New Results on Joint Channel and Impulsive Noise Estimation and Tracking in Underwater Acoustic OFDM Systems

Shuche Wang, Zhiqiang He, *Member, IEEE*, Kai Niu, Peng Chen, and Yue Rong, *Senior Member, IEEE*

***Abstract*— Impulsive noise can greatly affect the performance of underwater acoustic (UA) orthogonal frequency-division multiplexing (OFDM) systems. In this paper, by utilizing the sparsity of the UA channel impulse response and impulsive noise, we first propose a novel sparse Bayesian learning (SBL) based expectation maximization (EM) algorithm for joint channel estimation and impulsive noise mitigation in UA OFDM systems. Secondly, considering that the UA channel and impulsive noise are fast time-varying, we develop a new approach which combines the SBL with the forward-backward Kalman filtering to track the UA channel and impulsive noise. To further improve the system performance, we utilize the information available on data subcarriers for joint time-varying channel estimation and data detection, based on the SBL algorithm and the Kalman filter. The performance of our proposed algorithms is verified through both numerical simulations and by data collected during a UA communication experiment conducted in the estuary of the Swan River, Perth, Australia. The results demonstrate that compared with existing approaches, the proposed algorithms achieve a better system bit-error-rate and frame-error-rate performance.**

***Index Terms*— Kalman filter, impulsive noise, OFDM, sparse Bayesian learning, underwater acoustic communication.**

I. Introduction

IT IS well known that the underwater acoustic (UA) channel is one of the most challenging channels for wireless communication because of severe fading, extremely limited bandwidth, significant Doppler spread, and strong multipath interference [1]. In the past decades, the orthogonal frequency-division multiplexing (OFDM) technology has been applied to mitigate the inter-symbol interference (ISI) in high-rate UA communication systems [2]. However, impulsive noise introduced by human activities and natural sources can greatly affect the performance of UA OFDM systems [3]–[5]. Thus, it is essential to estimate and mitigate impulsive noise in UA communication systems.

One type of commonly used impulsive noise mitigation methods firstly find the samples of the received signals possibly contaminated by impulsive noise through threshold testing, and then use nonlinear blanking or clipping methods to adjust impulsive noise dominated samples [6]–[8]. However, these methods may destroy the received signals. Moreover, a fixed threshold is usually used in these methods which is not adaptive to the time-varying received signals and noise. Another type of impulsive noise reduction techniques utilize the structure of OFDM signals and the sparsity of impulsive noise [9]–[15]. Recently, a joint channel estimation and impulsive noise mitigation method has been proposed in [16], which has an improved performance compared with traditional least-squares (LS) based channel estimation methods and blanking based impulsive noise mitigation approaches.

In this paper, we study joint channel and impulsive noise estimation and tracking in UA OFDM systems. The contributions of this paper over existing works [9]–[16] are summarized below.

- The performance of existing algorithms on UA channel and impulsive noise estimation (eg. [12] and [15]) degrades in low signal-to-noise ratio (SNR) and/or highly impulsive noise environments. To improve the performance of the estimator in these environments, we propose a sparse Bayesian learning (SBL) [17] based expectation maximization (EM) algorithm to jointly estimate the sparse UA channel impulse response and the impulsive noise, by utilizing the received signals at the pilot subcarriers.
- Different to existing works in [11]–[16], we explicitly consider that the UA channel and impulsive noise are fast time-varying and develop a new algorithm to track the UA channel and impulsive noise by combining the SBL approach with the forward-backward Kalman filtering. This new algorithm has not been presented in [11]. Compared with conventional Kalman filtering, the proposed algorithm exploits the sparsity of the UA channel and impulsive noise during tracking [18]. Moreover, the

Manuscript received August 13, 2019; revised November 22, 2019; accepted January 10, 2020. Date of publication January 21, 2020; date of current version April 9, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1801101, in part by the National Natural Science Foundation of China under Grant 61671080, and in part by the Australian Research Council's Discovery Projects funding scheme under Grant DP140102131. This article was presented in part at the MTS/IEEE Oceans, Kobe, Japan, May 28-31, 2018 [11]. The associate editor coordinating the review of this article and approving it for publication was D. Matolak. *(Corresponding author: Yue Rong.)*

Shuche Wang, Zhiqiang He, and Kai Niu are with the Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: wangsc@bupt.edu.cn; hezq@bupt.edu.cn; niukai@bupt.edu.cn).

Peng Chen and Yue Rong are with the School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Bentley, WA 6102, Australia (e-mail: peng.ch@outlook.com; y.rong@curtin.edu.au).

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

Digital Object Identifier 10.1109/TWC.2020.2966622

proposed UA channel tracking algorithm is more general compared with [18].

- The existing works [11]–[14] only use pilot subcarriers to estimate the channel and the impulsive noise. Considering the limited number of pilot subcarriers in UA OFDM systems, to further improve the system performance, we propose a new joint channel estimation and data detection algorithm based on the SBL method where the information available on all subcarriers is utilized for a more accurate channel and impulsive noise estimation.
- The method above using all subcarriers is further extended to jointly detect the data and track the time-varying UA channel and impulsive noise based on the SBL algorithm and the Kalman filter.
- The proposed algorithms are assessed and compared with existing methods through both numerical simulations and by real data recorded during a UA communication experiment conducted in the estuary of the Swan River, Perth, Australia. Both simulation and experiment results demonstrate that compared with existing approaches, the proposed algorithms are more effective in mitigating the impulsive noise in UA OFDM systems, and achieve an improved system bit-error-rate (BER) and frame-error-rate (FER) performance.

We would like to note that SBL has been used in [19] for impulsive noise mitigation in powerline communications. Recently, SBL has been applied for channel estimation in OFDM systems [20], multiple-input multiple-output (MIMO)-OFDM systems [21], millimeter wave hybrid MIMO systems [22], and UA OFDM systems [23]. It has been shown in [19]– [23] that the SBL-based approaches perform better than traditional compressed sensing based algorithms in sparse channel environments. However, [19]–[23] only perform either channel estimation or impulsive noise mitigation *separately*. To the best of our knowledge, SBL is applied for the first time in this paper on *joint channel and impulsive noise estimation and tracking* in UA OFDM systems.

The rest of this paper is organized as follows. The model of a UA OFDM system is shown in Section II. The proposed joint channel and impulsive noise estimation and tracking algorithms are developed in Section III. Numerical simulation and river experiment results are presented in Section IV and Section V, respectively. We draw conclusions in Section VI.

II. System Model

We study a coded frame based UA OFDM system with $N_{\rm c}$ subcarriers [16], which contains $N_{\rm s}$ subcarriers for data transmission, $N_{\rm p}$ uniformly spaced pilot subcarriers and $N_{\rm u}$ null subcarriers at edges of the passband. The $N_{\rm c}$ subcarriers are located at the frequencies of

$$f_k = f_c + \frac{k}{T}, \quad k = -\frac{N_{\rm c}}{2} + 1, \ldots, \frac{N_{\rm c}}{2}$$

where f_c is the center carrier frequency and T is the length of one OFDM symbol. At the transmitter, an $L_{\rm b}$ long binary source data stream $\mathbf{b}$ is encoded, interleaved, and (possibly) punctured to generate a coded bit sequence $\mathbf{c}$ with a length of $L_{\rm c} = R_{\rm m} N_{\rm s} N_{\rm b}$, where $R_{\rm m}$ is the modulation order and $N_{\rm b}$ is the number of OFDM symbols in one data frame. Then the coded bit sequence $\mathbf{c}$ is mapped into $N_{\rm s}N_{\rm b}$ data symbols by either the phase-shift keying (PSK) or the quadrature amplitude modulation (QAM) constellations. The bandwidth of the OFDM symbols is $B = N_{\rm c}/T$. To prevent the ISI caused by multipath fading, the transmitter inserts a cyclic prefix (CP) with the length of $T_{\rm cp}$ to each OFDM symbol. Thus, the passband transmitted signal in one OFDM block can be presented as

$$\tilde{x}(t) = \begin{cases} 2\mathrm{Re}\left\{ \dfrac{e^{j2\pi f_c t}}{\sqrt{N_{\rm c}}} \displaystyle\sum_{k=-\frac{N_{\rm c}}{2}+1}^{\frac{N_{\rm c}}{2}} d[k+\frac{N_{\rm c}}{2}] e^{j2\pi \frac{k}{T} t} \right\}, & 0 \le t \le T; \\ \tilde{x}(t+T), & -T_{\rm cp} \le t < 0 \end{cases} \tag{1}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number, $d[n]$, $n = 1, \ldots, N_{\rm c}$, is the symbol of the nth subcarrier.

We adopt the following commonly applied assumptions on UA channels [15], [24], [25]

1) The delay of the ith path is a first order function of t as $\tau_i(t) = \tau_i - at$, where a is the Doppler scaling factor. All paths have the same Doppler scaling factor during one OFDM block.[1]
2) The gain of the ith path remains constant during one OFDM block as $A_i(t) = A_i$.

Thus, the impulse response of the time-varying UA channel can be expressed as

$$h(t) = \sum_{i=1}^{L} A_i \delta(t - (\tau_i - at)) \tag{2}$$

where L is the number of paths and $\delta(\cdot)$ denotes the Dirac delta function. From (2), the passband signal received through the UA channel can be written as

$$\tilde{r}(t) = \sum_{i=1}^{L} A_i \tilde{x}((1+a)t - \tau_i) + \tilde{v}(t) + \tilde{w}(t) \tag{3}$$

where $\tilde{v}(t)$ and $\tilde{w}(t)$ are the impulsive noise and the background Gaussian noise at the passband, respectively.

By cross-correlating $\tilde{r}(t)$ with the known preamble block in a data frame, the time duration of one received data frame $\hat{T}_{\rm rx}$ can be estimated. Then by comparing $\hat{T}_{\rm rx}$ with the known time duration of one transmitted data frame $T_{\rm tx}$, the Doppler scaling factor can be estimated as $\hat{a} = \frac{\hat{T}_{rx}}{T_{tx}} - 1$. The received signal is resampled with $\hat{a}$. Then after removing the CP, downshifting, and low-pass filtering (LPF), the baseband received signal can

[1]Under the assumption of an identical Doppler scaling factor for all paths, the inter-carrier interference (ICI) can be compensated after resampling with the estimated Doppler scaling factor. When this is not the case, part of useful signals will be treated as additive noise due to the presence of the ICI, which can increase the overall noise variance. Alternatively, a more complex receiver needs to be developed, which has been shown in [26] and [27], to deal with path-specific Doppler scaling factors.

