Joint Channel, Phase Noise, and Carrier Frequency Offset Estimation in Cooperative OFDM Systems

Omar H. Salim, Ali A. Nasir, Wei Xiang and Rodney A. Kennedy

***Abstract*—Cooperative communication systems employ cooperation among nodes in a wireless network to increase data throughput and robustness to signal fading. However, such advantages are only possible if there exist perfect synchronization among all nodes. Impairments like channel multipath, time varying phase noise (PHN) and carrier frequency offset (CFO) result in the loss of synchronization and diversity performance of cooperative communication systems. Joint estimation of these multiple impairments is necessary in order to correctly decode the received signal in cooperative systems. In this paper, we propose an iterative pilot-aided algorithm based on expectation conditional maximization (ECM) for joint estimation of multipath channels, Wiener PHNs, and CFOs in amplify-and-forward (AF) based cooperative orthogonal frequency division multiplexing (OFDM) system. Numerical results show that the proposed estimator achieves mean square error performance close to the derived hybrid Cramer-Rao lower bound (HCRB) for different PHN variances.**

I. Introduction

Cooperative communication has attracted considerable research interest due to its potential to reduce the size and power constraints caused by an increased number of antennas in multiple-input-multiple-output (MIMO)-mobile devices. Relay based cooperative communication schemes using single-antenna transceivers provide spatial diversity by forming virtual MIMO systems [1], [2]. Such diversity gains are however only possible if perfect synchronization, e.g., perfect estimation of channel, phase noises (PHNs), and carrier frequency offsets (CFOs) exists among all communication nodes [3].

In orthogonal frequency division multiplexing (OFDM) systems, which are employed to increase the transmission bandwidth efficiency and mitigate the effect of the frequency-selective fading in each link of the cooperative systems, CFOs and PHNs result in a common phase error (CPE) and inter-carrier interference (ICI) at the destination [4]. In addition, the estimation of the channel impulse response (CIR) for each link becomes challenging in the presence of CFOs and PHNs [3]. On the other hand, the accurate estimation of these *multiple impairments*, i.e., CIR, CFOs, and time-varying PHNs, is required for coherent detection of OFDM signals at the destination.

Most of the existing work in the literature focuses on estimating either CFOs while assuming perfect estimation of PHNs [5]–[7] or target the estimation PHN parameters while assuming perfect CFOs estimation [8]. More importantly, [5]–[8] do not provide the hybrid Cramér-Rao lower bound (HCRB) for joint estimation of multiple impairments in cooperative OFDM systems, which would provide essential information about the absolute performance of the estimation scheme.

Omar H. Salim and Wei Xiang are with the School of Mechanical and Electrical Engineering, University of Southern Queensland, Australia, and Ali A. Nasir and Rodney A. Kennedy are with Research School of Engineering, The Australian National University, Canberra, Australia. This research is partially supported under Australian Research Council's Discovery Projects funding scheme (project number DP110102548). Emails: omarhazim.salim@usq.edu.au, ali.nasir@anu.edu.au, wei.xiang@usq.edu.au, rodney.kennedy@anu.edu.au

A. Contributions

In this paper, we consider simple amplify-and-forward (AF) relaying protocol, where during the first time slot, source node broadcasts the information to the relay and destination node, while during the second time slot, the relay node amplifies and forwards the source information to the destination node. In order to guarantee the advantages of cooperative diversity, there is a need to estimate the CIR, time varying PHNs, and CFOs for the received signal at the destination node during both time slots. The overall contributions of this paper can be summarized as follows:

- This is a first paper to address the joint estimation of channel, PHN, and CFO in cooperative OFDM systems. We propose iterative pilot-aided algorithm based on the expectation conditional maximization (ECM) for AF cooperative OFDM networks in the presence of *unknown* channel gains, PHNs and CFOs. In the E-step, we propose *extended Kalman filter* (EKF) estimator that is shown to accurately track the time varying PHN over the transmission frame. In the M-step, we drive a closed form estimator to estimate the CFO and CIR parameters. It has been found through simulations that the proposed ECM based estimator only require few iterations to track the multiple impairments over the transmission frame.
- Simulations are carried out to investigate the performance of the proposed estimator. These simulation results demonstrate that the proposed estimators performance is close to the derived HCRB at moderate-to-high values of signal-to-noise ratios (SNRs).

B. 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 matlab notation $\mathbf{X}(n_1{:}n_2, m_1{:}m_2)$ is used to extract a submatrix within $\mathbf{X}$ 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 $\mathbf{x}$. $\mathbb{E}_{x,y}[\cdot]$ denotes the expectation over x and y, and $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary parts of a complex quantity, respectively. The imaginary unit is $j = \sqrt{-1}$, and $\otimes$ and $\star$ denote the circular convolution and normal convolution operators, respectively. Finally, $\mathcal{CN} \sim (\mu, \sigma^2)$ and $\mathcal{N} \sim (\mu, \sigma^2)$

978-1-4799-2003-7/14/$31.00 ©2014 IEEE

denote the complex and real Gaussian distributions with mean μ and variance σ^2.

C. 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. In Section III, the proposed estimator is derived while in Section IV, simulation results that investigate the performance of the proposed estimator are presented. Section V concludes the paper.

II. Signal Model

In this paper, the TDMA-based AF relay protocol is considered, i.e., the source node broadcasts the signal to the relay and the destination in the first hop (i.e., first transmission time slot). In the second hop, the relay amplifies the received signal and forwards it to the destination while the source is silent (see Fig.1). Since there is no collision between the received signals during the two consecutive hops at the destination, this transmission protocol maintains orthogonality at the expense of loss in spectral efficiency.

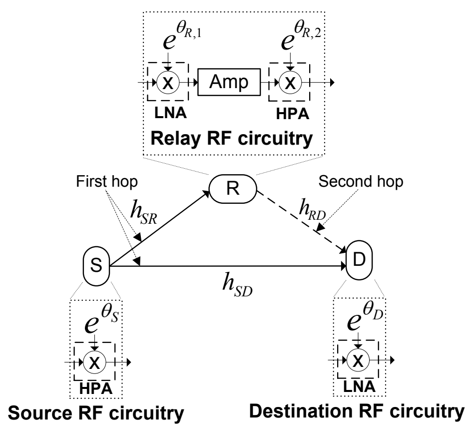

Fig. 1: **System block diagram**

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, and $\delta(x)$ is the unit impulse function. 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 $\mathbf{h}_{a,b} \triangleq [h_{a,b}(0), h_{a,b}(1), \ldots, h_{a,b}(L-1)]^T$. The frequency-domain channel coefficient matrix is $\mathbf{H}_{a,b} = \text{diag}\{H_{a,b}[0], H_{a,b}[1], \ldots, H_{a,b}[N-1]\}$, where $H_{a,b}[n] = \sum_{d=0}^{L-1} h_{a,b}(d)e^{-(j2\pi nd/N)}$ is the channel frequency response on the nth subcarrier and N is the number of subcarriers. 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 S transmits the modulated training symbol vector $\mathbf{d}_S = [d_S[0], d_S[1], \ldots, d_S[N-1]]^T$.

