Advantages of Multiple -Estimation in the Frequency-Selective Block Fading MIMO Environments

Hamid Nooralizadeh
Department of Electrical and Computer Engineering, Islamshahr Branch, Islamic Azad University, Tehran, Iran
E-mail addresses: nooralizadeh@iiau.ac.ir; h_n_alizadeh@yahoo.com(Corresponding author)

Received: Jan. 2015 Revised: Feb. 2015 Accepted: March 2015

ABSTRACT:
The performance of the multiple-estimation (ME) is examined in multiple-input multiple-output (MIMO) frequency-selective fading channels. The least squares (LS) technique, the shifted scaled least squares (SSLS) estimator, and the minimum mean square error (MMSE) estimator were probed in this paper. In the uniform and non-uniform power allocation, the closed form equations were obtained for total mean square error (TMSE) of the estimators. Analytical and numerical results showed that the LS estimator has lower error in the case of ME than single-estimation (SE). Moreover, it was seen that the performance of SSLS and MMSE channel estimators in the ME case is better than SE particularly at high signal to noise ratios (SNRs). It was shown that for small numbers of sub-blocks used for channel estimation, the SSLS and MMSE channel estimators are better than LS. However for large numbers of sub-blocks, inversely, the LS channel estimator is better than SSLS and MMSE. The un-equal power allocation was also examined analytically and numerically. Simulation results showed that exponential power allocation is proper for SSLS channel estimator in ME case.

KEYWORDS: Frequency selective fading, Multiple estimation, Multiple-input multiple-output, Least squares, Shifted scaled least squares, Minimum mean square error

1. INTRODUCTION

Multiple-input multiple-output (MIMO) technologies provide substantial benefits in improving the achievable capacity of the system and/or quality of service [1, 2]. The system's ability to approach the MIMO capacity heavily relies on the channel state information (CSI). In the coherent receivers [1], channel equalizers [3], transmit beamformers [4] and the perfect knowledge of the channel is usually needed.

Training-based channel estimation (TBCE) is widely used in practice for quasi-static or slow fading channels, e.g., indoor MIMO channels [5-11]. However, in outdoor MIMO channels where channels are under fast fading, the channel tracking and estimating algorithms as the wiener least mean squares (W-LMS) [12], Kalman filter [13, 14], recursive least squares (RLS) [15], generalized RLS (GRLS) [16], and generalized LMS (GLMS) [17] are used.

Using TBCE, it is shown in [5], [7] that the minimum mean square error (MMSE) channel estimator has the best performance because it employs more a-priori knowledge about the channel. For MIMO Rician flat fading channels, the new shifted scaled least squares (SSLS) channel estimator is presented in [8]. It is seen that this estimator has the best performance among the LS-based estimators in Rician channel model. Nevertheless, the MMSE channel estimator has lower error than that of SSLS in Rician fading channel model especially at high signal to noise ratios (SNRs) and spatial correlations [7].

In [9], the performances of the time-multiplexed (TM) and superimposed (SI) schemes have been compared in MIMO channel estimation. It is shown that in fast fading channels and/or for many receiver antennas, the SI scheme is better than TM but in other cases this scheme suffers from a higher estimation error. In part II of this paper [10], to improve the performance of the SI scheme, a decision directed approach is applied.

In [11], the problem of training optimization for estimating a MIMO flat fading channel in the presence of spatially and temporally correlated Gaussian noise is studied in an application-oriented setup. For the task of training sequence design in MIMO systems, a more general framework is introduced that can treat not only the minimization of channel estimator's mean square error (MSE) but also the optimization of a final performance metric of interest related to the use of the channel estimate in the communication system.

In order to perform the individual channel estimation at the destination, in [18], the SI training strategy is applied into the MIMO amplify-and-forward (AF) one-way relay network (OWRN). The discussion is restricted to the case of a slow, frequency-flat block fading model. A specific suboptimal channel estimation algorithm is applied in [18] using the optimal training sequences and to verify the Bayesian Cramér-Rao lower bound (CRLB) results, the normalized MSE performance for the estimation is provided.

In order to estimate MIMO frequency selective or MIMO inter-symbol interference (ISI) channels, training sequences should have both good autocorrelations and cross correlations. Furthermore, to separate the transmitted data and training symbols, one of the zero-padding (ZP) based guard period or cyclic prefix (CP) based guard period is inserted. In [19], a set of sequences with a zero correlation zone (ZCZ) is employed as optimal training signals. In [20], a novel transmit diversity scheme applicable to frequency selective channel is proposed. It is shown that with ZCZ complementary codes, both full space diversity and full frequency diversity can be obtained. In [21], different phases of a perfect poly-phase sequence such as the Frank sequence or Chu sequence are proposed. Furthermore, in [22-25], Golay complementary sets of poly-phase sequences have been used.

In [24], the performance of the best linear unbiased estimator (BLUE) and linear minimum mean square error (LMMSE) estimator is studied in the frequency selective Rayleigh fading MIMO channel. It is observed that the LMMSE estimator has better performance than the BLUE, because it can employ statistical knowledge about the channel.

In this paper, TBCE method is studied in the frequency-selective Rician fading MIMO channels using the new multiple-estimation (ME) method. In [26], SSLS and MMSE estimators are proposed that are suitable to estimate the above-mentioned channel model. Analytical results show that the proposed estimators achieve much better minimum possible Bayesian Cramér-Rao lower bounds (CRLBs) in the frequency selective Rician MIMO channels compared with those of Rayleigh one.

In this paper, the results of [26] in the single-estimation (SE) case are extended to the ME case. Here, the multiple estimates of the channel during received N sub-blocks are combined optimally. The optimal weight coefficients are achieved for the least squares (LS), SSLS, and MMSE channel estimators. Furthermore, the minimum total mean square error (TMSE) under optimal training is obtained for all estimators. Simulation results show that all estimators have better performance in the ME case than SE case especially at high SNRs. Increasing the number of sub-blocks N results in better performance with LS estimator than SSLS especially at medium SNRs. Therefore, the SSLS and MMSE estimators are mainly appropriate for Rician frequency selective fading channels with a short coherence time (fast fading). However the LS estimator is better than SSLS for channels with a long coherence time (slow fading).

The un-equal power allocation is also considered in this paper. Using the SSLS and MMSE estimators, it is shown that in linear power allocation the results are analogous to the uniform power allocation. Nevertheless, in exponential power allocation the channel estimation errors are lower than the uniform power allocation with SSLS estimator.

The rest of this paper is organized as follows. Section 2 introduces the channel and signal model. The SE and ME methods in the Rician frequency-selective fading MIMO channels are investigated in Section 3. Simulation results are presented in Section 4. Finally, concluding remarks are presented in Section 5.

2. CHANNEL AND SIGNAL MODEL

It is considered block transmission over block non-flat Rician fading MIMO channel with N_T transmit and N_R receiver antennas. The frequency selective fading sub-channels between each pair of Tx-Rx antenna elements are modeled by L+1 taps as $\mathbf{h}_{rt} = [h_{r,t}(0)\ h_{r,t}(1)\ \dots\ h_{r,t}(L)]^T$, $\forall r \in [1, N_R]$ and $t \in [1, N_T]$. It is assumed that all sub-channels have identical power delay profile (PDP) as (b_0, b_1, …, b_L). Then, the l^{th} taps of all the sub-channels have the same power b_l, $l \in [0, L]$, i.e., $E\{|h_{r,t}(l)|^2\} = b_l$; $\forall\, l, t, r$. It is also assumed unit power for each sub-channel, i.e., $\sum_{l=0}^{L} b_l = 1$.

