

# Phase Noise Estimation and Mitigation for OFDM Systems

Songping Wu, Pan Liu, and Yeheskel Bar-Ness

Corresponding Author: Songping Wu

Tel: (408)-222-0778

Email: swu@marvell.com

**Abstract**—OFDM suffers from severe performance degradation in the presence of phase noise. In particular, phase noise leads to common phase error (CPE) as well as intercarrier interference (ICI) in the frequency domain. Some approaches in the literature mitigate phase noise by directly evaluating and then compensating for CPE or ICI, while others choose to correct phase noise in the time domain.

A new parametric model of OFDM signals is proposed in this paper which shows that, in the presence of phase noise, each received frequency-domain subcarrier signal can be expressed as a sum of all subcarrier signals weighted by a vector parameter. Then, two reduced-complexity techniques are presented to estimate this weighting vector. The first is a maximum likelihood (ML) method whereas the second one is a linear minimum mean square error (LMMSE) technique. Using the obtained estimates, we also propose two approaches, i.e., a decorrelator and an interference canceler, to mitigate phase noise. It is shown that most conventional methods can be readily obtained from our approaches with some approximation or orthogonal transform.

Theoretical analysis and numerical results are provided to elaborate the proposed schemes. We show that the performance of both approaches is superior to that of conventional methods. Furthermore, LMMSE gives the best performance, while ML provides a much simpler yet effective way to mitigate phase noise.

**Index Terms**—OFDM, phase noise, LMMSE, ML.

## I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) has raised a lot of interest both in wireline and wireless communications. It has been deployed in many applications, including digital subscriber line (DSL), digital video/audio broadcasting (DVB/DAB), IEEE 802.11a wireless local area networks (WLAN) and European high performance local area networks (HIPERLAN/2) [1]–[3]. In comparison to single carrier transmission, OFDM is quite effective in combating channel multipath fading with relatively low complexity while providing high spectral efficiency. This makes it quite attractive for future high rate wireless multimedia communications.

One disadvantage of OFDM, however, is its sensitivity to frequency offset and phase noise, both of which destroy the orthogonalities among OFDM subcarriers, cause the leakage of digital Fourier transform (DFT), and eventually lead to common phase error (CPE) and intercarrier interference (ICI) among subcarriers. Frequency offset is a deterministic phenomenon and many approaches have been presented in the

literature to mitigate it [4]–[7]. On the other hand, phase noise, which is the focus of this paper, is a random process that proves to be more complicated than frequency offset.

Phase noise has been extensively analyzed in the literature [8]–[12]. Different approaches have been proposed to eliminate its effects using in-band pilots. These approaches can be categorized into two: the time-domain approach [13], [14], and the frequency-domain approach [15], [16].

The time-domain approach mitigates phase noise at the receiver before the digital Fourier transform (DFT). In particular, a time-domain approach presented in [13], estimates phase noise using a pilot tone surrounded by the guard band. This approach requires a special pilot pattern which, however, restricts its applicability. A more general approach was proposed in [14] which interprets the time domain phase noise by orthogonal transforms, such as discrete cosine transform (DCT) or DFT, and turns phase noise mitigation into the recovery of DCT/DFT-based waveforms. This approach is based on the least-squares (LS) method of [14, Eq. (9)], which requires a large number of pilots to guarantee a satisfactory bit error rate (BER) performance. Furthermore, the method in [14, Eq. (9)] was only considered in the absence of channel fading.

The frequency-domain approach involves the correction of CPE and ICI. The method introduced in [15] directly compensates for the CPE or for its phase. Such a method may not always be effective as it neglects ICI, an important contribution of phase noise. In [16], a new phase noise suppression (PNS) algorithm was proposed, which approximates ICI as random Gaussian noise and obtains the estimates of the CPE and the noise energy. These estimates are then applied to a minimum mean square error (MMSE) equalizer. By simultaneously mitigating both CPE and ICI, this method provides performance gain over the conventional one. An alternative MMSE approach was proposed in [17] to deal with phase noise in a AWGN channel. In this scheme, the equalization of the received signal is required before phase noise estimation. Since multiple OFDM symbols are needed for the MMSE estimation, its complexity may become an issue in practice.

In the presence of phase noise and AWGN noise, each received subcarrier signal can be expressed as the weighted sum of all transmitted subcarrier signals multiplied by the corresponding channel response on each subcarrier. Moreover, both CPE and ICI are functions of those weighting coefficients,

so that, once those weighting coefficients are obtained, phase noise can be readily mitigated. This idea was initially proposed in [18] and [19]. Based on this observation, we extend and propose two general approaches to estimate these phase noise parameters with significantly reduced complexity. Two phase noise mitigation schemes are then presented that employ the estimated phase noise parameters. It is shown that most conventional approaches based on in-band pilots, either in the time domain, or in the frequency domain, can be readily obtained from our proposed methods with some simple approximation or by an orthogonal transform. Numerical results are also provided to illustrate the effectiveness of the proposed schemes.

This paper is organized as follows. Section II presents the general phase noise and OFDM system models. The new approaches for phase noise estimation and mitigation are proposed in Section III, and their performance analysis is provided in Section IV. Section V provides the numerical results that demonstrate the effectiveness of the proposed schemes compared with previous approaches. Section VI concludes this paper.

## II. SYSTEM MODEL

We consider an  $N$ -subcarrier OFDM system with symbol period  $T$  and cyclic prefix length  $N_g$ . Moreover, perfect frequency and timing synchronization between transmitter and receiver is assumed.

### A. Phase Noise Model

Phase noise is typically modeled as a continuous Brownian motion or random Wiener process [9], given by

$$\phi(t) = \int_0^t u(t) dt \quad (1)$$

where  $u(t)$  is a zero-mean Gaussian random variable. For the  $n$ th data signal of the  $m$ th OFDM symbol, phase noise can then be described by a discrete Markov model which gives [14]

$$\phi_m(n) = \sum_{i=0}^{m(N+N_g)+N_g+n} u(i) \quad (2)$$

where the variance of i.i.d.  $u(i)$  is given by  $\sigma_u^2 = 2\pi\beta T/N$ , with  $\beta$  denoting the two-sided 3-dB linewidth<sup>1</sup>.

The correlation function of  $\phi_m(n)$  within the same OFDM symbol is shown to be [8] [11]

$$\begin{aligned} & E[\phi_m(n) \phi_m^*(l)] \\ &= \frac{2\pi\beta T}{N} \cdot \min(m(N+N_g) + n, m(N+N_g) + l) \end{aligned} \quad (3)$$

Consequently, the correlation function of  $e^{j\phi_m(n)}$  is given by

$$\begin{aligned} & E[e^{j\phi_m(n)} e^{-j\phi_m(l)}] = \\ & E\left[e^{j\sum_{i=m(N+N_g)+l+1}^{m(N+N_g)+n} u(i)}\right] = \prod_{i=1}^{n-l} \Psi(v_i|v_i=1) \end{aligned} \quad (4)$$

<sup>1</sup>The word "linewidth" is usually used in the literature to describe the 3-dB bandwidth of phase noise power spectrum.

where the characteristic function of the random Gaussian variable  $u(i)$ ,  $\Psi(v_i) = E[e^{jv_i u(i)}]$  is given by  $\Psi(v_i) = e^{-\pi\beta T v_i^2/N}$  [20], which further simplifies (4) as

$$E[e^{j\phi_m(n)} e^{-j\phi_m(l)}] = e^{-\pi\beta T(n-l)/N} \quad (5)$$

In the process of deriving (5), we have assumed  $n \geq l$ . As a matter of fact, we can readily extend (5) to a general case, namely

$$E[e^{j\phi_m(n)} e^{-j\phi_m(l)}] = e^{-\pi\beta T|n-l|/N} \quad (6)$$

which holds for arbitrary  $n$  and  $l$ . It follows that phase noise, as shown in (2) and (3), is a non-stationary random process with its correlation function changing with the time  $m$ . Nevertheless, the exponential process of random phase noise is stationary, as shown in (6).

### B. OFDM System Model

OFDM partitions the incoming data stream into  $N$  low-rate parallel substreams, which modulate a set of subcarriers using inverse digital Fourier transform (IDFT) so as to obtain the time domain signals. A cyclic prefix is then added to the time domain signal to eliminate inter-symbol interference (ISI) caused by channel multipath fading and enables simple channel equalization at the receiver.

