

# OPTIMAL TRAINING DESIGN AND PLACEMENT FOR CHANNEL ESTIMATION IN COOPERATIVE NETWORKS

Christos Mavrokefalidis<sup>†</sup>, Athanasios A. Rontogiannis\*, Kostas Berberidis<sup>†</sup>

<sup>†</sup>Dept. of Computer engineering and Informatics, University of Patras, Greece

Emails: {maurokef, berberid}@ceid.upatras.gr

\*Institute for Space Applications and Remote Sensing, National Observatory of Athens, Greece

Email: tronto@space.noa.gr

## ABSTRACT

In this paper, optimal training design and placement are studied for channel estimation in a frequency selective, single relay, amplify and forward cooperative system. Using a two-phase transmission protocol, the power distribution between the phases, among the pilot tones as well as the positioning of pilot tones within the transmitted OFDM symbols are examined. The LMMSE and two variants of the LS estimator are analyzed and the optimal training design and placement strategies are derived. Closed-form expressions are derived for the minimum achievable MSE in each case, which are precisely verified through simulations.

## 1. INTRODUCTION

The idea of cooperation has been introduced in communication systems in order to obtain spatial diversity gains, while only single-antenna terminals are used ([1]). In such systems, a relay node assists the communication between a transmitter and a receiver providing an additional transmission path. Most of the methods that have been proposed so far, can be divided into two categories depending on the procedure followed at the relay. Specifically, in Decode and Forward (DF) methods, the relay decodes, re-encodes and forwards the received signal while in Amplify and Forward (AF) methods, the relay multiplies the received signal with a constant and then forwards it.

In most of these methods, the involved channels are assumed to be known. Thus, in practice, a channel estimation procedure is required to provide the desired information. For flat fading channels, [2] studies channel estimation in the AF case assuming a single relay scenario. Multiple relays are considered in [3], [4] for the AF and DF cases, respectively. For frequency selective channels, a minimum variance unbiased (MVU) estimator for an AF single relay system is described in [5]. In [6], a linear minimum mean square

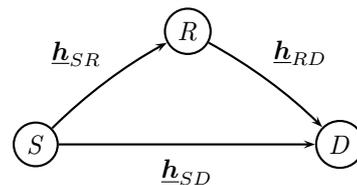


Fig. 1. The cooperative system

error (LMMSE) channel estimator is proposed assuming an AF multiple relay scenario. In [7], channel estimation in a two-way relay system is studied. Finally, in [8], the authors propose a single-carrier frequency-domain semi-blind channel estimation technique (as opposed to the aforementioned training-based ones) for a multiple relay network.

In this paper, we consider a frequency-selective single relay network, where transmission is realized through orthogonal frequency division multiplexing (OFDM). A two-phase transmission protocol is adopted [9] and the minimum required number of tones in each OFDM symbol is assigned to training. Assuming that a total power  $P$  is devoted to training at the source node, we investigate the optimal power allocation between the two phases of transmission and among the training symbols in each phase, as well as the optimal placement of pilots in the OFDM symbols. Both the LMMSE and the least squares (LS) channel estimators are studied. For the latter, two distinct cases, namely a deterministic and a stochastic, are analyzed separately. The optimality conditions for all three scenarios are derived and expressions for the minimum attainable mean squared error (MSE) are provided. The presented theoretical results are completely verified by simulations.

The works mostly related to the current paper are [5] and [7]. More specifically, the stochastic LS estimator mentioned above, turns out to be the MVU estimator of [5]. However, in [5] only the optimal training power allocation is studied, while the positioning of pilots is not investigated. Optimal pilot design and placement is also considered in [7]. However, the cooperative setting there is different.

This work was supported by the 03ED838 research project, implemented within the framework of the "Reinforcement Programme of Human Research Manpower" (PENED) and co-financed by National and EC Funds.

