

A Refined Channel Estimation Method for STBC/OFDM Systems in High-Mobility Wireless Channels

Meng-Lin Ku and Chia-Chi Huang

Abstract—In this paper, we investigate channel estimation for orthogonal frequency division multiplexing (OFDM) systems with space-time block code (STBC) in mobile wireless channels. Our proposed method consists of two-stage processing and is developed on the basis of the classical discrete Fourier transform (DFT)-based channel estimation method. In the initialization stage, we employ a multipath interference cancellation technique to estimate multipath delays and multipath complex gains. In the tracking stage, we develop a refined decision-feedback (DF) DFT-based channel estimation method in which a few pilot tones inserted in OFDM data symbols are applied to form an optimal gradient vector at the first iteration such that the error propagation effect is mitigated. In order to reduce computational complexity, an approximate weighting matrix is adopted to avoid matrix inversion. We demonstrate the proposed method through computer simulation of an STBC/OFDM system with two transmit antennas and a single receive antenna. The results show that our method outperforms the classical DFT-based method, the STBC-based minimum mean square error (MMSE) method, and the Kalman filtering method as well, and that significant signal-to-noise ratio (SNR) performance improvement can be achieved, especially when a high-level modulation scheme, e.g. 16QAM, is adopted in high-mobility environments.

Index Terms—Multiple-input multiple-output systems, space-time coding, orthogonal frequency division multiplexing, channel estimation, IEEE 802.16e.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) has been widely applied in wireless communication systems in recent years due to its capability of high-rate transmission and low-complexity implementation over frequency-selective fading channels. Space-time coding is another promising technique to provide diversity gain through the use of multiple transmit antennas, especially when receive diversity is too expensive to deploy. In particular, the space-time block code (STBC) has received a lot of attention because a simple linear decoder can be used at the receiver side [1][2]. These advantages make OFDM combined with STBC, known as STBC/OFDM,

an ideal choice for several applications such as wireless metropolitan area networks (WMANs) 802.16e [3], etc. However, a high-rate STBC/OFDM system employing multi-level modulation with non-constant envelope (e.g. 16QAM) generally requires accurate channel state information (CSI) to perform coherent detection. This in turn implies that dynamic channel estimation is a crucial factor in realizing a successful STBC/OFDM system over doubly selective channels.

Blind channel estimation, which merely relies on the received signals, is very attractive due to its bandwidth-saving advantage. Nevertheless, it requires a long data record, involves high computational complexity, and only applies to slowly time-varying channels. On the contrary, pilot-aided (PA) channel estimation, using pilot tones known to the receiver, shows great promise for applications in mobile wireless communication, even though the use of pilot tones ends up with lower data rate. Decision-feedback (DF) channel estimation offers an alternative way to track channel variations; nevertheless, it is vulnerable to decision error propagation in fast time-varying channels [4]–[9]. As a high quality channel estimator with low training overhead is needed for successful implementation of STBC/OFDM systems, we restrict our attention to the category of PA plus DF channel estimation methods in this paper.

Among a wide variety of PA plus DF channel estimation methods, the discrete Fourier transform (DFT)-based channel estimation method, derived from either minimum mean square error (MMSE) criterion or maximum likelihood (ML) criterion, has been intensively studied for OFDM systems with preambles [8]–[11]. It is shown in [7]–[10] that the DFT-based channel estimation method using the ML criterion is simpler to implement because it requires neither channel statistics nor operating signal-to-noise ratio (SNR). Furthermore, as presented in [10], the performance of the ML estimator is comparable to that of the MMSE estimator at intermediate or high SNR values when the number of pilot tones is sufficiently larger than the maximal channel length (in samples). Thus, we will focus on the ML estimator in this paper. In order to save bandwidth and improve system performance, DF data symbols are also used as pilots to track channel variations in subsequent OFDM data symbols, and this method is called the DF DFT-based channel estimation method. Recently, the mathematical equivalence between the DF DFT-based method and Newton's method has been studied in [12], and it is concluded that the DF DFT-based method is only applicable to slowly time-varying channels because of the local optimization capability of Newton's method.

Manuscript received June 1, 2007; revised January 1, 2008; accepted April 4, 2008. The associate editor coordinating the review of this paper and approving it for publication is X.-G. Xia.

This work is supported in part by the National Science Council under Grants NSC94-2220-E-009-033, NSC95-2220-E-009-010, and NSC96-2220-E-009-002.

M. L. Ku is with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, 300, Taiwan (e-mail: willyku.cm87@nctu.edu.tw).

C. C. Huang is with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, 300, Taiwan (e-mail: huangcc@faculty.nctu.edu.tw).

Digital Object Identifier 10.1109/T-WC.2008.070585

Most mobile wireless channels are characterized by channel impulse response (CIR) consisting of a few dominant paths. The multipath delays are usually slowly time-varying. The amplitude and the phase of each path, however, can vary relatively fast. In this paper, we propose a two-stage channel estimation method by utilizing these channel characteristics. In the initialization stage, we employ a multipath interference cancellation (MPIC)-based decorrelation method to identify significant paths. In the following tracking stage, we develop a refined DF DFT-based channel estimation method, in which we use a few pilot tones inserted in OFDM data symbols to form an optimal gradient vector at the first iteration. This optimal gradient vector helps the classical method jump out of the local optimum, thus reducing the error propagation effect. The classical DF DFT-based channel estimation method is then used at the following iterations. In addition, an approximate weighting matrix is adopted to reduce the computational complexity associated with the matrix inversion operation of the weighting matrix in the DF DFT-based channel estimation method.

The rest of this paper is organized as follows. In Section II, we describe the STBC/OFDM system. In Section III, we present the MPIC-based decorrelation method in the initialization stage. Next, the equivalence between the DF DFT-based channel estimation method and Newton's method is briefly reviewed, and we propose a refined DF DFT-based channel estimation method in the tracking stage. We then discuss the computational complexity of the proposed two-stage channel estimation method in Section IV. In Section V, we present our computer simulation and performance evaluation results. Finally, some conclusions are drawn in Section VI.

Notation: By convention, boldface letters are used for matrices, vectors, and sets. The superscripts $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ stand for complex conjugate, transpose and Hermitian, respectively. The notation $\Re(\cdot)$ takes the real part of (\cdot) . The notation $\mathbf{I}_{N \times N}$ presents an $N \times N$ identity matrix. We denote the Kronecker delta function as $\delta[n]$. Besides, we denote the dimension of the vector \mathbf{x} as $|\mathbf{x}|$. The notation $\{\cdot\}$ denotes a set, e.g. a set $\mathbf{x} = \{x_1, \dots, x_{|\mathbf{x}|}\}$, where $|\mathbf{x}|$ is cardinality of the set \mathbf{x} . Further, $\text{diag}\{\mathbf{x}\}$ denotes the diagonal matrix with vector \mathbf{x} on the diagonal, and $\text{diag}\{\mathbf{X}_1, \dots, \mathbf{X}_M\}$ denotes the block diagonal matrix with the submatrices $\mathbf{X}_1, \dots, \mathbf{X}_M$ on the diagonal; finally, $j = \sqrt{-1}$

