Joint Data Detection and Channel Estimation for OFDM Systems

Tao Cui, *Student Member, IEEE*, and Chintha Tellambura, *Senior Member, IEEE*

***Abstract*—We develop new blind and semi-blind data detectors and channel estimators for orthogonal frequency-division multiplexing (OFDM) systems. Our data detectors require minimizing a complex, integer quadratic form in the data vector. The semi-blind detector uses both channel correlation and noise variance. The quadratic for the blind detector suffers from rank deficiency; for this, we give a low-complexity solution. Avoiding a computationally prohibitive exhaustive search, we solve our data detectors using sphere decoding (SD) and V-BLAST and provide simple adaptations of the SD algorithm. We consider how the blind detector performs under mismatch, generalize the basic data detectors to nonunitary constellations, and extend them to systems with pilots and virtual carriers. Simulations show that our data detectors perform well.**

***Index Terms*—Channel estimation, orthogonal frequency-division multiplexing (OFDM), sphere decoding (SD).**

I. Introduction

ORTHOGONAL frequency-division multiplexing (OFDM) is used for high-data-rate wireless local area network (WLAN) standards, such as the Hiperlan and IEEE 802.11a, with data rates of up to 54 Mb/s, and it is being considered for fourth-generation (4G) mobile wireless systems [1]. A cyclic prefix (CP) and pilot tones for channel estimation [2]–[6] constitute a significant overhead or bandwidth loss, motivating the development of blind techniques for OFDM. They use statistical or deterministic properties of the transmit and receive signals; properties such as CP and pilot-induced redundancy, cyclostationarity, finite alphabets, and virtual carriers have been exploited [7]–[13]. Blind channel estimation for multiple-antenna OFDM is reported in [14].

Joint estimation of channel impuse response (CIR) and data symbols for OFDM has not been investigated extensively. A maximum-likelihood (ML) joint blind channel and data estimator [15] exploits the finite alphabet property of modulation symbols and the presence of virtual carriers (VCs). In [16], an ML joint estimator is derived, which requires pilot symbols for an initial estimate of the channel. In [17], a blind channel estimator for block fading channels is proposed using the super-trellis and the per-survivor algorithms, which require relatively high complexity. Note that many previous blind estimators typically use averaging over a large number of OFDM symbols (up to several thousands in some cases). These estimators thereby introduce a considerable latency into the overall system and require that the channel remains constant. Thus, estimators that require few OFDM symbols are preferable, as they can operate over nonzero Doppler channels without introducing an appreciable delay. The initial channel estimate can also be used for data detection over several OFDM symbols (provided that the channel variation is slow enough). Recently, [18] also proposed a deterministic blind joint channel and data estimator. The branch-and-bound principle is applied to solve a nonlinear integer problem associated with finding the curve that fits a subchannel in the least-square (LS) sense. However, the blind detector is not optimal and needs several OFDM symbols.

In this paper, we develop new blind and semi-blind data detectors and channel estimators for OFDM systems. Our data detectors require minimizing a complex, integer quadratic $\mathbf{x}^T\mathbf{G}\mathbf{x}^*$ where $\mathbf{x}$ is a data vector and $\mathbf{G}$ is a matrix. The blind detector is derived following the generalized likelihood ratio test (GLRT) approach [19]. The semi-blind detector uses both channel correlation and noise variance and is ML. The quadratic for the blind detector suffers from rank deficiency; for this, we give a low-complexity solution. Both detectors are obtained by posing the problem of the joint estimation of channel and data as a mixed discrete and continuous LS optimization problem. By eliminating the channel from it, we obtain a discrete integer LS problem (which has the same form for both detectors) for the data symbols. An exhaustive search of the solution space yields an optimal solution but has exponential complexity in the number of subcarriers and is computationally prohibitive. Avoiding this problem, we solve our data detectors using the sphere decoder (SD) [20]–[22] and Vertical Bell Laboratories Layered Space-Time (V-BLAST) [23] and provide simple adaptations of the SD. Our approach allows for substantial computational savings over an exhaustive search; for example, the total number of flops decreases by seven orders of magnitude compared with an exhaustive search. We also consider how the blind detector performs under mismatch, generalize the basic data detectors to nonunitary constellations, and extend them to systems with pilots and virtual carriers.

Paper approved by Y. Li, the Editor for Wireless Communications Theoryof the IEEE Communications Society. Manuscript received December 15, 2004; revised June 15, 2005. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, Informatics Circle of Research Excellence and Alberta Ingenuity Fund. This paper was presented in part at the IEEE Global Telecommunications Conference, November 2004, Dallas, TX.

T. Cui was with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada. He is now with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: taocui@caltech.edu).

C. Tellambura is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: chintha@ece.ualberta.ca).

Digital Object Identifier 10.1109/TCOMM.2006.873075

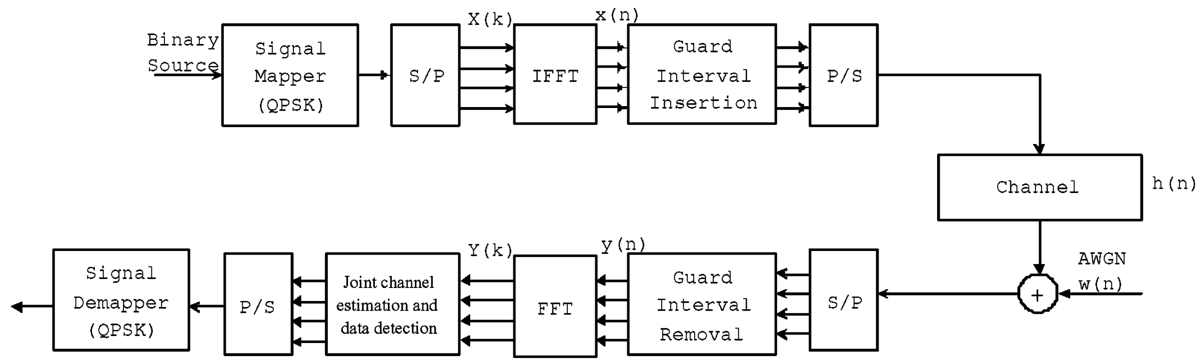

Fig. 1. Baseband OFDM system model.

This paper is organized as follows. Section II describes the basic baseband OFDM system model. Section III considers the joint channel estimation and data detection. In Section IV, we describe both V-BLAST and SD and provide several modifications for our joint data detectors. Section V generalizes the semi-blind detector considering mismatch and considers nonunitary constellations. Section VI gives simulation results, and Section VII concludes the paper.

Notation: $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^\dagger$ denote transpose, conjugate transpose, and Moore–Penrose pseudo-inverse, respectively. The set of all complex $K \times 1$ vectors is denoted by $\mathcal{C}^K$. A constellation $\mathcal{Q}$ is unitary if $|x| = 1$, for all $x \in \mathcal{Q}$, and the cardinality of $\mathcal{Q}$ is denoted by $|\mathcal{Q}|$. All $N \times 1$ vectors of elements of $\mathcal{Q}$ are denoted by $\mathcal{Q}^N$. For M-ary phase-shift keying (MPSK), the constellation $\mathcal{Q} = \{e^{j2\pi k/M}, k = 0, 1, \cdots, M-1\}$. A circularly complex Gaussian RV (CGRV) with mean μ and variance σ^2 is denoted by $z \sim \mathcal{CN}(\mu, \sigma^2)$. The discrete Fourier transform (DFT) matrix of size $N \times N$ is given by $\mathbf{F} = 1/\sqrt{N}[e^{j(2\pi/N)kl}]$, $k, l \in 0, 1, \cdots, N-1$ (where $j = \sqrt{-1}$).

II. OFDM Baseband Model

In the OFDM system given in Fig. 1, source data are grouped and/or mapped into multiphase symbols from $\mathcal{Q}$, which are modulated by the inverse DFT (IDFT) on N parallel subcarriers. Note that $X(k)$, $k = 0, 1, \ldots, N-1$ are called modulation symbols or transmit data symbols. The input symbol duration is T_s, and the OFDM symbol duration is NT_s. We assume that the composite CIR, which includes transmit and receive pulse shaping and the physical channel response between the transmitter and receiver, may be modeled as [24, p. 802]

$$h(\tau) = \sum_{l=0}^{L-1} h_l \delta(\tau - \tau_l) \tag{1}$$

where $h_l \sim \mathcal{CN}(0, E[h_l^2])$, and τ_l is the delay of the lth tap. Typically, it is assumed that $\tau_l = lT_s$, resulting in a finite impulse response filter with an effective length L. Assuming that the channel remains constant during each OFDM symbol, but may vary between OFDM symbols, and that the cyclic prefix is sufficiently long $(N_g > L)$, the post-DFT received samples $Y(k)$ are given by

$$Y(k) = H(k)X(k) + W(k), \qquad 0 \le k \le N-1 \tag{2}$$

where $H(k) = H(j2\pi k/N)$ is the complex channel frequency response at subcarrier k, $H(j\omega)$ is the Fourier transform of the CIR, and $W(k), k = 0, 1, \ldots, N-1$ are independent and identically distributed (i.i.d) CGRVs, each of which has zero mean and variance σ_n^2. Assuming $\tau_l = lT_s$, we find $\mathbf{H} = \mathbf{F}_L\mathbf{h}$, where $\mathbf{H} = [H(0), H(1), \cdots, H(N-1)]^T$, $\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}]^T$ is the CIR, and $\mathbf{F}_L$ is an $N \times L$ submatrix of the DFT matrix $\mathbf{F}$, which corresponds to each channel path. We can vectorize (2) as

$$\mathbf{Y} = \mathbf{X}_D\mathbf{F}_L\mathbf{h} + \mathbf{W} \tag{3}$$

where $\mathbf{X}_D = \mathrm{diag}\{X(0), X(1), \ldots, X(N-1)\}$ is a diagonal matrix. Note that (3) is the basis of our joint channel and data estimators.

