Joint Channel Estimation, Equalization, and Data Detection for OFDM Systems in the Presence of Very High Mobility

Erdal Panayırcı, *Fellow, IEEE*, Habib Şenol, *Member, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—This paper is concerned with the challenging and timely problem of joint channel estimation, equalization, and data detection for uplink orthogonal frequency division multiplexing (OFDM) systems in the presence of frequency selective and very rapidly time varying channels. The resulting algorithm is based on the space alternating generalized expectation maximization (SAGE) technique which is particularly well suited to multicarrier signal formats leading to a receiver structure that also incorporates interchannel interference (ICI) cancelation. In order to reduce the computational complexity of the algorithm, band-limited, discrete cosine orthogonal basis functions are employed to represent the rapidly time-varying fading channel by the discrete cosine serial expansion coefficients. It is shown that, depending on the normalized Doppler frequency, only a small number of expansion coefficients is sufficient to approximate the channel perfectly and there is no need to know the correlation function of the input signal. In this way, the resulting reduced dimensional channel coefficients are estimated and the data symbols detected iteratively with tractable complexity. The proposed SAGE joint detection algorithm updates the data sequences serially and the channel parameters are updated in parallel, leading to a receiver structure that also incorporates ICI cancelation. Computer simulations show that the cosine transformation represents the time-varying channel very effectively and the proposed algorithm has excellent symbol error rate and channel estimation performance even with a very small number of channel expansion coefficients employed in the algorithm, resulting in substantial reduction of the computational complexity.**

***Index Terms*—Intercarrier interference suppression, joint data detection and channel estimation, orthogonal frequency-division multiplexing (OFDM), rapidly varying wireless channels, space alternating generalized expectation maximization (SAGE) algorithm.**

I. Introduction

ORTHOGONAL frequency-division multiplexing has been shown to be an effective method to overcome intersymbol interference (ISI) caused by frequency-selective fading with a simple transceiver structure. Consequently, has become a key air interface for broadband high speed communication systems being standardized as the IEEE's 802.16 family—better known as Mobile Worldwide Interoperability Microwave Systems for Next-Generation Wireless Communication Systems (WiMAX)—and by the Third-Generation Partnership Project (3GPP) in the form of its Long-Term Evolution (LTE) project. Both systems employ orthogonal frequency division multiplexing/multiple access (OFDMA) as well as a new single-carrier frequency-division multiple access (SC-FDMA) format. To promote the IEEE 802.16 standards, recently, a high mobility feature has been introduced (IEEE 802.16m) to enable mobile broadband services at vehicular speeds beyond 120 km/h.

While many studies of OFDM receiver design have appeared in the literature for quasi-static channels [1]–[10], there is much left for investigation in the area of fast fading channel estimation, equalization, and data detection. Since mobility support is widely considered to be one of the key features in next generation wireless communication systems, OFDM transmission over very rapidly time varying multipath fading channels has been considered in a number of recent papers [11]–[26]. However, most of the these cited works have focused on finding low-complexity channel equalization and estimation techniques, separately, for OFDM receivers in the presence of highly mobile environments.

OFDM can drastically simplify the equalization problem by turning the frequency selective channel into a flat fading channel. A simple one-tap equalizer is needed to estimate the channel and recover the data. However, in fading channels with very high mobilities, the time variation of the channel over an OFDM symbol period results in a loss of subchannel orthogonality which leads to inter-channel interference (ICI) due to power leakage among OFDM subcarriers. In addition to this, interblock interference (IBI) arises when the channel delay spread is larger than the cyclic prefix (CP), which again results in ICI. Techniques for equalization in high mobility channels range from linear equalization, based on the zero-forcing (ZF) or the minimum mean-squared error (MMSE) criterion [11], [14]–[17], [19], [20], to nonlinear equalization based on decision-feedback or ICI cancelation [16], [17], [19], [20]. It has been shown that nonlinear equalizers based on ICI cancelation generally outperform linear approaches [17], [19], [20]. However, linear equalizers still preserve their importance mainly because they are less complex. In [19], the performance of matched filter (MF), least squares (LS), and MMSE successive

Manuscript received August 22, 2009; accepted March 27, 2010. Date of publication April 15, 2010; date current version July 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ta-Sung Lee. This work was supported by the NEWCOM++ Network of Excellence in Wireless Communications and WIMAGIC Strep projects funded through the EC 7th Framework Programs and by the U.S. National Science Foundation by Grant CNS-09-05398.

E. Panayırcı was with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA. He is now with Department of Electronics Engineering, Kadir Has University, Cibali 34083, Istanbul, Turkey (e-mail: eepanay@khas.edu.tr).

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

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Digital Object Identifier 10.1109/TSP.2010.2048317

detection with optimal ordering is investigated. However, since the number of subcarriers is usually very large in high-speed wide-band wireless standards, even the linear MMSE equalizer considered in [19] demands very high computational load. The specific structure of the Doppler-induced ICI in OFDM systems operating over highly mobile channels presents a distinctive feature of limited support of the Doppler spread that can be exploited by the receiver. References [14]–[17], and [20] exploit the banded character of the frequency-domain channel matrix to reach a complexity that is only linear in the number of subcarriers. In a certain sense, the assumption of a banded frequency-domain channel matrix is a natural extension of the time-invariant channel case, in which the frequency-domain channel matrix is diagonal and hence banded with the smallest possible bandwidth. In [14], using the banded structure of the channel matrix, a simple frequency domain equalizer is proposed that can compensate for these ICI coefficients, thereby significantly affecting the loss of subchannel orthogonality. However, the detection performance of the technique degrades substantially, since the data to be detected cannot fully use the contributing observation elements. The work presented in [20] combined [14] and [17] and derived a recursive decision feedback equalizer receiver for ICI suppression. The iterative MMSE serial linear equalizer (SLE) of [17], which takes the banded structure of the channel matrix into account, seems to be one of the most promising approaches to compensate for ICI. Iterative MMSE is then applied to estimate frequency-domain symbols. In [15], a block MMSE equalizer for OFDM systems over time-varying channels is presented. By exploiting the banded structure of the frequency-domain channel matrix, the complexity of the resulting algorithm turns out to be smaller than that of [17]. In [27], a new computationally feasible, maximum *a posteriori* probability (MAP)-based data symbol detection algorithm is proposed for OFDM systems operating in highly mobile channels, as an alternative to the existing suboptimal equalization/detection techniques summarized in the above paragraphs.

On the other hand, the recent work on the separate equalization and estimation for OFDM systems in a highly mobile environment can be summarized as follows. For a rapidly time-varying channel, the time-domain channel estimation method proposed in [19] is a potential candidate for the channel estimator, in order to mitigate the ICI. This technique estimates the fading channel by exploiting the time-varying nature of the channel as a provider of time diversity and reduces the computational complexity using the singular-value decomposition (SVD) method. However, the linear MMSE successive detection with optimal ordering proposed in [19] along with channel estimation demand very high computation, since the number of subcarriers is usually very large; thus it may not be feasible in practical systems. In [13], a low-complexity equalizer is designed first, assuming the channel is banded. Then, employing this equalizer, a pilot-aided MMSE channel estimation scheme for a time-varying wide-sense stationary uncorrelated scatters channel model is proposed. In [18] a general framework for a controlled removal of ICI and channel acquisition is proposed. A finite power series expansion for the time-varying frequency response is used and channel acquisition and ICI removal are accomplished in the frequency domain. In [23], to handle rapid variation within an OFDM symbol, a pilot-based estimation scheme using channel interpolation was proposed. Moreover, coupled with the proposed channel estimation scheme, a simple Doppler frequency estimation scheme was proposed. In [24], to compensate for the ICI, a modified Kalman filter (MKF) channel estimator for OFDM systems in a fast and frequency-selective Rayleigh fading channel was proposed. The time-varying channel was modeled as an autoregressive (AR) process and the proposed MKF was used to estimate the AR parameters. In addition, a channel predictor using regression analysis and an MMSE time-domain equalizer were also proposed to track the time-varying channel. The difference between the method described in [19] and the one proposed in [24] was that the former assumed the value of the fading parameter (or Doppler frequency) to be known in advance, whereas the latter estimated it by means of the MKF. In [25], two methods to mitigate ICI in an OFDM system with coherent channel estimation were proposed. Both methods employed a piece-wise linear approximation to estimate channel time-variations in each OFDM symbol. The first method extracted channel time-variation information from the cyclic prefix while the second method estimated these variations using the next symbol. Moreover, a closed-form expression for the improvement in average signal-to-interference ratio (SIR) was derived for a narrowband time-varying channel. However, piece-wise linear approximation does not hold any more in the presence of very high mobility. In [26], a decision-directed channel predictor for OFDM communications over time-varying channels was proposed. The channel prediction algorithm presented in [26] was capable of yielding up-to-date channel state information even without regular transmission of pilot symbols. Moreover, the proposed normalized least mean squares (NLMS) and recursive least squares (RLS) adaptive predictors do not require any prior statistical knowledge and are able to track nonstationary channels and noise statistics. However, to avoid error propagation, it was concluded that a certain minimum signal-to-noise ratio (SNR) was required and that the algorithm could not work for very high mobility.

In this paper, a computationally feasible space alternating generalized expectation maximization (SAGE) algorithm is proposed for the problem of joint multiuser data detection [28], channel estimation and equalization for OFDM systems operating in highly mobile and frequency selective channels. The channel variation over the duration of a data block is upper bounded by the maximum Doppler bandwidth which is determined by the maximum speed of the users. We exploit the band-limited discrete-cosine orthogonal basis functions to represent the time-varying fading channel through a discrete cosine serial expansion of low dimensionality. In this way, the resulting reduced dimensional channel coefficients are estimated and the data symbols detected iteratively with tractable complexity. The resulting SAGE-based receiver scheme comprises a channel estimator, interference canceler, equalizer and soft-input/hard-output serial data detector in each iteration.

This detection algorithm is compared with previously proposed algorithms in terms of both symbol error rate (SER) and complexity requirements. Computational complexity investiga-

tion as well as simulation results indicate that our algorithm performs quite well while having significant complexity advantages over the existing suboptimal detection and equalization algorithms described above.

The paper is organized as follows. Section II presents the system, including observation and time varying channel models, explains the discrete-cosine expansion of the time-varying channel and describes the pilot tone structures. In Section III, the proposed SAGE algorithm is presented for joint channel estimation, equalization and data detection. Furthermore we discuss initialization and computational complexity. In Section IV, the performance of the proposed receiver is evaluated via computer simulations. Finally Section V summarizes the main conclusions of the paper.

