



# CHALMERS

## Chalmers Publication Library

### **Bayesian Channel Estimation Techniques for AF MIMO Relaying Systems**

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

**IEEE VTS Vehicular Technology Conference Proceedings. VTC 2011 - Fall (ISSN: 1090-3038)**

Citation for the published paper:

Lioliou, P. ; Viberg, M. ; Matthaiou, M. (2011) "Bayesian Channel Estimation Techniques for AF MIMO Relaying Systems". IEEE VTS Vehicular Technology Conference Proceedings. VTC 2011 - Fall

<http://dx.doi.org/10.1109/VETECF.2011.6093283>

Downloaded from: <http://publications.lib.chalmers.se/publication/148497>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

# Bayesian Channel Estimation Techniques for AF MIMO Relaying Systems

Panagiota Lioliou, Mats Viberg, and Michail Matthaiou

Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden

E-mail: {panagiota.lioliou, viberg, michail.matthaiou}@chalmers.se

**Abstract**—In this paper, we consider the fundamental problem of channel estimation in multiple-input multiple-output (MIMO) amplify-and-forward (AF) relaying systems operating over random channels. Using the Bayesian framework, linear minimum mean square error (LMMSE) and expectation-maximization (EM) based maximum a posteriori (MAP) channel estimation algorithms are developed, that provide the destination with full knowledge of all channel parameters involved in the transmission. The performance of the proposed algorithms is evaluated in terms of the mean square error (MSE) as a function of the signal-to-noise ratio (SNR) during the training interval. Our simulation results show that the incorporation of prior knowledge into the channel estimation algorithm offers improved performance, especially in the low SNR regime.

## I. INTRODUCTION

Recently, the introduction of relaying nodes in wireless channels has triggered a significant research interest [1]. The deployment of relays has been identified as a suitable technique for providing broader coverage, higher transmission rates, and increased reliability. Moreover, when the source and destination are equipped with multiple antennas, the concept of relaying is combined with that of multiple-input multiple-output (MIMO) systems. Then, the system's performance can be further enhanced by exploiting the spatial dimension [2].

For different relaying strategies, the problem of power allocation and network beamforming has been well studied in the existing literature [3]. Several designs for the optimal relay amplifying factor have been proposed for maximizing the received signal-to-noise ratio (SNR) or some other performance measure, such as the channel capacity [3], [4]. Most existing relaying schemes, however, assume that perfect instantaneous channel state information (CSI) is available at the receiver. Therefore, it becomes clear that to exploit the advantages of MIMO relaying networks, an accurate CSI of all involved links is required. Despite the importance of this prerequisite, channel estimation is often ignored by either assuming perfect CSI or by considering only the estimation of the compound (from source to destination) channel [5]. While CSI of the compound channel guarantees feasible data detection at the destination, the knowledge of the individual channel responses can be utilized to further improve the overall system performance.

In this paper, we study stochastic channel estimation techniques for MIMO relaying systems with random channels. In particular, we develop linear minimum mean square error (LMMSE) and expectation-maximization (EM) based maximum a posteriori (MAP) channel estimation schemes for

a two-hop amplify-and-forward (AF) relaying system. We mainly focus on full-duplex AF relaying systems since their higher spectral efficiency and easy deployment make them attractive from a practical point of view [6]. However, we point out that the proposed algorithm can be straightforwardly extended to the half-duplex case. We construct the problem of channel estimation such that all channel parameters involved in the transmission are estimated by using measurements only at the destination. The clear advantage of this approach is that very simple, cost-efficient AF relaying nodes can be deployed, without any additional computational overhead at the relays, since the estimation and the resulting computational effort is centralized at the destination. The channel estimation problem would of course be simpler if the relays could also transmit pilot symbols [7]. However, this implies that the relays are aware of the structure of their received signal and capable of performing further signal processing. In addition, we remark that pilots at the relays also consume extra channel usage. Thus, there would be no gain in terms of training cost.

The rest of this paper is organized as follows: Section II presents the system and signal model. The channel estimation technique is presented in Sections III and IV. Numerical examples are provided in Section V while concluding remarks are given in Section VI.

**Notation.** We use bold upper case letters to denote matrices and bold lower case letters to denote vectors. The transpose, hermitian, and inverse of a matrix  $\mathbf{A}$  are denoted by  $\mathbf{A}^T$ ,  $\mathbf{A}^H$ , and  $\mathbf{A}^{-1}$ , respectively;  $\mathbf{I}_N$  is the identity matrix of size  $N$ . With  $\text{vec}(\mathbf{A})$ , we denote the vectorization of a matrix  $\mathbf{A}$ ;  $\text{Tr}(\mathbf{A})$  is the trace of the matrix  $\mathbf{A}$ . The operators  $\otimes$  and  $\diamond$  stand for the Kronecker and the Khatri-Rao product, respectively. Finally,  $E\{\}$  denotes expectation while  $\text{vecd}(\mathbf{A})$  stands for the vector formed from the diagonal elements of  $\mathbf{A}$ .

