A Low-Complexity Semi-Blind Joint CFO and Data Estimation Algorithm for OFDM systems

Kilbom Lee, Sung-Hyun Moon, and Inkyu Lee, *Senior Member, IEEE*

School of Electrical Engineering, Korea University, Seoul, Korea
Email: {bachhi, shmoon, inkyu}@korea.ac.kr

***Abstract*—In this paper, we propose a low-complexity semi-blind joint carrier frequency offset (CFO) and data estimation algorithm for orthogonal frequency division multiplexing (OFDM) systems. Given channel information, we first provide a new iterative algorithm which jointly estimates the CFO and data based on pilots by minimizing the mean square error between the received OFDM symbol and its regenerated signal. By using the matrix inversion lemma, the joint CFO and data estimator is divided into a CFO estimator and a data detector without loss of optimality, which significantly reduces the computational complexity. Also, we present a decision feedback strategy to select reliable data from previously detected data by adopting the probability metric which evaluates the reliability. Then, the simplified CFO estimator can utilize the selected reliable data as pilots in the next iteration step. Simulation results show that the simplified CFO estimator can achieve the average Cramer Rao bound in moderate and high signal to noise ratio (SNR) regions within a few iterations even for a small number of pilots with the help of the proposed decision feedback strategy.**

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is the most popular scheme for broadband communications because of its high spectral efficiency and simple receiver implementation [1]. However, it has been shown that the OFDM is sensitive to carrier frequency offset (CFO) which comes from Doppler shifts or transceiver oscillator instabilities. The CFO causes inter-carrier interference (ICI) among subcarriers, which results in severe bite error rate (BER) performance degradation [2]. Thus, many algorithms have been proposed to estimate the CFO in OFDM for decades [3]–[11].

In [3], the channel-independent CFO estimator is provided based on a preamble, but it is not preferable with respect to the spectral efficiency. In contrast, the pilot-aided algorithms in [4] and [5] can provide the trade-off between the performance and the spectral efficiency with low complexity based on the assumption that the channel and the CFO are static over two consecutive OFDM symbols. However, the schemes do not consider the ICI, which results in worse performance and slower convergence rate for large CFO.

Meanwhile, the blind CFO estimators need high computational complexity, while reducing pilot overhead. Recently, the authors in [9] proposed a blind CFO estimator under the assumption that channel information is known to the receiver, which shows much better performance over conventional blind algorithms in [7] and [8]. However, it requires grid search, and its theoretical lower bound is not provided.

In this paper, we propose a new iterative algorithm which jointly estimates the CFO and data based on pilots for OFDM systems. Here, we assume that perfect channel estimation is available at the receiver as in [9] and [10]. Then, it will be shown that the joint CFO and data estimator can be decoupled into a pure CFO estimator and a pure data detector without loss of optimality, which significantly reduces the computational complexity. Also, a decision feedback strategy is presented to select reliable decision data from previously detected data in an uncoded system by employing the probability metric which evaluates the reliability. Then, the simplified CFO estimator can utilize the selected reliable data as pilots in the next iteration step. Furthermore, we provide the Cramer Rao bound (CRB) for different numbers of pilots when perfect channel estimation is given. Simulation results show that the simplified CFO estimator aided by the proposed decision feedback strategy can achieve the average CRB of a preamble[1] in moderate and high signal to noise ratio (SNR) regions within a few iterations even with a small number of pilots. As a result, our proposed semi-blind algorithm provides the optimal performance with substantially low computational complexity at the expense of a small of spectral efficiency loss.

Throughout this paper, normal letters represent scalar quantities, boldface lowercase letters indicate vectors, and boldface uppercase letters designate matrices. We use $(\cdot)^T$, $(\cdot)^\dagger$ and $\|\cdot\|$ for transpose, complex conjugate transpose and the 2-norm operation, respectively. Also, $\Re\{c\}$ and $\Im\{c\}$ represent the real and imaginary components of c, respectively. The subscripts $[\cdot]_k$ and $[\cdot]_{i,j}$ designate the k-th element of a vector and the (i,j)-th entry of a matrix, respectively. $\mathbf{I}_l$ denotes an identity matrix of size $l \times l$. Moreover, diag$\{\mathbf{c}\}$ stands for a diagonal matrix whose diagonal elements are defined by $\mathbf{c}$.

II. SIGNAL MODEL

In this paper, we consider OFDM systems with N subcarriers and N_g cyclic prefix length. The CFO is defined as ϵ which is normalized by the subcarrier spacing. We define the time-domain channel impulse response vector $\mathbf{h}$

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2010-0017909).

[1]In this paper, we assume that the number of pilots in a preamble is the same as that of subcarriers.

as $\mathbf{h} = [h_0\ h_1\ \cdots\ h_{L-1}]^T$ where h_i has an independent and identically distributed (i.i.d.) complex Gaussian distribution and L represents the channel length. Let us denote $\mathbf{F} = [\mathbf{f}_0\ \mathbf{f}_1 \cdots \mathbf{f}_{N-1}]$ as the $N \times N$ discrete Fourier transform (DFT) matrix with $\mathbf{f}_i = \frac{1}{\sqrt{N}}[1\ e^{\frac{-j2\pi i}{N}} \cdots\ e^{\frac{-j2\pi i(N-1)}{N}}]^T$, and define $\mathbf{F}_L = [\mathbf{f}_0\ \mathbf{f}_1 \cdots \mathbf{f}_{L-1}]$. Then, the diagonal channel matrix $\mathbf{H}$ is given by $\mathbf{H} = \text{diag}\{\mathbf{F}_L\mathbf{h}\}$.