Assuming perfect frequency and timing synchronizations, we stay focused on the effects of phase noise on OFDM signals. For a high-rate OFDM system, channel is often assumed invariant within a block and modeled as  $L$ -tap  $T/N$ -spaced FIR filter  $\mathbf{g} = [g(0), g(1), \dots, g(L-1)]^T$ , where  $L$  is assumed to be less than  $N_g$ . Such a case can be found in the IEEE 802.11a WLAN standard [2].

After removing the cyclic prefix and taking the length- $N$  DFT at the receiver, the received  $k$ th subcarrier signal of the  $m$ th symbol can be expressed as [16]

$$\begin{aligned} y_m(k) &= x_m(k)h(k)c_m(0) \\ &+ \sum_{\substack{l=0 \\ l \neq k}}^{N-1} x_m(l)h(l)c_m(l-k) + z_m(k) \end{aligned} \quad (7)$$

where  $x_m(k)$  and  $h(k)$  are the frequency-domain signal and channel gain, respectively;  $z_m(k)$  denotes the AWGN noise with zero mean and variance  $\sigma^2$ . For the sake of simplicity, the transmitted signals  $x_m(k)$  are assumed to be mutually independent with  $E[|x_m(k)|^2] = E_x$ , while the energy of the channel gain  $h(k)$  is normalized to unity, namely  $E[|h_m(k)|^2] = 1$ . In (7),  $c_m(p)$  is given by

$$c_m(p) = \frac{1}{N} \sum_{n=0}^{N-1} e^{j2\pi np/N + j\phi_m(n)} \quad (8)$$

with  $\phi_m(n)$  denoting the random phase noise.

Equation (7) suggests that the received signal  $y_m(k)$  is the weighted sum of  $x_m(k)h(k)$  plus AWGN noise with the weighting coefficients denoted by  $\{c_m(p)\}_{p=0}^{N-1}$ . As a result, two detrimental effects occur:

- 1) Common phase error (CPE) denoted by  $c_m(0)$ ;

2) Intercarrier interference (ICI) represented by the second term of (7) as a function of  $\{c_m(p)\}_{p=1}^{N-1}$ .

With (8) being the IDFT of  $e^{j\phi_m(n)}$ ,  $c_m(p) = c_m(p \bmod N)$  is simply the frequency-domain expression of phase noise. Therefore, both effects are the corresponding results of phase noise, which eventually leads to system performance degradation.

### III. PHASE NOISE MITIGATION

#### A. Parameter Estimation

1) *Frequency Domain Parameters*: Equation (7) can be expressed, in a matrix form, as<sup>2</sup>

$$\mathbf{y} = \underbrace{\begin{pmatrix} \mathbf{w}_0 & \mathbf{w}_1 & \cdots & \mathbf{w}_{N-1} \end{pmatrix}}_{\mathbf{W}} \cdot \mathbf{c} + \mathbf{z} \quad (9)$$

where  $\mathbf{y} = [y_0, y_1, \dots, y_{N-1}]^T$ ; vectors  $\mathbf{z}$  and  $\mathbf{c}$  take the same form as  $\mathbf{y}$ ; matrix  $\mathbf{W}$  consists of the frequency-domain transmitted signals  $x(k)$  and the channel response  $h(k)$ . The first column vector of matrix  $\mathbf{W}$  is defined as  $\mathbf{w}_0 = (a_0 \ \cdots \ a_{k-1})^T$  with  $a_k = x(k)h(k)$ ,  $k = 0, 1, \dots, N-1$ . It follows that  $\mathbf{w}_k$  is just a left circulant  $k$ th-entry shift of  $\mathbf{w}_0$  and therefore  $\mathbf{W}$ , which is a circulant matrix determined by  $\mathbf{w}_0$ , is usually nonsingular in practice. The frequency-domain approach of phase noise mitigation is to find the vector  $\mathbf{c}$  and compensate for it.

2) *Time Domain Parameters*: Equation (8) can be rewritten in vector form as

$$\mathbf{c} = \frac{1}{\sqrt{N}} \mathbf{F}^H \mathbf{t} \quad (10)$$

where  $\mathbf{t} = [e^{j\phi(0)}, e^{j\phi(1)}, \dots, e^{j\phi(N-1)}]^T$ , and  $\mathbf{F}$  is the DFT matrix denoted by

$$\mathbf{F} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j2\pi/N} & \cdots & e^{-j2\pi(N-1)/N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j2\pi(N-1)/N} & \cdots & e^{-j2\pi(N-1)(N-1)/N} \end{pmatrix} \quad (11)$$

From equation (10), we have

$$\mathbf{t} = \sqrt{N} \mathbf{F} \mathbf{c} \quad (12)$$

which suggests that vector  $\mathbf{t}$  is actually the FFT of the desired frequency-domain vector  $\mathbf{c}$ . Therefore, once we have obtained the vector  $\mathbf{c}$  via a certain estimator, the time-domain vector  $\mathbf{t}$  can be obtained without any loss in estimation accuracy. Hereinafter, we will focus on different estimation techniques in the frequency domain. Using (12), it is quite straightforward to extend those methods to the time domain.

<sup>2</sup>Hereinafter, the symbol index  $m$  is omitted since it does not affect our analysis.

#### B. Maximum Likelihood Estimation (MLE)

We notice that in (9),  $\mathbf{y}$  is a Gaussian random variable with mean  $\mathbf{W}\mathbf{c}$  and variance  $\sigma_z^2 \mathbf{I}$ . Maximum likelihood estimation (MLE) is a method that asymptotically achieves the Cramer-Rao lower bound (CRLB) by minimizing the conditional probability distribution function  $p(\mathbf{y}|\mathbf{c})$  [21]. In additive white Gaussian noise  $\mathbf{z}$ , this is equivalent to minimizing the squared Euclidean distance<sup>3</sup>. Likewise, we use this superscript for all the other estimation results as well.

$$\check{\mathbf{c}} = \arg \min_{\mathbf{c}} \|\mathbf{y} - \mathbf{W}\mathbf{c}\|^2 \quad (13)$$

which leads to

$$\check{\mathbf{c}}_{ml} = (\mathbf{W}^H \mathbf{W})^{-1} \mathbf{W}^H \mathbf{y} = \mathbf{W}^{-1} \mathbf{y} \quad (14)$$

This estimator was first introduced in [18] for phase noise estimation where it was termed as simultaneous CPE and ICI correction (SCIC). Note that, both [18] and [19] proposed to estimate only a subset of phase noise vector  $\mathbf{c}$ . These are actually the simplified versions of the MLE in (14). For the time domain approach, (12) simply gives

$$\check{\mathbf{t}}_{ml} = \sqrt{N} \mathbf{F} \check{\mathbf{c}}_{ml} \quad (15)$$

Note that the time-domain approach proposed in [14] actually leads to the same estimate as (15) when the number of pilots equals the number of subcarriers  $N$ . In other words, the difference between (14) and the approach in [14] is just an orthogonal DFT transform.

#### C. Linear Minimum Mean Square Error Estimation (LMMSEE)

For the received signals in (9), the linear minimum mean square error estimation (LMMSEE) gives [21]

$$\check{\mathbf{c}}_{lmmse} = \mathbf{R}_{cy} \mathbf{R}_{yy}^{-1} \mathbf{y} \quad (16)$$

Since

$$\mathbf{R}_{cy} = E[\mathbf{c}\mathbf{y}^H] = \mathbf{R}_c \mathbf{W}^H \quad (17)$$

and

$$\mathbf{R}_{yy} = E[\mathbf{y}\mathbf{y}^H] = \mathbf{W} \mathbf{R}_c \mathbf{W}^H + \sigma_z^2 \mathbf{I} \quad (18)$$

where  $\mathbf{R}_c = E[\mathbf{c}\mathbf{c}^H]$ . Using (5), we denote  $\mathbf{R}_c$  by

$$[\mathbf{R}_c]_{mn} = e^{-\pi\beta T|m-n|/N}, \text{ with } m, n \in [1, N]$$

Substituting (17) and (18) into (16) yields

$$\check{\mathbf{c}}_{lmmse} = \mathbf{R}_c \left[ \mathbf{R}_c + \sigma_z^2 (\mathbf{W}^H \mathbf{W})^{-1} \right]^{-1} \mathbf{W}^{-1} \mathbf{y} \quad (19)$$

Likewise, in terms of (12), the time-domain approach gives

$$\check{\mathbf{t}}_{lmmse} = \sqrt{N} \mathbf{F} \check{\mathbf{c}}_{lmmse} \quad (20)$$

As can be seen in (19), the LMMSEE requires the statistics of phase noise as well as AWGN noise. In other words, the LMMSEE depends on the phase noise linewidth  $\beta$  and the AWGN noise variance  $\sigma_z^2$ . In practice, once the receiver design is done, the phase noise linewidth is known. In cases where a variation in temperature causes the change of the linewidth

<sup>3</sup>The estimation result is indicated by the superscript  $\check{\cdot}$ .

