This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2007 proceedings.

MIMO OFDM Frequency Offset Estimator with Low Computational Complexity

Yanxiang Jiang, Xiaohu You, Xiqi Gao
National Mobile Communications Research Laboratory,
Southeast University, Nanjing 210096, China.
E-mail: {yxjiang, xhyu, xqgao}@seu.edu.cn

Hlaing Minn
Department of Electrical Engineering,
University of Texas at Dallas, TX 75083-0688, USA.
E-mail: hlaing.minn@utdallas.edu

Abstract— **This paper addresses a low complexity frequency offset estimator for multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems over frequency selective fading channels. By exploiting the good correlation property of the training sequences, which are constructed from the Chu sequence, carrier frequency offset (CFO) estimation is obtained with great complexity reduction through factor decomposition for the derivative of the cost function. The CFO estimate's variance and Cramer-Rao bound (CRB) are developed to optimize the parameter of the simplified estimator and also to evaluate the estimation performance. Simulation results verify the good performance of the training-assisted CFO estimator.**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) is a leading modulation technique for wide-band wireless communications. Combining it with multiple-input multiple-output (MIMO) multi-antenna technique [1] promises a significant increase in the practically achievable throughput over wireless media. The performance of OFDM systems, however, is sensitive to carrier frequency offset (CFO) between transmitter and receiver oscillators [2]. Accurate estimation and compensation of CFO is therefore very important in order to realize the advantages of MIMO-OFDM.

CFO estimation is a well-studied problem for single-antenna OFDM systems [3]–[9], but a relatively new one for MIMO or MIMO OFDM systems [10]–[16]. Numerical calculations of the CFO estimators in [12] [13] require a large point discrete Fourier transform (DFT) operation and a time consuming line search. To reduce complexity, computationally efficient CFO estimators are introduced in [14]–[16]. However, the estimator in [14] is only applied to flat-fading MIMO channels, whereas the estimator in [15] still needs relatively heavy computational burden due to the exploitation of eigen-decomposition and polynomial root-finding procedures.

In this paper, with the aid of the training sequences constructed from the Chu sequence in [15], we propose a low complexity CFO estimator for MIMO OFDM systems, which has almost the same estimation accuracy in comparison with the one in [15]. To optimize the parameter of the proposed CFO estimator, the variance of the CFO estimation is derived. To evaluate the estimation performance, the Cramer-Rao Bound (CRB) is developed.

The rest of this paper is organized as follows. In Section II, the signal model is briefly described. The simplified CFO estimator and the complexity analysis are presented in Section III. The performance analysis is given in Section IV. Simulation results are shown in Section V. Final conclusions are drawn in Section VI.

Notations: Upper (lower) bold-face letters are used for matrices (column vectors). Superscripts $*$, T and H denote conjugate, transpose and Hermitian transpose, respectively. $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of the enclosed parameters, respectively. $(\cdot)_P$ denotes the residue of the number within the brackets modulo P. $\lfloor\cdot\rfloor$, $\|\cdot\|^2$, $\odot$ and $E[\cdot]$ denote the floor, Euclidean norm, Schur-Hadamard (i.e., element-wise) product and expectation operators, respectively. $[\boldsymbol{X}]_{m,n}$ denotes the (m,n)-th entry of a matrix $\boldsymbol{X}$. $[\boldsymbol{x}]_m$ denotes the m-th entry of a column vector $\boldsymbol{x}$. $[\boldsymbol{x}]^{(m)}$ denotes the m-cyclic-right-shift version of $\boldsymbol{x}^T$. $\text{diag}\{\boldsymbol{x}\}$ denotes a diagonal matrix with the elements of $\boldsymbol{x}$ on its diagonal. $\boldsymbol{F}_N$ and $\boldsymbol{I}_N$ denote the $N\times N$ unitary DFT matrix and the $N\times N$ identity matrix, respectively. $\boldsymbol{f}_N^k$ and $\boldsymbol{e}_N^k$ denote the k-th column vectors of $\boldsymbol{F}_N$ and $\boldsymbol{I}_N$, respectively. Unless otherwise stated, we assume $(N)_{2P}=0$, $Q=N/P$, $0\le p\le P-1$, $0\le q\le Q-2$, $0\le\mu\le N_t-1$ and $0\le\nu\le N_r-1$.

II. Signal Model

Consider a MIMO OFDM system with N subcarriers, N_t transmit antennas and N_r receive antennas. The training sequences and the baseband equivalent signal model are the same as in [15].

Let

$$0\le i_0<i_1<\cdots<i_\mu<\cdots<i_{N_t-1}<Q,$$

$$\boldsymbol{\Theta}^{(\mu)}=[\boldsymbol{e}_N^{i_\mu},\boldsymbol{e}_N^{i_\mu+Q},\cdots,\boldsymbol{e}_N^{i_\mu+(P-1)Q}],\ M=\lfloor P/N_I\rfloor,\ N_I\ge N_t.$$

Let $\boldsymbol{s}_P$ denote a length-P Chu sequence [17]. Define $\tilde{\boldsymbol{s}}_P^{(\mu)}=\sqrt{Q/N_t}\boldsymbol{F}_P[\boldsymbol{s}_P]^{(\mu M)}$. Then the training sequence vector at the μ-th transmit antenna is $\tilde{\boldsymbol{s}}^{(\mu)}=\boldsymbol{\Theta}^{(\mu)}\tilde{\boldsymbol{s}}_P^{(\mu)}$. For convenience, we henceforth refer to $\{\tilde{\boldsymbol{s}}^{(\mu)}\}_{\mu=0}^{N_t-1}$ as the Chu sequence based training sequences (CBTS). Note that other base sequences with zero auto-correlation property can also be employed in our training sequences (e.g., [18] and references therein).

