Transactions Papers

Joint CFO and Channel Estimation for OFDM-Based Two-Way Relay Networks

Gongpu Wang, Feifei Gao, Yik-Chung Wu, and Chintha Tellambura

Abstract—**Joint estimation of the carrier frequency offset (CFO) and the channel is developed for a two-way relay network (TWRN) that comprises two source terminals and an** *amplify-and-forward* **(AF) relay. The terminals use orthogonal frequency division multiplexing (OFDM). New zero-padding (ZP) and cyclic-prefix (CP) transmission protocols, which maintain the carrier orthogonality and ensure low estimation and detection complexity, are proposed. Both protocols lead to the same estimation problem which can be solved by the nulling-based least square (LS) algorithm and perform identically when the block length is large. We present detailed performance analysis by proving the unbiasedness of the LS estimators at high signal-to-noise ratio (SNR) and by deriving the closed-form expression of the mean-square-error (MSE). Simulation results are provided to corroborate our findings.**

Index Terms—Carrier frequency offset, channel estimation, two-way relay network, OFDM, performance analysis.

I. INTRODUCTION

RELAY network has become an active research field since the pioneer work [1], which developed low-complexity cooperative diversity strategies. In [1], data flows unidirectionally from the source to the relay and then to the destination. It is an example of one-way relay network (OWRN). However, for bidirectional flows, the source and destination exchange their roles dynamically. They both behave as source terminals, resulting in a two-way relay network (TWRN). By utilizing a "network coding"-like setup [2], the relay can help signal recovery at both source terminals.

Reference [3] reported that the overall communication rate between two source terminals in TWRN is approximately twice that achieved in OWRN, making TWRN particularly attractive to bidirectional systems. The capacity analysis and

the achievable rate region for *amplify-and-forward* (AF) and *decode-and-forward* (DF) based TWRN were explored in [4], [5]. In [6], the optimal relay mapping function that minimizes the bit-error rate (BER) was proposed. In [7], distributed space-time codes (STC) were designed for both AF and DF TWRN. Moreover, the optimal beamforming at the multi-antenna relay that maximizes the capacity of AF-based TWRN was developed in [8], and suboptimal resource allocation for orthogonal frequency division multiplexing (OFDM) based TWRN was derived in [9].

Most previous works [3]–[9] assume perfect channel state information (CSI) at the relay and/or the source terminals. But CSI estimation in TWRN is a complicated issue. Although conventional estimation methods are effective for DF-based TWRN, they do not work for AF based TWRN. For example, channel estimation for TWRN was studied in [10], [11] for frequency-flat and frequency-selective environments, respectively. These studies showed that AF TWRN systems require very different estimation techniques from conventional point-to-point systems.

Moreover, frequency mismatches that cause carrier frequency offset (CFO) require compensation before data detection [12]. The problem of CFO estimation in TWRN is even more difficult because one needs to cope with the mismatch between two source terminals as well as that between the source terminals and the relay. The estimation of CFO and channels with low complexity is, therefore, a challenging task. Moreover, maintaining carrier orthogonality in order to facilitate data detection is critically important.

Our earlier study of joint CFO and channel estimation in TWRN [13] considered general frequency selective channels. In this paper, we take a further step by analyzing OFDM TWRN. Our contributions are summarized as follows:

- Cyclic-prefix (CP)-based and zero-padding (ZP)-based OFDM protocols are proposed for TWRN to facilitate the joint CFO and channel estimation as well as data detection;
- Nulling-based least square (LS) algorithm is developed for the joint estimation problem.
- Detailed performance analysis is provided for the nulling-based LS estimator: (i) proving its unbiasedness at the
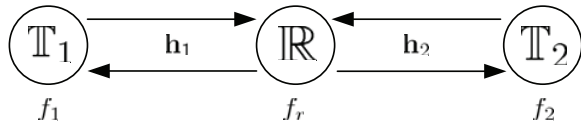
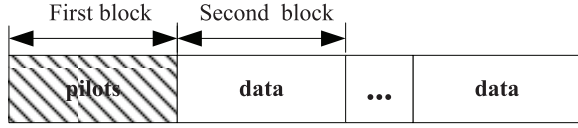Fig. 1.   System configuration for a three-node TWRN.

Fig. 2.   One data frame that contains training symbols in the first OFDM block and information symbols in the rest blocks.

high signal-to-noise ratio (SNR); (ii) deriving the mean-square-errors (MSEs) in closed-form;

- The numerical examples that verify our theoretical results are presented.

The rest of this paper is organized as follows. Section II presents ZP-based and CP-based OFDM protocols. Section III discusses the LS estimator and its performance analysis. Simulation results are provided in Section IV. Finally, conclusions are drawn in Section V.

Notations: Vectors and matrices are typed in boldface letters; the transpose, complex conjugate, Hermitian, and inverse of $\mathbf{A}$ are $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$, and $\mathbf{A}^{-1}$, respectively; $\Re\{\mathbf{A}\}$ and $\Im\{\mathbf{A}\}$ are the real and the imaginary parts of $\mathbf{A}$; $\dot{\mathbf{A}}$ and $\ddot{\mathbf{A}}$ represent the first-order and the second-order derivatives of $\mathbf{A}$; diag$\{\mathbf{a}\}$ is the diagonal matrix formed by $\mathbf{a}$, and $\mathbf{I}$ is the identity matrix; E$\{\cdot\}$ is the statistical expectation, and $\jmath = \sqrt{-1}$ is the imaginary unit.

II. PROBLEM FORMULATION

A. Two-Way Transmission Model

Consider a TWRN with two source nodes $\mathbb{T}_1$ and $\mathbb{T}_2$, and one relay node $\mathbb{R}$ (Fig. 1). Each node has only one half-duplex antenna. The baseband channel between $\mathbb{T}_i$, $i = 1, 2$ and $\mathbb{R}$ is denoted by $\mathbf{h}_i = [h_{i,0}, h_{i,1}, \ldots, h_{i,L}]^T$, where L is the order of the corresponding channel.[1] The elements in $\mathbf{h}_i$ are assumed to be independently zero-mean circularly symmetric complex Gaussian (CSCG) random variables. The variance of the lth element in $\mathbf{h}_i$ is denoted by $\sigma_{i,l}^2$. Following [2]– [9], we assume time-division-duplexing (TDD), which leads to reciprocal channels, i.e., the channel from $\mathbb{R}$ to $\mathbb{T}_i$ is $\mathbf{h}_i$ as well. The proposed strategies can be straightforwardly extended to more general situations including non-reciprocal channels. The average transmission powers of $\mathbb{T}_1$, $\mathbb{T}_2$, and $\mathbb{R}$ are denoted as P_1, P_2, and P_r, respectively. We further denote the oscillator frequencies as f_1, f_2, and f_r, respectively. Moreover, perfect time synchronization is assumed [12], [21]. The data frame structure is shown in Fig. 2, where the first block is devoted for training while the remaining blocks are used for information transmission.

In [11], OFDM with CP of length L was adapted to TWRN under perfect time and frequency synchronization.

These perfect conditions ensure that circular convolution between two frequency selective channels, i.e., $\mathbf{h}_1$ and $\mathbf{h}_2$, in time are equivalently converted to multiplication between two flat fading channels over different subcarriers. However, with CFOs, the inter-carrier interference (ICI) will be introduced in the frequency domain at different nodes. The joint estimation problem then becomes much complicated if the same OFDM based TWRN strategy is applied [11]. Even with perfect CSI, the data detection is complicated.

In this work, we adapt the conventional CP and ZP- based OFDM to TWRN and design two new transmission protocols for OFDM TWRN systems under the non-zero CFOs. Although they incur redundancy, these protocols facilitate channel estimation, frequency synchronization, as well as data detection. Our main contribution is the introduction of these protocols and their performance analysis.

B. ZP-based OFDM for TWRN

1) OFDM at terminals: We denote the mth OFDM block from $\mathbb{T}_i$ as $\tilde{\mathbf{s}}_i(m) = [\tilde{s}_{i,0}(m), \ldots, \tilde{s}_{i,N-1}(m)]^T$, where N is the block length. For conciseness, we omit the block index m unless otherwise mentioned. The corresponding time-domain signal block is obtained from the normalized inverse discrete Fourier transformation (IDFT) as

$$\mathbf{s}_i = \mathbf{F}^H \tilde{\mathbf{s}}_i = [s_{i,0}, s_{i,1}, \ldots, s_{i,N-1}]^T, \qquad (1)$$

where $\mathbf{F}$ is the normalized DFT matrix with the (p,q)-th entry given by $\frac{1}{\sqrt{N}} e^{-\jmath 2\pi(p-1)(q-1)/N}$. To avoid the inter-block interference (IBI) in the first transmission phase, at least L zeros must be padded at the end of $\mathbf{s}_i$. However, we propose to add $2L$ zeros in order to keep the phase continuity in the following transmission blocks. This point will be further explained later. As in the traditional ZP-OFDM system [14], the power of $\tilde{s}_{i,n}$, $n = 0, \ldots, N-1$ is $\frac{N+2L}{N} P_i$, so the average power constraint P_i is kept at $\mathbb{T}_i$.[2]

In Phase I, $\mathbb{T}_1$ and $\mathbb{T}_2$ up-convert the baseband signals by the carriers $e^{\jmath 2\pi f_i t}$ and send them to $\mathbb{R}$ simultaneously.[3]

2) Relay processing: The relay $\mathbb{R}$ down-converts the pass-band signal by $e^{-\jmath 2\pi f_r t}$. Of the mth OFDM block, the last L symbols are purely noise while the first $N + L$ symbols can be expressed as

$$\mathbf{r}_{zp} = \sum_{i=1}^{2} e^{\jmath 2\pi(m-1)(N+2L)(f_i-f_r)T_s} \mathbf{\Gamma}^{(N+L)}[f_i - f_r]$$
$$\times \mathbf{H}_{zp}^{(N)}[\mathbf{h}_i]\mathbf{s}_i + \mathbf{n}_r, \quad (2)$$

where

$$\mathbf{\Gamma}^{(K)}[f] = \text{diag}\{1, e^{\jmath 2\pi f T_s}, \ldots, e^{\jmath 2\pi f(K-1)T_s}\} \qquad (3)$$

---

[1] We assume the same channel length for $\mathbf{h}_i$ for the sake of notational simplicity. Our discussion can be straightforwardly extended to the more general case.

