

Simultaneous Estimation of Multi-Relay MIMO Channels

Hani Mehrpouyan, *Member, IEEE*, Steven D. Blostein, *Senior Member, IEEE*, and Björn Ottersten, *Fellow, IEEE*

Abstract—This paper addresses training-based channel estimation in distributed *amplify-and-forward* (AF) multi-input multi-output (MIMO) multi-relay networks. To reduce channel estimation overhead and delay, a training algorithm that allows for simultaneous estimation of the entire MIMO cooperative network’s channel parameters at the destination node is proposed. The exact *Cramér-Rao lower bound* (CRLB) for the problem is presented in closed-form. Channel estimators that are capable of estimating the overall source-relay-destination channel parameters at the destination are also derived. Numerical results show that while reducing delay, the proposed channel estimators are close to the derived CRLB over a wide range of signal-to-noise ratio values and outperform existing channel estimation methods. Finally, extensive simulations demonstrate that the proposed training method and channel estimators can be effectively deployed in combination with cooperative optimization algorithms to significantly enhance the performance of AF relaying MIMO systems in terms of average-bit-error-rate.

I. INTRODUCTION

Cooperative communications has attracted considerable research interest due to its potential for multiplexing and diversity gain through resource sharing amongst nodes within the network. Pioneering contributions can be found in [1], [2] and results on *multi-input multi-output* (MIMO) broadcast and multiple-access channels have been reported in [3], [4]. Note that almost all of the proposed algorithms require knowledge of channel state information (CSI) to deliver the promised performance enhancements. Therefore, accurate and efficient channel estimators are key to future deployments of MIMO cooperative networks.

In [5], a novel and iterative expectation conditional maximization channel estimator for *single-relay* MIMO-orthogonal frequency division multiplexing (OFDM) two-way relaying cooperative networks is analyzed. However, the results in [5] are limited to the case of *single-relay* networks, the CRLB for estimation of channel parameters is not derived, the performance of the proposed estimator is not analyzed, and the estimator in [5] requires 3-4 iterations to estimate the overall network’s channel parameters. In [6], a new and innovative TS design approach for separate estimation of channels from source to relay and relay to destination in *single-relay* two-hop cooperative networks is proposed. Nevertheless, the proposed algorithm in [6] is based on the assumption that the overall CSI corresponding to source-relay-destination links is *known* at the destination. In [7], channel estimators and training sequence design guidelines are proposed for correlated *single-relay* MIMO cooperative networks. However, the approach in [7] is only capable of estimating

a *small portion* of the cooperative network’s CSI. As a result, the effect of the estimator in [7] on the performance of MIMO cooperative networks is not investigated. Finally, in [8] and [9] the authors devise new channel estimators for full-duplex multi-relay MIMO cooperative networks, i.e., where all nodes within the network can transmit and receive at the same time. However, the estimators in [8] and [9] are only applicable to the case of single antenna relays and require a total of $K + 1$ time slots, where K denotes the number of relays in the network, to estimate the network’s channels. Moreover, the full-duplex assumption in [8], [9] may not be realistic from a practical point of view.

Decode-and-forward relaying cooperative networks require channel estimation and signal detection at both the relays and destination terminals [10], where algorithms similar to that of point-to-point MIMO systems can be used to estimate the CSI [11]. However, in the case of fixed-gain *amplify-and-forward* (AF) cooperative networks, the relays do not need to perform channel estimation, since they do not decode the received signal. Therefore, the network’s overall CSI can be estimated at the destination node [10]. Similarly, in the AF approach here, relays are not required to estimate and forward their CSI to the destination which improves bandwidth efficiency, reduces power consumption, and avoids further distortion by transmission over the relay-to-destination link.

Under the considerations of half-duplex cooperative networks, this paper proposes a new training method for AF MIMO multi-relay cooperative networks that allows for simultaneous estimation of source-relay-destination channels at the destination node. The contributions of this paper can be summarized as follows:

- Unlike existing channel estimators that require the training to be carried out, relay by relay, we propose a novel training scheme that enables simultaneous estimation of the overall CSI from source to relays to destination. The proposed training scheme can significantly reduce the delay associated with the channel estimation process in multi-relay MIMO cooperative networks while enabling the application of synchronization algorithms in these networks [12]–[14].
- The estimation problem is parameterized and a new closed-form CRLB expression for channel estimation for AF multi-relay MIMO cooperative systems is derived to benchmark the performance of the proposed channel estimators.
- Low complexity *least squares* (LS) and weighted LS channel estimators are derived that applies the proposed training method to estimate the overall CSI from source to the k th relay to destination, for $k = 1, \dots, K$, simultaneously.
- Numerical results are presented showing that the performances of the proposed estimators are close to the CRLB over a wide range of SNR values. It is shown that in addition to

This research has been supported by NSERC Discovery Grant 41731 and SES. Hani Mehrpouyan and Björn Ottersten are with the Interdisciplinary Centre for Security, Reliability and Trust at the University of Luxembourg, Luxembourg and Steven D. Blostein is with the Department of Electrical and Computer Engineering, Queen’s University, Kingston, ON Canada, Emails: hani.mehr@ieee.org, steven.blostein@queensu.ca, and bjorn.ottersten@uni.lu.

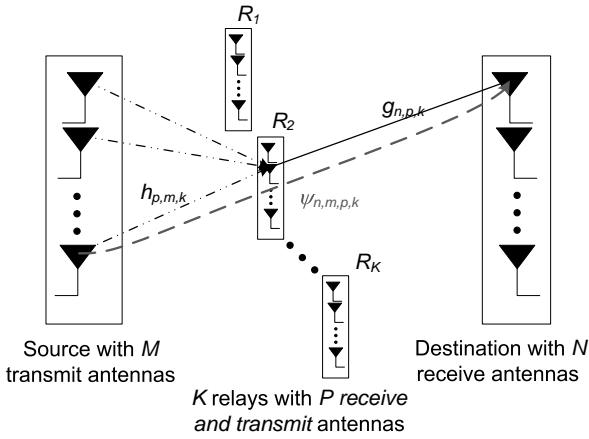


Fig. 1. System model for the multi-relay MIMO two-hop cooperative network.

enabling coherent detection, knowledge of the estimated CSI can be applied to optimize the *average bit-error rate (ABER)* performance of AF MIMO multi-relay cooperative networks.

The remainder of the paper is organized as follows: in Section II the signal model and the proposed training method is presented. Section III derives the new CRLB for the estimation problem. In Section IV, we derive the proposed LS and ML channel estimators. Section V, presents the simulation results while Section VI concludes the paper .

