

Blind Estimation of OFDM Carrier Frequency Offset via Oversampling

Biao Chen, *Member, IEEE*, and Hao Wang, *Student Member, IEEE*

Abstract—Blind deterministic estimation of the orthogonal frequency division multiplexing (OFDM) frequency offset via oversampling is proposed in this paper. This method utilizes the intrinsic phase shift of neighboring sample points incurred by the frequency offset that is *common* among all subcarriers. The proposed method is data efficient—it requires only a single OFDM symbol to achieve reliable estimation, hence making it more suitable to systems with stringent delay requirement and mobility-induced channel variation. The proposed scheme is devised to perfectly retrieve frequency offset in the absence of noise. Quite remarkably, we show that in the presence of channel noise, this intuitive scheme is indeed the maximum likelihood estimate of the carrier frequency offset. The possible presence of virtual carriers are also accommodated in the system model, and some interesting observations are obtained. The Cramér–Rao lower bound is derived for the oversampling-based signal model, and we show through numerical simulation that the proposed algorithm is efficient. Practical issues such as identifiability, the front-end filter bandwidth, and the possible presence of correlated noises are also carefully addressed.

Index Terms—Carrier frequency offset, maximum likelihood estimate, orthogonal frequency division multiplexing (OFDM), oversampling.

I. INTRODUCTION

CURRENT and future broadband wireless communication systems aim to provide high data rate services. As a result, multipath fading becomes a major concern as systems with high data rate are more liable to intersymbol interference (ISI) due to the increase in the normalized delay spread. It is, therefore, imperative to use modulation schemes that are robust to multipath fading. Orthogonal frequency division multiplexing (OFDM), because of the split of a high rate data stream into a parallel of lower rate data streams, provides the much needed resistance to multipath fading and has attracted increasing interest in recent years. In Europe, OFDM was first standardized for digital audio broadcasting (DAB) in 1995 [1] and terrestrial digital video broadcasting in 1997 [2]. It has also been proposed for high data rate packet transmission, *a.k.a.*, wireless LAN, as specified in IEEE 802.11a [3] and HIPERLAN/2 [4].

Although OFDM is inherently immune to frequency-selective fading, it is more sensitive to frequency offset compared

with single carrier systems. The presence of the carrier frequency offset (CFO) will introduce severe inter-carrier interference (ICI), which, if not properly compensated, would significantly degrade the system performance [5]. In order to mitigate this effect, various techniques have been proposed to estimate the CFO for OFDM systems [6]–[15]. In [6], Moose proposed a maximum likelihood (ML) estimator using repeated data symbol. Although the symbol values need not be known, repetition of OFDM symbols is in essence tantamount to a training-based scheme as it utilizes extra bandwidth. Data-assisted frequency acquisition and tracking were proposed in [7], where periodically inserted known symbols were explicitly used. In [8], Schmidl and Cox proposed a training symbol-based timing/frequency synchronization that utilized an OFDM symbol with identical halves. This was later generalized to a training symbol with multiple identical parts [9]. Various blind techniques have also been proposed. In [11], van de Beek *et al.* developed an ML estimator by exploiting the redundancy in the cyclic prefix (CP). This method, however, is developed based on the nondispersive channel model and suffers the error floor effect in the presence of frequency-selective fading channels. Schmidl and Cox proposed in [12] a blind estimation method that is only suitable to recover CFO values that are multiples of the carrier spacing. In [13], Choi *et al.* developed an ML estimator by assuming that the OFDM signal is complex Gaussian distributed, which is asymptotically true for circularly modulated OFDM symbols [16], [17]. The proposed algorithm, however, requires perfect knowledge of the second-order statistics about the channel noise and the Rayleigh fading channel. In [14] and [15], Liu and Tureli took advantage of the presence of virtual carriers in OFDM signaling and proposed blind estimation methods reminiscent of spectral analysis techniques in array processing, *i.e.*, MUSIC and ESPRIT. It was later shown that the proposed MUSIC algorithm is indeed the ML estimate of the CFO with a virtual carrier present signal model [18]–[20]. The drawback is that it usually requires multiple OFDM symbols to achieve desirable performance, thus introducing an extra delay at the receiver. Receiver diversity-based CFO estimation was given in [21], where multiple receive antennas were employed. Again, the presence of virtual carriers was exploited and the proposed estimator is indeed an extension of the MUSIC algorithm much in the same way as its extension to the case when multiple OFDM blocks are used with a single antenna.

In this paper, we exploit the intrinsic phase shift between neighboring samples due to the CFO. As a motivation, consider the two time instances t_1 and t_2 , as in Fig. 1, where only two subcarriers are plotted. The phase shift between t_1 and t_2 comes

Manuscript received July 1, 2001; revised August 28, 2003. This work was supported in part by the CASE center of Syracuse University. Part of the research results in this manuscript was presented at the 35th Asilomar Conference on Signals, Systems, and Computers, Monterey, CA, November 2001. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Joseph Tabrikian.

The authors are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: bichen@ecs.syr.edu; hwang08@ecs.syr.edu).

Digital Object Identifier 10.1109/TSP.2004.828899

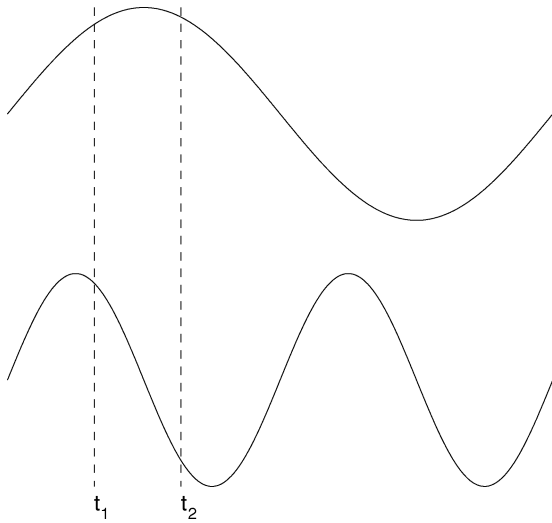


Fig. 1. Phase shift between t_1 and t_2 are attributed to two sources. One is the carrier frequencies that are different from carrier to carrier. The other is due to the frequency offset, which is common among all subcarriers.

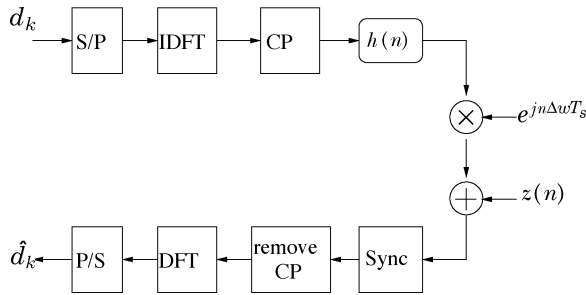


Fig. 2. Discrete-time baseband model of OFDM systems with carrier frequency offset. In the block diagram, S/P and P/S are serial-to-parallel and parallel-to-serial converter, respectively.

from two sources. The first is attributed to the carriers themselves. Notice that because of the frequency difference among subcarriers, different subcarriers will contribute different phase shifts. The other source is the CFO that causes a phase shift that is common across subcarriers as they all experience the same frequency offset (see Fig. 2). The proposed method intends to utilize the common phase shift due to the CFO to achieve high-resolution frequency offset estimation. One of the major advantages of the new method is that not only is it blind, hence bandwidth/power efficient, but it is also data efficient—only one OFDM symbol is needed for reliable estimation. Thus, it is more suitable for situations with strict delay requirement as in DAB/DVB applications. Further, its CFO acquisition range is the entire $[0, 2\pi)$ and, hence, is larger than the carrier spacing. The proposed method is applicable to both the frequency-nonsselective and frequency-selective fading channels and does not require the presence of virtual carriers. The extra cost of the proposed scheme comes with the increase of the receiver complexity due to the oversampling of the received signal by a factor of two. The benefit of the new method, however, significantly outweighs its cost: Not only does it relax most of the restrictions imposed in the blind frequency offset estimation schemes mentioned above, but its performance is superior than those previous methods for a wide range of SNR, as verified through extensive numerical simulations.

The algorithm is first developed as an intuitive scheme to perfectly retrieve the CFO in the absence of noise. We then derive rigorously the ML estimate of CFO using the oversampled signal model. Quite remarkably, we show that the former is just a special case of the latter. The organization of the paper is as follows. In the next section, we briefly introduce the oversampled OFDM system model with an oversampling factor of two. In Section III, a deterministic CFO estimation method is proposed using the oversampled signal. Simplification using the root-finding algorithm is also presented for fast computation. We derive in Section IV the oversampling-based ML estimator. We show that the deterministic CFO estimation proposed in Section III is indeed an ML estimate in the absence of virtual carriers. The Cramér–Rao lower bound (CRLB) is derived for CFO estimation using oversampled signal model. Numerical examples are given in Section V to demonstrate the performance advantage of the proposed method. Discussions are also given regarding the effect of possible noise correlation due to the front-end filter, and we conclude in Section VI.

The notations used in this paper are mostly standard and are summarized below, along with some of the OFDM system parameters.

