A Consistent OFDM Carrier Frequency Offset Estimator Based on Distinctively Spaced Pilot Tones

Jing Lei and Tung-Sang Ng, *Fellow, IEEE*

Abstract—A pilot-tone-based maximum likelihood estimator (PBMLE) for carrier frequency offset (CFO) in orthogonal frequency-division multiplexing (OFDM) systems is proposed. To obtain a consistent estimate of the CFO over a frequency-selective fading channel, the proposed method employs a preamble comprising distinctively spaced pilot tones. As a result of this preamble configuration, a large estimation range equal to the bandwidth of the OFDM signal can be achieved. Different from previous *ad hoc* pilot-tone-based CFO estimators, the PBMLE exploits the relationship between the CFO and the periodogram of the preamble. Analysis shows that the proposed PBMLE is asymptotically unbiased and efficient. To realize this PBMLE in practice, a suboptimal estimator is also introduced, in which a zero-padded fast Fourier transform is invoked and the CFO estimation is split into two phases: coarse and fine estimation. Coarse estimation is obtained through the correlation between the received preamble and its original pattern, whereas fine estimation is obtained by exploiting the magnitude attenuation in the vicinities of those CFO-shifted pilot tones. Both analytical investigations and computer simulations indicate that the accuracy of this simplified suboptimal estimator is proportional to the oversize ratio of zero padding. When the oversize ratio is sufficiently high, the performance of the suboptimal estimator approaches that of the proposed PBMLE.

Index Terms—Frequency synchronization, maximum likelihood estimate, orthogonal frequency-division multiplexing (OFDM), periodogram, pilot tone.

I. INTRODUCTION

AS an effective transmission scheme, orthogonal frequency-division multiplexing (OFDM) is robust against many channel impairments [1], [2], for instance, cochannel interference, frequency-selective fading, and impulsive parasitic noise. However, OFDM is vulnerable to the carrier frequency offset (CFO) arising from transceiver oscillator mismatches and/or Doppler shifts. Depending on the application, the CFO may amount to many tens of the subcarrier spacing, while even the CFO of a small fraction of the subcarrier spacing can cause serious performance degradation [3]. Therefore, frequency synchronization is one of the most important tasks performed by an OFDM receiver and has received significant attention in recent years [4]–[7], [9]–[24].

The authors are with the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong.

CFO estimators based on the CFO-induced phase rotation are presented in [4]–[7], [9]–[15], and [22]–[24]. The methods in [14] and [15] are bandwidth efficient as they do not require extra overhead and simply utilize the redundancy of the cyclic prefix (CP). However, they can only deal with a CFO less than half the subcarrier spacing and fail to render correct results when the length of CP is close to that of the multipath delay spread. Techniques in [4]–[6], [9]–[13], and [22]–[24] employ a preamble with repetitive time-domain structure, and their estimation range is proportional to the number of periods. Although the estimation range can be increased by shortening the period's duration, this is achieved at the cost of decreased estimation accuracy. Another problem associated with these techniques is that the preamble format is inflexible, so any offset beyond the predetermined estimation range cannot be tracked. The estimator in [7] exploits the second-order statistics [8] of the OFDM signal. To obtain a large estimation range, it relies on the use of subcarrier weighting and periodic precoding, which will result in the decrease of minimum constellation distance and the reduction of transmission power efficiency.

Based on the orthogonality between the guard bands (consecutive null subcarriers padded to the ends of OFDM data block) and the information-bearing subcarriers, subspace-based estimators are developed in [16]–[18]. However, the solution to their proposed cost function is channel dependent, which may lead to ambiguous estimates in the scenario of frequency-selective fading. In an attempt to solve this ambiguity, [19] discards the use of consecutive null subcarriers and distributes the null subcarriers to nonconsecutive locations. Unfortunately, if the pattern of channel nulls resulted from frequency selective fading is similar to that of the inserted null subcarriers, the modification proposed by [19] cannot yield consistent estimate for the true CFO either. Moreover, the guard band is usually adopted for the purpose of avoiding adjacent-channel interference (ACI), as shown in the HIPERLAN/2 and the IEEE802.11a standards; therefore, the application of [19] may not be justified in some practical situations. An additional difficulty with [19] lies in its realization, in which an adaptive algorithm is used to search for the minima of the cost function. Due to the nonconvexity of this cost function, the adaptive algorithm may not converge to the desired value when the initial estimate is far away from the true CFO.

The CFO estimators in [20] and [22] employ a preamble composed of pilot tones. The basic idea of these methods is to estimate the CFO by comparing the spectrum of the received preamble with the original pattern of the pilot tones. To cope with large CFO, nonuniform spacing of pilot tones [20] is shown to be a better choice than uniform spacing [22]. However, the influ-
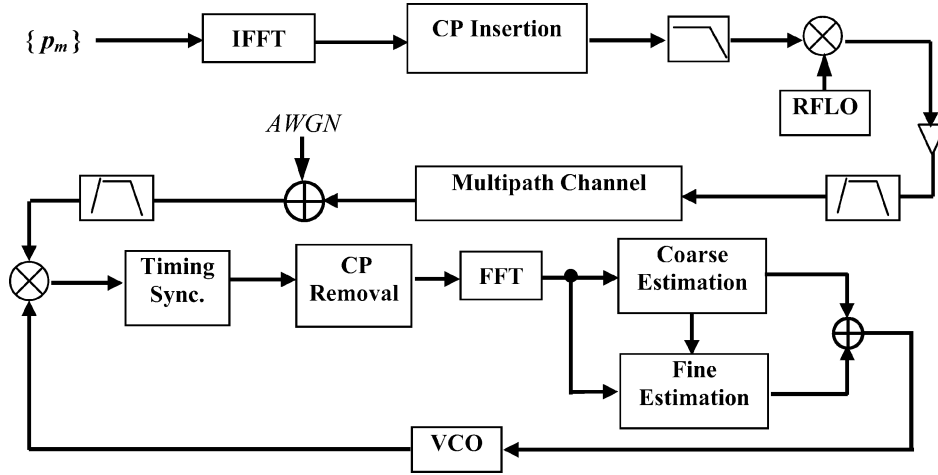
Fig. 1.   System model of the proposed PBMLE.

ence of frequency selectivity is not taken into account by these estimators, and the consistency of their CFO estimation is not guaranteed when the pilot tones undergo deep fading.

To obtain a consistent estimate of the true CFO over frequency-selective fading channel, this paper proposes a maximum likelihood estimator (MLE) based on a preamble comprising distinctively-spaced pilot tones. The proposed CFO estimator is robust against frequency selectivity because it is able to yield consistent result provided at least two pilot tones are not occupied by channel nulls. Different from previous pilot-tone based (PB) estimators [20], [22] that are *ad hoc*, the analytical expression of this PBMLE is systematically derived, which reveals the relationship between the CFO and the periodogram of the received preamble. To realize this estimator in practice, a suboptimal estimator is presented, which can be implemented by invoking a zero-padded fast Fourier transform (FFT). Furthermore, a fine estimator of CFO is developed by exploiting the relative magnitude attenuation around those CFO-shifted pilot tones. Analysis carried out in this paper indicates that the proposed estimator is asymptotically unbiased and efficient as its average bias approaches zero and its mean-square error approaches the Cramer–Rao bound (CRB).

The rest of this paper is organized as follows. The system model of the proposed PBMLE is described in Section II. The analytical expression for the PBMLE is derived in Section III. The consistency of this estimator and its practical realization are addressed in Section IV. A performance analysis is carried out in Section V, in which the estimation range and the statistical properties of the PBMLE are investigated. Computer simulation results are presented in Section VI. Finally, Section VII concludes this paper.

II. SYSTEM MODEL

We start with the system model of the proposed PBMLE, which is shown in Fig. 1. Let $\boldsymbol{P}_M = [p_0 \quad p_1 \quad \cdots \quad p_{M-1}]^T$ denote the preamble to be transmitted, where the superscript T denotes transpose. The preamble $\boldsymbol{P}_M$ comprises K pilot tones and $M-K$ nulls. The indexes of the pilot tones and the nulls are represented by sets $\{t_k\}_{k=1}^K$ and $\{n_k\}_{k=1}^{M-K}$, respectively. Thus, $|p_{n_k}| = 0$ and $\{t_k\}_{k=1}^K \cup \{n_k\}_{k=1}^{M-K} = \{0, \ 1, \ldots M-1\}$.

Without loss of generality, we assume $|p_{t_k}| = 1$. The spacing between the pilot tones is distinctive [19], which can be expressed as

$$t_{k_1} - t_{k_2} = t_{k_3} - t_{k_4} \Rightarrow k_1 = k_3 \text{ and } k_2 = k_4. \quad (1)$$

The motivation for such an arrangement is to guarantee consistency of the CFO estimator over frequency selective fading channel, which will be discussed in Section IV. Generally, a guard band is used to avoid adjacent-channel interference (ACI), so we assume subcarriers numbered from 0 to $M-1$ are employed for preamble transmission and subcarriers M to $N-1$ are used as the guard band, yielding $\boldsymbol{P}_N = [\boldsymbol{P}_M^T \quad p_{n_{M-K+1}} \quad p_{n_{M-K+2}} \quad \cdots \quad p_{n_{N-K}}]^T$, where $p_{n_{M-K+l}} = 0$ and $n_{M-K+l} = M + l, 1 \le l \le N - M$. OFDM modulation is implemented by invoking N-point IFFT over $\boldsymbol{P}_N$. Denote $\boldsymbol{W}_N$ as the $N \times N$ IFFT matrix, and $\boldsymbol{W}_K = [\boldsymbol{w}_{t_1} \quad \boldsymbol{w}_{t_2} \quad \cdots \quad \boldsymbol{w}_{t_K}]$ is the $N \times K$ submatrix of $\boldsymbol{W}_N$, where $\boldsymbol{w}_{t_i}$ is the t_ith column of $\boldsymbol{W}_N$. The resultant N time-domain samples can then be represented by

$$\begin{aligned} \boldsymbol{S} &= N^{-1}\boldsymbol{W}_N\boldsymbol{P}_N \\ &= N^{-1}\boldsymbol{W}_K\boldsymbol{P}_K \end{aligned} \quad (2)$$

where $\boldsymbol{P}_K = [p_{t_1} \quad p_{t_2} \quad \cdots \quad p_{t_K}]^T$.