Notations: Italic letters (x) are scalars, bold lower case letters (\mathbf{x}) are vectors, bold upper case letters (\mathbf{X}) are matrices, and $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $\text{Tr}(\cdot)$, $\mathbb{E}(\cdot)$, denote conjugate, transpose, conjugate transpose (hermitian), trace, and expectation, respectively. $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ represent real and imaginary parts, respectively. $\mathbf{I}_{X \times X}$ and $\mathbf{0}_{X \times X}$ denote the $X \times X$ identity and all zero matrices, respectively. $\text{diag}(\mathbf{x})$ is used to denote a diagonal matrix with elements given by vector \mathbf{x} . The operator, \otimes denotes Kronecker product while \diamond represents the *Khatri-Rao* or column-wise Kronecker product. $\text{vec}(\cdot)$ denotes the vectorization operator. Finally, $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote a multivariate complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, respectively.

II. SYSTEM MODEL AND PROPOSED TRAINING METHOD

A half-duplex space-division multiple access cooperative network consisting of a source, destination pair and a cluster of K relay nodes is considered, where the relays are assumed to be distributed throughout the network as shown in Fig. 1. The source, relays, and destination are equipped with M , P , and N antennas, respectively. Throughout this paper indices $m = 1, \dots, M$, $p = 1, \dots, P$, $n = 1, \dots, N$, and $k = 1, \dots, K$, are used to denote source antennas, relay antennas, destination antennas, and number of relays, respectively. As depicted in Fig. 1, the channel from the m th source antenna to the p th antenna of the k th relay is denoted by $h_{p,m,k}$, the channel from the p th antenna of the k th relay to the n th destination antenna is denoted by $g_{n,p,k}$, and the overall channels from source to relays to destination are denoted by $\psi_{n,m,p,k} \triangleq h_{p,m,k}g_{n,p,k}$, $\forall n, p, m, k$. As shown in Fig. 2, transmission is divided temporally into two intervals: *training* and *data transmission*. During the *training interval*, a TS of L symbols is used to estimate the cooperative network's CSI, $\psi_{n,m,p,k}$, $\forall n, p, m, k$, at the destination. In the *data transmission*

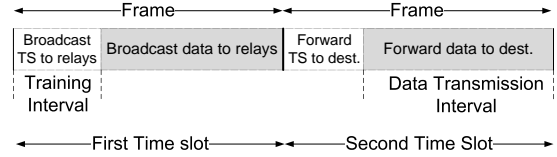


Fig. 2. Timing diagram for transmission of training sequences and data symbols within each frame.

interval, the estimated CSI can be used for optimization and coherent detection. Moreover, the proposed relaying scenario is *unidirectional* and is not based on the assumption of channel *reciprocity* as encountered in the case of time division duplex systems. In this paper, the following set of assumptions is adopted:

- A1. In both *training* and *data transmission intervals*, the signal from the source is transmitted to the destination over two time slots: in the first time slot, the source broadcasts its training and data symbols referred to as a frame, to the relays and in the second time slot, the relays *simultaneously* transmit their training and data symbols to the destination (see Fig. 1 and 2).
- A2. The TSs at source antennas and relays are assumed to be mutually *orthogonal*.
- A3. Quasi-static and frequency flat-fading channels are considered, where the source-relay-destination channels are assumed to remain constant over two time slots, i.e., over the length of a frame, the channels are modeled as *unknown non-random* parameters. However, the channels are assumed to change from frame to frame and each channel parameter is assumed to be independently and identically distributed (i.i.d), $\mathcal{CN}(0, \sigma^2)$. Similar to previous work in the literature, e.g., [10], the variance, σ^2 , is assumed to be known at the receiver.

Assumption A3 is used in previous channel estimation analyses [8]–[10], [15], [16]. Moreover, A3 is justifiable if applied to *orthogonal frequency division multiplexing (OFDM)* systems, which are used to combat frequency-selective fading in modern wireless systems.

A. Training Method and Model at the Relays

Let L denote the length of the TSs. Based on the above assumptions, the $L \times P$ matrix of received training symbols at the k th relay, $\mathbf{R}_k \triangleq [\mathbf{r}_{1,k}, \dots, \mathbf{r}_{P,k}]$ with $\mathbf{r}_{p,k} \triangleq [r_{p,k}(0), \dots, r_{p,k}(L-1)]^T$, is given by

$$\mathbf{R}_k = \mathbf{T}^{[s]} \mathbf{H}_k + \mathbf{V}_k, \quad k = 1, \dots, K \quad (1)$$

where

- $\mathbf{T}^{[s]} \triangleq [\mathbf{t}_1^{[s]}, \dots, \mathbf{t}_M^{[s]}]$ is the $L \times M$ matrix of the transmitted TSs such that $(\mathbf{T}^{[s]})^H \mathbf{T}^{[s]} = L \mathbf{I}_{M \times M}$, $\mathbf{t}_m^{[s]} \triangleq [t_m^{[s]}(0), \dots, t_m^{[s]}(L-1)]^T$ denotes the *known* TS transmitted from the m th source antenna,
- $\mathbf{H}_k \triangleq [\mathbf{h}_{1,k}, \dots, \mathbf{h}_{P,k}]$ denotes the *unknown* deterministic $M \times P$ channel matrix from source to the k th relay, $\mathbf{h}_{p,k} \triangleq [h_{p,1,k}, \dots, h_{p,M,k}]^T$, $h_{p,m,k}$ is the channel from the m th source antenna to the p th antenna of the k th relay, which is assumed to not change, i.e., *deterministic*, over a frame

but to change from frame to frame according to $h_{p,m,k} \sim \mathcal{CN}(0, \sigma_{h_{p,m,k}}^2)$, and

- $\mathbf{V}_k \triangleq [\mathbf{v}_{1,k}, \dots, \mathbf{v}_{P,k}]$ is the $L \times P$ matrix of additive white Gaussian noise (AWGN) at the k th relay and $\mathbf{v}_{p,k} \triangleq [v_{p,k}(0), \dots, v_{p,k}(L-1)]^T$ such that $\mathbf{v}_{p,k} \sim \mathcal{CN}(\mathbf{0}_{L \times 1}, \sigma_{v_k}^2 \mathbf{I}_{L \times L})$, $\forall p, k$.

