# ADAPTIVE BLIND SIMO IDENTIFICATION: DERIVATION OF AN OPTIMAL STEP SIZE FOR THE UNCONSTRAINED MULTICHANNEL LMS ALGORITHM

*Yiteng (Arden) Huang*†*, Jacob Benesty*‡*, and Jingdong Chen*†

† Bell Laboratories, Lucent Technologies
600 Mountain Avenue
Murray Hill, New Jersey 07974, USA
{arden, jingdong}@research.bell-labs.com

‡ Université du Québec, INRS-EMT
800 de la Gauchetière Ouest, Suite 6900
Montréal, Québec, H5A 1K6, Canada
benesty@inrs-emt.uquebec.ca

**ABSTRACT**

Adaptive algorithms for blindly identifying SIMO systems are appealing because of their computational efficiency and capability of continuously tracking a time-varying system. Adaptive multichannel LMS (MCLMS) algorithms (with and without the unit-norm constraint) are analyzed and the optimal step size is derived. A simple yet effective variable step-size unconstrained MCLMS algorithm is proposed and its performance is evaluated with simulations.

## 1. INTRODUCTION

In this paper, we consider blind identification of a single-input multiple-output (SIMO) system. In a SIMO FIR linear system as depicted in Fig. 1, the $i$-th observation $x_i(n)$ is expressed as follows:

$$x_i(n) = h_{\mathrm{t},i} * s(n) + b_i(n), \;\; i = 1, 2, \cdots, M, \tag{1}$$

where $s(n)$ represents the common source signal, $h_{\mathrm{t},i}$ stands for the true (subscript t) impulse response of the $i$-th channel, $b_i(n)$ is the additive noise signal captured by the $i$-th sensor, the symbol $*$ denotes the linear convolution operator, and $M$ is the number of channels. In a vector/matrix form, such a relationship (1) becomes:

$$\mathbf{x}_i(n) = \mathbf{H}_{\mathrm{t},i} \cdot \mathbf{s}(n) + \mathbf{b}_i(n), \tag{2}$$

where

$$\begin{aligned}
\mathbf{x}_i(n) &= \begin{bmatrix} x_i(n) & x_i(n-1) & \cdots & x_i(n-L+1) \end{bmatrix}^T, \\
\mathbf{H}_{\mathrm{t},i} &= \begin{bmatrix} h_{\mathrm{t},i,0} & \cdots & h_{\mathrm{t},i,L-1} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & h_{\mathrm{t},i,0} & \cdots & h_{\mathrm{t},i,L-1} \end{bmatrix}, \\
\mathbf{s}(n) &= \begin{bmatrix} s(n) & s(n-1) & \cdots & s(n-2L+2) \end{bmatrix}^T, \\
\mathbf{b}_i(n) &= \begin{bmatrix} b_i(n) & b_i(n-1) & \cdots & b_i(n-L+1) \end{bmatrix}^T,
\end{aligned}$$

$(\cdot)^T$ denotes a vector/matrix transpose, and $L$ is set to the length of the longest channel impulse response by assumption. Additive noise components in different channels are assumed to be uncorrelated with the source signal even though they might be mutually dependent. The channel parameter matrix $\mathbf{H}_{\mathrm{t},i}$ is of dimension $L \times (2L-1)$ and is constructed from the channel's impulse response:

$$\mathbf{h}_{\mathrm{t},i} = \begin{bmatrix} h_{\mathrm{t},i,0} & h_{\mathrm{t},i,1} & \cdots & h_{\mathrm{t},i,L-1} \end{bmatrix}^T. \tag{3}$$

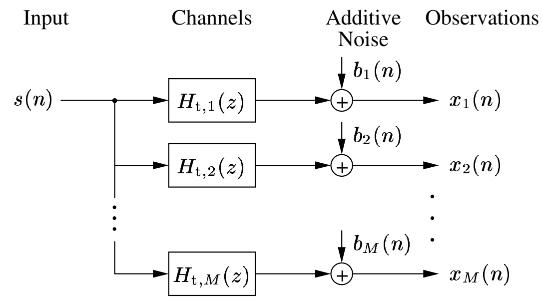


Figure 1: Illustration of the relationships between the input $s(n)$ and the observations $x_i(n)$ in a single-input multiple-output FIR system.