II. STBC/OFDM SYSTEM

A. Transmitted Signals

As shown in Fig. 1, we consider an STBC/OFDM system with N_T transmit and N_R receive antennas, employing K subcarriers among which M subcarriers are used to transmit data symbols plus pilot tones and the other $K - M$ subcarriers are used as either a DC subcarrier or virtual subcarriers at the two edges to avoid the spectrum overlapping problem at the receiver. Assume that \mathbf{Q} and \mathbf{J} denote the sets of subcarrier indices assigned to transmit data symbols and pilot tones, respectively. At subcarrier $k \in \mathbf{Q}$ (or $k \in \mathbf{J}$) and after symbol mapping, P modulated data symbols (or pilot tones) $\{s_1[k], \dots, s_P[k]\}$ are encoded by a $T \times N_T$ STBC encoder $\mathbf{X}[k]$ to generate N_T signal sequences of length T ,

each of which is denoted by $\{X^{(i)}[1, k], \dots, X^{(i)}[T, k]\}$, for $i = 1, \dots, N_T$ [1][2]. These N_T signal sequences possess the orthogonal property given by $\mathbf{X}^H[k]\mathbf{X}[k] = C[k]\mathbf{I}_{N_T}$, where $C[k] = C^{(i)}[k] = \sum_{t=1}^T |X^{(i)}[t, k]|^2$ is the total transmitted signal energy at the k th subcarrier for each transmit antenna, and they are simultaneously transmitted by the N_T transmit antennas within the duration of T OFDM data symbols. At the t th OFDM data symbol and the i th transmit antenna, after insertion of $K - M$ zeros for DC and virtual subcarriers, STBC encoded data symbols and pilot tones $X^{(i)}[t, k]$, for $k \in \mathbf{Q}$ and $k \in \mathbf{J}$, respectively, are modulated onto M subcarriers via a K -point inverse DFT (IDFT) unit to produce time domain samples. A cyclic prefix (CP) is then added in front of each OFDM data symbol to eliminate intersymbol interference caused by multipath channels. As shown in Fig. 1, each OFDM frame starts with a CP-added preamble which occupies one OFDM symbol and is followed by N_D consecutive OFDM data symbols. In order to avoid inter-antenna interference at the receiver side, pilot tones are alternatively inserted into the available subcarriers of the preamble with antenna-specific subcarrier shifts [3], and the preamble in frequency domain is denoted by $P^{(i)}[k]$, for $k \in \mathbf{Q} \cup \mathbf{J}$ and $i = 1, \dots, N_T$.

B. Channel Model

The complex baseband representation of impulse response for a mobile wireless channel (in discrete expression) between the i th transmit antenna and the j th receive antenna can be described by

$$h^{(j,i)}[t, \tau] = \sum_{l=1}^{L^{(j,i)}} \mu_l^{(j,i)}(t) \delta[\tau - \tau_l^{(j,i)}] \quad (1)$$

where $L^{(j,i)}$ is the number of resolvable paths, $\tau_l^{(j,i)}$ is the time delay in samples of the l th path, and $\mu_l^{(j,i)}(t)$ is the complex Gaussian fading gain of the l th path. All paths $\mu_l^{(j,i)}(t)$, for $l = 1, \dots, L^{(j,i)}$, are assumed to be independent of each other. Thus, with proper cyclic extension, the channel frequency response can be expressed as

$$H^{(j,i)}[t, k] = \sum_{l=1}^{L^{(j,i)}} \mu_l^{(j,i)}(t) \exp\left\{-j2\pi k \tau_l^{(j,i)} / K\right\} \quad (2)$$

where k is the subcarrier index.

C. Received Signals

We assume that both timing and carrier frequency synchronization are perfect, and that the length of channel impulse response is always smaller than the length of the CP. Another assumption here is that the channel is quasi-static over the duration of each time slot. The time index t in $H^{(j,i)}[t, k]$ is dropped hereafter. Hence, after the OFDM demodulator in Fig. 1, the received preamble at the j th receive antenna can be expressed as

$$\check{R}^{(j)}[k] = \sum_{i=1}^{N_T} H^{(j,i)}[k] P^{(i)}[k] + Z^{(j)}[k] \quad (3)$$

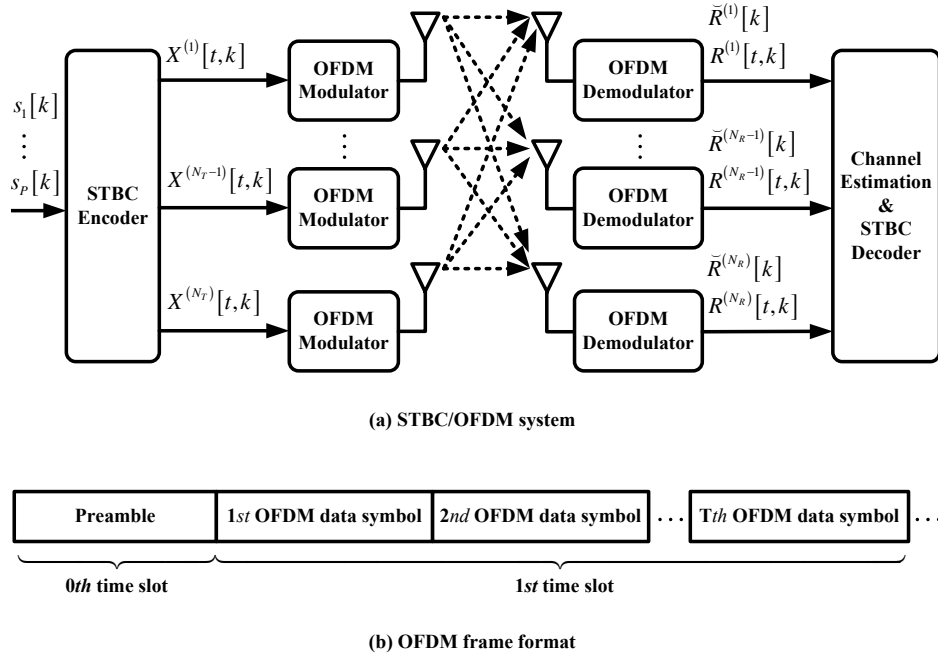


Fig. 1. (a) STBC/OFDM system (b) OFDM frame format.

for $k \in \mathbf{Q} \cup \mathbf{J}$. Moreover, the T successively received OFDM data symbols at the j th receive antenna are given by

$$R^{(j)}[t, k] = \sum_{i=1}^{N_T} H^{(j,i)}[k] X^{(i)}[t, k] + Z^{(j)}[t, k] \quad (4)$$

for $t = 1, \dots, T$ and $k \in \mathbf{Q} \cup \mathbf{J}$. Note that $Z^{(j)}[k]$ in (3) or $Z^{(j)}[t, k]$ in (4) are uncorrelated additive white Gaussian noise (AWGN) with zero-mean and variance σ_Z^2 on the j th receive antenna.

III. PROPOSED CHANNEL ESTIMATION METHOD

In this section, we first present the MPIC-based decorrelation method for the initialization stage. Before introducing the refined DF DFT-based channel estimation method for the tracking stage, we briefly review the equivalence between the DF DFT-based channel estimation method and Newton's method.

