RESEARCH

Open Access



Joint channel and impulse noise estimation based on compressed sensing and Kalman filter for OFDM system

Yiting Zhao¹, Youming Li^{1*}, Shoudong Shi¹ and Jianding Yu¹

*Correspondence: liyouming@nbu.edu.cn

¹ Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China

Abstract

Impulse noise (IN) widely exists in many communication systems, which seriously affects the performance of OFDM communication systems. A joint channel and IN estimation method based on all subcarriers is designed. This method uses a sparse Bayesian learning (SBL) algorithm incorporating forward-backward Kalman filter (FB-Kalman) to tackle the problem of joint channel and IN estimation and data detection for OFDM systems. Firstly, the channel impulse response and IN are regarded as unknown sparse vectors, and a SBL framework using all subcarriers is proposed to estimate the unknown vector. The SBL theory is used based on the prior distribution of variables, and then the forward-backward joint system is established, which applies the data detection simultaneously. We also propose the FB-Kalman implementation algorithm by using the expectation maximization updates. Explicit expressions of mean and covariance matrix of the posterior distribution are derived in the E-step. Simulation results show that the proposed algorithm improves the normalized mean square error and bit error rate performance of OFDM system in the presence of IN communication environment.

Keywords: Impulse noise, OFDM, Channel estimation, Sparse Bayesian learning

1 Introduction

IN always exists and is more harmful especially in OFDM system such as digital subscriber lines [1], power line communication (PLC) [2], wireless communication [3], and underwater acoustic communication [4]. IN model such as Gaussian mixture (GM), Middleton Class A [5], and Bernoulli-Gaussian [6] can cover most scenarios where IN occurs in PLC systems, and α -stable distribution model can be used to describe the some short-term spikes existed in shallow water acoustic channel [7].

Although OFDM is inherently more resistant to IN than single carrier modulation, its influence extends to all subcarriers and if the power of IN exceeds a certain threshold, it will lead to a sharp decline in system performance [8]. It follows that in practical applications, the IN should be properly suppressed in the receiver side before detection.

Numerous algorithms have been proposed. One type of commonly used IN mitigation algorithms is nonlinear filtering in which the received signal samples contaminated



© The Author(s) 2023. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/

by IN is then adjusted by using nonlinear blanking or clipping method. The drawback of such algorithms is that it not only suppresses the interference, but also destroys the original signal. Moreover, such algorithms need to use the prior statistics of IN to derive the optimal threshold, which is very difficult to obtain in practice, and the suboptimal threshold is usually set by using an empirical value. Another kind of suppression algorithms is based on compressed sensing (CS) theory. These algorithms are proposed base on the sparsity of IN and CIR [9-14], and their performances are superior to traditional blanking and clipping algorithms [15]. However, these algorithms have high computational complexity. In [16], the generalized approximate message passing (AMP) algorithm is used to solve the problems of joint channel coefficients, IN, and data symbols estimation. However, the algorithms based on prior information about CIR and IN, which is difficult to obtain. In [17], the distribution of CIR and IN are modeled as Gaussian Mixture (GM) model, which is the prior information of AMP algorithm. Nevertheless, this assumption does not apply to other channels. In [18], a SBL algorithm combined with Kalman filter is proposed to solve the joint channel and IN estimation problem in OFDM systems. Compared with SBL algorithm, it improves the estimation performance of the system. However, the estimation performance is limited by the number of pilots, increasing the number of pilots will reduce the spectral efficiency of OFDM system.

The problem of joint channel and IN estimation based on all subcarriers in the presence of IN for OFDM system is discussed. Initially, we regard the CIR and IN as an unknown sparse vector and use a SBL framework that employs all subcarriers to estimate unknown vector. The SBL theory is used based on the prior distribution of variables, and then the forward backward joint system is established, which apply the data detection simultaneously [19]. In addition, we develop an FB-Kalman implementation algorithm using EM updates to iteratively estimate the unknown parameters. In the E-step, explicit expressions of mean and covariance matrices of the posterior distribution are derived. In the M-step the unknown parameters are iteratively estimated. Simulation results show that the proposed algorithm improves the channel estimation performance and BER performance of OFDM system in the presence of IN communication environment.

2 System model

Consider a coded frame based OFDM system with N subcarriers, which contains N_d subcarriers for data transmission, N_p pilot subcarriers, and N_n null subcarriers. At the transmitting side, an OFDM frame is composed of N_b OFDM symbols. At the receiver side, the *m*th OFDM symbol $\mathbf{r}_m \in \mathbb{C}^{N \times 1}$ of the received frame in the frequency domain can be written in a vector form as

$$\boldsymbol{r}_m = \mathbf{D}_m \boldsymbol{h}_m + \boldsymbol{i}_m + \boldsymbol{w}_m \tag{1}$$

where $\mathbf{D}_{\mathbf{m}} = diag(\mathbf{d}_m)$ is a diagonal matrix with $\mathbf{d}_m \in \mathbb{C}^{N \times 1}$ as the main diagonal element which represents the *m*th transmitted OFDM symbol. $\mathbf{h}_m \in \mathbb{C}^{N \times 1}$ denotes the frequency domain CIR. $\mathbf{i}_m \in \mathbb{C}^{N \times 1}$ and $\mathbf{w}_m \in \mathbb{C}^{N \times 1}$ are the frequency domain IN and background Gaussian noise, respectively. The subscript $m(m = 1, 2, ..., N_b)$ represents the index of OFDM symbols in a frame. By introducing $\mathbf{h}_m = \sqrt{N} \mathbf{F}_L \mathbf{h}_{t,m}$ and $\mathbf{i}_m = \mathbf{F} \mathbf{i}_{t,m}$, where $\mathbf{F} \in \mathbb{C}^{N \times N}$ represents the normalized discrete Fourier transform matrix, $\mathbf{F}_L \in \mathbb{C}^{N \times L}$ is the submatrix selected by the first *L* columns of matrix \mathbf{F} , where *L* represents the channel length, $\mathbf{h}_{t,m} \in \mathbb{C}^{L \times 1}$ and $\mathbf{i}_{t,m} \in \mathbb{C}^{N \times 1}$ denote the time domain CIR and IN, respectively. Equation (1) can be rewritten as

$$\boldsymbol{r}_m = \sqrt{N} \boldsymbol{D}_m \boldsymbol{F}_L \boldsymbol{h}_{t,m} + \boldsymbol{F} \boldsymbol{i}_{t,m} + \boldsymbol{w}_m \tag{2}$$

