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

This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.

Copyright ©2013 by IEICE
absolute shrinkage and selection operator (LASSO) [9]. The performance of the proposed method will be evaluated by computer simulations.

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-TWNR in which two-time slots information exchange between terminal $T_i$ and terminal $T_j$ with the help of relay $R$. Both the two terminals and the relay have $N_i$ and $N_j$ antennas ($N_j \geq N_i$), respectively. Assume that $L$-length channel vectors between the $n_i$-th antenna of terminals $T_i$, $i=1,2$ and $n_i$-th antenna of relay $R$ are denoted by $h_{n_i} = [h_{n_i}(0), h_{n_i}(1), \ldots, h_{n_i}(L-1)]^T$ and $g_{n_i} = [g_{n_i}(0), g_{n_i}(1), \ldots, g_{n_i}(L-1)]^T$, respectively. Each channel vector is supported only by $K$ nonzero taps and $K \ll L$. Suppose that each nonzero tap is modeled as a complex Gaussian random variable with zero mean and variance $\sigma_{i,j}^2$, $i=0,1,\ldots, L$. In addition, $h_{n_i}$ and $g_{n_i}$ are assumed invariant in the two time slots information exchange. At time $t$, suppose that OFDM signal vectors are transmitted from $n_i$-th antenna of $T_i$, $i=1,2$ and $n_i$-th antenna of relay $R$ is given by

$$\begin{bmatrix}
H_{11} & H_{12} & \cdots & H_{1N_j} \\
H_{21} & H_{22} & \cdots & H_{2N_j} \\
\vdots & \vdots & \ddots & \vdots \\
H_{N_i1} & H_{N_i2} & \cdots & H_{N_iN_j}
\end{bmatrix} \in \mathbb{C}^{N_i \times N_j \times N_j}, \quad (3)
$$

and

$$\begin{bmatrix}
G_{11} & G_{12} & \cdots & G_{1N_j} \\
G_{21} & G_{22} & \cdots & G_{2N_j} \\
\vdots & \vdots & \ddots & \vdots \\
G_{N_i1} & G_{N_i2} & \cdots & G_{N_iN_j}
\end{bmatrix} \in \mathbb{C}^{N_j \times N_j \times N_j}, \quad (4)
$$

respectively, where $H_{ij}$ and $G_{ij}$ are channel matrices with the first columns of $[h_{n_i}^T, \alpha_{i+1,N_j-1}]^T$ and $[g_{n_j}^T, \alpha_{i,N_i-1}]^T$, respectively. The additive noise vector $z_n(t)$ satisfies $CN(0, \sigma_n^2 I_N)$. If we collect all received signal vectors $r_n$, $n=1,2,\ldots,N$, at $R$ to form a $N \times N$-length vector $r = [r_1^T, r_2^T, \ldots, r_N^T]^T$, then the received model in the MAC phase at relay is written as

$$r = Hs + Gx + z,$$

where

$$H = \begin{bmatrix}
H_{11} & H_{12} & \cdots & H_{1N_j} \\
H_{21} & H_{22} & \cdots & H_{2N_j} \\
\vdots & \vdots & \ddots & \vdots \\
H_{N_i1} & H_{N_i2} & \cdots & H_{N_iN_j}
\end{bmatrix} \in \mathbb{C}^{N_i \times N_j \times N_j}, \quad (3)
$$

and

$$G = \begin{bmatrix}
G_{11} & G_{12} & \cdots & G_{1N_j} \\
G_{21} & G_{22} & \cdots & G_{2N_j} \\
\vdots & \vdots & \ddots & \vdots \\
G_{N_i1} & G_{N_i2} & \cdots & G_{N_iN_j}
\end{bmatrix} \in \mathbb{C}^{N_j \times N_j \times N_j}, \quad (4)
$$

and

$$s = [s_1^T, s_2^T, \ldots, s_N^T]^T \in \mathbb{C}^{N \times N}, \quad (5)
$$

and

$$x = [x_1^T, x_2^T, \ldots, x_N^T]^T \in \mathbb{C}^{N \times N}, \quad (6)
$$

and

$$z = [z_1^T, z_2^T, \ldots, z_N^T]^T \in \mathbb{C}^{N \times N}. \quad (7)
$$

According to Eq. (2), the received signal vector $r_n$ is amplified by a positive coefficients $\beta$ which is given by

$$\beta = \sqrt{P_r \sum_{l=1}^{N} (\sigma_{l}^2 + \sigma_{l}^2 + \sigma_{l}^2)N_0}, \quad (8)
$$

where $P_r$ is relay’s amplify power which is given by $E(r_n^H r_n) = N P_r$.