A. First Time Slot

The received samples at the destination D and the relay R in time domain are given by

$$y_{D,1}(n) = e^{j(\theta_{S,D}(n)+2\pi n\epsilon_{S,D}/N)}(h_{S,D}(n) \otimes x(n)) + w_{D,1}(n) \tag{1}$$

$$y_{R,1}(n) = \sqrt{g_{S,R}}e^{j(\theta_{S,R_1}(n)+2\pi n\epsilon_{S,R}/N)}(h_{S,R}(n) \otimes x(n)) + w_R(n) \tag{2}$$

where

- $x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} d_S(k)e^{j2\pi kn/N}$ for $n = 0, 1, \ldots, N-1$, is complex baseband OFDM signal and *k* denotes the subcarrier index,
- $\{h_{S,D}(l)\}_{l=0}^{L-1}$ and $\{h_{S,R}(l)\}_{l=0}^{L-1}$ are the channel impulse response from $S \to D$ and $S \to R$, respectively,
- $g_{S,R}$=$(d_{S,D}/d_{S,R})^\gamma$ is the large-scale fading gain, where we assume that the distance between the source and the relay is smaller than the distance between the source and the destination (see Fig.1). $d_{S,D}$ and $d_{S,R}$ are the physical distances from $S \to D$ and $S \to R$, respectively, γ is the large-scale fading exponent,
- $\theta_{S,D}(n) = \theta_S(n) + \theta_D(n)$ and $\theta_{S,R_1}(n) = \theta_S(n) + \theta_{R_1}(n)$ are the PHN processes between S and D and S and R, respectively, during the first time slot, $\theta_k(n)$ for $k \in \{S, R, D\}$, is generated using a Wiener process, i.e., $\theta(n) = \theta(n-1) + \delta(n)$, $\forall$ n, where $\delta(n) \sim \mathcal{N}(0, \sigma_\delta^2)$ is the PHN innovation and σ_δ^2 is the variance of the innovation process [9],
- $\epsilon_{S,R}$ and $\epsilon_{R,D}$ denote the normalized CFOs between the S and R and R and D nodes, respectively,
- $w_{D,1}(n)$ and $w_R(n)$ are the additive white Gaussian noise (AWGN) samples with $\{w_{D,1}(n), w_R(n)\} \sim \mathcal{N}(0, \sigma_w^2)$.

Using (1) and (2) the received signals at D and R, $\mathbf{y}_{D,1} = [y_{D,1}(0), y_{D,1}(1), \ldots, y_{D,1}(N-1)]^T$ and $\mathbf{y}_{R,1} = [y_{R,1}(0), y_{R,1}(1), \ldots, y_{R,1}(N-1)]^T$, respectively, in vectorial form are given by

$$\mathbf{y}_{D,1} = \mathbf{E}_{S,D}\mathbf{P}_{S,D}\mathbf{F}^H\mathbf{D}_S\mathbf{F}_L\mathbf{h}_{S,D} + \mathbf{w}_{D,1} \tag{3}$$

$$\mathbf{y}_{R,1} = \sqrt{g_{SR}}\mathbf{E}_{S,R}\mathbf{P}_{S,R_1}\mathbf{F}^H\mathbf{D}_S\mathbf{F}_L\mathbf{h}_{S,R} + \mathbf{w}_R \tag{4}$$

where

- $\mathbf{P}_{S,D} \triangleq \text{diag}([e^{j\theta_{S,D}(0)}, e^{j\theta_{S,D}(1)}, \ldots, e^{j\theta_{S,D}(N-1)}]^T)$ and $\mathbf{P}_{S,R_1} \triangleq \text{diag}([e^{j\theta_{S,R_1}(0)}, e^{j\theta_{S,R_1}(1)}, \ldots, e^{j\theta_{S,R_1}(N-1)}]^T)$ are $N \times N$ PHN matrices from $S \to D$ and $S \to R$, respectively,
- $\mathbf{E}_{S,D}$ and $\mathbf{E}_{S,R}$ are $N \times N$ CFO matrices from $S \to D$, and from $S \to R$, respectively, i.e., $\mathbf{E}_{a,b} \triangleq \text{diag}([e^{(j2\pi\epsilon_{ab}/N)\times 0}, e^{(j2\pi\epsilon_{ab}/N)}, \ldots, e^{(j2\pi\epsilon_{ab}/N)\times(N-1)}]^T)$, $a, b \in \{S, R, D\}$,
- $\mathbf{F}$ is an $N \times N$ DFT matrix, i.e., $[\mathbf{F}]_{l,m} \triangleq (1/\sqrt{N})e^{-j(2\pi ml/N)}$ for $m, l = 0, 1, \cdots, N-1$,
- $\mathbf{D}_S \triangleq \text{diag}(\mathbf{d}_S[n])$,
- $\mathbf{F}_L$ is an $N \times L$ DFT matrix, i.e., $\mathbf{F}_L \triangleq \mathbf{F}(1:N, 1:L)$,
- $\mathbf{w}_{D,1} = [w_{D,1}(0), w_{D,1}(1), \ldots, w_{D,1}(N-1)]^T$ and $\mathbf{w}_R = [w_R(0), w_R(1), \ldots, w_R(N-1)]^T$ are AWGN vectors.

