

Maximum-Likelihood Carrier Frequency Offset Estimation for OFDM Systems Over Frequency-Selective Fading Channels

Tao Cui and Chintha Tellambura
 Department of Electrical and Computer Engineering
 University of Alberta
 Edmonton, AB, Canada T6G 2V4
 Email: {taocui, chintha}@ece.ualberta.ca

Abstract—This paper considers carrier frequency offset (CFO) estimation for OFDM systems over frequency-selective fading channels. We derive three new maximum-likelihood (ML) CFO estimators. The first estimator takes the pre-DFT signal at the receiver as Gaussian and averages the likelihood function over the Gaussian variable. The other two estimators first remove the channel impulse response (CIR) and average the resulting likelihood function over time-domain transmitted symbols. We demonstrate that the presence of virtual carriers is critically important for estimation. All our proposed estimators can achieve the asymptotic Cramér Rao Bound (ACRB) and can also be enhanced using embedded pilots.

I. INTRODUCTION

The considerable interest in Orthogonal Frequency-Division Multiplexing (OFDM), a bandwidth efficient signalling scheme for wireless communication, is mainly due to its remarkable resistance to frequency-selective fading and impulsive noise [1]. In OFDM, the available bandwidth is sliced into several parallel narrow subchannels so that each subchannel is only subject to flat fading, which simplifies equalization to mitigate frequency-selective fading. These features have made OFDM a standard for high-speed mobile communications. OFDM has been used in Europe for Digital Audio Broadcasting, ETSI-BRAN High Performance Local Area Networks (Hiperlan/2) and IEEE 802.11a. Additionally, OFDM forms the basis for the fourth-generation broad-band systems that will transmit multimedia to mobiles and portable personal communications.

While OFDM systems are robust to frequency selective fading, they are more sensitive to synchronization errors than single-carrier systems [2]. In particular, the presence of carrier frequency offset (CFO) introduces inter-carrier-interference (ICI) that would significantly degrade the system performance [2]. To mitigate this effect, various CFO estimation methods have been proposed. In [3], repeated symbols are used for frequency synchronization. The signal redundancy in the cyclic prefix (CP) is exploited for CFO estimation in [4]. In [5], a blind estimation method is developed but some restrictions on CFO values and carrier spacing are imposed. In [6], Liu *et al.* exploit the presence of virtual carriers in

OFDM and propose blind estimation methods reminiscent of spectral analysis techniques in array processing, i.e., MUSIC and ESPRIT. ML estimators are derived in [7]–[9]. In [8], [9], the CIR is first removed from the likelihood function and the frequency domain transmitted signals are removed by summing over all the possible signals.

In this paper, we derive three new ML estimators (MLE) for CFO estimation. The CIR and the transmitted symbols together are modelled as a Gaussian process for the first MLE, and they are removed by averaging the likelihood function over the resulting Gaussian variable. In MLE2, we first average the likelihood function over the CIR. Since the transmitted signal can be treated as Gaussian, we next average the likelihood function over the Gaussian variable, resulting in MLE2. In MLE3, the CIR is obtained by maximizing the likelihood function for a given CFO. The CIR is removed by substituting the CIR estimate back into the likelihood function. The MLE3 is then derived by averaging over the time domain signal. We show that the estimators in [6], [7] and all our proposed estimators are equivalent in high SNR, when only virtual carriers exist. Moreover, all the three estimators can use embedded pilots to improve their accuracy and to perform joint estimation of CIR and CFO.

Notation: boldface letters will be used for matrices and column vectors. $(\cdot)^H$ denotes Hermitian (conjugate transpose). \mathbf{I}_N denotes the $N \times N$ identity matrix. $\text{diag}\{\mathbf{x}\}$ stands for the diagonal matrix with the column vector \mathbf{x} on its diagonal. $\text{tr}(\mathbf{A}) = \sum_{i=1}^N a_{ii}$ is the trace of matrix \mathbf{A} . The Column-Circulant down matrix of a vector $\mathbf{a} = [a_1, a_2, \dots, a_n]^T$ (a_k can be scalar, vector or matrix) with ξ block columns is defined as

$$\mathcal{C}_D^{(\xi)}[\mathbf{a}] = \begin{bmatrix} a_1 & a_n & \cdots & a_{n-\xi+2} \\ a_2 & a_1 & \cdots & a_{n-\xi+3} \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n-1} & \cdots & a_{n-\xi+1} \end{bmatrix}. \quad (1)$$