A. Initialization Stage: The MPIC-Based Decorrelation Method

We all know that CIR can be estimated using the preamble placed at the beginning of each OFDM frame, while the difficulty is that for most wireless standards, the preamble does not have ideal auto-correlation due to the use of either guard band or non-equally spaced pilot tones. Fig. 2 outlines the MPIC-based decorrelation method to estimate CIR path-by-path by canceling out already known multipath interference. Since the preambles transmitted from different antennas do not interfere with each other at the receiver side, channel estimation can be independently performed for each transceiver antenna pair, and therefore the antenna indices j and i are omitted in the following. In step 1, we first define two parameters N_h and \mathbf{W}_b which represent a preassumed number of paths in

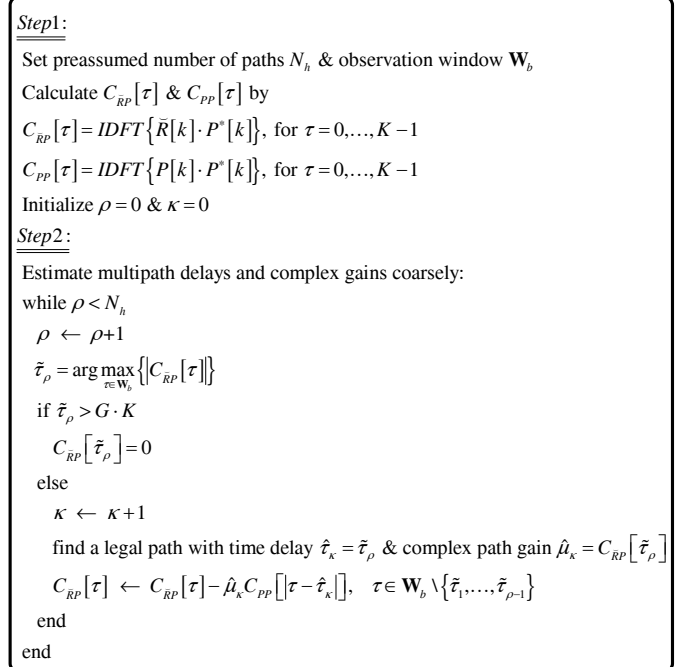


Fig. 2. The MPIC-based decorrelation method in the initialization stage. (G is the ratio of the CP length to the useful OFDM symbol time, and $\text{IDFT}\{\cdot\}$ is a K -point IDFT operation.)

a mobile radio channel and a multipath observation window, respectively. Next, we calculate the cyclic cross-correlation $C_{RP}[\tau]$ between the received and the transmitted preamble as well as the normalized cyclic auto-correlation $C_{PP}[\tau]$ of the transmitted preamble. Both ρ and κ , which stand for a path counting variable and the number of legal paths found by the MPIC-based decorrelation method, respectively, are initialized

to zero. In step 2, we start by increasing the value of the path counting variable, ρ , by one, and picking only one path whose time delay $\tilde{\tau}_\rho$ yields the largest value in $|C_{\tilde{R}P}[\tau]|$, for $\tau \in \mathbf{W}_b$. If the time delay $\tilde{\tau}_\rho$ is larger than the length of the CP, this path is treated as an illegal path, thereby discarded by setting $C_{\tilde{R}P}[\tilde{\tau}_\rho] = 0$. Otherwise, we increase the number of legal paths found, κ , by one, and then reserve this path as the κ th legal path with time delay $\hat{\tau}_\kappa = \tilde{\tau}_\rho$ and complex path gain $\hat{\mu}_\kappa = C_{\tilde{R}P}[\tilde{\tau}_\rho]$. The replica of the interference associated with this legal path is regenerated and subtracted from $C_{\tilde{R}P}[\tau]$ to obtain a refined cross-correlation function:

$$C_{\tilde{R}P}[\tau] \leftarrow C_{\tilde{R}P}[\tau] - \hat{\mu}_\kappa C_{PP}[|\tau - \hat{\tau}_\kappa|] \quad (5)$$

for $\tau \in \mathbf{W}_b \setminus \{\tilde{\tau}_1, \dots, \tilde{\tau}_{\rho-1}\}$, where " \leftarrow " is the assignment operation. We continue the iterative process of the step 2 until ρ reaches the preassumed value of N_h .

B. Equivalence between DF DFT-Based Method and Newton's Method

Through the initialization stage, we are able to obtain information on the number of paths $\kappa^{(j,i)} (\leq N_h)$, the multipath delays $\hat{\tau}_l^{(j,i)}$, the multipath complex gains $\hat{\mu}_l^{(j,i)}$, for $l = 1, \dots, \kappa^{(j,i)}$, and therefore the corresponding channel frequency response. With the assumption of the unchanged multipath delays over the duration of each OFDM frame, the DF DFT-based channel estimation method using the ML criterion (at the v th iteration) can be summarized as follows [8]–[10][12]. First, the re-encoded STBC matrix is used to obtain the least-square (LS) channel estimate [13]:

$$\delta_v^{(j)} = \hat{\mathbf{C}}_v^{-1} \hat{\mathbf{X}}_v^H \mathbf{R}^{(j)} \quad (6)$$

where v is the iteration number from 1 to V , $\mathbf{R}^{(j)} = [\mathbf{R}^{(j)T}[\Theta_1], \dots, \mathbf{R}^{(j)T}[\Theta_{|\Theta|}]]^T$ is the received signal vector at the j th receive antenna in which we have $\mathbf{R}^{(j)}[k] = [R^{(j)}[1, k], \dots, R^{(j)}[T, k]]^T$ and Θ is a subset of \mathbf{Q} used to track channel variations, $\hat{\mathbf{X}}_v = \text{diag}\{\hat{\mathbf{X}}_v[\Theta_1], \dots, \hat{\mathbf{X}}_v[\Theta_{|\Theta|}]\}$ consists of the re-encoded STBC matrix $\hat{\mathbf{X}}_v[k]$ with decision symbols $\hat{X}_v^{(i)}[t, k]$ which are obtained by applying the previously estimated channel frequency response to decode the received signal $\mathbf{R}^{(j)}[k]$ according to [1][2], $\hat{\mathbf{C}}_v = \text{diag}\{\hat{C}_v[\Theta_1] \mathbf{I}_{N_T}, \dots, \hat{C}_v[\Theta_{|\Theta|}] \mathbf{I}_{N_T}\}$ is the energy normalization factor in which $\hat{C}_v[k] = \hat{C}_v^{(i)}[k] = \sum_{t=1}^T |\hat{X}_v^{(i)}[t, k]|^2$. Thus, it is clear that the $((m-1)N_T + i)$ th entry of $\delta_v^{(j)}$ represents the LS estimate for the Θ_m th subcarrier at the i th transmit and the j th receive antenna. Second, a truncated DFT matrix is applied to improve the LS estimate as follows:

$$\mathbf{q}_v^{(j)} = \mathbf{F}^{(j)H} \Delta_v^{(j)} \quad (7)$$

$$\mathbf{E}^{(j)} = \mathbf{F}^{(j)H} \mathbf{F}^{(j)} \quad (8)$$

$$\mathbf{M}_v^{(j)} = \mathbf{F}^{(j)} \mathbf{E}^{(j)-1} \mathbf{q}_v^{(j)} \quad (9)$$

where $\Delta_v^{(j)} = \mathbf{\Pi} \delta_v^{(j)}$ and $\mathbf{\Pi}$ is an $N_T |\Theta| \times N_T |\Theta|$ permutation matrix obtained by making the $((i-1)|\Theta| + m)$ th row have all zeros except for a single one at the $((m-1)N_T + i)$ th column for $i = 1, \dots, N_T$ and $m = 1, \dots, |\Theta|$, $\mathbf{F}^{(j)} = \text{diag}\{\mathbf{F}^{(j,1)}, \dots, \mathbf{F}^{(j,N_T)}\}$, $\mathbf{F}^{(j,i)}$ is the $|\Theta| \times \kappa^{(j,i)}$

truncated DFT matrix whose (m, l) th entry is defined as $\exp\{-j2\pi\Theta_m \hat{\tau}_l^{(j,i)}/K\}$, and the $((i-1)|\Theta| + m)$ th entry of $\mathbf{M}_v^{(j)}$ represents the estimated channel frequency response for the Θ_m th subcarrier at the v th iteration, corresponding to the (j, i) th antenna pair. In order to initialize the channel estimator of (9), the CSI estimated in the previous time slot has to be taken as the initial value of the CSI for the current time slot.

