Research Article

Variable Is Better Than Invariable: Sparse VSS-NLMS Algorithms with Application to Adaptive MIMO Channel Estimation

Guan Gui,[1] Zhang-xin Chen,[2] Li Xu,[1] Qun Wan,[2] Jiyan Huang,[2] and Fumiyuki Adachi[3]

[1] *Department of Electronics and Information Systems, Akita Prefectural University, Akita 015-0055, Japan*
[2] *Department of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*
[3] *Department of Communications Engineering, Tohoku University, Sendai 980-8579, Japan*

Correspondence should be addressed to Guan Gui; guiguan@akita-pu.ac.jp

Channel estimation problem is one of the key technical issues in sparse frequency-selective fading multiple-input multiple-output (MIMO) communication systems using orthogonal frequency division multiplexing (OFDM) scheme. To estimate sparse MIMO channels, sparse *invariable step-size normalized least mean square* (ISS-NLMS) algorithms were applied to adaptive sparse channel estimation (ACSE). It is well known that step-size is a critical parameter which controls three aspects: algorithm stability, estimation performance, and computational cost. However, traditional methods are vulnerable to cause estimation performance loss because ISS cannot balance the three aspects simultaneously. In this paper, we propose two stable *sparse variable step-size* NLMS (VSS-NLMS) algorithms to improve the accuracy of MIMO channel estimators. First, ASCE is formulated in MIMO-OFDM systems. Second, different sparse penalties are introduced to VSS-NLMS algorithm for ASCE. In addition, difference between sparse ISS-NLMS algorithms and sparse VSS-NLMS ones is explained and their lower bounds are also derived. At last, to verify the effectiveness of the proposed algorithms for ASCE, several selected simulation results are shown to prove that the proposed sparse VSS-NLMS algorithms can achieve better estimation performance than the conventional methods via mean square error (MSE) and bit error rate (BER) metrics.

1. Introduction

High-rate data broadband transmission over multiple-input multiple-output (MIMO) channel has become one of the mainstream techniques for the next generation communication systems [1]. The major motivation is due to the fact that MIMO technology, as shown in Figure 1, is a way of using multiple antennas to simultaneously transmit multiple streams of data in wireless communications [2] and hence it can bring considerable improvements such as data rate, reliability, and energy efficiency. In fact, coherent receivers require accurate channel state information (CSI) since the received signals are distorted by multipath fading transmission. The accurate estimation of channel impulse response (CIR) is a crucial aspect and challenging issue in coherent

modulation and its accuracy has a significant impact on the overall performance of the communication system.

During last decades, there exist many channel estimation methods which were proposed for MIMO systems [3–11]. All these methods are categorized into two groups. The first group contains the linear channel estimation methods, for example, least squares (LS) algorithm, based on the assumption of dense CIRs. By applying these approaches, the performance of linear methods depends only on the size of MIMO channel. Note that narrowband MIMO channel may be modeled as dense channel model because of its very short time delay spread. Accurately, broadband MIMO channel is often modeled as sparse channel model [12–14]. A typical example of sparse channel is shown in Figure 2. It is well known that linear channel estimation methods are
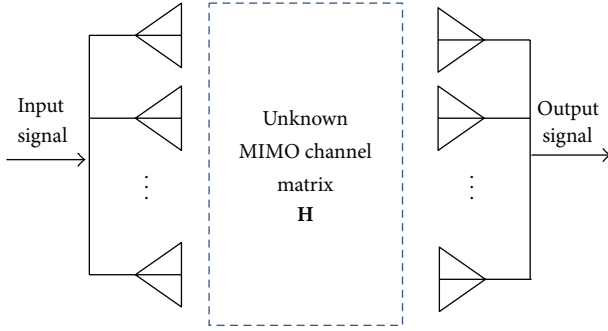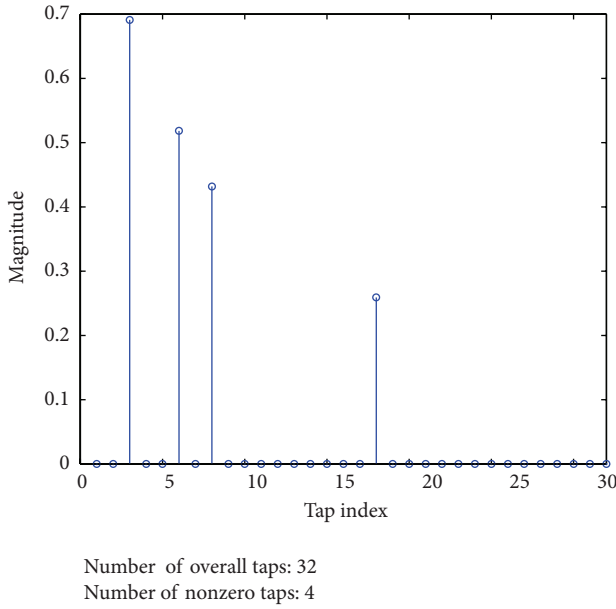
Figure 1: Signal transmission over MIMO channel.

Number of overall taps: 32
Number of nonzero taps: 4

Figure 2: A typical example of sparse multipath channel.

Figure 3: Adaptive algorithm for estimating MIMO channels.

relatively simple to implement due to their low computational complexity [3–8]. Unfortunately, their main drawback is the failure to exploit the inherent channel sparsity. The second group is the sparse channel estimation methods which use compressive sensing (CS) theory [15, 16]. Interested authors are recommended to refer to [17]. Basically, optimal sparse channel estimation often requires that its training signal satisfies restrictive isometry property (RIP) [18] in high probability. However, designing the RIP-satisfied training signal is a nonpolynomial (NP) hard problem [19]. Also, there exist some proposed methods which are stable with the cost of extra computational burden, especially in time-variant MIMO systems. For example, sparse channel estimation method using Dantzig selector was proposed for double-selective fading MIMO systems [10]. Indeed, the proposed method needs to be solved by linear programming which incurs high computational complexity. To reduce the computational cost, sparse channel estimation methods using greedy iterative algorithms were also proposed in [9, 11]. But their complexity still depends on the number of nonzero taps of MIMO channel.
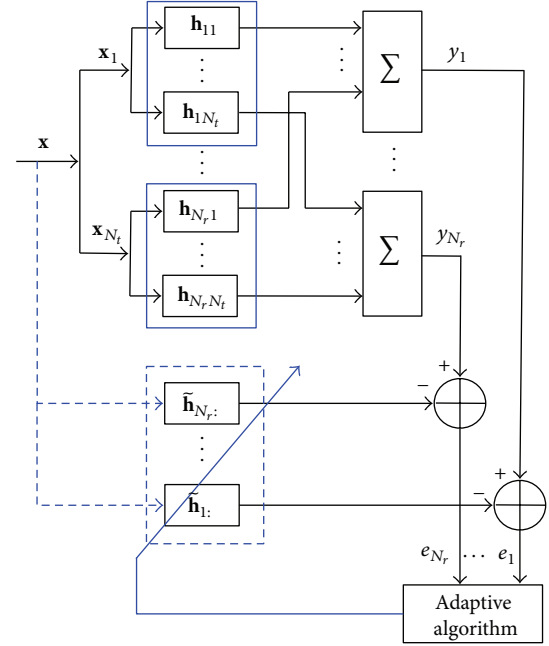
Unfortunately, the mentioned proposed methods do not have adaptive estimation capability. Adaptive sparse channel estimation (ASCE) methods using sparse *invariable step-size* (ISS) least mean square algorithms (ISS-LMS) were proposed in [20] for single-input single-output (SISO) channels. However, conventional ISS-LMS methods have two main drawbacks: (1) sensitive to random scale of training signal and (2) unstable in low signal-to-noise ratio (SNR) regime.

