Maximum Likelihood Timing and Carrier Frequency Offset Estimation for OFDM Systems With Periodic Preambles

Hung-Tao Hsieh, *Student Member, IEEE*, and Wen-Rong Wu, *Member, IEEE*

Abstract—Symbol timing offset (STO) and carrier frequency offset (CFO) estimation are two main synchronization operations in packet-based orthogonal frequency division multiplexing (OFDM) systems. To facilitate these operations, a periodic preamble is often placed at the beginning of a packet. CFO estimation has been extensively studied for the case of two-period preambles. In some applications, however, a preamble with more than two periods is available. A typical example is the IEEE802.11a/g wireless local area network system, which features a ten-period preamble. Recently, researchers have proposed a maximum likelihood (ML) CFO estimation method for such systems. This approach first estimates the received preamble using a least squares method and then maximizes the corresponding likelihood function. In addition to the standard calculations, this method requires an extra procedure to solve the roots of a polynomial function, which is disadvantageous for real-world implementations. In this paper, we propose a new ML method to solve the likelihood function directly and thereby perform CFO estimation. Our method can obtain a closed-form ML solution, without the need for the root-finding step. We further extend the proposed method to address the STO estimation problem as well as derive a lower bound on the estimation performance. Our simulations show that while the performance of the proposed method is either equal to or better than the existing method, the computational complexity is lower.

Index Terms—Frequency offset, maximum likelihood (ML), orthogonal frequency division multiplexing (OFDM), synchronization.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) is known as an efficient modulation technique [21], [22]. However, the performance of OFDM systems is sensitive to both symbol timing offset (STO) [19], [20] and carrier frequency offset (CFO). STO will reduce the effective cyclic prefix (CP) length and induce intersymbol interference, while CFO will damage the orthogonality among subcarriers and thereby induce intercarrier interference. For typical OFDM receivers, STO and CFO have to be estimated and compensated before data detection can be conducted.

Estimation methods for CFO and STO in OFDM systems can be classified into the following two categories: 1) data aided and 2) blind. The former is more suitable for packet-based transmission, while the latter is appropriate for continuous transmission such as broadcasting. Blind methods exploit the periodic structure of CPs to accomplish the estimation task [1]–[7]. Data-aided methods insert a known preamble, or pilot symbol, in front of each data packet such that it can easily be used by the receiver to achieve synchronization [9]–[18]. In this paper, we consider only the data-aided method.

CFO estimation usually consists of a fractional part and an integer part. Most researchers focus on how to estimate the fractional part, which is also the focus of this paper. For integer part estimation, see [9] and [10]. It has been shown that the performance of OFDM systems is greatly affected by CFO [30], and an accurate CFO estimation is required for real-world applications. A maximum likelihood (ML) CFO estimator using a preamble with two identical pilot symbols was first proposed in [11]. Using the same periodic preamble and taking null subcarriers into consideration, Huang and Letaief [12] propose a method that is able to estimate both fractional and integer CFOs. To avoid the extra overhead required in [12], Schmidl and Cox [13] introduce a preamble composed of two OFDM symbols: The first one has two identical periods (to estimate the fractional CFO and STO), and the second one has a special correlation with the first one (to estimate the integer CFO). To improve the performance, Morelli and Mengali [14] extend this area of research to treat preambles with periodicities of greater than two. Using the approach in [14], one can remove the second pilot symbol as required in [13]. As an improved version, Minn *et al.* [15] propose a CFO estimation based on the best linear unbiased estimation principle. Note that Morelli and Mengali [14] and Minn *et al.* [15] still use the same STO estimator as that in [13]. When the number of periods is greater than two, the method in [11] is no longer optimal. An ML CFO estimator for this problem was proposed in [16]. However, the required computational complexity is high. To alleviate this problem, a low-complexity approach was then proposed in [17]. Another simplified algorithm was also proposed in [18]. However, due to excessive approximation in the likelihood function, the performance of the CFO estimation in [18] does not approach the Cramér–Rao lower bound (CRLB) [25].

In this paper, we focus on CFO and STO estimation in the OFDM system with a periodic preamble. Specifically, we consider a preamble with more than two periods. The ML CFO

estimation for the system has been considered in [17]. The method in [17] is essentially a two-step approach: it first estimates the received preamble with a least squares (LS) method and then maximizes the corresponding likelihood function. In addition to regular computations, this method requires an extra procedure to solve for the roots of the derivative of the likelihood function. Thus, its computational complexity is higher, and the cost for real-world implementations is also increased.

In this paper, we develop a new ML method that solves the likelihood function directly for the CFO-estimation problem. Our method generates a closed-form ML solution, and the root-finding procedure is not required. As a result, the computational complexity and the implementation cost are lower than those in [17], while the performance of the proposed method is either equal to or better than that in [17]. The proposed method is further extended to STO estimation, and a theoretical lower performance bound is derived. Note that the performance bound for STO estimation has not previously been addressed in the literature. This paper is organized as follows. In Section II, the CFO-estimation method in [17] is briefly reviewed. The proposed CFO- and STO-estimation procedures are described in Sections III and IV. A lower bound on STO estimation performance is presented in Section V. Our simulation results are reported and discussed in Section VI. Our conclusions are presented in Section VII.

II. EXISTING APPROACH

In this section, we briefly review the algorithm proposed in [17]. Let the preamble in the OFDM system be periodic with period N and length QN. Denote the preamble signal as $s(k)$, where $k = 0, 1, \ldots, QN - 1$. The preamble is placed at the beginning of a packet and is subsequently transmitted through a wireless channel. Denote the channel response as $h(k)$ and the output signal as $x(k)$. Then, we have $x(k) = s(k) * h(k)$, where $*$ denotes the convolution operation. Assume that the maximum channel delay is N. Then, we can discard the first received N samples and retain the periodic property of the preamble $x(k)$. Thus, the received preamble can be expressed as [1]

$$y(k) = e^{\frac{j2\pi\varepsilon k}{N}} x(k) + w(k) \tag{1}$$

where $k = N, N+1, \ldots, QN - 1$, ε is CFO, and $w(k)$ represents additive white Gaussian noise with a variance of σ_w^2. We can perform an index transformation by letting $k = mN + n$, where $m = 1, \ldots, Q$ and $n = 0, \ldots, N - 1$ such that $x(k) = x(mN + n)$. For notational simplicity, we further let $x_m(n) = x(mN + n)$, denoting the nth sample of the mth period of $x(k)$. Due to periodicity, we have $x_p(n) = x_q(n)$ for $p, q \in \{1, \ldots, Q\}$. Similarly, we can define $y_m(n) = y(mN + n) = y(k)$, and $w_m(n) = w(mN + n) = w(k)$. Let $K = Q - 1$, and

$$\mathbf{y}(n) = \begin{bmatrix} y_1(n) & y_2(n) & \cdots & y_K(n) \end{bmatrix}^T$$
$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) & x_2(n) & \cdots & x_K(n) \end{bmatrix}^T$$
$$\mathbf{w}(n) = \begin{bmatrix} w_1(n) & w_2(n) & \cdots & w_K(n) \end{bmatrix}^T. \tag{2}$$

In addition, we define four matrices as follows:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}(0) & \mathbf{y}(1) & \cdots \mathbf{y}(N-1) \end{bmatrix}$$
$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(0) & \mathbf{x}(1)e^{\frac{j2\pi\varepsilon}{N}} & \cdots \mathbf{x}(N-1)e^{\frac{j2\pi\varepsilon(N-1)}{N}} \end{bmatrix}$$
$$\mathbf{W} = \begin{bmatrix} \mathbf{w}(0) & \mathbf{w}(1) & \cdots \mathbf{w}(N-1) \end{bmatrix}$$
$$\mathbf{A} = \begin{bmatrix} e^{j2\pi\varepsilon} & 0 & \cdots & 0 \\ 0 & e^{j2\pi\varepsilon \cdot 2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{j2\pi\varepsilon \cdot K} \end{bmatrix}. \tag{3}$$

The received preamble in (1) can then be rewritten as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{W}. \tag{4}$$

The method in [17] uses a two-step approach: it first estimates $\mathbf{X}$ using an LS method and then estimates CFO by maximizing the likelihood function. Since the noise is a Gaussian random variable, $\mathbf{y}(n)$ is a Gaussian random vector with a covariance matrix of $\sigma_w^2 \mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. For a given $\mathbf{A}$, the LS estimate of $\mathbf{X}$ can be expressed as $\mathbf{X}_{LS} = (1/K)\mathbf{A}^H\mathbf{Y} \equiv \mathbf{A}^+\mathbf{Y}$, where $(\cdot)^H$ denotes the Hermitian operation. Substituting $\mathbf{X}_{LS}$ back into (4), we can obtain the log-likelihood function as $\Lambda(\mathbf{A}) = \sum_{n=1}^{N} \| \mathbf{y}(n) - \mathbf{A}\mathbf{A}^+\mathbf{y}(n) \|^2 = N \cdot \text{trace}((\mathbf{I} - \mathbf{A}\mathbf{A}^+)\mathbf{R}_Y)$, where $\mathbf{R}_Y = E[\mathbf{Y}\mathbf{Y}^H]$. The (p,q)th entry of $\mathbf{R}_Y$ is $(1/N) \sum_{n=0}^{N-1} y_p(n)y_q^*(n), p, q \in [1, K]$ [28]. The desired CFO estimation can then be derived as

$$\hat{\varepsilon} = \arg\left\{ \min_\varepsilon \text{trace}\left((\mathbf{I} - \mathbf{A}\mathbf{A}^+)\mathbf{R}_Y\right) \right\}$$
$$= \arg\{\max_\varepsilon \mathbf{a}^H \mathbf{R}_Y \mathbf{a}\} \tag{5}$$

where $\mathbf{a}$ is a vector consisting of the diagonal elements of $\mathbf{A}$. It was shown in [26] that

$$\mathbf{a}^H \mathbf{R}_Y \mathbf{a} = \sum_{m=-(K-1)}^{K-1} b(m)e^{j2\pi m\varepsilon} \tag{6}$$

where $b(m) = \sum_{q-p=m}(1/N) \sum_{n=0}^{N-1} y_p(n)y_q^*(n)$. Taking the derivative of (6) with respect to ε and letting the result be zero, we obtain

$$\sum_{m=1}^{K-1} mb(m)z^m = \sum_{m=1}^{K-1} mb(-m)z^{-m} \tag{7}$$

where $z = e^{j2\pi\varepsilon}$. Equation (7) can be rewritten as

$$\text{Im}\left(\sum_{m=1}^{K-1} mb(m)z^m\right) = 0 \tag{8}$$

where $\text{Im}(\cdot)$ is an operator that isolates the imaginary part of a scalar value. Denote the set containing the roots of (8) by Ω. The CFO can then be estimated as follows [17]:

$$\hat{\varepsilon} = \frac{1}{j2\pi}\ln(\hat{z}) \tag{9}$$

where $\hat{z} = arg\{\max_{z \in \Omega}(\Lambda(z))\}$, and $|\hat{z}| = 1$.

