Multi-Relay Communications in the Presence of Phase Noise and Carrier Frequency Offsets

Omar H. Salim, *Member, IEEE*, Ali A. Nasir, *Member, IEEE*, Hani Mehrpouyan, *Member, IEEE*, and Wei Xiang, *Senior Member, IEEE*

***Abstract*— Impairments such as time varying phase noise (PHN) and carrier frequency offset (CFO) result in loss of synchronization and poor performance of multi-relay communication systems. Joint estimation of these impairments is necessary in order to correctly decode the received signal at the destination. In this paper, we address spectrally efficient multi-relay transmission scenarios where all the relays simultaneously communicate with the destination. We propose an iterative pilot-aided algorithm based on the expectation conditional maximization for joint estimation of multipath channels, Wiener PHNs, and CFOs in decode-and-forward-based multi-relay orthogonal frequency division multiplexing systems. Next, a new expression of the hybrid Cramér–Rao lower bound (HCRB) for the multi-parameter estimation problem is derived. Finally, an iterative receiver based on an *extended Kalman filter* for joint data detection and PHN tracking is employed. Numerical results show that the proposed estimator outperforms existing algorithms and its mean square error performance is close to the derived HCRB at different signal-to-noise ratios for different PHN variances. In addition, the combined estimation algorithm and the iterative receiver can significantly improve average bit-error rate (BER) performance compared with existing algorithms. In addition, the BER performance of the proposed system is close to the ideal case of perfect channel impulse responses, PHNs, and CFOs estimation.**

***Index Terms*— OFDM, multi-relay, channel estimation, phase noise, carrier frequency offset, hybrid Cramér-Rao lower bound, data detection.**

I. Introduction

A. Motivation and Related Works

MULTI-RELAY systems have attracted considerable research interests due to their potential to offer an effective solution to the issues faced by next generation (5G) cellular networks, such as significant path loss and shadowing at millimeter-wave (mmWave) frequencies [1], [2]. By employing multiple relay, one can enhance the range of mmWave links, while concurrently providing cooperative diversity to overcome shadowing due to obstacles and humans [3], [4]. In contrast to single-input single-output (SISO) systems, which may result in single phase noise (PHN) and carrier frequency offset (CFO), the multi-relay networks have multiple distributed nodes and each one has its own local oscillator. Thus, this gives rise to *multiple phase noises* (PHNs) and *multiple carrier frequency offsets* (CFOs) at the destination.

The motivation of adopting the low-cost oscillators at source, relays and destination, and providing high data rates in gigabits per second or even higher lead to the distortion of transmitted signal with different impairments such as CFO and PHN. In addition, PHN has more pronounced effect on system performance at higher frequencies, e.g., V-band/60 GHz and E-band/70-80 GHz [5]. Thus, it is increasingly important to develop efficient and accurate estimation algorithms for compensating the channels, PHNs, and CFOs to achieve an accurate synchronization amongst all communication nodes.

Orthogonal frequency division multiplexing (OFDM) is employed in multi-relay systems to increase the transmission bandwidth efficiency and mitigates the effect of frequency-selective fading. However, the presence of multiple PHN and CFO results in a common phase error (CPE) and inter-carrier interference (ICI) at the destination node, and the estimation of channel impulse response (CIR) for each link becomes challenging [6]. On the other hand, accurate estimation of CIRs in the presence of PHNs and CFOs is required for coherent detection at the destination. OFDMA technique can be used to assign different subcarriers to different relays. However, as mentioned in [7], OFDMA is restricted approach and may result in significant loss of spectral efficiency.

Many algorithms for a joint channel, PHN and/or CFO estimation in SISO and MIMO systems are proposed in [8]–[17]. However, the system models in [8]–[17] only consider a single oscillator at the transmitter and the receiver and thus requires the estimation of single PHN and/or single CFO parameter. In contrast, each relay in multi-relay systems has its own local oscillator and the received signal at the destination is affected by multiple PHN and CFO parameters. Thus, the estimation algorithms in [8]–[17] cannot be applied to estimate the required multiple PHN and CFO parameters at the destination of a multi-relay network.

Manuscript received March 17, 2016; revised August 11, 2016 and October 4, 2016; accepted October 20, 2016. Date of publication October 27, 2016; date of current version January 13, 2017. This research is in part support by the National Science Foundation EARS Grant award number 1642865. The associate editor coordinating the review of this paper and approving it for publication was T. Q. Duong.

O. H. Salim is with the School of Mechanical and Electrical Engineering, University of Southern Queensland, Toowoomba, QLD 4350, Australia (e-mail: omarhazim.salim@usq.edu.au).

A. A. Nasir is with the Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran 3126, Saudi Arabia (e-mail: anasir@kfupm.edu.sa).

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

W. Xiang is with the College of Science, Technology and Engineering, James Cook University, Cairns, QLD 4870, Australia (e-mail: wei.xiang@jcu.edu.au).

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/TCOMM.2016.2622704

In [18], the authors have presented general issues that need careful design for the successful implementation of OFDMA-based multi-hop networks. However, no estimation and detection algorithms are presented in [18]. In addition, the effects of PHN are not studied in [18]. In [19], channel estimation in the presence of CFOs is analyzed in decode-and-forward (DF) and amplify-and-forward (AF) cooperative systems. However, the authors in [19] do not take the effect of PHN into account. In [20] and [21], CFO estimation is investigated for DF and AF cooperative systems, respectively. However, the proposed algorithms in [20] and [21] are based on the assumption of perfect knowledge of channels. Moreover, in [21], a minimum mean square error (MMSE) equalizer is used to equalize the ICI, which is computationally very complex. In [22], channel estimation in the presence of PHN is investigated. However, the effect of CFOs is not taken into account. More importantly, [19]–[22] do not provide the hybrid Cramér-Rao lower bound (HCRB) for joint estimation of multiple impairments in multi-relay systems, which would provide essential information about the absolute performance of the estimation scheme and these bounds can be applied to obtain lower bounds on the performance of multi-relay network in the presence of imperfect CIR, PHN, and CFO estimation. The problem of joint channel, CFO, and PHN estimation is considered in the context of OFDM relay networks in [5]. However, the relaying approach in [5] is based on a single relay. In addition, the estimation approach in [5] is based on the *maximum a posteriori* (MAP) criterion, which is computationally very complex. Joint channel and CFO estimation based on the expectation-conditional maximization (ECM) approach was proposed in [23] for OFDMA uplink systems. However, in [23], the authors do not take the effect of multiple PHN parameters into account. In [7], the authors designed optimal training sequences for multi-user multi-input multi-output (MIMO)-OFDM systems and evaluated the performance of training sequences in the presence of residual PHN or residual CFO. However, the estimation approach in [7] depends on the orthogonality between the training sequences to reduce the effects of PHN or CFO. In addition, the estimation method in [7] does not provide any means of estimating or tracking multiple PHN and CFO parameters. Recently, we consider the problem of joint channel, PHN, and CFO estimation in OFDM AF and DF relay networks in [24] and [25]. However, in [24] and [25], the system model is based on time-division multiple-access (TDMA), which leads to a significant loss in spectral efficiency since each relay's parameters are estimated turn by turn. Moreover, the problem of joint data detection and PHN mitigation for multi-relay systems is not presented in [24] and [25]. In addition, the HCRB for joint channel, CFO, and PHN estimation, and the computational complexity of the estimation and detection for multi-relay networks are not addressed in [24] and [25]. Most recently, we consider the problem of joint PHN multi-parameter estimation and data detection for light field video transmission in MIMO-OFDM systems in [26]. However, the proposed algorithm in [26] is based on the assumption of accurate synchronization of CFOs. Moreover, the HCRB for joint estimation of mutiple impairments is not derived in [26]. In addition, the effects of joint channels, PHNs, and CFOs estimation on data detection is not addressed in [26]. Finally, in [27], a detection algorithm based on the Monte Carlo technique and the Bayesian approach for multi-user system in the presence of mutiple PHNs and CFOs has been proposed. However, the estimation algorithm in [27] requires the presence of multiple antennas at each transmitter and the application of STBCs with special trasnsmission structure. In our setup, that trasnmission is from multiple relays, where single antenna is employed at each relay to ensure implementation simplicity. Thus, the particular algorithm proposed in [27] is not applicable to our setup. In addition, no closed form expressions to estimate the channel, PHN, and CFO parameters are presented in [27]. More importantly, in [27], the HCRB for joint estimation of multiple channel, PHN and CFO parameters is not derived.

Given the time-varying nature of PHN, we need to track it not only during the training interval but also during the data transmission interval. Hence, following the training period, a receiver structure for joint data detection and PHN mitigation during the data transmission period is required. In the existing literature, joint data detection and mitigation of multiple PHN parameters is analyzed in [5] and [22]. However, the PHN tracking in [5] and [22] requires the application of pilots throughout an OFDM data symbol to compensate the CPE, which adversely affects the bandwidth efficiency and data detection performance. As will be explained in Section VII, the data detection approach of using the pilots to track the PHN over the data packet, as used in [5] and [22], has poor BER and lower PHN estimation performance compared to the *extended Kalman filter* (EKF) based detector in this paper.

B. Contributions

In this paper, a computationally efficient algorithm based on the ECM approach for joint estimation of channels, PHNs, and CFOs in OFDM-based DF relaying systems is presented. In the presence of time-varying PHN, an iterative data detection algorithm is also proposed to detect the data symbols. The major contributions of this paper are summarised as follows:

1) This paper addresses spectrally-efficient multi-relay transmission scenarios where the relays simultaneously send their signals to the destination, then the impairment parameters are estimated using an iterative pilot-aided algorithm based on the ECM algorithm at the destination. The proposed algorithm can estimate multiple *unknown* channel gains, PHNs and CFOs. In addition, we drive a closed-form estimator to obtaining the CFO and channel parameters. Based on simulation results, the proposed estimator is found only need few iterations to estimate the multiple impairments over the transmission packet.
2) We derive the HCRB for joint CIRs, PHNs, and CFOs estimation in DF-based multi-relay OFDM systems. Simulation results show that the mean square error (MSE) of the proposed algorithm is closer to the HCRB at different signal-to-noise ratios (SNRs).

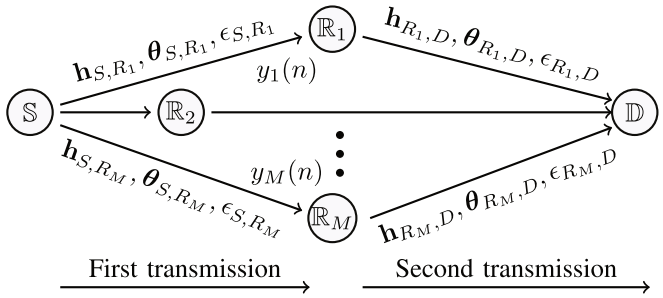


Fig. 1. The system block diagram for the multi-relay communication network.

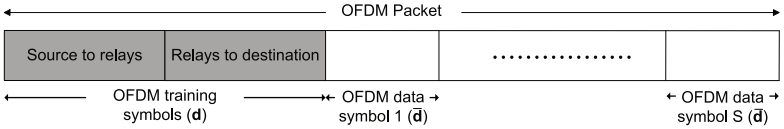


Fig. 2. Timing diagram for transmission of training and data symbols within an OFDM packet.

3) An iterative data detection algorithm based on the EKF for tracking the *unknown* time-varying PHNs throughout the OFDM data packet is presented. Simulations are carried out to investigate the performance of the proposed estimator and detector. Comparing with existing algorithms, the simulation results demonstrate that the combined estimation and detection algorithms significantly improve the MSE and the bit error rate (BER) performance. In addition, the BER performance of the proposed system is closer to the ideal case of perfect CIRs, PHNs, and CFOs estimation.

C. Notation

Superscripts $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ denote the conjugate, the conjugate transpose, and the transpose operators, respectively. Bold face small letters, e.g., $\mathbf{x}$, are used for vectors, bold face capital alphabets, e.g., $\mathbf{X}$, are used for matrices, and $[\mathbf{X}]_{x,y}$ represents the entry in row x and column y of $\mathbf{X}$. $\mathbf{I}_X$, $\mathbf{0}_{X\times X}$, and $\mathbf{1}_{X\times X}$ denote the $X \times X$ identity, all zero, and all 1 matrices, respectively. The notation $\mathbf{X}(n_1 : n_2, m_1 : m_2)$ is used to denote a submatrix of $\mathbf{X}$ from row n_1 to row n_2 and from column m_1 to column m_2. $|\cdot|$ is the absolute value operator, $|\mathbf{x}|$ denotes the element-wise absolute value of a vector $\mathbf{x}$, and diag($\mathbf{x}$) is used to denote a diagonal matrix, where the diagonal elements are given by vector $\mathbf{x}$. $\mathbf{X} \succeq \bar{\mathbf{X}}$ indicates that matrix $(\mathbf{X}-\bar{\mathbf{X}})$ is positive semi-definite. $\hat{\mathbf{X}}$, $\hat{\mathbf{x}}$, and $\hat{x}$ represent the estimate matrix, vector, and element, respectively. $\mathbb{E}_{x,y}[\cdot]$ denotes the expectation over x and y. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex quantity, respectively. $\nabla_{\mathbf{x}}$ and $\triangle^{\mathbf{x}}_{\mathbf{y}}$ represent the first and the second-order partial derivatives operator, i.e., $\nabla_{\mathbf{x}} = [\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_N}]^T$ and $\triangle^{\mathbf{x}}_{\mathbf{y}} = \nabla_{\mathbf{y}} \times \nabla^T_{\mathbf{x}}$. $\mathcal{N}(\mu, \sigma^2)$ and $\mathcal{CN}(\mu, \sigma^2)$ denote real and complex Gaussian distributions with mean μ and variance σ^2, respectively. $\otimes$ denotes circular convolution. Finally, $\dot{z}$ denotes the Jacobian of z.

D. Organization

The rest of this paper is organized as follows: Section II describes the system model, the scenario under consideration, and the assumptions in this work, Section III derives hybrid Cramér-Rao lower bound. Section IV derives the proposed estimator, Section V presents the joint data detection and PHN mitigation, Section VI illustrates complexity analysis of the proposed system while Section VII provides simulation results that investigate the performance of the proposed estimator and detector. Finally, Section VIII concludes the paper and summarizes its key findings.

II. Signal Model

We consider a half-duplex space-division multiple-access (SDMA) SISO multi-relay communication system with one source node, $\mathbb{S}$, M relays, $\mathbb{R}_1, \ldots, \mathbb{R}_M$, and a single destination node, $\mathbb{D}$, as shown in Fig. 1. An OFDM packet of $(S+2)$ symbols as shown in Fig. 2 is considered, which consists of two training symbol and S data symbols. The training symbols are known by the relays and destination, while the data symbols consist of modulated data, where *no pilots are included*. The two training symbols are used to separately estimate the unknown CIRs and CFOs in the presence of unknown PHN for both transmission phases of the source to relays and relays to destination. As shown in Figs. 1 and 2, during the training period, the source node broadcasts the training symbol in the first transmission phase to M relays, then the CIR and CFO in the presence of PHN are estimated at each relay. In the second transmission phase, M relay nodes simultaneously transmit the training symbols to the destination node and the estimation of multiple CIR and CFO parameters in the presence of multiple PHN parameters is performed at the destination node. Next, during the data period, the data symbols are transmitted from the source to M relays in the first transmission phase, then M relays simultaneously decode, re-encode, and forward the source information to the destination node during the second transmission phase. The constant CIR and CFO are compensated by using their estimates obtained during the training phase, while during the data transmission, we track the time varying PHN and decode the data. Therefore, in order to guarantee the advantages of multi-relay diversity, there is a need to estimate the channel gains, time varying

PHN, and CFO parameters for the received signals at the destination node during both transmission phases.

In this paper, the following set of assumptions are adopted:

A1. The channels are modeled as quasi-static Rayleigh fading channels, i.e., they are constant and *unknown* over the OFDM packet duration but change from packet to packet,

A2. CFO is modeled as an *unknown deterministic parameter* over a packet and is assumed to change from packet to packet.

A3. The time-varying PHN is assumed to change from symbol to symbol and is modeled by a Wiener process, i.e., $\theta_q(n) = \theta_q(n-1) + \delta(n)$, $\forall\ n$ and $q \in \{S, R_m, D\}$, where $\theta_q(n)$ is the PHN at the nth instant, $\delta(n) \sim \mathcal{N}(0, \sigma_\delta^2)$ is PHN innovation and σ_δ^2 is the variance of the innovation process.

