

Research on Channel Estimation and OFDM Signals Detection in Rapidly Time-Variant Channels

Min HUANG, Bingbing LI

The State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an, 710071, P.R.China

mhuang@mail.xidian.edu.cn, bbli01@126.com

Abstract. It is well known that iterative channel estimation and OFDM signals detection can significantly improve the performance of communication system. However, its performance is poor due to the modelling error of basis expansion model (BEM) being large enough and can not being ignored in rapidly time-variant channels. In this paper, channel estimation and OFDM signals detection are integrated into a real non-linear least squares (NLS) problem. Then the modified Broyden-Fletcher-Goldfarb-Shanno (MBFGS) algorithm is adopted to search the optimal solution. In addition, Cramer-Rao Bound (CRB) for our proposed approach is derived. Simulation results are presented to illustrate the superiority of the proposed approach.

Keywords

Channel estimation, OFDM signals detection, a real NLS problem, MBFGS, CRB.

1. Introduction

Increasing demand for high spectral efficiency and high performance has led to the development of fourth-generation (4G) broadband wireless systems. A potential transmission technique for 4G is orthogonal frequency-division multiplexing (OFDM) which has recently become one of the most popular modulation techniques and has been adopted as the transmission technology in many wireless communication standards such as Wireless Fidelity (Wi-Fi), Worldwide Interoperability for Microwave Access (WiMAX), Long Term Evolution (LTE) standards, and the Digital Video Broadcasting (DVB) Project [1]. In the OFDM communication systems, Channel estimation and signal detection are all key techniques. In the last decade, they have been extensively studied, respectively [2]-[23].

It is well known that coherent detection schemes are superior to differentially coherent or noncoherent schemes in terms of power efficiency, if channel information can be established perfectly. In time-variant channels, for computational convenience, some papers [2]-[4] assume the channel is static within one or more consecutive OFDM sym-

bols. In practice, this assumption will bring a certain amount of estimation error, especially in rapidly time-variant channels. However, the channel varies with time within a single OFDM symbol, leading to a big challenge. This is because the number of channel parameters which must be estimated is at least $N \times L$, where N and L denote the number of subcarriers and the number of channel taps, respectively, and it is larger than the number of the observation data in one OFDM symbol. Therefore, many existing works [5]-[11] resort to simplify channel model as a way of reducing the required number of channel parameters.

An alternative channel model is the Gauss-Markov model (GMM) [5], which models the time-variation of each tap by a Gauss-Markov process (usually only a first-order process is considered). However, it could only be appropriate for a slow-fading channel, not for a fast-fading channel [6]. Another popular channel model is the basis expansion model (BEM). In the BEM, the time-variation of each channel tap is expressed as a superposition of a few fixed basis functions, so that only $Q \times L$ BEM coefficients need to be estimated, where Q is the number of the basis functions. Several BEM variates are proposed in the literature, e.g., the complex-exponential BEM (CE-BEM) [7], the generalized CE-BEM (GCE-BEM) [8], the polynomial BEM (P-BEM) [9], the Karhunen-Loeve BEM (KL-BEM) [10], the discrete prolate spheroidal BEM (DPS-BEM) [11], and the others. Although the last two BEMs are closest to the true scenario, they require statistical channel knowledge, which has led to the model being usually unavailable in practice.

Based on the BEM models, the least squares estimator (LSE), the linear minimum mean square error estimator (LMMSEE) [12]-[14] and the best linear unbiased estimator (BLUE) are proposed in [6]. For the CE-BEM is constructed by a truncated Fourier series, it will cause the Gibbs phenomenon when it is used in these estimators [15]. In order to solve this problem, Hrycak proposes the orthogonal projection method [16] and the inverse reconstruction method [17]. Two methods based on P-BEM were proposed in [18] to estimate channel time-variations information in OFDM systems. The first one extracts these variations from the cyclic prefix (CP). The second one estimates that parameters by using the adjacent symbols. As the latter has more observation data, it is superior to the former. However, in these two schemes,

the pilots have been smeared by the inter-carrier interference (ICI). Tao et al [19] proposed a channel estimation technique using the ICI self-cancellation to resolve this problem. Nevertheless, this scheme requires the number of pilots being twice as much, leading to low spectral efficiency.

In terms of OFDM signals detection, the nonlinear equalizer MMSE with Successive Detection proposed in [20] outperformed the linear equalizers. For the purpose of improving the performance of detection, Wang et al [21] proposed a detection method which whitened the residual ICI and the noise, while Sebesta et al [22] proposed another scheme based on the cyclic autocorrelation function of CP.

Recently, in order to further improve system performance, many works [24]-[27] resort to the iterative strategy. In slowly time-variant channels, due to the exchange of information between channel estimation and OFDM signals detection, the iterative strategy has achieved good performance. However, in the rapidly time-variant channels, i.e., in the high speed railway or the low altitude aircraft, since the modelling error of BEM is large enough, the subsequent iterative operation is not adequate to compensate for this error.

In this paper, we propose a new approach from the point of view of global optimization. First, channel estimation and OFDM signals detection are equivalent to a complex non-linear least squares (NLS) problem. For convenience, it is transformed to a real NLS problem on the premise of that its characteristic remains the same. Then, the modified Broyden-Fletcher-Goldfarb-Shanno (MBFGS) algorithm [28], which is proved that it could achieve the global solution even for nonconvex unconstrained optimization problems, is adopted to solve the real NLS problem. Lastly, in order to evaluate the quality of our proposed approach, Cramer-Rao Bound (CRB) is derived.

In a word, the contributions of our paper are given as follows:

- We construct a real NLS problem, which perfectly represents the channel estimation and OFDM signals detection in Rapidly Time-Variant Channels.
- We propose a method, which mainly employs the MBFGS algorithm to solve the real NLS problem.
- We derive the CRB for our proposed approach.

The rest of the paper is structured as follows: Section 2 briefly introduces the system model which includes the OFDM system model, the basis expansion model and the OFDM system based on BEM. In Section 3, we depict our proposed method. CRB for our estimation is given in Section 4. Simulation results are exhibited in Section 5. Conclusions are presented in Section 6.

The following notations are used throughout the paper. Boldface lowercase and uppercase letters are used for vectors and matrices, respectively. Superscripts T , H and \dagger

denote transpose, conjugate transpose and pseudo inverse, respectively. The notation $\hat{\mathbf{x}}$, $\text{diag}(\mathbf{x})$ and $\mathbf{x}^{(k)}$ (or $X^{(k)}$) are reserved for the estimated \mathbf{x} , the diagonal matrix whose main diagonal equals \mathbf{x} and the vector \mathbf{x} (or the matrix X) in the k iteration, respectively. The matrix \mathbf{F} denotes the fast Fourier transform (FFT) matrix and the matrix \mathbf{F}^H denotes the inverse fast Fourier transform (IFFT) matrix. Furthermore, we denote the $x \times x$ identity matrix as \mathbf{I}_x .

2. System Model

2.1 OFDM System Model

It is assumed that the synchronizations of frequency and time are perfect. The received signal after removing the cyclic prefix (CP) is given by

$$\mathbf{y}(n) = \sum_{l=0}^{L-1} h_n^l d(n-l)_N + w(n), \quad 0 \leq n \leq N-1 \quad (1)$$

where h_n^l is the l th channel tap at the n th sample time, $d(n)$ is the n th transmitted sample, $(\cdot)_N$ represents a cyclic shift on the base of N , $w(n)$ is the additive white Gaussian noise (AWGN) with mean zero and variance σ_w^2 , L is the total number of channel taps. In order to avoid the inter-symbol interference (ISI), it is assumed that the highest values of path delays are always less than or equal to the length of CP in this paper.

