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

Compressed Channel Sensing for Two-way Relay Network

Guan GUI†‡ Qun WAN† and Fumiyuki ADACHI ‡

† Dept Electrical Engineering, University of Electrical Science and Technology of China, Chengdu, 611-731 China

‡ Dept. Electrical and Communication, Tohoku University, 6-6-05 Aza-Aoba, Aoba-ku, Sendai, 980-8579 Japan

E-mail: †‡gui@mobile.ecei.tohoku.ac.jp, †wanqun@uestc.edu.cn, ‡adachi@ecei.tohoku.ac.jp

Abstract Two-way relay network (TWRN) was introduced to realize high-data rate transmission and to improve spatial diversity over the frequency-selective fading channel. However, channel state information (CSI) is needed due to the requirement from coherent data detection and the self-data removal at terminals. Traditional training-based linear probing techniques are able to achieve the accurate CSI by using enough training sequence. However, these methods may lead to a low efficiency on the bandwidth due to that based on the implicit assumption of a rich underlying multipath, while neglected inherent channel sparsity in the TWRN. Unlike the previous methods, we propose a compressed channel sensing method (sometimes be called compressed sensing–based sparse channel estimation method) which exploits the sparse structure, and hence provide significant improvements in MSE performance when compared with conventional LS-based linear channel probing strategies in the TWRN. Simulation results confirm performance of the proposed method.

Keyword Compressed channel sensing (CCS), two-way relay network (TWRN), sparse channel estimation, Restricted Isometry Property (RIP).

1. Introduction

1.1. Background and motivation

As the increasing requirement of high system throughput and high quality wireless service, traditional single-antenna transmission techniques satisfy it due to low system capacity. It is natural that some people develop multiple antennas to boost the system capacity [1]. Unfortunately, packing many antennas onto a small mobile terminal often encounter difficulty of the size limitation.

In order to overcome this limitation, relay networks have drawn great attentions in recent years due to two aspects: first, relay node can provide high space diversity, that is to say, boosts high transmission capacity and second, terminal users packing single-antenna can obtain the increasing requirement of high network throughput [2] [3, 4]. Furthermore, relay nodes can either be provided by the telecommunication company or could be obtained from cooperating terminals of other idle users. For the latter situation, every user in this cooperative network can acted as a relay or relay candidate and exchange information for certain period.

In this paper, two-way relay network (TWRN) under amplify-to-forward (AF) protocol is investigated, where two terminals, $\mathbb{T}_1$ and $\mathbb{T}_2$, exchange information based on the assistance of a relay $\mathbb{R}$. TWRN has been intensively studied due to their capability of enhancing the transmission capacity and providing the spatial diversity for single-antenna wireless transceivers by employing the relay nodes as "virtual" antennas [5]. These kinds of relay network are usually divided into two time slots. During first time slot, terminal users transmit its own information to all relays. During second time slot, the relays amplify and retransmit the combined information to the terminal users.

Although these kinds of networks have many advantages, while accurate channel state information (CSI) are needed at the terminal users. Because of self-information removal and coherence detection of each terminal, not only the channel state information (CSI) needs to know from relay to itself but also the CSI from the other terminal to relay. In [6], the authors considered channel estimation problem at relay, and then broadcast two channel information to terminal users. This method has two disadvantages: cast computational burden to relay users and channel information was interfered again by noise and channel fading after broadcasted to terminal users. Training-based linear channel estimation methods have been proposed in [5, 7, 8] for TWRN under AF protocol. The authors considered the optimal training sequence design and linear probing methods based on the implicit assumption of a rich underlying multipath environment. In recent years, numerous physical channel measurements have shown that the multipath channels

Copyright ©20●● by IEICE

tend to exhibit cluster or sparse structures in which the majority of the channel taps end up being either zero or below the noise floor [9]. However, traditional training-based linear methods that rely on linear reconstruction strategies at the receiver seem incapable of exploiting the potential sparse multipath channels, thereby leading to the overutilization of the key communication resources such as energy and bandwidth. In other words, exploiting this channel sparsity will improve the efficiency on the bandwidth and energy with some effective channel estimation techniques. To be better understanding the concept of sparse channels, we give a group of figures to describe sparse channels both in P2P and TWRN which are shown in Fig.1. We assume amplitudes of channel impulse are exact sparse, that is to say, the channel amplitudes are supported by several dominant channel taps and the other taps are zero.