The procedure for CFO estimation in [17] can now be summarized as follows.

1) Construct the correlation matrix $\mathbf{R}_Y$.
2) Calculate the coefficient of (8) using $\mathbf{R}_Y$.
3) Find the nonzero roots of (8).
4) Substitute the roots into (6), find the maximum root, and calculate $\hat{\varepsilon}$ using (9).

As we see, (8) requires a root-finding operation. Thus, a set of suboptimum algorithms to address this issue was proposed in [17]. Unfortunately, these suboptimum methods cannot effectively reduce the computational complexity while still maintaining good performance.

III. PROPOSED ML CFO ESTIMATION

In this section, we develop a new CFO estimation method that solves the likelihood function directly. The signal model we use is the same as that in (4). We assume that each data packet is transmitted through a slow-fading channel with an impulse response of $h(k)$, $k = 0, \ldots, L-1$. Here, the $h(k)$s have Rayleigh distributions, and they are statistically independent. Note that the time-domain preamble signal is obtained from the discrete Fourier transform of the frequency-domain preamble signal, and the frequency-domain preamble signal is generally a white sequence. From the central limit theorem, the time-domain preamble signal can then be approximated as a white Gaussian sequence. Thus, the channel output $x(k)$, which equals $\sum_{l=0}^{L-1} h(l)s(k-l)$, and the received preamble $y(k)$ in (1) can be approximated as Gaussian sequences. Let the variance of the time-domain preamble signal, i.e., $s(k)$ be σ_s^2. Then, the variance of $x(n)$ equals $\sigma_x^2 = E\{\sum_{j=0}^{L-1} \sum_{l=0}^{L-1} h(j)s(k-j)h(l)^*s(k-l)^*\} = \sigma_s^2 \sum_{l=0}^{L-1} |h(l)|^2 = \sigma_s^2 \sigma_h^2$, and that of $y(k)$ equals $\sigma_x^2 + \sigma_w^2$. Note that $s(k)$ can be a psuedonoise sequence. In such a case, σ_s^2 indicates the averaged preamble power of $s(k)$.

Let $f(\cdot)$ be a probability density function. Then, we explicitly write out the log-likelihood function of ε as follows [1]:

$$
\begin{aligned}
\Lambda(\varepsilon) &= \ln\left\{\prod_{n\in\tilde{I}} f(\mathbf{y}(n))\right\} \\
&= \ln\left\{\frac{\prod_{n\in\tilde{I}} f(\mathbf{y}(n))}{\prod_{m\in[1,K]}\prod_{n\in\tilde{I}} f(y_m(n))}\right. \\
&\qquad \left. \cdot \prod_{m\in[1,K]}\prod_{n\in\tilde{I}} f(y_m(n))\right\} \\
&= \ln\left\{\prod_{n\in\tilde{I}} \frac{f(\mathbf{y}(n))}{f(y_1(n))\cdots f(y_K(n))}\right. \\
&\qquad \left. \cdot \prod_{m\in[1,K]}\prod_{n\in\tilde{I}} f(y_m(n))\right\}.
\end{aligned}
\tag{10}
$$

It is clear that the last term in (10), i.e., $\prod_{m\in[1,K],n\in\tilde{I}} \times f(y_m(n))$, is independent of ε [1]. As a result, this term can be dropped. Let

$$
\mathbf{u}(n) = e^{\frac{j2\pi\varepsilon n}{N}}\left[x_1(n)e^{j2\pi\varepsilon} \quad \cdots \quad x_K(n)e^{j2\pi\varepsilon\cdot K}\right]^T. \tag{11}
$$

We then rewrite (4) as $\mathbf{Y} = \mathbf{U} + \mathbf{W}$, where $\mathbf{U} = \mathbf{AX} = [\mathbf{u}(0), \mathbf{u}(1), \ldots, \mathbf{u}(N-1)]$. Then, $\mathbf{y}(n) = \mathbf{u}(n) + \mathbf{w}(n)$. Define $\mathbf{R}_u = E[\mathbf{u}(n)\mathbf{u}^H(n)]$ and $\mathbf{R}_y = E[\mathbf{y}(n)\mathbf{y}^H(n)]$. Then, we have

$$
\mathbf{R}_u = \sigma_x^2 \begin{bmatrix} 1 & e^{-j2\pi\varepsilon} & \cdots & e^{-j2\pi(K-1)\varepsilon} \\ e^{j2\pi\varepsilon} & 1 & \cdots & e^{-j2\pi(K-2)\varepsilon} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j2\pi(K-1)\varepsilon} & e^{j2\pi(K-2)\varepsilon} & \cdots & 1 \end{bmatrix}
\tag{12}
$$

and $\mathbf{R}_y = \mathbf{R}_u + \sigma_w^2\mathbf{I}$, where $\mathbf{I}$ is an identical matrix. Thus, we can express $f(\mathbf{y}(n))$ as [23], [24]

$$
f(\mathbf{y}(n)) = \left(\pi^K \det(\mathbf{R}_y)\right)^{-1} \exp\left[-\mathbf{y}(n)^H\mathbf{R}_y^{-1}\mathbf{y}(n)\right]. \tag{13}
$$

According to the matrix inversion lemma [8], we derive the inverse of $\mathbf{R}_y$ as

$$
\mathbf{R}_y^{-1} = \sigma_w^{-2}\mathbf{I} - \frac{\sigma_w^{-4}\mathbf{R}_u}{1 + \sigma_w^{-2}E\{\mathbf{u}^H\mathbf{u}\}}. \tag{14}
$$

Note that for $n \in \tilde{I}$, we have

$$
E\{y_p(n)y_q^*(n)\} = \begin{cases} \sigma_x^2 + \sigma_w^2, & \text{if } q-p = 0 \\ \sigma_x^2 e^{-j2\pi\varepsilon(q-p)}, & \text{if } q-p \neq 0 \end{cases}
\tag{15}
$$

where $p, q \in [1, K]$. As a result, $\mathbf{R}_y^{-1} = \sigma_w^{-2}\mathbf{I} - [\mathbf{R}_u/(\sigma_w^4 + K\sigma_w^2\sigma_x^2)]$, and

$$
f(y_p(n)) = \frac{\exp\left(-\frac{y_p(n)y_p^*(n)}{\sigma_x^2 + \sigma_w^2}\right)}{\pi(\sigma_x^2 + \sigma_w^2)}
\tag{16}
$$

where $p \in [1, K]$. Thus, the exponential term in (13) becomes

$$
\begin{aligned}
\mathbf{y}(n)^H\mathbf{R}_y^{-1}\mathbf{y}(n) &= \sigma_w^{-2}\sum_{p=1}^{K} y_p(n)y_p^*(n) \\
&\quad - C_0 \sum_{p=1}^{K}\sum_{q=1}^{K} y_p(n)y_q^*(n)e^{j2\pi(q-p)\varepsilon} \\
&= \left(\sigma_w^{-2} - C_0\right)\sum_{p=1}^{K} y_p(n)y_p^*(n) \\
&\quad - 2C_0 \text{Re}\left\{\sum_{p=1}^{K-1}\sum_{q>p}^{K} y_p(n)y_q^*(n)e^{j2\pi(q-p)\varepsilon}\right\}
\end{aligned}
\tag{17}
$$

where $C_0 = \sigma_x^2/(\sigma_w^4 + K\sigma_w^2\sigma_x^2)$, and $\text{Re}\{\cdot\}$ denotes the operation that isolates the real part of the indicated complex variable.

Dropping the superfluous terms and substituting (13)–(17) into (10), we finally express the log-likelihood function as

$$
\Lambda(\varepsilon) = \sum_{n=0}^{N-1} \ln \left\{ \frac{\left(\sigma_x^2 + \sigma_w^2\right)^K \exp\left[-\mathbf{y}(n)^H \mathbf{R}_y^{-1} \mathbf{y}(n)\right]}{\det(\mathbf{R}_y) \exp\left[-\frac{\sum\limits_{p=1}^{K} y_p(n) y_p^*(n)}{\sigma_x^2 + \sigma_w^2}\right]} \right\}
$$

$$(18)$$

$$
= C_1 + C_2 \phi + C_3 \sum_{p=1}^{K-1} \sum_{q>p}^{K} |\gamma_{pq}| \cos(\psi_{pq}) \tag{19}
$$

where

$$
\gamma_{pq} = \sum_{n=0}^{N-1} y_p(n) y_q^*(n) \quad (q \geq p \text{ and } p \geq 1) \tag{20}
$$

$$
\psi_{pq} = 2\pi\varepsilon(q-p) + \angle\gamma_{pq},
$$

$$
\phi = \sum_{p=1}^{K} \gamma_{pp} \tag{21}
$$

$$
C_1 = N \cdot \ln\left(\frac{\left(\sigma_x^2 + \sigma_w^2\right)^K}{\det(\mathbf{R}_y)}\right) \tag{22}
$$

$$
C_2 = (1-K)\frac{\rho^2}{\sigma_w^2\left(1 + (K-1)\rho\right)} \tag{23}
$$

$$
C_3 = \frac{2C_2}{(1-K)\rho} \tag{24}
$$

$$
\rho = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_w^2}. \tag{25}
$$

Note that ϕ is the received signal energy and that $\det(\mathbf{R}_y)$ is a constant, independent of ε. The detailed derivation of (19) is provided in Appendix A. Ignoring unrelated terms, we obtain the log-likelihood as

$$
\Lambda(\varepsilon) \propto \sum_{p=1}^{K-1} \sum_{q>p}^{K} |\gamma_{pq}| \cos(\psi_{pq}). \tag{26}
$$

To maximize the function, we first take a derivative of $\Lambda(\varepsilon)$ with respect to ε and obtain

$$
\frac{\partial}{\partial\varepsilon}\Lambda(\varepsilon) = -\sum_{p=1}^{K-1} \sum_{q>p}^{K} 2\pi(q-p)|\gamma_{pq}| \sin(\psi_{pq}). \tag{27}
$$

Thus, we have an alternative expression to that in (8). Now, the problem is how to solve (27). Since (27) involves a non-linear sine function, a closed-form solution will be difficult to

calculate. Here, we use a simple approximation method to overcome the problem. Using (20) and (1), we obtain

$$
\gamma_{pq} = e^{j2\pi\varepsilon(p-q)} \sum_{n=0}^{N-1} |x_1(n)|^2 + \sum_{n=0}^{N-1} w_p(n) w_q^*(n)
$$

$$
+ e^{j2\pi\varepsilon(pN+p)} \sum_{n=0}^{N-1} x_1(n) w_p^*(n)
$$

$$
+ e^{j2\pi\varepsilon(pN-q)} \sum_{n=0}^{N-1} x_1^*(n) w_q(n). \tag{28}
$$

In (28), we have used the periodic property that $x_1(n) = x_p(n) = x_q(n)$. Now, if the noise level is low, the noise related terms in (28) can be ignored. We then have

$$
\angle\gamma_{pq} \approx 2\pi\varepsilon(p-q). \tag{29}
$$

From (29), we write

$$
\psi_{pq} \approx 2\pi\varepsilon(q-p) + 2\pi\varepsilon(p-q) = 0. \tag{30}
$$

