Two Novel Iterative Joint Frequency-Offset and Channel Estimation Methods for OFDMA Uplink

Xiaoyu Fu, *Student Member, IEEE*, Hlaing Minn, *Senior Member, IEEE*, and Cyrus D. Cantrell, *Fellow, IEEE*

***Abstract*—We address joint estimation of frequency offsets and channel responses in OFDMA uplink. A cyclically equal-spaced, equal-energy interleaved pilot preamble is proposed by which two iterative estimators are developed. In the first estimator, we develop a modified SAGE (space alternating generalized expectation-maximization) method by incorporating multiuser interference cancellation in both time and frequency domains into the SAGE method. The proposed modified SAGE method yields a faster convergence rate, a better estimation performance, and a lower complexity than the existing conventional SAGE method from [10]. In the second method, we propose a very low complexity ad hoc method to replace the high complexity frequency offset estimator of the first method. The proposed ad hoc method is developed based on the time-domain characteristics of the pilot preamble and the frequency-domain power concentration property of OFDMA systems. It achieves almost the same estimation performance as the first proposed method but requires significantly lower complexity. Both proposed methods are not only better in estimation performance, convergence rate, and complexity, but also more robust to the number of active uplink users than the existing method while requiring only one OFDM training symbol and no restriction of the data subcarrier assignment scheme.**

***Index Terms*—OFDMA, SAGE, joint frequency offset and channel estimation, MUI.**

I. Introduction

ORTHOGONAL Frequency Division Multiple Access (OFDMA) is a promising technique for future high speed broadband wireless communication systems and it has recently been proposed or adopted in many industry standards (e.g., IEEE 802.16-2004[1], 802.16e, 802.20, [2]). OFDMA has several advantages such as multi-user diversity, scalibity/adaptability in modulation/data rate/spectral occupancy, modularity, and low complexity equalization. On the other hand, OFDMA uplink is associated with a more difficult task of synchronization and channel estimation due to the different time and frequency offsets of uplink users. Although timing offsets can be limited within the interference-free interval of the cyclic prefix by means of downlink synchronization or initial ranging process [3], (residual) frequency offsets result in inter-carrier interference (ICI), and could degrade the channel estimation and BER performance significantly [4]-[5]. Hence, reliable estimation of multiuser (residual) frequency offsets and channels is critical in realizing the OFDMA advantages.

Several methods have been recently proposed for the frequency offset estimation and/or channel estimation in OFDMA uplink [6]-[10]. The method in [6] considered a single unsynchronized user scenario which assumed that only the frequency offset of one new subscriber station (SS) was needed to be estimated while other SSs had already been perfectly synchronized to the base station (BS) references. Hence, this approach would not fully represent a practical situation. Moreover, it requires at least two OFDM training symbols. The method in [7] considered a grouped CAS (carrier assignment scheme) where each SS uses a group of adjacent sub-carriers. This sub-band type assignment has a good multiuser interference (MUI) suppression but it will lose frequency diversity if channel-aware CAS is not implemented. In [8], a multiple signal classification type blind frequency offset estimator based on the interleaved CAS was proposed. Due to the grid-search involved, its computation complexity is quite high and its estimation accuracy depends on the computation complexity and the number of SSs in the system. Recently two iterative methods, alternating projection (AP) and space alternating generalized expectation-maximization (SAGE), are presented in [9] and [10], respectively. Both of them involve iterative MUI cancellation in time-domain. However, they suffer from very high computation complexity and performance degradation due to MUI (a larger degradation for a larger number of SSs).

The existing methods are associated with one or more of the following unfavorable requirements: very high complexity, several OFDM training symbols, specific CAS, ideal system situation, and a small number of SSs. In this paper, aiming at addressing the above drawbacks, we propose two iterative joint estimators utilizing a cyclically equal-spaced, equal-energy interleaved pilot preamble which is based on the best channel identification conditions [11]. In the first estimator, we develop a modified SAGE (MSAGE) method by incorporating multiuser interference cancellation in both time and frequency domains into the SAGE method. In the second method, we propose a very low complexity ad hoc method to replace the high complexity frequency offset estimator of the first method. The proposed ad hoc method is developed based on the time-domain characteristics of the pilot preamble and the frequency-domain power concentration property of OFDMA systems. It achieves almost the same estimation performance

Paper approved by A. Anastasopoulos, the Editor for Iterative Detection, Estimation and Coding of the IEEE Communications Society. Manuscript received March 31, 2006; revised August 22, 2006 and October 30, 2006. This work was supported by the Erik Jonsson School Research Excellence Initiative, the University of Texas at Dallas, USA. This paper was presented in parts at IEEE Globecom 2006.

The authors are with the Department of Electrical Engineering, University of Texas at Dallas (e-mail: {xiaoyu.fu, hlaing.minn, cantrell}@utdallas.edu).

Digital Object Identifier 10.1109/TCOMM.2008.060206.

as the first proposed method but requires significantly lower complexity. Both proposed methods are not only better in estimation performance, convergence rate, and complexity, but also more robust to the number of active uplink users than the existing methods while requiring only one OFDM training symbol and no restriction of the data CAS scheme.

The rest of the paper is organized as follows. In the section II, we introduce the system description, the preamble design, and the signal model. Our proposed MSAGE is presented in section III. The proposed ad hoc method (AH) for complexity reduction is derived in section IV. The simulation results and discussions are presented in section V, and the paper is concluded in Section VI.

Notations: The superscripts T, $*$ and H represent the transpose, the conjugate and the Hermitian transpose, respectively. $\lfloor\cdot\rfloor$ is the floor operation. $\mathrm{diag}\{\boldsymbol{x}\}$ is a diagonal matrix with the vector $\boldsymbol{x}$ on its main diagonal. $E[\cdot]$ and $\Re(\cdot)$ represent the expectation and the real part of the enclosed parameters, respectively. $\boldsymbol{I}_N$ is the $N \times N$ identity matrix and $\boldsymbol{0}_m$ is the all-zero $m \times 1$ vector. $\mathrm{mod}(i, K)$ denotes i modulo K and $||\cdot||$ represents the Euclidean norm of the enclosed vector.

II. System Description, Preamble Design, and Signal Model

We consider uplink of an OFDMA system with N subcarriers. Each uplink packet contains one OFDM training symbol (preamble) for synchronization and channel estimation followed by several data symbols. Each OFDM symbol has $N_t = N + N_g$ samples where N is the length of useful part and N_g is the length of cyclic prefix (CP) part. For data transmission, the N sub-carriers are grouped into Q subchannels. Each subchannel has P subcarriers where $QP \leq N$. There are K SSs communicating with the BS simultaneously where $K \leq Q$. The k-th SS ($k \in \{1, \ldots, K\}$) has its own frequency offset v_k (normalized by the subcarrier spacing), channel impulse response (CIR) $\bar{\boldsymbol{h}}_k$ and timing offset d_k. We assume that each SS has already performed at least integer normalized frequency offset estimation and correction from the downlink, and we model v_k as a uniformly distributed random variable within the range $[-0.5, 0.5]$. The timing offset $d_k \in \{0, \ldots, d_{\max} - 1\}$ is due to the timing estimation error in initial ranging process where $d_{\max}$ is the maximum possible timing estimation error determined by the initial ranging process [1]. $\bar{\boldsymbol{h}}_k = [\bar{h}_k(0), \bar{h}_k(1), \ldots, \bar{h}_k(\bar{L} - 1)]^T$ is the sample-spaced CIR vector where $\bar{L}$ representing the maximum number of the sample-spaced channel taps is the same for all SSs. The CIR vectors are assumed unchanged during a packet interval. To avoid inter-symbol interference (ISI), N_g is chosen to be greater than or equal to $L' = \bar{L} + d_{\max}$.

In our system, the CASs for the preamble and for the data symbols are not necessary to be the same. We use an interleaved CAS for the preamble where each SS is assigned a disjoint set of N_p cyclically equal-spaced, equal-energy pilots. The pilot symbol on the m-th subcarrier of the k-th SS is given by

$$x_k(m) = \begin{cases} a_k c_k(u), & m = uD + \Delta_k; \quad (u = 0, \ldots, N_p - 1) \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where a_k is the amplitude factor and $\Delta_k = \lfloor \frac{(k-1)D}{K} \rfloor$ is the subcarrier index shift for the k-th SS. Note that our pilot design sets the inter-user adjacent pilot tone spacings as large as possible. For simplicity, $\{a_k\}_{k=1}^K$ are assumed to be the same for all SSs in the rest of the paper. $\boldsymbol{c}_k = [c_k(0), \ldots, c_k(N_p - 1)]^T$ is the non-zero pilot vector of the k-th SS where $|c_k(u)| = 1$ for $u = 0, \ldots, N_p - 1$ and $k = 1, \ldots, K$; $D = \frac{N}{N_p}$ is an integer representing the pilot spacing within each SS, and $N_p \geq L'$ [1]. The above preamble design satisfies the best channel identification condition [11] and gives reliable channel gain estimates of *all* subcarriers. This channel information is useful not only for the data detection in the current frame but also for the optimum resource allocation in the next uplink frame's data transmission. Our considered system does not have any restriction on the data CAS which could be updated through UL-map information message by BS before each uplink frame transmission.