## II. SYSTEM AND SIGNAL MODEL

We consider the communication between the source (S) and the destination (D) with the help of  $R$  intermediate relaying nodes. The source and the destination are equipped with  $N$  and  $M$  antennas, respectively, whereas the relaying nodes are equipped with one antenna each, see Figure 1. Let us denote the channel matrix from source to relays as  $\mathbf{H}_1 \in \mathcal{C}^{R \times N}$ , and from relays to destination as  $\mathbf{H}_2 \in \mathcal{C}^{M \times R}$ . It is assumed that there is no direct path between S and D due to the surrounding environment (e.g. in an urban environment). The two links experience independent, but not necessarily identical fading.

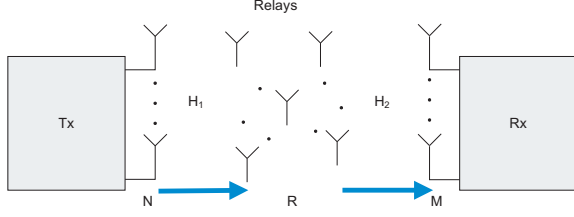


Fig. 1. Schematic illustration of a MIMO relaying network

In particular, we investigate an asymmetric scenario for the fading distributions of the S-R and R-D links: The S-R link is subject to Rician fading and the R-D link is subject to Rayleigh fading [8]. The proposed model can represent either an up or down link in a cellular network. In this case, a base station acts as S, a mobile node as R and another mobile node as D. Therefore,  $\mathbf{H}_1$  is a Rician fading channel with mean  $\frac{K}{K+1}\mathbf{H}_0$  and covariance  $\mathbf{R}_1 = \frac{1}{K+1}\mathbf{I}_{RN}$ , where  $\mathbf{H}_0$  represents the direct path component matrix, while  $K$  denotes the Rician  $K$ -factor defined as the ratio of the powers of the LoS component to the scattered components. The channel matrix between the relays and the destination,  $\mathbf{H}_2$ , is Rayleigh distributed containing i.i.d. zero-mean unit-variance complex Gaussian entries.

Assume that the source transmits the signal vector  $\mathbf{x}(n) \in \mathcal{C}^{N \times 1}$  to the relays and the destination, where  $n$  is the time index. The relays amplify and forward their received signal vector to the destination. We consider a narrowband channel and therefore, we can neglect any channel dynamics, including delays introduced by the relay. Hence, the received signal at the relays  $\mathbf{r}(n) \in \mathcal{C}^{R \times 1}$  and the destination  $\mathbf{y}(n) \in \mathcal{C}^{M \times 1}$  can be modeled as

$$\mathbf{r}(n) = \mathbf{H}_1 \mathbf{x}(n) + \mathbf{w}_R(n), \quad (1)$$

$$\mathbf{y}(n) = \mathbf{H}_2 \mathbf{G} \mathbf{r}(n) + \mathbf{w}_D(n), \quad (2)$$

where  $\mathbf{w}_R(n) \sim \mathcal{CN}(0, \sigma_r^2 \mathbf{I}_R)$  and  $\mathbf{w}_D(n) \sim \mathcal{CN}(0, \sigma_d^2 \mathbf{I}_M)$  are the independent white complex Gaussian noise contributions at the relays and the destination, respectively. If we assume that the relays operate in a distributed manner, then the amplification matrix  $\mathbf{G} \in \mathcal{C}^{R \times R}$  is diagonal, and its diagonal elements contain the complex amplifying factors,  $\{g_i\}_{i=1}^R$ , for each relaying node. We can now rewrite (2) as

$$\mathbf{y}(n) = \mathbf{H} \mathbf{x}(n) + \mathbf{w}(n), \quad (3)$$

where

$$\mathbf{H} = \mathbf{H}_2 \mathbf{G} \mathbf{H}_1, \quad (4)$$

$$\mathbf{w}(n) = \mathbf{H}_2 \mathbf{G} \mathbf{w}_R(n) + \mathbf{w}_D(n). \quad (5)$$

Thus,  $\mathbf{H} \in \mathcal{C}^{M \times N}$  and  $\mathbf{w}(n) \in \mathcal{C}^{M \times 1}$  correspond to the equivalent (compound) channel matrix and noise, respectively.

### III. CHANNEL ESTIMATION

In this section, we develop LMMSE and EM-based MAP algorithms to estimate the individual channel matrices ( $\mathbf{H}_1$ ,  $\mathbf{H}_2$ ) at the destination node. For the simplicity of explanation, we initially study the case of one relay ( $R = 1$ ) that assists the communication between the source and the destination.