To enable simultaneous estimation of the overall cooperative network's CSI, $\psi_{n,m,p,k}$, $\forall n, p, m, k$, at the destination and reduce delay, we propose that the k th relay forwards a linear transformation of its received training signal, denoted by $\tilde{\mathbf{R}}_k$, for $k = 1, \dots, K$, to the destination, given by

$$\begin{aligned} \tilde{\mathbf{R}}_k &= \mathbf{T}_k^{[r]} \Xi_k \left(\mathbf{I}_{P \times P} \diamond \left(\left(\mathbf{T}^{[s]} \right)^H \mathbf{R}_k / L \right) \right) \\ &= \mathbf{T}_k^{[r]} \Xi_k \left(\mathbf{I}_{P \times P} \diamond \mathbf{H}_k \right) + \tilde{\mathbf{V}}_k, \end{aligned} \quad (2)$$

where

- $\mathbf{T}_k^{[r]} \triangleq [\mathbf{T}_{1,k}^{[r]}, \dots, \mathbf{T}_{P,k}^{[r]}]$ is the $L \times MP$ matrix of the TSs at the k th relay such that

$$\begin{cases} \left(\mathbf{T}_k^{[r]} \right)^H \mathbf{T}_k^{[r]} = L \mathbf{I}_{MP \times MP}, & k = \bar{k} \\ \left(\mathbf{T}_k^{[r]} \right)^H \mathbf{T}_{\bar{k}}^{[r]} = \mathbf{0}_{MP \times MP}, & k \neq \bar{k} \end{cases}$$

- $\mathbf{T}_{p,k}^{[r]} \triangleq [\mathbf{t}_{p,1,k}^{[r]}, \dots, \mathbf{t}_{p,M,k}^{[r]}]$ is an $L \times M$ matrix formed by concatenating vectors of orthogonal TSs from M source antennas corresponding to the p th antenna of the k th relay,
- $\mathbf{t}_{p,m,k}^{[r]} \triangleq [t_{p,m,k}^{[r]}(0), \dots, t_{p,m,k}^{[r]}(L-1)]^T$,
- $\Xi_k \triangleq \text{diag}(\underbrace{\xi_{1,k}, \dots, \xi_{1,k}}_{\text{Repeats } M \text{ times}}, \dots, \underbrace{\xi_{P,k}, \dots, \xi_{P,k}}_{\text{Repeats } M \text{ times}})$ is an $MP \times MP$ diagonal matrix,

$$\xi_{p,k} = \sqrt{\frac{\rho_{p,k}^{[r]}}{\sum_{m=1}^M \rho_m^{[s]} \sigma_{h_{p,m,k}}^2 + \frac{M}{L} \sigma_{v_k}^2}}$$

maintains the k th relay's average transmission power constraint, $\rho_m^{[s]}$, $\rho_{p,k}^{[r]}$ denote the average transmit power from the m th source antenna and from the p th antenna of the k th relay, respectively, and

- $\tilde{\mathbf{V}}_k \triangleq \mathbf{T}_k^{[r]} \Xi_k \left(\mathbf{I}_{P \times P} \diamond \left(\mathbf{T}^{[s]H} \mathbf{V}_k / L \right) \right)$ is the $L \times P$ matrix of the transformed AWGN.

The second equality in (2) follows from Assumption A2, i.e., $(\mathbf{T}^{[s]})^H \mathbf{T}^{[s]} = L \mathbf{I}_{M \times M}$. The scaling factor $\xi_{p,k}$ is selected to keep the average power of the p th antenna of k th relay as $\rho_{p,k}^{[r]}$. In addition, $\xi_{p,k}$ follows from the assumption that unit amplitude symbols are transmitted from both source and relays, i.e., $(\mathbf{t}_m^{[s]})^H \mathbf{t}_m^{[s]} = L$ and $(\mathbf{t}_{p,m,k}^{[r]})^H \mathbf{t}_{p,m,k}^{[r]} = L$, $\forall p, m, k$ and since the i th symbol, for $i = 1, \dots, L$, of the transformed AWGN at the p th antenna of the k th relay, $\tilde{v}_{p,k}$, is distributed as $\tilde{v}_{p,k}(i) \sim \mathcal{CN}\left(0, \frac{M \sigma_{v_k}^2}{L}\right)$, $\forall p, k$.

The linear transformation in (2) ensures that the TSs corresponding to source-relay-destination antenna links are mutually orthogonal. This orthogonality of the TSs is used in Section IV to estimate the overall cooperative network's CSI, $\psi_{n,m,p,k}$, $\forall n, p, m, k$, simultaneously at the destination.

B. Training Signal Model at the Destination

The matrix of received training signals at the N antennas of the destination, $\mathbf{Y}_{L \times N} \triangleq [\mathbf{y}_1, \dots, \mathbf{y}_N]$ with $\mathbf{y}_n \triangleq$

$[y_n(0), \dots, y_n(L-1)]^T$ is given by

$$\begin{aligned} \mathbf{Y} &= \sum_{k=1}^K \tilde{\mathbf{R}}_k \mathbf{G}_k + \mathbf{W} \\ &= \sum_{k=1}^K \mathbf{T}_k^{[r]} \underbrace{\Xi_k \left(\mathbf{I}_{P \times P} \diamond \mathbf{H}_k \right) \mathbf{G}_k}_{\triangleq \Psi} + \sum_{k=1}^K \underbrace{\tilde{\mathbf{V}}_k \mathbf{G}_k}_{\triangleq \tilde{\mathbf{W}}} + \mathbf{W} \\ &= \mathbf{T}^{[r]} \Xi \Psi + \tilde{\mathbf{W}}, \end{aligned} \quad (3)$$

where

- $\mathbf{T}^{[r]} \triangleq [\mathbf{T}_1^{[r]}, \dots, \mathbf{T}_K^{[r]}]$ is an $L \times MPK$ matrix of TSs,
- $\mathbf{G}_k \triangleq [\mathbf{g}_{1,k}, \dots, \mathbf{g}_{P,k}]^T$ is a $P \times N$ matrix of *unknown* channels from the k th relay to the destination, $\mathbf{g}_{p,k} \triangleq [g_{1,p,k}, \dots, g_{N,p,k}]^T$, $g_{n,p,k}$ is the channel from the p th antenna of the k th relay to the n th destination antenna that is assumed to be *unknown* and *deterministic* over the length of a frame and to change from frame to frame according to $\mathcal{CN}(0, \sigma_{g_{n,p,k}}^2)$,
- $\Psi \triangleq [\Psi_{1,1}^T, \dots, \Psi_{P,K}^T]^T$ is the $MPK \times N$ matrix comprised of the entire collection of source-relay-destination channel coefficients, $\Psi_{p,k}$ is the *unknown* $M \times N$ source-relay-destination channel matrix corresponding to the p th antenna of the k th relay,
- $\Xi \triangleq \text{diag}(\underbrace{\xi_{1,1}, \dots, \xi_{1,1}}_{\text{Repeats } M \text{ times}}, \dots, \underbrace{\xi_{P,K}, \dots, \xi_{P,K}}_{\text{Repeats } M \text{ times}})$ is an $MPK \times MPK$ diagonal matrix,
- $\mathbf{W} \triangleq [\mathbf{w}_1, \dots, \mathbf{w}_N]$ with $\mathbf{w}_n \triangleq [w_n(0), \dots, w_n(L-1)]^T$, is the $L \times N$ AWGN noise matrix at the destination, where $\mathbf{w}_n \sim \mathcal{CN}(\mathbf{0}_{L \times 1}, \sigma_w^2 \mathbf{I}_{L \times L})$.