Then, the received OFDM symbol is represented in the frequency domain as

$$\begin{aligned} \mathbf{r} &= \mathbf{F}\boldsymbol{\Gamma}(\epsilon)\mathbf{F}^\dagger\mathbf{H}\mathbf{x} + \mathbf{w} \qquad (1) \\ &= \mathbf{C}(\epsilon)\mathbf{H}\mathbf{x} + \mathbf{w} \end{aligned}$$

where $\mathbf{x}$ represents the transmitted signal, $\boldsymbol{\Gamma}(\epsilon) = \text{diag}\{1,\ e^{\frac{j2\pi\epsilon}{N}}, \cdots, e^{\frac{j2\pi(N-1)\epsilon}{N}}\}$ equals a diagonal matrix whose diagonal entry stands for the phase shift of the received signal sample, $\mathbf{w}$ indicates the complex additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $\sigma_w^2\mathbf{I}_N$, and $\mathbf{C}(\epsilon) = \mathbf{F}\boldsymbol{\Gamma}(\epsilon)\mathbf{F}^\dagger$ is a circulant matrix with $\mathbf{C}^\dagger(\epsilon)\mathbf{C}(\epsilon) = \mathbf{I}_N$ which denotes the normalized interference matrix caused by CFO. Note that $\mathbf{C}(\epsilon)$ becomes an identity matrix when $\epsilon = 0$.

III. Proposed Estimation Algorithms

A. Overview of Conventional Estimation Algorithms

In this subsection, we overview conventional CFO estimation algorithms in OFDM based on the assumption that channel information is known to the receiver. Given channel information and a preamble, the linear minimum mean square error (LMMSE)-based CFO estimator in [10] is given by

$$\hat{\epsilon} = \frac{N}{2\pi\sum_{n=1}^{N-1} n^2|y_n|^2}\sum_{p=1}^{N-1} p|y_p|^2 \tan^{-1}\frac{\Im\{r_p y_p^*\}}{\Re\{r_p y_p^*\}}$$

where we have $\mathbf{x} = [x_1\ x_2\ \cdots\ x_N]^T$ and $\mathbf{y} = \mathbf{H}\mathbf{x} = [y_1\ y_2\ \cdots\ y_N]^T$. To alleviate the needs of the preamble, the authors in [9] proposed a iterative blind CFO estimator as

$$\begin{aligned} \hat{\epsilon}^{(i)} &= \arg\min_{\tilde{\epsilon}} \|\mathbf{r} - \mathbf{C}(\tilde{\epsilon})\mathbf{H}\hat{\mathbf{x}}\|^2 \\ &= \arg\min_{\tilde{\epsilon}} \|\mathbf{r} - \mathbf{C}(\tilde{\epsilon})\mathbf{H}\zeta\big(\mathbf{H}^{-1}\mathbf{C}(-\tilde{\epsilon})\mathbf{r}\big)\|^2 \end{aligned}$$

where $\hat{\epsilon}^{(i)}$ represents the i-th CFO estimate and $\hat{\mathbf{x}}$ denotes $\zeta\big(\mathbf{H}^{-1}\mathbf{C}(-\hat{\epsilon}^{(i)})\mathbf{r}\big)$ and ζ represents the ML detection operation which is defined as

$$[\zeta(\mathbf{a})]_n = \arg\min_{\rho_j} |[\mathbf{a}]_n - \rho_j|^2$$

where ρ_j stands for the j-th constellation symbol candidate for $j = 1, \cdots, K$ with K being the constellation size. Note that this algorithm needs ML detection of N transmitted symbols for each candidate $\tilde{\epsilon}$, which requires high computational complexity. To reduce the search size of $\tilde{\epsilon}$, the number of grid points M is set to a small value. Instead, for the i-th iteration and $\epsilon \in [-0.5,\ 0.5)$, the search range of $\tilde{\epsilon}$ is determined by $\left[\ \hat{\epsilon}^{(i-1)} - \frac{M-1}{2(M^i-1)},\ \hat{\epsilon}^{(i-1)} + \frac{M-1}{2(M^i-1)}\ \right]$ with M being an odd number and $\hat{\epsilon}^{(0)} = 0$. After N_{itr} iterations, the precision achieved is about $M^{-N_{itr}}$.

