Wireless Pers Commun (2015) 82:1161–1178
DOI 10.1007/s11277-015-2273-x

Joint Channel Estimation and Symbol Detection for OFDM Systems in Rapidly Time-Varying Sparse Multipath Channels

Habib Şenol

Published online: 14 January 2015

Abstract In this paper, we propose a space-alternating generalized expectation maximization (SAGE) based joint channel estimation and data detection algorithm in compressive sensing (CS) framework for orthogonal frequency-division multiplexing (OFDM) systems in rapidly time-varying sparse multipath channels. Using dynamic parametric channel model, the sparse multipath channel is parameterized by a small number of distinct paths, each represented by the path delays and path gains. In our model, we assume that the path gains rapidly vary within the OFDM symbol duration while the number of paths and path delays vary symbol by symbol. Since the convergency of the SAGE algorithm needs statistically independent parameter set of interest to be estimated, we specifically choose the discrete orthonormal Karhunen–Loeve basis expansion model (DKL-BEM) to provide statistically independent BEM coefficients within one OFDM symbol duration and use just a few significant BEM coefficients to represent the rapidly time-varying path gains. The resulting SAGE algorithm that also incorporates inter-channel interference cancellation updates the data sequences and the channel parameters serially. The computer simulations show that our proposed algorithm has better channel estimation and symbol error rate performance than that of the orthogonal matching pursuit algorithm that is commonly proposed in the CS literature.

Keywords Sparse multipath channel · OFDM · SAGE · Matching pursuit · Basis expansion

1 Introduction

Orthogonal frequency-division multiplexing (OFDM) has been shown to be an effective method to overcome inter-symbol interference (ISI) caused by frequency-selective fading with a simple transceiver structure. Due to its high data rate transmission capability, robustness against frequency selective fading channels and flexible spectrum allocation for different services, OFDM has been widely used in the current and future wireless communication

H. Şenol (✉)
Department of Computer Engineering, Kadir Has University, Fatih, 34083 Istanbul, Turkey
e-mail: hsenol@khas.edu.tr

systems, such as the digital audio/video broadcasting (DAB/DVB) systems, the asymmetric digital subscriber lines (ADSLs), the wireless local area networks (WLANs), the Mobile Worldwide Interoperability Microwave Systems for Next-Generation Wireless Communication Systems (WiMAX), and the Third-Generation Partnership Project (3GPP) long-term evolution (LTE) systems. Channel estimation task is required for coherent detection in OFDM systems and commonly achieved using pilot symbol transmissions. Channel estimation methods can be categorized basically into parametric/non-parametric channel model based estimation methods. There exists numerous pilot-aided channel estimation methods in the literature [1–5]. Most of them are nonparametric methods. Since non-parametric methods do not make any assumptions about the channel model, the dimension of the estimation problem can be quite large. However, in parametric channel estimation methods, the wireless radio channel is modeled with a few significant paths resulting usually in a sparse multipath channel model [5–8]. Consequently, parametric methods enable to reduce the dimension of the estimation task and the amount of pilot symbols needed in channel estimation, and therefore, as compared to the non-parametric channel model, the parametric channel model based estimator can achieve better performance [5]. Parametric channel modeling based channel estimation methods proposed in [5,9,10] use minimum description length (MDL) to detect the number of paths and then apply subspace methods such as the estimation of signal parameters using rotational invariance techniques (ESPRIT) and the multiple signal classification (MUSIC) to estimate the channel path delays. However, in time-varying channel scenarios the propagation delays, the number of delays and the tap coefficients vary over the time, and the static parametric channel model does not represent such a dynamic channel environment. A more realistic multipath channel model that allows the path number and the path delays to vary over the time is presented in [9]. In this work, we consider such a dynamic parametric channel model. Because of intensive computational complexity, the conventional subspace methods (e.g. ESPRIT, MUSIC) are no longer practical in time-varying channel estimation problem since the channel parameters vary over the time and their estimates need to be updated frequently.

More recently, compressive sensing (CS) techniques have been applied to sparse channel estimation [10–17]. The CS-based channel estimators in [11–16] assume that the channel is sparse in the equivalent discrete-time baseband representation. However, in practice, the sparsity assumption does not always hold due to the non-integer normalized path delays in the equivalent discrete-time baseband representation of the channel. Therefore, such an estimated channel may differ substantially from the original channel. The over-complete dictionaries with finer delay resolution are used for better modeling of sparse multipath channels while employing the CS-based channel estimators in [10,17,18]. In these works, as popular compressed sensing techniques, the matching pursuit (MP) and the orthogonal matching pursuit (OMP) algorithms were employed to deal with the time-varying sparse multipath channels for OFDM systems whereas [10] and [17,18] consider underwater acoustic and wireless channel environments, respectively. Focusing on works that consider wireless channel scenario, in [17], the multipath coefficients were assumed to be not changed during an OFDM symbol duration but they vary from symbol to symbol, and an OMP-like algorithm was proposed to estimate the path delays using adaptive delay grids to obtain a lower computational complexity. After estimation of the path delays, path gains were estimated using the polynomial basis expansion model (BEM). The proposed algorithm in [17] has the same performance with that of OMP algorithm. In [18], OMP algorithm based channel estimator is employed using so-called super-resolution dictionary matrix. First, ignoring ICI terms, the frequency-domain channel coefficients at pilot locations were estimated by least-square (LS) estimator and then unlike the classical OMP algorithm, the estimates of the frequency-domain channel

coefficients are used instead of the observations in the OMP algorithm while estimating the channel parameters. As a result, because of being ICI-ignorant, the channel estimator in [18] has limited performance.

In this paper, we propose a space-alternating generalized expectation maximization (SAGE) based joint channel estimation and data detection algorithm within the CS framework for OFDM systems in rapidly time-varying sparse multipath channels. Recently in our works, [19] and [20], the SAGE based channel estimation and data detection algorithms have been presented for OFDM systems operating over high mobility channels. However in [19] and [20], time-varying *non-sparse* multipath channels with known number of paths were considered. In this work, unlike [19] and [20], we consider a much more challenging time-varying multipath channel model; the multipath delay positions are sparse, unknown and randomly varying at non-integer multiples of the sampling duration, moreover, the number of the paths is unknown and randomly varying as well. Hence, the number of paths and the multipath delay positions are also the parameters of interest to be estimated. Consequently, derivation and structure of the proposed SAGE algorithm in this work are completely different from the SAGE algorithm in our earlier works [19,20]. In this paper, we model the random path delays within the guard interval duration of an OFDM symbol with the delay grid spaced at baseband sampling rate, and we achieve a substantial performance with a reasonable finer delay grid resolution. Since the SAGE algorithm needs independent parameter set, we specifically choose the discrete orthonormal Karhunen–Loeve basis expansion model (DKL-BEM) among other BEMs to provide statistically independent BEM coefficients within the one OFDM symbol duration. After employing DKL-BEM to represent the rapidly time-varying path gains, we estimate the DKL-BEM coefficients by applying our proposed SAGE algorithm that provides substantially much better estimation performance than that of OMP algorithm proposed in [10,17,18].

The remainder of this paper is organized as follows. Section 2 presents the system model including the observation and the time-varying sparse multipath channel models, and consequently, the basis expansion model of the time-varying channel. In Section 3, the proposed SAGE algorithm is presented for joint channel estimation and symbol detection. Furthermore, the initialization, the summary and the complexity analysis of the proposed SAGE algorithm are presented. Section 4 provides the performance results while Section 5 contains concluding remarks.

2 System Model

2.1 Signal and Channel Model

We consider a zero padded OFDM system with N subcarriers employing actively K subcarriers to transmit data symbols, and nothing is transmitted from the remaining N–K carriers for the purpose of zero-padding. During an OFDM symbol, each active subcarrier is modulated by a data symbol $b[k]$, where k represents the subcarrier index. After taking a K-point inverse fast Fourier transform (IFFT) of the data sequence and adding a cyclic prefix (CP) of duration T_{cp} before transmission to ISI, the transmitted continuous time-domain complex valued signal can be expressed as

$$s(t) = \frac{1}{\sqrt{N}} \sum_{k=-K/2}^{K/2-1} b[k]\, e^{j\,2\pi k\,\Delta f\,(t-T_{cp})}\, \zeta(t), \tag{1}$$

where $\Delta f = 1/T$ is the OFDM subcarrier spacing, T stands for OFDM symbol duration, $\zeta(t)$ denotes the unit pulse given by

$$\zeta(t) = \begin{cases} 1, & 0 \le t \le T_{sym} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

and $T_{sym} = T + T_{cp}$ is the duration of an entire OFDM symbol.

