Blind Receiver Design for OFDM Systems Over Doubly Selective Channels

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

***Abstract*—We develop blind data detectors for orthogonal frequency-division multiplexing (OFDM) systems over doubly selective channels by exploiting both frequency-domain and time-domain correlations of the received signal. We thus derive two blind data detectors: a time-domain data detector and a frequency-domain data detector. We also contribute a reduced complexity, suboptimal version of a time-domain data detector that performs robustly when the normalized Doppler rate is less than 3%. Our frequency-domain data detector and suboptimal time-domain data detector both result in *integer* least-squares (LS) problems. We propose the use of the V-BLAST detector and the sphere decoder. The time-domain data detector is not limited to the Doppler rates less than 3%, but cannot be posed as an integer LS problem. Our solution is to develop an iterative algorithm that starts from the suboptimal time-domain data detector output. We also propose channel estimation and prediction algorithms using a polynomial expansion model, and these estimators work with data detectors (decision-directed mode) to reduce the complexity. The estimators for the channel statistics and the noise variance are derived using the likelihood function for the data. Our blind data detectors are fairly robust against the parameter mismatch.**

***Index Terms*—Channel estimation, data detection, maximum likelihood (ML), orthogonal frequency-division multiplexing (OFDM).**

I. Introduction

INCREASING demand for high data rate and high performance has led to the development of fourth-generation (4G) broadband wireless systems. A potential transmission technique for 4G is orthogonal frequency-division multiplexing (OFDM). Although for time-invariant channels OFDM equalization is trivial with a one-tap equalizer, the OFDM performance is highly sensitive to time variations of the channel and carrier frequency offset, which destroy the subcarrier orthogonality and cause inter-carrier interference, resulting in an error floor. Frequency offset can be eliminated using either training-based or blind techniques[1] (see, e.g., [1] and references therein). Nevertheless, the elimination of inter-carrier interference due to time variations of the channel is essential for 4G applications. Conventional one-tap equalization without inter-carrier interference compensation results in an error floor. The effects of inter-carrier interference are analyzed in [2], where central-limit arguments are used to model inter-carrier interference as Gaussian and to quantify its impact on bit error rate (BER). A bound on signal-to-interference-and-noise ratio (SINR) (due to inter-carrier interference) is given in [3].

Several channel estimation algorithms have been proposed for time-varying channels. When the OFDM symbol duration is much smaller than the channel coherence time (i.e., relatively mild Doppler), the channel is approximately constant over an OFDM symbol and can be estimated by conventional techniques (see [4] and [5] (and references therein). If the channel varies linearly [6], it can be estimated by linear interpolation between two channel estimates acquired by training symbols. Such algorithms only compensate for the inter-carrier interference when the rate of time variations of the channel is relatively low. They cannot exploit the time diversity of a rapidly time-varying channel [7]. In [8], a minimum mean-square error (MMSE) channel estimator, and a successive interference cancellation detector with optimal ordering are proposed. Although this receiver achieves the time diversity and performs better with the increasing Doppler rate, it not only needs pilot symbols, which consumes bandwidth, but also has high complexity. However, note that [8] also develops a reduced complexity low-rank version. The motivation of this paper is therefore to derive blind data detectors without direct channel estimation for frequency- and time-selective (doubly selective) channels.

In this paper, we consider blind data detectors for OFDM systems over doubly selective channels by exploiting both frequency- and time-domain correlations of the received signal. We thus derive two blind data detectors: time-domain data detector and frequency-domain data detector. We also contribute a reduced complexity, suboptimal version of the time-domain data detector that performs robustly when the normalized Doppler rate is less than 3%. Our frequency-domain data detector and suboptimal time-domain data detector both result in *integer* least-squares (LS) problems. Although they can be optimally solved by brute-force exhaustive search, it has exponential complexity in the number of subcarriers and is hence computationally prohibitive. We thus use both the Vertical Bell Labs Space-Time (V-BLAST) detector [9] and the sphere decoder [10]–[12]. The time-domain data detector is not

Paper approved by Y. Li, the Editor for Wireless Communication Theory of the IEEE Communications Society. Manuscript received April 29, 2005; revised July 23, 2006. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, the Informatics Circle of Research Excellence, and the Alberta Ingenuity Fund. This paper was presented in part at the IEEE International Conference on Communications, May 2005, Seoul, Korea.

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 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.2007.894112

[1]However, to the best of our knowledge, previous blind techniques cannot be applied to doubly selective channels.

limited to the Doppler rates less than 3%, but cannot be posed as an integer LS problem. Our solution is to develop an iterative algorithm that starts from the suboptimal time-domain data detector output. We also propose channel estimation and prediction algorithms using a polynomial expansion model [14], which is a special case of the basis expansion model [15], and these estimators work with data detectors (decision-directed mode) to reduce the complexity. The estimators for the channel statistics and the noise variance, which are needed by our blind data detectors, are derived using the likelihood function for the data. Our blind data detectors are fairly robust against the residual mismatch of the proposed parameter estimators. The time-domain data detector is able to exploit the inherent time diversity of doubly selective channels.

This paper is organized as follows. Section II reviews the basic baseband OFDM system model and analyzes inter-carrier interference. Section III derives the new blind data detectors. Their efficient implementation using V-BLAST and the sphere decoder is considered in Section IV. In Section V, we introduce channel estimation and prediction algorithms using a polynomial expansion model. Section VI discusses parameter estimation and mismatch issues. Section VII presents simulation results, and Section VIII provides the conclusions.

Notation

Bold symbols denote matrices or vectors. The operators $\otimes$ and $\mathrm{tr}(\cdot)$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^\dagger$ denote the Kronecker product, trace, transpose, conjugate transpose, and Moore-Penrose pseudo-inverse, respectively. The set of K-dimensional vectors is $\mathcal{C}^K$. An M-ary phase shift keying (MPSK) signal constellation is given by $\mathcal{Q} = \{e^{j2\pi k/M},\ k = 0, 1, \ldots, M-1\}$, and the set of N-dimensional MPSK vectors is $\mathcal{Q}^N$. A circularly complex Gaussian variable with mean μ and variance σ_n^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 $[\mathbf{F}]_{k,l} = 1/\sqrt{N} e^{j(2\pi/N)kl}$, $k, l \in 0, 1, \ldots, N-1$. The diagonal matrix formed by a vector $\mathbf{A}$ is $\mathbf{A}_D$. $S^(N, M)[\mathbf{x}]$ is an $N \times M$ circulant matrix, whose rows are composed of cyclically shifted versions of $\mathbf{x}$.

II. System Model and Intercarrier Interference Analysis

A. OFDM and Intercarrier Interference Analysis

Input symbols are chosen from a finite constellation $\mathcal{Q}$ and modulated by inverse DFT (IDFT) on N subcarriers. The resulting time-domain samples are

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{j(2\pi kn/N)}, \quad n = 0, 1, \ldots, N-1 \tag{1}$$

where $X_k \in \mathcal{Q}$. In this paper, we only consider unitary (i.e., constant modulus) constellations, $|X_k|^2 = 1$. However, the receiver design in this paper may also be extended to nonunitary constellations. Note that $X_k, k = 0, 1, \ldots, N-1$, are typically called the modulation or input symbols, and the term "OFDM symbol" refers to the entire sequence $x_k|_{k=0}^{N-1}$. The input symbol duration is T_s and the OFDM symbol duration is $T = NT_s$. These discrete-time samples are appropriately filtered before transmission.

The guard interval includes a cyclic prefix of $\{x_{N_g-1}, \ldots, x_{N-1}\}$, where N_g is the number of samples in the guard interval (N_g is assumed to be larger than the delay spread of the channel). The composite response which includes transmit/receive pulse shaping and the physical channel response between the transmitter and receiver may be modeled as [5]

$$h(t, \tau) = \sum_{l=0}^{L-1} h(t, \tau_l)\delta(\tau - \tau_l) \tag{2}$$

where $h_l(t) \sim \mathcal{CN}(0, E\{h_l^2\})$ and τ_l is the delay of the lth tap. Typically, it is assumed that $\tau_l = lT_s$, and this results in a finite-impulse-response (FIR) filter with an effective length L. The received signal after sampling can be represented as

$$\begin{aligned} y_n &= \sum_{l=0}^{L-1} h(n, l)x_{n-l} + w_n \\ &= \frac{1}{\sqrt{N}} \sum_{l=0}^{L-1} h(n, l) \sum_{k=0}^{N-1} X_k e^{j(2\pi k(n-l)/N)} + w_n \end{aligned} \tag{3}$$

where $h(n, l)$ is the discrete-time baseband equivalent channel model and $w(t)$ is an additive white Gaussian noise (AWGN) process with mean zero and variance σ_n^2. At the receiver, the guard interval is removed and DFT demodulation is performed, resulting is

$$\begin{aligned} Y_k &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y_n e^{-j(2\pi kn/N)}, \quad k = 0, 1, 2, \ldots, N-1. \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} h(n, l) e^{-j(2\pi kn/N)} + W_k \\ &= X_k H_k + \alpha_k + W_k \end{aligned} \tag{4}$$

where $H_k = 1/N \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} h(n, l) e^{-j2\pi kl/N}$, $W_k = 1/\sqrt{N} \sum_{n=0}^{N-1} w_n e^{-j2\pi nk/N}$, $W_k \sim \mathcal{CN}(\mu, \sigma_n^2)$, and

$$\alpha_k = \frac{1}{N} \sum_{m=0, m\neq k}^{N-1} X_m \sum_{n=0}^{N-1} \left(\sum_{l=0}^{L-1} h(n, l) e^{-j2\pi ml/N} \right) e^{j2\pi n(m-k)/N}.$$

The α_k's in (4) represent intercarrier interference due to time variations of the channel. If it is time-invariant, we find $\alpha_k = 0$, which means no intercarrier interference exists. In a time-variant channel, $\alpha_k \neq 0$ and the conventional one-tap equalization results in

$$\hat{X}_k = X_k + \frac{\alpha_k + W_k}{H_k}. \tag{5}$$

The data detector makes a hard decision on $\hat{X}_k$. Since, in high SNR, the intercarrier interference power $E\{|\alpha_k|^2\}$ is larger than the noise power σ_n^2, an irreducible error floor will appear.