$(\cdot)^H$	Hermitian (conjugate transpose);
$(\cdot)^T$	transpose;
$(\cdot)^*$	complex conjugate;
$\mathbf{x}, \mathbf{s}$	boldface lowercase letters represent vectors;
$\mathbf{A}, \mathbf{X}$	boldface upper case letters represent matrices;
$\text{diag}(\mathbf{x})$	diagonal matrix with $\mathbf{x}$ at its diagonal;
$\mathbf{I}_{N \times N}$	$N \times N$ identity matrix;
N	number of subcarriers;
M	number of subcarriers that carry information;
$N - M$	number of virtual carriers;
T	IDFT/DFT interval (OFDM block duration excluding the cyclic prefix);
$T_s = T/N$	symbol interval;
$\omega = 2\pi/N$	carrier spacing;
$\Delta\omega$	frequency offset;
$\phi = \Delta\omega T_s$	normalized frequency offset (phase shift per symbol interval due to $\Delta\omega$).

We also define the following matrix:

$$\mathbf{W} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j\omega} & \cdots & e^{j(M-1)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j(N-1)\omega} & \cdots & e^{j(M-1)(N-1)\omega} \end{bmatrix}. \quad (1)$$

Indeed, $\mathbf{W}$ is the first M columns of an unitary $N \times N$ IDFT matrix, and it is easy to verify that $\mathbf{W}^H \mathbf{W} = \mathbf{I}_{M \times M}$, whereas $\mathbf{W} \mathbf{W}^H \neq \mathbf{I}_{N \times N}$ if $M < N$.

II. OVERSAMPLED OFDM SIGNAL MODEL

In an OFDM system, the available bandwidth is partitioned into N subchannels. By doing so, the symbol duration is N times large as that of a single carrier system with the same symbol rate. This significantly reduces the normalized delay spread and makes it more robust to time dispersive channels.

Unlike conventional frequency division multiplexing, the N subchannels in OFDM are designed to be orthogonal to each other. This allows frequency overlapping and eliminates the need for a guard band, hence achieving high bandwidth efficiency.

Specifically, for OFDM system with N subcarriers, N information symbols are used to construct one OFDM block. Each of the N symbols is used to modulate a subcarrier, and the N modulated subcarriers are added together to form an OFDM symbol block. Orthogonality among subcarriers is achieved by carefully selecting the carrier frequencies such that each OFDM symbol interval contains an integer number of periods for all subcarriers. Using the discrete-time baseband signal model, OFDM systems use the inverse discrete-time Fourier transform—discrete-time Fourier transform (IDFT-DFT) pair in the transmitter-receiver to achieve orthogonality among subcarriers, as illustrated in Fig. 2. To totally eliminate the interblock interference, cyclically extended guard time is inserted that is assumed to be longer than the maximum delay spread. The insertion of CP maintains the orthogonality among subcarriers in the presence of timing offset and multipath channel [22]. If virtual carriers are used (analogous to zero padding in IDFT), only M out of N carriers are used to modulate information symbols. Without loss of generality, we assume that the first M carriers are used as information carrying carriers. Equivalently, the information symbols corresponding to the $N - M$ virtual carriers can be assumed to be zero. We note here that the position of virtual carriers does affect the estimation performance [18].

The above discrete-time baseband model, as in Fig. 2, is based on symbol rate sampling. To obtain the oversampled OFDM signal model, we first write the continuous time complex baseband OFDM symbol at the transmitter

$$s(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{M-1} d_k e^{j \frac{2\pi k}{T} (t-t_s)} \quad (2)$$

where each d_k is used to modulate the subcarrier $e^{j2\pi k/T}$, t_s is the starting time of the current OFDM symbol (excluding CP time), T is the OFDM block duration (DFT/IDFT interval), N is the total number of subcarriers, and M is the number of active subcarriers. The scalar factor $1/\sqrt{N}$ is introduced for the ease of notation in the following derivation. Notice that the above continuous-time complex baseband OFDM signal model is valid when a rectangular pulse is applied to the entire IDFT output block [23]. This, indeed, is the way current OFDM systems are implemented [3], [24], [25].

In the presence of a time-dispersive channel and additive noise, the received continuous time baseband signal is now

$$x(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{M-1} d_k H(k) e^{j \frac{2\pi k}{T} (t-t_s)} + z(t)$$

where $H(k)$ is the channel frequency response corresponding to subcarrier k , and $z(t)$ is complex Gaussian noise. Strictly speaking, this received signal model is true only if the complex exponential signal in (2) has infinite duration, thus forming an eigen signal for any linear filter. For OFDM systems, the duration of this complex exponential lasts the entire OFDM block.

Thus, with respect to the channel bandwidth, this complex exponential still acts as an impulse in the frequency domain, and the effect of bandlimited channel can be largely neglected.

We define $T_s = T/N$ as the symbol interval. Assuming that the sampling instances are at $t = t_s + nT_s + \tau$, where τ is the initial timing shift, and $n = 0, \dots, N-1$, we have two sets of observations, depending on the choice of τ .

$$\begin{aligned} x_1(n) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{M-1} H(k) d_k e^{j \left(\frac{2\pi k}{N} + \Delta\omega \cdot T_s \right) n} + z_1(n), \quad \tau = 0 \\ x_2(n) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{M-1} H(k) d_k e^{j \left(\frac{2\pi k}{N} + \Delta\omega \cdot T_s \right) \left(n + \frac{1}{2} \right)} + z_2(n), \quad \tau = \frac{T_s}{2} \end{aligned}$$

where $z_1(n)$ and $z_2(n)$ are additive complex Gaussian noise. Here, the initial phase due to frequency offset is assumed to be zero. The above two sets of observations are obtained through oversampling $x(t)$ by a factor of two. Notice that if we define $\phi = \Delta\omega \cdot T_s$, then ϕ and the frequency offset $\Delta\omega$ differ only by a constant scalar, and hence, CFO estimation is equivalent to the estimation of the normalized phase shift ϕ .

The above signal model can be written in a more compact matrix form as follows:

$$\mathbf{x}_1 = \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} + \mathbf{z}_1 \quad (3)$$

$$\mathbf{x}_2 = e^{j \frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} + \mathbf{z}_2 \quad (4)$$

where $\mathbf{W}$ consists of the first M columns of the IDFT matrix as defined in (1); $\tilde{\mathbf{d}}$ is the symbol vector $[\tilde{d}_1, \dots, \tilde{d}_{M-1}]^T$ with $\tilde{d}_k = H(k) d_k$; and

$$\mathbf{E} = \text{diag} \left(1 e^{j \frac{\pi}{N}} \dots e^{j \frac{(M-1)\pi}{N}} \right) \quad (5)$$

$$\mathbf{P} = \text{diag}(1 e^{j\phi} \dots e^{j(N-1)\phi}). \quad (6)$$

Notice that matrix $\mathbf{E}$ reflects the phase shift between $\mathbf{x}_1$ and $\mathbf{x}_2$ due to the difference in subcarrier frequencies, whereas the phase shift between $\mathbf{x}_1$ and $\mathbf{x}_2$ due to the CFO is a *scalar constant* $e^{j\phi/2}$.

The above signal model lays out the groundwork for the ensuing frequency offset estimation techniques described in the next section. This oversampling-based scheme can be justified using an intuitive argument. Assuming that N carriers are all used to carry information symbols, the unknown vector $\tilde{\mathbf{d}}$ is of length N . Together with the frequency offset, there are a total of $N+1$ unknown parameters. Symbol rate sampling will only produce N independent samples; hence, it is not sufficient to solve for the $N+1$ unknowns. This motivates the oversampling idea for frequency offset estimation. Notice that from this argument, we see that in the presence of the virtual carrier, where $\tilde{\mathbf{d}}$ is then of size $M \times 1$, symbol rate sampling can indeed provide enough information for blind frequency offset estimation using one OFDM symbol block, as is used in the MUSIC-like algorithm developed by Liu and Tureli [14].

III. BLIND FREQUENCY OFFSET ESTIMATION

Based on (3) and (4), we propose, in this section, a blind estimation algorithm to recover the carrier frequency offset. We

develop the algorithm using a heuristic yet very sensible argument—we seek to perfectly retrieve the CFO in the absence of noise. Remarkably, we show in Section IV that the resulting estimate is, in fact, the ML estimate of the frequency offset under the oversampling framework in the presence of white Gaussian noise.

A. Perfect Recovery of CFO in the Absence of Noise

From (3) and (4), we notice that although $\tilde{\mathbf{d}}$ is unknown, the fact that they are common to both vectors may facilitate the blind estimation of the frequency offset ϕ . Define

$$\mathbf{y}_1 = \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 \quad (7)$$

$$\mathbf{y}_2 = e^{-j\frac{\phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2. \quad (8)$$

It is straightforward to verify that *in the absence of noise*

$$\mathbf{y}_1 = \mathbf{y}_2 = \tilde{\mathbf{d}}. \quad (9)$$

Thus, it is intuitive to find ϕ such that the resulting $\mathbf{y}_1$ and $\mathbf{y}_2$ are equal. Indeed, we will show that in the noiseless case, the true frequency offset is the unique value that yields $\mathbf{y}_1 = \mathbf{y}_2$. To establish this result, we first give the following lemma.

Lemma 1: If there is $\phi_o \in [0, 2\pi)$ and $\phi \in [0, 2\pi)$ such that

$$\mathbf{W}^H \mathbf{P}(\phi)^H \mathbf{P}(\phi_o) \mathbf{W} = e^{j\frac{\phi_o - \phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}(\phi)^H \mathbf{P}(\phi_o) \mathbf{W} \mathbf{E}$$

where $\mathbf{P}(\phi)$ and $\mathbf{P}(\phi_o)$ are diagonal matrices defined as in (6) with parameters ϕ and ϕ_o , respectively, and $\mathbf{E}$ is defined in (5), then we must have $\phi = \phi_o$.

The proof is given in Appendix A. This immediately leads to the following theorem.