The signal $s(t)$ is transmitted over a wireless multipath channel with time-varying impulse response given by

$$g(t, \tau) = \sum_{\ell=1}^{L} h_\ell(t)\, \delta(\tau - \tau_\ell), \tag{3}$$

where L is the number of channel paths, $h_\ell(t)$ is the time-varying gain and τ_ℓ is the delay of the ℓth path. Path gains, $\{h_\ell(t)\}_{\ell=1}^{L}$, are wide-sense stationary uncorrelated scattering (WSSUS) complex Gaussian process with the Jakes' power spectrum with the following autocorrelation function

$$E\{h_\ell(t) h_{\ell'}^*(t')\} = \Omega_\ell J_0\big(2\pi f_{dopp}(t - t')\big)\delta(\ell - \ell'), \tag{4}$$

where $(\cdot)^*$ denotes the complex conjugate operator, $\{\Omega_\ell\}_{\ell=1}^{L}$, represent the normalized powers of the channel paths satisfying $\sum_{\ell=1}^{L} \Omega_\ell = 1$. $J_0(\cdot)$ is the zero-th order Bessel function of the first kind, f_{dopp} is the maximum Doppler shift due to the vehicle motion and $\delta(\cdot)$ is the Kronecker delta function. So, the time domain received signal can be obtained as

$$\begin{aligned} y(t) &= g(t, \tau) \star s(t) + w(t) \\ &= \sum_{\ell=1}^{L} h_\ell(t)\, s(t - \tau_\ell) + w(t) \\ &= \frac{1}{\sqrt{N}} \sum_{\ell=1}^{L} \sum_{k=-K/2}^{K/2-1} h_\ell(t)\, b[k]\, e^{j\frac{2\pi}{T}k(t-\tau_\ell-T_{cp})}\, \zeta(t - \tau_\ell) + w(t), \end{aligned} \tag{5}$$

where $\star$ shows the convolution operator and $w(t)$ is zero-mean complex additive white Gaussian noise (AWGN).

At the receiver, $y(t)$ is converted into the discrete-time signal by means of low-pass filtering and A/D conversion with the sampling interval $T_s = T/N$. Assuming that K active subcarriers are within the region of frequency response of both transmitter and receiver filters, and the number of channel paths and the path delays do not change during an OFDM symbol duration, it is sufficient to consider the channel estimation only symbol by symbol. Therefore, the nth time sample within any OFDM symbol after the CP removal can be expressed as

$$\begin{aligned} y[n] &= y(T_{cp} + nT_s) \\ &= \frac{1}{\sqrt{N}} \sum_{\ell=1}^{L} \sum_{k=-K/2}^{K/2-1} h_\ell[n]\, b[k]\, e^{j\frac{2\pi}{N}k(n-\check{\tau}_\ell)} + w[n],\ n = 0, 1, \ldots, N-1, \end{aligned} \tag{6}$$

where $\check{\tau}_\ell = \tau_\ell / T_s$ is the normalized delay of the ℓth path and $w[n] = w(T_{cp} + nT_s)$ denotes the AWGN sample at time n with $w[n] \sim \mathcal{CN}(0, \mathcal{N}_0)$. It is straightforward that the vector form of (6) can be expressed as

$$\mathbf{y} = \sum_{\ell=1}^{L} \mathrm{diag}\big[\mathbf{F}^{\dagger}\big(\boldsymbol{\nu}(\check{\tau}_\ell) \odot \mathbf{b}\big)\big]\,\mathbf{h}_\ell + \mathbf{w}, \tag{7}$$

where

$$\begin{aligned}
\mathbf{y} &= \big[y[0], y[1], \ldots, y[N-1]\big]^T \in \mathcal{C}^N,\\
\mathbf{b} &= \Big[b\big[-\frac{K}{2}\big], b\big[-\frac{K}{2}+1\big], \ldots, b\big[\frac{K}{2}-1\big]\Big]^T \in \mathcal{C}^K,\\
\mathbf{h}_\ell &= \big[h_\ell[0], h_\ell[1], \ldots, h_\ell[N-1]\big]^T \in \mathcal{C}^N,\\
\mathbf{w} &= \big[w[0], w[1], \ldots, w[N-1]\big]^T \in \mathcal{C}^N,
\end{aligned} \tag{8}$$

$\boldsymbol{\nu}(\check{\tau}_\ell) \in \mathcal{C}^K$ is a column vector with entries $e^{-j\frac{2\pi}{N}k\check{\tau}_\ell}$, $(\cdot)^T$ denotes the transpose operator and $\odot$ stands for the Hadamard product. $\mathbf{F}$ is the FFT matrix having orthonormal discrete basis function $\frac{1}{\sqrt{N}}\,e^{-j\frac{2\pi}{N}nk}$.

2.2 Basis Expansion Model of the Channel

In time-varying environments, the number of the unknown path gains, $\{\mathbf{h}_\ell\}_{\ell=1}^{L}$ in (8), within one OFDM symbol is NL. However, the number of the observations in vector $\mathbf{y}$ is N. So, the number of unknown parameters is much larger than the number of the observations, which makes the estimation of the path gains difficult. To reduce the number of the parameters to be estimated, the BEM is proposed to model the time variations of multipath channel gains [20–25]. Employing the BEM, our path gain estimation problem turns into the estimation problem of the BEM coefficients. These BEM methods are the DKL-BEM, the discrete prolate spherical BEM (DPS-BEM), the complex exponential BEM (CE-BEM), the polynomial BEM (P-BEM), and discrete Legendre polynomial BEM (DLP-BEM). In this work, we choose DKL–BEM, since the SAGE algorithm needs statistically independent parameter set, and the BEM methods mentioned above, the DKL-BEM provides statistically independent BEM coefficients within the one OFDM symbol duration. Applying DKL-BEM, the time variations of the channel within one OFDM symbol duration are well approximated by a linear combination of the orthonormal basis functions as follows:

$$\tilde{h}_\ell[n] = \sum_{d=1}^{D} c_{\ell,d}\, u_d[n],\ \ n = 0, 1, \ldots, N-1, \tag{9}$$

where $c_{\ell,d}$ and $u_d(n)$ are the DKL-BEM coefficients and DKL-BEM orthonormal basis functions, respectively. The vector form of (9) can be obtained easily as follows

$$\tilde{\mathbf{h}}_\ell = \sum_{d=1}^{D} \mathbf{u}_d\, c_{\ell,d} \quad \text{and} \quad c_{\ell,d} = \mathbf{u}_d^{\dagger}\mathbf{h}_\ell, \tag{10}$$

where $\mathbf{u}_d = \big[u_d[0], u_d[1], u_d[N-1]\big]^T$ is the dth DKL-BEM orthonormal base vector and $(\cdot)^{\dagger}$ denotes the conjugate transpose operator. Substituting (10) into (7), we have

$$\mathbf{y} = \sum_{\ell=1}^{L}\sum_{d=1}^{D} \mathbf{a}_d(\check{\tau}_\ell)\, c_{\ell,d} + \mathbf{w}, \tag{11}$$

where

$$\mathbf{a}_d(\check{\tau}_\ell) = \mathbf{u}_d \odot \Big[\mathbf{F}^{\dagger}\big[\boldsymbol{\nu}(\check{\tau}_\ell) \odot \mathbf{b}\big]\Big]. \tag{12}$$

In this work, we are mainly interested in estimation of time-varying sparse multipath channels based on the observation (11). The overall continuous-time channel impulse response is represented by a parametric model in which the ℓth time-varying distinct path is characterized by path delay, $\breve{\tau}_\ell$, and a few significant DKL-BEM coefficients, $\{c_{\ell,d}\}_{d=1}^{D}$. In practice, the sparsity assumption does not always hold due to the non-integer normalized path delays in the equivalent discrete-time baseband representation of the channel. Therefore, such an estimated channel may differ substantially from the original channel. To achieve a better channel estimation performance, the A/D conversion at the input of the OFDM receiver is implemented with a sampling period T_s/ρ, $\rho \in \{1, 2, \ldots\}$ leading to a finer delay resolution. Consequently, the continuous-valued normalized path delays $\breve{\tau}_\ell$, $\ell \in \{1, 2, \ldots, L\}$ can be discretized as $\eta_\ell = \lfloor \frac{\tau_\ell}{T_s/\rho} \rfloor = \lfloor \rho\, \breve{\tau}_\ell \rfloor$ and take values from the set of possible discrete path delays

$$\eta_\ell \in \{0, 1, \ldots, \rho L_{cp} - 1\}, \tag{13}$$

where $L_{cp} = \lfloor T_{cp}/T_s \rfloor$, and $\lfloor \cdot \rfloor$ denotes the floor operator. Based on the associated discrete random channel tap positions $\{\eta_\ell\}_{\ell=1}^{L}$, the received signal in (11) can be rewritten as

$$\mathbf{y} = \sum_{\ell=1}^{L} \sum_{d=1}^{D} \mathbf{a}_{\eta_\ell,\, d}\, c_{\ell,\, d} + \mathbf{w}, \tag{14}$$

where, following (12), the vector $\mathbf{a}_{\eta_\ell,d} \in \mathcal{C}^N$ is defined as