II. System Model

A. Signal and Channel Models

We consider an OFDM system with N subcarriers. At the transmitter, K out of N subcarriers are actively employed to transmit data symbols and nothing is transmitted from the remaining $N-K$ carriers. The frequency-domain transmitted data symbols are denoted as $s(n,k)$, where n is the OFDM symbol discrete-time index and $k \in \{0,1\cdots,K-1\}$ is the subcarrier index. A cyclic prefix of length L_c is then added. We assume a time-varying mobile radio channel with discrete-time impulse response $h(n,l)$, $l=0,1,\cdots,L-1$, where L is the maximum channel length with $L \le L_c$. The Fourier transform of the channel impulse response at time $n=0,1,\cdots$, is defined as $H(n,k) \triangleq \sum_{l=0}^{L-1} h(n,l)\exp(-j2\pi lk/N)$, where $k=0,1,\cdots,K-1$ are the indices of the discrete OFDM subchannel frequencies. For a classical OFDM system with cyclic prefix duration greater than the channel impulse response length, the received signal is not corrupted by previous symbols and therefore all OFDM symbols can be processed separately. At the receiver, after matched filtering, symbol-rate sampling and discarding the symbols falling in the cyclic prefix, the received signal at the input of the discrete Fourier transform (DFT) and during the transmission of the m OFDM symbol can be expressed as [19]

$$r(mN_g+p)=\sum_{q=0}^{K-1} s(m,q)\frac{1}{N}\sum_{l=0}^{L-1} h(mN_g+p,l) \times \exp\left(j\frac{2\pi q(p-l)}{N}\right)+w(mN_g+p) \quad (1)$$

for $p=0,1,\cdots,N-1$ and $m=0,1,\cdots,M-1$, where M represents one OFDM frame length consisting of M consecutive OFDM symbols. $N_g \triangleq N+L_c$ and $w(\cdot)$ is zero-mean complex additive Gaussian noise with variance N_0. Note that when the normalized Doppler frequency is sufficiently small, the time varying-channel impulse response can be assumed to be constant over the duration of one OFDM symbol; that is $h(mN_g+p,l)\approx h(m,l)$ for $p=0,1,\cdots N-1$. Then it can be easily shown from (1) that the received signal at the output of DFT takes the known form

$$R(n,k)=s(n,k)H(n,k)+W(n,k), \quad k=0,1,\cdots,K-1$$

where $R(n,k), W(n,k)$ and $H(n,k)$ are received signal, noise and channel coefficients, respectively, all represented in the frequency-domain, corresponding to the nth OFDM symbol and kth subchannel.

If we focus on the detection of the kth data symbol transmitted during the mth OFDM timing slot, the final expression for the received signal can be expressed in vector form as follows:

$$\mathbf{r}(m)=s(m,k)\mathbf{V}_k\mathbf{h}(m)+\sum_{q=0,q\neq k} s(m,q)\mathbf{V}_q\mathbf{h}(m)+\mathbf{w}(m) \quad (2)$$

where

$$\mathbf{r}(m)=[r(mN_g),r(mN_g+1),\cdots,r(mN_g+N-1)]^T\in\mathcal{C}^N$$
$$\mathbf{w}(m)=[w(mN_g),w(mN_g+1),\cdots,w(mN_g+N-1)]^T\in\mathcal{C}^N$$

and

$$\mathbf{V}_q=\mathbf{F}_L^T(q)\otimes\frac{1}{N}\mathrm{diag}\left(\mathbf{F}_N^{\dagger}(q)\right)\in\mathcal{C}^{N\times NL}, \quad q=0,1,\cdots,K-1 \quad (3)$$

together with

$$\mathbf{F}_L(q) \triangleq \left[1,\exp\left(-j\frac{2\pi q}{N}\right),\cdots,\exp\left(-j\frac{2\pi q(L-1)}{N}\right)\right]^T\in\mathcal{C}^L.$$

The vector $\mathbf{h}(m)$ denotes the time-varying channel impulse response during the mth OFDM symbol

$$\mathbf{h}(m)=\left[\mathbf{h}_0^T(m),\mathbf{h}_1^T(m),\cdots,\mathbf{h}_{L-1}^T(m)\right]^T\in\mathcal{C}^{NL} \quad (4)$$

where $\mathbf{h}_l(m)=[h(mN_g,l),h(mN_g+1,l),\cdots,h(mN_g+N-1,l]^T$, $l=0,1,\cdots,L-1$ represents L-path wide-sense stationary uncorrelated scattering (WSSUS) Rayleigh fading coefficients at the (mN_g+p)th discrete-times. Assuming the Jakes' model, the autocorrelation function of the channel is

$$E\{h(mN_g+p,l)h^*(m'N_g+p',l')\} = \sigma_l^2 J_0\left(2\pi f_d T_s\left((m-m')N_g+(p-p')\right)\right)\delta(l-l') \quad (5)$$

where σ_l^2, $l=0,1\cdots,L-1$, represents the normalized power of the lth path of the channel satisfying $\sum_l \sigma_l^2=1$. $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_d is the Doppler shift due to the vehicle motion and $\delta(\cdot)$ is the Kronecker delta. $T_s=T/(N+L_c)$, T being the OFDM symbol duration. Finally, $\mathbf{w}(m)$ is the complex white Gaussian noise vector with zero-mean and $E[\mathbf{w}(m)\mathbf{w}^{\dagger}(m)]=N_0\mathbf{I}_N$. The derivation of (2) is given in the Appendix A.

B. Channel Basis Expansion

The performance of the receiver depends critically on the estimate of the time-varying channel impulse response $\mathbf{h}=[\mathbf{h}^T(0),\mathbf{h}^T(1)\cdots,\mathbf{h}^T(M-1)]^T\in\mathcal{C}^{MNL}$ from the $MN(MN<MNL)$ dimensional received vector $\mathbf{r}=[\mathbf{r}^T(0),\mathbf{r}^T(1)\cdots,\mathbf{r}^T(M-1)]^T$. It seems the estimation of the $MNL\times 1$ channel vector $\mathbf{h}$ is impossible by means of $\mathbf{r}$ since there are more unknowns to be determined than known

equations. However, the banded property of the channel matrix [17], [29] enables us to reduce the number of parameters needed for channel estimation substantially, consequently reducing the computational complexity of the channel estimation step.

We first apply a suitable basis expansion which describes the time variations of the discrete-time channel impulse response $h(mN_g+p,l)$ over a data block consisting of M OFDM symbols. We do not make any assumption regarding the amount of time-variation (equivalently, Doppler frequency) in the channel. For notational simplicity, let $t \triangleq mN_g+p$. Then

$$\begin{pmatrix} m=0,1,\cdots,M-1 \\ \text{and} \\ p=0,1,\cdots,N_g-1 \end{pmatrix} \Leftrightarrow \quad t=0,1,\cdots,MN_g-1.$$

For each channel path $l=0,1,\cdots,L-1$, the channel coefficients $h(t,l)$ can be represented as weighted sums of N_gM orthogonal basis functions $\{\psi_d(t)\}$ in the interval $[0,N_gMT_s]$:

$$h(t,l)=\sum_{d=0}^{MN_g-1}\psi_d(t)c(d,l), \quad t=0,1,\cdots,MN_g-1 \quad (6)$$

where $\{c(d,l)\}$ represent the expansion coefficients. As $h(.,l)$ is essentially a lowpass process whose bandwidth is determined by the Doppler frequency, it can be well approximated by the weighted sum of a substantially fewer number $D(\ll MN_g)$ of suitable basis functions:

$$\widetilde{h}(t,l)=\sum_{d=0}^{D-1}\psi_d(t)c(d,l) \quad t=0,1,\cdots,MN_g-1. \quad (7)$$

Similarly, using the orthogonality property of the basis functions, the expansion coefficients can be evaluated by the inverse transformation as

$$c(d,l)=\sum_{t=0}^{MN_g-1}\psi_d(t)h(t,l) \quad d=0,1,\cdots,D-1. \quad (8)$$

Employing complex exponentials as basis functions has been widely considered in the literature due to the orthogonality among columns of the basis-expansion matrix error when the channel is varying very rapidly. Recently the basis expansion matrix has been considered using the Karhunen-Loève transformation (KLT) [30]. The KLT based expansion model achieves minimum mean square channel modeling error and the expansion coefficients are uncorrelated. However, the implementation of the KLT-based expansion algorithm is computationally expensive and requires knowledge of the channel statistics.

In our work, we make use of the orthonormal discrete cosine transform (DCT) basis functions defined as

$$\psi_d(t)=\begin{cases}\sqrt{1/MN_g} & \text{if } d=0,\\ \sqrt{2/MN_g}\cos\left[(\pi d/MN_g)(t+1/2)\right] & \text{if } d>0.\end{cases} \quad (9)$$

Hence, $c(d,l)$ is the dth DCT-coefficient of $h(t,l)$. The dimension D of the basis expansion satisfies $D_{lower} \le D \le MN_g$. The lower bound D_{lower} is given by $D_{lower}=\lceil 2(f_D)_{max}M+1\rceil$, where $(f_D)_{max}$ is the maximum (one-sided) normalized Doppler bandwidth defined by

$$(f_D)_{max}=\frac{v_{max}f_c}{c}T$$

with v_{max}, f_c and c the maximum supported speed, the carrier frequency and the speed of light, respectively, and T the OFDM symbol duration. As $\psi_d(t)$ has its energy concentrated near the frequencies $\pi d/2MN_gT$ and $-\pi d/2MN_gT$ for $d=0,1,\cdots,D-1$, the DCT basis functions are well suited to represent the low-pass equivalent of the channel by means of a small number of basis functions. Also, the DCT basis functions have the advantages of being independent of the channel statistics and having expansion coefficients that become uncorrelated as the number of observations MN_g gets larger, as proven in Appendix B. By choosing D, we control the channel modeling mean square error (MSE)

$$MSE=\frac{1}{LMN_g}\sum_{l=0}^{L-1}\sum_{t=0}^{MN_g-1}E\left\{\left|h(t,l)-\tilde{h}(t,l)\right|^2\right\}. \quad (10)$$

Thus, for each channel path $l(l=0,1,\cdots,L-1)$, the channel and the expansion coefficients can be expressed in matrix form

$$\widetilde{\mathbf{h}}_l=\mathbf{\Psi}\mathbf{c}_l \quad (11)$$

and

$$\mathbf{c}_l=\mathbf{\Psi}^{\dagger}\widetilde{\mathbf{h}}_l \quad (12)$$

where

$$\widetilde{\mathbf{h}}_l=\left[\widetilde{h}(0,l),\widetilde{h}(1,l),\cdots,\widetilde{h}(MN_g-1,l)\right]^T\in\mathcal{C}^{MN_g}$$
$$\mathbf{c}_l=[c(0,l),c(1,l),\cdots,c(D-1,l)]^T\in\mathcal{C}^D$$

and $\mathbf{\Psi}$ represents the DCT matrix expressed as

$$\mathbf{\Psi}=[\boldsymbol{\psi}(0),\boldsymbol{\psi}(1),\cdots,\boldsymbol{\psi}(MN_g-1)]^T\in\mathcal{R}^{MN_g\times D} \quad (13)$$

and

$$\boldsymbol{\psi}(t)=[\psi_0(t),\psi_1(t),\cdots,\psi_{D-1}(t)]^T, \quad t=0,1,\cdots,MN_g-1.$$

Furthermore after removing the CP, it can be shown from (7) and (2) that the dimension of the channel vectors $\widetilde{\mathbf{h}}_l$, $l=0,1,\cdots,L-1$, in (11) reduces from MN_g to MN. Then it can be easily seen that the channel vector $\widetilde{\mathbf{h}}(m)$ in (4) is related to $\widetilde{\mathbf{h}}_l$ in (11) as

$$\widetilde{\mathbf{h}}_l(m)=\widetilde{\mathbf{h}}_l[mN_g:mN_g+N-1]$$

where we have used the notation $\mathbf{x}[i:j]$ to denote entry i to entry j of a vector $\mathbf{x}$. Finally from (11) it follows that

$$\widetilde{\mathbf{h}}(m)=\mathbf{\Phi}(m)\mathbf{c} \quad (14)$$

where

$$\widetilde{\mathbf{h}}(m)=\left[\widetilde{\mathbf{h}}_0^T(m),\widetilde{\mathbf{h}}_1^T(m),\cdots,\widetilde{\mathbf{h}}_{L-1}^T(m)\right]^T\in\mathcal{C}^{NL}$$
$$\mathbf{c}=[\mathbf{c}_0,\mathbf{c}_1,\cdots,\mathbf{c}_{L-1}]^T\in\mathcal{C}^{DL}$$

and

$$\mathbf{\Phi}(m) = \text{diag}\underbrace{(\mathbf{\Psi}(m), \mathbf{\Psi}(m)\cdots, \mathbf{\Psi}(m))}_{L} \in \mathcal{C}^{NL\times DL}$$

with

$$\mathbf{\Psi}(m) = [\boldsymbol{\psi}(mN_g), \boldsymbol{\psi}(mN_g+1), \cdots, \boldsymbol{\psi}(mN_g+N-1)]^T \in \mathcal{R}^{N\times D}.$$

Finally, substituting (14) into (2), the received signal is expressed in terms of the reduced dimensional channel vector $\mathbf{c}$ as follows:

$$\mathbf{r}(m) = s(m,k)\mathbf{A}_k(m)\mathbf{c} + \sum_{q=0,q\neq k}^{K-1} s(q,m)\mathbf{A}_q(m)\mathbf{c} + \mathbf{w}(m) \tag{15}$$

where $\mathbf{A}_q(m) \triangleq \mathbf{V}_q\mathbf{\Phi}(m) \in \mathcal{C}^{N\times DL}$.

