www.ietdl.org

Published in IET Communications
Received on 28th July 2013
Revised on 3rd December 2013
Accepted on 29th December 2013
doi: 10.1049/iet-com.2013.0665

ISSN 1751-8628

Stable adaptive sparse filtering algorithms for estimating multiple-input–multiple-output channels

Guan Gui, Fumiyuki Adachi

Department of Communications Engineering, Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan
E-mail: gui@mobile.ecei.tohoku.ac.jp

Abstract: Channel estimation problem is one of the key technical issues for broadband multiple-input–multiple-output (MIMO) signal transmission. To estimate the MIMO channel, a standard least mean square (LMS) algorithm was often applied to adaptive channel estimation because of its low complexity and stability. The sparsity of the broadband MIMO channel can be exploited to further improve the estimation performance. This observation motivates us to consider adaptive sparse channel estimation (ASCE) methods using sparse LMS (ASCE-LMS) algorithms. However, conventional ASCE methods have two main drawbacks: (i) sensitivity to random scaling of training signal and (ii) poor estimation performance in low signal-to-noise ratio (SNR) regime. The former drawback is tackled by proposing novel ASCE-NLMS algorithms. ASCE-NLMS mitigates interference of random scale of training signal and therefore it improves its algorithm stability. It is well-known that stable sparse normalised least-mean fourth (NLMF) algorithms can achieve better estimation performance than sparse NLMS algorithms. Therefore the authors propose an improved ASCE method using sparse NLMF algorithms (ASCE-NLMF) to improve the estimation performance in low SNR regime. Simulation results show that the proposed ASCE methods are shown to achieve better performance than conventional methods, that is, ASCE-LMS by computer simulations. Also, the stability of the proposed methods is confirmed by theoretical analysis.

1 Introduction

To achieve high spectral efficiency, broadband multiple-input–multiple-output (MIMO) signal transmission is one of the mainstream techniques in the next generation cellular communications systems [1–3]. The MIMO technology utilises multiple antennas to increase the transmission rate by spatial multiplexing or improve the reliability of communication by spatial diversity [4]. A typical example is employing a very large number of antennas at base station to make highly reliable data communication possible with very low transmit power in a frequency-selective fading channel [5]. The accurate estimation of channel impulse response (CIR) is a crucial and challenging issue in coherent modulation and its accuracy has a significant impact on the overall performance of communication system. Therefore inaccurate channel state information (CSI) deteriorates the aforementioned benefits. Since broadband signal propagates over frequency-selective fading channel, one of the critical challenges of MIMO communications is accurate CSI estimation. The basic channel estimation problem is to estimate the multiple channels seen by each receive antenna. Besides, in a high mobility environment, the MIMO channel is subject to time-variant fading (i.e. doubly-selective fading).

In last decades, a number of channel estimation methods were proposed for MIMO-orthogonal frequency division multiplexing (OFDM) systems [6–14], which are grouped into two categories.

The first category encompasses the linear channel estimation methods, for example, least squares (LS) algorithm, which are based on the assumption of dense CIRs [15]. The performance of linear methods depends only on the size of MIMO channel. Note that narrowband MIMO channels are often modelled as dense channel model because of two main reasons: signal transmission over frequency-flat fading (short delay spread) and low-speed analog-to-digital converter (ADC) sampling rate at the receiver [16]. As the number of channel taps sampled is very small and most of taps are non-zero. In contrast to the narrowband MIMO channel, broadband MIMO channel is often modelled as sparse channel model [17–19] because of two main reasons: signal transmission over frequency-selective fading (long delay spread) and high-speed ADC sampling at the receiver [16]. As the number of channel taps sampled is very large and also high-speed ADC sampling generates much more channel taps than low-speed ADC sampling even in the same time delay-spread. However, most of channel taps are zeros or below the noise floor and only very few channel taps are non-zero. For a better understanding of the difference between dense channel model and sparse one, a typical example of two channel models is shown in Fig. 1. The linear channel estimation methods are known to be relatively simple to implement because of their low computation complexity [4–9].

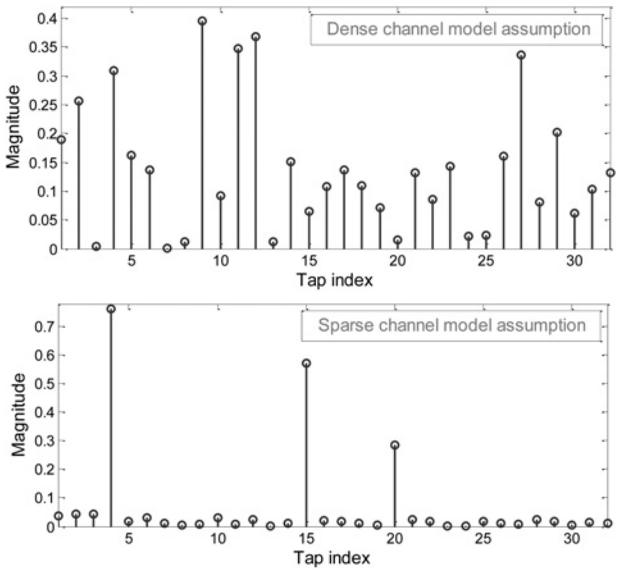

Fig. 1 *Two typical channel model assumptions: dense and sparse*

The second category encompasses the sparse channel estimation methods which use compressive sensing (CS) [20–22]. Optimal sparse channel estimation often requires its training signal to satisfy restrictive isometry property (RIP) [23]. However, designing the RIP-satisfied training signal is a non-polynomial (NP) hard problem [24]. The alternative suboptimal sparse channel estimation methods [12–14] have been proposed by utilising several empirical training matrices, for example, random Gaussian matrix, which satisfy RIP with a high probability [23]. However, there exist several stable estimation methods at the cost of extra high computational burden, especially in time-variant MIMO systems. For example, sparse channel estimation method using Dantzig selector was proposed for doubly-selective fading MIMO systems [13]. This method requires linear programming and hence it incurs high computational complexity. To reduce complexity, sparse channel estimation methods using greedy iterative algorithms were proposed in [12, 14]. However, their computational complexity depends on the number of non-zero taps of MIMO channel.

Unfortunately, the aforementioned methods can neither estimate the channel adaptively nor track the time-variant channel effectively according to system requirements. To estimate time-variant channel, adaptive sparse channel estimation (ASCE) methods using sparse least mean square (ASCE-LMS) algorithms were proposed in [25]. However, conventional ASCE-LMS methods have two main drawbacks: (i) sensitivity to random scale of training signal and (ii) instability in low signal-to-noise ratio (SNR) regime. To tackle the first drawback, in this paper, we first propose a kind of novel ASCE methods using normalised LMS (ASCE-NLMS) filtering algorithms for estimating broadband MIMO channels. In addition, since normalised least mean fourth (NLMF) filtering algorithm outperforms the well-known NLMS [26] by achieving a better balance between complexity and estimation accuracy. Likewise, we propose another kind of ASCE methods using sparse NLMF (ASCE-NLMF) filtering algorithms [27] to achieve much better estimation performance than ASCE-NLMS. In our previous research [28], a stable sparse NLMF algorithm was proposed, which achieves a better estimation performance than sparse NLMS algorithm [29] in single-input–single-output channel. Furthermore, simulation results were only considered to confirm the effectiveness of our proposed algorithms in [29]. Different from previous works, in this paper, we propose stable adaptive filtering algorithms for estimation MIMO channels and validate their effectiveness via theoretical analysis and computer simulation. The lower bounds which depend on the number of non-zero channel taps and the number of antennas are derived. Then, the superiority of the proposed algorithms is verified by computer simulations in terms of mean-square error (MSE) and bit error rate (BER).