3 Joint channel and impulsive noise estimation

In this section, based on SBL framework and Kalman filter, we propose the following new algorithm to jointly estimate impulse noise and channel. Let $\Phi_m = [\sqrt{N} \mathbf{D}_m \mathbf{F}_L, \mathbf{F}] \in \mathbb{C}^{N \times (L+N)}$ and $\theta_m = (\mathbf{h}_{t,m}^T, \mathbf{i}_{t,m}^T)^T \in \mathbb{C}^{L+N}$, Eq. (2) can be expressed as

$$\boldsymbol{r}_m = \boldsymbol{\Phi}_m \boldsymbol{\theta}_m + \boldsymbol{w}_m \tag{3}$$

Because the CIR $h_{t,m}$ and IN $i_{t,m}$ are regarded as unknown sparse vectors, θ_m formed by $h_{t,m}$ and $i_{t,m}$, can also be regarded as an sparse vector. Noting that the matrix Φ_m is an underdetermined matrix, estimating the unknown vector by Eq. (3) translates into a CS problem.

We assume that θ_m is unchanged in one OFDM symbol, but changes between symbol to symbol according to a state-space model. Then, taking Eq. (3) as the observation equation, the state equation can be expressed as

$$\boldsymbol{\theta}_m = \mathbf{A}\boldsymbol{\theta}_{m-1} + \boldsymbol{\nu}_m \tag{4}$$

where $\mathbf{A} \triangleq diag(\rho \mathbf{1}_L, \mathbf{0}_N)$, the $\mathbf{1}_L$ is a *L*-length vector with all elements as one, and $\mathbf{0}_N$ is a *N*-length vector with all elements as zero. $\boldsymbol{\theta}_m \sim \mathcal{CN}(0, \mathbf{B}\boldsymbol{\Gamma}_m)$, where $\boldsymbol{\Gamma}_m = diag(\boldsymbol{\Gamma}_{(h)}, \boldsymbol{\Gamma}_{(i)_m})$ and $\mathbf{B} \triangleq diag((1 - \rho^2)\mathbf{I}_L, \mathbf{I}_N)$. $\boldsymbol{\Gamma}_{(h)} = diag(\gamma_0, \gamma_1, ..., \gamma_{L-1})$, $\boldsymbol{\Gamma}_{(i)} = diag(\gamma_L, \gamma_{L+1}, ..., \gamma_{L+N-1})$ are the covariance matrices of the corresponding CIR and IN, respectively. The excitation noise $\boldsymbol{\nu}_m$ and the observation noise \boldsymbol{w}_m are independent Gaussian white noise with zero mean i.e., $\boldsymbol{\nu}_m \sim \mathcal{CN}(0, \mathbf{B}\boldsymbol{\Gamma}_m)$, $\boldsymbol{w}_m \sim \mathcal{CN}(0, \beta_m \mathbf{I}_N)$ and \mathbf{I}_N is an $N \times N$ identity matrix. ρ is the correlation coefficient of the state transition, and β_m is a scalar corresponding to the background noise.

We assume the following forward and backward systems

$$\boldsymbol{r}_m^f = \boldsymbol{\Phi}_m^f \boldsymbol{\theta}_m + \boldsymbol{w}_m^f \tag{5}$$

$$\boldsymbol{r}_m^b = \boldsymbol{\Phi}_m^b \boldsymbol{\theta}_m + \boldsymbol{w}_m^b \tag{6}$$

where the subscripts f denote the variable in the forward systems and b denote the variables in the backward systems, respectively. In the process of forward Kalman filter, the filtering process is initialized as $\theta_{0|0}^f = 0$, $\Sigma_{0|0}^f = \mathbf{I}_{L+N}$.

In the prediction step, we have

$$\boldsymbol{\theta}_{m|m-1}^{f} = \mathbf{A}\boldsymbol{\theta}_{m-1|m-1}^{f} \tag{7}$$

$$\boldsymbol{\Sigma}_{m|m-1}^{f} = \mathbf{A} \boldsymbol{\Sigma}_{m-1|m-1}^{f} \mathbf{A}^{\mathrm{T}} + \mathbf{B} \boldsymbol{\Gamma}_{m}^{f}$$
(8)

$$\boldsymbol{e}_{m}^{f} = \boldsymbol{r}_{m}^{f} - \boldsymbol{r}_{m|m-1}^{f} \tag{9}$$

$$\boldsymbol{r}_{m|m-1}^{f} = \boldsymbol{\Phi}_{m}^{f} \boldsymbol{\theta}_{m|m-1}^{f} \tag{10}$$

where $\theta_{m|m-1}^{f}$ denotes the *m*th state prediction obtained by using the result of the m-1th optimal state. $\theta_{m-1|m-1}^{f}$ is the result of the m-1th optimal state. e_{m}^{f} denotes the difference between the *m*th measured and predicted value. $\Sigma_{m|m-1}^{f}$ and $\Sigma_{m-1|m-1}^{f}$ are the *m*th prior and the m-1th posterior covariance matrices of estimation errors, respectively.

