IJICTR
International Journal of Information & Communication Technology Research

Volume 13- Number 1 – 2021 (25 -31)

Massive MIMO Slow-varying Channel Estimation Using Tensor Sparsity

Nasser Sadeghi
Laboratory of Wireless Communications and Signal Processing (WCSP), Faculty of Electrical Engineering, Sahand University of Technology
Tabriz, Iran
n_sadeghi99@sut.ac.ir

Masoumeh Azghani*
Laboratory of Wireless Communications and Signal Processing (WCSP), Faculty of Electrical Engineering, Sahand University of Technology
Tabriz, Iran
mazghani@sut.ac.ir

Received: 21 January 2021 - Accepted: 28 March 2021

Abstract— **In order to exploit the advantages of the massive MIMO systems, it is vital to apply the channel estimation task. The huge number of antennas at the base station of a massive MIMO system produces a large set of channel paths which requires to be estimated. Therefore, the channel estimation in such systems is more troublesome. In this paper, we propose to leverage the temporal joint sparsity of the massive MIMO channels to offer a more accurate channel estimation. To attain this goal, we would model the problem to exploit the spatial correlation among different antennas of the BS as well as the inter-user similarity of the channel supports. In addition, by assuming a slow time-varying channel, the supports of the channel matrices of various snapshots would be equal which enables us to impose the temporal joint sparsity on the channel submatrices. The simulation results validate the efficiency and superiority of the suggested scheme over its rivals.**

Keywords- Massive MIMO; Channel estimation; Sparisity; Joint sparsity.

I. INTRODUCTION

Massive MIMO systems have attracted a great deal of attention during recent years due to some outstanding features such as higher capacity and better communication gains. Therefore, they are recognized as a key technology for the next generation of communication system [1]. To leverage the higher degrees of freedom provided by the massive MIMO systems, it is essential to have the channel state information in the transmitter side (CSIT) [2-4]. In massive MIMO systems, due to the large number of antennas in the BS, the traditional channel estimation schemes would result in large pilot overhead. Thus, we need to look deeper into the challenge of massive MIMO channel estimation. There are some features in the massive MIMO systems which can be used to estimate the channel.

A number of recent works have investigated the massive MIMO channel estimation in TDD protocol [5] [6]. In TDD systems, the estimated uplink channel can be used to estimate the downlink channel owing to the reciprocity property. In FDD systems, however, there is no reciprocity and the downlink channel shall be estimated independently. Hence, massive MIMO channel estimation in the FDD mode is more challenging [4].

The conventional channel estimation methods such as Least square (LS) [7] and minimum mean square

* Corresponding Author

error (MMSE) [8] are not suitable for the massive MIMO systems due to the huge number of pilots required, in other words they will increase the overhead.

To deal with this issue, some special properties of the massive MIMO channels can be exploited. The massive MIMO channels can be modeled as sparse or structured sparse [9-12]. A signal is called sparse if most of its coefficients are zero in a domain [13]. Sparsity has found various applications in different fields [14-17]. In [18], the spatial correlation between various antennas of the BS and a user has been exploited to model the equivalent channel impulse response as a block sparse signal. Then, an iterative support detection scheme has been suggested to estimate the channels. In [19], the spatial temporal common sparsity of delay domain massive MIMO channels have been exploited to estimate the channel coefficient using a structured subspace pursuit algorithm. In [20], the individual and distributed joint sparsity have been leveraged to reduce the required pilot signals. In [21], joint sparsity of the users has been exploited to estimate the massive MIMO channel in a specific time. In [3], joint sparsity of the users in a non-ideal feedback model has been investigated.

In this paper, we have modeled the massive MIMO slow-varying channel estimation as a structured sparse signal recovery problem. The structured sparsity occurs due to the correlation of the channel paths between different BS antennas and users as well as the common support of the channel matrix over several time slots. The leveraged sparsity pattern which we call tensor sparsity reduces the pilot overhead to a desirable level and ameliorates the channel estimation accuracy at the same time.