The remainder of this paper is organised as follows. A MIMO-OFDM system model is described and problem formulation is given in Section 2. In Section 3, sparse NLMS and sparse NLMF algorithms are introduced and ASCE in MIMO-OFDM system is highlighted. Selected computer simulation results are given in Section 4. Section 5 closes the paper by summarising the main results and drawing some conclusions.

Notations: Throughout the paper, matrices and vectors are represented by boldface upper case letters and boldface lower case letters, respectively; the superscripts $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{-1}$ denote the transpose, the Hermitian transpose, and inverse operators, respectively; $\|\boldsymbol{h}\|_0$ is the ℓ_0-norm operator that counts the number of non-zero taps in $\boldsymbol{h}$ and $\|\boldsymbol{h}\|_p$ stands for the ℓ_p-norm operator which is computed by $\|\boldsymbol{h}\|_p = \left(\sum_i |h_i|^p\right)^{1/p}$, where $p \in (0, 2]$ is considered in this paper; $E\{\cdot\}$ denotes the expectation operator.

2 System model and problem formulation

We consider a time-variant MIMO-OFDM system. Let us denote the number of transmit antennas by N_t, the number of subcarriers by K, and the maximum delay of the channel by N. Frequency-domain signal vector at time t, $\bar{\boldsymbol{x}}_{n_t}(t) = \left[\bar{x}_{n_t}(t, 0), \bar{x}_{n_t}(t, 1), \ldots, \bar{x}_{n_t}(t, K-1)\right]^{\mathrm{T}}$, $n_t = 1, 2, \ldots, N_t$ is fed to inverse discrete Fourier transform (IDFT) at the n_tth antenna. Here we assume that the transmit power is $E\{\|\bar{\boldsymbol{x}}_{n_t}\|_2^2\} = KE_{\mathrm{s}}$ and E_{s} is the average symbol power. The resultant vector $\boldsymbol{x}_{n_t}(t) \triangleq \boldsymbol{F}^{\mathrm{H}}\bar{\boldsymbol{x}}_{n_t}(t)$ is padded with cyclic prefix (CP) of length $L_{\mathrm{CP}} \geq (N-1)$ to avoid inter-block interference, where $\boldsymbol{F}$ is a $K \times K$ DFT matrix with entries $[\boldsymbol{F}]_{kq} = 1/K\mathrm{e}^{-\mathrm{j}2\pi kq/K}$, $k, q = 0, 1, \ldots, K-1$. The signal is received by antenna after propagating the frequency-selective fading channel. To simplify the problem of ASCE for MIMO-OFDM systems, without loss of generality, only one receive antenna is considered in the following discussion, that is, $n_r = 1$. It is straightforward to extend to the multiple receive antenna case. Let us drop the time index t. The received signal y after CP removal and stacked input signal vector $\boldsymbol{x} = [\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}, \ldots, \boldsymbol{x}_{N_t}^{\mathrm{T}}]^{\mathrm{T}}$ are related to each other as

$$y = \sum_{n_t=1}^{N_t} \boldsymbol{h}_{n_t}^{\mathrm{T}}\boldsymbol{x}_{n_t} + z = \boldsymbol{h}^{\mathrm{T}}\boldsymbol{x} + z \qquad (1)$$

where $z \sim \mathcal{CN}(0, \sigma_n^2)$ is an additive noise Gaussian variable. The MIMO channel vector $\boldsymbol{h}$ is written as

$$\boldsymbol{h} = [\,\boldsymbol{h}_1^{\mathrm{T}} \quad \boldsymbol{h}_2^{\mathrm{T}} \quad \cdots \quad \boldsymbol{h}_{N_t}^{\mathrm{T}}\,]^{\mathrm{T}} \qquad (2)$$

where $\boldsymbol{h}_{n_t}(n_t = 1, 2, \ldots, N_t)$ is assumed to be an equal N-length sparse channel vector from the n_tth transmit antenna to the receive antenna. Each channel vector $\boldsymbol{h}_{n_t}$ is assumed to be only supported by T dominant channel taps. A typical example of sparse multipath channel with $N = 16$, which is supported by $T = 3$ dominant channel taps, is depicted in Fig. 2. The ASCE is performed in an iterative fashion as illustrated in Fig. 2. Let us consider the nth channel estimation update. Following the system model in (1), the corresponding channel estimation error $e(n)$ is defined as

$$e(n) := y - \tilde{y}(n) = y - \tilde{\boldsymbol{h}}^{\mathrm{T}}(n)\boldsymbol{x} \tag{3}$$

where $\tilde{\boldsymbol{h}}(n)$ denotes an estimate of MIMO channel $\boldsymbol{h}$ at the nth update and $\tilde{y}(n)$ is the output signal. The goal of ASCE is to estimate the MIMO channel vector $\boldsymbol{h}$ using error signal $e(n)$ and input training signal $\boldsymbol{x}$. Traditional ASCE methods using LMS (ASCE-LMS) algorithms were proposed to exploit channel sparsity. Specifically, the cost function of ASCE-LMS methods is formulated as [29]

$$L_s(n) = \frac{1}{2}e^2(n) + \lambda_{slp}\left\|\tilde{\boldsymbol{h}}(n)\right\|_p \tag{4}$$

where $0 \le p < 1$ and $\lambda_{slp} \ge 0$ denotes the sparse regulation parameter which trades off the MSE and sparsity of $\tilde{\boldsymbol{h}}(n)$. Fig. 3 shows a geometrical interpretation of (4). Sparse structure of channel vector $\tilde{\boldsymbol{h}}(n)$ could be exploited when $\lambda_{slp} > 0$ and it reduces to standard LMS algorithm when $\lambda_{slp} = 0$. It is worth mentioning that exploiting sparsity depends on the cost function in a sense that whether or not there exists a convex point (unique solution) between solution plane (many solutions) and sparse penalty function. In Fig. 3*a*, there is no unique solution as the ℓ_2-norm constraint cannot find the convex point in the solution plane. On the other hand, in Fig. 3*b*, unique solution is obtained by ℓ_1-norm constraint function. Without loss of generality, corresponding update equation of ASCE-LMS methods is written as

$$\begin{aligned}\tilde{\boldsymbol{h}}(n+1) &= \tilde{\boldsymbol{h}}(n) - \mu_s\frac{\partial L_s(n)}{\partial \tilde{\boldsymbol{h}}(n)} \\ &= \tilde{\boldsymbol{h}}(n) + \mu_s e(n)\boldsymbol{x} - \rho_{slp}\frac{\left\|\tilde{\boldsymbol{h}}(n)\right\|_p^{1-p}\operatorname{sgn}\left(\tilde{\boldsymbol{h}}(n)\right)}{\varepsilon + \left|\tilde{\boldsymbol{h}}(n)\right|^{1-p}}\end{aligned} \tag{5}$$