A4. The training symbol is assumed to be *known* at the relays and destination.

A5. The timing offsets are assumed to be perfectly estimated. Hence, it is not considered.

Note that assumptions A1, A2, A3, A4, and A5 are in line with previous studies and channel, PHN and CFO estimation algorithms in [6], [19]–[22], and [28]–[30]. Assumption A3 is also reasonable in many practical scenarios to describe the behavior of practical oscillators [6], [28]. In addition, assumption A4 is adopted in the IEEE 802.11ac/ad standards to estimate channel and CFO in [6] and [31]–[34].

The time-invariant composite CIR between any pair of nodes a and b is modeled as $h_{a,b}(\tau) = \sum_{l=0}^{L-1} h_{a,b}(l)\delta(\tau - lT_s)$, where $h_{a,b}(l)$ is the channel gain for the lth tap, $\delta(x)$ denotes the unit impulse function, and $a, b \in \{S, R_m, D\}$. L is the channel order, and $T_s = 1/B$, where B represents the total bandwidth. The channel order L is the same for any pair of nodes. For brevity, we define $\mathbf{h}_{a,b} \triangleq [h_{a,b}(0), h_{a,b}(1), \ldots, h_{a,b}(L-1)]^T$ and the channel gains $h_{a,b}(l)$ are modeled as complex Gaussian zero-mean random variables. The input data bits are first mapped to the complex symbols drawn from a signal constellation such as phase-shift keying (PSK) or quadrature amplitude modulation (QAM). Next, the source node, $\mathbb{S}$, transmits the modulated training symbol vector $\mathbf{d} \triangleq [d(0), d(1), \ldots, d(N-1)]^T$ to the destination during two transmission phases. Note that, in this paper as shown in Fig. 2, a symbol, $\mathbf{d}$, is used to indicate the training vector at the relays during training period, while $\bar{\mathbf{d}}$ is used to indicate the decoded data vector at the relays, that is transmitted further to the destination during data transmission period.

A. First Transmission Phase

The received signal at the mth relay, $\mathbb{R}_m$, is given by

$$\mathbf{z}_m = \mathbf{E}_{S,R_m}\mathbf{P}_{S,R_m}\mathbf{F}^H\mathbf{H}_{S,R_m}\mathbf{d} + \mathbf{v}_m, \tag{1}$$

where $\mathbf{z}_m \triangleq [z_m(0), z_m(1), \ldots, z_m(N-1)]^T$ is an $N\times 1$ vector, $\mathbf{E}_{S,R_m} \triangleq \text{diag}([e^{(j2\pi\epsilon_{S,R_m}/N)\times 0}, \ldots, e^{(j2\pi\epsilon_{S,R_m}/N)\times(N-1)}]^T)$ is the $N \times N$ CFO matrix, ϵ_{S,R_m} denotes the normalized CFO between $\mathbb{S} \longrightarrow \mathbb{R}_m$, $\mathbf{P}_{S,R_m} \triangleq \text{diag}([e^{j\theta_{S,R_m}(0)}, e^{j\theta_{S,R_m}(1)}, \ldots, e^{j\theta_{S,R_m}(N-1)}]^T)$ is the $N \times N$ PHN matrix, $\theta_{S,R_m}(n) \triangleq \theta_S(n) + \theta_{R_m}(n)$ for $n = 0, \ldots, N-1$ is the PHN at the nth instant between $\mathbb{S} \longrightarrow \mathbb{R}_m$, $\mathbf{H}_{S,R_m} \triangleq \text{diag}(\mathbf{F}_L\mathbf{h}_{S,R_m}) = \text{diag}(H_{S,R_m}[0], H_{S,R_m}[1], \ldots, H_{S,R_m}[N-1])$ is the $N \times N$ frequency-domain channel coefficient matrix, $\mathbf{F}_L$ is an $N \times L$ DFT matrix, i.e., $\mathbf{F}_L \triangleq \mathbf{F}(1:N, 1:L)$, $\mathbf{F}$ is an $N \times N$ DFT matrix, i.e., $[\mathbf{F}]_{l,v} \triangleq (1/\sqrt{N})e^{-j(2\pi vl/N)}$ for $v, l = 0, 1, \cdots, N-1$. Note that $H_{S,R_m}(n) \triangleq e^{j\bar{\theta}_{S,R_m}(0)}\bar{H}_{S,R_m}(n)$, and $\theta_{S,R_m}(n) \triangleq \bar{\theta}_{S,R_m}(n) - \bar{\theta}_{S,R_m}(0)$, this model helps to distinguish between the phase disturbance caused by PHN and the channel phase for the first sample, which in turn resolves the phase ambiguity in the joint estimation problem as indicated in Section IV, $\mathbf{D} \triangleq \text{diag}(\mathbf{d})$, $\mathbf{d} = [d(0), d(1), \ldots, d(N-1)]^T$ is the $N \times 1$ modulated training vector during training period and data vector during data transmission period, and $\mathbf{v}_m = [v_m(0), v_m(1), \ldots, v_m(N-1)]^T$ is AWGN vector at the mth relay, $\mathbb{R}_m$.

The estimation and detection problem between $\mathbb{S} \longrightarrow \mathbb{R}_m$ can be solved using source-to-relay estimation and detection techniques proposed for OFDM SISO systems in [35] and is not presented here to avoid repetition.

B. Second Transmission Phase

As shown in Fig. 2, in this phase, the relays apply the DF protocol on the received signals and forward them to the destination while the source is silent. [1]The received signal at the destination, D, for the training period can be defined as

$$\mathbf{r} = \sum_{m=1}^{M} \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D} + \mathbf{w}, \tag{2}$$

where $\mathbf{r} \triangleq [r(0), r(1), \ldots, r(N-1)]^T$ is an $N \times 1$ vector, $\mathbf{E}_{R_m,D} \triangleq \text{diag}([e^{(j2\pi\epsilon_{R_m,D}/N)\times 0}, e^{(j2\pi\epsilon_{R_m,D}/N)}, \ldots, e^{(j2\pi\epsilon_{R_m,D}/N)\times(N-1)}]^T)$ is an $N \times N$ CFO matrix, $\epsilon_{R_m,D}$ denotes the normalized CFO between $\mathbb{R}_m \longrightarrow \mathbb{D}$, $\mathbf{P}_{R_m,D} \triangleq \text{diag}([e^{j\theta_{R_m,D}(0)}, e^{j\theta_{R_m,D}(1)}, \ldots, e^{j\theta_{R_m,D}(N-1)}]^T)$ is the $N \times N$ PHN matrix, $\theta_{R_m,D}(n) \triangleq \theta_{R_m}(n) + \theta_D(n)$ for $n = 0, \ldots, N-1$, is the PHN at the nth instant between $\mathbb{R}_m \longrightarrow \mathbb{D}$, $\mathbf{h}_{R_m,D} \triangleq [h_{R_m,D}(0), h_{R_m,D}(1), \ldots, h_{R_m,D}(L-1)]^T$ is a $L \times 1$ CIR vector between $\mathbb{R}_m \longrightarrow \mathbb{D}$, $\mathbf{D}_m \triangleq \mathbf{\Lambda}_m\mathbf{D}$ is an $N \times N$ matrix, $\mathbf{D} \triangleq \text{diag}(\mathbf{d})$ is an $N \times N$ training matrix, $\mathbf{d}$ is an $N \times 1$ training symbol vector from the relay, $\mathbf{\Lambda}_m = \text{diag}(1, e^{j2\pi(L+1)m/N}, \ldots, e^{j2\pi(L+1)(N-1)m/N})$ is an $N \times N$ frequency modulation matrix and can be viewed as a frequency modulation and used to achieve the orthogonality between the training sequences at the destination [7]. $\mathbf{w}$ is an $N \times 1$ AWGN vector at the destination.

Equation (2) can be rewritten as follows

$$\mathbf{r} = \underbrace{\left[\mathbf{\Psi}_1, \mathbf{\Psi}_2, \ldots, \mathbf{\Psi}_M\right]}_{\mathbf{\Psi}} \mathbf{h}_{R,D} + \mathbf{w}, \tag{3}$$

where $\mathbf{\Psi} \triangleq \left[\mathbf{\Psi}_1, \mathbf{\Psi}_2, \ldots, \mathbf{\Psi}_M\right]$ is an $N \times ML$ matrix, $\mathbf{\Psi}_m \triangleq \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L$ is an $N \times L$ matrix, for $m = 1, \ldots, M$, and $\mathbf{h}_{R,D} \triangleq [\mathbf{h}_{R_1,D}^T, \ldots, \mathbf{h}_{R_M,D}^T]^T$ is a $ML \times 1$ CIR vector.

[1]The relays can also apply a precoding or beamforming approach at this stage to further enhance the system performance, but such approaches are beyond the scope of this work.

III. Derivation of the Hybrid Cramér-Rao Bound

In this section, the HCRB for joint estimation of channel, PHN, and CFO parameters is presented. Since the overall parameters of channel , $\mathbf{h}_{R,D}$, PHNs, $\boldsymbol{\theta}_{R,D}$, and CFOs, $\boldsymbol{\epsilon}_{R,D}$, need to be estimated, the parameter vector of interest, $\boldsymbol{\lambda}$, is given by

$$\boldsymbol{\lambda} = [\boldsymbol{\theta}_{R,D}^T \ \Re\{\mathbf{h}_{R,D}\}^T \ \Im\{\mathbf{h}_{R,D}\}^T \ \boldsymbol{\epsilon}_{R,D}^T]^T, \tag{4}$$

where $\boldsymbol{\lambda}$ comprises of both random and deterministic parameters, i.e., PHNs, $\boldsymbol{\theta}_{R,D}^T \triangleq [\boldsymbol{\theta}_{R_1,D}^T, \ldots, \boldsymbol{\theta}_{R_M,D}^T]^T$, $\boldsymbol{\theta}_{R_m,D}^T \triangleq [\theta_{R_m,D}(0), \ldots, \theta_{R_m,D}(N-1)]^T$ are random, while CIRs, $\mathbf{h}_{R,D} \triangleq [\mathbf{h}_{R_1,D}^T, \ldots, \mathbf{h}_{R_M,D}^T]^T$, and CFOs, $\boldsymbol{\epsilon}_{R,D}^T \triangleq [\epsilon_{R_1,D}, \ldots, \epsilon_{R_M,D}]^T$, are deterministic parameters. Thus, the HCRB instead of standard CRB is needed to be derived. The accuracy of estimating $\boldsymbol{\lambda}$ is lower bounded by the HCRB ($\boldsymbol{\Omega}$) as [36, pp. 1–85]

$$\mathbb{E}_{\mathbf{r},\boldsymbol{\theta}_{R,D}|\mathbf{h}_{R,D},\boldsymbol{\epsilon}_{R,D}} \left[(\hat{\boldsymbol{\lambda}}(\mathbf{r}) - \boldsymbol{\lambda})(\hat{\boldsymbol{\lambda}}(\mathbf{r}) - \boldsymbol{\lambda})^T \right] \succeq \boldsymbol{\Omega}. \tag{5}$$

Let us define $\boldsymbol{\Omega} \triangleq \mathbf{B}^{-1}$. Here, $\mathbf{B}$ is an $M(N+2L) \times M(N+2L)$ hybrid information matrix (HIM), which is determined according to the following theorem.

Theorem 1: The closed-form HIM for joint estimation of CIR, PHN, and CFO is given by

$$\mathbf{B} = \frac{2}{\sigma_w^2} \Re \left\{ \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} & \mathbf{B}_{14} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} & \mathbf{B}_{24} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} & \mathbf{B}_{34} \\ \mathbf{B}_{41} & \mathbf{B}_{42} & \mathbf{B}_{43} & \mathbf{B}_{44} \end{bmatrix} \right\}, \tag{6}$$

where $\mathbf{B}_{11} \triangleq \bar{\mathbf{Q}}_1^H \bar{\mathbf{Q}}_1 + \boldsymbol{\Lambda}$ is the $M(N-1) \times M(N-1)$ HIM for the estimation of $\boldsymbol{\theta}_{R,D}$, $\bar{\mathbf{Q}}_1 = \left[\bar{\boldsymbol{\Phi}}_1, \ldots, \bar{\boldsymbol{\Phi}}_M\right]$, $\bar{\boldsymbol{\Phi}}_m \triangleq \boldsymbol{\Phi}_m(2:N, 2:N)$, $\boldsymbol{\Phi}_m = j\mathrm{diag}(\mathbf{E}_{R_m D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m D})$ for $m = 1, \ldots, M$, $\boldsymbol{\Lambda}$ is an $M(N-1) \times M(N-1)$ tridiagonal matrix with diagonal elements given by $\frac{\sigma_w^2}{2\sigma_\delta^2}[1, 2, \ldots, 2, 1]$ and off-diagonal elements given by $\frac{-\sigma_w^2}{2\sigma_\delta^2}[1, \ldots, 1]$, $\mathbf{B}_{22} \triangleq \mathbf{Q}_2^H\mathbf{Q}_2$ is an $ML \times ML$ information matrix for the estimation of real part of $\mathbf{h}_{R,D}$, $\mathbf{Q}_2 = \left[\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_M\right]$, $\boldsymbol{\gamma}_m = \mathbf{E}_{R_m D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L$ for $m = 1, \ldots, M$, $\mathbf{B}_{33} \triangleq \mathbf{Q}_2^H\mathbf{Q}_2$ is an $ML \times ML$ information matrix for the estimation of imaginary part of $\mathbf{h}_{R,D}$, $\mathbf{B}_{44} \triangleq \mathbf{Q}_3^H\mathbf{Q}_3$ represents the information for the estimation of CFOs, $\boldsymbol{\epsilon}_{R,D}$, $\mathbf{Q}_3 = \left[\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_M\right]$, $\boldsymbol{\beta}_m = \acute{\mathbf{E}}_{Rm D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{Rm D}$, $\acute{\mathbf{E}}_{Rm D} \triangleq \mathrm{diag}([0, \frac{j2\pi}{N}e^{(j2\pi\epsilon_{Rm D}/N)}, \ldots, \frac{j2\pi(N-1)}{N}e^{(j2\pi\epsilon_{Rm D}/N)\times(N-1)}]^T)$ for $m = 1, \ldots, M$, $\mathbf{B}_{12} = \mathbf{B}_{21}^H \triangleq -j\bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_2$, $\bar{\mathbf{Q}}_2 = \left[\bar{\boldsymbol{\gamma}}_1, \ldots, \bar{\boldsymbol{\gamma}}_M\right]$, $\bar{\boldsymbol{\gamma}}_m = \boldsymbol{\gamma}_m(2:N, 1:L)$ for $m = 1, \ldots, M$, $\mathbf{B}_{13} = \mathbf{B}_{31}^H \triangleq \bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_2$, $\mathbf{B}_{14} = \mathbf{B}_{41}^H \triangleq \bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_3$, $\bar{\mathbf{Q}}_3 = \left[\bar{\boldsymbol{\beta}}_1, \ldots, \bar{\boldsymbol{\beta}}_M\right]$, $\bar{\boldsymbol{\beta}}_m = \boldsymbol{\beta}_m(2:N)$ for $m = 1, \ldots, M$, $\mathbf{B}_{23} = \mathbf{B}_{32}^H \triangleq j\mathbf{Q}_2^H\mathbf{Q}_2$, $\mathbf{B}_{24} = \mathbf{B}_{42}^H \triangleq j\mathbf{Q}_2^H\mathbf{Q}_3$, and $\mathbf{B}_{34} = \mathbf{B}_{43}^H \triangleq \mathbf{Q}_2^H\mathbf{Q}_3$.

Proof: The proof is given in Appendix A.

Finally, the HCRB, $\boldsymbol{\Omega}$, is given by the inverse of the HIM. i.e., $\boldsymbol{\Omega} = \mathbf{B}^{-1}$. Note that the HCRB of the channel, $\mathbf{h}_{R,D}$, is obtained by adding the HCRB for real and imaginary parts of channels, i.e., $\mathrm{HCRB}\{\mathbf{h}_{R,D}\} = \mathrm{HCRB}\{\Re\{\mathbf{h}_{R,D}\}\} + \mathrm{HCRB}\{\Im\{\mathbf{h}_{R,D}\}\}$ [37].