With the theoretical development of compressed sensing, a number of researches have proposed sparse channel probing schemes on point-to-point (P2P) communication systems either single-antenna [10] or multiple-antenna [11]. These studies classify numerical methods and theoretical analysis. The former schemes have been verified directly the feasibility of the proposed schemes but have not studied from theoretic perspective. While the latter focus on quantitative theoretical analysis of the their performance in terms of the reconstruction error while neglect the acceptable computational complexity [11] in practical applications. All of the contributions above sparse channel estimation schemes are limited in P2P communication systems.

1.2. Objective

In this paper, we first propose a sparse channel estimation technique with compressed sensing method for TWRN which differs from the previous estimation problem in P2P system. The proposed method divided three steps: First, probe position information of dominant channel taps; second, estimate the dominant channel taps and last, prone the non-dominant channel taps as zero. We can find that the proposed method can capture the sparse structure at first step in two-way relay channel (TWRC). When comparing to linear channel estimation methods, the proposed method has a higher spectral efficiency to obtain the same estimation performance. Next, we present a theoretical analysis on sparse channel estimation method in TWRN and confirm it by computer simulations.

1.3. Scope

The remainder of the paper is organized as follows. In section II, we first review the P2P channel model and then model the sparse TWRC problem. In section III, we review the compressed sensing and restate the CoSaMP [12], and then give the detailed steps on sparse two-way relay channel estimation problem. In section IV, we give a variety of simulation results and discussions on the presented method. Finally, conclusions are given in Section V.

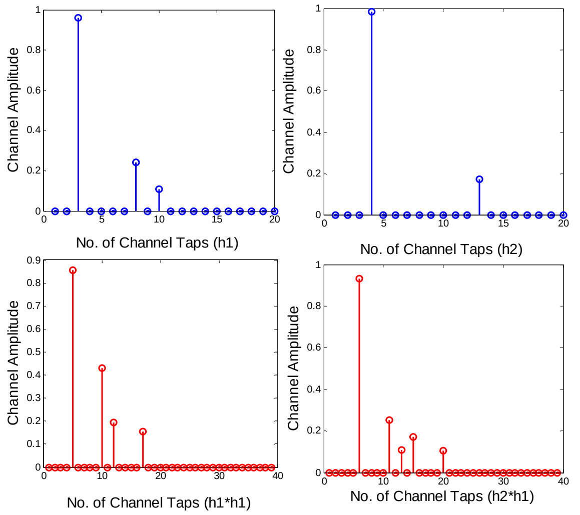

Figure 1. Example of channels amplitude with sparse behavior. Above the first row where two sub-figures are shown as channel vectors $\mathbf{h}_1$ and $\mathbf{h}_2$ in P2P communication systems, where have three dominant channel taps while channel length are 20. The second row where two subfigures (composite channel vectors $\mathbf{h}_1 * \mathbf{h}_1$ and $\mathbf{h}_2 * \mathbf{h}_1$) have four and five dominant channel taps, respectively, both the channels length are 39.

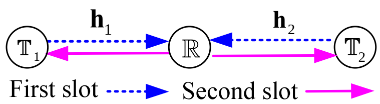

Figure 2. Typical system model in two-way relay networks.

2. Sparse Channel Model

To begin discuss the sparse channel model in the TWRN, we first introduce its network model.