It has been shown in [12] that the DF DFT-based channel estimation method can be equivalently expressed in the framework of Newton's method as [14]:

$$\bar{\delta}_v^{(j)} = \bar{\mathbf{M}}_{v-1}^{(j)} - \hat{\mathbf{C}}_v^{-1} \hat{\mathbf{X}}_v^H \mathbf{R}^{(j)} \quad (10)$$

$$\bar{\mathbf{q}}_v^{(j)} = \mathbf{F}^{(j)H} \bar{\Delta}_v^{(j)} \quad (11)$$

$$\mathbf{M}_v^{(j)} = \mathbf{M}_{v-1}^{(j)} - \mathbf{F}^{(j)} \mathbf{E}^{(j)-1} \bar{\mathbf{q}}_v^{(j)} \quad (12)$$

where $\bar{\mathbf{M}}_{v-1}^{(j)} = \mathbf{\Pi}^{-1} \mathbf{M}_{v-1}^{(j)}$ and $\bar{\Delta}_v^{(j)} = \mathbf{\Pi} \delta_v^{(j)}$. According to [12], the purpose of calculating the difference between $\bar{\mathbf{M}}_{v-1}^{(j)}$ and $\hat{\mathbf{C}}_v^{-1} \hat{\mathbf{X}}_v^H \mathbf{R}^{(j)}$ in (10), followed by the IDFT matrix $\mathbf{F}^{(j)H}$ in (11), is to form the gradient vector $\bar{\mathbf{q}}_v^{(j)}$ in Newton's method, as observed in (12). Furthermore, it is also proved in [12] that the role of the weighting matrix $\mathbf{E}^{(j)-1}$ in (12) is in fact the inverse of the Hessian matrix in Newton's method.

C. Refined DF DFT-Based Channel Estimation Method

Through the equivalence relation described in Section III.B., it is concluded that the classical DF DFT-based channel estimation method (called method I, hereafter) is limited by the local search capability of Newton's method and only applicable to slowly time-varying channels. In the previous studies [8][9], pilot tones as well as decision data symbols are simultaneously adopted to perform channel estimation at each iteration. This is, however, not a good solution in fast time-varying channels because decision data symbols easily induces the error propagation effect, whereas pilot tones are much more reliable than decision data symbols. From the viewpoint of optimization, the pilot tones inserted in each OFDM data symbol can play a more important role in providing a global search direction at the first iteration of the method I. Fig. 3 shows the refined DF DFT-based channel estimation method in the tracking stage. With the help from a few pilot tones to form a gradient vector according to (10) and (11), the refined channel tracking method (called method II, hereafter) is proposed here by only modifying the first iteration ($v = 1$) of (12) as follows

$$\mathbf{M}_1^{(j)} = \mathbf{M}_0^{(j)} - \mathbf{B}^{(j)} \mathbf{F}^{(j)} \bar{\mathbf{q}}_p^{(j)} \quad (13)$$

where $\bar{\mathbf{q}}_p^{(j)}$ is the gradient vector calculated according to (10) and (11) by only utilizing the set \mathbf{J} , instead of Θ , $\mathbf{B}^{(j)}$ is a block diagonal matrix defined as $\text{diag}\{\beta^{(j,1)} \mathbf{I}_{|\Theta|}, \dots, \beta^{(j,N_T)} \mathbf{I}_{|\Theta|}\}$, and $\beta^{(j,i)}$ is a real-valued

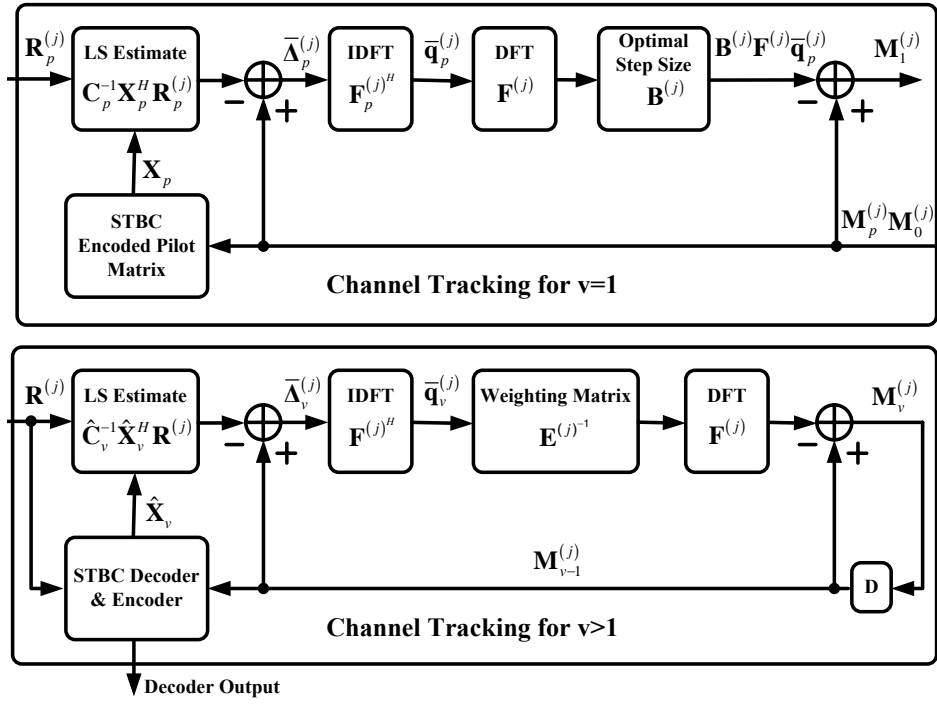


Fig. 3. Block diagram of the refined DF DFT-based channel estimation method in the tracking stage. (The subscript "p" is to indicate that the calculation is only associated with the pilot subcarrier set.)

