A Pilot-aided Frequency Offset Estimation Algorithm for OFDMA Uplink 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 pilot aided carrier frequency offset (CFO) estimation algorithm for orthogonal frequency division multiplexing access (OFDMA) uplink systems based on two consecutive received OFDMA symbols. Assuming that the channels and the CFOs are static over the two consecutive symbols, we express the second received OFDMA symbol in terms of the CFOs and the first OFDMA symbol. Based on this signal model, a new estimation algorithm which obtains the CFOs by minimizing the mean square distance between the received OFDMA symbol and its regenerated signal is provided. The simulation results show that the proposed algorithm approaches the average Cramer Rao bound for moderate and high signal to noise ratio regions. Moreover, the algorithm can be applied for any carrier assignment schemes.**

I. INTRODUCTION

Orthogonal frequency division multiplexing access (OFDMA) is a promising transmission system due to its high spectrum efficiency and robustness to inter-symbol interference [1]. In the OFDMA, mobile users (MUs) can be simultaneously served by allocating a group of subcarriers according to a carrier assignment scheme (CAS), which includes a subband-based CAS (SCAS), an interleaved CAS (ICAS) and a generalized CAS (GCAS) [2]. However, the OFDMA is sensitive to carrier frequency offset (CFO) which arises from Doppler shifts or transceiver oscillator instabilities. The CFO results in severe inter-carrier interference (ICI) and multiple access interference (MAI) among subcarriers, which significantly degrade the bit error rate (BER) performance.

A CFO estimation problem can be classified into an acquisition stage and a tracking stage. In the acquisition stage, the CFO estimates are usually obtained at the beginning of each new frame by utilizing a training sequence which comprises only symbols known to the base station (BS). There have been several papers devoted to the estimation of CFOs based on the training sequence for uplink OFDMA systems [3]–[7]. The main advantage of using the training sequence is that the BS can easily regenerate the received signal based on the estimated CFOs and channels [4]–[6].

In general, a pilot-aided algorithm is preferable in terms of spectral efficiency compared to a train sequence based algorithm. A gain in the spectral efficiency is more apparent when the acquisition step is needed frequently. However, it is difficult for a pilot-aided estimation algorithm to regenerate the received signal because the BS does not know all of the transmitted data. Therefore, it is more challenging to estimate the CFOs by using pilots for uplink OFDMA systems. For this reason, only a few works have been reported to solve the problem in uplink OFDMA systems [8]–[10]. Recently, the authors in [9] and [10] provided a pilot-aided acquisition scheme and a tracking algorithm for uplink OFDMA systems, respectively, based on two consecutive OFDMA symbols which have a tile structure[1].

In this paper, we will develop a new pilot-aided algorithm for uplink OFDMA systems with the GCAS. Assuming that the channels and the CFOs of MUs are static over two consecutive OFDMA symbols, we represent the second OFDMA symbol as the signal which includes the CFOs and the first OFDMA symbol. Then, we approximate this nonlinear model by a linear one in order to adopt a line search method. Thereby, an estimation algorithm is provided which obtains the CFOs by minimizing the mean square distance between the received second OFDMA symbol and its regenerated signal. The simulation results show that the proposed scheme approaches the average Cramer Rao bound (CRB) for moderate and high signal to noise ratio (SNR) regions for any CAS.

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 $\|\cdot\|^2$ for transpose, complex conjugate transpose and the 2-norm operation, respectively. In addition, $\Re\{c\}$ 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 in a vector and the (i,j)-th entry in a matrix, respectively. $\mathbf{I}_l$ and $\mathbf{0}_{l\times n}$ denote an identity matrix of size $l\times l$ and a zero matrix of size $l\times n$, respectively. Also, diag$\{\mathbf{x}\}$ stands for a diagonal matrix whose diagonal elements are defined by $\mathbf{x}$.

II. SIGNAL MODEL

In this paper, we consider uplink OFDMA systems with N subcarriers, N_g cyclic prefix length and K MUs. Let us define the time-domain channel impulse response vector $\mathbf{h}_k$ as $\mathbf{h}_k = [h_{0,k}\ h_{1,k}\ \cdots\ h_{L-1,k}]^T$ where $h_{i,k}$ has an independent and identically distributed (i.i.d.) complex Gaussian distribution and L stands for the channel length. We suppose that

[1]A small group of adjacent subcarriers contains pilots and data symbols from one user [9].

This work was supported by Seoul R&BD Program (WR080951).

978-1-4673-1881-5/12/$31.00 ©2012 IEEE

the channel is static over two consecutive OFDMA symbols, which is reasonable in practical situations [9] [10]. The set $\mathcal{C}_k$ consists of the subcarrier indices of the k-th MU and $\mathcal{C}_k \cap \mathcal{C}_l = \phi$ for $k \neq l$. The CFO of the k-th MU is defined as ϵ_k which is normalized by the subcarrier spacing and assumed to be static over two consecutive OFDMA symbols. We 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 $\mathbf{H}_k$ is given as $\mathbf{H}_k = \sqrt{N}\text{diag}\{\mathbf{F}_L \mathbf{h}_k\} = \text{diag}\{H_{0,k}, H_{1,k}, \cdots, H_{N-1,k}\}$ where $H_{i,k}$ is the channel gain of the i-th subcarrier for the k-th MU. Moreover, the l-th transmitted signal of the k-th MU can be obtained as $\mathbf{x}_{k,l} = [X_{0,l}^k\ \cdots\ X_{N-1,l}^k]^T$ where $X_{i,l}^k = 0$ for $i \notin \mathcal{C}_k$.

