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]–[3]. Amongst the different relaying protocols [4], 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 [5]. 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 [6], [7]. The distributed space-time block code (DSTBC) cooperative strategy allows relays to transmit in the same time slot using the DSTBC structure [8]–[10]. Given its potential, the combination of AF relaying and DSTBC has attracted recent research interest as an efficient means of providing cooperative diversity [11]–[13]. 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 [7]. It has been shown that time asynchronism can destroy the elaborate structure of the DSTBCs, reducing the achievable diversity gain [14], [15]. 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 [16].

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 [11], [12], [17], [18]. Others have focused on the design of delay tolerant DSTBC [19], [20] and carrier frequency offset tolerant space-frequency codes [21] for decode-and-forward (DF) systems. However, the algorithms in [19]–[21] 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 offset. More importantly, [19] and [21] provide no means of estimating MTOs and MCFOs, even though they require these parameters to be obtained at the destination for successful decoding.

Many algorithms have been proposed for estimation and
compensation of channel gains [13], [22], MTOs [23]–[25], or MCFOs [26], [27] in DSTBC-orthogonal frequency division multiplexing (OFDM) AF relaying systems. However, the algorithms in [13], [22]–[27] 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 [28]–[30], 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 [13], [22]–[26], [28]–[30] 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 [13], [22]–[26], [28]–[30] is very limited.

For general AF relaying with DSTBC, channel estimation and equalization [11], [31], [32], joint channel and MCFO [33]–[35], or joint channel and MTO estimation and compensation [14] 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 [27], [33]–[35] or vice versa [14]. 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, [7] addresses the problem of timing and carrier synchronization in AF and DF cooperative networks. However, the transceiver designs and estimation algorithms proposed in [7] 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 [7] 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 [36]. 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:

- 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 $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^\prime$, and $(\cdot)^\prime\prime$ 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. $E_x\{\cdot\}$ 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 $(\cdot)^{[\text{SR}]}$ and $(\cdot)^{[\text{DTP}]}$ denote the source to relay and relay to destination parameters, respectively. Symbols with superscripts $(\cdot)^{[\text{TP}]}$ and $(\cdot)^{[\text{DTP}]}$ denote the signals in training and data transmission periods, respectively. The operator, $\hat{\cdot}$ 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 $\mu$ and variance $\sigma^2$, respectively. $\mathcal{U}(-x, x)$ denotes a uniform distribution between $-x$ and $x$. Boldface small letters, $\mathbf{x}$ and $\mathbf{y}$ denote the covariance matrix of $\mathbf{x}$, where its diagonal elements are $\mathbf{x}_i, i = 1, \ldots, K$, and columns, $x$, to $y$, of the matrix $\mathbf{X}$. $\mathbf{1}_{X \times X}$ is the identity matrix, all zero, and all one matrices, respectively. $||\mathbf{x}||_2$ represents the $\ell_2$ norm of a vector $\mathbf{x}$ and $\text{Tr} \{ \mathbf{X} \}$ denotes the trace of $\mathbf{X}$. $\text{diag} (\mathbf{x})$ is used to denote a diagonal matrix, where its diagonal elements are given by the vector $\mathbf{x}$ and $\text{diag} (\mathbf{X}_1, \mathbf{X}_2, \ldots)$ represents a block diagonal matrix, where $\mathbf{X}_1, \mathbf{X}_2, \ldots$ are the diagonal matrices. Finally, $\Sigma_x, \Phi_x \triangleq \mathbb{E} \{ \mathbf{x} \mathbf{y}^H \}$ and $\Phi_x \triangleq \mathbb{E} \{ \mathbf{x} \mathbf{x}^H \}$ 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 TRANSEIVER DESIGN

As shown in Fig. 1, a half-duplex AF cooperative system with one source node, $S$, $K$ relays, $R_1, \ldots, R_K$, and a single destination node, $D$ is considered. Each node is equipped with a single omnidirectional antenna. The channels from the source to the $k$th relay and the $k$th 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, $\tau_k$ and $\nu_k$ 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 $k$th relay and from the $k$th 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], [14], [33], [37]–[41].

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], [14], [33], [37]–[41]. 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], [33], and [37].

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 $k$th 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 $k$th relay, $r_k(t), t = 0, 1, \ldots, T_n$, is down-converted by oscillator frequency $\omega_k$. The received signal is oversampled by a factor $N$, such that $T = T_nN$, where $T$ is the symbol duration and $T_n$ 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) Training Model at Relays: During TP, the sampled received signal vector at the $k$th relay, $r_k = [r_k(0), \ldots, r_k(LN - 1)]^T$, is given by

$$r_k = h_k \mathbf{A}_k^{[sr]} \mathbf{G}_k^{[sr]} \mathbf{k}_k^{[sr]} + \mathbf{u}_k, \quad k = 1, \ldots, K \quad (1)$$

The detailed design of the training sequences is outside the scope of this paper.
where

- $h_k$ denotes the unknown channel gain from source to the $k$th relay that changes from one frame to another frame following the distribution, $h_k \sim \mathcal{CN}(0, \sigma_h^2)$,
- $\Lambda_k \triangleq \text{diag} \left( e^{j2\pi \nu_1^s(0)/N}, \ldots, e^{j2\pi \nu_{L-1}^s(L-1)/N} \right)$ is a diagonal $LN \times LN$ matrix, $\nu_k^s$ denotes the unknown carrier frequency offset, normalized by the symbol duration $T_s$, between source and the $k$th relay,
- $G_k^{[sr]}$ is the $LN \times L$ matrix of the samples of the pulse shaping filter such that $G_k^{[sr]} = g_{rc} (iT_s - nT - \tau_k^{[sr]}), \tau_k^{[sr]}$ denotes the normalized unknown timing offset between source and the $k$th relay, $g_{rc}(t)$ stands for the root raised-cosine pulse shaping function,
- $t^{[sr]} \triangleq \left[ t^{[sr]}(0), \ldots, t^{[sr]}(L-1) \right]^T$ is the TS transmitted from source to relays, and
- $u_k \triangleq \left[ u_k(0), \ldots, u_k(LN-1) \right]^T$, $u_k(i) \forall i$, denotes the zero-mean complex additive white Gaussian noise (AWGN) at the $i$th sample of the received signal, i.e., $u_k(i) \sim \mathcal{CN}(0, \sigma_u^2)$.

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

The estimates of the timing and frequency offsets at the $k$th relay, $\hat{\tau}_k^{[sr]}$ and $\hat{\nu}_k^{[sr]}$, respectively, are used for matched filtering and frequency offset correction at the $k$th 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., [42], [43], can be applied to obtain $\hat{\tau}_k^{[sr]}$ and $\hat{\nu}_k^{[sr]}$, $\forall k$, and will not be discussed here.