In this paper, we propose to exploit the structured sparsity of the slow-varying massive MIMO channels. The joint sparsity of the various BS antennas and users as well as the common-support of the channel matrices in several consecutive time slots leads to a structured sparsity pattern which we call tensor sparse. The joint sparsity of the channels of various users and BS antennas exists due to the limitted scattereres at the BS side, especially when the users are close to each other [20]. Furthermore, the common support of channels of consecutive time slots is an outcome of the slow-variation of the channel over time. We have suggested a tensor sparse recovery algorithm called Tensor Orthogonal Matching Pursuit (TOMP) to reconstruct the mentioned structured sparse signal. The suggested algorithm is a greedy scheme based on support estimation, support pruning, and residual update. The support is estimated using the correlation of the measurements and the sensing matrix. The pruning step is done based on thresholding. The channel elements are estimated using the least squares step.

The proposed method has two advantages: The number of required pilots to achieve a specific channel estimation error is decreased compared to the benchmarks. Also, the channel estimation task has been speeded up since a joint estimation is applied for several snapshots.

The BS sends pilots to different users. Then, all of the users feedback their received signal to the BS to perform the channel estimation. In practice, the sparsity level of the channel depends on local scatterers in the environment. The channel sparsity level (the number of non-zero coefficients) is usually available as the preliminary information for the channel estimation [22, 23].

Notations: Uppercase and lowercase boldface denote the matrices and vectors, respectively. $(.)^H$ and $(.)^\dagger$ are conjugate transpose and Moore-Penrose pseudoinverse. $\mathbf{A}_\Omega$ denotes the submatrix composed of the columns of $\mathbf{A}$ whose indices belong to the set Ω. $\|\mathbf{A}\|_F$ indicates the Frobenius norm.

II. SYSTEM MODEL

Suppose a massive MIMO system with N_t antennas at the Base Station (BS) and K single antenna users $(N_t >> K)$ that use OFDM modulation at the BS. The received signal at the k^{th} user and t^{th} time slot can be considered in the frequency domain as:

$$\mathbf{y}_k(t) = \sum_{i=1}^{N_t} \mathbf{X}_i \mathbf{F}_L \bar{\mathbf{h}}_{k,i}(t) + \mathbf{n}_k(t) \qquad (1)$$

where $\mathbf{X}_i = diag\{\mathbf{x}_i\}$ and $\mathbf{x}_i \in \mathbb{C}^{N\times 1}$ denotes the transmitted pilots from the i^{th} antenna and N is the OFDM symbol length. $\mathbf{F}_L \in \mathbb{C}^{N\times L}$ is a sub-matrix consisting of the first L columns of the normalized discrete fourier transform (DFT) matrix of size $N \times N$. $\bar{\mathbf{h}}_{k,i}(t) = [\bar{h}^1_{k,i}(t), \bar{h}^2_{k,i}(t), \cdots, \bar{h}^L_{k,i}(t)]^T \in \mathbb{C}^{L\times 1}$ is the L -tap channel vector between the i^{th} antenna of the BS and the k^{th} user at the t^{th} snapshot. $\mathbf{n}_k(t) \in \mathbb{C}^{N\times 1}$ denotes the additive complex Gaussian noise vector that consists of independent and identically distributed (i.i.d) entries with zero mean and unit variance. Rewriting (1) in a matrix form, we have:

$$\mathbf{y}_k(t) = \bar{\mathbf{P}}\bar{\mathbf{h}}_k(t) + \mathbf{n}_k(t) \quad (2)$$

where $\bar{\mathbf{P}} = [\mathbf{X}_{(1)}\mathbf{F}_L, \mathbf{X}_{(2)}\mathbf{F}_L, \ldots, \mathbf{X}_{(\mathbf{N}_t)}\mathbf{F}_L]$ and $\bar{\mathbf{h}}_k(t) = [(\bar{\mathbf{h}}_{k,1}(t)^T, (\bar{\mathbf{h}}_{k,2}(t))^T, \ldots, (\bar{\mathbf{h}}_{k,N_t}(t))^T]^T \in \mathbb{C}^{N_t L\times 1}$

Due to the fact that there are a limited number of significant scatterers in the out-door environments, the delay-domain channel vectors, $\bar{\mathbf{h}}_{k,i}(t)$, would be sparse [19]. In other words, only a small number of the L taps of the channel vector would be non-zero which makes the channel sparse.