The Column-Circulant up matrix with ξ block columns is

defined as

$$\mathcal{C}_U^{(\xi)}[\mathbf{a}] = \begin{bmatrix} a_1 & a_2 & \cdots & a_\xi \\ a_2 & a_3 & \cdots & a_{\xi+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_1 & \cdots & a_{\xi-1} \end{bmatrix}. \quad (2)$$

II. SYSTEM MODEL

In an OFDM system, source data are grouped and mapped into d_k , which is selected from a complex signal constellation \mathcal{Q} with unitary energy. Complex data are modulated by inverse discrete Fourier transform (IDFT) on N parallel subcarriers. The symbol interval and block interval are denoted by T_s and NT_s . The resulting OFDM symbol during the m th block interval (for brevity we omit m) that comprises N samples is given by

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j(2\pi kn/N)}, \quad n = 0, 1, 2, \dots, N-1 \quad (3)$$

where

$$X(k) = \begin{cases} d_k & k \in I_d \\ p_k & k \in I_p \\ 0 & k \in I_v \end{cases} \quad (4)$$

and I_d is the index set of data subcarriers, I_p is the index set of subcarriers reserved for pilot symbols with N_p elements, I_v is the index set of subcarriers reserved for VCs with N_v elements and the number of data subcarriers is N_d . We have $N_d + N_v + N_p = N$.

The guard interval, inserted to prevent inter-block-interference, includes a cyclic prefix which replicates the end of the IFFT output samples. The number of samples in the guard interval N_g is assumed to be larger than the delay spread of the channel. The signal is transmitted over a frequency selective fading channel modelled as

$$h(t) = \sum_{l=0}^{L-1} \alpha_l \delta(t - \tau_l) \quad (5)$$

where $\alpha_l \sim \mathcal{CN}(0, \sigma_l^2)$ and τ_l is the delay of the l th tap. When a CFO exists, the received signal after sampling is given by

$$y(n) = e^{j2\pi\epsilon\frac{n}{N}} \sum_{l=0}^{L-1} h(l)x(n-l) + w(n) \quad (6)$$

where $w(n)$ is an Additive White Gaussian Noise (AWGN) sample, $w(n) \sim \mathcal{CN}(0, \sigma_n^2)$. Channel taps h_l ($l = 0, \dots, L-1$) represent the sampled overall channel impulse response (which comprises the transmit/receive filters and the physical channel $h(t)$). L is the total number of propagation paths. The CFO normalized by the block interval (NT_s) is denoted by ϵ . As usual, we assume the channel stays constant within each OFDM symbol. For convenience, (6) can be written in vector form as

$$\mathbf{y} = \mathbf{\Gamma}(\epsilon) \mathcal{C}_D^{(L)}[\mathbf{x}] \mathbf{h} + \mathbf{w} \quad (7)$$

or equivalently

$$\mathbf{y} = \mathbf{\Gamma}(\epsilon) \mathbf{F}^H \mathbf{X}_D \mathbf{F}_L \mathbf{h} + \mathbf{w} \quad (8)$$

where $\mathbf{x} = [x(0), x(1), \dots, x(N-1)]^T$, $\mathbf{y} = [y(0), y(1), \dots, y(N-1)]^T$, $\mathbf{h} = [h_0, h_1, \dots, h_{L-1}]$ and $\mathbf{w} = [w(0), w(1), \dots, w(N-1)]$ denote transmitted vector, received vector, channel vector and additive noise, respectively. The $N \times N$ discrete Fourier transform (DFT) matrix \mathbf{F} has the (i, j) -th entry

$$[\mathbf{F}]_{i,j} = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}(i-1)(j-1)} \quad (9)$$

and $\mathbf{F}_L = [\mathbf{f}_{I_h(0)}, \mathbf{f}_{I_h(1)}, \dots, \mathbf{f}_{I_h(L-1)}]$ is the relevant $N \times L$ submatrix of \mathbf{F} , where I_h is the index set of channel taps and \mathbf{f}_i is the i th column of \mathbf{F} . The diagonal matrices $\mathbf{\Gamma}(\epsilon)$

$$\mathbf{\Gamma}(\epsilon) = \text{diag}[1, \xi, \dots, \xi^{N-1}] \quad (10)$$

where $\xi = \exp(j2\pi\epsilon/N)$ and

$$\mathbf{X}_D = \text{diag}[X(0), X(1), \dots, X(N-1)] \quad (11)$$

Let $\mathbf{X} = [X(0), X(1), \dots, X(N-1)]^T$. $\mathbf{x} = \mathbf{F}^H \mathbf{X} / \sqrt{N}$ and $\mathcal{C}_D^{(L)}[\mathbf{x}] = \mathbf{F}^H \mathbf{X}_D \mathbf{F}_L$.