step size which can be determined by minimizing the following ML cost function over the pilot subcarrier indices:

$$\begin{aligned} \beta_{opt} &= \arg \min_{\beta} \Omega(\beta) \\ &= \arg \min_{\beta} \sum_{j=1}^{N_R} \left| \mathbf{R}_p^{(j)} - \mathbf{X}_p \mathbf{\Pi}_p^{-1} \right. \\ &\quad \left. \times \left(\mathbf{M}_p^{(j)} - \mathbf{B}_p^{(j)} \mathbf{F}_p^{(j)} \bar{\mathbf{q}}_p^{(j)} \right) \right|^2 \quad (14) \end{aligned}$$

where $\beta = [\beta^{(1)T}, \dots, \beta^{(N_R)T}]^T$ in which we define $\beta^{(j)} = [\beta^{(j,1)}, \dots, \beta^{(j,N_T)}]^T$, the notations $\mathbf{R}_p^{(j)}$, \mathbf{X}_p , $\mathbf{M}_p^{(j)}$, $\mathbf{B}_p^{(j)}$, $\mathbf{F}_p^{(j)}$ and $\mathbf{\Pi}_p$ are defined similar to $\mathbf{R}^{(j)}$, $\hat{\mathbf{X}}_v$, $\mathbf{M}_v^{(j)}$, $\mathbf{B}^{(j)}$, $\mathbf{F}^{(j)}$ and $\mathbf{\Pi}$, respectively, but here the set \mathbf{J} is used instead of Θ in the definition. Note that $\mathbf{M}_p^{(j)}$ is the estimated channel frequency obtained from the previous time slot. Let $M^{(j,i)}[k]$ and $\xi^{(j,i)}[k]$ for $k = J_m$ denote the $((i-1)|\mathbf{J}| + m)$ th entries of $\mathbf{M}_p^{(j)}$ and $\mathbf{F}_p^{(j)} \bar{\mathbf{q}}_p^{(j)}$, respectively. Furthermore, we define $\gamma^{(j)}[t, k] = R^{(j)}[t, k] - \sum_{i=1}^{N_T} M^{(j,i)}[k] X^{(i)}[t, k]$ and $\phi^{(j,i)}[t, k] = \xi^{(j,i)}[k] X^{(i)}[t, k]$. By expanding (14) and taking $\partial \Omega(\beta) / \partial \beta = \mathbf{0}$, after straightforward manipulations, the optimum value of $\beta^{(j)}$ is given by

$$\beta_{opt}^{(j)} = -\Phi^{(j)^{-1}} \Gamma^{(j)} \quad (15)$$

where the (m, u) th entry of the matrix $\Phi^{(j)}$ is calculated by $\sum_{t=1}^T \sum_{k \in \mathbf{J}} \Re\{\phi^{(j,m)}[t, k] \phi^{(j,u)*}[t, k]\}$, the u th entry of the vector $\Gamma^{(j)}$ is given by $\sum_{t=1}^T \sum_{k \in \mathbf{J}} \Re\{\phi^{(j,u)}[t, k] \gamma^{(j)*}[t, k]\}$. It is noted that, after the first iteration, we execute the channel tracking process of (10)–(12) for the second and subsequent iterations until a stopping criterion holds. The stopping criterion is to check whether the absolute value of each entry

in $\mathbf{F}^{(j)} \mathbf{E}^{(j)^{-1}} \bar{\mathbf{q}}_v^{(j)}$ is less than a prespecified threshold ε or the iteration number v reaches the maximum value of V . The channel tracking process for the current time slot will be stopped when either of the above two conditions holds.

In order to reduce computational complexity in the method II, we further propose method III to avoid the matrix inverse of the weighting matrix, $\mathbf{E}^{(j)^{-1}}$, by taking into account the strongly diagonal property of $\mathbf{E}^{(j)}$ which is originally proposed for reducing the complexity of multiuser detection in code division multiple access (CDMA) systems [15]. Define $\mathbf{E}^{(j)} = |\Theta|(\mathbf{I}_{\kappa^{(j)}} + \mathbf{O}_{off})$, where $\kappa^{(j)} = \sum_{i=1}^{N_T} \kappa^{(j,i)}$ and \mathbf{O}_{off} is a zero-diagonal matrix. Then, it follows that if $|\Theta|$ is large enough, an approximate weighting matrix of $\mathbf{E}^{(j)^{-1}}$ takes the form:

$$\begin{aligned} \mathbf{E}^{(j)^{-1}} &= \frac{1}{|\Theta|} (\mathbf{I}_{\kappa^{(j)}} + \mathbf{O}_{off})^{-1} \\ &\approx \frac{1}{|\Theta|} (\mathbf{I}_{\kappa^{(j)}} - \mathbf{O}_{off}) \\ &= \frac{1}{|\Theta|^2} \left(2|\Theta| \mathbf{I}_{\kappa^{(j)}} - \mathbf{E}^{(j)} \right) \quad (16) \end{aligned}$$

IV. COMPUTATIONAL COMPLEXITY

Now let us look at the computational complexity of the three methods, in terms of the number of real multiplications per transceiver antenna pair. In general, the operations of K -point IDFT and $K \times K$ matrix inversion need $4K \log_2 K$ and $4K^3$ real multiplications, respectively. Besides, the weighting matrix $\mathbf{E}^{(j)^{-1}}$ only needs to be calculated once in each OFDM frame as it is only related to the multipath delays

TABLE I
COMPUTATIONAL COMPLEXITY FOR THE SYSTEM PARAMETERS
GIVEN IN SECTION V.A.

Initialization Stage			
Method I and Method II	14128		
Method III	12208		
Tracking Stage ($ \mathbf{J} = 8$)			
$ \Theta $	48	96	192
The first iteration of the Method II and the Method III	16806	16902	17094
Each iteration of the Method I or each of the second and subsequent iterations of the Method II and the Method III	17312	17888	19040

$\hat{\tau}_l^{(j,i)}$. Therefore, the complexity considered in the initialization stage is mainly due to the operations of $C_{\hat{R}P}[\tau]$ and $\mathbf{E}^{(j)-1}$. The calculation and update of $C_{\hat{R}P}[\tau]$ require at most $4(|\mathbf{Q}| + |\mathbf{J}|)/N_T + 4K \log_2 K + 4N_h |\mathbf{W}_b|$ real multiplications. Moreover, the calculation of $\mathbf{E}^{(j)-1}$ needs at most $4N_h^3$ real multiplications, but it needs at most $2N_h^2$ real multiplications if the approximate weighting matrix of $\mathbf{E}^{(j)-1}$ in (16) is used instead. In the tracking stage, the computation for each iteration of the method I (or each of the second and subsequent iterations of the method II and the method III) involves the calculation of $\mathbf{F}^{(j)} \mathbf{E}^{(j)-1} \bar{\mathbf{q}}_v^{(j)}$, in total requiring at most $|\Theta|(4T + 2) + 2|\Theta|T/N_T + 8K \log_2 K + 4N_h^2$ real multiplications. For the method II and the method III, an optimal gradient vector is formed at the first iteration, in which the computation of $\mathbf{B}^{(j)} \mathbf{F}^{(j)} \bar{\mathbf{q}}_p^{(j)}$ and the optimum $\beta^{(j)}$ at most requires $|\mathbf{J}|(4T + 2) + 8K \log_2 K + 2|\Theta| + 10|\mathbf{J}|T + N_T^2 + N_T(2|\mathbf{J}|T + 1)$ real multiplications. The computational complexity for the system parameters given in Section V.A. is listed in Table I. The values of the parameters $|\mathbf{W}_b|$, N_h , T , N_T , K , $|\mathbf{Q}|$, and $|\mathbf{J}|$ are set as 109, 8, 2, 2, 256, 192, and 8, respectively, while the value of the parameter $|\Theta|$ can be 48, 96, or 192. For a fair performance comparison, both the data subcarrier set Θ and the pilot subcarrier set \mathbf{J} are used for tracking channel variations, except that only the pilot subcarrier set is adopted at the first iteration of the method II and the method III. Hence, we use $|\Theta| + |\mathbf{J}|$ to replace $|\Theta|$ in the calculation of the complexity in Table I. As observed in Table I, the complexity of the method II and the method III is a little bit lower than that of the method I. It can also be noticed that the complexity in the tracking stage is mostly due to the operations of the DFT and the IDFT which in total require $8K \log_2 K = 16384$ real multiplications. Some complexity gain can be achieved by using partial DFT processing such as in [9]. If $|\Theta|$ is larger than N_h , the computation of the partial DFT processing mainly depends on the size of Θ . As a result, the complexity of the tracking stage is basically dominated by $|\Theta|$ since $|\mathbf{J}|$, N_T and T are usually much smaller than $|\Theta|$.

