This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

Sparse Channel Estimation for MIMO-OFDM Systems in High-Mobility Situations

Xu Ma, *Student Member, IEEE*, Fang Yang, *Senior Member, IEEE*, Sicong Liu, *Member, IEEE*, Jian Song, *Fellow, IEEE*, and Zhu Han, *Fellow, IEEE*

***Abstract*—A new approach of channel estimation for multi-antenna systems is put forward in this article, which can be adopted in high-mobility situations such as high speed trains. The channel impulse response (CIR) is abstracted as three domains to improve the modeling accuracy. Both the time-domain preamble and the frequency-domain pilot are adopted in the OFDM frame. Firstly, the training in time domain is exploited to obtain the partial common support (PCS) of the channel. Then, the pilot location optimized by the genetic algorithm is employed to build the framework of structured compressive sensing (SCS) and recover the channel. A novel compressive recovery algorithm called adaptive support-aware block orthogonal matching pursuit (ASA-BOMP) for MIMO-OFDM systems is proposed to solve the problem. It is manifested in the simulation that the scheme in this article outperforms the traditional ones in both recovery probability, mean square error (MSE), and bit error rate (BER) over the doubly selective channel with low computational complexity.**

***Index Terms*—MIMO-OFDM, channel estimation, doubly selective channel, structured compressive sensing (SCS).**

I. Introduction

As the communication requirement is growing by leaps and bounds, multiple input multiple output (MIMO), which benefits from the property of large capacity, has exerted tremendous fascination to both academia and industry. Simultaneously, the system capacity can also be increased by adopting the technique of orthogonal frequency division multiplexing (OFDM). For these reasons, the afore-mentioned two techniques are usually combined as MIMO-OFDM, which has advantages to the performance of communications. MIMO-OFDM is not only one of the most paramount techniques currently and a potential technique aftertime.

To maintain the system performance, it is very important to perform the channel estimation with high precision. However, accurate channel estimation for the MIMO-OFDM system is challenging due to the multiple antennas. For every receive antenna in a MIMO-OFDM system, a great deal of channel information has to be recovered from all of the transmit ends. Consequently, the time and frequency domain overhead is ordinarily much larger than that in a single antenna system. The training usually includes the time-domain preamble [1] or frequency-domain pilot [2]. Most of the traditional schemes employ orthogonal training with low capacity, whose number raises with the transmit antennas linearly. Apart from that, a high spectral efficiency channel estimation scheme was proposed [3], whose training is non-orthogonal.

Nevertheless, channel estimation schemes in high-mobility situations such as high speed trains [4], [5] are not deeply investigated in most of the conventional works and it is a new field. Apart from the frequency selectivity caused by the multipath effect, the channel demonstrates the time selectivity as well when the mobile terminal is moving at high speed. The channel estimation in this kind of complex channel, which is known as doubly selective channel, is quite arduous, particularly for MIMO-OFDM. A multitude of communication systems may suffer from the doubly selective channel, such as long-term evolution rail [6] for high speed train and the railroad cognitive radio [7], etc.

In this condition, the channel distinguishes from each other among different samples. Therefore, the channel coefficients to be recovered is exceedingly large in amount. Under these circumstances, lots of time or frequency training are needed to perform channel estimation compared with the time-invariant channel, which will lead to the spectral efficiency decrease. When we perform the doubly selective channel estimation for single input single output (SISO) system, literature [8] considered the channel variations as a difference model, while literature [9] abstracted the channel in sample domain by utilizing a basis expansion model (BEM).

As the study of compressive sensing (CS) is becoming a research hotspot [10]–[13], a new perspective for channel estimation is to take advantage of the channel sparsity [14]–[22]. Literatures [14]–[17] utilized CS theory to perform the channel estimation in SISO systems. Although the time and frequency training are considered in these literatures, the

Xu Ma, Fang Yang, and Jian Song are with the Electronic Engineering Department & Research Institute of Information Technology, Tsinghua University, Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing 100084, P. R. China. Email: maxu14@mails.tsinghua.edu.cn.

Fang Yang and Jian Song are also with Key Laboratory of Digital TV System of Guangdong Province and Shenzhen City, Research Institute of Tsinghua University in Shenzhen, Shenzhen 518057, China. Email: {fangyang, jsong}@tsinghua.edu.cn.

Sicong Liu is with the Department of Communications Engineering, School of Information Science and Technology, Xiamen University, Xiamen, Fujian Province, P.R. China. Email: liusc1028@gmail.com.

Zhu Han is with the University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul, South Korea. Email: zhan2@uh.edu.

This work was supported in part by Science, Technology and Innovation Commission of Shenzhen Municipality (Grant No. JCYJ20160331171115521), National Key Research and Development Program of China (YS2017YFGH000376), Guangdong Key Laboratory Project (2017B030314147), Natural Science Foundation of Guangdong Province (2015A030312006), Young Elite Scientist Sponsorship Program by CAST (YESS20150120), and the US NSF CNS-1717454, CNS-1731424, CNS-1702850, CNS-1646607, ECCS-1547201, CMMI-1434789, CNS-1443917, and ECCS-1405121. (Corresponding Author: Fang Yang)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

process can only be performed under time-invariant channel for SISO systems. Doubly selective channel estimation problems in MIMO systems cannot be solved directly using these methods. For multiple antenna approaches, multiple input single output (MISO) channel estimation scheme based on the non-orthogonal pilots in frequency domain is investigated in [18]. Both the pilots location and power were designed by minimizing the coherence of the associated Fourier submatrix. Two different relaxations were proposed to solve the non-convex problem, which was the first fully deterministic pilot design in a MISO/multi-user scenario with shared pilot subcarriers [18]. Literatures [19]–[21] considered the MIMO-OFDM channel estimation problem, which extended the framework of CS to the structured CS (SCS). The spatial correlation among different transmit antennas are used, which indicates a set of common nonzero support in the SCS framework. Nevertheless, only frequency selective channel is addressed. In [22] and [23], the doubly selective channel estimation was performed using the SCS model, while a differential simultaneous orthogonal matching pursuit (DSOMP) [24] algorithm and structured distributed compressive sensing (SDCS) [25] were further proposed. A channel estimation scheme within multiple symbols was provided in [26]. However, the pilot pattern and the frame structure were not fully studied in these schemes which can be further optimized. Literature [27] utilized the atomic norm minimization to perform doubly selective channel estimation in SISO systems, while literature [28] adopted the block-sparse Bayesian learning (BSBL) to solve SISO doubly selective channel estimation problem. Literature [29] used the measurements outside the inter block interference (IBI) free region to solve the channel estimation problem, however the frequency-domain pilots were not considered. In general, most existing works conduct non-comprehensive study on doubly selective channel estimation for MIMO-OFDM system. Therefore, the relevant research work is highly significant.

In this article, a new approach for channel estimation is put forward in high-mobility situations. Both the time- and frequency-domain training (TFDT) are adopted in the MIMO-OFDM system. The channel is abstracted as three domains. Firstly, a set of block sparse vectors are acquired for the tap-based CIRs due to the identical supports of the channel impulse responses (CIRs) for different samples. Secondly, the sample-based CIRs can be expanded on a few bases. Thirdly, the antenna-based CIRs are uniformly zeros or nonzeros. The time-domain pseudo-random noise (PN) sequence provides the information of partial common support (PCS), while the frequency-domain pilot location optimized by a genetic algorithm (GA) is put forward to obtain the SCS framework by taking advantage of the spatial correlation. After cyclic reconstruction of the OFDM frame, the channel information can be acquired using the adaptive support-aware block orthogonal matching pursuit (ASA-BOMP) algorithm. It can be validated in the simulation that our scheme has superior performance than the traditional ones with lower complexity.

The article is organized as follows. The models of MIMO-TFDT-OFDM and the doubly selective channel are presented in Section II. The approach for channel estimation put forward in this article is investigated in Section III. The proposed pilot location design method and sparse signal recovery algorithm for MIMO-OFDM systems are introduced in Section IV and V, respectively. Simulation results and discussion are addressed in Section VI. Finally, the article is concluded in Section VII.

Notation: In this article, lower and upper boldface letters denote column vectors and matrices, respectively. $(\cdot)^T$, $(\cdot)^H$, $\otimes$, $\mathrm{E}(\cdot)$, $\|\cdot\|_2$, $\mathrm{mod}()$, and $\mathrm{diag}(\cdot)$ denote transpose, conjugate transpose, the circular convolution, the expectation of a random variable, the ℓ_2-norm of a vector, the modulo operation, and changing a vector into a diagonal matrix, respectively. $[\cdot]_{i,j}$ and $[\cdot]_{\mathbf{p},\mathbf{q}}$ denote the (i,j)-th element of a matrix, and a submatrix with row indices $\mathbf{p}$ and column indices $\mathbf{q}$, respectively. $\mathbf{I}_N^{\langle k \rangle}$ is an $N \times N$ permutation matrix, whose elements in i-th row and $(i+k)$-th column are filled with 1 $(0 \le i < N)$ and others are all 0s.

II. System Model

A. Frame Structure for MIMO-TFDT-OFDM Systems

The structure for one OFDM frame in time domain is denoted as $\boldsymbol{t} = [\boldsymbol{y}^T, \boldsymbol{x}^T]^T$. It consists of the data block $\boldsymbol{x}$ and the guard interval (GI) $\boldsymbol{y}$. The main purpose of the GI is to prevent the data block from being interfered by the previous frame. In this article, the GI, which is filled with the preamble, has another crucial function of channel estimation. In the scenario of MIMO, OFDM frames from variant transmit antennas can be denoted by using an extra subscript, e.g.,

$$\boldsymbol{t}_i = [\boldsymbol{y}_i^T, \boldsymbol{x}_i^T]^T, \tag{1}$$

where $\boldsymbol{t}_i$ is the OFDM frame from the i-th transmit antenna and $0 \le i < N_t$. The corresponding length-M preamble and length-N data block can be represented as $\boldsymbol{y}_i = [y_{i,0}, y_{i,1}, \ldots, y_{i,M-1}]^T$, and $\boldsymbol{x}_i = [x_{i,0}, x_{i,1}, \ldots, x_{i,N-1}]^T$, respectively.