III. Joint Channel Estimation and Data Detection

We next derive blind and semi-blind joint CIR estimators and data detectors. The semi-blind detector is derived by assuming the availability of the exact knowledge of the channel correlation matrix and noise variance.[1] We later show how the detector will behave under imperfect knowledge (i.e., mismatch conditions).

A. Blind Detector

Since the noise vector $\mathbf{W}$ in (3) is i.i.d Gaussian, the ML estimators of the CSI $(\mathbf{h})$ and transmitted symbols $(\mathbf{X}_D)$ are given by

$$(\hat{\mathbf{h}}, \hat{\mathbf{X}}_D) = \underset{\tilde{\mathbf{h}} \in \mathcal{C}^L, \tilde{\mathbf{X}}_D \in \mathcal{Q}^N}{\arg\min} \|\mathbf{Y} - \tilde{\mathbf{X}}_D\mathbf{F}_L\tilde{\mathbf{h}}\|^2. \tag{4}$$

[1]We hasten to add that our use of the term semi-blind is somewhat unconventional. Typically, semi-blind refers to the use of one or more pilots. We use the term semi-blind to indicate that the detector needs the knowledge of channel correlation and noise variance.

The minimization in (4) is a complex LS problem for $\hat{\mathbf{h}}$ and an integer LS problem for $\hat{\mathbf{X}}_D$. Given $\hat{\mathbf{X}}_D$ (we assume that $\hat{\mathbf{X}}_D = \mathbf{X}_D$), the channel response $\hat{\mathbf{h}}$ that minimizes (4) is given by the LS estimate

$$\hat{\mathbf{h}} = \left[(\hat{\mathbf{X}}_D\mathbf{F}_L)^H(\hat{\mathbf{X}}_D\mathbf{F}_L)\right]^{-1}(\hat{\mathbf{X}}_D\mathbf{F}_L)^H\mathbf{Y}. \tag{5}$$

Substituting (5) into (4), we obtain

$$\begin{aligned}\hat{\mathbf{X}}_D &= \arg\min_{\mathbf{X}_D}\left\|\mathbf{Y}-\mathbf{X}_D\mathbf{F}_L\left[(\mathbf{X}_D\mathbf{F}_L)^H(\mathbf{X}_D\mathbf{F}_L)\right]^{-1}\right.\\ &\qquad\left.\times(\mathbf{X}_D\mathbf{F}_L)^H\mathbf{Y}\right\|^2 &&(6a)\\ &= \arg\min_{\mathbf{X}_D}\\ &\quad\left\|\left[\mathbf{I}_N-\mathbf{X}_D\mathbf{F}_L\left(\mathbf{F}_L^H\mathbf{F}_L\right)^{-1}\mathbf{F}_L^H\mathbf{X}_D^H\right]\mathbf{Y}\right\|^2 &&(6b)\\ &= \arg\min_{\mathbf{X}_D}\mathbf{Y}^H\\ &\quad\times\left[\mathbf{I}_N-\mathbf{X}_D\mathbf{F}_L\left(\mathbf{F}_L^H\mathbf{F}_L\right)^{-1}\mathbf{F}_L^H\mathbf{X}_D^H\right]\mathbf{Y} &&(6c)\\ &= \arg\min_{\mathbf{X}_D}\mathbf{Y}^H\mathbf{X}_D\\ &\quad\times\left[\mathbf{I}_N-\mathbf{F}_L\left(\mathbf{F}_L^H\mathbf{F}_L\right)^{-1}\mathbf{F}_L^H\right]\mathbf{X}_D^H\mathbf{Y} &&(6d)\\ &= \arg\min_{\mathbf{x}}\mathbf{x}^T\mathbf{Y}_D^H\left[\mathbf{I}_N-\mathbf{F}_L\mathbf{F}_L^H\right]\mathbf{Y}_D\mathbf{x}^* &&(6e)\end{aligned}$$

where $\mathbf{Y}_D = \mathrm{diag}\{Y(0), Y(1), \cdots, Y(N-1)\}$, and $\mathbf{x} \in \mathcal{Q}^N$ is the vector whose elements are the diagonal elements of matrix $\mathbf{X}_D$. Equation (6b) is due to the use of the constant modulus constellation MPSK; (6c) follows from the fact that the matrix $\mathbf{I}_N - \mathbf{X}_D\mathbf{F}_L\mathbf{F}_L^H\mathbf{X}_D^H$ is an orthogonal projection matrix onto $null(\mathbf{X}_D\mathbf{F}_L)$, and the projection matrix has the property $P_\perp^2 = P_\perp$ and $P_\perp^H = P_\perp$.

The rank of the matrix $\mathbf{B} = \mathbf{Y}_D^H(\mathbf{I}_N - \mathbf{F}_L\mathbf{F}_L^H)\mathbf{Y}_D$ is only $N-L$. Note that $\mathbf{B}$ can be QR factorized as $\mathbf{B} = \mathbf{QR}$, where $\mathbf{Q}$ and $\mathbf{R}$ are unitary and upper triangular, respectively. Since the last L rows of $\mathbf{R}$ are zero, both the standard V-BLAST and SD fail here. We next modify (6e) so that both SD and V-BLAST can be applied.

Using the constant modulus property (e.g., $|X(k)|^2 = 1$ for $X(k) \in \mathcal{Q}$), $\mathbf{x}^T\mathbf{Y}_D^H\mathbf{Y}_D\mathbf{x}^* = \sum_{k=0}^{N-1}|Y(k)|^2$ is a constant. Therefore, the optimization problem (6e)[2] is equivalent to

$$\begin{aligned}\hat{\mathbf{X}}_D &= \arg\min_{\mathbf{x}\in\mathcal{Q}^N}\eta\mathbf{x}^T\mathbf{Y}_D^H\mathbf{Y}_D\mathbf{x}^* + \mathbf{x}^T\mathbf{B}\mathbf{x}^*\\ &= \arg\min_{\mathbf{x}\in\mathcal{Q}^N}\mathbf{x}^T\mathbf{Y}_D^H\left[(\eta+1)\mathbf{I}_N-\mathbf{F}_L\mathbf{F}_L^H\right]\mathbf{Y}_D\mathbf{x}^*. \quad (7)\end{aligned}$$

Since $\mathbf{F}_L\mathbf{F}_L^H$ is a positive semi-definite matrix with nonzero eigenvalue 1, $(\eta+1)\mathbf{I}_N - \mathbf{F}_L\mathbf{F}_L^H$ is a positive definite matrix if $\eta > 0$. For simplicity, we let $\eta = \sigma_n^2$.

The detector (7) can be solved via an exhaustive search over all M^N possible data sequences, a search whose complexity is exponential in N and is prohibitive for all but small N. Therefore, we develop two efficient algorithms: the first one uses V-BLAST (suboptimal), and the second one uses the SD [20] (optimal). Both algorithms exploit the Cholesky factorization of a positive definite matrix, which can be used for the blind detector (7) and the semi-blind detectors developed next. The algorithms are hence developed in detail in Section IV.

Remarks:

- The blind detector (6) is known as the GLRT detector [19]. A similar approach has been used for joint ML channel estimation and signal detection for single-input and multiple-output systems in [25], where the SD is also used for signal detection. In [18], blind and semi-blind detectors are obtained by using GLRT and a regression model. However, GLRT-based detectors are not ML. The ML joint detector will be given in the next subsection.
- After $\hat{\mathbf{X}}_D$ is estimated from (7), the LS estimate $\hat{\mathbf{h}}$ can be obtained by substituting $\hat{\mathbf{X}}_D$ into (5). Assuming that the channel remains constant over K OFDM symbols, the joint estimator is only performed every K symbols, which reduces the total complexity of the receiver.
- Both $\mathbf{x}$ and $\mathbf{x}e^{j\phi}$, where $\phi \in (0, 2\pi)$, satisfies (7), which shows that the blind detector (7) exhibits a phase ambiguity. This can be solved by using a pilot tone (see Section V-D).
- The blind detector needs knowledge of the channel length L, which can be obtained using the cyclic prefix-based algorithm [26]. If L is overestimated, we find from simulation that the performance degradation is negligible (see Section VI-D for a detailed discussion).