$\beta$ , some adaptive method could be introduced to track this change. The AWGN noise variance can be determined from the receiver bandwidth and noise figure.

Note that the method proposed in [16] can be obtained directly from an approximation of this method with significantly reduced computational complexity. We have the following remarks for the results we obtained:

*Remark 1:* We show that conventional frequency-domain approaches, where the ICI is approximated by a Gaussian random variable, are the direct results of our methods. Rewrite (9) as

$$\mathbf{y} = \mathbf{w}_0 \mathbf{c}_0 + \dot{\mathbf{z}} \quad (21)$$

where  $\dot{\mathbf{z}} = (\mathbf{w}_1 \ \cdots \ \mathbf{w}_{N-1}) \cdot \mathbf{c}^1 + \mathbf{z}$ , with  $\mathbf{c}^1$  denoting the vector  $\mathbf{c}$  without the first element. For small phase noise and a large number of subcarriers  $N$ , the CPE factor  $c_0$  contains most of the effects of phase noise. Therefore, it may be sufficient to estimate  $c_0$  using the ML method [15], or using the MMSE method for better performance [16]. This suggests that conventional frequency-domain methods, such as those in [15] and [16], are just the simplified versions of the methods proposed in the paper by estimating a subset of the weighting coefficient vector.

*Remark 2:* Conventional time-domain approaches, such as the one proposed in [14], give the corresponding time-domain expression of our estimated vector by applying IDFT to (14) or (19), as shown in (15) or (20). However, it is worth mentioning that [14] claims that the  $N$  parameters of the time-domain phase noise could be transformed by DCT/DFT into  $L$  ( $L < N$ ) parameters, which can then be estimated by using  $M$  pilot signals. Nevertheless, it is readily understood that one can not guarantee that such a transform does always exist. In other words, an approximation might be needed in some situations. Therefore,  $L$  can not be much smaller than  $N$  in order to make the approximation applicable. On the other hand, the condition  $L \leq M \leq N$  must be satisfied in order to obtain the solution. Therefore, a large number of pilots  $M$  is necessary, which significantly decreases spectral efficiency.

When using decision feedback, our approach has a high spectral efficiency. Moreover, unlike [14] where an AWGN channel is assumed, we explicitly deal with fading channel.

#### D. Complexity Reduction

Both the MLE and LMMSEE methods require the inversion of the matrix  $\mathbf{W}$ , which subsequently needs the computational complexity of the order of  $N^3$ . Similar to the shift-forward circulant matrix introduced in [22],  $\mathbf{W}$  is a shift-backward circulant matrix. Likewise,  $\mathbf{W}$  can also be readily decomposed into (see Appendix)

$$\mathbf{W} = \mathbf{F}^H \mathbf{V} \mathbf{F} \mathbf{P} \quad (22)$$

where  $\mathbf{F}$  is the DFT matrix,  $\mathbf{V}$  is a diagonal matrix defined in (A-8) and  $\mathbf{P}$  is a permutation matrix defined in (A-3). The inversion of  $\mathbf{W}$  is then given by

$$\mathbf{W}^{-1} = \mathbf{P}^{-1} \mathbf{F}^H \mathbf{V}^{-1} \mathbf{F} = \mathbf{P} \mathbf{F}^H \mathbf{V}^{-1} \mathbf{F} \quad (23)$$

where we used  $\mathbf{P}^{-1} = \mathbf{P}$ . Substituting (23) into the MLE (14) and the LMMSEE (19), we then have

$$\check{\mathbf{c}}_{ml} = \mathbf{P} \mathbf{F}^H \mathbf{V}^{-1} \mathbf{F} \mathbf{y} \quad (24)$$

On the other hand, using the Matrix Inversion Lemma

$$(\mathbf{P}^{-1} + \mathbf{A}^H \mathbf{B} \mathbf{A})^{-1} = \mathbf{P} - \mathbf{P} \mathbf{A}^H (\mathbf{A} \mathbf{P} \mathbf{A}^H + \mathbf{B}^{-1})^{-1} \mathbf{A} \mathbf{P} \quad (25)$$

leads to

$$\begin{aligned} & \mathbf{R}_c \left[ \mathbf{R}_c + \sigma_z^2 (\mathbf{W}^H \mathbf{W})^{-1} \right]^{-1} \\ &= \mathbf{I} - \left[ \mathbf{I} + \sigma_z^{-2} \mathbf{R}_c \mathbf{W}^H \mathbf{W} \right]^{-1} \end{aligned} \quad (26)$$

By substituting (23) and (26) into (19),  $\check{\mathbf{c}}_{lmmse}$  becomes

$$\begin{aligned} \check{\mathbf{c}}_{lmmse} &= \left\{ \mathbf{I} - \left[ \mathbf{I} + \sigma_z^{-2} \mathbf{R}_c \mathbf{W}^H \mathbf{W} \right]^{-1} \right\} \mathbf{W}^{-1} \mathbf{y} \\ &= \left\{ \mathbf{I} - \left[ \mathbf{I} + \sigma_z^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P} \right]^{-1} \right\} \\ & \quad \cdot \mathbf{P} \mathbf{F}^H \mathbf{V}^{-1} \mathbf{F} \mathbf{y} \end{aligned} \quad (27)$$

We notice that

- 1) The inversion of the diagonal matrix  $\mathbf{V}$  is quite simple;
- 2)  $\mathbf{P}$  is a simple permutation matrix;
- 3)  $\mathbf{F}^H$  or  $\mathbf{F}$  can be implemented with fast algorithms;

Therefore, the computational complexity is greatly reduced with respect to the inversion of  $\mathbf{W}$ . It is worth mentioning that the LMMSE still needs the matrix inversion of the term  $\mathbf{R}_c + \sigma_z^2 \mathbf{P} \mathbf{F}^H (\mathbf{V}^H \mathbf{V})^{-1} \mathbf{F} \mathbf{P}$ . Even though the computational complexity might be further reduced by taking advantage of its symmetric structure, LMMSE is still much more complex than MLE.

#### E. Phase Noise Mitigation

As discussed above, we have proposed two approaches to estimate phase noise parameters. In order to mitigate phase noise, we further present in this section two approaches for phase noise mitigation: decorrelation and ICI cancellation.

- 1) *Decorrelation:* Rewrite (9) as

$$\begin{aligned} \mathbf{y} &= \mathbf{C} \mathbf{w}_0 + \mathbf{z} \\ &= \mathbf{C} \mathbf{H} \mathbf{x} + \mathbf{z} \end{aligned} \quad (28)$$

with  $\mathbf{H} = \text{diag}(h(0), h(1), \dots, h(N-1))$ ,  $\mathbf{x} = [x(0), x(1), \dots, x(N-1)]^T$ , and accordingly

$$\mathbf{C} = \begin{pmatrix} c_0 & c_1 & \cdots & c_{N-1} \\ c_{N-1} & c_0 & \cdots & c_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & \cdots & c_{N-1} & c_0 \end{pmatrix} \quad (29)$$

Similar to  $\mathbf{W}$ ,  $\mathbf{C}$  is a circulant Matrix, each row of which has the identical elements as the previous row, but moved one position to the right and wrapped around. Then, with the estimated  $\check{c}_i, i = 0, \dots, N-1$ , from the above methods, we obtain, after phase noise mitigation and channel equalization, the estimated data

$$\hat{\mathbf{x}} = \mathbf{H}^{-1} \check{\mathbf{C}}^{-1} \mathbf{y} \quad (30)$$

The inversion of the  $N \times N$  matrix  $\check{\mathbf{C}}$  requires a computational complexity of  $O(N^3)$ , which might not be acceptable in many cases. We notice that the circulant matrix  $\mathbf{C}$  can be readily diagonalized by [22]

$$\mathbf{C} = \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \quad (31)$$

where  $\mathbf{F}$  denotes the DFT matrix defined by (11);  $\mathbf{\Lambda} = \text{diag} \left[ f(\mathbf{c}, 1), f\left(\mathbf{c}, e^{j\frac{2\pi}{N}}\right), \dots, f\left(\mathbf{c}, e^{j\frac{2\pi(N-1)}{N}}\right) \right]$  with the polynomial function  $f(\mathbf{c}, x)$  given in (A-7).