The timing offsets can be absorbed into the CIR as

$$\boldsymbol{h}_k = [\boldsymbol{0}_{d_k}^T, \bar{\boldsymbol{h}}_k^T, \boldsymbol{0}_{Ng-\bar{L}-d_k}^T]^T \tag{2}$$

where for a robust design we have set the length of the above CIR to $L = N_g$. If we have the exact knowledge of L', we can simply set $L = L'$. After the CP removal, the time-domain low-pass-equivalent channel output preamble vector $\boldsymbol{s}_k$ for the k-th SS is given by

$$\boldsymbol{s}_k = \boldsymbol{F}^H \bar{\boldsymbol{X}}_k (\sqrt{N} \boldsymbol{F}_L \boldsymbol{h}_k) \triangleq \boldsymbol{F}^H \bar{\boldsymbol{X}}_k \boldsymbol{H}_k \triangleq \boldsymbol{A}_k \boldsymbol{h}_k \tag{3}$$

where $\boldsymbol{H}_k = [H_k(0), H_k(1), \ldots, H_k(N-1)]^T$ is the channel frequency response vector for the k-th SS, $\boldsymbol{F}$ is the $N \times N$ unitary discrete Fourier transform matrix, $\boldsymbol{F}_L$ is the $N \times L$ matrix containing the first L columns of $\boldsymbol{F}$, $\bar{\boldsymbol{X}}_k = \mathrm{diag}\{\boldsymbol{x}_k\}$, $\boldsymbol{x}_k = [x_k(0), \ldots, x_k(N-1)]^T$, and $\boldsymbol{A}_k = \boldsymbol{F}^H \bar{\boldsymbol{X}}_k \sqrt{N} \boldsymbol{F}_L$. The n-th sample of $\boldsymbol{s}_k$ can be expressed as

$$s_k(n) = \frac{1}{\sqrt{N}} \sum_{u=0}^{N_P - 1} [X_k(uD + \Delta_k) H_k(uD + \Delta_k)] e^{\frac{j2\pi(uD+\Delta_k)n}{N}}. \tag{4}$$

The corresponding time domain received preamble vector is given by

$$\boldsymbol{r} = \sum_{k=1}^{K} \boldsymbol{\Gamma}(v_k) \boldsymbol{s}_k + \boldsymbol{z} \tag{5}$$

where

$$\begin{aligned} \boldsymbol{r} &= [r(0), r(1), \ldots, r(N-1)]^T \\ \boldsymbol{\Gamma}(v_k) &= \mathrm{diag}\{[1, e^{\frac{j2\pi v_k}{N}}, \ldots, e^{\frac{j2\pi(N-1)v_k}{N}}]\}, \end{aligned} \tag{6}$$

and $\boldsymbol{z}$ is the circularly-symmetric complex Gaussian noise vector with zero mean and covariance matrix $\sigma_z^2 \boldsymbol{I}_N$. The signal to noise ratio (SNR) for the k-th SS is defined as

$$\mathrm{SNR}_k = \frac{E[\boldsymbol{s}_k^H \boldsymbol{s}_k]}{N \sigma_z^2}. \tag{7}$$

For channels with independent $\{h_k(n)\}$ for all k and n, we have

$$\mathrm{SNR}_k = \frac{N_p a_k^2 \sigma_h^2}{N \sigma_z^2} \tag{8}$$

[1]Since N is usually an integer multiple of N_g (IEEE 802.16-2004, 802.16e and [2]) and $N_g \geq L'$, we simply use $N_p = N_g$ in the rest of paper.

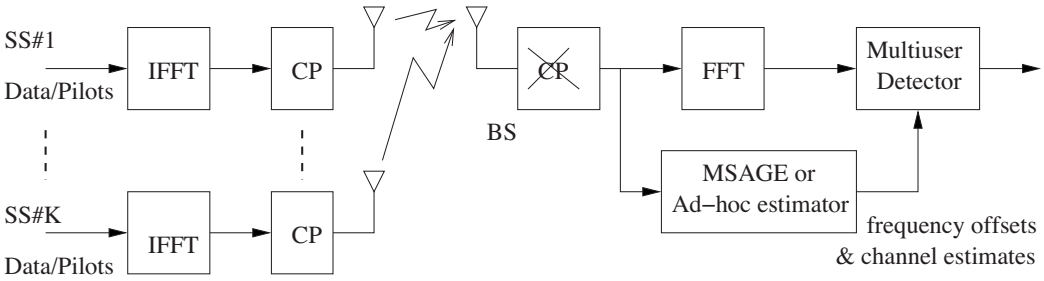

Fig. 1. The system block diagram.

where σ_h^2 is the average total energy of all channel taps of each SS and is defined as

$$\sigma_h^2 = E[\boldsymbol{h}^H \boldsymbol{h}]. \tag{9}$$

III. The Proposed Modified SAGE Method For Joint Estimation of Frequency Offsets and Channels

First, we briefly present how the MUI affects the conventional (CSAGE) method from [10] in term of the signal to interference plus noise ratio (SINR). Then we develop the MSAGE method by incorporating MUI cancelation in both time and frequency domains into the SAGE method and hence by improving the SINR. A system block diagram is shown in Fig. 1.

A. Brief overview of SAGE method

The CSAGE algorithm was first proposed in [12] for parameter estimation of superimposed signals, and has been applied to joint frequency offset and channel estimation of OFDMA uplink in [10]. The CSAGE algorithm is an iterative lower-complexity realization of joint ML estimation of multiple sets of parameters where in each iteration only one parameter set is updated.

Without loss of generality, let us consider the k-th SS. In the CSAGE method, the 'complete-data' $\boldsymbol{r}_k$ (we will call it 'conventional complete-data') is given by

$$\boldsymbol{r}_k = \boldsymbol{\Gamma}(v_k)\boldsymbol{s}_k + \boldsymbol{z} = \boldsymbol{r} - \sum_{m=1, m\neq k}^{K} \boldsymbol{\Gamma}(v_m)\boldsymbol{A}_m \boldsymbol{h}_m. \tag{10}$$

Hence, at the i-th iteration, the expectation of $\boldsymbol{r}_k$ based on the previous estimates $\{\hat{v}_k^{(i-1)}, \hat{\boldsymbol{h}}_k^{(i-1)}\}_{k=1}^K$ is obtained by

$$\begin{aligned} \tilde{\boldsymbol{r}}_k^{(i)} &= \boldsymbol{r} - \sum_{m=1, m\neq k}^{K} \boldsymbol{\Gamma}(\hat{v}_m^{(i-1)})\boldsymbol{A}_m \hat{\boldsymbol{h}}_m^{(i-1)} \\ &= \boldsymbol{\Gamma}(v_k)\boldsymbol{s}_k + \sum_{m=1, m\neq k}^{K} \boldsymbol{e}_m^{(i)} + \boldsymbol{z} \end{aligned} \tag{11}$$

where $(\cdot)^{(i)}$ denotes the i-th iteration and the interference $\boldsymbol{e}_m^{(i)}$ introduced by the m-th SS at the i-th iteration is due to the estimation error and is given by

$$\boldsymbol{e}_m^{(i)} = \boldsymbol{\Gamma}(v_m)\boldsymbol{A}_m \boldsymbol{h}_m - \boldsymbol{\Gamma}(\hat{v}_m^{(i-1)})\boldsymbol{A}_m \hat{\boldsymbol{h}}_m^{(i-1)}. \tag{12}$$

Then the joint ML estimation of the parameters for the k-th SS is performed based on $\tilde{\boldsymbol{r}}_k^{(i)}$ as follows.

$$\{\hat{v}_k^{(i)}, \hat{\boldsymbol{h}}_k^{(i)}\} = \arg\min_{\tilde{v}_k,\ \tilde{\boldsymbol{h}}_k} \left\{ ||\tilde{\boldsymbol{r}}_k^{(i)} - \{\boldsymbol{\Gamma}(\tilde{v}_k)\boldsymbol{A}_k \tilde{\boldsymbol{h}}_k\}||^2 \right\} \tag{13}$$

For $m \neq k$, set

$$\{\hat{v}_m^{(i)}, \hat{\boldsymbol{h}}_m^{(i)}\} = \{\hat{v}_m^{(i-1)}, \hat{\boldsymbol{h}}_m^{(i-1)}\}. \tag{14}$$

The CSAGE method is composed of E-step and M-step which are respectively given by (11) and (13). In addition to the desired signal and the Gaussian noise, MUI is also involved in $\tilde{\boldsymbol{r}}_k$. Since MUI at each iteration is introduced by the previous snap-shot estimation errors (as obvious from (12)), the SINR at the i-th iteration also depends on the snap-shot estimation errors and is given by

$$\begin{aligned} &\mathrm{SINR}_k^{(i)} \\ &= \frac{E[\boldsymbol{s}_k^H \boldsymbol{s}_k]}{\sum_{m=1, m\neq k}^{K} E_e|_{\Delta v_m^{(i)}, \rho_{h,m}^{(i)}} + N\sigma_z^2} \\ &= \frac{1}{\sum_{m=1, m\neq k}^{K} \left[2 + \rho_{h,m}^{(i)} - 2\Re(\alpha(\Delta v_m^{(i)}))\right] + 1/\mathrm{SNR}_k} \end{aligned} \tag{15}$$

where

$$\begin{aligned} E_e|_{\Delta v_m^{(i)}, \rho_{h,m}^{(i)}} &\triangleq E\left[(\boldsymbol{e}_m^{(i)})^H (\boldsymbol{e}_m^{(i)}) | \Delta v_m^{(i)}, \rho_{h,m}^{(i)}\right] \\ &\approx a_m^2 \sigma_h^2 N_p \left[2 + \rho_{h,m}^{(i)} - 2\Re(\alpha(\Delta v_m^{(i)}))\right] \\ \Delta v_m^{(i)} &= v_m - \hat{v}_m^{(i-1)} \\ \rho_{h,m}^{(i)} &= \frac{\sum_{n=0}^{L-1} \left| h_m(n) - \hat{h}_m^{(i-1)}(n) \right|^2}{\sigma_h^2}. \end{aligned} \tag{16}$$

Note that the average interference power from the m-th SS $E_e|_{\Delta v_m^{(i)}, \rho_{h,m}^{(i)}}$ is averaged over the frequency offset v_m and channel response $\boldsymbol{h}_m$ but is conditioned on the snap-shot estimation errors $\Delta v_m^{(i)}$ and $\rho_{h,m}^{(i)}$ (see the Appendix-A for more details). $\alpha(v+n-m) = \frac{1}{N}\sum_{p=0}^{N-1} e^{j2\pi(v+n-m)p/N}$ is the ICI coefficient corresponding to the interference from the n-th tone to the m-th tone when the normalized frequency offset is equal to v, and $\Delta v_m^{(i)}$ and $\rho_{h,m}^{(i)}$ are the snap-shot frequency offset estimation error (the residual frequency offset) and the normalized channel estimation square error, respectively, at the beginning of the i-th iteration.