B. Proposed Joint CFO and Data Estimator

In this subsection, we propose a joint CFO and data estimation algorithm based on channel information, which provides a trade-off between complexity and spectral efficiency. First, we define the index sets of pilots and data as $\mathcal{C}_p$ and $\mathcal{C}_d$, respectively. Also, the size of $\mathcal{C}_p$ and $\mathcal{C}_d$ are given by N_p and N_d, respectively. Thereby, we denote a diagonal matrix $\tilde{\boldsymbol{\Phi}}_p$ where $[\tilde{\boldsymbol{\Phi}}_p]_{i,i} = 1$ for $i \in \mathcal{C}_p$ and $[\tilde{\boldsymbol{\Phi}}_p]_{i,i} = 0$ otherwise. Extracting the columns whose indices are included in $\mathcal{C}_p$ from $\tilde{\boldsymbol{\Phi}}_p$ yields $\boldsymbol{\Phi}_p \in \mathbb{C}^{N\times N_p}$. Similarly, $\boldsymbol{\Phi}_d \in \mathbb{C}^{N\times N_d}$ is defined with $\mathcal{C}_d$.

Then, the received OFDM symbol $\mathbf{r}$ can be rewritten as

$$\begin{aligned} \mathbf{r} &= \mathbf{C}(\epsilon)\mathbf{H}\mathbf{x} + \mathbf{w} \\ &= \mathbf{C}(\epsilon)\mathbf{H}\boldsymbol{\Phi}_p\mathbf{x}_p + \mathbf{C}(\epsilon)\mathbf{H}\boldsymbol{\Phi}_d\mathbf{x}_d + \mathbf{w} \qquad (2) \end{aligned}$$

where $\mathbf{x}_p$ and $\mathbf{x}_d$ stand for the transmitted pilot vector of length N_p and the transmitted data vector of length N_d, respectively. Note that the LMMSE-based CFO estimator in [10] works only when $\boldsymbol{\Phi}_p = \mathbf{I}_N$, while the blind CFO estimator in [9] can be used regardless of $\boldsymbol{\Phi}_p$. Here, we see that equation (2) has $N_d + 1$ unknown parameters $\{\epsilon, \mathbf{x}_d\}$.

Now, we develop an iterative algorithm to jointly estimate the CFO and data $\mathbf{x}_d$ based on a line search method. Let us define the unknown parameter vector and the vector estimated at the i-th iteration as $\mathbf{v} = [\epsilon\ \mathbf{x}_d^T]^T$ and $\hat{\mathbf{v}}^{(i)} = [\hat{\epsilon}^{(i)}\ (\hat{\mathbf{x}}_d^{(i)})^T]^T$, respectively. By using $\mathbf{r}$ in (2) and $\hat{\mathbf{v}}^{(i)}$, we can restore $\mathbf{r}^{(i)}$ as

$$\mathbf{r}^{(i)} = \mathbf{C}(\hat{\epsilon}^{(i)})\mathbf{H}\hat{\mathbf{x}}^{(i)}$$

where $\hat{\mathbf{x}}^{(i)} = \boldsymbol{\Phi}_p\mathbf{x}_p + \boldsymbol{\Phi}_d\hat{\mathbf{x}}_d^{(i)}$.

It can be seen in (1) that $\mathbf{r}$ is a nonlinear function of $\mathbf{v}$ because of ϵ. By applying the linear approximation presented in [12], we have an approximated expression as a linear function of $\mathbf{v}$

$$\mathbf{r} \approx \mathbf{r}^{(i)} + \mathbf{G}(\mathbf{v} - \hat{\mathbf{v}}^{(i)}) + \mathbf{w} \qquad (3)$$

where the gradient matrix $\mathbf{G} \in \mathbb{C}^{N\times(N_d+1)}$ is defined as

$$\mathbf{G} \triangleq \left[\frac{\partial\mathbf{r}}{\partial\epsilon}\ \frac{\partial\mathbf{r}}{\partial\mathbf{d}}\right]_{\mathbf{v}=\hat{\mathbf{v}}^{(i)}} = \big[\mathbf{g}_1\ \mathbf{G}_2\big].$$

Here, $\mathbf{g}_1$ and $\mathbf{G}_2$ are derived, respectively, as

$$\begin{aligned} \mathbf{g}_1 &\triangleq \left.\frac{\partial\mathbf{r}}{\partial\epsilon}\right|_{\mathbf{v}=\hat{\mathbf{v}}^{(i)}} = \frac{j2\pi}{N}\mathbf{F}\mathbf{M}\boldsymbol{\Gamma}(\hat{\epsilon}^{(i)})\mathbf{F}^\dagger\mathbf{H}\hat{\mathbf{x}}^{(i)} \\ \mathbf{G}_2 &\triangleq \left.\frac{\partial\mathbf{r}}{\partial\mathbf{d}}\right|_{\mathbf{v}=\hat{\mathbf{v}}^{(i)}} = \mathbf{F}\boldsymbol{\Gamma}(\hat{\epsilon}^{(i)})\mathbf{F}^\dagger\mathbf{H}\boldsymbol{\Phi}_d \end{aligned}$$

where we have $\mathbf{M} = \text{diag}\{0,\ \cdots, N-1\}$. Then, equation (3) can be iteratively solved by a line search method as [12] [13]

$$\hat{\mathbf{v}}^{(i+1)} = \hat{\mathbf{v}}^{(i)} + \big(\mathbf{G}^\dagger\mathbf{G}\big)^{-1}\mathbf{G}^\dagger(\mathbf{r} - \mathbf{r}^{(i)}). \qquad (4)$$

Here, the proposed solution (4) requires matrix by matrix multiplications and an inverse operation of size $N_d + 1$.