Collecting the samples of the received signal to form a vector $\mathbf{y} = [y(0), \dots, y(N-1)]^T$ yields the following model

$$\mathbf{y} = \mathbf{H}_t \mathbf{d} + \mathbf{w} \quad (2)$$

where $\mathbf{d} = [d(0), \dots, d(N-1)]^T$, $\mathbf{w} = [w(0), \dots, w(N-1)]^T$, \mathbf{H}_t is an $N \times N$ channel impulse response matrix in the time domain. Using $h_n^l = 0$ for $N > l \geq L$, \mathbf{H}_t can be expressed as

$$\begin{aligned} \mathbf{H}_t &= \begin{bmatrix} h_0^0 & 0 & \cdots & 0 & h_0^{L-1} & \cdots & h_0^1 \\ h_1^1 & h_1^0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & h_{L-2}^{L-1} \\ h_{L-1}^{L-1} & \ddots & h_{L-1}^1 & h_{L-1}^0 & 0 & \ddots & 0 \\ 0 & h_L^{L-1} & \ddots & h_L^1 & h_L^0 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_{N-1}^{L-1} & \cdots & h_{N-1}^1 & h_{N-1}^0 \end{bmatrix} \\ &= \sum_{l=0}^{L-1} \text{diag}(\mathbf{h}_t^l) \mathbf{A}^l \end{aligned} \quad (3)$$

where $\mathbf{h}_t^l = [h_0^l, \dots, h_{N-1}^l]^T$ represents the l th channel tap within an OFDM symbol duration, and \mathbf{A} is the $N \times N$ cyclic permutation matrix given by

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

2.2 Basis Expansion Model

As can be seen from (3), it is very difficult to implement channel estimation in rapidly time-variant channels, since the number of parameters which need to be estimated is much larger than that of the observed data. A BEM [6] regarded as a simplified channel is employed, so that the number of estimated parameters is considerably reduced. Then, the l th channel tap \mathbf{h}_l^t can be presented as

$$\begin{aligned}\mathbf{h}_l^t &= \sum_{q=0}^{Q-1} h_{q,l} \mathbf{b}_q + \boldsymbol{\varepsilon}_l \\ &= \mathbf{B} \mathbf{h}_l + \boldsymbol{\varepsilon}_l\end{aligned}\quad (4)$$

where \mathbf{b}_q is the q th basis function, e.g., $\mathbf{b}_q = [e^{(j2\pi 0/N)(q-Q/2)} \dots e^{(j2\pi(N-1)/N)(q-Q/2)}]^T$ for CE-BEM [7] or $\mathbf{b}_q = [1^q \dots N^q]^T$ for P-BEM [9], etc. $\mathbf{B} = [\mathbf{b}_0, \dots, \mathbf{b}_{Q-1}]$ is an $N \times Q$ matrix that collects Q ($Q \ll N$) orthonormal basis function \mathbf{b}_q as columns, $\mathbf{h}_l = [h_{0,l}, \dots, h_{Q,l}]^T$ represents the BEM coefficients for the l th tap and $\boldsymbol{\varepsilon}_l$ represents the corresponding modeling error. As shown in (4), due to the BEM, the l th tap channel \mathbf{h}_l^t which needs to be estimated can be equivalent to \mathbf{h}_l . Therefore, the number of parameters for the l th tap decreases from N to Q .

2.3 OFDM System Model Based on BEM

Substituting (4) in (3), we can obtain

$$\begin{aligned}\mathbf{H}^t &= \sum_{l=0}^{L-1} \text{diag}(\sum_{q=0}^{Q-1} h_{q,l} \mathbf{b}_q + \boldsymbol{\varepsilon}_l) \mathbf{A}^l \\ &= \sum_{q=0}^{Q-1} \text{diag}(\mathbf{b}_q) \sum_{l=0}^{L-1} h_{q,l} \mathbf{A}^l + \sum_{l=0}^{L-1} \text{diag}(\boldsymbol{\varepsilon}_l) \mathbf{A}^l \\ &= \sum_{q=0}^{Q-1} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \Delta_q \mathbf{F} + \sum_{l=0}^{L-1} \text{diag}(\boldsymbol{\varepsilon}_l) \mathbf{A}^l\end{aligned}\quad (5)$$

with

$$\Delta_q = \text{diag}(\mathbf{F}_L [h_{q,0}, \dots, h_{q,L-1}]^T)$$

where \mathbf{F}_L stands for the first L columns of $\sqrt{N} \mathbf{F}$.

In the light of (5), (2) can be written as

$$\mathbf{y} = \sum_{q=0}^{Q-1} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \Delta_q \mathbf{F} \mathbf{d} + \sum_{l=0}^{L-1} \text{diag}(\boldsymbol{\varepsilon}_l) \mathbf{A}^l \mathbf{d} + \mathbf{w}. \quad (6)$$

After carrying out an N -point FFT, (6) becomes

$$\begin{aligned}\mathbf{r} &= \mathbf{F} \sum_{q=0}^{Q-1} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \Delta_q \mathbf{F} \mathbf{F}^H \mathbf{s} + \mathbf{F} \sum_{l=0}^{L-1} \text{diag}(\boldsymbol{\varepsilon}_l) \mathbf{A}^l \mathbf{F}^H \mathbf{s} + \mathbf{F} \mathbf{w} \\ &= \sum_{q=0}^{Q-1} \mathbf{F} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \Delta_q \mathbf{s} + \sum_{l=0}^{L-1} \mathbf{F} \text{diag}(\boldsymbol{\varepsilon}_l) \mathbf{A}^l \mathbf{F}^H \mathbf{s} + \mathbf{F} \mathbf{w} \\ &= \mathbf{H} \mathbf{s} + \boldsymbol{\psi} + \boldsymbol{\omega}\end{aligned}\quad (7)$$

with

$$\mathbf{H} = \sum_{q=0}^{Q-1} \mathbf{F} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \Delta_q, \boldsymbol{\psi} = \sum_{l=0}^{L-1} \mathbf{F} \text{diag}(\boldsymbol{\varepsilon}_l) \mathbf{A}^l \mathbf{F}^H \mathbf{s}$$

or

$$\begin{aligned}\mathbf{r} &= \sum_{q=0}^{Q-1} \mathbf{F} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \Delta_q \mathbf{s} + \boldsymbol{\psi} + \boldsymbol{\omega} \\ &= \sum_{q=0}^{Q-1} \mathbf{F} \text{diag}(\mathbf{b}_q) \mathbf{F}^H \text{diag}(\mathbf{s}) \mathbf{F}_L \mathbf{h} + \boldsymbol{\psi} + \boldsymbol{\omega} \\ &= \mathbf{P} \mathbf{h} + \boldsymbol{\psi} + \boldsymbol{\omega}\end{aligned}\quad (8)$$

with

$$\begin{aligned}\mathbf{P} &= [\mathbf{F} \text{diag}(\mathbf{b}_0) \mathbf{F}^H \text{diag}(\mathbf{s}) \mathbf{F}_L, \dots, \mathbf{F} \text{diag}(\mathbf{b}_{Q-1}) \mathbf{F}^H \text{diag}(\mathbf{s}) \mathbf{F}_L] \\ \mathbf{h} &= [h_{0,0}, \dots, h_{0,L-1}, \dots, h_{Q-1,0}, \dots, h_{Q-1,L-1}]^T\end{aligned}$$

where $\mathbf{s} = [s(0), \dots, s(N-1)]^T$, which is the N -point FFT of \mathbf{d} , represents the transmit data in the frequency domain. $\boldsymbol{\omega} = [\omega(0), \dots, \omega(N-1)]^T$, which is the N -point FFT of \mathbf{w} , represents the AWGN in the frequency domain.

