Pilot Decontamination in Massive MIMO Systems: Exploiting Channel Sparsity With Pilot Assignment

Zhilin Chen, and Chenyang Yang
School of Electronics and Information Engineering, Beihang University, China
Email: czl@sae.buaa.edu.cn, cyyang@buaa.edu.cn

***Abstract*—This paper deals with the pilot contamination problem for massive MIMO systems. Considering the sparse nature of channel impulse response inherent in wideband systems, the paths of channels of the desired and interference users hardly arrive at the same time, which allows most paths of desired channel to be distinguishable from the paths of interference channels in time-domain. Based on this observation, we first estimate the power delay profile (PDP) of the desired channel with the contaminated channel estimate, from which we acquire the delay of each path of the desired channel. By extracting the corresponding channel components from the contaminated channel estimate, a clean channel estimate can be obtained. To reduce the impact of pilot contamination on the estimated PDP, we propose a pilot assignment method among adjacent cells to randomize the interference over successive uplink frames. Simulation results demonstrate substantial sum rate gain of the proposed approach over existing methods.**

***Index Terms*—Channel estimation, pilot contamination, massive MIMO, sparse channel**

I. Introduction

Multiple-input multiple-output (MIMO) techniques are widely employed in wireless networks. By installing a large number of antennas at a base station (BS), *massive MIMO* can provide high data rate [1], which has become a promising candidate for the fifth-generation (5G) cellular networks [2].

The achievable rate of MIMO systems largely depends on the quality of the instantaneous channel information available at the BS. Considering the overhead to obtain channel information, time-division duplex (TDD) is widely recognized as a proper mode for massive MIMO systems, where the downlink channel is obtained through uplink training by exploiting the channel reciprocity. To reduce the overhead and avoid the coordination among multiple cells, orthogonal pilot sequences for multiple users in each cell are preferred to be reused among adjacent cells. As a result, the channel estimation is severely degraded by the interference in training from the users in neighbor cells with the same pilot sequence [1], [3]. As the number of antennas at each BS goes to infinity, the so-called *pilot contamination* becomes a bottleneck in achieving the promising performance gain of massive MIMO [3], [4].

Recently, significant research efforts have been made to circumvent the pilot contamination problem, and valuable results are obtained. Under the assumption of spatially correlated channels, the authors in [5] found that the pilot contamination vanishes when the minimum mean square error (MMSE) estimation is used if the subspaces of the covariance matrices of the desired and interference channels are non-overlapping. A coordinated pilot assignment method was proposed to meet such a condition. When the channels are assumed independent identically distributed (i.i.d.), the asymptotic orthogonality between the channels of the desired and interference users was exploited to mitigate the pilot contamination [6], [7]. The authors in [6] proposed a blind channel estimation method, where the desired channel was identified from the eigenvectors of the covariance matrix of the received samples with the help of the prior known average channel gains. Unlike [6], the authors in [7] introduced a power control protocol to identify the subspace of the signal of interest from the eigenvalue decomposition (EVD) of the sample covariance matrix.

In this paper, we propose an alternative solution for pilot decontamination. Considering that 5G systems will inevitably be wideband to support high rate, in contrast to previous works [5], [6], [7] that consider the channel feature in space domain, we exploits the channel feature of wideband systems in time domain by taking orthogonal frequency division multiplexing (OFDM) system as an example. Since the multipath channels are usually characterized by a few dominant paths [8], it is almost unlikely that the paths of the desired and interference channels arrive at the same time, which allows the paths of desired and interference channels to be distinguishable. To identify the desired paths from the contaminated channel estimate, we design a pilot assignment method to randomize the pilot contamination over successive uplink training frames, and estimate the power delay profile (PDP), from which the delays of desired paths can be estimated. By extracting the channel components corresponding to the delays from the contaminated channel estimate, an almost interference-free channel estimate can be obtained. Simulations results show that the proposed method is superior to existing methods.

