一般社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

Sparse Channel Estimation for MIMO-OFDM Amplify-and-Forward Two-Way Relay Networks

Guan GUI† and Fumiyuki ADACHI‡

Department of Communications Engineering, Graduate School of Engineering, Tohoku University
6-6-05 Aza-Aoba, Aramaki, Aoba-ku, Sendai, 980-8579 Japan
E-mail: † gui@mobile.ecei.tohoku.ac.jp, ‡ adachi@ ecei.tohoku.ac.jp

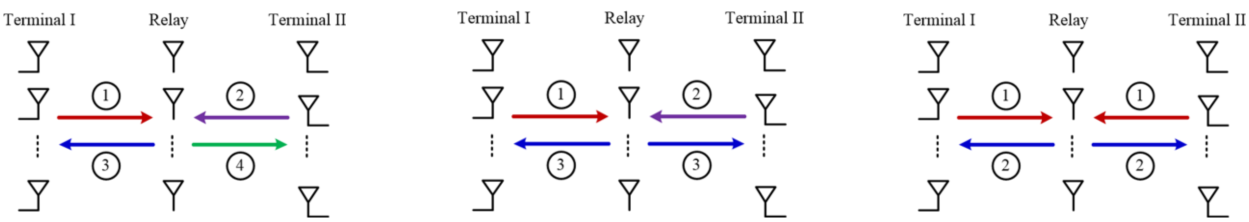
Abstract Accurate channel impulse response (CIR) is required for coherent detection and it can also help improve communication quality of service in next-generation wireless communication systems. One of the advanced systems is multi-input multi-output orthogonal frequency-division multiplexing (MIMO-OFDM) amplify and forward two-way relay networks (AF-TWRN). Linear channel estimation methods, e.g., least square (LS), have been proposed to estimate the CIR. However, these methods never take advantage of channel sparsity and then cause performance loss. In this paper, we propose a sparse channel estimation method to exploit the sparse structure information in the CIR at each end user. Sparse channel estimation problem is formulated as compressed sensing (CS) using sparse decomposition theory and the estimation process is implemented by LASSO algorithm. Computer simulation results are given to confirm the superiority of proposed method over the LS-based channel estimation method.

Keyword Sparse Channel Estimation, MIMO-OFDM, AF-TWRN, Compressed Sensing (CS)

1. Introduction

It is well known that wireless communication technologies are developing rapidly due to that the huge market is promoted by the skyrocketing number of wireless users is in last decades [1]. Until now, there have three promising techniques for broadband wireless communications. The first technique is multiple antenna transmission over multi-input multi-output (MIMO) that is becoming one of the prevail techniques for enhancing system capacity and combating multipath channel fading. The second technique is orthogonal frequency division multiplexing (OFDM) modulation which provides high spectral efficient and robustness mitigates frequency-selective channel fading [1]. The third technique is two-way relay network (TWRN) that implements information exchange in two time slots. When comparing with four-time slots traditional TWRN (see Fig. 1(a)) and three time slots physical-layer TWRN (see Fig. 1(b)) which achieve information exchange, two time slots TWRN (see Fig. 3(c)) can enhances system capacity 66.7% and 100%, respectively. In addition, TWRN can also improve transmission range with limited transmitted power [2]. To take advantage of three techniques fully, combine them into an advanced wireless communication system is one of promising candidate techniques. However, one of the key challenges is how to obtain accurate channel state information (CSI) which is applied for self-interference removal and coherent detection at each terminal.

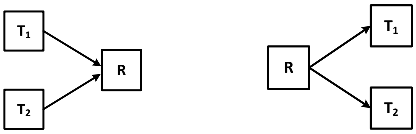
Traditional linear channel estimation methods, e.g, LS [3], have been proposed for MIMO-OFDM AF-TWRN . However, these methods cannot take the advantage of inherent channel sparsity and hence cause performance loss. In this paper, we propose sparse channel approach to exploit such channel sparsity. Sparse channel estimation problem in MIMO-OFDM AF-TWRN is formulated as CS problem. At each terminal, equivalent training signal is constructed to probe equivalent channel vector using least



(a) Four time slots. (b) Three time slots. (c) Two time slots.

Fig. 1. Information exchange using different time slots in MIMO-OFDM AF-TWRN.

absolute shrinkage and selection operator (LASSO) [9]. The performance of propose method will be evaluated by computer simulations.



(a) MAC phase. (b)BC phase.

Fig. 2. Information exchanges under TWNR.

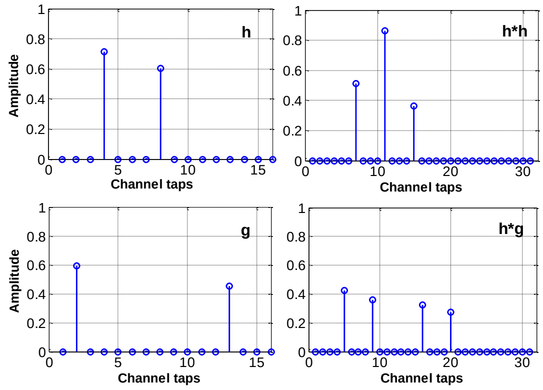


Fig. 3. Example of two individual channels and their cascaded ones.