V. COMPUTER SIMULATION

A. System Parameters

We demonstrate the performance of the proposed channel estimation methods through computer simulation of an

STBC/OFDM system with two transmit antennas and a single receive antenna. The parameters are set the same as those in the 802.16e OFDM standard [3]. The system occupies a bandwidth of $5MHz$ and operates in the $2.3GHz$ frequency band. The entire bandwidth is divided into $K = 256$ subcarriers among which $|\mathbf{Q}| + |\mathbf{J}| = 200$ subcarriers are used to transmit data symbols and pilot tones, and the remaining $M = 56$ subcarriers are used as virtual subcarriers at the two edges and a DC subcarrier. In the simulation, the modulation schemes for the data symbols are QPSK and 16QAM, while the BPSK modulation scheme is adopted for the pilot tones. Each pilot subcarrier transmits the same power as each data subcarrier. Each OFDM frame is composed of one OFDM preamble and $N_D = 40$ OFDM data symbols. The length of the CP is 64 sample periods, i.e., one quarter of the useful symbol time. The preambles transmitted from the first and second antennas use even and odd subcarriers respectively with a $3dB$ power boost, and the values of those subcarriers are set according to [3]. Both a conventional two-path channel and an International Telecommunication Union (ITU) Veh-A channel are simulated with path delays uniformly distributed from 0 to 50 sample periods, where the relative path power profiles are set as 0, 0 (dB) for the two-path channel and 0, -1 , -9 , -10 , -15 , -20 (dB) for the ITU Veh-A channel [16]. The vehicle speed v_e of $240km/hr$ is used to simulate mobile radio environments, for which Rayleigh fading is generated by Jakes' model [17]. Moreover, the multipath observation window \mathbf{W}_b is set as $[0, 108]$. The preassumed number of paths N_h is set as 4 and 8 in the two-path channel and the ITU Veh-A channel, respectively. Both the data subcarrier set Θ and the pilot subcarrier set \mathbf{J} are used in the tracking stage. The subcarrier indices of \mathbf{J} are uniformly assigned within the available subcarriers. The set Θ is uniformly selected from \mathbf{Q} , and the parameter $|\Theta|$ could be 192, 96, 48, 24, 12, 6 or 3. The values of the maximum iteration number V are set as 5 and 7 for QPSK and 16QAM modulation, respectively. For the stopping criterion, the prespecified threshold ε is set as 10^{-4} . The entire simulations are conducted in the equivalent baseband, and we assume both symbol synchronization and carrier synchronization are perfect. Finally, throughout the simulation, the parameter Eb/No is defined as a ratio of received bit energy to the power spectral density of noise.

For comparison purpose, the performance curve with ideal channel estimation, denoted as perfect CSI, is provided for reference and served as a performance lower bound. We also compare the proposed methods with both the STBC-based MMSE method [13] and the Kalman filtering method [18][19] where the decision-feedback methodology is employed under the assumption of ideal channel estimation in the initialization stage. Some statistical information such as Doppler frequency, auto-covariance of channels, and noise power is assumed to be known for these two existing methods. It is noted that the Kalman filtering method is mainly based on [18] and the received signals within a time slot are used to perform channel estimation according to the decision-feedback steps in [19].

B. Simulation Results

Fig. 4 shows the BER performance for QPSK modulation in the two-path channel at $v_e = 240km/hr$ with $|\mathbf{J}| = 8$ and

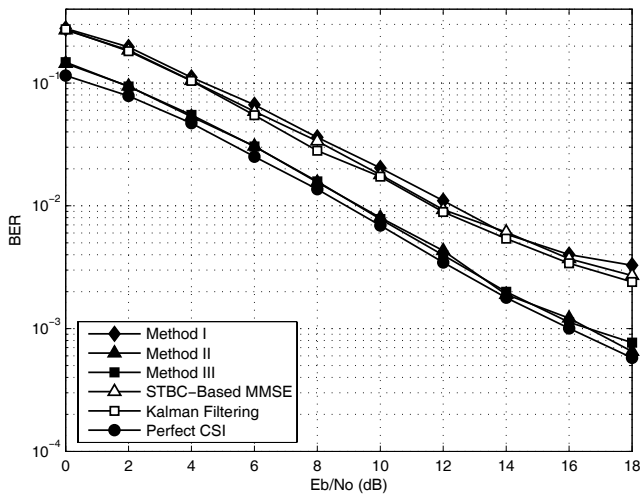


Fig. 4. BER performance for QPSK modulation in the two-path channel at $v_e = 240\text{km/hr}$. ($|\mathbf{J}| = 8$ and $|\Theta| = |\mathbf{Q}| = 192$).

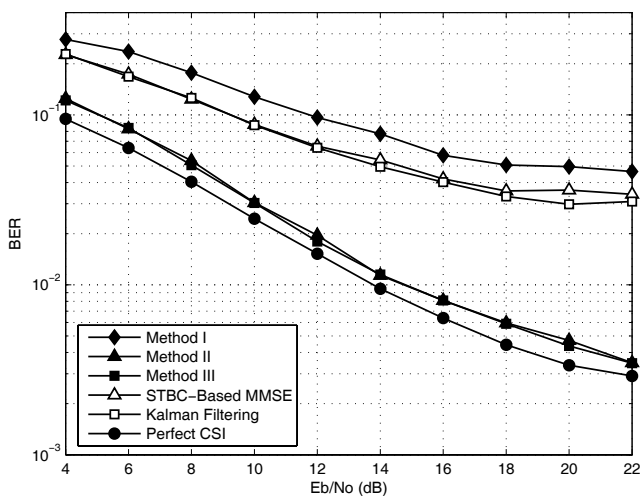


Fig. 5. BER performance for 16QAM modulation in the two-path channel at $v_e = 240\text{km/hr}$. ($|\mathbf{J}| = 8$ and $|\Theta| = |\mathbf{Q}| = 192$).

$|\Theta| = 192$. As can be observed in Fig. 4, at $\text{BER} = 10^{-2}$, the required E_b/N_0 for both the method II and the method III is about 3dB less than that for the method I, and at most 0.2dB more than that for the perfect CSI case. Compared with the STBC-based MMSE method and the Kalman filtering method, the method II and the method III also yield much better BER performance. Fig. 5 and Fig. 6 give the BER performance for 16QAM modulation in the two-path channel and the ITU Veh-A channel, respectively, at $v_e = 240\text{km/hr}$ with $|\mathbf{J}| = 8$ and $|\Theta| = 192$. As shown in Fig. 5, it is clear that with the help from the pilot gradient vector at the first iteration, both the method II and the method III provide a substantial gain in E_b/N_0 , and there is only 0.8dB degradation in the required E_b/N_0 at $\text{BER} = 10^{-2}$ compared with the performance curve of the perfect CSI case. However, for the method I, a higher error floor is clearly visible at a BER of 4×10^{-2} . Obviously, the two existing methods also suffer from severe degradation in the BER performance due to the error propagation effect, even though they are a little better than the method I. Similarly to

