Channel Estimation and Carrier Recovery in the Presence of Phase Noise in OFDM Relay Systems

Rui Wang*, Hani Mehrpouyan†, Meixia Tao**, and Yingbo Hua‡
Department of Information and Communications, Tongji University, Shanghai, China.
†Department of Computer and Electrical Engineering and Computer Science, California State University, Bakersfield, CA, USA.
**Department of Electronic Engineering at Shanghai Jiao Tong University, Shanghai, China.
‡Department of Electrical Engineering at the University of California, Riverside, CA, USA.
Emails: liouxingrui@gmail.com, hani.mehr@ieee.org, mxtao@sjtu.edu.cn, yhua@ee.ucr.edu.

***Abstract*—In this paper, we analyze joint channel, *carrier frequency offset (CFO)*, and *phase noise* estimation in *orthogonal frequency division multiplexing (OFDM)* relaying networks. To achieve this goal, a detailed transmission framework involving both training and data symbols is first presented. Next, a novel algorithm that applies the training symbols to jointly estimate the channel responses, CFO, and phase noise parameters based on the maximum a posteriori criterion is proposed. Additionally, to evaluate the performance of the proposed channel estimation and carrier recovery algorithms, we analyze the ambiguities among the estimated parameters. Based on this analysis, a new Hybrid Cramér-Rao Lower Bound (HCRLB) is derived, which can effectively avoid such ambiguities. The simulation results show that the proposed estimation algorithm can achieve a performance close to the derived HCRLB.**

I. Introduction

Application of relaying has been identified as a suitable approach for combating the long-distance channel distortion and the small-scale channel fading in wireless systems. Various physical layer techniques, such as distributed space-time block coding, relay precoding, etc., for relay systems have been extensively studied recently. From these works, it can be deduced that to deliver the advantages of relay networks, the network's *channel state information (CSI)* needs to be accurately obtained [1]. Moreover, due to oscillator imperfections and Doppler shifts, wireless communication systems are affected by *carrier frequency offset (CFO)* and *phase noise (PN)*, which need to be accurately estimated and compensated.

Joint estimation of the channel and CFO in single carrier relay systems has been studied in [2]. However, this work ignores the effect of PN. Although both CFO and PN result in an unknown rotation of the signal constellation and inter carrier inference (ICI), the estimation of PN parameters is more difficult due to the time varying nature of this parameter.

Due to its capability of combating frequency selectivity in the wireless channel, OFDM techniques have been extensively adopted in the latest wireless communication standards, e.g., Long Term Evolution, IEEE 802.11ac, Bluetooth, etc. The deteriorating effect of PN on the performance of point-to-point OFDM systems are analyzed in [3]. Undoubtedly, this effect can also be observed in OFDM based relay systems. Hence, conducting accurate channel and CFO estimation in the presence of PN is important for maintaining the quality of service in high-speed OFDM relay networks. Joint estimation of CFO and channel in OFDM relay systems is considered in [4], [5]. In particular, a two-time-slot cooperative estimation protocol has been proposed in [4] for both amplify-and-forward (AF) and decode-and-forward (DF) OFDM relaying systems. Moreover, in [5] the maximum likelihood (ML) and least squares based algorithms for joint CFO and channel estimation in relay networks are studied. However, none of the approaches in [4], [5] consider the effect of PN on channel and CFO estimation or the overall relaying performance.

In this paper, different from [4], [5], the problem of joint CFO and channel estimation in the presence of PN for OFDM relay systems is considered. Although joint CFO, PN, and channel estimation has been studied for point-to-point OFDM systems [6], [7], to our best knowledge, this problem has rarely been considered in the context of relay systems. The contributions of this paper can be summarized as follows:

- A training and data transmission framework for OFDM relay networks is proposed that enables joint estimation of channel, CFO, and PN parameters at the destination.
- The ambiguities amongst the estimated PN, CFO, and channels are analyzed. Based on this analysis, a new *hybrid Cramer-Rao lower bound (HCRLB)* for analyzing the performance of joint channel, CFO, and PN estimators in OFDM relay networks is derived, which can effectively avoid the estimation ambiguities.
- An iterative joint channel, CFO, and PN estimator based on the *maximum a posteriori (MAP)* criterion is proposed that exploits the correlation between PN parameters to significantly reduce estimation overhead. Moreover, the estimator's mean square error (MSE) performance is shown to be close to the derived HCRLB at moderate signal-to-noise ratios (SNRs).

Notations: $\odot$ denotes the Hadamard product. $\text{Diag}(\mathbf{a})$ denotes a diagonal matrix with $\mathbf{a}$ being its diagonal entries. $\text{Blkdiag}(\mathbf{A}_0, \mathbf{A}_1, \cdots, \mathbf{A}_{N-1})$ denotes a block diagonal matrix with $\mathbf{A}_0, \mathbf{A}_1, \cdots, \mathbf{A}_{N-1}$ as its diagonal matrices. $\Re(z)$ and $\Im(z)$ denote the real and imaginary operators, respectively.

II. System Model

We consider an AF OFDM relaying system with N subcarriers. A source node transmits its signal to a destination node through a relay node. Quaisi-static channels are considered, i.e., the CSI is assumed to be constant over the duration of a

This work is supported by the National 973 project of China under grant 2012CB316100, and the NSF of China under grants 61322102 and 61221001.

single packet. Each packet consists of two OFDM training symbols, which are followed by multiple data symbols as shown in Fig. 1. The two training symbols are used to separately estimate the channel responses and CFO in the presence of unknown PN for each hop: one from the source and one from the relay node. The proposed signal model can be applied to both full-duplex and half-duplex relaying networks based on the following system setups and assumptions: 1) *Full-duplex relaying*: In this setup, the proposed signal model is applicable to relaying networks that utilize highly directional transmit and receive antennas with large antenna gains at the relay, e.g., microwave and millimeter-wave systems [8]. This approach minimizes or eliminates the effect of self-interference at the relay.[1] Moreover, it is assumed that the relay forwards its signal to the destination in passband without converting it to baseband. This assumption is practical since there are various radio frequency (RF) amplifiers that can operate at high carrier frequencies and can be utilized in full-duplex relaying networks, e.g., Mini-Circuits AVA-24+ with a frequency range of 5–20 GHz [10]. 2) *Half-duplex relaying*: In this setup, it is assumed that the relay forwards its received signal on a different carrier frequency and does not convert it to baseband, i.e., the relay applies on-frequency/on-channel RF relaying [11]. Moreover, the difference between the receive and transmit carrier frequencies are assumed to be small to enable the application of a low PN oscillator at the relay. An example of such an oscillator is ROS-209-319+ ultra low noise voltage controlled oscillator that has a very small PN factor of -133 dBc/Hz at an offset frequency of 10 KHz [12]. As such, in this setup, it is assumed that the signal forwarded from the relay is not affected by PN.

A. Signal Transmission from Source to Destination

Let $\mathbf{s}^{[\text{s}]} \triangleq [s_0^{[\text{s}]}, s_1^{[\text{s}]}, \cdots, s_{N-1}^{[\text{s}]}]^T$ denote the frequency domain modulated training at the source node as shown in Fig. 2. This signal is then transformed into a set of parallel symbols $s_k^{[\text{s}]}$, for $k = 0, \cdots, N-1$. By conducting an inverse fast Fourier transform (IFFT), we obtain the time domain signal vector $\mathbf{x}^{[\text{s}]}$ as $\mathbf{x}^{[\text{s}]} = \mathbf{F}^H \mathbf{s}^{[\text{s}]}$, where $\mathbf{x}^{[\text{s}]} \triangleq [x_0^{[\text{s}]}, x_1^{[\text{s}]}, \cdots, x_{N-1}^{[\text{s}]}]^T$, and $\mathbf{F}$ is the normalized discrete Fourier transform (DFT) matrix. After adding the cyclic prefix (CP), the parallel signal vector is transformed into a time domain sequence denoted by $x^{[\text{s}]}(n)$, for $n = -L, \cdots, N-1$. Assume that the transmitted baseband continuous signal from the source is $\tilde{x}^{[\text{s}]}(t)$, the received baseband signal at the destination is denoted as

$$\begin{aligned} y^{[\text{s}]}(t) =& \alpha e^{j\theta^{[\text{s-d}]}(t)} e^{j\phi^{[\text{s-d}]}(t)} \big[g(t) \star h(t) \star \tilde{x}^{[\text{s}]}(t) \\ &+ g(t) \star v(t) \big] + w(t), \end{aligned} \tag{1}$$

where $\star$ denotes the linear convolution, α is the constant and scalar amplification factor at the relay, $h(t)$ and $g(t)$ are the frequency-selective fading channels from source to relay and the relay to destination, respectively, and $v(t)$ and $w(t)$ are

[1]Application of sophisticated transceivers has also been shown to minimize or eliminate the impact of self-interference at the relay [9].

