received through a noise-free sparse OFDM channel reveals the sparsity and other properties of the correlation functions of the received signal. These properties lead to a direct relationship between the positions of the MSTs of the sparse channel and the most significant lags of the correlation functions, which is then used in conjunction with a pilot-assisted LS estimation to detect the MSTs in a semi-blind fashion. It os also shown that the new MST detection algorithm can be extended for the estimation of multiple-input–multiple-output (MIMO)–OFDM channels. A number of computer-simulation-based experiments for various sparse channels are carried out to confirm the effectiveness of the proposed semi-blind approach.**

***Index Terms*—Most significant tap (MST) detection, multiple-input–multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), semi-blind, sparse-channel estimation.**

NOTATIONS

Throughout this paper, we adopt the following notations.

$\dagger$	Pseudoinverse.
$\otimes$	Kronecker delta.
T	Transpose.
H	complex-conjugate transpose;
$\circledast$	Circular convolution;
$\|\|\ \|\|_F$	Frobenius norm;
$\delta()$	Delta function;
vec()	A stacking of the columns of the involved matrix into a vector.

I. INTRODUCTION

IT IS WELL known that channel estimation is of crucial importance to orthogonal frequency-division multiplexing (OFDM) and multiple-input-multiple-output (MIMO)–OFDM systems [1]. Broadly speaking, channel-estimation techniques for OFDM or MIMO–OFDM systems can be categorized into three classes: training-based, blind, and semi-blind methods. First, training-based methods, such as the least square (LS), employ known training signals to render an accurate channel estimation and thus require a consumption of spectral resources [2]. Blind channel-estimation algorithms, such as those proposed in[3]–[5], which exploit the second-order stationary statistics, correlative coding, and other properties normally have a better spectral efficiency. With the idea of both the training-based and blind algorithms, semi-blind channel-estimation techniques can potentially enhance the quality of channel estimation [6]–[8].

A wireless channel can often be modeled as a sparse channel in which the delay spread could be very large, but the number of significant paths is normally very small [9]–[18]. Broadly speaking, there are two kinds of approaches for the sparse channel estimation. The first one estimates the complex amplitude and the delay of each path based on a nonsampling spaced parametrical channel modeling [9]. The second kind is based on the sparsity assumption of the equivalent discrete-time channel [11]–[17], [19] in which only a few taps in the long-tapped delay line are considered most significant. By exploiting the sparse structure of the channel, some improved channel-estimation algorithms have been developed for OFDM systems [10]–[12], [17], [20], [21] and code division multiple access systems [15], [16]. In this paper, we are concerned with the development of a very efficient sparse-channel-estimation approach for OFDM and MIMO–OFDM systems that requires only a few OFDM symbols. As the nonsampling-spaced sparse channel estimation requires a large number of OFDM symbols for the estimation of the delay subspace, we focus only on the sampling-spaced approach.

It should be mentioned that almost all of the sampling-spaced sparse-channel-estimation methods in the literature utilize a training sequence and follow two steps: 1) detect the positions of the most significant taps (MSTs), which are also referred to as the nonzero taps in some of the literature, and (2) obtain an improved channel estimate by exploiting the position of the MSTs. The key to these methods is the MST detection. Several techniques based on initial LS estimation have been developed for MST detection [10]–[12]. In these methods, an unstructured channel, i.e., a channel presumably having all nonzero taps, is considered in the initial LS estimation. Based on the preliminary estimate, the MST detection can simply be conducted by choosing a number of the largest amplitude channel taps [10]. To improve the detection accuracy, a generalized Akaike information criterion is utilized to estimate the MSTs in an iterative fashion [11], which is then simplified to a noniterative scheme [12]. The matching-pursuit (MP)-based MST detection method has also extensively been investigated by many researchers [13]–[16]. In the MP method, the MST is detected by sequentially selecting the column in a mixture

Manuscript received March 26, 2008; revised September 26, 2008. First published June 02, 2009; current version published March 10, 2010. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and in part by the Regroupement Stratégique en Microsystémes du Quebéc (ReSMiQ). This paper was recommended by Associate Editor W. Utschick.

The authors are with the Center for Signal Processing and Communications, Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: f_wan@ece.concordia.ca; weiping@ece.concordia.ca; swamy@ece.concordia.ca).

Digital Object Identifier 10.1109/TCSI.2009.2023765

matrix of the pilot signal that matches best the residual vector of the received signal. Another MST detection, which employs an active tap-detection criterion involving the correlation between the training pilots and the received signal, has been proposed in [19].

Once the MST information has been acquired, a refined channel estimate can be obtained in the second step. This is usually implemented with the structured version of the training-based approaches such as the LS method [11]–[13], [15] and the LMS method [19]. It has been shown that the performance of these sparse-channel-estimation methods is much better than that of the unstructured channel-estimation alternatives.

The common problem of the aforementioned sparse-channel-estimation methods is that a large number of pilots is needed in order to render an accurate MST detection and channel estimation. To increase the spectral efficiency, the available information of user data could be applied to both the MST detection and the channel estimation. Unfortunately, very little work on blind MST detection is found in the existing literature. To the best knowledge of the authors, there is only one cyclic-prefix (CP)-based blind method dealing with the MST detection of the sparse channel estimation for OFDM systems [17]. However, this detection scheme needs a large number of OFDM symbols as well as a large CP length in order to obtain precise MST positions.

In this paper, our main objective is to develop an efficient semi-blind MST detection approach for sparse OFDM channel estimation. First, the analysis of the second-order statistics of the received signal passing through a sparse channel is conducted for OFDM systems, leading to a relationship between the positions of the MSTs and the most significant lags (MSLs) of the correlation functions of the received signal. Based on this relationship, an efficient MST detection algorithm, which requires only a small number of OFDM symbols and a very small number of pilots, is proposed. Then, by using the acquired MST positions, an LS estimation of the sparse channel vector with respect to the MSTs are derived for OFDM systems. Furthermore, it is shown that the semi-blind MST detection algorithm can readily be extended to the sparse channel estimation of MIMO–OFDM systems.

The rest of this paper is organized as follows. Section II formulates the sparse LS channel-estimation problem. Section III presents a semi-blind MST detection approach for the estimation of sparse OFDM channels, including the analysis of the second-order statistics of the received signal through a sparse channel and the development of a new MST detection algorithm. Section IV extends the new semi-blind MST detection algorithm to the MIMO–OFDM channel estimation. Section V comprises a number of experiments validating the proposed approach and showing the significant advantage of the sparse solution over the regular LS method. Finally, Section VI concludes this paper by highlighting some of the contributions presented.

II. Problem Formulation

A. OFDM Signal Model

For an OFDM system, the transmitted mth OFDM symbol can be written as a vector of the frequency-domain signals, namely

$$\mathbf{X}(m) \triangleq [X(m,0), X(m,1), \ldots, X(m,K-1)]^{\mathrm{T}}$$

where K denotes the number of subcarriers. The inverse discrete Fourier transform (IDFT) processing gives the time-domain OFDM signal, denoted as

$$\mathbf{x}_T(m) \triangleq [x(m,0), x(m,1), \ldots, x(m,K-1)]^{\mathrm{T}}.$$

After adding a CP, each OFDM symbol is then sent out by the transmit antenna. At the receiver, after removing the CP, the received signal can be described as

$$\mathbf{y}_T(m) \triangleq [y(m,0), y(m,1), \ldots, y(m,K-1)]^{\mathrm{T}}. \quad (1)$$

Then, the received frequency-domain signal after the DFT processing is given by

$$\mathbf{Y}(m) \triangleq [Y(m,0), Y(m,1), \ldots, Y(m,K-1)]^{\mathrm{T}}.$$

Considering that OFDM systems are designed for broadband wireless communications, the signal bandwidth is always larger than the coherence bandwidth, implying that the channel is frequency selective. Most of the existing channel-estimation methods work on the equivalent discrete-time channel, i.e., the sampled version of the continuous-time channel response[22]. Thus, the discrete-time channel can be modeled by an L-tap finite-impulse response (FIR) filter. Assuming that the channel is constant during a number of consecutive OFDM symbols, the channel can be described by

$$h(l) \in C, \qquad (l = 0, 1, \ldots, L-1).$$

If the length of the CP is not less than the channel length L, the time-domain signal model for the frequency-selective fading channel is given by

$$y(m,n) = h(n) \circledast x(m,n) + v(m,n), \qquad m \in \{0, \ldots, g-1\} \quad (2)$$

where g is the number of OFDM symbols within which the channel remains unchanged and $v(m,n) \in C$ is a spatiotemporally uncorrelated noise with zero mean and varianceσ_v^2.

It is known that a wireless channel can often be modeled as a sparse channel that contains many zero taps in the uniform delay line [11]–[17], [19]. In this case, the channel with respect to the dth$(d = 0, 1, \ldots, D-1)$ MST (namely, the nonzero tap) can be expressed as

$$z(d) = h(l_d) \quad (3)$$

where $l_d\,(d = 0, 1, \ldots, D-1)$ are integers with $0 = l_0 < l_1 < \cdots < l_D - 1$. To distinguish from $h(l)$, $z(d)$ is referred to as the effective channel throughout this paper.

