

SPECIAL ISSUE PAPER

Normalized least mean square-based adaptive sparse filtering algorithms for estimating multiple-input multiple-output channels

Guan Gui^{1*}, Li Xu¹ and Fumiyuki Adachi²¹ Department of Electronics and Information Systems, Akita Prefectural University, Akita, Japan² Department of Communications Engineering, Graduate School of Engineering, Tohoku University, Sendai, Japan

ABSTRACT

This paper studies normalized least mean square-based adaptive sparse filtering algorithms for estimating multiple-input multiple-output (MIMO) channels. Although the MIMO channel is often modeled as sparse, traditional normalized least mean square-based filtering algorithm never takes the advantage of the inherent sparse structure information and thus causes some performance loss. Unlike the traditional method, the proposed two adaptive sparse channel estimation methods exploit the sparse structure information of MIMO channels. To validate the effectiveness of proposed MIMO channel estimates, theoretical analysis and simulation results are provided. We derive steady-state mean-square deviations of the proposed MIMO channel estimates and theoretically show that it is better than the traditional one. Moreover, their performance advantages are confirmed by computer simulations. Copyright © 2014 John Wiley & Sons, Ltd.

KEYWORDS

adaptive filtering algorithm; normalized least mean square (NLMS); ℓ_p -norm NLMS; ℓ_0 -norm NLMS; adaptive sparse channel estimation (ASCE); compressed sensing (CS)

*Correspondence

Guan Gui, Department of Electronics and Information Systems, Akita Prefectural University, Akita, Japan.

E-mail: guiguan@akita-pu.ac.jp

1. INTRODUCTION

The use of multiple-input multiple-output (MIMO) transmission (as shown in Figure 1) and orthogonal frequency division multiplexing (OFDM) makes high data communications over frequency-selective fading channels [1–3]. The accurate estimation of finite impulse response channel is a crucial and challenging issue in coherent modulation, and its accuracy has a significant impact on the overall system performance.

During the last decades, a number of channel estimation methods have been proposed for MIMO-OFDM systems [4–12]. These methods can be categorized into two types. The first type is linear channel estimation methods, for example, least squares algorithm [5,6], which is based on the assumption of dense channel impulse responses (CIRs). The second type is sparse channel estimation methods [11–13] using compressive sensing [14,15], which is based on the assumption of sparse CIRs.

In the linear channel estimation methods, the mean square error (MSE) performance depends on size of MIMO channel matrix only [11]. Note that the narrowband MIMO

channel may be modeled as the dense CIR because of its very short time delay spread; however, the broadband MIMO channel is often modeled as a sparse CIR [13,16–19]. A typical example of sparse CIR is shown in Figure 2. It is well-known that linear channel estimation methods are relatively simple to implement because of its low computational complexity [4–9]. However, the main drawback of linear channel estimation methods is the inability to exploit the inherent channel sparsity. Different from the linear channel estimation methods, the sparse channel estimation methods take advantage of the sparsity of the channel [11,20,21]. The optimal sparse channel estimation often requires circulant matrix of training signal to satisfy restrictive isometry property [22]. However, designing the restrictive isometry property-satisfied training matrix is a nonpolynomial hard problem [23]. Although some compressive sensing algorithms achieve stable sparse channel estimation in high probability [11,20,21], these algorithms often incur extra high computational burden, especially in fast fading communication systems. For example, one of the typical sparse channel estimations methods, using Dantzig selector algorithm [24], has been

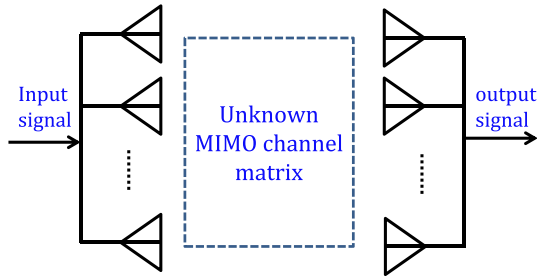


Figure 1. Signal transmission over a MIMO channel.

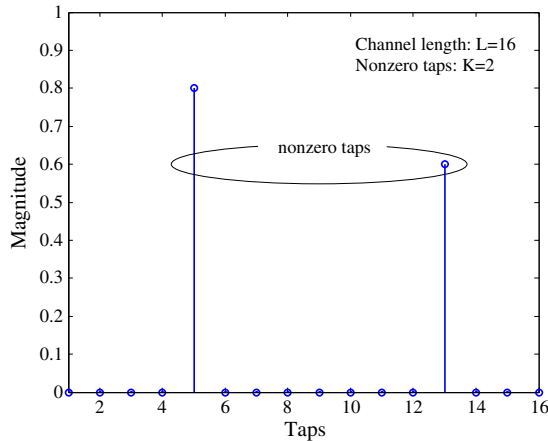


Figure 2. A typical example of sparse channel.

proposed for double-selective fading MIMO systems in [11]. However, Dantzig selector algorithm needs to be solved by linear programming (LP), and hence, it requires high computational complexity [11]. To reduce the complexity, sparse channel estimation methods using greedy iterative algorithms have been proposed in [10,12]. However, their complexity still depends on the number of nonzero taps of the MIMO channel due to the larger number of nonzero taps in the MIMO channel.

To exploit the channel sparsity while without sacrificing complexity, Chen *et al.* have proposed an effective sparse least mean square (LMS) algorithm using an ℓ_1 -norm sparse penalty [25]. Taheri *et al.* have proposed an ℓ_p -norm LMS (LP-LMS)-based adaptive sparse channel estimation (ASCE) method to further exploit the channel sparsity in single-antenna systems [26]. However, ASCE using the sparse LMS filtering algorithm is vulnerable to the random scaling of input signal. To fully take advantage of channel sparsity and to improve stability of estimation method, we have proposed ASCE that combines normalized LMS (NLMS) filtering algorithms and sparse constraints, for example, ℓ_p -norm and ℓ_0 -norm, for estimating single-antenna time-variant channels [27]. They are termed as ℓ_p -norm NLMS (LP-NLMS) and ℓ_0 -norm NLMS (L0-NLMS), respectively. To the best of our knowledge, ASCE methods for estimating MIMO channels have not been developed. To estimate the MIMO channel, in