2.2. BC phase

Because of system symmetrical in TWNR, without loss of generality, we consider the broadcasting (BC) phase at $T_i$, as shown in Fig.2(b). Let $y_n$ denote the received signal vectors at the $n_i$-th antenna at time $(t+T)$. If we collect $N_i$ received vectors $y_n$, as $y = [y_1^T, y_2^T, \ldots, y_{N_i}^T]^T$, then received signal model can be written as

$$y = \beta HHs + \beta H\beta Gx + \beta H\beta z + v, \quad (9)$$

where $v$ is the noise vector.
then, we have \( \mathbf{H} \). Therefore, the product of \( \mathbf{H} \), especially in the case of a 2-dimensional MIMO channel, can be decomposed as \( \mathbf{H}_{n,n'} = \mathbf{F}^\dagger \mathbf{W}_{n,n} \mathbf{F} \) and \( \mathbf{G}_{n,n'} = \mathbf{F}^\dagger \mathbf{U}_{n,n} \mathbf{F} \), respectively, where \((\cdot)^\dagger\) denotes Hermitian transpose operation and above diagonal matrices are given by
\[
\mathbf{W}_{n,n} = \text{Diag} \left[ H_{n,n}(0), ..., H_{n,n}(n), ..., H_{n,n}(N-1) \right],
\]
\[
\mathbf{U}_{n,n} = \text{Diag} \left[ G_{n,n}(0), ..., G_{n,n}(n), ..., G_{n,n}(N-1) \right],
\]
respectively. Based on the above analysis, it is easy found that the \( n \)-th diagonal entries \( H_{n,n}(n) \) in Eq. (14) and \( G_{n,n}(n) \) in Eq. (15) are obtained by
\[
H_{n,n}(n) = \sum_{l=0}^{N-1} h_{n,n}(l) e^{-j2\pi l n/N},
\]
\[
G_{n,n}(n) = \sum_{l=0}^{N-1} g_{n,n}(l) e^{-j2\pi l n/N},
\]
respectively. Therefore, the product of \( \beta \mathbf{H}_{n,n} \mathbf{H}_{n',n'}^\dagger \) and \( \beta \mathbf{G}_{n,n} \mathbf{G}_{n',n'}^\dagger \) with respect to \( n,n' = 1,2, ..., N_c \) and \( n,n' = 1,2, ..., N_c \) can also be written as
\[
\beta \mathbf{H}_{n,n} \mathbf{H}_{n',n'}^\dagger = \mathbf{F}^\dagger \mathbf{W}_{n,n} \mathbf{W}_{n,n'} \mathbf{F},
\]
\[
\beta \mathbf{G}_{n,n} \mathbf{G}_{n',n'}^\dagger = \mathbf{F}^\dagger \mathbf{U}_{n,n} \mathbf{U}_{n,n'} \mathbf{F},
\]
respectively. Hence, both \( \beta \mathbf{H}_{n,n} \mathbf{H}_{n',n'}^\dagger \) and \( \beta \mathbf{G}_{n,n} \mathbf{G}_{n',n'}^\dagger \) are two circulant matrices which have first columns of \( \beta \left( h_{n,n} * h_{n,n}^* \right) a_{0,0}, ..., a_{0,2L-1} \right] \) and \( \beta \left( g_{n,n} * g_{n,n}^* \right) a_{0,0}, ..., a_{0,2L-1} \right] \), respectively, where \( \ast \) denotes convolution operator between two channel vectors. Based on this observation, when the \( n \)-th column partitions of \( \mathbf{H} \), \( n,n' = 1,2, ..., N_c \), we can obtain an equivalent composite (linear convolution) channel vector \( \mathbf{q}_{n,n'} \triangleq \left[ q_{n,n'}(0), ..., q_{n,n'}(l), ..., q_{n,n'}(2L-1) \right] \) which is given by
\[
\mathbf{q}_{n,n'} \triangleq \beta \sum_{n=1}^{N_c} h_{n,n} * h_{n,n'}.
\]
Because of the symmetry of two MIMO channel matrices, we can easily find their symmetry relationship, that is, \( \mathbf{q}_{n,n'} = \mathbf{q}_{n',n} \). Hence, the product \( \beta \mathbf{H} \mathbf{H}^\dagger \) is equivalent to provide \( (N_c^2 + N_c)/2 \) independent \( (2L-1) \)-length composite channel vectors \( \mathbf{q}_{n,n'} \) with \( n,n' = 1,2, ..., N_c \). Note that \( (N_c^2 + N_c)/2 < N_c^2 \) if \( N_c > 1 \). By virtue of the duplication matrix property [5] on channel estimation, it can reduce some amount of complexity which relates to the number of antenna \( N_c \), especially in the case of a relatively large number antennas’ communication system.