To cope with the intersymbol interference (ISI) caused by multipath fading, a cyclic prefix (CP) of length L_{CP}, which is longer than the multipath delay spread, is inserted ahead of $\boldsymbol{S}$, yielding $[s_{N-L_{CP}} \quad s_{N+1-L_{CP}} \quad \cdots \quad s_{N-1} \quad \boldsymbol{S}^T]^T$. After that, these baseband samples are pulse shaped, up-converted to the radio frequency (RF), and transmitted over the multipath channel. Additive white Gaussian noise (AWGN) is also added into this channel. At the receiver, the signal is down converted by a local voltage-controlled oscillator (VCO) and f_d is assumed to be the normalized CFO, which equals the actual CFO (in Hertz) divided by Δf, where Δf denotes the spacing between adjacent subcarriers. Symbol timing synchronization (represented by the block of "Timing Sync." in Fig. 1) is carried out next, and the CP is removed. From the derivation in Section III, we shall see that the proposed method depends on the magnitude of pilot tones only, while a timing error will

simply phase shift the tones and thus will have no effect on the CFO estimate. In other words, the proposed method is insensitive to the timing synchronization error, as long as the timing offset plus the multipath delay spread does not exceed the length of CP so that no ISI is introduced. Finally, the received samples are given by

$$
\begin{aligned}
\boldsymbol{R} &= [\, r_0 \quad r_1 \quad \cdots \quad r_{N-1} \,]^T \\
&= N^{-1} e^{j\theta} \boldsymbol{\Omega}(f_d) \boldsymbol{W}_N \boldsymbol{H}_N \boldsymbol{P}_N + \boldsymbol{\xi} \\
&= N^{-1} e^{j\theta} \boldsymbol{\Omega}(f_d) \boldsymbol{W}_K \boldsymbol{H}_K \boldsymbol{P}_K + \boldsymbol{\xi}
\end{aligned}
\tag{3}
$$

where θ is a constant representing the phase shift, $\boldsymbol{\Omega}(f_d) = \mathrm{diag}\left\{ 1 \quad e^{j2\pi N^{-1} f_d} \quad \cdots \quad e^{j2\pi N^{-1}(N-1)f_d} \right\}$ is a diagonal matrix depending on f_d, $\boldsymbol{H}_N = \mathrm{diag}\{ h_1 \quad h_2 \quad \ldots \quad h_N \}$ is a diagonal matrix with h_i representing the channel gain on the ith subcarrier, $\boldsymbol{H}_K = \mathrm{diag}\{ h_{t_1} \quad h_{t_2} \quad \cdots \quad h_{t_K} \}$ is the submatrix of $\boldsymbol{H}_N$ which only includes the channel gains on the subcarriers of pilot tones, and vector $\boldsymbol{\xi} = [\, \varepsilon_0 \quad \varepsilon_1 \quad \ldots \quad \varepsilon_{N-1} \,]^T$ denotes complex white Gaussian noise with zero mean and covariance matrix

$$
\boldsymbol{C_\xi} = E\{\boldsymbol{\xi}\boldsymbol{\xi}^H\} = \sigma^2 \boldsymbol{I}_N
\tag{4}
$$

where the superscript H represents Hermitian and $\boldsymbol{I}_N$ is an $N \times N$ identity matrix.

III. DERIVATION OF PBMLE

Matrices $\boldsymbol{\Omega}(f_d)$ and $\boldsymbol{H}_N$ are both unknown but deterministic. Let $\boldsymbol{\Xi} = \boldsymbol{\Omega}(f_d)\boldsymbol{W}_K$ and $\boldsymbol{B}_K = \boldsymbol{H}_K \boldsymbol{P}_K$. Given $\boldsymbol{P}_K$, $\boldsymbol{B}_K$ is a linear function of $\boldsymbol{H}_K$, and $\boldsymbol{\Xi}$ is determined solely by f_d, which can be rewritten as

$$
\boldsymbol{\Xi}(f_d) = [\, \boldsymbol{e}_{t_1}(f_d) \quad \boldsymbol{e}_{t_2}(f_d) \quad \cdots \quad \boldsymbol{e}_{t_K}(f_d) \,]
\tag{5}
$$

where $\boldsymbol{e}_{t_k}(f_d) = \left[\, 1 \quad e^{j2\pi N^{-1}(t_k+f_d)} \quad \cdots \quad e^{j2\pi N^{-1}(N-1)(t_k+f_d)} \,\right]^T$. For a fixed pair of $(f_d, \boldsymbol{H}_K)$, the probability density function of $\boldsymbol{R}$ conditioned on $\boldsymbol{\Xi}(f_d)$ and $\boldsymbol{B}_K$ is

$$
\begin{aligned}
&f\{\boldsymbol{R} \mid \boldsymbol{\Xi}(f_d), \boldsymbol{B}_K\} \\
&= (\pi\sigma^2)^{-N} \exp\left\{ -\sigma^{-2} \left\| \boldsymbol{R} - N^{-1}e^{j\theta}\boldsymbol{\Xi}(f_d)\boldsymbol{B}_K \right\|^2 \right\}
\end{aligned}
\tag{6}
$$

where $\| \cdot \|^2$ is the Euclidean norm. Taking $\boldsymbol{\Lambda}(f_d, \boldsymbol{B}_K) = f\{\boldsymbol{R} \mid \boldsymbol{\Xi}(f_d), \boldsymbol{B}_K\}$ as the likelihood function of f_d and $\boldsymbol{B}_K$, their joint maximum likelihood (ML) estimates can be obtained by

$$
(f_d^\#, \boldsymbol{B}_K^\#) = \underset{(\hat{f}_d, \hat{\boldsymbol{B}}_K)}{\arg\max} \left\{ \boldsymbol{\Lambda}(\hat{f}_d, \hat{\boldsymbol{B}}_K) \right\}
\tag{7}
$$

where $f_d^\#$ and $\boldsymbol{B}_K^\#$ stand for the ML estimates of f_d and $\boldsymbol{B}_K$, while $\hat{f}_d$ and $\hat{\boldsymbol{B}}_K$ denote the trial values of f_d and $\boldsymbol{B}_K$, respectively. Since f_d and $\boldsymbol{B}_K$ are decoupled [27], [28], estimating f_d and $\boldsymbol{B}_K$ is a separable estimation problem that can be solved in two steps. Firstly, $\hat{f}_d$ is regarded as a constant and it is found that $\boldsymbol{\Lambda}(\hat{f}_d, \hat{\boldsymbol{B}}_K)$ reaches its maxima when

$$
\hat{\boldsymbol{B}}_K = \boldsymbol{\Xi}^H(\hat{f}_d) \boldsymbol{R}.
\tag{8}
$$

Secondly, (8) is substituted into $\boldsymbol{\Lambda}(\hat{f}_d, \hat{\boldsymbol{B}}_K)$ and the likelihood function becomes

$$
\begin{aligned}
&\boldsymbol{\Lambda}\left(\hat{f}_d, \hat{\boldsymbol{B}}_K\right)\Big|_{\hat{\boldsymbol{B}}_K = \boldsymbol{\Xi}^H(\hat{f}_d)\boldsymbol{R}} \\
&= (\pi\sigma^2)^{-N}\exp\left\{ \sigma^{-2}\left[N^{-1}\boldsymbol{R}^H\boldsymbol{\Xi}\left(\hat{f}_d\right)\boldsymbol{\Xi}^H\left(\hat{f}_d\right)\boldsymbol{R} - \boldsymbol{R}^H\boldsymbol{R} \right] \right\}.
\end{aligned}
\tag{9}
$$

As the right-hand side of (9) is independent of $\hat{\boldsymbol{B}}_K$, it coincides with the marginal likelihood function of $\hat{f}_d$. Obviously, maximizing (9) is equivalent to maximizing the inner term of

$$
O\left(\hat{f}_d\right) = \boldsymbol{R}^H \boldsymbol{\Xi}\left(\hat{f}_d\right) \boldsymbol{\Xi}^H\left(\hat{f}_d\right) \boldsymbol{R}.
\tag{10}
$$

Therefore, the ML estimation of $f_d^\#$ is given by

$$
f_d^\# = \underset{\hat{f}_d}{\arg\max} \left\{ O\left(\hat{f}_d\right) \right\}.
\tag{11}
$$

Furthermore, it is worth noting that different columns of $\boldsymbol{\Xi}(\hat{f}_d)$ are orthogonal, i.e.,

$$
\boldsymbol{e}_{t_{k_1}}^H\left(\hat{f}_d\right) \boldsymbol{e}_{t_{k_2}}\left(\hat{f}_d\right) = N\delta(t_{k_1} - t_{k_2}), \quad k_1, k_2 \in \{1, 2 \cdots K\}
\tag{12}
$$

where $\delta(\cdot)$ is the *Kronecker delta* function. Define $\boldsymbol{P}_{\boldsymbol{\Xi}}(\hat{f}_d) = \boldsymbol{\Xi}(\hat{f}_d)\boldsymbol{\Xi}^H(\hat{f}_d)$ and $\boldsymbol{P}_{\boldsymbol{\Xi}}(\hat{f}_d)$ can be decomposed into

$$
\boldsymbol{P}_{\boldsymbol{\Xi}}\left(\hat{f}_d\right) = \sum_{k=1}^{K} \boldsymbol{e}_{t_k}\left(\hat{f}_d\right) \boldsymbol{e}_{t_k}^H\left(\hat{f}_d\right).
\tag{13}
$$

The periodogram of signal $\boldsymbol{R}$ is given by

$$
\boldsymbol{\rho}(f) = \left| \sum_{m=0}^{N-1} \boldsymbol{r}_m e^{-j2\pi N^{-1}mf} \right|^2.
\tag{14}
$$

Substituting (13) and (14) into (11) yields

$$
f_d^\# = \arg\max_{\hat{f}_d} \left\{ \sum_{k=1}^{K} \rho\left(\hat{f}_{d,k}\right) \right\}
\tag{15}
$$

where $\hat{f}_{d,k} = t_k + \hat{f}_d$. The relationship between the periodogram of the received preamble $\boldsymbol{R}$ and the maximum likelihood estimates of f_d is therefore clearly established. When $N = K = 1$, (15) reduces to the MLE#1 method of [26], which is proposed for a single carrier system. An expression similar to (15) appears in [21], but the preamble is replaced by the data block and the channel frequency response matrix $\boldsymbol{H}_N$ is required to be positive definite. Even in the case when channel nulls are not considered, the method in [21] cannot produce a consistent estimate for f_d when $|h_i| \neq \mathrm{constant}$ and/or multilevel modulation such as QAM is employed. To lessen the effects of channel and data pattern, [21] uses multiple data blocks (≥ 50 OFDM symbols) for averaging, which is obviously impractical for burst mode applications.