For channel estimation purposes, a known training sequence  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_L]$  of length  $N$  is transmitted by the source, where  $L \geq N$ . We assume that the training sequence is orthogonal across all transmitting antennas [9]. The output of the training block  $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_L]$  can then be expressed as

$$\mathbf{Y} = \mathbf{H} \mathbf{X} + \mathbf{W}, \quad (6)$$

where  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_L]$ . After vectorizing each term in (6), we obtain

$$\text{vec}(\mathbf{Y}) = \text{vec}(\mathbf{H} \mathbf{X}) + \text{vec}(\mathbf{W}). \quad (7)$$

By using the identity  $\text{vec}(\mathbf{A} \mathbf{B} \mathbf{C}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$ , we can rewrite (7) as

$$\mathbf{y} = \left\{ (g \mathbf{h}_1 \mathbf{X})^T \otimes \mathbf{I}_M \right\} \mathbf{h}_2 + \mathbf{w}, \quad (8)$$

where  $g$  is the relay amplification factor while  $\mathbf{h}_1 \in \mathcal{C}^{1 \times N}$  and  $\mathbf{h}_2 \in \mathcal{C}^{M \times 1}$  are the unknown S-R and R-D channel vectors, respectively. Equivalently, (8) can also be expressed as

$$\mathbf{y} = (\mathbf{X}^T \otimes \mathbf{h}_2 g) \mathbf{h}_1^T + \mathbf{w}. \quad (9)$$

#### A. Suboptimal LMMSE-based Channel Estimation

Under the Bayesian framework, a typical procedure for jointly selecting the channel estimators  $\hat{\mathbf{h}}_1$  and  $\hat{\mathbf{h}}_2$  is to minimize the Bayesian MSE leading to the well known MMSE estimator [10]

$$\hat{\boldsymbol{\theta}} = E\{\boldsymbol{\theta} | \mathbf{y}\}, \quad (10)$$

where  $\hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \end{bmatrix}$ . However, since the computational complexity involved in finding the MMSE estimator is too high, we propose a simpler two step approach. In particular, the LMMSE estimate of  $\mathbf{h}_2$  is first computed as

$$\hat{\mathbf{h}}_2 = \tilde{\mathbf{H}}_0^H (\tilde{\mathbf{H}}_0 \tilde{\mathbf{H}}_0^H + |g|^2 \mathbf{R}_1 \otimes \mathbf{I}_M + \mathbf{R}_{ww})^{-1} \mathbf{y}, \quad (11)$$

where  $\tilde{\mathbf{H}}_0 = E\{(g \mathbf{h}_1 \mathbf{X})^T \otimes \mathbf{I}_M\} = \sqrt{\frac{K}{K+1}} (g \mathbf{h}_0 \mathbf{X})^T \otimes \mathbf{I}_M$ ,  $\mathbf{h}_0 = \text{vec}(\mathbf{H}_0)$  and  $\mathbf{R}_{ww} = (\sigma_R^2 |g|^2 + \sigma_D^2) \mathbf{I}_{ML}$  is the covariance matrix of the compound noise  $\mathbf{w}$ .

Given the LMMSE estimate of  $\mathbf{h}_2$ , the LMMSE estimator of  $\mathbf{h}_1$  can be similarly calculated as

$$\hat{\mathbf{h}}_1 = \mathbf{h}_0 + \left\{ \mathbf{R}_1 \tilde{\mathbf{H}}_2^H (\tilde{\mathbf{H}}_2 \mathbf{R}_1 \tilde{\mathbf{H}}_2^H + \hat{\mathbf{R}}_{ww})^{-1} (\mathbf{y} - \tilde{\mathbf{H}}_2 \mathbf{h}_0^T) \right\}^T, \quad (12)$$

where  $\tilde{\mathbf{H}}_2 = \mathbf{X}^T \otimes \hat{\mathbf{h}}_2 g$  and  $\hat{\mathbf{R}}_{ww} = \sigma_R^2 (\mathbf{I}_L \otimes \hat{\mathbf{h}}_2 |g|^2 \hat{\mathbf{h}}_2^H) + \sigma_D^2 \mathbf{I}_{ML}$ .

## B. EM-based MAP Channel Estimation

Alternatively, one may choose to employ the MAP estimator which is obtained by minimizing the Bayes risk for a ‘‘hit-or-miss cost function’’ [10]. The MAP estimator is obtained by maximizing the posterior density

$$\{\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2\} = \arg \max_{\mathbf{h}_1, \mathbf{h}_2} \left\{ \log p(\mathbf{y}|\mathbf{h}_1, \mathbf{h}_2) + \log p(\mathbf{h}_1, \mathbf{h}_2) \right\}. \quad (13)$$

Straightforward calculations show that the MAP estimates are the solution to the maximization problem

$$\{\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2\} = \arg \min_{\mathbf{h}_1, \mathbf{h}_2} \left\{ (\mathbf{y} - \tilde{\mathbf{H}}_2 \mathbf{h}_1^T)^H \hat{\mathbf{R}}_{ww}^{-1} (\mathbf{y} - \tilde{\mathbf{H}}_2 \mathbf{h}_1^T) + \log \det\{\hat{\mathbf{R}}_{ww}\} + \|\mathbf{h}_2\|^2 + (K+1)\|\mathbf{h}_1 - \mathbf{h}_0\|^2 \right\}. \quad (14)$$

The optimization problem in (14) is nonlinear and nonconvex. In other words, closed-form solutions for the optimal MAP estimates  $\hat{\mathbf{h}}_1$  and  $\hat{\mathbf{h}}_2$  cannot be derived. Therefore, we adopt the EM algorithm for solving the MAP estimation problem. The EM algorithm is an efficient iterative technique for finding the MAP estimates especially when the model depends on unobserved latent data.