Theorem 1: In the absence of noise, there is a unique value $\phi \in [0, 2\pi)$, corresponding to the true frequency offset, such that $\mathbf{y}_1 = \mathbf{y}_2$, where $\mathbf{y}_1$ and $\mathbf{y}_2$ are constructed according to (7) and (8).

Proof: Denote by ϕ_o the true frequency offset. Assume that there is a ϕ such that $\mathbf{y}_1(\phi) = \mathbf{y}_2(\phi)$, where $\mathbf{y}_1(\phi)$ and $\mathbf{y}_2(\phi)$ are constructed according to (7) and (8) using ϕ . Thus, we have, in the absence of noise

$$\begin{aligned} \mathbf{y}_1(\phi) &= \mathbf{W}^H \mathbf{P}^H(\phi) \mathbf{x}_1 = \mathbf{W}^H \mathbf{P}^H(\phi) \mathbf{P}(\phi_o) \mathbf{W} \tilde{\mathbf{d}} \\ \mathbf{y}_2(\phi) &= e^{-j\frac{\phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2 \\ &= e^{j\frac{\phi_o - \phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H(\phi) \mathbf{P}(\phi_o) \mathbf{W} \tilde{\mathbf{d}} \end{aligned}$$

where $\mathbf{P}(\phi)$ and $\mathbf{P}(\phi_o)$ are the diagonal matrices as in Lemma 1. Since $\tilde{\mathbf{d}}$ can be arbitrary, the above equality implies

$$\mathbf{W}^H \mathbf{P}(\phi)^H \mathbf{P}(\phi_o) \mathbf{W} = e^{j\frac{\phi_o - \phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}(\phi)^H \mathbf{P}(\phi_o) \mathbf{W} \mathbf{E}.$$

From Lemma 1, we must have $\phi = \phi_o$. Q.E.D.

Theorem 1 describes an easy way to perfectly retrieve the frequency offset in the absence of noise. That is, we search for ϕ in $[0, 2\pi)$ such that $\mathbf{y}_1$ and $\mathbf{y}_2$ constructed using (7) and (8) are identical. Further, we notice here that the CFO range can vary in the entire $[0, 2\pi)$, instead of a fraction of carrier spacing.

B. CFO Estimation in the Presence of Noise

In the presence of noise, it is unlikely that (9) will be true for any value of ϕ . However, in light of Theorem 1, we propose, intuitively, to minimize the distance between $\mathbf{y}_1$ and $\mathbf{y}_2$.

$$\hat{\phi} = \arg \min_{\phi} S_1(\phi) = (\mathbf{y}_1 - \mathbf{y}_2)^H (\mathbf{y}_1 - \mathbf{y}_2). \quad (10)$$

Justification of the above minimization criterion using the ML principle is provided in the next section. In the following, we present a simple way to solve the above minimization problem by means of root finding that is more computationally efficient than the brute force searching of ϕ from 0 to 2π .

Given (7) and (8), it is straightforward to show that $\mathbf{y}_1^H \mathbf{y}_1 = \mathbf{x}_1^H \mathbf{x}_1$ and $\mathbf{y}_2^H \mathbf{y}_2 = \mathbf{x}_2^H \mathbf{x}_2$. That is, $\mathbf{y}_1^H \mathbf{y}_1$ and $\mathbf{y}_2^H \mathbf{y}_2$ are independent of ϕ . Hence, the minimization is now equivalent to the following maximization problem

$$\max_{\phi} Q(\phi) = \mathbf{y}_1^H \mathbf{y}_2 + \mathbf{y}_2^H \mathbf{y}_1. \quad (11)$$

Since $\mathbf{y}_2^H \mathbf{y}_1 = (\mathbf{y}_1^H \mathbf{y}_2)^*$, $Q(\phi) = 2\text{real}(\mathbf{y}_2^H \mathbf{y}_1)$, where $\text{real}(\cdot)$ is the real part of its argument. Therefore, maximizing $Q(\phi)$ is equivalent to maximizing the real part of $\mathbf{y}_2^H \mathbf{y}_1$, which is the inner product of the constructed vectors from oversampled data. From (7) and (8), we have

$$\mathbf{y}_2^H \mathbf{y}_1 = e^{j\frac{\phi}{2}} \mathbf{x}_2^H \mathbf{P}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1.$$

By the matrix identity

$$\mathbf{X} \mathbf{y} = \mathbf{Y} \mathbf{x}$$

where $\mathbf{X} = \text{diag}(\mathbf{x})$ and $\mathbf{Y} = \text{diag}(\mathbf{y})$ with $\mathbf{x}$ and $\mathbf{y}$ are any arbitrary vectors of same size, we can write

$$\mathbf{y}_2^H \mathbf{y}_1 = e^{j\frac{\phi}{2}} \mathbf{x}_2^H \mathbf{P}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 = e^{j\frac{\phi}{2}} \mathbf{p}^H \mathbf{X}_2^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 \mathbf{p}$$

where $\mathbf{p} = [1 \ e^{-j\phi} \ \dots \ e^{-j(N-1)\phi}]^T$, $\mathbf{X}_1 = \text{diag}(\mathbf{x}_1)$, and $\mathbf{X}_2 = \text{diag}(\mathbf{x}_2)$. Thus, we have the following proposition.

Proposition 1: The minimization problem as in (10) is equivalent to maximizing

$$\text{real}(\mathbf{y}_2^H \mathbf{y}_1) = \text{real}\left(e^{j\frac{\phi}{2}} \mathbf{p}^H \mathbf{G} \mathbf{p}\right) \quad (12)$$

where $\mathbf{G} \triangleq \mathbf{X}_2^H \mathbf{W}^H \mathbf{P}^H \mathbf{X}_1$.

Thus, minimizing (10) is simplified to maximizing (12), i.e., the real part of a quadratic form. Therefore, no explicit construction of $\mathbf{y}_1$ and $\mathbf{y}_2$ is required. However, since $\mathbf{y}_2^H \mathbf{y}_1$ is still complex in general, maximizing its real part still involves an exhaustive search from 0 to 2π . To further simplify this, we note that $Q(\phi)$, as defined in (11), is always a real number. This allows us to achieve maximization by setting its derivative equal to zero. To do this, from Proposition 1, we can rewrite

$$Q(\phi) = e^{j\frac{\phi}{2}} \mathbf{p}^H \mathbf{G} \mathbf{p} + e^{-j\frac{\phi}{2}} \mathbf{p}^H \mathbf{G}^H \mathbf{p}.$$

Notice further that

$$\mathbf{p}^H \mathbf{G} \mathbf{p} = \sum_{m=1}^N \sum_{n=1}^N g_{m,n} p^*(m) p(n) = \sum_{m=1}^N \sum_{n=1}^N g_{m,n} e^{j(m-n)\phi}$$

where $g_{m,n}$ is the m -row n -column element of $\mathbf{G}$, and $p(m)$ is the m^{th} element of the vector $\mathbf{p}$. We can then rewrite the above quadratic form as

$$\mathbf{p}^H \mathbf{G} \mathbf{p} = \sum_{k=-N+1}^{N-1} G_k e^{jk\phi}$$

where

$$G_k = \begin{cases} \sum_{l=1}^{N-k} g_{l+k,l}, & k \geq 0 \\ \sum_{l=1}^{N-k} g_{l,l+k}, & k < 0. \end{cases} \quad (13)$$

Similarly, we can verify

$$\mathbf{p}^H \mathbf{G}^H \mathbf{p} = \sum_{k=-N+1}^{N-1} A_{-k}^* e^{jk\phi}.$$

Thus

$$Q(\phi) = e^{j\frac{\phi}{2}} \sum_{k=-N+1}^{N-1} G_k e^{jk\phi} + e^{-j\frac{\phi}{2}} \sum_{k=-N+1}^{N-1} G_{-k}^* e^{jk\phi}.$$

Taking the derivative with respect to ϕ and setting it equal to zero, we have

$$\begin{aligned} \frac{d}{d\phi} Q(\phi) &= e^{j\frac{\phi}{2}} \sum_{k=-N+1}^{N-1} j \left(\frac{2k+1}{2} \right) G_k e^{jk\phi} \\ &\quad + e^{-j\frac{\phi}{2}} \sum_{k=-N+1}^{N-1} j \left(\frac{2k-1}{2} \right) G_{-k}^* e^{jk\phi} = 0. \end{aligned}$$

Multiplying the above equation by $e^{j(N-1/2)\phi}$ and defining $t = e^{j\phi}$, we get the following equation:

$$g(t) = \sum_{k=0}^{2N-1} j \left(k - N + \frac{1}{2} \right) (G_{-k+N-1}^* + G_{k-N}) t^k = 0 \quad (14)$$

with $G_{-N} = 0$. Thus, finding ϕ that maximizes $Q(\phi)$ is equivalent to finding the root on the unit circle for $g(t)$ defined in (14). The procedure is summarized in the following.

- 1) Calculate $\mathbf{G} = \mathbf{X}_2^H \mathbf{W} \mathbf{E} \mathbf{W}^H \mathbf{X}_1$. Note that $\mathbf{W} \mathbf{E} \mathbf{W}^H$ can be precalculated, and $\mathbf{X}_2$ and $\mathbf{X}_1$ are both diagonal matrices; hence, we have the computational complexity is only $2N^2$.
- 2) Compute G_k according to (13).
- 3) Solve (14), and find the root on the unit circle whose angle corresponds to the frequency offset estimate.

In the case of multiple roots on the unit circle, we can simply compute the corresponding cost function $Q(\phi)$ and choose the one with the minimum cost.

