

Research Article

Superimposed Training-Based Joint CFO and Channel Estimation for CP-OFDM Modulated Two-Way Relay Networks

Gongpu Wang,¹ Feifei Gao,² Xin Zhang,³ and Chintla Tellambura¹

¹Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G 2V4

²School of Engineering and Science, Jacobs University, 28759 Bremen, Germany

³School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Gongpu Wang, gongpu@ece.ualberta.ca

Received 30 January 2010; Revised 23 May 2010; Accepted 16 June 2010

Academic Editor: Petar Popovski

Copyright © 2010 Gongpu Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Joint carrier frequency offset (CFO) and channel estimation is considered for two-way relay networks (TWRNs). Existing estimators provide only the convolved channel parameters and the mixed CFO values. In contrast, estimators using a superimposed training strategy are developed for the individual frequency and channel parameters. Depending on the number of pilots, three different estimators are developed. An iterative estimator with low complexity is also developed to further improve the estimation accuracy. The Cramér-Rao Bounds (CRBs) are derived. The simulations show that the iterative estimator converges rapidly, and the resultant estimation mean square error (MSE) approaches the CRB. For the special case of small CFO between the two source terminals, the MSE achieves the CRB at high SNRs, and the iterative algorithm is not necessary. However, for the general case, the gap between the MSE and the CRB indicates that there is room for further improvement of the estimation accuracy.

1. Introduction

Two-way relay networks (TWRNs) [1] have received much attention due to their improved spectral efficiency over a one-way relay network (OWRN) [2]. Unlike the OWRN where the data flow is unidirectional from the source to the relay and then to the destination, in a TWRN, two source terminals exchange information via the relay similar to network coding [3]. The overall communication rate of this setup is approximately twice as that achieved in a OWRN [4], making the TWRN concept particularly attractive to bidirectional systems. The achievable rate regions for *amplify-and-forward*- (AF-) and *decode-and-forward*- (DF-) based TWRNs were derived in [5, 6]. In [7], the optimal mapping function at the relay that minimizes the bit-error rate (BER) was proposed, and in [8], distributed space-time codes (STCs) that achieve full diversity were developed for both AF and DF TWRN. The optimal beamforming at the multiantenna relay that maximizes the capacity of AF-based TWRN was designed in [9], and suboptimal resource allocation for orthogonal frequency division multiplexing- (OFDM-) based TWRN was derived in [10].

Most TWRN works [4–10] assume perfect synchronization and channel state information (CSI) at the relay node and/or the source terminals. However, since these assumptions are not realistic, for the first time, [11, 12] designed the joint carrier frequency offset (CFO) and channel estimation algorithms for an OFDM-based TWRN. As the first effort to tackle this problem, [11, 12] only proposed to estimate the cascaded channels and cascaded CFO of the uplink and the downlink phases, but not the individual frequency and channel parameters. However, individual parameters are also important for TWRN systems. For example, they are required for the optimal beamforming and the carrier permutations [9, 10]. Moreover, if detailed information about CFOs is available at both source nodes, the two nodes can cooperate to eliminate CFOs, which may yield a better detection performance (this fact will be seen from the special case in Section 3 and from the corresponding simulation results Figures 4 and 5).

In this paper, we introduce the superimposed pilots at the relay to enable the estimation of all the individual channel and CFO parameters. The idea of superimposed pilots was first proposed in [13] for analog communication systems

and was later extended to digital communication systems for both synchronization and channel estimation [14–16]. Different from the traditional concept that the pilots are superimposed on the data symbols, we propose that the relay node superimposes its own pilots over the pilots received from the two terminals. Depending on the number of pilots, three different estimators are then developed for the initial parameter estimation. An iterative estimator is also proposed to improve the estimation accuracy. Moreover, the Cramér-Rao Bound (CRB) is derived and compared with the estimation mean square error (MSE). We find that, when the CFO between the two source terminals is small, the MSE approaches CRB in the high SNR region, indicating the optimality of the third proposed estimator. However, for more general cases, the MSE deviates from the CRB, which indicates the room for improving the estimation accuracy. This forms an open problem for the future research.

The remainder of this paper is organized as follows. In Section 2, the two-way communication systems and our joint CFO and channel estimation model are introduced. In Section 3, three estimators are proposed for initial CFO and channel estimation. The estimators vary depending on the number of pilots. An iterative estimator is also designed to improve the estimation accuracy. The CRBs of all parameters are derived in Section 4. Simulation results are presented in Section 5, and conclusions are made in Section 6.

1.1. Notations. Vectors and matrices are boldface small and capital letters; the transpose, complex conjugate, Hermitian, inverse, and pseudoinverse of the matrix \mathbf{A} are denoted by \mathbf{A}^T , \mathbf{A}^* , \mathbf{A}^H , \mathbf{A}^{-1} , and \mathbf{A}^\dagger , respectively; $\text{tr}(\mathbf{A})$ and $\|\mathbf{A}\|_F$ are the trace and the Frobenius norm of \mathbf{A} ; $[\mathbf{A}]_{ij}$ is the (i, j) th entry of \mathbf{A} , and $\text{diag}\{\mathbf{a}\}$ denotes a diagonal matrix with the diagonal element constructed from \mathbf{a} ; \otimes represents the linear convolution between two vectors. MATLAB notations for rows and columns of a matrix are adopted; $\mathbf{A}_{[:,i:j]}$ denotes the i th column to the j th columns of \mathbf{A} , and $\mathbf{A}_{[:,i:j]}^H$ denotes the i th column to the j th columns of \mathbf{A}^H ; the entry index of vector and matrix starts from 0; \mathbf{I}_P is the $P \times P$ identity matrix for any positive integer P ; $\mathbf{0}_{n \times m}$ represents the $n \times m$ matrix with all zero entries; $\text{E}\{\cdot\}$ is the average; j is the imaginary unit $\sqrt{-1}$ and \mathbf{e}_i is the basis vector of all-zeros except for the i th element of 1.

2. Problem Formulation

2.1. System Model. Consider a classical two-way relay network (TWRN) with two terminal nodes \mathbb{T}_j , $j = 1, 2$ and one relay node \mathbb{R} (Figure 1). Each node has only one half-duplex antenna. The baseband channel from \mathbb{T}_j to \mathbb{R} is denoted as $\mathbf{h}_j = [h_{j,0}, \dots, h_{j,L}]^T$, whose elements are independent with zero means and variances $\sigma_{j,i}^2$. The channel from \mathbb{R} to \mathbb{T}_j is also \mathbf{h}_j . (This is true for reciprocal channels. There may be a phase difference but is nevertheless ignored in this paper.) The training block length is set as N , which may or may not be the same as the data block length. The average powers of \mathbb{T}_j and \mathbb{R} are denoted as P_j and P_r , respectively. Furthermore, we denote the carrier frequency of \mathbb{T}_j as f_j and that of \mathbb{R} as f_r .

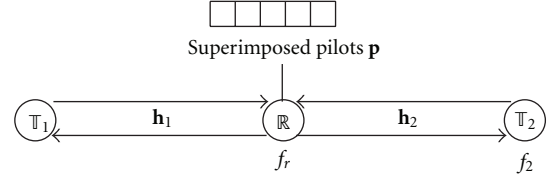


FIGURE 1: System configuration for two-way relay network.

In real applications, Doppler shifts and oscillator instabilities may result in CFOs such as $f_2 - f_1$ and $f_r - f_j$.

The target of this work is to separately estimate CFOs and channels \mathbf{h}_1 and \mathbf{h}_2 . It is important to achieve this goal within two transmission phases such that the training is compatible with the two-phase data transmission and can be embedded in the data frame. To do so, we modify the OFDM transmission scheme and introduce superimposed training at the relay node. In the following discussion, perfect time synchronization is assumed.