It is shown in Fig. 1 that both the pilots and PN sequence are configured in each TFDT-OFDM frame. The PN sequence $\mathbf{c}_i$ can be regarded as one category of the time-domain preamble, which is the same for different frames. On the one hand, the PN sequence is utilized to obtain PCS due to its superior characteristic of auto-correlation. On the other hand, a small quantity of frequency subcarriers are padded with pilots, which is used for obtaining accurate information of the channel subsequently.

It can be seen in Fig. 1 that the frequency-domain pilots are divided into G pilot groups. There are $2Q-1$ subcarriers for each pilot group. The Q subcarriers in the middle of the group are configured with nonzero pilots, while the rest $Q-1$ are configured with zero pilots in order to avoid the inter carrier interference (ICI). Q is set as an odd number in this article. The position of the pilot groups are designed and optimized by a GA for the accuracy of sparse signal recovery, whose simulation results will be illustrated in Section VI. The nonzero pilots with identical index from all of the pilot groups are combined as a vector $\mathbf{p}_q \in \mathbb{C}^G$ for $0 \le q < Q$.

The pilot group above is non-orthogonal because the pilot subcarriers are identical for all of the antennas. To distinguish different antennas, the amplitude of nonzero pilots is random with fixed phase in this article. Under these circumstances,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

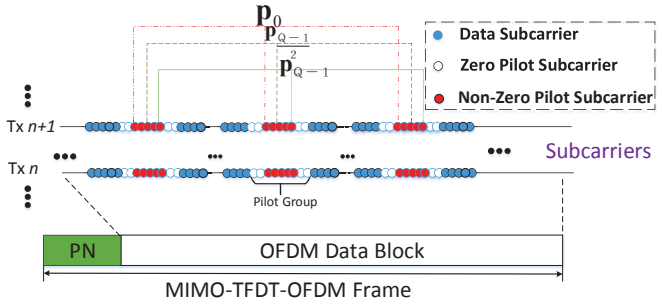


Fig. 1. The MIMO-TFDT-OFDM frame with the PN sequence and pilots.

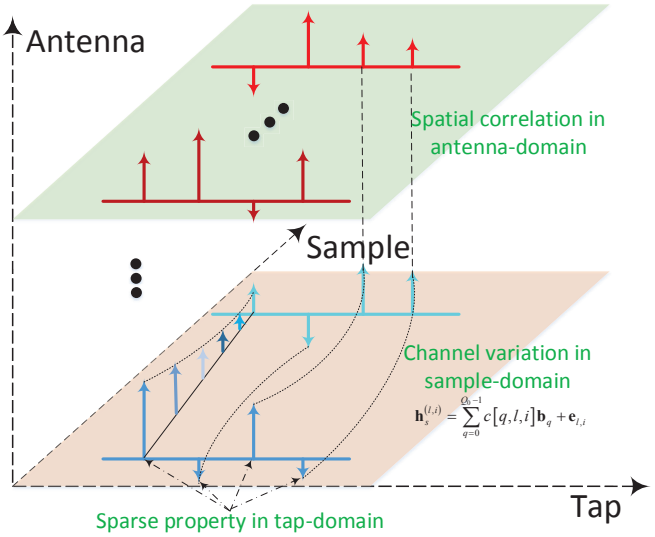


Fig. 2. The channel modeling in three domains.

the nonzero pilots are the measurement when performing the sparse signal recovery. Although the pilot powers in the nonzero pilot subcarriers are different, the average power for one nonzero pilot group will be normalized, which is a constraint to the pilot power. In comparison with the separate nonzero pilot scheme [22], which concentrates the power to the center pilot, our pilot pattern has an average distribution for the pilots. Consequently, the frequency-domain noise can be suppressed while the local signal-to-noise ratio (SNR) can increase visibly. In this way, the proposed scheme will have an outstanding performance.

B. The Channel Model for Three Domains

As shown in Fig. 2, the CIR in this article is characterized to be sparse for MIMO-OFDM. The channel is abstracted as three domains, i.e., sample, tap, and antenna, respectively. The sets of arrows are CIR gains for the channels. Each arrow represents a nonzero element for the CIR. Consequently, the figure illustrates a set of four-tap channels for a MIMO system.

In this article, the indexes of the receive antennas are left out on account of the same signal processing for every receive antenna. In this condition, the l-th channel tap at sample m and the i-th transmit antenna is represented as $h[m,l,i]$, where $0 \le l < L$ and L is the length of the channel. The tap-based CIR at sample m for the i-th antenna is represented as

$$\mathbf{h}_{\mathrm{t}}^{(m,i)} = [h[m,0,i], h[m,1,i], \cdots, h[m,L-1,i]]^T. \quad (2)$$

$\mathbf{h}_{\mathrm{t}}^{(m,i)}$ is assumed to be sparse with S nonzero elements since the channel length is usually much larger than the number of distinguishable taps [30]. S is commonly referred to as the sparsity level. The nonzero supports for $\mathbf{h}_{\mathrm{t}}^{(m,i)}$ are deemed to be the same because the channel changes slowly during an OFDM frame length [31]. Moreover, the channels among different antennas tend to have a property of spatial correlation [3], [30], [32] because the time delays for all transmit-receive antennas are approximately identical [33], [34]. The spatial correlation in this paper is defined as that there are uniform nonzero supports, and distinct nonzero amplitudes and phases for the *tap-based* CIRs. Accordingly, a jointly sparse property is shown for the *tap-based* CIR vectors $\mathbf{h}_{\mathrm{t}}^{(m,i)}(0 \le m < N, 0 \le i < N_t)$.

The *sample-based* CIR is written as

$$\mathbf{h}_{\mathrm{s}}^{(l,i)} = [h[0,l,i], h[1,l,i], \cdots, h[N-1,l,i]]^T. \quad (3)$$

It is composed of the l-th delays from every tap-based CIR over the OFDM frame for the i-th transmit antenna. As the the channel changes slowly during the OFDM frame length, $\mathbf{h}_{\mathrm{s}}^{(l,i)}(0 \le l < L)$ can be expanded by $Q_0(Q_0 \ll L)$ bases [9],

$$\mathbf{h}_{\mathrm{s}}^{(l,i)} = \sum_{q=0}^{Q_0-1} c[q,l,i]\mathbf{b}_q + \mathbf{e}_{l,i}, \quad (4)$$

where $c[q,l,i]$, $\mathbf{b}_q \in \mathbb{C}^N$, and $\mathbf{e}_{l,i} \in \mathbb{C}^N$ denote the q-th BEM coefficient, the corresponding BEM basis for the l-th channel tap and the i-th transmit antenna, and the modeling error, respectively. Utilizing the BEM, the number of the estimated coefficients is reduced from NSN_t to Q_0SN_t. In this article, the number of BEM bases and nonzero subcarriers are the same, i.e., $Q_0 = Q$.

The BEM vector is defined as

$$\mathbf{c}_{q,i} = [c[q,0,i], c[q,1,i], \ldots, c[q,L-1,i]]^T, \quad (5)$$

where $0 \le q < Q$. Based on the description in (4) and (5), $\mathbf{c}_{q,i}$ are jointly sparse for different q and i if the tap-based CIR vectors $\mathbf{h}_{\mathrm{t}}^{(m,i)}$ are jointly sparse. Moreover, the nonzero supports for $\mathbf{c}_{q,i}$ and $\mathbf{h}_{\mathrm{t}}^{(m,i)}$ are the same.

In this article, the complex exponential BEM (CE-BEM) [9] is adopted to reconstruct the sample-based CIR, which can be written as as

$$\begin{aligned}
\mathbf{b}_0^{\mathrm{CE}} &= (1, \cdots, e^{j\frac{2\pi}{N}n(-\frac{Q-1}{2})}, \cdots, e^{j\frac{2\pi}{N}(N-1)(-\frac{Q-1}{2})})^T \\
\mathbf{b}_1^{\mathrm{CE}} &= (1, \cdots, e^{j\frac{2\pi}{N}n(1-\frac{Q-1}{2})}, \cdots, e^{j\frac{2\pi}{N}(N-1)(1-\frac{Q-1}{2})})^T \\
&\vdots \\
\mathbf{b}_{Q-1}^{\mathrm{CE}} &= (1, \cdots, e^{j\frac{2\pi}{N}n(Q-1-\frac{Q-1}{2})}, \cdots, e^{j\frac{2\pi}{N}(N-1)(Q-1-\frac{Q-1}{2})})^T
\end{aligned} \quad (6)$$

The *antenna-based* CIR is denoted as

$$\mathbf{h}_{\mathrm{a}}^{(m,l)} = [h[m,l,0], h[m,l,1], \cdots, h[m,l,N_t-1]]^T, \quad (7)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

whose entries are simultaneously zeros or nonzeros due to the assumption of spatial correlation among the antennas. The representations in (2), (3), and (7) are the channel models for three domains in this article.

III. The Proposed Scheme

A. Step 1: Acquisition of PCS

Although the time-variant channel is considered in this article, the nonzero supports among different samples in one OFDM frame tend to be identical. In our scheme, the PCS Ω_0 of the sparse tap-based CIR is firstly obtained by adopting the PN sequence. As the PN sequence has superior auto-correlation characteristic, Ω_0 can be obtained by performing the correlation between the local PN sequence $\mathbf{y}$ and the received PN sequence $\mathbf{y}_r$, which can be written as

$$\boldsymbol{\alpha} = \frac{1}{M}\mathbf{y} \otimes \mathbf{y}_r, \tag{8}$$

where $\boldsymbol{\alpha} \in \mathbb{C}^M$ represents the coarse estimation of the channel. Although the CIR is time-variant, $\boldsymbol{\alpha}$ will imply the information of the PCS of the channel.