In the following, bold underlined small and capital letters denote vectors at the time and frequency domain, respectively. Also, bold capital letters are used for matrices.  $\mathbf{F}$  denotes the  $N \times N$  Fourier matrix whose  $(p, q)$ -th element is given by  $[\mathbf{F}]_{p,q} = \frac{1}{\sqrt{N}} e^{-j2\pi pq/N}$ .  $\mathbf{A}^T$  and  $\mathbf{A}^H$  denote transposition and conjugate transposition of  $\mathbf{A}$ . Also,  $\text{diag}\{\underline{\mathbf{a}}\}$  produces a diagonal matrix with  $\underline{\mathbf{a}}$  on its main diagonal and  $\text{vect}\{\mathbf{A}\}$  produces a vector whose elements are the diagonal elements of  $\mathbf{A}$ .  $\mathbf{I}_N$  is the identity matrix of size  $N$ ,  $\text{Tr}\{\mathbf{A}\}$  is the trace of  $\mathbf{A}$ ,  $\|\cdot\|$  is the 2-norm of a vector,  $|\cdot|$  the absolute value of a scalar and  $\mathcal{E}\{\cdot\}$  denotes statistical expectation. Finally,  $\underline{\mathbf{x}} \sim CN(\underline{\boldsymbol{\mu}}, \boldsymbol{\Sigma})$  denotes a complex Gaussian random vector with mean  $\underline{\boldsymbol{\mu}}$  and covariance matrix  $\boldsymbol{\Sigma}$ .

In Section 2, the system model is presented. In Sections 3, 4, the LMMSE and the LS channel estimators along with their corresponding performance optimization are described, respectively. In Section 5, simulation results are presented and Section 6 concludes the paper.

## 2. SYSTEM MODEL

We consider the cooperative system shown in Fig. 1. A two-phase transmission protocol proposed in [9] is adopted. As shown in [10], this protocol offers the optimal diversity / multiplexing trade-off among the AF half-duplex protocols. Specifically, at the first phase,  $S$  broadcasts a signal to  $R$  and  $D$ . At the second phase,  $S$  sends a new signal and, concurrently,  $R$  amplifies and forwards to  $D$  the signal received at the previous phase. The transmissions are performed in blocks of  $N$  symbols by utilizing OFDM and  $S$ ,  $R$  are assumed to be synchronized. To avoid interblock interference, a cyclic prefix (CP) of appropriate length is appended to the transmitted signals ([11]). The frequency selective channels  $\underline{\mathbf{h}}_{SD}$ ,  $\underline{\mathbf{h}}_{SR}$  and  $\underline{\mathbf{h}}_{RD}$  are modeled as vectors of lengths  $L_{SD}$ ,  $L_{SR}$  and  $L_{RD}$ , respectively.

At *Phase I*, after discarding the CP, the received signal vectors at  $R$  and  $D$  are given by

$$\underline{\mathbf{y}}_R = \mathbf{H}_{SR}\underline{\mathbf{x}}_1 + \underline{\mathbf{w}}_R, \quad (1)$$

$$\underline{\mathbf{y}}_1 = \mathbf{H}_{SD}\underline{\mathbf{x}}_1 + \underline{\mathbf{w}}_1, \quad (2)$$

respectively. Similarly, at *Phase II*, the received signal vector at  $D$  is expressed as

$$\underline{\mathbf{y}}_2 = \mathbf{H}_{SD}\underline{\mathbf{x}}_2 + \alpha\mathbf{H}_{RD}\underline{\mathbf{y}}_R + \underline{\mathbf{w}}_2. \quad (3)$$

In (1)-(3), the  $N \times N$  matrices  $\mathbf{H}_i$  are circulant, having as first columns the vectors  $[\underline{\mathbf{h}}_i^T \mathbf{0} \dots \mathbf{0}]^T$ , where  $i \in \{SD, SR, RD\}$ . Also,  $\underline{\mathbf{x}}_i = \mathbf{F}^H \mathbf{X}_i$ ,  $i = 1, 2$ , are the signal vectors transmitted by the source in each phase and  $\mathbf{X}_i$  denote the corresponding  $N \times 1$  information symbol blocks. The relay terminal normalizes the received vector  $\underline{\mathbf{y}}_R$  by a fixed gain factor  $\alpha$  ([2]), so that the average power per symbol remains constant. Finally, we assume that  $\underline{\mathbf{w}}_R \sim CN(\mathbf{0}, \sigma_R^2 \mathbf{I}_N)$ ,  $\underline{\mathbf{w}}_1$ ,

$\underline{\mathbf{w}}_2 \sim CN(\mathbf{0}, \sigma_D^2 \mathbf{I}_N)$ , where  $\sigma_R^2$ ,  $\sigma_D^2$  denote the noise variances at  $R$  and  $D$ , respectively.