For Rician frequency selective fading channels, the elements of the matrix $\mathbf{H}_l$, $\forall l \in [0, L]$, are defined similar to [27, 28] in the following form:

$$\mathbf{H}_l = \sqrt{b_l \frac{\kappa}{\kappa+1}}\ \tilde{\mathbf{M}}_l + \sqrt{\frac{b_l}{\kappa+1}}\ \tilde{\mathbf{H}}_l \tag{1}$$

Where κ is the channel Rice factor. The matrices $\tilde{\mathbf{M}}_l$ and $\tilde{\mathbf{H}}_l$ describe the line of sight (LOS) and scattered components, respectively. It is assumed that the elements of $\tilde{\mathbf{M}}_l$, $\forall\, l$, are complex as $(1+j)/\sqrt{2}$ and the elements of the matrix $\tilde{\mathbf{H}}_l$, $\forall\, l$, are independently and identically distributed (i.i.d.) complex Gaussian random variables with the zero mean and the unit variance. The frequency selective fading MIMO channel can be defined as the $N_R \times N_T(L+1)$ matrix $\mathbf{H}$ = {$\mathbf{H}_0$, $\mathbf{H}_1$, …, $\mathbf{H}_L$}, where $\mathbf{H}_l$ has the following structure:

$$\mathbf{H}_l = \begin{bmatrix} h_{11}(l) & h_{12}(l) & ... & h_{1N_T}(l) \\ h_{21}(l) & h_{22}(l) & ... & h_{2N_T}(l) \\ \vdots & \vdots & ... & \vdots \\ h_{N_R1}(l) & h_{N_R2}(l) & ... & h_{N_RN_T}(l) \end{bmatrix}, \forall l \in [0, L] \quad (2)$$

Moreover, it is assumed that the elements of matrices $\tilde{\mathbf{H}}_{l_1}$ and $\tilde{\mathbf{H}}_{l_2}$, $\forall l_1, l_2$, are independent of each other. Hence, the elements of the matrix **H** are also independent of each other.

Suppose that $\mathbf{h} = vec(\mathbf{H})$. The $N_R N_T (L+1) \times N_R N_T (L+1)$, so the covariance matix of **h** can be obtained as follows:

$$\mathbf{C}_h = \mathbf{R}_h - E\{\mathbf{h}\}E\{\mathbf{h}\}^H = \mathbf{C}_\Sigma \otimes \mathbf{I}_{N_R N_T} \quad (3)$$

where

$$\mathbf{C}_\Sigma = \frac{1}{1+\kappa}\begin{bmatrix} b_\circ & 0 & 0 & ... & 0 \\ 0 & b_1 & 0 & ... & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & ... & b_L \end{bmatrix} \quad (4)$$

Each transmitted block has N sub-blocks which contain training and data symbols as shown in Figure 1. The frame structure is the same for all Tx antennas. Training and data symbols are located in the first and end part of the sub-blocks, respectively. In practice, the channel is estimated using training symbols in the training phase. Then, the results are used for data detection. In order to estimate the channel matrix **H**, the $N_P \geq N_T (L+1) + L$ symbols are transmitted from each Tx antenna. The L first symbols are CP guard period that are used to avoid the interference from symbols before the first training symbols. At the receiver, because of their pollution by data, due to interference, these symbols are discarded. Hence, by collecting the last $N_P - L$ received vectors of (1) into the $N_R \times (N_P - L)$ matrix $\mathbf{Y} = [\mathbf{y}(L+1), \mathbf{y}(L+2), \ldots, \mathbf{y}(N_P)]$, the compact matrix form of received training symbols can be represented in a linear model as

$$\mathbf{Y} = \mathbf{H}\,\mathbf{X} + \mathbf{V} \quad (5)$$

Where **X** is the $N_T (L+1) \times (N_P - L)$ training matrix. The matrix **X** is constructed by the N_P-vector of transmitted symbols in the form of $\mathbf{x}(i) = [x_1(i), x_2(i), ..., x_{N_T}(i)]^T$ as follows:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}(L+1) & \mathbf{x}(L+2) & ... & \mathbf{x}(N_P) \\ \mathbf{x}(L) & \mathbf{x}(L+1) & ... & \mathbf{x}(N_P-1) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{x}(2) & \mathbf{x}(3) & \cdots & \mathbf{x}(N_P-L+1) \\ \mathbf{x}(1) & \mathbf{x}(2) & ... & \mathbf{x}(N_P-L) \end{bmatrix} \quad (6)$$

Note that $x_t(i)$ is the transmitted symbol by the t^{th} Tx antenna at symbol time i. The matrix **V** in (5) is the complex N_R-vector of additive Rx noise. The elements of the noise matrix are i.i.d. complex Gaussian random variables with zero-mean and σ^2 variance, and have the following correlation matrix:

$$\mathbf{R}_V = E\{\mathbf{V}^H\mathbf{V}\} = \sigma^2 N_R\, \mathbf{I}_{N_P - L} \quad (7)$$

The elements of **H** and noise matrix are independent of each other. The elements of the columns of **H** have the following $N_T (L+1) \times N_T (L+1)$ co-variance matrix:

$$\begin{aligned} \mathbf{C}_H &= \mathbf{R}_H - \mathbf{M}^H\mathbf{M} = E\{\mathbf{H}^H\mathbf{H}\} - \mathbf{M}^H\mathbf{M} \\ &= N_R\,(\mathbf{C}_\Sigma \otimes \mathbf{I}_{N_T}) \end{aligned} \quad (8)$$

In a particular case, when the uniform PDP is used, i.e., $b_0 = b_1 = \ldots = b_L = 1/(L+1)$, the result is

$$\mathbf{C}_H = \frac{N_R}{(1+\kappa)(1+L)}\,\mathbf{I}_{N_T(L+1)} \quad (9)$$

3. MULTIPLE CHANNEL ESTIMATION

In order to improve the performance of the estimators, the multiple estimates of the channel during received N sub-blocks are combined. It is assumed that the channel response is fixed within N sub-blocks. In other words, the coherent time of the channel is enough to use N sub-blocks for channel estimation. Such a channel is proper for indoor MIMO channels with low mobility. Suppose that N estimates $\hat{\mathbf{H}}_1, ..., \hat{\mathbf{H}}_N$ of the MIMO channel are obtained based on the training matrices $\mathbf{X}_1, ..., \mathbf{X}_N$, respectively. The results are combined in the following linear method:

$$\hat{\mathbf{H}}_{\text{ME}} = \sum_{n=1}^{N} a_n\, \hat{\mathbf{H}}_n \quad (10)$$

Where the optimal weight coefficients $a_1, ..., a_N$ are obtained so that the TMSE (11) is minimized subject to $\sum_{n=1}^{N} a_n = 1$.

$$J_{ME} = \mathrm{E}\left\{ \left\| \mathbf{H} - \sum_{n=1}^{N} a_n\, \hat{\mathbf{H}}_n \right\|_F^2 \right\} \quad (11)$$

Then, the optimization problem is

$$\min_{a_1,...,a_N} \mathrm{E}\left\{ \left\| \mathbf{H} - \sum_{n=1}^{N} a_n\, \hat{\mathbf{H}}_n \right\|_F^2 \right\} \quad S\,T \quad \sum_{n=1}^{N} a_n = 1 \quad (12)$$

In this section, the problem (12) will be solved considering the LS, the SSLS, and the MMSE channel estimators.

3.1. Multiple LS (MLS) estimator

For linear model of (5), the LS estimator which minimizes $tr\{(\mathbf{Y}\text{-}\mathbf{HX})^H(\mathbf{Y}\text{-}\mathbf{HX})\}$ is [29]:

$$\hat{\mathbf{H}}_{LS} = \mathbf{Y}\mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1} \quad 1\square$$

Under optimal training, it is shown that the error of the estimator is minimized as follows [26]

$$(J_{LS})_{\min} = \frac{\sigma^2 (N_T (L+1))^2 N_R}{P} \tag{14}$$

Where P is a given constant value as the total power of training matrix **X**. Using (5), the LS estimator (13) can be rewritten as

$$\hat{\mathbf{H}}_{LS} = \mathbf{H} + \mathbf{V}\mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1} \tag{15}$$

Using (15) and the constraint $\sum_{n=1}^{N} a_n = 1$, the error of the MLS estimation will be written as

$$\begin{aligned} J_{MLS} &= \mathrm{E}\left\{ \left\| \mathbf{H} - \sum_{n=1}^{N} a_n \hat{\mathbf{H}}_n \right\|_F^2 \right\} \\ &= \mathrm{E}\left\{ \left\| \mathbf{H} - \sum_{n=1}^{N} a_n (\mathbf{H} + \mathbf{V}_n \mathbf{X}_n^H (\mathbf{X}_n \mathbf{X}_n^H)^{-1}) \right\|_F^2 \right\} \\ &= \mathrm{E}\left\{ \left\| \sum_{n=1}^{N} a_n \mathbf{V}_n \mathbf{X}_n^H (\mathbf{X}_n \mathbf{X}_n^H)^{-1} \right\|_F^2 \right\} \\ &= \mathrm{E}\{tr\{(\sum_{n=1}^{N} a_n \mathbf{V}_n \mathbf{X}_n^H \mathbf{E}_n)^H (\sum_{m=1}^{N} a_m \mathbf{V}_m \mathbf{X}_m^H \mathbf{E}_m)\}\} \\ &= tr\{\sum_{n=1}^{N}\sum_{m=1}^{N} a_n^* a_m \mathbf{E}_n \mathbf{X}_n \mathrm{E}\{\mathbf{V}_n^H \mathbf{V}_m\} \mathbf{X}_m^H \mathbf{E}_m\} \\ &= \sigma^2 N_R \, tr\{\sum_{n=1}^{N} |a_n|^2 \mathbf{E}_n\} \end{aligned} \tag{16}$$

