Uplink channel estimation for massive MIMO systems exploring joint channel sparsity

Chenhao Qi and Lenan Wu

The joint sparsity of uplink channels in massive multi-input–multi-output (MIMO) systems is explored and a block sparse model is proposed for joint channel estimation. The block coherence of this model is analysed. It is indicated that as the number of antennas at the base station grows to be infinity, the block coherence will be zero. Then a block optimised orthogonal matching pursuit (BOOMP) algorithm is proposed. Simulation results verify the analysis and show that the joint estimation using the BOOMP algorithm can significantly improve the channel estimation performance.

Introduction: The use of massive multiple-input-multiple-output (MIMO) systems where the base station (BS) is equipped with tens or even hundreds of antennas have attracted much interest [1]. It is shown that as the number of antennas at the BS grows to a unprecedented number, the additive noise and Rayleigh fading effect will be negligible, and what dominates is the inter-cell interference caused by pilot contamination. To reduce the pilot overhead, one potential choice is to explore the inherent sparsity of wireless channels and to use sparse channel estimation [2–3]. In [2], a superimposed pilot design for downlink frequency-division duplex (FDD) massive MIMO systems is proposed based on structured compressed sensing (CS). In [3], sparse channel estimation with structured CS is proposed for multi-input–single-output systems. However, no work has been reported on uplink channel estimation in a time-division duplex (TDD) massive MIMO system and explored joint channel sparsity. As another option for a long-term evolution advanced standard, TDD is different to FDD in that the downlink channel state information is obtained by uplink channel estimation via channel reciprocity. In TDD systems, the burden of channel estimation is delivered from the user terminal to the BS, where the battery power of the user terminal to perform channel estimation can be saved.

In this Letter, we explore the joint sparsity of uplink channels from the user terminal to the BS. We propose a block sparse model for uplink channels sharing common support. Then the block coherence is analysed. As the number of antennas grows to be infinity, the block coherence will be zero. We then propose a block optimised orthogonal matching pursuit (BOOMP) algorithm for joint sparse channel estimation. The notations used in this Letter are defined as follows. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, $\boldsymbol{I}_M$, $\|\cdot\|_2$, $\mathcal{CN}$, $/$, $\cup$ and $\varnothing$ denote the matrix transpose, conjugate transpose (Hermitian), the identity matrix of size M, ℓ_2-norm, the complex Gaussian distribution, the set exclusion, the set union and the empty set, respectively.

Problem formulation: We consider a massive MIMO system including a BS equipped with M antennas and several user terminals each equipped with one antenna. We use orthogonal frequency-division multiplexing (OFDM) for uplink transmission. Suppose the total number of OFDM subcarriers is N. The user terminal employs $K(0<K\le N)$ subcarriers with the corresponding indices $\boldsymbol{p}=[P_1, P_2, \ldots, P_K](1\le P_1<P_2<\cdots<P_K\le N)$ to transmit pilot symbols for pilot-assisted channel estimation. The transmit pilot vectors are denoted as $\boldsymbol{x}=[x(P_1), x(P_2), \ldots, x(P_K)]^{\mathrm{T}}$. Then the BS will receive M different pilot vectors, denoted as $\boldsymbol{y}^{(i)}=[y^{(i)}(P_1), y^{(i)}(P_2), \ldots, y^{(i)}(P_K)]^{\mathrm{T}}$, $i=1, 2, \ldots, M$, each experiencing different multipath fading. We denote the channel impulse response (CIR) of each multipath channel as $\boldsymbol{h}^{(i)}=[h^{(i)}(1), h^{(i)}(2), \ldots, h^{(i)}(L)]^{\mathrm{T}}$, $i=1, 2, \ldots, M$. We have

$$\boldsymbol{y}^{(i)}=\boldsymbol{DFh}^{(i)}+\boldsymbol{\eta}^{(i)}, \quad i=1, 2, \ldots, M \tag{1}$$

