



Title	Semi-blind CFO, channel estimation and data detection for ofdm systems over doubly selective channels
Author(s)	He, L; Ma, S; Wu, YC; Ng, TS
Citation	The IEEE International Symposium on Circuits and Systems (ISCAS), Paris, France, 30 May-2 June 2010. In Proceedings of ISCAS, 2010, p. 1887-1890
Issued Date	2010
URL	http://hdl.handle.net/10722/126058
Rights	IEEE International Symposium on Circuits and Systems Proceedings. Copyright © IEEE.

Semi-Blind CFO, Channel Estimation and Data Detection for OFDM Systems over Doubly Selective Channels

Lanlan He, Shaodan Ma, Yik-Chung Wu and Tung-Sang Ng

Department of Electrical and Electronic Engineering, The University of Hong Kong

Email: {llhe, sdma, ycwu, tsng}@eee.hku.hk

Abstract—Semi-blind joint CFO, channel estimation and data detection for OFDM systems over doubly selective channels (DSCs) is investigated in this work. A joint iterative algorithm is developed based on the maximum a posteriori expectation-maximization (MAP-EM) algorithm. In addition, a novel algorithm is also proposed to obtain the initial estimates of CFO and channels. Simulation results show that the performance of the proposed CFO and channel estimators approaches to that of the estimators with full training at high SNRs. Moreover, after convergence, the performance of data detection is close to the ideal case with perfect CFO and channel state information.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) has been widely recognized as an efficient transmission technique for wireless communications. Recently, there is an increasing demand for OFDM systems operating in high mobility environment, such as Digital Multimedia Broadcasting (DMB), DVB-H (Digital Video Broadcast-Handheld), Media Forward Link Only (MediaFLO) and Wireless Metropolitan Area Networks (WiMAX) [1]–[3]. For broadband OFDM systems, high speed movement of mobile terminals causes Doppler spread and results in multi-path time-varying channels [4], i.e., doubly selective channels (DSCs). Due to DSCs, the number of channel parameters in one OFDM symbol significantly increases, which makes the channel estimation difficult. Meanwhile, this doubly-selectivity in the channel destroys the orthogonality among subcarriers and induces intercarrier interference (ICI) in OFDM systems, which also complicates the data detection. Moreover, OFDM systems are known to be sensitive to carrier frequency offset (CFO), and frequency synchronization is always a critical part in the design of OFDM receivers. Therefore, CFO, channel estimation and data detection for OFDM systems over DSCs are very challenging.

Over DSCs, a frequency synchronization method exploiting cyclic prefix (CP) has been introduced in [5]. However, the acquisition range of the proposed CFO estimator is only half of the subcarrier spacing thus limiting its applicability to large CFOs. Furthermore, channel estimation and data detection are not considered in [5]. On the other hand, training sequence based CFO and channel estimators have been proposed in

[6], [7]. Unfortunately, a whole OFDM symbol is required for training, which decreases the transmission efficiency.

In this paper, the problem of joint CFO, channel estimation and data detection for OFDM systems over DSCs is addressed. By taking into account the statistics of the parameters to be estimated, a semi-blind maximum a posteriori expectation-maximization (MAP-EM) algorithm is proposed. Using limited number of pilot subcarriers, the MAP-EM algorithm iteratively estimates the CFO, channel and recovers the unknown data. The tentatively recovered data is then exploited to aid the CFO and channel estimations in the next iteration until convergence. In addition to the MAP-EM algorithm, efficient initial CFO and channel estimators are also derived. Simulation results show that, with the efficient initialization developed in this paper, the performance of the proposed CFO and channel estimators approaches to that of the estimators with full training at high SNRs. Moreover, after convergence, the performance of data detection is close to the ideal case with perfect CFO and channel state information (CSI).

Notation: Boldface uppercase and lowercase letters will be used for matrices and vectors. Superscripts H and T denotes Hermitian and transpose respectively. The symbol \mathbf{I}_N denotes an $N \times N$ identity matrix, with \mathbf{e}_l denoting the l^{th} column of \mathbf{I}_N . $\text{diag}\{\mathbf{x}\}$ stands for the diagonal matrix with vector \mathbf{x} on its diagonal. The symbol \otimes denotes the Kronecker product and \odot denotes the Hadamard product. $\mathbb{E}\{\cdot\}$ denotes the expectation. $\text{Tr}\{\mathbf{X}\}$ and $|\mathbf{X}|$ are the trace and the determinant of a square matrix \mathbf{X} respectively. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary parts respectively. The matrix \mathbf{F} is the FFT matrix with $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{N}} e^{-j2\pi mn/N}$.

