Transceiver Design for Distributed STBC Based AF Cooperative Networks in the Presence of Timing and Frequency Offsets

Ali A. Nasir, *Student Member, IEEE*, Hani Mehrpouyan, *Member, IEEE*, Salman Durrani, *Senior Member, IEEE*, Steven D. Blostein, *Senior Member, IEEE*, Rodney A. Kennedy, *Fellow, IEEE*, and Björn Ottersten, *Fellow, IEEE*

***Abstract*—In multi-relay cooperative systems, the signal at the destination is affected by impairments such as multiple channel gains, *multiple timing offsets (MTOs)*, and *multiple carrier frequency offsets (MCFOs)*. In this paper we account for all these impairments and propose a new transceiver structure at the relays and a novel receiver design at the destination in *distributed space-time block code (DSTBC)* based *amplify-and-forward (AF)* cooperative networks. The Cramér-Rao lower bounds and a *least squares (LS)* estimator for the multi-parameter estimation problem are derived. In order to significantly reduce the receiver complexity at the destination, a *differential evolution (DE)* based estimation algorithm is applied and the initialization and constraints for the convergence of the proposed DE algorithm are investigated. In order to detect the signal from multiple relays in the presence of unknown channels, MTOs, and MCFOs, novel optimal and sub-optimal minimum mean-square error receiver designs at the destination node are proposed. Simulation results show that the proposed estimation and compensation methods achieve full diversity gain in the presence of channel and synchronization impairments in multi-relay AF cooperative networks.**

***Index Terms*—Cooperative communication, amplify-and-forward (AF), Cramér-Rao lower bound (CRLB), differential evolution (DE), distributed space-time block code (DSTBC), receiver design.**

I. Introduction

A. Motivation and Literature Survey

COOPERATIVE communication using relays has been recognized as a promising technique to improve the performance of wireless networks, reducing the need for physical deployment of multiple co-located antennas at the individual users [1], [2]. Amongst the different *relaying protocols* [3], amplify-and-forward (AF) is attractive due to its implementation simplicity, i.e., every relay node linearly processes the received signal before forwarding an appropriately scaled version to the destination.

Different relaying protocols have different cooperation strategies. The *repetition cooperative strategy* allows relays to transmit sequentially in different time slots, which leads to inefficient bandwidth utilization [4]. In order to improve bandwidth efficiency, this approach can be modified to allow for the relays to transmit simultaneously in multi-relay cooperative networks. However, depending on how the relays' signals superimpose at the destination, achieving full diversity gain is not guaranteed [5], [6]. The *distributed space-time block code (DSTBC) cooperative strategy* allows relays to transmit in the same time slot using the DSTBC structure [7]. Given its potential, the combination of AF relaying and DSTBC has attracted recent research interest as an efficient means of providing cooperative diversity [8], [9]. However, the performance benefits of DSTBC schemes depend on accurate synchronization amongst cooperating nodes, which increases the overhead at the relays.

The distributed nature of cooperative communication systems makes synchronization for AF relaying with DSTBC a non-trivial task. Unlike conventional multiple-input multiple-output systems, where co-located antenna elements may result in a single timing and a single carrier frequency offset, asynchronism amongst multiple distributed relays in cooperative networks gives rise to *multiple timing offsets (MTOs)* and *multiple carrier frequency offsets (MCFOs)* at the destination [6]. It has been shown that time asynchronism can destroy the elaborate structure of the DSTBCs, reducing the achievable diversity gain [10], [11]. Similarly, carrier frequency asynchronism can lead to a symbol to symbol channel time variation and to a violation of the essential requirement for successful DSTBC decoding and achieving spatial diversity [12].

Most of the literature on AF relaying with DSTBC has focused on analyzing the end-to-end system performance in terms of signal-to-noise ratio (SNR), outage probability, and symbol error probability while assuming perfect synchronization [8], [9]. Others have focused on the design of delay tolerant DSTBC [13] and carrier frequency offset tolerant space-frequency codes [14] for decode-and-forward (DF) systems. However, the algorithms in [13], [14] do not consider the joint estimation and compensation of MTOs and MCFOs and cannot be applied to AF relaying due to the codes that rely on the assumption that the received signals at the relays are free of any frequency and timing

Manuscript received May 25, 2012; revised December 21, 2012; accepted February 19, 2013. Date of publication April 12, 2013; date of current version May 21, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Milica Stojanovic.

A. A. Nasir, S. Durrani, and R. A. Kennedy are with the Research School of Engineering, The Australian National University, Australia (e-mail: ali.nasir@anu.edu.au; salman.durrani@anu.edu.au; rodney.kennedy@anu.edu.au).

H. Mehrpouyan is with the Department of Electrical and Computer Engineering, California State University, Bakersfield, CA 93309 USA (e-mail: hani.mehr@ieee.org).

S. D. Blostein is with the Department of Electrical and Computer Engineering, Queen's University, Kingston, Canada (e-mail: steven.blostein@queensu.ca).

B. Ottersten is with the Department of Interdisciplinary Center for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg (e-mail: bjorn.ottersten@uni.lu).

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/TSP.2013.2258015

offset. More importantly, [13] and [14] provide no means of estimating MTOs and MCFOs, even though they require these parameters to be obtained at the destination for successful decoding.

Algorithms have been proposed for estimation and compensation of channel gains [15], MTOs [16], or MCFOs [17] in DSTBC-orthogonal frequency division multiplexing (OFDM) AF relaying systems. However, the algorithms in [15]–[17] cannot obtain both timing and carrier frequency offsets *jointly* at the destination. Even though joint synchronization schemes for obtaining and compensating multiple channel gains, MTOs, and MCFOs for DF relaying DSTBC-OFDM based cooperative systems are available in the literature [18], [19], these solutions cannot be applied to AF relaying systems due to the fundamental differences in the relaying protocols and transceiver structures at the relays and destination. In addition, the algorithms in [15]–[19] exploit the cyclic prefix and the frequency domain structure of the signal, which is specific to OFDM systems and depending on the number of sub-carriers used, the carrier frequency offset acquisition range of the algorithms in [15]–[19] is very limited.

For general AF relaying with DSTBC, channel estimation and equalization [8], joint channel and MCFO [20]–[22], or joint channel and MTO estimation and compensation [10] are addressed in the literature. However, these schemes do not jointly estimate and compensate multiple channel gains, MTOs and MCFOs, i.e., they estimate and compensate either MCFOs while assuming perfect timing synchronization [20], [21] or vice versa [10]. As shown in Section VI-B in this paper, such an approach does not allow for successful decoding of the received signal at the destination in the presence of both MTOs and MCFOs. Finally, [6] addresses the problem of timing and carrier synchronization in AF and DF cooperative networks. However, the transceiver designs and estimation algorithms proposed in [6] cannot be applied to DSTBC-AF cooperative networks due to the particular processing required at the relays and destination to enable transmission and decoding of DSTBC in AF cooperative networks. Consequently, the transceiver design in [6] does not achieve full cooperative diversity.

B. Contributions

In this paper, complete synchronization, i.e., joint estimation and compensation of multiple channel gains, MTOs, and MCFOs, in DSTBC-AF relaying cooperative networks is proposed. The processing at each relay's transceiver consists of estimating and then compensating the source to relay timing and carrier frequency offsets before forwarding the linearly processed signals to the destination using DSTBC. At the destination, the channel gains, MTOs, and MCFOs are jointly estimated using transmitted training sequences and the proposed *least squares (LS)* or *differential evolution (DE)* based estimators [23]. The DE algorithm, which is an iterative method, significantly reduces computational complexity at the destination. Next, a novel minimum mean-square error (MMSE) receiver for compensating the effect of these impairments at the destination is derived. In summary, the main contributions of this work are as follows:

- A new transceiver structure at the relays and a receiver design at the destination for achieving synchronization in DSTBC-AF relaying cooperative network are proposed.
- New *Cramér-Rao lower bounds (CRLBs)* for joint estimation of multiple system parameters at the destination are derived. Aside from benchmarking the performance of the proposed estimators, these bounds are applied to determine the statistics of estimation errors for timing offset, frequency offset, and channels, which is in turn used in the design of the MMSE receivers at the destination.
- An LS estimator for joint estimation of channels, MTOs, MCFOs is formulated. Next, in order to significantly reduce the computational complexity associated with the joint estimation problem, a DE based algorithm is applied. A parameterization of the proposed DE estimator is provided that achieves fast convergence. Simulation results show that the mean square error (MSE) performances of both the LS and the DE estimators are close to the CRLB at moderate-to-high SNRs.
- An MMSE receiver for compensating the effect of impairments and detecting the signal from the relays at the destination is derived. In order to reduce overhead, a low complexity version denoted by *L-MMSE* is also proposed that can detect the received signals using only the overall channel gains, MTOs, and MCFOs estimates obtained at the destination.
- Extensive simulations are carried out to investigate the performance of the overall proposed transceiver design in DSTBC-AF cooperative networks for different numbers of relays and system design parameters. It is shown that the proposed DSTBC-AF relaying system achieves full spatial diversity gain for 2 and 4 relay networks and outperforms existing cooperative strategies in the presence of MTOs, MCFOs, and unknown channels.

C. Organization

The remainder of the paper is organized as follows: Section II details the system model for the proposed relay transceiver and destination receiver. Section III derives the CRLBs for the multiple parameter estimation problem. Section IV presents the proposed LS and DE based estimators. Section V derives the compensation algorithms denoted by MMSE and L-MMSE. Section VI presents numerical and simulation results that study the MSE and bit-error rate (BER) performances of the proposed estimators and compensation algorithms, respectively. Finally, Section VII concludes the paper and summarizes its key findings.

D. Notations

Superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)'$, and $(\cdot)''$ denote the conjugate, the transpose, the conjugate transpose, the first derivative, and second derivative operators, respectively. $\odot$ and $\otimes$ stand for the Hadamard and Kronecker products, respectively. $\mathbb{E}_x\{\cdot\}$ denotes the expectation operator with respect to the variable x. $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex quantity. Symbols with superscripts $(\cdot)^{[\mathrm{sr}]}$ and $(\cdot)^{[\mathrm{rd}]}$ denote the source to relay and relay to destination parameters, respectively. Symbols with superscripts $(\cdot)^{[\mathrm{TP}]}$ and $(\cdot)^{[\mathrm{DTP}]}$ denote the signals in training and data transmission periods, respectively. The operator, $\hat{x}$ represents the estimated value of x. $\mathcal{N}(\mu,\sigma^2)$ and $\mathcal{CN}(\mu,\sigma^2)$ denote the real and complex Gaussian distributions with mean μ and variance σ^2, respectively. $\mathcal{U}(-x,x)$ denotes a uniform

distribution between $-x$ and x. Boldface small letters, $\mathbf{x}$ and boldface capital letters, $\mathbf{X}$ are used for vectors and matrices, respectively. $[\mathbf{X}]_{x,y}$ represents the entry in row x and column y of $\mathbf{X}$ and $[\mathbf{X}]_{a:b,x:y}$ denotes a submatrix formed by the rows, a to b, and columns, x to y, of the matrix $\mathbf{X}$. $\mathbf{I}_X$, $\mathbf{0}_{X\times X}$, and $\mathbf{1}_{X\times X}$ denote $X \times X$ identity, all zero, and all one matrices, respectively. $\|\mathbf{x}\|$ represents the ℓ_2 norm of a vector $\mathbf{x}$ and $\mathrm{Tr}\{\mathbf{X}\}$ denotes the trace of $\mathbf{X}$. $\mathrm{diag}(\mathbf{x})$ is used to denote a diagonal matrix, where its diagonal elements are given by the vector $\mathbf{x}$ and $\mathrm{diag}([\mathbf{X_1}, \mathbf{X_2}, \ldots])$ represents a block diagonal matrix, where $[\mathbf{X_1}, \mathbf{X_2}, \ldots]$ are the diagonal matrices. Finally, $\mathbf{\Sigma_x}$, $\mathbf{\Phi_{x,y}} \triangleq \mathbb{E}\{\mathbf{x}\mathbf{y}^H\}$, and $\mathbf{\Phi_x} \triangleq \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ denote the covariance matrix of $\mathbf{x}$, the cross correlation matrix between the vectors $\mathbf{x}$ and $\mathbf{y}$, and the correlation matrix of $\mathbf{x}$, respectively.

II. System Model and Transceiver Design

As shown in Fig. 1, a half-duplex AF cooperative system with one source node, $\mathbb{S}$, K relays, $\mathbb{R}_1, \ldots, \mathbb{R}_K$, and a single destination node, $\mathbb{D}$ is considered. Each node is equipped with a single omnidirectional antenna. The channels from the source to the kth relay and the kth relay to the destination are denoted by h_k and f_k, respectively. The index $k = 1, \ldots, K$ is used for the K relays. In Fig. 1, τ_k and ν_k are used to denote timing and carrier frequency offsets, respectively. Throughout this paper, the following set of system design assumptions is considered:

A1. Each transmission frame from source to relays and relays to destination is comprised of two periods: a *training period (TP)* followed by a *data transmission period (DTP)*. Without loss of generality, it is assumed that during the TP, unit-amplitude phase shift keying (PSK) *training signals (TSs)* are transmitted from the source to the kth relay and from the kth relay to the destination, $\forall\, k$. The TSs from all the relays are linearly independent. Such TSs have also been considered previously, e.g., in [10], [20].[1]

A2. Quasi-static and frequency flat fading channels are considered, i.e., the channel gains do not change over the length of a frame but change from frame to frame according to a complex Gaussian distribution, $\mathcal{CN}(0, \sigma^2)$. The use of such channels is motivated by prior research in this field [10], [20], [24], [25]. Moreover, the assumption of frequency flat channels can be broadened to frequency selective channels by employing OFDM.

A3. The timing and carrier frequency offsets are modeled as *unknown* deterministic parameters over the frame length, which is similar to the approach adopted in [10], [20], and [24].

The proposed system model at the relays and destination is detailed below.

A. Proposed Transceiver Model at Relays

The block diagram of the AF transceiver at the kth relay is shown in Fig. 2. During the first time slot, the source transmits the training and data symbols to all relays. The received signal at the kth relay, $r_k(t)$, is down converted by oscillator frequency ω_k. The received signal is oversampled by a factor N, such that $T = T_s N$, where T is the symbol duration and T_s is the sampling period. L denotes the number of training symbols during TP and the number of data symbols during DTP. For clarity, here, the indices $n = 0, 1, \ldots, L-1$ and $i = 0, 1, \ldots, LN-1$ are used to denote the symbols and the T_s-spaced samples, respectively. Each relay uses the training signal from the source node to estimate the source-to-relay carrier frequency and timing offsets. The proposed transceiver design and signal model corresponding to the TP and DTP are detailed in the following two subsections.

[1]The detailed design of the training sequences is outside the scope of this paper.

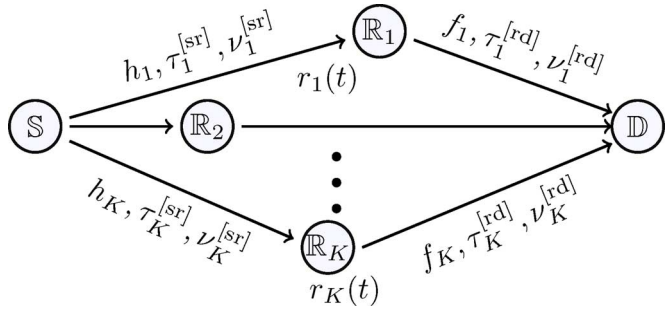


Fig. 1. System model for AF cooperative network.