That is to say, the computational complexity is reduce to \( \mathcal{O}(N_c^2 + N_c)/2 \) rather than \( \mathcal{O}(N_c^2) \), where \( \mathcal{O}(\cdot) \) denotes the calculation metric of complexity. Due to independent between the two MIMO channel matrices \( \mathbf{H} \) and \( \mathbf{G} \), hence, \( \beta \mathbf{H} \mathbf{G}^\dagger \) is equivalent to generate \( N_c^2 \) independent \( (2L-1) \)-length composite channel vectors \( \mathbf{p}_{n,n'} \triangleq \left[ p_{n,n'}(0), ..., p_{n,n'}(l), ..., p_{n,n'}(2L-1) \right] \) with respect to \( n,n' = 1,2, ..., N_c \), where
\[
\mathbf{p}_{n,n'} \triangleq \beta \sum_{n=1}^{N_c} h_{n,n} * g_{n,n'}.
\]
If we define \( \tilde{\mathbf{F}} = \mathbf{I}_N \otimes \mathbf{F} \in \mathbb{C}^{N_c N_c \times N_c} \), where \( \otimes \) denotes Kronecker product and \( \mathbf{I}_N \) denotes an \( N_c \times N_c \) identity matrix, the received signal \( \mathbf{y} \) in Eq. (9) is transformed to frequency-domain using DFT matrix \( \tilde{\mathbf{F}} \), then, we have
\[
\mathbf{y} = \tilde{\mathbf{F}} \beta \mathbf{H} \mathbf{H}^\dagger \mathbf{x} + \tilde{\mathbf{F}} \beta \mathbf{H} \mathbf{G}^\dagger \mathbf{x} + \mathbf{v},
\]
where \( \mathbf{v} = \tilde{\mathbf{F}} \mathbf{H} \mathbf{z} + \tilde{\mathbf{F}} \mathbf{v} \) denotes composite noise vector at the \( \mathbb{T}_c \). According to Eq. (18) and (19), \( \tilde{\mathbf{F}} \beta \mathbf{H} \mathbf{H}^\dagger \mathbf{H} \) and \( \tilde{\mathbf{F}} \beta \mathbf{H} \mathbf{G}^\dagger \mathbf{H} \) can be given in Eq. (23) and (24), respectively. If we define \( \mathbf{S}_t = \text{Diag}(\mathbf{x}) \) and \( \mathbf{X}_t = \text{Diag}(\mathbf{x}) \) as \( N_c \times N_c \) diagonal matrices, and collect all composited channel vectors as \( \mathbf{q} = \left[ q_{1,1}^T, q_{2,2}^T, ..., q_{N_c,1}^T, q_{N_c,2}^T, ..., q_{1,N_c}^T \right] \) and \( \mathbf{p} = \left[ p_{1,1}^T, p_{2,2}^T, ..., p_{N_c,1}^T, p_{N_c,2}^T, ..., p_{1,N_c}^T \right] \) respectively, then two equivalent training signal matrices can be written in Eq. (25) and (26) respectively, where \( \mathbf{F}_{L+1} \) is partial DFT matrix by extracting the first \( (2L-1) \)-columns of \( \mathbf{F} \).
Then received signal model in Eq. (22) can be rewritten concisely as
\[
\tilde{y} = S q + X p + \bar{v} = D b + \bar{v},
\]
where \( D = [S, X] \) denotes an equivalent training matrix combined two training signal matrices \( S \) and \( X \); and \( b = [q^T, p^T]^T \) denotes overall channel vector including \( q \) and \( p \). At the receive side of \( T_t \), channel estimator \( q \) is used to remove self-data interference and channel estimator \( p \) is applied to extract other users’ data information at \( T_t \).