From (30), we can then assume that $\sin(\psi_{pq}) \approx \psi_{pq}$ and approximate the expression in (27) by

$$
\frac{\partial}{\partial\varepsilon}\Lambda(\varepsilon) \simeq -\sum_{p=1}^{K-1} \sum_{q>p}^{K} 2\pi(q-p)|\gamma_{pq}|(\psi_{pq}). \tag{31}
$$

Setting the result in (31) to zero, we can estimate CFO as

$$
\hat{\varepsilon} = -\frac{\sum\limits_{p=1}^{K-1} \sum\limits_{q>p}^{K} |\gamma_{pq}|(q-p)\angle\gamma_{pq}}{2\pi \sum\limits_{p=1}^{K-1} \sum\limits_{q>p}^{K} |q-p|^2 |\gamma_{pq}|}. \tag{32}
$$

Note that the approximation in (30) will become exact if noise is not present and if ε is the true CFO. In other words, (27) and (31) will have the same zero-crossing point although the two functions are different, indicating that (31) and (27) will yield the same optimum solution. If noise is present, however, (31) and (27) will not have the same optimum solution. The accuracy of the solution in (31) depends on the signal-to-noise ratio (SNR) in (28). We define the SNR in (28) as SNR_γ and that in (1) as SNR. Then, $\text{SNR} = \sigma_x^2/\sigma_w^2$, as typically defined. From (28), it is simple to see that

$$
\text{SNR}_\gamma = \frac{N^2\sigma_x^4}{N\sigma_w^4 + 2N\sigma_x^2\sigma_w^2} = \frac{N \cdot \text{SNR}^2}{1 + 2\text{SNR}}. \tag{33}
$$

From (33), we can see that SNR_γ can be much larger than SNR as long as N is reasonably large and SNR is not very low. Subsequently, the approximation in (31) will introduce only a small error for a wide SNR range. As a simple example, let $N = 16$ and SNR = 0 dB. From (33), we obtain $\text{SNR}_\gamma = 7.27$ dB, which is much higher than SNR.

Note that the proposed estimate requires that we extract the phase from γ_{pq}. It is simple to see that the result is only unambiguous when $|\angle\gamma_{pq}| < \pi$. For a particular combination of p and q, the estimation range for CFO is $|\varepsilon| \leq 1/[2(q-p)]$.

TABLE I
COMPUTATIONAL COMPLEXITY COMPARISON FOR THE ALGORITHM IN [17] AND FOR THE PROPOSED ALGORITHMS

	Algorithm A in [17]	Algorithm A' in [17]	Algorithm B in [17]	Proposed algorithm I	Proposed algorithm II
No. of MUL	$2NK^2 + 7K - 10$	$2N(5K-6)+11$	$2K(NK+3)-6$	$(2NK + \frac{K}{2} + 3)(K-1) - 2$	$NQ(2N + \frac{7Q}{2} + \frac{1}{2}) - 4$
No. of ADD	$K(2NK+3)-5$	$2N(5K-6)+4$	$2NK^2+2K-3$	$K(K-1)(2N + \frac{K}{6} - \frac{1}{3}) + K - 5$	$QN(2N + \frac{Q^2}{6} + 4Q-6)-(Q^2 - Q + 8) + 5N$
No. of LN	1	1	1	0	0
No. of ABS	0	0	0	$\frac{K(K-1)}{2}$	$\frac{NQ(Q-1)}{2}$
No. of PH	0	0	0	$\frac{K(K-1)}{2}$	$\frac{NQ(Q-1)}{2}$
No. of DIV	1	1	2	1	N

Since the maximum value for $q - p$ is $K - 1$, the estimation range for CFO is $|\varepsilon| \leq 1/[2(K-1)]$. When K is large, the range becomes small. In the following, we propose a method to remedy this problem. The basic idea is to apply the phase-unwrapping procedure. We first calculate the phase angle for each γ_{pq}. Then, for each p, we calculate the phase difference of $\angle\gamma_{pq}$, $q = p + 1, p + 2, \ldots, K$. Let $d_{r,s}$ denote the phase difference, i.e., $d_{r,s} = \angle\gamma_{rs} - \angle\gamma_{r(s-1)}$, $r = 1, 2, \ldots, K - 2$ and $s = r + 2, r + 3, \ldots, K$. Since the maximum value of $|d_{r,s}|$ is π, whenever $|d_{r,s}| > \pi$, the phase need to be unwrapped. This can be performed with the following operation:

$$d_{r,s} = \begin{cases} d_{r,s} - 2\pi & \text{if } d_{r,s} > \pi \\ d_{r,s} + 2\pi & \text{if } d_{r,s} < -\pi. \end{cases} \quad (34)$$

For a value of r, the $d_{r,s}$ values should have the same signs. We can use this property to further correct occasional errors. Let g be the sum of all $d_{r,s}$ values, i.e., $g = \sum_{r=1}^{K-2} \sum_{s=r+2}^{K} d_{r,s}$. Then, we use the sign of g to determine the sign of $d_{r,s}$ and to evaluate $\angle\gamma_{pk}$, $k = p + 1, \ldots, K$. Finally, the unwrapped $\angle\gamma_{pq}$ can be written (with $q \geq p + 2$) as

$$\angle\gamma_{pq} = \angle\gamma_{p(p+1)} + \sum_{s=p+2}^{q} d_{p,s}. \quad (35)$$

Substituting (35) into (32), we can estimate CFO. Using our proposed procedure, the CFO estimation range can be greatly extended up to $|\varepsilon| < 1/2$.

Now, the procedure for our proposed ML CFO estimation can be summarized as follows.

1) Construct all γ_{pq}'s, where $p \in [1, K - 1]$ and $q \in [p + 1, K]$, and calculate their amplitude.
2) Use the phase unwrapping scheme to estimate the phase of γ_{pq}.
3) Substitute the results into (32), and calculate the ML estimate.

Clearly, the proposed estimate does not require the root-finding procedure, and this, in turn, effectively reduces the computational complexity. Step 1) above is similar to the calculation of $\mathbf{R}$ in Section II. However, our method is easier since we only have to compute γ_{pq} for $q > p$.

In this paragraph, we compare the computational complexity of the proposed ML estimate with that of the algorithm in [17]. Three algorithms are proposed in [17], which are referred to as algorithms A, A', and B. While algorithm A is optimal,

algorithms A' and B are suboptimal. Table I summarizes this result. In Table I, MUL, ADD, LN, ABS, PH, and DIV denote the multiplication, addition, natural logarithm, absolute value, phase derivation, and division operations, respectively. In addition, the algorithm proposed in this section is referred to as proposed algorithm I, and the one in Section IV is termed proposed algorithm II. For the proposed algorithms, we consider the worst case in which all the phase differences $d_{r,s}$ need to be unwrapped. Fig. 1 shows several examples of how Q and N affect the complexity. Note that the computational complexity for the root-finding procedure in [17] is not included here. For convenience, we treat all operations other than addition as multiplications. As we can see, the computational complexity for the proposed algorithm is slightly lower than that for algorithms A and B in [17], and algorithm A' in [17] is the lowest. However, algorithm A' truncates the polynomials with order higher than two in (6), i.e., $\Lambda(z) = \sum_{m=-2}^{2} b(m)z^m$. This impacts the estimation accuracy. Note that we can always truncate the summation terms in (32) and thereby reduce the computational complexity of proposed algorithm I. Since suboptimum approaches are not our focus, we will not consider the details here. We will now discuss the computational complexity of the root-finding procedure. As shown in [27] and [29], the root-finding procedure requires $O(K^3)$ multiplications. Table I shows that the computational complexity of algorithm A is $O(NK^2)$. Thus, the computational complexity of the root-finding procedure will be high when K is large. Furthermore, its implementation cost will also be higher, since we may need dedicated electronic circuitry to implement this function.

It is well known that the performance of an unbiased estimator is bounded by the CRLB [25]. If the variance of an unbiased estimator reaches the CRLB, we consider the estimator efficient. Following the procedure to derive performance bounds in [25], we can calculate the CRLB for our CFO estimation procedure. Let $\hat{\varepsilon}$ be an estimate of ε. The CRLB for our CFO estimation is then

$$\text{CRLB}(\hat{\varepsilon}) = -\frac{1}{E\left[\frac{\partial^2}{\partial\varepsilon^2}\Lambda(\varepsilon)\right]}$$

$$= \frac{(8\pi^2\rho)^{-1}\sigma_w^2\left(1 + (K-1)\rho\right)}{E\left[\sum_{p=1}^{K-1}\sum_{q>p}^{K}(q-p)^2\text{Re}\left\{\gamma_{pq}e^{j2\pi\varepsilon(q-p)}\right\}\right]}$$
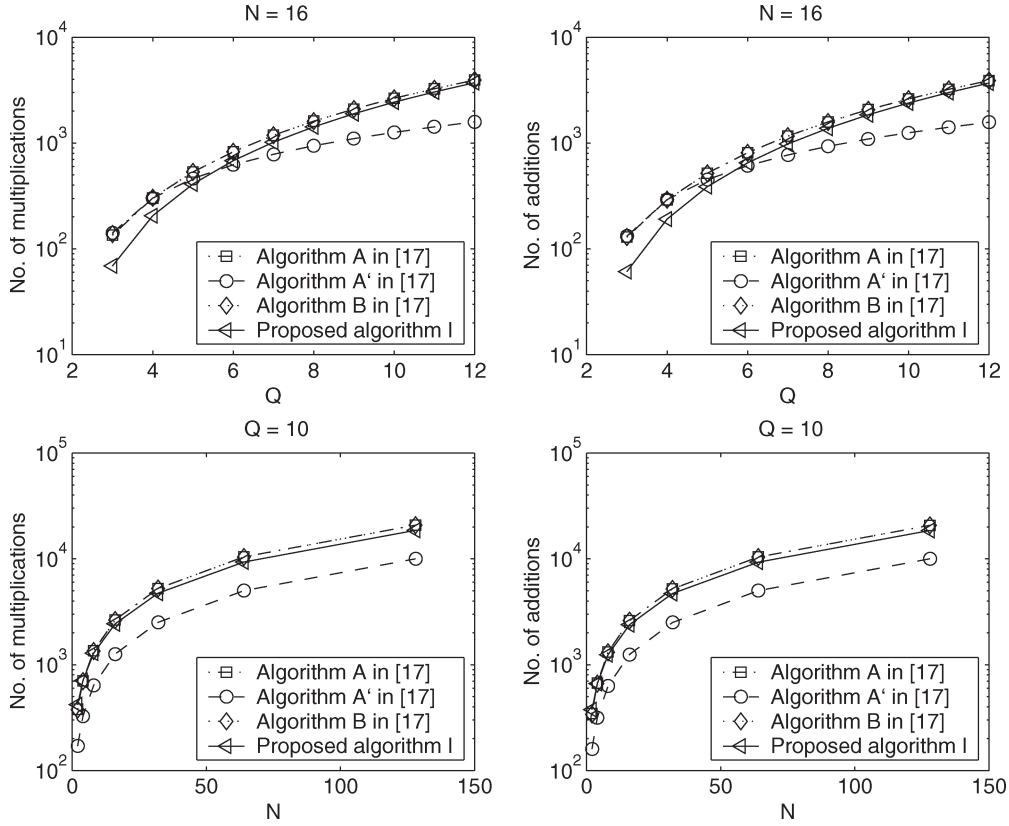
Fig. 1. Computational complexity comparison for the algorithm in [17] and proposed algorithm I. Note that the complexity of the root-finding procedure is not considered in [17].

$$
= \frac{\sigma_w^2 \left(1 + (K-1)\rho \right)}{8\pi^2 \rho N \sigma_x^2 \sum\limits_{p=1}^{K-1} \sum\limits_{q>p}^{K} (q-p)^2}
$$

$$
= \frac{1 + K \cdot \text{SNR}}{8\pi^2 N \cdot \text{SNR}^2} \frac{1}{\sum\limits_{p=1}^{K-1} \sum\limits_{q>p}^{K} (q-p)^2} \tag{36}
$$

where $E[\cdot]$ denotes the expectation.

