

RESEARCH

Open Access

Iterative joint frequency offset and channel estimation for OFDM systems using first and second order approximation algorithms

Rainfield Y Yen¹, Hong-Yu Liu² and Chia-Sheng Tsai^{3*}

Abstract

To implement an algorithm for joint estimation of carrier frequency offset (CFO) and channel impulse response (CIR) in orthogonal frequency division multiplexing (OFDM) systems, the maximum-likelihood criterion is commonly adopted. A major difficulty arises from the highly nonlinear nature of the log-likelihood function which renders local extrema or multiple solutions for the CFO and CIR estimators. Use of an approximation method coupled with an adaptive iteration algorithm has been a popular approach to ease problem solving. The approximation used in those existing methods is usually of the first order level. Here, in addition to a new first order approximation method, we also propose a second order approximation method. Further, for the part of the adaptive iteration algorithm, we adopt a new technique which will enable performance improvement. Our first order approximation method is found to outperform the existing ones in terms of estimation accuracies, tracking range, computation complexity, and convergence speed. As expected, our second order approximation method provides an even further improvement at the expense of higher computation complication.

Keywords: Carrier frequency offset, Orthogonal Frequency Division Multiplexing (OFDM), Synchronization, Channel estimation, Maximum-likelihood estimation

1. Introduction

For orthogonal frequency division multiplexing (OFDM) communications over mobile wireless channels, after an initial acquisition is completed for time and frequency synchronization [1,2], residual or fine carrier frequency offset (CFO) and time variation of the channel impulse response (CIR) will continue to exist due to Doppler shift, multipath fading, and/or local oscillator instability [3]. Therefore, fine CFO and CIR variation must be frequently tracked [3]. A majority of the methods proposed in the literature for CFO and CIR estimation either singly addresses CFO or CIR estimation. That is, either perfect CFO synchronization is assumed in deriving CIR estimations [4-14] or perfect CIR estimation is assumed in deriving CFO synchronization algorithms [15-17]. A more practical implementation is to jointly track CFO and CIR simultaneously. The commonly adopted

scheme for tracking CFO and/or CIR is to use a maximum likelihood (ML) estimation algorithm coupled with the aid of an alternate and adaptive iteration algorithm [3,18-20]. But there exists a major difficulty, namely, the local extrema or multiple solution complication arising from the highly nonlinear nature of the log-likelihood function. A most direct and straightforward approach is the use of the steepest descent or the gradient method as given in [18]. Two hurdles in the gradient method are the starting point and the adaptive step size for the adaptive iteration process. For the algorithm to work well so that the adaptive process will converge to the global solutions, proper or careful choice must be made of the starting point as well as the value or range of the step size. Often when the environment or the channel is changed, these choices will become different and hence must be re-established by a new search. Usually, the search is by trial and error and can be a nuisance. Therefore, theoretically the gradient method may be perfectly workable, but practically it may not be so conveniently implementable. An alternative is to use

* Correspondence: s087676@gmail.com

³Department of Computer Science and Engineering, Tatung University, No.40, Sec. 3, Zhongshan N. Rd., Taipei City 104, Taiwan
Full list of author information is available at the end of the article

mathematical approximations. For example, for what is called the joint maximum likelihood channel and frequency estimation (JML-CFE) algorithm given in [19], the authors used the frequency-domain expression of the discrete Fourier transform (DFT) output at the OFDM receiver where two terms containing CFO are approximated by truncating Taylor series expansions beyond the second order terms, one being the desired carrier term and the other the inter-carrier interference (ICI) term. The resultant approximate output expression was used to form an approximate log-likelihood function. They then took partial differentiations of this approximate log-likelihood function respectively with respect to CIR and CFO to obtain a set of gradients and set them to equal zeros thus establishing a system of simultaneous equations in terms of CIR and CFO (Note that, for frequency-selective channels, the CIR is actually a multi-dimensional vector whose dimension equals the channel length). Solving the simultaneous equations leads to approximate solutions for the CFO and the CIR estimators. The log-likelihood function is a second order polynomial in CFO after the Taylor series truncation. Therefore, partial differentiation with respect to CFO will result in a first order polynomial in CFO. So the approximation is eventually of the first order level. Based on the estimator solutions as obtained above, adaptive iterations can then be performed alternately to jointly estimate data, CFO, and CIR in a decision-directed fashion. However, their solutions are very lengthy and cumbersome making the whole algorithmic process extremely tedious and complex. Another example of joint CFO and CIR estimation technique for OFDM based on iterative approximation is the expectation-maximization (EM) algorithm given in [20]. There, for what is called the expectation step, a minimum-mean-square error (MMSE) estimate of CIR conditioned on CFO was first derived from the log-likelihood function for the time-domain received signal. Then, for what is called the maximization step, conditioned on the MMSE channel estimate, an approximate expression was obtained for the log-likelihood function by truncating the Taylor series expansion of a term involving CFO beyond the second order term. Setting to zero the partial differentiation of this approximate log-likelihood function with respect to CFO yielded the CFO estimate. For the same reason as argued above, this approximation is also of the first order level. Adaptive iterations were then performed alternately between the expectation and maximization steps to produce the final CFO and CIR estimators. There are some shortcomings in the above EM-based approach. First, the MMSE channel estimate in the expectation step requires the knowledge of noise power and the channel covariance information. Second, the MMSE criterion for estimating the channel is not

geared to maximizing the log-likelihood function. Thus, it is expected that the MMSE channel estimator will not perform as well as an ML channel estimator.

In this paper, we shall propose a new first order approximation method as well as a second order approximation method for joint CFO and CIR estimation in OFDM systems. As will be seen, our first order approximation method is much succinct and easily comprehensible. Additionally, with an ingenious design of the adaptive iteration algorithm, we are able to obtain extremely fast and excellent convergence results. In terms of estimation accuracies as well as tracking range, our algorithm offers comparatively more satisfactory performance than those of existing algorithms. Indubitably and expectedly, our second order approximation algorithm provides an even further improvement in frequency tracking and channel estimation performance over the first order counterpart.