where $\boldsymbol{D}\triangleq\mathrm{diag}\{\boldsymbol{x}\}$ denotes a diagonal matrix with the kth diagonal entries being $x(P_k)$, $k=1, 2, \ldots, K$, $\boldsymbol{F}$ is a K by L submatrix indexed by $\boldsymbol{p}=[P_1, P_2, \ldots, P_K]$ in row and $[1, 2, \ldots, L]$ in column from a standard N by N discrete Fourier transform matrix and $\boldsymbol{\eta}^{(i)}\sim\mathcal{CN}(0, \sigma^2\boldsymbol{I}_K)$ is an additive white Gaussian noise term of the ith uplink channel. We define the measurement matrix $\boldsymbol{A}\triangleq\boldsymbol{DF}$, then (1) is rewritten as

$$\boldsymbol{y}^{(i)}=\boldsymbol{Ah}^{(i)}+\boldsymbol{\eta}^{(i)}, \quad i=1, 2, \ldots, M \tag{2}$$

It has often been pointed out in the literature that the wireless channel is essentially sparse, where the number of non-zero taps of the channel, denoted as S, is much smaller than the channel length $L(0<S\ll L)$. Then the CS techniques can be applied for sparse channel estimation. Furthermore, it is shown in [4] that the CIR of different uplink channels shares a common support because the time of arrival at different antennas is similar while the paths amplitudes and phases are distinct. In other words, the non-zero positions of $\boldsymbol{h}^{(i)}$ are the same for $i=1, 2, \ldots, M$, whereas their non-zero coefficients are different. We can jointly consider the M equations in (2) and explore their joint sparsity. We define a stack vector of received pilots as $\boldsymbol{z}\triangleq\left[\boldsymbol{z}_1^{\mathrm{T}}, \boldsymbol{z}_2^{\mathrm{T}}, \ldots, \boldsymbol{z}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ where $\boldsymbol{z}_l\triangleq\left[y^{(1)}(l), y^{(2)}(l), \ldots, y^{(M)}(l)\right]^{\mathrm{T}}$ denotes the lth block of $\boldsymbol{z}$, $l=1, 2, \ldots, K$. In the same way, we define a stack vector of noise terms as $\boldsymbol{n}\triangleq\left[\boldsymbol{n}_1^{\mathrm{T}}, \boldsymbol{n}_2^{\mathrm{T}}, \ldots, \boldsymbol{n}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ where $\boldsymbol{n}_l\triangleq\left[\eta^{(1)}(l), \eta^{(2)}(l), \ldots, \eta^{(M)}(l)\right]^{\mathrm{T}}$ denotes the lth block of $\boldsymbol{n}$, $l=1, 2, \ldots, K$. We denote the entry at the ith-row jth-column of $\boldsymbol{A}$ as $A(i, j)$. We generate a new measurement matrix $\boldsymbol{B}$ by substituting $A(i, j)$ with an M by M diagonal matrix $A(i, j)\boldsymbol{I}_M$. Therefore, $\boldsymbol{B}$ is composed of totally KL blocks of diagonal matrix $\boldsymbol{B}(i, j)\triangleq A(i, j)\boldsymbol{I}_M$, $i=1, 2, \ldots, K$, $j=1, 2, \ldots, L$, where $\boldsymbol{B}(i, j)$ denotes the (i, j)th block of $\boldsymbol{B}$. Then the M equations in (2) can be combined together in one equation for the massive MIMO system as