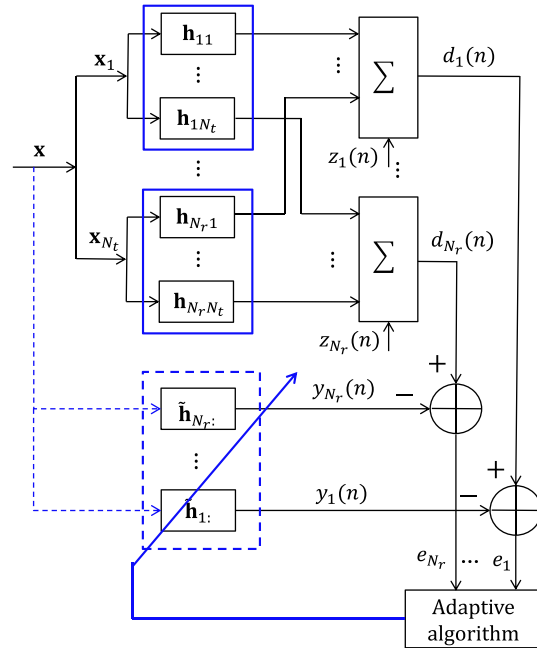


Figure 3. ASCE for estimating MIMO channels.

this paper, we propose MIMO-ASCE methods with LP-NLMS and L0-NLMS [27]. First, as shown in Figure 3, a typical MIMO system model is formulated so that each multiple-input single-output (MISO) channel vector can be estimated by ASCE methods. Second, steady-state mean square deviation (MSD) performance of proposed channel estimate is derived. Later, computer simulation results are presented to confirm the effectiveness of our proposed methods.

The remainder of the paper is organized as follows. A MIMO-OFDM system model is described and problem formulation is given in Section 2. In Section 3, the NLMS-based adaptive sparse filtering algorithm is introduced, and the proposed ASCE using sparse NLMS filtering algorithms for estimating MIMO channels is highlighted. In addition, performances of ASCE methods are compared analytically. Computer simulation results are given in Section 4 in order to evaluate and compare performances of the ASCE methods. Finally, we conclude the paper in Section 5.

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

2. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a time-variant MIMO-OFDM communication system as shown in Figure 1. At time index t , frequency-domain signal vector at the n_r -th antenna $\bar{\mathbf{x}}_{n_r}(t) = [\bar{x}_{n_r}(t, 0), \dots, \bar{x}_{n_r}(t, N-1)]^T$, $n_r = 1, 2, \dots, N_r$ is fed to inverse discrete Fourier transform, where N is the number of subcarriers. Assume that the transmit power is $E\{\bar{\mathbf{x}}_{n_r}(t)^2\} = NE_0$, where E_0 denotes unit power. The resultant vector $\mathbf{x}_{n_r}(t) \triangleq \mathbf{F}^H \bar{\mathbf{x}}_{n_r}(t)$ is padded with cyclic prefix of length $L_{CP} \geq (N-1)$ to avoid inter-block interference, where \mathbf{F} is a $N \times N$ discrete Fourier transform matrix with entries $[\mathbf{F}]_{cq} = 1/Ne^{-j2\pi cq/N}$, $c, q = 0, 1, \dots, N-1$. The time-domain signal is transmitted through length N channel and received by multiple antennas at the receiver. After cyclic prefix removal, the signal vector received by the n_r -th antenna at time t is written as y_{n_r} . Then, the ideal received signal vector $\mathbf{d} = [d_1, d_2, \dots, d_{N_r}]^T$ and input signal $\mathbf{x}(n)$ are related by

$$\mathbf{d} = \mathbf{H}\mathbf{x}(n) + \mathbf{z}(n) \quad (1)$$

where the MIMO channel matrix \mathbf{H} can be written as

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_{11}^T & \mathbf{h}_{12}^T & \cdots & \mathbf{h}_{1N_r}^T \\ \mathbf{h}_{21}^T & \mathbf{h}_{22}^T & \cdots & \mathbf{h}_{2N_r}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{N_r 1}^T & \mathbf{h}_{N_r 2}^T & \cdots & \mathbf{h}_{N_r N_r}^T \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{1:}^T \\ \mathbf{h}_{2:}^T \\ \vdots \\ \mathbf{h}_{N_r:}^T \end{bmatrix} \quad (2)$$

Notice that N is the size of filtering memory of each single channel between each antenna pair $h_{n_r n_r}$. Then, the ideal received signal at n_r -th antenna can be written as

$$d_{n_r} = \sum_{n_t=1}^{N_r} \mathbf{h}_{n_r n_t}^T \mathbf{x}_{n_t}(n) + z_{n_r}(n) = \mathbf{h}_{n_r:}^T \mathbf{x}(n) + z_{n_r}(n) \quad (3)$$

where $\mathbf{h}_{n_r:}^T = [\mathbf{h}_{n_r 1}^T, \dots, \mathbf{h}_{n_r n_r}^T, \dots, \mathbf{h}_{n_r N_r}^T] \in \mathbb{C}^{1 \times N_r N}$, $n_r = 1, 2, \dots, N_r$ is a MISO channel vector that consists of N_r single-input single-output subchannels $\mathbf{h}_{n_r n_t}$, ($n_t = 1, 2, \dots, N_r$). We assume that the $\mathbf{h}_{n_r n_t}$ is only supported by K -dominant channel taps whose positions are randomly determined. A typical example of sparse multipath channel is depicted in Figure 2. Hereby, at the n_r -th receive antenna, the corresponding signal estimation error e_{n_r} for $n_r = 1, 2, \dots, N_r$ at time t can be defined as

$$e_{n_r}(n) = d_{n_r} - y_{n_r}(n) = d_{n_r} - \mathbf{h}_{n_r:}^T(n) \mathbf{x}(n) \quad (4)$$

where $\mathbf{h}_{n_r:}^T(n)$ denotes the n_r -th adaptive updating estimator of $\mathbf{h}_{n_r:}^T$ and $y_{n_r}(n)$ is the output signal from NLMS filter as it is shown in Figure 3. By collecting all of the error signals $e_{n_r}(n)$, $n_r = 1, 2, \dots, N_r$, Equation (4) can be rewritten as matrix-vector form as

$$\begin{aligned} \mathbf{e}(n) &= [e_1(n), e_2(n), \dots, e_{N_r}(n)]^T \\ &= \mathbf{d} - \mathbf{y}(n) \\ &= \mathbf{d} - \mathbf{H}(n) \mathbf{x}(n) \end{aligned} \quad (5)$$

where $\mathbf{y}(n) = [y_1(n), \dots, y_{N_r}(n)]^T$ denotes estimate of the output signal; $\mathbf{H}(n)$ is the n -th adaptive estimate channel matrix \mathbf{H} . According to Equation (5), MIMO channel estimation problem is equivalent to estimating different individual MISO channel $\mathbf{h}_{n_r:}$ using error signal $e_{n_r}(n)$ and input training signal $\mathbf{x}(n)$. Estimating the MISO channel vector $\mathbf{h}_{n_r:}$, the standard LMS filtering algorithm [28] constructs the corresponding cost function:

