Int. J. Communications, Network and System Sciences, **2010, 3, 962-971**
doi:10.4236/ijcns.2010.312131 Published Online December 2010 (http://www.SciRP.org/journal/ijcns)

Performance Improvement in Estimation of Spatially Correlated Rician Fading MIMO Channels Using a New LMMSE Estimator

Hamid Nooralizadeh[1], Shahriar Shirvani Moghaddam[2]
[1]*Electrical Engineering Department, Islamshahr Branch, Islamic Azad University, Islamshahr, Iran*
[2]*DCSP Research Laboratory, Department of Electrical and Computer Engineering,*
Shahid Rajaee Teacher Training University, Tehran, Iran
E-mail: h_n_alizadeh@yahoo.com, sh_shirvani@srttu.edu
Received August 18, 2010; *revised September* 23, 2010; *accepted November* 2, 2010

Abstract

In most of the previous researches on the multiple-input multiple-output (MIMO) channel estimation, the fading model has been assumed to be Rayleigh distributed. However, the Rician fading model is suitable for microcellular mobile systems or line of sight mode of WiMAX. In this paper, the training based channel estimation (TBCE) scheme in the spatially correlated Rician flat fading MIMO channels is investigated. First, the least squares (LS) channel estimator is probed. Simulation results show that the Rice factor has no effect on the performance of this estimator. Then, a new linear minimum mean square error (LMMSE) technique, appropriate for Rician fading channels, is proposed. The optimal choice of training sequences with mean square error (MSE) criteria is investigated for these estimators. Analytical and numerical results show that the performance of proposed estimator in the Rician channel model compared with Rayleigh one is much better. It is illustrated that when the channel Rice factor and/or the correlation coefficient increase, the performance of the proposed estimator significantly improves.

Keywords: Channel Estimation, MIMO, Spatially Correlated Rician Fading, Optimal Training Sequences, LS, Generalized LMMSE

1. Introduction

Due to high capacity and diversity gain, multiple input multiple output (MIMO) systems have received considerable attention in wireless communications. It has been demonstrated that when the fades connecting pairs of transmit and receive antenna elements are independent, identically distributed (i.i.d.), the capacity of a Rayleigh distributed flat fading channel increases almost linearly with the minimum number of transmitter and receiver antennas [1-3]. Moreover, in [3] it is indicated that Rician fading can improve the capacity of a multiple antenna system, especially if the transmitter knows the value of the Rice factor.

In order to attain the advantages of MIMO systems, it is necessary that the receiver and/or transmitter have access channel state information (CSI). One of the most usual approaches to identify MIMO CSI is training based channel estimation (TBCE). This class of estimation is attractive especially when it decouples symbol detection from channel estimation and thus simplifies the receiver implementation and relaxes the required identification conditions.

The optimal choice of training signals is usually investigated by minimizing mean square error (MSE) of the linear MIMO channel estimator. In the literature, it is perceived that optimal design of training sequences for MIMO channel estimation is connected with the channel statistical characteristics, e.g., fading model and the channel noise model. For example, in [4], a sub-matrix of the discrete Fourier transform (DFT) matrix has been used to identify the Rayleigh distributed flat fading MIMO channel. In [5-7], in order to estimate MIMO inter symbol interference (ISI) channel, the delta sequence is used as optimal training. Further studies are reported in [8-13] using optimal training and considering a few aspects, e.g., the peak to average power ratio (PAPR) constraint on training sequences.

In [4,14-19] the spatially correlated fading MIMO channel is considered. In [14], the frequency offset and channel gain estimation is considered for MIMO ISI correlated fading channels. In [15,16], it is investigated that the impacts of spatial correlation are helpful not only to improve channel estimation but also to decrease the training length. It is noteworthy that the spatial correlation harms channel capacity [20].

In [4], the performance of the least squares (LS), scaled LS (SLS), minimum mean square error (MMSE), and relaxed MMSE (RMMSE) estimators is studied in the Rayleigh fading MIMO channel. The MMSE estimator has the best performance among the estimators, because it can employ more a priori knowledge about the channel.

In most previous works on the MIMO channel estimation, the channel fading is assumed to be Rayleigh distributed. Of course, the Rayleigh fading model is known to be a reasonable assumption for fading encountered in many wireless communications systems. However, Rician fading model is suitable for suburban areas where a line of sight (LOS) path often exists. This may also be true for microcellular or picocellular systems with cells of less than several hundred meters in radius.

In [21], the TBCE scheme is investigated in MIMO Rician flat fading channels. By the new method of shifted scaled least squares (SSLS), it is shown that increasing the channel Rice factor improves the performance of channel estimation. However, the SSLS channel estimator is only appropriate for uncorrelated Rician channels because this estimator cannot exploit the knowledge of spatial correlation of the MIMO channels. In uncorrelated fading, it is assumed that antenna elements are placed sufficiently apart. However, it is not always realized in practice due to insufficient antenna spacing when the channel estimation is used in compact terminals. The linear MMSE (LMMSE) channel estimator of [4] is appropriate for spatially correlated channels. Nevertheless, this estimator cannot benefit from the Rice factor of the Rician fading channels.

In this paper, a general form of the LMMSE channel estimator is proposed that is appropriate for spatially correlated Rician fading MIMO channels. We extend the results of [4] in the Rayleigh fading model to the more general Rician fading case. It is shown that this estimator can exploit the knowledge of both spatial correlation and Rician fading of the MIMO channels.

First, the traditional LS method is examined. It is demonstrated that the performance of this estimator is independent of the Rice factor. Then, the proposed MIMO channel estimator is introduced; we refer to it as generalized linear minimum mean square error (GLMMSE) estimator. It is shown that in the spatially correlated Rician fading MIMO channel when the Rice factor and/or the correlation coefficient increase, the accuracy of the GLMMSE estimator improves. Note that the performance of the proposed estimator in the Rician fading channels improves because:

1) We consider the effect of Rice factor and suggest a new formulation comparing with other references as [4] and [22].

2) We design the optimal training sequence appropriate for spatially correlated Rician channel model.

It means that the optimal training sequence and the LMMSE formulation in the Rayleigh channel estimation [4] are not suitable for Rician channel estimation.

The rest of this paper is organized as follows. Section 2 introduces the channel model. The performance of the LS and GLMMSE channel estimators and the optimal training sequence design are investigated in Sections 3 and 4. Simulation results are presented in Section 5. Finally, Section 6 concludes this paper.

