Structured-Compressed-Sensing-Based Impulsive Noise Cancelation for MIMO Systems

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

***Abstract*—In this paper, the powerful signal processing theory of structured compressed sensing (SCS) is exploited to overcome the challenge of impulsive noise (IN) cancelation in multiple-input multiple-output (MIMO) systems. To the best of the authors' knowledge, the SCS theory is adopted for the first time for IN elimination, bridging IN mitigation and MIMO systems for its potential applications in vehicular-related communications. In achieving the SCS-based IN cancelation, the measurements matrix of the IN is first obtained from the null subcarriers in the MIMO system specified by the IEEE 802.11 standards series. The SCS optimization framework is then formulated through the proposed spatially multiple measuring method, by fully exploiting the spatial correlation of the IN signals at different receive antennas. To efficiently reconstruct the IN signal, an enhanced SCS-based greedy algorithm, structured *a priori* aided sparsity adaptive matching pursuit, is proposed, which significantly improves the accuracy and robustness compared with the state-of-the-art methods. Theoretical analysis is presented to guarantee the convergence and the performance error bound of the proposed greedy algorithm. Computer simulations validate that the proposed scheme outperforms the conventional ones over the wireless MIMO channel.**

***Index Terms*—Impulsive noise (IN), multiple-input multiple-output (MIMO), noise mitigation, orthogonal frequency division multiplexing (OFDM), structured compressed sensing (SCS).**

Manuscript received May 10, 2016; revised August 25, 2016 and December 30, 2016; accepted January 19, 2017. Date of publication January 25, 2017; date of current version August 11, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61401248, in part by the New Generation Broadband Wireless Mobile Communication Network of the National Science and Technology Major Projects under Grant 2015ZX03002008, in part by the Tsinghua University Initiative Scientific Research Program under Grant 2014Z06098, and in part by the National Science Foundation under Grant CNS-1646607, Grant ECCS-1547201, Grant CCF-1456921, Grant CNS-1443917, and Grant ECCS-1405121. The review of this paper was coordinated by Prof. W. A. Hamouda. *(Corresponding author: Fang Yang.)*

S. Liu is with the Research Institute of Information Technology and Electronic Engineering Department, Tsinghua National Laboratory for Information Science and Technology,Tsinghua University, Beijing 100084, China (e-mail: liu-sc12@mails.tsinghua.edu.cn).

F. Yang and J. Song are with the Research Institute of Information Technology and Electronic Engineering Department, Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China, and also with the Shenzhen City (Guangdong Province) Key Laboratory of Digital TV System, Shenzhen 518057, China (e-mail: fangyang@tsinghua.edu.cn; jsong@tsinghua.edu.cn).

X. Wang is with the Department of Electrical and Computer Engineering, University of Western Ontario, London, ON N6A 5B9, Canada (e-mail: wang@eng.uwo.ca).

Z. Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005, USA (e-mail: hanzhu22@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2017.2657696

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM) has been widely adopted in multiple-input multiple-output (MIMO) systems, such as the wireless local area network specified in the IEEE 802.11n standard [1], and the wireless access in vehicular environments specified in the IEEE 802.11p Wireless Access in Vehicular Environments (WAVE) standard [2], due to its high spectral efficiency and superior performance. Despite its advantages, OFDM-based MIMO systems suffer from non-Gaussian and nonlinear impulsive noise (IN), which is commonly seen in wireless and wired systems, caused by the switches of electrical devices, ignition noise in vehicles, or strong bursty radio frequency emission, etc. [3]–[5]. The spark IN caused by the ignition sparks and engine rotation in vehicles, for example, has long been very common impact in vehicular communications such as the IEEE 802.11p system [6], [7]. The IN causes severe degradation and is a serious bottleneck of the performance of MIMO-OFDM systems [8]. One difficulty of IN mitigation in OFDM system is that the impulse noise shares similar statistics in time domain with OFDM signal, when the impulse noise is active. The difference between impulse noise and OFDM signal is that IN has extremely wide coherence bandwidth, whereas the different subcarriers of OFDM signal are independently modulated. The time-domain bursty IN contaminates all the subcarriers in the OFDM block, so it is very difficult to mitigate its influence on all the contaminated subcarriers if the IN energy exceeds a certain threshold [9].

To mitigate the IN, some conventional methods have been proposed. Some researches focus on the clipping and/or blanking method to produce nonlinear processors for receivers against the IN [10]. A multiple channel selection combining scheme to suppress the IN defined by the Gaussian mixture model is proposed in [11]. Precoding and frequency algebraic interpolation techniques are implemented to detect the positions of IN [12]. However, the influence of IN with high intensity cannot be effectively eliminated by conventional methods through suppressing the power of IN or clipping/blanking the useful data, which might result in data loss and performance degradation.

Another difficulty is that the time-domain IN signal is high dimensional and mixed up with the information data component, and it is difficult to acquire enough measurement data for IN estimation using conventional methods of parametric signal estimation. Fortunately, the recently emerged theory of compressed sensing (CS) [13]–[15] can be introduced to recover the IN signal to overcome these conventional difficulties. According

to the CS theory, the unknown high-dimensional sparse signal can be recovered from a measurement vector of much smaller size. Since the IN is naturally a sparse signal in the time domain, it is likely to be recovered based on CS. Nevertheless, the research on CS-based IN mitigation is inadequate yet, among which the sparse convex optimization (SCO) method [16] is to utilize the unused subcarriers as measurement data. In our previous work [17], the classical CS-based greedy algorithm of sparsity adaptive matching pursuit (SAMP) [18] is adopted to eliminate IN.

However, state-of-the-art CS-based methods are only aimed at the single-input single-output (SISO) system, and the research on IN cancelation for the MIMO system is insufficient. In the MIMO system, there exists the inherent spatial correlation between the IN signals at different receive antennas [19], [20]. The spatial correlation can be utilized using the extended theory of CS, i.e., the theory of structured compressed sensing (SCS) [21], to further improve the performance of conventional CS-based methods, especially when the intensity of the IN is fluctuating and its sparsity level (number of nonzero entries) is large. As an emerging and breakthrough theory of sparse approximation, the SCS theory has drawn plenty of attention for the high recovery accuracy in sparse approximation and robustness to large sparsity levels [21]. However, to the best of the authors' knowledge, there is no related research on the IN mitigation based on the SCS theory for MIMO systems. To fill this gap and improve the robustness and accuracy of state of the arts, a novel SCS-based IN recovery and elimination scheme for MIMO-OFDM systems is proposed in this paper, taking full advantage of the spatial correlation of the IN in different receive antennas. The contributions are twofold.

1) The SCS theory is introduced to the area of IN recovery for the first time, to derive the proposed method of spatially multiple measuring (SMM) and formulate the SCS convex optimization framework, fully exploiting the inherent spatial correlation of the IN in MIMO systems.
2) The classical CS-based SAMP algorithm is significantly improved and an enhanced SCS-based greedy algorithm, structured *a priori* aided SAMP (SPA-SAMP), is proposed, capable of recovering the IN with higher accuracy and robustness. The solution existence, convergence, and estimation error bound of the proposed algorithm are derived and guaranteed by theoretical analysis and proofs.

The rest of this paper is organized as follows. The models of the IN and the signal in the MIMO system are described in Section II. Section III presents the proposed SCS optimization formulation through the SMM method, and the SPA-SAMP algorithm for the IN estimation in MIMO-OFDM systems, constituting the main contribution of the paper. Simulation results are reported in Section IV to validate the proposed approach, followed by the conclusion in Section V.

$Notation$: Matrices and column vectors are denoted by boldface letters; $(\cdot)^\dagger$ and $(\cdot)^H$ denote the pseudoinversion operation and conjugate transpose; $\|\cdot\|_r$ represents the ℓ_r-norm operation; $\lceil x \rceil$ denotes the smallest integer no less than x; $\odot$ represents the circular convolution operator; $\otimes$ is the Kronecker product operation of matrices; $|\Pi|$ denotes the cardinality of the set Π; $\mathbf{v}|_\Pi$ denotes the entries of the vector $\mathbf{v}$ in the set of Π; $\mathbf{A}_\Pi$ represents the submatrix comprised of the Π columns of the matrix $\mathbf{A}$; $\mathbf{A}_j$ denotes the jth column of $\mathbf{A}$; $(\mathbf{A})_{i,j}$ denotes the entry at the ith row and jth column of matrix $\mathbf{A}$; Π^c denotes the complementary set of Π; $\mathrm{vec}(\mathbf{A})$ denotes the vectorization operation, which returns a vector consisting of all the columns of $\mathbf{A}$; and $\mathrm{Max}(\mathbf{v}, T)$ denotes the indices of the T largest entries of the vector $\mathbf{v}$.

II. System Model

A. Statistical Model of IN

The IN vector corresponding to the ith OFDM symbol is denoted by $\boldsymbol{\xi}_i = [\xi_{i,0}, \xi_{i,1}, \ldots, \xi_{i,N-1}]^T$ of length N. One of the fundamental features of IN is that the IN vector is sparse, with the support $\Pi = \{j \,|\xi_{i,j} \neq 0\ ,\ j = 0, 1, \ldots, N-1\}$ and sparsity level $K = |\Pi|$. The interference-to-noise ratio (INR) γ of the IN is represented by $\gamma = \mathbb{E}\{P\}/\sigma^2$, where $P = \sum_{j\in\Pi} |\xi_{i,j}|^2/K$ is the IN average power and σ^2 is the variance of additive white Gaussian noise (AWGN). The statistical properties of the instantaneous amplitude and the random occurrences of the IN have been empirically modeled in the literature, mainly including the Gaussian mixture model [22] and the Middleton's Class A model [23].

The instantaneous amplitude of the time-domain asynchronous IN can be modeled by the Gaussian mixture distribution [22], with the probability density function (pdf) given by

$$p_Z(z) = \sum_{j=0}^{J_m} \beta_j \cdot g_j(z) \tag{1}$$

where $g_j(z)$ is the pdf of Gaussian distribution with zero mean and variance of σ_j^2, β_j is the mixture coefficient of the corresponding Gaussian pdf, and J_m is the number of Gaussian components.