2.1. Two-way relay system model

Consider a typical TWRN system model where the two terminal users, $\mathbb{T}_1$ and $\mathbb{T}_2$, exchange information with the assistant of an relay as shown in Fig.1. The relay and two terminals are assumed to have one antenna each. Let channel impulse responses $\mathbf{h}_1$ and $\mathbf{h}_2$ be the deterministic (but unknown) K-sparse impulse responses of the frequency-selective fading channels between terminals to relay. Both $\mathbf{h}_1$ and $\mathbf{h}_2$ have same physical property modeled in Eq. (1) and they are assumed to be zero-mean circularly symmetric complex Gaussian random variables with variances $\sigma_{\mathbf{h}_1}^2$ and $\sigma_{\mathbf{h}_2}^2$, respectively, and are independent of each other. The average transmission powers of $\mathbb{T}_1$, $\mathbb{T}_2$ and relay are P_1, P_2, and P_r, respectively. For the time being, we assume perfect synchronization among three terminals.

In the TWRN, at first time slot, two terminals $\mathbb{T}_1$ and $\mathbb{T}_2$ send out data symbols $\mathbf{d}_1$ and $\mathbf{d}_2$, respectively, the power constraints of the transmission is $E\{\|\mathbf{d}_i\|_2^2\} = P_i$, $i = 1,2$, where $E[.]$ is the ensemble average operation. Hence, the relay R receives

$$\mathbf{y}_r = \sqrt{2E/T}\left(\mathbf{d}_1 * \mathbf{h}_1 + \mathbf{d}_2 * \mathbf{h}_2\right) + \mathbf{n}_r , \tag{1}$$

where E and T are the data symbol energy and time duration, respectively, and $*$ denotes discrete-time convolution operator, $\mathbf{n}_r$ is a complex Gaussian noise with zero mean and covariance matrix $E\{\mathbf{n}_r\mathbf{n}_r^H\} = \mathbf{I}_{N+L-1}$.

To minimize the computational burden on $\mathbb{R}$, which broadcasts the scaled version of this superposition signal, i.e., $\mathbb{R}$ works under the amplify-and-forward (AF) protocol. The received signal in (3) is then amplified by a relay factor

$$\alpha = \sqrt{\frac{P_r}{\sigma_{\mathbf{h}_1}^2 P_1 + \sigma_{\mathbf{h}_2}^2 P_2 + \sigma_n^2}} , \tag{2}$$

and broadcasts it. Without loss of generality, we consider only the channel estimation problem at $\mathbb{T}_1$, while the discussion over $\mathbb{T}_2$, can be made correspondingly. The received signal at $\mathbb{T}_1$ can be written as follows:

$$\begin{aligned}\mathbf{y}_1 &= \alpha \mathbf{y}_r * \mathbf{h}_1 + \mathbf{n}_1 \\ &= \alpha\sqrt{2E/T}\left(\mathbf{d}_1 * \mathbf{h}_1 * \mathbf{h}_1 + \mathbf{d}_2 * \mathbf{h}_2 * \mathbf{h}_1\right) + \alpha \mathbf{n}_r * \mathbf{h}_1 + \mathbf{n}_1 , \\ &= \alpha\sqrt{2E/T}\left(\mathbf{d}_1 * \mathbf{b} + \mathbf{d}_2 * \mathbf{c}\right) + \mathbf{n}\end{aligned} \tag{3}$$

where $\mathbf{b} = \mathbf{h}_1 * \mathbf{h}_1$ and $\mathbf{c} = \mathbf{h}_2 * \mathbf{h}_1$ denote convolution channel vector, respectively; And $\mathbf{n} = \alpha \mathbf{n}_r * \mathbf{h}_1 + \mathbf{n}_1$ is a complex Gaussian noise with $\mathcal{CN}(\mathbf{0}_{\tilde{N}}, \sigma_n^2(\alpha^2|\mathbf{h}_1|^2 + \mathbf{I}_{\tilde{N}})$. Assuming the perfect information of $\hat{\mathbf{b}}$, the self-data removal is done as

$$\hat{\mathbf{y}}_1 = \mathbf{y}_1 - \alpha\sqrt{2E/T}\left(\mathbf{d}_1 * \hat{\mathbf{b}}\right). \tag{4}$$