2. Channel Model

For flat MIMO channels, the block fading model is assumed. It means that the channel response is fixed within one block and changes from one block to another randomly. The transmitter and receiver are equipped with N_T and N_R antennas, respectively. During the training period, the received signal in such a system can be written in the matrix form as

$$\mathbf{Y} = \mathbf{HX} + \mathbf{V} \tag{1}$$

where $\mathbf{Y}$, $\mathbf{X}$ and $\mathbf{V}$ are the complex N_R-vector of received signals on the N_R receive antennas, the possibly complex N_T-vector of transmitted signals on the N_T transmit antennas, and the complex N_R-vector of additive receiver noise, respectively. The elements of noise matrix are i.i.d. complex Gaussian random variables with zero-mean and σ_n^2 variance, and the correlation matrix of $\mathbf{V}$ is then given by

$$\mathbf{R}_{\mathrm{V}} = E\left\{\mathbf{V}^H\mathbf{V}\right\} = \sigma_n^2 N_R \mathbf{I}_{N_P} \tag{2}$$

where N_P is the number of transmitted training symbols by each transmitter antenna, $\mathbf{I}_{N_P}$ is the $N_P \times N_P$ identity matrix, $(\cdot)^H$ denotes the matrix Hermitian, and $E\{\cdot\}$ is the mathematical expectation.

The channel matrix $\mathbf{H}$ in the model (1) is the $N_R \times N_T$ matrix of complex fading coefficients. The $(r, t)^{\text{th}}$ element of the matrix $\mathbf{H}$ denoted by $h_{r,t}$ represents the fading coefficient value between the r^{th} receiver antenna and the t^{th} transmitter antenna. The elements of $\mathbf{H}$ are Gaussian with independent real and imaginary parts each distributed as $\mathcal{N}\left(\mu/\sqrt{2},\sigma^2\right)$. So, the elements $h_{r,t}$ of **H** are identically distributed complex Gaussian random variables $h_{r,t} \sim \mathcal{CN}\left(\mu(1+j)/\sqrt{2},2\sigma^2\right)$ for $r = 1, 2,\cdots, N_R$ and $t = 1, 2,\cdots, N_T$. The magnitude of the elements of **H** has the Rician distribution

$$f_A(a) = 2(1+\kappa)ae^{-(1+\kappa)a^2-\kappa} I_\circ\left(2\sqrt{\kappa(1+\kappa)}\,a\right) \quad (3)$$

where $I_\circ$ is the modified Bessel function of first kind, of order zero, and the Rice factor, κ, can be defined as

$$\kappa = \frac{\mu^2}{2\sigma^2} \quad (4)$$

For notational convenience, we have also presented the normalization $\mu^2 + 2\sigma^2 = 1$. Note that (3) reduces to the Rayleigh probability density function (pdf) when $\kappa = 0$. If elements of **H** are distributed as described above, **H** will be a complex normally distributed matrix, denoted as $\mathbf{H} \sim \mathcal{CN}(\mathbf{M}, \mathbf{C}_H)$ where $\mathbf{C}_H$ and **M** are the Hermitian covariance matrix and the mathematical expectation matrix of the **H**, respectively. The matrix **M** can be written as follows:

$$\mathbf{M} = \frac{\mu}{\sqrt{2}}(1+j)\mathbf{1}_{N_R\times N_T} \quad (5)$$

Here, $\mathbf{1}_{N_R\times N_T}$ is an $N_R \times N_T$ matrix whose entries are all 1. We assume that the elements $h_{r,t}$ of **H** are correlated. Suppose that $\rho^{|m-n|}\left(0 \le |\rho| \le 1\right)$ is the correlation coefficient of elements in the columns m^{th} and n^{th} of the **H**. Therefore, the correlation of any two elements from m^{th} and n^{th} columns of **H** is expressed in the following form:

$$\begin{aligned} E\left\{h_{im}^* h_{in}\right\} &= 2\sigma^2\rho^{|m-n|} + \mu^2 \\ &= \frac{\rho^{|m-n|}}{1+\kappa} + \frac{\kappa}{1+\kappa} = \frac{\rho^{|m-n|}+\kappa}{1+\kappa} \end{aligned} \quad (6)$$

where m, $n = 1, 2,\cdots, N_T$, $i = 1, 2,\cdots, N_R$, and $(\cdot)^*$ denotes the complex conjugate. Then, the $(m, n)^{th}$ element of the channel correlation matrix can be written as

$$r_{mn} = N_R\frac{\rho^{|m-n|}+\kappa}{1+\kappa} \quad (7)$$

Therefore, the correlation matrix of the spatially correlated Rician fading MIMO channel can be expressed in Equation (8).

Note that when $\rho = 0$, (8) reduces to the special case of (12) in [3] and when $\kappa = 0$, it reduces to the spatially correlated Rayleigh fading channel introduced in [4]. Using (5), the covariance matrix of the described Rician fading model can be written as (9).

If the elements $h_{r,t}$ of **H** are uncorrelated ($\rho = 0$), we can write the result as

$$\mathbf{C}_H = \frac{N_R}{1+\kappa}\mathbf{I}_{N_T} \quad (10)$$

The elements of **H** and noise matrix are independent of each other. In order to estimate the channel matrix, it is required that $N_P \ge N_T$ training symbols are transmitted by each transmitter antenna. The function of a channel estimation algorithm is to recover the channel matrix **H** based on the knowledge of **Y** and **X**.

3. LS Channel Estimator

Consider that **H** is an unknown deterministic matrix. To identify it from (1), the LS approach minimizes $tr\left\{(\mathbf{Y}-\mathbf{HX})^H(\mathbf{Y}-\mathbf{HX})\right\}$ which results in

$$\hat{\mathbf{H}}_{LS} = \mathbf{YX}^H\left(\mathbf{XX}^H\right)^{-1} \quad (11)$$

where $tr\{\cdot\}$ denotes the trace of a matrix, $(\cdot)^{-1}$ denotes the matrix inverse. The LS error criterion (MSE) is defined by

$$J_{LS} = E\left\{\left\|\mathbf{H}-\hat{\mathbf{H}}_{LS}\right\|_F^2\right\} \quad (12)$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm. Let us write from (1) and (11)

$$\begin{aligned} \mathbf{H}-\hat{\mathbf{H}}_{LS} &= \mathbf{H}-(\mathbf{HX}+\mathbf{V})\mathbf{X}^H\left(\mathbf{XX}^H\right)^{-1} \\ &= -\mathbf{VX}^H\left(\mathbf{XX}^H\right)^{-1} \end{aligned} \quad (13)$$

$$\mathbf{R}_H = \frac{N_R}{1+\kappa}\begin{bmatrix} 1+\kappa & \rho+\kappa & \rho^2+\kappa & \cdots & \rho^{N_T-1}+\kappa \\ \rho+\kappa & 1+\kappa & \rho+\kappa & \cdots & \rho^{N_T-2}+\kappa \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho^{N_T-1}+\kappa & \rho^{N_T-2}+\kappa & \rho^{N_T-3}+\kappa & \cdots & 1+\kappa \end{bmatrix} \quad (8)$$

$$\mathbf{C}_H = \mathbf{R}_H - \mathbf{M}^H\mathbf{M} = \frac{N_R}{1+\kappa}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N_T-1} \\ \rho & 1 & \rho & \cdots & \rho^{N_T-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho^{N_T-1} & \rho^{N_T-2} & \rho^{N_T-3} & \cdots & 1 \end{bmatrix} \quad (9)$$