B. Second Time Slot

In this time slot, the relay simply amplifies and forwards the received signal to the destination. By adopting the low phase noise variance assumption, as explained in [8], the approximated received singal at D is given by

$$\begin{aligned}\mathbf{y}_{D,2}&=\sqrt{g_{R,D}\frac{\mathbb{E}[|\mathbf{d}_S|^2]}{\mathbb{E}[|\mathbf{y}_{R,1}|^2]}}\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{F}\mathbf{y}_{R,1}+\mathbf{w}_{D,2}\\&=\sqrt{\frac{g_{R,D}g_{S,R}\xi_D}{g_{S,R}\xi_D\sigma_h^2+\sigma_w^2}}\mathbf{E}_{R,D}\mathbf{E}_{S,R}\mathbf{P}_{R_2,D}\mathbf{P}_{S,R_1}\mathbf{F}^H\mathbf{H}_{S,R}\mathbf{H}_{R,D}\mathbf{d}_S\\&+\sqrt{\frac{g_{R,D}\xi_D}{g_{S,R}\xi_D\sigma_h^2+\sigma_w^2}}\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{F}\mathbf{w}_R+\mathbf{w}_{D,2}\end{aligned}\qquad(5)$$

Equation (5) can further be simplified as

$$\begin{aligned}\mathbf{y}_{D,2}&=\sqrt{\frac{g_{R,D}g_{S,R}\xi_D}{g_{S,R}\xi_D\sigma_h^2+\sigma_w^2}}\mathbf{E}_{S,R,D}\mathbf{P}_{S,R,D}\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\mathbf{h}_{S,R,D}\\&+\sqrt{\frac{g_{R,D}\xi_D}{g_{S,R}\xi_D\sigma_h^2+\sigma_w^2}}\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{F}\mathbf{w}_R+\mathbf{w}_{D,2}\end{aligned}\qquad(6)$$

where

- σ_h^2=$\mathbb{E}[|h_{S,R}|^2]$ is the variance of CIR vector,
- $g_{R,D}$=$(d_{S,D}/d_{R,D})^\gamma$, $d_{R,D}$ is the distance from $R\to D$,
- $\mathbf{w}_{D,2}$ is the AWGN vector at the destination,
- $\mathbf{E}_{S,R,D}$=$\mathbf{E}_{R,D}\mathbf{E}_{S,R}$ is the effective CFO matrix during the first and the second hops from $S\to R\to D$,
- $\mathbf{P}_{S,R,D}\triangleq\text{diag}([e^{j\theta_{S,R,D}(0)},\ldots,e^{j\theta_{S,R,D}(N-1)}]^T)$, $\theta_{S,R,D}$=$\theta_S+\theta_{R,1}+\theta_{R,2}+\theta_D$ is the effective PHN matrix during the first and the second hops from $S\to R\to D$,
- $g_{R,D}$=$(d_{S,D}/d_{R,D})^\gamma$, $d_{R,D}$ is the distance from $R\to D$,
- $\mathbf{F}_{(2L-1)}$ is an $N\times(2L-1)$ DFT matrix, i.e., $\mathbf{F}_{(2L-1)}\triangleq\mathbf{F}(1:N,1:(2L-1))$, L is the channel length,
- $\mathbf{h}_{S,R,D}\triangleq\mathbf{h}_{S,R}\star\mathbf{h}_{R,D}$ is a $2L-1\times1$ vector,
- $\mathbf{h}_{R,D}$ is CIR between the relay and destination,
- $\mathbf{H}_{R,D}\triangleq\text{diag}\,(\mathbf{F}_L\mathbf{h}_{R,D})$, and $\mathbf{H}_{S,D}\triangleq\text{diag}\,(\mathbf{F}_L\mathbf{h}_{S,D})$.

The received signal vector at D, $\mathbf{y}_{D,2}$ in (6), is a circularly symmetric complex Gaussian random variable, i.e., $\mathbf{y}_{D,2}\sim\mathcal{CN}(\boldsymbol{\mu}_{\mathbf{y}_{D,2}},\boldsymbol{\Sigma}_{\mathbf{y}_{D,2}})$, with mean $\boldsymbol{\mu}_{\mathbf{y}_{D,2}}=\text{q}_1\mathbf{E}_{S,R,D}\mathbf{P}_{S,R,D}\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\mathbf{h}_{S,R,D}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}_{D,2}}=(\text{q}_2^2\sigma_w^2\sigma_h^2+\sigma_w^2)\mathbf{I}_N$, where $\text{q}_1\triangleq\sqrt{\frac{g_{R,D}g_{S,R}\xi_D}{g_{S,R}\xi_D\sigma_h^2+\sigma_w^2}}$ and $\text{q}_2\triangleq\sqrt{\frac{g_{R,D}\xi_D}{g_{S,R}\xi_D\sigma_h^2+\sigma_w^2}}$ ($\boldsymbol{\mu}_{\mathbf{y}_{D,2}}$ and $\boldsymbol{\Sigma}_{\mathbf{y}_{D,2}}$ are derived in Appendix A).

C. Problem Formulation

In order to gain the advantages of cooperative diversity, the destination receiver needs to decode the received signals, $\mathbf{y}_{D,1}$ in (3), and $\mathbf{y}_{D,2}$ in (6), during first and second time slots, respectively. This in turn requires the estimation of $\boldsymbol{\theta}_{S,D}\triangleq[\theta_{S,D}(0),\ldots,\theta_{S,D}(N-1)]^T$, $\mathbf{h}_{S,D}\triangleq[h_{S,D}(0),\ldots,h_{S,D}(L-1)]^T$ and $\epsilon_{S,D}$ during first time slot and the estimation of $\boldsymbol{\theta}_{S,R,D}\triangleq[\theta_{S,R,D}(0),\ldots,\theta_{S,R,D}(N-1)]^T$, $\mathbf{h}_{S,R,D}\triangleq[h_{S,R,D}(0),\ldots,h_{S,R,D}(L-1)]^T$ and $\epsilon_{S,R,D}$ during second time slot. For brevity, the following section derives an algorithm for joint estimation of $\boldsymbol{\theta}_{S,R,D}$, $\mathbf{h}_{S,R,D}$, and $\epsilon_{S,R,D}$, during the second time slot only. The estimation of the respective impairments during the first time slot can be easily obtained by following the same approach. Let us define $\boldsymbol{\lambda}\triangleq[\boldsymbol{\theta}_{S,R,D}^T\ \mathbf{h}_{S,R,D}^T\ \epsilon_{S,R,D}]^T$

III. Proposed ECM based Estimator

The ECM algorithm at the destination receiver iterates between the expectation step (E-step) and the maximization step (M-step). In E-step, EKF is proposed to update the PHN vector at $(i+1)$th iteration, $\boldsymbol{\theta}_{S,R,D}^{[i+1]}$, using the CIR and CFO estimates, $\hat{\mathbf{h}}_{S,R,D}^{[i]}$ and $\hat{\epsilon}_{S,R,D}^{[i]}$, respectively, obtained from the previous (ith) iteration, while in M-step, closed-form estimators are derived to update the CIR and CFO estimates, $\hat{\mathbf{h}}_{S,R,D}^{[i+1]}$ and $\hat{\epsilon}_{S,R,D}^{[i+1]}$, respectively. The proposed ECM algorithm at the ith iteration is given as follows.