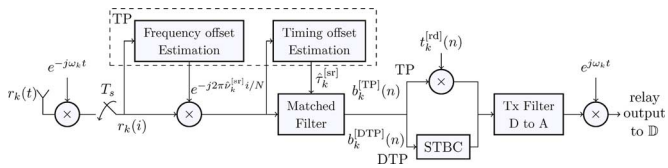


Fig. 2. Proposed AF transceiver design at the kth relay transceiver.

1) Training Model at Relays: During TP, the sampled received signal vector at the kth relay, $\mathbf{r}_k \triangleq [r_k(0), \ldots, r_k(LN-1)]^T$, is given by

$$\mathbf{r}_k = h_k \mathbf{\Lambda}_k^{[\mathrm{sr}]} \mathbf{G}_k^{[\mathrm{sr}]} \mathbf{t}^{[\mathrm{sr}]} + \mathbf{u}_k, \quad k = 1, \ldots, K \tag{1}$$

where

- h_k denotes the *unknown* channel gain from source to the kth relay that changes from one frame to another frame following the distribution, $h_k \sim \mathcal{CN}(0, \sigma_h^2)$,
- $\mathbf{\Lambda}_k^{[\mathrm{sr}]} \triangleq \mathrm{diag}([e^{j2\pi\nu_k^{[\mathrm{sr}]}(0)/N}, \ldots, e^{j2\pi\nu_k^{[\mathrm{sr}]}(LN-1)/N}])$ is a diagonal $LN \times LN$ matrix, $\nu_k^{[\mathrm{sr}]}$ denotes the *unknown* carrier frequency offset, normalized by the symbol duration T, between source and the kth relay,
- $\mathbf{G}_k^{[\mathrm{sr}]}$ is the $LN \times L$ matrix of the samples of the pulse shaping filter such that $[\mathbf{G}_k^{[\mathrm{sr}]}]_{i,n} \triangleq g_{rrc}(iT_s - nT - \tau_k^{[\mathrm{sr}]}T)$, $\tau_k^{[\mathrm{sr}]}$ denotes the normalized *unknown* timing offset between source and the kth relay, $g_{rrc}(t)$ stands for the root raised-cosine pulse shaping function,
- $\mathbf{t}^{[\mathrm{sr}]} \triangleq [t^{[\mathrm{sr}]}(0), \ldots, t^{[\mathrm{sr}]}(L-1)]^T$ is the TS transmitted from source to relays, and
- $\mathbf{u}_k \triangleq [u_k(0), \ldots, u_k(LN-1)]^T$, $u_k(i) \forall\, i$, denotes the zero-mean complex additive white Gaussian noise (AWGN) at the ith sample of the received signal, i.e., $u_k(i) \sim \mathcal{CN}(0, \sigma_{u_k}^2)$.

Note that the noise terms at the relays are assumed to be mutually uncorrelated, i.e., $\mathbb{E}\{\mathbf{u}_k^H \mathbf{u}_{\bar{k}}\} = 0$, for $k \neq \bar{k}$.

The estimates of the timing and frequency offsets at the kth relay, $\hat{\tau}_k^{[\mathrm{sr}]}$ and $\hat{\nu}_k^{[\mathrm{sr}]}$, respectively, are used for matched filtering and frequency offset correction at the kth relay, as shown in the Fig. 2. Standard estimation and synchronization techniques for

carrier and timing synchronization in point-to-point communication systems, e.g., [26], can be applied to obtain $\hat{\tau}_k^{[\mathrm{sr}]}$ and $\hat{\nu}_k^{[\mathrm{sr}]}$, $\forall\, k$, and will not be discussed here.

Let the timing and frequency offset estimation errors between source and the kth relay be denoted by $\delta_{\tau_k^{[\mathrm{sr}]}} = \tau_k^{[\mathrm{sr}]} - \hat{\tau}_k^{[\mathrm{sr}]}$ and $\delta_{\nu_k^{[\mathrm{sr}]}} = \nu_k^{[\mathrm{sr}]} - \hat{\nu}_k^{[\mathrm{sr}]}$, respectively. According to the results in [10], $\delta_{\tau_k^{[\mathrm{sr}]}}$ and $\delta_{\nu_k^{[\mathrm{sr}]}}$ are Gaussian distributed, i.e., $\delta_{\tau_k^{[\mathrm{sr}]}} \sim \mathcal{N}(0, \sigma^2_{\tau_k^{[\mathrm{sr}]}})$ and $\delta_{\nu_k^{[\mathrm{sr}]}} \sim \mathcal{N}(0, \sigma^2_{\nu_k^{[\mathrm{sr}]}})$ [10], where $\sigma^2_{\tau_k^{[\mathrm{sr}]}}$ and $\sigma^2_{\nu_k^{[\mathrm{sr}]}}$ are timing and frequency offset estimation error variances, respectively. In this paper, $\sigma^2_{\tau_k^{[\mathrm{sr}]}}$ and $\sigma^2_{\nu_k^{[\mathrm{sr}]}}$ are set to their respective lower bounds in point-to-point systems given in ([26], Ch. 6). Taking into account the timing and frequency offset estimation errors, the vector of imperfectly synchronized received signals at the kth relay, $\mathbf{b}_k^{[\mathrm{TP}]} \triangleq [b_k^{[\mathrm{TP}]}(0), \ldots, b_k^{[\mathrm{TP}]}(L-1)]^T$, as shown in Fig. 2, can be written as [26]

$$\mathbf{b}_k^{[\mathrm{TP}]} = h_k \mathbf{x}_k^{[\mathrm{TP}]} + \mathbf{n}_k, \tag{2}$$

where

- $\mathbf{x}_k^{[\mathrm{TP}]} \triangleq \mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} \mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} \mathbf{t}^{[\mathrm{sr}]}$, $\mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} \triangleq \mathrm{diag}([e^{j2\pi\delta_{\nu_k^{[\mathrm{sr}]}}(0)}, \ldots, e^{j2\pi\delta_{\nu_k^{[\mathrm{sr}]}}(L-1)}])$ is an $L \times L$ diagonal matrix, $\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}$ is an $L \times L$ matrix, such that $[\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}]_{n,\bar{n}} \triangleq g_{rc}(nT - \bar{n}T - \delta_{\tau_k^{[\mathrm{sr}]}} T)$, for $n, \bar{n} = 0, \ldots, L-1$, $g_{rc}(t)$ stands for the raised-cosine function, which results from the convolution of $g_{rrc}(t)$ and root raised-cosine matched filter,
- $\mathbf{n}_k \triangleq \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]} \hat{\mathbf{\Lambda}}_k^{[\mathrm{sr}]} \mathbf{u}_k$ is the matched filtered noise, $\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}$ is an $L \times LN$ matrix such that $[\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}]_{n,i} \triangleq g_{rrc}(nT - iT_s - \hat{\tau}_k^{[\mathrm{sr}]} T)$, and $\hat{\mathbf{\Lambda}}_k^{[\mathrm{sr}]}$ is an $LN \times LN$ matrix such that $\hat{\mathbf{\Lambda}}_k^{[\mathrm{sr}]} \triangleq \mathbf{\Lambda}_k^{[\mathrm{sr}]}|_{\nu_k^{[\mathrm{sr}]} = \hat{\nu}_k^{[\mathrm{sr}]}}$.

Subsequently, the kth relay applies its unique training signal, $\mathbf{t}_k^{[\mathrm{rd}]} \triangleq [t_k^{[\mathrm{rd}]}(0), \ldots, t_k^{[\mathrm{rd}]}(L-1)]^T$ to the matched filtered output, $\mathbf{b}_k^{[\mathrm{TP}]}$, performs the pulse shaping operation, and up converts the analog signal by the oscillator frequency ω_k (see Fig. 2). Note that the proposed AF transceiver in Fig. 2 is less complex than a DF relaying transceiver since there is no channel estimation, equalization, and decoding blocks after the matched filtering block.

2) Data Transmission Model at Relays: During DTP, the modulated data symbol vector, $\mathbf{s} \triangleq [s(0), s(1), \ldots, s(L-1)]^T$, is transmitted from the source to the relays. The sampled received signal vector at the kth relay during the DTP is given by (1), where $\mathbf{t}^{[\mathrm{sr}]}$ is replaced by the data symbol vector, $\mathbf{s}$. After performing timing and frequency synchronization, the matched filtered signal vector at the kth relay in the DTP, $\mathbf{b}_k^{[\mathrm{DTP}]} \triangleq [b_k^{[\mathrm{DTP}]}(0), \ldots, b_k^{[\mathrm{DTP}]}(L-1)]^T$, is given by [26]

$$\mathbf{b}_k^{[\mathrm{DTP}]} = h_k \mathbf{x}_k^{[\mathrm{DTP}]} + \mathbf{n}_k, \tag{3}$$

where $\mathbf{x}_k^{[\mathrm{DTP}]} \triangleq \mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} \mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} \mathbf{s}$ and $\mathbf{n}_k$ are defined below (2). During the DTP, before forwarding its signal to the destination, each relay applies a DSTBC to the corresponding matched filter output, performs the pulse shaping operation, and up converts the analog signal by the oscillator frequency ω_k (see Fig. 2).

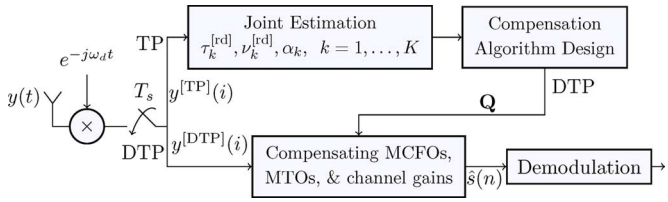


Fig. 3. Proposed AF receiver design at the destination.

B. Proposed Receiver Model at Destination

The block diagram for the proposed receiver at the destination is shown in Fig. 3. As described in Section II-A, each relay applies its independent training and DSTBC. Consequently, the received signal at the destination, $y(t)$, is a superposition of signals from different relays. At the destination, $y(t)$ is down converted by the oscillator frequency ω_d and then oversampled by the factor N. Furthermore, multiple channel gains, MTOs, and MCFOs are estimated at the receiver using the known and linearly independent TSs transmitted from the relays. The proposed receiver design and system model at the destination during the TP and the DTP is outlined in the following two subsections.

1) Training Model at Destination: During the TP, the sampled received signal vector at the destination, $\mathbf{y}^{[\mathrm{TP}]} \triangleq [y^{[\mathrm{TP}]}(1), \ldots, y^{[\mathrm{TP}]}(LN-1)]^T$, is given by

$$\mathbf{y}^{[\mathrm{TP}]} = \sum_{k=1}^{K} \zeta_k f_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \left(\mathbf{b}_k^{[\mathrm{TP}]} \odot \mathbf{t}_k^{[\mathrm{rd}]} \right) + \mathbf{w}, \tag{4}$$

where

- f_k denotes the *unknown* channel gain from the kth relay to the destination, which changes from frame to frame following the distribution $f_k \sim \mathcal{CN}(0, \sigma_f^2)$,
- $\zeta_k = 1/\sqrt{\sigma_h^2 + \sigma_{u_k}^2}$ satisfies the kth relay's power constraint,
- $\mathbf{\Lambda}_k^{[\mathrm{rd}]} \triangleq \mathrm{diag}([e^{j2\pi\nu_k^{[\mathrm{rd}]}(0)/N}, \ldots, e^{j2\pi\nu_k^{[\mathrm{rd}]}(LN-1)/N}])$ is an $LN \times LN$ matrix, $\nu_k^{[\mathrm{rd}]}$ denotes the normalized *unknown* frequency offset from the kth relay to the destination,
- $[\mathbf{G}_k^{[\mathrm{rd}]}]_{i,n} \triangleq g_{rrc}(iT_s - nT - \tau_k^{[\mathrm{rd}]} T)$ is an $LN \times L$ matrix, $\tau_k^{[\mathrm{rd}]}$ denotes the normalized fractional *unknown* timing offset between the kth relay and destination, and
- $\mathbf{w} \triangleq [w(0), \ldots, w(LN-1)]^T$ and $w(i) \forall\, i$, denotes the zero-mean complex AWGN at the ith sample of the received signal, i.e., $w(i) \sim \mathcal{CN}(0, \sigma_w^2)$.

Substituting (1) into (4), $\mathbf{y}^{[\mathrm{TP}]}$ can be written as

$$\mathbf{y}^{[\mathrm{TP}]} = \sum_{k=1}^{K} \alpha_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \mathbf{p}_k + \sum_{k=1}^{K} \beta_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \mathbf{v}_k + \mathbf{w}, \tag{5}$$

where $\mathbf{p}_k \triangleq \mathbf{x}_k^{[\mathrm{TP}]} \odot \mathbf{t}_k^{[\mathrm{rd}]}$, $\mathbf{v}_k \triangleq \mathbf{n}_k \odot \mathbf{t}_k^{[\mathrm{rd}]}$, $\alpha_k \triangleq \zeta_k f_k h_k$ and $\beta_k \triangleq \zeta_k f_k$. The general model in (5) can be written compactly in matrix form as

$$\mathbf{y}^{[\mathrm{TP}]} = \mathbf{\Xi}\mathbf{A}\mathbf{p} + \mathbf{\Xi}\mathbf{B}\mathbf{v} + \mathbf{w}. \tag{6}$$

where

- $\mathbf{\Xi} \triangleq [\mathbf{\Xi}_1, \ldots, \mathbf{\Xi}_K]$ is an $LN \times LK$ matrix, $\mathbf{\Xi}_k \triangleq \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]}$ is the $LN \times L$ matrix of the kth relay's frequency offset, $\nu_k^{[\mathrm{rd}]}$, and timing offset, $\tau_k^{[\mathrm{rd}]}$,

- $\mathbf{A} \triangleq \mathrm{diag}(\alpha_1, \ldots, \alpha_K) \otimes \mathbf{I}_L$ is the $LK \times LK$ matrix of the overall channel gains from source-to-relays-to-destination,
- $\mathbf{B} \triangleq \mathrm{diag}(\beta_1, \ldots, \beta_K) \otimes \mathbf{I}_L$ is the $LK \times LK$ matrix of relay-to-destination channels,
- $\mathbf{p} \triangleq [\mathbf{p}_1^T, \ldots, \mathbf{p}_K^T]^T$, and $\mathbf{v} \triangleq [\mathbf{v}_1^T, \ldots, \mathbf{v}_K^T]^T$.

Using the training received signal, $\mathbf{y}^{[\mathrm{TP}]}$, and known TSs, $\mathbf{t}^{[\mathrm{sr}]}$ and $\mathbf{t}_k^{[\mathrm{rd}]}$, $\forall\, k$, the unknown parameters α_k, $\tau_k^{[\mathrm{rd}]}$ and $\nu_k^{[\mathrm{rd}]}$, $\forall\, k$, can be jointly estimated at the destination (see Section IV).