Using (2) and (13), the MSE (12) can be rewritten as

$$J_{LS} = E\left\{ \left\| -\mathbf{V}\mathbf{X}^H \left(\mathbf{X}\mathbf{X}^H\right)^{-1} \right\|_F^2 \right\} = \sigma_n^2 N_R tr\left\{ \left(\mathbf{X}\mathbf{X}^H\right)^{-1} \right\} \tag{14}$$

Let us find **X** which minimizes (14) subject to a transmitted power constraint. This is equivalent to the following optimization problem:

$$\min_{\mathrm{X}} tr\left\{ \left(\mathbf{X}\mathbf{X}^H\right)^{-1} \right\} \;\; S.T\; tr\left\{\mathbf{X}\mathbf{X}^H\right\} = p \tag{15}$$

where p is a given constant value considered as the total power of training matrix **X**. To solve (15), the Lagrange multiplier method is used. The problem can be written as

$$L\left(\mathbf{X}\mathbf{X}^H, \eta\right) = tr\left\{ \left(\mathbf{X}\mathbf{X}^H\right)^{-1} \right\} + \eta\left[tr\left\{\mathbf{X}\mathbf{X}^H\right\} - p \right] \tag{16}$$

where η is the Lagrange multiplier. By differentiating (16) with respect to $\mathbf{X}\mathbf{X}^H$ and setting the result equal to zero, it is obtained that the optimal training matrix should satisfy the Equation (17)

$$\mathbf{X}\mathbf{X}^H = \frac{1}{\sqrt{\eta}} \mathbf{I}_{N_T} \tag{17}$$

Equation (17) can be expressed in the following form using the constraint $tr\left\{\mathbf{X}\mathbf{X}^H\right\} = p$,

$$\mathbf{X}\mathbf{X}^H = \frac{p}{N_T} \mathbf{I}_{N_T} \tag{18}$$

Therefore, any training matrix with orthogonal rows of the same norm $\sqrt{p/N_T}$ is optimal. Let us dictate PAPR constraint on **X** that is considered in [4,11,12]. To satisfy this constraint, a properly normalized sub-matrix of the DFT matrix can be used

$$\mathbf{X} = \sqrt{\frac{p}{N_P N_T}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_{N_P} & \cdots & W_{N_P}^{(N_P-1)} \\ \vdots & \vdots & & \vdots \\ 1 & W_{N_P}^{(N_T-1)} & \cdots & W_{N_P}^{(N_T-1)(N_P-1)} \end{bmatrix} \tag{19}$$

where $W_k = \exp\left(j2\pi/k\right)$. Substituting (18) back into (14), the channel estimation error under optimal training is given by

$$\left(J_{LS}\right)_{\min} = \frac{\sigma_n^2 N_T^2 N_R}{p} \tag{20}$$

In a particular case that $N_T = N_R = 1$, single-input single-output (SISO) channel, the MSE of (20) is minimum. Then, increasing the number of antennas results in higher MSE. On the other hand, the capacity of an MIMO channel increases when the number of antennas is increased. Note that the error in (20) is proportional to the square of N_T. This causes a certain restriction in the number of transmit antennas as compared with the number of receive antennas used. For optimal training which satisfies (18), the LS channel estimator (11) yields

$$\hat{\mathbf{H}}_{LS} = \frac{N_T}{p} \mathbf{Y}\mathbf{X}^H \tag{21}$$

These results are the same as [4], because the LS estimator cannot exploit any statistical knowledge about the Rayleigh or Rician fading channels. In the next section, we derive new results in the Rician channel model by the new GLMMSE estimator.

4. Proposed GLMMSE Channel Estimator

For linear model (1), the MMSE and LMMSE estimators are identical [23]. So, let us obtain a general form of linear estimator, appropriate for Rician fading channels, that minimizes the estimation MSE of **H**. It can be expressed in the following form:

$$\hat{\mathbf{H}}_{GLMMSE} = E\{\mathbf{H}\} + \left(\mathrm{Y} - E\{\mathbf{Y}\}\right)\mathbf{A}_{\circ} = \mathbf{M} + \left(\mathbf{Y} - \mathbf{M}\mathbf{X}\right)\mathbf{A}_{\circ} \tag{22}$$

Here, $\mathbf{A}_{\circ}$ has to be obtained so that the following MSE is minimized:

$$J_{GLMMSE} = E\left\{ \left\| \mathbf{H} - \hat{\mathbf{H}}_{GLMMSE} \right\|_F^2 \right\} \tag{23}$$

The optimal $\mathbf{A}_{\circ}$ can be found from $\partial J_{GLMMSE} / \partial \mathbf{A}_{\circ} = 0$ and it is given by

$$\mathbf{A}_{\circ} = \left(\mathbf{X}^H \mathbf{C}_H \mathbf{X} + \sigma_n^2 N_R \mathbf{I}_{N_P} \right)^{-1} \mathbf{X}^H \mathbf{C}_H \tag{24}$$

Proof: See the Appendix.

Substituting $\mathbf{A}_{\circ}$ back into (22), the GLMMSE channel estimator of **H** can be rewritten as

$$\hat{\mathbf{H}}_{GLMMSE} = \mathbf{M} + \left(\mathbf{Y} - \mathbf{M}\mathbf{X}\right)\left(\mathbf{X}^H \mathbf{C}_H \mathbf{X} + \sigma_n^2 N_R \mathbf{I}_{N_P} \right)^{-1} \mathbf{X}^H \mathbf{C}_H \tag{25}$$

Note that in the Rayleigh fading channel, $\mathbf{M} = 0$, $\mathbf{C}_H = \mathbf{R}_H$. This estimator not only utilizes received and transmitted signals but also takes the advantages of the channel first and second-order statistics. The required knowledge of the channel statistics can be estimated by some methods. For instance, the problem of estimating the MIMO channel covariance, based on limited amounts of training sequences, is treated in [24]. Moreover, in [25], estimation of the channel autocorrelation matrix is performed by an instantaneous autocorrelation estimator where only one channel estimate (obtained by a very low complexity channel estimator) has been used as input.