A blind channel identification (BCI) algorithm is to estimate the channel impulse responses $\mathbf{h}_{\mathrm{t},i}$, $i = 1, 2, \cdots, M$, from the observations $\mathbf{x}_i(n)$ without utilizing any knowledge about the source signal $\mathbf{s}(n)$.

The idea of BCI was first proposed by Sato [1]. The technology has a variety of potential applications in wireless communications and other signal processing systems. From the pioneering work of Tong et al. [2], it is well know that a SIMO system can be blindly identified using only second-order statistics (SOS) of the outputs if the following two conditions are met (assumed throughout this paper):

1. The polynomials formed from $\mathbf{h}_{\mathrm{t},i}$, $i = 1, 2, \cdots, M$, are co-prime, i.e., the channel transfer functions $H_{\mathrm{t},i}(z)$ do not share any common zeros;
2. The autocorrelation matrix $\mathbf{R}_{ss} = E\left\{\mathbf{s}(n)\mathbf{s}^T(n)\right\}$ of the source signal is of full rank (such that the SIMO system can be fully excited from a perspective of system identification), where $E\{\cdot\}$ denotes mathematical expectation.

As research in this area advances and demands for efficient implementation emerge, developing BCI algorithms becomes imperative. Two important proposals, based only on the SOS of the system outputs, were proposed. One is the adaptive multichannel LMS (MCLMS) algorithm (with a unit-norm constraint on the channel impulse response vector) [3], and the other is the unconstrained MCLMS (UMCLMS) algorithm [4]. Both algorithms work well for an identifiable, slowly time-varying SIMO system of moderately long channels like most wireless communication systems. But the step size governs the rate of convergence and the steady-state misalignment error. A fixed step size usually cannot meet the conflicting requirement of fast convergence and low misalignment. Moreover, in order to prevent the algorithms from diverging, several trials need to be conducted before a proper step size is found. This drawback obviously will obstruct the use of these adaptive algorithms in practice.

In this paper, we will derive the optimal step size for the UMCLMS algorithm, which minimizes the misalignment error in each step of adaptation. Using this discovery, we will develop a variable step-size UMCLMS (VSS-UMCLMS) algorithm. The effectiveness of this step-size control scheme will be justified by simulations.

## 2. BCI FUNDAMENTALS AND ADAPTIVE MULTICHANNEL LMS ALGORITHM

For a SIMO system, the vector of channel impulse responses lies in the null space of the cross-correlation like matrix of channel outputs [6]:

$$\mathbf{R}_x \mathbf{h}_{\mathrm{t}} = \mathbf{0}, \tag{4}$$

where

$$\begin{aligned}
\mathbf{R}_x &= \begin{bmatrix} \sum_{i\neq 1}\mathbf{R}_{x_ix_i} & -\mathbf{R}_{x_2x_1} & \cdots & -\mathbf{R}_{x_Mx_1} \\ -\mathbf{R}_{x_1x_2} & \sum_{i\neq 2}\mathbf{R}_{x_ix_i} & \cdots & -\mathbf{R}_{x_Mx_2} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbf{R}_{x_1x_M} & -\mathbf{R}_{x_2x_M} & \cdots & \sum_{i\neq M}\mathbf{R}_{x_ix_i} \end{bmatrix}, \\
\mathbf{R}_{x_ix_j} &= E\left\{\mathbf{x}_i(n)\mathbf{x}_j^T(n)\right\}, \ \ i,j=1,2,\cdots,M, \\
\mathbf{h}_{\mathrm{t}} &= \begin{bmatrix} \mathbf{h}_{\mathrm{t},1}^T & \mathbf{h}_{\mathrm{t},2}^T & \cdots & \mathbf{h}_{\mathrm{t},M}^T \end{bmatrix}^T.
\end{aligned}$$

For a blindly identifiable SIMO system with the two assumptions made in the previous section, Matrix $\mathbf{R}_x$ is rank deficient by 1 in the absence of noise and channel impulse responses can be uniquely determined from $\mathbf{R}_x$, which contains only the SOS of the system outputs. When noise is present, $\mathbf{h}_{\mathrm{t}}$ would be the eigenvector of $\mathbf{R}_x$ corresponding to its smallest eigenvalue.