2. System Model and Problem Formulation

As shown in Fig. 1(c), we consider a MIMO-OFDM AF-TWRN in which two-time slots information exchange between terminal $\mathbb{T}_1$ and terminal $\mathbb{T}_2$ with the help of relay $\mathbb{R}$. Both the two terminals and the relay have N_t and N_r antennas ($N_r \geq N_t$), respectively. Assume that L-length channel vectors between the n_t-th antenna of terminals $\mathbb{T}_i, i=1,2$ and n_r-th antenna of relay $\mathbb{R}$ are denoted by $\boldsymbol{h}_{n_t n_r} = [h_{n_t n_r}(0), h_{n_t n_r}(1), ..., h_{n_t n_r}(L-1)]^T$ and $\boldsymbol{g}_{n_t n_r} = [g_{n_t n_r}(0), g_{n_t n_r}(1), ..., g_{n_t n_r}(L-1)]^T$, respectively. Each channel vector is supported only by K nonzero taps and $K << L$. Suppose that each the nonzero tap is modeled as a complex Gaussian random variable with zero mean and variance $\sigma_{h,l}^2$, and $\sigma_{g,l}^2$ $l=0,1,...,L$. In addition, $\boldsymbol{h}_{n_t n_r}$ and $\boldsymbol{g}_{n_t n_r}$ are assumed invariant in the two time slots information exchange. At time t, suppose that OFDM signal vectors are transmitted from n_t-th antenna of $\mathbb{T}_i$, $i=1,2$ are $\bar{\boldsymbol{s}}_{n_t} = [\bar{s}_{n_t}(0), ..., \bar{s}_{n_t}(n), ..., \bar{s}_{n_t}(N-1)]^T$ and $\bar{\boldsymbol{x}}_{n_t} = [\bar{x}_{n_t}(0), ..., \bar{x}_{n_t}(n), ..., \bar{x}_{n_t}(N-1)]^T$ respectively, where N is the number of subcarriers and $n_r = 1,2,...,N_r$. At the same time, two transmitted power is assumed $E\{\bar{\boldsymbol{s}}_{n_t}^H \bar{\boldsymbol{s}}\} = NP_1$ and $E\{\bar{\boldsymbol{x}}_{n_t}^H \bar{\boldsymbol{x}}\} = NP_2$, respectively.

2.1. MAC phase

In the multi-access (MAC) phase as shown in Fig. 2(a), inverse discrete Fourier transform (IDFT) is computed for frequency-domain signal vectors $\bar{\boldsymbol{s}}_{n_t}$ and $\bar{\boldsymbol{x}}_{n_t}$. The resultant vectors, $\boldsymbol{s}_{n_t} = \boldsymbol{F}^H \bar{\boldsymbol{s}}_{n_t}$ and $\boldsymbol{x}_{n_t} = \boldsymbol{F}^H \bar{\boldsymbol{x}}_{n_t}$, are then cyclic prefix (CP) padded with length $L_{CP} \geq (L-1)$ to avoid inter-block interference (IBI). Here, $\boldsymbol{F}$ is a $N \times N$ discrete Fourier transform (DFT) matrix where entries $f_{mn} = 1/N e^{-j2\pi mn/N}$, $m,n = 0,1,...,N$. After removed the CP, the received signal vector at the n_r-th antenna of $\mathbb{R}$ for $t=1,2,...,T$ is written as

$$\boldsymbol{r}_{n_r} = \sum_{n_t=1}^{N_t} \boldsymbol{H}_{n_t n_r} \boldsymbol{s}_{n_t} + \sum_{n_t=1}^{N_t} \boldsymbol{G}_{n_t n_r} \boldsymbol{x}_{n_t} + \boldsymbol{z}_{n_r}, \tag{1}$$

for $n_r = 1,2,...,N_r$, where $\boldsymbol{H}_{n_t n_r}$ and $\boldsymbol{G}_{n_t n_r}$ are circultant matrices with the first columns of $[\boldsymbol{h}_{n_t n_r}^T \; \boldsymbol{0}_{1\times(N-L)}]^T$ and $[\boldsymbol{g}_{n_t n_r}^T \; \boldsymbol{0}_{1\times(N-L)}]^T$, respectively. The additive noise vector $\boldsymbol{z}_{n_r}(t)$ satisfies $\mathcal{CN}\left(\boldsymbol{0}_{N\times 1}, \sigma_n^2 \boldsymbol{I}_N\right)$. If we collect all received signal vectors $\boldsymbol{r}_{n_r}$, $n_r = 1,2,...,N_r$ at $\mathbb{R}$ to form a $N_r N$-length vector $\boldsymbol{r} = [\boldsymbol{r}_1^T, ..., \boldsymbol{r}_{n_r}^T, ..., \boldsymbol{r}_{N_r}^T]^T$, then the received model in the MAC phase at relay is written as

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{G}\boldsymbol{x} + \boldsymbol{z}, \tag{2}$$

where

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{21} & \cdots & \boldsymbol{H}_{N_t 1} \\ \boldsymbol{H}_{12} & \boldsymbol{H}_{22} & \cdots & \boldsymbol{H}_{N_t 2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}_{1N_r} & \boldsymbol{H}_{2N_r} & \cdots & \boldsymbol{H}_{N_t N_r} \end{bmatrix} \in \mathbb{C}^{N_r N \times N_t N}, \tag{3}$$

