

# Channel Estimation, Carrier Recovery, and Data Detection in the Presence of Phase Noise in OFDM Relay Systems

Rui Wang, *Member, IEEE*, Hani Mehrpouyan, *Member, IEEE*, Meixia Tao, *Senior Member, IEEE*, and Yingbo Hua, *Fellow, IEEE*

**Abstract**—Due to its time-varying nature, oscillator phase noise can significantly degrade the performance of the channel estimation, carrier recovery, and data detection blocks in high-speed wireless communication systems. In this paper, we propose a new data-aided joint channel, carrier frequency offset (CFO) and phase noise estimator for orthogonal frequency division multiplexing (OFDM) relay systems. For the data transmission phase, we propose a new iterative receiver that tracks phase noise and detects the transmitted symbols. Additionally, we derive the hybrid Cramér–Rao lower bound for evaluating the performance of channel estimation and carrier recovery algorithms in OFDM relay networks. Extensive simulations demonstrate that the application of the proposed estimation and receiver blocks significantly improves the performance of OFDM relay networks in the presence of phase noise and CFO.

**Index Terms**—Relay, orthogonal frequency division multiplexing (OFDM), channel estimation, phase noise.

## I. INTRODUCTION

APPLICATION of relaying has been identified as a suitable approach for combating long-distance channel distortion and small-scale fading in wireless communication systems [2]. Various physical layer techniques, such as distributed space-time block coding [3], [4], precoding [5], [6], power scheduling [7], etc., for relay systems have been extensively studied. From these works, it can be deduced that to deliver the advantages of relay networks, the network's

Manuscript received November 7, 2014; revised April 21, 2015 and August 12, 2015, accepted September 20, 2015. Date of publication October 5, 2015; date of current version February 8, 2016. This work was supported in part by the NSF of China under Grant 61401313. The work of R. Wang was supported by the Shanghai Pujiang Program under Grant 15PJD037. The work of H. Mehrpouyan was supported by the Department of Defense under grant number W911NF-15-1-0033. The work of M. Tao was supported by the National 973 project of China under Grant 2012CB316100 and in part by the NSF of China under Grant 61322102. Part of this work was presented at the IEEE Global Communications Conference 2014 [1]. The associate editor coordinating the review of this paper and approving it for publication was Luca Sanguinetti.

R. Wang is with the College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China (e-mail: ruiwang@tongji.edu.cn).

H. Mehrpouyan is with the Department of Computer and Electrical Engineering, Boise State University, Boise, ID 83725-2075 USA (e-mail: hani.mehr@ieee.org).

M. Tao is with the Department of Electronic Engineering and Department of Electrical and Computer Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: mxtao@sjtu.edu.cn).

Y. Hua is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: yhua@ece.ucr.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2015.2487268

channel state information (CSI) needs to be accurately obtained [8]–[14], while the negative impact of impairments such as carrier frequency offset (CFO) and phase noise (PN) caused by Doppler shifts and oscillator imperfections needs to be mitigated [15].

In single carrier communication systems, CFO and PN are multiplicative and result in a rotation of the signal constellation from symbol to symbol and erroneous data detection [16], [17]. On the other hand, in the case of orthogonal frequency division multiplexing (OFDM) systems, CFO and PN are convolved with the data symbols, resulting in the rotation of the signal constellation and inter-carrier interference (ICI), which can significantly deteriorate the overall performance of an OFDM system [18]–[22]. Thus, extensive research has been recently carried out to find carrier recovery schemes that complement traditional approaches, e.g., those based on the phase-locked loop (PLL). More importantly, as demonstrated in [18], [23], [24], to accurately obtain the channel, CFO, and PN parameters in communications systems, these parameters need to be jointly estimated. However, the prior art on channel and CFO estimation in relay networks has not taken into consideration the detrimental impact of PN. In fact, unlike in single carrier systems [25], PN in OFDM systems cannot be modeled as an additive noise, since such an approach significantly deteriorates the performance of algorithms for estimating and mitigating the impact of this impairment [24].

Due to the presence of multiple hops between source and destination, channel estimation in relay systems is quite different from traditional point-to-point systems. For the amplify-and-forward (AF) relaying strategy, one approach is to only estimate source to destination channels [8], [9]. However, to further enhance cooperative system performance by enabling relay precoding/beamforming or relay resource allocation, the channel response of each hop needs to be separately estimated [9]–[12]. Furthermore, since the channel response from relay to destination affects the destination noise covariance matrix, estimating individual channel responses is generally required for more accurate signal detection at the destination. It is worth noting that the contributions in [8]–[12] only focus on channel estimation while ignoring the effect of CFO and PN.

Joint estimation of the channel responses and CFO in single carrier relay systems has been considered in [23], [26]. In [26], the Gauss-Hermite integration and approximate Rao-Blackwellization based joint CFO and channel estimators are proposed, while in [23] joint CFO and channel estimation via

the MUSIC algorithm is analyzed. However, the works in [23], [26] ignore the effect of PN. In fact, although both CFO and PN result in an unknown rotation of the signal constellation, PN is a time-varying parameter compared to the CFO and can be more difficult to track [17], [18]. More importantly, the negative impact of CFO and PN may be greater in the case of OFDM systems compared to single carrier systems [27], [28].

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 is analyzed in [27], [28]. Undoubtedly, this effect can also be observed in OFDM based cooperative 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 [29], [30]. In particular, a two-time-slot cooperative estimation protocol to obtain these parameters has been proposed in [29], while in [30] the authors studied the maximum likelihood (ML) and the least squares based joint CFO and channel estimation algorithms. More importantly, none of the approaches in [29], [30] consider the effect of PN on channel and CFO estimation or the overall relaying performance. While ignoring the effect of CFO, joint channel and PN estimation in OFDM relay networks is analyzed in [31]. However, none of the approaches in [29]–[31] consider the effect of PN on joint channel and CFO estimation. In [32], the authors consider PN, CFO, and channel estimation in a relay based cooperative network. However, the work presented in [32] only focuses on the derivation of an estimator that obtains the cascaded channel parameters for both hops, does not present a detector for signal reception in the presence of PN at the destination node.

In this paper, we consider the problem of joint CFO, PN, and individual channel estimation for OFDM 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 among the estimated PN, CFO, and channels are analyzed. Based on this analysis, a *hybrid Cramér-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. The approach proposed here can be also applied to point-to-point systems to reduce PN estimation and carrier recovery 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) when the phase noise variance is not large.
- A comb-type OFDM symbol containing both pilots and data symbols is proposed to track the time-varying PN parameters during the data transmission interval. Next, an

iterative receiver that applies the proposed OFDM symbol to perform joint data detection and PN tracking at the destination node is derived.

The rest of the paper is organized as follows. Section II presents the system model and assumptions in this paper. The joint estimation algorithm is presented in Section III. In Section IV, the HCRLB for the proposed joint estimation problem is derived. The proposed iterative receiver for joint data detection and PN tracking is presented in Section V. Extensive simulation results are illustrated in Section VI. Finally, we conclude the paper in Section VII.

*Notations:*  $\mathbb{E}(\cdot)$  denotes the expectation of its argument.  $\odot$ ,  $\star$ , and  $*$  denote the Hadamard product, linear, and circular convolutions, respectively.  $\text{Tr}(\mathbf{A})$ ,  $\mathbf{A}^{-1}$ , and  $\det(\mathbf{A})$  denote the trace, inverse, and determinant of matrix  $\mathbf{A}$ , respectively.  $\text{Diag}(\mathbf{a})$  denotes a diagonal matrix with  $\mathbf{a}$  being its diagonal entries.  $\text{Blkdiag}(\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_{N-1})$  denotes a block diagonal matrix with  $\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_{N-1}$  as its diagonal matrices.  $\mathbf{A}(N : M, :)$  and  $\mathbf{A}(:, N : M)$  denote a submatrix containing the  $N$ -th to  $M$ -th rows of  $\mathbf{A}$  and a submatrix containing the  $N$ -th to  $M$ -th columns of  $\mathbf{A}$ , respectively. Superscripts  $(\cdot)^T$ ,  $(\cdot)^*$  and  $(\cdot)^H$  denote the transpose, conjugate, and conjugate transpose, respectively.  $\mathbf{0}_{N \times M}$ ,  $\mathbf{I}_N$ , and  $\mathbf{1}_N$  denote the  $N \times M$  zero matrix,  $N \times N$  identity matrix, and  $N \times 1$  vector of ones, respectively.  $\Re(z)$  and  $\Im(z)$  denote the real and imaginary operators.  $\mathbb{C}^{x \times y}$  and  $\mathbb{R}^{x \times y}$  denote spaces of  $x \times y$  matrices with complex and real entries, respectively.  $\Delta_{\mathbf{x}}^2 f(\cdot) \triangleq \left[ \frac{\partial^2 f}{\partial \mathbf{x}^2} \right]^T$  denotes the second order partial derivative of function  $f(\cdot)$  with respect to vector  $\mathbf{x}$ . Finally,  $\mathcal{CN}(\mathbf{x}, \mathbf{\Sigma})$  and  $\mathcal{N}(\mathbf{x}, \mathbf{\Sigma})$  denote complex and real Gaussian distributions, respectively, with mean  $\boldsymbol{\mu}$  and covariance  $\mathbf{\Sigma}$ . Moreover, to improve the presentation of this paper, we summarize the notations in Table I.

## II. SYSTEM MODEL

An AF relaying OFDM system is considered, where a source node transmits its signal to a destination node through a relay. Here, it is assumed that there is no direct link from the source to the destination due to physical impairments, such as deep fading and shadowing effect. Similar to prior work in this field, e.g., [18], [19], quasi static fading channels are considered, i.e., the CSI is assumed to be constant over the duration of a 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 both the source to relay and relay to destination hops. Moreover, the signal received during the data transmission interval is used for data detection and compensation of phase noise as elaborated in Section V. Although time synchronization is an important issue in relay systems, it has been addressed in work like [53]. Hence, we assume that the overall network is synchronized in time. We consider the following two different relay operations.

### A. Two Relay Operations

1) *Analog relay processing:* Unlike [31], it is assumed that the relay node simply forwards the received signal without

TABLE I  
NOTATIONS USED IN THIS WORK

Notation	Stand for	Notation	Stand for
[s-d]	index from source to destination	$\theta^{[r-d]}(n)$	the $n$ -th PN sample for relay-destination link
[r-d]	index from relay to destination	$\mathbf{y}^{[s]}$	received training signal transmitted from the source
$\mathbf{h}$	source to relay channel	$\phi^{[s-d]}$	normalized CFO for source-destination link
$\mathbf{g}$	relay to destination channel	$\phi^{[r-d]}$	normalized CFO for relay-destination link
$\mathbf{c}$	$\mathbf{h} * \mathbf{g}$	$\mathbf{y}^{[r]}$	received training signal transmitted from the relay
$L_h$	length of $\mathbf{h}$	$\mathbf{v}$	additive noise at the relay node
$L_g$	length of $\mathbf{g}$	$\mathbf{w}$	additive noise at the destination node
$L$	length of $\mathbf{c}$	$\mathbf{F}$	normalized discrete Fourier transform matrix
$\theta^{[s-d]}(n)$	the $n$ -th PN sample for source-destination link	$\mathbf{F}_{[L]}$	$\sqrt{N}\mathbf{F}(:, 0 : L - 1)$

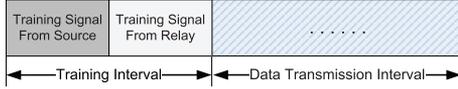


Fig. 1. The proposed timing diagram of the OFDM relay system, in which we assume that the channel responses and CFO remain time-invariant during the entire packet while PN is assumed to vary with time.

removing the CP corresponding to the source-to-relay link and appending a new CP for the relay-to-destination link. This assumption ensures a considerably simpler relaying structure as the relay does not need to convert the received signal from passband to baseband. The proposed signal model can be applied to both full-duplex and half-duplex relaying networks based on the following system setups and assumptions:

- *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 [33], [34]. This approach may also minimize or eliminate the effect of self-interference at the relay.<sup>1</sup> 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 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 [36].
- *Half-duplex relaying*: In this setup, for analog relay processing, it is assumed that the relay forwards its received signal on a different carrier frequency and may not need to convert it to baseband, i.e., the relay applies on-frequency/on-channel RF relaying [37]. 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 [38]. As such, in this setup, it is assumed that the signal forwarded from the relay is not affected by PN.

<sup>1</sup>Application of sophisticated transceivers has also been shown to minimize or eliminate the impact of self-interference at the relay [35]. In a case that the self-interference cannot be completely eliminated, the corresponding discussion can be found in Remark 1.

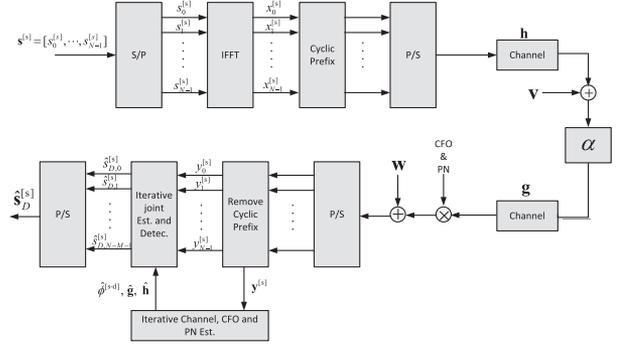


Fig. 2. Illustration of transceiver structure of the OFDM relay system.

2) *Digital relay processing*: In this setup, the relay performs a digital processing. That is, the relay first converts the received passband signal to the baseband. Subsequently, the relay removes the CP and adds a new CP. Then, after scaling the new generated signals to satisfy the relay's power constraint, they are converted to the passband and delivered to the destination.

### B. Signal Transmission from Source to Destination

The overall transmission and reception structure of each OFDM symbol from the source to the destination node is illustrated in Fig. 2. We assume that  $N$  subcarriers are used for OFDM transmission. Let  $\mathbf{s}^{[s]} \triangleq [s_0^{[s]}, s_1^{[s]}, \dots, s_{N-1}^{[s]}]^T$  denote the frequency domain modulated training or data signal sequence at the source node, which is then transformed into a set of parallel symbols  $s_k^{[s]}$ , for  $k = 0, \dots, N - 1$ . By conducting an inverse fast Fourier transform (IFFT), we obtain the time domain signal vector  $\mathbf{x}^{[s]}$  as  $\mathbf{x}^{[s]} = \mathbf{F}^H \mathbf{s}^{[s]}$ , where  $\mathbf{x}^{[s]} \triangleq [x_0^{[s]}, x_1^{[s]}, \dots, x_{N-1}^{[s]}]^T$ , and  $\mathbf{F}$  is the normalized discrete Fourier transform (DFT) matrix with  $F_{n,k} = \frac{1}{\sqrt{N}} \exp(-j \frac{2\pi(n-1)(k-1)}{N})$ . After adding the CP, the parallel signal vector is transformed into a time domain sequence denoted by  $x^{[s]}(n)$ , for  $n = -N_{CP}, \dots, N - 1$  with  $N_{CP}$  being a length of CP. Subsequently, the transmitted baseband continuous signal from the source,  $\tilde{x}^{[s]}(t)$ , can be written as

$$\tilde{x}^{[s]}(t) = \sum_{n=-L}^{N-1} x^{[s]}(n)q(t - nT_s), \quad 0 \leq t \leq T + T_{CP} \quad (1)$$

where  $T_s = T/N$  with  $T$  denoting an OFDM symbol duration,  $q(t)$  is the pulse shaping filter,  $T_{CP}$  is the duration of the CP, and  $x^{[s]}(-n) = x^{[s]}(N-n)$ , for  $n = -N_{CP}, \dots, -1$ , is the added CP symbol.

1) *Analog relay processing*: At the destination, the baseband received signal,  $y^{[s]}(t)$ , is given by

$$\begin{aligned} y^{[s]}(t) &= \alpha e^{j\theta^{[s-d]}(t)} e^{j\phi^{[s-d]}(t)} g(t) \star z(t) + w(t) \\ &= \alpha e^{j\theta^{[s-d]}(t)} e^{j\phi^{[s-d]}(t)} \left[ g(t) \star h(t) \star \tilde{x}^{[s]}(t) \right. \\ &\quad \left. + g(t) \star v(t) \right] + w(t), \end{aligned} \quad (2)$$

where  $z(t) = h(t) \star \tilde{x}^{[s]}(t) + v(t)$  and  $\alpha$  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 the additive noises at the relay and at the destination, respectively. Using a similar approach to point-to-point systems [18], [39], [40],  $\theta^{[s-d]}(t)$  is the PN corresponding to source-relay-destination link, while  $\phi^{[s-d]}(t) \triangleq 2\pi \Delta f^{[s-d]} t$  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 is determined as

$$\begin{aligned} y^{[s]}(nT_s) &= \alpha e^{j\theta^{[s-d]}(nT_s)} e^{j2\pi \Delta f^{[s-d]} nT_s} \underbrace{[g(nT_s) \star h(nT_s)]}_{\triangleq c(nT_s)} \\ &\quad \star \tilde{x}^{[s]}(nT_s) + g(nT_s) \star v(nT_s) + w(nT_s), \\ n &= 0, 1, \dots, N-1 \end{aligned} \quad (3)$$

where circular convolution appears in (3) due to the added CP at the source node. Note that to avoid ICI, the length of CP, i.e.,  $N_{CP} = T_{CP}/T_s$ , should be not less 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. (3) can be written in vector form as<sup>2</sup>

$$\mathbf{y}^{[s]} = \alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} [\mathbf{C}\mathbf{x}^{[s]} + \mathbf{G}\mathbf{v}] + \mathbf{w}, \quad (4)$$

where

- $\mathbf{y}^{[s]} \triangleq [y^{[s]}(0), y^{[s]}(1), \dots, y^{[s]}(N-1)]^T$ ,
- $\mathbf{\Lambda}_{\theta^{[s-d]}} \triangleq \text{Diag} \left[ e^{j\theta^{[s-d]}(0)}, e^{j\theta^{[s-d]}(1)}, \dots, e^{j\theta^{[s-d]}(N-1)} \right]$ , where  $\theta^{[s-d]}(n) = \theta^{[s-d]}(n-1) + \Delta^{[s-d]}(n)$  with  $\Delta^{[s-d]}(n)$  being a real Gaussian variable and elaborated later in Section II,
- $\mathbf{\Lambda}_{\phi^{[s-d]}} \triangleq \text{Diag} \left[ 1, e^{j2\pi\phi^{[s-d]}/N}, \dots, e^{j2\pi\phi^{[s-d]}(N-1)/N} \right]$ ,  $\phi^{[s-d]} \triangleq \Delta f^{[s-d]} T$  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  $\mathbf{c} \triangleq [c(0), c(1), \dots, c(L-1)]^T$ ,
- $\mathbf{v} \triangleq [v(-L_g+1), \dots, v(0), \dots, v(N-1)]^T$  and  $\mathbf{w} \triangleq [w(0), w(1), \dots, w(N-1)]^T$  are the sampled additive noise at the relay and destination nodes, respectively, and

$$\mathbf{G} = \begin{bmatrix} g(L_g-1) & g(L_g-2) & \dots & 0 \\ 0 & g(L_g-1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & g(0) \end{bmatrix} \quad (5)$$

<sup>2</sup>For notational convenience, we discard the term  $T_s$  in (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. Finally, 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.

