Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **2**, March **2018**

Channel Estimation for Multi-user Massive MIMO Systems based on Compressive Sensing with Truncated-HOSVD

Qing-Zhu Wang[1], Meng-Ying Wei[1,*] and Yi-Hai Zhu[2]

[1]School of Information Engineering, Northeast Electric Power University, 169 Changchun Road, Jilin, China
[2]Engineering Technology Center CRRC Changchun Railway Vehicles Co.,Ltd
*Corresponding author: 383532294@qq.com

Received March, 2017; revised December, 2017

Abstract. *Compressive sensing (CS) theory is widely employed in channel estimation of massive multiple-input multiple-output (MIMO) systems to reduce training pilot and feedback overhead. Compressive sensing based on Tucker model (Tucker-CS) is an emerging approach for high-order data representation. In this paper, we propose an improved estimation algorithm based on Tucker-CS for noisy channels. We first introduce truncated higher-order singular value decomposition (T-HOSVD) and hard threshold selection strategy. The optimal hard threshold of singular value is then used to reconstruct channel state information (CSI) from noise. Finally, we apply this algorithm to downlink CSI estimation for frequency-division duplexing (FDD) multi-user massive MIMO systems. Simulation results demonstrate that this proposed system outperforms state of the art.*
Keywords: Multi-user massive MIMO, Channel estimation, Tucker model, Compressive sensing, Optimal threshold

1. **Introduction.** A higher demand for communications has followed the rise of commercial 4G. Massive MIMO is the new wireless network technology proposed in this background. Compared with traditional MIMO, massive MIMO has higher spectrum utilization composed of hundreds of antennas on the base station (BS) side [1-2]. It provides users with high transmission rate and satisfies the high traffic density demand of the network. Precise knowledge of reliable CSI at the transmitter is necessary to fully utilize the spatial multiplexing gains and the array gains of massive MIMO [3]. However, due to the large number of antennas at BS side, the traditional pilot-based channel estimation algorithm used in multi-user massive MIMO systems leads to excessive pilot consumption and takes up too much bandwidth. Thus, acquiring CSI with a small amount of training pilot is currently a huge challenge for massive MIMO technology.

In time-division duplexing (TDD) massive MIMO systems, CSIT can be obtained by exploiting the channel reciprocity using uplink pilots [4-5]. Most cellular systems today employ FDD because it is considered more effective for systems with traffic and delay-sensitive applications [6]. It is thus important to explore effective channel estimation methods in massive MIMO systems with FDD. However, the coefficients of unknown downlink CSI are very high due to the large number of transmitting antenna in BS [7]. Direct channel estimation leads to high computational complexity, and causes excessive pilot and feedback overhead. Least square (LS) and minimum mean square error (MMSE) are two kinds of conventional CSI estimation algorithms [8]. These algorithms, however, are not suitable for massive MIMO systems, because the number of training pilots required increases as the number of transmitting antenna increases, and the noise in channel is completely ignored. Thus, an applicable CSI estimation approach is needed for FDD multi-user massive MIMO systems.

Compressive sensing technique has been known to recover sparse signals from a small number of linear measurements [9]. The application of CS to wireless communication and networks has been extensively studied, including wireless channel estimation, wireless sensor network, and network tomography. In massive MIMO system, the channel matrices tend to be sparse as the transmitting antenna increases at the BS side [10]. A number of CSI estimation algorithms based on CS have been proposed to improve estimation precision and reduce pilot and feedback overhead. For example, Nguyen et al. [11] built a low-rank matrix approximation based on CS and solved it via a quadratic semidefine programming (SDP). However, this algorithm applies only to TDD massive MIMO systems. Some channel estimation algorithms for FDD massive MIMO systems have been proposed in recent years. A modified subspace pursuit (SP) algorithm was proposed to solve conventional CS based CSI estimation problems by exploiting the prior support adaptively based on quality information in reference [12]. Rao et al. [13] proposed a new CSI feedback and estimation scheme for FDD massive MIMO system made up of three parts: 1) BS broadcasts training pilot to all users; 2) Each user obtains the compressive measurement and feeds back to BS; 3) BS jointly recovers the CSI based on measurements. In addition, joint orthogonal matching pursuit (Joint-OMP) is presented for CSI reconstruction. In a recent study [14], a weighted block L1- minimization based estimation approach is proposed for the same sparsity structure and CSI feedback and estimation protocol as [13]. Whether based on greedy algorithm like OMP, SP or convex optimization algorithm in CS, the CSI estimation algorithms for large and high-order channel coefficient matrices in multi-user massive MIMO cause excessive pilot and feedback consumption due to the fact that each individual user must be considered. They are further inadequate because they do not consider channel noise. Moreover, the two CS algorithms also result in high computational complexity as the iterations progress.