Notations: $\mathbf{Z}[n,m]$ and $\mathbf{z}[n]$ denote the (n,m)th entry of matrix $\mathbf{Z}$ and the nth entry of vector $\mathbf{z}$, and $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote transpose and conjugate transpose, respectively. $|\cdot|$ is the magnitude of a complex variable. $\mathbb{E}\{\cdot\}$ is the expectation.

II. System Model

We consider a downlink TDD OFDM cellular network with B full frequency reused cells. In each cell, a BS equipped with

This work was supported by the National High Technology Research and Development Program of China (No. 2014AA01A705).

978-1-4799-7088-9/14/$31.00 ©2014 IEEE

M antennas serves K single-antenna users on N subcarriers.

We assume block fading channel, which is constant during each frame of uplink training and downlink transmission but may be independent from the one in another frame. The downlink channel for precoding is obtained through uplink training by exploiting the channel reciprocity. To facilitate channel estimation for each user, N_p pilots are inserted in the N subcarriers for uplink training [9]. To eliminate the multiuser interference, the pilots of multiple users within the same cell should be orthogonal. For ease of exposition, we consider time-division orthogonality, where each user occupies one OFDM symbol for training and hence totally K OFDM symbols are required for each cell. We assume that the K OFDM training symbols are reused by all cells, which causes the pilot contamination.

To demonstrate how to exploit channel sparsity to eliminate pilot contamination, we consider spatially uncorrelated frequency selective channel, where only the sparsity in time domain is taken into account. The discrete channel impulse response (CIR) between the mth antenna at the bth BS and the kth user in the jth cell can be expressed as

$$h_{b_m j_k}(n) = \sum_{l=0}^{L_{b_m j_k}-1} \alpha^l_{b_m j_k}\delta(n-\tau^l_{b_m j_k}),\ 0 \le n \le N_\mathrm{h}-1 \quad (1)$$

where $L_{b_m j_k}$ is the number of non-zero resolvable paths, the integer $\tau^l_{b_m j_k}$ is the delay of the lth path, $\alpha^l_{b_m j_k}$ is the corresponding amplitude, and N_h is the length of the CIR, which is assumed less than the length of cyclic prefix (CP) of the OFDM system and $L_{b_m j_k} \le N_\mathrm{h}$.

Denote $\mathbf{h}_{b_m j_k} \triangleq [h_{b_m j_k}(0), \cdots, h_{b_m j_k}(N_\mathrm{h}-1)]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{h}\times 1}$ as the CIR vector, and $\mathbf{a}_{b_m j_k} \in \mathbb{C}^{N_\mathrm{h}\times 1}$ as the power delay profile (PDP) vector of $\mathbf{h}_{b_m j_k}$, whose nth entry is $\mathbf{a}_{b_m j_k}[n] = \mathbb{E}\{|\mathbf{h}_{b_m j_k}[n]|^2\}$ and is time-invariant. For massive MIMO with co-located antennas at each BS, the PDPs seen at different antennas are identical, thus $\mathbf{a}_{b_m j_k}$ can be simplified to $\mathbf{a}_{b j_k}$.

In wideband systems, the multipath channel exhibits a sparse nature [8], i.e., $L_{b_m j_k} \ll N_\mathrm{h}$. In other words, most components of $\mathbf{h}_{b_m j_k}$ and $\mathbf{a}_{b j_k}$ are approximately zero. The number of non-zero paths $L_{b_m j_k}$ depends on the system bandwidth and propagation environments. When the antennas of a BS are mounted at a high place, the channel is usually modeled by several dominant paths, typically seven to nine [10]. An example of sparse channel is illustrated in Fig. 1.