Then, the received consecutive OFDMA symbols are represented in the frequency domain as

$$\mathbf{r}_l = \mathbf{F}\sum_{k=1}^{K} \boldsymbol{\Gamma}_l(\epsilon_k)\mathbf{F}^{\dagger}\mathbf{H}_k\mathbf{x}_{k,l} + \mathbf{w}_l \quad \text{for} \quad l = 1, 2 \tag{1}$$

where $\boldsymbol{\Gamma}_1(\epsilon_k) = \text{diag}\{1,\ \exp\left(\frac{j2\pi\epsilon_k}{N}\right), \cdots, \exp\left(\frac{j2\pi(N-1)\epsilon_k}{N}\right)\}$ equals a diagonal matrix whose diagonal entry stands for the phase shift of the corresponding received signal sample, we define $\boldsymbol{\Gamma}_2(\epsilon_k) = e^{j2\pi\epsilon_k N_t/N}\boldsymbol{\Gamma}_1(\epsilon_k)$ with $N_t = N + N_g$, and $\mathbf{w}_l = [w_{0,l}\ w_{1,l} \cdots\ w_{N-1,l}]^T$ indicates the complex additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $\sigma_w^2 \mathbf{I}_N$.

Let us define a diagonal matrix $\boldsymbol{\Psi}_k$ where $[\boldsymbol{\Psi}_k]_{i,i} = 1$ for $i \in \mathcal{C}_k$ and $[\boldsymbol{\Psi}_k]_{i,i} = 0$ otherwise. Then, we denote the composite transmitted data for MUs as $\bar{\mathbf{x}}_l = \sum_{k=1}^{K} \boldsymbol{\Psi}_k \mathbf{x}_{k,l} = [\bar{X}_{0,l}\ \bar{X}_{1,l}\ \cdots\ \bar{X}_{N-1,l}]^T$ where $\bar{X}_{i,l} = X_{i,l}^k$ for $i \in \mathcal{C}_k$. Similar to $\bar{\mathbf{x}}_l$, the composite channel frequency response is defined as $\bar{\mathbf{H}} = \sum_{k=1}^{K} \boldsymbol{\Psi}_k \mathbf{H}_k = \text{diag}\{\bar{H}_0,\ \bar{H}_1,\ \cdots,\ \bar{H}_{N-1}\}$ where $\bar{H}_i = H_{i,k}$ for $i \in \mathcal{C}_k$. Then, the expression in (1) can be rewritten as [11]

$$\begin{aligned} \mathbf{r}_l &= \left(\sum_{k=1}^{K} \mathbf{C}_l(\epsilon_k)\boldsymbol{\Psi}_k\right)\bar{\mathbf{H}}\bar{\mathbf{x}}_l + \mathbf{w}_l \quad \text{for} \quad l = 1, 2 \\ &= \mathbf{Q}_l\bar{\mathbf{H}}\bar{\mathbf{x}}_l + \mathbf{w}_l \end{aligned} \tag{2}$$

where $\mathbf{C}_l(\epsilon_k) = \mathbf{F}\boldsymbol{\Gamma}_l(\epsilon_k)\mathbf{F}^{\dagger}$ is a circulant matrix and $\mathbf{Q}_l = \sum_{k=1}^{K} \mathbf{C}_l(\epsilon_k)\boldsymbol{\Psi}_k$ stands for the interference matrix caused by CFOs. Note that $\mathbf{Q}_l$ becomes an identity matrix if all CFOs are zero. By applying the zero-forcing filter based on the estimated CFOs, the compensated received signal can be obtained as $\mathbf{g}_l = \mathbf{Q}_l^{-1}\mathbf{r}_l$. If the estimated CFOs are so accurate that interference of subcarriers are sufficiently suppressed, $\bar{\mathbf{x}}_l$ can be detected from $\mathbf{g}_l$ by using an one-tap equalizer.

III. Proposed Pilot-aided Estimation Algorithms

In this section, we derive a new CFO estimation algorithm which utilizes the two consecutive OFDMA symbols based on a line search method. To obtain the initial CFOs, a simple method is also provided.