It's interesting to see that, in fact, the diagonal elements of  $\mathbf{\Lambda}$  are just the IDFT of  $\mathbf{c}$ , i.e.,  $\mathbf{\Lambda} = \sqrt{N} \text{diag}(\mathbf{F}^H \mathbf{c})$ .

Therefore, in terms of (31), (30) yields

$$\hat{\mathbf{x}} = \mathbf{H}^{-1} \mathbf{F}^H \check{\mathbf{\Lambda}}^{-1} \mathbf{F} \mathbf{y} \quad (32)$$

Equation (32) shows that  $\hat{\mathbf{x}}$  depends on the diagonal matrix operations and DFT transforms. Hence, it's quite straightforward to see that the computational complexity is greatly reduced.

2) *ICI and CPE Cancellation*: As shown in (21), the received data can be expressed as

$$\begin{aligned} \mathbf{y} &= \mathbf{w}_0 c_0 + \mathbf{W}^1 \mathbf{c}^1 + \mathbf{z} \\ &= c_0 \mathbf{H} \mathbf{x} + \mathbf{W}^1 \mathbf{c}^1 + \mathbf{z} \end{aligned} \quad (33)$$

where  $\mathbf{W}^1$  denotes the matrix  $\mathbf{W}$  without the first column vector. Then, the estimated data is given by

$$\hat{\mathbf{x}} = \mathbf{H}^{-1} \check{c}_0^{-1} [\mathbf{y} - \mathbf{W}^1 \check{\mathbf{c}}^1] \quad (34)$$

The ICI and CPE cancellation method reduces computational complexity even more than the decorrelator, due to the fact that it does not require a matrix inversion. Moreover, compared with the decorrelator, the ICI plus CPE canceller does not enhance AWGN noise. This makes it perform better in case of medium to low estimation error levels.

#### F. Practical Considerations

Both ML and LMMSE techniques proposed in this paper require *a priori* knowledge of the matrix  $\mathbf{W}$ , which is determined by the transmitted signals and the channel gains in the frequency domain. In other words, we have to know:

- 1) The channel frequency response;
- 2) All transmitted signals within an OFDM symbol.

In a high-rate OFDM system, the channel usually varies slightly within an OFDM symbol. Therefore, as in [2], all subcarriers of the preamble of each block are used for channel estimation. In practice, phase noise is limited to small to medium levels such that the major portion of phase noise energy is contained in the CPE, i.e., the CPE is dominant over the ICI. Since pilot-assisted channel estimation actually corrects the CPE in the preamble, and the effect of the ICI on the estimation accuracy is determined by the energy level of the ICI itself, these channel estimates are reliable enough to provide adequate information for our purpose [23], [24].

The knowledge of the transmitted signals can be replaced by the tentative data, i.e., the initial estimate of the data signals. These initial estimates can be obtained by applying one of the conventional CPE correction methods [15], [16], and channel

equalization to the received signals in the frequency domain. It is well known that such a decision feedback results in a better performance since the tentative data brings more information for better decisions. However, error propagation might occur with such a method, especially for low SNRs, since the first decision under low SNRs usually contains many errors, which might not improve the final decision or, sometimes, might even make it worse. In this case, our simulations show that the AWGN noise determines the performance, and that phase noise and its mitigation does not affect the performance significantly.

With the estimated channel and the tentative data applied to (14) or (19), phase noise is therefore mitigated using (32) or (34).

#### IV. PERFORMANCE ANALYSIS

In this section, we discuss the performance of different estimation methods on system performance. We focus on the frequency-domain approach since the extension of the analysis to the time domain is quite straightforward.

##### A. General Comparisons

- 1) It is well known that LMMSEE performs better than MLE by minimizing the mean square error of the estimation result. However, LMMSEE requires *a priori* knowledge of the phase noise linewidth  $\beta$  and the AWGN noise variance  $\sigma_z^2$ , which, however, are not necessary for MLE. In this regard, MLE may be a better choice than LMMSE.
- 2) Both methods involve a matrix inversion which requires the computations of the order of  $O(N^3)$ . Nevertheless, the mathematical methods described earlier can be used to significantly reduce the complexity. It is clear that LMMSEE is more complex than MLE since it requires extra computations for  $\mathbf{I} - \left[ \mathbf{I} + \sigma^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P} \right]^{-1}$ . We notice that, although the phase noise  $\phi_m(n)$  is nonstationary, the resulting random variable  $e^{j\phi_m(n)}$ , as shown in (5), is stationary, and so is  $\mathbf{c}$ . Hence, once we know the phase noise linewidth  $\beta$ , the DFT length  $N$  and the OFDM symbol period  $T$ ,  $\mathbf{R}_c$  can be precomputed and stored for the subsequent data processing.
- 3) For high signal to noise ratios (SNRs), i.e.,  $\sigma_z^2 \rightarrow 0$ , (19) reduces to

$$\left\{ \mathbf{I} - \left[ \mathbf{I} + \sigma^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P} \right]^{-1} \right\} \mathbf{W}^{-1} \mathbf{y} = \mathbf{W}^{-1} \mathbf{y} \quad (35)$$

which is the same as (14). This implies that, for high SNRs, the performance of MLE and LMMSE will be similar.

##### B. Computational Complexity

We first keep in mind that each complex multiplication is equivalent to four real multiplications plus two real additions, and each complex addition is equivalent to two real additions.

1) *MLE*: This technique requires the inversion of  $\mathbf{W}$ , whose computation is of the order of  $O(N^3)$ . By taking advantage of the properties of the circulant matrix, we can factorize  $\mathbf{W}$  into four matrices with one being diagonal and the remaining three being unitary (some are DFT matrices). Since the inversion of a diagonal matrix is very simple, and the inversion of unitary matrices in (23) is known, i.e.,  $\mathbf{P}^{-1} = \mathbf{P}$ ,  $\mathbf{F}^{-1} = \mathbf{F}^H$ , it is readily seen the significant reduction on the computations. More specifically, the following features are helpful in reducing complexity:

- 1)  $\mathbf{P}$  is a simple permutation matrix that doesn't require complex multiplications or additions;
- 2)  $\mathbf{F}^H$  or  $\mathbf{F}$  can be implemented using fast algorithm;
- 3)  $\mathbf{V}$  is a diagonal matrix whose elements are related to the DFT of vector  $\mathbf{w}_0$ , or, more specifically,  $\{a_k\}_{k=0}^{N-1}$ .

The approximate numbers of complex multiplications and complex additions for (24) are  $3N^2$  and  $3N^2 - 2N$ , respectively. Since  $\mathbf{F}$  is a DFT matrix and FFT algorithms are available, the complexity can be reduced to  $\frac{3N}{2} \log_2 N + N$  complex multiplications and  $3N \log_2 N$  complex additions, respectively.

2) *LMMSEE*: Equation (27) requires the additional computation of  $\mathbf{I} - \left[ \mathbf{I} + \sigma^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P} \right]^{-1}$  in comparison to (14).