IV. ML ESTIMATE OF THE CFO USING OVERSAMPLED SIGNAL

The above blind frequency offset estimation is motivated by trying to perfectly recover ϕ in the absence of noise. It turns out, quite remarkably, that this intuitive scheme coincides with the ML estimation of the CFO using the noisy data when no virtual carriers are used. We start with the general oversampled signal model in the presence of virtual carriers and derive the ML estimate assuming uncorrelated noise vectors.

Assuming that $\mathbf{z}_1$ and $\mathbf{z}_2$ in (3) and (4) are both complex Gaussian with covariance matrix $\sigma^2 \mathbf{I}$ and uncorrelated with each other,¹ we can write the likelihood function as

$$L(\phi, \tilde{\mathbf{d}}) = \frac{1}{(\pi\sigma^2)^N} \exp \left\{ -\frac{1}{\sigma^2} \left[(\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}})^H (\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}}) + \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} \right)^H \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} \right) \right] \right\}. \quad (15)$$

The ML estimate for ϕ and $\tilde{\mathbf{d}}$ is

$$(\phi_{ML}, \tilde{\mathbf{d}}_{ML}) = \arg \max_{\phi, \tilde{\mathbf{d}}} L(\phi, \tilde{\mathbf{d}})$$

or equivalently, we are to minimize

$$S(\phi, \tilde{\mathbf{d}}) = (\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}})^H (\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}}) + \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} \right)^H \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} \right).$$

Taking the gradient of $S(\phi, \tilde{\mathbf{d}})$ with respect to $\tilde{\mathbf{d}}$ and setting it to zero, we get

$$\begin{aligned} \nabla_{\tilde{\mathbf{d}}} S(\phi, \tilde{\mathbf{d}}_{ML}(\phi)) &= -2 \mathbf{W}^H \mathbf{P}^H (\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}}) \\ &\quad - 2 e^{-j\frac{\phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{d}} \right) = 0 \end{aligned}$$

from which we can solve for $\tilde{\mathbf{d}}_{ML}$

$$\tilde{\mathbf{d}}_{ML} = \frac{1}{2} \left(\mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 + e^{-j\frac{\phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2 \right).$$

Notice that this is true for any ϕ , thus allowing us to plug it back into $S(\phi, \tilde{\mathbf{d}})$ to solve for ϕ . We have

$$\begin{aligned} S(\phi) &= \left(\mathbf{x}_1 - \frac{1}{2} \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 - \frac{1}{2} e^{-j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2 \right)^H \\ &\quad \times \left(\mathbf{x}_1 - \frac{1}{2} \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 - \frac{1}{2} e^{-j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \mathbf{E}^H \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2 \right) \\ &\quad + \left(\mathbf{x}_2 - \frac{1}{2} \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2 - \frac{1}{2} e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \mathbf{E} \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 \right)^H \\ &\quad \times \left(\mathbf{x}_2 - \frac{1}{2} \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H \mathbf{x}_2 - \frac{1}{2} e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \mathbf{E} \mathbf{W}^H \mathbf{P}^H \mathbf{x}_1 \right) \\ &= \left((\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_1 + \frac{1}{2} \mathbf{P} \mathbf{W} (\mathbf{y}_1 - \mathbf{y}_2) \right)^H \\ &\quad \times \left((\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_1 + \frac{1}{2} \mathbf{P} \mathbf{W} (\mathbf{y}_1 - \mathbf{y}_2) \right) \\ &\quad + \left((\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_2 + \frac{1}{2} e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} (\mathbf{y}_2 - \mathbf{y}_1) \right)^H \\ &\quad \times \left((\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_2 + \frac{1}{2} e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} (\mathbf{y}_2 - \mathbf{y}_1) \right) \\ &= \mathbf{x}_1^H (\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_1 + \frac{1}{4} (\mathbf{y}_1 - \mathbf{y}_2)^H (\mathbf{y}_1 - \mathbf{y}_2) \\ &\quad + \mathbf{x}_2^H (\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_2 + \frac{1}{4} (\mathbf{y}_2 - \mathbf{y}_1)^H (\mathbf{y}_2 - \mathbf{y}_1) \\ &= \mathbf{x}_1^H (\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_1 + \mathbf{x}_2^H (\mathbf{I} - \mathbf{P} \mathbf{W} \mathbf{W}^H \mathbf{P}^H) \mathbf{x}_2 \\ &\quad + \frac{1}{2} (\mathbf{y}_1 - \mathbf{y}_2)^H (\mathbf{y}_1 - \mathbf{y}_2) \end{aligned} \quad (16)$$

¹This assumption relies on the increased bandwidth of the front-end filter compared with critically sampled systems. This and the effect of correlated noise resulting from Nyquist front-end filter will be discussed later.

where we have used the fact that

$$(\mathbf{I} - \mathbf{P}\mathbf{W}\mathbf{W}^H\mathbf{P}^H)\mathbf{P}\mathbf{W} = \mathbf{0}$$

which is a direct consequence of the fact that $\mathbf{P}\mathbf{W}\mathbf{W}^H\mathbf{P}^H$ is a projection matrix onto the space spanned by $\mathbf{P}\mathbf{W}$. Hence, we have the following theorem.

Theorem 2: Assume that the noise vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ in (3) and (4) are both complex white Gaussian with covariance matrix $\sigma^2\mathbf{I}$. Further, $\mathbf{z}_1$ and $\mathbf{z}_2$ are uncorrelated with each other. Then, the ML estimation for ϕ is equivalent to finding ϕ that minimizes $S(\phi)$, as in (16).

In the special case when no virtual carriers are present, $M = N$, and $\mathbf{W}$ is the IDFT matrix itself, we have $\mathbf{W}\mathbf{W}^H = \mathbf{I}_{N \times N}$. Thus, the first two terms in (16) are both zero, and we have the following proposition.

Proposition 2: Assume that the noise vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ in (3) and (4) are both complex white Gaussian with covariance matrix $\sigma^2\mathbf{I}$. Further, $\mathbf{z}_1$ and $\mathbf{z}_2$ are uncorrelated with each other. In addition, we assume that $M = N$, i.e., there is no virtual carrier involved. Then, the ML estimation for ϕ is equivalent to finding ϕ that minimizes the distance between $\mathbf{y}_1$ and $\mathbf{y}_2$, as in (10).

The form of (16) also deserves further discussion. In [26], it was shown that the MUSIC algorithm in [14] is essentially the ML estimation of the CFO with the virtual carrier present OFDM model and is equivalent to minimizing

$$S_2(\phi) = \mathbf{x}^H(\mathbf{I} - \mathbf{P}\mathbf{W}\mathbf{W}^H\mathbf{P}^H)\mathbf{x}. \quad (17)$$

Clearly, S in (16) is a weighted sum of S_1 in (10) and S_2 in (17). This, of course, is not surprising given that (16) is derived using the oversampled model in the presence of virtual carriers. Extension to oversampling by a factor greater than two can also be obtained in a similar fashion. The corresponding ML estimation is equivalent to minimizing, for oversampling by a factor of L

$$S(\phi) = \sum_{i=1}^L \left[\mathbf{x}_i^H(\mathbf{I} - \mathbf{P}\mathbf{W}\mathbf{W}^H\mathbf{P}^H)\mathbf{x}_i + \left(\mathbf{y}_i - \frac{1}{L} \sum_{j=1}^L \mathbf{y}_j \right)^H \left(\mathbf{y}_i - \frac{1}{L} \sum_{j=1}^L \mathbf{y}_j \right) \right]$$

where $\mathbf{y}_i$ is constructed in a similar fashion from $\mathbf{x}_i$. The proof is fairly lengthy, although it is straightforward, and we skip the details. It is easy to show that this reduces to (16) for $L = 2$.

A. CRLBs for CFO Estimation

In [19] and [26], the CRLB for CFO estimation using the OFDM symbol in the presence of virtual carriers was derived. The obtained CRLB also sheds light on the identifiability issue using the virtual carrier-based OFDM signal [19]. In this section, we derive the CRLB for CFO estimation using the oversampled OFDM model. The detailed derivation is in Appendix B, and we give the result below.

Theorem 3: The CRLB for the oversampled signal model is

$$\text{CRLB}(\phi) = \frac{\frac{\sigma^2}{2}}{C - D} \quad (18)$$

where

$$\begin{aligned} \mathbf{A} &= \text{diag}(0, 1, 2, \dots, N-1) \\ \mathbf{B} &= \text{diag}\left(\frac{1}{2}, \frac{3}{2}, \dots, \left(\frac{N-1}{2}\right)\right) \\ C &= |\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 + |\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2 \\ D &= \frac{1}{2}|\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}} + \mathbf{E}^H\mathbf{W}^H\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2. \end{aligned}$$

Some interesting observations can be obtained from the above CRLB. The CRLB using the virtual carrier present signal model with symbol rate sampling is [26]

$$\text{CRLB}_0(\phi) = \frac{\frac{\sigma^2}{2}}{|\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 - |\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2}.$$

We show through simple algebra that this CRLB is lower bounded by the CRLB that uses oversampled signal model, i.e.,

$$\frac{\frac{\sigma^2}{2}}{C - D} \leq \frac{\frac{\sigma^2}{2}}{|\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 - |\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2}.$$

This leads to

$$|\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2 + |\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 \geq \frac{1}{2}|\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}} + \mathbf{E}^H\mathbf{W}^H\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2.$$

From the triangular inequality, we always have

$$\begin{aligned} |\mathbf{E}^H\mathbf{W}^H\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2 + |\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 \\ \geq \frac{1}{2}|\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}} + \mathbf{E}^H\mathbf{W}^H\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2. \end{aligned} \quad (19)$$

Therefore, we only have to prove

$$|\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2 \geq |\mathbf{E}^H\mathbf{W}^H\mathbf{B}\mathbf{W}\tilde{\mathbf{e}}|^2 \quad (20)$$

which is

$$(\mathbf{B}\mathbf{W}\tilde{\mathbf{e}})^H(\mathbf{I} - \mathbf{W}\mathbf{W}^H)\mathbf{B}\mathbf{W}\tilde{\mathbf{e}} \geq 0 \quad (21)$$

where we have used the fact that $\mathbf{W}\mathbf{E}\mathbf{E}^H\mathbf{W}^H = \mathbf{W}\mathbf{E}(\mathbf{W}\mathbf{E})^H = \mathbf{W}\mathbf{W}^H$. Since the IDFT matrix is unitary and $\mathbf{W}$ is composed of M columns of the IDFT matrix, $\mathbf{I} - \mathbf{W}\mathbf{W}^H$ is therefore non-negative definite. Hence, (21) is always true.