Most of the development of channel estimation based on CS is focused on the 2D channel coefficient matrix. However, the multi-user massive MIMO system channel coefficient has higher dimensional tensors resulting from the large-scale antennas and users. Channel estimation based on traditional CS calculates each user's channel state information respectively, resulting in pilot waste and high computational complexity. Tensor decomposition makes it possible to extend CS to a higher dimensional. Recently, Cesar et al. [15] provided a Tucker-model based CS algorithm that multiplies the data tensor by a different sensing matrix in each mode, then recovers the original tensor from multilinear projections. Compared to existing sparsity-based CS methods [16], this Tucker-CS algorithm does not require assuming sparsity or a dictionary based representation. Furthermore, due to involve no iterations, it is fast and that makes it suitable for high-order data problems.

Current channel estimation methods based on CS require that each individual user be accounted for resulting in training pilot waste and excessive feedback overhead. In addition to, these methods do not consider the effects of channel noise. This paper thus proposes a novel CSI estimation approach that regards the channel coefficient in FDD multi-user massive MIMO systems as a 3D tensor, then truncated higher-order singular value decomposition (T-HOSVD) and hard threshold selection strategy are introduced. Finally, the optimal hard threshold of singular value is used for the reconstruction of CSI from the noise.

The remainder of this paper is organized as follows: Section II introduces the related notation, system model and basic theory used throughout the paper. In Section III, the truncated higher-order singular value decomposition (T-HOSVD) and selection strategy of hard threshold are proposed, then the application of this truncated Tucker based compressive downlink CSI recovery is presented. Section IV provides several numerical results which verify our theoretical results and our evaluation of the proposed algorithms performance. In Section V, the main conclusions are outlined.

2. Basic theory and system model.

2.1. **Compressive Sensing.** Traditional CS is an approach for reconstruction of sparse signals or signals with sparse representation in some domain [17]. A signal $x \in R^M$ is called k-sparse if it only has k nonzero entries. The signal is then measured not via standard point samples but rather through the projection by a measurement matrix $\Phi \in R^{N\times M}$ where $N < M$ and the measurement value can be written as:

$$y = \Phi x \tag{1}$$

The final goal of CS is to recover the signal x from the fewest possible measurements y.

2.2. **Tensor notations.** A tensor is a multi-dimensional matrix, e.g. $\mathcal{X} \in R^{I_1\times I_2\times\cdots\times I_N}$ is an N-th order tensor. A vector and a matrix can be regarded as a one-order tensor and a second-order tensor respectively, e.g.$x \in R^I$ and $X \in R^{I_1\times I_2}$. The order of a tensor is the number of modes[18]. For instance, tensor $\mathcal{X} \in R^{I_1\times I_2\times\cdots\times I_N}$ has order N and the dimension of its n-th mode is I_n. The element of a tensor is referred as $x_{i_1,i_2,\cdots,i_n}$. Tensor is high-order complex problem and cannot be calculated by general methods. It is necessary to unfolding the tensor to the matrix to apply the higher-order singular value decomposition technique to the tensor and simplify the product of tensor and measurement matrix. Namely, the tensor should be rearranged to matrices according to the different modes.