- 1)  $\mathbf{F} \mathbf{P}$  involves permutation only.  $\mathbf{V} \mathbf{F} \mathbf{P}$  requires  $N^2$  complex multiplications and no additional computations for  $\mathbf{P} \mathbf{F}^H \mathbf{V}^H$ . Note also that the elements of  $\mathbf{V}$  have already been formed before computing  $\mathbf{P} \mathbf{F}^H \mathbf{V}^{-1} \mathbf{F} \mathbf{y}$ ;
- 2) The multiplication of  $\mathbf{P} \mathbf{F}^H \mathbf{V}^H$  with  $\mathbf{V} \mathbf{F} \mathbf{P}$ , when the FFT algorithm is applied, requires  $\frac{N^2}{2} \log_2 N$  complex multiplications and  $N^2 \log_2 N$  complex additions;
- 3) The multiplication of  $\sigma^{-2} \mathbf{R}_c$  with  $\mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P}$  requires  $N^3/2$  complex multiplications and  $N^3/2$  complex additions;
- 4) The addition of  $\mathbf{I}$  to  $\sigma^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P}$  requires  $N$  additions;
- 5) The inversion of  $\mathbf{R}_c (\mathbf{W}^H \mathbf{W}) + \sigma_z^2 \mathbf{I}$  requires  $N^3$  complex multiplications and  $N^3$  complex additions, respectively;
- 6) The addition of  $\mathbf{I}$  with  $\left[ \mathbf{I} + \sigma^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P} \right]^{-1}$  requires  $N$  complex additions;
- 7) The multiplication of  $\mathbf{I} - \left[ \mathbf{I} + \sigma^{-2} \mathbf{R}_c \mathbf{P} \mathbf{F}^H \mathbf{V}^H \mathbf{V} \mathbf{F} \mathbf{P} \right]^{-1}$  and  $\mathbf{W}^{-1} \mathbf{y}$  requires  $N^2$  complex multiplications and  $N^2$  complex additions, respectively.

In summary, LMMSE requires an additional  $(3N^3 + N^2 \log_2 N + 2N^2)/2$  complex multiplications and  $\frac{3N^3}{2} + N^2 \log_2 N + N^2 + 2N$  complex additions approximately, on top of  $\frac{3N}{2} \log_2 N + N$  complex multiplications and  $3N \log_2 N$  complex additions required by MLE. Hence, the complexity associated with the LMMSE is much larger in comparison to that of the MLE.

### C. Mean Square Error

For the two estimation methods, MLE and LMMSEE, it's readily understood that LMMSEE would yield better perfor-

mance. The mean square errors (MSEs) of the two methods can be obtained as in [21].

Assuming  $\hat{\mathbf{c}} = \mathbf{c} + \mathbf{e}$ , where  $\mathbf{e}$  denotes the Gaussian distributed estimation error with zero mean and variance  $\sigma_e^2$ ,  $\check{\mathbf{\Lambda}} = \text{diag} \left( f_{\check{c}}(1), f_{\check{c}} \left( e^{j \frac{2\pi}{N}} \right), \dots, f_{\check{c}} \left( e^{j \frac{2\pi(N-1)}{N}} \right) \right)$ , whereas  $\mathbf{\Lambda} = \text{diag} \left( f_c(1), f_c \left( e^{j \frac{2\pi}{N}} \right), \dots, f_c \left( e^{j \frac{2\pi(N-1)}{N}} \right) \right)$ . In the following, we compare the performance of different phase noise mitigation methods based on the estimation error.

1) *Decorrelator*: From (28) and (32), we can show that

$$\begin{aligned} & E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] \\ &= E \left[ \left\| \mathbf{x} - \mathbf{H}^{-1} \mathbf{F}^H \check{\mathbf{\Lambda}}^{-1} \mathbf{F} (\mathbf{C} \mathbf{H} \mathbf{x} + \mathbf{z}) \right\|^2 \right] \\ &= E \left[ \left\| \mathbf{H}^{-1} \mathbf{F}^H \mathbf{\Lambda}_B \mathbf{F} \mathbf{H} \mathbf{x} - \mathbf{H}^{-1} \mathbf{F}^H \check{\mathbf{\Lambda}}^{-1} \mathbf{F} \mathbf{z} \right\|^2 \right] \end{aligned} \quad (36)$$

where  $\mathbf{\Lambda}_B = \text{diag} \left( b_{\check{c}}(1), b_{\check{c}} \left( e^{j \frac{2\pi}{N}} \right), \dots, b_{\check{c}} \left( e^{j \frac{2\pi(N-1)}{N}} \right) \right)$  with  $b_{\check{c}}(x) = \frac{\sum_{i=0}^{N-1} e_i x^i}{\sum_{i=0}^{N-1} \check{c}_i x^i} = \frac{\sum_{i=0}^{N-1} e_i x^i}{\sum_{i=0}^{N-1} (c_i + e_i) x^i}$ . Equation (36) further gives

$$\begin{aligned} & E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] \\ &= E \left[ \left\| \mathbf{H}^{-1} \mathbf{F}^H \mathbf{\Lambda}_B \mathbf{F} \mathbf{H} \mathbf{x} - \mathbf{H}^{-1} \mathbf{F}^H \check{\mathbf{\Lambda}}^{-1} \mathbf{F} \mathbf{z} \right\|^2 \right] \\ &= \sigma_x^2 \cdot \text{tr} \left[ \mathbf{D}^{-1} \mathbf{\Lambda}_B \mathbf{D} \mathbf{D}^H \mathbf{\Lambda}_B^H (\mathbf{D}^{-1})^H \right] \\ &\quad + \sigma_z^2 \cdot \text{tr} \left[ \mathbf{H}^{-1} \left( \mathbf{F}^H (\check{\mathbf{\Lambda}} \check{\mathbf{\Lambda}}^H)^{-1} \mathbf{F} \right) (\mathbf{H}^{-1})^H \right] \\ &= \sigma_x^2 \cdot \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} \left| [\mathbf{D}^{-1} \mathbf{\Lambda}_B \mathbf{D}]_{ik} \right|^2 \\ &\quad + \sigma_z^2 \cdot \sum_{i=0}^{N-1} \left\{ \left| f_{\check{c}} \left( e^{j \frac{2\pi}{N} i} \right) \right|^{-2} \right\} \sum_{k=0}^{N-1} \left\{ |h_k|^{-2} \right\} \end{aligned} \quad (37)$$

where  $\text{tr}(\cdot)$  denotes the trace of a matrix;  $\mathbf{D}$  is defined as  $\mathbf{D} = \mathbf{F} \mathbf{H}$ .

2) *Interference Canceler*: From (9) and (34), we have

$$\begin{aligned} & E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] \\ &= E \left[ \left\| \mathbf{x} - \mathbf{H}^{-1} \check{c}_0^{-1} [\mathbf{W} \mathbf{c} + \mathbf{z} - \mathbf{W}^1 \check{c}^1] \right\|^2 \right] \\ &= E \left[ \left\| \check{c}_0^{-1} \mathbf{H}^{-1} \mathbf{E} \mathbf{H} \mathbf{x} - \check{c}_0^{-1} \mathbf{H}^{-1} \mathbf{z} \right\|^2 \right] \\ &= \sigma_x^2 |\check{c}_0^{-1}|^2 \cdot \text{tr} \left[ \mathbf{H}^{-1} \mathbf{E} \mathbf{H} \mathbf{H}^H \mathbf{E} (\mathbf{H}^{-1})^H \right] \\ &\quad + |\check{c}_0^{-1}|^2 \sigma_z^2 \cdot \text{tr} \left[ \mathbf{H}^{-1} (\mathbf{H}^{-1})^H \right] \end{aligned} \quad (38)$$

where matrix  $\mathbf{E}$  is a circulant matrix determined by the error vector  $\mathbf{e}$ , and has the same form as (29). Following the diagonalization method described in (31),  $\mathbf{E}$  can be written as

$$\mathbf{E} = \mathbf{F}^H \mathbf{\Lambda}_E \mathbf{F} \quad (39)$$

which, when substituted into (38), yields

$$\begin{aligned}
& E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] \\
&= |\check{c}_0^{-1}|^2 \sigma_x^2 \cdot \text{tr} \left[ \mathbf{D}^{-1} \mathbf{\Lambda}_E \mathbf{D} \mathbf{D}^H \mathbf{\Lambda}_E^H (\mathbf{D}^{-1})^H \right] \\
&\quad + |\check{c}_0^{-1}|^2 \sigma_z^2 \cdot \text{tr} \left[ \mathbf{H}^{-1} (\mathbf{H}^{-1})^H \right] \\
&= |\check{c}_0^{-1}|^2 \sigma_x^2 \cdot \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} \left| [\mathbf{D}^{-1} \mathbf{\Lambda}_E \mathbf{D}]_{ik} \right|^2 \\
&\quad + |\check{c}_0^{-1}|^2 \sigma_z^2 \cdot \sum_{k=0}^{N-1} \left\{ |h_k|^{-2} \right\} \tag{40}
\end{aligned}$$

Notice that when CPE accounts for the majority part of phase noise power, or  $\frac{|c_0|^2}{\sum_{i=0}^{N-1} |c_i|^2} \rightarrow 1$ , it can be readily seen that both MSEs in (40) and (37) can be approximated by

$$\begin{aligned}
E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] &= \sigma_x^2 \cdot \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} \left| [\mathbf{D}^{-1} \mathbf{\Lambda}_E \mathbf{D}]_{ik} \right|^2 \\
&+ \sigma_z^2 \cdot \sum_{k=0}^{N-1} \left\{ |h_k|^{-2} \right\} \tag{41}
\end{aligned}$$

The interference canceler may be more cost effective than the decorrelator when complexity is of concern since the decorrelator requires a matrix inversion. On the other hand, as the proposed phase noise mitigation methods are effective for small to moderate phase noise levels with variance less than  $10^{-1}$ , (41) holds for both methods, and the performance will be similar regardless of the method (decorrelator or interference canceler) that is used for phase noise mitigation.