$$L_{n_r}(n) = \frac{1}{2} e_{n_r}^2(n) \quad (6)$$

for $n_r = 1, 2, \dots, N_r$. It is obvious that LMS-based adaptive channel estimation (ACE) can be derived as

$$\begin{aligned} \mathbf{h}_{n_r:}(n+1) &= \mathbf{h}_{n_r:}(n) - \mu \frac{\partial L_{n_r:}(n)}{\partial \mathbf{h}_{n_r:}(n)} \\ &= \mathbf{h}_{n_r:}(n) + \mu \mathbf{x}(n) e_{n_r}(n) \end{aligned} \quad (7)$$

where $\mu \in (0, \gamma_{\max}^{-1})$ is the step size of LMS gradient descend and γ_{\max} is the maximum eigenvalue of the $N_r N \times N_r N$ covariance matrix, which is calculated as $\mathbf{R} = E\{\mathbf{x}(n) \mathbf{x}^T(n)\}$. The stability of LMS-based method is vulnerable to random scaling of training signal [29]. To improve the stability, NLMS filtering algorithm is considered as standard method for estimating MISO channels [30]. Hence, the update equation is modified as

$$\mathbf{h}_{n_r:}(n+1) = \mathbf{h}_{n_r:}(n) + \frac{\mu}{\underbrace{\mathbf{x}^T(n) \mathbf{x}(n)}_{\mu(n)}} \mathbf{x}(n) e_{n_r}(n) \quad (8)$$

where $\mu(n) = \mu / (\mathbf{x}^T(n) \mathbf{x}(n))$ is termed as variable step size (VSS) that depends on the random input signal $\mathbf{x}(n)$. The advantage of the NLMS filtering algorithm over the LMS filtering algorithm using invariable step size μ is briefly discussed in the following. Interested authors are recommended to refer to [31] for detailed derivation process. The $(n+1)$ -th adaptive channel estimation $\mathbf{h}_{n_r:}(n+1)$ is obtained by solving

$$\begin{aligned} \text{minimize} \quad & \mathbf{w}_{n_r:}(n) = \mathbf{h}_{n_r:}(n+1) - \mathbf{h}_{n_r:}(n) \\ \text{subject to} \quad & \mathbf{h}_{n_r:}^T(n+1) \mathbf{x}(n) = d_{n_r}(n) \end{aligned} \quad (9)$$

To solve the aforementioned equality constrained optimization problem, Lagrange duality theory is adopted [31], and then the optimal solution is obtained as

$$\tilde{\mathbf{h}}_{n_r:}(n+1) = \tilde{\mathbf{h}}_{n_r:}(n) + \frac{1}{2} \mu^* \mathbf{x}(n) \quad (10)$$

where μ^* is a Lagrange multiplier. Substituting Equation (10) into Equation (9), the ideal received signal can be rewritten as

$$\begin{aligned} d_{n_r} &= \tilde{h}_{n_r}:(n+1)x(n) \\ &= \tilde{h}_{n_r}:(n)x(n) + \frac{1}{2}\mu^*x^T(n)x(n) \end{aligned} \quad (11)$$

Then, the VSS of NLMS filtering algorithm is easily obtained as

$$\mu^* = \frac{2e(n)}{\mathbf{x}^T(n)\mathbf{x}(n)} \quad (12)$$

From Equation (12), one can find that VSS of NLMS filtering algorithm is set so that the algorithm can achieve the flexible trade-off between steady-state MSE performance and convergence speed. The positive real factor VSS μ controls the update scale from one iteration to the next without changing the direction. To ensure the stability of the NLMS filtering algorithm, the VSS μ in Equation (8) is bounded as follows [31]:

$$0 < \mu < \frac{2}{(N_t N + 2)P_x} \quad (13)$$

where $P_x = E\{x^2(n)\}$ denotes input signal power. Under two independent assumptions, (i) input signal and noise are independent, and (ii) multiple antennas are uncorrelated, steady-state MSD of MIMO channel estimate using NLMS filtering algorithm is derived as

$$\begin{aligned} \text{MSD}_{NLMS}(\infty) &= \lim_{n \rightarrow \infty} E\{(\mathbf{H}(n) - \mathbf{H})^T(\mathbf{H}(n) - \mathbf{H})\} \\ &= N_r \lim_{n \rightarrow \infty} E\{(\mathbf{h}_{n_r}:(n) - \mathbf{h}_{n_r}:)^T \\ &\quad \times (\mathbf{h}_{n_r}:(n) - \mathbf{h}_{n_r}:)\} \\ &= \frac{\mu N_r N_t N}{2 - \mu(N_t N + 2)} \cdot \frac{P_0}{P_x} \end{aligned} \quad (14)$$

where $P_0 = E\{z^2(n)\}$ denotes the additive noise power. However, traditional methods in either Equation (7) or (8) never exploit the inherent sparsity in MIMO channels and then incur waste of resources that could be utilized by advanced signal processing method. Therefore, in the next section, we propose two sparse NLMS filtering algorithms so as to exploit sparse structure information.

3. PROPOSED SPARSE NLMS FILTERING ALGORITHMS FOR ESTIMATING MIMO CHANNELS

3.1. Proposed sparse NLMS filtering algorithms

Consider an ℓ_p -norm sparse penalty on NLMS cost function to produce sparse channel estimate as this penalty

term forces the channel tap values of $h_{n_r}:$ to approach zero. It is termed as LP-NLMS, which was proposed for single-antenna systems in [26]. For the n_r -th MISO channel vector, the cost function of the LP-NLMS is given by

$$L_{lp,n_r}(n) = \frac{1}{2}e_{n_r}^2(n) + \lambda_{lp,n_r} \|\mathbf{h}_{n_r}:(n)\|_p \quad (15)$$

where $\|\cdot\|_p$ is the ℓ_p -norm operator and λ_{lp,n_r} is a regularization parameter that balances the MSE and the sparsity. According to Equation (15), the update equation for LP-NLMS-based ASCE can be derived as

$$\begin{aligned} \mathbf{h}_{n_r}:(n+1) &= \mathbf{h}_{n_r}:(n) + \underbrace{\frac{\mu}{\mathbf{x}^T(n)\mathbf{x}(n)}}_{\mu(n)} \cdot e_{n_r}(n)\mathbf{x}(n) \\ &\quad \underbrace{- \rho_{lp,n_r} \frac{\|\mathbf{h}_{n_r}:(n)\|_p^{1-p} \text{sgn}(\mathbf{h}_{n_r}:(n))}{\sigma + |h_{n_r}:(n)|^{1-p}}}_{g_1(\mathbf{h}_{n_r}:(n))} \end{aligned} \quad (16)$$

where $\rho_{lp,n_r} = \mu\lambda_{lp,n_r}$ depends on the gradient descend step size μ and the regularization parameter λ_{lp,n_r} ; $\sigma > 0$ is a given positive parameter, and $g_1(\mathbf{h}_{n_r}:(n))$ is the sparse penalty strength function.