IV. PROPOSED JOINT ML STO AND CFO ESTIMATION

In this section, we extend the method developed in Section III to solve the STO-estimation problem. The core idea is to apply a sliding data window for the received $(Q+1)N$ samples; each window covers the preamble in the context of a particular timing offset. We perform the ML CFO estimation for data in each window and store the estimated CFO and the corresponding maximum log-likelihood. Thereafter, the estimated CFO with the largest log-likelihood is selected as the ML CFO estimate. The corresponding window position is taken as the ML STO estimate. Let the window size be QN, and define the set $V_i = \{y(i), y(i+1), \ldots, y(i+QN-1)\}$ to be the received data in window i. Since the maximum delay is shorter than N, it is clear that $0 \leq i \leq N-1$. If we let the STO be θ, V_θ will cover the complete preamble. In Appendix B,

we show that the log-likelihood function for V_i can be expressed by

$$
\Lambda^i(\varepsilon) = C_1^i + C_2^i \phi^i + C_3^i \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} \left| \gamma_{pq}^i \right| \cos\left(\psi_{pq}^i \right) \tag{37}
$$

where the superscript i indicates that all the variables are calculated within V_i, and C_1^i, C_2^i, and C_3^i can be treated as window independent. Thus, we can simplify the above log-likelihood function using

$$
\Lambda^i(\varepsilon) \approx C_2 \phi^i + C_3 \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} \left| \gamma_{pq}^i \right| \cos\left(\psi_{pq}^i \right) \tag{38}
$$

where C_2 and C_3 are the same as those in (23) and (24). Since the received signal power ϕ^i is independent of CFO, we can estimate CFO using (32) as

$$
\hat{\varepsilon}^i = -\frac{\sum\limits_{p=0}^{Q-2} \sum\limits_{q>p}^{Q-1} \left| \gamma_{pq}^i \right| (q-p) \angle \gamma_{pq}^i}{2\pi \sum\limits_{p=0}^{Q-2} \sum\limits_{q>p}^{Q-1} |q-p|^2 \left| \gamma_{pq}^i \right|}. \tag{39}
$$

Note that the upper bound in the summation terms of (39) is Q instead of K. The estimated STO is then

$$
\hat{\theta} = arg \left\{ \max_i \left(\Lambda^i(\hat{\varepsilon}^i) \right) \right\} = i_{\text{opt}}. \tag{40}
$$

Now, the procedure for the proposed joint ML STO and CFO estimation can be summarized as follows.

1) Calculate γ_{pq}^i and its amplitude, where $i \in [1, N]$, $p \in [0, Q-2]$, and $q \in [p+1, Q-1]$.
2) Use the phase unwrapping procedure outlined above to calculate $\angle\gamma_{pq}^i$.
3) Substitute the results into (38) and (39), and calculate $\Lambda^i(\hat{\varepsilon}^i)$ and $\hat{\varepsilon}^i$.
4) Find i_{opt} such that $\Lambda^{i_{\text{opt}}}(\hat{\varepsilon}^{i_{\text{opt}}}) > \Lambda^i(\hat{\varepsilon}^i)$, $i \neq i_{\text{opt}}$.
5) The ML STO estimate is i_{opt}, and the ML CFO estimate is then $\hat{\varepsilon}^{i_{\text{opt}}}$.

As we can see from the above procedure, the computational complexity of the algorithm will be N times higher than that in Section III. Note also that the upper limit of p is $Q-2$ instead of $K-2$. In other words, we have an extra period for CFO estimation. By leveraging the sliding window structure, we can effectively reduce the computational complexity in calculating γ_{pq}^i. Similar to the definition of γ_{pq}, we obtain $\gamma_{pq}^i = \sum_{n=i}^{i+N-1} y_p(n)[y_q(n)]^*$. Then, it is simple to show that

$$\gamma_{pq}^i = \gamma_{pq}^{i-1} + y_p(i+N-1)\left[y_q(i+N-1)\right]^* \\ -y_p(i-1)\left[y_q(i-1)\right]^*. \quad (41)$$

From (41), we can see that except for $i=0$, the calculation of γ_{pq}^i requires only two complex multiplications and two complex additions. This will greatly reduce the required computational complexity in the scenario of joint STO and CFO estimation. The required computational complexity has been summarized in Table I.

We can also obtain the CRLB for the CFO estimate. All we have to do is to replace K with Q in (36). Since $Q = K+1$, the CRLB is lower than that in (36). Note that the STO is a discrete value. No performance lower bounds have been reported to date in the literature. In Section V, we will derive a lower bound to address this omission.

V. PERFORMANCE ANALYSIS OF STO ESTIMATION

In this section, we analyze the performance of the proposed STO estimation method. We first redefine (38) as $\Lambda^i(\varepsilon) = C_2\phi^i + C_3\xi^i$, where

$$\phi^i = \sum_{p=0}^{K}\sum_{n=i}^{i+N-1} y_p(n)y_p^*(n)$$
$$= \sum_{p=0}^{K}\sum_{n=0}^{N-1} x_p(n)x_p^*(n) + w_p(n)w_p^*(n)$$
$$+ 2\text{Re}\left\{x_p(n)w_p^*(n)\exp\left(j2\pi\varepsilon\frac{pN+n}{N}\right)\right\} \quad (42)$$

$$\xi^i = \sum_{p=0}^{K-1}\sum_{q>p}^{K}|\gamma_{pq}^i|\cos(\psi_{pq}^i)$$
$$= \sum_{p=0}^{K-1}\sum_{q>p}^{K}\sum_{n=0}^{N-1} x_p(n)w_q^*(n)\exp\left(j2\pi\varepsilon\frac{qN+n}{N}\right)$$
$$+ w_p(n)x_q^*(n)\exp\left(-j2\pi\varepsilon\frac{pN+n}{N}\right)$$
$$+ w_p(n)w_q^*(n)\exp(j2\pi\varepsilon(q-p)) + x_p(n)x_q^*(n). \quad (43)$$

Note here that ϕ^i and ξ^i are random variables. The mean value of $\Lambda^i(\varepsilon)$, which is denoted by μ_Λ^i, is equal to $C_2\mu_\phi^i + C_3\mu_\xi^i$, where μ_ϕ^i and μ_ξ^i are the mean of ϕ^i and ξ^i, respectively. The variance of Λ^i can be expressed by $\nu_\Lambda^i = C_2^2\nu_\phi^i + C_3^2\nu_\xi^i + 2C_2C_3\kappa_{\phi\xi}^i$, where ν_ϕ^i and ν_ξ^i denote the variance of ϕ^i and ξ^i, respectively, and $\kappa_{\phi\xi}^i$ the covariance between ϕ^i and ξ^i. The whole set of V_i, $0 \leq i \leq N-1$, has $(Q+1)N$ samples, and it may cover three regions. The first region consists of the noise samples, the second region the periodic preamble samples, and the third region the data samples. We denote these regions by I_N, I_P, and I_D. Thus, the signal variance in I_N is σ_w^2, that in I_P is $\sigma_x^2 + \sigma_w^2$, and that in I_D is $\sigma_d^2 + \sigma_w^2$, where σ_d^2 represents the variance of data samples. Recall that θ is the actual STO in the system. Using θ as a reference, we can have the following three cases for the value of i: 1) $i=\theta$; 2) $i<\theta$; and 3) $i>\theta$ ($0 \leq i \leq N-1$). The statistics of ϕ^i and ξ^i are different across these three cases. In Appendix C, we provide a detailed derivation of μ_ϕ^i, μ_ξ^i, ν_ϕ^i, ν_ξ^i, and $\kappa_{\phi\xi}^i$.

For the proposed STO-estimation algorithm, an error occurs when $i_{\text{opt}} \neq \theta$. Thus, we can define the error probability of STO estimation as $P(\cup_{i,i\neq\theta}\{\Lambda^\theta < \Lambda^i\})$, where $P(\cdot)$ denotes the probability of a certain event. Note that the evaluation of $P(\Lambda^\theta < \Lambda^i)$ only requires 1-D integration. If the log-likelihood functions for all i's are independent and identically distributed, we have $P(\cup_{i,i\neq\theta}\{\Lambda^\theta < \Lambda^i\}) = \sum_{i,i\neq\theta} P(\Lambda^\theta < \Lambda^i)$. Unfortunately, the log-likelihood functions are not independent. As a result, we have to conduct multidimensional integration, which is both complex and difficult. Therefore, we propose a simple alternative to overcome the problem. Instead of the exact error probability, we attempt to derive a lower bound.