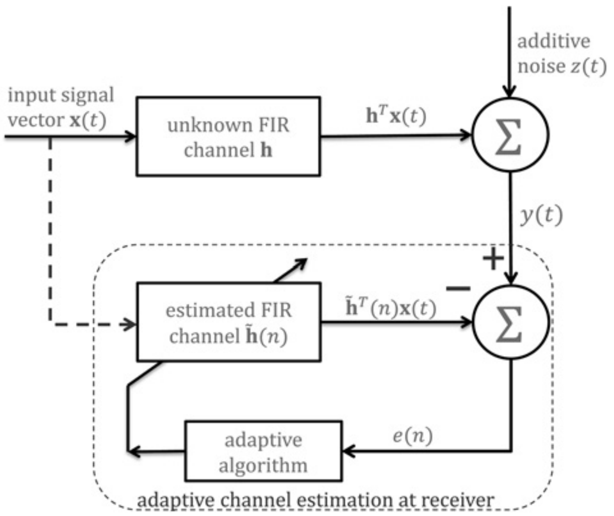

Fig. 2 *Diagram of adaptive MIMO channel estimation method*

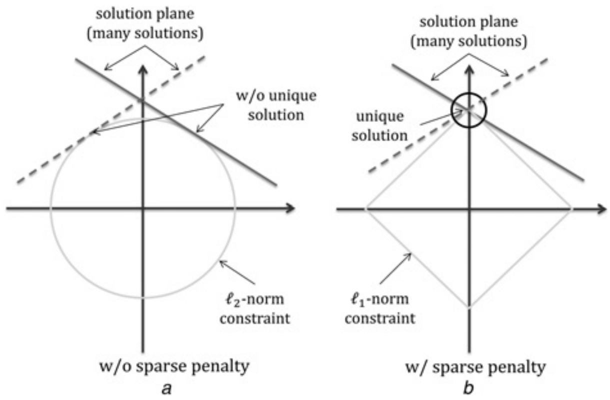

Fig. 3 *Geometrical interpretation of standard LMS and sparse LMS algorithm*

a Case with many solutions without sparse penalty ($\lambda_{slp} = 0$)
b Case with unique solution with sparse penalty ($\lambda_{slp} > 0$)

where $\rho_{slp} = \mu_s\lambda_{slp}$ and $\mu_s \in (0, \gamma_{\max}^{-1})$ is the step size of LMS gradient descend; where ε is a positive parameter to ensure bound of last term in (5); and $\gamma_{\max}$ is the maximum eigenvalue of the covariance matrix $\boldsymbol{R} = E\{\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\}$.

3 Estimating MIMO channels using adaptive filtering algorithms

3.1 *ASCE-NLMS methods*

In this subsection, we introduce two ASCE methods using sparse NLMS filtering algorithms, which are termed as ASCE-NLMS. To exploit channel sparsity, two different sparse penalties (i.e. ℓ_p-norm and ℓ_0-norm) are considered for cost functions. Basically, sparse penalties can compel the small channel coefficients to zero with high probability [19] so that the ASCE-NLMS exploits chanenl sparsity.

Firstly, we introduce ASCE using ℓ_p-norm NLMS (ASCE-LP-NLMS) filtering algorithm. Based on (4) and (5), update equation of ASCE-LP-NLMS is given by

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) + \mu_s\frac{e(n)\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} - \rho_{slp}\frac{\left\|\tilde{\boldsymbol{h}}(n)\right\|_p^{1-p}\operatorname{sgn}\left(\tilde{\boldsymbol{h}}(n)\right)}{\varepsilon + \left|\tilde{\boldsymbol{h}}(n)\right|^{1-p}} \tag{6}$$

Setting $p = 0$, it is degenerates into the ℓ_0-norm NLMS (L0-NLMS) [29] and the cost function of ASCE-L0-NLMS becomes

$$L_{sl0}(n) = \frac{1}{2}e^2(n) + \lambda_{sl0}\left\|\tilde{\boldsymbol{h}}(n)\right\|_0 \tag{7}$$

where λ_{sl0} is a regularisation parameter to balance the estimation error and sparse penalty. Since solving the ℓ_0-norm minimisation is an NP-hard problem [24], we replace the ℓ_0-norm with an approximate continuous function as [30]

$$\|\boldsymbol{h}\|_0 \simeq \sum_{l=0}^{N_tN-1} \left(1 - e^{-\beta|h_l|}\right) \tag{8}$$

According to the approximate function, L0-LMS cost function can be revised as

$$L_{sl0}(n) = \frac{1}{2}e^2(n) + \lambda_{sl0} \sum_{l=0}^{N_tN-1} \left(1 - e^{-\beta|h_l|}\right) \tag{9}$$

Then, the update equation of ASCE-L0-NLMS can be derived as

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) + \mu_s e(n)\boldsymbol{x} - \rho_{sl0}\beta \mathrm{sgn}\left(\tilde{\boldsymbol{h}}(n)\right) e^{-\beta|\tilde{\boldsymbol{h}}(n)|} \tag{10}$$

where $\rho_{sl0} = \mu_s\lambda_{sl0}$. It is worth mentioning that the exponential function in (10) has high computational complexity. To reduce the computational complexity, its first-order Taylor series expansion is considered as [30]

$$e^{-\beta|h|} \simeq \begin{cases} 1 - \beta|h|, & \text{when } |h| \le 1/\beta \\ 0, & \text{others} \end{cases} \tag{11}$$

where h is any element of channel vector $\boldsymbol{h}$. Finally, the update equation of L0-NLMS filtering algorithm based ASCE can be derived as

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) + \frac{e(n)\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} - \rho_{sl0} G_{l0}\left(\tilde{\boldsymbol{h}}(n)\right) \tag{12}$$

where $G_{l0}(\tilde{\boldsymbol{h}}(n))$ is defined by

$$G_{l0}(\boldsymbol{h}) = \begin{cases} 2\beta^2 h - 2\beta \mathrm{sgn}(h), & \text{when } |h| \le 1/\beta \\ 0, & \text{others} \end{cases} \tag{13}$$

3.2 ASCE-NLMF methods

An improved ASCE-NLMF filtering algorithm is introduced in this subsection. The proposed method is motivated by the fact that standard LMF filtering algorithm can achieve lower error performance bound than LMS [31]. Following the pioneering work [31], a cost function $L_f(n)$ of standard LMF filtering algorithm is constructed as

$$L_f(n) = \frac{1}{4}e^4(n) \tag{14}$$

The update equation of ACE using LMF filtering algorithms is given by

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) - \mu_f \frac{\partial L_f(n)}{\partial \tilde{\boldsymbol{h}}(n)} = \tilde{\boldsymbol{h}}(n) + \mu_f e^3(n)\boldsymbol{x} \tag{15}$$

where $\mu_f \in (0, 2)$ is a gradient descend step-size controlling convergence speed and steady-state performance. However, the LMF filtering algorithm has several stability problems that may put a limitation to its use in applications. The stability of the LMF filtering algorithm depends highly on several uncertain factors: initial setting of the adaptive filter weights, input power of the adaptive filter, noise power and unbounded regressors [27]. To improve the stability of the LMF algorithm, Eweda proposed a NLMF filtering algorithm [27]. Then, the update equation of (ACE-NLMF) filtering algorithm is given by

$$\begin{aligned} \tilde{\boldsymbol{h}}(n+1) &= \tilde{\boldsymbol{h}}(n) + \mu_f \frac{e^3(n)\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2\left(\|\boldsymbol{x}\|_2^2 + e^2(n)\right)} \\ &= \tilde{\boldsymbol{h}}(n) + \mu_f(n)\frac{e(n)\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} \end{aligned} \tag{16}$$

where $\mu_f(n) = \mu_f e^2(n)/(\|\boldsymbol{x}\|_2^2 + e^2(n))$ denotes a variable step-size. For a better understanding the property of the variable step-size, detailed discussion is given. Here, we observe that when $e^2(n) \gg \|\boldsymbol{x}\|_2^2$, then $\mu_f(n) \to \mu_f$; when $e^2(n) \simeq \|\boldsymbol{x}\|_2^2$, then $\mu_f(n) \to \mu_f/2$; when $e^2(n) \ll \|\boldsymbol{x}\|_2^2$, then $\mu_f \to 0$. Hence, NLMF algorithm in (16) is stable which is equivalent to NLMS algorithm in (6). The detail interpretation about equivalence between them is given in the Appendix. Assuming the statistical time is sufficiently large so that $\|\boldsymbol{x}\|_2^2 = N\sigma_x^2$, variable step-size $\mu_f(n)$ can be rewritten as

$$\mu_f(n) = \frac{\mu_f}{N\sigma_x^2/e^2(n) + 1} \tag{17}$$

Above equation describes relationship between $\mu_f(n)$ and square error of the received signal $e^2(n)$. Equation (17) shows that $\mu_f(n)$ is an increasing function over $e^2(n)$ as illustrated in Fig. 4. The variable step-size $\mu_f(n)$ decreases to ensure stability of sparse NLMF as $e^2(n)$ decreases.