To overcome the two harmful factors on channel estimation and extend their applications to estimate MIMO channels, sparse ISS normalized least mean square (ISS-NLMS) algorithms, for example, zero-attracting ISS-NLMS (ZA-ISS-NLMS) and reweight ZA-ISS-NLMS (RZA-ISS-NLMS), were proposed in [21]. It is well known that *step-size* is a critical parameter which controls the estimation performance, convergence rate, and computational cost. Different from conventional sparse ISS-NLMS algorithms [21], *zero-attracting variable step-size NLMS* (ZA-VSS-NLMS) algorithm was proposed for ASCE to improve estimation performance in sparse multipath single-input single-output (SISO) systems [22]. Unlike the previous works, this paper proposes two sparse VSS-NLMS algorithms for estimating sparse MIMO channels. The main contribution of this paper is summarized as follows. First, we derive the lower bound of proposed MIMO channel estimator for introducing the research motivation. Second, we extend the proposed VSS-ZA-NLMS for estimating SISO channels in [22] to MIMO channels. Third, a reweighted ZA-VSS-NLMS (RZA-VSS-NLMS) is proposed to further improve the estimation performance of MIMO channels. In addition, we explain the reason why sparse VSS-NLMS algorithms can achieve better performance than conventional sparse ISS-NLMS ones. Finally, Monte Carlo based computer simulations are conducted to confirm the
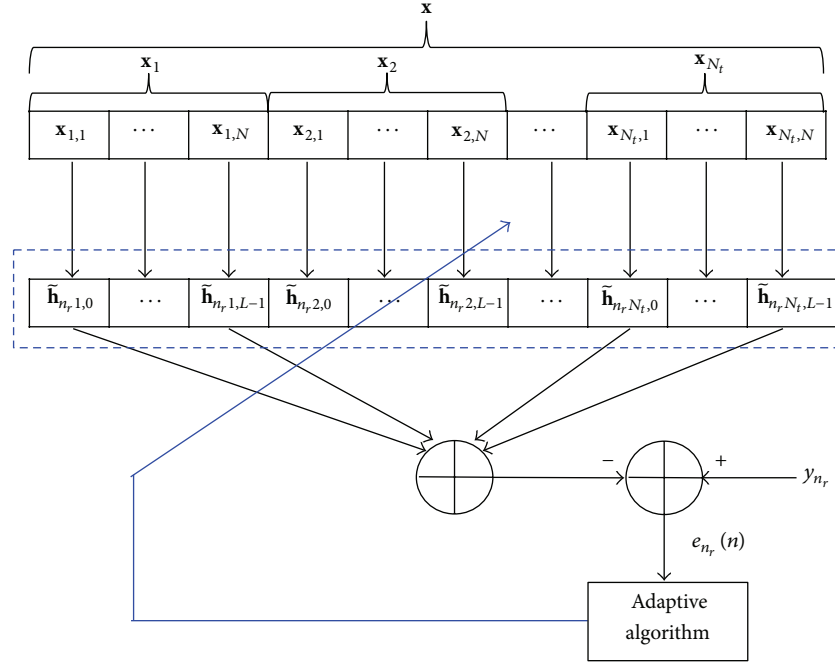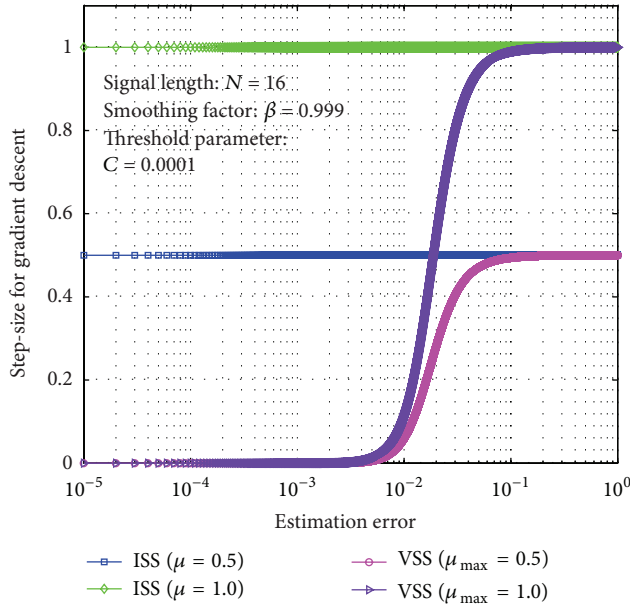
FIGURE 4: MISO channel estimation at n_rth antenna of the receiver.

FIGURE 5: ISS and VSS versus updating estimation error.

effectiveness of our proposed algorithms via two metrics: bit error rate (BER) and mean square error (MSE).

The remainder of this paper is organized as follows. A baseband MIMO system model is described and problem formulation is presented in Section 2. In Section 3, sparse ISS-NLMS algorithms are overviewed. In Section 4, sparse VSS-NLMS algorithms are proposed and a figure example is also given to explain the difference between ISS and VSS based algorithms. Simulation results are presented in

Section 5 in order to assess the proposed methods. Finally, we conclude the paper in Section 6.

Notations. Capital bold letters and small bold letters denote matrices and row/column vectors, respectively. The discrete FOURIER transform (DFT) matrix is denoted by $\mathbf{F}$ with entries $[\mathbf{F}]_{kn} = \mathbf{1}/K e^{-j2\pi kq/K}$, $k, q = 0, 1, \ldots, K - 1$; $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$ and $|\cdot|$ denote the transpose, conjugate transpose, matrix inversion, and absolute operations, respectively; $E\{\cdot\}$ denotes the expectation operator; assume any vector $\mathbf{h} = [h_0, \ldots, h_l, \ldots, h_{L-1}]^T$; $\|\mathbf{h}\|_1$ and $\|\mathbf{h}\|_2$ denote ℓ_1-norm, that is, $\|\mathbf{h}\|_1 = \sum_l |h_l|$, and ℓ_2-norm, that is, $\|\mathbf{h}\|_2 = (\sum_l |h_l|^2)^{1/2}$; $\mathrm{sgn}(\mathbf{h})$ is a component-wise function which is defined as $\mathrm{sgn}(h) = 1$ for $h > 0$, $\mathrm{sgn}(h) = 0$ for $h = 0$, and $\mathrm{sgn}(h) = 1$ for $h < 0$, where h denotes any component in vector $\mathbf{h}$; $\widetilde{\mathbf{h}}$ represents the channel estimator of $\mathbf{h}$.