To develop an adaptive BCI implementation, a simple way is to take advantage of the cross relations among the outputs, as we did in an earlier study of the adaptive multi-channel LMS (MCLMS) algorithm [3]. By following the fact that

$$x_i * h_{\mathrm{t},j} = s * h_{\mathrm{t},i} * h_{\mathrm{t},j} = x_j * h_{\mathrm{t},i}, \quad i,j=1,2,\cdots,M,\ i\neq j, \tag{5}$$

we have, in the absence of noise, the following cross relation at time $n$:

$$\mathbf{x}_i^T(n)\mathbf{h}_{\mathrm{t},j} = \mathbf{x}_j^T(n)\mathbf{h}_{\mathrm{t},i}, \ \ i,j=1,2,\cdots,M,\ i\neq j. \tag{6}$$

When noise is present and/or the estimate of channel impulse responses deviates from the true value, an *a priori* error signal is produced:

$$e_{ij}(n+1) = \mathbf{x}_i^T(n+1)\mathbf{h}_j(n) - \mathbf{x}_j^T(n+1)\mathbf{h}_i(n), \quad i,j=1,2,\cdots,M, \tag{7}$$

where $\mathbf{h}_i(n)$ is the model filter for the $i$-th channel at time $n$. In order to avoid the trivial estimate of all zero elements, a unit-norm constraint is imposed on

$$\mathbf{h}(n) = \begin{bmatrix} \mathbf{h}_1^T(n) & \mathbf{h}_2^T(n) & \cdots & \mathbf{h}_M^T(n) \end{bmatrix}^T,$$

leading to the normalized error signal $\epsilon_{ij}(n+1) = e_{ij}(n+1)/\|\mathbf{h}(n)\|$. Accordingly, the cost function is formulated as:

$$J(n+1) = \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} \epsilon_{ij}^2(n+1), \tag{8}$$

and the update equation of the MCLMS algorithm is deduced as follows:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \mu\nabla J(n+1), \tag{9}$$

where $\mu$ is a small positive step size,

$$\begin{aligned}
\nabla J(n+1) &= \frac{\partial J(n+1)}{\partial \mathbf{h}(n)} \\
&= \frac{2\left[\tilde{\mathbf{R}}_x(n+1)\mathbf{h}(n) - J(n+1)\mathbf{h}(n)\right]}{\|\mathbf{h}(n)\|^2},
\end{aligned} \tag{10}$$

$$\tilde{\mathbf{R}}_x(n) = \begin{bmatrix} \sum\limits_{i\neq 1}\tilde{\mathbf{R}}_{x_ix_i}(n) & -\tilde{\mathbf{R}}_{x_2x_1}(n) & \cdots & -\tilde{\mathbf{R}}_{x_Mx_1}(n) \\ -\tilde{\mathbf{R}}_{x_1x_2}(n) & \sum\limits_{i\neq 2}\tilde{\mathbf{R}}_{x_ix_i}(n) & \cdots & -\tilde{\mathbf{R}}_{x_Mx_2}(n) \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{\mathbf{R}}_{x_1x_M}(n) & -\tilde{\mathbf{R}}_{x_2x_M}(n) & \cdots & \sum\limits_{i\neq M}\tilde{\mathbf{R}}_{x_ix_i}(n) \end{bmatrix},$$

and

$$\tilde{\mathbf{R}}_{x_ix_j}(n) = \mathbf{x}_i(n)\mathbf{x}_j^T(n), \ \ i,j=1,2,\cdots,M.$$