[2] This power scaling is also necessary to make a fair comparison with CP-OFDM as did in the point-to-point system [15].

[3] Note that the oscillator may have an initial phase, but it is omitted for brevity since the constant phase can be absorbed into the channel effects.

$$
\begin{aligned}
\mathbf{y}_{zp} =& \alpha_{zp} e^{j2\pi(m-1)(f_r-f_1)(N+2L)T_s} \mathbf{\Gamma}^{(N+2L)}[f_r - f_1] \mathbf{H}_{zp}^{(N+L)}[\mathbf{h}_1] \mathbf{r}_{zp} \\
=& \alpha_{zp} \mathbf{\Gamma}^{(N+2L)}[f_r - f_1] \mathbf{H}_{zp}^{(N+L)}[\mathbf{h}_1] \sum_{i=1}^{2} \mathbf{\Gamma}^{(N+L)}[f_i - f_r] \mathbf{H}_{zp}^{(N)}[\mathbf{h}_i] \mathbf{s}_i \Big) \\
& + \underbrace{\alpha_{zp} e^{j2\pi(m-1)(N+2L)(f_2-f_1)T_s} \mathbf{\Gamma}^{(N+2L)}[f_r - f_1] \mathbf{H}_{zp}^{(N+L)}[\mathbf{h}_1] \mathbf{n}_r + \mathbf{n}_1}_{\mathbf{n}_e}, \quad (6)
\end{aligned}
$$

with T_s representing the sampling period, and

$$
\mathbf{H}_{zp}^{(K)}[\mathbf{x}] \triangleq \underbrace{\begin{bmatrix} x_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ x_L & \ddots & x_0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & x_L \end{bmatrix}}_{K \text{ columns}} \quad (4)
$$

for any vector $\mathbf{x} = [x_0, x_1, \dots, x_L]^T$. Moreover, $\mathbf{n}_r$ is the $(N + L) \times 1$ noise vector, each entry with variance σ_n^2.

Remark 1: Though only the first L ZP from $\mathbb{T}_i$ contribute to $\mathbf{r}_{zp}$, the last L ZP is still necessary in that the accumulated CFO of the mth OFDM block starts from $(m - 1)(N + 2L)(f_i - f_r)$. This value is necessary to keep the phase continuity as will be seen later.

Then, $\mathbb{R}$ add L zeros at the end of $\mathbf{r}_{zp}$ and scales the whole block of $(N + 2L)$ by the factor

$$
\begin{aligned}
\alpha_{zp} =& \sqrt{\frac{(N + 2L)P_r}{\mathrm{E}\{\|\mathbf{r}_{zp}\|^2\}}} \\
=& \sqrt{\frac{N + 2L}{N + L} \cdot \frac{P_r}{\sum_{i=1}^{2} \sum_{l=0}^{L} \sigma_{i,l}^2 P_i + \sigma_n^2}}
\end{aligned} \quad (5)
$$

to keep the average relay power constraint. The resultant signal blocks will be up-converted to passband by $e^{j2\pi f_r t}$.

3) Signal reformulation at terminals: Due to symmetry, we only study $\mathbb{T}_1$ during the second phase. After the down-conversion of the passband signal by $e^{-j2\pi f_1 t}$, the mth OFDM block is expressed as[4] (6) on the top of this page, where $\mathbf{n}_1$ is the $(N + 2L) \times 1$ noise vector at $\mathbb{T}_1$ with the variance σ_n^2, and $\mathbf{n}_e$ defines the overall noise component. The covariance of $\mathbf{n}_e$ is computed as

$$
\begin{aligned}
\mathbf{R}_{zp} =& \sigma_n^2 (\alpha_{zp}^2 \mathbf{\Gamma}^{(N+2L)}[f_r - f_1] \mathbf{H}_{zp}^{(N+L)}[\mathbf{h}_1] \\
& \times (\mathbf{H}_{zp}^{(N+L)}[\mathbf{h}_1])^H (\mathbf{\Gamma}^{(N+2L)}[f_r - f_1])^H + \mathbf{I}). \quad (7)
\end{aligned}
$$

In most practical communications,[5] there is $N \gg L$ and $\mathbf{R}_{zp}$ can be approximated by its expectation over $\mathbf{h}_1$:

$$
\mathbf{R}_{zp} \approx \underbrace{\sigma_n^2 \left(\alpha_{zp}^2 \sum_{l=0}^{L} \sigma_{h_{1,l}}^2 + 1 \right)}_{\sigma_{ne}^2} \mathbf{I}, \quad (8)
$$

where σ_{ne}^2 denotes the equivalent noise variance.

Lemma 1: The following two equalities hold for any $\mathbf{\Gamma}^{(\cdot)}[f]$ in (3) and $\mathbf{H}_{zp}^{(\cdot)}[\mathbf{x}]$ in (4), where $(\cdot)$ represents the appropriate dimensions:

$$
\mathbf{H}_{zp}^{(N)}[\mathbf{x}] \mathbf{\Gamma}^{(N)}[f] = \mathbf{\Gamma}^{(N+L)}[f] \mathbf{H}_{zp}^{(N)} \left[\mathbf{\Gamma}^{(L+1)}[-f]\mathbf{x} \right], \quad (9)
$$

and conversely

$$
\mathbf{\Gamma}^{(N+L)}[f] \mathbf{H}_{zp}^{(N)}[\mathbf{x}] = \mathbf{H}_{zp}^{(N)} \left[\mathbf{\Gamma}^{(L+1)}[f]\mathbf{x} \right] \mathbf{\Gamma}^{(N)}[f]. \quad (10)
$$

Proof: Proved from the straightforward manipulations. ∎

Lemma 1 says that $\mathbf{\Gamma}^{(\cdot)}[f]$ can be switched from the right (left) side of $\mathbf{H}_{zp}^{(\cdot)}[\mathbf{h}_i]$ to the left (right) side by changing the dimension of $\mathbf{\Gamma}^{(\cdot)}[f]$ and rotating $\mathbf{h}_i$.

From Lemma 1, $\mathbf{y}_{zp}$ can be rewritten as

$$
\begin{aligned}
\mathbf{y}_{zp} =& \alpha_{zp} \mathbf{H}_{zp}^{(N+L)}[\mathbf{\Gamma}^{(L+1)}[f_r - f_1]\mathbf{h}_1] \mathbf{H}_{zp}^{(N)}[\mathbf{h}_1] \mathbf{s}_1 \\
& + \alpha_{zp} e^{j2\pi(m-1)(N+2L)(f_2-f_1)T_s} \mathbf{\Gamma}^{(N+2L)}[f_2 - f_1] \\
& \times \mathbf{H}_{zp}^{(N+L)}[\mathbf{\Gamma}^{(L+1)}[f_r - f_2]\mathbf{h}_1] \mathbf{H}_{zp}^{(N)}[\mathbf{h}_2] \mathbf{s}_2 + \mathbf{n}_e.
\end{aligned} \quad (11)
$$

We further note that

$$
\mathbf{H}_{zp}^{(N+L)}[\mathbf{x}_1] \mathbf{H}_{zp}^{(N)}[\mathbf{x}_2] = \mathbf{H}_{zp}^{(N)}[\mathbf{x}_1 \otimes \mathbf{x}_2], \quad (12)
$$

where $\otimes$ denotes the linear convolution between the two vectors. Hence $\mathbf{y}_{zp}$ can be written as

$$
\begin{aligned}
\mathbf{y}_{zp} =& \alpha_{zp} \mathbf{H}_{zp}^{(N)} \Big[\underbrace{(\mathbf{\Gamma}^{(L+1)}[f_r - f_1]\mathbf{h}_1) \otimes \mathbf{h}_1}_{\mathbf{a}_{zp}} \Big] \mathbf{s}_1 \\
& + \alpha_{zp} e^{j2\pi(m-1)(N+2L)(f_2-f_1)T_s} \mathbf{\Gamma}^{(N+2L)}[\underbrace{f_2 - f_1}_{v}] \\
& \times \mathbf{H}_{zp}^{(N)} \Big[\underbrace{(\mathbf{\Gamma}^{(L+1)}[f_r - f_2]\mathbf{h}_1) \otimes \mathbf{h}_2}_{\mathbf{b}_{zp}} \Big] \mathbf{s}_2 + \mathbf{n}_e, \quad (13)
\end{aligned}
$$

where $\mathbf{a}_{zp}$, $\mathbf{b}_{zp}$ are the $(2L+1) \times 1$ equivalent channel vectors and v is the equivalent CFO.