Equation (4) can be expressed in a vector form as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W} \tag{6}$$

where $\mathbf{Y} = [Y_0, \ldots, Y_{N-1}]^T$, $\mathbf{X} = [X_0, \ldots, X_{N-1}]^T$, $\mathbf{W} = [W_0, \ldots, W_{N-1}]^T$, and $[\mathbf{H}]_{k,m} = 1/N \sum_{n=0}^{N-1} (\sum_{l=0}^{L-1} h(n,l) e^{-j2\pi ml/N}) e^{j2\pi n(m-k)/N}$. The model (6) is similar to a multiple-input multiple-output (MIMO) system. If the channel $\mathbf{H}$ is perfectly known at the receiver, MMSE equalization [8] eliminates the error floor due to one-tap equalization.

B. Channel Correlations

A wide-sense stationary uncorrelated scattering (WSSUS) channel (2) is characterized by its delay power spectrum and scattering function. We assume that $h(t,\tau_l)$ $(l = 0, \ldots, L-1)$ have the same normalized correlation function (scattering function) $r_t(\Delta t)$. Hence

$$r_{h_l}(\Delta t) = E\{h(t+\Delta t, \tau_l) h_l^*(t, \tau_l)\} = \sigma_l^2 r_t(\Delta t) \tag{7}$$

where $r_t(\Delta t)$ is an even function (wide-sense stationary). Therefore, the autocorrelation function of the channel is

$$E\{h(t+\Delta t, \tau_{l_1}) h(t, \tau_{l_2})\} = \sigma_{l_1}^2 r_t(\Delta t) \delta(l_1 - l_2). \tag{8}$$

Specifically, for Jakes' model [13], $r_t(\Delta t) = J_0(2\pi f_D \Delta t)$, where $J_0(\cdot)$ denotes the zeroth-order Bessel function of the first kind, and f_D is the Doppler frequency. The normalized Doppler frequency is $f_d = N f_D T_s$ and the frequency domain correlation function is

$$r_f(\Delta f) = \sum_{l=0}^{L-1} \sigma_l^2 e^{-j2\pi \Delta f \tau_l}. \tag{9}$$

Note that $r_f(\Delta f)$ depends on the power delay profile (PDP), $\sigma_0^2, \ldots, \sigma_{L-1}^2$.

III. Blind Data Detectors

A. Time-Domain Detector

The time-domain data detector exploits the correlation of the received time-domain signal (pre-DFT correlation). Given x_n or X_k, the time-domain received signal y_n (3) is zero-mean complex Gaussian. The correlation between y_{n_1} and y_{n_2} in (3) is

$$\begin{aligned} E\{y_{n_1} y_{n_2}^*\} &= \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} E\{h(n_1, l_1) h^*(n_2, l_2)\} x_{n_1 - l_1} x_{n_2 - l_2}^* \\ &\quad + \sigma_n^2 \delta(n_1 - n_2) \\ &= r_t[(n_1 - n_2)T_s] \sum_{l=0}^{L-1} \sigma_l^2 x_{n_1 - l} x_{n_2 - l}^* + \sigma_n^2 \delta(n_1 - n_2) \end{aligned} \tag{10}$$

where the second equality comes from (8). Substituting (1) into (10), we obtain

$$\begin{aligned} &E\{y_{n_1} y_{n_2}^*\} \\ &= \frac{r_t[(n_1 - n_2)T_s]}{N} \\ &\quad \times \sum_{l=0}^{L-1} \sigma_l^2 \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} X_{k_1} X_{k_2}^* e^{j2\pi k_1 (n_1 - l)/N} e^{-j2\pi k_2 (n_2 - l)/N} \\ &\quad + \sigma_n^2 \delta(n_1 - n_2) \\ &= \frac{r_t[(n_1 - n_2)T_s]}{N} \\ &\quad \times \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} X_{k_1} X_{k_2}^* e^{j2\pi k_1 n_1/N} e^{-j2\pi k_2 n_2/N} \sum_{l=0}^{L-1} \sigma_l^2 e^{-j2\pi (k_1 - k_2) l/N} \\ &\quad + \sigma_n^2 \delta(n_1 - n_2) \\ &= \frac{r_t[(n_1 - n_2)T_s]}{N} \\ &\quad \times \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} X_{k_1} X_{k_2}^* e^{j2\pi k_1 n_1/N} e^{-j2\pi k_2 n_2/N} r_f\left(\frac{k_1 - k_2}{N T_s}\right) \\ &\quad + \sigma_n^2 \delta(n_1 - n_2) \\ &= r_t[(n_1 - n_2)T_s] \mathbf{f}_{n_1}^H \mathbf{X}_D \mathbf{R}_f \mathbf{X}_D^H \mathbf{f}_{n_2} + \sigma_n^2 \delta(n_1 - n_2) \end{aligned} \tag{11}$$

where $\mathbf{X}_D = \mathrm{diag}\{X_0, X_1, \ldots, X_{N-1}\}$, the frequency-domain correlation matrix $[\mathbf{R}_f]_{ij} = r_f[(i-j)/(NT_s)]$, for $i = 0, 1, \ldots, N-1$, $j = 0, 1, \ldots, N-1$, and $\mathbf{f}_i$ is the ith column of the DFT matrix $\mathbf{F}$. The third equality comes from (9). The autocorrelation matrix of $\mathbf{y} = [y_0, \ldots, y_{N-1}]^T$ can thus be written as

$$\begin{aligned} \mathbf{R}_{yy} &= \mathcal{F}^H \mathbf{R}_t \otimes (\mathbf{X}_D \mathbf{R}_f \mathbf{X}_D^H) \mathcal{F} + \sigma_n^2 \mathbf{I}_N \\ &= \mathcal{F}^H (\mathbf{I}_N \otimes \mathbf{X}_D)(\mathbf{R}_t \otimes \mathbf{R}_f)(\mathbf{I}_N \otimes \mathbf{X}_D^H) \mathcal{F} + \sigma_n^2 \mathbf{I}_N \end{aligned} \tag{12}$$

where $[\mathbf{R}_t]_{i,j} = r_t[(i-j)T_s]$ is the time-domain correlation matrix $\mathcal{F} = \mathrm{diag}\{\mathbf{f}_0, \ldots, \mathbf{f}_{N-1}\}$ is block diagonal, and $\otimes$ denotes the Kronecker product with $(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$.

The probability density function (pdf) of $\mathbf{y}$ conditional on the transmitted data $\mathbf{X}$ is therefore

$$p(\mathbf{y}|\mathbf{X}) = (\pi^N \det(\mathbf{R}_{yy}))^{-1} \exp(-\mathbf{y}^H \mathbf{R}_{yy}^{-1} \mathbf{y}). \tag{13}$$

Since all data vectors are equally likely, the maximum-likelihood (ML) estimate of $\mathbf{X}$ is given by minimizing the cost metric

$$g(\mathbf{X}) = \mathbf{y}^H \mathbf{R}_{yy}^{-1} \mathbf{y} + \ln[\det(\mathbf{R}_{yy})]. \tag{14}$$

Although this cost function is exact, the presence of the determinant term is problematic. Fortunately, that the determinant term can be ignored without a significant performance loss is evidenced by the following reasons. Since $\det(\mathbf{R}_{yy})$ is the sum of product of $X_0, \ldots, X_{N-1}$, it is independent of $\mathbf{X}$ when N is large using the law of large numbers. This result is also verified by simulation. Moreover, when the channel is time-invariant, $r_t(\Delta t) = 1$. Therefore, $\mathbf{R}_{yy}$ reduces to

$$\mathbf{R}_{yy} = \mathbf{F}^H \mathbf{X}_D \mathbf{R}_f \mathbf{X}_D^H \mathbf{F} + \sigma_n^2 \mathbf{I}_N. \tag{15}$$

$\det(\mathbf{R}_{yy})$ can thus be expressed as

$$\begin{aligned}\det(\mathbf{R}_{yy}) &= \det(\mathbf{X}_D)\det(\mathbf{R}_f+\sigma_n^2\mathbf{I}_N)\det(\mathbf{X}_D^H)\\ &= \det(\mathbf{R}_f+\sigma_n^2\mathbf{I}_N).\end{aligned} \tag{16}$$

When X_k's belong to a unitary constellation, $\det(\mathbf{R}_{yy})$ is independent of $\mathbf{X}_D$.