$$\mathbf{a}_{\eta_\ell,d} = \mathbf{a}_d(\breve{\tau}_\ell)\Big|_{\breve{\tau}_\ell = \frac{\eta_\ell}{\rho}} = \mathbf{u}_d \odot \Big[\mathbf{F}^{\dagger}\big[\boldsymbol{\nu}_{\eta_\ell} \odot \mathbf{b}\big]\Big], \tag{15}$$

and $\boldsymbol{\nu}_{\eta_\ell} = \boldsymbol{\nu}(\breve{\tau}_\ell)\Big|_{\breve{\tau}_\ell = \frac{\eta_\ell}{\rho}} \in \mathcal{C}^K$ is the column vector with entries $e^{-j\frac{2\pi}{\rho N} k \eta_\ell}$. Eventually for $d \in \{1, 2, \ldots, D\}$ and $\eta_\ell \in \{0, 1, \ldots, \rho L_{cp} - 1\}$, and defining the mapping $(\eta_\ell, d) \mapsto r_m$ so as to be

$$r_m = \eta_\ell\, D + d, \tag{16}$$

the received signal in (14) can be rewritten as

$$\mathbf{y} = \sum_{m=1}^{M} \mathbf{a}_{r_m} c_m + \mathbf{w} = \mathbf{A}\mathbf{c} + \mathbf{w}, \tag{17}$$

where $\mathbf{a}_{r_m} \in \mathcal{C}^N$ is the r_mth column vector of the so-called *dictionary matrix* $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{\rho D L_{cp}}] \in \mathcal{C}^{N \times \rho D L_{cp}}$, corresponding to the η_ℓth discrete multipath channel tap delay and dth Karhunen Loeve base, respectively. Vector $\mathbf{c}$ is the sparse BEM coefficient vector with non-zero elements $\{c_m\}_{m=1}^{M}$, where $M = LD$. Reversely, using (16), for a given random column index $r_m \in \{1, 2, \ldots, \rho D L_{cp}\}$, the corresponding discrete tap delay and basis function indices are obtained as follows

$$\begin{aligned} \eta_\ell &= \left\lfloor \frac{r_m - 1}{D} \right\rfloor, \\ d &= (r_m - 1) \bmod (D) + 1. \end{aligned} \tag{18}$$

The estimation problem of non-zero elements of the sparse coefficient vector $\mathbf{c}$ in (17) can be solved by a sparse signal recovery problem. The MP algorithm and its variants are very popular sparse recovery methods and very commonly used for such type of estimation problems. In this paper, initial estimates of the channel gains and channel delays are performed

by the MP algorithm, and they are updated together with data symbols within the proposed SAGE algorithm iterations to improve their estimation performance.

3 Sparse Channel Estimation and Data Detection with SAGE Algorithm

We now derive the SAGE algorithm regarding the signal model given by (17). The SAGE algorithm, proposed by Fessler et al. [26], is a two fold generalization of the so-called *expectation maximization* (EM) algorithm that provides updated estimates for an unknown parameter set $\boldsymbol{\theta}$. First, rather than updating all parameters simultaneously at ith iteration, only a subset of $\boldsymbol{\theta}_{\mathcal{S}}$ indexed by $\mathcal{S} = \mathcal{S}[i]$ is updated while keeping the parameters in the complement set $\boldsymbol{\theta}_{\bar{\mathcal{S}}}$ fixed; and second, the concept of the complete data $\boldsymbol{\chi}$ is extended to that of the so-called *admissible hidden data* $\boldsymbol{\chi}_{\mathcal{S}}$ to which the observed signal $\mathbf{y}$ is related by means of a possibly nondeterministic mapping $\boldsymbol{\chi}_{\mathcal{S}} \mapsto \mathbf{y}(\boldsymbol{\chi}_{\mathcal{S}})$. The convergence rate of the SAGE algorithm is usually higher than that of the EM algorithm, because the conditional Fisher information matrix given for each set of parameters is likely smaller than that of the complete data, given for the entire space. At the ith iteration, the expectation-step (*E-step*) of the SAGE algorithm is defined as

$$Q_{\mathcal{S}}(\boldsymbol{\theta}_{\mathcal{S}}|\boldsymbol{\theta}^{[i]}) = E\left\{\log p\left(\boldsymbol{\chi}_{\mathcal{S}}|\boldsymbol{\theta}_{\mathcal{S}}, \boldsymbol{\theta}_{\bar{\mathcal{S}}}^{[i]}\right)\middle|\mathbf{y}, \boldsymbol{\theta}^{[i]}\right\}. \tag{19}$$

In the maximization step (*M-Step*), only $\boldsymbol{\theta}_{\mathcal{S}}$ is updated, i.e.,

$$\boldsymbol{\theta}_{\mathcal{S}}^{[i+1]} = \arg\max_{\boldsymbol{\theta}_{\mathcal{S}}} Q_{\mathcal{S}}(\boldsymbol{\theta}_{\mathcal{S}}|\boldsymbol{\theta}^{[i]}) \tag{20}$$

We now give the details of the SAGE algorithm as follows:

The unknown parameter set to be estimated in our problem is

$$\boldsymbol{\theta} = \{\mathbf{r}, \mathbf{c}, \mathbf{b}^{\mathcal{D}}\}, \tag{21}$$

where $\mathbf{r} = [\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_M}]^{\mathbf{T}}$ is the DKL-BEM coefficient index vector having entries as defined in (16) , $\mathbf{c} = [c_1, c_2, \ldots, c_M]^T$ is the DKL-BEM coefficient vector, $\mathbf{b}^{\mathcal{D}} = \left[b[k_1], b[k_2], \ldots b[k_{K_{\mathcal{D}}}]\right]^T$ is the non-pilot OFDM data vector where $k_q \in \left\{-\frac{K}{2}, (-\frac{K}{2}+1), \ldots, (\frac{K}{2}-1)\right\}$ is the qth data location in discrete frequency domain, and $K_{\mathcal{D}}$ represents the number of data symbols in one OFDM symbol. To obtain a receiver architecture that iterates between soft-data and channel estimation in the SAGE algorithm, we decompose the unknown parameter set $\boldsymbol{\theta}$ into $M + K_{\mathcal{D}}$ subsets, representing the parameters, $\mathbf{r}$, $\mathbf{c}$, and $\mathbf{b}^{\mathcal{D}}$, as follows

- The first M subsets of $\boldsymbol{\theta}$ are chosen as $\boldsymbol{\theta}_m = \{r_m, c_m\}$, $m = 1, 2, \ldots, M$. For each subset we define $\bar{\boldsymbol{\theta}}_m = \boldsymbol{\theta} \setminus \boldsymbol{\theta}_m = \{\bar{\mathbf{r}}_m, \bar{\mathbf{c}}_m, \mathbf{b}^{\mathcal{D}}\}$, where $\bar{\mathbf{r}}_m = \mathbf{r} \setminus \mathbf{r_m}$, $\bar{\mathbf{c}}_m = \mathbf{c} \setminus c_m$, and $\setminus$ denotes the set exclusion operator.
- The $(M+q)$th subset of $\boldsymbol{\theta}$ is chosen as by $\theta_{M+q} = b[k_q]$, $q = 1, 2, \ldots, K_{\mathcal{D}}$, and $\bar{\boldsymbol{\theta}}_{M+q} = \boldsymbol{\theta} \setminus \theta_{M+q} = \{\mathbf{r}, \mathbf{c}, \bar{\mathbf{b}}_q^{\mathcal{D}}\}$, where $\bar{\mathbf{b}}_q^{\mathcal{D}} = \mathbf{b}^{\mathcal{D}} \setminus b[k_q]$.

At the ith iteration of the SAGE algorithm, only the parameters in one set are updated, whereas the other parameters are kept fixed, and this process is repeated until all parameters are updated. According to the above parameter subset definitions, each iteration of the SAGE algorithm for our problem has two steps:

1. $\boldsymbol{\theta}_m = \{r_m, c_m\}, m = 1, 2, \ldots, M$ is updated with SAGE algorithm while $\bar{\boldsymbol{\theta}}_m = \{\bar{\mathbf{r}}_m, \bar{\mathbf{c}}_m, \mathbf{b}^{\mathcal{D}}\}$ is fixed.

2. $\theta_{M+q} = b[k_q]\,, q = 1, 2, \ldots, K_{\mathcal{D}}$ is updated with SAGE algorithm while $\bar{\boldsymbol{\theta}}_{M+q} = \{\mathbf{r}, \mathbf{c}, \bar{\mathbf{b}}_q^{\mathcal{D}}\}$ is fixed.

We now derive the SAGE algorithm below by also specifying the corresponding admissible hidden data sets.