2. Signal and system model

Define $\mathbf{X} = \text{diag}\{X_0, X_1, \dots, X_{N-1}\}$ as the diagonal matrix with $\{X_k, k = 0, 1, \dots, N-1\}$ as the set of transmitted baseband frequency-domain data symbols over an OFDM symbol block of length N (in symbol units); $\mathbf{r} = [r_0, r_1, \dots, r_{N-1}]^T$ as the time-domain baseband received signal vector; and $\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]^T$ as the time-domain baseband noise vector where $\{w_n\}$ are independent, identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance σ_w^2 and can usually be denoted by $w_n \sim N(0, \sigma_w^2)$ for convenience. For a frequency-selective channel of dispersion length ν , let $\mathbf{h} = [h_0, h_1, \dots, h_{\nu-1}]^T$ be the channel impulse response (CIR) vector with $\{h_m, m = 0, 1, \dots, \nu-1\}$ being spatially uncorrelated and $\mathbf{H} = [H_0, H_1, \dots, H_{N-1}]^T$ be the channel frequency response (CFR) vector. Assuming the frequency offset normalized to subcarrier spacing is δ . Then, at the receiver, upon demodulation and discarding the cyclic prefix, the baseband sample at the n th time slot in the received time-domain OFDM block can be expressed as [3]

$$r_n = e^{j2\pi n\delta/N} y_n + w_n, \quad n = 0, 1, \dots, N-1 \quad (1)$$

where

$$y_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H_k X_k e^{j2\pi nk/N}. \quad (2)$$

Here we have adopted unitary discrete Fourier transform (DFT) for signal data. We can readily combine (1) and (2) into a vector form as [18]

$$\mathbf{r} = \frac{1}{\sqrt{N}} \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{H} + \mathbf{w} = \frac{1}{\sqrt{N}} \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_\nu \mathbf{h} + \mathbf{w}, \quad (3)$$

where

$$\mathbf{F}_L = \begin{bmatrix} 1 & 1 & \cdots & \cdots & 1 \\ 1 & e^{-j2\pi/N} & \cdots & \cdots & e^{-j2\pi(L-1)/N} \\ 1 & e^{-j4\pi/N} & \ddots & \cdots & e^{-j4\pi(L-1)/N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j2\pi(N-1)/N} & \cdots & \cdots & e^{-j2\pi(N-1)(L-1)/N} \end{bmatrix} \quad (4)$$

is an $N \times L$ matrix and

$$\mathbf{D}_\delta = \text{diag}\{1, e^{j2\pi\delta/N}, \dots, e^{j2\pi(N-1)\delta/N}\}. \quad (5)$$

From (3), a log-likelihood function can be derived as

$$\begin{aligned} \ln \Lambda &= -N \ln(\pi\sigma_w^2) - \frac{1}{\sigma_w^2} \left\| \mathbf{r} - \frac{1}{\sqrt{N}} \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h} \right\|^2 \\ &= -N \ln(\pi\sigma_w^2) - \frac{1}{\sigma_w^2} \left(\mathbf{r}^H \mathbf{r} - \frac{2}{\sqrt{N}} \text{Re}\{\mathbf{r}^H \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h}\} \right. \\ &\quad \left. + \frac{1}{N} \mathbf{h}^H \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_\delta^H \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h} \right) \\ &= -N \ln(\pi\sigma_w^2) - \frac{1}{\sigma_w^2} \left(\mathbf{r}^H \mathbf{r} - \frac{2}{\sqrt{N}} \text{Re}\{\mathbf{r}^H \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h}\} \right. \\ &\quad \left. + \mathbf{h}^H \mathbf{F}_v^H \mathbf{X}^H \mathbf{X} \mathbf{F}_v \mathbf{h} \right), \end{aligned} \quad (6)$$

where $\text{Re}\{\}$ means real part. Now, by setting $\partial \ln \Lambda / \partial \mathbf{h}^* = \mathbf{0}$, we can obtain a solution for \mathbf{h} that will render a maximum $\ln \Lambda$ for a fixed δ . This is just an ML estimate of \mathbf{h} at a fixed δ given by

$$\hat{\mathbf{h}} = \frac{1}{\sqrt{N}} (\mathbf{F}_v^H \mathbf{X}^H \mathbf{X} \mathbf{F}_v)^{-1} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_\delta^H \mathbf{r}. \quad (7)$$

Constant modulus training sequence has been proven optimal for channel estimation [21]. Chu sequence [22], for example, falls onto this category. We shall use a Chu sequence given by $\{X_k = e^{j\pi m k^2 / N}, m \text{ being any integer relatively prime to } N\}$. This results in $\mathbf{X}^H \mathbf{X} = \mathbf{I}_N$. Then, (7) can be simplified to

$$\hat{\mathbf{h}} = \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_\delta^H \mathbf{r}. \quad (8)$$

Next, setting $\frac{\partial \ln \Lambda}{\partial \delta} = 0$ leads to

$$\text{Im}\{\mathbf{r}^H \mathbf{Q} \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h}\} = 0 \quad (9)$$

where $\mathbf{Q} = \text{diag}\{0, 1, \dots, N-1\}$ and $\text{Im}\{\}$ means imaginary part. Replacing the \mathbf{h} of (9) by the $\hat{\mathbf{h}}$ of (8), we find

$$\text{Im}\{\mathbf{r}^H \mathbf{Q} \mathbf{D}_\delta \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_\delta^H \mathbf{r}\} = \text{Im}\{\mathbf{r}^H \mathbf{D}_\delta \mathbf{G} \mathbf{D}_\delta^H \mathbf{r}\}$$

$$= \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im}\{r_m^* r_n g_{m,n} e^{j2\pi\delta(m-n)/N}\} = 0, \quad (10)$$

where $g_{m,n}$ is the (m,n) th element of an $N \times N$ matrix \mathbf{G} given by

$$\begin{aligned} \mathbf{G} &= \mathbf{Q} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N = \{g_{m,n}\} \\ &= \begin{bmatrix} 0 & & & & \\ \sum_{l=0}^{v-1} \sum_{k=0}^{N-1} X_k e^{-j2\pi(l-1)k/N} \sum_{k'=0}^{N-1} X_{k'}^* e^{j2\pi l k' / N} & & & & \\ & \vdots & & & \\ (N-1) \sum_{l=0}^{v-1} \sum_{k=0}^{N-1} X_k e^{-j2\pi(l-N+1)k/N} \sum_{k'=0}^{N-1} X_{k'}^* e^{j2\pi l k' / N} & & & & \\ & & 0 & & \\ \sum_{l=0}^{v-1} \sum_{k=0}^{N-1} X_k e^{-j2\pi(l-1)k/N} \sum_{k'=0}^{N-1} X_{k'}^* e^{j2\pi(l-1)k' / N} & & & & \\ & \vdots & & & \\ (N-1) \sum_{l=0}^{v-1} \sum_{k=0}^{N-1} X_k e^{-j2\pi(l-N+1)k/N} \sum_{k'=0}^{N-1} X_{k'}^* e^{j2\pi(l-N+1)k' / N} & & & & \\ \cdots & & & & \\ \cdots & \sum_{l=0}^{v-1} \sum_{k=0}^{N-1} X_k e^{-j2\pi(l-1)k/N} \sum_{k'=0}^{N-1} X_{k'}^* e^{j2\pi(l-1)k' / N} & & & \\ \cdots & \vdots & & & \\ \cdots & \sum_{l=0}^{v-1} \sum_{k=0}^{N-1} X_k e^{-j2\pi(l-N+1)k/N} \sum_{k'=0}^{N-1} X_{k'}^* e^{j2\pi(l-N+1)k' / N} & & & \\ \cdots & & & & \end{bmatrix}. \end{aligned} \quad (11)$$