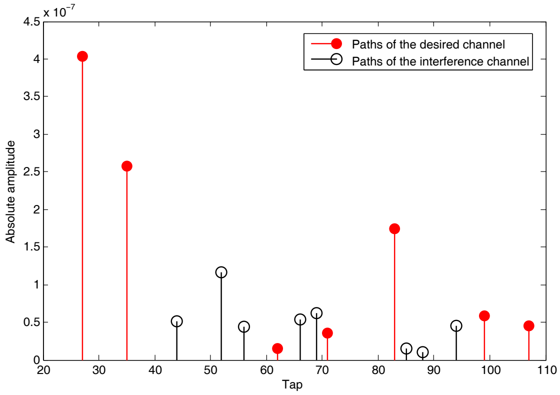

Fig. 1. Sparsity of the multipath channels. The paths of the desired channel are distinguishable from the paths of the interference channel in time domain.

III. Channel Estimation Assisted With Pilot Assignment

In multipath channels with time-domain sparsity, the paths of channels from different users hardly arrives at the same time. This makes the resolvable paths distinguishable in time-domain with high probability. If we can obtain the delays of paths of the desired channels, the desired CIR vector can be estimated without interference. The delays of the desired paths can be obtained from the estimated PDP vector. Since the CIR vectors of the desired user at the M co-located antennas have the same statistics, the PDP vector can be simply estimated from the M CIR vector estimates. However, because these CIR estimates are contaminated by those from interference users, the resulting PDP vector estimate is also contaminated. To identify which paths are from the desired users, we resort to pilot assignment to randomize the interference.

A. CIR Estimation With Pilot Contamination

Consider that both the kth user in the bth cell and the κ_jth user in the jth ($j \neq b$) cell transmit the pilot signals $x_0, \cdots, x_{N_\mathrm{p}-1}$ on the $d_0, \cdots, d_{N_\mathrm{p}-1}$th subcarriers in an OFDM symbol. After dropping the CP and applying N-point fast Fourier transformation, the received pilot signal on the mth antenna at the bth BS can be expressed as

$$\mathbf{y}_{b_m} = \mathbf{X}\mathbf{F}_\mathrm{p}\mathbf{h}_{b_m b_k} + \sum_{j=0,j\neq b}^{B-1} \mathbf{X}\mathbf{F}_\mathrm{p}\mathbf{h}_{b_m j_{\kappa_j}} + \mathbf{w}_{b_m} \quad (2)$$

where $\mathbf{y}_{b_m} \in \mathbb{C}^{N_\mathrm{p}\times 1}$, $\mathbf{X} \in \mathbb{C}^{N_\mathrm{p}\times N_\mathrm{p}}$ is a diagonal matrix with diagonal entries equal to $x_0, \cdots, x_{N_\mathrm{p}-1}$, $\mathbf{F}_\mathrm{p} \in \mathbb{C}^{N_\mathrm{p}\times N_\mathrm{h}}$ is a partial Fourier matrix extracted from the N-point full Fourier matrix with row indices corresponding to $\{d_0, \cdots, d_{N_p-1}\}$ and column indices corresponding to $\{0, \cdots, N_\mathrm{h}-1\}$, i.e., $\mathbf{F}_\mathrm{p}[i,n] = e^{-j2\pi\frac{d_i n}{N}}, 0 \le i \le N_\mathrm{p}-1, 0 \le n \le N_\mathrm{h}-1$, $\mathbf{h}_{b_m b_k}$ is the CIR vector of the desired user seen at the mth antenna at the bth BS, $\mathbf{h}_{b_m j_{\kappa_j}}$ is the CIR vector of the interference user, and $\mathbf{w}_{b_m} \in \mathbb{C}^{N_p\times 1}$ is the additive Gaussian noise vector whose entries are with zero mean and variance σ^2_n.