For the given problem, let us define $\mathbf{s}\triangleq\text{q}_1\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\mathbf{h}_{S,R,D}=[s(0),\ldots,s(N-1)]^T$ is $N\times1$ vector and re-write $\mathbf{y}_{D,2}$ in (6) as follows

$$\mathbf{u}=\mathbf{E}_{S,R,D}\mathbf{P}_{S,R,D}\mathbf{s}+\mathbf{v};\qquad(7)$$

where the overall noise vector $\mathbf{v}\triangleq[v(0),\ldots,v(N-1)]^T=\text{q}_2\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{F}\mathbf{w}_R+\mathbf{w}_{D,2}$ is distributed as $\mathbf{v}\sim\mathcal{CN}(\mathbf{0},\tilde{\sigma}_w^2\mathbf{I}_N)$ and $\tilde{\sigma}_w^2\triangleq\text{q}_2^2\sigma_w^2\sigma_h^2+\sigma_w^2$. Following [10], the complete data is defined as $\mathbf{z}\triangleq[\mathbf{u}^T\ \boldsymbol{\theta}_{S,R,D}^T]^T$ and the negative log likelihood function of the complete data, $\log p(\mathbf{z};\epsilon_{S,R,D})$, is given by

$$\begin{aligned}\log p(\mathbf{z};\epsilon_{S,R,D})=\text{C1}+\frac{1}{\tilde{\sigma}_w^2}\sum_{n=0}^{N-1}&\parallel u(n)-e^{j2\pi\epsilon_{S,R,D}n/N}\\&\times e^{j\theta_{S,R,D}(n)}s(n)\parallel^2+\log p(\theta_{S,R,D}(0))\\&+\sum_{n=0}^{N-1}\log p(\theta_{S,R,D}(n)|\theta_{S,R,D}(n-1)),\end{aligned}\qquad(8)$$

where C1 is a constant. The detailed E-step and M-step for estimating the CIR, PHN, and CFO are as follows:

E-step: In this step, the received signal $u(n)$ is first multiplied by $e^{-j2\pi\hat{\epsilon}_{S,R,D}^{[i]}n/N}$. Next, the signal $x(n)\triangleq e^{-j2\pi n\epsilon_{S,R,D}^{[i]}/N}u(n)$ is used to estimate the PHN vector, where $\hat{\epsilon}_{S,R,D}^{[i]}$ is the latest CFO estimate obtained from the previous iteration. First, we propose to use EKF during E-step to estimate the PHN samples $\boldsymbol{\theta}_{S,R,D}$. The signal $x(n)$ can be written as

$$\begin{aligned}x(n)&=e^{-j2\pi n\epsilon_{S,R,D}^{[i]}/N}u(n)\\&=e^{j2\pi n\Delta\hat{\epsilon}_{S,R,D}/N}e^{j\theta_{S,R,D}(n)}s^{[i]}(n)+\tilde{v}(n),\end{aligned}\qquad(9)$$

where $s^{[i]}(n)$ is nth symbol of the vector $\mathbf{s}^{[i]}\triangleq\text{q}_1\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\mathbf{h}_{S,R,D}^{[i]}$, $\Delta\hat{\epsilon}_{S,R,D}\triangleq\epsilon_{S,R,D}-\hat{\epsilon}_{S,R,D}^{[i]}$, and $\tilde{v}(n)\triangleq v(n)e^{-j2\pi n\hat{\epsilon}_{S,R,D}^{[i]}/N}$. The state and observation equations at time n are given by

$$\theta_{S,R,D}(n)=\theta_{S,R,D}(n-1)+\delta_{S,R,D}(n),\qquad(10)$$

$$u(n)=z(n)+\tilde{v}(n)=e^{j\theta_{S,R,D}(n)}s(n)+\tilde{v}(n).\qquad(11)$$

Since the observation equation in (11) is a non-linear function of the unknown state vector $\boldsymbol{\theta}_{S,R,D}$, the EKF is used instead of simple Kalman filtering. The EKF uses Taylor series expansion to linearize the non-linear observation equation in (11) about the current estimates [11]. Thus, the Jacobian of $z(n)$ is evaluated by computing the first order partial derivative of $z(n)$ with respect to $\theta_{S,R,D}(n)$ as

$$\dot{z}(n) = \frac{\partial z(\theta_{S,R,D}(n))}{\partial \theta_{S,R,D}(n)}\Big|_{\theta_{S,R,D}(n)=\hat{\theta}_{S,R,D}(n|n-1)} \quad (12)$$
$$= jz(\hat{\theta}_{S,R,D}(n|n-1)) \quad (13)$$
$$= je^{j\hat{\theta}^{[i]}_{S,R,D}(n|n-1)}\hat{s}(n),$$

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

$$\hat{\theta}^{[i+1]}_{S,R,D}(n|n-1) = \hat{\theta}^{[i+1]}_{S,R,D}(n-1|n-1), \quad (14)$$
$$M^{[i+1]}(n|n-1) = M^{[i+1]}(n-1|n-1) + \sigma^2_{\delta_{S,R,D}}, \quad (15)$$

Given the observation $u(n)$, the Kalman gain $K(n)$, posteriori state estimate $\hat{\theta}^{[i+1]}_{S,R,D}(n|n)$, and the filtering error covariance, $M^{[i+1]}(n|n)$ are given by

$$K(n) = M^{[i+1]}(n|n-1)\dot{z}^*(\theta_{S,R,D}(n|n-1)) \times \Big(\dot{z}(\theta_{S,R,D}(n|n-1)) \times M^{[i+1]}(n-1|n-1) \times \dot{z}^*(\theta_{S,R,D}(n|n-1)) + \tilde{\sigma}^2_w\Big)^{-1}. \quad (16)$$

$$\hat{\theta}^{[i+1]}_{S,R,D}(n|n) = \hat{\theta}^{[i+1]}_{S,R,D}(n|n-1) + \Re\big\{K_n\big(u(n) - e^{j\hat{\theta}^{[i+1]}_{S,R,D}(n|n-1)}\hat{s}^{[i]}(n)\big)\big\} \quad (17)$$

$$M^{[i+1]}(n|n) = \Re\big\{M^{[i+1]}(n|n-1) - K(n)\dot{z}(\theta_{S,R,D}(n|n-1)) \times M^{[i+1]}(n|n-1)\big\}, \quad (18)$$

where $\hat{s}^{[i]}(n)$ is nth symbol of the vector $\hat{\mathbf{s}}^{[i]} \triangleq \mathrm{q}_1\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\hat{\mathbf{h}}^{[i]}_{S,R,D}$. Before starting the EKF recursion (12)-(18), $\hat{\theta}^{[1]}_{S,R,D}(1|0)$ and $M^{[1]}(1|0)$ are initialized by $\hat{\theta}^{[1]}_{S,R,D}(1|0) = 0$ and $M^{[1]}(1|0) = \sigma^2_{\delta_{S,R,D}}$.