## 3. OPTIMAL STEP SIZE AND THE PROPOSED VARIABLE STEP-SIZE UMCLMS ALGORITHM

It was shown in [3] that the MCLMS algorithm is able to converge in the mean to the true channel impulse response vector $\mathbf{h}_{\mathrm{t}}$ if the step size $\mu$ is properly specified. However, there is no guide on how to choose $\mu$ in practice. In order to avoid divergence, a conservatively small $\mu$ is usually used, which inevitably sacrifices the convergence speed of the adaptive algorithm. In this section we will show what is the optimal step size for the UMCLMS algorithm and propose a variable step-size UMCLMS algorithm.

We begin with re-examining the update equation (9). As the adaptive algorithm proceeds, the cost function $J(n+1)$ diminishes and its gradient with respect to $\mathbf{h}(n)$ can be approximated as

$$\nabla J(n+1) \approx \frac{2\tilde{\mathbf{R}}_x(n+1)\mathbf{h}(n)}{\|\mathbf{h}(n)\|^2}. \tag{11}$$

If we remove the unit-norm constraint, a simplified UMCLMS adaptive algorithm is deduced:

$$\mathbf{h}(n+1) = \mathbf{h}(n) - 2\mu\tilde{\mathbf{R}}_x(n+1)\mathbf{h}(n), \tag{12}$$

which is theoretically equivalent to the adaptive algorithm proposed in [4] although the cost functions are defined in different ways in these two adaptive BCI algorithms.

With such a simplified adaptive algorithm, the primary concern is whether it would converge to the trivial all-zero estimate. Fortunately this will not happen as long as the initial estimate $\mathbf{h}(0)$ is not orthogonal to the true channel impulse response vector $\mathbf{h}_{\mathrm{t}}$, as shown in [4]. This can be easily demonstrated by pre-multiplying (12) with $\mathbf{h}_{\mathrm{t}}^T$:

$$\mathbf{h}_{\mathrm{t}}^T\mathbf{h}(n+1) = \mathbf{h}_{\mathrm{t}}^T\mathbf{h}(n) - 2\mu\mathbf{h}_{\mathrm{t}}^T\tilde{\mathbf{R}}_x(n+1)\mathbf{h}(n). \tag{13}$$

Using the cross relation (6), we know $\mathbf{h}_\mathrm{t}^T\tilde{\mathbf{R}}_x(n+1) = \mathbf{0}^T$ in the absence of noise. This implies that the gradient $\nabla J(n+1)$ is orthogonal to $\mathbf{h}_\mathrm{t}$ at any time $n$. As a result, (13) turns out to be

$$\mathbf{h}_\mathrm{t}^T\mathbf{h}(n+1) = \mathbf{h}_\mathrm{t}^T\mathbf{h}(n). \tag{14}$$

This indicates that $\mathbf{h}_\mathrm{t}^T\mathbf{h}(n)$ is time-invariant for the UMCLMS algorithm. Provided that $\mathbf{h}_\mathrm{t}^T\mathbf{h}(0) \neq 0$, $\mathbf{h}(n)$ would not converge to zero.

Decompose the model filter $\mathbf{h}(n)$ as follows:

$$\mathbf{h}(n) = \mathbf{h}_\perp(n) + \mathbf{h}_\|(n), \tag{15}$$

where $\mathbf{h}_\perp(n)$ and $\mathbf{h}_\|(n)$ are perpendicular and parallel to $\mathbf{h}_\mathrm{t}$, respectively. Since the gradient $\nabla J(n+1)$ is orthogonal to $\mathbf{h}_\mathrm{t}$ and $\mathbf{h}_\mathrm{t}$ is parallel to $\mathbf{h}_\|(n)$, obviously $\nabla J(n+1)$ is orthogonal to $\mathbf{h}_\|(n)$ as well. Therefore, the update equation (12) of the UMCLMS algorithm can be decomposed into the following two separate equations:

$$\begin{aligned}\mathbf{h}_\perp(n+1) &= \mathbf{h}_\perp(n) - \mu\nabla J(n+1), & (16)\\ \mathbf{h}_\|(n+1) &= \mathbf{h}_\|(n). & (17)\end{aligned}$$

From (16) and (17), it is clear that the UMCLMS algorithm adapts the model filter only in the direction that is perpendicular to $\mathbf{h}_\mathrm{t}$. The component $\mathbf{h}_\|(n)$ is not altered in the process of adaptation.