be obtained from (1) and (3) as[2]

$$r(t) = \mathrm{LPF}\left[\tilde{r}\left(\frac{t}{1+\hat{a}}\right)e^{-j2\pi f_c t}\right] \approx \frac{e^{j2\pi\Delta_f t}}{\sqrt{N_\mathrm{c}}}\sum_{k=-\frac{N_\mathrm{c}}{2}+1}^{\frac{N_\mathrm{c}}{2}}\left\{d\Big[k+\frac{N_\mathrm{c}}{2}\Big]e^{j2\pi\frac{k}{T}t}\sum_{i=1}^{L}A_i e^{-j2\pi f_k\tau_i}\right\} + v(t) + w(t) \quad (4)$$

where $\Delta_f = \frac{a-\hat{a}}{1+\hat{a}}f_c$ is the carrier frequency offset (CFO) introduced by the Doppler shift, $v(t)$ and $w(t)$ are the baseband impulsive noise and background Gaussian noise, respectively. Similar to [15, p. 8192] and [24, p. 200], (4) is obtained by approximating $e^{j2\pi\frac{k}{T}(1+a)t/(1+\hat{a})}$ as $e^{j2\pi\frac{k}{T}t}$, as a is very close to $\hat{a}$. By sampling the baseband received signal $r(t)$ at the rate of $1/B$, the discrete time samples of one OFDM symbol can be obtained from (4) as

$$r_t[n] = \frac{e^{j2\pi\Delta_f n/B}}{\sqrt{N_\mathrm{c}}}\sum_{k=-\frac{N_\mathrm{c}}{2}+1}^{\frac{N_\mathrm{c}}{2}} d\Big[k+\frac{N_\mathrm{c}}{2}\Big]e^{j2\pi\frac{kn}{N_\mathrm{c}}}\,h\Big[k+\frac{N_\mathrm{c}}{2}\Big] + v_t[n] + w_t[n], \quad n = 1,\ldots,N_\mathrm{c} \quad (5)$$

where $h[n] = \sum_{i=1}^{L} A_i e^{-j2\pi\tau_i f_{n-\frac{N_\mathrm{c}}{2}}}$, $n = 1,\ldots,N_\mathrm{c}$, is the channel frequency response at the nth subcarrier. Hereafter, the subscript "t" is used to denote samples obtained in the time-domain. The matrix-vector form of (5) is given by

$$\mathbf{r}_t = \mathbf{\Psi}(\Delta_f)\mathbf{F}^\mathrm{H}\mathbf{D}\mathbf{h} + \mathbf{v}_t + \mathbf{w}_t \quad (6)$$

where $(\cdot)^\mathrm{H}$ stands for the Hermitian transpose, $\mathbf{F}$ is an $N_\mathrm{c}\times N_\mathrm{c}$ normalized discrete Fourier transform (DFT) matrix with the (n,k)th element of $1/\sqrt{N_\mathrm{c}}e^{-j2\pi n(k-N_\mathrm{c}/2)}$, and

$$\begin{aligned}\mathbf{\Psi}(\Delta_f) &= \mathrm{diag}(e^{j2\pi\Delta_f/B},\ldots,e^{j2\pi\Delta_f N_\mathrm{c}/B})\\ \mathbf{D} &= \mathrm{diag}(d[1],\ldots,d[N_\mathrm{c}])\\ \mathbf{r}_t &= (r_t[1],\ldots,r_t[N_\mathrm{c}])^\mathrm{T}\\ \mathbf{v}_t &= (v_t[1],\ldots,v_t[N_\mathrm{c}])^\mathrm{T}\\ \mathbf{w}_t &= (w_t[1],\ldots,w_t[N_\mathrm{c}])^\mathrm{T}\\ \mathbf{h} &= (h[1],\ldots,h[N_\mathrm{c}])^\mathrm{T}.\end{aligned}$$

Here $\mathrm{diag}(\cdot)$ stands for a diagonal matrix and $(\cdot)^\mathrm{T}$ denotes the matrix transpose.

To obtain an accurate estimation of the channel impulse response and the impulsive noise, the CFO Δ_f needs to be properly estimated and compensated [15], [24]. In this paper, we apply the "clipping-blanking and Doppler" method in [15] to estimate and compensate Δ_f by using the null subcarriers. After the CFO estimation and compensation, the residual CFO is very small. We will show in Section IV-B and Section V-A that such small residual CFO has negligible influence on the system performance.

After the CFO compensation and omitting the residual CFO, from (6), the time domain and frequency domain received signal vectors are given respectively by

$$\mathbf{r}_t = \mathbf{F}^\mathrm{H}\mathbf{D}\mathbf{F}\mathbf{h}_t + \mathbf{v}_t + \mathbf{w}_t \quad (7)$$

$$\mathbf{r} = \mathbf{D}\mathbf{h} + \mathbf{v} + \mathbf{w} \quad (8)$$

[2]We assume that there is no phase offset between the transmitter and receiver carriers.

where $\mathbf{h}_t = \mathbf{F}^\mathrm{H}\mathbf{h}$ denotes the time domain channel impulse response, $\mathbf{v} = \mathbf{F}\mathbf{v}_t$, and $\mathbf{w} = \mathbf{F}\mathbf{w}_t$.

III. The Proposed Algorithms

In this section, we propose the following new algorithms based on the SBL framework and the Kalman filter to jointly estimate the UA channel and impulsive noise.

1) By exploiting the sparsity of the UA channel impulse response and the impulsive noise, we propose a novel SBL based channel and impulsive noise estimation algorithm using the pilot subcarriers.
2) Considering that the UA channel and impulsive noise are time-varying, we develop a new algorithm to track the time-varying UA channel and impulsive noise by combining the SBL algorithm with the forward-backward Kalman filtering utilizing the pilot subcarriers.
3) Noting the above algorithms only use pilot subcarriers, we propose a novel algorithm for joint channel and impulsive noise estimation and data detection by exploiting all subcarriers.
4) Combining the advantages of the algorithms above, we propose a new algorithm for joint time-varying channel and impulsive noise tracking and data detection by employing all subcarriers.

A. Sparse Bayesian Learning Based Channel and Impulsive Noise Estimation Algorithm Using Pilot Subcarriers

We introduce an $N_\mathrm{p}\times N_\mathrm{c}$ matrix $\mathbf{P}$ which selects N_p pilot subcarriers out of total N_c subcarriers. From (8), the received signal vector at the pilot subcarriers is given by

$$\mathbf{r}_p = \mathbf{D}_p\mathbf{h}_p + \mathbf{v}_p + \mathbf{w}_p \quad (9)$$

where $\mathbf{D}_p = \mathrm{diag}(\mathbf{d}_p)$, $\mathbf{d}_p$ is the pilot sequence, $\mathbf{h}_p$ is the channel frequency response at the pilot subcarriers, $\mathbf{v}_p = \mathbf{P}\mathbf{v}$, and $\mathbf{w}_p = \mathbf{P}\mathbf{w}$.

By introducing $\mathbf{h}_{p,t} = \mathbf{F}_p^\mathrm{H}\mathbf{h}_p$ and $\mathbf{v}_{p,t} = \mathbf{F}_p^\mathrm{H}\mathbf{v}_p$, where $\mathbf{F}_p$ is an $N_\mathrm{p}\times N_\mathrm{p}$ DFT matrix, (9) can be rewritten as

$$\begin{aligned}\mathbf{r}_p &= \mathbf{D}_p\mathbf{F}_p\mathbf{h}_{p,t} + \mathbf{F}_p\mathbf{v}_{p,t} + \mathbf{w}_p\\ &= \mathbf{M}_p\boldsymbol{\theta}_p + \mathbf{w}_p\end{aligned} \quad (10)$$

where $\mathbf{M}_p = (\mathbf{D}_p\mathbf{F}_p, \mathbf{F}_p)$ and $\boldsymbol{\theta}_p = (\mathbf{h}_{p,t}^\mathrm{T}, \mathbf{v}_{p,t}^\mathrm{T})^\mathrm{T}$. Since (10) represents an under-determined system, $\boldsymbol{\theta}_p$ cannot be estimated from (10) using the conventional LS method [16]. However, as the UA channel is typically sparse with only a few non-zero entries in $\mathbf{h}_{p,t}$, and $\mathbf{v}_{p,t}$ is a 'fold-and-add' of $\mathbf{v}_t$ which is also sparse,[3] $\boldsymbol{\theta}_p$ is a sparse vector, and therefore can be estimated by the SBL method. Using the estimated $\boldsymbol{\theta}_p$, the impulsive noise $\mathbf{v}_t$ can be estimated and subtracted from the received signals, and $\mathbf{h}_t$ can be estimated and used for the equalization and demodulation operations.

We assume that $\mathbf{w}_p$ has independent and identically distributed complex Gaussian entries with zero mean and variance $1/\beta$. Thus, the likelihood of (10) can be written as

$$\mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p,\beta) = (\beta/\pi)^{N_\mathrm{p}}\exp(-\beta\|\mathbf{r}_p - \mathbf{M}_p\boldsymbol{\theta}_p\|^2) \quad (11)$$

[3]As the pilot subcarriers are uniformly spaced among all subcarriers, $\mathbf{v}_p$ can be viewed as under-sampled $\mathbf{v}$. As a result, $\mathbf{v}_{p,t}$ is a time-aliased version of $\mathbf{v}_t$, which is a "fold-and-add" of $\mathbf{v}_t$.

where $\|\cdot\|$ denotes the Euclidean norm of a vector. To avoid over-fitting, we assume that $\boldsymbol{\theta}_p$ has a complex Gaussian prior distribution with zero mean and variance $1/\alpha[i]$ for its ith entry as

$$\mathrm{p}(\boldsymbol{\theta}_p|\boldsymbol{\alpha}) = \pi^{-2N_\mathrm{p}} \prod_{i=1}^{2N_\mathrm{p}} \alpha_i \exp(-\boldsymbol{\theta}_p^\mathrm{H} \mathrm{diag}(\boldsymbol{\alpha})\boldsymbol{\theta}_p) \tag{12}$$

where $\boldsymbol{\alpha} = (\alpha[1], \ldots, \alpha[2N_\mathrm{p}])^\mathrm{T}$. We treat the parameters $\boldsymbol{\alpha}$ and β as random variables with Gamma distributions as

$$\mathrm{p}(\boldsymbol{\alpha}) = \prod_{i=1}^{2N_\mathrm{p}} \gamma(\alpha[i]; a, b), \quad \mathrm{p}(\beta) = \gamma(\beta; c, d) \tag{13}$$

where $\gamma(\cdot)$ denotes the Gamma distribution and a, b, c, d are parameters of the Gamma distributions. The reason of choosing the Gamma distribution is that under the Bayesian framework, Gamma distribution is the conjugate prior for Gaussian parameters $\boldsymbol{\alpha}$ and β. Moreover, with a Gamma prior distribution, the posterior is more flexible to represent a heavy-tail distribution and can have a more sparse solution.

