IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 21, 2019, accepted September 10, 2019, date of publication September 13, 2019,
date of current version September 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2941233

Approximate Message Passing-Aided Iterative Channel Estimation and Data Detection of OFDM-IM in Doubly Selective Channels

LI WEI AND JIANPING ZHENG

State Key Laboratory of Integrated Services Network, Collaborative Innovation Center of Information Sensing and Understanding, School of Telecommunications Engineering, Xidian University, Xi'an 710071, China

Corresponding author: Jianping Zheng (jpzheng@xidian.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671340.

ABSTRACT The channel estimation and data detection of orthogonal frequency-division multiplexing with index modulation in doubly selective channels is a challenging problem due to the large number of channel parameters and the severe inter-carrier interference. In this paper, an iterative channel estimation and data detection scheme is proposed with the aid of the recent compressed sensing algorithm of approximate message passing (AMP). The time-varying channel is first approximated by basis expansion model (BEM) to reduce the number of channel parameters to be estimated, and AMP is utilized to estimate the BEM coefficients based on the pilot signals and soft information of data from the decoder. The channel is then reconstructed from the estimated BEM coefficients, and AMP is also employed to recover the transmit data based on the estimated channel and soft information of data from the decoder. The soft information of AMP detector is further sent to the decoder, which feeds soft information to both AMP estimator and AMP detector. To address the undesired measurement matrix in both channel estimation and data detection, two variants of AMP, i.e., damped AMP and vector AMP (VAMP), are utilized as practical estimator and/or detector here. Moreover, in both damped AMP/VAMP estimator and detector, the efficient utilization of soft data information is presented. Finally, simulation results are given to verify the performance of the proposed receiver.

INDEX TERMS OFDM, index modulation, approximate message passing, doubly selective channel, basis expansion model, pilot-assisted estimation, turbo receiver.

I. INTRODUCTION

Orthogonal frequency division multiplexing with index modulation (OFDM-IM) is a novel multi-carrier modulation technique [1]–[3]. In OFDM-IM, the N subcarriers of one OFDM symbol is partitioned into multiple blocks with IM being performed independently in each block. By IM, the information is conveyed by both specified subcarrier activation patterns (APs) and conventional signal constellation symbols, and the resulted signal is sparse. Due to this sparse signal character and the additional information conveying by subcarrier AP, OFDM-IM shows advantages over OFDM in several aspects such as energy efficiency and robustness to Doppler spread. Due to these potential advantages, OFDM-IM has attracted much attention. The recent research progresses include the performance analysis [4], [5], the methods and schemes to achieve coding gain [6], transmit diversity gain [7], [8], higher energy and/or spectral efficiency [9]–[11], to reduce the peak-to-average power ratio [12], the efficient signal detection [13], [14], and the extensions and variants to the parallel channel [15], the rapidly time-varying (RTV) channel [16], [17], and the multiple-input multiple-output (MIMO) channel [18]–[21]. Interested readers are referred to the recent surveys in [22]–[24] for further information.

Here, we focus on the performance of OFDM-IM in the doubly selective channel (DSC), i.e., channel with both frequency-selective and time-selective property. DSC is used typically to model high-mobility wireless communication and underwater acoustic communication. In this channel, channel estimation and data detection are two challenging problems. Firstly, due to the large amounts of channel parameters resulted by the time-varying channel property, channel

The associate editor coordinating the review of this manuscript and approving it for publication was Kezhi Wang.

estimation is tricky to handle. Further, the RTV will introduce Doppler effects which destroy the orthogonality between subcarriers when multicarrier modulation such as OFDM is utilized to overcome the frequency-selectivity [25]. Therefore, the channel matrix in the frequency domain is full matrix instead of a diagonal matrix, which increases the difficulty of signal detection. On the other hand, the AP constraint (part information is conveyed by specific APs) of IM signal also enhances the difficulty of efficient signal detection. In this paper, we study the challenging channel estimate and signal detection of OFDM-IM in the DSC.

A. RELATED WORK

Although the channel estimate of OFDM-IM over DSCs has not been reported in the literature, the channel estimate of OFDM over DSCs has attracted extensive research. An effective method to reduce the complexity of channel estimation is to introduce basis expansion model (BEM) [25]–[28]. Based on BEM, some effective channel estimation algorithms have been reported [27]–[36], [49]. In [26], several estimators were presented, including linear minimum mean square (LMMSE) estimator, least squares estimator and the best linear unbiased estimator (BLUE). In [29], the maximum-likelihood (ML) approach was applied to estimate the BEM coefficients. In [30], BEM was utilized with Bayesian approaches to capture the time-variation of the channel in OFDM transmissions over DSCs at low computational cost. In [31], an expectation maximization (EM)-based blind channel estimator with the framework of BEM for the OFDM system was proposed by utilizing the historical information of high speed railways. In [32], the authors considered the autoregressive model and BEM to estimate the carrier frequency offset and the channel based on EM algorithms. In [33], the authors adopted different kinds of BEM basis in sparse massive MIMO channels and acquire the channel side information (CSI) through prior information based block sparse Bayesian learning scheme. In [34], a sparse channel estimation scheme for massive MIMO-OFDM downlink transmission over time-varying channels was presented. This proposed scheme exploited the sparsity in the delay domain and the high correlation in the spatial domain. Furthermore, turbo iteration approach is usually utilized to improve the performance of channel estimation. In [35], a turbo iterative receiver consisting of LMMSE BEM channel estimation, LMMSE data detection and maximum a posterior (MAP) decoding was proposed. By utilizing the soft data information from decoder, the channel estimation becomes more accurate than that only use the pilot signals. In [36], more efficient belief propagation (BP) algorithm was utilized as the channel estimator in the turbo iterative approach. However, the complexity of BP is still high, since the factor graph constructed based on BEM is usually very dense. In [49], an efficient Turbo message passing framework for joint channel estimation and data detection was proposed, which made more efficient use of the prior information about the channel and the data for inter-carrier interference cancellation.

The signal detection of OFDM-IM over DSCs was first studied in [3], where the MMSE log-likelihood ratio (MMSE-LLR) detector, the block cancellation detector based on the successive interference cancellation (SIC) strategy, and the ordered SIC-based signal-power detector were proposed. In [37], we proposed the group-based SIC semidifinite relaxation detection to achieve a flexible performance-complexity tradeoff by integrating the AP constraint in the convex constraints of the equivalent convex programming problem. Moreover, in [37], a two-stage detection strategy was also proposed. In [38], we further proposed to utilize the recent compressed sensing algorithm of approximate message passing (AMP). The AP constraint is transferred into the prior probability of IM signal block and integrated in the posterior mean and variance update procedure of AMP detector. With a proper damping, the AMP detector shows superior performance in terms of both bit error rate (BER) and computational complexity.

B. OUR CONTRIBUTIONS

AMP algorithm [39]–[42] is a principal approximated framework for sum-product algorithm (SPA). By leveraging central limit theorem and other Taylor expansions, messages in AMP are manipulated associated with nodes instead of edges on the factor graph. This makes AMP very computationally efficient with predictable performance for a large Gaussian measurement matrix with independent and identically distributed (i.i.d.) entries [43], [44]. However, for general measurement matrix, the convergence and good performance of AMP can not be guaranteed [43], [44]. To deal with this issue, two feasible approaches are proposed. In [45], the damping strategy was introduced in AMP to address the compressive phase retrieval problem. In [46], [47], vector AMP (VAMP) algorithm was proposed, which extends AMP's guarantees from i.i.d. sub-Gaussian measurement matrix to the larger class of right-rotationally invariant matrix.

In the literature, the combination of AMP and BEM for channel estimation in OFDM-IM has not been reported. Motivated by the good signal recovery performance and computationally efficient implementation of damped AMP and VAMP for general measurement matrix, in this paper, we propose to extend our early work in [38] to turbo receiver [35], [36] consisting of channel estimator, data detector, and channel decoder. Specifically, damped AMP and VAMP are taken as either channel estimation or data detection. In the presented turbo receiver, the soft information transmitted among estimator, detector and decoder, the efficient utilization of soft information in estimator and detector, and message scheduling are studied.

The main contributions of this paper can be summarized as follows:

1) AMP-based channel estimator for OFDM-IM in DSCs. To the best of our knowledge, the application of AMP algorithm in the BEM-based channel estimation over DSCs has not been reported in the literature. Specifically, the implementations of damped AMP and VAMP

channel estimators are derived for BEM-based DSCs. It is found that the proposed AMP estimator can achieve nearly the same performance as LMMSE estimator with much lower computational complexity.

2) The VAMP data detector for OFDM-IM in DSCs is first introduced. The AMP-based turbo receiver of OFDM-IM over DSCs is studied. The soft information transmitted among estimator, detector and decoder, is properly designed. The efficient utilization of soft information in estimator and detector is given. Moreover, the message scheduling is also presented.
3) Computer simulation shows that the proposed AMP-based turbo iterative receiver with proper damping factors can achieve superior performance.

The rest of the paper is organized as follows. In Section II, signal model is first given. In Section III, the AMP and VAMP algorithms are briefly introduced. In Section IV, the proposed AMP-based turbo receiver is presented. Simulation results are given in Section V. Finally, Section VI concludes this paper.

Notation: $\mathbb{C}$ represents the field of complex numbers, $|\mathcal{S}|$ denotes the size of set $\mathcal{S}$, the superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\dagger$ indicate complex conjugate, transpose, Hermitian transpose, and Moore-Penrose inverse of the argument, respectively, $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian random variable with mean μ and variance σ^2, $C(n,k)$ is the binomial coefficient, and $\lfloor\cdot\rfloor$ denotes the floor function. The notation $\otimes$ represents the Kronecker product, $\mathrm{E}\{\cdot\}$ and $\mathrm{var}\{\cdot\}$ denote the expectation and variance, respectively, $diag\{\mathbf{x}\}$ denotes the diagonal matrix with the vector $\mathbf{x}$ as its diagonal. Further, we use x_n or $x(n)$ to indicate the n-th element of the vector $\mathbf{x}$, and $\mathbf{X}_{m,n}$ or $[\mathbf{X}]_{m,n}$ to indicate the (m,n)-th entry of the matrix $\mathbf{X}$, $\langle\mathbf{x}\rangle$ is the empirical averaging operation defined as $\langle\mathbf{x}\rangle := \frac{1}{N}\sum_{n=1}^{N} x_n$ for the N-dimensional argument $\mathbf{x}$. For two vectors $\mathbf{x}$ and $\mathbf{y}$ with the same size, $\mathbf{x}/\mathbf{y}$ denotes the component-wise division.