B. Semi-Blind Detector

This requires knowledge of the autocorrelation matrix $\mathbf{R}_h$ of the CIR $\mathbf{h}$ and the noise variance σ_n^2. We classify it as a semi-blind detector. From (3), $\mathbf{h}$ and $\mathbf{W}$ are zero-mean complex Gaussian random vectors. The received samples $Y(k)$ can be modeled as i.i.d. zero-mean CGRVs that are conditional on the data vector. The autocorrelation matrix of the received signal is thus given by

$$\begin{aligned}\mathbf{R}_Y &= E\{\mathbf{Y}\mathbf{Y}^H\} = \mathbf{X}_D\mathbf{F}_L\mathbf{R}_h\mathbf{F}_L^H\mathbf{X}_D^H + \sigma_n^2\mathbf{I}_N\\ &= \mathbf{X}_D\left(\mathbf{F}_L\mathbf{R}_h\mathbf{F}_L^H + \sigma_n^2\mathbf{I}_N\right)\mathbf{X}_D^H. \end{aligned}\tag{8}$$

The determinant of $\mathbf{R}_Y$ can be expressed as

$$\begin{aligned}\det(\mathbf{R}_Y) &= \det(\mathbf{X}_D)\det\left(\mathbf{F}_L\mathbf{R}_h\mathbf{F}_L^H + \sigma_n^2\mathbf{I}_N\right)\det\left(\mathbf{X}_D^H\right)\\ &= \det\left(\mathbf{F}_L\mathbf{R}_h\mathbf{F}_L^H + \sigma_n^2\mathbf{I}_N\right). \end{aligned}\tag{9}$$

Note that the determinant of $\mathbf{R}_Y$ is independent of $\mathbf{X}_D$. Ignoring terms that are independent of $\mathbf{X}_D$, the log-likelihood function is given by

$$\Lambda(\mathbf{Y}|\mathbf{X}_D) = -\mathbf{Y}^H\mathbf{R}_Y^{-1}\mathbf{Y}. \tag{10}$$

As with (6e), maximizing the log-likelihood function is equivalent to solving

$$\hat{\mathbf{X}}_D = \arg\min_{\mathbf{x}\in\mathcal{Q}^N}\mathbf{x}^T\mathbf{Y}_D^H\left(\mathbf{F}_L\mathbf{R}_h\mathbf{F}_L^H + \sigma_n^2\mathbf{I}_N\right)^{-1}\mathbf{Y}_D\mathbf{x}^*. \tag{11}$$

[2]In [15], an approximate iterative LS projection algorithm is developed to solve optimization problems similar to (6e). However, the convergence of that algorithm is not guaranteed.

As the matrix $(\mathbf{F}_L\mathbf{R}_h\mathbf{F}_L^H+\sigma_n^2\mathbf{I}_N)^{-1}$ is positive definite, we propose the use of V-BLAST [23] and the SD algorithm [20] to solve (11). This is developed in detail next.

IV. Data Detection Algorithms

In this section, we review the SD and V-BLAST algorithm, which efficiently solve our blind and semi-blind detectors. We also give a new efficient variant of SD that can handle any constellations.

A. V-BLAST Detection

The blind detector (7) and the semi-blind detector (11) can be solved via the V-BLAST algorithm [23]. These detectors can be written in a general form as

$$\hat{\mathbf{X}}_D = \arg\min_{\mathbf{x}\in\mathcal{Q}^N} \mathbf{x}^T\mathbf{G}\mathbf{x}^* = \|\mathbf{M}\mathbf{x}^*\|^2 \tag{12}$$

where $\mathbf{G}$ is a positive definite matrix, which can be Cholesky factored as $\mathbf{G}=\mathbf{M}^H\mathbf{M}$. The V-BLAST ordering finds a permutation matrix $\mathbf{\Pi}$ such that the QR decomposition $\mathbf{QR}=\mathbf{M}'=\mathbf{M\Pi}$, where $\mathbf{Q}$ is unitary and $\mathbf{R}=[r_{ij}]$ is upper triangular, has the property that $\min_{1\le i\le N} r_{ii}$ is maximized over all column permutations. For $k=N,N-1,\ldots,1$, the algorithm chooses $\pi(k)$ such that

$$\pi(k) = \arg\min_{j\notin\{\pi(1),\ldots,\pi(k-1)\}} \|(\mathbf{G}_k)_j\|^2 \tag{13}$$

where $(\mathbf{G}_k)_j$ is the jth row of $\mathbf{G}_k$, $\mathbf{G}_k$ is the pseudo inverse of $\mathbf{M}_k$, and $\mathbf{M}_k$ denotes the matrix obtained by zeroing columns $\pi(1),\ldots,\pi(k-1)$ of $\mathbf{M}$. Equation (12) can be expressed as

$$\hat{\mathbf{X}}_D = \arg\min_{\mathbf{x}\in\mathcal{Q}^N} \|\mathbf{R}\mathbf{x}^*\|^2. \tag{14}$$

Since $\mathbf{R}$ is upper triangular, the kth element of (14) is given by

$$\hat{x}_k = \arg\min_{x_k,\ldots,x_N}\left(r_{kk}x_k^* + \sum_{i=k+1}^{N} r_{ki}x_i^*\right)^2 \tag{15}$$

where r_{ki} is the (k,i)th entry of matrix $\mathbf{R}$, and the estimate is free of interference from subcarriers $1,2,\ldots,k-1$. Thus, x_k can be estimated by minimizing (15). Proceeding in the order $x_N,x_{N-1},\ldots,x_1$ and assuming correct previous decisions, we can cancel the interference between subcarriers in each step and estimate $\mathbf{x}$. This sequential detection suffers from error propagation, even with optimal ordering. Note that the V-BLAST solution may be used as a starting point for the SD.

B. SD Algorithm

1) Algorithm: The original Fincke and Phost [20] SD (FPSD) has been used for ML detection of lattice codes [21], for detection in multiple-antenna wireless communication systems [27]. The SD is an efficient method to find the optimal solution to an integer LS problem [20], [21], [28]. This is the problem of finding the closest lattice point in N dimensions to a given point $\mathbf{x}\in\mathcal{C}^N$. In our case, the search space has M^N lattice points, so the exhaustive search is only possible for small N only. To the best of the authors' knowledge, the SD has not been applied for joint channel estimation and data detection in OFDM systems.

While the FPSD can only handle real systems, complex systems can readily be decoupled to formulate real systems. The complex problem (12) can be transformed into the real matrix equation

$$\tilde{\mathbf{X}}_D = \arg\min_{\tilde{\mathbf{x}}\in\mathcal{A}^{2N}} \tilde{\mathbf{x}}^T\tilde{\mathbf{M}}^T\tilde{\mathbf{M}}\mathbf{x} \tag{16}$$

where

$$\begin{aligned}\tilde{\mathbf{x}} &= \begin{bmatrix}\mathrm{Re}\{\mathbf{x}^T\} & -\mathrm{Im}\{\mathbf{x}^T\}\end{bmatrix}\\ \tilde{\mathbf{M}} &= \begin{bmatrix}\mathrm{Re}\{\mathbf{M}^T\} & -\mathrm{Im}\{\mathbf{M}^T\}\\ \mathrm{Im}\{\mathbf{M}^T\} & \mathrm{Re}\{\mathbf{M}^T\}\end{bmatrix}\end{aligned} \tag{17}$$

and $\mathcal{A}=\{\mathrm{Re}(\mathcal{Q}),\mathrm{Im}(\mathcal{Q})\}$. Thus, (16) can be solved via the SD. For example, if $\mathbf{x}$ belongs to 4QAM, each entry of $\tilde{\mathbf{x}}$ belongs to BPSK. Therefore, the 4QAM complex system becomes a BPSK real system. Any M-QAM can be decoupled similarly. Thus, (12) can be reduced to a real system (16).