M-step: In this step, the CIR and CFO from $S \to R \to D$ are estimated by minimizing the log likelihood function in (8). In order to further reduce the complexity associated with the *M*-step, the minimization in (8) is done with respect to one of the estimation parameter while keeping the other parameter at its most recently updated value [12]. We first minimize the log likelihood function in (8) with respect to $\epsilon_{S,R,D}$ to update the CFO estimate for $(i+1)$th iteration, $\hat{\epsilon}^{[i+1]}_{S,R,D}$, while channel is kept constant at its ith iteration value, $\hat{\mathbf{h}}^{[i]}_{S,R,D}$ and updated PHN vector, $\hat{\boldsymbol{\theta}}^{[i+1]}_{S,R,D}$, is obtained from the E-step. Thus, the CFO estimate update, $\hat{\epsilon}^{[i+1]}_{S,R,D}$, is given by

$$\hat{\epsilon}^{[i+1]}_{S,R,D} = \arg\min_{\epsilon_{S,R,D}} \sum_{n=0}^{N-1} \parallel u(n) - e^{j2\pi\epsilon_{S,R,D}n/N}e^{j\theta_{S,R,D}(n)} s(n) \parallel^2 \Big|_{\theta_{S,R,D}(n)=\hat{\theta}^{[i+1]}_{S,R,D}(n),\mathbf{h}_{S,R,D}=\hat{\mathbf{h}}^{[i]}_{S,R,D}} \quad (19)$$

After simplifying (19), we have

$$\hat{\epsilon}^{[i+1]}_{S,R,D} = \arg\max_{\epsilon_{S,R,D}} \sum_{n=0}^{N-1} \Re\{(u(n))^*\hat{S}^{[i]}(n)e^{j2\pi\epsilon_{S,R,D}n/N}\} \quad (20)$$

where $\hat{S}^{[i]}(n) = e^{j\hat{\theta}^{[i+1]}_{S,R,D}(n)}\hat{s}^{[i]}(n)$. In order to handle the non-linearity of (20), we can approximate the term $e^{j2\pi\epsilon_{S,R,D}n/N}$ using Taylor series expansion around the pervious CFO estimate, $\hat{\epsilon}^{[i]}_{S,R,D}$, up to the second order term as

$$e^{j2\pi\epsilon_{S,R,D}n/N} = e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N} + (\epsilon_{S,R,D} - \hat{\epsilon}^{[i]}_{S,R,D})\left(j\frac{2\pi}{N}n\right) \times e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N} + \frac{1}{2}(\epsilon_{S,R,D} - \hat{\epsilon}^{[i]}_{S,R,D})^2\left(j\frac{2\pi}{N}n\right)^2 \times e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N} \quad (21)$$

Substituting (21) into (20), $\hat{\epsilon}^{[i+1]}_{S,R,D}$ is given by

$$\hat{\epsilon}^{[i+1]}_{S,R,D} = \arg\max_{\epsilon_{S,R,D}} \Bigg\{ \sum_{n=0}^{N-1}\Re\big\{(u(n))^*\hat{S}^{[i+1]}(n)e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N} + (\epsilon_{S,R,D} - \hat{\epsilon}^{[i]}_{S,R,D})\sum_{n=0}^{N-1}\Re\big\{(u(n))^*\hat{S}^{[i+1]}(n)\left(j\frac{2\pi}{N}n\right) e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N}\big\} + \frac{1}{2}(\epsilon_{S,R,D} - \hat{\epsilon}^{[i]}_{S,R,D})^2 \sum_{n=0}^{N-1}\Re\big\{(u(n))^*\hat{S}^{[i+1]}(n)\left(j\frac{2\pi}{N}n\right)^2 e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N}\big\}\Bigg\} \quad (22)$$

Taking the derivative of (22) with respect to $\epsilon_{S,R,D}$ and equating the result to zero, the estimate of $\epsilon_{S,R,D}$ at the $(i+1)$th iteration is given by:

$$\hat{\epsilon}^{[i+1]}_{S,R,D} = \hat{\epsilon}^{[i]}_{S,R,D} + \frac{N}{2\pi}\frac{\sum_{n=0}^{N-1} n\Im\big\{(u(n))^*\hat{S}^{[i+1]}(n)e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N}\big\}}{\sum_{n=0}^{N-1} n^2\Re\big\{(u(n))^*\hat{S}^{[i+1]}(n)e^{j2\pi\hat{\epsilon}^{[i]}_{S,R,D}n/N}\big\}}, \quad (23)$$

Next, by setting $\boldsymbol{\theta}_{S,R,D}$ and $\epsilon_{S,R,D}$ to their latest updated values, the updated value of $\hat{\mathbf{h}}_{S,R,D}$ at the $(i+1)$th iteration, $\hat{\mathbf{h}}^{[i+1]}_{S,R,D}$, is calculated. Based on the vectorial form of received signal in (7), the negative log likelihood function, in (8), can be written as

$$\log p(\mathbf{z}; \epsilon_{S,R,D}) = \mathrm{C1} + \parallel \mathbf{u} - \mathbf{E}_{S,R,D}\mathbf{P}_{S,R,D}\boldsymbol{\Gamma}\mathbf{h}_{S,R,D} \parallel^2 + \log p(\boldsymbol{\theta}_{S,R,D}), \quad (24)$$

where $\boldsymbol{\Gamma} \triangleq \mathrm{q}_1\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}$. Taking the derivative of (24) with respect to $\mathbf{h}_{S,R,D}$ and equating the result to zero, the estimate of $\mathbf{h}_{S,R,D}$ at the $(i+1)$th iteration is given by:

