

RESEARCH

Open Access

On the optimality of training signals for MMSE channel estimation in MIMO-OFDM systems

Junho Jo¹ and Ilsoo Sohn^{2*}

Abstract

In this paper, we investigate the optimality of training signals for linear minimum mean square error (LMMSE) channel estimation in multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) with frequency-selective fading channels. This is a very challenging problem due to its mathematical intractability and has not been analytically solved in the literature. Using the Lagrange multiplier method, we derive the optimality conditions for training signal design. Important findings revealed on optimal training signals are twofold: (i) the energies of the training signals on each subcarrier are equal, and (ii) on each subcarrier, the training signals transmitted from the different antennas are orthogonal and of equal energy. We verify that our results are in line with the design principles that have been derived in single-carrier MIMO systems. Two types of optimal training signal examples that satisfy the optimality conditions are presented for practical implementations in MIMO-OFDM systems. Simulation results show that the training signals based on the optimality conditions outperform other non-optimal training signals in terms of channel estimation performance.

Keywords: Optimal training signal; MMSE channel estimation; MIMO; OFDM; Frequency-selective fading

1 Introduction

Recently, increasing interest has been concentrated on multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) for broadband wireless communication. The combination of OFDM with MIMO exploits the benefits from both techniques, i.e., the robustness to combat multipath delay spread and an increase in system capacity [1-5]. For a practical implementation of MIMO-OFDM systems, channel estimation becomes very important for the system performance. Imperfect channel estimation typically leads to the increase of error rates and reduces transmission efficiency. In this study, we focus on the problem of designing MIMO-OFDM training signals for channel estimation, a critical component in many modern wireless communication systems.

Several approaches on optimal training signal design have been proposed in the literature. In the absence of prior statistical information about the channel, simple least-squares channel estimation is used. In [6], the

optimal placement of training signals is studied for a single-input single-output (SISO) OFDM case. The optimal training design is extended to the MIMO-OFDM case in [7]. In [8], more general training structures for MIMO-OFDM are presented that utilize frequency division, time division, and code division multiplexing.

In the presence of prior statistical information about the channel, a more efficient channel estimation technique can be used. Linear minimum mean square error (LMMSE) estimators that incorporate prior knowledge to improve channel estimation are known to be optimal for this case. In [9,10], optimal conditions for training signals are studied for SISO frequency-selective fading channels. In [11], the results are extended to MIMO frequency-selective fading channels, and it is revealed that training signals across transmit antennas should be orthogonal and training signals should be equi-powered. The results are very simple but effective and thus have been widely used as design guidelines for recent wireless systems. One can intuitively generalize the main principles to multi-carrier systems while the optimality of the principles has remained unproved for the multi-carrier systems, i.e., OFDM. For example, downlink reference signals in commercial LTE systems are designed with the

*Correspondence: ilsoo.sohn@gachon.ac.kr

²Department of Electronic Engineering, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam 461-701, Korea
Full list of author information is available at the end of the article

same principles. In [12], an optimal training design for both least-squares estimators and LMMSE estimators is studied assuming cyclic delay diversity OFDM systems.

Meanwhile, research interests have been shifted to more practical issues on MIMO channel estimations. In [13,14], a training signal design in the existence of inphase and quadrature imbalances is considered for SISO-OFDM and MIMO-OFDM systems. A robust training signal design for LMMSE channel estimator in case of imperfect knowledge of second-order characteristics of channels is studied in [15-17]. Efficient algorithms that exploit the spatial correlation of MIMO channels are proposed for training signal design in [18,19].

In this paper, we aim to complete the puzzle with the missing piece. We directly tackle a multi-carrier system model and solve the optimality conditions of training signals for LMMSE estimator in MIMO-OFDM systems. This is quite challenging because simultaneous considerations on all MIMO dimensions, channel statistics, multi-carriers, and multisymbols lead to an extremely complex modeling and mathematically intractable problem. This has never been solved in the literature to the best of our knowledge. We analytically derive the optimality conditions using the Lagrange multiplier method, which is the principle contribution of this study.

The remainder of the paper is organized as follows. Section 2 describes the system model of our work. After Section 3 analytically derives the optimality conditions, the optimal training signal design is discussed in Section 4. Section 5 presents two types of optimal training signal examples for practical MIMO-OFDM systems. Simulation results are provided in Section 6. Finally, Section 7 draws conclusions.

Notations: Uppercase and lowercase boldface letters are used for matrices and vectors, respectively. The superscript ‘*’ denotes the conjugate transpose, superscript ‘T’ denotes the transpose, and superscript ‘-1’ denotes the matrix/vector inverse. We will use $\mathbb{E}[\cdot]$ for expectation, $\text{vec}(\cdot)$ for matrix vectorization, $\text{tr}[\cdot]$ for the matrix trace, \otimes for the Kronecker product, and \mathbf{I}_N for the $N \times N$ identity matrix.

2 System model

We consider a MIMO-OFDM system with Q transmit and P receive antennas and perform an analysis in the frequency domain to search for the properties of an optimal training signal. For the best analytical tractability, we will work directly in the frequency domain. The number of OFDM subcarriers is N , and we consider a block of M OFDM symbols transmitted across the channel. We use the notation $x_{n,q,m}$ to denote the training signal transmitted on the q th transmit antenna in the m th symbol and on the n th subcarrier. For the n th subcarrier, we may, therefore, consider that we transmit a matrix of training signals,

\mathbf{X}_n , where $x_{n,q,m}$ is the element in the q th row and m th column. The following notations are used in this paper:

- \mathbf{X}_n : Training signal matrix for the n th subcarrier ($Q \times M$),
- \mathbf{Y}_n : Received signal matrix for the n th subcarrier ($P \times M$),
- \mathbf{H}_n : MIMO channel matrix for the n th subcarrier ($P \times Q$),
- \mathbf{W}_n : Received noise matrix for the n th subcarrier ($P \times M$).