The algorithm consists of two main steps: an expectation step followed by a maximization step. During the expectation step, the algorithm computes the expectation of the posterior density. The expectation is taken with respect to  $\mathbf{h}_1$  (treated as the missing data), conditioned on the current estimate of  $\mathbf{h}_2$  and the received data  $\mathbf{y}$

$$Q(\mathbf{h}_2, \hat{\mathbf{h}}_2^k) = E\{\log p(\mathbf{y}, \mathbf{h}_1, \mathbf{h}_2|\mathbf{y}; \hat{\mathbf{h}}_2^k)\}. \quad (15)$$

During the maximization step, the updated MAP estimate of  $\mathbf{h}_2$  is obtained by maximizing  $Q(\mathbf{h}_2, \hat{\mathbf{h}}_2^k)$

$$\hat{\mathbf{h}}_2^{k+1} = \arg \max_{\mathbf{h}_2} \{Q(\mathbf{h}_2, \hat{\mathbf{h}}_2^k)\}. \quad (16)$$

After some algebraic manipulations, the expression for the re-estimate  $\hat{\mathbf{h}}_2^{k+1}$  can be obtained as follows

$$\hat{\mathbf{h}}_2^{k+1} = (\Phi_2 + \mathbf{I}_M)^{-1} \Phi_1^H \mathbf{y}, \quad (17)$$

where  $\Phi_1 = (E\{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k\} \mathbf{X})^T \otimes \Psi$ ,  $\Phi_2 = (\text{Tr}(\mathbf{C}_{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k}) + E\{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k\} (E\{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k\})^H) \otimes \Psi$  and  $\Psi = (\sigma_R^2 \hat{\mathbf{h}}_2^k |g|^2 (\hat{\mathbf{h}}_2^k)^H + \sigma_D^2 \mathbf{I}_{MN})^{-1}$ .

According to the Bayesian linear model given in (9), the conditional mean  $E\{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k\}$  is the LMMSE estimator of  $\mathbf{h}_1$  and can be evaluated from (12) by using  $\tilde{\mathbf{H}}_2 = \mathbf{X}^T \otimes \hat{\mathbf{h}}_2^k g$  and  $\hat{\mathbf{R}}_{ww} = \sigma_R^2 (\mathbf{I}_L \otimes \hat{\mathbf{h}}_2^k |g|^2 \hat{\mathbf{h}}_2^k)^H + \sigma_D^2 \mathbf{I}_{ML}$ . Then, the covariance  $\mathbf{C}_{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k}$  can be evaluated as

$$\mathbf{C}_{\mathbf{h}_1|\mathbf{y}; \hat{\mathbf{h}}_2^k} = \mathbf{R}_1 - \mathbf{R}_1 \tilde{\mathbf{H}}_2^H (\tilde{\mathbf{H}}_2 \mathbf{R}_1 \tilde{\mathbf{H}}_2^H + \hat{\mathbf{R}}_{ww})^{-1} \tilde{\mathbf{H}}_2 \mathbf{R}_1. \quad (18)$$

The choice of the starting point  $\hat{\mathbf{h}}_2^0$  for the initialization of the EM algorithm is very important since it affects the convergence point of the algorithm. Hence, we suggest to use the LMMSE estimates evaluated in (11) and (12) as the initial guess. In summary, after selecting the initial starting point, the EM algorithm iteratively refines the channel estimates until convergence.

## IV. EXTENSION TO MULTIPLE RELAYS

In this section we extend our analysis to the case of multiple relays ( $R > 1$ ) between the source and the destination. In this case, we propose to estimate the individual channels  $\mathbf{H}_1$  and  $\mathbf{H}_2$  via multiple steps. Since we include  $R > 1$  relaying nodes, there are  $R$  different channel pairs  $\mathbf{h}_{1,i}, \mathbf{h}_{2,i}, i = 1, \dots, R$ , to be estimated. The vectors  $\mathbf{h}_{1,i}$  and  $\mathbf{h}_{2,i}$  refer to the  $i$ th row of  $\mathbf{H}_1$  and the  $i$ th column of  $\mathbf{H}_2$ , respectively. As shown in [11], in order to obtain channel knowledge of the individual links from source to relays,  $\mathbf{H}_1$ , and from relays to destination,  $\mathbf{H}_2$ , the total channel training interval should be divided into  $R$  frames. For each such frame, we create different compound channels by varying the amplifying factors at the relays. Therefore, by varying the amplifying matrix  $\mathbf{G}$  within  $R$  consecutive time slots, we can establish  $R$  independent equations with respect to  $\mathbf{H}_1$  and  $\mathbf{H}_2$  that can guarantee the successful estimation of the channel pairs  $\mathbf{h}_{1,i}, \mathbf{h}_{2,i}, i = 1, \dots, R$ .