2. System Model

A frequency-selective fading MIMO communication system using OFDM modulation scheme is considered in Figure 3. Initially, frequency domain signal vector $\overline{\mathbf{x}}_{n_t}(t) = [\overline{x}_{n_t}(t, 0), \ldots, \overline{x}_{n_t}(t, K - 1)]^T$, $n_t = 1, 2, \ldots, N_t$, is fed to inverse discrete Fourier transform (IDFT) at the n_tth antenna, where K is the number of subcarriers and N_t is the number of transmit antennas. Assume that the transmit power is normalized as $E\{\|\overline{\mathbf{x}}_{n_t}(t)\|_2^2\} = 1$. The resultant vector $\mathbf{x}_{n_t}(t) \triangleq \mathbf{F}^H \overline{\mathbf{x}}_{n_t}(t)$ is padded with cyclic prefix (CP) of length $L_{\mathrm{CP}} \geq (K - 1)$ to avoid interblock interference (IBI). After CP removal, the received signal vector at the n_rth antenna for time t is written as $y_{n_r}(t)$, where $n_r = 1, 2, \ldots, N_r$. As shown in

Figure 4, the received signal vector $\mathbf{y}$ and input signal vector $\mathbf{x}(t)$ are related by

$$y_{n_r}(t) = \sum_{n_t=1}^{N_t} \mathbf{h}_{n_r n_t}^T \mathbf{x}_{n_t}(t) + z_{n_r}(t)$$

$$= \mathbf{h}_{n_r:}^T \mathbf{x}(t) + z_{n_r}(t), \tag{1}$$

where $\mathbf{x}(t) = [\mathbf{x}_1^T(t), \mathbf{x}_2^T(t), \ldots, \mathbf{x}_{N_t}^T(t)]^T$ collects all of the input signal vectors from different antennas at the transmitter, $z_{n_r}(t)$ is an additive white Gaussian noise (AWGN) variable with distribution $\mathscr{CN}(0, \sigma_n^2)$, and n_rth received multiple-input single-output (MISO) channel vector $\mathbf{h}_{n_r:}$ is written as

$$\mathbf{h}_{n_r:} := \left[\underbrace{h_{n_r1,0}, \ldots, h_{n_r1,L-1}}_{\mathbf{h}_{n_r1}^T}, \ldots, \underbrace{h_{n_rn_t,0}, \ldots, h_{n_rn_t,L-1}}_{\mathbf{h}_{n_rn_t}^T}, \ldots, \right.$$

$$\left. \underbrace{h_{n_rN_t,0}, \ldots, h_{n_rN_t,L-1}}_{\mathbf{h}_{n_rN_t}^T} \right]^T \tag{2}$$

and the matrix-vector form of system model (1) is also written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}, \tag{3}$$

where received signal vector $\mathbf{y}$, noise vector $\mathbf{z}$, and channel matrix $\mathbf{H}$ can be represented, respectively, as follows:

$$\mathbf{y} = \left[y_1(t), y_2(t), \ldots, y_{N_r}(t) \right]^T \in \mathbb{C}^{N_r \times 1},$$

$$\mathbf{z} = \left[z_1(t), z_2(t), \ldots, z_{N_r}(t) \right]^T \in \mathbb{C}^{N_r \times 1},$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_{1:}^T \\ \mathbf{h}_{2:}^T \\ \vdots \\ \mathbf{h}_{N_r:}^T \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{11}^T & \mathbf{h}_{12}^T & \cdots & \mathbf{h}_{1N_t}^T \\ \mathbf{h}_{21}^T & \mathbf{h}_{22}^T & \cdots & \mathbf{h}_{2N_t}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{N_r1}^T & \mathbf{h}_{N_r2}^T & \cdots & \mathbf{h}_{N_rN_t}^T \end{bmatrix} \in \mathbb{C}^{N_r \times N_t L}, \tag{4}$$

where $\mathbf{h}_{n_r n_t}$, $n_r = 1, 2, \ldots, N_t$, is assumed to be equal L-length sparse channel vector from receiver to n_tth antenna. In addition, we also assume that each channel vector $\mathbf{h}_{n_r n_t}$ is only supported by T dominant channel taps.

3. Overview of Sparse ISS-NLMS Algorithms

According to the system model in (1), the n_tth updating estimation error $e_{n_r}(n)$ can be written as

$$e_{n_r}(n) = y_{n_r}(t) - y_{n_r}(n)$$

$$= y_{n_r}(t) - \widetilde{\mathbf{h}}_{n_r:}^T(n) \mathbf{x}(t), \tag{5}$$

where $\widetilde{\mathbf{h}}_{n_r:}(n)$ denotes an MISO channel estimator of the $\mathbf{h}_{n_r:}$; $\mathbf{e}(n) = [e_1(n), e_1(n), \ldots, e_{N_r}(n)]^T$ denotes receive error vector at the nth adaptive update; and $y_{n_r}(t)$ is the received signal at the n_rth receive antenna.

3.1. ISS-ZA-NLMS.
According to (5), the cost function of ISS-ZA-LMS [23] at the n_rth antenna of the receiver can be constructed as

$$G_{\mathrm{ZA},n_r}(n) = \frac{1}{2}e_{n_r}^2(n) + \lambda_{\mathrm{ZA}} \left\| \widetilde{\mathbf{h}}_{n_r:}(n) \right\|_1, \tag{6}$$

where λ_{ZA} is a regularization parameter to balance the square estimation error $e_{n_r}^2(n)$ and sparse penalty of $\widetilde{\mathbf{h}}_{n_r:}(n)$. Hence, the corresponding update equation of ISS-ZA-LMS [23] for MIMO channel estimation is derived as

$$\widetilde{\mathbf{h}}_{n_r:}(n+1) = \widetilde{\mathbf{h}}_{n_r:}(n) - \mu \frac{\partial G_{\mathrm{ZA},n_r}(n)}{\partial \widetilde{\mathbf{h}}_{n_r:}(n)}$$

$$= \widetilde{\mathbf{h}}_{n_r:}(n) + \mu e_{n_r} \mathbf{x}(t) - \gamma_{\mathrm{ZA}} \mathrm{sgn}\left(\widetilde{\mathbf{h}}_{n_r:}(n)\right), \tag{7}$$

for $n_r = 1, 2, \ldots, N_r$, where $\gamma_{\mathrm{ZA}} = \mu \lambda_{\mathrm{ZA}}$ and μ is the ISS. To mitigate random scaling of input signal $\mathbf{x}(t)$, based on the ISS-ZA-LMS algorithm in (7), the update equation of improved ISS-ZA-NLMS [20, 23] was proposed as

$$\widetilde{\mathbf{h}}_{n_r:}(n+1) = \widetilde{\mathbf{h}}_{n_r:}(n) + \frac{\mu e_{n_r} \mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)} - \gamma_{\mathrm{ZA}} \mathrm{sgn}\left(\widetilde{\mathbf{h}}_{n_r:}(n)\right). \tag{8}$$

3.2. ISS-RZA-NLMS.
It is well known that ISS-ZA-LMS cannot distinguish between zero taps and nonzero taps as it gives the same penalty to all the taps which are often forced to be zero with the same probability; therefore, its performance will degrade in less sparse systems. Motivated by the reweighted ℓ_1-norm minimization recovery algorithm [24], Chen et al. have proposed a heuristic approach to reinforce the zero attractor which was termed as the ISS-RZA-LMS [25]. The cost function of ISS-RZA-LMS is given by

$$G_{\mathrm{RZA},n_r}(n) = \frac{1}{2}e_{n_r}^2(n)$$

$$+ \lambda_{\mathrm{RZA}} \sum_{n_t=1}^{N_t} \sum_{l=0}^{L-1} \log\left(1 + \varepsilon_{\mathrm{RZA}}\left|\widetilde{h}_{n_rn_t,l}(n)\right|\right), \tag{9}$$

where $\lambda_{\mathrm{RZA}} > 0$ is the regularization parameter and reweighted factor $\varepsilon_{\mathrm{RZA}} > 0$ is the positive threshold. In computer simulation, the threshold is set as $\varepsilon_{\mathrm{RZA}} = 20$ which is also suggested in previous papers [26, 27]. The lth coefficient $\widetilde{h}_{n_r,l}(n)$ of ISS-RZA-LMS channel estimator $\widetilde{\mathbf{h}}_{n_r}(n)$ is then updated as

$$\widetilde{h}_{n_rn_t,l}(n+1) = \widetilde{h}_{n_rn_t,l}(n) - \mu \frac{\partial G_{\mathrm{RZA},n_r}}{\partial \widetilde{h}_{n_rn_t,i}(n)}$$