Remark 2: With the insertion of $2L$ ZP at $\mathbb{T}_i$, only the unique *effective* CFO $v = f_2 - f_1$ will be involved in the received signal block (13).

4) Data detection at terminals: If the cascaded channel $\mathbf{a}_{zp}$ is estimated at $\mathbb{T}_1$, then the first term on the right-hand side (RHS) of (13) can be removed since $\mathbb{T}_1$ knows its own signal $\mathbf{s}_1$.[6] If the CFO v is also known, then $e^{j2\pi(m-1)(N+2L)vT_s} \mathbf{\Gamma}^{(N+2L)}[v]$ can be compensated for, and the remaining signal is

$$
\mathbf{z}_{zp} = \alpha_{zp} \mathbf{H}_{zp}^{(N)}[\mathbf{b}_{zp}] \mathbf{s}_2 + \mathbf{n}_e. \quad (14)
$$

---

[4]In fact, the down-conversion carrier is $e^{-j2\pi f_1(t-\tau)}$, where τ includes the overall time delay. Nonetheless, the non-zero delay only causes a constant phase rotation and can be absorbed into the channel estimation. For notation simplicity, we omitted τ in our paper.

[5]In IEEE 802.11a standards [16], $N \geq 4L$ is adopted.

[6]Since α_{zp} does not depend on the instant channel knowledge, one can always assume it is known at $\mathbb{T}_1$ from past statistics or from a scalar feedback.

$$\mathbf{r}_{cp} = \sum_{i=1}^{2} e^{j2\pi(f_i - f_r)[(m-1)(N+2L)+L]T_s} \mathbf{\Gamma}^{(N+L)}[f_i - f_r]\mathbf{H}_{cv}^{(N+L)}[\mathbf{h}_i]\mathbf{T}_{cp}^{(2L)}\mathbf{s}_i + \mathbf{n}_r$$

$$= \sum_{i=1}^{2} e^{j2\pi(f_i - f_r)[(m-1)(N+2L)+L]T_s} \mathbf{\Gamma}^{(N+L)}[f_i - f_r]\mathbf{T}_{cp}^{(L)}\mathbf{H}_{cp}^{(N)}[\mathbf{h}_i]\mathbf{s}_i + \mathbf{n}_r. \tag{19}$$

$$\mathbf{y}_{cp} = \alpha_{cp}e^{j2\pi(f_r - f_1)[(m-1)(N+2L)+2L]T_s}\mathbf{\Gamma}^{(N)}[f_r - f_1]\mathbf{H}_{cv}^{(N)}[\mathbf{h}_1]\mathbf{r}_{cp}$$

$$= \alpha_{cp}e^{j2\pi(f_r - f_1)LT_s}\mathbf{\Gamma}^{(N)}[f_r - f_1]\mathbf{H}_{cv}^{(N)}[\mathbf{h}_1]\mathbf{\Gamma}^{(N+L)}[f_1 - f_r]\mathbf{T}_{cp}^{(L)}\mathbf{H}_{cp}^{(N)}[\mathbf{h}_1]\mathbf{s}_1$$

$$+ \alpha_{cp}e^{j2\pi(f_r - f_1)LT_s}e^{j2\pi(f_2 - f_1)[(m-1)(N+2L)+L]T_s}\mathbf{\Gamma}^{(N)}[f_r - f_1]\mathbf{H}_{cv}^{(N)}[\mathbf{h}_1]\mathbf{\Gamma}^{(N+L)}[f_2 - f_r]\mathbf{T}_{cp}^{(L)}\mathbf{H}_{cp}^{(N)}[\mathbf{h}_2]\mathbf{s}_2$$

$$+ \underbrace{\alpha_{cp}e^{j2\pi(f_r - f_1)[(m-1)(N+2L)+2L]T_s}\mathbf{\Gamma}^{(N)}[f_r - f_1]\mathbf{H}_{cv}^{(N)}[\mathbf{h}_1]\mathbf{n}_r + \mathbf{n}_1}_{\mathbf{n}_e}. \tag{24}$$

Moreover, since $\mathbf{H}_{zp}^{(N)}[\mathbf{b}_{zp}]$ is the $(N+2L) \times N$ Toeplitz matrix following the structure in (4), we can apply the same approach in the conventional ZP-based OFDM system [14] to efficiently detect the desired signal. Adding the last $2L$ elements of $\mathbf{z}_{zp}$ to its first $2L$ elements, we obtain

$$\mathbf{w}_{zp} = \alpha_{zp}\mathbf{H}_{cp}^{(N)}[\mathbf{b}_{zp}]\mathbf{s}_2 + \tilde{\mathbf{n}}_e, \tag{15}$$

where $\mathbf{H}_{cp}^{(N)}[\mathbf{b}_{zp}]$ is the $N \times N$ circulant matrix with the first column $[\mathbf{b}_{zp}^T, \mathbf{0}_{1\times(N-2L-1)}]^T$, and $\tilde{\mathbf{n}}_e$ is the resultant noise vector. Implicitly, we need $N \geq 2L + 1$ in (15), which is normally satisfied by practical systems [16]. As long as $\mathbf{b}_{zp}$ has been estimated, conventional OFDM detection can be performed. Note that the effective fading coefficient on each subcarrier is the DFT of $\mathbf{b}_{zp}$, a combination of CFO $f_r - f_2$ and channels $\mathbf{h}_1$, $\mathbf{h}_2$.

5) Joint CFO and channel Estimation : The above discussion shows that the task is to estimate $\mathbf{a}_{zp}$, $\mathbf{b}_{zp}$, and v. For the first training block, we can rewrite (13) as

$$\mathbf{y}_{zp} = \mathbf{S}_1^{(N+2L)}\mathbf{a}_{zp} + \mathbf{\Gamma}^{(N+2L)}[v]\mathbf{S}_2^{(N+2L)}\mathbf{b}_{zp} + \mathbf{n}_e, \tag{16}$$

where $\mathbf{S}_i^{(N+2L)}$ is the $(N+2L) \times (2L+1)$ circulant matrix with the first column $[\alpha_{zp}\mathbf{s}_i^T, \mathbf{0}_{1\times 2L}^T]^T$. Obviously, (16) is different from the conventional work [12] in that only a part of the signal component is accompanied with the CFO matrix. Moreover, $N \geq 2L + 3$ is required to estimate all the unknown parameters. The detailed estimation algorithms will be presented in the next section.

C. CP-based OFDM for TWRN

1) OFDM modulation at terminals: Each terminal $\mathbb{T}_i$ first obtains the time domain OFDM signal $\mathbf{s}_i$ from its frequency domain information block $\tilde{\mathbf{s}}_i$ via the normalized IDFT approach. We propose that $\mathbb{T}_i$ adds a CP of length $2L$ in the front of $\mathbf{s}_i$. Define

$$\mathbf{T}_{cp}^{(P)} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_P \\ \hline \mathbf{I}_N \end{bmatrix}, \tag{17}$$

for any $P \leq N$. Then the baseband signal sent out from $\mathbb{T}_i$ is mathematically expressed as $\mathbf{T}_{cp}^{(2L)}\mathbf{s}_i$, which is up-converted to the passband signal by $e^{j2\pi f_i t}$.

2) Relay processing: Relay $\mathbb{R}$ first down-converts the passband signal by $e^{-j2\pi f_r t}$ and obtains signal blocks of length $N + 2L$. Define the convolution matrix

$$\mathbf{H}_{cv}^{(K)}[\mathbf{x}] = \left.\begin{bmatrix} x_P & \cdots & x_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & x_P & \cdots & x_0 \end{bmatrix}\right\} K \text{ rows}, \tag{18}$$

for any $\mathbf{x} = [x_0, x_1, \ldots, x_P]^T$. Of the mth received block, the first L symbols is corrupted by the IBI, while the last $N + L$ symbols can be written as (19) on the top of this page, where the properties

$$\mathbf{H}_{cv}^{(N+L)}[\mathbf{h}_i]\mathbf{T}_{cp}^{(2L)} = \mathbf{T}_{cp}^{(L)}\mathbf{H}_{cv}^{(N)}[\mathbf{h}_i]\mathbf{T}_{cp}^{(L)} \tag{20}$$

$$\mathbf{H}_{cv}^{(N)}[\mathbf{h}_i]\mathbf{T}_{cp}^{(L)} = \mathbf{H}_{cp}^{(N)}[\mathbf{h}_i] \tag{21}$$

are used.

Importantly, we replace the first L symbols of each OFDM block by zeros instead of removing them, in order to keep the phase continuity of later transmission. To keep the power constraint, $\mathbb{R}$ scales $\mathbf{r}$ by a factor

$$\alpha_{cp} = \sqrt{\frac{(N+2L)P_r}{\mathrm{E}\{\|\mathbf{r}_{cp}\|^2\}}} \tag{22}$$

$$= \sqrt{\frac{N+2L}{N+L} \cdot \frac{P_r}{\sum_{i=1}^{2}\sum_{l=0}^{L}\sigma_{i,l}^2 P_i + \sigma_n^2}},$$

where the property that

$$\left(\mathbf{T}_{cp}^{(L)}\right)^H \mathbf{T}_{cp}^{(L)} = \begin{bmatrix} \mathbf{I}_{N-L} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{I}_L \end{bmatrix} \tag{23}$$

is used during the computation. Relay $\mathbb{R}$ then up-converts the baseband signal by $e^{j2\pi f_r t}$ and broadcasts it to both terminals.