Although this is intuitively true as the information regarding ϕ due to the presence of virtual carriers is still preserved in the oversampling model, we want to point out that the above comparison requires that the noise variances for oversampled and critically sampled signals are the same. This issue will be further discussed in the next section.

Another important observation is about the identifiability using the oversampled signal model. Since identifiability implies that the Fisher information matrix is full rank (hence, finite CRLB), we need to find out the condition such that $C > D$ (i.e., $\text{CRLB}(\phi) < \infty$). First

$$|\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 - |\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 = \tilde{\mathbf{d}}^H\mathbf{W}^H\mathbf{A}(\mathbf{I} - \mathbf{W}\mathbf{W}^H)\mathbf{A}\mathbf{W}\tilde{\mathbf{d}} \geq 0.$$

The above inequality comes from the fact that $\mathbf{I} - \mathbf{W}\mathbf{W}^H$ is non-negative definite. Therefore

$$|\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2 \geq |\mathbf{W}^H\mathbf{A}\mathbf{W}\tilde{\mathbf{d}}|^2.$$

Combine this with (19) and (20), and we have $C \geq D$. Further, from the triangular inequality in (19), equality holds only if $\tilde{\mathbf{d}} =$

0. Thus, if $\tilde{d} \neq 0$ or, equivalently, not all subcarrier frequencies coincide with channel nulls (notice $\tilde{d}_k = H(k)d_k$), then the OFDM CFO is identifiable from the oversampled data. This is also easy to see from the signal model in (3) and (4). Indeed, if at least one of the nonvirtual subcarriers is not “nulled” by the channel frequency response, the phase shift due to the CFO between $\mathbf{x}_1$ and $\mathbf{x}_2$ is still preserved and can be identified.

V. PERFORMANCE EVALUATION

In this section, numerical examples are presented to study the performance of the proposed blind estimation scheme. The parameter setting is as follows. The total number of carriers is $N = 64$; among them, $M = 52$ are information carrying subcarriers, as specified in the current wireless LAN standards [3], [4]. The discrete time carrier spacing is therefore $\omega = 2\pi/N = 0.0982$, with normalized symbol interval $T_s = 1$. The guard time, which is cyclically extended, has a length of 11 symbol intervals. The time dispersive channel is generated using the following model. The maximum delay spread is assumed to be of $\tau_{max} = 9$, which is shorter than the guard time. A six-tap multipath channel is generated with delays uniformly distributed between 0 and τ_{max} . The channel gain for each path is assumed to be independent complex Gaussian with unit variance. The obtained multipath channel is then appropriately normalized. Thus, the channel impulse response is now

$$h(t) = \sum_{k=0}^5 \alpha_k \delta(t - \tau_k). \quad (22)$$

A. Numerical Simulation Compared With CRLB

Fig. 3 gives the numerical results of CRLB as a function of SNR. We use a randomly generated channel with the following parameter:

$$\begin{aligned} \tau &= [5.85, 8.66, 0.40, 6.15, 8.44, 1.42] \\ \alpha &= [-1.09 - 0.71i, -0.23 - 0.25i, -0.73 - 0.41i \\ &\quad 0.74 + 1.32i, -0.29 + 0.10i, 0.11 - 0.60i]. \end{aligned} \quad (23)$$

From Fig. 3, it is seen that the SNR gain for the CRLB using oversampled data is about 3 dB compared with that using the symbol rate sampling that exploits only the presence of virtual carriers. The mean square error (MSE) of the frequency offset estimation using the two ML estimators that minimize $S(\phi)$ and $S_2(\phi)$, respectively, are also plotted in the figure. The MSE is defined as

$$\text{MSE} = \frac{1}{M_c} \sum_{k=1}^{M_c} (\phi - \hat{\phi}(k))^2$$

where M_c is the total Monte Carlo runs, and $\hat{\phi}(k)$ is the frequency offset estimate of the k th Monte Carlo run. The total number of Monte Carlo runs is 500. We see that both estimators can achieve their respective CRLBs fairly closely.

B. Performance Comparison

In this section, we compare the performance of various CFO estimation schemes, including the CP correlation [11]. We use

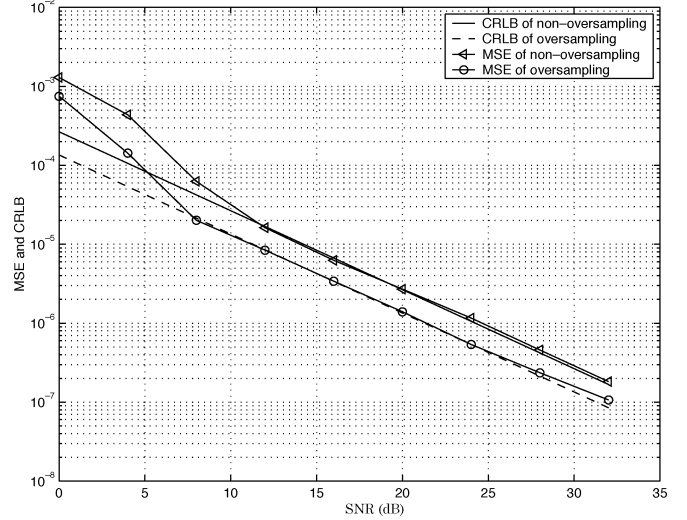


Fig. 3. CRLBs of the two signal models and the normalized MSE of the two ML estimators with and without oversampling.

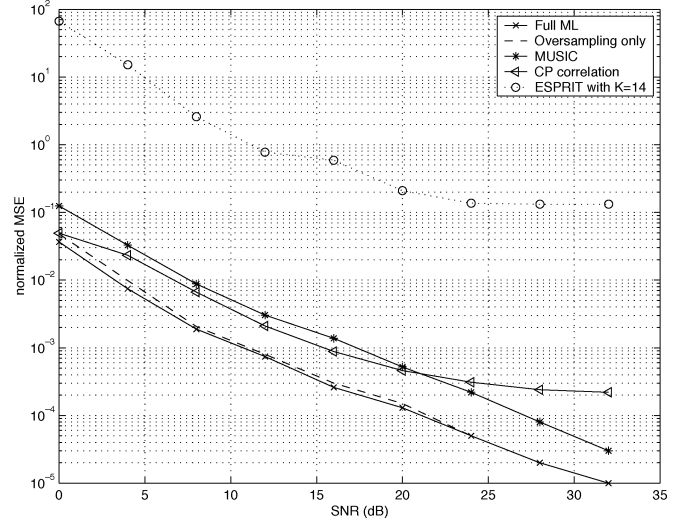


Fig. 4. Normalized MSE of frequency offset estimation of various blind methods for the channel in (23). K is the number of OFDM symbols used in the ESPRIT-based blind estimation algorithm.

the normalized mean square error of the CFO estimate as the performance measure.

$$\text{NMSE} = \frac{1}{M_c} \sum_{k=1}^{M_c} \left(\frac{\phi - \hat{\phi}(k)}{\omega} \right)^2$$

where $\omega = 2\pi/N$ is the carrier spacing. Throughout all examples, the true frequency offset is assumed to be $0.1\omega = 0.0098$, and we use $M_c = 500$ Monte Carlo runs.

Fig. 4 gives the normalized MSEs using the same channel parameter as above. In the legend, “full ML” refers to minimizing $S(\phi)$, whereas “oversampling” refers to minimizing $S_1(\phi)$. Notice that for ESPRIT, multiple blocks have to be used, and we choose $K = 14$ in this example. From Fig. 4, we see that for almost the entire SNR range, the new estimation algorithms (both $S(\phi)$ and $S_1(\phi)$) based on the oversampled model consistently perform better than all other blind schemes. It is also noticeable

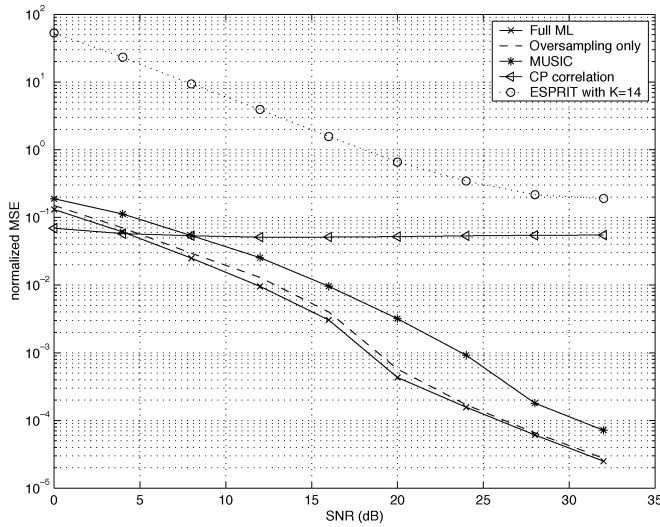


Fig. 5. Average normalized MSE of frequency offset estimation of various blind methods for 30 randomly generated channels.

that the improvement of using $S(\phi)$ over $S_1(\phi)$ is almost negligible, which makes $S_1(\phi)$ more appealing because of the fast root-finding algorithm developed in Section III. We also remark that the performance the CP-based blind algorithm is insensitive to the channel SNR—this is due to the fact that the CP-based algorithm is developed under the assumption of nondispersive channel, e.g., we have a “clean” cyclic prefix that, apart from the channel noise, should be an exact replica of the corresponding data block. In the presence of multipath channel, the interblock interference renders this assumption invalid, and the algorithm exhibits a residual error floor at large SNR.