IV. Iterative Estimation

The parameter vector of interest in (4) can be rearranged into groups , i.e., $\boldsymbol{\lambda} \triangleq [\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_M^T]^T$, where $\boldsymbol{\lambda}_m^T \triangleq [\boldsymbol{\theta}_{R_m,D}^T, \mathbf{h}_{R_m,D}^T, \epsilon_{R_m,D}]^T$, for $m = \{1, \ldots, M\}$. During the estimation process using ECM, each group, $\boldsymbol{\lambda}_m$, is updated while keeping the remaining groups fixed at their latest updated values. [2] In addition, for each group *a hidden data set* is selected [38]. In this case, the hidden data set denoted by $\mathbf{y}_m$ for $\boldsymbol{\lambda}_m$ is given by

$$\mathbf{y}_m = \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D} + \mathbf{w}_m, \tag{7}$$

where $\mathbf{w}_m$ is the $N \times 1$ AWGN vector. The updating process for $\boldsymbol{\lambda}_m$ at the ith iteration in the proposed ECM estimator consists of the *E- and M-Steps*.

A. E-Step

In this step, a hidden data set is calculated from the received signal, $\mathbf{r}$, in (2) and depends on the latest CIR, PHN, and CFO estimates obtained from the previous iteration. Thus, while setting $\boldsymbol{\lambda}_\ell = \hat{\boldsymbol{\lambda}}_\ell^{[i]}, \forall \ell \neq m$, the expectation of the *log-likelihood function (LLF)* of the hidden data set for the parameter $\boldsymbol{\lambda}_m$, $\mathrm{N}(\boldsymbol{\lambda}_m|\hat{\boldsymbol{\lambda}}_\ell^{[i]})$ is determined as

$$\mathrm{N}(\boldsymbol{\lambda}_m|\hat{\boldsymbol{\lambda}}^{[i]}) \triangleq \mathbb{E}\left\{ \log p(\mathbf{y}_m|\boldsymbol{\lambda}_m, \left\{\hat{\boldsymbol{\lambda}}_\ell^{[i]}\right\}_{\ell\neq m}) \middle| \mathbf{r}, \hat{\boldsymbol{\lambda}}_\ell^{[i]} \right\}, \tag{8}$$

where

$$\begin{aligned} p(\mathbf{y}_m|\boldsymbol{\lambda}_m, \left\{\hat{\boldsymbol{\lambda}}_\ell^{[i]}\right\}_{\ell\neq m}) &= p(\mathbf{y}_m|\boldsymbol{\lambda}_m) \\ &= \frac{1}{\pi\sigma_w^2}\exp\left\{ \frac{\|\mathbf{y}_m - \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D}\|^2}{\sigma_w^2} \right\}. \end{aligned} \tag{9}$$

Substituting (9) into (8), we obtain

$$\begin{aligned} &\mathrm{N}(\boldsymbol{\lambda}_m|\hat{\boldsymbol{\lambda}}^{[i]}) \\ &= \mathrm{C}_1 - \frac{1}{\sigma_w^2}\mathbb{E}\left\{ \|\mathbf{y}_m - \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D}\|^2 \middle| \mathbf{r}, \hat{\boldsymbol{\lambda}}_\ell^{[i]} \right\} \\ &= \mathrm{C}_1 - \frac{1}{\sigma_w^2}\mathbb{E}\left\{ \left\|\mathbf{y}_m^{[i]} - \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D}\right\|^2 \right\}, \end{aligned} \tag{10}$$

where $\mathrm{C}_1 = \log(\pi\sigma_w^2)$ is a constant and

$$\begin{aligned} \mathbf{y}_m^{[i]} &\triangleq \mathbb{E}\left\{ \mathbf{y}_m|\mathbf{r}, \hat{\boldsymbol{\lambda}}^{[i]} \right\} = \hat{\mathbf{E}}_{R_m,D}\hat{\mathbf{P}}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\hat{\mathbf{h}}_{R_m,D} \\ &\quad + \left(\mathbf{r} - \sum_{\ell=1}^{M} \hat{\mathbf{E}}_{R_\ell,D}\hat{\mathbf{P}}_{R_\ell,D}\mathbf{F}^H\mathbf{D}_\ell\mathbf{F}_L\hat{\mathbf{h}}_{R_\ell,D} \right) \\ &= \mathbf{r} - \sum_{\substack{\ell=1 \\ \ell\neq m}}^{M} \hat{\mathbf{E}}_{R_\ell,D}^{[i]}\hat{\mathbf{P}}_{R_\ell,D}^{[i]}\mathbf{F}^H\mathbf{D}_\ell\mathbf{F}_L\hat{\mathbf{h}}_{R_\ell,D}^{[i]}. \end{aligned} \tag{11}$$

[2] The convergence is analyzed later.

B. M-Step

In this step, the CIR, CFO, and PHN parameters between the mth relay and the destination are estimated, i.e., $\hat{\boldsymbol{\theta}}_{R_m,D}, \hat{\mathbf{h}}_{R_m,D}, \hat{\epsilon}_{R_m,D}$. $\boldsymbol{\lambda}_m$ in the $(i+1)$th iteration, $\hat{\boldsymbol{\lambda}}_m^{[i+1]}$, is determined as

$$\begin{aligned}\hat{\boldsymbol{\lambda}}_m^{[i+1]} &= \arg\max_{\boldsymbol{\lambda}_m} \mathrm{N}(\boldsymbol{\lambda}_m|\hat{\boldsymbol{\lambda}}^{[i]}) \\ &= \arg\min_{\boldsymbol{\lambda}_m} \left\|\hat{\mathbf{y}}_m^{[i]} - \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D}\right\|^2. \end{aligned} \tag{12}$$

In order to further reduce the complexity associated with the *M-step* of the EM algorithm, the ECM scheme [39] is applied in this section, where the cost function in (12) is minimized with respect to one of the parameters of interest while keeping the remaining parameters at their most recently updated values [39], [40]. The steps of the ECM approach as follows.

1) PHN Estimation: In this step, $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i+1]}$ can be determined as follows. The nth symbol of the signal vector, $\hat{y}_m^{[i]}(n)$ in (11), is first multiplied by $e^{-j2\pi\hat{\epsilon}_{R_m,D}^{[i]}n/N}$ for $n = \{0, 1, \ldots, N-1\}$, where $\hat{\epsilon}_{R_m,D}^{[i]}$ is the latest CFO estimate obtained from the previous iteration. Next, the signal $u_m(n) \triangleq e^{-j2\pi n\hat{\epsilon}_{R_m,D}^{[i]}/N}\hat{y}_m^{[i]}(n)$ is used to estimate the PHN vector. The signal $u_m(n)$ can be written as

$$\begin{aligned} u_m(n) &= e^{-j2\pi n\epsilon_{R_m,D}^{[i]}/N}\hat{y}_m^{[i]}(n) \\ &= e^{j2\pi n\Delta\hat{\epsilon}_{R_m,D}/N}e^{j\theta_{R_m,D}(n)}s_m^{[i]}(n) + \tilde{\alpha}_m(n), \end{aligned} \tag{13}$$

where $s_m^{[i]}(n)$ is the nth symbol of the vector, $\mathbf{s}_m^{[i]} \triangleq \mathbf{F}^H\mathbf{D}\mathbf{F}_L\hat{\mathbf{h}}_{R_m,D}^{[i]}$, $\Delta\hat{\epsilon}_{R_m,D} \triangleq \epsilon_{R_m,D} - \hat{\epsilon}_{R_m,D}^{[i]}$, $\tilde{\alpha}_m(n) \triangleq \alpha_m(n)e^{-j2\pi n\hat{\epsilon}_{R_m,D}^{[i]}/N}$ and $\alpha_m(n)$ is the nth symbol of the overall noise vector, $\boldsymbol{\alpha}_m$. $\boldsymbol{\alpha}_m$ is the result of thermal noise and residual interference from the relays and as shown in [23], it is nearly Gaussian distributed with zero-mean and some variance σ_w^2. For the proposed problem, the state and observation equations at time n are given by

$$\theta_{R_m,D}(n) = \theta_{R_m,D}(n-1) + \delta_{R_m,D}(n), \tag{14}$$

$$u_m(n) = g_m(n) + \alpha_m(n) = e^{j\theta_{R_m,D}(n)}s_m(n) + \tilde{\alpha}_m(n), \tag{15}$$

respectively. Since the observation equation in (15) is a non-linear function of the unknown state vector $\boldsymbol{\theta}_{R_m,D}$, the EKF is used instead of a simple Kalman filter. Based on Taylor series expansion, the EKF can linearize the non-linear observation equation in (15) about the current estimates [41]. Thus, the Jacobian of $g_m(n)$ is evaluated by computing the first order partial derivative of $g_m(n)$ with respect to $\theta_{R_m,D}(n)$ as

$$\begin{aligned}\dot{g}_m(n) &= \left.\frac{\partial g_m(\theta_{R_m,D}(n))}{\partial\theta_{R_m,D}(n)}\right|_{\theta_{R_m,D}(n)=\hat{\theta}_{R_m,D}(n|n-1)} \\ &= jg(\hat{\theta}_{R_m,D}(n|n-1)) = je^{j\hat{\theta}_{R_m,D}^{[i]}(n|n-1)}\hat{s}_m(n), \end{aligned} \tag{16}$$

where $\dot{g}_m$ denotes the Jacobian of g_m evaluated at $\theta_{R_m,D}(n)$. The first and second moments of the state vector at the $(i+1)$th iteration denoted by $\hat{\theta}_{R_m,D}^{[i+1]}(n|n-1)$ and $M^{[i+1]}(n|n-1)$, respectively, are given by

$$\hat{\theta}_{R_m,D}^{[i+1]}(n|n-1) = \hat{\theta}_{R_m,D}^{[i+1]}(n-1|n-1), \tag{17}$$

$$M_m^{[i+1]}(n|n-1) = M_m^{[i+1]}(n-1|n-1) + \sigma_{\delta_{R_m,D}}^2, \tag{18}$$

repectively. Given the observation $u_m(n)$, the Kalman gain $K_m(n)$, posteriori state estimate $\hat{\theta}_{R_m,D}^{[i+1]}(n|n)$, and the filtering error covariance, $M_m^{[i+1]}(n|n)$ are given by

$$\begin{aligned} K_m(n) &= M_m^{[i+1]}(n|n-1)\dot{g}_m^*(\theta_{S,R,D}(n|n-1)) \\ &\quad\times\Big(\dot{g}_m(\theta_{R_m,D}(n|n-1)) \times M_m^{[i+1]}(n-1|n-1) \\ &\quad\times\dot{g}_m^*(\theta_{R_m,D}(n|n-1)) + \sigma_w^2\Big)^{-1}, \end{aligned} \tag{19}$$

$$\begin{aligned}\hat{\theta}_{R_m,D}^{[i+1]}(n|n) &= \hat{\theta}_{R_m,D}^{[i+1]}(n|n-1) \\ &\quad+\Re\big\{K_m(n)\big(u_m(n) - e^{j\hat{\theta}_{R_m,D}^{[i+1]}(n|n-1)}\hat{s}_m^{[i]}(n)\big)\big\}, \end{aligned} \tag{20}$$

$$\begin{aligned} M_m^{[i+1]}(n|n) &= \Re\big\{M_m^{[i+1]}(n|n-1) \\ &\quad- K_m(n)\dot{g}_m(\theta_{R_m,D}(n|n-1)) \\ &\quad\times M_m^{[i+1]}(n|n-1)\big\}, \end{aligned} \tag{21}$$

respectively. Before starting the EKF recursion (16)-(21), $\hat{\theta}_{R_m,D}^{[1]}(1|0)$ and $M_m^{[1]}(1|0)$ are initialized by $\hat{\theta}_{R_m,D}^{[1]}(1|0) = 0$ and $M_m^{[1]}(1|0) = \sigma_{\delta_{R_m,D}}^2$, respectively.

Note that the EKF based estimation works regardless of the selected PHN model. This is because the choice of a different PHN model only changes the generation of PHN vector.

2) CFO Estimation: In this step, $\hat{\epsilon}_{R_m,D}^{[i+1]}$ can be determined as follows. By setting $\boldsymbol{\theta}_{R_m,D}$ and $\mathbf{h}_{R_m,D}$ to their latest updated values, $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i+1]}$ and $\mathbf{h}_{R_m,D}^{[i]}$, respectively, the updated value of $\epsilon_{R_m,D}$ at the $(i+1)$th iteration, $\hat{\epsilon}_{R_m,D}^{[i+1]}$, can be determined as

$$\begin{aligned}&\hat{\epsilon}_{R_m,D}^{[i+1]} \\ &= \arg\min_{\epsilon_{R_m,D}}\left\|\hat{\mathbf{y}}_m^{[i]} - \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D}\right\|^2\Big|_{\substack{\boldsymbol{\theta}_{R_m,D}=\hat{\boldsymbol{\theta}}_{R_m,D}^{[i+1]} \\ \mathbf{h}_{R_m,D}=\hat{\mathbf{h}}_{R_m,D}^{[i]}}} \\ &= \arg\min_{\epsilon_{R_m,D}}\sum_{n=0}^{N-1}\left\|\hat{y}_m^{[i]}(n) - e^{j2\pi\epsilon_{R_m,D}n/N}e^{j\hat{\theta}_{R_m,D}^{[i+1]}(n)}\hat{s}_m^{[i]}(n)\right\|^2 \\ &= \arg\max_{\epsilon_{R_m,D}}\sum_{n=0}^{N-1}\Re\{(\hat{y}_m^{[i]}(n))^*\hat{S}_m^{[i]}(n)e^{j2\pi\epsilon_{R_m,D}n/N}\}, \end{aligned} \tag{22}$$

where $\hat{S}_m^{[i]}(n) = e^{j\hat{\theta}_{R_m,D}^{[i+1]}(n)}\hat{s}_m^{[i]}(n)$. In order to handle the nonlinearity of (22), Taylor series expansion can be used to approximate the term $e^{j2\pi\epsilon_{R_m,D}n/N}$ around the pervious CFO estimate, $\hat{\epsilon}_{R_m,D}^{[i]}$, up to the second order term as

$$\begin{aligned}&e^{j2\pi\epsilon_{R_m,D}n/N} \\ &= e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]}n/N} + (\epsilon_{R_m,D} - \hat{\epsilon}_{R_m,D}^{[i]})\left(j\frac{2\pi}{N}n\right) \\ &\quad\times e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]}n/N} + \frac{1}{2}(\epsilon_{R_m,D} - \hat{\epsilon}_{R_m,D}^{[i]})^2\left(j\frac{2\pi}{N}n\right)^2 \\ &\quad\times e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]}n/N}. \end{aligned} \tag{23}$$

Substituting (23) into (22), $\hat{\epsilon}_{R_m,D}^{[i+1]}$ is given by

$$\begin{aligned}\hat{\epsilon}_{R_m,D}^{[i+1]} &= \arg\max_{\epsilon_{R_m,D}} \Bigg\{ \sum_{n=0}^{N-1} \Re\big\{(\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N} \\ &\quad + (\epsilon_{R_m,D} - \hat{\epsilon}_{R_m,D}^{[i]}) \sum_{n=0}^{N-1} \Re\big\{(\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) \\ &\quad \left(j\frac{2\pi}{N}n\right) e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N}\big\} + \frac{1}{2}(\epsilon_{R_m,D} - \hat{\epsilon}_{R_m,D}^{[i]})^2 \\ &\quad \sum_{n=0}^{N-1} \Re\big\{(\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) \left(j\frac{2\pi}{N}n\right)^2 e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N}\big\}\Bigg\}. \end{aligned} \tag{24}$$

Taking the derivative of (24) with respect to $\epsilon_{R_m,D}$ and equating the result to zero, the estimate of $\epsilon_{R_m,D}$ at the $(i+1)$th iteration is given by

$$\hat{\epsilon}_{R_m,D}^{[i+1]} = \hat{\epsilon}_{R_m,D}^{[i]} - \frac{N}{2\pi} \frac{\sum_{n=0}^{N-1} n\Im\big\{(\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N}\big\}}{\sum_{n=0}^{N-1} n^2\Re\big\{(\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N}\big\}}. \tag{25}$$

