ADAPTIVE BLIND CHANNEL IDENTIFICATION: MULTI-CHANNEL LEAST MEAN SQUARE AND NEWTON ALGORITHMS

Yiteng (Arden) Huang and Jacob Benesty

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

ABSTRACT

The problem of identifying a single-input multiple-output FIR system without a training signal, the so-called blind system identification, is addressed and two adaptive multi-channel approaches, least mean square (LMS) and Newton algorithms, are proposed. In contrast to the existing batch blind channel identification schemes, the proposed algorithms construct an error signal based on the cross relations between different channels in a novel, systematic way. The corresponding cost (error) function is easy to manipulate and facilitates the use of adaptive filtering methods for an efficient blind channel identification scheme. It is theoretically shown and practically demonstrated by numerical studies that the proposed algorithms converge in the mean to the desired channel impulse responses for an identifiable system.

1. INTRODUCTION

The desire for blind channel identification and estimation technique arises from a variety of potential applications in signal processing and communications, e.g. dereverberation, separation of speech from multiple sources, time delay estimation, speech enhancement, image deblurring, wireless communications, etc. In all these applications, *a priori* knowledge of the source signal is either inaccessible or very expensive to aquire, making the blind method a necessity.

In this paper, we consider the blind estimation of the impulse responses of a single-input multiple-output (SIMO) FIR system as presented in Fig. 1. The i-th observation $x_i(n)$ is the result of a linear convolution between the source signal $s(n)$ and the corresponding channel response h_i, corrupted by an additive noise $b_i(n)$:

$$x_i(n) = h_i * s(n) + b_i(n), \;\; i = 1, 2, ..., M, \tag{1}$$

where the $*$ symbol is the linear convolution operator and M is the number of channels. In a vector form, the relationship of the input and the observation for the i-th channel is written as:

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

where

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

and $(\cdot)^T$ denotes a vector/matrix transpose. The 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 $\mathbf{H}_i$ is of dimension $L \times (2L-1)$ and is constructed from the channel's impulse response:

$$\mathbf{h}_i = [h_{i,0}\; h_{i,1}\; \cdots\; h_{i,L-1}]^T, \tag{3}$$

where L is set to the length of the longest channel impulse response by assumption.

Early studies [1] of blind channel identification and equalization focused primarily on higher (than second) order statistics-based schemes. These schemes suffer from slow convergence and local minima, and therefore are unsatisfactory in tracking a fast time-varying system. In 1991, Tong *et al.* [2] demonstrated the possibility of using only second-order statistics (SOS) of multichannel system outputs to solve the channel identification problem. Since then, many SOS-based approaches have been proposed, such as the subspace (SS) algorithm [3] and the cross relation (CR) algorithm [4] (see [5] for a review on the SOS methods and the references therein). These batch methods can accurately estimate an identifiable SIMO system using a finite number of samples when the signal-to-noise ratio (SNR) is high. However, while these algorithms are able to yield a good estimate of the channel impulse responses, they are in general computationally intensive and are difficult to implement in an adaptive mode [5].

For blind channel identification to be practically useful in real-time applications, it is imperative that the algorithm should be computationally simple and can be adaptively implemented. As a result of our effort in this direction, two adaptive algorithms, namely the multi-channel least mean square (MCLMS), which is a generalization of the adaptive eigenvalue decomposition algorithm [6], and the multi-channel Newton (MCN) algorithms, are proposed in this paper.

2. ADAPTIVE MULTI-CHANNEL LMS AND NEWTON ALGORITHMS

2.1. Multi-Channel LMS Algorithm

When the input signal is unknown, the cross-relation between the sensor outputs can be exploited to estimate the channel impulse responses. By following the fact

$$x_i * h_j = s * h_i * h_j = x_j * h_i, \;\; i, j = 1, 2, ..., M, \; i \neq j, \tag{4}$$

0-7803-7402-9/02/$17.00 ©2002 IEEE

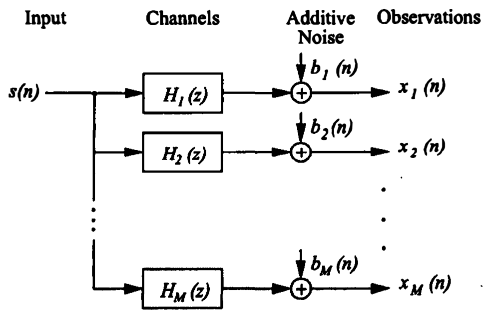

Figure 1: Illustration of the relationships between the input $s(n)$ and the observations $x_i(n)$ in a single-input multi-channel FIR system.