In the update step, we can compute

$$\mathbf{K}_{\mathbf{m}}^{\mathbf{f}} = \boldsymbol{\Sigma}_{m|m-1}^{f} (\boldsymbol{\Phi}_{m}^{f})^{H} \left((\boldsymbol{\beta}_{m}^{f}) \mathbf{I}_{\mathbf{N}} + \boldsymbol{\Phi}_{m}^{f} \boldsymbol{\Sigma}_{m|m-1}^{f} (\boldsymbol{\Phi}_{m}^{f})^{H} \right)^{-1}$$
(11)

$$\boldsymbol{\theta}_{m|m}^{f} = \boldsymbol{\theta}_{m|m-1}^{f} + \mathbf{K}_{\mathbf{m}}^{\mathbf{f}} \boldsymbol{e}_{m}^{f}$$
(12)

$$\boldsymbol{\Sigma}_{m|m}^{f} = (\mathbf{I} - \mathbf{K}_{\mathbf{m}}^{\mathbf{f}} \boldsymbol{\Phi}_{m}^{f}) \boldsymbol{\Sigma}_{m|m-1}^{f}$$
(13)

where $\theta_{m|m}^{f}$ is the best estimate of the *m*th state, $\mathbf{K}_{\mathbf{m}}^{\mathbf{f}}$ is the Kalman gain, and e_{m}^{f} is given by (9).

The process of backward Kalman filtering is similar to the forward Kalman filtering in (7)–(13), except that the former is from time $m = N_b$ to m = 1.

For (5) and (6), we can apply the linear minimum mean-squared estimator (LMMSE) for each linear system to estimate θ_m in each linear system. For the forward system, the estimated value of θ_m is

$$\hat{\boldsymbol{\theta}}_{m}^{f} = \boldsymbol{\Sigma}_{m}^{f} (\boldsymbol{\Phi}_{m}^{f})^{H} \mathbf{R}_{\boldsymbol{w}_{m}^{f}}^{-1} \boldsymbol{r}_{m}^{f}$$
(14)

where $\Sigma_m^f = (\mathbf{R}_{\theta_m}^{-1} + (\Phi_m^f)^H \mathbf{R}_{w_m^f}^{-1} \Phi_m^f)^{-1}$ is the *m*th estimation error matrix, \mathbf{R}_{θ_m} and $\mathbf{R}_{w_m^f}$ are the covariance matrices of θ_m and w_m^f , respectively. Similarly, the estimated value of θ_m from the backward system is

$$\hat{\boldsymbol{\theta}}_{m}^{b} = \boldsymbol{\Sigma}_{m}^{b} (\boldsymbol{\Phi}_{m}^{b})^{H} \mathbf{R}_{\boldsymbol{w}_{\mathbf{b}}^{\mathbf{b}}}^{-1} \boldsymbol{r}_{m}^{b}$$
(15)

where $\Sigma_m^b = (\mathbf{R}_{\theta_m}^{-1} + (\Phi_m^b)^H \mathbf{R}_{w_m^b}^{-1} \Phi_m^b)^{-1}$ is the *m*th estimation error matrix, $\mathbf{R}_{w_m^b}$ is the covariance matrix of w_m^b .

Combining (5) and (6), we have

$$\begin{pmatrix} \boldsymbol{r}_m^f \\ \boldsymbol{r}_m^b \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Phi}_m^f \\ \boldsymbol{\Phi}_m^b \end{pmatrix} \boldsymbol{\theta}_m + \begin{pmatrix} \boldsymbol{w}_m^f \\ \boldsymbol{w}_m^b \end{pmatrix}$$
(16)

The LMMSE estimation of θ from (16) is given by

$$\hat{\boldsymbol{\theta}}_{m}^{fb} = \left(\left[(\boldsymbol{\Phi}_{m}^{f})^{H} \ (\boldsymbol{\Phi}_{m}^{b})^{H} \right] \begin{bmatrix} \mathbf{R}_{w_{m}^{f}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{w_{m}^{b}}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{m}^{f} \\ \boldsymbol{\Phi}_{m}^{b} \end{bmatrix} + \mathbf{R}_{\boldsymbol{\theta}_{m}}^{-1} \right)^{-1} \\ \times \left[(\boldsymbol{\Phi}_{m}^{f})^{H} \ (\boldsymbol{\Phi}_{m}^{b})^{H} \right] \begin{bmatrix} \mathbf{R}_{w_{m}^{f}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{w_{m}^{b}}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{r}_{m}^{f} \\ \boldsymbol{r}_{m}^{b} \end{bmatrix} \\ = ((\boldsymbol{\Phi}_{m}^{f})^{H} \mathbf{R}_{w_{m}^{f}}^{-1} \boldsymbol{\Phi}_{m}^{f} + (\boldsymbol{\Phi}_{m}^{b})^{H} \mathbf{R}_{w_{m}^{b}}^{-1} \boldsymbol{\Phi}_{m}^{b} + \mathbf{R}_{\boldsymbol{\theta}_{m}}^{-1})^{-1} \\ \times ((\boldsymbol{\Phi}_{m}^{f})^{H} \mathbf{R}_{w_{m}^{f}}^{-1} \boldsymbol{r}_{m}^{f} + (\boldsymbol{\Phi}_{m}^{b})^{H} \mathbf{R}_{w_{m}^{b}}^{-1} \boldsymbol{r}_{m}^{b})$$
(17)