## V. NUMERICAL RESULTS

The proposed approaches are evaluated in this section for Rayleigh fading channels by Monte Carlo trials. The OFDM data of each user are constructed based on the IEEE 802.11a standard, where part of the packet preamble is specifically designed for channel estimation. Channel coding is not used in our simulations since we focus on symbol error rate (SER) of an uncoded OFDM system and the SNR is denoted by the energy per sample (subcarrier signal). The phase noise variance is defined as  $2\pi\beta T/N$ , as in (3). 16QAM modulation is employed in the simulation unless otherwise specified. The cyclic prefix has the length larger than the channel delay spread throughout the simulations. The simulation parameters are summarized in Table I.

TABLE I  
SIMULATION PARAMETERS

Parameter	Value
bandwidth for each user	20 MHz
number of subcarriers	$N = 64$
number of pilots	$N_p = 4$
Modulation method	16QAM

Fig. 1-2 demonstrate the computational complexity needed for MLE and LMMSE respectively. For comparison, we have

added the performance of SCIC in [18]. We see from the figures that, in order to get the best performance from LMMSE, we need to trade the computational complexity. As a result, LMMSE has the highest complexity among all schemes, even slightly exceeding the conventional SCIC scheme in [18].

In Fig. 3, we plot the estimated real and imaginary parts of the CPE and the ICI energy obtained by averaging all the estimates from Monte Carlo trials. These estimates are compared with their actual values. The estimated CPE, including real and imaginary parts, and the ICI energy, are quite close to the theoretical values if the phase noise level is less than  $10^{-1}$ , and therefore, it is straightforward to see that phase noise mitigation is accurate within this range. For large phase noise with variance above  $10^{-1}$ , mitigation might not be very effective since the estimation errors become large. When the phase noise variance is of the order of  $10^{-2}$  or less, we observe that the imaginary part of CPE is very small compared with the real part which is close to unity. This explains the fact that, in some applications, when the phase noise is small, the CPE can be approximated as  $1 + j\varepsilon$ , where  $\varepsilon$  denotes the imaginary part of the CPE, and we only need to get its imaginary part to correct the CPE affecting the received signals.

Fig. 4-5 show the SER performance versus the SNR for both the MLE and LMMSE approaches when the phase noise variance is equal to  $10^{-2}$ . It is shown that the proposed approaches outperform the conventional CPEC technique for both the decorrelator and the interference canceler cases. For the LMMSE approach, a performance gain of 1-2 dB is observed in comparison to the MLE approach. This gain clearly comes from the fact that the LMMSE uses the phase noise statistics to minimize the overall estimation errors. These two figures also indicate that there is no error propagation when using the proposed methods. For sufficiently high SNRs, the proposed methods have better performance after using decision feedback. For low SNRs less than 10dB, the AWGN noise is dominant. The performance with or without phase noise is indistinguishable, and none of mitigation methods affect the performance significantly.

Fig. 6 illustrates the MSEs for both the MLE and the LMMSE approaches, and confirms the results shown in Fig. 4-5. It clearly demonstrates the better performance when the LMMSE is used. This is not a surprise since we are aware that LMMSE achieves minimum estimation error and therefore should have better mitigation results than MLE. Furthermore, we can hardly differentiate the performance between the decorrelator and the interference canceler. This is because, at the phase noise level of  $10^{-2}$ , as shown in Fig. 3, the CPE is quite dominant over the ICI in terms of energy. In this regard, the approximation of equation (41) holds for both the decorrelator and the interference canceler. Therefore, these two approaches should have a similar MSE performance. Due to this, if we compare Fig. 4 with Fig. 5, it is not hard to see that the SER performance for both methods is also comparable.

## VI. CONCLUSIONS

OFDM is quite sensitive to phase noise, which gives rise to common phase error as well as intercarrier interference, and

leads to performance loss. Many approaches in the literature mitigate phase noise either in the frequency domain by evaluating and then compensating for CPE or ICI, or in the time domain by directly correcting phase noise.

In the presence of phase noise, a new parametric model of OFDM signals was proposed in this paper. In terms of this model, at the receiver side, any frequency-domain subcarrier signal is described as a sum of all subcarrier signals weighted by a parameter vector. Both the maximum likelihood (ML) and the linear minimum mean square error (LMMSE) techniques were presented to estimate this weighting vector. Then phase noise was mitigated through different approaches based on the obtained estimates. It was shown that most conventional methods can be readily obtained from our approach with some approximation or an orthogonal transform. The data structure of the proposed model has further been analyzed to provide a much simpler method for implementation.

To compare the two proposed techniques, we presents a theoretical analysis which was further verified by computer simulations. We showed that, when used for phase noise estimation and mitigation, both the MLE and the LMMSE approaches result in superior performance over conventional methods. Furthermore, LMMSE gives the best performance of all methods, while ML provides a much simpler yet effective way to estimate and mitigate phase noise.

## REFERENCES

- [1] J. Bingham, "Multicarrier modulation for data transmission: an idea for whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.
- [2] IEEE Std 802.11a-1999, *Supplement to IEEE standard for information technology - telecommunications and information exchange between systems - local and metroplitan area networks - specific requirements. Part 11: wireless LAN medium access control (MAC) and physical layer (PHY) specifications: high-speed physical layer in the 5GHz band.* <http://www.ieee.org>, Dec. 1999.
- [3] ETSI TS 101 475 V1.3.1 (2001-12), *Broadband radio access networks (BRAN); HIPERLAN type 2; physical (PHY) layer.* <http://www.etsi.org>, Dec. 2001.
- [4] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, pp. 2908–2914, Oct. 1994.
- [5] H. Sari, G. Karam, and I. Jeanclaude, "Channel equalization and carrier synchronization in OFDM systems," in *1993 Tirrenia Int. Workshop on Digital Communications*, (Tirrenia, Italy), Sep. 1993.
- [6] M. A. Visser and Y. Bar-Ness, "OFDM frequency offset correction using an adaptive decorrelator," in *Proc. PIMRC'98*, (Boston, MA), pp. 816–820, Sep. 1998.
- [7] M. A. Visser, P. Zong, and Y. Bar-Ness, "A novel method for blind frequency offset correction in OFDM systems," in *Proc. CISS'32*, (Princeton, NJ), pp. 483–488, Mar. 1998.
- [8] T. Pollet, M. 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. 1995.
- [9] L. Tomba, "On the effect of wiener phase noise in OFDM systems," *IEEE Trans. Commun.*, vol. 46, pp. 580–583, May 1998.
- [10] A. G. Armada, "Understanding the effects of phase noise in orthogonal frequency division multiplexing (OFDM)," *IEEE Trans. Broadcast.*, vol. 47, pp. 153–159, Jun. 2001.
- [11] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Trans. Fundamental Theory and Applications.*, vol. 47, pp. 655–674, May 2000.
- [12] S. Wu and Y. Bar-Ness, "Performance analysis on the effect of phase noise in OFDM systems," in *Proc. ISSSTA'02*, (Prague, Czech), pp. 133–138, Sep. 2002.