The estimator $f_d^\#$ is said to be consistent if there is only one $\hat{f}_d$ that maximizes $O(\hat{f}_d)$. In the next section, we shall prove consistency of the proposed CFO estimator over the frequency-selective fading channel and introduce a practical realization for (15). Since $\rho(f)$ has a period N, i.e., $\rho(f) \equiv \rho(f + N)$, in the sequel we shall consider the scenario of $f_d \in [-N/2+1, N/2]$ only.

IV. CONSISTENCY AND REALIZATION

A. Consistency of the Proposed PBMLE

Over frequency-selective fading channels, deep fading or channel nulls may occur on some subcarriers, which may deteriorate the performance of the PB CFO estimators. For instance, consider the extreme case that channel nulls occupy all the subcarriers of the pilot tones, i.e., $\boldsymbol{H}_K = 0$, which lead to $\boldsymbol{B}_K = 0$ and, hence, $O(f_d) \equiv 0, \forall f_d$. In this situation, the PB CFO estimators fail to produce any meaningful result. Let Γ denote the number of pilot tones occupied by the channel nulls and assume at least two pilot tones do not coincide with the channel nulls, i.e., $\Gamma \leq K - 2$. We shall prove that the proposed PBMLE produces consistent estimate of f_d under this assumption regardless of the form of the channel.

First, we make the following definitions:

$$\boldsymbol{\alpha} = \boldsymbol{W}_K^T \boldsymbol{\Omega}(\hat{f}_d - f_d) \boldsymbol{W}_K^* \boldsymbol{B}_K^*$$
$$= [a_1 \quad a_2 \quad \cdots \quad a_K]^T \qquad (16)$$

and

$$\boldsymbol{\beta} = \boldsymbol{W}_N^T \boldsymbol{\Omega}(\hat{f}_d - f_d) \boldsymbol{W}_N^* \boldsymbol{B}_N^*$$
$$= [b_1 \quad b_2 \quad \cdots \quad b_N]^T \qquad (17)$$

where the superscript $*$ denotes complex conjugate and $\boldsymbol{B}_N = \boldsymbol{H}_N \boldsymbol{P}_N$. From the configuration of $\boldsymbol{P}_K$ and $\boldsymbol{P}_N$, we obtain

$$\boldsymbol{B}_K^H \boldsymbol{B}_K = \boldsymbol{B}_N^H \boldsymbol{B}_N \qquad (18)$$

and

$$\boldsymbol{B}_K^H \boldsymbol{W}_K^H = \boldsymbol{B}_N^H \boldsymbol{W}_N^H. \qquad (19)$$

As a result, the elements of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are related by

$$a_k = b_{t_k}, \quad 1 \leq k \leq K \qquad (20)$$

and, consequently

$$\boldsymbol{\alpha}^H \boldsymbol{\alpha} \leq \boldsymbol{\beta}^H \boldsymbol{\beta}. \qquad (21)$$

After omitting the noise term and the constant factor independent of f_d, (3) can be rewritten as

$$\boldsymbol{R} = \boldsymbol{\Omega}(f_d) \boldsymbol{W}_K \boldsymbol{B}_K. \qquad (22)$$

Substituting (22) into (10), we have

$$O\left(\hat{f}_d\right) = \boldsymbol{\alpha}^H \boldsymbol{\alpha}. \qquad (23)$$

Using (16)–(18) and (21), it follows that

$$O\left(\hat{f}_d\right) \leq \boldsymbol{\beta}^H \boldsymbol{\beta} = O\left(f_d\right). \qquad (24)$$

Let $d(\hat{f}_d)$ represent the difference between $O\left(\hat{f}_d\right)$ and $O\left(f_d\right)$, which can be expressed by

$$d(\hat{f}_d) = O(f_d) - O(\hat{f}_d)$$
$$= \boldsymbol{B}_K^H \boldsymbol{W}_K^H \boldsymbol{\Omega}(\hat{f}_d - f_d) \boldsymbol{W}_{N-K} \boldsymbol{W}_{N-K}^H \boldsymbol{\Omega}(f_d - \hat{f}_d) \boldsymbol{W}_K \boldsymbol{B}_K$$
$$= \boldsymbol{\gamma}^H \boldsymbol{\gamma} \qquad (25)$$

where $\boldsymbol{\gamma} = \boldsymbol{W}_{N-K}^H \boldsymbol{\Omega}\left(f_d - \hat{f}_d\right) \boldsymbol{W}_K \boldsymbol{B}_K$ and $\boldsymbol{W}_{N-K} = [\boldsymbol{w}_{n_1} \quad \boldsymbol{w}_{n_2} \quad \cdots \quad \boldsymbol{w}_{n_{N-K}}]$. Particularly, when $\hat{f}_d = f_d$, we have $O\left(\hat{f}_d\right) = \boldsymbol{B}_K^H \boldsymbol{B}_K$ and $\gamma = 0$.

Remark: To prove consistency of $f_d^\#$ is equivalent to proving that $\hat{f}_d = f_d$ is the only value to maximize $O\left(\hat{f}_d\right)$. It is evident from (24) and (25) that the necessary and sufficient condition for $O\left(\hat{f}_d\right)$ to reach its maxima is $\boldsymbol{\gamma} = 0$. In the following, we shall prove $\hat{f}_d = f_d$ is the only value that can make $\boldsymbol{\gamma} = 0$, provided $\Gamma \leq K - 2$.

For the convenience of expression, let $\Delta = \hat{f}_d - f_d$, then $\boldsymbol{\gamma}$ can be expanded into

$$\boldsymbol{\gamma} = [\boldsymbol{w}_{n_1 - \Delta} \quad \boldsymbol{w}_{n_2 - \Delta} \quad \cdots \quad \boldsymbol{w}_{n_{N-K} - \Delta}]^H \left(\sum_{k=1}^{K} p_{t_k} h_{t_k} \boldsymbol{w}_{t_k}\right) \qquad (26)$$

where $\boldsymbol{w}_{n_m - \Delta} = [1 \quad e^{j2\pi N^{-1}(n_m - \Delta)} \quad \cdots \quad e^{j2\pi N^{-1}(N-1)(n_m - \Delta)}]^T, 1 \leq m \leq N - K$. Next, we define the following sets: $\overline{A} = \{\boldsymbol{w}_{t_k}, 1 \leq k \leq K\}$, $\overline{B} = \{\boldsymbol{w}_{n_m}, 1 \leq m \leq N - K\}$, $\overline{C} = \{\boldsymbol{w}_{n_m - \Delta}, 1 \leq m \leq N - K\}$ and $\overline{D} = \{\boldsymbol{w}_{t_k} \mid h_{t_k} = 0, 1 \leq k \leq K\}$. Because vectors in $\overline{A}, \overline{B},$ and $\overline{C}$ are orthogonal, i.e., $\boldsymbol{w}_{t_l}^H \boldsymbol{w}_{t_i} = N\delta(t_l - t_i)$, $\boldsymbol{w}_{n_l}^H \boldsymbol{w}_{n_i} = N\delta(n_l - n_i)$ and $\boldsymbol{w}_{n_l - \Delta}^H \boldsymbol{w}_{n_i - \Delta} = N\delta(n_l - n_i)$, they form the basis of the subspaces that correspond to sets $\overline{A}, \overline{B},$ and $\overline{C}$, respectively. Since $\boldsymbol{\gamma} = 0$ holds if and only if the vectors within $\overline{C}$ belong to the null space of $\overline{A}$ [25], and $\overline{B} \cup \overline{D}$ contains the basis of the null space of $\overline{A}$, it follows that

$$\ddot{S}\{\overline{C}\} \subset \ddot{S}\{\overline{B} \cup \overline{D}\} \qquad (27)$$

where $\ddot{S}\{\cdot\}$ denotes the linear space spanned by the vectors inside the bracket. Define $n_m' = (n_m - \Delta) \bmod N$ and $\{t_{\tilde{k}}\} = \{t_k \mid h_{t_k} = 0\}$, where "mod" denotes the modular operator, $\{t_{\tilde{k}}\}$ include the indexes of those pilot tones occupied by the channel nulls, and the size of $\{t_{\tilde{k}}\}$ is Γ. Due to the particular form of the vectors within $\overline{B}, \overline{C},$ and $\overline{D}$, their orthogonality can equivalently be represented by their indexes; therefore, (27) equals

$$\{n_m'\}_{m=1}^{N-K} \subset \{n_m\}_{m=1}^{N-K} \cup \{t_{\tilde{k}}\}. \qquad (28)$$

In the following, we shall consider the cases $\{n_m'\}_{m=1}^{N-K} \cap \{t_{\tilde{k}}\} = \{\emptyset\}$ and $\{n_m'\}_{m=1}^{N-K} \cap \{t_{\tilde{k}}\} \neq \{\emptyset\}$ separately when (28) holds, where $\{\emptyset\}$ denotes the empty set.