III. MAXIMUM LIKELIHOOD ESTIMATION OF CFO

The received symbol vector \mathbf{y} is Gaussian with mean $\mathbf{\Gamma}(\epsilon) \mathcal{C}_D^{(L)}[\mathbf{x}] \mathbf{h}$ and covariance matrix $\sigma_n^2 \mathbf{I}_N$. The likelihood function for the unknown parameters \mathbf{x} , \mathbf{h} and ϵ is given by

$$\Lambda(\mathbf{y}|\mathbf{x}, \mathbf{h}, \epsilon) = \exp \left\{ -\frac{1}{\sigma_n^2} \left\| \mathbf{y} - \mathbf{\Gamma}(\epsilon) \mathcal{C}_D^{(L)}[\mathbf{x}] \mathbf{h} \right\|^2 \right\}. \quad (12)$$

We next consider ML CFO estimation using the likelihood function.

A. MLEI

When the number of subcarriers N and the channel length L is large, the signal $\mathbf{g} = \mathcal{C}_D^{(L)}[\mathbf{x}] \mathbf{h}$ can be modelled as a zero-mean complex Gaussian process via the central limit theorem. The autocorrelation matrix of \mathbf{g} can be evaluated as

$$\begin{aligned} \mathbf{R}_g &= E\{\mathbf{g}\mathbf{g}^H\} = E\left\{ \mathcal{C}_D^{(L)}[\mathbf{x}] \mathbf{h} \mathbf{h}^H \left(\mathcal{C}_D^{(L)}[\mathbf{x}] \right)^H \right\} \\ &= E\left\{ \mathbf{F}^H \mathbf{X}_D \mathbf{F}_L \mathbf{h} \mathbf{h}^H \mathbf{F}_L^H \mathbf{X}_D^H \mathbf{F} \right\} \\ &= \mathbf{F}^H (r \mathbf{\Lambda}_d + \mathbf{\Lambda}_p \mathbf{F}_L \mathbf{R}_h \mathbf{F}_L^H \mathbf{\Lambda}_p^H) \mathbf{F} \end{aligned} \quad (13)$$

where $r = \sum_{l=0}^{L-1} \sigma_l^2$, $\mathbf{R}_h = E\{\mathbf{h}\mathbf{h}^H\}$ is the covariance matrix of \mathbf{h} and

$$\mathbf{\Lambda}_d = \text{diag}[\lambda_0, \lambda_1, \dots, \lambda_{N-1}], \lambda_k = \begin{cases} 1 & k \in I_d \\ 0 & \text{otherwise} \end{cases}. \quad (14)$$

$$\mathbf{\Lambda}_p = \text{diag}[\lambda_0, \lambda_1, \dots, \lambda_{N-1}], \lambda_k = \begin{cases} 1 & k \in I_p \\ 0 & \text{otherwise} \end{cases}. \quad (15)$$

The average of $\Lambda(\mathbf{y}|\mathbf{x}, \mathbf{h}, \epsilon)$ over \mathbf{g} gives the marginal likelihood function $\Lambda(\mathbf{y}|\epsilon)$, which removes the likelihood function dependence on \mathbf{h} and \mathbf{x} . The marginal likelihood function can be expressed as

$$\Lambda(\mathbf{y}|\epsilon) = E_{\mathbf{g}} \left\{ \exp \left\{ -\frac{1}{\sigma_n^2} \|\mathbf{y} - \Gamma(\epsilon)\mathbf{g}\|^2 \right\} \right\} \quad (16)$$

where $E_{\mathbf{X}}\{\cdot\}$ denotes the statistical average with respect to the random vector \mathbf{X} . Eq. (16) can be interpreted as the characteristic function (CF) $\lambda_Q(s) = E\{\exp(sQ)\}$ of the quadratic form

$$Q = \|\mathbf{y} - \Gamma(\epsilon)\mathbf{g}\|^2 \quad (17)$$

of the zero-mean Gaussian vector \mathbf{g} evaluated for $s = -1/(\sigma_n^2)$. If \mathbf{R}_g is non-singular (e.g., no VCs), the CF is given by [10, p. 595, eq. (B-3-20)]. When VC exists, \mathbf{R}_g becomes singular, which means some of the random variables in \mathbf{g} are linearly dependent. To make the CF approach still applicable in this case, we assume

$$\mathbf{\Lambda}_v(\zeta) = \text{diag}[\lambda_0, \lambda_1, \dots, \lambda_{N-1}], \lambda_k = \begin{cases} \zeta & k \in I_v \\ 1 & \text{otherwise} \end{cases} \quad (18)$$