The Middleton's Class A model is a common statistical model of the IN with the parameters of the overlapping factor A and the background-to-IN power ratio ω [23]. Gaussian mixture distribution can generate the special case of Middleton's Class A distribution when the parameters $\beta_j = e^{-A}A^j/j!$ and $\sigma_j^2 = (j/A + \omega)/(1+\omega)$ as $J_m \to \infty$.

The arrival rate of the IN bursts, i.e., the number of the IN bursts per second, follows Poisson process, and is given by the probability

$$P(\Lambda) = \lambda^\Lambda e^{-\lambda}/\Lambda! \tag{2}$$

where λ denotes the rate of IN arrivals [24]. The Gaussian mixture model and the Poisson arrival process are adopted in this paper.

The IN also has an important characteristic among different receivers in the MIMO system, the spatial correlation, because the IN signals at different receive antennas occur in a burst manner in the time domain. They are highly correlated and will simultaneously have an impact on different receive antennas in the MIMO system [19], [20]. Since the spatial scale of receive antennas are small enough, the IN contaminations on each of

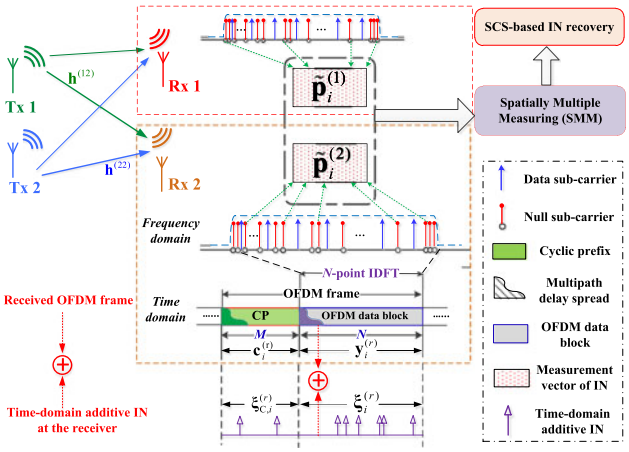

Fig. 1. Time-frequency OFDM frame structure exploited by the SMM method for the SCS-based IN recovery and cancelation.

them take place at the same time samples. Hence, the spatial correlation indicates that the supports (the positions of the nonzero entries) of the IN signals at different receive antennas are the same, i.e. $\Pi_{(1)} = \Pi_{(2)} = \cdots = \Pi_{(N_\mathrm{r})} = \Pi$, where $\Pi_{(j)}$ denotes the support of the IN at the jth receive antenna, although their amplitudes might be different from each other due to different channel conditions of different antennas.

B. Signal Model of MIMO-OFDM System

The signal frame and the frequency-domain structure of a typical MIMO system specified in the IEEE 802.11p (WAVE) standard [2] are shown in Fig. 1. Without loss of generality, the 2×2 MIMO system is illustrated in Fig. 1 and investigated in this paper, whereas the proposed scheme is also applicable in arbitrary $N_t \times N_\mathrm{r}$ MIMO systems, where N_t and N_r are the number of transmit and receive antennas, respectively. The null subcarriers, including the reserved tones and the virtual subcarrier masks specified in various communications standards such as the IEEE 802.11p WAVE standard, are utilized by the proposed SMM method to acquire the measurement matrix for SCS-based IN recovery. In the time domain, the ith transmitted frame at the tth transmit antenna consists of the OFDM block $\mathbf{x}_i^{(t)} = [x_{i,0}^{(t)}, x_{i,1}^{(t)}, \ldots, x_{i,N-1}^{(t)}]^T$ of length N and its cyclic prefix (CP) $\mathbf{c}_i^{(t)}$ of length-M. The ith time-domain OFDM block $\mathbf{x}_i$ is the inverse discrete Fourier transform (DFT) of the corresponding frequency-domain data $\tilde{\mathbf{X}}_i^{(t)}$ in the N subcarriers, which contains some reserved null subcarriers whose indices are denoted by the set Ω. All the N_t transmit antennas are sending the OFDM frames simultaneously, then passing through the wireless MIMO channel with the channel impulsive response $\mathbf{h}^{(tr)} = [h_0^{(tr)}, h_1^{(tr)}, \ldots, h_{L-1}^{(tr)}]^T$ between the tth transmit antenna and the rth receive antenna. At the receiver, the received ith time-domain OFDM frame at the rth receive antenna is denoted by $[(\mathbf{c}_i^{(r)})^T, (\mathbf{y}_i^{(r)})^T]^T$ consisting of the received CP $\mathbf{c}_i^{(r)}$ and the following OFDM data block $\mathbf{y}_i^{(r)}$, respectively, where the OFDM data block $\mathbf{y}_i^{(r)}$ is given by

$$\begin{aligned}\mathbf{y}_i^{(r)} &= [y_{i,0}^{(r)}, y_{i,1}^{(r)}, \ldots, y_{i,N-1}^{(r)}]^T \\ &= \sum_{t=1}^{N_t} \mathbf{h}^{(tr)} \odot \mathbf{x}_i^{(t)} + \boldsymbol{\xi}_i^{(r)} + \mathbf{w}_i^{(r)} \end{aligned} \tag{3}$$

where $\boldsymbol{\xi}_i^{(r)}$ and $\mathbf{w}_i^{(r)}$ denote the length-N time-domain IN signal additive to the rth receive antenna, and the time-domain AWGN vector, respectively. The operator $\odot$ denotes circular convolution. Note that the length-N time-domain IN signal is located at the received OFDM data block $\mathbf{y}_i^{(r)}$, but not related with the CP $\mathbf{c}_i^{(r)}$. Then, transforming the received signal to the frequency domain by DFT yields

$$\tilde{\mathbf{Y}}_i^{(r)} = \mathbf{F}_N \mathbf{y}_i^{(r)} = \sum_{t=1}^{N_t} \mathbf{H}^{(tr)} \tilde{\mathbf{X}}_i^{(t)} + \mathbf{F}_N \boldsymbol{\xi}_i^{(r)} + \tilde{\mathbf{w}}_i^{(r)} \tag{4}$$

where $\mathbf{F}_N$ and $\tilde{\mathbf{w}}_i^{(r)}$ denote the N-point DFT matrix and the frequency-domain AWGN vector at the rth receive antenna, respectively. The $N \times N$ channel frequency response matrix between tth transmit and rth receive antennas is denoted by $\mathbf{H}^{(tr)} = \mathrm{diag}\{\mathbf{F}_N \mathbf{h}^{(tr)}\}$. Recall that the $R = |\Omega|$ subcarriers of $\tilde{\mathbf{X}}_i^{(t)}$ corresponding to the set Ω are reserved to zero for all the transmit antennas, so we can select the R subcarriers corresponding to the set Ω out of $\tilde{\mathbf{Y}}_i^{(r)}$ using a selection matrix $\mathbf{S}_R$, yielding the measurement vector $\tilde{\mathbf{p}}_i^{(r)}$ as given by

$$\begin{aligned}\tilde{\mathbf{p}}_i^{(r)} &= \mathbf{S}_R \tilde{\mathbf{Y}}_i^{(r)} \\ &= \sum_{t=1}^{N_t} \mathbf{S}_R \mathbf{H}^{(tr)} \tilde{\mathbf{X}}_i^{(t)} + \mathbf{F}_\mathrm{R} \boldsymbol{\xi}_i^{(r)} + \mathbf{S}_R \tilde{\mathbf{w}}_i^{(r)} \\ &= \mathbf{0} + \mathbf{F}_\mathrm{R} \boldsymbol{\xi}_i^{(r)} + \tilde{\mathbf{w}}_{R,i}^{(r)} \end{aligned} \tag{5}$$

where the $R \times N$ selection matrix $\mathbf{S}_R$ is composed of the R rows from the $N \times N$ identity matrix $\mathbf{I}_N$ that correspond to the set Ω. The last equation in (5) holds because the entries of $\tilde{\mathbf{X}}_i^{(t)}$ corresponding to the set Ω are zero, so $\mathbf{S}_R \mathbf{H}^{(tr)} \tilde{\mathbf{X}}_i^{(t)} = \mathbf{0}$. The $R \times N$ partial DFT matrix $\mathbf{F}_\mathrm{R}$ is composed of the R rows of the N-point DFT matrix $\mathbf{F}_N$ that correspond to the set Ω. Then, the received data at the rth receive antenna in the null subcarriers set Ω can be rewritten briefly as

$$\tilde{\mathbf{p}}_i^{(r)} = \mathbf{F}_\mathrm{R} \boldsymbol{\xi}_i^{(r)} + \tilde{\mathbf{w}}_{R,i}^{(r)} \tag{6}$$

where the vector $\tilde{\mathbf{p}}_i^{(r)} = [\tilde{p}_{i,0}^{(r)}, \tilde{p}_{i,1}^{(r)}, \ldots, \tilde{p}_{i,R-1}^{(r)}]^T$ of length R is the measurement vector of the IN at the null subcarriers at the rth receive antenna, and $\tilde{\mathbf{w}}_{R,i}^{(r)}$ denotes the corresponding length-R frequency-domain AWGN vector. Consequently, the observation matrix utilized for the following SCS-based IN recovery is the partial DFT matrix $\mathbf{F}_\mathrm{R}$ given by

$$\mathbf{F}_\mathrm{R} = \frac{1}{\sqrt{N}} [\boldsymbol{\chi}_0 \quad \boldsymbol{\chi}_1 \quad \cdots \quad \boldsymbol{\chi}_{N-1}] \tag{7}$$

where the kth entry of $\boldsymbol{\chi}_m$ is $\exp(-j2\pi mk/N), k \in \Omega, m = 0, \ldots, N-1$. Note that the measurement vector $\tilde{\mathbf{p}}_i^{(r)}$ at the set

Ω only contains the components of IN and AWGN, whereas the information data component is not included since the null subcarriers are set to zeros.