$$\boldsymbol{z}=\boldsymbol{Bw}+\boldsymbol{n} \tag{3}$$

where $\boldsymbol{w}$ is defined as a stack vector $\boldsymbol{w}\triangleq\left[\boldsymbol{w}_1^{\mathrm{T}}, \boldsymbol{w}_2^{\mathrm{T}}, \ldots, \boldsymbol{w}_K^{\mathrm{T}}\right]^{\mathrm{T}}$ with $\boldsymbol{w}_l\triangleq\left[h^{(1)}(l), h^{(2)}(l), \ldots, h^{(M)}(l)\right]^{\mathrm{T}}$ denoting the lth block of $\boldsymbol{w}$, $l=1, 2, \ldots, L$. Since the non-zero positions of $\boldsymbol{h}^{(i)}$ are the same for $i=1, 2, \ldots, M$, the entries of $\boldsymbol{w}_i$ will be either all zero or all non-zero, exhibiting the 'block sparsity'. Correspondingly, we represent $\boldsymbol{B}$ as a concatenation of column-blocks $\boldsymbol{B}_l$, $l=1, 2, \ldots, L$, as

$$\boldsymbol{B}\triangleq\Big[\underbrace{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_M}_{\boldsymbol{B}_1}, \underbrace{\boldsymbol{b}_{M+1}, \boldsymbol{b}_{M+2}, \ldots, \boldsymbol{b}_{2M}}_{\boldsymbol{B}_2}, \ldots, \underbrace{\boldsymbol{b}_{LM-M+1}, \boldsymbol{b}_{LM-M+2}, \ldots, \boldsymbol{b}_{LM}}_{\boldsymbol{B}_L}\Big] \tag{4}$$

where $\boldsymbol{b}_j$ denotes the jth column of $\boldsymbol{B}$, $j=1, 2, \ldots, LM$. It is observed that the columns within each block $\boldsymbol{B}_l$ are orthogonal to each other, meaning that the rank of each block is M.

Analysis of block coherence: Assume that each column of $\boldsymbol{B}$ is normalised. This assumption is reasonable because we can normalise $\boldsymbol{B}$ by simply decomposing it into a normalised matrix $\boldsymbol{Q}$ and a diagonal matrix $\boldsymbol{G}$ so that $\boldsymbol{B}=\boldsymbol{QG}$. After the sparse recovery, we can obtain the solution to the original problem by multiplying the results with $\boldsymbol{G}^{-1}$.

We define the 'coherence' of $\boldsymbol{A}$ in a manner consistent with the literature as

$$\mu(\boldsymbol{A})=\max_{l\neq k}|\boldsymbol{a}_l^{\mathrm{H}}\boldsymbol{a}_k| \tag{5}$$

where $\boldsymbol{a}_l$ denotes the lth column of $\boldsymbol{A}$, $l=1, 2, \ldots, L$. To improve the sparse channel estimation performance for each uplink channel, it is better to minimise $\mu(\boldsymbol{A})$ [5]. Similarly, we define the 'coherence' of $\boldsymbol{B}$ as

$$\mu(\boldsymbol{B})=\max_{l\neq k}|\boldsymbol{b}_l^{\mathrm{H}}\boldsymbol{b}_k| \tag{6}$$

where the definition of $\boldsymbol{b}_l$ is given in (4). Obviously, we have $\mu(\boldsymbol{B})=\mu(\boldsymbol{A})$. We further define the 'block coherence' of $\boldsymbol{B}$ according to [6] as

$$\mu_B(\boldsymbol{B})=\frac{1}{M}\max_{l\neq k}\rho(\boldsymbol{B}_l^{\mathrm{H}}\boldsymbol{B}_k) \tag{7}$$

where we denote the spectrum norm of a given matrix $\boldsymbol{R}$ as $\rho(\boldsymbol{R})\triangleq\lambda_{\max}^{1/2}(\boldsymbol{R}^{\mathrm{H}}\boldsymbol{R})$, with $\lambda_{\max}(\boldsymbol{R}^{\mathrm{H}}\boldsymbol{R})$ representing the largest eigenvalue of the positive-semi-definite matrix $\boldsymbol{R}^{\mathrm{H}}\boldsymbol{R}$. It can be derived that