The performance of the GLMMSE channel estimator is measured by the error matrix $\boldsymbol{\varepsilon} = \mathbf{H} - \hat{\mathbf{H}}_{GLMMSE}$, whose pdf is Gaussian with zero mean and the following covariance matrix:

$$\mathbf{C}_{\varepsilon} = \mathbf{R}_{\varepsilon} = E\left\{\boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon}\right\} = \left(\mathbf{C}_H^{-1} + \frac{1}{\sigma_n^2 N_R}\mathbf{X}\mathbf{X}^H\right)^{-1} \quad (26)$$

Therefore, the estimation error can be computed as

$$\begin{aligned} J_{GLMMSE} &= E\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}_{GLMMSE}\right\|_F^2\right\} = E\left\{tr\left(\boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon}\right)\right\} \\ &= tr\left\{\mathbf{C}_{\varepsilon}\right\} = tr\left\{\left(\mathbf{C}_H^{-1} + \frac{1}{\sigma_n^2 N_R}\mathbf{X}\mathbf{X}^H\right)^{-1}\right\} \end{aligned} \quad (27)$$

Let us find $\mathbf{X}$ which minimizes the channel estimation error subject to a transmitted power constraint. Thus, we seek the matrix $\mathbf{X}$ that is the solution to the optimization problem (28)

$$\min_{\mathbf{X}} tr\left\{\left(\mathbf{C}_H^{-1} + \frac{1}{\sigma_n^2 N_R}\mathbf{X}\mathbf{X}^H\right)^{-1}\right\} S.T\; tr\left\{\mathbf{X}\mathbf{X}^H\right\} = p \quad (28)$$

To solve (28), the Lagrange multiplier method is applied. The problem can be written as

$$\begin{aligned} L\left(\mathbf{X}\mathbf{X}^H, \eta\right) &= tr\left\{\left(\mathbf{C}_H^{-1} + \frac{1}{\sigma_n^2 N_R}\mathbf{X}\mathbf{X}^H\right)^{-1}\right\} \\ &+ \eta\left[tr\left\{\mathbf{X}\mathbf{X}^H\right\} - p\right] \end{aligned} \quad (29)$$

By differentiating (29) with respect to $\mathbf{X}\mathbf{X}^H$ and equating to zero, we have

$$\mathbf{X}\mathbf{X}^H = \sqrt{\frac{\sigma_n^2 N_R}{\eta}}\mathbf{I}_{N_T} - \sigma_n^2 N_R \mathbf{C}_H^{-1} \quad (30)$$

Using the constraint $tr\left\{\mathbf{X}\mathbf{X}^H\right\} = p$, (30) can be expressed as

$$\mathbf{X}\mathbf{X}^H = \frac{p + \sigma_n^2 N_R tr\left\{\mathbf{C}_H^{-1}\right\}}{N_T}\mathbf{I}_{N_T} - \sigma_n^2 N_R \mathbf{C}_H^{-1} \quad (31)$$

By applying (31) in (27), $\mathbf{C}_H^{-1} + \left(1/\sigma_n^2 N_R\right)\mathbf{X}\mathbf{X}^H$ will be a diagonal matrix. Therefore, according to the lemma 1 in [4], we obtain that the MSE (27) will be minimized as

$$\left(J_{GLMMSE}\right)_{\min} = \frac{N_T^2}{p/\left(N_R \sigma_n^2\right) + tr\left\{\mathbf{C}_H^{-1}\right\}} \quad (32)$$

In a particular case that the elements $h_{r,t}$ of $\mathbf{H}$ are uncorrelated ($\rho = 0$), the covariance matrix of $\mathbf{H}$ is diagonal and from (10)

$$tr\left\{\mathbf{C}_H^{-1}\right\} = \frac{1+\kappa}{N_R}N_T \quad (33)$$

Using (33), it is observed that (31) is the same as (18). It means that in the case of $\rho = 0$, both the LS and GLMMSE approaches have the same condition on the optimal training matrices.

Using (10) and (18), the GLMMSE channel estimator (25) reduces to

$$\hat{\mathbf{H}}_{GLMMSE} = \left(1-\beta\right)\mathbf{M} + \alpha \mathbf{Y}\mathbf{X}^H \quad (34)$$

where

$$\alpha = \frac{N_P}{p + \sigma_n^2\left(1+\kappa\right)N_P},\; \beta = \frac{pN_P}{pN_T + \sigma_n^2\left(1+\kappa\right)N_P N_T} \quad (35)$$

Substituting (33) back into (32), MSE in the particular case of $\rho = 0$ is given by (36)

$$\left(J_{GLMMSE}\right)_{\min} = \frac{N_R N_T^2}{\left(p/\sigma_n^2\right) + \left(1+\kappa\right)N_T} \quad (36)$$

Equation (36) shows that when the Rice factor, κ, increases, the MSE considerably decreases. In other words, in the Rician fading channel model compared with Rayleigh one, the obtained MSE improves. Increasing the channel Rice factor causes decreasing the MSE, and for higher values of κ, the MSE is proportional to $1/\kappa$. When $\kappa = 0$, (36) is identical to the acquired result in [4] for RMMSE channel estimator.

In general, the covariance matrix of the $\mathbf{H}$ is given by (9) and the condition for the optimal training matrix of the GLMMSE channel estimator is different from that of the LS estimator. We seek the matrix $\mathbf{X}$ that is the solution to (31). Thus, the eigen-value decomposition (EVD) of $\mathbf{C}_H$ in the form of $\mathbf{C}_H = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$ is used, where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the nonnegative eigenvalues of the $\mathbf{C}_H$ as its diagonal elements and $\mathbf{Q}$ is a unitary matrix containing the eigenvectors of the $\mathbf{C}_H$ in its columns. Using this notation, (27) can be rewritten as

$$\begin{aligned} J_{GLMMSE} &= tr\left\{\left(\mathbf{Q}\boldsymbol{\Lambda}^{-1}\mathbf{Q}^H + \frac{1}{\sigma_n^2 N_R}\mathbf{X}\mathbf{X}^H\right)^{-1}\right\} \\ &= tr\left\{\left(\boldsymbol{\Lambda}^{-1} + \frac{1}{\sigma_n^2 N_R}\mathbf{Q}^H\mathbf{X}\mathbf{X}^H\mathbf{Q}\right)^{-1}\right\} \end{aligned} \quad (37)$$

Equation (37) can be reduced by replacing $\left(1/\sqrt{\sigma_n^2 N_R}\right)\mathbf{Q}^H\mathbf{X}$ by $\tilde{\mathbf{X}}$

$$J_{GLMMSE} = tr\left\{\left(\boldsymbol{\Lambda}^{-1} + \tilde{\mathbf{X}}\tilde{\mathbf{X}}^H\right)^{-1}\right\} \quad (38)$$

Also, the total transmitted training power constraint in (28) can be rewritten in the following form:

$$tr\left\{\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H\right\} = \frac{1}{\sigma_n^2 N_R}tr\left\{\mathbf{Q}^H\mathbf{X}\mathbf{X}^H\mathbf{Q}\right\} = \frac{p}{\sigma_n^2 N_R} \quad (39)$$