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_{11} & \boldsymbol{G}_{21} & \cdots & \boldsymbol{G}_{N_t 1} \\ \boldsymbol{G}_{12} & \boldsymbol{G}_{22} & \cdots & \boldsymbol{G}_{N_t 2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{G}_{1N_r} & \boldsymbol{G}_{2N_r} & \cdots & \boldsymbol{G}_{N_t N_r} \end{bmatrix} \in \mathbb{C}^{N_r N \times N_t N}, \tag{4}$$

$$\boldsymbol{s} = [\boldsymbol{s}_1^T, ..., \boldsymbol{s}_{n_t}^T, ..., \boldsymbol{s}_{N_t}^T]^T \in \mathbb{C}^{N_t N \times 1}, \tag{5}$$

$$\boldsymbol{x} = [\boldsymbol{x}_1^T, ..., \boldsymbol{x}_{n_t}^T, ..., \boldsymbol{x}_{N_t}^T]^T \in \mathbb{C}^{N_t N \times 1}, \tag{6}$$

$$\boldsymbol{z} = [\boldsymbol{z}_1^T, ..., \boldsymbol{z}_{n_r}^T, ..., \boldsymbol{z}_{N_r}^T]^T \in \mathbb{C}^{N_r N \times 1}. \tag{7}$$

According to Eq. (2), the received signal vector $\boldsymbol{r}_{n_r}$ is amplified by a positive coefficients β which is given by

$$\beta = \sqrt{\frac{P_r}{N_t \sum_{l=0}^{L-1} (\sigma_{h,l}^2 P_1 + \sigma_{g,l}^2 P_2) + N_0}}, \tag{8}$$

where P_r is relay's amplify power which is given by $E\{\boldsymbol{r}_{n_r}^H \boldsymbol{r} n_r\} = NP_r$.

2.2. BC phase

Because of system symmetrical in TWRN, without loss of generality, we consider the broadcasting (BC) phase at $\mathbb{T}_1$, as shown in Fig.2(b). Let $\boldsymbol{y}_{n_t}$ denote the received signal vectors at the n_t-th antenna at time $(t+T)$. If we collect N_t received vectors $\boldsymbol{y}_{n_t}$ as $\boldsymbol{y} = [\boldsymbol{y}_1^T, ..., \boldsymbol{y}_{n_t}^T, ..., \boldsymbol{y}_{N_t}^T]^T$, then received signal model can be written as

$$\boldsymbol{y} = \beta \tilde{\boldsymbol{H}} \boldsymbol{H} \boldsymbol{s} + \beta \tilde{\boldsymbol{H}} \boldsymbol{G} \boldsymbol{x} + \beta \tilde{\boldsymbol{H}} \boldsymbol{z} + \boldsymbol{v}, \tag{9}$$

where

$$\tilde{\boldsymbol{H}}=\begin{bmatrix}\boldsymbol{H}_{11} & \boldsymbol{H}_{12} & \cdots & \boldsymbol{H}_{1N_r}\\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} & \cdots & \boldsymbol{H}_{2N_r}\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{H}_{N_t1} & \boldsymbol{H}_{N_t2} & \cdots & \boldsymbol{H}_{N_tN_r}\end{bmatrix}\in\mathbb{C}^{N_tN\times N_rN}, \tag{10}$$

$$\boldsymbol{v}=[\boldsymbol{v}_1,...,\boldsymbol{v}_{n_t},...,\boldsymbol{v}_{N_t}]^T\in\mathbb{C}^{N_tN\times 1}, \tag{11}$$

where $\boldsymbol{v}_{n_t}$ is a noise vector at n_r-th antenna of $\mathbb{T}_1$, satisfying $\boldsymbol{v}_{n_t}\in\mathcal{CN}(\boldsymbol{0}_{N\times 1},\sigma_n^2\boldsymbol{I}_N)$. According to matrix theory [4], circulant matrices $\boldsymbol{H}_{n_tn_r}$ and $\boldsymbol{G}_{n_tn_r}$, $n_r=1,2,\ldots,N_r$, $n_t=1,2,\ldots,N_t$, can be decomposed as

$$\boldsymbol{H}_{n_tn_r}=\boldsymbol{F}^H\boldsymbol{W}_{n_tn_r}\boldsymbol{F}, \tag{12}$$

$$\boldsymbol{G}_{n_tn_r}=\boldsymbol{F}^H\boldsymbol{U}_{n_tn_r}\boldsymbol{F}, \tag{13}$$

respectively, where $(\cdot)^H$ denotes matrix Hermitian transition operation and above diagonal matrices are given by

$$\boldsymbol{W}_{n_tn_r}=Diag\ H_{n_tn_r}(0),...,H_{n_tn_r}(n),...,H_{n_tn_r}(N-1)\ , \tag{14}$$

$$\boldsymbol{U}_{n_tn_r}=Diag\ G_{n_tn_r}(0),...,G_{n_tn_r}(n),...,G_{n_tn_r}(N-1)\ , \tag{15}$$

respectively. Based on the above analysis, it is easy found that the n-th diagonal entries $H_{n_tn_r}(n)$ in Eq. (14) and $G_{n_tn_r}(n)$ in Eq. (15) are obtained by

$$H_{n_tn_r}(n)=\sum\nolimits_{l=0}^{L-1}h_{n_tn_r}(l)e^{-j2nl/N}, \tag{16}$$

$$G_{n_tn_r}(n)=\sum\nolimits_{l=0}^{L-1}g_{n_tn_r}(l)e^{-j2nl/N}, \tag{17}$$

respectively. Therefore, the product of $\beta\boldsymbol{H}_{n_tn_r}\boldsymbol{H}_{n_t'n_r'}$ and $\beta\boldsymbol{G}_{n_tn_r}\boldsymbol{H}_{n_t'n_r'}$ with respect to $n_t,n_t'=1,2,...,N_t$ and $n_r,n_r'=1,2,...,N_r$ can also be written as