$$\begin{aligned}\mu_B(\boldsymbol{B})&=\frac{1}{M}\max_{l\neq k}\rho(\boldsymbol{a}_k^{\mathrm{H}}\boldsymbol{a}_l\boldsymbol{I}_M)=\frac{1}{M}\max_{l\neq k}\lambda_{\max}^{1/2}(|\boldsymbol{a}_k^{\mathrm{H}}\boldsymbol{a}_l|^2\boldsymbol{I}_M)\\&=\frac{1}{M}\max_{l\neq k}|\boldsymbol{a}_k^{\mathrm{H}}\boldsymbol{a}_l|=\frac{1}{M}\mu(\boldsymbol{A})\end{aligned} \tag{8}$$

If M grows to be infinity, $\mu_B(\boldsymbol{B})$ will be zero, which means that the blocks B_l, $l=1, 2, \ldots, L$ in (4) will be orthogonal to each other, leading to the unique recovery of blocks. Therefore, as the number of antennas at the BS in massive MIMO systems grows to be large, the probability to jointly recover the positions of non-zero channel taps will be very high. In this way, we can reduce the pilot overhead and therefore relieve the pilot contamination in massive MIMO systems.

Block optimised orthogonal matching pursuit: Now we propose a BOOMP algorithm for the proposed model in (3). Note that the BOOMP algorithm presented in this Letter is based on the optimised OMP algorithm [7] instead of the basic OMP algorithm.

Algorithm 1: Block optimised orthogonal matching pursuit

1: Input: $\boldsymbol{B}$, z, M, L, σ.
2: Initialisations: $\boldsymbol{r} \Leftarrow z$. $T \Leftarrow 0$. $\boldsymbol{\Lambda} \Leftarrow \emptyset$
3: **while** $\|\boldsymbol{r}\|_2 > M\sigma$ and $T \leq L$
4: $T \Leftarrow T + 1$.
5: Obtain J via (9).
6: $\boldsymbol{\Lambda} \Leftarrow \boldsymbol{\Lambda} \cup \{J\}$.
7: $\boldsymbol{r} \Leftarrow z - \boldsymbol{B}_{\boldsymbol{\Lambda}}(\boldsymbol{B}_{\boldsymbol{\Lambda}}^{\mathrm{H}}\boldsymbol{B}_{\boldsymbol{\Lambda}})^{-1}\boldsymbol{B}_{\boldsymbol{\Lambda}}^{\mathrm{H}}z$.
8: **end**
9: Output: $\hat{\boldsymbol{h}}_{\boldsymbol{\Lambda}}^{(i)} \Leftarrow (\boldsymbol{A}_{\boldsymbol{\Lambda}}^{\mathrm{H}}\boldsymbol{A}_{\boldsymbol{\Lambda}})^{-1}\boldsymbol{A}_{\boldsymbol{\Lambda}}^{\mathrm{H}}\boldsymbol{y}^{(i)}, \quad i = 1, 2, \ldots, M$.

First, we initialise a residue vector $\boldsymbol{r} \Leftarrow z$ and a loop counter $T \Leftarrow 0$. At each iteration, we obtain an index of the non-zero entry of $\boldsymbol{h}^{(i)}$ by

$$J = \arg \max_{j \in \{1, 2, \ldots, L\} \setminus \boldsymbol{\Lambda}} \|(\boldsymbol{B}_j^{\mathrm{H}}\boldsymbol{B}_j)^{-1}\boldsymbol{B}_j^{\mathrm{H}}\boldsymbol{r}\|_2 \tag{9}$$