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

$$\mathbf{x}_i^T(n)\mathbf{h}_j = \mathbf{x}_j^T(n)\mathbf{h}_i, \quad i,j = 1,2,...,M, \; i \neq j. \tag{5}$$

But in the presence of noise, this cross relation no longer holds and an error signal can be defined as follows:

$$e_{ij}(n) = \begin{cases} \mathbf{x}_i^T(n)\mathbf{h}_j - \mathbf{x}_j^T(n)\mathbf{h}_i, & i \neq j, \; i,j = 1,2,...,M \\ 0 & i = j, \; i,j = 1,2,...,M \end{cases}. \tag{6}$$

Here we have $(M-1)M/2$ distinct error signals $e_{ij}(n)$, which exclude the case $e_{ii}(n) = 0$ and count the $e_{ij}(n) = -e_{ji}(n)$ pair only once. Assuming that these error signals are equally important, we now define a cost function as follows:

$$\chi(n) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} e_{ij}^2(n). \tag{7}$$

Therefore, the channel impulse responses will be determined by minimizing this error function. In order to avoid a trivial estimate with all zero elements, a unit-norm constraint is imposed on $\mathbf{h} \triangleq \left[\mathbf{h}_1^T \; \mathbf{h}_2^T \; \cdots \; \mathbf{h}_M^T\right]^T$ at all times such that the error signal becomes:

$(M\times L) \times 1$

$$\epsilon_{ij}(n) = \frac{e_{ij}(n)}{\|\mathbf{h}(n)\|}, \tag{8}$$

and the corresponding cost function is given by:

$$J(n) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \epsilon_{ij}^2(n) = \frac{\chi(n)}{\|\mathbf{h}\|^2}. \tag{9}$$

Therefore, the desired solution for $\mathbf{h}$ is determined by minimizing the mean value of the cost function $J(n)$:

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h}} E\{J(n)\}. \tag{10}$$

Direct minimization is computationally intensive and may be even intractable when the channel impulse responses are long and the number of channels is large. Here, an LMS algorithm is proposed to solve this minimization problem efficiently:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) - \mu \, \nabla J(n)|_{\mathbf{h}=\hat{\mathbf{h}}(n)}, \tag{11}$$

where μ is a small positive step size and ∇ is a gradient operator.

In order to determine the gradient in (11), we take a derivative of $J(n)$ with respect to $\mathbf{h}$:

$$\nabla J(n) = \frac{\partial J(n)}{\partial \mathbf{h}} = \frac{1}{\|\mathbf{h}\|^2}\left[\frac{\partial \chi(n)}{\partial \mathbf{h}} - 2J(n)\mathbf{h}\right], \tag{12}$$

where

$$\frac{\partial \chi(n)}{\partial \mathbf{h}} = \left[\left(\frac{\partial \chi(n)}{\partial \mathbf{h}_1}\right)^T \; \left(\frac{\partial \chi(n)}{\partial \mathbf{h}_2}\right)^T \; \cdots \; \left(\frac{\partial \chi(n)}{\partial \mathbf{h}_M}\right)^T\right]^T.$$

Let us now evaluate the partial derivative of $\chi(n)$ only with respect to the coefficients of the k-th ($k = 1,2,...,M$) channel impulse response:

$$\begin{aligned} \frac{\partial \chi(n)}{\partial \mathbf{h}_k} &= \frac{\partial}{\partial \mathbf{h}_k}\left[\sum_{i=1}^{M-1}\sum_{j=i+1}^{M} e_{ij}^2(n)\right] \\ &= \sum_{i=1}^{k-1} 2e_{ik}\mathbf{x}_i(n) + \sum_{j=k+1}^{M} 2e_{kj}(n)\left(-\mathbf{x}_j(n)\right) \\ &= \sum_{i=1}^{k-1} 2e_{ik}\mathbf{x}_i(n) + \sum_{j=k+1}^{M} 2e_{jk}(n)\mathbf{x}_j(n) \\ &= \sum_{i=1}^{M} 2e_{ik}(n)\mathbf{x}_i(n), \end{aligned} \tag{13}$$