II. SYSTEM MODEL

In an OFDM system, the source data in frequency domain \mathbf{x} is modulated onto N parallel subcarriers to obtain the time domain signal \mathbf{s} . In general, the elements of \mathbf{x} can be categorized into two sub-vectors \mathbf{x}_p and \mathbf{x}_d , where \mathbf{x}_p contains pilot and \mathbf{x}_d contains data. Assuming pilots and data occupy different subcarriers, the time domain signal \mathbf{s} can be written as

$$\mathbf{s} = \mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbf{x}_d \quad (1)$$

where \mathbf{F}_p collects those rows of \mathbf{F} corresponding to pilot subcarriers, while \mathbf{F}_d collects those rows of \mathbf{F} corresponding to data subcarriers.

The work was supported in part by the Hong Kong Research Grants Council (Grant No.: 7154/08E) and the HKU Seed Funding Programme, Project No. 200811159094.

A cyclic prefix (CP) with length longer than the delay spread of the channel, is inserted at the beginning of each OFDM symbol to prevent intersymbol interference (ISI). The signal is then transmitted through a multi-path time-varying channel which has L independent taps with average power of the l^{th} tap denoted by σ_l^2 . The auto-correlation of the l^{th} channel tap follows the classical Jakes' model [6] given by $\mathbb{E}\{h_l(mT_s)h_l(nT_s)\} = \sigma_l^2 J_0(2\pi f_D(m-n)T_s)$, where $J_0(\cdot)$ represents the zero-order Bessel function of the first kind, f_D represents the maximum Doppler shift, and T_s is the sample interval.

At the receiver side, assuming perfect timing synchronization is achieved, after discarding the CP, the received signal $\mathbf{y} = [y(0), y(1), \dots, y(N-1)]^T$ can be written as

$$\mathbf{y} = \Phi(\varepsilon)\mathbf{H}(\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbf{x}_d) + \mathbf{w} \quad (2)$$

where $\Phi(\varepsilon) = \text{diag}\{1, e^{j2\pi\varepsilon/N}, \dots, e^{j2\pi(N-1)\varepsilon/N}\}$ represents the effect of CFO ε , $\mathbf{w} = [w(0), w(1), \dots, w(N-1)]^T$ is the additive white Gaussian noise vector with zero-mean and covariance $\sigma_w^2 \mathbf{I}_N$, and the matrix \mathbf{H} is defined as

$$\mathbf{H} = \begin{bmatrix} h_0(0) & \mathbf{0} & h_{L-1}(0) & \cdots & h_1(0) \\ h_1(1) & h_0(1) & \mathbf{0} & h_{L-1}(1) & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \mathbf{0} & h_{L-1}(N-1) & \cdots & h_0(N-1) \end{bmatrix} \quad (3)$$

where for notation simplicity, $h_l(n)$ is used to denote $h_l(nT_s)$. Notice that, over DSCs, the channel varies sample by sample, and therefore the number of unknown channel parameter is $N \times L$.

Reversing the position of channel and data, the following equality holds

$$\mathbf{H}\mathbf{s} = \mathbf{D}[\mathbf{s}]\mathbf{h} \quad (4)$$

where $\mathbf{D}[\mathbf{s}] = [\text{diag}\{\Xi_0 \mathbf{s}\}, \dots, \text{diag}\{\Xi_{L-1} \mathbf{s}\}]$ with $\Xi_l = [\mathbf{e}_{l+1}, \dots, \mathbf{e}_N, \mathbf{e}_1, \dots, \mathbf{e}_l]$ and the channel vector \mathbf{h} is defined as $\mathbf{h} = [\mathbf{h}_0^T, \dots, \mathbf{h}_{L-1}^T]^T$ with $\mathbf{h}_l = [h_l(0), \dots, h_l(N-1)]^T$ being the channel coefficients of the l^{th} tap for the whole OFDM symbol. Since each channel tap is independent Gaussian distributed, the probability density function (pdf) of \mathbf{h} follows