$$= \widetilde{h}_{n_rn_t,l}(n) + \mu e_{n_r} x(t-l)$$

$$- \frac{\gamma_{\mathrm{RZA}} \mathrm{sgn}\left(\widetilde{h}_{n_rn_t,l}(n)\right)}{1 + \varepsilon_{\mathrm{RZA}}\left|\widetilde{h}_{n_rn_t,l}(n)\right|}, \tag{10}$$

for $n_r = 1, 2, \ldots, N_r$, where $\gamma_{\text{RZA}} = \mu \lambda_{\text{RZA}} \varepsilon_{\text{RZA}}$. According to (10), hence, ISS-RZA-NLMS [20, 23] was proposed as

$$
\widetilde{\mathbf{h}}_{n_r:} (n+1) = \widetilde{\mathbf{h}}_{n_r:} (n) + \frac{\mu e_{n_r} \mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}
$$
$$
- \gamma_{\text{RZA}} \frac{\text{sgn}\left(\widetilde{\mathbf{h}}_{n_r:}(n)\right)}{1 + \varepsilon_{\text{RZA}}\left|\widetilde{\mathbf{h}}_{n_r:}(n)\right|} . \tag{11}
$$

Note that the sparse penalty term $\text{sgn}(\widetilde{\mathbf{h}}_{n_r:}(n))/(1 + \varepsilon_{\text{RZA}}|\widetilde{\mathbf{h}}_{n_r:}(n)|)$ in (11) replaces these channel coefficients $\{\widetilde{h}_{n_r n_t, l}(n), l = 0, 1, \ldots, L-1, n_t = 1, 2, \ldots, N_t\}$ under the threshold $1/\varepsilon_{\text{RZA}}$ as zero.

3.3. Drawback of the Sparse ISS-LMS Algorithms. Comparing the standard ISS-NLMS algorithm [28], sparse ISS-NLMS algorithms have a common ability of exploiting channel sparsity. Without the loss of generality, we derive the steady-state mean square error (MSE) performance of the ISS-ZA-NLMS [23] as for the typical example to illustrate the drawbacks of the sparse ISS-NLMS algorithms. Assuming that $\widetilde{\mathbf{H}}_{\text{ISS}}(n)$ denotes the sparse MIMO channel estimator, under the independence assumption, in [25], the steady-state MSE of ISS-ZA-NLMS estimator $\widetilde{\mathbf{H}}_{\text{ISS}}(n)$ was derived as

$$
\Delta_{\text{ISS}}(\infty) = \lim_{n \to \infty} E\left\{\left\|\left(\widetilde{\mathbf{H}}_{\text{ISS}}(n) - \mathbf{H}\right)\mathbf{x}(t)\right\|_2^2\right\}
$$
$$
= \lim_{n \to \infty} E\left\{\sum_{n_r=1}^{N_r} \left\|\left(\widetilde{\mathbf{h}}_{n_r:}(n) - \mathbf{h}\right)\mathbf{x}(t)\right\|_2^2\right\}
$$
$$
= \frac{\text{Tr}\left[\mathbf{R}(\mathbf{I} - \mu\mathbf{R})^{-1}\right]\sigma_n^2 N_r}{2 - \text{Tr}\left[\mathbf{R}(\mathbf{I} - \mu\mathbf{R})^{-1}\right]}
$$
$$
+ \frac{\gamma_1 \rho_{\text{ZA}}\left(\rho_{\text{ZA}} - 2\gamma_2/\gamma_1\right)}{\left(2 - \text{Tr}\left[\mathbf{R}(\mathbf{I} - \mu\mathbf{R})^{-1}\right]\right)\mu}
$$
$$
\leq \frac{\text{Tr}\left[\mathbf{R}(\mathbf{I} - \mu\mathbf{R})^{-1}\right]\sigma_n^2 N_r}{2 - \text{Tr}\left[\mathbf{R}(\mathbf{I} - \mu\mathbf{R})^{-1}\right]}
$$
$$
\leq \frac{\lambda_{\max}\sigma_n^2 N_r}{2 - 3\mu\lambda_{\max}}, \tag{12}
$$

where $\gamma_1 = E[\text{sgn}(\widetilde{\mathbf{h}}_{n_r:}(n))(\mathbf{I} - \mu\mathbf{R})^{-1}\text{sgn}(\widetilde{\mathbf{h}}_{n_r:}(n))] > 0$ and $\gamma_2 = E\|\widetilde{\mathbf{h}}_{n_r:}(\infty)\|_1 - \|\mathbf{h}_{n_r:}\|_1$. To exploit the channel sparsity, ρ_{ZA} should be selected in the range $(0, 2\gamma_2/\gamma_1]$ so that $(\rho_{\text{ZA}} - 2\gamma_2/\gamma_1) \leq 0$. According to (12), the lower bound of $\Delta_{\text{ISS}}(\infty)$ depends on the three factors: $\{\lambda_{\max}, \sigma_n^2, \mu\}$. However, $\lambda_{\max}$ and σ_n^2 are determined by the input signal $\mathbf{x}(t)$ and additive noise $z(t)$, respectively. Only selecting the smaller step-size can further achieve better MSE performance. However, if small step-size μ is adopted, it will incur slow convergence speed (i.e., high computation complexity) on overall adaptive channel estimation. Hence, it is expected that large step-size is used in the case of large MSE to accelerate the convergence

speed, while small step-size is used in the case of smaller MSE to improve the steady-state MSE performance. Assume $\widetilde{\mathbf{H}}(n)$ denotes the nth update MIMO channel estimator using sparse VSS-NLMS algorithms. As $\mu \to 0$, the lower bound of steady-state MSE of sparse VSS-NLMS algorithms is derived as

$$
\Delta_{\text{VSS}}(\infty) = \lim_{n \to \infty} E\left\{\left\|\left(\widetilde{\mathbf{H}}(n) - \mathbf{H}\right)\mathbf{x}(t)\right\|_2^2\right\}
$$
$$
= \lim_{n \to \infty} E\left\{\sum_{n_r=1}^{N_r}\left\|\left(\widetilde{\mathbf{h}}_{n_r:}(n) - \mathbf{h}\right)\mathbf{x}(t)\right\|_2^2\right\}
$$
$$
\leq \lim_{\mu \to 0} \frac{\lambda_{\max}\sigma_n^2 N_r}{2 - 3\mu\lambda_{\max}} \tag{13}
$$
$$
= \frac{\lambda_{\max}\sigma_n^2 N_r}{2}
$$
$$
\leq \Delta_{\text{ISS}}(\infty).
$$

To simultaneously achieve higher convergence speed and lower steady-state MSE performance, we propose sparse VSS-NLMS algorithms for estimating MIMO channels in the next section.

4. Proposed Sparse VSS-NLMS Algorithms for Estimating MIMO Channels

Recall that the ISS-ZA-NLMS algorithm in (8) does not make use of the VSS rather than ISS. Inspired from the VSS-NLMS algorithm which has been proposed in [29], to improve estimation performance of MIMO channels, sparse VSS-NLMS algorithms are proposed. Unlike the sparse ISS-NLMS algorithm, sparse VSS-NLMS algorithms are time-variant with respect to the accuracy of updating estimators.