The estimation of $\boldsymbol{\theta}_p$ requires the calculation of the posterior distribution, which from the Bayes' rule is given by

$$\mathrm{p}(\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta|\mathbf{r}_p) = \frac{\mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta)\mathrm{p}(\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta)}{\mathrm{p}(\mathbf{r}_p)}$$

where $\mathrm{p}(\mathbf{r}_p) = \int \mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta)\mathrm{p}(\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta)\, d\boldsymbol{\theta}_p d\boldsymbol{\alpha} d\beta$ cannot be directly calculated. Interestingly, $\mathrm{p}(\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta|\mathbf{r}_p)$ can be decomposed as

$$\mathrm{p}(\boldsymbol{\theta}_p, \boldsymbol{\alpha}, \beta|\mathbf{r}_p) = \mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta)\mathrm{p}(\boldsymbol{\alpha}, \beta|\mathbf{r}_p) \tag{14}$$

where

$$\begin{aligned}\mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta) &= \frac{\mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \beta)\mathrm{p}(\boldsymbol{\theta}_p|\boldsymbol{\alpha})}{\mathrm{p}(\mathbf{r}_p|\boldsymbol{\alpha}, \beta)} \\ &= \frac{\mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \beta)\mathrm{p}(\boldsymbol{\theta}_p|\boldsymbol{\alpha})}{\int \mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \beta)\mathrm{p}(\boldsymbol{\theta}_p|\boldsymbol{\alpha})\, d\boldsymbol{\theta}_p}.\end{aligned} \tag{15}$$

Note that the numerator of (15) is the product of two Gaussian density functions and the denominator of (15) is the convolution of them. By substituting (11) and (12) into (15), we find that $\mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta)$ is a Gaussian density function as

$$\begin{aligned}&\mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta) \\ &\quad = \pi^{-2N_\mathrm{p}}|\boldsymbol{\Sigma}_p|^{-1}\exp(-(\boldsymbol{\theta}_p - \boldsymbol{\mu}_p)^\mathrm{H}\boldsymbol{\Sigma}_p^{-1}(\boldsymbol{\theta}_p - \boldsymbol{\mu}_p))\end{aligned} \tag{16}$$

where $(\cdot)^{-1}$ and $|\cdot|$ denote the matrix inversion and determinant, respectively, and

$$\boldsymbol{\Sigma}_p = (\beta\mathbf{M}_p^\mathrm{H}\mathbf{M}_p + \mathrm{diag}(\boldsymbol{\alpha}))^{-1}, \quad \boldsymbol{\mu}_p = \beta\boldsymbol{\Sigma}_p\mathbf{M}_p^\mathrm{H}\mathbf{r}_p. \tag{17}$$

Based on (16), the optimal $\boldsymbol{\theta}_p$ is given by

$$\boldsymbol{\theta}_p = \boldsymbol{\mu}_p \tag{18}$$

which depends on $\boldsymbol{\alpha}$ and β as can be seen from (17).

The optimal $\boldsymbol{\alpha}$ and β can be obtained by maximizing $\mathrm{p}(\boldsymbol{\alpha}, \beta|\mathbf{r}_p)$ in (14). However, $\mathrm{p}(\boldsymbol{\alpha}, \beta|\mathbf{r}_p)$ cannot be expressed in analytic form. To solve this issue, we apply the EM algorithm as follows. From (11)-(13), we obtain the log-likelihood function as

$$\mathcal{L} = \ln(\mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \beta)\mathrm{p}(\boldsymbol{\theta}_p|\boldsymbol{\alpha})\mathrm{p}(\boldsymbol{\alpha})\mathrm{p}(\beta)). \tag{19}$$

By ignoring the terms in (19) that are independent of $\boldsymbol{\alpha}$, we can obtain the optimal $\boldsymbol{\alpha}$ by maximizing

$$E_{\mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta)}[\ln(\mathrm{p}(\boldsymbol{\theta}_p|\boldsymbol{\alpha})\mathrm{p}(\boldsymbol{\alpha}))] \tag{20}$$

where $E_{\mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta)}[\cdot]$ stands for the expectation with respect to the distribution in (16). By setting the derivative of (20) with respect to $\alpha[i]$ to zero, we obtain

$$\alpha[i] = \frac{a}{b + \Sigma_p[i, i] + |\mu_p[i]|^2}, \quad i = 1, \ldots, 2N_\mathrm{p} \tag{21}$$

where $\Sigma_p[i, i]$ and $\mu_p[i]$ denote the (i, i)th element of $\boldsymbol{\Sigma}_p$ and the ith element of $\boldsymbol{\mu}_p$ in (17), respectively. Similarly, β can be optimized by maximizing $E_{\mathrm{p}(\boldsymbol{\theta}_p|\mathbf{r}_p, \boldsymbol{\alpha}, \beta)}[\ln(\mathrm{p}(\mathbf{r}_p|\boldsymbol{\theta}_p, \beta)\mathrm{p}(\beta))]$, leading to

$$\beta = \frac{c + N_\mathrm{p} - 1}{\tilde{d}} \tag{22}$$

where

$$\begin{aligned}\tilde{d} = d + \mathbf{r}_p^\mathrm{H}\mathbf{r}_p - \mathbf{r}_p^\mathrm{H}\mathbf{M}_p\boldsymbol{\mu}_p - \boldsymbol{\mu}_p^\mathrm{H}\mathbf{M}_p^\mathrm{H}\mathbf{r}_p \\ + \mathrm{tr}(\mathbf{M}_p\boldsymbol{\Sigma}_p\mathbf{M}_p^\mathrm{H}) + \boldsymbol{\mu}_p^\mathrm{H}\mathbf{M}_p^\mathrm{H}\mathbf{M}_p\boldsymbol{\mu}_p\end{aligned}$$

and $\mathrm{tr}(\cdot)$ denotes the matrix trace. We update $\boldsymbol{\Sigma}_p$, $\boldsymbol{\mu}_p$, $\boldsymbol{\alpha}$, and β iteratively following (17), (21), and (22). After the convergence of the algorithm, the sparse $\boldsymbol{\theta}_p$ is obtained as (18).

B. Forward-Backward Kalman Filtering Based Channel and Impulsive Noise Tracking Algorithm Using Pilot Subcarriers

In this subsection, we propose an SBL based forward-backward Kalman filtering algorithm to track the time-varying sparse channel and impulsive noise in each data frame. We assume that $\boldsymbol{\theta}_{p,m}$ remains fixed in one OFDM symbol, but varies from symbol to symbol according to a state-space model, where the subscript $m = 1, \ldots, N_\mathrm{b}$ denotes the index of OFDM symbols in one data frame. Thus, for the time-varying UA channel, the system model in (10) can be modified as

$$\mathbf{r}_{p,m} = \mathbf{M}_{p,m}\boldsymbol{\theta}_{p,m} + \mathbf{w}_{p,m}, \quad m = 1, \ldots, N_\mathrm{b} \tag{23}$$

$$\boldsymbol{\theta}_{p,m} = \boldsymbol{\theta}_{p,m-1} + \mathbf{q}_{p,m}, \quad m = 1, \ldots, N_\mathrm{b} \tag{24}$$

where $\mathbf{w}_{p,m} \sim \mathcal{CN}(\mathbf{0}, \beta_m^{-1}\mathbf{I}_{N_\mathrm{p}})$, $\mathbf{q}_{p,m} \sim \mathcal{CN}(\mathbf{0}, \mathbf{A}_m^{-1})$, $\mathbf{A}_m = \mathrm{diag}(\boldsymbol{\alpha}_m)$, $\mathbf{I}_n$ denotes the $n \times n$ identity matrix, and $\mathcal{CN}$ denotes the complex Gaussian distribution.

In the forward filtering, the tracking process is initialized with $\boldsymbol{\Omega}_{p,0|0} = \mathbf{I}_{2N_\mathrm{p}}$ and $\boldsymbol{\theta}_{p,0|0} = \mathbf{0}$. For $m = 1, \ldots, N_\mathrm{b}$, a prediction step and an update step are performed at each m. In particular, in the prediction step, we have

$$\boldsymbol{\theta}_{p,m|m-1} = \boldsymbol{\theta}_{p,m-1|m-1} \tag{25}$$

$$\boldsymbol{\Omega}_{p,m|m-1} = \boldsymbol{\Omega}_{p,m-1|m-1} + \mathbf{A}_m^{-1} \tag{26}$$

$$\mathbf{r}_{p,m|m-1} = \mathbf{M}_{p,m}\boldsymbol{\theta}_{p,m|m-1} \tag{27}$$

$$\mathbf{e}_{p,m} = \mathbf{r}_{p,m} - \mathbf{r}_{p,m|m-1} \tag{28}$$

where $m|m-1$ means the prediction of the mth symbol given the previous measurements $[\mathbf{r}_{p,1}, \ldots, \mathbf{r}_{p,m-1}]$, $\boldsymbol{\Omega}_{p,m|m-1}$ and

$\mathbf{\Omega}_{p,m-1|m-1}$ are covariance matrices of estimation errors. In the update step, we compute

$$\mathbf{K}_m=\mathbf{\Omega}_{p,m|m-1}\mathbf{M}_{p,m}^{\mathrm{H}}(\beta_m^{-1}\mathbf{I}_{N_\mathrm{p}}+\mathbf{M}_{p,m}\mathbf{\Omega}_{p,m|m-1}\mathbf{M}_{p,m}^{\mathrm{H}})^{-1} \tag{29}$$

$$\boldsymbol{\theta}_{p,m|m}=\boldsymbol{\theta}_{p,m|m-1}+\mathbf{K}_m\mathbf{e}_{p,m} \tag{30}$$

$$\mathbf{\Omega}_{p,m|m}=(\mathbf{I}_{2N_\mathrm{p}}-\mathbf{K}_m\mathbf{M}_{p,m})\mathbf{\Omega}_{p,m|m-1} \tag{31}$$

where $\mathbf{K}_m$ is the Kalman gain.

Note that the value of $\boldsymbol{\alpha}_m$ and β_m are needed in (26) and (29), respectively. From (23)-(28), we have

$$\mathbf{e}_{p,m}=\mathbf{M}_{p,m}(\mathbf{q}_{p,m}+\boldsymbol{\theta}_{p,m-1}-\boldsymbol{\theta}_{p,m-1|m-1})+\mathbf{w}_{p,m}. \tag{32}$$

We can apply the SBL algorithm developed in Section III-A to obtain $\boldsymbol{\alpha}_m$ and β_m from (32) by considering the log-likelihood function expressed as

$$\mathcal{L}=\ln(\mathrm{p}(\mathbf{e}_{p,m}|\mathbf{q}_{p,m},\beta_m;\mathbf{\Omega}_{p,m-1|m-1}) \times\mathrm{p}(\mathbf{q}_{p,m}|\boldsymbol{\alpha}_m)\mathrm{p}(\boldsymbol{\alpha}_m)\mathrm{p}(\beta_m)). \tag{33}$$

Then $\boldsymbol{\alpha}_m$ can be obtained by maximizing

$$E_{\mathrm{p}(\mathbf{q}_{p,m}|\mathbf{e}_{p,m},\boldsymbol{\alpha}_m,\beta_m;\mathbf{\Omega}_{p,m-1|m-1})}[\ln(\mathrm{p}(\mathbf{e}_{p,m}|\boldsymbol{\alpha}_m)\mathrm{p}(\boldsymbol{\alpha}_m))] \tag{34}$$

where $\mathrm{p}(\mathbf{q}_{p,m}|\mathbf{e}_{p,m},\boldsymbol{\alpha}_m,\beta_m;\mathbf{\Omega}_{p,m-1|m-1})$ can be calculated similar to (15) and turns out to be a Gaussian probability density function with the mean vector $\boldsymbol{\nu}_{p,m}$ and the covariance matrix $\mathbf{C}_{p,m}$ as below

$$\boldsymbol{\nu}_{p,m}=\mathbf{C}_{p,m}\mathbf{M}_{p,m}^{\mathrm{H}}\mathbf{B}_{p,m}\mathbf{e}_{p,m} \tag{35}$$

$$\mathbf{C}_{p,m}=(\mathbf{M}_{p,m}^{\mathrm{H}}\mathbf{B}_{p,m}\mathbf{M}_{p,m}+\mathbf{A}_m)^{-1}. \tag{36}$$

Here $\mathbf{B}_{p,m}=(\mathbf{M}_{p,m}\mathbf{\Omega}_{p,m-1|m-1}\mathbf{M}_{p,m}^{\mathrm{H}}+\beta_m^{-1}\mathbf{I}_{N_\mathrm{p}})^{-1}$.