As shown in [13], the likelihood function is approximately Gaussian. We denote the distribution of Λ^i using $G(\mu_\Lambda^i, \nu_\Lambda^i)$, where $G(\cdot)$ denotes the Gaussian distribution. Consider the joint density function of Λ^i and Λ^j. Using the Gaussian assumption, we write the bivariate Gaussian distribution as

$$P(\Lambda^i, \Lambda^j) = \frac{1}{2\pi \cdot \nu_\Lambda^i \cdot \nu_\Lambda^j \cdot \sqrt{1-C_c(i,j)}} \\ \cdot \exp\left(-\frac{z_{ij}}{2(1-C_c(i,j))}\right) \quad (44)$$

where $1 \leq i, j \leq N$, and

$$z_{ij} = \frac{\left(\Lambda^i - \mu_\Lambda^i\right)^2}{\nu_\Lambda^i} + \frac{\left(\Lambda^j - \mu_\Lambda^j\right)^2}{\nu_\Lambda^j} \\ - \frac{2C_c(i,j)\left(\Lambda^i - \mu_\Lambda^i\right)\left(\Lambda^j - \mu_\Lambda^j\right)}{\sqrt{\nu_\Lambda^i \cdot \nu_\Lambda^j}} \quad (45)$$

$$C_c(i,j) = \frac{E\left\{\Lambda^i(\Lambda^j)^*\right\} - \mu_\Lambda^i\mu_\Lambda^{j^*}}{\sqrt{\nu_\Lambda^i \cdot \nu_\Lambda^j}}. \quad (46)$$
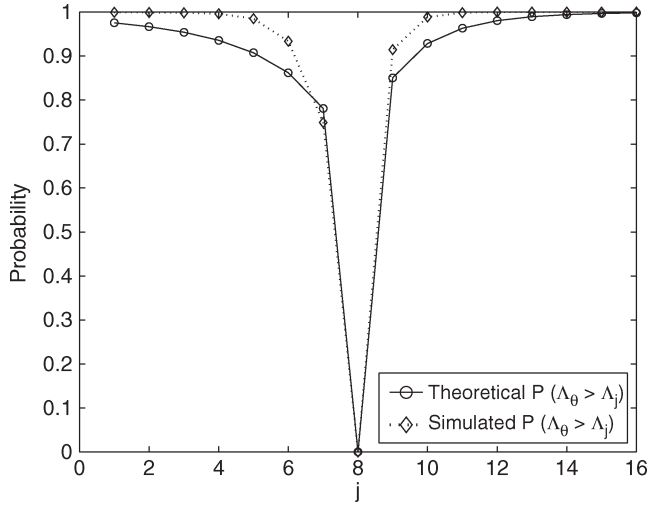
Fig. 2. Comparison of simulated and theoretical $P(\Lambda_\theta > \Lambda_j)$.

Note that $C_c(i, j)$ is the corresponding correlation coefficient. The numerator of $C_c(i, j)$ is expressed as

$$E\left\{\Lambda^i(\Lambda^j)^*\right\} = \mu_\Lambda^i \mu_\Lambda^{j*} + C_2^2 \kappa_{\phi\phi}^{ij} + C_3^2 \kappa_{\xi\xi}^{ij}$$
$$+ C_2 C_3 \kappa_{\phi\xi}^{ij} + C_2 C_3 \kappa_{\xi\phi}^{ij} \quad (47)$$

where κ_{ab}^{ij} denotes the covariance of a^i and b^{j*} ($a^i, b^j \in \{\phi^i, \phi^j, \xi^i, \xi^j\}$). The main idea here is only to calculate $P(\Lambda^\theta > \Lambda^i)$ for all i's (except for $i = \theta$) and then use the result to derive a lower bound. Thus, we only have to consider $C_c(i, \theta)$ as

$$\kappa_{\phi\phi}^{i\theta} = 2\sigma_x^2 \sigma_w^2 \left(QN - |i - \theta|\right) \quad (48)$$

$$\kappa_{\xi\xi}^{i\theta} = Q(Q-1)\sigma_x^2 \sigma_w^2 \left[\frac{N}{3}(2Q-1) - \frac{1}{2}|i - \theta|\right]$$
$$+ \frac{1}{2}QN(Q-1)\sigma_w^4 \quad (49)$$

$$\kappa_{\phi\xi}^{i\theta} = \kappa_{\xi\phi}^{i\theta} = (Q-1)^2 N \sigma_x^2 \sigma_w^2$$
$$+ (Q-1)\left(N - |i - \theta|\right)\sigma_x^2 \sigma_w^2. \quad (50)$$

Substituting (45)–(50) into (44), we can then evaluate $P(\Lambda^\theta > \Lambda^i)$. Given this definition, we have $P(\Lambda^\theta > \Lambda^i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\Lambda^\theta} P(\Lambda^i, \Lambda^\theta) d\Lambda^i d\Lambda^\theta$. Simulations have been conducted to evaluate the validity of our theoretical results. Using the scenario depicted in Section VI, we compare the theoretical and simulated $P(\Lambda^\theta > \Lambda^i)$ in Fig. 2. In the figure, we see that the theoretical $P(\Lambda^\theta > \Lambda^i)$ is close to the simulated result. If we let $P_{\min} = \min_{i \neq \theta} P(\Lambda^\theta > \Lambda^i)$, we can then treat $P_{\min}$ as an upper bound for the correct probability of STO estimation (i.e., $i_{\text{opt}} = \theta$). Thus, we can then have a lower bound for the error probability of STO estimation (LBSTO) as $1 - P_{\min}$.

VI. SIMULATIONS AND DISCUSSIONS

In this section, we report our simulation results, which evaluate the performance of the proposed algorithms. We adopt a Rayleigh multipath channel with an exponential power decay and five channel taps. The preamble, which is generated from a
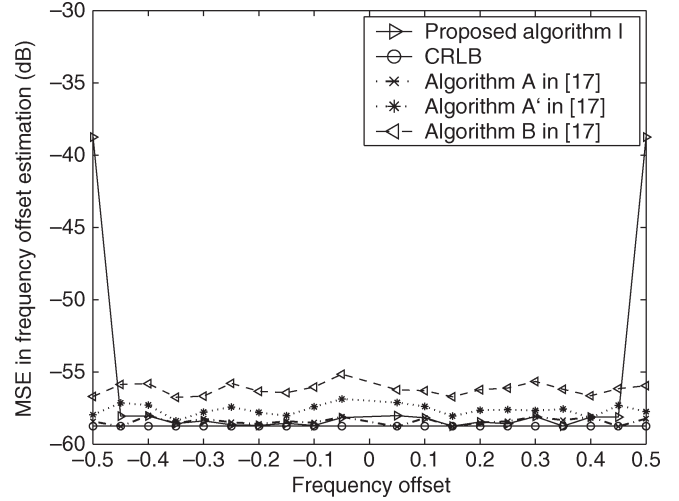

Fig. 3. Performance comparison of CFO estimation, the algorithm in [17], and proposed algorithm I; SNR = 10 dB.
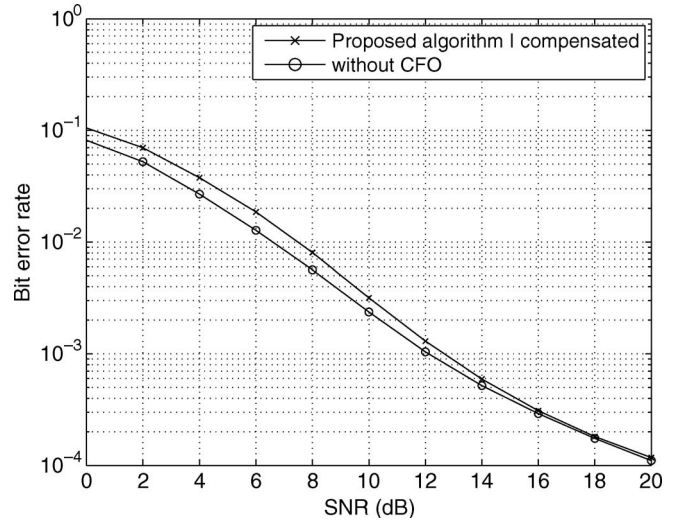

Fig. 4. BER comparison for systems with and without CFO.

frequency-domain binary-phase-shift-keying-modulated signal, has ten periods, and each period has 16 samples. The data following the preamble are transmitted using a 16-quadratic-amplitude-modulation scheme. The mean square error (MSE) of the estimated CFO is used as a performance measure. We first consider the CFO-only estimation problem. In this case, the first received N samples are discarded. As previously mentioned, we term the proposed approach for this scenario as algorithm I (as described in Section III). We compare the proposed ML estimator with that in [17]. One optimum algorithm (algorithm A) and two suboptimum algorithms (algorithm A′ and B) in [17] are simulated. Fig. 3 shows the simulation result for SNR at 10 dB. In the figure, we can see that the performances of algorithms A′ and B are poorer. Algorithm A and the proposed algorithm offer a similar level of performance that is very close to the CRLB. To evaluate the impact of CFO on system performance, we conduct simulations for systems with and without CFO. For the system with CFO, we first use the proposed method to estimate CFO and then conduct CFO compensation. Fig. 4 shows the BER comparison for $\varepsilon = 0.2$. As we can see in
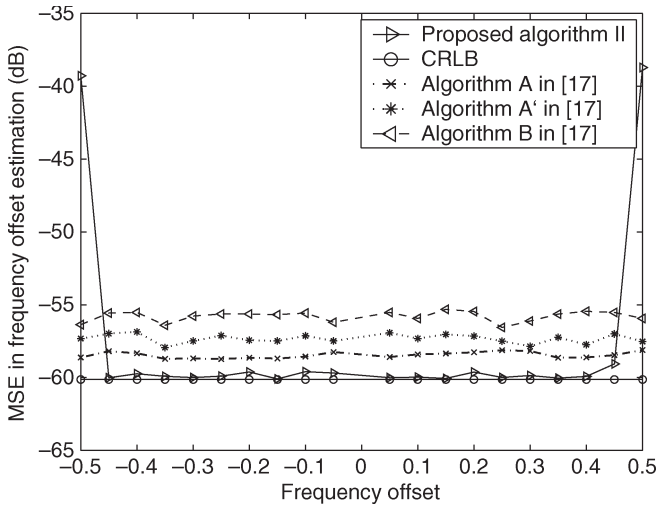
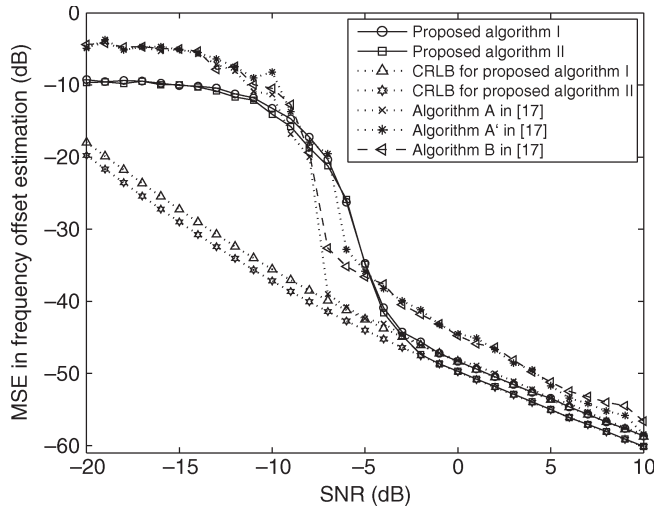Fig. 5. Performance comparison for CFO estimation, the algorithm in [17], and proposed algorithm II; SNR = 10 dB.

Fig. 7. Error probability of STO estimation (proposed algorithm II).