Using the well-known identity  $\mathbf{H}_j = \mathbf{F}^H \boldsymbol{\Lambda}_j \mathbf{F}$  for circulant matrices, where  $j \in \{SD, RD, SR\}$  and  $\boldsymbol{\Lambda}_j$  are diagonal matrices, it can be easily shown from (1) and (2) that the frequency domain vectors  $\underline{\mathbf{Y}}_i = \mathbf{F} \underline{\mathbf{y}}_i$ ,  $i = 1, 2$  can be expressed as

$$\begin{bmatrix} \underline{\mathbf{Y}}_1 \\ \underline{\mathbf{Y}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{X}_1 \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{\Lambda}}_{SD} \\ \underline{\boldsymbol{\Lambda}}_R \end{bmatrix} + \begin{bmatrix} \underline{\mathbf{W}}_1 \\ \underline{\mathbf{W}}_2 \end{bmatrix} \quad (4)$$

In the last equation,  $\mathbf{X}_i = \text{diag}\{\underline{\mathbf{X}}_i\}$ ,  $\underline{\boldsymbol{\Lambda}}_{SD} = \text{vect}\{\boldsymbol{\Lambda}_{SD}\}$  and  $\underline{\boldsymbol{\Lambda}}_R = \text{vect}\{\boldsymbol{\Lambda}_R\}$  where  $\boldsymbol{\Lambda}_R = \alpha \boldsymbol{\Lambda}_{SR} \boldsymbol{\Lambda}_{RD}$ . It can be also verified that,  $\underline{\mathbf{W}}_1 \sim CN\{\mathbf{0}, \sigma_D^2 \mathbf{I}_N\}$  and  $\underline{\mathbf{W}}_2$  has zero mean and  $\mathbf{C} = \sigma_D^2 \mathbf{I}_N + \sigma_R^2 |\alpha|^2 \mathcal{E}\{\boldsymbol{\Lambda}_{RD}^H \boldsymbol{\Lambda}_{RD}\}$  as a covariance matrix. It should be noted that  $\underline{\boldsymbol{\Lambda}}_{SD}$  in (4) contains the frequency response of the source-to-destination channel in  $N$  frequency tones. Similarly,  $\underline{\boldsymbol{\Lambda}}_R$  contains the frequency response of the overall source-relay-destination channel expressed in the time-domain as  $\underline{\mathbf{h}}_R = \alpha \underline{\mathbf{h}}_{SR} * \underline{\mathbf{h}}_{RD}$ , where  $*$  denotes convolution.

In the following, we assume without loss of generality that  $L_{SD} = L_{SR} + L_{RD} - 1 = L$  and a minimum number of  $L$  tones in each  $\underline{\mathbf{X}}_i$  are devoted to training. If  $\underline{\boldsymbol{\Lambda}}_{i,L}$  stands for the vector formed by any  $L$  randomly selected elements of  $\underline{\boldsymbol{\Lambda}}_i$ ,  $i \in \{SD, R\}$ , then  $\underline{\boldsymbol{\Lambda}}_{i,L} = \sqrt{N} \mathbf{F}_L \underline{\mathbf{h}}_i$  ([12]). The  $L \times L$  matrix  $\mathbf{F}_L$  is produced by the corresponding  $L$  rows of  $\mathbf{F}$  and its first  $L$  columns. Then, from Eq. (4) we get

$$\underline{\mathbf{Y}} = \mathbf{A} \underline{\mathbf{h}} + \underline{\mathbf{W}} \quad (5)$$

where  $\underline{\mathbf{Y}} = [\underline{\mathbf{Y}}_{1,L}^T \underline{\mathbf{Y}}_{2,L}^T]^T$ ,  $\underline{\mathbf{W}} = [\underline{\mathbf{W}}_{1,L}^T \underline{\mathbf{W}}_{2,L}^T]^T$ ,  $\underline{\mathbf{h}} = [\underline{\mathbf{h}}_{SD}^T \underline{\mathbf{h}}_R^T]^T$  and

$$\mathbf{A} = \sqrt{N} \begin{bmatrix} \mathbf{X}_{1,L} & \mathbf{0} \\ \mathbf{X}_{2,L} & \mathbf{X}_{1,L} \end{bmatrix} \begin{bmatrix} \mathbf{F}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_L \end{bmatrix} = \sqrt{N} \mathbf{X} \mathbf{F}_d. \quad (6)$$