The maximum likelihood data detection (MLD) is done at the terminal $\mathbb{T}_1$ as

$$\begin{aligned}\hat{\mathbf{d}}_2 &= \arg\max_{\mathbf{d}_2} P(\hat{\mathbf{y}}_1|\mathbf{d}_2) \\ &= \arg\max_{\mathbf{d}_2} \frac{1}{\pi\sigma_n^2\left(\alpha|\mathbf{h}_1|^2+1\right)} \exp\left\{-\frac{\left|\hat{\mathbf{y}}_1 - \alpha\sqrt{2E/T}\left(\mathbf{d}_2 * \hat{\mathbf{c}}\right)\right|^2}{\pi\sigma_n^2\left(\alpha^2|\mathbf{h}_1|^2+1\right)}\right\} \\ &= \arg\min_{\mathbf{d}_2} \left|\hat{\mathbf{y}}_1 - \alpha\sqrt{2E/T}\left(\mathbf{d}_2 * \hat{\mathbf{c}}\right)\right|^2 .\end{aligned} \tag{5}$$

Hence, from Eq. (5), it can be understood that the exact knowledge of channel $\mathbf{h}_1$ and $\mathbf{h}_2$ are not required for data self-information removal and coherent detection. Only composite channel vectors $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ are needed. Hence, accurate channel estimation of $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ are most important.

2.3. Sparse channel model

Suppose that $\mathbf{x}_1$ and $\mathbf{x}_2$ are training sequences which are perfect separated from the data $\mathbf{d}_1$ and $\mathbf{d}_2$, respectively. Hence, training-based sparse channel model and received signal can be given by

$$\mathbf{y}_{tr} = \alpha\sqrt{2E/T}\left(\mathbf{x}_1 * \mathbf{b} + \mathbf{x}_2 * \mathbf{c}\right) + \mathbf{n} . \tag{6}$$

According to the random matrix theorem [13], circulant matrices have convolutional structure inherent to linear systems identification problems. Thus, system model (8) can be rewritten as linear matrix-vector product form

$$\mathbf{y}_{tr} = \alpha\sqrt{2E/T}\left(\mathbf{X}_1\mathbf{b} + \mathbf{X}_2\mathbf{c}\right) + \mathbf{n} = \alpha\sqrt{2E/T}\mathbf{X}\boldsymbol{\theta} + \mathbf{n} . \tag{7}$$

Define $\tilde{N} = N + 2L - 2$ where $\mathbf{X}_{i=1,2}$ are $\tilde{N} \times (2L-1)$ partial circulant channel matrix whose first row and first column can be written as $[x_i(0), \mathbf{0}_{2L-2}^T]$ and $[\mathbf{x}_i, \mathbf{0}_{2L-2}^T]^T$, respectively. Where $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2] \in \mathbb{C}^{\tilde{N}\times(4L-2)}$ and $\boldsymbol{\theta} = [\mathbf{b}^T, \mathbf{c}^T]^T \in \mathbb{C}^{4L-2}$ denote composite training sequence and composite channel vector.

3. Compressed Channel Sensing

3.1. Review of compressed sensing

Compressed sensing (CS) [14, 15] is a novel sampling theory that lies at the interdiscipline of a multiple research fields such as statistics signal processing computational harmonic analysis. In order to review the theoretical underpinnings of CS, we consider the following linear measurement model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \mathbf{n} , \tag{8}$$

where $\mathbf{X}$ denotes the training signal which is often underdetermined, $\mathbf{y}$ denotes the observation vector and $\mathbf{n}$ is noise; $\boldsymbol{\theta}$ denotes unknown sparse vector. According to the CS [14, 15], if an unknown signal vector satisfies sparse or approximates sparse, then a designed measurement matrix $\mathbf{X}$ can accurate capture most of its dominant information. One central property of $\mathbf{X}$ that has been very useful in proving the optimality of a number of CS reconstruction procedures is so-called restricted isometry property (RIP) [16] due to E. Candès.