Since the matrix $\tilde{\mathbf{M}}^T\tilde{\mathbf{M}}$ is positive definite, Cholesky factorization of it yields $\tilde{\mathbf{M}}^T\tilde{\mathbf{M}}=\mathbf{R}^H\mathbf{R}$, where $\mathbf{R}$ is an upper triangular matrix with r_{ii} real and positive. We then have

$$\mathbf{s}^T\mathbf{R}^T\mathbf{R}\mathbf{s} = \sum_{i=1}^{N}\left| r_{ii}s_i + \sum_{j=i+1}^{N} r_{ij}s_j\right|^2 \le r^2 \tag{18}$$

where $\mathbf{s}=[\mathrm{Re}\{\mathbf{x}^T\},\mathrm{Im}\{\mathbf{x}^T\}]^T$. The SD searches the lattice inside a hypersphere of radius r instead of searching the whole lattice. We will discuss the choice of r later. Substituting $q_{ii}=r_{ii}^2$ for $i=1,\cdots,N$ and $q_{ij}=r_{ij}/r_{ii}$ for $j=i+1,\cdots,N$, we can get

$$\mathbf{s}^T\mathbf{R}^T\mathbf{R}\mathbf{s} = \sum_{i=1}^{N} q_{ii}\left| s_i + \sum_{j=i+1}^{N} q_{ij}s_j\right|^2 \le r^2. \tag{19}$$

The key idea of the SD is to generate $s_N,\ldots,s_1$ so that the bound in (19) is progressively satisfied, i.e., a range for s_N is determined by ignoring $i=2,\ldots,N$ terms in (19). Once this range is established, it is used to find the range for s_{N-1} by ignoring $i=3,\ldots,N$ terms in (19). This process continues until all of the candidate vectors are generated; starting from s_N and working backward, for s_i, we have

$$q_{ii}\left| s_i + \sum_{j=i+1}^{N} q_{ij}s_j\right|^2 \le r^2 - \sum_{l=i+1}^{N} q_{ll}\left| s_l + \sum_{j=l+1}^{N} q_{lj}s_j\right|^2. \tag{20}$$

If s_i and q_{ij} are real, the bound of s_i can be obtained as (21), which is shown at the bottom of the following page, where $\lceil\cdot\rceil$

denotes the smallest integer greater than or equal to its argument. $\lfloor\cdot\rfloor$ denotes the largest integer less than or equal to its argument. Therefore, s_i is selected from

$$s_i \in [\mathrm{LB}_i, \mathrm{UB}_i] \bigcap \mathcal{A} \quad i = 1, 2, \ldots, N. \tag{22}$$

Let $\mathcal{L}$ be the list of all candidates generated at the end of this process. If $\mathcal{L}$ is empty, r is increased, and the process repeats. Note that $|\mathcal{L}| \ll |\mathcal{A}|^N$. Now we can compute $\mathbf{s}^T\mathbf{R}^T\mathbf{R}\mathbf{s}$ for all $\mathbf{s} \in \mathcal{L}$, and the resulting minimum yields the ML estimate. This basically sums up the original SD idea. One can immediately try to improve this. For example, when going through the list $\mathcal{L}$, say, at $\mathbf{s}'$, if $C = \mathbf{s}'^T\mathbf{R}^T\mathbf{R}\mathbf{s}'$ is less than r^2, then the radius r is updated to $\sqrt{C}$. In [27], if $C < r^2$, the bounds of all s_i's (21) are updated. Although this accelerates the original SD, its complexity still depends on the initial radius. To further reduce the complexity, one can use [28] and [29]. In the FPSD, the search is started at the surface of the sphere, which is inefficient. This improved algorithm [28], [29] begins its search near the center of the sphere. As above, the radius r can be reduced more rapidly compared with the FPSD. Furthermore, the complexity of the improved algorithm is not highly sensitive to the initial radius [28]. Adapting their idea to the current problem, we need to arrange the search order according to the value of $T_i = |s_l + \sum_{j=l+1}^{N} q_{lj}s_j|^2$. We search the lattice point with smaller T_i first. This is the modified SD algorithm for OFDM systems.

However, not all M-PSK constellations can be decoupled into real systems (e.g., 8-PSK). Hochwald and Brink [30] show a modified sphere decoder that handles complex constellations, but it involves the computationally inefficient $\cos^{-1}$ operation, slowing down the SD. We show here that the decouple algorithm can still be used to handle M-PSK, when $\mathcal{A}$ is not an integer set. However, each element of $\mathbf{x}$ is constrained depending on the constellation. For a given $\mathrm{Re}(x_i)$, the candidates for $\mathrm{Im}(x_i)$ are hence constrained. Let $\mathcal{Q}_R = \{\mathrm{Re}(x) | x \in \mathcal{Q}\}$. Let $\mathcal{Q}_I(x) = \{\mathrm{Im}(x) | x \in \mathcal{Q}, x \in \mathcal{Q}_R\}$. Therefore, s_i is selected from

$$s_i \in \begin{cases} [\mathrm{LB}_i, \mathrm{UB}_i] \cap \mathcal{Q}_R & i = 2k, k = 0, 1, \ldots, N-1 \\ [\mathrm{LB}_i, \mathrm{UB}_i] \cap \mathcal{Q}_I(s_{i-1}) & i = 2k+1, \\ & k = 0, 1, \ldots, N-1 \end{cases}. \tag{23}$$

Since $Q_I(s)$ can be precomputed for each s from $\mathcal{Q}_R$ and be stored in memory, additional computational complexity is avoided. This simple idea can be used to handle any constellations. The decoupling for M-QAM can be viewed as a special case of our generalization because $Q_I(s)$'s are the same for any s from $\mathcal{Q}_R$, and $\mathcal{Q}_R = \mathcal{Q}_I$.

2) Initial Radius: An important problem is how to choose the initial radius. If it is too small, no candidate point would be found inside the sphere. However, if it is too large, too many candidates would be found, and the overall complexity would become exponential. In [22], an initial radius is chosen according to the variance of the additive noise. In contrast, we use a relaxation approach to choose the initial radius. For OFDM symbols with MPSK, we relax (7) and (11) as $\min \mathbf{x}^H\mathbf{G}\mathbf{x}$, $\mathbf{x}^H\mathbf{x} = N$, where the vector $\mathbf{x} \in \mathcal{C}^N$. The Lagrangian $\mathcal{L}(\mathbf{x}, \lambda)$ for this minimization problem is $\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{x}^H\mathbf{G}\mathbf{x} + \lambda(\mathbf{x}^H\mathbf{x} - N)$. The optimal λ here is the maximum eigenvalue of matrix $\mathbf{G}$, and $\tilde{\mathbf{x}}$ is the eigenvector corresponding to λ. We quantize $\tilde{\mathbf{x}}$ into a point in $\mathcal{Q}^N$ as $\hat{\mathbf{x}}$. By substituting $\hat{\mathbf{x}}$ into (12), the initial radius is given by $r^2 = \hat{\mathbf{x}}^H\mathbf{G}\hat{\mathbf{x}}$.

V. Generalizations

A. Semi-Blind Detector Under Mismatch

The detector (11) needs knowledge of the channel covariance matrix $\mathbf{R}_h$ and the noise variance σ_n^2. In [31], it is shown that $\mathbf{R}_h$ remains constant during 300 OFDM symbols for a given OFDM system. In [26], we give a cyclic prefix-based algorithm to estimate $\mathbf{R}_h$ and σ_n^2 without pilots. If $\mathbf{R}_h$ and σ_n^2 are not available or there exists a residual estimation error, we design the estimators for $\tilde{\mathbf{R}}_h$ and $\tilde{\sigma}_n^2$, while the true values are $\mathbf{R}_h$ and σ_n^2. In [3] and [5], an estimator that is suficiently robust to the mismatch is designed for the worst case, which is taken to be the uniform power delay profile (UPDP). We follow their approach and consider a suboptimal criterion as follows:

$$\begin{aligned} &\mathrm{E}\left\{\mathbf{Y}^H\hat{\mathbf{X}}_D\left(\mathbf{F}_L\tilde{\mathbf{R}}_h\mathbf{F}_L^H + \tilde{\sigma}_n^2\mathbf{I}\right)^{-1}\hat{\mathbf{X}}_D^H\mathbf{Y}\right\} \\ &\quad= \mathrm{E}\Bigg\{\mathbf{h}^H\mathbf{F}_L^H\mathbf{X}_D^H\hat{\mathbf{X}}_D\left(\mathbf{F}_L\tilde{\mathbf{R}}_h\mathbf{F}_L^H + \tilde{\sigma}_n^2\mathbf{I}\right)^{-1}\hat{\mathbf{X}}_D^H\mathbf{X}_D\mathbf{F}_L\mathbf{h} \\ &\quad\quad+ \mathbf{W}^H\hat{\mathbf{X}}_D\left(\mathbf{F}_L\tilde{\mathbf{R}}_h\mathbf{F}_L^H + \tilde{\sigma}_n^2\mathbf{I}\right)^{-1}\hat{\mathbf{X}}_D^H\mathbf{W}\Bigg\}. \end{aligned} \tag{24}$$