2) *Digital relay processing*: In this setup, the received signal at the relay after removing CP is given by

$$z(t) = e^{j\theta^{[s-r]}(t)} e^{j\phi^{[s-r]}(t)} h(t) \star \tilde{x}^{[s]}(t) + v(t), \quad (6)$$

where  $\theta^{[s-r]}(t)$  is the PN corresponding to the source-relay link, and  $\phi^{[s-r]}(t) \triangleq 2\pi \Delta f^{[s-r]} t$  is the CFO. After sampling at a sampling rate of  $1/T_s$  and removing the CP, the received signal at the relay is given by

$$z(n) = e^{j\theta^{[s-r]}(n)} e^{j2\pi \Delta \phi^{[s-r]} n/N} h(n) \star \tilde{x}^{[s]}(n) + v(n) \quad (7)$$

where  $\phi^{[s-r]} \triangleq \Delta f^{[s-r]} T$  is the normalized CFO in the source-to-relay link. Eq. (7) can be written in vector form as

$$\mathbf{z} = \mathbf{\Lambda}_{\theta^{[s-r]}} \mathbf{\Lambda}_{\phi^{[s-r]}} \mathbf{H}\mathbf{x}^{[s]} + \mathbf{v}, \quad (8)$$

where

- $\mathbf{z} \triangleq [z(0), z(1), \dots, z(N-1)]^T$ ,
- $\mathbf{\Lambda}_{\theta^{[s-r]}} \triangleq \text{Diag} \left[ e^{j\theta^{[s-r]}(0)}, e^{j\theta^{[s-r]}(1)}, \dots, e^{j\theta^{[s-r]}(N-1)} \right]$  with  $\theta^{[s-r]}(n) = \theta^{[s-r]}(n-1) + \Delta^{[s-r]}(n)$  and  $\Delta^{[s-r]}(n)$  being a real Gaussian random variable,
- $\mathbf{\Lambda}_{\phi^{[s-r]}} \triangleq \text{Diag} \left[ 1, e^{j2\pi\phi^{[s-r]}/N}, \dots, e^{j2\pi\phi^{[s-r]}(N-1)/N} \right]$ ,
- $\mathbf{H} \triangleq \mathbf{F}^H \mathbf{\Lambda}_{\tilde{\mathbf{h}}} \mathbf{F}$ ,  $\mathbf{\Lambda}_{\tilde{\mathbf{h}}} \triangleq \text{Diag}(\tilde{\mathbf{h}})$  with  $\tilde{\mathbf{h}} \triangleq \sqrt{N} \mathbf{F} [\mathbf{c}^T, \mathbf{0}_{N-L_h,1}^T]^T$ ,
- $\mathbf{v} \triangleq [v(0), \dots, v(N-1)]^T$  is the sampled additive noise at the relay.

Note that in (8), since the matrix  $\mathbf{\Lambda}_{\theta^{[s-r]}} \mathbf{\Lambda}_{\phi^{[s-r]}}$  is unitary, vector  $\mathbf{\Lambda}_{\theta^{[s-r]}}^{-1} \mathbf{\Lambda}_{\phi^{[s-r]}}^{-1} \mathbf{v}$  has the same statistic property as the noise vector  $\mathbf{v}$ . Hence, without loss of generality, we rewrite (8) as

$$\mathbf{z} = \mathbf{\Lambda}_{\theta^{[s-r]}} \mathbf{\Lambda}_{\phi^{[s-r]}} (\mathbf{H}\mathbf{x}^{[s]} + \mathbf{v}). \quad (9)$$

Then, after adding a new CP to  $\mathbf{z}$  and scaling it by  $\alpha$ ,  $\mathbf{z}$  is delivered to the destination. The discrete vector-form received signal at the destination can be written as

$$\begin{aligned} \mathbf{y}^{[s]} &= \alpha \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \bar{\mathbf{G}} \mathbf{z} + \mathbf{w} \\ &= \alpha \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \bar{\mathbf{G}} \mathbf{\Lambda}_{\theta^{[s-r]}} \mathbf{\Lambda}_{\phi^{[s-r]}} (\mathbf{H}\mathbf{x}^{[s]} + \mathbf{v}) + \mathbf{w} \end{aligned} \quad (10)$$

where

- $\mathbf{\Lambda}_{\theta^{[r-d]}} \triangleq \text{Diag} \left[ e^{j\theta^{[r-d]}(0)}, e^{j\theta^{[r-d]}(1)}, \dots, e^{j\theta^{[r-d]}(N-1)} \right]$  with  $\theta^{[r-d]}(n)$  denoting the  $n$ th phase noise sample and  $\theta^{[r-d]}(n) = \theta^{[r-d]}(n-1) + \Delta^{[r-d]}(n)$  with  $\Delta^{[r-d]}(n)$  being a real Gaussian random variable,
- $\mathbf{\Lambda}_{\phi^{[r-d]}} \triangleq \text{Diag} \left[ 1, e^{j2\pi\phi^{[r-d]}/N}, \dots, e^{j2\pi\phi^{[r-d]}(N-1)/N} \right]$ ,  $\phi^{[r-d]} \triangleq \Delta f^{[s-d]} T$  is the normalized CFO corresponding to the relay-to-destination link,
- $\bar{\mathbf{G}}$  is a circulant channel matrix given by  $\bar{\mathbf{G}} \triangleq \mathbf{F}^H \mathbf{\Lambda}_{\tilde{\mathbf{g}}} \mathbf{F}$  with  $\mathbf{\Lambda}_{\tilde{\mathbf{g}}} \triangleq \text{Diag}(\tilde{\mathbf{g}})$ ,  $\tilde{\mathbf{g}} \triangleq \sqrt{N} \mathbf{F} [\mathbf{g}^T, \mathbf{0}_{N-L_g,1}^T]^T$ ,  $\mathbf{g} \triangleq [g(0), g(1), \dots, g(L_g-1)]^T$ , and

- $\mathbf{w}$  is the sampled additive noise at the destination as defined in (4).

When the channel coherence bandwidth is large, using the same approximation outlined in [31], [32], and [41, page 113], (10) can be approximated as

$$\begin{aligned} \mathbf{y}^{[s]} &\approx \alpha \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \mathbf{\Lambda}_{\theta^{[s-r]}} \mathbf{\Lambda}_{\phi^{[s-r]}} \bar{\mathbf{G}} (\mathbf{H}\mathbf{x}^{[s]} + \mathbf{v}) + \mathbf{w} \\ &= \alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} (\bar{\mathbf{G}}\mathbf{H}\mathbf{x}^{[s]} + \bar{\mathbf{G}}\mathbf{v}) + \mathbf{w}, \end{aligned} \quad (11)$$

where  $\mathbf{\Lambda}_{\theta^{[s-d]}} \triangleq \text{Diag} [e^{j\theta^{[s-d]}(0)}, e^{j\theta^{[s-d]}(1)}, \dots, e^{j\theta^{[s-d]}(N-1)}]$  with  $\theta^{[s-d]}(n) = \theta^{[s-r]}(n) + \theta^{[r-d]}(n)$ , and  $\mathbf{\Lambda}_{\phi^{[s-d]}} \triangleq \text{Diag} [1, e^{j2\pi\phi^{[s-d]}/N}, \dots, e^{j2\pi\phi^{[s-d]}(N-1)/N}]$  with  $\phi^{[s-d]} = \phi^{[s-r]} + \phi^{[r-d]}$ . Comparing (4) and (11), we can see that the received signals at the destination for analog relay processing and digital relay processing have similar forms. The only difference is that the relay-to-destination channel,  $g(n)$ , is linearly convolved with  $\tilde{x}^{[s]}(n)$  and  $v(n)$  for the analog relay processing case, while it is circularly convolved with  $\tilde{x}^{[s]}(n)$  and  $v(n)$  for the digital relay processing scenario. Note that the presence of the circular convolution in the digital relay processing signal model does not significantly change the structure of channel matrices  $\mathbf{C}$  and  $\mathbf{G}$  in (4). Accordingly, all the steps in the proposed estimation algorithm and performance analysis for analog relay processing can be also applied to the digital relay processing case. In fact, the four-step iterative estimator proposed in Section III has a similar form for both analog and digital relay processings. The proposed signal models for analog and digital relay processings in (4) and (11), respectively, are valid for narrowband systems. Under the assumption of wideband signal transmission, phase noise at the transmitter and receiver may not be combined in to a single term and need to be treated separately. This will result in further analyses that can determine the impact of the quality of oscillators at the source, relay, and destination on the overall performance of the relay network. Furthermore, future investigations can determine which one of these oscillators and their quality will act as a major bottleneck on system performance.

3) *Relay power constraint and scaling factor  $\alpha$* : For both analog and digital relay operations, the power scaling,  $\alpha$ , should be selected such that the transmitted power from the relay in passband is less than or equal to total available power at the relay,  $P_T^{[r]}$ , i.e., the transmitted signal from the relay in (2) or (6) should satisfy

$$\mathbb{E}(z(t)z^*(t)) \leq P_T^{[r]}. \quad (12)$$

In our scheme, we assume that the scaling factor  $\alpha$  at the relay in (2) is approximately determined in a long-term fashion.

Let us first consider the case where the CP is removed. In this case, the baseband received signal vector at the relay in the frequency domain,  $\mathbf{z}$ , is given by

$$\mathbf{z} \triangleq \tilde{\mathbf{h}} \odot \mathbf{s}^{[s]} + \mathbf{v} \in \mathbb{C}^{N \times 1}, \quad (13)$$

where  $\tilde{\mathbf{h}} \triangleq [\tilde{h}_1, \tilde{h}_2, \dots, \tilde{h}_N]^T$  with  $\tilde{h}_k = \sum_{n=0}^{L_h-1} \exp^{-j2\pi kn} h(n)$ , for  $k = 0, \dots, N-1$ . In addition, it is

assumed that  $\tilde{h}_k \sim \mathcal{CN}(0, \sum_{n=0}^{L_h-1} \sigma_h^2(n))$  with  $\sigma_h^2(n)$  denoting the variance of  $h(n)$ . Hence, with  $\mathbb{E}(\mathbf{s}^{[s]}\mathbf{s}^{[s]H}) = P_T^{[s]}\mathbf{I}$ , the received signal power of  $\mathbf{z}$ ,  $P_z$ , of  $\mathbf{z}$  is given by  $P_z = \mathbb{E}_{\mathbf{h}, \mathbf{s}}(\|\mathbf{z}\|_2^2) = N P_T^{[s]} (\sum_{n=0}^{L_h-1} \sigma_h^2(n)) + N \sigma_R^2$ . By considering the added CP, the total power of the received signal at the relay node can be approximated as  $\bar{P}_z = P_z \frac{N_{CP} + N}{N}$ . Thus, the relay scaling factor,  $\alpha$ , can be determined as  $\alpha = \sqrt{P_T^{[r]} / \bar{P}_z}$ .

### C. Training Signal Transmission from Relay to Destination

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

$$\mathbf{y}^{[r]} = \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \bar{\mathbf{G}} \mathbf{x}^{[r]} + \mathbf{w}, \quad (14)$$

where

- $\mathbf{x}^{[r]} \triangleq [x^{[r]}(0), x^{[r]}(1), \dots, x^{[r]}(N-1)]^T = \mathbf{F}^H \mathbf{s}^{[r]}$ ,  $\mathbf{s}^{[r]}$  is the frequency domain relay training signal,
- $\mathbf{\Lambda}_{\theta^{[r-d]}} \triangleq \text{Diag} [e^{j\theta^{[r-d]}(0)}, e^{j\theta^{[r-d]}(1)}, \dots, e^{j\theta^{[r-d]}(N-1)}]$ ,  $\theta^{[r-d]}(n)$  is the  $n$ -th PN sample corresponding to relay-destination link, and  $\theta^{[r-d]}(n) = \theta^{[r-d]}(n-1) + \Delta^{[r-d]}(n)$  with  $\Delta^{[r-d]}(n)$  being a real Gaussian variable and elaborated later in Section II-D
- $\mathbf{\Lambda}_{\phi^{[r-d]}} \triangleq \text{Diag} [1, e^{j2\pi\phi^{[r-d]}/N}, \dots, e^{j2\pi\phi^{[r-d]}(N-1)/N}]$ ,
- $\phi^{[r-d]}$  is the normalized CFO generated by the mismatch between the relay and destination carrier frequencies,

### D. Statistical Model of Phase Noise

Similar to [18] and based on the properties of PN in practical oscillators, PN is modeled by a Wiener process, i.e.,

$$\theta^{[i]}(n) = \theta^{[i]}(n-1) + \Delta^{[i]}(n), \quad i = [s-d], [r-d] \quad (15)$$

where  $\Delta^{[i]}(n)$  is a real Gaussian variable following  $\Delta^{[i]}(n) \sim \mathcal{N}(0, \sigma_{\Delta^{[i]}}^2)$ . Here  $\sigma_{\Delta^{[i]}}^2 = 2\pi\beta^{[i]}T_s$  with  $\beta^{[i]}$  denoting the two-sided 3-dB bandwidth of the Lorentzian spectrum of the oscillator [42], [43]. As in [18], [19], it is assumed that  $\theta^{[i]}(-1) = 0$  since the residual PN at the start of the frame is estimated as part of the channel parameters. From (15), it can be concluded that the PN vector,  $\boldsymbol{\theta}^{[i]} \triangleq [\theta^{[i]}(0), \theta^{[i]}(1), \dots, \theta^{[i]}(N-1)]^T$ , follows a Gaussian distribution, i.e.,  $\boldsymbol{\theta}^{[i]} \sim \mathcal{N}(0, \boldsymbol{\Psi}^{[i]})$ , where the covariance matrix  $\boldsymbol{\Psi}^{[i]}$  is given by

$$\boldsymbol{\Psi}^{[i]} = \sigma_{\Delta^{[i]}}^2 \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 2 & 2 & \dots & 2 & 2 \\ 1 & 2 & 3 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & N-1 & N \end{bmatrix}. \quad (16)$$

In obtaining the covariance matrix in (16), similar to prior results in this field [19], it is assumed that the PN variances are small enough such that  $\theta^{[i]}(n)$  does not reach its maximum value of  $\pi$ . This assumption is justifiable since practical oscillators have a very small PN variance as shown in [25].

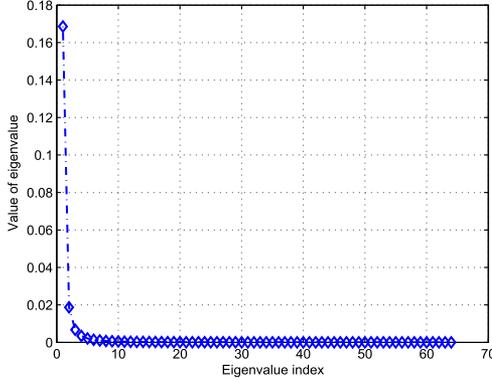


Fig. 3. Illustration of the eigenvalues of  $\Psi^{[i]}$  at  $N = 64$  and  $\sigma_{\Delta}^2 = 10^{-4}$ .

Based on the signal model in (4) and (14), it can be observed that a large number of channel, CFO, and PN parameters need to be jointly estimated, which increases the computational complexity of the receiver at the destination. Thus, to reduce estimation overhead, we take advantage of the correlation amongst the PN parameters to reduce the number of unknown parameters that need to be estimated. The eigenvalues of the covariance matrix,  $\Psi^{[i]}$  are illustrated in Fig. 3. From this figure it can be deduced that most eigenvalues of the matrix  $\Psi^{[i]}$  are close to zero. Thus, the PN vector,  $\theta^{[i]}$ , can be represented as

$$\theta^{[i]} = \Pi^{[i]} \eta^{[i]}, \quad i = [\text{s-d}], [\text{r-d}] \quad (17)$$

where  $\eta^{[i]} \sim \mathcal{N}(0, \mathbf{I}_M) \in \mathbb{C}^{M \times 1}$  is the shortened unknown PN vector containing  $M \leq N$  PN parameters, while  $\Pi^{[i]} \in \mathbb{C}^{N \times M}$  is the transformation matrix that allows for obtaining  $\eta^{[i]}$  from  $\theta^{[i]}$ . Moreover, the singular value decomposition of  $\Psi^{[i]}$  is given by  $\Psi^{[i]} = \mathbf{U}^{[i]} \mathbf{D}^{[i]} \mathbf{U}^{[i]T}$ , where  $\mathbf{U}^{[i]}$  is the  $N \times N$  eigenvector matrix of  $\Psi^{[i]}$  and  $\mathbf{D}^{[i]} = \text{Diag}(\mathbf{v}^{[i]})$ . Here,  $\mathbf{v}^{[i]} \triangleq [v_{i,0}, v_{i,1}, \dots, v_{i,N-1}]^T$  is the vector of the eigenvalues of  $\Psi^{[i]}$  arranged in decreasing order. Subsequently, the matrix  $\Pi^{[i]}$  in (17) can be selected as  $\Pi^{[i]} = \tilde{\mathbf{U}}^{[i]} \tilde{\mathbf{D}}^{[i]}$ , where  $\tilde{\mathbf{U}}^{[i]} = \mathbf{U}^{[i]}(:, 0 : M - 1)$  and  $\tilde{\mathbf{D}}^{[i]} = \text{Diag}(\tilde{\mathbf{v}}^{[i]})$  with  $\tilde{\mathbf{v}}^{[i]} \triangleq [\sqrt{v_{i,0}}, \sqrt{v_{i,1}}, \dots, \sqrt{v_{i,M-1}}]^T$ . In the subsequent sections,  $\eta^{[i]}$ , for  $i = [\text{s-d}], [\text{r-d}]$  is estimated instead of  $\theta^{[i]}$ . A suitable choice of  $M$  that allows for accurate PN tracking is presented in Section VI.

### III. PROPOSED JOINT CHANNEL, CFO AND PHASE NOISE ESTIMATION

In this section, we aim to estimate the channel, CFO and phase noise during the training phase. It is worth noting that the estimated channel and CFO can be used for the following data detection. However, as the phase noise is time-varying, it has to be re-estimated during the data detection phase.

From (4) and (14), it is found that the received signals can also be transformed into the frequency domain to conduct the estimation. However, due to the presence of phase noise and CFO, the inter-carrier interference (ICI) is introduced which makes the estimation harder in the frequency domain. Hence, in this work, we try to conduct the joint estimation in the time

domain where the number of the unknown parameters is less. To proceed, we reformulate (4) as

$$\mathbf{y}^{[\text{s}]} = \alpha \Lambda_{\theta^{[\text{s-d}]}} \Lambda_{\phi^{[\text{s-d}]}} \left( \mathbf{F}^H \Lambda_{\mathbf{s}^{[\text{s}]}} \mathbf{F}_{[L]} \mathbf{c} + \mathbf{G} \mathbf{v} \right) + \mathbf{w}, \quad (18)$$

where  $\mathbf{s}^{[\text{s}]}$  denotes the training symbol transmitted from source such that  $\mathbb{E}(\mathbf{s}^{[\text{s}]} \mathbf{s}^{[\text{s}]H}) = P_T^{[\text{s}]} \mathbf{I}_N$ ,  $P_T^{[\text{s}]}$  is the transmit power from the source,  $\Lambda_{\mathbf{s}^{[\text{s}]}} \triangleq \text{Diag}(\mathbf{s}^{[\text{s}]})$ , and  $\mathbf{F}_{[L]} \triangleq \sqrt{N} \mathbf{F}(:, 0 : L - 1)$ . Similarly, the received signal  $\mathbf{y}^{[\text{r}]}$  in (14) can be rewritten as

$$\mathbf{y}^{[\text{r}]} = \Lambda_{\theta^{[\text{r-d}]}} \Lambda_{\phi^{[\text{r-d}]}} \mathbf{F}^H \Lambda_{\mathbf{s}^{[\text{r}]}} \mathbf{F}_{[L_g]} \mathbf{g} + \mathbf{w}, \quad (19)$$

where  $\Lambda_{\mathbf{s}^{[\text{r}]}} \triangleq \text{Diag}(\mathbf{s}^{[\text{r}]})$  denotes the training symbol transmitted from relay such that  $\mathbb{E}(\mathbf{s}^{[\text{r}]} \mathbf{s}^{[\text{r}]H}) = P_T^{[\text{r}]} \mathbf{I}_N$ ,  $P_T^{[\text{r}]}$  is the transmit power from the relay, and  $\mathbf{F}_{[L_g]} \triangleq \sqrt{N} \mathbf{F}(:, 0 : L_g - 1)$ . As in [18], it is assumed that  $\mathbf{s}^{[\text{s}]}$  and  $\mathbf{s}^{[\text{r}]}$  are known constant-modulus training symbols. It is worth noting that the constant-modulus symbols are only used for the training phase, while the symbols for the data transmission can be of any modulation format.