3.1 Estimation of Channel Parameters, $\boldsymbol{\theta}_m = \{r_m, c_m\}, m = 1, 2, \ldots, M$

A suitable approach for applying the SAGE algorithm for estimation of $\boldsymbol{\theta}_m = \{r_m, c_m\}$ is to decompose the receive vector in (17) into the sum

$$\mathbf{y} = \mathbf{x}^{(m)} + \bar{\mathbf{x}}^{(m)}, \tag{22}$$

where

$$\mathbf{x}^{(m)} = \mathbf{a}_{r_m} c_m + \mathbf{w} \quad \text{and} \quad \bar{\mathbf{x}}^{(m)} = \sum_{m'=1,\, m' \neq m}^{M} \mathbf{a}_{r_{m'}} c_{m'}. \tag{23}$$

We define the admissible hidden data as $\boldsymbol{\chi}_m = \{\mathbf{x}^{(m)}\}$. To perform the *E-Step* of the SAGE algorithm, the conditional expectation is taken over $\boldsymbol{\chi}_m$ given the observation $\mathbf{y}$ and given that $\boldsymbol{\theta}$ equals its estimate calculated at ith iteration

$$\begin{aligned} Q_m(\boldsymbol{\theta}_m|\boldsymbol{\theta}^{[i]}) &= E\big\{\log p(\boldsymbol{\chi}_m|\boldsymbol{\theta}_m, \bar{\boldsymbol{\theta}}_m^{[i]})\big|\mathbf{y}, \boldsymbol{\theta}^{[i]}\big\} \\ &= E\big\{\log p(\mathbf{x}^{(m)}|r_m, c_m, \bar{\mathbf{r}}_m^{[i]}, \bar{\mathbf{c}}_m^{[i]}, \mathbf{b}^{\mathcal{D}[i]})\big|\mathbf{y}, \mathbf{r}^{[\mathbf{i}]}, \mathbf{c}^{[\mathbf{i}]}, \mathbf{b}^{\mathcal{D}[\mathbf{i}]}\big\}, \end{aligned} \tag{24}$$

where

$$\log p(\mathbf{x}^{(m)}|r_m, c_m, \bar{\mathbf{r}}_m^{[i]}, \bar{\mathbf{c}}_m^{[i]}, \mathbf{b}^{\mathcal{D}[i]}) \sim -\frac{1}{\mathcal{N}_0} (\mathbf{x}^{(m)} - \mathbf{a}_{r_m}^{[i]} c_m)^{\dagger} (\mathbf{x}^{(m)} - \mathbf{a}_{r_m}^{[i]} c_m), \tag{25}$$

and, following (16), $\mathbf{a}_{r_m}^{[i]}$ is calculated from (15) for given symbol vector $\mathbf{b}^{[i]}$ and the DKL-BEM coefficient index r_m. Inserting (25) into (24) we obtain

$$Q_m(\boldsymbol{\theta}_m|\boldsymbol{\theta}^{[i]}) = \frac{1}{\mathcal{N}_0}\big(2\,\mathfrak{Re}\{c_m^* \mathbf{a}_{r_m}^{[i]^{\dagger}} \widehat{\mathbf{x}^{(m)}}^{[i]}\} - \big\|\mathbf{a}_{r_m}^{[i]}\big\|^2 |c_m|^2\big), \tag{26}$$

where $\mathfrak{R}\{\cdot\}$ denotes the real part, $\|\cdot\|$ stands for the ℓ_2-norm of a vector and $\widehat{\mathbf{x}^{(m)}}^{[i]}$ is defined as

$$\widehat{\mathbf{x}^{(m)}}^{[i]} = E\big\{\mathbf{x}^{(m)}|\mathbf{y}, \mathbf{r}^{[\mathbf{i}]}, \mathbf{c}^{[\mathbf{i}]}, \mathbf{b}^{\mathcal{D}[\mathbf{i}]}\big\} = \mathbf{y} - \sum_{\mathbf{m}'=\mathbf{1},\, \mathbf{m}' \neq \mathbf{m}}^{\mathbf{M}} \mathbf{a}_{\mathbf{r}_{\mathbf{m}'}^{[\mathbf{i}]}}^{[\mathbf{i}]} \mathbf{c}_{\mathbf{m}'}^{[\mathbf{i}]} \tag{27}$$

In the *M-Step* of the SAGE algorithm, the estimates of $\boldsymbol{\theta}_m = \{r_m, c_m\}$ are updated sequentially for $m = 1, 2, \ldots, M$ at the $(i+1)$st iteration according to following maximization

$$\boldsymbol{\theta}_m^{[i+1]} = \arg\max_{\boldsymbol{\theta}_m} Q_m(\boldsymbol{\theta}_m|\boldsymbol{\theta}^{[i]}), \tag{28}$$

where $Q_m(\boldsymbol{\theta}_m|\boldsymbol{\theta}^{(i)})$ is given by (26). So, taking the derivative of $Q_m(\boldsymbol{\theta}_m|\boldsymbol{\theta}^{[i]})$ with respect to c_m^* and equating to zero, we find the final SAGE estimates of the mth parameter set, $\boldsymbol{\theta}_m = \{r_m, c_m\}$, at $(i+1)$st iteration as follows:

$$r_m^{[i+1]} = \arg\max_r \frac{\left|\mathbf{a}_r^{[i]\dagger}\widehat{\mathbf{x}^{(m)}}^{[i]}\right|^2}{\left\|\mathbf{a}_r^{[i]}\right\|^2}, \quad r \in \{1, 2, \ldots, \rho D L_{cp}\},\ r \notin \{r_1, r_2, \ldots, r_{m-1}\}$$

$$c_m^{[i+1]} = \frac{\mathbf{a}_{r_m^{[i+1]}}^{[i]\ \dagger}\ \widehat{\mathbf{x}^{(m)}}^{[i]}}{\left\|\mathbf{a}_{r_m^{[i+1]}}^{[i]}\right\|^2}. \tag{29}$$

3.2 Detection of Data Symbols, $\theta_{M+q} = b[k_q],\ q = 1, 2, \ldots, K_{\mathcal{D}}$

In order to obtain the SAGE algorithm for detection of data symbols, using (14), (15), and (16), we obtain the following alternative form of the observation equation in (17)

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{b} + \mathbf{w}, \tag{30}$$

where

$$\boldsymbol{\Phi} = \sqrt{N}\left(\mathbf{F}^{\dagger} \odot \mathbf{H}\right) \in \mathcal{C}^{N\times K}, \tag{31}$$

and $\mathbf{H} \in \mathcal{C}^{N\times K}$ is the channel matrix with entries $H[n, k]$ representing the frequency response of the channel at discrete frequency k and time n. So, the channel matrix that is well approximated by DKL-BEM is obtained as

$$\tilde{\mathbf{H}} = \frac{1}{\sqrt{N}} \sum_{m=1}^{M} \boldsymbol{\Psi}_{r_m} c_m, \tag{32}$$

where recalling the demapping $r_m \mapsto (\eta_\ell, d)$ in (18), $\boldsymbol{\Psi}_{r_m}$ matrix is defined as $\boldsymbol{\Psi}_{r_m} = \mathbf{u}_d\, \boldsymbol{v}_{\eta_\ell}^T \in \mathcal{C}^{N\times K}$. In order to derive the SAGE algorithm for estimation of $\theta_{M+q} = b[k_q]$, the receive vector in (30) is decomposed into the sum

$$\mathbf{y} = \mathbf{z}^{(q)} + \bar{\mathbf{z}}^{(q)}, \tag{33}$$

where

$$\mathbf{z}^{(q)} = \boldsymbol{\psi}[k_q] b[k_q] + \mathbf{w} \quad , \quad \bar{\mathbf{z}}^{(q)} = \sum_{k=-K/2,\, k\neq k_q}^{K/2-1} \boldsymbol{\psi}[k] b[k], \tag{34}$$

and $\boldsymbol{\psi}[k]$ denotes the column vector of the matrix $\boldsymbol{\Phi}$ in (30) at discrete frequency k such that $\boldsymbol{\Phi} = \left[\boldsymbol{\psi}[-\frac{K}{2}], \boldsymbol{\psi}[-\frac{K}{2}+1], \ldots, \boldsymbol{\psi}[\frac{K}{2}-1]\right]$. We define the admissible hidden data $\boldsymbol{\chi}_{M+q} = \{\mathbf{z}_q\}$ to detect the qth data symbol $b[k_q]$. Now, let us derive the SAGE algorithm. To perform the *E-Step* of the SAGE algorithm, the conditional expectation is taken over $\boldsymbol{\chi}_{M+q}$ given the observation $\mathbf{y}$ and given that $\boldsymbol{\theta}$ equals its estimate calculated at ith iteration

$$\begin{aligned} Q_{M+q}(\theta_{M+q}|\boldsymbol{\theta}^{[i]}) &= E\left\{\log p(\boldsymbol{\chi}_{M+q}|\theta_{M+q}, \bar{\boldsymbol{\theta}}_{M+q}^{[i]})\big|\mathbf{y}, \boldsymbol{\theta}^{[i]}\right\} \\ &= E\left\{\log p(\mathbf{z}^{(q)}|b[k_q], \bar{\mathbf{b}}_q^{\mathcal{D}}[i], \mathbf{r}^{[\mathbf{i}]}, \mathbf{c}^{[\mathbf{i}]})\big|\mathbf{y}, \mathbf{r}^{[\mathbf{i}]}, \mathbf{c}^{[\mathbf{i}]}, \mathbf{b}^{\mathcal{D}^{[\mathbf{i}]}}\right\}, \end{aligned} \tag{35}$$

where

$$\log p(\mathbf{z}^{(q)}|b[k_q], \bar{\mathbf{b}}_q^{\mathcal{D}}[i], \mathbf{r}^{[\mathbf{i}]}, \mathbf{c}^{[\mathbf{i}]}) \sim -\frac{\mathbf{1}}{\mathcal{N}_{\mathbf{0}}} \left\|\mathbf{z}^{(\mathbf{q})} - \boldsymbol{\psi}^{[\mathbf{i}]}[\mathbf{k_q}]\, \mathbf{b}[\mathbf{k_q}]\right\|^{\mathbf{2}}, \tag{36}$$