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 \( b \) using the composite training signal matrix \( D \). With respect to Eq. (27), LS based channel estimator \( b_{LS} \) can be computed by
\[
b_{LS} = (D^T D)^{-1} D^T \tilde{y} = b + (D^T D)^{-1} D^T \bar{v}.
\]
Since the noise variance of \( \bar{v} \) is given by
\[
E[\bar{v}^T \bar{v}] = N_0 (\beta^2 N_s \sum_{i=0}^{L-1} \sigma_{\lambda,i}^2 + 1),
\]
the average MSE of LS channel estimator \( b_{LS} \) can be given by
\[
MSE[b_{LS}] = N_0 (\beta^2 N_s \sum_{i=0}^{L-1} \sigma_{\lambda,i}^2 + 1) \text{Trace}
\[
\left[(D^T D)^{-1}\right].
\]
It is well known that the training matrix \( D \) has \( N_c (3N_s + 1)(2L-1)/2 \) columns that are normalized in a way such that \( \|D\|_F^2 = N_s (3N_s + 1)(2L-1)/2 \), where \( \|D\|_F \) denotes the Frobenius norm. Optimal training design for LS-based channel estimation method is the one that subjects to \( D^T D = I_{N_s (3N_s + 1)(2L-1)/2} \). Hence, we can obtain
\[
\text{Trace}(D^T D) = \|D\|_F^2 = N_s (3N_s + 1)(2L-1)/2.
\]

3. Sparse Channel Estimation

According to the CS [6], [7], accurate sparse channel estimation requires that training signal matrix \( D \) be satisfied restricted isometry property (RIP) [8] in high probability. Hence, according to the system model in Eq. (27), optimal sparse channel estimator \( b_{opt} \) can be given by
\[
b_{opt} = \arg \min_{b} \left\{ \frac{1}{2} \|Db - \tilde{y}\|_F^2 + \lambda \|b\|_0 \right\},
\]
where \( \|b\|_0 \) denotes Euclidean norm which is given by \( \|b\|_0 = \sum |b_i| \); \( \|b\|_F \) denote zero-norm operator which counts their nonzero taps and \( \lambda \) is regularization parameter which trades off the estimation error and sparseness of the channel. Assume the positions set of all channel taps of \( b \) is \( \Omega \) and its nonzero taps set is \( \Gamma \). The number of nonzero taps of \( b \) is \( T \), then the lower bound of sparse channel estimator can be derived as
MSE[\mathbf{b}_{opt}]=N_s\left(\beta^2\mathbf{N}\sum_{i=0}^{L-1}\sigma_i^2+1\right)\text{Trace}\left((\mathbf{D}_r^T\mathbf{D}_r)^{-1}\right).
\geq N_s\left(\beta^2\mathbf{N}\sum_{i=0}^{L-1}\sigma_i^2+1\right)\text{Trace}\left((\mathbf{D}_r^T\mathbf{D}_r)^{-1}\right)
= N_s\beta^2\mathbf{N}\sum_{i=0}^{L-1}\sigma_i^2+1\right)\text{Trace}\left((\mathbf{D}_r^T\mathbf{D}_r)^{-1}\right).
\tag{34}

where \text{Trace}\left((\mathbf{D}_r^T\mathbf{D}_r)^{-1}\right)=I_T$ denotes the optimal signal training for sparse channel estimation. Comparing Eq. (34) to Eq. (32), we can find that the lower bound of optimal channel estimator depends on $T$ rather than overall channel length $N/(3N_r+1)(2L-1)/2$ 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 $\mathbf{y}$, LASSO based sparse channel estimator $\mathbf{b}_{\text{sparse}}$ can be obtained

$$
\mathbf{b}_{\text{sparse}} = \arg\min_\mathbf{b}\left\{ \frac{1}{2}\|\mathbf{D}\mathbf{b}-\mathbf{y}\|^2 + \lambda\|\mathbf{b}\|_1 \right\},
$$

where $\|\mathbf{b}\|_1$ denotes L1-norm which is given by $\|\mathbf{b}\|_1 = \sum |b_i|$. 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 $(N_r, N_r)$ 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 $10\log(P_1/\sigma_n^2)$ at relay and $10\log(P_1/\sigma_n^2)$ at transmitter.

![Fig. 4. Performance comparison versus SNR.](image)

![Fig. 5. Performance comparison versus SNR.](image)
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 sparse, 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.

![Graph](image_url)

Fig. 6. Performance comparison versus SNR.

![Graph](image_url)

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