$$p(\mathbf{h}) = \frac{1}{(\pi)^{NL} |\mathbf{R}_h|} \exp(-\mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h}) \quad (5)$$

where the covariance matrix is $\mathbf{R}_h = \mathbf{R}_L \otimes \mathbf{J}$ with $\mathbf{R}_L = \text{diag}\{\sigma_0^2, \dots, \sigma_{L-1}^2\}$ and the $(k, m)^{\text{th}}$ entry of \mathbf{J} being $[\mathbf{J}]_{k,m} = J_0(2\pi f_D(m-k)T_s)$. The equality given by (4) will be used in the derivation of our proposed algorithm.

III. SEMI-BLIND CFO, CHANNEL ESTIMATION AND DATA DETECTION

First, notice that $(\varepsilon, \mathbf{h}, \mathbf{x}_d)$ are the unknown parameters that should be estimated. Because the noise is Gaussian distributed, based on (2), the likelihood function for $(\varepsilon, \mathbf{h}, \mathbf{x}_d)$ is given by

$$p(\mathbf{y}|\mathbf{h}, \varepsilon, \mathbf{x}_d) = \frac{1}{(\pi\sigma_w^2)^N} \exp\left(-\frac{1}{\sigma_w^2} \|\mathbf{y} - \Phi(\varepsilon)\mathbf{H}(\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbf{x}_d)\|^2\right). \quad (6)$$

Here we exploit the fact that \mathbf{H} is completely characterized by \mathbf{h} and therefore conditioning on \mathbf{H} is equivalent to conditioning on \mathbf{h} . The joint maximum likelihood (ML) CFO, channel estimation and data detection problem over DSC based on (6) is an ill-posed one. A direct ML solution based on (6) is not available.

To overcome the problem discussed above, we propose an iterative joint CFO, channel estimation and data detection method based on MAP-EM algorithm [8]. For the derivation, the received signal \mathbf{y} is referred as incomplete data and we take $(\mathbf{y}, \mathbf{x}_d)$ as the complete data. The MAP-EM algorithm alternates between the E-step and the M-step until convergence, and these two steps at the i^{th} iteration are [8]:

E-step: Compute

$$Q(\varepsilon, \mathbf{h}|\hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}) = \mathbb{E}\{\log p(\varepsilon, \mathbf{h}|\mathbf{y}, \mathbf{x}_d)|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\};$$

M-step: Solve $(\hat{\varepsilon}^i, \hat{\mathbf{h}}^i) = \arg \max Q(\varepsilon, \mathbf{h}|\hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$.

The expectation in the E-step is with respect to the conditional pdf $p(\mathbf{x}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$. The symbol $\hat{\varepsilon}^{i-1}$ and $\hat{\mathbf{h}}^{i-1}$ are the $(i-1)^{\text{th}}$ estimate of the CFO and channel respectively. Notice that a data estimator will be obtained as a byproduct of the E-step [9]. The derivations of E-step and M-step are detailed as follows.

E-step:

Using Bayes' rules, we have

$$p(\varepsilon, \mathbf{h}|\mathbf{y}, \mathbf{x}_d) = \frac{p(\mathbf{y}|\varepsilon, \mathbf{h}, \mathbf{x}_d)p(\mathbf{h})p(\varepsilon)}{p(\mathbf{y}|\mathbf{x}_d)} \quad (7)$$

where the pdf $p(\mathbf{y}|\varepsilon, \mathbf{h}, \mathbf{x}_d)$ and $p(\mathbf{h})$ are specified in (6) and (5) respectively. The prior pdf of the CFO is represented by $p(\varepsilon)$, and if there is no prior information on CFO, we can set $p(\varepsilon) = \text{constant}$. Notice that $p(\mathbf{y}|\mathbf{x}_d)$ is generally not available, however, it is independent of ε and \mathbf{h} . Putting (5) and (6) into (7), and dropping those terms independent of ε and \mathbf{h} , we have

$$\begin{aligned} \log p(\varepsilon, \mathbf{h}|\mathbf{y}, \mathbf{x}_d) &\propto 2\Re\{\mathbf{y}^H \Phi(\varepsilon)\mathbf{H}(\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbf{x}_d)\} \\ &\quad - (\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbf{x}_d)^H \mathbf{H}^H \mathbf{H} (\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbf{x}_d) \\ &\quad - \sigma_w^2 \mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h} + \sigma_w^2 \log p(\varepsilon). \end{aligned} \quad (8)$$