Case 1: $\{n_m'\}_{m=1}^{N-K} \cap \{t_{\tilde{k}}\} = \{\emptyset\}$: If $m_1 \neq m_2$, it is obvious that $n_{m_1} \neq n_{m_2}$ and $n_{m_1}' \neq n_{m_2}'$. Since $\{t_k\}_{k=1}^{K} \cup \{n_m\}_{m=1}^{N-K} = \{0, 1, \ldots N - 1\}$ and $\{t_k\}_{k=1}^{K} \cap \{n_m\}_{m=1}^{N-K} = \{\emptyset\}$, when (28) holds and $\{n_m'\}_{m=1}^{N-K} \cap \{t_{\tilde{k}}\} = \{\emptyset\}$, we obtain $\{n_m'\}_{m=1}^{N-K} = \{n_m\}_{m=1}^{N-K}$. Moreover, due to the insertion of pilot tones at subcarriers $\{t_k\}_{k=1}^{K}$, null subcarriers' indexes $\{n_m\}_{m=1}^{N-K}$ are separated into K subsets, and the indexes within each subset are consecutive (because of the modula-N operation, index $N - 1$ is regarded as adjacent to index 0). The size of subset k is given by

$$L_k = \begin{cases} t_{k+1} - t_k - 1, & 1 \leq k \leq K - 1 \\ N - 1 + t_1 - t_k, & k = K. \end{cases} \qquad (29)$$

Considering (1), we have

$$L_{k_1} \neq L_{k_2}, \quad \forall k_1, k_2 \in [1, K] \text{ and } k_1 \neq k_2. \qquad (30)$$
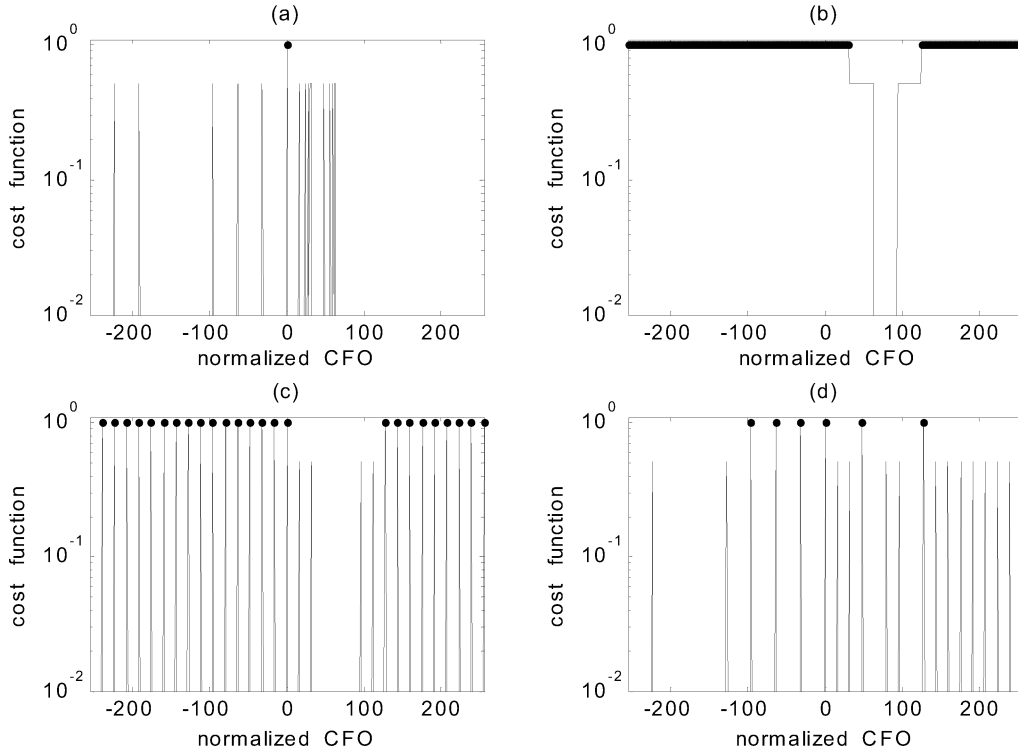
Fig. 2.   Performance of the normalized cost function employed by different PB CFO estimators over a frequency selective fading channel, under the assumption of $f_d = 0$ and all but two pilot tones are occupied by channel nulls ($\Gamma = K - 2$). Global maxima of the cost function are indicated by solid dots. (a) Cost function employed by proposed PBMLE. (b) Cost function employed by Ge's estimator [21]. (c) Cost function employed by Keller's estimator [22]. (d) Cost function employed by Nogami's estimator [20].

Letting $L_{\overline{K}} = \max_k \{L_k\}$, the indexes of subset $\overline{K}$ are given by

$$\{n_{m_l}\}_{l=1}^{L_{\overline{K}}} = \begin{cases} [t_{\overline{K}} + 1, t_{\overline{K}+1} - 1], & 1 \leq \overline{K} \leq K-1 \\ [t_{\overline{K}} + 1, N - 1] \cup [0, t_1 - 1], & \overline{K} = K. \end{cases} \quad (31)$$

On the other hand, $\{(n_{m_l} - \Delta) \bmod N\}_{l=1}^{L_{\overline{K}}}$ is a subset within $\{n'_m\}_{m=1}^{N-K}$ that has $L_{\overline{K}}$ consecutive indexes. Because (30) and $\{n'_m\}_{m=1}^{N-K} = \{n_m\}_{m=1}^{N-K}$, we have

$$\{(n_{m_l} - \Delta) \bmod N\}_{l=1}^{L_{\overline{K}}} = \{n_{m_l}\}_{l=1}^{L_{\overline{K}}}. \quad (32)$$

Equation (32) holds and only holds when $\Delta = 0$. Therefore, we have proved that when (28) holds and $\{n'_m\}_{m=1}^{N-K} \cap \{t_{\widetilde{k}}\} = \{\emptyset\}$, $\gamma = 0$ if and only if $\hat{f}_d = f_d$.

Case 2: $\{n'_m\}_{m=1}^{N-K} \cap \{t_{\widetilde{k}}\} \neq \{\emptyset\}$: Let $t'_k = (t_k - \Delta) \bmod N$, then $\{t'_k\}_{k=1}^K \cup \{n'_m\}_{m=1}^{N-K} = \{t_k\}_{k=1}^K \cup \{n_m\}_{m=1}^{N-K} = \{0, 1, \ldots N - 1\}$. Since the size of $\{t_{\widetilde{k}}\}$ equals Γ, when (28) holds and $\Gamma \leq K - 2$, there are at least two common elements belonging to both $\{t'_k\}_{k=1}^K$ and $\{t_k\}_{k=1}^K$. Assuming $\{t'_{k_1}, t'_{k_2}\} \subset \{t'_k\}_{k=1}^K \cup \{t_k\}_{k=1}^K$, from the definition of t'_k, we find there exists $\{t_{k_1}, t_{k_2}\} \subset \{t_k\}_{k=1}^K$ which satisfies

$$\begin{cases} t'_{k_1} = (t_{k_1} - \Delta) \bmod N \\ t'_{k_2} = (t_{k_2} - \Delta) \bmod N \end{cases} \quad (33)$$

and it follows that $t'_{k_1} - t'_{k_2} = t_{k_1} - t_{k_2}$. Using (1) yields

$$t'_{k_1} = t_{k_1} \text{ and } t'_{k_2} = t_{k_2}. \quad (34)$$

Substituting (34) back into (33), we obtain that (33) holds if and only if $\Delta = 0$. Hence, we have also proved that when (28) holds and $\{n'_m\}_{m=1}^{N-K} \cap \{t_{\widetilde{k}}\} \neq \{\emptyset\}$, $\gamma = 0$ if and only if $\hat{f}_d = f_d$.

Combining cases 1 and 2, and following the *Remark* after (25), we have proved that $\gamma = 0$ if and only if $\hat{f}_d = f_d$. Therefore, consistency of $f_d^{\#}$ is guaranteed if at least two pilot tones are not occupied by channel nulls. It is worth noting that consistency of the proposed PBMLE is independent of the form of the channel as the proof above does not involve any assumption about $\boldsymbol{H}_N$ except for the condition of $\Gamma \leq K - 2$. We remark that the previous CFO estimator [19]–[22] does not have this property. For example, in [19], nulls are inserted in regular OFDM symbols and its consistency is still influenced by the channel form. To elaborate, consider the case that the pattern of channel nulls coincides with that of inserted nulls, that is $f_{\widecheck{i}_k} = f_{i_k} + f_{\text{shift}}$, where $f_{\widecheck{i}_k}$ and f_{i_k} denote the frequencies of channel nulls and inserted null subcarriers, respectively, and $f_{\text{shift}} \neq 0$ is a constant. Obviously, this CFO estimator is not consistent since it cannot distinguish the nulls generated by $\{f_{\widecheck{i}_k}\}$ and $\{f_{i_k}\}$. As a final remark, we also note the PB estimators in [20]–[22] (for the purpose of comparison, the subcarriers of the data blocks of [21] are regarded as pilot tones) do not take into account the effect of frequency selectivity and may lead to multiple extrema of the cost function and nonuniqueness of the CFO estimation. To illustrate, the normalized cost function of CFO (original cost function divided by its maxima) employed in the proposed PBMLE and [20]–[22] are shown in Fig. 2(a)–(d) over a frequency selective fading channel with channel nulls. We assume $N = 512$, $f_d = 0$, and channel nulls occupy all the pilot tones except those on subcarriers 32 and 64 (i.e., $\Gamma = K - 2$). The maxima of the cost function are indicated by solid dots. Fig. 2(a) shows the cost function used by the PBMLE, in which

the subcarriers of the pilot tones are given by $t_k = 2^{k-1}$ and the condition in (1) is satisfied. Fig. 2(b) plots the cost function used by Ge's estimator [21], in which the subcarriers of the data block are given by $[-200, 200]$, and the modulation format is assumed to be QPSK. Fig. 2(c) gives the cost function employed in Keller's estimator [22], in which the subcarriers of the uniformly spaced pilot tones are given by $t_k = -200 + 4k$. Fig. 2(d) depicts the cost function used by Nogami's estimator [20], in which the PN sequence used to assign the pilot tones is $\{0\,101\,001\,100\,000\}$ and the subcarriers of the nonuniformly spaced pilot tones are given by

$$\{-176, -128, -112, -64, -48, -16,$$
$$32, 64, 80, 112, 144, 176, 208\}.$$

It is clear from these figures that except for the proposed PBMLE, the other PB CFO estimators all suffer from ambiguity resulting from the multiple extrema of the cost function and, consequently, their consistency is lost under such a channel condition.