2) Data Transmission Model at Destination: During the DTP, the sampled received signal at the destination receiver, $\mathbf{y}^{[\mathrm{DTP}]} \triangleq [y^{[\mathrm{DTP}]}(1), \ldots, y^{[\mathrm{DTP}]}(LN-1)]^T$, is given by

$$\mathbf{y}^{[\mathrm{DTP}]} = \sum_{k=1}^{K} \beta_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \mathbf{\Omega}_k \mathbf{b}_k^{[\mathrm{DTP}]} + \mathbf{w}, \tag{7}$$

where $\mathbf{\Omega}_k$ is the predefined $L \times L$ STBC matrix at the kth relay. Substituting (3) into (7), the signal model in (7) can be written in matrix form as

$$\mathbf{y}^{[\mathrm{DTP}]} = \mathbf{\Xi}\mathbf{A}\mathbf{\Omega}\mathbf{x}^{[\mathrm{DTP}]} + \mathbf{\Xi}\mathbf{B}\mathbf{\Omega}\mathbf{n} + \mathbf{w}, \tag{8}$$

where $\mathbf{\Omega} \triangleq \mathrm{diag}(\mathbf{\Omega}_1, \ldots, \mathbf{\Omega}_K)$, $\mathbf{x}^{[\mathrm{DTP}]} \triangleq [(\mathbf{x}_1^{[\mathrm{DTP}]})^T, \ldots, (\mathbf{x}_K^{[\mathrm{DTP}]})^T]^T$, and $\mathbf{n} \triangleq [\mathbf{n}_1^T, \ldots, \mathbf{n}_K^T]^T$. The estimates obtained during the TP are used to design matrix, $\mathbf{Q}$, which is used to compensate the effect of multiple channel gains, MTOs, and MCFOs and enable the detection of signals from relays at the destination (see Section V).

Note that unlike [10], the proposed processing structure at the relays in Fig. 2 is not based on the assumption of perfect timing and frequency offset estimation and perfect matched filtering at the relays [27]. More importantly, the authors in [10] did not consider the presence of MCFOs in their system.

III. Cramer-Rao Lower Bounds

In this section, *Fisher information matrix (FIM)* and CRLBs for joint estimation of multiple channel gains, MTOs, and MCFOs for AF-DSTBC cooperative systems are derived in closed-form.

Based on the assumptions in Section II, the AWGN at the relays, $\mathbf{u}_k$, $\forall\, k$, and destination, $\mathbf{w}$ in (4), are mutually independent. Accordingly, the received training signal at the destination, $\mathbf{y}^{[\mathrm{TP}]}$ in (6), is a circularly symmetric complex Gaussian random variable, $\mathbf{y}^{[\mathrm{TP}]} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}, \mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}})$, with mean $\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}$ and covariance matrix $\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}$ that are given by (see Appendix A)

$$\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}} = \sum_{k=1}^{K} \alpha_k \mathbf{\Xi}_k \mathbf{t}_k = \mathbf{\Xi}\mathbf{A}\mathbf{t}, \quad \text{and} \tag{9a}$$

$$\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}} = \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{\Xi}^H + \mathbf{\Sigma}_{\mathbf{w}}, \tag{9b}$$

respectively. In (9), $\mathbf{t}_k \triangleq (\mathbf{t}^{[\mathrm{sr}]} \odot \mathbf{t}_k^{[\mathrm{rd}]})$, $\mathbf{t} \triangleq [\mathbf{t}_1^T, \ldots, \mathbf{t}_K^T]^T$, $\mathbf{\Sigma}_{\mathbf{w}} \triangleq \sigma_w^2 \mathbf{I}_{LN}$, $\mathbf{\Sigma}_{\mathbf{p}} \triangleq \mathbb{E}\{(\mathbf{p}-\mathbf{t})(\mathbf{p}-\mathbf{t})^H\}$, $\mathbf{\Sigma}_{\mathbf{v}} \triangleq \mathbb{E}\{\mathbf{n}\mathbf{n}^H\} \odot \mathbf{t}^{[\mathrm{rd}]}(\mathbf{t}^{[\mathrm{rd}]})^H$, $\mathbf{t}^{[\mathrm{rd}]} \triangleq [(\mathbf{t}_1^{[\mathrm{rd}]})^T, \ldots, (\mathbf{t}_K^{[\mathrm{rd}]})^T]^T$, and $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{\Xi}$ are defined below (6).

The first step in determining the FIM is to formulate the parameter vector of interest. The destination node must estimate the MTOs and MCFOs from the relays to the destination, $\boldsymbol{\tau}^{[\mathrm{rd}]} \triangleq [\tau_1^{[\mathrm{rd}]}, \ldots, \tau_K^{[\mathrm{rd}]}]^T$ and $\boldsymbol{\nu}^{[\mathrm{rd}]} \triangleq [\nu_1^{[\mathrm{rd}]}, \ldots, \nu_K^{[\mathrm{rd}]}]^T$, respectively, and the overall channel gains from source to relays to destination, $\boldsymbol{\alpha} \triangleq [\alpha_1, \ldots, \alpha_K]^T$. As a result, the parameter vector of interest, $\boldsymbol{\lambda}$, is given by

$$\boldsymbol{\lambda} \triangleq \left[\Re\{\boldsymbol{\alpha}\}^T, \Im\{\boldsymbol{\alpha}\}^T, \left(\boldsymbol{\nu}^{[\mathrm{rd}]}\right)^T, \left(\boldsymbol{\tau}^{[\mathrm{rd}]}\right)^T\right]^T. \tag{10}$$

Note that $\mathbf{B}$ in (6) does not need to be estimated since it only scales the AWGN forwarded from the relays.

Theorem 1: Based on the proposed AF-DSTBC relaying scheme, the FIM for the estimation of $\boldsymbol{\lambda}$ is given by (11), shown at the bottom of the page, where

- $\mathbf{\Psi} \triangleq [\mathbf{\Xi}_1\mathbf{t}_1, \ldots, \mathbf{\Xi}_K\mathbf{t}_K]$ is an $LN \times K$ matrix,
- $\mathbf{\Gamma} \triangleq [\mathbf{\Lambda}_1(\mathbf{G}_K^{[\mathrm{rd}]})'\mathbf{t}_1, \ldots, \mathbf{\Lambda}_K(\mathbf{G}_K^{[\mathrm{rd}]})'\mathbf{t}_K]$ is an $LN \times K$ matrix,
- $(\mathbf{G}_k^{[\mathrm{rd}]})' \triangleq \frac{\partial \mathbf{G}_k^{[\mathrm{rd}]}}{\partial \tau_k^{[\mathrm{rd}]}}$, $\forall\, k$, is an $LN \times L$ matrix,
- $\mathbf{D} \triangleq 2\pi \times \mathrm{diag}\{0, 1, 2, \ldots, LN-1\}$ is an $LN \times LN$ matrix,
- $\boldsymbol{\mathcal{A}} \triangleq \mathrm{diag}\{\alpha_1, \ldots, \alpha_K\}$ is a $K \times K$ matrix,
- $\mathbf{\Pi}$ is a $4K \times 4K$ matrix, such that $[\mathbf{\Pi}]_{m,\bar{m}} = \mathrm{Tr}\{\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},\bar{m}}\}$ for $m, \bar{m} = 1, \ldots, 4K$,
- $\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m}$ is an $LN \times LN$ matrix that is defined at the bottom of the page,

$$\mathbf{F} = \begin{bmatrix} 2\Re\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & -2\Im\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & -2\Im\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Psi}\boldsymbol{\mathcal{A}}\right\} & 2\Re\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Gamma}\boldsymbol{\mathcal{A}}\right\} \\ 2\Im\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & 2\Re\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & 2\Re\left\{\mathbf{\Psi}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Psi}\boldsymbol{\mathcal{A}}\right\} & 2\Im\left\{\mathbf{\Psi}^H\Sigma_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Gamma}\boldsymbol{\mathcal{A}}\right\} \\ 2\Im\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Psi}^H\mathbf{D}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & 2\Re\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Psi}^H\mathbf{D}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & 2\Re\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Psi}^H\mathbf{D}\Sigma_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Psi}\boldsymbol{\mathcal{A}}\right\} & 2\Im\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Psi}^H\mathbf{D}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Gamma}\boldsymbol{\mathcal{A}}\right\} \\ 2\Re\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Gamma}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & -2\Im\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Gamma}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Psi}\right\} & -2\Im\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Gamma}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Psi}\boldsymbol{\mathcal{A}}\right\} & 2\Re\left\{\boldsymbol{\mathcal{A}}^H\mathbf{\Gamma}^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Gamma}\boldsymbol{\mathcal{A}}\right\} \end{bmatrix} + \mathbf{\Pi}, \tag{11}$$

$$\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m} \triangleq \begin{cases} \mathbf{\Xi}\mathbf{U}_{\alpha_m}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{U}_{\alpha_m}^H\mathbf{\Xi}^H, & m = 1, \ldots, K \\ j\mathbf{\Xi}\mathbf{U}_{\alpha_{m-K}}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H - j\mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{U}_{\alpha_{m-K}}^H\mathbf{\Xi}^H, & m = K+1, \ldots, 2K \\ \mathbf{U}_{\nu_{m-2K}^{[\mathrm{rd}]}}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{U}_{\nu_{m-2K}^{[\mathrm{rd}]}}^H + \mathbf{U}_{\nu_{m-2K}^{[\mathrm{rd}]}}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{U}_{\nu_{m-2K}^{[\mathrm{rd}]}}^H, & m = 2K+1, \ldots, 3K \\ \mathbf{U}_{\tau_{m-3K}^{[\mathrm{rd}]}}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{U}_{\tau_{m-3K}^{[\mathrm{rd}]}}^H + \mathbf{U}_{\tau_{m-3K}^{[\mathrm{rd}]}}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{U}_{\tau_{m-3K}^{[\mathrm{rd}]}}^H, & m = 3K+1, \ldots, 4K \end{cases}$$

- $\mathbf{U}_{\alpha_k} \triangleq \mathrm{diag}(\mathbf{0}_{L\times L(k-1)}, \mathbf{I}_L, \mathbf{0}_{L\times L(K-k)})$ is an $L \times LK$ matrix, and $\mathbf{U}_{\nu_k^{[\mathrm{rd}]}} \triangleq [\mathbf{0}_{LN\times L(k-1)}, j\mathbf{D}\Xi_k, \mathbf{0}_{LN\times L(K-k)}]$ and $\mathbf{U}_{\tau_k^{[\mathrm{rd}]}} \triangleq [\mathbf{0}_{LN\times L(k-1)}, \mathbf{\Lambda}_k^{[\mathrm{rd}]}(\mathbf{G}_k^{[\mathrm{rd}]})', \mathbf{0}_{LN\times L(K-k)}]$ are $LN \times LK$ matrices.

  Proof: See Appendix A.

Combining the real and imaginary parts of the channel, the new set of parameters, $\bar{\boldsymbol{\lambda}} \triangleq [\boldsymbol{\alpha}^T, (\boldsymbol{\nu}^{[\mathrm{rd}]})^T, (\boldsymbol{\tau}^{[\mathrm{rd}]})^T]^T$, is given by

$$\bar{\boldsymbol{\lambda}} = \underbrace{\begin{bmatrix} \mathbf{I}_K & j\mathbf{I}_K & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{I}_K & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{I}_K \end{bmatrix}}_{\triangleq \mathbf{M}} \boldsymbol{\lambda}. \quad (12)$$

The CRLB matrix, $\mathbf{C}$, for the complex-valued estimation vector, $\bar{\boldsymbol{\lambda}}$ is evaluated as $\mathbf{C} = \mathbf{M}\mathbf{F}^{-1}\mathbf{M}^H$, and can be written in block matrix form as

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\nu}^{[\mathrm{rd}]}} & \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\tau}^{[\mathrm{rd}]}} \\ \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} & \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}} \\ \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} & \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}} \end{bmatrix}, \quad (13)$$

where $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$ are $K \times K$ matrices and the CRLBs for the estimation of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\tau}^{[\mathrm{rd}]}$, and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, are given by the diagonal elements of $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, respectively.

Remark 1: Equation (11) shows that the FIM for the estimation of $\boldsymbol{\lambda}$ is not block diagonal. Thus, there exists coupling between the estimation errors of channel gains, MTOs, and MCFOs, i.e., the estimation performance of one parameter is affected by the presence of the remaining parameters. This shows the importance of jointly estimating channel gains, MTOs, and MCFOs in AF-DSTBC relaying cooperative systems. More importantly, this result indicates that the previously proposed methods that assume perfect timing or frequency synchronization while estimating MCFOs and MTOs in [20] and [10], respectively, cannot be applied to jointly estimate and compensate these impairments in AF-DSTBC relaying cooperative networks. This finding is also corroborated by the simulation results in Section VI-B, where it is shown that application of the algorithms in [10], [20] cannot achieve synchronization in the presence of both MTOs and MCFOs in AF-DSTBC cooperative networks.

IV. Joint Parameter Estimation at the Destination

In this section, the LS estimator for joint estimation of multiple channel gains, MTOs and MCFOs is derived. Subsequently, it is shown that using DE, the computational complexity for obtaining these impairments in DSTBC-AF relaying networks can be significantly reduced.

A. LS Estimation

Based on the signal model in (6), the LS estimates of the parameters, $\boldsymbol{\alpha}$, $\boldsymbol{\tau}^{[\mathrm{rd}]}$, and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, can be determined by minimizing the cost function, $\mathcal{J}_{\boldsymbol{\alpha},\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, according to

$$\begin{aligned} &\mathcal{J}_{\boldsymbol{\alpha},\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} \\ &= \left\| \mathbf{y}^{[\mathrm{TP}]} - \left[\Lambda_1^{[\mathrm{rd}]}\mathbf{G}_1^{[\mathrm{rd}]}, \ldots, \Lambda_K^{[\mathrm{rd}]}\mathbf{G}_K^{[\mathrm{rd}]}\right] \mathbf{A}\mathbf{t} \right\|^2 \\ &= \left\| \mathbf{y}^{[\mathrm{TP}]} - \mathbf{\Psi}\boldsymbol{\alpha} \right\|^2. \end{aligned} \quad (14)$$

Given $\boldsymbol{\tau}^{[\mathrm{rd}]}$ and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, it is straightforward to show that the LS estimate of $\boldsymbol{\alpha}$, denoted by $\hat{\boldsymbol{\alpha}}$, can be determined as

$$\hat{\boldsymbol{\alpha}} = (\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\mathbf{\Psi}^H\mathbf{y}^{[\mathrm{TP}]}. \quad (15)$$

Substituting (15) in (14), the estimates of MTOs and MCFOs, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, respectively, are obtained via

$$\begin{aligned} &\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}, \hat{\boldsymbol{\nu}}^{[\mathrm{rd}]} \\ &= \arg \min_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} \underbrace{-\left(\mathbf{y}^{[\mathrm{TP}]}\right)^H \mathbf{\Psi}\left(\mathbf{\Psi}^H\mathbf{\Psi}\right)^{-1}\mathbf{\Psi}^H\mathbf{y}^{[\mathrm{TP}]}}_{\triangleq \chi(\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]})}, \end{aligned} \quad (16)$$

where $\arg\min$ denotes the arguments, $\boldsymbol{\tau}^{[\mathrm{rd}]}$ and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, that minimize the expression $\chi(\boldsymbol{\tau}^{[\mathrm{rd}]}, \boldsymbol{\nu}^{[\mathrm{rd}]})$ and $\mathbf{y}^{[\mathrm{TP}]}$, defined in (6), is a function of timing and frequency offset estimation errors from the source-to-relay link. The channel estimates, $\hat{\boldsymbol{\alpha}}$, are obtained by substituting $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$ back into (15).

The minimization in (16) requires a multidimensional exhaustive search over the discretized set of possible timing and frequency offset values, which is inherently very computationally complex. Furthermore, to reach the CRLB (see Fig. 5 in Section VI), the exhaustive search in (16) needs to be carried out with very high resolution[2], which significantly increases the sets of possible values for both frequency and timing offsets, $\boldsymbol{\nu}^{[\mathrm{rd}]}$ and $\boldsymbol{\tau}^{[\mathrm{rd}]}$, respectively, and in turn, further increases the complexity of the proposed LS estimator.