Considering frequency-selective fading, a signal model is given as

$$\mathbf{Y}_n = \mathbf{H}_n \mathbf{X}_n + \mathbf{W}_n, \tag{1}$$

where all elements of \mathbf{H}_n are uncorrelated, and all noise variables are independent, i.e., $\mathbb{E}[\mathbf{W}_n \mathbf{W}_n^*] = \sigma^2 \mathbf{I}$. Aggregating the matrices for all subcarriers gives

$$\mathbf{x} = \begin{bmatrix} \text{vec}(\mathbf{X}_1) \\ \text{vec}(\mathbf{X}_2) \\ \vdots \\ \text{vec}(\mathbf{X}_N) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \text{vec}(\mathbf{Y}_1) \\ \text{vec}(\mathbf{Y}_2) \\ \vdots \\ \text{vec}(\mathbf{Y}_N) \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix},$$

$$\mathbf{h} = \begin{bmatrix} \text{vec}(\mathbf{H}_1) \\ \text{vec}(\mathbf{H}_2) \\ \vdots \\ \text{vec}(\mathbf{H}_N) \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_N \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} \text{vec}(\mathbf{W}_1) \\ \text{vec}(\mathbf{W}_2) \\ \vdots \\ \text{vec}(\mathbf{W}_N) \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_N \end{bmatrix}, \tag{2}$$

where \mathbf{x} , \mathbf{y} , \mathbf{h} , \mathbf{w} are respectively $MNQ \times 1$, $MNP \times 1$, $NPQ \times 1$, and $MNP \times 1$ matrices. By using a well-known result regarding the vec operation, i.e., $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A}) \text{vec}(\mathbf{X})$, (1) becomes

$$\begin{aligned} \text{vec}(\mathbf{Y}_n) &= \text{vec}(\mathbf{H}_n \mathbf{X}_n) + \text{vec}(\mathbf{W}_n) \\ &= \text{vec}(\mathbf{I}_P \mathbf{H}_n \mathbf{X}_n) + \text{vec}(\mathbf{W}_n) \\ &= (\mathbf{X}_n^T \otimes \mathbf{I}_P) \text{vec}(\mathbf{H}_n) + \text{vec}(\mathbf{W}_n), \end{aligned}$$

or can be rewritten as

$$\mathbf{y}_n = (\mathbf{X}_n^T \otimes \mathbf{I}_P) \mathbf{h}_n + \mathbf{w}_n. \tag{3}$$

Combining all the N equations in (3) yields

$$\mathbf{y} = \begin{bmatrix} (\mathbf{X}_1^T \otimes \mathbf{I}_P) & & & \\ & (\mathbf{X}_2^T \otimes \mathbf{I}_P) & & \\ & & \ddots & \\ & & & (\mathbf{X}_N^T \otimes \mathbf{I}_P) \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_N \end{bmatrix} + \mathbf{w},$$

or shortly denoted as

$$\mathbf{y} = \mathbf{X} \mathbf{h} + \mathbf{w}, \tag{4}$$

where \mathbf{X} ($MNP \times NPQ$) is defined as

$$\mathbf{X} = \begin{bmatrix} (\mathbf{X}_1^T \otimes \mathbf{I}_p) & & & \\ & (\mathbf{X}_2^T \otimes \mathbf{I}_p) & & \\ & & \ddots & \\ & & & (\mathbf{X}_N^T \otimes \mathbf{I}_p) \end{bmatrix}. \quad (5)$$

3 Analysis on optimality conditions

In this section, the problem formulation for minimizing the LMMSE channel estimation errors in MIMO-OFDM systems is presented first. Then, optimal conditions for the training signals are derived for an optimal training signal design.

3.1 Problem formulation

The standard LMMSE estimate of \mathbf{h} , based upon the observation of \mathbf{y} , becomes

$$\hat{\mathbf{h}} = \mathbb{E}[\mathbf{h}\mathbf{y}^*] (\mathbb{E}[\mathbf{y}\mathbf{y}^*])^{-1} \mathbf{y} \quad (6)$$

By using (4), this becomes

$$\begin{aligned} \hat{\mathbf{h}} &= \mathbb{E}[\mathbf{h}\mathbf{h}^*] \mathbf{X}^* (\mathbf{X}\mathbb{E}[\mathbf{y}\mathbf{y}^*]\mathbf{X}^* + \sigma_w^2 \mathbf{I})^{-1} \mathbf{y} \\ &= \mathbf{R}_{\mathbf{h}\mathbf{h}} \mathbf{X}^* (\mathbf{X}\mathbf{R}_{\mathbf{h}\mathbf{h}}\mathbf{X}^* + \sigma_w^2 \mathbf{I})^{-1} \mathbf{y}. \end{aligned} \quad (7)$$

The covariance matrix of the error can be expressed as

$$\begin{aligned} \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}})(\mathbf{h} - \hat{\mathbf{h}})^*] &= \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}})\mathbf{h}^*] \\ &= \mathbf{R}_{\mathbf{h}\mathbf{h}} - \mathbf{R}_{\mathbf{h}\mathbf{h}}\mathbf{X}^* (\mathbf{X}\mathbf{R}_{\mathbf{h}\mathbf{h}}\mathbf{X}^* + \sigma_w^2 \mathbf{I})^{-1} \mathbf{X}\mathbf{R}_{\mathbf{h}\mathbf{h}}. \end{aligned} \quad (8)$$

Using the well-known matrix identity,

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1},$$

(8) can be rewritten as

$$\begin{aligned} \mathbb{E}[(\mathbf{h} - \hat{\mathbf{h}})(\mathbf{h} - \hat{\mathbf{h}})^*] &= \left(\mathbf{R}_{\mathbf{h}\mathbf{h}}^{-1} + \frac{\mathbf{X}^*\mathbf{X}}{\sigma_w^2} \right)^{-1} \\ &= \sigma_w^2 \left(\sigma_w^2 \mathbf{R}_{\mathbf{h}\mathbf{h}}^{-1} + \mathbf{X}^*\mathbf{X} \right)^{-1}. \end{aligned} \quad (9)$$

The MMSE becomes [20,21]

$$\text{MMSE} = \sigma_w^2 \text{tr} \left[\left(\sigma_w^2 \mathbf{R}_{\mathbf{h}\mathbf{h}}^{-1} + \mathbf{X}^*\mathbf{X} \right)^{-1} \right]. \quad (10)$$

Our goal is to find the matrix \mathbf{X} in the form of (5) that minimizes (10), subject to the total transmit energy constraint.