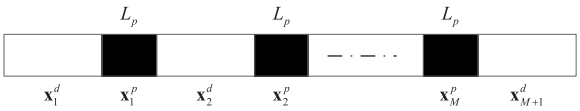


FIGURE 1. OFDM frame consisting of data signals and equidistant pilot clusters.

II. SIGNAL MODEL

Consider OFDM symbol with N subcarriers. The transmit signal $\mathbf{x}$ is composed of N_d-dimensional data symbol $\mathbf{x}^d$ and N_p-dimensional pilot symbol $\mathbf{x}^p$ with $N = N_d + N_p$. As shown in Fig. 1, we adopt the equidistant pilot cluster schedule, there are M pilot clusters of length L_p denoted as $\mathbf{x}_m^p$, $m = 0, \ldots, M-1$.

At the receiver, the received frequency domain (FD) singal $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \mathbf{W}_N \underbrace{\mathbf{R}_{CP}\mathbf{H}\mathbf{T}_{CP}}_{\mathbf{H}_t} \mathbf{W}_N^H \mathbf{x} + \boldsymbol{\omega} = \mathbf{H}_f \mathbf{x} + \boldsymbol{\omega} \tag{1}$$

Here, $\mathbf{W}_N$ is the normalized discrete Fourier transfer (DFT) matrix with $[\mathbf{W}_N]_{p,q} = 1/\sqrt{N}\exp(-j2\pi pq/N), p, q = 0, \cdots, N-1$, $T_{CP} \triangleq \left(\mathbf{I}_{CP}^T \mathbf{I}_N\right)^T$ is the $(N+L)\times N$ matrix that inserts the cyclic prefix (CP) with $\mathbf{I}_{CP}$ consisting of the last L rows of the identity matrix $\mathbf{I}_N$, $R_{CP} \triangleq (\mathbf{0}_{N\times L}\mathbf{I}_N)$ is the $N \times (N+L)$ matrix that removes CP, $\mathbf{H}_f \in \mathbb{C}^{N\times N}$ is FD channel matrix, and $\boldsymbol{\omega} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}\right) \in \mathbb{C}^N$ is additive noise. Furthermore, the time-domain (TD) channel $\mathbf{H} \in \mathbb{C}^{(N+L)\times(N+L)}$ is

$$\mathbf{H} = \begin{pmatrix} h_0^1 & & & & & & \\ h_1^2 & h_0^2 & & & & & \\ \vdots & & \ddots & & & & \\ h_{L-1}^L & & \ddots & h_0^L & & & \\ & \ddots & & & \ddots & & \\ & & h_{L-1}^{N+L} & h_{L-2}^{N+L} & \cdots & h_0^{N+L} & \end{pmatrix} \tag{2}$$

with $h_l^j \sim \mathcal{CN}\left(0, \sigma_l^2\right)$ denoting the channel gain of l-th tap at j-th time slot with $l = 0, \ldots, L-1, j = 1, \ldots, N+L$. $\mathbf{H}_t \in \mathbb{C}^{N\times N}$ is the TD channel matrix incorporating CP processing at transceiver. Here, we assumption that the maximum channel order is equal to the CP length, both denoted by L.

A. OFDM-IM

Here, the data signal $\mathbf{x}^d$ is generated by OFDM-IM. Specifically, the N_d data subcarriers in one OFDM symbol is divided into $G_0 = N_d/N_0$ blocks, each containing N_0 subcarriers. The length-m_0 corresponding information sequence is partitioned into $G_0 = m_0/p_0$ blocks as well, each having p_0 bits. In each block, only n_0 out of N_0 subcarriers are activated to transmit signals, the other $N_0 - n_0$ subcarriers are silent. The first p_1 bits are used to determine the indicies of the active subcarriers and the remaining $p_2 = p_0 - p_1$ bits are mapped into n_0 symbols of QAM/PSK constellation $\mathcal{X}$. Obviously, it has $p_1 = \left\lfloor \log_2 C(N_0, n_0) \right\rfloor$ and $p_2 = n_0 \log_2 |\mathcal{X}|$.

Denote the set of the indices of n_0 active subcarriers in the β-th, β= 1,2, $\cdots$,G_0 block by

$$I_\beta = \{i_{\beta,1}, i_{\beta,2}, \ldots, i_{\beta,n_0}\} \tag{3}$$

with $i_{\beta,\gamma} \in \{1,2,\ldots,N_0\}$, γ= 1,2,...,n_0. Correspondingly, the set of n_0 QAM/PSK symbols is denoted by

$$S_\beta = \{s_{\beta,1}, s_{\beta,2}, \ldots, s_{\beta,n_0}\} \tag{4}$$

with $s_{\beta,\gamma} \in \mathcal{X}$, γ= 1,2,...,n_0. Here, $s_{\beta,\gamma}$ is assumed to be energy-normalized, i.e., $\mathrm{E}\left\{|s_{\beta,\gamma}|^2\right\} = 1$. Then, the FD signal in this block can be expressed by

$$\mathbf{x}_\beta^d = \sum_{\gamma=1}^{n_0} s_{\beta,\gamma}\mathbf{e}_{i_{\beta,\gamma}} \tag{5}$$

where $\mathbf{e}_n$ is the N_0-dimensional all-zero vector except a one in position n. Finally, the data signal in one OFDM symbol

can be expressed by

$$\mathbf{x}^d = \left[\mathbf{x}_1^{d,T}, \ldots, \mathbf{x}_\beta^{d,T}, \ldots, \mathbf{x}_{G_0}^{d,T}\right]^T \tag{6}$$

with $\mathbf{x}_\beta^d \in \mathbb{C}^{N_0}$.

B. BEM MODEL

Consider the time-varying behavior of the l-th channel tap within one OFDM symbol in the $N \times 1$ vector, and stack all the channel taps within the block in a single $NL \times 1$ vector $\mathbf{h}^t := \left[h_0^1, \cdots, h_{L-1}^1, \cdots, h_0^N, \cdots, h_{L-1}^N\right]^T$. The BEM permits to express the vector as

$$\mathbf{h}^t = (\mathbf{B} \otimes \mathbf{I}_L)\,\mathbf{h} \tag{7}$$

where $\mathbf{B} := \left[\mathbf{b}_1, \ldots, \mathbf{b}_Q\right]$ is a $N \times Q$ matrix that collects Q ($Q \ll N$) orthonormal basis functions $\mathbf{b}_q$ as columns, $\mathbf{h}$ is the $QL \times 1$ vector that collects all the BEM coefficients of all the channel taps, which is represented as

$$\mathbf{h} := \left[h_{1,0}, \cdots, h_{1,L-1}, \cdots, h_{Q,0}, \cdots, h_{Q,L-1}\right]^T \tag{8}$$

with $\mathbf{h}_l := \left[h_{1,l}, \cdots, h_{Q,l}\right]^T$ representing the BEM coefficients for the l-th tap.

In light of the BEM, the input-output equation (1) can be rewritten as

$$\mathbf{y} = \sum_{q=1}^{Q} \mathbf{D}_q \mathbf{C}_q \mathbf{x} + \boldsymbol{\omega} \tag{9}$$

Here, $\mathbf{D}_q$ is a circulant matrix whose first column is the frequency response of the q-th basis function, i.e.,

$$\mathbf{D}_q = \mathbf{W}_N diag\{\mathbf{b}_q\}\mathbf{W}_N^H \tag{10}$$

and $\mathbf{C}_q$ is a diagonal matrix whose diagonal is the frequency response of the BEM coefficients corresponding to the q-th basis function, i.e.,

$$\mathbf{C}_q = diag\{\mathbf{F}_L[h_{q,0}, \ldots, h_{q,L-1}]^T\} \tag{11}$$

Here, $\mathbf{F}_L$ stands for the first L columns of the DFT matrix $\sqrt{N}\mathbf{W}_N$.

III. AMP AND VAMP

In this section, we briefly introduce the AMP algorithm and VAMP algorithm.

A. PROBLEM FORMULATION

Consider the problem of recovering an unknown vector $\mathbf{x} \in \mathbb{C}^N$ given the observation $\mathbf{y} \in \mathbb{C}^M$ and known measurement matrix $\mathbf{A}$, i.e.,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w} \tag{12}$$

where $\mathbf{w}$ is the noise vector. This recovery problem is known as compressed sensing when $M \ll N$ and $\mathbf{x}$ is sparse.

B. AMP ALGORITHM

AMP algorithm [39]–[42] is an iterative compressed sensing approach to recover $\mathbf{x}$ from measurements of the form (12). Algorithm 1 gives a brief summarization of AMP algorithm. In Algorithm 1, $\mathbf{g}_1(\cdot, \gamma)$ is the denoiser function, $\hat{\mathbf{x}}_k$ is the estimate of $\mathbf{x}$ in the k-th iteration, $\mathbf{v}_k$ is the residual vector. The term $\frac{N}{M}\alpha_{k-1}\mathbf{v}_{k-1}$ is known as the Onsager term, which makes AMP achieve guaranteed performance.