The estimation error matrix is

$$\boldsymbol{\Sigma}_{m}^{\hat{\boldsymbol{\theta}}_{fb}} = \left(\begin{bmatrix} (\boldsymbol{\Phi}_{m}^{f})^{H} & (\boldsymbol{\Phi}_{m}^{b})^{H} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{w_{m}^{f}}^{-1} & 0\\ 0 & \mathbf{R}_{w_{m}^{b}}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{m}^{f} \\ \boldsymbol{\Phi}_{m}^{b} \end{bmatrix} + \mathbf{R}_{\boldsymbol{\theta}_{m}}^{-1} \right)^{-1}$$

$$= ((\boldsymbol{\Phi}_{m}^{f})^{H} \mathbf{R}_{w_{m}^{f}}^{-1} \boldsymbol{\Phi}_{m}^{f} + (\boldsymbol{\Phi}_{m}^{b})^{H} \mathbf{R}_{w_{m}^{b}}^{-1} \boldsymbol{\Phi}_{m}^{b} + \mathbf{R}_{\boldsymbol{\theta}_{m}}^{-1})^{-1}$$

$$(18)$$

According to equation (14) and (15), (17) and (18) can be simplified, respectively, as

$$\hat{\boldsymbol{\theta}}_{m}^{fb} = ((\boldsymbol{\Sigma}_{m}^{f})^{-1} + (\boldsymbol{\Sigma}_{m}^{b})^{-1} - (\mathbf{R}_{\boldsymbol{\theta}_{m}})^{-1})^{-1} \\ ((\boldsymbol{\Sigma}_{m}^{f})^{-1}\hat{\boldsymbol{\theta}}_{m}^{f} + (\boldsymbol{\Sigma}_{m}^{b})^{-1}\hat{\boldsymbol{\theta}}_{m}^{b})$$
(19)

$$\Sigma_{\hat{\theta}_{m}^{fb}} = ((\Sigma_{m}^{f})^{-1} + (\Sigma_{m}^{b})^{-1} - (\mathbf{R}_{\theta_{m}})^{-1})^{-1}$$
(20)

As Kalman filtering is an extension of sequential LMMSE estimation, the forward and backward estimations can be combined following (19) and (20), respectively, as

$$\boldsymbol{\theta}_{m|m} = \left((\boldsymbol{\Sigma}_{m|m}^{f})^{-1} + (\boldsymbol{\Sigma}_{m|m}^{b})^{-1} \right)^{-1} \\ \left((\boldsymbol{\Sigma}_{m|m}^{f})^{-1} \boldsymbol{\theta}_{m|m}^{f} + ((\boldsymbol{\Sigma}_{m|m}^{b})^{-1} \boldsymbol{\theta}_{m|m}^{b} \right)$$
(21)

$$\Sigma_{m|m} = \left((\Sigma_{m|m}^{f})^{-1} + (\Sigma_{m|m}^{b})^{-1} \right)^{-1}$$
(22)

where $\theta_{m|m}^{f}$ and $\Sigma_{m|m}^{f}$ are given by (12) and (13), respectively, $\theta_{m|m}^{b}$ and $\Sigma_{m|m}^{b}$ obtained in a similar way. Since there is no prior auto-correlation information about θ , we set $\mathbf{R}_{\theta}^{-1} = 0$ in (19) and (20).

In (7)–(13), we see that in addition to the unknown parameters Γ_m and β , there are some unknown matrix elements of the matrix Φ which just based on unknown data symbol **D**.

Let $\boldsymbol{\xi} \triangleq \{\{\boldsymbol{\Gamma}_m\}, \boldsymbol{\beta}, \mathbf{D}_m\}$ as the set of unknown parameters, the steps of EM algorithm are as follows:

In the E-step, the expected value of joint probability density $p(\{r_m\}, \{\theta_m\}, \{\Gamma_m\}, \beta)$ under posterior probability distribution $p(\theta|r)$ is obtained from

$$Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(k)}) = E_{\{\boldsymbol{\theta}\}|\{\boldsymbol{r}\}} \Big\{ \log p(\{\boldsymbol{r}\}, \{\boldsymbol{\theta}\}, \{\boldsymbol{\Gamma}_m^{(k)}\}, \boldsymbol{\beta}^{(k)}) \Big\}$$
(23)

where *k* represents the number of iterations.

In the M-step, first fix one parameter, solve another parameter, and then solve the optimal parameter alternately as

$$\boldsymbol{\xi}^{(k+1)} = \operatorname*{arg\,max}_{\boldsymbol{\xi}} Q(\boldsymbol{\xi}|\boldsymbol{\xi}^{(k)}) \tag{24}$$

The likelihood function of (2) can be written

$$p(\lbrace \boldsymbol{r}\rbrace, \lbrace \boldsymbol{\theta}\rbrace, \lbrace \boldsymbol{\Gamma}_m\rbrace, \beta) = p(\boldsymbol{\theta}_1; \boldsymbol{\Gamma}_1) \prod_{m=2}^{N_b} p(\boldsymbol{\theta}_m | \boldsymbol{\theta}_{m-1}; \boldsymbol{\Gamma}_m) \times \prod_{m=1}^{N_b} p(\boldsymbol{r}_m | \boldsymbol{\theta}_m; \beta)$$
(25)

Form (24), we obtain the log-likelihood function as

$$\log p(\lbrace \boldsymbol{r} \rbrace, \lbrace \boldsymbol{\theta} \rbrace) \propto -\sum_{m=1}^{N_b} \frac{\|\boldsymbol{r}_m - \boldsymbol{\Phi}_m \boldsymbol{\theta}_m\|_2^2}{\beta}$$
$$-NN_b \log \beta - \sum_{m=1}^{N_b} \log |\boldsymbol{\Gamma}_m|$$
$$-\sum_{m=2}^{N_b} (\boldsymbol{\theta}_m - \mathbf{A} \boldsymbol{\theta}_{m-1})^H (\mathbf{B} \boldsymbol{\Gamma}_m)^{-1} (\boldsymbol{\theta}_m - \mathbf{A} \boldsymbol{\theta}_{m-1}) - \boldsymbol{\theta}_1^H \boldsymbol{\Gamma}_1^{-1} \boldsymbol{\theta}_1$$
(26)