$\boldsymbol{\psi}^{[i]}[k_q]$ is the column vector of the matrix $\boldsymbol{\Phi}^{[i]}$ at discrete frequency k_q, and following (18) and the definitions between (30) and (33), $\boldsymbol{\Phi}^{[i]}$ is easily obtained for given $\{\mathbf{r}^{[i]}, \mathbf{c}^{[i]}\}$. Inserting (36) into (35) we obtain

$$Q_{M+q}(\theta_{M+q}|\boldsymbol{\theta}^{[i]}) = \frac{1}{\mathcal{N}_0}\left(2\,\mathfrak{Re}\left\{b^*[k_q]\boldsymbol{\psi}^{[i]^\dagger}[k_q]\,\widehat{\mathbf{z}^{(q)}}^{[i]}\right\} - \left\|\boldsymbol{\psi}^{[i]}[k_q]\right\|^2 |b[k_q]|^2\right), \quad (37)$$

and $\widehat{\mathbf{z}^{(q)}}^{[i]}$ is defined as

$$\widehat{\mathbf{z}^{(q)}}^{[i]} = E\left\{\mathbf{z}^{(q)}|\mathbf{y}, \mathbf{r}^{[\mathbf{i}]}, \mathbf{c}^{[\mathbf{i}]}, \mathbf{b}^{\mathcal{D}[\mathbf{i}]}\right\} = \mathbf{y} - \sum_{\mathbf{k}=-\mathbf{K/2}\,,\,\mathbf{k}\neq\mathbf{k_q}}^{\mathbf{K/2-1}} \boldsymbol{\psi}^{[\mathbf{i}]}[\mathbf{k}]\,\mathbf{b}^{[\mathbf{i}]}[\mathbf{k}]. \quad (38)$$

In the *M-Step* of the SAGE algorithm, the estimate of $\theta_{M+q} = b[k_q]$ is updated sequentially for $q = 1, 2, \ldots, K_{\mathcal{D}}$ at the $(i+1)$st iteration according to following maximization

$$\theta_{M+q}^{[i+1]} = \arg\max_{\theta_{M+q}} Q_{M+q}(\theta_{M+q}|\boldsymbol{\theta}^{[i]}), \quad (39)$$

where $Q_{M+q}(\theta_{M+q}|\boldsymbol{\theta}^{[i]})$ is given by (37). Moreover, it follows from (37) that the qth data symbol, $b[k_q]$, can be obtained in the continuous domain by maximizing the right-hand side expression of (37). However, since $b[k_q]$ is discrete, belonging to a signal constellation point, we must quantize to its nearest constellation point. Consequently, the data update rule of the SAGE algorithm takes the following form

$$b^{[i+1]}[k_q] = \text{Quant}\left(\frac{\boldsymbol{\psi}^{[i]^\dagger}[k_q]\,\widehat{\mathbf{z}^{(q)}}^{[i]}}{\left\|\boldsymbol{\psi}^{[i]}[k_q]\right\|^2}\right), \quad (40)$$

where Quant$(\cdot)$ denotes the quantization process that quantizes its argument to its nearest data symbol constellation point. Note that, in (40), the quantized term $\boldsymbol{\psi}^{[i]^\dagger}[k_q]\,\widehat{\mathbf{z}^{(q)}}^{[i]}/\left\|\boldsymbol{\psi}^{[i]}[k_q]\right\|^2$ can be interpreted as a joint channel equalization and ICI cancellation process generated at the ith iteration step of the SAGE algorithm.

3.3 Initialization of the Algorithm

As the initialization procedure, we can apply one of the MP algorithms to obtain initial estimate of the DKL-BEM coefficient indices and the corresponding DKL-BEM coefficients, $\left\{\mathbf{r}^{[0]}, \mathbf{c}^{[0]}\right\}$, considering the observation model in (17). The MP algorithm sequentially selects the dominant BEM coefficient index that maximizes the projection of the residual vector onto the corresponding column vector of the so-called dictionary matrix. While applying the MP algorithm, in order to obtain the dictionary matrix $\mathbf{A}^{(P)}$ using (15), (16), and (17), we use the pilot symbols in their respective positions and set the unknown data symbols to zero [10]. As a first step in the MP algorithm, the column in the matrix $\mathbf{A}^{(P)} = [\mathbf{a}_1^{(P)}, \mathbf{a}_2^{(P)}, \ldots, \mathbf{a}_{\rho DL_{cp}}^{(P)}]$ which is best aligned with the residue vector, $\mathbf{R}_0 = \mathbf{y}$, is found and denoted as $\mathbf{a}_{r_1^{[0]}}^{(P)}$. Then the projection of $\mathbf{R}_0$ along this direction is removed from $\mathbf{R}_0$ and the residual $\mathbf{R}_1$ is obtained. The algorithm proceeds by sequentially choosing the column which is the best matches until termination criterion is met. At the mth iteration, the index of the vector from $\mathbf{A}^{(P)}$ most closely aligned with the residual vector $\mathbf{R}_{m-1}$ is obtained as follows

$$r_m^{[0]} = \arg\max_r \frac{|\mathbf{a}_r^{(P)^\dagger}\mathbf{R}_{m-1}|^2}{\|\mathbf{a}_r^{(P)}\|^2}, \quad r = 1, \ldots, \rho DL_{cp} \text{ and } r \notin \{r_1^{[0]}, \ldots, r_{m-1}^{[0]}\}, \quad (41)$$

and the DKL-BEM coefficient at index $r_m^{[0]}$ is

$$c_m^{[0]} = \frac{\mathbf{a}_{r_m^{[0]}}^{(P)\dagger} \mathbf{R}_{m-1}}{\|\mathbf{a}_{r_m^{[0]}}^{(P)}\|^2}. \tag{42}$$

Subsequently, the new residual vector is computed as follows

$$\mathbf{R}_m = \mathbf{R}_{m-1} - c_m^{[0]} \mathbf{a}_{r_m^{[0]}}^{(P)}. \tag{43}$$

The MP algorithm repeats (41), (42) and (43) until the termination criterion is met. We chose a termination criterion $\left|c_{\widehat{M}}^{[0]}\right| < \epsilon \sqrt{\sum_{m=1}^{\widehat{M}-1} \left|c_m^{[0]}\right|^2}$ ($\epsilon = 0.01, 0.05$, etc.). At the end of the MP algorithm, we obtain $\mathbf{c}^{[0]}$, $\mathbf{r}^{[0]}$, $\widehat{M}$, the initial estimates of the coefficient vector, DKL-BEM coefficient index vector, and the number of the DKL-BEM coefficients, respectively.

3.4 Summary of the SAGE Algorithm

Based on the results in Sections 3.1 and 3.2, $\{r_m, c_m\}_{m=1}^{\widehat{M}}$ and $\{b[k_q]\}_{q=1}^{K_\mathcal{D}}$ can be estimated and detected sequentially in the first and second stage of the SAGE algorithm, respectively, incorporating the previous estimation and detection results in the the SAGE algorithm as follows:

Initial: For $i = 0$, determine the initial estimates, $\left\{r_m^{[0]}, c_m^{[0]}\right\}_{m=1}^{\widehat{M}}$ from the MP algorithm as described in Section 3.3.