Algorithm 1 AMP [47]

1: **Initialization:** initialize $\mathbf{r}_0$, γ_0 as the mean and variance of $\mathbf{x}$, respectively, set $\mathbf{v}_{-1} = \mathbf{0}$ and $k = 0$.
2: $\hat{\mathbf{x}}_k = \mathbf{g}_1(\mathbf{r}_k, \gamma_k)$
3: $\alpha_k = \langle \mathbf{g}'_1(\mathbf{r}_k, \gamma_k) \rangle$
4: $\mathbf{v}_k = \mathbf{y} - \mathbf{A}\hat{\mathbf{x}}_k + \frac{N}{M}\alpha_{k-1}\mathbf{v}_{k-1}$
5: $\mathbf{r}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{A}^H\mathbf{v}_k$
6: Select γ_{k+1}
7: $k \leftarrow k + 1$

One of the important limitations of AMP is the susceptibility to the measurement matrix. AMP shows superiority with the large i.i.d. sub-Gaussian measurement matrix, however, it diverges under the more general ill-conditioned $\mathbf{A}$ [43], [44]. To overcome this problem partly, two feasible approaches are proposed. One is to introduce damping, and the resulted damped AMP [45] usually has improved performance over AMP for general measurement matrix. The other is the VAMP algorithm, which extends AMP's guarantees from i.i.d. sub-Gaussian measurement matrix to the larger class of right-rotationally invariant matrix.

C. VAMP ALGORITHM

VAMP algorithm shows stronger robustness to measurement matrix, it has many similarities with AMP algorithm. The main difference is the decoupling part, which relates to the vector and matrix computation instead of scalar computation as in AMP. In VAMP, it has to perform an initial singular value decomposition (SVD) of measurement matrix, and the per-iteration complexity of VAMP can be made similar to that of AMP [47].

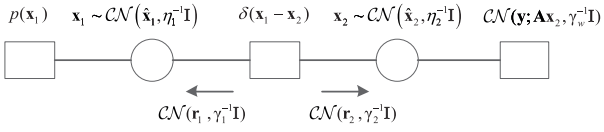


FIGURE 2. Factor graph presentation of VAMP algorithm based on variable splitting (13).

In VAMP, the unknown vector $\mathbf{x}$ is splitted into two identical variables $\mathbf{x}_1 = \mathbf{x}_2$, thus the equivalent factorization of equation (12) is represented as

$$p(\mathbf{y}, \mathbf{x_1}, \mathbf{x_2}) = p(\mathbf{x_1})\delta(\mathbf{x_1} - \mathbf{x_2})\mathcal{CN}(\mathbf{y}; \mathbf{A}\mathbf{x}_2, \gamma_w^{-1}\mathbf{I}) \tag{13}$$

where $\delta(\cdot)$ is the Dirac delta distribution and γ_w is the noise precision. The VAMP can be derived based on the factor graph constructed from (13), as shown in Fig. 2.

In Fig. 2, the belief of $\mathbf{x}_1 \sim \mathcal{CN}\left(\hat{\mathbf{x}}_1, \eta_1^{-1}\mathbf{I}\right)$ can be recognized as the posterior probability under likelihood

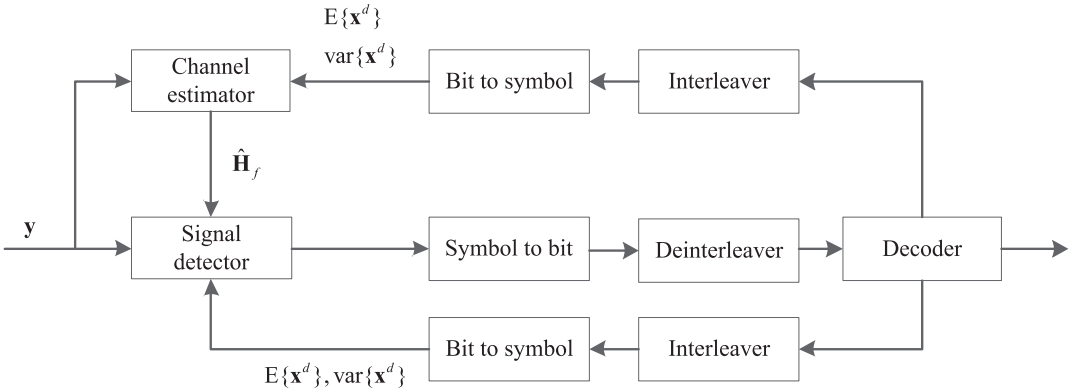


FIGURE 3. Architecture of turbo iterative receiver.

$\mathbf{x}_1 \sim \mathcal{CN}(\mathbf{r}_1, \gamma_1^{-1}\mathbf{I})$ and prior $\mathbf{x}_1 \sim \mathcal{CN}(\mathbf{x}_1; \mathbf{x}_{pri}, v_{pri}^{-1}\mathbf{I})$, which is similar to the AMP. The belief of $\mathbf{x}_2 \sim \mathcal{CN}\left(\hat{\mathbf{x}}_2, \eta_2^{-1}\mathbf{I}\right)$ can be recognized as the MMSE estimate of a random vector $\mathbf{x}_2$ under likelihood $\mathcal{CN}(\mathbf{y}; \mathbf{A}\mathbf{x}_2, \gamma_w^{-1}\mathbf{I})$ and prior $\mathbf{x}_2 \sim \mathcal{CN}(\mathbf{r}_2, \gamma_2^{-1}\mathbf{I})$. The damped VAMP is summarized in Algorithm 2, where $\mathbf{g}_1(\cdot, \gamma)$ is the denoiser function, and the MMSE estimation $\mathbf{g}_2(\mathbf{r}_{2,k}, \gamma_{2,k}) = (\gamma_w \mathbf{A}^H\mathbf{A} + \gamma_{2,k}\mathbf{I})^{-1}(\gamma_w \mathbf{A}^H\mathbf{y} + \gamma_{2,k}\mathbf{r}_{2,k})$.

Algorithm 2 Damped VAMP [47]

1: **Initialization:** initialize $\mathbf{x}_{pri}$ and v_{pri}. Select initial $\mathbf{r}_{1,0}$, $\gamma_{1,0}$ and set $k = 0$;
2: **Denoising**
3: $\hat{\mathbf{x}}_{1,k} = \mathbf{g}_1(\mathbf{r}_{1,k}, \gamma_{1,k})$
4: $\hat{\mathbf{x}}_{1,k} = \rho\hat{\mathbf{x}}_{1,k} + (1-\rho)\hat{\mathbf{x}}_{1,k-1}$
5: $\alpha_{1,k} = \langle \mathbf{g}'_1(\mathbf{r}_{1,k}, \gamma_{1,k})\rangle$
6: $\eta_{1,k} = \gamma_{1,k}/\alpha_{1,k}$
7: $\gamma_{2,k} = \eta_{1,k} - \gamma_{1,k}$
8: $\mathbf{r}_{2,k} = \gamma_{2,k}^{-1}\left(\hat{\mathbf{x}}_{1,k}\eta_{1,t} - \mathbf{r}_{1,k}\gamma_{1,k}\right)$
9: **LMMSE estimation**
10: $\hat{\mathbf{x}}_{2,k} = \mathbf{g}_2(\mathbf{r}_{2,k}, \gamma_{2,k})$
11: $\alpha_{2,k} = \langle \mathbf{g}'_2(\mathbf{r}_{2,t}, \gamma_{2,k})\rangle$
12: $\eta_{2,k} = \gamma_{2,k}/\alpha_{2,k}$
13: $\gamma_{1,k+1} = \eta_{2,k} - \gamma_{2,k}$
14: $\gamma_{1,k+1} = \rho\gamma_{1,k+1} + (1-\rho)\gamma_{1,k}$
15: $\mathbf{r}_{1,k+1} = \gamma_{1,k+1}^{-1}\left(\hat{\mathbf{x}}_{2,k}\eta_{2,k} - \mathbf{r}_{2,k}\gamma_{2,k}\right)$
16: $k \leftarrow k+1$

IV. AMP-BASED TURBO RECEIVER

A. TURBO RECEIVER STRUCTURE

The turbo receiver structure is shown in Fig. 3. It consists of three components: channel estimator, data detector and channel decoder. The information is exchanged among these components. The channel estimator leverages the received signals and soft estimates from the decoder to estimate the BEM coefficients. The estimated coefficients are used to reconstruct the FD channel. In conjunction with the reconstructed channel, the data detector leverages the received signals and soft estimates from the decoder to perform the refined data detection. Finally, the channel decoder leverages the soft information from data detector to recover the transmitted data, and the mean and variance of the data are sent back to channel estimator and data detector for further iteration. This procedure repeats till convergence or a maximum iteration number is achieved.

B. CHANNEL ESTIMATOR

Based on BEM model, we will only need to estimate QL BEM coefficients in $\mathbf{h}$ with the aid of pilots instead of directly estimating tricky $\mathbf{H}_t$, this achieves great computational convenience due to $QL \ll NL$. As shown in the Fig.3, the means and the variances of information symbols are updated in each iteration. The estimated mean of data symbol $\mathbf{x}^d$ is denoted as $\hat{\mathbf{x}}$= $\mathrm{E}\{\mathbf{x}^d\} = [\hat{x}_1, \ldots, \hat{x}_{N_d}]^T$ and the estimated variance is denoted as $\hat{\mathbf{v}}$= $\mathrm{var}\{\mathbf{x}^d\} = [\hat{v}_1, \ldots, \hat{v}_{N_d}]^T$. In each turbo iteration, they are updated using soft information from the channel decoder.

For m-th pilot cluster $\mathbf{x}_m^p = [x_{\mathcal{I}_m}^p, \cdots x_{\mathcal{I}_m+L_P-1}^p]^T$, $\mathcal{I}_m$ is the begin position of m-th pilot cluster. Considering $L_P + 2\Delta$ observation samples: $\mathbf{y}_m^{\mathcal{O}} = [y_{\mathcal{I}_m-\Delta}^p, \cdots y_{\mathcal{I}_m+L_P+\Delta-1}^p]^T$, where Δ is the smoothing parameter used to control the amount of interference taken into consideration for channel estimation. Obviously, the received FD samples include the information of both the pilots and the unknown data symbols, and the larger Δ is, the more data symbols are taken into consideration.