For all of these reasons and for reducing computational complexity, we ignore the determinant term in (14) and obtain the cost function

$$g_T(\mathbf{X}) = \mathbf{y}^H\mathbf{R}_{yy}^{-1}\mathbf{y} \tag{17}$$

where the subscript T denotes the time-domain data detector. Unfortunately, it is impossible to move $\mathbf{X}_D$ out of $\mathbf{R}_{yy}^{-1}$ and write $g_T(\mathbf{X})$ as a quadratic form of $\mathbf{X}$ due to the existence of $\mathbf{R}_t$ or time selectivity. However, starting from the suboptimal time-domain data detector in the following, we solve (17) by using a greedy detector in the next section.

To derive the suboptimal time-domain data detector, we use the Taylor-series expansion of $r_t(\Delta t) = r_0 + r_1\Delta t + r_2(\Delta t)^2 + \cdots$. Since the time correlation function $r_t(\Delta t)$ is an even function (wide-sense stationary), the coefficients of the odd terms are zero, or $r_{2k+1} = 0$ for $k = 0, 1, \ldots$. Specifically, for Jakes' model, the zeroth-order Bessel function of the first kind can be expanded as $J_0(2\pi x) \approx 1 - (\pi x)^2$. Hence

$$\begin{aligned}r_t\left[(n_1-n_2)T_s\right] &= J_0(2\pi f_D T_s(n_1-n_2))\\ &= J_0\left(\frac{2\pi f_D N T_s(n_1-n_2)}{N}\right)\\ &\approx 1-\left(\frac{\pi f_d(n_1-n_2)}{N}\right)^2\\ &= 1-r_2\left(\frac{n_1-n_2}{N}\right)^2+r_4\left(\frac{n_1-n_2}{N}\right)^4\\ &\quad + o\left(\left(\frac{n_1-n_2}{N}\right)^4\right)\end{aligned} \tag{18}$$

where $r_2 = (\pi f_d)^2$ and $r_4 = (1/4)(\pi f_d)^4$. When $\pi f_d < 0.1$ or $f_d < 0.03$, $r_2 < 0.01$, $r_4 < 0.0001$ and all terms except the first term can be neglected. Therefore, $r_t\left[(n_1-n_2)T_s\right] = 1$ and $\mathbf{R}_{yy}$ is given in (15). The cost function for the suboptimal time-domain data detector is given by

$$g_{ST}(\mathbf{X}) = \mathbf{y}^H\mathbf{R}_{yy}^{-1}\mathbf{y} = \mathbf{y}^H\mathbf{F}^H\mathbf{X}_D(\mathbf{R}_f+\sigma_n^2\mathbf{I}_N)^{-1}\mathbf{X}_D^H\mathbf{F}\mathbf{y} \tag{19}$$

where the subscript ST denotes the suboptimal time-domain detector. Since $\mathbf{Y} = \mathbf{F}\mathbf{y}$, (19) can be further simplified as

$$\begin{aligned}g_{ST}(\mathbf{X}) &= \mathbf{Y}^H\mathbf{X}_D(\mathbf{R}_f+\sigma_n^2\mathbf{I}_N)^{-1}\mathbf{X}_D^H\mathbf{Y}\\ &= \mathbf{X}^T\mathbf{Y}_D^H(\mathbf{R}_f+\sigma_n^2\mathbf{I}_N)^{-1}\mathbf{Y}_D\mathbf{X}^*\\ &= \mathbf{X}^H\mathbf{Y}_D^T(\mathbf{R}_f^T+\sigma_n^2\mathbf{I}_N)^{-1}\mathbf{Y}_D^*\mathbf{X}\end{aligned} \tag{20}$$

where the second equality comes from the commutative property of the multiplication between a vector and a diagonal matrix. The suboptimal time-domain detector output is given by

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}\in\mathcal{Q}^N}\mathbf{X}^H\mathbf{G}_1\mathbf{X} \tag{21}$$

where $\mathbf{G}_1 = \mathbf{Y}_D^T(\mathbf{R}_f^T+\sigma_n^2\mathbf{I}_N)^{-1}\mathbf{Y}_D^*$.

Remarks:

- Ignoring the second term in (18) results in an suboptimal time-domain detector that is the same as a data detector for time-invariant channels, suggesting that, when the normalized Doppler rate is less than 3%, time selectivity is not severe. Therefore, the suboptimal time-domain detector does not need to know the Doppler rate exactly and can handle any normalized Doppler rate up to 3%.
- Both the time-domain data detector and the suboptimal time-domain detector need the knowledge of the PDP and the noise variance σ_n^2, which the receiver may not know exactly. In Section VI, we investigate the resulting parameter mismatch problem and contribute new algorithms to estimate these parameters.

B. Frequency-Domain Detector

In the frequency-domain data detector, we make use of the correlation of the frequency-domain received signal Y_k in (4). Using the central limited theorem, the intercarrier interference term α_k can be modeled as a complex Gaussian variable. Its mean can be computed as

$$\begin{aligned}E\{\alpha_k\} =& \frac{1}{N}\sum_{m=0,m\neq k}^{N-1}E\{X_m\}\sum_{n=0}^{N-1}\\ &\times\left(\sum_{l=0}^{L-1}E\{h(n,l)\}e^{-j2\pi ml/N}\right)e^{j2\pi n(m-k)/N}\\ =&0\end{aligned} \tag{22}$$

and the autocorrelation is given by (23), shown at the bottom of the page, where the expectation is taken for both X_k[2] and

[2]A rigorous derivation should take expectation only for $h(n,l)$. However, this will make the detector complicated and difficult to solve. Therefore, the frequency-domain detector derived in this way is only a suboptimal. This is the reason for the error floor of the frequency-domain detector in Section VII.

$$\begin{aligned}E\{\alpha_{k_1}\alpha_{k_2}^*\} =& \frac{1}{N^2}\sum_{\substack{m_1=0\\ m_1\neq k_1}}^{N-1}\sum_{\substack{m_2=0\\ m_2\neq k_2}}^{N-1}E\{X_{m_1}X_{m_2}^*\}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}\left(\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1}E\{h(n_1,l_1)h^*(n_2,l_2)\}e^{-j2\pi m_1l_1/N}e^{j2\pi m_2l_2/N}\right)\\ &\times e^{j2\pi n_1(m_1-k_1)/N}e^{-j2\pi n_2(m_2-k_2)/N}\\ =&\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\left\{\delta(k_1-k_2)-\frac{1}{N^2}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}r_t[(n_1-n_2)T_s]\left[e^{-j2\pi n_1(k_1-k_2)/N}+\left(1-\delta(k_1-k_2)\right)e^{-j2\pi n_2(k_1-k_2)/N}\right]\right\}\end{aligned} \tag{23}$$

$h(n,l)$, and we have assumed that X_m, $m=0,\ldots,N-1$ are i.i.d. ($E\{X_{m_1}X_{m_2}^*\}$ for $m_1 \neq m_2$). Therefore, given X_k, Y_k is also complex Gaussian with mean

$$E\{Y_k\} = X_k E\{H_k\} + E\{\alpha_k\} + E\{W_k\} = 0. \quad (24)$$

The correlation between Y_{k_1} and Y_{k_2} can thus be written as

$$E\{Y_{k_1}Y_{k_2}^*\} = E\{X_{k_1}H_{k_1}H_{k_2}^*X_{k_2}^* + X_{k_1}H_{k_1}\alpha_{k_2}^* + \alpha_{k_1}H_{k_2}^*X_{k_2}^* + \alpha_{k_1}\alpha_{k_2}^* + W_{k_1}W_{k_2}^*\}. \quad (25)$$

Substituting (4), (8), and (9) into (25), we can show

$$E\{H_{k_1}H_{k_2}^*\} == \frac{1}{N^2} r_f\left(\frac{k_1-k_2}{NT_s}\right)\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} r_t\left[(n_1-n_2)T_s\right] = \sigma_{\mathrm{ICI}}^2 r_f\left(\frac{k_1-k_2}{NT_s}\right) \quad (26)$$

where $\sigma_{\mathrm{ICI}}^2 = 1/N^2\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} r_t\left[(n_1-n_2)T_s\right]$. Assuming that data symbols are statistically independent or equivalently $E\{X_{k_1}X_{k_2}\} = \delta(k_1-k_2)$, we obtain

$$E\{\alpha_{k_1}H_{k_2}^*X_{k_2}^*\} = \frac{1}{N}\sum_{m=0,m\neq k_1}^{N-1} E\{X_mX_{k_2}^*\} \times \sum_{n=0}^{N-1}\left(\sum_{l=0}^{L-1} E\{h(n,l)H_{k_2}^*\}e^{-j2\pi ml/N}\right) \times e^{j2\pi n(m-k_1)/N}. \quad (27)$$

Substituting H_k into (27), we get (28), shown at the bottom of the page.