The function $Q(\varepsilon, \mathbf{h}|\hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$ to be maximized is then computed as

$$\begin{aligned} Q(\varepsilon, \mathbf{h}|\hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}) &= 2\Re\{\mathbf{y}^H \Phi(\varepsilon)\mathbf{H}(\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \mathbb{E}\{\mathbf{x}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\})\} \\ &\quad - 2\Re\{\mathbf{x}_p^H \mathbf{F}_p^H \mathbf{H}^H \mathbf{H} \mathbf{F}_d^H \mathbb{E}\{\mathbf{x}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\}\} \\ &\quad - \text{Tr}\{\mathbf{H}^H \mathbf{H} \mathbf{F}_d^H \mathbb{E}\{\mathbf{x}_d \mathbf{x}_d^H|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\} \mathbf{F}_d\} \\ &\quad - \mathbf{x}_p^H \mathbf{F}_p^H \mathbf{H}^H \mathbf{H} \mathbf{F}_p^H \mathbf{x}_p - \sigma_w^2 \mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h} + \sigma_w^2 \log p(\varepsilon) \end{aligned} \quad (9)$$

where $\mathbb{E}\{\mathbf{x}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\}$ is the conditional mean of \mathbf{x}_d and $\mathbb{E}\{\mathbf{x}_d \mathbf{x}_d^H|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\}$ is its conditional correlation matrix.

Notice that

$$\begin{aligned} \mathbb{E}\{\mathbf{x}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\} &= \sum_{\tilde{\mathbf{x}}_d \in C^{N_d}} p(\tilde{\mathbf{x}}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}) \tilde{\mathbf{x}}_d \\ &\approx \hat{\mathbf{x}}_d^i \end{aligned} \quad (10)$$

where $\hat{\mathbf{x}}_d^i = \arg \max_{\tilde{\mathbf{x}}_d \in C^{N_d}} p(\tilde{\mathbf{x}}_d|\mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$ and the approximation is accurate at high signal-to-noise ratio (SNR).

Here $\hat{\mathbf{x}}_d^i$ is in fact the MAP estimate of the data [9] which has a prohibitively high computational complexity. In order to reduce the complexity, the MAP estimator is replaced by $\hat{\mathbf{x}}_d^i = \text{demod}(\bar{\mathbf{x}}_d^i)$, where $\text{demod}(\cdot)$ represents the operation of constellation mapping and $\bar{\mathbf{x}}_d^i$ is the MMSE estimate of \mathbf{x}_d given $\hat{\varepsilon}^{i-1}$ and $\hat{\mathbf{h}}^{i-1}$ as

$$\bar{\mathbf{x}}_d^i = (\sigma_w^2 \mathbf{\Lambda}_d^{-1} + \mathbf{F}_d \mathbf{H}_{i-1}^H \mathbf{H}_{i-1} \mathbf{F}_d^H)^{-1} \mathbf{F}_d (\mathbf{H}_{i-1}^H \mathbf{\Phi}^H(\hat{\varepsilon}^{i-1}) \mathbf{y} - \mathbf{H}_{i-1}^H \mathbf{H}_{i-1} \mathbf{F}_p^H \mathbf{x}_p) \quad (11)$$

where $\mathbf{\Lambda}_d$ is a diagonal matrix whose elements depend on the average power of \mathbf{x}_d . Similarly, we have

$$\mathbb{E}\{\mathbf{x}_d \mathbf{x}_d^H | \mathbf{y}, \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}\} \approx \hat{\mathbf{x}}_d^i (\hat{\mathbf{x}}_d^i)^H. \quad (12)$$