Various channel estimation methods such as least square (LS) and MMSE can be employed to estimate the CIR vector. We take LS channel estimation as an example, which does not assume known statistical channel information. The CIR vector of the desired user at the mth ($m = 0, \cdots, M-1$) antenna of the bth BS is estimated as

$$\hat{\mathbf{h}}_{b_m b_k} = (\mathbf{X}\mathbf{F}_\mathrm{p})^\dagger \mathbf{y}_{b_m} = \mathbf{h}_{b_m b_k} + \sum_{j=0,j\neq b}^{B-1} \mathbf{h}_{b_m j_{\kappa_j}} + \tilde{\mathbf{w}}_{b_m} \quad (3)$$

where the second term is the pilot contamination, $\tilde{\mathbf{w}}_{b_m} = (\mathbf{X}\mathbf{F}_\mathrm{p})^\dagger \mathbf{w}_{b_m}$, and $(\cdot)^\dagger$ is the pseudo-inverse operation.

B. Multipath Delay Estimation and Refined CIR Estimation

The nth entry of the PDP vector can be estimated as

$$\hat{\mathbf{a}}_{bb_k}[n] = \frac{1}{M}\sum_{m=0}^{M-1} |\hat{\mathbf{h}}_{b_m b_k}[n]|^2$$

$$\overset{(a)}{\approx} \mathbb{E}\{|\hat{\mathbf{h}}_{b_m b_k}[n]|^2\} \tag{4}$$

$$\overset{(b)}{=} \mathbf{a}_{bb_k}[n] + \sum_{j=0,j\neq b}^{B-1} \mathbf{a}_{bj_{\kappa_j}}[n] + \mathbf{v}[n] \tag{5}$$

where step (a) is obtained by using sample average to approximate the ensemble average, step (b) is derived by substituting (3) to (4) and by considering that different CIR vectors are uncorrelated, and $\mathbf{v}[n] \triangleq \sigma_{\mathrm{n}}^2(\mathbf{F}_{\mathrm{p}}^{\mathrm{H}}\mathbf{X}^{\mathrm{H}}\mathbf{X}\mathbf{F}_{\mathrm{p}})^{\dagger}[n,n]$ is the noise term with $(\mathbf{F}_{\mathrm{p}}^{\mathrm{H}}\mathbf{X}^{\mathrm{H}}\mathbf{X}\mathbf{F}_{\mathrm{p}})^{\dagger}[n,n]$ being the (n,n)th entry of $(\mathbf{F}_{\mathrm{p}}^{\mathrm{H}}\mathbf{X}^{\mathrm{H}}\mathbf{X}\mathbf{F}_{\mathrm{p}})^{\dagger}$.

Note that $\mathbf{v}[n]$ is constant that can be easily computed, which can be eliminated from $\hat{\mathbf{a}}_{bb_k}[n]$. Therefore, in the sequel we drop the noise term for notational simplicity.

The *coarse estimate* of the PDP vector in (5) is contaminated by an interference term $\sum_{j=0,j\neq b}^{B-1} \mathbf{a}_{bj_{\kappa_j}}[n]$. To identify which components in $\hat{\mathbf{a}}_{bb_k}$ correspond to the desired channel, we exploit the stationarity of the PDP vectors over successive frames of uplink training and downlink transmission.

Consider T successive frames, during each frame the CIR estimates for users in the bth cell are obtained with (3), and then the *coarse estimates* of the PDP vectors of each of the desired users are obtained with (5). In the tth frame, the *coarse estimation* of the nth entry of the PDP vector for the channel from the kth user in the bth cell is

$$\hat{\mathbf{a}}_{bb_k}^t[n] = \mathbf{a}_{bb_k}[n] + \sum_{j=0,j\neq b}^{B-1} \mathbf{a}_{bj_{\kappa_j^t}}[n] \tag{6}$$

where $\mathbf{a}_{bj_{\kappa_j^t}}$ is the PDP vector from the interference user κ_j^t in the jth cell during the tth frame.

To remove the interference term in (6), we can randomize the interference users in successive frames, which can be accomplished by pilot assignment designed in next subsection.