Since (8) is now channel independent, we have decoupled δ from \mathbf{h} and (10) can thus be solved for δ alone. However, (10) is highly nonlinear in δ and contains infinite number of solutions. We only desire the one solution that yields the global maximum of $\ln \Lambda$ at a given \mathbf{h} . The task is not possible by analytical means. The most direct and straightforward approach is the steepest descent or the gradient method which has been treated in [18]. But the problem with the gradient method lies in the proper selection of the starting point and the adaptive step size of the adaptive iteration process. An improper selection of the starting point or the step size or both may either lead to an undesired solution (a local extremum) or result in process divergence. Moreover, when the environment or the channel is changed, these choices will also be changed and must be reestablished by a new search. Often such a search is by trial and error and can be a nuisance. For the above reasons, the joint frequency tracking and channel estimation using the gradient method may not be practically implementable. An alternative is to simplify (10) by approximation so that (11) can be reduced to an

analytically solvable form. We shall present two approximation methods.

2.1 First order approximation

Assuming δ is small, we can approximate $e^{j2\pi(m-n)\delta/N}$ by $1 + j2\pi(m-n)\delta/N$. We call this the first order approximation since we have retained only the terms up to the first order. Then, (10) becomes

$$\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \left\{ r_m^* r_n g_{m,n} e^{j2\pi(m-n)\delta/N} \right\} \quad (12)$$

$$\approx \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \left\{ r_m^* r_n g_{m,n} [1 + j2\pi(m-n)\delta/N] \right\} \approx 0.$$

Now, we can easily solve (12) for δ to get the CFO estimator as

$$\hat{\delta} = -\frac{N \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \{ r_m^* r_n g_{m,n} \}}{2\pi \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Re} \{ (m-n) r_m^* r_n g_{m,n} \}} = -\frac{N \text{Im} \{ \mathbf{r}^H \mathbf{G} \mathbf{r} \}}{2\pi \text{Re} \{ \mathbf{r}^H \mathbf{F} \mathbf{r} \}}, \quad (13)$$

where $\mathbf{F} = \mathbf{Q}\mathbf{G} - \mathbf{G}\mathbf{Q}$. But (13) is only an approximate solution and hence is not very accurate. We can improve the accuracy by iterative corrections as follows. Start with iteration 1 by labeling $\hat{\delta}$ and \mathbf{r} in (13) as $\hat{\delta}_1$ and \mathbf{r}_1 . We then make the correction $\mathbf{r}_2 = \mathbf{D}_{\hat{\delta}_1}^H \mathbf{r}_1$. After this correction, a residual CFO is assumed to remain and is estimated by using (13) again. Label this estimate as $\hat{\delta}_2 = -\frac{N \text{Im} \{ \mathbf{r}_2^H \mathbf{G} \mathbf{r}_2 \}}{2\pi \text{Re} \{ \mathbf{r}_2^H \mathbf{F} \mathbf{r}_2 \}}$. Next, a new correction is made as $\mathbf{r}_3 = \mathbf{D}_{\hat{\delta}_2}^H \mathbf{r}_2$. This corrective process is iteratively continued until eventually the residual CFO approaches zero. The final CFO estimator is obtained by summing all the residual CFO estimates.

When noise is large or when signal-to-noise ratio (SNR) is small, it is found that the above iterative process may become divergent. To remedy, we may introduce a step size to avoid large corrective steps. The algorithm can therefore be summarized as follows:

$$i = 1, 2, \dots, L$$

(for a total of L iterations)

$$\hat{\delta}_i = \frac{\lambda N \text{Im} \{ \mathbf{r}_i^H \mathbf{G} \mathbf{r}_i \}}{2\pi \text{Re} \{ \mathbf{r}_i^H \mathbf{F} \mathbf{r}_i \}}, \quad (14a)$$

where λ is a step size.

$$\mathbf{r}_{i+1} = \mathbf{D}_{\hat{\delta}_i}^H \mathbf{r}_i. \quad (14b)$$

Eventually at an L th iteration, we will have $\hat{\delta}_L \rightarrow 0$. Then, CFO estimate

$$\hat{\delta} = \sum_{i=1}^L \hat{\delta}_i, \quad (14c)$$

CIR estimate

$$\hat{\mathbf{h}} = \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_{\hat{\delta}}^H \mathbf{r} = \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_{\hat{\delta}_L}^H \dots \mathbf{D}_{\hat{\delta}_2}^H \mathbf{D}_{\hat{\delta}_1}^H \mathbf{r}_1$$

$$= \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{r}_L. \quad (14d)$$

Some remarks of interest are in order.

- 1) Unlike the gradient method where the step size needs be carefully searched and usually varies with changing environment, here λ is easily picked. Under moderate conditions when the initial CFO is small, we can simply choose $\lambda = 1$. For stringent conditions, e.g., when initial CFO is large, we can choose $\lambda = 0.5$ or $\lambda = 0.2$.
- 2) We wish to point out that it is the corrective action of (14b) that makes our algorithm better. Normally, one would substitute the new CFO estimate obtained from (13) in each iteration into (8) to find a new $\hat{\mathbf{h}}$ without the corrective action of (14b). This new $\hat{\mathbf{h}}$ is substituted into (9) to obtain a renewed gradient of (10) and then a renewed solution of CFO estimate given by (13) is obtained, and on and on. This latter approach is also what is used in [19] and [20] for their adaptive iteration algorithms. Computer simulations have been performed and results show that the algorithm given by (14a) ~ (14d) indeed performs much better than the latter approach.
- 3) Equation (8) or (14d) is an ML channel estimator which is optimal in the sense that it will yield a maximum likelihood function for a given CFO. As noted earlier, we expect the ML channel estimator will outperform the MMSE channel estimator adopted in [20]. Simulation results indeed confirm this fact.

As argued earlier, the CFO estimator of (13) or (14a) is an approximation of the first order level. It is possible to find a CFO estimator using second order approximation as to be presented next.