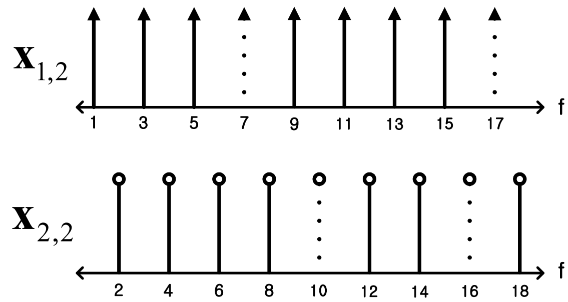

Figure 1. Transmitted OFDM symbols of each MU at the second time slot in the frequency domain

A. New representation of the second received OFDMA symbol

Here, we describe some preliminary definitions and mathematical results related to $\mathbf{r}_1$ and $\mathbf{r}_2$. Let us define $\mathbf{b}$ and $\mathbf{B}$ as $\mathbf{b} = [b_0\ \cdots\ b_{N-1}]$ with $b_m = \bar{X}_{m,2}/\bar{X}_{m,1}$ and $\mathbf{B} = \text{diag}\{\mathbf{b}\}$, respectively, where b_m equals zero if the m-th subcarrier is null. By using $\mathbf{r}_1$, $\mathbf{B}$ and the CFOs, the second received OFDMA symbol $\mathbf{r}_2$ can be rewritten as

$$\begin{aligned} \mathbf{r}_2 &= \mathbf{Q}_2\bar{\mathbf{H}}\bar{\mathbf{x}}_2 + \mathbf{w}_2 = \mathbf{Q}_2\bar{\mathbf{H}}\mathbf{B}\bar{\mathbf{x}}_1 + \mathbf{w}_2 \\ &= \mathbf{Q}_2\mathbf{B}\mathbf{Q}_1^{-1}\mathbf{Q}_1\bar{\mathbf{H}}\bar{\mathbf{x}}_1 + \mathbf{w}_2 = \mathbf{Q}_2\mathbf{B}\mathbf{Q}_1^{-1}(\mathbf{r}_1 - \mathbf{w}_1) + \mathbf{w}_2 \\ &= \mathbf{Q}_2\mathbf{B}\mathbf{Q}_1^{-1}\mathbf{r}_1 + \mathbf{n} \end{aligned} \tag{3}$$

where $\mathbf{n}$ is expressed as $\mathbf{n} = \mathbf{w}_2 - \mathbf{Q}_2\mathbf{B}\mathbf{Q}_1^{-1}\mathbf{w}_1$.

In (3), an unknown parameter set is given by $\{\epsilon_1, \cdots, \epsilon_K,\quad b_0, \cdots, b_{N-1}\}$. Since the number of the unknown elements is $N + K$ and the size of equation (3) is N, at least K elements in $\mathbf{b}$ should be known to the BS in advance. For this reason, we assume that $\bar{X}_{m,2}$ is equal to $\bar{X}_{m,1}$ (i.e. $b_m = 1$) for certain subcarriers where the number of those is equal to or greater than K. Based on this knowledge, we will jointly estimate the CFOs and the remaining unknown part of $\mathbf{b}$. With some abuse of terminology, the known and unknown elements in $\mathbf{b}$ are referred to as pilots and differential data, respectively. Then, we denote $\mathbf{p} = [1\ \cdots\ 1]^T$ and $\mathcal{C}_p$ as the vector which consists of N_p pilots of all MUs with $N_p \geq K$ and the set which contains the indices of pilots, respectively. Similarly, we define $\mathbf{d} = [d_1\ \cdots\ d_{N_d}]^T$ and $\mathcal{C}_d$ as the vector which comprises N_d unknown differential data of all MUs with $N_d = N - N_p$ and the corresponding set, respectively.

Fig. 1 illustrates[2] an example of $\mathbf{x}_{1,2}$ and $\mathbf{x}_{2,2}$ in the frequency domain where $\mathcal{C}_1$ and $\mathcal{C}_2$ are defined as $\{1, 3, \cdots, 17\}$ and $\{2, 4, \cdots, 18\}$, respectively. Moreover, $\mathcal{C}_p$ and $\mathcal{C}_d$ are defined as $\{7, 10, 16, 17\}$ and $\{\mathcal{C}_1 \cup \mathcal{C}_2\} \setminus \mathcal{C}_p$, respectively. In this case, $\mathbf{b}$ is given by $[d_1\ \cdots\ d_6\ 1\ d_7\ d_8\ 1\ d_9\ \cdots\ d_{13}\ 1\ 1\ d_{14}]^T$. Then, $\bar{X}_{m,2}$, denoted by a dotted line, is equal to $\bar{X}_{m,1}$ for $m \in \mathcal{C}_p$, which leads to a reduction of the data rate as $\frac{2N - N_p}{2N}$ in two OFDMA symbols[3].

[2]For simple illustration, the ICAS is employed in Fig. 1. Note that our proposed algorithm works regardless of the employed CAS.

[3]For our proposed scheme, it is not necessary to know the transmitted data $\bar{X}_{m,1}$ for $m \in \mathcal{C}_p$.