As far as a general system identification algorithm is concerned, the most important performance measure apparently should be the difference between the true channel impulse response and the estimate. With a BCI method, the SIMO FIR system can be blindly identified up to a scale. Therefore, the misalignment of an estimate $\mathbf{h}(n)$ with respect to the true channel impulse response vector $\mathbf{h}_\mathrm{t}$ would be:

$$d(n) = \min_\alpha \|\mathbf{h}_\mathrm{t} - \alpha\mathbf{h}(n)\|^2, \tag{18}$$

where $\alpha$ is an arbitrary scale. Substituting (15) into (18) and finding the minimum produces

$$\begin{aligned}d(n) &= \min_\alpha \left[\|\mathbf{h}(n)\|^2\alpha^2 - 2\|\mathbf{h}_\|\|\|\mathbf{h}_\mathrm{t}\|\alpha + \|\mathbf{h}_\mathrm{t}\|^2\right]\\ &= \frac{\|\mathbf{h}_\mathrm{t}\|^2}{1 + \left(\|\mathbf{h}_\|(n)\|/\|\mathbf{h}_\perp(n)\|\right)^2}. & (19)\end{aligned}$$

Clearly the ratio of $\|\mathbf{h}_\|(n)\|$ over $\|\mathbf{h}_\perp(n)\|$ reflects how close the estimate is from the desired solution. With this feature in mind, the optimal step size $\mu_\mathrm{o}(n+1)$ for the UMCLMS algorithm at time $n+1$ would be the one that makes $\mathbf{h}_\perp(n+1)$ have a minimum norm, i.e.,

$$\begin{aligned}\mu_\mathrm{o}(n+1) &= \arg\min_\mu \|\mathbf{h}_\perp(n+1)\|\\ &= \arg\min_\mu \|\mathbf{h}_\perp(n) - \mu\nabla J(n+1)\|. & (20)\end{aligned}$$

Since $\mathbf{h}_\|(n+1)$ is time-invariant and $\mathbf{h}_\|(n+1)$ is orthogonal to $\mathbf{h}_\perp(n+1)$, minimizing the norm of $\mathbf{h}_\perp(n+1)$ is equivalent to minimizing the norm of $\mathbf{h}(n+1)$. As such, we have:

$$\begin{aligned}\mu_\mathrm{o}(n+1) &= \arg\min_\mu \|\mathbf{h}(n+1)\|\\ &= \arg\min_\mu \|\mathbf{h}(n) - \mu\nabla J(n+1)\|. & (21)\end{aligned}$$

In order to minimize the norm of $\mathbf{h}(n+1) = \mathbf{h}(n) - \mu(n+1)\nabla J(n+1)$, as illustrated in Fig. 2, $\mu(n+1)$ should be chosen such that $\mathbf{h}(n+1)$ is orthogonal to $\nabla J(n+1)$. Therefore, we project $\mathbf{h}(n)$ onto $\nabla J(n+1)$ and obtain the optimal step size:

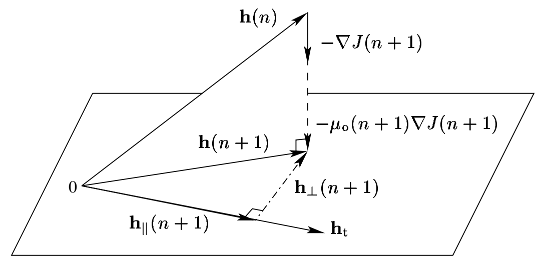


Figure 2: Illustration of the optimal step size $\mu_\mathrm{o}(n+1)$ for the unconstrained MCLMS BCI algorithm in a 3-dimensional space.

$$\mu_\mathrm{o}(n+1) = \frac{\mathbf{h}^T(n)\nabla J(n+1)}{\|\nabla J(n+1)\|^2}. \tag{22}$$

Finally, this new adaptive algorithm with the optimal step size is referred to as the variable step-size unconstrained MCLMS (VSS-UMCLMS) for BCI.