2.2. CP-Based OFDM Modulation at Terminals. Denote one OFDM block from \mathbb{T}_i as $\tilde{\mathbf{s}}_i = [\tilde{s}_{i,0}, \dots, \tilde{s}_{i,N-1}]^T$. The corresponding time-domain signal block is obtained from the normalized inverse discrete Fourier transformation (IDFT) as

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

where \mathbf{F} is the $N \times N$ normalized DFT matrix with the (p, q) th entry given by $(1/\sqrt{N})e^{-j2\pi(p-1)(q-1)/N}$. To maintain the subcarrier orthogonality during the overall transmission, we propose to add the cyclic prefix of length $2L$ as did in [11]. This implicitly requires $N \geq 2L$ which is nevertheless satisfied by most OFDM systems.

Define the matrix

$$\mathbf{T}_{\text{cp}}^{(P)} = \begin{bmatrix} \mathbf{0}_{P \times (N-P)} & \mathbf{I}_P \\ \text{---} & \text{---} \\ & \mathbf{I}_N \end{bmatrix}, \quad (2)$$

for any $P \leq N$. The baseband signal from \mathbb{T}_i including the CP can now be mathematically expressed as $\mathbf{T}_{\text{cp}}^{(2L)} \mathbf{s}_i$ of length $N + 2L$. The signal is up-converted to the passband signal by the carrier $e^{j2\pi f_i t}$. 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.

2.3. Relay Processing. With the assumption of the perfect time synchronization, the signals from \mathbb{T}_1 and \mathbb{T}_2 arrive at \mathbb{R} simultaneously. Relay then down-converts the received signal by the carrier $e^{-j2\pi f_r t}$. It is important to mention that \mathbb{R} removes only the first L symbols in each block.

Denote T_s as the symbol duration and define

$$\mathbf{\Gamma}^{(K)}[f] = \text{diag}\{1, e^{j2\pi f T_s}, \dots, e^{j2\pi f (K-1)T_s}\},$$

$$\mathbf{H}_{\text{cv}}^{(K)}[\mathbf{x}] = \left\{ \begin{bmatrix} x_p & \cdots & x_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & x_p & \cdots & x_0 \end{bmatrix} \right\} \quad K \text{ rows}, \quad (3)$$

for $\mathbf{x} = [x_0, x_1, \dots, x_p]^T$. The resultant baseband signal block at \mathbb{R} is of length $N + L$ and can be expressed as

$$\begin{aligned} \mathbf{r} &= \sum_{i=1}^2 e^{j2\pi(f_i - f_r)LT_s} \mathbf{\Gamma}^{(N+L)} [f_i - f_r] \mathbf{H}_{\text{cv}}^{(N+L)} [\mathbf{h}_i] \mathbf{T}_{\text{cp}}^{(2L)} \mathbf{s}_i + \mathbf{n}_r \\ &= \sum_{i=1}^2 e^{j2\pi(f_i - f_r)LT_s} \mathbf{\Gamma}^{(N+L)} [f_i - f_r] \mathbf{T}_{\text{cp}}^{(L)} \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_i] \mathbf{s}_i + \mathbf{n}_r, \end{aligned} \quad (4)$$

where $\mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_i]$ is the $N \times N$ circulant matrix with the first column $[\mathbf{h}_i^T, \mathbf{0}_{1 \times (N-L-1)}^T]^T$, and we use the property that

$$\mathbf{H}_{\text{cv}}^{(N+L)} [\mathbf{h}_i] \mathbf{T}_{\text{cp}}^{(2L)} = \mathbf{T}_{\text{cp}}^{(L)} \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_i]. \quad (5)$$

Moreover, each element in the noise vector \mathbf{n}_r is assumed to be independent and identically distributed (i.i.d) zero-mean complex white Gaussian with variance σ_n^2 . The estimation of the CFO and channels \mathbf{h}_1 and \mathbf{h}_2 at \mathbb{R} can be easily done as in the 2×1 multi-input single-output (MISO) system [17, 18].

The relay then superimposes a time-domain training signal \mathbf{p} over \mathbf{r} and obtain

$$\mathbf{t} = \alpha \mathbf{r} + \mathbf{p}, \quad (6)$$

where the scaling factor α and the superimposed training should satisfy the following power constraint:

$$\begin{aligned} &\alpha^2 \mathbb{E}\{\|\mathbf{r}\|^2\} + \|\mathbf{p}\|^2 \\ &= \alpha^2 (N + L) \sum_{j=1}^2 \sum_{l=0}^L \sigma_{j,l}^2 P_j + \alpha^2 (N + L) \sigma_n^2 + \|\mathbf{p}\|^2 \quad (7) \\ &\leq (N + L) P_r. \end{aligned}$$

Note that (7) is a constraint on the average power because the instant channel values are unknown before the estimation.

It can be easily shown that the range of α is $(0, \sqrt{P_r / (\sum_{j=1}^2 \sum_{l=0}^L \sigma_{j,l}^2 P_j + \sigma_n^2)})$ that balances the power between the training from \mathbb{T}_j and the superimposed training from \mathbb{R} . Note that the training signal \mathbf{p} is generated from N training signals \mathbf{p}_0 and a cyclic prefix of length L . By using our definition (2), the training signal can be expressed as $\mathbf{p} = \mathbf{T}_{\text{cp}}^{(L)} \mathbf{p}_0$. The pilots \mathbf{p}_0 , \mathbf{s}_1 , and \mathbf{s}_2 are pre-designed at both source terminals for channel estimation.

Finally, \mathbb{R} up-converts the resultant signal \mathbf{t} to passband by the carrier $e^{j2\pi f_r t}$.

2.4. Signal Reformulation at Terminals. Due to symmetry, we only illustrate the process at \mathbb{T}_1 . After down-converting the passband signal by $e^{-j2\pi f_1 t}$, \mathbb{T}_1 obtains the baseband block of

length $N + L$. It then removes the first L elements and the remaining signal is written as

$$\begin{aligned} \mathbf{y} &= e^{j2\pi(f_r - f_1)LT_s} \mathbf{\Gamma}^{(N)} [f_r - f_1] \mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_1] (\alpha \mathbf{r} + \mathbf{p}) + \mathbf{n}_1 \\ &= \alpha \mathbf{\Gamma}^{(N)} [f_r - f_1] \mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_1] \mathbf{\Gamma}^{(N+L)} [f_1 - f_r] \mathbf{T}_{\text{cp}}^{(L)} \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_1] \mathbf{s}_1 \\ &\quad + \alpha e^{j2\pi(f_2 - f_1)LT_s} \mathbf{\Gamma}^{(N)} [f_r - f_1] \mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_1] \mathbf{\Gamma}^{(N+L)} [f_2 - f_r] \\ &\quad \times \mathbf{T}_{\text{cp}}^{(L)} \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_2] \mathbf{s}_2 + e^{j2\pi(f_r - f_1)LT_s} \mathbf{\Gamma}^{(N)} [f_r - f_1] \\ &\quad \times \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_1] \mathbf{p}_0 \\ &\quad + \underbrace{\alpha e^{j2\pi(f_r - f_1)LT_s} \mathbf{\Gamma}^{(N)} [f_r - f_1] \mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_1] \mathbf{n}_r + \mathbf{n}_1}_{\mathbf{n}_e}, \end{aligned} \quad (8)$$

where the property

$$\mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_1] \mathbf{T}_{\text{cp}}^{(L)} = \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_1] \quad (9)$$

is used and, without loss of generality, the noise vector \mathbf{n}_1 is assumed to have the same statistics as \mathbf{n}_r . The equivalent Gaussian noise \mathbf{n}_e has the covariance

$$\begin{aligned} \mathbf{R}_n &= \sigma_n^2 \left(\alpha^2 \mathbf{\Gamma}^{(N)} [f_r - f_1] \mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_1] \right. \\ &\quad \left. \times \left(\mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_1] \right)^H \left(\mathbf{\Gamma}^{(N)} [f_r - f_1] \right)^H + \mathbf{I} \right). \end{aligned} \quad (10)$$

Define $\mathbf{\Omega}^{(K)} [f] = \text{diag}\{e^{j2\pi f(K-1)T_s}, \dots, e^{j2\pi f T_s}, 1\}$. The following properties hold

$$\begin{aligned} \mathbf{H}_{\text{cv}}^{(N)} [\mathbf{h}_i] \mathbf{\Gamma}^{(N+L)} [f] &= \mathbf{\Gamma}^{(N)} [f] \mathbf{H}_{\text{cv}}^{(N)} \left[\mathbf{\Omega}^{(L+1)} [f] \mathbf{h}_i \right], \\ \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{x}_1] \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{x}_2] &= \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{x}_1 \circledast \mathbf{x}_2], \end{aligned} \quad (11)$$

where \circledast denotes the N -point circular convolution but reduces to linear convolution if N is greater than or equal to the length of $\mathbf{x}_1 \circledast \mathbf{x}_2$.