We would like to note that in [18], it is assumed that $\boldsymbol{\theta}_{p,m-1}=\boldsymbol{\theta}_{p,m-1|m-1}$ in (32), i.e., $\mathbf{\Omega}_{p,m-1|m-1}=\mathbf{0}$. Interestingly, in this case, there is $\mathbf{B}_{p,m}=\beta_m\mathbf{I}_{N_\mathrm{p}}$, thus (35) and (36) are similar to (17). Therefore, our algorithm is more general than the approach in [18]. By setting the derivative of (34) with respect to $\alpha_m[i]$ to zero, we obtain

$$\alpha_m[i]=\frac{a}{b+C_{p,m}[i,i]+|\nu_{p,m}[i]|^2},\quad i=1,\ldots,2N_\mathrm{p}. \tag{37}$$

We can obtain β_m by maximizing

$$E_{\mathrm{p}(\mathbf{q}_{p,m}|\mathbf{e}_{p,m},\boldsymbol{\alpha}_m,\beta_m;\mathbf{\Omega}_{p,m-1|m-1})}[\ln(\mathrm{p}(\beta_m) \times\mathrm{p}(\mathbf{e}_{p,m}|\mathbf{q}_{p,m},\beta_m;\mathbf{\Omega}_{p,m-1|m-1}))]. \tag{38}$$

By setting the derivative of (38) with respect to β_m to zero, we obtain

$$\frac{c-1}{\beta_m}+\frac{\partial\ln|\mathbf{B}_{p,m}|}{\partial\beta_m} =d+E_{\mathrm{p}(\mathbf{q}_{p,m}|\mathbf{e}_{p,m},\boldsymbol{\alpha}_m,\beta_m;\mathbf{\Omega}_{p,m-1|m-1})}[(\mathbf{e}_{p,m}-\mathbf{M}_{p,m})^{\mathrm{H}} \times\mathbf{B}_{p,m}(\mathbf{e}_{p,m}-\mathbf{M}_{p,m})]. \tag{39}$$

Let us introduce the eigenvalue decomposition of $\mathbf{M}_{p,m}\mathbf{\Omega}_{p,m-1|m-1}\mathbf{M}_{p,m}^{\mathrm{H}}=\mathbf{U}_m\mathbf{\Lambda}_m\mathbf{U}_m^{\mathrm{H}}$. It can be shown from (39) that β_m can be obtained through solving the following equation

$$\sum_{i=1}^{N_\mathrm{p}}\left(\frac{\lambda_m[i,i]}{\beta_m\lambda_m[i,i]+1}+\frac{Z_m[i,i]}{(\beta_m\lambda_m[i,i]+1)^2}\right)+d =\frac{N_\mathrm{p}+c-1}{\beta_m} \tag{40}$$

where $\lambda_m[i,i]$, $i=1,\ldots,N_\mathrm{p}$, is the ith diagonal element of $\mathbf{\Lambda}_m$ and $Z_m[i,i]$ is the ith diagonal element of

$$\mathbf{Z}_m=\mathbf{U}_m^{\mathrm{H}}(\mathbf{e}_{p,m}\mathbf{e}_{p,m}^{\mathrm{H}}-\mathbf{e}_{p,m}\boldsymbol{\nu}_{p,m}^{\mathrm{H}}\mathbf{M}_{p,m}^{\mathrm{H}}-\mathbf{M}_{p,m}\boldsymbol{\nu}_{p,m}\mathbf{e}_{p,m}^{\mathrm{H}} +\mathbf{M}_{p,m}(\boldsymbol{\nu}_{p,m}\boldsymbol{\nu}_{p,m}^{\mathrm{H}}+\mathbf{C}_{p,m})\mathbf{M}_{p,m}^{\mathrm{H}})\mathbf{U}_m. \tag{41}$$

For the case of $\mathbf{\Omega}_{p,m-1|m-1}=\mathbf{0}$ [18], there are $\lambda_m[i,i]=0$, $i=1,\ldots,N_\mathrm{p}$, and (40) becomes

$$\begin{aligned}\frac{N_\mathrm{p}+c-1}{\beta_m}&=d+\mathrm{tr}(\mathbf{Z}_m)\\&=d+\mathbf{e}_{p,m}^{\mathrm{H}}\mathbf{e}_{p,m}-\boldsymbol{\nu}_{p,m}^{\mathrm{H}}\mathbf{M}_{p,m}^{\mathrm{H}}\mathbf{e}_{p,m}\\&\quad-\mathbf{e}_{p,m}^{\mathrm{H}}\mathbf{M}_{p,m}\boldsymbol{\nu}_{p,m}+\boldsymbol{\nu}_{p,m}^{\mathrm{H}}\mathbf{M}_{p,m}^{\mathrm{H}}\mathbf{M}_{p,m}\boldsymbol{\nu}_{p,m}\\&\quad+\mathrm{tr}(\mathbf{M}_{p,m}\mathbf{C}_{p,m}\mathbf{M}_{p,m}^{\mathrm{H}})\end{aligned}$$

which is similar to (22). Therefore, our algorithm is more general than [18]. Finally, $\boldsymbol{\alpha}_m$ and β_m are obtained iteratively from (37), (40), and (35), (36).

The procedure of the forward Kalman filtering in (25)-(31) can be similarly applied to the backward Kalman filtering from time $m=N_\mathrm{b}$ to time $m=1$. The results of the forward and backward Kalman filtering of $\boldsymbol{\theta}_{p,m}$ can be optimally combined as shown below. We assume that we have the following forward and backward systems

$$\mathbf{r}_{p,f}=\mathbf{M}_{p,f}\boldsymbol{\mu}+\mathbf{w}_{p,f} \tag{42}$$

$$\mathbf{r}_{p,b}=\mathbf{M}_{p,b}\boldsymbol{\mu}+\mathbf{w}_{p,b} \tag{43}$$

where $\boldsymbol{\mu}$ needs to be estimated, the subscripts f and b denote variables in the forward and backward systems, respectively, the noise vectors $\mathbf{w}_{p,f}$ and $\mathbf{w}_{p,b}$ are independent of each other. We can apply the linear minimum mean-squared (LMMSE) estimator for each linear system to estimated $\boldsymbol{\mu}$. For the forward system (42), $\boldsymbol{\mu}$ can be estimated as

$$\hat{\boldsymbol{\mu}}_f=\mathbf{\Sigma}_f\mathbf{M}_{p,f}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,f}}^{-1}\mathbf{r}_{p,f} \tag{44}$$

where $\mathbf{\Sigma}_f=(\mathbf{R}_\mu^{-1}+\mathbf{M}_{p,f}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,f}}^{-1}\mathbf{M}_{p,f})^{-1}$ is the estimation error matrix, $\mathbf{R}_\mu$ and $\mathbf{R}_{\mathbf{w}_{p,f}}$ are the covariance matrices of $\boldsymbol{\mu}$ and $\mathbf{w}_{p,f}$, respectively. Similar to (44), we can estimate $\boldsymbol{\mu}$ from the backward system (43) as

$$\hat{\boldsymbol{\mu}}_b=\mathbf{\Sigma}_b\mathbf{M}_{p,b}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,b}}^{-1}\mathbf{r}_{p,b}$$

where $\mathbf{\Sigma}_b=(\mathbf{R}_\mu^{-1}+\mathbf{M}_{p,b}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,b}}^{-1}\mathbf{M}_{p,b})^{-1}$ and $\mathbf{R}_{\mathbf{w}_{p,b}}$ is the covariance matrix of $\mathbf{w}_{p,b}$.

Combining (42) and (43), we have

$$\begin{pmatrix}\mathbf{r}_{p,f}\\\mathbf{r}_{p,b}\end{pmatrix}=\begin{pmatrix}\mathbf{M}_{p,f}\\\mathbf{M}_{p,b}\end{pmatrix}\boldsymbol{\mu}+\begin{pmatrix}\mathbf{w}_{p,f}\\\mathbf{w}_{p,b}\end{pmatrix}. \tag{45}$$

The LMMSE estimation of $\boldsymbol{\mu}$ from (45) is given by

$$\begin{aligned}\hat{\boldsymbol{\mu}}_c &= \left(\begin{bmatrix}\mathbf{M}_{p,f}^{\mathrm{H}} & \mathbf{M}_{p,b}^{\mathrm{H}}\end{bmatrix}\begin{bmatrix}\mathbf{R}_{\mathbf{w}_{p,f}}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_{\mathbf{w}_{p,b}}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{M}_{p,f}\\ \mathbf{M}_{p,b}\end{bmatrix}+\mathbf{R}_{\mu}^{-1}\right)^{-1}\\ &\quad\times\begin{bmatrix}\mathbf{M}_{p,f}^{\mathrm{H}} & \mathbf{M}_{p,b}^{\mathrm{H}}\end{bmatrix}\begin{bmatrix}\mathbf{R}_{\mathbf{w}_{p,f}}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_{\mathbf{w}_{p,b}}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{r}_{p,f}\\ \mathbf{r}_{p,b}\end{bmatrix}\\ &= (\mathbf{M}_{p,f}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,f}}^{-1}\mathbf{M}_{p,f}+\mathbf{M}_{p,b}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,b}}^{-1}\mathbf{M}_{p,b}+\mathbf{R}_{\mu}^{-1})^{-1}\\ &\quad\times(\mathbf{M}_{p,f}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,f}}^{-1}\mathbf{r}_{p,f}+\mathbf{M}_{p,b}^{\mathrm{H}}\mathbf{R}_{\mathbf{w}_{p,b}}^{-1}\mathbf{r}_{p,b})\\ &= (\boldsymbol{\Sigma}_f^{-1}+\boldsymbol{\Sigma}_b^{-1}-\mathbf{R}_{\mu}^{-1})^{-1}(\boldsymbol{\Sigma}_f^{-1}\hat{\boldsymbol{\mu}}_f+\boldsymbol{\Sigma}_b^{-1}\hat{\boldsymbol{\mu}}_b). \end{aligned}\tag{46}$$

Since Kalman filtering is a generalization of sequential LMMSE estimation, the forward and backward estimations can be combined following (46) as

$$\begin{aligned}\boldsymbol{\theta}_{p,m}^{c} = \Big(\big(\boldsymbol{\Omega}_{p,m|m}^{f}\big)^{-1}+\big(\boldsymbol{\Omega}_{p,m|m}^{b}\big)^{-1}\Big)^{-1}\Big(&\big(\boldsymbol{\Omega}_{p,m|m}^{f}\big)^{-1}\boldsymbol{\theta}_{p,m|m}^{f}\\ &+\big(\boldsymbol{\Omega}_{p,m|m}^{b}\big)^{-1}\boldsymbol{\theta}_{p,m|m}^{b}\Big)\end{aligned}$$

where the subscripts f and b denote the values from the forward and backward filtering, respectively, with $\boldsymbol{\theta}_{p,m|m}^{f}$ and $\boldsymbol{\Omega}_{p,m|m}^{f}$ given by (30) and (31), respectively, and $\boldsymbol{\theta}_{p,m|m}^{b}$ and $\boldsymbol{\Omega}_{p,m|m}^{b}$ obtained in a similar way. Note that we set $\mathbf{R}_{\mu}^{-1}=\mathbf{0}$ in (46) as there is no prior auto-correlation information about $\boldsymbol{\mu}$.