## 4. SIMULATIONS

In this section, we will evaluate the performance of the proposed VSS-UMCLMS algorithm by simulations. A comparison to the UMCLMS and MCLMS algorithms with a number of different pre-specified step sizes is also presented.

Similar to our earlier studies on BCI, we use the normalized projection misalignment (NPM) as a performance measure in this paper, which is given by:

$$\mathrm{NPM}(n) \triangleq \frac{\|\boldsymbol{\varepsilon}(n)\|}{\|\mathbf{h}_\mathrm{t}\|}, \tag{23}$$

where

$$\boldsymbol{\varepsilon}(n) = \mathbf{h}_\mathrm{t} - \frac{\mathbf{h}_\mathrm{t}^T\mathbf{h}(n)}{\mathbf{h}^T(n)\mathbf{h}(n)}\mathbf{h}(n)$$

is the projection misalignment vector. By projecting $\mathbf{h}_\mathrm{t}$ onto $\mathbf{h}(n)$ and defining a projection error, we take into account only the undesirable misalignment of the channel estimate, disregarding an arbitrary gain factor inherently associated with it [7].

The SIMO FIR system to be identified consists of $M = 3$ channels. The impulse response of each channel has $L = 32$ taps and their coefficients are randomly generated. Fig. 3 plots these three impulse responses, which have been checked to ensure that they do not share any common zeros. The source and additive noise signals are uncorrelated and both are white Gaussian random sequences. The sampling rate is 8 kHz. The model filter is initialized as $\mathbf{h}(0) = \mathbf{1}$ for all investigated adaptive algorithms. Note that $\mathbf{h}_\mathrm{t}^T\mathbf{h}(0) = 0.3915 \neq 0$.

Fig. 4 shows the convergence in terms of the NPM for all the algorithms. In Panel (a) noise is absent and in Panel (b) the signal-to-noise ratio (SNR) is 30 dB. Regarding the UMCLMS and MCLMS algorithms, a number of different step sizes were tried and two sets of results with $\mu = 0.025$ and $0.01$ are presented here. We see that increasing the step size would accelerate the UMCLMS and MCLMS algorithms to converge. But this

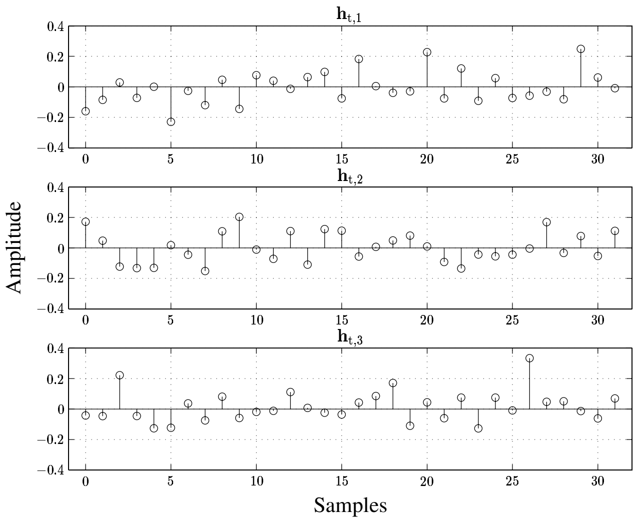


Figure 3: Impulse responses of a single-input three-output system used in the simulation for BCI.

trend fails when $\mu$ is greater than 0.025, around where those two algorithms start diverging. The difference between them in these simulations is insignificant. The MCLMS perferms better than the UMCLMS only when their step sizes are close to the critical value of divergence. The proposed VSS-UMCLMS algorithm converges much faster than the UMCLMS and MCLMS algorithms both in the absence and presence of noise. Furthermore, the final NPM for the VSS-UMCLMS algorithm is also smaller.