Note that in order to reduce the complexity associated with the exhaustive search in (16), alternating projection (AP) has been used to transform the multidimensional minimization problem into a series of one-dimensional optimizations that are carried out sequentially [28]. However, in our extensive simulations we have observed that AP does not converge to the true estimates. This may be attributed to the source-to-relay timing and frequency offset estimation errors, which are not known at the destination and significantly affect the initialization of AP. Therefore, in this paper *differential evolution (DE)* is employed as a computationally efficient algorithm to carry out the minimization in (16) [23]. The initialization and parameterization of the proposed DE estimator are outlined below.

B. Differential Evolution Based Estimation

DE and genetic algorithms are considered as a subclass of *evolutionary algorithms* since they attempt to evolve the solution for a problem through recombination, mutation, and survival of the fittest. More specifically, DE is an optimization algorithm, where a number of parameter vectors are generated and updated at each iteration in order to reach the solution [23]. Let d, for $d = 1, \ldots, D_I$, denote the number of iterations[3]. The detailed steps of the proposed DE algorithm for carrying out the minimization in (16) are as follows:

Step 1) *Initialization*: A *population* comprising of D_P *target vectors* is generated. The first iteration's ℓth target vector, $\boldsymbol{\theta}_{\ell,1} \triangleq [\tau_{1,\ell,1}^{[\mathrm{rd}]}, \ldots, \tau_{K,\ell,1}^{[\mathrm{rd}]}, \nu_{1,\ell,1}^{[\mathrm{rd}]}, \ldots, \nu_{K,\ell,1}^{[\mathrm{rd}]}]^T$, for $\ell = 1, \ldots, D_P$, is generated by selecting $\nu_{k,\ell,1}^{[\mathrm{rd}]}$ and $\tau_{k,\ell,1}^{[\mathrm{rd}]}$, $\forall\, k, \ell$, uniformly and randomly from the

[2]Step sizes of 10^{-2} and 10^{-4} for MTOs and MCFOs, respectively.

[3]Also called "*generation*" in DE terminology.

set of possible values for frequency and timing offsets, respectively, i.e., $\nu_{k,\ell,1}^{[\mathrm{rd}]}, \tau_{k,\ell,1}^{[\mathrm{rd}]} \sim \mathcal{U}(-0.5, 0.5)$, $\forall\, k, \ell$. Subsequently, at each iteration, the D_P target vectors in the population are updated according to Steps 2–4 below.

Step 2) *Mutation*: At the dth iteration for each individual target vector, $\boldsymbol{\theta}_{\ell,d}$, three distinct and different target vectors denoted by $\boldsymbol{\theta}_{r1,d}$, $\boldsymbol{\theta}_{r2,d}$, and $\boldsymbol{\theta}_{r3,d}$, are randomly selected from the population. Subsequently, a *mutant* vector corresponding to the ℓth target vector and dth iteration, $\boldsymbol{\varpi}_{\ell,d} = [\varpi_{1,\ell,d}, \ldots, \varpi_{2K,\ell,d}]^T$, is created via

$$\boldsymbol{\varpi}_{\ell,d} = \boldsymbol{\theta}_{r1,d} + (\boldsymbol{\theta}_{r2,d} - \boldsymbol{\theta}_{r3,d}) \cdot D_F, \tag{17}$$

where D_F is a real positive scaling factor that controls the rate at which the population evolves.

Step 3) *Crossover*: In this step, the DE employs a uniform crossover between each target vector, $\boldsymbol{\theta}_{\ell,d}$, and the mutant vector, $\boldsymbol{\varpi}_{\ell,d}$, to create a *trial* vector, $\boldsymbol{\vartheta}_{\ell,d} \triangleq [\vartheta_{1,\ell,d}, \ldots, \vartheta_{2K,\ell,d}]^T$, such that

$$\vartheta_{q,\ell,d} = \begin{cases} \varpi_{q,\ell,d}, & \iota \le D_R \\ \theta_{q,\ell,d}, & \text{otherwise} \end{cases}, \quad q = 1, \ldots, 2K \tag{18}$$

where $\iota \sim \mathcal{U}(0,1)$ and the crossover probability, $D_R \in [0,1]$ controls the fraction of parameter values that are copied from the mutant vector, $\boldsymbol{\varpi}_{\ell,d}$, to the trial vector, $\boldsymbol{\vartheta}_{\ell,d}$, $\forall\, \ell$.

Step 4) *Selection*: If compared to the ℓth target vector, $\boldsymbol{\theta}_{\ell,d}$, the ℓth trial vector, $\boldsymbol{\vartheta}_{\ell,d}$, results in a smaller objective function, χ, in (16), it replaces the target vector in the next iteration, i.e.,

$$\boldsymbol{\theta}_{\ell,d+1} = \begin{cases} \boldsymbol{\vartheta}_{\ell,d}, & \chi(\boldsymbol{\vartheta}_{\ell,d}) \le \chi(\boldsymbol{\theta}_{\ell,d}) \\ \boldsymbol{\theta}_{\ell,d}, & \text{otherwise.} \end{cases} \tag{19}$$

Step 5) *Stopping Criteria*: There are various stopping criteria for the DE. However, our extensive simulations show that to reach the CRLB and obtain accurate estimates for different channel realizations, the algorithm can be terminated if the target vector that results in the lowest objective function, χ, i.e., the estimates of timing and frequency offsets, remains unchanged for a predefined number, ς, of consecutive iterations.

By executing Steps 1–5 above, the estimates of timing and frequency offsets, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, respectively, can be obtained. Substituting these estimates in (15) also generates the desired channel estimates. The performance and convergence of the proposed DE algorithm to the true estimates is highly dependent on the values of the DE parameters, e.g., the population size D_P, the scaling factor D_F, the crossover probability D_R, and the stopping criteria ς. As a matter of fact, each unique optimization problem requires the DE parameters to be initialized differently [23], where finding the appropriate values for these parameters can be a difficult and a non-trivial task [29]. However, by applying the results in Section III and by following the general guidelines in [23], [29], we are able to determine an appropriate parameterization for the DE algorithm.

Note that D_P is typically selected to be 10 times the dimension of the target vector. Since, we are solving a $2K$-dimensional minimization problem in (16), we set $D_P = 20K$. D_F is generally selected between 0.5 and 1.0 and extensive simulations in Section VI-A show that $D_F = 0.85$ results in the fastest convergence while maintaining estimation performance close to that of the CRLB. Moreover, from the FIM in Section III, we know that the estimation of parameters of vector, $\boldsymbol{\lambda}$, is mutually coupled. As a result, D_R is set to $D_R = 0.9$ [23]. Finally, in order to reach the CRLB, the stopping criteria is set to $\varsigma = 80$.

In DE terminology, the above procedure can be classified as "DE-rand", since to obtain the mutant vector, $\boldsymbol{\varpi}_{\ell,d}$, the target vector, $\boldsymbol{\theta}_{r1,d}$ in (17), is randomly selected from the population. In another variant of the DE, denoted by "DE-best", at each iteration, the target vector that minimizes the objective function, χ, is selected as $\boldsymbol{\theta}_{r1,d}$ in (17). However, in our problem formulation, "DE-best" does not always converge to the true values of timing and frequency offsets.[4] Similar findings are also observed in [23], where it is shown that "DE-rand" is more effective at reaching the global solution than "DE-best" for different optimization problems ([23], p. 154).

Remark 2: Even though the general conditions for global convergence of evolutionary algorithms are established in [30], it cannot be analytically shown that DE meets these conditions [31] and converges to the global solution of (16). Moreover, in [30], no specific algorithm that meets these conditions is proposed. Although in [31], a variation of DE is proposed, it is indicated in [49] that this new approach also does not guarantee the global convergence of the DE algorithm. Nevertheless, by appropriately selecting the DE parameters, in our extensive simulations, we have observed the proposed DE estimator to always converge to the true values of timing and frequency offsets as described in Sections IV-C and VI-A.

In order to compare the performance of the proposed estimators against the CRLB, it needs to be shown that they are unbiased estimators. Moreover, as shown in Section V, the statistical properties of estimates can be used to design the destination receiver such that it is more robust to the MTO, MCFO, and multiple channel estimation errors. Even though it is difficult to analytically determine the statistics of the estimates obtained using the proposed LS and DE estimators, we numerically found through simulations that the proposed estimators are unbiased and their second order statistics can be accurately approximated by the CRLB derived in Section III, i.e., expected value of channel estimation errors, $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$, timing offset estimation errors, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} = \boldsymbol{\tau}^{[\mathrm{rd}]} - \hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$, and frequency offset estimation errors, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} = \boldsymbol{\nu}^{[\mathrm{rd}]} - \hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, is $\mathbf{0}_{K\times 1}$ and they can be modeled as a multi-variate Gaussian distribution, i.e., $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}})$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}})$, and $\boldsymbol{\delta}_{\boldsymbol{\alpha}} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}})$, where $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$, $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, and $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$ are given in (13).

C. Complexity Analysis

In this subsection, the computational complexities of the LS and the DE based estimation techniques are presented. Computational complexity is defined as the number of additions plus multiplications [6]. In addition, comparing two values or generating a random number is considered as a single arithmetic operation.

[4]This is observed through extensive simulations, which are not included here due to space limitation.

The computational complexity of the LS algorithm, denoted by C_{LS} is calculated as given by (20) at the bottom of the page, where κ_τ and κ_ν denote the step sizes or resolutions for timing and frequency offsets, respectively, in the exhaustive search in (16) and $C_\alpha = 2K^3 + (4LN-1)K^2 + (LN-1)K$ is the complexity for obtaining $\hat{\boldsymbol{\alpha}}$ via (15).

The computational complexity of the proposed DE algorithm, denoted by C_{DE} is calculated as

$$C_{\mathrm{DE}} = D_I \left\{ \underbrace{6D_P K}_{(17)} + \underbrace{4D_P K}_{(18)} + \underbrace{D_P(C_\chi+1)}_{(19)} + \underbrace{D_P(C_\chi+1)}_{\text{stopping criteria}} \right\} + C_\alpha, \quad (21)$$

where D_I is the total number of iterations required for the DE to converge and $C_\chi = 2K^3 + (4LN-1)K^2 + (2(LN)^2 - LN)K + (LN)^2 + LN - 1$ is the computational complexity to evaluate the objective function χ in (16). Note that extensive simulations demonstrate that the proposed DE algorithm on average converges to the true estimates after approximately $D_I = 220$ and $D_I = 530$ iterations for $K = 2$ and $K = 4$, respectively, for a wide range of SNRs.

Remark 3: By evaluating C_{LS} and C_{DE} in (20) and (21), it is observed that impairment estimation using the DE is 5.6×10^7 and 1.2×10^{19} times more computationally efficient compared to the proposed LS algorithm, for $K = 2$ and $K = 4$ relays, respectively. Note that this comparison is carried out by evaluating the LS estimates using exhaustive search in (16) and by setting $L = 80$, $\kappa_\tau = 10^{-2}$, $\kappa_\nu = 10^{-4}$, $D_P = 20K$, $D_R = 0.9$, $D_F = 0.85$, and $D_I = 220$ and 530 for $K = 2$ and $K = 4$, respectively. These values are selected to ensure that both estimators reach the CRLB over a wide range of SNR values. This comparison serves to illustrate the significantly lower complexity of the proposed DE based estimator.

V. Compensation Algorithms at the Destination

The received signal at the destination, $\mathbf{y}^{[\mathrm{DTP}]}$ in (7), is the superposition of the relays' transmitted signals that are attenuated differently, are no longer aligned with each other in time, and are experiencing phase rotations at different rates due to different channels, MTOs, and MCFOs, respectively. In this section, we derive the MMSE and the low complexity MMSE (L-MMSE) receivers to compensate the effect of these impairments.

A. MMSE Compensation Design

The goal of the compensation algorithm is to design an $L \times LN$ matrix, $\mathbf{Q}$, as shown in Fig. 3, which is applied to the received signal, $\mathbf{y}^{[\mathrm{DTP}]}$, to remove the effect of MTOs, MCFOs, and multiple channels. The MMSE compensation matrix, $\mathbf{Q}$, is given by minimizing the cost function $\mathcal{J}_{\mathbf{s}}$ below

$$\mathcal{J}_{\mathbf{s}} = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u},\mathbf{w},\mathbf{s}} \left\{ \left\| \mathbf{Q}\mathbf{y}^{[\mathrm{DTP}]} - \mathbf{s} \right\|^2 \right\}, \quad (22)$$

where $\mathbf{s}$ is defined below (2), and the expectation in (22) is taken with respect to the statistics of the estimation errors $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} = \boldsymbol{\tau}^{[\mathrm{sr}]} - \hat{\boldsymbol{\tau}}^{[\mathrm{sr}]}$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} = \boldsymbol{\nu}^{[\mathrm{sr}]} - \hat{\boldsymbol{\nu}}^{[\mathrm{sr}]}$, $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$ and $\mathbf{s}$, $\mathbf{u}$, and $\mathbf{w}$. Furthermore, to meet the power constraints at the destination, it is assumed that $\mathrm{Tr}\{\mathbb{E}_{\mathbf{s}}\{\mathbf{s}\mathbf{s}^H\}\} \leq L$.

Assuming that the estimators used to obtain these synchronization and channel impairments are unbiased and asymptotically optimal, i.e., reach the CRLB for high SNR or very large number of observation symbols, the estimation errors, $\boldsymbol{\delta}_{\boldsymbol{\alpha}}$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$, can be modeled by $\boldsymbol{\delta}_{\boldsymbol{\alpha}} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}})$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{sr}]},\boldsymbol{\tau}^{[\mathrm{sr}]}})$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{sr}]},\boldsymbol{\nu}^{[\mathrm{sr}]}})$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}})$, and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}})$, respectively, [32]. Note that $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, are given in (13), and $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{sr}]},\boldsymbol{\tau}^{[\mathrm{sr}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{sr}]},\boldsymbol{\nu}^{[\mathrm{sr}]}}$ can be determined from the results in [33]. In the following, knowledge of the statistics of the estimation errors for channels and synchronization parameters are applied to design a compensation algorithm at the destination receiver that is robust to imperfect channel and synchronization parameter estimation.