2.2 Second order approximation

By retaining terms up to the second order, we get the second order approximation given as $e^{j2\pi(m-n)\delta/N} \approx 1 + j2\pi(m-n)\delta/N - 2\pi^2(m-n)^2\delta^2/N^2$. Then, (10) can be approximated as

$$\begin{aligned} & \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \left\{ r_m^* r_n g_{m,n} e^{j2\pi\delta(m-n)/N} \right\} \\ & \approx \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \left\{ r_m^* r_n g_{m,n} [1 + j2\pi(m-n)\delta/N \right. \\ & \quad \left. - 2\pi^2(m-n)^2\delta^2/N^2] \right\} \approx 0. \end{aligned} \quad (15)$$

Solving (15), we get

$$\delta = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \quad (16)$$

where

$$A = \frac{2\pi^2}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \left\{ r_m^* r_n g_{m,n} (m-n)^2 \right\} = \frac{2\pi^2}{N^2} \text{Im} \{ \mathbf{r}^H \mathbf{E} \mathbf{r} \}, \quad (17a)$$

$$B = -\frac{2\pi}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Re} \left\{ r_m^* r_n g_{m,n} (m-n) \right\} = -\frac{2\pi}{N} \text{Re} \{ \mathbf{r}^H \mathbf{F} \mathbf{r} \}, \quad (17b)$$

$$C = -\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \text{Im} \left\{ r_m^* r_n g_{m,n} \right\} = -\text{Im} \{ \mathbf{r}^H \mathbf{G} \mathbf{r} \}, \quad (17c)$$

with $\mathbf{E} = \mathbf{Q}\mathbf{F} - \mathbf{F}\mathbf{Q} = \mathbf{Q}^2\mathbf{G} - 2\mathbf{Q}\mathbf{G}\mathbf{Q} + \mathbf{G}\mathbf{Q}^2$.

As before, the approximate solution of (16) can attain better accuracy by iterative corrections. We summarize the adaptive iteration algorithm below.

$$i = 1, 2, \dots, L$$

(for a total of L iterations)

$$\begin{aligned} \hat{\delta}_{i,1} &= \frac{-B_i + \sqrt{B_i^2 - 4A_i C_i}}{2A_i}, \\ \hat{\delta}_{i,2} &= \frac{-B_i - \sqrt{B_i^2 - 4A_i C_i}}{2A_i}. \end{aligned} \quad (18a)$$

Choose between $\hat{\delta}_{i,1}$ and $\hat{\delta}_{i,2}$ the one that will give a larger log-likelihood function, or equivalently,

$$\begin{aligned} \hat{\delta}_i &= \arg \min_{\delta_{i,p}} \left\| \mathbf{r}_i - \frac{1}{N^2} \mathbf{D}_{\delta_{i,p}} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_{\delta_{i,p}}^H \mathbf{r}_i \right\|^2, \\ p &= 1, 2, \end{aligned} \quad (18b)$$

$$\mathbf{r}_{i+1} = \mathbf{D}_{\delta_i}^H \mathbf{r}_i, \quad (18c)$$

where

$$A_i = \frac{2\pi^2}{N^2} \text{Im} \{ \mathbf{r}_i^H \mathbf{E} \mathbf{r}_i \}, \quad (18d)$$

$$B_i = -\frac{2\pi}{N} \text{Re} \{ \mathbf{r}_i^H \mathbf{F} \mathbf{r}_i \}, \quad (18e)$$

$$C_i = -\text{Im} \{ \mathbf{r}_i^H \mathbf{G} \mathbf{r}_i \}, \quad (18f)$$

and

$$\begin{aligned} & \frac{1}{N^2} \mathbf{D}_{\delta_{i,p}} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_{\delta_{i,p}}^H \mathbf{r}_i \\ &= \frac{1}{\sqrt{N}} \mathbf{D}_{\delta_{i,p}} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \hat{\mathbf{h}}_i, \end{aligned} \quad (18g)$$

with

$$\hat{\mathbf{h}}_i = \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_{\delta_{i,p}}^H \mathbf{r}_i. \quad (18h)$$

Finally, at an L th iteration, we shall have $\hat{\delta}_L \rightarrow 0$. Then, CFO estimate

$$\hat{\delta} = \sum_{i=1}^L \hat{\delta}_i, \quad (18i)$$

CIR estimate

$$\hat{\mathbf{h}} = \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{D}_{\hat{\delta}}^H \mathbf{r} = \frac{1}{N\sqrt{N}} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{r}_L. \quad (18j)$$

For this second order approximation algorithm, no step size is needed as no divergence problem will occur in the small SNR region.

Simulation results show that both the above first order and second order approximation algorithms have faster convergence rate as well as better estimation accuracies than other existing algorithms including the gradient algorithm. Detailed performance comparisons will later be presented between prevalent existing algorithms and our algorithms.

3. Cramer-Rao bounds

For the current case of joint CFO and CIR estimation in OFDM, we can readily use a result in [3] to find the Fisher information matrix as

$$\mathbf{J} = \frac{2}{\sigma_w^2} \text{Re} \left\{ \begin{array}{l} \left[\begin{array}{ccc} \mathbf{F}_v^H \mathbf{X}^H \mathbf{X} \mathbf{F}_v & j\mathbf{F}_v^H \mathbf{X}^H \mathbf{X} \mathbf{F}_v & \frac{j2\pi}{N^2} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{Q} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h} \\ -j\mathbf{F}_v^H \mathbf{X}^H \mathbf{X} \mathbf{F}_v & \mathbf{F}_v^H \mathbf{X}^H \mathbf{X} \mathbf{F}_v & \frac{2\pi}{N^2} \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{Q} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h} \\ \frac{j2\pi}{N^2} \mathbf{h}^H \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{Q} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v & -\frac{2\pi}{N^2} \mathbf{h}^H \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{Q} \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v & \frac{4\pi^2}{N^3} \mathbf{h}^H \mathbf{F}_v^H \mathbf{X}^H \mathbf{F}_N \mathbf{Q}^2 \mathbf{F}_N^H \mathbf{X} \mathbf{F}_v \mathbf{h} \end{array} \right] \end{array} \right\}. \quad (19)$$

Then the CRBs can be obtained from the diagonal elements of \mathbf{J}^{-1} . Note that the CRBs obtained here are for the ML estimates of CIR \mathbf{h} and CFO δ (i.e., $\hat{\mathbf{h}}$ and $\hat{\delta}_{LC}$).

4. Simulation results

We shall first use a static channel model. That is to assume channel fading remains unchanged over a great many OFDM blocks (slow quasi-static fading). This implies that the maximum Doppler frequency must satisfy $f_M = v f_c / c \ll 1/T$ corresponding to a mobile speed $v \ll c/(f_c T)$, where c is the speed of light, T is one OFDM block length in seconds, and f_c is the carrier frequency in Hz. Taking an 802.11a standard with $f_c = 5$ GHz and $\Delta f = 1/T = 312.5$ kHz, this requires $v \ll 67,500$ km/hr. Apparently, this requirement for slow quasi-static fading is easily met in practice. For example, if $v = 60$ km/hr, then $f_M \approx 10^{-3}/T$, it is thus reasonable to assume fading to remain unchanged over several hundred OFDM blocks.