Let $\boldsymbol{h}_L^{(\nu,\mu)}$ denote the length-L channel impulse response vector with L being the maximum channel length. Assume L

1-4244-0353-7/07/$25.00

is shorter than the length of cyclic prefix (CP) N_g. Let $\tilde{\varepsilon}$ denote the frequency offset normalized by the subcarrier spacing. Let $\boldsymbol{y}^{(\nu)}$ denote the length-N received vector at the ν-th receive antenna after CP removal. Define

$$\begin{aligned} \boldsymbol{y} &= [\boldsymbol{y}^{(0),T}, \boldsymbol{y}^{(1),T}, \cdots, \boldsymbol{y}^{(N_r-1),T}]^T, \\ \boldsymbol{h}^{(\nu)} &= [\boldsymbol{h}_L^{(\nu,0),T}, \boldsymbol{h}_L^{(\nu,1),T}, \cdots, \boldsymbol{h}_L^{(\nu,N_t-1),T}]^T, \\ \boldsymbol{h} &= [\boldsymbol{h}^{(0),T}, \boldsymbol{h}^{(1),T}, \cdots, \boldsymbol{h}^{(N_r-1),T}]^T, \\ \boldsymbol{D}_{\bar{N}}(\tilde{\varepsilon}) &= \text{diag}\{[1, e^{j2\pi\tilde{\varepsilon}/N}, \cdots, e^{j2\pi\tilde{\varepsilon}(\bar{N}-1)/N}]^T\}. \end{aligned}$$

Then, the cascaded received vector $\boldsymbol{y}$ over the N_r receive antennas is given by

$$\boldsymbol{y} = \sqrt{N} e^{j\frac{2\pi\varepsilon N_g}{N}} [\boldsymbol{I}_{N_r} \otimes (\boldsymbol{D}_N(\tilde{\varepsilon})\boldsymbol{S})]\boldsymbol{h} + \boldsymbol{w}, \tag{1}$$

where

$$\begin{aligned} \boldsymbol{S} = \bar{\boldsymbol{F}}_{\mathcal{P}}^* \check{\boldsymbol{S}} \check{\boldsymbol{F}}, \ \check{\boldsymbol{S}} &= \text{diag}\{[\tilde{\boldsymbol{s}}_P^{(0),T}, \tilde{\boldsymbol{s}}_P^{(1),T}, \cdots, \tilde{\boldsymbol{s}}_P^{(N_t-1),T}]^T\}, \\ \bar{\boldsymbol{F}}_{\mathcal{P}} &= \boldsymbol{F}_N[\boldsymbol{\Theta}^{(0)}, \boldsymbol{\Theta}^{(1)}, \cdots, \boldsymbol{\Theta}^{(N_t-1)}], \\ \check{\boldsymbol{F}} &= [\boldsymbol{e}_{N_t}^0 \otimes \check{\boldsymbol{F}}_L^{(0)}, \boldsymbol{e}_{N_t}^1 \otimes \check{\boldsymbol{F}}_L^{(1)}, \cdots, \boldsymbol{e}_{N_t}^{N_t-1} \otimes \check{\boldsymbol{F}}_L^{(N_t-1)}], \\ \check{\boldsymbol{F}}_L^{(\mu)} &= \boldsymbol{\Theta}^{(\mu),T} \boldsymbol{F}_N [\boldsymbol{I}_L, \boldsymbol{0}_{L\times(N-L)}]^T, \end{aligned}$$

and $\boldsymbol{w}$ is a length-N_rN vector of zero-mean, uncorrelated complex Gaussian noise samples with equal variance of σ_w^2.

III. Low Complexity CFO Estimator for MIMO OFDM

By exploiting the periodicity property of the training sequences, the received vector $\boldsymbol{y}$ can be stacked into a $Q \times N_rP$ matrix $\boldsymbol{Y} = [\boldsymbol{Y}^{(0)}, \boldsymbol{Y}^{(1)}, \cdots, \boldsymbol{Y}^{(N_r-1)}]$, where

$$\begin{aligned} \boldsymbol{Y}^{(\nu)} &= [\boldsymbol{y}_0^{(\nu)}, \boldsymbol{y}_1^{(\nu)}, \cdots, \boldsymbol{y}_{P-1}^{(\nu)}], \\ \boldsymbol{y}_p^{(\nu)} &= [\boldsymbol{e}_N^p, \boldsymbol{e}_N^{p+P}, \cdots, \boldsymbol{e}_N^{p+(Q-1)P}]^T \boldsymbol{y}^{(\nu)}. \end{aligned}$$

Then, by using the similar approach as shown in the derivation of the fractional CFO estimator in [15], $\boldsymbol{Y}$ can be expressed in the following form

$$\boldsymbol{Y} = \boldsymbol{B}(\tilde{\varepsilon})\boldsymbol{X} + \boldsymbol{W}, \tag{2}$$