Based on (3) the following remark is in order:

Remark 1: Based on the proposed training method the channel parameters can be estimated simultaneously and the training and data symbols are transmitted from the source terminal to the destination in two time slots. On the other hand, using existing channel estimation algorithms, e.g., [8]–[10], the training and data symbols need to be transmitted to the destination, relay by relay, or source-antenna by source-antenna, which for the *same diversity order* would reduce the overall throughput of the system by factors of K or M , respectively. Thus, by simultaneously estimating the overall networks channel parameters, the proposed training approach significantly increases the overall throughput of MIMO multi-relay cooperative networks. For example, for accurate channel estimation, the training sequence length is typically of length $L = 16$ or $L = 32$ symbols [10]. Using a relay by relay [8] or source antenna by source antenna training approach, for a typical network with $M = 2$, $P = 2$, $K = 4$, $N = 4$, $L = 16$, the training overhead as the total time spent in training in symbol duration is 80 and 64 symbols, respectively. On the other hand, using the proposed approach the overall training overhead can be reduced to 32 symbols.

III. CRAMÉR-RAO LOWER BOUND

In this section, the exact CRLB for joint estimation of channels in AF multi-relay MIMO cooperative networks is derived. Based on the received training signal model in (3), the vector of parameters of interest, $\boldsymbol{\lambda}$, is given by

$$\boldsymbol{\lambda} \triangleq [\boldsymbol{\lambda}_1^T, \dots, \boldsymbol{\lambda}_N^T]^T, \quad (4)$$

where $\boldsymbol{\lambda}_n \triangleq [\text{Re}\{\boldsymbol{\psi}_{n,1,1}^T\}, \dots, \text{Re}\{\boldsymbol{\psi}_{n,P,K}^T\}, \text{Im}\{\boldsymbol{\psi}_{n,1,1}^T\}, \dots, \text{Im}\{\boldsymbol{\psi}_{n,P,K}^T\}]^T$ is a $2MPK \times 1$ vector and $\boldsymbol{\psi}_{n,p,k} \triangleq [\psi_{n,1,p,k}, \dots, \psi_{n,M,p,k}]^T$ is an $M \times 1$ vector of channel parameters. According to the assumptions in Section II, $\mathbf{v}_k, \mathbf{v}_{\hat{k}}, \forall k \neq \hat{k}$, $\mathbf{w}_n, \mathbf{w}_{\hat{n}}, \forall n \neq \hat{n}$, and $\mathbf{v}_k, \mathbf{w}_n, \forall k, n$, are mutually independent.

Proposition: Based on the proposed training method, the $2MPKN \times 2MPKN$ Fisher's information matrix (FIM) corresponding to (4) is given by

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{1,1} & \cdots & \mathbf{F}_{1,N} \\ \vdots & \ddots & \vdots \\ \mathbf{F}_{N,1} & \cdots & \mathbf{F}_{N,N} \end{bmatrix}, \quad (5)$$

where $\mathbf{F}_{n,\bar{n}}$, for $n, \bar{n} = 1, \dots, N$, are $2MPK \times 2MPK$ submatrices of the FIM and are given by

$$\mathbf{F}_{n,\bar{n}} = \begin{bmatrix} 2\Re\{\mathbf{U}_n^H \boldsymbol{\Sigma}_y^{-1} \mathbf{U}_{\bar{n}}\} & -2\Im\{\mathbf{U}_n^H \boldsymbol{\Sigma}_y^{-1} \mathbf{U}_{\bar{n}}\} \\ 2\Im\{\mathbf{U}_n^H \boldsymbol{\Sigma}_y^{-1} \mathbf{U}_{\bar{n}}\} & 2\Re\{\mathbf{U}_n^H \boldsymbol{\Sigma}_y^{-1} \mathbf{U}_{\bar{n}}\} \end{bmatrix}, \quad (6)$$

In (6), $\mathbf{U}_n \triangleq [\mathbf{0}_{(n-1)L \times MPK}^T, (\mathbf{T}^{[r]} \boldsymbol{\Xi})^T, \mathbf{0}_{(N-n)L \times MPK}^T]^T$ is an $LN \times MPK$ matrix, the submatrices of the $LN \times LN$ covariance matrix of the received signal vector at all the destination receive antennas, $\mathbf{y} \triangleq [\mathbf{y}_1^T, \dots, \mathbf{y}_N^T]^T$, $\boldsymbol{\Sigma}_y$, are given by

$$\boldsymbol{\Sigma}_{y_{n,\bar{n}}} = \begin{cases} \boldsymbol{\Theta} \boldsymbol{\Theta}_n^H + \sigma_w^2 \mathbf{I}_{L \times L}, & n = \bar{n} \\ \boldsymbol{\Theta} \boldsymbol{\Theta}_n^H, & n \neq \bar{n} \end{cases} \quad n, \bar{n} = 1, \dots, N, \quad (7)$$

and $\boldsymbol{\Theta}_n \triangleq [\boldsymbol{\Theta}_{n,1,1}, \dots, \boldsymbol{\Theta}_{n,P,K}]$ is the $L \times MPK$ matrix of concatenated TSs with $\boldsymbol{\Theta}_{n,p,k} \triangleq \frac{\sigma_{v_k} \xi_{p,k} g_{n,p,k}}{\sqrt{L}} \mathbf{T}_{p,k}^{[r]}$.

Proof: See Appendix A.

The CRLB is given by

$$\text{CRLB}(\boldsymbol{\lambda}) = \mathbf{F}^{-1}. \quad (8)$$

Moreover, in order to derive the CRLB for the estimation of the combined real and imaginary parts of channel parameters, the set of parameters of interest, $\boldsymbol{\lambda}$, is modified as [17]

$$\bar{\boldsymbol{\lambda}} \triangleq \underbrace{\begin{bmatrix} \mathbf{I}_{MPKN \times MPKN} & j\mathbf{I}_{MPKN \times MPKN} \end{bmatrix}}_{\triangleq \mathbf{J}} \boldsymbol{\lambda}. \quad (9)$$

The CRLB for the estimation of $\bar{\boldsymbol{\lambda}}$ is given by

$$\text{CRLB}(\bar{\boldsymbol{\lambda}}) = \mathbf{J} \text{CRLB}(\boldsymbol{\lambda}) \mathbf{J}^H. \quad (10)$$

Since the CRLB is the lower bound on estimation of deterministic parameters it is often the case that it depends on the set of parameters that are being estimated [17].