The basic definition is given as follows:

Definition 1 (Restricted Isometry Property) Suppose that $\mathbf{X}$ be an $n \times p$ complex-valued training signal has unit ℓ_2-norm columns. $\mathbf{X}$ satisfies the RIP of order S with parameter $\delta_S \in (0,1)$ ($\mathbf{X} \in \mathrm{RIP}(S,\delta_S)$) , if

$$(1-\delta_S)\|\boldsymbol{\theta}\|_2^2 \le \|\mathbf{X}\boldsymbol{\theta}\|_2^2 \le (1+\delta_S)\|\boldsymbol{\theta}\|_2^2 \qquad (9)$$

holds for all sparse vectors $\boldsymbol{\theta}$ having no more than S nonzero coefficients.

Where $\|\boldsymbol{\theta}\|_2^2$ denotes the ℓ_2 norm given by $\|\boldsymbol{\theta}\|_2^2 = \sum_{i=1}^{p}|\theta_i|^2$. To use RIP for compressed channel sensing (CCS), we of course need to determine what kinds of matrices have small restricted isometry constant (RIC), and how many measurements are needed. Although it is quite difficult to check whether a given matrix satisfies this condition, it has been shown that many matrices satisfy the RIP with high property and few measurements.

In particular, there are two main reasons that we have chosen CoSaMP algorithm to address the AF-TWRN channel estimation with CS. First, it is one of the few reconstructions to recover the sparse signal accurately and reduce the RIP requirement of mixed-norm optimization methods [17-19]. Second, unlike these methods, it is an efficient reconstruction method with fast runtime, which is close to the complexity of LS-based channel estimation method. The following theorem states the iterative reconstruction error of the CoSaMP, which is generalized from real-value [12] to complex-value space in this paper.

3.2. Compressed Channel Sensing

CCS in the TWRN, we not only need to consider the sparse channel estimation problem, but also amplify the received signal at the relay. According to the above (7), the detail of the CCS can be implemented by CoSaMP [12] as follows:

Compressed Channel Sensing (CCS)

Input: $\mathbf{X}_1$, $\mathbf{X}_2$, S *and* $\mathbf{y}$, *that are training sequences transmitted from* T_1 *and* T_2, *the number of dominant channel taps of* $\boldsymbol{\theta}$, *observation vector* $\mathbf{y}$ *at* T_1. *The superimposed training sequence* $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$ *is combined with* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *which satisfy RIP with overwhelming probability.*

Output: *Channel estimator* $\hat{\boldsymbol{\theta}}$ *which includes* $\hat{\mathbf{b}}$ *and* $\hat{\mathbf{c}}$.

Initialize the number of nonzero taps: $\Omega^0 = \varnothing$.

Initialize the residual: $\mathbf{r}^0 = \mathbf{y}$.

While *stopping criterion* $\{i : i \ge 4S \mid i = 100\}$ *is not satisfied,* ***Do***

Channel dominant taps identification: $P_i = \mathbf{X}^* \mathbf{r}^i$, *select* $\Omega_{P_i}^i = \mathrm{supp}(P_i, 2S)$.

Update the channel dominant taps: $\Omega^i = \Omega^{i-1} \bigcup \Omega_{P_i}^i$.

Channel estimation with LS: $\hat{\boldsymbol{\theta}}^i = \mathbf{X}_{\Omega^i}^{\dagger}\mathbf{y}$.

Prune non-dominant channel coefficients: $\Omega_D^i = \mathrm{supp}(\hat{\boldsymbol{\theta}}^i, S)$ *and* $\hat{\boldsymbol{\theta}}^i_{(\Omega_D^i)^c} = 0$.