Definition 1 (n-mode unfolding of tensor): n-mode unfolding of tensor is a process that the elements in n-mode of the tensor are arranged in a matrix of column vectors to obtain a new matrix. Given a tensor $\mathcal{X} \in R^{I_1\times I_2\times\cdots\times I_N}$, its mode-n fibers are the vectors obtained by fixing all indices except which correspond to columns (n=1), rows (n=2) and so on [15]. Tensor elements $(i_1, i_2, \cdots , i_n)$ maps to matrix elements (i_n, j), with $j = 1 + \sum_{k\neq n}(i_k - 1)J_k$ where $J_k = \prod_{m\neq n}^{k-1} I_m$.

Definition 2 (n-mode product of tensor and matrix): tensor cannot be multiplied by matrix directly. Thus, the product of tensor and matrix is the n-mode unfolding of the tensor multiplied by the matrix in same dimension. Assigning $\mathcal{X} \in R^{I_1\times I_2\times\cdots\times I_N}$ and $U \in R^{R\times I_n}$ as the tensor and the matrix respectively, the n-mode product of them is $\mathcal{Y} = \mathcal{X} \times_n U \in R^{I_1\times\cdots\times I_{n-1}\times R\times I_{n+1}\times\cdots\times I_N}$ defined by:

$$y_{i_1 i_2\cdots i_{n-1} r i_{n+1}\cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2\cdots i_n\cdots i_N} u_{r i_n} \tag{2}$$

2.3. **System model.** The existing channel estimation algorithms [13-14], consider a multi-user massive MIMO system with FDD consisting of M transmitting antennas in the BS side and K users, each including N receiving antennas. The BS broadcasts a sequence of T training pilots through M antennas to estimate the downlink channels. At each user side, the received measurement can be expressed as:

$$\mathrm{Y_i} = \mathrm{H_i X} + \mathrm{N_i}, i = 1, \cdots , K \tag{3}$$

Where $\mathrm{H_i} \in R^{N\times M}$ is downlink channel matrix from BS to each user side, $\mathrm{X} \in R^{M\times T}$ is training pilot and $\mathrm{N_i} \in R^{N\times T}$ is Gaussian random noise matrix with zero mean and variance σ_n^2.

In this paper, we propose a novel 3D channel model $\mathcal{H} \in R^{N\times M\times K}$ for multi-user massive MIMO system, in which modes are the number of receiving antennas N, transmitting antennas M, and users K. We have considered all channels as a whole tensor to measure and estimate. First, as discussed in [13-14], the BS broadcasts a sequence of T training pilot to the users. In our proposed system model, this process can be regarded as the mode-2 unfolding of the tensor CSI $\mathcal{H} \in R^{N\times M\times K}$ is measured by the training pilot from BS directly and the measurement is formulated as:

$$\mathrm{Y} = \mathcal{H} \times_2 \mathrm{X} + \mathrm{N} \tag{4}$$

Where $\mathrm{N} \in R^{NK\times T}$ is a Gaussian random noise matrix. Each user then feeds back the observed values which comprise the final result $\mathrm{Y} \in R^{NK\times T}$ to the BS side. Finally, other two Gaussian sensing matrices are utilized to measure the two remaining modes of the measured tensor. With this, all measurements for 3D channel model of multi-user massive MIMO system are completed. Considering this measurement process is done in the BS, no additional bandwidth or cost results from the training pilot or feedback.

3. **Proposed estimation algorithm.** In this section, an improved CSI reconstruction algorithm based on Tucker-CS is introduced. It consists of two parts: 1) Truncation in higher-order singular value decomposition; 2) Selection of optimal hard threshold.

3.1. **Truncated higher-order singular value decomposition.** Higher-order singular value decomposition (HOSVD) is a generalization of the matrix singular value decomposition (SVD) to higher order matrices. It is an excellent method to approximately decompose the tensor into a core tensor product with multiple matrices. In this paper, we only consider three-order tensor with modes M, N, and K, as described in Section II. For 3D CSI $\mathcal{H} \in R^{N\times M\times K}$ with noise, we utilize the Tucker-CS [17] to decompose and reconstruct $\mathcal{H}$:

$$\mathcal{Y} = \mathcal{H} \times_1 \Phi_1 \times_2 \Phi_2 \times_3 \Phi_3 \tag{5}$$

$$\widetilde{\mathcal{H}} = \mathcal{Y} \times_1 \mathrm{Z_1 Y_1^\dagger} \times_2 \mathrm{Z_2 Y_2^\dagger} \times_3 \mathrm{Z_3 Y_3^\dagger} \tag{6}$$