3) CIR Estimation: In this step, $\hat{\mathbf{h}}_{R_m,D}^{[i+1]}$ can be determined as follows. By setting $\boldsymbol{\theta}_{R_m,D}$ and $\epsilon_{R_m,D}$ to their latest updated values, $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i+1]}$ and $\hat{\epsilon}_{R_m,D}^{[i+1]}$, respectively, the updated value of $\mathbf{h}_{R_m,D}$ at the $(i+1)$th iteration, $\hat{\mathbf{h}}_{R_m,D}^{[i+1]}$ is calculated. The negative log likelihood function for mth relay can be written as

$$\log p(\mathbf{y}_m; \epsilon_{R_m,D}) = \mathrm{C}_2 + \left\|\mathbf{y}_m^{[i]} - \mathbf{E}_{R_m,D}^{[i+1]} \mathbf{P}_{R_m,D}^{[i+1]} \boldsymbol{\Gamma}_m \mathbf{h}_{R_m,D}\right\|^2, \tag{26}$$

where $\boldsymbol{\Gamma}_m \triangleq \mathbf{F}^H \mathbf{D}_m \mathbf{F}_L$ and C_2 is a constant. By taking the derivative of negative LLF in (26) with respect to $\mathbf{h}_{R_m,D}$, the estimate of $\mathbf{h}_{R_m,D}$ at the $(i+1)$th iteration, $\hat{\mathbf{h}}_{R_m,D}^{[i+1]}$ is given by

$$\hat{\mathbf{h}}_{R_m,D}^{[i+1]} = (\boldsymbol{\Gamma}_m^H \boldsymbol{\Gamma}_m)^{-1} \boldsymbol{\Gamma}_m^H \hat{\mathbf{P}}_{R_m,D}^{[i+1]H} \hat{\mathbf{E}}_{R_m,D}^{[i+1]H} \mathbf{y}_m^{[i]}, \tag{27}$$

where $\hat{\mathbf{E}}_{R_m,D}^{[i+1]} \triangleq \mathrm{diag}([e^{(j2\pi\hat{\epsilon}_{R_m,D}^{[i+1]}/N)\times 0}, e^{(j2\pi\hat{\epsilon}_{R_m,D}^{[i+1]}/N)}, \ldots, e^{(j2\pi\hat{\epsilon}_{R_m,D}^{[i+1]}/N)\times(N-1)}]^T)$, and $\hat{\epsilon}_{R_m,D}^{[i+1]}$ is obtained from (25), $\hat{\mathbf{P}}_{R_m,D}^{[i+1]} \triangleq \mathrm{diag}([e^{j\hat{\theta}_{R_m,D}^{[i+1]}(0)}, e^{j\hat{\theta}_{R_m,D}^{[i+1]}(1)}, \ldots, e^{j\hat{\theta}_{R_m,D}^{[i+1]}(N-1)}]^T)$, and $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i+1]} \triangleq [\hat{\theta}_{R_m,D}^{[i+1]}(0), \hat{\theta}_{R_m,D}^{[i+1]}(1), \ldots, \hat{\theta}_{R_m,D}^{[i+1]}(N-1)]^T$ are obtained from (20).

Using (20), (25), (27) and reapplying the above algorithm, for $m = \{1, \ldots, M\}$, estimates channel gains, multiple PHN, and CFO parameters for all the relays can be obtained at the destination. The iterations stop when the difference between LLFs of two iterations is smaller than a threshold ζ, i.e.

$$\left| \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D}^{[i+1]} \hat{\mathbf{P}}_{R_m,D}^{[i+1]} \boldsymbol{\Gamma}_m \hat{\mathbf{h}}_{R_m,D}^{[i+1]} \right\|^2 - \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D}^{[i]} \hat{\mathbf{P}}_{R_m,D}^{[i]} \boldsymbol{\Gamma}_m \hat{\mathbf{h}}_{R_m,D}^{[i]} \right\|^2 \right| \leq \zeta. \tag{28}$$

C. Initialization and Convergence

The appropriate initialization of CFOs and CIRs, i.e., $\hat{\epsilon}_{R,D}^{[0]} = [\hat{\epsilon}_{R_1,D}^{[0]}, \hat{\epsilon}_{R_2,D}^{[0]}, \ldots, \hat{\epsilon}_{R_M,D}^{[0]}]^T$ and $\hat{\mathbf{h}}_{R,D}^{[0]} = [\hat{\mathbf{h}}_{R_1,D}^{[0]}, \hat{\mathbf{h}}_{R_2,D}^{[0]}, \ldots, \hat{\mathbf{h}}_{R_M,D}^{[0]}]^T$ can help the proposed estimator to obtain the CIRs, PHNs, and CFOs parameters in a few iterations. The initialization process can be summarized as follows:

- The initial channel estimate, $\hat{\mathbf{h}}_{R,D}^{[0]}$, is obtained by $\hat{\mathbf{h}}_{R,D}^{[0]} = (\boldsymbol{\Psi}^{H[0]}\boldsymbol{\Psi}^{[0]})^{-1}\boldsymbol{\Psi}^{H[0]}\mathbf{r}$. Here, $\boldsymbol{\Psi}^{[0]} \triangleq \left[\boldsymbol{\Psi}_1^{[0]}, \boldsymbol{\Psi}_2^{[0]}, \ldots, \boldsymbol{\Psi}_M^{[0]}\right]$, $\boldsymbol{\Psi}_m^{[0]} \triangleq \hat{\mathbf{E}}_{R_m,D}^{[0]} \mathbf{F}^H \mathbf{D}_m \mathbf{F}_L$ and $\hat{\mathbf{E}}_{R_m,D}^{[0]} = \hat{\mathbf{E}}_{R_m,D}|_{\hat{\epsilon}_{R_m,D}=\hat{\epsilon}_{R_m,D}^{[0]}=0}$.
- The initial CFO estimate of mth relay is obtained by applying an exhaustive search for the value of $\epsilon_{R_m,D}$ that maximizes the function, $\mathbf{q}_m^H \boldsymbol{\Gamma}_m (\boldsymbol{\Gamma}_m^H \boldsymbol{\Gamma}_m)^{-1} \boldsymbol{\Gamma}_m^H \mathbf{q}_m$. Here, $\boldsymbol{\Gamma}_m \triangleq \hat{\mathbf{E}}_{R_m,D} \mathbf{F}^H \mathbf{D}_m \mathbf{F}_L$, and $\hat{\mathbf{q}}_m \triangleq \mathbf{r} - \sum_{\substack{\ell=1 \\ \ell\neq m}}^{M} \hat{\mathbf{E}}_{R_l,D} \mathbf{F}^H \mathbf{D}_m \mathbf{F}_L \hat{\mathbf{h}}_{R_l,D}$ with keeping the remaining parameters, i.e., $\hat{\epsilon}_{R_l,D}$ and $\hat{\mathbf{h}}_{R_l,D}$ for $\ell \neq m$ on the most recently updated values. Note that this exhaustive search needs to be only carried out at the system start up to initialize the estimation process. Simulations in Section VII indicate that an exhaustive search with a *coarse* step size of 10^{-2} is sufficient for the initialization of the proposed estimator. This coarse step size significantly reduces the overall complexity.
- Using $\hat{\epsilon}_{R_m,D}^{[0]}$ for $m = 1, \ldots, M$, the initial channel estimate, $\hat{\mathbf{h}}_{R,D}^{[0]}$, is obtained by $\hat{\mathbf{h}}_{R,D}^{[0]} = (\boldsymbol{\Psi}^{H[0]}\boldsymbol{\Psi}^{[0]})^{-1}\boldsymbol{\Psi}^{H[0]}\mathbf{r}$. Here, $\hat{\mathbf{E}}_{R_m,D}^{[0]} = \hat{\mathbf{E}}_{R_m,D}|_{\hat{\epsilon}_{R_m,D}=\hat{\epsilon}_{R_m,D}^{[0]}}$.

Simulation results in Section VI show that at SNR of 20 dB or higher the proposed ECM-based estimator always converges to the true estimates in only 2 iterations.

Remark 1: The convergence of proposed ECM algorithm to the global solution cannot be analytically shown [38]. However, in general, the ECM algorithm monotonically increase the LLF at every iteration and converge to a local maximum. Moreover, estimated parameters converge monotonically to the global solution, if the algorithm is initialized in a region suitably close to the global solution [38]. Based on the equivalent system model in (1) and the simulation results in Section VII, it can be concluded that the proposed ECM algorithm converges globally when the PHN vector $\hat{\boldsymbol{\theta}}_{R,D}$ is initialized as $\hat{\boldsymbol{\theta}}_{R,D}^{[0]} = [\underbrace{\mathbf{0}_{N-1\times 1}}_{m=1}, \underbrace{\mathbf{0}_{N-1\times 1}}_{2}, \ldots, \underbrace{\mathbf{0}_{N-1\times 1}}_{M}]^T$. Note that the initialization of PHN with zero vector because from initial training, the phase of the estimated channels incorporates the phase introduced from PHN.

V. Joint Data Detection and PHN Mitigation

In order to decode the received signal at $\mathbb{D}$ in the presence of PHNs and CFOs, an iterative detector based EKF for multi-relay cooperative systems is proposed.

At first, using the estimates of CIRs and CFOs, $\hat{\mathbf{h}}_{R,D}$, $\hat{\boldsymbol{\epsilon}}_{R,D}$, the received signal, $\mathbf{r}$ in (2), passes through an *iterative* algorithm of data detection and PHN mitigation. We propose

to use an EKF to track the PHN samples, $\boldsymbol{\theta}_{R_m,D}$, over the data symbols. The PHN estimation is similar to that in (16)-(21) and is not presented here to avoid repetition. However, instead of training-based PHN tracking, the process of PHN estimation is followed in decision-directed fashion for the received data symbols. In other words, the estimate of the data symbol in the previous iteration, $\hat{\bar{\mathbf{d}}}_m^{[i-1]}$, is used to update the symbol's PHN estimate at the current iteration $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i]}$. Particularly, $\mathbf{s}_m^{[i]}$ in (15), is calculated as $\mathbf{s}_m^{[i]} = \mathbf{F}^H \hat{\bar{\mathbf{D}}}_m^{[i]} \mathbf{F}_L \hat{\mathbf{h}}_{R_m,D}$, where $\hat{\mathbf{h}}_{R_m,D}$ is the CIR vector estimate obtained from the ECM estimator during the training interval, and $\hat{\bar{\mathbf{D}}}_m^{[i]} \triangleq \boldsymbol{\Lambda}_m \text{diag}(\hat{\bar{\mathbf{d}}}^{[i]})$. Next, the data vector estimate is updated for the ith iteration. Following [29] and based on the received signal in (2), the negative LLF for the received signal, $\mathbf{r}$, can be written as

$$\log p(\mathbf{r}, \hat{\mathbf{d}}, \hat{\boldsymbol{\theta}}_{R,D}) = \mathbf{C} + \frac{1}{2\sigma_w^2} \parallel \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D} \hat{\boldsymbol{\Upsilon}}_m \hat{\bar{\mathbf{d}}} \parallel^2 + \frac{1}{2\xi_d} \parallel \hat{\bar{\mathbf{d}}} \parallel^2 + \log p(\boldsymbol{\theta}_{R,D}), \tag{29}$$

where $\hat{\boldsymbol{\Upsilon}}_m \triangleq \mathbf{F}^H \hat{\mathbf{H}}_{R_m,D} \boldsymbol{\Lambda}_m$ is an $N \times N$ matrix, $\hat{\mathbf{H}}_{R_m,D} \triangleq \text{diag}(\mathbf{F}_L \hat{\mathbf{h}}_{R_m,D})$ is an $N \times N$ matrix of estimated channel frequency response, $\hat{\mathbf{E}}_{R_m,D} \triangleq \text{diag}([e^{(j2\pi\hat{\epsilon}_{R_m,D}/N)\times 0}, \ldots, e^{(j2\pi\hat{\epsilon}_{R_m,D}/N)\times(N-1)}]^T)$ is the $N \times N$ estimated CFO matrix of mth relay, $\hat{\mathbf{P}}_{R_m,D} \triangleq \text{diag}([e^{j\hat{\theta}_{R_m,D}(0)}, \ldots, e^{j\hat{\theta}_{R_m,D}(N-1)}]^T)$ is the $N \times N$ estimated PHN matrix, $\hat{\bar{\mathbf{d}}} \triangleq [\hat{\bar{d}}(0), \cdots, \hat{\bar{d}}(N-1)]^T$ is the $N \times N$ estimate vector of the modulated data vector, and ξ_d is the average transmitted symbol power and normalized to 1.

Taking the derivative of (29) with respect to $\bar{\mathbf{d}}$ and equating the result to zero, the estimate of $\hat{\bar{\mathbf{d}}}$ at the ith iteration, $\hat{\bar{\mathbf{d}}}^{[i]}$ is given by

$$\hat{\bar{\mathbf{d}}}^{[i]} = \left(\hat{\boldsymbol{\Omega}}^{H[i]} \hat{\boldsymbol{\Omega}}^{[i]} + \frac{\sigma_w^2}{\xi_d} \mathbf{I}_N \right)^{-1} \hat{\boldsymbol{\Omega}}^{H[i]} \mathbf{r}, \tag{30}$$

where $\hat{\boldsymbol{\Omega}}^{[i]} \triangleq \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D}^{[i]} \hat{\boldsymbol{\Upsilon}}_m$, $\hat{\mathbf{P}}_{R_m,D}^{[i]} \triangleq \text{diag}([e^{j\hat{\theta}_{R_m,D}^{[i]}(0)}, e^{j\hat{\theta}_{R_m,D}^{[i]}(1)}, \ldots, e^{j\hat{\theta}_{R_m,D}^{[i]}(N-1)}]^T)$, and $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i]} \triangleq [\hat{\theta}_{R_m,D}^{[i]}(0), \hat{\theta}_{R_m,D}^{[i]}(1), \ldots, \hat{\theta}_{R_m,D}^{[i]}(N-1)]^T$ are obtained via the EKF based estimator.

Using the EKF set of equations (16)-(21) and (30), the proposed algorithm iteratively updates the PHN and data estimates, respectively, and stops when the difference between likelihood functions of two iterations is smaller than a threshold ζ, i.e.,

$$\left| \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D}^{[i+1]} \hat{\boldsymbol{\Upsilon}}_m \hat{\bar{\mathbf{d}}}^{[i+1]} \right\|^2 - \sum_{n=0}^{N-1} \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D}^{[i]} \hat{\boldsymbol{\Upsilon}}_m \hat{\bar{\mathbf{d}}}^{[i]} \right\|^2 \right| \leq \zeta. \tag{31}$$

Let $\hat{\bar{\mathbf{d}}}^{[0]}$ denote the initial estimate of the transmitted data vector. Appropriate initialization of $\hat{\bar{\mathbf{d}}}^{[0]}$ results in the proposed iterative detector to converge quickly. In our algorithm, the initial data estimate is obtained using $\hat{\bar{\mathbf{d}}}^{[0]} = (\hat{\boldsymbol{\Omega}}^{H[j-1]} \hat{\boldsymbol{\Omega}}^{[j-1]} + \frac{\sigma_w^2}{\xi_d} \mathbf{I}_N)^{-1} \hat{\boldsymbol{\Omega}}^{H[j-1]} \mathbf{r}$, where $\hat{\boldsymbol{\Omega}}^{[j-1]} \triangleq \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D}^{[j-1]} \hat{\boldsymbol{\Upsilon}}_m$, and $\hat{\mathbf{P}}_{R,D}^{[j-1]}$ is the PHN matrix estimate obtained from the previous OFDM symbol. Simulation results in Section VII indicate that at SNR= 20 dB the proposed detector, on average, converges after 2 iterations. The overall estimation and detection algorithm is summarized in Algorithm 1.