Theorem 1. Suppose the number of iterations approaches infinity ($n \rightarrow \infty$) and $\mu(n)$ satisfies Equation (13). Then the n_r -th converged MISO channel $\mathbf{h}_{n_r}:(\infty)$ using LP-NLMS filtering algorithm can be derived as follows:

$$\begin{aligned} E\{\mathbf{h}_{n_r}:(\infty)\} &= \mathbf{h}_{n_r} - \frac{P_x \rho_{lp,n_r}}{\mu} \\ &\quad \times E\left[\frac{\|\mathbf{h}_{n_r}:(\infty)\|_p^{1-p} \text{sgn}(\mathbf{h}_{n_r}:(\infty))}{\sigma + |h_{n_r}:(\infty)|^{1-p}}\right] \end{aligned} \quad (17)$$

Proof. Let us define the channel error as $\mathbf{v}_{n_r}:(n) = \mathbf{h}_{n_r}:(n) - \mathbf{h}_{n_r}:$. According to Equation (16), $\mathbf{v}_{n_r}:(n)$ can be rewritten as

$$\begin{aligned} \mathbf{v}_{n_r}:(n) &= [\mathbf{I} - \mu(n)\mathbf{x}(n)\mathbf{x}^T(n)]\mathbf{v}_{n_r}:(n-1) + \mu(n)z_{n_r}:(n)x(n) \\ &\quad - \rho_{lp,n_r} \frac{\|h_{n_r}:(n)\|_p^{1-p} \text{sgn}(h_{n_r}:(n))}{\sigma + |h_{n_r}:(n)|^{1-p}} \end{aligned} \quad (18)$$

Taking expectation on both sides of Equation (18), one can obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} E \{ \mathbf{v}_{n_r} : (n) \} &= \lim_{n \rightarrow \infty} E \left\{ \left[\mathbf{I} - \mu(n) \mathbf{x}(n) \mathbf{x}^T(n) \right] \mathbf{v}_{n_r} : (n-1) + \mu(n) z_{n_r} : (n) \mathbf{x}(n) - \rho_{lp, n_r} \frac{\| \mathbf{h}_{n_r} : (n) \|_p^{1-p} \text{sgn}(\mathbf{h}_{n_r} : (n))}{\sigma + |h_{n_r} : (n)|^{1-p}} \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ \left[\mathbf{I} - \mu \mathbf{R} / P_x \right] E[\mathbf{v}_{n_r} : (n-1)] - \rho_{lp, n_r} E \left[\frac{\| \mathbf{h}_{n_r} : (n) \|_p^{1-p} \text{sgn}(\mathbf{h}_{n_r} : (n))}{\sigma + |h_{n_r} : (n)|^{1-p}} \right] \right\} \\ &= \left[\mathbf{I} - \mu \mathbf{R} / P_x \right] E[\mathbf{v}_{n_r} : (\infty)] - \rho_{lp, n_r} E \left[\frac{\| \mathbf{h}_{n_r} : (\infty) \|_p^{1-p} \text{sgn}(\mathbf{h}_{n_r} : (\infty))}{\sigma + |h_{n_r} : (\infty)|^{1-p}} \right] \end{aligned} \quad (19)$$

Because $E[\mathbf{h}_{n_r} : (n)] = E[\mathbf{v}_{n_r} : (n)] + \mathbf{h}_{n_r} :$, the convergence condition of $E[\mathbf{h}_{n_r} : (n)]$ is the same as $E[\mathbf{v}_{n_r} : (n)]$ in Equation (19), which is independent of ρ_{lp, n_r} , as well as the ℓ_p -norm sparse constraint function. However, according to Equation (19), an appropriate selection of ρ_{lp, n_r} and ℓ_p -norm functions for LP-NLMS filtering algorithm realizes lower MSD than NLMS [25].

By setting $p = 0$ (ℓ_0 -norm sparse constraint function) in the LP-NLMS-based ASCE, the zero-attracting forces the channel tap values of $h_{n_r} :$ to approach zero. It is termed as ℓ_0 -norm NLMS (L0-NLMS) [27], and whose cost function is given by

$$L_{l0, n_r} (n) = \frac{1}{2} e_{n_r}^2 (n) + \lambda_{l0, n_r} \| \mathbf{h}_{n_r} : (n) \|_0 \quad (20)$$

where λ_{l0, n_r} is a regularization parameter to balance the estimation error and sparse penalty. Because solving the L_0 -norm minimization is a nonpolynomial hard problem [23], we replace it with approximate continuous function

$$\| \mathbf{h}_{n_r} : \|_0 \approx \sum_{l=0}^{N_t N - 1} \left(1 - e^{-\beta |h_{n_r} : l|} \right) \quad (21)$$

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

$$L_{l0, n_r} (n) = \frac{1}{2} e_{n_r}^2 (n) + \lambda_{l0, n_r} \sum_{l=0}^{N_t N - 1} \left(1 - e^{-\beta |h_{n_r} : l|} \right) \quad (22)$$

Then, the update equation for L0-LMS-based ASCE can be derived as

$$\lim_{n \rightarrow \infty} E [g_2 (\mathbf{h}_{n_r} : (n))] = \begin{cases} 2\beta^2 h_{n_r} : l (\infty) - 2\beta E [\text{sgn}(h_{n_r} : l (\infty))], & \text{when } |h_{n_r} : l (\infty)| \leq 1/\beta \\ 0, & \text{others} \end{cases} \quad (28)$$

$$\begin{aligned} \mathbf{h}_{n_r} : (n+1) &= \mathbf{h}_{n_r} : (n) + \underbrace{\frac{\mu}{\mathbf{x}^T(n) \mathbf{x}(n)}}_{NLMS} \cdot e_{n_r} (n) \mathbf{x}(n) \\ &\quad - \rho_{l0, n_r} \beta \text{sgn} (\mathbf{h}_{n_r} : (n)) e^{-\beta |h_{n_r} : (n)|} \end{aligned} \quad (23)$$

where $\rho_{l0, n_r} = \mu \lambda_{l0, n_r}$. It is worth mentioning that the exponential function in Equation (21) will cause high computational complexity. To reduce the computational complexity, the first-order Taylor series expansion of exponential functions is taken into consideration as [32]

$$e^{-\beta |h_{n_r} : l|} \approx \begin{cases} 1 - \beta |h_{n_r} : l|, & \text{when } |h_{n_r} : l| \leq 1/\beta \\ 0, & \text{others} \end{cases} \quad (24)$$