$$\beta\boldsymbol{H}_{n_tn_r}\boldsymbol{H}_{n_t'n_r'}=\boldsymbol{F}^H\beta\boldsymbol{W}_{n_tn_r}\boldsymbol{W}_{n_tn_r}\boldsymbol{F}, \tag{18}$$

$$\beta\boldsymbol{H}_{n_tn_r}\boldsymbol{G}_{n_tn_r}=\boldsymbol{F}^H\beta\boldsymbol{W}_{n_tn_r}\boldsymbol{U}_{n_tn_r}\boldsymbol{F}, \tag{19}$$

respectively. Hence, both $\beta\boldsymbol{H}_{n_tn_r}\boldsymbol{H}_{n_tn_r}$ and $\beta\boldsymbol{G}_{n_tn_r}\boldsymbol{H}_{n_tn_r}$ are two circulant matrices which have first columns of $[\beta(\boldsymbol{h}_{n_tn_r}*\boldsymbol{h}_{n_tn_r})\ \ \boldsymbol{0}_{1\times(N-2L+1)}]^T$ and $[\beta(\boldsymbol{h}_{n_tn_r}*\boldsymbol{g}_{n_tn_r})\ \ \boldsymbol{0}_{1\times(N-2L+1)}]^T$, respectively, where '*' denotes convolution operator between two channel vectors. Based on this observation, when the n_1-th row partitions of $\beta\tilde{\boldsymbol{H}}$ multiplies with the n_2-th column partitions of $\boldsymbol{H}$, $n_t,n_t'=1,2,\ldots,N_t$, we can obtain an equivalent composite (linear convolution) channel vector $\boldsymbol{q}_{n_t'n_t}\triangleq[q_{n_t'n_t}(0),...,q_{n_t'n_t}(l),...,q_{n_t'n_t}(2L-1)]$ which is given by

$$\boldsymbol{q}_{n_t'n_t}\triangleq\beta\sum\nolimits_{n_r=1}^{N_r}\boldsymbol{h}_{n_t'n_r}*\boldsymbol{h}_{n_tn_r}. \tag{20}$$

Because of the symmetry of two MIMO channel matrices, we can easy find their symmetry relationship, that is, $\boldsymbol{q}_{n_t'n_t}=\boldsymbol{q}_{n_tn_t'}$. Hence, the product $\beta\tilde{\boldsymbol{H}}\boldsymbol{H}$ is equivalent to provide $(N_t^2+N_t)/2$ independent $(2L-1)$ -length composite channel vectors $\boldsymbol{q}_{n_tn_t'}$ with $n_t,n_t'=1,...,N_t$. Note that $(N_t^2+N_t)/2<N_t^2$ if $N_t>1$. By virtual of the duplication matrix property [5] on channel estimation, it can reduce some amount of complexity which relates to the number of antenna N_t, especially in the case of a relatively large number antennas' communication system. That is to say, the computational complexity is reduce to $\mathcal{O}((N_t^2+N_t)/2)$ rather than $\mathcal{O}(N_t^2)$, where $\mathcal{O}(\cdot)$ denotes the calculation metric of complexity. Due to independent between the two MIMO channel matrices $\tilde{\boldsymbol{H}}$ and $\boldsymbol{G}$, hence, $\beta\tilde{\boldsymbol{H}}\boldsymbol{G}$ is equivalent to generate N_t^2 independent $(2L\text{-}1)$-length composite channel vectors $\boldsymbol{p}_{n_tn_t'}\triangleq[p_{n_tn_t'}(0),....,p_{n_tn_t'}(l),....,p_{n_tn_t'}(2L-1)]^T$ with respect to $n_t,n_t'=1,...,N_t$, where

$$\boldsymbol{p}_{n_tn_t'}\triangleq\beta\sum\nolimits_{n_r=1}^{N_r}\boldsymbol{h}_{n_tn_r}*\boldsymbol{g}_{n_tn_r}. \tag{21}$$

If we define $\tilde{\boldsymbol{F}}=\boldsymbol{I}_{N_t}\otimes\boldsymbol{F}\in\mathbb{C}^{NN_t\times NN_t}$, where '$\otimes$' denotes Kronecker product and $\boldsymbol{I}_{N_t}$ denotes an $N_t\times N_t$ identity matrix, the received signal $\boldsymbol{y}$ in Eq. (9) is transformed to frequency-domain using DFT matrix $\tilde{\boldsymbol{F}}$, then, we have

$$\bar{\boldsymbol{y}}=\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{H}\tilde{\boldsymbol{F}}^H\bar{\boldsymbol{s}}+\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{G}\boldsymbol{F}^H\bar{\boldsymbol{x}}+\bar{\boldsymbol{v}}, \tag{22}$$

where $\bar{\boldsymbol{v}}=\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{z}+\tilde{\boldsymbol{F}}\boldsymbol{v}$ denotes composite noise vector at the $\mathbb{T}_1$. According to Eq. (18) and (19), $\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{H}\tilde{\boldsymbol{F}}^H$ and $\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{G}\tilde{\boldsymbol{F}}^H$ can be given in Eq. (23) and (24), respectively. If we define $\boldsymbol{S}_i=Diag(\bar{\boldsymbol{s}}_i)$ and $\boldsymbol{X}_i=Diag(\bar{\boldsymbol{x}}_i)$ as $N\times N$ diagonal matrices, and collect all composited channel vectors as $\boldsymbol{q}=[\boldsymbol{q}_{11}^T,...,\boldsymbol{q}_{1N_t}^T,\boldsymbol{q}_{22}^T,...,\boldsymbol{q}_{2N_t}^T,...,\boldsymbol{q}_{N_t^2}^T]^T$ and $\boldsymbol{p}=[\boldsymbol{p}_{11}^T,...,\boldsymbol{p}_{N_t1}^T,...,\boldsymbol{p}_{N_t1}^T,...,\boldsymbol{p}_{N_tN_t}^T]^T$, respectively, then two equivalent training signal matrices can be written in Eq. (25) and (26) respectively, where $\boldsymbol{F}_{2L-1}$ is partial DFT matrix by extracting the first $(2L-1)$-columns of $\boldsymbol{F}$.