Hijazi and Ros [24] proposed an iterative channel estimation and OFDM signals detection method, i.e., an alternating least square (ALS) algorithm. With (8), the detected information data which can be obtained by the LSE in the previous iterations, are used for refining the channel estimation by the least square (LS) equalizer [20] in the light of (7). In slowly time-variant channels, since the BEM modeling error in (4) is small enough, the interference item $\boldsymbol{\psi}$ in (7) and (8) can be negligible [6]. Consequently the LS method applied in (7) and (8) could be regarded as the optimal algorithm, which can be proved in Appendix A. In addition, according to the Lemma 1 proposed in [29], ALS algorithm could achieve a local optimum solution.

Lemma 1 *Supposing there are two resolved vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in the given equations, alternating optimization algorithm (which solves for the optimal $\boldsymbol{\alpha}$ by fixing $\boldsymbol{\beta}$, and then solves for the optimal $\boldsymbol{\beta}$ by fixing $\boldsymbol{\alpha}$ and iterates in this manner until convergence) could obtain a local optimum solution.*

However, in rapidly time-variant channels, the interference item $\boldsymbol{\psi}$ in (7) and (8) can not be negligible for the BEM modeling error being no longer small [6]. Moreover, being different from $\boldsymbol{\omega}$, the interference item $\boldsymbol{\psi}$ associated with the BEM modeling error and the transmitted data is non-Gaussian distributed. According to Appendix A, since LS method can not be guaranteed to be the optimal algorithm in non-Gaussian, the ALS applied in [24] can not be guaranteed to achieve even a local optimum solution. For solving this problem, a new method is proposed in the next section.

3. Proposed Method

3.1 Construction of the Real NLS Problem

We assume there are K equally spaced pilots, s_{l_i} , at subcarriers $l_i = (i \times N)/K$, for $0 \leq i < K$. All these pilots together form the pilot vector $\mathbf{s}_p = [s_{l_0}, \dots, s_{l_{K-1}}]^T$. The remaining subcarriers in single OFDM symbol are reserved for the information data, which can be collected in the information vector \mathbf{s}_d . For convenience, we collect the unknown parameters in (7) or (8) to form a vector $\mathbf{x} = [\mathbf{s}_d^H, \mathbf{h}^H]^H \in \mathbb{C}^{M \times 1}$ with $M = N - K + Q \times L$. Consequently, (7) and (8) can be written as

$$\mathbf{r} = \mathbf{f}(\mathbf{x}) + \boldsymbol{\psi} + \boldsymbol{\omega} \quad (9)$$

where \mathbf{f} is a nonlinear map from \mathbf{x} to \mathbf{Hs} or \mathbf{Ph} . For well estimating \mathbf{x} in (9), the maximum likelihood estimation (MLE) regarded as optimal estimator is adopted and given by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathbb{C}^{M \times 1}} \frac{1}{2} \|\mathbf{f}(\mathbf{x}) - \mathbf{r}\|_2^2. \quad (10)$$

Obviously, this is a complex NLS problem. In order to avoid the complex derivative in the later discussion, the problem should be transformed into a real NLS problem. Due to the unknown parameter vectors of the real and imaginary parts being independent of each other, we can define

$$\begin{aligned} \bar{\mathbf{x}} &= [\operatorname{Re}[\mathbf{s}_d]^T, \operatorname{Im}[\mathbf{s}_d]^T, \operatorname{Re}[\mathbf{h}]^T, \operatorname{Im}[\mathbf{h}]^T]^T \in \mathbb{R}^{2M \times 1} \\ \bar{\mathbf{r}} &= [\operatorname{Re}[\mathbf{r}]^T, \operatorname{Im}[\mathbf{r}]^T]^T \in \mathbb{R}^{2N \times 1}. \end{aligned}$$

Then, (10) becomes

$$\bar{\mathbf{x}} = \arg \min_{\bar{\mathbf{x}} \in \mathbb{R}^{2M \times 1}} \phi(\bar{\mathbf{x}}) \quad (11)$$

where $\phi(\bar{\mathbf{x}}) = \frac{1}{2} \|\mathbf{g}(\bar{\mathbf{x}}) - \bar{\mathbf{r}}\|_2^2$ with \mathbf{g} is a nonlinear map from $\bar{\mathbf{x}}$ to $[\operatorname{Re}[\mathbf{Hs}]^T, \operatorname{Im}[\mathbf{Hs}]^T]^T$ or $[\operatorname{Re}[\mathbf{Ph}]^T, \operatorname{Im}[\mathbf{Ph}]^T]^T$.

3.2 Solution of the Real NLS Problem

It is known that BFGS algorithm [30] is a considerable popular Quasi-Newton method for solving the convex optimization problems. However, we could not directly adopt this algorithm to solve the real NLS problem for the problem being not necessarily a convex optimization problem. Fortunately, the MBFGS algorithm proposed in [28], has been proved that it could achieve the global solution even for nonconvex unconstrained optimization problems. Therefore, we adopt the MBFGS algorithm to solve the real NLS problem. The MBFGS uses several simple update formulas to sequentially update $\bar{\mathbf{x}}$, starting from a random vector $\bar{\mathbf{x}}^{(0)}$, until a stable solution is obtained. The update formulas are given at the top of next page.

Obviously, the estimated information vector and the BEM coefficients vector can be extracted from the estimated

$\hat{\mathbf{x}}$. However, we have assumed that the information data are continuous variables in (11) for convenience. In practical, they only take the discrete constellation points. To overcome this problem, we do a hard decision of $\hat{\mathbf{s}}_d$. Then, together with the pilots to update \mathbf{P} in (8). Subsequently, we can achieve the more exact $\hat{\mathbf{h}}$ by the LS estimator [6]

$$\hat{\mathbf{h}} = \mathbf{P}^\dagger \mathbf{r}. \quad (14)$$

With the aid of $\hat{\mathbf{h}}$, renew the matrix \mathbf{H} in (7). Then, after eliminating the effect of the pilots on the information data, (7) can be written as

$$\tilde{\mathbf{r}} = \mathbf{r} - \mathbf{H}_p \mathbf{s}_p = \mathbf{H}_d \mathbf{s}_d + \boldsymbol{\omega} \quad (15)$$

where \mathbf{H}_p is a $N \times K$ matrix, which is carved out of \mathbf{H} corresponding to the pilot subcarriers.

Lastly, the more exact information vector $\hat{\mathbf{s}}_d$ can be easily obtained by LS equalizer [20]

$$\hat{\mathbf{s}}_d = \mathbf{H}_d^\dagger \tilde{\mathbf{r}}. \quad (16)$$

3.3 Complexity Analysis

In this paper, the computational complexity of the proposed method is evaluated by the number of real multiplications. Assume that the operation of $N \times N$ complex matrix inversion needs N^3 complex multiplications and a complex multiplication is equivalent to three real multiplications. For our proposed method, the main complexity is in solving MBFGS, (14), (15) and (16). The number of real multiplication required to calculate each step of MBFGS is listed in Tab. 1. The computation of $\hat{\mathbf{h}}$ in (14) requires $3QLN + 6(QL)^2N + 3(QL)^3$ real multiplications. After updating the matrix \mathbf{H} in (7) in the light of $\hat{\mathbf{h}}$, the computation of $\tilde{\mathbf{r}}$ in (15) requires $3NK$ real multiplications. Lastly, the computation of $\hat{\mathbf{s}}_d$ in (16) requires $9N^3 - 21KN^2 + 3N^2 + 15K^2N - 3KN - 3K^3$ real multiplications. Therefore, the whole algorithm requires $I_m[9QN^3 + 15QN^2 + 3QLN^2 + 27N^2 + 6QL^2N + 3LN + 4MN + 16M^3 + 4M^2 + 2M \log(2M) + 6M] + 9N^3 - 21KN^2 + 3N^2 + 15K^2N + 6(QL)^2N + 3QLN - 3K^3 + 3(QL)^3$ real multiplications. Here, I_m represents the number of iterations in MBFGS. Note that, for large N , the computational complexity of the proposed method is $\sim O([9I_m Q + 9]N^3)$.