For purposes that will be clear below, we also express (15) in a more compact matrix form as follows:

$$\mathbf{r} = \mathbf{Z}_\mathbf{s}\mathbf{c} + \mathbf{w} \tag{16}$$

where

$$\mathbf{r} = \left[\mathbf{r}^T(0), \mathbf{r}^T(1), \cdots, \mathbf{r}^T(M-1)\right]^T \in \mathcal{C}^{NM}$$
$$\mathbf{Z}_\mathbf{s} = \left[\mathbf{Z}_\mathbf{s}^T(0), \mathbf{Z}_\mathbf{s}^T(1), \cdots, \mathbf{Z}_\mathbf{s}^T(M-1)\right]^T \in \mathcal{C}^{NM\times DL}$$

and

$$\mathbf{Z}_\mathbf{s}(m) = \sum_{q=0}^{K-1} s(m,q)\mathbf{A}_q(m).$$

C. Pilot Symbol Selection

For channel estimation and especially for initialization of the SAGE algorithm as explained in the following section, insertion of pilot symbols is necessary. Although, several pilot patterns are possible, for ease of implementation we consider the scheme in which all subcarriers in a given time slot are dedicated to pilot symbols. Due to high-mobility, time-domain correlation is important and plays a key role in the initial reduced complexity MMSE channel estimation used in the SAGE algorithm. We assume there are P pilot OFDM symbols located at time-slots $m_1, m_2, \cdots, m_P$ where $m_p \in \{0,1,\cdots,M-1\}$. They can be arranged in a way so that pilot spacings can be adjusted with respect to the time variations of the channel. When the channel is slowly varying, the time-domain correlation decays at a slower rate and consequently, pilot spacing can be chosen wider.

III. Data Detection Using the SAGE Technique

The problem of interest is to derive an iterative algorithm based on the SAGE technique for data detection without complete channel state information, employing the signal model given by (2). Since the SAGE method has been studied and applied to a number of problems in communications over the years, the details of the algorithm will not be presented in this paper. The reader is referred to [31] for a general exposition of the SAGE algorithm and to [9] for its application to an estimation problem related to the work herein. A suitable approach for applying the SAGE algorithm for detection of the nonpilot data signals $s(m,k)$, $k = 0,1,\cdots,K-1$, $m \in \{0,1,\cdots,M-1\}$ and $m \neq m_1, m_2, \cdots, m_P$, is to decompose the received signal in (15) into the sum [32]

$$\mathbf{r}(m) = \mathbf{y}_k(m) + \overline{\mathbf{y}}_k(m) \tag{17}$$

where

$$\mathbf{y}_k(m) = s(m,k)\mathbf{A}_k(m)\mathbf{c} + \mathbf{w}(m) \tag{18}$$

and

$$\overline{\mathbf{y}}_k(m) = \sum_{q=0,q\neq k}^{K-1} s(m,q)\mathbf{A}_q(m)\mathbf{c}. \tag{19}$$

We now derive a SAGE algorithm for detecting the nonpilot OFDM data vectors in the set $\mathbf{s} \triangleq \{\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{K-1}\}$, where $\mathbf{s}_k = [s(0,k), s(1,k), \cdots, s(M-1,k)]^T$, transmitted within an observed frame consisting of M OFDM symbols, based on the received vector $\mathbf{r}$. To obtain a receiver architecture that iterates between soft-data and channel estimation, one might choose the parameter vector to be $\{\mathbf{s}\}$. At iteration (i), only the data symbol vector of the subchannel k, $\{\mathbf{s}_k\}$ is updated, while the symbol vectors of other subchannels $\overline{\mathbf{s}}_k = \mathbf{s} \setminus \mathbf{s}_k$ are kept fixed, where the notation "\" denotes the set exclusion operator.

In the SAGE algorithm, we view the observed data $\mathbf{r}$ as the *incomplete* data and since $\mathbf{c}$ is unknown, we incorporate $\mathbf{c}$ into the *admissible hidden* data set as $\chi_k = \{\mathbf{y}_k, \mathbf{c}\}$ to which the incomplete data $\mathbf{r}$ are related through a possibly nondeterministic mapping [31].

A. Expectation-Step (E-Step)

The SAGE algorithm is based on the expectation-maximization (EM) algorithm. The first step to implement the EM algorithm, called the *Expectation Step (E-Step)*, is the computation of the average log-likelihood function, averaged over $\mathbf{c}$. The conditional expectation is taken over $\boldsymbol{\chi}$ given the observation $\mathbf{r}$ and given that $\mathbf{s}$ equals its estimate calculated at ith iteration

$$Q_k\left(\mathbf{s}_k|\mathbf{s}^{(i)}\right) = E\left\{\ln p\left(\mathbf{y}_k|\mathbf{s}_k, \overline{\mathbf{s}}_k^{(i)}, \mathbf{c}\right)|\mathbf{r}, \mathbf{s}^{(i)}\right\}. \tag{20}$$

By neglecting the terms independent of $\mathbf{s}$, $\ln p(\mathbf{y}_k|\mathbf{s}_k, \overline{\mathbf{s}}_k^{(i)}, \mathbf{c})$ can be calculated from (20) as

$$\ln p\left(\mathbf{y}_k|\mathbf{s}_k, \overline{\mathbf{s}}_k^{(i)}, \mathbf{c}\right) \sim \sum_{m=0}^{M-1}\left[\Re\left\{s^*(m,k)\mathbf{c}^\dagger\mathbf{A}_k^\dagger(m)\mathbf{y}_k(m)\right\} - \frac{1}{2}|s(m,k)|^2\mathbf{c}^\dagger\mathbf{A}_k^\dagger(m)\mathbf{A}_k(m)\mathbf{c}\right] \tag{21}$$

where $\Re\{\cdot\}$ denotes the real part of its argument. Inserting (21) into (20), we have for $Q_k(\mathbf{s}_k|\mathbf{s}^{(i)})$

$$Q_k\left(\mathbf{s}_k|\mathbf{s}^{(i)}\right) = \sum_{m=0}^{M-1}\left[\Re\left\{s^*(m,k)\left(\mathbf{c}^\dagger\mathbf{A}_k^\dagger(m)\mathbf{y}_k(m)\right)^{(i)}\right\} - \frac{1}{2}|s(m,k)|^2\left(\mathbf{c}^\dagger\mathbf{A}_k^\dagger(m)\mathbf{A}_k(m)\mathbf{c}\right)^{(i)}\right] \tag{22}$$

where

$$\left(\mathbf{c}^\dagger \mathbf{A}_k^\dagger(m)\mathbf{y}_k(m)\right)^{(i)} \triangleq E\left\{\mathbf{c}^\dagger \mathbf{A}_k^\dagger(m)\mathbf{y}_k(m)|\mathbf{r},\mathbf{s}^{(i)}\right\} \quad (23)$$

and

$$\left(\mathbf{c}^\dagger \mathbf{A}_k^\dagger(m)\mathbf{A}_k(m)\mathbf{c}\right)^{(i)} \triangleq E\left\{\mathbf{c}^\dagger \mathbf{A}_k^\dagger(m)\mathbf{A}_k(m)\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right\}. \quad (24)$$

Eq. (23) can be calculated by applying the conditional expectation rule as

$$\left(\mathbf{c}^\dagger \mathbf{A}_k^\dagger(m)\mathbf{y}_k(m)\right)^{(i)} = E\left\{\mathbf{c}^\dagger \mathbf{A}_k^\dagger(m) E\left\{\mathbf{y}_k(m)|\mathbf{c},\mathbf{r},\mathbf{s}^{(i)}\right\}|\mathbf{r},\mathbf{s}^{(i)}\right\}. \quad (25)$$

The conditional distribution of $\mathbf{y}_k(m)$ given $\mathbf{r}$, $\mathbf{c}$ and $\mathbf{s}=\mathbf{s}^{(i)}$ is Gaussian with mean

$$E\left(\mathbf{y}_k(m)|\mathbf{r},\mathbf{c},\mathbf{s}^{(i)}\right) = s^{(i)}(m,k)\mathbf{A}_k(m)\mathbf{c} + \left(\mathbf{r}(m) - \sum_{q=0}^{K-1} s^{(i)}(m,q)\mathbf{A}_q(m)\mathbf{c}\right) \quad (26)$$

where $\mathbf{s}^{(i)}$ is the estimated value of the signal $\mathbf{s}$ at the ith iteration step and, if $m \in \{m_1, m_2, \cdots, m_P\}$, then $s^{(i)}(m,q)$ is taken equal to the corresponding pilot symbol. Inserting (26) in (25), and subsequently substituting (25) and (24) in (22), we can rewrite (22) as

$$Q_k\left(\mathbf{s}_k|\mathbf{s}^{(i)}\right) = \sum_{m=0}^{M-1}\left[\Re\left\{s^*(m,k)\Upsilon_k^{(i)}(m)\right\} - \frac{1}{2}|s(m,k)|^2 E\left\{\mathbf{c}^\dagger \mathbf{\Gamma}_{k,k}(m)\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right\}\right] \quad (27)$$

where

$$\Upsilon_k^{(i)}(m) = E\left\{\mathbf{c}^\dagger|\mathbf{r},\mathbf{s}^{(i)}\right\}\mathbf{A}_k^\dagger(m)\mathbf{r}(m) - \sum_{q=0,q\neq k}^{K-1} s^{(i)}(m,q) E\left\{\mathbf{c}^\dagger \mathbf{\Gamma}_{k,q}(m)\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right\} \quad (28)$$

and $\mathbf{\Gamma}_{k,q}(m) \triangleq \mathbf{A}_k^\dagger(m)\mathbf{A}_q(m)$.