III. SCS Formulation and Algorithm for IN Recovery

A. Existing CS-Based Measurements of IN for SISO Systems

According to the classical CS theory, it is feasible to recover the IN signal $\boldsymbol{\xi}_i$ for the SISO system using the measurement vector $\tilde{\mathbf{p}}_i$ given by (6) in the presence of background AWGN $\tilde{\mathbf{w}}_{R,i}$ [13]. To solve the CS measurement model in (6), one can equivalently relax it to a convex ℓ_1-norm minimization problem, which can be efficiently solved by state-of-the-art CS-based greedy algorithms, such as SAMP [18] and PA-SAMP [17]. However, these CS-based methods are only aimed at the SISO system and ignores the spatial correlation of the IN in the MIMO system, which might result in performance degradation, particularly when the intensity of the IN is fluctuating and its sparsity is large.

B. Proposed SCS Formulation for IN Recovery in MIMO-OFDM System: SMM of IN

To further improve the immunity to bad conditions, the SCS theory is introduced to extend our previous work [17] to MIMO-OFDM systems, leading to the proposed SMM method. In the $N_\mathrm{t} \times N_\mathrm{r}$ MIMO-OFDM system, the proposed SMM method exploits all the N_r measurement vectors given in (6) corresponding to the spatially correlated N_r receive antennas, yielding the SCS measurement model

$$\tilde{\mathbf{P}} = \left[\tilde{\mathbf{p}}_i^{(1)}, \tilde{\mathbf{p}}_i^{(2)}, \ldots, \tilde{\mathbf{p}}_i^{(N_\mathrm{r})}\right]_{R\times N_\mathrm{r}} = \mathbf{F}_\mathrm{R}\boldsymbol{\Xi}_0 + \tilde{\mathbf{W}} \tag{8}$$

where $\tilde{\mathbf{P}}$ is the measurements matrix composed of the N_r measurement vectors obtained by (6) from the N_r receive antennas. $\boldsymbol{\Xi}_0 = [\boldsymbol{\xi}_i^{(1)}, \boldsymbol{\xi}_i^{(2)}, \ldots, \boldsymbol{\xi}_i^{(N_\mathrm{r})}]_{N\times N_\mathrm{r}}$ is the spatially jointly sparse IN matrix whose columns share the same support Π due to the spatial correlation described in Section II-A, though the amplitudes of the nonzero values in the same row can be different for different receive antennas. $\tilde{\mathbf{W}} = [\tilde{\mathbf{w}}_{R,i}^{(1)}, \tilde{\mathbf{w}}_{R,i}^{(2)}, \ldots, \tilde{\mathbf{w}}_{R,i}^{(N_\mathrm{r})}]$ is the AWGN matrix.

The formulated mathematical model in (8) complies with the newly developed theory of SCS precisely [21]. Each column of the measurements matrix $\tilde{\mathbf{P}}$ is one measurement vector of the IN related with one receive antenna, and the spatial correlation is taken good advantage of by the SCS measurements model. Afterward, the multiple IN signals within $\boldsymbol{\Xi}_0$ that are jointly sparse can be simultaneously recovered by solving the nonlinear SCS optimization problem formulated as follows:

$$\hat{\boldsymbol{\Xi}} = \arg\min_{\boldsymbol{\Xi}\in\mathbb{C}^{N\times N_\mathrm{r}}} \|\boldsymbol{\Xi}\|_{0,q}, \quad \text{s.t.} \left\|\tilde{\mathbf{P}} - \mathbf{F}_\mathrm{R}\boldsymbol{\Xi}\right\|_{q,q} \le \varepsilon_S \tag{9}$$

where ε_S denotes the bound of the constraint in (9) due to the AWGN noise $\tilde{\mathbf{W}}$, and $\ell_{p,q}$-norm of the matrix $\boldsymbol{\Xi}$ is defined by

$$\|\boldsymbol{\Xi}\|_{p,q} = \left(\sum_m \left\|(\boldsymbol{\Xi}^T)_m\right\|_q^p\right)^{1/p} \tag{10}$$

with $(\boldsymbol{\Xi}^T)_m$ being the mth row of $\boldsymbol{\Xi}$. The SCS problem in (9) is a nonconvex and NP-hard one since the $\ell_{0,2}$-norm is adopted [21]. Fortunately, it can be relaxed to an equivalent convex optimization problem by adopting the mixed $\ell_{1,2}$-norm [21], yielding

$$\hat{\boldsymbol{\Xi}} = \arg\min_{\boldsymbol{\Xi}\in\mathbb{C}^{N\times N_\mathrm{r}}} \|\boldsymbol{\Xi}\|_{1,2}, \quad \text{s.t.} \left\|\tilde{\mathbf{P}} - \mathbf{F}_\mathrm{R}\boldsymbol{\Xi}\right\|_{2,2} \le \varepsilon_S \tag{11}$$

where based on (10) the mixed $\ell_{1,2}$-norm $\|\boldsymbol{\Xi}\|_{1,2}$ is explicitly given by

$$\|\boldsymbol{\Xi}\|_{1,2} = \sum_{m=1}^{N} \left\|(\boldsymbol{\Xi}^T)_m\right\|_2 \tag{12}$$

and thus it is derived that the error bound due to AWGN can be calculated by

$$\varepsilon_S = \|\tilde{\mathbf{W}}\|_{2,2} = \sqrt{RN_\mathrm{r}}\sigma. \tag{13}$$

Note that the previously described CS-based approach can be regarded as a special case of the proposed SCS-based framework with $N_\mathrm{r} = 1$ in (8) and (9). The SCS convex optimization problem (9) can be efficiently solved using the SCS-based greedy algorithms, such as simultaneous orthogonal matching pursuit (S-OMP) [25]. However, S-OMP requires the sparsity level to be known to reconstruct the sparse signal, which is impractical for the IN signal in realistic wireless MIMO systems. Moreover, there is no *a priori* information exploited in state-of-the-art methods. Hence, we propose the SCS-based greedy algorithm of SPA-SAMP to solve the mixed $\ell_{1,2}$-norm minimization problem and improve the performance against conventional algorithms.

C. Enhanced SCS-Based Greedy Algorithm for IN Recovery: SPA-SAMP

Unlike our previously proposed classical CS-based greedy algorithm [17] that copes with only one-dimensional measurement vector, and thus is not robust in severe conditions, the proposed SCS-based greedy algorithm SPA-SAMP further improves the performance by exploiting the spatial correlation of the IN signals at multiple receive antennas to maximize the accuracy of each iteration in the greedy pursuit process.

First, since the intensity of IN is normally much higher than that of data component or AWGN in the time domain, it is feasible to obtain a coarse *a priori* estimation of the partial support $\Pi^{(0)}$ of the IN signals at all the N_r receive antennas through thresholding. The indices of the time-domain samples in the ith received OFDM data blocks $\{\mathbf{y}_i^{(r)}\}_{r=1}^{N_\mathrm{r}}$ with the average power exceeding the given threshold λ_t are included in the partial support $\Pi^{(0)}$, which is given by

$$\Pi^{(0)} = \left\{ n \left| \frac{1}{N_\mathrm{r}} \sum_{r=1}^{N_\mathrm{r}} \left|y_{i,n}^{(r)}\right|^2 > \lambda_t, n = 0, 1, \ldots, N-1 \right.\right\} \tag{14}$$

where the power threshold λ_t is given by

$$\lambda_t = \frac{\alpha}{NN_\mathrm{r}} \sum_{r=1}^{N_\mathrm{r}} \sum_{n=0}^{N-1} \left|y_{i,n}^{(r)}\right|^2 \tag{15}$$

where α is a coefficient that can be configured large enough to ensure the accuracy of the time-domain partial support of the IN. Afterward, the *a priori* partial support estimation $\Pi^{(0)}$ can be exploited to facilitate the initialization and iterations of SPA-SAMP.

The SPA-SAMP algorithm is based on the relaxed convex SCS optimization problem in (11) by minimizing the $\ell_{1,2}$-norm of the IN spatially joint sparse matrix $\mathbf{\Xi}$. Its pseudocode is summarized in Algorithm 1. Specifically, the input includes the measurements matrix $\tilde{\mathbf{P}}$, the observation matrix $\mathbf{\Psi} = \mathbf{F}_{\mathrm{R}}$, and the *a priori* partial support $\Pi^{(0)}$, as well as the iteration step size Δs, which can be a compromise between the convergence rate and the accuracy. During the multiple iterations, the accuracy of the temporary support estimation $\Pi^{(k)}$ is improved at each iteration, and the testing sparsity level T is increased by Δs when the stage switches. The halting condition is determined by the $\ell_{2,2}$-norm of the residual matrix $\|\mathbf{R}\|_{2,2} \leq C_\varepsilon \cdot \varepsilon_S$ with C_ε given by (44) derived in the proof of Theorem 3.2. The output is the final support Π_{F} and the recovered IN jointly sparse matrix $\hat{\mathbf{\Xi}}$, s.t. $\hat{\mathbf{\Xi}}|_{\Pi_{\mathrm{F}}} = \mathbf{\Psi}^{\dagger}_{\Pi_{\mathrm{F}}}\tilde{\mathbf{P}}, \hat{\mathbf{\Xi}}|_{\Pi^c_{\mathrm{F}}} = \mathbf{0}$.

Afterward, the rth column of the recovered IN jointly sparse matrix $\hat{\mathbf{\Xi}}$ at the output of Algorithm 1, namely $\hat{\boldsymbol{\xi}}_i^{(r)}$, is exactly the estimation of the real time-domain IN signal $\boldsymbol{\xi}_i^{(r)}$ corresponding to the ith OFDM symbol at the rth receive antenna. Then, this recovered IN signal can be canceled out from the received OFDM symbol before the following demapping and decoding.

The overall structure and explanations of each step of Algorithm 1 are described as follows.

Phase 1—Input before the algorithm. The *a priori* estimated support $\Pi^{(0)}$, the initial sparsity level $K^{(0)} = |\Pi^{(0)}|$, and the multiple measurements matrix $\tilde{\mathbf{P}}$ are input into the algorithm.