Ω_0 is acquired by means of choosing the entries which are larger than the threshold value θ, e.g., $\Omega_0 = \{l : \|\alpha[l] \geq \theta\|\}_{l=0}^{L-1}$. The value of θ can be configured as $\theta = 3(\sum_{l=0}^{L-1} |\alpha[l]|^2)^{1/2}/L$ [30]. After Ω_0 is obtained, it will be utilized to perform accurate channel estimation afterwards.

B. Step 2: Frame Structure after Reconstruction

The OFDM data block in this article is interfered by the previous PN sequence. Hence, the process of frame reconstruction is indispensable to eliminate the contamination for the channel recovery.

Fig.3 demonstrates the frame reconstruction process. The idea is according to the method of overlapping and adding (OLA) [35]. It can be implemented by extending the OLA operations to every transmit antenna. The differences include the following three aspects. Firstly, the influence from every antenna must be considered in MIMO-OFDM scenario. Secondly, on account of the time-invariant channel, the IBI caused by the PN sequence in *Step 2.1* is not a simple linear convolution between CIR and PN sequence, but a complex effect considering the variant channel. Finally, in *Step 2.3*, the first part of the convolution between the PN and CIR should also take the time-variant channel into consideration.

Note that our scheme is derived from a MISO system. However, when the multiple receive antennas in a MIMO system perform signal processing as the single antenna does in a MISO system, the so-called MISO system is equivalent to and can be easily extended to the MIMO system. In this way, the time-domain OFDM frame after performing frame reconstruction is denoted as [8]

$$\mathbf{r}_\Sigma = \sum_{i=0}^{N_t-1} \mathbf{r}_i = \sum_{i=0}^{N_t-1} \mathbf{H}_i\mathbf{t}_i + \mathbf{w}_t, \tag{9}$$

where $\mathbf{r}_i, \mathbf{t}_i \in \mathbb{C}^N$ are the receive and transmit time-domain OFDM frames from the i-th transmit antenna, respectively. $\mathbf{r}_\Sigma$

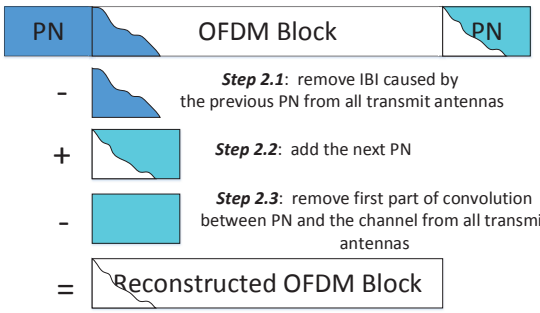


Fig. 3. Block cyclic reconstruction for the TFDT-OFDM system.

describes the receive frame superimposed by $\mathbf{r}_i (0 \leq i < N_t)$, while $\mathbf{w}_t \in \mathbb{C}^N$ represents the i.i.d. additive white Gaussian noise (AWGN). The matrix $\mathbf{H}_i \in \mathbb{C}^{N\times N}$ is made up of the entries $[\mathbf{H}_i]_{k,j} = h[k, \text{mod}(k-j, N)]$.

As the signal processing for different receive antennas are the same, a single receive antenna is considered in the following deduction. The i in the subscript is left out to facilitate reading. In this condition, the frequency-domain received frame is represented by

$$\widetilde{\mathbf{r}} = \mathbf{F}_N(\mathbf{H}(\mathbf{F}_N^H\widetilde{\mathbf{t}}) + \mathbf{w}_t) = \mathbf{H}_F\widetilde{\mathbf{t}} + \mathbf{w}_f, \tag{10}$$

where $\widetilde{\mathbf{t}}$ and $\widetilde{\mathbf{r}}$ are the transmit and received frames in the frequency domain, respectively. $\mathbf{F}_N \in \mathbb{C}^{N\times N}$ is the discrete Fourier transform matrix, $\mathbf{H}_F = \mathbf{F}_N\mathbf{H}\mathbf{F}_N^H$ is the frequency-domain channel matrix, and $\mathbf{w}_f = \mathbf{F}_N\mathbf{w}_t$ is the noise.

As the channel matrix $\mathbf{H}_F$ is related with the CE-BEM bases which has a series of good property like orthogonality, $\mathbf{H}_F$ is rewritten as [9]

$$\mathbf{H}_F = \sum_{k=0}^{Q-1} \mathbf{B}_k^{\text{CE}}\mathbf{C}_k, \tag{11}$$

where $\mathbf{B}_k^{\text{CE}} = \sqrt{N}\mathbf{F}_N\text{diag}(\mathbf{b}_k^{\text{CE}})\mathbf{F}_N^H$ and $\mathbf{C}_k = \text{diag}(\mathbf{F}_N(\mathbf{c}_{k,i}^T, \mathbf{0}_{1\times(N-L)})^T)$. Consequently, by taking (11) into (10), we can get

$$\widetilde{\mathbf{r}} = \sum_{k=0}^{Q-1} \mathbf{B}_k^{\text{CE}}\mathbf{C}_k\widetilde{\mathbf{t}} + \mathbf{w}_f. \tag{12}$$

C. Step 3: Accurate Doubly Selective Channel Estimation

For one thing, based on the characteristics of CE-BEM bases, $\mathbf{B}_k^{\text{CE}}$ is rewritten as [9]

$$\mathbf{B}_k^{\text{CE}} = \sqrt{N}\mathbf{F}_N\text{diag}(\mathbf{b}_k^{\text{CE}})\mathbf{F}_N^H = \sqrt{N}\mathbf{I}_N^{\langle k-\frac{Q-1}{2}\rangle}, \tag{13}$$

For another, $\mathbf{C}_i\widetilde{\mathbf{t}}$ is transformed to

$$\mathbf{C}_k\widetilde{\mathbf{t}} = \text{diag}(\mathbf{F}_N(\mathbf{c}_{k,i}^T, \mathbf{0}_{(N-L)\times 1})^T)\widetilde{\mathbf{t}} = \frac{1}{\sqrt{N}}\text{diag}(\widetilde{\mathbf{t}})\mathbf{F}'_N\mathbf{c}_{k,i}, \tag{14}$$

where $\mathbf{F}'_N = [\mathbf{F}_N]_{0:N-1,0:L-1}$. Hence, $\widetilde{\mathbf{r}}$ is denoted by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

$$\widetilde{\mathbf{r}} = \sum_{k=0}^{Q-1} \mathbf{I}_N^{\langle k-\frac{Q-1}{2}\rangle} \mathrm{diag}(\widetilde{\mathbf{t}}) \mathbf{F}'_N \mathbf{c}_{k,i} + \mathbf{w}_f. \tag{15}$$

The pilots with identical indices $\mathbf{p}_q (0 \le q < Q)$ are obtained by left multiplying $\mathbf{\Psi}_q = [\mathbf{I}_N]_{\mathbf{p}_q,0:N-1} \in \mathbb{C}^{G\times N}$, which is expressed as

$$\widetilde{\mathbf{r}}_{\mathbf{p}_q} = \mathbf{\Psi}_q \widetilde{\mathbf{r}} = \sum_{k=0}^{Q-1} \mathbf{\Psi}_q \mathbf{I}_N^{\langle k-\frac{Q-1}{2}\rangle} \mathrm{diag}(\widetilde{\mathbf{t}}) \mathbf{F}'_N \mathbf{c}_{k,i} + \mathbf{w}', \tag{16}$$

where $\mathbf{w}' = \mathbf{\Psi}_q \mathbf{w}_f$. $\mathbf{\Psi}_q \mathbf{I}_N^{\langle k-\frac{Q-1}{2}\rangle} \mathrm{diag}(\widetilde{\mathbf{t}}) = \mathbf{0}$ for $k+q > Q$ and $k+q < 1$ on the basis of the pilot groups in Fig. 1. Therefore, $\widetilde{\mathbf{r}}_{\mathbf{p}_q}$ is rewritten as

$$\widetilde{\mathbf{r}}_{\mathbf{p}_q} = \sum_{k=0}^{Q-1} \mathbf{\Phi}_{k-q+\frac{Q-1}{2}} \mathbf{c}_{k,i} + \mathbf{w}', \quad 0 \le q < Q, \tag{17}$$

where

$$\mathbf{\Phi}_k = \begin{cases} \mathbf{\Psi}_k \mathrm{diag}(\widetilde{\mathbf{t}}) \mathbf{F}'_N, & 0 \le k < Q \\ \mathbf{0}^{G\times L}, & \text{other} \quad k \end{cases}. \tag{18}$$

Combining the cumulation in (18), we can get

$$\widetilde{\mathbf{r}}_{\mathbf{p}_q} = \mathbf{\Phi} \widetilde{\mathbf{c}}_q + \mathbf{w}', \quad 0 \le q < Q, \tag{19}$$

where $\widetilde{\mathbf{c}}_q = [\mathbf{c}^T_{-\frac{Q-1}{2}+q}, \mathbf{c}^T_{-\frac{Q-1}{2}+q+1}, \ldots, \mathbf{c}^T_{\frac{Q+1}{2}+q}]^T$ and $\mathbf{\Phi} = [\mathbf{\Phi}_0, \mathbf{\Phi}_1, \cdots, \mathbf{\Phi}_{\mathrm{Q}-1}]$ are the sparse vector and the sensing matrix, respectively. $\mathbf{c}_k^T = \mathbf{0}^{L\times 1}$ when $k \ge Q$ and $k < 0$.

The next steps utilize $Q = 3$ and noiseless situation as an illustration for simple presentation without loss of generality. Other values of Q with noise share the same results, whose derivation is analogous.