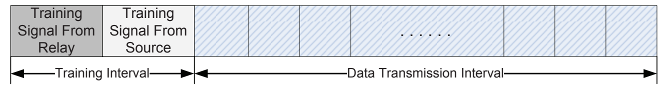

Fig. 1. Illustration of the timing diagram of the OFDM relay system.

the additive noises at the relay node and at the destination node, respectively. $\theta^{[\text{s-d}]}(t)$ is the PN corresponding to source-relay-destination link, while $\phi^{[\text{s-d}]}(t) \triangleq 2\pi \triangle f^{[\text{s-d}]} t$ is the CFO caused by the unmatched source and destination carrier frequencies.

After sampling at a sampling rate of $1/T_s$ and removing the CP, the received signal at the destination, $y^{[\text{s}]}(t)$, is given by

$$\begin{aligned} y^{[\text{s}]}(nT_s) =& \alpha e^{j\theta^{[\text{s-d}]}(nT_s)} e^{j2\pi\Delta f^{[\text{s-d}]} nT_s} \big[\underbrace{g(nT_s) \star h(nT_s)}_{\triangleq c(nT_s)} \\ & * \tilde{x}^{[\text{s}]}(nT_s) + g(nT_s) \star v(nT_s) \big] + w(nT_s) \end{aligned} \tag{2}$$

where circular convolution $*$ appears in (2) due to the added CP at the source node. To avoid ICI, the length of CP, N_{CP}, should be larger than $L = L_h + L_g - 1$ with L_h and L_g being the number of channel taps of $h(t)$ and $g(t)$, respectively. Eq. (2) can be written in vector form as

$$\mathbf{y}^{[\text{s}]} = \alpha \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \mathbf{\Lambda}_{\phi^{[\text{s-d}]}} \left[\mathbf{C}\mathbf{x}^{[\text{s}]} + \mathbf{G}\mathbf{v} \right] + \mathbf{w}, \tag{3}$$

where $\mathbf{y}^{[\text{s}]} \triangleq [y^{[\text{s}]}(0), y^{[\text{s}]}(1), \cdots, y^{[\text{s}]}(N-1)]^T$, $\mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \triangleq \text{Diag}\big[e^{j\theta^{[\text{s-d}]}(0)}, e^{j\theta^{[\text{s-d}]}(1)}, \cdots, e^{j\theta^{[\text{s-d}]}(N-1)}\big]$, $\mathbf{\Lambda}_{\phi^{[\text{s-d}]}} \triangleq \text{Diag}\big[1, e^{j2\pi\phi^{[\text{s-d}]}/N}, \cdots, e^{j2\pi\phi^{[\text{s-d}]}(N-1)/N}\big]$, $\phi^{[\text{s-d}]} \triangleq \Delta f^{[\text{s-d}]} T$ is the normalized CFO, $\mathbf{C} \triangleq \mathbf{F}^H \mathbf{\Lambda}_{\tilde{\mathbf{c}}} \mathbf{F}$, $\mathbf{\Lambda}_{\tilde{\mathbf{c}}} \triangleq \text{Diag}(\tilde{\mathbf{c}})$ with $\tilde{\mathbf{c}} \triangleq \sqrt{N}\mathbf{F}[\mathbf{c}^T, \mathbf{0}_{N-L,1}^T]^T$ and $\mathbf{c} \triangleq [c(0), c(1), \cdots, c(L-1)]^T$, $\mathbf{v} \triangleq [v(-L_g+1), \cdots, v(0), \cdots, v(N-1)]^T$ and $\mathbf{w}_1 \triangleq [w(0), w(1), \cdots, w(N-1)]^T$ are the sampled additive noise at the relay and destination, respectively, and

$$\mathbf{G} = \begin{bmatrix} g(L_g-1) & g(L_g-2) & \cdots & 0 \\ 0 & g(L_g-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(0) \end{bmatrix} \tag{4}$$

is an $N \times (N + L_g - 1)$ matrix. The additive noise at the relay and destination are distributed as $\mathbf{v} \sim \mathcal{CN}(0, \sigma_R^2 \mathbf{I}_{N+N_{CP}})$ and $\mathbf{w} \sim \mathcal{CN}(0, \sigma_D^2 \mathbf{I}_N)$, respectively. Although $\mathbf{C}$ is an $N \times N$ circulant matrix, $\mathbf{G}$ is a regular $N \times (N + L_g - 1)$ matrix, since no CP is added at the relay node.

B. Training Signal Transmission from Relay to Destination

Recall that the second OFDM training symbol is transmitted from the relay to separately estimate the channels corresponding to the second hop. Following similar steps as above, the vector of received training signal at the destination node from the relay, $\mathbf{y}^{[\text{r}]} \triangleq [y^{[\text{r}]}(0), y^{[\text{r}]}(1), \cdots, y^{[\text{r}]}(N-1)]^T$, is given by

$$\mathbf{y}^{[\text{r}]} = \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{r-d}]}} \mathbf{\Lambda}_{\phi^{[\text{r-d}]}} \bar{\mathbf{G}} \mathbf{x}^{[\text{r}]} + \mathbf{w}, \tag{5}$$

where $\mathbf{x}^{[\text{r}]} \triangleq [x^{[\text{r}]}(0), x^{[\text{r}]}(1), \cdots, x^{[\text{r}]}(N-1)]^T = \mathbf{F}^H \mathbf{s}^{[\text{r}]}$, $\mathbf{s}^{[\text{r}]}$ is the frequency domain relay training signal, $\mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{r-d}]}} \triangleq \text{Diag}\big[e^{j\theta^{[\text{r-d}]}(0)}, e^{j\theta^{[\text{r-d}]}(1)}, \cdots, e^{j\theta^{[\text{r-d}]}(N-1)}\big]$, $\theta^{[\text{r-d}]}(n)$ is the n-th PN sample corresponding to relay-destination link, $\mathbf{\Lambda}_{\phi^{[\text{r-d}]}} \triangleq$

$\text{Diag}\left[1, e^{j2\pi\phi^{[\text{r-d}]}/N}, \cdots, e^{j2\pi\phi^{[\text{r-d}]}(N-1)/N}\right]$, $\phi^{[\text{r-d}]}$ is the normalized CFO generated by the mismatch between the relay and destination carrier frequencies, $\bar{\mathbf{G}}$ is a circulant channel matrix given by $\bar{\mathbf{G}} \triangleq \mathbf{F}^H \mathbf{\Lambda}_{\tilde{\mathbf{g}}} \mathbf{F}$, with $\mathbf{\Lambda}_{\mathbf{g}} = \text{Diag}(\tilde{\mathbf{g}})$, $\tilde{\mathbf{g}} \triangleq \sqrt{N}\mathbf{F}[\mathbf{g}^T, \mathbf{0}_{N-L_g,1}^T]^T$, and $\mathbf{g} \triangleq [g(0), g(1), \cdots, g(L_g-1)]^T$.