On the other hand, in order to compare the complexity of different methods, the complexity of our proposed method, the LSE proposed in [6] and the ALS algorithm proposed in [24] are all presented in Tab. 2. Here, the symbol $I_{[24]}$ denotes the number of iterations in [24]. Assuming $I_{[24]} = I_m$, it is concluded that these methods can be ordered in terms of the complexity in an ascending manner as the LSE, the ALS algorithm, and finally, our proposed method. However, the theory analysis above and the simulation results given in Section 5 confirm that these methods are ordered in terms of the performance in opposite direction.

Algorithm MBFGS

Step 1: Initialize $k = 0$, $\Omega^{(0)} = \mathbf{I}_{2M}$, where Ω denotes an approximation of the inverse Hessian of $\phi(\bar{\mathbf{x}}^{(k)})$.

Step 2: In the light of (7) and (8), $\mathbf{P}^{(k)}$ and $\mathbf{H}^{(k)}$ can be constructed by $\bar{\mathbf{x}}^{(k)}$ and the pilot vector \mathbf{s}_p . Then, let $\nabla\phi(\bar{\mathbf{x}}^{(k)})$ denote the gradient of the $\phi(\bar{\mathbf{x}}^{(k)})$ with respect to $\bar{\mathbf{x}}^{(k)}$. And it can be easily obtained from

$$\nabla\phi(\bar{\mathbf{x}}^{(k)}) = \begin{bmatrix} \text{Re}[\mathbf{H}_d^{(k)}] & \text{Re}[\mathbf{P}^{(k)}] & -\text{Im}[\mathbf{H}_d^{(k)}] & -\text{Im}[\mathbf{P}^{(k)}] \\ \text{Im}[\mathbf{H}_d^{(k)}] & \text{Im}[\mathbf{P}^{(k)}] & \text{Re}[\mathbf{H}_d^{(k)}] & \text{Re}[\mathbf{P}^{(k)}] \end{bmatrix}^T \begin{bmatrix} \text{Re}[\mathbf{g}(\bar{\mathbf{x}}^{(k)}) - \bar{\mathbf{r}}] \\ \text{Im}[\mathbf{g}(\bar{\mathbf{x}}^{(k)}) - \bar{\mathbf{r}}] \end{bmatrix} \quad (12)$$

where $\mathbf{H}_d^{(k)}$ is a $N \times (N - K)$ matrix, which is carved out of $\mathbf{H}^{(k)}$ corresponding to the information data subcarriers.

Step 3: Compute the search direction $\mathbf{z}^{(k)} = -\Omega^{(k)}\nabla\phi(\bar{\mathbf{x}}^{(k)})$.

Step 4: Cast about the optimal step length $\lambda^{(k)} = \arg \min_{\lambda \geq 0} \phi(\bar{\mathbf{x}}^{(k)} + \lambda^{(k)}\mathbf{z}^{(k)})$.

Step 5: Update solution $\bar{\mathbf{x}}^{(k+1)} = \bar{\mathbf{x}}^{(k)} - \lambda^{(k)}\mathbf{z}^{(k)}$.

Step 6: Being similar to **Step 2**, $\mathbf{P}^{(k+1)}$ and $\mathbf{H}^{(k+1)}$ can be updated by $\bar{\mathbf{x}}^{(k+1)}$ and the pilot vector \mathbf{s}_p . Then, compute $\phi(\bar{\mathbf{x}}^{(k+1)})$ and check convergence. If $\|\phi(\bar{\mathbf{x}}^{(k+1)})\| < \varepsilon$ for sufficiently small values of ε , stop.

Step 7: Compute $\nabla\phi(\bar{\mathbf{x}}^{(k+1)})$ by (12).

Step 8: Set $\mathbf{p}^{(k+1)} = \lambda^{(k)}\mathbf{z}^{(k)}$ and $\mathbf{q}^{(k+1)} = \nabla\phi(\bar{\mathbf{x}}^{(k+1)}) - \nabla\phi(\bar{\mathbf{x}}^{(k)}) + t^{(k)}\|\nabla\phi(\bar{\mathbf{x}}^{(k)})\|\mathbf{p}^{(k)}$ with

$$t^{(k)} = 1 + \max \left\{ 0, -\frac{(\nabla\phi(\bar{\mathbf{x}}^{(k+1)}) - \nabla\phi(\bar{\mathbf{x}}^{(k)}))^T \mathbf{p}^{(k)}}{\|\mathbf{p}^{(k)}\|^2} \right\},$$

then update the approximation of the inverse Hessian by

$$\Omega^{(k+1)} = \left(\mathbf{I}_{2M} - \frac{\mathbf{p}^{(k)} [\mathbf{q}^{(k)}]^T}{[\mathbf{q}^{(k)}]^T \mathbf{p}^{(k)}} \right) \Omega^{(k)} \left(\mathbf{I}_{2M} - \frac{\mathbf{q}^{(k)} [\mathbf{p}^{(k)}]^T}{[\mathbf{q}^{(k)}]^T \mathbf{p}^{(k)}} \right) + \frac{\mathbf{p}^{(k)} [\mathbf{p}^{(k)}]^T}{[\mathbf{q}^{(k)}]^T \mathbf{p}^{(k)}}. \quad (13)$$

Step 9: Return to **Step 3** with $k = k + 1$.

The number of real multiplications	
Step 1	0
Step 2	$6QN^3 + 9QN^2 + 3QL^2N + 3QLN^2 + 3LN + 4NM + N^2$
Step 3	$4M^2$
Step 4	$2M \log(2M)$
Step 5	$2M$
Step 6	$3QN^3 + 6QN^2 + 3QL^2N + 4M$
Step 7	$6QN^3 + 9QN^2 + 3QL^2N + 3QLN^2 + 3LN + 4NM + N^2$
Step 8	$16M^3 + 26N^2$
Step 9	0

Tab. 1. Computation complexity of each step in MBFGS.

The number of real multiplications	
Our proposed method	$\sim O([9I_m Q + 9]N^3)$
LSE	$\sim O(3QN^3)$
The ALS algorithm	$\sim O(6I_{[24]}QN^3)$

Tab. 2. Computation complexity of each method.