$$\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{H}\tilde{\boldsymbol{F}}^H=\begin{bmatrix}\sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{1n_r}\boldsymbol{W}_{1n_r} & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{1n_r}\boldsymbol{W}_{2n_r} & \cdots & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{1n_r}\boldsymbol{W}_{N_tn_r}\\ \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{2n_r}\boldsymbol{W}_{1n_r} & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{2n_r}\boldsymbol{W}_{2n_r} & \cdots & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{2n_r}\boldsymbol{W}_{N_tn_r}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{N_tn_r}\boldsymbol{W}_{1n_r} & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{N_tn_r}\boldsymbol{W}_{2n_r} & \cdots & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{N_tn_r}\boldsymbol{W}_{N_tn_r}\end{bmatrix}\in\mathbb{C}^{NN_t\times NN_t}, \tag{23}$$

$$\tilde{\boldsymbol{F}}\beta\tilde{\boldsymbol{H}}\boldsymbol{G}\tilde{\boldsymbol{F}}^{H}=\begin{bmatrix}\sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{1n_r}\boldsymbol{G}_{1n_r} & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{1n_r}\boldsymbol{G}_{2n_r} & \cdots & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{1n_r}\boldsymbol{G}_{N_tn_r}\\ \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{2n_r}\boldsymbol{G}_{1n_r} & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{2n_r}\boldsymbol{G}_{2n_r} & \cdots & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{2n_r}\boldsymbol{G}_{N_tn_r}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{N_tn_r}\boldsymbol{G}_{1n_r} & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{N_tn_r}\boldsymbol{G}_{2n_r} & \cdots & \sum_{n_r=1}^{N_r}\beta\boldsymbol{W}_{N_tn_r}\boldsymbol{G}_{N_tn_r}\end{bmatrix}\in\mathbb{C}^{NN_t\times NN_t}. \quad (24)$$

$$\boldsymbol{S}=\begin{bmatrix}\boldsymbol{S}_1\boldsymbol{F}_{2L-1} & \boldsymbol{S}_2\boldsymbol{F}_{2L-1} & \boldsymbol{S}_3\boldsymbol{F}_{2L-1} & \cdots & \boldsymbol{S}_{N_t}\boldsymbol{F}_{2L-1} & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{0}_{N\times(2L-1)}\\ \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{S}_1\boldsymbol{F}_{2L-1} & \boldsymbol{S}_2\boldsymbol{F}_{2L-1} & \cdots & \boldsymbol{S}_{N_t-1}\boldsymbol{F}_{2L-1} & \boldsymbol{S}_{N_t}\boldsymbol{F}_{2L-1} & \boldsymbol{0}_{N\times(2L-1)} & \vdots\\ \vdots & \boldsymbol{0}_{N\times(2L-1)} & \ddots & \ddots & \vdots & \ddots & \ddots & \boldsymbol{0}_{N\times(2L-1)}\\ \boldsymbol{0}_{N\times(2L-1)} & \cdots & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{S}_1\boldsymbol{F}_{2L-1} & \boldsymbol{S}_2\boldsymbol{F}_{2L-1} & \cdots & \boldsymbol{S}_{N_t-1}\boldsymbol{F}_{2L-1} & \boldsymbol{S}_{N_t}\boldsymbol{F}_{2L-1}\end{bmatrix}\in\mathbb{C}^{NN_t\times(2L-1)N_t(N_t+1)/2} \quad (25)$$

$$\boldsymbol{X}=\begin{bmatrix}\boldsymbol{X}_1\boldsymbol{F}_{2L-1} & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{0}_{N\times(2L-1)} & \cdots & \boldsymbol{X}_{N_t}\boldsymbol{F}_{2L-1} & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{0}_{N\times(2L-1)}\\ \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{X}_1\boldsymbol{F}_{2L-1} & \boldsymbol{0}_{N\times(2L-1)} & \vdots & \cdots & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{X}_{N_t}\boldsymbol{F}_{2L-1} & \boldsymbol{0}_{N\times(2L-1)} & \vdots\\ \vdots & \boldsymbol{0}_{N\times(2L-1)} & \ddots & \boldsymbol{0}_{N\times(2L-1)} & \vdots & \vdots & \boldsymbol{0}_{N\times(2L-1)} & \ddots & \boldsymbol{0}_{N\times(2L-1)}\\ \boldsymbol{0}_{N\times(2L-1)} & \cdots & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{X}_1\boldsymbol{F}_{2L-1} & \cdots & \boldsymbol{0}_{N\times(2L-1)} & \cdots & \boldsymbol{0}_{N\times(2L-1)} & \boldsymbol{X}_{N_t}\boldsymbol{F}_{2L-1}\end{bmatrix}\in\mathbb{C}^{NN_t\times(2L-1)N_t^2} \quad (26)$$

Then received signal model in Eq. (22) can be rewritten concisely as

$$\bar{\boldsymbol{y}}=\boldsymbol{S}\boldsymbol{q}+\boldsymbol{X}\boldsymbol{p}+\bar{\boldsymbol{v}}=\boldsymbol{D}\boldsymbol{b}+\bar{\boldsymbol{v}}, \quad (27)$$

where $\boldsymbol{D}=[\boldsymbol{S},\boldsymbol{X}]$ denotes an equivalent training matrix combined two training signal matrices $\boldsymbol{S}$ and $\boldsymbol{X}$; and $\boldsymbol{b}=[\boldsymbol{q}^T,\boldsymbol{p}^T]^T$ denotes overall channel vector including $\boldsymbol{q}$ and $\boldsymbol{p}$. At the receive side of $\mathbb{T}_1$, channel estimator $\boldsymbol{q}$ is used to remove self-data interference and channel estimator $\boldsymbol{p}$ is applied to extract other users' data information at $\mathbb{T}_1$.