- [13] M. S. El-Tanany, Y. Wu, and L. Hazy, "Analytical modeling and simulation of phase noise interference in OFDM-based digital television terrestrial broadcasting systems," *IEEE Trans. Broadcast.*, vol. 47, pp. 20–31, Mar. 2001.
- [14] R. A. Casas, S. L. Biracree, and A. E. Youtz, "Time domain phase noise correction for OFDM signals," *IEEE Trans. Broadcast.*, vol. 48, pp. 230–236, Sep. 2002.
- [15] P. Robertson and S. Kaiser, "Analysis of the effects of phase noise in orthogonal frequency division multiplexing (OFDM) systems," in *Proc. ICC'95*, (Seattle, WA), pp. 1652–1657, 1995.
- [16] S. Wu and Y. Bar-Ness, "A phase noise suppression algorithm for OFDM based WLANs," *IEEE Commun. Lett.*, vol. 6, pp. 535–537, Dec. 2002.
- [17] F. Munier, T. Eriksson, and A. Svensson, "Receiver algorithms for OFDM systems in phase noise and AWGN," in *Proc. PIMRC*, (Barcelona, Spain), pp. 1998–2002, Sep. 2004.
- [18] S. Wu and Y. Bar-Ness, "A new phase noise mitigation method in OFDM systems with simultaneous CPE and ICI correction," in *Proc. MCSS'03*, (Oberpfaffenhofen, Germany), Sep. 2003.
- [19] D. Petrovic, W. Rave, and G. Fettweis, "Phase noise suppression in OFDM including intercarrier interference," in *Proc. International OFDM Workshop*, (Hamburg, Germany), pp. 219–224, Sep. 2003.
- [20] J. G. Proakis, *Digital Communications*. McGraw-Hill Inc., 3rd ed., 1995.
- [21] S. C. Kay, *Fundamentals of statistical signal processing: Estimation Theory*. NJ, Prentice Hall, 1993.
- [22] P. J. Davis, *Circulant Matrices*. New York, John Wiley and Sons, 1979.
- [23] O. Edfors, M. Sandell, J. V. D. Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931–939, Jul. 1998.
- [24] Y. Li, L. J. C. Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, Jul. 1998.

## APPENDIX

As can be easily deduced from (9), the vector  $\mathbf{w}_0$  uniquely determines the shift-backward circulant matrix  $\mathbf{W}$  by

$$\mathbf{W} = \sum_{i=0}^{N-1} a_i \mathbf{\Gamma}_i \quad (\text{A-1})$$

where the matrix  $\mathbf{\Gamma}_i$  is defined by

$$\mathbf{\Gamma}_i = \begin{pmatrix} 0 & \cdots & 1_{(1,i+1)} & \cdots & 0 \\ \vdots & \diagdown & 0 & 0 & \vdots \\ 1_{(i+1,1)} & 0 & \diagup & 0 & 1_{(N-i-1,N)} \\ \vdots & 0 & 0 & \diagdown & \vdots \\ 0 & \cdots & 1_{(N,N-i-1)} & \cdots & 0 \end{pmatrix} \quad (\text{A-2})$$

with the positions of all non-zero entries marked by the subscripts. Specifically, for  $i = 0$ ,  $\mathbf{\Gamma}_i$  becomes the identity matrix, i.e.,  $\mathbf{\Gamma}_0 = \mathbf{I}$ . In view of the structure of the matrix  $\mathbf{\Gamma}_i$ , it is clear that

$$\mathbf{\Gamma}_i = \underbrace{\begin{pmatrix} 0 \cdots 0 & 0 & 1 \\ 1 & 0 & \diagdown \\ 0 & \diagdown & \diagdown \\ \vdots & \diagdown & 0 \\ 0 \cdots 0 & 1 & 0 \end{pmatrix}}_{\mathbf{A}} \mathbf{A}^{i+1} \underbrace{\begin{pmatrix} 0 \cdots 0 & 1 \\ \vdots & \diagdown \\ 0 & \diagdown \\ 1 & 0 \cdots 0 \end{pmatrix}}_{\mathbf{P}} \quad (\text{A-3})$$

Substituting (A-3) into (A-1) yields

$$\mathbf{W} = \left( \sum_{i=0}^{N-1} a_i \mathbf{A}^{i+1} \right) \mathbf{P} \quad (\text{A-4})$$

$\mathbf{A}$  is a unitary matrix that performs the backward-shift permutation.  $\mathbf{A}$  can be diagonalized by

$$\mathbf{A} = \mathbf{F}^H \mathbf{U} \mathbf{F} \quad (\text{A-5})$$

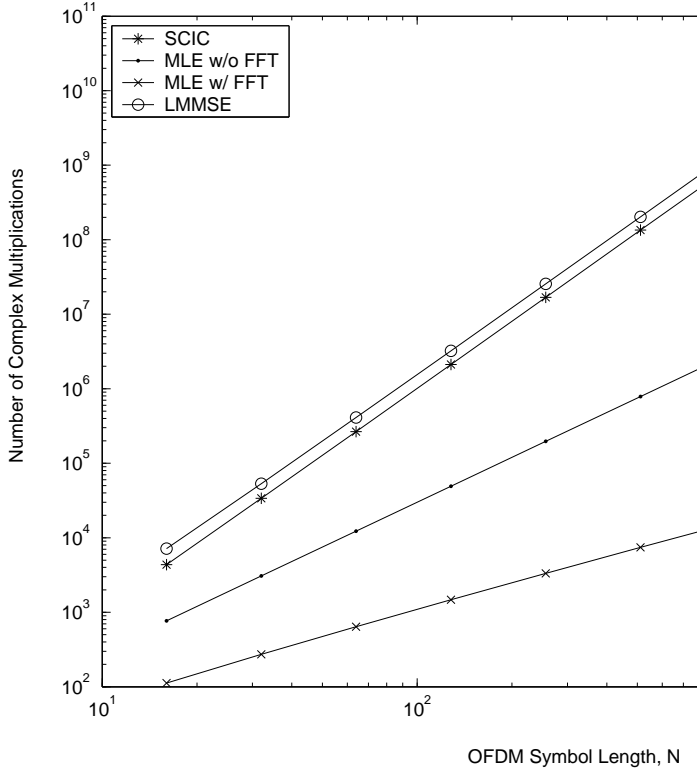


Fig. 1. Numbers of complex multiplications for different phase noise estimation approaches



Fig. 2. Numbers of complex additions for different phase noise estimation approaches

where  $\mathbf{F}$  is the DFT matrix defined by (11) and  $\mathbf{U} = \text{diag}\left(1, e^{j\frac{2\pi}{N}}, \dots, e^{j\frac{2\pi(N-1)}{N}}\right)$ . Substituting (A-5) into (A-4) gives

$$\mathbf{W} = \mathbf{F}^H \mathbf{V} \mathbf{F} \mathbf{P} \quad (\text{A-6})$$

Let  $\mathbf{y} = [y_0, y_1, \dots, y_{N-1}]^T$ . We define the polynomial function

$$f(\mathbf{y}, x) = \sum_{i=0}^{N-1} y_i x^i \quad (\text{A-7})$$

Then, the matrix  $\mathbf{V}$  in (A-6) can be expressed as

$$\mathbf{V} = \text{diag}[g_0, g_1, \dots, g_{N-1}] \quad (\text{A-8})$$

where  $g_k = e^{-j\frac{2\pi}{N}k} f_{\mathbf{w}_0}\left(e^{-j\frac{2\pi}{N}k}\right)$ . Equivalently, (A-8) can then be expressed in a matrix form

$$\mathbf{V} = \sqrt{N} \text{diag}(\mathbf{U} \mathbf{F} \mathbf{w}_0) \quad (\text{A-9})$$