Assuming $N > 2L$ and using the above properties, we can further simplify \mathbf{y} as

$$\begin{aligned} \mathbf{y} &= \alpha \mathbf{H}_{\text{cp}}^{(N)} \left[\left(\mathbf{\Omega}^{(L+1)} [f_1 - f_r] \mathbf{h}_1 \right) \circledast \mathbf{h}_1 \right] \mathbf{s}_1 \\ &\quad + \alpha e^{j2\pi(f_2 - f_1)LT_s} \mathbf{\Gamma}^{(N)} [f_2 - f_1] \mathbf{H}_{\text{cp}}^{(N)} \\ &\quad \times \left[\left(\mathbf{\Omega}^{(L+1)} [f_2 - f_r] \mathbf{h}_1 \right) \circledast \mathbf{h}_2 \right] \mathbf{s}_2 \\ &\quad + e^{j2\pi(f_r - f_1)LT_s} \mathbf{\Gamma}^{(N)} [f_r - f_1] \mathbf{H}_{\text{cp}}^{(N)} [\mathbf{h}_1] \mathbf{p}_0 + \mathbf{n}_e. \end{aligned} \quad (12)$$

Define

$$\begin{aligned} w &= f_r - f_1, \quad \mathbf{a} = \left(\mathbf{\Omega}^{(L+1)} [-w] \mathbf{h}_1 \right) \circledast \mathbf{h}_1, \\ v &= f_2 - f_1, \quad \mathbf{b} = \left(\mathbf{\Omega}^{(L+1)} [v - w] \mathbf{h}_1 \right) \circledast \mathbf{h}_2, \\ \mathbf{\Gamma}_L^{(K)} [f] &= e^{j2\pi f L T_s} \mathbf{\Gamma}^{(K)} [f]. \end{aligned} \quad (13)$$

Then (12) can be expressed as

$$\begin{aligned} \mathbf{y} &= \alpha \mathbf{H}_{\text{cp}}^{(N)}[\mathbf{a}]\mathbf{s}_1 + \alpha \Gamma_L^{(N)}[\nu]\mathbf{H}_{\text{cp}}^{(N)}[\mathbf{b}]\mathbf{s}_2 \\ &\quad + \Gamma_L^{(N)}[w]\mathbf{H}_{\text{cp}}^{(N)}[\mathbf{h}_1]\mathbf{p}_0 + \mathbf{n}_e \\ &= \alpha \mathbf{S}_1 \mathbf{a} + \alpha \Gamma_L^{(N)}[\nu]\mathbf{S}_2 \mathbf{b} + \Gamma_L^{(N)}[w]\mathbf{P}\mathbf{h}_1 + \mathbf{n}_e, \end{aligned} \quad (14)$$

where \mathbf{S}_j is the $N \times (2L + 1)$ circulant matrix with the first column \mathbf{s}_j , and \mathbf{P} is the $N \times (L + 1)$ circulant matrix with the first column \mathbf{p}_0 .

3. Joint Estimation Algorithms

Based on the new signal model (14), the task is to estimate the individual channels \mathbf{h}_1 , and \mathbf{h}_2 and the CFOs ν and w . We omit all redundant superscripts and subscripts for notation simplicity and rewrite (14) as

$$\mathbf{y} = \alpha \mathbf{S}_1 \mathbf{a} + \alpha \Gamma[\nu]\mathbf{S}_2 \mathbf{b} + \Gamma[w]\mathbf{P}\mathbf{h}_1 + \mathbf{n}_e. \quad (15)$$

Note that the number of parameters to be estimated is $2L + 4$. Furthermore, \mathbf{a} is a function of w and \mathbf{h}_1 , and \mathbf{b} is a function of ν , w , \mathbf{h}_1 and \mathbf{h}_2 . Depending on the number of pilots, we can thus develop three different estimation methods.

3.1. Estimation for Sufficiently Large N . When $N \geq 5L + 5$, there are sufficient degrees of freedom in the training signals, and ν , w , \mathbf{a} , \mathbf{b} and \mathbf{h}_1 can be simply treated as individual variables. That is, the above-mentioned relationships among the variables are ignored. Rewrite \mathbf{y} as

$$\mathbf{y} = \underbrace{\begin{bmatrix} \alpha \mathbf{S}_1 & \alpha \Gamma[\nu]\mathbf{S}_2 & \Gamma[w]\mathbf{P} \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{a} \\ \mathbf{h}_1 \\ \mathbf{d} \end{bmatrix}}_{\mathbf{d}} + \mathbf{n}_e, \quad (16)$$

where \mathbf{C} and \mathbf{d} are defined in (16). From the least-squares (LS) method, the two CFO estimates are

$$\{\hat{\nu}, \hat{w}\} = \arg \max_{\nu, w} \mathbf{y}^H \mathbf{C} (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H \mathbf{y}, \quad (17)$$

where $\hat{\nu}$ and \hat{w} can be obtained either from a two-dimensional search or from the alternating projection that converts the 2-dimensional maximization into a series of 1D maximization problems. Details on the implementation of the alternating projection method can be found from [19]. Then values of \mathbf{d} are estimated as

$$\hat{\mathbf{d}} = (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H \mathbf{y}, \quad (18)$$

where the value of \mathbf{C} is obtained by using the estimates $\hat{\nu}$ and \hat{w} .

We next explore the relationships among \mathbf{a} , \mathbf{b} , and \mathbf{h}_1 to improve the quality of the estimates. From (13), we note that

$$\mathbf{a} = \mathbf{H}_{\text{zp}}^{(L+1)}[\Omega^{(L+1)}[-w]\mathbf{h}_1]\mathbf{h}_1, \quad (19)$$

where $\mathbf{H}_{\text{zp}}^{(K)}[\mathbf{x}]$ is a tall Toeplitz matrix with the first column $[\mathbf{x}^T, \mathbf{o}_{1 \times (K-1)}^T]$. The estimate of $\mathbf{H}_{\text{zp}}^{(L+1)}[\Omega^{(L+1)}[-w]\mathbf{h}_1]$ can be expressed as $\mathbf{H}_{\text{zp}}^{(L+1)}[\Omega^{(L+1)}[-\hat{w}]\hat{\mathbf{h}}_1]$.

By subtracting the estimate of the second item in (15) from \mathbf{y} and by using (19), an improved estimate of \mathbf{h}_1 is obtained as

$$\begin{aligned} \hat{\mathbf{h}}_1 &= \left(\alpha \mathbf{S}_1 \mathbf{H}_{\text{zp}}^{(L+1)}[\Omega^{(L+1)}[-\hat{w}]\hat{\mathbf{h}}_1] + \Gamma[w]\mathbf{P} \right)^\dagger \\ &\quad \times (\mathbf{y} - \alpha \Gamma[\nu]\mathbf{S}_2 \hat{\mathbf{b}}). \end{aligned} \quad (20)$$

Similarly, from (13) we find that

$$\mathbf{b} = \mathbf{H}_{\text{zp}}^{(L+1)}[\Omega^{(L+1)}[\nu - w]\mathbf{h}_1]\mathbf{h}_2. \quad (21)$$

Thus \mathbf{h}_2 can be estimated as

$$\hat{\mathbf{h}}_2 = \left(\mathbf{H}_{\text{zp}}^{(L+1)}[\Omega^{(L+1)}[\hat{\nu} - \hat{w}]\hat{\mathbf{h}}_1] \right)^\dagger \hat{\mathbf{b}}. \quad (22)$$

In summary, (17), (20) and (22) provide estimates of all the parameters. These initial estimates can be further improved by the iterative estimator developed in Section 3.4.