4.1. VSS-ZA-NLMS. At time t, based on the previous research on the ISS-ZA-NLMS and VSS-NLMS algorithms, VSS-ZA-NLMS algorithm is proposed as follows:

$$
\widetilde{\mathbf{h}}_{n_r}(n+1) = \widetilde{\mathbf{h}}_{n_r}(n) + \mu_{n_r}(n)\frac{e_{\text{ZA}}(n)\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}
$$
$$
- \gamma_{\text{ZA}}\,\text{sgn}\left(\widetilde{\mathbf{h}}_{n_r}(n)\right), \tag{14}
$$

where $\mu_{n_r}(n)$ is the VSS which is given by

$$
\mu_{n_r}(n) = \mu_{\max} \cdot \frac{\mathbf{p}_{n_r}^T(n)\mathbf{p}_{n_r}(n)}{\mathbf{p}_{n_r}^T(n)\mathbf{p}_{n_r}(n) + C}, \tag{15}
$$

where C is a positive threshold parameter which is related to received signal-to-noise ratio (SNR), $C \sim \mathcal{O}(1/\text{SNR})$. According to (15), the range of VSS is given as $\mu_{n_r}(n) \in (0, \mu_{\max})$, where $\mu_{\max}$ is the maximal step-size of gradient descent. Theoretically, the maximal step-size is less than 2 to ensure the adaptive algorithm stability [28]. Please note that $\mathbf{p}_{n_r}(n)$ in (15) is given by

$$
\mathbf{p}_{n_r}(n) = \beta\mathbf{p}_{n_r}(n-1) + (1-\beta)\frac{\mathbf{x}(t)e_{n_r}(n)}{\mathbf{x}^T(t)\mathbf{x}(t)}, \tag{16}
$$

where $\beta \in [0, 1)$ is a smoothing factor to trade off VSS and estimation error.

4.2. VSS-RZA-NLMS. The lth channel coefficient $\widetilde{h}_{n_r,l}(n)$ of $\widetilde{\mathbf{h}}_{n_r}(n)$ is then updated by

$$\widetilde{h}_{n_r,l}(n+1) = \widetilde{h}_{n_r,l}(n) + \mu(n)\frac{e_{n_r}(n)\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}$$
$$- \gamma_{\text{RZA}}\frac{\text{sgn}\left(\widetilde{h}_{n_r,l}(n)\right)}{1 + \varepsilon_{\text{RZA}}\left|\widetilde{h}_{n_r,l}(n)\right|}. \tag{17}$$

Then, the matrix-vector form of (15) can also be expressed as

$$\widetilde{\mathbf{h}}_{n_r}(n+1) = \widetilde{\mathbf{h}}_{n_r}(n) + \mu(n)\frac{e_{n_r}(n)\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}$$
$$- \gamma_{\text{RZA}}\frac{\text{sgn}\left(\widetilde{\mathbf{h}}_{n_r}(n)\right)}{1 + \varepsilon_{\text{RZA}}\left|\widetilde{\mathbf{h}}_{n_r}(n)\right|}. \tag{18}$$

Please note that the second term in (16) attracts the channel coefficients $\widetilde{h}_{n_r,l}(n)$, $l = 0, 1, \ldots, L-1$, whose magnitudes are comparable to $1/\varepsilon_{\text{RZA}}$ to zeros. For estimating MIMO channels, two proposed sparse VSS-NLMS algorithms are summarized in Algorithm 1.

Remark 1. To better understand the difference between ISS and VSS, based on (8), (11), and (15), it is worth mentioning that step-size μ for sparse ISS-NLMS algorithm is invariable but the step-size $\mu_{\text{ZA}}(n)$ for sparse VSS-NLMS algorithm is variable as depicted in Figure 5, where the maximal step-size and ISS are set as $\mu_{\text{max}} \in \{0.5, 1.0\}$ and $\mu \in \{0.5, 1.0\}$, respectively. From the figure, one can easily find that ISS is kept invariant. Unlike the ISS, VSS $\mu(n)$ decreases adaptively as the estimation performance increases and vice versa. In other words, sparse VSS-NLMS algorithms adopting VSS for adaptive gradient descend; large step-size is adopted to speed up convergence rate for reducing computational complexity; small step-size is adopted to ensure algorithm stability in the case of high-accuracy estimator for further improving estimation performance.

5. Computer Simulations

To confirm the effectiveness of the proposed methods, two metrics, that is, MSE and BER, are adopted for performance evaluation. Channel estimators are evaluated by average MSE which is defined by

$$\text{Average MSE}\left\{\widetilde{\mathbf{H}}(n)\right\} = E\left\{\left\|\mathbf{H} - \widetilde{\mathbf{H}}(n)\right\|_2^2\right\}, \tag{19}$$

and system performance is evaluated by the BER metric which adopts different data modulation schemes, such as phase shift keying (PSK) and quadrature amplitude modulation (QAM). The results are averaged over 1000 independent Monte-Carlo runs. The length of each channel vector $\{\mathbf{h}_{n_r n_t}, n_r = 1, 2, \ldots, N_r, n_t = 1, 2, \ldots, N_t\}$ is set as equal length

TABLE 1: Simulation parameters.

| Parameters	Values
Number of transceivers (N_t, N_r)	(4, 4)
Channel length of each $\mathbf{h}_{n_r n_t}$	$L = 16$
Number of nonzero coefficients	$T \in \{1, 4\}$
Distribution of nonzero coefficient	Random Gaussian $\mathcal{CN}(0, 1)$
Threshold parameter for VSS-NLMS	$C \in \{10^{-4}, 10^{-5}\}$
Received SNR for channel estimation	{5 dB, 15 dB, 25 dB}
Received SNR E_0/N_0 for data trans.	12 dB~30 dB
Step-size	$\mu = 0.2$ and $\mu_{\text{max}} = 2$
Regularization parameter for ZA-NLMS	$\rho_{\text{ZA}} = 0.0006\sigma_n^2$ for $T = 1$ $\rho_{\text{ZA}} = 0.0002\sigma_n^2$ for $T = 4$
Regularization parameter for RZA-NLMS	$\rho_{\text{RZA}} = 0.006\sigma_n^2$ for $T = 1$ $\rho_{\text{RZA}} = 0.002\sigma_n^2$ for $T = 4$
Modulation schemes	QPSK, 8PSK, 16PSK, 16QAM, 64QAM, 128QAM

with $L = 16$ and corresponding number of dominant taps is set to $T \in \{1, 4\}$. Each dominant channel tap follows random Gaussian distribution as $\mathcal{CN}(0, \sigma_{\mathbf{h}}^2)$ and their positions are randomly decided within the length of $\mathbf{h}_{n_r n_t}$. In addition, MISO channel vector $\mathbf{h}_{n_r:}$ is subject to $E\{\|\mathbf{h}_{n_r:}\|_2^2\}=1$. The received SNR is defined as P_0/σ_n^2, where P_0 is the power of received signal. Computer simulation parameters are listed in Table 1. Based on the research work in [30], it is worth mentioning that threshold parameters of sparse VSS-NLMS algorithms are adopted $C = 10^{-4}$ for 5 dB and $C = 10^{-5}$ for 10 dB and 20 dB, respectively.

In the first example, average MSE performance of proposed methods is evaluated in the case of $T = 1$ and 4 in Figures 6, 7, 8, 9, 10, and 11 under three SNR regimes, that is, 5 dB, 10 dB, and 20 dB. To confirm the effectiveness of the proposed method, we compare it with previous methods, that is, ISS-NLMS [28], VSS-NLMS [29], and sparse ISS-NLMS [23, 25]. In addition, to achieve a better steady-state estimation performance, regularization parameters for sparse VSS-NLMS algorithms, that is, VSS-ZA-NLMS and VSS-RZA-NLMS, are adopted from [27], which depend on the number of nonzero taps of a channel. In the case of different SNR regimes, for example, 5 dB, 10 dB, and 20 dB, as shown in Figures 6, 7, 8, 9, 10, and 11, two proposed methods achieved better estimation performance than sparse ISS-NLMS ones.