where

$$\begin{aligned} \boldsymbol{B}(\tilde{\varepsilon}) &= [\boldsymbol{b}_0, \boldsymbol{b}_1, \cdots, \boldsymbol{b}_{N_t-1}], \\ \boldsymbol{b}_\mu &= [1, e^{j\frac{2\pi(\tilde{\varepsilon}+i_\mu)}{Q}}, \cdots, e^{j\frac{2\pi(\tilde{\varepsilon}+i_\mu)(Q-1)}{Q}}]^T, \\ \boldsymbol{X} &= [\boldsymbol{X}^{(0)}, \boldsymbol{X}^{(1)}, \cdots, \boldsymbol{X}^{(N_r-1)}], \\ \boldsymbol{X}^{(\nu)} &= [\boldsymbol{x}^{(\nu,0)}, \boldsymbol{x}^{(\nu,1)}, \cdots, \boldsymbol{x}^{(\nu,N_t-1)}]^T, \\ \boldsymbol{x}^{(\nu,\mu)} &= \sqrt{P} e^{j\frac{2\pi\tilde{\varepsilon}N_g}{N}} \boldsymbol{D}_P(\tilde{\varepsilon}+i_\mu)\boldsymbol{F}_P^H \text{diag}\{\tilde{\boldsymbol{s}}_P^{(\mu)}\}\check{\boldsymbol{F}}_L^{(\mu)}\boldsymbol{h}_L^{(\nu,\mu)}, \\ \boldsymbol{W} &= [\boldsymbol{W}^{(0)}, \boldsymbol{W}^{(1)}, \cdots, \boldsymbol{W}^{(N_r-1)}], \\ \boldsymbol{W}^{(\nu)} &= [\boldsymbol{w}_0^{(\nu)}, \boldsymbol{w}_1^{(\nu)}, \cdots, \boldsymbol{w}_{P-1}^{(\nu)}], \\ \boldsymbol{w}_p^{(\nu)} &= [\boldsymbol{e}_N^p, \boldsymbol{e}_N^{p+P}, \cdots, \boldsymbol{e}_N^{p+(Q-1)P}]^T(\boldsymbol{e}_{N_r}^{\nu,T} \otimes \boldsymbol{I}_N)\boldsymbol{w}. \end{aligned}$$

Since the elements of $\boldsymbol{W}$ are independently and identically distributed (i.i.d.) Gaussian random variable with zero mean and equal variance of σ_w^2, according to the multivariate statistical theory, the log-likelihood function of $\boldsymbol{Y}$ conditioned on $\boldsymbol{B}(\varepsilon)$ and $\boldsymbol{X}$ with ε denoting a candidate CFO can be obtained as

$$\begin{aligned} \ln\, p(\boldsymbol{Y}|\boldsymbol{B}(\varepsilon), \boldsymbol{X}) &= -\sigma_w^{-2} tr\{[\boldsymbol{Y} - \boldsymbol{B}(\varepsilon)\boldsymbol{X}][\boldsymbol{Y} - \boldsymbol{B}(\varepsilon)\boldsymbol{X}]^H\} \\ &= -\sigma_w^{-2} \sum_{\nu=0}^{N_r-1}\sum_{p=0}^{P-1} \left\{\|\boldsymbol{y}_p^{(\nu)} - \boldsymbol{B}(\varepsilon)\check{\boldsymbol{x}}_p^{(\nu)}\|^2\right\}, \end{aligned} \tag{3}$$

where $\check{\boldsymbol{x}}_p^{(\nu)} = \boldsymbol{X}^{(\nu)}\boldsymbol{e}_P^p$. For a deterministic ε, by maximizing (3) with respect to $\check{\boldsymbol{x}}_p^{(\nu)}$, the maximum likelihood (ML) estimation of $\check{\boldsymbol{x}}_p^{(\nu)}$ can be obtained as [19]

$$\hat{\boldsymbol{x}}_p^{(\nu)} = Q^{-1}\boldsymbol{B}^H(\varepsilon)\boldsymbol{y}_p^{(\nu)}, \tag{4}$$

where we have exploited the following condition $i_\mu = i_{\mu'}$ iff. $\mu = \mu'$. From (4), we can get the estimation of $\boldsymbol{X}$ by stacking $\hat{\boldsymbol{x}}_p^{(\nu)}$ into the $N_t \times N_rP$ matrix as follows

$$\hat{\boldsymbol{X}} = Q^{-1}\boldsymbol{B}^H(\varepsilon)\boldsymbol{Y}. \tag{5}$$

Substitute (5) into (3). Then, after some straightforward manipulations, the cost function with respect to ε can be obtained as

$$\ln\, p(\boldsymbol{Y}|\varepsilon) = tr[\boldsymbol{B}^H(\varepsilon)\hat{\boldsymbol{R}}_{\boldsymbol{YY}}\boldsymbol{B}(\varepsilon)], \tag{6}$$

where $\hat{\boldsymbol{R}}_{\boldsymbol{YY}} = \boldsymbol{YY}^H$. Direct ML CFO estimate from the cost function in (6) undertakes heavy computational burden. In order to simplify the computation, we will give a low complexity approach subsequently.

Define

$$z = e^{j2\pi\varepsilon/Q},\ z_\mu = e^{j2\pi i_\mu/Q},\ \check{\boldsymbol{b}}_Q(z) = [z, z^2, \cdots, z^Q]^T.$$

Then, by exploiting the Hermitian property of $\hat{\boldsymbol{R}}_{\boldsymbol{YY}}$, the cost function in (6) can be transformed into the following degree-$(Q-1)$ polynomial (ignoring the irrelevant item)

$$\begin{aligned} f(z) = \boldsymbol{c}^T &\left\{\left[\sum_{\mu=0}^{N_t-1}\check{\boldsymbol{b}}_{Q-1}(z_\mu)\right] \odot \check{\boldsymbol{b}}_{Q-1}(z)\right\} \\ &+ \boldsymbol{c}^H\left\{\left[\sum_{\mu=0}^{N_t-1}\check{\boldsymbol{b}}_{Q-1}(z_\mu^{-1})\right] \odot \check{\boldsymbol{b}}_{Q-1}(z^{-1})\right\}, \end{aligned} \tag{7}$$

where $\boldsymbol{c}$ is a length-$(Q-1)$ vector with its elements given by $[\boldsymbol{c}]_q = \sum_{j-i=q+1}[\hat{\boldsymbol{R}}_{\boldsymbol{YY}}]_{i,j}$. Taking the first derivative of $f(z)$ with respect to z yields

$$f'(z) = z^{-1}[g(z) - g^*(z)], \tag{8}$$

where

$$g(z) = \boldsymbol{c}^T\left\{\left[\sum_{\mu=0}^{N_t-1}\check{\boldsymbol{b}}_{Q-1}(z_\mu)\right] \odot \check{\boldsymbol{b}}_{Q-1}(z) \odot \boldsymbol{q}\right\},\ \boldsymbol{q} = [1, 2, \cdots, Q-1]^T.$$

From (6), (7) and (8), it can be seen that the ML estimation of $\tilde{\varepsilon}$ can be indirectly obtained through the roots of $f'(z) = 0$.