By analyzing the expression (26), the optimal Γ_m can be obtained as

$$\left\{ \boldsymbol{\Gamma}_{m}^{(k+1)} \right\} = \underset{\{\boldsymbol{\Gamma}_{m}\}}{\operatorname{arg\,min}} E_{\{\boldsymbol{r}\}|\{\boldsymbol{\theta}\}} \left\{ \sum_{m=1}^{N_{b}} \log |\boldsymbol{\Gamma}_{m}| + \boldsymbol{\theta}_{1}^{H} \boldsymbol{\Gamma}_{1}^{-1} \boldsymbol{\theta}_{1} + \sum_{m=2}^{N_{b}} (\boldsymbol{\theta}_{m} - \mathbf{A}\boldsymbol{\theta}_{m-1})^{H} (\mathbf{B}\boldsymbol{\Gamma}_{m})^{-1} (\boldsymbol{\theta}_{m} - \mathbf{A}\boldsymbol{\theta}_{m-1}) \right\}$$
$$= \underset{\{\boldsymbol{\Gamma}_{m}\}}{\operatorname{arg\,min}} \left\{ \sum_{m=1}^{N_{b}} \log |\boldsymbol{\Gamma}_{m}| + Tr(\boldsymbol{\Gamma}_{1}^{-1}\mathbf{M}_{1|1}) + \sum_{m=2}^{N_{b}} Tr(\boldsymbol{\Gamma}_{m}^{-1}\mathbf{B}^{-1}\mathbf{M}_{m|m}) \right\}$$
(27)

where $\mathbf{M}_{\mathbf{m}|\mathbf{m}} = \boldsymbol{\Sigma}_{m|m} + \hat{\boldsymbol{\theta}}_{m|m} \hat{\boldsymbol{\theta}}_{m|m}^{H}$. Similarly, β can be optimized

$$\beta^{(k+1)} = \arg\min_{\beta} E_{\{\theta\}|\{r\}} \left\{ \sum_{m=1}^{N_b} \frac{\|\boldsymbol{r}_m - \boldsymbol{\Phi}_m \boldsymbol{\theta}_{m|m}\|_2^2}{\beta} + NN_b \log \beta \right\}$$
(28)

By setting the derivative of (28) with respect to β as zero, we obtain

$$\beta^{(k+1)} = \frac{1}{NN_b} Tr \left[\sum_{m=1}^{N_b} (\boldsymbol{r}_m - \boldsymbol{\Phi}_m \boldsymbol{\theta}_{m|m}) (\boldsymbol{r}_m - \boldsymbol{\Phi}_m \boldsymbol{\theta}_{m|m})^H + \boldsymbol{\Phi}_m \boldsymbol{\Sigma}_{m|m} \boldsymbol{\Phi}_m^H \right]$$
(29)

After obtaining the estimated values of $\Gamma_m^{(k+1)}$ and $\beta^{(k+1)}$, respectively, the estimated values of data symbol **D**_m is derived as follows.

$$D_{m}^{(k+1)} = \arg \max_{\mathbf{D}} Q\left(\mathbf{D}^{(\mathbf{k}+1)}; \mathbf{\Phi}^{(k)}, \{\mathbf{\Gamma}_{m}^{(k+1)}\}, \beta^{(k+1)}\right)$$

$$= \arg \max_{\mathbf{D}} \left\{ c - E_{\{r\}|\{\theta\}} \left\{ \frac{\|\mathbf{r}_{m} - \mathbf{\Phi}_{m} \boldsymbol{\theta}_{m}\|_{2}^{2}}{\beta} \right\} \right\}$$

$$= \arg \max_{\mathbf{D}} \left\{ c - \beta^{-1} \left[\|\mathbf{r}_{m} - \mathbf{\Phi}_{m} \boldsymbol{\theta}_{m}\|_{2}^{2} + Tr(\mathbf{\Phi}_{m}^{H} \boldsymbol{\Sigma}_{m} \mathbf{\Phi}_{m}) \right] \right\}$$

$$= \arg \min_{\mathbf{D}} \|\mathbf{r}_{m} - \mathbf{\Phi}_{m} \boldsymbol{\theta}_{m}\|_{2}^{2} + Tr(\mathbf{\Phi}_{m}^{H} \boldsymbol{\Sigma}_{m} \mathbf{\Phi}_{m})$$
(30)

where *c* is a constant independent of *D*, and *Tr*() denotes the matrix trace. We have

$$\|\boldsymbol{r}_{m} - \boldsymbol{\Phi}_{m}\boldsymbol{\theta}_{m}\|_{2}^{2} + Tr(\boldsymbol{\Phi}_{m}^{H}\boldsymbol{\Sigma}_{m}\boldsymbol{\Phi}_{m})$$

$$= \|\boldsymbol{r}_{m} - \boldsymbol{\Phi}_{m}\boldsymbol{\theta}_{m}\|_{2}^{2} + Tr(N\boldsymbol{D}_{m}\boldsymbol{F}_{L}\boldsymbol{\Sigma}_{1,1}\boldsymbol{F}_{L}^{H}\boldsymbol{D}_{m}^{H}$$

$$+ \sqrt{N}\boldsymbol{F}\boldsymbol{\Sigma}_{2,1}\boldsymbol{F}_{L}^{H}\boldsymbol{D}_{m}^{H} + \sqrt{N}\boldsymbol{D}_{m}\boldsymbol{F}_{L}\boldsymbol{\Sigma}_{1,2}\boldsymbol{F}^{H} + \boldsymbol{F}\boldsymbol{\Sigma}_{2,2}\boldsymbol{F}^{H})$$
(31)

It is reasonable to assume that the CIR vector and IN vector are independent of each other, so that the covariance submatrices $\Sigma_{2,1}$ and $\Sigma_{1,2}$ in Eq. (31) can be set to zero matrix.