From (9), the m-th observation cluster $\mathbf{y}_m^{\mathcal{O}}$ can be expressed by, using BEM model,

$$\begin{aligned}\mathbf{y}_m^{\mathcal{O}} &= \sum_{q=1}^{Q}\mathbf{D}_{q,m}^p\mathbf{C}_q^p\mathbf{x}^p + \sum_{q=1}^{Q}\mathbf{D}_{q,m}^d\mathbf{C}_q^d\hat{\mathbf{x}}^d + \sum_{q=1}^{Q}\mathbf{D}_{q,m}^d\mathbf{C}_q^d\tilde{\mathbf{x}}^d + \boldsymbol{\omega}_m \\ &= \sum_{q=1}^{Q}\mathbf{D}_{q,m}^p\mathbf{C}_q^p\mathbf{x}^p + \sum_{q=1}^{Q}\mathbf{D}_{q,m}^d\mathbf{C}_q^d\hat{\mathbf{x}}^d + \tilde{\mathbf{d}}_m + \boldsymbol{\omega}_m \\ &= \sum_{q=1}^{Q}\mathbf{D}_{q,m}\mathbf{C}_q\hat{\mathbf{x}} + \mathbf{n}_m \end{aligned} \tag{14}$$

where $\mathbf{D}_{q,m}^{p}$ is a $(L_P+2\Delta)\times N_p$ matrix consisting of the $(L_P+2\Delta)$ rows of $\mathbf{D}_q$ with indices from $(\mathcal{I}_m-\Delta)$ to $(\mathcal{I}_m+L_p+\Delta-1)$ and columns corresponding to the positions of the pilot symbols $\mathbf{x}^p$. $\mathbf{D}_{q,m}^{d}$ is a $(L_P+2\Delta)\times N_d$ matrix consisting of the $(L_P+2\Delta)$ rows of $\mathbf{D}_q$ with indices from $(\mathcal{I}_m-\Delta)$ to $(\mathcal{I}_m+L_p+\Delta-1)$ and columns corresponding to the positions of the data symbols $\mathbf{x}^d$. $\mathbf{D}_{q,m}$ is an $(L_P+2\Delta)\times N$ matrix consisting of the $(L_P+2\Delta)$ rows of $\mathbf{D}_q$ with indices from $(\mathcal{I}_m-\Delta)$ to $(\mathcal{I}_m+L_p+\Delta-1)$. $\mathbf{C}_q^p$ and $\mathbf{C}_q^d$ are submatricies of the diagonal matrix $\mathbf{C}_q$ consisting of the diagonal elements on the positions of pilot and data symbols, respectively. $\tilde{\mathbf{x}}^d=\mathbf{x}^d-\hat{\mathbf{x}}^d$ is the residual interference. $\hat{\mathbf{x}}$ is the estimated symbols including the pilot symbols $\mathbf{x}^p$ and means of data symbols $\hat{\mathbf{x}}^d$. The equivalent noise $\mathbf{n}_m$ is the sum of residual data interference $\tilde{\mathbf{d}}_m=\sum_{q=1}^{Q}\mathbf{D}_{q,m}^{d}\mathbf{C}_q^d\tilde{\mathbf{x}}^d$ and noise $\boldsymbol{\omega}_m$.

With some algebra, the first term of (14) can be expressed by

$$\sum_{q=1}^{Q}\mathbf{D}_{q,m}\mathbf{C}_q\hat{\mathbf{x}}=\mathbf{D}_m\{\mathbf{I}_Q\otimes[diag(\hat{\mathbf{x}})\mathbf{F}_L]\}\mathbf{h}=\mathbf{A}_m\mathbf{h} \quad (15)$$

Analogously, it has $\tilde{\mathbf{d}}_m=\mathbf{D}_m^d\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}\mathbf{h}$ with $\mathbf{D}_m^d=[\mathbf{D}_{1,m}^d,\cdots,\mathbf{D}_{Q,m}^d]$. Stacking M observation clusters together, we obtain the received signal expressed by

$$\mathbf{y}^{\mathcal{O}}=\mathbf{D}^{\mathcal{O}}\{\mathbf{I}_Q\otimes[diag(\hat{\mathbf{x}})\mathbf{F}_L]\}\mathbf{h}+\mathbf{n}^{\mathcal{O}}=\mathbf{A}^{\mathcal{O}}\mathbf{h}+\mathbf{n}^{\mathcal{O}} \quad (16)$$

where $\mathbf{y}^{\mathcal{O}}$ is a $M(L_P+2\Delta)\times 1$ vector consisting of observation part in received signals, $\mathbf{A}^{\mathcal{O}}=\mathbf{D}^{\mathcal{O}}\{\mathbf{I}_Q\otimes[diag(\hat{\mathbf{x}})\mathbf{F}_L]$ with $\mathbf{D}^{\mathcal{O}}=[\mathbf{D}_1^T,\cdots,\mathbf{D}_M^T]^T$, $\mathbf{n}^{\mathcal{O}}=\tilde{\mathbf{d}}+\boldsymbol{\omega}^{\mathcal{O}}$ with $\tilde{\mathbf{d}}=\mathbf{D}^{d\mathcal{O}}\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}\mathbf{h}$, $\mathbf{D}^{d\mathcal{O}}=[\mathbf{D}_1^{dT},\cdots,\mathbf{D}_M^{dT}]^T$ and $\boldsymbol{\omega}^{\mathcal{O}}$ is additive noise. For notation simplicity, we set $\xi=M(L_p+2\Delta)$.

In (16), the distribution of equivalent noise is assumed to be $\mathbf{n}^{\mathcal{O}}\sim\mathcal{CN}(\mathbf{0},\mathbf{R}_{\mathbf{n}^{\mathcal{O}}})$ with

$$\mathbf{R}_{\mathbf{n}^{\mathcal{O}}}=\sigma^2\mathbf{I}_\xi+\mathbf{R}_{\tilde{d}} \quad (17)$$

where

$$\begin{aligned}\mathbf{R}_{\tilde{d}}&=\mathrm{E}_{\mathbf{h},\tilde{\mathbf{x}}^d}\{\tilde{\mathbf{d}}\tilde{\mathbf{d}}^H\}\\&=\mathrm{E}_{\mathbf{h},\tilde{\mathbf{x}}^d}\{\mathbf{D}^{d\mathcal{O}}\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}\mathbf{h}\mathbf{h}^H\\&\quad\times\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}^H\mathbf{D}^{d\mathcal{O}H}\}\\&=\mathbf{D}^{d\mathcal{O}}\mathbf{R}_{\tilde{\mathbf{x}}}\mathbf{D}^{d\mathcal{O}H} \quad (18)\end{aligned}$$

$$\begin{aligned}\mathbf{R}_{\tilde{x}}&=\mathrm{E}_{\mathbf{h},\tilde{\mathbf{x}}^d}\left\{\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}\mathbf{h}\mathbf{h}^H\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}^H\right\}\\&=\mathrm{E}_{\tilde{\mathbf{x}}^d}\left\{\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}\mathbf{R}_{\mathbf{h}}\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)\mathbf{F}_L]\}^H\right\}\\&=\mathrm{E}_{\tilde{\mathbf{x}}^d}\left\{\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)]\}\boldsymbol{\mathcal{H}}\{\mathbf{I}_Q\otimes[diag(\tilde{\mathbf{x}}^d)]\}^H\right\} \quad (19)\end{aligned}$$

Here, $\boldsymbol{\mathcal{H}}=(\mathbf{I}_Q\otimes\mathbf{F}_L)\mathbf{R}_h(\mathbf{I}_Q\otimes\mathbf{F}_L)^H$, and $\mathbf{R}_{\mathbf{h}}$ is the covariance matrix of BEM coefficients $\mathbf{h}$ which can be calculated by

$$\mathbf{R}_{\mathbf{h}}=\mathrm{E}_{\mathbf{h}}\{\mathbf{h}\mathbf{h}^H\}=\left(\mathbf{B}^\dagger\otimes\mathbf{I}_L\right)\mathbf{R}_{\mathbf{h}}^t\left(\mathbf{B}^\dagger\otimes\mathbf{I}_L\right)^H \quad (20)$$

with $\mathbf{R}_{\mathbf{h}}^t$ being the covariance matrix of the time domain channel $\mathbf{h}^t$. Furthermore, $\mathbf{R}_{\mathbf{h}}^t$ dependents on the channel statistic character and the normalized Doppler spread which are both assumed to be known at the receiver. In conjunction with the variance information $\hat{v}$ (var$\{x^d\}$ in Fig. 3) from channel decoder, it has

$$[\mathbf{R}_{\mathbf{x}}]_{m,n}=\begin{cases}\hat{v}_{\bmod(m,N_d)}[\boldsymbol{\mathcal{H}}]_{m,n}, & \text{if } \bmod(m-n,N_d)=0\\0, & \text{otherwise}\end{cases} \quad (21)$$

Given the equation (16), the tricky channel estimation problem is transformed to signal recovery problem. In [25] and [27], LMMSE channel estimation was derived as:

$$\hat{\mathbf{h}}=\mathbf{R}_{\mathbf{h}}\mathbf{A}^{\mathcal{O}H}(\mathbf{A}^{\mathcal{O}}\mathbf{R}_{\mathbf{h}}\mathbf{A}^{\mathcal{O}H}+\mathbf{R}_{\mathbf{n}^{\mathcal{O}}})^{-1}\mathbf{y}^{\mathcal{O}} \quad (22)$$

However, the LMMSE estimator requires to perform the complex matrix inverse computation and has a high computational complexity. Here, we present the application of computation-efficient AMP algorithm in this channel estimation problem.

1) AMP-BASED ESTIMATOR

AMP algorithm has been briefly introduced in Section III as Algorithm 1. Note that the entries of the measurement matrix $\mathbf{A}^{\mathcal{O}}$ in (16) are not i.i.d. Gaussian distributed, the damping strategy is employed to improve the performance of AMP estimator. Algorithm 3 summarizes the proposed damped AMP channel estimator, where $\rho\in(0,1]$ is the damping factor.

From Algorithm 3, the complexity of damped AMP channel estimator is $\mathcal{O}\left(T\left(12QL\xi+7QL+6\xi\right)\right)$ with $K,1\le K\le K_{\max}$ being the actual iteration number. On the other hand, the complexity of LMMSE estimator is $\mathcal{O}\left(\xi^3+(4QL-1)\,\xi^2+(4Q^2L^2-QL+1)\xi-QL\right)$ from (22). Therefore, in general, a numerous complexity reduction is achieved by using damped AMP estimator.