Where $\mathbf{E}_n = (\mathbf{X}_n \mathbf{X}_n^H)^{-1}$ and the latter one is obtained using the following equation:

$$\mathrm{E}\{\mathbf{V}_n^H \mathbf{V}_m\} = \begin{cases} \sigma^2 N_R \mathbf{I}_{N_P - L} & ; \quad n = m \\ \mathbf{0} & ; \quad n \neq m \end{cases} \tag{17}$$

Then, for MLS estimation, the problem (12) can be written as

$$\min_{a_1,\dots,a_N} tr\left\{ \sum_{n=1}^{N} |a_n|^2 \mathbf{E}_n \right\} \quad ST \quad \sum_{n=1}^{N} a_n = 1 \tag{18}$$

The LS estimator is unbiased. The constraint in (18) guarantees that the MLS estimation is also unbiased.
To solve (18), the Lagrange multiplier method is used. The problem can be written as

$$L(a_1,\dots,a_N,\eta) = tr\left\{ \sum_{n=1}^{N} |a_n|^2 \mathbf{E}_n \right\} + \eta \left\{ \sum_{n=1}^{N} a_n - 1 \right\} \tag{19}$$

To find $a_1,\dots,a_N$, the partial derivatives of (19) with respect to $a_i \ (i = 1,2,\dots,N)$ are computed. Then, the results are set equal to zero. Finally, the optimal weight coefficients in the MLS estimation are obtained from [5]:

$$a_n = \frac{1}{tr\{\mathbf{E}_n\} \sum_{l=1}^{N} 1/tr\{\mathbf{E}_l\}} \quad ; \quad n = 1,\dots,N \tag{20}$$

It is straightforward to show that under optimal training for LS estimator

$$tr\{\mathbf{E}_n\} = tr\{(\mathbf{X}_n \mathbf{X}_n^H)^{-1}\} = tr\{(N_T(L+1)/P_n)\mathbf{I}_{N_T(L+1)}\} = \frac{(N_T(L+1))^2}{P_n} \tag{21}$$

Where P_n is the total power of training matrix $\mathbf{X}_n$ which is used during the training phase in the sub-block n. Suppose that $P_n = k_n P$ is the transmitted power during the n-th $(n = 1,\dots,N)$ training period and $P_{tot} = \sum_{n=1}^{N} P_n = N P$ is the total transmitted power during the N training periods. Then $\sum_{n=1}^{N} k_n = N$ and using (21), the optimal weight coefficients (20) can be rewritten as

$$a_n = \frac{1}{((N_T(L+1))^2 / k_n P) \sum_{l=1}^{N} (P_l / (N_T(L+1))^2)} = \frac{k_n P}{\sum_{l=1}^{N} P_l} = \frac{k_n}{N} \tag{22}$$

Using (21) and (22), under optimal training, the TMSE (16) is minimized as follows

$$J_{MLS_{(\min)}} = \sigma^2 N_R \frac{(N_T(L+1))^2}{P N^2} \sum_{n=1}^{N} k_n = \frac{\sigma^2 N_R (N_T(L+1))^2}{P N} \tag{23}$$

Comparing (23) and (14), it is seen that in the MLS estimation the error reduces by the number of sub-blocks N which is used for channel estimation. It is notable that the error (23) is independent of P_n, the transmitted power during the n-th training period. It means that for uniform training powers and non-uniform training powers during N training periods, the error is the same.

3.2. Multiple SSLS (MSSLS) estimator

Consider (5), the SSLS channel estimator can be expressed in the following form [8]

$$\hat{\mathbf{H}}_{SSLS} = \gamma \hat{\mathbf{H}}_{LS} + (1-\gamma)\mathbf{M} \tag{24}$$

$$\gamma = \frac{tr\{\mathbf{C_H}\}}{tr\{\mathbf{C_H}\} + J_{LS}} \tag{25}$$

Generally speaking, the scaling factor in (25) is between 0 and 1. When the channel fading is weak ($\kappa \to \infty$ or AWGN) or the transmitted power is small, i.e., $tr\{\mathbf{C_H}\} << J_{LS}$, the scaling factor $\gamma \to 0$. Also, when the channel fading is strong ($\kappa \to 0$ or Rayleigh) or the transmitted power is large, i.e., $tr\{\mathbf{C_H}\} >> J_{LS}$, the scaling factor $\gamma \to 1$. Finally, in the Rician fading channel ($0 < \kappa < \infty$), we have $0 < \gamma < 1$.
According to [8], optimal training for LS and SSLS estimators is identical. Under optimal training, the TMSE minimizes as follows:

$$(J_{SSLS})_{\min} = \frac{\sigma^2 N_R N_T^2 (L+1)^2 tr\{\mathbf{C_H}\}}{P\, tr\{\mathbf{C_H}\} + \sigma^2 N_R N_T^2 (L+1)^2} \tag{26}$$

Using (15), the SSLS channel estimator (24) can be rewritten as

$$\hat{\mathbf{H}}_{SSLS} = \gamma\mathbf{H} + \gamma\mathbf{V}\mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1} + (1-\gamma)\mathbf{M} \tag{27}$$

Using (27), the TMSE of MSSLS estimator is expressed as

$$\begin{aligned} J_{MSSLS} &= \mathrm{E}\left\{\left\|\mathbf{H} - \sum_{n=1}^{N} a_n\,\hat{\mathbf{H}}_n\right\|_F^2\right\} \\ &= \mathrm{E}\left\{\left\|\mathbf{H} - \sum_{n=1}^{N} a_n(\gamma_n\mathbf{H} + \gamma_n\mathbf{V}_n\mathbf{X}_n^H\mathbf{E}_n + (1-\gamma_n)\mathbf{M})\right\|_F^2\right\} \\ &= \mathrm{E}\left\{\left\|(1-\sum_{n=1}^{N} a_n\gamma_n)\mathbf{H} - \sum_{n=1}^{N} a_n\gamma_n\mathbf{V}_n\mathbf{X}_n^H\mathbf{E}_n - \sum_{n=1}^{N} a_n(1-\gamma_n)\mathbf{M}\right\|_F^2\right\} \\ &= \mathrm{E}\left\{\left\|(1-\sum_{n=1}^{N} a_n\gamma_n)(\mathbf{H}-\mathbf{M}) - \sum_{n=1}^{N} a_n\gamma_n\mathbf{V}_n\mathbf{X}_n^H\mathbf{E}_n\right\|_F^2\right\} \\ &= \mathrm{E}\{tr\{((1-\sum_{n=1}^{N} a_n^*\gamma_n^*)(\mathbf{H}-\mathbf{M})^H - \sum_{n=1}^{N} a_n^*\gamma_n^*\mathbf{E}_n\mathbf{X}_n\mathbf{V}_n^H) \\ &\quad \times((1-\sum_{n=1}^{N} a_n\gamma_n)(\mathbf{H}-\mathbf{M}) - \sum_{n=1}^{N} a_n\gamma_n\mathbf{V}_n\mathbf{X}_n^H\mathbf{E}_n)\}\} \end{aligned} \tag{28}$$

Using (7), (17), $\sum_{n=1}^{N} a_n = 1$, and with some calculations the result is

$$\begin{aligned} J_{MSSLS} &= (1-\sum_{n=1}^{N} a_n\gamma_n)(1-\sum_{n=1}^{N} a_n^*\gamma_n^*)tr\{\mathbf{C_H}\} \\ &\quad + \sigma^2 N_R \sum_{n=1}^{N} |a_n|^2|\gamma_n|^2 tr\{\mathbf{E}_n\} \end{aligned} \tag{29}$$