3.2. Estimation with Not-So-Large N . In order to reduce the overhead, we will use fewer pilots than before. Define \mathcal{K}_1 , \mathcal{K}_2 , and \mathcal{K}_r as the frequency domain pilot index sets from \mathbb{T}_1 , \mathbb{T}_2 , and \mathbb{R} , with cardinality K_1 , K_2 , and K_r , respectively. We require $K_1 \geq L + 1$, $K_2 \geq L + 1$, $K_r \geq L + 1$, and $\mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_r = \{1, \dots, N\}$. Here, we do not assume the disjoint sets, and so $K_1 + K_2 + K_r \geq N$. From (15), we know that the frequency domain pilots are $\tilde{\mathbf{s}}_j = \mathbf{F}\mathbf{s}_j$, and $\tilde{\mathbf{p}}_0 = \mathbf{F}\mathbf{p}_0$.

Remark 1. As will be seen later that pilots in \mathcal{K}_r are used to estimate \mathbf{h}_1 , and those in \mathcal{K}_2 are used to estimate \mathbf{h}_2 at \mathbb{T}_1 . Due to symmetry, pilots in \mathcal{K}_1 are used to estimate \mathbf{h}_1 at \mathbb{T}_2 . This gives the above requirements on cardinality.

Let us collect nonzero pilots from \mathbb{T}_j into a $K_j \times 1$ vector $\check{\mathbf{s}}_j$ and nonzero pilots from \mathbb{R} into a $K_r \times 1$ vector $\check{\mathbf{p}}_0$. Since \mathbf{S}_j and \mathbf{P} are column-wise circulant matrices, they can be represented as

$$\begin{aligned} \mathbf{S}_j &= \mathbf{F}^H \text{diag}\{\check{\mathbf{s}}_j\} \mathbf{F}_{[:,1:2L+1]} = \mathbf{F}_{[:,\mathcal{K}_j]}^H \text{diag}\{\check{\mathbf{s}}_j\} \mathbf{F}_{[\mathcal{K}_j,1:2L+1]} \\ \mathbf{P} &= \mathbf{F}^H \text{diag}\{\check{\mathbf{p}}_0\} \mathbf{F}_{[:,1:L+1]} = \mathbf{F}_{[:,\mathcal{K}_r]}^H \text{diag}\{\check{\mathbf{p}}_0\} \mathbf{F}_{[\mathcal{K}_r,1:L+1]}. \end{aligned} \quad (23)$$

Define $\overline{\mathcal{K}}_1$ as the complement set of \mathcal{K}_1 . Multiplying both sides of (15) with $\mathbf{F}_{[\overline{\mathcal{K}}_1, :]}$ yields

$$\begin{aligned} \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{y} &= \underbrace{\left[\alpha \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \Gamma[\nu] \mathbf{F}_{[:,\mathcal{K}_2]}^H \text{diag}\{\check{\mathbf{s}}_2\} \quad \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \Gamma[w] \mathbf{P} \right]}_{\mathbf{C}_1} \\ &\quad \times \underbrace{\begin{bmatrix} \mathbf{b} \\ \mathbf{h}_1 \\ \mathbf{d}_1 \end{bmatrix}}_{\mathbf{d}_1} + \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{n}_e, \end{aligned} \quad (24)$$

where $\check{\mathbf{b}} = \mathbf{F}_{[\mathcal{K}_2, 1:2L+1]} \mathbf{b}$ is the DFT response of \mathbf{b} on the subcarrier set \mathcal{K}_2 , and \mathbf{C}_1 is an $(N - K_1) \times (K_2 + L + 1)$ matrix.

As long as $N - K_1 - K_2 - (L + 1) \geq 2$, namely, when there is sufficient degree of freedom to estimate the two unknown CFOs, they can be estimated as

$$\{\hat{\nu}, \hat{w}\} = \arg \max_{\nu, w} \mathbf{y}^H \mathbf{F}_{[\overline{\mathcal{K}}_1, :]}^H \mathbf{C}_1 (\mathbf{C}_1^H \mathbf{C}_1)^{-1} \mathbf{C}_1^H \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{y}. \quad (25)$$

Considering the range of K_j and K_r , the minimum number of N is $3L + 5$, when sets are disjoint, and $K_1 = K_2 = L + 1$, $K_r = L + 3$. Then,

$$\hat{\mathbf{d}}_1 = \left(\mathbf{C}_1^H \mathbf{C}_1 \right)^{-1} \mathbf{C}_1^H \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{y}. \quad (26)$$

By definition

$$\check{\mathbf{b}} = \mathbf{F}_{[\mathcal{K}_2, 1:2L+1]} \mathbf{H}_{zp}^{(L+1)} \left[\boldsymbol{\Omega}^{(L+1)} [\nu - w] \mathbf{h}_1 \right] \mathbf{h}_2. \quad (27)$$

Then, \mathbf{h}_2 can be estimated as

$$\hat{\mathbf{h}}_2 = \left(\mathbf{F}_{[\mathcal{K}_2, 1:2L+1]} \mathbf{H}_{zp}^{(L+1)} \left[\boldsymbol{\Omega}^{(L+1)} [\nu - w] \mathbf{h}_1 \right] \right)^\dagger \check{\mathbf{b}}. \quad (28)$$

Remark 2. Though the number of unknown parameters is $2L + 3$, we should allow certain redundancy to provide low complexity estimation, for example, linear estimation. In our work, the minimum training length is set as $N = 3L + 5$.

3.3. Joint Estimation with Minimum Training Length: A Special Case. In practical applications, the relay terminal is often a simple device while the two source terminals may employ high-precision synchronization circuits. Thus, it is reasonable to expect the CFO between the two source terminals to be negligible. In this case, $\nu \approx 0$ or $\nu \ll 1/N$, that is, one subcarrier spacing. This is also true at the CFO tracking stage when the frequency difference between two terminals is quite small. By taking advantage of the negligible CFO between the two source terminals, parameter estimation is achieved with the minimum training length $N = 2L + 3$, that is, the same number of the unknowns variables.

Let us choose the same frequency pilot sets for \mathbb{T}_1 and \mathbb{T}_2 , that is, $\mathcal{K}_1 = \mathcal{K}_2$. Left multiplying the received signal \mathbf{y} by $\mathbf{F}_{[\overline{\mathcal{K}}_1, :]}$ gives

$$\begin{aligned} \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{y} &= \underbrace{\alpha \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \boldsymbol{\Gamma} [\nu] \mathbf{F}_{[:, \mathcal{K}_2]}^H \text{diag}\{\check{\mathbf{s}}_2\} \mathbf{F}_{[\mathcal{K}_2, 1:2L+1]} \mathbf{b}}_{\approx \mathbf{0}} \\ &\quad + \underbrace{\mathbf{G}_{[\overline{\mathcal{K}}_1, :]} \boldsymbol{\Gamma} [w] \mathbf{P} \mathbf{h}_1 + \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{n}_e}_{\mathbf{C}_2}, \end{aligned} \quad (29)$$

where \mathbf{C}_2 is an $(N - K_1) \times (L + 1)$ matrix, and the first term is negligible because $\nu \approx 0$. As long as $(N - K_1 - L - 1) \geq 1$, the CFO between the relay and the first source terminal can be obtained as

$$\hat{w} = \arg \max_w \mathbf{y}^H \mathbf{F}_{[\overline{\mathcal{K}}_1, :]}^H \mathbf{C}_2 \left(\mathbf{C}_2^H \mathbf{C}_2 \right)^{-1} \mathbf{C}_2^H \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{y}, \quad (30)$$

$$\hat{\mathbf{h}}_1 = \mathbf{C}_2^\dagger \mathbf{F}_{[\overline{\mathcal{K}}_1, :]} \mathbf{y}.$$

Since \mathbf{a} can be estimated from \hat{w} and $\hat{\mathbf{h}}_1$, then \mathbf{h}_2 can be estimated as

$$\hat{\mathbf{h}}_2 = \mathbf{C}_3^\dagger \left(\mathbf{y} - \alpha \mathbf{S}_1 \hat{\mathbf{a}} - \boldsymbol{\Gamma} [\hat{w}] \mathbf{P} \hat{\mathbf{h}}_1 \right), \quad (31)$$

where $\mathbf{C}_3 = \alpha \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} \left[\boldsymbol{\Omega}^{(L+1)} [\hat{\nu} - \hat{w}] \hat{\mathbf{h}}_1 \right]$.