Fig. 6. Performance comparison for CFO estimation, the algorithm in [17], and proposed algorithms I and II; $N = 16$, and $Q = 10$.
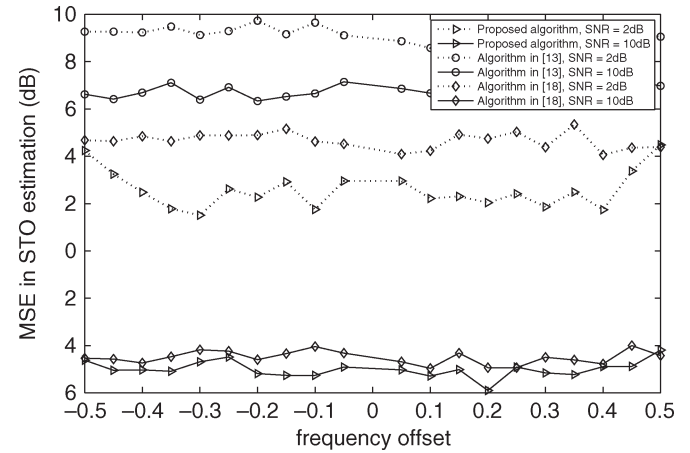
Fig. 8. Performance comparison for STO estimation (SNR = 2 and 10 dB).

the figure, the BER performance degrades slightly when CFO is present.

We then consider the case of the joint STO and CFO estimation process. In this case, discarding the first received N samples is not necessary. As a result, one additional preamble is available. This means that the proposed method may offer better performance compared with the previous scenario. However, the price we pay for the additional STO estimation is the increase in computational complexity. As mentioned, we name this approach proposed algorithm II (as explained in Section IV). Using a similar approach, the method in [17] can also be used to estimate STO. However, its computational complexity increases much more than our method. Fig. 5 shows the simulation result for the CFO estimate. The proposed method offers good performance. Only when CFO is very close to ± 0.5 does the performance of the proposed algorithms degrade. Fig. 6 shows the CFO estimation result for various SNRs. In the figure, we see that the proposed method still works well for SNRs as low as -5 dB. The algorithms in [17] perform
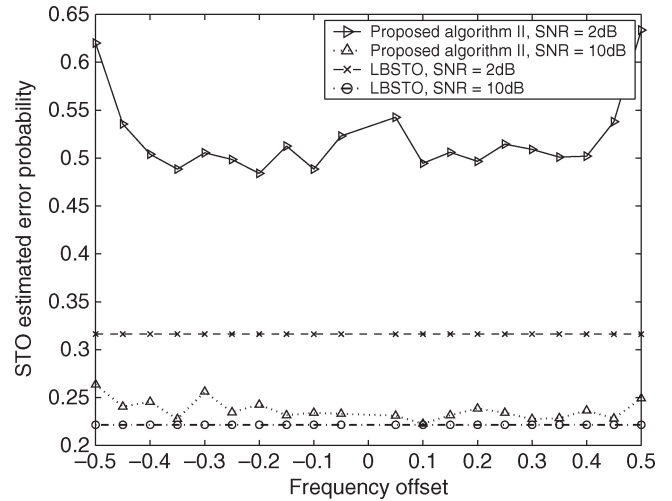
well until SNR reaches -7 dB, which is somewhat better than the proposed algorithms. However, when SNR falls below -8 dB, the proposed algorithms again outperform those in [17]. This may be because the correlation matrix in (6) is very noisy, and the roots therefore cannot be solved reliably. Fig. 7 shows the error probability for the STO estimation. We observe that the derived lower bound for the STO estimation is tight when the SNR is high. Note that the error probability we defined is only relevant to performance evaluation. If the channel response is shorter than the CP (which is the typical case), we can always has some tolerance for the STO estimation. Thus, there is no need to calculate the exact channel delay. In real-world applications, it is a common practice to reduce the estimated STO by a couple of samples when conducting STO compensation. Another property is that STO estimation performance is not particularly impacted when CFO is close to 0.5. In the literature, there exist a number of STO estimation methods. We select the two algorithms proposed in [13] and [18] for comparison. Fig. 8 shows the MSE curves for these approaches and for the proposed algorithms ($\theta = 8$). The figure confirms that the proposed method performs best.

VII. Conclusion

In this paper, we have developed new algorithms for ML STO and CFO estimation in OFDM systems with periodic preambles. The proposed algorithms do not have to calculate the roots of the derivative of the likelihood function. The operations are simple, and the computational complexity is low. With the proposed method, we can simultaneously solve the STO and CFO estimation problems. We also derive a lower bound for the STO estimation error. Simulations show that the proposed methods offer good performance, and the derived lower bound is tight when the SNR is high.

APPENDIX A
DERIVATION OF (19)

The likelihood function in (18) can be rewritten as

$$
\Lambda(\varepsilon) = \sum_{n=0}^{N-1} \ln \left\{ \frac{\left(\sigma_x^2 + \sigma_w^2\right)^K \exp\left[-\mathbf{y}(n)^H \mathbf{R}_y^{-1} \mathbf{y}(n)\right]}{\det(\mathbf{R}_y) \exp\left[-\frac{\sum_{p=1}^{K} y_p(n) y_p^*(n)}{\sigma_x^2 + \sigma_w^2}\right]} \right\}
$$

$$
= \sum_{n=0}^{N-1} \left\{ \ln\left[\left(\sigma_x^2 + \sigma_w^2\right)^K \left(\det(\mathbf{R}_y)\right)^{-1}\right] \right.
$$

$$
\left. + \frac{\sum_{p=1}^{K} y_p(n) y_p^*(n)}{\sigma_x^2 + \sigma_w^2} - \mathbf{y}(n)^H \mathbf{R}_y^{-1} \mathbf{y}(n) \right\}.
$$

(51)

Then, by substituting (17) into (51), we derive the log-likelihood function as

$$
\Lambda(\varepsilon) = \sum_{n=0}^{N-1} \left\{ \ln\left[\frac{\left(\sigma_x^2 + \sigma_w^2\right)^K}{\det(\mathbf{R}_y)}\right] + \frac{\sum_{p=1}^{K} y_p(n) y_p^*(n)}{\sigma_x^2 + \sigma_w^2} \right.
$$

$$
- \left(\sigma_w^{-2} - C_0\right) \sum_{p=1}^{K} y_p(n) y_p^*(n)
$$

$$
\left. + 2C_0 \mathrm{Re}\left\{ \sum_{p=1}^{K-1} \sum_{q>p}^{K} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\} \right\}
$$

$$
= \sum_{n=0}^{N-1} \left\{ \ln\left[\left(\sigma_x^2 + \sigma_w^2\right)^K \left(\det(\mathbf{R}_y)\right)^{-1}\right] \right.
$$

$$
+ \left[\frac{1}{\sigma_x^2 + \sigma_w^2} - \left(\sigma_w^{-2} - C_0\right)\right] \sum_{p=1}^{K} y_p(n) y_p^*(n)
$$

$$
\left. + 2C_0 \mathrm{Re}\left\{ \sum_{p=1}^{K-1} \sum_{q>p}^{K} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\} \right\}
$$

$$
= \frac{(1-K)\sigma_x^4}{\sigma_w^2 \left(\sigma_x^2 + \sigma_w^2\right)\left(K\sigma_x^2 + \sigma_w^2\right)} \sum_{p=1}^{K} \sum_{n=0}^{N-1} y_p(n) y_p^*(n)
$$

$$
+ 2C_0 \mathrm{Re}\left\{ \sum_{p=1}^{K-1} \sum_{q>p}^{K} \sum_{n=0}^{N-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}
$$

$$
+ N\left\{ \ln\left[\left(\sigma_x^2 + \sigma_w^2\right)^K \left(\det(\mathbf{R}_y)\right)^{-1}\right] \right\}
$$

(52)

where $C_0 = \sigma_x^2/(\sigma_w^4 + K\sigma_w^2\sigma_x^2)$. By substituting (21) and (25) into (52), we can express (52) as

$$
\Lambda(\varepsilon) = C_1 + C_2 \sum_{p=1}^{K} \gamma_{pp}
$$

$$
+ C_3 \mathrm{Re}\left\{ \sum_{p=1}^{K-1} \sum_{q>p}^{K} \gamma_{pq} e^{j2\pi(q-p)\varepsilon} \right\}
$$

$$
= C_1 + C_2 \sum_{p=1}^{K} \gamma_{pp}
$$

$$
+ C_3 \mathrm{Re}\left\{ \sum_{p=1}^{K-1} \sum_{q>p}^{K} \left(|\gamma_{pq}|e^{j\angle\gamma_{pq}}\right) e^{j2\pi(q-p)\varepsilon} \right\}
$$

$$
= C_1 + C_2\phi + C_3 \sum_{p=1}^{K-1} \sum_{q>p}^{K} |\gamma_{pq}|\cos(\psi_{pq})
$$

(53)

where ψ_{pq}, ϕ, C_1, C_2, and C_3 are defined as (22)–(24).

APPENDIX B
DERIVATION OF (38)

We assume that the channel noise, the received preamble, and the received data are statistically uncorrelated with one another. We define three column vectors $\mathbf{y}_1(n) = [y_0(n), \ldots, y_{Q-1}(n)]^T$, $\mathbf{y}_2(n) = [y_1(n), \ldots, y_{Q-1}(n)]^T$, and $\mathbf{y}_3(n) = [y_0(n), \ldots, y_{Q-2}(n)]^T$ and their autocorrelation matrix as $\mathbf{R}_{yk}$ for $k = 1, 2$, and 3. Note that i is the window index of (37), and θ is the real STO. Therefore, (19) can be derived for the following two cases: 1) $i \leq \theta$ and 2) $i > \theta$. Using the approach taken to derive (18), we obtain the log-likelihood function for the first case as

$$
\Lambda^{i\leq\theta}(\varepsilon) = \ln\left\{ \prod_{n=i}^{n=i+N-1} \frac{f(\mathbf{y}(n))}{\prod_{k=1}^{Q} f(y_{k-1}(n))} \right\}
$$

$$
= \sum_{n=i}^{\theta-1} \ln\left\{ \frac{f(\mathbf{y}_2(n))}{f(y_1(n))\cdots f(y_{Q-1}(n))} \right\}
$$

$$
+ \sum_{n=\theta}^{i+N-1} \ln\left\{ \frac{f(\mathbf{y}_1(n))}{f(y_0(n))\cdots f(y_{Q-1}(n))} \right\}
$$