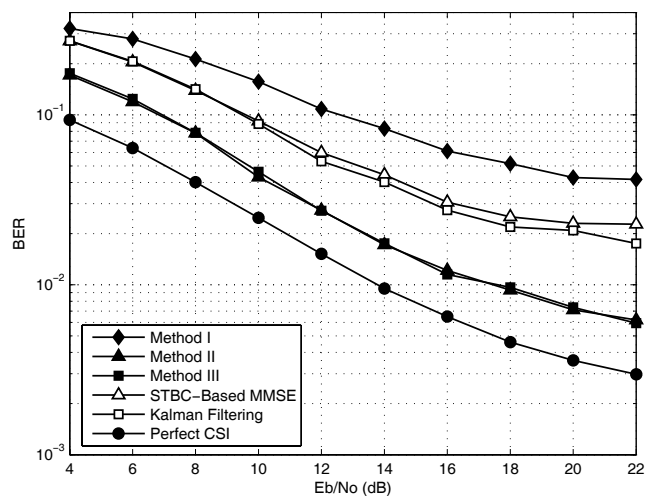


Fig. 6. BER performance for 16QAM modulation in the ITU Veh-A channel at $v_e = 240\text{km/hr}$. ($|\mathbf{J}| = 8$ and $|\Theta| = |\mathbf{Q}| = 192$).

the above results, Fig. 6 shows that both the method II and the method III significantly outperform the method I, whereas they still have 3.6dB gap in E_b/N_0 compared with the perfect CSI case. The two existing methods still have a significant E_b/N_0 gap compared with the method II and the method III. Fig. 7 and Fig. 8 show BER versus f_d in the ITU Veh-A channel for QPSK modulation and 16QAM modulation, respectively, where f_d is the maximum Doppler frequency, normalized by the subcarrier spacing. Clearly, the BER performance gap between the method I and the method II (or the method III) becomes larger as f_d increases up to 0.07 and 0.05 in the cases of QPSK and 16QAM modulation, respectively. It is worthy to mention that the method III is comparable to the method II at different maximum Doppler frequencies, even though an approximate weighting matrix is adopted in the method III. For calibration purpose, the BER performance for the three proposed methods without error propagation is also simulated in these two figures. We observe that the adoption of the optimal gradient vector effectively reduces the error propagation effect and even the BER performance without error propagation is attainable for f_d up to 0.05 and 0.03 in the cases of QPSK and 16QAM modulation, respectively. Fig. 9 shows BER versus $|\mathbf{J}|$ in the ITU Veh-A channel for 16QAM modulation at $v_e = 240\text{km/hr}$ with $|\Theta| = |\mathbf{Q}|$ and $E_b/N_0 = 20\text{dB}$. As can be observed from this figure, the BER performance of the method II (or the method III) converges faster than the method I, while it is improved continually until $|\mathbf{J}|$ increases to 8. Moreover, both the method II and the method III can still give better BER performance even when the number of pilot tones is small. Fig. 10 shows BER versus $|\Theta|$ in the ITU Veh-A channel for 16QAM modulation at $v_e = 240\text{km/hr}$ with $|\mathbf{J}| = 8$ and $E_b/N_0 = 20\text{dB}$. This figure indicates that there is a trade-off between BER and $|\Theta|$, and that one can reduce computational complexity by using a smaller $|\Theta|$ at the cost of slight performance degradation whenever $|\Theta|$ is larger than 48. Fig. 11 compares the average number of iterations in the ITU Veh-A channel for 16QAM modulation at $v_e = 240\text{km/hr}$ with $|\mathbf{J}| = 8$ and

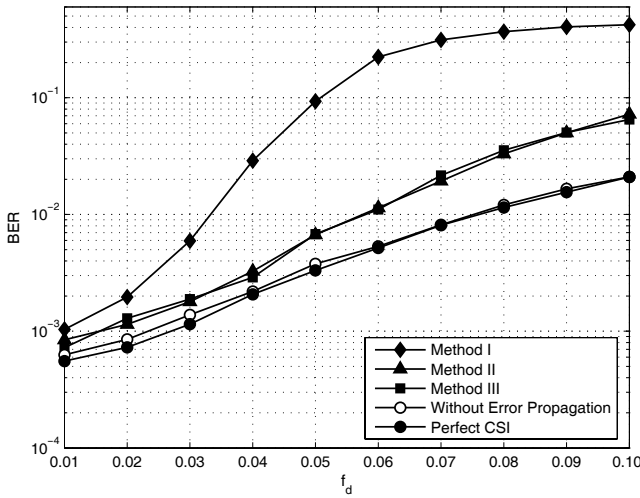


Fig. 7. BER versus normalized Doppler frequency in the ITU Veh-A channel for QPSK modulation. ($|\mathbf{J}| = 8$, $|\Theta| = |\mathbf{Q}| = 192$, and $E_b/N_0=16dB$).

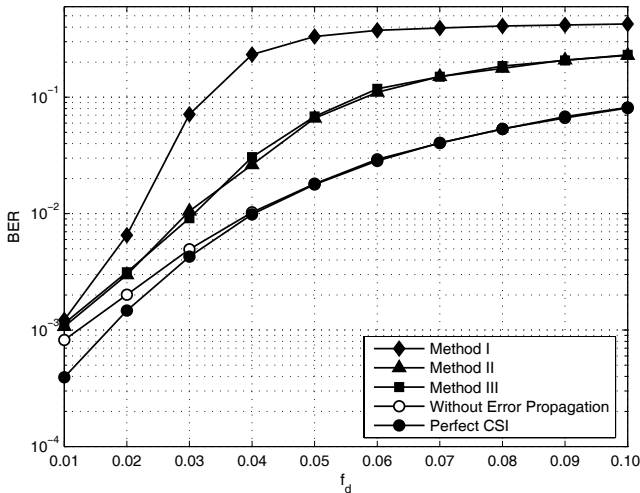


Fig. 8. BER versus normalized Doppler frequency in the ITU Veh-A channel for 16QAM modulation. ($|\mathbf{J}| = 8$, $|\Theta| = |\mathbf{Q}| = 192$, and $E_b/N_0=22dB$).

$E_b/N_0=20dB$. Due to the use of the approximate weighting matrix, when $|\Theta|$ is larger than 48, the method III requires one more iteration compared with the method II. However, the method III is still an attractive approach by decreasing V from 7 to 6 at the price of slight BER performance degradation when the computational complexity of matrix inverse is an issue.

VI. CONCLUSION

In this paper, we present a two-stage channel estimation method for STBC/OFDM systems in mobile wireless channels. In the initialization stage, an MPIC-based decorrelation method is used to estimate multipath delays and multipath complex gains. In the tracking stage, two refined DF DFT-based channel estimation methods are proposed by using a few pilot tones to form an optimal gradient vector at the first iteration, and the optimal step size is directly calculated from the received signals. Further, in order to reduce computational complexity of matrix inverse in the method II, an approximate

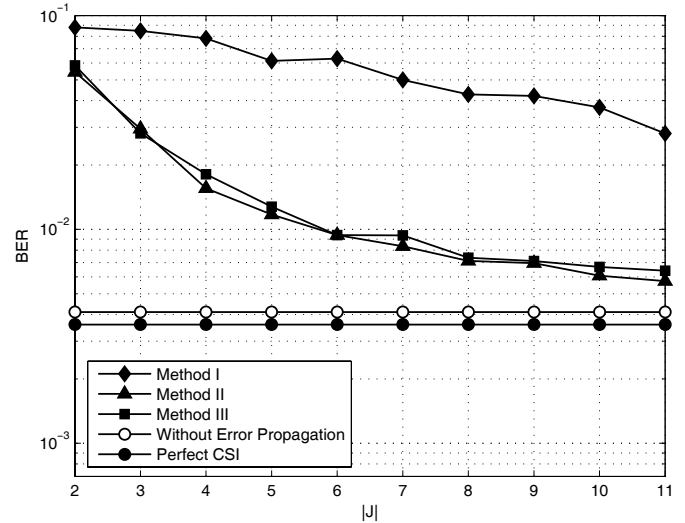


Fig. 9. BER versus number of pilot tones used in the ITU Veh-A channel for 16QAM modulation. ($v_e = 240km/hr$, $|\Theta| = |\mathbf{Q}|$, and $E_b/N_0=20dB$).