Since the expectation $E\{X_{k_1}H_{k_1}\alpha_{k_2}^*\} = E\{\alpha_{k_2}H_{k_1}^*X_{k_1}^*\}^*$ can be obtained similarly, we have

$$\begin{aligned} &E\{X_{k_1}H_{k_1}\alpha_{k_2}^* + \alpha_{k_1}H_{k_2}^*X_{k_2}^* + \alpha_{k_1}\alpha_{k_2}^*\} \\ &= \delta(k_1-k_2)\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\left\{1 - \frac{1}{N^2}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} r_t[(n_1-n_2)T_s]\right\} \\ &= \delta(k_1-k_2)\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\left(1-\sigma_{\mathrm{ICI}}^2\right). \end{aligned} \quad (29)$$

The autocorrelation matrix of $\mathbf{Y}$ can then be written as

$$\mathbf{R}_{YY} = \sigma_{\mathrm{ICI}}^2\mathbf{X}_D\mathbf{R}_f\mathbf{X}_D^H + \left(\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\left(1-\sigma_{\mathrm{ICI}}^2\right) + \sigma_n^2\right)\mathbf{I}_N. \quad (30)$$

The likelihood function of $\mathbf{Y}$ conditional on the transmitted data $\mathbf{X}$ can be expressed as

$$p(\mathbf{Y}|\mathbf{X}) = (\pi^N \det(\mathbf{R}_{YY}))^{-1}\exp(-\mathbf{Y}^H\mathbf{R}_{YY}^{-1}\mathbf{Y}). \quad (31)$$

Since

$$\det(\mathbf{R}_{YY}) = \det\left(\sigma_{\mathrm{ICI}}^2\mathbf{R}_f + \left(\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\left(1-\sigma_{\mathrm{ICI}}^2\right)+\sigma_n^2\right)\mathbf{I}_N\right)$$

is independent of $\mathbf{X}$, maximizing the likelihood function (31) is equivalent to minimizing

$$\begin{aligned} g_F(\mathbf{X}) &= \mathbf{Y}^H\mathbf{R}_{YY}^{-1}\mathbf{Y} \\ &= \sigma_{\mathrm{ICI}}^{-2}\mathbf{X}^H\mathbf{Y}_D^T\left(\mathbf{R}_f^T + \sigma_{en}^2\mathbf{I}_N\right)^{-1}\mathbf{Y}_D^*\mathbf{X} \end{aligned} \quad (32)$$

where

$$\sigma_{en}^2 = \frac{\left(\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\left(1-\sigma_{\mathrm{ICI}}^2\right)+\sigma_n^2\right)}{\sigma_{\mathrm{ICI}}^2}. \quad (33)$$

Compared with $g_{ST}(\mathbf{X})$ in (19), $g_F(\mathbf{X})$ replaces σ_n^2 with σ_{en}^2. Hence, σ_{en}^2 can be considered as the equivalent noise variance, which incorporates the effect of intercarrier interference. The equivalent noise variance σ_{en}^2 is determined by the PDP and the scattering function. If the channel is time-invariant, $\sigma_{\mathrm{ICI}}^2 = 1$ and $\sigma_{en}^2 = \sigma_n^2$. Equation (32) reduces to (19). Ignoring the constant $\sigma_{\mathrm{ICI}}^{-2}$, the frequency-domain data detector is given by

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}\in\mathcal{Q}^N}\mathbf{X}^H\mathbf{G}_2\mathbf{X} \quad (34)$$

where $\mathbf{G}_2 = \mathbf{Y}_D^T\left(\mathbf{R}_f^T + \sigma_{en}^2\mathbf{I}_N\right)^{-1}\mathbf{Y}_D^*$.

Remarks:

- Both the suboptimal time-domain detector and the frequency-domain data detector result in a quadratic form in $\mathbf{X}$. Since all X_k's belong to a finite discrete constellation $\mathcal{Q}$, both (21) and (34) form so-called integer LS problems. An exhaustive search has the exponential complexity of $|\mathcal{Q}|^N$, making it infeasible for practical systems. The next section gives two algorithms to efficiently solve (21) and (34).
- If $\tilde{\mathbf{X}}$ is a solution of the time-domain data detector, we find that $\tilde{\mathbf{X}}e^{j\theta}$ ($e^{j\theta} \in \mathcal{Q}$) is also feasible for the time-domain data detector. Similar phase ambiguity exists for the suboptimal time-domain detector and the frequency-domain data detector as well. This can be resolved by a few pilots.

$$\begin{aligned} E\{\alpha_{k_1}H_{k_2}^*X_{k_2}^*\} &= \frac{1}{N^2}\sum_{m=0,m\neq k_1}^{N-1} E\{X_mX_{k_2}^*\}\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}\sum_{l_1=0}^{L-1}\sum_{l_2=0}^{L-1} E\{h(n_1,l_1)h^*(n_2,l_2)\}e^{j2\pi(k_2l_2-ml_1)/N}e^{j2\pi n_1(m-k_1)/N} \\ &= \frac{1-\delta(k_1-k_2)}{N^2}\left(\sum_{l=0}^{L-1}\sigma_l^2\right)\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} r_t[(n_1-n_2)T_s]e^{-j2\pi n_1(k_1-k_2)/N} \end{aligned} \quad (28)$$

IV. Efficient Detection Algorithms

For the suboptimal time-domain detector and the frequency-domain data detector, (21) and (34) can be expressed in the common general form

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}\in\mathcal{Q}^N} \|\mathbf{S} - \mathbf{B}\mathbf{X}\|^2 \tag{35}$$

where $\mathbf{B}$ is the corresponding Cholesky decomposition matrix of $\mathbf{G}_1$ or $\mathbf{G}_2$, and $\mathbf{S} = 0$ for the suboptimal time-domain detector and frequency-domain data detector. If the channel $\mathbf{H}$ is perfectly known at the receiver, the ML detector can also be written in (35), where $\mathbf{S} = \mathbf{Y}$ and $\mathbf{B} = \mathbf{H}$ are given in (6). The well-known V-BLAST and sphere decoder algorithms can solve (35) and achieve the time diversity. Here, we do not consider the MMSE-BLAST detector developed in [8], since its complexity is $\mathcal{O}(N^4)$, which is higher than those of V-BLAST and the sphere decoder.

A. V-BLAST Detection

The V-BLAST detector [9] changes the order of symbol detection so that the error probability is reduced. To do so, one needs a permutation matrix $\mathbf{\Pi}$ such that the QR decomposition of $\mathbf{B}' = \mathbf{B}\mathbf{\Pi} = \mathbf{Q}\mathbf{R}$ has the property that $\min_{0\le i\le N-1} r_{ii}$ is maximized over all column permutations, where the r_{ii}'s are the diagonal terms of $\mathbf{R}$. For $k = N-1, N-2, \ldots, 0$, the algorithm chooses $\pi(k)$ such that

$$\pi(k) = \arg\min_{j\notin\{\pi(1),\ldots,\pi(k-1)\}} \|(\mathbf{G}_k)_j\|^2 \tag{36}$$

where $(\mathbf{G}_k)_j$ is the jth row of $\mathbf{G}_k$, $\mathbf{G}_k$ is the pseudo inverse of $\mathbf{B}_k$, and $\mathbf{B}_k$ denotes the matrix obtained by zeroing columns $\pi(1), \ldots, \pi(k-1)$ of $\mathbf{B}$. Equation (35) can be expressed as

$$\hat{\mathbf{X}}_D = \arg\min_{\mathbf{X}\in\mathcal{Q}^N} \|\mathbf{S}' - \mathbf{R}\mathbf{X}'\|^2 \tag{37}$$

where $\mathbf{S}' = \mathbf{Q}^H\mathbf{S}$ and $\mathbf{X}' = \mathbf{X}\mathbf{\Pi}^T$. Since $\mathbf{R}$ is upper triangular, the kth element of (37) is

$$\hat{X}'_k = \arg\min_{X'_k,\ldots,X'_{N-1}} \left| s'_k - \left(r_{kk}X'^*_k + \sum_{i=k+1}^{N-1} r_{ki}X'^*_i \right) \right|^2 \tag{38}$$

where the estimate is free of interference from subcarriers $0, 1, \ldots, k-1$. Thus, X'_k can be estimated by minimizing (38). Proceeding with $X'_{N-1}, \ldots, X'_0$ and assuming correct previous decisions, the interference between subcarriers can be cancelled successively.

B. Sphere Decoder

The sphere decoder [10] reduces decoding complexity by taking advantage of the lattice structure of transmitted signals to achieve ML performance with a moderate complexity. The interested reader is referred to [10]–[12].

C. Greedy Iterative Algorithm

When the normalized Doppler exceeds 3%, the suboptimal time-domain detector cost function (21) is no longer accurate. Moreover, since the time-domain data detector cannot be written in quadratic form in $\mathbf{X}$, the V-BLAST detector and the sphere decoder are not useful and hence we propose a greedy algorithm for the time-domain data detector. Note that, from the simulation results in Section VII., the solution quality of the suboptimal time-domain detector is usually good though it has an error floor in high SNR. We thus use the suboptimal time-domain detector output as the initial estimate denoted by $\mathbf{X}^{(0)}$, where the superscript $^{(i)}$ denotes the ith iteration. Let the vector $\mathbf{X}$ be partitioned into several groups. Let the group size be denoted by S and the number of groups is M, $N = SM$. In the ith iteration, when the algorithm processes $\mathbf{g}_m = \{X_{(m-1)S+1}, \ldots, X_{mS}\}$, all of the other $M-1$ groups are fixed and $\mathbf{g}_m$ is chosen as

$$\hat{\mathbf{g}}_m = \arg\min_{\mathbf{g}_m\in\mathcal{Q}^S} f(\mathbf{g}_m|\mathbf{g}_1, \ldots, \mathbf{g}_{m-1}, \mathbf{g}_{m+1}, \ldots, \mathbf{g}_M) \tag{39}$$

where $f(\cdot)$ is the time-domain data detector cost function in (17). It then proceeds to the $m+1$th group and so on. The same process continues until convergence, $\mathbf{X}^{(i)} = \mathbf{X}^{(i+1)}$. It is evident that, when $M = 1$, this detector reduces to an exhaustive search. If $M = N$, it becomes an iterative decision feedback detector. Therefore, the group detector varies between ML and decision feedback when S varies. Simulation results shown later indicates that $S = 1, 2$, ensures good performance.