Update the estimation residual: $\mathbf{y}_r^i = \mathbf{y} - \mathbf{X}_{\Omega_D^i}(\hat{\boldsymbol{\theta}}^i)_{\Omega_D^i}$

End while

Remark: Accurate channel estimator $\hat{\boldsymbol{\theta}}$ obtained by the proposed CCS method, and then separate the estimator $\hat{\boldsymbol{\theta}}$ as $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$. $\Omega_{P_i}^i = \mathrm{supp}(P_i, 2S)$ denotes that it selects the maximum $2S$ dominant channel coefficients in P_i. Ω^c denotes the complementary set of Ω.

4. Simulation Results

In this section, the mean square error (MSE) performance of the proposed method with CoSaMP algorithm [12] is evaluated by simulations. For the purpose of comparison, the MSE performance of other existing algorithms such as LS channel estimator and oracle channel estimator (LS-based known position of dominant taps, sometime is termed as oracle lower bound) will also be evaluated as for reference. The simulation parameters are given in Table I.

TABLE I. SIMULATION PARAMETERS

channel estimation	***Linear method***	LS
	CS method	CoSaMP
Channel fading	Frequency-selective block fading	
Power delay profile	Uniform	
Channel length ($\mathbf{h}_1$, $\mathbf{h}_2$)	$L = 16$	
Channel length ($\mathbf{b}$, $\mathbf{c}$)	$(2L-1) = 31$	
Taps coefficients	Random Gaussian independent variables	
SNR (dB)	0~30	
Length of training sequence	40, 60, and 80	
Transmit/relay power	$P_1 = P_2 = P_r$	
Monte Carlo	M=1000 trails	

We assume that the channel vectors ($\mathbf{h}_1$ and $\mathbf{h}_2$) have the same length L and hence the length of convolution channel vectors ($\mathbf{b}$ and $\mathbf{c}$) is $(2L-1)$. And we also model the two-way relay channel as frequency-selective block fading so that the signal can be transmitted in a data block. To effectively verify the proposed method, we compare the average MSE performances of the channel estimation via different condition and its computational complexity. At first, the average MSE of $\hat{\mathbf{b}}$ is defined as

$$\mathrm{MSE}(\hat{\mathbf{b}}) = 1/\left(M(2L-1)\right)\sum_{m=1}^{M}\left\|\mathbf{b} - \hat{\mathbf{b}}_m\right\|_2^2 . \qquad (10)$$

It is obvious that smaller MSE means to more accurate channel estimation and vice versa.

4.1. Average MSE versus SNR

We give average MSE performance comparisons of channel estimators versus different SNR. From Fig. 2 and

Fig. 3, we find that the average MSE performance becoming better as the SNR increases. The proposed method has a small gap to oracle bound but better than the LS-based linear channel estimation. On this comparison, we also consider the MSE performance of channel estimation effected by different length (40, 60 and 80) of training sequence.