3) Signal reformulation at terminals: After down-converting the passband signal by $e^{-j2\pi f_1 t}$, $\mathbb{T}_1$ obtains the baseband block of length $N + 2L$ and removes the first $2L$ elements. The remaining signal is written in (24) on the top of this page, where the notation $\mathbf{n}_e$ is slightly abused to denote the equivalent noise here, which has the covariance

$$\mathbf{R}_{cp} = \sigma_n^2(\alpha_{cp}^2\mathbf{\Gamma}^{(N)}[f_r - f_1]\mathbf{H}_{cv}^{(N)}[\mathbf{h}_1]$$
$$(\mathbf{H}_{cv}^{(N)}[\mathbf{h}_1])^H(\mathbf{\Gamma}^{(N)}[f_r - f_1])^H + \mathbf{I}). \tag{25}$$

$$\mathbf{y}_{cp} = \alpha_{cp}e^{j2\pi(f_r-f_1)LT_s}\mathbf{H}_{cv}^{(N)}[\mathbf{\Omega}^{(L+1)}[f_1-f_r]\mathbf{h}_1]\mathbf{T}_{cp}^{(L)}\mathbf{H}_{cp}^{(N)}[\mathbf{h}_1]\mathbf{s}_1$$
$$+ \alpha_{cp}e^{j2\pi v[(m-1)(N+2L)+L]T_s}e^{j2\pi(f_r-f_1)LT_s}\mathbf{\Gamma}^{(N)}[v]\mathbf{H}_{cv}^{(N)}[\mathbf{\Omega}^{(L+1)}[f_2-f_r]\mathbf{h}_1]\mathbf{T}_{cp}^{(L)}\mathbf{H}_{cp}^{(N)}[\mathbf{h}_2]\mathbf{s}_2 + \mathbf{n}_e$$
$$= \alpha_{cp}e^{j2\pi(f_r-f_1)LT_s}\mathbf{H}_{cp}^{(N)}[\mathbf{\Omega}^{(L+1)}[f_1-f_r]\mathbf{h}_1]\mathbf{H}_{cp}^{(N)}[\mathbf{h}_1]\mathbf{s}_1$$
$$+ \alpha_{cp}e^{j2\pi(f_r-f_1)LT_s}e^{j2\pi v[(m-1)(N+2L)+L]T_s}\mathbf{\Gamma}^{(N)}[v]\mathbf{H}_{cp}^{(N)}[\mathbf{\Omega}^{(L+1)}[f_2-f_r]\mathbf{h}_1]\mathbf{H}_{cp}^{(N)}[\mathbf{h}_2]\mathbf{s}_2 + \mathbf{n}_e, \tag{29}$$

$$\mathbf{y}_{cp} = \alpha_{cp}\mathbf{H}_{cp}^{(N)}[\underbrace{e^{j2\pi(f_r-f_1)LT_s}((\mathbf{\Omega}^{(L+1)}[f_1-f_r]\mathbf{h}_1)\otimes\mathbf{h}_1)}_{\mathbf{a}_{cp}}]\mathbf{s}_1$$
$$+ \alpha_{cp}e^{j2\pi v[(m-1)(N+2L)+L]T_s}\mathbf{\Gamma}^{(N)}[v]\mathbf{H}_{cp}^{(N)}[\underbrace{e^{j2\pi(f_r-f_1)LT_s}((\mathbf{\Omega}^{(L+1)}[f_2-f_r]\mathbf{h}_1)\otimes\mathbf{h}_2)}_{\mathbf{b}_{cp}}]\mathbf{s}_2 + \mathbf{n}_e. \tag{31}$$

When $N \gg L$, the following approximation can be made:

$$\mathbf{R}_n \approx \underbrace{\sigma_n^2\left(\alpha_{cp}^2\sum_{l=0}^{L}\sigma_{h_{1,l}}^2 + 1\right)}_{\sigma_{ne}^2}\mathbf{I}, \tag{26}$$

where, with slightly abuse of notations, σ_{ne}^2 is again used to denote the equivalent noise variance.

We observe that $(\mathbf{H}_{cv}^{(N)}[\mathbf{h}_i])^T$ has the same structure as $\mathbf{H}_{zp}^{(N)}[\mathbf{h}_i]$ but with the elements of $\mathbf{h}_i$ ordered in a reverse way. Therefore, the principle of Lemma 1 can be straightforwardly extended to derive the following equality:

$$\mathbf{H}_{cv}^{(N)}[\mathbf{h}_i]\mathbf{\Gamma}^{(N+L)}[f] = \mathbf{\Gamma}^{(N)}[f]\mathbf{H}_{cv}^{(N)}[\mathbf{\Omega}^{(L+1)}[f]\mathbf{h}_i], \tag{27}$$

where

$$\mathbf{\Omega}^{(K)}[f] = \mathrm{diag}\{e^{j2\pi f(K-1)T_s},\ldots,e^{j2\pi fT_s},1\}. \tag{28}$$

Hence, $\mathbf{y}_{cp}$ is rewritten as (29) where the property (21) is used when deriving the second equality.

We further note that

$$\mathbf{H}_{cp}^{(N)}[\mathbf{x}_1]\mathbf{H}_{cp}^{(N)}[\mathbf{x}_2] = \mathbf{H}_{cp}^{(N)}[\mathbf{x}_1 \circledast \mathbf{x}_2], \tag{30}$$

where $\circledast$ denotes the N-point circular convolution between two vectors. Since $N \geq 2L + 1$ is assumed, the N-point circular convolution between $\mathbf{h}_i$'s coincides with the linear convolution of their non-zero part plus $N - (2L+1)$ zeros at the end. Then $\mathbf{y}_{cp}$ can be expressed as (31) on the top of this page where $\mathbf{a}_{cp}$, $\mathbf{b}_{cp}$ are the $(2L+1) \times 1$ equivalent channel vectors and $v = f_2 - f_1$ is the equivalent CFO.

Remark 3: By keeping the signal block length as $N + 2L$ at $\mathbb{R}$, only the unique _effective_ CFO v will be involved in the received signal block (13).

4) _Data detection at terminals:_ Just as in ZP-based OFDM, $\mathbb{T}_1$ removes the self-signal component and compensates for the CFO matrix. The resultant signal is

$$\mathbf{w}_{cp} = \alpha_{cp}\mathbf{H}_{cp}^{(N)}[\mathbf{b}_{cp}]\mathbf{s}_2 + \mathbf{n}_e. \tag{32}$$

The conventional OFDM detection can be then performed, and the fading coefficient on each carrier is the DFT of $\mathbf{b}_{cp}$.

TABLE I
COMPARISON BETWEEN ZP- AND CP-OFDM MODULATED TWRN.

	ZP	CP
Transmitter activity	add $2L$ zeros suffix	add $2L$ cyclic prefix
Relay activity	set the last L symbols zero	set the first L symbols zero
Destination activity	none	remove $2L$ prefix
Received signal length	$N + 2L$	N
Required pilot length	$N \geq 2L + 3$	$N \geq 4L + 3$

5) _Joint CFO and channel estimation:_ From previous discussion in II-C we know that the task is to estimate $\mathbf{a}_{cp}$, $\mathbf{b}_{cp}$, and v. Since training sequences appear at the first block, i.e., $m = 1$, we can rewrite $\mathbf{y}_{cp}$ as

$$\mathbf{y}_{cp} = \mathbf{S}_1^{(N)}\mathbf{a}_{cp} + \mathbf{\Gamma}_{cp}^{(N)}[v]\mathbf{S}_2^{(N)}\mathbf{b}_{cp} + \mathbf{n}_e, \tag{33}$$

where $\mathbf{S}_i^{(N)}$ is the $N \times (2L+1)$ circulant matrix with the first column $\alpha_{cp}\mathbf{s}_i$ and $\mathbf{\Gamma}_{cp}^{(N)}[v] = e^{j2\pi vLT_s}\mathbf{\Gamma}^{(N)}[v]$. Moreover, $N \geq 4L + 3$ is required to estimate all the unknown parameters.

D. Comparison between ZP-based and CP-based OFDM

1) _Similarities:_ Both protocols require the same amount of the redundancy, that is $4L$ extra samples over two phases. The transmission efficiency can be defined as $N/(2N + 4L)$. Moreover, both the final models (16) and (33) have the same structure. Thus, the same type of estimator can be designed for both systems.

2) _Differences:_ For the ZP approach, the final received signal blocks at the terminals are of the length $N + 2L$. However, for the CP approach, the final received signal blocks at the terminals are of the length N. For both approaches, the number of unknowns to be estimated is $2(2L+1) + 1$. The equivalent channels are of different definitions. That is, $\mathbf{a}_{zp}$ and $\mathbf{a}_{cp}$ take different values, and so do $\mathbf{b}_{zp}$ and $\mathbf{b}_{cp}$. Since ZP-based OFDM has a longer observation interval, the related estimation is expected to be more accurate.