B. Sparse LS Estimation

Prior to the development of the sparse LS algorithm for the estimation of the effective channel, we first introduce the training-based LS channel-estimation algorithm for OFDM systems. Assume that the K_p subcarriers, say, from $i_{\mathrm{pilot}1}$ to $i_{\mathrm{pilot}K_p}$, of each OFDM symbol carry the pilot signal. The transmitted and the received pilot vectors can be defined as

$$\mathbf{X}_{\mathrm{pilot}}(m) \triangleq \left[X(m, i_{\mathrm{pilot}1}), \ldots, X(m, i_{\mathrm{pilot}K_p})\right]^{\mathrm{T}},$$
$$\mathbf{Y}_{\mathrm{pilot}}(m) \triangleq \left[Y(m, i_{\mathrm{pilot}1}), \ldots, Y(m, i_{\mathrm{pilot}K_p})\right]^{\mathrm{T}}.$$

It should be noted that the pilot signal might not be located at the same position in each OFDM symbol. Let $\mathbf{F}_1$ be a $K \times L$ matrix formed by the first L columns of a $K \times K$ DFT matrix $\mathbf{F}_0$. For the mth OFDM symbol, one can form a $K_p \times L$ matrix, for example, $\mathbf{F}(m)$, by taking only the rows of $\mathbf{F}_1$ associated with the K_p pilot subcarriers. It was shown in [2] that

$$\mathbf{Y}_{\text{pilot}}(m) = \mathbf{X}_{\text{pilot}-\text{diag}}(m)\mathbf{F}(m)\mathbf{h} + \boldsymbol{\xi}_{\text{pilot}}(m) \quad (4)$$

where $\mathbf{X}_{\text{pilot}-\text{diag}}(m) \triangleq \text{diag}(\mathbf{X}_{\text{pilot}}(m))$, $\mathbf{h} \triangleq [h(0), \ldots, h(L-1)]^T$, and $\boldsymbol{\xi}_{\text{pilot}}(m)$ represent the frequency-domain noise corresponding to $v_{i_R}(m,n)$ in (2). From (4), the received frequency-domain pilot signal with respect to g OFDM symbols can be obtained as

$$\mathbf{Y}_{\text{pilot}} = \mathbf{A}\mathbf{h} + \boldsymbol{\xi}_{\text{pilot}} \quad (5)$$

where

$$\mathbf{Y}_{\text{pilot}} \triangleq \left[\mathbf{Y}^{\mathrm{T}}_{\text{pilot}}(0), \ldots, \mathbf{Y}^{\mathrm{T}}_{\text{pilot}}(g-1)\right]^{\mathrm{T}}$$

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{X}_{\text{pilot}-\text{diag}}(0)\mathbf{F}(0) \\ \vdots \\ \mathbf{X}_{\text{pilot}-\text{diag}}(g-1)\mathbf{F}(g-1) \end{bmatrix}$$

$$\boldsymbol{\xi}_{i_R,\text{pilot}} \triangleq \left[\boldsymbol{\xi}^{\mathrm{T}}_{i_R,\text{pilot}}(0), \ldots, \boldsymbol{\xi}^{\mathrm{T}}_{i_R,\text{pilot}}(g-1)\right]^{\mathrm{T}}.$$

From (5), one can obtain an LS minimization problem, i.e.,

$$\min_{\hat{\mathbf{h}}} \|\mathbf{Y}_{\text{pilot}} - \mathbf{A}\hat{\mathbf{h}}\|^2 \quad (6)$$

whose solution is given by

$$\hat{\mathbf{h}} = \tilde{\mathbf{A}}^{\dagger}\mathbf{Y}_{\text{pilot}}. \quad (7)$$

In the aforementioned LS channel-estimation method as well as in many existing OFDM channel-estimation methods such as those in[23] and [24], the sparse case of the wireless channel has not been taken into consideration. Thus, the channel estimate obtained is not efficient when the channel is sparse. In the following, we would like to propose a sparse LS solution for OFDM channel estimation.

Using (3) and assuming that the MSTs are correctly estimated, (4) can be rewritten as

$$\mathbf{Y}_{\text{pilot}}(m) = \mathbf{X}_{\text{pilot}-\text{diag}}(m)\bar{\mathbf{F}}(m)\mathbf{z} + \boldsymbol{\xi}_{\text{pilot}}(m) \quad (8)$$

where $\mathbf{z} = [z(0), \ldots, z(D-1)]^{\mathrm{T}}$ and $\bar{\mathbf{F}}(m)$ is a $K_p \times D$ matrix whose dth column is the l_dth column of $\mathbf{F}(m)$, $(d = 0, 1, \ldots, D-1)$. Accordingly, (5) can be re-expressed as

$$\mathbf{Y}_{\text{pilot}} = \bar{\mathbf{A}}\mathbf{z} + \bar{\boldsymbol{\xi}}_{\text{pilot}} \quad (9)$$

where

$$\bar{\mathbf{A}} \triangleq \begin{bmatrix} \mathbf{X}_{\text{pilot}-\text{diag}}(0)\bar{\mathbf{F}}(0) \\ \vdots \\ \mathbf{X}_{\text{pilot}-\text{diag}}(g-1)\bar{\mathbf{F}}(g-1) \end{bmatrix}. \quad (10)$$

From (9), an LS estimate of the sparse channel with respect to the MSTs $l_d(d = 0, 1, \ldots, D-1)$ can be obtained as

$$\hat{\mathbf{z}} = (\bar{\mathbf{A}})^{\dagger}\mathbf{Y}_{\text{pilot}}. \quad (11)$$

Obviously, the key to the structured LS estimation of the effective channel lies in the detection of the MSTs. In the following section, we first analyze the second-order statistics of the signal received through the sparse channel and then propose a novel semi-blind MST detection algorithm.

III. Semi-Blind MST Detection Algorithm

A. *Second-Order Statistics of the Received Signal Through Sparse Channel*

It is well known that the correlation function of the received signal vector $y(m,n)$ plays a crucial role in blind or semi-blind channel estimation[5], [25], [26], which can be, in general, defined as

$$r(l) \triangleq \mathrm{E}\{y(m,n)y^*(m,n-l)\}, \quad l = 0, 1, \ldots, P. \quad (12)$$

In this section, we would like to express $r(l)$ in terms of the effective sparse channel $z(d), (d = 0, 1, \ldots, D-1)$, in the absence of noise, and show that $r(l)$ has only a few MSLs, i.e., most of the values of $r(l)$ are zero, due to the sparse feature of the channel.

Using (2) and (3) in (12), we obtain

$$r(l) = \mathbf{z}_A \mathbf{R}_{x,D}(l)\mathbf{z}_A^{\mathrm{H}} \quad (13)$$

where

$$\mathbf{z}_A \triangleq [z(0) \quad z(1) \quad \cdots \quad z(D-1)] \quad (14)$$

$$\mathbf{R}_{x,D}(l) \triangleq \mathrm{E}\left\{\begin{bmatrix} x(n) \\ x(n-l_1) \\ \vdots \\ x(n-l_{D-1}) \end{bmatrix}\begin{bmatrix} x(n) \\ x(n-l_1) \\ \vdots \\ x(n-l_{D-1}) \end{bmatrix}^{\mathrm{H}}\right\}. \quad (15)$$

Here, the index m has been omitted for the sake of notational convenience. Clearly, the value of $r(l)$ mainly depends on $\mathbf{R}_{x,D}(l)$. Due to the fact that $\mathrm{E}\{x(n-i)x^{\mathrm{H}}(n-j)\} = \sigma_x^2\delta(i-j)$, we can rewrite (15) as

$$\mathbf{R}_{x,D}(l) = \begin{bmatrix} \delta(l) & \delta(l+l_1) & \cdots & \delta(l+l_{D-1}) \\ \delta(l-l_1) & \delta(l) & \cdots & \delta(l+l_{D-1}-l_1) \\ \vdots & \vdots & \ddots & \vdots \\ \delta(l-l_{D-1}) & \delta(l-l_{D-1}+l_1) & \cdots & \delta(l) \end{bmatrix}. \quad (16)$$

Obviously, $\mathbf{R}_{x,D}(l)$ is a lower triangular matrix for $l \in [0, P]$. Moreover, the nonzero elements of $\mathbf{R}_{x,D}(l)$ occur only when $l = l_i - l_j, (i,j = 0, 1, \ldots, D-1; i \geq j)$. Since the channel is sparse, i.e., $D \ll L \leq P$, there is only a small number of choices of l which make $\mathbf{R}_{x,D}(l)$ a nonzero matrix. In other words, $\mathbf{R}_{x,D}(l)$ is a zero matrix for most of the values of l ranging from 0 to P. It is clear from (13) and (16) that as long as $\mathbf{R}_{x,D}(l)$ is a zero matrix, $r(l)$ is zero. By substituting (14) and (16) into (13), $r(l)$ can be expressed in terms of the effective channel $z(d)$ as

$$r(l) = \sum_{i=0}^{D-1}\sum_{j=0}^{D-1} \delta(l - l_i + l_j)z(i)z^*(j) \quad (17)$$