C. Statistical Model of Phase Noise

Based on the properties of PN in practical oscillators, PN in each OFDM symbol is modeled by a Wiener process, i.e.,

$$\theta^{[i]}(n) = \theta^{[i]}(n-1) + \Delta^{[i]}(n), \quad i = [\text{s-d}], [\text{r-d}] \tag{6}$$

where $\Delta^{[i]}(n-1)$ is a real Gaussian variable following $\Delta^{[i]}(n) \sim \mathcal{N}(0, \sigma^2_{\Delta^{[i]}})$. As in [6], it is assumed that $\theta^{[i]}(-1) = 0$. From (6), it can be concluded that the PN vector, $\boldsymbol{\theta}^{[i]} \triangleq [\theta^{[i]}(0), \theta^{[i]}(1), \cdots, \theta^{[i]}(N-1)]^T$, follows a Gaussian distribution, i.e. $\boldsymbol{\theta}^{[i]} \sim \mathcal{N}(0, \mathbf{\Psi}^{[i]})$, where the covariance matrix $\mathbf{\Psi}^{[i]}$ is given by

$$\mathbf{\Psi}^{[i]} = \sigma^2_{\Delta^{[i]}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 2 & \cdots & 2 & 2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & N-1 & N \end{bmatrix}. \tag{7}$$

Here we assume that in practice, $\Delta^{[i]}(n)$ is small enough such that $\theta^{[i]}(n)$ is always far smaller than π. To reduce estimation overhead of joint estimation of channel, CFO, and PN parameters, we utilize correlation amongst the PN parameters, as shown in (7), to reduce the number of unknown parameters that need to be estimated. It is not hard to verify that most eigenvalues of the matrix $\mathbf{\Psi}^{[i]}$ are close to zero. Thus, the PN vector, $\boldsymbol{\theta}^{[i]}$, can be represented as

$$\boldsymbol{\theta}^{[i]} = \mathbf{\Pi}^{[i]} \boldsymbol{\eta}^{[i]}, \qquad i = [\text{s-d}], [\text{r-d}] \tag{8}$$

where $\boldsymbol{\eta}^{[i]} \in \mathbb{C}^{M \times 1}$ is the shortened unknown PN vector with length of M following $\boldsymbol{\eta}^{[i]} \sim \mathcal{N}(0, \mathbf{I}_M)$ and $\mathbf{\Pi}^{[i]} \in \mathbb{C}^{N \times M}$ transforms $\boldsymbol{\eta}^{[i]}$ to $\boldsymbol{\theta}^{[i]}$. Moreover, the singular value decomposition of $\mathbf{\Psi}^{[i]}$ is given by $\mathbf{\Psi}^{[i]} = \mathbf{U}^{[i]} \mathbf{D}^{[i]} \left(\mathbf{U}^{[i]}\right)^T$, where $\mathbf{U}^{[i]}$ is the $N \times N$ eigenvector matrix of $\mathbf{\Psi}^{[i]}$ and $\mathbf{D}^{[i]} = \text{Diag}(\boldsymbol{\nu}^{[i]})$ with $\boldsymbol{\nu}^{[i]} \triangleq [\nu_{i,0}, \nu_{i,1}, \cdots, \nu_{i,N-1}]^T$ containing the eigenvalues of $\mathbf{\Psi}^{[i]}$ arranged in decreasing order. Thus, the matrix $\mathbf{\Pi}^{[i]}$ in (8) can be selected as $\mathbf{\Pi}^{[i]} = \tilde{\mathbf{U}}^{[i]} \tilde{\mathbf{D}}^{[i]}$, where $\tilde{\mathbf{U}}^{[i]} = \mathbf{U}^{[i]}(:, 0 : M-1)$ and $\tilde{\mathbf{D}}^{[i]} = \text{Diag}(\tilde{\boldsymbol{\nu}}^{[i]})$ with $\tilde{\boldsymbol{\nu}}^{[i]} \triangleq [\sqrt{\nu_{i,0}}, \sqrt{\nu_{i,1}}, \cdots, \sqrt{\nu_{i,M-1}}]^T$. In the subsequent section, $\boldsymbol{\eta}^{[i]}$, for $i = [\text{s-d}], [\text{r-d}]$ is estimated instead of $\boldsymbol{\theta}^{[i]}$ in both training and data transmission intervals.

III. Joint Channel, CFO and Phase Noise Estimation

To proceed, we first reformulate (3) as

$$\mathbf{y}^{[\text{s}]} = \alpha \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \mathbf{\Lambda}_{\phi^{[\text{s-d}]}} \left(\mathbf{F}^H \mathbf{\Lambda}_{s^{[\text{s}]}} \mathbf{F}_{[L]} \mathbf{c} + \mathbf{G}\mathbf{v}\right) + \mathbf{w}, \tag{9}$$

where $\mathbf{s}^{[\text{s}]}$ with $\mathbb{E}(\mathbf{s}^{[\text{s}]} \left[\mathbf{s}^{[\text{s}]}\right]^H) = P_{\text{T}}^{[\text{s}]} \mathbf{I}_N$ denotes the training symbol transmitted from the source, $\mathbf{\Lambda}_{s^{[\text{s}]}} \triangleq \text{Diag}(\mathbf{s}^{[\text{s}]})$, and $\mathbf{F}_{[L]} \triangleq \sqrt{N}\mathbf{F}(:, 0 : L-1)$. Similarly, the received signal $\mathbf{y}^{[\text{r}]}$ in (5) can be rewritten as

$$\mathbf{y}^{[\text{r}]} = \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{r-d}]}} \mathbf{\Lambda}_{\phi^{[\text{r-d}]}} \mathbf{F}^H \mathbf{\Lambda}_{s^{[\text{r}]}} \mathbf{F}_{[L_g]} \mathbf{g} + \mathbf{w}, \tag{10}$$

where $\mathbf{\Lambda}_{s^{[\text{r}]}} \triangleq \text{Diag}(\mathbf{s}^{[\text{r}]})$ with $\mathbb{E}(\mathbf{s}^{[\text{r}]} \left[\mathbf{s}^{[\text{r}]}\right]^H) = P_{\text{T}}^{[\text{r}]} \mathbf{I}_N$ denotes the training symbol transmitted from relay and $\mathbf{F}_{[L_g]} \triangleq \sqrt{N}\mathbf{F}(:, 0 : L_g - 1)$. As in [6], it is assumed that $\mathbf{s}^{[\text{s}]}$ and $\mathbf{s}^{[\text{r}]}$ are known constant-modulus training symbols.

Based on [6] and $\mathbf{y}^{[\text{r}]}$, the MAP estimates of $\phi^{[\text{r-d}]}$ is obtained. Using the estimated CFO $\hat{\phi}^{[\text{r-d}]}$, the PN vector $\boldsymbol{\theta}^{[\text{r-d}]}$ is estimated as

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{[\text{r-d}]} = \mathbf{\Pi}^{[\text{r-d}]} \big[[\mathbf{\Pi}^{[\text{r-d}]}]^T \Re(\hat{\mathbf{\Lambda}}_{\phi^{[\text{r-d}]}} \mathbf{A} \mathbf{A}^H \hat{\mathbf{\Lambda}}^H_{\phi^{[\text{r-d}]}}) \mathbf{\Pi}^{[\text{r-d}]} \\ + \frac{\sigma^2 P_{\text{T}}^{[\text{r}]}}{2} \mathbf{I}_M \big]^{-1} [\mathbf{\Pi}^{[\text{r-d}]}]^T \Im(\hat{\mathbf{\Lambda}}_{\phi^{[\text{r-d}]}} \mathbf{A} \mathbf{A}^H \hat{\mathbf{\Lambda}}^H_{\phi^{[\text{r-d}]}}) \mathbf{1}_N, \end{aligned} \tag{11}$$

where $[\hat{\mathbf{\Lambda}}_{\phi^{[\text{r-d}]}}]_{m,m} = \exp(\frac{j2\pi(m-1)\hat{\phi}^{[\text{r-d}]}}{N})$, $\mathbf{A} \triangleq [\mathbf{Y}^{[\text{r}]}]^H \mathbf{F}^H \mathbf{\Lambda}_{s^{[\text{r}]}} \mathbf{V}$, $\mathbf{Y}^{[\text{r}]} \triangleq \text{Diag}(\mathbf{y}^{[\text{r}]})$, and $\mathbf{V} \triangleq \mathbf{F}(:, L_g : N-1)$. Note that unlike [6], here, the PN parameters are estimated by first obtaining the shortened PN vector $\boldsymbol{\eta}^{[\text{r-d}]}$ and then reconstruct the complete PN vector $\boldsymbol{\theta}^{[\text{r-d}]}$.