The optimization problem is

$$\min_{a_1,\ldots,a_N} (1-\sum_{n=1}^{N} a_n\gamma_n)(1-\sum_{n=1}^{N} a_n^*\gamma_n^*)tr\{\mathbf{C_H}\} + \sigma^2 N_R \sum_{n=1}^{N} |a_n|^2|\gamma_n|^2 tr\{\mathbf{E}_n\} \quad ST \quad \sum_{n=1}^{N} a_n = 1 \tag{30}$$

The SSLS estimator is biased. The constraint in (25) results in that the MSSLS estimation is also biased. Using the Lagrange multiplier method, the result is

$$\begin{aligned} L(a_1,\ldots,a_N,\eta) &= (1-\sum_{n=1}^{N} a_n\gamma_n)(1-\sum_{n=1}^{N} a_n^*\gamma_n^*)tr\{\mathbf{C_H}\} \\ &\quad + \sigma^2 N_R \sum_{n=1}^{N} |a_n|^2|\gamma_n|^2 tr\{\mathbf{E}_n\} + \eta\left\{\sum_{n=1}^{N} a_n - 1\right\} \end{aligned} \tag{31}$$

By differentiating (31) with respect to a_i $(i = 1,2,\ldots,N)$ and setting the results equal to zero the result is

$$-\gamma_i\,(1-\sum_{n=1}^{N} a_n^*\gamma_n^*)tr\{\mathbf{C_H}\} + \sigma^2 N_R\, a_i^*|\gamma_i|^2 tr\{\mathbf{E}_i\} + \eta = 0 \tag{32}$$

In general, equation (32) cannot be solved analytically. Nevertheless, in the uniform power allocation $P_1 = \ldots = P_N = P_{tot}/N = P$ where $\gamma_1 = \ldots = \gamma_N = \gamma$ and $\mathbf{E}_1 = \ldots = \mathbf{E}_N = \mathbf{E}$, (32) can be rewritten as:

$$-\gamma(1-\gamma^*)tr\{\mathbf{C_H}\} + \sigma^2 N_R\, a_i^*|\gamma|^2 tr\{\mathbf{E}\} + \eta = 0 \tag{33}$$

Using $\sum_{n=1}^{N} a_n = 1$, the result for Lagrange multiplier will be

$$\eta = \gamma(1-\gamma^*)tr\{\mathbf{C_H}\} - \sigma^2 N_R \frac{|\gamma|^2}{N} tr\{\mathbf{E}\} \tag{34}$$

Substituting (34) back into (33), the result is

$$a_n = \frac{1}{N} \;;\quad n = 1,\ldots,N \tag{35}$$

Using (21), (25), and (35), it is shown that under optimal training the TMSE (29) is minimized as

$$\begin{aligned} J_{MSSLS\,(\min)} &= \left(\frac{\sigma^2 N_R N_T^2(1+L)^2}{P\,tr\{\mathbf{C_H}\} + \sigma^2 N_R N_T^2(1+L)^2}\right)^2 tr\{\mathbf{C_H}\} \\ &\quad + \frac{\sigma^2 N_R N_T^2(1+L)^2 P}{N}\left(\frac{tr\{\mathbf{C_H}\}}{P\,tr\{\mathbf{C_H}\} + \sigma^2 N_R N_T^2(1+L)^2}\right)^2 \end{aligned} \tag{36}$$

When N =1, (36) reduces to the special case of (26) for single channel estimation with the SSLS estimator. It is seen that the second term in (36) decreases when the number of sub-blocks N increases. In the non-uniform power allocation,

$P_n = k_n P$, $P_{tot} = \sum_{n=1}^{N} P_n = N\,P$, $\sum_{n=1}^{N} k_n = N$, suppose that $a_n = k_n / N$ for n = 1, 2, …, N. With some calculations, the TMSE (29) is minimized in this case as:

$$\begin{aligned} J_{MSSLS\,(\min)} &= tr\{\mathbf{C_H}\}\left(1 - P\frac{tr\{\mathbf{C_H}\}}{N}\sum_{n=1}^{N}\frac{k_n^2}{k_n P\,tr\{\mathbf{C_H}\} + \sigma^2 N_R N_T^2(1+L)^2}\right)^2 \\ &\quad + P\,\sigma^2 N_R N_T^2(1+L)^2\left(\frac{tr\{\mathbf{C_H}\}}{N}\right)^2 \sum_{n=1}^{N}\frac{k_n^3}{(k_n P\,tr\{\mathbf{C_H}\} + \sigma^2 N_R N_T^2(1+L)^2)^2} \end{aligned} \tag{37}$$

When $k_n = 1$, (37) reduces to (36).

3.3. Multiple MMSE (MMMSE) estimator

For linear model of (5), the MMSE channel estimator of **H** is given by [26]

$$\hat{\mathbf{H}}_{MMSE} = \mathbf{M} + (\mathbf{Y} - \mathbf{MX})\,\mathbf{A} \tag{38}$$

Where

$$\mathbf{A} = (\mathbf{X}^H\mathbf{C}_H\,\mathbf{X} + \sigma^2 N_R\,\mathbf{I}_{N_P-L})^{-1}\mathbf{X}^H\mathbf{C}_H \tag{39}$$

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

$$\mathbf{C}_{\boldsymbol{\varepsilon}} = \mathbf{R}_{\boldsymbol{\varepsilon}} = \mathrm{E}\{\boldsymbol{\varepsilon}^H\boldsymbol{\varepsilon}\} = (\mathbf{C_H}^{-1} + \frac{1}{\sigma^2 N_R}\mathbf{X}\mathbf{X}^H)^{-1} \tag{40}$$

Then, the MMSE estimation error is given by

$$J_{MMSE} = \mathrm{E}\left\{ \left\| \mathbf{H} - \hat{\mathbf{H}}_{MMSE} \right\|_F^2 \right\} = \mathrm{E}\{tr(\boldsymbol{\varepsilon}^H \boldsymbol{\varepsilon})\}$$
$$= tr\{(\mathbf{C}_\mathbf{H}^{-1} + \frac{1}{\sigma^2 N_R}\mathbf{X}\mathbf{X}^H)^{-1}\} \tag{41}$$

To minimize (41) subject to the transmitted power constraint $tr\{\mathbf{X}\mathbf{X}^H\} = P$, the Lagrange multiplier method is used. The problem can be written as follows:

$$L(\mathbf{X}\mathbf{X}^H, \eta) = tr\{(\mathbf{C}_\mathbf{H}^{-1} + \frac{1}{\sigma^2 N_R}\mathbf{X}\mathbf{X}^H)^{-1}\} + \eta\,[tr\{\mathbf{X}\mathbf{X}^H\} - P] \tag{42}$$

where η is the Lagrange multiplier. By differentiating (42) with respect to $\mathbf{X}$ and setting the result equal to zero, it is obtained that the optimal training matrix should satisfy the following equation:

$$\mathbf{X}\mathbf{X}^H == \frac{P + \sigma^2 N_R\, tr\{C_H^{-1}\}}{N_T}\mathbf{I}_{N_T(L+1)} - \sigma^2 N_R \mathbf{C}_\mathbf{H}^{-1} \tag{43}$$

Substituting (43) back into (41), the TMSE will be minimized as

$$(J_{MMSE})_{\min} = \frac{N_T^2 (L+1)}{(P/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}} \tag{44}$$

Using (5) and (38), the MMSE channel estimator can be rewritten as

$$\hat{\mathbf{H}}_{MMSE} = \mathbf{M} + (\mathbf{H} - \mathbf{M})\mathbf{X}\mathbf{A} + \mathbf{V}\mathbf{A} \tag{45}$$