Table I presents a detailed comparison between the two proposed OFDM protocols.

III. NULLING-BASED LEAST SQUARE ESTIMATION

Since (16) and (33) have the same structure, we will omit the superscript and the subscript, aiming to provide a unified estimation algorithm for both OFDM protocols.

A. Algorithm

Since $\mathbf{S}_1$ is a tall matrix, a matrix $\mathbf{J}$ can be found such that $\mathbf{J}^H\mathbf{S}_1 = \mathbf{0}$. We propose to select $\mathbf{J}$ with the orthogonal property that $\mathbf{J}^H\mathbf{J} = \mathbf{I}$, since it has the best condition number. A simple choice of $\mathbf{J}$ is the basis of the orthogonal complement space of $\mathbf{S}_1$.

Left-multiplying $\mathbf{y}$ by $\mathbf{J}^H$ gives

$$\mathbf{J}^H\mathbf{y} = \mathbf{0} + \underbrace{\mathbf{J}^H\mathbf{\Gamma}\mathbf{S}_2}_{\mathbf{G}}\mathbf{b} + \underbrace{\mathbf{J}^H\mathbf{n}_e}_{\mathbf{n}}, \tag{34}$$

where $\mathbf{G}$ and $\mathbf{n}$ are defined as the corresponding items. Due to the orthogonal property of $\mathbf{J}$, we know the statistics of $\mathbf{n}$ remain the same as that of $\mathbf{n}_e$ if the latter is approximated as white Gaussian noise.

The nulling-based least square (LS) estimate of $\mathbf{b}$ can be immediately found from (34) as:

$$\hat{\mathbf{b}} = (\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\mathbf{J}^H\mathbf{y}. \tag{35}$$

The CFO is estimated from

$$\hat{v} = \arg\max_v \ \mathbf{y}^H\mathbf{J}\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\mathbf{J}^H\mathbf{y}, \tag{36}$$

and $\hat{\mathbf{b}}$ is obtained from (35). Finally, the LS estimation of channel $\mathbf{a}$ is obtained from

$$\hat{\mathbf{a}} = (\mathbf{S}_1^H\mathbf{S}_1)^{-1}\mathbf{S}_1^H(\mathbf{y} - \hat{\mathbf{\Gamma}}\mathbf{S}_2\hat{\mathbf{b}}), \tag{37}$$

where

$$\hat{\mathbf{\Gamma}} = \hat{\mathbf{\Gamma}}[\hat{v}] = \text{diag}\{1, e^{j2\pi\hat{v}T_s}, ..., e^{j2\pi(M-1)\hat{v}T_s}\}. \tag{38}$$

We need to emphasize that the proposed nulling-based LS algorithm is a modification of the standard LS estimation. One can also apply any other type of standard estimation algorithms over (16) and (33).

B. Performance Analysis of Estimation Mean-Square Error

Due to the nulling process, the estimation in (16) and (33) is more complicated than the classical model in [12]. Hence, the analysis in [12] cannot be directly extended to our scenario, which demands for a new derivation.

In this subsection, we will prove that the nulling-based LS estimator is unbiased at high SNR and derive the closed-form MSE expressions by using the perturbation theory. Note that asymptotic analysis is commonly used in the joint estimation problems [17].

For notation simplicity, we denote

$$\mathbf{y}_n = \mathbf{J}^H\mathbf{y}, \qquad \mathbf{P}_G = \mathbf{G}\underbrace{(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H}_{\mathbf{\Phi}}, \tag{39}$$

where $\mathbf{\Phi}$ represents the corresponding item. Let v_0 and $\hat{v}_0$ be the true and the estimated CFO, respectively. The LS estimator (36) can be written as

$$\hat{v}_0 = \arg\max_v g(v) = \arg\max_v \mathbf{y}_n^H\mathbf{P}_G\mathbf{y}_n. \tag{40}$$
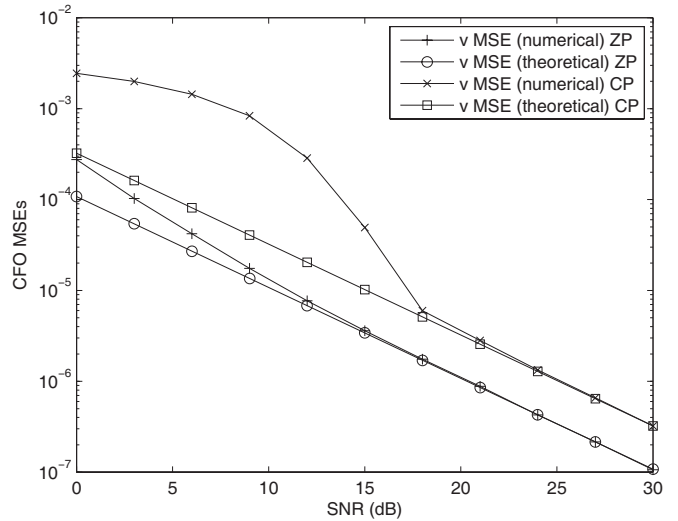

Fig. 3. CFO estimation MSEs version SNR for both ZP- and CP-based OFDM modulation; $N = 16$.

Lemma 2: At high SNR, the CFO estimation error from (40) can be approximated by

$$\Delta v \triangleq \hat{v}_0 - v_0 \approx -\frac{\dot{g}(v_0)}{\text{E}\{\ddot{g}(v_0)\}}, \tag{41}$$

and the CFO estimation is unbiased, i.e., $E\{\Delta v\} = 0$.

Proof: See Appendix I. ∎

Theorem 1: The MSE of the CFO estimation is

$$\text{E}\{\Delta v^2\} = \frac{\sigma_{ne}^2}{2\mathbf{b}^H\dot{\mathbf{G}}^H[\mathbf{I} - \mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H]\dot{\mathbf{G}}\mathbf{b}}. \tag{42}$$

Proof: See Appendix II. ∎

Theorem 2: The channel estimation $\hat{\mathbf{b}}$ is unbiased at high SNR, and its MSE is

$$\text{MSE}\{\mathbf{b}\} = (\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\dot{\mathbf{G}}\mathbf{b}\mathbf{b}^H\dot{\mathbf{G}}^H\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}\text{E}\{\Delta v^2\}$$
$$+ \sigma_{ne}^2(\mathbf{G}^H\mathbf{G})^{-1}. \tag{43}$$

Proof: See Appendix III. ∎

IV. SIMULATIONS RESULTS

In this section, the performance of our proposed joint CFO and channel estimation strategy is studied. A four-tap channel model with the exponential delay profile $\sigma_{il}^2 = e^{-l/10}, i = 1, 2, l = 0, 1, 2, 3$ is assumed. The variance of the noise is taken as $\sigma_n^2 = 1$. The SNR is defined as the ratio of symbol power to the noise power, i.e., E_s/N_0. The normalized frequencies f_1, f_r, and f_2 are set as 0.95, 1 and 1.05, respectively. Thus the CFO is as large as 0.1. The MSEs for CFO and the channel are chosen as the figure of merit, defined respectively by

$$\text{MSE}(v) = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{v}_i - v)^2,$$

$$\text{MSE}(\mathbf{x}) = \frac{1}{10000} \sum_{i=1}^{10000} \frac{1}{7}(\hat{\mathbf{x}}_i - \mathbf{x})^2,$$

where $\mathbf{x}$ represents a composite channel, such as $\mathbf{a}_{zp}$ or $\mathbf{b}_{zp}$ in (13), and $\mathbf{a}_{cp}$ or $\mathbf{b}_{cp}$ in (31). 10000 Monte-Carlo trials are
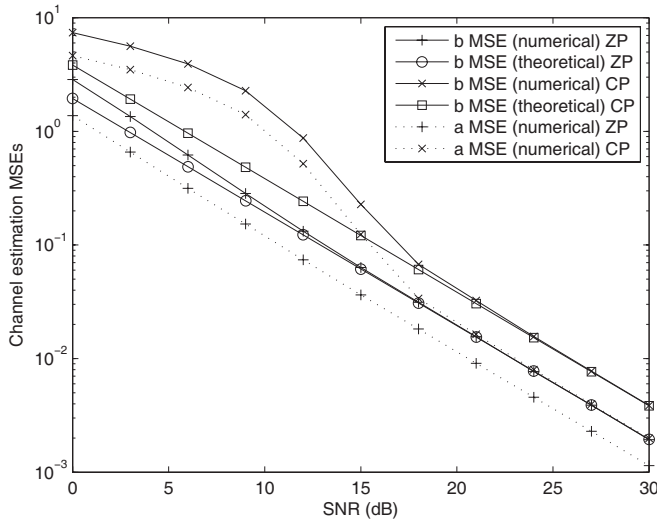
Fig. 4. Channel estimation MSEs version SNR for both ZP- and CP-based OFDM modulation; $N = 16$.
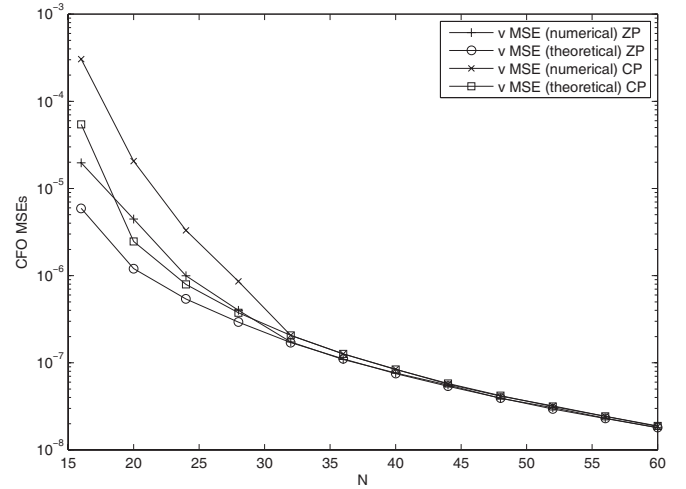