2) VAMP-BASED ESTIMATOR

As described in Section III-B, VAMP algorithm is another feasible method to recover the BEM coefficients from (16), which has a similar complexity with AMP algorithm [46], [47]. Analogously, to improve the performance for a general measurement matrix, the damping strategy can be utilized [47]. Here, a damping factor $\rho\in(0,1]$ is introduced in the VAMP channel estimation.

In the proposed damped VAMP channel estimator, firstly, compute the SVD $\mathbf{A}^{\mathcal{O}}=\mathbf{U}\mathbf{S}\mathbf{V}^H$ and the preconditioned vector $\tilde{\mathbf{y}}=\widetilde{\gamma}_w\mathbf{S}^H\mathbf{U}^H\mathbf{y}^{\mathcal{O}}$. Secondly, the belief of $\mathbf{h}_1$ is computed by employing the prior probability $\mathbf{h}_1\sim\mathcal{CN}(\mathbf{h}_1;\hat{\mathbf{h}}^0,\hat{\mathbf{v}}^0)$ and the likelihood probability of $\mathbf{h}_1\sim\mathcal{CN}(\mathbf{r}_1,\gamma_1^{-1}\mathbf{I})$ as

$$\mathbf{h}_1\sim\mathcal{CN}\left(\hat{\mathbf{h}},\eta_1^{-1}\mathbf{I}\right) \quad (23)$$

with

$$\eta_{1,k}=\frac{1}{QL}\sum\nolimits_i 1\Big/\hat{v}_i^0+\gamma_{1,k} \quad (24)$$

$$\hat{\mathbf{h}}_{1,k}=\eta_{1,k}^{-1}(\hat{\mathbf{h}}^0\Big/\hat{\mathbf{v}}^0+\mathbf{r}_{1,k}\gamma_{1,k}) \quad (25)$$

Algorithm 3 Damped AMP Estimator

1: **Initialization:** Initialize the mean and variance of $\mathbf{h}$: $\hat{h}_i^0 = 0$ and $\hat{v}_i^0 = [\mathbf{R}_h]_{ii}$, $i = 1, \ldots, QL$, $V_j^{-1} = 1$, $Z_j^{-1} = y_j^{\mathcal{O}}, j = 1, \ldots, \xi$, and $k = 0$.

2: **Decoupling step:** For $j = 1, \ldots, \xi$, compute

$$V_j^k = \rho \sum_i \left|\mathbf{A}_{ji}^{\mathcal{O}}\right|^2 \hat{v}_i^k + (1-\rho) V_j^{k-1}$$

$$Z_j^k = \sum_i \mathbf{A}_{ji}^{\mathcal{O}} \hat{h}_i^k - \frac{V_j^k}{[\mathbf{R}_{\mathbf{n}^{\mathcal{O}}}]_{jj} + V_j^{k-1}} (y_j^{\mathcal{O}} - Z_j^{k-1})$$

For $i = 1, \ldots, QL$, calculate

$$\Sigma_i^k = \left[\sum_j \frac{\left|\mathbf{A}_{ji}^{\mathcal{O}}\right|^2}{[\mathbf{R}_{\mathbf{n}^{\mathcal{O}}}]_{jj} + V_j^k} \right]^{-1}$$

$$\bar{h}_i^k = \rho \hat{h}_i^k + (1-\rho) \bar{h}_i^{k-1}$$

$$R_i^k = \bar{h}_i^k + \Sigma_i^k \sum_j \frac{\mathbf{A}_{ji}^{\mathcal{O}\,*} (y_j^{\mathcal{O}} - Z_j^k)}{[\mathbf{R}_{\mathbf{n}^{\mathcal{O}}}]_{jj} + V_j^k}.$$

Here $\rho = 1$ is used in the first iteration.

3: **Denoising step:** For $i = 1, \ldots, QL$, compute the posterior variance $\hat{v}_i^{k+1}$ and mean $\hat{h}_i^{k+1}$ using

$$\hat{v}_i^{k+1} = (1 \Big/ \hat{v}_i^0 + 1 \Big/ \Sigma_i^k)^{-1}$$

$$\hat{h}_i^{k+1} = (\hat{h}_i^0 \Big/ \hat{v}_i^0 + R_i^k \Big/ \Sigma_i^k) \hat{v}_i^{k+1}$$

4: Set $k \leftarrow k+1$ and proceed to step 2 until $k > K_{\max}$ or $\sum_i \left|\hat{h}_i^{k+1} - \hat{h}_i^k\right|^2 < \varepsilon_{toc} \sum_i \left|\hat{h}_i^k\right|^2$.

Thirdly, the prior probability of $\mathbf{h}_2$ is computed by $\mathbf{h}_2 \sim \mathcal{CN}(\mathbf{r}_2, \gamma_2^{-1}\mathbf{I})$, with

$$\gamma_{2,k} = \eta_{1,k} - \gamma_{1,k} \tag{26}$$

$$\mathbf{r}_{2,k} = \gamma_{2,t}^{-1} \left(\hat{\mathbf{h}}_{1,k} \eta_{1,k} - \mathbf{r}_{1,k} \gamma_{1,k} \right) \tag{27}$$

Fourthly, the belief of $\mathbf{h}_2 \sim \mathcal{CN}\left(\hat{\mathbf{h}}_{2,k}, \eta_2^{-1}\mathbf{I}\right)$ can be obtained through employing the prior probability and the likelihood information $\mathcal{CN}(\mathbf{y}^{\mathcal{O}}; \mathbf{A}^{\mathcal{O}}\mathbf{h}_2, \gamma_w^{-1}\mathbf{I})$ with $\widetilde{\gamma}_w = \frac{1}{\xi}\sum_i [\mathbf{R}_{n^{\mathcal{O}}}]_{ii}^{-1}$. Concretely, it has

$$\hat{\mathbf{h}}_{2,k} = \mathbf{V}(\widetilde{\gamma}_w \mathbf{S}^H \mathbf{S} + \gamma_{2,k}\mathbf{I})^{-1}(\tilde{\mathbf{y}} + \gamma_{2,k}\mathbf{V}^H \mathbf{r}_{2,k}) \tag{28}$$

$$\eta_2 = \frac{1}{R} \sum_r \widetilde{\gamma}_w s_r^2 + \gamma_{2,k}, s_r = [\mathbf{S}]_{rr} \tag{29}$$

where R is the rank of $\mathbf{A}^{\mathcal{O}}$, i.e., $R = \text{rank}(\mathbf{A}^{\mathcal{O}})$. Finally, update the likelihood probability of $\mathbf{h}_1 \sim \mathcal{CN}(\mathbf{r}_1, \gamma_1^{-1}\mathbf{I})$ using

$$\mathbf{r}_{1,k+1} = \left(\hat{\mathbf{x}}_{2,k} - \alpha_{2,k}\mathbf{r}_{2,k}\right) \Big/ \left(1 - \alpha_{2,k}\right) \tag{30}$$

$$\gamma_{1,k+1} = \gamma_{2,k}\left(1 - \alpha_{2,k}\right) \Big/ \alpha_{2,k} \tag{31}$$

$$\gamma_{1,k+1} = \rho\gamma_{1,k+1} + (1-\rho)\gamma_{1,k} \tag{32}$$

The detailed implementation of damped VAMP estimator is listed in Algorithm 4.

Algorithm 4 Damped VAMP Estimator

1: **Initialization:** initialize $\hat{h}_i^0 = 0$ and $\hat{v}_i^0 = [\mathbf{R}_{\mathbf{h}}]_{ii}$, $i = 1, \ldots, QL$,. Select initial $\mathbf{r}_{1,0}$, $\gamma_{1,0}$ and set $k = 0$;

2: Compute the SVD $\mathbf{A}^{\mathcal{O}} = \mathbf{U}\mathbf{S}\mathbf{V}^H$, $R = \text{rank}(\mathbf{A}^{\mathcal{O}})$ and the preconditioned vector $\tilde{\mathbf{y}} = \widetilde{\gamma}_w \mathbf{S}^H \mathbf{U}^H \mathbf{y}^{\mathcal{O}}$, $\widetilde{\gamma}_w = \frac{1}{\xi}\sum_i [\mathbf{R}_{n^{\mathcal{O}}}]_{ii}^{-1}$.

3: $\eta_{1,k} = \frac{1}{QL}\sum_i 1 \Big/ \hat{v}_i^0 + \gamma_{1,k}$

4: $\hat{\mathbf{h}}_{1,k} = \eta_{1,k}^{-1}(\hat{\mathbf{h}}^0 \Big/ \hat{\mathbf{v}}^0 + \mathbf{r}_{1,k}\gamma_{1,k})$

5: $\hat{\mathbf{h}}_{1,k} = \rho\hat{\mathbf{h}}_{1,k} + (1-\rho)\hat{\mathbf{h}}_{1,k-1}$

6: $\gamma_{2,k} = \eta_{1,k} - \gamma_{1,k}$

7: $\mathbf{r}_{2,k} = \gamma_{2,k}^{-1}\left(\hat{\mathbf{h}}_{1,k}\eta_{1,t} - \mathbf{r}_{1,k}\gamma_{1,k}\right)$

8: $\mathbf{D}_k = (\widetilde{\gamma}_w \mathbf{S}^H \mathbf{S} + \gamma_{2,k}\mathbf{I})^{-1}$

9: $\hat{\mathbf{h}}_{2,k} = \mathbf{V}\mathbf{D}_k(\tilde{\mathbf{y}} + \gamma_{2,k}\mathbf{V}^H \mathbf{r}_{2,k})$

10: $\alpha_{2,k} = \frac{1}{R}\sum_r \frac{\gamma_{2,k}}{\widetilde{\gamma}_w s_r^2 + \gamma_{2,k}}$, $s_r = [\mathbf{S}]_{rr}$

11: $\mathbf{r}_{1,k+1} = \left(\hat{\mathbf{h}}_{2,k} - \alpha_{2,k}\mathbf{r}_{2,k}\right) \Big/ \left(1 - \alpha_{2,k}\right)$

12: $\gamma_{1,k+1} = \gamma_{2,k}\left(1 - \alpha_{2,k}\right) \Big/ \alpha_{2,k}$

13: $\gamma_{1,k+1} = \rho\gamma_{1,k+1} + (1-\rho)\gamma_{1,k}$

14: Set $k \leftarrow k+1$ and proceed to line 3 until $k > K_{\max}$ or $\left\|\mathbf{r}_{1,k+1} - \mathbf{r}_{1,k}\right\|^2 < \varepsilon_{toc}\left\|\mathbf{r}_{1,k}\right\|^2$.