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

$$\begin{aligned} \mathbf{h}_{n_r} : (n+1) &= \mathbf{h}_{n_r} : (n) + \underbrace{\frac{\mu}{\mathbf{x}^T(n) \mathbf{x}(n)}}_{NLMS} \cdot e_{n_r} (n) \mathbf{x}(n) \\ &\quad - \rho_{l0, n_r} g_2 (\mathbf{h}_{n_r} : (n)) \end{aligned} \quad (25)$$

where $g_2 (\mathbf{h}_{n_r} : (n))$ is the sparse penalty strength function that is defined as

$$\begin{aligned} g_2 (\mathbf{h}_{n_r} : (n)) &= \begin{cases} 2\beta^2 h_{n_r} : l (n) - 2\beta \text{sgn}(h_{n_r} : l (n)), & \text{when } |h_{n_r} : l (n)| \leq 1/\beta \\ 0, & \text{others} \end{cases} \end{aligned} \quad (26)$$

Theorem 2. Suppose the number of iterations approaches to infinity ($n \rightarrow \infty$) and $\mu(n)$ satisfies Equation (13). Then, the n_r -th converged MISO channel $\mathbf{h}_{n_r} : (n+1)$ using L0-NLMS filtering algorithm can be derived as follows:

$$E \{ \mathbf{h}_{n_r} : (\infty) \} = \mathbf{h}_{n_r} : - \frac{P_x \rho_{l0, n_r}}{\mu} E [g_2 (\mathbf{h}_{n_r} : (\infty))] \quad (27)$$

where

Note that the proof of the Theorem 2 is similar to the derivation of Theorem 1 in Equation (17). Because very sparse channel (i.e., $K \ll N$) was considered in this paper, one can find that $(N_t K + 2) \mu P_x < (N_t N + 2) \mu P_x \ll 2$ according to Equation (13). The steady-state MSD of ASCE using sparse NLMS filtering algorithms, according to [33], can be derived as

Table I. Sparse NLMS filtering algorithms for estimating MIMO channels.

Input	Training signal vector: $\mathbf{x}(n) = [x_1^T(n), x_2^T(n), \dots, x_{N_t}^T(n)]^T$ Received signal vector: $\mathbf{d} = [d_1, d_2, \dots, d_{N_r}]^T$
Output	Channel estimate $\hat{\mathbf{H}}$
Initialize	Update times $n = 1$, Initial channel estimate $\tilde{\mathbf{h}}(0) = \mathbf{0}$
Algorithm formulation	Determine n_r : $n_r = \text{mod}(n-1, N_r) + 1$ Choose n_r -th channel estimate Received antenna: $\mathbf{h}_{n_r}(n) = \hat{\mathbf{H}}(n_r, :)$ Set desired signal: $d_{n_r} = \mathbf{y}(n_r)$
Calculate error	Update error: $e_{n_r}(n) = d_{n_r} - \mathbf{h}_{n_r}^T(n)\mathbf{x}(n)$
Adaptive filtering algorithms	LP-NLMS in Equation (16) or L0-NLMS in Equation (25)
Stop criterion	$\mathbf{H}(n_r, :) = \mathbf{h}_{n_r}(n+1)$, if $\mathbf{H}(n+1) - \mathbf{H}(n)_2^2 \leq 10^{-5}$ or $n > \text{threshold}$ is satisfied, then algorithm stop; otherwise, $n = n + 1$, run the algorithm from step 2.

$$\begin{aligned}
MSD_{\text{sparse}}(\infty) &= \lim_{n \rightarrow \infty} E \{ (\mathbf{H}(n) - \mathbf{H})^T (\mathbf{H}(n) - \mathbf{H}) \} \\
&= N_r \lim_{n \rightarrow \infty} E \{ (\mathbf{h}_{n_r}(n) - \mathbf{h}_{n_r})^T (\mathbf{h}_{n_r}(n) - \mathbf{h}_{n_r}) \} \\
&= \frac{\mu N_r N_t N}{2 - \mu(N_t N + 2)} \frac{P_0}{P_x} \left(1 - \frac{\gamma_2}{\gamma_1 + \gamma_2 + \sqrt{\gamma_1^2 + \frac{32\beta^2 N_t N}{\pi} G(\mathbf{h}_{n_r})}} \right) \leq \frac{\mu N_r N_t N}{2 - \mu(N_t N + 2)} \frac{P_0}{P_x}
\end{aligned} \tag{29}$$

where γ_1 , γ_2 , and $G(\mathbf{h}_{n_r})$ in Equation (29) are given by

$$\gamma_1 = 4\beta^2 P_x \mu + 2G(\mathbf{h}_{n_r}) N_t N \tag{30}$$

$$\gamma_2 = \frac{16\beta^2 N_t N}{\pi(2 - (N_t N + 2)\mu P_x)} \tag{31}$$

$$G(\mathbf{h}_{n_r}) = \sum_{l \in \Omega} g_l^2(h_l), \quad i = 1, 2 \text{ and } l = 0, 1, \dots, N_t N - 1 \tag{32}$$

where Ω denotes the position sets of zero or approximate zero taps. Choosing reasonable sparse penalty parameter can exploit the sparsity of MIMO channel. Hence, sparse NLMS filtering algorithm can achieve lower MSD than NLMS, that is, $MSD_{\text{sparse}}(\infty) \leq MSD_{\text{NLMS}}(\infty)$.

3.2. Adaptive MIMO channel estimation

Based on ℓ_p -norm and ℓ_0 -norm sparse constraints, two sparse NLMS filtering algorithms were proposed in Equations (16) and (23), respectively. For estimating MIMO channels, ASCE methods using sparse NLMS filtering algorithms are provided in Table I.

4. COMPUTER SIMULATIONS

In this section, we present the estimation performance of the proposed ASCE estimators. The 100 independent Monte Carlo runs are averaged for evaluation. The length of channel vector \mathbf{h}_{n_r} between each pair (n_t, n_r) is set as $N = 16$, and its number of dominant taps is set

as $K = 1$ and 4 , respectively. The values of dominant channel taps follow Gaussian distribution, which is subjected to $E \{ \|\mathbf{h}_{n_r}\|_2^2 = 1 \}$, and their positions are randomly allocated within the length of \mathbf{h}_{n_r} . The threshold of stopping criterion is set as $\text{threshold} = 1000$. Note that setting threshold according to specific system requirements, the received signal-to-noise ratio (SNR) is defined as $10 \log(E_0/\sigma_n^2)$, where $E_0 = 1$ is transmitted symbol power at each antenna. Here, simulation environment is set in two SNR regimes: 10 and 20 dB. All of the step sizes and regularization parameters are listed in Table II. The estimation performance is evaluated by average MSD, which is defined as

$$\text{Average MSD}\{\mathbf{H}(n)\} = E \{ \|\mathbf{H} - \mathbf{H}(n)\|_2^2 \} \tag{33}$$

where H and $H(n)$ are the actual MIMO channel and its n -th adaptive channel estimate, respectively.