The joint estimation of channel response, CFO, and PN is a hybrid estimation problem consisting of both deterministic parameters and random parameters. Thus, it can be conducted under the Bayesian framework. As a result, we need to find the joint estimates of $\phi^{[\text{s-d}]}$, $\boldsymbol{\eta}^{[\text{s-d}]}$, $\mathbf{h}$ and $\mathbf{g}$ by optimizing the following unconstrained function

$$\{\hat{\phi}^{[\text{s-d}]}, \hat{\boldsymbol{\eta}}^{[\text{s-d}]}, \hat{\mathbf{h}}, \hat{\mathbf{g}}\} \propto \arg\min \mathcal{L}(\phi^{[\text{s-d}]}, \boldsymbol{\eta}^{[\text{s-d}]}, \mathbf{h}, \mathbf{g}) \tag{12}$$

where $\mathcal{L}(\phi^{[\text{s-d}]}, \boldsymbol{\eta}^{[\text{s-d}]}, \mathbf{h}, \mathbf{g}) = \log\det(\mathbf{\Sigma}) + (\mathbf{y} - \boldsymbol{\mu})^H \mathbf{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}) + \frac{1}{2}[\boldsymbol{\eta}^{[\text{s-d}]}]^T \boldsymbol{\eta}^{[\text{s-d}]}$ with $\mathbf{y} \triangleq \left[[\mathbf{y}^{[\text{s}]}]^T, [\mathbf{y}^{[\text{r}]}]^T\right]^T$, $\boldsymbol{\mu} \triangleq \left[(\alpha \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \mathbf{\Lambda}_{\phi^{[\text{s-d}]}} \mathbf{F}^H \mathbf{\Lambda}_{s^{[\text{s}]}} \mathbf{F}_{[L]} \mathbf{c})^T, (\mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{r-d}]}} \mathbf{\Lambda}_{\phi^{[\text{r-d}]}} \mathbf{F}^H \mathbf{\Lambda}_{s^{[\text{r}]}} \mathbf{F}_{[L_g]} \mathbf{g})^T\right]^T$, and $\mathbf{\Sigma} \triangleq \text{Blkdiag}\left(\mathbf{\Sigma}^{[\text{r}]}, \mathbf{\Sigma}^{[\text{d}]}\right)$ with $\mathbf{\Sigma}^{[\text{r}]} = \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \mathbf{\Lambda}_{\phi^{[\text{s-d}]}} \mathbf{G} \mathbf{G}^H \mathbf{\Lambda}^H_{\phi^{[\text{s-d}]}} \mathbf{\Lambda}^H_{\boldsymbol{\theta}^{[\text{s-d}]}} + \sigma_D^2 \mathbf{I}_N$ and $\mathbf{\Sigma}^{[\text{d}]} = \sigma_D^2 \mathbf{I}_N$. Although the CFO, $\phi^{[\text{s-d}]}$, and PN vector, $\boldsymbol{\theta}^{[\text{s-d}]}$, are only contained in the received signal $\mathbf{y}^{[\text{s}]}$, the backward substitution method proposed in [6] cannot be exploited here to solve (12) due to the unknown noise covariance matrix $\mathbf{\Sigma}^{[\text{r}]}$. To make the non-convex problem (12) tractable, we propose to decouple (12) into serval subproblems that each can be solved separately in an iterative approach.

A. Phase Noise Estimation

In the first subproblem, we intend to obtain an estimate of the PN vector $\boldsymbol{\eta}^{[\text{s-d}]}$ at the $(k+1)$-th iteration, $[\hat{\boldsymbol{\eta}}^{[\text{s-d}]}]^{[k+1]}$, via the estimates of $[\boldsymbol{\phi}^{[\text{s-d}]}]$, $\mathbf{h}$, and $\mathbf{g}$ from the k-th iteration, $[\hat{\boldsymbol{\phi}}^{[\text{s-d}]}]^{[k]}$, $\hat{\mathbf{h}}^{[k]}$ and $\hat{\mathbf{g}}^{[k]}$, respectively, according to

$$[\hat{\boldsymbol{\eta}}^{[\text{s-d}]}]^{[k+1]} \propto \arg\min_{\boldsymbol{\eta}^{[\text{s-d}]}} \mathcal{L}_{\boldsymbol{\eta}^{[\text{s-d}]}}, \tag{13}$$

where $\mathcal{L}_{\boldsymbol{\eta}^{[\text{s-d}]}} = \log\det(\mathbf{\Sigma}^{[\text{r}]}) + (\mathbf{y}^{[\text{s}]} - \boldsymbol{\mu}^{[\text{s-d}]})^H [\mathbf{\Sigma}^{[\text{r}]}]^{-1} (\mathbf{y}^{[\text{s}]} - \boldsymbol{\mu}^{[\text{s-d}]}) + \frac{1}{2}[\boldsymbol{\eta}^{[\text{s-d}]}]^T \boldsymbol{\eta}^{[\text{s-d}]}$ with $\boldsymbol{\mu}^{[\text{s-d}]} \triangleq \alpha \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \hat{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}} \mathbf{F}^H \mathbf{\Lambda}_{s^{[\text{s}]}} \mathbf{F}_{[L]} \hat{\mathbf{c}}^{[k]}$, $[\hat{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}]_{m,m} = \exp(\frac{j2\pi(m-1)[\hat{\phi}^{[\text{s-d}]}]^{[k]}}{N})$ and $\hat{\mathbf{c}}^{[k]} \triangleq \hat{\mathbf{h}}^{[k]} \star \hat{\mathbf{g}}^{[k]}$, $\mathbf{\Sigma}^{[\text{r}]} = \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}} \hat{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}} \hat{\mathbf{G}}^{[k]} [\hat{\mathbf{G}}^{[k]}]^H [\hat{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}]^H \mathbf{\Lambda}^H_{\boldsymbol{\theta}^{[\text{s-d}]}} + \sigma_D^2 \mathbf{I}_N$, and $\hat{\mathbf{G}}^{[k]}$ is constructed from $\hat{\mathbf{g}}^{[k]}$ as shown in (4). The solution of $\boldsymbol{\eta}^{[\text{s-d}]}$ in (13) should be in general obtained through exhaustive search. To simplify the problem and obtain a closed-form solution, we first propose to approximate the covariance matrix $\mathbf{\Sigma}^{[\text{r}]}$ as $[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[k]}$. And then using the Taylor approximation $e^{j\theta^{[\text{s-d}]}(n)} \approx 1 + j\theta^{[\text{s-d}]}(n)$ as in [6], [13], $[\boldsymbol{\eta}^{[\text{s-d}]}]^{[k+1]}$, can be found as