Consider a frequency-selective channel of length $\nu=9$ having an exponential power profile with unit energy. Taking $N=64$ for the OFDM system and assuming an CFO of $\delta = 0.02$, then for several algorithms, Figures 1 and 2 respectively show plots of CFO and CIR estimator variance or mean square error (MSE) vs. SNR. The CFO estimator MSE is defined as $E\left[\left(\hat{\delta} - \delta\right)^2\right]$ with $E[\]$ denoting expectation, while the CIR estimator MSE is defined by $E\left[\|\hat{\mathbf{h}} - \mathbf{h}\|^2\right]/\nu$. Since both the Chu sequence and the channel are of unit power, the SNR is simply given by $1/\sigma_w^2$. Under the moderate condition of $\delta = 0.02$, we can simply use a step size of $\lambda=1$ (the value of λ will not be explicitly indicated in plots whenever it is unity) for our first order approximation algorithm, meanwhile we find that $L=1$ will suffice. Note that $L=1$ means that results are obtained in a single iteration. On the other hand, the gradient method requires about 15 to 20 iterations using the starting point at the origin and a step size which is found to be 10^{-6} . The EM-based algorithm is found to require about 20 to 25 iterations. For performance evaluations, we have incorporated corresponding CRB vs. SNR curves in both figures.

Apparently from the figures, the gradient method and our second order method have the best performance. Our first order method and JML-CFE come second with the former slightly better in that it outperforms the latter in channel estimation. However, if we had used $L=3$ for the former, the performance would be much improved to almost equal those of the gradient method and the second order algorithm. The EM-based algorithm shows the worst performance in CFO estimation among all, but it performs better than JML-CFE in CIR estimation. The reason that our approximation algorithms can outperform the EM-based and the JML-CFE algorithms is mainly because of the iterative correction action of (14b) or (18c). The iterative corrections greatly help refine the solution to better accuracy. The refining improvement of (18c) gets more pronounced with the second order approximation process as a second order approximation is certainly a better approximation than the first order approximation. The gradient method of [18] performs comparably to our approximation algorithms. However, as stated earlier, the gradient method

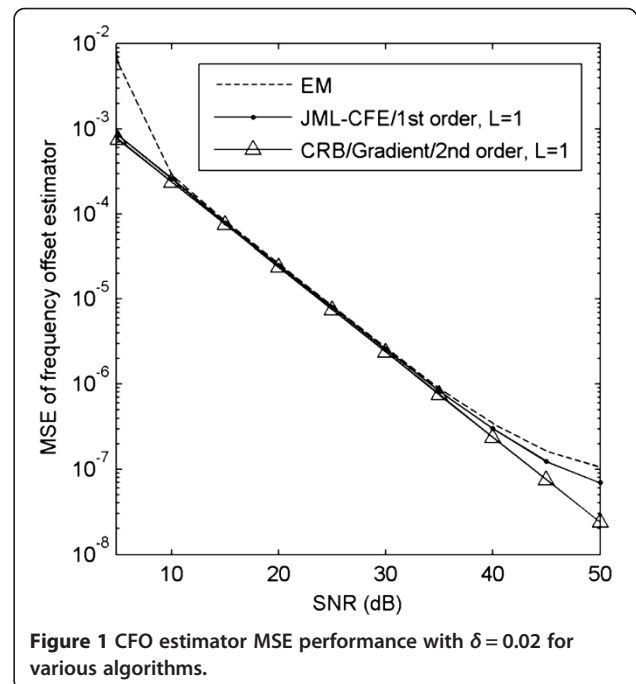
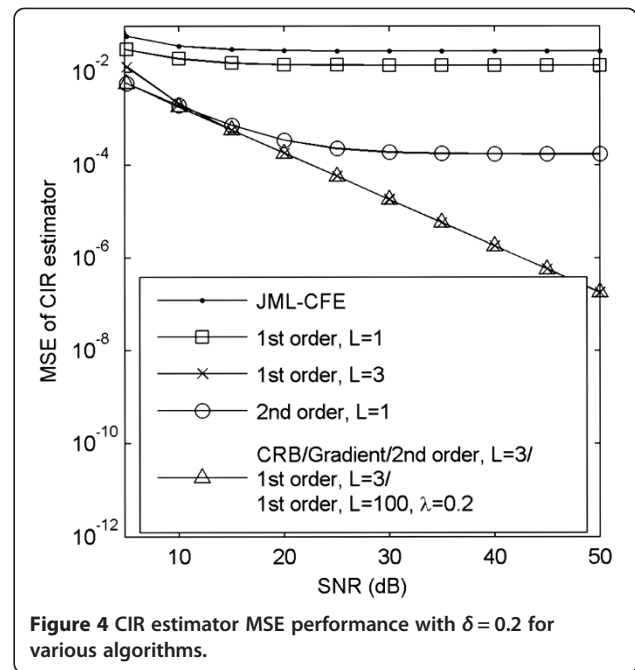
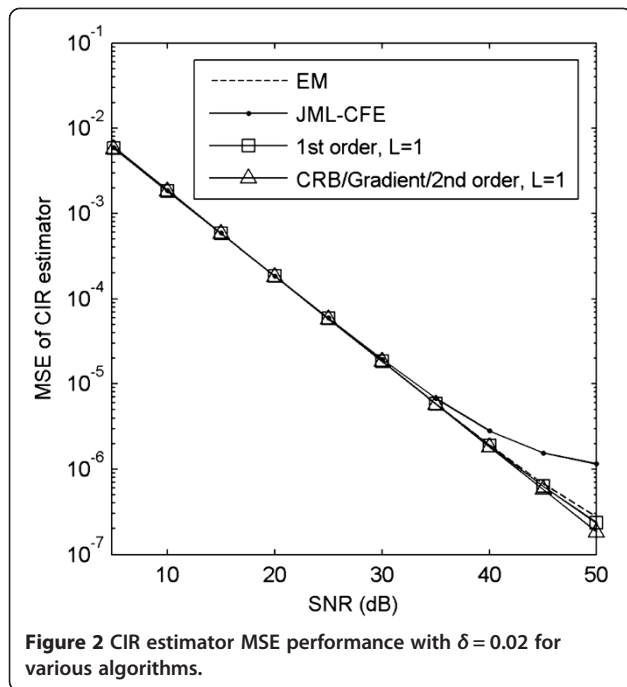
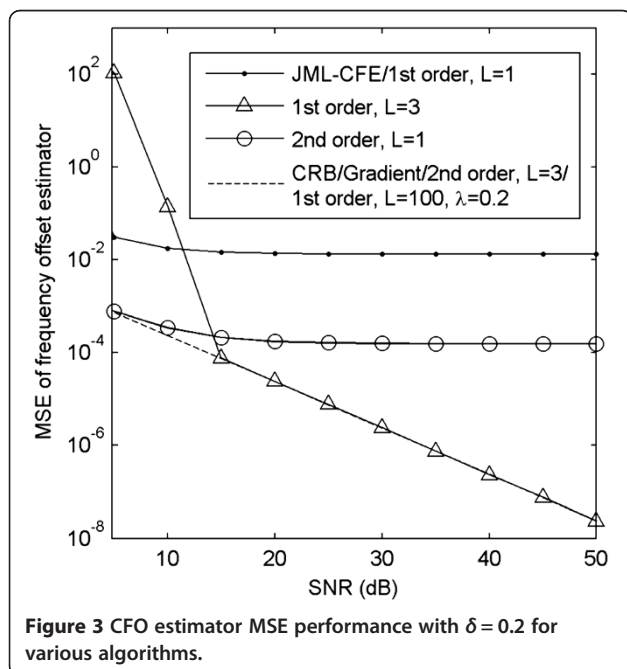