The estimates given by (30) and (31) need not be improved by iterations because they already achieve CRB under high SNR conditions.

3.4. Iterative Algorithm to Improve the Performance. With the initial estimates of all the parameters, an iterative approach can be applied to improve the estimation accuracy. Re-denote the initial estimate as $\nu^{(0)}$, $w^{(0)}$, $\mathbf{a}^{(0)}$, $\mathbf{b}^{(0)}$, $\mathbf{h}_1^{(0)}$, $\mathbf{h}_2^{(0)}$, respectively, with the superscript representing the number of the iterations. We will estimate $\nu^{(1)}$, $w^{(1)}$ simultaneously from the ML estimation process as

$$\begin{aligned} & \left[\nu^{(1)}, w^{(1)} \right] \\ &= \arg \min_{\nu, w} \left(\mathbf{y} - \alpha \mathbf{S}_1 \mathbf{a}^{(0)} - \alpha \boldsymbol{\Gamma} [\nu] \mathbf{S}_2 \mathbf{b}^{(0)} - \boldsymbol{\Gamma} [w] \mathbf{P} \mathbf{h}_1^{(0)} \right)^H \mathbf{R}_n^{-1} \\ & \quad \times \left(\mathbf{y} - \alpha \mathbf{S}_1 \mathbf{a}^{(0)} - \alpha \boldsymbol{\Gamma} [\nu] \mathbf{S}_2 \mathbf{b}^{(0)} - \boldsymbol{\Gamma} [w] \mathbf{P} \mathbf{h}_1^{(0)} \right), \end{aligned} \quad (32)$$

where \mathbf{R}_n^{-1} is always obtained by using the newest estimates of w and \mathbf{h}_1 in (10) expressed as

$$\begin{aligned} \mathbf{R}_n &= \sigma_n^2 \left(\alpha^2 \boldsymbol{\Gamma}^{(N)} [w^{(0)}] \mathbf{H}_{cv}^{(N)} [\mathbf{h}_1^{(0)}] \right. \\ & \quad \left. \times \left(\mathbf{H}_{cv}^{(N)} [\mathbf{h}_1^{(0)}] \right)^H \left(\boldsymbol{\Gamma}^{(N)} [w^{(0)}] \right)^H + \mathbf{I} \right). \end{aligned} \quad (33)$$

The complexity here is not significant even if the 2-dimensional search is applied, since the search region for fine estimation is around the initial estimation and is thus very small.

Then we can obtain $\mathbf{h}_2^{\{1\}}$ and $\mathbf{h}_1^{\{1\}}$ from

$$\begin{aligned} \mathbf{h}_2^{\{1\}} &= \arg \min_{\mathbf{h}_2} \left(\mathbf{y} - \alpha \mathbf{S}_1 \mathbf{a}^{(0)} - \boldsymbol{\Gamma} [w^{\{1\}}] \mathbf{P} \mathbf{h}_1^{\{0\}} \right. \\ & \quad \left. - \alpha \boldsymbol{\Gamma} [\nu^{\{1\}}] \mathbf{S}_2 \mathbf{H}_{12}^{\{1\}} \mathbf{h}_2 \right)^H \mathbf{R}_n^{-1} \\ & \quad \times \left(\mathbf{y} - \alpha \mathbf{S}_1 \mathbf{a}^{(0)}, \right. \\ & \quad \left. - \boldsymbol{\Gamma} [w^{\{1\}}] \mathbf{P} \mathbf{h}_1^{\{0\}} - \alpha \boldsymbol{\Gamma} [\nu^{\{1\}}] \mathbf{S}_2 \mathbf{H}_{12}^{\{1\}} \mathbf{h}_2 \right) \\ \mathbf{h}_1^{\{1\}} &= \arg \min_{\mathbf{h}_1} \left(\mathbf{y} - \left[\alpha \mathbf{S}_1 \mathbf{H}_{11}^{\{0\}} + \alpha \boldsymbol{\Gamma} [\nu^{\{1\}}] \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} \right] \mathbf{h}_2^{\{1\}} \right. \\ & \quad \times \boldsymbol{\Omega}^{L+1} [\nu^{\{1\}} - w^{\{1\}}] + \boldsymbol{\Gamma} [w^{\{1\}}] \mathbf{P} \mathbf{h}_1 \Big)^H \\ & \quad \times \mathbf{R}_n^{-1} \left(\mathbf{y} - \left[\alpha \mathbf{S}_1 \mathbf{H}_{11}^{\{0\}} + \alpha \boldsymbol{\Gamma} [\nu^{\{1\}}] \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} \right] \mathbf{h}_2^{\{1\}} \right. \\ & \quad \left. \times \boldsymbol{\Omega}^{L+1} [\nu^{\{1\}} - w^{\{1\}}] + \boldsymbol{\Gamma} [w^{\{1\}}] \mathbf{P} \mathbf{h}_1 \right), \end{aligned} \quad (34)$$

where $\mathbf{H}_{12}^{\{1\}} = \mathbf{H}_{zp}^{(L+1)} \left[\boldsymbol{\Omega}^{(L+1)} [\nu^{\{1\}} - w^{\{1\}}] \mathbf{h}_1^{\{0\}} \right]$, and $\mathbf{H}_{11}^{\{0\}} = \mathbf{H}_{zp}^{(L+1)} \left[\boldsymbol{\Omega}^{(L+1)} [-w^{\{1\}}] \mathbf{h}_1^{\{0\}} \right]$.

The interactive processing could gain the improvement from the fact that the initial estimation does not fully exploit the correlation between \mathbf{a} , \mathbf{b} , and \mathbf{h}_1 .

Remark 3. The superimposed training used in our paper is different from the traditional ones in two aspects. First, traditional superimposed pilots [14] are added on the data symbols while our superimposed pilots are added on the

pilots from the source by the relay. On the other hand, traditional superimposed pilots [15] select the first order statistics for channel estimation while we use the nulling-based LS method for joint CFO and channel estimation.

4. Cramér-Rao Bound

In this section, we derive the CRB that defines the theoretical bound of the estimation accuracy. The CRB is an important tool to study the performance of the estimation algorithms.