Furthermore, since the inter-antenna space is much lower compared to the distance between user and BS, the channels between different antennas of the BS and a user would have similar sparsity pattern which is called the spatial correlation of different channels from the BS to a user [18]. Therefore, the channels between different antennas of the BS and a user would have the same support which can be expressed as:

$$\boldsymbol{\Omega}_{\bar{\mathbf{h}}_{k,1}(t)} = \boldsymbol{\Omega}_{\bar{\mathbf{h}}_{k,2}(t)} = \ldots = \boldsymbol{\Omega}_{\bar{\mathbf{h}}_{k,N_t}(t)} \quad (3)$$

Where $\boldsymbol{\Omega}_{\bar{\mathbf{h}}_{k,i}(t)}$ indicates for the $support(\bar{\mathbf{h}}_{k,i}(t))$, i.e the indices of the non-zero channel coefficients. Moreover, in multi-user massive MIMO systems, a number of scatterers would be the same for all the users which causes a common block support between the channels of different users [20]. In addition, since the users are located in different places, there will be some different scatterers in the user-BS paths which causes the individual block support in the channel vector of various users. In summary, the supports of different users would have a common part as well as an individual part.

In order to exploit the common sparsity pattern of the channel vectors, we reshape the vector $\bar{\mathbf{h}}_k(t)$ and the columns of the matrix $\bar{\mathbf{P}}$ so that the similar channel taps of different $\bar{\mathbf{h}}_k(t)$ vectors appear in a block. Thus, $\bar{\mathbf{h}}_k(t)$ can be written as:

$$\begin{aligned} &\mathbf{h}_k(t) = reshape(\bar{\mathbf{h}}_k(t)) = \\ &[\bar{h}_{k,1}^1(t), \bar{h}_{k,2}^1(t) \ldots, \bar{h}_{k,N_t}^1(t), \bar{h}_{k,1}^2(t) \\ &, \ldots, \bar{h}_{k,N_t}^2(t), \ldots, \bar{h}_{k,1}^L(t), \ldots, \bar{h}_{k,N_t}^L(t)]^T \end{aligned} \quad (4)$$

The schematic diagram of our reshaping procedure has been depicted in Figure (1).

After reshaping the channel vector of all users, we concatenate them in a column-wise manner as:

$$\mathbf{H}_t = [\mathbf{h}_1(t), \mathbf{h}_2(t), \ldots, \mathbf{h}_K(t)] \quad (5)$$

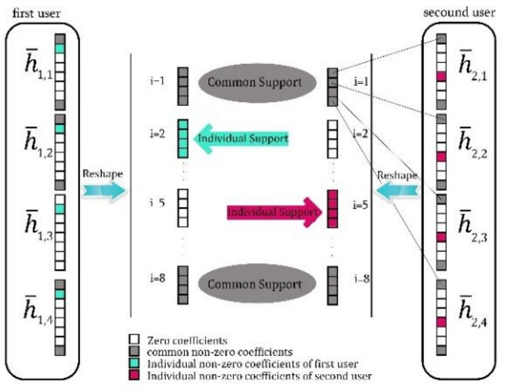


Figure 1. The schematic diagram of the channel between the BS antennas and the users for $N_t = 4$, $L = 8$ and $K = 2$.

The structured sparsity pattern of the channel matrix at t^{th} snapshot is represented as:

$$\mathbf{H}_t = \left[\begin{array}{ccccc} \bigstar & \bigstar & \bigstar & \cdots & \bigstar \\ \vdots & \vdots & \vdots & & \vdots \\ \bigstar & \bigstar & \bigstar & \cdots & \bigstar \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ \bigstar & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \bigstar & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \bigstar \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \bigstar \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{l=1} \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{l=2} \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{l=3} \\ \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} \text{l=L} \end{array} \quad (6)$$

Also, we define $\mathbf{P}_l$ as:

$$\mathbf{P}_l = [\mathbf{X}_1 \mathbf{F}_L(:,l), \mathbf{X}_2 \mathbf{F}_L(:,l), \ldots, \mathbf{X}_{N_t} \mathbf{F}_L(:,l)] \quad (7)$$

By concatenating the $N \times N_t$ sub-matrices, $\mathbf{P}_l$, we obtain the matrix $\mathbf{P}$ as:

$$\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_L] \quad (8)$$

As a result, the received matrix at the t^{th} snapshot in the BS can be written as:

$$\mathbf{Y}_t = \mathbf{P}\mathbf{H}_t + \mathbf{N}_t \quad (9)$$

where $\mathbf{Y}_t = [\mathbf{y}_1(t), \mathbf{y}_2(t), \ldots, \mathbf{y}_k(t)]$ and $\mathbf{N}_t = [\mathbf{n}_1(t), \mathbf{n}_2(t), \ldots, \mathbf{n}_k(t)]$

By concatenating the T consecutive time slots of the received signals, we will have:

$$\mathcal{Y}(:,:,t) = \mathbf{Y}_t \quad (10)$$

where $\mathcal{Y} \in \mathbb{C}^{N \times K \times T}$ is the received tensor. N, K, and T are the length, width, and depth of the tensor, respectively.