Also,  $\underline{\mathbf{W}}_{1,L} \sim CN\{\mathbf{0}, \sigma_D^2 \mathbf{I}_L\}$  and  $\underline{\mathbf{W}}_{2,L}$  is zero mean with covariance matrix  $\mathbf{C}_L = \sigma_D^2 \mathbf{I}_L + \sigma_R^2 |\alpha|^2 \mathcal{E}\{\boldsymbol{\Lambda}_{RD,L}^H \boldsymbol{\Lambda}_{RD,L}\}$ . Note that the meaning of subscript  $L$  in the above expressions is that  $L$  elements or rows have been retained from the respective vectors and matrices.

In the next section, we present optimal design and placement of pilots tones, so that the mean squared error (MSE) of the LMMSE and LS estimators of  $\underline{\mathbf{h}}$  in (5) is minimized. We assume that the total training power for both phases is  $P$  and that  $P_1 = \sum_{n=1}^L p_n$  and  $P_2 = \sum_{n=1}^L q_n$  are the power distributions between the two phases with  $P_1 + P_2 = P$ . Finally,  $\underline{\mathbf{p}} = [p_1, \dots, p_L]^T$  and  $\underline{\mathbf{q}} = [q_1, \dots, q_L]^T$  is the powers assigned to the training symbols (elements of  $\mathbf{X}_{1,L}$  and  $\mathbf{X}_{2,L}$  in (6)) in each phase.

## 3. LMMSE ESTIMATION

In this case, we assume that  $\underline{\mathbf{h}}_i \sim CN(\mathbf{0}_{L_i}, \text{diag}\{\underline{\boldsymbol{\sigma}}_i\})$  where  $\underline{\boldsymbol{\sigma}}_i = [\sigma_i^2(1), \dots, \sigma_i^2(L_i)]^T$ ,  $\sigma_i^2(n)$  is the variance of  $h_i(n)$

with  $i \in \{SD, SR, RD\}$  and  $n = 1, \dots, L_i$ .

### 3.1. Channel estimation

The LMMSE estimator of  $\mathbf{h}$  in (5) is given by ([13])

$$\hat{\mathbf{h}} = \mathbf{C}_{hh} \mathbf{A}^H (\mathbf{A} \mathbf{C}_{hh} \mathbf{A}^H + \mathbf{C}_{WW})^{-1} \mathbf{Y}, \quad (7)$$

where  $\mathbf{C}_{hh} = \mathcal{E}\{\mathbf{h}\mathbf{h}^H\}$  and  $\mathbf{C}_{WW} = \mathcal{E}\{\mathbf{W}\mathbf{W}^H\}$ . Its error covariance matrix is expressed as

$$\mathbf{C}_e = (\mathbf{C}_{hh}^{-1} + \mathbf{A}^H \mathbf{C}_{WW}^{-1} \mathbf{A})^{-1}. \quad (8)$$

The covariance matrices in (7) and (8) can be written as

$$\mathbf{C}_{hh} = \begin{bmatrix} \text{diag}\{\underline{\sigma}_{SD}\} & \mathbf{0} \\ \mathbf{0} & |\alpha|^2 \text{diag}\{\underline{\sigma}_{SR} * \underline{\sigma}_{RD}\} \end{bmatrix}, \quad (9)$$

and

$$\mathbf{C}_{WW} = \begin{bmatrix} \sigma_D^2 \mathbf{I}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_L \end{bmatrix}. \quad (10)$$

Moreover, it can be shown that  $\mathcal{E}\{|\Lambda_{RD,L}(k)|^2\} = \sigma^2$  and  $\sigma^2 = \sum_{n=1}^{L_{RD}} \sigma_{RD}^2(n)$  for every diagonal element  $k$  of  $\Lambda_{RD,L}$  ([13, p. 535]) and, thus,  $\mathbf{C}_L = (\sigma_D^2 + |\alpha|^2 \sigma_R^2 \sigma^2) \mathbf{I}_L$ .