To obtain a more robust performance comparison, we randomly generate 30 different channels according to model (22) and average the obtained MSE. The result is shown in Fig. 5. Clearly, the oversampling method provides superior performance over a wide range of SNR, as compared with other blind methods tested.

Finally, to test the robustness of the proposed algorithm, we use a channel whose length is longer than the guard interval. In this case, the observations suffer from interblock interference. Here, the channel length is assumed to be 16, whereas the guard time is still 11. Again, 30 randomly generated channels are used, and the average NMSE is obtained for the different blind algorithms tested. We see from Fig. 6 that even though all the algorithms suffer performance degradation, the oversampling-based methods provide the most robust performance in the presence of intersymbol interference.

Although, throughout the above examples, we assumed $\phi = 0.1\omega$, i.e., ϕ is only a fraction of the carrier spacing, the proposed method works equally well for frequency offset exceeding the carrier spacing, as observed through extensive simulations.

C. Front-End Filter and the Effect of Correlated Noise

In the above simulation, we have assumed that the noise variances for the oversampled OFDM signal and the symbol rate sampled OFDM signal are the same. This assumption is only reasonable if the front-end filters to both oversampled and critically sampled systems have the same bandwidth. Notice that it is

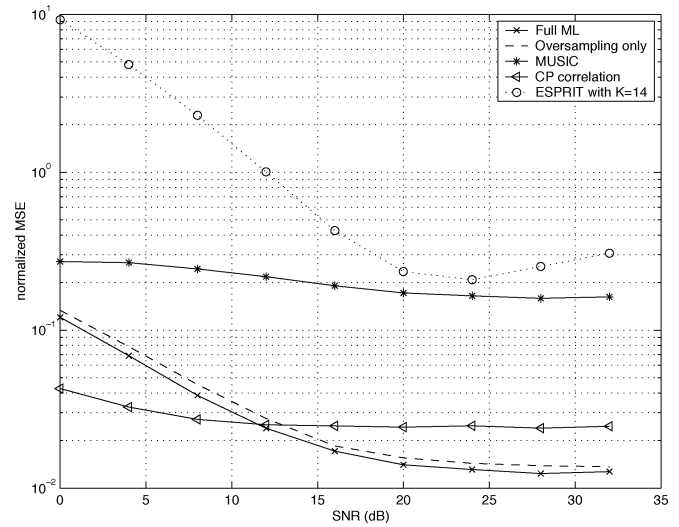


Fig. 6. Normalized MSE of frequency offset estimation of various blind methods averaged over 30 channels with channel length longer than the guard time.

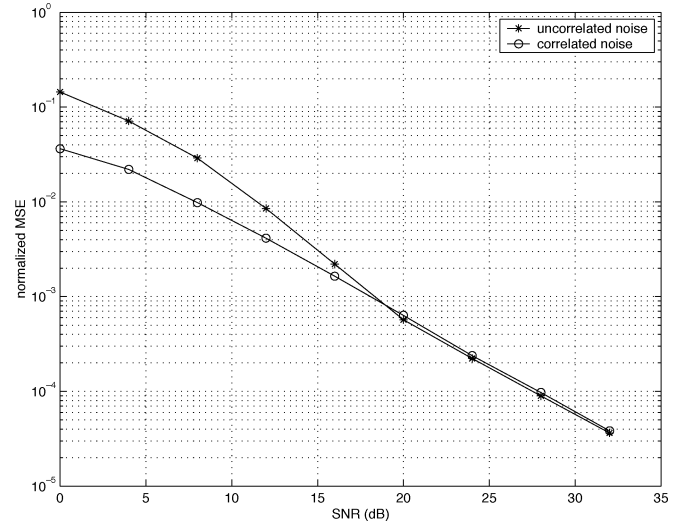


Fig. 7. Normalized MSE with correlated and uncorrelated noise vectors.

always desirable to minimize the front-end filter bandwidth for noise mitigation purposes; thus, if we use the Nyquist front-end filter for a critically sampled OFDM signal, it will introduce correlation among noise observations when oversampling is used. While it is still possible to obtain some performance gain even if we consider the possible 3-dB loss if we double the front-end filter bandwidth, it is more interesting to investigate the performance of the proposed algorithm with correlated noise vectors because of the Nyquist front-end filter. In this case, the noise variance remains the same as critically sampled OFDM signals.

With oversampling by a factor of two, it is easy to show that the noise vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are still white Gaussian vectors, yet $\mathbf{x}_1$ and $\mathbf{x}_2$ are now correlated with each other. The correlation structure can be easily obtained by sampling the sinc function (its autocorrelation function) at twice the Nyquist rate. Using this correlated noise vectors, we implement the proposed oversampling-based algorithm and compare it with the results when the noise vectors are assumed uncorrelated. Fig. 7 is obtained

through averaging over 30 randomly generated channels. From the plot, it can be seen that the performance using correlated noise vectors is as good as, if not better than, that which assumes uncorrelated noises and would certainly be better than the MUSIC-like algorithm [corresponding to minimizing $S_2(\phi)$] with the same noise variance. It should also be pointed out that with correlated noise vectors, the optimal estimation performance can be potentially better when the noise correlation is also utilized.

VI. CONCLUSIONS

In this paper, we investigate blind estimation of the CFO for OFDM systems. Oversampled data is utilized to fully exploit the intrinsic phase shift among neighboring samples due to the CFO. Perfect recovery of the CFO can be achieved in the noiseless case. In the presence of channel noise, the proposed method leads to a simple root-finding problem that allows efficient computation. We have showed the equivalence of such intuition-based scheme to an ML estimate of the CFO in the noisy case. The CRLB is derived, and some intuitive observations are given. Besides its performance advantage over many existing blind CFO estimation algorithms, a more important feature is its data efficiency. Using only a single OFDM block, the proposed algorithm can achieve reliable estimation performance, even in the absence of virtual carriers.

APPENDIX A PROOF OF LEMMA 1

We prove this by contradiction. Define

$$\mathbf{P}(\phi_o - \phi) = \mathbf{P}^H(\phi) \mathbf{P}(\phi_o) = \text{diag} \left[1 e^{j(\phi_o - \phi)} \dots e^{j(N-1)(\phi_o - \phi)} \right].$$

We assume that there exists $\phi \neq \phi_o$ with $\phi \in [0, 2\pi)$ and $\phi_o \in [0, 2\pi)$ such that

$$\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W} = e^{j\frac{\phi_o - \phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W} \mathbf{E}.$$

Should the above equality hold, the traces of both sides of the equality should also be the same. We have

$$\begin{aligned} \text{tr}(\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W}) &= \text{tr} \left(e^{j\frac{\phi_o - \phi}{2}} \mathbf{E}^H \mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W} \mathbf{E} \right) \\ &= e^{j\frac{\phi_o - \phi}{2}} \text{tr}(\mathbf{E}^H \mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W} \mathbf{E}) \\ &= e^{j\frac{\phi_o - \phi}{2}} \text{tr}(\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W} \mathbf{E} \mathbf{E}^H) \\ &= e^{j\frac{\phi_o - \phi}{2}} \text{tr}(\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W}). \end{aligned}$$

Since both ϕ and ϕ_o are in $[0, 2\pi)$, hence, $\phi_o - \phi \in (-2\pi, 2\pi)$. In order for the above equality to be true, we must have, for $\phi_o \neq \phi$ (hence $e^{j(\phi_o - \phi)/2} \neq 1$),

$$\text{tr}(\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W}) = 0 \quad (24)$$

but

$$\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W} = \sum_{k=0}^{N-1} e^{j(\phi_o - \phi)} \mathbf{v}_k \mathbf{v}_k^H$$

where $\mathbf{v}_k$ is the conjugate transpose (Hermitian) of the k^{th} row of matrix $\mathbf{W}$, i.e.,

$$\mathbf{v}_k = \frac{1}{\sqrt{N}} \left[1 e^{j(k-1)\omega} \dots e^{j(M-1)(k-1)\omega} \right]^H.$$

Thus

$$\begin{aligned} \text{tr}(\mathbf{W}^H \mathbf{P}(\phi_o - \phi) \mathbf{W}) &= \sum_{k=0}^{N-1} e^{j(\phi_o - \phi)} \text{tr}(\mathbf{v}_k \mathbf{v}_k^H) \\ &= e^{j(\phi_o - \phi)} \sum_{k=0}^{N-1} \text{tr}(\mathbf{v}_k^H \mathbf{v}_k) \\ &= M e^{j(\phi_o - \phi)} \\ &\neq 0. \end{aligned}$$

This contradicts (24). Hence, we must have $\phi = \phi_o$. This proves uniqueness.