This behaviour also coincides with standard adaptive signal processing theory [26]. Generally speaking, in the case of large error, large step-size is preferred to accelerate the gradient descend speed (convergence rate); in the case of small error, small step-size is adopted to obtain accurate estimation. Variable step-size used in our scheme can balance well between channel estimation performance and stability. According to the previous research in [29], if the standard NLMS algorithm is stable, then its corresponding ASCE method using sparse NLMS algorithm is also stable. The main reason is that stability of NLMS has no direct

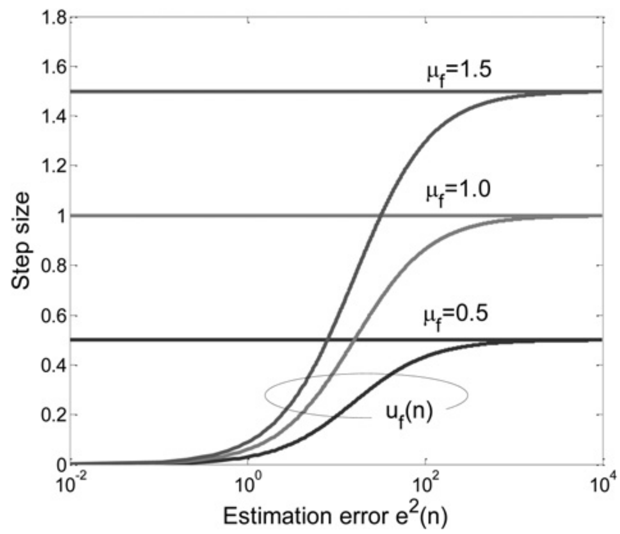

Fig. 4 *Step-size $\mu_f(n)$ is adaptive variable to ensure NLMF-based algorithms global stable $\mu_f \in \{0.5, 1.0, 1.5\}$ and the square of received signal error $e^2(n) \in (0.01, 10^4)$*

relation to sparse penalty strength. Indeed, its estimation performance depends on sparse penalty functions. Without loss of generality, if the standard NLMF algorithm is stable, then its corresponding sparse NLMF algorithm is also stable. The detail stability analysis of NLMF is given in the Appendix. In the sequel, we propose several reliable ASCE-NLMF approaches.

Different from ACE-NLMF [27] neglecting channel structure information, we propose a ASCE method using LP-NLMF (ASCE-LP-NLMS) filtering algorithm to exploit sparsity in MIMO channel so that it can further obtain performance gain. First, the cost function of ASCE-LP-NLMF is constructed as

$$L_{flp}(n) = \frac{1}{4}e^4(n) + \lambda_{flp}\left\|\tilde{\boldsymbol{h}}(n)\right\|_p \tag{18}$$

where λ_{flp} is a regularisation parameter which trades off the fourth-order mismatching estimation error and ℓ_p-norm sparse penalty of $\tilde{\boldsymbol{h}}(n)$. The update equation of ASCE-LP-NLMS can be derived as

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) + \mu_f(n)\frac{e(n)\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} - \rho_{flp}\frac{\left\|\tilde{\boldsymbol{h}}(n)\right\|_p^{1-p}\operatorname{sgn}\left(\tilde{\boldsymbol{h}}(n)\right)}{\varepsilon + \left|\tilde{\boldsymbol{h}}(n)\right|^{1-p}} \tag{19}$$

where $\rho_{flp} = \mu_f\lambda_{flp}$ depends on gradient descend step-size μ_f and regularisation parameter λ_{flp}. Similarly, cost function of ASCE method using L0-LMF algorithm is written as

$$L_{fl0}(n) = \frac{1}{4}e^4(n) + \lambda_{fl0}\left\|\tilde{\boldsymbol{h}}(n)\right\|_0 \tag{20}$$

where $\lambda_{fl0} > 0$ is a regularisation parameter which trades off the fourth-order mismatching estimation error and sparseness of MIMO channel. The corresponding updating equation of ASCE method using L0-NLMF filtering algorithm (ASCE-L0-NLMF) is given by

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) + \mu_f(n)\frac{e(n)\boldsymbol{x}}{\|\boldsymbol{x}\|_2^2} - \beta_2 g_{L0}\left(\tilde{\boldsymbol{h}}(n)\right) \tag{21}$$

where $\beta_2 = \mu_f\lambda_{fl0}$ and $g_{L0}(\tilde{\boldsymbol{h}}(n))$ is an approximate sparse ℓ_0-norm function which is defined in (15). The Cramér-Rao lower bound (CRLB) of ASCE is derived in Section 4. By utilising known channel position information, CRLB of ASCE becomes $\text{CRLB}_{\text{sparse}} \sim \mathcal{O}(T)$.

4 Simulation results and discussions

In this section, we evaluate our proposed ASCE estimators. The 1000 independent Monte Carlo runs is averaged for evaluation. The length of channel vector $\boldsymbol{h}_{n_t}$ between each transmit and receive antenna is set to $N = 16$ and the number of dominant taps is set to $T = 1$ and 3. The distribution of dominant channel taps follows Gaussian distribution, which is subjected to $E\{\|\boldsymbol{h}_{n_t}\|_2^2\} = 1$, and their positions are randomly selected within the length of $\boldsymbol{h}_{n_t}$. Received SNR is defined by E_S/N_0, where E_S is the average received power of symbol and N_0 is the noise power. Here, we set the SNR as 3, 6 and 9 dB. The step sizes and regularisation parameters are listed in Table 1. The estimation performance is evaluated by average MSE metric which is defined as Average $\text{MSE}\{\tilde{\boldsymbol{h}}(n)\} = E\{\|\boldsymbol{h} - \tilde{\boldsymbol{h}}(n)\|_2^2\}$, where $\boldsymbol{h}$ and $\tilde{\boldsymbol{h}}(n)$ are the actual MIMO channel vector and its estimate at the nth update, respectively. Note that the initial channel estimator $\tilde{\boldsymbol{h}}(0) = \boldsymbol{0}$ is considered in computer simulation.

Table 1 Simulation parameters and their values

Parameters	Values
number of transmit antennas	$N_t = 2$
signal transmission scheme	OFDM
channel fading	random AWGN channel
gradient descend step-size: μ_s	0.5
gradient descend step-size: μ_f	1.5
regularisation parameter: λ_{slp}	$2\times10^{-4}\sigma_n^2\log(N/T)$
regularisation parameter: λ_{flp}	$2\times10^{-6}\sigma_n^2\log(N/T)$
regularisation parameter: λ_{sl0}	$2\times10^{-3}\sigma_n^2\log(N/T)$
regularisation parameter: λ_{fl0}	$2\times10^{-5}\sigma_n^2\log(N/T)$

In the first example, the proposed methods are evaluated in Fig. 5 ($T = 1$) and Fig. 6 ($T = 3$) at SNR = 3 dB. The step-size of sparse NLMS algorithms and sparse NLMF algorithms are set as $\mu_s = 0.5$ [29] and $\mu_f = 1.5$, respectively. As the two