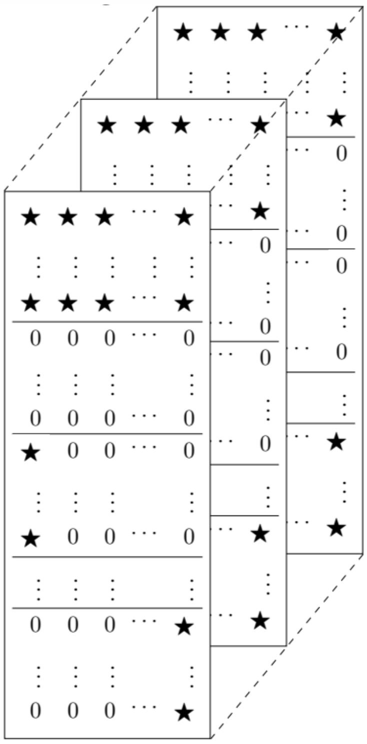
T shall be less than or equal to the channel coherence time. Since the channel has slow variation over time, the channel sparsity pattern would not change considerably during T consecutive snapshots. In other words, the path delays vary much slower than the path gains due to the temporal channel correlation, so the channel support is almost unchanged during the coherence time [24]. The concatenated channel matrix can be constructed in a similar way as:

$$\mathcal{H}(:,:,t) = \mathbf{H}_t \quad (11)$$

where $\mathcal{H} \in \mathbb{C}^{N_t L \times K \times T}$.

Hence, the channel model is presented in consecutive time slots as a tensor with non-zero 3D blocks for joint support between users and 2D blocks for each user's individual support in sequential time slots.

The following shows a schematic of the channel tensor.



III. The Proposed Channel Estimation Techniqe

In this section, we illustrate the proposed method to recover the tensor sparse signal for MIMO channel estimation. The aim is to recover the tensor $\mathbf{H}(\mathbf{t})$ from the received matrix $\mathbf{Y}(\mathbf{t})$ over T time slots. The suggested scheme is a greedy tensor sparse recovery technique which we call tensor orthogonal matching pursuit (TOMP).

At the first step, the algorithm estimates the common support between the users. The next step is to obtain the individual supports of the various users channel. At the last step, the channel matrix elements are recovered from the estimated support and the received matrices using the least squares technique. The details of the proposed TOMP algorithm has been given in Algorithm 1.

It should be noted that in this algorithm, Ω^c represents common support of the users. Ω_k^i represents individual support of the k^{th} user. The operator $I_{j\in\omega}$ has value 1 if j belongs to the set ω and is zero otherwise.