4. Cramer-Rao Bound

In order to evaluate the quality of our proposed approach, the CRB for our estimation is derived in this section. Assuming the interference item ψ can be negligible, the received vector \mathbf{r} in (7) or (8) is a complex circularly Gaussian process with mean $\boldsymbol{\mu} = \mathbf{H}\mathbf{s}$ or $\boldsymbol{\mu} = \mathbf{P}\mathbf{h}$ and variance $\mathbf{R} = \sigma_w^2 \mathbf{I}_N$. According to the conclusion in [12], [31], its Fisher Matrix

is given by

$$\begin{aligned} [\mathbf{J}]_{i,j} &= \text{trace} \left\{ \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \bar{\mathbf{x}}_i} \mathbf{R}^{-1} \frac{\partial \mathbf{R}}{\partial \bar{\mathbf{x}}_j} \right\} \\ &\quad + 2\text{Re} \left[\frac{\partial \boldsymbol{\mu}^H}{\partial \bar{\mathbf{x}}_i} \mathbf{R}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \bar{\mathbf{x}}_j} \right] \\ &= \frac{2}{\sigma_w^2} \text{Re} \left[\frac{\partial \boldsymbol{\mu}^H}{\partial \bar{\mathbf{x}}_i} \frac{\partial \boldsymbol{\mu}}{\partial \bar{\mathbf{x}}_j} \right]. \end{aligned} \quad (17)$$

For the sake of convenience, we can define

$$\bar{\mathbf{s}}_d = [\text{Re}[\mathbf{s}_d]^T, \text{Im}[\mathbf{s}_d]^T]^T, \bar{\mathbf{h}} = [\text{Re}[\mathbf{h}]^T, \text{Im}[\mathbf{h}]^T]^T.$$

As a result, (17) becomes

$$\mathbf{J} = \frac{2}{\sigma_w^2} \begin{bmatrix} \mathbf{J}_{\bar{\mathbf{s}}_d \bar{\mathbf{s}}_d} & \mathbf{J}_{\bar{\mathbf{s}}_d \bar{\mathbf{h}}} \\ \mathbf{J}_{\bar{\mathbf{h}} \bar{\mathbf{s}}_d} & \mathbf{J}_{\bar{\mathbf{h}} \bar{\mathbf{h}}} \end{bmatrix} \quad (18)$$

with

$$\begin{aligned} \mathbf{J}_{\bar{\mathbf{s}}_d \bar{\mathbf{s}}_d} &= \text{Re} \left[\frac{\partial \boldsymbol{\mu}^H}{\partial \bar{\mathbf{s}}_d} \frac{\partial \boldsymbol{\mu}}{\partial \bar{\mathbf{s}}_d} \right], \mathbf{J}_{\bar{\mathbf{s}}_d \bar{\mathbf{h}}} = \text{Re} \left[\frac{\partial \boldsymbol{\mu}^H}{\partial \bar{\mathbf{s}}_d} \frac{\partial \boldsymbol{\mu}}{\partial \bar{\mathbf{h}}} \right], \\ \mathbf{J}_{\bar{\mathbf{h}} \bar{\mathbf{s}}_d} &= \text{Re} \left[\frac{\partial \boldsymbol{\mu}^H}{\partial \bar{\mathbf{h}}} \frac{\partial \boldsymbol{\mu}}{\partial \bar{\mathbf{s}}_d} \right], \mathbf{J}_{\bar{\mathbf{h}} \bar{\mathbf{h}}} = \text{Re} \left[\frac{\partial \boldsymbol{\mu}^H}{\partial \bar{\mathbf{h}}} \frac{\partial \boldsymbol{\mu}}{\partial \bar{\mathbf{h}}} \right]. \end{aligned}$$

Since

$$\frac{\partial \boldsymbol{\mu}^H}{\partial \text{Re}[\mathbf{s}_d]} = \frac{\partial \mathbf{s}^H \mathbf{H}^H}{\partial \text{Re}[\mathbf{s}_d]} = \mathbf{H}_d^H \quad (19)$$

and

$$\frac{\partial \boldsymbol{\mu}^H}{\partial \text{Im}[\mathbf{s}_d]} = \frac{\partial \mathbf{s}^H \mathbf{H}^H}{\partial \text{Im}[\mathbf{s}_d]} = -j\mathbf{H}_d^H, \quad (20)$$

then

$$\frac{\partial \mu^H}{\partial \bar{s}_d} = \frac{\partial s^H H^H}{\partial \bar{s}_d} = \Xi_1 H_d^H, \quad (21)$$

with

$$\Xi_1 = [I_{N-K}, jI_{N-K}]^H.$$

Similarly, we have

$$\frac{\partial \mu^H}{\partial \bar{h}} = \frac{\partial h^H P^H}{\partial \bar{h}} = \Xi_2 P^H \quad (22)$$

with

$$\Xi_2 = [I_{LQ}, jI_{LQ}]^H.$$

Substituting (21) and (22) in (18), we obtain

$$J = \frac{2}{\sigma_w^2} \begin{bmatrix} \text{Re}[\Xi_1 H_d^H H_d \Xi_1^H] & \text{Re}[\Xi_1 H_d^H P \Xi_2^H] \\ \text{Re}[\Xi_2 P^H H_d \Xi_1^H] & \text{Re}[\Xi_2 P^H P \Xi_2^H] \end{bmatrix}. \quad (23)$$

Then, the Cramer-Rao Bound for \mathbf{h} can be expressed as

$$CRB_h = \sum_{i=2(N-K)}^{2M-1} [J^{-1}]_{i,i} \quad (24)$$

Note that the mean square error of the estimated \mathbf{h}_l^t is given by

$$\begin{aligned} MSE_{\mathbf{h}_l^t} &= E[(\hat{\mathbf{h}}_l^t - \mathbf{h}_l^t)^H (\hat{\mathbf{h}}_l^t - \mathbf{h}_l^t)] \\ &= E[(\mathbf{B}\hat{\mathbf{h}}_l - \mathbf{B}\mathbf{h}_l)^H (\mathbf{B}\hat{\mathbf{h}}_l - \mathbf{B}\mathbf{h}_l)] \\ &= E[(\hat{\mathbf{h}}_l - \mathbf{h}_l)^H \mathbf{B}^H \mathbf{B} (\hat{\mathbf{h}}_l - \mathbf{h}_l)] \\ &= E[(\hat{\mathbf{h}}_l - \mathbf{h}_l)^H (\hat{\mathbf{h}}_l - \mathbf{h}_l)] \\ &= MSE_{\mathbf{h}_l} \end{aligned} \quad (25)$$

where $MSE_{\mathbf{h}_l}$ represents the mean square error of the estimated \mathbf{h}_l , for the columns of \mathbf{B} being orthonormal.

Without loss of generality, we will consider the wide sense stationary uncorrelated scattering (WSSUS) model as the real channel mode, whose taps are independent with each other. Therefore, the mean square error of the estimated \mathbf{h} can be written as

$$MSE_{\mathbf{h}} = L \times MSE_{\mathbf{h}_l}. \quad (26)$$

From (24), (25) and (26), we have

$$MSE_{\mathbf{h}_l^t} \geq \frac{1}{L} \sum_{2(N-K)}^{2M-1} [J^{-1}]_{i,i}. \quad (27)$$

5. Simulation Results

In this section, by using Matlab, we present simulation results to assess the QPSK-OFDM system performance

based on the proposed algorithm. The parameters of the system selected are in concordance with the standard WiMAX IEEE 802.16e. The system operates with a 1.25 MHz bandwidth and is divided into 512 subcarriers. The carrier frequency is set to 3.5 GHz. The length of CP, as well as the number of pilots, is 64. The scheme given in [32], [33] will be used for generating the time-variant channels. We further assume the number of taps is 4. The mean power and the time delay of the l th path are $e^{-l/10}$ and $0.8 \cdot l \mu s$ for $l \in \{0, \dots, L-1\}$, respectively. Moreover, all channel taps have a Jakes Doppler spectrum. In our proposed algorithm, the P-BEM will be adopted to simplify the time-variant channels, leading to estimated parameters reduction. As a rule of thumb, the Q depicted in (5) should satisfy $(Q-1)/2 = \lceil f_n \rceil + 1$ [34]. Here, f_n represent the normalized Doppler frequency shift. Therefore, in our simulation, $Q = 5$ is selected for $f_n = 0.2$ representing slowly time-variant cahnnels or $f_n = 0.8$ representing rapidly time-variant channels.