C. Low-Complexity CFO Estimator

To reduce the complexity, we now propose a low-complexity solution without loss of optimality. By using the previously estimated CFO $\hat{\epsilon}^{(i)}$, we can compensate the CFO from $\mathbf{r}$ as

$$\begin{aligned}\mathbf{z} &= \mathbf{F}\mathbf{\Gamma}(-\hat{\epsilon}^{(i)})\mathbf{F}^{\dagger}\mathbf{r} \\ &= \mathbf{F}\mathbf{\Gamma}(\epsilon-\hat{\epsilon}^{(i)})\mathbf{F}^{\dagger}\mathbf{H}\mathbf{x}+\mathbf{F}\mathbf{\Gamma}(-\hat{\epsilon}^{(i)})\mathbf{F}^{\dagger}\mathbf{w} \\ &= \mathbf{C}(\Delta\epsilon^{(i+1)})\mathbf{H}\mathbf{x}+\mathbf{n} \end{aligned} \tag{5}$$

where we have $\Delta\epsilon^{(i+1)} = \epsilon - \hat{\epsilon}^{(i)}$ and $\mathbf{n} = \mathbf{C}(-\hat{\epsilon}^{(i)})\mathbf{w}$. Here, $\mathbf{n}$ has the same distribution as $\mathbf{w}$ since $\mathbb{E}[\mathbf{n}\mathbf{n}^{\dagger}] = \mathbf{C}(-\hat{\epsilon}^{(i)})\mathbb{E}[\mathbf{w}\mathbf{w}^{\dagger}]\mathbf{C}(\hat{\epsilon}^{(i)}) = \sigma_w^2\mathbf{I}_N$. Therefore, we can obtain $\hat{\epsilon}^{(i+1)} = \hat{\epsilon}^{(i)} + \Delta\hat{\epsilon}^{(i+1)}$ without any performance loss by estimating the residual CFO $\Delta\hat{\epsilon}^{(i+1)}$ from $\mathbf{z}$.

In this case, the solution (4) can be transformed as

$$\hat{\mathbf{b}}^{(i+1)} = \hat{\mathbf{b}}^{(i)} + \left(\mathbf{G}^{\dagger}\mathbf{G}\right)^{-1}\mathbf{G}^{\dagger}(\mathbf{z}-\mathbf{z}^{(i)}) \tag{6}$$

where $\hat{\mathbf{b}}^{(i)} = [0\ (\hat{\mathbf{x}}_d^{(i)})^T]^T$, $\hat{\mathbf{b}}^{(i+1)} = [\Delta\hat{\epsilon}^{(i+1)}\ (\hat{\mathbf{x}}_d^{(i+1)})^T]^T$ and $\mathbf{z}^{(i)} = \mathbf{H}\hat{\mathbf{x}}^{(i)}$. Especially, $\mathbf{g}_1$ and $\mathbf{G}_2$ are changed to

$$\mathbf{g}_1 = \frac{j2\pi}{N}\mathbf{F}\mathbf{M}\mathbf{F}^{\dagger}\mathbf{H}\hat{\mathbf{x}}^{(i)}, \quad \mathbf{G}_2 = \mathbf{H}\mathbf{\Phi}_d \tag{7}$$

where $\Delta\hat{\epsilon}^{(i)}$ is assumed to be zero.

To further simplify the solution (6), we employ the matrix inversion lemma [14]. Then, from the joint CFO and data estimator in (6), a pure CFO estimator is extracted as

$$\Delta\hat{\epsilon}^{(i+1)} = \mathbf{g}_1^{\dagger}\left[\mathbf{I}_N - \mathbf{G}_2(\mathbf{G}_2^{\dagger}\mathbf{G}_2)^{-1}\mathbf{G}_2^{\dagger}\right](\mathbf{z}-\mathbf{z}^{(i)})/\alpha \tag{8}$$

where $\alpha = \mathbf{g}_1^{\dagger}\left[\mathbf{I}_N - \mathbf{G}_2(\mathbf{G}_2^{\dagger}\mathbf{G}_2)^{-1}\mathbf{G}_2^{\dagger}\right]\mathbf{g}_1$. On the other hand, based on $\hat{\epsilon}^{(i+1)} = \hat{\epsilon}^{(i)} + \Delta\hat{\epsilon}^{(i+1)}$, we can obtain $\hat{\mathbf{x}}_d^{(i+1)}$ as

$$\hat{\mathbf{x}}_d^{(i+1)} = \mathbf{H}^{-1}\mathbf{\Gamma}(-\hat{\epsilon}^{(i+1)})\mathbf{r}.$$

In (8), the projection matrix to the null space of $\mathbf{G}_2$ can be rewritten as

$$\mathbf{I}_N - \mathbf{G}_2(\mathbf{G}_2^{\dagger}\mathbf{G}_2)^{-1}\mathbf{G}_2^{\dagger} = \mathbf{I}_N - \mathbf{\Phi}_d\mathbf{\Phi}_d^{\dagger} = \mathbf{\Phi}_p\mathbf{\Phi}_p^{\dagger}. \tag{9}$$