Figure 1 CFO estimator MSE performance with $\delta = 0.02$ for various algorithms.



has the shortcoming that the starting point and step size are not easily searched hence making practical implementation difficult.

We next increase the CFO to $\delta = 0.2$. Figures 3 and 4 show the corresponding performance curves for the same various algorithms. Here, the gradient method is found to work fine as well. Now with the increased CFO, the first order algorithms begin to perform poorly. As shown in the figures, the JML-CFE algorithm and



our first order method with $L=1$ and $\lambda=1$ yield unacceptable performance results. The EM-based algorithm totally fails to perform and therefore is out of the picture. However, it is to be noted that, for a given δ , three parameters can affect or determine our algorithm performance. They are: 1) approximation order, 2) number of process iterations L , and 3) adjustment step size λ . Therefore, by increasing the number of iterations and adjusting the step size, our first order algorithm can be made to work with acceptable results. For instance, if using $\lambda = 0.2$ with 100 iterations, the first order method can give a performance comparable to the second order algorithm with $L=3$ as well as to the gradient algorithm. For our second order algorithm, no adjustment step size needs be introduced as mentioned earlier (i.e., $\lambda = 1$), but the number of iterations must be at least $L=3$. An explanation is warranted for the first order performance in the low SNR region. As can be observed in Figure 3, the first order algorithm with $L=1$ is seen to do better than $L=3$ at low SNR. The reason for this odd phenomenon is because, when noise is high, the initial estimate given by (14a) may be in error. Therefore, further iterations from this erroneous point will direct the search to a wrong direction and lead to even worse estimates. Nonetheless, a remedy can be made by use of the adjustment step size of λ . As can be seen in Figure 3, when using a $\lambda = 0.2$ along with sufficient iterations $L=100$, the first order algorithm can achieve a comparable performance as that of a second order with $L=3$.

For the gradient method, we have used the origin as the starting point which works fine thus far. However,

when the CFO exceeds half the subcarrier spacing, the difficulty of choosing the starting point as mentioned earlier starts to show up. Using the origin as the starting point at large CFO values will not work so fine anymore and, in fact, the gradient method will usually fail to converge to the global solution. In Figures 5 and 6, performance curves of the gradient method and our 2nd order approximation method are given for $\delta = 1$. Here, the gradient method fails to deliver acceptable performance, nor the second order algorithm with $L=2$ and 3. But when increasing the iteration number to $L=4$, the second order algorithm can perform satisfactorily if SNR exceeds 20dB. Again, the same odd phenomenon for the first order case at low SNR also occurs to the second order case due to a poor initial estimate given by (18b). We note here that, for fine frequency synchronization, CFO is usually smaller than 0.5 [18-20]. Nonetheless, occasional large CFO that exceeds 0.5 does occur. This is when our proposed method becomes advantageous and useful. Furthermore, besides its major shortcoming on the choice of starting point and step size, the gradient method has a slower convergence rate than our second order algorithm which needs only a single digit iteration number ($L \leq 4$ in all cases of simulations given above) as against a two digit iteration number for the gradient method.

Numerous simulation tests indicate that, roughly, the EM-based algorithm has a tracking range less than ± 0.05 of the subcarrier spacing and both the JML-CFE algorithm and our first order algorithm with $L=1$ have a tracking range about ± 0.1 of the subcarrier spacing. Our

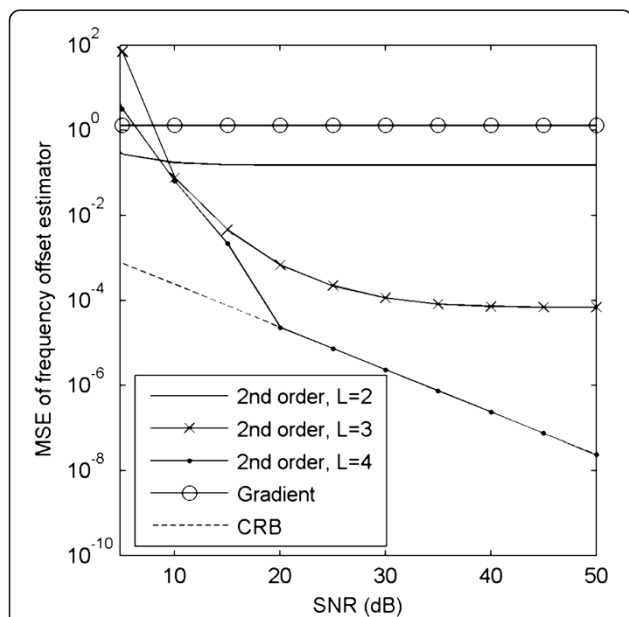


Figure 5 CFO estimator MSE performance with $\delta = 1$ for second order approximation and gradient method.