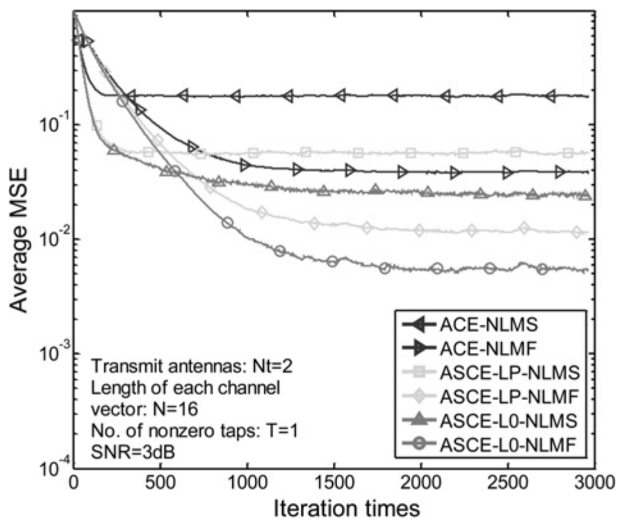

Fig. 5 *Performance comparison against iteration times*

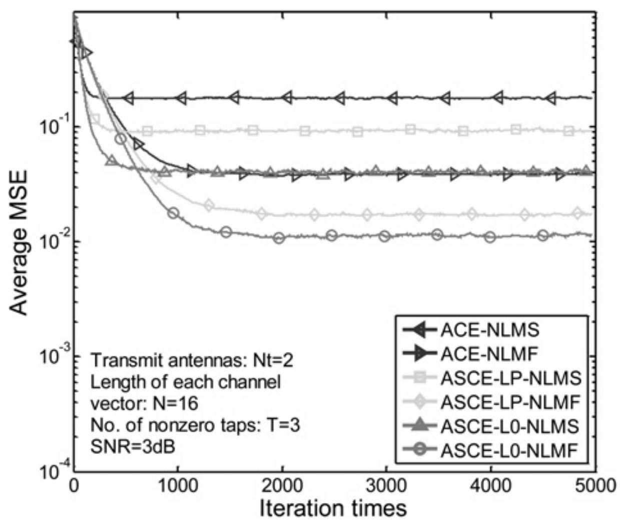

Fig. 6 *Performance comparison against iteration times*

doi: 10.1049/iet-com.2013.0665

figures show, ASCE-NLMS method achieves better estimation performance than ACE-NLMS. Similarly, ASCE-NLMF method also achieves better estimation performance than ACE-NLMF method. We can also observe that ASCE-NLMF method outperforms the ASCE-NLMS method significantly but at the cost of higher computational complexity (much more iteration times). Relatively, the computational complexity of ASCE-NLMS is very low [29].

In the second evaluation, the proposed methods are evaluated at SNR regimes 6 and 9 dB as shown in Figs. 7 and 8, respectively. Once again, we can noted improvement over conventional methods. Please note that computational complexity of ASCE-NLMF method increases with SNR. The main reason is that higher SNR provides smaller received signal error square which reduces the variable step-size $\mu_f(n)$ for adaptive updating.

In the third evaluation, in order to shed light on the entire system performance in terms of error probability, numerical simulation was adopted to evaluate the average BER in Fig. 9. Indeed, the evaluation of average BER can be quite cumbersome because it depends on the bit-to-symbol mapping used [1]. To avoid the high computation, here, a simple BER evaluation method via invertible exponential-type approximations is adopted [32]. For the multilevel phase shift keying (PSK) modulation and multilevel quadrature amplitude modulation (QAM) schemes, their average BERs can be computed by

$$\mathrm{BER}_{M-\mathrm{PSK}} = a_1 \mathrm{e}^{-b\gamma_s \sin^2(\pi/M)} + a_2 \mathrm{e}^{-2b\gamma_s \sin^2(\pi/M)}, \quad \text{for } M \geq 4 \tag{22}$$

and

$$\begin{aligned}\mathrm{BER}_{M-\mathrm{QAM}} =& 2ka_1 \mathrm{e}^{-(1.5b\gamma_s/M-1)} + (2ka_2 - k^2a_1^2)\mathrm{e}^{-(3b\gamma_s/M-1)} \\ & - k^2a_2^2\mathrm{e}^{-(6b\gamma_s/M-1)} - 2k^2a_1a_2\mathrm{e}^{-(4.5b\gamma_s/M-1)}, \\ & \text{for } k = (\sqrt{M}-1)/\sqrt{M}\end{aligned} \tag{23}$$

respectively, where $a_1 = 0.3017$, $a_2 = 0.438$ and $b = 1.0510$ are the optimal curve-fitting coefficients; M is the multilevel