Calculating the roots of $f'(z) = 0$ is somewhat complicated. In the following, we try to find an approach without the operation of direct polynomial root-calculating. Assume that the channel taps remain constant during the training period and P is not less than the maximum delay spread of the channels, i.e., $P \geq L$, and the channel energy is mainly concentrated in the first M taps. Then, with our CBTS training sequences, we have (see Appendix I for details)

$$g^*(z) = z^{-Q} g(z)\kappa(q), \quad (9)$$

where $\kappa(q) = (q+1)[\boldsymbol{c}]_q^* / [(Q-q-1)[\boldsymbol{c}]_{Q-q-2}]$. By exploiting (9), the polynomial $f'(z)$ can be decomposed as

$$f'(z) = z^{-(Q+1)} g(z)[z^Q - \kappa(q)]. \quad (10)$$

With the definition of z, we have $z^{-(Q+1)} \neq 0$. Let $\tilde{z} = e^{j2\pi\tilde{\varepsilon}/Q}$ and assume $Q > N_t$. Then, with (32) as shown in Appendix I and the CBTS training sequences, we have (see Appendix II for details)

$$g(\tilde{z}) > 0 \ \&\ f'(\tilde{z}) = 0. \quad (11)$$

Accordingly, by solving the following relatively simple polynomial equation $z^Q - \kappa(q) = 0$, the desired CFO can be obtained as

$$\hat{\varepsilon} = \arg\max_{\varepsilon \in \{\hat{\varepsilon}_{q'} | 0 \leq q' \leq Q-1\}} \{f(z)\}, \quad (12)$$

where $\hat{\varepsilon}_{q'} = 1/(2\pi) \cdot \arg\{\kappa(q)\} + q' - Q/2$.

The computational load of the proposed estimator mainly involves $(1/2 \cdot N_r N Q + 5/2 \cdot (Q-1)Q)$ complex additions and $(1/2 \cdot N_r N Q + 4(Q-1)Q)$ complex multiplications. In comparison with the direct CFO estimate from (6), which requires an exhaustive search over a large set of frequency grids, our proposed estimator's computational complexity is significantly lower.

IV. Performance Analysis

A. The Variance of the CFO Estimation

With the definition of $\hat{\boldsymbol{R}}_{\boldsymbol{YY}}$, we have

$$\hat{\boldsymbol{R}}_{\boldsymbol{YY}} \doteq N_r P \sigma_x^2 \boldsymbol{B}(\tilde{\varepsilon})\boldsymbol{B}^H(\tilde{\varepsilon}) + \hat{\boldsymbol{R}}_{\boldsymbol{YW}} + \hat{\boldsymbol{R}}_{\boldsymbol{WW}}, \quad (13)$$

where

$$\sigma_x^2 = E[|[\boldsymbol{x}^{(\nu,\mu)}]_p|^2],$$

$$\hat{\boldsymbol{R}}_{\boldsymbol{YW}} = \boldsymbol{B}(\tilde{\varepsilon})\boldsymbol{X}\boldsymbol{W}^H + \boldsymbol{W}\boldsymbol{X}^H\boldsymbol{B}^H(\tilde{\varepsilon}),\ \hat{\boldsymbol{R}}_{\boldsymbol{WW}} = \boldsymbol{W}\boldsymbol{W}^H.$$

Assume

$$E\{[\boldsymbol{h}_L^{(\nu,\mu)}]_l^* [\boldsymbol{W}]_{i,j}\} = 0$$

$$E\{[\boldsymbol{h}_L^{(\nu,\mu)}]_l^* [\boldsymbol{h}_L^{(\nu',\mu')}]_{l'}\} = 0,\ \forall(\nu,\mu) \neq (\nu',\mu').$$

Then, for $i \neq j$ and $i' \neq j'$, we have

$$E[[\hat{\boldsymbol{R}}_{\boldsymbol{YW}}]_{i,j}] = E[[\hat{\boldsymbol{R}}_{\boldsymbol{WW}}]_{i,j}] = E[[\hat{\boldsymbol{R}}_{\boldsymbol{YW}}]_{i,j}^* [\hat{\boldsymbol{R}}_{\boldsymbol{WW}}]_{i',j'}] = 0, \quad (14)$$

$$E[|[\hat{\boldsymbol{R}}_{\boldsymbol{YW}}]_{i,j}|^2] = 2N_t N_r P \sigma_x^2 \sigma_w^2,\ E[|[\hat{\boldsymbol{R}}_{\boldsymbol{WW}}]_{i,j}|^2] = N_r P \sigma_w^4. \quad (15)$$

Define $[\boldsymbol{\alpha}]_q = \sum_{j-i=q+1} [\hat{\boldsymbol{R}}_{\boldsymbol{YW}}]_{i,j}$, $[\boldsymbol{\beta}]_q = \sum_{j-i=q+1} [\hat{\boldsymbol{R}}_{\boldsymbol{WW}}]_{i,j}$. Then, we obtain

$$[\boldsymbol{c}]_q = N_r P \sigma_x^2 (Q-q-1)\tilde{z}^{-q-1} \sum_{\mu=0}^{N_t-1} z_\mu^{-q-1} + [\boldsymbol{\alpha}]_q + [\boldsymbol{\beta}]_q. \quad (16)$$