Define

$$\begin{aligned} \boldsymbol{\mu} &\triangleq \alpha \mathbf{S}_1 \mathbf{a} + \alpha \Gamma[v] \mathbf{S}_2 \mathbf{b} + \Gamma[w] \mathbf{P} \mathbf{h}_1, \\ \boldsymbol{\eta} &\triangleq [v, w, \Re\{\mathbf{h}_1\}^T, \Im\{\mathbf{h}_1\}^T, \Re\{\mathbf{h}_2\}^T, \Im\{\mathbf{h}_2\}^T]^T. \end{aligned} \quad (35)$$

According to [20], the (i, j) th entry of the Fisher Information Matrix (FIM) can be calculated as

$$[\mathbf{F}]_{i,j} = 2\Re\left[\frac{\partial \boldsymbol{\mu}^H}{\partial \eta_i} \mathbf{R}_n^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \eta_j}\right] + \text{tr}\left[\mathbf{R}_n^{-1} \frac{\partial \mathbf{R}_n}{\partial \eta_i} \mathbf{R}_n^{-1} \frac{\partial \mathbf{R}_n}{\partial \eta_j}\right]. \quad (36)$$

After some tedious simplifications, we derive

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}}{\partial v} &= j\alpha \mathbf{D}_1 \Gamma[v] \mathbf{S}_2 \mathbf{b} + j\Gamma[v] \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} \\ &\quad \times [\mathbf{h}_2] \mathbf{D}_0 \boldsymbol{\Omega}^{(L+1)} [v - w] \mathbf{h}_1, \\ \frac{\partial \boldsymbol{\mu}}{\partial w} &= \alpha \mathbf{S}_1 \mathbf{H}_{zp}^{(L+1)} [\mathbf{h}_1] (-j\mathbf{D}_0) \boldsymbol{\Omega}^{(L+1)} [-w] \mathbf{h}_1 \\ &\quad + \alpha \Gamma[v] \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} [\mathbf{h}_2] (-j\mathbf{D}_0) \boldsymbol{\Omega}^{(L+1)} [v - w] \mathbf{h}_1 \\ &\quad + j\mathbf{D}_1 \Gamma[w] \mathbf{P} \mathbf{h}_1, \\ \frac{\partial \boldsymbol{\mu}}{\partial \Re\{\mathbf{h}_1\}^T} &= \alpha \mathbf{S}_1 \mathbf{H}_{zp}^{(L+1)} [\boldsymbol{\Omega}^{(L+1)} [-w] \mathbf{h}_1] \\ &\quad + \alpha \mathbf{S}_1 \mathbf{H}_{zp}^{(L+1)} [\mathbf{h}_1] \boldsymbol{\Omega}^{(L+1)} [-w] \\ &\quad + \alpha \Gamma[v] \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} [\mathbf{h}_2] \boldsymbol{\Omega}^{(L+1)} [v - w] + \Gamma[w] \mathbf{P}, \\ \frac{\partial \boldsymbol{\mu}}{\partial \Re\{\mathbf{h}_2\}^T} &= \alpha \Gamma[v] \mathbf{S}_2 \mathbf{H}_{zp}^{(L+1)} [\boldsymbol{\Omega}^{(L+1)} [v - w] \mathbf{h}_1], \\ \frac{\partial \mathbf{R}_n}{\partial w} &= \sigma_n^2 \left(j\alpha^2 \mathbf{D}_N \boldsymbol{\Gamma}^{(N)} [w] \mathbf{H}_{cv}^{(N)} [\mathbf{h}_1] \left(\mathbf{H}_{cv}^{(N)} [\mathbf{h}_1] \right)^H \right. \\ &\quad \times \boldsymbol{\Gamma}^{(N)} [-w] - j\alpha^2 \boldsymbol{\Gamma}^{(N)} [w] \mathbf{H}_{cv}^{(N)} [\mathbf{h}_1] \\ &\quad \left. \times \left(\mathbf{H}_{cv}^{(N)} [\mathbf{h}_1] \right)^H \mathbf{D}_N \boldsymbol{\Gamma}^{(N)} [-w] \right) \\ \frac{\partial \mathbf{R}_n}{\partial \Re\{\mathbf{h}_{1i}\}^T} &= \sigma_n^2 \alpha^2 \boldsymbol{\Gamma}^{(N)} [w] \mathbf{H}_{cv}^{(N)} [\mathbf{e}_{1i}] \left(\mathbf{H}_{cv}^{(N)} [\mathbf{h}_1] \right)^H \boldsymbol{\Gamma}^{(N)} [-w] \\ \frac{\partial \mathbf{R}_n}{\partial \Re\{\mathbf{h}_{2i}\}^T} &= \frac{\partial \mathbf{R}_n}{\partial v} = \mathbf{0}_{N \times N}, \end{aligned} \quad (37)$$

where

$$\begin{aligned} \mathbf{D}_0 &= 2\pi T_s \text{diag}\{L, (L-1), \dots, 1, 0\}, \\ \mathbf{D}_1 &= 2\pi T_s \text{diag}\{L, \dots, (L+N-1)\}, \\ \mathbf{D}_N &= 2\pi T_s \text{diag}\{0, 1, \dots, (N-1)\}, \end{aligned} \quad (38)$$

and \mathbf{e}_i is an $(L+1) \times 1$ vector whose i th element equals 1 and others 0.

The CRB of $\boldsymbol{\eta}$ is then obtained by inverting the FIM \mathbf{F} . Since

$$\begin{bmatrix} v \\ w \\ \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1, 0, & \mathbf{0}^T, & \mathbf{0}^T, & \mathbf{0}^T, & \mathbf{0}^T \\ 0, 1, & \mathbf{0}^T, & \mathbf{0}^T, & \mathbf{0}^T, & \mathbf{0}^T \\ \mathbf{0}, \mathbf{0}, & \mathbf{I}_{L+1}, & j\mathbf{I}_{L+1}, & 0 \cdot \mathbf{I}_{L+1}, & 0 \cdot \mathbf{I}_{L+1} \\ \mathbf{0}, \mathbf{0}, & 0 \cdot \mathbf{I}_{L+1}, & 0 \cdot \mathbf{I}_{L+1}, & \mathbf{I}_{L+1}, & j\mathbf{I}_{L+1} \end{bmatrix}}_{\boldsymbol{\Xi}} \times \boldsymbol{\eta}, \quad (39)$$

the CRB of v , w , \mathbf{h}_1 , and \mathbf{h}_2 can be expressed as

$$\text{CRB} = \boldsymbol{\Xi} \cdot \mathbf{F}^{-1} \cdot \boldsymbol{\Xi}^H. \quad (40)$$

5. Simulation Results

The performance of the proposed three estimation algorithms along with the iterative estimator is investigated. A four-tap model for both \mathbf{h}_1 and \mathbf{h}_2 is assumed, and each tap is assumed complex Gaussian with unit variance as did in [21]. The variance of the noise is taken as $\sigma_n^2 = 1$. The normalized frequencies f_1 , f_r , and f_2 are set as 0.94, 1, and 1.06, respectively. The MSE is chosen as the figure of merit, defined by

$$\begin{aligned} \text{MSE}(v) &= \frac{1}{10000} \sum_{i=1}^{10000} (\hat{v}_i - v)^2, \\ \text{MSE}(w) &= \frac{1}{10000} \sum_{i=1}^{10000} (\hat{w}_i - w)^2, \\ \text{MSE}(\mathbf{x}) &= \frac{1}{10000} \sum_{i=1}^{10000} \frac{1}{4} (\hat{\mathbf{x}}_i - \mathbf{x})^2, \end{aligned} \quad (41)$$

where \mathbf{x} represents \mathbf{h}_1 or \mathbf{h}_2 , and 10000 is the number of the Monte-Carlo trials used for average. In all the following simulations, α is set as half the maximum value, that is, $\alpha = 0.5\sqrt{P_r / (\sum_{j=1}^2 \sum_{l=0}^L \sigma_{j,l}^2 P_j + \sigma_n^2)}$.

5.1. Sufficiently Large N . In this case, we chose $N = 24$, which is greater than $5L + 5 = 20$. The received signal \mathbf{y} at \mathbb{T}_1 is generated according to (9). Initial CFO and channel estimates are obtained from \mathbf{y} through (17) and (18). The estimate $\hat{\mathbf{h}}_1$ of \mathbf{h}_1 is updated as (20). Finally, the iterative estimator in Section 3.4 is applied and is found to converge in three iterations.