Fig. 5. CFO estimation MSEs version block length for both ZP- and CP-based OFDM modulation; SNR= 10 dB.
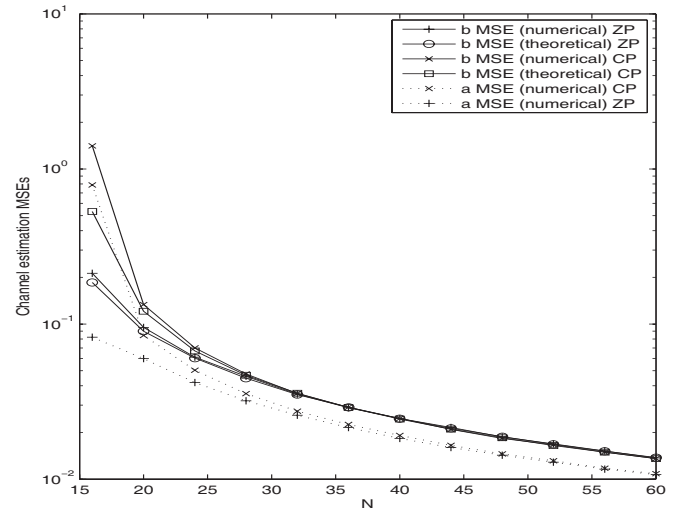

Fig. 6. Channel estimation MSEs version block length for both ZP- and CP-based OFDM modulation; SNR= 10 dB.

used for the averaging. For each example, the ZP-based and CP-based OFDM schemes are compared fairly, i.e., the same CFO, channel, noise realization and training sequence for each Monte-Carlo run.

In the first example, the block length $N = 16$ and the CFO MSEs are plotted versus SNR for both ZP- and CP-based OFDM TWRN (Fig. 3). The theoretical MSEs are also displayed in the same figure. Fig. 3 reveals that for both protocols, the CFO MSEs approach their theoretical values at high SNR. The mismatch in the low SNR region, generally known as the *outlier* [19], [20], occurs because of the estimation ambiguity in several Monte-Carlo runs, which distorts the average performance. Moreover, as discussed in Section II-D, CFO estimation from the ZP-based OFDM is better than that from the CP-based OFDM. The difference is observed to be 5 dB. Clearly, the reason is the longer signal block for the ZP-based OFDM.

The channel estimation MSEs versus SNR is shown in Fig. 4. The estimation MSEs of $\mathbf{b}_{zp}$ and $\mathbf{b}_{cp}$ approach their corresponding theoretical values much faster than those for CFO estimation because the errors in the estimated phase have less effect on the channel estimation but a more severe effect on the CFO estimation. Again, channel estimation of both $\mathbf{a}_{zp}$ and $\mathbf{b}_{zp}$ is better than that of $\mathbf{a}_{cp}$ and $\mathbf{b}_{cp}$ and the difference is about 3 dB.

The first two examples demonstrate the superiority of the ZP-based TWRN estimates because of the longer observation block. Intuitively, the difference should decrease when block length N gets larger. It is then of interest to examine how the difference changes with N. Therefore, the CFO estimation MSEs versus the OFDM block length is plotted in Fig. 5 for a fixed SNR= 10 dB. Fig. 5 reveals that with larger N, the estimations from both the ZP-based and CP-based OFDM become more accurate. Performance difference decreases as N increases and they converge for $N \geq 48$.

The channel estimation MSEs versus the OFDM block length is shown in Fig. 6, which suggests a similar observation made from Fig. 5. However, the difference between the ZP-

OFDM and CP-OFDM diminishes at $N = 40$.

Finally, we examine the symbol decoding errors for both ZP- and CP-based OFDM protocols. To demonstrate the difference, we choose a small value of N as 16. The symbol constellations from both $\mathbb{T}_1$ and $\mathbb{T}_2$ are set as QPSK. The symbol-error rate (SER) versus SNR is shown in Fig. 7. We also include the detection performance with perfect synchronization and channel information for comparison. Interestingly, ZP- and CP-based OFDM exhibit the similar detection performance under perfect synchronization and channel estimation. This phenomenon is similar to that observed in the point-to-point OFDM systems [15].[7] However, with the estimated CFO and channels, we identify a 2 dB performance gain in ZP-OFDM as compared to its CP counterpart. This gain immediately derives from the improved estimation accuracy.

---

[7]Although we scale the symbol power in the ZP-based OFDM, the noise power will also increase when the last $2L$ symbols are added to the first $2L$ symbols in (15).
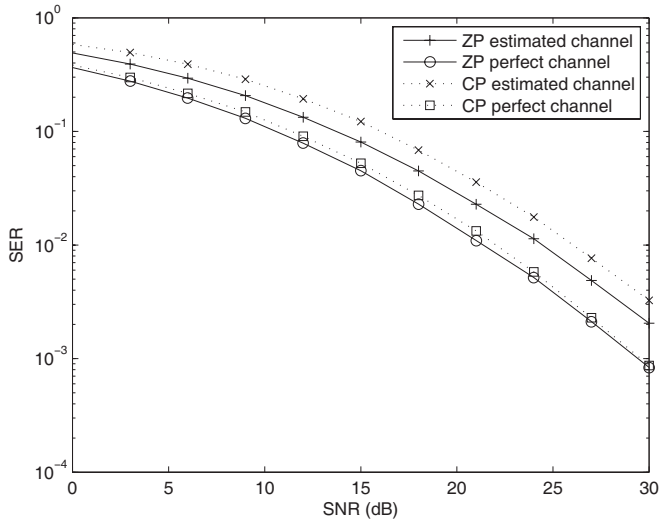
Fig. 7. Performance SER versus SNR for both ZP- and CP-based OFDM modulation; $N = 16$.

V. CONCLUSION

In this paper, we studied joint CFO and channel estimation for TWRN over frequency selective channels. Our main contribution is the proposal and analysis of two OFDM transmission protocols for TWRN, which allow low-complexity joint estimation and data detection. The performance of the nulling-based LS estimator was studied by proving that it is unbiased at high SNR and by deriving the closed-form expressions of the MSEs. Finally, the simulation results demonstrated the effectiveness of the proposed schemes. Interestingly, although ZP-based OFDM performs better than the CP-based OFDM, the performance difference diminishes when the block length gets large. This practically useful finding suggests that CP-based OFDM could be preferred due to the compatibility with the existing OFDM standards.

Remark 4: Practically one may prefer to use CP-OFDM since it is more compatible with currently adopted standards, e.g., IEEE 802.11a [16]. One important insight gained from Fig. 5 and Fig. 6 is that the CP-OFDM will not suffer from much performance loss compared to the ZP-based OFDM when N is large. This makes our CP-OFDM protocol a good choice for practical standards, e.g., $N = 64, 1024, 2048$.

APPENDIX I
PROOF OF LEMMA 2

From [20], we know that

$$\Delta v \approx - \frac{\frac{\partial g(v)}{\partial v}\big|_{v=v_0}}{\frac{\partial^2 g(v)}{\partial v^2}\big|_{v=v_0}} = -\frac{\dot{g}(v_0)}{\ddot{g}(v_0)}. \quad (44)$$

The first order derivative of $\mathbf{G}$ can be calculated as

$$\dot{\mathbf{G}} = \frac{\partial \mathbf{G}}{\partial v} = j\mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2. \quad (45)$$

Applying the equality

$$\frac{\partial \boldsymbol{\Phi}^{-1}}{\partial v} = -\boldsymbol{\Phi}^{-1}\frac{\partial \boldsymbol{\Phi}}{\partial v}\boldsymbol{\Phi}^{-1}$$
$$= -\boldsymbol{\Phi}^{-1}(\dot{\mathbf{G}}^H\mathbf{G} + \mathbf{G}^H\dot{\mathbf{G}})\boldsymbol{\Phi}^{-1}, \quad (46)$$

we get

$$\dot{g}(v) = \mathbf{y}_n^H \dot{\mathbf{P}}_G \mathbf{y}_n = \underbrace{\mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n}_{M_1} + \underbrace{\mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n}_{M_2}$$
$$- \underbrace{\mathbf{y}_n^H (\mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H\mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H + \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H\dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H)\mathbf{y}_n}_{M_3}. \quad (47)$$