Substituting (11) and (12) into (9), it follows

$$\begin{aligned} Q(\varepsilon, \mathbf{h} | \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}) &= 2\Re\{\mathbf{y}^H \mathbf{\Phi}(\varepsilon) \mathbf{H}(\mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \hat{\mathbf{x}}_d^i)\} \\ &- 2\Re\{\mathbf{x}_p^H \mathbf{F}_p \mathbf{H}^H \mathbf{H} \mathbf{F}_d^H \hat{\mathbf{x}}_d^i\} \\ &- \text{Tr}\{\mathbf{H}^H \mathbf{H} \mathbf{F}_d^H \hat{\mathbf{x}}_d^i (\hat{\mathbf{x}}_d^i)^H \mathbf{F}_d\} - \mathbf{x}_p^H \mathbf{F}_p \mathbf{H}^H \mathbf{H} \mathbf{F}_p^H \mathbf{x}_p \\ &- \sigma_w^2 \mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h} + \sigma_w^2 \log p(\varepsilon). \end{aligned} \quad (13)$$

M-step:

In the M-step, we maximize $Q(\varepsilon, \mathbf{h} | \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$ in (13) with respect to ε and \mathbf{h} . However, notice that \mathbf{H} depends on \mathbf{h} in a nonlinear way, direct maximization of (13) with respect to \mathbf{h} is difficult. In the following we derive an alternative expression for $Q(\varepsilon, \mathbf{h} | \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$ from which a closed-form solution of \mathbf{h} can be obtained.

We first rewrite (13) as

$$\begin{aligned} Q(\varepsilon, \mathbf{h} | \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}) &= 2\Re\{\mathbf{y}^H \mathbf{\Phi}(\varepsilon) \mathbf{H} \hat{\mathbf{s}}^i\} - (\hat{\mathbf{s}}^i)^H \mathbf{H}^H \mathbf{H} \hat{\mathbf{s}}^i \\ &- \sigma_w^2 \mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h} + \sigma_w^2 \log p(\varepsilon) \end{aligned} \quad (14)$$

with $\hat{\mathbf{s}}^i = \mathbf{F}_p^H \mathbf{x}_p + \mathbf{F}_d^H \hat{\mathbf{x}}_d^i$. Due to the fact that $\mathbf{H} \hat{\mathbf{s}}^i = \mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{h}$, we have

$$\begin{aligned} Q(\varepsilon, \mathbf{h} | \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1}) &= 2\Re\{\mathbf{y}^H \mathbf{\Phi}(\varepsilon) \mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{h}\} - \mathbf{h}^H \mathbf{D}^H[\hat{\mathbf{s}}^i] \mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{h} \\ &- \sigma_w^2 \mathbf{h}^H \mathbf{R}_h^{-1} \mathbf{h} + \sigma_w^2 \log p(\varepsilon). \end{aligned} \quad (15)$$

By setting the first derivative of $Q(\varepsilon, \mathbf{h} | \hat{\varepsilon}^{i-1}, \hat{\mathbf{h}}^{i-1})$ in (15) with respect to \mathbf{h} to zero, \mathbf{h} minimizing the function above for any ε follows as

$$\hat{\mathbf{h}}^i = (\mathbf{D}^H[\hat{\mathbf{s}}^i] \mathbf{D}[\hat{\mathbf{s}}^i] + \sigma_w^2 \mathbf{R}_h^{-1})^{-1} \mathbf{D}^H[\hat{\mathbf{s}}^i] \mathbf{\Phi}^H(\varepsilon) \mathbf{y}. \quad (16)$$

Notice that the inverse of \mathbf{R}_h is required but does not exist. Using matrix inverse lemma, we have

$$\hat{\mathbf{h}}^i = \mathbf{R}_h \mathbf{D}^H[\hat{\mathbf{s}}^i] (\mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{R}_h \mathbf{D}^H[\hat{\mathbf{s}}^i] + \sigma_w^2 \mathbf{I}_N)^{-1} \mathbf{\Phi}^H(\varepsilon) \mathbf{y} \quad (17)$$

where the inverse of \mathbf{R}_h is avoided.