With the definition of $\kappa(q)$, we further obtain

$$\kappa(q) = \frac{\tilde{z}^Q + \xi}{|\tilde{z}^{-Q+q+1} + [\boldsymbol{\eta}]_q|^2}, \quad (17)$$

where

$$\xi = \tilde{z}^{Q-q-1}[\boldsymbol{\zeta}]_q + \tilde{z}^{q+1}[\boldsymbol{\eta}]_q^* + [\boldsymbol{\zeta}]_q[\boldsymbol{\eta}]_q^*,$$

$$[\boldsymbol{\zeta}]_q = \{[\boldsymbol{\alpha}]_q^* + [\boldsymbol{\beta}]_q^*\} / [N_r P (Q-q-1)\sigma_x^2 \sum_{\mu=0}^{N_t-1} z_\mu^{q+1}],$$

$$[\boldsymbol{\eta}]_q = \{[\boldsymbol{\alpha}]_{Q-q-2} + [\boldsymbol{\beta}]_{Q-q-2}\} / [N_r P (q+1)\sigma_x^2 \sum_{\mu=0}^{N_t-1} z_\mu^{q+1}].$$

Note that the variances of $[\boldsymbol{\zeta}]_q$ and $[\boldsymbol{\eta}]_q$ are given by

$$\mathrm{var}\{[\boldsymbol{\zeta}]_q\} = \frac{2N_t\gamma_{xw}^{-1} + \gamma_{xw}^{-2}}{N_r P (Q-q-1)|\sum_{\mu=0}^{N_t-1} z_\mu^{q+1}|^2}, \quad (18)$$

$$\mathrm{var}\{[\boldsymbol{\eta}]_q\} = \frac{2N_t\gamma_{xw}^{-1} + \gamma_{xw}^{-2}}{N_r P (q+1)|\sum_{\mu=0}^{N_t-1} z_\mu^{q+1}|^2}, \quad (19)$$

where $\gamma_{xw} = \sigma_x^2/\sigma_w^2$. Without loss of generality, when $\gamma_{xw} \gg 1$, we have

$$\mathrm{var}\{[\boldsymbol{\zeta}]_q\}, \mathrm{var}\{[\boldsymbol{\eta}]_q\} \gg \mathrm{var}\{[\boldsymbol{\zeta}]_q[\boldsymbol{\eta}]_q^*\}. \quad (20)$$

Then, the variance of ξ can be obtained as

$$\mathrm{var}\{\xi\} \doteq \frac{2[N_t Q + \rho(q)]\gamma_{xw}^{-1} + Q\gamma_{xw}^{-2}}{N_r P (q+1)(Q-q-1)|\sum_{\mu=0}^{N_t-1} z_\mu^{q+1}|^2}, \quad (21)$$

where

$$\rho(q) = \begin{cases} 2(q+1)\frac{\Re\{(\sum_{\mu=0}^{N_t-1} z_\mu^{2(q+1)})(\sum_{\mu=0}^{N_t-1} z_\mu^{-q-1})^2\}}{|\sum_{\mu=0}^{N_t-1} z_\mu^{q+1}|^2}, & 0 \leq q \leq Q/2-1 \\ 2(Q-q-1)\frac{\Re\{(\sum_{\mu=0}^{N_t-1} z_\mu^{2(q+1)})(\sum_{\mu=0}^{N_t-1} z_\mu^{q+1})^2\}}{|\sum_{\mu=0}^{N_t-1} z_\mu^{q+1}|^2}, & Q/2 \leq q \leq Q-2 \end{cases}.$$

When $\mathrm{var}\{\xi\} \ll 1$, from (17) we obtain

$$\begin{aligned} \hat{\varepsilon} &= \tilde{\varepsilon} + \frac{1}{2\pi}\arg\{1 + e^{-j2\pi\tilde{\varepsilon}}\xi\} \\ &\doteq \tilde{\varepsilon} + \frac{1}{2\pi}\Im\{e^{-j2\pi\tilde{\varepsilon}}\xi\}. \end{aligned} \quad (22)$$

Then, we have

$$\mathrm{var}\{\hat{\varepsilon}\} \doteq \frac{2[N_t Q + \rho(q)]\gamma_{xw}^{-1} + Q\gamma_{xw}^{-2}}{8\pi^2 N_r P (q+1)(Q-q-1)|\sum_{\mu=0}^{N_t-1} z_\mu^{q+1}|^2}. \quad (23)$$

It follows immediately that $\mathrm{var}\{\hat{\varepsilon}\}$ depends on q for fixed $\{i_\mu\}_{\mu=0}^{N_t-1}$, N_t, N_r, P, Q and γ_{xw}. The choice of q can be optimized based on (23).

B. The Cramer-Rao Bound

Define $\boldsymbol{\theta} = [\tilde{\varepsilon}, \Re(\boldsymbol{h}^T), \Im(\boldsymbol{h}^T)]^T$. Then, the Fisher information matrix can be obtained [19] as

$$\mathcal{F} = -E\left[\frac{\partial^2 \ln p(\boldsymbol{y}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right] = \frac{2N}{\sigma_w^2}\begin{bmatrix} \alpha & \boldsymbol{\gamma}^T \\ \boldsymbol{\gamma} & \boldsymbol{\Xi} \end{bmatrix}, \quad (24)$$

where

$$\alpha = \frac{4\pi^2}{N^2}\boldsymbol{h}^H \mathcal{X}^H \mathcal{B}^2 \mathcal{X}\boldsymbol{h},$$

$$\gamma = \frac{2\pi}{N}\left[\Im[\boldsymbol{h}^H\mathcal{X}^H\mathcal{B}\mathcal{X}], \Re[\boldsymbol{h}^H\mathcal{X}^H\mathcal{B}\mathcal{X}]\right]^T,$$

$$\boldsymbol{\Xi} = \begin{bmatrix} \Re[\mathcal{X}^H\mathcal{X}] & -\Im[\mathcal{X}^H\mathcal{X}] \\ \Im[\mathcal{X}^H\mathcal{X}] & \Re[\mathcal{X}^H\mathcal{X}] \end{bmatrix},$$

$$\mathcal{X} = \boldsymbol{I}_{N_r} \otimes \boldsymbol{S},$$

$$\mathcal{B} = \boldsymbol{I}_{N_r} \otimes \mathrm{diag}\{[N_g, N_g+1, \cdots, N_g+N-1]^T\}.$$