By performing damped AMP/VAMP channel estimation, we get the estimated BEM coefficient denoted as $\hat{\mathbf{h}}$. The estimated BEM coefficient is utilized to reconstruct the TD channel through $\mathbf{h}^t = (\mathbf{B} \otimes \mathbf{I}_L)\,\hat{\mathbf{h}}$, which is further transferred into the FD matrix $\hat{\mathbf{H}}_f$ by (1). Finally, the reconstructed FD channel is sent to detector to help the data detection.

C. DATA DETECTOR

Based on the signal system (1), we can present the input-output relation of data subcarriers as:

$$\mathbf{y}^d = \hat{\mathbf{H}}_f^d \mathbf{x}^d + \boldsymbol{\omega}^d \tag{33}$$

where $\mathbf{y}^d = \mathbf{y} - \hat{\mathbf{H}}_f^p \mathbf{x}^p$, $\hat{\mathbf{H}}_f^p$ is a $N \times N_p$ matrix consisting of the N_p columns corresponding to the positions of the pilot symbols in $\mathbf{x}$, $\hat{\mathbf{H}}_f^d$ is a $N \times N_d$ matrix consisting of the N_d columns corresponding to the positions of the pilot symbols in $\mathbf{x}$, and $\boldsymbol{\omega}^d$ is the corresponding additive noise.

1) AMP-BASED DETECTOR

Here, we extend the proposed AMP detector in [38] from the uncoded system to the turbo receiver with iterative estimator, detector, and decoder. The proposed damped AMP detector is listed in Algorithm 5.

In the initialization step, the initial mean and variance are obtained as follows. In the first global iteration, the initial values are computed according to the IM signal character (AP constraint) as presented in [38], since no soft information

can be obtained from channel decoder in this initial iteration. In the subsequent global iterations, the initial values are computed by the soft information from the channel decoder through a bit-to-symbol mapping as shown in Fig. 3.

In the denoising step, the posterior mean and variance are updated as follows. The posterior probability utilized to compute posterior mean and variance can be obtained by the product of prior probability and likelihood probability. In the first global iteration, the prior probability is computed according to the IM signal character (AP constraint) as presented in [38]. In the subsequent global iterations, the soft information from the channel decoder is taken as the updated prior information, and the posterior distribution of x_i can be computed by

$$q^k(x_i) \propto \mathcal{CN}(x_i; \hat{x}_i^0, \hat{v}_i^0)\mathcal{CN}(x_i; R_i^k, \Sigma_i^k) \tag{34}$$

where $\hat{x}_i^0$ and $\hat{v}_i^0$ are the prior mean and variance obtained from the soft information from the decoder. Then, the posterior variance and mean estimates of x_i can be calculated as

$$\hat{v}_i^{k+1} = (1/\hat{v}_i^0 + 1/\Sigma_i^k)^{-1} \tag{35}$$

$$\hat{x}_i^{k+1} = (\hat{x}_i^0/\hat{v}_i^0 + R_i^t/\Sigma_i^k)\hat{v}_i^{k+1} \tag{36}$$

Algorithm 5 Damped AMP Detector

1: **Initialization:** Initialize $\hat{x}_i^0$ and $\hat{v}_i^0$, $i = 1, \ldots, N_d$ by (6) and (7) in [38] for the first global iteration, and then initialized by the soft estimates from the decoder for the subsequent iterations, $V_j^{-1} = 1$, $Z_j^{-1} = y_j^d$, $j = 1, \ldots, N$, and $k = 0$.

2: **Decoupling step:** For $j = 1, \ldots, N$, compute

$$V_j^k = \rho \sum\nolimits_i |h_{ji}|^2 \hat{v}_i^k + (1-\rho)V_j^{k-1}$$

$$Z_j^k = \sum\nolimits_i h_{ji}\hat{x}_i^k - V_j^k(y_j^d - Z_j^{k-1}) \Big/ (\sigma^2 + V_j^{k-1})$$

For $i = 1, \ldots, N_d$, calculate

$$\Sigma_i^k = \left[\sum_j \frac{|h_{ji}|^2}{\sigma^2 + V_j^k}\right]^{-1}, \bar{x}_i^k = \rho\hat{x}_i^k + (1-\rho)\bar{x}_i^{k-1}$$

$$R_i^k = \bar{x}_i^k + \Sigma_i^k \sum\nolimits_a h_{ai}^*(y_a^d - Z_a^k) \Big/ (\sigma^2 + V_a^k)$$

Here $\rho = 1$ is used in the first global iteration.

3: **Denoising step:** For $i = 1, \ldots, N_d$, compute $\hat{x}_i^{k+1}$ and $\hat{v}_i^{k+1}$ using (15) and (16) in [38] for the first global iteration, and (35) and (36) for the subsequent iterations, respectively.

4: Set $k \leftarrow k+1$ and proceed to step 2) until $k > K_{\max}$ or $\sum_i \left|\hat{x}_i^{k+1} - \hat{x}_i^k\right|^2 < \varepsilon_{toc} \sum_i \left|\hat{x}_i^k\right|^2$.

2) VAMP-BASED DETECTOR

Here, we present the damped VAMP detector for the signal recovery problem of (33).

In the damped VAMP detector, the initialization is the same as that in damped AMP detector. The subsequent steps are as follows. Firstly, compute the SVD $\hat{\mathbf{H}}_f^d = \mathbf{USV}^H$ and preconditioned vector $\tilde{\mathbf{y}} = \gamma_w \mathbf{S}^H \mathbf{U}^H \mathbf{y}^d$. Given $\mathbf{r}_{1,t}$, $\gamma_{1,t}$ and prior probability $\mathbf{x}_1^d \sim \mathcal{CN}(\hat{\mathbf{x}}^0, \hat{\mathbf{v}}^0)$, compute the belief of $\mathbf{x}_1^d \sim \mathcal{CN}\left(\hat{\mathbf{x}}_1, \eta_1^{-1}\mathbf{I}\right)$ according to

$$\eta_{1,k} = \frac{1}{N_d}\sum\nolimits_i 1 \Big/ \hat{v}_i^0 + \gamma_{1,k} \tag{37}$$

$$\hat{\mathbf{x}}_{1,k} = \eta_{1,k}^{-1}(\hat{\mathbf{x}}^0 \Big/ \hat{\mathbf{v}}^0 + \mathbf{r}_{1,k}\gamma_{1,k}) \tag{38}$$

Secondly, compute the message from variable node $\mathbf{x}_1$ to factor node f_2: $\mu_{\mathbf{x}_1^d \to f_2}(\mathbf{x}_1^d) \sim \mathcal{CN}(\mathbf{r}_2, \gamma_2^{-1}\mathbf{I})$, with

$$\gamma_{2,k} = \eta_{1,k} - \gamma_{1,k} \tag{39}$$

$$\mathbf{r}_{2,k} = \gamma_{2,t}^{-1}\left(\hat{\mathbf{x}}_{1,k}\eta_{1,k} - \mathbf{r}_{1,k}\gamma_{1,k}\right) \tag{40}$$

Thirdly, compute the belief of $\mathbf{x}_2^d \sim \mathcal{CN}\left(\hat{\mathbf{x}}_2, \eta_2^{-1}\mathbf{I}\right)$. This can be recognized as the MMSE estimate of a random vector $\mathbf{x}_2^d$ under likelihood $\mathcal{CN}(\mathbf{y}; \mathbf{H}_f\mathbf{x}_2, \gamma_w^{-1}\mathbf{I})$, $\gamma_w^{-1} = \sigma^2$ and prior $\mathbf{x}_2^d \sim \mathcal{CN}(\mathbf{r}_2, \gamma_2^{-1}\mathbf{I})$. Thus, it has

$$\hat{\mathbf{x}}_{2,k} = \mathbf{V}(\gamma_w \mathbf{S}^H\mathbf{S} + \gamma_{2,k}\mathbf{I})^{-1}(\tilde{\mathbf{y}} + \gamma_{2,k}\mathbf{V}^H\mathbf{r}_{2,k}) \tag{41}$$

$$\eta_2 = \frac{1}{R}\sum\nolimits_r \gamma_w s_r^2 + \gamma_{2,k}, s_r = [\mathbf{S}]_{rr} \tag{42}$$

with $R = \text{rank}(\hat{\mathbf{H}}_f^d)$.

Finally, update the likelihood probability of $\mathbf{x}_1^d$ using

$$\mathbf{r}_{1,k+1} = \left(\hat{\mathbf{x}}_{2,k} - \alpha_{2,k}\mathbf{r}_{2,k}\right) \Big/ \left(1 - \alpha_{2,k}\right) \tag{43}$$

$$\gamma_{1,k+1} = \gamma_{2,k}\left(1 - \alpha_{2,k}\right) \Big/ \alpha_{2,k} \tag{44}$$

To improve the detection for general measurement matrix, as suggested in [47], the damping factor $\rho \in (0, 1]$ is introduced. The detailed implementation of damped VAMP detector is presented in Algorithm 6.

V. SIMULATION RESULTS

In this section, the bit error rate (BER) performance of the proposed damped AMP/VAMP channel estimator and AMP/VAMP detector is given by computer simulations.