Phase 2—Initialization phase. The initially estimated IN matrix is set by $\hat{\mathbf{\Xi}}^{(0)}|_{\Pi^{(0)}} \leftarrow \mathbf{\Psi}^{\dagger}_{\Pi^{(0)}}\tilde{\mathbf{P}}$, and the initial testing sparsity level is set by $T \leftarrow K^{(0)} + \Delta s$.

Phase 3—Main iterations. The outer loop is a repetition of multiple stages, with each stage being composed of multiple iterations and a different testing sparsity level, and the outer loop terminates until the halting condition of the algorithm is met, i.e., $\|\mathbf{R}\|_{2,2} \leq C_\varepsilon \cdot \varepsilon_S$. Within the inner loop of each stage, the iterations process includes the following.

1) Preliminary test (Lines 5 and 6), where the atoms corresponding to the largest $T - K^{(0)}$ entries are generated by the residue matrix projection onto the dictionary $\left(\mathbf{\Psi}^H\mathbf{R}^{(k-1)}\right)_{l,j}$, is chosen as the preliminary list S_k.
2) Candidate list (Line 7), where the candidate support list L_k for the current iteration is made by aggregating the preliminary test list S_k and the temporary final support of the previous iteration $\Pi^{(k-1)}$.
3) Temporary final list (Line 8), where the temporary final support for the current iteration $\Pi^{(k)}$ is formed by selecting the largest T entries out of the projection of the measurements matrix $\tilde{\mathbf{P}}$ onto the plane spanned by the subset of the dictionary $\mathbf{\Psi}_{L_k}$ corresponding to the candidate list L_k.
4) Computing estimation and residue (Lines 10 and 11), where the estimated IN matrix is calculated based on the least squares principle implemented on the temporary final support $\Pi^{(k)}$, and the residue matrix $\mathbf{R}$ is calculated using the estimated IN matrix.
5) Stage switching (Lines 12–17), where the stage is switched and the testing sparsity level T is increased by Δs when the $\ell_{2,2}$-norm of the residue matrix is greater than that of the previous iteration, otherwise the stage keeps the same and the iteration goes into the next.

Algorithm 1 (*SPA-SAMP*): Structured a Priori Aided Sparsity Adaptive Matching Pursuit for IN Recovery.

Input:
1) The a priori estimated support $\Pi^{(0)}$
2) Initial sparsity level $K^{(0)} = |\Pi^{(0)}|$
2) Measurements matrix $\tilde{\mathbf{P}}$
3) Observation matrix $\mathbf{\Psi} = \mathbf{F}_{\mathrm{R}}$
4) Step size Δs.

Initialization:

1: $\hat{\mathbf{\Xi}}^{(0)}\Big|_{\Pi^{(0)}} \leftarrow \mathbf{\Psi}^{\dagger}_{\Pi^{(0)}}\tilde{\mathbf{P}}$ (initially estimated IN matrix)
2: $\mathbf{R}^{(0)} \leftarrow \tilde{\mathbf{P}} - \mathbf{\Psi}\hat{\mathbf{\Xi}}^{(0)}$ (initial residue matrix)
3: $T \leftarrow K^{(0)} + \Delta s$ (initial testing sparsity level); $k \leftarrow 1$

Iterations:

4: **repeat**
5: $\quad \mathbf{v} \in \mathbb{C}^N \quad \text{s.t.} \quad v_l = \sum_{j=1}^{N_{\mathrm{r}}} \left|\left(\mathbf{\Psi}^H\mathbf{R}^{(k-1)}\right)_{l,j}\right|$
6: $\quad S_k \leftarrow \mathrm{Max}\{\mathbf{v}, T - K^{(0)}\}$ {Preliminary test}
7: $\quad L_k \leftarrow \Pi^{(k-1)} \cup S_k$ {Make candidate list}
8: $\quad \mathbf{u} \in \mathbb{C}^{|L_k|} \quad \text{s.t.} \quad u_l = \sum_{j=1}^{N_{\mathrm{r}}} \left|\left(\mathbf{\Psi}^{\dagger}_{L_k}\tilde{\mathbf{P}}\right)_{l,j}\right|$
9: $\quad \Pi^{(k)} \leftarrow \mathrm{Max}\{\mathbf{u}, T\}$ {Temporary final list}
10: $\quad \hat{\mathbf{\Xi}}^{(k)}\Big|_{\Pi^{(k)}} \leftarrow \mathbf{\Psi}^{\dagger}_{\Pi^{(k)}}\tilde{\mathbf{P}}; \ \ \hat{\mathbf{\Xi}}^{(k)}\Big|_{\Pi^{(k)^c}} \leftarrow \mathbf{0}$
11: $\quad \mathbf{R} \leftarrow \tilde{\mathbf{P}} - \mathbf{\Psi}_{\Pi^{(k)}}\mathbf{\Psi}^{\dagger}_{\Pi^{(k)}}\tilde{\mathbf{P}}$ {Compute residue}
12: $\quad$ **if** $\|\mathbf{R}\|_{2,2} \geq \left\|\mathbf{R}^{(k-1)}\right\|_{2,2}$ **then**
13: $\quad\quad T \leftarrow T + \Delta s$ {Stage switching}
14: $\quad$ **else**
15: $\quad\quad \Pi_{\mathrm{F}} \leftarrow \Pi^{(k)}; \ \ \mathbf{R}^{(k)} \leftarrow \mathbf{R}$
16: $\quad\quad k \leftarrow k + 1$ {Same stage, next iteration}
17: $\quad$ **end if**
18: **until** $\|\mathbf{R}\|_{2,2} \leq C_\varepsilon \cdot \varepsilon_S$

Output:
1) Final support Π_{F}
2) Recovered jointly sparse IN matrix $\hat{\mathbf{\Xi}}$, s.t. $\hat{\mathbf{\Xi}}\Big|_{\Pi_{\mathrm{F}}} = \mathbf{\Psi}^{\dagger}_{\Pi_{\mathrm{F}}}\tilde{\mathbf{P}}, \quad \hat{\mathbf{\Xi}}\Big|_{\Pi^c_{\mathrm{F}}} = \mathbf{0}$

Phase 4—Output. The output of the algorithm includes the final support and the recovered jointly sparse IN matrix corresponding to the final support.

It should be observed from Algorithm 1 that the spatial correlation of the IN signals in the MIMO system is fully exploited in SPA-SAMP by considering the aggregated contributions of all the N_{r} projections associated with N_{r} receive antennas when selecting the optimal candidate list in each iteration, instead of making the candidate list from only one measurement vector

as in state-of-the-art algorithms. Hence, the performance of IN recovery is significantly improved, which is reported in computer simulations. Furthermore, SPA-SAMP is more applicable in realistic wireless transmission scenarios than state-of-the-art SCS-based greedy algorithms, such as S-OMP that requires the sparsity level to be known [25].

The complexity of the proposed SPA-SAMP algorithm is analyzed as follows. For each iteration, two steps have to be considered, i.e., the inner product between the observation matrix $\mathbf{\Psi}$ and the residue $\mathbf{R}$ that has complexity $\mathcal{O}(N_{\rm r}RN)$, and the equivalent LS problem $\hat{\mathbf{\Xi}}^{(k)}|_{\Pi^{(k)}} \leftarrow \mathbf{\Psi}^{\dagger}_{\Pi^{(k)}}\tilde{\mathbf{P}}$ with complexity $\mathcal{O}(N_{\rm r}RK)$. The total number of iterations is reduced to $K - K^{(0)}$, so that the overall complexity of SPA-SAMP is in the order of $\mathcal{O}\left((K-K^{(0)})N_{\rm r}R(N+K)\right)$, which is not high in common cases since $K \ll N$ due to the sparse property of the IN signal.

D. Solution Existence and Convergence of the SPA-SAMP Algorithm

Before going into the details of the proofs in this section, one should be familiar with the basic definitions and terminologies in the CS and SCS theory, including but not limited to the following: K-sparse, restricted isometry property (RIP) and RIP-constant δ_K, block-K-sparse, block-RIP and block-RIP-constant $\delta_K^{(\rm B)}$, etc. [13]. These useful definitions have already been given in related literature, so the details of them are omitted in this paper for simplicity.

We first provide Lemma 3.1 to show that the output estimated support of any stage in SPA-SAMP is equivalent to that of the SP algorithm with corresponding sparsity level. Then, we provide two theorems to prove the solution existence and convergence of the proposed SPA-SAMP algorithm based on the lemma, with the RIP condition of the observation matrix $\mathbf{F}_{\rm R}$ and the block-RIP condition of the generated block-sparse observation matrix $\mathbf{\Psi}_{\rm B} = (\mathbf{F}_{\rm R} \otimes \mathbf{I}_{N_{\rm r}})$, where $\otimes$ is the Kronecker product operation of matrices.

Lemma 3.1: The output estimated final support of the ℓth stage in SPA-SAMP $\forall \ell \in [1, \ldots, \lceil K/\Delta s \rceil]$, is equivalent to that of the SP algorithm [26] with the sparsity level $K_\ell = \ell \cdot \Delta s$.

The proof of Lemma 3.1 is given in Appendix A. After this lemma, we derive the following two important theorems concerning about the convergence and performance error bound of the proposed SPA-SAMP algorithm. At the beginning, a lemma concerning about the iteration number within each stage is given as follows.