As it can be observed from Figures 5 and 6, since sparse VSS-NLMS algorithms take advantage of the channel sparsity as for prior information, hence they achieve better estimation performance than standard VSS-NLMS algorithm, especially in a very sparse channel case, for example, $T = 1$. The sparse VSS-NLMS algorithms can exploit much more sparse information for sparser channel. It is obviously observed that the performance gaps between proposed methods and

Input	(1) $\mathbf{x}(t)$ and $\mathbf{y}(t)$;
	(2) $\mu_{\max} = 1$ and C;
	(3) γ_{ZA} for VSS-ZA-NLMS;
	(4) γ_{RZA} and $\varepsilon_{\mathrm{RZA}}$ for VSS-RZA-NLMS;
Output	Channel estimator $\widetilde{\mathbf{H}}$.
Initialization	$n \leftarrow 1$;
	$\mathbf{p}_{n_r}(0) \leftarrow \mathbf{0}$;
	$\widetilde{\mathbf{h}}(0) \leftarrow \mathbf{0}$.
While	$\left\|\widetilde{\mathbf{H}}(n+1) - \widetilde{\mathbf{H}}(n)\right\|_2^2 \le 10^{-5}$ or $n \ge 5000$ **Do**
	$n \leftarrow n+1$;
	$n_r \leftarrow \mathrm{mod}(n-1, N_r) + 1$;
	$\widetilde{\mathbf{h}}_{n_r}(n) \leftarrow \widetilde{\mathbf{h}}(n_r; :)$;
	$d_{n_r}(n) \leftarrow y(n_r)$;
	$e_{n_r}(n) \leftarrow d_{n_r}(n) - \widetilde{\mathbf{h}}_{n_r}^T \mathbf{x}(n)$;
	$\mathbf{p}_{n_r}(n) \leftarrow \beta \mathbf{p}_{n_r}(n-1) + (1-\beta)\mathbf{x}(t)e_{n_r}(n)/\mathbf{x}^T(t)\mathbf{x}(t)$;
	$\mu_{n_r}(n) \leftarrow \mu_{\max} \cdot \mathbf{p}_{n_r}^T(n)\mathbf{p}_{n_r}(n)/(\mathbf{p}_{n_r}^T(n)\mathbf{p}_{n_r}(n) + C)$;
	$\widetilde{\mathbf{h}}_{n_r}(n+1) \leftarrow \widetilde{\mathbf{h}}_{n_r}(n) + \mu_{n_r}(n)e_{\mathrm{ZA}}(n)\mathbf{x}(t)/\mathbf{x}^T(t)\mathbf{x}(t) - \gamma_{\mathrm{ZA}} \mathrm{sgn}(\widetilde{\mathbf{h}}_{n_r}(n))$ for VSS-ZA-NLMS in (14) or
	$\widetilde{\mathbf{h}}_{n_r}(n+1) \leftarrow \widetilde{\mathbf{h}}_{n_r}(n) + \mu(n)e_{n_r}(n)\mathbf{x}(t)/\mathbf{x}^T(t)\mathbf{x}(t) - \gamma_{\mathrm{RZA}} \mathrm{sgn}(\widetilde{\mathbf{h}}_{n_r}(n))/(1 + \varepsilon_{\mathrm{RZA}}|\widetilde{\mathbf{h}}_{n_r}(n)|)$
	for VSS-RZA-NLMS in (18);
End	$\widetilde{\mathbf{H}}(n_r; :) \leftarrow \widetilde{\mathbf{h}}_{n_r}(n+1)$

ALGORITHM 1: Sparse VSS-NLMS algorithms for estimating MIMO channels.
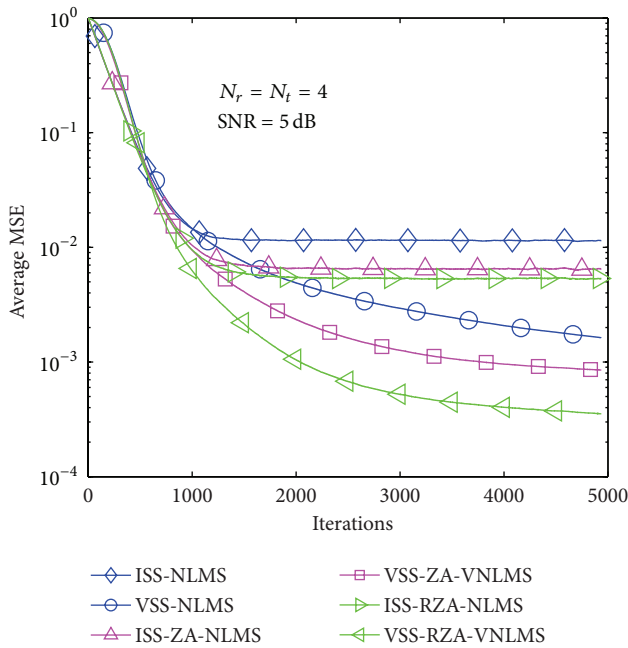
FIGURE 6: Average MSE performance versus received iterations ($T = 1$).
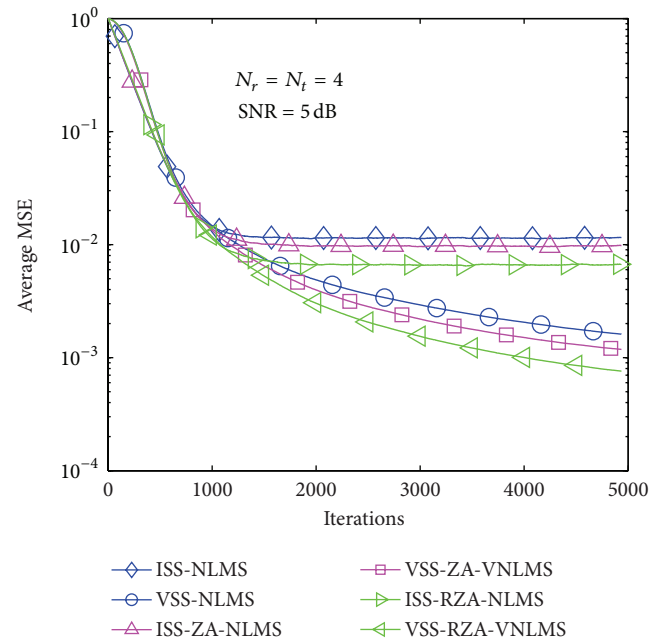
FIGURE 7: Average MSE performance versus received iterations ($T = 4$).

previous methods in Figure 6 ($T = 1$) are bigger than the gaps in Figure 7 ($T = 4$).