The second equality comes from (3). If the first term of (24) is independent of $\tilde{\mathbf{R}}_h$, then the bit error rate (BER) may be less dependent on the mismatch. Ignoring the noise variance σ_n^2 and letting $\Delta\mathbf{X} = \hat{\mathbf{X}}_D^H\mathbf{X}_D$, the first term of (24) can be written as

$$\begin{aligned} &\mathrm{E}\left\{\mathbf{h}^H\mathbf{F}_L^H\mathbf{X}_D^H\hat{\mathbf{X}}_D\left(\mathbf{F}_L\tilde{\mathbf{R}}_h\mathbf{F}_L^H + \tilde{\sigma}_n^2\mathbf{I}\right)^{-1}\hat{\mathbf{X}}_D^H\mathbf{X}_D\mathbf{F}_L\mathbf{h}\right\} \\ &\quad\simeq \mathrm{E}\left\{\mathbf{h}^H\mathbf{F}_L^H\Delta\mathbf{X}^H\left(\mathbf{F}_L\tilde{\mathbf{R}}_h\mathbf{F}_L^H\right)^{-1}\Delta\mathbf{X}\mathbf{F}_L\mathbf{h}\right\} \\ &\quad= \mathrm{E}\left\{\mathbf{h}^H\mathbf{F}_L^H\Delta\mathbf{X}^H\mathbf{F}_L(\tilde{\mathbf{R}}_h)^{-1}\mathbf{F}_L^H\Delta\mathbf{X}\mathbf{F}_L\mathbf{h}\right\}. \end{aligned} \tag{25}$$

$$\underbrace{\left\lceil -\sqrt{\frac{r^2 - \sum_{l=i+1}^{N} q_{ll}\left|s_l + \sum_{j=l+1}^{N} q_{lj}s_j\right|^2}{q_{ii}}} - \sum_{j=i+1}^{N} q_{ij}s_j \right\rceil}_{\mathrm{LB}_i} \le s_i \le \underbrace{\left\lfloor \sqrt{\frac{r^2 - \sum_{l=i+1}^{N} q_{ll}\left|s_l + \sum_{j=l+1}^{N} q_{lj}s_j\right|^2}{q_{ii}}} - \sum_{j=i+1}^{N} q_{ij}s_j \right\rfloor}_{\mathrm{UB}_i} \tag{21}$$

Since $\mathbf{F}_L^H \Delta \mathbf{X} \mathbf{F}_L$ is a circulant matrix, let the first row be $a(0), a(1), \cdots, a(L-1)$. Equation (25) becomes

$$\sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{|a((i+j) \bmod L)|^2}{\tilde{\mathbf{R}}_h(j)} \mathbf{R}_h(i). \tag{26}$$

If $\tilde{\mathbf{R}}_h(j)$ is UPDP or $\tilde{\mathbf{R}}_h(j) = 1/L$, (26) is equal to $L \sum_{i=0}^{L-1} |a(i)|^2$, which is independent of $\tilde{\mathbf{R}}_h$. Hence, the detector is robust to mismatch with this choice of $\tilde{\mathbf{R}}_h$.

B. Nonunitary Constellations

The blind and semi-blind detectors can be extended to nonunitary constellations. Instead of solving the optimal detector (4), the suboptimal blind detector solves

$$(\hat{\mathbf{h}}, \hat{\mathbf{X}}_D) = \underset{\tilde{\mathbf{h}} \in \mathcal{C}^L, \tilde{\mathbf{X}}_D \in \mathcal{Q}^N}{\arg\min} \left\| \tilde{\mathbf{X}}_D^{-1} \mathbf{Y} - \mathbf{F}_L \tilde{\mathbf{h}} \right\|^2 \tag{27}$$

where $\mathcal{Q}$ denotes the nonunitary constellation. The LS estimate of $\mathbf{h}$ is given by

$$\hat{\mathbf{h}} = \mathbf{F}_L^H \hat{\mathbf{X}}_D^{-1} \mathbf{Y}. \tag{28}$$

Substituting (28) into (27), we obtain

$$\begin{aligned} \hat{\mathbf{X}}_D &= \underset{\mathbf{X}_D}{\arg\min}\, \mathbf{Y}^H \mathbf{X}_D^{-H} \left(\mathbf{I}_N - \mathbf{F}_L \mathbf{F}_L^H \right) \mathbf{X}_D^{-1} \mathbf{Y} \\ &= \underset{\mathbf{x} \in (\mathcal{Q}^{-1})^N}{\arg\min}\, \mathbf{x}^H \mathbf{Y}_D^H \left(\mathbf{I}_N - \mathbf{F}_L \mathbf{F}_L^H \right) \mathbf{Y}_D \mathbf{x} \end{aligned} \tag{29}$$

where $\mathcal{Q}^{-1} = \{x^{-1} | x \in \mathcal{Q}\}$. Equation (29) can also be solved using V-BLAST and SD, but there is a performance loss due to the suboptimal nature of (27).

To adapt the semi-blind detector to the nonunitary case, we note that (8) is

$$\mathbf{R}_Y = \mathbf{X}_D \mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H \mathbf{X}_D^H + \sigma_n^2 \mathbf{I}_N. \tag{30}$$

Maximizing the log-likelihood function is equivalent to solving

$$\begin{aligned} \hat{\mathbf{X}}_D &= \underset{\mathbf{X}_D}{\arg\min}\, \mathbf{Y}^H \left(\mathbf{X}_D \mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H \mathbf{X}_D^H + \sigma_n^2 \mathbf{I}_N \right)^{-1} \mathbf{Y} \\ &= \underset{\mathbf{X}_D}{\arg\min}\, \mathbf{Y}^H \mathbf{X}_D^{-H} \\ &\quad \times \left(\mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H + \sigma_n^2 \left(\mathbf{X}_D \mathbf{X}_D^H \right)^{-1} \right)^{-1} \mathbf{X}_D^{-1} \mathbf{Y}. \end{aligned} \tag{31}$$

As in [3], we derive a suboptimal detector by replacing the term $(\mathbf{X}_D \mathbf{X}_D^H)^{-1}$ in (31) with its expectation $E\{(\mathbf{X}_D \mathbf{X}_D^H)^{-1}\}$. Assuming the same constellation on all of the subcarriers, we have $E\{(\mathbf{X}_D \mathbf{X}_D^H)^{-1}\} = \rho \mathbf{I}_N$, where $\rho = E\{1/|x_k|^2\}$. Therefore, the suboptimal semi-blind detector is given by

$$\hat{\mathbf{X}}_D = \underset{\mathbf{x} \in (\mathcal{Q}^{-1})^N}{\arg\min}\, \mathbf{x}^H \mathbf{Y}_D^H \left(\mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H + \rho \sigma_n^2 \mathbf{I}_N \right)^{-1} \mathbf{Y}_D \mathbf{x}. \tag{32}$$

However, for an optimal solution of (31), we note that $\mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H$ is positive semi-definite. It can be readily verified that $\mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H + \sigma_n^2 (\mathbf{X}_D \mathbf{X}_D^H)^{-1}$ is positive definite. We use the following definition from [32, p. 469].

Definition: Let $\mathbf{A}$ and $\mathbf{B}$ be Hermitian matrices. We write $\mathbf{A} \succeq \mathbf{B}$ if the matrix $\mathbf{A} - \mathbf{B}$ is positive semi-definite. Similarly, $\mathbf{A} \succ \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive definite.

Let $\mathbf{A} = \mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H + \sigma_n^2 (\mathbf{X}_D \mathbf{X}_D^H)^{-1}$ $\mathbf{B} = \mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H + \rho_m \sigma_n^2 \mathbf{I}_N$, where $\rho_m = \max\{1/|x_k|^2\}$. It can be readily verified that $\mathbf{B} \succeq \mathbf{A}$. Using Corollary 7.7.4 of [32, p. 471], we can obtain $\mathbf{A}^{-1} \succeq \mathbf{B}^{-1}$. Therefore, for any $\mathbf{x} \in (\mathcal{Q}^{-1})^N$, we have

$$\mathbf{x}^H \mathbf{Y}_D^H \mathbf{B}^{-1} \mathbf{Y}_D \mathbf{x} \leq \mathbf{x}^H \mathbf{Y}_D^H \mathbf{A}^{-1} \mathbf{Y}_D \mathbf{x}. \tag{33}$$

When applying sphere decoding for (31), one finds the minimum point among all of the points satisfying

$$\mathbf{x}^H \mathbf{Y}_D^H \mathbf{A}^{-1} \mathbf{Y}_D \mathbf{x} \leq r^2. \tag{34}$$

Using (33), we find the point minimizing (31) among all of the points satisfying

$$\mathbf{x}^H \mathbf{Y}_D^H \mathbf{B}^{-1} \mathbf{Y}_D \mathbf{x} \leq r^2. \tag{35}$$

Hence, the SD should be modified when the search goes to the bottom of the search tree; the radius r is updated according to $\mathbf{x}^H \mathbf{Y}_D^H \mathbf{A}^{-1} \mathbf{Y}_D \mathbf{x}$ instead of $\mathbf{x}^H \mathbf{Y}_D^H \mathbf{B}^{-1} \mathbf{Y}_D \mathbf{x}$. This gives the optimal solution.