Let the timing and frequency offset estimation errors between source and the $k$th relay be denoted by $\delta_{\tau_k^{[sr]}} = \hat{\tau}_k^{[sr]} - \tau_k^{[sr]}$ and $\delta_{\nu_k^{[sr]}} = \hat{\nu}_k^{[sr]} - \nu_k^{[sr]}$, respectively. According to the results in [14], $\delta_{\tau_k^{[sr]}}$ and $\delta_{\nu_k^{[sr]}}$ are Gaussian distributed, i.e., $\delta_{\tau_k^{[sr]}} \sim \mathcal{N}(0, \sigma_{\tau_k^{[sr]}}^2)$ and $\delta_{\nu_k^{[sr]}} \sim \mathcal{N}(0, \sigma_{\nu_k^{[sr]}}^2)$ [14], where $\sigma_{\tau_k^{[sr]}}^2$ and $\sigma_{\nu_k^{[sr]}}^2$ are timing and frequency offset estimation error variances, respectively. In this paper, $\sigma_{\tau_k^{[sr]}}^2$ and $\sigma_{\nu_k^{[sr]}}^2$ are set to their respective lower bounds in point-to-point systems given in [43, Ch. 6]. Taking into account the timing and frequency offset estimation errors, the vector of imperfectly synchronized received signals at the $k$th relay, $b_k^{[TP]} \triangleq \left[ b_k^{[TP]}(0), \ldots, b_k^{[TP]}(L-1) \right]^T$, as shown in Fig. 2, can be written as [43,44],

$$b_k^{[TP]} = h_k x_k^{[TP]} + n_k,$$

where

- $x_k^{[TP]} \triangleq \Delta_{\Lambda_k^{[sr]}} \Delta_{G_k^{[sr]}} t^{[sr]}, \Delta_{\Lambda_k^{[sr]}} \Delta_{G_k^{[sr]}} \triangleq \text{diag} \left( e^{j2\pi \delta_{\tau_k^{[sr]}}(0)/N}, \ldots, e^{j2\pi \delta_{\tau_k^{[sr]}}(L-1)/N} \right)$ is an $L \times L$ diagonal matrix, $\Delta_{G_k^{[sr]}}$ is an $L \times L$ matrix such that $\Delta_{G_k^{[sr]}} = g_{rc} (nT - \bar{n}T - \delta_{\nu_k^{[sr]}}), \text{ for } n, \bar{n} = 0, \ldots, L-1, g_{rc}(t)$ stands for the raised-cosine function, which results from the convolution of $g_{rc}(t)$ and root raised-cosine matched filter,
- $n_k \triangleq G_k^{[sr]} \Delta_k^{[sr]} u_k$ is the matched filtered noise, $G_k^{[sr]}$ is an $LN \times LN$ matrix such that $G_k^{[sr]} = g_{rc}(nT - iT_s - \hat{\nu}_k^{[sr]}), \text{ and } G_k^{[sr]}$ is an $LN \times LN$ matrix such that $\Delta_{\Lambda_k^{[sr]}} \Delta_{G_k^{[sr]}} \triangleq \left[ \tau_k^{[sr]}(n), \nu_k^{[sr]}(n), \alpha_k^{[sr]}, \nu_k^{[sr]}(L-1) \right]$. 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 $\omega_d$ and then oversampled by the factor $N$. Furthermore, multiple channel

![Fig. 3. Proposed AF receiver design at the destination.](image-url)
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, $y^{[TP]} \triangleq [y^{[TP]}(1), \ldots, y^{[TP]}(LN - 1)]^T$, is given by

$$y^{[TP]} = \sum_{k=1}^{K} \zeta_k f_k A_k^{[rd]} G_k^{[rd]} (b_k^{[TP]} \otimes t_k^{[rd]}) + w,$$

where

- $f_k$ denotes the unknown channel gain from the $k$th relay to the destination, which changes from frame to frame following the distribution $f_k \sim \mathcal{CN}(0, \sigma_f^2)$,
- $\zeta_k = 1/\sqrt{\sigma_h^2 + \sigma_w^2}$ satisfies the $k$th relay’s power constraint,
- $\Lambda_k^{[rd]} \triangleq \text{diag} \left( e^{i2\pi\nu_k^{[rd]}(0)/N}, \ldots, e^{i2\pi\nu_k^{[rd]}(L-1)/N} \right)$ is an $LN \times LN$ matrix, $\nu_k^{[rd]}$ denoted the normalized unknown frequency offset from the $k$th relay to the destination,
- $G_k^{[rd]} \triangleq \begin{bmatrix} g_k & 0 & \cdots & 0 \\
0 & g_k & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & g_k \end{bmatrix} \in \mathbb{C}^{LN \times LN}$

$g_k$ denotes the channel gain from the $k$th relay to the destination, and $\pi_n \triangleq \{\pi_k^{[rd]}(nT) : n = 0, \ldots, LN-1\}$ is the known training signal for the $k$th relay.

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

$$y^{[TP]} = \sum_{k=1}^{K} \alpha_k A_k^{[rd]} G_k^{[rd]} p_k + \sum_{k=1}^{K} \beta_k A_k^{[rd]} G_k^{[rd]} v_k + w,$$

where $p_k \triangleq x_k^{[TP]} \otimes t_k^{[rd]}$, $v_k \triangleq n_k \otimes t_k^{[rd]}$, $\alpha_k \triangleq \zeta_k f_k h_k$ and

$\beta_k \triangleq \zeta_k f_k$. The general model in (5) can be written compactly in matrix form as

$$y^{[TP]} = \Xi x^{[TP]} + \Xi B v + w,$$

where

- $\Xi \triangleq [\Xi_1, \ldots, \Xi_K]$ is an $LN \times K$ matrix, $\Xi_k \triangleq A_k^{[rd]} G_k^{[rd]}$ is the $LN \times L$ matrix of the $k$th relay’s frequency offset, $\nu_k^{[rd]}$, and timing offset, $\tau_k^{[rd]}$,
- $\Lambda \triangleq \text{diag}(\alpha_1, \ldots, \alpha_K) \otimes I_L$ is the $L \times L$ matrix of the overall channel gains from source-to-relays-to-destination,
- $\mathbf{B} \triangleq \text{diag}(\beta_1, \ldots, \beta_K) \otimes I_L$ is the $L \times L$ matrix of relay-to-destination channels,
- $\mathbf{w} \triangleq [w(1), \ldots, w(LN-1)]^T$ denotes the zero-mean complex AWGN at the $k$th sample of the received signal, i.e., $w(i) \sim \mathcal{CN}(0, \sigma_w^2)$.

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

$$y^{[DTP]} = \sum_{k=1}^{K} \beta_k A_k^{[rd]} G_k^{[rd]} \Omega_k b_k^{[DTP]} + w,$$

where $\Omega_k$ is the predefined $L \times L$ STBC matrix at the $k$th relay. Substituting (3) into (7), the signal model in (7) can be written in matrix form as

$$y^{[DTP]} = \Xi A \Omega x^{[DTP]} + \Xi B v + w,$$

where

- $\Omega \triangleq \text{diag}(\Omega_1, \ldots, \Omega_K)$, $x^{[DTP]} \triangleq \begin{bmatrix} x_1^{[DTP]} \ldots x_K^{[DTP]} \end{bmatrix}^T$, and $\mathbf{n} \triangleq [n_1^T, \ldots, n_K^T]^T$.

The estimates obtained during the TP are used to design matrix, $Q$, 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 [14], 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 [45]. More importantly, the authors in [14] 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, $u_k$, $\nu_k$, and destination, $w$ in (4), are mutually independent. Accordingly, the received training signal at the destination, $y^{[TP]}$ in (6), is a circularly symmetric complex Gaussian random variable, $y^{[TP]} \sim \mathcal{CN}(\mu_y^{[TP]}, \Sigma_y^{[TP]})$, with mean $\mu_y^{[TP]}$ and covariance matrix $\Sigma_y^{[TP]}$ that are given by (see Appendix A)

$$\mu_y^{[TP]} = \sum_{k=1}^{K} \alpha_k \Xi_t k, \Xi_t k = \Xi A t, \quad \text{and}$$

$$\Sigma_y^{[TP]} = \Xi A \Sigma_p A^H \Xi + \Xi A \Sigma_r A^H \Xi + \Sigma_w,$$

respectively. In (9), $t_k \triangleq (t_k^{[rd]} \otimes t_k^{[rd]})$, $t \triangleq (t_1^{[rd]} \ldots t_K^{[rd]})^T$, $\Sigma_p \triangleq \sigma_p^2 I_{LN}$, $\Sigma_r \triangleq \mathbb{E}\{ (p - t)^H (p - t) \}$, $\Sigma_v \triangleq \mathbb{E}\{ n^H \otimes t \}$ and $\Xi A$, $\Xi B$, and $\Xi$ 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, $\tau^{[rd]} \triangleq \begin{bmatrix} \tau_1^{[rd]} \ldots \tau_K^{[rd]} \end{bmatrix}^T$ and $\nu^{[rd]} \triangleq \begin{bmatrix} \nu_1^{[rd]} \ldots \nu_K^{[rd]} \end{bmatrix}^T$, respectively, and the overall channel gains from source to relays to destination, $\alpha \triangleq [\alpha_1, \ldots, \alpha_K]^T$. As a result, the parameter vector of interest, $\lambda$, is given by

$$\lambda \triangleq \begin{bmatrix} \Re(\alpha)^T, \Im(\alpha)^T, (\nu^{[rd]})^T, (\tau^{[rd]})^T \end{bmatrix}^T.$$
Theorem 1: Based on the proposed AF-DSTBC relaying scheme, the FIM for the estimation of the channel is given by (11) at the bottom of this page, where