where the last step follows from the fact $e_{kk}(n) = 0$. We may express this equation concisely in matrix form as follows:

$$\begin{aligned} \frac{\partial \chi(n)}{\partial \mathbf{h}_k} &= 2\mathbf{X}(n)\mathbf{e}_k(n) \\ &= 2\mathbf{X}(n)[\mathbf{C}_k(n) - \mathbf{D}_k(n)]\mathbf{h}, \end{aligned} \tag{14}$$

where we have defined, for convenience,

$$\begin{aligned} \mathbf{X}(n) &= [\underset{L\times 1}{\mathbf{x}_1(n)} \; \underset{L\times 1}{\mathbf{x}_2(n)} \; \cdots \; \underset{L\times 1}{\mathbf{x}_M(n)}]_{L\times M}, \\ \mathbf{e}_k(n) &= [e_{1k}(n) \; e_{2k}(n) \; \cdots \; e_{Mk}(n)]^T \quad M\times 1 \\ &= \begin{bmatrix} \mathbf{x}_1^T(n)\mathbf{h}_k - \mathbf{x}_k^T(n)\mathbf{h}_1 \\ \mathbf{x}_2^T(n)\mathbf{h}_k - \mathbf{x}_k^T(n)\mathbf{h}_2 \\ \vdots \\ \mathbf{x}_M^T(n)\mathbf{h}_k - \mathbf{x}_k^T(n)\mathbf{h}_M \end{bmatrix} \\ &= [\mathbf{C}_k(n) - \mathbf{D}_k(n)]\mathbf{h}, \\ \mathbf{C}_k(n) &= \begin{bmatrix} \underbrace{\begin{matrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{matrix}}_{(k-1)L} & \begin{matrix} \mathbf{x}_1^T(n) \\ \mathbf{x}_2^T(n) \\ \vdots \\ \mathbf{x}_M^T(n) \end{matrix} & \underbrace{\begin{matrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{matrix}}_{(M-k)L} \end{bmatrix}_{M\times ML} \\ &= [\mathbf{0}_{M\times(k-1)L} \; \mathbf{X}^T(n) \; \mathbf{0}_{M\times(M-k)L}]_{M\times ML}, \\ \mathbf{D}_k(n) &= \begin{bmatrix} \mathbf{x}_k^T(n) & 0 & \cdots & 0 \\ 0 & \mathbf{x}_k^T(n) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_k^T(n) \end{bmatrix}_{M\times ML} \end{aligned}$$

Continuing, we evaluate the two matrix products in (14) individually as follows:

$$\mathbf{X}(n)\mathbf{C}_k(n) = \left[\mathbf{0}_{L\times(k-1)L}\ \sum_{i=1}^{M}\tilde{\mathbf{R}}_{x_ix_i}(n)\ \mathbf{0}_{L\times(M-k)L}\right], \quad (15)$$

$$\mathbf{X}(n)\mathbf{D}_k(n) = \left[\tilde{\mathbf{R}}_{x_1x_k}(n)\ \tilde{\mathbf{R}}_{x_2x_k}(n)\ \cdots\ \tilde{\mathbf{R}}_{x_Mx_k}(n)\right], \quad (16)$$

where

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

Next, substituting (15) and (16) into (14) yields

$$\frac{\partial\chi(n)}{\partial\mathbf{h}_k} = 2\left[-\tilde{\mathbf{R}}_{x_1x_k}(n)\cdots\sum_{i\neq k}\tilde{\mathbf{R}}_{x_ix_i}(n)\cdots-\tilde{\mathbf{R}}_{x_Mx_k}(n)\right]\mathbf{h}. \quad (17)$$