The MSEs and CRBs of CFO estimation as a function of SNR are shown in Figure 2. The iterative algorithm improves the estimation accuracy significantly. Specifically,

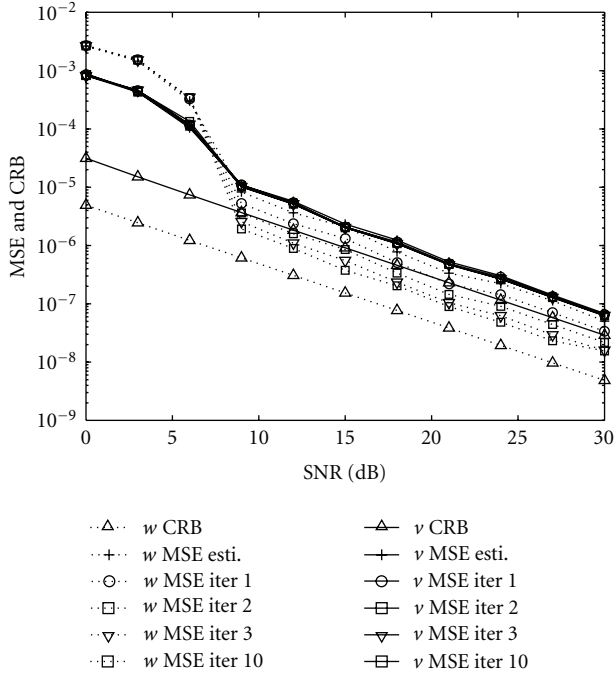


FIGURE 2: CFO estimation MSE versus SNR: $N = 24$.

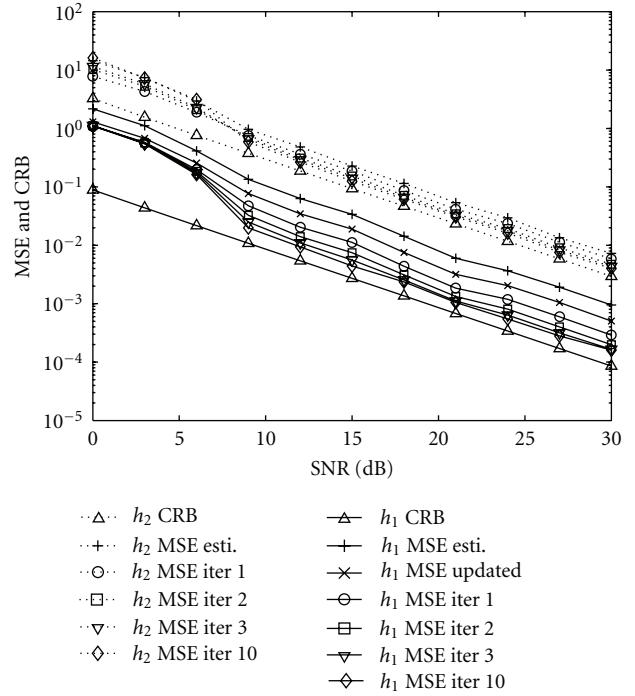


FIGURE 3: Channel estimation MSE versus SNR: $N = 24$.

the improvement of w , which is the CFO between \mathbb{R} and \mathbb{T}_1 , is much more significant than that of v , the CFO between \mathbb{T}_2 and \mathbb{T}_1 . The reason is due to the fact that the received signals contain more information about w than that about v . In (14), all components of \mathbf{y} contain information about w while only the second term of \mathbf{y} does so of v .

The channel estimation MSEs and CRBs versus SNR are shown in Figure 3. It is observed that the gaps between the MSE and the CRB are smaller compared to those in the CFO estimation. The reason is that phase errors have less effect on the channel estimation than on the CFO estimation. Similarly to CFO estimation, iteration improves the estimation accuracy, and \mathbf{h}_1 improves more than \mathbf{h}_2 since most components of \mathbf{y} contain the information of \mathbf{h}_1 .

5.2. Not-So-Large N . Next we choose $N = 3L + 5 = 14$, and $\mathcal{K}_1 = 4, \mathcal{K}_2 = 4, \mathcal{K}_r = 6$. The initial CFO and channel estimates are obtained from (25), (26), and (28). The estimates are iteratively updated. We find that ten iterations reach convergence. The MSEs and CRBs versus SNRs for both CFO and channel estimation are displayed in Figures 4 and 5, respectively.

Since smaller training length is applied and the average symbol power is kept the same, the performance here is a little worse than that in Figures 2 and 3. It is observed that in the high SNR region, the MSEs approach CRBs after ten iterations. The iterative estimator improves the estimation of \mathbf{h}_1 for all SNRs while, for \mathbf{h}_2 , it only works at the high SNR region. Conversely, the iterative estimator degrades the estimation of \mathbf{h}_2 at low SNRs. A possible reason is as follows. Since only the second item of \mathbf{y} in (14) contains \mathbf{h}_2 , and iterations require reconstruction of \mathbf{a} from the initial

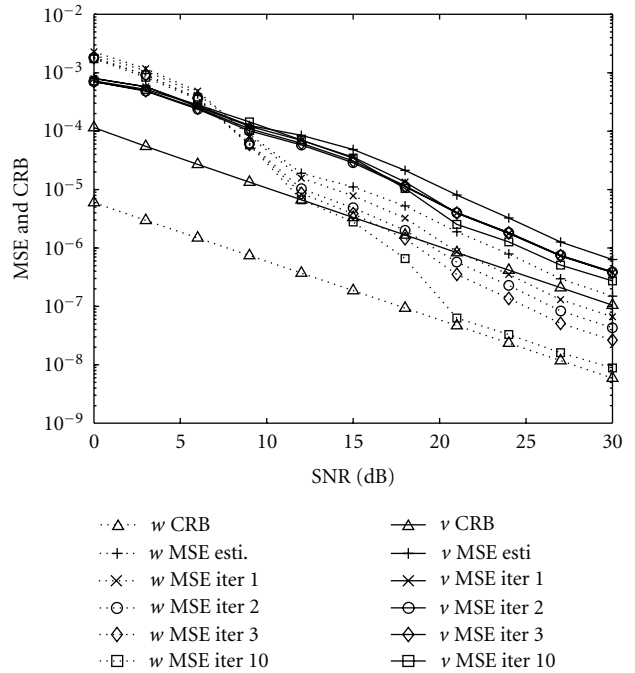
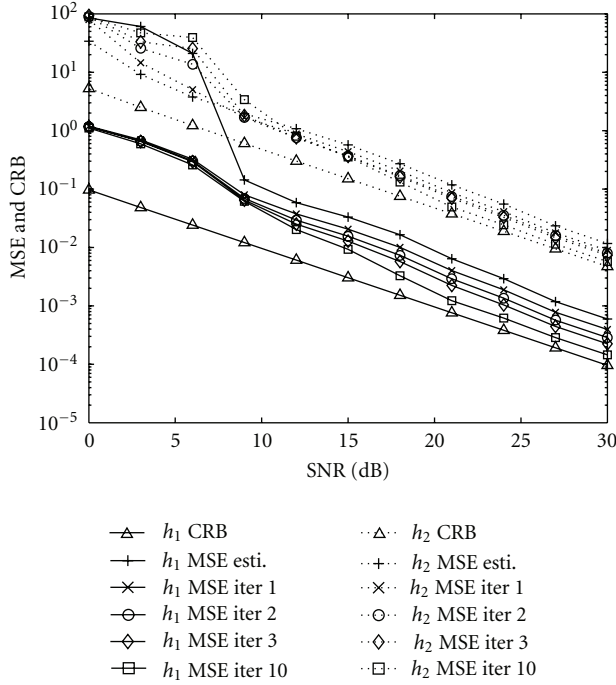


FIGURE 4: CFO estimation MSE versus SNR: $N = 14$.

estimate \hat{w} and $\hat{\mathbf{h}}_1$, the ambiguity of $\hat{\mathbf{h}}_2$ increased at the low SNR region from errors of all factors. However, in the case of large training length, for example, $N = 24$, \mathbf{a} is directly estimated, which avoids the error propagation from erroneous \hat{w} and $\hat{\mathbf{h}}_1$. Therefore similar phenomenon is not observed in Figure 3.