Theorem 2: The average MSE of the recovered data with respect to the source signal, $\mathcal{J}_{\mathbf{s}}$, is given by

$$\begin{aligned}\mathcal{J}_{\mathbf{s}} = \mathrm{Tr}\Big\{ & \mathbf{Q}\hat{\boldsymbol{\Xi}}\mathbf{R}\hat{\boldsymbol{\Xi}}^H\mathbf{Q}^H + \mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{Q}^H \\ & + \boldsymbol{\Phi}_{\mathbf{s}} - \mathbf{Q}\hat{\boldsymbol{\Xi}}\hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}} \\ & - \boldsymbol{\Phi}^H_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H\hat{\boldsymbol{\Xi}}^H\mathbf{Q}^H \Big\} \\ & + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{sr}]}}\right) + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{rd}]}}\right), \end{aligned} \quad (23)$$

where

- $\hat{\boldsymbol{\Xi}} \triangleq \boldsymbol{\Xi}|_{\boldsymbol{\tau}^{[\mathrm{rd}]}=\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}=\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}}$ is an $LN \times LK$ matrix, $\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}} = [\underbrace{\boldsymbol{\Phi}_{\mathbf{s}},\ldots,\boldsymbol{\Phi}_{\mathbf{s}}}_{K}]^T$ is an $LK \times L$ matrix, $\boldsymbol{\Phi}_{\mathbf{s}} = \mathbf{I}_L$,
- $\mathbf{R} = \hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H + (\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \otimes \mathbf{I}_L) \odot (\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H) + \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H$ is an $LK \times LK$ matrix,
- $\hat{\mathbf{A}} \triangleq \mathrm{diag}(\hat{\boldsymbol{\alpha}}) \otimes \mathbf{I}_L$ is an $LK \times LK$ matrix, $\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \triangleq \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}}$ is a $K \times K$ matrix,
- $\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}} = \mathbf{1}_{K\times K} \otimes \boldsymbol{\Phi}_{\mathbf{s}}$ is an $LK \times LK$ matrix,
- $\boldsymbol{\Sigma}_{\mathbf{n}}$ is the $LK \times LK$ covariance matrix of the matched filtered noise at the relays, $\mathbf{n}$, with its kth $L \times L$ block diagonal matrix, $[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L:kL,(k-1)L:kL}$, given by $[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L:kL,(k-1)L:kL} \triangleq \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\boldsymbol{\Sigma}_{\mathbf{u}_k}(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]})^H$,

$$C_{\mathrm{LS}} = \underbrace{\left(\frac{1}{\kappa_\tau\kappa_\nu}\right)^K \left[2K^3 + (4LN-1)K^2 + (2(LN)^2 - LN)K + (LN)^2 + LN - 1\right]}_{\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]},\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]} \text{ in (16)}} + C_\alpha, \quad (20)$$

$\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \triangleq \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}|_{\hat{\tau}_k^{[\mathrm{sr}]}=\tau_k^{[\mathrm{sr}]}}$ is an $L \times LN$ matrix, and $\boldsymbol{\Sigma}_{\mathbf{u}_k} \triangleq \sigma_{u_k}^2 \mathbf{I}_{LN}$, and

- $\mathcal{O}_p(\cdot)$ denotes the big omicron function for stochastic parameters [34].

Proof: See Appendix B.

Proposition 1: Using (23), the MMSE compensation matrix, $\mathbf{Q}^{[\mathrm{MMSE}]}$, is given by

$$\mathbf{Q}^{[\mathrm{MMSE}]} = \boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}^H \boldsymbol{\Omega}^H \hat{\mathbf{A}}^H \hat{\boldsymbol{\Xi}}^H \left(\hat{\boldsymbol{\Xi}}\mathbf{R}\hat{\boldsymbol{\Xi}}^H + \boldsymbol{\Sigma}_{\mathbf{w}}\right)^{-1}, \quad (24)$$

where $\mathbf{R}$, $\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}$, $\hat{\mathbf{A}}$, and $\hat{\boldsymbol{\Xi}}$ are defined below (23), and $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}_{\mathbf{w}}$ are defined below (8) and (9), respectively.

Proof: Equation (24) follows from Theorem 1 by taking the derivative of $\mathcal{J}_{\mathbf{s}}$ in (23) with respect to $\mathbf{Q}$, setting the results to zero, and carrying out straightforward algebraic manipulations.

Remark 4: Evaluation of the MMSE compensation matrix, $\mathbf{Q}^{[\mathrm{MMSE}]}$, requires knowledge of the overall channel gains, $\hat{\boldsymbol{\alpha}}$, relay to destination timing and frequency offsets, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, respectively, relay to destination channel gains $\boldsymbol{\beta}$, and source to relay timing offsets, $\boldsymbol{\tau}^{[\mathrm{sr}]} \triangleq [\hat{\tau}_1^{[\mathrm{sr}]}, \ldots, \tau_K^{[\mathrm{sr}]}]^T$, at the destination. Even though, $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$, and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$ can be jointly estimated via the estimators proposed in Section IV and are known at the destination, $\boldsymbol{\beta}$ and $\boldsymbol{\tau}^{[\mathrm{sr}]}$ are not known at the destination. Thus, the proposed compensation algorithm requires the relays to feed forward the estimates of source to relay timing offsets and channel gains, $\hat{\tau}_k^{[\mathrm{sr}]}$ and $\hat{h}_k$, respectively, to the destination. Note that using $\hat{h}_k$, an estimate of relay to destination channels, $\hat{\beta}_k$, can be obtained according to $\hat{\beta}_k = \frac{\hat{\alpha}_k}{\hat{h}_k}$.

Remark 5: According to the results in Appendix B, it can be concluded that due to the presence of channel, timing, and frequency offset estimation errors, $\mathbf{Q}^{[\mathrm{MMSE}]}$ is not the optimal MMSE compensation matrix at the destination. However, one can derive this optimal MMSE matrix by assuming perfect knowledge of channels, timing, and frequency offsets from source to relays and relays to destination, i.e., $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} = \mathbf{0}_{K\times 1}$.

Proposition 2: Based on the assumption of perfect synchronization and channel estimation, the optimal MMSE compensation matrix, $\mathbf{Q}^{[\mathrm{OPT}]}$, can be determined as

$$\mathbf{Q}^{[\mathrm{OPT}]} = \boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}^H \boldsymbol{\Omega}^H \mathbf{A}^H \boldsymbol{\Xi}^H \left(\boldsymbol{\Xi}\mathbf{R}^{[\mathrm{OPT}]}\boldsymbol{\Xi}^H + \boldsymbol{\Sigma}_{\mathbf{w}}\right)^{-1}, \quad (25)$$

where $\mathbf{R}^{[\mathrm{OPT}]} = \mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H\mathbf{A}^H + \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H$.

Proof: By assuming $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} = \mathbf{0}_{K\times 1}$ in (22) and carrying out the steps in Appendix B, the result in (25) follows.

Even though the proposed optimal compensation approach has limited practical applications, it can be used to benchmark the performance of DSTBC-AF cooperative networks. As a matter of fact, $\mathbf{Q}^{[\mathrm{OPT}]}$ is applied in Section VI-B to determine a lower bound on the BER performance of DSTBC-AF cooperative networks in the presence of channel and synchronization impairments.

B. L-MMSE Compensation Design

In order to reduce synchronization overhead, in this subsection, a low complexity approach that does not require relays to feed forward synchronization and channel estimates to the destination is proposed.

The compensation matrix $\mathbf{Q}^{[\mathrm{MMSE}]}$ in (24) is a function of the source to relay channels and timing offsets since the matrix $\mathbf{R}$ defined below (23) is a function of $\mathbf{B}$ and $\boldsymbol{\Sigma}_{\mathbf{n}}$, respectively. However, over a large number of channel realizations, the second order channel statistics can be used to approximate source to relay channel gains on average. As a result, $\mathbf{B}$ in (24) can be replaced by $\mathbf{B}^{[\mathrm{L-MMSE}]} \approx \mathrm{diag}(\sigma_f\zeta_1, \ldots, \sigma_f\zeta_K) \otimes \mathbf{I}_L$. Moreover, the dependence on timing offset estimates from source to relay can be eliminated by approximating the block diagonal matrix $\boldsymbol{\Sigma}_{\mathbf{n}}$, defined below (23), with the diagonal matrix $\boldsymbol{\Sigma}_{\mathbf{n}}^{[\mathrm{L-MMSE}]}$ with its kth diagonal block given by $[\boldsymbol{\Sigma}_{\mathbf{n}}^{[\mathrm{L-MMSE}]}]_{(k-1)L:kL,(k-1)L:kL} = \sigma_{u_k}^2 \mathbf{I}_L$ at high SNR. The latter follows from the definition of $\boldsymbol{\Sigma}_{\mathbf{n}}$ and extensive simulations that show that the off-diagonal elements of $\boldsymbol{\Sigma}_{\mathbf{n}}$ that are dependent on source to relay timing offsets vanish much more quickly compared to its diagonal elements as SNR increases. Using the above approximations and (24), a low complexity compensation matrix, $\mathbf{Q}^{[\mathrm{L-MMSE}]}$, for mitigating the effect of impairments can be found as

$$\mathbf{Q}^{[\mathrm{L-MMSE}]} = \boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}^H \boldsymbol{\Omega}^H \hat{\mathbf{A}}^H \hat{\boldsymbol{\Xi}}^H \times \left(\hat{\boldsymbol{\Xi}}\mathbf{R}^{[\mathrm{L-MMSE}]}\hat{\boldsymbol{\Xi}}^H + \boldsymbol{\Sigma}_{\mathbf{w}}\right)^{-1}, \quad (26)$$

where

$$\mathbf{R}^{[\mathrm{L-MMSE}]} = \hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H + (\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \otimes \mathbf{I}_L) \odot (\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H) + \mathbf{B}^{[\mathrm{L-MMSE}]}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}^{[\mathrm{L-MMSE}]}\boldsymbol{\Omega}^H(\mathbf{B}^{[\mathrm{L-MMSE}]})^H$$

.

Simulation results in Section VI-B show that the BER performance of a multi-relay DSTBC-AF cooperative network using the proposed L-MMSE approach in (26) is very close to that of the MMSE receiver in (24) over a wide range of SNRs. This result indicates that the approximations used to arrive at (26) hold for practical scenarios of interest. Thus, the proposed L-MMSE approach can be effectively applied to compensate the effect of MTOs, MCFOs, and channels in DSTBC-AF cooperative networks, while significantly reducing synchronization overhead.

VI. Simulation Results

In this section, we investigate the receiver performance at the destination, where multiple channel gains, MTOs, and MCFOs are jointly estimated and compensated in order to decode the received signal. In our simulation setup, we consider $K = 2$ and 4 relays in DSTBC-AF cooperative systems. Quadrature phase-shift keying modulation is employed for data transmission. Length of the training signals, $\mathbf{t}^{[\mathrm{sr}]}$ and $\mathbf{t}_k^{[\mathrm{rd}]}$, $\forall k$, are set to $L = 80$ symbols during the TP and length of the source data vector $\mathbf{s}$, is set to $L = 400$ symbols during the DTP, resulting in a synchronization overhead of 16%. Oversampling factor is set to $N = 2$ and a root-raised cosine filter with a roll-off factor of 0.22 is employed. At each relay, the DSTBC is generated randomly based on an isotropic distribution on the space of $L \times L$ unitary matrices, which is a benchmark method for

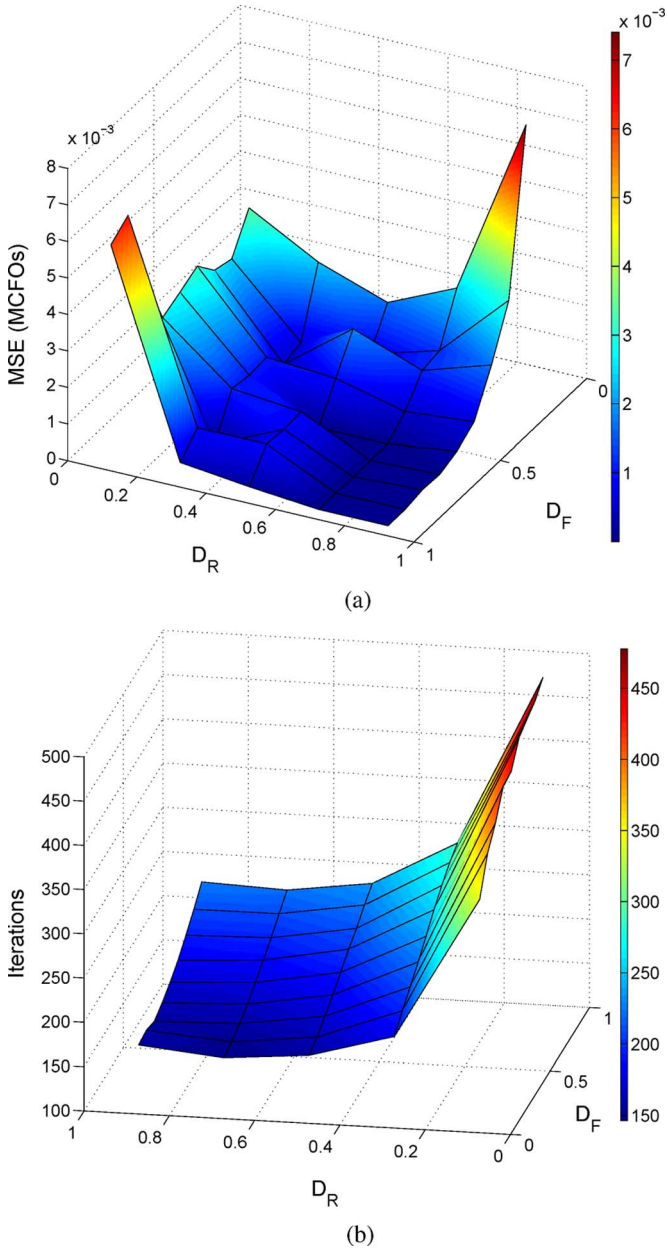


Fig. 4. (a) CRLB and MSE performance of MCFOs' estimation for DE based algorithm with different parameter settings, as a function of SNR (dB). (b) Number of generations (iterations) required for the convergence of DE algorithm with different parameter settings.

generating DSTBC in AF cooperative networks [7]. The propagation loss is modeled as $\eta = (\varepsilon/\varepsilon_0)^{-\rho}$, where ε is the distance between transmitter and receiver, ε_0 is the reference distance, and ρ is the path loss exponent [26]. We set $\varepsilon_0 = 1$ km, and $\rho = 2.7$, which corresponds to an urban area cellular network [26]. The timing and frequency offsets at relays and destination, $\boldsymbol{\tau}^{[\mathrm{sr}]}$, $\boldsymbol{\tau}^{[\mathrm{rd}]}$, $\boldsymbol{\nu}^{[\mathrm{sr}]}$, and $\boldsymbol{\nu}^{[\mathrm{rd}]}$ are uniformly drawn from the full acquisition range, $(-0.5, 0.5)$. Without loss of generality, it is assumed that the AWGN at all relays has the same variance, i.e., $\sigma_u^2 = \sigma_{u_1}^2 = \cdots = \sigma_{u_K}^2$ and $\sigma_u^2 = \sigma_w^2 = 1/\mathrm{SNR}$. The MSE performance of the proposed estimators and the end-to-end BER of DSTBC-AF cooperative networks using the proposed compensation algorithms are detailed below.

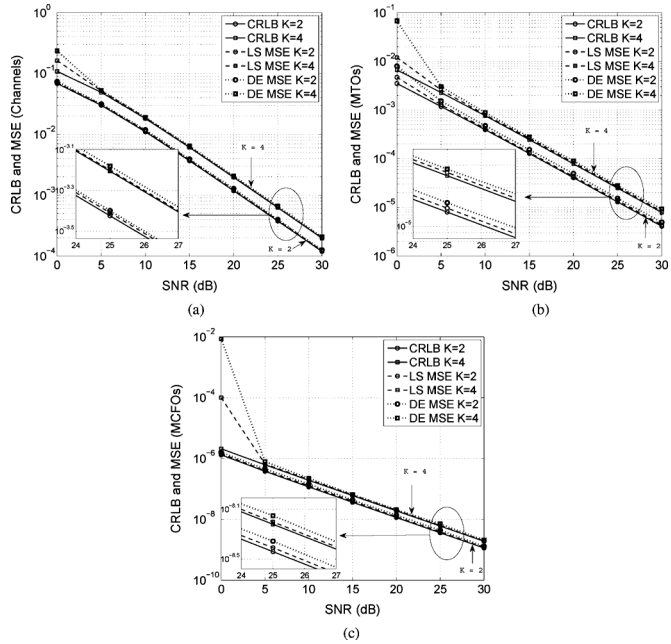


Fig. 5. CRLBs in (13) and MSE for the estimation of (a) channel gains, (b) MTOs, and (c) MCFOs as a function of SNR (dB).