The prior probability density function (pdf) of $\mathbf{c} = [\mathbf{c}_0^T, \mathbf{c}_1^T, \cdots, \mathbf{c}_{L-1}^T]^T$ is chosen as $\mathbf{c} \sim N(\mathbf{0}, \mathbf{\Sigma}_\mathbf{c}^{(0)})$. The covariance matrix of $\mathbf{c}$ can be determined as

$$\mathbf{\Sigma}_\mathbf{c}^{(0)} = \mathrm{diag}\left(\mathbf{R}_\mathbf{c}(0), \mathbf{R}_\mathbf{c}(1), \cdots, \mathbf{R}_\mathbf{c}(L-1)\right) \in \mathcal{C}^{LD\times LD}$$

where

$$\mathbf{R}_\mathbf{c}(l) = \mathbf{\Psi}^T \mathbf{R}_\mathbf{h}(l)\mathbf{\Psi} \in \mathcal{C}^{D\times D}.$$

$\mathbf{R}_\mathbf{h}(l)$, the covariance matrix of $\mathbf{h}_l$, can be obtained from (5) as

$$\mathbf{R}_\mathbf{h}(l) = \sigma_l^2 \begin{bmatrix} r(0) & r(1) & \cdots & r(MN_g-1) \\ r(1) & r(0) & \cdots & r(MN_g-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(MN_g-1) & r(MN_g-2) & \cdots & r(0) \end{bmatrix} \quad (29)$$

with $r(k) = J_0(2\pi f_d k T_s)$. However, for sufficiently large length MN_g of the observation frame, it can be shown (see Appendix B) that $\mathbf{R}_\mathbf{c}(l)$, the covariance matrix of $\mathbf{c}$, for each channel path $l = 0, 1, \cdots, L-1$ becomes diagonal as

$$\mathbf{R}_\mathbf{c}(l) = \sigma_l^2 \mathrm{diag}\left(\lambda(l,0), \lambda(l,1), \cdots, \lambda(l,D-1)\right) \quad (30)$$

where $\lambda(l,d) = S_h(l, d/2N_gMT_s)$, $d = 0, 1, \cdots, D-1$, and $S_h(\cdot,\cdot)$ is the channel's scattering function defined by the Fourier transform of $r(k)$. For Jakes' Doppler profile it is given by $S_h(l,f) = \sigma_l^2/\pi\sqrt{f_D^2 - f^2}$ for $|f| < f_D$, [33].

On the other hand, since $\mathbf{w} \sim N(\mathbf{0}, N_0\mathbf{I})$, using the observation equation for $\mathbf{r}$ in (16), we can write the conditional pdf of $\mathbf{c}$ given $\mathbf{r}$ and $\mathbf{s}^{(i)}$ as

$$p\left(\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right) \sim p\left(\mathbf{r}|\mathbf{c},\mathbf{s}^{(i)}\right)p(\mathbf{c}) \sim \exp\left(-\frac{1}{N_0}\left(\mathbf{r} - \mathbf{Z}_{\mathbf{s}^{(i)}}\mathbf{c}\right)^\dagger\left(\mathbf{r} - \mathbf{Z}_{\mathbf{s}^{(i)}}\mathbf{c}\right) - \mathbf{c}^\dagger\left(\mathbf{\Sigma}_\mathbf{c}^{(0)}\right)^{-1}\mathbf{c}\right). \quad (31)$$

After some algebra it can be shown that [34]

$$p\left(\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right) \sim N\left(\boldsymbol{\mu}_\mathbf{c}^{(i)}, \mathbf{\Sigma}_\mathbf{c}^{(i)}\right)$$

where

$$\boldsymbol{\mu}_\mathbf{c}^{(i)} = \frac{1}{N_0}\mathbf{\Sigma}_\mathbf{c}^{(i)}\mathbf{Z}_{\mathbf{s}^{(i)}}^\dagger\mathbf{r} \quad (32)$$

and

$$\mathbf{\Sigma}_\mathbf{c}^{(i)} = \left(\left(\mathbf{\Sigma}_\mathbf{c}^{(0)}\right)^{-1} + \frac{1}{N_0}\mathbf{Z}_{\mathbf{s}^{(i)}}^\dagger\mathbf{Z}_{\mathbf{s}^{(i)}}\right)^{-1}. \quad (33)$$

Now let us compute the expectations on the right hand side of (28). The first expectation can be computed from (32) as follows:

$$E\left\{\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right\} = \boldsymbol{\mu}_\mathbf{c}^{(i)}. \quad (34)$$

By defining $\boldsymbol{\xi}_k(m) \triangleq \mathbf{A}_k(m)\mathbf{c}$, the second expectation in (28) can be determined as follows:

$$E\left\{\mathbf{c}^\dagger \mathbf{\Gamma}_{k,q}(m)\mathbf{c}|\mathbf{r},\mathbf{s}^{(i)}\right\} = \mathrm{tr}\left(\mathbf{\Xi}_{k,q}^{(i)}(m)\right) \quad (35)$$

where $\mathbf{\Xi}_{k,q}^{(i)}(m) \triangleq E\{\boldsymbol{\xi}_q(m)\boldsymbol{\xi}_k^\dagger(m)|\mathbf{r},\mathbf{s}^{(i)}\} = \mathbf{A}_q(m)(\mathbf{\Sigma}_\mathbf{c}^{(i)} + \boldsymbol{\mu}_\mathbf{c}^{(i)}(\boldsymbol{\mu}_\mathbf{c}^{(i)})^\dagger)\mathbf{A}_k^\dagger(m)$, and $\mathrm{tr}(\cdot)$ denotes the trace of its argument.

In summary of the above, the $\Upsilon_k^{(i)}(m)$ function in (28) can be expressed as

$$\Upsilon_k^{(i)}(m) = \left(\boldsymbol{\mu}_{\mathbf{c}}^{(i)}\right)^{\dagger} \mathbf{A}_k^{\dagger}(m)\mathbf{r}(m) - \sum_{q=0,q\neq k}^{K-1} s^{(i)}(m,q)\mathrm{tr}\left(\boldsymbol{\Xi}_{k,q}^{(i)}(m)\right). \quad (36)$$

B. Maximization-Step (M-Step)

In the M-step of the SAGE algorithm, the estimates of the data sequence are updated at the $(i+1)$th iteration according to

$$\mathbf{s}_k^{(i+1)} = \arg\max_{\mathbf{s}_k} Q_k\left(\mathbf{s}_k|\mathbf{s}^{(i)}\right), \quad \text{and}$$
$$\overline{\mathbf{s}}_k^{(i+1)} = \overline{\mathbf{s}}_k^{(i)}$$

where $Q(\mathbf{s}_k|\mathbf{s}^{(i)})$ is given by (27).

Keeping in mind (35) and substituting (27) into the above equation yields the following for $\mathbf{s}_k^{(i+1)}$:

$$s_k^{(i+1)} = \arg\max_{\mathbf{s}_k} \sum_{m=0}^{M-1} \left[\Re\left\{s^*(m,k)\Upsilon_k^{(i)}(m)\right\} - \frac{1}{2}|s(m,k)|^2 \mathrm{tr}\left(\boldsymbol{\Xi}_{k,k}^{(i)}(m)\right)\right], \text{ and}$$
$$\overline{\mathbf{s}}_k^{(i+1)} = \overline{\mathbf{s}}_k^{(i)} \quad (37)$$

where the maximization should be taken over the nonpilot data symbols $s(m,k)$ for $m \in \{0,1,\cdots,M-1\}$ and $m \neq m_1, m_2, \cdots, m_P$.

Moreover, when no coding is used, it follows from (37) that each component of $s^{(i+1)}(m,k)$ can be separately obtained in the continuous-domain by maximizing the corresponding summation in the right-hand side of (37) as

$$\tilde{s}^{(i+1)}(m,k) = \frac{\Upsilon_k^{(i)}(m)}{\mathrm{tr}\left(\boldsymbol{\Xi}_{k,k}^{(i)}(m)\right)}.$$

However, since $s(m,k)$ is discrete, belonging to a signal constellation point, we must quantize $\tilde{s}^{(i)}(m,k)$ to its nearest constellation point in each iteration. Consequently, substituting (36) into the above equation, the data update rule of the SAGE algorithm takes the following form:

$$s^{(i+1)}(m,k) = \mathrm{Quant}\left(\frac{\boldsymbol{\mu}_{\mathbf{c}}^{(i)\dagger}\mathbf{A}_k^{\dagger}(m)\mathbf{r}(m) - \sum_{q=0,q\neq k}^{K-1} s^{(i)}(m,q)\mathrm{tr}\left(\boldsymbol{\Xi}_{k,q}^{(i)}(m)\right)}{\mathrm{tr}\left(\boldsymbol{\Xi}_{k,k}^{(i)}(m)\right)}\right) \quad (38)$$

where $\mathrm{Quant}(\cdot)$ denotes the quantization process that quantizes its argument to its nearest data symbol constellation point.

Note that, as shown in Fig. 1, (36) can be interpreted as joint equalization and ICI cancelation performed right after the channel estimation step, implemented in the time-domain, immediately following the analog-to-digital (A/D) conversion and cyclic prefix deletion processes at the OFDM receiver. Consequently, we can think of the quantities $\Upsilon_k^{(i)}(m)$ in (36) as the outputs of an ICI canceler, generated at the ith iteration step of the SAGE algorithm. After the SAGE algorithm converges at some iteration step, say the Ith step, the original data is detected by conventional coherent detection in the frequency domain using the interference-cleaned signal vector $\mathbf{y} = [y(0), y(1), \cdots, y(N-1)]^T$, where $y(k) \triangleq \Upsilon^{(I)}(k)$.

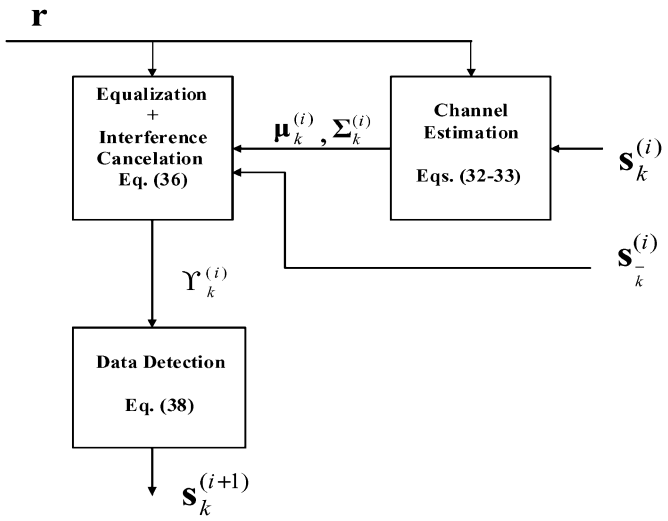


Fig. 1. Channel estimation, equalization, interference cancelation and data detection performed for the kth subcarrier at the ith iteration step.

C. Initialization

The performance of the receiver operating with the data detection and channel estimation algorithm proposed above is closely related to the quality of the initial values of $\mathbf{c}^0$ and the data $\mathbf{s}^{(0)}$, which are computed as follows.