Denotes I_d as the indices of subcarriers with data symbols. By substituting (31) into (30), the iterative update formula of each data symbol $d_m[i]$, i.e., the *i*th diagonal element of d_m , $i \in I_d$, can be obtained by solving the following problem

$$(d_m[i])^{(k+1)} = \underset{d_m[i]}{\arg\min} |r_m[i] - \sqrt{N} d_m[i] \mathbf{F}_L[i, :] \boldsymbol{\theta}_{m|m,1} - \mathbf{F}[i, :] \boldsymbol{\theta}_{m|m,2}|^2 + \mathbf{C}_{\mathbf{m},\mathbf{1}}[i, i] |d_m[i]|^2$$
(32)

where

$$\mathbf{C}_{\mathbf{m},\mathbf{1}}[i,i] = N\mathbf{F}\boldsymbol{\Sigma}_{m|m,1,1}\mathbf{F}^H$$
(33)

$$\boldsymbol{\Sigma}_{m|m} = \begin{pmatrix} \boldsymbol{\Sigma}_{m|m,1,1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{m|m,2,2} \end{pmatrix}$$
(34)

$$\boldsymbol{\theta}_{m|m} = \begin{bmatrix} \boldsymbol{\theta}_{m|m,1}^T, \boldsymbol{\theta}_{m|m,2}^T \end{bmatrix}^T$$
(35)

and $\mathbf{F}[i,:]$ is the *i*th row of \mathbf{F} , $\boldsymbol{\theta}_{m|m}$ and $\boldsymbol{\Sigma}_{m|m}$ can be obtained, respectively, by (21) and (22).

The entire algorithm of FB-Kalman is summarized in Algorithm 1.

```
Algorithm 1 FB-Kalman
Input: r_m, \Phi_m, 1 \le m \le N_b; k_{max}
Output: \hat{\theta}_m, 1 \le m \le N_b
  Give the initial value: \Gamma_m^{(0)} = \mathbf{I} for k = 1, ..., N_b and \beta^{(0)} = 1
   while k \leq k_{max} do
       E-step:
       Set \hat{\theta}_{0|0} = 0, \Sigma_{0|0} = \Gamma_1^{(k)}
       for m = 1 to N_b do f
Update \theta_{m|m}^f and \Sigma_{m|m}^f based on Eq.(7)-(13)
       end for
       for m = N_b to 1 do
Update \theta^b_{m|m} and \Sigma^b_{m|m}
       end for
       Update \theta_{m|m} and \Sigma_{m|m} based on Eq.(21)-(22)
       M-step
        Update M_{m|m} = \Sigma_{m|m} + \hat{\theta}_{m|m} \hat{\theta}_{m|m}^{H}
       Update \Gamma_m^{(k+1)} based on Eq.(27)
        Update \beta based on Eq.(29)
        Update D based on Eq.(32)-(35)
  end while
   return \theta_{m|m};
```

4 Complexity analysis

The FB-Kalman algorithm includes three major steps: forward filtering, backward filtering, and EM parameters estimation. In the forward filtering step, most of the calculations are used to calculate the Kalman gain \mathbf{K}_m , which has a complexity of $O(N^3)$ for each OFDM symbol. The backward filtering step has a computational complexity of $O(N^3)$ for each OFDM symbol. Moreover, the complexity of the EM step is $O(I_1N^3)$, where I_1 is the average number of iterations of the EM process till convergence. Thus, for each OFDM symbol, the computational complexity order of the FB-Kalman algorithm is $O((I_1 + 2)N^3)$.

The JCI-Kalman algorithm uses pilot subcarriers, the complexity of the filtering step is $O(N_p^3)$ for each OFDM symbol, and the complexity of the EM step is $O(I_2N_p^3)$ for each OFDM symbol, where I_2 is the average number of iterations of the EM process till convergence. So, the computational complexity of this algorithm is $O((I_2 + 1)N_p^3)$ per OFDM symbol.

The SBL-LS algorithm uses SBL method to remove IN by using null subcarriers, and the CIR is estimated by least square method. The EM step has a computational complexity of $O(I_3N_n^3)$ for each OFDM symbol, where I_3 is the average number of iterations of the EM process till convergence. And the complexity of the LS step is $O(N_p^3)$ for each OFDM symbol. Thus the SBL-LS algorithm has a computational complexity of $O(I_3N_n^3 + N_p^3)$.

From the above analysis, we see that the FB-Kalman algorithm using all subcarriers has higher complexity than the JCI-Kalman algorithm using only pilot subcarriers and the SBL-LS algorithm using pilot subcarriers and null subcarriers. However, the simulation results show that the proposed FB-Kalman algorithm provides best performance in NMSE and BER. Similarly, compared with the original SBL based algorithm [10], the proposed FB-Kalman algorithm has higher computational complexity and better

performance by using filtering and smoothing operations to track time-varying channels and IN. These trade-offs of performance and complexity are meaningful for practical OFDM systems.

5 Results and discussion

In this section, we demonstrate the performance of the proposed joint channel and IN estimation algorithms. An OFDM system with N = 256, $N_p = 44$, $N_d = 162$, and $N_n = 50$ is simulated. The Rayleigh-fading uncorrelated-scattering model with sparse impulse response [20] is adopted. Each OFDM frame is composed of $N_b = 7$ OFDM symbols.

The noise including IN and background noise is realized by the publicly available software [21], which uses a Gaussian mixture model to simulate the IN distribution, where the probability of each noise component is 0.9, 0.07, 0.03, and the corresponding power of each noise component is 1, 100, 1000. In the noise generated in the simulation, 7% of the IN components exceed the background noise power by 20 dB, and about 3% of the IN components exceed the background noise power by 30 dB.

In addition, we also consider IN environment with α -stable distribution. To verify the performance of channel and IN estimation, α -stable IN model with characteristic exponent $\alpha = 1$, skewness parameter $\beta = 0$, scale parameter $\gamma = 0.05$, and location parameter $\delta = 0$ is used in the simulation.