Based on the assumptions of the channel, it can be observed that

$$\mathbf{R}_{\mathbf{h}\mathbf{h}} = \mathbf{R} \otimes \mathbf{I}_Q \otimes \mathbf{I}_p, \quad (11)$$

where \mathbf{R} ($N \times N$) is a Toeplitz Hermitian matrix of the form

$$\mathbf{R} = \begin{bmatrix} r_0 & r_1 & \cdots & r_{N-1} \\ r_{N-1}^* & r_0 & \cdots & r_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_1^* & r_2 & \cdots & r_0 \end{bmatrix}. \quad (12)$$

We also have

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} (\mathbf{X}_1^T \otimes \mathbf{I}_p) & & & \\ & (\mathbf{X}_2^T \otimes \mathbf{I}_p) & & \\ & & \ddots & \\ & & & (\mathbf{X}_N^T \otimes \mathbf{I}_p) \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{X}_1^T & & & \\ & \mathbf{X}_2^T & & \\ & & \ddots & \\ & & & \mathbf{X}_N^T \end{bmatrix} \otimes \mathbf{I}_p \\ &= \mathbf{Z} \otimes \mathbf{I}_p, \end{aligned} \quad (13)$$

where \mathbf{Z} ($MN \times NQ$) is defined as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{X}_1^T & & & \\ & \mathbf{X}_2^T & & \\ & & \ddots & \\ & & & \mathbf{X}_N^T \end{bmatrix}. \quad (14)$$

Furthermore,

$$\begin{aligned} \mathbf{X}^*\mathbf{X} &= (\mathbf{Z} \otimes \mathbf{I}_p)^* (\mathbf{Z} \otimes \mathbf{I}_p) \\ &= (\mathbf{Z}^* \otimes \mathbf{I}_p) (\mathbf{Z} \otimes \mathbf{I}_p) \\ &= \mathbf{Z}^*\mathbf{Z} \otimes \mathbf{I}_p. \end{aligned} \quad (15)$$

Using some basic properties regarding \otimes operations and expressions in (11)~(15), the MMSE in (10) can be rewritten as

$$\begin{aligned} \text{MMSE} &= \sigma_w^2 \text{tr} \left[\left(\sigma_w^2 \mathbf{R}_{\mathbf{h}\mathbf{h}}^{-1} + \mathbf{X}^*\mathbf{X} \right)^{-1} \right] \\ &= \sigma_w^2 \text{tr} \left[\left(\sigma_w^2 (\mathbf{R} \otimes \mathbf{I}_Q \otimes \mathbf{I}_p)^{-1} + \mathbf{Z}^*\mathbf{Z} \otimes \mathbf{I}_p \right)^{-1} \right] \\ &= \sigma_w^2 \text{tr} \left[\left(\sigma_w^2 (\mathbf{R}^{-1} \otimes \mathbf{I}_Q \otimes \mathbf{I}_p) + \mathbf{Z}^*\mathbf{Z} \otimes \mathbf{I}_p \right)^{-1} \right] \\ &= \sigma_w^2 \text{tr} \left[\left((\sigma_w^2 (\mathbf{R}^{-1} \otimes \mathbf{I}_Q) + \mathbf{Z}^*\mathbf{Z}) \otimes \mathbf{I}_p \right)^{-1} \right] \\ &= \sigma_w^2 \text{tr} \left[\left(\sigma_w^2 (\mathbf{R}^{-1} \otimes \mathbf{I}_Q) + \mathbf{Z}^*\mathbf{Z} \right)^{-1} \otimes \mathbf{I}_p \right] \\ &= P \sigma_w^2 \text{tr} \left[\left(\sigma_w^2 (\mathbf{R} \otimes \mathbf{I}_Q)^{-1} + \mathbf{Z}^*\mathbf{Z} \right)^{-1} \right]. \end{aligned} \quad (16)$$

In addition, it is known that, for any matrices \mathbf{A} and \mathbf{B} , there exists a permutation matrix Π such that $\mathbf{A} \otimes \mathbf{B} = \Pi (\mathbf{B} \otimes \mathbf{A}) \Pi^T$ [22]. Let Π be a permutation matrix such that $\mathbf{R} \otimes \mathbf{I}_Q = \Pi (\mathbf{I}_Q \otimes \mathbf{R}) \Pi^T$. Using this relation in (16) yields

$$\begin{aligned}
 \text{MMSE} &= P\sigma_w^2 \text{tr} \left[\left(\sigma_w^2 \Pi (\mathbf{I}_Q \otimes \mathbf{R})^{-1} \Pi^T + \mathbf{Z}^* \mathbf{Z} \right)^{-1} \right] \\
 &= P\sigma_w^2 \text{tr} \left[\left(\sigma_w^2 \Pi (\mathbf{I}_Q \otimes \mathbf{R})^{-1} \Pi^T + \Pi \Pi^T \mathbf{Z}^* \mathbf{Z} \Pi \Pi^T \right)^{-1} \right] \\
 &= P\sigma_w^2 \text{tr} \left[\left(\Pi \left(\sigma_w^2 (\mathbf{I}_Q \otimes \mathbf{R})^{-1} + \Pi^T \mathbf{Z}^* \mathbf{Z} \Pi \right) \Pi^T \right)^{-1} \right]. \tag{17}
 \end{aligned}$$

Note that row permutations by Π followed by column permutations by the same Π does not change the trace of the matrix because any diagonal parameter may be permuted but still remains in the diagonal position. Thus, we have

$$\text{MMSE} = P\sigma_w^2 \text{tr} \left[\left(\sigma_w^2 (\mathbf{I}_Q \otimes \mathbf{R})^{-1} + \Pi^T \mathbf{Z}^* \mathbf{Z} \Pi \right)^{-1} \right]. \tag{18}$$

Finally, our optimization problem is formulated as a constrained optimization problem as

$$\begin{aligned}
 &\text{Minimize } P\sigma_w^2 \text{tr} \left[\left(\sigma_w^2 (\mathbf{I}_Q \otimes \mathbf{R})^{-1} + \Pi^T \mathbf{Z}^* \mathbf{Z} \Pi \right)^{-1} \right] \\
 &\text{Subject to } \text{tr} [\mathbf{Z}^* \mathbf{Z}] = E_{\text{total}}, \tag{19}
 \end{aligned}$$

where E_{total} is the total transmit energy.