Using (11) and (45), the TMSE of MMMSE channel estimator is expressed as

$$J_{MMMSE} = \mathrm{E}\left\{ \left\| \mathbf{H} - \sum_{n=1}^{N} a_n \hat{\mathbf{H}}_n \right\|_F^2 \right\}$$
$$= \mathrm{E}\left\{ \left\| \mathbf{H} - \sum_{n=1}^{N} a_n (\mathbf{M} + (\mathbf{H}-\mathbf{M})\,\mathbf{X}_n\mathbf{A}_n + \mathbf{V}_n\mathbf{A}_n) \right\|_F^2 \right\}$$
$$= \mathrm{E}\left\{ \left\| (\mathbf{H}-\mathbf{M})(\mathbf{I}_{N_P-L} - \sum_{n=1}^{N} a_n\,\mathbf{X}_n\mathbf{A}_n) - \sum_{n=1}^{N} a_n \mathbf{V}_n\mathbf{A}_n \right\|_F^2 \right\}$$
$$= tr\{(\mathbf{I}_{N_P-L} - \sum_{n=1}^{N} a_n^* \mathbf{A}_n^H \mathbf{X}_n^H)\mathbf{C}_\mathbf{H}(\mathbf{I}_{N_P-L} - \sum_{m=1}^{N} a_m\,\mathbf{X}_m\mathbf{A}_m)$$
$$+ \sum_{n=1}^{N}\sum_{m=1}^{N} a_n^* a_m \mathbf{A}_n^H\, \mathrm{E}\{\mathbf{V}_n^H \mathbf{V}_m\}\mathbf{A}_m\} \tag{46}$$

Where

$$\mathbf{A}_n = (\mathbf{X}_n^H \mathbf{C}_H\, \mathbf{X}_n + \sigma^2 N_R\, \mathbf{I}_{N_P-L})^{-1}\mathbf{X}_n^H \mathbf{C}_H \tag{47}$$

Using (17), (47), and with some calculations, the TMSE (46) can be expressed as

$$J_{MMMSE} = tr\{\mathbf{C}_\mathbf{H}\} - \sum_{n=1}^{N} a_n\, tr\{\mathbf{C}_\mathbf{H}\mathbf{X}_n\mathbf{A}_n\}$$
$$+ \sum_{n=1}^{N} (|a_n|^2 - a_n^*)\, tr\{\mathbf{A}_n^H \mathbf{X}_n^H \mathbf{C}_\mathbf{H}\} \tag{48}$$
$$+ \sum_{m=1}^{N}\sum_{\substack{n=1\\ n\neq m}}^{N} a_n^* a_m\, tr\{\mathbf{A}_n^H \mathbf{X}_n^H \mathbf{C}_\mathbf{H}\mathbf{X}_m\,\mathbf{A}_m\}$$

The optimization problem is

$$\min_{a_1,\dots,a_N} J_{MMMSE} \quad S.T \quad \sum_{n=1}^{N} a_n = 1 \tag{49}$$

The MMSE estimator is biased. The constraint in (49) results in that the multiple MMSE estimation is also biased. The Lagrange multiplier method is used as

$$L(a_1,\dots,a_N,\eta) = J_{MMMSE} + \eta\left\{\sum_{n=1}^{N} a_n - 1\right\} \tag{50}$$

The partial derivatives of (50) are obtained with respect to $a_i\ (i = 1,2,\dots,N)$, then, the result is set equal to zero as

$$\frac{\partial L}{\partial a_i} = -tr\{\mathbf{C}_\mathbf{H}\mathbf{X}_i\mathbf{A}_i\} + a_i^*\, tr\{\mathbf{A}_i^H \mathbf{X}_i^H \mathbf{C}_\mathbf{H}\} + \sum_{\substack{n=1\\ n\neq i}}^{N} a_n^*\, tr\{\mathbf{A}_n^H \mathbf{X}_n^H \mathbf{C}_\mathbf{H}\mathbf{X}_i\mathbf{A}_i\} + \eta = 0 \tag{51}$$

Using the optimal training condition in MMSE channel estimator

$$\mathbf{X}_i\mathbf{X}_i^H = ((P + \sigma^2 N_R\, tr\{\mathbf{C}_\mathbf{H}^{-1}\})/N_T)\,\mathbf{I}_{N_T(L+1)} - \sigma^2 N_R \mathbf{C}_\mathbf{H}^{-1}$$

and with some calculations, (51) reduces to

$$N_T^2\left(\frac{tr\{\mathbf{C}_\mathbf{H}^{-1}\}}{(P_i/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}} - (L+1)\right)\sum_{\substack{n=1\\ n\neq i}}^{N}\frac{a_n^*}{(P_i/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}} + \eta = 0 \tag{52}$$

In the uniform power allocation, $P_1 = \dots = P_N = P_{tot}/N = P$, using $\sum_{n=1}^{N} a_n = 1$, (52) reduces to

$$N_T^2\,\frac{(P(L+1)/\sigma^2 N_R) + L\, tr\{\mathbf{C}_\mathbf{H}^{-1}\}}{[(P/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}]^2}\,(a_i^* - 1) + \eta = 0 \tag{53}$$

Using (53) and $\sum_{n=1}^{N} a_n = 1$, the Lagrange multiplier can be obtained as

$$\eta = N_T^2\left(\frac{N-1}{N}\right)\frac{(P(L+1)/\sigma^2 N_R) + L\, tr\{\mathbf{C}_\mathbf{H}^{-1}\}}{[(P/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}]^2} \tag{54}$$

Substituting (54) back into (53), it is shown that in the uniform power allocation a_n is same as (35). Using (35) and under optimal training, the TMSE (48) is minimized in the uniform power allocation as

$$J_{MMMSE\,(\min)} = \frac{N_T^2}{(P/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}}\left(\frac{L+1}{N} + \frac{N-1}{N}\,\frac{tr\{\mathbf{C}_\mathbf{H}^{-1}\}}{(P/\sigma^2 N_R) + tr\{\mathbf{C}_\mathbf{H}^{-1}\}}\right) \tag{55}$$

When N =1, (55) reduces to the special case of (44) for single channel estimation with the MMSE estimator. According to (55), it is seen that the error decreases when the number of sub-blocks N increases.

In the non-uniform power allocation, $P_n = k_n P$, $P_{tot} = \sum_{n=1}^{N} P_n = N \times P$, $\sum_{n=1}^{N} k_n = N$, suppose that $a_n = k_n / N$ for n = 1, 2, …, N. With some calculations, it is shown that the TMSE (48) is minimized as

$$J_{MMMSE_{(\min)}} = \frac{N_T^2}{N^2}(tr\{\mathbf{C_H^{-1}}\}(\sum_{n=1}^{N}\frac{k_n}{((k_n P/\sigma^2 N_R)+tr\{\mathbf{C_H^{-1}}\})})^2 + (L+1)\sum_{n=1}^{N}\frac{k_n^2}{((k_n P/\sigma^2 N_R)+tr\{\mathbf{C_H^{-1}}\})} - tr\{\mathbf{C_H^{-1}}\}\sum_{n=1}^{N}\frac{k_n^2}{((k_n P/\sigma^2 N_R)+tr\{\mathbf{C_H^{-1}}\})^2}) \quad (56)$$

When $k_n = 1$, (56) reduces to (55).

4. SIMULATION RESULTS

In this section, the performance of the MLS, MSSLS, and MMMSE estimators is numerically examined. It is assumed that each sub-channel has the exponential PDP as

$$b_l = \frac{(1-e^{-1})\, e^{-l}}{1-e^{-L-1}} \quad ; \quad l = 0,1,\ldots,L \quad (57)$$

As a performance measure, it is considered that the channel TMSE is normalized by the average channel energy as

$$NTMSE = \frac{E\left\{ \left\| \mathbf{H} - \hat{\mathbf{H}} \right\|_F^2 \right\}}{E\left\{ \left\| \mathbf{H} \right\|_F^2 \right\}} \quad (58)$$

As optimal training signals, the ZCZ sequences of [26] are employed. Figure 2 shows normalized TMSE (NTMSE) of the LS channel estimator versus SNR in the case of SE (N =1) and ME (N =2, 3, 4, 5, 10). In this figure, it is seen that increasing the number of the sub-blocks *N* results in a lower error of the estimation. It means that the performance of the LS estimator in the frequency selective MIMO channel in ME case is better than SE case.

In Figure 3, NTMSE of the SSLS channel estimator is demonstrated in the case of SE and ME. According to this figure, the SSLS estimator has better performance in ME case than SE especially at high SNRs. On the other hand, at low SNRs, the NTMSEs of the estimator for various numbers of sub-blocks *N* are analogous.

In Figure 4, NTMSE of the MMSE channel estimator is also shown in the case of SE and ME. The results are same as Figure 3 for SSLS estimator. Furthermore, the performance of MMSE estimator is better than SSLS especially for low SNRs and high numbers of *N*. This result is also confirmed by Figures 5-8. In these figures, the performance of the LS, SSLS and MMSE estimators is compared for various SNRs and the number of sub-blocks *N*. As depicted in these figures, for small numbers of *N*, the performance of the MMSE and SSLS estimators in the frequency selective Rician MIMO channel is better than LS particularly at low SNRs. However, for large numbers of *N*, the LS estimator has lower NTMSE than SSLS especially at medium SNRs.