By substituting (9) to (8), the CFO estimator can be simplified to

$$\Delta\hat{\epsilon}^{(i+1)} = \mathbf{g}_1^{\dagger}\mathbf{\Phi}_p\mathbf{\Phi}_p^{\dagger}(\mathbf{z}-\mathbf{z}^{(i)})/\beta \tag{10}$$

where $\beta = \|\mathbf{\Phi}_p^{\dagger}\mathbf{g}_1\|^2$. It can be seen that $\mathbf{G}_2$ is not involved in equation (10), and the computational complexity reduces from $O(N^2(N_d+1)+(N_d+1)^3)$ to $O(N\log_2 N)$ which comes from the computation of $\mathbf{z}$ in (5) and $\mathbf{g}_1$ in (7). We note that the solution (10) yields the same $\Delta\hat{\epsilon}^{(i+1)}$ as in the solution (6) for given $\hat{\mathbf{b}}^{(i)}$.

Now, we derive the CRB for the CFO for the joint estimation of CFO ϵ and data $\mathbf{x}_d$. By applying the results in [14] and [15] to (1), the CRB for the CFO is obtained as

$$\mathrm{CRB}(\epsilon) = \frac{\sigma_w^2}{2}\Big(\Re\big\{\mathbf{g}_1^{\dagger}\big(\mathbf{I}_N - \mathbf{G}_2(\mathbf{G}_2^{\dagger}\mathbf{G}_2)^{-1}\mathbf{G}_2^{\dagger}\big)\mathbf{g}_1\big\}\Big)^{-1} \tag{11}$$

where $\mathbf{g}_1 = \frac{j2\pi}{N}\mathbf{F}\mathbf{M}\mathbf{\Gamma}(\epsilon)\mathbf{F}^{\dagger}\mathbf{H}\mathbf{x}$ and $\mathbf{G}_2 = \mathbf{F}\mathbf{\Gamma}(\epsilon)\mathbf{F}^{\dagger}\mathbf{H}\mathbf{\Phi}_d$. Similar to (9), the term inside in (11) can be denoted as

$$\begin{aligned}\mathbf{g}_1^{\dagger}\big(\mathbf{I}_N - \mathbf{C}(\epsilon)\mathbf{\Phi}_d\mathbf{\Phi}_d^{\dagger}\mathbf{C}^{\dagger}(\epsilon)\big)\mathbf{g}_1 &= \mathbf{g}_1^{\dagger}\mathbf{C}(\epsilon)\mathbf{\Phi}_p\mathbf{\Phi}_p^{\dagger}\mathbf{C}^{\dagger}(\epsilon)\mathbf{g}_1 \\ &= \|\mathbf{\Phi}_p^{\dagger}\mathbf{C}^{\dagger}(\epsilon)\mathbf{g}_1\|^2. \end{aligned} \tag{12}$$

By substituting (12) to (11), the CRB can be expressed as a simplified form

$$\mathrm{CRB}(\epsilon) = \frac{\sigma_w^2}{2\|\mathbf{\Phi}_p^{\dagger}\mathbf{t}_1\|^2} = \frac{\sigma_w^2}{2\sum_{k\in\mathcal{C}_p}|[\mathbf{t}_1]_k|^2} \tag{13}$$

where $\mathbf{t}_1 = \mathbf{C}^{\dagger}(\epsilon)\mathbf{g}_1$. Here, it is clearly shown that the CRB decreases as N_p grows. Therefore, if the previous decision data is used as pilots in the next iteration step, we can expect more improved performance. Note that the blind CFO estimator in [9] considers all previous decision data as pilots, but it may result in error propagation due to incorrect decision data. In contrast, the algorithm in [6] employs the channel decoder output to obtain reliable decision data. However, channel decoding and re-encoding are required, which causes high computational complexity. Thus, we introduce a way to select reliable decision data in an uncoded system in the following subsection.

D. Decision Feedback Strategy

In this subsection, we present a decision feedback strategy to select the reliable data from the previous decision data $\zeta(\hat{\mathbf{x}}_d^{(i)})$. Note that the simplified CFO estimator in (10) uses the reliable data as pilots in the next iteration step.

For notation simplicity, we define $[\mathbf{x}_d]_n$, $[\hat{\mathbf{x}}_d^{(i)}]_n$ and $[\zeta(\mathbf{x}_d^{(i)})]_n$ as s_n, $\hat{s}_n^{(i)}$ and $\hat{\rho}_n^{(i)}$, respectively. Thereby, the normalized conditional probability for the candidate symbol $\hat{\rho}_n^{(i)}$ at the n-th data s_n is represented as [16]

$$\eta_n^{(i)} = \frac{p(\hat{s}_n^{(i)}|s_n = \hat{\rho}_n^{(i)})}{\sum_{j=1}^{K} p(\hat{s}_n^{(i)}|s_n = \rho_j)} \tag{14}$$

where $0 < \eta_n^{(i)} < 1$ and the conditional probability density function (PDF) of $\hat{s}_n^{(i)}$ is given by

$$p(\hat{s}_n^{(i)}|s_n = \rho_j) = \frac{1}{\pi\sigma_w^2}\exp\Big(-\frac{|\hat{s}_n^{(i)}-\rho_j|^2}{\sigma_w^2}\Big).$$