After all T estimates of the PDP vector with randomized contamination are available, we can obtain a *refined estimate* by taking their geometric average to reduce the interference, whose nth entry is given as follows

$$\bar{\mathbf{a}}_{bb_k}[n] = (\hat{\mathbf{a}}_{bb_k}^0[n] \times \hat{\mathbf{a}}_{bb_k}^1[n] \times \cdots \times \hat{\mathbf{a}}_{bb_k}^{T-1}[n])^{\frac{1}{T}} \tag{7}$$

The following proposition indicates that $\bar{\mathbf{a}}_{bb_k}$ is an accurate estimate of the desired PDP vector by judiciously designing the pilot assignment.

Prop. 1: If the nth entry of the PDP vector corresponds to the delay of a desired path, i.e., $\mathbf{a}_{bb_k}[n] \neq 0$, then $\bar{\mathbf{a}}_{bb_k}[n] \geq \mathbf{a}_{bb_k}[n]$. If the nth entry does not correspond to the delay of a desired path, i.e., $\mathbf{a}_{bb_k}[n] = 0$, then $\bar{\mathbf{a}}_{bb_k}[n]$ converges to zero with probability one as T increases, if the contamination terms $\sum_{j=0,j\neq b}^{B-1} \mathbf{a}_{bj_{\kappa_j^t}}[n], t = 0,\cdots,T-1$ are uncorrelated.

Proof: When $\mathbf{a}_{bb_k}[n] \neq 0$, by substituting (6) to (7) it is easy to see that the first part of the proposition is obtained since the contamination terms are non-negative. When $\mathbf{a}_{bb_k}[n] = 0$, $\hat{\mathbf{a}}_{bb_k}^t[n]$ in (6) may also equal to zero if no interference paths arrive at the delay of n. Denote $q_t \triangleq \mathrm{Prob}(\hat{\mathbf{a}}_{bb_k}^t[n] = 0) < 1$. If the contamination terms are uncorrelated, we have $\mathrm{Prob}(\bar{\mathbf{a}}_{bb_k}[n] = 0) = 1-(1-q_0)\times\cdots\times(1-q_{T-1})$, which approaches one as T increases. Then, the second part of the proposition is proved. ■

The proposition suggests that we can decide whether index n corresponds to a delay of the desired path based on the value of each component in $\bar{\mathbf{a}}_{bb_k}$. If $\bar{\mathbf{a}}_{bb_k}[n]$ is non-zero, n corresponds to a delay of path of the desired channel.

In practice, the *refined estimate* of the PDP vector can be obtained via uplink training iteratively based on current CIR estimate and the PDP estimates in previous frames. The performance of the delay estimation based on the *refined estimate* of PDP is improved after each iteration.

With the estimated delays of the paths, we can obtain a refined estimate of the CIR vector for the desired user by selecting the corresponding channel components from the contaminated CIR estimate in (3). Denote $\mathcal{L}$ as the index set of the estimated delays. Then, the nth entry of the desired CIR vector can be expressed as

$$\hat{\mathbf{g}}_{b_m b_k}[n] = \begin{cases} \hat{\mathbf{h}}_{b_m b_k}[n], & n \in \mathcal{L} \\ 0, & n \notin \mathcal{L} \end{cases} \tag{8}$$

Then, $\hat{\mathbf{g}}_{b_m b_k} \in \mathbb{C}^{N_{\mathrm{h}}\times 1}, n = 0,1,\cdots,N_{\mathrm{h}}-1$, is the finally obtained refined estimate of the CIR vector.

C. Pilot Assignment Strategy

In what follows, we propose a pilot assignment strategy to decorrelate the contamination terms in the *coarse estimates* of the PDP vectors during different frames.

From (6), we can observe that the contamination in the tth frame depends on the PDP vectors of the $B-1$ interference users $\kappa_j^t, 0 \leq j \leq B-1, j \neq b$. Therefore, if we can make the interference users in all T frames totally different, i.e.,

$$\kappa_j^{t_1} \neq \kappa_j^{t_2}, 0 \leq t_1, t_2 \leq T-1, t_1 \neq t_2, \tag{9}$$

for $0 \leq j \leq B-1, j \neq b$, then the contamination terms in all T frames will be uncorrelated.