Where Φ_n is measurement matrix, and $Y_{(n)}^\dagger$ is truncated MP pseudo-inverse matrix of $Y_{(n)}$ which is mode-n unfolding matrix of core tensor. The other parameters mentioned in (6) are as follows:

$$\mathcal{Z}^{(n)} = \begin{cases} \mathcal{H} \times_2 \Phi_2 \times_3 \Phi_3, & for\ n{=}1 \\ \mathcal{H} \times_1 \Phi_1 \times_3 \Phi_3, & for\ n{=}2 \\ \mathcal{H} \times_1 \Phi_1 \times_2 \Phi_2, & for\ n{=}3 \end{cases} \tag{7}$$

$$Z_n = (\mathcal{Z}^{(n)})_{(n)} \tag{8}$$

The most important part of truncated higher-order singular value decomposition in Tucker-CS are to eliminate noise and keep channel coefficients. After the SVD, $Y_{(n)}$ and $Y_{(n)}^\dagger$ can be described as:

$$Y_{(n)} = U_n S_n V_n^T \tag{9}$$

$$Y_{(n)}^\dagger = V_n \widetilde{S}_i^n U_n^T \tag{10}$$

With U_n and V_n are unitary matrices, S_n is the diagonal matrix composed by singular values s_i^n and $\widetilde{S}_i^n$ in (10) is defined as:

$$\widetilde{S}_i^n = \begin{cases} \frac{1}{s_i^n}, & for \quad s_i^n > \tau_n \\ 0, & for \quad s_i^n \le \tau_n \end{cases} \tag{11}$$

Where τ_n is the hard threshold. How to choose the optimal hard threshold of singular values is discussed below.

3.2. **Optimal hard threshold.** In recent years, a few different approaches have been proposed regarding optimal hard threshold selection in SVD [19-20]. Gavish et al [20] presented a method to choose the optimal hard threshold for matrix A $\in R^{N\times M}$ in unknown noise level with coefficient $\beta = \frac{N}{M}$. We extend this method into 3D tensor as follows:

$$\tau_n = \lambda_n(\beta_n)\sqrt{n}\sigma \tag{12}$$

$$\widetilde{\sigma}(Y_n) = \frac{y^n_{med}}{\sqrt{n \times \mu_n}} \tag{13}$$

Where y^n_{med} is the median singular value of $Y_{(n)}$ and μ_n is the median of the Marcenko-Pastur distribution, namely, the unique solution in $\beta_{n,-} \leq x \leq \beta_{n,+}$ to the equation:

$$\int_{\beta_{n,-}}^{x} \frac{\sqrt{(\beta_{n,+} - t)(t - \beta_{n,-})}}{2\pi\beta_n t}\, dt = \frac{1}{2} \tag{14}$$

With $\beta_{n,\pm} = (1 \pm \sqrt{\beta_n})^2$. As discussed in the system model from Section II and the first part of Section III, the original 3D CSI $\mathcal{H} \in R^{N\times M\times K}$ is measured and the core tensor is $\mathcal{Y} \in R^{N\times T\times K}$. Thus, the dimension of $Y_{(n)}$ is

$$Y_{(n)} \in \begin{cases} R^{N\times TK}, & for\ n{=}1 \\ R^{T\times NK}, & for\ n{=}2 \\ R^{K\times NT}, & for\ n{=}3 \end{cases} \tag{15}$$

The corresponding parameter β_n is as follows:

$$\beta_n = \begin{cases} \frac{N}{TK}, & for\ n{=}1 \\ \frac{T}{NK}, & for\ n{=}2 \\ \frac{K}{NT}, & for\ n{=}3 \end{cases} \tag{16}$$