The effects of the number of SSs K, the estimation errors of the previous iteration, and SNR on the $\mathrm{SINR}_k^{(i)}$ of the CSAGE method are presented in Fig. 2(a). For illustration

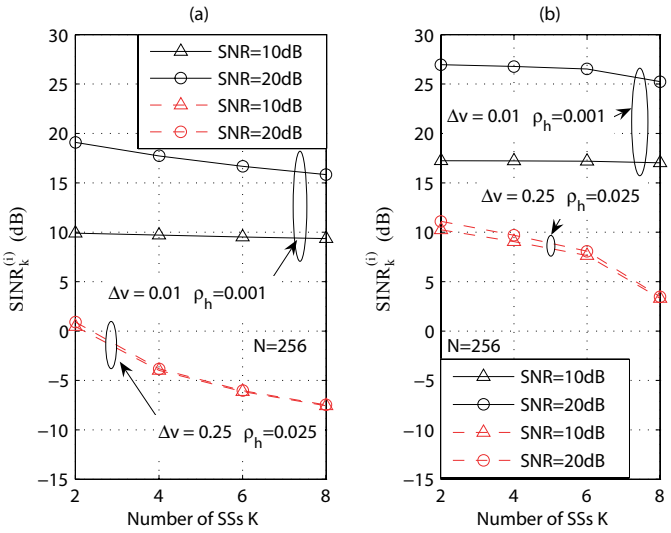

Fig. 2. The impact of K, SNR, Δv, and $\rho_{\boldsymbol{h}}$ on SINR for (a) the CSAGE method, (b) the MSAGE method ($B = 3$).

purpose, we assume that all SSs have the same prevous iteration estimation error statistics $\Delta v_m^{(i)}$ and $\rho_{h,m}^{(i)}$ (denoted by Δv and ρ_h in the figure). $\mathrm{SINR}_k^{(i)}$ of the CSAGE method degrades significantly when K and/or the previous iteration estimation errors increase. In particular, when $\Delta v = 0.25$, $\rho_h = 0.025$ and $K = 8$, $\mathrm{SINR}_k^{(i)}$ is about 20 dB lower than SNR. On the other hand, increasing SNR does not improve $\mathrm{SINR}_k^{(i)}$ much when the previous iteration estimation errors are relatively large (say, $\Delta v = 0.25$ and $\rho_h = 0.025$), which usually occur at the first several iterations of the CSAGE algorithm resulting in a slower convergence rate as well as unreliable estimation performance. This fact explains why the choice of the initial estimates has a very significant effect on the SAGE method's convergence rate and performance [12].

B. The proposed Modified SAGE Method

Different from the CSAGE method which improves SINR iteratively through MUI cancelation in time-domain, we propose a modified method which improves SINR by suppressing MUI in both time and frequency domain as follows.

We perform two steps prior to the estimation of the k-th SS's parameters in order to reduce the interference energy from other SSs. First, we perform frequency offset compensation of the k-th SS in the time-domain using the frequency offset estimate of the previous iteration so that the transmitted energy of each pilot is more concentrated at the corresponding subcarrier. Next, based on the proposed interleaved pilot structure, we perform frequency-domain windowing centered at each of the N_p pilots of the k-th SS in order to suppress interference from other SSs while keeping most of k-th SS's signal energy within the window. Each window has a bandwidth of B subcarriers[2] and is symmetric at the corresponding subcarrier. After these two steps, we define a 'new complete data' in frequency-domain as

$$\boldsymbol{Y}_k = \boldsymbol{\Phi}_k \boldsymbol{F}\boldsymbol{\Gamma}^H(v_k)\boldsymbol{r}_k = \boldsymbol{\Phi}_k \boldsymbol{F}\boldsymbol{s}_k + \boldsymbol{\Phi}_k \boldsymbol{F}\boldsymbol{z} \qquad (17)$$

[2] B is a positive odd integer.

where $\boldsymbol{\Phi}_k$ represents the frequency-domain windowing operation for the k-th SS and its m-th row, n-th column element is given by (18). Note that the corresponding time-domain signal of the new complete data is not the same as the 'conventional complete-data' $\boldsymbol{r}_k$.

The expectation step of the modified SAGE method at the i-th iteration is given by

$$\begin{aligned}\tilde{\boldsymbol{Y}}_k^{(i)} &= \boldsymbol{\Phi}_{\boldsymbol{k}}\boldsymbol{F}\boldsymbol{\Gamma}^H(\hat{v}_k^{(i-1)})[\boldsymbol{r} - \sum_{m=1,m\neq k}^{K} \boldsymbol{\Gamma}(\hat{v}_m^{(i-1)})\boldsymbol{A}_m\hat{\boldsymbol{h}}_m^{(i-1)}] \\ &= \boldsymbol{F}\boldsymbol{\Gamma}(\Delta v_k^{(i)})\boldsymbol{s}_k + \underbrace{[\boldsymbol{\Phi}_{\boldsymbol{k}}\boldsymbol{F}\boldsymbol{\Gamma}(\Delta v_k^{(i)}) - \boldsymbol{F}\boldsymbol{\Gamma}(\Delta v_k^{(i)})]\boldsymbol{s}_k}_{\tilde{\boldsymbol{S}}_k^{(i)}} \\ &+ \sum_{m=1,m\neq k}^{K} \underbrace{\boldsymbol{\Phi}_{\boldsymbol{k}}\boldsymbol{F}\boldsymbol{\Gamma}^H(\hat{v}_k^{(i-1)})\boldsymbol{e}_m^{(i)}}_{\tilde{\boldsymbol{I}}_m^{(i)}} + \underbrace{\boldsymbol{\Phi}_{\boldsymbol{k}}\boldsymbol{F}\boldsymbol{\Gamma}^H(\hat{v}_k^{(i-1)})\boldsymbol{z}}_{\boldsymbol{W}_k^{(i)}} \qquad (19)\end{aligned}$$

where $\tilde{\boldsymbol{S}}_k^{(i)}$ is the signal leakage outside the k-th SS's frequency-domain windows, $\tilde{\boldsymbol{I}}_m^{(i)}$ is the interference introduced by the m-th SS within the k-th SS's windows, and $\boldsymbol{W}_k^{(i)}$ is Gaussian noise within the k-th SS's windows. From the above equation, we observe that $\boldsymbol{Y}_k$ is no longer the 'complete-data' for v_k but for the frequency offset estimation error Δv_k.

The average energies of the interference and the signal leakage are, respectively, given by (20) and (21),(see the Appendix-A for more details) where

$$\beta(p) \triangleq \sum_{q=-\frac{B-1}{2}}^{\frac{B-1}{2}} e^{\frac{-j2\pi qp}{D}} = \begin{cases} B, & p = 0 \\ \frac{\sin(\frac{\pi Bp}{D})}{\sin(\frac{\pi p}{D})}, & \text{otherwise} \end{cases} \qquad (22)$$

$$\gamma(p) \triangleq E[e^{\frac{-j2\pi pv}{D}}] = \begin{cases} 1, & p = 0 \\ \frac{D}{p\pi}\sin(\frac{\pi p}{D}), & \text{otherwise} \end{cases} \qquad (23)$$

$$J_n \triangleq \{\lceil -n/N_P\rceil, \lceil -n/N_P\rceil + 1, \ldots, \lfloor (N-1-n)/N_P\rfloor\}. \qquad (24)$$

The SINR of the new complete data $\tilde{\boldsymbol{Y}}_k^{(i)}$ at the i-th iteration is given by

$$\widetilde{\mathrm{SINR}}_k^{(i)} = \frac{E[\boldsymbol{s}_k^H\boldsymbol{s}_k]}{\sum_{m=1,m\neq k}^{K} E_{\tilde{I}}|_{\Delta v_m^{(i)},\rho_{h,m}^{(i)}} + N\sigma_z^2 B/D + E_{\tilde{S}}|_{\Delta v_k^{(i)}}}. \qquad (25)$$

Different from (15), $\widetilde{\mathrm{SINR}}_k^{(i)}$ depends on not only the estimation errors and SNR but also the window bandwidth B. The optimum choice of B at the i-th iteration is obtained by maximizing (25) as

$$B^{\dagger,(i)} = \arg\max_{B}\{\widetilde{\mathrm{SINR}}_k^{(i)}\}. \qquad (26)$$