Using lemma 1 of [4], the minimum of (38) will be obtained if $\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H$ has the following diagonal structure:

$$\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H = diag\left(|\tilde{x}_1|^2, |\tilde{x}_2|^2, \cdots, |\tilde{x}_{N_T}|^2\right) \quad (40)$$

Then, the optimal training matrix for the GLMMSE channel estimation method can be found by solving the following constrained optimization problem:

$$\min_{\tilde{\mathbf{X}}} tr\left\{\left(\mathbf{\Lambda}^{-1} + \tilde{\mathbf{X}}\tilde{\mathbf{X}}^H\right)^{-1}\right\} \; S.T \; tr\left\{\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H\right\} = \frac{p}{\sigma_n^2 N_R} \quad (41)$$

Using Lagrange multiplier method and taking into account (40), the optimal training matrix of the GLMMSE method can be found by minimizing the following function:

$$\begin{aligned} L\left(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H, \eta\right) &= tr\left\{\left(\mathbf{\Lambda}^{-1} + \tilde{\mathbf{X}}\tilde{\mathbf{X}}^H\right)^{-1}\right\} \\ &\quad + \eta\left[tr\left\{\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H\right\} - \left(p/\sigma_n^2 N_R\right)\right] \\ &= \sum_{i=1}^{N_T}\left(\lambda_i^{-1} + \left|\tilde{x}_i\right|^2\right)^{-1} \\ &\quad + \eta\sum_{i=1}^{N_T}\left(\left|\tilde{x}_i\right|^2 - \left(p/\sigma_n^2 N_R N_T\right)\right) \end{aligned} \quad (42)$$

where λ_i for $i = 1, 2, \cdots, N_T$ are the nonnegative eigenvalues of the $\mathbf{C}_H$. Differentiating (42) with respect to $|\tilde{x}_i|^2$ for $i = 1, 2, \cdots, N_T$ and setting the results equal to zero yields

$$\frac{1}{\left(\lambda_i^{-1} + \left|\tilde{x}_i\right|^2\right)^2} = \eta, \quad i = 1, 2, \cdots, N_T \quad (43)$$

The water-filling-type solution of this problem is

$$\tilde{x}_i = \begin{cases} \sqrt{\eta_\circ - \lambda_i^{-1}}, & if\ \eta_\circ > \lambda_i^{-1} \\ 0, & if\ \eta_\circ < \lambda_i^{-1} \end{cases} \quad (44)$$

The constant $\eta_\circ = \eta^{-0.5}$ should be adjusted so that the transmitted power constraint (39) is satisfied. If $N_P = N_T$, then the optimal $\tilde{\mathbf{X}}$ can be written in the following matrix form:

$$\tilde{\mathbf{X}} = \left(\left[\eta_\circ \mathbf{I}_{N_T} - \mathbf{\Lambda}^{-1}\right]^+\right)^{1/2} \quad (45)$$

where the operator $[\cdot]^+$ is interpreted as meaning that all negative entries of a real matrix are replaced by zeros and all nonnegative entries are leaved unchanged. Finally, the optimal training matrix can be written as

$$\mathbf{X} = \sqrt{\sigma_n^2 N_R}\,\mathrm{Q}\left(\left[\eta_\circ \mathbf{I}_{N_T} - \mathbf{I}^{-1}\right]^+\right)^{1/2} \quad (46)$$

The matrices $\mathbf{Q}$ and $\mathbf{\Lambda}$ are obtained from the EVD of $\mathbf{C}_H$ and the constant $\eta_\circ$ should be adjusted so that the transmitted power constraint in (28) is satisfied.

5. Simulation Results

In this section, our goal is to compare the performance of the LS and GLMMSE channel estimators in the Rayleigh and Rician flat fading channels, numerically. Also, we contrast the results with the LMMSE channel estimator of [4] and SSLS channel estimator of [21]. For the sake of simplicity and without loss of generality, we assume a 2 × 2 MIMO channel, *i.e.*, $N_T = N_R = 2$. It is also supposed that the spatially correlated Rician fading MIMO channel has the covariance matrix (9). Hence, the elements of the covariance matrix of the channel can be written in the following form:

$$\left[\mathbf{C}_H\right]_{k,l} = \frac{N_R}{1+\kappa}\rho^{|k-l|}; \; 0 \le |\rho| \le 1 \quad (47)$$

where k, l are the indexes of the array sensors. As a performance measure, we consider the channel MSE, normalized with the average channel energy as:

$$NMSE = \frac{E\left\{\left\|\mathbf{H} - \hat{\mathbf{H}}\right\|_F^2\right\}}{E\left\{\left\|\mathbf{H}\right\|_F^2\right\}} \quad (48)$$

The signal to noise ratio (SNR) is defined as:

$$SNR = \frac{p}{\sigma_n^2} \quad (49)$$

Figure 1 shows the normalized MSE (NMSE) of the LS channel estimator with optimal training versus SNR for various Rice factors of the channel. As it is expected, this estimator cannot exploit the knowledge of the channel Rice factor; a phenomenon that is confirmed by this figure. In [4] and [9], it is demonstrated that the LS estimator does not require any knowledge about the channel. Hence, it is also clear that the performance of this estimator is independent of ρ and the type of channel fading.

The numerical and analytical results coincide when the number of independent simulation runs reaches to 5000. Using (20), the NMSE of the LS channel estimator is plotted in **Figure 1**. As depicted in this figure, the analytical and numerical results are almost identical.

Figures 2 and **3** indicate the NMSE of the LS, LMMSE of [4] and GLMMSE channel estimators with orthogonal training of (19) versus SNR in the case of $\rho = 0.1$ and $\rho = 0.8$, respectively. It is observed that the proposed GLMMSE estimator has the best performance among the methods tested. Increasing the channel Rice factor and/or the correlation coefficient of the array sensor elements improves the performance of this estimator especially at low SNRs compared with the fixed performance of the LS estimator. Moreover, increasing ρ improves the performance of the LMMSE channel estimator. These results are due to the fact that the Bayesian

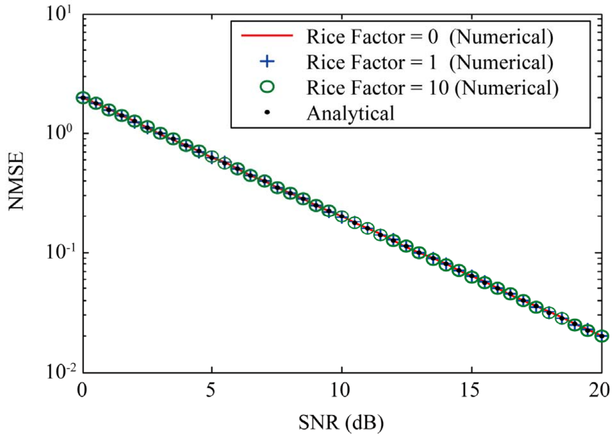

Figure 1. NMSE of the LS channel estimator for various Rice factors of the channel, $N_R = N_T = 2$ (Numerical and analytical results).

estimators, e.g., the proposed GLMMSE channel estimator can employ more a priori knowledge about the channel. As depicted in **Figures 2** and **3**, at high SNRs, the performances of the LS, LMMSE and GLMMSE channel estimators are nearly identical, particularly for low Rice factors and spatial correlations. However, at higher κ and ρ, the performance of the GLMMSE estimator is still better than that of the LS estimator. Note that in the special case, $\kappa = 0$, the proposed estimator is the same as the LMMSE estimator of [4]. However, in the presence of LOS paths, it is obvious that the proposed GLMMSE channel estimator outperforms the LMMSE estimator of [4].