By using $\mathcal{C}_p$ and $\mathcal{C}_d$, we represent a diagonal matrix $\tilde{\mathbf{\Phi}}_p$ where $[\tilde{\mathbf{\Phi}}_p]_{i,i} = 1$ for $i \in \mathcal{C}_p$ and $[\tilde{\mathbf{\Phi}}_p]_{i,i} = 0$ otherwise. Extracting the columns whose indices are contained in $\mathcal{C}_p$ from $\tilde{\mathbf{\Phi}}_p$ yields $\mathbf{\Phi}_p$ of size $N \times N_p$. Similarly, $\mathbf{\Phi}_d$ of size $N \times N_d$ can be defined with $\mathcal{C}_d$. Then, we check that $\mathbf{\Phi}_p^T \mathbf{\Phi}_p = \mathbf{I}_{N_p}$, $\mathbf{\Phi}_d^T \mathbf{\Phi}_d = \mathbf{I}_{N_d}$ and $\mathbf{\Phi}_p^T \mathbf{\Phi}_d = \mathbf{0}_{N_p \times N_d}$. Thereby, the vector $\mathbf{b}$ is given by

$$\mathbf{b} = \mathbf{\Phi}_p \mathbf{p} + \mathbf{\Phi}_d \mathbf{d} \tag{4}$$

where $\mathbf{p}$ and $\mathbf{d}$ can be obtained as $\mathbf{\Phi}_p^T \mathbf{b}$ and $\mathbf{\Phi}_d^T \mathbf{b}$, respectively.

By utilizing (4), the received signal $\mathbf{r}_2$ in (3) can be rewritten as

$$\begin{aligned} \mathbf{r}_2 &= \mathbf{Q}_2 \tilde{\mathbf{R}}_1 \mathbf{b} + \mathbf{n} \\ &= \mathbf{Q}_2 \tilde{\mathbf{R}}_1 \mathbf{\Phi}_p \mathbf{p} + \mathbf{Q}_2 \tilde{\mathbf{R}}_1 \mathbf{\Phi}_d \mathbf{d} + \mathbf{n} \end{aligned} \tag{5}$$

where $\tilde{\mathbf{R}}_1 = \text{diag}\{\mathbf{Q}_1^{-1}\mathbf{r}_1\}$. Unfortunately, its rigorous application to our estimation algorithm turns out to be difficult, since $\mathbf{n}$ has a correlated covariance matrix $\mathbf{C}_n = \sigma_w^2 \big(\mathbf{I}_N + \mathbb{E}[\mathbf{Q}_2 \mathbf{B} \mathbf{Q}_1^{-1} (\mathbf{Q}_1^{-1})^\dagger \mathbf{B}^\dagger \mathbf{Q}_2^\dagger]\ \big)$. To facilitate the analysis, we assume that all MUs' CFOs are adequately small so as to justify an uncorrelated approximation as[4] [10]

$$\mathbf{C}_n \simeq 2\sigma_w^2 \mathbf{I}_N. \tag{6}$$

B. Proposed estimation algorithm

Now, we develop an one-shot algorithm to jointly estimate the CFOs and the unknown part of $\mathbf{b}$ from the signal model in (5) based on a line search method. Let us define the unknown parameter vector and the initial estimated corresponding vector as $\mathbf{u} = [\boldsymbol{\xi}^T\ \mathbf{d}^T]^T$ with $\boldsymbol{\xi} = [\epsilon_1\ \cdots\ \epsilon_K]^T$ and $\hat{\mathbf{u}}^{(0)} = [(\hat{\boldsymbol{\xi}}^{(0)})^T\ (\hat{\mathbf{d}}^{(0)})^T]^T$ with $\hat{\boldsymbol{\xi}}^{(0)} = [\hat{\epsilon}_1^{(0)}\ \cdots\ \hat{\epsilon}_K^{(0)}]^T$ and $\hat{\mathbf{d}}^{(0)} = [\hat{d}_1^{(0)}\ \cdots\ \hat{d}_{N_d}^{(0)}]^T$, respectively. From our new expression for $\mathbf{r}_2$ in (3), we can regenerate $\mathbf{r}_2^{(0)}$ by using $\mathbf{r}_1$ and $\hat{\mathbf{u}}^{(0)}$ as

$$\mathbf{r}_2^{(0)} = \mathbf{Q}_2^{(0)} \mathbf{B}^{(0)} (\mathbf{Q}_1^{(0)})^{-1} \mathbf{r}_1 \tag{7}$$

where $\mathbf{B}^{(0)}$, $\mathbf{Q}_1^{(0)}$ and $\mathbf{Q}_2^{(0)}$ are obtained from the initial estimated vector $\hat{\mathbf{u}}^{(0)}$.

Note that $\mathbf{r}_2$ in (5) is a nonlinear function of $\mathbf{u}$ due to the CFOs. By applying the first-order Taylor series expansion, we can approximate $\mathbf{r}_2$ as a linear function of $\mathbf{u}$

$$\mathbf{r}_2 \approx \mathbf{r}_2^{(0)} + \mathbf{G}(\mathbf{u} - \hat{\mathbf{u}}^{(0)}) + \mathbf{n} \tag{8}$$

where the gradient matrix $\mathbf{G}$ of size $N \times (N_d + K)$ is defined as

$$\begin{aligned} \mathbf{G} &\triangleq \left[\frac{\partial \mathbf{r}_2}{\partial \epsilon_1}\ \cdots\ \frac{\partial \mathbf{r}_2}{\partial \epsilon_K}\ \frac{\partial \mathbf{r}_2}{\partial \mathbf{d}} \right]_{\mathbf{u}=\hat{\mathbf{u}}^{(0)}} \\ &= \left[\mathbf{t}(\hat{\epsilon}_1^{(0)})\ \cdots\ \mathbf{t}(\hat{\epsilon}_K^{(0)})\ \mathbf{G}_2 \right]. \end{aligned}$$

[4]If all CFOs are zero and the constant-magnitude constellation such as 4-QAM is employed, we have $\mathbf{C}_n = 2\sigma_w^2 \mathbf{I}_N$.