APPENDIX B DERIVATION OF THE CRLB

From (15), the log-likelihood function, neglecting the constant term, can be obtained as

$$\ln L = -(\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}})^H (\mathbf{x}_1 - \mathbf{P} \mathbf{W} \tilde{\mathbf{d}}) - \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{e}} \right)^H \left(\mathbf{x}_2 - e^{j\frac{\phi}{2}} \mathbf{P} \mathbf{W} \tilde{\mathbf{e}} \right).$$

The unknown parameters are $(\bar{\mathbf{d}}, \hat{\mathbf{d}}, \phi)$, where $\bar{\mathbf{d}}$ and $\hat{\mathbf{d}}$ are the real and imaginary parts of the symbol vector $\tilde{\mathbf{d}}$. In the following, j is reserved for $\sqrt{-1}$.

The first step is to find the Fisher information matrix (FIM). To do so, we need to derive the second-order derivatives of the log likelihood functions with respect to all the unknown parameters. Denote by Q_1 and Q_2 the first and second terms in the above log likelihood function. For Q_1 , we have

$$\begin{aligned} \frac{\partial Q_1}{\partial \bar{d}_i} &= \frac{1}{\sigma} \frac{\partial}{\partial \bar{d}_i} [\mathbf{d}^H (\mathbf{P} \mathbf{W})^H \mathbf{x}_1 + \mathbf{x}_1^H \mathbf{P} \mathbf{W} \mathbf{d} - \mathbf{d}^H \mathbf{d}] \\ &= \frac{1}{\sigma} (\mathbf{p}_i^H \mathbf{x}_1 + \mathbf{x}_1^H \mathbf{p}_i - 2\bar{d}_i) \end{aligned} \quad (25)$$

where $\mathbf{p}_i$ is the i th column of matrix $\mathbf{P} \mathbf{W}$. Similarly

$$\frac{\partial Q_1}{\partial \hat{d}_i} = \frac{1}{\sigma} (-j \mathbf{p}_i^H \mathbf{x}_1 + j \mathbf{x}_1^H \mathbf{p}_i - 2\hat{d}_i).$$

From the above two equations, we can get

$$\begin{aligned} \frac{\partial^2 Q_1}{\partial \bar{d}_i \partial \hat{d}_k} &= \frac{\partial^2 Q_1}{\partial \hat{d}_i \partial \bar{d}_k} = 0, \quad \text{for } i \neq k \\ \frac{\partial^2 Q_1}{\partial \bar{d}_i^2} &= \frac{\partial^2 Q_1}{\partial \hat{d}_i^2} = -\frac{2}{\sigma^2}. \end{aligned}$$

In addition, from (25), and recognizing that matrix $\mathbf{P}$ as a function of ϕ as in (6), we get

$$\begin{aligned} \frac{\partial^2 Q_1}{\partial \bar{d}_i \partial \phi} &= \frac{1}{\sigma^2} [(j \mathbf{A} \mathbf{p}_i)^H \mathbf{x}_1 + \mathbf{x}_1^H (j \mathbf{A} \mathbf{p}_i)] \\ \frac{\partial^2 Q_1}{\partial \hat{d}_i \partial \phi} &= \frac{1}{\sigma^2} [-j (j \mathbf{A} \mathbf{p}_i)^H \mathbf{x}_1 + j \mathbf{x}_1^H (j \mathbf{A} \mathbf{p}_i)] \end{aligned}$$

where $\mathbf{A} = \text{diag}(0, 1, 2, \dots, N-1)$ as in Theorem 3. Written in more compact matrix forms, we have

$$\begin{aligned}\frac{\partial^2 Q_1}{\partial \phi \partial \tilde{\mathbf{d}}} &= \frac{2}{\sigma^2} \text{Re} \{ (j\mathbf{A}\mathbf{P}\mathbf{W})^H \mathbf{x}_1 \} \\ \frac{\partial^2 Q_1}{\partial \phi \partial \hat{\mathbf{d}}} &= \frac{2}{\sigma^2} \text{Im} \{ (j\mathbf{A}\mathbf{P}\mathbf{W})^H \mathbf{x}_1 \}.\end{aligned}$$

The remaining term is the second-order derivative with respect to ϕ . We have

$$\begin{aligned}\frac{\partial Q_1}{\partial \phi} &= -\frac{\partial}{\partial \phi} \left\{ \frac{1}{\sigma^2} [(\mathbf{x}_1 - \mathbf{P}\mathbf{W}\tilde{\mathbf{d}})^H (\mathbf{x}_1 - \mathbf{P}\mathbf{W}\tilde{\mathbf{d}})] \right\} \\ &= -\frac{1}{\sigma^2} \frac{\partial}{\partial \phi} \left\{ -(\mathbf{P}\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{x}_1 - \mathbf{x}_1^H \mathbf{P}\mathbf{W}\tilde{\mathbf{d}} \right\} \\ &= -\frac{1}{\sigma^2} \frac{\partial}{\partial \phi} \left\{ -(\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{P}^H \mathbf{x}_1 - \mathbf{x}_1^H \mathbf{P}\mathbf{W}\tilde{\mathbf{d}} \right\} \\ &= -\frac{1}{\sigma^2} \left\{ -(\mathbf{W}\tilde{\mathbf{d}})^H (j\mathbf{A}\mathbf{P})^H \mathbf{x}_1 - \mathbf{x}_1^H (j\mathbf{A}\mathbf{P}) \mathbf{W}\tilde{\mathbf{d}} \right\}.\end{aligned}$$

From the above, we get

$$\begin{aligned}\frac{\partial^2 Q_1}{\partial \phi^2} &= \frac{-1}{\sigma^2} \left\{ -(\mathbf{W}\tilde{\mathbf{d}})^H (j^2 \mathbf{A}^2 \mathbf{P})^H \mathbf{x}_1 - \mathbf{x}_1^H (\mathbf{A}^2 \mathbf{P}) \mathbf{W}\tilde{\mathbf{d}} \right\} \\ &= \frac{-2}{\sigma^2} \text{Re} \{ (\mathbf{A}^2 \mathbf{P}\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{x}_1 \}.\end{aligned}$$

The partial derivatives of Q_2 with respect to $\tilde{\mathbf{d}}$, $\hat{\mathbf{d}}$, and ϕ can be derived in a similar fashion and is summarized below.

$$\begin{aligned}\frac{\partial^2 Q_2}{\partial \hat{d}_i \partial \tilde{d}_k} &= \frac{\partial^2 Q_2}{\partial \tilde{d}_i \partial \hat{d}_k} = 0 \quad (\text{for } i \neq k.) \\ \frac{\partial^2 Q_2}{\partial \tilde{d}_i^2} &= \frac{\partial^2 Q_2}{\partial \hat{d}_i^2} = -\frac{2}{\sigma^2} \\ \frac{\partial^2 Q_2}{\partial \phi \partial \tilde{\mathbf{d}}} &= \frac{2}{\sigma^2} \text{Re} \left\{ \left(j e^{\frac{j\phi}{2}} \mathbf{B}\mathbf{P}\mathbf{W}\mathbf{E} \right)^H \mathbf{x}_2 \right\} \\ \frac{\partial^2 Q_2}{\partial \phi \partial \hat{\mathbf{d}}} &= \frac{2}{\sigma^2} \text{Im} \left\{ \left(j e^{\frac{j\phi}{2}} \mathbf{B}\mathbf{P}\mathbf{W}\mathbf{E} \right)^H \mathbf{x}_2 \right\} \\ \frac{\partial^2 Q_2}{\partial \phi^2} &= -\frac{2}{\sigma^2} \text{Re} \left\{ \left(e^{\frac{j\phi}{2}} \mathbf{B}^2 \mathbf{P}\mathbf{W}\mathbf{E}\tilde{\mathbf{d}} \right)^H \mathbf{x}_2 \right\}\end{aligned}$$

where $\mathbf{B} = \text{diag}(1/2, 3/2, \dots, (N-1/2))$, as in Theorem 3.

Since the log likelihood function is a sum of Q_1 and Q_2 , the FIM is, therefore, as shown in the equation at the bottom of the page, where

$$\mathbf{u} = (j\mathbf{A}\mathbf{P}\mathbf{W})^H \mathbf{x}_1 + \left(j e^{\frac{j\phi}{2}} \mathbf{B}\mathbf{P}\mathbf{W}\mathbf{E} \right)^H \mathbf{x}_2$$

and

$$\begin{aligned}\mathbf{v} &= E(\mathbf{u}) \\ &= (j\mathbf{A}\mathbf{P}\mathbf{W})^H (\mathbf{P}\mathbf{W}\tilde{\mathbf{d}}) + \left(j e^{\frac{j\phi}{2}} \mathbf{B}\mathbf{P}\mathbf{W}\mathbf{E} \right)^H \left(e^{\frac{j\phi}{2}} \mathbf{P}\mathbf{W}\mathbf{E}\tilde{\mathbf{d}} \right).\end{aligned}$$