Thereafter, we incorporate (17) into (12) and obtain

$$\frac{\partial\chi(n)}{\partial\mathbf{h}} = 2\tilde{\mathbf{R}}(n)\mathbf{h}, \quad (18)$$

$$\nabla J(n) = \frac{1}{\|\mathbf{h}\|^2}\left[2\tilde{\mathbf{R}}(n)\mathbf{h} - 2J(n)\mathbf{h}\right], \quad (19)$$

where

$$\tilde{\mathbf{R}}(n) = \begin{bmatrix} \sum_{i\neq1}\tilde{\mathbf{R}}_{x_ix_i}(n) & -\tilde{\mathbf{R}}_{x_2x_1}(n) & \cdots & -\tilde{\mathbf{R}}_{x_Mx_1}(n) \\ -\tilde{\mathbf{R}}_{x_1x_2}(n) & \sum_{i\neq2}\tilde{\mathbf{R}}_{x_ix_i}(n) & \cdots & -\tilde{\mathbf{R}}_{x_Mx_2}(n) \\ \vdots & \vdots & \ddots & \vdots \\ -\tilde{\mathbf{R}}_{x_1x_M}(n) & -\tilde{\mathbf{R}}_{x_2x_M}(n) & \cdots & \sum_{i\neq M}\tilde{\mathbf{R}}_{x_ix_i}(n) \end{bmatrix}.$$

Finally, we substitute (19) into (11) and have the update equation:

$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) - \frac{2\mu}{\|\hat{\mathbf{h}}(n)\|^2}\left[\tilde{\mathbf{R}}(n)\hat{\mathbf{h}}(n) - J(n)\hat{\mathbf{h}}(n)\right]. \quad (20)$$

If the channel estimate is always normalized after each update, then we have the simplified algorithm

$$\hat{\mathbf{h}}(n+1) = \frac{\hat{\mathbf{h}}(n) - 2\mu\left[\tilde{\mathbf{R}}(n)\hat{\mathbf{h}}(n) - \chi(n)\hat{\mathbf{h}}(n)\right]}{\left\|\hat{\mathbf{h}}(n) - 2\mu\left[\tilde{\mathbf{R}}(n)\hat{\mathbf{h}}(n) - \chi(n)\hat{\mathbf{h}}(n)\right]\right\|}. \quad (21)$$

Assuming that the independence assumption [7] holds, it can be easily shown that the LMS algorithm converges in the mean if the step size satisfies

$$0 < \mu < \frac{1}{\lambda_{\max}}, \quad (22)$$

where $\lambda_{\max}$ is the largest eigenvalue of the matrix $E\{\tilde{\mathbf{R}}(n) - J(n)\mathbf{I}_{ML\times ML}\}$.

After convergence, taking the expectation of (20) gives

$$\mathbf{R}\frac{\hat{\mathbf{h}}(\infty)}{\|\hat{\mathbf{h}}(\infty)\|} = E\{J(\infty)\}\frac{\hat{\mathbf{h}}(\infty)}{\|\hat{\mathbf{h}}(\infty)\|}, \quad (23)$$

which is the desired result: $\hat{\mathbf{h}}$ converges in the mean to the eigenvector of $\mathbf{R} \triangleq E\{\tilde{\mathbf{R}}(n)\}$ corresponding to the smallest eigenvalue $E\{J(\infty)\}$.

2.2. Multi-Channel Newton Algorithm

In order to accelerate the convergence of the MCLMS algorithm, we present here a multi-channel Newton (MCN) algorithm with a variable step size during adaptation:

$$\hat{\mathbf{h}}_{n+1} = \hat{\mathbf{h}}_n - E^{-1}\left\{\nabla^2 J(n)\right\}\nabla J(n)\Big|_{\mathbf{h}=\hat{\mathbf{h}}_n}, \quad (24)$$

where $\nabla^2 J(n)$ is the Hessian matrix of $J(n)$ with respect to $\mathbf{h}$. Taking derivative of (19) with respect to $\mathbf{h}$ and using the unit norm constraint $\|\mathbf{h}\| = 1$ yields

$$\begin{aligned}\nabla^2 J(n) &= 2\left\{\tilde{\mathbf{R}}(n) - \mathbf{h}[\nabla J(n)]^T - J(n)\mathbf{I}_{ML\times ML}\right\} - \\ &\quad 4\left[\tilde{\mathbf{R}}(n)\mathbf{h} - J(n)\mathbf{h}\right]\mathbf{h}^T. \end{aligned} \quad (25)$$