Here, $\mathbf{t}(\hat{\epsilon}_k^{(0)})$ and $\mathbf{G}_2$ are derived as, respectively,

$$\begin{aligned} \mathbf{t}(\hat{\epsilon}_k^{(0)}) &\triangleq \left. \frac{\partial \mathbf{r}_2}{\partial \epsilon_k} \right|_{\mathbf{u}=\mathbf{u}^{(0)}} \\ &= \frac{j2\pi}{N} \Big\{ \mathbf{F}(\mathbf{M} + N_t \mathbf{I}) \mathbf{\Gamma}_2(\hat{\epsilon}_k^{(0)}) \mathbf{F}^\dagger \mathbf{\Psi}_k \mathbf{B}^{(0)} \mathbf{g}_1^{(0)} \\ &\quad - \mathbf{Q}_2^{(0)} \mathbf{B}^{(0)} (\mathbf{Q}_1^{(0)})^{-1} \mathbf{F} \mathbf{M} \mathbf{\Gamma}_1(\hat{\epsilon}_k^{(0)}) \mathbf{F}^\dagger \mathbf{\Psi}_k \mathbf{g}_1^{(0)} \Big\} \\ \mathbf{G}_2 &\triangleq \left. \frac{\partial \mathbf{r}_2}{\partial \mathbf{d}} \right|_{\mathbf{u}=\mathbf{u}^{(0)}} = \mathbf{Q}_2^{(0)} \tilde{\mathbf{R}}_1^{(0)} \mathbf{\Phi}_d \end{aligned}$$

where we have $\mathbf{M} = \text{diag}\{0, \cdots, N-1\}$, $\mathbf{g}_1^{(0)} = (\mathbf{Q}_1^{(0)})^{-1} \mathbf{r}_1$ and $\tilde{\mathbf{R}}_1^{(0)} = \text{diag}\{(\mathbf{Q}_1^{(0)})^{-1} \mathbf{r}_1\}$. Note that the accuracy of the approximation in (8) depends on the initial estimated vector $\hat{\mathbf{u}}^{(0)}$.

Now, we define a trial vector of $\mathbf{u}$ as $\tilde{\mathbf{u}}^{(1)} = [(\tilde{\boldsymbol{\xi}}^{(1)})^T\ (\tilde{\mathbf{d}}^{(1)})^T]^T$. Based on (8) and $\tilde{\mathbf{u}}^{(1)}$, the regenerated signal $\tilde{\mathbf{r}}_2^{(1)}$ is obtained as

$$\tilde{\mathbf{r}}_2^{(1)} = \mathbf{r}_2^{(0)} + \mathbf{G}(\tilde{\mathbf{u}}^{(1)} - \hat{\mathbf{u}}^{(0)}). \tag{9}$$

By minimizing the mean square distance between $\mathbf{r}_2$ in (8) and $\tilde{\mathbf{r}}_2^{(1)}$ in (9), we find the new estimated vector $\hat{\mathbf{u}}^{(1)}$ which is formulated as

$$\begin{aligned} \hat{\mathbf{u}}^{(1)} &= \arg\min_{\tilde{\mathbf{u}}^{(1)}} \left\{ \mathbb{E} \| \mathbf{r}_2 - \tilde{\mathbf{r}}_2^{(1)} \|^2 \right\} \\ &\approx \arg\min_{\tilde{\mathbf{u}}^{(1)}} \left\{ \mathbb{E} \| \mathbf{G}(\mathbf{u} - \tilde{\mathbf{u}}^{(1)}) + \mathbf{n} \|^2 \right\}. \end{aligned} \tag{10}$$

Then, the problem (10) can be solved by a line search method as [4] [5]

$$\hat{\mathbf{u}}^{(1)} = \hat{\mathbf{u}}^{(0)} + \big(\mathbf{G}^\dagger \mathbf{G} \big)^{-1} \mathbf{G}^\dagger (\mathbf{r}_2 - \mathbf{r}_2^{(0)}). \tag{11}$$

Since the performance of the solution in (11) depends on $\hat{\mathbf{u}}^{(0)}$, we will address an initial estimation method for $\hat{\mathbf{u}}^{(0)}$ in the following subsection.