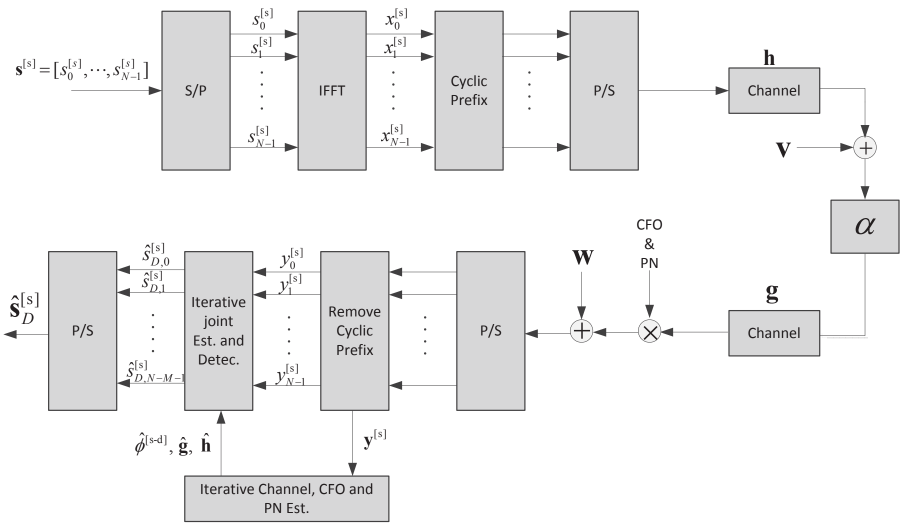

Fig. 2. Illustration of transceiver structure of the OFDM relay system.

$$\begin{aligned}[\hat{\boldsymbol{\eta}}^{[\text{s-d}]}]^{[k+1]} =& \big[\Re(\mathbf{B}^H[[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}]^{-1}\mathbf{B}) + \frac{1}{2}\mathbf{I}_M\big]^{-1} \\ &\times \Re(\mathbf{B}^H[[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}]^{-1}\bar{\mathbf{y}}^{[\text{s}]}),\end{aligned} \tag{14}$$

where $\mathbf{B} \triangleq j\text{Diag}(\alpha\hat{\boldsymbol{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{F}^H\boldsymbol{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\mathbf{c}}^{[k]})\boldsymbol{\Pi}^{[\text{s-d}]}$. Using (14), the unshortened PN estimates at the $(k+1)$-th iteration, $[\hat{\boldsymbol{\theta}}^{[\text{s-d}]}]^{[k+1]}$, can be determined as $[\hat{\boldsymbol{\theta}}^{[\text{s-d}]}]^{[k+1]} = \boldsymbol{\Pi}^{[\text{s-d}]}[\hat{\boldsymbol{\eta}}^{[\text{s-d}]}]^{[k+1]}$. Finally, the noise covariance matrix, $[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}$, is updated via the estimated PN, $[\hat{\boldsymbol{\theta}}^{[\text{s-d}]}]^{[k+1]}$.

B. Relay to Destination Channel Estimation

In the second subproblem, the channel response $\mathbf{g}$ is updated by applying the estimated CFO, source-to-relay channel, and PN vector, $[\hat{\phi}^{[\text{s-d}]}]^{[k]}$, $\hat{\mathbf{h}}^{[k]}$ and $[\hat{\boldsymbol{\theta}}^{[\text{s-d}]}]^{[k+1]}$, respectively. To proceed, the combined channel $\mathbf{c}$ is first rewritten as

$$\mathbf{c} \triangleq \tilde{\mathbf{G}}\mathbf{h} = \tilde{\mathbf{H}}\mathbf{g}, \tag{15}$$

where $\tilde{\mathbf{G}} \in \mathbb{C}^{L\times L_h}$ is denoted as

$$\tilde{\mathbf{G}} = \begin{bmatrix} g(0) & 0 & \cdots & 0 & 0 \\ g(1) & g(0) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & g(1) & g(0) \end{bmatrix}, \tag{16}$$

and $\tilde{\mathbf{H}} \in \mathbb{C}^{L\times L_g}$ has a similar form as $\tilde{\mathbf{G}}$. Subsequently, the optimization problem for updating $\mathbf{g}$ is given by

$$\begin{aligned}\hat{\mathbf{g}}^{[k+1]} \propto& \arg\min_{\mathbf{g}} \mathcal{L}_{\mathbf{g}} \\ \propto& \arg\min_{\mathbf{g}} \log\det(\boldsymbol{\Sigma}) + (\mathbf{y}-\mathbf{C}\mathbf{g})^H\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{C}\mathbf{g}),\end{aligned} \tag{17}$$

where $\mathbf{C} \triangleq [(\alpha\hat{\boldsymbol{\Lambda}}^{[k+1]}_{\boldsymbol{\theta}^{[\text{s-d}]}}\hat{\boldsymbol{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{F}^H\boldsymbol{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\tilde{\mathbf{H}}}^{[k]})^T, (\hat{\boldsymbol{\Lambda}}_{\boldsymbol{\theta}^{[\text{r-d}]}}\hat{\boldsymbol{\Lambda}}_{\boldsymbol{\phi}^{[\text{r-d}]}}\mathbf{F}^H\boldsymbol{\Lambda}_{s^{[\text{r}]}}\mathbf{F}_{[L_g]})^T]^T$ with $\hat{\tilde{\mathbf{H}}}^{[k]}$ being formed by using the estimate of the source-to-relay channel in the k-th iteration $\hat{\mathbf{h}}^{[k]}$ according to (15), and $\boldsymbol{\Sigma} \triangleq \text{Blkdiag}\big(\boldsymbol{\Sigma}^{[\text{r}]}, \sigma_D^2\mathbf{I}_N\big)$ with $\boldsymbol{\Sigma}^{[\text{r}]} = \alpha^2\sigma_R^2\hat{\boldsymbol{\Lambda}}^{[k+1]}_{\boldsymbol{\theta}^{[\text{s-d}]}}\hat{\boldsymbol{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{G}\mathbf{G}^H[\hat{\boldsymbol{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}]^H[\hat{\boldsymbol{\Lambda}}^{[k+1]}_{\boldsymbol{\theta}^{[\text{s-d}]}}]^H + \sigma_D^2\mathbf{I}_N$. It is worth noting that the covariance matrix $\boldsymbol{\Sigma}$ is also dependent on the channel response $\mathbf{g}$ as shown in (12), which makes solving (17) difficult. However, due to the iterative nature of the proposed algorithm, while updating the value of $\mathbf{g}$ at the $(k+1)$-th iteration, the covariance matrix from the k-th, i.e, previous, iteration is used instead. By equating the gradient of $\mathcal{L}_{\mathbf{g}}$ to zero, a closed-form solution for the relay-to-destination channel at the $(k+1)$-th iteration, $\hat{\mathbf{g}}^{[k+1]}$, can be found as

$$\hat{\mathbf{g}}^{[k+1]} = \big(\mathbf{C}^H\big[\hat{\boldsymbol{\Sigma}}^{[k]}\big]^{-1}\mathbf{C}\big)^{-1}\mathbf{C}^H[\hat{\boldsymbol{\Sigma}}^{[k]}]^{-1}\mathbf{y}. \tag{18}$$

Then, using $\hat{\mathbf{g}}^{[k+1]}$, the noise covariance $[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}$ is updated.

C. Source to Relay Channel Estimation

In the third subproblem, we intend to update the estimate of the source to relay channel based on the estimates $[\phi^{[\text{s-d}]}]^{[k]}$, $[\boldsymbol{\theta}^{[\text{s-d}]}]^{[k+1]}$ and $\mathbf{g}^{[k+1]}$ via the following optimization problem

$$\hat{\mathbf{h}}^{[k+1]} \propto \arg\min_{\mathbf{h}} (\mathbf{y}^{[\text{s}]}-\mathbf{D}\mathbf{h})^H[[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}]^{-1}(\mathbf{y}^{[\text{s}]}-\mathbf{D}\mathbf{h}), \tag{19}$$

where $\mathbf{D} \triangleq \alpha\hat{\boldsymbol{\Lambda}}^{[k+1]}_{\boldsymbol{\theta}^{[\text{s-d}]}}\hat{\boldsymbol{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{F}^H\boldsymbol{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\tilde{\mathbf{G}}}^{[k+1]}$. In (19), $\hat{\tilde{\mathbf{G}}}^{[k+1]}$ is formed as indicated in (16) by using $\hat{\mathbf{g}}^{[k+1]}$. Similar to the relay to destination channel, $\mathbf{g}$, the closed-form solution of $\mathbf{h}$ in (19) can be obtained as

$$\hat{\mathbf{h}}^{[k+1]} = \big(\mathbf{D}^H\big[[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}\big]^{-1}\mathbf{D}\big)^{-1}\mathbf{D}^H\big[[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}\big]^{-1}\mathbf{y}^{[\text{s}]}. \tag{20}$$