For every  $\mathbf{G}^{(i)}$ , a known training sequence  $\mathbf{X} \in \mathcal{C}^{N \times L}$ , of length  $N$  ( $L \geq N$ ) is transmitted. Again, we assume that the training sequence is orthogonal across all transmitting antennas. The output of the  $i$ th training block  $\mathbf{Y}^{(i)} = [\mathbf{y}_1, \dots, \mathbf{y}_L]$  can then be expressed as

$$\mathbf{Y}^{(i)} = \mathbf{H}^{(i)} \mathbf{X} + \mathbf{W}^{(i)}, \quad (19)$$

where  $\mathbf{W}^{(i)} = [\mathbf{w}_1, \dots, \mathbf{w}_L]$ . After vectorizing each term in (19), we obtain

$$\mathbf{y}^{(i)} = \tilde{\mathbf{X}} \mathbf{h}^{(i)} + \mathbf{w}^{(i)}, \quad (20)$$

where  $\tilde{\mathbf{X}} = \mathbf{X}^T \otimes \mathbf{I}_M$ ,  $\mathbf{h}^{(i)} = \text{vec}(\mathbf{H}_2 \mathbf{G}^{(i)} \mathbf{H}_1)$ , and  $\mathbf{w}^{(i)} = \text{vec}(\mathbf{W}^{(i)})$ . Then, the LMMSE estimator is employed for the estimation of the  $i$ th compound channel matrix  $\mathbf{h}^{(i)}$

$$\hat{\mathbf{h}}^{(i)} = \mathbf{R}_{hh}^{(i)} \tilde{\mathbf{X}}^H (\tilde{\mathbf{X}} \mathbf{R}_{hh}^{(i)} \tilde{\mathbf{X}}^H + (\mathbf{R}_{ww}^{(i)})^{-1})^{-1} \mathbf{y}^{(i)}, \quad (21)$$

where  $\mathbf{R}_{hh}$  and  $\mathbf{R}_{ww}$  denote the covariance matrices of the compound channel and the compound noise, respectively, and can be calculated as

$$\mathbf{R}_{hh}^{(i)} = \mathbf{Q}_1 \otimes \mathbf{I}_M, \quad (22)$$

$$\mathbf{R}_{ww}^{(i)} = (\sigma_R^2 \text{Tr}\{\mathbf{G}^{(i)} (\mathbf{G}^{(i)})^H\} + \sigma_D^2) \mathbf{I}_{ML}, \quad (23)$$

where  $\mathbf{Q}_1 = \frac{K}{K+1} (\mathbf{H}_0^H \text{Tr}\{\mathbf{G}^{(i)} (\mathbf{G}^{(i)})^H \mathbf{G}^{(i)}\} \mathbf{H}_0)^T + \frac{1}{K+1} \text{Tr}\{\mathbf{G}^{(i)} (\mathbf{G}^{(i)})^H \mathbf{G}^{(i)}\} \mathbf{I}_N$ . The performance of the estimator is measured by the estimation error  $\epsilon^{(i)} = \hat{\mathbf{h}}^{(i)} - \mathbf{h}^{(i)}$ , whose mean is zero and whose covariance matrix is [10]

$$\mathbf{C}_{\epsilon\epsilon}^{(i)} = \left( (\mathbf{R}_{hh}^{(i)})^{-1} + \tilde{\mathbf{X}}^H (\mathbf{R}_{ww}^{(i)})^{-1} \tilde{\mathbf{X}} \right)^{-1}. \quad (24)$$

After calculating  $R$  vector estimates of the compound channel by using (21) and aligning the estimated vectors along the columns of a new matrix  $\tilde{\mathbf{H}}$ , we obtain

$$\tilde{\mathbf{H}} = \Xi \tilde{\mathbf{G}} + \mathbf{E}, \quad (25)$$

where  $\tilde{\mathbf{H}} = [\hat{\mathbf{h}}^{(1)}, \dots, \hat{\mathbf{h}}^{(R)}]$ ,  $\Xi = \mathbf{H}_1^T \diamond \mathbf{H}_2$ , and  $\tilde{\mathbf{G}} = [\text{vecd}(\mathbf{G}^{(1)}), \dots, \text{vecd}(\mathbf{G}^{(R)})]$ ,  $\mathbf{E} = [\epsilon^{(1)}, \dots, \epsilon^{(R)}]$ ; thus  $\tilde{\mathbf{H}} \in \mathcal{C}^{MN \times R}$ ,  $\Xi \in \mathcal{C}^{MN \times R}$ , and  $\tilde{\mathbf{G}} \in \mathcal{C}^{R \times R}$ . Since the relay amplification factors can be designed a priori, one can guarantee that the matrix  $\tilde{\mathbf{G}}$  is full rank. For instance, the amplifying

factors during the estimation phase can be generated by using the DFT matrix, scaled in such a way that the relay transmit power constraint is satisfied. One could also choose  $\tilde{\mathbf{G}} = \mathbf{I}_R$ , which in turn implies that during the  $i$ th estimation interval all but the  $i$ th relaying node are switched off.