TABLE I
EXPRESSION OF CORRELATION FUNCTION $r(l)$ FOR $D = 4$

Case D4.1:	l	l_1	l_2			
$l_1 : l_2 : l_3 = 1 : 2 : 3$	$r(l)$	$r_{1,0} + r_{2,1} + r_{3,2}$	$r_{2,0} + r_{3,1}$			
Case D4.2:	l	l_1	l_2	$l_3 - l_1$		
$l_1 : l_2 : l_3 = 1 : 2 : 4$	$r(l)$	$r_{1,0} + r_{2,1}$	$r_{2,0} + r_{3,2}$	$r_{3,1}$		
Case D4.3:	l	l_1	l_2	$l_2 - l_1$		
$l_1 : l_2 : l_3 = 2 : 3 : 4$	$r(l)$	$r_{1,0} + r_{3,1}$	$r_{2,0}$	$r_{2,1} + r_{3,2}$		
Case D4.4: $l_3 = l_1 + l_2$	l	l_1	l_2	$l_2 - l_1$		
$l_2 \neq 2l_1$, $l_1 : l_2 \neq 2 : 3$	$r(l)$	$r_{1,0} + r_{3,2}$	$r_{2,0} + r_{3,1}$	$r_{2,1}$		
Case D4.5: $l_3 \neq l_1 + l_2$	l	l_1	l_2	$l_3 - l_1$	$l_3 - l_2$	
$l_2 = 2l_1$, $l_3 \neq 2l_2$	$r(l)$	$r_{1,0} + r_{2,1}$	$r_{2,0}$	$r_{3,1}$	$r_{3,2}$	
Case D4.6: $l_3 \neq l_1 + l_2$	l	l_1	l_2	$l_2 - l_1$	$l_3 - l_1$	
$l_2 \neq 2l_1$, $2l_2 = l_3 + l_1$	$r(l)$	$r_{1,0}$	$r_{2,0}$	$r_{2,1} + r_{3,2}$	$r_{3,1}$	
Case D4.7: $l_1 : l_3 = 1 : 2$	l	l_1	l_2	$l_2 - l_1$	$l_3 - l_2$	
$l_1 : l_2 \neq 2 : 3$	$r(l)$	$r_{1,0} + r_{3,1}$	$r_{2,0}$	$r_{2,1}$	$r_{3,2}$	
Case D4.8: $l_3 \neq l_1 + l_2$, $l_2 \neq 2l_1$	l	l_1	l_2	$l_2 - l_1$	$l_3 - l_1$	$l_3 - l_2$
$l_1 : l_2 \neq 2 : 3$, $2l_2 \neq l_3 + l_1$	$r(l)$	$r_{1,0}$	$r_{2,0}$	$r_{2,1}$	$r_{3,1}$	$r_{3,2}$

In obtaining (17), a unit signal variance, i.e., $\sigma_x^2 = 1$, has been assumed without loss of generality. In the following, we disclose the relationship between the MSTs with the MSLs.

Let us consider the simplest case when $D = 2$. In this case, there are only two nonzero effective channel elements $z(0)$ and $z(1)$, corresponding to $h(0)$ and $h(l_1)$, respectively. From (17), one can find that there are only two nonzero values of $r(l)$, i.e.,

$$r(0) = z(0)z^*(0) + z(1)z^*(1) \tag{18}$$

$$r(l_1) = z(1)z^*(0) \tag{19}$$

which means that the MSL positions of $r(l)$ are$l = 0, l_1$.

When $D = 3$, it is clear from (17) that $r(l)$ has different expressions depending on the relationship between l_1 and l_2. Using (17), one can obtain the following results. *Case D3.1*: if $l_2 = 2l_1$

$$r(l) = \begin{cases} \sum_{i=0}^{2} z(i)z^{\mathrm{H}}(i), & \text{if } l = 0 \\ z(1)z^*(0) + z(2)z^*(1), & \text{if } l = l_1 \\ z(2)z^*(0), & \text{if } l = l_2 \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

Case D3.2: if $l_2 \neq 2l_1$

$$r(l) = \begin{cases} \sum_{i=0}^{2} z(i)z^*(i), & \text{if } l = 0 \\ z(1)z^*(0), & \text{if } l = l_1 \\ z(2)z^*(1), & \text{if } l = l_2 - l_1 \\ z(2)z^*(0), & \text{if } l = l_2 \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

It is seen that the main difference between the two cases *D3.1* and *D3.2* lies in the number of MSLs in addition to the expression of $r(l)$. The two common MSLs of the two cases are $r(0)$ and $r(l_2)$. Other MSLs depend actually on the relationship between l_1 andl_2. If l_1 is not identical to $l_2 - l_1$, for any value of l, $\delta(l - l_1)$ and $\delta(l - l_2 + l_1)$ in (16) cannot be unity at the same time, leading to two different MSLs, i.e., $r(l_1) = z(1)z^*(0)$ and $r(l_2 - l_1) = z(2)z^*(1)$. However, if l_1 equals $l_2 - l_1$, both $\delta(l - l_1)$ and $\delta(l - l_2 + l_1)$ could be unity simultaneously. Then, the two presumable MSLs overlap, yielding the only MSL, i.e., $r(l = l_1 = l_2 - l_1) = z(1)z^*(0) + z(2)z^*(1)$.

In the case of $D = 4$, one can easily obtain $r(l)$ with respect to the first and the last MSLs as

$$r(0) = \sum_{i=0}^{3} z(i)z^*(i) \tag{22}$$

$$r(l_3) = z(3)z^*(0) \tag{23}$$

which is similar to the cases of $D = 2$ and $D = 3$. All other MSLs can be determined from the relationship among l_1, l_2, $l_2 - l_1$, $l_3 - l_1$, and $l_3 - l_2$. As some of these values could be identical as discussed for the case of $D = 3$, the number of the MSLs of $r(l)$ can be different. As such, the position of the MSLs and the expression of $r(l)$ are dependent on the values of l_1, l_2, and l_3, and their differences. Using (17), one can have a total of eight possible cases as summarized in Table I, where $r_{i,j} \triangleq z(i)z^*(j)$. Note that the correlation matrices with respect to the first and the last MSLs, which are common to all the eight cases, are not included in Table I for notational simplicity. Evidently, *Case D4.8* indicates a possibility of having a maximum of seven MSLs of $r(l)$ including the first and the last ones. When some of the values of l_1, l_2, $l_2 - l_1$, $l_3 - l_1$, and $l_3 - l_2$ happen to be identical, the number of the MSLs would be reduced. For example, *Case D4.2* corresponds to a possibility that the MSLs atl_1 and $l_2 - l_1$ are overlapped as well as the MSLs at l_2 and $l_3 - l_2$ are overlapped.

We have obtained earlier the expressions of $r(l)$, $(l = 0, 1, \ldots, l_{D-1})$, in terms of the effective channel matrix $z(d), (d = 0, 1, \ldots, D-1)$ for the cases $D = 2, 3, 4$. When $D > 4$, a table similar to Table I can easily be designed by simple computer programming. Noting that $l_0 = 0$, all the potential MSLs can be written as $\epsilon_{ij} = l_i - l_j (i > j)$, each corresponding to $r_{i,j} = r(l = l_i - l_j)$. If any two of ϵ_{ij}'s are identical, the corresponding two potential MSLs are overlapped, and, thus, the corresponding two $r_{i,j}$'s are combined. As seen in the next subsection, the structure of $r(l)$ can be exploited to detect the MSTs of the sparse channel.

B. Detection of MSTs

We now propose an MST detection algorithm based on the structure of $r(l)$ as analyzed in the previous subsection. The idea is to determine the MST position based on the MSL position of $r(l)$. It will be shown through simulation studies that, by using a small number of OFDM symbols and pilot subcarriers, our new MST detection method has a high detection accuracy.

Prior to developing the new method, we first introduce the estimated version $\hat{r}(l)$ of $r(l)$. The correlation function of $y(m,n)$ can be estimated as

$$\hat{r}(l) = \frac{1}{gK}\sum_{m=0}^{g-1}\sum_{n=0}^{K-1} y(m,n)y^*(m,n-l) \tag{24}$$

where $y(m,n) = y(m,K+n)$ for $n < 0$. By letting

$$\mathbf{x}_D(m,n) \triangleq [x(m,n-l_0), x(m,n-l_1), \ldots, x(m,n-l_{D-1})]^{\mathrm{T}}, \quad (n=0,1,\ldots,K-1)$$

where $x(m,n) = x(m,K+n)$ for $n < 0$, the circular convolution (2) in the noisy case can be rewritten in the matrix form as

$$y(m,n) = \mathbf{z}_A \mathbf{x}_D(m,n) + v(m,n). \tag{25}$$

Substituting (25) into (24) and using (13), we have

$$\hat{r}(l) = r(l) + \mathbf{z}_A \Delta\mathbf{R}_{x,D}(l)\mathbf{z}_A^{\mathrm{H}} + r_v \tag{26}$$

where

$$\Delta\mathbf{R}_{x,D}(l) \triangleq \frac{1}{gK}\sum_{m=0}^{g-1}\sum_{n=0}^{K-1} \mathbf{x}_D(m,n)\mathbf{x}_D^{\mathrm{H}}(m,n-l) - \mathbf{R}_{x,D}(l) \tag{27}$$

represents the signal perturbation and r_v is the perturbation error introduced by the noise.