Lemma 3.2 (iteration number within stage): Let $n_{\rm it}(K_\ell)$ denote the number of iterations consumed by the ℓth iteration of SPA-SAMP, which has the testing sparsity level K_ℓ. One has

$$n_{\rm it}(K_\ell) \le \frac{1.5K_\ell}{-\log c_{K_\ell}} \tag{16}$$

where

$$c_{K_\ell} = \frac{2\delta_{3K_\ell}(1+\delta_{3K_\ell})}{(1-\delta_{3K_\ell})^3}. \tag{17}$$

Proof: According to [26, Th. 8], the inequality (16) holds for the SP with sparsity level K_ℓ, with (17) given by [26, eq. (8)]. Then, based on the proof of Lemma 3.1, the ℓth stage of SPA-SAMP is equivalent to SP with sparsity level K_ℓ, so the conclusion (16) also holds for the ℓth stage of SPA-SAMP, which completes the proof. ■

Theorem 3.1 (convergence without noise): If $\mathbf{F}_{\rm R}$ satisfies RIP with the RIP constant $\delta_{3K_s} < 0.165$, then SPA-SAMP converges to the exact IN matrix $\mathbf{\Xi}_0$ to be recovered as in $\tilde{\mathbf{P}} = \mathbf{F}_{\rm R}\mathbf{\Xi}_0$, which is the noiseless case of (8), after $\ell_{\max} = \lceil K/\Delta s \rceil$ stages of iterations, and the total number of iterations $N_{\rm IT}$ is bounded by

$$N_{\rm IT} \le \lceil K/\Delta s \rceil \cdot \left(\frac{1.5K_s}{-\log c_{K_s}}\right) \tag{18}$$

where

$$K_s = \lceil K/\Delta s \rceil \cdot \Delta s. \tag{19}$$

The proof of Theorem 3.1 is given in Appendix B. Theorem 3.1 guarantees the convergence of SPA-SAMP to the exact desired solution without background noise. When in the presence of noise, the convergence is guaranteed by the following theorem.

Theorem 3.2 (convergence with noise): If $\mathbf{F}_{\rm R}$ satisfies RIP with the constant $\delta_{3K_s} < 0.083$, $K_s = \Delta s\lceil K/\Delta s \rceil$, let the output estimated IN matrix at the end of the final $\ell_{\max} = \lceil K/\Delta s \rceil$th stage of SPA-SAMP be denoted by $\hat{\mathbf{\Xi}}_{(\ell_{\max})}$, then one has that $\hat{\mathbf{\Xi}}_{(\ell_{\max})}$ converges to the real IN matrix $\mathbf{\Xi}_0$ as in the multiple measurements model $\tilde{\mathbf{P}} = \mathbf{F}_{\rm R}\mathbf{\Xi}_0 + \tilde{\mathbf{W}}$ in (8), with the estimation error bounded by

$$\left\|\mathbf{\Xi}_0 - \hat{\mathbf{\Xi}}_{(\ell_{\max})}\right\|_{2,2} \le C'_{K_s}\left\|\tilde{\mathbf{W}}\right\|_{2,2} \tag{20}$$

where

$$C'_{K_s} = \frac{1+\delta_{3K_s}+\delta_{3K_s}^2}{\delta_{3K_s}(1-\delta_{3K_s})} \tag{21}$$

and the total number of iterations of SPA-SAMP $N_{\rm IT}$ is upper bounded by (18).

The proof of Theorem 3.2 is given in Appendix C. Till this end, the solution existence and the convergence of the proposed SPA-SAMP algorithm in both noiseless and noise cases are theoretically proved and guaranteed. The performance error bound of the estimated IN in the presence of background noise is also derived in closed form.

IV. Simulation Results and Discussions

The performance of the proposed SCS-based SMM method with the SPA-SAMP algorithm for IN recovery and cancelation in the MIMO wireless vehicular systems is evaluated through simulations. The simulation setup is basically configured in a wireless vehicular transmission scenario. The OFDM subcarrier number $N = 1024$ and the number of null subcarriers $R = 128$, and the length of CP is $M = 1/8N = 128$. The 16QAM (16 quadratic-amplitude modulation) scheme in low-speed vehicular scenario with the relative receiver velocity of 30 km/h is considered. To make the vehicular transmission more dependable,

TABLE I
PARAMETER PROFILE OF ITU-R VEHICULAR B MULTIPATH CHANNEL

Tap	Relative delay (ns)	Average power (dB)
1	0	−2.5
2	300	0
3	8900	−12.8
4	12 900	−10.0
5	17 100	−25.2
6	20 000	−16.0

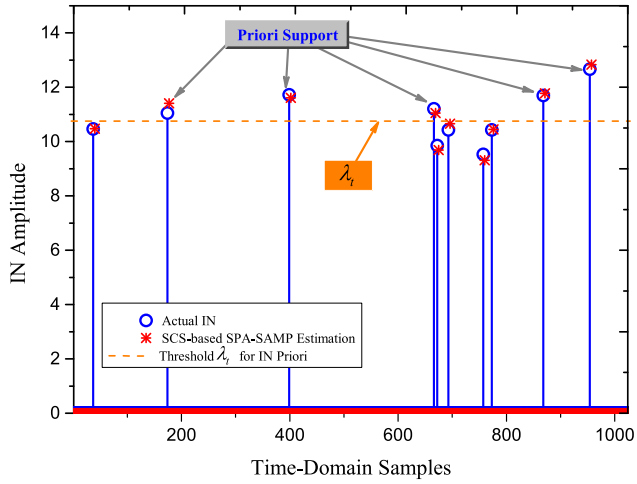

Fig. 2. Graphical visualization of the IN recovery using the proposed SCS-based SMM method with the SPA-SAMP algorithm.

the low density parity check code with code length of 1944 b and code rate of 1/2 as specified in [2] is adopted. The 2×2 MIMO system with spatial correlation at the receive antennas [27] in the presence of IN is adopted, where $N_t = N_r = 2$. The typical six-tap multipath channel model named by Vehicular-B specified by ITU-R is adopted, whose parameters are listed in Table I [28]. The coefficient for the partial support acquisition is $\alpha = 5.0$. The Gaussian mixture model of the IN is adopted with the parameters configured as $A = 0.15, \omega = 0.02, K = 10, \lambda = 50/\text{sec}$, and $\gamma = 30$ dB.

The performance of one realization of the IN recovery using the proposed SCS-based SMM method with the SPA-SAMP algorithm for the 2×2 MIMO system is depicted in Fig. 2. The partial support is first obtained using the threshold λ_t in (15). From the null subcarriers sets at the N_r receive antennas, we can get the measurements matrix from which the accurate IN is recovered using SPA-SAMP. It is observed from Fig. 2 that the final IN estimation matches the actual IN very well.

The mean square error (mse) performances of the proposed SMM method with the SCS-based SPA-SAMP algorithm, and the state-of-the-art CS-based algorithms (PA-SAMP [17] and SAMP [18]) for IN recovery in the 2×2 MIMO system are shown in Fig. 3. The theoretical Cramer–Rao lower bound (CRLB) $2\sigma^2 \cdot (N \cdot K/R)$ [29] is illustrated as benchmark. It can be observed that the proposed SCS-based SPA-SAMP algorithm achieves the mse of 10^{-3} at the INR of 23.2 dB and 32.0 dB with $K = 8$ and $K = 16$, respectively, which

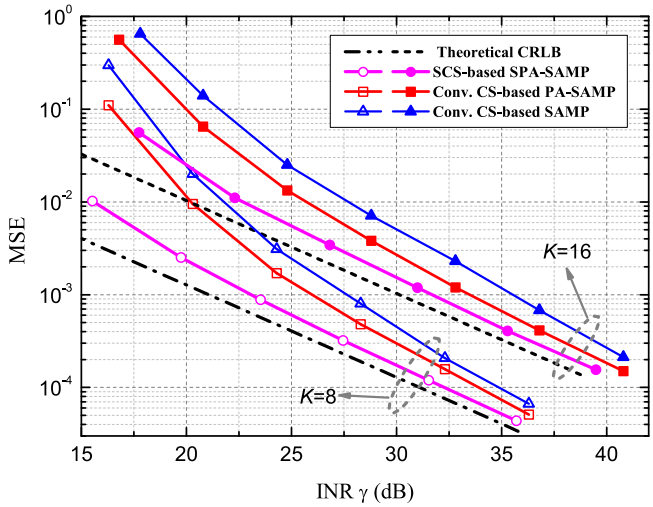

Fig. 3. MSE performance of the IN reconstruction using the proposed SCS-based SMM method with the SPA-SAMP algorithm compared with the state-of-the-art CS-based methods in the 2×2 MIMO system.

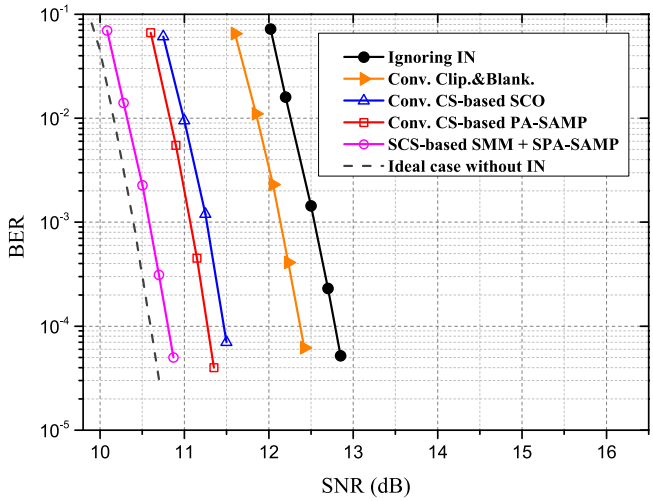

Fig. 4. BER performance of different schemes for IN mitigation and cancelation for the 2×2 MIMO system.

outperforms conventional PA-SAMP and SAMP algorithms by approximately 2.2 dB and 3.8 dB, respectively. It is noted from Fig. 3 that the mse of the proposed SPA-SAMP is asymptotically approaching the theoretical CRLB with the increase of the INR, verifying the high recovery accuracy. Moreover, the introduction of SCS will reduce the required length of measurement vector from $\mathcal{O}(K \log_2(N/K))$ in the standard CS down to $\mathcal{O}(K)$ in SCS [30], leading to slacker measurement requirement and higher spectral efficiency.