The previous relation (25) suggests an LMMSE estimate of  $\Xi$  [9] as

$$\hat{\Xi} = \tilde{\mathbf{H}}(\tilde{\mathbf{G}}^H \mathbf{R}_{\Xi} \tilde{\mathbf{G}} + \mathbf{C}_E)^{-1} \tilde{\mathbf{G}}^H \mathbf{R}_{\Xi}. \quad (26)$$

The correlation matrices  $\mathbf{R}_{\Xi}$  and  $\mathbf{C}_E$  can be calculated as

$$\mathbf{R}_{\Xi} = E\{\Xi^H \Xi\} = \mathbf{J}^H (\mathbf{Q}_2 \otimes \mathbf{I}_M) \mathbf{J}, \quad (27)$$

$$\mathbf{C}_E = E\{\mathbf{E}^H \mathbf{E}\} = \text{diag}\left(\text{Tr}(\mathbf{C}_{\varepsilon\varepsilon}^{(1)}), \dots, \text{Tr}(\mathbf{C}_{\varepsilon\varepsilon}^{(R)})\right), \quad (28)$$

where  $\mathbf{J} = [\mathbf{e}_1 \dots \mathbf{e}_R]$  is a selection matrix of order  $R^2 \times R$ ,  $\mathbf{e}_k$  is a  $R^2 \times 1$  column vector with a unity element in the  $k$ th position and zeros elsewhere, and  $\mathbf{Q}_2 = \frac{K}{K+1}(\mathbf{H}_0^H \mathbf{H}_0)^T + \frac{1}{K+1} \mathbf{I}_N$ .

The Khatri-Rao product in (25) is the column-wise Kronecker product between the channel matrices  $\mathbf{H}_1^T$  and  $\mathbf{H}_2$  and can be inverted only up to a scalar ambiguity per column. Let  $\hat{\xi}_i$  denote the  $i$ th column of the estimate  $\hat{\Xi}$ . This can be expressed as

$$\hat{\xi}_i = \mathbf{h}_{1,i}^T \otimes \mathbf{h}_{2,i} + \varepsilon_{\xi}^{(i)}, \quad (29)$$

where  $\varepsilon_{\xi}^{(i)}$  is the  $i$ th column of the error matrix  $\mathbf{E}_{\Xi} = \Xi - \hat{\Xi}$  with zero mean and correlation matrix  $\mathbf{C}_{\Xi\xi} = E\{\mathbf{E}_{\Xi}^H \mathbf{E}_{\Xi}\} = \{\mathbf{R}_{\Xi}^{-1} + \tilde{\mathbf{G}} \mathbf{C}_E^{-1} \tilde{\mathbf{G}}^H\}^{-1}$  [9].

The relation in (29) is very similar to (8) and (9), that describe the vector form of the output data after the training block for the single relay case. Therefore, each channel pair  $\mathbf{h}_{1,i}$ ,  $\mathbf{h}_{2,i}$ , can be independently estimated by applying the proposed LMMSE and EM-based algorithms developed in the previous sections for the single relay case, for  $i = 1, \dots, R$ .

In particular, the LMMSE estimate of  $\mathbf{h}_{2,i}$  is first computed as

$$\hat{\mathbf{h}}_{2,i} = \tilde{\mathbf{H}}_0^H (\tilde{\mathbf{H}}_0 \tilde{\mathbf{H}}_0^H + \mathbf{R}_1 \otimes \mathbf{I}_M + \mathbf{C}_{\xi\xi}^{(i)})^{-1} \hat{\xi}_i, \quad (30)$$

where  $\tilde{\mathbf{H}}_0 = E\{\mathbf{h}_{1,i}^T \otimes \mathbf{I}_M\} = \sqrt{\frac{K}{K+1}} \mathbf{h}_{0,i}^T \otimes \mathbf{I}_M$ ,  $\mathbf{h}_{0,i}$  is the  $i$ th row of  $\mathbf{H}_0$ , and  $\mathbf{C}_{\xi\xi}^{(i)} = E\{\varepsilon_{\xi}^{(i)} (\varepsilon_{\xi}^{(i)})^H\}$ .

Given the LMMSE estimate of  $\mathbf{h}_{2,i}$ , the LMMSE estimator of  $\mathbf{h}_{1,i}$  can be similarly calculated as

$$\hat{\mathbf{h}}_{1,i} = \mathbf{h}_{0,i} + \left\{ \mathbf{R}_1 \tilde{\mathbf{H}}_2^H (\tilde{\mathbf{H}}_2 \mathbf{R}_1 \tilde{\mathbf{H}}_2^H + \mathbf{C}_{\xi\xi}^{(i)})^{-1} (\hat{\xi}_i - \tilde{\mathbf{H}}_2 \mathbf{h}_{0,i}^T) \right\}^T, \quad (31)$$

where  $\tilde{\mathbf{H}}_2 = \mathbf{I}_N \otimes \hat{\mathbf{h}}_{2,i}$ . The previous expressions in (30) and (31) are very similar to the LMMSE estimates in (11) and (12), respectively, that were obtained for the single relay case.

Finally, after calculating the initial LMMSE estimates, the proposed EM-based MAP algorithm can be applied for iteratively refining the channel estimates for each channel pair  $\mathbf{h}_{2,i}$  and  $\mathbf{h}_{1,i}^T$ , for  $i = 1, \dots, R$ .