First, the impact of SNR on the average MSD performance is evaluated. Figures 4–6 show that LP-NLMS-based ASCE methods can achieve better estimation performance than standard NLMS-based ACE. Because L0-NLMS-based ASCE methods take advantage of the sparsity of MIMO channel, the estimation performance is better than NLMS. In addition, the three figures also indicate that the estimation performance of ACE using NLMS is insensitive to the number of nonzero channel taps. On

Table II. Simulation parameters.

Parameters	μ	λ_{lp}	λ_{l0}
Values	0.5, 1.0 and 1.5	$(1e-4)\sigma_n^2$	$(1e-3)\sigma_n^2$

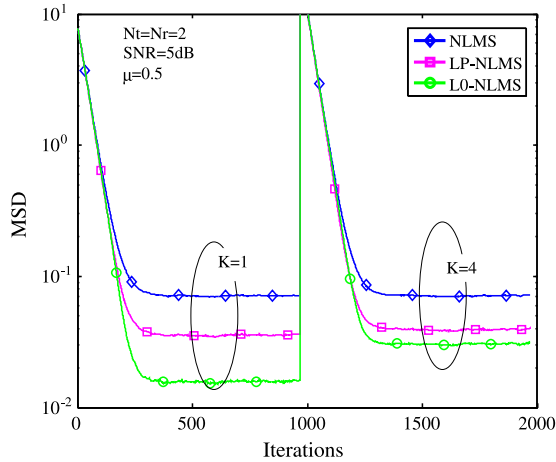


Figure 4. Performance comparison at SNR = 5 dB and $\mu = 0.5$.

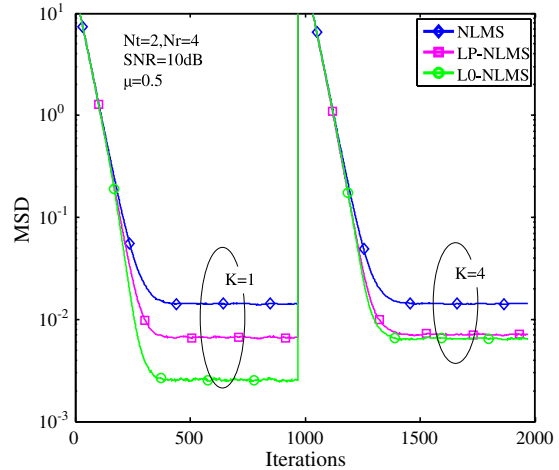


Figure 7. Performance comparison at SNR = 10 dB and $\mu = 0.5$.

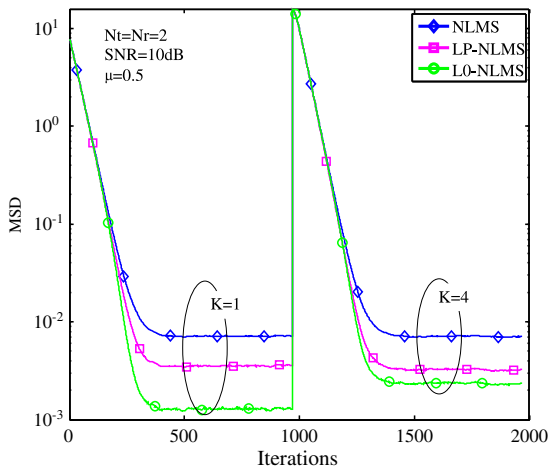


Figure 5. Performance comparison at SNR = 10 dB and $\mu = 0.5$.

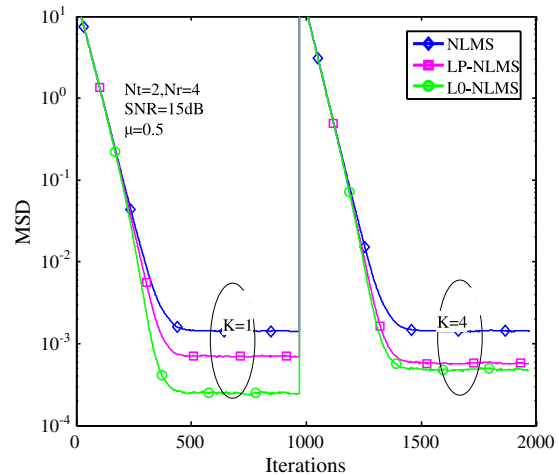


Figure 8. Performance comparison at SNR = 15 dB and $\mu = 0.5$.

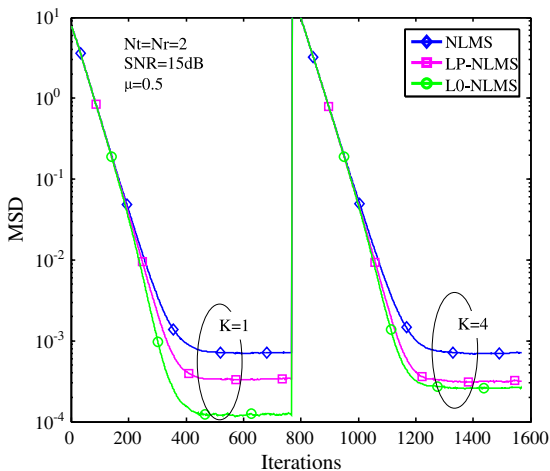


Figure 6. Performance comparison at SNR = 15 dB and $\mu = 0.5$.

the other hand, the proposed ASCE methods depend on the number of nonzero taps. The proposed ASCE methods can achieve better estimation performance for sparser channel.

Second, the impact of the number of transmit/receive antennas on the estimation performance of the proposed methods is evaluated in Figures 7 and 8 where the number of transmit/receive antennas is set to $(N_t, N_r) = (2, 4)$. The figures show that the performance advantage of proposed ASCE is still better than traditional ACE method even if the number of receive antenna was changed.