Obviously, the first term in the right-hand side (RHS) of (26) is an ideal correlation matrix of the received signal vector $\mathbf{y}(n)$ with neither the signal perturbation nor the noise corruption, while the second and the third terms are the errors introduced by the signal and noise perturbations, respectively. Note that in the noisy case, $\hat{r}(l)$ is, in general, a nonzero value even if $r(l) = 0$. However, the MSLs can be detected from the norm of $\hat{r}(l)$ since the first term in the RHS of (26), $r(l)$, would be dominant provided that the signal-to-noise ratio (SNR) is not extremely low. Accordingly, the first step of our method is to detect the MSLs of $\hat{r}(l)$ by comparing its absolute value with the threshold, i.e.,

$$\eta = \frac{K_e}{P_L+1}\sum_{l=0}^{P_L} |\hat{r}(l)| \tag{28}$$

where P_L is a predetermined length, and the coefficient K_e is used to adjust the average absolute value of $r(l)$. Once the MSLs of $\hat{r}(l)$ are acquired, the information on their positions can be utilized to determine the MST position l_d of the sparse channel as discussed in the previous subsection. The second step of the new MST detection algorithm can be described as follows.

Assume that a total of W MSLs of $\hat{r}(l)$ have been detected, whose positions are denoted as q_i with$q_0 < q_1 < \cdots < q_{W-1}$. We now determine the possible l_d's according to the structures of $r(l)$ for different values of D and W. Although we have in general $D \le W$, it is clear that the positions of the first and the last MSTs are simply given by

$$l_0 = q_0 = 0 \tag{29}$$

$$l_{D-1} = q_{W-1}. \tag{30}$$

Obviously, (29) and (30) give the only two nonzero taps if $W = 2$. In the following, we determine $l_d\,(d = 1,2,\ldots,D-2)$ from $q_w\,(w = 1,2,\ldots,W-2)$ where $W > 2$.

If $W = 3$, we have only the case *D3.1* according to (20). Then, there is only one additional tap l_d to be determined, which is readily given by

$$l_1 = q_1 = \frac{q_2}{2}.$$

If $W = 4$, we should have case *D3.2* from (21) or *D4.1* in Table I. Then, one can have the following three possible solutions.

Case W4.1: When $D = 3$ with $l_2 > 2l_1$, we have $q_1 = l_1$ and $q_2 = l_2 - l_1$, which gives$l_1 = q_1$ and $l_2 = q_1 + q_2 = q_3$.

Case W4.2: When $D = 3$ with $l_2 < 2l_1$, we obtain$q_1 = l_2 - l_1$ and $q_2 = l_1$, which yields $l_1 = q_2$ and$l_2 = q_1 + q_2 = q_3$.

Case W4.3: When $D = 4$ with $l_1 : l_2 : l_3 = 1 : 2 : 3$, we readily have $l_1 = q_1$ and $l_2 = q_2$.

In order to make a decision among the three choices, we present an LS criterion similar to the sparse LS channel-estimation algorithm proposed in Section II-B. Let $l'_d\,(d = 0,1,\ldots,D'-1)$ be the potential MSTs of the sparse channel and $\bar{\mathbf{F}}'(m)$ is a $K_p \times D'$ matrix, whose dth column is the l_dth column of $\mathbf{F}(m), (d = 0,1,\ldots,D'-1)$. Using the idea of joint channel estimation and zero detection proposed in[27], a cost function for the new MST detection method can be established from (9) as

$$J(l'_0,\ldots,l'_{D'-1}) = \|\mathbf{Y}_{\mathrm{pilot}} - \bar{\mathbf{A}}'\hat{\mathbf{z}}'\|_F^2 \tag{31}$$

where $\hat{\mathbf{z}}' = (\bar{\mathbf{A}}')^{\dagger}\mathbf{Y}_{\mathrm{pilot}}$ and

$$\bar{\mathbf{A}}' \triangleq \begin{bmatrix} \mathbf{X}_{\mathrm{pilot-diag}}(0)\bar{\mathbf{F}}'(0) \\ \vdots \\ \mathbf{X}_{\mathrm{pilot-diag}}(g-1)\bar{\mathbf{F}}'(g-1) \end{bmatrix}. \tag{32}$$

Evidently, the cost J reaches its minimum when $l'_d(d = 0,1,\ldots,D'-1)$ gives true MSTs. In the case of$W = 4$, our MST detection only needs to calculate a few costs as given by (31) with respect to Cases *W4.1, W4.2*, and *W4.3*, denoted as$J_{4,1}$, $J_{4,2}$ and $J_{4,3}$, respectively. The complete scheme is shown in Fig. 1, which gives three possible MST detection results, namely: *Case W4.1:* $l_1 = q_1$; *Case W4.2:* $l_1 = q_2$; and *Case W4.3:* $l_1 = q_1, l_2 = q_2$.

It should be stressed that the detection scheme in Fig. 1 is already the improved version of a straightforward implementation. First, instead of computing all the three costs $J_{4,1}$, $J_{4,2}$, and $J_{4,3}$ and making a detection among the three values, a test on $q_1 : q_2 : q_3 = 1 : 2 : 3$ is performed to possibly reduce the three candidates to two, which could not only save the calculation of one cost but also enhance the detection performance by excluding the false MST candidate *Case W4.3* from comparison. Second, when the test is "yes," we have found that $J_{4,3}$ could be very close to $J_{4,1}$ ($J_{4,2}$) if *Case W4.1* (*Case W4.2*) is true, indicating a possibility of misdetection of *Case W4.1* (*Case W4.2*) as *Case W4.3*, at a low level of SNR. On the other hand, when *Case W4.3* is true, $J_{4,3}$ is found to be much smaller than

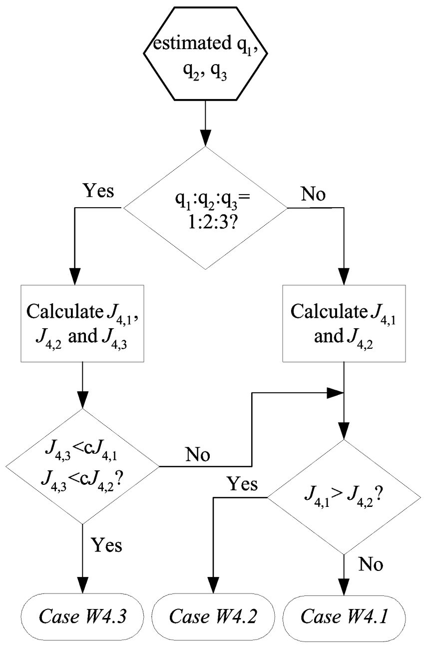

Fig. 1. MST detection scheme for $W = 4$.

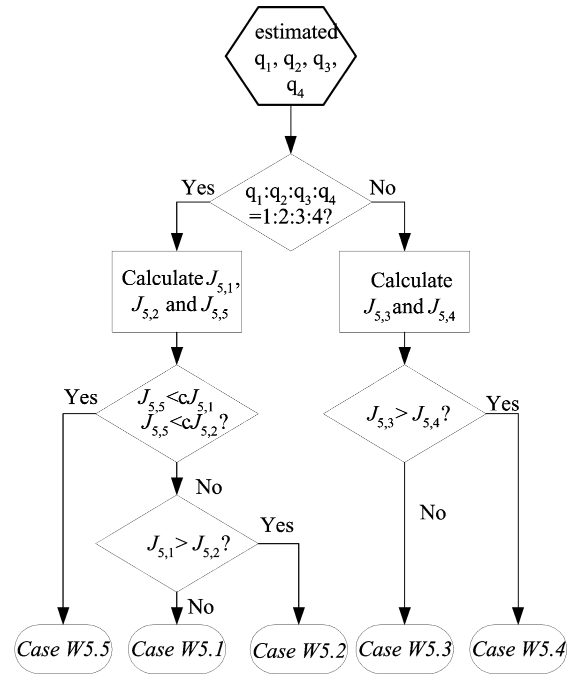

Fig. 2. MST detection scheme for $W = 5$.

$J_{4,1}$ and $J_{4,2}$. In the improved detection scheme, therefore, we have employed a scaling coefficient c to distinguish *W4.3* from *W4.1* and *W4.2*. The value of c should be chosen according to the level of SNR. For example, for a moderate SNR, we have found that $c = 0.7–0.9$ is a proper choice.

Similarly, for $W = 5$, one can have the following possible situations:

Case W5.1: $q_1 = l_1, q_2 = l_2$, and $q_3 = l_3 - l_1$, when $D = 4$ and $l_1 : l_2 : l_3 = 1 : 2 : 4$;

Case W5.2: $q_1 = l_2 - l_1, q_2 = l_1$, and $q_3 = l_2$, when $D = 4$ and $l_1 : l_2 : l_3 = 2 : 3 : 4$;

Case W5.3: $q_1 = l_1, q_2 = l_2 - l_1$, and $q_3 = l_2$, when $D = 4$, $l_3 = l_1 + l_2$, and $l_2 > 2l_1$;

Case W5.4: $q_1 = l_2 - l_1, q_2 = l_1$, and $q_3 = l_2$, when $D = 4$, $l_3 = l_1 + l_2$, and $l_2 < 2l_1$;

Case W5.5: $q_1 = l_1, q_2 = l_2$, and $q_3 = l_3$, when $D = 5$ and $l_1 : l_2 : l_3 : l_4 = 1 : 2 : 3 : 4$.