### 3.2. Training design

To optimize the performance of the LMMSE estimator the following minimization problem must be solved

$$\begin{aligned} \min_{P_1, P_2, \underline{p}, \underline{q}, \{i_k\}} \quad & \left\{ \frac{1}{2L} \text{Tr}\{\mathbf{C}_e\} \right\} \\ \text{s.t.} \quad & P_1 + P_2 = P \end{aligned} \quad (11)$$

where  $\{i_k, k = 0, 1, \dots, L\}$  is the set of the  $L$  pilot positions in  $\mathbf{X}_1, \mathbf{X}_2$ , in which pilot symbols will be placed.

It is known that for a  $K \times K$  positive definite matrix  $\mathbf{B}$ ,  $\text{Tr}\{\mathbf{B}^{-1}\} \geq \sum_i [\mathbf{B}]_{i,i}^{-1}$ , with equality when  $\mathbf{B}$  is diagonal. Hence, from (11)

$$\frac{\text{Tr}\{\mathbf{C}_e\}}{2L} \geq \frac{\sum_{i=1}^{2L} [\mathbf{C}_{hh}^{-1} + \mathbf{A}^H \mathbf{C}_{WW}^{-1} \mathbf{A}]_{ii}^{-1}}{2L} = f(P_1, P_2), \quad (12)$$

where

$$f(P_1, P_2) = \sum_{i=1}^L \frac{(2L)^{-1}}{c_i + \sigma_D^{-2} P_1 + u^{-1} P_2} + \sum_{i=L+1}^{2L} \frac{(2L)^{-1}}{c_i + u^{-1} P_1} \quad (13)$$

and  $u = \sigma_D^2 + |\alpha|^2 \sigma_R^2 \sigma^2$ ,  $c_i = [\mathbf{C}_{hh}^{-1}]_{ii}$ . As it can be observed from (13), the lower bound in (12) depends only on the distribution of power  $P$  between the two phases. However, in order to achieve this bound, the positions and the power distribution for the training symbols should also be determined.

By substituting  $P_1 = P - P_2$  in  $f(P_1, P_2)$ , we get that  $f(P_1, P_2) = g(P_1)$ , where  $g(P_1)$  is monotonically decreasing, attaining its minimum value when  $P_1$  is maximized, i.e.  $P_1 =$

$P$  or  $P_2 = 0$ . Thus, the *minimum* value of the lower bound in (12) is expressed as

$$g(P) = \sum_{i=1}^L \frac{(2L)^{-1}}{c_i + \sigma_D^{-2} P} + \sum_{i=L+1}^{2L} \frac{(2L)^{-1}}{c_i + u^{-1} P} \quad (14)$$

As already mentioned, to achieve this lower bound,  $\mathbf{C}_e$  must be diagonal, or from (12),  $\mathbf{T} = \mathbf{A}^H \mathbf{C}_{WW}^{-1} \mathbf{A}$  must be diagonal. Having  $\mathbf{X}_{2,L} = \mathbf{0}$  in (6), it can be shown that  $\mathbf{T}$  becomes diagonal when  $p_i = P/L, \forall i$  and  $\mathbf{F}_L^H \mathbf{F}_L = \frac{L}{N} \mathbf{I}_L$ . The last two conditions are satisfied when the pilot tones in the first transmission phase are equispaced and equipowered.<sup>1</sup>

Overall, the performance of the estimator of Eq. (7) is optimized when (a)  $P_1 = P, P_2 = 0$ , (b)  $p_i = P/L, q_i = 0$  and (c)  $i_k = k \frac{N}{L}$ . The minimum channel estimation MSE is then given by Eq. (14).

## 4. LS ESTIMATION

### 4.1. Channel estimation

The LS estimator of  $\mathbf{h}$  in (5), and the error covariance matrix  $\mathbf{C}_{LS} = \mathcal{E}\{(\hat{\mathbf{h}} - \mathbf{h})(\hat{\mathbf{h}} - \mathbf{h})^H\}$  that describes its performance, are given by

$$\hat{\mathbf{h}} = \mathbf{A}^{-1} \mathbf{Y}, \quad \mathbf{C}_{LS} = \mathbf{A}^{-1} \mathbf{C}_{WW} \mathbf{A}^{-H}, \quad (15)$$

where  $\mathbf{C}_{WW}$  is defined in (10). However, depending on whether the channel  $\mathbf{h}_{RD}$  is assumed stochastic or deterministic, matrix  $\mathbf{C}_L$  is different, leading to different solutions for the optimization problem, as explained below.