Remark 2: Based on (8), it can be concluded that the CRLB and channel estimation performance are a function of TS length, L , channels from relays to destination, SNR, and number of source antennas, M , due to $\xi_{p,k}$. In addition, the CRLB in (8) demonstrates that due to the proposed training approach, unlike the existing channel estimators for AF relaying networks, e.g., [10], channel estimation performance of the proposed scheme is not affected by the number of relays, K , number of antennas per relay, P , nor by the number of destination antennas, N .

IV. CHANNEL ESTIMATION

Based on the parameterizations of the estimation problem in Section III, in this section the LS and ML estimators for the estimation of the source-relays-destination channels are derived.

A. Least Squares Estimator

From (3), the optimal LS estimates of the channel matrix $\boldsymbol{\Psi}$, denoted by $\hat{\boldsymbol{\Psi}}^{[\text{LS}]}$, can be determined as

$$\begin{aligned} \hat{\boldsymbol{\Psi}}^{[\text{LS}]} &= \frac{\boldsymbol{\Xi}^{-1} (\mathbf{T}^{[r]})^H \mathbf{Y}}{L} \\ &= \boldsymbol{\Psi} + \underbrace{\frac{\boldsymbol{\Omega} + \boldsymbol{\Xi}^{-1} (\mathbf{T}^{[r]})^H \mathbf{W}}{L}}_{\triangleq \Delta \boldsymbol{\Psi}^{[\text{LS}]}} \end{aligned} \quad (11)$$

where $[\boldsymbol{\Omega}]_{MPK \times N} \triangleq \left[\left((\mathbf{I}_{P \times P} \diamond (\mathbf{T}^{[s]H} \mathbf{V}_1)) \mathbf{G}_1 \right)^T, \dots, \left(\mathbf{I}_{P \times P} \diamond (\mathbf{T}^{[s]H} \mathbf{V}_K) \mathbf{G}_K \right)^T \right]^T$ and $\mathbf{T}^{[s]} \triangleq [\mathbf{t}_1^{[s]}, \dots, \mathbf{t}_M^{[s]}]$. The second equality in (11) follows from (3), and the proposed training method in (2) and the third equality in (11) follows from the definition of $\tilde{\mathbf{V}}_k$ in (2).

Based on the assumptions in Section II, $\mathbf{v}_k, \mathbf{v}_{\hat{k}}, \forall k \neq \hat{k}$, $\mathbf{w}_n, \mathbf{w}_{\hat{n}}, \forall n \neq \hat{n}$, and $\mathbf{v}_k, \mathbf{w}_n, \forall k, n$, are mutually independent. Thus, the covariance matrix of the LS estimation error, $\Delta \boldsymbol{\Psi}^{[\text{LS}]}$, can be calculated as

$$\begin{aligned} \boldsymbol{\Sigma}_{\Delta \boldsymbol{\Psi}^{[\text{LS}]}} &= \frac{1}{L^2} \mathbb{E} \left[\left(\boldsymbol{\Omega} + \boldsymbol{\Xi}^{-1} (\mathbf{T}_{p,k}^{[r]})^H \mathbf{W} \right) \right. \\ &\quad \left. \times \left(\boldsymbol{\Omega} + \boldsymbol{\Xi}^{-1} (\mathbf{T}_{p,k}^{[r]})^H \mathbf{W} \right)^H \right] \end{aligned} \quad (12a)$$

$$= \frac{1}{L^2} \mathbb{E} [\boldsymbol{\Omega} \boldsymbol{\Omega}^H] + \frac{\sigma_w^2}{L} N \boldsymbol{\Xi}^{-2} \mathbf{I}_{MPK \times MPK} \quad (12b)$$

$$= \frac{1}{L} (\boldsymbol{\Phi} + N \sigma_w^2 \boldsymbol{\Xi}^{-2}) \mathbf{I}_{MPK \times MPK}, \quad (12c)$$

where $\boldsymbol{\Phi} \triangleq \text{diag} \left(\underbrace{\sigma_{v_1}^2 \|\mathbf{g}_{1,1}\|^2, \dots, \sigma_{v_1}^2 \|\mathbf{g}_{1,1}\|^2}_{\text{Repeats } M \text{ times}}, \dots, \right.$

$\left. \underbrace{\sigma_{v_K}^2 \|\mathbf{g}_{P,K}\|^2, \dots, \sigma_{v_K}^2 \|\mathbf{g}_{P,K}\|^2}_{\text{Repeats } M \text{ times}} \right)$ is an $MPK \times MPK$ diagonal

matrix. Eq. (12b) follows from (12a), since the noise at the relays and destination are assumed to be mutually independent and due to the orthogonality of the transmitted TSs.

Finally, it can be shown that the proposed LS estimator is an unbiased estimator, since,

$$\mathbb{E} [\hat{\boldsymbol{\Psi}}^{[\text{LS}]}] = \mathbb{E} \left[\boldsymbol{\Psi} + \frac{\boldsymbol{\Omega} + \boldsymbol{\Xi}^{-1} (\mathbf{T}^{[r]})^H \mathbf{W}}{L} \right] \quad (13a)$$

$$= \boldsymbol{\Psi} + \frac{1}{L} \mathbb{E} [\boldsymbol{\Omega} + \boldsymbol{\Xi}^{-1} (\mathbf{T}^{[r]})^H \mathbf{W}] = \boldsymbol{\Psi}, \quad (13b)$$

where the second equality in (13b) follows from the fact that $\boldsymbol{\Omega}$ and \mathbf{W} denote the zero-mean Gaussian noise at the relays and destination, respectively.

B. Weighted Least Squares

Since \mathbf{y} defined below (6) is a Gaussian observation vector and since

, the log-likelihood function (LLF) of the vector of channel parameters, the covariance matrix, $\Sigma_{\mathbf{y}}$, is positive semi-definite¹, using Cholesky decomposition, the log-likelihood function (LLF) of the channel parameters $\psi = [\psi_{1,1,1}^T, \dots, \psi_{N,P,K}^T]^T$, up to an additive constant, $\varrho(\mathbf{y})$, is given by

$$\varrho(\mathbf{y}) = \left(\Pi \mathbf{y} - \Pi \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \psi \right)^H \times \left(\Pi \mathbf{y} - \Pi \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \psi \right). \quad (14)$$

According to (14), it is well-known that the ML estimator for ψ denoted by $\hat{\psi}^{[\text{ML}]}$ is given by

$$\hat{\psi}^{[\text{ML}]} = \left(\left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right)^H \Sigma_{\mathbf{y}}^{-1} \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \right)^{-1} \times \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \Sigma_{\mathbf{y}}^{-1} \mathbf{y}. \quad (15)$$