It can be found that

$$\mathrm{E}(M_1(v_0)) = j\mathrm{E}[(\mathbf{Gb} + \mathbf{n})^H \mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2\boldsymbol{\Phi}^{-1}\mathbf{G}^H(\mathbf{Gb} + \mathbf{n})]$$
$$= j\mathrm{E}[\mathbf{b}^H \mathbf{G}^H \mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2\mathbf{b}]$$
$$+ j\mathrm{E}[\mathbf{n}^H \mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2\boldsymbol{\Phi}^{-1}\mathbf{G}^H\mathbf{n}], \quad (48)$$

$$\mathrm{E}(M_2(v_0)) = -j\mathrm{E}[\mathbf{b}^H \mathbf{S}_2^H \mathbf{D}\boldsymbol{\Gamma}^H \mathbf{J}\mathbf{Gb}]$$
$$- j\mathrm{E}[\mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{S}_2^H \mathbf{D}\boldsymbol{\Gamma}^H \mathbf{J}\mathbf{n}], \quad (49)$$

$$\mathrm{E}(M_3(v_0)) = \mathrm{E}[\mathbf{b}^H(-j\mathbf{S}_2\mathbf{D}\boldsymbol{\Gamma}^H \mathbf{J}\mathbf{G} + j\mathbf{G}^H \mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2)\mathbf{b}]$$
$$+ \mathrm{E}[\mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}(-j\mathbf{S}_2\mathbf{D}\boldsymbol{\Gamma}^H \mathbf{J}\mathbf{G})\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n}]$$
$$+ \mathrm{E}[\mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}(j\mathbf{G}^H \mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2)\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n}]. \quad (50)$$

Combining (48), (49), (50) and (47), we get

$$\mathrm{E}[\dot{g}(v_0)] = j\mathrm{E}[\mathbf{n}^H(\mathbf{I} - \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H)\mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n}]$$
$$- j\mathrm{E}[\mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{S}_2^H \mathbf{D}\boldsymbol{\Gamma}^H \mathbf{J}(\mathbf{I} - \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H)\mathbf{n}]$$
$$= -2\sigma_{ne}^2 \Im\{tr(\underbrace{\mathbf{G}^H(\mathbf{I} - \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H)}_{\mathbf{0}})\mathbf{J}^H \mathbf{D}\boldsymbol{\Gamma}\mathbf{S}_2\boldsymbol{\Phi}^{-1}\}$$
$$= 0. \quad (51)$$

In order to get $\ddot{g}(v_0)$, we need to compute the first-order derivative of M_1, M_2, and M_3 as

$$\dot{M}_1 = \mathbf{y}_n^H \ddot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n - \mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n$$
$$+ \mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n, \quad (52)$$

$$\dot{M}_2 = \mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n - \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n$$
$$+ \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\ddot{\mathbf{G}}^H \mathbf{y}_n, \quad (53)$$

$$\dot{M}_3 = \mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n + \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\ddot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n$$
$$- \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n + \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n$$
$$- \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n. \quad (54)$$

Thus we obtain

$$\ddot{g}(v_0) = \mathbf{y}_n^H \ddot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n + \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\ddot{\mathbf{G}}^H \mathbf{y}_n$$
$$+ 2\mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n - 2\mathbf{y}_n^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n$$
$$- 2\mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{y}_n - \mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\ddot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n$$
$$+ 2\mathbf{y}_n^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{y}_n$$
$$= \mathbf{b}^H \mathbf{G}^H \ddot{\mathbf{G}}\mathbf{b} + \mathbf{n}^H \ddot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n} + \mathbf{b}^H \mathbf{G}^H \ddot{\mathbf{G}}\mathbf{G}^H \mathbf{b}$$
$$+ \mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}\ddot{\mathbf{G}}^H \mathbf{n} + 2\mathbf{b}^H \mathbf{G}^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{G}\mathbf{b}$$
$$+ 2\mathbf{n}^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{n} - 2\mathbf{b}^H \mathbf{G}^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\mathbf{b}$$
$$- 2\mathbf{n}^H \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n} - 2\mathbf{b}^H \dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{G}\mathbf{b}$$
$$- 2\mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H \mathbf{n} - \mathbf{b}^H \ddot{\boldsymbol{\Phi}}\mathbf{b}$$
$$- \mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}\ddot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n} + 2\mathbf{b}^H \dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\mathbf{b}$$
$$+ 2\mathbf{n}^H \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H \mathbf{n}. \quad (55)$$

After some tedious simplification, it can be obtained that

$$\mathrm{E}[\ddot{g}(v_0)] = 2\mathbf{b}^H \dot{\mathbf{G}}^H(\mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H - \mathbf{I})\dot{\mathbf{G}}\mathbf{b}, \quad (56)$$

and $\ddot{g}(v_0)$ can be written as

$$\ddot{g}(v_0) = \mathrm{E}\{\ddot{g}(v_0)\} + \mathcal{O}_2(\mathbf{n}) + \mathcal{O}_2(\mathbf{n}^2), \quad (57)$$

where $\mathcal{O}_2(\mathbf{n})$ and $\mathcal{O}_2(\mathbf{n}^2)$ represent the linear and quadrature functions of $\mathbf{n}$ in $\ddot{g}(v_0)$, whose explicit forms are omitted for the sake of brevity.

Similarly, $\dot{g}(v_0)$ can be expressed as

$$\dot{g}(v_0) = \mathcal{O}_1(\mathbf{n}) + \mathcal{O}_1(\mathbf{n}^2), \quad (58)$$

where $\mathcal{O}_1(\mathbf{n})$ and $\mathcal{O}_1(\mathbf{n}^2)$ represent the linear and quadrature functions of $\mathbf{n}$ existing in $\dot{g}(v_0)$. Substituting (58) and (57) into (44) gives

$$\Delta v \approx - \frac{\mathcal{O}_1(\mathbf{n}) + \mathcal{O}_1(\mathbf{n}^2)}{\mathrm{E}\{\ddot{g}(v_0)\} + \mathcal{O}_2(\mathbf{n}) + \mathcal{O}_2(\mathbf{n}^2)}$$
$$\approx - \frac{\mathcal{O}_1(\mathbf{n}) + \mathcal{O}_1(\mathbf{n}^2)}{\mathrm{E}\{\ddot{g}(v_0)\}} = - \frac{\dot{g}(v_0)}{\mathrm{E}\{\ddot{g}(v_0)\}}. \quad (59)$$

From (51), (56) and (59), we can find $E\{\Delta v\} = 0$.

APPENDIX II
PROOF FOR THEOREM 1

According to Lemma 2, the MSE of the CFO estimation is

$$\mathrm{E}\{\Delta v^2\} = \frac{\mathrm{E}\{\dot{g}(v_0)^2\}}{\mathrm{E}\{\ddot{g}(v_0)\}^2}. \quad (60)$$

The numerator can be computed as

$$\begin{aligned} \mathrm{E}[\dot{g}(v)^2] &= \mathrm{E}[\mathbf{y}_n^H \dot{\mathbf{P}}_G \mathbf{y}_n \mathbf{y}_n^H \dot{\mathbf{P}}_G \mathbf{y}_n] \\ &= \mathrm{E}[\mathbf{y}_n^H \dot{\mathbf{P}}_G (\mathbf{n}\mathbf{n}^H + \mathbf{G}\mathbf{b}\mathbf{b}^H \mathbf{G}^H) \dot{\mathbf{P}}_G \mathbf{y}_n] \\ &= \sigma_{ne}^2 \mathrm{E}[\mathbf{n}^H \dot{\mathbf{P}}_G \dot{\mathbf{P}}_G \mathbf{n}] + \sigma_{ne}^2 \mathrm{E}[\mathbf{b}^H \mathbf{G}^H \dot{\mathbf{P}}_G \dot{\mathbf{P}}_G \mathbf{G}\mathbf{b}] \\ &\quad + \mathrm{E}[\mathbf{n}^H \dot{\mathbf{P}}_G \mathbf{G}\mathbf{b}\mathbf{b}^H \mathbf{G}^H \dot{\mathbf{P}}_G \mathbf{n}] \\ &\quad + \mathrm{E}[\mathbf{b}^H \mathbf{G}^H \dot{\mathbf{P}}_G \mathbf{G}\mathbf{b}\mathbf{b}^H \mathbf{G}^H \dot{\mathbf{P}}_G \mathbf{G}\mathbf{b}], \quad (61) \end{aligned}$$

where

$$\dot{\mathbf{P}}_G = \dot{\mathbf{G}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H + \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\mathbf{G}}^H - \mathbf{G}\boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\mathbf{G}^H. \quad (62)$$

At high SNR, the first term in (61) can be neglected, and the last term is $\mathbf{0}$ because

$$\mathbf{G}^H \dot{\mathbf{P}}_G \mathbf{G} = \mathbf{0}. \quad (63)$$

Moreover, the second and the third term are the same. After some tedious computation, we obtain

$$\mathbf{G}^H \dot{\mathbf{P}}_G \dot{\mathbf{P}}_G \mathbf{G} = \dot{\mathbf{G}}^H [\mathbf{I} - \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H]\dot{\mathbf{G}}. \quad (64)$$

Therefore, (61) can be rewritten as

$$\mathrm{E}[\dot{g}(v)^2] = 2\sigma_{ne}^2 \mathbf{b}^H \dot{\mathbf{G}}^H [\mathbf{I} - \mathbf{G}\boldsymbol{\Phi}^{-1}\mathbf{G}^H]\dot{\mathbf{G}}\mathbf{b}. \quad (65)$$

Substituting (56) and (65) into (60), we proved Theorem 1.