C. SBL Based Joint Data Detection and Channel and Impulsive Noise Estimation Algorithm Using All Subcarriers

The performance of the SBL based channel and impulsive noise estimation algorithm proposed in Section III-A can be improved by increasing the number of pilot subcarriers. However, the number of pilot subcarriers is limited in UA OFDM systems due to bandwidth and spectral efficiency constraints. In this subsection, we develop a joint data detection and channel estimation algorithm which utilizes all subcarriers to improve the accuracy of channel and impulsive noise estimation. From (8), the received signals on all N_{c} subcarriers can be expressed as

$$\mathbf{r}=\mathbf{M}\boldsymbol{\theta}+\mathbf{w}\tag{47}$$

where

$$\mathbf{M}=(\mathbf{DF},\mathbf{F}),\quad \boldsymbol{\theta}=(\mathbf{h}_t^{\mathrm{T}},\mathbf{v}_t^{\mathrm{T}})^{\mathrm{T}}.\tag{48}$$

As $\mathbf{D}$ contains both the known pilot information and the unknown data, we estimate the unknown part of $\mathbf{D}$ using the EM procedure. In the E-step, similar to (17), we have

$$\begin{aligned}\boldsymbol{\Sigma} &= (\beta\mathbf{M}^{\mathrm{H}}\mathbf{M}+\mathrm{diag}(\boldsymbol{\alpha}))^{-1}\\ &= \begin{pmatrix}\boldsymbol{\Sigma}_{1,1} & \boldsymbol{\Sigma}_{1,2}\\ \boldsymbol{\Sigma}_{1,2}^{\mathrm{H}} & \boldsymbol{\Sigma}_{2,2}\end{pmatrix}\end{aligned}\tag{49}$$

$$\boldsymbol{\mu}=\beta\boldsymbol{\Sigma}\mathbf{M}^{\mathrm{H}}\mathbf{r}=[\boldsymbol{\mu}_1^{\mathrm{T}},\boldsymbol{\mu}_2^{\mathrm{T}}]^{\mathrm{T}}\tag{50}$$

where $\boldsymbol{\Sigma}_{i,j}$, $i,j=1,2$, are $N_{\mathrm{c}}\times N_{\mathrm{c}}$ sub-matrices of $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ contain the first and the last N_{c} elements of $\boldsymbol{\mu}$, respectively. In the M-step, similar to (19), the log-likelihood function of (47) can be written as

$$\mathcal{L}=\ln(\mathrm{p}(\mathbf{r}|\boldsymbol{\theta},\beta;\mathbf{D})\mathrm{p}(\boldsymbol{\theta}|\boldsymbol{\alpha})\mathrm{p}(\boldsymbol{\alpha})\mathrm{p}(\beta)).\tag{51}$$

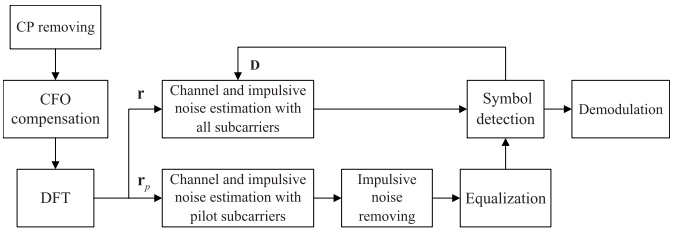


Fig. 1. Block diagram of the proposed SBL based joint data detection and channel and impulsive noise estimation algorithm with all subcarriers.

By ignoring the terms in (51) that are independent of $\mathbf{D}$, we can obtain the optimal $\mathbf{D}$ by maximizing

$$E_{\mathrm{p}(\boldsymbol{\theta}|\mathbf{r},\boldsymbol{\alpha},\beta;\mathbf{D})}[\ln(\mathrm{p}(\mathbf{r}|\boldsymbol{\theta},\beta;\mathbf{D}))]$$

which is equivalent to solving the following problem

$$\min_{\mathbf{D}}\ \|\mathbf{r}-\mathbf{M}\boldsymbol{\mu}\|^2+\mathrm{tr}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^{\mathrm{H}}).\tag{52}$$

By substituting (48), (49), and (50) into (52), we have

$$\begin{aligned}&\|\mathbf{r}-\mathbf{M}\boldsymbol{\mu}\|^2+\mathrm{tr}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^{\mathrm{H}})\\ &\quad=\|\mathbf{r}-\mathbf{DF}\boldsymbol{\mu}_1-\mathbf{F}\boldsymbol{\mu}_2\|^2+\mathrm{tr}(\mathbf{DF}\boldsymbol{\Sigma}_{1,1}\mathbf{F}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}\\ &\qquad+\mathbf{F}\boldsymbol{\Sigma}_{1,2}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}\mathbf{D}^{\mathrm{H}}+\mathbf{DF}\boldsymbol{\Sigma}_{1,2}\mathbf{F}^{\mathrm{H}}+\mathbf{F}\boldsymbol{\Sigma}_{2,2}\mathbf{F}^{\mathrm{H}}).\end{aligned}\tag{53}$$

Let us introduce $\mathcal{I}_d$ as the indices of subcarriers with data symbols. From (53), the data symbol $d[i]$, i.e., the ith diagonal element of $\mathbf{D}$, $i\in\mathcal{I}_d$, can be obtained by solving the problem of

$$\begin{aligned}\min_{d[i]}\ &|r[i]-d[i]\mathbf{F}[i,:]\boldsymbol{\mu}_1-\mathbf{F}[i,:]\boldsymbol{\mu}_2|^2+C_1[i,i]|d[i]|^2\\ &+d[i]C_2[i,i]+d^*[i]C_2^*[i,i]\end{aligned}\tag{54}$$

where $\mathbf{C}_1=\mathbf{F}\boldsymbol{\Sigma}_{1,1}\mathbf{F}^{\mathrm{H}}$, $\mathbf{C}_2=\mathbf{F}\boldsymbol{\Sigma}_{1,2}\mathbf{F}^{\mathrm{H}}$, $\mathbf{F}[i,:]$ is the ith row of $\mathbf{F}$, and $(\cdot)^*$ stands for complex conjugate.

Similar to (21) and (22), $\boldsymbol{\alpha}$ and β are optimized as

$$\alpha[i]=\frac{a}{b+\Sigma[i,i]+|\mu[i]|^2},\quad i=1,\ldots,2N_{\mathrm{c}}\tag{55}$$

$$\beta=\frac{c+N_{\mathrm{p}}-1}{\bar{d}}\tag{56}$$

where $\Sigma[i,i]$ and $\mu[i]$ denote the (i,i)th element of $\boldsymbol{\Sigma}$ in (49) and the ith element of $\boldsymbol{\mu}$ in (50), respectively, and

$$\bar{d}=d+\mathbf{r}^{\mathrm{H}}\mathbf{r}-\mathbf{r}^{\mathrm{H}}\mathbf{M}\boldsymbol{\mu}-\boldsymbol{\mu}^{\mathrm{H}}\mathbf{M}^{\mathrm{H}}\mathbf{r}+\mathrm{tr}(\mathbf{M}\boldsymbol{\Sigma}\mathbf{M}^{\mathrm{H}})+\boldsymbol{\mu}^{\mathrm{H}}\mathbf{M}^{\mathrm{H}}\mathbf{M}\boldsymbol{\mu}.$$

Therefore, we update $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$, $\mathbf{D}$, $\boldsymbol{\alpha}$, and β iteratively following (49), (50), (54), (55), and (56). Note that the unknown data symbols in $\mathbf{D}$ need to be initialized. Towards this end, we can apply the SBL joint estimation algorithm using pilot subcarriers in Section III-A to obtain an estimation of $\mathbf{h}_f$ and $\mathbf{v}_f$. Then the initial value of data symbols can be obtained after channel equalization and symbol detection. Fig. 1 shows the block diagram of the proposed SBL based joint data detection and channel and impulsive noise estimation algorithm with all subcarriers. Compared with the algorithm in Section III-A, it will be shown in Section IV and Section V that by using all subcarriers, the algorithm in this section has better BER and FER performance, at the price of a higher computational complexity.

D. SBL and Kalman Filtering Based Joint Data Detection and Channel Tracking Algorithm Using All Subcarriers

In this subsection, we combine advantages of the Kalman filtering and the algorithm in Section III-C, and propose a new algorithm for joint time-varying channel and impulsive noise estimation/tracking and data detection using all subcarriers. From (47), the time-varying system model with all $N_{\rm c}$ subcarriers can be written as

$$\mathbf{r}_m = \mathbf{M}_m\boldsymbol{\theta}_m + \mathbf{w}_m, \quad m = 1, \ldots, N_{\rm b} \tag{57}$$

$$\boldsymbol{\theta}_m = \boldsymbol{\theta}_{m-1} + \mathbf{q}_m, \quad m = 1, \ldots, N_{\rm b} \tag{58}$$

where $\mathbf{M}_m = [\mathbf{D}_m\mathbf{F},\ \mathbf{F}]$.

This algorithm is initialized with $\boldsymbol{\Omega}_{0|0} = \mathbf{I}_{2N_{\rm c}}$, $\boldsymbol{\theta}_{0|0} = \mathbf{0}$, and $\mathbf{M}_1$ is obtained using the algorithm in Section III-C. The prediction and update processes are similar to (25)-(31) but without the subscript 'p'. It is also essential to estimate the hyperparameters $\boldsymbol{\alpha}_m$ and β_m and the unknown data symbols in $\mathbf{D}_m$ in the M-step. Similar to (32), from (57) and (58) we have

$$\mathbf{e}_m = \mathbf{M}_m(\mathbf{q}_m + \boldsymbol{\theta}_{m-1} - \boldsymbol{\theta}_{m-1|m-1}) + \mathbf{w}_m. \tag{59}$$

Based on (59), $\boldsymbol{\alpha}_m$ and β_m can be estimated following the steps from (33) to (40) without the subscript 'p'. Similar to (54), the unknown data symbols in $\mathbf{D}_m$ can be obtained by solving the optimization problem of

$$\min_{d_m[i]} \ |r_m[i] - d_m[i]\mathbf{F}[i,:]\boldsymbol{\theta}_{m|m,1} - \mathbf{F}[i,:]\boldsymbol{\theta}_{m|m,2}|^2 + C_{m,1}[i,i]|d_m[i]|^2 + d_m[i]C_{m,2}[i,i] + d_m^*[i]C_{m,2}^*[i,i]$$

where

$$\mathbf{C}_{m,1} = \mathbf{F}\boldsymbol{\Omega}_{m|m,1,1}\mathbf{F}^{\rm H}$$

$$\mathbf{C}_{m,2} = \mathbf{F}\boldsymbol{\Omega}_{m|m,1,2}\mathbf{F}^{\rm H}$$

$$\boldsymbol{\Omega}_{m|m} = \begin{pmatrix} \boldsymbol{\Omega}_{m|m,1,1} & \boldsymbol{\Omega}_{m|m,1,2} \\ \boldsymbol{\Omega}_{m|m,1,2}^{\rm H} & \boldsymbol{\Omega}_{m|m,2,2} \end{pmatrix}$$

$$\boldsymbol{\theta}_{m|m} = [\boldsymbol{\theta}_{m|m,1}^{\rm T}, \boldsymbol{\theta}_{m|m,2}^{\rm T}]^{\rm T}.$$

Here $\boldsymbol{\theta}_{m|m}$ is the estimated $\boldsymbol{\theta}_m$ and $\boldsymbol{\Omega}_{m|m}$ is the covariance matrix of the estimation error.