In addition, it can be also shown that the proposed ML estimator is unbiased since

$$\mathbb{E} \left[\hat{\psi}^{[\text{ML}]} \right] = \mathbb{E} \left[\left(\left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right)^H \Sigma_{\mathbf{y}}^{-1} \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \right)^{-1} \times \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \Sigma_{\mathbf{y}}^{-1} \times \left(\left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \psi + \tilde{\mathbf{w}} \right) \right] \quad (16a)$$

$$= \psi + \left(\left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right)^H \Sigma_{\mathbf{y}}^{-1} \times \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \right)^{-1} \left(\mathbf{T}^{[r]} \Xi \otimes \mathbf{I}_{N \times N} \right) \times \Sigma_{\mathbf{y}}^{-1} \mathbb{E} [\tilde{\mathbf{w}}] \quad (16b)$$

$$= \psi. \quad (16c)$$

Since it is well known that multiplication of an $X \times W$ by a $W \times Z$ matrix and the inverse of an $X \times X$ matrix require $\mathcal{O}(XWZ)$ and $\mathcal{O}(X^3)$ multiplications plus additions, respectively, with $\mathcal{O}(\cdot)$ denoting the big Omicron notation [18], the computational requirements of the proposed LS and ML algorithms are $\mathcal{O}(MPKLN)$ and $\mathcal{O}(M^3P^3K^3N^3 + M^2P^2K^2N^2L)$, respectively. The following remark is in order:

Remark 3: The covariance matrix of the observation vector at all the receive antennas of the destination, $\Sigma_{\mathbf{y}}$ defined in (7), is a function of the *unknown* channels from relays to destination, $\mathbf{g}_{p,k} = [g_{1,p,k}, \dots, g_{N,p,k}]^T$, for $p = 1, \dots, P$ and $k = 1, \dots, K$. As a result, the proposed ML estimator in (15) has little practical application in estimating the source-relays-destination channels. However, note that $g_{n,p,k}$ is distributed as $\mathcal{CN}(0, \sigma_{g_{n,p,k}}^2)$. Thus, $g_{n,p,k}$ can be replaced by $\sigma_{g_{n,p,k}}$ in (7), to derive a *weighted least squares (WLS)* estimator for obtaining the source-relays-destination channels.²

¹The covariance matrix of a vector of random variables is always positive semi-definite.

²Note that $\sigma_{g_{n,p,k}}^2$ can be estimated at the destination using routine estimation algorithms as outlined in [17].

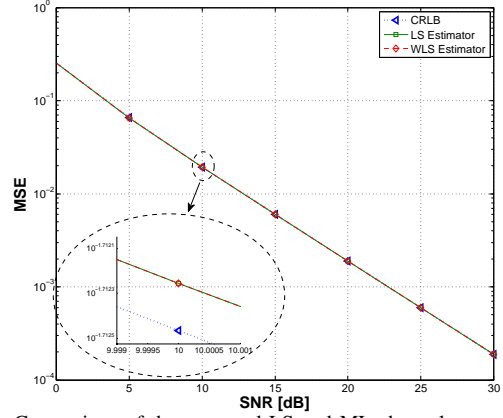


Fig. 3. Comparison of the proposed LS and ML channel estimators' MSE and the CRLB in (8) vs. SNR for $L = 16$, $K = M = N = P = 2$.

V. NUMERICAL RESULTS AND DISCUSSIONS

Throughout this section, propagation loss is modeled as $\beta = (d/d_0)^{-m}$ [19], where d represents the distance between transmitter and receiver, d_0 is the reference distance, and m is the path loss exponent. In this paper $d_0 = 1\text{km}$ and $m = 2.3$, which corresponds to urban area cellular networks. The distances from source to relays and relays to destination are denoted by $d^{[\text{sr}]}$ and $d^{[\text{rd}]}$, respectively, with $d^{[\text{sr}]} + d^{[\text{rd}]} = 2\text{km}$. The channels from source antennas to relays and from relays to destination antennas are modeled as i.i.d complex Gaussian random variables with $\mathcal{CN}(0, 1)$. Walsh-Hadamard codes combined with binary phase-shift keying (BPSK) modulation are used during the training interval. Finally, $\rho_m^{[s]} = \rho_{p,k}^{[r]} = 1$, $\forall m, p, k$, the noise variances at the relays, $\sigma_{v_k}^2$, $\forall k$, and destination, σ_w^2 is set to the inverse of source to relay and relay to destination SNRs, respectively.

Fig. 3 presents the CRLB and MSE of the LS and WLS channel estimators vs. SNR. Note that the proposed LS and WLS estimators both are close to the CRLB and their MSEs are very close. The latter follows from the fact that the error corresponding to the LS estimator is white as shown in (12) due to the application of proposed training algorithm. Therefore, there are no correlations in the residual channel estimation errors that can be exploited by the WLS estimator for better estimation performance.

In Fig. 4 the performance of the proposed LS estimator is compared against the estimator in [10], where time division multiple access (TDMA) is used at the source antennas to extend the results in [10] to the case of MIMO cooperative networks. Due to the application of TDMA, the algorithm in [10] requires a total of $2M$ time slots to estimate the source-relays-destination channels (the proposed scheme requires 2 time slots). Therefore, to maintain the same channel estimation overhead and ensure a fair comparison the TS length, L , for the proposed algorithm and the algorithm in [10] is set to $L = 32$ and $L/M = 8$, respectively ($M = 4$ in this case). Fig. 4 illustrates that the LS algorithm proposed in this paper outperform the estimator in [10] by an average margin of 3.2 dB. Moreover, it should be pointed out that due to the overhead required by the algorithm in [10], its application to AF MIMO cooperative networks is not straightforward. Finally, comparison of the MSE of the proposed estimators against other references in addition to [10] is made difficult by the fact that the parameterizations

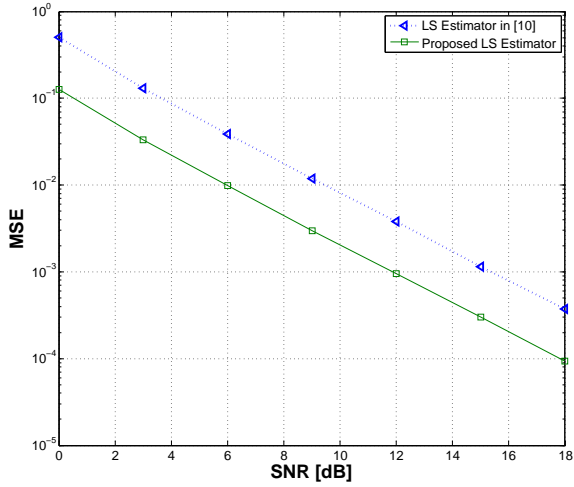


Fig. 4. Comparison of the proposed LS estimator and the estimator in [10] for $M = 4$, $N = 2$, $K = 4$, $L = 32$, and $P = 1$.

and training methods applied in many references significantly vary and are not applicable to the proposed channel estimation problem. However, in this paper the CRLB for the estimation of channels is derived and used as a performance benchmark for the proposed estimators. This is one of the main advantages of the analysis presented in this paper.