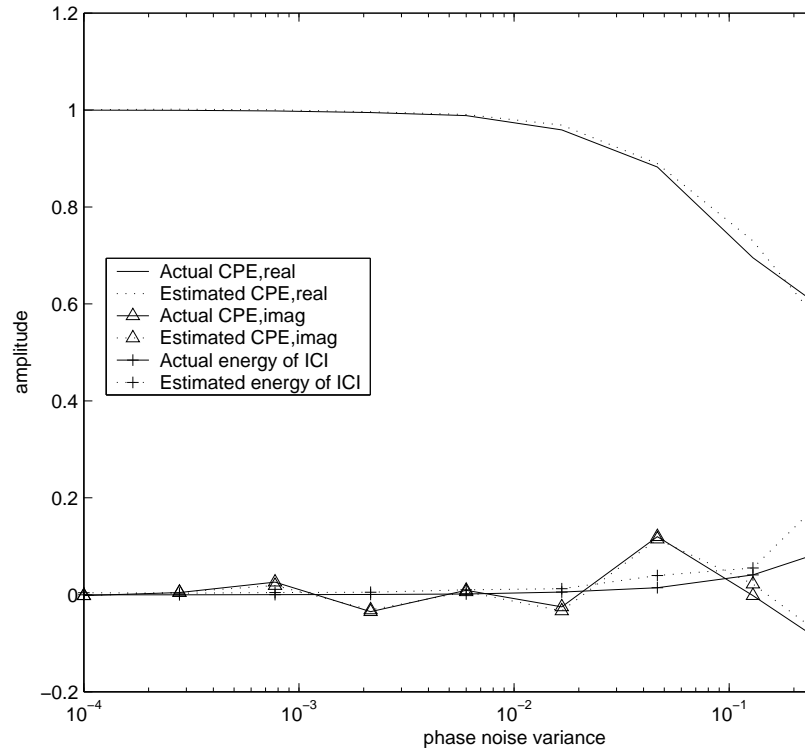


Fig. 3. Estimated CPE and ICI compared with their actual values

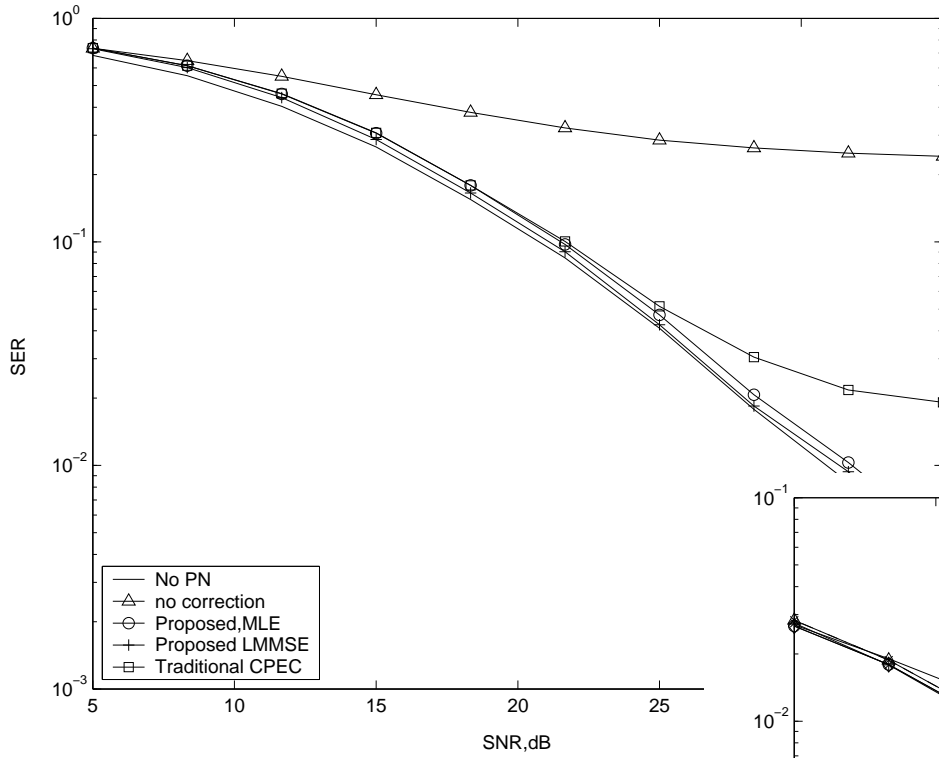


Fig. 4. The SER performance of the proposed generalized approaches using decorrelator versus the conventional CPEC approach, where the phase noise variance equals  $10^{-2}$ , and the number of pilots equals 4.

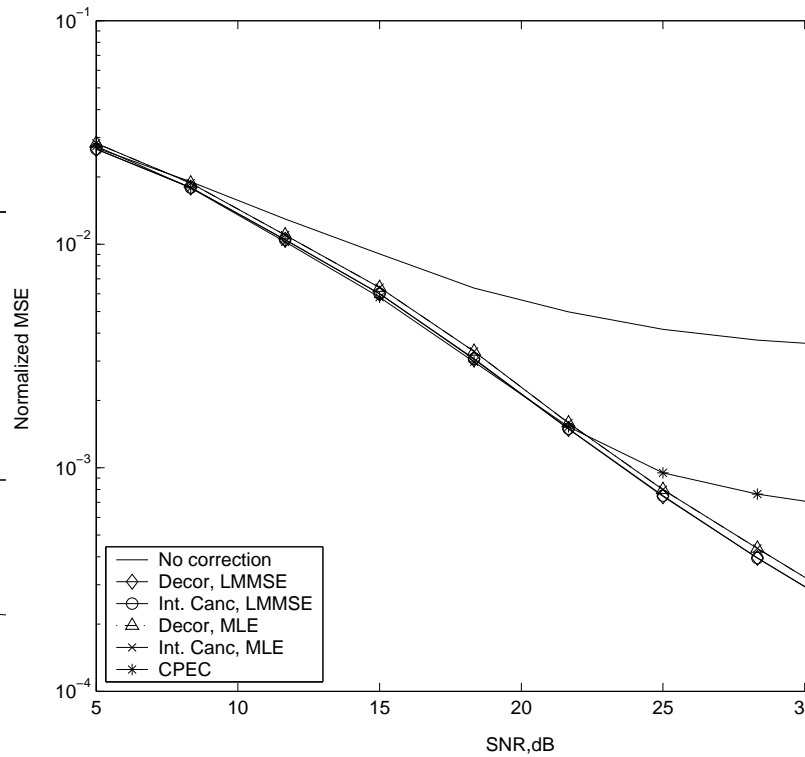


Fig. 6. Mean square error of LMMSE and MLE, using interference canceler, with the phase noise variance of  $10^{-2}$ .

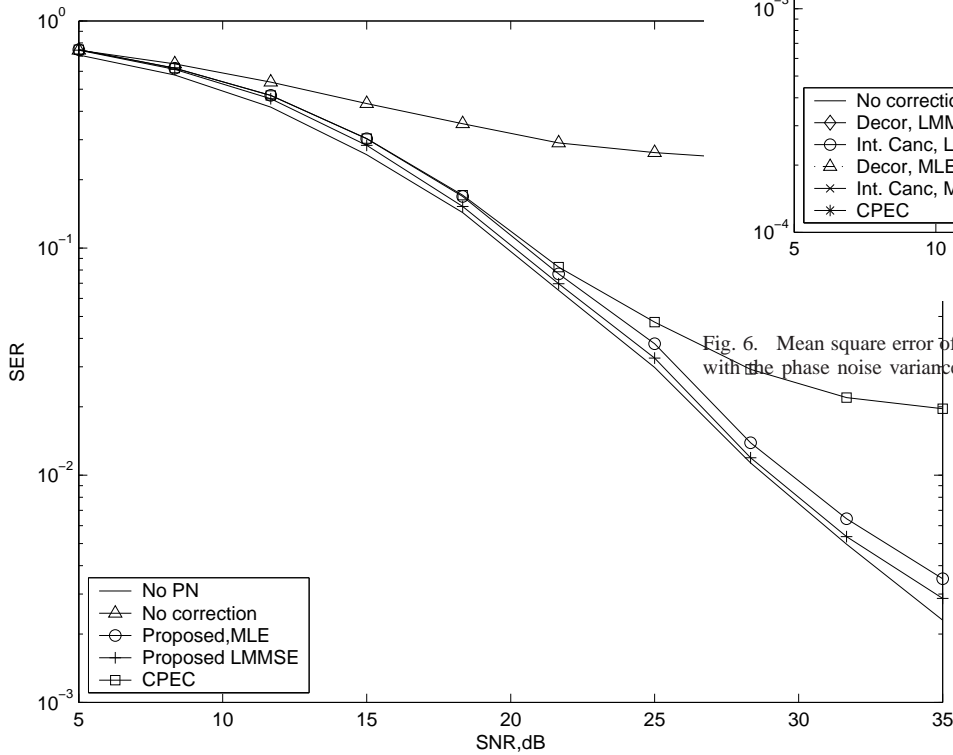


Fig. 5. The SER performance of the proposed generalized approaches using interference canceler versus the conventional CPEC approach, where the phase noise variance equals  $10^{-2}$ , and the number of pilots equals 4.