The corresponding MST detection scheme is described in Fig. 2, which gives the following five detection results.

Case W5.1: $l_1 = q_1, l_2 = q_2$; *Case W5.2:* $l_1 = q_2, l_2 = q_3$; *Case W5.3:* $l_1 = q_1, l_2 = q_3$;

Case W5.4: $l_1 = q_2, l_2 = q_3$; and *Case W5.5:* $l_1 = q_1$, $l_2 = q_2, l_3 = q_3$.

Clearly, the key to the aforementioned detection scheme is to design a table listing all the possible MST candidates corresponding to the same value of W using the expression of $r(l)$ obtained in the previous subsection. This table can easily be constructed even if W has a large value, since the expression of $r(l)$ can simply be derived by computer programming as stated before. The proposed sparse LS channel-estimation approach can be summarized as follows.

1) MSLs are first estimated by comparing the absolute value of the correlation function with the threshold calculated from (28).
2) The estimated MSLs are utilized, together with a small number of pilots, for the estimation of MSTs.
3) Based on the estimated MSTs, the channel coefficients are estimated in an LS fashion as given by (7).

Note that, in some extremely noisy cases, the MSLs might be incorrectly detected, which would lead to the cases that there are no corresponding MST detected due to the logic in the MST detection scheme, as shown in Fig. 1. In this case, we can use a joint MSL and MST detection scheme. First, a relatively small threshold η is used to get more candidate lags, leading to all possible combinations of different values of W and positions of lags. For each of these combinations, a minimum cost can then be obtained by using the aforementioned proposed scheme. The final decision can be made based on the minimum of all the combinations. Overall, in comparison with many of the MST detection methods available in the literature, which normally require searching all the possible MST candidates over the entire channel length L, the proposed detection method reduces the search range to a very small number of MST candidates, by exploiting the MSL information of $\hat{r}(l)$. More importantly, the conventional methods need a large number of pilots or OFDM symbols to conduct a training-based or CP-assisted MST detection. In contrast, our technique uses only a small number of OFDM symbols and pilots to implement an accurate detection. It will be shown in Section V that the performance of the proposed sparse LS channel-estimation algorithm is significantly superior to that of the regular LS method.

IV. Sparse-Channel Estimation of MIMO–OFDM Systems

In this section, we would like to extend the proposed semi-blind MST detection algorithm for sparse channel estimation of MIMO–OFDM systems. Consider a MIMO–OFDM system with the vertical Bell Laboratories layered space-time

(V-BLAST) structure, which consists of N_T transmit and N_R receive antennas. Then, the lth tap of the MIMO channel can be described by a channel matrix as

$$\mathbf{H}(l) \triangleq \begin{bmatrix} h_{1,1}(l) & h_{1,2}(l) & \cdots & h_{1,N_T}(l) \\ h_{2,1}(l) & h_{2,2}(l) & \cdots & h_{2,N_T}(l) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R,1}(l) & h_{N_R,2}(l) & \cdots & h_{N_R,N_T}(l) \end{bmatrix} \notin C^{N_R \times N_T}$$

where $h_{i_R,i_T}(l), (0 \le l \le L-1)$ represents the channel response between the i_Rth receive antenna and the i_Tth transmit antenna. Accordingly, the time-domain signal model (2) in single-input–single-output (SISO) OFDM systems can be modified for the MIMO case as

$$y_{i_R}(m,n) = \sum_{i_T=1}^{N_T} h_{i_R,i_T}(n) \circledast x_{i_T}(m,n) + v_{i_R}(m,n), \quad m \in \{0,\ldots,g-1\} \tag{33}$$

where $x_{i_T}(m,n)$ is the time-domain signal with respect to the i_Tth transmit antenna and $y_{i_R}(m,n)$ and $v_{i_R}(m,n)$ are the time-domain signal and noise with respect to the i_Rth receive antenna. Here, we assume a sparse MIMO channel, in which all $N_R \times N_T$ elements of $\mathbf{H}(l)$ are assumed to have the same MSTs. In this case, the channel matrix with respect to the dth $(d = 0,1,\ldots,D-1)$ MST (namely, the nonzero tap) can be expressed as

$$\mathbf{Z}(d) = \mathbf{H}(l_d) \tag{34}$$

where $l_d \, (d = 0,1,\ldots,D-1)$ are integers with$0 = l_0 < l_1 < \cdots < l_{D-1}$. In the following, we first derive a sparse LS algorithm for the effective channel matrix$\mathbf{Z}(d), (d = 0,1,\ldots,D-1)$.

The MIMO version of (4) can be easily obtained as

$$\mathbf{Y}_{i_R,\text{pilot}}(m) = \sum_{i_T=1}^{N_T} \mathbf{X}_{i_T,\text{pilot-diag}}(m)\mathbf{F}(m)\mathbf{h}_{i_R,i_T} + \boldsymbol{\xi}_{i_R,\text{pilot}}(m) \tag{35}$$

where $\mathbf{X}_{i_T,\text{pilot-diag}}(m)$ and $\mathbf{Y}_{i_R,\text{pilot}}(m)$ are the transmit and received signal vector in the link between the i_Tth transmit and the i_Rth receive antennas, respectively, and $\mathbf{h}_{i_R,i_T} \triangleq [h_{i_R,i_T}(0),\ldots,h_{i_R,i_T}(L-1)]^T$. Using (34) and (35), one can obtain

$$\mathbf{Y}_{i_R,\text{pilot}}(m) = \sum_{i_T=1}^{N_T} \mathbf{X}_{i_T,\text{pilot-diag}}(m)\bar{\mathbf{F}}(m)\mathbf{z}_{i_R,i_T} + \boldsymbol{\xi}_{i_R,\text{pilot}}(m) \tag{36}$$

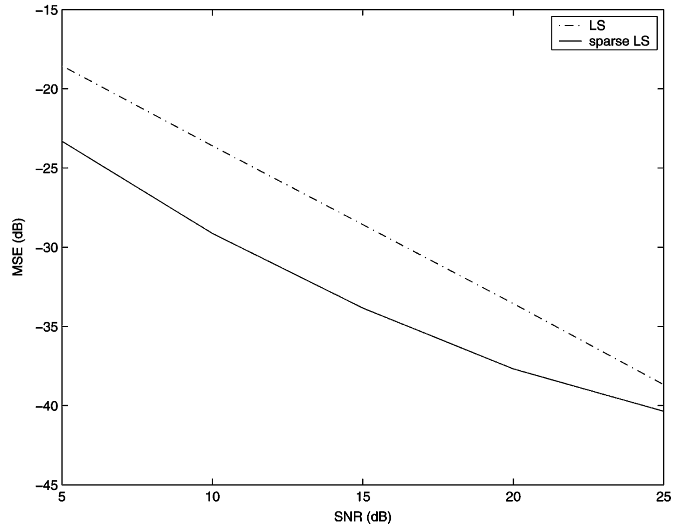

Fig. 3. MSE versus SNR.

where $\mathbf{z}_{i_R,i_T} = [z_{i_R,i_T}(0),\ldots,z_{i_R,i_T}(D-1)]^{\mathrm{T}}$. By defining

$$\tilde{\mathbf{Y}}_{\text{pilot}} \triangleq [\mathbf{Y}_{1,\text{pilot}} \ldots, \mathbf{Y}_{N_R,\text{pilot}}]$$
$$\mathbf{Z} \triangleq \begin{bmatrix} \mathbf{z}_{1,1} & \cdots & \mathbf{z}_{N_R,1} \\ \vdots & \ddots & \vdots \\ \mathbf{z}_{1,N_T} & \cdots & \mathbf{z}_{N_R,N_T} \end{bmatrix} \tag{37}$$

and using a manner similar to the derivation of (9), it can be proved that

$$\tilde{\mathbf{Y}}_{\text{pilot}} = \tilde{\mathbf{A}}\mathbf{Z} + \bar{\boldsymbol{\xi}}_{\text{pilot}} \tag{38}$$

where the relationship of the terms are given in (39) at the bottom of the page.