From the detection point of view, it may appear that one only needs to estimate the CFO,  $\phi^{[\text{s-d}]}$ , and the *combined* source-relay-destination channel,  $\mathbf{c}$ , in the presence of PN,  $\theta^{[\text{s-d}]}$ . However, as shown in (18), the relay-to-destination channel,  $\mathbf{g}$ , affects the statistic of the additive noise at the destination node. Hence, it needs to be known to develop a joint PN estimation and data detection algorithm based on the MAP criterion. Consequently, here, the parameters of interest are: the CFO,  $\phi^{[\text{s-d}]}$ , the channel from source to relay,  $\mathbf{h} \triangleq [h(0), h(1), \dots, h(L_h - 1)]^T$  and the relay to destination channel,  $\mathbf{g}$ . Moreover, in addition to the parameters of interest, there are also unknown nuisance parameters, e.g., the CFO and PN from relay to destination,  $\phi^{[\text{r-d}]}$  and  $\theta^{[\text{r-d}]}$ , respectively, that also need to be jointly estimated. Using the approach in [18] and the received training signal from the relay node,  $\mathbf{y}^{[\text{r}]}$ , the MAP estimates of the CFO from relay to destination,  $\hat{\phi}^{[\text{r-d}]}$ , can be obtained as

$$\begin{aligned} \hat{\phi}^{[\text{r-d}]} &= \arg \min_{\phi^{[\text{r-d}]}} -\mathbf{1}_N^T \Im(\Lambda_{\phi^{[\text{r-d}]}} \mathbf{A} \mathbf{A}^H \Lambda_{\phi^{[\text{r-d}]}}^H)^T \\ &\times \left[ \Re(\Lambda_{\phi^{[\text{r-d}]}} \mathbf{A} \mathbf{A}^H \Lambda_{\phi^{[\text{r-d}]}}^H) + \frac{\sigma^2 P_T^{[\text{r}]}}{2} [\Psi^{[\text{r-d}]}]^{-1} \right]^{-1} \\ &\times \Im(\Lambda_{\phi^{[\text{r-d}]} \mathbf{A} \mathbf{A}^H \Lambda_{\phi^{[\text{r-d}]}^H} \mathbf{1}_N + \mathbf{1}_N^T \Lambda_{\phi^{[\text{r-d}]} \mathbf{A} \mathbf{A}^H \Lambda_{\phi^{[\text{r-d}]}^H} \mathbf{1}_N, \end{aligned} \quad (20)$$

where  $\mathbf{A} \triangleq \mathbf{Y}^{[\text{r}]} \mathbf{H} \mathbf{F}^H \Lambda_{\mathbf{s}^{[\text{r}]}} \mathbf{V}$ ,  $\mathbf{Y}^{[\text{r}]} \triangleq \text{Diag}(\mathbf{y}^{[\text{r}]})$ , and  $\mathbf{V} \triangleq \mathbf{F}(:, 0 : L_g - 1)$ . Using the estimated CFO from relay to destination,  $\hat{\phi}^{[\text{r-d}]}$ , the PN vector  $\theta^{[\text{r-d}]}$  is estimated as

$$\begin{aligned} \hat{\theta}^{[\text{r-d}]} &= \Pi^{[\text{r-d}]} \left[ \Pi^{[\text{r-d}]T} \Re \left( \hat{\Lambda}_{\phi^{[\text{r-d}]} \mathbf{A} \mathbf{A}^H \hat{\Lambda}_{\phi^{[\text{r-d}]}^H} \right) \Pi^{[\text{r-d}]} \right. \\ &\left. + \frac{\sigma^2 P_T^{[\text{r}]}}{2} \mathbf{I}_M \right]^{-1} \Pi^{[\text{r-d}]T} \Im \left( \hat{\Lambda}_{\phi^{[\text{r-d}]} \mathbf{A} \mathbf{A}^H \hat{\Lambda}_{\phi^{[\text{r-d}]}^H} \right) \mathbf{1}_N, \end{aligned} \quad (21)$$

where  $[\hat{\Lambda}_{\phi^{[r-d]}}]_{m,m} = \exp\left(\frac{j2\pi(m-1)\hat{\phi}^{[r-d]}}{N}\right)$ . Unlike the approach in [18], in (21), the shortened PN vector  $\eta^{[r-d]}$  is estimated first which reduces the complexity of the estimator by requiring the calculation of a considerably smaller matrix inverse. Based on the estimated  $\hat{\phi}^{[r-d]}$  and  $\hat{\theta}^{[r-d]}$ , the remaining parameters of interest can be estimated via  $\mathbf{y}^{[s]}$  and  $\mathbf{y}^{[r]}$ .

From (18) and (19), it can be observed that the joint estimation of channel response, CFO, and PN is a hybrid estimation problem consisting of both deterministic parameters,  $\phi^{[s-d]}$ ,  $\mathbf{h}$ ,  $\mathbf{g}$ , and random parameters,  $\eta^{[s-d]}$ . The posterior distribution of the parameters of interests given the received signals,  $\mathbf{y}^{[s]}$  and  $\mathbf{y}^{[r]}$ , can be written as

$$p\left(\phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g} | \mathbf{y}^{[s]}, \mathbf{y}^{[r]}\right) = p\left(\mathbf{y}^{[s]}, \mathbf{y}^{[r]} | \phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g}\right) \times p\left(\eta^{[s-d]}\right) / p\left(\mathbf{y}^{[s]}, \mathbf{y}^{[r]}\right). \quad (22)$$

Maximizing the posterior distribution in (22) is equivalent to minimizing the negative log-likelihood function (LLF)  $\mathcal{L}(\phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g}) = -\log(p(\mathbf{y}^{[s]}, \mathbf{y}^{[r]} | \phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g})) - \log(p(\eta^{[s-d]}))$ . Our objective is to find the joint estimates of  $\phi^{[s-d]}$ ,  $\eta^{[s-d]}$ ,  $\mathbf{h}$ , and  $\mathbf{g}$  by optimizing the following unconstrained function

$$\{\hat{\phi}^{[s-d]}, \hat{\eta}^{[s-d]}, \hat{\mathbf{h}}, \hat{\mathbf{g}}\} \propto \arg \min_{\phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g}} \mathcal{L}(\phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g}) \quad (23)$$

where  $\mathcal{L}(\phi^{[s-d]}, \eta^{[s-d]}, \mathbf{h}, \mathbf{g}) = \log \det(\Sigma) + (\mathbf{y} - \mu)^H \Sigma^{-1} (\mathbf{y} - \mu) + \frac{1}{2} \eta^{[s-d]T} \eta^{[s-d]}$ ,  $\mathbf{y} \triangleq [\mathbf{y}^{[s]T}, \mathbf{y}^{[r]T}]^T$ ,  $\mu \triangleq [(\alpha \Lambda_{\theta^{[s-d]}} \Lambda_{\phi^{[s-d]}} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]}) \mathbf{c}^T, (\Lambda_{\theta^{[r-d]}} \Lambda_{\phi^{[r-d]}} \mathbf{F}^H \Lambda_{s[r]} \mathbf{F}_{[L_g]}) \mathbf{g}^T]^T$ ,  $\Sigma \triangleq \text{Blkdiag}(\Sigma^{[r]}, \Sigma^{[d]})$ ,  $\Sigma^{[r]} = \alpha^2 \sigma_R^2 \Lambda_{\theta^{[s-d]}} \Lambda_{\phi^{[s-d]}} \mathbf{G} \mathbf{G}^H \Lambda_{\phi^{[s-d]}}^H \Lambda_{\theta^{[s-d]}}^H + \sigma_D^2 \mathbf{I}_N$ , and  $\Sigma^{[d]} = \sigma_D^2 \mathbf{I}_N$ . Although the CFO,  $\phi^{[s-d]}$ , and PN vector,  $\theta^{[s-d]}$ , are only contained in the received signal,  $\mathbf{y}^{[s]}$ , the backward substitution method proposed in [18] cannot be exploited here to solve (23) due to the unknown noise covariance matrix  $\Sigma^{[r]}$ . Such a problem is mathematically intractable and, in the followings subsection, we seek suboptimal solutions in minimizing (23) by decoupling (23) into several subproblems that can be each solved separately in an iterative approach.

*Remark 1:* For the full-duplex relaying case, if self-interference cannot be completely canceled at the relay, the received signal at the destination node can be rewritten as

$$\mathbf{y}^{[s]}(t) = \alpha e^{j\theta^{[s-d]}(t)} e^{j\phi^{[s-d]}(t)} [g(t) \star h(t) \star \tilde{x}^{[s]}(t) + g(t) \star (v(t) + q(t))] + w(t), \quad (24)$$

where  $q(t)$  is the unknown self-interference, which usually depends on the channel from relay transmitter antenna to the receiver antenna, the transmitted signal from the source and the noise introduced by the relay circuit [21], [22]. With (24), (4) becomes

$$\mathbf{y}^{[s]} = \alpha \Lambda_{\theta^{[s-d]}} \Lambda_{\phi^{[s-d]}} [\mathbf{C} \mathbf{x}^{[s]} + \mathbf{G} \bar{\mathbf{v}}] + \mathbf{w}, \quad (25)$$

where  $\bar{\mathbf{v}} = \mathbf{v} + \mathbf{q}$  with  $\mathbf{q} \triangleq [q(-L_g + 1), \dots, q(0), \dots, q(N - 1)]^T$ . Basically, introducing the self-interference  $\mathbf{q}$  changes the covariance matrix  $\Sigma^{[r]}$ , i.e., the original covariance matrix  $\Sigma^{[r]} = \alpha^2 \sigma_R^2 \Lambda_{\theta^{[s-d]}} \Lambda_{\phi^{[s-d]}} \mathbf{G} \mathbf{G}^H \Lambda_{\phi^{[s-d]}}^H \Lambda_{\theta^{[s-d]}}^H + \sigma_D^2 \mathbf{I}_N$  is changed to

$$\Sigma^{[r]} = \alpha^2 \Lambda_{\theta^{[s-d]}} \Lambda_{\phi^{[s-d]}} \mathbf{Q} \mathbf{Q}^H \Lambda_{\phi^{[s-d]}}^H \Lambda_{\theta^{[s-d]}}^H + \sigma_D^2 \mathbf{I}_N \quad (26)$$

where  $\mathbf{Q} = \sigma_R^2 \mathbf{I} + \mathbb{E}(\mathbf{q} \mathbf{q}^H)$ . In this case, the new covariance matrix  $\Sigma^{[r]}$  given in (26) should be used instead. Considering that we have assumed that the channel from relay transmitter antenna to the receive antenna changes very slowly, generally the statistics of  $\mathbf{q}$  can be estimated and updated in a long-term fashion. That is, the covariance matrix  $\mathbf{Q}$  is assumed to be known during the estimation process. In this case, our proposed estimator is still applicable to scenario that the self-interference is present at the relay.

It is worth noting that in (24) we have assumed that the self-interference link is free of phase noise. Considering the impact of phase noise in the self-interference link is an interesting topic of research and worth of future endeavor.

#### A. Phase Noise Estimation

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

$$[\hat{\eta}^{[s-d]}]^{[k+1]} \propto \arg \min_{\eta^{[s-d]}} \mathcal{L}_{\eta^{[s-d]}} \quad (27)$$

where  $\mathcal{L}_{\eta^{[s-d]}} = \log \det(\Sigma^{[r]}) + (\mathbf{y}^{[s]} - \mu^{[s-d]})^H [\Sigma^{[r]}]^{-1} (\mathbf{y}^{[s]} - \mu^{[s-d]}) + \frac{1}{2} \eta^{[s-d]T} \eta^{[s-d]}$  with  $\mu^{[s-d]} \triangleq \alpha \Lambda_{\theta^{[s-d]}} \hat{\Lambda}_{\phi^{[s-d]}}^{[k]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{c}}^{[k]}$ ,  $[\hat{\Lambda}_{\phi^{[s-d]}}^{[k]}]_{m,m} = \exp\left(\frac{j2\pi(m-1)[\hat{\phi}^{[s-d]}]^{[k]}}{N}\right)$ ,  $\hat{\mathbf{c}}^{[k]} \triangleq \hat{\mathbf{h}}^{[k]} \star \hat{\mathbf{g}}^{[k]}$ ,  $\Sigma^{[r]} = \alpha^2 \sigma_R^2 \Lambda_{\theta^{[s-d]}} \hat{\Lambda}_{\phi^{[s-d]}}^{[k]} \hat{\mathbf{G}}^{[k]} \hat{\mathbf{G}}^{[k]H} \hat{\Lambda}_{\phi^{[s-d]}}^{[k]H} \Lambda_{\theta^{[s-d]}}^H + \sigma_D^2 \mathbf{I}_N$ , and  $\hat{\mathbf{G}}^{[k]}$  is constructed from  $\hat{\mathbf{g}}^{[k]}$  based on (5). As shown in Appendix A, a closed-form solution for the PN estimate at the  $(k + 1)$ -th iteration,  $[\eta^{[s-d]}]^{[k+1]}$ , can be found as

$$[\hat{\eta}^{[s-d]}]^{[k+1]} = \left[ \Re \left( \mathbf{B}^H \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \mathbf{B} \right) + \frac{1}{2} \mathbf{I}_M \right]^{-1} \times \Re \left( \mathbf{B}^H \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \bar{\mathbf{y}}^{[s]} \right), \quad (28)$$

where  $[\hat{\Sigma}^{[r]}]^{[k]}$  is the estimate of the noise covariance matrix at the  $k$ -th iteration. Using (28), the un-shortened PN estimates at the  $(k + 1)$ -th iteration,  $[\hat{\theta}^{[s-d]}]^{[k+1]}$ , can be determined as  $[\hat{\theta}^{[s-d]}]^{[k+1]} = \Pi^{[s-d]} [\hat{\eta}^{[s-d]}]^{[k+1]}$  (see Section II-D). Finally, the noise covariance matrix,  $[\hat{\Sigma}^{[r]}]^{[k]}$ , is updated via  $[\hat{\theta}^{[s-d]}]^{[k+1]}$ .

#### 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,  $[\hat{\phi}^{[s-d]}]^{[k]}$ ,  $\hat{\mathbf{h}}^{[k]}$  and  $[\hat{\theta}^{[s-d]}]^{[k+1]}$ , respectively. To proceed, the combined channel  $\mathbf{c}$  is first rewritten as

$$\mathbf{c} = \tilde{\mathbf{G}}\mathbf{h} = \tilde{\mathbf{H}}\mathbf{g}, \quad (29)$$

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

$$\tilde{\mathbf{G}} = \begin{bmatrix} g(0) & 0 & \dots & 0 & 0 \\ g(1) & g(0) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & g(0) & 0 \\ 0 & 0 & \dots & g(1) & g(0) \end{bmatrix}, \quad (30)$$

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 the relay-to-destination channel,  $\mathbf{g}$ , is given by

$$\begin{aligned} \hat{\mathbf{g}}^{[k+1]} &\propto \arg \min_{\mathbf{g}} \mathcal{L}_{\mathbf{g}} \\ &\propto \arg \min_{\mathbf{g}} \log \det(\boldsymbol{\Sigma}) + (\mathbf{y} - \mathbf{C}\mathbf{g})^H \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{C}\mathbf{g}), \end{aligned} \quad (31)$$

where  $\mathbf{C} \triangleq \left[ \left( \alpha \hat{\Lambda}_{\theta^{[s-d]}]^{[k+1]} \hat{\Lambda}_{\phi^{[s-d]}]^{[k]} \mathbf{F}^H \boldsymbol{\Lambda}_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{H}}^{[k]} \right)^T, \left( \hat{\Lambda}_{\theta^{[r-d]}]^{[k]} \hat{\Lambda}_{\phi^{[r-d]}]^{[k]} \mathbf{F}^H \boldsymbol{\Lambda}_{s[r]} \mathbf{F}_{[L_g]} \right)^T \right]^T$  with  $\hat{\mathbf{H}}^{[k]}$  being formed by using the estimate of the source-to-relay channel in the  $k$ -th iteration  $\hat{\mathbf{h}}^{[k]}$  according to (29), and  $\boldsymbol{\Sigma} \triangleq \text{Blkdiag}(\boldsymbol{\Sigma}^{[r]}, \sigma_D^2 \mathbf{I}_N)$  with  $\boldsymbol{\Sigma}^{[r]} = \alpha^2 \sigma_R^2 \hat{\Lambda}_{\theta^{[s-d]}]^{[k+1]} \hat{\Lambda}_{\phi^{[s-d]}]^{[k]} \mathbf{G} \mathbf{G}^H \hat{\Lambda}_{\phi^{[s-d]}]^{[k]} \hat{\Lambda}_{\theta^{[s-d]}]^{[k+1]} + \sigma_D^2 \mathbf{I}_N$ . Since the covariance matrix  $\boldsymbol{\Sigma}$  is dependent on the channel response  $\mathbf{g}$  as shown in (23), it is impossible to find a closed-form solution for  $\mathbf{g}$  based on (31). Thus, we propose to use the channel covariance matrix at the  $k$ -th (previous iteration),  $\hat{\boldsymbol{\Sigma}}^{[k]}$ , to obtain an estimate of  $\mathbf{g}$  at the  $(k+1)$ -th iteration. Using this approach and by equating the gradient of  $\mathcal{L}_{\mathbf{g}}$  in (31) 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 derived as

$$\hat{\mathbf{g}}^{[k+1]} = \left( \mathbf{C}^H \left[ \hat{\boldsymbol{\Sigma}}^{[k]} \right]^{-1} \mathbf{C} \right)^{-1} \mathbf{C}^H \left[ \hat{\boldsymbol{\Sigma}}^{[k]} \right]^{-1} \mathbf{y}. \quad (32)$$

Subsequently, using  $\hat{\mathbf{g}}^{[k+1]}$ , the noise covariance  $[\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k]}$  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  $[\phi^{[s-d]}]^{[k]}$ ,  $[\theta^{[s-d]}]^{[k+1]}$ , and  $\mathbf{g}^{[k+1]}$  via the following optimization problem

$$\begin{aligned} \hat{\mathbf{h}}^{[k+1]} &\propto \arg \min_{\mathbf{h}} \mathcal{L}_{\mathbf{h}} \\ &\propto \arg \min_{\mathbf{h}} (\mathbf{y}^{[s]} - \mathbf{D}\mathbf{h})^H \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} (\mathbf{y}^{[s]} - \mathbf{D}\mathbf{h}), \end{aligned} \quad (33)$$

where  $\mathbf{D} \triangleq \alpha \hat{\Lambda}_{\theta^{[s-d]}]^{[k+1]} \hat{\Lambda}_{\phi^{[s-d]}]^{[k]} \mathbf{F}^H \boldsymbol{\Lambda}_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{G}}^{[k+1]}$ . In (33),  $\hat{\mathbf{G}}^{[k+1]}$  is formed as indicated in (30) by using  $\hat{\mathbf{g}}^{[k+1]}$ . Similar to the relay

to destination channel,  $\mathbf{g}$ , the closed-form solution of  $\mathbf{h}$  in (33) can be obtained as

$$\hat{\mathbf{h}}^{[k+1]} = \left( \mathbf{D}^H \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} \mathbf{D} \right)^{-1} \mathbf{D}^H \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} \mathbf{y}^{[s]}. \quad (34)$$