Putting (17) into (15), we have

$$\begin{aligned} Q(\varepsilon | \hat{\varepsilon}^{i-1}) &= \mathbf{y}^H \mathbf{\Phi}(\varepsilon) \mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{R}_h \mathbf{D}^H[\hat{\mathbf{s}}^i] (\mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{R}_h \mathbf{D}^H[\hat{\mathbf{s}}^i] \\ &+ \sigma_w^2 \mathbf{I}_N)^{-1} \mathbf{\Phi}^H(\varepsilon) \mathbf{y} + \sigma_w^2 \log p(\varepsilon). \end{aligned} \quad (18)$$

The i^{th} estimate of the CFO $\hat{\varepsilon}^i$ can be obtained by one-dimension search.

In summary, the proposed MAP-EM algorithm iterates among the following equations until $|\hat{\varepsilon}^i - \hat{\varepsilon}^{i-1}| < K_\varepsilon$ and $\|\hat{\mathbf{h}}^i - \hat{\mathbf{h}}^{i-1}\| < K_h$. For $i = 1, 2, \dots$

$$\hat{\mathbf{x}}_d^i = \text{demod}(\bar{\mathbf{x}}_d^i);$$

$$\hat{\varepsilon}^i = \arg \max Q(\varepsilon | \hat{\varepsilon}^{i-1});$$

$$\hat{\mathbf{h}}^i = \mathbf{R}_h \mathbf{D}^H[\hat{\mathbf{s}}^i] (\mathbf{D}[\hat{\mathbf{s}}^i] \mathbf{R}_h \mathbf{D}^H[\hat{\mathbf{s}}^i] + \sigma_w^2 \mathbf{I}_N)^{-1} \mathbf{\Phi}^H(\hat{\varepsilon}^i) \mathbf{y}$$

where $\bar{\mathbf{x}}_d^i$ and $Q(\varepsilon | \hat{\varepsilon}^{i-1})$ are defined in (11) and (18) respectively, and K_ε and K_h are predefined thresholds to terminate the algorithm.

Remark 1: In case there is no prior information of CFO and all the subcarriers are training, the E-step is not necessary and we can set $\hat{\mathbf{s}}^i = \mathbf{s}$ in (17) and (18). The resultant channel and CFO estimators are equivalent to those proposed in [6].

IV. INITIALIZATION

A good initialization is essential to EM algorithm, and therefore, the problem now becomes how to obtain the initial estimates of CFO and channel. Notice that the received signal is $y(n) = e^{j2\pi\varepsilon n/N} \sum_{l=0}^{L-1} h_l(n) s(n-l) + w(n)$. Since each tap of the channel is independent identically distributed, when L is large enough, with the central limit theory, the undistorted signal (i.e., $\sum_{l=0}^{L-1} h_l(n) s(n-l)$) is approximately Gaussian, and accordingly $p(\mathbf{y} | \varepsilon)$ is approximately Gaussian.

Now we compute the mean and covariance matrix of \mathbf{y} given ε . Since the channel is zero-mean, $\mathbb{E}\{\mathbf{y} | \varepsilon\} = \mathbf{0}$. After lengthy but straightforward manipulations (omitted due to limited space), the covariance matrix can be given by

$$\begin{aligned} \mathbb{E}\{\mathbf{y} \mathbf{y}^H | \varepsilon\} &= \mathbf{\Phi}(\varepsilon) (\mathbf{D}[\mathbf{F}_p^H \mathbf{x}_p] \mathbf{R}_h \mathbf{D}^H[\mathbf{F}_p^H \mathbf{x}_p] \\ &+ \mathbf{J} \odot (\mathbf{F}_d^H \mathbf{\Lambda}_d \mathbf{F}_d) + \sigma_w^2 \mathbf{I}_N) \mathbf{\Phi}^H(\varepsilon). \end{aligned} \quad (19)$$

Therefore, based on the approximate Gaussian distribution $p(\mathbf{y} | \varepsilon)$, according to (19) and the fact that $\mathbb{E}\{\mathbf{y} | \varepsilon\} = \mathbf{0}$, the initial CFO estimator follows ML criteria as

$$\begin{aligned} \hat{\varepsilon}^0 &= \arg \min_{\varepsilon} \mathbf{y}^H \mathbf{\Phi}(\varepsilon) (\mathbf{D}[\mathbf{F}_p^H \mathbf{x}_p] \mathbf{R}_h \mathbf{D}^H[\mathbf{F}_p^H \mathbf{x}_p] \\ &+ \mathbf{J} \odot (\mathbf{F}_d^H \mathbf{\Lambda}_d \mathbf{F}_d) + \sigma_w^2 \mathbf{I}_N)^{-1} \mathbf{\Phi}^H(\varepsilon) \mathbf{y}. \end{aligned} \quad (20)$$