Therefore, in the Rician frequency selective MIMO channels with a long coherence time and hence large *N*, the LS estimator is generally an appropriate method but in channels with a short coherence time and hence small *N*, the SSLS and MMSE are mainly better than LS.

The results in Figures. 9, 10 and 11 are obtained considering non-uniform power allocation during *N* sub-blocks that are used for channel estimation. In these figures, the SSLS and MMSE channel estimators are examined. The proposed non-uniform power allocations are linear and exponential schemes as follows:

$$P_n = k_n P = \frac{2n}{N+1} P, \quad n = 1, \ldots, N \quad (59)$$

$$P_n = k_n P = \frac{N(e-1)}{e^{-1}(1-e^{-N})} e^{-n} P, \quad n = 1, \ldots, N \quad (60)$$

It means that the optimal weight coefficients, a_n, have the linear and exponential distribution, respectively. In Figures. 9, 10 and 11, the results are compared with uniform power allocation. It is seen that the errors with linear power allocation and uniform power allocation are analogous. However, the exponential power allocation has lower error than the uniform power allocation with SSLS channel estimator. On the other hand, the exponential power allocation has higher error than the uniform power allocation with MMSE channel estimator particularly at low SNRs and large values of L in multi-path MIMO channel

In practice, to obtain the best result in channel estimation, one of the LS, SSLS, or MMSE methods can be used considering the channel statistics, the number of antennas, SNR, and *N* (or channel coherent time) in (23), (37), and (56). In order to choose the best estimator among the LS, SSLS, and MMSE channel estimators, the NTMSEs of (23), (37), and (56) can be computed and compared at the receiver.

5. CONCLUSIONS

The advantages of ME have been probed in the Rician frequency selective fading MIMO channels using LS, SSLS, and MMSE estimators. In the both SE and ME cases, the channel estimation errors have been obtained under optimal training. In the case of ME, the optimal weight coefficients and TMSE were achieved for aforementioned estimators with uniform and non-uniform power allocations.

Analytical and numerical results showed that the performance of all estimators in the ME case is remarkably better than SE case. For small values of *N*, suitable for estimation of the channel with fast fading, the MMSE estimator is better than SSLS (LS).

However, for large values of *N*, proper for estimation of the channel with slow fading, the LS estimator is better than SSLS particularly at medium SNRs. It was also shown that the performance of the SSLS estimator in the un-equal power allocation is remarkably better than equal power allocation.

ACKNOWLEDGEMENTS
The author gratefully acknowledges the financial and other supports of this research, provided by the Islamic Azad University, Eslamshahr Branch, Tehran, Iran.

The list of symbols:

$(\cdot)^H$	The Hermitian of a matrix (vector)
$(\cdot)^{-1}$	The inverse of a matrix (vector)
$(\cdot)^T$	The transpose of a matrix (vector)
$tr\{\cdot\}$	The trace of a matrix
$E\{\cdot\}$	Mathematical expectation
$\mathbf{I}_m$	The $m\times m$ identity matrix
$\|\cdot\|_F^2$	The Frobenius norm
$(\cdot)^*$	The complex conjugate
$\otimes$	The Kronecker product
$\mathrm{vec}(\cdot)$	Vectorizing a matrix

REFERENCES

1. D. Tse, P. **"Viswanath, Fundamentals of Wireless Communication"**, *Cambridge University Press*, Cambridge, 2005.
2. S. K. Jayaweera and H.V. Poor, **"On the capacity of multiple-antenna systems in Rician fading"**, *IEEE Trans., Wirel. Commun.*, Vol. 4, No. 3, pp. 1102-1111, 2005.
3. J. Zhang and Y.R. Zheng, **"Improved frequency domain channel estimation and equalization for MIMO wireless communications"**, *Springer Netherlands, Int. J. Wirel. Inf. Netw.*, Vol. 16, No. 1-2, pp. 12-21. 2009.
4. A. K. Sadek, W. Su, and K.J.R. Liu, **"Eigen-selection approach for joint beamforming and space-frequency coding in MIMO-OFDM systems with spatial correlation feedback"**, in: *Proceedings of the 6th IEEE Workshop on Signal Processing Advances in Wireless Communications*, pp. 565-569, 2005.
5. M. Biguesh, A. B. 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, 2006.
6. E. Bjornson, B. Ottersten, **"A framework for training based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance"**, *IEEE Trans. Signal Process.*, Vol. 58 No. 3, 2, pp. 1807–1820, 2010.
7. H. Nooralizadeh, S. Shirvani Moghaddam, **"Performance improvement in estimation of spatially correlated Rician fading MIMO channels using a new LMMSE estimator"**, *Int. J. Commun. Netw. Syst. Sci*., Vol. 3, No. 12, pp. 962-971, 2010.
8. H. Nooralizadeh, S. Shirvani Moghddam, **"A novel shifted type of SLS estimator for estimation of Rician flat fading MIMO channels"**, *Signal Process.*, Vol. 90, No. 6, pp. 1887-1894, 2010.
9. M. Coldrey, and P. Bohlin, **"Training-based MIMO systems: Part I- performance comparison"**, *IEEE Trans. Signal Process*., Vol. 55, No. 11, pp. 5464-5476, 2007.
10. M. Coldrey, and P. Bohlin, **"Training-based MIMO systems: Part II- improvements using detected symbol information"**, *IEEE Trans.* Signal Process, Vol. 56, No. 1, pp. 296-303, 2008.
11. D. Katselis and etal, **"Training sequence design for MIMO channels: an application-oriented approach"**, *EURASIP J. Wirel. Commun. Networking*, 2013:245, October 2013.
12. L. Lindbom, M. Sternad, and A. Ahlen, **"Tracking of time-varying mobile radio channels .I. the Wiener LMS algorithm"**, *IEEE Trans. Commun*., Vol. 49, No. 12, pp. 2207-2217, 2001.
13. J. Zhu, M. F. Siyau, K. K. Loo, **"Channel tracking scheme using block-based Kalman algorithm for MIMO wireless systems"**, in: *Proceedings of 8th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Helsinki, pp. 1-5, June 2007.
14. B. S. Chen, Y. Chang-Yi, and L. Wei-Ji, **"Robust fast time-varying multipath fading channel estimation and equalization for MIMO-OFDM systems via a fuzzy method"**, *IEEE Trans. Veh. Tech*., Vol. 61, No. 4, pp. 1599-1609, May 2012.
15. E. Karami and M. Shiva, **"Decision-directed recursive least squares MIMO channels tracking"**, *EURASIP J. Wirel. Commun. Networking*, Vol. 2, pp. 1-10, 2006.
16. Y. H. Kho and D. P. Taylor, **"MIMO channel estimation and tracking based on polynomial prediction with application to equalization"**, *IEEE Trans. Veh. Tech*., Vol. 57, No. 3, pp. 1585-1595, 2008.
17. Y. H. Kho and D. P. Taylor, **"Reduced complexity MIMO channel estimation and equalization using a polynomial-predictor-based vector GLMS algorithm"**, in: *Proceedings of the 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Athens, pp. 1-5, 2007.
18. X. Xu, J. Wu, S. Ren, L. Song, and H. Xiang, **"Superimposed training-based channel estimation for MIMO relay networks"**, *International Journal of Antennas and Propagation*, pp. 1-11, 2012.
19. W. Yuan, P. Wang, and P. Fan, **"Performance of multi-path MIMO channel estimation based on ZCZ training sequences"**, in *Proceedings of IEEE International Symposium on Microwave*, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE 2005), Vol. 2, pp. 1542-1545, Beijing, August 2005.