Note that, the receiver velocity v is related to the normalized Doppler frequency f_n by the formula [6], [16]

$$v = f_n \frac{cB}{fcN}$$

where c , B and f_c represent the speed of light, the OFDM signal bandwidth and the carrier frequency, respectively. Hence, we can easily find that $f_n = 0.2$ and $f_n = 0.8$ are equivalent to a user moving at the speed of 151 km/h and 603 km/h with the WiMAX system parameters, respectively.

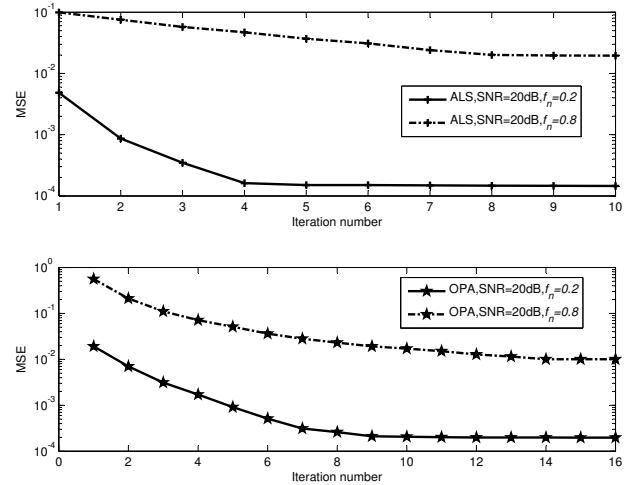


Fig. 1. Convergence characteristic of the ALS algorithm and OPA under various normalized Doppler shift. Solid curves: $f_n = 0.2$. Dashed curves: $f_n = 0.8$.

Fig. 1 shows the MSE of channel estimation for the ALS algorithm [24] and our proposed algorithm (OPA) under various numbers of iterations. Note, the solid lines and the dashed lines represent the simulations under the normalized Doppler frequency shift $f_n = 0.2$ and $f_n = 0.8$, respectively. From the graph, it can be seen that the ALS algorithm has to go through 4 iterations before convergence for

$f_n = 0.2$ while 8 iterations are needed for $f_n = 0.8$. OPA has to go through 9 iterations before convergence for $f_n = 0.2$ while 14 iterations are needed for $f_n = 0.8$. According to the complexity analysis in part 3 of Section 3, it can easily be found that the convergence speed of OPA is slower than that of the ALS algorithm, and consequently the complexity of OPA is about 3.5 times more than that of the ALS algorithm for $f_n = 0.2$ while about 2.7 times for $f_n = 0.8$. However, the simulation results given in Fig. 2 and Fig. 3 illustrate that OPA outperforms the ALS algorithm in rapidly time-variant channels.

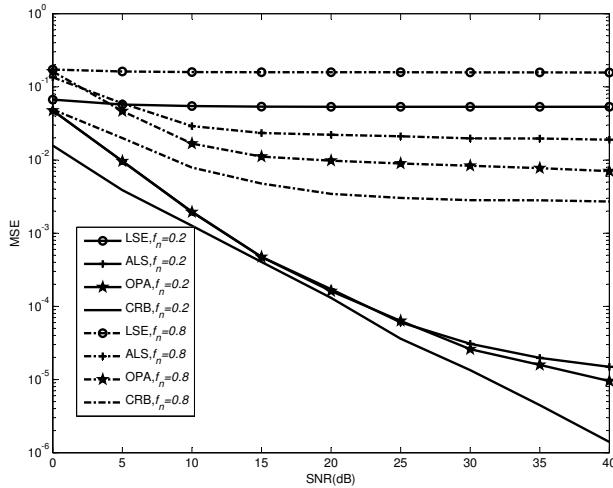


Fig. 2. MSE versus SNR for LSE, the ALS algorithm and OPA.
Solid curves: $f_n = 0.2$. Dashed curves: $f_n = 0.8$.

In the condition of convergence, Fig. 2 shows the MSE of channel estimation for LSE, the ALS algorithm and OPA under various SNR. As a reference, we also plot the CRB for our proposed algorithm. The simulation results indicate that both the ALS algorithm and OPA are much better than LSE. This is due to LSE applying to the FDKD [35] pilot structure only and LSE not using the transmitted information data. Furthermore, when $f_n = 0.2$, the MSE of OPA is substantially the same as that of the ALS algorithm, which is relatively close to the CRB. When $f_n = 0.8$, the MSE of OPA is superior to that of the ALS algorithm. However, due to the BEM modeling error in (4) being relatively large under the rapidly time-variant channel, the MSE of OPA is still a little worse than that of CRB.

Fig. 3 shows the bit error rate (BER) performance with respect to the SNR for $f_n = 0.2$ and $f_n = 0.8$. It is obvious that the detection performance is almost in consistence with the corresponding channel estimation performance for the three methods. Especially, it is more apparent that the performance of our method is much better than that of the iterative method in the rapidly time-variant channel.

Furthermore, we operate our simulation in the Typical Urban (TU) channel [36] whose parameters are summarized in Tab. 3. The performance of convergence, estimation and detection for different algorithms are shown in Fig. 4, Fig. 5

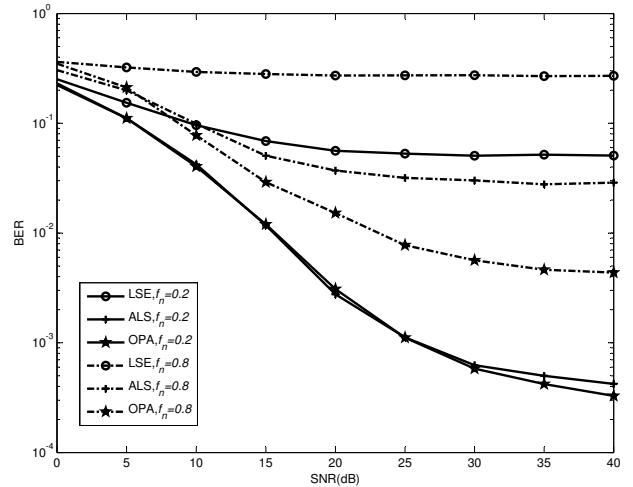


Fig. 3. BER versus SNR for LSE, the ALS algorithm and OPA.
Solid curves: $f_n = 0.2$. Dashed curves: $f_n = 0.8$.

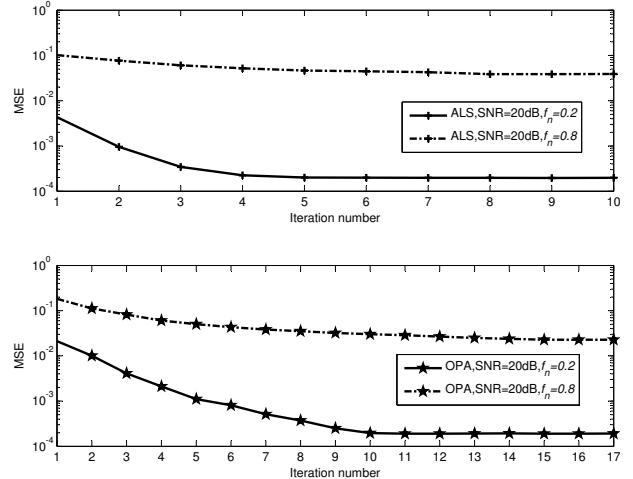


Fig. 4. Convergence characteristic of the ALS algorithm and OPA under various normalized Doppler shift values in the Typical Urban (TU) channel. Solid curves: $f_n = 0.2$. Dashed curves: $f_n = 0.8$.