Algorithm 1 The proposed TOMP method

1: **input:**
2: The received tensor at the BS: $\mathcal{Y}$.
3: The pilot matrix: $\mathbf{P}$.
4: The individual support size: S_i.
5: The common support size: S_c.
6: The number of time slots: T.
7: The predefined threshold value: η.
8: **output:**
9: The estimated channel tensor: $\widehat{\mathcal{H}}$.
10: **procedure** The proposed method
11: $\widehat{\mathcal{H}} \leftarrow \mathbf{0}$
12: Common Support Identification:
13: $\Omega^c \leftarrow \emptyset$.
14: $\mathbf{R}_t \leftarrow \mathbf{Y}_t$
15: **for** i=1:S_c **do**
16: $\mathbf{a} \leftarrow \mathbf{0}$
17: $\mathbf{a}(l) \leftarrow \Sigma_{t=1}^{T} \|\mathbf{R}_t^H \mathbf{P}_l\|_F, \quad \forall_l$
18: $c \leftarrow sort[\mathbf{a}, descend]$
19: $d \leftarrow c[1:2\times(S_c - i)]$
20: $d_k^{'} \leftarrow \{j : j \in d, \Sigma_{t=1}^{T} \|\mathbf{R}_t^H(:,k)\mathbf{P}_d\|_F \geqq T\times\eta\}\,\forall k$
21: $\Omega^c \leftarrow \Omega^c \cup \left\{argmax_j \Sigma_{k=1}^{K} I_{\{j\in d_k^{'}\}}\right\}$
22: $\mathbf{R}_t \leftarrow \mathbf{R}_t - (\mathbf{P}_{\Omega^C}\mathbf{P}_{\Omega^C}^{\dagger}\mathbf{R}_t), \forall_t$
23: **end for**
24: Individual Support Identification:
25: **for** k=1:K **do**
26: $\mathbf{r}_t \leftarrow \mathbf{R}_t(:,k)$
27: $\Omega_k^i = \emptyset$, $\mathbf{b} \leftarrow \mathbf{0}$
28: **for** j=1:S_i **do**
29: $\mathbf{b}(l) \leftarrow \Sigma_{t=1}^{T} \|\mathbf{r}_t^H \mathbf{P}_l\|_F, \forall_l$
30: $\Omega_k^i \leftarrow \Omega_k^i \cup argmax(\mathbf{b})$
31: $\mathbf{r}_t \leftarrow \mathbf{r}_t - (\mathbf{P}_{\Omega^i}\mathbf{P}_{\Omega^i}^{\dagger})\mathbf{r}_t$
32: **end for**
33: $\Omega_k \leftarrow [\Omega^c, \Omega_k^i]$
34: **for** t=1:T **do**
35: $\mathbf{z} \leftarrow \mathcal{Y}(:,k,t)$
36: $\widehat{\mathcal{H}}_{\Omega_k}(:,k,t) \leftarrow \mathbf{P}_{\Omega_k}^{\dagger}\mathbf{z}$
37: **end for**
38: **end for**
39: **end procedure**

The algorithm consists of two parts: In the first part, the common support of the users is estimated. At the second part, the individual supports of the users are estimated. In Lines 17-19, the candidate support values have been detected based on the correlation of the residual and sensing matrices. In Line 20, the support pruning is conducted using a thresholding operator. The common support element is identified in Line 21 using the maximum number of occurrences among all the users. In Line 22, the effect of the identified common support element has been removed from the residual matrix to produce an updated residual matrix. After estimating the common support, the individual support elements have been detected. For each of the users, the correlation of the residual vector and the pilot matrix is calculated (Line 29). The corresponding element of the maximum correlation is selected as the individual support element of the user (Line 30). The effect of the chosen support element has been removed from the residual vector (Line 31). After estimating both the common support and individual supports, the channel matrix is estimated based on the least squares technique (Line 36).

IV. Simulation Results

In this section, the simulation results are reported. We consider a single-cell massive MIMO system with

$N_t = 16$ antennas at BS and $K = 8$ single-antenna users. The channel vectors are assumed to have $L = 16$ taps and T time slots. The thresholding value of the proposed TOMP algorithm has been set to $\eta = 21.4$. The non-zero channel locations are selected uniformly at random. The pilot signals are random values generated from the standard normal distribution. The proposed method is compared with three related schemes (OMP [25, 26], Joint OMP [20], and Block ISD [18]) as benchmarks in different scenarios. In addition, the Exact least squares (LS) algorithm has been simulated as the lower bound of channel estimation error. In the Exact LS algorithm, it is assumed that the BS has the prior knowledge of the channel supports and estimates the channel vectors by least squares technique.

The normalized mean square error (NMSE) has been considered as the performance measure which is defined as:

$$NMSE = \frac{\| \mathbf{H} - \mathbf{H} \|_F^2}{\| \mathbf{H} \|_F^2} \quad (12)$$

The simulations have been repeated for 10 times and the average NMSE has been reported in different experiments.

At the first scenario, we examine the effect of the number of time slots on the accuracy of channel estimation using the proposed algorithm. Figure 2 shows the average NMSE curves versus the number of time slots (T). In this simulation, the value of SNR is $12dB$, the number of pilots is $N = 100$, the common support size is $S_c = 3$ and the size of individual support is $S_i = 1$ for each of the users.

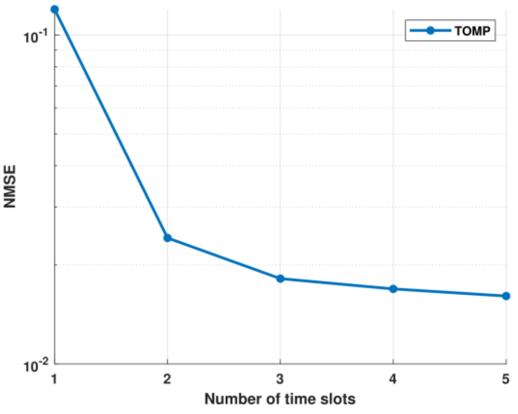


Figure 2. NMSE for $S_c = 3$, $S_i = 1$, $SNR = 12dB$ and $N = 100$ versus T.