### D. CFO Estimation

In order to find an estimate of the source-destination CFO at the  $(k+1)$ -th iteration,  $[\hat{\phi}^{[s-d]}]^{[k+1]}$ , similar to the steps in (27), we approximate the covariance matrix,  $\boldsymbol{\Sigma}^{[r]}$  with  $[\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k]}$  and solve the unconstrained problem

$$\begin{aligned} [\hat{\phi}^{[s-d]}]^{[k+1]} &\propto \arg \min_{\phi^{[s-d]}} \mathcal{L}_{\phi^{[s-d]}} \\ &\propto \arg \min_{\phi^{[s-d]}} (\mathbf{y}^{[s]} - \boldsymbol{\mu}_{\phi^{[s-d]}})^H \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} \\ &\quad \times (\mathbf{y}^{[s]} - \boldsymbol{\mu}_{\phi^{[s-d]}}), \end{aligned} \quad (35)$$

where  $\boldsymbol{\mu}_{\phi^{[s-d]}} \triangleq \alpha \boldsymbol{\Lambda}_{\phi^{[s-d]}} \hat{\Lambda}_{\theta^{[s-d]}]^{[k+1]} \mathbf{F}^H \boldsymbol{\Lambda}_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{c}}^{[k+1]}$ . To make the problem in (35) more tractable and find a closed-form solution, a Taylor series approximation similar to that in (27) is applied here. Accordingly,  $e^{\frac{j2\pi m \phi^{[s-d]}}{N}}$  can be approximated as

$$\begin{aligned} e^{\frac{j2\pi m \phi^{[s-d]}}{N}} &\approx e^{\frac{j2\pi m [\hat{\phi}^{[s-d]}]^{[k]}}{N}} + \left( \phi^{[s-d]} - [\hat{\phi}^{[s-d]}]^{[k]} \right) \frac{j2\pi m}{N} \\ &\quad \times e^{\frac{j2\pi m [\hat{\phi}^{[s-d]}]^{[k]}}{N}}, \end{aligned} \quad (36)$$

where  $[\hat{\phi}^{[s-d]}]^{[k]}$  is the estimated CFO at the  $k$ -th iteration. Using (36),  $\mathcal{L}_{\phi^{[s-d]}}$  in (35) can be approximated as

$$\begin{aligned} \mathcal{L}_{\phi^{[s-d]}} &\approx \left( \mathbf{y}^{[s]} - [\hat{\Lambda}_{\phi^{[s-d]}]^{[k]} + (\phi^{[s-d]} - [\hat{\phi}^{[s-d]}]^{[k]}) \tilde{\Lambda}_{\phi^{[s-d]}]^{[k]}] \mathbf{d}^{[s-d]} \right)^H \\ &\quad \times \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} \left( \mathbf{y}^{[s]} - [\hat{\Lambda}_{\phi^{[s-d]}]^{[k]} + (\phi^{[s-d]} - [\hat{\phi}^{[s-d]}]^{[k]}) \tilde{\Lambda}_{\phi^{[s-d]}]^{[k]}] \right) \\ &\quad \times [\tilde{\Lambda}_{\phi^{[s-d]}]^{[k]}] \mathbf{d}^{[s-d]} \right), \end{aligned} \quad (37)$$

where  $\mathbf{d}^{[s-d]} \triangleq \alpha \hat{\Lambda}_{\theta^{[s-d]}]^{[k+1]} \mathbf{F}^H \boldsymbol{\Lambda}_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{c}}^{[k+1]}$  and  $\tilde{\Lambda}_{\phi^{[s-d]}]^{[k]}$  is a diagonal matrix where its  $m$ -th diagonal element is given by  $[\tilde{\Lambda}_{\phi^{[s-d]}]^{[k]}}]_{m,m} = \frac{j2\pi(m-1)}{N} e^{\frac{j2\pi(\hat{\phi}^{[s-d]}]^{[k]}(m-1)}{N}}$ . By setting  $\frac{\partial \mathcal{L}_{\phi^{[s-d]}}}{\partial \phi^{[s-d]}} = 0$  and solving for  $\phi^{[s-d]}$ , a closed-form solution for the CFO estimate at the  $(k+1)$ -th iteration,  $[\phi^{[s-d]}]^{[k+1]}$ , can be found as

$$\begin{aligned} [\hat{\phi}^{[s-d]}]^{[k+1]} &= \\ &[\hat{\phi}^{[s-d]}]^{[k]} + \frac{\Re \left( (\mathbf{y}^{[s]} - \hat{\Lambda}_{\phi^{[s-d]}]^{[k]} \mathbf{d}^{[s-d]})^H \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} \tilde{\Lambda}_{\phi^{[s-d]}]^{[k]} \mathbf{d}^{[s-d]} \right)}{\mathbf{d}^{[s-d]H} \tilde{\Lambda}_{\phi^{[s-d]}]^{[k]} \left[ \hat{\boldsymbol{\Sigma}}^{[r]} \right]^{-1} \tilde{\Lambda}_{\phi^{[s-d]}]^{[k]} \mathbf{d}^{[s-d]}}. \end{aligned} \quad (38)$$

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

The overall iterative joint estimation algorithm can be summarized as follows:

**Algorithm 1.**

- **Solve**  $\hat{\phi}^{[r-d]}$  and  $\hat{\theta}^{[r-d]}$  using (20) and (21) and initialize  $\phi^{[s-d]}$ ,  $\mathbf{g}$ ,  $\mathbf{h}$ ,  $\Sigma^{[r]}$ .
- **Repeat**
  - Update  $[\hat{\theta}^{[s-d]}]^{[k+1]}$  with  $[\hat{\phi}^{[s-d]}]^{[k]}$ ,  $\hat{\mathbf{h}}^{[k]}$  and  $\hat{\mathbf{g}}^{[k]}$  being fixed by using (28) and then update  $[\hat{\Sigma}^{[r]}]^{[k]}$ ;
  - Update  $\hat{\mathbf{g}}^{[k+1]}$  with  $[\hat{\phi}^{[s-d]}]^{[k]}$ ,  $\hat{\mathbf{h}}^{[k]}$  and  $[\hat{\theta}^{[s-d]}]^{[k+1]}$  being fixed by using (32) and then update  $[\hat{\Sigma}^{[r]}]^{[k]}$ ;
  - Update  $\hat{\mathbf{h}}^{[k+1]}$  with  $[\hat{\phi}^{[s-d]}]^{[k]}$ ,  $\hat{\mathbf{g}}^{[k+1]}$  and  $[\hat{\theta}^{[s-d]}]^{[k+1]}$  being fixed by using (34);
  - Update  $[\hat{\phi}^{[s-d]}]^{[k+1]}$  with  $\hat{\mathbf{h}}^{[k+1]}$ ,  $\hat{\mathbf{g}}^{[k+1]}$  and  $[\hat{\theta}^{[s-d]}]^{[k+1]}$  being fixed by using (38) and then update  $[\hat{\Sigma}^{[r]}]^{[k]}$  as  $[\hat{\Sigma}^{[r]}]^{[k+1]}$ ;
- **Until**  $e(n+1) - e(n) \leq \epsilon$  where  $e(n)$  denotes the obtained value of objective function in (23) after the  $n$ -th iteration and  $\epsilon$  is a pre-set convergence accuracy, we set  $\epsilon = 10^{-5}$ .

*E. Initialization of the Proposed Iterative Algorithm*

In *Algorithm 1*, initial estimates of the CFO, relay-to-destination channel, source-to-relay channel, and  $\hat{\Sigma}^{[r]}$ , which are denoted by  $[\hat{\phi}^{[s-d]}]^{[0]}$ ,  $\hat{\mathbf{g}}^{[0]}$ ,  $\hat{\mathbf{h}}^{[0]}$  and  $[\hat{\Sigma}^{[r]}]^{[0]}$ , respectively, are required. Thus, we present the initialization steps for the proposed iterative estimator. Simulations in Section VI show that the proposed estimator converges to the true values of the parameters of interest for this choice of initialization.

Since the relay-to-destination CFO and PN parameters,  $\hat{\phi}^{[r-d]}$  and  $\hat{\theta}^{[r-d]}$ , respectively, are estimated via (20) and (21), respectively, the initial relay-to-destination channel estimates,  $\hat{\mathbf{g}}^{[0]}$ , can be obtained from the received signal  $\mathbf{y}^{[r]}$  via [18]

$$\hat{\mathbf{g}}^{[0]} = \frac{1}{NP_T^{[r]}} \mathbf{F}_{[L_g]}^H \Lambda_{s[r]}^H \mathbf{F} \Lambda_{\theta[r-d]}^H \hat{\Lambda}_{\phi[r-d]}^H \mathbf{y}^{[r]}. \quad (39)$$

Next, we seek to obtain the initial estimates of the source-to-destination CFO,  $[\hat{\phi}^{[s-d]}]^{[0]}$ , and source-to-relay channel,  $\hat{\mathbf{h}}^{[0]}$ . By ignoring the PN terms, (18) can be approximated as

$$\mathbf{y}^{[s]} \approx \alpha \Lambda_{\phi[s-d]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{G}}^{[0]} \mathbf{h} + \alpha \Lambda_{\phi[s-d]} \hat{\mathbf{G}}^{[0]} \mathbf{v} + \mathbf{w},$$

where  $\hat{\mathbf{G}}^{[0]}$  and  $\hat{\mathbf{G}}^{[0]}$  are formed via  $\hat{\mathbf{g}}^{[0]}$  according to (30) and (5), respectively. Subsequently, using the ML criterion the initial estimates of the CFO,  $[\hat{\phi}^{[s-d]}]^{[0]}$ , and channel,  $\hat{\mathbf{h}}^{[0]}$ , can be obtained by minimizing

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

where  $\Sigma^{[r]} = \alpha^2 \sigma_R^2 \Lambda_{\phi[s-d]} \hat{\mathbf{G}}^{[0]} \hat{\mathbf{G}}^{[0]H} \Lambda_{\phi[s-d]}^H + \sigma_D^2 \mathbf{I}_N$ . Accordingly,  $[\hat{\phi}^{[s-d]}]^{[0]}$  and  $\hat{\mathbf{h}}^{[0]}$  can be determined as [44]

$$\begin{aligned} \hat{\mathbf{h}}^{[0]} &= \left( \alpha \hat{\mathbf{G}}^{[0]H} \mathbf{F}_{[L]}^H \Lambda_{s[s]}^H \mathbf{F} \Lambda_{\phi[s-d]}^H [\Sigma^{[r]}]^{-1} \Lambda_{\phi[s-d]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \right. \\ &\quad \left. \times \hat{\mathbf{G}}^{[0]} \right)^{-1} \hat{\mathbf{G}}^{[0]H} \mathbf{F}_{[L]}^H \Lambda_{s[s]}^H \mathbf{F} \Lambda_{\phi[s-d]}^H [\Sigma^{[r]}]^{-1} \mathbf{y}^{[s]}, \quad (40) \end{aligned}$$

$$\begin{aligned} [\hat{\phi}^{[s-d]}]^{[0]} &= \min_{\phi^{[s-d]}} (\mathbf{y}^{[s]} - \alpha \Lambda_{\phi[s-d]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{G}}^{[0]} \hat{\mathbf{h}}^{[0]})^H \\ &\quad \times [\Sigma^{[r]}]^{-1} (\mathbf{y}^{[s]} - \alpha \Lambda_{\phi[s-d]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{G}}^{[0]} \hat{\mathbf{h}}^{[0]}), \quad (41) \end{aligned}$$

where the minimization in (41) is carried out through a one-dimensional exhaustive search. As the exhaustive search is only required for the initial setup, since for subsequent OFDM packets, the previous CFO estimates can be applied to initialize the proposed iterative estimator. Also, our simulation results show that an exhaustive search with a coarse step size of  $10^{-2}$  is sufficient for the initialization process. As for the additive noise covariance matrix,  $[\hat{\Sigma}^{[r]}]^{[0]}$ , using the Taylor approximation in Section III-A, we have

$$\begin{aligned} [\hat{\Sigma}^{[r]}]^{[0]} &= \alpha^2 \sigma_R^2 \Lambda_{\theta[s-d]} \hat{\Lambda}_{\phi[s-d]} \hat{\mathbf{G}}^{[0]} \hat{\mathbf{G}}^{[0]H} \hat{\Lambda}_{\phi[s-d]}^H \Lambda_{\theta[s-d]}^H + \sigma_D^2 \mathbf{I}_N \\ &\approx \mathbf{\Omega} + \mathbf{\Omega} \odot \left( \boldsymbol{\theta}^{[s-d]} \boldsymbol{\theta}^{[s-d]H} \right) + \sigma_D^2 \mathbf{I}_N \\ &\approx \mathbf{\Omega} + \mathbf{\Omega} \odot \boldsymbol{\Psi}^{[s-d]} + \sigma_D^2 \mathbf{I}_N, \quad (42) \end{aligned}$$

where  $\mathbf{\Omega} \triangleq \alpha^2 \sigma_R^2 \hat{\Lambda}_{\phi[s-d]} \hat{\mathbf{G}}^{[0]} \hat{\mathbf{G}}^{[0]H} \hat{\Lambda}_{\phi[s-d]}^H$ . In (42), since  $\boldsymbol{\theta}^{[s-d]}$  is not known, we use the expectation  $\mathbb{E}(\boldsymbol{\theta}^{[s-d]} \boldsymbol{\theta}^{[s-d]H}) = \boldsymbol{\Psi}^{[s-d]}$  instead of the term  $\boldsymbol{\theta}^{[s-d]} \boldsymbol{\theta}^{[s-d]H}$ . This allows for a closed-form expression for obtaining the source-to-relay channel estimates.

*Remark 2:* Similar to point-to-point systems [19], [39], [40], while jointly estimating the channel, CFO, and PN parameters in OFDM relay systems, a residual ambiguity may exist amongst these parameters. In what follows, we demonstrate the impact of this ambiguity on evaluating the performance of the proposed estimators.

The negative LLF in (23) can be rewritten as

$$\begin{aligned} &\{\hat{\phi}^{[s-d]}, \hat{\theta}^{[s-d]}, \hat{\phi}^{[r-d]}, \hat{\theta}^{[r-d]}, \hat{\mathbf{h}}, \hat{\mathbf{g}}\} \\ &\propto \arg \min \log \det(\Sigma) + (\mathbf{y} - \boldsymbol{\mu})^H \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \\ &\quad + \frac{1}{2} \boldsymbol{\theta}^{[s-d]T} [\boldsymbol{\Psi}^{[s-d]}]^{-1} \boldsymbol{\theta}^{[s-d]} + \frac{1}{2} \boldsymbol{\theta}^{[r-d]T} [\boldsymbol{\Psi}^{[r-d]}]^{-1} \boldsymbol{\theta}^{[r-d]}. \quad (43) \end{aligned}$$

Eq. (43) is similar to (23) with the exception that  $\phi^{[r-d]}$  and  $\boldsymbol{\theta}^{[r-d]}$  are also treated as parameters of interest and  $\boldsymbol{\eta}^{[i]}$  is replaced with  $\boldsymbol{\theta}^{[i]}$ . At very high SNR, i.e.,  $\sigma_R^2 \rightarrow 0$  and  $\sigma_D^2 \rightarrow 0$ , (43) can be further simplified as

$$\begin{aligned} &\{\hat{\phi}^{[s-d]}, \hat{\theta}^{[s-d]}, \hat{\phi}^{[r-d]}, \hat{\theta}^{[sr-d]}, \hat{\mathbf{h}}, \hat{\mathbf{g}}\} \\ &\propto \arg \min \log \det(\Sigma) + (\mathbf{y} - \boldsymbol{\mu})^H \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}). \quad (44) \end{aligned}$$

From (44) it can be concluded that the metric for estimation of parameters of interest is solely dependent on the received signal instead of the prior information at high SNR [44]. Moreover, it can be straightforwardly shown that the received training symbols, e.g.,  $\mathbf{y}^{[r]}$ , are not altered under a common

phase rotation,  $\varphi_g$ , between the channel response,  $\hat{\mathbf{g}}$ , and PN parameters  $\hat{\boldsymbol{\theta}}^{[r-d]}$ , i.e.,

$$\hat{\mathbf{g}} \rightarrow \exp(-j\varphi_g)\mathbf{g}, \hat{\boldsymbol{\theta}}^{[r-d]} \rightarrow \boldsymbol{\theta}^{[r-d]} + \varphi_g \mathbf{1}_N. \quad (45)$$

Thus, the common phase rotation,  $\varphi_g$ , can be considered as a phase ambiguity amongst the channel and PN parameters that cannot be estimated. Using a similar approach, it can also be shown that there exists a phase ambiguity between the estimate of the source-to-relay channel,  $\hat{\mathbf{h}}$ , and the estimate of the source-to-destination PN parameter,  $\hat{\boldsymbol{\theta}}^{[s-d]}$  given by

$$\hat{\mathbf{h}} \rightarrow \exp(-j\varphi_h)\mathbf{h}, \hat{\boldsymbol{\theta}}^{[s-d]} \rightarrow \boldsymbol{\theta}^{[s-d]} + (\varphi_h + \varphi_g)\mathbf{1}_N, \quad (46)$$

where  $\varphi_h$  is the phase ambiguity associated with channel  $\mathbf{h}$ . In addition to the ambiguity between channel and PN, a phase ambiguity may also exist between the PN and CFO as:

$$\begin{aligned} \hat{\phi}^{[s-d]} &\rightarrow \phi^{[s-d]} - \epsilon^{[s-d]}, \quad \hat{\phi}^{[r-d]} \rightarrow \phi^{[r-d]} - \epsilon^{[r-d]}, \\ \hat{\boldsymbol{\theta}}^{[s-d]} &\rightarrow \boldsymbol{\theta}^{[s-d]} + \boldsymbol{\epsilon}^{[s-d]}, \quad \hat{\boldsymbol{\theta}}^{[r-d]} \rightarrow \boldsymbol{\theta}^{[r-d]} + \boldsymbol{\epsilon}^{[r-d]}, \end{aligned} \quad (47)$$

where  $[\boldsymbol{\epsilon}^{[s-d]}]_m = \frac{2\pi(m-1)\epsilon^{[s-d]}}{N}$  and  $[\boldsymbol{\epsilon}^{[r-d]}]_m = \frac{2\pi(m-1)\epsilon^{[r-d]}}{N}$ . 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 the channel responses  $\hat{\mathbf{h}}$  and  $\hat{\mathbf{g}}$ , can be computed as<sup>3</sup>

$$\text{MSE}_{\mathbf{g}} = \mathbb{E}_{\mathbf{g}} \left( \|\hat{\mathbf{g}} - \mathbf{g}\|_2^2 \right), \text{MSE}_{\mathbf{h}} = \mathbb{E}_{\mathbf{h}} \left( \|\hat{\mathbf{h}} - \mathbf{h}\|_2^2 \right), \quad (48)$$

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

$$\text{MSE}_{\phi^{[s-d]}, \boldsymbol{\theta}^{[s-d]}} = \mathbb{E}_{\boldsymbol{\theta}^{[s-d]}(m), \phi^{[s-d]}} \left( \|\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}\|_2^2 \right), \quad (49)$$

where  $\boldsymbol{\delta} = \boldsymbol{\delta} - \delta_0 \mathbf{1}$ ,  $\hat{\boldsymbol{\delta}} = \hat{\boldsymbol{\delta}} - \hat{\delta}_0 \mathbf{1}$ ,  $\boldsymbol{\delta} = [\delta_0, \delta_1, \dots, \delta_{N-1}]^T$  with  $\delta_m = \theta^{[s-d]}(m) + \frac{2\pi(m-1)\phi^{[s-d]}}{N}$ , and  $\hat{\boldsymbol{\delta}} = [\hat{\delta}_0, \hat{\delta}_1, \dots, \hat{\delta}_{N-1}]^T$  with  $\hat{\delta}_m = \hat{\theta}^{[s-d]}(m) + \frac{2\pi(m-1)\hat{\phi}^{[s-d]}}{N}$ .