1) Initialization of Channel Coefficients: The initial channel estimate, $\mathbf{c}^{(0)}$, can be determined with the aid of the pilot symbols. As mentioned in Section II-C, we consider all subcarriers in a given time slot dedicated to pilot symbols. Assuming that there are P pilot OFDM symbols located at time-slots $m_1, m_2, \cdots, m_P$ where $m_p \in \{0,1,\cdots,M-1\}$, they can be arranged in an $NP \times LD(NP > LD)$ pilot-symbol matrix $\mathbf{Z}_{\mathbf{s}_P} = [\mathbf{Z}_{\mathbf{s}}^T(m_1), \mathbf{Z}_{\mathbf{s}}^T(m_2), \cdots, \mathbf{Z}_{\mathbf{s}}^T(m_P)]^T$ where $\mathbf{Z}_{\mathbf{s}}(m_p) = \sum_{q=0}^{K-1} s(m_p,q)\mathbf{A}_q(m_p)$. Accordingly, from (16) the received vector corresponding to the one OFDM frame of length M can be expressed as

$$\mathbf{r}_P = \mathbf{Z}_{\mathbf{s}_P}\mathbf{c} + \mathbf{w}_P. \quad (39)$$

The MMSE estimate of the initial the channel parameter $\mathbf{c}^{(0)}$ can be determined from (39) as follows:

$$\mathbf{c}^{(0)} \equiv \widehat{\mathbf{c}}_{\mathrm{MMSE}} = \left(\mathbf{Z}_{\mathbf{s}_P}^{\dagger}\mathbf{Z}_{\mathbf{s}_P} + N_0\left(\boldsymbol{\Sigma}_{\mathbf{c}}^{(0)}\right)^{-1}\right)^{-1}\mathbf{Z}_{\mathbf{s}_P}^{\dagger}\mathbf{r}_P \quad (40)$$

where N_0 is the variance of the Gaussian noise.

2) Initialization of Data Symbols: The initial value of the data symbols $\mathbf{s}^{(0)}(m)$ for each $m \in \{0,1,\cdots,M-1\}$, $m \neq m_1, m_2, \cdots, m_P$, can be determined as follows. From (2) and (14) the received signal $\mathbf{r}(m)$ can be expressed as

$$\mathbf{r}(m) = \mathbf{H}^{(0)}(m)\mathbf{s}^{(0)}(m) + \mathbf{w}(m) \tag{41}$$

where $\mathbf{r}^{(0)}(m)$ and $\mathbf{w}^{(0)}(m)$ are defined in (2) and

$$\begin{aligned}\mathbf{s}^{(0)}(m) &= \left[s^{(0)}(m,0), s^{(0)}(m,1), \cdots, s^{(0)}(m,K-1)\right]^T \\ \mathbf{H}^{(0)}(m) &= \left[\mathbf{H}^{(0)}(m,0), \mathbf{H}^{(0)}(m,1)\cdots, \mathbf{H}^{(0)}(m,K-1)\right] \\ \mathbf{H}^{(0)}(m,q) &= \mathbf{V}_q \mathbf{h}^{(0)}(m)\end{aligned} \tag{42}$$

and

$$\mathbf{h}^{(0)}(m) = \mathbf{\Psi}^{\dagger}(m)\mathbf{c}^{(0)}.$$

Note that from (3) and (42) it follows that

$$\mathbf{H}^{(0)}(m,k) = \frac{1}{N}\left[H_0^{(0)}(m,k), H_1^{(0)}(m,k)e^{j\frac{2\pi}{N}}, \cdots, H_{N-1}^{(0)}(m,k)e^{j\frac{2\pi(N-1)}{N}}\right]^T$$

where $H_p^{(0)}(m,k)$ represents the time-varying channel coefficient in the frequency domain at the discrete time $t = mN_g + p$ and at the discrete frequency k.

In order to detect the initial data symbols $s^{(0)}(m)$, $m \in \{0,1,\cdots,M-1\}$, $m \neq m_1, m_2, \cdots, m_P$, several options are possible. We focus on a low complexity linear MMSE detection approach here. We omit the time-index m from now on for notational simplicity since the detection process is the same for each m. From the observation (41), the linear MMSE detection of data can be expressed as

$$\mathbf{s}^{(0)} \equiv \widehat{\mathbf{s}}_{\mathrm{MMSE}} = \mathbf{H}^{(0)\dagger}\left(\mathbf{H}^{(0)}\mathbf{H}^{(0)\dagger} + N_0\mathbf{I}_N\right)^{-1}\mathbf{r}. \tag{43}$$

The matrix inversion in (43) requires $O(N^3)$ floating point operations (flops), which is computationally very expensive when N is large such as for the IEEE 802.11 and IEEE 802.16 families. However, as is known, time-varying channels produce a nearly-banded channel matrix whose only main diagonal, Q subdiagonal and Q superdiagonals are nonzero. The bandwidth $2Q$ is a parameter to be adjusted according to the mobility-rate of the channel. It has been determined that $Q = \lfloor f_d N \rfloor + 1$ is an appropriate choice for Rayleigh fading [30]. The banded property of the channel can be exploited to reduce the computational complexity by means of low complexity decompositions such as the Cholesky or $\mathrm{LL}^{\dagger}$ factorization [35] of Hermitian banded matrices.

Since $\mathbf{T} = \mathbf{H}^{(0)}\mathbf{H}^{(0)\dagger} + N_0\mathbf{I}_N$ in (43) is a Hermitian banded matrix with lower and upper bandwidth $2Q$, we choose the $\mathrm{LL}^{\dagger}$ factorization to obtain $\mathbf{T}^{-1}$ which has the advantage of not requiring the computation of a square root. The basic steps of the data detection are given later as presented in [15]

- STEP-1: Construct the banded matrix $\mathbf{T} = \mathbf{H}^{(0)}\mathbf{H}^{(0)\dagger} + N_0\mathbf{I}_N$
- STEP-2: Noting that $\mathbf{T}$ is a positive definite matrix, perform the $\mathrm{LL}^{\dagger}$ Cholesky factorization of $\mathbf{T}$, as expressed by $\mathbf{T} = \mathbf{L}\mathbf{L}^{\dagger}$, where the triangular factor $\mathbf{L}$ has lower bandwidth $2Q$;
- STEP-3: Solve the system $\mathbf{T}\mathbf{d} = \mathbf{r}$ for $\mathbf{d}$ by the following steps:
  - (i) solve the triangular system $\mathbf{L}\mathbf{f} = \mathbf{r}$ for $\mathbf{f}$;
  - (ii) solve the triangular system $\mathbf{L}^{\dagger}\mathbf{d} = \mathbf{f}$ for $\mathbf{d}$.
- STEP-4: Calculate $\widehat{\mathbf{s}}_{\mathrm{MMSE}} = \mathbf{H}^{(0)\dagger}\mathbf{d}$

D. Computational Complexity

In this subsection we evaluate the computational cost of the proposed algorithm in terms of complex multiplications (CMs) and complex additions (CAs) that are required per each data symbol to be detected. The computational complexity of the algorithm is determined by the parameters N, K D, L, M, and Q. The particular structures of the diagonal and Hermitian as well as banded matrices have been exploited to keep the number of computations to a minimum. Let us first determine the number of computations required to compute the quantities $\mathbf{A}_k(m)$ and $\mathrm{tr}(\mathbf{\Xi}_{k,q}^{(i)}(m))$ in (38) as follows.

Keeping in mind $\mathbf{A}_k(m) = \mathbf{V}_k\mathbf{\Phi}(m) \in \mathcal{C}^{N\times DL}$ and $\mathbf{\Phi}(m) = \mathbf{I}_L \otimes \mathbf{\Psi}(m)$ with $\mathbf{\Psi}(m) \in \mathcal{C}^{N\times D}$ we can show that

$$\begin{aligned}\mathbf{A}_k(m) &= \left(\mathbf{F}_L^T(k) \otimes \frac{1}{N}\mathrm{diag}\left(\mathbf{F}_N^{\dagger}(k)\right)\right)(\mathbf{I}_L \otimes \mathbf{\Psi}(m)) \\ &= \mathbf{F}_L^T(k) \otimes \left(\frac{1}{N}\mathrm{diag}\left(\mathbf{F}_N^{\dagger}(k)\right)\mathbf{\Psi}(m)\right) \\ &= \mathbf{F}_L^T(k) \otimes \left(\underbrace{\left[\frac{1}{N}\mathbf{F}_N^*(k), \cdots, \frac{1}{N}\mathbf{F}_N^*(k)\right]}_{N\times D} \odot \mathbf{\Psi}(m)\right)\end{aligned}$$

where $\mathbf{F}_L(k)$ is defined by (3). The notations $\otimes$ and $\odot$ denote Kronecker and element-by-element products, respectively. Note that $[(1/N)\mathbf{F}_N^*(k), \cdots, (1/N)\mathbf{F}_N^*(k)] \odot \mathbf{\Psi}(m)$ requires ND complex multiplications. Consecutively, it is easily seen that the computational load of $\mathbf{A}_k(m)$ is $ND(L+1)$ CM. Note that for $k = 0,1,\cdots,K-1$ and $m = 0,1,\cdots,M-1$, the $\mathbf{A}_k(m)$'s can be precomputed and used throughout the operation of the system. Therefore, it will not be taken into account in the complexity calculations.

Similarly, to compute $\mathrm{tr}(\mathbf{\Xi}_{k,q}^{(i)}(m))$, where $\mathbf{\Xi}_{k,q}^{(i)}(m) = \mathbf{A}_q(m)(\mathbf{\Sigma}_{\mathbf{c}}^{(i)} + \boldsymbol{\mu}_{\mathbf{c}}^{(i)}(\boldsymbol{\mu}_{\mathbf{c}}^{(i)})^{\dagger})\mathbf{A}_k^{\dagger}(m)$, consider vectors $\mathbf{u}_j^T$ and $\mathbf{v}_j$ of length DL which are the jth row of $\mathbf{A}_q(m)$ and the jth column of $\mathbf{A}_k^{\dagger}(m)$, respectively. It then follows that

$$\mathrm{tr}\left(\mathbf{\Xi}_{k,q}^{(i)}(m)\right) = \sum_{j=1}^{N}\left(\mathbf{u}_j^T\mathbf{\Sigma}_{\mathbf{c}}^{(i)}\mathbf{v}_j + \mathbf{u}_j^T\boldsymbol{\mu}_{\mathbf{c}}^{(i)}\left(\boldsymbol{\mu}_{\mathbf{c}}^{(i)}\right)^{\dagger}\mathbf{v}_j\right).$$

Since the reduced-dimensional channel covariance matrix $\mathbf{\Sigma}_{\mathbf{c}}^{(i)}$ is diagonal, it is easy to see that there are $2DL$ CMs and DL CAs in calculation of $\mathbf{u}_j^T\mathbf{\Sigma}_{\mathbf{c}}^{(i)}\mathbf{v}_j$. Similarly, $2DL$ CMs and CAs are needed to calculate $\mathbf{u}_j^T\boldsymbol{\mu}_{\mathbf{c}}^{(i)}(\boldsymbol{\mu}_{\mathbf{c}}^{(i)})^{\dagger}\mathbf{v}_j$. Therefore, the total computational load of the trace operation is roughly $4N(4DL+1)$ CMs and $N(3DL+2)$ CAs leading to a total of $N(7DL+3)$ complex operations.

By means of the above results we can evaluate the total number of multiplications and additions required to

implement the SAGE algorithm in (38) as follows. Assuming that the SAGE algorithm converges in $I_{\rm iter}$ iterations, the computations needed to detect each data symbol $s(k,m)$ in (38) is $I_{\rm iter}(9NDL + K + DL - 1)$ CMs and $I_{\rm iter}(N(7DL + 3) + K + DL - 1)$ CAs which is roughly $O(16I_{\rm iter}DLN)$ complex operations for nonconstant envelope signal constellations, since $N \gg DL$.