Using [10, p. 595, eq. (B-3-20)], we can obtain

$$\Lambda(\mathbf{y}|\epsilon, \zeta) = \exp \left\{ \frac{1}{\sigma_n^2} \mathbf{y}^H \Gamma(\epsilon) \mathbf{G}(\zeta) \Gamma(\epsilon)^H \mathbf{y} \right\} \quad (19)$$

Let $\zeta \rightarrow 0$, we have

$$\Lambda_1(\mathbf{y}|\epsilon) = \lim_{\zeta \rightarrow 0} \Lambda(\mathbf{y}|\epsilon, \zeta) = \exp \left\{ \frac{1}{\sigma_n^2} \mathbf{y}^H \Gamma(\epsilon) \mathbf{G} \Gamma(\epsilon)^H \mathbf{y} \right\} \quad (20)$$

where $\mathbf{G} = \mathbf{R}_g(\mathbf{R}_g + \sigma_n^2 \mathbf{I}_N)^{-1}$. Eq. (20) follows from that the limitation and integral operations are interchangeable. The ML estimate of ϵ can be obtained by maximizing

$$\hat{\epsilon} = \arg \max_{\epsilon} \mathbf{y}^H \Gamma(\epsilon) \mathbf{G} \Gamma(\epsilon)^H \mathbf{y} \quad (21)$$

B. MLE2

In MLE2 and MLE3, we assume that no pilot exists. The existence of pilots will be discussed in the full journal version. As [9], we first evaluate the marginal likelihood function $\Lambda(\mathbf{y}|\mathbf{x}, \epsilon)$, which is the average of $\Lambda(\mathbf{y}|\mathbf{h}, \mathbf{x}, \epsilon)$ with respect to \mathbf{h} and can be expressed as

$$\Lambda(\mathbf{y}|\mathbf{x}, \epsilon) = E_{\mathbf{h}} \{ \Lambda(\mathbf{y}|\mathbf{h}, \mathbf{x}, \epsilon) \}. \quad (22)$$

Using [10, p. 595, eq. (B-3-20)] and dropping irrelevant factors, (22) becomes

$$\Lambda(\mathbf{y}|\mathbf{x}, \epsilon) = \exp \left\{ -\frac{1}{\sigma_n^2} \left(\mathbf{y}^H \mathbf{y} - \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(L)}[\mathbf{x}] \times \mathbf{G} \left(\mathcal{C}_D^{(L)}[\mathbf{x}] \right)^H \Gamma(\epsilon)^H \mathbf{y} \right) \right\} \quad (23)$$

where $\mathbf{G} = \mathbf{R}_h(\mathbf{R}_h + \sigma_n^2 \mathbf{I}_N)^{-1}$. We define $N \times L$ matrix Ξ

$$\Xi_{i,j} = \begin{cases} 1 & i = I_h(j), j = 0, 1, \dots, L-1 \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

and

$$\mathbf{\Omega} = \text{diag}[\omega_0, \omega_1, \dots, \omega_{N-1}], \omega_i = \begin{cases} 1 & i = I_h(k), \\ 0 & k = 0, 1, \dots, L-1 \\ & \text{otherwise} \end{cases} \quad (25)$$

It can be readily verified that $\mathcal{C}_D^{(L)}[\mathbf{x}] = \mathcal{C}_D^{(N)}[\mathbf{x}] \Xi$, $\Xi \Xi^H = \mathbf{\Omega}$ and

$$\begin{aligned} \mathbf{y}^H \mathbf{y} &= \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}] \left(\mathcal{C}_D^{(N)}[\mathbf{x}] \right)^H \Gamma^H(\epsilon) \mathbf{y} \\ &+ \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}_v] \left(\mathcal{C}_D^{(N)}[\mathbf{x}] \right)^H \Gamma^H(\epsilon) \mathbf{y} \\ &+ \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}] \left(\mathcal{C}_D^{(N)}[\mathbf{x}_v] \right)^H \Gamma^H(\epsilon) \mathbf{y} \\ &+ \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}_v] \left(\mathcal{C}_D^{(N)}[\mathbf{x}_v] \right)^H \Gamma^H(\epsilon) \mathbf{y} \end{aligned} \quad (26)$$

where $\mathbf{x}_v = \mathbf{F}^H \mathbf{\Lambda}_v / \sqrt{N}$, $\mathcal{C}_D^{(N)}[\mathbf{x}_v] = \mathbf{F}^H \mathbf{\Lambda}_v \mathbf{F}$ and

$$\mathbf{\Lambda}_v = \text{diag}[\lambda_0, \lambda_1, \dots, \lambda_{N-1}], \lambda_k = \begin{cases} 1 & k \in I_v \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