After CFO estimation, initial channel estimation is now considered. The system model in (2) can be rewritten as

$$\mathbf{y} = \mathbf{\Phi}(\varepsilon) \mathbf{D}[\mathbf{F}_p^H \mathbf{x}_p] \mathbf{h} + \mathbf{\Phi}(\varepsilon) \mathbf{H} \mathbf{F}_d \mathbf{x}_d + \mathbf{w}. \quad (21)$$

By treating \mathbf{x}_d as interference, the MMSE channel estimate is then given by

$$\begin{aligned} \hat{\mathbf{h}}^0 &= \mathbf{R}_h \mathbf{D}^H[\mathbf{F}_p^H \mathbf{x}_p] (\mathbf{D}[\mathbf{F}_p^H \mathbf{x}_p] \mathbf{R}_h \mathbf{D}^H[\mathbf{F}_p^H \mathbf{x}_p] \\ &+ \mathbf{J} \odot (\mathbf{F}_d^H \mathbf{\Lambda}_d \mathbf{F}_d) + \sigma_w^2 \mathbf{I}_N)^{-1} \mathbf{\Phi}^H(\hat{\varepsilon}^0) \mathbf{y}. \end{aligned} \quad (22)$$

Remark 2: In the extreme case that all data are training, \mathbf{s} becomes known and $p(\mathbf{y} | \varepsilon)$ is definitely Gaussian. The initial CFO and channel estimators in (20) and (22) reduce to those proposed in [6].

V. SIMULATION RESULTS AND DISCUSSIONS

In this section, the performance of the proposed algorithm for joint CFO, channel estimation and data detection is demonstrated by Monte Carlo simulations, where each point is obtained by averaging over $M = 10^4$ runs. Each OFDM symbol has 128 subcarriers ($N=128$) and the length of CP is 8.

Carrier frequency is $f_c = 2$ GHz, the sample interval $T_s = 2\mu s$ and the speed of vehicle is $v = 219$ km/hr, which results in the normalized maximal Doppler shift $N f_D T_s = 0.1$. The channel has six taps ($L = 6$) with an exponential power delay profile, which is generated according to Jakes' model. The CFO is assumed to follow a uniform distribution on the interval $[-0.5, 0.5]$. We set thresholds $K_\epsilon = 0.001$ and $K_h = 0.01$. The pilot structure in [10] is adopted. More specifically, ten pilot clusters are used with each cluster occupying three pilots, where only one non-zero pilot is transmitted in the middle of the cluster. This means that roughly 23.4% of the subcarriers are occupied by pilots. The non-zero pilots are generated following complex Gaussian distribution and the data are chosen from 16 QAM constellation.

In Fig. 1, the MSEs of CFO and channel estimates versus the number of iterations are depicted for different SNRs. It can be seen that the performance of CFO and channel estimation improves significantly in the first iteration. Furthermore, the MSEs converge to stable values quickly and there is no observable improvement after about five iterations.

Fig. 2 shows the MSE of the channel estimate as a function of SNR. The proposed initial channel estimator is labeled as 'Iter=0'. The channel estimator derived for full training [6] is also shown for comparison. As can be seen, though the density of pilots is low (only 23.4% of the subcarriers are occupied by pilots), as the number of iteration increases, the performance of the proposed channel estimator improves continually. After five iterations, it performs almost the same as that with full training at high SNRs. Similar results are also observed for CFO estimation and the figure is not presented here due to space limitation.

Fig. 3 shows the BER performance of the proposed algorithm as a function of SNR. For comparison, the ideal case which assumes perfect CFO and CSI is also depicted. It is clear that, after convergence, the BER performance is close to the ideal case.