3.2 Optimality conditions

For simplicity of notations, we define \mathbf{D} ($NQ \times NQ$) and \mathbf{A} ($NQ \times NQ$) as

$$\mathbf{D} \equiv \Pi^T \mathbf{Z}^* \mathbf{Z} \Pi, \tag{20}$$

and

$$\mathbf{A} \equiv \left(\sigma_w^2 (\mathbf{I}_Q \otimes \mathbf{R})^{-1} + \mathbf{D} \right), \tag{21}$$

where $\mathbf{Z}^* \mathbf{Z}$ is Hermitian and block diagonal. We solve the constrained optimization problem using the Lagrange multiplier method. By letting μ be the Lagrange multiplier, the Lagrangian is expressed as

$$J(\mathbf{D}, \mu) = P\sigma_w^2 \text{tr} [\mathbf{A}^{-1}] + \mu (\text{tr} [\mathbf{D}] - E_{\text{total}}). \tag{22}$$

To obtain the optimal solution, we set the derivatives of the Lagrangian function $J(\mathbf{D}, \mu)$ to zeros as

$$\frac{\partial}{\partial x} J(\mathbf{D}, \mu) = 0, \tag{23}$$

where x is an arbitrary parameter in \mathbf{D} . We split our approaches into two cases: diagonal parameters and off-diagonal parameters in matrix \mathbf{D} . To find the required derivatives, we will use the following property. For any matrix \mathbf{A} depending on a parameter x [22],

$$\frac{\partial}{\partial x} \mathbf{A}^{-1} = -\mathbf{A}^{-1} \left(\frac{\partial}{\partial x} \mathbf{A} \right) \mathbf{A}^{-1}, \tag{24}$$

which directly follows from the fact that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

We first consider the diagonal parameters x in \mathbf{D} and derive the following lemma.

Lemma 1. *By denoting the k th column of \mathbf{A}^{-1} as \mathbf{a}_k , the following condition should be satisfied in order to achieve $\frac{\partial}{\partial x} J(\mathbf{D}, \mu) = 0$.*

$$\text{C1: } \|\mathbf{a}_k\|^2 = c, \quad \text{for all } k, \tag{25}$$

where c is a constant.

Proof. Let x be the k th diagonal parameter in \mathbf{D} . We define the derivative of \mathbf{A} with respect to x as

$$\frac{\partial \mathbf{A}}{\partial x} = \mathbf{E} \equiv \text{diag} \{0, \dots, 0, 1, 0, \dots, 0\}, \tag{26}$$

where \mathbf{E} is a diagonal matrix in which only the k th diagonal element has a value of one. Using associating (24) with (22), we have

$$\begin{aligned}
 \frac{\partial}{\partial x} J(\mathbf{D}, \mu) &= -P\sigma_w^2 \text{tr} [\mathbf{A}^{-1} \mathbf{E} \mathbf{A}^{-1}] + \mu \\
 &= -P\sigma_w^2 \text{tr} [\mathbf{A}^{-1} \mathbf{E} (\mathbf{A}^{-1})^*] + \mu \\
 &= -P\sigma_w^2 \|\mathbf{a}_k\|^2 + \mu \tag{27}
 \end{aligned}$$

where we use the fact that \mathbf{A}^{-1} is a Hermitian matrix in the second line. Applying $\frac{\partial}{\partial x} J(\mathbf{D}, \mu) = 0$ finally gives

$$\|\mathbf{a}_k\|^2 = \frac{\mu}{P\sigma_w^2}, \tag{28}$$

which holds for all diagonal elements in \mathbf{D} without loss of generality. This completes the proof. \square

We then consider the off-diagonal parameters χ in \mathbf{D} and derive the following lemma. Note that off-diagonal parameters are complex numbers. We split each parameter into two real parameters as $\chi = x + iy$.

Lemma 2. *By denoting the k th column of \mathbf{A}^{-1} as \mathbf{a}_k , the following condition should be satisfied in order to achieve $\frac{\partial}{\partial x} J(\mathbf{D}, \mu) = 0$.*

$$\text{C2: } \mathbf{a}_k^* \mathbf{a}_l = 0, \quad \text{for all } k, l \ (k \neq l). \tag{29}$$

Proof. Let χ appear in the k th column and in the l th row in \mathbf{D} . We first focus on the real part x . We define the derivative of \mathbf{A} with respect to x as $\frac{\partial \mathbf{A}}{\partial x} = \mathbf{E}$, where \mathbf{E} is defined as a sparse matrix in which only the element in the k th column l th row and the element in the l th column k th row have a value of one. Using associating (24) with (22), we have

$$\begin{aligned}
 \frac{\partial}{\partial x} J(\mathbf{D}, \mu) &= -P\sigma_w^2 \text{tr} [\mathbf{A}^{-1} \mathbf{E} \mathbf{A}^{-1}] \\
 &= -P\sigma_w^2 \text{tr} [\mathbf{A}^{-1} \mathbf{E} (\mathbf{A}^{-1})^*] \\
 &= -P\sigma_w^2 (\mathbf{a}_k^* \mathbf{a}_l + \mathbf{a}_l^* \mathbf{a}_k) \tag{30}
 \end{aligned}$$

where we use the fact that \mathbf{A}^{-1} is a Hermitian matrix in the second line. Applying $\frac{\partial}{\partial x} J(\mathbf{D}, \mu) = 0$ finally gives

$$\text{Re}[\mathbf{a}_k^* \mathbf{a}_l] = 0, \tag{31}$$

where $\text{Re}[\cdot]$ denotes the real part of a complex number. Then, similar derivations with respect to the imaginary part y yield

$$\text{Im}[\mathbf{a}_k^* \mathbf{a}_l] = 0, \tag{32}$$

where $\text{Im}[\cdot]$ denotes the imaginary part of a complex number. Combining both results provides

$$\mathbf{a}_k^* \mathbf{a}_l = 0, \tag{33}$$

which holds for all off-diagonal elements in \mathbf{D} without loss of generality. This completes the proof. \square

4 Optimal training signal design

In this section, we provide an optimal training signal satisfying the optimality conditions derived in the previous section.