Now, we derive the CRB of the CFO for the joint estimation of CFOs and the unknown differential data $\mathbf{d}$. Based on (3) and (6), the likelihood function of $\mathbf{r}_2$ conditioned on $\mathbf{u}$ and $\mathbf{r}_1$ is written by

$$f(\mathbf{r}_2 | \mathbf{u}, \mathbf{r}_1) = \frac{1}{(2\pi\sigma_w^2)^N} \exp\Big(-\frac{1}{2\sigma_w^2} \| \mathbf{r}_2 - \mathbf{Q}_2 \mathbf{B} \mathbf{Q}_1^{-1} \mathbf{r}_1 \|^2 \Big). \tag{12}$$

By utilizing (12) and the results in [12], the CRB for the CFO of the k-th MU is obtained as

$$\text{CRB}(\epsilon_k) = \sigma_w^2 \left[\Big(\Re\big\{ \mathbf{Z}^\dagger \big(\mathbf{I} - \mathbf{P}(\mathbf{P}^\dagger \mathbf{P})^{-1} \mathbf{P}^\dagger \big) \mathbf{Z} \big\} \Big)^{-1} \right]_{k,k} \tag{13}$$

where $\mathbf{Z} = [\mathbf{t}(\epsilon_1)\ \cdots\ \mathbf{t}(\epsilon_K)]$ and $\mathbf{P} = \mathbf{Q}_2 \tilde{\mathbf{R}}_1 \mathbf{\Phi}_d$.

C. Initialization of unknown parameters

To ensure the accuracy of the algorithm in (11), we describe a method to estimate the initial CFOs and the unknown differential data $\hat{\mathbf{u}}^{(0)} = [(\hat{\boldsymbol{\xi}}^{(0)})^T\ (\hat{\mathbf{d}}^{(0)})^T]^T$ with pilots. We assume that $\mathcal{C}_k \cap \mathcal{C}_p$ is not an empty set for $k = 1, \cdots, K$, which indicates that each MU has at least one pilot. To expand

(2), let us define the m-th element of $\mathbf{r}_l$ as $R_{m,l}$. Then, we can obtain the initial estimate of ϵ_k as

$$\hat{\epsilon}_k^{(0)} = \frac{N}{2\pi N_t} \arg\Big\{ \sum_{m\in\mathcal{C}_k\cap\mathcal{C}_p} R_{m,1}^* R_{m,2} \Big\} \tag{14}$$
$$\approx \frac{N}{2\pi N_t} \arg\Big\{ e^{j2\pi\epsilon_k N_t/N} \sum_{m\in\mathcal{C}_k\cap\mathcal{C}_p} |\Upsilon(\epsilon_k)|^2 |\bar{H}_m \bar{X}_{m,2}|^2 \Big\}$$

where $\Upsilon(\epsilon) \triangleq \frac{1}{N} e^{j\pi\epsilon\frac{N-1}{N}} \frac{\sin(\pi\epsilon)}{\sin(\pi\epsilon/N)}$, and the approximation comes from ignoring the noise, the ICI and the MAI.

Let us define $\mathbf{g}_2^{(0)}$ as $\mathbf{g}_2^{(0)} = \big(\mathbf{Q}_2^{(0)}\big)^{-1}\mathbf{r}_2$. Here, $\mathbf{g}_1^{(0)}$ and $\mathbf{g}_2^{(0)}$ are expected to have lower ICI and MAI than $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively. Therefore, more accurate initial estimates of ϵ_k can be obtained by repeating the step in (14) based on $\mathbf{g}_1^{(0)}$ and $\mathbf{g}_2^{(0)}$. Then, $\mathbf{g}_1^{(0)}$ and $\mathbf{g}_2^{(0)}$ are recalculated by using the second initial CFO estimates.

On the other hand, the initial estimate $\hat{\mathbf{d}}^{(0)}$ is computed as

$$\hat{\mathbf{d}}^{(0)} = (\mathrm{diag}\{\mathbf{\Phi}_d^T \mathbf{g}_1^{(0)}\})^{-1} \mathbf{\Phi}_d^T \mathbf{g}_2^{(0)}. \tag{15}$$

Then, we obtain $\hat{\mathbf{u}}^{(0)} = [(\hat{\boldsymbol{\xi}}^{(0)})^T \ (\hat{\mathbf{d}}^{(0)})^T]^T$.

IV. Simulation Results

In this section, we compare the computational complexity and the mean square error (MSE) performance of our proposed algorithm with those of other conventional algorithms in OFDMA uplink systems with 128 subcarriers and $N_g = 16$. We assume a 5-tap Rayleigh fading exponential decaying channel and perform $1,000$ simulation runs. In addition, $\{\epsilon_1, \cdots, \epsilon_K\}$ are independently chosen from a uniform distribution within a range $[-0.3, 0.3]$. For simple simulations, each MU is randomly assigned to $P = N/K$ subcarriers. Also, the number of pilots and unknown differential data of each MU is set to N_p/K and N_d/K $(N = N_p + N_d)$, respectively. In addition, we employ differential quadrature phase shift keying (DQPSK).