Unfortunately, no closed-form solution for the exact optimum B could be obtained due to the intractable optimization problem and the lack of the knowledge of the exact estimation errors. However, for given system parameters (N, N_p, K) and fixed design values of the estimation errors, the exact value of $B^{\dagger,(i)}$ could be numerically obtained. Fig. 3 shows the effect of B on $\widetilde{\mathrm{SINR}}$ for several values of K, SNR and the estimation errors when N is set to 256. Smaller K or/and larger estimation errors of the previous iteration give a larger optimum B. For a small K, the optimum B increases with SNR. If a fixed B is used throughout the proposed algorithm

$$\boldsymbol{\Phi}_k(m,n) = \begin{cases} 1, & (m=n)\&(m=\mathrm{mod}(uD+\Delta_k \pm b, N),\ u=0,\ldots,N_p-1;\ b=0,\ldots,\frac{B-1}{2}) \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

$$E_{\tilde{I}}|_{\Delta v_m^{(i)},\rho_{h,m}^{(i)}} \triangleq E\left[(\tilde{\boldsymbol{I}}_m^{(i)})^H \tilde{\boldsymbol{I}}_m^{(i)} | \Delta v_m^{(i)}, \rho_{h,m}^{(i)}\right] \approx \frac{a_m^2\sigma_h^2}{D^2}\sum_{n=0}^{N-1}\sum_{p\in J_n}\beta(p)\gamma^2(p)e^{\frac{j2\pi(\Delta_m-\Delta_k)p}{D}} \times \left[1+(1+\rho_{h,m}^{(i)})e^{\frac{-j2\pi\Delta v_m^{(i)}p}{D}} - e^{\frac{-j2\pi(n+pN_p)\Delta v_m^{(i)}}{N}} - e^{\frac{j2\pi n\Delta v_m^{(i)}}{N}}\right] \tag{20}$$

$$\begin{aligned} E_{\tilde{S}}|_{\Delta v_k^{(i)}} &\triangleq E\left[(\tilde{\boldsymbol{S}}_k^{(i)})^H \tilde{\boldsymbol{S}}_k^{(i)} | \Delta v_k^{(i)}\right] = N_p a_k^2\sigma_h^2 - E\left[\sum_{q=-\frac{B-1}{2}}^{\frac{B-1}{2}}\sum_{m=0}^{N_p-1}\sum_{u=0}^{N_p-1}|x_k(uD+\Delta_k)H_k(uD+\Delta_k)\right. \\ &\quad \left.\alpha(\Delta v_k^{(i)}+(u-m)D+q)\right|^2\Bigg] \\ &= a_k^2\sigma_h^2\left[N_p - \frac{1}{D^2}\sum_{n=0}^{N-1}\sum_{p\in J_n}\beta(p)e^{j2\pi\Delta v_k^{(i)}p/D}\right] \end{aligned} \tag{21}$$

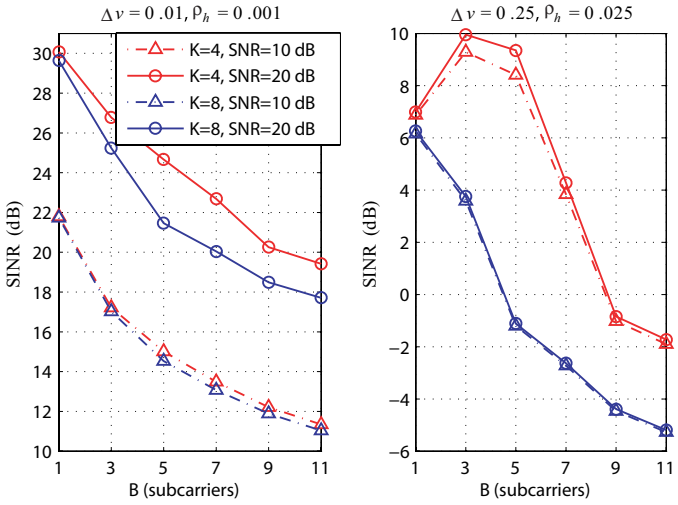

Fig. 3. The effect of B on SINR under different SNR, K, Δv, and $\rho_{\boldsymbol{h}}$ where $N = 256$.

under $N = 256$ condition, $B = 3$ would be a good choice since for most of the conditions of SNR, K, Δv and ρ_h, the differences between the maximum $\widetilde{\mathrm{SINR}}$ and the $\widetilde{\mathrm{SINR}}$ for $B = 3$ are quite small. Alternatively, when the system parameters changed (e.g., for large N and small K), the differences between the maximum $\widetilde{\mathrm{SINR}}$ and the $\widetilde{\mathrm{SINR}}$ for small B could be large, and then, an adaptive B can be used. We can set the initial value of B by utilizing the knowledge of K and the operating SNR that the BS has, together with (relatively large) designed values of initial estimation errors. Then after a few iterations as the estimation errors decrease, we can choose a smaller B.

Fig. 2(b) shows the effects of K on the $\widetilde{\mathrm{SINR}}_k^{(i)}$ of the MSAGE method for several values of SNR and estimation errors. For larger estimation errors ($\Delta v = 0.25$ and $\rho_h = 0.025$), an increase in K from 2 to 6 results in a negligible SINR degradation for the MSAGE method as shown in Fig. 2(b) but a considerable SINR degradation for the CSAGE method as presented in Fig. 2(a). This fact implies that the MSAGE method is more robust to the number of active SSs than the CSAGE method. In addition to this robustness advantage, we observe from Fig. 2(a) and (b) that the SINR improvement of our proposed MSAGE method over the existing CSAGE method ranges from 7 to 10 dB.

Our proposed MSAGE algorithm using the 'new complete-data' is summarized as follows:

- E-step: For $i = 1, 2, ...,$ and $k = \mathrm{mod}(i, K)$, compute (27).
- M-step: Compute

$$\{\Delta\hat{v}_k^{(i)}, \hat{\boldsymbol{h}}_k^{(i)}\} = \arg\min_{\Delta\tilde{v}_k,\ \tilde{\boldsymbol{h}}_k}\left\{||\tilde{\boldsymbol{Y}}_k^{(i)} - \mathfrak{B}_{\boldsymbol{\Phi}_k}\{\boldsymbol{F}\boldsymbol{\Gamma}(\Delta\tilde{v}_k)\boldsymbol{A}_k\tilde{\boldsymbol{h}}_k\}||^2\right\} \tag{28}$$

$$\hat{v}_k^{(i)} = \hat{v}_k^{(i-1)} + \Delta\hat{v}_k^{(i)}. \tag{29}$$

For $m \neq k$, set

$$\{\hat{v}_m^{(i)}, \hat{\boldsymbol{h}}_m^{(i)}\} = \{\hat{v}_m^{(i-1)}, \hat{\boldsymbol{h}}_m^{(i-1)}\} \tag{30}$$

where $\mathfrak{B}_{\boldsymbol{\Phi}_k}\{\boldsymbol{Z}\}$ represents the operation which returns a $BN_p \times N_{\mathrm{col}}$ matrix by picking BN_p rows from the $N \times N_{\mathrm{col}}$ matrix $\boldsymbol{Z}$ according to the indices of non-zero elements in the main diagonal of $\boldsymbol{\Phi}_k$. Solving the above optimization in (28) gives

$$\Delta\hat{v}_k^{(i)} = \arg\max_{\Delta\tilde{v}_k} \Re\left\{(\tilde{\boldsymbol{Y}}_k^{(i)})^H\boldsymbol{\Omega}(\Delta\tilde{v}_k)\boldsymbol{\Upsilon}(\Delta\tilde{v}_k)\tilde{\boldsymbol{Y}}_k^{(i)}\right\} \tag{31}$$

and

$$\hat{\boldsymbol{h}}_k^{(i)} = \boldsymbol{\Upsilon}(\Delta\hat{v}_k^{(i)})\tilde{\boldsymbol{Y}}_k^{(i)} \tag{32}$$

where

$$\boldsymbol{\Omega}(\Delta\tilde{v}_k) = \mathfrak{B}_{\boldsymbol{\Phi}_k}\{\boldsymbol{F}\boldsymbol{\Gamma}(\Delta\tilde{v}_k)\boldsymbol{A}_k\} \tag{33}$$

$$\tilde{\boldsymbol{Y}}_k^{(i)} = \mathfrak{B}_{\boldsymbol{\Phi}_k}\left\{\boldsymbol{F}\boldsymbol{\Gamma}^H(\hat{v}_k^{(i-1)})\left[\boldsymbol{r} - \sum_{m=1,m\neq k}^{K}\boldsymbol{\Gamma}(\hat{v}_m^{(i-1)})\boldsymbol{A}_m\hat{\boldsymbol{h}}_m^{(i-1)}\right]\right\}. \quad (27)$$

$$\boldsymbol{\Upsilon}(\Delta\hat{v}_k^{(i)}) = \left(\boldsymbol{\Omega}^H(\Delta\hat{v}_k^{(i)})\boldsymbol{\Omega}(\Delta\hat{v}_k^{(i)})\right)^{-1}\boldsymbol{\Omega}^H(\Delta\hat{v}_k^{(i)}). \quad (34)$$

The above proposed estimators have smaller complexity than the counterparts in the CSAGE method due to the smaller sizes of matrices and vectors involved. In our algorithm, the initial estimates $\{\hat{v}_k^{(0)}\}$ are set to zeros (the mean value of $\{\hat{v}_k^{(0)}\}$), and $\{\hat{\boldsymbol{h}}_k^{(0)}\}$ are obtained from (32) with

$$\begin{aligned}\Delta\hat{v}_k^{(0)} &\triangleq 0\\ \tilde{\boldsymbol{Y}}_k^{(0)} &\triangleq \mathfrak{B}_{\boldsymbol{\Phi}_k}\{\boldsymbol{F}\boldsymbol{r}\}.\end{aligned} \quad (35)$$

Note that if we know $\{d_k\}$ or their estimates (i.e., fine timing offset estimator is applied) and/or the exact number of channel taps, then we can easily incorporate them into our approach (e.g., by using the significant channel taps selection approach in [16] or [17]) to get enhanced channel estimation performance which may further reduce the number of iterations required in our approach.

IV. The Proposed Ad Hoc Iterative Method

Our proposed MSAGE method converges much faster than the CSAGE method due to the SINR improvement and also has lower complexity. But the MSAGE's frequency offset estimator in (31) still requires a large complexity. In this section, we present a much lower complexity estimator set to replace the M-step ((31) and (32)) of the MSAGE method.