FIGURE 5: Channel estimation MSE versus SNR: $N = 14$.

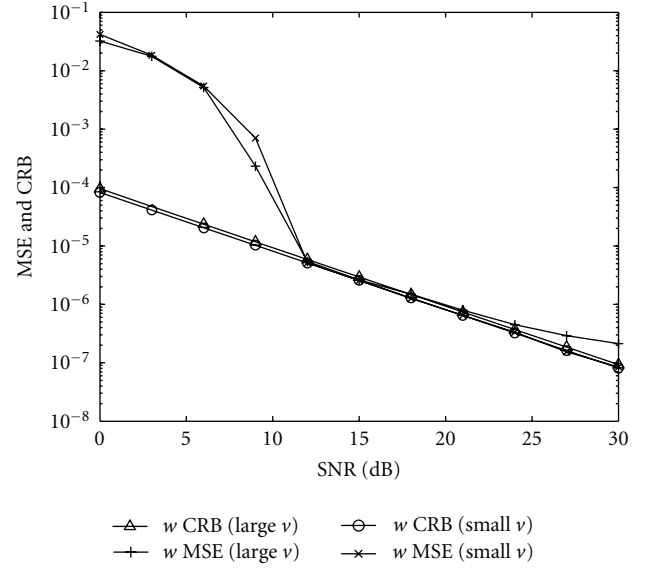
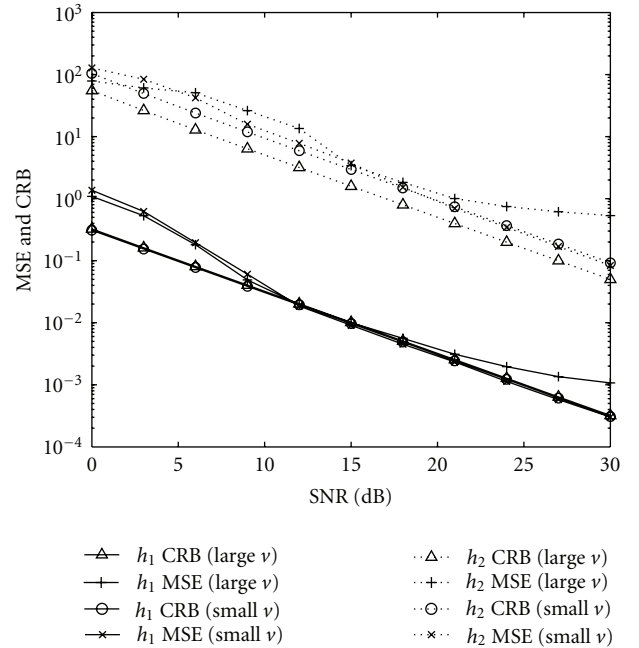
From Figure 2 to Figure 5, one key conclusion is that there is room for new algorithms to improve the performance, due to the gap between MSEs and CRBs. This open question is an interesting topic for the future research.

5.3. Minimum N : A Special Case. In the last example, we set f_2 as 0.95 and 0.9401, such that the CFOs between the two terminals are $\nu = 0.01$ and $\nu = 0.0001$, respectively. The minimum training length is chosen as $N = 2L + 3 = 9$. The CFO and channel estimation results can be obtained from (30) and (31). As mentioned before, these cases do not require iterations because ν is nearly zero. The estimation MSEs of w and channels, as well as their corresponding CRBs are shown in Figures 6 and 7, respectively.

These figures show that the estimation accuracy is quite close to CRB. The reason is that the interference due to the pilots is negligible in this case as the CFO between the two source terminals is negligible. For $\nu = 0.01$, a relatively larger value, there exists an error floor for both CFO and channel estimation at the high SNR region. When ν is as small as 0.0001, the best estimation performance can be achieved since the MSE attaches the CRB.

6. Conclusions

In this paper, superimposed pilot-based CFO and channel estimation was investigated for CP-OFDM modulated TWRN. Three direct estimation algorithms as well as the iteration algorithm to improve the performance were developed. We also derived the analytical CRBs as the benchmark for the designed algorithms. With superimposed training, all the individual parameters can be estimated at all three nodes.

FIGURE 6: CFO estimation MSE versus SNR: $N = 9$.FIGURE 7: Channel estimation MSE versus SNR: $N = 9$.

From the simulations, it is found that although the iterative estimator improves the performance, gaps remain between the MSE and CRB, indicating room for further improving the performance.

Acknowledgment

This work was supported in part by the German Research Foundation (DFG) under Grant GA 1654/1-1.

References

- [1] S. G. S. Katti and D. Katabi, "Embracing wireless interference: analog network coding," Tech. Rep., Computer Science and Artificial Intelligence Laboratory, February 2007.
- [2] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.
- [3] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.
- [4] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Proceedings of the Annual Conference on Signals, Systems and Computers*, vol. 2005, pp. 1066–1071, Pacific Grove, Calif, USA, October 2005.
- [5] B. Rankov and A. Wittneben, "Achievable rate regions for the two-way relay channel," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '06)*, pp. 1668–1672, Seattle, Wash, USA, July 2006.
- [6] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions for bi-directional relaying," submitted to *IEEE Transactions on Information Theory*.
- [7] T. Cui, T. Ho, and J. Kliewer, "Memoryless relay strategies for two-way relay channels: performance analysis and optimization," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 1139–1143, Beijing, China, May 2008.
- [8] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 3888–3892, Beijing, China, May 2008.
- [9] 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 Journal on Selected Areas in Communications*, vol. 27, no. 5, Article ID 5072356, pp. 699–712, 2009.
- [10] K. H. Chin, R. Zhang, and Y.-C. Liang, "Two-way relaying over OFDM: optimized tone permutation and power allocation," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 3908–3912, Beijing, China, May 2008.
- [11] G. Wang, F. Gao, Y. Wu, and C. Tellambura, "Joint CFO and channel estimation for CP-OFDM modulated two-way relay networks," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 1–5, Cape Town, South Africa, June 2010.
- [12] G. Wang, F. Gao, Y. Wu, and C. Tellambura, "Joint CFO and channel estimation for ZP-OFDM modulated two-way relay networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '10)*, pp. 1–5, Sydney, Australia, April 2010.
- [13] C. Kastenholz and W. Birkemeier, "A simultaneous information transfer and channel-sounding modulation technique for wide-band channels," *IEEE Transactions on Communications*, vol. 13, no. 2, pp. 162–165, 1965.
- [14] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Communications Letters*, vol. 7, no. 9, pp. 413–415, 2003.
- [15] M. Ghogho, D. McLernon, E. Alameda-Hernandez, and A. Swami, "Channel estimation and symbol detection for block transmission using data-dependent superimposed training," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 226–229, 2005.
- [16] G. Wang and C. Tellambura, "Super-imposed pilot-aided channel estimation and power allocation for relay systems," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, pp. 1–6, Budapest, Hungary, April 2009.
- [17] Z. Wang, Y. Xin, and G. Mathew, "Iterative carrier-frequency offset estimation for generalized OFDMA uplink transmission," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1373–1383, 2009.
- [18] X. N. Zeng and A. Ghayeb, "Joint CFO and channel estimation for OFDMA uplink: an application of the variable projection method," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2306–2311, 2009.
- [19] M.-O. Pun, M. Morelli, and C.-C. J. Kuo, "Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 726–736, 2006.
- [20] M. Kay, *Fundamentals of Statistical Signal Processing*, vol. 1 of *Estimation Theory*, Prentice Hall, New York, NY, USA, 1993.
- [21] F. Gao, R. Zhang, and Y.-C. Liang, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4443–4455, 2009.