As expected, as the time passes, the error decreases. Because larger number of time slots leads to higher accuracy in detecting non-zero blocks.

In the next simulation, we investigate the performances of methods for various SNR values ranging from $2dB$ to $20dB$. Figure 3 shows the average NMSE curves versus SNR. In this simulation, the number of pilots is $N = 90$, the number of time slots is $T = 10$, the common support size is set to $S_c = 2$, and the size of individual support is $S_i = 1$.

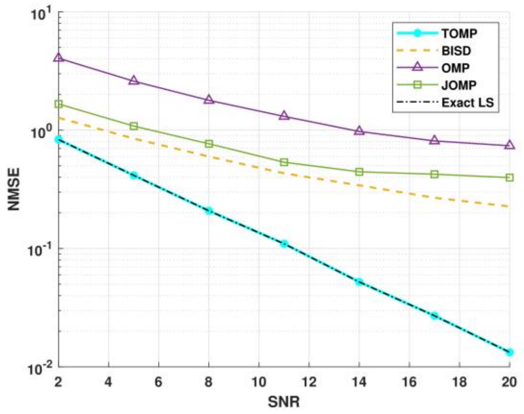


Figure 3. NMSE Comparison of the methods for $S_c = 2$, $S_i = 1$, $N = 90$ and $T = 10$ versus SNR.

According to this figure, we see that the channel estimation error of the methods decreases with the SNR value. Moreover, the suggested TOMP method outperforms the other methods and its average NMSE curve coincides with that of the exact LS scheme which has been selected as a performance bound. This indicates that all non-zero blocks have been correctly detected and the support estimation is perfect.

In the third simulation, we study the performances of algorithms in the case of changing the number of pilots. The average NMSE curves of the methods have been depicted versus the number of pilots in Figure 4. We have considered $S_c = 2$, $S_i = 1$, $T = 10$ and $SNR = 12dB$.

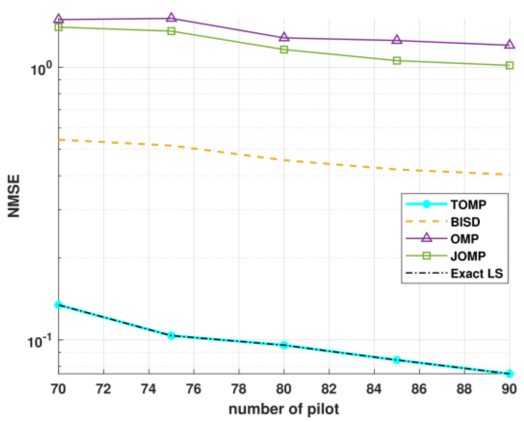


Figure 4. Comparison of the methods for $S_c = 2$, $S_i = 1$, $T = 10$ and $SNR = 12dB$ for different number of pilots.

By increasing the number of pilots, the channel estimation error of all the methods is reduced. This decreasing trend is more outstanding for the TOMP method compared to the others. The proposed method outperforms its counterparts. It achieves lower NMSE with lower number of pilots.

It should be noted that the number of pilots required for the reliable estimation of the conventional methods is $N_tL = 16 \times 16 = 256$ [18] which is much more than that of the sparsity-based schemes.

In another experiment, we investigate the effect of the common support size, S_c, in the estimation accuracy of the methods. To this end, we depict the average NMSE of the methods versus S_c ranging from 1 to 4 in Figure 5. We have considered $SNR = 12dB$, $N = 90$, $T = 5$ and $S_i = 1$ for this simulation.