First, the expectation of the 'new complete-data' in the time-domain at the i-th iteration can be expressed as

$$\begin{aligned}&\tilde{\boldsymbol{y}}_k^{(i)}\\ &= \boldsymbol{F}^H\boldsymbol{\Phi}_{\boldsymbol{k}}\boldsymbol{F}\boldsymbol{\Gamma}^H(\hat{v}_k^{(i-1)})\left[\boldsymbol{r} - \sum_{m=1,m\neq k}^{K}\boldsymbol{\Gamma}(\hat{v}_m^{(i-1)})\boldsymbol{A}_m\hat{\boldsymbol{h}}_m^{(i-1)}\right]\\ &= \boldsymbol{\Gamma}(\Delta v_k^{(i)})\boldsymbol{s}_k + \underbrace{\boldsymbol{F}^H\left[\sum_{m=1,m\neq k}^{K}\tilde{\boldsymbol{I}}_m^{(i)} + \tilde{\boldsymbol{S}}_k^{(i)} + \boldsymbol{W}_k^{(i)}\right]}_{\tilde{\boldsymbol{w}}_k^{(i)}}.\end{aligned} \quad (36)$$

Therefore, the n-th sample of $\tilde{\boldsymbol{y}}_k^{(i)}$ can be given by

$$\tilde{y}_k^{(i)}(n) = e^{\frac{j2\pi n\Delta v_k^{(i)}}{N}}[s_k(n) + \bar{w}_k^{(i)}(n)]. \quad (37)$$

where $\bar{w}_k^{(i)}(n) = e^{\frac{-j2\pi n\Delta v_k^{(i)}}{N}}\tilde{w}_k^{(i)}(n)$. Note that due to our proposed pilot design, $\boldsymbol{s}_k$ contains D identical (if certain known phase shifts are compensated) parts of length N_p each. We utilize this periodicity property of the training signals together with SINR $\gg 1$ in developing our ad hoc method. Define the correlation terms as

$$\begin{aligned}R_k^{(i)}(m) &= \frac{1}{N-mN_p}\sum_{n=0}^{N-mN_p-1}(\tilde{y}_k^{(i)}(n))^*\tilde{y}_k^{(i)}(n+mN_p)\\ &= e^{\frac{j2\pi m(\Delta_k+\Delta v_k^{(i)})}{D}}E_{s,k}[1+\gamma_k^{(i)}(m)]\ ,\ m=0,1,\ldots,M-1\end{aligned} \quad (38)$$

where the values of $E_{s,k}$ and $\gamma_k^{(i)}(m)$ are given in (39) and (40), respectively, and M is a design parameter satisfying $M \leq \frac{D}{2}$. Next, define

$$\begin{aligned}\varphi_k^{(i)}(m) &= \arg\left\{R_k^{(i)}(m)(R_k^{(i)}(m-1))^*\right\}\\ &\approx \frac{2\pi(\Delta v_k^{(i)}+\Delta_k)}{D} + \gamma_{I,k}^{(i)}(m) - \gamma_{I,k}^{(i)}(m-1)\end{aligned} \quad (41)$$

where $\gamma_{I,k}^{(i)}(m)$ is the imaginary part of $\gamma_k^{(i)}(m)$. The above approximation holds when $\widetilde{\mathrm{SINR}}_k^{(i)} \gg 1$ (our method satisfies this requirement, c.f. Fig. 2). Developing the best linear unbiased estimator (e.g., [13],[14]) is infeasible in this case because the exact values of the frequency offsets are required in the co-variance matrix calculation. Hence, we apply the least-squares (LS) principle to (41) and obtain our ad hoc frequency offset estimator as

$$\Delta\hat{v}_k^{(i)} = \frac{D}{2M\pi}\sum_{m=1}^{M}\varphi_k^{(i)}(m) - \Delta_k. \quad (42)$$

Performing the frequency offset compensation using the above estimate and applying the LS channel estimation, we obtain the channel estimate for the k-th SS at the i-th iteration where $k = \mathrm{mod}(i,k)$ as

$$\hat{\boldsymbol{h}}_k^{(i)} = (\boldsymbol{A}_k^H\boldsymbol{A}_k)^{-1}\boldsymbol{A}_k^H\boldsymbol{\Gamma}^H(\Delta\hat{v}_k^{(i)})\tilde{\boldsymbol{y}}_k^{(i)}. \quad (43)$$

In summary, our proposed ad hoc iterative method is obtained by replacing (27), (31) and (32) with (36), (42) and (43), respectively.

V. Simulation Results

A. Simulation Setup

The simulation parameters are selected from [2]. Uplink bandwidth is 2.5 MHz corresponding to sampling period $T_s = 0.35\mu s$. The sub-carrier spacing is 11.16 kHz. The system has $N = 256$ sub-carriers unless mentioned otherwise. The maximum timing offset $d_{\max}$ is set to 8 samples[3]. N_g is set to 16. Actual timing offsets (in samples) $\{d_k\}$ are independently generated with a discrete uniform distribution within the range [0, 8].[4] The frequency offsets $\{v_k\}$ are independently generated with a uniform distribution within $[-0.5, 0.5]$. The SUI-3 channel model with 3 paths [15] is used. The CIR including the combined transmit and receive filter effect (before incorporating the timing offset) for the k-th user is given by

$$\bar{h}_k(l) = \sum_{i=0}^{2}\zeta_i g_T(lT_s - \tau_i - t_0), l = 0,\ldots,\bar{L}-1 \quad (44)$$

[3]It satisfies the synchronization requirement ($\pm 1/4$ of the minimum CP) defined in the IEEE 802.16-2004.

[4]Non-integer part of the timing offset can be absorbed into the CIR and hence, we just use integer timing offsets (in samples) in the simulation.

$$E_{s,k} = \frac{1}{N_p}\sum_{n=0}^{N_p-1}|s_k(n)|^2 \tag{39}$$

$$\gamma_k^{(i)}(m) = \frac{1}{(N-mN_P)E_{s,k}}e^{\frac{-j2\pi m\Delta_k}{D}}\sum_{n=0}^{N-mN_p-1}\Big[s_k^*(n)\bar{w}_k^{(i)}(n+mN_p) + s_k(n+mN_p)(\bar{w}_k^{(i)}(n))^* + (\bar{w}_k^{(i)}(n))^*\bar{w}_k^{(i)}(n+mN_p)\Big] \tag{40}$$

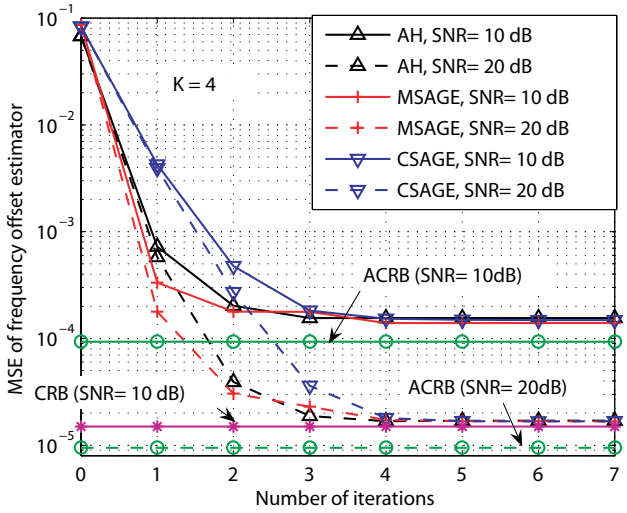

Fig. 4. MSE comparison of frequency offset estimators ($K = 4$).

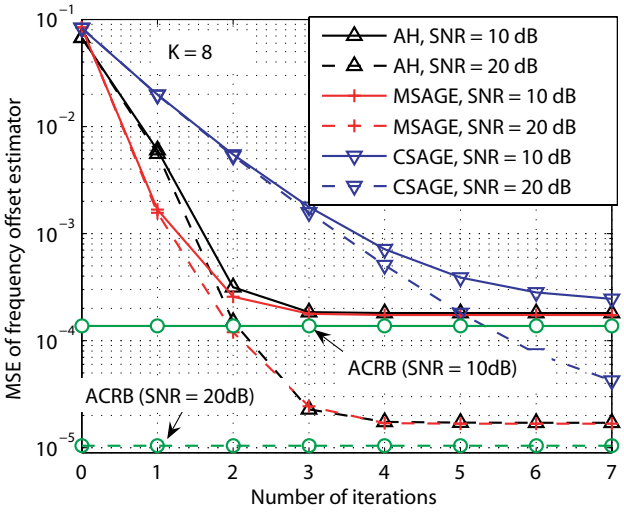

Fig. 5. MSE comparison of frequency offset estimators ($K = 8$).

where $g_T(t)$ is a spectral raised-cosine filter with a roll-off factor of 0.5, $\{\zeta_i\}$ are the complex Gaussian channel path gains, $\{\tau_i\}$ are the the channel path delays, and t_0 is a time shift for causality. The average total energy of the channel taps σ_h^2 for each user is set to unity and $\bar{L}$ is set to 7.

The proposed preamble used in the AH and MSAGE methods is defined by equation (1) with $N_p = 16$. During data transmission of the proposed methods, arbitrary (random) CAS is used where N sub-carriers are divided into $Q = K$ sub-channels and each sub-channel has $P = N/K$ sub-carriers (corresponding to a fully-loaded system). For the reference CSAGE method, an arbitrary CAS is used (each SS uses P sub-carriers too) but is kept the same for both the preamble and the data transmission. For the fixed B approach, we use $B = 3$ sub-carriers. For the ad hoc frequency estimator in (42), M is set to $\frac{N}{2N_p}$ for all SSs. For the CSAGE and MSAGE methods, the number of grid points N_s for frequency offset estimator is set to 201 within the range of $[-0.5, 0.5]$. Unless mentioned otherwise, all the simulation results are based on the 10000 simulation runs.

B. Simulation Results and Discussions

Figs. 4-9 show the mean-square error (MSE) and bit error rate (BER) performance of AH, MSAGE and CSAGE methods versus the number of iterations for different values of the number of SSs K.