### 4.2. Training design: Stochastic case

In this framework,  $\mathbf{h}_{RD}$  is assumed to be a complex Gaussian random vector, as in the LMMSE case and  $\mathbf{C}_L$  is defined by the expression given after Eq. (10). Moreover, the optimization problem is defined as in (11) by replacing  $\mathbf{C}_e$  with  $\mathbf{C}_{LS}$ . The trace of the error covariance matrix will be now lower bounded as follows

$$\frac{\text{Tr}\{\mathbf{C}_{LS}\}}{2L} \geq \frac{\sum_{i=1}^{2L} [\mathbf{A}^H \mathbf{C}_{WW}^{-1} \mathbf{A}]_{ii}^{-1}}{2L} = f_{LS}(P_1, P_2), \quad (16)$$

where

$$f_{LS}(P_1, P_2) = \frac{1}{2(\sigma_D^{-2} P_1 + u^{-1} P_2)} + \frac{1}{2u^{-1} P_1}, \quad (17)$$

and  $u$  is defined as before.

Following the same procedure as in the LMMSE case, the conditions under which the performance of the LS estimator is optimized turn out to be identical. In this case, the achievable lower bound is given by

$$g_{LS}(P) = \frac{\sigma_D^2 + u}{2P} = \frac{2\sigma_D^2 + |\alpha|^2 \sigma_R^2 \sigma^2}{2P}, \quad (18)$$

<sup>1</sup>It is also assumed that  $N/L$  is an integer

where  $\sigma^2$  is defined as in the LMMSE case.

It can be seen that  $g_{LS}(P)$  is obtained from  $g(P)$  by setting  $c_i = 0$  in (14). Since  $c_i \geq 0, \forall i$ ,  $g_{LS}(P) \geq g(P)$ , i.e. as expected, the performance of the LMMSE estimator is superior than that of the LS estimator provided that the second order statistics of the involved channels are available.

### 4.3. Training design: Deterministic case

In this case,  $\underline{h}_{RD}$  is considered to be deterministic and, so,  $\mathbf{C}_L = \sigma_D^2 + |\alpha|^2 \sigma_R^2 \mathbf{\Lambda}_{RD,L}^H \mathbf{\Lambda}_{RD,L}$ . In the following we define  $\mathbf{\Lambda}_{RD,L}^H \mathbf{\Lambda}_{RD,L} = \text{diag}\{\sigma_1^2, \dots, \sigma_L^2\}$ , where  $\sigma_i^2, i = 1, \dots, L$  are the squared amplitudes of the relay-to-destination channels in  $L$  carriers. Such a modeling is useful, for example, when the channel between  $R$  and  $D$  is constant for long periods of time [14] (e.g., when the relay and the destination are static nodes such as base stations). In such cases,  $\sigma_i^2, i = 1, \dots, L$ , can be considered to be known at  $S$ .

The minimization problem is identical to the one defined for the stochastic case. However, here, we can not follow the same optimization procedure because the related bound is not attainable. After some algebraic manipulations, the following relation is derived

$$\begin{aligned} \frac{\text{Tr}\{\mathbf{C}_{LS}\}}{2L} &= \frac{1}{2LN} \text{Tr}\{(\mathbf{X}^H \mathbf{C}_{WW}^{-1} \mathbf{X})^{-1} (\mathbf{F}_d \mathbf{F}_d^H)^{-1}\} \\ &= \frac{1}{2LN} \text{Tr}\{\mathbf{Q}\mathbf{S}\} \\ &\geq \frac{1}{2LN} \sum_i \lambda_i \mu_i, \end{aligned} \quad (19)$$

where  $\lambda_i$ 's are the eigenvalues of  $\mathbf{Q}$  in non-decreasing order and  $\mu_i$ 's are the eigenvalues of  $\mathbf{S}$  in non-increasing order [15, p. 249]. Equality holds when both matrices are diagonal.