E. Computational Complexity of the Proposed Algorithms

In this subsection, we analyze the computational complexity of the four proposed algorithms.

- The complexity of the SBL algorithm using pilot subcarriers proposed in Section III-A is dominated by (17) in the EM process. Thus, the computational complexity of this algorithm for each OFDM symbol can be estimated as $\mathcal{O}(I_1\ N_{\rm p}^3)$, where I_1 is the average number of iterations of the EM process till convergence.
- As all subcarriers are used for joint data detection and channel and impulsive noise estimation, the number of subcarriers used is increased from $N_{\rm p}$ to $N_{\rm c}$. Therefore, for each OFDM symbol, the complexity order of the algorithm in Section III-C is $\mathcal{O}(I_2\ N_{\rm c}^3)$, where I_2 is the average number of iterations required by the EM process till convergence.
- The SBL based Kalman filtering algorithm in Section III-B includes three major steps: filtering (predicting and updating), smoothing, and EM parameters estimation. In the filtering step, most of the computation is spent on calculating the Kalman gain $\mathbf{K}_m$ (29), which has a complexity of $\mathcal{O}(N_{\rm p}^3)$ for each OFDM symbol. Based on the analysis above, the complexity of the EM step is $\mathcal{O}(I_1\ N_{\rm p}^3)$. Moreover, the backward filtering in the smoothing step has a computational complexity of $\mathcal{O}(N_{\rm p}^3)$ for each OFDM symbol. Thus, for each OFDM symbol, the computational complexity order of the algorithm in Section III-B is $\mathcal{O}((I_1+2)N_{\rm p}^3)$.
- As all $N_{\rm c}$ subcarriers are used for the SBL based forward-backward Kalman filtering algorithm in Section III-D, the computational complexity of this algorithm is $\mathcal{O}((I_2+2)N_{\rm c}^3)$ per OFDM symbol.

From the analysis above, we can see that the computational complexity of the proposed algorithms is practical considering that UA communication systems usually have a much lower data rate than terrestrial radio systems. Moreover, compared with the estimators using only pilot subcarriers, estimators using all subcarriers have a higher complexity. However, it will be shown in Sections IV and V that the latter algorithms have a better BER and FER performance. Similarly, compared with the original SBL based algorithms, the algorithms with forward-backward Kalman filtering have a higher computational complexity and a better performance, by using filtering and smoothing operations for tracking the time-varying channel and impulsive noise. These performance-complexity tradeoffs are interesting for practical UA OFDM systems.

IV. Numerical Simulation Results

In this section, we present numerical simulation results of the four proposed algorithms.

A. Simulation Setup

An OFDM system with $N_{\rm c} = 512$, $N_{\rm s} = 325$, $N_{\rm p} = 128$, and $N_{\rm u} = 59$ is simulated. The data symbols are modulated by $1/3$ rate turbo encoded QPSK constellations, while the pilot symbols are modulated by QPSK constellations without channel coding. Considering the code puncturing, the number of information-carrying bits in each frame is $L_{\rm b} = 1088$.

The multipath channel consists of 15 discrete paths, where the time delay between two adjacent paths follows the exponential distribution with a mean value of 1 ms. The phase of each path is uniformly distributed between $-\pi$ and π. The paths amplitudes have Rayleigh distribution whose variances have an exponentially decreasing profile with a total decay of 20 dB from $t = 0$ to $t = T_{\rm cp}$. The system bandwidth B is 4 kHz. Thus, the subcarrier spacing f_{sc} is 7.8 Hz. The CFO Δ_f is randomly generated in $[-f_{sc}/2, f_{sc}/2]$.

We apply the Gaussian mixture model [15] to simulate the composite noise $\mathbf{u} = \mathbf{v}_t + \mathbf{w}_t$ in (7) with

$$u[i] \sim (1-\eta)\mathcal{CN}(0, \sigma_b^2) + \eta\mathcal{CN}(0, \sigma_I^2), \quad i = 1, \ldots, N_{\rm c} \tag{60}$$

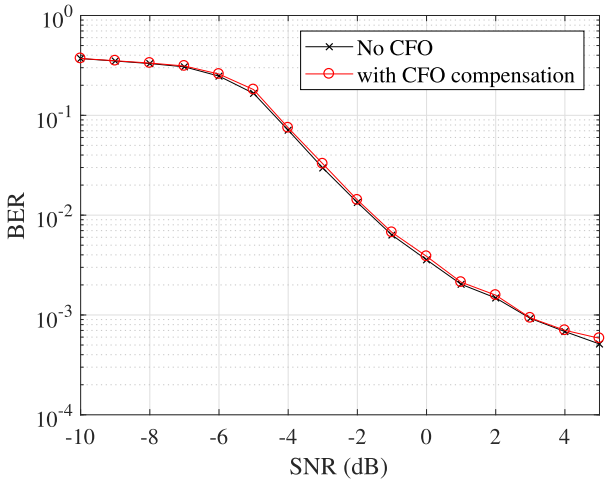


Fig. 2. The effect of residual CFO on the system BER performance.

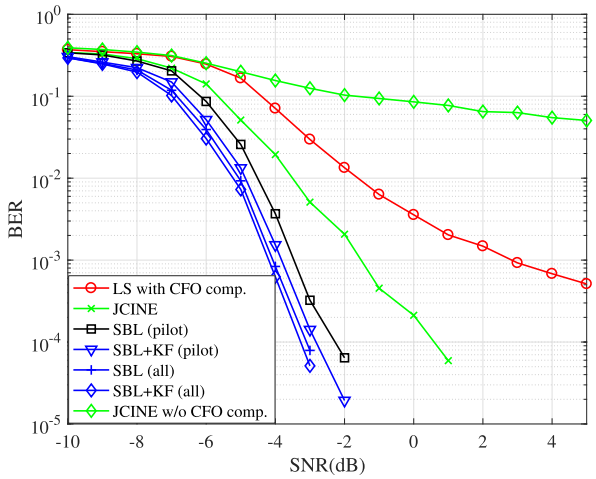


Fig. 3. BER performance of various algorithms versus SNR.

where η is the probability of the impulsive noise, σ_b^2 and σ_I^2 are the variance of the background noise and impulsive noise, respectively. The SNR is defined as

$$\text{SNR} = \frac{\sigma_s^2}{\eta\sigma_I^2 + (1-\eta)\sigma_b^2}$$

where σ_s^2 is the signal power, and the impulsive noise to background noise ratio (INR) is defined as $\text{INR} = \sigma_I^2/\sigma_b^2$. Similar to [12], we choose $\eta = 0.02$ and $\text{INR} = 26$ dB.

B. Residual Carrier Frequency Offset

In this subsection, we study the impact of the residual CFO on the system BER after the CFO estimation and compensation using the null subcarriers. We adopt the "clipping-blanking and Doppler" method in [15] to estimate the CFO. The "LS + blanking" approach is applied to detect the received signals, where the blanking method is used to suppress the impulsive noise, followed by the LS method for channel estimation. We compare the system BER in the following two scenarios to analysis the influence of the residual CFO.

1) The OFDM signals are transmitted through the multipath UA channel with Gaussian mixture noise (60), but the CFO is set to zero. We consider the BER of such system without any Doppler shift as a benchmark for comparison.
2) We add random CFO in $[-f_{sc}/2, f_{sc}/2]$ to the OFDM symbols received by the system above. A one-dimensional search method [15], [24] is used to estimate the CFO with a step size of 0.25 Hz.

Fig. 2 shows the system BER versus SNR for the two scenarios above with $\text{INR} = 26$ dB. It can be seen that the Doppler shift in the second scenario can be mitigated effectively, since after the CFO estimation and compensation, the system BER is almost identical to the first scenario without any CFO. Moreover, Fig. 2 indicates that the small residual CFO after the CFO compensation is not a limiting factor of the BER performance of UA OFDM systems [25].

C. BER Performance Comparison

In this subsection, we compare the BER performance of the following seven algorithms versus the SNR with $\text{INR} = 26$ dB. Except for the first algorithm, the other six algorithms perform the CFO estimation and compensation before channel estimation.

- *JCINE + LS INC without CFO compensation (JCINE w/o CFO comp.):* This receiver uses the joint channel and impulsive noise estimation (JCINE) and LS-based impulsive noise cancelation (INC) algorithm [16], but bypasses the CFO estimation and compensation step.
- *LS + blanking with CFO compensation (LS with CFO comp.):* This algorithm performs the CFO estimation and compensation followed by the LS + blanking channel estimation.
- *JCINE + LS INC algorithm (JCINE):* This receiver uses the JCINE algorithm for the channel and impulsive noise estimation, followed by the LS impulsive noise cancelation [16].
- *SBL algorithm using pilot subcarriers (SBL (pilot)):* This algorithm is developed in Section III-A.
- *SBL+Kalman filtering (KF) using pilot subcarriers (SBL+KF (pilot)):* This algorithm is proposed in Section III-B.
- *SBL based joint channel estimation and data detection using all subcarriers (SBL (all)):* This algorithm is presented in Section III-C.
- *SBL+KF joint channel tracking and data detection using all subcarriers (SBL+KF (all)):* Details of this algorithm are explained in Section III-D.

It can be seen from Fig. 3 that the UA OFDM system yields a poor BER performance when the CFO compensation is not performed. Although the LS + blanking algorithm carries out the CFO compensation, it has a higher BER than the other five algorithms. In fact, the LS + blanking algorithm suffers from an obvious BER floor with an increasing SNR. Compared with the JCINE + LS INC algorithm, the proposed SBL (pilot) algorithm has a 3 dB gain at high SNRs. Moreover, we can see that the SBL+KF (pilot) algorithm further

Fig. 4. Placement of the transmitter and receiver during the experiment.

reduces the system BER, because the Kalman filter exploits the knowledge of the time-varying channel to improve the accuracy of channel estimation. We can also observe from Fig. 3 that by utilizing the data subcarriers, the SBL based joint channel estimation and data detection algorithm can obtain a 1 dB gain compared with the SBL algorithm using pilot subcarriers. Finally, the proposed SBL + KF using all subcarriers algorithm outperforms the other six algorithms by combining the advantages of the Kalman filter and all subcarriers for channel and impulsive noise estimation.

V. Experiment Results

We apply the proposed algorithms to process the signals recorded during a UA communication experiment conducted in the estuary of the Swan River, Perth, Australia, which has brackish water with varying salinity. The transmitter-receiver distance was around 936 meters measured from Google Maps as shown in Fig. 4. Since the receiver hydrophone was placed in shallow warm water close to a jetty, there was a significant amount of highly impulsive snapping shrimp noise during the experiment. Another source of impulsive noise was from waves breaking at the jetty piers whose intensity increases with the wind speed.