and keep J in an active set $\boldsymbol{\Lambda}$. Since $\boldsymbol{h}^{(i)}$ shares a common support for $i = 1, 2, \ldots, M$, we need only one active set to keep the common support. We denote the submatrix indexed by $\boldsymbol{\Lambda}$ in blocks from $\boldsymbol{B}$ and the submatrix indexed by $\boldsymbol{\Lambda}$ in columns from $\boldsymbol{A}$ as $\boldsymbol{B}_{\boldsymbol{\Lambda}}$ and $\boldsymbol{A}_{\boldsymbol{\Lambda}}$, respectively. We iteratively update the residue $\boldsymbol{r}$ by the least squares (LS) estimation in step 7 of algorithm 1, where $(\boldsymbol{B}_{\boldsymbol{\Lambda}}^{\mathrm{H}}\boldsymbol{B}_{\boldsymbol{\Lambda}})^{-1}\boldsymbol{B}_{\boldsymbol{\Lambda}}^{\mathrm{H}}$ is the pseudo inverse of $\boldsymbol{B}_{\boldsymbol{\Lambda}}$. Once the power of the residue is comparable to the noise or the number of iterations is greater than L, we stop the iterations. Meanwhile, we output the estimated CIR as $\hat{\boldsymbol{h}}^{(i)}$, with the coefficients of non-zero taps denoted as $\hat{\boldsymbol{h}}_{\boldsymbol{\Lambda}}^{(i)}$, $i = 1, 2, \ldots, M$.

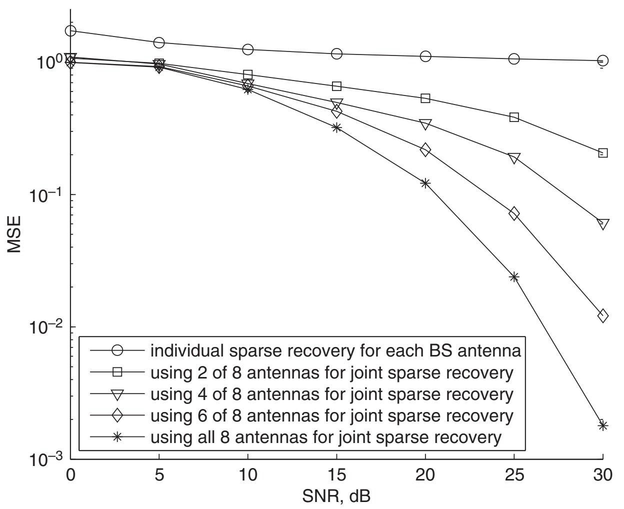

Fig. 1 *Comparisons of individual sparse channel estimation and joint sparse channel estimation in terms of MSE*

Simulation results: We consider a massive MIMO system using $N = 256$ OFDM subcarriers for uplink transmission, where $K = 16$ subcarriers are used to transmit pilot symbols. The indices of pilot subcarriers after pilot optimisation is $\boldsymbol{p} = [8, 40, 48, 52, 72, 82, 99, 142, 145, 154, 158, 161, 183, 209, 212, 230]$, achieving a very small coherence $\mu(\boldsymbol{A}) = 4.7021$ [5]. Suppose the channel length is $L = 60$. Now we compare the individual sparse recovery using the OMP and the joint sparse recovery using the BOOMP, where we set $S = 12$ and $M = 8$. Since $K \leq 2S$, the individual sparse recovery for each link cannot succeed, from an information theoretical point of view. As shown in Table 1, individually estimated positions of non-zero entries of $\boldsymbol{h}^{(i)}$ for each BS antenna are all incorrect. Then we use 2 of 8, 4 of 8, 6 of 8 and all 8 BS antennas, respectively, for joint sparse recovery. It is seen that we can obtain the true positions exactly as those of the original CIR if we use all 8 antennas for joint sparse recovery, whereas the results using 6 of 8 antennas are very close to the original CIR. Therefore, the estimation performance will increase if we use more antennas for joint sparse recovery. In Fig. 1 we compare the individual sparse channel estimation and joint sparse channel estimation in terms of mean square errors (MSEs). It is seen that the joint estimation using algorithm 1 performs much better than the individual estimation using the OMP; and the latter exhibits the floor effect at the high SNR region while the former does not. Moreover, we can further improve the MSE performance by employing more BS antennas for joint sparse channel estimation, which shows that our scheme is typically beneficial for massive MIMO systems.