Instead of using (11), we use (10) to compute $\mathbf{R}_{yy}$. We first take the IDFT of $\mathbf{X}$ and then substitute the time-domain signals $\mathbf{x}$ into (10), which involves only L additions and multiplications, compared with the $N \times N$ matrix multiplication in (11). In the greedy detector, IDFT is not necessary every time. When the mth group is changed from $\breve{\mathbf{g}}_m$ to $\tilde{\mathbf{g}}_m$, $\breve{x}_n$ is updated to yield

$$\tilde{x}_n = \breve{x}_n + \frac{1}{\sqrt{N}} \sum_{k=(m-1)S+1}^{mS} (\tilde{X}_k - \breve{X}_k) e^{j2\pi kn/N}. \tag{40}$$

This update has complexity $\mathcal{O}(N)$, which is less than the complexity of IFFT $\mathcal{O}(N \log N)$.

V. Channel Estimation and Decision Directed Detection

A channel estimation algorithm for doubly selective channels has been given in [8]. However, its complexity is very high and it needs several pilot symbols, which is bandwidth-inefficient. Moreover, the time delay introduced by this algorithm may be problematic for a practical system. To alleviate these drawbacks, we propose channel estimation and prediction algorithms using a polynomial expansion model (PEM) [14], which is a special case of basis expansion model [15], and, more importantly, these estimation algorithms can be used in conjunction with our data detectors to reduce the overall computational complexity.

A. Polynomial Expansion Channel Model and Time Diversity Analysis

In [14], the doubly selective channel is modeled as a polynomial in time. For our OFDM problem, we assume that the PEM coefficients are constant within each OFDM symbol. We

denote the channel in the kth OFDM symbol as $h^{(k)}(t,\tau)$. Using Taylor's theorem, we expand $h^{(k)}(t,\tau)$ in t at NT_s as

$$h^{(k)}(t,\tau)=\sum_{l=0}^{L-1}\sum_{q=0}^{Q}h_q^{(k)}(\tau_l)\left(\frac{t-NT_s}{T_s}\right)^q\delta(\tau-\tau_l) \quad (41)$$

where Q is the order of the polynomial. We expand the function at NT_s because we will use the same coefficients to predict the channel in the $k+1$th symbol. The equivalent discrete channel model can then be written as

$$h^{(k)}(n,l)=\sum_{q=0}^{Q}h_q^{(k)}(l)\,(n-N)^q \quad (42)$$

for $n=0,2,\ldots,N-1, l=0,1,\ldots,L-1$. Equation (42) can be further written into matrix form as

$$\mathbf{h}^{(k)}(l)=\mathbf{\Phi}\mathbf{h}_p^{(k)}(l) \quad (43)$$

where $\mathbf{h}^{(k)}(l)=[h^{(k)}(0,l),\ldots,h^{(k)}(N-1,l)]^T$, $\mathbf{h}_p^{(k)}(l)=[h_0^{(k)}(l),\ldots,h_Q^{(k)}(l)]^T$ and $\mathbf{\Phi}(n,:)=[1,\ldots,(n-N)^Q]$. Here, we use the Matlab notation and subscript p denotes PEM. If $N\geq Q+1$, $\mathbf{\Phi}$ is invertible, and we can obtain

$$E\{\mathbf{h}_p^{(k)}(l)\}=E\{\mathbf{\Phi}^{\dagger}\mathbf{h}^{(k)}(l)\}=\mathbf{0} \quad (44)$$

and the autocorrelation matrix of $\mathbf{h}_p^{(k)}(l)$ is

$$\mathbf{R}_b(l)=E\{\mathbf{h}_p^{(k)}(l)(\mathbf{h}_p^{(k)}(l))^H\}=\sigma_l^2\mathbf{\Phi}^{\dagger}\mathbf{R}_t\left(\mathbf{\Phi}^{\dagger}\right)^H. \quad (45)$$

Note that, after using PEM, $h^{(k)}(n,l)$ becomes a nonstationary process.

Note that the PEM immediately reveals how time diversity occurs. We ignore the superscript (k) for simplicity. Substituting (42) into (3), we have

$$\begin{aligned}y_n&=\sum_{l=0}^{L-1}\sum_{q=0}^{Q}h_q(l)(n-N)^q x_{n-l}+w_n\\&=\frac{1}{\sqrt{N}}\sum_{q=0}^{Q}(n-N)^q\sum_{l=0}^{L-1}h_q(l)\sum_{k=0}^{N-1}X_k e^{j2\pi k(n-l)/N}+w_n.\end{aligned} \quad (46)$$

The vector form of (46) is

$$\mathbf{y}=\sum_{q=0}^{Q}\mathbf{\Omega}_q S^{(N,N)}[\mathbf{h}_q]\mathbf{x}+\mathbf{w} \quad (47)$$

where $\mathbf{\Omega}_q=\mathrm{diag}\{(-N)^q,\ldots,1^q\}$ and $\mathbf{h}_q=[h_q(0),\ldots,h_q(L-1)]^T$. After performing FFT on both sides of (47), we obtain

$$\mathbf{Y}=\sum_{q=0}^{Q}\mathbf{\Psi}_q\mathbf{H}_q\mathbf{X}+\mathbf{W} \quad (48)$$

where $\mathbf{\Psi}_q=\mathbf{F}\mathbf{\Omega}_q\mathbf{F}^H$ and $\mathbf{H}_q=\mathrm{diag}\{H_q(0),\ldots,H_q(N-1)\}$, $H_q(k)=\sum_{l=0}^{L-1}h_q(l)e^{-j2\pi kl/N}$. Clearly, the received vector $\mathbf{Y}$ in (48) has $Q+1$ replicas of $\mathbf{X}$, resulting in time diversity.

B. Channel Estimation and Prediction

Using the PEM, (47) can be rewritten as

$$\mathbf{y}=\sum_{q=0}^{Q}\mathbf{\Omega}_q S^{(N,L)}[\mathbf{x}]\mathbf{h}_q+\mathbf{w}=\mathbf{T}\mathbf{h}_p+\mathbf{w} \quad (49)$$

where $\mathbf{h}_p=[\mathbf{h}_0^T,\ldots,\mathbf{h}_Q^T]^T$ and $\mathbf{T}=[\mathbf{\Omega}_0 S^{(N,L)}[\mathbf{x}],\ldots,\mathbf{\Omega}_Q S^{(N,L)}[\mathbf{x}]]$. If data are detected using the detectors in Section III, the PEM coefficients can be estimated using the MMSE estimator as

$$\hat{\mathbf{h}}_p=\left(\sigma_n^2\mathbf{R}_p^{-1}+\mathbf{T}^H\mathbf{T}\right)^{-1}\mathbf{T}^H\mathbf{y} \quad (50)$$

where $\mathbf{R}_p=E\{\mathbf{h}_p\mathbf{h}_p^H\}$ is the autocorrelation matrix of $\mathbf{h}_p$. It can be computed from (45). The polynomial order determines the performance of the channel estimator and Q depends on the normalized Doppler f_d. We note an overestimate of Q also reduces the accuracy of PEM due to numerical instability. For normalized Doppler values as large as 10%, we find $Q=2$ is enough to achieve good performance.

Assuming that the PEM coefficients $\mathbf{h}_p$ are constant in two symbols, $\hat{\mathbf{h}}_p$ estimated in the kth symbol can be used to predict the channel in the $k+1$th symbol as

$$\hat{\mathbf{h}}^{(k+1)}(l)=\mathbf{\Phi}\hat{\mathbf{h}}_p^{(k)}(l) \quad (51)$$

where $\mathbf{\Phi}(n,:)=[1,\ldots,(n+N_g)^Q]$.

To improve the performance in the $k+1$th symbol, a decision-directed technique can be applied. In the $k+1$th symbol, the data symbols $\mathbf{X}^{(k+1)}$ is first detected using

$$\hat{\mathbf{X}}^{(k+1)}=\underset{\mathbf{X}^{(k+1)}\in\mathcal{Q}^N}{\arg\min}\left\|\mathbf{Y}^{(k+1)}-\hat{\mathbf{H}}^{(k+1)}\mathbf{X}^{(k+1)}\right\|^2 \quad (52)$$

where $\hat{\mathbf{H}}^{(k+1)}$ is computed from $\hat{\mathbf{h}}^{(k+1)}(l)$ as (50). The detection algorithms in Section IV can be used to solve (52). After obtaining $\hat{\mathbf{X}}^{(k+1)}$, it is substituted back into (50) to estimate $\hat{\mathbf{h}}_p^{(k+1)}$. While iterations can continue until convergence, we find just one iteration guarantees good performance. The same process can be applied to $k+2$-th symbol using $\hat{\mathbf{h}}_p^{(k+1)}$ and so on until there appears significant error propagation caused by the decision-directed operation. Our data detectors are then used to truncate the error propagation.