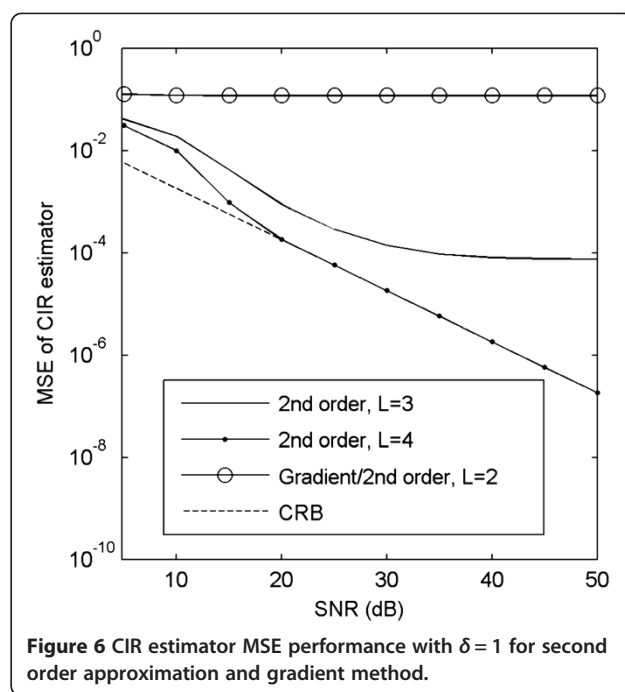


Figure 6 CIR estimator MSE performance with $\delta = 1$ for second order approximation and gradient method.

first order algorithm with $L=3$ can achieve a tracking range about ± 0.2 of the subcarrier spacing and it is found that further increasing L will not help the matter better. The gradient method can cover the range of half the subcarrier spacing. For our second order algorithm, when using $L=4$, the tracking range can go up to $[-1, 1]$ of the subcarrier spacing.

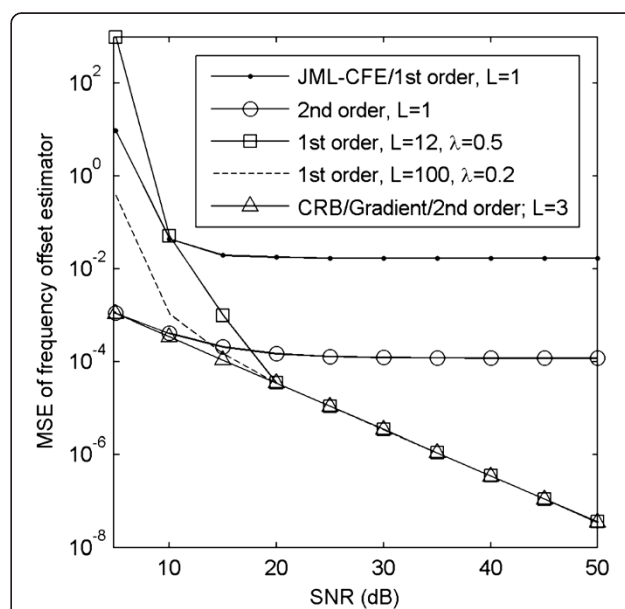
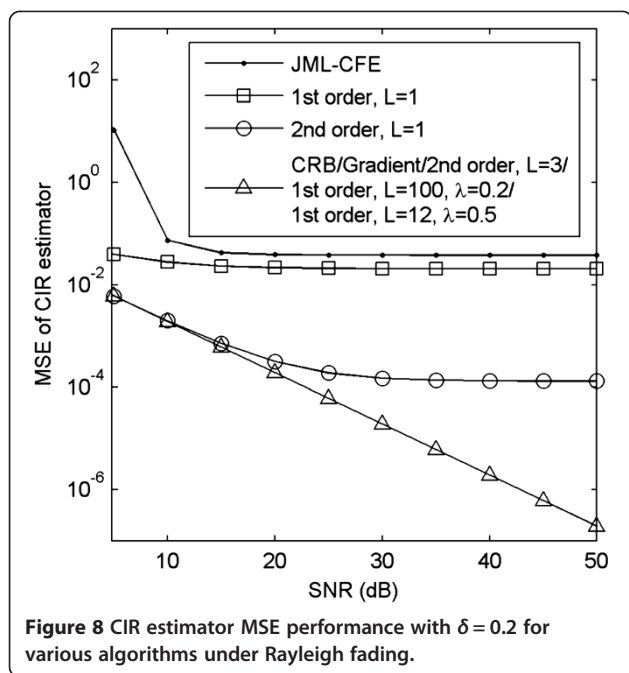


Figure 7 CFO estimator MSE performance with $\delta = 0.2$ for various algorithms under Rayleigh fading.



Over a long term period when fading is to be considered, simulations for a frequency-selective Rayleigh fading channel show more or less similar performance behaviors as for the static channel case. The impact of fading on the estimate performance for our first and second order algorithms includes reduced tracking range and the need for more iterations. Figures 7 and 8 present the performances of various algorithms under Rayleigh fading with $\delta = 0.2$. Here again, the EM-based algorithm is found to yield unsatisfactory results and hence its performance is not shown in the figure. In view of Figures 7 and 8, we see that, for the same value $L=1$, the second order algorithm outperforms the first order algorithm as it should. However, when the iteration number is increased sufficiently with the aid of an adjustment step size, the first order algorithm can be enhanced to outperform a second order algorithm with a low iteration number. This can be observed in the figures for the cases of first order algorithm with $L=12, \lambda = 0.5$ and $L=100, \lambda = 0.2$ as both these first order cases outperform

the second order with $L=1$. However, as shown in the figures, the powerful second order algorithm with only $L=3$ can easily compete with the above two cases of first order algorithm.

Finally, computation complexity is compared for various algorithms. Table 1 displays the arithmetic operations needed in terms of real multiplication, addition, division, and square roots. One complex multiplication is equivalent to four real multiplications. One complex addition is equivalent to two real additions. Some composite data matrixes can be calculated in advanced and stored in memory before performing estimation algorithm. \mathbf{G} in (11), for example, can be pre-computed and stored in memory for later use in adaptive iterations. The needed arithmetic operations displayed in Table 1 exclude such situations. Moreover, L represents the number of iterative adaption needed.

As an example, we take $L=3$ for both the 1st and the 2nd order approximations, $L=20$ for the EM algorithm, $L=15$ for the gradient algorithm, and $\nu = N/4$. Then, the EM algorithm requires $40N^2 + 260N - 20$ real multiplications, the JML-CFE needs $27N^2 + 16N$ real multiplications, and the gradient method takes $61N^2 + 80N + 1$ real multiplications. In comparisons, the 1st order approximation requires the least computational complexity of $25N^2 + 39N - 3$ real multiplications, while the 2nd order approximation needs the most computation complication of $85N^2 + 78N + 6$ real multiplications.

5. Conclusion

We show that, for ML-based joint CFO and CIR estimation in OFDM systems, a proper approximation simplification of the log-likelihood function coupled with the aid of a good iterative correction algorithm can result in satisfactory estimation performance. We first present a first order approximation algorithm and compare it with prevalent existing algorithms (also of first order). Our first order algorithm is found to be superior to others with faster convergence, wider tracking range, least computation requirement, and estimation accuracies. Then a second order approximation algorithm is presented with further performance improvement with higher computational cost.