- $\Psi \triangleq [\Xi_1 t_1, \ldots, \Xi_K t_k]$ is an $LN \times K$ matrix,
- $\Gamma \triangleq \left| A_1 \left( G_{[d]}^{[rd]} \right)^T t_1, \ldots, A_K \left( G_{[d]}^{[rd]} \right)^T t_K \right|$ is an $LN \times K$ matrix,
- $\left( G_{[d]}^{[rd]} \right) \triangleq \frac{\partial G_{[d]}^{[rd]}}{\partial \tau}, \forall k$, is an $LN \times L$ matrix,
- $D \triangleq 2\pi \times \text{diag}\{0, 1, 2, \ldots, LN - 1\}$ is an $LN \times LN$ matrix,
- $\mathcal{A} \triangleq \text{diag}\{\alpha_1, \ldots, \alpha_K\}$ is a $K \times K$ matrix,
- $\Pi$ is a $4K \times 4K$ matrix, such that
  
  $$[\Pi]_{m, m} = \text{Tr} \left\{ \Sigma_{\psi, \nu}^{-1} \Sigma_{\psi, \nu}^{-1} \right\} \text{for} \ m, m = 1, \ldots, 4K,$$

- $\Sigma_{\psi, \nu}^{-1}$ is an $LN \times LN$ matrix that is defined at the bottom of this page,
- $U_{\alpha_\nu} \triangleq \text{diag}(0_{L \times L(k-1)}, I_L, 0_{L \times L(L-k)})$ is an $L \times LK$ matrix, and
- $U_\nu \triangleq \left[ 0_{L \times L(L-k)}, \frac{jD}{2} \Xi_k, 0_{L \times L(L-k)} \right]$ and $U_\nu \triangleq \left[ 0_{L \times L(L-k)}, \frac{jD}{2} \Xi_k, 0_{L \times L(L-k)} \right]$ are $LN \times LK$ matrices.

Proof: See Appendix A. Combining the real and imaginary parts of the channel, the new set of parameters, $\hat{\lambda} \triangleq \left[ \alpha^T, (\nu^{[rd]})^T, (\tau^{[rd]})^T \right]^T$, is given by

$$\hat{\lambda} = \left[ \begin{array}{cccc} I_K & jI_K & 0_{K \times K} & 0_{K \times K} \\ 0_{K \times K} & I_K & 0_{K \times K} & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & I_K & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & 0_{K \times K} & I_K \end{array} \right] \lambda. \quad (12)$$

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

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

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

Remark 1: Eq. (11) shows that the FIM for the estimation of $\lambda$ 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 relay cooperation systems. More importantly, this result indicates that the previously proposed methods that assume perfect timing or frequency synchronization while estimating MCFOs and MTOs in [33] and [14], respectively, cannot be applied to jointly estimate and compensate these impairments in AF-DSTBC relay cooperative networks. Finding is also corroborated by the simulation results in Section VI-B, where it is shown that application of the algorithms in [14], [33] 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, $\alpha$, $\tau^{[rd]}$, and $\nu^{[rd]}$, can be determined by minimizing the cost function, $J_{\alpha, \tau^{[rd]}, \nu^{[rd]}}$, according to

$$J_{\alpha, \tau^{[rd]}, \nu^{[rd]}} = \| y^{[TP]} - \left[ \Lambda_{[d]}^{[rd]} B_{[d]}^{[rd]} \right] \Gamma \|_2^2 = \| y^{[TP]} - \psi \alpha \|_2^2. \quad (14)$$

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

$$\hat{\alpha} = \left( \psi^H \psi \right)^{-1} \psi^H y^{[TP]}. \quad (15)$$
Substituting (15) in (14), the estimates of MTOs and MCFOs, \( \hat{\tau}^{[\text{rd}]} \) and \( \hat{\nu}^{[\text{rd}]} \), respectively, are obtained via

\[
\hat{\tau}^{[\text{rd}]}(\nu^{[\text{rd}]}) = \arg \min_{\tau^{[\text{rd}]}(\nu^{[\text{rd}]})} \left( -\frac{1}{2} \text{Tr} (Y_T^P H \Upsilon (H^\dagger \Upsilon)^{-1} H^\dagger Y_T^P) \right),
\]

where \( \arg \min \) denotes the arguments, \( \tau^{[\text{rd}]}(\nu^{[\text{rd}]}) \), that minimize the expression \( \chi (\tau^{[\text{rd}]}(\nu^{[\text{rd}]}) , \nu^{[\text{rd}]}) \) and \( Y_T^P \), defined in (6), is a function of timing and frequency offset estimation errors from the source-to-relay link. The channel estimates, \( \hat{\alpha} \), are obtained by substituting \( \hat{\tau}^{[\text{rd}]}(\nu^{[\text{rd}]}) \) and \( \hat{\nu}^{[\text{rd}]} \) 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. 6 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, \( \nu^{[\text{rd}]} \) and \( \tau^{[\text{rd}]} \), 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 [46]. 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) [36]. 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 [36]. Let \( d \), for \( d = 1, \ldots, D_1 \), 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 \( 4 \)th target vector, \( \theta_{1,1} = [\tau_{1,1}^{[\text{rd}]} , \ldots , \tau_{K,1}^{[\text{rd}]} , \nu_{1,1}^{[\text{rd}]} , \ldots , \nu_{K,1}^{[\text{rd}]} ]^T \), for \( \ell = 1, \ldots , D_P \), is generated by selecting \( \nu_{1,1}^{[\text{rd}]} \) and \( \tau_{k,1}^{[\text{rd}]} \) uniformly and randomly from the set of possible values for frequency and timing offsets, respectively, i.e., \( \nu_{1,1}^{[\text{rd}]} , \tau_{k,1}^{[\text{rd}]} \sim \mathcal{U}(-0.5, 0.5), \forall k, \ell \).