By using $\widetilde{\sigma}(Y_n)$ instead of σ in (12):

$$\tau_n = \lambda_n(\beta_n)\sqrt{n}\sigma = \lambda_n(\beta_n)\sqrt{n}\widetilde{\sigma}(Y_n) = \frac{\lambda_n(\beta_n) y^n_{med}}{\sqrt{\mu_n}} \tag{17}$$

Defining $\Omega(\beta_n) = \frac{\lambda_n(\beta_n)}{\sqrt{\mu_n}}$, (17) can be expressed as:

$$\tau_n = \Omega(\beta_n) y^n_{med} \tag{18}$$

According to optimal threshold coefficient for matrices [20], by making available a Matlab script, the optimal threshold coefficient $\Omega(\beta_n)$ for 3D tensor can be evaluated approximately as follows:

$$\Omega(\beta_n) \approx 0.56\beta_n^3 - 0.95\beta_n^2 + 1.82\beta_n + 1.43 \tag{19}$$

The pseudo-code used for the proposed algorithm is presented below.

4. **Proposed estimation algorithm.** Our numerical simulations were conducted in Matlab2010b on a work station with a 2.1GHz Intel Celeron CPU and 6 GB RAM. We compare the performance of the proposed algorithm to the Joint-OMP algorithm [13], Modified L1-minimization algorithm [14] and the Tucker-CS algorithm for multi-user massive MIMO system with FDD mode.

Fig. 1 shows the normalized mean square error (NMSE) of four algorithms in different SNRs and the parameters are set as [14], the number of BS antennas $M = 160$, the number of user antennas $N = 2$, the number of users $K = 40$, and the number of training pilot symbols $T = 45$. The proposed algorithm and Tucker-CS algorithm have higher

Algorithm 1
Input:
Core tensor $\mathcal{Y} \in R^{N\times T\times K}$
Output:
Optimal threshold of singular values τ_n, $n = 1, 2, 3$
Start:
for $n = 1 : 3$, do:
Unfold core tensor $\mathcal{Y}$ to $Y_{(n)}$ in n-mode;
Compute y^n_{med} according to (9);
Compute β_n as (16);
Compute $\Omega(\beta_n)$ as (19);
Compute τ_n as (18);
end for
End

Algorithm 2
Input:
(1) Tensor CSI $\mathcal{H} \in R^{N\times M\times K}$
(2) Training pilot $\mathrm{X} \in R^{M\times T}$
(3) Sensing matrices $\Phi_1 \in R^{N\times N}$, $\Phi_3 \in R^{K\times K}$
Output:
Reconstruction of CSI $\widetilde{\mathcal{H}}$
Start:
(1) Compute core tensor $\mathcal{Y}$ according to (5);
(2) for $n = 1 : 3$, do:
Compute Z_n as (7)-(8);
Compute the optimal threshold τ_n corresponding to Z_n by Algorithm 1;
Compute the truncated MP pseudo-inverse matrix $Y^{\dagger}_{(n)}$ according to τ_n by (9)-(11);
end for
(3) Reconstruct $\widetilde{\mathcal{H}}$ according to (6);
End

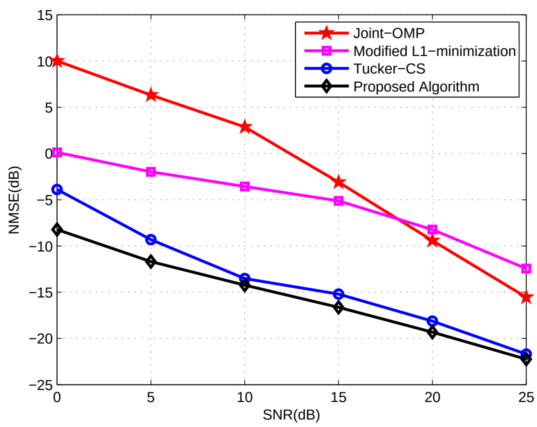

FIGURE 1. NMSE of the proposed algorithm, Joint-OMP, Modified L1-minimization and Tucker-CS for different SNRs