Hence, (22) becomes

$$\Lambda(\mathbf{y}|\mathbf{x}, \epsilon) = \exp \left\{ -\frac{1}{\sigma_n^2} (\beta_1 + \beta_2 + \beta_3 + \beta_4) \right\} \quad (28)$$

where

$$\begin{aligned} \beta_1 &= \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}] (\mathbf{I}_N - \Xi \mathbf{G} \Xi) \left(\mathcal{C}_D^{(N)}[\mathbf{x}] \right)^H \Gamma^H(\epsilon) \mathbf{y} \\ \beta_2 &= \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}_v] \left(\mathcal{C}_D^{(N)}[\mathbf{x}] \right)^H \Gamma^H(\epsilon) \mathbf{y} \\ \beta_3 &= \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}] \left(\mathcal{C}_D^{(N)}[\mathbf{x}_v] \right)^H \Gamma^H(\epsilon) \mathbf{y} \\ \beta_4 &= \mathbf{y}^H \Gamma(\epsilon) \mathcal{C}_D^{(N)}[\mathbf{x}_v] \left(\mathcal{C}_D^{(N)}[\mathbf{x}_v] \right)^H \Gamma^H(\epsilon) \mathbf{y} \end{aligned} \quad (29)$$

Let $\mathbf{z} = \Gamma^H(\epsilon) \mathbf{y}$ and note that for two vectors \mathbf{a} and \mathbf{b} of length N the following relationship holds

$$\mathbf{a}^H \mathcal{C}_D^{(N)}[\mathbf{b}] = \mathbf{b}^T \mathcal{C}_U^{(N)}[\mathbf{a}^*]. \quad (30)$$

Therefore $\beta_1 - \beta_4$ can be rewritten as

$$\begin{aligned} \beta_1 &= \mathbf{x}^H \mathcal{C}_U^{(N)}[\mathbf{z}] (\mathbf{I}_N - \Xi \mathbf{G} \Xi) \left(\mathcal{C}_U^{(N)}[\mathbf{z}] \right)^H \mathbf{x} \\ \beta_2 &= \mathbf{x}^H \mathcal{C}_U^{(N)}[\mathbf{z}] \left(\mathcal{C}_U^{(N)}[\mathbf{z}] \right)^H \mathbf{x}_v \\ \beta_3 &= \mathbf{x}_v^H \mathcal{C}_U^{(N)}[\mathbf{z}] \left(\mathcal{C}_U^{(N)}[\mathbf{z}] \right)^H \mathbf{x} \\ \beta_4 &= \mathbf{x}_v^H \mathcal{C}_U^{(N)}[\mathbf{z}] \left(\mathcal{C}_U^{(N)}[\mathbf{z}] \right)^H \mathbf{x}_v. \end{aligned} \quad (31)$$

Without pilots, the time domain transmitted signal \mathbf{x} can be modelled as a zero-mean complex Gaussian process by using the central limit theorem when the number of subcarriers N is large. The covariance matrix of \mathbf{x} can be obtained as

$$\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^H\} = \frac{1}{N} \mathbf{F}^H \mathbf{\Lambda}_d \mathbf{F} \quad (32)$$

where Λ_d is given in (14). The average of $\Lambda(\mathbf{y}|\mathbf{x}, \epsilon)$ with respect to \mathbf{x} generates the marginal likelihood function, which is given by

$$\Lambda(\mathbf{y}|\epsilon) = E_{\mathbf{x}}\{\Lambda(\mathbf{y}|\mathbf{x}, \epsilon)\}. \quad (33)$$