20. L. Bingchao, F. Li, and L. Daoben, **"Transmit diversity design for channels with inter symbol interference,"** in *Proceedings of 78th IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1-5, Las Vegas, NV, Sept. 2013.
21. J. K. Hwang, R. L. Chung and M. F. Tsai, **"Highly efficient sparse multi-path channel estimator with optimal Chu-sequence preamble for MIMO frequency-domain DFE receiver"**, in *Proceedings of IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2006)*, pp. 755-759, Vancouver, BC, August 2006.
22. S. Wang and A. Abdi, "**Aperiodic complementary sets of sequences- based MIMO frequency selective channel estimation",** *IEEE Communication Letters*, Vol. 9, No. 10, pp. 891-893, 2005.
23. S. Wang, A. Abdi, **"Low-complexity optimal estimation of MIMO ISI channels with binary training sequences,"** *IEEE Signal Processing Letters*, Vol. 13, No. 11, pp. 657-660, 2006.
24. S. Wang, A. Abdi, **"**MIMO ISI channel estimation using uncorrelated Golay complementary sets of poly-phase sequences,**"** *IEEE Transactions on Vehicular Technology*, Vol. 56, No. 5, pp. 3024-3039, Sep. 2007.
25. R. Rodriguez-Avila, G. Nunez-Vega, R. Parra-Michel, and A. Mendez-Vazquez, **"Frequency selective joint Tx/Rx I/Q imbalance estimation using Golay complementary sequences",** *IEEE Trans. Wirel. Commun.*, Vol. 12, No. 5, pp. 2171-2179, 2013.
26. H. Nooralizadeh, S. Shirvani Moghddam, **"Appropriate algorithms for estimating frequency selective Rician fading MIMO channels and channel Rice factor: substantial benefits of Rician model and estimator tradeoffs",** *EURASIP J. Wirel. Commun. Networking 2010*, 2010:753637.
27. Y. J. Huang, B. S. Chen, **"Decision feedback equalization of frequency selective Rician fading channels",** *Signal Process.* Vol. 76, No. 2, pp. 117-133, 1999.
28. P. Ciblat, P. Bianchi, and M. Ghogho, **"Optimal training for frequency offset estimation in correlated Rice frequency selective channel",** in: Proceedings of the 9th *IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2008)*, pp. 6-10, 2008.
29. S. M. Kay, "**Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory",** *Prentice-Hall*, Upper Saddle River, NJ, 1993.

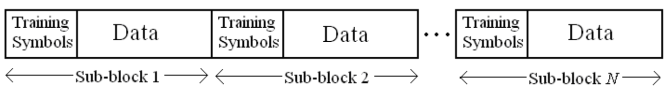

Fig. 1. Frame structure for each Tx antenna in a MIMO channel

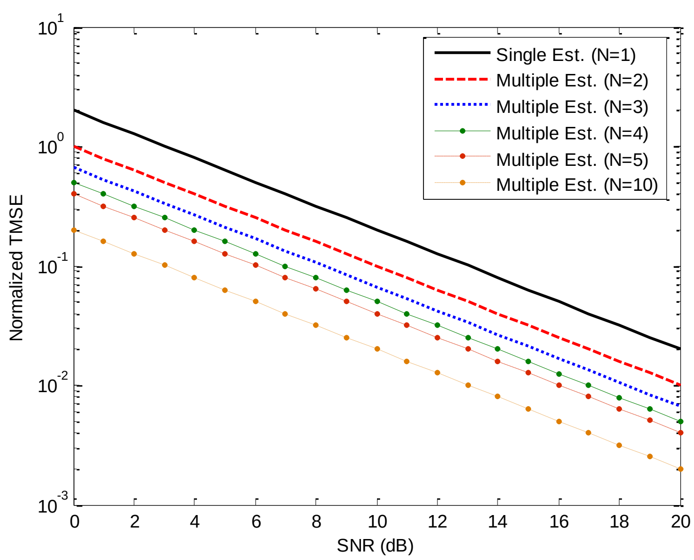

Fig. 2. NTMSE of the LS estimator in the case of single-estimation and multiple-estimation (N = 2, 3, 4, 5, 10, $N_T = N_R$ = 2, L=1, N_P =5)

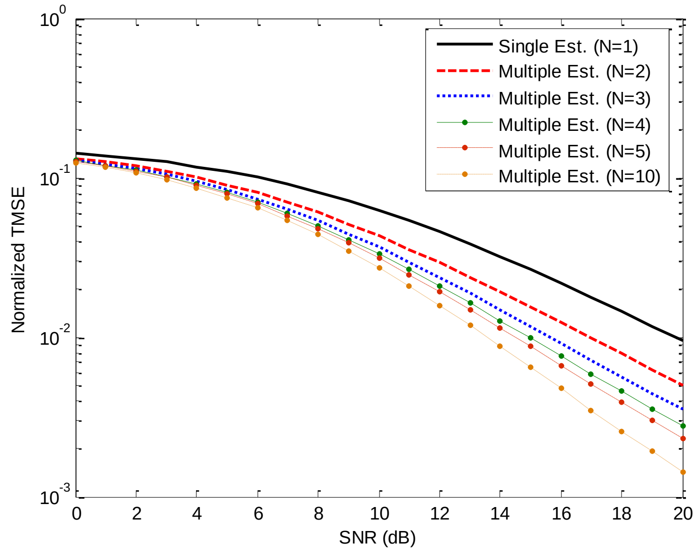

Fig. 3. NTMSE of the SSLS estimator in the case of single-estimation and multiple-estimation (N = 2, 3, 4, 5, 10, $N_T = N_R$ = 2, L=1, N_P =5, κ = 5)

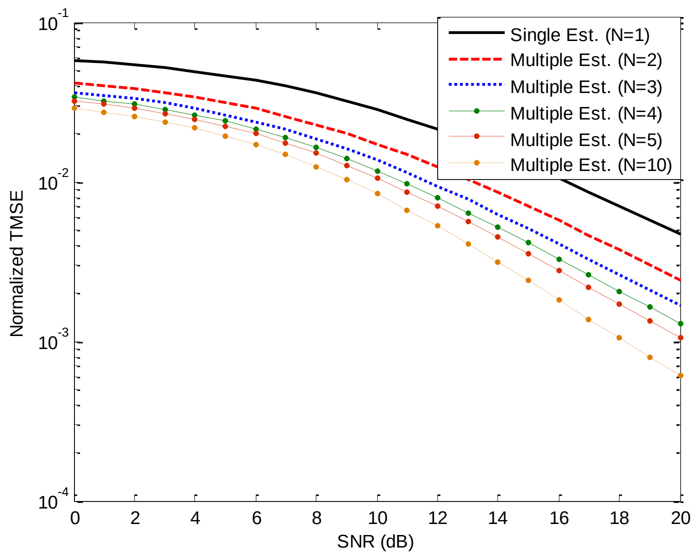

Fig. 4. NTMSE of the MMSE estimator in the case of single-estimation and multiple-estimation (N = 2, 3, 4, 5, 10, $N_T = N_R$ = 2, L=1, N_P =5, κ = 5)

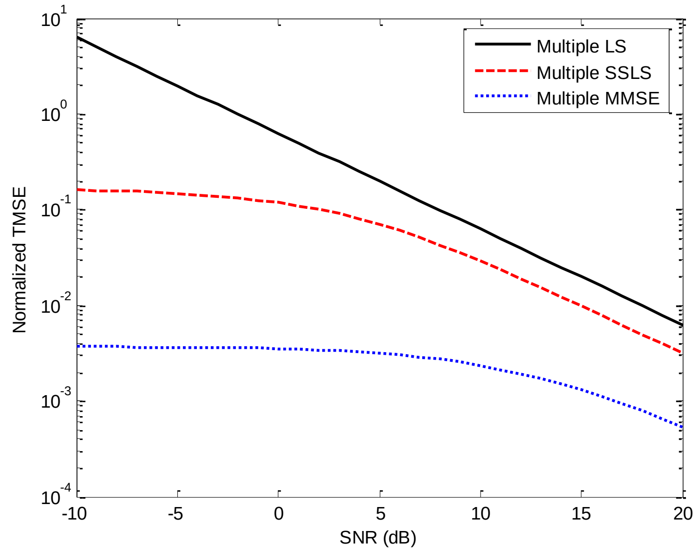

Fig. 5. NTMSE of the LS, SSLS, and MMSE estimators in the case of multiple-estimation (N = 2, $N_T = N_R$ = 2, L=4, N_P =20, κ = 5)