Using the fact that $\mathbf{A}$ and $\mathbf{B}$ are both diagonal, and thus, $\mathbf{P}^H \mathbf{A}^2 \mathbf{P} = \mathbf{A}^2$ and $\mathbf{P}^H \mathbf{B}^2 \mathbf{P} = \mathbf{B}^2$, a simpler expression is obtained as

$$J = \frac{2}{\sigma^2} \begin{bmatrix} J_{11} & \text{Re}\{\mathbf{w}^T\} & \text{Im}\{\mathbf{w}^T\} \\ \text{Re}\{\mathbf{w}\} & 2\mathbf{I}_M & \mathbf{O}_M \\ \text{Im}\{\mathbf{w}\} & \mathbf{O}_M & 2\mathbf{I}_M \end{bmatrix}$$

where

$$J_{11} = [(\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{A}^2 (\mathbf{W}\tilde{\mathbf{d}}) + (\mathbf{W}\mathbf{E}\tilde{\mathbf{d}})^H \mathbf{B}^2 (\mathbf{W}\mathbf{E}\tilde{\mathbf{d}})]$$

and $\mathbf{w} \triangleq (j\mathbf{A}\mathbf{W})^H (\mathbf{W}\tilde{\mathbf{d}}) + (j\mathbf{B}\mathbf{W}\mathbf{E})^H (\mathbf{W}\mathbf{E}\tilde{\mathbf{d}}) = \tilde{\mathbf{w}} + j\hat{\mathbf{w}}$, with $\tilde{\mathbf{w}}$ and $\hat{\mathbf{w}}$ being the real and imaginary parts of $\mathbf{w}$, respectively. To find the CRLB for ϕ , we need to invert the FIM and find the element corresponding to ϕ , which is $[J^{-1}]_{1,1}$, i.e., the element of the first row and column. Using inverse matrix for a partitioned matrix [27, pp. 247–248], we get

$$\begin{aligned}[J^{-1}]_{11} &= \frac{1}{J_{11} - [\tilde{\mathbf{w}}^T, \hat{\mathbf{w}}^T] \begin{bmatrix} \tilde{\mathbf{w}} \\ \hat{\mathbf{w}} \end{bmatrix}} \\ &= \frac{\frac{\sigma^2}{2}}{(\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{A}^2 (\mathbf{W}\tilde{\mathbf{d}}) + (\mathbf{W}\mathbf{E}\tilde{\mathbf{d}})^H \mathbf{B}^2 (\mathbf{W}\mathbf{E}\tilde{\mathbf{d}}) - |\mathbf{W}|^2}\end{aligned}$$

which can be verified easily to be equivalent to that in (18).

REFERENCES

- [1] Eur. Telecommun. Stand., *Radio Broadcasting Systems: Digital Audio Broadcasting to Mobile, Portable and Fixed Receivers*, Feb. 1995.
- [2] Eur. Telecommun. Stand., *Digital Video Broadcasting: Framing Structure, Channel Coding, and Modulation for Digital Terrestrial Television*, Aug. 1997.
- [3] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std 802.11a, 1999.
- [4] Broadband Radio Access Networks (BRAN): HIPERLAN Type 2 Technical Specification Part 1—Physical Layer, Oct. 1999.
- [5] T. Pollet, M. van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, pp. 191–193, Feb./Mar./Apr. 1993.
- [6] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, pp. 2908–2914, Oct. 1994.
- [7] M. Luise and R. Reggiannini, "Carrier frequency acquisition and tracking for OFDM systems," *IEEE Trans. Commun.*, vol. 44, pp. 1590–1598, Nov. 1996.
- [8] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, pp. 1613–1621, Dec. 1997.

$$\begin{aligned}J &= -E \left\{ \begin{bmatrix} -\frac{2}{\sigma^2} \text{Re} \left\{ (\mathbf{A}^2 \mathbf{P}\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{x}_1 + \left(e^{\frac{j\phi}{2}} \mathbf{B}^2 \mathbf{P}\mathbf{W}\mathbf{E}\tilde{\mathbf{d}} \right)^H \mathbf{x}_2 \right\} & \frac{2}{\sigma^2} \text{Re}\{\mathbf{u}\}^T & \frac{2}{\sigma^2} \text{Im}\{\mathbf{u}\}^T \\ \frac{2}{\sigma^2} \text{Re}\{\mathbf{u}\} & -\frac{2}{\sigma^2} \mathbf{I}_M - \frac{2}{\sigma^2} \mathbf{I}_M & \mathbf{O}_M \\ \frac{2}{\sigma^2} \text{Im}\{\mathbf{u}\} & \mathbf{O}_M & -\frac{2}{\sigma^2} \mathbf{I}_M - \frac{2}{\sigma^2} \mathbf{I}_M \end{bmatrix} \right\} \\ &= \frac{2}{\sigma^2} \begin{bmatrix} (\mathbf{P}\mathbf{W}\tilde{\mathbf{d}})^H \mathbf{A}^2 (\mathbf{P}\mathbf{W}\tilde{\mathbf{d}}) + (\mathbf{P}\mathbf{W}\mathbf{E}\tilde{\mathbf{d}})^H \mathbf{B}^2 (\mathbf{P}\mathbf{W}\mathbf{E}\tilde{\mathbf{d}}) & -\text{Re}\{\mathbf{v}^T\} & -\text{Im}\{\mathbf{v}^T\} \\ -\text{Re}\{\mathbf{v}\} & 2\mathbf{I}_M & \mathbf{O}_M \\ -\text{Im}\{\mathbf{v}\} & \mathbf{O}_M & 2\mathbf{I}_M \end{bmatrix}\end{aligned}$$

- [9] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Commun. Lett.*, vol. 3, pp. 75–77, Mar. 1999.
- [10] S. Zazo and J. M. Paez-Borralló, "Analysis of a new frequency synchronization scheme in OFDM systems," *Signal Process.*, vol. 81, pp. 1695–1704, 2001.
- [11] J. van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 1800–1805, July 1997.
- [12] T. M. Schmidl and D. C. Cox, "Blind synchronization for OFDM," *Electron. Lett.*, vol. 33, pp. 113–114, Feb. 1997.
- [13] Y. 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, pp. 664–655, Mar. 2001.
- [14] H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," *IEEE Commun. Lett.*, vol. 2, pp. 104–106, Apr. 1998.
- [15] U. Tureli, H. Liu, and M. Zoltowski, "OFDM blind carrier offset estimation: ESPRIT," *IEEE Trans. Commun.*, vol. 48, pp. 1459–1461, Sept. 2000.
- [16] S. Wei, D. L. Goeckel, and P. E. Kelly, "A modern extreme value theory approach to calculating the distribution of the peak-to-average power ratio in OFDM systems," in *Proc. IEEE Int. Conf. Commun.*, New York, Apr. 2002, pp. 156–159.
- [17] H. Wang and B. Chen, "On the correlation of OFDM symbol powers: some observations, derivations, and applications," in *Proc. Conf. Inform. Sci., Syst.*, Baltimore, MD, Mar. 2003.
- [18] X. Ma and G. B. Giannakis, "Unifying and optimizing null-subcarrier based frequency-offset estimators for OFDM," in *Proc. Int. Conf. Inform. Commun., Signal Process.*, Singapore, Oct. 2001.
- [19] M. Ghogho, A. Swami, and G. B. Giannakis, "Optimizing null-subcarrier selection for CFO estimation in OFDM over frequency-selective fading channels," in *Proc. GLOBECOM*, Nov. 2001.
- [20] B. Chen, "Maximum likelihood estimation of OFDM carrier frequency offset," *IEEE Signal Processing Lett.*, vol. 9, pp. 123–126, Apr. 2002.
- [21] U. Tureli, D. Kivanc, and H. Liu, "Multicarrier synchronization with diversity," in *Proc. Veh. Technol. Conf.*, Atlantic City, NJ, Oct. 1995, pp. 952–956.
- [22] R. van Nee and R. Prasad, *OFDM For Multimedia Wireless Communications*. Boston, MA: Artech House, 2000.
- [23] H. Bolcskei, P. Duhamel, and R. Hleiss, "Design of pulse shaping OFDM/OQAM systems for high data-rate transmission over wireless channels," in *Proc. IEEE Int. Conf. Commun.*, vol. 1, Vancouver, BC, Canada, June 1999, pp. 559–564.
- [24] H. Ochiai and H. Imai, "On the distribution of the peak-to-average power ratio in OFDM signals," *IEEE Trans. Commun.*, vol. 49, pp. 282–289, Feb. 2001.
- [25] W. Y. Zou and Y. Wu, "COFDM: An overview," *IEEE Trans. Broadcasting*, vol. 41, pp. 1–8, Mar. 1995.
- [26] B. Chen and H. Wang, "Maximum likelihood estimation of OFDM carrier frequency offset," in *Proc. IEEE Int. Conf. Commun.*, New York, Apr. 2002.
- [27] J. Schott, *Matrix Analysis for Statistics*. New York: Wiley, 1996.

Biao Chen (S'98–M'99) received the M.S. degree in statistics and the Ph.D. degree in electrical engineering from the University of Connecticut, Storrs, in 1998 and 1999, respectively.

From 1999 to 2000, he was with Cornell University, Ithaca, NY, as a Post-Doctoral research associate. Since 2000, he has been with Syracuse University, Syracuse, NY, as an assistant professor with the Department of Electrical Engineering and Computer Science. His area of interest mainly focuses on statistical signal processing with applications in wireless communications and distributed sensor networks.

Hao Wang (S'03) received the B.S. degree from Tsinghua University, Beijing, China, in 1996 and the M.S. degree from the University of Connecticut, Storrs, in 1998, both in electrical engineering. Since January 2001, he has been pursuing the Ph.D. degree with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY.

From 1999 to 2000, he was with Schlumberger Technologies Asian Ltd., Beijing. His current research interests are statistical signal processing and multicarrier communications, specifically, channel estimation, timing/frequency synchronization, and peak power analysis for OFDM systems.