Key parameters of the experimental system[4] are: $N_{\rm b} = 5$, bandwidth $= 4$ kHz, carrier frequency $= 12$ kHz, $N_{\rm c} = 512$, and length of CP $= 25$ ms. Each frame contains a preamble block in front of 5 OFDM data blocks. The information bits are coded by 1/2 or 1/3 rate turbo codes and then modulated by QPSK constellations. Each transmission contains 500 frames with 250 frames for each coding rate. The data files recorded at the receiver during three transmissions were named T83, T84, and T85, respectively.

A. Channel Conditions

In this subsection, we analyze the channel conditions and the features of received signals during the experiment.

[4]There are short-range UA modems which can achieve a higher data rate than this experimental system. Nevertheless, the purpose of this experimental system is to test the performance of the proposed algorithms, not particularly aiming at short-range high-rate transmissions.

TABLE I
Estimated SBR and SIR in Three Files

File name	Coding rate	SBR (dB)	SIR (dB)
T83	1/3	11.0	-7.6
	1/2	11.5	-7.3
T84	1/3	4.8	-14.2
	1/2	4.6	-13.4
T85	1/3	7.2	-13.8
	1/2	6.5	-13.6

Fig. 5 demonstrates the Doppler shift of the received signals in the three recorded files estimated by the preamble block in each frame through calculating the phase difference in the time domain between halves of the received synchronization training signals. This method is capable of estimating a small frequency shift. It can be seen that depending on the wind and wave conditions, the range of Doppler shifts is different in the three recorded files. In particular, for the T83 and T85 files, only a few frames of received signals have a Doppler shift greater than 0.25 Hz. As shown in Section IV-B, the step size of the one-dimensional search method used for CFO estimation is set to 0.25 Hz, and the residue CFO does not affect the system BER, the step of CFO estimation and compensation can be skipped to reduce the computational complexity at the receiver. For the T84 file, as around 10% of the frames have a Doppler shift larger than 0.25 Hz, the CFO estimation and compensation step is applied to process the received signals.

Fig. 6 shows the amplitude of the received signals in one typical data frame of each recorded file. The amplitude of the impulsive noise in the corresponding data frames estimated by the algorithm proposed in Section III-A is illustrated in Fig. 7. By comparing Fig. 6 with Fig. 7, it can be seen that the impulsive noise is correctly detected. Fig. 8 demonstrates the amplitude of the channel impulse response estimated by the algorithm in Section III-A for two consecutive OFDM blocks in the T83 file. It can be seen that the dominant multipath components of the channel essentially remain static for two consecutive OFDM blocks, which supports the assumption on UA channels in Section II. Table I illustrates the estimated signal-to-background (non-impulsive) noise ratio (SBR) and signal-to-impulsive noise ratio (SIR) of the received signals in three recorded files. The SBR and SIR are defined as SBR $= \sigma_s^2/\sigma_b^2$ and SIR $= \sigma_s^2/\sigma_I^2$. From Figs. 6 and 7 and Table I, it can be seen that the signals in the T83 file are only slightly disturbed by impulsive noise during transmission. The signals in the T84 file are contaminated by a significant amount of impulsive noise, as they have the lowest estimated SIR among the three recorded files. Moreover, it can be seen that the signals in the T85 file undergo a medium level of impulsive noise compared with the other two files.

B. Receiver Performance

In this subsection, the performance of the six algorithms listed in Section IV-C except for the JCINE + LS INC without CFO compensation algorithm is compared in terms of the BER (uncoded and coded) and the frame error rate (FER).

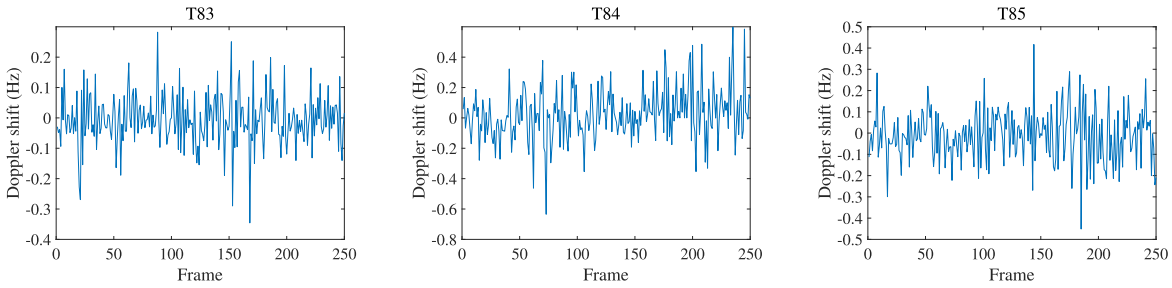


Fig. 5. Estimated Doppler shift from the three recorded files.

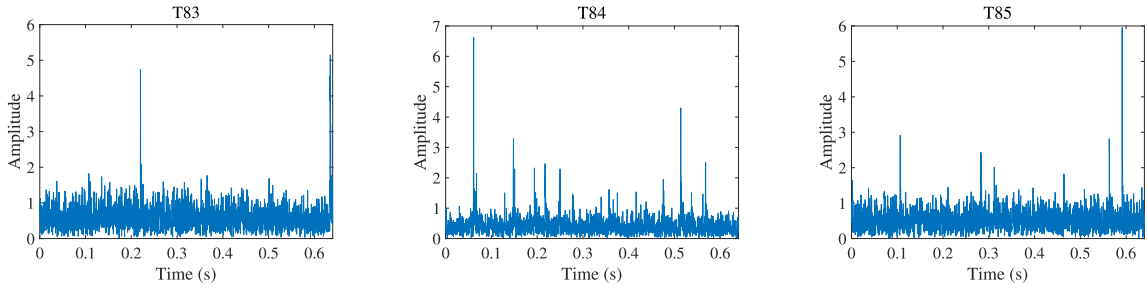


Fig. 6. Amplitude of the received signals in one frame of the three recorded files.

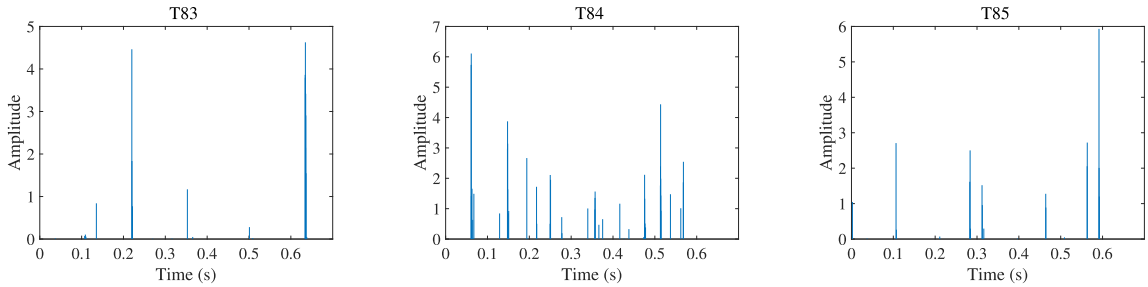


Fig. 7. Amplitude of the estimated impulsive noise in one frame of the three recorded files.

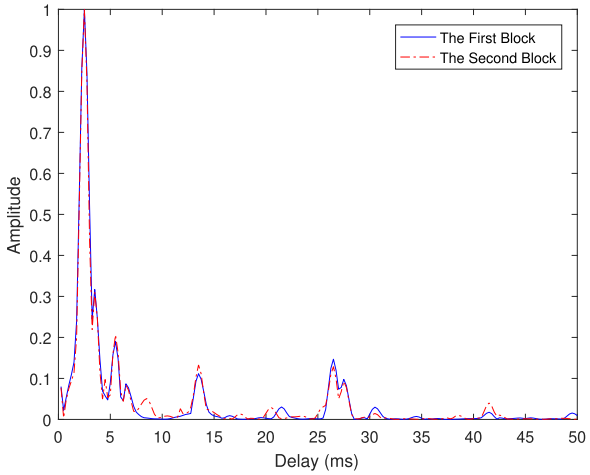


Fig. 8. Amplitude of the channel impulse response over two consecutive OFDM blocks in the T83 file.

TABLE II
PERFORMANCE OF SIX ALGORITHMS FOR THE T83 FILE

Coding Rate	Algorithm	Uncoded BER	Coded BER	FER
1/3	LS+blanking	5.2%	0	0
	JCINE+LS INC	3.5%	0	0
	SBL (pilot subcarriers)	1.8%	0	0
	SBL+KF (pilot subcarriers)	1.5%	0	0
	SBL (all subcarriers)	1.5%	0	0
	SBL+KF (all subcarriers)	1.3%	0	0
1/2	LS+blanking	4.7%	0	0
	JCINE+LS INC	3.3%	0	0
	SBL (pilot subcarriers)	1.9%	0	0
	SBL+KF (pilot subcarriers)	1.5%	0	0
	SBL (all subcarriers)	1.4%	0	0
	SBL+KF (all subcarriers)	1.4%	0	0

The results are shown in Tables II-IV for the three data files, respectively.

We can observe from Table II that since the signals in the T83 file are only slightly contaminated by impulsive noise, all six algorithms obtain zero coded BER and FER for the recorded T83 file. The proposed SBL and SBL+KF algorithms using pilot subcarriers yield a lower uncoded BER than the existing LS+blanking and JCINE+LS INC methods. Moreover, the proposed SBL and SBL+KF algorithms using all subcarriers can obtain a further slight improvement in the uncoded BER performance.

As signals in the T84 file are severely affected by impulsive noise, we can see from Table III that the existing LS+blanking

TABLE III
PERFORMANCE OF SIX ALGORITHMS FOR THE T84 FILE

Coding Rate	Algorithm	Uncoded BER	Coded BER	FER
1/3	LS+blanking	15.5%	1.3%	7.3%
	JCINE+LS INC	14.7%	0.5%	4.1%
	SBL (pilot subcarriers)	12.0%	0.07%	0.4%
	SBL+KF (pilot subcarriers)	11.4%	0	0
	SBL (all subcarriers)	11.0%	0	0
	SBL+KF (all subcarriers)	10.7%	0	0
1/2	LS+blanking	14.6%	15.9%	84.7%
	JCINE+LS INC	13.5%	11.1%	62.9%
	SBL (pilot subcarriers)	11.4%	3.9%	27.8%
	SBL+KF (pilot subcarriers)	10.6%	2.5%	16.5%
	SBL (all subcarriers)	10.5%	2.1%	13.7%
	SBL+KF (all subcarriers)	10.3%	1.0%	8.1%

TABLE IV
PERFORMANCE OF SIX ALGORITHMS FOR THE T85 FILE

Coding Rate	Algorithm	Uncoded BER	Coded BER	FER
1/3	LS+blanking	11.2%	0	0
	JCINE+LS INC	9.1%	0	0
	SBL (pilot subcarriers)	5.6%	0	0
	SBL+KF (pilot subcarriers)	5.2%	0	0
	SBL (all subcarriers)	5.0%	0	0
	SBL+KF (all subcarriers)	4.9%	0	0
1/2	LS+blanking	11.7%	3.9%	24.8%
	JCINE+LS INC	9.8%	0.5%	5.2%
	SBL (pilot subcarriers)	6.8%	0.2%	1.6%
	SBL+KF (pilot subcarriers)	6.3%	0.01%	0.4%
	SBL (all subcarriers)	6.2%	0	0
	SBL+KF (all subcarriers)	6.0%	0	0

and JCINE+LS INC algorithms have a very high FER with the 1/2 rate signals. Interestingly, compared with the JCINE+LS INC algorithm, the proposed SBL algorithm using pilot subcarriers has a 2.1% reduction in the uncoded BER, 7% reduction in the coded BER, and 35% reduction in the FER for the 1/2 rate signals. Moreover, for the 1/3 rate signals, the SBL algorithm using pilot subcarriers reduces the FER from 4.1% to 0.4%. The proposed SBL+KF algorithm using the pilot subcarriers obtains zero coded BER and FER for the 1/3 rate signals recorded, and 11% reduction in the FER over the SBL algorithm with pilot subcarriers for the 1/2 rate signals. After utilizing the data subcarriers, the proposed SBL+KF algorithm has a 1.5% reduction in the coded BER and 8% reduction in the FER compared with the SBL+KF algorithm only using pilot subcarriers in the 1/2 coding rate.