According to the formulated system model in Eq. (27), it is easy found that main object of this paper is to estimate the overall channel vector $\boldsymbol{b}$ using the composite training signal matrix $\boldsymbol{D}$. With respect to Eq. (27), LS based channel estimator $\boldsymbol{b}_{LS}$ can be computed by

$$\boldsymbol{b}_{LS}=\left(\boldsymbol{D}^H\boldsymbol{D}\right)^{-1}\boldsymbol{D}^H\bar{\boldsymbol{y}}=\boldsymbol{b}+(\boldsymbol{D}^H\boldsymbol{D})^{-1}\boldsymbol{D}^H\bar{\boldsymbol{v}}. \quad (28)$$

Since the noise variance of $\bar{\boldsymbol{v}}$ is given by

$$E\{\bar{\boldsymbol{v}}^H\bar{\boldsymbol{v}}\}=N_0(\beta^2N_r\sum_{l=0}^{L-1}\sigma_{h,l}^2+1), \quad (29)$$

then the average MSE of LS channel estimator $\boldsymbol{b}_{LS}$ can be given by

$$MSE\{\boldsymbol{b}_{LS}\}=N_0\left(\beta^2N_r\sum_{l=0}^{L-1}\sigma_{h,l}^2+1\right)Trace\left\{(\boldsymbol{D}^H\boldsymbol{D})^{-1}\right\}. \quad (30)$$

It is well known that the training matrix $\boldsymbol{D}$ has $N_t(3N_t+1)(2L-1)/2$ columns that are normalized in a way such that $\|\boldsymbol{D}\|_F^2=N_t(3N_t+1)(2L-1)/2$, where $\|\cdot\|_F$ denotes the Frobenius norm. Optimal training design for LS-based channel estimation method is the one that subjects to $\boldsymbol{D}^H\boldsymbol{D}=\boldsymbol{I}_{N_t(3N_t+1)(2L-1)/2}$. Hence, we can obtain

$$Trace(\boldsymbol{D}^H\boldsymbol{D})=\|\boldsymbol{D}\|_F^2=N_t(3N_t+1)(2L-1)/2, \quad (31)$$

where $Trace(\boldsymbol{A})$ is defined to be the sum of the elements on the main diagonal of matrix $\boldsymbol{A}$. According to arithmetic–harmonic means inequality, lower bound for the LS channel estimation error can be derived as

$$\begin{aligned}MSE\{\boldsymbol{b}_{LS}\}&\geq\frac{N_0\left(\beta^2N_r\sum_{l=0}^{L-1}\sigma_{h,l}^2+1\right)\left(N_t(3N_t+1)(2L-1)/2\right)^2}{Trace\left\{(\boldsymbol{D}^H\boldsymbol{D})\right\}}\\&=\frac{N_0\left(\beta^2N_r\sum_{l=0}^{L-1}\sigma_{h,l}^2+1\right)\left(N_t(3N_t+1)(2L-1)/2\right)^2}{N_t(3N_t+1)(2L-1)/2}\\&=N_0(\beta^2N_r\sum_{l=0}^{L-1}\sigma_{h,l}^2+1)(3N_t+1)(2L-1)/2.\end{aligned} \quad (32)$$

From the derivation in Eq. (32), the lower bound of LS can be written as $MSE\{\boldsymbol{b}_{LS}\}\sim\mathcal{O}(N_0,\sum_{l=0}^{L-1}\sigma_{h,l}^2,\beta,N_r,N_t,L)$. Generally, linear channel estimation methods, e.g., LS, emphasize on optimal training designing to improve estimation performance while neglect the inherent sparsity of channel.

3. Sparse Channel Estimation

According to the CS [6], [7], accurate sparse channel estimation requires that training signal matrix $\boldsymbol{D}$ be satisfied restricted isometry property (RIP) [8] in high probability. Hence, according to the system model in Eq. (27), optimal sparse channel estimator $\boldsymbol{b}_{opt}$ can be given by

$$\boldsymbol{b}_{opt}=\arg\min_{\boldsymbol{b}}\left\{\frac{1}{2}\|\boldsymbol{D}\boldsymbol{b}-\bar{\boldsymbol{y}}\|_2^2+\lambda\|\boldsymbol{b}\|_0\right\}, \quad (33)$$

where $\|\boldsymbol{b}\|_2$ denotes Euclidean norm which is given by $\|\boldsymbol{b}\|_2^2=\sum_i|b_i|^2$; $\|\boldsymbol{b}\|_0$ denote zero-norm operator which counts their nonzero taps and λ is regularization parameter which trades off the estimation error and sparseness of the channel. Assume the positions set of all channel taps of $\boldsymbol{b}$ is Ω and its nonzero taps set is Γ. The number of nonzero taps of $\boldsymbol{b}$ is T, then the lower bound of sparse channel estimator can be derived as