The bit error rate (BER) performances of different IN mitigation and cancelation schemes in the 2×2 MIMO system under the ITU-R Vehicular-B multipath channel are illustrated in Fig. 4, including the proposed SCS-based SMM method with the SPA-SAMP algorithm, as well as the conventional non-CS-based method (clipping and blanking [10]), and the state-of-the-art CS-based methods (SCO [16] and our previously proposed PA-SAMP [17]). The worst case ignoring IN

and the ideal case without IN are also depicted as benchmarks. It can be found that at the target BER of 10^{-4}, the proposed SCS-based scheme outperforms the conventional CS-bsed methods (SCO and PA-SAMP), the conventional non-CS-based method (clipping and blanking), and the case ignoring IN by approximately 0.7, 1.6, and 2.1 dB, respectively. Furthermore, the gap between the proposed method and the ideal case without IN is only about 0.2 dB, validating the accuracy and effectiveness of the proposed IN recovery scheme.

V. Conclusion

A novel SCS-based scheme of IN cancelation for MIMO systems is proposed in this paper. The gap of the lack of research on the IN elimination for MIMO systems is filled by the proposed SMM method along with the enhanced SCS-based greedy algorithm SPA-SAMP, which introduces the powerful emerging SCS theory into the area of IN mitigation for the first time. The spatial correlation of the IN signals at different receive antennas are fully exploited, leading to the measurements matrix and the formulation of the SCS convex optimization model. The IN signal is then reconstructed by the proposed SPA-SAMP algorithm, which is specifically designed and improved for the MIMO system. Furthermore, the proposed scheme is applicable and can be easily extended to other MIMO-OFDM-based broadband transmission systems in the presence of IN.

Appendix

A. Proof of Lemma 3.1

Proof: The key of this proof is to prove the equivalence of the key steps, i.e., the preliminary test (Lines 5 and 6 in Algorithm 1) and making temporary final list (Lines 8 and 9 in Algorithm 1), to those of the SP algorithm in [26], under the condition of the spatial correlation of the IN signals at different receive antennas. We have the following two propositions.

Proposition A.1 (preliminary test): The process of the preliminary test (Lines 5 and 6 in Algorithm 1) in SPA-SAMP as given by

$$S_k \leftarrow \text{Max}\{\mathbf{v} \in \mathbb{C}^N, T - K^{(0)}\} \tag{22}$$

$$\text{s.t. } v_l = \sum_{j=1}^{N_\text{r}} |(\mathbf{\Psi}^H \mathbf{R}^{(k-1)})_{l,j}|, \;\; l \in [1, \ldots, N] \tag{23}$$

is equivalent to the step 1) of the SP algorithm in [26], i.e.

$$S_k^{(j)} \leftarrow \text{Max}\{\mathbf{v}^{(j)}, T - K^{(0)}\} \tag{24}$$

$$\text{s.t. } v_l^{(j)} = |(\mathbf{\Psi}^H \mathbf{r}_j^{(k-1)})_l| \;\; \forall j \in [1, \ldots, N_\text{r}] \tag{25}$$

where $\mathbf{r}_j^{(k-1)} = \mathbf{R}_j^{(k-1)}$ is the jth column of $\mathbf{R}^{(k-1)}$.

Proposition A.2 (temporary final list): The process of making the temporary final list (Lines 8 and 9 in Algorithm 1) in SPA-SAMP as given by

$$\Pi^{(k)} \leftarrow \text{Max}\{\mathbf{u} \in \mathbb{C}^{|L_k|}, T\} \tag{26}$$

$$\text{s.t. } u_l = \sum_{j=1}^{N_\text{r}} |(\mathbf{\Psi}_{L_k}^\dagger \tilde{\mathbf{P}})_{l,j}|, \;\; l \in [1, \ldots, |L_k|] \tag{27}$$

is equivalent to the steps 2) and 3) of the SP algorithm in [26], i.e.

$$\Pi_{(j)}^{(k)} \leftarrow \text{Max}\{\mathbf{u}^{(j)}, T\} \tag{28}$$

$$\text{s.t. } u_l^{(j)} = |(\mathbf{\Psi}_{L_k^{(j)}}^\dagger \tilde{\mathbf{p}}_i^{(j)})_l|, \;\; \forall j \in [1, \ldots, N_\text{r}] \tag{29}$$

where $\Pi_{(j)}^{(k)}$ is the temporary final list of kth iteration for jth receive antenna.

Now we prove the above two propositions. For Proposition A.1, based on the *orthogonality property* of the least squares principle, the process of (24) and (25) can pick out the best atoms set (columns of $\mathbf{\Psi}$), i.e., $S_k^{(j)}$, which is the most probable to be included in the real support of the corresponding IN vector $\boldsymbol{\xi}_i^{(j)}$ out of all the atoms in $\mathbf{\Psi}$, through choosing the largest inner products between the residue vector $\mathbf{r}_j^{(k-1)}$ and the dictionary matrix $\mathbf{\Psi}$. By doing so, the selected preliminary test of kth iteration for jth receive antenna list is given by

$$S_k^{(j)} = \Pi_{(j)} - \Pi_{(j)}^{(k-1)} \bigcap \Pi_{(j)} \tag{30}$$

where $\Pi_{(j)}$ is the real support of the IN signal at jth receive antenna. Because of the spatial correlation $\Pi_{(j)} = \Pi, \forall j \in [1, \ldots, N_\text{r}]$, and $\Pi_{(j)}^{(k-1)} = \Pi^{(k-1)}$ due to the mathematical induction principle as presented later in this proof, one has

$$S_k^{(j)} = S_k \;\; \forall j \in [1, \ldots, N_\text{r}]. \tag{31}$$

Then, based on (31), it can be noted that the preliminary test list calculated from each process of (24) associated with the jth receive antenna yields the same atoms set S_k, hence, the process of (22), which picks out the atoms set by accumulating all the N_r residue vector inner products, is equivalent to (24). Thus, Proposition A.1 holds, which guarantees the equivalence of the preliminary test lists of the two algorithms.

For Proposition A.2, the process of solving the problem (28) and (29) is picking out the best T atoms set $\Pi_{(j)}^{(k)}$ out of the enlarged candidate list $L_k^{(j)}$, so that the selected set $\Pi_{(j)}^{(k)}$ (temporary final list) is corresponding to the largest T projections (coordinates) of $\tilde{\mathbf{p}}_i^{(j)}$ onto the spanned plane of the atoms (columns) in $\mathbf{\Psi}_{L_k^{(j)}}$. Hence, one has

$$\Pi_{(j)}^{(k)} = \Pi_{(j)} \bigcap L_k^{(j)} \tag{32}$$

where $L_k^{(j)} = S_k^{(j)} \bigcup \Pi_j^{(k-1)}$ is the enlarged candidate list with $|L_k^{(j)}| > T$, and from (31) along with the spatial correlation, one has

$$L_k^{(j)} = L_k, \;\; \forall j \in [1, \ldots, N_\text{r}] \tag{33}$$

$$\Pi_{(j)}^{(k)} = \Pi^{(k)}, \;\; \forall j \in [1, \ldots, N_\text{r}] \tag{34}$$

where (34) can be further induced in detail using the principle of mathematical induction for argument k, based on the margin fact that $\Pi_{(j)}^{(0)} = \Pi^{(0)}, \forall j$, due to the spatial correlation, along

with the inductions of (30) through (32). In fact, for $k = 1$, (30) through (34) hold obviously; for $k \geq 1$, if they hold for $k-1$, it is obvious that they also hold for k. Based on (34), it can be noted that the temporary final list calculated from each process of (28) associated with the jth receive antenna yields the same atoms set $\Pi^{(k)}$, hence, the process of (26), which picks out the atoms set by accumulating each of the $N_{\rm r}$ measurement vector projections, is equivalent to (28). Thus, Proposition A.2 holds, which guarantees the equivalence of making temporary final lists for the two algorithms.

From these two propositions, one can infer that, for the ℓth stage of SPA-SAMP (the corresponding sparsity level $K_\ell = \ell \cdot \Delta s$), the preliminary and temporary final lists for each iteration are the same with those of SP with K_ℓ, so the output estimated final supports of them are equal. Consequently, we reach the conclusion that the output estimated final support of the SPA-SAMP at the ℓth stage should be the same with that of SP with sparsity level $K_\ell = \ell \cdot \Delta s$. ■

B. Proof of Theorem 3.1

Proof: From [26, Th. 2], it was proved that the residue norm $\|\mathbf{r}^{(k)}\|_2 < \|\mathbf{r}^{(k-1)}\|_2, \forall k$ for any sparsity level K_ℓ if $\mathbf{F}_{\rm R}$ satisfies RIP with constant $\delta_{3K_s} < 0.165$. Based on Lemma 3.1, the ℓth stage of SPA-SAMP is equivalent to SP with sparsity level K_ℓ, so after each iteration of SPA-SAMP in the ℓth stage, the residue vector $\|\mathbf{r}_j^{(k)}\|_2 < \|\mathbf{r}_j^{(k-1)}\|_2, \forall k, j$. Thus, the $\ell_{2,2}$-norm of the residue matrix satisfies $\|\mathbf{R}^{(k)}\|_{2,2} < \|\mathbf{R}^{(k-1)}\|_{2,2}$ *within each stage* before the stage switching.

Recall that the condition of stage switching for each stage of SPA-SAMP is

$$\|\mathbf{R}^{(k)}\|_{2,2} \geq \|\mathbf{R}^{(k-1)}\|_{2,2}. \tag{35}$$

Based on Lemma 3.2, the number of iterations of the ℓth stage in SPA-SAMP is bounded by (16), so the stage switching condition (35) will be met for each stage in SPA-SAMP and each stage will definitely end at that time. Then, concerning about the solution convergence, we have the following proposition.

Proposition B.1: The global halting condition of SPA-SAMP without noise, which is given by

$$\left\|\mathbf{R}_{(\ell)}\right\|_{2,2} = \mathbf{0} \tag{36}$$

cannot be met within the stage $\ell, \ell < \lceil K/\Delta s \rceil$, and thus the stage will finally switch to next one. Until $\ell = \lceil K/\Delta s \rceil$, SPA-SAMP converges to the exact desired solution $\mathbf{\Xi}_0$ and the global halting condition in (36) is met.