A. MSE Performance

Following the approach in [6], [20], [24], [25] and for ease of reproducing the CRLBs and MSE curves in Figs. 4 and 5, specific channels are used. For $K = 2$, $\mathbf{h} = [.279 - .9603j, .8837 + .4681j]^T$ and $\mathbf{f} = [.7820 + .6233j, .9474 - .3203j]^T$ and for $K = 4$, $\mathbf{h} = [.279 - .9603j, .8837 + .4681j, -.343 + .732i, -.734 - .451i]^T$ and $\mathbf{f} = [.7820 + .6233j, .9474 - .3203j, -.2413 + .724i, .5141 - .893i]^T$. Note that in the next subsection, we employ random Rayleigh fading channels for analyzing the end-to-end BER performance of the overall system. As stated in Section IV-C and to reach the CRLB, the step sizes, κ_τ and κ_ν, for the exhaustive search in (16) for the LS estimator are set to 10^{-2} and 10^{-4}, respectively. Without loss of generality, the CRLBs and the MSE estimation performance for the first relay node is presented only, where similar results are observed for the other relays.

Fig. 4(a) shows the MSE of frequency offset estimation[5] as a function of the DE parameters: the crossover probability, D_R, and the scaling factor, D_F. It is shown that the lowest MSE is achieved for $D_F = 0.85$ and $D_R = 0.9$. Fig. 4(b) shows the number of iterations required for the convergence of the DE algorithm for different D_R and D_F values. It can be seen that the DE algorithm converges more quickly as D_R increases from 0.1 to 0.9. As a matter of fact, Fig. 4(b) demonstrates that the proposed DE algorithm converges with the fewest number of iterations when $D_F = 0.25$ and $D_R = 0.9$. However, as shown in Fig. 4(a), this fast convergence comes at the cost of MSE performance. The results in Figs. 4(a) and (b) demonstrate that the proposed DE algorithm provides an effective trade-off between synchronization performance and complexity that can be

[5]Similar results can be obtained for timing offset estimation, however, are not included here due to lack of space.

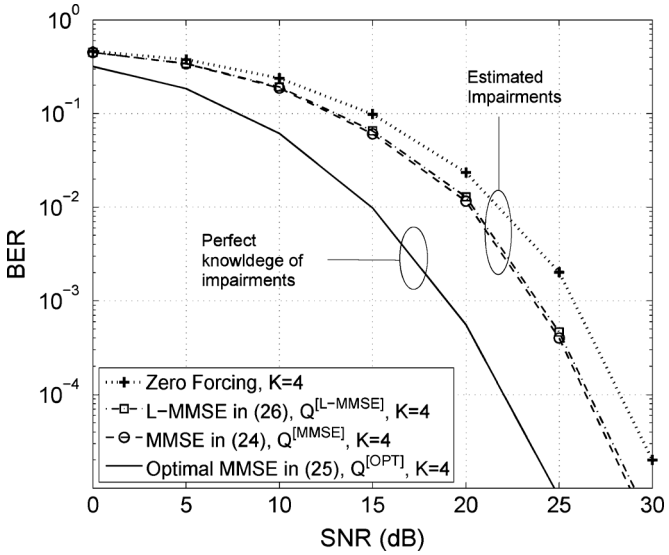


Fig. 6. BER of ZF, L-MMSE, MMSE, and optimal receivers for 4-relay AF-DSTBC cooperative systems.

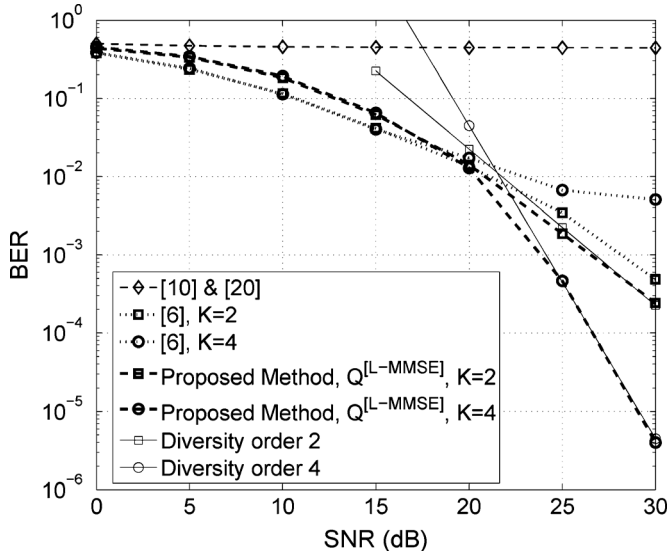


Fig. 7. BER of the proposed L-MMSE receiver while using the DE-based estimator against the approaches in [10], [20], and [6].

used to meet the performance requirements of cooperative communication systems. In this scenario, faster convergence can be achieved by selecting a smaller value for D_F, e.g., $D_F = 0.25$, while compromising estimation accuracy. In order to achieve the best estimation accuracy, in the remainder of this section, we set $D_F = 0.85$ and $D_R = 0.9$.

Figs. 5(a), (b), and (c) show the CRLBs and the MSEs for the estimation of channel gains, MTOs, and MCFOs for different SNRs and $K = 2$ and 4 relays. Fig. 5 shows that the CRLBs and the MSEs for a 2-relay DSTBC-AF cooperative network are lower compared to that of a network equipped with 4 relays. This is because as the number of relays increases, more parameters need to be jointly estimated at the destination and the effect of amplified and forwarded AWGN from the relays is also more prominent at the destination. Moreover, it is demonstrated that at low SNR ($\mathrm{SNR} < 5$ dB), the proposed estimators exhibit poor performance due to the noise at the relays, which is amplified and forwarded to the destination. However at moderate-to-high SNRs ($\mathrm{SNR} \geq 5$ dB), Fig. 5 shows that the MSEs of the proposed LS and DE estimators are close to the CRLB. Note while showing comparable estimation accuracy, the proposed DE algorithm is significantly less complex than the LS estimator as shown in *Remark 3*.

B. BER Performance

In this subsection the BER performance of a DSTBC-AF cooperative system that employs the proposed DE estimator and compensation algorithms is investigated. The channels from source to relays and from relays to destination are modeled as independent and identically distributed complex Gaussian random variables with $\mathcal{CN}(0, 1)$.

Fig. 6 shows the end-to-end BER performance of a 4-relay DSTBC-AF cooperative system, using the proposed MMSE, L-MMSE, and optimal compensation algorithms in (24), (26), and (25), respectively. The performances of the proposed compensation algorithms are also compared to that of the zero-forcing (ZF) receiver, which obtains an estimate of the transmitted signal via $\hat{\mathbf{s}} = (\boldsymbol{\Upsilon}^H \boldsymbol{\Upsilon})^{-1} \boldsymbol{\Upsilon}^H \mathbf{y}^{[\mathrm{DTP}]}$, where $\boldsymbol{\Upsilon} \triangleq \sum_{k=1}^{K} \hat{\alpha}_k \hat{\boldsymbol{\Xi}}_k \boldsymbol{\Omega}_k$. For the MMSE, L-MMSE, and ZF methods, the compensation matrices, $\mathbf{Q}^{[\mathrm{MMSE}]}$ in (24), $\mathbf{Q}^{[\mathrm{L-MMSE}]}$ in (26), and $\boldsymbol{\Upsilon}$ defined above are evaluated using the estimated values, $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$, and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, obtained in the training period using the DE algorithm[6] while the optimal compensation matrix $\mathbf{Q}^{[\mathrm{OPT}]}$ in (25) is evaluated based on perfect knowledge of channels and synchronization parameters. Fig. 6 shows that both the proposed MMSE and L-MMSE receivers outperform the ZF receiver with a performance gain of 2 dB or more at moderate-to-high SNRs. Furthermore, Fig. 6 shows that the BER of a DSTBC-AF cooperative system based on estimated impairments is close that of the ideal scenario with perfect knowledge of impairments and optimal compensation (a gap of 4–5 dBs). This result demonstrates the robustness of the proposed transceiver designs to estimation errors. Finally, Fig. 6 demonstrates that even though the proposed L-MMSE receiver does not require relays to feed forward source to relay parameter estimates to the destination, its BER is very close to that of the MMSE receiver with a performance gap of only 0.1–0.3 dB.

Fig. 7 depicts the BER results for a cooperative system that first employs the re-synchronization filter in [10] to compensate MTOs and then attempts to remove MCFOs by employing the algorithm in [20]. This plot, which is labeled as " [10] & [20]", shows that such an approach fails to equalize the effect of impairments at the destination since the re-synchronization filter in [10] fails to compensate MTOs in the presence of MCFOs. Subsequently, the algorithm in [20] fails to nullify MCFOs, since the input signal is corrupted by MTOs. In addition, even though Fig. 7 illustrates the BER performance of a cooperative system applying the transceiver design in [6] for 2 and 4 relays, it is important to consider that the proposed DSTBC-AF transceiver design in this paper and the scheme in [6] vastly differ. In [6], a less complex AF

[6]Estimation using LS estimator results in similar BER performance as obtained using the DE algorithm.

transceiver design is used that does not allow for the application of DSTBCs, and as a result, sacrifices spatial diversity to achieve lower complexity. On the other hand, as shown in Fig. 7, the DSTBC-AF transceiver design proposed in this paper achieves full diversity gain at high SNR, while requiring additional synchronization overhead at the relays to allow for the application of DSTBCs.

Fig. 7 shows that at low SNR, the proposed algorithm exhibits a higher BER performance compared to that of [6]. This follows from the inherent relay transceiver structure for the DSTBC-AF scheme proposed here, where the relays need to estimate and compensate the effect of timing and frequency offsets before applying DSTBCs and forwarding their signals to the destination. As a result, the forwarded signals from the relays are affected by timing and frequency offset estimation errors that negatively affect cooperative performance at the destination node. This effect is particularly more prominent at low SNR, where the timing and frequency offset estimation errors at the relays have larger variances. On the other hand, the algorithm in [6] does not apply a matched filter at each relay, cannot apply DSTBCs, and does not require the relays to achieve carrier synchronization before forwarding their signal to the destination. Hence, the approach in [6] is not significantly influenced by the timing and frequency offset estimation errors at the relays at low SNR. At high SNR, the proposed algorithm exhibits large performance gain compared to the results in [6], e.g., at $\mathrm{SNR} = 30$ dB, the BER of the proposed algorithm is 4×10^{-6} for $K = 4$ relays while the BER of the algorithm in [6] is 5×10^{-2}. Moreover, due to the application of DSTBCs in the proposed cooperative system, the BER of a 4-relay system achieves a large diversity gain over that of a 2-relay network while the algorithm in [6] provides little to no spatial diversity gain as the number of relays increases.

VII. Conclusion

This article introduces a novel DSTBC based AF relaying cooperative network in the presence of MTOs, MCFOs, and unknown channels. New CRLBs for the multi-parameter estimation problem are derived and LS and DE based estimators for joint estimation of multiple channel gains, MTOs, and MCFOs are proposed. The convergence, parametrization, and complexity of the proposed DE algorithm are analyzed and it is empirically shown that the proposed DE estimator is 5.6×10^7 or 1.2×10^{19} times more computationally efficient than the proposed LS estimator when considering 2 or 4 relay AF-DSTBC networks, respectively. Simulation results show that the MSEs of the proposed estimators are close to the CRLB at moderate-to-high SNRs ($\mathrm{SNR} \geq 5$ dB). Finally, a novel MMSE receiver is proposed to decode the received signal at the destination by compensating for the effect of multiple channel gains, MTOs and MCFOs. To reduce the overhead associated with synchronization in AF-DSTBC networks, an L-MMSE receiver is also proposed that does not require the relays to feed forward any estimates to the destination. Simulations show that the BER plots of the proposed MMSE and L-MMSE receivers are only 0.1–0.3 dB apart while showing a performance gap of 4–5 dB from the derived optimal MMSE receiver that has access to perfect knowledge of impairments. More importantly, these results demonstrate that the combination of the proposed estimators and compensation algorithms ensures that a multi-relay AF-DSTBC cooperative network may achieve full spatial diversity in the presence of MTOs, MCFOs, and multiple unknown channels.

Note that the results in this paper can be extended to the case of frequency selective channels by considering OFDM transmission. Similar to single carrier systems, MTOs and MCFOs result in inter-symbol, inter-block, and inter-carrier interferences in OFDM systems. As a matter of fact, joint compensation of MCFOs and MTOs in DSTBC-AF cooperative OFDM systems is still an open research area and is the subject of future research.

Appendix A
Derivation of FIM in (11)

The elements of the $4K \times 4K$ FIM are given by [32]

$$[\mathbf{F}]_{m,\bar{m}} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}^H}{\partial \lambda_m}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \lambda_{\bar{m}}}\right\} + \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \lambda_m}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \lambda_{\bar{m}}}\right\}, \quad m, \bar{m} = 1, \ldots, 4K. \tag{A.1}$$

The mean of the received signal in (6) during TP, $\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}$, is calculated as

$$\begin{aligned}\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}} &= \mathbb{E}\left\{\boldsymbol{\Xi}\mathbf{A}\mathbf{p} + \boldsymbol{\Xi}\mathbf{B}\mathbf{v} + \mathbf{w}\right\} \\ &= \mathbb{E}\left\{\boldsymbol{\Xi}\mathbf{A}\left(\mathbf{x}^{[\mathrm{TP}]} \odot \mathbf{t}^{[\mathrm{rd}]}\right) + \boldsymbol{\Xi}\mathbf{B}\left(\mathbf{n} \odot \mathbf{t}^{[\mathrm{rd}]}\right) + \mathbf{w}\right\} &\text{(A.2a)} \\ &= \boldsymbol{\Xi}\mathbf{A}\left(\mathbb{E}\left\{\mathbf{x}^{[\mathrm{TP}]}\right\} \odot \mathbf{t}^{[\mathrm{rd}]}\right) + \boldsymbol{\Xi}\mathbf{B}\left(\mathbb{E}\left\{\mathbf{n}\right\} \odot \mathbf{t}^{[\mathrm{rd}]}\right) + \mathbb{E}\{\mathbf{w}\} &\text{(A.2b)} \\ &= \boldsymbol{\Xi}\mathbf{A}\mathbf{t}, &\text{(A.2c)}\end{aligned}$$

where $\mathbf{x}^{[\mathrm{TP}]} \triangleq [(\mathbf{x}_1^{[\mathrm{TP}]})^T, \ldots, (\mathbf{x}_K^{[\mathrm{TP}]})^T]^T$ and $\mathbf{x}_k^{[\mathrm{TP}]}$ is defined below (2). (A.2b) follows from (A.2a) because $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\nu}^{[\mathrm{rd}]}$ and $\boldsymbol{\tau}^{[\mathrm{rd}]}$ are assumed to be deterministic parameters. Furthermore, (A.2c) follows from (A.2b) since according to the assumptions in Section II-A, timing and frequency offset estimation errors from the source to the kth relay, $\delta_{\tau_k^{[\mathrm{sr}]}}$ and $\delta_{\nu_k^{[\mathrm{sr}]}}$, respectively, are modeled as zero-mean Gaussian random variables. Consequently, $\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}|_{\delta_{\tau_k^{[\mathrm{sr}]}}=\delta_{\nu_k^{[\mathrm{sr}]}}=0} = \mathbf{I}_L$, $\mathbb{E}\{\mathbf{w}\} = \mathbf{0}_{LN\times 1}$, $\mathbb{E}\{\mathbf{n}_k\} = \hat{\mathbf{G}}_k^{[\mathrm{sr}]}\hat{\boldsymbol{\Lambda}}_k^{[\mathrm{sr}]}\mathbb{E}\{\mathbf{u}_k\} = \mathbf{0}_{L\times 1}$, and $\mathbb{E}\{\mathbf{x}_k^{[\mathrm{TP}]}\} = \mathbb{E}\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\}\mathbf{t}^{[\mathrm{sr}]} = \mathbf{t}^{[\mathrm{sr}]}$. Finally $\mathbf{t}$ in (A.2c) is defined as $\mathbf{t} \triangleq [\underbrace{(\mathbf{t}^{[\mathrm{sr}]})^T, \ldots, (\mathbf{t}^{[\mathrm{sr}]})^T}_{K}]^T \odot \mathbf{t}^{[\mathrm{rd}]}$. The $LN \times LN$