$$\hat{\mathbf{h}}^{[i+1]}_{S,R,D} = (\boldsymbol{\Gamma}^H\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^H\hat{\mathbf{P}}^H_{S,R,D}\hat{\mathbf{E}}^H_{S,R,D}\mathbf{u}, \quad (25)$$

where $\hat{\mathbf{E}}_{S,R,D} \triangleq \mathrm{diag}([e^{(j2\pi\hat{\epsilon}^{[i+1]}_{S,R,D}/N)\times 0}, e^{(j2\pi\hat{\epsilon}^{[i+1]}_{S,R,D}/N)}, \ldots, e^{(j2\pi\hat{\epsilon}^{[i+1]}_{S,R,D}/N)\times(N-1)}]^T)$, and $\hat{\epsilon}^{[i+1]}_{S,R,D}$ is obtained from (23), $\hat{\mathbf{P}}_{S,R,D} \triangleq \mathrm{diag}([e^{j\hat{\theta}^{[i+1]}_{S,R,D}(0)}, e^{j\hat{\theta}^{[i+1]}_{S,R,D}(1)}, \ldots, e^{j\hat{\theta}^{[i+1]}_{S,R,D}(N-1)}]^T)$, and $\hat{\boldsymbol{\theta}}^{[i+1]}_{S,R,D} \triangleq [\hat{\theta}^{[i+1]}_{S,R,D}(0), \hat{\theta}^{[i+1]}_{S,R,D}(1), \ldots, \hat{\theta}^{[i+1]}_{S,R,D}(N-1)]^T$ is obtained from (17).

Using (17), (23) and (25), the proposed algorithm iteratively updates the PHN and CFO and CIR parameters in E-step and M-step of the algorithm, respectively, and stops when the difference between likelihood functions of two iterations is smaller than a threshold ζ, i.e.,

$$\left| \sum_{n=0}^{N-1} \left\| u(n) - e^{j2\pi\hat{\epsilon}_{S,R,D}^{[i+1]} n/N} e^{j\hat{\theta}_{S,R,D}^{[i+1]}(n)} s^{[i+1]}(n) \right\|^2 - \sum_{n=0}^{N-1} \left\| u(n) - e^{j2\pi\hat{\epsilon}_{S,R,D}^{[i]} n/N} e^{j\hat{\theta}_{S,R,D}^{[i]}}(n) s^{[i]}(n) \right\|^2 \right| \leq \zeta. \tag{26}$$

The appropriate initialization of CFO and CIR, i.e., $\hat{\epsilon}_{S,R,D}^{[0]}$ and $\hat{\mathbf{h}}_{S,R,D}^{[0]}$, respectively, can help the proposed estimator to estimate the CIR, CFO, and PHN parameters in a few iterations. The initial CFO estimate is obtained by applying a linear search for the value of $\epsilon_{S,R,D}$ that minimizes the cost function, $\sum_{n=0}^{N-1} \parallel u(n) - e^{j2\pi\hat{\epsilon}_{S,R,D} n/N} \hat{s}(n) \parallel^2$, where $\hat{s}(n)$ is nth symbol of the vector $\hat{\mathbf{s}} \triangleq \mathrm{q}_1 \mathbf{F}^H \mathbf{D}_S \mathbf{F}_{(2L-1)} \hat{\mathbf{h}}_{S,R,D}$, $\hat{\mathbf{h}}_{S,R,D} \triangleq (\mathbf{\Gamma}^H \mathbf{\Gamma})^{-1} \mathbf{\Gamma}^H \hat{\mathbf{E}}_{S,R,D}^H \mathbf{u}$ and linear search is made with a coarse step size of 10^{-2}. Next, using the initial CFO estimate $\hat{\epsilon}_{S,R,D}^{[0]}$, initial channel estimate, $\hat{\mathbf{h}}_{S,R,D}^{[0]}$, is obtained by applying the equation, $\hat{\mathbf{h}}_{S,R,D}^{[0]} \triangleq (\mathbf{\Gamma}^H \mathbf{\Gamma})^{-1} \mathbf{\Gamma}^H (\hat{\mathbf{E}}_{S,R,D}^{[0]})^H \mathbf{u}$, where $\hat{\mathbf{E}}_{S,R,D}^{[0]} = \hat{\mathbf{E}}_{S,R,D}|_{\hat{\epsilon}_{S,R,D} = \hat{\epsilon}_{S,R,D}^{[0]}}$. The simulation results show that the proposed estimator always converges to true estimates, e.g., at $SNR = 20$ dB and with threshold $\zeta = 10^{-3}$, on average, the estimator converges after 2 iterations only.

IV. Hybrid Cramér-Rao bound

In this section, the HCRB for joint estimation of $\boldsymbol{\lambda} = [\boldsymbol{\theta}_{S,R,D}^T \ \mathbf{h}_{S,R,D}^T \ \epsilon_{S,R,D}]^T$ is presented. Note that the vector of parameters of interest, $\boldsymbol{\lambda}$, comprises both random and deterministic parameters, e.g., PHN, $\boldsymbol{\theta}_{S,R,D}$, is random while CIR, $\mathbf{h}_{S,R,D}$, and CFO, $\epsilon_{S,R,D}$, are deterministic parameters. Thus, HCRB instead of standard CRB is needed to be derived. The accuracy of estimating $\boldsymbol{\lambda}$ is lower bounded by the HCRB ($\mathbf{\Omega}$) as [13]

$$\mathbb{E}_{\mathbf{u}, \boldsymbol{\theta}_{S,R,D} | \mathbf{h}_{S,R,D}, \epsilon_{S,R,D}} \left[(\hat{\boldsymbol{\lambda}}(\mathbf{u}) - \boldsymbol{\lambda})(\hat{\boldsymbol{\lambda}}(\mathbf{u}) - \boldsymbol{\lambda})^T \right] \succeq \mathbf{\Omega}. \tag{27}$$

Let us define $\mathbf{\Omega} = \mathbf{B}^{-1}$ and $\mathbf{B}$ is an $(N + (2L-1) + 1) \times (N + (2L-1) + 1)$ hybrid information matrix (HIM), which is given in the following theorem.