$$MSE\{\mathbf{b}_{opt}\} = N_0\left(\beta^2 N_r \sum_{l=0}^{L-1}\sigma_{h,l}^2+1\right)Trace\left\{(\mathbf{D}_\Gamma^H\mathbf{D}_\Gamma)^{-1}\right\}.$$
$$\geq \frac{N_0\left(\beta^2 N_r \sum_{l=0}^{L-1}\sigma_{h,l}^2+1\right)T}{Trace\left\{(\mathbf{D}_\Gamma^H\mathbf{D}_\Gamma)\right\}} \quad (34)$$
$$= N_0(\beta^2 N_r \sum_{l=0}^{L-1}\sigma_{h,l}^2+1)T.$$

where $Trace\{(\mathbf{D}_\Gamma^H\mathbf{D}_\Gamma)\}=\mathbf{I}_T$ denotes the optimal signal training for sparse channel estimation. Comparing Eq. (34) to Eq. (32), we can found that the lower bound of optimal channel estimator depends on T rather than overall channel length $N_t(3N_t+1)(2L-1)/2$ of $\mathbf{b}$. If we can estimate positions of nonzero taps of $\mathbf{b}$, then sparse channel estimation performance could be improved. Since solving the optimal sparse channel estimation in Eq. (33) is NP hard problem [7]. Hence, it is necessary to develop alternative suboptimal sparse channel estimation method.

In this paper, we propose a sparse channel estimation method for MIMO-OFDM AF-TWRN and it is implemented by LASSO algorithm [9]. Given a equivalent training matrix $\mathbf{D}$ and a received signal $\bar{\mathbf{y}}$, LASSO based sparse channel estimator $\mathbf{b}_{lasso}$ can be obtained

$$\mathbf{b}_{lasso} = \arg\min_{b}\left\{\frac{1}{2}\|\mathbf{D}\mathbf{b}-\bar{\mathbf{y}}\|_2^2+\lambda\|\mathbf{b}\|_1\right\}, \quad (35)$$

where $\|\mathbf{b}\|_1$ denotes L1-norm which is given by $\|\mathbf{b}\|_1=\sum_i|b_i|$. In a practical system, accurate number of nonzero channel taps is unknown. Hence, to obtain accurate sparse channel estimation, effective training signal design is required. In accordance with the CS [6], [7], two kinds of training design methods, i.e., random Gaussian and random binary, are considered for computer simulation to evaluate our proposed method.

4. Computer Simulation

In this section, we present the simulation results to evaluate sparse channel estimation method in MIMO-OFDM AF-TWRN. Here we compare the performance of the proposed estimator with LS-based channel estimator and adopt 100 independent Monte-Carlo runs for average. The number (N_t, N_r) of transmitter/relay pairs are considered three cases: $(2,2)$, $(2,4)$ and $(4,4)$. All of the channel vectors have same length $L=16$ and $K=1,2,3,4,5,6$, and its positions of nonzero channel taps are randomly generated. Training signal length of each antenna is set as $N=32$ to ensure $N\geq 2L-1$. Transmit power is set as $P_1=P_2=P$ and relay power is allocated as $P_r=2P$. The signal to noise ratio (SNR) is defined as $10\log(P_r/\sigma_n^2)$ at relay and $10\log(P_r/\sigma_n^2)$ at transmitter, respectively.

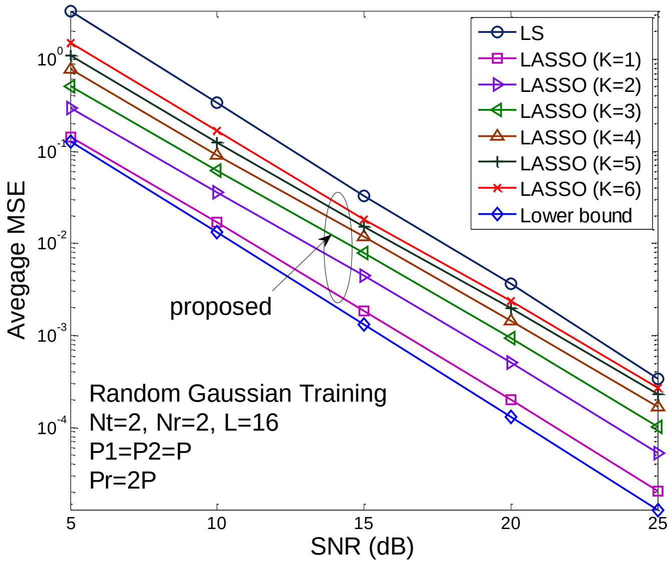


Fig. 4. Performance comparison versus SNR.

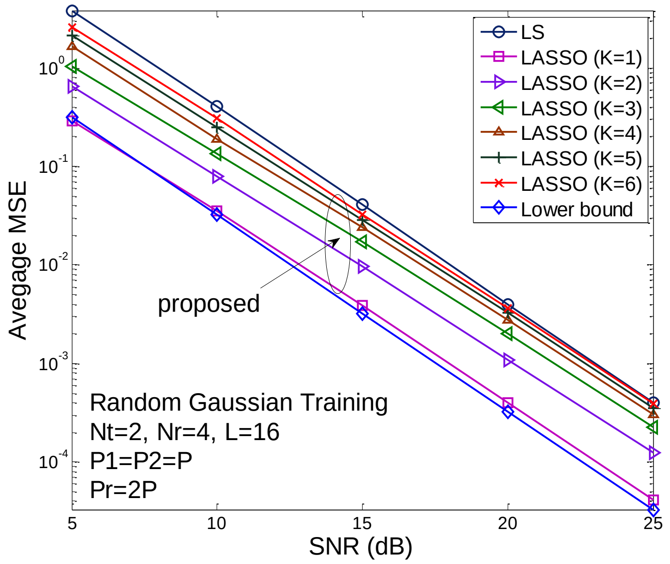


Fig. 5. Performance comparison versus SNR.