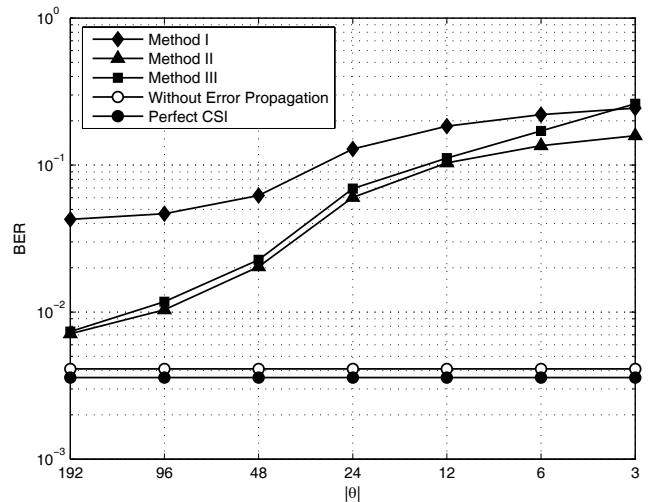


Fig. 10. BER versus number of data subcarriers used in the ITU Veh-A channel for 16QAM modulation. ($v_e = 240km/hr$, $|\mathbf{J}| = 8$, and $E_b/N_0=20dB$).

weighting matrix is proposed and used in the method III. The simulation results show that both the method II and the method III can effectively alleviate the error propagation effect and thus significantly improve the performance of the method I (i.e., the classical DF DFT-based channel estimation method). The two refined methods also outperform the STBC-based MMSE method and the Kalman filtering method, especially when a high-level modulation scheme, e.g. 16QAM, is adopted in high-mobility environments.

REFERENCES

- [1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.
- [2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.
- [3] "Local and metropolitan area networks part 16: air interface for fixed broadband wireless access systems," *IEEE Std 802.16-2004*, Oct. 2004.

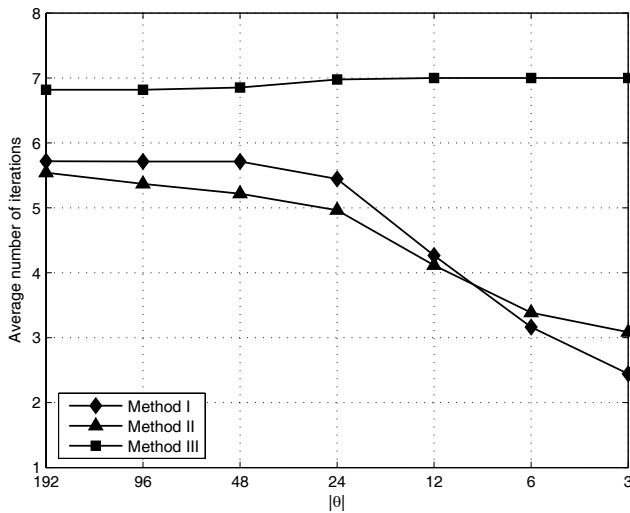


Fig. 11. Average number of iterations versus $|\Theta|$ in the ITU Veh-A channel for 16QAM modulation. ($v_e = 240\text{km/hr}$, $|\mathbf{J}| = 8$, and $E_b/N_0=20\text{dB}$).

- [4] S. Colieri, M. Ergen, A. Puri, and A. Bahai, "A study of channel estimation in OFDM systems," in *Proc. 56th IEEE Veh. Technol. Conf.*, Sept. 2002, pp. 894-898.
- [5] K. S. Ahn and H. K. Baik, "Decision feedback detection for space-time block coding over time-selective fading channels," in *Proc. Personal, Indoor and Mobile Radio Commun.*, Sept. 2003, pp. 1983-1987.
- [6] A. Chini, Y. Wu, M. El-Tanany, and S. Mahmoud, "Filtered decision feedback channel estimation for OFDM-based DTV terrestrial broadcasting system," *IEEE Trans. Broadcasting*, vol. 44, no. 1, pp. 2-11, Mar. 1998.
- [7] S. Werner, M. Enescu, and V. Koivunen, "Low-complexity time-domain channel estimators for mobile wireless OFDM systems," in *Proc. IEEE Workshop on Signal Processing Syst. Design and Implementation*, Nov. 2005, pp. 245-250.
- [8] J.-H. Park, M.-K. Oh, and D.-J. Park, "New channel estimation exploiting reliable decision-feedback symbols for OFDM Systems," in *Proc. Int. Conf. on Commun.*, June 2006, pp. 3046-3051.
- [9] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinx, and M. Engels, "A low-complexity ML channel estimator for OFDM," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 135-140, Feb. 2003.
- [10] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3065-3073, Dec. 2001.

- [11] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "Analysis of DFT-based channel estimators for OFDM," *Wireless Personal Commun.*, vol. 12, Jan. 2000.
- [12] M.-L. Ku and C.-C. Huang, "A derivation on the equivalence between Newton's method and DF DFT-based method for channel estimation in OFDM systems," submitted to *IEEE Trans. Wireless Commun.*
- [13] Y. Gong and K. B. Letaief, "Low complexity channel estimation for space-time coded wideband OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 876-882, Sept. 2003.
- [14] J. S. Arora, *Introduction to Optimum Design*. Elsevier, 2004.
- [15] R. R. Muller and S. Verdu, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 8, pp. 1429-1441, Aug. 2001.
- [16] J. Laiho, A. Wacker, and T. Novosad, *Radio Network Planning and Optimisation for UMTS*. New York: Wiley, 2002.
- [17] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.
- [18] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. IEEE GLOBECOM 2004*, Dec. 2004, pp. 2487-2492.
- [19] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183-186, Feb. 2002.



Meng-Lin Ku was born in Taoyuan, Taiwan. He received the B.S. and M.S. degrees in communication engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2002 and 2003, respectively, where he is currently working towards the Ph.D. degree in the Department of Communication Engineering. His research interests are in the area of digital receiver design and optimization for wireless communication systems.



Chia-Chi Huang was born in Taiwan, R.O.C. He received the B.S. degree in electrical engineering from National Taiwan University in 1977 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1980 and 1984, respectively.

From 1984 to 1988, he was an RF and Communication System Engineer with the Corporate Research and Development Center, General Electric Co., Schenectady, NY, where he worked on mobile radio communication system design. From 1989 to 1992, he was with the IBM T.J. Watson Research Center, Yorktown Heights, NY, as a Research Staff Member, working on indoor radio communication system design. Since 1992, he has been with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan, currently as a Professor.