In the second example, system performance using proposed channel estimators is also evaluated with respect to BER performance. Two kinds of signal modulation schemes, that is, multiple PSK and multiple QAM, are considered. Received SNR is defined by E_0/N_0, where E_0 is the average

received power of symbol and N_0 is the noise power. In Figure 12, multiple PSK schemes, that is, QPSK, 8PSK, and 16PSK, are considered for data modulation and system performance was evaluated. One can find that there is no big performance difference using QPSK and 8PSK due to high transmission. If the higher-order data modulation is adopted, much better BER performance will be achieved
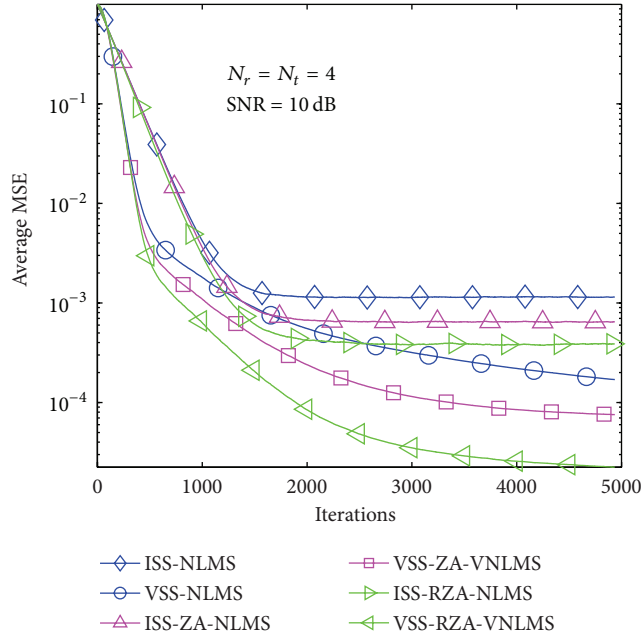
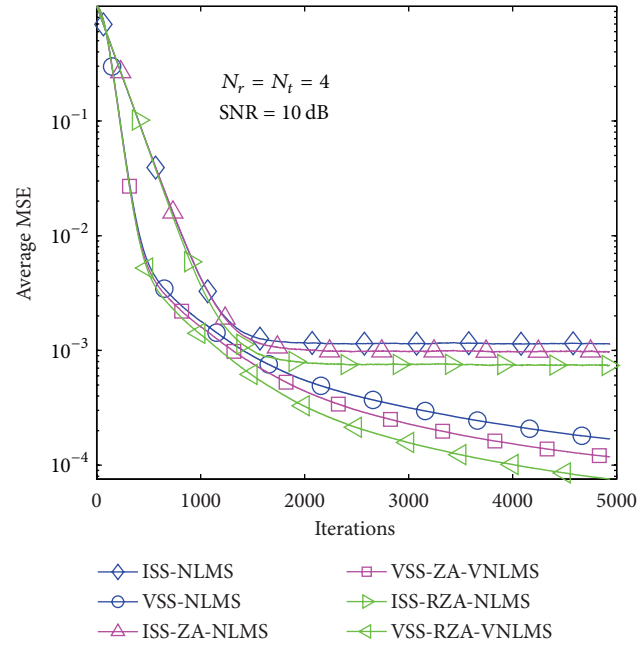FIGURE 8: Average MSE performance versus received iterations ($T = 1$).

FIGURE 9: Average MSE performance versus received iterations ($T = 4$).
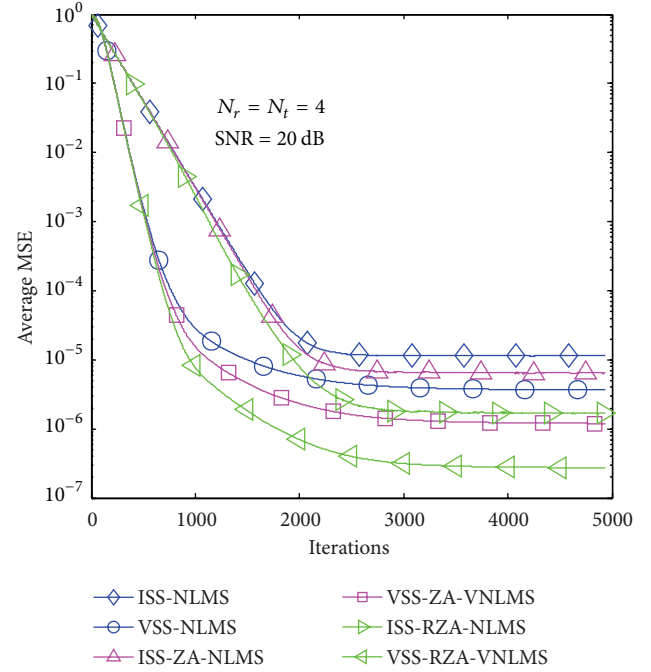

FIGURE 10: Average MSE performance versus received iterations ($T = 1$).

adopted as for a typical example of performance evaluation. We can find that 16QAM based system performance is better than 16PSK based system using the same channel estimators.

6. Conclusion

Traditional adaptive MIMO channel estimation methods utilize sparse ISS-NLMS algorithms using ISS. One of the main disadvantages of the traditional methods is the inability to balance the convergence speed and the estimation accuracy. In this paper, two sparse VSS-NLMS algorithms were proposed for estimating MIMO channels. Unlike the traditional sparse ISS-NLMS algorithms, the proposed algorithms utilized VSS which can change adaptively as the estimation error. Simulation results were provided to confirm the effectiveness of the proposed methods in three aspects: convergence speed, estimation performance, and system performance. First, convergence speed of the proposed methods using VSS is faster than ISS-NLMS based methods due to the fact that VSS for adaptive gradient descent is more efficient than ISS. In other words, VSS can well balance that fast convergence speed is dominant in the case of large estimation error while high accuracy is dominant in the case of small estimation error. Second, the proposed adaptive estimators can achieve better MSE gain than previous methods in different SNR regimes especially for sparser channels. At last, system performance using the proposed channel estimators can also achieve better BER performance than previous methods especially in high-order modulation signal based systems.
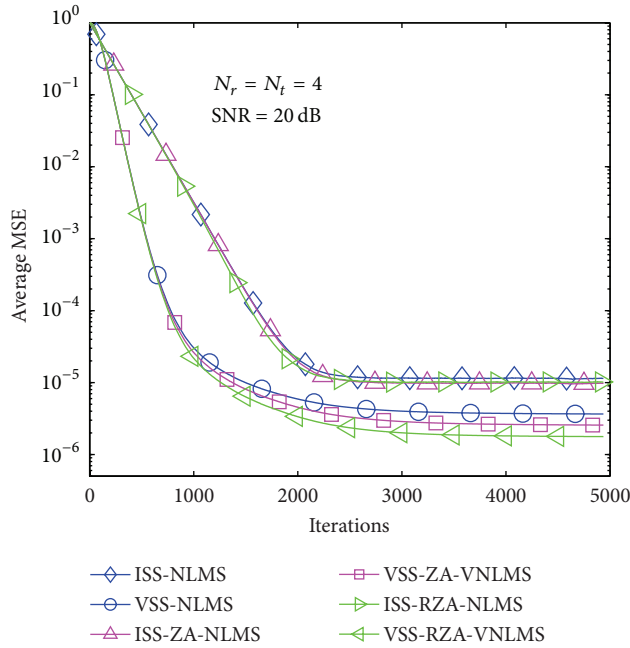
when compared with previous methods, that is, ISS-NLMS, VSS-NLMS, and sparse ISS-NLMS algorithms. In Figure 13, multiple QAM schemes, that is, 16QAM, 64QAM, and 128QAM, are considered for data modulation. One can easily find that the proposed method can achieve a better estimation than previous methods. In addition, we also compare the system performance with respect to different modulation schemes, PSK and QAM. In Figure 14, 16PSK and 16QAM are

FIGURE 11: Average MSE performance versus received iterations ($T = 4$).

FIGURE 13: Average BER performance versus received SNR (QAM).