From (38), an LS estimate of the sparse channel can be obtained as

$$\hat{\mathbf{z}} = (\tilde{\mathbf{A}})^{\dagger}\tilde{\mathbf{Y}}_{\text{pilot}}. \tag{40}$$

Now, we consider the extension of the proposed semi-blind MST detection algorithm for the MIMO–OFDM systems. In contrast to the correlation function for OFDM systems, the second-order statistics of the received signal for the MIMO–OFDM systems is a correlation matrix, which can be, in general, written as

$$\mathbf{R}(l) \triangleq \mathrm{E}\{\mathbf{y}(n)\mathbf{y}^{\mathrm{H}}(n-l)\}, \qquad (l = 0,1,\ldots,P) \tag{41}$$

where

$$\mathbf{y}(n) \triangleq [y_1(n), y_2(n), \ldots, y_{N_R}(n)]^{\mathrm{T}}. \tag{42}$$

$$\tilde{\mathbf{A}} \triangleq \begin{bmatrix} \mathbf{X}_{1,\text{pilot-diag}}(0)\mathbf{F}(0) & \cdots & \mathbf{X}_{N_T,\text{pilot-diag}}(0)\mathbf{F}(0) \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{1,\text{pilot-diag}}(g-1)\mathbf{F}(g-1) & \cdots & \mathbf{X}_{N_T,\text{pilot-diag}}(g-1)\mathbf{F}(g-1) \end{bmatrix}. \tag{39}$$

In the previous two equations, the symbol index m has been omitted. Substituting (42) into (41) and using (33) and (34), we obtain

$$\mathbf{R}(l) = \mathbf{Z}_A \bar{\mathbf{R}}_{x,D}(l) \mathbf{Z}_A^{\mathrm{H}} \tag{43}$$

where

$$\mathbf{Z}_A \triangleq [\mathbf{Z}(0) \quad \mathbf{Z}(1) \quad \cdots \quad \mathbf{Z}(D-1)] \tag{44}$$

$$\bar{\mathbf{R}}_{x,D}(l) \triangleq \mathrm{E}\left\{ \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{x}(n-l_1) \\ \vdots \\ \mathbf{x}(n-l_{D-1}) \end{bmatrix} \begin{bmatrix} \mathbf{x}(n-l) \\ \mathbf{x}(n-l_1-l) \\ \vdots \\ \mathbf{x}(n-l_{D-1}-l) \end{bmatrix}^H \right\} \tag{45}$$

with $\mathbf{x}(n) \triangleq [x_1(n), x_2(n), \ldots, x_{N_T}(n)]^{\mathrm{T}}$. Noting that $\mathrm{E}\{\mathbf{x}(n-i)\mathbf{x}^{\mathrm{H}}(n-j)\} = \sigma_x^2 \delta(i-j)\mathbf{I}$, (45) can be rewritten as

$$\bar{\mathbf{R}}_{x,D}(l) = \mathbf{R}_{x,D}(l) \otimes \mathbf{I}. \tag{46}$$

Using (44), (46), and (16), one can prove that (43) can be rewritten as

$$\mathbf{R}(l) = \sum_{i=0}^{D-1} \sum_{j=0}^{D-1} \delta(l - l_i + l_j) \mathbf{Z}(i) \mathbf{Z}^{\mathrm{H}}(j) \tag{47}$$

By comparing (47) with (17), one can find that the difference between the two equations is that the scalers $r(l)$ and $z(i)$ are replaced by the matrices $\mathbf{R}(l)$ and $\mathbf{Z}(i)$. Therefore, the proposed semi-blind MST detection algorithm in Section III can be applied to channel estimation of MIMO–OFDM systems after modifications, as shown in the following.

1) The symbols $\hat{r}(l)$ and $z(d)$ in the SISO OFDM case should be replaced by $\hat{\mathbf{R}}(l)$ and $\mathbf{Z}(d)$, respectively, for MIMO–OFDM channel estimation. Accordingly, $r_{i,j} = z(i)z^*(j)$ in Table I is replaced with $\mathbf{R}_{i,j} = \mathbf{Z}(i)\mathbf{Z}^{\mathrm{H}}(j)$.
2) Instead of using (28), the threshold used for MSL estimation is calculated by

$$\eta = \frac{K_e}{P_L + 1} \sum_{l=0}^{P_L} \|\hat{\mathbf{R}}(l)\|_F. \tag{48}$$

3) Instead of using (31), the cost function in the MST detection can be established from (38) as

$$J\left(l_0', \ldots, l_{D'-1}'\right) = \|\tilde{\mathbf{Y}}_{\mathrm{pilot}} - \tilde{\mathbf{A}}'\hat{\mathbf{z}}'\|_F^2 \tag{49}$$

where $\hat{\mathbf{z}}' = (\tilde{\mathbf{A}}')^{\dagger}\tilde{\mathbf{Y}}_{\mathrm{pilot}}$, and $\tilde{\mathbf{A}}'$ is given in the equation shown at the bottom of the page, where $\bar{\mathbf{F}}'(m)$ has the same definition as in Section III-B.

Once the MSTs are obtained by using the proposed semi-blind algorithm, the MIMO channels can be estimated in the LS sense, as given by (40), yielding a sparse LS estimate of the effective channel matrix $\mathbf{Z}(d), (d = 0, 1, \ldots, D-1)$.

V. Simulation Results

A. OFDM Channel Estimation

We consider an OFDM system in which the number of subcarriers is set to 1024, and the length of the CP is 30. In our simulation, the QPSK modulation is used, and a sparse Rayleigh channel modeled by a three-nonzero tap FIR filter is assumed, in which each tap is independent identically distributed (i.i.d.) complex normal distributed. The MSTs are $l_0 = 0$, $l_1 = 4$, and $l_2 = 11$. The constant K_e in (28) is set to 0.8. The estimation performance is evaluated in terms of the MSE of the channel estimate as given by

$$\mathrm{MSE} = \frac{1}{N_{\mathrm{MC}}} \sum_{n=1}^{N_{\mathrm{MC}}} \|\hat{\mathbf{h}}_n - \mathbf{h}_n\|^2$$

where N_{MC} is the number of Monte Carlo iterations and $\mathbf{h}_n$ and $\hat{\mathbf{h}}_n$ are the true and the estimated channel vectors with respect to the nth Monte Carlo iteration, respectively.

Experiment A1: MSE Versus SNR: Here, the MSE performance versus the SNR is investigated. The simulation involves 2000 Monte Carlo runs of the transmission of ten OFDM symbols with 30 pilot subcarriers. Fig. 3 shows the MSE plots of the proposed sparse and original LS methods. It is seen that the sparse LS method performs significantly better than the original LS method, and, in particular, it achieves a gain of about 5.5 dB over the regular LS version when the SNR is 10 dB.

Experiment A2: MSE Versus Pilot Length: In this experiment, we investigate the channel-estimation performance versus the number of pilot subcarriers. Fig. 4 shows the MSE plots from 2000 Monte Carlo iterations for five OFDM symbols with pilot subcarriers varying from 15 to 35 at an SNR of 10 dB. One can find that the sparse LS method can achieve about 6-dB gain over the regular LS method.

Experiment A3: MSE Versus the Number of OFDM Symbols: In this experiment, we investigate the channel-estimation performance versus the number of OFDM symbols. Fig. 5 shows the MSE plots from 1000 Monte Carlo iterations for 30 pilot subcarriers with OFDM symbols varying from 5 to 25 at an SNR of 10 dB. One can find that when only five OFDM symbols are used, the sparse LS method can achieve a gain of 6 dB over the regular LS method. When 25 OFDM symbols are used, it is observed that the performance improvement from the sparse LS method becomes 4.9 dB. Obviously, the proposed method is more advantageous for a smaller number of OFDM symbols used.

$$\tilde{\mathbf{A}}' \triangleq \begin{bmatrix} \mathbf{X}_{1,\mathrm{pilot-diag}}(0)\bar{\mathbf{F}}'(0) & \cdots & \mathbf{X}_{N_T,\mathrm{pilot-diag}}(0)\bar{\mathbf{F}}'(0) \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{1,\mathrm{pilot-diag}}(g-1)\bar{\mathbf{F}}'(g-1) & \cdots & \mathbf{X}_{N_T,\mathrm{pilot-diag}}(g-1)\bar{\mathbf{F}}'(g-1) \end{bmatrix}. \tag{50}$$

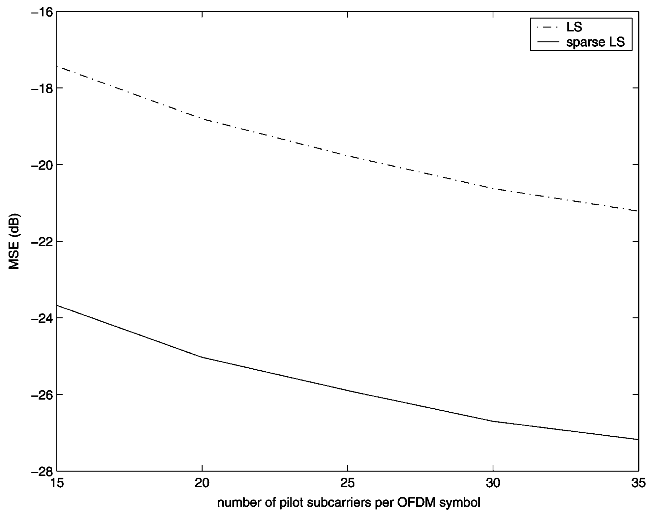

Fig. 4. MSE versus the number of pilot subcarriers.