The computational cost of the initialization step of the algorithm can be determined as follows. To compute the initial MMSE estimate of the reduced dimensional channel vector $\mathbf{c}^{(0)}$, given by (40), $NLPD/MK \approx LPD/M$ CMs and CAs are necessary per detected data symbol, since $N \approx K$ and the term $(\mathbf{Z}_{\mathbf{s}_P}^{\dagger}\mathbf{Z}_{\mathbf{s}_P} + N_0(\mathbf{\Sigma}_{\mathbf{c}}^{(0)})^{-1})^{-1}\mathbf{Z}_{\mathbf{s}_P}^{\dagger}$ in (40) can be precomputed and need not be recomputed again during the detection stage.

The computational load of the low-complexity linear MMSE algorithm to compute the initial data $\mathbf{s}^{(0)}$ in (43) can be determined as follows. Taking into account that $\mathbf{H}^{(0)}\mathbf{H}^{(0)\dagger}$ is nearly banded with bandwidth $4Q$, the computation of $\mathbf{T} = \mathbf{H}^{(0)}\mathbf{H}^{(0)\dagger} + N_0\mathbf{I}_N$ in Step-1, requires $KN(2Q+1)$ CMs and CAs. For Step-2, it was shown in [15] that computation of $\mathrm{LL}^{\dagger}$ factorization algorithm requires $(2Q^2 + 3Q)N$ CMs and $(2Q^2 + Q)N$ CAs. Furthermore, solving each of the two banded triangular equation systems of Step-3 requires $4QN$ CMs and CAs. Finally, the last step needs NK CMs and CAs to calculate $\mathbf{s}^{(0)} = \mathbf{H}^{(0)\dagger}\mathbf{d}$. Therefore, the entire initialization algorithm requires roughly $2N(Q+1) + LPD/M$ CMs and $2N(Q+1) + LPD/M$ CAs per detected symbol leading to a total of $4N(Q+1) \sim O(4NQ)$ complex operations, since $K \approx N$, $N \gg Q$ and $NQ \gg LPD/M$. Consequently, the total computational complexity to implement our equalization/detection algorithm per detected data symbol is $O(N(4Q + 16I_{\rm iter}QDL))$.

On the other hand, a number of suboptimal equalizers/detectors have been proposed to reduce complexity, such as MMSE detectors [17] and BLAST detectors [36]. However all these schemes perform far from the maximum likelihood (ML) detector. The VBLAST equalization/detection algorithm yields substantially better error performance than conventional linear equalizer/detectors. Therefore, we would like to compare the computational complexity of our algorithm with that of VBLAST.

The VBLAST algorithm performs four operations: nulling, slicing, canceling, and ordering [36]. Estimation of the symbols is done at the slicing stage and a new received signal is calculated by subtracting the estimated symbols. If we assume that the receiver has the complete channel state information and we do not count slicing operation that corresponds to demodulation, the nulling and canceling processes require $2N$ CMs and $2N-1$ CAs for each iteration and thus for all iterations requires a total of $(4N-1)N \sim O(N^2)$ complex operations. Also, ordering of the symbols require $N \times N$ matrix inversion and sorting in each iteration loop. The repeated inverse matrix computation in the ordering step is the main computational bottleneck of the algorithm and it requires total a total of $O(N^4)$ complex operations for detection of the N data symbols. However by using the banded structure of the channel matrix this complexity can be reduced to $O(Q^2N^2)$, using Rugini's approach [15]. In addition to this, computation of the signal-to-interference-plus-noise ratio $(SINR_k)$ at each iteration step k, requires $N^2 + 4N + 2$ CMs and $N^2 - 2N + 3$ CAs leading to a total of $O(N^3)$ complex operations. As a result, the complete VBLAST algorithm needs approximately $(O(N^3) + O(N^2Q^2) + O(N^2))/N \sim O(N^2)$ complex operations per detected data symbol. Therefore, the computational complexity of the VBLAST receiver is substantially greater than that of our algorithm, (i.e., $O(N^2)$ versus $O(N)$ per data symbol) even under the assumption that the channel is known by the receiver, and increases rapidly with the number of subcarriers which makes its real-time implementation prohibitive for OFDM based WiMAX and LTE types of systems.

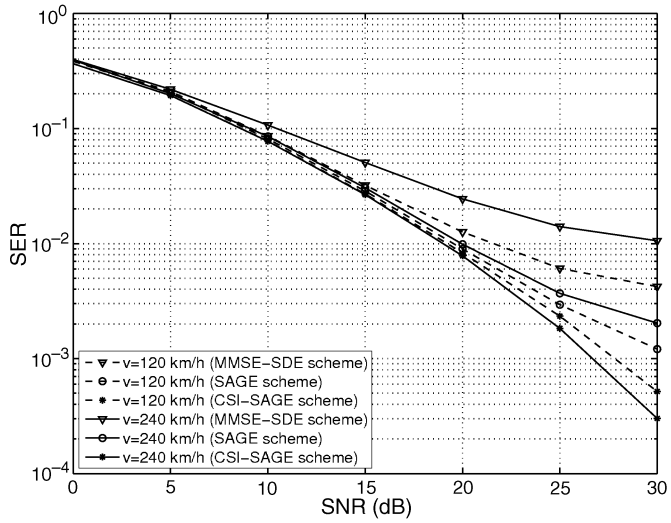


Fig. 2. SER versus SNR simulation results for different detection schemes: $N = K = 1024$, $L = 3$, $M = 50$, QPSK signaling.

Based on the above discussions on computational complexities, we conclude that the complexity of our algorithm is much smaller than that of the VBLAST algorithm. On the other hand, our algorithm has the potential to achieve better performance than the VBLAST since it is known that the SAGE algorithm achieves the optimal ML detection performance if it converges.

IV. Computer Simulations

In this section, we present simulation results to assess the performance of OFDM systems based on the proposed receiver. The system operates over a 10-MHz bandwidth with 1024 subchannels operating on a carrier frequency of 2.5 GHz and with quadrature phase shift keying (QPSK) signaling. A multipath wireless channel having an exponentially decaying power delay profile with the normalized powers [19], $\sigma_0^2 = 0.448$, $\sigma_1^2 = 0.321$, and $\sigma_2^2 = 0.230$ is chosen.

In Fig. 2, the SER performance of the proposed algorithm is presented as a function of SNR for the normalized Doppler frequencies $f_DT = 0.0284$ and $f_DT = 0.0569$ corresponding to a mobile terminal moving at speeds of 120 km/h and 240 km/h, respectively, and for the OFDM frame length $M = 50$. The lower bounds D_{lower} for the dimension of the DCT basis expansion coefficients, the chosen dimensions D, corresponding to the channel modeling $\mathrm{MSE} = 10^{-3}$ [defined in (10)], as well as the pilot spacings are given in Table I.

TABLE I
SIMULATION PARAMETERS

v (km/h)	D_{lower}	D chosen for MSE=10^{-3}	Pilot Spacing
120	4	7	8
240	7	11	7

Our extensive computer simulations have shown that the pilot spacings (PSs) chosen according to $\mathrm{PS} = M/D_{lower}$ resulted in the best SER performance. The solid and the dashed curves in Fig. 2 represent the SER performance curves when we have perfect channel state information (CSI) and for the proposed SAGE algorithm including channel estimation, corresponding to mobilities $f_DT = 0.0284$ ($v = 120$ km/h) and $f_DT = 0.0569$ ($v = 240$ km/h). Note that as the velocity increases, the rapidly varying channel not only destroys the orthogonality but also provides the receiver with time-diversity. As can be seen from (2), the data $s(m,k)$ is carried on N components of the channel vector $\mathbf{h}_k = \mathbf{V}_k\mathbf{h}$ in one symbol duration. When the channel is perfectly known (CSI is available), the CSI-SAGE scheme is able to make good use of the time diversity. Consequently, in Fig. 2, the performance for the case with $v = 240$ km/h (solid line) is better than for the case with $v = 120$ km/h (dashed line). However, when the channel CSI is not available, the gain from diversity may be overwhelmed by the channel estimation errors. This effect is clearly seen from the performance curves in Fig. 2 which show that the SER performance of the SAGE detection scheme for $v = 120$ km/h is better than the case of $v = 240$ km/h when the CSI is obtained through channel estimation.

In Fig. 3, the SER is plotted as a function of the number of SAGE iterations for several SNR values and for 16-ary quadrature amplitude modulation (16QAM) modulation. It can be seen that three to four iterations are sufficient in order for the SAGE algorithm to converge. The initial estimates of the channel and the transmitted data are performed by the reduced-complexity MMSE estimation techniques based on the pilot symbols as explained at the end of Section III. We refer to this method for obtaining the initial channel and data estimates as the MMSE separate detection and estimation (MMSE-SDE) scheme. We conclude from these curves that even when the number of DCT coefficients is chosen to be fairly small as compared to the total number of coefficients NM, the performance loss in SER is not significant when CSI is not available. We also can see that the SER performance of the SAGE algorithm obtained at the end of the third iteration step is much lower than that of the MMSE-SDE.

We also investigated the average MSE performance of the channel estimation as part of our algorithm. The average MSE, here, is defined as the combination of the channel estimation error with the channel modeling error due to the truncation of the DCT expansion coefficients. In Fig. 4, the average MSE curves are plotted for several SNR values employing the same parameter values as given in Table I. As can be seen in Fig. 3, the algorithm achieves excellent MSE performance even when the number of DCT coefficients is truncated at 11.

The effects of channel estimation on the average MSE and on the SER performance are investigated as functions of the pilot

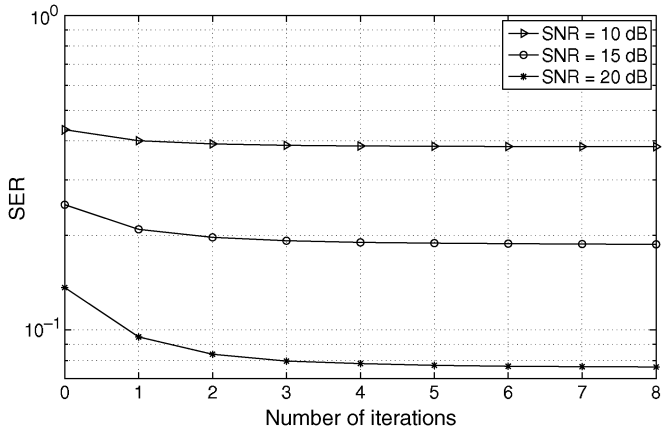


Fig. 3. SER versus Number of Iterations for 16QAM signaling scheme, $f_DT = 0.0569$ ($v = 240$ km/h), $N = K = 1024$, $L = 3$ and $M = 50$.