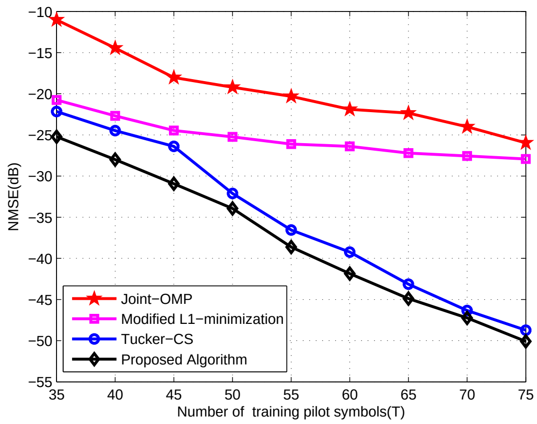

FIGURE 2. NMSE of the proposed algorithm, Joint-OMP, Modified L1-minimization and Tucker-CS with different number of training pilot symbols (T)

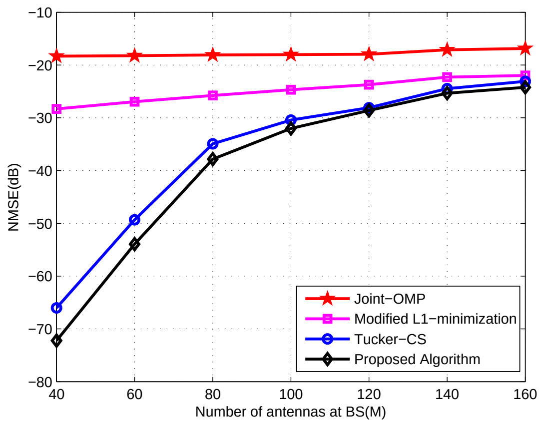

FIGURE 3. NMSE of the proposed algorithm, Joint-OMP, Modified L1-minimization and Tucker-CS with different number of antennas at BS (M)

estimation precision than the other two algorithms. Additionally, the proposed algorithm outperforms Tucker-CS because it considers channel noise and conducts a de-noising process to reduce the influence of noise.

As shown in Fig. 2, we compared the NMSE of estimated CSI with the number of training pilot symbols T with the parameter settings $M = 100$, $N = 2$, $K = 40$, and transmit SNR=30dB. It is obviously the proposed algorithm outperforms other algorithms and we found that the estimation accuracy of CSI increases as T increases.

In addition, we compared the NMSE of estimated CSI with the number of antennas at the BS under the system as $N = 2$, $K = 40$, $T = 45$, and SNR=30dB. The proposed algorithm achieved better performance than the other algorithms, as shown in Fig. 3. However, we found that CSI estimation accuracy decreases as M increases. Because the dimensions of the channel coefficient matrices grow larger as M increases, more measurements are required. Which of the four algorithms is utilized, the estimation accuracy decreases.

In Table I, we compare average computation time with different numbers of antennas at BS. The multi-user massive MIMO system parameter settings are $N = 2$, $K = 40$, $T =$

45, and SNR=30dB. The proposed algorithm exhibits similar computation time to Tucker-CS, which is much shorter than the Modified L1-minimization algorithm but a little longer than the Joint-OMP. In general, the proposed algorithm has lower computational complexity by involving no iteration.

Table 1. Comparison of computation time (s)

Massive MIMO System	M	Methods	Computation Time(s)
SNR=30dB $N=2$ $K=40$ $T=45$	50	Joint-OMP	0.05
		Modified L1-minimization	257.66
		Tucker-CS	0.36
		Proposed algorithm	0.39
	100	Joint-OMP	0.18
		Modified L1-minimization	425.05
		Tucker-CS	1.09
		Proposed algorithm	1.13
	150	Joint-OMP	0.23
		Modified L1-minimization	762.38
		Tucker-CS	1.66
		Proposed algorithm	1.68

5. **Conclusions.** The proposed algorithm is a novel channel estimation algorithm for multi-user massive MIMO systems with FDD mode. It extends optimal hard threshold selection in SVD to a higher-order and uses truncated HOSVD in Tucker-CS to eliminate noise. The proposed algorithm was applied and compared with current channel estimation methods. Experimental results demonstrate that the proposed algorithm outperforms other estimation methods for noisy channel in FDD multi-user massive MIMO systems.