A simple way to do this is using circular shifting pilot assignment. Based on the assumption that the pilot signals for the users in the same cell is time-division orthogonal, the pilot assignment is equivalent to the training symbol assignment. Suppose that each cell employs K OFDM symbols for the training of the K users. Denote $\mathbf{A}_b, b = 0,\cdots,B-1$ as the $T \times K$ pilot assignment matrices for all B cells. Denote its (t,k)th element $\mathbf{A}_b[t,k]$ as the index of the OFDM training symbol that assigned to the kth user in the bth cell during the tth frame, whose value ranges from 0 to $K-1$.

We generate $\mathbf{A}_b$ by circularly shifting of the preceding row vectors: the first row vector of $\mathbf{A}_b$ is given as $[0,1,\cdots,K-1]$, and the rth row vector is rotated b elements to the left relative to the $(r-1)$th row vector. An example of $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2$ with $B = 3, K = 3, T = 3$ are respectively given as

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix} \tag{10}$$

The following proposition shows that such a pilot assignment strategy satisfies (9) under mild conditions. Owing to the space limitation, we only give a sketch of the proof.

Prop. 2: If K is a prime number with $K \geq B$ and $K \geq T$, the indices of interference users $\kappa_j^0, \cdots, \kappa_j^{T-1}$ are totally different.

Proof: Based on the assignment $\mathbf{A}_0, \cdots, \mathbf{A}_{B-1}$, after regular manipulations we can show that the index of the interference user in the tth frame is,

$$\kappa_j^t = (k - t(j-b))(\bmod K) \tag{11}$$

where $(\cdot)(\bmod K)$ denotes the remainder of $(\cdot)$ divided by K. For any pair of frames $t_1 \neq t_2, 0 \leq t_1, t_2 \leq T-1$, we have

$$\kappa_j^{t_1} - \kappa_j^{t_2} = ((t_2 - t_1)(j-b))(\bmod K) \tag{12}$$

Note that $0 < |t_2 - t_1| < T$ and $0 < |j - b| < B$. When $T, B < K$ and K is a prime number, $|(t_2 - t_1)(j - b)|$ is indivisible by K. Thus the right hand side of (12) is non-zero, then $\kappa_j^{t_1} \neq \kappa_j^{t_2}$ holds. ■

The condition $K \geq T$ in the proposition implies that the proposed CIR estimate strategy should converge rapidly. Another condition $K \geq B$ is easy to be satisfied in practice for massive MIMO systems. For the third condition, when K is not a prime number, we can chose a prime number $K^{'}$ slightly larger than K and generate the $T \times K^{'}$ assignment matrices. Then, we choose the first K columns to form the final assignment matrices. Note that such a pilot assignment strategy does not need the coordination among multiple cells.

IV. Simulation Results

In this section, we evaluate the performance of the proposed method of channel estimation and pilot assignment by comparing with existing methods with simulations.

We consider a network of seven cells, each consists of one BS to serve 10 uniformly located single-antenna users. To remove the boundary effect, wrap around is considered. The radius of each cell is 250 m, and the path-loss is modeled as $\mathrm{PL}^{\mathrm{dB}} = 35.3 + 37.6 \log_{10}(d)$, where d (in meter) is the distance between the BS and the user. The transmit powers of each user and each BS are 23 dBm and 46 dBm, respectively. The OFDM parameters are set according to the LTE systems with 20 MHz bandwidth [9]. We consider the extended vehicular A (EVA) channel model [10], which has a maximum delay of 2510 ns with nine non-zero paths. The zero-forcing (ZF) beamforming with equal power allocation is employed for downlink transmission. The simulation results are averaged over 500 random drops of users, where in different drops the small scale channels are independently generated.