Suppose that the matrix \mathbf{D} is a multiple of the identity, i.e., $\mathbf{D} = c\mathbf{I}$, where c is a real constant. Because \mathbf{R} is a Toeplitz Hermitian matrix, it can be written as

$$\mathbf{R} = \mathbf{F}\mathbf{\Lambda}\mathbf{F}^{-1}, \tag{34}$$

where \mathbf{F} and $\mathbf{\Lambda}$ are a unitary matrix and a diagonal matrix, respectively. In addition, we may write

$$\mathbf{I}_Q \otimes \mathbf{R} = \begin{bmatrix} \mathbf{F} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{\Lambda} \end{bmatrix} \begin{bmatrix} \mathbf{F}^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F}^{-1} \end{bmatrix}, \tag{35}$$

$$(\mathbf{I}_Q \otimes \mathbf{R})^{-1} = \begin{bmatrix} \mathbf{F} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{\Lambda}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{F}^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F}^{-1} \end{bmatrix}. \tag{36}$$

For shorter notations,

$$(\mathbf{I}_Q \otimes \mathbf{R})^{-1} = \bar{\mathbf{F}}\bar{\mathbf{\Lambda}}^{-1}\bar{\mathbf{F}}^{-1}, \tag{37}$$

where $\bar{\mathbf{F}}$ and $\bar{\mathbf{\Lambda}}^{-1}$ are respectively defined as

$$\bar{\mathbf{F}} = \begin{bmatrix} \mathbf{F} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F} \end{bmatrix}, \bar{\mathbf{\Lambda}}^{-1} = \begin{bmatrix} \mathbf{\Lambda}^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{\Lambda}^{-1} \end{bmatrix} \tag{38}$$

Associating (37) in (21) gives

$$\begin{aligned} \mathbf{A}^{-1} &= (\sigma_w^2 (\mathbf{I}_Q \otimes \mathbf{R})^{-1} + \mathbf{D})^{-1} \\ &= (\sigma_w^2 \bar{\mathbf{F}}\bar{\mathbf{\Lambda}}^{-1}\bar{\mathbf{F}}^{-1} + c\mathbf{I})^{-1} \\ &= (\sigma_w^2 \bar{\mathbf{F}}\bar{\mathbf{\Lambda}}^{-1}\bar{\mathbf{F}}^{-1} + c\bar{\mathbf{F}}\bar{\mathbf{F}}^{-1})^{-1} \\ &= \bar{\mathbf{F}} (\sigma_w^2 \bar{\mathbf{\Lambda}}^{-1} + c\mathbf{I})^{-1} \bar{\mathbf{F}}^{-1} \\ &= \begin{bmatrix} \mathbf{F} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F} \end{bmatrix} \begin{bmatrix} (\sigma_w^2 \mathbf{\Lambda}^{-1} + c\mathbf{I})^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & (\sigma_w^2 \mathbf{\Lambda}^{-1} + c\mathbf{I})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{F}^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{F} (\sigma_w^2 \mathbf{\Lambda}^{-1} + c\mathbf{I})^{-1} \mathbf{F}^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{F} (\sigma_w^2 \mathbf{\Lambda}^{-1} + c\mathbf{I})^{-1} \mathbf{F}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{C} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{C} \end{bmatrix}, \end{aligned} \tag{39}$$

where \mathbf{C} is a Toeplitz Hermitian matrix. Thus, it is clear that $\mathbf{D} = c\mathbf{I}$ satisfies both the optimality conditions $\mathbf{C1}$ in (25) and $\mathbf{C2}$ in (29). Solving $\mathbf{D} = \mathbf{\Pi}^T \mathbf{Z}^* \mathbf{Z} \mathbf{\Pi} = c\mathbf{I}$ for $\mathbf{Z}^* \mathbf{Z}$ yields

$$\mathbf{Z}^* \mathbf{Z} = \mathbf{\Pi} (c\mathbf{I}) \mathbf{\Pi}^T = c\mathbf{I}. \tag{40}$$

From (14) and (40), we finally reach the following theorem on the optimal training signal.

Theorem (Optimal training signal). *In MIMO-OFDM systems where a sequence of a training signal is transmitted at the transmitter through Q transmit antennas, N subcarriers, M OFDM symbols with the total transmit power of E_{total} and LMMSE channel estimation is performed at the receiver upon the reception of M OFDM symbols from P received antennas, the training signal is 'optimal' in terms of minimizing channel estimation errors if the training signal satisfies the following conditions:*

- The energy of the training signal in each subcarrier is equal, i.e.,

$$\text{tr}[\mathbf{X}_n \mathbf{X}_n^*] = \frac{E_{total}}{N}, \text{ for } n = 1, 2, \dots, N. \tag{41}$$

- On each subcarrier, the training signals transmitted from the different antennas are orthogonal and of equal energy, i.e.,

$$\mathbf{X}_n \mathbf{X}_n^* = \frac{E_{total}}{NQ} \mathbf{I}, \text{ for } n = 1, 2, \dots, N. \tag{42}$$

5 Examples of optimal training signals

In this section, we use the theorem revealed in the previous section as a design guideline and present two examples of optimal training signal implementations. These designs are practical owing to its simple structure. Note

that optimal designs of training signals are not limited to the following cases.