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 \( d \)th iteration for each individual target vector, \( \theta_{1,\ell} \), three distinct and different target vectors denoted by \( \theta_{1,\ell}^{[\text{rd}]} \), \( \theta_{2,\ell}^{[\text{rd}]} \), and \( \theta_{3,\ell}^{[\text{rd}]} \), are randomly selected from the population. Subsequently, a mutant vector corresponding to the \( \ell \)th target vector and \( d \)th iteration, \( \varpi_{\ell,d} = [\varpi_{1,\ell,d} , \ldots , \varpi_{2K,\ell,d} ]^T \), is created via

\[
\varpi_{\ell,d} = \theta_{1,\ell} + ( \theta_{2,\ell} - \theta_{3,\ell} ) \cdot D_F,
\]

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, \( \theta_{1,\ell} \), and the mutant vector, \( \varpi_{\ell,d} \), to create a trial vector, \( \vartheta_{\ell,d} \), defined in (17) and (19).

\[
\vartheta_{\ell,d} = \theta_{1,\ell} + ( \theta_{1,\ell} - \varpi_{\ell,d} ) \cdot D_C, \quad q = 1, \ldots, 2K
\]

**Step 4 Selection: If compared to the \( \ell \)th target vector, \( \theta_{1,\ell} \), the \( \ell \)th trial vector, \( \vartheta_{\ell,d} \), results in a smaller objective function, \( \chi \), in (16), it replaces the target vector in the next iteration, i.e.,

\[
\theta_{\ell,d+1} = \begin{cases} 
\theta_{\ell,d}, & \chi (\theta_{\ell,d}) \leq \chi (\theta_{\ell,d}) \\
\vartheta_{\ell,d}, & \text{otherwise}
\end{cases}
\]

By executing Steps 1-5 above, the estimates of timing and frequency offsets, \( \hat{\tau}^{[\text{rd}]} \) and \( \hat{\nu}^{[\text{rd}]} \), 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 \( D_C \), and the stopping criteria \( \varsigma \). As a matter of fact, each unique optimization problem requires the DE parameters to be initialized differently [36], where finding the appropriate values for these parameters can be a difficult and a non-trivial task [47]. However, by applying the results in Section III and by following the general guidelines in [36], [47], we are able to determine an appropriate parameterization for the DE algorithm.

\(^2\)Step sizes of \( 10^{-2} \) and \( 10^{-4} \) for MTOs and MCFOs, respectively.

\(^3\)Also called "generation" in DE terminology.
Note that $D_F$ 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 $D_F = 20K$. $D_F$ 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, $\lambda$, is mutually coupled. As a result, $D_R$ is set to $D_R = 0.9$ [36]. Finally, in order to reach the CRLB, the stopping criteria is set to $\zeta = 80$.

In DE terminology, the above procedure can be classified as “DE-rand”, since to obtain the mutant vector, $\sigma_{i,d}$, the target vector, $\theta_{1,d}$ 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, $\gamma$, is selected as $\theta_{1,d}$ in (17). However, in our problem formulation, “DE-best” does not always converge to the true values of timing and frequency offsets. Similar findings are also observed in [36], where it is shown that “DE-rand” is more effective at reaching the global solution than “DE-best” for different optimization problems [36, p. 154].

Remark 2: Even though the general conditions for global convergence of evolutionary algorithms are established in [48], it cannot be analytically shown that DE meets these conditions [49] and converges to the global solution of (16). Moreover, in [48], no specific algorithm that meets these conditions is proposed. Although in [49], 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.

C. Statistical Properties of Estimates

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, in this subsection we numerically show that the proposed estimators are unbiased and their second order statistics can be accurately approximated by the CRLB derived in Section III.

Figs. 4(a) and 4(b) show the histogram of the relay-to-destination timing and frequency offset estimation errors, i.e., $\delta_{p[t]} = \tau_{[t]} - \hat{\tau}_{[t]}$ and $\delta_{v[t]} = \nu_{[t]} - \hat{\nu}_{[t]}$, for 3600 realizations, $K = 2$ relays, $\text{SNR} = 20 \text{\,dB}$, and $\tau_{[r,d]}, \nu_{[r,d]} \sim \mathcal{U}(-0.5, -0.5)$. It can be observed that the means of timing and frequency offset estimation errors, $\delta_{p[t]}$ and $\delta_{v[t]}$, respectively, are very close to zero. As a matter of fact, the average values of timing and frequency offset estimates are observed to be $\mathbb{E} \{\delta_{p[t]}\} = [2.25 \times 10^{-5}, -8.85 \times 10^{-5}]^T$ and $\mathbb{E} \{\delta_{v[t]}\} = [2.49 \times 10^{-5}, 1.09 \times 10^{-6}]^T$. Thus, it can be empirically concluded that the proposed DE and LS algorithms for estimation of MTOs and MCFOs are asymptotically unbiased estimators. Subsequently, it straightforwardly follows from (15) that the estimates of channel gains $\alpha$ are also unbiased. More importantly, the numerical results in Figs. 4(a) and 4(b) demonstrate that the estimation errors for timing and frequency offsets can be modeled as a multi-variate Gaussian distribution, i.e., $\delta_{p[t]} \sim \mathcal{N}(0_{K \times 1}, \mathbf{C}_{\tau[t], \tau[t]})$, $\delta_{v[t]} \sim \mathcal{N}(0_{K \times 1}, \mathbf{C}_{\nu[t], \nu[t]})$, where $\mathbf{C}_{\tau[t], \tau[t]}$ and $\mathbf{C}_{\nu[t], \nu[t]}$, are given in (13)$^5$. More specifically, the average MSEs between the numerically evaluated covariance matrices of $\delta_{p[t]}$ and $\delta_{v[t]}$, and the analytical covariance matrices, $\mathbf{C}_{\tau[t], \tau[t]}$ and $\mathbf{C}_{\nu[t], \nu[t]}$, are $1.264 \times 10^{-10}$ and $7.2 \times 10^{-18}$, respectively.

D. 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 [7]. In addition, comparing two values or generating a random number is considered as a single arithmetic operation.

The computational complexity of the LS algorithm, denoted by $C_{LS}$ is calculated as given by (20) at the bottom of this page, where $\kappa_{\tau}$ and $\kappa_{\nu}$ 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 $\hat{\alpha}$ via (15). The computational complexity of the proposed DE algorithm, denoted by $C_{DE}$ is calculated as

$$C_{DE} = D_I \left\{ 6D_{\rho}K + 4D_{\rho}K + D_{\rho}(C_\chi + 1) \right\} \frac{1}{\kappa_{\tau} \kappa_{\nu}} + D_{\rho}(C_\chi + 1) + C_\alpha, \quad (21)$$

where $D_I$ is the total number of iterations required for the DE to converge and $C_\chi = 2K^3 + (4LN - 1)K^2 + (2LN)^2 -$ $^5$ $^5$ Similar results are obtained for the channel gains, $\alpha$, i.e., the distribution of $\delta_\alpha = \alpha - \hat{\alpha}$ is given by $\delta_\alpha \sim \mathcal{CN}(0_{K \times 1}, \mathbf{C}\alpha, \alpha)$ but are not included here due to lack of space.

$$C_{LS} = \left( \frac{1}{\kappa_{\tau} \kappa_{\nu}} \right)^K \left[ 2K^3 + (4LN - 1)K^2 + (2LN)^2 - LN \right] + C_\alpha, \quad (20)$$
received signal, respectively, for a wide range of SNRs. The MMSE compensation matrix, is given by minimizing the cost function compared to the proposed LS algorithm, for and by setting by evaluating the LS estimates using exhaustive search in (16) relays, respectively. Note that this comparison is carried out LN 5 observing that impairment estimation using the DE is lower complexity of the proposed DE based estimator.

V. COMPENSATION ALGORITHMS AT THE DESTINATION

The received signal at the destination, , 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 matrix, , as shown in Fig. 3, which is applied to the received signal, , to remove the effect of MTOs, MCFOs, and multiple channels. The MMSE compensation matrix, , is given by minimizing the cost function below

where is defined below (2), and the expectation in (22) is taken with respect to the statistics of the estimation errors , , , , and , and by setting . Furthermore, to meet the power constraints at the destination, it is assumed that

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, , , , , and , can be modeled by , for , and , for , respectively, [50]. Note that , , and , are given in (13), and and can be determined from the results in [51]. 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, , is given by

where

- is an matrix, 
- is an matrix,
- is an matrix,
- is an matrix,
be concluded that due to the presence of channel, timing, and frequency offsets from source to relay and relay to destination, i.e., knowledge of channels, timing, and frequency offsets from source to relay and relay to destination can derive this optimal MMSE matrix by assuming perfect MMSE compensation matrix at the destination. However, one can still justify the optimality of this compensation matrix, and are known at the destination, jointly estimated via the estimators proposed in Section IV.