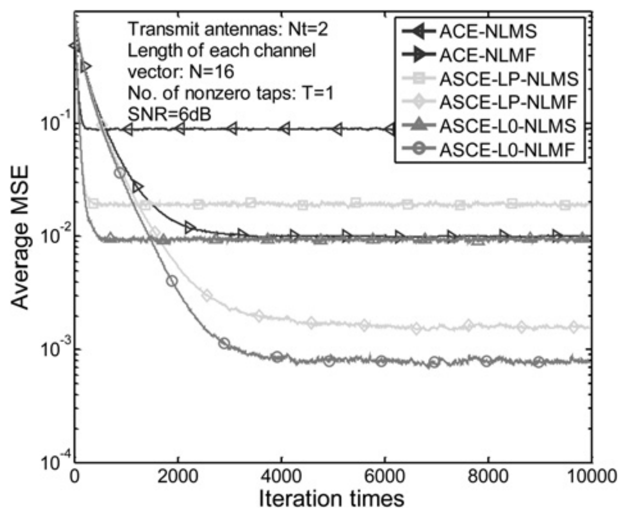

Fig. 7 *Performance comparison against iteration times*

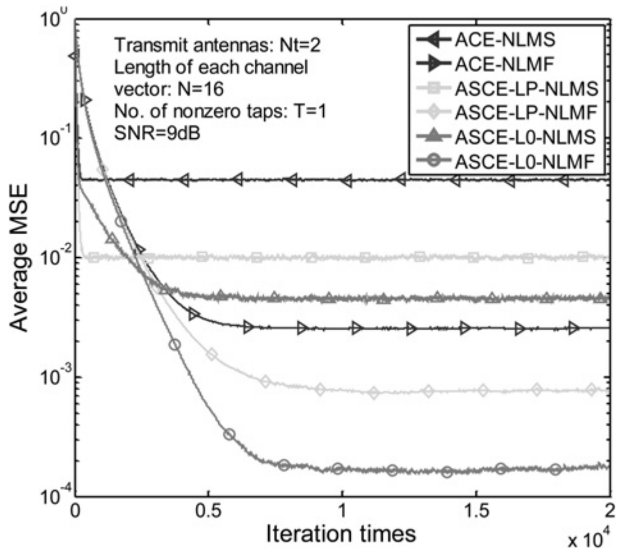

Fig. 8 *Performance comparison against iteration times*

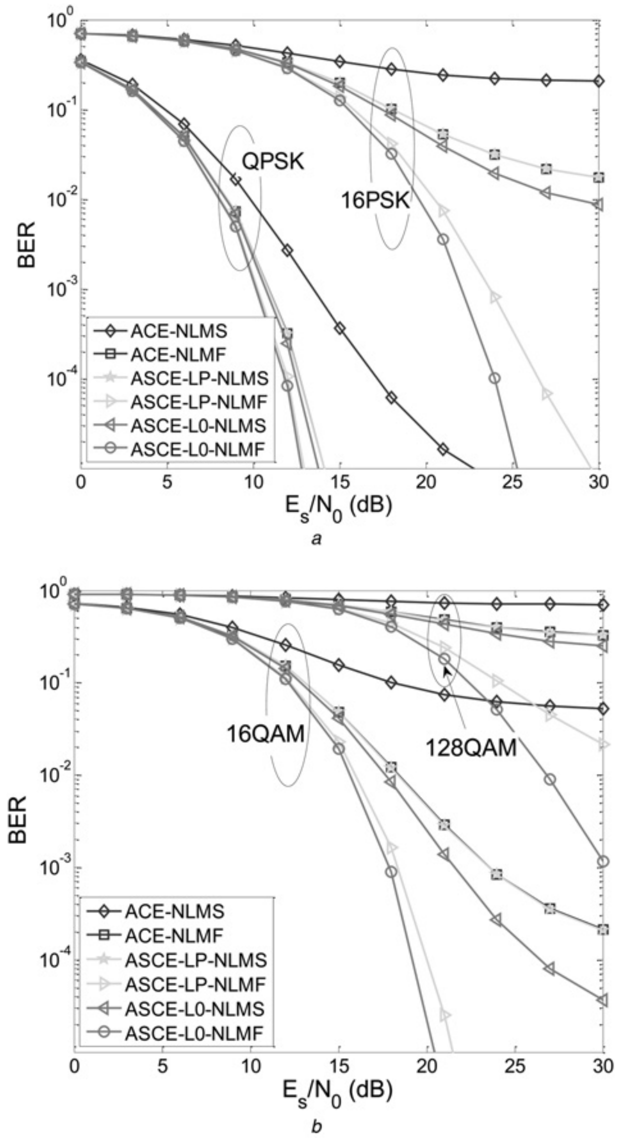

Fig. 9 *BER performance in different SNRs*

a Multiple PSK modulations
b Multiple QAM modulations

of modulation and γ_s is defined as

$$\gamma_s = \frac{10^{\mathrm{SNR}/10}(1 - \mathrm{MSE})}{10^{\mathrm{SNR}/10}\mathrm{MSE} + 1} \tag{24}$$

Different signal modulation schemes, that is, multiple PSKs, multiple QAMs, are considered. In addition, steady channel estimators are considered as in the case of SNR = 6 dB and number non-zero taps of each subchannel vector (e.g. $\boldsymbol{h}_{n_t}$) is 1. It is found that proposed sparse NLMF-type methods outperforms sparse NLMS-type methods. According to Fig. 9, this result could be predicted even if different modulation schemes are used. Let us take the 16-PSK modulation scheme, for example, as shown in Fig. 9*a*. Since the steady channel estimators are obtained in low SNR regime, that is, SNR = 6 dB, average BER can achieve 10^{-5} by using the proposed steady channel estimators. Hence, the proposed methods are effective in various modulation schemes based communication systems even in the low SNR environment.

5 Conclusion and future works

In this paper, we proposed ASCE methods using sparse NLMS and sparse NLMF algorithms for frequency-selective fading MIMO-OFDM systems. Firstly, system model was formulated. Secondly, cost functions of the two proposed methods were constructed using different sparse penalties, that is, ℓ_p-norm and ℓ_0-norm. Simulation results indicate that the proposed methods achieve a better MSE performance than the standard ACE-NLMS method without much increase in computational complexity. Furthermore, the average BER performance of ASCE-NLMS and ASCE-NLMF methods was evaluated assuming different modulation schemes. The simulation results also confirmed that the proposed ASCE-NLMF methods are even better than ASCE-NLMS methods in terms of MSE and BER metrics but have a higher amount of computation complexity in terms of iteration times.

6 Acknowledgment

The authors would like to think Dr. Koichi Adachi of Institute for Infocomm Research (I2R) for his valuable comments and suggestions. This work was supported by grant-in-aid for the Japan Society for the Promotion of Science (JSPS) fellows grant number 24-02366.

7 References

1 Raychaudhuri, D., Mandayam, N.B.: 'Frontiers of wireless and mobile communications', *Proc. IEEE*, 2012, **100**, (4), pp. 824–840
2 Adachi, F., Kudoh, E.: 'New direction of broadband wireless technology', *Wirel. Commun. Mob. Comput.*, 2007, **7**, (8), pp. 969–983
3 Dai, L., Wang, Z., Yang, Z.: 'Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems', *IEEE J. Sel. Areas Commun.*, 2013, **31**, (2), pp. 251–263
4 Sayeed, A.M.: 'Deconstructing multiantenna fading channels', *IEEE Trans. Signal Process.*, 2002, **50**, (10), pp. 2563–2579
5 Persson, D.B., Lau, K., Lau, B.K., *et al.*: 'Scaling up MIMO: Opportunities and challenges with very large arrays', *IEEE Signal Process. Mag.*, 2013, **30**, (1), pp. 40–60
6 Gibson, J.D., Iltis, R.A.: 'Channel estimation and data detection for MIMO-OFDM systems'. IEEE Global Telecommunications Conf. (GLOBECOM), San Francisco, USA, December 2003, pp. 581–585
7 Biguesh, M., Gershman, B.: 'Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals', *IEEE Trans. Signal Process.*, 2006, **54**, (3), pp. 884–893
8 Chang, T., Chiang, W., Hong, Y.P., *et al.*: 'Training sequence design for discriminatory channel estimation in wireless MIMO systems', *IEEE Trans. Signal Process.*, 2010, **58**, (12), pp. 6223–6237
9 He, S., Tugnait, J.K., Meng, X.: 'On superimposed training for MIMO channel estimation and symbol detection', *IEEE Trans. Signal Process.*, 2007, **55**, (6), pp. 3007–3021
10 Pham, T.-H., Liang, Y., Nallanathan, A.: 'A joint channel estimation and data detection receiver for multiuser MIMO IFDMA systems', *IEEE Trans. Commun.*, 2009, **57**, (6), pp. 1857–1865
11 Wang, Z.J., Han, Z., Liu, K.J.R.: 'A MIMO-OFDM channel estimation approach using time of arrivals', *IEEE Trans. Wirel. Commun.*, 2005, **4**, (3), pp. 1207–1213
12 Taubock, G., Hlawatsch, F., Eiwen, D., *et al.*: 'Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing', *IEEE J. Sel. Top. Signal Process.*, 2010, **4**, (2), pp. 255–271
13 Bajwa, W.U., Haupt, J., Sayeed, A.M., *et al.*: 'Compressed channel sensing: a new approach to estimating sparse multipath channels', *Proc. IEEE*, 2010, **98**, (6), pp. 1058–1076
14 Wang, N., Gui, G., Zhang, Z., *et al.*: 'A novel sparse channel estimation method for multipath MIMO-OFDM systems'. IEEE 74th Vehicular Technology Conf. (VTC-Fall), San Francisco, California, USA, September 2011, pp. 1–5
15 Lee, S.J.: 'On the training of MIMO-OFDM channels with least square channel estimation and linear interpolation', *IEEE Commun. Lett.*, 2008, **12**, (2), pp. 100–102
16 Gui, G., Peng, W., Adachi, F.: 'Sub-Nyquist rate ADC sampling-based compressive channel estimation', *Wirel. Commun. Mob. Comput.*, 2013, article first publihed online, doi.10.1002/wcm2372
17 Zhou, Y., Sayeed, A.M.: 'Experimental study of MIMO channel statistics and capacity via virtual channel representation'. UW Tech. Report, http//www.dune.ece.wisc.edu/pdfs/zhoumeas.pdf
18 Czink, N., Yin, X., Özcelik, H., *et al.*: 'Cluster characteristics in a MIMO indoor propagation environment', *IEEE Trans. Wirel. Commun.*, 2007, **6**, (4), pp. 1465–1475
19 Vuokko, L., Kolmonen, V.-M., Salo, J., *et al.*: 'Measurement of large-scale cluster power characteristics for Geometric channel models', *IEEE Trans. Antennas Propag.*, 2007, **55**, (11), pp. 3361–3365
20 Candes, E.J., Romberg, J., Tao, T.: 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information', *IEEE Trans. Inf. Theory*, 2006, **2**, (2), pp. 489–509
21 Donoho, D.L.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 1289–1306
22 Dai, L., Wang, Z., Yang, Z.: 'Compressive sensing based time domain synchronous OFDM transmission for vehicular communications', *IEEE J. Sel. Areas Commun.*, 2013, **31**, (9), pp. 460–469
23 Candes, E.J.: 'The restricted isometry property and its implications for compressed sensing', *Comptes Rendus Math.*, 2008, **1**, (346), pp. 589–592
24 Amini, A., Unser, M., Marvasti, F.: 'Compressibility of deterministic and random infinite sequences', *IEEE Trans. Signal Process.*, 2011, **59**, (99), pp. 5193–5201
25 Taheri, O., Vorobyov, S.A.: 'Sparse channel estimation with Lp-norm and reweigthted L1-norm penalized least mean square'. IEEE Int. Conf. on Acoustics Speech and Signal Processing (ICASSP), Prague, Czech Republic, May 2011, pp. 2864–2867
26 Widrow, B., Stearns, D.: 'Adaptive signal processing' (Prentice-Hall, New Jersey, 1985)
27 Eweda, E.: 'Global stabilization of the least mean fourth algorithm', *IEEE Trans. Signal Process.*, 2012, **60**, (3), pp. 1473–1477
28 Gui, G., Adachi, F.: 'Adaptive sparse system identification using normalized least-mean fourth algorithm', *Int. J. Commun. Syst.*, 2013, article first published online, doi: 10.1002/dac.2637
29 Gui, G., Peng, W., Adachi, F.: 'Improved adaptive sparse channel estimation based on the least mean square algorithm'. IEEE Wireless Communications and Networking Conf. (WCNC), Shanghai, China, April 2013, pp. 3130–3134
30 Gu, Y., Jin, J., Mei, S.: 'L0-norm constraint LMS algorithm for sparse system identification', *IEEE Signal Process. Lett.*, 2009, **16**, (9), pp. 774–777
31 Walach, E., Widrow, B.: 'The least mean fourth (LMF) adaptive algorithm and its family', *IEEE Trans. Inf. Theory*, 1984, **30**, (2), pp. 275–283
32 Gui, G., Adachi, F.: 'Sparse LMF algorithm for adaptive channel estimation in low SNR region', *Int. J. Commun. Syst.*, 2013, article first published online, doi: 10.1002/dac.2531