Tap	Tap delay (μs)	Tap gain (dB)	Doppler spectrum
0	0	-3	Jakes
1	0.2	0	Jakes
2	0.5	-2	Jakes
3	1.6	-6	Gauss I
4	2.3	-8	Gauss II
5	5.0	-10	Gauss II

Tab. 3. Typical Urban (TU) channel parameters.

and Fig. 6, respectively. Fig. 4 indicates that the MSE of the ALS algorithm converges to the optimal point in 5 iterations for $f_n = 0.2$ and in 8 iterations for $f_n = 0.8$. Further, the MSE of OPA converges to the optimal point in 10 iterations for $f_n = 0.2$ and in 14 iterations for $f_n = 0.8$. As well as the description in Fig. 1, this shows that OPA keeps worse convergence performance than the ALS algorithm in

TU channel, leading to high computational complexity. Nevertheless, as shown in Fig. 2, Fig. 5 illustrates that the MSE performance of OPA in rapidly time-variant channels is still better than that of the ALS algorithm. In addition, comparing the results in Fig. 6 with those in Fig. 3, it is obvious that OPA also has lower BER than the ALS algorithm in the worse channel, i.e., the TU channel.

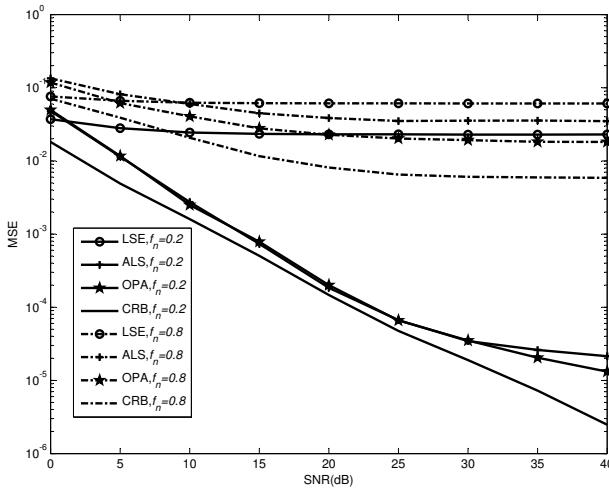


Fig. 5. MSE versus SNR for LSE, the ALS algorithm and OPA in the Typical Urban (TU) channel. Solid curves: $f_n = 0.2$. Dashed curves: $f_n = 0.8$.

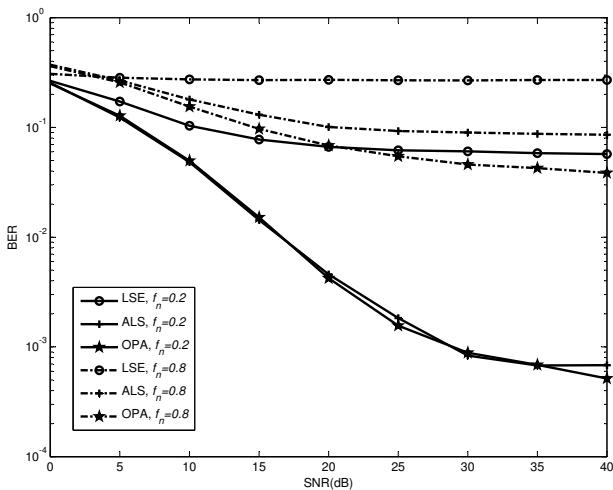


Fig. 6. BER versus SNR for LSE, the ALS algorithm and OPA in the Typical Urban (TU) channel. Solid curves: $f_n = 0.2$. Dashed curves: $f_n = 0.8$.

6. Conclusion

In this paper, we propose a novel algorithm for joint channel estimation and signal detection, which are considered a real NLS problem. Then, the MBFGS algorithm is adopted to solve the problem. Moreover, the CRB is derived

for evaluating the quality of our proposed algorithm. Simulation results show that our proposed algorithm achieves better performance than the iterative method in rapidly time-variant channel.

Acknowledgements

This work was supported by the 111 Project (B08038), the National Natural Science Foundation of China under Grant No. 61271299 and the Shenzhen Kongqie talent program under Grant YFZZ20111013.

Appendix A

LS method can be regarded as the optimal algorithm in slowly time-variant channels.

First, we will prove that the LS method applied in (8) can be regarded as the optimal algorithm in slowly time-variant channels.

Since the BEM modeling error in (4) is very small in slowly time-variant channels, the interference item ψ in (8) can be negligible. Therefore, (8) can be rewritten as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \boldsymbol{\omega} \quad (28)$$

With \mathbf{H} , the transmitted data \mathbf{s} in (28) can be estimated by maximum likelihood estimator (MLE) which is considered the optimal estimator [12]. The MLE estimate is obtained by maximizing the likelihood function of the received data \mathbf{r} given transmitted data \mathbf{s} , which is given by

$$p(\mathbf{r}|\mathbf{s}) = \frac{1}{(2\pi)^{N/2} |\mathbf{R}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{r} - \mathbf{H}\mathbf{s})^H \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}\mathbf{s}) \right]. \quad (29)$$

For computational convenience, the MLE estimate can be obtained by the log-likelihood function, which can be expressed as

$$\ln p(\mathbf{r}|\mathbf{s}) = \ln \frac{1}{(2\pi)^{N/2} |\mathbf{R}|^{1/2}} - \frac{1}{2} (\mathbf{r} - \mathbf{H}\mathbf{s})^H \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}\mathbf{s}). \quad (30)$$

By setting the derivative of the log-likelihood function with respect to \mathbf{s} to zero, we have

$$\begin{aligned} \frac{\partial \ln p(\mathbf{r}|\mathbf{s})}{\partial \mathbf{s}} &= -\frac{1}{2} \frac{\partial (\mathbf{r} - \mathbf{H}\mathbf{s})^H \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}\mathbf{s})}{\partial \mathbf{s}} \\ &= -\frac{1}{2} \mathbf{H}^H \mathbf{R}^{-1} (\mathbf{r} - \mathbf{H}\mathbf{s}) \\ &= 0 \end{aligned} \quad (31)$$

which leads to the MLE solution for \mathbf{s} given by

$$\begin{aligned} \mathbf{s} &= (\mathbf{H}^H \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^H \mathbf{R}^{-1} \mathbf{r} \\ &= (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{r} \\ &= \mathbf{H}^\dagger \mathbf{r}. \end{aligned} \quad (32)$$

As can be seen in the equation above, the solution for MLE is the same as that for LS estimator. Hence, the LS

method applied in (8) can be regarded as the optimal algorithm in slowly time-variant channels.

Similarly, the LS method applied in (7) also can be regarded as the optimal algorithm in slowly time-variant channels.