APPENDIX III
PROOF FOR THEOREM 2

From (35), we know

$$\hat{\mathbf{b}} = (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^H \mathbf{J}^H \mathbf{y} = (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^H (\mathbf{Gb} + \mathbf{n}), \quad (66)$$

where $\hat{\mathbf{G}} = \mathbf{J}^H \hat{\boldsymbol{\Gamma}} \mathbf{S}_2$, and $\hat{\boldsymbol{\Gamma}}$ is defined in (38).

From Taylor's expansion, we know

$$e^{jm\hat{v}_0} = e^{jmv_0} + jme^{jmv_0}\Delta v - m^2 e^{jmv_0}\Delta v^2 + ... \quad (67)$$

Then, (38) can be expressed as

$$\hat{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma} + j\mathbf{D}\boldsymbol{\Gamma}\Delta v - \mathbf{D}^2\boldsymbol{\Gamma}\Delta v^2 + ... \quad (68)$$

At high SNR, the higher order statistics can be omitted, and $\hat{\mathbf{G}}$ can be rewritten as

$$\hat{\mathbf{G}} \approx \mathbf{J}^H (\boldsymbol{\Gamma} + j\mathbf{D}\boldsymbol{\Gamma}\Delta v)\mathbf{S}_2 = \mathbf{G} + \dot{\mathbf{G}}\Delta v. \quad (69)$$

Substituting (69) into (66), we obtain that

$$\hat{\mathbf{b}} = \mathbf{b} - (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^H \dot{\mathbf{G}}\Delta v\mathbf{b} + (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^H \mathbf{n}. \quad (70)$$

At high SNR, using the approximation $(\mathbf{I} + \Delta\mathbf{X})^{-1} \approx \mathbf{I} - \Delta\mathbf{X}$ for a positive semi-definite matrix [18] and omitting the higher order statistics, we obtain

$$(\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \approx (\mathbf{G}^H \mathbf{G})^{-1} - \boldsymbol{\Phi}^{-1}\dot{\boldsymbol{\Phi}}\boldsymbol{\Phi}^{-1}\Delta v. \quad (71)$$

Then we can rewrite (70) as

$$\hat{\mathbf{b}} \approx \mathbf{b} - (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \dot{\mathbf{G}}\Delta v\mathbf{b} + (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^H \mathbf{n}. \quad (72)$$

Therefore,

$$\mathrm{E}\{\Delta\mathbf{b}\} = \mathrm{E}\{\hat{\mathbf{b}} - \mathbf{b}\} \quad (73)$$
$$= \mathrm{E}\{-(\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \dot{\mathbf{G}}\Delta v\mathbf{b} + (\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1} \hat{\mathbf{G}}^H \mathbf{n}\} = \mathbf{0},$$
$$\mathrm{E}\{\Delta\mathbf{b}\Delta\mathbf{b}^H\} = \mathrm{E}\{(\hat{\mathbf{b}} - \mathbf{b})(\hat{\mathbf{b}} - \mathbf{b})^H\}$$
$$= \mathrm{E}\{(\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \dot{\mathbf{G}}\mathbf{b}\mathbf{b}^H \dot{\mathbf{G}}^H \mathbf{G}(\mathbf{G}^H \mathbf{G})^{-1}(\Delta v)^2\}$$
$$+ \sigma_{ne}^2 \mathrm{E}\{(\hat{\mathbf{G}}^H \hat{\mathbf{G}})^{-1}\}. \quad (74)$$

Using (71), we can rewrite (74) as

$$\mathrm{E}\{\Delta\mathbf{b}\Delta\mathbf{b}^H\} = (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \dot{\mathbf{G}}\mathbf{b}\mathbf{b}^H \dot{\mathbf{G}}^H \mathbf{G}(\mathbf{G}^H \mathbf{G})^{-1}$$
$$\times \mathrm{E}\{\Delta v^2\} + \sigma_{ne}^2 (\mathbf{G}^H \mathbf{G})^{-1}. \quad (75)$$

REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[2] S. G. S. Katti and D. Katabi, "Embracing wireless interference: analog network coding," in *Computer Science and Artificial Intelligence Laboratory Technical Report*, Feb. 2007.

[3] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Proc. Annual Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Oct. 2005, pp. 1066–1071.

[4] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proc. IEEE ISIT*, Seattle, USA, July 2006, pp. 1668–1672.

[5] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions for bi-directional relaying," submitted to *IEEE Trans. Inf. Theory*, avalaible at arXiv:0808.0954v1.

[6] T. Cui, T. Ho, and J. Kliewer, "Memoryless relay strategies for two-way relay channels: performance analysis and optimization," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 1139–1143.

[7] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 3888–3892.

[8] R. Zhang, Y.-C. Liang, C. C. Chai, and S. G. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, June 2009.

[9] C. K. Ho, R. Zhang, and Y.-C. Liang, "Two-way relaying over OFDM: optimized tone permutation and power allocation," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 3908–3912.

[10] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.

[11] ——, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4443–4455, Nov. 2009.

[12] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1580–1589, Sep. 2000.

[13] G. Wang, F. Gao, and C. Tellambura, "Joint frequency offset and channel estimation methods for two-way relay networks," in *Proc. IEEE GLOBECOM*, Honolulu, Hawaii, Nov./Dec. 2009.

[14] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.

[15] F. Gao, Y. Zeng, A. Nallanathan, and T. S. Ng, "Robust subspace blind channel estimation for cyclic prefixed MIMO OFDM systems: algorithm, identifiability and performance analysis," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 378–388, Feb. 2008.

[16] "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: high speed physical layer in the 5 GHZ band," IEEE 802.11a, 1999.

[17] U. Tureli, D. Kivanc, and H. Liu, "Experimental and analytical studies on a high resolution OFDM carrier estimator," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 629–643, Mar. 2001.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[19] F. Gao and A. Nallanathan, "Blind maximum likelihood CFO estimation for OFDM systems via polynomial rooting," *IEEE Signal Process. Lett.*, vol. 13, no. 2, pp. 73–76, Feb. 2006.

[20] F. Gao, T. Cui, and A. Nallanathan, "Scattered pilots and virtual carriers based frequency offset tracking for OFDM systems: algorithms, identifiability, and performance analysis," *IEEE Trans. Commun.*, vol. 56, no. 4, pp. 619–629, Apr. 2008.

[21] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1910–1917, Nov. 2003.
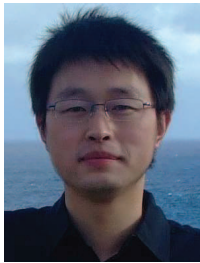
Gongpu Wang received the B.Eng. degree in communication engineering from Anhui University, Hefei, Anhui China, in 2001, the M.Sc. degree from Beijing University of Posts and Telecommunications, Beijing China, in 2004. From 2004 to 2007, he was a teacher in Beijing University of Posts and Telecommunications. He is currently working towards a Ph.D. degree at the Electrical and Computer Engineering Department, University of Alberta, AB, Canada. His research interests are in wireless communication theory and signal processing for communications.

Feifei Gao (S'05-M'09) received the B.Eng. degree in information engineering from Xi'an Jiaotong University, Xi'an, Shaanxi China, in 2002, the M.Sc. degree from the McMaster University, Hamilton, ON, Canada in 2004, and the Ph.D. degree from National University of Singapore in 2007. He was a Research Fellow at Institute for Infocomm Research, A*STAR, Singapore in 2008. He joined the School of Engineering and Science at Jacobs University, Bremen, Germany in 2009, where he is currently an Assistant Professor. His research interests are in communication theory, broadband wireless communications, signal processing for communications, MIMO systems, and array signal processing. Mr. Gao has co-authored more than 70 refereed IEEE journal and conference papers and has served as a TPC member for IEEE ICC, IEEE GLOBECOM, IEEE VTC and IEEE PIMRC.

Yik-Chung Wu obtained the B.Eng. (EEE) degree in 1998 and the M.Phil. degree in 2001 from The University of Hong Kong (HKU), and Ph.D. degree in 2005 from Texas A&M University, USA. From Aug. 2005 to Aug. 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since Sep. 2006, he has been with the University of Hong Kong as an Assistant Professor. He was a TPC member for IEEE VTC Fall 2005, Spring 2011, Globecom 2006, 2008, ICC 2007, 2008 and 2011. He is currently serving as an associate editor for the IEEE COMMUNICATIONS LETTERS. Yik-Chung's research interests are in general area of signal processing and communication systems, and in particular receiver algorithm design, synchronization techniques, channel estimation and equalization.

Chintha Tellambura (SM'02) received the B.Sc. degree (with first-class honors) from the University of Moratuwa, Moratuwa, Sri Lanka, in 1986, the M.Sc. degree in electronics from the University of London, London, U.K., in 1988, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1993.

He was a Postdoctoral Research Fellow with the University of Victoria (1993-1994) and the University of Bradford (1995-1996). He was with Monash University, Melbourne, Australia, from 1997 to 2002. Presently, he is a Professor with the Department of Electrical and Computer Engineering, University of Alberta. His research interests include Diversity and Fading Countermeasures, Multiple-Input Multiple-Output (MIMO) Systems and Space-Time Coding, and Orthogonal Frequency Division Multiplexing (OFDM).

Prof. Tellambura is an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the Area Editor for Wireless Communications Systems and Theory in the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was Chair of the Communication Theory Symposium in Globecom'05 held in St. Louis, MO.