Table 1 Computation complexity

	Real multiplication	Real addition	Real division	Square roots
EM	$L(8vN + 13N - 1)$	$2L(2vN + N - v - 2)$	L	0
JML-CFE	$2N(13N + 2v + 8)$	$12N^2 + 2vN - 3N - 2v - 6$	1	0
Gradient	$LN(4N + 5) + N(4v + 5) - 1$	$L(2N^2 - 1) + 2v(N - 1)$	0	0
The 1st order algorithm	$L(8N^2 + 13N - 1) + 4vN$	$L(4N^2 - 3) + 2v(N - 1) - 1$	L	0
The 2nd order algorithm	$2L(14N^2 + 13N + 1) + 4vN$	$6LN(N + 1) - 3L - 1$	$2L$	$2L$

Abbreviations

CFO: Carrier frequency offset; CIR: Channel impulse response; OFDM: Orthogonal frequency division multiplexing; DFT: Discrete Fourier transform; ICI: Inter-carrier interference; EM: Expectation-maximization; MMSE: Minimum-mean-square error; i.i.d.: Identically distributed; RVs: Random variables; CFR: Channel frequency response; SNR: Signal-to-noise ratio.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Electrical Engineering, Tamkang University, No.151, Yingzhan Rd., Tamsui Dist., New Taipei City 25137, Taiwan. ²Department of Electrical Engineering, Fu Jen Catholic University, No.510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 24205, Taiwan. ³Department of Computer Science and Engineering, Tatung University, No.40, Sec. 3, Zhongshan N. Rd., Taipei City 104, Taiwan.

Received: 9 June 2012 Accepted: 29 October 2012

Published: 16 November 2012

References

1. TM Schmidl, DC Cox, Robust frequency and timing synchronization for OFDM. *IEEE Trans. Commun.* **45**(12), 1613–1621 (1997)
2. JJ van de Beek, M Sandell, PO Borjesson, ML estimation of time and frequency offset in OFDM systems. *IEEE Trans. Signal Process.* **45**(7), 1800–1805 (1997)
3. HY Liu, RY Yen, Effective adaptive iteration algorithm for frequency tracking and channel estimation in OFDM systems. *IEEE Trans. Veh. Technol.* **59**(4), 2093–2097 (2010)
4. JJ van de Beek, O Edfors, M Sandell, SK Wilson, PO Borjesson, On channel estimation in OFDM systems, in *Proceedings of the IEEE 45th Vehicular Technology Conference, Chicago, IL*, vol. 2, 1995, pp. 815–819
5. O Edfors, M Sandell, JJ van de Beek, SK Wilson, PO Borjesson, OFDM channel estimation by singular value decomposition. *IEEE Trans. Commun.* **46**(7), 931–939 (1998)
6. Y Li, N Seshadri, S Ariyavisitakul, Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels. *IEEE J. Sel. Areas Commun.* **17**(3), 461–471 (1999)
7. Y Li, Pilot-symbol-aided channel estimation for OFDM in wireless systems. *IEEE Trans. Veh. Technol.* **49**(4), 1207–1215 (2000)
8. TL Tung, K Yao, RE Hudson, Channel estimation and adaptive power allocation for performance and capacity improvement of multiple-antenna OFDM systems, in *Proceedings of the IEEE SPAWC'01, Taoyuan, Taiwan*, 2001, pp. 82–85
9. Y Li, Simplified channel estimation for OFDM systems with multiple transmit antennas. *IEEE Trans. Wireless Commun.* **1**(1), 67–75 (2002)
10. Z Cheng, D Dahlhaus, Time versus frequency domain channel estimation for OFDM systems with antenna arrays, in *Proceedings of the IEEE 6th International Conference on Signal Processing, Beijing, China*, vol. 2, 2002, pp. 1340–1343
11. Y Li, JH Winters, NR Sollenberger, MIMO-OFDM for wireless communications: signal detection with enhanced channel estimation. *IEEE Trans. Commun.* **50**(9), 1471–1477 (2002)
12. S Coleri, M Ergen, A Puri, A Bahai, Channel estimation techniques based on pilot arrangement in OFDM systems. *IEEE Trans. Broadcast.* **48**(3), 223–229 (2002)
13. X Wang, Y Wu, JY Chouinard, Modified channel estimation algorithms for OFDM systems with reduced complexity, in *Proceedings of the IEEE 7th International Conference on Signal Processing, Beijing, China*, vol. 2, 2004, pp. 1747–1751
14. MZA Khan, Low-complexity ML channel estimation schemes for OFDM, in *Proceedings of the IEEE 13th International Conference on Network, Kuala Lumpur, Malaysia*, vol. 2, 2005, pp. 607–612
15. KW Kang, J Ann, HS Lee, Decision-directed maximum-likelihood estimation of frame synchronization offset. *Electron. Lett.* **30**(25), 2153–2154 (1994)
16. PH Moose, A technique for orthogonal frequency division multiplexing frequency offset correction. *IEEE Trans. Commun.* **42**(10), 2908–2914 (1994)
17. T Keller, L Piazzi, P Mandarini, L Hanzo, Orthogonal frequency division multiplex synchronization techniques for frequency-selective fading channels. *IEEE J. Sel. Areas Commun.* **19**(6), 999–1008 (2001)

18. X Ma, H Kobayashi, SC Schwartz, Joint frequency offset and channel estimation for OFDM, in *Proceedings of the IEEE GLOBECOM 2003, San Francisco, USA*, vol. 1, 2003, pp. 15–19
19. FZ Merli, GM Vitetta, Iterative ML-based estimation of carrier frequency offset, channel impulse response and data in OFDM transmissions. *IEEE Trans. Commun.* **56**(3), 497–506 (2008)
20. JH Lee, JC Han, SC Kim, Joint carrier frequency synchronization and channel estimation for OFDM systems via the EM algorithm. *IEEE Trans. Veh. Technol.* **55**(1), 167–172 (2006)
21. P Stoica, O Besson, Training sequence design for frequency offset and frequency-selective channel estimation. *IEEE Trans. Commun.* **51**(11), 1910–1917 (2003)
22. DC Chu, Polyphase codes with good periodic correlation properties. *IEEE Trans. Inform. Theory.* **18**(4), 531–532 (1972)

doi:10.1186/1687-1499-2012-341

Cite this article as: Yen et al.: Iterative joint frequency offset and channel estimation for OFDM systems using first and second order approximation algorithms. *EURASIP Journal on Wireless Communications and Networking* 2012 **2012**:341.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com