Random Gaussian training is considered in Figs.4 and 5, and random binary training is considered in Figs. 6 and 7. From the four figures, we can find that the proposed sparse channel estimator is better than LS one. In addition, the four figures show that LS channel estimator depends on channel length while LASSO one relies on nonzero number K of channel. Note that the lower bound is given by ideal LS channel estimator which is known nonzero taps position of channel. In four experiments, the proposed sparse method works well on different number of nonzero taps of channel. However, for sparser channel estimation, more sparsity can be exploited. In other words, much better performance can be improved. Take the $K=1$ for example, the proposed sparse channel estimator approach to lower bound. On the contrary, channel is approximate parse, e.g., $K=6$, the performance advantage of the proposed method is no longer obvious. When the $K=L=16$, then the proposed sparse channel estimator reduce to LS one. Because single channel vector between each pair of antennas is not exact sparse, it will incur much number of nonzero taps in their cascaded channel. Hence, the proposed method can works well in very sparse channel.

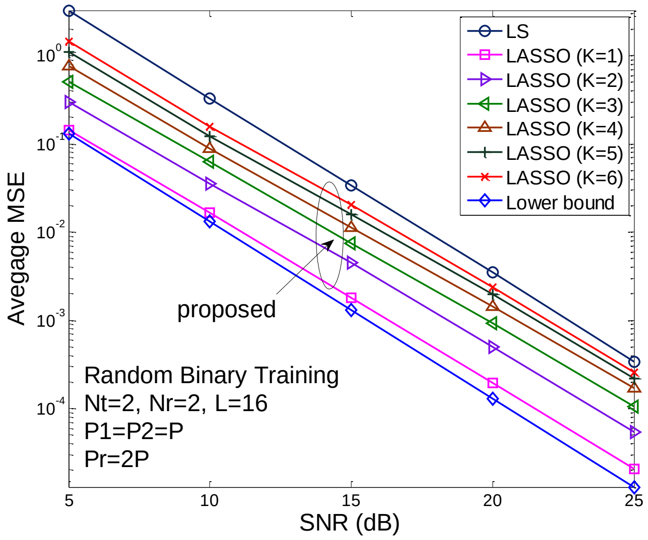


Fig. 6. Performance comparison versus SNR.

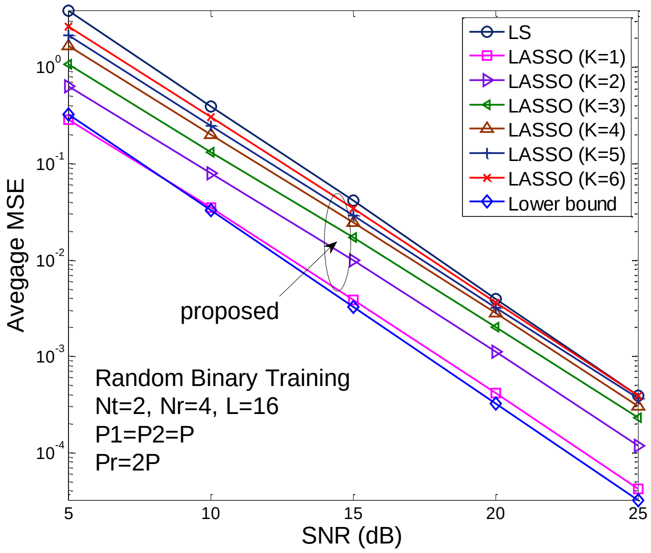


Fig. 7. Performance comparison versus SNR.

5. Conclusion

In this paper, we proposed a sparse channel estimation method which can exploit the extra knowledge of sparse structure as for prior information and hence it can increase spectral efficient or enhance estimation performance when compared with traditional methods. Computer simulation results were showed the performance advantages of our proposed method than LS using MSE standard

Reference

[1] D. Raychaudhuri and N. B. Mandayam, "Frontiers of Wireless and Mobile Communications," *Proceedings of the IEEE*, vol. 100, no. 4, pp. 824–840, Apr. 2012.

[2] F. Gao, R. Zhang, Y. Liang, and S. Member, "Optimal Channel Estimation and Training Design for Two-Way Relay Networks," *IEEE Transactions On Signal Processing*, vol. 57, no. 11, pp. 4443–4455, 2009.

[3] T.-H. Pham and Y.-C. Liang, "Channel estimation and training design for MIMO-OFDM two-way relay systems," in *2012 IEEE International Conference on Communications (ICC)*, 2012, pp. 3703–3707.

[4] R. M. Gray, "Toeplitz and Circulant Matrices: A review," *Foundations and Trends in Communications and Information Theory*, vol. 3, no. 2, pp. 155–239, 2006.

[5] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, vol. 31, no. 4. John Wiley & Sons, Ltd, 1989.

[6] E. J. Candes, J. Romberg, and T. Tao, "Robust Uncertainty Principles : Exact Signal Reconstruction From Highly Incomplete Frequency Information," *IEEE Transctions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[7] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[8] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9–10, pp. 589–592, May 2008.

[9] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society (B)*, vol. 58, no. 1, pp. 267–288, 1996.