B. Realization

Suboptimal ML Estimator: Although (15) gives the ML estimate of f_d, its evaluation is not straightforward. Nevertheless, it is worth noting from (14) that $\rho(f)$ can be obtained through FFT when $f \in \overline{\mathbf{Z}}$, where $\overline{\mathbf{Z}}$ denotes the set of integers. When $f \notin \overline{\mathbf{Z}}$, interpolation can be employed to approximate $\rho(f)$. Based on this observation, a suboptimal MLE is presented in the following. To begin with, we rewrite the CFO as $f_d = \lambda/L + q$, where $\lambda \in \overline{\mathbf{Z}}$, $L = 2^l$, $l > 0$, and $l \in \overline{\mathbf{Z}}$, and $|q| \le (2L)^{-1}$. Then, we build a zero-padded sequence $\{r'_m\}$

$$r'_m = \begin{cases} r_m, & 0 \le m \le N-1 \\ 0, & N \le m \le LN-1. \end{cases} \quad (35)$$

Next, we invoke the LN-point FFT over $\{r'_m\}$ yielding

$$\Re(v) = \sum_{m=0}^{LN-1} r'_m \exp[-j2\pi m v(LN)^{-1}], \quad \frac{-LN}{2} < v \le \frac{LN}{2}. \quad (36)$$

The suboptimal MLE of f_d then takes the form of

$$\tilde{f}_d^{\#} = L^{-1} \arg\max_{\hat{v}} \left\{ \sum_{k=1}^{K} |\Re(\hat{v}+Lt_k)|^2 \right\}. \quad (37)$$

Using (3) and ignoring the white noise term ξ, $\Re(v)$ can be rewritten as

$$\Re(v) = \sum_{m=0}^{N-1} \sum_{k=1}^{K} h_{t_k} p_{t_k} e^{j2\pi m(LN)^{-1}[L(f_d+t_k)-v]}$$
$$= \sum_{k=1}^{K} h_{t_k} p_{t_k} e^{j\omega_{t_k,v}(N-1)} \frac{\sin(N\omega_{t_k,v})}{\sin(\omega_{t_k,v})} \quad (38)$$

where $\omega_{t_k,v} = \pi(LN)^{-1}[L(f_d+t_k)-v]$. When $q = 0$, it turns out that

$$\arg\max_{\hat{f}_d} \left\{ \sum_{k=1}^{K} \rho(\hat{f}_{d,k}) \right\} = L^{-1} \arg\max_{\hat{v}} \left\{ \sum_{k=1}^{K} |\Re(\hat{v}+Lt_k)|^2 \right\} \quad (39)$$

and it follows $\tilde{f}_d^{\#} = f_d^{\#}$. In this case, the suboptimal estimator (37) agrees with the optimal estimator (15). If $q \ne 0$, (39) no longer holds but (37) still produces the best approximation to (15) from the set of $\{-N/2+1, -N/2+1+1/L, \ldots N/2\}$, i.e., $\tilde{f}_d^{\#} = \lambda/L$. As a result, we have $\left| \tilde{f}_d^{\#} - f_d \right| \le (2L)^{-1}$. Evidently, the estimation accuracy is proportional to L and the suboptimal estimator reduces to the optimal estimator when L tends to infinity. On the other hand, increasing L also implies increasing the size of FFT, which is undesirable for practical implementation. Consequently, there is a tradeoff between the desired precision and the computation simplicity.

Refined Suboptimal Estimator: When L is limited, (37) can only produce a coarse estimate of f_d. Due to the sensitivity of the OFDM system to CFO, another procedure is needed to refine the performance of (37). As observed from (38), the FFT bins around $\Re(\tilde{f}_d^{\#} + Lt_k)$ contain side information about q, which can be exploited to refine the coarse estimate. In the following, we will introduce a fine CFO estimator based on the FFT bins around $\Re(\tilde{f}_d^{\#} + Lt_k)$.

Suppose the spacing between adjacent pilot tones is large enough ($|t_k - t_{k+1}| \ge 3$) so that the ACI resulting from adjacent tones can be ignored, thereby for $\lambda + L(t_k - 1) < k < \lambda + L(t_k + 1)$, the magnitude of (38) can be approximated by

$$|\Re(v)| \approx \left| \exp[j\omega_{t_k,v}(N-1)] \frac{\sin(N\omega_{t_k,v})}{\sin(\omega_{t_k,v})} h_{t_k} \right|$$
$$= \left| \frac{\sin(N\omega_{t_k,v})}{\sin(\omega_{t_k,v})} \right| \cdot |h_{t_k}|. \quad (40)$$

Particularly, when $v = \lambda + L(t_k - 1/2)$ and $\lambda + L(t_k + 1/2)$, we have

$$\left| \Re\left[\lambda + L\left(t_k - \frac{1}{2} \right) \right] \right| \approx \frac{N\cos(\pi q)}{\pi\left(q + \frac{1}{2}\right)} |h_{t_k}| \quad (41)$$

and

$$\left| \Re\left[\lambda + L\left(t_k + \frac{1}{2} \right) \right] \right| \approx \frac{N\cos(\pi q)}{\pi\left(\frac{1}{2} - q\right)} |h_{t_k}|. \quad (42)$$

After some simple arithmetic manipulations, q can be approximated by

$$\hat{q} = \frac{1 - \left| \Re\left[\lambda + L\left(t_k - \frac{1}{2} \right) \right] \right| \cdot \left| \Re\left[\lambda + L\left(t_k + \frac{1}{2} \right) \right] \right|^{-1}}{2\left(1 + \left| \Re\left[\lambda + L\left(t_k - \frac{1}{2} \right) \right] \right| \cdot \left| \Re\left[\lambda + L\left(t_k + \frac{1}{2} \right) \right] \right|^{-1} \right)} \quad (43)$$

where $\hat{q}$ represents the estimation of q. Taking into account all the pilot tones, a smoother estimator of q takes the form of

$$\tilde{q} = \frac{1}{K} \sum_{k=1}^{K} \frac{1 - \left| \Re\left[\lambda + L\left(t_k - \frac{1}{2} \right) \right] \right| \cdot \left| \Re\left[\lambda + L\left(t_k + \frac{1}{2} \right) \right] \right|^{-1}}{2\left(1 + \left| \Re\left[\lambda + L\left(t_k - \frac{1}{2} \right) \right] \right| \cdot \left| \Re\left[\lambda + L\left(t_k + \frac{1}{2} \right) \right] \right|^{-1} \right)}. \quad (44)$$

Considering (37) again after ignoring the ACI from neighboring pilot tones, we have

$$\arg\max_{\hat{v}} \left\{ \sum_{k=1}^{K} |\Re(\hat{v}+Lt_k)|^2 \right\}$$
$$\approx \arg\max_{\hat{v}} \left\{ \left| \frac{\sin(N\hat{\omega}_{\hat{v}})}{\sin(\hat{\omega}_{\hat{v}})} \right|^2 \sum_{k=1}^{K} |h_{t_k}|^2 \right\} \quad (45)$$

TABLE I
COMPUTATIONAL LOAD OF THE PROPOSED METHOD, NNE, AND MTGBE

CFO ESTIMATOR	NUMBER OF REAL PRODUCTS	NUMBER OF REAL ADDITIONS
PROPOSED METHOD	$\approx 2LN\log_2^{LN} + 6K + KM$	$\approx 2LN\log_2^{LN} + 2K + KM$
NNE	$\approx N^2 + 2N + 2N\log_2^{N} + 5K$	$\approx N^2 + 2N\log_2^{N} - N + 6K$
MTGBE	$\approx 22N_\alpha N^2 + 10N_\alpha N$	$\approx 2N_\alpha N^3 - 2kN_\alpha N^2 + 2N_\alpha N^2$

where $\widehat{\varpi}_{\hat{v}} = \pi(LN)^{-1}(Lf_d - \hat{v})$. It can be found from the right-hand side of (45) that the factor $\sum_{k=1}^{K}|h_{t_k}|^2$ is a constant independent of $\hat{v}$; therefore

$$\arg\max_{\hat{v}}\left\{\left|\frac{\sin(N\varpi_{\hat{v}})}{\sin(\varpi_{\hat{v}})}\right|^2\sum_{k=1}^{K}|h_{t_k}|^2\right\}$$

$$= \arg\max_{\hat{v}}\left\{\left|\frac{\sin(N\widehat{\varpi}_{\hat{v}})}{\sin(\widehat{\varpi}_{\hat{v}})}\right|\sum_{k=1}^{K}|h_{t_k}|\right\}. \quad (46)$$

Thus, the square operator involved in calculating the periodogram is dropped, and (37) can be simplified into

$$\tilde{f}_d^{\#} = L^{-1}\arg\max_{\hat{v}}\left\{\sum_{k=1}^{K}|\Re(\hat{v} + Lt_k)|\right\}. \quad (47)$$

Combining (44) and (47) finally leads to the refined suboptimal estimate of f_d

$$\tilde{f}_d = \tilde{f}_d^{\#} + \tilde{q}. \quad (48)$$

Computational Load: It can be seem from (44) and (47) that the computation complexity of (48) mainly depends on obtaining $\tilde{f}_d^{\#}$. Despite the simplification given by (45)–(47), the brute-force implementation of (47) still requires a computationally burdensome exhaustive search for $\tilde{f}_d^{\#}$ over the possible range of CFO. To lessen the computation load, we introduce a strategy to simplify the implementation of (47) in the following.