(54)

$$
= \frac{\theta-i}{N} C_{12} + \frac{i+N-\theta}{N} C_1'
$$

$$
+ \sum_{n=i}^{\theta-1} C_2 \sum_{p=1}^{Q-1} y_p(n) y_p^*(n)
$$

$$+ \sum_{n=i}^{\theta-1} C_3 \mathrm{Re} \left\{ \sum_{p=1}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ \sum_{n=\theta}^{i+N-1} C_3' \mathrm{Re} \left\{ \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ \sum_{n=\theta}^{i+N-1} C_2' \sum_{p=0}^{Q-1} y_p(n) y_p^*(n) \tag{55}$$

where

$$C_{12} = N \cdot \ln \left(\frac{(\sigma_x^2 + \sigma_w^2)^K}{\det(\mathbf{R}_{y2})} \right) \tag{56}$$

$$C_1' = N \cdot \ln \left(\frac{(\sigma_x^2 + \sigma_w^2)^Q}{\det(\mathbf{R}_{y1})} \right) \tag{57}$$

$$C_2' = (1-Q) \frac{\rho^2}{\sigma_w^2 \left(1 + (Q-1)\rho\right)} \tag{58}$$

$$C_3' = =, \frac{2C_2'}{(1-Q)\rho}. \tag{59}$$

Similarly, we can derive the log-likelihood function for $i > \theta$ as

$$\Lambda^{i>\theta}(\varepsilon) = \sum_{n=\theta+N}^{i+N-1} \ln \left\{ \frac{f(\mathbf{y}_3(n))}{f(y_0(n)) \cdots f(y_{Q-2}(n))} \right\}$$
$$+ \sum_{n=i}^{\theta+N-1} \ln \left\{ \frac{f(\mathbf{y}_1(n))}{f(y_0(n)) \cdots f(y_{Q-1}(n))} \right\} \tag{60}$$
$$= \frac{i-\theta}{N} C_{13} + \frac{\theta+N-i}{N} C_1'$$
$$+ \sum_{n=\theta+N}^{i+N-1} C_2 \sum_{p=0}^{Q-2} y_p(n) y_p^*(n)$$
$$+ \sum_{n=\theta+N}^{i+N-1} C_3 \mathrm{Re} \left\{ \sum_{p=0}^{Q-3} \sum_{q>p}^{Q-2} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ \sum_{n=i}^{\theta+N-1} C_2' \sum_{p=0}^{Q-1} y_p(n) y_p^*(n)$$
$$+ \sum_{n=i}^{\theta+N-1} C_3' \mathrm{Re} \left\{ \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\} \tag{61}$$

where

$$C_{13} = N \cdot \ln \left(\frac{(\sigma_x^2 + \sigma_w^2)^K}{\det(\mathbf{R}_{y3})} \right).$$

Since $\mathbf{y}_2(n)$, $i \le n \le \theta - 1$ in (54) and $\mathbf{y}_3(n)$, $\theta + N \le n \le i + N - 1$ in (60) contain $Q - 1$ periods of the preamble, $\det(\mathbf{R}_{y2})$ and $\det(\mathbf{R}_{y3})$ will be the same as $\det(\mathbf{R}_y)$ [see (22)]. Consequently, $C_{12} = C_{13} = C_1$. From (22)–(24), we see that C_1, C_2, and C_3 can be calculated by replacing Q and $\mathbf{R}_{y1}$ with

K and $\mathbf{R}_y$, respectively, in (57)–(59). When Q is reasonably large, we obtain $C_1' \approx C_1$, $C_2' \approx C_2$, and $C_3' \approx C_3$. Thus, we rewrite (55) and (61) as

$$\Lambda^{i\le\theta}(\varepsilon) \simeq \sum_{n=i}^{\theta-1} C_3 \mathrm{Re} \left\{ \sum_{p=1}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ \sum_{n=\theta}^{i+N-1} C_3 \mathrm{Re} \left\{ \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ C_1 + \sum_{n=i}^{\theta-1} C_2 \sum_{p=1}^{Q-1} y_p(n) y_p^*(n)$$
$$+ \sum_{n=\theta}^{i+N-1} C_2 \sum_{p=0}^{Q-1} y_p(n) y_p^*(n) \tag{62}$$

$$\Lambda^{i>\theta}(\varepsilon) \simeq \sum_{n=i}^{\theta+N-1} C_3 \mathrm{Re} \left\{ \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ \sum_{n=\theta+N}^{i+N-1} C_3 \mathrm{Re} \left\{ \sum_{p=0}^{Q-3} \sum_{q>p}^{Q-2} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}$$
$$+ C_1 + \sum_{n=\theta+N}^{i+N-1} C_2 \sum_{p=0}^{Q-2} y_p(n) y_p^*(n)$$
$$+ \sum_{n=i}^{\theta+N-1} C_2 \sum_{p=0}^{Q-1} y_p(n) y_p^*(n). \tag{63}$$

We now approximate $\sum_{p=1}^{Q-1} y_p(n) y_p^*(n)$ and $\sum_{p=1}^{Q-2} \sum_{q>p}^{Q-1} \times y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon}$ in (62) with $\sum_{p=0}^{Q-1} y_p(n) y_p^*(n)$ and $\sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon}$, respectively. Similarly, we also approximate $\sum_{p=0}^{Q-2} y_p(n) y_p^*(n)$ and $\sum_{p=0}^{Q-3} \sum_{q>p}^{Q-2} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon}$ in (63) with $\sum_{p=0}^{Q-1} y_p(n) y_p^*(n)$ and $\sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon}$, respectively. Given these approximations, $\Lambda^{i\le\theta}(\varepsilon)$ and $\Lambda^{i>\theta}(\varepsilon)$ can be identically written as

$$\Lambda^i(\varepsilon) \simeq C_1 + C_2 \sum_{n=i}^{i+N-1} \sum_{p=0}^{Q-1} y_p(n) y_p^*(n)$$
$$+ C_3 \sum_{n=i}^{i+N-1} \mathrm{Re} \left\{ \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} y_p(n) y_q^*(n) e^{j2\pi(q-p)\varepsilon} \right\}. \tag{64}$$

Using the approach that is similar to that in Appendix A, we finally obtain

$$\Lambda^i(\varepsilon) \simeq C_1 + C_2 \phi^i + C_3 \mathrm{Re} \left\{ \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} |\gamma_{pq}^i| \cos\left(\psi_{pq}^i\right) \right\} \tag{65}$$

where $\gamma_{pq}^i = \sum_{n=i}^{i+N-1} y_p(n) y_q^*(n)$, $\phi^i = \sum_{p=0}^{Q-1} \gamma_{pp}^i$, and $\psi_{pq}^i = 2\pi\varepsilon(q-p) + \angle\gamma_{pq}^i$. Note that the approximations we made are equivalent to adding $|\theta - i|$ samples (noise or data)

in calculating the likelihood functions. Since the number of samples in the ith sliding data window QN is usually much larger than the number of added samples $|\theta - i|$, the added samples will not change the likelihood functions too much. The approximation errors also depend on the distance between the window position and the actual STO, i.e., $|\theta - i|$. When the distance is larger, the error is also larger. However, if the distance is larger, the likelihood function tends to be smaller, and a larger error is then tolerable. Finally, we note that the added samples, either noise or data, are uncorrelated with the preamble samples.

APPENDIX C
DERIVATIONS OF μ_ϕ^i, μ_ξ^i, ν_ϕ^i, ν_ξ^i, AND $\kappa_{\phi\xi}^i$

We first note that θ is the real STO in the system. Using θ as a reference, we have the following three cases for the value of i: 1) $i = \theta$; 2) $i < \theta$; and 3) $i > \theta$ ($0 \le i \le N - 1$). For the first case, the window covers the preamble data only (I_P). Thus, (42) and (43) can be simplified to

$$\phi^\theta = \sum_{p=0}^{Q-1} \sum_{n=\theta}^{\theta+N-1} x_p(n)x_p^*(n) + w_p(n)w_p^*(n)$$
$$+ 2\mathrm{Re}\left\{ x_p(n)w_p^*(n) \exp\left(j2\pi\varepsilon \frac{pN+n}{N} \right) \right\} \quad (66)$$

$$\xi^\theta = \sum_{p=0}^{Q-2} \sum_{q>p}^{Q-1} \sum_{n=\theta}^{\theta+N-1} x_p(n)x_q^*(n)$$
$$+ x_p(n)w_q^*(n) \exp\left(j2\pi\varepsilon \frac{qN+n}{N} \right)$$
$$+ w_p(n)x_q^*(n) \exp\left(-j2\pi\varepsilon \frac{pN+n}{N} \right)$$
$$+ w_p(n)w_q^*(n) \exp\left(j2\pi\varepsilon(q-p) \right). \quad (67)$$

The mean values of ϕ^i and ξ^i for the first case are then

$$\mu_{\phi,1}^\theta = QN \left(\sigma_x^2 + \sigma_w^2 \right) \quad (68)$$

$$\mu_{\xi,1}^\theta = \frac{QN(Q-1)}{2} \sigma_x^2. \quad (69)$$

The corresponding variance values are

$$\nu_{\phi,1}^\theta = 2QN \left(\sigma_x^2\sigma_w^2 \right) \quad (70)$$

$$\nu_{\xi,1}^\theta = \frac{QN(Q-1)}{2} \sigma_w^4$$
$$+ \sigma_x^2\sigma_w^2 \frac{QN(Q-1)(2Q-1)}{3}. \quad (71)$$

The corresponding covariance value is

$$\kappa_{\phi\xi,1}^\theta = QN(Q-1) \left(\sigma_x^2\sigma_w^2 \right). \quad (72)$$

Here, $\kappa_{\phi\xi,j}^i$ denotes $\kappa_{\phi\xi}^i$ in the jth case discussed. For the second case, the window covers the sets I_N and I_P. Thus, ϕ^i

and ξ^i can be expressed as

$$\phi^i = \sum_{n=i}^{\theta-1} w_0(n)w_0^*(n)$$
$$+ \sum_{n=\theta}^{i+N-1} x_0(n)x_0^*(n) + w_0(n)w_0^*(n)$$
$$+ 2\mathrm{Re}\left\{ x_0(n)w_0^*(n) \exp\left(j2\pi\varepsilon \frac{n}{N} \right) \right\}$$
$$+ \sum_{p=1}^{Q-1} \sum_{n=i}^{i+N-1} x_p(n)x_p^*(n) + w_p(n)w_p^*(n)$$
$$+ 2\mathrm{Re}\left\{ x_p(n)w_p^*(n) \exp\left(j2\pi\varepsilon \frac{pN+n}{N} \right) \right\} \quad (73)$$

$$\xi^i = \sum_{q>0}^{Q-1} \sum_{n=i}^{\theta-1} w_0(n)x_q^*(n) \exp\left(-j2\pi\varepsilon \frac{n}{N} \right)$$
$$+ w_0(n)w_q^*(n) \exp(j2\pi\varepsilon q)$$
$$+ \sum_{q>0}^{Q-1} \sum_{n=\theta}^{N+i-1} x_0(n)x_q^*(n)$$
$$+ x_0(n)w_q^*(n) \exp\left(j2\pi\varepsilon \frac{qN+n}{N} \right)$$
$$+ w_0(n)x_q^*(n) \exp\left(-j2\pi\varepsilon \frac{n}{N} \right)$$
$$+ w_0(n)w_q^*(n) \exp(j2\pi\varepsilon q)$$
$$+ \sum_{p=1}^{Q-2} \sum_{q>p}^{Q-1} \sum_{n=i}^{i+N-1} x_p(n)x_q^*(n)$$
$$+ x_p(n)w_q^*(n) \exp\left(j2\pi\varepsilon \frac{qN+n}{N} \right)$$
$$+ w_p(n)x_q^*(n) \exp\left(-j2\pi\varepsilon \frac{pN+n}{N} \right)$$
$$+ w_p(n)w_q^*(n) \exp\left(j2\pi\varepsilon(q-p) \right). \quad (74)$$