We adopt the complex-exponential (CE) BEM [27] to model the time-varying channel. The number of subcarriers is $N = 128$, the Jakes model [48] is used to generate the RTV Rayleigh fading channels. The exponential power-delay profile is considered, i.e., the variance of the l-th channel tap is $\sigma_l^2 \propto \exp\left(-l/L\right)$. The OFDM-IM parameters are $N_0 = 4$, $n_0 = 2$, and QPSK modulator is used. We use code rate-1/3 $(133, 171, 165)_8$ convolutional code as channel code, and the code length is 6900. In both damped AMP/VAMP estimator and detector, $K_{max} = 15$ and $\varepsilon_{toc} = 10^{-12}$. The SNR is defined as E_b/σ_0^2 where $E_b = (N+L)/(G_0B_0)$ and $\sigma_0^2 = n_0\sigma^2/N_0$ is the noise variance in the time domain.

We adopt the frequency domain Kronecker delta (FDKD) pilot structure with equidistant placed pilot clusters. In each pilot cluster, there is only one nonzero symbol in the middle and two zero symbols surrounding it. That is, there are guard

Algorithm 6 Damped VAMP Detector

1: **Initialization:** Initialize $\hat{x}_i^0$ and $\hat{v}_i^0$, $i = 1, \ldots, N_d$ by (6) and (7) in [38] for the first global iteration, and then initialized by the soft estimates from the decoder for the subsequent iterations. Select initial $\mathbf{r}_{1,0}$, $\gamma_{1,0}$ and set $k = 0$;
2: Compute the SVD $\tilde{\mathbf{H}}_f^d = \mathbf{USV}^H$, $R = \text{rank}(\tilde{\mathbf{H}}_f^d)$ and the preconditioned vector $\tilde{\mathbf{y}} = \gamma_w \mathbf{S}^H \mathbf{U}^H \mathbf{y}^d$, $\gamma_w = 1/\sigma^2$.
3: $\eta_{1,k} = \frac{1}{N_d} \sum_i 1/\hat{v}_i^0 + \gamma_{1,k}$
4: $\hat{\mathbf{x}}_{1,k} = \eta_{1,k}^{-1}(\mathbf{x}^0/\mathbf{v}^0 + \mathbf{r}_{1,k}\gamma_{1,k})$
5: $\hat{\mathbf{x}}_{1,k} = \rho\hat{\mathbf{x}}_{1,k} + (1-\rho)\hat{\mathbf{x}}_{1,k-1}$
6: $\gamma_{2,k} = \eta_{1,k} - \gamma_{1,k}$
7: $\mathbf{r}_{2,k} = \gamma_{2,k}^{-1}\left(\hat{\mathbf{x}}_{1,k}\eta_{1,k} - \mathbf{r}_{1,k}\gamma_{1,k}\right)$
8: $\mathbf{D}_k = (\gamma_w \mathbf{S}^H\mathbf{S} + \gamma_{2,k}\mathbf{I})^{-1}$
9: $\hat{\mathbf{h}}_{2,k} = \mathbf{V}\mathbf{D}_k(\tilde{\mathbf{y}} + \gamma_{2,k}\mathbf{V}^H\mathbf{r}_{2,k})$
10: $\alpha_{2,k} = \frac{1}{R}\sum_r \frac{\gamma_{2,k}}{\gamma_w s_r^2 + \gamma_{2,k}}$, $s_r = [\mathbf{S}]_{rr}$
11: $\mathbf{r}_{1,k+1} = \left(\hat{\mathbf{x}}_{2,k} - \alpha_{2,k}\mathbf{r}_{2,k}\right)/\left(1-\alpha_{2,k}\right)$
12: $\gamma_{1,k+1} = \gamma_{2,k}\left(1-\alpha_{2,k}\right)/\alpha_{2,k}$
13: $\gamma_{1,k+1} = \rho\gamma_{1,k+1} + (1-\rho)\gamma_{1,k}$
14: Set $k \leftarrow k+1$ and proceed to step 2) until $k > K_{\max}$ or $\left\|\mathbf{r}_{1,k+1} - \mathbf{r}_{1,k}\right\|^2 < \varepsilon_{toc}\left\|\mathbf{r}_{1,k}\right\|^2$.

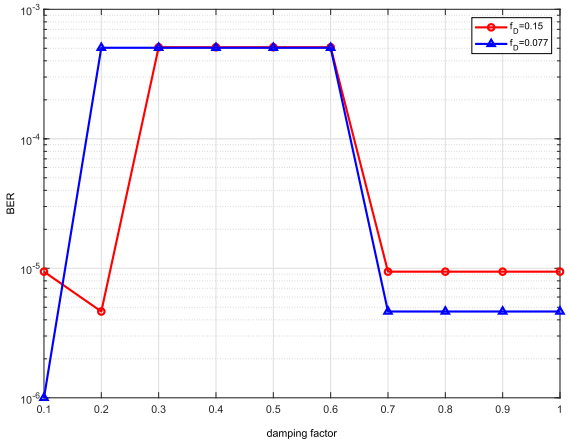


FIGURE 4. BER performance of damped AMP estimator with varying damping factors.

bands around the nonzero pilot in each cluster. $N_p = 36$ pilot symbols are inserted in the transmitted signal with $M = 12$, $L_p = 3$. In one OFDM symbol, only 92 subcarriers are used for data transmission. The number of BEM tap is $Q = 3$, the smoothing parameter is $\Delta = 0$ for the first iteration and $\Delta = 2$ for the subsequent iterations.

A. IMPACT OF DAMPING FACTOR

We adopt the $\rho = 0.95$ in both VAMP estimator and VAMP detector, which is suggested in [47] and also performs well in our simulation. Similarly, in damped AMP detector, the damping factor is set to be 0.3, which is used in [38] and performs well here. To design proper damping factor in damped AMP estimator, extensive simulations are performed. Fig. 4 gives the performance of damped AMP estimator with different damping factor at SNR = 8dB. From the figure, it can be observed that $\rho = 0.2$ is the best damping factor for the normalized Doppler spread $f_D = 0.15$ and $\rho = 0.1$ for $f_D = 0.077$. These damping factors are employed in the later simulations.

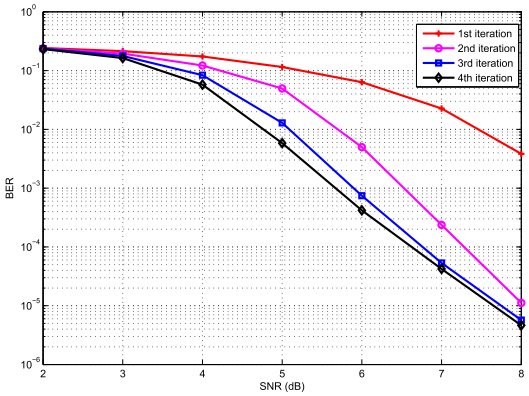


FIGURE 5. BER curves of turbo receiver for varying global iteration numbers at $f_D = 0.15$. Here, damped AMP estimator and damped AMP detector are used.

B. IMPACT OF GLOBAL ITERATION NUMBER

Fig. 5 gives BER curves of turbo receiver for varying global iteration numbers at $f_D = 0.15$. Here, damped AMP estimator and damped AMP detector are used. From this figure, a further iteration will improve the performance obviously when the global iteration numbers are 1 and 2, and the further increase of iteration after 3 iterations achieves only slight gain. Thus, in the later simulation, we set the global iteration number to be 3.

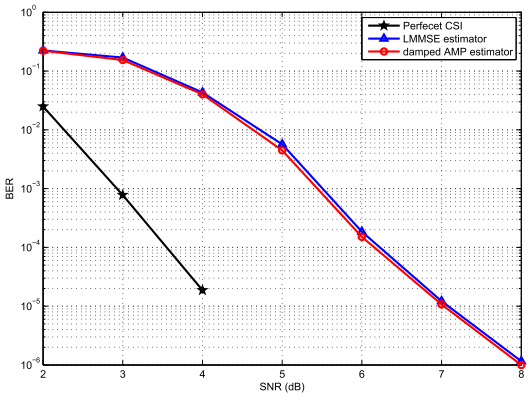


FIGURE 6. Performance comparison of LMMSE estimator and damped AMP estimator at $f_D = 0.077$.

C. COMPARISON OF DIFFERENT ESTIMATORS

In Fig. 6, the BER performance of LMMSE estimator and damped AMP estimator is given. As a reference, the performance with perfect channel is also presented. In this simulation, damped AMP detector is adopted and normalized Doppler frequency is 0.077. From the figure, the BER performance of the damped AMP estimator is close to the

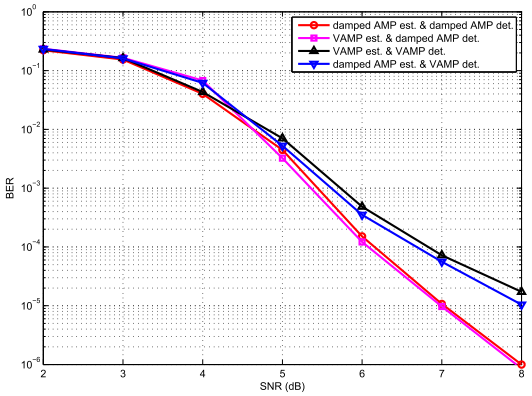


FIGURE 7. BER performance of four AMP-based estimator/detector combinations at $f_D = 0.077$.

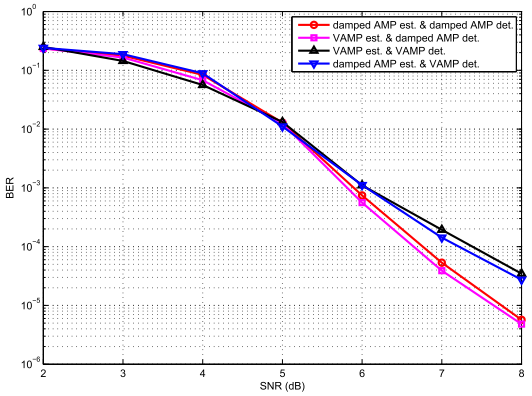


FIGURE 8. BER performance of four AMP-based estimator/detector combinations at $f_D = 0.15$.