Motivated by the fact that a more reliable candidate has higher $\eta_n^{(i)}$, we address how to choose reliable decision data from $\zeta(\hat{\mathbf{x}}_d^{(i)})$.

for $n = 1, 2, \cdots, N_d^{(i)}$

1) Find $\hat{\rho}_n^{(i)}$ closest to $\hat{s}_n^{(i)}$.
2) Compute (14) based on $\hat{\rho}_n^{(i)}$ and $\hat{s}_n^{(i)}$.
3) Choose $\hat{\rho}_n^{(i)}$ as a pilot which satisfies $\eta_n^{(i)} > \gamma$ where γ stands for the threshold value with $0 < \gamma < 1$.

end for

4) Delete the chosen elements from $\hat{\mathbf{x}}_d^{(i)}$.

Let us denote the new set[2] which includes the indices of the i-th chosen decision data as $\mathcal{B}_p^{(i)}$. Thereby, we update $\mathcal{C}_p$ and $\mathcal{C}_d$ as

$$\mathcal{C}_p^{(i)} = \mathcal{C}_p^{(i-1)} \cup \mathcal{B}_p^{(i)}, \quad \mathcal{C}_d^{(i)} = \mathcal{C}_d^{(i-1)} \setminus \mathcal{B}_p^{(i)}$$

where $\mathcal{C}_p^{(0)} = \mathcal{C}_p$ and $\mathcal{C}_d^{(0)} = \mathcal{C}_d$. Moreover, $N_p^{(i)}$ and $N_d^{(i)}$ represent the size of $\mathcal{C}_p^{(i)}$ and $\mathcal{C}_d^{(i)}$, respectively. Then, $\mathbf{\Phi}_p^{(i)}$ and $\mathbf{\Phi}_d^{(i)}$ are newly defined by $\mathcal{C}_p^{(i)}$ and $\mathcal{C}_d^{(i)}$, respectively. By using $\mathbf{\Phi}_p^{(i)}$ and the i-th chosen decision data, we can update $\hat{\mathbf{x}}_p^{(i)}$. It is noted that $\hat{\mathbf{x}}_p^{(i)}$ should be updated by elements chosen at the i-th iteration such as $\hat{\rho}_n^{(i)}$ not $\hat{s}_n^{(i)}$ and its size increases to $N_p^{(i)}$, while the size of $\hat{\mathbf{x}}_d^{(i)}$ reduces from $N_d^{(i-1)}$ to $N_d^{(i)}$ in the step 4) in the above algorithm. Then, we perform the CFO estimator (10) to obtain $\Delta\hat{\epsilon}^{(i+1)}$ based on $\hat{\mathbf{x}}_p^{(i)}$ and $\hat{\mathbf{x}}_d^{(i)}$.

IV. Simulation Results

In this section, we show the BER and the mean square error (MSE) of the CFO estimator of our proposed algorithms in OFDM systems with $N = 128$ and $N_g = 16$. We assume an 8-tap Rayleigh fading exponential decaying channel, with uncoded QPSK, and $2,000$ simulation runs are performed. In addition, ϵ is chosen from a uniform distribution within a range $[-0.4, 0.4]$. Also, $\mathcal{C}_p$ is randomly generated with $\mathcal{C}_d = \{1, \cdots, N\} \setminus \mathcal{C}_p$. For brevity, we refer to the LMMSE-based CFO estimator in [10] and the blind CFO estimator in [9] which minimizes reconstruction error as LMMSE and MRE, respectively. Moreover, the estimator based on joint weighted least square in [4] is referred to as JWLS. Also, we call the simplified line-search based CFO estimator in (10) as SLS and refer to that with the proposed decision feedback strategy (DFS) as DFS-SLS. For the MRE, the number of grid points M and iterations N_{itr} are set to 41 and 3, respectively. Moreover, we set the iteration number of the SLS and the JWLS[3] to 2 and 4, respectively. In addition, for the DFS-SLS, the threshold value γ and N_{itr} are set to 0.99 and 3, respectively, to achieve the CRB of $N_p = 128$ at SNR 35 dB.

Fig. 1 provides the comparison of computational complexity of our proposed algorithms and conventional algorithms where the line search based joint CFO and data estimator (LS) indicates the solution (4). Here, we assume $N_p = N/16$, $N_d = N - N_p$, $M = 40$ and QPSK modulation. It can be seen that the MRE has the highest computational complexity next to the LS, because the ML detection of N transmitted symbols should be performed for each candidate $\tilde{\epsilon}$. In contrast, the computational complexity of the LMMSE is lowest, but a preamble is required. For the JWLS, the computational complexity is slightly higher than that of the LS. As expected, the LS has extremely high computational complexity due to a matrix inverse operation, while the SLS and the DFS-SLS show significantly reduced complexity.