Taking mathematical expectation of (25) and invoking the independence assumption [7] produces

$$\begin{aligned}E\{\nabla^2 J(n)\} &= 2\mathbf{R} - 4\mathbf{h}\mathbf{h}^T\mathbf{R} - 4\mathbf{R}\mathbf{h}\mathbf{h}^T - \\ &\quad 2E\{J(n)\}\left[\mathbf{I}_{ML\times ML} - 4\mathbf{h}\mathbf{h}^T\right]. \end{aligned} \quad (26)$$

In practice, $\mathbf{R}$ and $E\{J(n)\}$ are not known such that we have to estimate their values. Since $J(n)$ decreases as adaptation proceeds and is relative small particularly after convergence, we can neglect the term $E\{J(n)\}$ in (26) for simplicity and with appropriate accuracy, as suggested by simulations. The matrix $\mathbf{R}$ is estimated recursively in a conventional way as follows:

$$\begin{aligned}\hat{\mathbf{R}}(0) &= \left(\sum_{i=1}^{M}\frac{\mathbf{x}_i^T\mathbf{x}_i}{L}\right)\mathbf{I}_{ML\times ML}, \\ \hat{\mathbf{R}}(n) &= \lambda\hat{\mathbf{R}}(n-1) + \tilde{\mathbf{R}}(n),\ \text{for } n \geq 1, \end{aligned} \quad (27)$$

where λ $(0 < \lambda < 1)$ is an exponential forgetting factor.

By using these approximations, we finally deduce the multichannel Newton algorithm:

$$\hat{\mathbf{h}}(n+1) = \frac{\hat{\mathbf{h}}(n) - \rho\Delta\hat{\mathbf{h}}(n)}{\left\|\hat{\mathbf{h}}(n) - \rho\Delta\hat{\mathbf{h}}(n)\right\|}, \quad (28)$$

where ρ is a new step size, close to but less than 1, and

$$\begin{aligned}\Delta\hat{\mathbf{h}}(n) = &\left[\hat{\mathbf{R}}(n) - 2\hat{\mathbf{h}}(n)\hat{\mathbf{h}}(n)^T\hat{\mathbf{R}}(n) - 2\hat{\mathbf{R}}(n)\hat{\mathbf{h}}(n)\hat{\mathbf{h}}(n)^T\right]^{-1} \\ &\left[\tilde{\mathbf{R}}(n)\hat{\mathbf{h}}(n) - \chi(n)\hat{\mathbf{h}}(n)\right]. \end{aligned}$$

3. SIMULATIONS

Monte Carlo simulations were carried out to evaluate the proposed algorithms using short channel impulse responses that are common in digital communication problems. For comparison, the cross relation (CR) algorithm is also studied.

The normalized root mean square error (NRMSE) in dB is used as a performance measure in this paper and is defined as,

$$\text{NRMSE} \triangleq 20\log_{10}\left[\frac{1}{\|\mathbf{h}\|}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\|\epsilon^{(i)}\|^2}\right], \quad (29)$$

where N is the number of Monte-Carlo runs, $(\cdot)^{(i)}$ denotes a value obtained for the i-th run, and $\epsilon = \mathbf{h} - [(\mathbf{h}^T\hat{\mathbf{h}})/(\hat{\mathbf{h}}^T\hat{\mathbf{h}})]\hat{\mathbf{h}}$ is a projection error vector. By projecting $\mathbf{h}$ onto $\hat{\mathbf{h}}$ and defining a projection error, we take only the misalignment of the channel estimate into account [8].

In all simulations, the source signal was an uncorrelated binary phase shift keying sequence and the additive noise is i.i.d. zero mean Gaussian. The specified SNR is defined as SNR $\triangleq 10\log_{10}[\sigma_s^2\|\mathbf{h}\|^2/(M\sigma_b^2)]$, where σ_s^2 and σ_b^2 are the signal and noise powers, respectively.