**Remark 1:** Note that  $\mathbf{H}_1$  and  $\mathbf{H}_2$  can be determined only up to scaling, no matter how many different training intervals are utilized. We can prove that  $\hat{\mathbf{H}}_1 \rightarrow \mathbf{D}^{-1} \mathbf{H}_1$  and  $\hat{\mathbf{H}}_2 \rightarrow \mathbf{H}_2 \mathbf{D}$  as either  $\text{SNR} \rightarrow \infty$  or  $L \rightarrow \infty$ , where  $\mathbf{D}$  is an arbitrary diagonal matrix. Fortunately, this ambiguity causes no problems when

optimizing the relay channel because the effect of the  $i$ th relay is seen only via the product  $\mathbf{h}_{2,i} g_i \mathbf{h}_{1,i}^T$ . This becomes clear if we rewrite the compound channel in (4) as

$$\mathbf{H} = \sum_{i=1}^R g_i \mathbf{H}_i, \quad (32)$$

where  $\mathbf{H}_i = \mathbf{h}_{2,i} \mathbf{h}_{1,i}^T$ . Thus, only knowledge of the matrices  $\{\mathbf{H}_i\}_{i=1}^R$  is necessary for determining the impact of a certain choice of gain factors  $\{g_i\}_{i=1}^R$ .

## V. PERFORMANCE EVALUATION

In this section, we present some numerical results to illustrate the performance of our proposed algorithm. In our simulations a  $2 \times 2$  MIMO system with 2 AF relaying nodes is considered ( $N = M = R = 2$ ). In particular, we investigate an asymmetric fading scenario where  $\mathbf{H}_1$  is a Rician fading channel with mean  $\frac{K}{K+1} \mathbf{H}_0$  and covariance  $\mathbf{R}_1 = \frac{1}{K+1} \mathbf{I}_{RN}$ , and  $\mathbf{H}_2$ , is a Rayleigh fading channel containing i.i.d. zero-mean unit-variance complex Gaussian entries. The training sequence  $\mathbf{X} \in \mathcal{C}^{N \times L}$  in (19) is obtained from the first  $N$  rows of a DFT matrix of size  $L \times L$ . We use the smallest possible value for the training length  $L$ , that is  $L = N = 2$ . The relay amplifying factors during the estimation phase are generated by using the DFT matrix, scaled in such a way that the relay transmit power constraint is satisfied. We use a fixed transmit power of 0 dB at the source and the relays and vary the noise power at the relays and the destination. For simplicity, we assume that the noise power is the same at the relays and the destination, i.e.  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ .

To evaluate the channel estimation accuracy we compute the average mean square error (MSE); due to the ambiguity of the model (see Remark 1), however, some care must be taken in this averaging. If the first column of  $\mathbf{H}_1$  is normalized, in effect by dividing by the corresponding elements of  $\hat{\mathbf{H}}_1$ , then the average mean-square error matrix may not exist. This is certainly the case when the elements of  $\mathbf{H}_1$  are zero-mean Gaussian distributed, since the elements of the first column of  $\mathbf{H}_1$  will then occasionally take on very small values. However, for the scenario under investigation, we can avoid this somewhat artificial situation, since  $\mathbf{H}_1$  is a Ricean fading channel with non-zero mean. Then, the normalized MSE for the first channel matrix  $\mathbf{H}_1$  can be computed as

$$\text{NMSE}_1 = \frac{\|\mathbf{H}_1 - \mathbf{D} \hat{\mathbf{H}}_1\|_{\text{F}}^2}{\|\mathbf{H}_1\|_{\text{F}}^2}, \quad (33)$$

where  $\mathbf{D}$  is a diagonal matrix. The  $i$ th diagonal element of  $\mathbf{D}$  is given by

$$\mathbf{D}(i, i) = \frac{\mathbf{H}_{1,1}(i)}{\hat{\mathbf{H}}_{1,1}(i)}, \quad (34)$$

where  $\mathbf{H}_{1,1}(i)$  is the  $i$ th element of the first column of  $\mathbf{H}_1$  and  $\hat{\mathbf{H}}_{1,1}(i)$  is the  $i$ th element of the first column of  $\hat{\mathbf{H}}_1$ . In a similar way, we can define the normalized MSE for  $\mathbf{H}_2$

$$\text{NMSE}_2 = \frac{\|\mathbf{H}_2 - \hat{\mathbf{H}}_2 \mathbf{D}^{-1}\|_{\text{F}}^2}{\|\mathbf{H}_2\|_{\text{F}}^2}. \quad (35)$$

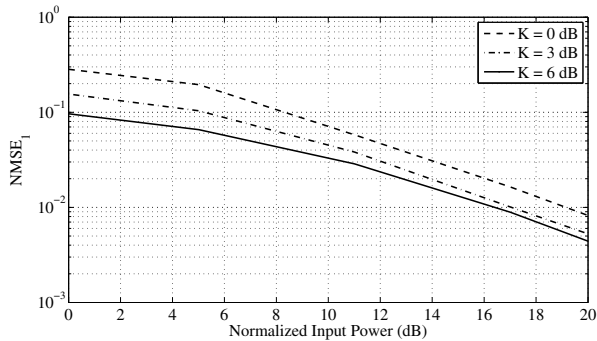


Fig. 2. Channel estimation performance for  $\mathbf{H}_1$  for a  $2 \times 2$  MIMO channel with 2 AF relaying nodes for  $K = 0, 3, 6$  dB.