The formulas for $0 \le q < 3$ in (19) can be integrated as a whole, which is

$$\widetilde{\mathbf{r}}_p = \mathbf{\Phi}_\Lambda \widetilde{\mathbf{c}}, \tag{20}$$

where $\widetilde{\mathbf{r}}_p = [\widetilde{\mathbf{r}}^T_{\mathbf{p}_0}, \widetilde{\mathbf{r}}^T_{\mathbf{p}_1} \widetilde{\mathbf{r}}^T_{\mathbf{p}_2}]^T$, $\mathbf{\Phi}_\Lambda = \mathrm{diag}(\mathbf{\Phi}, \mathbf{\Phi}, \mathbf{\Phi}) \in \mathbb{C}^{3G\times 9L}$. $\widetilde{\mathbf{c}} = [\widetilde{\mathbf{c}}_0^T, \widetilde{\mathbf{c}}_1^T, \widetilde{\mathbf{c}}_2^T]^T = [\mathbf{0}^T, \mathbf{c}_0^T, \mathbf{c}_1^T, \mathbf{c}_0^T, \mathbf{c}_1^T, \mathbf{c}_2^T, \mathbf{c}_1^T, \mathbf{c}_2^T, \mathbf{0}^T]^T$ includes both $\mathbf{0}^T \in \mathbb{C}^{L\times 1}$ and nonzero parts of $\mathbf{c}_0^T$, $\mathbf{c}_1^T$, and $\mathbf{c}_2^T$. For simplicity, the zero parts are ignored by deleting the corresponding columns in $\mathbf{\Phi}_\Lambda$. The nonzero parts are integrated by summing up the corresponding columns. Therefore, (20) can be simplified as

$$\widetilde{\mathbf{r}}_p = \hat{\mathbf{\Phi}}_\Lambda \check{\mathbf{c}}, \tag{21}$$

where $\check{\mathbf{c}} = [\mathbf{c}_0^T, \mathbf{c}_0^T, \mathbf{c}_1^T, \mathbf{c}_1^T, \mathbf{c}_1^T, \mathbf{c}_2^T, \mathbf{c}_2^T]^T$. $\hat{\mathbf{\Phi}}_\Lambda \in \mathbb{C}^{3G\times 7L}$ is the matrix after columns delete and replaced from $\mathbf{\Phi}_\Lambda$. Rewrite $\hat{\mathbf{\Phi}}_\Lambda$ as $\hat{\mathbf{\Phi}}_\Lambda = [\hat{\mathbf{\Phi}}_{\Lambda 0}, \hat{\mathbf{\Phi}}_{\Lambda 1}, \ldots, \hat{\mathbf{\Phi}}_{\Lambda 6}]$ by seperating it into 7 isometric submatrix, where $\hat{\mathbf{\Phi}}_{\Lambda k} \in \mathbb{C}^{3G\times L}$ for $0 \le k < 7$. The ultimate formula in the CS framework can be represented as

$$\widetilde{\mathbf{r}}_p = \hat{\mathbf{\Phi}} \hat{\mathbf{c}}, \tag{22}$$

where $\hat{\mathbf{\Phi}} = [\hat{\mathbf{\Phi}}_{\Lambda 0} + \hat{\mathbf{\Phi}}_{\Lambda 1}, \hat{\mathbf{\Phi}}_{\Lambda 2} + \hat{\mathbf{\Phi}}_{\Lambda 3} + \hat{\mathbf{\Phi}}_{\Lambda 4}, \hat{\mathbf{\Phi}}_{\Lambda 5} + \hat{\mathbf{\Phi}}_{\Lambda 6}]$ and $\hat{\mathbf{c}} = [\mathbf{c}_0^T, \mathbf{c}_1^T, \mathbf{c}_2^T]^T$.

In (22), only one transmit antenna is considered and we leave out i in the subscript. In general cases, the subscript i is appended and the received signal is superimposed by $\widetilde{\mathbf{r}}_p$, which is

$$\widetilde{\mathbf{r}}_{p,\text{total}} = \sum_{i=0}^{N_t-1} \widetilde{\mathbf{r}}_{p,i} = \sum_{i=0}^{N_t-1} \hat{\mathbf{\Phi}}_i \hat{\mathbf{c}}_i. \tag{23}$$

Define $\overline{\boldsymbol{c}} = [\hat{\mathbf{c}}_1^T, \hat{\mathbf{c}}_2^T, \ldots, \hat{\mathbf{c}}_{N_t-1}^T]^T$ and $\overline{\mathbf{\Phi}} = [\hat{\mathbf{\Phi}}_1, \hat{\mathbf{\Phi}}_{2,\ldots,} \hat{\mathbf{\Phi}}_{N_t-1}]$, then (23) can be simplified as

$$\widetilde{\mathbf{r}}_{p,\text{total}} = \overline{\mathbf{\Phi}} \overline{\boldsymbol{c}}, \tag{24}$$

where $\overline{\boldsymbol{c}} \in \mathbb{C}^{QLN_t}$ and $\overline{\mathbf{\Phi}} \in \mathbb{C}^{QG\times QLN_t}$ are the sparse signal to be recovered and the sensing matrix, respectively. $\overline{\boldsymbol{c}}$ can be divided into L blocks because $\mathbf{c}_q$ have the characteristics of jointly sparse for $0 \le q < Q$. The i-th block includes the entries $[\widehat{\mathbf{c}}]_{i:L:(QN_t-1)L+i}$, which are uniformly zeros or nonzeros. Accordingly, equation (24) is a quintessential block sparse model and using SCS methods $\mathbf{c}_q$ will be estimated.

IV. Pilot Pattern Design

In the CS theroy, the sparse signal recovery is greatly determined by the sensing matrix $\hat{\mathbf{\Phi}}$. The sensing matrix is determined by the pilot group location of all transmit antennas according to the derivation from (12) to (24) in this article. Consequently, a pilot group location, which is appropriately designed, is beneficial to the sparse signal recovery and the channel estimation.

For assessing the sensing matrix, the restricted isometry property (RIP) [36] is a crucial criterion. The RIP is fulfilled for the sensing matrix $\hat{\mathbf{\Phi}}$ if a constant $\delta (0 < \delta < 1)$ satisfying

$$(1-\delta) \|\boldsymbol{\eta}\|_2^2 \le \left\| \hat{\mathbf{\Phi}} \boldsymbol{\eta} \right\|_2^2 \le (1+\delta) \|\boldsymbol{\eta}\|_2^2, \tag{25}$$

for all *S*-sparse vectors $\boldsymbol{\eta} \in C^L$. Nevertheless, the identification algorithm with polynomial time complexity for RIP does not exist [37]. Fortunately, the mutual incoherence property (MIP), which is the sufficient condition for RIP, can be evaluated easily instead. The optimized parameter, i.e., the *coherence* is the maximum coherence between different columns and is defined as

$$\mu = \max_{0\le l,k<L, l\ne k} \frac{|\langle \boldsymbol{\varphi}_l, \boldsymbol{\varphi}_k \rangle|}{\|\boldsymbol{\varphi}_l\|_2 \cdot \|\boldsymbol{\varphi}_k\|_2}, \tag{26}$$

where $\boldsymbol{\varphi}_l$ and $\boldsymbol{\varphi}_k$ are the *l*-th and the *k*-th column vector in $\hat{\mathbf{\Phi}}$, respectively. The minimization of coherence has been exploited in lots of works to increase the accuracy of sparse signal recovery [38]–[40].

Considering MIP in SCS scenario, the performance of coherence deteriorates due to quantitative growth of the sensing matrix column vectors. By utilizing the characteristic of block sparsity, the column vectors corresponding to identical taps are integrated to blocks. In this condition, the blockcorrelation, which is also called as *blockcoherence* [41], is adopted to assess the correlation in one block. The blockcoherence is the optimized parameter in this article and is represented by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

$$\mu_{\mathrm{B}}(\boldsymbol{\Phi}) = \max_{0\le l,k<n,l\neq k} \frac{\rho\left(\boldsymbol{\Phi}_l^H \boldsymbol{\Phi}_k\right)}{N_t}, \tag{27}$$

where $\rho(\boldsymbol{A})$ is the spectrum norm for $\boldsymbol{A}$, which is equal to the maximum singular value. The column vectors of the blocks are normalized in the sensing matrix. The blockcoherence is the parameter to describe the coherence among the blocks, which is similar to the coherence. When the size of the block is 1, the blockcoherence is degraded to the coherence. Accordingly, the blockcoherence is an extension of the coherence. Instead of the column coherence, the coherence between different blocks is evaluated, which is the feature of the blockcoherence. With a small blockcoherence, it can be easy for the blocks to be discriminated from each other, which will increase the accuracy of the signal estimation. Therefore, the blockcoherence should be minimized in this article, i.e.,

$$\boldsymbol{\Phi}_{\text{opt}} = \arg\min_{\boldsymbol{\Phi}} \left(\mu_{\mathrm{B}}(\boldsymbol{\Phi})\right). \tag{28}$$

It is arduous to find an analytic solutions for the problem above in (29). Moreover, it is also impractical to do the brute force search due to the enormous search space. Consequently, a suboptimal search algorithm is needed.

The uniformly-spaced pilot pattern will lead to a large value of blockcoherence [42], which will result in the failure of the sparse signal recovery. Accordingly, the random location of pilots is more suitable for sparse signal recovery. However, its performance can be unpredictable. To design a sub-optimal pilots location, the conventional optimization methods includes the stochastic sequential search in [42], the cross entropy optimization in [43], and the greedy pilot allocation method in [44]. However, all these algorithms are based on the pilot optimization scheme in SISO system under static channel. In that condition, the optimization problem can be simplified to the cyclic difference set (CDS) or almost difference sets (ADS). For the doubly selective channel in MIMO condition in this article, the GA for pilot pattern design is proposed, which is shown in Algorithm **1**.