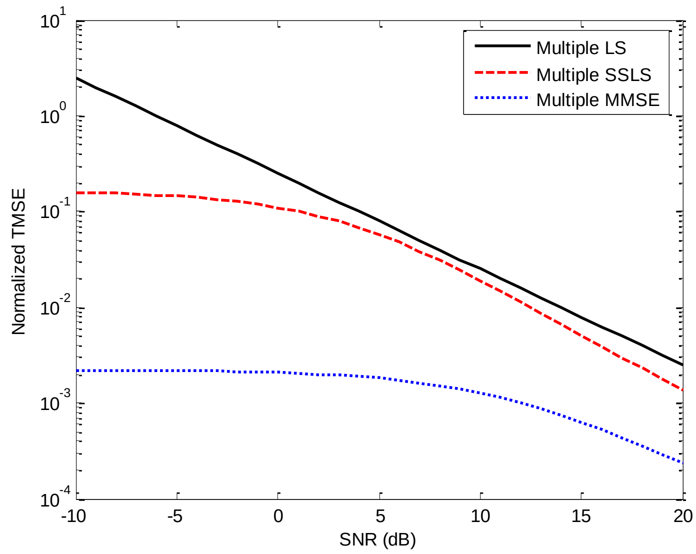

Fig. 6. NTMSE of the LS, SSLS, and MMSE estimators in the case of multiple-estimation (N = 5, $N_T = N_R$ = 2, L=4, N_P =20, κ = 5)

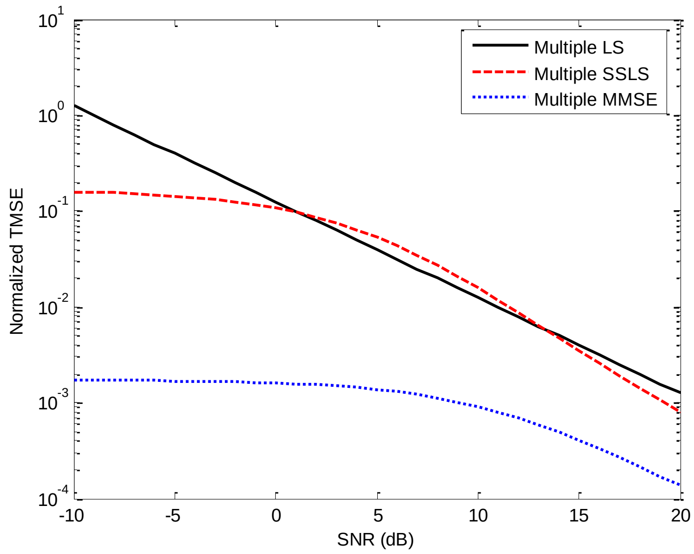

Fig. 7. NTMSE of the LS, SSLS, and MMSE estimators in the case of multiple-estimation (N = 10, $N_T = N_R$ = 2, L=4, N_P =20, κ = 5)

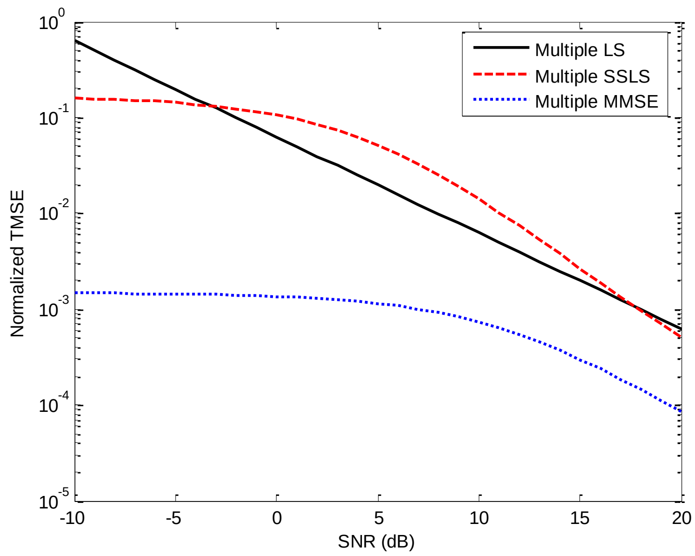

Fig. 8. NTMSE of the LS, SSLS, and MMSE estimators in the case of multiple-estimation (N = 20, $N_T = N_R$ = 2, L=4, N_P =20, κ = 5)

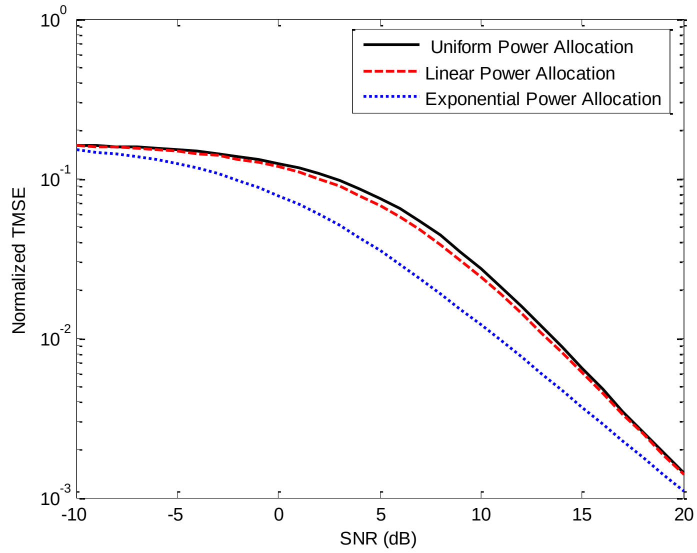

Fig. 9. NTMSE of the SSLS estimator in uniform power and non-uniform power allocation (N = 10, $N_T = N_R$ = 2, L=1, N_P =5, κ = 5)

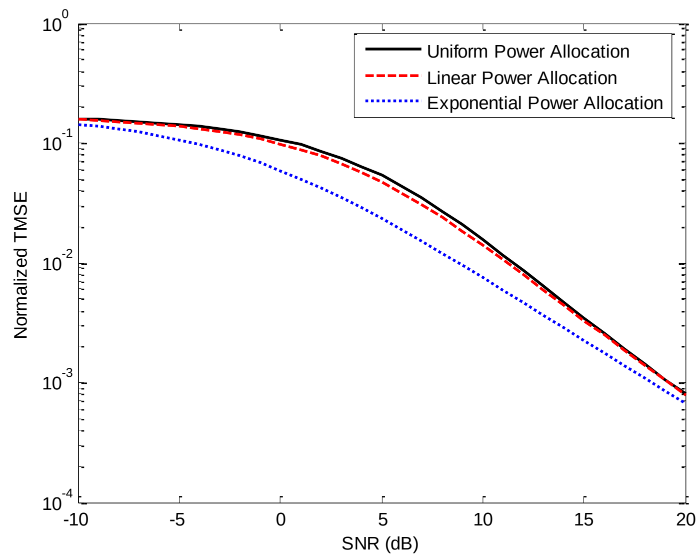

Fig. 10. NTMSE of the SSLS estimator in uniform power and non-uniform power allocation (N = 10, $N_T = N_R$ = 2, L=4, N_P =20, κ = 5)

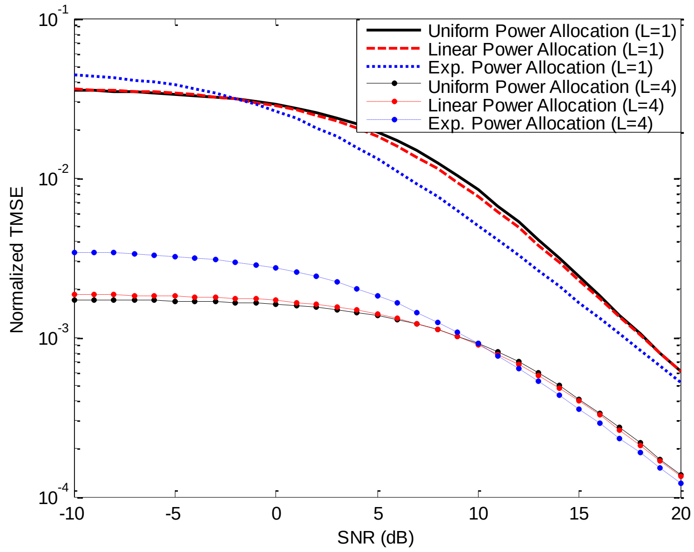

Fig. 11. NTMSE of the MMSE estimator in uniform power and non-uniform power allocation for L=1, 4 (N = 10, $N_T = N_R = 2$, $\kappa = 5$)