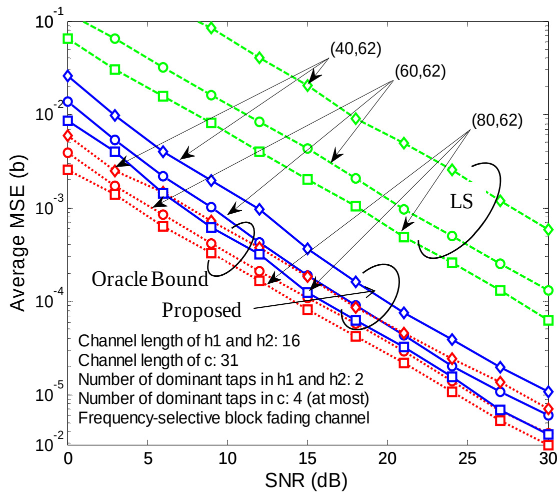

Figure 3. MSE performance of (b) versus SNR

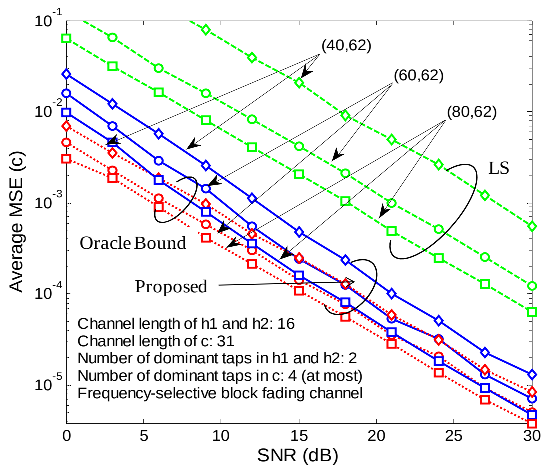

Figure 4. MSE performance of (c) versus SNR

4.2. Average MSE versus channel sparsity

We give MSE performance comparisons of channel estimators versus different channel sparsity. In this part, we suppose that the number of dominant channel taps of $\mathbf{h}_1$ and $\mathbf{h}_2$ are classified four group: (2,2), (2,4), (4,2) and (4,4). From Figure (5) and (6), we can find that LS-based average MSE performances are not changed due to the linear channel estimation method neglected channel sparsity. While the proposed method utilized channel structure as for prior information, and hence has a better MSE performance then linear method. We can find that the less number of nonzero tap in channel the better MSE performance obtained. It is worth noting that if the channel is not sparse but dense, then CS-based channel estimators have the same MSE performance to LS-based linear channel estimators.

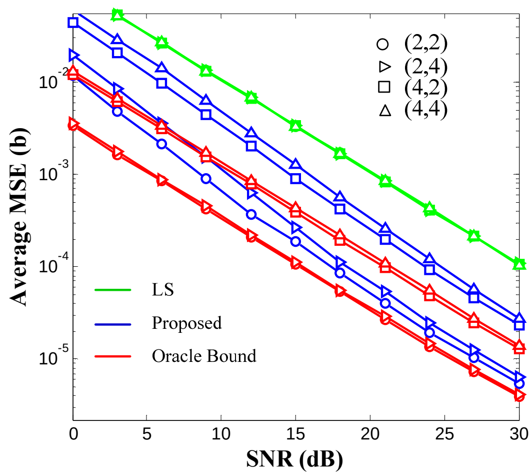

Figure 5. MSE performance of (b) via different channel sparsity.

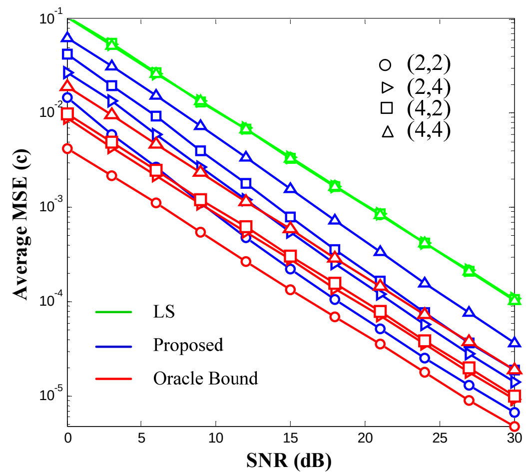

Figure 6. MSE performance of (c) via different channel sparsity.

4.3. Computational complexity

To study the computational complexity of the proposed method, we have evaluated the complexity ratio between the proposed method and LS-based channel estimation methods. We can find that the computational complexity of the proposed method and LS are very close due to the

$$Ratio = \frac{\mathrm{O}(Proposed)}{\mathrm{O}(LS)} \approx 3.5\,, \tag{11}$$

close to a constant, where the operator $\mathcal{O}(.)$ [20] is complexity measure operator . Hence, it is shown that the proposed method is also computational efficient.