Finally, the impact of the step sizes, that is, $\mu = 0.5, 1,$ and 1.5 on the estimation performance of the proposed method, is shown in Figures 9 and 10. The SNR is set to 10 dB, and the number of transmit/receive antennas is set to $(N_t, N_r) = (2, 2)$ in Figure 9 and $(N_t, N_r) = (2, 4)$ in Figure 10. The two figures clearly show that the proposed ASCE methods can achieve better performance

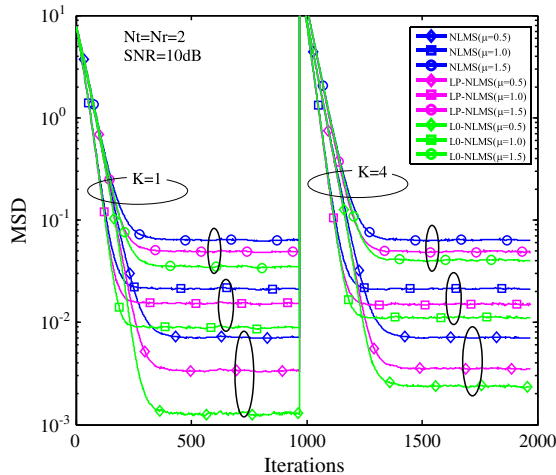


Figure 9. Performance comparison versus step size of gradient descends.

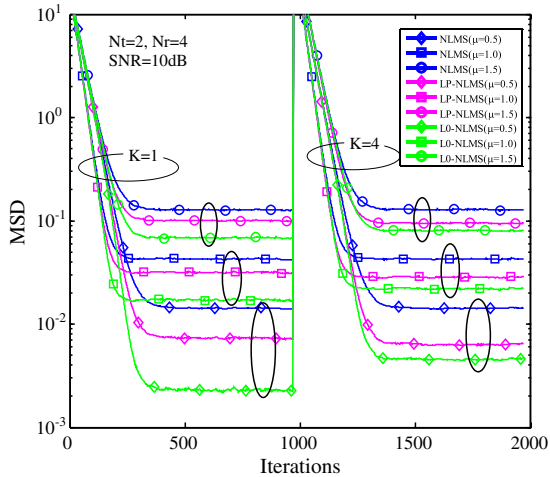


Figure 10. Performance comparison versus step size of gradient descend.

than standard ACE method. Here, note that each proposed ASCE method using smaller gradient descend step size can achieve better estimation accuracy than the one using bigger step size, at the cost of slightly higher computational complexity. For a practical time-variant MIMO-OFDM system, both performance and simplicity of the ASCE method are required. In different SNR region, hence, different step size could be applied to trade off the performance and computational complexity of the proposed ASCE methods.

5. CONCLUSION

In this paper, we proposed ASCE methods using NLMS-based adaptive sparse filtering algorithms for estimating MIMO channels. First, system model was formulated

to ensure that each MISO channel vector can be estimated independently. Second, cost function of the two proposed methods was constructed using sparse constraint functions, that is, ℓ_p -norm and ℓ_0 -norm. Third, MIMO channel matrix was estimated using proposed adaptive sparse filtering algorithm-based ASCE methods. In addition, steady-state MSD of proposed methods was derived, and their MSD performance was proved lower than traditional method. At last, computer simulations were shown that proposed ASCE methods (with sparse constraint) achieved better performance than standard ACE method (without sparse constraint) without scarifying extra computational complexity.

ACKNOWLEDGEMENTS

The authors would like to appreciate Dr. Koichi Adachi and the anonymous reviewers for their valuable comments and suggestion to improve the quality of this paper.

REFERENCES

1. Raychaudhuri D, Mandayam NB. Frontiers of wireless and mobile communications. *Proceedings of the IEEE* 2012; **100**(4): 824–840.
2. Adachi F, Grag D, Takaoka S, Takeda K. Broadband CDMA techniques. *IEEE Wireless Communications* 2005; **12**(2): 8–18.
3. Adachi F, Kudoh E. New direction of broadband wireless technology. *Wireless Communications & Mobile Computing* 2007; **7**(8): 969–983.
4. Gibson JD, Iltis RA. Channel estimation and data detection for MIMO-OFDM systems. In *IEEE Global Telecommunications Conference (GLOBECOM'03)*, San Francisco, USA, December 2003; 581–585.
5. Biguesh M, Gershman aB. Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals. *IEEE Transactions on Signal Processing* 2006; **54**(3): 884–893.
6. Chang T, Chiang W, Hong YP. Training sequence design for discriminatory channel estimation in wireless MIMO systems. *IEEE Transactions on Signal Processing* 2010; **58**(12): 6223–6237.
7. He S, Tugnait JK, Meng X. On superimposed training for MIMO channel estimation and symbol detection. *IEEE Transactions on Signal Processing* 2007; **55**(6): 3007–3021.
8. Pham TH, Liang Y, Nallanathan A. A joint channel estimation and data detection receiver for multiuser MIMO IFDMA systems. *IEEE Transactions on Communications* 2009; **57**(6): 1857–1865.
9. Wang ZJ, Han Z, Liu KJR. A MIMO-OFDM channel estimation approach using time of arrivals. *IEEE*