In GA, the items to be optimized are named individuals. An individual is translated into binary code and contains all the pilot center location information. There are N subcarriers in total. However, the eligible pilot center location number is less than N because different pilot groups cannot overlap. Actually, there are $Y = N - [2Q + (2Q-1)(G-1)]$ eligible pilot centers in total. Therefore, the individuals can be generated by randomly obtaining G numbers from 1 to Y. Then the G different numbers which contains the center location information will be translated into binary code. When performing the decoding process, the binary code can be translated into G numbers according to the coding rule. And then, sort the numbers in ascending order. For the i-th number Y_i, it corresponds the $Y_i + Q + (2Q-1)(i-1)$-th subcarrier.

The individuals can be mapped to a sensing matrix $\hat{\boldsymbol{\Phi}}$ according to *Step 3* in Section III. The fitness value of the individual, which is the parameter to be optimized, is the blockcoherence of the mapping sensing matrix.

In the algorithm, the crossover probability p_c can be set as 0.5 for a balance of exchanging elements of individuals. The probability distribution $\boldsymbol{P}$ is discrete and divided into n_d parts, where the non-uniform distribution can be available. After n_d individuals are generated, each individual is distributed with a selection probability according to the fitness value and $\boldsymbol{P}$. After selection, we perform the crossover for every element of the two individuals in order. On the basis of the individual definition, the values of individual elements are in the set of $\{1, -1\}$, which is convenient for the crossover. The crossover is an operation of exchanging an element of the two individuals. Appropriate crossover probability is indispensable to avoid the fitness value from going to local convergence.

The loop of the proposed GA ends after the individuals of a new generation are all the same or the simulation generation reaches n_g. At that time, an optimized individual with small blockcoherence will be obtained, which can be decoded to a sub-optimal locations of pilot centers.

Algorithm 1 GA for Pilot Pattern Design.

Inputs:

1) Subcarrier number N;
2) BEM coefficient Q;
3) Pilot groups number G
4) Number of transmit antennas N_t;
5) The crossover probability p_c;
6) Individual number for each generation n_d;
7) Maximum simulation generations number n_g;
8) A discrete probability distribution $\boldsymbol{P}$ for individual selection;

1: First generation initialization for n_d individuals.
2: **while** simulation generation does not reach n_g **do**
3: Calculate the fitness values of current generation individuals (using equation (27)) and sort them descending.
4: Distribute the selection probability through $\boldsymbol{P}$ according to the fitness value of the individuals.
5: **for** $i_{\text{iter}} = 0 : n_d/2 - 1$ **do**
6: Select two individuals based on the selection probability.
7: For the chromosome pairs selected, determine whether to be exchanged based on p_c.
8: If the new individuals are both available, they will comprise part of the next generation. Else, return to Step 6.
9: **end for**
10: Obtain a new generation.
11: **end while**

Output: Individual with smallest fitness value.

V. The Proposed ASA-BOMP for MIMO-OFDM Systems

A. Algorithm Description

Utilizing the designed pilot pattern and location, the receiver can perform channel estimation based on (24). A novel recovery algorithm called ASA-BOMP is proposed, which is detailed in Algorithm **2**. The matrix $\mathbf{c}_{\Omega}^{(s)}$ represents the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

Algorithm 2 Proposed ASA-BOMP for Doubly Selective Channel Estimation in MIMO-OFDM systems

Inputs:

1) Initial channel PCS $\Omega_0 = \emptyset$;
2) Noisy measurements $\widetilde{\mathbf{r}}_{p,\text{total}} \in \mathbb{C}^{QG}$;
3) Sensing matrix $\hat{\mathbf{\Phi}} \in \mathbb{C}^{QG \times QLN_t}$

Initialization:

1: $\Omega \leftarrow \Omega_0$
2: $S = \|\Omega_0\|_0 - 1$

Iterations:

3: **while** $\|\mathbf{r}^{(S-\|\Omega_0\|_0)}\|_2 > \epsilon^2$ **do**
4: $S \leftarrow S + 1$
5: **for** $s = 0 : S - \|\Omega_0\|_0 - 1$ **do**
6: $\mathbf{c}^{(s)} \leftarrow \mathbf{0}$; $\mathbf{c}_{\Omega}^{(s)} \leftarrow \hat{\mathbf{\Phi}}_{\Omega}^{\dagger} \widetilde{\mathbf{r}}_{p,\text{total}}$
7: $\mathbf{r}^{(s)} \leftarrow \widetilde{\mathbf{r}}_{p,\text{total}} - \hat{\mathbf{\Phi}} \mathbf{c}^{(s)}$
8: $\mathbf{g} \leftarrow \hat{\mathbf{\Phi}}^H \mathbf{r}^{(s)}$
9: $\Omega \leftarrow \Omega \cup \arg\max_j \sum_{i=0}^{QN_t-1} |\mathbf{g}_{iL+j}|$
10: **end for**
11: $\mathbf{c}^{(S-\|\Omega_0\|_0)} \leftarrow \mathbf{0}$; $\mathbf{c}_{\Omega}^{(S-\|\Omega_0\|_0)} \leftarrow \hat{\mathbf{\Phi}}_{\Omega}^{\dagger} \widetilde{\mathbf{r}}_{p,\text{total}}$
12: $\mathbf{r}^{(S-\|\Omega_0\|_0)} \leftarrow \widetilde{\mathbf{r}}_{p,\text{total}} - \hat{\mathbf{\Phi}} \mathbf{c}^{(S-\|\Omega_0\|_0)}$
13: **end while**

Output: $\hat{\mathbf{c}} = \mathbf{c}^{(S-\|\Omega_0\|_0)}$

submatrix of $\mathbf{c}^{(s)}$ with blocks in Ω, while $\hat{\mathbf{\Phi}}_{\Omega}$ is a submatrix of $\hat{\mathbf{\Phi}}$ with columns in Ω. The acquired PCS Ω_0 is the additional priori information in ASA-BOMP compared with the traditional algorithm. Moreover, the computational complexity is lower because only $S - \|\Omega_0\|_0$ iterations are required instead of S in BOMP. The adoption of the PCS can also improve the recovery performance of the algorithm. Furthermore, the algorithm is more feasible because it assumes no priori information to S and will adaptively detect the sparsity level.

After the estimation of $\mathbf{c}_{q,i} (0 \le q < Q, 0 \le i < N_t)$, the elements $c[q,l,i]$ can be acquired by equation (5). And then, by utilizing (4) the doubly selective channel can be recovered.

B. Computational Complexity

For the inputs of the algorithm, Ω_0 and S are constants, while $\widetilde{\mathbf{r}}_{p,\text{total}}$ can be acquired according to the pilot pattern, which is obtained offline using Algorithm **1**. The sensing matrix $\hat{\mathbf{\Phi}}$ can be obtained based on (18) to (24). The computational complexity mainly lies in the generation of $\mathbf{\Phi}_k$ in (18), where Q diagonal matrix multiplications are needed with the computational complexity of $O(QNL)$.

For the body part of the algorithm, there are two kinds of iterations: an outer iteration (from step 3 to step 13) and an inner iteration (from step 5 to step 10), respectively. The outer iteration is utilized to try different values of the sparsity level of the channel, which is intensely essential for the adaptive property of the algorithm. The inner iteration is the main part of the algorithm, which will run $S - \|\Omega_0\|_0$ times to get the nonzero support. For each inner iteration, the computational complexity consists of two main parts: the equivalent least square problem $\mathbf{c}_{\Omega}^{(s)} \leftarrow \hat{\mathbf{\Phi}}_{\Omega}^{\dagger} \widetilde{\mathbf{r}}_{p,\text{total}}$ in step 5, and the inner product between the observation matrix $\hat{\mathbf{\Phi}}^H$ and the residue matrix $\mathbf{r}^{(s)}$ in step 7. The former has the computational complexity of $O(Q^2 LGN_t)$, while the latter has the computational complexity of $O(Q^2 LGN_t(QG+1))$. The inner iterations are performed $S - \|\Omega_0\|_0$ times for each outer iteration. The outer is performed S times at most. Compared with these, the computational complexity for the inputs can be omitted. Therefore, the total computational complexity of the ASA-BOMP is on the order of $O((S-\|\Omega_0\|_0)SQ^2 LGN_t(QG+2))$.

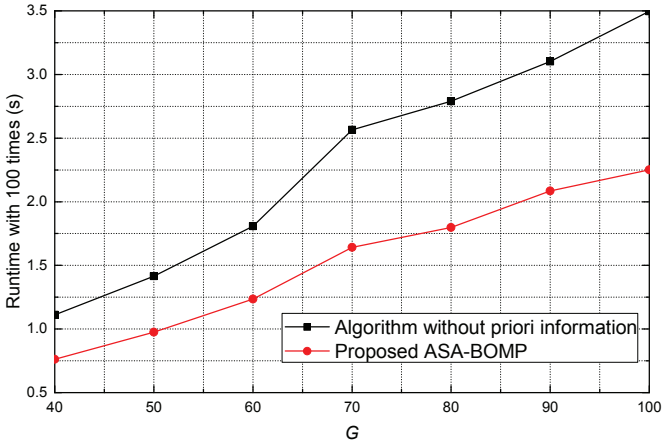


Fig. 4. Performance of the proposed ASA-BOMP and algorithm without priori information.

C. Performance Analysis

The performance of the ASA-BOMP is evaluated in Fig. 4. The simulation was run on the MATLAB 2012b via the CPU of Intel(R) Core(TM) i7-4770 with 3.40 GHz and memory size of 10.0 GB. The runtime versus the value of G is simulated. For other coefficients, $Q = 3$, $L = 256$, and $N_t = 8$ are configured. Note that the scale of the sensing matrix is $\hat{\mathbf{\Phi}} \in \mathbb{C}^{QG \times QLN_t}$. The algorithm without priori information of PCS is also simulated as comparison. In the figure, it can be seen that the ASA-BOMP runs faster than the counterpart, which suggests that the ASA-BOMP is an efficient algorithm to perform sparse signal recovery problems.