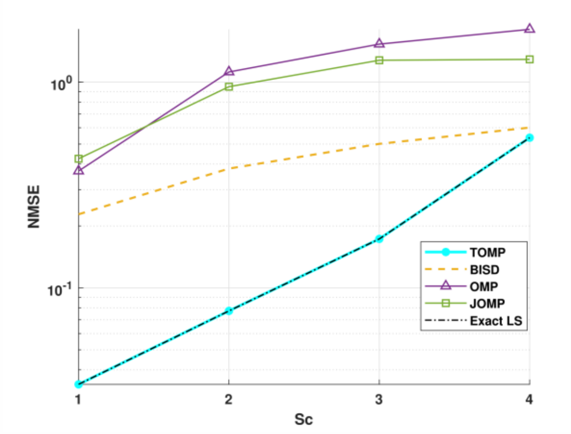


Figure 5. Average NMSE versus common support size, S_c, for $SNR = 12dB$, $N = 90$, $T = 5$ and $S_i = 1$.

According to this figure, we observe that the estimation error of the methods increases with the common support size. This increasing trend is due to the fact that for larger values of S_c, the channel sparsity is reduced (the channel becomes denser). Therefore, the performance of the sparsity-based schemes deteriorates. Furthermore, owing to the usage of tensor sparsity property of the channel in the proposed algorithm, this scheme still works better than the other methods and coincides with the exact LS method.

In the last scenario, we study the effect of the individual support size, S_i. The average NMSE curves of the methods versus different values of S_i have been shown in Figure 6. In this test, we have set $SNR = 12dB$, $N = 90$, $T = 5$ and $S_c = 1$.

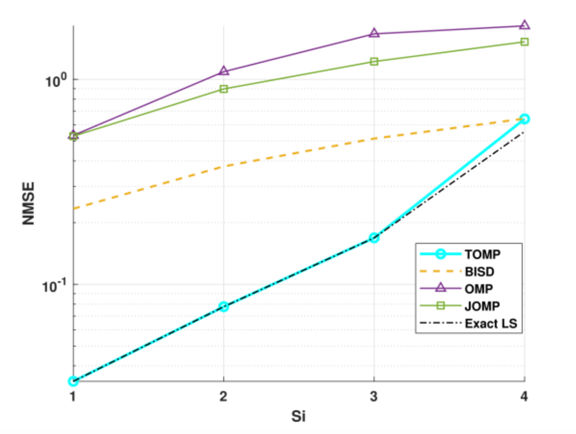


Figure 6. Comparison of the methods for $SNR = 12dB$, $N = 90$, $S_c = 1$ and different values of S_i.

The results of this figure also validates that the suggested method has better performance. In addition, similar to the previous figure, increasing the support size deteriorates the estimation accuracy of all of the methods.

As a general conclusion, the performance of the proposed method in all of the simulated scenarios is much better than those of the benchmarks and coincides with the exact LS algorithm. It should be emphasized that the exact LS algorithm has the ideal prior information of the channel support. Therefore, this indicates that the proposed method has achieved accurate channel support estimation. This advantage is due to the fact that the proposed method exploits the tensor sparsity property which leads to high accuracy in channel estimation.

V. CONCLUSION

This paper considers the challenge of channel estimation in slow- varying massive MIMO systems. At first, the channel of the consecutive time slots has been modeled as a tensor with a few non-zero 3D blocks which we called tensor sparsity pattern. In the next step, a tensor sparse recovery algorithm called TOMP has been proposed to estimate the channel. The simulation results have confirmed the outperformance of the suggested TOMP scheme in comparison with its rivals.

REFERENCES

[1] Dahlman, E., et al., *3G evolution: HSPA and LTE for mobile broadband*. 2010: Academic press.
[2] Lee, B., et al. *Antenna grouping based feedback reduction for FDD-based massive MIMO systems*. in *2014 IEEE International Conference on Communications (ICC)*. 2014. IEEE.
[3] Sadeghi, N. and M. Azghani, *Multi-user massive MIMO channel estimation using joint sparsity and non-ideal feedback modeling*. Digital Signal Processing, 2020. **100**: p. 102640.
[4] Gao, Z., et al., *Spatially common sparsity based adaptive channel estimation and feedback for FDD massive*