As the signals in the T85 file suffer from a medium level of impulsive noise, we can see from Table IV that for the 1/3 rate signals, all six algorithms achieve zero coded BER and FER for the data recorded. For the 1/2 rate signals, the SBL algorithm using pilot subcarriers introduces around 3% reduction in the uncoded BER and FER. Moreover, the SBL+KF algorithm using pilot subcarriers achieves 0.01% coded BER and 0.4% FER. The proposed algorithms with all subcarriers can achieve zero coded BER and FER in the 1/2 rate signals.

From Tables II-IV, we can conclude that the proposed SBL and SBL+KF algorithms effectively reduce the system BER and FER compared with existing methods. Moreover, results on the experimental data demonstrate that after using both pilot and data subcarriers, the system BER and FER performance can be further improved.

VI. CONCLUSION

In this paper, we have developed four new joint channel and impulsive noise estimation and tracking algorithms for UA OFDM systems. The proposed algorithms efficiently exploit the sparsity of the UA channel and impulsive noise through the sparse Bayesian learning framework. Furthermore, to track the time-varying UA channel, we have proposed the SBL+Kalman filtering algorithm to capture the temporal correlation of the sparse time-varying channel. A joint data detection and channel estimation algorithm has also been proposed to improve the accuracy of channel and impulsive noise estimation. Finally, we have proposed a novel algorithm for joint time-varying channel tracking and data detection with all subcarriers. The performance of the proposed algorithms has been verified through both numerical simulations and data collected during a UA communication experiment. Compared with existing methods, the proposed algorithms can greatly improve the robustness of UA OFDM systems against impulsive noise.

REFERENCES

[1] D. Kilfoyle and A. Baggeroer, "The state of the art in underwater acoustic telemetry," *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 4–27, Jan. 2000.

[2] M. Stojanovic, "Low complexity OFDM detector for underwater acoustic channels," in *Proc. OCEANS*, Boston, MA, USA Sep. 2006.

[3] J. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean," *Mar. Ecol. Prog. Ser.*, vol. 395, pp. 5–20, Dec. 2009.

[4] M. Chitre, S. Kuselan, and V. Pallayil, "Ambient noise imaging in warm shallow waters; Robust statistical algorithms and range estimation," *J. Acoust. Soc. Amer.*, vol. 132, no. 2, pp. 838–847, Aug. 2012.

[5] M. Chitre, S. Ong, and J. Potter, "Performance of coded OFDM in very shallow water channels and snapping shrimp noise," in *Proc. OCEANS*, Washington, DC, USA, vol. 2, Jul. 2006, pp. 996–1001.

[6] T. Suzuki, H. M. Tran, and T. Wada, "An underwater acoustic OFDM communication system with shrimp (impulsive) noise cancelling," in *Proc. Int. Conf. Comput., Manage., Telecommun.*, Da Nang, Vietnam, Apr. 2014, pp. 152–156.

[7] Y. Ma, P. So, and E. Gunawan, "Performance analysis of OFDM systems for broadband power line communications under impulsive noise and multipath effects," *IEEE Trans. Power Del.*, vol. 20, no. 2, pp. 674–682, Apr. 2005.

[8] S. Dimitrov, S. Sinanovic, and H. Haas, "Clipping noise in OFDM-based optical wireless communication systems," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1072–1081, Apr. 2012.

[9] M. Chitre, J. Potter, and S.-H. Ong, "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise," *IEEE J. Ocean. Eng.*, vol. 31, no. 2, pp. 497–503, Apr. 2006.

[10] H. Sun, W. Shen, Z. Wang, S. Zhou, X. Xu, and Y. Chen, "Joint carrier frequency offset and impulse noise estimation for underwater acoustic OFDM with null subcarriers," in *Proc. Oceans*, Hampton, VA, USA, Oct. 2012, pp. 1–4.

[11] S. Wang, Z. He, K. Niu, P. Chen, and Y. Rong, "A sparse Bayesian learning based joint channel and impulsive noise estimation algorithm for underwater acoustic OFDM systems," in *Proc. IEEE Kobe Techno-Oceans (OTO)*, Kobe, Japan, May 2018, pp. 1–5.

[12] P. Chen, Y. Rong, S. Nordholm, and Z. He, "Joint channel and impulsive noise estimation in underwater acoustic OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10567–10571, Nov. 2017.

[13] P. Chen, Y. Rong, and S. Nordholm, "Pilot-subcarrier based impulsive noise mitigation for underwater acoustic OFDM systems," in *Proc. 11th ACM Int. Conf. Underwater Netw. Syst. (WUWNet)*, Shanghai, China, Oct. 2016, p. 20.

[14] P. Chen, Y. Rong, S. Nordholm, A. Duncan, and Z. He, "Compressed sensing based channel estimation and impulsive noise cancelation in underwater acoustic OFDM systems," in *Proc. IEEE Region Conf.*, Singapore, Nov. 2016, pp. 2541–2544.

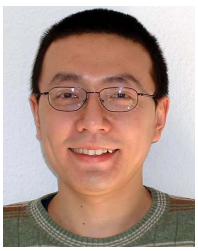
[15] X. Kuai, H. Sun, S. Zhou, and E. Cheng, "Impulsive noise mitigation in underwater acoustic OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8190–8202, Oct. 2016.
[16] P. Chen, Y. Rong, S. Nordholm, Z. He, and A. J. Duncan, "Joint channel estimation and impulsive noise mitigation in underwater acoustic OFDM communication systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6165–6178, Sep. 2017.
[17] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.
[18] E. Karseras, K. Leung, and W. Dai, "Tracking dynamic sparse signals using Hierarchical Bayesian Kalman filters," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Vancouver, BC, Canada, May 2013, pp. 6546–6550.
[19] J. Lin, M. Nassar, and B. L. Evans, "Impulsive noise mitigation in powerline communications using sparse Bayesian learning," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1172–1183, Jul. 2013.
[20] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint approximately sparse channel estimation and data detection in OFDM systems using sparse Bayesian learning," *IEEE Trans. Signal Process.*, vol. 62, no. 14, pp. 3591–3603, Jul. 2014.
[21] A. Mishra, N. Yashaswini, and A. K. Jagannatham, "SBL-based joint sparse channel estimation and maximum likelihood symbol detection in OSTBC MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4220–4232, May 2018.
[22] S. Srivastava, A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Quasi-static and time-selective channel estimation for block-sparse millimeter wave hybrid MIMO systems: Sparse Bayesian learning (SBL) based approaches," *IEEE Trans. Signal Process.*, vol. 67, no. 5, pp. 1251–1266, Mar. 2019.
[23] G. Qiao, Q. Song, L. Ma, S. Liu, Z. Sun, and S. Gan, "Sparse Bayesian learning for channel estimation in time-varying underwater acoustic OFDM communication," *IEEE Access*, vol. 6, pp. 56675–56684, 2018.
[24] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE J. Ocean. Eng.*, vol. 33, no. 2, pp. 198–209, Apr. 2008.
[25] S. Mason, C. Berger, S. Zhou, and P. Willett, "Detection, synchronization, and Doppler scale estimation with multicarrier waveforms in underwater acoustic communication," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 9, pp. 1638–1649, Dec. 2008.
[26] C. Berger, S. Zhou, J. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1708–1721, Mar. 2010.
[27] Y. Huang, L. Wan, S. Zhou, Z. Wang, and J. Huang, "Comparison of sparse recovery algorithms for channel estimation in underwater acoustic OFDM with data-driven sparsity learning," *Phys. Commun.*, vol. 13, pp. 156–167, Dec. 2014.

Shuche Wang received the B.E. degree in information engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2017, where he is currently pursuing the M.Sc. degree in information and telecommunication engineering. His research interests include wireless communications, signal processing, and communication networks.

Zhiqiang He (Member, IEEE) received the B.E. and Ph.D. degrees (Hons.) from the Beijing University of Posts and Telecommunications, China, all in signal and information processing, in 1999 and 2004, respectively. Since July 2004, he has been with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, where he is currently a Professor and the Director of the Center of Information Theory and Technology. His research interests include signal and information processing in wireless communications, networking architecture and protocol design, machine learning, and underwater acoustic communications.

Kai Niu received the B.S. degree in information engineering and the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications (BUPT), China, in 1998 and 2003, respectively. He is currently a Professor with the School of Information and Communication Engineering, BUPT. His research interests include channel coding theory and its applications, such as polar codes and LDPC/turbo codes. He has published over 200 research papers in the fields of 4G/5G, MIMO detection, and channel coding. He is serving as a Senior Member of the Chinese Institute of Electronics (CIE) and the Vice Chairman of the Information Theory Chapter of CIE. He is also an Associate Editor of the IEEE COMMUNICATIONS LETTERS.

Peng Chen received the B.E. degree in information engineering and the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications, China, in 2007 and 2013, respectively. He was a System Engineer with the Mobile Communication Division of Datang Mobile Communications Equipment Co., Ltd., Beijing, China, from April 2013 to September 2014. From October 2014 to November 2018, he was with the Department of Electrical and Computer Engineering, Curtin University, Bentley, Australia, as a Research Associate. His research interests include signal processing, wireless communication systems, algorithm design, and channel estimation.



Yue Rong (Senior Member, IEEE) received the Ph.D. degree *(summa cum laude)* in electrical engineering from the Darmstadt University of Technology, Darmstadt, Germany, in 2005.

He was a Post-Doctoral Researcher with the Department of Electrical Engineering, University of California, Riverside, CA, USA, from February 2006 to November 2007. Since December 2007, he has been with Curtin University, Bentley, Australia, where he is currently a Professor. His research interests include signal processing for communications, wireless communications, underwater acoustic communications, underwater optical wireless communications, applications of linear algebra and optimization methods, and statistical and array signal processing. He has published over 160 journal and conference papers in these areas. He was also a TPC member of the IEEE ICC, IEEE GlobalSIP, EUSIPCO, IEEE ICCC, WCSP, IWCMC, and ChinaCom. He was a recipient of the Best Paper Award at the 2011 International Conference on Wireless Communications and Signal Processing, the Best Paper Award at the 2010 Asia–Pacific Conference on Communications, and the Young Researcher of the Year Award of the Faculty of Science and Engineering at Curtin University in 2010. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2014 to 2018, an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS from 2012 to 2014, and a Guest Editor of the IEEE Journal on Selected Areas in Communications special issue on theories and methods for advanced wireless relays.