VI. Simulation Results

In the following, we simulate our proposed doubly selective channel estimation scheme and evaluate the performance. For the antennas, the azimuth angles of arrival (AoA) and angles of departure (AoD) are assumed to take continuous values, i.e., not quantized, and are uniformly distributed in $[0, 2\pi]$. The parameters in the simulations are listed in Table I.

Using the selectivity of both time and frequency domains, the doubly selective channel for a single antenna can be simulated. For multiple antennas, firstly, the non-zero supports are identical for different channels according to the property of spatial correlation. Secondly, the amplitudes of the taps are independent. Therefore, for the specific sample index, independent channel taps are configured for the channels from different antennas.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

TABLE I
PARAMETERS IN THE SIMULATIONS

Length of the PN sequences M	256
Length of the OFDM frames N	4096
Center frequency	634 MHz
Bandwidth W	7.56 MHz
Channel frequency selectivity	6-tap ITU vehicular-B channel model [45]
Channel time selectivity	Normalized Doppler frequency $f_{\max} = 0.08$ corresponding to the speed of 240 km/h
BEM coefficient Q	3

A. Influence of Pilot Group Location

The blockcoherence of different types of pilot patterns are illustrated in Fig. 5. Apart from the pilot group location designed by the proposed GA in Algorithm **1** ($p_c = 0.5, n_d = n_g = 100$, and $\boldsymbol{P}$ is the Geometric distribution with the common ratio of 0.95), the random pilot groups of pilot pattern is also illustrated as a benchmark. $N_t = 2, 4$, and 8 are considered, respectively. The blockcoherence for uniformly-spaced pilot pattern is extremely greater than that of the random one, which is not illustrated in the figure. It can be seen that the blockcoherence for the random pilot pattern is much larger than the proposed one, which indicates that the designed pilot pattern is more suitable for the sparse channel recovery. In general, the proposed GA can efficiently reduce the value of blockcoherence.

It should be noted that the sensing matrix with larger N_t tends to have a smaller value of blockcoherence, which is counter-intuitive. As the number of antennas rises, the additional column vectors will break the orthogonality of the blocks, which may lead to the increase of the local blockcoherence (between two blocks). However, the blockcoherence is the maximum of the local blockcoherences, as shown in (27). For the maximum value, there is more relevance than orthogonality. This means when the size of the blocks increases, the additional column vector will be more likely to reduce the relevance. Therefore, the sensing matrix with the largest N_t has the smallest blockcoherence, and Fig. 5 illustrates this observation.

B. Performance of PAPR

The average peak-to-average power ratio (PAPR) for 10000 different frames is simulated. $G = 80$ and $Q = 3$ are configured. The average PAPR for the proposed scheme is 8.9751, while the average PAPR for the conventional scheme in [20] is 9.2333. The proposed scheme even outperforms the conventional one in PAPR because the number of the nonzero pilots is larger and the power of the pilots is less fluctuating. It indicates that the proposed pilot pattern scheme will not degrade the PAPR performance, which is intensely essential for the system stability.

C. Performance of Channel Estimation

The accurate recovery probability is simulated in Fig. 6 between the ASA-BOMP and the traditional BOMP [21] for

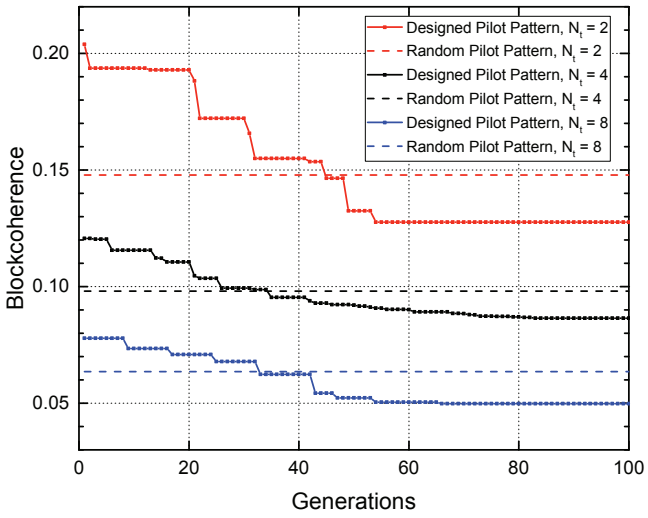


Fig. 5. Blockcoherence for different types of pilot patterns

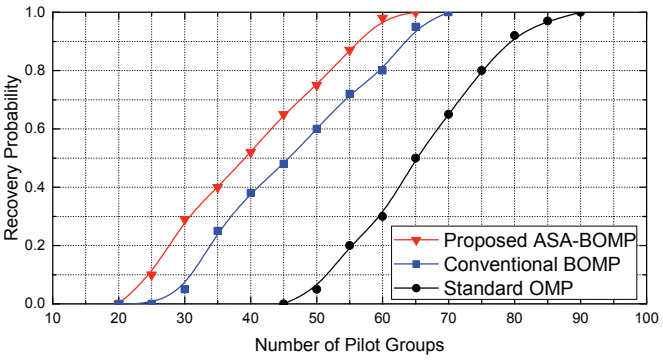


Fig. 6. Accurate recovery probabilities comparison with SNR = 20 dB in a 4×4 MIMO system.

4×4 MIMO systems. Meanwhile, the performance of OMP is also shown as a reference. The accurate recovery is the correct recovery of the support. The SNR is set to be 20 dB. Fig. 6 illustrates that the ASA-BOMP algorithm needs only 65 pilot groups to maintain the accurate CIR recovery, which saves 5 and 25 groups compared with BOMP and OMP, respectively. The improvements originate from the priori information of the obtained PCS and the utilization of the block sparse characteristic.

Fig. 7 demonstrates a quintessential channel estimation result for 4×4 MIMO systems when SNR is configured as 20 dB. For the sake of time-variant channel, only channel estimation results at sample 0 and L are demonstrated in the figure. It can be seen that the PCS can be obtained by selecting the entries larger than the threshold value. By using the obtained PCS, the nonzero support for the channel can be efficiently estimated for different antennas. When SNR = 20 dB, the nonzero support can be accurately estimated, while the amplitude and angle will have a tiny bias.

Fig. 8 illustrates the mean square error (MSE) simulations for $N_t = 2, 4$, and 8. $G = 80$ is configured for accurate estimation of the channel. The MSE is defined as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

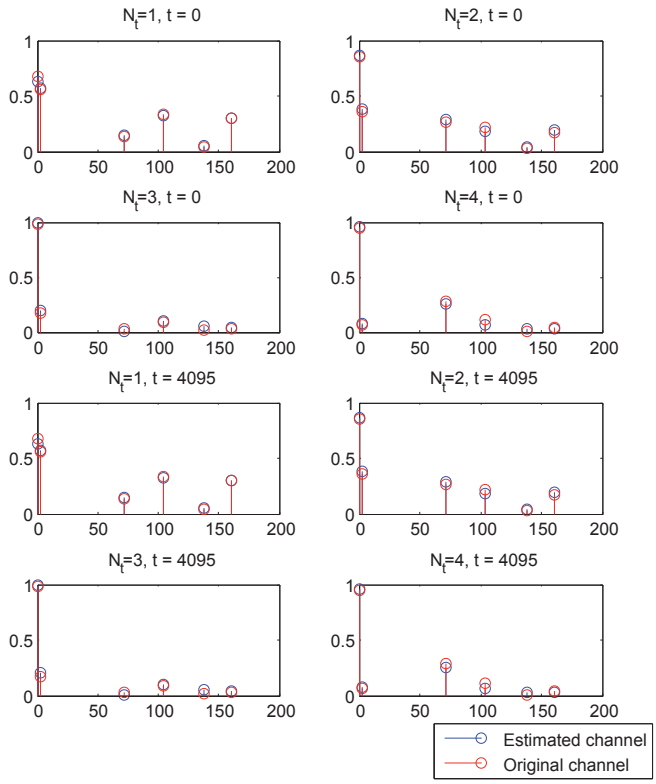


Fig. 7. Illustration of channel support recovery in a 4×4 systems with SNR = 20 dB.

$$\mathrm{E}(\|\widehat{\mathbf{h}} - \mathbf{h}\|_2^2) = \sum_{l=0}^{N-1} \| \sum_{i=0}^{N_t-1} (\widehat{\mathbf{h}}_s^{(l,i)} - \mathbf{h}_s^{(l,i)})\|_2^2. \qquad (29)$$

The conventional BOMP [21] is simulated by contrast, while the Cramer-Rao lower bound (CRLB) ($\eta = SN_t\sigma^2/G$, whose proof can be seen in Appendix) in the ideal case is also evaluated as a reference. Our scheme has better performance than the traditional BOMP because of the auxiliary information of nonzero support. Moreover, the power of the measurements is much larger than that in [22] due to more nonzero pilots, which further improves the performance. It can be seen that the ASA-BOMP algorithm has 1.5, 1.8, and more than 5 dB gains compared the conventional BOMP when the MSE of 10^{-2} is considered for $N_t = 2, 4$, and 8, respectively. The MSE is approaching the CRLB and is less than 0.5 dB away from it above the SNR level of 25 dB. The conventional TFDT-MIMO-OFDM time-invariant channel estimation technique in [3] is also simulated. For better presentation, only $N_t = 2$ is illustrated. It can be seen that the conventional TFDT-OFDM channel estimation scheme failed to recover the doubly selective channel, which proves the importance and the advantage of our proposed scheme. To summarize, the proposed scheme outperforms the traditional ones for MIMO-OFDM systems.