Note that we count K iterations (i-loop in the summarized description of the MSAGE algorithm) as one iteration in the figures since only one SS is updated each time and there are K SSs.

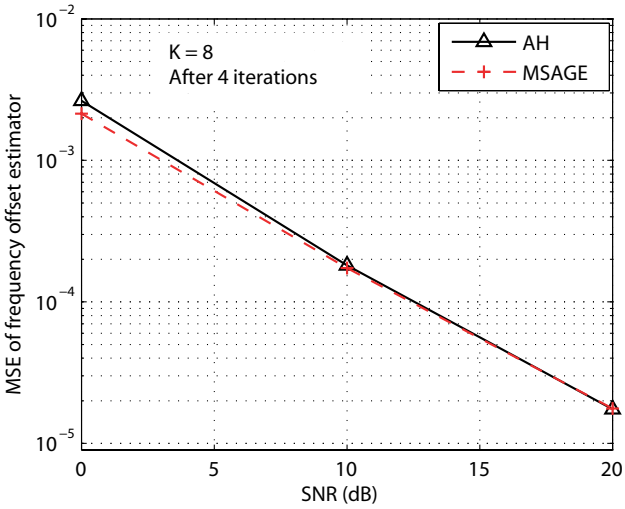

Fig. 6. MSE of the proposed frequency offset estimators(K=8, 4 iterations).

1) MSE Performance of Frequency Offset Estimators: Fig. 4 and Fig. 5 show the MSE performances of the frequency offset estimators from the CSAGE, MSAGE, and AH methods for $K = 4$ and $K = 8$, respectively. Both the convergence rate and the MSE performance of the CSAGE method degrade significantly when K increases from 4 to 8. The reason is that the increase in K enlarges the MUI and results in a lower SINR for the CSAGE method. On the other hand, by suppressing the MUI ahead of the parameters estimation,

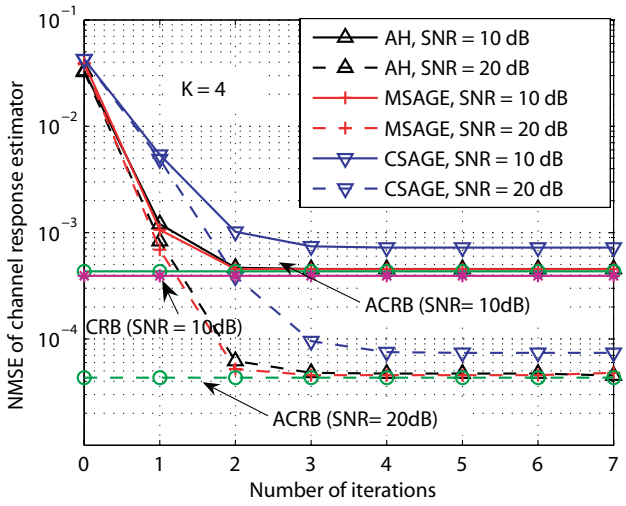

Fig. 7. NMSE comparison of channel estimators (K = 4).

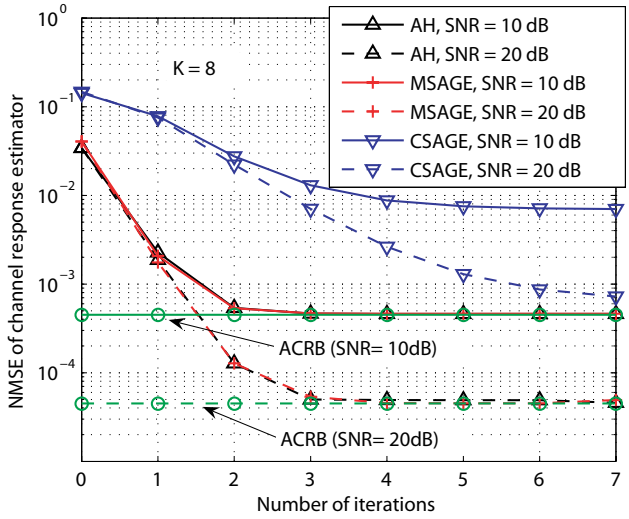

Fig. 8. NMSE comparison of channel estimators (K = 8).

AH and MSAGE methods experience no noticeable MSE performance degradation when K increases. The increase in K only makes the proposed two methods converge slower at the first two iterations. The proposed MSAGE and AH methods need less number of iterations than the CSAGE method to achieve the same estimaton performance. The MSEs of the proposed methods approach to the average CRB (ACRB) (see Appendix-B) after 3 iterations. But the CSAGE's MSE may not approach to the average CRB even after 7 iterations. Fig. 6 shows the MSE performance curves of the two proposed frequency offset estimators versus SNR when $K = 8$. We observe that the two proposed frequency offset estimators achieve approximately the same performance after 4 iterations. Note that the performance of the two proposed iterative methods also depends on the SINR and the number of iterations.

2) NMSE Performance of Channel Estimators: We use the normalized mean-square error (NMSE) defined by

$$\mathrm{NMSE_h} = \frac{1}{KN_g} E\left[\sum_{k=1}^{K}\sum_{l=0}^{N_g-1} |h_k(l) - \hat{h}_k(l)|^2\right] \qquad (45)$$

as the CIR estimation performance measure. The NMSE performances of the three channel estimators for $K = 4$ and $K = 8$ are plotted in Fig. 7 and Fig. 8, respectively. The CSAGE's channel estimator is very sensitive to the MUI, similar to its frequency offset estimator. For example, if all methods stop after 3 iterations, an increase of K from 4 to 8 at SNR = 20dB degrades the NMSE of the CSAGE method from 10^{-4} to 4×10^{-2}. But the NMSEs of both AH and MSAGE remain almost the same at 4×10^{-5}. Hence, both proposed methods give not only a better robustness against MUI but also a better MSE performance than the CSAGE method. The NMSE of the proposed methods approach to the average CRB after 2 or 3 iterations. But there is a significant gap between the CSAGE's NMSE and the average CRB even after 7 iterations.

3) Uncoded BER Performance: QPSK modulation is used for data transmission in our simulation, and all BER results are obtained from 1000 simulation runs. Fig. 9 shows the

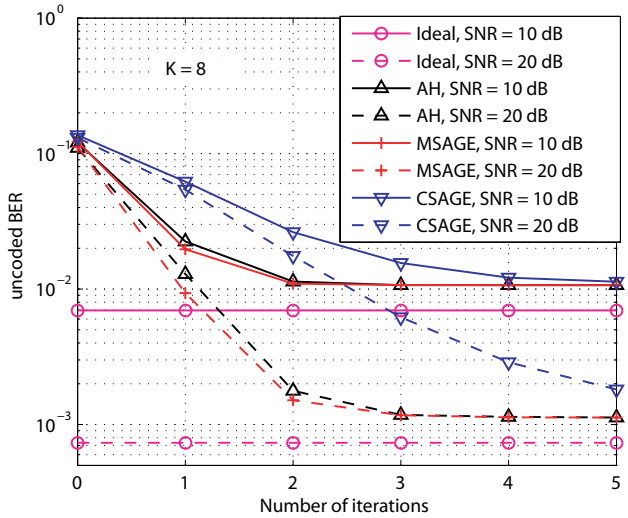

Fig. 9. Uncoded BER comparison (K=8).

BER performance curves for $K = 8$. The BERs of the proposed AH and MSAGE methods are much closer to the ideal performance obtained with known frequency offsets and channel responses than that of the CSAGE method due to the better synchronization and channel estimation performance. After 3 iterations at SNR = 20 dB, the BER of the CSAGE method is about seven times larger than that of the proposed methods.

4) Adaptive B Effect: Under the condition of large N and small K (e.g., $N = 1024$ and $K = 2$), when the estimation errors decrease, both optimum value of B as well as the corresponding value of the maximum $\widetilde{\mathrm{SINR}}$ change significantly. The estimation errors at the first iteration could be large but they will decrease at later iterations. Hence, the value of B should be adjusted adaptively according to the estimation error variation. Fig. 10 shows the performance of the proposed AH frequency offset estimator using an adaptive B for $N = 1024$ and $K = 2$. We set the parameter $B_{\mathrm{adpt}} = 33$ at the first iteration and change it to 3 at the beginning of the second iteration. For the considered conditions, using a

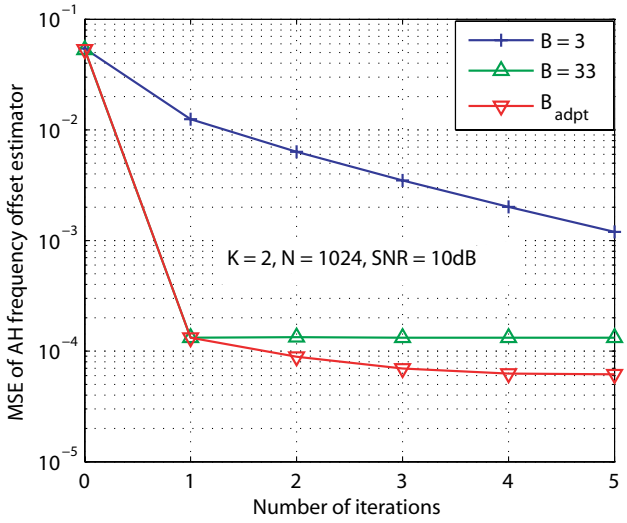

Fig. 10. MSE of the proposed frequency offset estimator using an adaptive B.

small fixed B converges very slowly while using a large fixed B gives a fast convergence rate but its MSE can be affected since it cannot fully exploit the power concentration and noise suppression property in the frequency domain. Using B_{adpt} yields not only a fast convergence (hence low complexity) but also a better MSE.