VI. CONCLUSIONS

Semi-blind CFO, channel estimation and data detection for OFDM systems over DSCs was developed based on the MAP-EM algorithm. For initialization, an approximate ML CFO estimator and a MMSE channel estimator were derived. Simulation results demonstrated that the proposed MAP-EM algorithm converges in a few iterations and moreover, at high SNRs, the performance of CFO and channel estimators approaches to that of the estimators with full training. Moreover, after convergence, the BER performance is close to the ideal case with perfect CFO and channel state information.

REFERENCES

- [1] J. Gozalvez, "Moving beyond 3G," *IEEE Vehicular Technology Magazine*, vol. 1, no. 2, pp. 48-55, 2006.
- [2] C. Yeun, "Mobile TV technologies," in *Proc. IEEE ICICT*, pp.2-9, Dec. 2007.
- [3] J. Zhou, Z. Ou, M. Rautiainen, T. Koskela, M. Ylianttila, "Digital television for mobile devices," *IEEE Multimedia*, vol. 16, no. 1, pp. 60-71, Jan.-March 2009.
- [4] Y. Q. Zhou, J. Wang and M. Sawhashi, "Downlink transmission of broadband OFCDM systems-part II: effect of Doppler shift," *IEEE Trans. Commun.*, vol.54, no.6, pp. 1097-1108, Jun. 2006.

- [5] T. Lv, H. Li, and J. Chen, "Joint estimation of symbol timing and carrier frequency offset of OFDM signals over fast time-varying multipath channels," *IEEE Trans. Signal Processing*, vol.53, no.12, pp. 4526-4536, Dec. 2005.
- [6] N. Ricklin and J. Zeidler, "Data-aided joint estimation of carrier frequency offset and frequency-selective time-varying channel," in *Proc. IEEE ICC*, pp. 5098 - 5102, May 2008.
- [7] Q. Jiang, J. Speidel, C. Zhao, "A novel carrier frequency offset estimation for OFDM systems over time-varying multipath channels," *Journal Wireless Personal Commun.*, vol 49, no. 4, pp. 587-596, June 2009.
- [8] G. McLachlan, *The EM algorithm and extensions*. U.S.: Wiley, 2007.
- [9] J. Choi, *Adaptive and iterative signal processing in communications*. U.K.: Cambridge, 2006.
- [10] A. Kannu and P. Schniter, "Design and analysis of MMSE pilot-aided cyclic-prefixed block transmissions for doubly selective channels," *IEEE Trans. Signal Processing*, vol. 56, no.3, pp. 1148-1160, Mar. 2008.

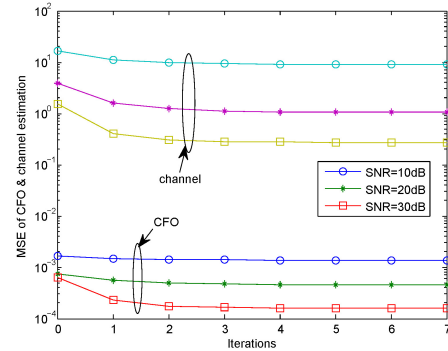


Fig. 1. MSE of CFO and channel estimation versus number of iterations

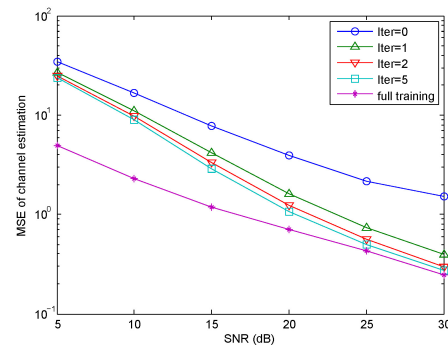


Fig. 2. MSE of channel estimation versus different SNRs

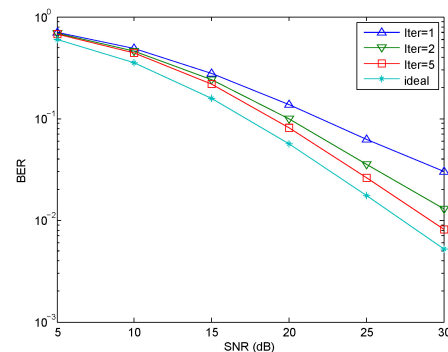


Fig. 3. BER versus different SNRs