#### F. Complexity Analysis

In this subsection, we analyze the computational complexity of the proposed MAP based joint estimation algorithm, which is defined as the total number of complex multiplications plus number of additions required by Algorithm 1 [45]. We denote the computational complexity by  $C = C_M + C_A$ , where  $C_M$  denotes the number of multiplications and  $C_A$  denotes the number of additions. The number of iterations required by the proposed algorithm is denoted by  $\mathcal{N}$  (refer to the simulation results shown in Fig. 6 to see the required iteration number). Subsequently,  $C_M$  and  $C_A$  can be determined as

<sup>3</sup>Note that for a deterministic value, the expectation of the square error is taken over multiple Monte-Carlo simulations.

$$\begin{aligned} C_M &= \mathcal{N} \left[ C_M^{(28)} + C_M^{(32)} + C_M^{(34)} + C_M^{(38)} \right] + C_M^I, \\ C_A &= \mathcal{N} \left[ C_A^{(28)} + C_A^{(32)} + C_A^{(34)} + C_A^{(38)} \right] + C_A^I, \end{aligned} \quad (50)$$

where  $C_M^{(n)}$  and  $C_A^{(n)}$  denote the required numbers of multiplications and additions for equation (n), respectively, and the specific number of  $C_M^{(n)}$  and  $C_A^{(n)}$  are given in (B.1);  $C_M^I$  and  $C_A^I$  denote the required numbers of multiplications and additions for algorithm initialization, and are expressed as

$$\begin{aligned} C_M^I &= C_M^{I(20)} + C_M^{I(21)} + C_M^{I(39)} + C_M^{I(40)} + C_M^{I(41)}, \\ C_A^I &= C_A^{I(20)} + C_A^{I(21)} + C_A^{I(39)} + C_A^{I(40)} + C_A^{I(41)}, \end{aligned} \quad (51)$$

where we use  $C_M^{I(n)}$  and  $C_A^{I(n)}$  to denote the required numbers of multiplications and additions for equation (n) during initialization, respectively, and the specific number of  $C_M^{I(n)}$  and  $C_A^{I(n)}$  are given in (B.2).

In comparison, it is worth noting that directly solving the original MAP problem in (23) is impossible, since we need to exhaustively search the solution of  $\phi^{[s-d]}$ ,  $\boldsymbol{\theta}^{[s-d]}$ ,  $\mathbf{h}$ , and  $\mathbf{g}$ . In (23), these parameters are coupled with each other. Also, the elements of  $\mathbf{h}$  and  $\mathbf{g}$  are complex value without bounded range. Thus, the complexity of directly solving the original MAP problem is infinite.

#### IV. THE HYBRID CRAMÉR-RAO LOWER BOUND

In this section, the HCRLB for joint estimation of channel, CFO, and PN in OFDM relay networks is derived.

As stated in *Remark 2*, due to the ambiguities between the estimation of channel responses, CFO, and PN, (18) and (19) are first rewritten as

$$\begin{aligned} \mathbf{y}^{[s]} &= \alpha \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{[s-d]}} \boldsymbol{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{F}^H \underline{\boldsymbol{\Lambda}}_{s[s]} \mathbf{F}_{[L]}\mathbf{c} + \mathbf{G}\mathbf{v} \right) + \mathbf{w}, \\ \mathbf{y}^{[r-d]} &= \boldsymbol{\Lambda}_{\boldsymbol{\theta}^{[r-d]}} \boldsymbol{\Lambda}_{\phi^{[r-d]}} \mathbf{F}^H \underline{\boldsymbol{\Lambda}}_{s[r]} \mathbf{F}_{[L_g]}\mathbf{g} + \mathbf{w}, \end{aligned} \quad (52)$$

where  $\mathbf{c} \triangleq \mathbf{h} \star \mathbf{g}$  with  $\mathbf{h}$  and  $\mathbf{g}$  defined in (48),  $[\underline{\boldsymbol{\Lambda}}_{s[s]}]_{m,m} \triangleq s_{m-1}^{[s]} \exp(j\mathcal{L}c(0))$ ,  $[\underline{\boldsymbol{\Lambda}}_{s[r]}]_{m,m} \triangleq s_{m-1}^{[r]} \exp(j\mathcal{L}g(0))$  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  $\mathbf{G}$  is constructed using  $\mathbf{g}$  similar to (5). Accordingly, the HCRLB for the estimation problem is given by [46]

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

where  $\boldsymbol{\lambda} \triangleq \left[ \phi^{[s-d]}, \boldsymbol{\theta}^{[s-d]T}, \phi^{[r-d]}, \boldsymbol{\theta}^{[r-d]T}, \mathbf{g}_0, \mathfrak{R}(\tilde{\mathbf{g}})^T, \mathfrak{I}(\tilde{\mathbf{g}})^T, \mathfrak{R}(\tilde{\mathbf{h}})^T, \mathfrak{I}(\tilde{\mathbf{h}})^T \right]^T$  denotes the vector of parameters of interest,  $\tilde{\mathbf{g}} \triangleq \mathbf{g}(1:L_g-1)$ ,  $\tilde{\mathbf{h}} \triangleq \mathbf{h}(1:L_h-1)$ , and  $\mathbf{B}$  is the Bayesian information matrix (BIM) that is given by

$$\begin{aligned} \mathbf{B} &= \mathbb{E}_{\boldsymbol{\theta}^{[s-d]}, \boldsymbol{\theta}^{[r-d]}} [\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})] + \mathbb{E}_{\boldsymbol{\theta}^{[s-d]}, \boldsymbol{\theta}^{[r-d]}} \left[ -\Delta_{\boldsymbol{\lambda}}^{\lambda} \log p(\boldsymbol{\theta}^{[s-d]}) \right] \\ &\quad + \mathbb{E}_{\boldsymbol{\theta}^{[s-d]}, \boldsymbol{\theta}^{[r-d]}} \left[ -\Delta_{\boldsymbol{\lambda}}^{\lambda} \log p(\boldsymbol{\theta}^{[r-d]}) \right]. \end{aligned} \quad (53)$$

In (53),  $\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda}) = \mathbb{E}_{\mathbf{y}} \left[ -\Delta_{\boldsymbol{\lambda}}^{\lambda} \log p(\mathbf{y}; \boldsymbol{\lambda}) \right]$  denotes the Fisher's information matrix (FIM). In the following subsection the BIM in (53) is derived in detail.

### A. Derivation of $\mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} [\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})]$

In order to derive  $\mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} [\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})]$ , we first derive the FIM for the parameters of interest  $\boldsymbol{\lambda}$ .

*Theorem 1:* The  $Q \times Q$  Fisher's information matrix  $\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})$  with  $Q = 2(N + L_g + L_h)$  for the joint estimation problem is given by

$$\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda}) = \begin{bmatrix} \text{FIM}_{1,1} + \Upsilon_{1,1} & \dots & \text{FIM}_{1,Q} + \Upsilon_{1,Q} \\ \vdots & \ddots & \vdots \\ \text{FIM}_{Q,1} + \Upsilon_{Q,1} & \dots & \text{FIM}_{Q,Q} + \Upsilon_{Q,Q} \end{bmatrix}. \quad (54)$$

In (54),  $\text{FIM}_{i,j}$ , for  $i, j = 1, 2, \dots, Q$ , is determined as

$$\text{FIM}_{i,j} = 2\Re \left( \boldsymbol{\rho}_i^H \boldsymbol{\Sigma}^{-1} \boldsymbol{\rho}_j \right), \quad (55)$$

where  $\boldsymbol{\Sigma} = \text{Blkdiag}(\boldsymbol{\Sigma}^{[r]}, \sigma_D^2 \mathbf{I}_N)$  with  $\boldsymbol{\Sigma}^{[r]} = \alpha^2 \sigma_R^2 \boldsymbol{\Lambda}_{\theta^{[s-d]}} \boldsymbol{\Lambda}_{\phi^{[s-d]}} \mathbf{G} \mathbf{G}^H \boldsymbol{\Lambda}_{\phi^{[s-d]}}^H \boldsymbol{\Lambda}_{\theta^{[s-d]}}^H + \sigma_D^2 \mathbf{I}_N$ , and  $\boldsymbol{\rho}_i$  is given in (C.1) and (C.2) in Appendix C. Moreover, in (54),  $\Upsilon_{i,j}$  is given by (56), shown at the bottom of the page, where  $\mathbf{Q}_i \triangleq \text{Blkdiag}(\mathbf{W}_i, \mathbf{0}_{N \times N})$  with  $\mathbf{W}_i$  being given in (C.3) in Appendix C.

*Proof:* See Appendix D.  $\blacksquare$

Although the FIM can be obtained in closed-form, a closed-form expression for  $\mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} [\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})]$  cannot be obtained due to the presence of a complex multidimensional integration. Hence, here,  $\mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} [\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})]$  is numerically evaluated.

### B. Derivation of $\mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} \left[ -\Delta_{\lambda}^{\lambda} \log p(\boldsymbol{\theta}^{[s-d]}) \right]$ and $\mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} \left[ -\Delta_{\lambda}^{\lambda} \log p(\boldsymbol{\theta}^{[r-d]}) \right]$

Since  $p(\boldsymbol{\theta}^{[s-d]})$  and  $p(\boldsymbol{\theta}^{[r-d]})$  are independent of  $\phi^{[s-d]}$ ,  $\phi^{[r-d]}$ ,  $\mathbf{g}$ , and  $\mathbf{h}$ , we can straightforwardly obtain

$$\begin{aligned} & \mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} \left[ -\Delta_{\lambda}^{\lambda} \log p(\boldsymbol{\theta}^{[s-d]}) \right] \\ &= \text{Blkdiag} \left( 0, [\boldsymbol{\Psi}^{[s-d]}]^{-1}, 0, \mathbf{0}_{N \times N}, \mathbf{0}_{(2L_g-1) \times (2L_g-1)}, \right. \\ & \quad \left. \mathbf{0}_{(2L_h-1) \times (2L_h-1)} \right), \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E}_{\theta^{[s-d]}, \theta^{[r-d]}} \left[ -\Delta_{\lambda}^{\lambda} \log p(\boldsymbol{\theta}^{[r-d]}) \right] \\ &= \text{Blkdiag} \left( 0, \mathbf{0}_{N \times N}, 0, [\boldsymbol{\Psi}^{[r-d]}]^{-1}, \mathbf{0}_{(2L_g-1) \times (2L_g-1)}, \right. \\ & \quad \left. \mathbf{0}_{(2L_h-1) \times (2L_h-1)} \right). \end{aligned}$$

Finally, the BIM in (53) can be calculated using the results in Sections IV-A and IV-B.

*Remark 3:* Since we consider to jointly estimate three channel parameters, PN, CFO, and the channels at the destination

node, the considered estimation problem in this paper is significantly more complex than the one considered in [47], our analysis has shown that the derivation of the asymptotic CRB is very challenging to derive, which is beyond the scope of this work and will be considered in our future work.

### C. Derivation of the Transformed HCRLB

As shown in Remark 2, due to the ambiguities in the estimation of parameters of interest, the MSE of the CFO and PN is computed jointly as shown in (49). Consequently, the parameters of interests,  $\boldsymbol{\lambda}$  needs to be transformed to  $\boldsymbol{\lambda}_{\text{mod}} = \left[ \boldsymbol{\delta}^T, \phi^{[r-d]}, \boldsymbol{\theta}^{[r-d]T}, \mathbf{g}_0, \Re(\tilde{\mathbf{g}})^T, \Im(\tilde{\mathbf{g}})^T, \mathbf{h}_0, \Re(\tilde{\mathbf{h}})^T, \Im(\tilde{\mathbf{h}})^T \right]^T$ . Since  $\delta_m = \theta^{[s-d]}(m) + \frac{2\pi(m-1)\phi^{[s-d]}}{N}$ , this transformation can be written in matrix form as

$$\boldsymbol{\lambda}_{\text{mod}} = \boldsymbol{\Xi} \boldsymbol{\lambda},$$

where  $\boldsymbol{\Xi} \triangleq \boldsymbol{\Xi}_2 \boldsymbol{\Xi}_1$ ,  $\boldsymbol{\Xi}_1 \triangleq \text{Blkdiag}(0, \tilde{\boldsymbol{\Xi}}_1, 1, \mathbf{I}_{N \times N}, \mathbf{I}_{(2L_g-1) \times (2L_g-1)}, \mathbf{I}_{(2L_h-1) \times (2L_h-1)})$ ,  $\boldsymbol{\Xi}_2 \triangleq \text{Blkdiag}(\tilde{\boldsymbol{\Xi}}_2, 1, \mathbf{I}_{N \times N}, \mathbf{I}_{(2L_g-1) \times (2L_g-1)}, \mathbf{I}_{(2L_h-1) \times (2L_h-1)})$ , and

$$\begin{aligned} \tilde{\boldsymbol{\Xi}}_1 &\triangleq \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{N \times N}, \\ \tilde{\boldsymbol{\Xi}}_2 &\triangleq \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ \frac{2\pi}{N} & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{2\pi(N-1)}{N} & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{N \times (N+1)}. \end{aligned}$$

Thus, the HCRLB for the transformed parameters of interest  $\boldsymbol{\lambda}_{\text{mod}}$  is given by  $\mathbf{HCRLB}_{\text{mod}} = \boldsymbol{\Xi} \mathbf{B}^{-1} \boldsymbol{\Xi}^T$  [44].

## V. DATA DETECTION IN PRESENCE OF PHASE NOISE

In this section, a receiver structure for data detection at the destination in the presence of PN is proposed. Since the PN parameters vary over an OFDM symbol, they need to be accurately tracked over the length of each symbol. Hence, we propose the transmission of comb-type data symbols from the source node, i.e., each transmitted symbol consists of both pilot and data subcarriers (see Fig. 4). As discussed in Section II-D, in each OFDM data symbol, it is sufficient to estimate the shortened PN vector of length  $M$ . Thus, the maximum number of subcarriers utilized for data transmission is  $N - M$ . Based on that, if the coherence time of the OFDM channel is  $L_c$  symbols, the overhead of pilots in both the estimation phase and the data transmission phase is  $\frac{2N+(L_c-2)M}{L_c N}$ .

$$\Upsilon_{i,j} = \begin{cases} \Upsilon_{i,j} = \text{Tr} \left[ \boldsymbol{\Sigma}^{-1} \mathbf{Q}_i \boldsymbol{\Sigma}^{-1} \mathbf{Q}_j \right] & i = 1, 2, \dots, N+1 \\ & j = 2N+3, 2N+4, \dots, 2N+2L_g+1, \\ \Upsilon_{i,j} = 0 & \text{Otherwise} \end{cases} \quad (56)$$

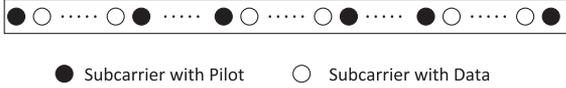


Fig. 4. Illustration of the comb-type data symbol.

The received signal at the destination during the data transmission interval is given by

$$\begin{aligned} \mathbf{y}^{[s]} &= \alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} (\mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}} \mathbf{s}^{[s]} + \mathbf{G} \mathbf{v}) + \mathbf{w} \\ &= \mathbf{T}_T \mathbf{s}_T^{[s]} + \mathbf{T}_D \mathbf{s}_D^{[s]} + \alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{G} \mathbf{v} + \mathbf{w}, \end{aligned} \quad (57)$$

where  $\mathbf{s}^{[s]}$  denotes the comb-type signal transmitted during the data transmission interval with  $\mathbb{E}(\mathbf{s}^{[s]} \mathbf{s}^{[s]H}) = P_T^{[s]} \mathbf{I}_N$ ,  $\mathbf{s}_T^{[s]}$  and  $\mathbf{s}_D^{[s]}$  denote the pilot and data vector contained in  $\mathbf{s}^{[s]}$ , respectively, and  $\mathbf{T}_T$  and  $\mathbf{T}_D$  are the associated sub-matrices of the combined channel,  $\alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}}$ , corresponding to  $\mathbf{s}_T^{[s]}$  and  $\mathbf{s}_D^{[s]}$ , respectively. Since in (57), the unknown PN vector  $\theta^{[s-d]}$  and data vector  $\mathbf{s}_D^{[s]}$  are coupled with each other, similar to the estimation part, an iterative method is applied here. By using the MAP criterion as in (22), the joint estimation of PN parameters and data can be formulated as

$$\begin{aligned} \{\hat{\theta}^{[s-d]}, \mathbf{s}_D^{[s]}\} &= \arg \min_{\theta^{[s-d]}, \mathbf{s}_D^{[s]}} \log \det(\mathbf{\Sigma}^{[r]}) \\ &\quad + (\mathbf{y}^{[s]} - \boldsymbol{\mu})^H [\mathbf{\Sigma}^{[r]}]^{-1} (\mathbf{y}^{[s]} - \boldsymbol{\mu}) \\ &\quad + \frac{1}{2} \eta^{[s-d]T} \eta^{[s-d]}, \end{aligned} \quad (58)$$

where  $\boldsymbol{\mu} \triangleq \alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}} \mathbf{s}^{[s]}$  with  $\mathbf{\Lambda}_{\hat{\phi}^{[s-d]}}$  and  $\mathbf{\Lambda}_{\hat{\mathbf{c}}}$  being determined based on the estimated CFO and channels,  $\hat{\phi}^{[s-d]}$  and  $\hat{\mathbf{c}}$ , respectively, and  $\mathbf{\Sigma}^{[r]} = \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \hat{\mathbf{G}} \hat{\mathbf{G}}^H \mathbf{\Lambda}_{\phi^{[s-d]}}^H \mathbf{\Lambda}_{\theta^{[s-d]}}^H + \sigma_D^2 \mathbf{I}_N$  is the noise covariance matrix that is calculated via the estimated channels,  $\hat{\mathbf{g}}$ , and CFO,  $\hat{\phi}^{[s-d]}$ . First, the data symbols at the  $k$ -th iteration,  $[\mathbf{s}_D^{[s]}]^{[k]}$ , are used to estimate the PN at the  $(k+1)$ -th iteration,  $[\theta^{[s-d]}]^{[k+1]}$ . To obtain a closed-form solution, as in Section III-A, (57) is approximated by

$$\begin{aligned} \mathbf{y}^{[s]} &\approx \alpha \mathbf{\Lambda}_{\hat{\phi}^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}} [\mathbf{s}^{[s]}]^{[k]} + \text{Diag} \left( j \alpha \mathbf{\Lambda}_{\hat{\phi}^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}} [\mathbf{s}^{[s]}]^{[k]} \right) \\ &\quad \times \boldsymbol{\Pi}^{[s-d]} \boldsymbol{\eta}^{[s-d]} + \alpha \hat{\boldsymbol{\Lambda}}_{\theta^{[s-d]}}^{[k]} \mathbf{\Lambda}_{\hat{\phi}^{[s-d]}} \hat{\mathbf{G}} \mathbf{v} + \mathbf{w}, \end{aligned}$$

where  $\boldsymbol{\eta}^{[s-d]}$  denotes the shorten PN vector. By equating the gradient of (58) to zero,  $[\hat{\theta}^{[s-d]}]^{[k+1]}$  can be determined as

$$\begin{aligned} [\hat{\theta}^{[s-d]}]^{[k+1]} &= \boldsymbol{\Pi}^{[s-d]} \left( \Re \left( \mathbf{M}^H [\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k]} \right)^{-1} \mathbf{M} \right) + \frac{1}{2} \mathbf{I}_M \Big)^{-1} \\ &\quad \Re \left( \mathbf{M}^H \times [\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k]} \right)^{-1} (\mathbf{y}^{[s]} - \alpha \mathbf{\Lambda}_{\hat{\phi}^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}} [\mathbf{s}^{[s]}]^{[k]}), \end{aligned} \quad (59)$$

where  $\mathbf{M} \triangleq \text{Diag} \left( j \alpha \mathbf{\Lambda}_{\hat{\phi}^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}} [\mathbf{s}^{[s-d]}]^{[k]} \right) \boldsymbol{\Pi}^{[s]}$  and  $[\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k]} = \alpha^2 \sigma_R^2 \hat{\boldsymbol{\Lambda}}_{\theta^{[s-d]}}^{[k]} \mathbf{\Lambda}_{\hat{\phi}^{[s-d]}} \hat{\mathbf{G}} \hat{\mathbf{G}}^H \mathbf{\Lambda}_{\phi^{[s-d]}}^H \hat{\boldsymbol{\Lambda}}_{\theta^{[s-d]}}^{[k]H} + \sigma_D^2 \mathbf{I}_N$ . Secondly, using  $[\hat{\theta}^{[s-d]}]^{[k+1]}$  and the noise covariance matrix at the  $(k+1)$ -th

---

**Algorithm 2.**


---

- **Initialize**  $\mathbf{s}_D^{[s]}$  and  $\mathbf{\Sigma}^{[r]}$
  - **Repeat**
    - Update  $[\theta^{[s-d]}]^{[k+1]}$  with the estimated  $[\mathbf{s}_D^{[s]}]^{[k]}$  by using (59) and then update  $[\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k]}$  as  $[\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k+1]}$ ;
    - Update  $[\mathbf{s}_D^{[s]}]^{[k+1]}$  with the estimated  $[\theta^{[s-d]}]^{[k+1]}$  by using (60);
  - **Until**  $q(n+1) - q(n) \leq \epsilon$  where  $q(n)$  denotes the obtained value of objective function in (58) after the  $n$ -th iteration and  $\epsilon$  is a pre-set convergence accuracy.
- 

iteration,  $[\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k+1]}$ , an estimate of the transmitted symbols at the  $(k+1)$ -th iteration can be obtained as

$$\begin{aligned} [\mathbf{s}_D^{[s]}]^{[k+1]} &= \left( \hat{\mathbf{T}}_D^{[k+1]H} [\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k+1]} \right)^{-1} \hat{\mathbf{T}}_D^{[k+1]} \\ &\quad \times [\hat{\boldsymbol{\Sigma}}^{[r]}]^{[k+1]} \left( \mathbf{y}^{[s]} - \hat{\mathbf{T}}_T^{[k+1]} \mathbf{s}_T^{[s]} \right). \end{aligned} \quad (60)$$

In (60), although  $\hat{\mathbf{T}}_T^{[k+1]}$  and  $\hat{\mathbf{T}}_D^{[k+1]}$  are defined similar to  $\mathbf{T}_T$  and  $\mathbf{T}_D$  in (57), they are obtained via the estimates  $[\hat{\theta}^{[s-d]}]^{[k+1]}$ ,  $\hat{\phi}^{[s-d]}$ , and  $\hat{\mathbf{c}}$ . The overall iterative detector is given below.