5.1 Sequential transmission on antennas

Assume that the number of OFDM symbols transmitted is equal to the number of transmit antennas, $M = Q$, and let

$$\mathbf{X}_n = \begin{bmatrix} z_1^{(n)} & 0 & \cdots & 0 \\ 0 & z_2^{(n)} & & \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & z_Q^{(n)} \end{bmatrix}, \text{ for } n = 1, 2, \dots, N, \tag{43}$$

where $z_i^{(n)}$ are arbitrary complex numbers but satisfying $|z_i^{(n)}|^2 = \frac{E_{\text{total}}}{NQ}$. This implementation implies that each successive OFDM symbol is transmitted on a different antenna in a round-robin fashion. Figure 1 illustrates an example of the optimal training signal design in this type of optimal training signal implementation.

5.2 Interlaced transmission on antennas

Assume that the number of OFDM symbols transmitted is equal to the number of transmit antennas, $M = Q$, and N is a multiple of Q . In this implementation, we transmit simultaneously from every antenna in each symbol interval, but each antenna uses only every Q th subcarrier. Let matrix Φ represent a cyclic shift operation, which causes a cyclic shift by one element in the upward direction. For example, if $Q = 3$, we have

$$\Phi = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \tag{44}$$

and it operates as

$$\Phi \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix}. \tag{45}$$

We define the training signal matrix by

$$\mathbf{X}_n = \Phi^{(n-1)}\Psi, \text{ for } n = 1, 2, \dots, N, \tag{46}$$

where Φ^n is the n th power of Φ ,

$$\Psi = \begin{bmatrix} z_1 & 0 & \cdots & 0 \\ 0 & z_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & z_Q \end{bmatrix}, \tag{47}$$

and z_i are arbitrary complex numbers but satisfying $|z_i^{(n)}|^2 = \frac{E_{\text{total}}}{NQ}$. This implementation implies that each antenna always transmit in every OFDM symbol but using only every Q th subcarrier. During the first symbol, for example, Antenna 1 uses subcarriers 1, $Q + 1, 2Q + 1, \dots$, Antenna Q uses subcarriers 2, $Q + 2, 2Q + 2, \dots$, etc. During the second symbol, Antenna 1 uses subcarriers 2, $Q + 2, 2Q + 2, \dots$, and Antenna Q uses subcarriers 3, $Q + 3, 2Q + 3, \dots$, and so on. Figure 2 illustrates an example of the optimal training signal design in this type of optimal training signal implementation.

6 Simulation results

In this section, we verify the optimality of the training signals through extensive computer simulations. We consider a MIMO-OFDM system where the number of TX antennas is $Q = 4$, the number of RX antennas is $P = 4$, the number of OFDM subcarriers is $N = 128$, and the length of the cyclic prefix is 32. The number of OFDM symbols for a training signal is set to $M = 4$. A wide sense stationary uncorrelated scattering (WSSUS) model is considered for a multipath channel [23]. A multipath intensity profile of an exponential distribution is used where the number of delay taps is $L = 128$ and an exponentially decaying factor is α . Doppler frequency is assumed to be zero. The optimal training signal shown in Figure 1 is used in the simulations. Five non-optimal training signals are also generated for performance comparisons. Unlike the optimal training signal that satisfies the conditions in (41)

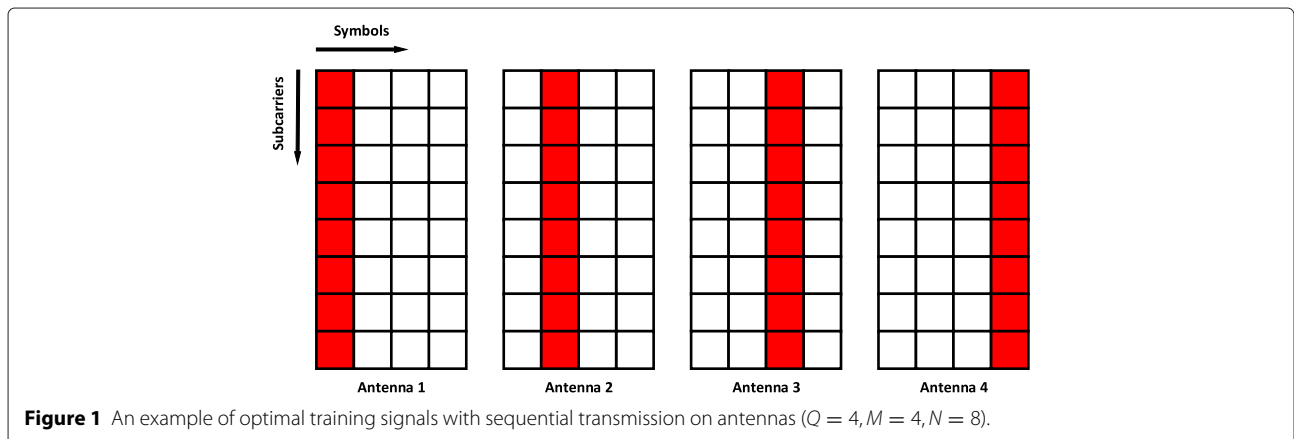


Figure 1 An example of optimal training signals with sequential transmission on antennas ($Q = 4, M = 4, N = 8$).

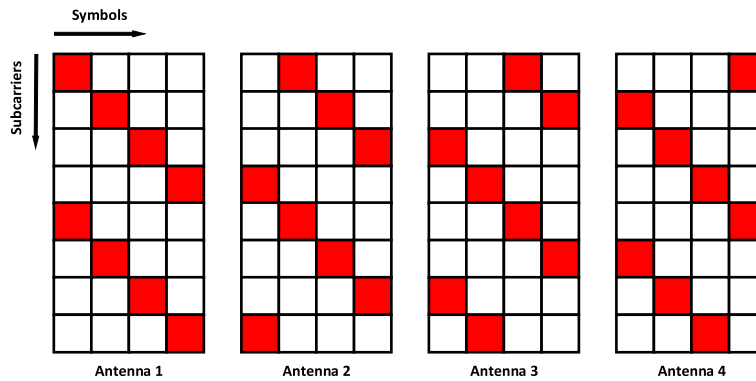


Figure 2 An example of optimal training signals with interlaced transmission on antennas ($Q = 4, M = 4, N = 8$).

and (42), the non-optimal training signals are created at random but they all satisfy the total transmit power E_{total} .