Using [10, p. 595, eq. (B-3-20)] and following the same approach as (18)-(20) for the singular case, (33) yields

$$\Lambda_2(\mathbf{y}|\epsilon) = \frac{1}{\det\left(\mathbf{I}_N + \frac{\mathbf{R}_x \mathbf{A}}{\sigma_n^2}\right)} \exp\left\{-\frac{1}{\sigma_n^2} \mathbf{x}_v^H (\mathbf{B} + \mathbf{C}) \mathbf{x}_v\right\}. \quad (34)$$

where

$$\begin{aligned} \mathbf{A} &= \mathcal{C}_U^{(N)}[\mathbf{z}] (\mathbf{I}_N - \Xi \mathbf{G} \Xi) \left(\mathcal{C}_U^{(N)}[\mathbf{z}]\right)^H \\ \mathbf{B} &= \mathcal{C}_U^{(N)}[\mathbf{z}] \left(\mathcal{C}_U^{(N)}[\mathbf{z}]\right)^H \\ \mathbf{C} &= \mathbf{B} \left(\mathbf{I}_N + \frac{\mathbf{R}_x \mathbf{A}}{\sigma_n^2}\right)^{-1} \mathbf{R}_x \mathbf{B}. \end{aligned} \quad (35)$$

Therefore, MLE2 is given by

$$\hat{\epsilon} = \arg \max_{\epsilon} \Lambda_2(\mathbf{y}|\epsilon). \quad (36)$$

C. MLE3

Two possible approaches may be taken to estimate ϵ according to the ML criterion. One is the Bayesian approach adopted in MLE2. We use the alternative approach to derive MLE3. If we keep ϵ and \mathbf{x} fixed and let \mathbf{h} vary in (12), $\Lambda(\mathbf{y}|\mathbf{x}, \mathbf{h}, \epsilon)$ achieves a maximum for

$$\begin{aligned} \mathbf{h}(\mathbf{x}, \epsilon) &= \left[\left(\mathcal{C}_D^{(L)}[\mathbf{x}]\right)^H \mathcal{C}_D^{(L)}[\mathbf{x}] \right]^{-1} \left(\mathcal{C}_D^{(L)}[\mathbf{x}]\right)^H \mathbf{\Gamma}^H(\epsilon) \mathbf{y} \\ &= \left(\mathbf{F}_L^H \Lambda_d \mathbf{F}_L\right)^{-1} \left(\mathcal{C}_D^{(L)}[\mathbf{x}]\right)^H \mathbf{\Gamma}^H(\epsilon) \mathbf{y} \end{aligned} \quad (37)$$

Substituting (37) into (12), we can obtain

$$\begin{aligned} &\Lambda(\mathbf{y}|\mathbf{x}, \epsilon) \\ &= \exp\left\{-\frac{1}{\sigma_n^2} \left\| \mathbf{y} - \mathbf{\Gamma}(\epsilon) \mathcal{C}_D^{(L)}[\mathbf{x}] \left(\mathbf{F}_L^H \Lambda_d \mathbf{F}_L\right)^{-1} \right. \right. \\ &\quad \left. \left. \times \left(\mathcal{C}_D^{(L)}[\mathbf{x}]\right)^H \mathbf{\Gamma}^H(\epsilon) \mathbf{y} \right\|^2\right\} \\ &= \exp\left\{-\frac{1}{\sigma_n^2} \left(\mathbf{y}^H \mathbf{y} - \mathbf{y}^H \mathbf{\Gamma}(\epsilon) \mathcal{C}_D^{(L)}[\mathbf{x}] \left(\mathbf{F}_L^H \Lambda_d \mathbf{F}_L\right)^{-1} \right. \right. \\ &\quad \left. \left. \times \left(\mathcal{C}_D^{(L)}[\mathbf{x}]\right)^H \mathbf{\Gamma}^H(\epsilon) \mathbf{y}\right)\right\}. \end{aligned} \quad (38)$$

Following the same procedure as (24)-(35), the marginal likelihood function can be derived as

$$\Lambda_3(\mathbf{y}|\epsilon) = \frac{1}{\det\left(\mathbf{I}_N + \frac{\mathbf{R}_x \mathbf{A}}{\sigma_n^2}\right)} \exp\left\{-\frac{1}{\sigma_n^2} \mathbf{x}_v^H (\mathbf{B} + \mathbf{C}) \mathbf{x}_v\right\}. \quad (39)$$

where

$$\begin{aligned} \mathbf{A} &= \mathcal{C}_U^{(N)}[\mathbf{z}] \left(\mathbf{I}_N - \Xi \left(\mathbf{F}_L^H \Lambda_d \mathbf{F}_L\right)^{-1} \Xi\right) \left(\mathcal{C}_U^{(N)}[\mathbf{z}]\right)^H \\ \mathbf{B} &= \mathcal{C}_U^{(N)}[\mathbf{z}] \left(\mathcal{C}_U^{(N)}[\mathbf{z}]\right)^H \\ \mathbf{C} &= \mathbf{B} \left(\mathbf{I}_N + \frac{\mathbf{R}_x \mathbf{A}}{\sigma_n^2}\right)^{-1} \mathbf{R}_x \mathbf{B}. \end{aligned} \quad (40)$$