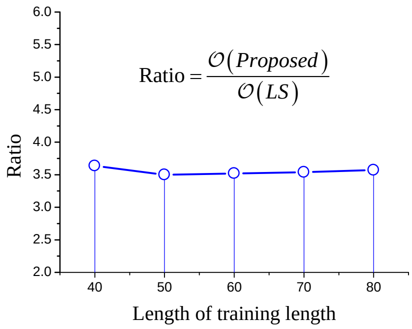

Figure 7. Complexity comparion between proposed method and LS

5. Conclusion

The channel estimation problem in TWRN is very important issue since that the channel information not only for the coherent data detection but also for the self-data removal at terminals. Hence, we proposed a CS-based sparse channel estimation method to address the problem in this paper. Both the theoretic analysis and computer simulations have confirmed performance of the proposed method. Furthermore, the computational complexity of our proposed algorithm has a low complexity that of the LS-based channel estimation. According to the above studies, we conclude that the presented sparse two-way relay channel estimation method combined the robust to noise and efficient to complexity.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," European Trans. Telecomm., pp. 585-596, Nov. 1999.

[2] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," IEEE Trans. Inform. Theory, vol. IT-25, pp. 572-584, Sept. 1979.

[3] H. Gacanin and F. Adachi, "Broadband analog network coding " IEEE Transactions on Wireless Communications, vol. 9(5), pp. 1577 - 1583, 2010

[4] F. Gao, T. Cui and A. Nallanathan, "On Channel Estimation and Optimal Training Design for Amplify and Forward Relay Networks," IEEE Transactions on Wireless Communications, vol. 7（5）, pp. 1907-1916, May 2008.

[5] F. Gao, R. Zhang and Y.-C. Liang, "Optimal Channel Estimation and Training Design for Two-Way Relay Networks," IEEE Transactions on Communications, vol. 57(10), Oct. 2009.

[6] T. Sjodin, H. Gacanin and F. Adachi, "Two-Slot Channel Estimation for Analog Network Coding Based on OFDM in a Frequency-Selective Fading Channel " VTC 2010-Spring, pp. 1 - 5, 2010.

[7] F. Gao, R. Zhang and Y.-C. Liang, "Channel Estimation for OFDM Modulated Two-Way Relay Networks," IEEE Transactions on signal processing, vol. 57(11), Nov. 2009

[8] F. Gao, T. Cui and A. Nallanathan, "Optimal Training Design for Channel Estimation in Decode-and-Forward Relay Networks With Individual and Total Power Constraints," IEEE Transactions on Signal Processing, vol. 56(12), pp. 5937-5949, Dec 2008.

[9] Z. Yan, M. Herdin, A. M. Sayeed and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," Available http://dune.ece.wisc.edu/pdfs/zhoumeas.pdf., Feb. 2007.

[10] S. F. Cotter and B. D. Rao, "Sparse channel estimation via matching pursuit with application to equalization," IEEE Transactions on Communications, vol. 50(3), pp. 374-377, MAR. 2002.

[11] W. U. Bajwa, J. Haupt, A. Sayeed and R. Nowak, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," Proceedings of IEEE, vol. 98(6), pp. 1058-1076, June 2010.

[12] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," Appl. Comput. Harmon. Anal., pp. 301-321, 2009.

[13] R. M. Gray, "Toeplitz and Circulant Matrices: A Review," Foundations and Trends in Communications and Information Theory,, vol. 3(2), pp. 155-239, 2006.

[14] E. Candès, J. Romberg and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Transaction on Information Theory, vol. 52(2), pp. 489-509, Feb. 2006.

[15] D. L. Donoho, "Compressed sensing," IEEE Transaction on Information Theory, vol. 52(4), pp. 1289-1306, Apr. 2006

[16] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," Comptes Rendus Mathematique, vol. 346(9-10), pp. 589-592, May 2008.

[17] S. S. Chen, D. L. Donoho and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM J. Sci. Comput., vol. 20(1), pp. 33-61, 1999.

[18] T. Ribshirani, "Regression shrinkage and selection via the lasso," J. R. Stat. Soc. B, vol. 58(1), pp. 267-288, 1996.

[19] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," Annals of Statistics, vol. 35, pp. 2392-2404, 2006.

[20] D. Needell, "Topics in Compressed Sensing," PhD Dissertation, Mathematics, Univ. of California, Davis, May 2009.