It is worth mentioning that the proposed estimation and detection algorithms in this paper can be extended to a more complicated system setup such as multi-user MIMO-OFDM systems. If we assume that multiple antenna at each user are

Algorithm 1 Proposed Estimation and Detection Algorithms

ESTIMATION

$\hat{\theta}_{R_m,D}^{[0]}(1|0) = 0$ and $M_{R_m,D}^{[0]}(1|0) = \sigma_{\delta_{R_m,D}}^2$ and obtain $\hat{\epsilon}_{R,D}^{[0]}$ and $\hat{\mathbf{h}}_{R,D}^{[0]}$ using an exhaustive search with coarse step size i.e., 10^{-2}

while $\left| \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D}^{[i+1]} \hat{\mathbf{P}}_{R_m,D}^{[i+1]} \boldsymbol{\Gamma}_m \hat{\mathbf{h}}_{R_m,D}^{[i+1]} \right\|^2 - \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D}^{[i]} \hat{\mathbf{P}}_{R_m,D}^{[i]} \boldsymbol{\Gamma}_m \hat{\mathbf{h}}_{R_m,D}^{[i]} \right\|^2 \right| \leq \zeta$ **do**

  for $m = 1, \ldots, M$ **do**

    $\mathbf{y}_m^{[i]} = \mathbf{r} - \sum_{\substack{\ell=1 \\ \ell \neq m}}^{M} \hat{\mathbf{E}}_{R_\ell,D}^{[i]} \hat{\mathbf{P}}_{R_\ell,D}^{[i]} \mathbf{F}^H \mathbf{D}_\ell \mathbf{F}_L \hat{\mathbf{h}}_{R_\ell,D}^{[i]}$

    for $n = 0, 1, \ldots, N-1$ **do**

      $(17) - (21)$

    end for

    for $n = 0, 1, \ldots, N-1$ **do**

    $\hat{\epsilon}_{R_m,D}^{[i+1]} = \hat{\epsilon}_{R_m,D}^{[i]} - \frac{N}{2\pi} \frac{\sum_{n=0}^{N-1} n \Im\left\{ (\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N} \right\}}{\sum_{n=0}^{N-1} n^2 \Re\left\{ (\hat{y}_m^{[i]}(n))^* \hat{S}_m^{[i+1]}(n) e^{j2\pi\hat{\epsilon}_{R_m,D}^{[i]} n/N} \right\}}$

    end for

    $\hat{\mathbf{h}}_{R_m,D}^{[i+1]} = (\boldsymbol{\Gamma}_m^H \boldsymbol{\Gamma}_m)^{-1} \boldsymbol{\Gamma}_m^H \hat{\mathbf{P}}_{R_m,D}^{[i+1]H} \hat{\mathbf{E}}_{R_m,D}^{[i+1]H} \mathbf{y}_m^{[i]}$

    $\hat{\mathbf{h}}_{R_m,D}^{[i]} = \hat{\mathbf{h}}_{R_m,D}^{[i+1]}$, $\hat{\boldsymbol{\theta}}_{R_m,D}^{[i]} = \hat{\boldsymbol{\theta}}_{R_m,D}^{[i+1]}$, $\hat{\epsilon}_{R_m,D}^{[i]} = \hat{\epsilon}_{R_m,D}^{[i+1]}$

  end for

end while

DATA DETECTION

for $j = 1, \ldots, J$ **do**

  Obtain $\hat{\bar{\mathbf{d}}}^{[0]} = (\hat{\boldsymbol{\Omega}}^{H[j-1]} \hat{\boldsymbol{\Omega}}^{[j-1]} + \frac{\sigma_w^2}{\xi_d} \mathbf{I}_N)^{-1} \hat{\boldsymbol{\Omega}}^{H[j-1]} \mathbf{r}$, and Replace $\hat{\bar{\mathbf{d}}}^{[0]}$ by its hard decision.

  while $\left| \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D}^{[i+1]} \hat{\boldsymbol{\Upsilon}}_m \hat{\bar{\mathbf{d}}}^{[i+1]} \right\|^2 - \sum_{n=0}^{N-1} \left\| \mathbf{r} - \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D} \hat{\mathbf{P}}_{R_m,D}^{[i]} \hat{\boldsymbol{\Upsilon}}_m \hat{\bar{\mathbf{d}}}^{[i]} \right\|^2 \right| \leq \zeta$ **do**

    for $m = 1, \ldots, M$ **do**

      $\mathbf{y}_m^{[i]} = \mathbf{r} - \sum_{\substack{\ell=1 \\ \ell \neq m}}^{M} \hat{\mathbf{E}}_{R_\ell,D} \hat{\mathbf{P}}_{R_\ell,D}^{[i]} \hat{\boldsymbol{\Upsilon}}_m \hat{\bar{\mathbf{d}}}^{[i]}$

      Using the EKF set of equation in Section IV to estimate the PHN parameters,

    end for

    $\hat{\bar{\mathbf{d}}}^{[i+1]} = \left(\hat{\boldsymbol{\Omega}}^{H[i]} \hat{\boldsymbol{\Omega}}^{[i]} + \frac{\sigma_w^2}{\xi_d} \mathbf{I}_N \right)^{-1} \hat{\boldsymbol{\Omega}}^{H[i]} \mathbf{r}$, and Replace $\hat{\bar{\mathbf{d}}}^{[i+1]}$ by its hard decision.

    $\hat{\bar{\mathbf{d}}}^{[i]} = \hat{\bar{\mathbf{d}}}^{[i+1]}$

  end while

end for

fed by a single oscillator, as generally considered in MIMO setup [8]–[10], [27], the received signal at the BS, during uplink transmission, is affected by multiple PHN and CFO parameters. Therefore, the proposed algorithm in our paper can be easily modified and applied to a multi-user MIMO-OFDM setup, where BS can estimate multiple CFO and PHN parameters by employing our estimation and detection algorithms.

VI. Complexity Analysis

In this section, the computational complexities of the proposed ECM algorithm and iterative receiver for data detection in multi-relay cooperative systems are analyzed. Throughout this section, computational complexity is defined as the number of complex additions plus number of multiplications [5]. Let us denote the computational complexity, from $\mathbb{S} \rightarrow \mathbb{R} \rightarrow \mathbb{D}$, of the ECM algorithm by $C_{\text{EST}} = \underbrace{C_{\text{EST}}^{[M]}}_{\mathbb{S}\rightarrow\mathbb{R}} + \underbrace{C_{\text{EST}}^{[A]}}_{\mathbb{S}\rightarrow\mathbb{R}} + \underbrace{C_{\text{EST}}^{[M]}}_{\mathbb{R}\rightarrow\mathbb{D}} + \underbrace{C_{\text{EST}}^{[A]}}_{\mathbb{R}\rightarrow\mathbb{D}}$. The notations $C_{\text{EST}}^{[M]}$ and $C_{\text{EST}}^{[A]}$ are used to denote the number of complex multiplications and additions, respectively. Since the link from $\mathbb{S} \rightarrow \mathbb{R}_m$ is similar to a SISO system, then by following [35, eqs. (28) and (29)], the $\underbrace{C_{\text{EST}}^{[M]}}_{\mathbb{S}\rightarrow\mathbb{R}}$ and $\underbrace{C_{\text{EST}}^{[A]}}_{\mathbb{S}\rightarrow\mathbb{R}}$ of the ECM estimator from $\mathbb{S} \rightarrow \mathbb{R}$, are determined, respectively, and not presented here to avoid repetition. The $\underbrace{C_{\text{EST}}^{[M]}}_{\mathbb{R}\rightarrow\mathbb{D}}$ and $\underbrace{C_{\text{EST}}^{[A]}}_{\mathbb{R}\rightarrow\mathbb{D}}$ from $\mathbb{R} \rightarrow \mathbb{D}$, are determined as

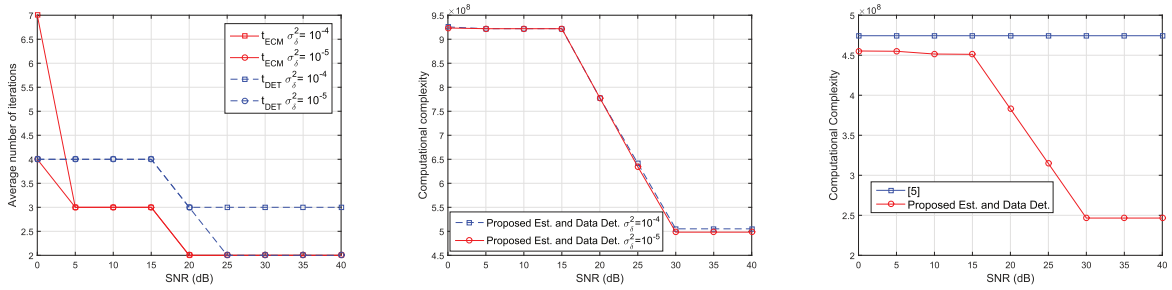
$$
\begin{aligned}
&\underbrace{C_{\text{EST}}^{[M]}}_{\mathbb{R}\rightarrow\mathbb{D}} \\
&= M\Bigg[\underbrace{(M-1)\big[N^2(3N+L)+NL\big]}_{(11)} + \big[\underbrace{N}_{(16)} + \underbrace{5N}_{(19)} \\
&\quad + \underbrace{2N}_{(20)} + \underbrace{2N}_{(21)} + \underbrace{7N}_{(25)} + \underbrace{LN(2N+1)}_{(27)} \\
&\quad + \underbrace{N(N^2+L(N+1))}_{\mathbf{s}_m \text{in}(15)}\big] t_{\text{ECM}} \\
&\quad + \underbrace{(M-1)\big[N(N^2+L(N+1))\big]}_{\hat{\mathbf{q}}_m \triangleq \bar{\mathbf{r}} - \sum_{\substack{\ell=1 \\ \ell\neq m}}^{M} \hat{\mathbf{E}}_{R_l,D}\mathbf{F}^H\bar{\mathbf{D}}_m\mathbf{F}_L\hat{\mathbf{h}}_{R_l,D}} \\
&\quad + \underbrace{M^2L^2(ML+N^2+N)+MLN}_{\hat{\mathbf{h}}_{R,D}^{[0]} = (\boldsymbol{\Psi}^{H[0]}\boldsymbol{\Psi}^{[0]})^{-1}\boldsymbol{\Psi}^{H[0]}\mathbf{r}} \\
&\quad + \big[\underbrace{L^3+L^2(N+1)+N(2L+1)}_{\mathbf{q}_m^H\boldsymbol{\Gamma}_m(\boldsymbol{\Gamma}_m^H\boldsymbol{\Gamma}_m)^{-1}\boldsymbol{\Gamma}_m^H\mathbf{q}_m} \\
&\quad + \underbrace{N(2N+L)}_{\boldsymbol{\Gamma}_m \triangleq \hat{\mathbf{E}}_{R_m,D}\mathbf{F}^H\bar{\mathbf{D}}_m\mathbf{F}_L}\big] t_{\text{initialize}}\Bigg]. \qquad (32)
\end{aligned}
$$

$$
\begin{aligned}
&\underbrace{C_{\text{EST}}^{[A]}}_{\mathbb{R}\rightarrow\mathbb{D}} \\
&= M\Bigg[\underbrace{(M-1)\big[N(N-1)(3N+L)+N(L-1)+N\big]}_{(11)} \\
&\quad + \big[\underbrace{N}_{(18)} + \underbrace{N}_{(19)} + \underbrace{2N}_{(20)} + \underbrace{N}_{(21)} + \underbrace{2N+1}_{(25)} + \underbrace{L(N-1)(2N+1)}_{(27)} \\
&\quad + \underbrace{N(N-1)(L+1)+N(L-1)}_{\mathbf{s}_m \text{in}(15)}\big] t_{\text{ECM}} \\
&\quad + \underbrace{(M-1)\big[N(N-1)(L+2N)+NL\big]}_{\hat{\mathbf{q}}_m \triangleq \bar{\mathbf{r}} - \sum_{\substack{\ell=1 \\ \ell\neq m}}^{M} \hat{\mathbf{E}}_{R_l,D}\mathbf{F}^H\bar{\mathbf{D}}_m\mathbf{F}_L\hat{\mathbf{h}}_{R_l,D}} \\
&\quad + \underbrace{M^3L^3+ML(N-1)(ML+1)+MNL(ML-1)}_{\hat{\mathbf{h}}_{R,D}^{[0]} = (\boldsymbol{\Psi}^{H[0]}\boldsymbol{\Psi}^{[0]})^{-1}\boldsymbol{\Psi}^{H[0]}\mathbf{r}} \\
&\quad + \big[\underbrace{L^3+(N-1)(L^2+L+1)+(L-1)(N+L)}_{\mathbf{q}_m^H\boldsymbol{\Gamma}_m(\boldsymbol{\Gamma}_m^H\boldsymbol{\Gamma}_m)^{-1}\boldsymbol{\Gamma}_m^H\mathbf{q}_m} \\
&\quad + \underbrace{N(N-1)(L+2)+N(L-1)}_{\boldsymbol{\Gamma}_m \triangleq \hat{\mathbf{E}}_{R_m,D}\mathbf{F}^H\bar{\mathbf{D}}_m\mathbf{F}_L}\big] t_{\text{initialize}}\Bigg]. \qquad (33)
\end{aligned}
$$

where t_{ECM} is the number of iterations required by the ECM algorithm and $t_{\text{initialize}}$ is the number of iterations need for coarse estimation step to obtain the initial estimates of the CFOs. Similarly, the computational complexity of the proposed EKF based data detection algorithm is denoted by $C_{\text{DET}} = \underbrace{C_{\text{DET}}^{[M]}}_{\mathbb{S}\rightarrow\mathbb{R}} + \underbrace{C_{\text{DET}}^{[A]}}_{\mathbb{S}\rightarrow\mathbb{R}} + \underbrace{C_{\text{DET}}^{[M]}}_{\mathbb{R}\rightarrow\mathbb{D}} + \underbrace{C_{\text{DET}}^{[A]}}_{\mathbb{R}\rightarrow\mathbb{D}}$, where $C_{\text{DET}}^{[M]}$ and $C_{\text{DET}}^{[A]}$ denote the number of complex multiplications and additions used by the detector. Following [35, eqs. (30) and (31)], $\underbrace{C_{\text{DET}}^{[M]}}_{\mathbb{S}\rightarrow\mathbb{R}}$ and $\underbrace{C_{\text{DET}}^{[A]}}_{\mathbb{S}\rightarrow\mathbb{R}}$ are determined and not presented here to avoid repetition. $\underbrace{C_{\text{DET}}^{[M]}}_{\mathbb{R}\rightarrow\mathbb{D}}$ and $\underbrace{C_{\text{DET}}^{[A]}}_{\mathbb{R}\rightarrow\mathbb{D}}$ are determined as

$$
\begin{aligned}
&\underbrace{C_{\text{DET}}^{[M]}}_{\mathbb{R}\rightarrow\mathbb{D}} \\
&= M\Bigg[\underbrace{(M-1)\big[N^2(3N+L)+NL\big]}_{(11)} + \big[\underbrace{N}_{(16)} + \underbrace{5N}_{(19)} \\
&\quad + \underbrace{2N}_{(20)} + \underbrace{2N}_{(21)} + \underbrace{N(N^2+(2L-1)(N+1))}_{\mathbf{s}_m \text{in}(15)} \\
&\quad + \underbrace{N^2(3N+1)+N}_{(30)} + \underbrace{4N^3}_{\hat{\boldsymbol{\Omega}}^{[i]} = \sum_{m=1}^{M} \hat{\mathbf{E}}_{R_m,D}\hat{\mathbf{P}}_{R_m,D}^{[i]}\hat{\boldsymbol{\Upsilon}}_m} \\
&\quad + \underbrace{N(2L-1)}_{\hat{\boldsymbol{\Upsilon}}_m = \mathbf{F}^H\hat{\mathbf{H}}_{R_m,D}\boldsymbol{\Lambda}_m}\big] t_{\text{DET}}\Bigg] \\
&\quad + \underbrace{4N^3+N^2(3N+1)+N}_{\hat{\mathbf{d}}^{[0]} = (\hat{\boldsymbol{\Omega}}^{H[j-1]}\hat{\boldsymbol{\Omega}}^{[j-1]} + \frac{\sigma_w^2}{\varsigma_d}\mathbf{I}_N)^{-1}\hat{\boldsymbol{\Omega}}^{H[j-1]}\bar{\mathbf{r}}}, \qquad (34)
\end{aligned}
$$

$$
\begin{aligned}
&\underbrace{C_{\text{DET}}^{[A]}}_{\mathbb{R}\rightarrow\mathbb{D}} \\
&= M\Bigg[\underbrace{(M-1)\big[N(N-1)(3N+L)+N(L-1)+N\big]}_{(11)} \\
&\quad + \big[\underbrace{N}_{(18)} + \underbrace{N}_{(19)} + \underbrace{2N}_{(20)} + \underbrace{N}_{(21)}
\end{aligned}
$$



(a) Average number of iterations **(b)** The computational complexity **(c)** Comparison with [5]

Fig. 3. Average number of iterations and the computational complexity of proposed algorithms with the comparison with the approach in [5] for phase noise variance $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2, $L = 4$, $M = 2$ relays, and 16-QAM modulation.