Here, we denote the residue matrix and the estimated IN matrix at the end of the ℓth stage as $\mathbf{R}_{(\ell)}$ and $\hat{\mathbf{\Xi}}_{(\ell)}$, respectively. To prove Proposition B.1, first we use the proof of contradiction to prove the first part of it. Assume that $\exists \ell < \lceil K/\Delta s \rceil$ such that the residue matrix norm of the ℓth stage, $\left\|\mathbf{R}_{(\ell)}\right\|_{2,2} = \mathbf{0}$ is met during the ℓth stage. Since the sparsity level $K_\ell = \ell \cdot \Delta s < K_s$, ($K_s$ is the real sparsity level), so

$$\exists \mathbf{\Xi}' \neq \mathbf{0}, \text{ s. t. } \tilde{} \mathbf{\Xi}_0 - \hat{\mathbf{\Xi}}_{(\ell)} = \mathbf{\Xi}'. \tag{37}$$

Thus,

$$\begin{aligned}\left\|\mathbf{R}_{(\ell)}\right\|_{2,2} &= \left\|\tilde{\mathbf{P}} - \mathbf{F}_{\rm R}(\mathbf{\Xi}_0 - \mathbf{\Xi}')\right\|_{2,2} \\ &= \left\|(\tilde{\mathbf{P}} - \mathbf{F}_{\rm R}\mathbf{\Xi}_0) + \mathbf{F}_{\rm R}\mathbf{\Xi}'\right\|_{2,2} \\ &= \|\mathbf{F}_{\rm R}\mathbf{\Xi}'\|_{2,2} > 0\end{aligned}$$

which is contradictory to the assumption. Hence, the global halting condition in (36) cannot be met for $\ell < \lceil K/\Delta s \rceil$.

Then, based on [26, Th. 1], the SP algorithm with sparsity level K_s converges to the exact desired solution without noise since $\mathbf{F}_{\rm R}$ satisfies RIP with constant $\delta_{3K_s} < 0.165$. Thus, based on Lemma 3.1, the $\lceil K/\Delta s \rceil$th stage with testing sparsity level of K_s is equivalent to SP with sparsity level K_s, and it is guaranteed that SPA-SAMP converges to the exact solution $\mathbf{\Xi}_0$, i.e.

$$\hat{\mathbf{\Xi}}_{(\ell_{\max})} = \mathbf{\Xi}_0 \tag{38}$$

where $\ell_{\max} = \lceil K/\Delta s \rceil$, and $\hat{\mathbf{\Xi}}_{(\ell_{\max})}$ denotes the output of the final ($\ell_{\max}$th) stage. Then, it is evident that

$$\begin{aligned}\left\|\mathbf{R}_{(l_{\max})}\right\|_{2,2} &= \left\|\tilde{\mathbf{P}} - \mathbf{F}_{\rm R}\hat{\mathbf{\Xi}}_{(\ell_{\max})}\right\|_{2,2} \\ &= \left\|\tilde{\mathbf{P}} - \mathbf{F}_{\rm R}\mathbf{\Xi}_0\right\|_{2,2} = 0\end{aligned}$$

so Proposition B.1 holds.

It is evident that the total number of stages of SPA-SAMP is $\lceil K/\Delta s \rceil$ to reach the sparsity level K_s such that $K_s = \lceil K/\Delta s \rceil \cdot \Delta s \geq K$. For each stage $\ell \leq \lceil K/\Delta s \rceil$, due to Lemma 3.2 one has

$$n_{\rm it}(K_\ell) \leq \frac{1.5K_\ell}{-\log c_{K_\ell}} \leq \frac{1.5K_s}{-\log c_{K_s}} \; \forall \ell \leq \lceil K/\Delta s \rceil \tag{39}$$

where the second inequality holds because $K_\ell \leq K_s$, $\forall l \leq \lceil K/\Delta s \rceil$, and since the RIP constant δ_{3K_ℓ} in (17) is monotonically increasing with K_ℓ (which is easy to verify using a proof of contradiction based on the definition of RIP), we have that c_{K_ℓ} is monotonically increasing with ℓ. Hence, the total number $N_{\rm IT}$ of iterations of SPA-SAMP, including all $\ell_{\max}$ stages, is bounded by

$$N_{\rm IT} = \sum_{\ell=1}^{\ell_{\max} = \lceil K/\Delta s \rceil} n_{\rm it}(K_\ell) \leq \lceil K/\Delta s \rceil \cdot \left(\frac{1.5K_s}{-\log c_{K_s}}\right). \tag{40}$$

Then (18) holds, which completes the proof. ■

C. Proof of Theorem 3.2

Proof: Similarly to the proof of Theorem 3.1, according to [26, Th. 10], it was proved that the residue vector norm $\|\mathbf{r}^{(k)}\|_2 < \|\mathbf{r}^{(k-1)}\|_2, \forall k$ for any sparsity level K_ℓ, if $\mathbf{F}_{\rm R}$ satisfies RIP with constant $\delta_{3K_s} < 0.083$ in the presence of background noise. Then similarly, the residue matrix satisfies $\|\mathbf{R}^{(k)}\|_{2,2} < \|\mathbf{R}^{(k-1)}\|_{2,2}$ within each stage. The number of iterations of the ℓth stage is still bounded by (16) based on Lemma 3.2, so the stage switching condition in (35) will be met for each stage in SPA-SAMP.

According to Lemma 3.1, the $\ell_{\max}$th stage of SPA-SAMP is equivalent to SP with sparsity level of $K_{\ell_{\max}} = K_s$. Then, since $\mathbf{F}_{\mathrm{R}}$ satisfies RIP with the constant $\delta_{3K_s} < 0.083$, based on [26, Th. 9], one has

$$\left\|\boldsymbol{\xi}_i^{(j)} - \hat{\boldsymbol{\xi}}_{(l_{\max})}^{(j)}\right\|_2 \le C'_{K_s}\left\|\tilde{\mathbf{w}}_i^{(j)}\right\|_2 \quad \forall j \in [1,\ldots,N_{\mathrm{r}}] \tag{41}$$

where $\boldsymbol{\xi}_i^{(j)}$ and $\hat{\boldsymbol{\xi}}_{(l_{\max})}^{(j)}$ denote the jth column of matrices $\boldsymbol{\Xi}_0$ and $\hat{\boldsymbol{\Xi}}_{(\ell_{\max})}$, respectively, and the constant C'_{K_s} is given by (21), which is derived by [26, Th. 9]. Equation (41) indicates that the error between any IN vector of the estimated IN matrix and the real one is bounded by the noise power, and it holds for any receive antenna $j \in [1,\ldots,N_{\mathrm{r}}]$. Then, one has the error of the estimated IN matrix bounded by

$$\left\|\boldsymbol{\Xi}_0 - \hat{\boldsymbol{\Xi}}_{(\ell_{\max})}\right\|_{2,2} \le C'_{K_s}\left\|\tilde{\mathbf{W}}\right\|_{2,2} \tag{42}$$

since the $\ell_{2,2}$-norm of a matrix calculates the square root of the sum of all entries of the matrix as defined by (10), and thus (42) is derived from (41). Next, we raise a proposition similar to that in Theorem 3.1 as follows.

Proposition C.1: The global halting condition of SPA-SAMP with noise, which is given by

$$\left\|\mathbf{R}_{(\ell)}\right\|_{2,2} \le C_\varepsilon \cdot \varepsilon_S \tag{43}$$

where

$$C_\varepsilon = 1 + \sqrt{1+\delta_{2K_s}^{(\mathrm{B})}} \cdot C'_{K_s} \tag{44}$$

is met at the end of the final stage of $\ell_{\max} = \lceil K/\Delta s \rceil$, and the SPA-SAMP algorithm converges to the exact desired solution $\boldsymbol{\Xi}_0$ with error bounded by (42) after the $\ell_{\max}$th stage ($\delta_{2K_s}^{(\mathrm{B})}$ is the block-RIP constant).

To prove this proposition, we have known that the stage $\ell_{\max} = \lceil K/\Delta s \rceil$ of SPA-SAMP is equivalent to SP with sparsity level K_s based on Lemma 3.1, so after $\ell_{\max}$ stages of SPA-SAMP, the residue matrix $\mathbf{R}_{(\ell_{\max})}$ yields the following:

$$\left\|\mathbf{R}_{(\ell_{\max})}\right\|_{2,2} = \left\|\tilde{\mathbf{P}} - \mathbf{F}_{\mathrm{R}}\hat{\boldsymbol{\Xi}}_{(\ell_{\max})}\right\|_{2,2} \tag{45}$$

$$= \left\|\mathbf{F}_{\mathrm{R}}\left(\boldsymbol{\Xi}_0 - \hat{\boldsymbol{\Xi}}_{(\ell_{\max})}\right) + \tilde{\mathbf{W}}\right\|_{2,2} \tag{46}$$

$$\le \left\|\tilde{\mathbf{W}}\right\|_{2,2} + \left\|\mathbf{F}_{\mathrm{R}}\left(\boldsymbol{\Xi}_0 - \hat{\boldsymbol{\Xi}}_{(\ell_{\max})}\right)\right\|_{2,2} \tag{47}$$

$$\le \left\|\tilde{\mathbf{W}}\right\|_{2,2} + \sqrt{1+\delta_{2K_s}^{(\mathrm{B})}}\left\|\boldsymbol{\Xi}_0 - \hat{\boldsymbol{\Xi}}_{(\ell_{\max})}\right\|_{2,2} \tag{48}$$

$$\le \underbrace{\left[1 + \sqrt{1+\delta_{2K_s}^{(\mathrm{B})}} \cdot C'_{K_s}\right]}_{:=C_\varepsilon}\left\|\tilde{\mathbf{W}}\right\|_{2,2} \tag{49}$$

$$= C_\varepsilon \cdot \varepsilon_S \tag{50}$$

where (47) holds due to *Triangle Inequality*, (49) holds because of (42), and (50) is derived due to (13). The reason why (48) holds is as follows.