REFERENCES

[1] E. G. Larsson, F. Tufvesson, O. Edfors and T. L. Marzetta, Massive MIMO for next generation wireless systems, *IEEE Commun. Mag.*, vol.52, no.2, pp.186–195, 2014.

[2] F. Rusek, D. Persson, B. K. Lau, and E. G. Larsson, Scaling up MIMO: Opportunities and challenges with very large arrays, *IEEE Signal Process. Mag.*, vol.30, no.1, pp.40–60, 2013.

[3] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, A coordinated approach to channel estimation in large-scale multiple-antenna systems, *IEEE J. Sel. Areas Commun.*, vol.31, no.2, pp.264–273, 2013.

[4] P. W. C. Chan, E. S. Lo, and R. R. Wang, The evolution path of 4G networks: FDD or TDD, *IEEE Commun. Mag.*, vol.44, no.12, pp.42–50, 2007.

[5] T. L. Marzetta, G. Caire, M. Debbah, I. Chih-Lin, and S. K. Mohammed, Special issue on massive MIMO, *Journal of Commun. and Networks*, vol.15, no.4, pp.333–337, 2013.

[6] J. Choi, D. J. Love, and P. Bidigare, Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory, *IEEE J. Sel. Topics Signal Process.*, vol.8, no.5, pp.802–814, 2014.

[7] M. Biguesh and A. B. Gershman, Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals, *IEEE Trans. Signal Process.*, vol.54, no.3, pp.884–893, 2006.

[8] L. Kewen and X. Ke, Research of MMSE and LS channel estimation in OFDM systems, *IEEE International Conference on Information Science and Engineering*, pp.2308–2311, 2010.

[9] D. L. Donoho, Compressed Sensing, *IEEE Trans. Inf. Theory*, vol.52, no.4, pp.1289–1306, 2006.

[10] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, Channel measurements for large antenna arrays, *IEEE Int. Symp. Wireless Commun. Systems*, pp.811–815, 2012.

[11] S. L. H. Nguyen and A. Ghrayeb, Compressive sensing based channel estimation for massive multiuser MIMO systems, *IEEE Wireless Commun. Networking Conf.*, pp.2890–2895, 2013.

[12] X. Rao and V. K. N. Lau, Compressive sensing with prior support quality information and application to massive MIMO channel estimation with temporal correlation, *IEEE Trans. Signal Process.*, vol.63, no.18, pp.3261–3271, 2015.

[13] X. Rao and V. K. N. Lau, Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems, *IEEE Trans. Signal Process.*, vol.62, no.12, pp.3261–3271, 2014.

[14] C. C. Tseng, J. Y. Wu, and T. S. Lee, Enhanced compressive downlink CSI recovery for FDD massive MIMO systems using weighted block L1-Minimization, *IEEE Trans. Commun.*, vol.64, no.3, pp.1055–1067, 2016.

[15] F. C. Cesar, and C. Andrzej, Stable, robust, and superfast reconstruction of tensors using multi-way projections, *IEEE Trans. Signal Process.*, vol.63, no.3, pp.780–793, 2015.

[16] S. Friedland, Q. Li, and D. Schofeld, Compressive Sensing of Sparse Tensors, *IEEE Trans. Image Process.*, vol.23, no.10, pp.4438–4446, 2014.

[17] R. G. Baraniuk, Compressive sensing, *IEEE Signal Process. Mag.*, vol.24, no.4, pp.118–124, 2007.

[18] N. Vannieuwenhoven, A new truncation strategy for the higher-order singular value decomposition, *Siam Journal on Scientific Computing*, vol.34, no.2, pp.1027–1052, 2012.

[19] R. R. Nadakuditi, OptShrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage, *IEEE Trans. Inf. Theory*, vol.60, no.5, pp.3002–3018, 2014.

[20] G. Matan, and L. D. David, The optimal hard threshold for singular values is $4/\sqrt{3}$, *IEEE Trans. Inf. Theory*, vol.60, no.8, pp.5040–5053, 2014.