Step-1) Update the channel parameters $\{r_m^{[i+1]}, c_m^{[i+1]}\}$ sequentially for $m = 1, 2, \ldots, \widehat{M}$ from (29) computing $\widehat{\mathbf{x}^{(m)}}^{[i]}$ in (27) recursively as

$$\widehat{\mathbf{x}^{(m)}}^{[i]} = \widehat{\mathbf{x}^{(m-1)}}^{[i]} - \left(\mathbf{a}_{r_{m-1}^{[i+1]}}^{[i]} c_{m-1}^{[i+1]} - \mathbf{a}_{r_m^{[i]}}^{[i]} c_m^{[i]}\right), \tag{44}$$

where $\widehat{\mathbf{x}^{(0)}}^{[i]} = \mathbf{y} - \sum_{m'=1}^{\widehat{M}} \mathbf{a}_{r_{m'}^{[i]}}^{[i]} c_{m'}^{[i]}$ and $\mathbf{a}_{r_0^{[i+1]}}^{[i]} = \mathbf{0}$, $c_0^{[i+1]} = 0$ for all i. So, it is straightforward from (43) that $\widehat{\mathbf{x}^{(0)}}^{[0]} = \mathbf{R}_{\widehat{M}} = \mathbf{y} - \sum_{m'=1}^{\widehat{M}} \mathbf{a}_{r_{m'}^{[0]}}^{(P)} c_{m'}^{[0]}$.

Step-2) If $m = \widehat{M}$ go to *Step-3*.

Step-3) Using the estimates $\{r_m^{[i+1]}, c_m^{[i+1]}\}_{m=1}^{\widehat{M}}$ obtained in *Step-1*, update the channel matrix $\boldsymbol{\Phi}^{[i+1]}$ from (31) and (32).

Step-4) Update $b^{[i+1]}[k_q]$ sequentially for $q = 1, 2, \ldots, K_\mathcal{D}$ from (40) computing $\widehat{\mathbf{z}^{(q)}}^{[i]}$ in (38) recursively as

$$\widehat{\mathbf{z}^{(q)}}^{[i]} = \widehat{\mathbf{z}^{(q-1)}}^{[i]} - \left(\boldsymbol{\psi}^{[i+1]}[k_{q-1}]\, b^{[i+1]}[k_{q-1}] - \boldsymbol{\psi}^{[i+1]}[k_q]\, b^{[i]}[k_q]\right), \tag{45}$$

where $\widehat{\mathbf{z}^{(0)}}^{[i]} = \mathbf{y} - \sum_{k=-K/2}^{K/2-1} \boldsymbol{\psi}^{[i+1]}[k]\, b^{[i]}[k]$ and $\boldsymbol{\psi}^{[i+1]}[k_0] = \mathbf{0}$, $b^{[i+1]}[k_0] = 0$ for all i.

Step-5) If $q = K_\mathcal{D}$ go to *Step-6*

Step-6) Using the data symbols $\{b[k_q]\}_{q=1}^{K_\mathcal{D}}$ obtained in *Step-4*, update the dictionary matrix $\mathbf{A}^{[i+1]}$ from (15), (16) and (17).

Table 1 Computational complexity details

Eq. no.	Variable	CMs	CAs
Initialization			
(41), (42) and (43)	$\{r_m^{[0]}, c_m^{[0]}, \mathbf{R}_m\}_{m=1}^{M=LD}$	$\approx \rho LD^2 L_{cp}$	$\approx \rho LD^2 L_{cp}$
SAGE iteration			
In the next line of (44)	$\widehat{\mathbf{x}^{(0)}}^{[i]}$	$\approx LD$	$\approx LD$
(29) and (44)	$\{r_m^{[i+1]}, c_m^{[i+1]}, \widehat{\mathbf{x}^{(m)}}^{[i]}\}_{m=1}^{M=LD}$	$\approx \rho LD^2 L_{cp}$	$\approx \rho LD^2 L_{cp}$
(31) and (32)	$\boldsymbol{\Phi}^{[i+1]}$	$\approx NLD$	$\approx N(LD-1)$
In the next line of (45)	$\widehat{\mathbf{z}^{(0)}}^{[i]}$	$\approx N$	$\approx N$
(40) and (45)	$\{b^{[i+1]}[k_q], \widehat{\mathbf{z}^{(q)}}^{[i+1]}\}_{q=1}^{K_{\mathcal{D}}}$	$\approx 2K_{\mathcal{D}}$	$\approx 2K_{\mathcal{D}}$
(15), (16) and (17)	$\mathbf{A}^{[i+1]}$	$\approx \rho NDL_{cp}$	$\approx \rho NDL_{cp}$

Step-7) Take $i \leftarrow (i+1)$ and then continue the SAGE iterations from *Step-1* until the maximum number of the SAGE iterations is reached.

Step-8) END.

3.5 Complexity Analysis

The computational complexity is presented in Table 1 under the assumption that $N \approx K$ and $\widehat{M} \approx M = LD$ for simplicity of the complexity analysis. The initial values, $\{r_m^{[0]}, c_m^{[0]}, \mathbf{R}_m\}_{m=1}^{M}$ in (41), (42) and (43), are obtained by the MP algorithm and require approximately $\rho LD^2 L_{cp}$ complex multiplications (CMs) and $\rho LD^2 L_{cp}$ complex additions (CAs) per OFDM subcarrier as given in Table 1. In each iteration of the SAGE algorithm, $\widehat{\mathbf{x}^{(0)}}^{[i]}$, $\{r_m^{[i+1]}, c_m^{[i+1]}, \widehat{\mathbf{x}^{(m)}}^{[i+1]}\}_{m=1}^{M}$, $\boldsymbol{\Phi}^{[i+1]}$, $\widehat{\mathbf{z}^{(0)}}^{[i]}$, $\{b^{[i+1]}[k_q], \widehat{\mathbf{z}^{(q)}}^{[i+1]}\}_{q=1}^{K_{\mathcal{D}}}$ and $\mathbf{A}^{[i+1]}$ are updated, respectively, and the SAGE algorithm needs approximately $(\rho DL_{cp}+1)(N+LD)+2K_{\mathcal{D}}$ CMs and $\rho NDL_{cp}+(\rho DL_{cp}+1)LD+2K_{\mathcal{D}}$ CAs per OFDM subcarrier in each SAGE iteration. As a result, assuming that the SAGE algorithm converges in I_{iter} iterations and taking $\rho DL_{cp} \gg 1$ for simplicity, it follows from Table 1 that the computational complexity per iteration of the proposed SAGE algorithm presented in this work is approximately $(\rho DL_{cp}+1)N + \big(\rho DL_{cp}(1+\frac{1}{I_{iter}})+1\big)LD + 2K_{\mathcal{D}}$ CMs and $\rho NDL_{cp} + \big(\rho DL_{cp}(1+\frac{1}{I_{iter}})+1\big)LD + 2K_{\mathcal{D}}$ CAs per OFDM subcarrier, consequently it is in the order of $\mathcal{O}(\rho NDL_{cp})$. We compare the computation load of our algorithm with that of [19] that proposes a SAGE based joint channel estimation and data detection algorithm for OFDM systems under the assumption of rapidly time-varying non-sparse wireless multipath channel having path delay positions at integer multiple of baseband sampling duration. Thus, taking $\rho = 1$ for path delay positions at integer multiple of baseband sampling duration and $L_{cp} = L$ for non-sparsity, the order of the computational complexity of our algorithm becomes $\mathcal{O}(NDL)$ that is much less than that of [19].

4 Computer Simulations

In this section, we present computer simulations to evaluate the performance of the proposed joint channel estimation and data detection algorithm. Simulation parameters are summarized in Table 2.We assume that the channel delays, τ_ℓ, $\ell \in \{1, 2, \ldots, L\}$, are independent with

Table 2 Simulation parameters

Number of subcarriers (N)	512
Number of occupied subcarriers (K)	360
Bandwidth (BW)	5 MHz
Subcarrier spacing (Δf)	15 KHz
Sampling frequency (f_s)	7.68 MHz
Normalized Doppler frequency ($f_{dopp}T$)	0.02, 0.04
Carrier frequency (f_c)	2.5 GHz
Cyclic prefix length (T_{cp})	$40 \times T_s$
Mean of the number of the multipaths ($\bar{L}$)	5
Modulation formats	BPSK, QPSK, 16-QAM
Pilot spacing	$8 \times \Delta f$
Delay resolution (ρ)	1, 2, 4, 8
Number of the BEM coefficients (D)	3
Maximum number of the SAGE iterations	5

respect to each other and uniformly distributed within the interval $[0, T_{cp}]$. The number of paths L obeys a Poisson distribution with mean $\bar{L}$ having the following probability density function with random variable ξ

$$p_L(\xi) = \frac{(\bar{L}-1)^{(\xi-1)}\, e^{-(\bar{L}-1)}}{(\xi-1)!}, \quad \xi \in \mathbb{N}^+, \quad \bar{L} \in \mathbb{N}^+, \tag{46}$$

where $\mathbb{N}^+$ denotes the set of positive integer numbers. We chose a multipath wireless channel having an exponentially decaying power delay profile, $\Omega_\ell = Ce^{-\tau_\ell / T_{cp}}$, where C is the power normalization constant such that $\sum_{\ell=1}^{L} \Omega_\ell = 1$. We consider a comb-type pilot structure with the equally spaced pilot subcarriers. We measure the performance of the system in terms of the frequency-domain normalized mean squared error (MSE) and the symbol error rate (SER). We define the frequency-domain normalized MSE metric as follows

$$\text{MSE} = E\left\{ \frac{\sum_{n=0}^{N-1} \sum_{k=-K/2}^{K/2-1} \left| H(n,k) - \widehat{H}(n,k) \right|^2}{\sum_{n=0}^{N-1} \sum_{k=-K/2}^{K/2-1} \left| H(n,k) \right|^2} \right\}, \tag{47}$$

where the expectation is computed by the Monte Carlo method.