Therefore, MLE3 is given by

$$\hat{\epsilon} = \arg \max_{\epsilon} \Lambda_3(\mathbf{y}|\epsilon). \quad (41)$$

IV. PERFORMANCE OF ESTIMATORS

Since the MLE1 is equivalent to the one in [6] without pilots, we expect that MLE1 is unbiased. Hence MLE2 and MLE3 are asymptotically unbiased. If pilots exist, we assess the performance of our proposed CFO estimator using the approach in [11], which show that the expectation of the ML estimator in high SNR is given by

$$E\{\hat{\epsilon}\} \approx \epsilon - \frac{E\{\dot{g}(\epsilon)\}}{E\{\ddot{g}(\epsilon)\}} \quad (42)$$

where $g(\epsilon)$ denotes the cost function of the estimator given by (21) in this case. The first and second derivatives of $g(\epsilon)$ are given by $\dot{g}(\epsilon)$ and $\ddot{g}(\epsilon)$. It can be proved that

$$E\{\dot{g}(\epsilon)\} = \epsilon. \quad (43)$$

MLE1 is thus unbiased.

From (21), (34) and (39), $g(\epsilon)$ is periodic of period N . The range of MLE1 is $[0, N)$ if $L < N_v$. The integer CFO and fractional CFO do not need to perform separately. To assure that MLE1 only has a unique minimum, the VCs are placed distinctly and $N_v = L + 2$ in [12]. MLE2 and MLE3 asymptotically follow the same properties.

If neither pilot nor VC exists, it can be readily verified that \mathbf{G} in (20) becomes a diagonal matrix. In this case, Eq. (20) becomes a constant, which means that MLE1 cannot estimate CFO without VC. Likewise, MLE2, MLE3 and the estimators in [6], [7] cannot work without VC, which means all the ML CFO estimators are VC driven.

In MLE1-MLE3, pilots embedded for channel estimation can be used to assist CFO estimation. This suggests the possibility of joint estimation of CIR and CFO.

V. NUMERICAL RESULTS

We now present numerical results to illustrate the effectiveness of the proposed ML estimators for a practical OFDM system. We assume the following system specifications:

- Both the data and pilot symbols are chosen from the QPSK alphabet, denoted by \mathcal{Q} .
- The carrier frequency of the OFDM system is 5GHz and the data rate is 12MHz.

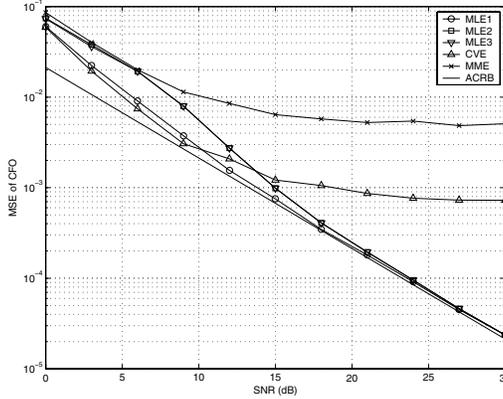


Fig. 1. MSE of CFO versus SNR in a QPSK OFDM system with $\epsilon = 0.25$, $N = 64$, $N_p = 0$ and $N_v = 12$.

- The 6-ray COST 207 TU model with the PDP $[0.189, 0.379, 0.239, 0.095, 0.061, 0.037]$ and delay profile $[0.0, 0.2, 0.5, 1.6, 2.3, 5.0]\mu s$ is considered. The channel remains constant during each OFDM data block but varies from one block to another. All the paths are rounded to integer and we do not consider leakage. The PDP is assumed perfectly known at the receiver.
- A normalized CFO of 0.25 is considered.
- The number of subcarriers $N = 64$, the number of VCs $N_v = 12$ and VCs are distinctly placed [12] to preserve the uniqueness of the estimator.