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

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{b}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{b}_{23} \\ \mathbf{b}_{31} & \mathbf{b}_{32} & \mathrm{b}_{33} \end{bmatrix}, \tag{28}$$

where

- $\mathbf{B}_{11} = \mathbf{\Xi}_{D_{11}} + \mathbf{\Xi}_{P_{11}}$ is an $N \times N$ HIM for the estimation of $\boldsymbol{\theta}_{S,R,D}$,
- $\mathbf{\Xi}_{D_{11}} = \mathrm{diag}(\mathbf{h}_{S,R,D}^H \mathbf{F}_{2L-1}^H \mathbf{D}_S^H \mathbf{F}) \mathbf{\Sigma}^{-1} \mathrm{diag}(\mathbf{F}^H \mathbf{D}_S \mathbf{F}_{2L-1} \mathbf{h}_{S,R,D})$, $\mathbf{\Sigma} \triangleq \tilde{\sigma}_w^2 \mathbf{I}_N$, and $\mathbf{\Xi}_{P_{11}}$ is defined as in [14, eq.(19)].
- $\mathbf{B}_{22} = \mathbf{F}_{2L-1}^H \mathbf{D}_S^H \mathbf{\Sigma}^{-1} \mathbf{D}_S \mathbf{F}_{2L-1}$ is an $(2L-1) \times (2L-1)$ HIM for the estimation of $\mathbf{h}_{S,R,D}$,
- $\mathrm{b}_{33} = \mathbf{h}_{S,R,D}^H \mathbf{F}_{2L-1}^H \mathbf{D}_S^H \mathbf{F} \mathbf{\Sigma}^{-1} \mathbf{M} \mathbf{F}^H \mathbf{D}_S \mathbf{F}_{2L-1} \mathbf{h}_{S,R,D}$ is a scalar representing the hybrid information for the estimation of $\epsilon_{S,R,D}$,
- $\mathbf{B}_{12} = \mathbf{B}_{21}^T = \frac{-j}{\tilde{\sigma}_w^2} \mathrm{diag}(\mathbf{h}_{S,R,D}^H \mathbf{F}_{2L-1}^H \mathbf{D}_S^H \mathbf{F}) \mathbf{F}^H \mathbf{D}_S \mathbf{F}_{2L-1}$,
- $\mathbf{b}_{13} = \mathbf{b}_{31}^T = \frac{1}{\tilde{\sigma}_w^2} \mathrm{diag}(\mathbf{h}_{S,R,D}^H \mathbf{F}_{2L-1}^H \mathbf{D}_S^H \mathbf{F} \sqrt{\mathbf{M}}) \mathbf{F}^H \mathbf{D}_S \mathbf{F}_{2L-1} \mathbf{h}_{S,R,D}$,
- $\mathbf{b}_{23} = \mathbf{b}_{32}^T = j \mathbf{F}_{2L-1}^H \mathbf{D}_S^H \mathbf{F} \mathbf{\Sigma}^{-1} \sqrt{\mathbf{M}} \mathbf{F}^H \mathbf{D}_S \mathbf{F}_{2L-1} \mathbf{h}_{S,R,D}$,
- $\mathbf{M} \triangleq \mathrm{diag}\left(\left[(2\pi \frac{0}{N})^2, (2\pi \frac{1}{N})^2, \ldots, (2\pi \frac{N-1}{N})^2 \right]^T \right)$.

Finally, the HCRB, $\mathbf{\Omega}$, is given by the inverse of HIM, i.e., $\mathbf{\Omega} = \mathbf{B}^{-1}$. The detailed derivation is omitted due to space limitation. However, the derivation can be pursued by using the similar steps given in [15], where we derived the HCRB for joint estimation of PHN and CFO.

V. Simulation results and discussions

In this section, the performance of the proposed estimator is examined against the derived HCRB. 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 with different PHN variances, e.g. $\sigma_\delta^2 = [10^{-3}, 10^{-4}]$rad^2, where $\sigma_{\delta_S}^2 = \sigma_{\delta_D}^2 = \sigma_{\delta_{R,1}}^2 = \sigma_{\delta_{R,2}}^2 = \sigma_\delta^2$. Note that, $\sigma_\delta^2 = 10^{-3}$ rad^2, corresponds to a very high phase noise variance [16]. Since carrier frequency offsets from source to relays, $\epsilon_{S,R}$, are carried over to the destination, $\epsilon_{S,R}$ and $\epsilon_{R,D}$ 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). Similar to the parameter setting adopted in [8], the large-scale channel fading parametrization is set as γ=2, $d_{S,D}$=1, $d_{S,R}$=0.5, $d_{R,D}$=0.72, $g_{S,R}$=4 and $g_{R,D}$=1.9.

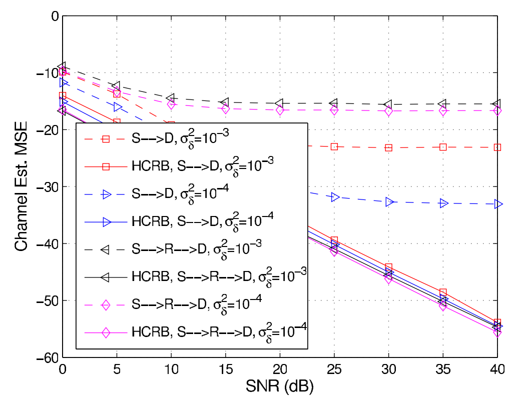

Fig. 2: MSE of channel estimation for the proposed estimator compared to HCRB for phase noise variance $\sigma_\delta^2 = [10^{-3}, 10^{-4}]$ rad^2.

Figs. 2, 3 and 4 plot the HCRB and mean-square error (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 estimators MSE are dependent on the variance of the PHN process and are lower for a lower PHN variance; 2) Figs. 2, 3 and 4 show that CIR, CFO and PHN 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. 3) The estimation performance and HCRB for the direct link, $S \to D$, are better than those of the single hop link (relay link), i.e., $S \to R \to D$. This due to the noise at the relays which is amplified and forwarded to the destination. 4) It is shown that the

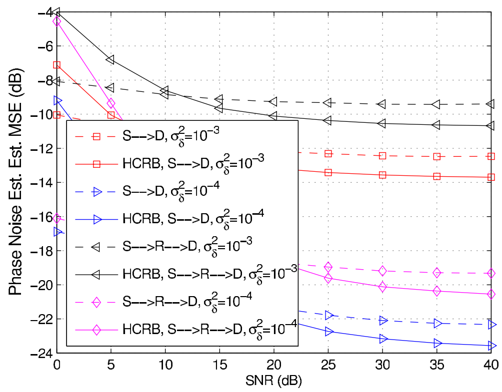

Fig. 3: MSE of phase noise estimation for the proposed estimator compared to HCRB for phase noise variance $\sigma_\delta^2 = [10^{-3}, 10^{-4}]$ rad^2.