By straight-forward calculation, the CRB for the CFO estimation is obtained

$$\mathrm{CRB} = \frac{N\sigma_w^2}{8\pi^2\boldsymbol{h}^H\mathcal{X}^H\mathcal{B}[\boldsymbol{I}_{N_rN} - \mathcal{X}(\mathcal{X}^H\mathcal{X})^{-1}\mathcal{X}^H]\mathcal{B}\mathcal{X}\boldsymbol{h}}. \quad (25)$$

For frequency selective fading channels, the above CRB represents the bound for CFO estimation with particular (snapshot) channel realization. To obtain the average CRB, we simply average the above CRB over independent channel realizations. Note that the average CRB for CFO estimation corresponds to the extended Miller and Chang bound (EMCB) [20], [21].

V. Simulation Results

A number of simulations are executed to evaluate the performance of the proposed CFO estimator with the CBTS training sequences. Throughout the simulations, a MIMO OFDM system of bandwidth 20MHz operating at 5GHz with $N = 1024$ and $N_g = 64$ is used. For the training sequences, we set $P = 64$ and $Q = 16$. Each channel has four independent Rayleigh fading taps, whose relative average-powers and propagation delays are $\{0, -9.7, -19.2, -22.8\}$dB and $\{0, 0.1, 0.2, 0.4\}\mu$s, respectively. The normalized CFO $\tilde{\varepsilon}$ is generated within the range $(-Q/2, Q/2]$. In the sequel, unless otherwise stated, we assume $N_t = 3$ and $N_r = 2$.

In Figs. 1 and 2, we show the variance of the proposed CFO estimator as a function of q with $\{i_\mu\}_{\mu=0}^{N_t-1} = \{3, 5, 11\}$ and $\{3, 7, 14\}$, respectively. The solid and dotted curves are the results from Monte Carlo simulations and analysis, respectively. We observe that the results from analysis agree quite well with those from simulations except at some non-optimal q values. We also observe that the variance of the estimation achieves its minimum for $q = 5, 7, 9$ with $\{i_\mu\}_{\mu=0}^{N_t-1} = \{3, 5, 11\}$ and for $q = 6, 8$ with $\{i_\mu\}_{\mu=0}^{N_t-1} = \{3, 7, 14\}$. These observations imply that we can obtain the optimal value of q from the analytical results after $\{i_\mu\}_{\mu=0}^{N_t-1}$ is determined.

Depicted in Fig. 3 is the performance comparison between the proposed CFO estimator ($q = 6$, $\{i_\mu\}_{\mu=0}^{N_t-1} = \{3, 7, 14\}$) and the estimator in [15]. In order to substantiate that the training sequences constructed from the Chu sequence do help to improve the estimation accuracy, the performance of

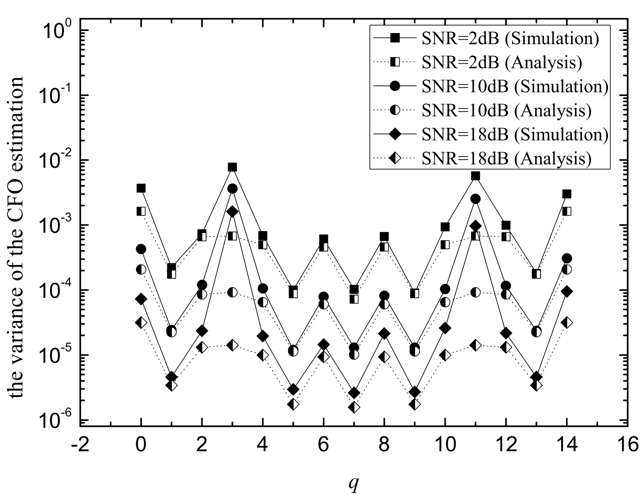

Fig. 1. The variance of the proposed CFO estimator as a function of q with $\{i_\mu\}_{\mu=0}^{N_t-1} = \{3, 5, 11\}$.

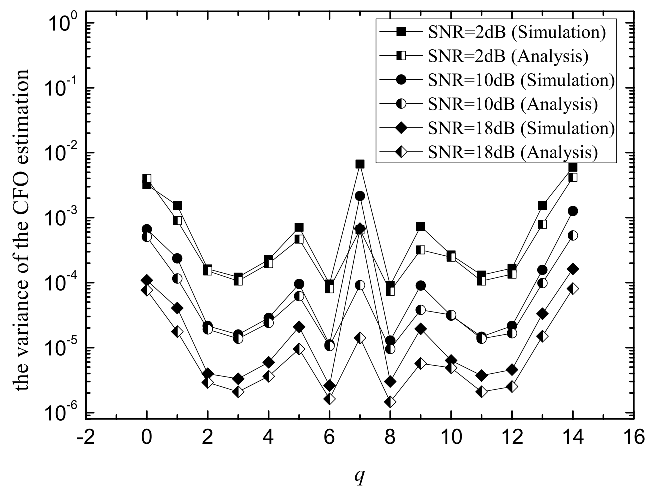

Fig. 2. The variance of the proposed CFO estimator as a function of q with $\{i_\mu\}_{\mu=0}^{N_t-1} = \{3, 7, 14\}$.