33 Olabiyi, O., Annamalai, A.: 'Invertible exponential-type approximations for the Gaussian probability integral', *IEEE Wirel. Commun. Lett.*, 2012, **1**, (5), pp. 544–547
34 Eweda, E., Bershad, N.J.: 'Stochastic analysis of a stable normalized least mean fourth algorithm for adaptive noise canceling with a white Gaussian reference', *IEEE Trans. Signal Process.*, 2012, **60**, (12), pp. 6235–6244
35 Hübscher, P.I., Bermudez, J.C.M., Nascimento, V.H.: 'A mean-square stability analysis of the least mean fourth adaptive algorithm', *IEEE Trans. Signal Process.*, 2007, **55**, (8), pp. 4018–4028

8 Appendix

8.1 Comparisons between NLMS-type and NLMF-type algorithms

To improve the stability of LMF, Eweda and Bershad in [27] proposed a global stable NLMF algorithm with update equation constructed by

$$\tilde{\boldsymbol{h}}(n+1) = \tilde{\boldsymbol{h}}(n) + \frac{\mu_f e^3(n)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}\left(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} + e^2(n)\right)} \tag{25}$$

where step-size μ_f should satisfy $\mu_f \in (0, 2)$ to ensure the proposed NLMF stability. Please note that the behaviour of NLMF depends highly on the $e^2(n)$. If $e^2(n) \gg \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$, then the behaviour of NLMF is reduced to the behaviour of standard NLMS, that is,

$$\lim_{e^2(n) \gg \boldsymbol{x}^{\mathrm{H}}(t)\boldsymbol{x}(t)} \frac{\mu_f e^3(n)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}\left(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} + e^2(n)\right)} = \frac{\mu_f e(n)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}} \tag{26}$$

where $e^2(n)$ dominates in the $(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} + e^2(n))$. Likewise, if $e^2(n) \ll \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$, then the behaviour of NLMF is reduce to the behaviour of traditional NLMS in (3), that is,

$$\lim_{e^2(n) \ll \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}} \frac{\mu_f e^3(n)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}\left(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} + e^2(n)\right)} = \frac{\mu_f(n) e(n)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}} \tag{27}$$

where $\mu_f(n) = \mu_f e^2(n)/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$ is a variable step-size of gradient descend. As the $e^2(n) \ll \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$, $\mu_f(n) = e^2(n)/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} \to 0$. That is to say, the variable step-size $\mu_f(n)$ can trade off stability and convergence rate of proposed algorithm in each update. Combining (26) and (27), the stability of NLMF-type algorithms approaches to the well-known NLMS-type algorithms. Note that this stable behaviour of NLMF coincides with our proposed algorithm in [33]. In the case of $\mu_f = \mu_s = \mu$, it is worth mentioning that variable step-size $\mu_f(n)$ of NLMF is always smaller than μ_s due to the fact that

$$\mu_f(n) = \frac{\mu_f e^2(n)}{\|\boldsymbol{x}\|_2^2 + e^2(n)} = \frac{\mu_f}{\|\boldsymbol{x}\|_2^2/e^2(n) + 1} < \mu \tag{28}$$

for any $\|\boldsymbol{x}\|_2^2/e^2(n) > 0$. Hence, one can find that the computational complexity of stable NLMF is always higher than either LMF or NLMS. □

8.2 Lower bound of ASCE-NLMF channel estimator

To derive the approximate CRLB of ASCE-NLMF channel estimator $\tilde{\boldsymbol{h}}(n)$, assuming the position set Ω of non-zero channel taps is known. First of all, we define the nth adaptive updating error $\boldsymbol{v}_\Omega(n)$ as

$$\boldsymbol{v}_\Omega(n) = \tilde{\boldsymbol{h}}_\Omega(n) - \boldsymbol{h}_\Omega \tag{29}$$

According to (3), ideal received signal error of non-zero channel taps can be written as

$$\begin{aligned} e_\Omega(n) &= y - y(n) \\ &= y - \tilde{\boldsymbol{h}}_\Omega(n)\boldsymbol{x}_\Omega \\ &= z - \tilde{\boldsymbol{h}}_\Omega(n)\boldsymbol{x}_\Omega \end{aligned} \tag{30}$$

Then $(n+1)$th updating channel error $\boldsymbol{v}_\Omega(n+1)$ of optimal sparse NLMF can be written as

$$\begin{aligned} \boldsymbol{v}_\Omega(n+1) &= \boldsymbol{v}_\Omega(n) + \frac{e_\Omega^2(n)}{\|\boldsymbol{x}_\Omega\|_2^2 + e_\Omega^2(n)} \cdot \frac{\mu_f e_\Omega(n)\boldsymbol{x}_\Omega}{\|\boldsymbol{x}_\Omega\|_2^2} \\ &\simeq \boldsymbol{v}_\Omega(n) + \gamma\left(e_\Omega^2(n)\right)\frac{\mu_f e_\Omega(n)\boldsymbol{x}_\Omega}{\|\boldsymbol{x}_\Omega\|_2^2} \end{aligned} \tag{31}$$

where $\|\boldsymbol{x}_\Omega\|_2^2 = T\sigma_x^2$ and T is the number of non-zero channel taps of $\tilde{\boldsymbol{h}}(n)$. According to (31), the lower bound can be achieved when $e_\Omega^2(n) \ll \|\boldsymbol{x}_\Omega\|_2^2$, that is, $\gamma(e_\Omega^2(n)) \simeq e_\Omega^2(n)/T\sigma_x^2$. Taking the expectation, we have