Fig. 1 shows the mean-square error (MSE) of the proposed MMSE channel estimator with different polynomial order Q in a channel with the PDP given in Section VII and a 10% normalized Doppler, where the MSE is defined as

$$\mathrm{MSE}=\frac{1}{NM}\sum_{k=1}^{M}\sum_{l=0}^{L-1}\|\mathbf{h}^k(l)-\hat{\mathbf{h}}^k(l)\|^2 \quad (53)$$

and M is the total number of Monte Carlo simulations. We assume perfect data in the kth symbol. The model with $Q=0$ exhibits an error floor in high SNR. For the kth symbol, the models with $Q=1$ and $Q=2$ almost achieve the same performance, although $Q=2$ performs slightly better than $Q=1$ in high SNR. If $\mathbf{h}_p^{(k)}$ is used to predict the channel in the $k+1$th symbol, $Q=2$ performs much better than $Q=1$, but both

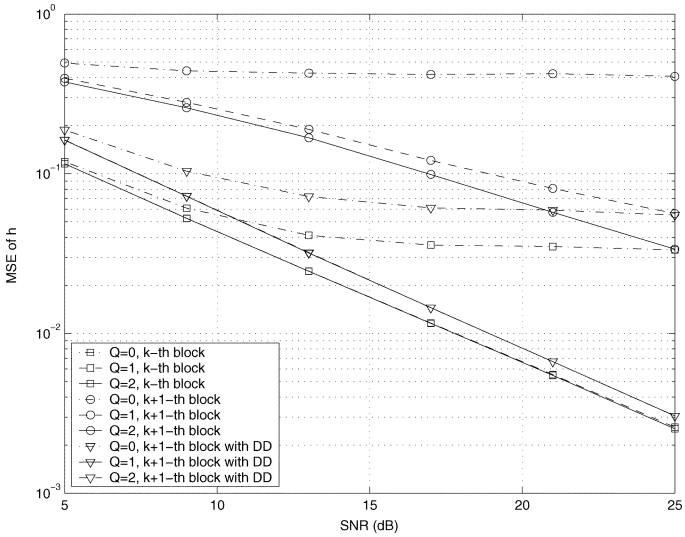

Fig. 1. MSE of channel estimator and prediction versus SNR using PEM with different polynomial order Q in a BPSK-OFDM system with $N = 32$ subcarriers and a normalized Doppler rate of 10% ($f_d = 0.1$).

of them have large performance loss compared with the perfect data case. However, if the decision-directed technique is used, both $Q = 1$ and $Q = 2$ improve much; for instance, at $\mathrm{MSE} = 3 \times 10^{-3}$, the performance loss is less than 1 dB.

VI. Parameters Estimation and Mismatch Analysis

Our detectors need to know the PDP, scattering function $r_t(\cdot)$, and the noise variance σ_n^2. These quantities remain constant over long periods and need not be estimated every OFDM symbol. The block diagram of the full receiver structure incorporating the time-domain data detector and the parameter estimators is given in Fig. 2.

A. Channel Covariance Matrix Estimation

After $\hat{\mathbf{h}}_p$ is estimated using (50), the channel in the kth OFDM symbol and the nth time slot $\mathbf{h}(kN+n) = [h(kN+n,0),\ldots,h(kN+n,L-1)]^T$ can be obtained via (43). The channel autocorrelation matrix can thus be estimated as

$$\hat{\mathbf{R}}_h = \frac{1}{KN}\sum_{k=1}^{K}\sum_{n=0}^{N-1}\mathbf{h}(kN+n)\mathbf{h}^H(kN+n) \tag{54}$$

where K is the number of OFDM symbols used and $\mathbf{R}_h = \mathrm{diag}\{\sigma_0^2,\ldots,\sigma_{L-1}^2\}$. Extensive simulation results indicate that the MSE of this estimate is almost independent of the SNR and the normalized MSE is about 1%, which may be due to the use of the PEM. The normalized MSE is defined as

$$\mathrm{NMSE} = \frac{\sum_{k=1}^{I}\sum_{l=0}^{L-1}|\hat{\mathbf{R}}_h(l,l)-\sigma_l^2|^2}{I\sum_{l=0}^{L-1}\sigma_l^2} \tag{55}$$

where I is the total number of Monte Carlo runs. In the simulation, we set $f_dT = 0.1$ and $K = 300$.

B. Noise Variance Estimation

Most of the noise variance estimation algorithms [16], [17] are valid only for flat and slow fading channels. To the best of our knowledge, no noise variance estimation algorithm has ever been proposed for OFDM systems over doubly selective channels. We note that if the data symbols are known, the pdf of $\mathbf{y}$ (13) becomes conditional on σ_n^2 as

$$p(\mathbf{y}|\sigma_n^2) = (\pi^N \det(\mathbf{R}_{yy}))^{-1}\exp(-\mathbf{y}^H\mathbf{R}_{yy}^{-1}\mathbf{y}). \tag{56}$$

To proceed, we first apply SVD decomposition to the matrix $\mathcal{F}^H\mathbf{R}_t \otimes (\mathbf{X}_D\mathbf{R}_f\mathbf{X}_D^H)\mathcal{F} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, where $\mathbf{U}$ is a unitary matrix and $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_0,\ldots,\lambda_{N-1}\}$. Ignoring irrelevant terms, the log-likelihood function of (56) can be simplified as

$$g(\sigma_n^2) = -\sum_{i=0}^{N-1}\log(\lambda_i+\sigma_n^2) - \frac{|y_i'|^2}{\lambda_i+\sigma_n^2} \tag{57}$$

where $\mathbf{y}' = \mathbf{U}^H\mathbf{y}$. The variance estimator is therefore

$$\hat{\sigma}_n^2 = \arg\min_{\sigma_n^2}\left[\sum_{i=0}^{N-1}\log(\lambda_i+\sigma_n^2) + \frac{|y_i'|^2}{\lambda_i+\sigma_n^2}\right]. \tag{58}$$

Let $\hat{\sigma}_n^2(k)$ be the estimate of σ_n^2 in the kth OFDM symbol. The final variance estimate is

$$\hat{\sigma}_n^2 = \frac{1}{K}\sum_{k=1}^{K}\hat{\sigma}_n^2(k) \tag{59}$$

where K is the number of OFDM symbols used.

Fig. 3 shows the normalized MSE (NMSE) $E\{[(\hat{\sigma}_n^2 - \sigma_n^2)/\sigma_n^2]^2\}$ of the proposed estimator with perfect data and Doppler frequency. The estimator performance is fairly insensitive to the operating SNR and gets better with the number of OFDM symbols.

C. Doppler Frequency Estimation

In [18], a Doppler frequency estimation algorithm is proposed, but it needs noise variance. We instead propose a new Doppler frequency estimator. Like the noise variance estimator, (13) can be written as the conditional pdf on f_d assuming perfect data and noise variance. However $p(\mathbf{y}|f_d)$ depends on f_d in a more complicated way than (58) on σ_n^2. We use Brent's method [19] to solve the unconstrained optimization problem in f_d. The initial point is set to the value in the last estimation. Finally f_d is estimated as

$$\hat{f}_d = \frac{1}{K}\sum_{k=1}^{K}\hat{f}_d(k) \tag{60}$$

where $\hat{f}_d(k)$ is Doppler estimate in the kth symbol and K is the number of OFDM symbols.

The NMSE of f_d ($E\{[(f_d-\hat{f}_d)/f_d]^2\}$) with different K is given in Fig. 4. We assume perfect data and noise variance. The estimator performance increases with the SNR for 8 dB, after which it remains more or less flat. As expected, the performance also increases with the number of OFDM symbols. We note that

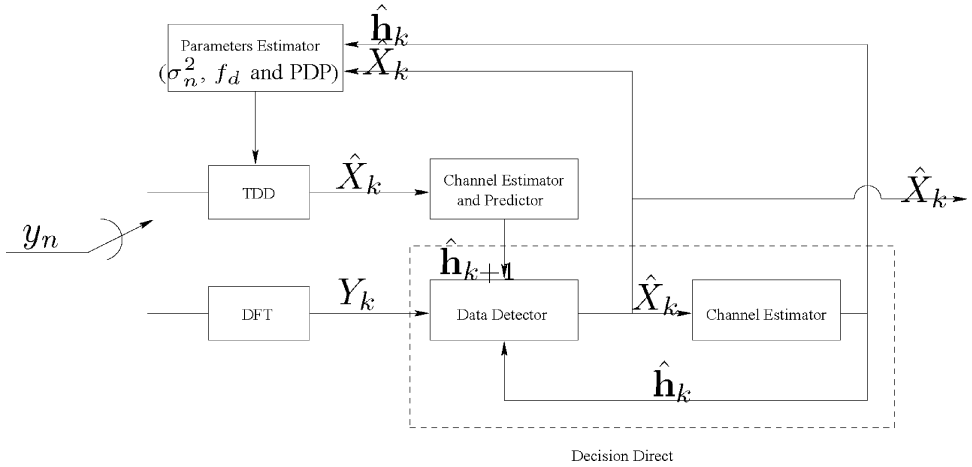

Fig. 2. Blind receiver structure for OFDM systems.