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

$$Q^{[\text{MMSE}]} = \Phi_{x_{\text{DTP}},s}^H \Omega^H \hat{A}^H \Xi^H \left( \Xi R^{[\text{MMSE}]} \Xi^H + \Sigma_w \right)^{-1},$$

where $R$, $\Phi_{x_{\text{DTP}},s}$, $\hat{A}$, and $\Xi$ are defined below (23), and $\Omega$ and $\Sigma_w$ are defined below (8) and (9), respectively.

**Proof:** Eq. (24) follows from Theorem 1 by taking the derivative of $J_{\Phi}$ in (23) with respect to $Q$, setting the results to zero, and carrying out straightforward algebraic manipulations.

**Remark 4:** Evaluation of the MMSE compensation matrix, $Q^{[\text{MMSE}]}$, requires knowledge of the overall channel gains, $\hat{\alpha}$, to relay destination timing and frequency offsets, $\tilde{\nu}[\tilde{\beta}]$, and $\hat{\nu}[\tilde{\beta}]$, respectively, to destination channel gains $\beta$, and source to relay timing offsets, $\tau[\tilde{\beta}]$, at the destination. Even though, $\alpha$, $\tau[\tilde{\beta}]$, and $\tilde{\nu}[\tilde{\beta}]$ can be jointly estimated via the estimators proposed in Section IV and are known at the destination, $\beta$ and $\tau[\tilde{\beta}]$ 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, $\tilde{\nu}[\tilde{\beta}]$ and $\hat{\tau}[\tilde{\beta}]$, respectively, to the destination. Note that using $h_k$, an estimate of relay to destination channels, $\beta_k$, can be obtained according to $\beta_k = \frac{\tau_k}{\hat{\tau}_k}$.

**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, $Q^{[\text{MMSE}]}$ 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., $\delta_{\alpha} = \delta_{\tilde{\nu}[\tilde{\beta}]} = \delta_{\nu[\tilde{\beta}]} = 0_{K \times 1}$.

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

$$Q^{[\text{OPT}]} = \Phi_{x_{\text{DTP}},s}^H \Omega^H \hat{A}^H \Xi^H \left( \Xi R^{[\text{OPT}]} \Xi^H + \Sigma_w \right)^{-1},$$

where $R^{[\text{OPT}]} = \Delta \Omega \Phi_{x_{\text{DTP}},s}^H \Omega^H \hat{A}^H + \Delta \Omega \Sigma_w \Omega^H B H$.

**Proof:** By assuming $\delta_{\alpha} = \delta_{\tilde{\nu}[\tilde{\beta}]} = \delta_{\nu[\tilde{\beta}]} = 0_{K \times 1}$ 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, $Q^{[\text{OPT}]}$ 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 $Q^{[\text{L-MMSE}]}$ in (24) is a function of the source to relay channels and timing offsets since the matrix $R$ defined below (23) is a function of $B$ and $\Sigma_n$, 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, $B$ in (24) can be replaced by $B^{[\text{L-MMSE}]} \approx \text{diag} \left( \sigma_{f[1]}, \ldots, \sigma_{f[K]} \right) \otimes \hat{L}$. Moreover, the dependence on timing offset estimates from source to relay can be eliminated by approximating the block diagonal matrix $\Sigma_n$, defined below (23), with the diagonal matrix $\Sigma_n^{[\text{L-MMSE}]}$ with its $k$th diagonal block given by $[\Sigma_n^{[\text{L-MMSE}]})_{(k-1)L:kL,(k-1)L:kL}] = \sigma_{u_k} I_{L}$ at high SNR. The latter follows from the definition of $\Sigma_n$ and extensive simulations that show that the off-diagonal elements of $\Sigma_n$ 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, $Q^{[\text{L-MMSE}]}$, for mitigating the effect of impairments can be found as

$$Q^{[\text{L-MMSE}]} = \Phi_{x_{\text{DTP}},s}^H \Omega^H \hat{A}^H \Xi^H \left( \Xi R^{[\text{L-MMSE}]} \Xi^H + \Sigma_w \right)^{-1},$$

where $R^{[\text{L-MMSE}]} = \hat{A} \Omega \Phi_{x_{\text{DTP}},s}^H \Omega^H \hat{A}^H + (C_{\alpha,\alpha} \otimes \hat{L}) \otimes (\Phi_{x_{\text{DTP}},s}^H \Omega^H B^{[\text{L-MMSE}]} \otimes \Sigma_n^{[\text{L-MMSE}]} \Omega^H B^{[\text{L-MMSE}]} H)$. 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, $t[\tilde{\beta}]$ and $t_k[\tilde{\beta}]$, $\forall k$, are set to $L = 80$ symbols during the TP and length of the source data vector $s$, 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 generating DSTBC in AF cooperative networks [8], [9]. The propagation loss is modeled as $\eta = (\varepsilon/\varepsilon_0)^{-\rho}$, where $\varepsilon$ is the distance between transmitter and receiver, $\varepsilon_0$ is the reference distance, and $\rho$ is the path loss exponent [43]. We set $\varepsilon_0 = 1$ km, and $\rho = 2.7$, which corresponds to an urban area cellular network [43]. The timing and frequency offsets at relays and destination, $\tau^{[\text{sr}]}$, $\tau^{[\text{rd}]}$, $\nu^{[\text{sr}]}$, and $\nu^{[\text{rd}]}$ 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., $\sigma^2_u = \sigma^2_v = \ldots = \sigma^2_K$ and $\sigma^2_w = 1/\text{SNR}$. 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.

A. MSE performance

Following the approach in [7], [33], [37], [41] and for ease of reproducing the CRLBs and MSE curves in Fig. 5 and Fig. 6, specific channels are used. For $K = 2$, $h = [.279 - .9603 j, .8837 + .4681 j]^T$ and $f = [.7820 + .6233 j, .9474 - .3203 j]^T$ and for $K = 4$, $h = [.279 - .9603 j, .8837 + .4681 j, -.343 + .7321 i, -.734 - .4511 i]^T$ and $f = [.7820 + .6233 j, .9474 - .3203 j, -.2413 + .7241 i, .5141 - .8933 i]^T$. 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-D and to reach the CRLB, the step sizes, $\kappa_T$ and $\kappa_p$, 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. 5(a) shows the MSE of frequency offset estimation 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. 5(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. 5(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. 5(a), this fast convergence comes at the cost of MSE performance. The results in Figs. 5(a) and 5(b) demonstrate that the proposed DE algorithm provides an effective trade-off between synchronization performance and complexity that can be 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. 6(a), 6(b), and 6(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. 6 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 (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 (SNR $\geq$ 5 dB), Fig. 6 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 $CN(0, 1)$.

Fig. 7 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 $\hat{s} = (\Upsilon^H \Upsilon)^{-1} \Upsilon^H \nu^{[DPT]}$, where $\Upsilon \triangleq \sum_{k=1}^{K} \beta_k \xi_k \Omega_k$. For the MMSE, L-MMSE, and ZF methods, the compensation matrices, $Q^{[\text{MMSE}]}$ in (24), $Q^{[\text{L-MMSE}]}$ in (26), and $\Upsilon$ defined above are evaluated using the estimated values, $\alpha_k$, $\tau_k^{[\text{rd}]}$, and $\nu_k^{[\text{rd}]}$, obtained in the training period using the DE algorithm\footnote{Estimation using LS estimator results in similar BER performance as obtained using the DE algorithm.} while the optimal compensation matrix $Q^{[\text{OPT}]}$ in (25) is evaluated based on perfect knowledge of channels and synchronization parameters. Fig. 7 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. 7 shows that the BER of a DSTBC-AF cooperative system based on estimated impairments is close to 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. 7 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. 8 depicts the BER results for a cooperative system that first employs the re-synchronization filter in [14] to compensate MTOs and then attempts to remove MCFOs by employing the algorithm in [33]. This plot, which is labeled as “[14] & [33]”, shows that such an approach fails to equalize the effect of impairments at the destination since the re-synchronization filter in [14] fails to compensate MTOs in the presence of MCFOs. Subsequently, the algorithm in [33] fails to nullify MCFOs, since the input signal is corrupted by MTOs. In addition, even though Fig. 8 illustrates the BER performance of a cooperative system applying the transceiver design in [7] for 2 and 4 relays, it is important to consider that the proposed DSTBC-AF transceiver design in this paper and the scheme in [7] vastly differ. In [7], a less complex AF 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. 8, 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. 8 shows that at low SNR, the proposed algorithm exhibits a higher BER performance compared to that of [7]. 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 [7] 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 [7] 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 [7], e.g., at SNR $= 30$ dB, the BER of the
The proposed algorithm is $4 \times 10^{-6}$ for $K = 4$ relays while the BER of the algorithm in [7] is $5 \times 10^{-2}$. 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 [7] provides little to no spatial diversity gain as the number of relays increases.