FIGURE 12: Average BER performance versus received SNR (PSK).
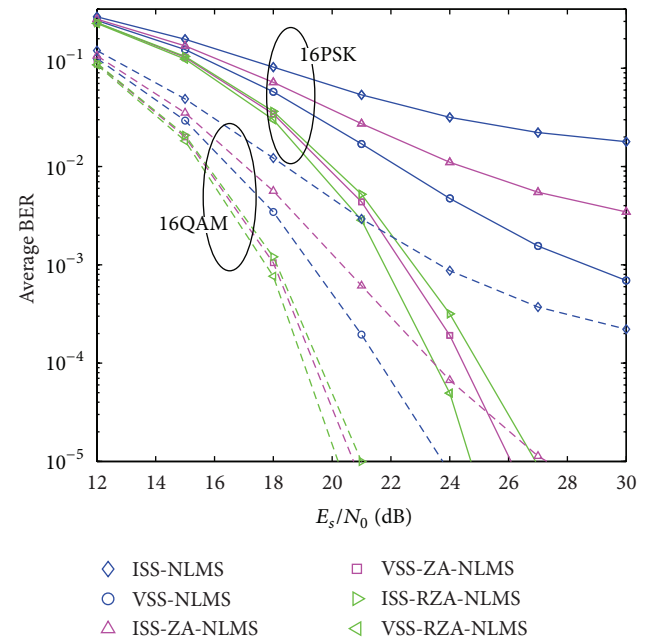
FIGURE 14: Comparison between 16PSK and 16QAM modulation schemes.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper. Authors of the paper do not have a direct financial relation that might lead to a conflict of interests for any of the authors.
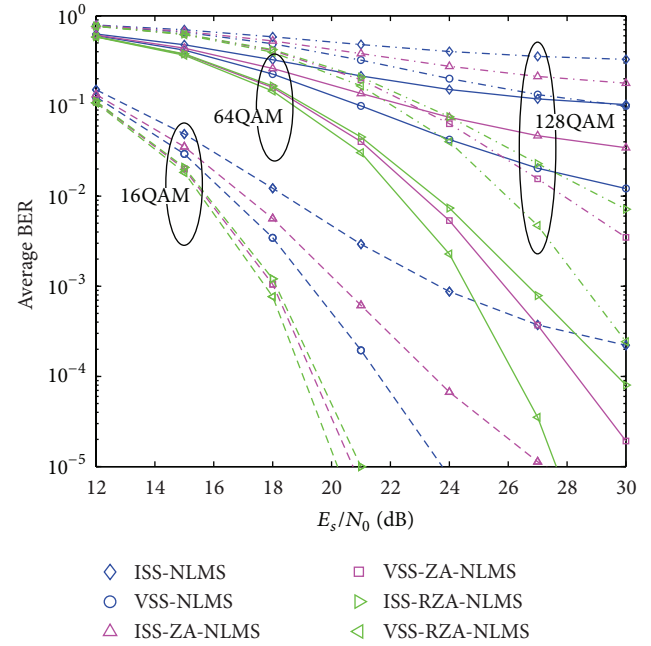
Acknowledgments

References

[1] F. Adachi and E. Kudoh, "New direction of broadband wireless technology," *Wireless Communications and Mobile Computing*, vol. 7, no. 8, pp. 969–983, 2007.

[2] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, 2002.

[3] J. Yue, K. J. Kim, J. D. Gibson, and R. A. Iltis, "Channel estimation and data detection for MIMO-OFDM systems," in *Proceeding of the IEEE Global Telecommunications Conference (GLOBECOM '03)*, vol. 2, pp. 581–585, San Francisco, Calif, USA, December 2003.

[4] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, 2006.

[5] T.-H. Chang, W.-C. Chiang, Y.-W. P. Hong, and C.-Y. Chi, "Training sequence design for discriminatory channel estimation in wireless MIMO systems," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6223–6237, 2010.

[6] S. He, J. K. Tugnait, and X. Meng, "On superimposed training for MIMO channel estimation and symbol detection," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 3007–3021, 2007.

[7] T.-H. Pham, Y.-C. Liang, and A. Nallanathan, "A joint channel estimation and data detection receiver for multiuser MIMO IFDMA systems," *IEEE Transactions on Communications*, vol. 57, no. 6, pp. 1857–1865, 2009.

[8] Z. J. Wang, Z. Han, and K. J. R. Liu, "A MIMO-OFDM channel estimation approach using time of arrivals," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1207–1213, 2005.

[9] G. Tauböck, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255–271, 2010.

[10] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: a new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[11] N. Wang, G. Gui, Z. Zhang, T. Tang, and J. Jiang, "A novel sparse channel estimation method for multipath MIMO-OFDM systems," in *Proceeding of the 74th IEEE Vehicular Technology Conference (VTC-Fall '11)*, pp. 1–5, San Francisco, Calif, USA, September 2011.

[12] L. Dai, Z. Wang, and Z. Yang, "Compressive sensing based time domain synchronous OFDM transmission for vehicular communications," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 460–469, 2013.

[13] L. Dai, Z. Wang, and Z. Yang, "Next-generation digital television terrestrial broadcasting systems: key technologies and research trends," *IEEE Communications Magazine*, vol. 50, no. 6, pp. 150–158, 2012.

[14] N. Czink, X. Yin, H. Özcelik, M. Herdin, E. Bonek, and B. H. Fleury, "Cluster characteristics in a MIMO indoor propagation environment," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1465–1475, 2007.

[15] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[16] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[17] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," *IEICE Transactions on Communications*, vol. 96, no. 3, pp. 685–712, 2013.

[18] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[19] A. Amini, M. Unser, and F. Marvasti, "Compressibility of deterministic and random infinite sequences," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5193–5201, 2011.

[20] G. Gui, W. Peng, and F. Adachi, "Improved adaptive sparse channel estimation based on the least mean square algorithm," in *Proceeding of the IEEE Wireless Communications and Networking Conference (WCNC '13)*, pp. 3105–3109, Shanghai, China, April 2013.

[21] G. Gui and F. Adachi, "Adaptive sparse channel estimation for time-variant MIMO-OFDM systems," in *Proceeding of the 9th International Wireless Communications & Mobile Computing Conference (IWCMC '13)*, pp. 1–6, Cagliari, Italy, July 2013.

[22] G. Gui, S. Kumagai, A. Mehbodniya, and F. Adachi, "Variable is good: adaptive sparse channel estimation using VSS-ZA-NLMS algorithm," in *Proceeding of the International Conference on Wireless Communications and Signal Processing (WCSP '13)*, pp. 291–295, Hangzhou, China, October 2013.

[23] G. Gui and F. Adachi, "Improved adaptive sparse channel estimation using least mean square algorithm," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, pp. 1–18, 2013.

[24] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[25] Y. Chen, Y. Gu, and A. O. Hero III, "Sparse LMS for system identification," in *Proceeding of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09)*, pp. 3125–3128, Taipei, Taiwan, April 2009.

[26] Z. Huang, G. Gui, A. Huang, D. Xiang, and F. Adachi, "Regularization selection method for LMS-type sparse multipath channel estimation," in *Proceeding of the 19th Asia-Pacific Conference on Communications (APCC '13)*, pp. 655–660, Bali Island, Indonesia, August 2013.

[27] G. Gui, A. Mehbodniya, and F. Adachi, "Least mean square/fourth algorithm for adaptive sparse channel estimation," in *Proceeding of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '13)*, pp. 1–5, London, UK, September 2013.

[28] B. Widrow and P. N. Stearns, *Adaptive Signal Processing*, Prentice Hall, New Jersey, NJ, USA, 4 edition, 1985.

[29] H.-C. Shin, A. H. Sayed, and W.-J. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 132–135, 2004.

[30] H. Cho and S. W. Kim, "Variable step-size normalized LMS algorithm by approximating correlation matrix of estimation error," *Signal Processing*, vol. 90, no. 9, pp. 2792–2799, 2010.