Since $\boldsymbol{\Xi}_0$ and $\hat{\boldsymbol{\Xi}}_{(\ell_{\max})}$ are block-K_s-sparse after vectorization (the number of rows with nonzero ℓ_2-norm are no more than K_s), then $(\boldsymbol{\Xi}_0 - \hat{\boldsymbol{\Xi}}_{(\ell_{\max})})$ is block-$2K_s$-sparse after vectorization because there are at most $2K_s$ rows with nonzero ℓ_2-norm (when the supports of $\boldsymbol{\Xi}_0$ and $\hat{\boldsymbol{\Xi}}_{(\ell_{\max})}$ are completely different). Let $\boldsymbol{\Xi}' = \boldsymbol{\Xi}_0 - \hat{\boldsymbol{\Xi}}_{(\ell_{\max})}$, then $\|\boldsymbol{\Xi}'\|_{2,2} = \left\|\mathrm{vec}\left(\boldsymbol{\Xi}'^T\right)\right\|_2$ since they both calculate the square root of the sum of all entries' powers. Assuming that the block observation matrix $\boldsymbol{\Psi}_{\mathrm{B}} = (\mathbf{F}_{\mathrm{R}} \otimes \mathbf{I}_{N_{\mathrm{r}}})$ satisfies the block-RIP with block-RIP constant $\delta_{2K}^{(\mathrm{B})}$, then one has

$$\|\mathbf{F}_{\mathrm{R}}\boldsymbol{\Xi}'\|_{2,2} = \left\|\mathrm{vec}\left[\left(\mathbf{F}_{\mathrm{R}}\boldsymbol{\Xi}'\right)^T\right]\right\|_2 \tag{51}$$

$$= \left\|(\mathbf{F}_{\mathrm{R}} \otimes \mathbf{I}_{N_{\mathrm{r}}})\mathrm{vec}\left(\boldsymbol{\Xi}'^T\right)\right\|_2 \tag{52}$$

$$\le \sqrt{1+\delta_{2K_s}^{(\mathrm{B})}}\left\|\mathrm{vec}\left(\boldsymbol{\Xi}'^T\right)\right\|_2 \tag{53}$$

$$= \sqrt{1+\delta_{2K_s}^{(\mathrm{B})}}\,\|\boldsymbol{\Xi}'\|_{2,2} \tag{54}$$

where (51) holds since both sides calculate the square root of the summation of the powers of all entries, and the reason is the same with (54); (52) holds due to the definitions of the vectorization process and the matrix Kronecker production; (53) holds because of the block-RIP of $(\mathbf{F}_{\mathrm{R}} \otimes \mathbf{I}_{N_{\mathrm{r}}})$.

Till now, we have proved Proposition C.1. Due to Proposition C.1, the SPA-SAMP algorithm with noise is guaranteed to reach the global halting condition (43) and will end at the $\ell_{\max}$th stage, and the estimated IN matrix $\hat{\boldsymbol{\Xi}}_{(\ell_{\max})}$ converges to the real IN $\boldsymbol{\Xi}_0$ with error bounded by (42). Besides, the total number of iterations N_{IT} is upper bounded by (18) with the similar proof to that in Theorem 3.1, thus omitted. Till now, the proof of this Theorem has been completed. ■

REFERENCES

[1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11n, Oct. 2009.

[2] *IEEE Standard for Information Technology—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments*, IEEE Standard 802.11p, Jul. 2010.

[3] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 1, pp. 249–258, Feb. 2002.

[4] R. Shepherd, J. Gaddie, and D. Nielson, "New techniques for suppression of automobile ignition noise," *IEEE Trans. Veh. Technol.*, vol. VT-25, no. 1, pp. 2–12, Feb. 1976.

[5] H. Suraweera and J. Armstrong, "Noise bucket effect for impulse noise in OFDM," *Electron. Lett.*, vol. 40, no. 18, pp. 1156–1157, Sep. 2004.

[6] D. Thompson and J. Dixon, "Vehicle noise," in *Advanced Applications in Acoustics, Noise and Vibration*. London, U.K.: Spon Press, 2004, pp. 236–291.

[7] A. B. Vallejo-Mora, J. J. Sanchez-Martinez, F. J. Canete, J. A. Cortes, and L. Diez, "Characterization and evaluation of in-vehicle power line channels," in *Proc. IEEE Global Telecommun. Conf.*, 2010, pp. 1–5.

[8] F. Juwono, Q. Guo, D. Huang, and K. P. Wong, "Deep clipping for impulsive noise mitigation in OFDM-based power-line communications," *IEEE Trans. Power Del.*, vol. 29, no. 3, pp. 1335–1343, Jun. 2014.

[9] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 145–147, Feb. 1996.

[10] S. Zhidkov, "Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers," *IEEE Trans. Commun.*, vol. 56, no. 1, pp. 5–9, Jan. 2008.

[11] A. Dubey, R. Mallik, and R. Schober, "Performance analysis of a power line communication system employing selection combining in correlated log-normal channels and impulsive noise," *IET Commun.*, vol. 8, no. 7, pp. 1072–1082, May 2014.
[12] F. Abdelkefi, P. Duhamel, and F. Alberge, "Impulsive noise cancellation in multicarrier transmission," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 94–106, Jan. 2005.
[13] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[14] Z. Han, H. Li, and W. Yin, *Compressive Sensing for Wireless Networks*. Cambridge, U.K.: Cambridge Univ., 2013.
[15] 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.
[16] T. Naffouri, A. Quadeer, and G. Caire, "Impulse noise estimation and removal for OFDM systems," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 976–989, Mar. 2014.
[17] S. Liu, F. Yang, W. Ding, and J. Song, "Double kill: Compressive-sensing-based narrow-band interference and impulsive noise mitigation for vehicular communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5099–5109, Jul. 2016.
[18] T. Do, G. Lu, N. Nguyen, and T. Tran, "Sparsity adaptive matching pursuit algorithm for practical compressed sensing," in *Proc. Conf. Signals, Syst. Comput.*, Asilomar, CA, USA, Oct. 2008, pp. 581–587.
[19] K. A. Saaifan and W. Henkel, "A receiver design for MIMO systems over Rayleigh fading channels with correlated impulse noise," in *Proc. IEEE Global Commun. Conf.*, 2012, pp. 2481–2486.
[20] P. Gao and C. Tepedelenlioglu, "Space-time coding over fading channels with impulsive noise," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 220–229, Jan. 2007.
[21] M. Duarte and Y. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.
[22] M. Nassar, K. Gulati, Y. Mortazavi, and B. Evans, "Statistical modeling of asynchronous impulsive noise in powerline communication networks," in *Proc. IEEE Global Telecommun. Conf.*, Houston, TX, USA, Dec. 2011, pp. 1–6.
[23] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. EMC-19, no. 3, pp. 106–127, Aug. 1977.
[24] N. Andreadou and F. Pavlidou, "Modeling the noise on the OFDM powerline communications system," *IEEE Trans. Power Del.*, vol. 25, no. 1, pp. 150–157, Jan. 2010.
[25] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part ii: Convex relaxation," *Signal Process.*, vol. 86, no. 3, pp. 589–602, Mar. 2006.
[26] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[27] G. J. Byers and F. Takawira, "Spatially and temporally correlated MIMO channels: Modeling and capacity analysis," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 634–643, May 2004.
[28] *Guideline for Evaluation of Radio Transmission Technology for IMT-2000*, Standard Recommendation ITU-R M. 1225, 1997.
[29] S. Liu, F. Yang, and J. Song, "Narrowband interference cancelation based on priori aided compressive sensing for DTMB systems," *IEEE Trans. Broadcast.*, vol. 61, no. 1, pp. 66–74, Mar. 2015.
[30] E. V. D. Berg and M. Friedlander, "Theoretical and empirical results for recovery from multiple measurements," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2516–2527, May 2010.

Sicong Liu (S'15) received the B.S.E. degree in electronic engineering from Tsinghua University, Beijing, China in 2012, where he is currently working toward the Ph.D. degree in electronic engineering in the DTV Technology R&D Center.

His research interests include broadband transmission techniques, noise and interference mitigation, and wireless and power line 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 an Associate Professor with the DTV Technology R&D Center, Tsinghua University. His research interests include power line communication, visible light communication, and digital television terrestrial broadcasting.

Xianbin Wang (S'98–M'99–SM'06) received the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2001.

He is a Professor and Canada Research Chair at Western University, London, ON, Canada. From July 2002 to December 2007, he was with the Communications Research Centre Canada. From January 2001 to July 2002, he was a System Designer with STMicroelectronics. His research interests include 5G networks, adaptive wireless systems, communications security, and locationing technologies.

Dr. Wang is an IEEE Distinguished Lecturer. He has received many awards and recognition, including Canada Research Chair, CRC President's Excellence Award, Canadian Federal Government Public Service Award, Ontario Early Researcher Award and three IEEE Best Paper Awards. He currently serves as an Editor/Associate Editor of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE WIRELESS COMMUNICATIONS LETTERS, and the IEEE TRANSACTIONS ON BROADCASTING. He was also an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS between 2007 and 2011. He was involved in a number of IEEE conferences including GLOBECOM, ICC, WCNC, VTC, ICME, and CWIT, in different roles such as symposium chair, tutorial instructor, track chair, session chair and TPC chair.

Jian Song (M'06–SM'10–F'16) received the B.Eng. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1990 and 1995, respectively.

He worked for Tsinghua University after his graduation. He was with The Chinese University of Hong Kong, Hong Kong, and the University of Waterloo, Waterloo, ON, Canada, in 1996 and 1997, respectively. He has been with the Hughes Network Systems in the United States for seven years before joining the faculty team in Tsinghua in 2005 as a Professor. He is currently the Director of Tsinghua's DTV Technology R&D Center, Tsinghua University. He has been working in different areas of fiber-optic, satellite, and wireless communications and power line communications. His current research focuses on digital TV broadcasting. He has published more than 200 peer-reviewed journals and conference papers. He holds two U.S. and more than 40 Chinese patents.

Dr. Song is a Fellow of IET.

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

From 2000 to 2002, he was an R&D Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. He is currently a Professor with the Electrical and Computer Engineering Department and the Computer Science Department at the University of Houston, Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid.

Dr. Han was the recipient of 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, the 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. He is currently an IEEE Communications Society Distinguished Lecturer.