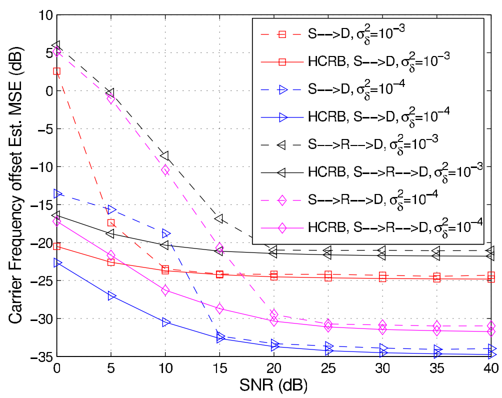

Fig. 4: MSE of frequency offset estimation for the proposed estimator compared to HCRB for phase noise variance $\sigma_\delta^2 = [10^{-3}, 10^{-4}]$ rad^2.

MSEs of the proposed estimator for both direct and relay links are close to their HCRLBs at mid-to-high SNRs; 5) In Fig. 3, the MSE of the proposed estimator is lower than the HCRB at lower SNR. This is due to the fact that the HCRB cannot be derived in closed-form while taking into account the range of CFO values, i.e., (-0.25. 0.25), while the estimator takes the advantage of this known prior estimation range. Thus, the HCRB is higher than the MSE of the proposed estimator at lower SNR.

VI. Conclusion

This is a first paper to address the joint estimation of channel, PHN, and CFO in cooperative OFDM systems. In this paper, a new iterative estimator, that jointly estimates the unknown channel gains, PHNs, and CFOs, for AF relaying cooperative OFDM systems has been proposed. The proposed estimator is 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 moderate-to-high SNRs.

Appendix A
Derivation of the mean and covariance matrix in (6)

Given $\mathbb{E}\{\mathbf{w}_R\} = \mathbf{0}_{N\times 1}$ and $\mathbb{E}\{\mathbf{w}_{D,2}\} = \mathbf{0}_{N\times 1}$, the mean of the received signal in (6), $\boldsymbol{\mu}_{\mathbf{y}_{D,2}}$, is calculated as

$$\begin{aligned}\boldsymbol{\mu}_{\mathbf{y}_{D,2}} &= \mathbb{E}\{\mathrm{q}_1\mathbf{E}_{S,R,D}\mathbf{P}_{S,R,D}\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\mathbf{h}_{S,R,D} \\ &\quad+\mathrm{q}_2\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{F}\mathbf{w}_R + \mathbf{w}_{D,2}\} \\ &= \mathrm{q}_1\mathbf{E}_{S,R,D}\mathbf{P}_{S,R,D}\mathbf{F}^H\mathbf{D}_S\mathbf{F}_{(2L-1)}\mathbf{h}_{S,R,D}. \qquad \text{(A.1)}\end{aligned}$$

The covariance matrix, $\boldsymbol{\Sigma}_{\mathbf{y}_{D,2}}$, as

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{y}_{D,2}} &= \mathbb{E}\{(\mathbf{y}_{D,2} - \boldsymbol{\mu}_{\mathbf{y}_{D,2}})(\mathbf{y}_{D,2} - \boldsymbol{\mu}_{\mathbf{y}_{D,2}})^H\} \\ &= \mathrm{q}_2^2\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{F}\ \mathbb{E}\{\mathbf{w}_R\mathbf{w}_R^H\}\ \mathbf{F}^H\mathbf{H}_{R,D}^H\mathbf{F}\mathbf{P}_{R_2,D}^H \\ &\quad \mathbf{E}_{R,D}^H + \mathbb{E}\{\mathbf{w}_{D,2}\mathbf{w}_{D,2}^H\} \\ &= \mathrm{q}_2^2\sigma_w^2\mathbf{E}_{R,D}\mathbf{P}_{R_2,D}\mathbf{F}^H\mathbf{H}_{R,D}\mathbf{H}_{R,D}^H\mathbf{F}\mathbf{P}_{R_2,D}^H\mathbf{E}_{R,D}^H + \sigma_w^2\mathbf{I}_N \\ &= (\mathrm{q}_2^2\sigma_w^2\sigma_h^2 + \sigma_w^2)\mathbf{I}_N, \qquad \text{(A.2)}\end{aligned}$$

where $\mathbb{E}\{\mathbf{H}_{R,D}\mathbf{H}_{R,D}^H\} = \sigma_h^2\mathbf{I}_N$, $\mathbf{F}^H\mathbf{F} = \mathbf{I}_N$, $\mathbf{P}_{R_2,D}\mathbf{P}_{R_2,D}^H = \mathbf{I}_N$ and $\mathbf{E}_{R_2,D}\mathbf{E}_{R_2,D}^H = \mathbf{I}_N$.

References

[1] Y. Jing and B. Hassibi, "Cooperative diversity in wireless relay networks with multiple-antenna nodes," in *International Symposium on Information ISIT*, Sept. 2005, pp. 815–819.

[2] N. J. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 3062–3080, Dec. 2004.

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

[4] A. Garcia Armada, "Understanding the effects of phase noise in orthogonal frequency division multiplexing (OFDM)," *IEEE Trans. Broadcast.*, vol. 47, no. 2, pp. 153 –159, Jun 2001.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] A. M. A. Demir and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Biomed. Circuits Syst.*, vol. 47, pp. 655–674, May 2000.

[10] E. P. Simon, L. Ros, H. Hijazi, and M. Ghogho, "Joint carrier frequency offset and channel estimation for OFDM systems via the EM algorithm in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 754 –765, Feb. 2012.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*. Signal Processing Series: Prentice Hall, 1993.

[12] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. Hoboken, NJ: Wiley, 2008.

[13] H. L. van Trees and K. L. Bell, *Bayesian bounds for parameter estimation and nonlinear filtering/tracking*. New York, USA: John Wiley and Sons, 2007.

[14] 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.

[15] O. H. Salim, A. A. Nasir, H. Mehrpouyan, and W. Xiang, "Phase noise and carrier frequency offset in OFDM systems: Joint estimation and hybrid Cramér-Rao lower bound," in *Proc. IEEE SPAWC*, Darmstadt, Germany, June 2013, pp. 649 – 653.

[16] H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. K. Karagiannidis, and T. Svensson, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sept. 2012.