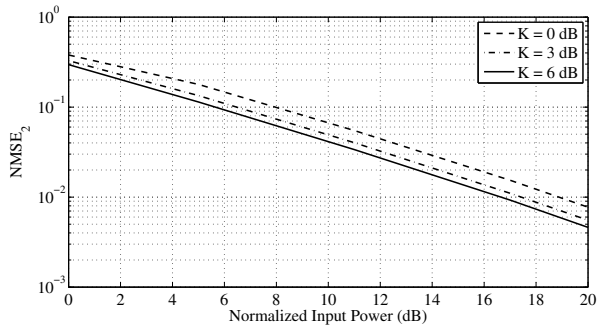


Fig. 3. Channel estimation performance for  $\mathbf{H}_2$  for a  $2 \times 2$  MIMO channel with 2 AF relaying nodes for  $K = 0, 3, 6$  dB.

Figures 2 and 3 show the NMSE for the estimation of  $\mathbf{H}_1$  and  $\mathbf{H}_2$ , respectively, as a function of the input power normalized to the noise power during the training interval for different values of the Rician  $K$ -factor. As expected, the estimation performance is significantly improved for higher values of  $K$ , especially in the low SNR regime. This is a well-known property of Bayesian channel estimators since they utilize any available a priori information about the unknown parameters to be estimated. The higher the value of the  $K$ -factor is, the stronger the impact of the available prior knowledge about  $\mathbf{H}_1$  becomes. Figure 4 shows the NMSE curves of the estimated channel  $\mathbf{H}_2$  for  $K = 0$  dB. The application of the EM-based MAP algorithm improves the estimation accuracy by iteratively refining the channel estimates. The algorithm is numerically stable and convergence is achieved even after 2 iterations. Finally, we point out that, as a starting point for the EM-based MAP algorithm, we use the LMMSE estimates of  $\mathbf{h}_{2,i}$  that are given by (30).

## VI. CONCLUSION

In this paper, we have investigated the problem of channel estimation in MIMO AF relaying systems with random channels. We were particularly interested in jointly estimating the individual channels, from source to relays and from relays to destination, by using only measurements at the destination node. Under the Bayesian framework, LMMSE and EM-based MAP algorithms were developed for estimating the individual channel responses. The clear advantage of the proposed estimation schemes is that they are applicable to

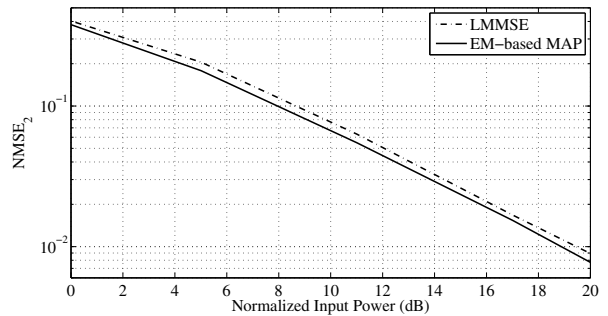


Fig. 4. Channel estimation performance for  $\mathbf{H}_2$  for a  $2 \times 2$  MIMO channel with 2 AF relaying nodes for  $K = 0$  dB.

minimum-complexity relays that simply amplify and forward their received signal. Our theoretical results were supported by simulations for illustrating the effectiveness of the proposed algorithms. Specifically, the performance of the proposed algorithms was evaluated in terms of MSE as a function of the SNR during the training interval. We demonstrated that by incorporating the available prior knowledge of the channel statistics into our estimation algorithms the MSE is reduced, especially in the low SNR regime. The application of the EM-based MAP algorithm improves the estimation accuracy by iteratively refining the channel estimates.

## ACKNOWLEDGMENTS

This work has been supported in part by VINNOVA within the VINN Excellence Center Chase, and in part by SSF within the Strategic Research Center Charmant.

## REFERENCES

- [1] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [2] S. Jin, M. McKay, C. Zhong, and K.-K. Wong, "Ergodic capacity analysis of amplify-and-forward mimo dual-hop systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2204–2224, May 2010.
- [3] 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.
- [4] O. Munoz-Medina, J. Vidal, and A. Agustin, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2593–2604, Jun. 2007.
- [5] 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.
- [6] A. Wolfgang, M. Coldrey, and P. Bohlin, "Advanced repeaters in cellular communication systems," in *Proc. European Conference on Antennas and Propagation (EuCap)*, Berlin, Germany, Mar. 2009, pp. 1504–1508.
- [7] T. Kong and Y. Hua, "Optimal channel estimation and training design for MIMO relays," in *Proc. IEEE Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, Nov. 2010, pp. 663–667.
- [8] H. A. Suraweera, G. K. Karagiannidis, and P. J. Smith, "Performance analysis of the dual-hop asymmetric fading channel," *IEEE Trans. Commun.*, vol. 8, no. 6, pp. 2783–2788, Jun. 2009.
- [9] M. Biguesh and A. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
- [10] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall PTR, 1993.
- [11] P. Lioliou, M. Viberg, and M. Coldrey, "Performance analysis of relay channel estimation," in *Proc. IEEE Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, USA, Nov. 2009, pp. 1533–1537.