D. CFO Estimation

In order to find an estimate of the source-destination CFO at the $(k+1)$-th iteration, $[\hat{\phi}^{[\text{s-d}]}]^{[k+1]}$, similar to the steps in (13), we approximate the covariance matrix, $\boldsymbol{\Sigma}^{[\text{r}]}$ with $[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}$ and solve the unconstrained problem

$$\begin{aligned}[\hat{\phi}^{[\text{s-d}]}]^{[k+1]} \propto& \arg\min_{\phi^{[\text{s-d}]}} \mathcal{L}_{\phi^{[\text{s-d}]}} \propto \arg\min_{\phi^{[\text{s-d}]}} \\ &(\mathbf{y}^{[\text{s}]}-\boldsymbol{\mu}_{\phi^{[\text{s-d}]}})^H\big[[\hat{\boldsymbol{\Sigma}}^{[\text{r}]}]^{[k]}\big]^{-1}(\mathbf{y}^{[\text{s}]}-\boldsymbol{\mu}_{\phi^{[\text{s-d}]}}),\end{aligned} \tag{21}$$

where $\boldsymbol{\mu}_{\phi^{[\text{s-d}]}} \triangleq \alpha\boldsymbol{\Lambda}_{\boldsymbol{\phi}^{[\text{s-d}]}}\hat{\boldsymbol{\Lambda}}^{[k+1]}_{\boldsymbol{\theta}^{[\text{s-d}]}}\mathbf{F}^H\boldsymbol{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\mathbf{c}}^{[k+1]}$. To make the problem in (21) more tractable and find a closed-form solution, a Taylor series approximation similar to that in (13) is applied here. By setting $\frac{\partial \mathcal{L}_{\phi^{[\text{s-d}]}}}{\partial \phi^{[\text{s-d}]}} = 0$ and solving for $\phi^{[\text{s-d}]}$, a closed-form solution for the CFO estimate at the $(k+1)$-th iteration, $[\phi^{[\text{s-d}]}]^{k+1}$, is found as

$$[\hat{\phi}^{[\text{s-d}]}]^{k+1} = [\hat{\phi}^{[\text{s-d}]}]^{k} + \frac{\Re\big((\mathbf{y}^{[\text{s}]} - \hat{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{d}^{[\text{s-d}]})^H[[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[k]}]^{-1}\tilde{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{d}^{[\text{s-d}]}\big)}{[\mathbf{d}^{[\text{s-d}]}]^H[\tilde{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}]^H[[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[k]}]^{-1}\tilde{\mathbf{\Lambda}}^{[k]}_{\phi^{[\text{s-d}]}}\mathbf{d}^{[\text{s-d}]}}. \tag{22}$$

Then the noise covariance matrices $\mathbf{\Sigma}^{[\text{r}]}$ and $\hat{\mathbf{\Sigma}}^{[k]}$ are updated using $[\phi^{[\text{s-d}]}]^{k+1}$ as $[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[k+1]}$ and $\hat{\mathbf{\Sigma}}^{[k+1]}$, respectively.

Remark 1: For the initial points $[\hat{\phi}^{[\text{s-d}]}]^{[0]}$, $\hat{\mathbf{g}}^{[0]}$, $\hat{\mathbf{h}}^{[0]}$ and $[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[0]}$ in *Algorithm 1*, based on $\hat{\phi}^{[\text{r-d}]}$ and $\hat{\boldsymbol{\theta}}^{[\text{r-d}]}$, $\hat{\mathbf{g}}^{[0]}$ can be obtained from the received signal $\mathbf{y}^{[\text{r}]}$ as [6]: $\hat{\mathbf{g}}^{[0]} = \frac{1}{NP_{\text{T}}^{[\text{r}]}}\mathbf{F}^H_{[L_g]}\mathbf{\Lambda}^H_{s^{[\text{r}]}}\mathbf{F}\hat{\mathbf{\Lambda}}^H_{\theta^{[\text{r-d}]}}\hat{\mathbf{\Lambda}}^H_{\phi^{[\text{r-d}]}}\mathbf{y}^{[\text{r}]}$. By ignoring the PN terms, (9) can be approximated as $\mathbf{y}^{[\text{s}]} \approx \alpha\mathbf{\Lambda}_{\phi^{[\text{s-d}]}}\mathbf{F}^H\mathbf{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\tilde{\mathbf{G}}}^{[0]}\mathbf{h} + \alpha\mathbf{\Lambda}_{\phi^{[\text{s-d}]}}\hat{\mathbf{G}}^{[0]}\mathbf{v} + \mathbf{w}$, where $\hat{\tilde{\mathbf{G}}}^{[0]}$ and $\hat{\mathbf{G}}^{[0]}$ are formed via $\hat{\mathbf{g}}^{[0]}$ according to (16) and (4), respectively. Subsequently, using the ML criterion the initial estimates of the CFO, $[\hat{\phi}^{[\text{s-d}]}]^{[0]}$, and channel, $\hat{\mathbf{h}}^{[0]}$, can be obtained by minimizing

$$\{\hat{\mathbf{h}}^{[0]}, [\hat{\phi}^{[\text{s-d}]}]^{[0]}\} = \min_{\mathbf{h},\phi^{[\text{s-d}]}} \ (\mathbf{y}^{[\text{s}]} - \alpha\mathbf{\Lambda}_{\phi^{[\text{s-d}]}}\mathbf{F}^H\mathbf{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\tilde{\mathbf{G}}}^{[0]}\mathbf{h})^H \times [\mathbf{\Sigma}^{[\text{r}]}]^{-1}(\mathbf{y}^{[\text{s}]} - \alpha\mathbf{\Lambda}_{\phi^{[\text{s-d}]}}\mathbf{F}^H\mathbf{\Lambda}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\hat{\tilde{\mathbf{G}}}^{[0]}\mathbf{h}) + \log\det(\mathbf{\Sigma}^{[\text{r}]}),$$

where $\mathbf{\Sigma}^{[\text{r}]} \triangleq \alpha^2\sigma_R^2\mathbf{\Lambda}_{\phi^{[\text{s-d}]}}\hat{\mathbf{G}}^{[0]}[\hat{\mathbf{G}}^{[0]}]^H\mathbf{\Lambda}^H_{\phi^{[\text{s-d}]}} + \sigma_D^2\mathbf{I}_N$. As for $[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[0]}$, using the Taylor approximation we have

$$[\hat{\mathbf{\Sigma}}^{[\text{r}]}]^{[0]} \approx \mathbf{\Omega} + \mathbf{\Omega} \odot \mathbf{\Psi}^{[\text{s-d}]} + \sigma_D^2\mathbf{I}_N \tag{23}$$

where $\mathbf{\Omega} \triangleq \alpha^2\sigma_R^2\hat{\mathbf{\Lambda}}_{\phi^{[\text{s-d}]}}\hat{\mathbf{G}}^{[0]}[\hat{\mathbf{G}}^{[0]}]^H\hat{\mathbf{\Lambda}}^H_{\phi^{[\text{s-d}]}}$. In (23), we use the expectation $\mathbb{E}(\boldsymbol{\theta}^{[\text{s-d}]}[\boldsymbol{\theta}^{[\text{s-d}]}]^H) = \mathbf{\Psi}^{[\text{s-d}]}$ to replace the term $\boldsymbol{\theta}^{[\text{s-d}]}[\boldsymbol{\theta}^{[\text{s-d}]}]^H$. This allows for a closed-form expression for obtaining the source-to-relay channel estimates.