The NMSE of the LS, LMMSE and GLMMSE channel estimators with optimal training in the case of $\rho = 0.1$ and $\rho = 0.8$ is shown in **Figures 4** and **5**, respectively. **Figures 6** and **7** compare the NMSE of the LS and

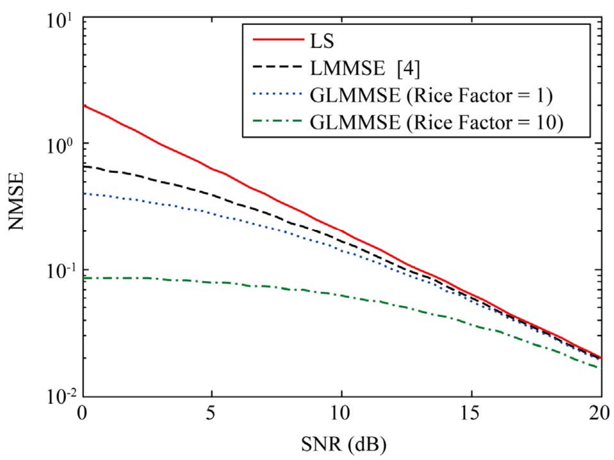

Figure 2. NMSE of the LS, LMMSE [4] and GLMMSE (κ = 1, 10) channel estimators in the case of orthogonal training signals ($N_R = N_T = 2$, $\rho = 0.1$).

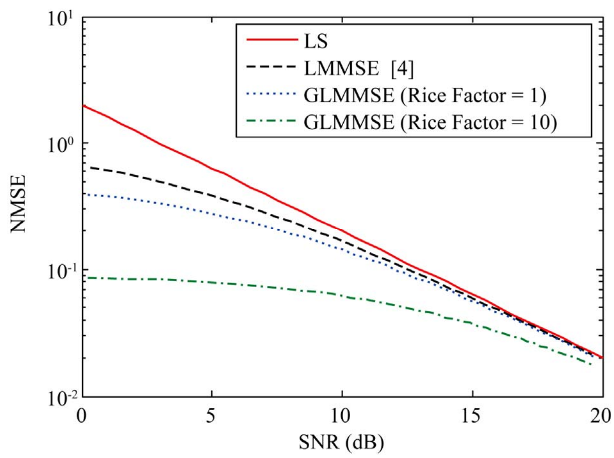

Figure 4. NMSE of the LS, LMMSE [4] and GLMMSE (κ = 1, 10) channel estimators in the case of optimal training signals ($N_R = N_T = 2$, $\rho = 0.1$).

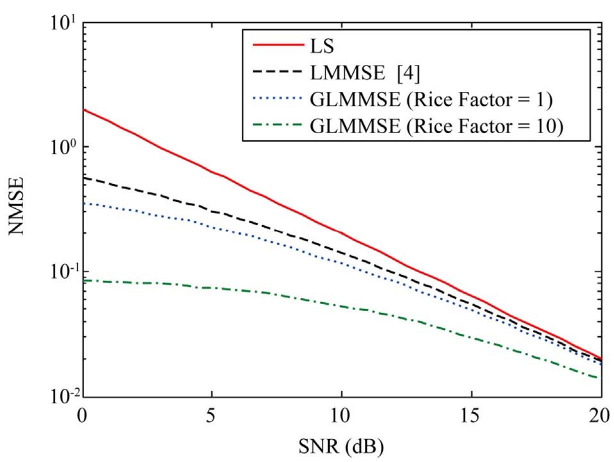

Figure 3. NMSE of the LS, LMMSE [4] and GLMMSE (κ = 1, 10) channel estimators in the case of orthogonal training signals ($N_R = N_T = 2$, $\rho = 0.8$).

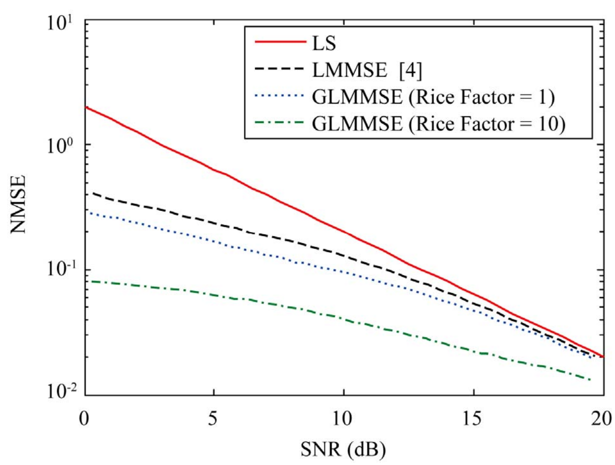

Figure 5. NMSE of the LS, LMMSE [4] and GLMMSE (κ = 1, 10) channel estimators in the case of optimal training signals ($N_R = N_T = 2$, $\rho = 0.8$).

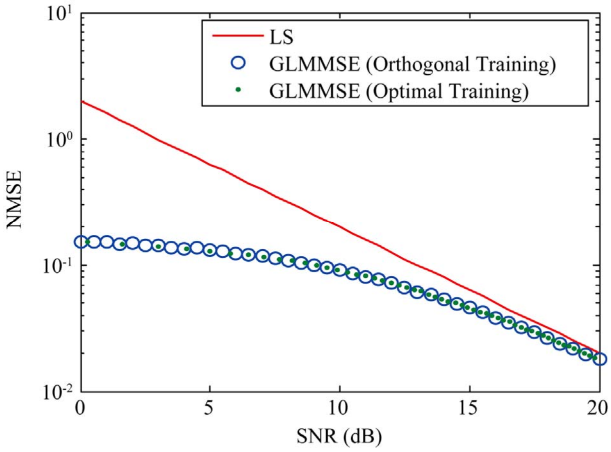

Figure 6. NMSE of the GLMMSE and LS channel estimators for κ = 5 with optimal and orthogonal training signals ($N_R = N_T = 2, \rho = 0.1$).