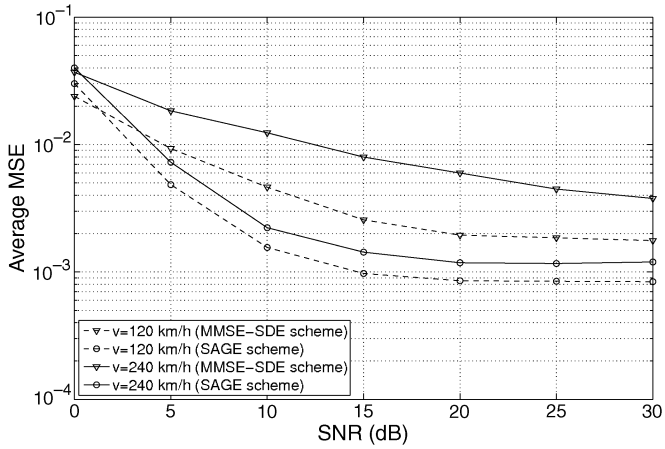


Fig. 4. Average MSE versus SNR simulation results for different detection schemes: $N = K = 1024$, $L = 3$, $M = 50$, QPSK signaling.

spacing (Δ) with different numbers of DCT truncation coefficients (D). The results are shown in Fig. 5 and Fig. 6, respectively, when the normalized Doppler frequency $f_DT = 0.0569$ ($v = 240$ km/h), SNR = 30 dB and the OFDM frame length is chosen as $M = 50$. Our extensive computer simulations have shown that the average MSE performance depends heavily on the pilot spacings and number of pilot symbols employed within an OFDM observation frame. During the preliminary simulations we always placed the first pilot tone at the beginning of the frame and the others equally spaced over the frame, and observed that the MSE performance degrades significantly if the last pilot symbol cannot be placed at the end of the frame as result of the equally spaced placement. Based on these observations, we have modified our computer simulations in such a way that the first and the last pilots are always placed at the beginning and at the end of the OFDM frame. The rest of the pilots are placed almost equally spaced over the OFDM frame. From Fig. 5 it is observed that the performance stays almost the constant up until pilot spacing 12, and then degrades substantially from pilot spacing 12 to 16. This is mainly due to the fact that the number of pilot symbols drops from five to four at this point.

Similarly, it is concluded from Fig. 6 that the SER performance does not change significantly for pilot spacings up to

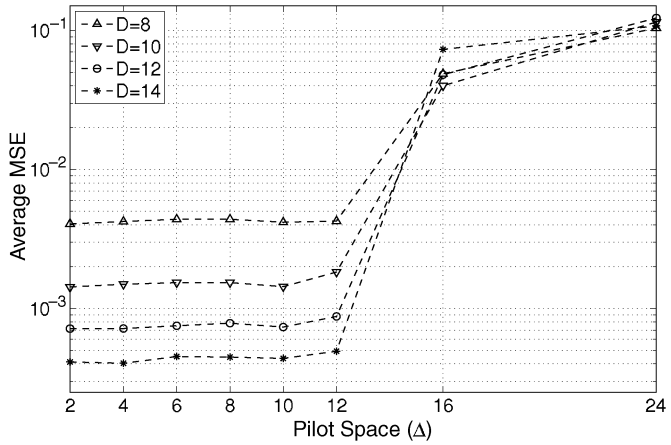


Fig. 5. Average MSE versus Pilot Spacing (Δ) for $f_DT = 0.0569$ ($v = 240\,\mathrm{km/h}$) and $M = 50$, QPSK signaling.

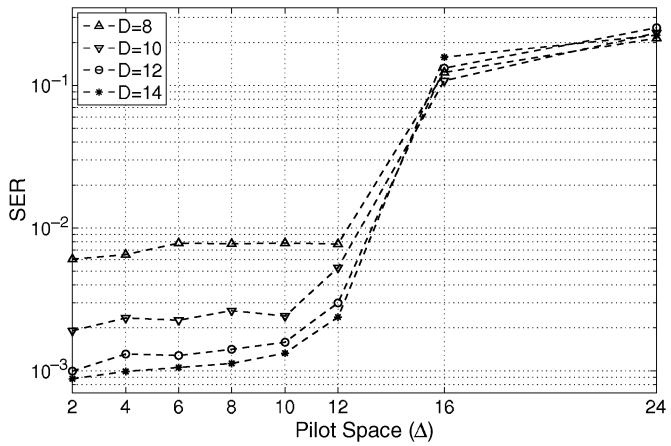


Fig. 6. SER versus Pilot Spacing (Δ) for $f_DT = 0.0569$ ($v = 240\,\mathrm{km/h}$), $N = K = 1024$, $L = 3$ and $M = 50$, QPSK signaling.

a certain value determined by the Doppler frequency. For example, from Fig. 6 it is seen that the best SER performance ($\approx 5.5{\times}10^{-3}$) is achieved when $\Delta = 12$ and $D = 8$ for the normalized Doppler frequency $f_DT = 0.0569$ ($v = 240$ km/h). Whereas, SER $= 1.7 \times 10^{-3}$ is achieved for $\Delta = 10$ and $D = 12$. Consequently, the best pilot spacings and number of pilots can be determined from Figs. 5 and 6 for the given SER performance values based on the optimal pilot placement described above.

Finally, we compare the SER performance of our algorithm with separate data detection, equalization and channel estimation techniques. Fig. 7 presents the SER as a function of SNR for the matched filter (MF), MMSE-SDE and the linear MMSE successive detection with optimal ordering and channel estimation (MMSE-SucDE) proposed in [19], and the joint SAGE detection and channel estimation, when the Doppler frequency $f_DT = 0.1$ ($v = 450$ km/h) and when 16-ary phase shift keying (16PSK) modulation is employed. For comparison of our results with the ones presented in [19], the same pilot spacing was chosen as $\Delta = 3$. The MF suffers from severe ICI. The MMSE-SDE is slightly worse than the MMSE-SucDE especially for SNR values greater than 15 dB. It is seen that our joint

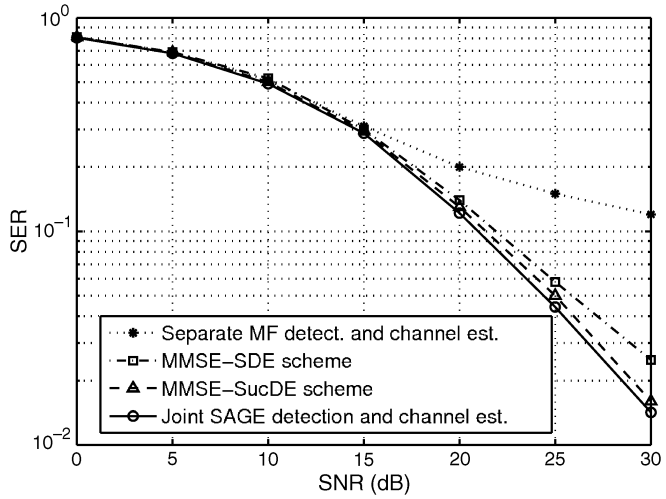


Fig. 7. SER comparison of different detection and channel estimation techniques $\Delta = 3$ for $f_DT = 0.1$ ($v = 450\,\mathrm{km/h}$), $N = K = 1024$, $L = 3$, and $M = 25$, 16PSK signaling.

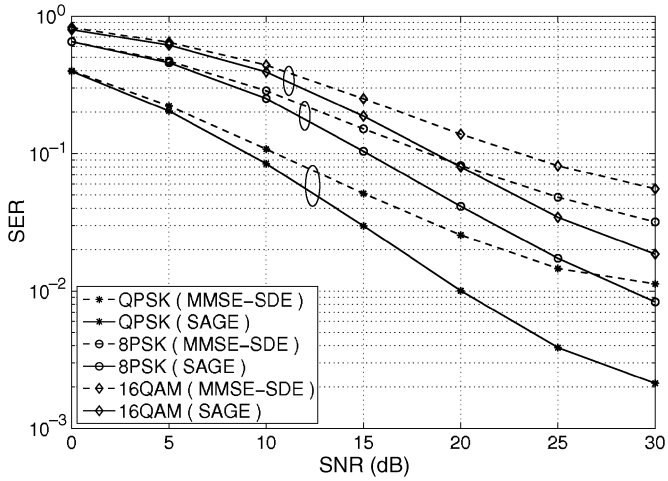


Fig. 8. SER comparison of QPSK, 8PSK and 16QAM signaling schemes $f_DT = 0.0569$ ($v = 240\,\mathrm{km/h}$), $N = K = 1024$, $L = 3$, and $M = 50$.

detection and channel estimation algorithm has the best performance, and that it is slightly better than the MMSE-SucDE scheme. Note that since the number of subcarriers N is usually very large, e.g., $N = 1024$ in WiMAX and LTE systems, even only the linear MMSE-based equalization-detection part of the scheme proposed in [19] demands very high computation load, and it may not be feasible in a practical system.

Finally, in Fig. 8, the SER performances of our algorithm are compared with that of the MMSE-SDE for QPSK, 8PSK and 16QAM signalling schemes with parameters $f_DT = 0.0569$ ($v = 240$ km/h), $N = K = 1024$, $L = 3$, and $M = 50$. The performance curves shown in Fig. 8 indicate that the SAGE algorithm clearly outperforms the MMSE-SDE technique also for different type of modulation formats.

V. Conclusion

The problem of joint channel estimation, equalization and data detection for OFDM systems operating in the presence of frequency selective and very rapidly time-varying channels has

been investigated in this paper. We have presented an iterative approach based on the SAGE algorithm and closed form expressions have been derived for data detection which incorporate channel estimation as well as partial interference cancelation. The cosine orthogonal basis functions have been applied to describe the time-varying channel. It has been shown by computer simulations that, depending on the normalized Doppler frequency, only a small number of expansion coefficients is sufficient to approximate the channel perfectly and that there is no need to know the statistics of the input signal. The proposed algorithm has excellent symbol error rate and channel estimation performance even with a very small number of channel expansion coefficients, resulting in substantial reduction of the computational complexity.

Appendix A
Derivation of (2)

Recalling $\mathbf{h}_l(m)$ from (4), the received signal in (1) can be expressed in vector form as follows:

$$\mathbf{r}(m)=\sum_{q=0}^{K-1} s(m,q)\mathrm{diag}\left(\mathbf{F}_N^{\dagger}(q)\right)\frac{1}{N} \times\sum_{l=0}^{L-1}\mathbf{h}_l(m)\exp\left(-j\frac{2\pi ql}{N}\right)+\mathbf{w}(m) \quad (44)$$

where $\otimes$ denotes the Kronecker product and

$$\mathbf{F}_L(q)\triangleq\left[1,\exp\left(-j\frac{2\pi q}{N}\right),\cdots,\exp\left(-j\frac{2\pi q(L-1)}{N}\right)\right]^T\in\mathcal{C}^L. \quad (45)$$

Note that

$$\sum_{l=0}^{L-1}\mathbf{h}_l(m)\exp(-j2\pi ql/N)=\left(\mathbf{F}_L^T(q)\otimes\mathbf{I}_N\right)\mathbf{h}(m) \quad (46)$$

where $\mathbf{h}(m)$ is given in (4). Finally, substituting (46) into (44) we obtain

$$\mathbf{r}(m)=\sum_{q=0}^{K-1}s(m,q)\underbrace{\left(\mathbf{F}_L^T(q)\otimes\frac{1}{N}\mathrm{diag}\left(\mathbf{F}_N^{\dagger}(q)\right)\right)}_{\mathbf{v}_q}\mathbf{h}(m)+\mathbf{w}(m) \quad (47)$$

from which (2) can be obtained.