In Algorithm 2, initial estimates of  $\mathbf{s}_D^{[s]}$  and  $\mathbf{\Sigma}^{[r]}$  are obtained similar to that of the training interval.

*Remark 4:* As indicated here, the ambiguities associated with calculating the MSE for channel response, CFO, and PN parameters do not affect the data transmission interval. Let us denote the ambiguities of the channels and CFO in the training phase as  $\hat{\mathbf{g}} \rightarrow \exp(-j\varphi_g) \mathbf{g}$ ,  $\hat{\mathbf{c}} \rightarrow \exp(-j(\varphi_h + \varphi_g)) \mathbf{c}$  and  $\hat{\phi}^{[s-d]} \rightarrow \phi^{[s-d]} - \epsilon^{[s-d]}$ . These ambiguities can be combined during the data transmission phase in the overall estimate of the PN parameters  $\theta^{[s-d]}$  in (57), which can be written as  $\hat{\theta}^{[s-d]} \rightarrow \theta^{[s-d]} + (\varphi_g + \varphi_h) \mathbf{1} + \epsilon^{[s-d]}$  ( $\epsilon^{[s-d]}$  is defined in (47)). It can be clearly observed that these ambiguities do not affect the overall channel response,  $\alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F} \mathbf{\Lambda}_{\hat{\mathbf{c}}}$ , and the received signal in (57).

*Remark 5:* In Algorithm 2, we apply a suboptimal linear detection approach by first conducting the equalization. The complexity of the linear detection scheme includes estimating  $\theta^{[s-d]}$  in (59) and estimating  $\mathbf{s}_D^{[s]}$  in (60). Assume that the numbers of subcarriers used for training and data transmission in each OFDM symbol are  $M$  and  $N - M$ , respectively. Denote the required numbers of multiplications and additions for equation (n) by  $C_M^{(n)}$  and  $C_A^{(n)}$ , respectively, the complexity of the linear detection is given by

$$C = \mathcal{N} (C_M^{(58)} + C_M^{(59)} + C_A^{(58)} + C_A^{(59)}) \quad (61)$$

where  $\mathcal{N}$  denotes the number of iterations required by the proposed linear detection and

$$\begin{aligned} C_M^{(58)} &= 2N^3 + N^2M + NM^2 + 2NM + M^2 \\ C_M^{(59)} &= N^3 + (N - M)^3 + (N - M)N^2 + N(N - M)^2 \\ &\quad + N(N - M) + (N - M)^2 \end{aligned}$$

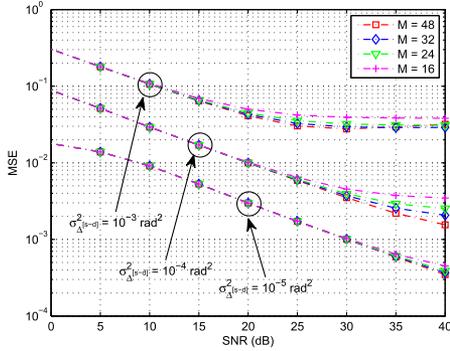


Fig. 5. The MSE of phase noise estimation with different  $M$ .

$$\begin{aligned}
 C_A^{(58)} &= N + N(N-1) + 2NM(N-1) + M + M(N-1) \\
 &\quad + M(M-1) + N(M-1), \\
 C_A^{(59)} &= N + N(M-1) + 2N(N-1)(N-M) \\
 &\quad + (N-M)(N-1) + (N-M)(N-M-1). \quad (62)
 \end{aligned}$$

It is worth noting that the maximum-likelihood (ML) detector may be used to further improve the detection performance. However, in ML detection, we need to exhaustively search  $\theta^{[s-d]}$  and  $\mathbf{s}_D^{[s]}$  in (58). Consider that the elements of  $\theta^{[s-d]}$  are complex value without a bounded range, its high computational complexity may not be affordable in practical systems in our case.

## VI. SIMULATION RESULTS

In this section, extensive simulations are carried out to evaluate the performance of the proposed algorithms. Without loss of generality, it is assumed that the noise powers at relay and destination nodes are the same, i.e.,  $\sigma_R^2 = \sigma_D^2 = 1$ . In all the simulations, the subcarriers are modulated in quadrature phase shift keying (QPSK) format for both training and data transmission phases. Moreover, the following simulation parameters are considered:

### A. General Channel Setup

It is assumed that the multi-path channels exhibit unit-variance Rayleigh fading characteristics. The multipath fading channels from relay-to-destination and source-to-relay,  $\mathbf{g}$  and  $\mathbf{h}$ , respectively, are assumed to consist of 6 taps, i.e.,  $L_g = L_h = 6$ .  $N = 64$  subcarriers are used in each OFDM symbol. The normalized CFOs,  $\phi^{[s-d]}$  and  $\phi^{[r-d]}$ , are uniformly drawn from  $[-0.4, 0.4]$  and  $[-0.2, 0.2]$ , respectively, and  $\sigma_{\Delta[s-d]}^2 = \sigma_{\Delta[r-d]}^2 = \sigma_{\Delta}^2$ . Without loss of generality, in the remainder of this section, we assume  $\alpha = 1$  by letting  $P_T^{[r]} = \bar{P}_Z$ . Moreover, it is assumed that  $P_T^{[s]} = P_T^{[r]} = P_T = \text{SNR}$ .

Fig. 5 depicts the MSE of PN estimation, when estimating the shortened phase vector  $\eta^{[s-d]}$  for different values of  $M$  (see Section II-D). For ease of comparison and to isolate the effect of CFO and channel estimation, it is assumed that the channel response,  $\mathbf{c}$ , and the CFO,  $\phi^{[s-d]}$ , are perfectly known. Fig. 5 shows that when the phase noise variance is large, the

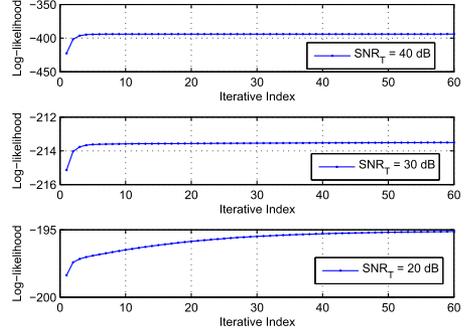


Fig. 6. The convergence of the proposed joint estimation algorithm at different SNR ( $\sigma_{\Delta[s-d]}^2 = 10^{-4} \text{rad}^2$ ).

choice of  $M$  makes a difference in estimation performance even when the SNR is small. Moreover, when the SNR is large, the choice of  $M$  can make a difference even when the phase noise variance is small. In fact, when the phase noise variance or SNR are large, the small eigenvalues still contribute greatly to the estimation accuracy of the phase noise parameters. Hence, even though many eigenvalues in matrix  $\Psi$  are close to zero, choosing a larger  $M$  can still improve the performance of phase noise estimation. However, the choice of  $M$  introduces a trade-off between estimation performance and complexity. From the plots in Fig. 5, it can be concluded that when the PN innovation variance  $\sigma_{\Delta[s-d]}^2$  is small, i.e.,  $\sigma_{\Delta[s-d]}^2 = 10^{-5} \text{rad}^2$ , PN parameters can be accurately estimated using  $M = 16$  compared to  $M = 64$ . Such an approach greatly reduces the PN estimation overhead. For scenarios with higher innovation variances, i.e.,  $\sigma_{\Delta[s-d]}^2 = 10^{-4} \text{rad}^2$  and  $\sigma_{\Delta[s-d]}^2 = 10^{-3} \text{rad}^2$ , it can be deduced that a larger value of  $M$  is needed to ensure accurate PN estimation, e.g.,  $M = 32$ . However, even for these larger PN variances, using the proposed scheme, the number of PN parameters that need to be tracked is reduced by one half. Accordingly, in the remainder of this section,  $M = 32$  is used.

In Fig. 6, the convergence of the proposed joint estimation algorithm is plotted for different SNRs. It can be observed that on average less than 50 iterations are needed for the proposed algorithm to converge to the true estimates for a wide range of SNR values. More importantly, the results in Fig. 6 show that as the SNR increases the proposed algorithm converges more quickly, e.g., for SNR = 30 dB less than 10 iterations are needed for the proposed estimator to converge.

Figs. 7 and 8 illustrate the estimation MSE of relay-to-destination channel,  $\mathbf{g}$  (defined in Remark 2), for PN variances,  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$  and  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ , respectively, while the estimation MSE of the source-to-relay channel,  $\mathbf{h}$  (defined in Remark 2), is presented in Figs. 9 and 10. As a comparison, the channel estimation performance while ignoring the effect of PN on the received signal is also presented in these figures. Finally, the proposed estimation algorithms performance is benchmarked using the derived HCRLB in Section IV. Figs. 7–10 indicate that by including the PN parameters in the joint estimation problem, channel estimation performance in relay networks can be significantly enhanced. At moderate SNR, Figs. 7–10 also show that the proposed algorithm has a constant performance gap with respect to the derived HCRLB

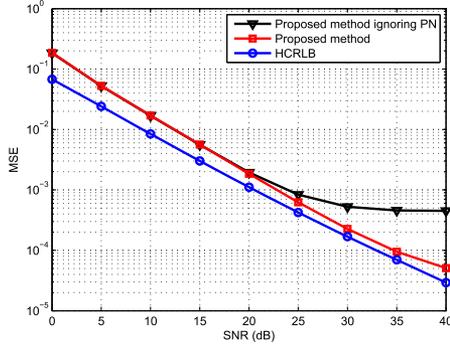


Fig. 7. The MSE of  $\underline{g}$  estimation at  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$ .

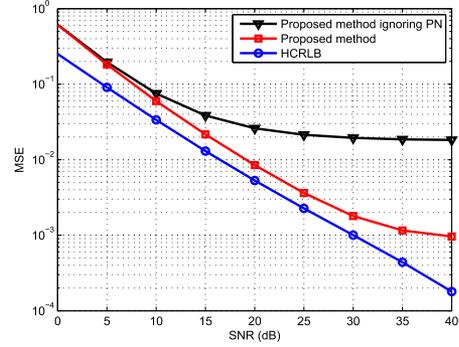


Fig. 10. The MSE of  $\underline{h}$  estimation at  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ .

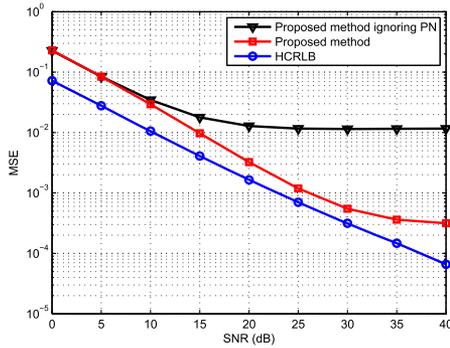


Fig. 8. The MSE of  $\underline{g}$  estimation at  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ .

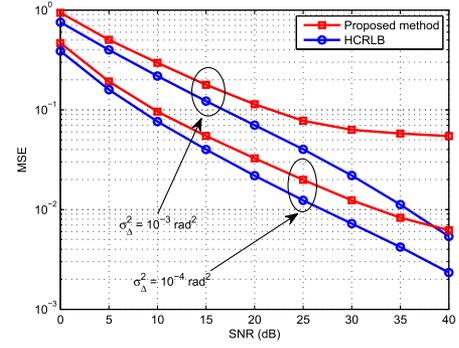


Fig. 11. The MSE of CFO plus PN estimation at  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$  and  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$ .

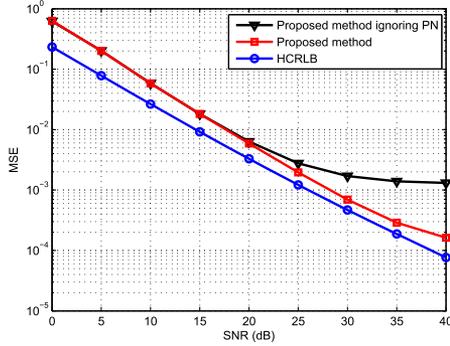


Fig. 9. The MSE of  $\underline{h}$  estimation at  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$ .

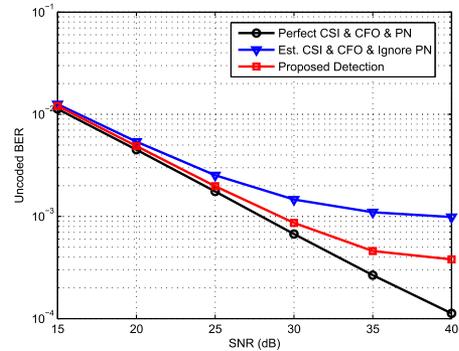


Fig. 12. The BER performance for the proposed joint data detection algorithm at  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$ .

bound for both PN innovation variances of  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$  and  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ . This is due to the inherent structure of the HCRLB, which is not necessarily a very tight bound as stated in [48]. Nevertheless, the performance of the proposed estimator is close to the derived HCRLB for moderate SNR. Finally, the results in Figs. 7–10 indicate that for large PN innovation variances, e.g.,  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ , the channel estimation performance suffers from an MSE error-floor at high SNR. This error-floor is caused by the time-varying PN parameters that cannot be perfectly estimated. We also observe that at low SNR, the performance of the proposed algorithm is close to the traditional estimation by ignoring the PN. This is because at low SNR, the overall estimation performance of the estimator is dominated by the additive noise at the destination node. This is different from the case at high SNR in which the algorithm’s estimation performance is dominated by the PN.

Fig. 11 illustrates the MSE for estimation of combined CFO and PN,  $\underline{\delta}$  for different PN variances. Similar to the results for channel estimation, the overall estimation performance suffers from an error floor for large PN variances, e.g.,  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ . This phenomenon can be similarly justified due to the imperfect estimation of PN parameters. Moreover, there is a 5 dB gap between the CFO and PN estimation MSE and the derived HCRLB at medium SNRs.

Figs. 12 and 13 illustrate the end-to-end BER of an uncoded OFDM and a coded OFDM<sup>4</sup> relay networks when applying the combination of the proposed iterative estimator and detector at  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$ . It is observed that significant performance

<sup>4</sup>Here we apply the convolutional coding with a rate of  $\frac{1}{2}$  with trellis structure given by (3, [6, 7]).

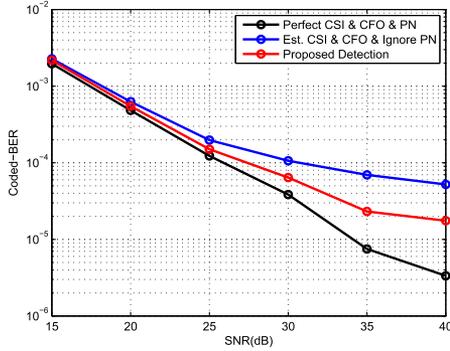


Fig. 13. The coded-BER performance for the proposed joint data detection algorithm at  $\sigma_{\Delta}^2 = 10^{-4} \text{rad}^2$ .

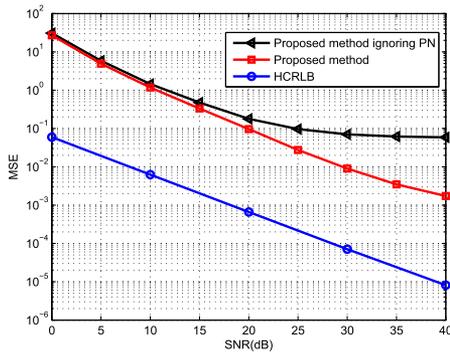


Fig. 14. The MSE of  $\mathbf{h}$  estimation at  $\sigma_{\Delta}^2 = 10^{-2} \text{rad}^2$ .

gains can be achieved by using the proposed joint data detection and PN estimation algorithm compared to a scheme that ignores the impact of PN. However, compared to the case with perfect channel, CFO, and PN, the proposed data detection algorithm still suffers from an error-floor at high SNR regime. This can be again attributed to imperfect PN estimation, where at high SNR, the overall BER of the OFDM relay system is dominated by PN and not the additive noise. This result indicates the importance of considering the impact of PN when determining the link budget, throughput, and coverage of wireless relay networks.

### B. 3GPP Channel Setup

Following the 3GPP-LTE system setting [49], the multipath channel  $\mathbf{g}$  and  $\mathbf{h}$  consist of  $L_g = L_h = 5$  channel taps and follow the exponentially decaying power-delay profile with a decay constant of 0.5136 [50]. In addition, we consider a 3GPP-LTE system with 256 subcarriers. The sampling frequency is  $f_s = 1.92$  MHz and the carrier frequency is  $f_c = 2$  GHz. The choice of the carrier frequency is to motivate the use of the 3GPP-LTE channel model. However, the proposed algorithm is applicable to higher carrier frequencies in the millimeter-wave band. The CP length of each OFDM symbol is set to 10 samples, i.e.,  $N_{CP} = 10$ . The unknown normalized CFO is assumed to be uniformly distributed over a range of  $[-0.5, 0.5]$ , which equivalently means that the CFO is uniformly distributed over a range of  $[-3.75, 3.75]$  MHz. Moreover, it is assumed that  $P_T^{[s]} = \text{SNR}$  and  $P_T^{[r]} = 2 \times \text{SNR}$ .

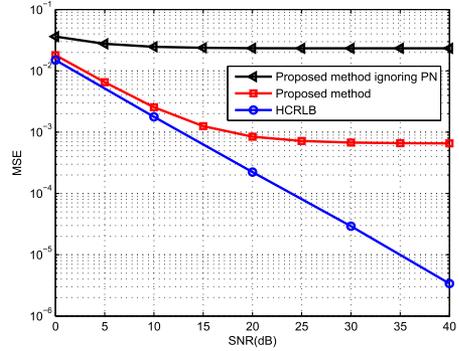


Fig. 15. The MSE of  $\mathbf{g}$  estimation at  $\sigma_{\Delta}^2 = 10^{-2} \text{rad}^2$ .