Figure 3 compares the LMMSE channel estimation performances of different training signals. As the SNR increases, the channel estimation error decreases. The optimal training signal shows a considerable performance gap compared to the other non-optimal training signals. For example, a SNR gain using the optimal training signal is more than 5 dB in most cases. Figure 4 shows how the LMMSE channel estimation performance varies with the multipath intensity profile. With a small exponentially decaying factor, large number of multipaths become dominant, resulting in high frequency selectivity and relatively low LMMSE channel estimation performance. As a small exponentially decaying factor increases, the number of dominant multipaths

decreases, resulting in low frequency selectivity and relatively high LMMSE channel estimation performance. Again, the optimal training signal provides a huge performance gap compared to the other non-optimal training signals in all cases. Figure 5 shows the LMMSE channel estimation performance with respect to the antenna dimension. As the number of antennas increases, the MMSE of the optimal training signal does not increase while those of non-optimal training signals continuously increases. Accordingly, the performance gap between the optimal and non-optimal training signals also increase. This is because the optimal training signal is designed such that training signals transmitted from the different antennas are orthogonal. In non-optimal training signal, training signals simultaneously transmitted from different antennas collide and interfere each

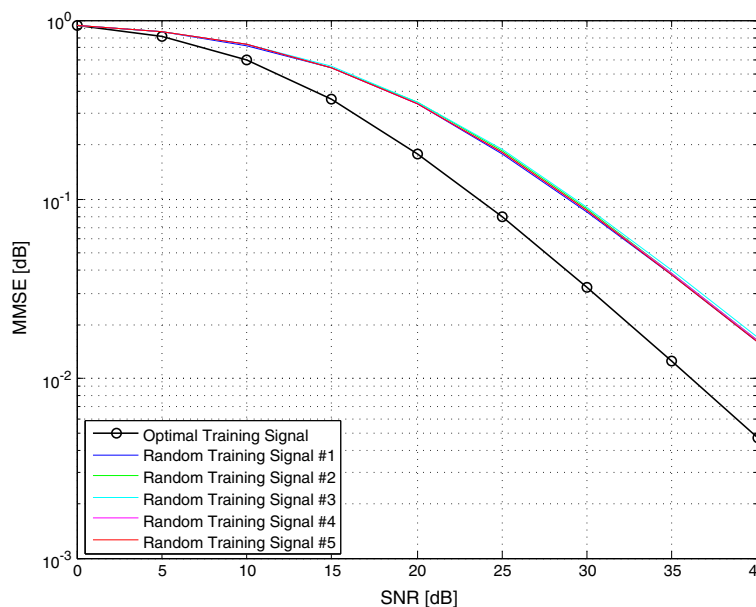


Figure 3 MMSE versus SNR, where $P = 4, Q = 4, M = 4, N = 128, L = 128$, and $\alpha = 10$.

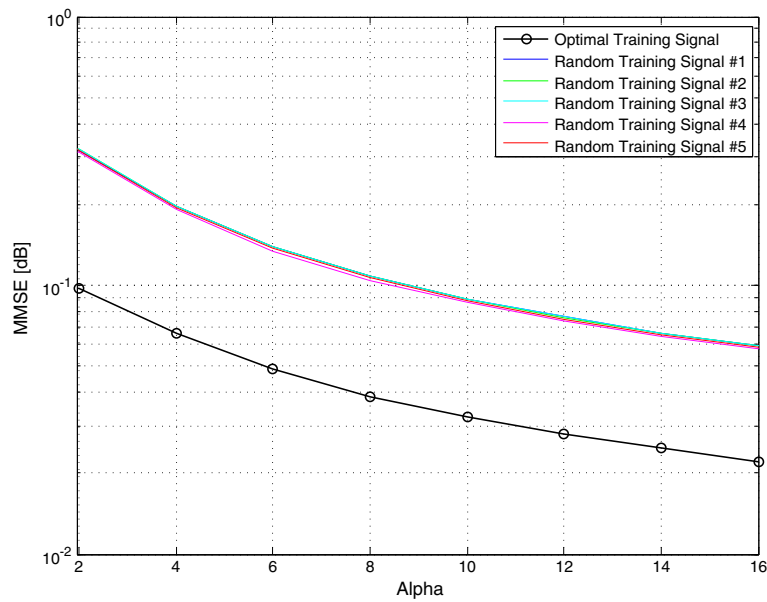


Figure 4 MMSE versus exponentially decaying factor, where $P = 4, Q = 4, M = 4, N = 128, L = 128$, and $SNR = 30$ dB.

other, which degrades the LMMSE channel estimation performance.

7 Conclusions

In this paper, optimality conditions are analytically derived and design guidelines for the optimal training signals are provided for LMMSE channel estimation for MIMO-OFDM. On the basis of the analysis, we clearly reveal that the training signal that satisfies the

following is optimal: (i) the energy of the training signal on each subcarrier is equal, and (ii) on each subcarrier, the training signals transmitted from the different antennas are orthogonal and of equal energy. Interestingly, the optimality conditions of training signals for LMMSE estimator in MIMO-OFDM systems are basically in line with design principles known from single-carrier MIMO systems; training signals across transmit antennas should be orthogonal and training signals should be