Remark 2: Due to the joint parameters estimation, the residual phase ambiguities between the channel, CFO, and PN and between the PN and CFO may exist. These ambiguities make it difficult to assess the estimation accuracy of the proposed iterative estimator. Thus, here, a new approach for determining the MSE of the estimated parameters is proposed. The MSE of $\hat{\mathbf{h}}$ and $\hat{\mathbf{g}}$, can be computed as

$$\text{MSE}_{\mathbf{g}} = ||\hat{\underline{\mathbf{g}}} - \underline{\mathbf{g}}||_2^2, \quad \text{MSE}_{\mathbf{h}} = ||\hat{\underline{\mathbf{h}}} - \underline{\mathbf{h}}||_2^2, \tag{24}$$

where $\hat{\underline{\mathbf{g}}} \triangleq \exp(-j\angle\hat{g}(0))\hat{\mathbf{g}}$, $\hat{\underline{\mathbf{h}}} \triangleq \exp(-j\angle\hat{h}(0))\hat{\mathbf{h}}$, $\underline{\mathbf{g}} \triangleq \exp(-j\angle g(0))\mathbf{g}$ and $\underline{\mathbf{h}} \triangleq \exp(-j\angle h(0))\mathbf{h}$. Using this approach, the phase ambiguity between the PN and channels, does not affect the MSE of channel estimation. Similarly, for the CFO and PN, the overall MSE is calculated as

$$\text{MSE}_{\phi^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{s-d}]}} = ||\underline{\boldsymbol{\delta}} - \hat{\underline{\boldsymbol{\delta}}}||_2^2, \tag{25}$$

where $\underline{\boldsymbol{\delta}} = \boldsymbol{\delta} - \delta_0\mathbf{1}$, $\hat{\underline{\boldsymbol{\delta}}} = \hat{\boldsymbol{\delta}} - \hat{\delta}_0\mathbf{1}$, $\boldsymbol{\delta} = [\delta_0, \delta_0, \cdots, \delta_{N-1}]^T$ with $\delta_m = \theta^{[\text{s-d}]}(m) + \frac{2\pi(m-1)\phi^{[\text{s-d}]}}{N}$, and $\hat{\boldsymbol{\delta}} = [\hat{\delta}_0, \hat{\delta}_1, \cdots, \hat{\delta}_{N-1}]^T$ with $\hat{\delta}_m = \hat{\theta}^{[\text{s-d}]}(m) + \frac{2\pi(m-1)\hat{\phi}^{[\text{s-d}]}}{N}$.

IV. The Hybrid Cramér-Rao Lower Bound

As stated in *Remark 2*, due to the ambiguities between the estimation of channel responses, CFO, and PN, (9) and (10) are first rewritten as

$$\begin{aligned}\mathbf{y}^{[\text{s}]} &= \alpha\mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{s-d}]}}\mathbf{\Lambda}_{\phi^{[\text{s-d}]}}\left(\mathbf{F}^H\underline{\mathbf{\Lambda}}_{s^{[\text{s}]}}\mathbf{F}_{[L]}\underline{\mathbf{c}} + \underline{\mathbf{G}}\mathbf{v}\right) + \mathbf{w},\\ \mathbf{y}^{[\text{r-d}]} &= \mathbf{\Lambda}_{\boldsymbol{\theta}^{[\text{r-d}]}}\mathbf{\Lambda}_{\phi^{[\text{r-d}]}}\mathbf{F}^H\underline{\mathbf{\Lambda}}_{s^{[\text{r}]}}\mathbf{F}_{[L_g]}\underline{\mathbf{g}} + \mathbf{w},\end{aligned} \tag{26}$$

where $\underline{\mathbf{c}} \triangleq \underline{\mathbf{h}} \star \underline{\mathbf{g}}$ with $\underline{\mathbf{h}}$ and $\underline{\mathbf{g}}$ defined in (24), $[\underline{\mathbf{\Lambda}}_{s^{[\text{s}]}}]_{m,m} \triangleq s^{[\text{s}]}_{m-1}\exp(j\angle c(0))$, $[\underline{\mathbf{\Lambda}}_{s^{[\text{r}]}}]_{m,m} \triangleq s^{[\text{r}]}_{m-1}\exp(j\angle g(0))$ are known diagonal training signal matrices that are rotated by the phases of the first elements of the channels, $\mathbf{c}$ and $\mathbf{g}$, respectively, and matrix $\underline{\mathbf{G}}$ is constructed using $\underline{\mathbf{g}}$ similar to (4). Accordingly, the HCRLB for the estimation problem is given by

$$\mathbb{E}_{\mathbf{y},\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}|\phi^{[\text{s-d}]},\phi^{[\text{r-d}]},\underline{\mathbf{g}},\underline{\mathbf{h}}}\left[(\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}})(\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}})^T\right] \succeq \mathbf{B}^{-1},$$

where $\boldsymbol{\lambda} \triangleq \Big[\phi^{[\text{s-d}]}, (\boldsymbol{\theta}^{[\text{s-d}]})^T, \phi^{[\text{r-d}]}, (\boldsymbol{\theta}^{[\text{r-d}]})^T, \underline{g}_0, \Re(\tilde{\underline{\mathbf{g}}})^T, \Im(\tilde{\underline{\mathbf{g}}})^T, \underline{h}_0, \Re(\tilde{\underline{\mathbf{h}}})^T, \Im(\tilde{\underline{\mathbf{h}}})^T\Big]^T$ with $\tilde{\underline{\mathbf{g}}} = \underline{\mathbf{g}}(1 : L_g - 1)$ and $\tilde{\underline{\mathbf{h}}} = \underline{\mathbf{h}}(1 : L_h - 1)$ denoting the vector of parameters of interest, and $\mathbf{B}$ denoting the Bayesian information matrix (BIM), which is given by

$$\begin{aligned}\mathbf{B} =& \mathbb{E}_{\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}}\left[\mathbf{FIM}(\mathbf{y};\boldsymbol{\lambda})\right] + \mathbb{E}_{\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}}\left[-\triangle^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}}\log p(\boldsymbol{\theta}^{[\text{s-d}]})\right]\\ &+ \mathbb{E}_{\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}}\left[-\triangle^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}}\log p(\boldsymbol{\theta}^{[\text{r-d}]})\right].\end{aligned} \tag{27}$$

In (27), $\mathbf{FIM}(\mathbf{y};\boldsymbol{\lambda}) = \mathbb{E}_{\mathbf{y}}\left[-\triangle^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}}\log p(\mathbf{y};\boldsymbol{\lambda})\right]$ denotes the Fisher's information matrix (FIM). Note that due to the space limitation, we omit the derivations of $\mathbb{E}_{\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}}\left[\mathbf{FIM}(\mathbf{y};\boldsymbol{\lambda})\right]$, $\mathbb{E}_{\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}}\left[-\triangle^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}}\log p(\boldsymbol{\theta}^{[\text{s-d}]})\right]$, and $\mathbb{E}_{\boldsymbol{\theta}^{[\text{s-d}]},\boldsymbol{\theta}^{[\text{r-d}]}}\left[-\triangle^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}}\log p(\boldsymbol{\theta}^{[\text{r-d}]})\right]$. The details can be found in [14].

V. Simulation Results

We assume that the multi-path channels exhibit Rayleigh fading characteristics, i.e., following $\mathcal{N}(0,1)$. The noise powers at relay and destination are assumed to be the same, i.e., $\sigma_R^2 = \sigma_D^2 = 1$. The relay amplification factor is simply set as $\alpha = 1$. Moreover, it is assumed that $P_{\text{T}}^{[\text{s}]} = P_{\text{T}}^{[\text{r}]} = P_{\text{T}}$. The SNR during the training phase is denoted by $\text{SNR}_T = P_{\text{T}}$. The system parameters of the OFDM relay model are set up as: $L_h = L_g = 6$, $N = 64$, all the subcarriers are modulated in quadrature phase shift keying (QPSK) format. The normalized CFOs, $\phi^{[\text{s-d}]}$ and $\phi^{[\text{r-d}]}$, are uniformly drawn from $[-0.4, 0.4]$ and $[-0.2, 0.2]$, respectively, and the PN innovation variances for, $\boldsymbol{\theta}^{[\text{s-d}]}$ and $\boldsymbol{\theta}^{[\text{r-d}]}$ are assumed to be the same, i.e., $\sigma^2_{\Delta^{[\text{s-d}]}} = \sigma^2_{\Delta^{[\text{r-d}]}} = \sigma^2_{\Delta}$. As our considered joint estimation problem has not been studied by other works, the comparison with existing works is not available.