5) Computation Complexity: The overall computation load of iterative methods consist of two parts: computation complexity at each iteration and the number of iterations used. At each iteration, the CSAGE method needs $\mathcal{O}(KN_sNN_g)$ complex operations due to $N \times N_g$ dimension matrix multiplication at each grid point. For the MSAGE method, only $BN_p \times N_g$ matrix multiplication is needed at each grid point (the matrix $\mathbf{\Omega}_k(v)$ and $\mathbf{\Upsilon}_k(v)$ can be precomputed and stored) so that the corresponding computation complexity order is reduced to $\mathcal{O}(KN_sBN_pN_g)$ at each iteration. The AH method needs the smallest computation complexity order per iteration which is $\mathcal{O}(KMN)$. For our simulation parameters, the order of complex operations for the CSAGE, MSAGE and AH methods are approximately $823296K$, $154368K$, and $2048K$, respectively, where K is the number of users. Moreover, the CSAGE method needs more iterations to converge. Hence, both proposed methods have much less computation load than the CSAGE method. In particular, the proposed AH method achieves almost the same performance as the proposed MSAGE method while requiring significantly lower complexity and hence, the AH mehtod is very appealing for practical implementation.

VI. Conclusions

We have presented two iterative methods for the joint estimation of frequency offsets and channels in OFDMA uplink by using disjoint sets of cyclically equal-spaced, equal energy pilot tones assigned to the uplink users. The first method (MSAGE) incorporates multiuser interference cancellation in both time and frequency domains into the SAGE method. The second method (AH) replaces the complexity-intensive M-step of the MSAGE method with a significantly lower complexity version. By utilizing the time-domain characteristics of the proposed pilot signals, the frequency-domain power concentration property of OFDMA systems, and the multiuser interference cancellation in both time and frequency domains, both proposed methods achieve faster convergence rate, better estimation performance, lower complexity, and better robustness against the number of users than the existing conventional SAGE method. The AH method is very appealing for practical implementation since it achieves almost the same peformance as the MSAGE method while requiring significantly lower complexity.

Appendix-A

In this appendix, we present the derivations of (16) and (20). First, the n-th sample of $\boldsymbol{e}_m^{(i)}$ is given by (46).

Recalling v_m is a random variable with uniform distribution in $[-0.5, 0.5]$, we obtain the (n,g)-th element of the covariance matrix of $\boldsymbol{e}_m^{(i)}$ as shown in (47), where we have used the following two assumptions. First, the channel frequency responses at pilot tones are uncorrelated since the pilot distance are large enough to be greater than the channel coherence bandwidth. Second, the channel estimation error $\Delta H(g) = H_m(g) - \hat{H}_m(g)$ is uncorrelated with $H_m(g)$ since the LS channel estimator's MSE does not depend on the channel characteristics.

Next, by using $\tilde{\boldsymbol{I}}_m^{(i)}$ from (19), we calculate (20) as shown in (48), where $\beta(p)$, $\gamma(p)$ and J_n are defined in (22), (23) and (24), respectively.

Appendix-B

This section presents the average CRB of the joint estimation of frequency offsets and channel coefficients in OFDMA uplink which was used in evaluating our proposed method's performance.

We denote the real-valued estimation parameter vector as

$$\boldsymbol{\theta} = [\boldsymbol{v}^T, \boldsymbol{h_R}^T, \boldsymbol{h_I}^T]^T, \tag{49}$$

where

$$\boldsymbol{v} = [v_1, v_2, \ldots, v_K]^T \tag{50}$$

$$\boldsymbol{h_R} = [\Re(\boldsymbol{h}_1^T, \boldsymbol{h}_2^T, \ldots, \boldsymbol{h}_K^T)]^T \tag{51}$$

$$\boldsymbol{h_I} = [\Im(\boldsymbol{h}_1^T, \boldsymbol{h}_2^T, \ldots, \boldsymbol{h}_K^T)]^T. \tag{52}$$

The Fisher information matrix is given by

$$\boldsymbol{\mathcal{F}} = -E\left[\frac{\partial^2 ln[f(\boldsymbol{r};\boldsymbol{\theta})]}{\partial^2 \boldsymbol{\theta}}\right] \tag{53}$$

where $f(\boldsymbol{r};\boldsymbol{\theta})$ is the probability density function of the received signal conditioned on $\boldsymbol{\theta}$ and is given by

$$f(\boldsymbol{r};\boldsymbol{\theta}) = \frac{1}{(\pi\sigma_z^2)^N} e^{-\frac{1}{\sigma_z^2}\|\boldsymbol{r} - \sum_{k=1}^{K} \mathbf{\Gamma}(v_k)\boldsymbol{A}_k\boldsymbol{h}_k\|^2}. \tag{54}$$

The snap-shot CRB for the covariance matrix of the estimation parameter vector is given by the inverse of the Fisher information matrix and the snap-shot CRB for the variance of n-th variable of the parameter vector is given by the n-th diagonal element of the snap-shot CRB for the covariance matrix. Define $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_K]$ and $\boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_K]$ where $\boldsymbol{u_k} = \mathbf{\Gamma}(v_k)\boldsymbol{A}_k$, $\boldsymbol{q}_k = \boldsymbol{D}\mathbf{\Gamma}(v_k)\boldsymbol{A}_k\boldsymbol{h}_k$ and $\boldsymbol{D} =$

$$e_m^{(i)}(n) = \frac{1}{\sqrt{N}} \sum_{u=0}^{N_p-1} x_m(uD+\Delta_m) \Big[H_m(uD+\Delta_m) e^{j2\pi n(uD+\Delta_m+v_m)/N} - \hat{H}_m^{(i-1)}(uD+\Delta_m) e^{j2\pi n(uD+\Delta_m+\hat{v}_m^{(i-1)})/N} \Big] \tag{46}$$

$$\begin{aligned}
&E[(e_m^{(i)}(n))^* e_m^{(i)}(g)|_{\Delta v_m^{(i)}, \rho_{h,m}^{(i)}}] \\
&= \frac{1}{N} \sum_{u=0}^{N_p-1} \sum_{l=0}^{N_p-1} E\Big[x_m^*(uD+\Delta_m) x_m(lD+\Delta_m) \Big(H_m(uD+\Delta_m) e^{j2\pi n(uD+\Delta_m+v_m)/N} \\
&- \hat{H}_m^{(i-1)}(uD+\Delta_m) e^{j2\pi n(uD+\Delta_m+\hat{v}_m^{(i-1)})/N} \Big)^* \\
&\times \Big(H_m(lD+\Delta_m) e^{j2\pi g(lD+\Delta_m+v_m)/N} - \hat{H}_m^{(i-1)}(lD+\Delta_m) e^{j2\pi g(lD+\Delta_m+\hat{v}_m^{(i-1)})/N} \Big) \Big] \\
&\approx \frac{a_m^2}{N} \sum_{u=0}^{N_p-1} E\Big[e^{j2\pi(uD+\Delta_m+v_m)(g-n)/N} \Big] \Big\{ E\left[|H_m(uD+\Delta_m)|^2 \right] \\
&+ E\Big[|\hat{H}_m^{(i-1)}(uD+\Delta_m)|^2 \Big] e^{-j2\pi\Delta v_m^{(i)}(g-n)/N} - E\Big[(H_m(uD+\Delta_m))^* \hat{H}_m^{(i-1)}(uD+\Delta_m) \Big] e^{-j2\pi\Delta v_m^{(i)} g/N} \\
&- E\Big[H_m(uD+\Delta_m) (\hat{H}_m^{(i-1)}(uD+\Delta_m))^* \Big] e^{j2\pi\Delta v_m^{(i)} n/N} \Big\} \\
&\approx \frac{a_m^2 \sigma_h^2}{N} \sum_{u=0}^{N_p-1} \Big\{ e^{j2\pi(uD+\Delta_m)(g-n)/N} E\Big[e^{j2\pi v_m(g-n)/N} \Big] \Big[1 + (1+\rho_{h,m}^{(i)}) e^{-j2\pi\Delta v_m^{(i)}(g-n)/N} \\
&- e^{-j2\pi\Delta v_m^{(i)} g/N} - e^{j2\pi\Delta v_m^{(i)} n/N} \Big] \Big\}
\end{aligned} \tag{47}$$

$$\begin{aligned}
&E\Big[(\tilde{\boldsymbol{I}}_m^{(i)})^H \tilde{\boldsymbol{I}}_m^{(i)}|_{\Delta v_m^{(i)}, \rho_{h,m}^{(i)}} \Big] = \\
&\frac{1}{N} \sum_{q=-\frac{B-1}{2}}^{\frac{B-1}{2}} \sum_{u=0}^{N_p-1} E[\sum_{n=0}^{N-1} e^{j2\pi(\hat{v}_k^{(i-1)}+uD+\Delta_k+q)n/N} (e_m^{(i)}(n))^* \sum_{g=0}^{N-1} e^{-j2\pi(\hat{v}_k^{(i-1)}+uD+\Delta_k+q)g/N} e_m^{(i)}(g)] \\
&= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{g=0}^{N-1} e^{\frac{j2\pi\Delta_k(n-g)}{N}} E\left[e^{\frac{j2\pi\hat{v}_k^{(i-1)}(n-g)}{N}} \right] E\Big[(e_m^{(i)}(g))^* e_m^{(i)}(n) \Big] \sum_{u=0}^{N_p-1} e^{\frac{j2\pi u(n-g)}{N_p}} \sum_{q=-\frac{B-1}{2}}^{\frac{B-1}{2}} e^{\frac{j2\pi q(n-g)}{N}} \\
&= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{p\in J_n} e^{\frac{-j2\pi\Delta_k p}{D}} E\left[e^{\frac{-j2\pi\hat{v}_k^{(i-1)} p}{D}} \right] E\Big[(e_m^{(i)}(n))^* e_m^{(i)}(n+pN_p) \Big] N_p \beta(p) \\
&\approx \frac{a_m^2 \sigma_h^2}{D^2} \sum_{n=0}^{N-1} \sum_{p\in J_n} \beta(p) \gamma^2(p) e^{\frac{j2\pi(\Delta_m-\Delta_k)p}{D}} \\
&\times \left[1 + (1+\rho_{h,m}^{(i)}) e^{\frac{-j2\pi\Delta v_m^{(i)} p}{D}} - e^{\frac{-j2\pi(n+pN_p)\Delta v_m^{(i)}}{N}} - e^{\frac{j2\pi n\Delta v_m^{(i)}}{N}} \right]
\end{aligned} \tag{48}$$