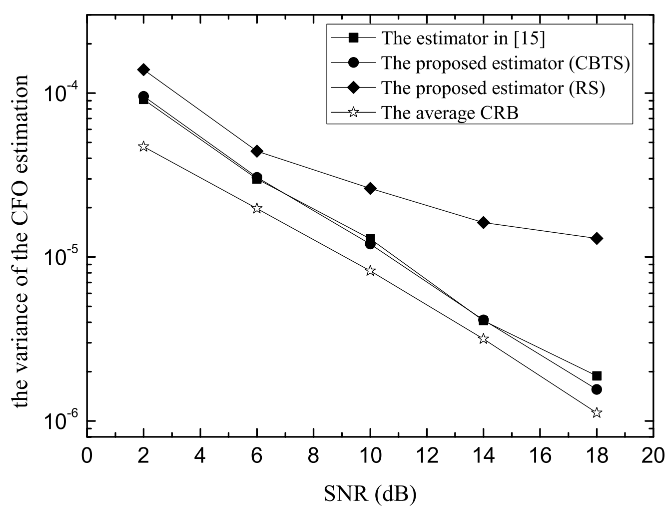

Fig. 3. CFO estimation performance for the different CFO estimators.

the proposed estimator with random sequences (RS), whose elements are generated randomly, is included. Also included as a performance benchmark is the average CRB. It can be observed that the performance of the proposed estimator with CBTS is almost the same as in [15] and approaches the average CRB, which verify its high estimation accuracy. It can also be observed that the performance of the proposed CFO estimator with CBTS is far better than that with RS, which should be attributed to the good correlation property of our CBTS training sequences.

VI. Conclusions

In this paper, we have presented a low complexity training sequence assisted CFO estimator for MIMO OFDM systems over frequency selective fading channels. We have developed the variance of the CFO estimation and the CRB to evaluate the estimation performance. By exploiting the optimized parameter from the estimation variance, our CFO estimator with CBTS yields good performance.

Acknowledgment

This work was supported in part by National Natural Science Foundation of China under Grant 60496311, and in part by the Erik Jonsson School Research Excellence Initiative, the University of Texas at Dallas, USA.

Appendix I

This appendix presents the proof of (9). It follows immediately that the polynomials on both sides of (9) are $(Q-1)$-degree. Therefore, we only need to prove that the corresponding polynomial coefficients are equal.

Without considering noise, $\hat{\boldsymbol{R}}_{YY}$ can be expressed as $\hat{\boldsymbol{R}}_{YY} = \boldsymbol{B}(\tilde{\varepsilon})\hat{\boldsymbol{R}}_{XX}\boldsymbol{B}^H(\tilde{\varepsilon})$, where $\hat{\boldsymbol{R}}_{XX} = \boldsymbol{X}\boldsymbol{X}^H$. Define

$$\varpi_{\mu,\mu'} = (i_\mu - i_{\mu'})/Q,\ \boldsymbol{s}_P^{(\mu)} = \sqrt{N_t/Q}\boldsymbol{F}_P^H\tilde{\boldsymbol{s}}_P^{(\mu)}.$$

Then, with the assumptions that $P \geq L$ and the channel taps remain constant during the training period, we obtain

$$[\hat{\boldsymbol{R}}_{XX}]_{\mu,\mu'} = \frac{N}{N_t}\sum_{\nu=0}^{N_r-1}\sum_{l=0}^{L-1}\sum_{l'=0}^{L-1}\left\{h_l^{(\nu,\mu)}h_{l'}^{(\nu,\mu'),*}[\boldsymbol{A}^{(\mu,\mu')}]_{l,l'}\right\}, \qquad (26)$$

where $[\boldsymbol{A}^{(\mu,\mu')}]_{l,l'} = [\boldsymbol{s}_P^{(\mu)}]^{(l),T}\boldsymbol{D}_P(\varpi_{\mu,\mu'}Q)[\boldsymbol{s}_P^{(\mu')}]^{(l'),*}$. With the aim of complexity reduction, we replace $[\hat{\boldsymbol{R}}_{XX}]_{\mu,\mu'}$ with its expected value. Then, exploiting the good correlation property of CBTS which is inherited from the Chu sequence, we obtain

$$[\hat{\boldsymbol{R}}_{XX}]_{0,0} = [\hat{\boldsymbol{R}}_{XX}]_{1,1} = \cdots = [\hat{\boldsymbol{R}}_{XX}]_{N_t-1,N_t-1}. \qquad (27)$$

Let $[\boldsymbol{s}_P]_p = e^{j\pi\upsilon p^2/P}$ with υ being coprime with P. Define $p_{\mu,l} = \mu M + l$. Then, we have

$$[\boldsymbol{A}^{(\mu,\mu')}]_{l,l'} = (-1)^{\upsilon(p_{\mu,l}-p_{\mu',l'})+1}e^{j\pi\upsilon(p_{\mu,l}^2-p_{\mu',l'}^2)/P}e^{-j\pi(P-1)[\upsilon(p_{\mu,l}-p_{\mu',l'})-\varpi_{\mu,\mu'}]/P} \times\sin(\pi\varpi_{\mu,\mu'})/\sin\{\pi[\upsilon(p_{\mu,l}-p_{\mu',l'})-\varpi_{\mu,\mu'}]/P\}. \qquad (28)$$

It follows immediately that

$$|[\boldsymbol{A}^{(\mu,\mu')}]_{l,l'}|_{p_{\mu,l}-p_{\mu',l'}\neq 0} \ll |[\boldsymbol{A}^{(\mu,\mu')}]_{l,l'}|_{p_{\mu,l}-p_{\mu',l'}=0}. \qquad (29)$$