By simple inspection, matrix  $\mathbf{Q} = (\mathbf{X}^H \mathbf{C}_{WW}^{-1} \mathbf{X})^{-1}$  becomes diagonal when  $\mathbf{X}_{2,L} = 0$  or, else, when  $P_1 = P, P_2 = 0$  and, consequently  $\underline{\mathbf{q}} = \mathbf{0}$ . Similarly, matrix  $\mathbf{S} = (\mathbf{F}_d \mathbf{F}_d^H)^{-1}$  becomes diagonal only when  $\mathbf{F}_L \mathbf{F}_L^H = \frac{L}{N} \mathbf{I}_L$  which requires that the pilot tones are equispaced. This means that the set  $\{i_k\}$  is determined as in the LMMSE case. Using these results in Eq. (19), the trace becomes equal to  $\frac{1}{2L} \text{Tr}\{\mathbf{C}_{LS}\} = \frac{1}{2LN} \sum_i [\mathbf{Q}]_{ii} [\mathbf{S}]_{ii} = h(\underline{\mathbf{p}})$ , where

$$h(\underline{\mathbf{p}}) = \frac{1}{2L^2} \sum_{i=1}^L \frac{\sigma_D^2 + u_i}{p_i}, \quad (20)$$

and  $u_i = \sigma_D^2 + |\alpha|^2 \sigma_R^2 \sigma_i^2, i = 1, \dots, L$ . Finally, the optimal power distribution of  $\underline{\mathbf{p}}$  is found by minimizing Eq. (20) under the constraint  $\sum_{i=1}^L p_i = P$  using Lagrange multipliers. The solution of this optimization problem turns out to be

$$p_i = \frac{\sqrt{\sigma_D^2 + u_i}}{\sum_{i=1}^L \sqrt{\sigma_D^2 + u_i}} P, \quad i = 1, \dots, L. \quad (21)$$

and the minimum MSE of the LS estimator for a deterministic  $\underline{h}_{RD}$  is given by (20).

## 5. EXPERIMENTS

In this section, the theoretical results are verified through simulations. More specifically, the MSE  $\frac{1}{2L} E\{\|\underline{\mathbf{h}} - \hat{\underline{\mathbf{h}}}\|^2\}$  versus the signal to noise ratio (SNR) at the destination is plotted for the estimators presented in the previous sections. We assume that the variances of each channel  $\underline{h}_i$  follow a simple exponentially decaying model, i.e.  $\sigma_i^2(n) = e^{-n} + 1$  where  $n = 0, 1, \dots, L_i - 1$  and  $i \in \{SD, SR, RD\}$ . Also, we set for simplicity  $\alpha = 1, P = 1$  and  $\sigma_D^2 = \sigma_R^2$ .

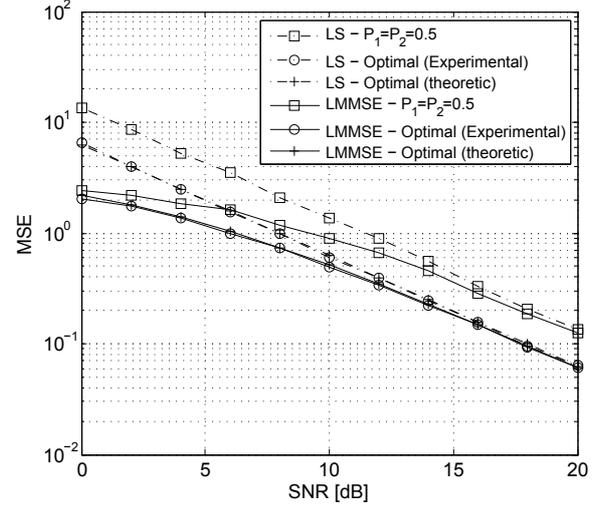


Fig. 2. LMMSE vs LS (stochastic case)