References

- [1] POLAK, L., KRATOCHVIL, T. DVB-H and DVB-SH-A performance and evaluation of transmission in fading channels. In *Proceedings of the 34th International Conference on Telecommunication and Signal Processing (TSP2011)*. Budapest (Hungary), 2011, p. 549 - 553.
- [2] SANZI, F., JELTING, S., SPEIDEL, J. A comparative study of iterative channel estimators for mobile OFDM systems. *IEEE Transactions on Wireless Communications*, 2003, vol. 2, no. 5, p. 849 - 859.
- [3] NEVAT, I., YUAN, J. Error propagation mitigation for iterative channel tracking, detection and decoding of BICM-OFDM systems. In *4th International Symposium on Wireless Communication Systems (ISWCS)*. Trondheim (Norway), 2007, p. 75 - 80.
- [4] AL-NAFFOURI, T. An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM. *IEEE Transactions on Signal Processing*, 2007, vol. 55, no. 7, p. 3924 - 3930.
- [5] DONG, M., TONG, L., SADLER, B. Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels. *IEEE Transactions on Signal Processing*, 2004, vol. 52, no. 5, p. 1403 - 1418.
- [6] TANG, Z., CANNIZARO, R. C., LEUS, G., BANELLI, P. Pilot-assisted time-varying channel estimation for OFDM systems. *IEEE Transactions on Signal Processing*, 2007, vol. 55, no. 5, p. 2226 - 2238.
- [7] TSATSANIS, M. K., GIANNAKIS, G. B. Modeling and equalization of rapidly fading channels. *International Journal of Adaptive Control and Signal Processing*, 1996, vol. 10, no. 2 - 3, p. 159 - 176.
- [8] LEUS, G. On the estimation of rapidly time-varying channels. In *Proceedings of 12th European Signal Processing Conference (EUSIPCO)*. Vienna (Austria), 2004, p. 2227 - 2230.
- [9] TOMASIN, S., GOROKHOV, A., YANG, H., LINNARTZ, J. P. Iterative interference cancellation and channel estimation for mobile OFDM. *IEEE Transactions on Wireless Communications*, 2005, vol. 4, no. 1, p. 238 - 245.
- [10] VISINTIN, M. Karhunen-Loeve expansion of a fast Rayleigh fading process. *IEEE Electronics Letters*, 1996, vol. 32, no. 8, p. 1712 - 1713.
- [11] ZEMEN, T., MECKLEMBRÄUKER, C. F. Time-variant channel estimation using discrete prolate spheroidal sequences. *IEEE Transactions on Signal Processing*, 2005, vol. 53, no. 9, p. 3597 - 3607.
- [12] KEY, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs (NJ, USA): Prentice-Hall, 1993.
- [13] HUANG, L., LONG, T., MAO, E., SO, H. C. MMSE-based MDL method for robust estimation of number of sources without eigen-decomposition. *IEEE Transactions on Signal Processing*, 2009, vol. 57, no. 10, p. 4135 - 4142.
- [14] HUANG, L., LONG, T., MAO, E., SO H. C. MMSE-based MDL method for accurate source number estimation. *IEEE Signal Processing Letters*, 2009, vol. 16, no. 9, p. 798 - 801.
- [15] TADMOR, E. Filters, mollifiers and the computation of the Gibbs phenomenon. *Acta Numerica*, 2007, vol. 16, p. 305 - 378.
- [16] HRYCAK, T., DAS, S., MATZ, G., FEICHTINGER, H. Practical estimation of rapidly varying channels for OFDM systems. *IEEE Transactions on Communications*, 2011, vol. 59, no. 11, p. 3040 - 3048.
- [17] HRYCAK, T., DAS, S., MATZ, G. Inverse methods for reconstruction of channel taps in OFDM systems. *IEEE Transactions on Signal Processing*, 2012, vol. 60, no. 5, p. 2666 - 2671.
- [18] TAO, C., QIU, J., LIU, L. A novel OFDM channel estimation algorithm with ICI mitigation over fast fading channels. *Radioengineering*, 2010, vol. 19, no. 2, p. 347 - 355.
- [19] MOSTOFI, Y., COX, D. ICI mitigation for pilot-aided OFDM mobile systems. *IEEE Transactions on Wireless Communications*, 2005, vol. 4, no. 2, p. 765 - 774.
- [20] CHOI, Y. S., VOLTZ, P. J., CASSARA, F. A. On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels. *IEEE Transactions on Communications*, 2001, vol. 49, no. 8, p. 1375 - 1387.
- [21] WANG, H. W., LIN, D. W., SANG, T. H. OFDM signal detection in doubly selective channels with whitening of residual intercarrier interference and noise. *IEEE Journal on Selected Areas in Communications*, 2012, vol. 30, no. 4, p. 684 - 694.
- [22] SEBESTA, V., MARSALEK, R., FEDRA, Z. OFDM signal detector based on cyclic autocorrelation function and its properties. *Radioengineering*, 2011, vol. 20, no. 4, p. 926 - 931.
- [23] LI, R., HUANG, L., SHI, Y., SO, H. C. Gershgorin disk-based robust spectrum sensing for cognitive radio. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Florence (Italy), 2014, p. 7328 - 7332.
- [24] HIJAZI, H., ROS, L. Polynomial estimation of time-varying multipath gains with intercarrier interference mitigation in OFDM systems. *IEEE Transactions on Vehicular Technology*, 2009, vol. 58, no. 1, p. 140 - 151.
- [25] PANAYIRCI, E., SENOL, H., POOR, H. V. Joint channel estimation, equalization, and data detection for OFDM systems in the presence of very high mobility. *IEEE Transactions on Signal Processing*, 2010, vol. 58, no. 8, p. 4225 - 4238.
- [26] ABOUTORAB, N., HARDJAWANA, W., VUCETIC, B. A new iterative Doppler-assisted channel estimation joint with parallel ICI cancellation for high-mobility MIMO-OFDM systems. *IEEE Transactions on Vehicular Technology*, 2012, vol. 61, no. 4, p. 1577 - 1589.
- [27] HARDJAWANA, W., LI, R., VUCETIC, B., LI, Y. A new iterative channel estimation for high mobility MIMO-OFDM systems. In *Proceedings of IEEE 71st Vehicular Technology Conference (VTC 2010 - Spring)*. Taipei (Taiwan), 2010, p. 1 - 5.
- [28] LI, D. H., FUKUSHIMA, M. A modified BFGS method and its global convergence in nonconvex minimization. *Journal of Computational and Applied Mathematics*, 2001, vol. 129, p. 15 - 35.
- [29] AGRAWAL, R., CIOFFI, J. M. Beamforming design for the MIMO downlink for maximizing weighted sum-rate. In *Proceedings of International Symposium on Information Theory and Its Applications (ISITA)*. Auckland (New Zealand), 2008, p. 1 - 6.
- [30] NOCEDAL, J., WRIGHT, S. J. *Numerical Optimization*. New York: Springer-Verlag, 2006.
- [31] HUANG, J. Y., WANG, P., SHEN, X. F. CRLBs for pilot-aided channel estimation in OFDM system under Gaussian and non-Gaussian mixed noise. *Radioengineering*, 2012, vol. 21, no. 4, p. 1117 - 1124.
- [32] ZHENG, Y. R., XIAO, C. Simulation models with correct statistical properties for Rayleigh fading channels. *IEEE Transactions on Communications*, 2003, vol. 51, no. 6, p. 920 - 928.
- [33] HUANG, L., SO, H. C., QIAN, C. Volume-based method for spectrum sensing. *Digital Signal Processing*, 2014, vol. 28, p. 48 - 56.

- [34] SCHNITER, P. Low-complexity equalization of OFDM in doubly selective channels. *IEEE Transactions on Signal Processing*, 2004, vol. 52, p. 1002 - 1011.
- [35] KANNU, A. P., SCHNITER, P. MSE-optimal training for linear time-varying channels. *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 2005, p. 789 - 792.
- [36] PÄTZOLD, M. *Mobile Fading Channels*. New York: Wiley-IEEE Press, 2002.

About Authors ...

Min HUANG was born in 1985. He received the B.S. and M.E. degrees in Electrical Communication from Xidian Uni-

versity, China in 2008 and 2010, respectively. Now, he is working toward the doctor degree in the State Key Laboratory of Integrated Service Networks, Xidian University. His research interest includes wireless communication, digital signal processing.

Bingbing LI was born in 1955. He received the Ph.D. degree in Communication and Information Systems from Xidian University in 1995. He is now a professor of State Key Laboratory of ISN at the Department of Communication Engineering in Xidian University. He has worked at the DTV standardization as a DTV technical expert. His research interests include digital video broadcasting systems and mobile communication techniques.