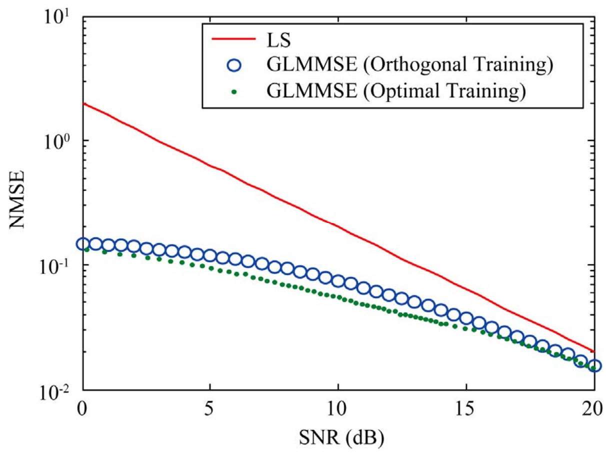

Figure 7. NMSE of the GLMMSE and LS channel estimators for κ = 5 with optimal and orthogonal training signals ($N_R = N_T = 2, \rho = 0.8$).

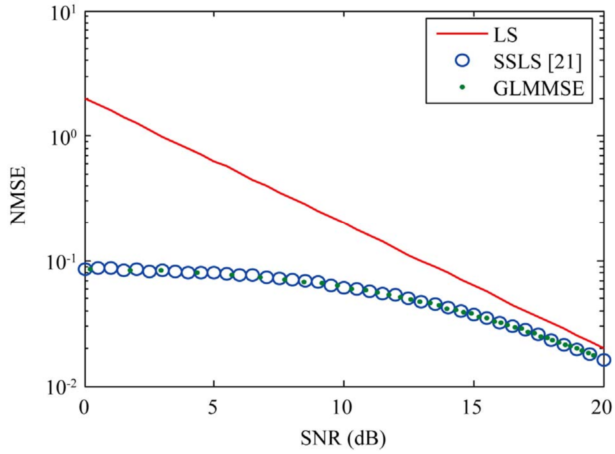

Figure 8. NMSE of the LS, SSLS [21] and GLMMSE channel estimators in the case of optimal training signals ($N_R = N_T = 2, \rho = 0.1, \kappa = 10$).

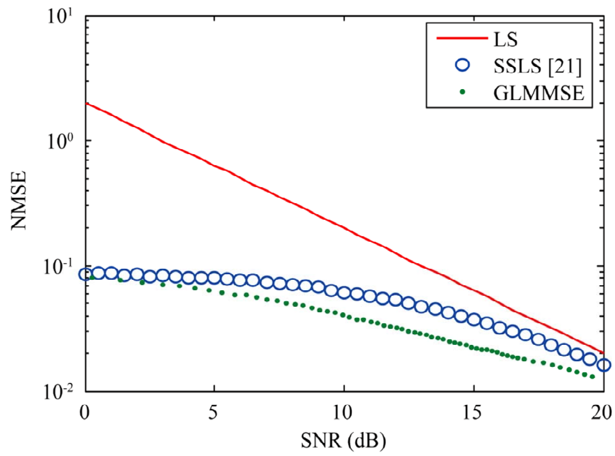

Figure 9. NMSE of the LS, SSLS [21] and GLMMSE channel estimators in the case of optimal training signals ($N_R = N_T = 2, \rho = 0.8, \kappa = 10$).

GLMMSE estimators with optimal and orthogonal training for $\kappa = 5$ in the case of $\rho = 0.1$ and $\rho = 0.8$, respectively. It is observed that at high spatial correlation the performance of the GLMMSE channel estimator with optimal training is better than that of orthogonal training. At low spatial correlation, the performance of the GLMMSE channel estimator with optimal training and orthogonal training is closely identical. In order to obtain the advantage of the optimal training sequence design, long-term statistics of the channel need to be estimated at the receiver and fed back to the transmitter. Hence, when the GLMMSE channel estimator is used to estimate MIMO channel with low spatial correlation, the transmitter has no need to the channel knowledge.

Finally, we compared the performance of the LS, SSLS of [21] and GLMMSE channel estimators in **Figures 8** and **9**. Clearly, the NMSE of the proposed channel estimator is smaller than that of the SSLS channel estimator, particularly at higher spatial correlations, because the GLMMSE estimator can employ more a priori knowledge about the channel than the SSLS estimator. It is notable that the NMSE of the SSLS channel estimator is independent of ρ.

6. Conclusions

We have proposed a new channel estimator (GLMMSE) that is suitable for spatially correlated Rician fading MIMO channel estimation. This estimator has better performance than the SSLS estimator of [21] and LMMSE estimator of [4]. Analytical and numerical results confirm the superiority of the GLMMSE estimator in the mentioned channel model. It is demonstrated that increasing κ and/or ρ decreases the NMSE of the offered estimator. Hence, to obtain the given value of MSE, the required SNR can be reduced in the Rician channel estimation. Clearly, increasing the number of antennas in MIMO systems leads to decreasing the performance of estimators. In the Rician fading MIMO channel, the unfavorable effect of increasing the number of antennas on the performance of GLMMSE channel estimator can be compensated. In other words, for the given values of SNR and MSE, the number of antennas possibly increases. Therefore, the Rician fading MIMO channels result in a higher capacity than the Rayleigh fading MIMO channels without increasing MSE. Moreover, training length can be reduced in the presence of the spatially correlated channel and/or Rician model to improve the bandwidth efficiency without increasing MSE. It is noteworthy that Rician fading is known as a more appropriate model for wireless environments with a dominant direct LOS path; and, in the microcellular mobile systems, this model is better than the Rayleigh one.

7. References

[1] D. Tse and P. Viswanath, "Fundamentals of Wireless Communication," Cambridge University Press, Cambridge, 2005.

[2] I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transactions on Telecommunications*, Vol. 10, No. 6, November 1999, pp. 585-595.

[3] S. K. Jayaweera and H. V. Poor, "On the Capacity of Multiple-Antenna Systems in Rician Fading," *IEEE Transactions on Wireless Communications*, Vol. 4, No. 3, May 2005, pp. 1102-1111.

[4] M. Biguesh and A. B. Gershman, "Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals," *IEEE Transactions on Signal Processing*, Vol. 54, No. 3, March 2006, pp. 884-893.

[5] X. Ma, L. Yang and G. B. Giannakis, "Optimal Training for MIMO Frequency-Selective Fading Channels," *IEEE Transactions on Wireless Communications*, Vol. 4, No. 2, March 2005, pp. 453-466.

[6] G. Leus and A.-J. von der Veen, "Optimal Training for ML and LMMSE Channel Estimation in MIMO Systems," *Proceedings of* 13*th IEEE Workshop on Statistical Signal Processing*, Bordeaux, 17-20 July 2005, pp. 1354-1357.