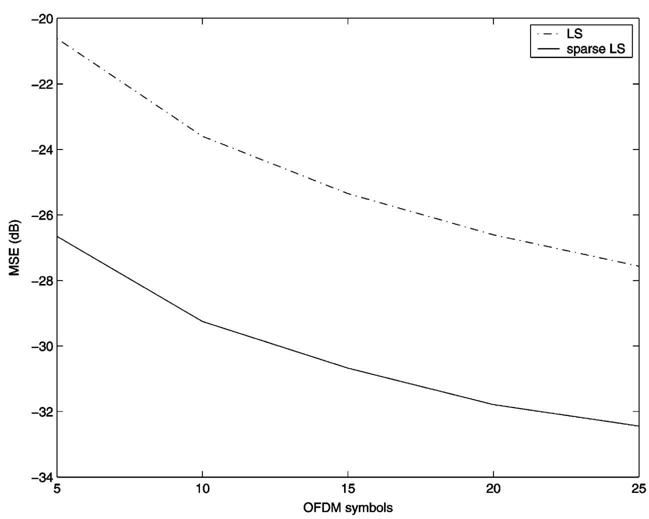

Fig. 5. MSE versus the number of OFDM symbols.

TABLE II
ERROR PROBABILITY OF THE MST DETECTION VERSUS DIFFERENT SNR

SNR(dB)	Error probability
5	2.7×10^{-3}
9	4×10^{-4}
13	3×10^{-4}
17	10^{-4}
21	10^{-4}

B. MIMO–OFDM Channel Estimation

Based on the same condition for the OFDM system as in the previous subsection, we consider now a MIMO–OFDM system with two transmit and four receive antennas. The training sequence is generated using the method in [2]. For the purpose of comparison, the channel is first estimated by the regular LS method and our previously developed semi-blind method[7], [8]. As for the estimation of the effective channel, we consider the proposed sparse LS method with the MST detection. An ideal sparse LS method with the knowledge of the true MST information is also simulated for comparison.

Experiment B1: MSE Versus SNR: In this experiment, the channel estimation performance in terms of the MSE as a function of the SNR is investigated for a three-tap sparse channel with $l_0 = 0$, $l_1 = 4$, and $l_2 = 11$. First, the channel is considered as Rayleigh distributed, i.e., each tap corresponds to a 4×2 random matrix whose elements are i.i.d. complex Gaussian variables with zero mean and unit variance. The simulation involves 20 000 Monte Carlo runs of the transmission of one OFDM symbol at 1024 subcarriers, of which 30 are used as pilot for training purpose. The error probability of the MST detection versus SNR is shown in Table II. It is clear that a high performance of the MST detection is achieved despite the fact that only one OFDM symbol is used. In particular, when $\mathrm{SNR} \geq$ 17 dB, the error probability is 10^{-4}. Fig. 6 shows the MSE plots from the proposed as well as the ideal sparse methods along with two regular methods. It is seen that the sparse LS method is highly consistent with its ideal version. It is also noted that the proposed sparse method significantly outperforms the two regular methods. Specifically, the sparse LS method is consistently superior to the original method by nearly 6.5 dB. It is seen that the ideal and the proposed sparse LS methods deliver almost the same estimation results. It is interesting to see such a phenomenon. This is because the MSE plot is the averaged result from 20 000 Monte Carlo iterations; meanwhile, the error probability is very small, which does not considerably affect the MSE of the channel estimation for most of the SNR levels. It is of interest to mention that, by utilizing the proposed MST detection algorithm in conjunction with our previously developed semi-blind channel-estimation method in [8], an additional gain in the channel-estimation performance can be expected.

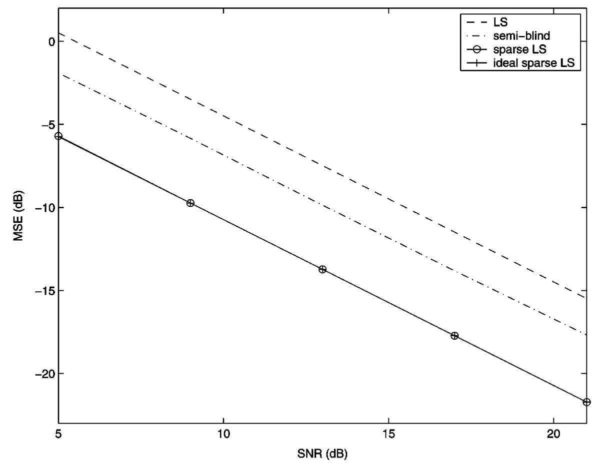

Fig. 6. MSE versus SNR for a Rayleigh fading channel.

We now show the estimation result of another channel, which has a similar sparsity as in the previous channel, but undergoes an exponential decay with variances of the three significant taps as$\sigma_0 = 1$, $\sigma_1 = \mathrm{sqrt}(10^{-0.2})$, and $\sigma_2 = \mathrm{sqrt}(10^{-0.55})$. It is seen from Fig. 7, for this exponential channel, that the MSE of the sparse method is slightly worse than that of the ideal sparse method. However, the LS sparse solution is significantly better than the two regular methods.

Experiment B2: BER Versus SNR: Now, the bit error rate (BER) performance of the MIMO–OFDM system is investigated by using the estimated channel matrix and an ordered V-BLAST decoder. The channel is modeled by a four-tap sparse Rayleigh channel with $l_0 = 0$, $l_1 = 4$, $l_2 = 8$, and $l_3 = 12$. The

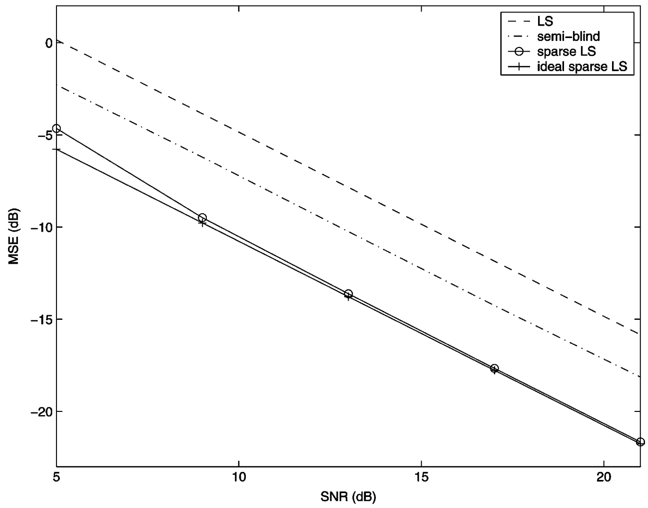

Fig. 7. MSE versus SNR for a channel with exponential decay.

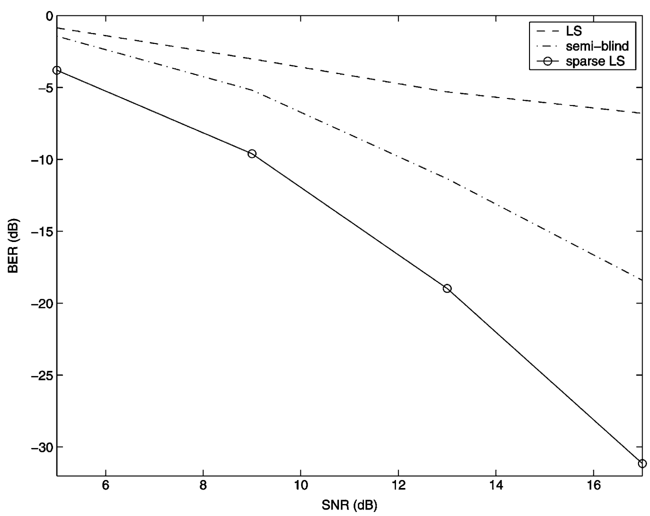

Fig. 8. BER versus SNR.

simulation involves 4000 Monte Carlo runs of the transmission of one OFDM symbol with 20 pilot subcarriers. Fig. 8 shows the BER performance of the proposed sparse LS method as well as its original version. Clearly, the performance improvement provided by the proposed sparse LS method gets more prominent compared with the regular one as SNR increases. For example, when SNR is increased to 17 from 5 dB, the performance gain of the proposed sparse LS method is boosted to 24 from 3 dB.

VI. Conclusion

A semi-blind MST detection approach has been proposed for the sparse channel estimation of OFDM and MIMO–OFDM systems. A relationship between the MSTs of the sparse channel and the MSLs of the correlation function of the OFDM signal received in the absence of noise has first been revealed. This relation has then been exploited to develop a highly efficient semi-blind MST detection algorithm that requires only a small number of OFDM symbols and pilot subcarriers for an LS-based detection. The semi-blind MST detection algorithm has been extended to MIMO–OFDM channel estimation. As the new approach does not require estimating all the channel taps, it has saved a large amount of computations compared with regular channel-estimation methods. Computer simulations based on various sparse channels have confirmed that the proposed sparse-channel-estimation approach significantly outperform the regular OFDM and MIMO–OFDM channel-estimation techniques without utilizing the MST detection.

Acknowledgment

The authors would like to thank the anonymous reviewers, the Associate Editor, and the Editor-in-Chief for their careful review of this paper. Their valuable comments and suggestions have greatly helped improve the presentation of this paper.

References

[1] Y. Jung, J. Kim, S. Lee, H. Yoon, and J. Kim, "Design and implementation of MIMO–OFDM baseband processor for high-speed wireless LANs," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 54, no. 7, pp. 631–635, Jul. 2007.

[2] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO–OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.