[2]Note that s_n denotes the n-th element of $\mathbf{x}_d$ with $1 \le n \le N_d$, while $m \in \mathcal{B}_p^{(i)}$ indicates the m-th index in $\hat{\mathbf{x}}^{(i)}$ with $1 \le m \le N$.

[3]Similar to the SLS, the JWLS obtains the CFO by iteratively compensating the previously estimated CFO and estimating the residual CFO from two consecutive received OFDM symbols.

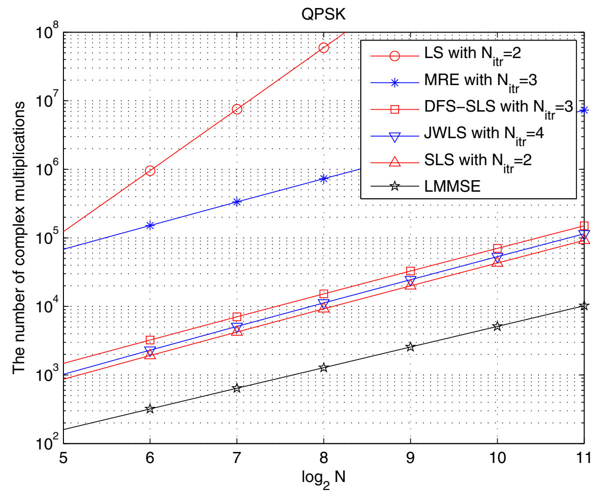

Figure 1. Computation Complexity of CFO estimation algorithms

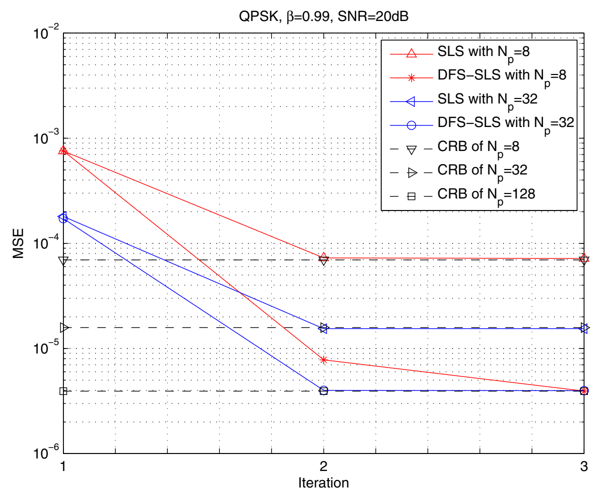

Figure 2. MSE of SLS and DFS-SLS with respect to iterations

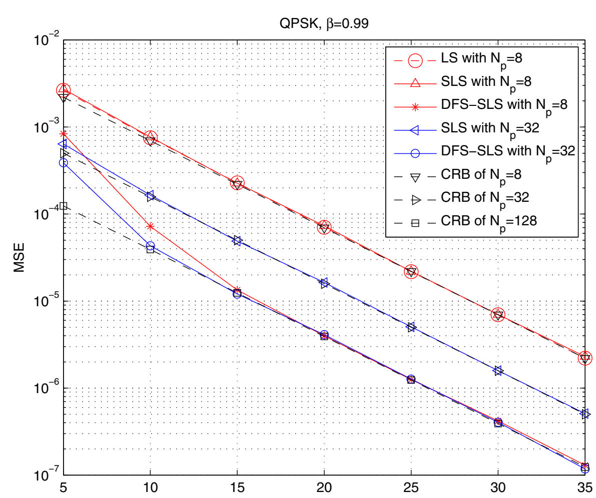

Figure 3. MSE of SLS and DFS-SLS with respect to SNR

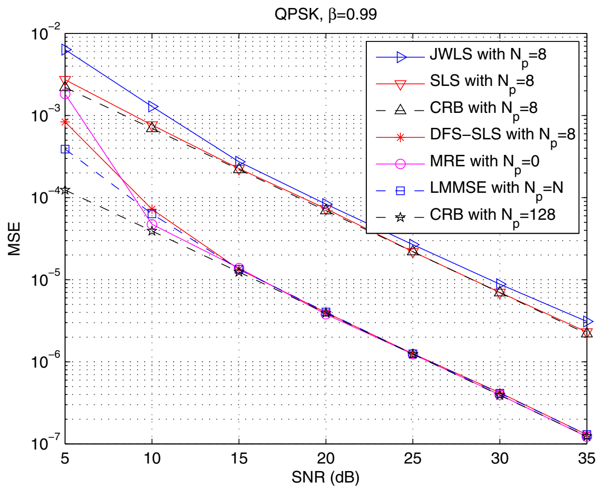

Figure 4. MSE comparison of algorithms

Fig. 2 illustrates the MSE of the SLS and the DFS-SLS with different numbers of iterations and SNR of 20 dB. Here, we present the average CRB with different numbers of pilots which is obtained by averaging the CRB in (13) over several simulation runs with different random CFOs and channel gains. We see that the DFS-SLS achieves the CRB of $N_p = 128$ within only two iterations regardless of N_p, which means that sufficiently large number of reliable decision data is obtained after just one iteration. Note that the DFS-SLS with $N_p = 8$ has the similar performance and the convergence rate as that with $N_p = 32$. Therefore, the DFS-SLS with $N_p = 8$ is more preferable in terms of the spectral efficiency.