$$+ \underbrace{N(N-1)(L+1)+N(L-1)}_{\mathbf{s}_m \text{ in (15)}} + \underbrace{(N-1)(N^2+N)}_{(30)} + \underbrace{4N^2(N-1)}_{\hat{\mathbf{\Omega}}^{[i]}=\sum_{m=1}^{M}\hat{\mathbf{E}}_{R_m,D}\hat{\mathbf{P}}_{R_m,D}^{[i]}\hat{\mathbf{\Upsilon}}_m} + \underbrace{N(L-1)}_{\hat{\mathbf{\Upsilon}}_m=\mathbf{F}^H\hat{\mathbf{H}}_{R_m,D}\mathbf{\Lambda}_m} \Big] t_{\text{DET}} + \underbrace{4N^2(N-1)+(N-1)(2N^2+N)}_{\hat{\tilde{\mathbf{d}}}^{[0]}=(\hat{\mathbf{\Omega}}^{H[j-1]}\hat{\mathbf{\Omega}}^{[j-1]}+\frac{\sigma_w^2}{\xi_d}\mathbf{I}_N)^{-1}\hat{\mathbf{\Omega}}^{H[j-1]}\bar{\mathbf{r}}}. \quad (35)$$

where t_{DET} is the number of iterations required by the proposed data detection algorithm. It can be observed from the results in Fig. 3-(a) that: (i) at low SNR, i.e., SNR $<$ 20 dB, on average, the proposed detector converges after t_{DET} more than 2 iterations, (ii) the number of iterations decreases to $t_{\text{ECM}} = t_{\text{DET}} = 2$ at SNR $\geq$ 20 dB, and (iv) the proposed ECM algorithm converges to the true estimates when the CFO estimates are initialized with a step size of 10^{-2}, i.e., $t_{\text{initialize}} = 10^2$. Using these values for the number of iterations, we get the computational complexity of the proposed algorithms for multi-relay networks with $M = 2$ relays as shown in Fig. 3-(b). The results in Fig. 3-(b) show that (i) at low SNR, i.e., SNR $<$ 20 dB, the computational complexity of the proposed algorithms dependent on the variance of the PHN process, since at low SNR the performance of the performance of the proposed estimator and detector is dominated by AWGN and PHN variance, while at moderate-to-high SNR, i.e., SNR $>$ 20 dB the performance of the system is limited by residual PHN and CFO, (ii) at moderate-to-high SNR compared to low SNR, the proposed estimation and data detection algorithms are computationally more efficient. These results are anticipated, since the proposed estimation and data detection algorithms require few iterations at moderate-to-high SNRs as shown in Fig. 3-(a). In Fig. 3-(c), we compare the proposed algorithms with the one in [5]. It is worth noting in [5] performs the estimation and detection using a single relay. Hence, for fairness, we compare the estimation and detection by using a single relay, i.e., $M = 1$, and the that of [5]. We observe from Fig. 3-(c) that for different SNRs, the computational complexity of the proposed algorithms outperforms [5], which maybe of practical interest for multi-relay applications with stringent performance requirements.

VII. Simulation Results and Discussions

In this section, we present simulation results to evaluate the performance of the proposed estimation and data detection algorithms. A multipath Rayleigh fading channel with a delay of $L = 4$ taps and an exponentially decaying power delay profile is assumed between each pair of nodes. A training symbol size of $N = 64$ subcarriers is used, where each subcarrier is modulated using quadrature phase-shift keying (QPSK) scheme. The Wiener PHN is generated in each node with different PHN variances, e.g. $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$rad^2. Note that, $\sigma_\delta^2 = 10^{-4}$ rad^2, corresponds to a high phase noise variance [22]. Since carrier frequency offsets from source to relays, ϵ_{S,R_m}, are carried over to the destination, ϵ_{S,R_m} and $\epsilon_{R_m,D}$ have the range $(-0.25,0.25)$ in order to limit the total frequency offset from source to destination, $\epsilon_{S,D}$ to the range $(-0.5, 0.5)$. The data symbols are drawn from normalized 4, 16, or 64 quadrature amplitude modulation (QAM). The simulation results are averaged over 1×10^5 Monte Carlo simulation runs. Finally, the *mean-square error* (MSE) performance of ECM estimator and the *bit error rate* (BER) performance of the overall multi-relay network.

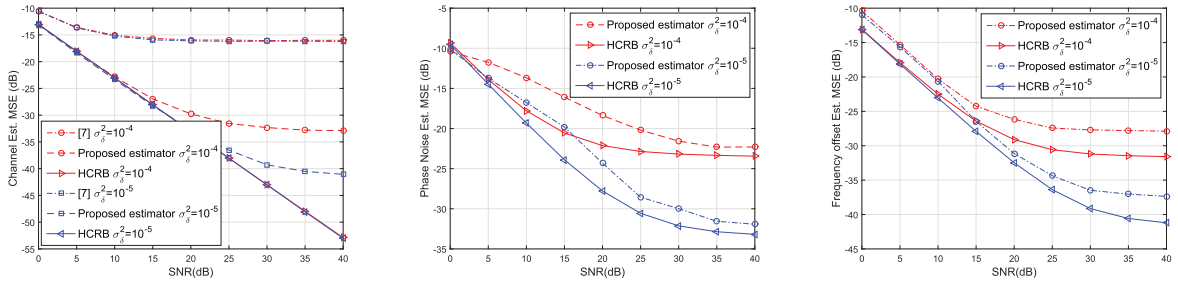
A. Estimation Performance

In this subsection, we compare the performance of the proposed ECM estimator with the HCRB in Theorem 1 and the estimation approach based on MMSE-optimal training sequences in [7]. Fig. 4 plots the HCRB and MSE for estimating the CIR, PHN, and CFO, respectively, using the proposed algorithm.

The results lead to the following observations:

1) The HCRB and the proposed estimatorÂ's MSE are dependent on the variance of the PHN process and are lower for a lower PHN variance;

2) The results in Fig. 4 show that CIRs, CFOs and PHNs estimation performances suffer from an error floor, which is directly related to the variance of the PHN process. This follows from the fact that at low SNR the performance of the system is dominated by AWGN, while at high SNR the performance of the proposed estimator is limited by PHN and the resulting ICI;



(a) MSE of channel estimation compared to the estimation approach in [7]

(b) MSE of phase noise estimation

(c) MSE of frequency offset estimation

Fig. 4. MSE of CIRs, PHNs, and CFOs estimation for the proposed estimator compared to HCRB for phase noise variance $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 with M =2.

3) Fig. 4-(a) shows at different SNRs, the proposed estimator significantly outperforms the estimator in [7]. This result is anticipated, since the orthogonality of the training sequences proposed in [7] could not be achieved in the presence of PHN and CFO. Therefore, the estimation approach in [7] may not be used in the presence of PHN and CFO;

4) The results in Fig. 4 show that the MSEs of the proposed estimator are close to their HCRLBs at moderate-to-high SNRs.

Note that since the PHN vector is initialized with zeros, the MSE of phase noise estimation of first sample represents the MSE of channel phase estimation.

B. Impact of PHN on the Cooperative Performance

In the following, we examine the combined estimation and data detection performance in terms of the BER. The following system setups are considered for comparison:

(i) Cooperative systems that applied the proposed estimation and data detection algorithms (labelled as "Proposed Est. and Data Det.").

(ii) The data detection based on pilots in [5] and [22] (labelled as "Data Det. based Pilots [5,22]").

(iii) As a lower-bound on the BER performance, a system assuming perfect channels, PHNs, and CFOs estimation (labelled "Perf. CIRs, PHNs & CFOs est.").

Note that the BER performance of proposed algorithm is compared with that one in [5] and [22] since the detection approach based on pilots in [5] and [22] could be used to mitigate the PHNs by estimating the CPE which is similar to all subcarriers. Moreover, no BER compersion with other basic relevant works is presented in the paper since the system model in our paper considers multiple PHN and CFO parameters estimation and the existing system models only consider a single PHN and CFO parameter estimation.

Fig. 5 shows the BER performance with $M = 2$ for PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 and 16-QAM modulation. The following observations can be made from Fig. 5:

1) The BER performance using the proposed algorithm significantly outperforms the existing data detection based on pilots in [5] and [22] at different SNRs. This result is anticipated, since the detection method in [5] and [22] depends on the pilots, which are effected by the ICI and

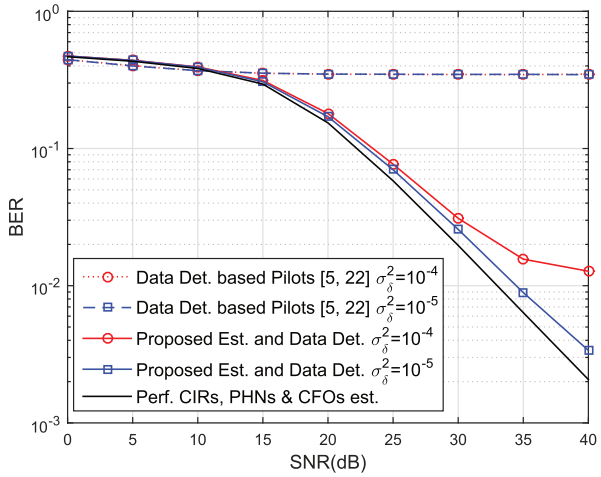


Fig. 5. BER performance for PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 and 16-QAM modulation with M =2.

the interference signals between the antennas. Therefore, the detection approach in [5] and [22] maybe only used for cooperative systems with a single relay as in [5] or for multi-relay systems based on TDMA transmission as in [22]. Thus, the pilot approach for data detection in [5], [22] may not be used for the joint estimation and data detection for multi-relay systems based on SDMA transmission.

2) Compared to an ideal case of perfect CIR, PHN and CFO estimation, the BER performance using the proposed algorithms is close to ideal case of perfect CIR, PHN and CFO estimation when $\sigma_\delta^2 = 10^{-5}$ rad^2. However, at high PHN variance, i.e., $\sigma_\delta^2 = 10^{-4}$ rad^2, the BER performance suffer from an error floor at high SNR. This result is anticipated, since at high PHN variance, the performance of a cooperative OFDM system is dominated by PHN, which cannot be completely eliminated.

C. Impact of Increase Number Relays on Cooperative Performance

In this subsection, we examine the performance of the proposed algorithms compared to "Perf. CIRs, PHNs & CFOs est." performance with the increase of number of relays and subcarriers in the multi-relay network. Moreover, we compare

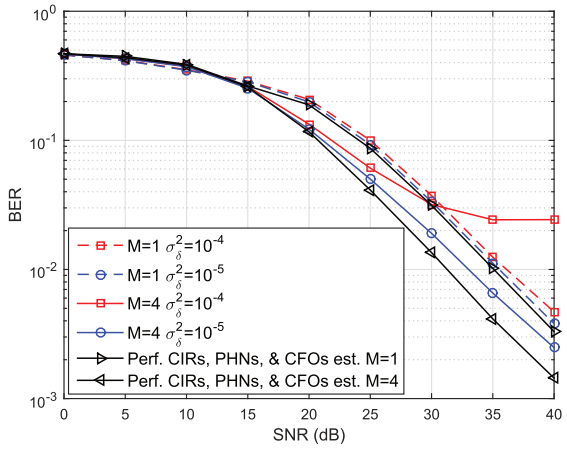


Fig. 6. BER performance at different number of relays, $M = [1, 4]$ for PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$rad^2 and 16-QAM modulation.

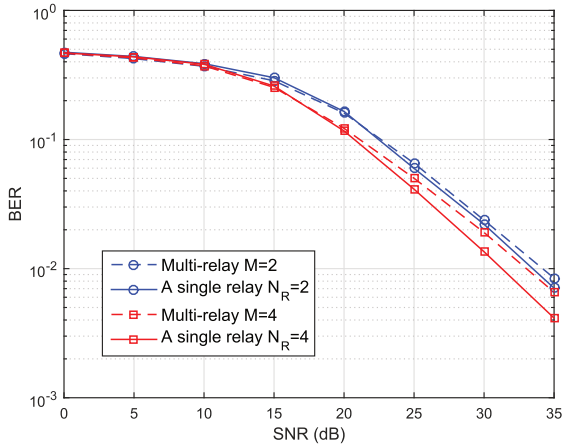


Fig. 7. BER performance of the multi-relay systems and a single relay equipped with multi-antennas for PHN variance $\sigma_\delta^2 = 10^{-5}$ rad^2 and 16-QAM modulation.

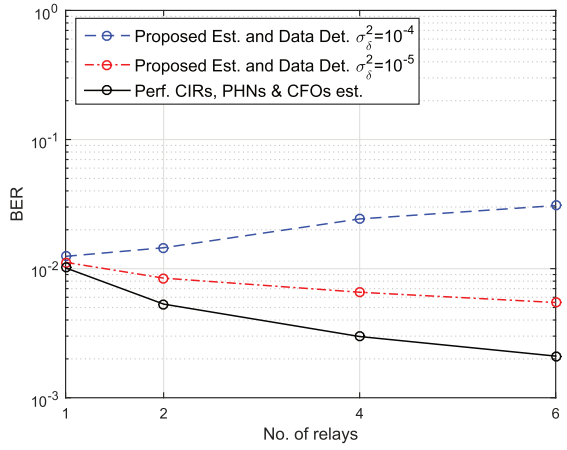


Fig. 8. BER performance at different number of relays at SNR = 35 dB for PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 and 16-QAM modulation.

the performance of multi-relay systems with a single relay equipped with mutiple antennas (N_R).

Figs. 6, 7, 8, and 9 respectively show (i) the BER performance at different number of relays, $M = [1, 2, 4]$,

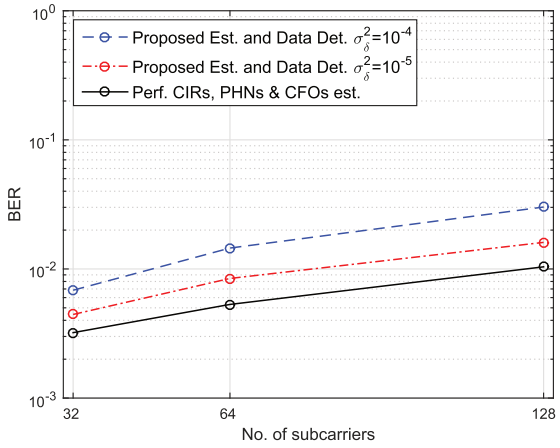


Fig. 9. BER performance at different number of subcarriers at SNR = 35 dB for $M = 2$, PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 and 16-QAM modulation.

(ii) the BER performance of the multi-relay systems and a single relay equipped with multiple antennas (N_R), (iii) different number of relays, $M = [1, 2, 4, 6]$, and $SNR = 35$ dB, and (iv) different number of subcarriers, $N = [64, 128, 256]$, and $SNR = 35$ dB, for PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 and 16-QAM modulation. The following observations can be made from Figs. 6, 7, 8, and 9:

1) The results in Figs. 6 and 8 show that at low PHN variance, i.e., $\sigma_\delta^2 = 10^{-5}$ rad^2, the multi-relay systems using the combination of the proposed estimation and data detection algorithms outperforms a single relay system. More importantly, the BER performance improves as the number of relays increases. For instance, at BER = 10^{-2}, the SNR gain for the multirelay systems is almost 3 dB compared to the performance of a single relay system. In addition, the performance of multi-relay is closer to the ideal case of perfect CIRs, PHNs, and CFOs estimation. For example, in Fig. 6, a performance gap of 1.6 dB at BER = 10^{-2}.

2) In Fig. 6, at high PHN variance, i.e., $\sigma_\delta^2 = 10^{-4}$ rad^2, the BER performance degrades with increase of number of relays. This result is anticipated, since at high PHN variance, the proposed ECM estimator demonstrate poor performance due to the considerable residual PHN and CFO estimation error from source to relays, which is forwarded to the destination. Therefore, in the presence of high PHN variance, the cooperative system can achieve significant BER performance by combining the proposed estimation and data detection algorithms and using few relays. However, this approach maintains higher BER performance at the expense of a loss in the range of network links which could be enhanced by using multi-relay to overcome the blockage issues in some communication systems such as mmWave systems [42]. Therefore, in the presence of PHN and CFO, the increasing of number of relays could enhance the network range at the expense of degradation in the BER performance.