Fig. 9 illustrates the performance in moderate mobility scenarios, where 30 km/h is simulated. The conventional moderate mobility scheme [3] is compared as a benchmark. It can be seen that the performances of the two schemes are

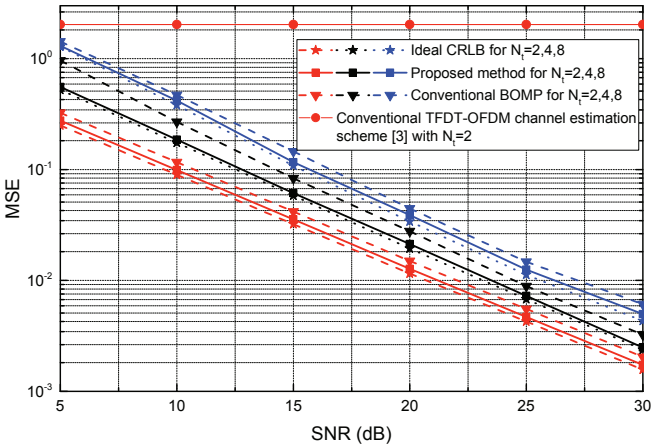


Fig. 8. The MSE performance comparison under speed of 240 km/h among different schemes for different number of transmit antennas.

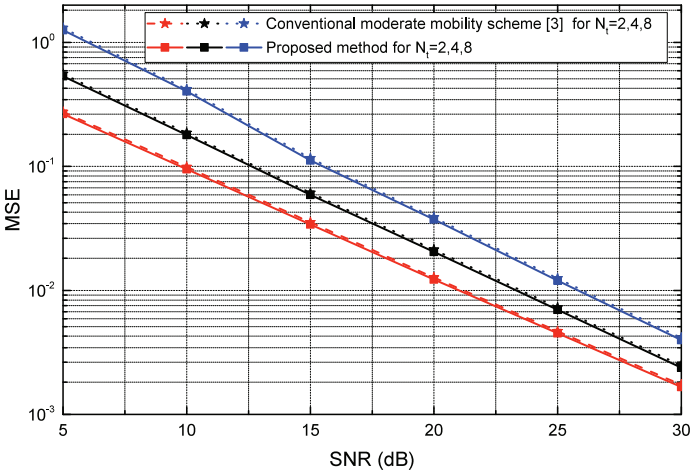


Fig. 9. The MSE performance comparison under speed of 30 km/h among different schemes for different number of transmit antennas.

almost the same, which suggests that the proposed scheme has advantage mainly in high mobility scenarios and can also be adopted in moderate mobility scenarios with an outstanding performance.

The BER is considered in Fig. 10 when low density parity check code (LDPC) is adopted [46] in a 4×4 MIMO system. The code length is 7493, code rates are 0.4 and 0.8, while the constellations are 256QAM and 16QAM. The ideal BER is also considered to indicate the lower bound. On the one hand, high-order modulations like 256QAM can be supported well in our scheme under doubly selective channel. On the other hand, low-order modulations such as 16QAM has excellent performance with low SNR demands. Our approach has 0.9 dB and 0.7 dB gains compared with the traditional BOMP when the BER is 10^{-3} for 256QAM and 16QAM, respectively. Moreover, our scheme is approaching and only around 0.2 dB away from the lower bound.

VII. Conclusions

A new channel estimation scheme in high-mobility situations is considered for multiple antenna systems in this article. The channel is abstracted as 3 domains to describe the doubly

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

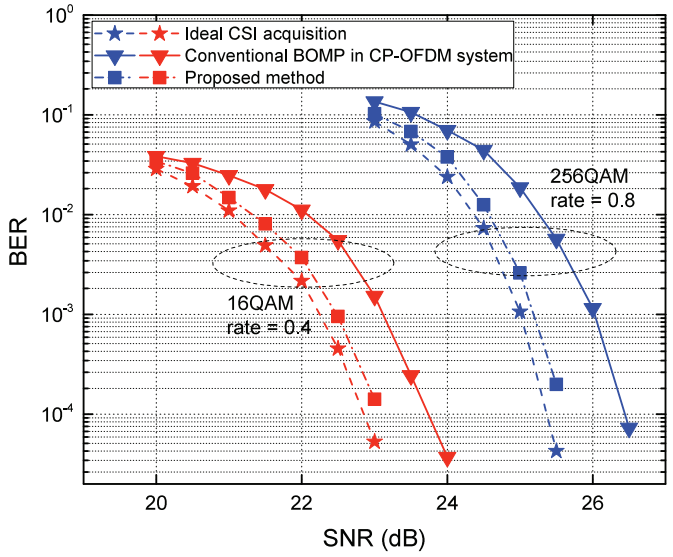


Fig. 10. BER performance comparison with 256QAM&16QAM constellation and LDPC rate of 0.8&0.4.

selective property. Both the PN sequence and the pilots are employed in our scheme. The PN sequence is adopted to obtain the PCS, while the pilots finally recover the channel based on the SCS model by utilizing the optimized pilot location via GA and the spatial correlation among antennas. An improved SCS algorithm ASA-BOMP for MIMO-OFDM systems is proposed for accurate CIR acquisition. It is demonstrated in the simulations that our scheme has better performance than the traditional ones with lower complexity and will be a promising technique in the future MIMO-OFDM systems.

APPENDIX
THE PROOF OF CRLB OF MSE

In the ideal case, the nonzero support is estimated correctly, which can be denoted as $D(\|D\|_0 = QSN_t)$. Consequently, the entries outside the support D are set to zeros. Equation (24) is simplified as

$$\tilde{\mathbf{r}}_{p,\text{total}} = \overline{\boldsymbol{\phi}}_D \bar{\mathbf{c}}_D, \tag{30}$$

which can be estimated by solving an over-determined equation under the maximum likelihood (ML) criterion.

$$\bar{\mathbf{c}}_{\text{est}} = \overline{\boldsymbol{\phi}}_D^{+} \tilde{\mathbf{r}}_{p,\text{total}} = (\overline{\boldsymbol{\phi}}_D^{H} \overline{\boldsymbol{\phi}}_D)^{-1} \overline{\boldsymbol{\phi}}_D \tilde{\mathbf{r}}_{p,\text{total}}. \tag{31}$$

Then, the CRLB of $\bar{\mathbf{c}}$ can be denoted as

$$\eta_c = \mathrm{E}\{\|\bar{\mathbf{c}}_{\text{est}} - \bar{\mathbf{c}}_D\|_2\} = \frac{QSN_t}{QG}\sigma^2 = \frac{SN_t}{G}\sigma^2. \tag{32}$$

Utilizing (4), the CRLB of the proposed channel estimation method can be derived as

$$\begin{aligned}
&\sqrt{N}\eta_h = \mathrm{E}\{\|\widehat{\mathbf{h}} - \mathbf{h}\|_2\} \\
&= \mathrm{E}\{\sqrt{\sum_{l=0}^{N-1} \| \sum_{i=0}^{N_t-1} (\widehat{\mathbf{h}}_s^{(l,i)} - \mathbf{h}_s^{(l,i)})\|_2^2}\} \\
&= \mathrm{E}\{\sqrt{\sum_{l=0}^{N-1} \| \sum_{i=0}^{N_t-1} (\sum_{q=0}^{Q-1} \widehat{c}[q,l,i]\mathbf{b}_q - \sum_{q=0}^{Q-1} c[q,l,i]\mathbf{b}_q)\|_2^2}\} \\
&= \mathrm{E}\{\sqrt{\sum_{l=0}^{N-1} \|(\sum_{i=0}^{N_t-1} \sum_{q=0}^{Q-1} (\widehat{c}[q,l,i] - c[q,l,i]))\mathbf{b}_q\|_2^2}\} \\
&= \mathrm{E}\{\sqrt{\sum_{l=0}^{N-1} [(\sum_{q=0}^{Q-1} (\widehat{c}[q,l] - c[q,l])b_q)^H \sum_{q=0}^{Q-1} (\widehat{c}[q,l] - c[q,l])b_q]}\} \\
&= \mathrm{E}\{\sqrt{\sum_{l=0}^{N-1} \sum_{q=0}^{Q-1} \sum_{i=0}^{N_t-1} \mathbf{b}_q^H \mathbf{b}_q (\widehat{c}[q,l,i] - c[q,l,i])^2}\} \\
&= \sqrt{N}\eta_c \\
&= \frac{\sqrt{N}SN_t}{G}\sigma^2
\end{aligned} \tag{33}$$

The derivation utilizes the orthogonality of the expansion basis, and the ℓ_2-norm of the basis is $\sqrt{\mathbf{b}_q^H \mathbf{b}_q} = \sqrt{N}$ for $q = 0, 1, ..., Q-1$. Accordingly, the CRLB is rewritten as

$$\eta_h = \frac{SN_t}{G}\sigma^2. \tag{34}$$

REFERENCES

[1] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO-OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 219–224, May 2006.

[2] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[3] W. Ding, F. Yang, S. Liu, X. Wang, and J. Song, "Non-orthogonal time-frequency training-sequence-based CSI acquisition for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5714–5719, Jul. 2016.

[4] Y. Fu, C. X. Wang, A. Ghazal, e. H. M. Aggoune, and M. M. Alwakeel, "Performance investigation of spatial modulation systems under non-stationary wideband high-speed train channel models," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6163–6174, Sept. 2016.

[5] H. Ghazzai, T. Bouchoucha, A. Alsharoa, E. Yaacoub, M. S. Alouini, and T. Y. Al-Naffouri, "Transmit power minimization and base station planning for high-speed trains with multiple moving relays in OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 175–187, Jan. 2017.

[6] "LTE/SAE–the future railway mobile radio system? long-term visions on railway mobile radio technologies draft." Int. Union Railways, Tech. Rep., Nov. 2009.

[7] A. Amanna, M. Gadnhiok, M. J. Price, J. H. Reed, W. P. Siriwongpairat, and T. K. Himsoon, "Railway cognitive radio," *IEEE Veh. Technol. Mag.*, vol. 5, no. 3, pp. 82–89, Sept. 2010.