## 5. CONCLUSIONS

The optimal step size of the adaptive multichannel LMS (MCLMS) algorithm for blind SIMO identification was derived and a variable step-size unconstrained MCLMS algorithm was proposed. Compared with the conventional unconstrained MCLMS algorithm, the proposed method converges much faster and yields more accurate estimate of the system's channel impulse responses, as demonstrated by simulations. In addition, the proposed method is much easier to use in practice since the step size does not have to be specified in advance.

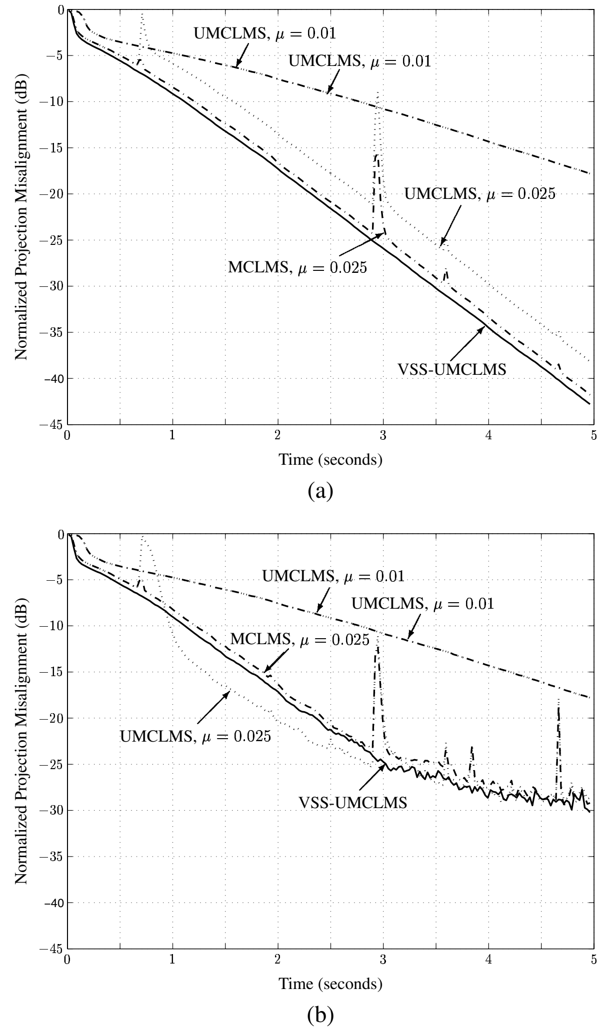


Figure 4: Normalized projection misalignment of the VSS-UMCLMS, UMCLMS, and MCLMS algorithms (a) in the absence of noise, and (b) at 30-dB SNR.

## 6. REFERENCES


[1] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation," *IEEE Trans. Commun.*, vol. COM-23, no. 6, pp. 679–682, Jun. 1975.

[2] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conf. on Signals, Systems, and Computers*, 1991, vol. 2, pp. 856–860.

[3] Y. Huang and J. Benesty, "Adaptive multi-channel least mean square and Newton algorithms for blind channel identification," *Elsevier Science Signal Processing*, vol. 82, pp. 1127–1138, Aug. 2002.

[4] H. Chen, X. Cao, and J. Zhu, "Convergence of stochastic-approximation-based algorithms for blind channel identification," *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1214–1225, May 2002.

[5] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[6] C. Avendano, J. Benesty, and D. R. Morgan, "A least squares component normalization approach to blind channel identification," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1999, vol. 4, pp. 1797–1800.

[7] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Processing Lett.*, vol. 5, no. 7, pp. 174–176, Jul. 1998.