Therefore, with the assumption that the channel energy is mainly concentrated in the first M taps, $|[\hat{\boldsymbol{R}}_{XX}]_{\mu,\mu'}|_{\mu\neq\mu'}$ can be made very small with CBTS, which yields

$$|[\hat{\boldsymbol{R}}_{XX}]_{\mu,\mu'}|_{\mu\neq\mu'} \ll [\hat{\boldsymbol{R}}_{XX}]_{\mu,\mu}. \qquad (30)$$

Then, we have $\hat{\boldsymbol{R}}_{XX} \doteq N_rP\sigma_x^2\boldsymbol{I}_{N_t}$, where $\sigma_x^2 = E[|[\boldsymbol{x}^{(\nu,\mu)}]_p|^2]$. It follows immediately that $\hat{\boldsymbol{R}}_{YY} \doteq N_rP\sigma_x^2\boldsymbol{B}(\tilde{\varepsilon})\boldsymbol{B}^H(\tilde{\varepsilon})$. Then, with the definition of $[\boldsymbol{c}]_q$, we have

$$[\boldsymbol{c}]_q \doteq N_rP\sigma_x^2(Q-q-1)\tilde{z}^{-q-1}\sum_{\mu=0}^{N_t-1}z_\mu^{-q-1}, \qquad (31)$$

where $\tilde{z} = e^{j2\pi\tilde{\varepsilon}/Q}$. By substituting the above result into (9), the polynomial coefficient corresponding to z^{-q} at both sides of (9) can be calculated to be

$$N_rP\sigma_x^2q(Q-q)\tilde{z}^q\sum_{\mu=0}^{N_t-1}\sum_{\mu'=0}^{N_t-1}\left(z_\mu/z_{\mu'}\right)^q, \qquad (32)$$

where we have utilized the following property $z_\mu^Q = 1$. This completes the proof.

Appendix II

In this appendix, we will prove (11). With CBTS and (32), we can obtain

$$g(\tilde{z}) = N_rP\sigma_x^2(\boldsymbol{q}\odot\boldsymbol{\Psi})^H[(Q-\boldsymbol{q})\odot\boldsymbol{\Psi}], \qquad (33)$$

where

$$\boldsymbol{\Psi} = [\Psi(1),\Psi(2),\cdots,\Psi(Q-1)]^T,\ \Psi(q) = e^{-j2\pi qi_0/Q}\sum_{\mu=0}^{N_t-1}e^{j2\pi qi_\mu/Q}.$$

From the definition of $\boldsymbol{\Psi}$, we have $\boldsymbol{\Psi} = \boldsymbol{\Phi}\boldsymbol{1}_{N_t}$, where

$$\boldsymbol{\Phi} = [\boldsymbol{\phi}_0,\boldsymbol{\phi}_1,\cdots,\boldsymbol{\phi}_{N_t-1}],$$

$$\boldsymbol{\phi}_\mu = [e^{j2\pi(i_\mu-i_0)/Q},e^{j4\pi(i_\mu-i_0)/Q},\cdots,e^{j2(Q-1)\pi(i_\mu-i_0)/Q}]^T.$$

Since $\boldsymbol{\Phi}$ is a Vandermonde matrix, it is of full-rank (rank N_t) with $Q > N_t$. Therefore, $\boldsymbol{\Psi}$ cannot be the all zero vector and then $g(\tilde{z}) > 0$, $f'(\tilde{z}) = 0$. This completes the proof.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, pp. 191–193, Feb./Mar./Apr. 1995.

[3] 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, pp. 1800–1805, July 1997.

[4] U. Tureli, H. Liu, and M. D. Zoltowski, "OFDM blind carrier offset estimation: ESPRIT," *IEEE Trans. Commun.*, vol. 48, pp. 1459–1461, Sept. 2000.

[5] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, pp. 2908–2914, Oct. 1994.

[6] T. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, pp. 1613–1621, Dec. 1997.

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

[8] D. Huang and K. B. Letaief, "Carrier frequency offset estimation for OFDM systems using null subcarriers," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 813–823, May 2006.

[9] H. Minn and P. Tarasak, "Improved maximum likelihood frequency offset estimation based on likelihood metric design," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 2076–2086, June 2006.

[10] Y. Yao and G. B. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO and multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 173–183, Jan. 2005.

[11] A. N. Mody and G. L. Stuber, "Synchronization for MIMO OFDM systems," in *Proc. IEEE Globecom'01*, vol. 1, Nov. 2001, pp. 509–513.

[12] O. Besson and P. Stoica, "On parameter estimation of MIMO flat-fading channels with frequency offsets," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 602–613, Mar. 2003.

[13] X. Ma, M. K. Oh, G. B. Giannakis, and D. J. Park, "Hopping pilots for estimation of frequency-offset and multi-antenna channels in MIMO OFDM," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 162–172, Jan. 2005.

[14] F. Simoens and M. Moeneclaey, "Reduced complexity data-aided and code-aided frequency offset estimation for flat-fading MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1558–1567, June 2006.

[15] Y. X. Jiang, X. Q. Gao, X. H. You, and W. Heng, "Training sequence assisted frequency offset estimation for MIMO OFDM," in *Proc. IEEE ICC'06*, vol. 12, June 2006, pp. 5371–5376.

[16] Y. X. Jiang, H. Minn, X. Q. Gao, X. H. You, and Y. Li, "Frequency offset estimation and training sequence design for MIMO OFDM," *to appear in IEEE Trans. Wireless Commun.*

[17] D. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inform. Theory*, vol. 18, pp. 531–532, July 1972.

[18] W. H. Mow, "A new unified construction of perfect root-of-unity sequences," in *Proc. IEEE Spread-Spectrum Techniques and Applications*, vol. 3, Sept. 1996, pp. 955–959.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Englewood Cliffs ed. NJ: Prentical-Hall, 1993.

[20] 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.

[21] 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.