[8] J. Fu, C. Pan, Z. Yang, and L. Yang, "Low-complexity equalization for TDS-OFDM systems over doubly selective channels," *IEEE Trans. Broadcast.*, vol. 51, no. 3, pp. 401–407, Sept. 2005.

[9] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2226–2238, May 2007.

[10] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[11] Z. Han, H. Li, and W. Yin, "Compressive sensing for wireless networks." Cambridge University Press, UK, 2013.

[12] J. Meng, W. Yin, Y. Li, N. T. Nguyen, and Z. Han, "Compressive sensing based high-resolution channel estimation for OFDM system," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 1, pp. 15–25, Feb. 2012.

[13] X. Ma, F. Yang, S. Liu, J. Song, and Z. Han, "Doubly selective channel estimation for MIMO systems based on structured compressive sensing," in *Proc. 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Jun. 2017, pp. 610–615.

[14] F. Yang, W. Ding, L. Dai, and J. Song, "Joint time-frequency channel estimation method for OFDM systems based on compressive sensing," in *2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*, Aug. 2014, pp. 1–4.

[15] W. Ding, F. Yang, C. Zhang, L. Dai, and J. Song, "Simultaneous time-frequency channel estimation based on compressive sensing for OFDM system," in *2014 IEEE Global Communications Conference*, Dec. 2014, pp. 3573–3578.

[16] W. Ding, F. Yang, C. Pan, L. Dai, and J. Song, "Compressive sensing based channel estimation for OFDM systems under long delay channels," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 313–321, Jun. 2014.

[17] L. Dai, J. Wang, Z. Wang, P. Tsiaflakis, and M. Moonen, "Spectrum- and energy-efficient OFDM based on simultaneous multi-channel reconstruction," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6047–6059, Dec. 2013.

[18] R. Mohammadian, A. Amini, and B. H. Khalaj, "Deterministic pilot design for sparse channel estimation in a MISO/multi-user OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 129–140, Jan. 2017.

[19] W. Ding, F. Yang, S. Liu, and J. Song, "Structured compressive sensing-based non-orthogonal time-domain training channel state information acquisition for multiple input multiple output systems," *IET Communications*, vol. 10, no. 6, pp. 685–690, Apr. 2016.

[20] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 601–617, Feb. 2016.

[21] W. Hou and C. W. Lim, "Structured compressive channel estimation for large-scale MISO-OFDM systems," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 765–768, May 2014.

[22] P. Cheng, Z. Chen, Y. Rui, Y. J. Guo, L. Gui, M. Tao, and Q. T. Zhang, "Channel estimation for OFDM systems over doubly selective channels: A distributed compressive sensing based approach," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4173–4185, Oct. 2013.

[23] A. Kalakech, M. Berbineau, I. Dayoub, and E. P. Simon, "Time-domain lmmse channel estimator based on sliding window for OFDM systems in high-mobility situations," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5728–5740, Dec. 2015.

[24] X. Zhang, L. Gui, Q. Qin, and B. Gong, "Dynamic sparse channel estimation over doubly selective channels: Differential simultaneous orthogonal matching pursuit," in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Jun. 2016, pp. 1–6.

[25] Q. Qin, B. Gong, L. Gui, X. Ren, and W. Chen, "Structured distributed sparse channel estimation for high mobility OFDM systems," in *2015 International Workshop on High Mobility Wireless Communications (HMWC)*, Oct. 2015, pp. 56–60.

[26] Q. Qin, L. Gui, B. Gong, X. Ren, and W. Chen, "Structured distributed compressive channel estimation over doubly selective channels," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 521–531, Sept. 2016.

[27] S. Pejoski and V. Kafedziski, "Estimation of sparse time dispersive channels in pilot aided OFDM using atomic norm," *IEEE Wireless Communications Letters*, vol. 4, no. 4, pp. 397–400, Aug. 2015.

[28] O. E. Barbu, C. N. Manchn, C. Rom, T. Balercia, and B. H. Fleury, "OFDM receiver for fast time-varying channels using block-sparse bayesian learning," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 053–10 057, Dec. 2016.

[29] G. Dziwoki and J. Izydorczyk, "Iterative identification of sparse mobile channels for TDS-OFDM systems," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 384–397, Jun. 2016.

[30] W. Ding, F. Yang, W. Dai, and J. Song, "Time frequency joint sparse channel estimation for MIMO-OFDM systems," *IEEE Commun. Lett.*, vol. 19, no. 1, pp. 58–61, Jan. 2015.

[31] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1384–1400, Jul. 2000.

[32] S. Liu, F. Yang, X. Wang, J. Song, and Z. Han, "Structured-compressed-sensing-based impulsive noise cancelation for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6921–6931, Aug 2017.

[33] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104–118, Jan. 2015.

[34] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse MIMO channels with common support," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3705 – 3716, Dec. 2011.

[35] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2136–2148, Dec. 2002.

[36] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[37] Z. Ben-Haim, Y. C. Eldar, and M. Elad, "Coherence-based performance guarantees for estimating a sparse vector under random noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5030–5043, Oct. 2010.

[38] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.

[39] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[40] X. Ma, F. Yang, W. Ding, and J. Song, "Novel approach to design time-domain training sequence for accurate sparse channel estimation," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 512–520, Sept. 2016.

[41] Y. C. Eldar and H. Bolcskei, "Block-sparsity: Coherence and efficient recovery," *in Proc. International Conference on Acoustics Speech and Signal Processing (ICASSP), Taipei, Taiwan*, pp. 2885–2888, Apr. 2009.

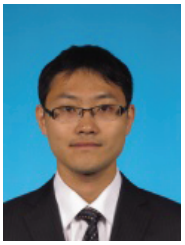
[42] C. Qi, L. Wu, Y. Huang, and A. Nallanathan, "Joint design of pilot power and pilot pattern for sparse cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5384–5390, Nov. 2015.

[43] E. Manasseh, S. Ohno, and M. Nakamoto, "Pilot symbol assisted channel estimation for OFDM-based cognitive radio systems," *EURASIP J. Adv. Signal Process.*, vol. 51, no. 1, pp. 1–11, Mar. 2013.

[44] R. Mohammadian, A. Amini, and B. H. Khalaj, "Compressive sensing-based pilot design for sparse channel estimation in OFDM systems," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 4–7, Jan. 2017.

[45] *Guidelines for evaluation of radio transmission technology for IMT-2000*, Recommendation ITU-R M. 1225 Std., Dec. 1997.

[46] J. Song, Z. Yang, L. Yang, K. Gong, C. Pan, J. Wang, and Y. Wu, "Technical review on Chinese digital terrestrial television broadcasting standard and measurements on some working modes," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 1–7, Mar. 2007.



Xu Ma received his B.S.E. in 2014 from the Department of Electronic Engineering in Tsinghua University, Beijing China. Currently, he is pursing the Ph. D. degree at the DTV Technology R&D Center, Tsinghua University. His research interests lie in the field of the power line communication (PLC), visible light communication (VLC), and sparse signal processing theory in wireless communications.

Fang Yang (M'11-SM'13) received the B.S.E. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2005 and 2009, respectively. He is currently working as an Associate Professor with the Research Institute of Information Technology, Tsinghua University. He has published over 120 peer-reviewed journal and conference papers. He holds over 40 Chinese patents and 2 PCT patents. His research interests lie in the fields of channel coding, channel estimation, interference cancelation, and signal processing techniques for communication system, especially in power line communication, visible light communication, and digital television terrestrial broadcasting. Dr. Yang received the IEEE Scott Helt Memorial Award (best paper award in IEEE Transactions in Broadcasting) in 2015. He is the Secretary General of Sub-Committee 25 of the China National Information Technology Standardization (SAC/TC28/SC25).

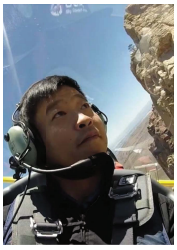
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2018.2811368, IEEE Transactions on Vehicular Technology

Sicong Liu (S'15-M'17) received his B.S.E. and the PhD degree both in electronic engineering from Tsinghua University, Beijing, China in 2012 and 2017 (with the highest honor). From 2010 to 2011, he was a visiting scholar in City University of Hong Kong. From 2017 to 2018, he served as a senior research engineer in Huawei Technologies Co. Ltd. Currently, he is an assistant professor in the Department of Communications Engineering, School of Information Science and Technology, Xiamen University, China.

Sicong Liu has published over 35 journal and conference research papers. He owns 7 Chinese invention patents. He is one of the core members that draft the Broadband Power Line Communications Standard in China. He has won the Best Doctoral Dissertation Award of Tsinghua University. He is a reviewer of many top journals and served as the guest editor of the Future Internet Journal, and the TPC member of several international conferences.

His research interests lie in sparse signal processing, interference mitigation, 5G emerging technologies, and machine learning.

Jian Song (M'06-SM'10-F'16) received his B. Eng and Ph. D. degrees in Electrical Engineering both from Tsinghua University, Beijing China in 1990 and 1995, respectively and worked for the same university upon his graduation. He has worked at The Chinese University of Hong Kong and University of Waterloo, Canada in 1996 and 1997, respectively. He has been with Hughes Network Systems in USA for 7 years before joining the faculty team in Tsinghua in 2005 as a professor. He is now the director of Tsinghua's DTV Technology R&D center. He has been working in quite different areas of fiber-optic, satellite and wireless communications, as well as the power line communications. His current research interest is in the area of digital TV broadcasting. Dr. Song has published more than 200 peer-reviewed journal and conference papers. He holds 2 US and more than 40 Chinese patents. He is Fellow of both IEEE and IET.



Zhu Han (S'01-M'04-SM'09-F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. Currently, Dr. Han is an IEEE Communications Society Distinguished Lecturer. In 2017, Dr. Han is 1% highly cited researcher according to Web of Science