$$\begin{aligned} E\left\{\gamma\left(e_\Omega^2(n)\right)\right\} &= \frac{E\{e_\Omega^2(n)\}}{T\sigma_x^2} \\ &= \frac{E\{z^2\} + E\{\boldsymbol{x}_\Omega^{\mathrm{H}}\boldsymbol{x}_\Omega\}E\{\boldsymbol{v}_\Omega^{\mathrm{H}}(n)\boldsymbol{v}_\Omega(n)\}}{T\sigma_x^2} \\ &= \frac{\sigma_z^2}{T\sigma_x^2} + MSE(n) \ll 1 \end{aligned} \tag{32}$$

where $\mathrm{MSE}(n) \ll 1 - \sigma_z^2/T\sigma_x^2$. For simplification, we consider $\|\boldsymbol{x}_\Omega\|_2^2 = 1$. Hence, the $(n+1)$th adaptive updating MSE performance of ASCE-NLMF channel estimator is given by (see equation (33) at the bottom of the next page) Since $\boldsymbol{x}$ and $\boldsymbol{v}(n)$ are independent, therefore, $E\{\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{v}_\Omega(n)\} = 0$ and $E\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{v}(n)\} = 0$ according to [34, 35], we can obtain

$$E\left\{\left[\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{v}_\Omega(n)\right]^2\right\} \simeq \sigma_x^2 E\left\{\boldsymbol{v}_\Omega^{\mathrm{T}}(n)\boldsymbol{v}_\Omega(n)\right\} \tag{34}$$

$$\begin{aligned} &E\left\{\left[\boldsymbol{x}_\Omega^{\mathrm{T}}(t)\boldsymbol{v}_\Omega(n)\right]^2\boldsymbol{x}_\Omega^{\mathrm{T}}(t)\boldsymbol{x}_\Omega(t)\right\} \\ &\quad \simeq (T+2)\sigma_x^4 E\left\{\boldsymbol{v}_\Omega^{\mathrm{T}}(n)\boldsymbol{v}_\Omega(n)\right\} \end{aligned} \tag{35}$$

$$E\left\{\left[\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{v}_\Omega(n)\right]^4\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{x}_\Omega\right\} \simeq (3T+12)\sigma_x^6 E^2\left\{\boldsymbol{v}_\Omega^{\mathrm{T}}(n)\boldsymbol{v}_\Omega(n)\right\} \tag{36}$$

$$E\left\{\left[\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{v}_\Omega(n)\right]^6\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{x}_\Omega\right\} \simeq (15T+90)\sigma_x^8 E^3\left\{\boldsymbol{v}_\Omega^{\mathrm{T}}(n)\boldsymbol{v}_\Omega(n)\right\} \tag{37}$$

$$E\left\{\boldsymbol{x}_\Omega^{\mathrm{T}}\boldsymbol{x}_\Omega\right\} \simeq T\sigma_x^2 \tag{38}$$

For any random Gaussian-distribution noise z, according to the paper [34], $E\{z^4\} = 3\sigma_z^4$ and $E\{z^6\} = 15\sigma_z^6$. Then

MSE$(n+1)$ in (33) is simplified to

$$\begin{aligned}\text{MSE }(n+1) \simeq & \left[1-6\mu_f\sigma_z^2\sigma_x^2+45(T+2)\mu_f^2\sigma_z^4\sigma_x^4\right]\text{MSE }(n)\\ & +\left[135(T+4)\mu_f^2\sigma_z^2\sigma_x^6-6\mu_f\sigma_x^4\right]\text{MSE}^2(n)\\ & +15(T+6)\mu_f^2\sigma_x^8\text{MSE}^3(n)+15T\mu_f^2\sigma_z^6\sigma_x^2\end{aligned} \tag{39}$$

Since nth adaptive update error MSE(n) is very small, using high-order approximations $\lim_{n\to\infty}\text{MSE}^2(n)=0$ and $\lim_{n\to\infty}\text{MSE}^3(n)=0$. Then, we can obtain

$$\begin{aligned}\text{CRLB}_{\text{Sparse}} & =\lim_{n\to\infty}\text{MSE }(n+1)\\ & =\left[1-6\mu_f\sigma_z^2\sigma_x^2+45(T+2)\mu_f^2\sigma_z^4\sigma_x^4\right]\text{CRLB}_{\text{Sparse}}\\ & \quad +15T\mu_f^2\sigma_z^6\sigma_x^2\end{aligned} \tag{40}$$

Based on (37), the $\text{CRLB}_{\text{Sparse}}$ is derived by

$$\text{CRLB}_{\text{Sparse}}=\frac{5T\mu_f\sigma_z^4}{2-15(T+2)\mu_f\sigma_x^2\sigma_z^2} \tag{41}$$

Please note that the $\text{CRLB}_{\text{Sparse}}$ of ASCE-NLMF channel estimator depends on the number of non-zero channel taps T, noise variance σ_z^2, input signal variance σ_x^2 and step-size μ_f. It is well known that when the step-size $\mu_f\to 0$, $\text{CRLB}_{\text{Sparse}}$ can be further derived by

$$\text{CRLB}_{\text{Sparse}}\geq 2.5T\mu_f\sigma_z^4\tilde{\mathcal{O}}(T) \tag{42}$$

Note that $\lim_{\mu_f\to 0}(2-15(T+2)\mu_f\sigma_x^2\sigma_z^2)=2$. Likewise, when the CIR satisfies dense distribution then its $\text{CRLB}_{\text{Dense}}$ is derived by

$$\text{CRLB}_{\text{Dense}}\geq 2.5N\mu_f\sigma_z^4\tilde{\mathcal{O}}(N) \tag{43}$$

Comparing (42) with (43), we can understand that the channel estimator has a direct relation with the number of non-zero channel taps. The two lower bounds can also explain well why sparse methods can improve estimation performance better than the traditional methods which neglect the channel sparse structure.

$$\begin{aligned}\text{MSE}(n+1) & =E\left\{\boldsymbol{v}_\Omega^{\text{T}}(n+1)\boldsymbol{v}_\Omega(n+1)\right\}\\ & =E\left\{\left[\boldsymbol{v}_\Omega(n)+\mu_f e_\Omega^3(n)\boldsymbol{x}_\Omega\right]^{\text{T}}\left[\boldsymbol{v}_\Omega(n)+\mu_f e_\Omega^3(n)\boldsymbol{x}_\Omega\right]\right\}\\ & =E\left\{\left[\boldsymbol{v}_\Omega(n)+\mu_f\left(z-\boldsymbol{v}_\Omega^{\text{T}}(n)\boldsymbol{x}_\Omega\right)^3\boldsymbol{x}_\Omega(n)\right]^{\text{T}}\left[\boldsymbol{v}_\Omega(n)+\mu_f\left(z-\boldsymbol{v}_\Omega^{\text{T}}(n)\boldsymbol{x}_\Omega\right)^3\boldsymbol{x}_\Omega\right]\right\}\\ & =E\left\{\boldsymbol{v}_\Omega^{\text{T}}(n)\boldsymbol{v}_\Omega(n)\right\}-2\mu_f E\left\{\left[\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{v}_\Omega(n)\right]^4\right\}-6\mu_f E\left\{z^2\right\}E\left\{\left[\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{v}_\Omega(n)\right]^2\right\}\\ & \quad+\mu_f^2 E\left\{\left[\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{v}_\Omega(n)\right]^6\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{x}_\Omega\right\}+15\mu_f^2 E\left\{z^2\right\}E\left\{\left[\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{v}_\Omega(n)\right]^4\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{x}_\Omega\right\}\\ & \quad+15\mu_f^2 E\left\{z^4\right\}E\left\{\left[\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{v}_\Omega(n)\right]^2\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{x}_\Omega\right\}+\mu_f^2 E\left\{z^6\right\}E\left\{\boldsymbol{x}_\Omega^{\text{T}}\boldsymbol{x}_\Omega\right\}\end{aligned} \tag{33}$$

IET Commun., 2014, Vol. 8, Iss. 7, pp. 1032–1040
doi: 10.1049/iet-com.2013.0665