In Figs. 3 and 4, we illustrate the estimation MSE of $\underline{\mathbf{g}}$ and $\underline{\mathbf{h}}$ at $\sigma^2_{\Delta} = 10^{-3}$ rad^2, respectively. As a comparison, the channel estimation performance while ignoring the effect of PN on the received signal is also presented. Finally, the proposed estimation algorithms performance is benchmarked

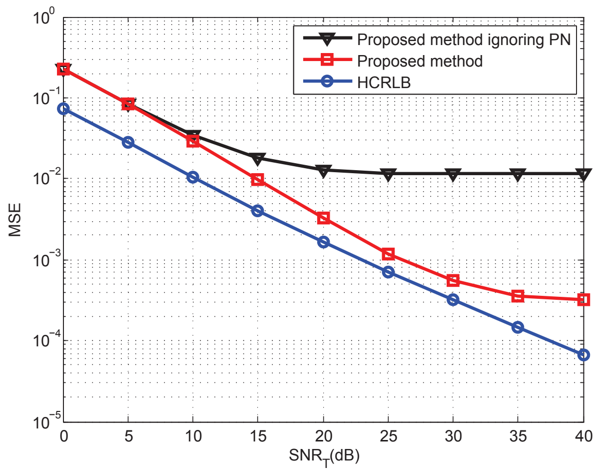

Fig. 3. The MSE of $\underline{\mathbf{g}}$ estimation at $\sigma_{\Delta}^{2} = 10^{-3}\text{rad}^{2}$.

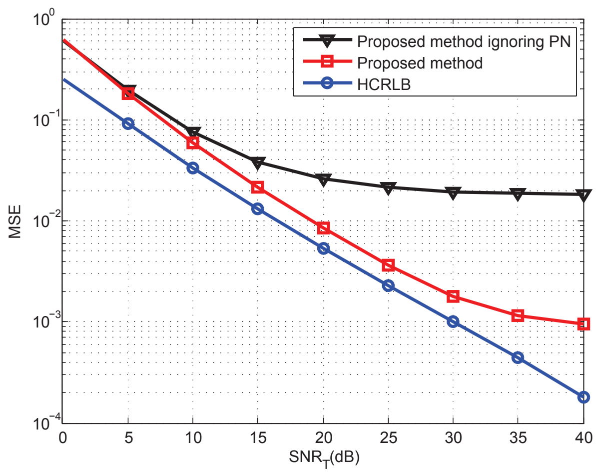

Fig. 4. The MSE of $\underline{\mathbf{h}}$ estimation at $\sigma_{\Delta}^{2} = 10^{-3}\text{rad}^{2}$.

using the derived HCRLB. Figs. 3–4 indicate that by including the PN parameters in the joint estimation problem, CFO and channel estimation performance in relay networks can be significantly enhanced, and the proposed algorithm has almost a constant performance gap with the derived HCRLB bound. This may be due to the loose of the HCRLB [15]. Nevertheless, the performance of the proposed estimator is close to the derived HCRLB for moderate to high SNRs. Fig. 5 illustrate the MSE for estimation of combined CFO and PN, $\underline{\boldsymbol{\delta}}$ at $\sigma_{\Delta}^{2} = 10^{-3}$ rad^{2}. Similar to the results for channel estimation, the overall estimation performance suffers from an error floor. This phenomenon can be similarly justified due to the imperfect estimation of PN parameters.

VI. Conclusions

In this paper, joint channel, CFO, and PN estimation and data detection in OFDM relay networks is analyzed. Based on the proposed training and data transmission frame, a new joint CFO, channel, and PN estimation algorithm that iteratively estimates these impairments is derived. Meanwhile, a new PN estimation approach that reduces the dimensionality of the estimation problem is proposed. Simulation results show that the proposed estimation algorithm can achieve a performance close to the derived HCRLB.

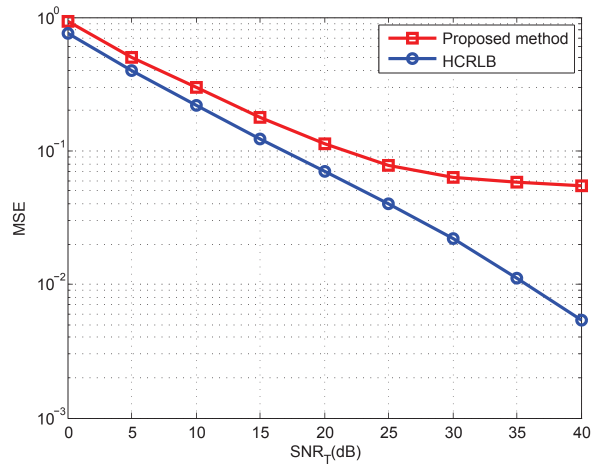

Fig. 5. The MSE of CFO plus PN estimation at $\sigma_{\Delta}^{2} = 10^{-3}\text{rad}^{2}$.

References

[1] F. Gao, T. Cui, and A. Nallanathan, “On channel estimation and optimal training design for amplify and forward relay networks,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[2] H. Mehrpouyan and S. Blostein, “Bounds and algorithms for multiple frequency offset estimation in cooperative networks,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1300–1311, Apr. 2011.

[3] L. Tomba, “On the effect of Wiener phase noise in OFDM systems,” *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 580–583, May 1998.

[4] Z. Zhang, W. Zhang, and C. Tellambura, “Cooperative OFDM channel estimation in the presence of frequency offsets,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3447–3459, Sep. 2009.

[5] L. Thiagarajan, S. Sun, and T. Quek, “Joint carrier frequency offset and channel estimation in OFDM based non-regenerative wireless relay networks,” in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Apr. 2009, pp. 2569–2572.

[6] D. D. Lin, R. Pacheco, T. J. Lim, and D. Hatzinakos, “Joint estimation of channel response, frequency offset, and phase noise in OFDM,” *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3542–3554, Sep. 2006.

[7] F. Septier, Y. Delignon, A. Menhaj-Rivenq, and C. Garnier, “Monte carlo methods for channel, phase noise, and frequency offset estimation with unknown noise variances in OFDM systems,” *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3613–3626, Aug. 2008.

[8] H. Mehrpouyan, M. Khanzadi, M. Matthaiou, A. Sayeed, R. Schober, and Y. Hua, “Improving bandwidth efficiency in e-band communication systems,” *IEEE Trans. Commun.*, vol. 52, no. 3, pp. 121–128, March 2014.

[9] Y. Lee, J. Ha, and J. Choi, “Design of a wideband indoor repeater antenna with high isolation for 3g systems,” *IEEE Antennas and Wireless Propagation Letters*, vol. 9, pp. 697–700, 2010.

[10] Mini-Circuits, “Wideband microwave amplifier ava-24+.” [Online]. Available: http://www.minicircuits.com/pdfs/AVA-24+.pdf

[11] K. Salehian, M. Guillet, B. Caron, and A. Kennedy, “On-channel repeater for digital television broadcasting service,” *IEEE Trans. on Broadcasting*, vol. 48, no. 2, pp. 97–102, 2002.

[12] Mini-Circuits, “Ultra low noise voltage conrtolled ocillator ros-209-319+.” [Online]. Available: http://www.minicircuits.com/pdfs/ROS-209-319+.pdf

[13] H. Mehrpouyan *et al.*, “Joint estimation of channel and oscillator phase noise in MIMO systems,” *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.

[14] R. Wang, H. Mehrpouyan, M. Tao, and Y. Hua, “Channel estimation, carrier recovery, and data detection in the presence of phase noise in ofdm relay systems,” available at: http://arxiv.org/abs/1309.7564.

[15] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. John Wiley and Sons Inc., 2007.