[7] H. Vikalo, B. Hassibi, B. Hochwald and T. Kailath, "On the Capacity of Frequency-Selective Channels in Training-Based Transmission Schemes," *IEEE Transactions on Signal Processing*, Vol. 52, No. 9, September 2004, pp. 2572-2583.

[8] S. A. Yang and J. Wu, "Optimal Binary Training Sequence Design for Multiple-Antenna Systems over Dispersive Fading Channels," *IEEE Transactions on Vehicular Technology*, Vol. 51, No. 5, September 2002, pp. 1271-1276.

[9] W. Yuan, P. Wang and P. Fan, "Performance of Multi-Path MIMO Channel Estimation Based on ZCZ Training Sequences," *Proceedings of IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, Vol. 2, Beijing, 8-12 August 2005, pp. 1542-1545.

[10] S. Wang and A. Abdi, "Aperiodic Complementary Sets of Sequences-Based MIMO Frequency Selective Channel Estimation," *IEEE Communication Letters*, Vol. 9, No. 10, October 2005, pp. 891-893.

[11] S. Wang and A. Abdi, "Low-Complexity Optimal Estimation of MIMO ISI Channels with Binary Training Sequences," *IEEE Signal Processing Letters*, Vol. 13, No. 11, November 2006, pp. 657-660.

[12] S. Wang and A. Abdi, "MIMO ISI Channel Estimation Using Uncorrelated Golay Complementary Sets of Poly-Phase Sequences," *IEEE Transactions on Vehicular Technology*, Vol. 56, No. 5, September 2007, pp. 3024-3039.

[13] H. M. Wang, X. Q. Gao, B. Jiang, X. H. You and W. Hong, "Efficient MIMO Channel Estimation Using Complementary Sequences," *IET Communications*, Vol. 1, No. 5, October 2007, pp. 962-969.

[14] W. Dong, J. Li and Z. Lu, "Parameter Estimation for Correlated MIMO Channels with Frequency-Selective Fading," *Wireless Personal Communications*, Vol. 52, No. 4, March 2010, pp. 813-828.

[15] J. Pang, J. Li, L. Zhao and Z. Lu, "Optimal Training Sequences for MIMO Channel Estimation with Spatial Correlation," *Proceedings of* 66*th IEEE Vehicular Technology Conference*, Baltimore, 30 September-3 October 2007, pp. 651-655.

[16] J. Pang, J. Li, L. Zhao and Z. Lu, "Optimal Training Sequences for Frequency-Selective MIMO Correlated Fading Channels," *Proceedings of* 21*st IEEE International Conference on Advanced Information Networking and Applications*, Niagara Falls, 21-23 May 2007, pp. 820-824.

[17] M. Kiessling, J. Speidel and Y. Chen, "MIMO Channel Estimation in Correlated Fading Environments," *Proceedings of* 58*th IEEE Vehicular Technology Conference*, Orlando, Vol. 2, 6-9 October 2003, pp. 1187-1191.

[18] G. Xie, X. Fang, A. Yang and Y. Liu, "Channel Estimation with Pilot Symbol and Spatial Correlation Information," *Proceedings of IEEE International Symposium on Communications and Information Technologies*, Sydney, 16-19 October 2007, pp. 1003-1006.

[19] E. Björnson and B. Ottersten, "Training-Based Bayesian MIMO Channel and Channel Norm Estimation," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Taipei, 19-24 April 2009, pp. 2701-2704.

[20] D. Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn, "Fading Correlation and Its Effect on the Capacity of Multi-Element Antenna Systems," *IEEE Transactions on Communications*, Vol. 48, No. 3, March 2000, pp. 502-513.

[21] H. Nooralizadeh and S. S. Moghaddam, "A Novel Shifted Type of SLS Estimator for Estimation of Rician Flat

Fading MIMO Channels," *Signal Processing*, Vol. 90, No. 6, June 2010, pp. 1887-1894.

[22] L. Huang, G. Mathew and J. W. M. Bergmans, "Pilot-Aided Channel Estimation for Systems with Virtual Carriers," *Proceedings of IEEE International Conference on Communications*, Istanbul, 11-15 June 2006, pp. 3070-3075.

[23] S. M. Kay, "Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory," Prentice-Hall, Upper Saddle River, 1993.

[24] K. Werner and M. Jansson, "Estimating MIMO Channel Covariances from Training Data under the Kronecker Model," *Signal Processing*, Vol. 89, No. 1, January 2009, pp. 1-13.

[25] C. Mehlführer and M. Rupp, "Novel Tap-Wise LMMSE Channel Estimation for MIMO W-CDMA," *Proceedings of IEEE Global Telecommunications Conference*, New Orleans, 30 November-4 December 2008, pp. 1-5.

Appendix

Proof of Equation (24):

Using (22), the MSE (23) can be written as follows:

$$\begin{aligned} J_{GLMMSE} &= E\left\{\left\|\mathbf{H}-\mathbf{M}-(\mathbf{Y}-\mathbf{MX})\mathbf{A}_{\circ}\right\|_F^2\right\} \\ &= E\left\{tr\left\{\left[\mathbf{H}-\mathbf{M}-\left(\mathbf{Y}-\mathbf{MX}\right)\mathbf{A}_{\circ}\right]^H \right.\right. \\ &\quad \left.\left.\times\left[\mathbf{H}-\mathbf{M}-\left(\mathbf{Y}-\mathbf{MX}\right)\mathbf{A}_{\circ}\right]\right\}\right\} \end{aligned} \quad \text{(A-1)}$$

With some calculations, the MSE (A-1) is given by

$$\begin{aligned} J_{GLMMSE} &= tr\left\{\left(\mathbf{I}_{N_T}-\mathbf{A}_{\circ}^H\mathbf{X}^H\right)\mathbf{C}_H\left(\mathbf{I}_{N_T}-\mathbf{X}\mathbf{A}_{\circ}\right)\right\} \\ &\quad +\sigma_n^2 N_R\, tr\left\{\mathbf{A}_{\circ}^H\mathbf{A}_{\circ}\right\} \end{aligned} \quad \text{(A-2)}$$

The optimal $\mathbf{A}_{\circ}$ can be found from

$$\frac{\partial J_{GLMMSE}}{\partial \mathbf{A}_{\circ}} = -\mathbf{X}^T\mathbf{C}_H + \mathbf{X}^T\mathbf{C}_H\,\mathbf{X}^*\mathbf{A}_{\circ}^* + \sigma_n^2 N_R \mathbf{A}_{\circ}^* = 0 \quad \text{(A-3)}$$

where $(\cdot)^T$ denotes the matrix transpose. Finally, we have

$$\mathbf{A}_{\circ} = \left(\mathbf{X}^H\mathbf{C}_H\,\mathbf{X} + \sigma_n^2 N_R \mathbf{I}_{N_P}\right)^{-1}\mathbf{X}^H\mathbf{C}_H \quad \text{(A-4)}$$