For simple notation, we call the proposed line search based algorithm in (11) as LS. Also, the grid search algorithm [9] and the ad hoc CFO estimator [10] are referred to as GS and AHE, respectively. For the GS, the number of grid points G is chosen to be 601 over the range of $[-0.3, 0.3]$ to achieve the precision of 10^{-3}. Also, the iteration number of the GS and the AHE is set to 3.

The initial estimates of the proposed algorithms are obtained from the initial stages in (14) and (15). In this case, the computational complexity of the LS is as follows. For obtaining $\hat{\mathbf{d}}^{(0)}$, two inverse operations of $\mathbf{Q}_l^{(0)}$ are required in (15). Also, we need the computation of $(\mathbf{G}^\dagger\mathbf{G})^{-1}$ in (11). Note that the computation of $(\mathbf{Q}_l^{(0)})^{-1}$ in (7) is not necessary because the computation is already performed at the initial step in (15). Then, the complexity of the initial estimator in (15) and the line search based estimator in (11) are given by $O(4N^3)$ and $O((N_d + K)^3)$, respectively. For the AHE, its computational complexity is about $O(6N^3)$. In contrast, the overall complexity of the GS is $O(2N_i(GNN_p + N^3))$ where N_i denotes the iteration number [10]. Note that the presented computational complexity includes the process required for data detection. For example, when $N = 128$, $K = 4$, $N_i = 3$ and $N_p = 24$, the computational complexity of the LS is about 21% of the GS[5]. Meanwhile, the training sequence based algorithms in [4]–[6] have much lower computational complexity. This is due to the fact that those algorithms do not need computationally intensive CFO compensation as $\mathbf{Q}_l^{-1}\mathbf{r}_l$ at the expense of the spectral efficiency. Nevertheless, the complexity issue of the pilot-aided algorithms can be alleviated by a low-complexity implementation of $\mathbf{Q}_l^{-1}\mathbf{r}_l$, as presented in [13].

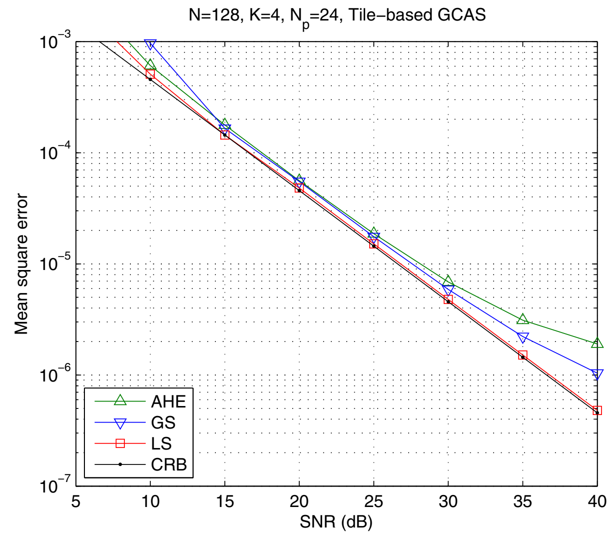

Figure 2. MSE performance of CFO estimators with different algorithms

Fig. 2 compares the MSE performance of the proposed algorithm with those of the conventional algorithms. Here, we employ the tile-based GCAS with the tile size of $Z = 4$. The average CRB is obtained by averaging the CRB in (13) over multiple simulation runs with different random CFOs and channel gains. It can be seen that all algorithms yield higher MSE at low SNR. This can be explained by the fact that the algorithms produce large CFO errors at low SNR. In the SNR higher than 10 dB, the LS approaches the average CRB. Compared to the LS, the GS and the AHE produce higher performance loss in moderate and high SNR regions, and the performance gap are about 2.5 dB and 5 dB at the SNR of 40 dB, respectively.

Fig. 3 illustrates the MSE performance of our proposed algorithm according to the number of pilots. Contrary to the previous simulation, the GCAS with $Z = 1$ and $K = 2$ is employed. We see that the LS with $N_p = 24$ achieves a 3 dB performance gain over that with $N_p = 12$. It is clear that our proposed algorithm can effectively obtain a good trade-off between the MSE performance and the spectral efficiency for the GCAS.

[5]Gaussian elimination is assumed to assess the computational complexity of an inverse operation with $O(N^3) \approx \frac{1}{3}N^3$.