The performance of the FB-Kalman algorithm is compared with the other two algorithms in this section:

- **SBL-LS:** assume that the IN is mitigated by using the SBL method with null subcarriers [13], the tap-aware LS method is used to estimate the CIR.
- **JCI-Kalman:** assume that the IN and CIR are jointly estimated by using the JCI-Kalman method with pilot subcarriers [18].

Figure 1 compares the channel estimation performance of all the above mitigation algorithms for convolutional coded systems with 1/2 rate using 4-QAM. The figure shows that the NMSE of the three algorithms gradually decreases with the increase in signal-to-noise ratio (SNR). The proposed algorithm performs better than SBL-LS and JCI-Kalman. Compared with SBL-LS, the SNR improvement of the proposed algorithm is more than 10 dB, compared with JCI-Kalman algorithm, the SNR improvement is about 7 dB. Compared with SBL-LS, JCI-Kalman has a gain of about 4 dB in SNR. It can be seen that the joint channel and IN estimation algorithm has better performance than the single estimation. The reason why the proposed algorithm has better performance is that the forward message and the backward message can provide more useful information.

Figure 2 provides the BER performance for convolutional coded systems using 4-QAM. As shown in the figure, the BER of the three algorithms gradually decreases with the increase in SNR. The proposed algorithm is always better than the other two algorithms, and until the SNR approaches 10 dB, the BER of JCI-Kalman and FB-Kalman gradually approaches and then to zero. When BER = 10^{-3} , the proposed algorithm achieves about 2 dB gain in SNR over JCI-Kalman and about 7 dB gain in SNR over SBL-LS.



Fig. 1 Channel estimation performance comparison for various mitigation methods in coded 4-QAM system



Fig. 2 BER performance comparison for various mitigation methods in coded 4-QAM system

Figure 3 compares the channel estimation performance of three mitigation algorithms for uncoded systems using 4-QAM. Similar to the coded system, the NMSE of three algorithms decreases with the increase in SNR. The proposed algorithm performs better than the other two algorithms. Compared with SBL-LS, the SNR improvement of the proposed algorithm is more than 8 dB, compared with JCI-Kalman, the SNR improvement is more than 5 dB. Compared with SBL-LS algorithm, JCI-Kalman algorithm has a gain of about 3 dB in SNR. Compared with the coded system, the uncoded system has higher NMSE. It shows that channel coding can effectively improve the signal channel estimation performance.



Fig. 3 Channel estimation performance comparison for various mitigation methods in uncoded 4-QAM system



Fig. 4 Channel estimation performance comparison for various mitigation methods in uncoded 4-QAM system

Figure 4 provides the BER performance for uncoded systems using 4-QAM. Similar to the coded system, the figure also shows that the BER of three algorithms decreases with the increase in SNR. It can be seen that the proposed algorithm performs better than the other two algorithms. The uncoded system exhibits higher BER compared to the coded system. It shows that channel coding can effectively improve system performance.

In the IN environment with α -stable distribution, the results of NMSE and BER performance for coded system are shown in Figs. 5 and 6, respectively, while the results for uncoded system are shown in Figs. 7 and 8, respectively. For coded system, Fig. 5



Fig. 5 Channel estimation performance comparison of various mitigation methods for coded 4-QAM system with α -stable distribution model



Fig. 6 BER performance comparison of various mitigation methods for coded 4-QAM system with α -stable distribution model

compares the channel estimation performance of three algorithms. It is obviously seen that the proposed algorithm performs better than SBL-LS and JCI-Kalman. Compared with SBL-LS algorithm, the SNR improvement of the proposed algorithm is more than 10 dB, compared with JCI-Kalman algorithm, the SNR improvement is more than 8 dB. Figure 6 provides the BER performance of three algorithms. The figure shows that the proposed algorithm outperforms the other two algorithms. When BER = 10^{-3} , the proposed algorithm achieves about 5 dB gain in SNR over JCI-Kalman and about 6 dB gain in SNR over SBL-LS. For uncoded system, Fig. 7 compares the channel estimation



Fig. 7 Channel estimation performance comparison of various mitigation methods for uncoded 4-QAM system with α -stable distribution model



Fig. 8 BER performance comparison of various mitigation methods for uncoded 4-QAM system with α -stable distribution model

performance of three algorithms. Similar to the coded system, the proposed algorithm performs better than the other algorithms. Compared with SBL-LS algorithm, the SNR improvement of the proposed algorithm is more than 12 dB, compared with JCI-Kalman algorithm, the SNR improvement is about 10 dB. Figure 8 provides the BER performance of three algorithms. Similar to the coded system, the proposed algorithm outperforms the other two algorithms. When BER = 10^{-3} , the proposed algorithm achieves about 5 dB gain in SNR over JCI-Kalman and about 9 dB gain in SNR over SBL-LS.



Fig. 9 Channel estimation performance variation of three algorithms with respect to the number of iterations when SNR at 5 dB



Fig. 10 Channel estimation performance variation of three algorithms with respect to the number of iterations when SNR at 20 dB

Figures 9 and 10 show the NMSE of three algorithms versus the number of iterations in uncoded system with SNRs of 5 dB and 20 dB, respectively. The figures show that both the FB-Kalman and JCI-Kalman algorithms exhibit a rapid convergence rate during the first 15 iterations, with a significant decrease in NMSE, and then the convergence rate gradually slows down and NMSE approaches a stable value. After 40 iterations, both algorithms have reached an almost stable state. While the SBL-LS algorithm converges fast, but the NMSE is the largest among three algorithms. The FB-Kalman achieves better and more reliable performance than the JCI-Kalman and SBL-LS because it can jointly estimate the channel, IN, and symbols using all subcarriers in the received OFDM symbol. Moreover, the FB-Kalman filter algorithm captures the time correlation of sparse time-varying channels, so it has better BER performance.