C. Iterative Improvement of Channel Estimates Via Decision Feedback

Our proposed detectors may also be used for channel tracking if the channel remains constant for K OFDM symbols. For $p = 1$ (the first symbol), the initial channel estimate $\hat{\mathbf{h}}^1$ is given by (5). For the remaining OFDM symbols ($p = 2, \cdots, K$), channel estimation may not be necessary. Instead, $\hat{\mathbf{h}}^1$ is used to detect $p = 2, \ldots, K$ OFDM symbols. Decision-feedback-type iterations can also be used to track a slowly varying channel. From (3), the channel estimate $\mathbf{h}^{p+1}$ in the $p+1$th iteration of the decision feedback loop is given by

$$\hat{\mathbf{h}}^{p+1} = \left[\left(\hat{\mathbf{X}}_D^p \mathbf{F}_L \right)^H \left(\hat{\mathbf{X}}_D^p \mathbf{F}_L \right) \right]^{-1} \left(\hat{\mathbf{X}}_D^p \mathbf{F}_L \right)^H \mathbf{Y}_p \tag{36}$$

where $\hat{\mathbf{X}}_D^p$ and $\mathbf{Y}_p$ denote the estimated and received symbols in the pth iteration. $\hat{\mathbf{X}}_D^{p+1}$ is obtained by blind or semiblind detectors.

D. Extension to OFDM Systems With Pilot Symbols and Virtual Carriers

Existing OFDM standards such as IEEE802.11a incorporate pilot symbols [33]. These pilots can be used to reduce the search space and to solve the phase ambiguity.

With the presence of pilots, some elements of the sequence $\mathbf{x} = \{X(0), X(1), \ldots, X(N-1)\}$ are known *a priori*. Let

$X(k)$, $k \in I_p$ denote the pilot symbols, and I_p is the index set of pilot symbols with $|I_p| = P$. Let $\mathbf{x}_p = X(k)\ k \in I_p$ denote the pilot symbols vector, and $\mathbf{x}_d$ denote the data symbols vector. We rearrange $\mathbf{x}$ so that the last P symbols are $\mathbf{x}_p$ and form a new vector $\mathbf{x} = [\mathbf{x}_d^T, \mathbf{x}_p^T]^T$. Since the last P symbols are known *a priori*, the search space is limited to $|\mathcal{Q}|^{N-P}$. The ambiguity of the detectors is automatically eliminated by the pilot symbols.

We consider an OFDM system with N subcarriers, of which N_d are modulated by the data symbols, and the remaining $N_v = N - N_d$ are unmodulated virtual carriers. For simplicity, we assume that the subcarriers 1 to N_d are used for data. The blind detector (7) can be rewritten as

$$\hat{\mathbf{X}}_D = \arg\min_{\mathbf{x} \in \mathcal{Q}^{N_d}} \mathbf{x}^T \mathbf{Y}_D^H \left(\mathbf{F}_{N_d,L} \mathbf{R}_h \mathbf{F}_{N_d,L}^H + \sigma_n^2 \mathbf{I}_{N_d} \right)^{-1} \mathbf{Y}_D \mathbf{x}^*. \tag{37}$$

The semi-blind detector (1) can be rewritten as

$$\hat{\mathbf{X}}_D = \arg\min_{\mathbf{x} \in \mathcal{Q}^{N_d}} \mathbf{x}^T \mathbf{Y}_D^H \left(\mathbf{F}_{N_d,L} \mathbf{R}_h \mathbf{F}_{N_d,L}^H + \sigma_n^2 \mathbf{I}_{N_d} \right)^{-1} \mathbf{Y}_D \mathbf{x}^*. \tag{38}$$

The SD can also be applied to the virtual carrier cases.

VI. Simulation Results

We consider a frequency-selective slow Rayleigh fading channel with L Gaussian complex taps h_l with $\sigma_l^2 = E[|h_l|^2] = \sigma_0^2 e^{-l/5}$ for $l = 1, \cdots, L$, and the six-tap COST 207 TU channel model [34], which has the delay profile {0.0, 0.2, 0.5, 1.6, 2.3, 5.0} μs and power profile {0.189,0.379,0.239, 0.095,0.061,0.037}. The channel output SNR is E_b/N_0. An OFDM system with 32 subcarriers and BPSK is simulated (The carrier frequency of the OFDM system is 5 GHz and the data rate is 3 MHz. The guard interval is $N_g = 8$). A training symbol is transmitted at the Nth subcarrier to solve the scaling ambiguity. The performance of one-tap equalization with perfect knowledge of the CIR—perfect channel state information (CSI)—provides the benchmark.

A. Estimators' BER, Mean Square Error, and Complexity Dependence on SNR

Proposed detectors are tested on OFDM systems with the above simulation parameters under different SNR over a 6-ary exponential PDP channel. Fig. 2 shows the mean square error (MSE) of channel estimation, which is defined as

$$\text{MSE} = E\left\{ \sum_{l=1}^{L} |h_l - \hat{h}_l|^2 \right\}. \tag{39}$$

The semi-blind detector (11) with the SD has MSE performance identical to that of the blind detector (7) with the SD. In high SNR, the semi-blind detector with V-BLAST performs close to that with the SD, while the blind detector with V-BLAST still has a 1.2-dB gap over that with the SD at MSE $= 5 \times 10^{-4}$.

In Fig. 3, the BER performance of the OFDM system is compared with that of the benchmark. Both detectors with the SD are within 0.5 dB of the benchmark in high SNR. The performance

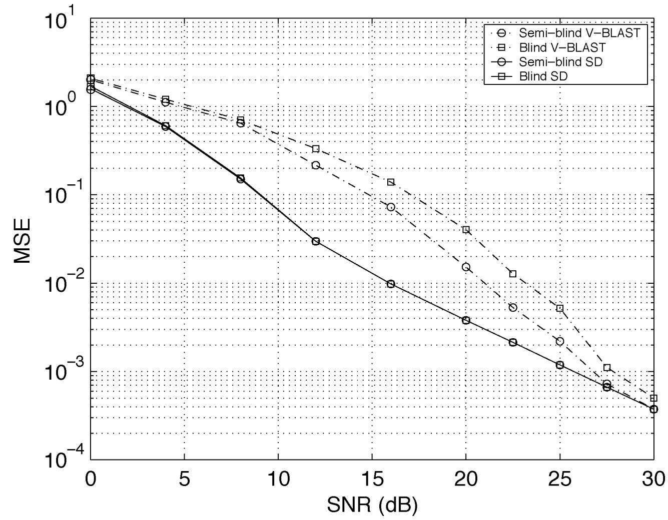

Fig. 2. MSE of the joint ML estimation of the channel response versus SNR for an OFDM system with $N = 32$ and BPSK in a 6-ray exponential PDP channel.

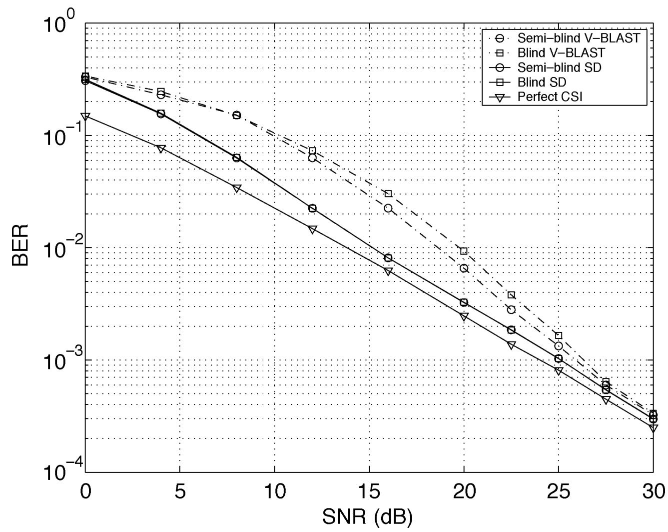

Fig. 3. BER of the joint ML estimation algorithm versus SNR for an OFDM system with $N = 32$ and BPSK in a 6-ray exponential PDP channel.

of V-BLAST detection for the semi-blind detector is comparable with that of SD in high SNR. In low SNR, the gap between SD and V-BLAST can be as large as 5 dB. The average computational complexity as a function of the SNR is given in Fig. 4. Note that the complexity of the exhaustive search is 1.81×10^{13} flops while, even for an SNR of 10 dB, all detectors' complexities are within 10^6 flops by using SD. The computational time can be saved significantly. The complexity of both detectors increases with the increase of SNR. At 0 dB, the semi-blind detector is 32 times faster than the blind detector. The blind detector has higher complexity than the semi-blind detector in low SNR. This is possibly due to the inherent rank deficiency in (6e) while the complexity is greatly reduced compared with the exhaustive search in the first $N - L$ variables [27] in the blind detector. When the SNR is larger than 25 dB, both the semi-blind detector and the blind detector have identical complexity. The

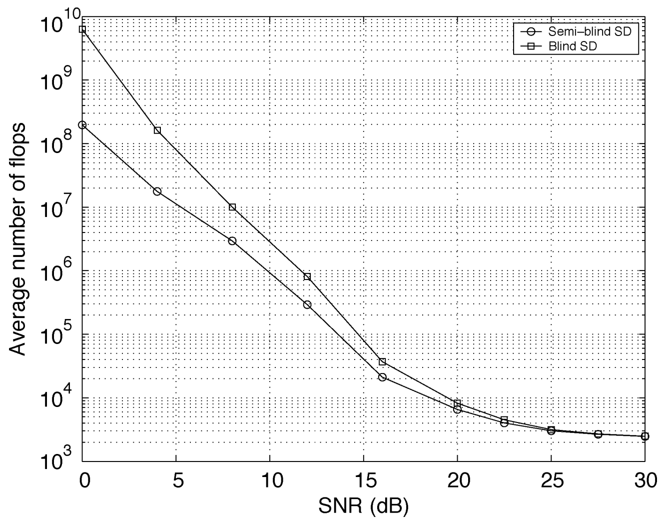

Fig. 4. Computational complexity versus SNR for an OFDM system with $N = 32$ and BPSK in a 6-ray exponential PDP channel.