- Transactions on Wireless Communications* 2005; **4**(3): 1207–1213.
10. Taubock G, Hlawatsch F, Eiwen D, Rauhut H. Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing. *Journal of Selected Topics in Signal Processing* 2010; **4**(2): 255–271.
 11. Bajwa WU, Haupt J, Sayeed AM, Nowak R. Compressed channel sensing: a new approach to estimating sparse multipath channels. *Proceedings of the IEEE* 2010; **98**(6): 1058–1076.
 12. Wang N, Gui G, Zhang Z, Tang T. A novel sparse channel estimation method for multipath MIMO-OFDM systems. In *IEEE 74th Vehicular Technology Conference (VTC'11-Fall)*, San Francisco, California, USA, 2011; 1–5.
 13. Dai L, Wang Z, Yang Z. Spectrally efficient time-frequency training OFDM for mobile large-scale MIMO systems. *IEEE Journal on Selected Areas in Communications* 2013; **31**(2): 251–263.
 14. Candes EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory* 2006; **52**(2): 489–509.
 15. Donoho DL. Compressed sensing. *IEEE Transactions on Information Theory* 2006; **52**(4): 1289–1306.
 16. Zhou Y, Sayeed AM. Experimental study of MIMO channel statistics and capacity via virtual channel representation, *UW Tech. Report*, University of Wisconsin-Madison, Madison, USA. (Available from: <http://dune.ece.wisc.edu/pdfs/zhoumeas.pdf>).
 17. Herdin M, Bonek E, Member S, Fleury BH, Czik N, Yin X, Ozcelik H. Cluster characteristics in a MIMO indoor propagation environment. *IEEE Transactions on Wireless Communications* April 2007; **6**(4): 1465–1475.
 18. Vuokko L, Kolmonen VM, Salo J, Vainikainen P. Measurement of large-scale cluster power characteristics for geometric channel models. *IEEE Transactions on Antennas and Propagation* 2007; **55**(11): 3361–3365.
 19. Dai L, Wang Z, Yang Z. Next-generation digital television terrestrial broadcasting systems: key technologies and research trends. *IEEE Communications Magazine* Sep. 2012; **50**(6): 150–158.
 20. Gui G, Mehdodniya A, Adachi F. Bayesian sparse channel estimation and data detection for OFDM communication Systems. In *IEEE 78th Vehicular Technology Conference (VTC'13-Fall)*, Las Vegas, September 2013; 1–5.
 21. Gui G, Wan Q, Huang AM, Jiang CG. Partial sparse multipath channel estimation using ℓ_1 -regularized LS algorithm. In *IEEE TENCON, Hyderabad, India*, Nov. 2008; 2–5.
 22. Candes EJ. The restricted isometry property and its implications for compressed sensing. *Comptes rendus Mathematique* 2008; **1**(346): 589–592.
 23. Amini A, Unser M, Marvasti F, Member S. Compressibility of deterministic and random infinite sequences. *IEEE Transactions on Signal Processing* 2011; **59**(99): 1–1.
 24. Candés EJ, Tao T. Rejoinder: the Dantzig selector: statistical estimation when p is much larger than n . *Ann. Stat.* 2007; **35**: 2392–2404.
 25. Chen Y, Gu Y, Hero III AO. Sparse LMS for system identification. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, April 19–24, 2009; 3125–3128.
 26. Taheri O, Vorobyov SA. Sparse channel estimation with L_p -norm and reweighted L_1 -norm penalized least mean square. In *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, 2011, Vol. 22–27; 2864–2867.
 27. Gui G, Peng W, Adachi F. Improved adaptive sparse channel estimation based on the least mean square algorithm. In *IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, China, April 7–10 2013; 3130–3134.
 28. Widrow B, Stearns D. *Adaptive Signal Processing*, Vol. 4. Prentice Hall: New Jersey, 1985.
 29. Gui G, Adachi F. Improved adaptive sparse channel estimation using least mean square algorithm. *EURASIP Journal on Wireless Communications and Networking* 2013; **2013**(1): 1–18.
 30. Ahmed MA, Jimaa SA, Abualhaol IY. Enhanced channel estimation technique in MIMO-OFDM system. In *IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Barcelona, Spain, October 2012; 545–549.
 31. Haykin S. *Adaptive Filter Theory* (3rd edn). Prentice-Hall: Englewood Cliffs, NJ, 1996.
 32. Gu Y, Jin J, Mei S. L_0 -norm constraint LMS algorithm for sparse system identification. *IEEE Signal Processing Letters* 2009; **16**(9): 774–777.
 33. Su G, Jin J, Gu Y, Wang J. Performance analysis of ℓ_0 norm constraint least mean square algorithm. *IEEE Transactions on Signal Processing* 2012; **60**(5): 2223–2235.

AUTHORS' BIOGRAPHIES



Guan Gui (M'11) received the DrEng degree in Information and Communication Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2011. From October 2009 to March 2012, with the financial support from the China Scholarship Council (CSC) and the Global Center of Education (ECOE) of Tohoku University, he joined

the wireless signal processing and network laboratory (Prof. Adachi's laboratory), Department of Communications Engineering, Graduate School of Engineering, Tohoku University, as research assistant and postdoctoral research fellow, respectively. From September 2012 to March 2014, he was supported by the Japan Society for the Promotion of Science (JSPS) fellowship as postdoctoral research fellow at the same laboratory. Since April 2014, he has been with Akita Prefectural University, Akita, Japan, where he is a special assistant professor. He is currently engaged in research of multidimensional system control, adaptive filter, compressive sensing, sparse dictionary designing, channel estimation, and advanced wireless techniques. He is a member of Institute of Electrical and Electronics Engineers (IEEE). Dr. Gui has been an associate editor for the Wiley Journal of Security and Communication Networks since 2012. He received IEEE International Conference on Communications (ICC) Best Paper Award as well as IEEE Vehicular Technology Conference (VTC-Spring) Best Student Paper Award in year 2014.



Li Xu (M'94–SM'08) received the BEng degree from the Huazhong University of Science and Technology, Wuhan, China, in 1982, and the MEng and DrEng degrees from Toyohashi University of Technology, Toyohashi, Japan, in 1990 and 1993, respectively. From April 1993 to March 1998, he was an assistant professor at the

Department of Knowledge-Based Information Engineering, Toyohashi University of Technology. From April 1998 to March 2000, he was a lecturer at the Department of Information Management, Asahi University, Gifu, Japan.

Since April 2000, he has been with the Faculty of Systems Science and Technology, Akita Prefectural University, Akita, Japan, where he is currently a professor at the Department of Electronics and Information Systems. His research interests include multidimensional system theory, signal processing, and the applications of computer algebra to system theory. Prof. Xu has been an associate editor for the international journal of Multidimensional Systems and Signal Processing (MSSP) since 2000.



Fumiyuki Adachi received the BS and DrEng degrees in Electrical Engineering from Tohoku University, Sendai, Japan, in 1973 and 1984, respectively. In April 1973, he joined the Electrical Communications Laboratories of Nippon Telegraph and Telephone Corporation (now NTT) and conducted various types of research

related to digital cellular mobile communications. From July 1992 to December 1999, he was with NTT Mobile Communications Network, Inc. (now NTT DoCoMo, Inc.), where he led a research group on wideband/broadband CDMA wireless access for IMT-2000 and beyond. Since January 2000, he has been with Tohoku University, Sendai, Japan, where he is a professor of Communication Engineering at the Graduate School of Engineering. He is currently engaged in research of gigabit wireless communication technology with a data rate above one gigabit per second, with the aim to realize the next-generation frequency and energy-efficient broadband mobile communication systems. He has been serving as the Institute of Electrical and Electronics Engineers (IEEE) VTS Distinguished Lecturer since 2011. From October 1984 to September 1985, he was a United Kingdom SERC Visiting Research Fellow in the Department of Electrical Engineering and Electronics at Liverpool University. He is an IEICE Fellow and was a co-recipient of the IEICE Transactions best paper of the year award in 1996, 1998, and 2009 and also a recipient of Achievement award 2003. He is an IEEE Fellow and was a co-recipient of the IEEE Vehicular Technology Transactions best paper of the year award in 1980 and again in 1990 and also a recipient of Avant Garde award in 2000. He was a recipient of Thomson Scientific Research Front Award in 2004, Ericsson Telecommunications Award in 2008, and Telecom System Technology Award in 2010.