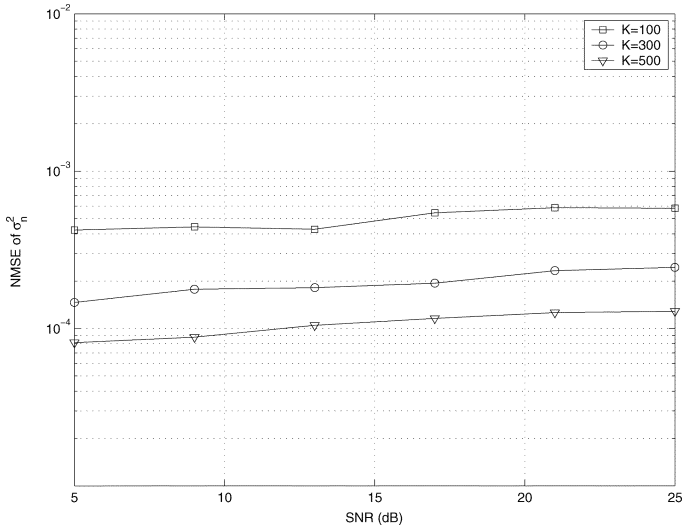

Fig. 3. NMSE of noise variance estimator with different number of symbols K in a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 3%.

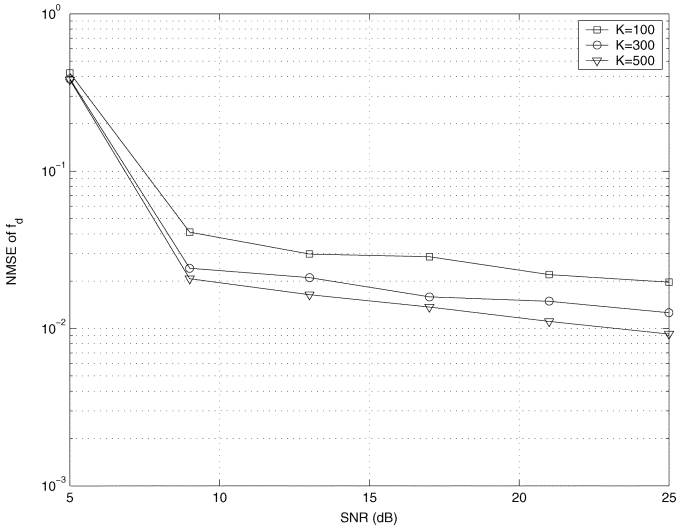

Fig. 4. NMSE of Doppler frequency estimator with different number of symbols K in a BPSK-OFDM system with 32 subcarriers and a 3% normalized Doppler.

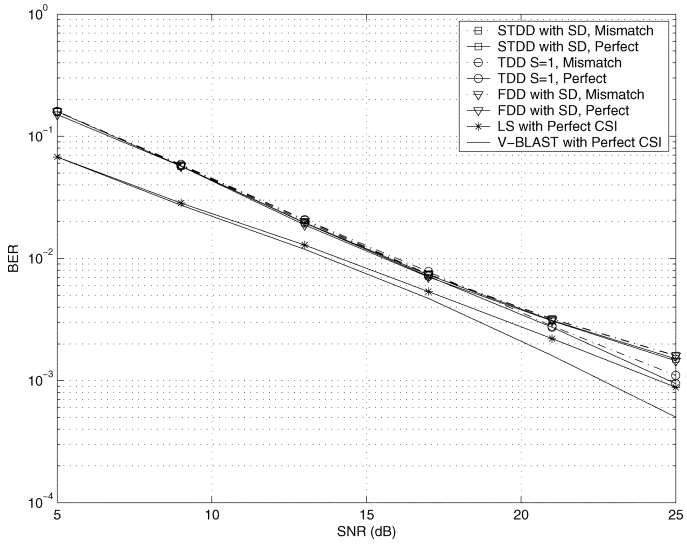

Fig. 5. BER of the mismatched blind data detectors for a BPSK-OFDM system with $N = 32$ and a 3% normalized Doppler. The data detectors are designed for an actual SNR of 25 dB, 1.5% normalized Doppler and a uniform PDP.

the noise variance and Doppler frequency estimators can work iteratively until convergence.

D. Mismatch Effects

We now investigate the performance loss when there is a mismatch between the true statistical values and the predetermined values. To do that, we use an idea shown in [4], [8], [20] that an estimator designed for the uniform power delay profile (UPDP)—the worst case—is robust to the channel statistics mismatch. Thus, we design our data detectors for the UPDP, an SNR of 25 dB and a Doppler rate of 1.5%, but the actual channel parameters are given at the beginning of Section VII. Fig. 5 presents the effect of parameter mismatch on our data detectors. The performance degradation at low SNR is negligible. At high SNR, the performance difference is within 0.5 dB. When the data detectors are designed for a Doppler rate less than the actual one, they achieve smaller time diversity than the detectors designed for the actual Doppler rate.

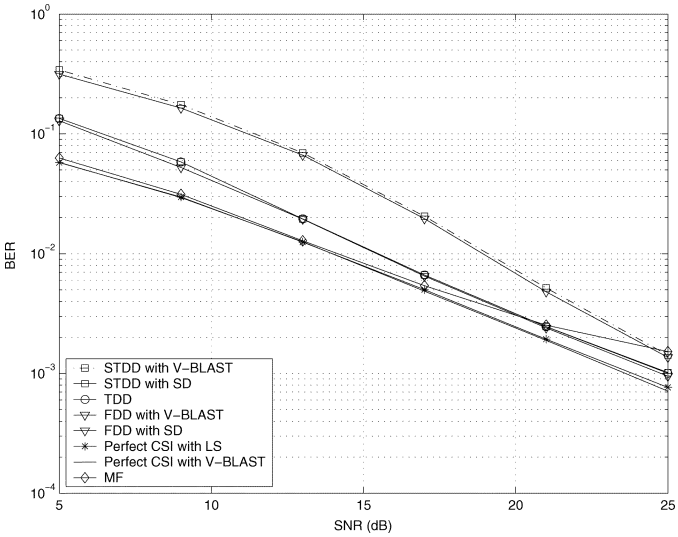

Fig. 6. BER performance versus SNR for a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 0.75%.

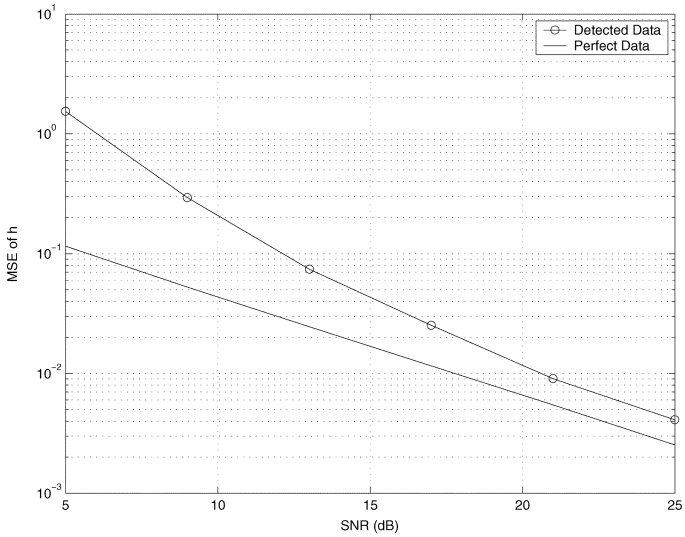

Fig. 7. MSE of channel estimation with perfect data and estimated data symbols for a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 0.75%.

VII. Simulation Results

We next give computer simulation results for the following system specifications.

- Both data and pilot symbols are chosen from binary phase-shift keying (BPSK).
- A 32-subcarrier OFDM system with the data rate 3 Mb/s is considered and $N_g = 8$.
- A six-ray channel with the PDP [0.189, 0.379, 0.239, 0.095, 0.061, 0.037] taken from COST 207 TU model [21] is considered. Each path is a complex Gaussian random process, independently generated with the classical Jakes' spectrum and the IDFT method [22]. The channel statistics and noise variance are known at the receiver.
- The number of subcarriers is thirty two ($N = 32$). Just one pilot is used, i.e., X_{N-1} is known/fixed *a priori* so that V-BLAST and sphere decoder return a unique solution.

We set $S = 1, 2$ for the time-domain data detector iterative detector and compare our data detectors with ideal LS, matched filter (MF) [8, Eq. (17)], and V-BLAST detectors, which are assumed to know the channel perfectly.

Fig. 6 compares the BER performance of different detectors for a 0.75% normalized Doppler. The time-domain data detector is detected using the greedy algorithm with the group size being one ($S = 1$) (the same performance observed for $S = 2$). The time-domain data detector, suboptimal time-domain detector, and frequency-domain data detector with sphere decoder perform almost the same and are 1 dB worse than the two reference detectors. Increasing SNR, V-BLAST performs close to sphere decoder. The performance gap is 2 dB at a BER of 2×10^{-3}. The suboptimal time-domain detector performs robustly when the normalized Doppler is less than 3%. Fig. 7 shows the MSE of channel estimation with perfect data and detected data using proposed detectors with the sphere decoder. A first-order PEM ($Q = 1$) is used.

Fig. 8 shows the BER of different detectors for a 3% normalized Doppler rate. The suboptimal time-domain detector and frequency-domain data detector with both V-BLAST and sphere decoder have error floors at high SNR since the former detector is derived for normalized Doppler rates below 3% and the latter detector is derived suboptimally. The time-domain data detectors with group sizes 1 and 2 perform identically and approach the LS reference detector. The decision-directed technique reduces the error floor of both suboptimal time-domain detector and frequency-domain data detector. V-BLAST with perfect CSI performs better than the LS detector, which shows that V-BLAST exploits the time diversity. A second-order PEM ($Q = 2$) is used for channel estimation. All the data detectors have roughly the same MSE. In Fig. 9, detected data channel estimate has a 3.8-dB loss over its reference counterpart at an MSE of 10^{-2}.