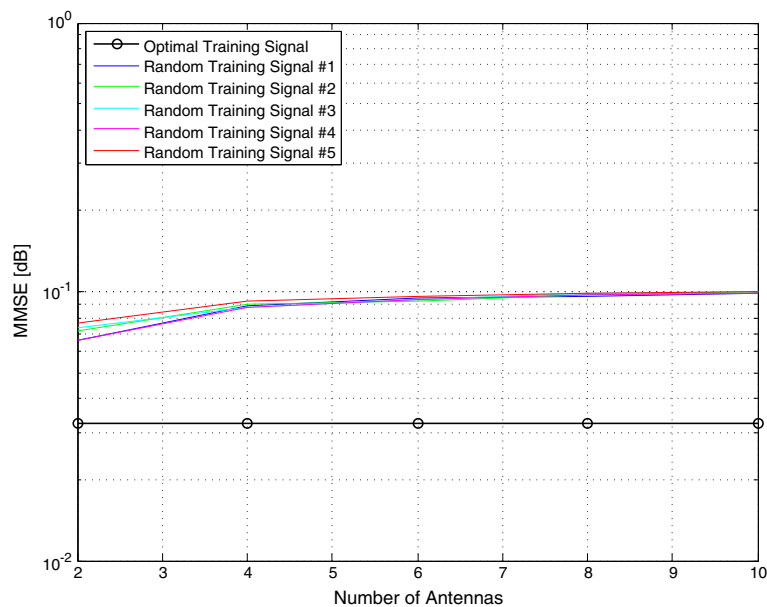


Figure 5 MMSE versus Number of Antennas, where $P = Q = M, N = 128, L = 128, \alpha = 10$, and $SNR = 30$ dB.

equi-powered. We mathematically prove that the simple generalization of the design principles with an additional dimension, i.e., multi-carriers, still holds optimality. This work is important because the optimality conditions of training signals for LMMSE estimation in MIMO-OFDM systems have been mathematically proved. Future research may include an extension of the results to more practical channel statistics, e.g., correlated channels and time-varying channels.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Network Strategy, KT, 90 Bulljeong-ro, Bundang-gu, Seongnam 463-711, Korea. ²Department of Electronic Engineering, Gachon University, 1342 Seongnam-daero, Sujeong-gu, Seongnam 461-701, Korea.

Received: 23 December 2014 Accepted: 26 March 2015

Published online: 16 April 2015

References

- GJ Foschini, Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical J.* **1**(2), 41–59 (1996)
- GJ Foschini, MJ Gans, On limits of wireless communications in a fading environment when using multi-element antennas. *Wireless Personal Commun.* **6**(3), 311–335 (1998)
- IE Telatar, Capacity of multi-antenna Gaussian channels. *Eur Trans. Telecommunications.* **10**(6), 585–595 (1999)
- SB Weinstein, PM Ebert, Data transmission by frequency division multiplexing using the discrete Fourier transform. *IEEE Trans. Commun.* **19**(5), 628–634 (1971)
- I Sohn, JY Ahn, Joint processing of ZF detection and MAP decoding for MIMO-OFDM system. *ETRI J.* **26**(5), 384–390 (2004)
- R Negi, J Cioffi, Pilot tone selection for channel estimation in a mobile OFDM system. *IEEE Trans. Consumer Electron.* **44**(3), 1122–1128 (1998)
- I Barhum, G Leus, M Moonen, Optimal training design for MIMO OFDM systems in mobile wireless channels. *IEEE Trans. Signal Process.* **51**(6), 1615–1624 (2003)
- H Minn, N Al-Dhahir, Optimal training signals for MIMO OFDM channel estimation. *IEEE Trans. Wireless Commun.* **5**(5), 1158–1168 (2006)
- Y-S Choi, PJ Voltz, FA Cassara, On channel estimation and detection for multicarrier signals in Fast and selective Rayleigh. *IEEE Trans. Commun.* **49**(8), 1375–1387 (2001)
- S Ohno, GB Giannakis, Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels. *IEEE Trans. Inf. Theory.* **50**(9), 2138–2145 (2004)
- X Ma, L Yang, GB Giannakis, Optimal training for MIMO frequency-selective fading channels. *IEEE Trans. Wireless Commun.* **4**(2), 453–466 (2005)
- W-C Huang, C-P Li, H-J Li, Optimal pilot sequence design for channel estimation in CDD-OFDM systems. *IEEE Trans. Wireless Commun.* **11**(11), 4006–4016 (2012)
- H Minn, D Munoz, in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*. Pilot designs for channel estimation of OFDM systems with frequency-dependent I/Q imbalances (Budapest, 2009), pp. 1–6
- H Minn, D Munoz, Pilot designs for channel estimation of MIMO OFDM systems with frequency-dependent I/Q imbalances. *IEEE Trans. Commun.* **58**(8), 2252–2264 (2010)
- N Shariati, J Wang, M Bengtsson, in *Proc. IEEE Asilomar Conference on Signals, Systems, and Computers (ACSSC)*. On robust training sequence design for correlated MIMO channel estimation (Pacific Grove, 2012), pp. 504–507
- N Shariati, J Wang, M Bengtsson, in *Proc. IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*. A robust MISO training sequence design (Vancouver, 2013), pp. 4564–4568
- N Shariati, J Wang, M Bengtsson, Robust training sequence design for correlated MIMO channel estimation. *IEEE Trans. Signal Process.* **62**(1), 107–120 (2014)
- C-C Cheng, Y-C Chen, YT Su, in *Proc. IEEE International Conference on Communications (ICC)*. Modelling and estimation of correlated MIMO-OFDM fading channels (Kyoto, 2011), pp. 1–5
- C-T Chiang, CC Fung, Robust training sequence design for spatially correlated MIMO channel estimation. *IEEE Trans. Vehicular Technol.* **60**(7), 2882–2894 (2011)
- SM Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. (Prentice-Hall, Englewood Cliffs, NJ, USA, 1993)
- M Biguesh, AB Gershman, Downlink channel estimation in cellular systems with antenna arrays at base stations using channel probing with feedback. *EURASIP J. Appl. Signal Process.* **9**, 1330–1339 (2004)
- H Lutkepohl, *Handbook of matrices*. (John Wiley & Sons, NY, USA, 1996)
- PA Bello, Characterization of randomly time-variant linear channels. *IEEE Trans. Commun. Syst.* **11**, 360–393 (1963)

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com