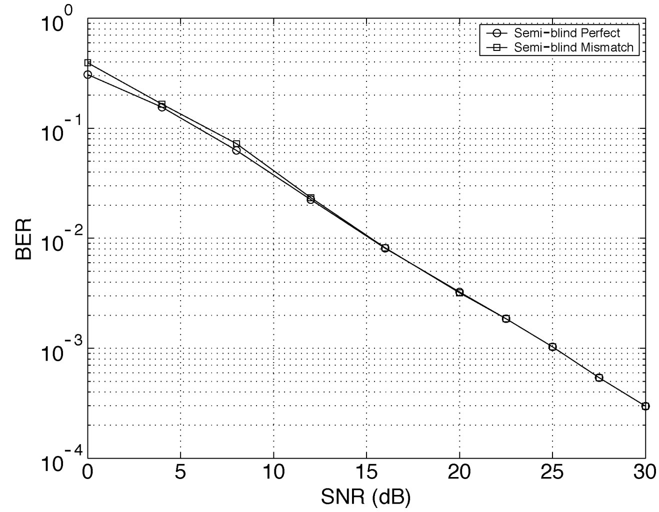

Fig. 6. Effects of semi-blind detector design mismatch in an OFDM system with $N = 32$ and BPSK. The channel is simulated using an exponential PDP. However, in the semi-blind detector, the uniform PDP is assumed.

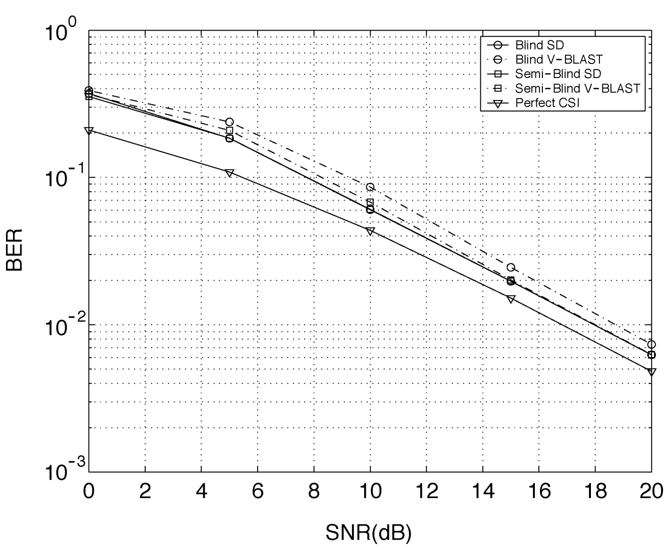

Fig. 5. BER of the joint ML estimation algorithm versus SNR for an OFDM system with $N = 32$ and BPSK in a TU channel.

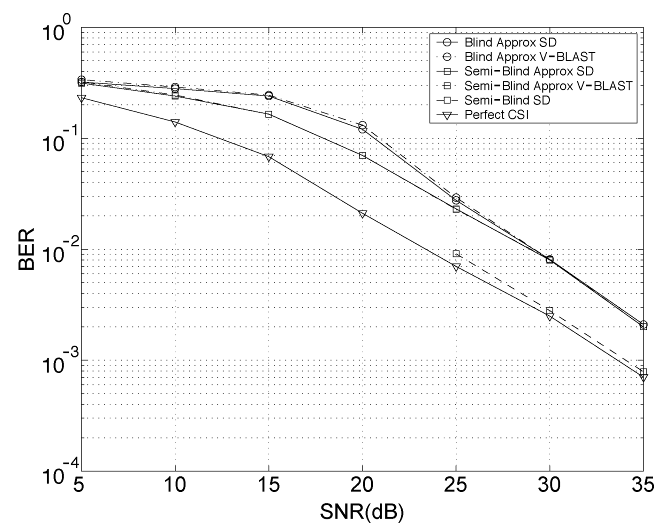

Fig. 7. BER of the joint ML estimation algorithm versus SNR for an OFDM system with $N = 32$ and 4PAM in a 3-ray exponential PDP channel.

semi-blind detector is preferable in low SNR when the channel statistics are known at the receiver.

We also compare the different algorithms over the COST 207 TU channel model, described above. The channel is assumed to be constant for 100 OFDM symbols. The channel is estimated using the first OFDM symbol, and the remaining 99 OFDM symbols are detected using the channel estimate. Fig. 5 shows the BER of V-BLAST and SD detection for the semi-blind detector and SD detection for the blind detector. The blind and semi-blind detectors almost achieve the bound given by one-tap equalization with perfect CIR. As shown in Fig. 3, the V-BLAST detection for the semi-blind detector is comparable to that of SD. This result seems to contradict the results given in [35], where great performance improvement is achieved by using the SD. The reason may be that the order of the constellation is not very high.

B. Mismatch

The effect of semi-blind design mismatch is shown in Fig. 6. The semi-blind detector is designed for UPDP and SNR $=$ 20 dB and evaluated for a 6-ary exponentially decaying power-delay profile. The BER of the robust design is compared with perfect $\mathbf{R}_h$ and σ_n^2. From the figure, the BER performance of the two detectors are almost the same. This figure confirms the robust design criteria.

C. Nonunitary Constellation

The performance of the blind suboptimal detector with SD and V-BLAST, the semi-blind suboptimal detector with SD and V-BLAST, and the semi-blind optimal detector with modified SD are compared with that of the benchmark in Fig. 7 for an OFDM system with $N = 32$ and the 4PAM constellation. The suboptimal detectors are denoted by "Approx" in the figure. The suboptimal blind and semi-blind detectors with SD and V-BLAST perform close in high SNR. However, all of them

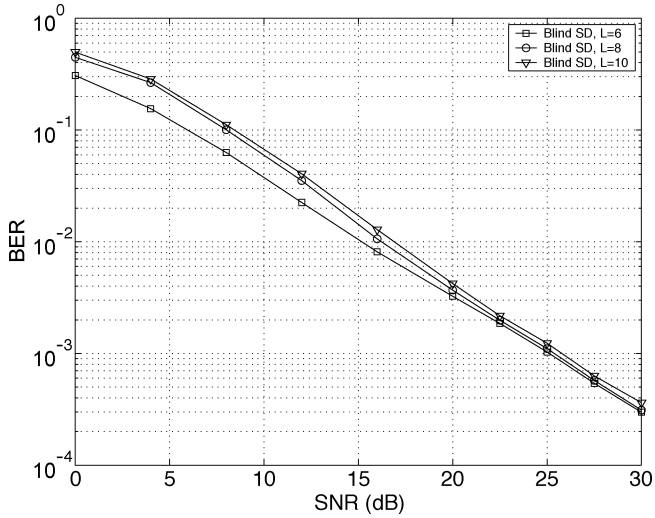

Fig. 8. BER versus SNR for a BPSK OFDM system with $N = 32$ and assuming different channel length L.

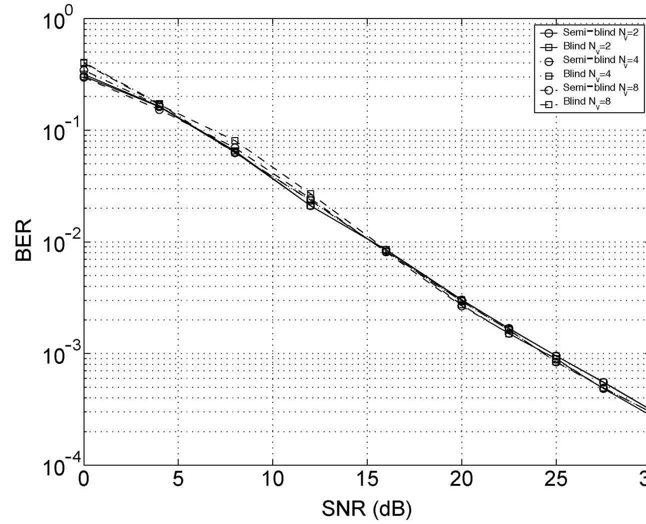

Fig. 9. BER versus SNR for an BPSK OFDM system with $N = 32$ and different number of VCs.

have a 4-dB performance loss at BER $= 2 \times 10^{-3}$. The optimal semi-blind detector performs close to the benchmark in high SNR. At BER $= 2 \times 10^{-3}$, it has only a 0.5-dB loss. In the figure, we only plot the optimal semi-blind BER curve above 25 dB. This is due to the fact that the bound for $\mathbf{x}^H \mathbf{Y}_D^H \mathbf{B}^{-1} \mathbf{Y}_D \mathbf{x}$ given by $\mathbf{x}^H \mathbf{Y}_D^H \mathbf{A}^{-1} \mathbf{Y}_D \mathbf{x}$ becomes weak in low SNR, and the complexity becomes exponential in N.