The first simulation studied an ill-conditioned two-channel FIR system whose impulse responses in each channel are second order and are given by:

L=3 →
$$\mathbf{h}_1 = [1\ \ -2\cos(\pi/10)\ 1]^T, \quad \mathbf{h}_2 = [1\ \ -2\cos(\pi/5)\ 1]^T. \quad (30)$$

$N = 100$ Monte Carlo runs were performed. For the CR method, 50 samples of observations from each channel were used. For the proposed MCLMS and MCN algorithms, the step sizes $\mu = 0.01$ and $\rho = 0.95$ were fixed, respectively. The results are presented in Fig. 2. Panel (a) shows a comparison of NRMSE vs. SNR among the CR, MCLMS and MCN algorithms and Panel (b) compares convergence of the two proposed adaptive approaches. It is remarkable that all algorithms can determine the channel impulse responses with comparable accuracy. It cannot be concluded that the proposed MCLMS and MCN algorithms perform significantly better than the CR method particularly when the SNR is greater than 35 dB because the adaptive algorithms uses more samples in this simulation. As clearly seen, the MCN algorithm dramatically accelerates the convergence.

In the second simulation, we consider a three-channel system whose impulse responses are randomly determined and are longer ($L = 15$) than those in the first simulation. However, since more channels are used, it is less likely for all of the channels to share a common zero which would invalidate the system's identifiability. The results are shown in Fig. 3 with the same layout as that in the first experiment. The NRMSEs were obtained by averaging the results of $N = 200$ Monte Carlo runs. The CR method used 120 samples and the step sizes for the MCLMS and MCN algorithms were $\mu = 0.001$ and $\rho = 0.95$, respectively. In this case, the MCN method performed still better than the MCLMS algorithm.

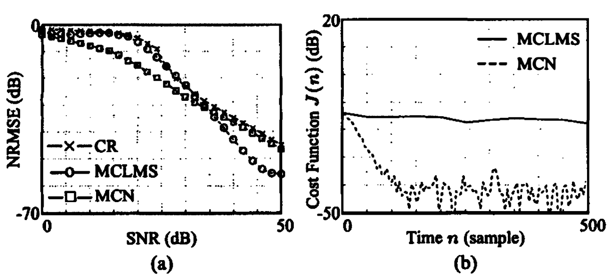

Figure 2: Comparison of performance among the CR, MCLMS, and MCN algorithms for identifying the ill-conditioned two-channel FIR system. (a) NRMSE vs. SNR and (b) trajectories of the cost function $J(n)$ for one typical run of the MCLMS and MCN algorithms at 50 dB SNR.

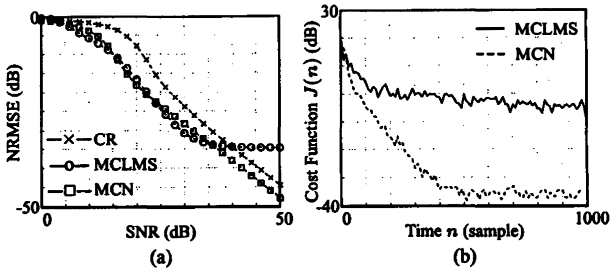

Figure 3: Comparison of performance among the CR, MCLMS, and MCN algorithms for identifying the three-channel FIR system. (a) NRMSE vs. SNR and (b) trajectories of the cost function $J(n)$ for one typical run of the MCLMS and MCN algorithms at 50 dB SNR.

4. CONCLUSIONS

The blind channel identification and estimation problem is studied. An error function based on the cross relations between different channels is constructed in a systematic way for a multichannel FIR system. The resulting cost function is concise and potentially facilitates the use of adaptive filtering techniques in the future development of efficient adaptive blind channel identification schemes. As an example, an LMS approach was proposed and its convergence in the mean to the desired channel impulse responses was theoretically shown. Furthermore, in order to accelerate convergence of the LMS adaptive algorithm, a Newton method was proposed. Simulation results justified our analysis and the proposed MCLMS and MCN algorithms performed well for an identifiable system.

5. 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] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[4] 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.

[5] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.

[6] Y. Huang, J. Benesty, and G. W. Elko, "Adaptive eigenvalue decomposition algorithm for real-time acoustic source localization system," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1999, vol. 2, pp. 937–940.

[7] S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1996.

[8] 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.