VII. CONCLUSIONS

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 unbiased and is $5.6 \times 10^7$ or $1.2 \times 10^{19}$ 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 (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 [50]

$$[F]_{m, \tilde{m}} = 2\Re\left\{ \frac{\partial \mu_{y|m+1}^H}{\partial \lambda_m} \sum_{l=1}^{K} \frac{\partial \mu_{y|m+1}}{\partial \lambda_{\tilde{m}}} \right\} + \text{Tr}\left\{ \sum_{l=1}^{K} \frac{\partial \Sigma_{y|m+1}}{\partial \lambda_m} \sum_{l=1}^{K} \frac{\partial \Sigma_{y|m+1}}{\partial \lambda_{\tilde{m}}} \right\},$$

$m, \tilde{m} = 1, \ldots, 4K.$ \hspace{1cm} (A.1)

The mean of the received signal in (6) during TP, $\mu_{y|m+1}$, is calculated as

\begin{align*}
\mu_{y|m+1} &= E\{ E_A p + E_B v + w \} \\
&= E\{ E_A (x^{TP} \circ t^{[rd]}) + E_B (n \circ t^{[rd]}) + w \} \hspace{1cm} (A.2a) \\
&= E_A E\{ x^{TP} \circ t^{[rd]} \} + E_B E\{ n \circ t^{[rd]} \} \\
&= E_A t, \hspace{1cm} (A.2b)
\end{align*}

where $x^{TP} = \left[ (x_1^{TP})^T, \ldots, (x_K^{TP})^T \right]^T$ and $x_k^{TP}$ is defined below (2). (A.2b) follows from (A.2a) because $\alpha, \beta, \nu^{[rd]}$ and $\tau^{[rd]}$ 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 $k$th relay, $\delta_{\nu}[w]$ and $\delta_{\nu}[w]$, respectively, are modeled as zero-mean Gaussian random variables. Consequently, $\Delta_{\nu}[\delta_{\nu}[w]]\Lambda^{\nu}[\delta_{\nu}[w]] = \delta_{\nu}[w] = 0 = \mathbf{I}_L$, $\mathbb{E}\{w\} = 0_{L \times 1}$, $\mathbb{E}\{n_k\} = G_{\nu}[\delta_{\nu}[w]]\mathbb{E}\{u_k\} = 0_{L \times 1}$, and $\mathbb{E}\{\hat{x}_k[TP]\} = \mathbb{E}\{\Delta_{\nu}[\delta_{\nu}[w]]\Lambda^{\nu}[\delta_{\nu}[w]]\} = t^{[\nu]}$. Finally $t$ in (A.2c) is defined as $t \triangleq \begin{bmatrix} t^{[\nu]} \end{bmatrix}^T$. The $LN \times LN$ received signal's covariance matrix during TP, $\Sigma_{\nu}[\nu]$, can be determined as

$$\Sigma_{\nu}[\nu] = \mathbb{E}(p - (t)H)^\dagger A^{H}\Xi$$

$$+ \mathbb{E}(m^{H}\circ t^{[\nu]})B^{H}\Xi$$

$$= \Xi A_{\nu}A^{H}\Xi + \mathbb{E}\Sigma_{\nu}B^{H}\Xi + \Sigma_{\nu},$$

where $\Sigma_{\nu} \triangleq \mathbb{E}\{(p - (t)^H\}^{H}$ and $\Sigma_{\nu} \triangleq \mathbb{E}\{m^{H}\} \circ t^{[\nu]}$ are $LN \times LN$ and $LN \times LN$ matrices, and $\Sigma_{\nu} \triangleq \mathbb{E}\{mm^{H}\}$. The derivatives of $\Sigma_{\nu}[\nu]$ in (A.1) can be calculated as

$$\frac{\partial \Sigma_{\nu}[\nu]}{\partial \mathbb{R}\{\nu\}} \triangleq \Xi A_{\nu}A^{H}\Xi +$$

$$+ \Xi A_{\nu}B^{H}\Xi,$$

where $\Xi \triangleq \mathbb{E}\{mm^{H}\}$ and $\mathbb{R}\{\nu\} \triangleq \{0, 1, 2, \ldots, LN - 1\}$. Moreover, the derivative of $\Sigma_{\nu}[\nu]$ in (A.1) with respect to the parameters of interest is given by

$$\Sigma_{\nu}'[\nu], k \triangleq \frac{\partial \Sigma_{\nu}[\nu]}{\partial \mathbb{R}\{\nu\}} \triangleq \Xi U_{\alpha_k}A^{H}\Xi$$

$$+ \Xi A_{\nu}A^{H}\Xi,$$

where $U_{\alpha_k} = \text{diag}(L_{L \times (k-1)}, I_L, L_{L \times (L-k)})$ is an $LN \times LN$ matrix, and $U_{\nu}[\nu] = \{0_{L \times L_{(k-1)}}, jDB_{\nu}[\nu], 0_{L \times (L-k)}\}$ are $LN \times LN$ 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

$$F_{k, \bar{k}} = 2R\{t^{[\nu]} B^{H}\Xi^{\nu} A_{k}^{H} X^{\nu} \} + \Pi_{k, \bar{k}},$$

and

$$F_{k, \bar{k} + K} = -23\{t^{[\nu]} B^{H}\Xi^{\nu} A_{k}^{H} X^{\nu} \} + \Pi_{k, \bar{k} + K},$$

$$F_{k + K, \bar{k} + K} = 2R\{t^{[\nu]} B^{H}\Xi^{\nu} A_{k}^{H} X^{\nu} \} + \Pi_{k + K, \bar{k} + K},$$

$$F_{k + 2K, \bar{k} + 2K} = 2R\{t^{[\nu]} B^{H}\Xi^{\nu} A_{k}^{H} X^{\nu} \} + \Pi_{k + 2K, \bar{k} + 2K},$$

$$F_{k + 3K, \bar{k} + 3K} = 2R\{t^{[\nu]} B^{H}\Xi^{\nu} A_{k}^{H} X^{\nu} \} + \Pi_{k + 3K, \bar{k} + 3K},$$

where $\Sigma_{\nu}[\nu]$ are the elements of $4K \times 4K$ matrix, $\Pi$, such that $\Pi_{m, \bar{m}} = \text{Tr}\{\Sigma_{\nu}[\nu] \Sigma_{\nu}[\nu]^{H}\}$. Note that the lower triangular elements of $F$ can be easily obtained by simple manipulation of (A.6a)-(A.6j) and are not included here for brevity. Using (A.6a)-(A.6j), $F$ can be determined as given in (11).