Fig. 5 shows the ABERs of AF relaying and the cooperative relaying in [20] for different numbers of antennas and relays when considering perfect and estimated CSI. To maximize throughput, during the data transmission interval the *Vertical Bell Laboratories Layered Space-Time Architecture* [21] in combination with a *minimum-mean square error (MMSE)* equalizer and *successive interference cancellation (SIC)* are applied. Quadrature phase shift keying (QPSK) modulation is used with $L = 16$ and a frame length of 128 symbols, resulting in an overall estimation overhead of 12.5%. The LS estimator proposed in Section IV is used, where it is assumed that channel estimation is performed at an SNR=30dB. Relays are uniformly distributed throughout the network. New channels are generated from frame to frame. Two scenarios are provided: 1. the relays simply amplify and forward the received signal to the destination (denoted by *AF relaying*); 2. to demonstrate that the estimated channels can be applied to improve the overall networks performance, the AF relaying approach in [20] in combination with estimated or perfect CSI is used to optimize the performance of an AF relaying MIMO cooperative network (denoted by *AF relaying in [20]*). The results in Fig. 5 illustrate that in the case of AF relaying the proposed channel estimation algorithm can be effectively combined with optimization algorithms, e.g., in this case the algorithm in [20], to significantly improve the performance of multi-relay AF MIMO cooperative network. For example, compared to AF, by applying the LS estimated CSI, the AF relaying algorithm in [20] can result in a 5 dB performance gain for a network equipped with 2 relays and 2 antennas at all the terminals. This demonstrates in addition to enabling coherent detection at the destination, the estimated channels parameters can be applied to cooperative optimization. Fig. 5 establishes the importance of considering the effect of imperfect channel estimation on the performance of AF MIMO cooperative networks.

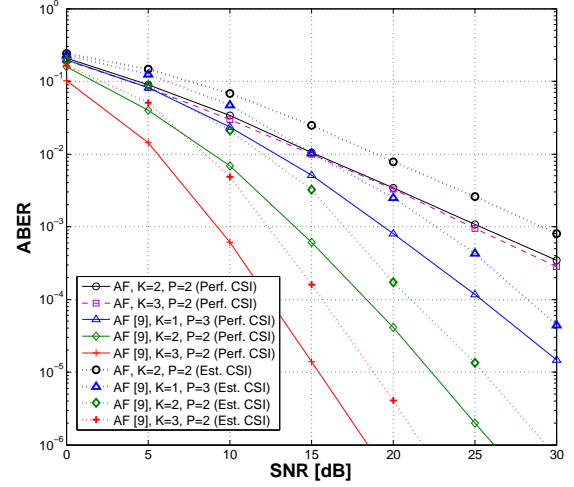


Fig. 5. BER performance of AF and the AF relaying in [20] for perfect and estimated CSI for $M = 2$, and $N = 2$.

VI. CONCLUSION

In this paper a new training algorithm is proposed that can simultaneously estimate a MIMO relaying network's complete CSI at the destination node. Next, an LS estimator is derived and it is shown that its performance is close to the derived CRLB and to that of the derived WLS estimator over a wide range of SNR values. Unlike estimators proposed in the literature, the proposed estimators do not require the relays to perform channel estimation. Furthermore, the proposed estimators are shown to outperform the channel estimator in [10] by an average of 3.2 dB. Moreover, it is demonstrated that the estimated source-relays-destination channels can be effectively applied to enhance cooperative communication, e.g., [20], to significantly enhance the performance of MIMO AF multi-relay cooperative networks in terms of ABER.

APPENDIX A DERIVATION OF CRLB

In this appendix, we provide the derivation steps for the FIM for channel estimation in MIMO AF multi-relay networks. Based on the proposed training method and assumptions in Section II, the received training vector at all receive antennas, $\mathbf{y} \triangleq [\mathbf{y}_1^T, \dots, \mathbf{y}_N^T]^T$, is distributed as $\mathcal{CN}(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)$, where the mean of \mathbf{y} is given by $\boldsymbol{\mu}_y \triangleq [\boldsymbol{\mu}_{y_1}^T, \dots, \boldsymbol{\mu}_{y_N}^T]^T$ with $\boldsymbol{\mu}_{y_n} \triangleq \sum_{k=1}^K \sum_{p=1}^P \xi_{p,k} \mathbf{T}_{p,k}^{[r]} \boldsymbol{\psi}_{n,p,k}$ and the $LN \times LN$ covariance matrix of \mathbf{y} , $\boldsymbol{\Sigma}_y$, is given by

$$\boldsymbol{\Sigma}_{\mathbf{y}_{n,\bar{n}}} = \begin{cases} \sum_{k=1}^K \sigma_{v_k}^2 \sum_{p=1}^P \frac{\xi_{p,k}^2 |g_{n,p,k}|^2}{L} \mathbf{T}_{p,k}^{[r]} \left(\mathbf{T}_{p,k}^{[r]} \right)^H + \sigma_w^2 \mathbf{I}_{L \times L}, & n = \bar{n} \\ \sum_{k=1}^K \sigma_{v_k}^2 \sum_{p=1}^P \frac{\xi_{p,k}^2 |g_{n,p,k}|^2}{L} \mathbf{T}_{p,k}^{[r]} \left(\mathbf{T}_{p,k}^{[r]} \right)^H, & n \neq \bar{n} \end{cases} \quad (\text{A.1a})$$

$$= \begin{cases} \boldsymbol{\Theta} \boldsymbol{\Theta}^H + \sigma_w^2 \mathbf{I}_{L \times L}, & n = \bar{n} \\ \boldsymbol{\Theta} \boldsymbol{\Theta}_n^H, & n \neq \bar{n}. \end{cases} \quad (\text{A.1b})$$

where $\boldsymbol{\Theta}_n \triangleq [\boldsymbol{\Theta}_{1,1}, \dots, \boldsymbol{\Theta}_{P,K}]$ is the $L \times MPK$ matrix of concatenated TSs and $\boldsymbol{\Theta}_{n,p,k} \triangleq \frac{\sigma_{v_k} \xi_{p,k} g_{n,p,k}}{\sqrt{L}} \mathbf{T}_{p,k}^{[r]}$. Note that (A.1) follows from the assumption that the noise terms at the