received signal's covariance matrix during TP, $\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}$, can be determined as

$$\begin{aligned}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}} &= \mathbf{\Xi}\mathbf{A}\mathbb{E}\left\{(\mathbf{p}-\mathbf{t})(\mathbf{p}-\mathbf{t})^H\right\}\mathbf{A}^H\mathbf{\Xi}^H \\ &\quad + \mathbf{\Xi}\mathbf{B}\left(\mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\}\odot \mathbf{t}^{[\mathrm{rd}]}\left(\mathbf{t}^{[\mathrm{rd}]}\right)^H\right)\mathbf{B}^H\mathbf{\Xi}^H \\ &\quad + \mathbb{E}\left\{\mathbf{w}\mathbf{w}^H\right\} \qquad \text{(A.3a)} \\ &= \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H + \mathbf{\Xi}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{\Xi}^H + \mathbf{\Sigma}_{\mathbf{w}}, \qquad \text{(A.3b)}\end{aligned}$$

where $\mathbf{\Sigma}_{\mathbf{p}} \triangleq \mathbb{E}\{(\mathbf{p}-\mathbf{t})(\mathbf{p}-\mathbf{t})^H\}$ and $\mathbf{\Sigma}_{\mathbf{v}} \triangleq \mathbb{E}\{\mathbf{n}\mathbf{n}^H\}\odot \mathbf{t}^{[\mathrm{rd}]}(\mathbf{t}^{[\mathrm{rd}]})^H$ are $LK\times LK$ matrices, and $\mathbf{\Sigma}_{\mathbf{w}} \triangleq \mathbb{E}\{\mathbf{w}\mathbf{w}^H\}$ is an $LN\times LN$ matrix.

The derivatives of $\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}$ in (A.1) can be calculated as

$$\begin{aligned}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Re\{\alpha_k\}} &= -j\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Im\{\alpha_k\}} = \mathbf{\Xi}_k\mathbf{t}_k, \\ \frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \tau_k^{[\mathrm{rd}]}} &= \alpha_k\mathbf{\Lambda}_k^{[\mathrm{rd}]}\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)'\mathbf{t}_k, \\ \frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \nu_k^{[\mathrm{rd}]}} &= j\alpha_k\mathbf{D}\mathbf{\Xi}_k\mathbf{t}_k, \qquad \text{(A.4)}\end{aligned}$$

where $(\mathbf{G}_k^{[\mathrm{rd}]})' \triangleq \frac{\partial \mathbf{G}_k^{[\mathrm{rd}]}}{\partial \tau_k^{[\mathrm{rd}]}}$ and $\mathbf{D} \triangleq 2\pi\times\mathrm{diag}\{0,1,2,\ldots,LN-1\}$. Moreover, the derivative of $\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}$ in (A.1) with respect to the parameters of interest is given by

$$\begin{aligned}\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k} &\triangleq \frac{\partial \mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Re\{\alpha_k\}} = \mathbf{\Xi}\mathbf{U}_{\alpha_k}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H \\ &\quad + \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{U}_{\alpha_k}^H\mathbf{\Xi}^H, \qquad \text{(A.5a)} \\ \mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k+K} &\triangleq \frac{\partial \mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Im\{\alpha_k\}} = j\mathbf{\Xi}\mathbf{U}_{\alpha_k}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H \\ &\quad - j\mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{U}_{\alpha_k}^H\mathbf{\Xi}^H, \qquad \text{(A.5b)} \\ \mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k+2K} &\triangleq \frac{\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \nu_k^{[\mathrm{rd}]}} = \mathbf{U}_{\nu_k^{[\mathrm{rd}]}}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H \\ &\quad + \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{U}_{\nu_k^{[\mathrm{rd}]}}^H \\ &\quad + \mathbf{U}_{\nu_k^{[\mathrm{rd}]}}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{\Xi}^H \\ &\quad + \mathbf{\Xi}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{U}_{\nu_k^{[\mathrm{rd}]}}^H, \qquad \text{(A.5c)} \\ \mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k+3K} &\triangleq \frac{\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \tau_k^{[\mathrm{rd}]}} = \mathbf{U}_{\tau_k^{[\mathrm{rd}]}}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{\Xi}^H \\ &\quad + \mathbf{\Xi}\mathbf{A}\mathbf{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{U}_{\tau_k^{[\mathrm{rd}]}}^H \\ &\quad + \mathbf{U}_{\tau_k^{[\mathrm{rd}]}}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{\Xi}^H \\ &\quad + \mathbf{\Xi}\mathbf{B}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{U}_{\tau_k^{[\mathrm{rd}]}}^H, \qquad \text{(A.5d)}\end{aligned}$$

where $\mathbf{U}_{\alpha_k} \triangleq \mathrm{diag}(\mathbf{0}_{L\times L(k-1)}, \mathbf{I}_L, \mathbf{0}_{L\times L(K-k)})$ is an $LK\times LK$ matrix, and $\mathbf{U}_{\nu_k^{[\mathrm{rd}]}} \triangleq [\mathbf{0}_{LN\times L(k-1)}, j\mathbf{D}\mathbf{\Xi}_k, \mathbf{0}_{LN\times L(K-k)}]$ and $\mathbf{U}_{\tau_k^{[\mathrm{rd}]}} \triangleq [\mathbf{0}_{LN\times L(k-1)}, \mathbf{\Lambda}_k^{[\mathrm{rd}]}(\mathbf{G}_k^{[\mathrm{rd}]})', \mathbf{0}_{LN\times L(K-k)}]$ are $LN\times LK$ matrices. After substituting (A.4), (A.5a), and (A.5b) in (A.1), the upper triangular elements of FIM, for $k,\bar{k}=1,\ldots,K$, can be obtained as

$$\begin{aligned}&\mathbf{F}_{k,\bar{k}} \\ &\quad = 2\Re\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Xi}_{\bar{k}}\mathbf{t}_{\bar{k}}\right\} + \mathbf{\Pi}_{k,\bar{k}}, \qquad \text{(A.6a)} \\ &\mathbf{F}_{k,\bar{k}+K} \\ &\quad = -2\Im\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Xi}_{\bar{k}}\mathbf{t}_{\bar{k}}\right\} + \mathbf{\Pi}_{k,\bar{k}+K}, \qquad \text{(A.6b)} \\ &\mathbf{F}_{k+K,\bar{k}+K} \\ &\quad = 2\Re\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Xi}_{\bar{k}}\mathbf{t}_{\bar{k}}\right\} + \mathbf{\Pi}_{k+K,\bar{k}+K}, \qquad \text{(A.6c)} \\ &\mathbf{F}_{k,\bar{k}+2K} \\ &\quad = -2\Im\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Xi}_{\bar{k}}\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} + \mathbf{\Pi}_{k,\bar{k}+2K}, \qquad \text{(A.6d)} \\ &\mathbf{F}_{k+K,\bar{k}+2K} \\ &\quad = 2\Re\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Xi}_{\bar{k}}\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} + \mathbf{\Pi}_{k+K,\bar{k}+2K}, \qquad \text{(A.6e)} \\ &\mathbf{F}_{k+2K,\bar{k}+2K} \\ &\quad = 2\Re\left\{\alpha_k^*\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{D}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{D}\mathbf{\Xi}_{\bar{k}}\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} + \mathbf{\Pi}_{k+2K,\bar{k}+2K}, \\ &\qquad \text{(A.6f)} \\ &\mathbf{F}_{k,\bar{k}+3K} \\ &\quad = 2\Re\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Lambda}_{\bar{k}}^{[\mathrm{rd}]}\left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)'\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} \\ &\qquad + \mathbf{\Pi}_{k,\bar{k}+3K}, \qquad \text{(A.6g)} \\ &\mathbf{F}_{k+K,\bar{k}+3K} \\ &\quad = 2\Im\left\{\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Lambda}_{\bar{k}}^{[\mathrm{rd}]}\left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)'\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} \\ &\qquad + \mathbf{\Pi}_{k+K,\bar{k}+3K}, \qquad \text{(A.6h)} \\ &\mathbf{F}_{k+2K,\bar{k}+3K} \\ &\quad = 2\Im\left\{\alpha_k^*\mathbf{t}_k^H\mathbf{\Xi}_k^H\mathbf{D}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Lambda}_{\bar{k}}^{[\mathrm{rd}]}\left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)'\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} \\ &\qquad + \mathbf{\Pi}_{k+2K,\bar{k}+3K}, \qquad \text{(A.6i)} \\ &\mathbf{F}_{k+3K,\bar{k}+3K} \\ &\quad = 2\Re\left\{\alpha_k^*\mathbf{t}_k^H\left(\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)'\right)^H\left(\mathbf{\Lambda}_k^{[\mathrm{rd}]}\right)^H\right. \\ &\qquad \left.\times\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Lambda}_{\bar{k}}^{[\mathrm{rd}]}\left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)'\mathbf{t}_{\bar{k}}\alpha_{\bar{k}}\right\} + \mathbf{\Pi}_{k+3K,\bar{k}+3K}, \qquad \text{(A.6j)}\end{aligned}$$

where $[\mathbf{\Pi}]_{m,\bar{m}}$ for $m,\bar{m} = 1,\ldots,4K$ are the elements of $4K\times 4K$ matrix, $\mathbf{\Pi}$, such that $[\mathbf{\Pi}]_{m,\bar{m}} = \mathrm{Tr}\{\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m}\mathbf{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\mathbf{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},\bar{m}}\}$. Note that the lower triangular elements of $\mathbf{F}$ can be easily obtained by simple manipulation of (A.6a)–(A.6j) and are not included here for brevity. Using (A.6a)–(A.6j), $\mathbf{F}$ can be determined as given in (11).

Appendix B
Derivation of $\mathcal{J}_{\mathbf{s}}$ in (23)

For notational brevity, let us denote $\mathbf{x}^{[\mathrm{DTP}]}$ by $\mathbf{x}$ throughout this section. Substituting (8) in (22), we get

$$\begin{aligned}\mathcal{J}_{\mathbf{s}} = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u},\mathbf{w},\mathbf{s}} \\ \times\left\{\left\|\mathbf{Q}\left(\mathbf{\Xi}\mathbf{A}\mathbf{\Omega}\mathbf{x}+\mathbf{\Xi}\mathbf{B}\mathbf{\Omega}\mathbf{n}+\mathbf{w}\right)-\mathbf{s}\right\|^2\right\}. \qquad \text{(B.1)}\end{aligned}$$

Since, $\mathbf{s}$, $\mathbf{w}$ and $\mathbf{u}$ are uncorrelated, (B.1) can be expanded as

$$\begin{aligned}\mathcal{J}_{\mathbf{s}} = \mathrm{Tr}\Big\{ & \mathbf{Q}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}} \\ & \times \left\{\boldsymbol{\Xi}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\boldsymbol{\Xi}^H\right\}\mathbf{Q}^H \\ & + \mathbf{Q}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H\boldsymbol{\Xi}^H\right\}\mathbf{Q}^H \\ & - \mathbf{Q}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\boldsymbol{\Xi}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}\right\}\right\} \\ & - \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}^H\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\boldsymbol{\Xi}^H\right\}\mathbf{Q}^H \\ & + \boldsymbol{\Phi}_{\mathbf{s}} + \mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{Q}^H\Big\}, \end{aligned} \tag{B.2}$$

where $\boldsymbol{\Phi}_{\mathbf{x}} \triangleq \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{s}}\{\mathbf{x}\mathbf{x}^H\}$, $\boldsymbol{\Sigma}_{\mathbf{n}} \triangleq \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u}}\{\mathbf{n}\mathbf{n}^H\}$, $\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}} \triangleq \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{s}}\{\mathbf{x}\mathbf{s}^H\}$, $\boldsymbol{\Phi}_{\mathbf{s}} \triangleq \mathbb{E}_{\mathbf{s}}\{\mathbf{s}\mathbf{s}^H\}$ and $\boldsymbol{\Sigma}_{\mathbf{w}} \triangleq \mathbb{E}_{\mathbf{w}}\{\mathbf{w}\mathbf{w}^H\}$. These expectations are calculated as follows.

1) Calculation of $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}\}$: Since $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$, the channel matrix $\mathbf{A}$, given below (5), can be written as $\mathbf{A} = \hat{\mathbf{A}} + \boldsymbol{\Delta}_{\mathbf{A}}$, where $\hat{\mathbf{A}} \triangleq \mathrm{diag}(\hat{\boldsymbol{\alpha}}) \otimes \mathbf{I}_L$ and $\boldsymbol{\Delta}_{\mathbf{A}} \triangleq \mathrm{diag}(\boldsymbol{\delta}_{\boldsymbol{\alpha}}) \otimes \mathbf{I}_L$ are $LK \times LK$ matrices. According to the assumption in Section V, $\boldsymbol{\delta}_{\boldsymbol{\alpha}}$ is distributed as $\boldsymbol{\delta}_{\boldsymbol{\alpha}} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}})$. Thus, $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}\}$ are evaluated as

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\} \\ &\quad = \hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H + (\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \otimes \mathbf{I}_L) \\ &\qquad \odot (\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H), \end{aligned} \tag{B.3a}$$

$$\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}\} = \hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}. \tag{B.3b}$$

2) Calculation of $\boldsymbol{\Phi}_{\mathbf{s}}$ & $\boldsymbol{\Phi}_{\mathbf{x}}$: Let us assume that the transmitted data symbols are uncorrelated such that $\boldsymbol{\Phi}_{\mathbf{s}} \triangleq \mathbb{E}_{\mathbf{s}}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_L$. Furthermore, $\boldsymbol{\Phi}_{\mathbf{x}}$ is an $LK \times LK$ correlation matrix, whose $L \times L$ blocks are given by

$$\begin{aligned}&[\boldsymbol{\Phi}_{\mathbf{x}}]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} \\ &\quad = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}}\left\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\boldsymbol{\Phi}_{\mathbf{s}}\boldsymbol{\Delta}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}^H\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}^H\right\}, \\ &\qquad\qquad k,\bar{k} = 1,\ldots,K, \end{aligned} \tag{B.4}$$