LMMSE estimator. On the other hand, from Section IV-B, the damped AMP estimator has a numerous complexity gain over LMMSE estimator. Thus, the damped AMP estimator can achieve the LMMSE estimation performance with a much lower complexity.

D. COMPARISON OF DIFFERENT ESTIMATOR/ DETECTOR COMBINATIONS

Fig. 7 and Fig. 8 show the BER performance of four AMP-based estimator/detector combinations at $f_D = 0.077$ and $f_D = 0.15$, respectively. From these figures, the damped AMP detector has better performance than damped VAMP detector. When damped AMP detector is used, damped AMP estimator and damped VAMP estimator show similar performance. This can be partly interpreted by the fact that the AP constraint of OFDM-IM is not taken into account fully in the first global iteration in the damped VAMP detector. In fact, the damped AMP detector utilizes the AP constraint of OFDM-IM fully by computing the posterior mean and variance according to (15) and (16) in [38], as shown in denoising step of Algorithm 5.

VI. CONCLUSION

In this paper, the AMP algorithm was utilized in the iterative channel estimation and data detection of OFDM-IM in DSCs. Note that the guaranteed performance of AMP can not be achieved due to the measurement matrices in both BEM coefficient estimate and data detection are general (not i.i.d sub-Gaussian). The damped AMP and VAMP algorithms which have stronger robustness to the measurement matrix were taken as the practical channel estimator and/or data detector here. In the proposed turbo receiver, the message propagated among channel estimator, data detector and channel decoder was designed, and the corresponding message scheduling was presented. Furthermore, the efficient integration of soft input information in damped AMP/VAMP estimator and detector was proposed. Finally, the superior performance of the proposed AMP-based turbo receiver was demonstrated by extensive computer simulations.

REFERENCES

[1] R. Abu-Alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. IEEE 20th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Tokyo, Japan, Sep. 2009, pp. 177–181.

[2] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *Proc. IEEE GLOBECOM*, Dec. 2011, pp. 728–732.

[3] E. Başar, U. Aygölü, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.

[4] M. Wen, X. Cheng, M. Ma, B. Jiao, and H. V. Poor, "On the achievable rate of OFDM with index modulation," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1919–1932, Apr. 2016.

[5] Y. Ko, "A tight upper bound on bit error rate of joint OFDM and multicarrier index keying," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1763–1766, Oct. 2014.

[6] Y. Xiao, S. Wang, L. Dan, X. Lei, P. Yang, and W. Xiang, "OFDM with interleaved subcarrier-index modulation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1447–1450, Aug. 2014.

[7] E. Başar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 381–384, Aug. 2015.

[8] J. Zheng and R. Chen, "Achieving transmit diversity in OFDM-IM by utilizing multiple signal constellations," *IEEE Access*, vol. 5, pp. 8978–8988, 2017.

[9] R. Fan, Y. J. Yu, and Y. L. Guan, "Generalization of orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5350–5359, Oct. 2015.

[10] B. Zheng, F. Chen, M. Wen, F. Ji, H. Yu, and Y. Liu, "Low complexity ML detector and performance analysis for OFDM with in-phase/quadrature index modulation," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1893–1896, Nov. 2015.

[11] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 50–60, 2017.

[12] J. Zheng and H. Lv, "Peak-to-average power ratio reduction in OFDM index modulation through convex programming," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1505–1508, Jul. 2017.

[13] J. Zheng and Q. Liu, "Low-complexity soft-decision detection of coded OFDM with index modulation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7759–7763, Aug. 2018.

[14] Z. Hu, F. Chen, M. Wen, F. Ji, and H. Yu, "Low-complexity LLR calculation for OFDM with index modulation," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 618–621, Aug. 2018.

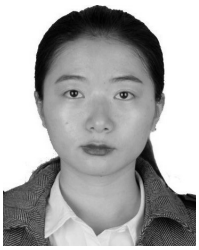
[15] J. Zheng, "Adaptive index modulation for parallel Gaussian channels with finite alphabet inputs," *IEEE Trans. Veh. Techol.*, vol. 65, no. 8, pp. 6821–6827, Aug. 2016.

[16] M. Wen, X. Cheng, L. Yang, Y. Li, X. Cheng, and F. Ji, "Index modulated OFDM for underwater acoustic communications," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 132–137, May 2016.

[17] J. Zheng and R. Chen, "Linear processing for intercarrier interference in OFDM index modulation based on capacity maximization," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 683–687, May 2017.

[18] E. Başar, "Multiple-input multiple-output OFDM with index modulation," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2259–2263, Dec. 2015.

[19] E. Basar, "On multiple-input multiple-output OFDM with index modulation for next generation wireless networks," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 3868–3878, Aug. 2016.
[20] T. Datta, H. S. Eshwaraiah, and A. Chockalingam, "Generalized space-and-frequency index modulation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 4911–4924, Jul. 2016.
[21] R. Chen and J. Zheng, "Index-modulated MIMO-OFDM: Joint space-frequency signal design and linear precoding in rapidly time-varying channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 7067–7079, Oct. 2018.
[22] N. Ishikawa, S. Sugiura, and L. Hanzo, "Subcarrier-index modulation aided OFDM-will it work?" *IEEE Access*, vol. 4, pp. 2580–2593, 2016.
[23] E. Başar, "Index modulation techniques for 5G wireless networks," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 168–175, Jul. 2016.
[24] M. Wen, X. Cheng, and L. Yang, *Index Modulation for 5G Wireless Communications (Wireless Networks)*, Berlin, Germany: Springer, 2017.
[25] F. Hlawatsch and G. Matz (Editors), *Wireless Communications Over Rapidly Time-Varying Channels*, Amsterdam, The Netherlands: Elsevier, 2011.
[26] A. P. Kannu and P. Schniter, "MSE-optimal training for linear time-varying channels," in *Proc. IEEE ICASSP*, Philadelphia, PA, USA, Mar. 2005, pp. 789–792.
[27] 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.
[28] D. K. Borah and B. T. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 862–873, Jun. 1999.
[29] H. Abdzadeh-Ziabari, W.-P. Zhu, and M. N. S. Swamy, "Joint maximum likelihood timing, frequency offset, and doubly selective channel estimation for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2787–2791, Mar. 2018.
[30] H. Nguyen-Le and T. Le-Ngoc, "Pilot-aided joint CFO and doubly-selective channel estimation for OFDM transmissions," *IEEE Trans. Broadcast.*, vol. 56, no. 4, pp. 514–522, Dec. 2010.
[31] X. Wang, G. Wang, R. Fan, and B. Ai, "Channel estimation with expectation maximization and historical information based basis expansion model for wireless communication systems on high speed railways," *IEEE Access*, vol. 6, pp. 72–80, 2018.
[32] E. P. Simon, L. Ros, H. Hijazi, and M. Ghogho, "Joint carrier frequency offset and channel estimation for OFDM systems via the EM algorithm in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 754–765, Feb. 2012.
[33] X. Wang, J. Wang, L. He, and J. Song, "Basis expansion model based spectral efficient channel estimation scheme for massive MIMO systems," in *Proc. IEEE IWCMC*, Valencia, Spain, Jun. 2017, pp. 2062–2067.
[34] Q. Qin, L. Gui, B. Gong, and S. Luo, "Sparse channel estimation for massive MIMO-OFDM systems over time-varying channels," *IEEE Access*, vol. 6, pp. 33740–33751, 2018.
[35] K. Fang, L. Rugini, and G. Leus, "Block transmissions over doubly selective channels: Iterative channel estimation and turbo equalization," *EURASIP J. Adv. Signal Process.*, vol. 2010, pp. 1–13, Apr. 2010.
[36] L. Martínez, P. Ochandiano, I. Sobrón, and M. Mendicute, "Iterative BEM channel estimation and BP detection for ICI cancellation in DVB systems with very high mobility," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast.*, Seoul, South Korea, Jun. 2012, pp. 1–6.
[37] J. Zheng and H. Lv, "A semidefinite relaxation approach to OFDM-IM detection in rapidly time-varying channels," *IEEE Access*, vol. 6, pp. 16806–16815, 2018.
[38] L. Wei, J. Zheng, and Q. Liu, "Approximate message passing detector for index modulation with multiple active resources," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 972–976, Jan. 2019.
[39] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.
[40] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *Proc. IEEE Inf. Theory Workshop Inf. Theory*, Cairo, Egypt, Jan. 2010, pp. 1–5.
[41] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.
[42] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, Aug. 2011, pp. 2168–2172.
[43] S. Rangan, P. Schniter, and A. Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun./Jul. 2014, pp. 236–240.
[44] F. Caltagirone, L. Zdeborová, and F. Krzakala, "On convergence of approximate message passing," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jul. 2014, pp. 1812–1816.
[45] P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1043–1055, Feb. 2015.
[46] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 2017, pp. 1588–1592.
[47] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 6664–6684, Oct. 2019. [Online]. Available: https://arxiv.org/abs/1610.03082
[48] Y. R. Zheng and C. Xiao, "Improved models for the generation of multiple uncorrelated Rayleigh fading waveforms," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 256–258, Jun. 2002.
[49] X. Kuai, X. Yuan, and Y.-C. Liang, "Message-passing based OFDM receiver for time-varying sparse multipath channels," *IEEE Trans. Veh. Tech.*, vol. 67, no. 10, pp. 10097–10101, Oct. 2018.



LI WEI received the B.S. degree in communication engineering from Southwest Jiaotong University, in 2015, and the M.S. degree in electronic and communication engineering from Xidian University, in 2019. Her research interest includes signal processing for wireless communications.

JIANPING ZHENG received the B.S., M.S., and the Ph.D. degrees in information and communication engineering from Xidian University, in 2003, 2006, and 2008, respectively. From 2010 to 2011, he was a Research Fellow with the PWTC Laboratory, Nanyang Technological University, Singapore. He joined the School of Telecommunications Engineering, Xidian University, in 2006, where he is currently an Associate Professor. His current research interests include MIMO, 5G techniques, and signal processing for wireless communications.