Step 1 : Setting the threshold as $\widehat{T}_r = \widehat{C}\sum_{v=1-LN/2}^{LN/2}|\Re(v)|$, where $0 < \widehat{C} < K^{-1}$ is a design parameter.

Step 2 : Defining $\widehat{R}(v) = \begin{cases} |\Re(v)|, & |\Re(v)| \geq \widehat{T}_r \\ 0, & |\Re(v)| < \widehat{T}_r \end{cases}$, letting $\widehat{S} = \left\{v_iL^{-1}\big|\widehat{R}(v_i) > 0\right\}_{i=1}^{\widehat{M}}$ and collecting the elements of $\widehat{S}$ in vector $\boldsymbol{P} = L^{-1}\begin{bmatrix} v_1 & v_2 & \cdots & v_{\widehat{M}} \end{bmatrix}$ according to the increasing order of v_i, where $\widehat{M}$ denotes the size of set $\widehat{S}$.

Step 3 : Correlating $\widehat{\boldsymbol{P}}$ with the shifted pattern of $\boldsymbol{P}_M$, finding the shift amount that corresponds to the peak correlation value, and denoting it as $\widehat{f}_d$.

Step 4 : Replacing $\tilde{f}_d^{\#}$ by $\widehat{f}_d$ in **(48)**.

In Table I, we have evaluated the computational load of (48) based on the simplification strategy above. As a benchmark, the computational load of the PB CFO estimator proposed by [20] (NNE) and the null-subcarrier-based nonhopping CFO estimator employed by [19] (the approach given by Section IV-A)

are also shown. We assume the proposed method and the NNE employ a preamble comprising K pilot tones, and the MTGBE inserts K nonconsecutive nulls in each OFDM symbol. Since both the proposed method and the NNE have feedforward structures, they can yield a close-form expression for f_d based on one OFDM symbol only. On the other hand, the realization of the MTGBE depends on a feedback structure and its computation complexity is proportional to the convergence time N_α, which denotes the number of adaptive iterations required to make f_d sufficiently small. As a consequence, the computational complexity of (48) and the NNE is dominated by the FFT and the correlation manipulation used for peak picking, while the complexity of the MTGBE depends mainly on the updating of covariance matrix and the convergence time N_α, which is a function of the initial CFO, the equivalent loop bandwidth, the SNR level, and the correlation factor. When $L \leq 16$, $N \geq 2^7$, and $N_\alpha > 5$, the computational load of the proposed method will be lower than that of the MTGBE. In addition, the proposed method also has less computational complexity than the NNE because generally we can arrange the distinctively spaced pilot tones to make $K \ll N$ and $\widehat{M} \ll N$ (for example, when $N = 2^9$ letting $t_k = 2^{k-1}$ leads to $K = 10$).

V. PERFORMANCE ANALYSIS

A. Estimation Range

It is evident from (14) and (15) that the proposed cost function is periodic, and N is one of its periods. From the discussion on the consistency of the PBMLE, it is found that for $f_d \in [-N/2 + 1, N/2]$, if at least two pilot tones are not occupied by channel nulls, there exists a unique $f_d^{\#}$ to maximize the cost function of CFO. Therefore, the estimation range of the PBMLE spans the entire bandwidth of the OFDM signal, which is equivalent to $[(-N/2 + 1)\Delta f, N\Delta f/2]$ Hz.

B. Statistical Properties

For simplicity of analysis, here we only consider the statistical properties of (15). Following [29], when $\mathrm{SNR} \gg 1$, the expectation and variance of $f_d^{\#}$ can be given by

$$E\{f_d^{\#}\} \approx f_d - \left.\frac{E\{O'(\hat{f}_d)\}}{E\{O''(\hat{f}_d)\}}\right|_{\hat{f}_d = f_d} \quad (49)$$

and

$$\mathrm{var}\{f_d^{\#}\} \approx \left.\frac{E\{[O'(\hat{f}_d)]^2\}}{[E\{O''(\hat{f}_d)\}]^2}\right|_{\hat{f}_d = f_d} \quad (50)$$

respectively, where we have (51) and (52), shown at the bottom of the page. $\overline{T} = \text{diag}\{0, 1, \ldots N - 1\}$ and $Q_K = \sum_{k=1}^{K} w_{it_k} w_{t_k}^H$. Inserting (3) into (51) and (52), and noting $W_K^H W_K = NI_K$, $E\{\xi\} = 0_{N\times 1}$ and $E\{\Omega(-f_d)\xi\xi^H\Omega(f_d)\} = E\{\xi\xi^H\} = \sigma^2 I_N$, we obtain

$$E[O'(f_d)] = 0 \qquad (53)$$

and

$$E\{O''(f_d)\} = 8\pi^2 N^{-2} B_K^H W_K^H \overline{T}(Q_K - NI_N)\overline{T}W_K B_K. \qquad (54)$$

Assuming SNR $\gg 1$, we have

$$E\{[O'(f_d)]^2\} \approx 8\pi^2 N^{-1}\sigma^2 B_K^H W_K^H \overline{T}(NI_N - Q_K)\overline{T}W_K B_K. \qquad (55)$$

Employing (53)–(55), (49) and (50) can be reduced to

$$E\{f_d^\#\} \approx f_d \qquad (56)$$

and

$$\text{var}\{f_d^\#\} \approx \frac{N^3\sigma^2}{8\pi^2 B_K^H W_K^H \overline{T}(NI_N - Q_K)\overline{T}W_K B_K}. \qquad (57)$$

In [27], the general formula of computing the Fisher information matrix is given, which can be used to derive the CRB for the estimation of a vector parameter. Since the CRB places a lower bound on the variance of each element to be estimated, it can be used as a benchmark to evaluate the performance of an unbiased estimator. For this reason, we collect the parameters of (6) in vector $Y = [\text{Re}(H_K), \text{Im}(H_K), f_d]^T$, where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real and imaginary parts of the enclosed quantity, respectively. Accordingly, the Fisher information matrix F is given by

$$[F]_{i_1,i_2} = E\left\{\frac{\partial^2 \ln f[R|\Xi(f_d), B_K]}{\partial Y_{i_1} Y_{i_2}}\right\}, 1 \le i_1, i_2 \le 2K + 1 \qquad (58)$$

where $[\cdot]_{i_1,i_2}$ represents the entry on the i_1th row and the i_2th column of the matrix. Inserting (6) into (58) yields

$$F = 2\sigma^{-2}\begin{bmatrix} NI_K & 0 & 0 \\ 0 & NI_K & 0 \\ \text{Im}(\eta^H W_K) & \text{Re}(\eta^H W_K) & \eta^H \eta \end{bmatrix} \qquad (59)$$

where

$$\eta = 2\pi N^{-1}\overline{T}W_K B_K. \qquad (60)$$

The CRB of parameter f_d is given by

$$\text{CRB}(f_d) = [F^{-1}]_{2K+1,2K+1}. \qquad (61)$$

From [26] and [29], the elements in the last column of F^{-1} can be expressed by

$$[F^{-1}]_{i,2K+1} = \frac{\sigma^2}{2\eta^H(I_N - N^{-1}Q_K)\eta} \cdot [\text{Im}(W_K^H\eta)^T \quad -\text{Re}(W_K^H\eta)^T \quad 1]_i^T. \qquad (62)$$

Therefore, (61) becomes

$$\text{CRB}(f_d) = \frac{\sigma^2}{2\eta^H(I_N - N^{-1}Q_K)\eta} = \frac{N^3\sigma^2}{8\pi^2 B_K^H W_K^H \overline{T}(NI_N - Q_K)\overline{T}W_K B_K}. \qquad (63)$$

Obviously, (63) coincides with (57), indicating the variance of the proposed PBMLE approaches the CRB and the proposed estimator is therefore asymptotically efficient.

VI. SIMULATION RESULTS

The following assumptions are made in the computer simulations conducted in this paper: 1) Carrier frequency = 1 GHz, channel bandwidth = 5 MHz, sampling interval $T_s = 0.2 \mu s$,

---

$$O'(\hat{f}_d) = \left.\frac{\partial[O(\hat{f}_d)]}{\partial(\hat{f}_d)}\right] = j2\pi N^{-1}$$

$$\cdot \sum_{k=1}^{K}[R^H\Omega(\hat{f}_d)\text{diag}\{0, 1, 2\cdots N - 1\}w_{t_k}w_{t_k}^H\Omega(-\hat{f}_d)R - R^H\Omega(\hat{f}_d) \times w_{t_k}w_{t_k}^H\text{diag}\{0, 1, 2\cdots N - 1\}\Omega(-\hat{f}_d)R]$$

$$= j2\pi N^{-1}R^H\Omega(\hat{f}_d)(\overline{T}Q_K - Q_K\overline{T})\Omega(-\hat{f}_d)R \qquad (51)$$

$$O''(\hat{f}_d) = \frac{\partial^2[O(\hat{f}_d)]}{\partial^2(\hat{f}_d)}$$

$$= (j2\pi N^{-1})^2\sum_{k=1}^{K}[R^H\Omega(\hat{f}_d)\text{diag}\{0^2, 1^2, 2^2\cdots(N-1)^2\}w_{t_k}w_{t_k}^H\Omega(-\hat{f}_d)R + R^H\Omega(\hat{f}_d)\times$$

$$w_{t_k}w_{t_k}^H\text{diag}\{0^2, 1^2, 2^2\cdots(N-1)^2\}\Omega(-\hat{f}_d)R - 2R^H\text{diag}\{0, 1, 2\cdots(N-1)\}\times$$

$$\Omega(\hat{f}_d)w_{t_k}w_{t_k}^H\text{diag}\{0, 1, 2\cdots(N-1)\}\Omega(-\hat{f}_d)R]$$