MIMO. IEEE Transactions on Signal Processing, 2015. **63**(23): p. 6169-6183.
[5] Larsson, E.G., et al., *Massive MIMO for next generation wireless systems*. IEEE communications magazine, 2014. **52**(2): p. 186-195.
[6] Nguyen, S.L.H. and A. Ghrayeb. *Compressive sensing-based channel estimation for massive multiuser MIMO systems*. in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*. 2013. IEEE.
[7] Biguesh, M. and A.B. Gershman, *Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals*. IEEE transactions on signal processing, 2006. **54**(3): p. 884-893.
[8] Yin, H., et al., *A coordinated approach to channel estimation in large-scale multiple-antenna systems*. IEEE Journal on selected areas in communications, 2013. **31**(2): p. 264-273.
[9] Zhou, Y., et al., *Experimental study of MIMO channel statistics and capacity via the virtual channel representation*. Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep, 2007. **5**: p. 10-15.
[10] Kyritsi, P., et al., *Correlation analysis based on MIMO channel measurements in an indoor environment*. IEEE Journal on Selected areas in communications, 2003. **21**(5): p. 713-720.
[11] Kaltenberger, F., et al. *Correlation and capacity of measured multi-user MIMO channels*. in *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*. 2008. IEEE.
[12] Hoydis, J., et al. *Channel measurements for large antenna arrays*. in *2012 International Symposium on Wireless Communication Systems (ISWCS)*. 2012. IEEE.
[13] Azghani, M. and F. Marvasti, *Sparse signal processing*, in *New perspectives on approximation and sampling theory*. 2014, Springer. p. 189-213.
[14] Afkhaminia, F. and M. Azghani, *2D off-grid DOA estimation using joint sparsity*. IET Radar, Sonar & Navigation, 2019. **13**(9): p. 1580-1587.
[15] Azghani, M., et al., *Missing low-rank and sparse decomposition based on smoothed nuclear norm*. IEEE Transactions on Circuits and Systems for Video Technology, 2019. **30**(6): p. 1550-1558.
[16] Sedghi, R. and M. Azghani, *Sparsity-based MIMO interference suppression technique in the presence of imperfect channel state information*. IET Communications, 2019. **13**(19): p. 3201-3206.
[17] Maleki, N. and M. Azghani, *Sparse Mixed Norm Adaptive Filtering Technique*. Circuits, Systems, and Signal Processing, 2020. **39**(11): p. 5758-5775.
[18] Shen, W., et al. *Massive MIMO channel estimation based on block iterative support detection*. in *2016 IEEE Wireless Communications and Networking Conference*. 2016. IEEE.
[19] Gao, Z., et al., *Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO*. IEEE Transactions on Communications, 2015. **64**(2): p. 601-617.
[20] Rao, X. and V.K. Lau, *Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems*. IEEE Transactions on Signal Processing, 2014. **62**(12): p. 3261-3271.
[21] Sadeghi, N. and M. Azghani. *Channel Estimation using Block Sparse Joint Orthogonal Matching Pursuit in Massive MIMO Systems*. in *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*. 2021. IEEE.
[22] Bottou, L. and N. Murata, *Stochastic approximations and efficient learning*. The Handbook of Brain Theory and Neural Networks, Second edition,. The MIT Press, Cambridge, MA, 2002.
[23] Barbotin, Y., et al., *Estimation of sparse MIMO channels with common support*. IEEE Transactions on Communications, 2012. **60**(12): p. 3705-3716.
[24] Telatar, I.E. and D.N.C. Tse, *Capacity and mutual information of wideband multipath fading channels*. IEEE transactions on information theory, 2000. **46**(4): p. 1384-1400.
[25] Berger, C.R., et al., *Application of compressive sensing to sparse channel estimation*. IEEE Communications Magazine, 2010. **48**(11): p. 164-174.
[26] Bajwa, W.U., et al., *Compressed channel sensing: A new approach to estimating sparse multipath channels*. Proceedings of the IEEE, 2010. **98**(6): p. 1058-1076.

Nasser Sadeghi received his B.Sc. and M.Sc. degrees in Communication Engineering from Bonab University, Tabriz, Iran, in 2015 and Sahand University of Technology, Tabriz, Iran, in 2018, respectively. He is currently pursuing his PhD in Sahand University of Technology. His research interests include Signal Processing, Massive MIMO, and wireless communication.

Masoumeh Azghani received the B.Sc. and M.Sc. degrees in Electrical Engineering from Tabriz University in 2009 and 2011, respectively, and the Ph.D. degree in communications from the Sharif University of Technology, Tehran, Iran, in 2014. She was a Visiting Researcher with the Institute for Info Comm Research (I2R)-ASTAR, Singapore, in 2015. In September 2015, she joined the Electrical Engineering Department of the Sahand University of Technology as an Assistant Professor. Since December 2019, she is an Associate Professor in Sahand University of Technology. Her research interests include various fields of Signal Processing and Wireless Communications.