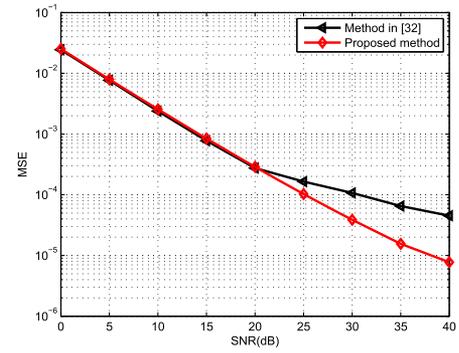


Fig. 16. The performance comparison between the proposed estimation and [32] at  $\sigma_{\Delta}^2 = 10^{-3} \text{rad}^2$ .

Figs. 14 and 15 illustrate the estimation MSE of the source-to-relay channel  $\mathbf{h}$  and the estimation MSE of the relay-to-destination channel  $\mathbf{g}$ , respectively, with PN variances of  $\sigma_{\Delta}^2 [s-d] = \sigma_{\Delta}^2 [r-d] = \sigma_{\Delta}^2 = 10^{-2} \text{rad}^2$ .<sup>5</sup> We observe that for the 3GPP-LTE setting with higher PN variances, the performance gain of the proposed estimation over the one ignoring the phase noise is more significant. However, we see from Figs. 14 and 15 that the performance gap between the derived HCRLB and the obtained MSE becomes larger with a higher phase noise variance. In Fig. 16, we compare the proposed estimation with the one in [32]. It is worth noting that [32] aims to estimate the cascaded channel parameters for the source-to-destination link, in which the channel estimation is similar to that of point-to-point systems. Hence, for fairness, we compare the estimation MSE of the relay-to-destination link by using our proposed estimator and the that of [32]. We observe from Fig. 16 that for low and moderate SNRs, the MSE performance of the proposed approach and that of [32] are similar. However, our proposed estimator outperforms [32] at high SNR values, which maybe of practical interest for backhaul microwave and millimeter-wave applications with stringent performance requirements.

<sup>5</sup>In practice, this is a very high PN variance. As shown in [51], [52], the phase noise innovation variance is small, e.g., using the measurement results in [52], [Fig. 16], and [52, Eq. (10)] for a free-running oscillator operating at 2.8 GHz with  $T_S = 10^{-6}$  the phase noise rate is calculated to be  $\sigma_{\Delta}^2 = 10^{-4}$ .

## VII. CONCLUSIONS

In this paper, joint channel, CFO, and PN estimation and data detection in OFDM relay networks is analyzed. Due to its time-varying nature, new algorithms for tracking the PN parameters in both the training and data transmission intervals are proposed. During the training interval, a new joint CFO, channel, and PN estimation algorithm that iteratively estimates these impairments is derived. To reduce estimation overhead, the proposed algorithm applies the correlation amongst the PN parameters to reduce the dimensionality of the estimation problem. Simulations show that the proposed estimator significantly enhances channel estimation performance in presence of PN, converges quickly, and performs close to the derived HCRLB at medium SNRs. Moreover, an iterative joint PN estimation and data detection receiver based on the MAP criterion at the destination node is proposed. The combination of the proposed estimation and data detection algorithms is shown to result in 5–10 dB performance gains over schemes that ignore the deteriorating effect of PN. Future research interests include extending the proposed estimation and detection scheme to OFDMA networks and single-hop multi-relay scenarios. Considering that in an OFDMA network, each user can only use a portion of the available subcarriers. To obtain an accurate estimates of phase noise, CFO, and channel parameters, multiple OFDM training symbols may be required. For the single-hop multi-relay scenario, as each relay node has an independent oscillator, more phase noise and CFO errors may be introduced to the network, which makes the estimation and detection process more challenging. For full-duplex relaying networks, a more detailed analysis of the impact of self interference on the performance of the estimator and the overall system is needed. Moreover, enhancing phase noise estimation and deriving a tight HCRLB for very strong phase noise variances should also be subject of future research.

APPENDIX A  
DERIVATION OF (28)

In this section, an expression for the optimization in (27) is derived. It is straightforward to determine that the optimization in (27) is a nonlinear and non-convex problem. Thus, the solution of  $\eta^{[s-d]}$  in (27) should be in general obtained through exhaustive search. To simplify the problem and obtain a closed-form solution, we first approximate the covariance matrix  $\Sigma^{[r]}$  as  $[\hat{\Sigma}^{[r]}]^{[k]} = \alpha^2 \sigma_R^2 \hat{\Lambda}_{\theta^{[s-d]}}^{[k]} \hat{\Lambda}_{\phi^{[s-d]}}^{[k]} \hat{\mathbf{G}}^{[k]} \hat{\mathbf{G}}^{[k]H} \hat{\Lambda}_{\phi^{[s-d]}}^{[k]H} \hat{\Lambda}_{\theta^{[s-d]}}^{[k]H} + \sigma_D^2 \mathbf{I}_N$ , where  $[\hat{\Lambda}_{\theta^{[s-d]}}^{[k]}]_{m,m} = e^{j[\hat{\theta}^{[s-d]}(m)]^{[k]}}$  is obtained from the previous iteration. Moreover, since the PN innovation variance of practical oscillators is usually small, the elements in  $\Lambda_{\theta^{[s-d]}}$  can be approximated by a Taylor series expansion as  $e^{j\theta^{[s-d]}(n)} \approx 1 + j\theta^{[s-d]}(n)$ . This small angle approximation has also been used in [18], [25], [39] for PN estimation. Hence, the PN matrix,  $\Lambda_{\theta^{[s-d]}}$ , can be approximated as  $\Lambda_{\theta^{[s-d]}} \approx \mathbf{I}_N + j\text{Diag}(\theta^{[s-d]})$  and  $\mathcal{L}_{\eta^{[s-d]}}$  in (27) can be rewritten as

$$\begin{aligned} \mathcal{L}_{\eta^{[s-d]}} &\approx (\bar{\mathbf{y}}^{[s]} - \mathbf{B}\eta^{[s-d]})^H \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} (\bar{\mathbf{y}}^{[s]} - \mathbf{B}\eta^{[s-d]}) \\ &\quad + \frac{1}{2} \eta^{[s-d]T} \eta^{[s-d]} \end{aligned}$$

$$\begin{aligned} &\approx \bar{\mathbf{y}}^{[s]H} \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \bar{\mathbf{y}}^{[s]} - 2\Re \left( \bar{\mathbf{y}}^{[s]H} \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \mathbf{B} \right) \\ &\quad \times \eta^{[s-d]} + \eta^{[s-d]T} \Re \left( \mathbf{B}^H \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \mathbf{B} \right) \eta^{[s-d]} \\ &\quad + \frac{1}{2} \eta^{[s-d]T} \eta^{[s-d]}, \end{aligned} \quad (\text{A.1})$$

where  $\bar{\mathbf{y}}^{[s]} \triangleq \mathbf{y}^{[s]} - \alpha \hat{\Lambda}_{\phi^{[s-d]}}^{[k]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{c}}^{[k]}$  and  $\mathbf{B} \triangleq j\text{Diag}(\alpha \hat{\Lambda}_{\phi^{[s-d]}}^{[k]} \mathbf{F}^H \Lambda_{s[s]} \mathbf{F}_{[L]} \hat{\mathbf{c}}^{[k]}) \Pi^{[s-d]}$ . Next, by equating the gradient of (A.1) to zero, i.e.,

$$\begin{aligned} \frac{\partial \mathcal{L}_{\eta^{[s-d]}}}{\partial \eta^{[s-d]}} &= -2\Re \left( \mathbf{B}^H \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \bar{\mathbf{y}}^{[s]} \right) \\ &\quad + 2\Re \left( \mathbf{B}^H \left[ [\hat{\Sigma}^{[r]}]^{[k]} \right]^{-1} \mathbf{B} \right) \eta^{[s-d]} + \eta^{[s-d]} \\ &= \mathbf{0}_{M \times 1}, \end{aligned}$$

then we obtain (28).

APPENDIX B  
COMPLEXITY ANALYSIS OF ALGORITHM 1

In (50),  $C_M^{(n)}$  and  $C_A^{(n)}$  are determined as

$$\begin{aligned} C_M^{(28)} &= N^3 + 2N^2M + M^3 + NM + M^2 \\ &\quad + \underbrace{MN + N}_{\text{update } [\hat{\theta}^{[s-d]}]^{[k+1]} \text{ and } [\hat{\Sigma}^{[r]}]^{[k]}}, \\ C_M^{(32)} &= \underbrace{NLL_g + 3NL_g + 2N^2L_g + N + (2N)^3}_{\text{update C in (32)}} \\ &\quad + N^2L_g + NL_g^2 + L_g^3 + 2NL_g, \\ C_M^{(34)} &= \underbrace{NLL_h + 2NL_h + N^2L_g + N + NL_h}_{\text{update D in (34)}} \\ &\quad + N^3 + N^2 + N, \\ C_M^{(38)} &= \underbrace{N + 3NL + N^2L}_{\text{update } \mathbf{d}^{[s-d]} \text{ in (38)}} + N^3 + N^2 + 3N, \\ C_A^{(28)} &= M(N-1) + 2M(M-1) + M \\ &\quad + NM(N-1) + NM(N-1), \\ C_A^{(32)} &= 4NL_g(2N-1) + 2N(2N-1) + L_g(2N-1), \\ C_A^{(34)} &= 2NL_h(N-1) + N(N-1) + L_h(N-1), \\ C_A^{(38)} &= 1 + 2N(N-1)2(N-1) + N. \end{aligned} \quad (\text{B.1})$$

Now we determine  $C_M^I$  and  $C_A^I$ . Note that the initialization of Algorithm 1 requires us to compute  $\hat{\phi}^{[r-d]}$ ,  $\hat{\theta}^{[r-d]}$ ,  $\hat{\mathbf{g}}^{[0]}$ ,  $\hat{\mathbf{h}}^{[0]}$  and  $[\hat{\phi}^{[s-d]}]^{[0]}$  in (20), (21), (39), (40) and (41), respectively. Hence,  $C_M^{I(n)}$  and  $C_A^{I(n)}$  can be determined as

$$\begin{aligned} C_M^{I(20)} &= N_e(3N^2 + N), \\ C_A^{I(20)} &= N_e \left( N + N^2(N-1) + 2N(N-1) + 2(N-1) \right), \end{aligned} \quad (\text{B.2})$$

where  $N_e$  is the number determining the precision in exhaustively searching  $\phi^{[r-d]}$  in (20), and

$$\begin{aligned} C_M^{I(21)} &= 4N^3 + 2N^2 + M^3, \\ C_A^{I(21)} &= N^2(N-1) + 2NM(N-1) + M + 2N(N-1) \\ &\quad + M(N-1) + M(M-1) + N(M-1) \\ &\quad + L_g(N-1), \\ C_M^{I(39)} &= 3N^2 + N^3, \\ C_A^{I(39)} &= N(N-1) + L_g(N-1), \\ C_M^{I(40)} &= 6N^3 + 2N^2, \\ C_A^{I(40)} &= L(L_h-1) + 2N(L-1) + N + N(N-1) \\ &\quad + (N-1), \end{aligned}$$

and

$$\begin{aligned} C_M^{I(41)} &= N_e(2N^3 + 4N^2 + N), \\ C_A^{I(41)} &= N_e(L(L_h-1) + 2N(L-1) + N + N(N-1) \\ &\quad + (N-1)), \end{aligned}$$

where  $N_e$  is the number determining the precision in exhaustively searching as defined in (B.2).

#### APPENDIX C DEFINITIONS OF $\rho_i$ AND $\mathbf{W}_i$

For  $\rho_i$ , we have

$$\rho_i = \begin{cases} \left[ (\alpha \mathbf{\Lambda} \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L]}\mathbf{c})^T, \mathbf{0}_{N \times 1}^T \right]^T, & i = 1 \\ \left[ \mathbf{a}_{i-2}^T \text{Diag} \left( \alpha \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L]}\mathbf{c} \right), \mathbf{0}_{N \times 1}^T \right]^T, & i = 2, 3, \dots, N+1 \\ \left[ \mathbf{0}_{N \times 1}^T, (\mathbf{\Lambda} \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[r]}} \mathbf{F}_{[L_g]}\mathbf{g})^T \right]^T, & i = N+2 \\ \left[ \mathbf{0}_{N \times 1}^T, \mathbf{b}_{i-2N-3}^T \text{Diag} \left( \mathbf{\Lambda}_{\phi^{[r-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[r]}} \mathbf{F}_{[L_g]}\mathbf{g} \right) \right]^T, & i = N+3, N+4, \dots, 2N+2 \\ \mathbf{E}(:, 1), & i = 2N+3 \\ \mathbf{E}(:, i-2N-2), & i = 2N+4, 2N+5, \dots, 2N+L_h+2 \\ j\mathbf{E}(:, i-2N-L_h-1), & i = 2N+L_h+3, \dots, 2N+2L_h+1 \\ \mathbf{K}(:, 1), & i = 2N+2L_h+2 \end{cases} \quad (C.1)$$

and

$$\rho_i = \begin{cases} \mathbf{K}(:, i-2N-2L_h-1), \\ i = 2N+2L_h+3, \dots, 2N+2L_h+L_g+1, \\ j\mathbf{K}(:, i-2N-2L_h-2L_g), \\ i = 2N+2L_h+L_g+2, \dots, Q \end{cases} \quad (C.2)$$

where  $\mathbf{\Lambda}$  is a diagonal matrix with  $[\mathbf{\Lambda}]_{m,m} = \frac{j2\pi(m-1)}{N}$ ,  $\mathbf{a}_m = [\mathbf{0}_{1,m-1}, j \exp(j\theta^{[s-d]}(m)), \mathbf{0}_{1,N-m}]^T$ ,  $\mathbf{b}_m = [\mathbf{0}_{1 \times (m-1)}, j \exp(j\theta^{[r-d]}(m)), \mathbf{0}_{1 \times (N-m)}]^T$ ,  $\mathbf{E} = \begin{bmatrix} \alpha \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L]}\mathbf{c} \\ \mathbf{\Lambda}_{\phi^{[r-d]}} \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[r]}} \mathbf{F}_{[L_g]}\mathbf{g} \end{bmatrix} \tilde{\mathbf{H}}$  with  $\tilde{\mathbf{H}}$  being constructed via  $\mathbf{h}$  as in (30),  $\mathbf{K} \triangleq [(\alpha \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L]}\mathbf{c})^T, \mathbf{0}_{N \times 1}^T]^T$  with  $\tilde{\mathbf{G}}$  being constructed using  $\mathbf{g}$  as in (29).

For  $\mathbf{W}_i$ , we have

$$\mathbf{W}_i = \begin{cases} \left( \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{G} \mathbf{G}^H \mathbf{\Lambda}_{\theta^{[s-d]}}^H \right) \odot \left( \mathbf{\Lambda} \boldsymbol{\vartheta} \boldsymbol{\vartheta}^H + \boldsymbol{\vartheta} \boldsymbol{\vartheta}^H \mathbf{\Lambda}^H \right), & i = 1 \\ \left( \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{G} \mathbf{G}^H \mathbf{\Lambda}_{\theta^{[s-d]}}^H \right) \odot \left( \mathbf{a}_{i-2} \boldsymbol{\theta}^{[s-d]H} + \boldsymbol{\theta}^{[s-d]} \mathbf{a}_{i-2}^H \right), & i = 2, \dots, N+1 \\ \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{D}_0 \mathbf{G}^H + \mathbf{G} \mathbf{D}_0^H \right) \mathbf{\Lambda}_{\phi^{[s-d]}}^H \mathbf{\Lambda}_{\theta^{[s-d]}}^H, & i = 2N+3 \\ \alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{D}_{i-2N-3} \mathbf{G}^H + \mathbf{G} \mathbf{D}_{i-2N-3}^H \right) \\ \times \mathbf{\Lambda}_{\phi^{[s-d]}}^H \mathbf{\Lambda}_{\theta^{[s-d]}}^H, & i = 2N+4, \dots, 2N+L_g+2 \\ j\alpha^2 \sigma_R^2 \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{D}_{i-L_g-2N-2} \mathbf{G}^H - \mathbf{G} \mathbf{D}_{i-L_g-2N-2}^H \right) \\ \times \mathbf{\Lambda}_{\phi^{[s-d]}}^H \mathbf{\Lambda}_{\theta^{[s-d]}}^H, & i = 2N+L_g+3 \dots 2N+2L_g+1 \end{cases} \quad (C.3)$$

where  $\boldsymbol{\vartheta} \triangleq \left[ 1, e^{\frac{j2\pi\phi^{[s-d]}}{N}}, \dots, e^{\frac{j2\pi(N-1)\phi^{[s-d]}}{N}} \right]$  and  $\mathbf{D}_m \triangleq [\mathbf{0}_{N \times (L_g-m-1)}, \mathbf{I}_N, \mathbf{0}_{N \times m}]$ ,  $\forall m$ .

#### APPENDIX D DERIVATION OF FIM

In this section, the FIM for joint estimation of channels, CFO, and PN parameters, i.e.,  $\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})$ , is derived. First, note that the combined received signal vector at the destination node in (52),  $\mathbf{y} \triangleq [\mathbf{y}^{[s]T}, \mathbf{y}^{[r]T}]^T$  is a multivariate Gaussian random variable, i.e.,  $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  with mean  $\boldsymbol{\mu} = [(\alpha \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L]}\mathbf{c})^T, (\mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[r]}} \mathbf{F}_{[L_g]}\mathbf{g})^T]^T$  and covariance  $\boldsymbol{\Sigma} = \text{Blkdiag}(\boldsymbol{\Sigma}^{[r]}, \sigma_D^2 \mathbf{I}_N)$ . As a result, the  $(i, j)$ -th element of  $\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})$  can be determined as [44]

$$\begin{aligned} [\mathbf{FIM}(\mathbf{y}; \boldsymbol{\lambda})]_{i,j} &= 2\text{Re} \left[ \frac{\partial \boldsymbol{\mu}^H}{\partial \lambda_i} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \lambda_j} \right] \\ &\quad + \text{Tr} \left[ \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \lambda_i} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \lambda_j} \right]. \end{aligned} \quad (D.1)$$