In our simulation plots, we compare our proposed SAGE algorithm with the OMP algorithm that is commonly proposed in the CS literature as a very popular signal recovery method [10–13,17,18]. In Figs. 1 and 2, the MSE and SER performance of our algorithm is compared with that of the OMP algorithm for two different normalized Doppler frequencies: $f_{dopp}T = 0.02 (v = 130\,\text{km/h})$, $f_{dopp}T = 0.04 (v = 260\,\text{km/h})$, employing the quadrature phase shift keying (QPSK) signaling format and delay resolution $\rho = 4$. Also, the performance curves corresponding to perfect channel state information (CSI) are included in Fig. 2 for comparison purposes, and exhibit that the performance loss in SER is not significant when perfect CSI is not available. The performance curves shown in Figs. 1 and 2 indicate that the SAGE algorithm clearly outperforms the OMP algorithm above the low SNR levels. The performance degradation at low SNR levels is because of the sensitivity of the SAGE algorithm to initial values of the parameters to be updated within the SAGE iterations. The initial estimates of the BEM coefficients obtained by the MP algorithm cannot be improved sufficiently

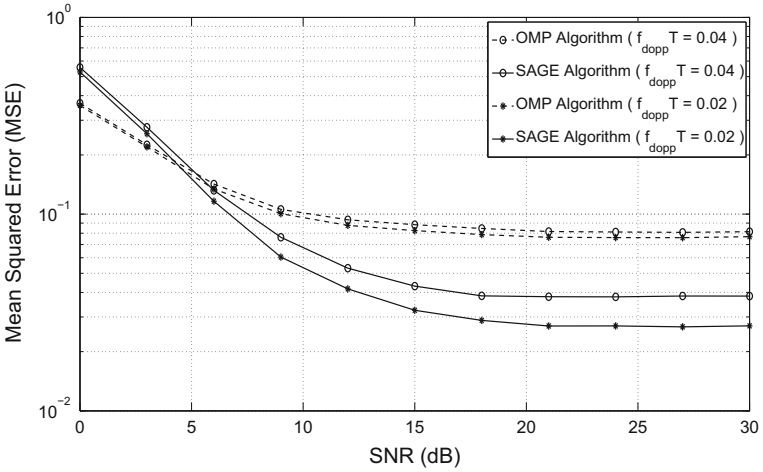

Fig. 1 MSE performance comparisons for normalized Doppler frequencies $f_{dopp}T = 0.02$ and $f_{dopp}T = 0.04$ ($\rho = 4$, QPSK signaling)

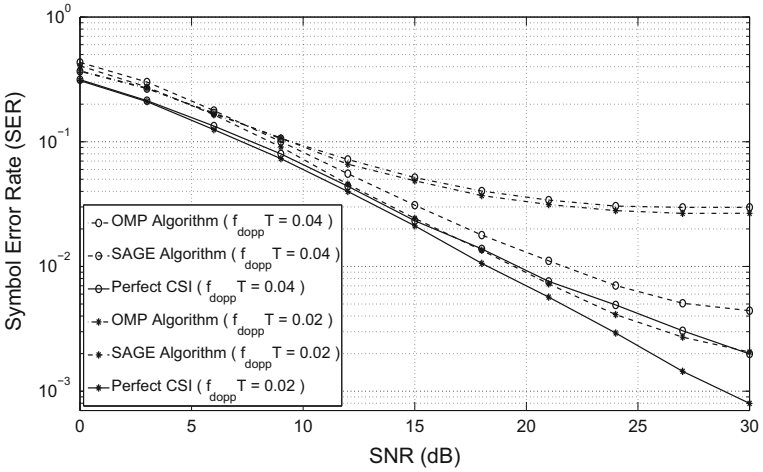

Fig. 2 SER performance comparisons for normalized Doppler frequencies $f_{dopp}T = 0.02$ and $f_{dopp}T = 0.04$ ($\rho = 4$, QPSK signaling)

at low SNR levels, consequently the SAGE algorithm cannot converge to better estimates of the BEM coefficients. Therefore, the SAGE algorithm exhibits a worse performance than that of OMP algorithm at low SNR levels.

Figures 3 and 4 show the MSE and SER performance curves of the SAGE algorithm for binary phase shift keying (BPSK), QPSK, 16 quadrature amplitude modulation (16-QAM) signaling formats. The SER performance curves for perfect CSI case are also included in Fig. 4. We conclude from the curves in Figs. 3 and 4 that the SAGE algorithm substantially outperforms the OMP algorithm for different signaling formats, as well, above the low SNR levels.

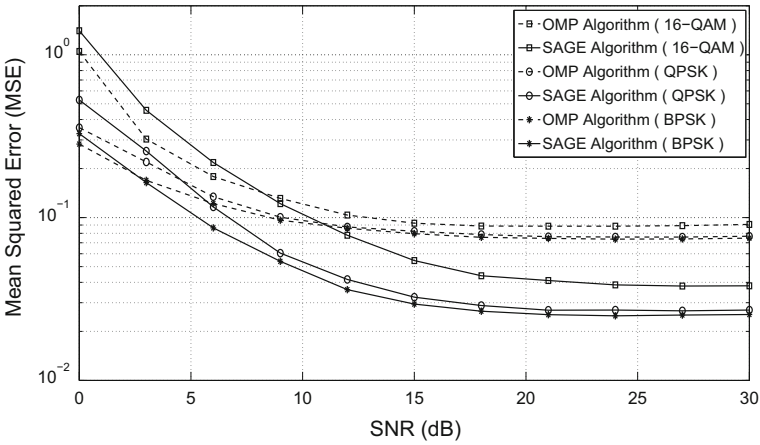

Fig. 3 MSE performance comparisons for BPSK, QPSK, and 16-QAM signaling schemes ($f_{dopp}T = 0.02$, $\rho = 4$)

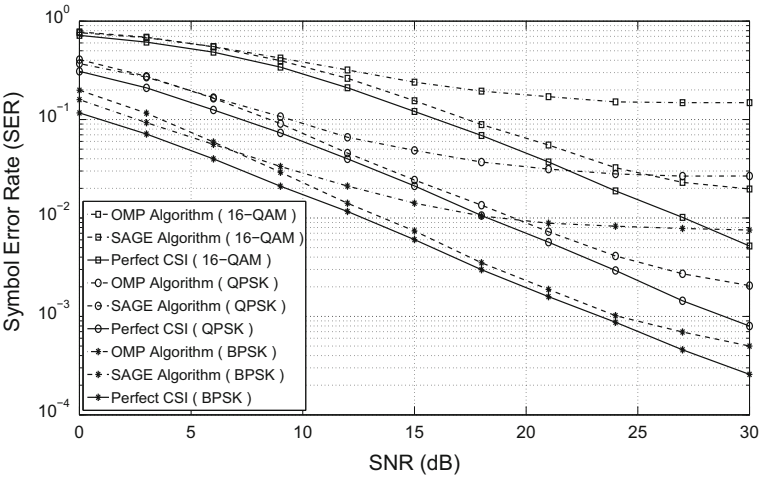

Fig. 4 SER performance comparisons for BPSK, QPSK, and 16-QAM signaling schemes ($f_{dopp}T = 0.02$, $\rho = 4$)

In order to investigate the convergency of the SAGE algorithm, the SER performance curves are plotted for each SAGE iteration in Fig. 5. As shown in Fig. 5, the SAGE algorithm converges in at most 4 iterations.

In Fig. 6, we also evaluate the SER performance of the SAGE algorithm for higher delay resolutions ($\rho = 1, 2, 4, 8$). The SER curves in Fig. 6 clearly exhibit that 4 times the delay resolution ($\rho = 4$) is sufficient for better performance of the SAGE algorithm.