6 Conclusion

In this paper, we discuss the problems of joint sparse channel estimation, IN mitigation, and data detection in the presence of IN for OFDM systems. A joint channel estimation and IN suppression algorithm by using all subcarriers based on SBL and FB-Kalman is proposed. An efficient implementation algorithm based on EM is utilized to estimate the unknown parameters. Simulation results verify the efficiency of the proposed algorithm.

Although the proposed algorithm improves the estimation performance of OFDM systems in IN background, it comes at the cost of sacrificing algorithm complexity. Our future work includes fast implementation algorithms with robust performance.

Abbreviations

AMP	Approximate message passing
BER	Bit error rate
CS	Compressed sensing
CIR	Channel impulse response
EM	Expectation maximization
FB-Kalman	Forward–backward Kalman filter
GM	Gaussian mixture
IN	Impulse noise
LMMSE	Linear mean-squared estimator
NMSE	Normalized mean square error
OFDM	Orthogonal frequency division multiplexing
PLC	Power line communication
SBL	Sparse Bayesian learning
SNR	Signal-to-noise Ratio

Acknowledgements

This work was supported in part by the International Cooperation Project of the Ministry of Science and Technology under Grant 2018YFE0206500; in part by the National Natural Science Foundation of China under Grant 61571250; and in part by the Zhejiang Natural Science Foundation under Grant LY22F010018.

Author Contributions

All the authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials Not applicable.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Received: 27 March 2023 Accepted: 9 October 2023 Published online: 21 October 2023

References

- T. Bai et al., Discrete multi-tone digital subscriber loop performance in the face of impulsive noise. IEEE Access. 5, 10478–10495 (2017). https://doi.org/10.1109/ACCESS.2017.2712359
- H. Zhang, L. Yang, L. Hanzo, Compressed impairment sensing-assisted and interleaved-double-FFT-aided modulation improves broadband power line communications subjected to asynchronous impulsive noise. IEEE Access. 4, 81–96 (2016)
- K.L. Blackard, T.S. Rappaport, C.W. Bostian, Measurements and models of radio frequency impulsive noise for indoor wireless communications. IEEE J. Sel. Areas Commun. 11, 991–1001 (1993)
- 4. X. Kuai, H. Sun, S. Zhou, E. Cheng, Impulsive noise mitigation in underwater acoustic OFDM systems. IEEE Trans. Veh. Technol. 65, 8190–8202 (2016)
- D. Middleton, Non-Gaussian noise models in signal processing for telecommunications: new methods an results for class A and class B noise models. IEEE Trans. Inf. Theory 45(4), 1129–1149 (1999). https://doi.org/10.1109/18.761256
- M. Ghosh, Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems. IEEE Trans. Commun. 44(2), 145–147 (1996). https://doi.org/10.1109/26.486604
- S.P. Talebi, S.J. Godsill, D.P. Mandic, Filtering structures for α-stable systems. IEEE Control Syst. Lett. 7, 553–558 (2023). https://doi.org/10.1109/LCSYS.2022.3202827
- M. Ghosh, Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems. IEEE Trans. Commun. 44, 145–147 (1996)
- K. Yan, H. Zhang, H.C. Wu, Robust multipath channel estimation in the presence of impulsive noise. IET Commun. 12, 228–235 (2018)
- S. Dimitrov, S. Sinanovic, H. Haas, Clipping noise in OFDM-based optical wireless communication systems. IEEE Trans. Commun. 60, 1072–1081 (2012)
- 11. Mandar Chitre, S. Kuselan, V. Pallayil, Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation. J. Acoust. Soc. Am. **132**, 838–847 (2012)
- 12. D.L. Donoho, Compressed sensing. IEEE Trans. Inf. Theory 52, 1289–1306 (2006)
- J. Lin, M. Nassar, B.L. Evans, Impulsive noise mitigation in powerline communications using sparse Bayesian learning. IEEE J. Sel. Areas Commun. 31, 1172–1183 (2013)
- P. Chen, Y. Rong, S. Nordholm, A.J. Duncan, Z. He, Compressed sensing based channel estimation and impulsive noise cancelation in underwater acoustic OFDM systems, in 2016 IEEE Region 10 Conference (TENCON) (2016), p. 2539–2542
- T. Suzuki, H.M. Tran, T. Wada, An underwater acoustic OFDM communication system with shrimp (impulsive) noise cancelling, in 2014 International Conference on Computing, Management and Telecommunications (ComManTel) (2014), p.152–156
- M. Nassar, P. Schniter, B.L. Evans, A factor graph approach to joint OFDM channel estimation and decoding in impulsive noise environments. IEEE Trans. Signal Process. 62, 1576–1589 (2014)
- S. Wu, S. Wang, Z. He, K. Niu, Y. Rong, An Approximate Message Passing Algorithm for Channel and Impulsive Noise Estimation in Underwater Acoustic OFDM Systems. OCEANS 2019 - Marseille. 1–5 (2019). https://doi.org/10.1109/ OCEANSE
- X. Lv, Y. Li, Y. Wu, H. Liang, Kalman filter based recursive estimation of slowly fading sparse channel in impulsive noise environment for OFDM systems. IEEE Trans. Veh. Technol. 69, 2828–2835 (2020)
- S. Wang, Z. He, K. Niu, P. Chen, Y. Rong, New results on joint channel and impulsive noise estimation and tracking in underwater acoustic OFDM systems. IEEE Trans. Wireless Commun. 19, 2601–2612 (2020)
- 20. P. Schniter, Belief-propagation-based joint channel estimation and decoding for spectrally efficient communication over unknown sparse channels. Phys. Commun. 5, 91101 (2012)
- 21. N. Okafor, M. Deyoung, UT Austin interference modeling and mitigation toolbox. http://users.ece.utexas.edu/ bevans/projects/rfi/software/. Accessed 15 October 2008

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com