D. Impact of Channel-Length Overestimation

The blind detector needs knowledge of the channel length L. If L is overestimated, the effect of channel-length overestimation is presented in Fig. 8. The simulation is performed over a 6-ary exponential PDP channel. The blind detector is evaluated at $L = 6$, 8, 10. The overestimation of L causes a performance loss in low SNR. However, in high SNR, the performance loss is negligible. At BER $= 4 \times 10^{-4}$, the detector with $L = 8$ has less than 0.1 dB loss over that with perfect L. When L increases to 10, the gap is still less than 0.5 dB. Therefore, our blind detector is insensitive to the overestimation of L.

E. Impact of Virtual Carriers

The dependence of the proposed detectors' performance on the number of virtual carriers (VCs) is highlighted in Figs. 9 and 10. The number of VCs is varied from 2 to 8 (correspondingly, the number of the data subcarriers P goes from 30 to 24). Fig. 9 shows that both the semi-blind and blind detectors are insensitive to the number of VCs. However, the MSE performance is degraded by increasing the number of VCs, since the number of information symbols for channel estimation is reduced. In Fig. 10, due to the use of VCs, the complexity of our semi-blind and blind detectors is greatly reduced. Figs. 9 and 10 are plotted assuming that no transmit power normalization is employed (i.e., a transmitted OFDM symbol has constant energy regardless of the number of VCs). For power normalization, the increase in the number of VCs will result in more power allocation to each of the data subcarriers, thus improving performance.

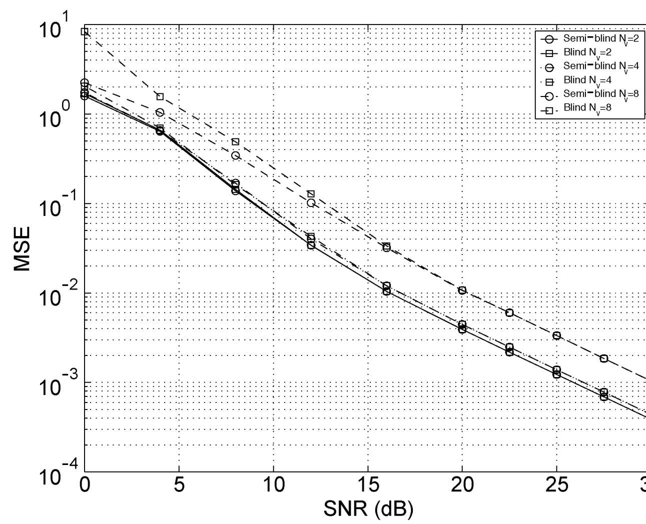

Fig. 10. MSE versus SNR for an BPSK OFDM system with $N = 32$ and different number of VCs.

VII. Conclusion

We have developed new blind and semi-blind data detectors and channel estimators for OFDM systems. Our data detectors are ML and require minimizing a complex, integer quadratic function. The semi-blind detector uses both channel correlation and noise level. The quadratic for the blind detector suffers from rank deficiency, to which we gave an efficient solution. We have also provided simple adaptations of the SD algorithm to handle M-PSK constellations and to achieve reduced complexity. We considered how the blind detector performs under mismatch, generalized the basic data detectors to nonunitary constellations, and extended them to systems with pilots and VCs. Simulation results show that the proposed detectors perform close to the ideal case. They may also be extended to MIMO-OFDM systems and OFDM over fast fading channels. These applications are currently being investigated.

REFERENCES

[1] G. Stuber, J. Barry, S. Mclaughlin, Y. Li, M. Ingram, and T. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.

[2] J. Rinne and M. Renfors, "Pilot spacing in orthogonal frequency division multiplexing systems on practical channels," *IEEE Trans. Consum. Electron.*, vol. 42, no. 4, pp. 959–962, Nov. 1996.

[3] O. Edfors, M. Sandell, J.-J. van de Beek, S. Wilson, and P. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[4] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[5] Y. Li, J. Cimini, and N. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, Jul. 1998.

[6] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[7] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1699–1712, Jul. 2002.

[8] J. Heath, R. W. , and G. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.

[9] S. Zhou and G. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1402–1414, Aug. 2001.

[10] 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.

[11] S. Roy and C. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 572–579, Oct. 2002.

[12] N. Chotikakamthorn and H. Suzuki, "On identifiability of OFDM blind channel estimation," in *Proc. Veh. Technol. Conf.*, Sep. 1999, vol. 4, pp. 2358–2361.

[13] M. Necker and G. Stuber, "Totally blind channel estimation for OFDM over fast varying mobile channels," in *Proc. IEEE Int. Conf. Commun.*, May 2002, pp. 421–425.

[14] C. Gao, M. Zhao, S. Zhou, and Y. Yao, "Blind channel estimation algorithm for MIMO-OFDM systems," *Electron. Lett.*, vol. 39, pp. 1420–1422, Sep. 2003.

[15] Y. Song, S. Roy, and L. Akers, "Joint blind estimation of channel and data symbols in OFDM," in *Proc. Veh. Technol. Conf.*, May 2000, vol. 1, pp. 46–50.

[16] P. Chen and H. Kobayashi, "Maximum likelihood channel estimation and signal detection for OFDM systems," in *Proc. IEEE Int. Conf. Commun.*, Apr. 2002, vol. 3, pp. 1640–1645.

[17] M. Luise, R. Reggiannini, and G. M. Vitetta, "Blind equalization/detection for OFDM signals over frequency-selective channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1568–1578, Oct. 1998.

[18] M.-X. Chang and Y. T. Su, "Blind and semiblind detections of OFDM signals in fading channels," *IEEE Trans. Commun.*, vol. 52, no. 5, pp. 744–754, May 2004.

[19] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 651–668, Mar. 2002.

[20] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Computat.*, vol. 44, pp. 463–471, Apr. 1985.

[21] E. Viterbo and J. Bouros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[22] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm—Part I: The expected complexity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pt. 1, pp. 2806–2810, Aug. 2005.

[23] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–15, Jan. 1999.

[24] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY: McGraw-Hill, 2001.

[25] P. Stoica, H. Vikalo, and B. Hassibi, "Joint maximum-likelihood channel estimation and signal detection for SIMO channels," in *Proc. IEEE ICASSP*, Apr. 2003, vol. 4, pp. 13–16.

[26] T. Cui and C. Tellambura, "Cyclic-prefix based noise variance and power delay profile estimation for OFDM systems," in *Proc. IEEE Pacific Rim Conf. Commun., Comput., Signal Process.*, Aug. 2005, pp. 518–521.

[27] M. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[28] A. Chan and I. Lee, "A new reduced-complexity sphere decoder for multiple antenna systems," in *Proc. IEEE Int. Conf. Commun.*, Apr. 2002, vol. 1, pp. 460–464.

[29] C. P. Schnorr and M. Euchner, "Lattice basis reduction: improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, pp. 181–191, 1994.

[30] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[31] A. Szabo, T. Kurpjuhn, and W. Utschick, "Linearly prefiltered OFDM based on long-term properties of the channel covariance matrix," in *Proc. Veh. Technol. Conf.*, 2002, pp. 2066–2070.

[32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[33] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-Speed Physical Layer in the 5 GHz Band*, IEEE Std 802.11a, Sept. 1999.

[34] G. L. Stuber, *Principles of Mobile Communication*, 2nd ed. Norwell, MA: Kluwer, 2001.

[35] O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161–163, May 2000.

Tao Cui (S'04) received the B.Eng. degree in information engineering from Xian Jiaotong University, Xian, China, in 2003, the M.Sc. degree from the University of Alberta, Edmonton, AB, Canada in 2005, and is currently working toward the Ph.D. degree at the Department of Electrical Engineering, California Institute of Technology, Pasadena.

His research interests are in communication theory, broadband wireless communications, space-time coding, MIMO systems, and wireless networks.

Mr. Cui was a recipient of postgraduate scholarships from the Alberta Ingenuity Fund and the Alberta Informatics Circle of Research Excellence (iCORE).

Chintha Tellambura (M'97–SM'02) received the B.Sc. degree (with first-class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, in 1986, the M.Sc. degree in electronics from the University of London, London, U.K., in 1988, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1993.

He was a Postdoctoral Research Fellow with the University of Victoria (1993–1994) and the University of Bradford (1995–1996). He was with Monash University, Melbourne, Australia, from 1997 to 2002. Presently, he is an Associate Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. His research interests include coding, communication theory, modulation, equalization, and wireless communications.

Prof. Tellambura is an Associate Editor for both the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was a Chair of the Communication Theory Symposium for Globecom 2005, St. Louis, MO.