$$= -4\pi^2 N^{-2}R^H\Omega(\hat{f}_d)(\overline{T}^2 Q_K - 2\overline{T}Q_K\overline{T} + Q_K\overline{T}^2)\Omega(-\hat{f}_d)R \qquad (52)$$
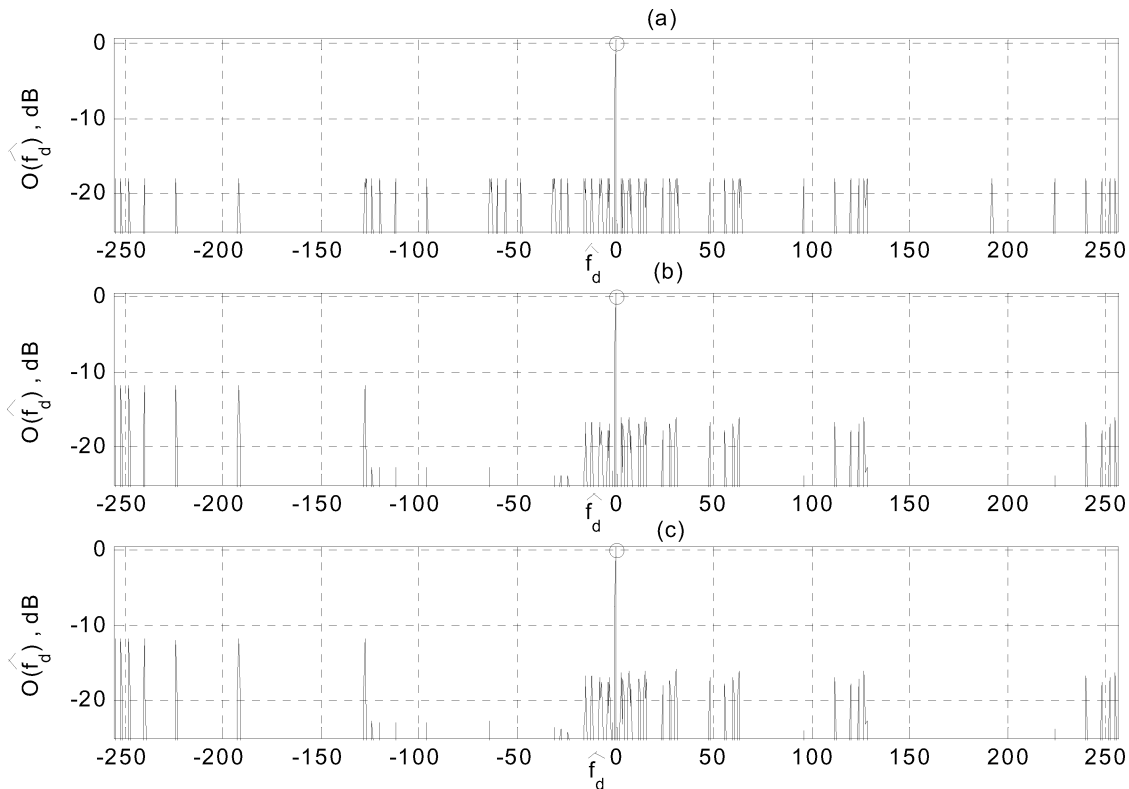
Fig. 3.   Normalized likelihood function $O(\hat{f}_d)$ versus $\hat{f}_d$ over different channels for $f_d = 0$. Global maxima of the likelihood function are indicated by circles. (a) Ideal channel, no AWGN. (b) Frequency selective fading channel, no AWGN. (c) Frequency selective fading channel, $\mathrm{SNR} = 5$ dB.

TABLE II
COARSE ESTIMATION RESULTS OF (47) FOR DIFFERENT L AND f_d AT $\mathrm{SNR} = 10$ dB

f_d L	0.63	−9	21.44	−58.185	110.773	−127.099	−255	256
2	0.5	−9	21.5	−58	111	−127	−255	256
4	0.75	−9	21.5	−58.25	110.75	−127	−255	256
8	0.625	−9	21.5	−58.125	110.75	−127.125	−255	256
16	0.625	−9	21.4375	−58.1875	110.75	−127.125	−255	256

$N = 512, M = 500, K = 7, t_k = 2^{k+2} - 7$ and $L_{CP} = 50$; 2) Pulse shaping is performed through a raised cosine filter with the roll-off factor of 0.5; 3) The impulse response of the frequency-selective fading channel is modeled as the sum of 20 paths with exponential power delay profiles [5], [26], i.e.,

$$h(t) = \sum_{i=1}^{20} A_i \delta(t - \tau_i - \tau_0) \tag{64}$$

where A_i and τ_i represent the complex gain and the delay of path i, respectively, and τ_0 is a constant denoting the initial delay which satisfies $|\tau_0 + \tau_{20}| < L_{CP}T_s$; in particular, $\tau_i = iT_s$ and $\{A_i\}_{i=1}^{20}$ are Rayleigh random variables with $E\left\{|A_i|^2\right\} = \exp\left[-\tau_i \left(5T_s\right)^{-1}\right]$ and $E\left\{A_iA_j\right\} = \delta(i - j)$; 4) The SNR is defined by $|B_K|^2 \sigma^{-2}$; and 5) Timing offset is negligible.

Fig. 3(a)–(c) illustrate the normalized likelihood function for three different cases when $f_d = 0$. Case (a) shows $O(\hat{f}_d)$ versus $\hat{f}_d$ over the ideal channel ($\boldsymbol{H}_N = \boldsymbol{I}_N$) in the absence of AWGN. Case (b) shows $O(\hat{f}_d)$ versus $\hat{f}_d$ over the selective fading channel given by (64) in the absence of AWGN. Case (c) demonstrates the effects of both AWGN and the fading channel of (64), and the SNR level is set to 5 dB. It can be seen from these figures that even at the low SNR level of case (c), there is only one global maxima of $O(\hat{f}_d)$ corresponding to $\hat{f}_d = 0$ over the entire range of $[-N/2 + 1, N/2]$. Therefore the ML estimation of CFO leads to $f_d^{\#} = 0$ and $f_d^{\#} = f_d$, the desired result. Similarly, when $f_d \neq 0$, the global maxima of the likelihood function will occur at the frequency location of f_d.

Assuming $\mathrm{SNR} = 10$ dB, Table II gives the coarse estimation results of (47) corresponding to $f_d = \{0.63, -9, 21.44, -58.185, 110.773, -127.099, -255, 256\}$ and $L = \{2, 4, 8, 16\}$. It can be observed from this table that the estimation range covers $[-N/2 + 1, N/2]$, which is
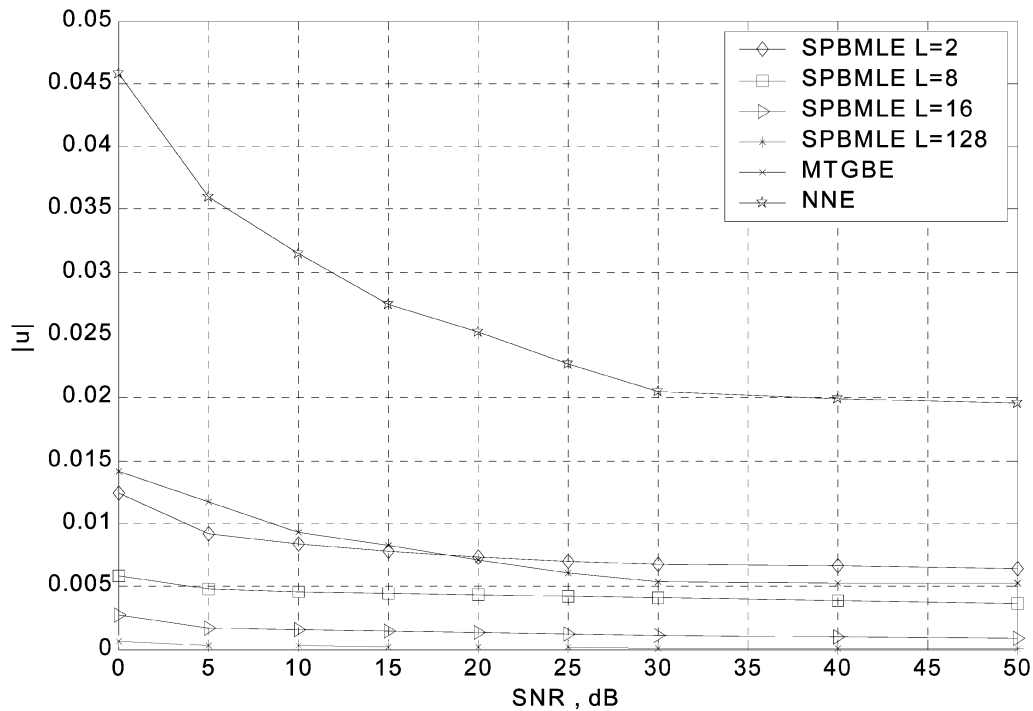
Fig. 4. Magnitude of the average bias for the proposed suboptimal CFO estimator (SPBMLE) as a function of SNR, with L as a parameter. As a benchmark, the performance of the CFO estimators MTGBE [19] and NNE [20] are also shown.
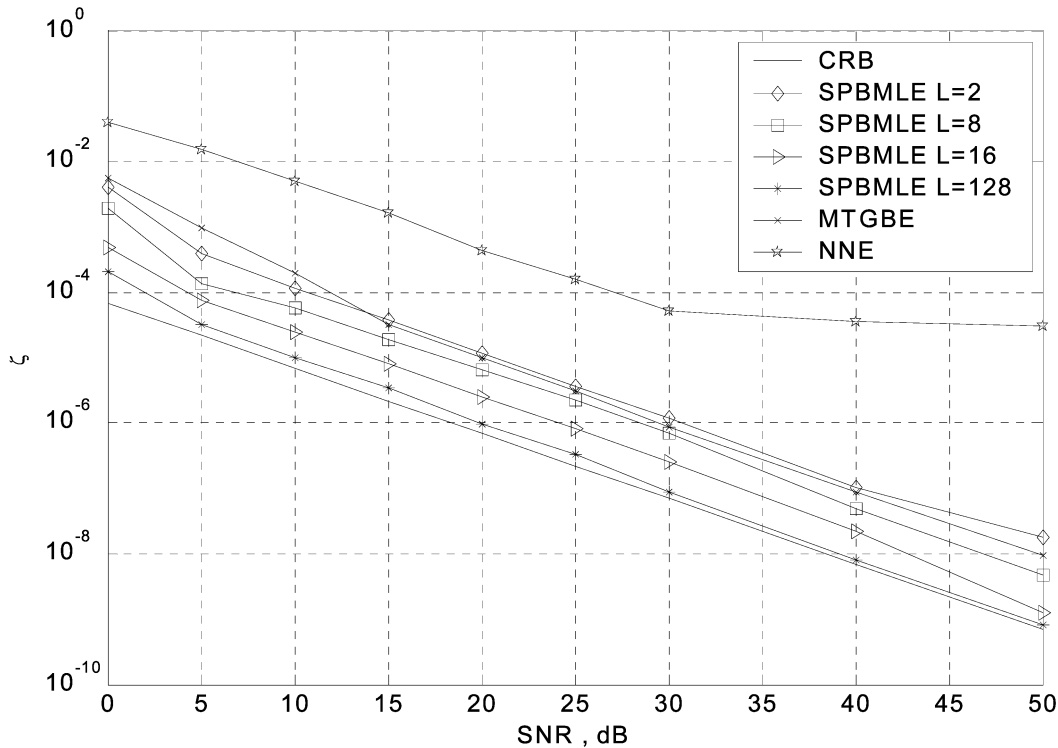