[3] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO–OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 670–685, Mar. 2007.

[4] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO–OFDM systems via nonredundant linear precoding," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 784–789, Jan. 2007.

[5] Y.-S. Chen and C.-A. Lin, "Blind-channel identification for MIMO single-carrier zero-padding block-transmission systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 6, pp. 1571–1579, Jul. 2008.

[6] Y. Zeng, W. H. Lam, and T. S. Ng, "Semi-blind channel estimation and equalization for MIMO space-time coded OFDM," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 2, pp. 463–474, Feb. 2006.

[7] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "Linear prediction based semi-blind channel estimation for MIMO–OFDM system," in *Proc. IEEE ISCAS*, 2007, pp. 3239–3242.

[8] F. Wan, W.-P. Zhu, and M. N. S. Swamy, "A semi-blind channel estimation approach for MIMO–OFDM systems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2821–2834, Jul. 2008.

[9] M. R. Raghavendra, E. Lior, S. Bhashyam, and K. Giridhar, "Parametric channel estimation for pseudo-random tile-allocation in uplink OFDMA," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5370–5381, Nov. 2007.

[10] H. Minn and V. K. Bhargava, "An investigation into time-domain approach for OFDM channel estimation," *IEEE Trans. Broadcast.*, vol. 46, no. 4, pp. 240–248, Dec. 2000.

[11] M. R. Raghavendra and K. Giridhar, "Improving channel estimation in OFDM systems for sparse multipath channels," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 52–55, Jan. 2005.

[12] J. K. Hwang, R. L. Chung, M. F. Tsai, and J. H. Deng, "Highly efficient sparse multipath channel estimator with Chu-sequence preamble for frequency-domain MIMO DFE receiver," *IEICE Trans. Commun.*, vol. E90-B, no. 8, pp. 2103–2110, Aug. 2007.

[13] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374–377, Mar. 2002.

[14] Y. Liu and D. K. Borah, "Estimation of fading channels with large possible delay spreads," *Electron. Lett.*, vol. 39, no. 1, pp. 130–131, Jan. 2003.

[15] S. Kim and R. A. Iltis, "A matching-pursuit/GSIC-based algorithm for DS-CDMA sparse-channel estimation," *IEEE Signal Process. Lett.*, vol. 11, no. 1, pp. 12–15, Jan. 2004.

[16] D. K. Borah, "Estimation of frequency-selective CDMA channels with large possible delay and Doppler spreads," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1126–1136, Jul. 2006.

[17] C. R. N. Athaudage and A. D. S. Jayalath, "Delay-spread estimation using cyclic-prefix in wireless OFDM systems," *Proc. Inst. Elect. Eng. Commun.*, vol. 151, no. 6, pp. 559–566, Dec. 2004.

[18] G. E. Pfander, H. Rauhut, and J. Tanner, "Identification of matrices having a sparse representation," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5376–5388, Nov. 2008.

[19] J. Homer, I. Mareels, and C. Hoang, "Enhanced detection-guided NLMS estimation of sparse FIR-modeled signal channels," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 8, pp. 1783–1791, Aug. 2006.

[20] J. Oliver, R. Aravind, and K. M. M. Prabhu, "Sparse channel estimation in OFDM systems by threshold-based pruning," *Electron. Lett.*, vol. 44, no. 13, pp. 830–832, Jun. 19, 2008.

[21] I. C. Wong and B. L. Evans, "Sinusoidal modeling and adaptive channel prediction in mobile OFDM systems," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1601–1615, Apr. 2008.

[22] J. G. Proakis*, Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[23] C. Li and S. Roy, "Subspace-based blind channel estimation for OFDM by exploiting virtual carriers," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 141–150, Jan. 2003.

[24] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 315–325, Jan. 2004.

[25] C.-A. Lin and Y.-S. Chen, "Blind identification of MIMO channels using optimal periodic precoding," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 4, pp. 901–911, Apr. 2007.

[26] K. Kohno, Y. Inouye, and M. Kawamoto, "A matrix pseudo-inversion lemma for positive semidefinite Hermitian matrices and its application to adaptive blind deconvolution of MIMO systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 1, pp. 424–435, Feb. 2008.

[27] C. Carbonelli, S. Vedantam, and U. Mitra, "Sparse channel estimation with zero tap detection," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1743–1763, May 2007.

Feng Wan (S'09) received the B.E. degree in electrical engineering from Wuhan University, Wuhan, China, in 2000, the M.E. degree in electrical engineering from the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, in 2003, and the Ph.D. degree in electrical engineering from Concordia University, Montreal, QC, Canada, in 2009.

Since 2004, he has been a Research Assistant with Concordia University. His research interests are in the area of digital signal processing, statistical estimation theory and its applications in MIMO, OFDM, and smart antenna wireless communications.

Dr. Wan is the recipient of the Best Paper Award at the 2008 IEEE International Conference on Neural Networks and Signal Processing. He was also the recipient of the Doctoral Thesis Completion Award, the Graduate Fellowship, and the International Tuition Fee Remission Award from Concordia University, the Chinese Government Award for Outstanding Self-Financed Students Abroad from China Scholarship Council, and the Postdoctoral Fellowship from Natural Sciences and Engineering Research Council of Canada.

Wei-Ping Zhu (SM'97) received the B.E. and M.E. degrees in electrical engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical engineering from Southeast University, Nanjing, in 1991.

He was a Postdoctoral Fellow from 1991 to 1992 and a Research Associate from 1996 to 1998 with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada. During 1993–1996, he was an Associate Professor with the Department of Information Engineering, Nanjing University of Posts and Telecommunications. From 1998 to 2001, he worked in hi-tech companies in Ottawa, Canada, including Nortel Networks and SR Telecom Inc. Since July 2001, he has been with Concordia University's Electrical and Computer Engineering Department as an Associate Professor. His research interests include the areas of digital filters, filter banks, speech and image processing, and signal processing for wireless communication.

Dr. Zhu was an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I: FUNDAMENTAL THEORY AND APPLICATIONS from October 2001 to December 2003. He currently serves as an Associate Editor of *Circuits, Systems and Signal Processing*.

M. N. S. Swamy (S'59–M'62–SM'74–F'80) received the B.Sc. (Hons.) degree in mathematics from Mysore University, Mysore, India, in 1954, the Diploma degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1957, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Saskatchewan, Saskatoon, Canada, in 1960 and 1963, respectively. In August 2001, He received the Doctor of Science (*Honoris Causa*) degree in engineering by Ansted University "in recognition of his exemplary contributions to the research in Electrical and Computer Engineering and to Engineering Education, as well as his dedication to the promotion of Signal Processing and Communications Applications."

Since July 2001, he holds the Concordia Chair (Tier I) in Signal Processing, Concordia University, Montreal, QC, Canada. He has also taught with the Electrical Engineering Department, Technical University of Nova Scotia, Halifax, NS, and with the University of Calgary, Calgary, AB, Canada, as well as with the Department of Mathematics, University of Saskatchewan. He is currently a Research Professor and the Director of the Center for Signal Processing and Communications with the Department of Electrical and Computer Engineering, Concordia University, where he served as the Founding Chair of the Department of Electrical Engineering from 1970 to 1977, and Dean of Engineering and Computer Science from 1977 to 1993, during which time, he developed the faculty into a research oriented faculty from what was primarily an undergraduate one. In 2008, Concordia University instituted the *M. N. S. Swamy Research Chair in Electrical Engineering* as a recognition of his research contributions. In May 2009, he was awarded the title of *Honorary Professor* by the National Chiao Tong University, Taiwan. He has published extensively in the areas of number theory, circuits, systems and signal processing, and is the holder of five patents. He is the coauthor of two book chapters and four books: *Graphs, Networks and Algorithms* (Wiley, 1981; Russian translation—Mir Publishers, 1984; Chinese translation—Education Press, 1987), *Graphs: Theory and Algorithms* (Wiley, 1992), *Switched Capacitor Filters: Theory, Analysis and Design* (Prentice Hall International U.K. Ltd., 1995), and*Neural Networks in a Softcomputing Framework* (New York: Springer, 2006).

Dr. Swamy is a Fellow of a number of professional societies including the Institute of Electrical Engineers (United Kingdom), the Engineering Institute of Canada, the Institution of Engineers (India), and the Institution of Electronic and Telecommunication Engineers (India). He has served the IEEE in various capacities such as the President-Elect in 2003, President in 2004, Past-President in 2005, Vice President (Publications) during 2001–2002, Vice-President in 1976, Editor-in-Chief of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I from June 1999 to December 2001, Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS during June 1985 to May 1987, Program Chair for the 1973 IEEE Circuits and Systems (CAS) Symposium, General Chair for the 1984 IEEE CAS Symposium, Vice-Chair for the 1999 IEEE CAS Symposium, and a member of the Board of Governors of the CAS Society. He was a founding member of Micronet from its inception in 1999 as a Canadian Network of Centers of Excellence until its expiration in 2004, and was also its coordinator for Concordia University. He is the recipient of many IEEE-CAS Society awards, including the Education Award in 2000, Golden Jubilee Medal in 2000, and the 1986 Guillemin-Cauer Best Paper Award.