3) We can see from Figs. 6 and 7 that a single relay network outperforms a multi-relay diversity network in the presence of strong phase noise variance, which implies that relay selection may outperform the multi-relay diversity network in the presence of strong phase noise variance. However, as

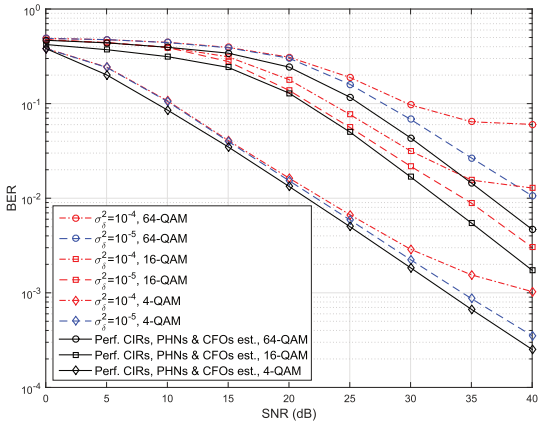


Fig. 10. BER performance for a multi-relay system at different modulations, 4-QAM, 16-QAM, and 64-QAM, PHN variance, $\sigma_\delta^2 = [10^{-4}, 10^{-5}]$ rad^2 and M =2.

clearly mentioned to motivate the importance of our proposed algorithm, if a relay is badly blocked due to shadowing, our distributed system is the best option to achieve the relay diversity compared to a single relay since the other relays can still make a link to the receiver. Note that blockage is not such a major issue in 4G where omni directional antennas are used for the user and access point terminals. However, as we migrate to 5G and beyond and use more directional antennas blockage can more easily happen between the channel estimation instances.

4) The results in Fig. 8 show that at moderate PHN variance, i.e., $\sigma_\delta^2 = 10^{-5}$ rad^2, a multi-relay system has better BER performance than a single relay system. As per expectation from general multi-relay system which assumes perfect CFO, PHN, and CHN estimation, the diversity gain is achieved by adding relays. Our particular contribution is that even in the presence of multiple impairments and moderate PHN variance (10^{-5} rad^2), the application of our proposed estimation and detection algorithms succeed to achieve the diversity gain. However, at high PHN variance, i.e., $\sigma_\delta^2 = 10^{-4}$ rad^2, the BER performance deteriorates as we increase the number of relays. This implies that presence of strong phase noise hinders the achievement of multi-relay diversity gain. This result leads to a very important research opening as millimeter wave communication technology in 5G expects strong phase noise due to high frequency transmission. Our findings show that millimeter wave communication cannot easily enjoy the multi-relay diversity gain unless very sophisticated phase noise estimation and tracking algorithm is employed.

5) Fig. 9 shows that BER performance deteriorates by increasing the number of subcarriers due to decrease in the effective subcarrier spacing.

D. Impact of Modulation on Cooperative Performance

Fig. 10 evaluates the BER performance of the multi-relay system at different modulations, i.e., 4-QAM, 16-QAM and 64-QAM. The following observations can be made from Fig. 10:

1) Even for denser constellation, i.e., 64-QAM, the proposed estimation and data detection algorithms achieve BER performance that is closer to the ideal case of perfect CIRs, PHNs, and CFOs estimation. For example, as shown in Fig. 10, at BER of 10^{-2} and a PHN variance of 10^{-5} rad^2 and 64-QAM, the performance of multi-relay system is close with 2.5 dB to the ideal case of perfect CIRs, PHNs, and CFOs estimation.

2) The results in Fig. 10 of 64-QAM modulation in the presence of strong PHN variance of 10^{-4} rad^2 show that even at high SNRs where the channel noise could be neglected, the overall BER performance of a multi-relay system suffers from an error floor which is higher than 10^{-2}. This is because the subcarriers in 64-QAM become closely spaced and more sensitive to the noise caused by the residual PHN. Meanwhile, at high PHN, i.e., $\sigma_\delta^2 = 10^{-4}$, the application of 16-QAM modulation still yield high BER since the BER is affected by the ICI from the residual PHN which cannot be completely eliminated.

3) The proposed algorithm achieves an overall BER performance lower than 10^{-2} at SNR $>$ 20 dB if the modulation order is reduced to 4-QAM. This is anticipated since 4-QAM has lower sensitivity to the noise caused by the channel and ICI.

VIII. Conclusion

In this paper, we address the joint estimation of unknown multiple channel, phase noise (PHN), and carrier frequency offset (CFO) parameters for DF-relaying cooperative OFDM systems. A new iterative estimator is proposed and found to be computationally efficient since it estimate the desired parameters in few iterations. Simulation results show that the performance of the proposed estimator is close to the derived HCRB at different signal-to-noise ratios (SNRs). Next, an *iterative* algorithm for joint data detection and PHN mitigation is proposed for the OFDM data symbols. The proposed algorithm employs an EKF based approach to track the time-varying PHN parameters throughout the OFDM data symbols. Numerical results show that the combination of proposed ECM based estimator and the iterative data detection algorithm can enhances the performance of cooperative systems to be closer to the ideal case of perfect CIRs, PHNs and CFOs estimation in terms of BER. The performance analysis for the multi-relay OFDM system in the presence of multiple PHN, CFO and channel estimation is an open future research problem. It is observed that in the presence of high PHN varaiance, relay selection scheme may provide better diversity gains. However, addressing the specific relay selection problem in the presence of multiple impairments is outside the scope of this paper and can be considered for future work.

Appendix
Derivation of the HCRB

The hybrid information matrix $\mathbf{B}$ can be written as [36, pp. 1–85]

$$\mathbf{B} = \mathbf{\Xi}_D + \mathbf{\Xi}_P, \tag{A.1}$$

where $\mathbf{\Xi}_D \triangleq \mathbb{E}_{\boldsymbol{\theta}}\left[\boldsymbol{\Psi}(\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D})\right]$ with $\boldsymbol{\Psi}(\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D}) \triangleq$

$$\frac{\partial \boldsymbol{\mu}^H}{\partial \boldsymbol{\lambda}_i} = \begin{cases} \left[\tilde{\boldsymbol{\Phi}}_1, \ldots, \tilde{\boldsymbol{\Phi}}_M\right]^H, \text{ is an } MN \times N \text{ matrix}, & (\boldsymbol{\lambda}_i = \boldsymbol{\theta}_{R,D}) \\ \left[\tilde{\boldsymbol{\gamma}}_1, \ldots, \tilde{\boldsymbol{\gamma}}_M\right]^H, \text{ is an } ML \times N \text{ matrix}, & (\boldsymbol{\lambda}_i = \Re\{\mathbf{h}_{R,D}\}) \\ j\left[\tilde{\boldsymbol{\gamma}}_1, \ldots, \tilde{\boldsymbol{\gamma}}_M\right]^H, \text{ is an } ML \times N \text{ matrix}, & (\boldsymbol{\lambda}_i = \Im\{\mathbf{h}_{R,D}\}) \\ \left[\tilde{\boldsymbol{\beta}}_1, \ldots, \tilde{\boldsymbol{\beta}}_M\right]^H, \text{ is an } M \times N \text{ matrix}, & (\boldsymbol{\lambda}_i = \boldsymbol{\epsilon}_{R,D}) \end{cases} \tag{A.4}$$

$$\begin{aligned} \boldsymbol{\Xi}_P &= \mathbb{E}_{\boldsymbol{\theta}|\mathbf{h},\epsilon}\left[-\Delta^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}} \log p(\boldsymbol{\theta}|\mathbf{h},\epsilon)|\epsilon\right] \triangleq \begin{bmatrix} \boldsymbol{\Xi}_{P_{11}} & \boldsymbol{\Xi}_{P_{12}} & \boldsymbol{\Xi}_{P_{13}} & \boldsymbol{\xi}_{P_{14}} \\ \boldsymbol{\Xi}_{P_{21}} & \boldsymbol{\Xi}_{P_{22}} & \boldsymbol{\Xi}_{P_{23}} & \boldsymbol{\xi}_{P_{24}} \\ \boldsymbol{\Xi}_{P_{31}} & \boldsymbol{\Xi}_{P_{32}} & \boldsymbol{\Xi}_{P_{33}} & \boldsymbol{\xi}_{P_{34}} \\ \boldsymbol{\xi}_{P_{41}} & \boldsymbol{\xi}_{P_{42}} & \boldsymbol{\xi}_{P_{43}} & \boldsymbol{\xi}_{P_{44}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\boldsymbol{\theta}}_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Re\{\mathbf{h}\}}_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Im\{\mathbf{h}\}}_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\epsilon}_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})\right] \\ \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\boldsymbol{\theta}}_{\Re\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Re\{\mathbf{h}\}}_{\Re\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Im\{\mathbf{h}\}}_{\Re\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\epsilon}_{\Re\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] \\ \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\boldsymbol{\theta}}_{\Im\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Re\{\mathbf{h}\}}_{\Im\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Im\{\mathbf{h}\}}_{\Im\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\epsilon}_{\Im\{\mathbf{h}\}} \log p(\boldsymbol{\theta})\right] \\ \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\boldsymbol{\theta}}_{\epsilon} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Re\{\mathbf{h}\}}_{\epsilon} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\Im\{\mathbf{h}\}}_{\epsilon} \log p(\boldsymbol{\theta})\right] & \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\epsilon}_{\epsilon} \log p(\boldsymbol{\theta})\right] \end{bmatrix}, \end{aligned} \tag{A.6}$$

$\mathbb{E}_{\mathbf{r}|\boldsymbol{\theta}_{R,D},\Re\{\mathbf{h}_{R,D}\},\Im\{\mathbf{h}_{R,D}\},\boldsymbol{\epsilon}_{R,D}}\big[- \Delta^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}} \log p(\mathbf{r}|\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D})|, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D}\big]$ denoting the Fisher's information matrix (FIM) and $\boldsymbol{\Xi}_P \triangleq \mathbb{E}_{\boldsymbol{\theta}|,\Re\{\mathbf{h}_{R,D}\},\Im\{\mathbf{h}_{R,D}\},\boldsymbol{\epsilon}_{R,D}}\big[- \Delta^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}} \log p(\boldsymbol{\theta}_{R,D}|, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D})|, \boldsymbol{\epsilon}_{R,D}\big]$ is the prior information matrix with $p(\boldsymbol{\theta}_{R,D}|\mathbf{h}_{R,D}, \boldsymbol{\epsilon}_{R,D})$ denoting the prior distribution of PHN vector given the CIR and CFO. Thus, we first obtain expressions for matrices $\boldsymbol{\Xi}_D$ and $\boldsymbol{\Xi}_P$.

A. Computation of $\boldsymbol{\Xi}_D \triangleq \mathbb{E}_{\boldsymbol{\theta}}\big[\boldsymbol{\Psi}(\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D})\big]$

To compute FIM, first, the likelihood function $p(\mathbf{r}|\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D})$ is given by

$$\begin{aligned} &p(\mathbf{r}|\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \quad \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D}) \\ &\quad = \mathrm{C} \exp\left[\frac{-1}{\sigma_w^2}(\mathbf{r}-\boldsymbol{\mu})^H(\mathbf{r}-\boldsymbol{\mu})\right], \end{aligned} \tag{A.2}$$

where $\mathrm{C} \triangleq (\pi\sigma_w^2)^{-N}$. Given $\boldsymbol{\theta}_{R,D}$, $\mathbf{h}_{R,D}$, and $\boldsymbol{\epsilon}_{R,D}$, $\mathbf{r}$ is a complex Gaussian vector with mean vector $\boldsymbol{\mu} = \sum_{m=1}^{M} \mathbf{E}_{R_m,D}\mathbf{P}_{R_m,D}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_m,D}$ and covariance matrix $\sigma_w^2\mathbf{I}_N$. The FIM, $\boldsymbol{\Psi}(\boldsymbol{\theta}_{R,D}, \Re\{\mathbf{h}_{R,D}\}, \Im\{\mathbf{h}_{R,D}\}, \boldsymbol{\epsilon}_{R,D})$, will be $M(N+2L) \times M(N+2L)$ matrix for joint estimation of $M(N-1)$ PHNs parameters $\boldsymbol{\theta}_{R,D}$, $2ML$ channels parameters $\Re\{\mathbf{h}_{R,D}\}$ and $\Im\{\mathbf{h}_{R,D}\}$ and M CFOs parameters $\boldsymbol{\epsilon}_{R,D}$. Using (A.2), the (i,j)th entry of $\boldsymbol{\Psi}$ can be written as [41]

$$[\boldsymbol{\Psi}]_{i,j} = \frac{2}{\sigma_w^2}\Re\left\{\frac{\partial \boldsymbol{\mu}^H}{\partial \boldsymbol{\lambda}_i}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\lambda}_j}\right\}, \tag{A.3}$$

where $\tilde{\boldsymbol{\Phi}}_m = \mathrm{diag}(\mathbf{E}_{R_mD}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_mD})\mathbf{a}_i$, $\tilde{\boldsymbol{\gamma}}_m = \mathbf{E}_{R_mD}\mathbf{P}_{R_mD}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{e}_l$, $\tilde{\boldsymbol{\beta}}_m = \acute{\mathbf{E}}_{RmD}\mathbf{P}_{R_mD}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{RmD}$, $\mathbf{a}_i = [0, 0_{1\times i-1}, je^{j\theta_{R,D_i}}, 0_{1\times N-i}]^T$ for $i = 1, \ldots, N-1$, $\mathbf{e}_l = [0_{1\times l-1}, 1, 0_{1\times L-1}]^T$ for $l = 1, \ldots, L$, $\acute{\mathbf{E}}_{RmD} \triangleq \mathrm{diag}([0, \frac{j2\pi}{N}e^{(j2\pi\epsilon_{RmD}/N)}, \ldots, \frac{j2\pi(N-1)}{N}e^{(j2\pi\epsilon_{RmD}/N)\times(N-1)}]^T)$, and $m = 1, \ldots, M$.

Substituting (A.4), as shown at the top of this page, into (A.5), and calculating the explicit expectation over $\boldsymbol{\theta}_{R,D}$, the matrix $\boldsymbol{\Xi}_D$ is obtained as

$$\begin{aligned} &\boldsymbol{\Xi}_D \\ &= \frac{2}{\sigma_w^2}\Re\left\{\begin{bmatrix} \bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_1 & -j\bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_2 & \bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_2 & \bar{\mathbf{Q}}_1^H\bar{\mathbf{Q}}_3 \\ j\bar{\mathbf{Q}}_2^H\bar{\mathbf{Q}}_1 & \mathbf{Q}_2^H\mathbf{Q}_2 & j\mathbf{Q}_2^H\mathbf{Q}_2 & j\mathbf{Q}_2^H\mathbf{Q}_3 \\ \bar{\mathbf{Q}}_2^H\bar{\mathbf{Q}}_1 & -j\mathbf{Q}_2^H\mathbf{Q}_2 & \mathbf{Q}_2^H\mathbf{Q}_2 & \mathbf{Q}_2^H\mathbf{q}_5 \\ \bar{\mathbf{Q}}_3^H\bar{\mathbf{Q}}_1 & -j\bar{\mathbf{Q}}_3^H\mathbf{Q}_2 & \mathbf{Q}_3^H\mathbf{Q}_2 & \mathbf{Q}_3^H\mathbf{Q}_3 \end{bmatrix}\right\}, \end{aligned} \tag{A.5}$$

where $\bar{\mathbf{Q}}_1 = \left[\bar{\boldsymbol{\Phi}}_1, \ldots, \bar{\boldsymbol{\Phi}}_M\right]$, $\bar{\boldsymbol{\Phi}}_m = \boldsymbol{\Phi}_m(2:N, 2:N)$, $\boldsymbol{\Phi}_m = j\mathrm{diag}(\mathbf{E}_{R_mD}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{R_mD})$, $\mathbf{Q}_2 = \left[\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_M\right]$, $\boldsymbol{\gamma}_m = \mathbf{E}_{R_mD}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L$, $\bar{\mathbf{Q}}_2 = \left[\bar{\boldsymbol{\gamma}}_1, \ldots, \bar{\boldsymbol{\gamma}}_M\right]$, $\bar{\boldsymbol{\gamma}}_m = \boldsymbol{\gamma}_m(2:N, 1:L)$, $\mathbf{Q}_5 = \left[\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_M\right]$, $\boldsymbol{\beta}_m = \acute{\mathbf{E}}_{RmD}\mathbf{F}^H\mathbf{D}_m\mathbf{F}_L\mathbf{h}_{RmD}$, $\bar{\mathbf{Q}}_5 = \left[\bar{\boldsymbol{\beta}}_1, \ldots, \bar{\boldsymbol{\beta}}_M\right]$, $\bar{\boldsymbol{\beta}}_m = \boldsymbol{\beta}_m(2:N)$, $\acute{\mathbf{E}}_{RmD} \triangleq \mathrm{diag}([0, \frac{j2\pi}{N}e^{(j2\pi\epsilon_{RmD}/N)}, \ldots, \frac{j2\pi(N-1)}{N}e^{(j2\pi\epsilon_{RmD}/N)\times(N-1)}]^T)$, and $m = 1, \ldots, M$.