**APPENDIX B**

**DERIVATION OF $J_n$ IN (23)**

For notational brevity, let us denote $\nu[\text{DPTP}]$ by $x$ throughout this section. Substituting (8) in (22), we get

$$J_n = \mathbb{E}(\delta_{\nu}[\nu][\nu], \delta_{\nu}[\nu][\nu], \delta_{\nu}[\nu][\nu], \delta_{\nu}[\nu][\nu], \nu[\text{DPTP}], w, s)$$

$$\{\|\mathbb{E}(A\nu[\text{DPTP}]x + \Xi B\nu[\text{DPTP}]n + w) - s\|^{2}\}. \quad (B.1)$$
Since, s, w and u are uncorrelated, (B.1) can be expanded as

\[
\mathcal{J}_n = \text{Tr} \left\{ Q \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \Xi \mathbb{E}_{\delta_{n}} \left\{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H A^H \right\} \Xi^H \right\} Q^H \right. \\
+ Q \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \Xi \mathbb{E}_{\delta_{n}} \left\{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H \tilde{B}^H \Xi^H \right\} Q^H \right. \\
- Q \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \mathbb{E}_{\delta_{n}} \left\{ A \Omega \tilde{\Phi}_x \right\} \right\} Q^H \\
- \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \mathbb{E}_{\delta_{n}} \left\{ \tilde{\Phi}_x \tilde{\Phi}_x^H A^H \right\} \Xi^H \right\} Q^H \\
+ \tilde{\Phi}_x + Q \Sigma_n w \right\}, \tag{B.2}
\]

where \( \tilde{\Phi}_x \triangleq \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \{ xx^H \}, \quad \Sigma_n \triangleq \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}, u} \{ nn^H \}, \quad \tilde{\Phi}_{x,s} \triangleq \mathbb{E}_{\delta_{s[n]} \delta_{s[n], s}} \{ xs^H \}, \quad \Phi_{x,s} \triangleq \mathbb{E}_{\delta_{s[n]} \delta_{s[n], s}} \{ ss^H \} \) and \( \Sigma_w \triangleq \mathbb{E}_w \{ ww^H \} \). These expectations are calculated as follows.

1) Calculation of \( \mathbb{E}_{\delta_{s[n]}} \{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H A^H \} \) and \( \mathbb{E}_{\delta_{n}} \{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H \} \): Since \( \delta_{s[n]} = \alpha - \tilde{\alpha} \), the channel matrix \( A \), given below (5), can be written as \( A = \hat{A} + \Delta_A \), where \( \hat{A} \triangleq \text{diag} (\tilde{\alpha}) \otimes I_L \) and \( \Delta_A \triangleq \text{diag} (\delta_{s[n]}) \otimes I_L \) are \( L \times L \) matrices. According to the assumption in Section V, \( \delta_{s[n]} \) is distributed as \( \delta_{\alpha} \sim \mathcal{N} (0_{K \times 1}, \Sigma_{\alpha, \alpha}) \). Thus, \( \mathbb{E}_{\delta_{s[n]}} \{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H A^H \} \) and \( \mathbb{E}_{\delta_{n}} \{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H \} \) are evaluated as

\[
\mathbb{E}_{\delta_{s[n]}} \{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H A^H \} = \hat{A} \Omega \tilde{\Phi}_x \tilde{\Omega}^H \hat{A}^H + (C_{\alpha, \alpha} \otimes I_L) \odot (\tilde{\Psi}_x \tilde{\Omega}^H), \tag{B.3a}
\]

\[
\mathbb{E}_{\delta_{n}} \{ A \Omega \tilde{\Phi}_x \tilde{\Omega}^H \} = \hat{A} \Omega \tilde{\Phi}_x \tilde{\Omega}^H. \tag{B.3b}
\]

2) Calculation of \( \tilde{\Phi}_s \) & \( \tilde{\Phi}_x \): Let us assume that the transmitted data symbols are uncorrelated such that \( \tilde{\Phi}_s \triangleq \mathbb{E}_s \{ ss^H \} = I_L \). Furthermore, \( \tilde{\Phi}_x \) is an \( L \times L \) correlation matrix, whose \( L \times L \) blocks are given by

\[
[\tilde{\Phi}_x](k-1)L+1:kL,(k-1)L+1:kL = \Delta^{[0]}_{A_k} \Delta^{[0]}_{G_k} \tilde{\Phi}_s \left( \Delta^{[0]}_{G_k} \right)^H \left( \Delta^{[0]}_{A_k} \right)^H + \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \Delta^{[0]}_{A_k} \delta^{[0]}_{G_k} \tilde{\Phi}_s \left( \Delta^{[0]}_{G_k} \right)^H \right\} + \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \delta^{[0]}_{G_k} \tilde{\Phi}_s \left( \Delta^{[0]}_{G_k} \right)^H \right\} + \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \tilde{\Phi}_s \right\}, \tag{B.6a}
\]

\[
[\Sigma_n](k-1)L+1:kL,(k-1)L+1:kL = \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}, u} \left\{ \tilde{G}_k^{[s]} \Delta^{[s]}_{A_k} \tilde{u}_k \tilde{u}_k^H \left( \Delta^{[s]}_{A_k} \right)^H \left( \tilde{G}_k^{[s]} \right)^H \right\} + \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}, u} \left\{ \tilde{G}_k^{[s]} \Sigma_{u_k} \left( \tilde{G}_k^{[s]} \right)^H \right\}, \tag{B.8a}
\]

Based on the assumption in Section V, \( \delta_{\tau[n]} \sim \mathcal{N} (0_{K \times 1}, \Sigma_{\alpha, \alpha}) \). Thus, by applying Taylor series expansion around their means and by ignoring the second and higher order terms, the non-linear terms \( \Delta A[k] \) and \( \Delta G[k] \), which are functions of \( \delta_{\tau[k]} \) and \( \delta_{\tau[k]} \), respectively, can be written as

\[
\Delta A[k] = \Delta_0 A[k] + \Delta_0 A[k]^{'} \delta_{\tau[k]} + \mathcal{O}_p \left( \sigma_{\tau[k]}^2 \right), \tag{B.5a}
\]

\[
\Delta G[k] = \Delta_0 G[k] + \Delta_0 G[k]^{'} \delta_{\tau[k]} + \mathcal{O}_p \left( \sigma_{\tau[k]}^2 \right), \tag{B.5b}
\]

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 \( 0_{L \times L} \) according to the distributions of \( \delta_{\tau[n]} \) and \( \delta_{\tau[n]} \).

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

\[
[\tilde{\Phi}_x](k-1)L+1:kL,(k-1)L+1:kL = \mathbb{E}_{\delta_{s[n]} \delta_{s[n]}} \left\{ \Delta^{[s]}_{A_k} \Delta^{[s]}_{G_k} \tilde{\Phi}_s \right\} \approx \Delta^{[s]}_{A_k} \Delta^{[s]}_{G_k} \tilde{\Phi}_s = \Phi_s. \tag{B.7}
\]

4) Calculation of \( \Sigma_n \): The \( L \times L \) blocks of the \( L \times L \) matrix \( \Sigma_n \) can be calculated as shown in (B.8) at the bottom of this page, where \( \Sigma_{u_k} \triangleq \sigma_{\tau[k]}^2 I_{L \times L} \) and (B.8b) follows from
(B.8a) since the noise at different relays, $u_k$ and $u_{k'}, \forall k \neq k'$, are uncorrelated. Note that $\hat{G}^{[sr]}_k$ is a non-linear function of $\hat{\tau}^{[sr]}_k \pm \delta_{\hat{\tau}^{[sr]}_k} + \delta_{\tau^{[sr]}_k}$. Thus, in order to evaluate (B.8b), the Taylor series expansion of the non-linear term $\hat{G}^{[sr]}_k$ around $\tau^{[sr]}_k$ is evaluated as

$$\hat{G}^{[sr]}_k = G^{[sr]}_k + \left(G^{[sr]}_k\right)' \delta_{\tau^{[sr]}_k} + O_p\left(\delta_{\tau^{[sr]}_k}^2 \right),$$