In Fig. 1, the MSE ($E\{(\hat{\epsilon} - \epsilon)^2\}$) of proposed estimators with no pilot are compared with those of the Morelli and Mengali Estimator (MME) [8], the Chiavaccini and Vitetta Estimator (CVE) [9] and the asymptotic Cramer-Rao bound (ACRB) given by [9] as

$$\text{ACRB} = \sigma_n^2 \left[\frac{\pi^2 N_v (N^2 - 1)}{3N^2} - \frac{2\pi^2}{N^2} \sum_{i \in I_v} \sum_{j \in I_v, i \neq j} \frac{1}{|\exp[j2\pi(i-j)/N] - 1|^2} \right]. \quad (44)$$

In low SNR, all the estimators perform similarly and CVE performs slightly better than the other estimators. Both MME and CVE show error floors in high SNR. The error floors of MME and CVE are 0.0051 and 0.0007, respectively. MLE2 and MLE3 perform almost identically for any SNR and both performances approach that of MLE1. Our proposed estimators asymptotically approach the ACRB.

Fig. 2 shows the performance of MLE1 with different number of pilots. The optimal pilots for CFO estimator can be obtained by minimizing the CRB for MLE1. Here, the pilots are equispaced and are randomly chosen from \mathcal{Q} . They vary each frame. MLE1 with $N_p = 4$ gains 2 dB over that of $N_p = 0$. The gain can be as large as 4 dB when $N_p = 16$. These verify the performance enhancement using pilots. The CIR and CFO can thus be jointly estimated using these pilots.

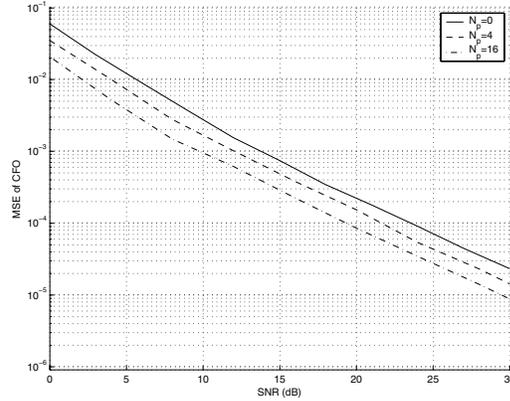


Fig. 2. MSE of CFO versus SNR with $N_p = 0, 4, 16$ in a QPSK OFDM system with $\epsilon = 0.25$, $N = 64$ and $N_v = 12$.

VI. CONCLUSION

Three new ML CFO estimators have been derived in this paper. MLE2 and MLE3 perform close to MLE1 in high SNR and all of them achieve the ACRB. We show that the estimators in [6], [7] and MLE1 are equivalent and are VC driven estimators. All our proposed estimators can utilize the pilots to enhance CFO estimation, enabling joint channel and CFO estimation. The proposed estimators can be implemented by an efficient FFT algorithm.

REFERENCES

- [1] B. LeFloch, M. Alard, and C. Berrou, "Coded orthogonal frequency division multiplex," *Proc. IEEE*, vol. 83, pp. 982–996, June 1995.
- [2] T. Pollet and M. Moeneclaey, "The effect of carrier frequency offset on the performance of band limited single carrier and ofdm signals," in *Proceeding of IEEE GLOBECOM96*, London, U.K., 1996, pp. 719–723.
- [3] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*
- [4] J. J. van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Processing*, vol. 45, no. 7, pp. 1800–1805, Jul. 1997.
- [5] M. Schmidl and D. C. Cox, "Blind synchronization for OFDM," *Electron. Lett.*, vol. 33, pp. 113–114, Feb. 1997.
- [6] H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," *IEEE Commun. Lett.*, vol. 2, no. 4, pp. 104–106, April 1998.
- [7] F. Daffara and A. Chouly, "Maximum likelihood frequency detectors for orthogonal multicarrier systems," in *Proceeding of IEEE International Conference on Communications*, Geneva, May 1993, pp. 766–771.
- [8] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over frequency selective channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.
- [9] E. Chiavaccini and G. M. Vitetta, "Maximum-likelihood frequency recovery for OFDM signals transmitted over multipath fading channels," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 244–251, Feb. 2004.
- [10] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.
- [11] M. H. Meyers and L. Franks, "Joint carrier phase and symbol timing recovery for PAM systems," *IEEE Trans. Commun.*, vol. 28, no. 8, pp. 1121–1129, Aug. 1980.
- [12] X. Ma, C. Tepedelenioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: identifiability, algorithms, and performance," *IEEE J. Select. Areas Commun.*, vol. 19, no. 12, pp. 2504–2515, Dec. 2001.