Table 1: Comparisons of individual sparse recovery and joint sparse recovery for $M = 8$

	Positions of non-zero entries
Original	2, 13, 21, 24, 29, 33, 41, 42, 43, 53, 54, 60
1st antenna	2, 3, 13, 21, 29, 33, 42, 53, 54
2nd antenna	2, 9, 20, 21, 29, 30, 39, 42, 49, 55, 56, 60
3rd antenna	2, 6, 7, 21, 24, 29, 33, 37, 43, 46, 53, 54, 59
4th antenna	2, 12, 13, 15, 21, 23, 25, 29, 34, 44, 45, 53
5th antenna	2, 5, 7, 8, 13, 21, 41, 43, 49, 53, 54, 56, 59
6th antenna	3, 6, 14, 19, 21, 25, 40, 41, 43, 47, 53, 54, 55
7th antenna	1, 2, 14, 15, 17, 24, 30, 31, 37, 44, 50, 59
8th antenna	1, 2, 6, 13, 21, 24, 32, 33, 42, 47, 50, 53, 54, 55
Joint 2 antennas	1, 3, 13, 21, 22, 24, 36, 39, 43, 45, 54, 55, 58, 60
Joint 4 antennas	2, 21, 24, 29, 33, 41, 42, 43, 53, 54, 60
Joint 6 antennas	2, 13, 21, 24, 33, 41, 42, 43, 49, 53, 54, 60
Joint 8 antennas	2, 13, 21, 24, 29, 33, 41, 42, 43, 53, 54, 60

Conclusion: We have proposed a block sparse model for uplink channels in massive MIMO systems. We have also proposed a BOOMP algorithm for joint sparse channel estimation. Simulation results have shown that the joint estimation using the BOOMP can significantly improve the channel estimation performance.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (NSFC) under grant 61302097 and the Ph.D. Programs Foundation of the Ministry of Education of China under grant 20120092120014.

29 July 2014
doi: 10.1049/el.2014.2769

Chenhao Qi and Lenan Wu (*School of Information Science and Engineering, Southeast University, Nanjing 210096, People's Republic of China*)

E-mail: qch@seu.edu.cn

References

1 Rusek, F., Persson, D., Lau, B.K., Larsson, E.G., Marzetta, T.L., Edfors, O., and Tufvesson, F.: 'Scaling up MIMO: opportunities and challenges with very large array', *IEEE Signal Process. Mag.*, 2012, **30**, (1), pp. 40–60
2 Gao, Z., Dai, L., and Wang, Z.: 'Structured compressive sensing based superimposed pilot design in downlink large-scale MIMO systems', *Electron. Lett.*, 2014, **50**, (12), pp. 896–898
3 Hou, W., and Lim, C.: 'Structured compressive channel estimation for large-scale MISO-OFDM systems', *IEEE Commun. Lett.*, 2014, **18**, (5), pp. 765–768
4 Barbotin, Y., Hormati, A., Rangan, S., and Vetterli, M.: 'Estimation of sparse MIMO channels with common support', *IEEE Trans. Commun.*, 2012, **60**, (12), pp. 3705–3716
5 Qi, C., and Wu, L.: 'Tree-based backward pilot generation for sparse channel estimation', *Electron. Lett.*, 2012, **48**, (9), pp. 501–503
6 Eldar, Y.C., Kuppinger, P., and Bolcskei, H.: 'Block-sparse signals: uncertainty relations and efficient recovery', *IEEE Trans. Signal Process.*, 2010, **58**, (6), pp. 3042–3054
7 Rebollo-Neira, L., and Lowe, D.: 'Optimized orthogonal matching pursuit approach', *IEEE Signal Process. Lett.*, 2002, **9**, (4), pp. 137–140