(B.9)

where $G^{[sr]}_k \triangleq \hat{G}^{[sr]}_k - \langle G^{[sr]}_k \rangle^{[sr]}$, $\langle G^{[sr]}_k \rangle^{[sr]} = \frac{\partial \hat{G}^{[sr]}_k}{\partial \tau^{[sr]}_k} |_{\tau^{[sr]}_k = \langle \tau^{[sr]}_k \rangle}$, and since the noise at the relays are uncorrelated, $\Sigma_n$ is an $L \times L$ block diagonal matrix. Using (B.9) in (B.8), the $L \times L$ diagonal matrices of $\Sigma_n$ are given by

$$\Sigma_n = \left[ \Sigma_n \right]_{(k-1)L+1:kL} \left[ \Sigma_n \right]_{(k-1)L+1:kL} + \varepsilon_{\delta^{[sr]}_n} \left\{ \left( A \Omega \Phi_x \Omega H A^H \right) \Xi^H \right\}$$

and

$$\varepsilon_{\hat{G}^{[sr]}_k} \left\{ \left( \Xi B \Omega \Sigma_n \Omega B^H \Xi^H \right) \right\};$$

where using similar steps as in (B.5), (B.6), (B.9), and (B.10), the expectations over $\delta_{\hat{\tau}^{[sr]}_k}$ and $\delta_{\tau^{[sr]}_k}$ in (B.2) can be approximated by

$$\mathbb{E}_{\delta^{[sr]}_n} \left\{ \sum_{k=1}^{L} \left( A \Omega \Phi_x \Omega H A^H \right) \Xi^H \right\} = \mathbb{E}_{\delta^{[sr]}_n} \left\{ \sum_{k=1}^{L} \left( A \Omega \Phi_x \Omega H A^H \right) \Xi^H \right\} + O_p\left(\delta_{\tau^{[sr]}_k}^2 \right),$$

(B.11a)

$$\varepsilon_{\hat{G}^{[sr]}_k} \left\{ \left( \Xi B \Omega \Sigma_n \Omega B^H \Xi^H \right) \right\} = \mathbb{E}_{\delta^{[sr]}_n} \left\{ \sum_{k=1}^{L} \left( A \Omega \Phi_x \Omega H A^H \right) \Xi^H \right\} + O_p\left(\delta_{\tau^{[sr]}_k}^2 \right),$$

(B.11b)

where

$$\Xi = \mathbb{E}_{\left[ \delta^{[sr]}_n \right]} \left\{ \left( A \Omega \Phi_x \Omega H A^H \right) \Xi^H \right\}$$

and

$$\mathbb{E}_{\delta^{[sr]}_n} \left\{ \left( A \Omega \Phi_x \Omega H A^H \right) \Xi^H \right\}$$

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


Ali A. Nasir

Ali A. Nasir received his B.Sc. (1st class hons.) degree in Electrical Engineering from the University of Engineering and Technology (UET), Lahore, Pakistan in 2007. He is currently pursuing his PhD degree at the Research School of Engineering, the Australian National University (ANU), Canberra, Australia. Prior to joining ANU, he worked as a Lecturer in UET for nine months and then as a Design Engineer in the Centre for Advanced Research in Engineering (CARE), Islamabad, Pakistan for a year. He was a research visitor in Chalmers University of Technology, Gothenburg, Sweden for six months from April till October of 2011.

Ali was awarded a University Gold Medal for outstanding performance during the final year of his undergraduate studies. He is a recipient of an ANU International Ph.D. scholarship for the duration of his Ph.D. He was also awarded an ANU Vice Chancellor’s Higher Degree Research (HDR) travel grant in 2011. His research interests are in the area of synchronization and channel estimation in cooperative communication, MIMO and OFDM systems.

Hani Mehrpouyan

Hani Mehrpouyan (S’05-M’10) received his B.Sc. honours degree in Computer Engineering from Simon Fraser University, Burnaby, Canada in 2004. In 2005 he joined the Department of ECE at Queens University, Kingston, Canada as a Ph.D. candidate, where he received his Ph.D. in Electrical Engineering in 2010. From September 2010 to March of 2012 he was a Post-Doc at the Department of Signal and Systems at Chalmers University of Technology where he lead the MIMO aspects of the microwave backhauling for next generation wireless networks project. He next joined the University of Luxembourg as a Research Associate from April 2012 to August of 2012, where he was responsible for new interference cancelation and synchronization schemes for next generation satellite communication links. Since August of 2012 he has been an Assistant Professor at the Department of Computer and Electrical Engineering and Computer Science at California State University, Bakersfield.

Hani has received more than 10 scholarships and awards, e.g., IEEE Globecom Early Bird Student Award, NSERC-IRDF, NSERC PGS-D, NSERC CGS-M Alexander Graham Bell, B.C. Wireless Innovation, and more. He has more than 20 publications in prestigious IEEE Journals and Conferences. He has served as a TPC member for IEEE Globecom 2012. Hani has also been involved with industry leaders such as Ericsson AB, Research in Motion (RIM), Alcatel etc. His current research interests lie in the area of applied signal processing and physical layer of wireless communication systems, including synchronization, channel estimation, interference cancelation, and performance optimization. For more information refer to www.mehrpouyan.info.
Salman Durrani (S’00-M’05-SM’10) received the B.Sc. (1st class honours) degree in Electrical Engineering from the University of Engineering & Technology, Lahore, Pakistan in 2000. He received the PhD degree in Electrical Engineering from the University of Queensland, Brisbane, Australia in Dec. 2004. Since March 2005, he has been a Lecturer in the Research School of Engineering, College of Engineering & Computer Science, The Australian National University, Canberra, Australia. His current research interests are in wireless communications and signal processing, including synchronization in cooperative communication systems, connectivity of ad-hoc networks and vehicular networks and signal processing on the unit sphere. He serves as a Technical Program Committee Member for international conferences such as ICC ’12, PIMRC ’11 and AusCTW ’12. He was a recipient of an International Postgraduate Research Scholarship from the Australian Commonwealth during 2001-2004. He was awarded a University Gold Medal during his undergraduate studies. He has 47 publications to date in refereed international journals and conferences. He is a Member of Institution of Engineers, Australia and a Senior Member of IEEE.

Steven D. Blostein (S’83-M’88-SM’96) received his 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 in the Department of Electrical and Computer Engineering Queen’s University since 1988 and currently holds the position of Professor. From 2004-2009 he was Department Head. From 1999-2003, he was the leader of the Multi-Rate Wireless Data Access Major Project sponsored by the Canadian Institute for Telecommunications Research. He has also been a consultant to industry and government in the areas of image compression, target tracking, radar imaging and wireless communications. He spent sabbatical leaves at Lockheed Martin Electronic Systems and at Communications Research Centre in Ottawa. His current interests lie in the application of signal processing to problems in wireless communications systems, including synchronization, network MIMO and physical layer optimization for multimedia transmission. He has been a member of the Samsung 4G Wireless Forum as well as an invited distinguished speaker. He served as Chair of IEEE Kingston Section (1994), Chair of the Biennial Symposium on Communications (2000,2006,2008), associate editor for IEEE Transactions on Image Processing (1996-2000), and Publications Chair for IEEE ICASSP 2004. For a number of years has been serving on technical program committees for IEEE Communications Society conferences that include ICC, Globecom and WCNC. He has been serving as an editor of IEEE Transactions on Wireless Communications since 2007. He is a registered professional engineer in Ontario and a Senior Member of IEEE.

Rodney A. Kennedy (S’86-M’88-SM’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, Canberra. He is currently a Professor in the Research School of Engineering, Australian National University. He is a Fellow of the IEEE. His research interests include digital signal processing, digital and wireless communications, and acoustical signal processing.