In Fig. 2, we show the convergence of the refined CIR estimation with the pilot assignment during successive frames. We can see that after three or four frames for training, the per-cell sum rate achieved by the proposed method can approach that of the perfect CIR vector, where the gap is mainly caused by the residual contamination on the estimated paths of the desired CIR vector. The proposed CIR estimation method is superior to the contaminated LS estimate in (3) (with legend "LS method") even without any iteration, since the noise exists in (3) is reduced after the CIR construction in (8).

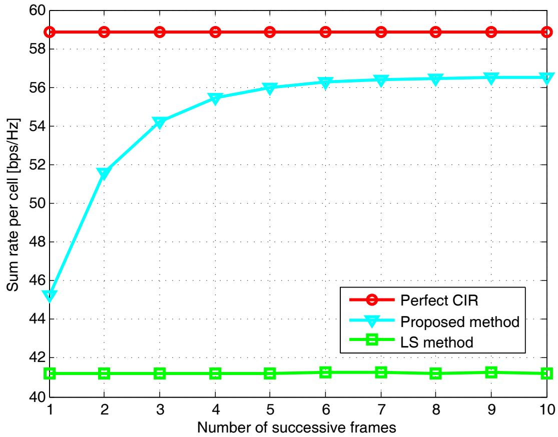

Fig. 2. Sum rate per-cell vs. the number of frames, $M = 100$.

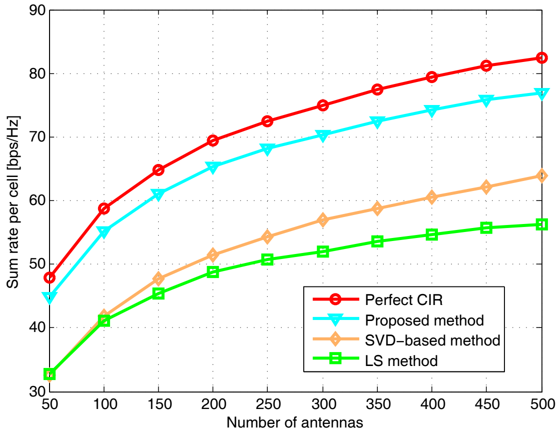

Fig. 3. Sum rate per-cell vs. the number of antennas at each BS, $T = 4$.

In Fig. 3, we compare the sum rate of the proposed method with the SVD-based method [7], where the performance under perfect CIR and that achieved by the LS estimate with pilot contamination are also shown. We can see that our method is superior in all cases. By contrast, the SVD-based method only performs well with large number of antennas.

V. Conclusions

In this paper, we resolved the pilot contamination problem for massive MIMO systems by exploiting the inherent sparsity nature in time domain of the channels of wideband systems. To distinguish the channel impulse response of a desired user from the channels of the interference users, we proposed to estimate the delays of paths after randomize the interference by judiciously assigning pilots in successive uplink frames. Simulation results showed that the proposed scheme converges rapidly and is superior to existing methods.

REFERENCES

[1] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

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

[3] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[4] J. Hoydis, S. T. Brink, and M. Debbah, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?” *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

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

[6] H. Q. Ngo and E. G. Larsson, “EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays,” in *Proc. IEEE ICASSP*, 2012.

[7] R. Müller, L. Cottatellucci, and M. Vehkaperä, “Blind pilot decontamination,” *IEEE J. Select. Topics Signal Process.*, early access.

[8] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, “Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling,” *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–479, Mar. 2001.

[9] 3GPP TS 36.211, “Physical channels and modulation (Release 10),” 3rd Generation Partnership Project (3GPP), Tech. Rep., 2010.

[10] 3GPP TS 36.104, “Base Station (BS) radio transmission and reception (Release 10),” 3rd Generation Partnership Project (3GPP), Tech. Rep., Mar. 2012.