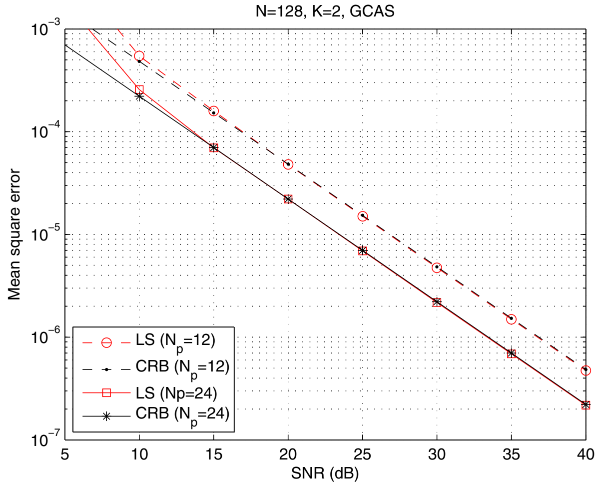

Figure 3. MSE performance of proposed algorithm with different numbers of pilots

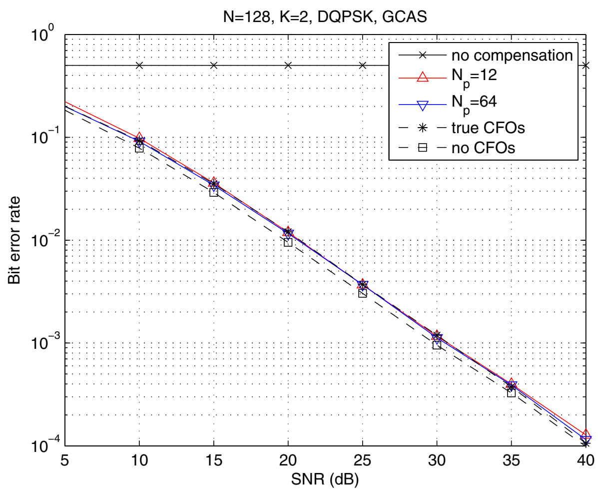

Figure 4. BER performance of proposed algorithm with different numbers of pilots

Fig. 4 shows the BER performance according to N_p. In the case of true CFOs, the BS determines the unknown differential data $\mathbf{d}$ from $\mathbf{g}_l = \mathbf{Q}_l^{-1}\mathbf{r}_l$ $(l = 1, 2)$ where $\mathbf{Q}_l^{-1}$ is computed by the true CFOs. Since the LS jointly estimates $\mathbf{d}^{(1)}$ and $\epsilon_k^{(1)}(k = 1, \cdots, K)$, we do not need to obtain $\mathbf{d}^{(1)}$ from $\mathbf{g}_l^{(1)}$. Thus, for DQPSK, the computational complexity becomes reduced by avoiding two inverse operations[6] of $\mathbf{Q}_l^{(1)}$. The BER performance of the LS has the similar performance of the case of true CFOs, while an 1 dB loss is observed compared to the case of no CFOs. Note that the BER performance with $N_p = 12$ is similar to that with $N_p = 64$, which indicates that $N_p = 12$ is sufficient for DQPSK systems in the presence of CFOs. As a result, our proposed scheme can provide higher spectral efficiency with a negligible BER performance loss.

[6]For the algorithms in [9] and [10], two additional inverse operations for $(\mathbf{Q}_l^{(1)})^{-1}$ are still required.

V. Conclusions

In this paper, we have proposed a pilot-aided joint estimation method for CFOs and the unknown differential data for OFDMA uplink systems. Based on the first received OFDMA symbol, we represent the second OFDMA symbol as the signal which includes the CFOs, the pilots and the unknown differential data. Then, by approximating this nonlinear model with a linear function, we have proposed a new joint estimation algorithm motivated by the line search method. We confirm from simulation results that the joint estimation algorithm aided by the initial estimation method can achieve the average CRB for moderate and high SNR regions. As a result, our proposed algorithms effectively provide a good trade-off between the MSE performance and the spectral efficiency for general carrier assignment schemes.

References

[1] "IEEE standard for local and metropolitan area networks, Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile operation in Licensed Bands," *IEEE Std. 802.16e*, February 2006.

[2] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multicarrier OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 479–483, October 1999.

[3] Z. Cao, U. Tureli, and Y.-D. Yao, "Deterministic Multiuser CFO Estimation for Interleaved OFDMA Uplink," *IEEE Transactions on Communications*, vol. 52, pp. 1585–1594, September 2004.

[4] 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.

[5] 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.

[6] K. Lee, S.-H. Moon, and I. Lee, "Low-Complexity Leakage-Based Carrier Frequency Offset Estimation Techniques for OFDMA Uplink Systems," in *Proc. IEEE Globecom '10*, December 2010.

[7] ——, "Robust Pilot Designs for Consistent Frequency Offset Estimation in OFDM systems," in *Proc. IEEE Globecom '10*, December 2010.

[8] Y. Bhargava and K. Giridhar, "Efficient Synchronization and Frequency tracking for cellular reuse-1 OFDMA systems," in *Proc. IEEE WCNC '07*, August 2007, pp. 2409–2413.

[9] P. Sun and L. Zhang, "Low Complexity Pilot Aided Frequency Synchronization for OFDMA Uplink Transmission," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 3758–3769, July 2009.

[10] P. Sun, M. Michele, and L. Zhang, "Carrier Frequency Offset Tracking in the IEEE 802.16e OFDMA Uplink," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3613–3619, December 2010.

[11] Z. Cao, U. Tureli, and Y.-D. Yao, "Frequency synchronization for generalized OFDMA uplink," in *Proc. IEEE Globecom*, vol. 2, November 2004, pp. 1071–1075.

[12] 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.

[13] K. Lee and I. Lee, "CFO Compensation for Uplink OFDMA Systems with Conjugated Gradient," in *Proc. IEEE ICC '11*, June 2011.