Appendix B
Proof of Asymptotic Convergence of the Channel Covariance Matrix

The correlation between $c(d,l)$ and $c(d',l)$ can be computed from (8) as

$$E\left\{c(d,l)c^*(d',l)\right\}=\sum_{t=0}^{N_gM-1}\psi_d(t)\sum_{t'=0}^{N_gM-1}E\left\{h(t,l)h^*(t',l)\right\}\psi_{d'}(t') \quad (48)$$

where $E\{h(t,l)h^*(t',l)\}=R_h(t-t',l)$.

We now show that for sufficiently large block size M

$$\sum_{t'=0}^{N_gM-1}R_h(t-t',l)\psi_{d'}(t')\approx\lambda(d',l)\psi_{d'}(t). \quad (49)$$

Consequently, since $\psi_d(\cdot)$'s are orthonormal, it follows from (49) that

$$E\left\{c(d,l)c^*(d',l)\right\}=\begin{cases}\sigma_l^2\lambda(d,l) & \text{if } d=d'\\ 0 & \text{if } d\neq d'.\end{cases} \quad (50)$$

Using the definition of the cosine transform given by (9), the left-hand side of (49) can be expressed as

$$I\triangleq\sum_{t'=0}^{N_gM-1}R_h(t-t',l)\psi_{d'}(t') =\sqrt{\frac{2}{N_gM}}\sum_{t'=0}^{N_gM-1}R_h(t-t',l)\cos\left(\frac{\pi d'}{N_gM}\left(t'+\frac{1}{2}\right)\right), \text{ for } d>0. \quad (51)$$

Since $R_h(t,l)$ is the inverse Fourier transformation of $S_h(f,l)$, the channel's scattering function, it follows that

$$R_h(t,l)=T_s\int_{-1/2T_s}^{1/2T_s}S_h(f,l)\exp(j2\pi ftT_s)df, \quad l=0,1,\cdots,L-1. \quad (52)$$

On substituting (52) into (51) and after some algebra we have

$$I=T_s\sqrt{\frac{2}{N_gM}}\int_{-1/2T_s}^{1/2T_s}S_h(f,l)df \times\left(\frac{1}{2}e^{j\frac{\pi d'}{2N_gM}}\sum_{t'=0}^{N_gM-1}e^{j2\pi t'\left(\frac{d'}{2N_gM}-fT_s\right)} +\frac{1}{2}e^{-j\frac{\pi d'}{2N_gM}}\sum_{t'=0}^{N_gM-1}e^{-j2\pi t'\left(\frac{d'}{2N_gM}+fT_s\right)}\right). \quad (53)$$

It can be easily shown that as $MN_g\to\infty$ we can use the Poisson sum formula to obtain

$$\sum_{t'=0}^{N_gM-1}\exp\left(\pm j2\pi t'\left(\frac{d'}{2N_gM}-fT_s\right)\right)=\sum_{k=-\infty}^{+\infty}\delta\left(\frac{d'}{2N_gM}-fT_s-k\right) \quad (54)$$

where $\delta(\cdot)$ is the Dirac delta function. Substituting (54) into (53), and integrating the resulting integrals with respect to f we have

$$I=\sqrt{\frac{2}{N_gM}}S_h\left(\frac{\pi d'}{N_gMT_s}\right)\cos\left(\frac{\pi d'}{N_gM}\left(t+\frac{1}{2}\right)\right). \quad (55)$$

Finally, defining $\lambda(d',t)\triangleq S_h(\pi d'/N_gMT_s)$ and $\psi_{d-}(t)\triangleq\sqrt{2/N_gM}\cos((\pi d'/N_gM)(t+(1/2)))$, the claim in (50) is proved.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers of this manuscript for their valuable and constructive comments which helped to improve the paper considerably.

REFERENCES

[1] Y. G. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–914, Jul. 1998.

[2] O. Edfors, M. Sandell, J. van de Beek, S. Wilson, and P. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[3] Y. G. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, 2000.

[4] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[5] M.-H. Ng and S.-W. Cheung, "Bandwidth-efficient pilot-symbol-aided technique," *Electron. Lett.*, vol. 34, no. 16, pp. 1548–1550, 1998.

[6] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.

[7] H. Senol, H. A. Cirpan, E. Panayirci, and M. Cevik, "A low complexity time-domain MMSE channel estimator for space-time/frequency block coded OFDM systems," *EURASIP J. Appl. Signal Process.*, vol. 2006, no. 19, pp. 1–14, 2006, 10.1155/ASP/2006/39026, Article ID 39026.

[8] H. A. Cirpan, E. Panayirci, and H. Dogan, "Non-data-aided channel estimation for OFDM systems with space-frequency transmit diversity," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 449–457, Mar. 2006.

[9] H. Dogan, H. A. Cirpan, and E. Panayirci, "An efficient joint channel estimation and decoding algorithm for turbo-coded space-time orthogonal frequency division multiplexing receivers," *IET Commun.*, vol. 2, no. 7, pp. 886–894, Aug. 2008.

[10] H. Dogan, H. A. Cirpan, and E. Panayirci, "Iterative channel estimation and decoding of turbo coded SFBC-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3090–3101, Aug. 2007.

[11] I. Barhumi, G. Leus, and M. Moonen, "Equalization for OFDM over doubly-selective channels," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1445–1458, Apr. 2006.

[12] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2451–2464, Oct. 2002.

[13] X. Huang and H.-C. Wu, "Robust and efficient intercarrier interference mitigation for OFDM systems in time-varying fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2517–2528, Sep. 2007.

[14] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 27–32, Jan. 1999.

[15] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Commun. Lett.*, vol. 9, no. 7, pp. 619–621, Jul. 2005.

[16] L. Rugini, P. Banelli, and G. Leus, "Low-complexity banded equalizers for OFDM systems in Doppler spread channels," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–13, 2006, no. Article ID 67404.

[17] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1002–1011, Apr. 2004.

[18] A. Gorokhov and J. P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: Equalization and channel acquisition," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 572–583, Apr. 2004.

[19] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.

[20] X. Cai and G. B. Giannakis, "Bounding performance and suppressing intercarrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2047–2056, Dec. 2003.

[21] S. Tomasin, A. Gorokhov, H. Yang, and J.-P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 238–245, Jan. 2005.

[22] S.-J. Hwang and P. Schniter, "Efficient sequence detection of multicarrier transmissions over doubly dispersive channels," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–17, 2006, Article ID 93638.

[23] W. Song and J. Lim, "Pilot-symbol aided channel estimation for OFDM with fast fading channels," *IEEE Tran. Broadcast.*, vol. 49, no. 4, pp. 398–402, Dec. 2003.

[24] K. Han, S. Lee, J. Lim, and K. Sung, "Channel estimation for OFDM with fast fading channels by modified Kalman filter," *IEEE Trans. Consumer Electron.*, vol. 50, no. 2, pp. 443–449, 2004.

[25] Y. Mostofi and D. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765–774, Feb. 2005.

[26] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time varying channels for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 593–602, Mar. 2005.

[27] H. Dogan, E. Panayirci, and H. V. Poor, "A Gibbs sampling based MAP detection algorithm for OFDM over rapidly varying mobile radio channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM2009)*, Honolulu, HI, Nov. 30–Dec. 4 2009.

[28] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inf. Theory*, vol. 44, pp. 677–690, Mar. 1998.

[29] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 27–32, Jan. 1999.

[30] T. Zemen, C. F. Mecklenbrauker, J. Wehinger, and R. R. Muller, "Iterative joint time-variant channel estimation and multi-user detection for MC-CDMA," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1469–1478, Jun. 2006.

[31] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2664–2677, 1994.

[32] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[33] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. New York: Wiley-IEEE, 1994.

[34] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer-Verlag, 1994.

[35] G. Golub and C. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: John Hopkins Univ. Press, 1996.

[36] G. J. Foschni, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

Erdal Panayırcı (S'73–M'80–SM'91–F'03) received the Diploma Engineering degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey, and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing.

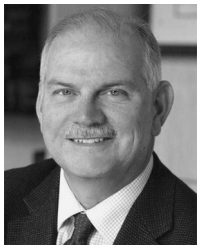
Until 1998, he was with the Faculty of Electrical and Electronics Engineering, Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is Professor of Electrical Engineering and Head of the Electronics Engineering Department at Kadir Has University, Istanbul. He spent two years (1980–1981) with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. Between 1990–1991, he was with the Center for Communications and Signal Processing, New Jersey Institute of Technology, Newark, as Visiting Professor, and in 1998–2000, he was Visiting Professor with the Department of Electrical Engineering, Texas A&M University, College Station. During 2008–2009, he was a Research Scholar with the Department of Electrical Engineering, Princeton University, Princeton, NJ. He has been the principal coordinator of a 6th and 7th Frame European project called Network of Excellent on Wireless Communications (NEWCOM) representing Kadir Has University for five years and WIMAGIC Strep project for two years. His recent research interests include communication theory, synchronization, advanced signal processing techniques and their applications to wireless communications, coded modulation and interference cancelation with array processing. He published extensively in leading scientific journals and international conferences. He has coauthored the book *Principles of Integrated Maritime Surveillance Systems* (Boston, MA: Kluwer Academic, 2000).

Dr. Panayırcı was an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS in the areas of synchronizations and equalizations during 1995–1999. He served as a Member of IEEE Fellow Committee during 2005–2008. He was the Technical Program Chair of the IEEE International Conference on Communications (ICC-2006) held in Istanbul in 2006. He is Technical Program Chair of the upcoming IEEE PIMRC to be held in Istanbul in 2010. Presently, he is head of the Turkish Scientific Commission on Signals and Systems of International Union of Radio Science (URSI).

Habib Şenol (S'04–M'07) was born in Nazilli, Turkey, in 1971. He received the B.S. and M.S. degrees from Istanbul University, Istanbul, Turkey, in 1993 and 1999, respectively, both in electronics engineering. From 1996 to 1999, he was a research assistant with Istanbul University. He received the Ph.D. degree in electronics engineering from Işık University, Istanbul, in 2006.

He is currently a faculty member of computer engineering with Kadir Has University, Istanbul. He was a Postdoctoral researcher with the Department of Electrical Engineering, Arizona State University, Tempe, in 2007. He has been a researcher of a 6th and 7th Frame European projects called Network of Excellent on Wireless Communications (NEWCOM) and WIMAGIC STREP. His research interests cover statistical signal processing, estimation and equalization algorithms for wireless communications, multicarrier (OFDM) systems, distributed detection, and estimation.

Dr. Şenol has served as reviewer for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE TRANSACTIONS ON COMMUNICATIONS.



H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Michael Henry Strater University Professor of Electrical Engineering and Dean of the School of Engineering and Applied Science. His research interests are in the areas of stochastic analysis, statistical signal processing and their applications in wireless networks, and related fields. Among his publications in these areas are the recent books *MIMO Wireless Communications* (Cambridge, U.K.: Cambridge Univ. Press, 2007) and *Quickest Detection* (Cambridge, U.K.: Cambridge Univ. Press, 2009).

Dr. Poor is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U.K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, and during 2004–2007, he served as the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY. He was the recipient of the 2005 IEEE Education Medal. Recent recognition of his work includes the 2007 Technical Achievement Award of the IEEE Signal Processing Society, and the 2008 Aaron D. Wyner Award of the IEEE Information Theory Society, and the 2009 Edwin Howard Armstrong Award of the IEEE Communications Society.