B. Computation of $\boldsymbol{\Xi}_P \triangleq \mathbb{E}_{\boldsymbol{\theta}|\mathbf{h}_{R,D},\boldsymbol{\epsilon}_{R,D}}\left[-\Delta^{\boldsymbol{\lambda}}_{\boldsymbol{\lambda}} \log p(\boldsymbol{\theta}_{R,D}|\mathbf{h}_{R,D}, \boldsymbol{\epsilon}_{R,D})|\mathbf{h}_{R,D}, \boldsymbol{\epsilon}_{R,D}\right]$

The second factor in HIM, defined in (A.1), can be written as: where $p(\boldsymbol{\theta})$ is the prior distribution of $\boldsymbol{\theta}$.

1) Computation of $\boldsymbol{\Xi}_{P_{11}} \triangleq \mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\boldsymbol{\theta}}_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})\right]$: From [43, eq. (19)], we obtain the $M(N-1) \times M(N-1)$ matrix $\mathbb{E}_{\boldsymbol{\theta}}\left[-\Delta^{\boldsymbol{\theta}}_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})\right]$ as

$$\boldsymbol{\Xi}_{P_{11}} = \frac{-1}{\sigma_\delta^2}\begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & 1 & -2 & 1 \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}. \tag{A.7}$$

2) Computation of remaining terms in (A.6): Since CFO is a deterministic parameter and no prior knowledge of $\mathbf{h}$ is assumed, we have $\boldsymbol{\Xi}_{P_{12}} = \boldsymbol{\Xi}_{P_{21}}^T = \mathbf{0}_{M(N-1)\times ML}$, $\boldsymbol{\Xi}_{P_{13}} =$

$\boldsymbol{\Xi}_{P_{31}}^T = \mathbf{0}_{M(N-1)\times ML}$, $\boldsymbol{\Xi}_{P_{22}} = \boldsymbol{\Xi}_{P_{33}} = \boldsymbol{\Xi}_{P_{23}} = \boldsymbol{\Xi}_{P_{32}} = \mathbf{0}_{ML\times ML}$, $\boldsymbol{\Xi}_{P_{14}} = \boldsymbol{\Xi}_{P_{14}}^T = \mathbf{0}_{M(N-1)\times 1}$, $\boldsymbol{\Xi}_{P_{23}} = \boldsymbol{\Xi}_{P_{32}}^T = \mathbf{0}_{ML\times M}$, $\boldsymbol{\Xi}_{P_{24}} = \boldsymbol{\Xi}_{P_{42}}^T = \boldsymbol{\Xi}_{P_{34}} = \boldsymbol{\Xi}_{P_{43}}^T = \mathbf{0}_{ML\times M}$, and $\boldsymbol{\Xi}_{P_{44}} = \mathbf{0}_{M\times M}$.

Using the above results, we can evaluate the HIM in (6), as shown at the top of the previous page, since $\mathbf{B}_{11} = \boldsymbol{\Xi}_{D_{11}} + \boldsymbol{\Xi}_{P_{11}}$, $\mathbf{B}_{22} = \boldsymbol{\Xi}_{D_{22}} + \boldsymbol{\Xi}_{P_{22}} = \boldsymbol{\Xi}_{D_{22}}$, $\mathbf{B}_{33} = \boldsymbol{\Xi}_{D_{33}} + \boldsymbol{\Xi}_{P_{33}} = \boldsymbol{\Xi}_{D_{33}}$, $\mathbf{B}_{44} = \boldsymbol{\Xi}_{D_{44}} + \boldsymbol{\Xi}_{P_{44}} = \boldsymbol{\Xi}_{D_{33}}$, $\mathbf{B}_{12} = \mathbf{B}_{21}^H = \boldsymbol{\Xi}_{D_{12}} + \boldsymbol{\Xi}_{P_{12}} = \boldsymbol{\Xi}_{D_{12}}$, $\mathbf{B}_{13} = \mathbf{B}_{31}^H = \boldsymbol{\Xi}_{D_{13}} + \boldsymbol{\Xi}_{P_{13}} = \boldsymbol{\Xi}_{D_{13}}$, $\mathbf{B}_{23} = \mathbf{B}_{32}^H = \boldsymbol{\Xi}_{D_{23}} + \boldsymbol{\Xi}_{P_{23}} = \boldsymbol{\Xi}_{D_{23}}$, $\mathbf{B}_{14} = \mathbf{B}_{41}^H = \boldsymbol{\Xi}_{D_{14}} + \boldsymbol{\Xi}_{P_{14}} = \boldsymbol{\Xi}_{D_{14}}$, $\mathbf{B}_{24} = \mathbf{B}_{24}^H = \boldsymbol{\Xi}_{D_{24}} + \boldsymbol{\Xi}_{P_{24}} = \boldsymbol{\Xi}_{D_{24}}$, and $\mathbf{B}_{34} = \mathbf{B}_{43}^H = \boldsymbol{\Xi}_{D_{34}} + \boldsymbol{\Xi}_{P_{34}} = \boldsymbol{\Xi}_{D_{34}}$.

REFERENCES

[1] Z. Lan *et al.*, "Relay with deflection routing for effective throughput improvement in Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1453–1465, Oct. 2009.

[2] Z. Genc, G. H. Olcer, E. Onur, and I. Niemegeers, "Improving 60 GHz indoor connectivity with relaying," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Cape Town, South Africa, May 2010, pp. 1–6.

[3] Y. Jing and B. Hassibi, "Cooperative diversity in wireless relay networks with multiple-antenna nodes," in *Proc. Int. Symp. Inf. (ISIT)*, Sep. 2005, pp. 815–819.

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

[5] R. Wang, H. Mehrpouyan, M. Tao, and Y. Hua, "Channel estimation, carrier recovery, and data detection in the presence of phase noise in OFDM relay systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1186–1205, Feb. 2016.

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

[7] Y. Chi, A. Gomaa, N. Al-Dhahir, and A. R. Calderbank, "Training signal design and tradeoffs for spectrally-efficient multi-user MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2234–2245, Jul. 2011.

[8] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "Reduced-complexity joint baseband compensation of phase noise and I/Q imbalance for MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3450–3460, Nov. 2010.

[9] H. Minn, N. Al-Dhahir, and Y. Li, "Optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and phase noise," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1754–1759, Oct. 2006.

[10] R. Hamila, O. Özdemir, and N. Al-Dhahir, "Beamforming OFDM performance under joint phase noise and I/Q imbalance," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 2978–2989, May 2016.

[11] F. Z. Merli and G. M. Vitetta, "A factor graph approach to the iterative detection of OFDM signals in the presence of carrier frequency offset and phase noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 868–877, Mar. 2008.

[12] X. Zhou, X. Yang, R. Li, and K. Long, "Efficient joint carrier frequency offset and phase noise compensation scheme for high-speed coherent optical OFDM systems," *J. Lightw. Technol.*, vol. 31, no. 11, pp. 1755–1761, Jun. 1, 2013.

[13] 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.

[14] L. Tomba and W. A. Krzymien, "Sensitivity of the MC-CDMA access scheme to carrier phase noise and frequency offset," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1657–1665, Sep. 1999.

[15] 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, nos. 2–4, pp. 191–193, Feb./Mar./Apr. 1995.

[16] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Phase noise estimation in OFDM: Utilizing its associated spectral geometry," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1999–2012, Apr. 2016.

[17] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Performance analysis of OFDM with Wiener phase noise and frequency selective fading channel," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1321–1331, May 2011.

[18] B. Can, M. Portalski, H. S. D. Lebreton, S. Frattasi, and H. A. Suraweera, "Implementation issues for OFDM-based multihop cellular networks," *IEEE Commun. Mag.*, vol. 45, no. 9, pp. 74–81, Sep. 2007.

[19] 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.

[20] Q. Huang, M. Ghogho, J. Wei, and P. Ciblat, "Practical timing and frequency synchronization for OFDM-based cooperative systems," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3706–3716, Jul. 2010.

[21] Y. Yao and X. Dong, "Multiple CFO mitigation in amplify-and-forward cooperative OFDM transmission," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3844–3854, Dec. 2012.

[22] 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.

[23] M. O. Pun, M. Morelli, and C. C. J. Kuo, "Iterative detection and frequency synchronization for OFDMA uplink transmissions," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 629–639, Feb. 2007.

[24] O. H. Salim, A. A. Nasir, W. Xiang, and R. A. Kennedy, "Joint channel, phase noise, and carrier frequency offset estimation in cooperative OFDM systems," in *Proc. Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 4384–4389.

[25] O. H. Salim, "Cooperative systems based signal processing techniques with applications to three-dimensional video transmission," Ph.D. dissertation, School Mech. Elect. Eng., Univ. Southern Queensland, Toowoomba, QLD, Australia, Sep. 2015. [Online]. Available: http://eprints.usq.edu.au/28859/

[26] O. H. Salim, W. Xiang, A. A. Nasir, G. Wang, and H. Mehrpouyan, "MIMO-OFDM systems in the presence of phase noise with application to light field video transmission," *IEEE Trans. Veh. Commun.*, submitted for publication. [Online]. Available: http://arxiv.org/abs/1602.02834

[27] F. Z. Merli, X. Wang, and G. M. Vitetta, "A Bayesian multiuser detection algorithm for MIMO-ODFM systems affected by multipath fading, carrier frequency offset, and phase noise," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 3, pp. 506–516, Apr. 2008.

[28] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 47, no. 5, pp. 655–674, May 2000.

[29] 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.

[30] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "A non-iterative technique for phase noise ICI mitigation in packet-based OFDM systems," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5945–5950, Nov. 2010.

[31] F. Munier, T. Eriksson, and A. Svensson, "An ICI reduction scheme for OFDM system with phase noise over fading channels," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1119–1126, Jul. 2008.

[32] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1607–1616, Aug. 2007.

[33] T. Paul and T. Ogunfunmi, "Wireless LAN comes of age: Understanding the IEEE 802.11n amendment," *IEEE Circuits Syst. Mag.*, vol. 8, no. 1, pp. 28–54, Mar. 2008.

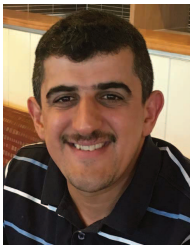
[34] X. Zhu, A. Doufexi, and T. Kocak, "Throughput and coverage performance for IEEE 802.11ad millimeter-wave WPANs," in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, Budapest, Hungary, May 2011, pp. 1–5.

[35] O. H. Salim, A. A. Nasir, H. Mehrpouyan, W. Xiang, S. Durrani, and R. A. Kennedy, "Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid Cramér–Rao lower bound," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3311–3325, Sep. 2014.

[36] H. L. van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. New York, NY, USA: Wiley, 2007.

[37] M. A. Ruder, R. Schober, and W. H. Gerstacker, "Cramer–Rao lower bound for channel estimation in a MUROS/VAMOS downlink transmission," in *Proc. IEEE Pers. Indoor Mobile Radio Commun. Symp. (PIMRC)*, Toronto, ON, Canada, Sep. 2011, pp. 1433–1437.

[38] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2664–2677, Oct. 1994.
[39] X.-L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: A general framework," *Biometrika*, vol. 80, no. 2, pp. 267–278, Jun. 1993.
[40] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. Hoboken, NJ, USA: Wiley, 2008.
[41] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Signal Processing Series). Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[42] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.
[43] S. Bay, C. Herzet, J.-M. Brossier, J.-P. Barbot, and B. Geller, "Analytic 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.



Omar H. Salim (S'10–M'15) received the B.Sc. and M.Sc. degrees in electronics and communication engineering from Mosul University, Mosul, Iraq, in 2001 and 2004, respectively, and the Ph.D. degree in signal processing for communication engineering from the University of Southern Queensland, Toowoomba, Australia, where he is currently a Post-Doctoral Research Fellow with the School of Electrical and Mechanical. He has been a Lecturer with the Department of Computer Engineering, University of Mosul, since 2004. He has authored several high-quality papers. His research interests include signal processing for communication and OFDM systems and digital signal processing implementation for 3-D video communication systems.

Ali A. Nasir (S'09–M'13) received the Ph.D. degree in telecommunications engineering from Australian National University, Australia, in 2013. He was a Research Fellow with Australian National University from 2012 to 2015. He was an Assistant Professor with the School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Paksitan, from 2015 to 2016. He is currently an Assistant Professor with the Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. His research interests are in the area of signal processing in wireless communication systems. He was a Co-Chair of the WCNC Workshop on wireless energy harvesting in 2016. He is an Associate Editor of the IEEE CANADIAN JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING.

Hani Mehrpouyan (S'05–M'10) received the B.Sc. degree (Hons.) in electrical engineering from Simon Fraser University, Burnaby, BC, Canada, in 2004, and the Ph.D. degree in electrical engineering from Queens University, Kingston, ON, Canada, in 2011. From 2011 to 2012, he held a post-doctoral position with the Chalmers University of Technology, where he led the MIMO aspects of the microwave backhauling for the Next Generation Wireless Networks Project. He joined the University of Luxembourg as a Research Associate in 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. He has over 50 publications in prestigious IEEE journals and conferences. His current research interests lie in the area of applied signal processing and application of high performance computing on wireless communication systems, including millimeter-wave systems, hybrid computing for simulating complex communication systems, reconfigurable antennas, and energy harvesting systems, synchronization, and channel estimation. He was a recipient of the National Science Foundation EARS grant for his research in the area of millimeterwave communications. He was also a recipient of many scholarships and awards, such as the IEEE Globecom Early Bird Student Award, the NSERC-IRDF, the NSERC PGSD, the NSERC CGS-M Alexander Graham Bell, and the B.C. Wireless Innovation. He is an Associate Editor of the IEEE COMMUNICATION LETTERS and he has also served as a TPC member of the IEEE Globecom, ICC, and VTC. He has also been involved with industry leaders, such as Ericsson AB and Blackberry.

Wei Xiang (S'00–M'04–SM'10) received the B.Eng. and M.Eng. degrees, in electronic engineering, from the University of Electronic Science and Technology of China, Chengdu, China, in 1997 and 2000, respectively, and the Ph.D. degree in telecommunications engineering from the University of South Australia, Adelaide, SA, Australia, in 2004.

From 2004 to 2015, he was with the School of Mechanical and Electrical Engineering, University of Southern Queensland, Toowoomba, QLD, Australia. In 2008, he was a Visiting Scholar with Nanyang Technological University, Singapore. From 2010 to 2011, he was a Visiting Scholar with the University of Mississippi, Oxford, MS, USA. From 2012 to 2013, he was an Associate Professor with The University of Hong Kong. He has authored or co-authored over 170 papers in peer-reviewed journal and conference papers. He is currently a Foundation Professor and the Head of Discipline Internet of Things Engineering with the College of Science and Engineering, James Cook University, Cairns, QLD, Australia. His research interests are in the broad area of communications and information theory, particularly coding and signal processing for multimedia communications systems. He is a Fellow of the IET and a Fellow of Engineers Australia. He was a co-recipient of three Best Paper Awards at the 2015 WCSP, the 2011 IEEE WCNC, and the 2009 ICWMC. He was a recipient of several prestigious fellowship titles. He was named a Queensland International Fellow (2010–2011) by the Queensland Government of Australia, an Endeavour Research Fellow (2012–2013) by the Commonwealth Government of Australia, a Smart Futures Fellow (2012–2015) by the Queensland Government of Australia, and a JSPS Invitational Fellow jointly by the Australian Academy of Science and the Japanese Society for Promotion of Science (2014–2015). He is an Editor of the IEEE COMMUNICATIONS LETTERS, and an Associate Editor for *Telecommunication Systems* (Springer).