destination and relays are mutually independent, the definition of $\tilde{\mathbf{v}}_{p,k}$, and the assumption that the noise terms at the antennas of the relays are mutually independent. For parameter estimation from a complex Gaussian observation sequence, the entries of the mpk th row and $\bar{m}\bar{p}\bar{k}$ th column of the submatrix $\mathbf{F}_{n,\bar{n}}$, for $m, \bar{m} = 1, \dots, M$, $p, \bar{p} = 1, \dots, P$, $k, \bar{k} = 1, \dots, K$, and $n, \bar{n} = 1, \dots, N$, are given by [17]

$$\begin{aligned} [\mathbf{F}_{n,\bar{n}}]_{mpk,\bar{m}\bar{p}\bar{k}} &= 2\text{Re} \left\{ \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}^H}{\partial \lambda_{n,m,p,k}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \lambda_{\bar{n},\bar{m},\bar{p},\bar{k}}} \right\} \\ &+ \text{Tr} \left(\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \lambda_{n,m,p,k}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \lambda_{\bar{n},\bar{m},\bar{p},\bar{k}}} \right), \end{aligned} \quad (\text{A.2})$$

where $\boldsymbol{\lambda}$ is defined in (4). The derivative components of the CRLB are computed as

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \text{Re}\{\psi_{n,m,p,k}\}} &= -j \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \text{Im}\{\psi_{n,m,p,k}\}} \\ &= \underbrace{\left[\mathbf{0}_{1 \times L(n-1)}, \xi_{p,k} \left(\mathbf{t}_{p,m,k}^{[r]} \right)^T, \mathbf{0}_{1 \times L(N-n)} \right]^T}_{\triangleq \mathbf{u}_{n,m,p,k}}, \end{aligned} \quad (\text{A.3})$$

and using (A.1a)

$$\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \text{Re}\{\psi_{n,m,p,k}\}} = \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \text{Im}\{\psi_{n,m,p,k}\}} = 0. \quad (\text{A.4})$$

After substituting the derivatives in (A.3) and (A.4) into (A.2) and carrying out straightforward algebraic manipulations, the elements of $\mathbf{F}_{n,\bar{n}}$ can be obtained as

$$[\mathbf{F}_{n,\bar{n}}]_{mpk,\bar{m}\bar{p}\bar{k}} = 2\text{Re} \left\{ \mathbf{u}_{n,m,p,k}^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{u}_{\bar{n},\bar{m},\bar{p},\bar{k}} \right\}, \quad (\text{A.5a})$$

$$[\mathbf{F}_{n,\bar{n}}]_{mpk,(\bar{m}+M)(\bar{p}+P)(\bar{k}+K)} = -2\text{Im} \left\{ \mathbf{u}_{n,m,p,k}^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{u}_{\bar{n},\bar{m},\bar{p},\bar{k}} \right\}, \quad (\text{A.5b})$$

$$[\mathbf{F}_{n,\bar{n}}]_{(m+M)(p+P)(k+K),\bar{m}\bar{p}\bar{k}} = 2\text{Im} \left\{ \mathbf{u}_{n,m,p,k}^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{u}_{\bar{n},\bar{m},\bar{p},\bar{k}} \right\}, \quad (\text{A.5c})$$

$$[\mathbf{F}_{n,\bar{n}}]_{(m+M)(p+P)(k+K),(\bar{m}+M)(\bar{p}+P)(\bar{k}+K)} = 2\text{Re} \left\{ \mathbf{u}_{n,m,p,k}^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{u}_{\bar{n},\bar{m},\bar{p},\bar{k}} \right\}. \quad (\text{A.5d})$$

Using (A.5a)-(A.5d), $\mathbf{FIM}_{k,\bar{k}}$, for $k, \bar{k} = 1, \dots, N_r$, can be obtained as shown in (6) in Section III.

APPENDIX II

Proof of Theorem 2

REFERENCES

- [1] N. J. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 3, pp. 3062–3080, Dec. 2004.
- [2] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, Part I and Part II," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1948, Nov. 2003.
- [3] B. Khoshnevis, W. Yu, and R. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, Oct. 2008.
- [4] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.
- [5] T.-H. Pham, Y.-C. Liang, H. K. Garg, and A. Nallanathan, "Joint channel estimation and data detection for MIMO-OFDM two-way relay networks," in *Proc. IEEE Global Commun. Conf. (GC)*, Miami, USA, Dec. 2010, pp. 1–5.
- [6] J. Ma, P. Orlik, J. Zhang, and G.-Y. Li, "Pilot matrix design for interim channel estimation in two-hop MIMO AF relay systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dresden, Germany, Jun. 2009, pp. 14–18.
- [7] J. Pang, G. Shen, D. Wang, L. Jiang, and W. Wang, "Channel estimation and optimal training design for amplify and forward MIMO relay channel under spatial fading correlation," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Ottawa, Canada, Sep. 2010, pp. 1–5.
- [8] P. Lioliou and M. Viberg, "Least-squares based channel estimation for MIMO relays," in *Proc. Int. ITG Workshop Smart Antennas (WSA)*, Darmstadt, Germany, Feb. 2008, pp. 90–95.
- [9] P. Lioliou, M. Viberg, and M. Coldrey, "Performance analysis of relay channel estimation," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, USA, May 2010, pp. 1533–1538.
- [10] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.
- [11] S. Serbetli and A. Yener, "MMSE transmitter design for correlated MIMO systems with imperfect channel estimates: Power allocation trade-offs," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2295–2304, Aug. 2006.
- [12] 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.
- [13] X. Li, C. Xing, Y.-C. Wu, and S. C. Chan, "Timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Trans. Signal Processing*, vol. 58, no. 4, pp. 2218–2229, Apr. 2010.
- [14] A. A. Nasir, H. Mehrpouyan, S. D. Blostein, S. Durrani, and R. A. Kennedy, "Timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Trans. Signal Processing*, vol. 60, no. 2, pp. 793–811, Feb. 2012.
- [15] C.S. Patel and G. L. Stüber, "Channel estimation for amplify and forward relay based cooperation diversity systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2348–2356, Jun. 2007.
- [16] B. Gedik and M. Uysal, "Impact of imperfect channel estimation on the performance of amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1468–1479, Mar. 2009.
- [17] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*. Prentice Hall, Signal Processing Series, 1993.
- [18] P. Borwein, "On the complexity of calculating factorials," *Journal of Algorithms*, vol. 6, pp. 376–380, 1985.
- [19] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. Wiley-InterScience, John Wiley & Sons, Inc., 1997.
- [20] Y. Zheng, H. Mehrpouyan, and S. D. Blostein, "Application of phase shift in coherent multi-relay MIMO communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dresden, Germany, Jun. 2009, pp. 1–5.
- [21] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI Int. Symposium on Signals, Systems, and Elect. (ISSSE)*, Pisa Italy, Oct 1998, pp. 295–300.