Fig. 5. Mean-square error of the proposed suboptimal CFO estimator (SPBMLE) as a function of the SNR, with L as a parameter. As a benchmark, the CRB given by (63) and the performance of the CFO estimators MTGBE [19] and NNE [20] are also shown.

equal to the bandwidth of the OFDM signal. In addition, when f_d is an integer (cases $f_d = -9, -255, 256$), $\widetilde{f}_d^{\#} = f_d$; also, when f_d has a fractional part (cases $f_d = 0.63$, $21.44, -58.185, 110.773, -127.099$), the coarse estimate produced by (47) satisfies the relationship of $\left|\widetilde{f}_d^{\#} - f_d\right| \leq$

$(2L)^{-1}$. If L can be made infinitely large, the coarse estimator alone is expected to yield the unbiased estimate of f_d.

Once the coarse estimation $\widetilde{f}_d^{\#}$ has been obtained, the remaining part of $f_d - \widetilde{f}_d^{\#}$ can be corrected by (44). To measure the accuracy of the refined suboptimal estimator (SPBMLE) given

by (48), we introduce two new variables: the average bias u and the mean-square error ζ, which are defined by

$$u = f_d - (N_S)^{-1} \sum_{i=1}^{N_S} (\tilde{f}_d)_i \tag{65}$$

and

$$\zeta = (N_S)^{-1} \sum_{i=1}^{N_S} [(\tilde{f}_d)_i - f_d]^2 \tag{66}$$

respectively, where N_S denotes the total number of simulations and $(\tilde{f}_d)_i$ represents the outcome of the ith run. With L as a parameter, Figs. 4 and 5 illustrate the magnitude of u and ζ versus SNR, respectively. As a benchmark, the performance of the estimators proposed in [19] (MTGBE) and [20] (NNE) are also shown in Figs. 4 and 5, and the CRB given by (63) is plotted against the curves of mean square error in Fig. 5. The normalized CFO is set to $f_d = -58.185$ and $N_S = 10^3$. The system parameters employed by the MTGBE and the NNE are the same as those in obtaining Fig. 2, and the step size and correlation factor adopted by the MTGBE equal those used by [19, Example 3]. It is clear from these two figures that both the SPBMLE and the MTGBE outperform the NNE, and the SPBMLE also outperforms the MTGBE for $L \geq 8$. Furthermore, it is obvious that the accuracy of the proposed SPBMLE is proportional to L, the oversize ratio. When L is sufficiently large, it is found that the SPBMLE is asymptotically unbiased and its variance approaches the CRB closely. For most practical applications [1]–[3], $L = 8$ is sufficient for SNR levels of interest.

VII. CONCLUSION

In this paper, a PBMLE has been proposed to estimate the CFO of OFDM systems over a frequency-selective fading channel. The derived likelihood function has revealed the relationship between the CFO and the periodogram of the preamble. Due to the distinctive spacing of the pilot tones, the proposed PBMLE is robust against frequency-selective fading and is independent of the form of the channel, as long as at least two pilot tones are not occupied by channel nulls. Analysis has shown that the PBMLE is asymptotically unbiased and efficient. To realize this PBMLE in practice, a suboptimal estimator has been developed, in which zero-padded FFT is invoked for implementation and the estimation process is split into two phases: the coarse estimate, obtained by the correlation between the received spectrum and the original pattern of the pilot tones, and the fine estimate, obtained by investigating the magnitude attenuation of the frequency bins around those CFO-shifted pilot tones. Both analytical analysis and computer simulations have indicated that the accuracy of this simplified estimator is proportional to the oversize ratio of zero-padded FFT and its estimation range equals the bandwidth of the OFDM signal. When the oversize ratio is sufficiently high, the performance of the suboptimal estimator approaches that of the PBMLE.

ACKNOWLEDGMENT

The authors wish to thank the editor and the anonymous reviewers for their helpful comments and suggestions in improving the quality of this paper.

REFERENCES

[1] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, pp. 665–675, July 1985.

[2] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

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

[4] F. Classen and H. Meyr, "Frequency synchr onization algorithms for OFDM systems suitable for communication over frequency selective fading channels," in *Proc. VTC'94*, 1994, pp. 1655–1659.

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

[6] M. Morelli, A. N. Andrea, and U. Mengali, "Frequency ambiguity resolution in OFDM systems," *IEEE Commun. Lett.*, vol. 4, pp. 134–136, Apr. 2000.

[7] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Trans. Commun.*, vol. 48, pp. 988–999, June 2001.

[8] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat fading channels: A cyclostationary approach," *IEEE Trans. Commun.*, vol. 46, pp. 400–411, Mar. 1998.

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

[10] M. Luise and R. Reggiannini, "Carrier frequency acquisition and tracking for OFDM systems," *IEEE Trans. Commun.*, vol. 44, pp. 1590–1598, Nov. 1996.

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

[12] G. Santella, "Frequency and symbol synchronization system of OFDM signals: Architecture and simulation results," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 254–275, Jan. 2000.

[13] H. K. Song, Y. H. You, J. H. Paik, and Y. S. Cho, "Frequency-offset synchronization and channel estimation for OFDM-based transmissions," *IEEE Commun. Lett.*, vol. 4, pp. 95–97, Mar. 2000.

[14] J. Van de Beek, M. Sandel, and P. O. Borjesson, "ML estimation of timing and frequency offset in OFDM systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 1800–1805, July 1997.

[15] N. Lashkarian and S. Kiaei, "Class of cyclic-based estimators for frequency-offset estimation of OFDM systems," *IEEE Trans. Commun.*, vol. 48, pp. 1590–1598, Dec. 2000.

[16] H. Liu and U. Tureli, "A high efficiency estimator for OFDM communications," *IEEE Commun. Lett.*, vol. 2, pp. 104–106, Apr. 1998.

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

[18] U. Tureli, D. Kivanc, and H. Liu, "Experimental and analytical studies on a high-resolution OFDM carrier frequency offset estimator," *IEEE Trans. Veh. Technol.*, vol. 50, Mar. 2001.

[19] X. Ma, C. Tepedelenlioglu, 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, pp. 2504–2515, Dec. 2001.

[20] H. Nogami and T. Nagashima, "A frequency and timing period acquisition technique for OFDM systems," in *Proc. PIMRC'95*, Sept. 1995, pp. 1010–1015.

[21] H. Ge and K. Wang, "Efficient method for carrier offset correction in OFDM system," in *Proc. ICASSP'99*, 1999, pp. 2467–2470.

[22] T. Keller, L. Piazzo, P. Mandarini, and L. Hanzo, "Orthogonal frequency division multiplex synchronization techniques for frequency-selective fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 999–1008, June 2001.

[23] A. J. Coulson, "Maximum likelihood synchronization for OFDM using a pilot symbol: Algorithms," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 2486–2494, Dec. 2001.

[24] ——, "Maximum likelihood synchronization for OFDM using a pilot symbol: Analysis," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 2495–2503, Dec. 2001.

[25] J. Gilbert and L. Gilbert, *Linear Algebra and Matrix Theory*. San Diego, CA: Academic, 1995.

[26] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, pp. 1580–1589, Sept. 2000.

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

[28] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least square problems whose variables separate," *J. Soc. Indust. Appl. Math., Series B: Numerical Anal.*, vol. 10, no. 2, pp. 413–432, Apr. 1973.

[29] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers*. New York: Wiley, 1997.

Jing Lei received the B.Eng degree in biomedical engineering and the M.Eng degree in electrical engineering from Southeast University, Nanjing, China.

She is now a Research Fellow with the University of Hong Kong, Hong Kong. Her research interests include the areas of wireless communications and signal processing.

Tung-Sang Ng (S'74–M'78–SM'90–F'03) received the B.Sc. (Eng.) degree from the University of Hong Kong, Hong Kong, and the M.Eng. and Ph.D. degrees from the University of Newcastle, Australia, respectively, all in electrical engineering.

He worked for BHP Steel International and the University of Wollongong, Australia, after graduation for 14 years. He then returned to the University Hong Kong in 1991 as a Professor and Chair of the Electronic Engineering Department. His research interests include wireless communication systems, spread spectrum techniques, CDMA and digital signal processing. He has published over 200 international journal and conference papers.

Dr. Ng was the General Chair of ISCAS'97 and the VP-Region 10 of the IEEE CAS Society (1999 and 2000). He was an Executive Committee Member and a Board Member of the IEE Informatics Divisional Board (1999–2001) and was a member of IEE Council (1999–2001). He was awarded the Honorary Doctor of Engineering Degree by the University of Newcastle, Australia, in 1997 for his services to higher education generally and to engineering education specifically. He received the Senior Croucher Foundation Fellowship in 1999 and the IEEE Third Millenium medal in February 2000.