When the normalized Doppler frequency increases to 6%, suboptimal time-domain detector and frequency-domain data detector with both V-BLAST and sphere decoder have large

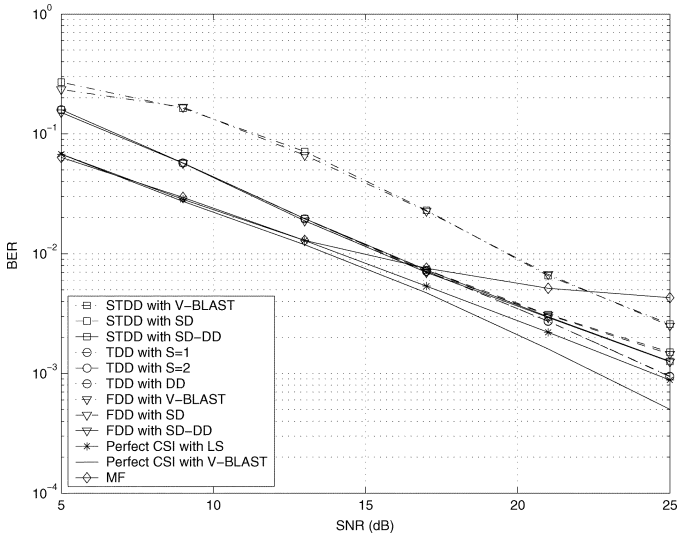

Fig. 8. BER performance versus SNR for a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 3%.

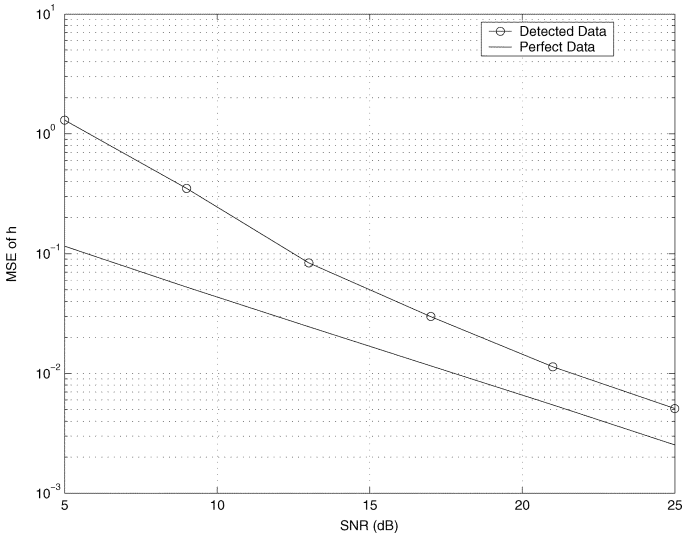

Fig. 9. MSE of channel estimation with perfect data and estimated data symbols for a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 3%.

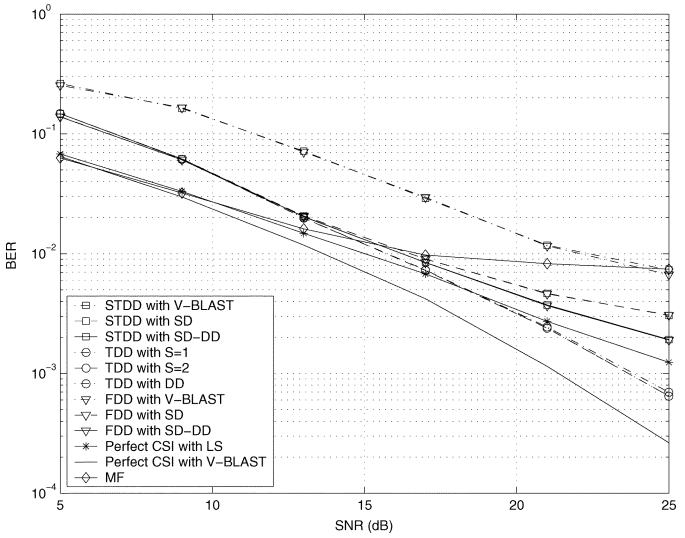

Fig. 10. BER performance versus SNR for a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 6%.

error floors at high SNR (Fig. 10). The error floor of V-BLAST is larger than that of sphere decoder. However, the time-domain data detector performs better than the LS reference detector and the performance loss is 2 dB over the V-BLAST reference detector, which shows that the time-domain data detector exploits the time-diversity induced by fast fading [8]. The decision-directed technique improves the performance of all the detectors and greatly reduces the error floors. Note however that all the benchmark detectors (LS, MF, and V-BLAST) know the channel matrix $\mathbf{H}$ [see (6)] completely (i.e., N^2 channel coefficients). Despite their total channel knowledge, some of the benchmark detectors perform worse than our proposed detectors in high SNR. Fig. 11 compares the MSE of channel estimation of different data detectors. The polynomial order is set to $Q = 2$.

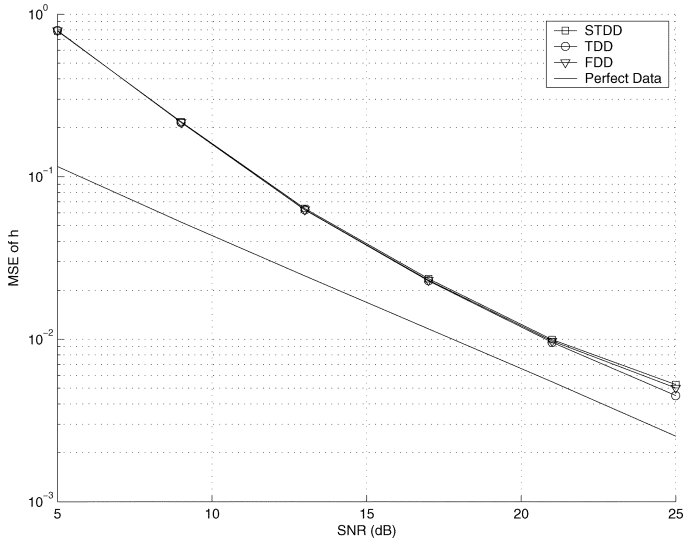

Fig. 11. MSE of channel estimation with detected data symbols from different blind data detectors for a BPSK-OFDM system with 32 subcarriers and a normalized Doppler of 6%.

VIII. Conclusion

We derived time-domain and frequency-domain data detectors for OFDM systems in doubly selective channels, which exploit the correlation among time-domain and frequency-domain received signal samples. Our detectors avoid channel estimation, which is a difficult task in fast fading channels, and use the low-complexity V-BLAST detector and the near-optimum sphere decoder. The polynomial expansion model has been derived to estimate and predict the doubly selective channel. A decision-direct technique has been proposed to reduce the data detector complexity. Our data detectors are only used to stop the decision-directed error propagation. We also derived parameter estimators and analyzed the effect of residual mismatch. The suboptimal time-domain detector and frequency-domain data detector perform robustly for normalized Doppler rates up to 3% and the time-domain data detector can exploit the time diversity.

References

[1] U. Tureli, D. Kivanc, and H. Liu, "Experimental and analytical studies on a high-resolution OFDM carrier frequency offset estimator," *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 629–643, Mar. 2001.

[2] M. Russell and G. Stuber, "Interchannel interference analysis of OFDM in a mobile environment," in *Proc. IEEE Veh. Technol. Conf.*, Jul. 1995, vol. 2, pp. 820–824.

[3] Y. Li and L. J. Cimini, Jr, "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 401–404, Mar. 2001.

[4] Y. Li, L. J. Cimini, Jr, 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.

[5] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 7, pp. 1207–1215, Jul. 2000.

[6] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 27–32, Jan. 1999.

[7] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 123–132, Jan. 1999.

[8] Y.-S. Choi, P. Voltz, and F. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.
[9] 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.
[10] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, inlcuding a complexity analysis," *Math. Computation*, vol. 44, pp. 463–471, Apr. 1985.
[11] E. Viterbo and J. Bouros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 1639–1642, Jul. 1999.
[12] 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.
[13] J. W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.
[14] D. K. Borah and B. T. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 862–873, Jun. 1999.
[15] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1351–1366, May 2003.
[16] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWG channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.
[17] N. C. Beaulieu, A. S. Toms, and D. R. Pauluzzi, "Comparison of four SNR estimators for QPSK modulations," *IEEE Commun. Lett.*, vol. 4, no. 2, pp. 43–45, Feb. 2000.
[18] J. Cai, W. Song, and Z. Li, "Doppler spread estimation for mobile OFDM systems in Rayleigh fading channels," *IEEE Trans. Consumer Electron.*, vol. 49, no. 5, pp. 973–977, Nov. 2003.
[19] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2002.
[20] 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.
[21] G. L. Stuber, *Principles of Mobile Communication*, 2nd ed. Norwell, MA: Kluwer, 2001.
[22] D. Young and N. Beaulieu, "The generation of correlated Rayleigh random variates by inverse discrete Fourier transform," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1114–1127, Jul. 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 a 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. He 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.