5 Conclusions

In this work, a SAGE based channel estimation and symbol detection algorithm is proposed in the CS framework for OFDM systems operating over rapidly time-varying sparse multipath

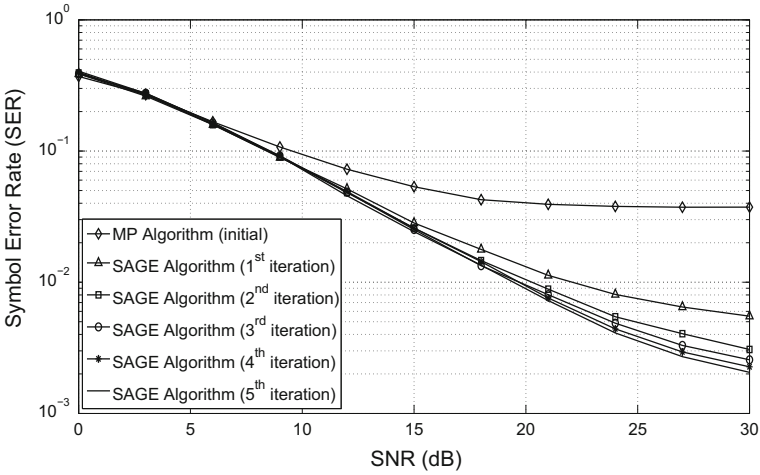

Fig. 5 SER versus SNR simulation results for each number of SAGE iterations ($f_{dopp}T = 0.02$, $\rho = 4$, QPSK signaling)

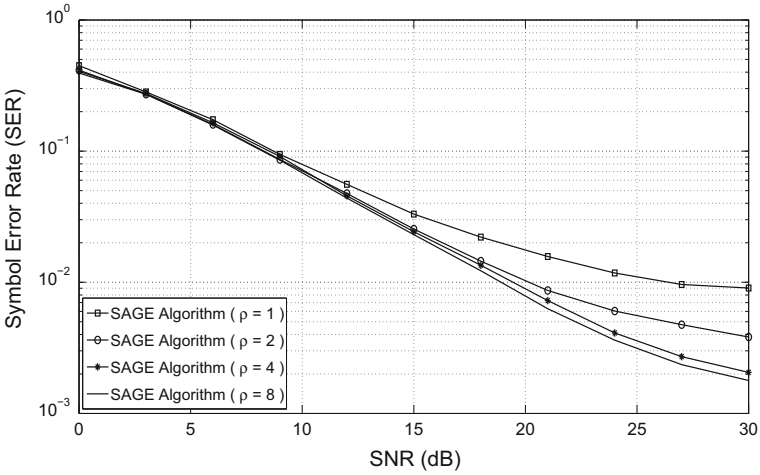

Fig. 6 SER versus SNR simulation results for each of delay resolution ($f_{dopp}T = 0.02$, QPSK signaling)

channels. For better modeling of rapidly time-varying sparse multipath channels, we use the over-complete dictionary with finer delay resolution to be able to represent sparse multipath delay positions and employ the DKL-BEM to represent the rapidly time-varying path gains within one OFDM symbol duration. The initial estimates of the BEM coefficients and the corresponding coefficient indices are obtained by the MP algorithm and they are updated together with data symbols within the proposed SAGE algorithm iterations to improve their estimation performance. The computer simulations have demonstrated that the proposed algorithm has substantially much better symbol error rate and channel estimation performance than that of the OMP algorithm that is commonly proposed in the CS literature as a very popular signal recovery method.

References

1. Li, Y., Cimini, L. J., & Sollenberger, N. R. (1998). Robust channel estimation for OFDM systems with rapid dispersive fading channels. *IEEE Transactions on Communications*, *46*(7), 902–915.
2. Edfors, O., Sandell, M., van de Beek, J. J., Wilson, S. K., & Böjesson, P. O. (1998). OFDM channel estimation by singular value decomposition. *IEEE Transactions on Communications*, *46*, 931–939.
3. Li, Y. (2000). Pilot-symbol-aided channel estimation for OFDM in wireless systems. *IEEE Transactions on Vehicular Technology*, *49*(4), 1207–1215.
4. Morelli, M., & Mengali, U. (2001). A comparison of pilot-aided channel estimation methods for OFDM systems. *IEEE Transactions on Signal Processing*, *49*(12), 3065–3073.
5. Yang, B., Letaief, K. B., Cheng, R. S., & Cao, Z. (2001). Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling. *IEEE Transactions on Communications*, *49*(12), 467–479.
6. Chen, J.-T., Paulraj, A., & Reddy, U. (2001). Multichannel maximum-likelihood sequence estimation (MLSE) equalizer for GSM using a parametric channel model. *IEEE Transactions on Communications*, *47*(1), 53–63.
7. Chen, N., Zhang, J., & Zhang, P. (2008). Improved channel estimation based on parametric channel approximation modeling for OFDM systems. *IEEE Transactions on Broadcasting*, *54*(2), 217–225.
8. Liu, S., Wang, F., Zhang, R., & Liu, Y. (2008). A simplified parametric channel estimation scheme for OFDM systems. *IEEE Transactions on Wireless Communication*, *7*(12), 5082–5090.
9. Jakobsen, M. L., Laugesen, K., Manchon, C. N., Kirkelund, G. E., Rom, C., & Fleury, B. (2010). *Parametric modeling and pilot-aided estimation of the wireless multipath channel in OFDM systems*. Proceedings of IEEE ICC 2010, Cape Town, South Africa.
10. Berger, C. R., Zhou, S., Preisig, J., & Willett, P. (2010). Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing. *IEEE Transactions on Signal Processing*, *58*(3), 1708–1721.
11. Cotter, S. F., & Rao, B. D. (2002). Sparse channel estimation via matching pursuit with application to equalization. *IEEE Transactions on Communications*, *50*(3), 374–377.
12. Wu, C.-J., & Lin, D.W. (2006) *A group matching pursuit algorithm for sparse channel estimation for OFDM transmission*. Proceedings of IEEE ICASSP 2006, Toulouse, France.
13. Li, W., & Preisig, J. C. (2007). Estimation of rapidly time-varying sparse channels. *IEEE Transactions on Oceanic Engineering*, *32*(4), 927–939.
14. Sharp, M., & Scaglione, A., (2008) *Application of sparse signal recovery to pilot-assisted channel estimation*. Proceedings of the IEEE ICASSP, Las Vegas, USA.
15. Taubock, G., Hlawatsch, F., Eiwen, D., & Rauhut, H. (2010). Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing. *IEEE Journal of Selected Topics in Signal Processing*, *4*(2), 255–271.
16. Cheng, P., Chen, Z., Rui, Y., Guo, Y. J., Gui, L., Tao, M., et al. (2013). Channel estimation for OFDM systems over doubly selective channels: A distributed compressive sensing based approach. *IEEE Transactions on Communications*, *61*(10), 4173–4185.
17. Hu, D., Wang, X., & He, L. (2013). A new sparse channel estimation and tracking method for time-varying OFDM systems. *IEEE Transactions on Vehicular Technology*, *62*(9), 4648–4653.
18. Zhou, F., Tan, J., Fan, X., & Zhang, L. (2014). A novel method for sparse channel estimation using super-resolution dictionary. *EURASIP Journal on Advances in Signal Processing*, *2014*(29), 1–11.
19. Panayirci, E., Senol, H., & Poor, H. V. (2010). Joint channel estimation, equalization and data detection for OFDM systems in the presence of very high mobility. *IEEE Transactions on Signal Process*, *58*(8), 4225–4238.
20. Şenol, H., Panayırcı, E., & Poor, H. V. (2012). Non-data-aided joint channel estimation and equalization for OFDM systems in very rapidly varying mobile channels. *IEEE Transactions on Signal Process*, *60*(8), 4236–4253.
21. Şenol, H., Çırpan, H. A., & Panayırcı, E. (2005). A low complexity KL-expansion based channel estimator for OFDM systems. *EURASIP Journal on Wireless Communication Networking*, *2005*(2), 163–174.
22. Zemen, T., & Mecklenbrauker, C. F. (2005). Time-variant channel estimation using discrete prolate spheroidal sequences. *IEEE Transactions on Signal Process*, *53*(9), 3597–3607.
23. Giannakis, G. B., & Tepedelenlioglu, C. (1998). Basis expansion models and diversity techniques for blind equalization of time-varying channels. *Proceedings of the IEEE*, *86*(10), 1969–1986.
24. Tang, Z., Cannizzaro, R. C., Leus, G., & Banelli, P. (2007). Pilot-assisted time-varying channel estimation for OFDM systems. *IEEE Transactions on Signal Process*, *55*(5), 2226–2238.
25. Hijazi, H., & Ros, L. (2009). Polynomial estimation of time-varying multipath gains with intercarrier interference mitigation in OFDM systems. *IEEE Transactions on Vehicular Technology*, *58*(1), 140–151.

26. Fessler, J. A., & Hero, A. O. (1994). Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Process*, *42*(10), 2664–2677.

Habib Şenol received the B.Sc. and M.Sc. degrees from Istanbul University, Istanbul, Turkey, in 1993 and 1999, respectively, both in Electronics Engineering. From 1996 to 1999, he was a research assistant at Istanbul University. He received the Ph.D. degree in Electronics Engineering from Işık University, Istanbul, Turkey, in 2006. He is currently faculty member of Computer Engineering at Kadir Has University, Istanbul, Turkey. Dr. Şenol was a post-doctoral researcher at the Department of Electrical Engineering, Arizona State University, USA, in 2007. Dr. Şenol's research interests cover statistical signal processing, estimation and equalization algorithms for wireless communications, multicarrier (OFDM) systems, distributed detection and estimation.