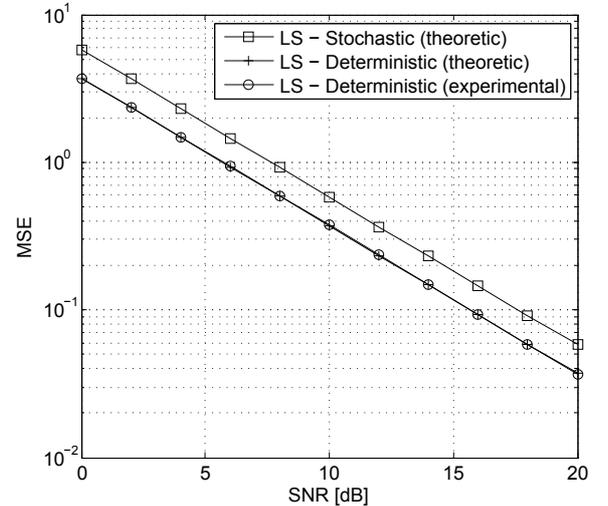


Fig. 3. LS: Deterministic vs Stochastic case

In Fig. 2, MSE versus SNR curves are presented for the LMMSE and LS (stochastic case) channel estimation. We observe that Eqs. (14), (18), describing the optimal performance of the two estimators, are verified. Moreover, the performance degradation of the estimation is presented in the case  $P_1 = P_2 = P/2$ , i.e. when we deviate from the con-

dition  $P_1 = P$ ,  $P_2 = 0$ . Similar results are obtained for the other conditions, as well.

In Fig. 3, the two LS scenaria are compared. We can see from the figure that not only the theoretical MSE of Eq. (20) is verified but also the superiority of the deterministic LS over the stochastic LS estimator is corroborated.

## 6. CONCLUSIONS

In this paper, optimal channel estimation has been investigated for a single relay cooperative system in a frequency selective environment. Based on the LMMSE and the LS estimation criteria, specific conditions are provided for optimizing the estimation performance at the destination node.

It has been shown that, in all cases, the source node should assign the total training power at the first transmission phase to avoid interference with the relay during the second phase. An important consequence of this approach is that channel estimation is decoupled in the two phases and can, thus, be performed independently.

For LMMSE and stochastic LS estimation, the optimal solution calls for training symbols equipowered and equispaced. However, when the relay-to-destination channel is assumed deterministic, equispaced but not equipowered training symbols result in the minimum MSE. The presented theoretical results have been completely verified by simulation experiments.

## 7. REFERENCES

- [1] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative Communications in Wireless Networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, Oct. 2004.
- [2] C. S. Patel and G. L. Studer, "Channel estimation for amplify and forward relay based cooperation diversity systems," *IEEE Trans. on Wireless Communications*, vol. 6, no. 6, pp. 2348–2356, July 2007.
- [3] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. on Wireless Communications*, vol. 7, no. 5, pp. 1907–1916, May 2008.
- [4] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. on Signal Processing*, vol. 56, no. 12, pp. 5937–5949, Dec. 2008.
- [5] Kwanghoon Kim, Haelyong Kim, and Hyuncheol Park, "Ofdm channel estimation for the amply-and-forward cooperative channel," in *Proc. of IEEE VTC*, 2007, pp. 1642–1646.
- [6] K. Yan, S. Ding, Y. Qiu, Y. Wand, and H. Liu, "A low-complexity LMMSE channel estimation method for OFDM-based cooperative diversity systems with multiple Amplify-and-forward relays," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, no. 149803.
- [7] F. Gao, R. Zhang, and Y.-C. Liang, "Channel Estimation for OFDM Modulated Two-Way Relay Networks," *IEEE Trans. on Signal Processing*, vol. 57, no. 11, pp. 4443–4455, Nov. 2009.
- [8] A. S. Lalos, A. A. Rontogiannis, and K. Berberidis, "Channel estimation techniques in amplify and forward relay networks," in *Proc. of IEEE SPAWC*, 2008, p. 446450.
- [9] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading Relay Channels: Performance Limits and Space-Time Signal Design," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.
- [10] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.
- [11] Z. Wang and G. B. Giannakis, "Wireless Multicarrier Communications: where Fourier meets Shannon," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 29–48, May 2000.
- [12] I. Barhumi, G. Leus, and M. Moonen, "Optimal Training Design for MIMO OFDM Systems in Mobile Wireless Channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1615–1624, June 2003.
- [13] S. M. Kay, *Fundamentals of statistical signal processing - Estimation theory*, Prentice Hall, 1993.
- [14] H. Yomo and E. de Carvalho, "A CSI estimation method for wireless relay networks," *IEEE Communications letters*, vol. 11, no. 6, pp. 480–482, June 2007.
- [15] I. Olkin A. W. Marshall, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, 1979.