Based on the assumption in Section V, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{sr}]},\boldsymbol{\tau}^{[\mathrm{sr}]}})$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} \sim \mathcal{N}(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{sr}]},\boldsymbol{\nu}^{[\mathrm{sr}]}})$. Thus, by applying Taylor series expansion around their means and by ignoring the second and higher order terms, the non-linear terms $\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}$ and $\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}$, which are functions of $\delta_{\nu_k^{[\mathrm{sr}]}}$ and $\delta_{\tau_k^{[\mathrm{sr}]}}$, respectively, can be written as

$$\begin{aligned}\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} &= \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]} + \left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\right)' \delta_{\nu_k^{[\mathrm{sr}]}} + \mathcal{O}_p\left(\sigma_{\nu_k^{[\mathrm{sr}]}}^2\right), \\ \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} &= \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]} + \left(\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{sr}]}}^2\right), \end{aligned} \tag{B.5}$$

where $\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]} \triangleq \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}|_{\delta_{\nu_k^{[\mathrm{sr}]}}=0}$, $\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]} \triangleq \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}|_{\delta_{\tau_k^{[\mathrm{sr}]}}=0}$, $(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]})' = \frac{\partial \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}}{\partial \delta_{\nu_k^{[\mathrm{sr}]}}}|_{\delta_{\nu_k^{[\mathrm{sr}]}}=0}$, $(\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]})' = \frac{\partial \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}}{\partial \delta_{\tau_k^{[\mathrm{sr}]}}}|_{\delta_{\tau_k^{[\mathrm{sr}]}}=0}$. Substituting (B.5) into (B.4), we arrive at (B.6), where (B.6b) follows from (B.6a) because the expectations in (B.6b) are equal to $\mathbf{0}_{L\times L}$ according to the distributions of $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}$ and $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}$.

$$\begin{aligned}&[\boldsymbol{\Phi}_{\mathbf{x}}]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} \\ &\quad = \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Phi}_{\mathbf{s}}\left(\boldsymbol{\Delta}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H \\ &\qquad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}}\left\{\left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\right)'\delta_{\nu_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Phi}_{\mathbf{s}}\right. \\ &\qquad\qquad \left.\times\left(\boldsymbol{\Delta}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\right\} \\ &\qquad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}}\left\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\left(\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\right)'\delta_{\tau_k^{[\mathrm{sr}]}}\boldsymbol{\Phi}_{\mathbf{s}}\right. \\ &\qquad\qquad \left.\times\left(\boldsymbol{\Delta}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\right\} \\ &\qquad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}}\left\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Phi}_{\mathbf{s}}\left(\left(\boldsymbol{\Delta}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)'\right)^H\right. \\ &\qquad\qquad \left.\times\delta_{\tau_k^{[\mathrm{sr}]}}\left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\right\} \\ &\qquad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}}\left\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Phi}_{\mathbf{s}}\left(\boldsymbol{\Delta}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)^H\right. \\ &\qquad\qquad \left.\times\left(\left(\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}^{[0]}\right)'\right)^H\delta_{\nu_{\bar{k}}^{[\mathrm{sr}]}}\right\} + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{sr}]}}^2\right), \\ &\qquad\qquad k,\bar{k} = 1,\ldots,K, \end{aligned} \tag{B.6a}$$

$$\begin{aligned}&[\boldsymbol{\Phi}_{\mathbf{x}}]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} \approx \boldsymbol{\Phi}_{\mathbf{s}}, \\ &\qquad\qquad k,\bar{k} = 1,\ldots,K, \end{aligned} \tag{B.6b}$$

3) Calculation of $\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}$: Using (B.5) and (B.6), the $L \times L$ blocks of the $LK \times LK$ matrix $\boldsymbol{\Phi}_{\mathbf{x}}$ are given by

$$\begin{aligned}[\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}]_{(k-1)L+1:kL,:} &= \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}}\left\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\boldsymbol{\Phi}_{\mathbf{s}}\right\} \\ &\approx \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}^{[0]}\boldsymbol{\Phi}_{\mathbf{s}} = \boldsymbol{\Phi}_{\mathbf{s}}. \end{aligned} \tag{B.7}$$

4) Calculation of $\boldsymbol{\Sigma}_{\mathbf{n}}$: The $L \times L$ blocks of the $LK \times LK$ matrix $\boldsymbol{\Sigma}_{\mathbf{n}}$ can be calculated as shown in (B.8),

$$\begin{aligned}&[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} \\ &\quad = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u}}\left\{\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}\hat{\boldsymbol{\Lambda}}_k^{[\mathrm{sr}]}\mathbf{u}_k\mathbf{u}_{\bar{k}}^H\left(\hat{\boldsymbol{\Lambda}}_{\bar{k}}^{[\mathrm{sr}]}\right)^H\left(\hat{\boldsymbol{\mathcal{G}}}_{\bar{k}}^{[\mathrm{sr}]}\right)^H\right\} \end{aligned} \tag{B.8a}$$

$$\begin{aligned}&= \begin{cases}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}}\left\{\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}\boldsymbol{\Sigma}_{\mathbf{u}_k}\left(\hat{\boldsymbol{\mathcal{G}}}_{\bar{k}}^{[\mathrm{sr}]}\right)^H\right\}, & k = \bar{k} \\ \mathbf{0}_{L\times L}, & k \neq \bar{k}\end{cases} \\ &\qquad\qquad k,\bar{k} = 1,\ldots,K, \end{aligned} \tag{B.8b}$$

where $\boldsymbol{\Sigma}_{\mathbf{u}_k} \triangleq \sigma_{u_k}^2\mathbf{I}_{LN}$ and (B.8b) follows from (B.8a) since the noise at different relays, $\mathbf{u}_k$ and $\mathbf{u}_{\bar{k}}$, $\forall\, k \neq \bar{k}$, are uncorrelated. Note that $\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}$ is a non-linear function of $\hat{\tau}_k^{[\mathrm{sr}]} \triangleq \tau_k^{[\mathrm{sr}]} + \delta_{\tau_k^{[\mathrm{sr}]}}$.

Thus, in order to evaluate (B.8b), the Taylor series expansion of the non-linear term $\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}$ around $\tau_k^{[\mathrm{sr}]}$ is evaluated as

$$\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]} = \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} + \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{sr}]}}\right), \tag{B.9}$$

where $\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \triangleq \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}\big|_{\hat{\tau}_k^{[\mathrm{sr}]}=\tau_k^{[\mathrm{sr}]}}$, $(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]})' = \frac{\partial \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}}{\partial \hat{\tau}_k^{[\mathrm{sr}]}}\big|_{\hat{\tau}_k^{[\mathrm{sr}]}=\tau_k^{[\mathrm{sr}]}}$, and since the noise at the relays are uncorrelated, $\boldsymbol{\Sigma}_{\mathbf{n}}$ is an $LK \times LK$ block diagonal matrix. Using (B.9) in (B.8), the $L \times L$ diagonal matrices of $\boldsymbol{\Sigma}_{\mathbf{n}}$ are given by

$$\begin{aligned}
&[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L+1:kL,(k-1)L+1:kL} \\
&\quad = \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \boldsymbol{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H \\
&\qquad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} \boldsymbol{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H \right\} \\
&\qquad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \boldsymbol{\Sigma}_{\mathbf{u}_k} \left(\left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' \right)^H \delta_{\tau_k^{[\mathrm{sr}]}} \right\} \\
&\qquad + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{sr}]}}\right) \\
&\quad = \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \boldsymbol{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{sr}]}}\right) \\
&\quad \approx \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \boldsymbol{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H.
\end{aligned} \tag{B.10}$$

5) Calculation of $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\{\boldsymbol{\Xi}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\}\boldsymbol{\Xi}^H\}$ *and* $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\{\boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H\boldsymbol{\Xi}^H\}$*:* Using similar steps as in (B.5), (B.6), (B.9), and (B.10), the expectations over $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$ in (B.2) can be approximated by

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}} \left\{\boldsymbol{\Xi}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\boldsymbol{\Xi}^H\right\} \\
&\quad = \hat{\boldsymbol{\Xi}}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\hat{\boldsymbol{\Xi}}^H + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{rd}]}}\right) \\
&\quad \approx \hat{\boldsymbol{\Xi}}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\hat{\boldsymbol{\Xi}}^H,
\end{aligned} \tag{B.11a}$$

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}} \{\boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H\boldsymbol{\Xi}^H\} \\
&\quad = \hat{\boldsymbol{\Xi}}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H\hat{\boldsymbol{\Xi}}^H + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{rd}]}}\right) \\
&\quad \approx \hat{\boldsymbol{\Xi}}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H\hat{\boldsymbol{\Xi}}^H,
\end{aligned} \tag{B.11b}$$

where $\hat{\boldsymbol{\Xi}} \triangleq \boldsymbol{\Xi}|_{\boldsymbol{\nu}^{[\mathrm{rd}]}=\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}=\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\}$ is given in (B.3a).

By combining the results derived in (B.3), (B.6), (B.7), (B.10), (B.11a), and (B.11b) in (B.2), the cost function in (23) is obtained.

REFERENCES

[1] Z. Sheng, K. K. Leung, and Z. Ding, "Cooperative wireless networks: From radio to network protocol designs," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 64–69, May 2011.

[2] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperative diversity—Part I: System description; Part II: Implementation and performance analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.

[3] 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.

[4] S. Ma, Y. Yang, and H. Sharif, "Distributed MIMO technologies in cooperative wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 78–82, May 2011.

[5] M.-K. Chang and S.-Y. Lee, "Performance analysis of cooperative communication system with hierarchical modulation over Rayleigh fading channel," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2848–2852, June 2009.

[6] A. A. Nasir, H. Mehrpouyan, S. D. Blostein, S. Durrani, and R. A. Kennedy, "Timing and carrier synchronization with channel estimation in multi-relay cooperative networks," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 793–811, Feb. 2012.

[7] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[8] F. Khan, Y. Chen, and M. Alouini, "Novel receivers for AF relaying with distributed STBC using cascaded and disintegrated channel estimation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1370–1379, Apr. 2012.

[9] T. Q. Duong, G. C. Alexandropoulos, H. Zepernick, and T. A. Tsiftsis, "Orthogonal space-time block codes with CSI-assisted amplify-and-forward relaying in correlated Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 882–889, Mar. 2011.

[10] X. Li, C. Xing, Y.-C. Wu, and S. C. Chan, "Timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2218–2229, Apr. 2010.

[11] J. Harshan and B. S. Rajan, "Distributed space-time block codes for two-hop wireless relay networks," *J. Commun.*, vol. 5, no. 4, pp. 282–296, Apr. 2010.

[12] H. M. Wang and X. G. Xia, "Asynchronous cooperative communication systems: A survey on signal designs," *Sci. China Inf. Sci.*, vol. 54, no. 8, pp. 1547–1561, Apr. 2011.

[13] Z. Zhong, S. Zhu, and A. Nallanathan, "Delay-tolerant distributed linear convolutional space-time code with minimum memory length under frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3944–3949, Aug. 2009.

[14] H. Wang, Q. Yin, and X.-G. Xia, "Full diversity space-frequency codes for frequency asynchronous cooperative relay networks with linear receivers," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 236–247, Jan. 2011.

[15] D. Wang and S. Fu, "Asynchronous cooperative communications with STBC coded single carrier block transmission," in *Proc. IEEE GLOBECOM*, 2007.

[16] Z. Li, X.-G. Xia, and M. H. Lee, "A simple orthogonal space-time coding scheme for asynchronous cooperative systems for frequency selective fading channels," *IEEE Trans. Commun.*, vol. 58, no. 8, pp. 2219–2224, Aug. 2010.

[17] Y. Yao and X. Dong, "On the detection of distributed STBC AF cooperative OFDM signal in the presence of multiple CFOs," in *Proc. IEEE ICC*, 2010.

[18] Z. Li and X.-G. Xia, "An alamouti coded OFDM transmission for cooperative systems robust to both timing errors and frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1839–1844, May 2008.

[19] 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, July 2010.

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

[21] H. Wang, X.-G. Xia, and Q. Yin, "Computationally efficient equalization for asynchronous cooperative communications with multiple frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 648–655, Feb. 2009.

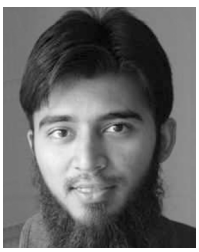
[22] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Achieving cooperative diversity with multiple frequency offset estimation," in *Proc. IEEE ICC*, 2011.

[23] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*, G. Rozenberg, T. Bäck, J. Kok, H. Spaink, and A. E. Eiben, Eds. Berlin, Germany: Springer-Verlag, 2005.

[24] O. Besson and P. Stoica, "On parameter estimation of MIMO flat-fading channels with frequency offsets," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 602–613, Mar. 2003.

[25] H. Mehrpouyan and S. Blostein, "Estimation, training, and effect of timing offsets in distributed cooperative networks," in *Proc. IEEE GLOBECOM*, 2010.

[26] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*, ser. Wiley Series in Telecommun. and Signal Process., J. G. Proakis, Ed. New York, NY, USA: Wiley, 1998.
[27] H. Mehrpouyan and S. D. Blostein, "Comments on "timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Signal Process. Lett.*, vol. 59, pp. 4047–4048, Aug. 2011.
[28] I. Ziskand and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.
[29] F. Neri and V. Tirronen, "Recent advances in differential evolution: A survey and experimental analysis," *Artificial Intelligence Review*, vol. 33, no. 1–2, pp. 61–106, Oct. 2009.
[30] G. Rudolph, "Convergence of evolutionary algorithms in general search space," in *Proc. IEEE Int. Conf. Evol. Computat.*, May 1996, pp. 50–54.
[31] M. Vasile, E. Minisci, and M. Locatelli, "An inflationary differential evolution algorithm for space trajectory optimization," *IEEE Trans. Evolution. Computat.*, vol. 15, no. 2, pp. 267–281, Apr. 2011.
[32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[33] Y. Tian, X. Lei, Y. Xiao, and S. Li, "SAGE based joint timing-frequency offsets and channel estimation in distributed MIMO systems," *Elsevier J. Comput. Commun.*, vol. 33, no. 17, pp. 2125–2131, July 2010.
[34] Y. M. Bishop, S. E. Fienberg, and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice*. New York, NY, USA: Springer, 2007.



Ali A. Nasir (S'11) received the B.Sc. (1st class honors) degree in electrical engineering from the University of Engineering and Technology (UET), Lahore, Pakistan, in 2007. He received the Ph.D. degree in electrical engineering from the Research School of Engineering, Australian National University (ANU), Canberra, in 2013.

He is currently a Postdoctoral Research Fellow in the Applied Signal Processing Group, ANU.

Hani Mehrpouyan (S'05–M'10) received the B.Sc. honors degree in computer engineering from Simon Fraser University, Burnaby, Canada, in 2004. He received the Ph.D. degree in electrical engineering in 2010 from the Department of ECE, Queen's University, Kingston, Canada.

He is now an Assistant Professor with the Department of Computer and Electrical Engineering and Computer Science, California State University, Bakersfield. For more information refer to www.mehrpouyan.info.

Salman Durrani (S'00–M'05–SM'10) received the B.Sc. (1st class honors) degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2000. He received the Ph.D. degree in electrical engineering from the University of Queensland, Brisbane, Australia, in December 2004.

He has been with the Australian National University, Canberra, since 2005, where he is currently Senior Lecturer in the Research School of Engineering, College of Engineering and Computer Science.

Steven D. Blostein (S'83–M'88–SM'96) received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, in 1983, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois, Urbana-Champaign, in 1985 and 1988, respectively.

He has been on the faculty of the Department of Electrical and Computer Engineering, Queen's University, Kingston, Canada, since 1988 and currently holds the position of Professor.

Rodney A. Kennedy (S'86–M'88–S'01–F'05) received the B.E. degree from the University of New South Wales, Sydney, Australia, the M.E. degree from the University of Newcastle, and the Ph.D. degree from the Australian National University (ANU), Canberra.

He is currently a Professor in the Research School of Engineering, ANU. His research interests include digital signal processing, digital and wireless communications, and acoustical signal processing.

Björn Ottersten (F'04) was born in Stockholm, Sweden, in 1961. He received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986. He received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1989.

Currently, he is Director for the Interdisciplinary Centre for Security, Reliability and Trust at the University of Luxembourg.