In Fig. 3, we exhibit the MSE of the SLS and the DFS-SLS according to SNR. It is verified that the SLS has the same performance of the LS, as expected. Similar to the results in Fig. 2, the DFS-SLS achieves the CRB with $N_p = 128$ in moderate and high SNR regions. In contrast, at low SNR, there is a small gap between DFS-SLS and CRB with $N_p = 128$. This is due to the fact that sufficiently large number of reliable decision data can not be obtained at low SNR. Nevertheless, the DFS-SLS yields substantially improved performance over the SLS even at low SNR.

Fig. 4 compares the MSE of the SLS and DFS-SLS with conventional algorithms. It can be seen that the JWLS yields worse performance over the SLS. This can be explained by the fact that the JWLS ignores the ICI due to the residual CFO unlike the SLS. Similar to the DFS-SLS, the conventional algorithms also produce worse performance at low SNR, while the CRB is achieved in moderate and high SNR regions. However, note that the LMMSE is simpler than the MRE, but a preamble is required. In conclusion, our proposed semi-blind scheme produces almost the same performance as conventional algorithms with low complexity at the expense of a small spectral efficiency loss.

V. Conclusions

In this paper, we have developed a joint CFO and data estimation algorithm for OFDM systems. By applying the matrix inversion lemma, a pure CFO estimator has been provided without loss of optimality. Moreover, we have presented a decision feedback strategy to choose reliable decision data from previous decision data by utilizing the probability metric which evaluates the reliability. We have observed from simulation results that the simplified CFO estimator with small numbers of pilots can achieve the average CRB for a preamble in moderate and high SNR regions within a few iterations with the help of the proposed decision feedback strategy. As a result, our proposed semi-blind algorithm provides the optimal performance in moderate and high SNR regions, while ensuring high spectral efficiency with substantially low computational complexity.

References

[1] J. A. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communications Magazine*, pp. 5–14, May 1990.

[2] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Transactions on Communications*, vol. 43, pp. 191–193, Feburary/March/April 1995.

[3] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Communications Letters*, vol. 3, pp. 75–77, March 1999.

[4] P.-Y. Tsai, H.-Y. Kang, and T.-D. Chiueh, "Joint Weighted Least-Squares Estimation of Carrier-Frequency Offset and Timing Offset for OFDM Systems Over Multipath Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 1, pp. 211–223, January 2005.

[5] P.-H. Chiang, D.-B. Lin, H.-J. Li, and G. L. Stuber, "Joint Estimation of Carrier-Frequency and Sampling-Frequency Offsets for SC-FDE Systems on Multipath Fading Channels," *IEEE Transactions on Comunnications*, vol. 56, pp. 1231–1235, August 2008.

[6] F. Z. Merli and G. M. Vitetta, "Iterative ML-based estimation of carrier frequency offset, channel impulse response and data in OFDM transmissions," *IEEE Transactions on Communications*, vol. 56, pp. 497–506, March 2008.

[7] X. N. Zeng and A. Ghrayeb, "A blind carrier frequency offset estimation scheme for OFDM systems with constant modulus signaling," *IEEE Transactions on Communications*, vol. 56, pp. 1032–1037, July 2008.

[8] L. Wu, X.-D. Zhang, P.-S. Li, and Y.-T. Su, "A blind CFO estimator based on smoothing power spectrum for OFDM systems," *IEEE Transactions on Communications*, vol. 57, pp. 1924–1927, July 2009.

[9] X. Li, E. Like, Z. Wu, and M. Temple, "Highly Accurate Blind Carrier Frequency Offset Estimator for Mobile OFDM Systems," in *Proc. IEEE ICC '10*, pp. 1–5, May 2010.

[10] H.-Y. Liu and R. Y. Yen, "Effective Adaptive Iteration Algorithm for Frequency Tracking and channel Estimation in OFDM Systems," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 2093–2097, May 2010.

[11] K. Lee, S.-H. Moon, and I. Lee, "Robust Pilot Designs for Consistent Frequency Offset Estimation in OFDM systems," in *Proc. IEEE GLOBECOM '10*, pp. 1–5, December 2010.

[12] Y. Na and H. Minn, "Line Search Based Iterative Joint Estimation of Channels and Frequency Offsets for Uplink OFDMA Systems," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 4374–4382, December 2007.

[13] X. N. Zeng, A. Ghrayeb, and D. Mao, "CFO Estimation for Uplink OFDM systems: An application of the variable projection method," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 2306–2311, May 2009.

[14] G. A. F. Seber and C. J. Wild, *Nonlinear Regression*. Wiley, New York, 1989.

[15] O. Besson and P. Stoica, "On parameter estimation of MIMO flat-fading channels with frequency offsets," *IEEE Transactions on Signal Processing*, vol. 51, pp. 602–613, March 2003.

[16] J.-S. Kim, S.-H. Moon, and I. Lee, "A New Reduced Complexity ML Detection Scheme for MIMO Systems," *IEEE Transactions on Communications*, vol. 58, pp. 1302–1310, April 2010.