To obtain (D.1), the following derivatives are evaluated as

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}}{\partial \phi^{[s-d]}} &= [(\alpha \mathbf{\Lambda} \mathbf{\Lambda}_{\theta^{[s-d]}} \mathbf{\Lambda}_{\phi^{[s-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L]}\mathbf{c})^T, \mathbf{0}_{N \times 1}^T]^T, \\ \frac{\partial \boldsymbol{\mu}}{\partial \phi^{[r-d]}} &= [\mathbf{0}_{N \times 1}^T, (\mathbf{\Lambda} \mathbf{\Lambda}_{\theta^{[r-d]}} \mathbf{\Lambda}_{\phi^{[r-d]}} \mathbf{F}^H \underline{\mathbf{\Lambda}}_{s^{[r]}} \mathbf{F}_{[L_g]}\mathbf{g})^T]^T, \end{aligned}$$

$$\begin{aligned}\frac{\partial \underline{\boldsymbol{\mu}}}{\partial \theta^{[s-d]}(m)} &= \left[ \mathbf{a}_m^T \text{Diag} \left( \alpha \boldsymbol{\Lambda}_{\phi^{[s-d]}} \mathbf{F}^H \underline{\boldsymbol{\Lambda}}_{s^{[s]}} \mathbf{F}_{[L] \underline{\mathbf{c}}} \right), \mathbf{0}_{N \times 1}^T \right]^T, \\ \frac{\partial \underline{\boldsymbol{\mu}}}{\partial \theta^{[r-d]}(m)} &= \left[ \mathbf{0}_{N \times 1}^T, \mathbf{b}_m^T \text{Diag} \left( \boldsymbol{\Lambda}_{\phi^{[r-d]}} \mathbf{F}^H \underline{\boldsymbol{\Lambda}}_{s^{[r]}} \mathbf{F}_{[L_g] \underline{\mathbf{g}}} \right) \right]^T,\end{aligned}\quad (\text{D.2})$$

where  $\mathbf{a}_m$  and  $\mathbf{b}_m$  are defined below (55). Moreover, for channel responses  $\underline{\mathbf{g}}$  and  $\underline{\mathbf{h}}$ , for  $m = 1, 2, \dots, L_g$ , we have  $\frac{\partial \underline{\boldsymbol{\mu}}}{\partial (g(0))} = \mathbf{E}(:, 1)$  and

$$\frac{\partial \underline{\boldsymbol{\mu}}}{\partial \Re(g(m))} = \mathbf{E}(:, m), \quad \frac{\partial \underline{\boldsymbol{\mu}}}{\partial \Im(g(m))} = j\mathbf{E}(:, m), \quad (\text{D.3})$$

and for  $m = 1, 2, \dots, L_h$ , we have  $\frac{\partial \underline{\boldsymbol{\mu}}}{\partial (h(0))} = \mathbf{K}(:, 1)$  and

$$\frac{\partial \underline{\boldsymbol{\mu}}}{\partial \Re(h(m))} = \mathbf{K}(:, m), \quad \frac{\partial \underline{\boldsymbol{\mu}}}{\partial \Im(h(m))} = j\mathbf{K}(:, m), \quad (\text{D.4})$$

where  $\mathbf{E}$  and  $\mathbf{K}$  are defined as in (C.1) and (C.2), respectively. Since  $\phi^{[r-d]}$ ,  $\theta^{[r-d]}$ ,  $\underline{\mathbf{h}}$  are irrelevant to the noise covariance matrix  $\boldsymbol{\Sigma}$ , it is straightforward to determine that  $\frac{\partial \boldsymbol{\Sigma}}{\partial \phi^{[r-d]}} = \frac{\partial \boldsymbol{\Sigma}}{\partial \theta^{[r-d]}(m)} = \frac{\partial \boldsymbol{\Sigma}}{\partial h(0)} = \frac{\partial \boldsymbol{\Sigma}}{\partial \Re(h(m))} = \frac{\partial \boldsymbol{\Sigma}}{\partial \Im(h(m))} = \mathbf{0}$ ,  $\forall m$ . Moreover, for the CFO and PN parameters,  $\phi^{[s-d]}$  and  $\theta^{[s-d]}$ , we can obtain that  $\frac{\partial \boldsymbol{\Sigma}}{\partial \phi^{[s-d]}} = \text{Blkdiag} \left( \frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial \phi^{[s-d]}} \right)$ , where  $\frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial \phi^{[s-d]}} = \left( \alpha^2 \sigma_R^2 \boldsymbol{\Lambda}_{\theta^{[s-d]}} \underline{\mathbf{G}} \underline{\mathbf{G}}^H \boldsymbol{\Lambda}_{\theta^{[s-d]}}^H \right) \odot \left( \boldsymbol{\Lambda} \boldsymbol{\vartheta} \boldsymbol{\vartheta}^H + \boldsymbol{\vartheta} \boldsymbol{\vartheta}^H \boldsymbol{\Lambda}^H \right)$ , and

$$\begin{aligned}\frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial \theta^{[s-d]}(m)} &= \left( \alpha^2 \sigma_R^2 \boldsymbol{\Lambda}_{\theta^{[s-d]}} \underline{\mathbf{G}} \underline{\mathbf{G}}^H \boldsymbol{\Lambda}_{\theta^{[s-d]}}^H \right) \\ &\odot \left( \mathbf{a}_m \boldsymbol{\theta}^{[s-d]H} + \boldsymbol{\theta}^{[s-d]} \mathbf{a}_m^H \right), \quad \forall m\end{aligned}\quad (\text{D.5})$$

with  $\boldsymbol{\vartheta} \triangleq \left[ 1, e^{\frac{j2\pi\phi^{[s-d]}}{N}}, \dots, e^{\frac{j2\pi(N-1)\phi^{[s-d]}}{N}} \right]$ . For channel response  $\underline{\mathbf{g}}$ , based on the structure of  $\underline{\mathbf{G}}$  as shown in (5), we have

$$\begin{aligned}\frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial g(0)} &= \alpha^2 \sigma_R^2 \boldsymbol{\Lambda}_{\theta^{[s-d]}} \boldsymbol{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{D}_0 \underline{\mathbf{G}}^H + \underline{\mathbf{G}} \mathbf{D}_0^H \right) \\ &\quad \times \boldsymbol{\Lambda}_{\phi^{[s-d]}}^H \boldsymbol{\Lambda}_{\theta^{[s-d]}}^H, \\ \frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial \Re(g(m))} &= \alpha^2 \sigma_R^2 \boldsymbol{\Lambda}_{\theta^{[s-d]}} \boldsymbol{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{D}_m \underline{\mathbf{G}}^H + \underline{\mathbf{G}} \mathbf{D}_m^H \right) \\ &\quad \times \boldsymbol{\Lambda}_{\phi^{[s-d]}}^H \boldsymbol{\Lambda}_{\theta^{[s-d]}}^H, \quad m = 1, 2, \dots, L_g - 1, \\ \frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial \Im(g(m))} &= j\alpha^2 \sigma_R^2 \boldsymbol{\Lambda}_{\theta^{[s-d]}} \boldsymbol{\Lambda}_{\phi^{[s-d]}} \left( \mathbf{D}_m \underline{\mathbf{G}}^H - \underline{\mathbf{G}} \mathbf{D}_m^H \right) \\ &\quad \times \boldsymbol{\Lambda}_{\phi^{[s-d]}}^H \boldsymbol{\Lambda}_{\theta^{[s-d]}}^H, \quad m = 1, 2, \dots, L_g - 1,\end{aligned}\quad (\text{D.6})$$

where  $\mathbf{D}_m \triangleq \left[ \mathbf{0}_{N \times (L_g - m - 1)}, \mathbf{I}_N, \mathbf{0}_{N \times m} \right]$ . Subsequently, the derivatives of the covariance matrix with respect to the relay and imaginary parts of the relay-to-destination channel parameters are given by  $\frac{\partial \boldsymbol{\Sigma}}{\partial g(0)} = \text{Blkdiag} \left( \frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial g(0)}, \mathbf{0}_{N \times N} \right)$  and  $\frac{\partial \boldsymbol{\Sigma}}{\partial \Im(g(m))} = \text{Blkdiag} \left( \frac{\partial \boldsymbol{\Sigma}^{[r]}}{\partial \Im(g(m))}, \mathbf{0}_{N \times N} \right)$ , respectively. By combining (D.2)–(D.6) together, the results in Theorem 1 are derived.

## REFERENCES

- [1] R. Wang, H. Mehrpouyan, M. Tao, and Y. Hua, "Channel estimation and carrier recovery in the presence of phase noise in OFDM relay systems," in *Proc. IEEE Globecom*, Austin, TX, USA, Dec. 8–12, 2015.
- [2] Y. Hua, D. W. Bliss, S. Gazor, Y. Rong, and Y. Sung, Eds., "Theories and methods for advanced wireless relays: Issue I," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, Sep. 2012.
- [3] S. Yiu, R. Schober, and L. Lampe, "Distributed space time block coding," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1195–1206, Jul. 2006.
- [4] Y. Hua, Y. Mei, and Y. Chang, "Wireless antennas—Making wireless communications perform like wireline communications," in *Proc. IEEE Top. Conf. Wireless Commun. Technol.*, Honolulu, HI, USA, Oct. 15–17, 2003, pp. 47–73.
- [5] R. Wang and M. Tao, "Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1352–1365, Mar. 2012.
- [6] S. Xu and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way MIMO relay system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1645–1655, May 2011.
- [7] Y. Ma, A. Liu, and Y. Hua, "A dual-phase power allocation scheme for multicarrier relay system with direct link," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 5–16, Jan. 2014.
- [8] 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.
- [9] Y. Jing and X. Yu, "ML-based channel estimations for non-regenerative relay networks with multiple transmit and receive antennas," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1428–1439, Sep. 2012.
- [10] P. Lioliou, M. Viberg, and M. Matthaiou, "Bayesian approach to channel estimation for AF MIMO relaying systems," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1440–1451, Sep. 2012.
- [11] J. Ma, P. Orlik, J. Zhang, and G. Li, "Pilot matrix design for estimating cascaded channels in two-hop MIMO amplify-and-forward relay systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1956–1965, Jun. 2011.
- [12] Y. Rong, M. Khandaker, and Y. Xiang, "Channel estimation of dual-hop MIMO relay system via parallel factor analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2224–2233, Jun. 2012.
- [13] X. Xie, M. Peng, B. Zhao, W. Wang, and Y. Hua, "Maximum a posteriori based channel estimation strategy for two-way relaying channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 450–463, Jan. 2014.
- [14] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4438–4446, Sep. 2011.
- [15] M. Dohler, R. W. Heath, A. Lozano, C. B. Papadias, and R. A. Valenzuela, "Is the PHY layer dead?," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 159–165, Apr. 2011.
- [16] S. Bay, C. Herzet, J.-M. Brossier, J.-P. Barbot, and B. Geller, "Analytical and asymptotic analysis of Bayesian Cramér-Rao bound for dynamical phase offset estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 61–70, Jan. 2008.
- [17] N. Noels, H. Steendam, M. Moeneclaey, and H. Bruneel, "Carrier phase and frequency estimation for pilot-symbol assisted transmission: Bounds and algorithms," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4578–4587, Dec. 2005.
- [18] 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.
- [19] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1862–1874, May 2007.
- [20] T. C. W. Schenck, X.-J. Tao, P. F. M. Smulders, and E. R. Fledderus, "On the influence of phase noise induced ICI in MIMO OFDM systems," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 682–684, Aug. 2005.
- [21] V. Syrjala, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 2977–2990, Jun. 2014.
- [22] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancellation in wireless full-duplex," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4494–4510, Nov. 2013.
- [23] H. Mehrpouyan and S. Blotstein, "Bounds and algorithms for multiple frequency offset estimation in cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1300–1311, Apr. 2011.

- [24] O. Salim, A. Nasir, H. Mehrpouyan, W. Xiang, S. Durrani, and R. Kennedy, "Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid Cramer-Rao lower bound," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3311–3325, Sep. 2014.
- [25] 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.
- [26] K. J. Kim, R. Iltis, and H. Poor, "Frequency offset and channel estimation in cooperative relay networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3142–3155, Sep. 2011.
- [27] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 191–193, Feb. 1995.
- [28] L. Tomba, "On the effect of Wiener phase noise in OFDM systems," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 580–583, May 1998.
- [29] 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.
- [30] L. Thiagarajan, S. Sun, and T. Quek, "Joint carrier frequency offset and channel estimation in OFDM based non-regenerative wireless relay networks," in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP'09)*, Taipei, Taiwan, Apr. 19–24, 2009, pp. 2569–2572.
- [31] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "On the performance of OFDM-based amplify-and-forward relay networks in the presence of phase noise," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1458–1466, May 2011.
- [32] O. Salim, A. Nasir, W. Xiang, and R. Kennedy, "Joint channel, phase noise, and carrier frequency offset estimation in cooperative OFDM systems," in *Proc. IEEE Int. Conf. Commun. (ICC'14)*, Sydney, NSW, Australia, Jun. 10–14, 2014, pp. 4384–4389.
- [33] J. Wells, *Multi-Gigabit Microwave and Millimeter-Wave Wireless Communications*, 1st ed. Norwood, MA, USA: Artech House, 2010.
- [34] K.-C. Huang and D. J. Edwards, *Millimetre Wave Antennas for Gigabit Wireless Communications: A Practical Guide to Design and Analysis in a System Context*. Hoboken, NJ, USA: Wiley, 2008.
- [35] Y. Lee, J. Ha, and J. Choi, "Design of a wideband indoor repeater antenna with high isolation for 3G systems," *IEEE Antennas Wireless Propag. Lett.*, vol. 9, pp. 697–700, Jul. 2010.
- [36] Mini-Circuits. *Wideband Microwave Amplifier AVA-24+* [Online]. Available: <http://www.minicircuits.com/pdfs/AVA-24+.pdf>
- [37] K. Salehian, M. Guillet, B. Caron, and A. Kennedy, "On-channel repeater for digital television broadcasting service," *IEEE Trans. Broadcast.*, vol. 48, no. 2, pp. 97–102, Jun. 2002.
- [38] Mini-Circuits. *Ultra Low Noise Voltage Controlled Oscillator ROS-209-319+* [Online]. Available: <http://www.minicircuits.com/pdfs/ROS-209-319+.pdf>
- [39] J. Tao, J. Wu, and C. Xiao, "Estimation of channel transfer function and carrier frequency offset for OFDM systems with phase noise," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4380–4387, Oct. 2009.
- [40] 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.
- [41] A. Georgiadis, H. Rogier, and L. Roselli, Ed., *Microwave and Millimeter Wave Circuits and Systems: Emerging Design, Technologies and Applications*. Hoboken, NJ, USA: Wiley, 2012.
- [42] A. Chorti and M. Brookes, "A spectral model for RF oscillators with power-law phase noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 9, pp. 1989–1999, Sep. 2006.
- [43] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I, Fund. Theory Appl.*, vol. 47, no. 5, pp. 655–674, May 2000.
- [44] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
- [45] A. Nasir, H. Mehrpouyan, R. Schober, and Y. Hua, "Phase noise in MIMO systems: Bayesian Cramer-Rao bounds and soft-input estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2675–2692, May 2013.
- [46] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*. Hoboken, NJ, USA: Wiley, 2001.
- [47] S. Sezginer, P. Bianchi, and W. Hachem, "Asymptotic Cramer-Rao bounds and training design for uplink MIMO-OFDMA systems with frequency offsets," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3606–3622, Jul. 2007.
- [48] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Hoboken, NJ, USA: Wiley, 2007.
- [49] R. Love, R. Kuchibhotla, A. Ghosh, R. Ratasuk, B. Classon, and Y. Blankenship, "Downlink control channel design for 3GPP LTE," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC'08)*, Las Vegas, NV, USA, Mar. 31/Apr. 3, 2008, pp. 813–818.
- [50] D. Morgan, "Analysis and realization of an exponentially-decaying impulse response model for frequency-selective fading channels," *IEEE Signal Process. Lett.*, vol. 15, pp. 441–444, May 2008.
- [51] J. A. McNeill, "Jitter in ring oscillators," *IEEE J. Solid-State Circuits*, vol. 32, no. 6, pp. 870–879, Jun. 1997.
- [52] A. Hajimiri, S. Limotyrakis, and T. H. Lee, "Jitter and phase noise in ring oscillators," *IEEE J. Solid-State Circuits*, vol. 34, no. 6, pp. 790–804, Jun. 1999.
- [53] A. A. Nasir, H. Mehrpouyan, S. D. Blostein, S. Durani, and R. A. Kenedy, "Timing and carrier synchronization with channel estimation in multi-relay cooperative networks," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 793–811, Mar. 2012.
- [54] M. Martalo, C. Tripodi, and R. Raheli, "On the information rate of phase noise-limited communications," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2015, pp. 1–7.
- [55] H. Ghozlan and G. Kramer, "Models and information rates for Wiener phase noise channels," <http://arxiv.org/abs/1503.03130>, Mar. 2015.



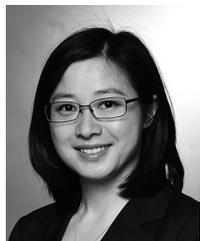
**Rui Wang** (M'14) received the B.S. degree from Anhui Normal University, Wuhu, China, in 2006, and the M.S. degree from Shanghai University, Shanghai, China, in 2009, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2013, all in electronic engineering. From August 2012 to February 2013, he was a Visiting Ph.D. Student at the Department of Electrical Engineering, University of California, Riverside, CA, USA. From October 2013 to October 2014, he was with the Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong, as a Postdoctoral Research Associate. Since October 2014, he has been with the College of Electronics and Information Engineering, Tongji University, Shanghai, China, as an Assistant Professor. His research interests include wireless cooperative communications, MIMO technique, network coding, and OFDM.



**Hani Mehrpouyan** (S'05–M'10) received the B.Sc. degree (Hons.) from Simon Fraser University, Burnaby, Canada, in 2004, and the Ph.D. degree in electrical engineering from Queens University, Kingston, ON, Canada, in 2010.

From September 2010 to March 2012, he was a Postdoc with Chalmers University of Technology, Gothenburg, Sweden, where he led the MIMO aspects of the microwave backhauling and physical layer of next generation wireless networks project. Then, he joined the University of Luxembourg, Luxembourg, as a Research Associate from April 2012 to August 2012, where he was responsible for new interference cancellation and synchronization schemes for satellite communication links. From 2012 to 2015, he was an Assistant Professor with California State University, Bakersfield, CA, USA. Since 2015, he has been an Assistant Professor with Boise State University, Boise, ID, USA. His research interests include applied signal processing and physical layer of wireless communication systems, including millimeter-wave systems, reconfigurable antennas, energy harvesting systems, synchronization, and channel estimation.

Dr. Mehrpouyan has close to 50 publications in prestigious IEEE journals and conferences. He is an Associate Editor of the IEEE COMMUNICATION LETTERS and he has also served as a TPC member for the IEEE Globecom, ICC, VTC, etc. He has also been involved with industry leaders such as Ericsson AB and Blackberry. He has received many scholarships and awards including the IEEE Globecom Early Bird Student Award, NSERC-IRDF, NSERC PGS-D, NSERC CGS-M Alexander Graham Bell, and B.C. Wireless Innovation.



**Meixia Tao** (S'00–M'04–SM'10) received the B.S. degree in electronic engineering from Fudan University, Shanghai, China, in 1999, and the Ph.D. degree in electrical and electronic engineering from Hong Kong University of Science and Technology, Hong Kong, in 2003.

She is currently a Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University, China. Prior to that, she was a Member of Professional Staff with Hong Kong Applied Science and Technology Research Institute, Hong Kong, from 2003 to 2004, and a Teaching Fellow then an Assistant Professor with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, from 2004 to 2007. Her research interests include content-centric wireless networks, resource allocation, interference management, and physical layer security.

Dr. Tao is a member of the Executive Editorial Committee of the IEEE TRANSACTION ON WIRELESS COMMUNICATIONS since January 2015. She serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE WIRELESS COMMUNICATIONS LETTERS. She was on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2007 to 2011 and the IEEE COMMUNICATIONS LETTERS from 2009 to 2012. She has also served as the TPC chair of IEEE/CIC ICC 2014 and as Symposium Co-Chair of IEEE ICC 2015.

She was the recipient of the IEEE Heinrich Hertz Award for Best Communications Letters in 2013, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2009, and the International Conference on Wireless Communications and Signal Processing (WCSP) Best Paper Award in 2012.



**Yingbo Hua** (S'86–M'88–SM'92–F'02) received the B.S. degree from Southeast University, Nanjing, China, in 1982, and the Ph.D. degree from Syracuse University, Syracuse, NY, USA, in 1988.

He held a faculty position with the University of Melbourne, Melbourne, Australia, from 1990 to 2000, where he was promoted to the rank of Reader and Associate Professor in 1996. After taking a leave as a Visiting Professor at Hong Kong University of Science and Technology, Hong Kong, from 1999 to 2000, and a Consultant with Microsoft Research in summer 2000, he joined the University of California, Riverside, CA, USA, in 2001, where he is a Senior Full Professor. He has authored/coauthored over 300 articles and coedited three volumes of books, with thousands of citations, in the fields of sensing, signal processing, and communications.

Dr. Hua has served as Editor, Guest Editor, Member of Editorial Board and/or Member of Steering Committee for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE SIGNAL PROCESSING LETTERS, *EURASIP Signal Processing*, the *IEEE Signal Processing Magazine*, the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, and the IEEE WIRELESS COMMUNICATION LETTERS. He has served on Technical and/or Organizing Committees for over 50 international conferences and workshops. He is a Fellow of AAAS from 2011.