diag$\{[0, 1, \cdots, N-1]\}$. Then, by straight-forward calculation, we obtain the snap-shot CRB for the frequency offset vector as

$$\mathrm{CRB}(\boldsymbol{v}) = \mathrm{diag}[(\boldsymbol{F}_{11} - \boldsymbol{F}_{12}\boldsymbol{F}_{22}^{-1}\boldsymbol{F}_{21})^{-1}], \tag{55}$$

and the snap-shot CRB for the channel coefficient vector as

$$\mathrm{CRB}([\Re(\boldsymbol{h}^T), \Im(\boldsymbol{h}^T)]^T) = \mathrm{diag}[(\boldsymbol{F}_{22} - \boldsymbol{F}_{21}\boldsymbol{F}_{11}^{-1}\boldsymbol{F}_{12})^{-1}], \tag{56}$$

where

$$\begin{aligned}
\boldsymbol{F}_{11} &= \left[(\tfrac{2\pi}{N})^2 \Re(\boldsymbol{Q}^H\boldsymbol{Q}) \right] \\
\boldsymbol{F}_{12} &= \left[\tfrac{2\pi}{N}\Im(\boldsymbol{Q}^H\boldsymbol{U}) \quad \tfrac{2\pi}{N}\Re(\boldsymbol{Q}^H\boldsymbol{U}) \right] \\
\boldsymbol{F}_{21} &= \begin{bmatrix} -\frac{2\pi}{N}\Im(\boldsymbol{U}^H\boldsymbol{Q}) \\ \frac{2\pi}{N}\Re(\boldsymbol{U}^H\boldsymbol{Q}) \end{bmatrix} \\
\boldsymbol{F}_{22} &= \begin{bmatrix} \Re(\boldsymbol{U}^H\boldsymbol{U}) & -\Im(\boldsymbol{U}^H\boldsymbol{U}) \\ \Im(\boldsymbol{U}^H\boldsymbol{U}) & \Re(\boldsymbol{U}^H\boldsymbol{U}) \end{bmatrix}.
\end{aligned} \tag{57}$$

The CRBs derived in (53), (55) and (56) follow the classical CRB (page 925 in [18] or page 30 in [19]) where the expectation in (53) is taken with respect to $f(\boldsymbol{r};\boldsymbol{\theta})$ in (54). For random frequency offset and fading channel, these CRBs represent the

bounds for joint estimation of particular (snapshot) frequency offset and channel realization. Hence, the above snap-shot CRB is only for a particular realization of the parameter vector. We obtain the average CRB (ACRB) by averaging the above snap-shot CRB over the random frequency offsets and the random channel gains.

Note that the CRBs of random frequency offsets and random channel coefficients can be obtained by first averaging $\mathcal{F}$ in (53) over the random frequency offsets and random channel gains, and then taking the inverse of it. This CRB is also shown in Fig. 4 and Fig. 7 where we observe that ACRB gives a tighter bound than the CRB. In fact, ACRB of frequency offset estimation corresponds to the extended Miller and Chang bound (EMCB) and its tightness has been reported in [20] [21].

REFERENCES

[1] IEEE LAN/MAN Standards Committee, "Broadband wireless access: IEEE MAN standard," IEEE 802.16-2004, 2004.

[2] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," *Intel Technology J.*, vol. 8, no. 3, pp. 201-212, Aug. 2004.

[3] H. Minn and X. Fu, "A new ranging method for OFDMA systems," in *Proc. IEEE Globecom 2005,* vol. 3, pp. 1435-1440.

[4] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 191-193, Feb./Mar./Apr. 1995.

[5] M. Gudmundson and P. O. Anderson, "Adjacent channel interference in an OFDM system," in *Proc. IEEE VTC*, 1996, pp. 918-922.

[6] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA System," *IEEE Trans. Commun.,* no. 2, pp. 296-306, Feb. 2004.

[7] J. J.van de Beek, P. O. Borjesson, M. L. Bouchest, D. Landstram, J. M. Arenas, P. Odling, C. Ostberg, M. Wahlqvist, and S. K. Wilson, "A time and frequency synchronization scheme for multiuser OFDM," *IEEE J. Select. Areas Commun.,* vol. 17, pp. 1900-1914, Nov. 1999.

[8] Z. Cao, U. Tureli, and Y. D. Yao, "Efficient structure-based carrier frequency offset estimation for interleaved OFDMA uplink," in *Proc. IEEE ICC'03*, pp. 3361-3365.

[9] M.-O. Pun, S.-H. Tsai, and C.-C. J. Kuo, "Joint maximum likelihood estimation of carrier frequency offset and channel in uplink OFDMA systems," in *Proc. IEEE Globecom 2004,* pp. 3748-3752.

[10] M.-O. Pun, S.-H. Tsai, and C.-C. J. Kuo, "An EM-based joint maximum likelihood estimation of carrier frequency offset and channel for uplink OFDMA systems," in *Proc. IEEE VTC 2004,* pp. 598-602.

[11] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158-1168, May 2006. (Also see *Proc. IEEE Globecom 2004*, pp. 219-224.)

[12] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Processing,* pp. 2664-2677, Oct. 1994

[13] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Commun. Lett.,* vol. 3, pp. 75-77, Mar. 1999

[14] H. Minn, P. Tarasak, and V. K. Bhargava, "OFDM frequency offset estimation based on BLUE principle," in *Proc. IEEE VTC 2002,* pp. 1230-1234.

[15] IEEE LAN/MAN Standards Committee, "Channel models for fixed wireless applications," Document IEEE.802.16.3c-01/29r4.

[16] H. Minn, V. K. Bhargava, and K. B. Letaief, "A robust timing and frequency synchronization for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 822-839, July 2003.

[17] H. Minn, V. K. Bhargava, and K. B. Letaief, "A combined timing and frequency synchronization and channel estimation for OFDM," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 416-422, Mar. 2006.

[18] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Wiley-Interscience, 2002.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall PTR, 1993.

[20] F. Gini and R. Reggiannini, "On the use of Cramer-Rao-like bounds in the presence of random nuisance parameters," *IEEE Trans. Commun.,* vol. 48, no. 12, pp. 2120-2126, Dec. 2000.

[21] H. Minn, X. Fu, and V. K. Bhargava, "Optimal periodic training signal for frequency offset estimation in frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 1081-1096, June 2006.

Xiaoyu Fu (S'04) received the B.S. and M.S. degree in electrical engineering from University of Electronic Science and Technology of China, Sichuan, China, in 1999 and 2002, respectively, and the Ph.D. degree at the University of Texas at Dallas (UTD), Dallas, USA, in 2006. From 2003 to 2006, he was a Research Assistant with the Information and Communication laboratory at the Department of Electrical Engineering, UTD. Since 2007, he has been with Newport Media Inc, CA, USA, where he is currently engaged in Mobile Digital-TV receiver design and implementation.

Hlaing Minn (S'99-M'01-SM'07) received his B.E. degree in Electronics from Yangon Institute of Technology, Yangon, Myanmar, in 1995, M.Eng. degree in Telecommunications from Asian Institute of Technology (AIT), Pathumthani, Thailand, in 1997 and Ph.D. degree in Electrical Engineering from the University of Victoria, Victoria, BC, Canada, in 2001.

He was with the Telecommunications Program in AIT as a laboratory supervisor during 1998. He was a research assistant from 1999 to 2001 and a post-doctoral research fellow during 2002 in the Department of Electrical and Computer Engineering at the University of Victoria. Since September 2002, he has been with the Erik Jonsson School of Engineering and Computer Science, the University of Texas at Dallas, USA, as an Assistant Professor. His research interests include wireless communications, statistical signal processing, error control, detection, estimation, synchronization, signal design, cross-layer design, and cognitive radios. He is an Editor for the *IEEE Transactions on Communications*.

Cyrus D. Cantrell received the Bachelor's degree from Harvard in 1962, the Master's degree from Princeton University in 1964, and the Ph.D. degree from Princeton University in 1968, all in physics. Currently Dr. Cantrell is Professor of Electrical Engineering and Physics, Director of PhoTEC (the Photonic Technology and Engineering Center), and Associate Dean for Academic Affairs in the Erik Jonsson School of Engineering and Computer Science at the University of Texas at Dallas. Dr. Cantrell is a licensed Professional Engineer in the State of Texas.

Prior to joining the University of Texas at Dallas, he was a Staff Member at Los Alamos National Laboratory, working on engineering and scientific aspects of separation of isotopes using lasers. Prior to Los Alamos, he was an Assistant Professor of Physics (later Associate Professor with tenure) at Swarthmore College, Swarthmore, Pennsylvania. In 1985 he was elected a Fellow of the IEEE, and in 2000 he was awarded an IEEE Third Millennium Medal. Dr. Cantrell is also a Fellow of the Optical Society of America and the American Physical Society.

Dr. Cantrell's research career at academic institutions and Los Alamos has produced 30 Ph.D. graduates, over 100 journal, conference and book-chapter publications, and 3 U.S. patents. Dr. Cantrell's active research areas are photonics, optical networks, RF and microwaves, and electromagnetics.