Their mean values are

$$\mu_{\phi,2}^i = (QN + i - \theta) \left(\sigma_x^2 + \sigma_w^2 \right) + (i - \theta)\sigma_x^2 \quad (75)$$

$$\mu_{\xi,2}^i = (Q-1)(N + i - \theta)\sigma_x^2$$
$$+ \frac{(Q-1)(Q-2)}{2} N\sigma_x^2 \quad (76)$$

the corresponding variance values are

$$\nu_{\phi,2}^i = 2\sigma_x^2\sigma_w^2[QN + i - \theta] \quad (77)$$

$$\nu_{\xi,2}^i = \frac{QN(Q-1)}{2} \sigma_w^4 + \sigma_x^2\sigma_w^2 N(Q-1)$$
$$\cdot \left[(Q-1)\left(2 + \frac{i-\theta}{N} \right) + \frac{(Q-2)(2Q-3)}{3} \right]$$
$$\quad (78)$$

and the covariance value is

$$\kappa_{\phi\xi,2}^i = (QN + i - \theta)(Q-1) \left(\sigma_x^2\sigma_w^2 \right). \quad (79)$$

For the third case, the window covers sets I_P and I_D, and we write ϕ^i and ξ^i as

$$
\begin{aligned}
\phi^i = & \sum_{p=0}^{Q-2} \sum_{n=i}^{i+N-1} \left\{ x_p(n)x_p^*(n) + w_p(n)w_p^*(n) + 2\mathrm{Re} \right. \\
& \left. \times \left[x_p(n)w_p^*(n) \exp\left(j2\pi\varepsilon\frac{pN+n}{N}\right) \right] \right\} \\
& + \sum_{n=i}^{\theta+N-1} \left\{ x_{Q-1}(n)x_{Q-1}^*(n) + w_{Q-1}(n)w_{Q-1}^*(n) \right. \\
& \quad + 2\mathrm{Re}\left[x_{Q-1}(n)w_{Q-1}^*(n) \right. \\
& \quad \left. \left. \times \exp\left(j2\pi\varepsilon\frac{(Q-1)N+n}{N}\right) \right] \right\} \\
& + \sum_{n=\theta+N}^{i+N-1} \left\{ x_{Q-1}(n)x_{Q-1}^*(n) + w_{Q-1}(n)w_{Q-1}^*(n) \right. \\
& \quad + 2\mathrm{Re}\left[x_{Q-1}(n)w_{Q-1}^*(n) \right. \\
& \quad \left. \left. \times \exp\left(j2\pi\varepsilon\frac{(Q-1)N+n}{N}\right) \right] \right\}
\end{aligned}
\tag{80}
$$

$$
\begin{aligned}
\xi^i = & \sum_{p=0}^{Q-3} \sum_{q>p}^{Q-2} \sum_{n=i}^{i+N-1} x_p(n)x_q^*(n) \\
& + x_p(n)w_q^*(n) \exp\left(j2\pi\varepsilon\frac{qN+n}{N}\right) \\
& + w_p(n)x_q^*(n) \exp\left(-j2\pi\varepsilon\frac{pN+n}{N}\right) \\
& + w_p(n)w_q^*(n) \exp\left(j2\pi\varepsilon(q-p)\right) \\
& + \sum_{p=0}^{Q-2} \sum_{n=i}^{\theta+N-1} x_p(n)x_{Q-1}^*(n) \\
& + x_p(n)w_{Q-1}^*(n) \exp\left(j2\pi\varepsilon\frac{(Q-1)N+n}{N}\right) \\
& + w_p(n)x_{Q-1}^*(n) \exp\left(-j2\pi\varepsilon\frac{pN+n}{N}\right) \\
& + w_p(n)w_{Q-1}^*(n) \exp\left(j2\pi\varepsilon(Q-p-1)\right) \\
& + \sum_{p=0}^{Q-2} \sum_{n=\theta+N}^{i+N-1} x_p(n)x_{Q-1}^*(n) \\
& + x_p(n)w_{Q-1}^*(n) \exp\left(j2\pi\varepsilon\frac{(Q-1)N+n}{N}\right) \\
& + w_p(n)x_{Q-1}^*(n) \exp\left(-j2\pi\varepsilon\frac{pN+n}{N}\right) \\
& + w_p(n)w_{Q-1}^*(n) \exp\left(j2\pi\varepsilon(Q-p-1)\right).
\end{aligned}
\tag{81}
$$

Thus, the mean values are

$$
\mu_{\phi,3}^i = (QN+\theta-i)\left(\sigma_x^2+\sigma_w^2\right) + (i-\theta)\left(\sigma_d^2+\sigma_w^2\right) \tag{82}
$$

$$
\mu_{\xi,3}^i = (Q-1)(N-i+\theta)\sigma_x^2 + \frac{(Q-1)(Q-2)}{2}N\sigma_x^2 \tag{83}
$$

the variance values are

$$
\nu_{\phi,3}^i = 2\sigma_x^2\sigma_w^2(QN-i+\theta) + 2(i-\theta)\sigma_d^2\sigma_w^2 \tag{84}
$$

$$
\begin{aligned}
\nu_{\xi,3}^i = & \sigma_x^2\sigma_w^2 N(Q-1) \\
& \times \left[(2Q-3)\left(1+\frac{\theta-i}{N}\right) + \frac{(Q-2)(2Q-3)}{3} + 1 \right] \\
& + \frac{QN(Q-1)}{2}\sigma_w^4 + (Q-1)^2(i-\theta)(\sigma_d^2+\sigma_w^2) \tag{85}
\end{aligned}
$$

and the covariance value is

$$
\kappa_{\phi\xi,3}^i = (Q-1)\sigma_x^2\sigma_w^2\left[QN+2(\theta-i)\right]. \tag{86}
$$

REFERENCES

[1] J. J. van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1800–1805, Jul. 1997.

[2] N. Lashkarian, "Class of cyclic-based estimators for frequency-offset estimation of OFDM systems," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2139–2149, Dec. 2000.

[3] Y.-C. Wu and E. Serpedin, "Comments on 'Class of cyclic-based estimators for frequency-offset estimation of OFDM systems,'" *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 413–414, Mar. 2005.

[4] M. Tanda, "Blind symbol-timing and frequency-offset estimation in OFDM systems with real data symbols," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1609–1612, Oct. 2004.

[5] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing.* New York: Wiley, 1998.

[6] 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.

[7] D. Lee and K. Cheun, "Coarse symbol synchronization algorithms for OFDM systems in multipath channels," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 446–448, Oct. 2002.

[8] S. Haykin, *Adaptive Filter Theory 4E*. Englewood Cliffs, NJ: Prentice–Hall, 2002, p. 440.

[9] M. Morelli, A. N. D'Andrea, and U. Mengali, "Frequency ambiguity resolution in OFDM systems," *IEEE Commun. Lett.*, vol. 4, no. 4, pp. 134–136, Apr. 2000.

[10] K.-T. Lee and J.-S. Seo, "Pilot-aided iterative frequency offset estimation for digital video broadcasting (DVB) systems," *IEEE Trans. Consum. Electron.*, vol. 53, no. 1, pp. 11–16, Feb. 2007.

[11] P. H. Moose, "A technique of orthogonal frequency division multiplexing frequency offset correlation," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct. 1994.

[12] 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.

[13] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

[14] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Commun. Lett.*, vol. 3, no. 3, pp. 75–77, Mar. 1999.

[15] H. Minn, P. Tarasak, and V. K. Bhargava, "OFDM frequency offset estimation methods based on BLUE principle," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, BC, Canada, Sep. 2002, pp. 1230–1234.

[16] J. Li, G. Liu, and G. B. Giannakis, "Carrier frequency offset estimation for OFDM-based WLANs," *IEEE Signal Process. Lett.*, vol. 8, no. 3, pp. 80–82, Mar. 2001.

[17] J. H. Yu and Y. T. Su, "Pilot-assisted maximum-likelihood frequency-offset estimation for OFDM systems," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1997–2008, Nov. 2004.

[18] M. H. Cheng and C. C. Chou, "Maximum-likelihood estimation of frequency and time offsets in OFDM systems with multiple sets of identical data," *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2848–2852, Jul. 2006.

[19] W. D. Warner and C. Leung, "OFDM/FM frame synchronization for mobile radio data communication," *IEEE Trans. Veh. Technol.*, vol. 42, no. 3, pp. 302–313, Aug. 1993.

[20] L. Hanzo, "Bandwidth-efficient wireless multimedia communications," *Proc. IEEE*, vol. 86, no. 7, pp. 1342–1382, Jul. 1998.

[21] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Boston, MA: Artech House, 2002.

[22] J. Heiskala and J. Terry, *OFDM Wireless LANs: A Theoretical and Practical Guide*. Indianapolis, IN: Sams, 2002.

[23] Y. S. Choi, P. J. Voltz, and F. A. Cassara, "ML estimation of carrier frequency offset for multicarrier signals in Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 644–655, Mar. 2001.

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

[25] H. V. Poor, *An Introduction to Signal Detection and Estimation 2E*. Berlin, Germany: Springer-Verlag, 1994.

[26] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer–Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[27] Y. Jian, H. Minn, X. Gao, X. You, and Y. Li, "Frequency offset estimation and training sequence design for MIMO OFDM," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1244–1254, Apr. 2008.

[28] S. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 832–835, Nov. 1985.

[29] W. S. McCormick and L. L. James, "Efficient parallel rooting of complex polynomials on the unit circle," *IEEE Trans. Signal Process.*, vol. 39, no. 10, pp. 2347–2351, Oct. 1991.

[30] L. Rugini and P. Banelli, "BER of OFDM systems imparied by carrier frequency offset in multipath fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2279–2288, Sep. 2008.
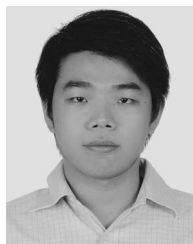
Hung-Tao Hsieh (S'08) received the B.S. degree in physics in 2000 from the National Central University, Taoyuan, Taiwan, and the M.S. degree in electro-physics in 2002 from the National Chiao Tung University (NCTU), Hsinchu, Taiwan, where he is currently working toward the Ph.D. degree with the Department of Communication Engineering.

His research interests include detection/estimation theories and communication signal processing.

Wen-Rong Wu (M'89) received the B.S. degree in mechanical engineering from Tatung Institute of Technology, Taipei, Taiwan, in 1980 and the M.S. degrees in mechanical and electrical engineering and the Ph.D. degree in electrical engineering from the State University of New York, Buffalo, in 1985, 1986, and 1989, respectively.

Since August 1989, he has been a faculty member with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan. His research interests include statistical signal processing and digital communication.