Correspondence

Data-Recursive Algorithms for Blind Channel Identification in Oversampled Communication Systems

Dennis L. Goeckel, Alfred O. Hero, and Wayne E. Stark

***Abstract*— Data-recursive algorithms are presented for performing blind channel identification in oversampled communication systems. Novel on-line solutions with complexities that are only linear in the oversampling rate are considered, and mean convergence conditions are provided. Numerical results are presented for a binary phase-shift keyed (BPSK) system.**

***Index Terms*— Adaptive signal processing, equalizers, identification, pseudonoise coded communication.**

I. Introduction

Much work has appeared recently on channel identification/equalization in communication systems when the channel output is oversampled [1]–[5]. The common base of this work is the exploitation of the cyclostationarity of the oversampled process, which allows for blind channel identification based on only second-order statistics. Because estimation of second-order statistics requires fewer samples than that of higher order statistics for a given level of estimation accuracy, we expect algorithms based on second-order statistics to exhibit faster convergence.

The algorithms of [1], [3], and [6] depend on the determination of the eigenvector corresponding to the minimum eigenvalue of a matrix of cross correlations of observations. This makes the complexity of these algorithms cubic in the product of the oversampling rate and the number of symbol periods spanned by the channel response. Thus, these algorithms can be prohibitively complex, particularly when the oversampling rate is large. For example, the oversampling rate can be large in many spread-spectrum systems; in particular, it has been observed independently by other researchers [8], [9] and the authors [10] that a chip-sampled spread spectrum system can be interpreted as an oversampled communications system, and thus, second-order methods can be used for blind channel identification. In such spread-spectrum systems, the oversampling rate is at least as large as the processing gain. In other applications, such as underwater communications [11], the number of symbol periods spanned by the channel response can be large. For real-time operation in each of the above situations, data-recursive algorithms of much lower complexity are desirable. Thus, in this work, on-line algorithms are presented that are only linear in *either* the oversampling rate or the number of symbol periods spanned by the channel response.

Manuscript received December 1, 1995; revised November 7, 1997. This work was supported in part by a National Science Foundation Graduate Fellowship, a Rackham Pre-Doctoral Fellowship, the National Science Foundation under Grants BCS-90243370 and NCR-9115869, and the Army Research Office under Contract DAAH04-95-I-0246. The associate editor coordinating the review of this paper and approving it for publication was Prof. Georgios B. Giannakis.

D. L. Goeckel is with the Electrical and Computer Engineering Department, University of Massachusetts, Amherst, MA 01003-5110 USA.

A. O. Hero and W. E. Stark are with the Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, MI 48109-2122 USA.

Publisher Item Identifier S 1053-587X(98)05231-3.

An outline of this correspondence is as follows. In Section II, the system and channel model are briefly defined, and the off-line subchannel response matching (SRM) algorithm is presented. Section III develops the proposed adaptive algorithms for second-order identification of oversampled systems. Finally, numerical results for the proposed algorithms are shown in Section IV.

II. Problem Formulation and Subchannel Response Matching

Let

$\{s_k\}_{k=-\infty}^{\infty}$	transmitted symbol stream;
T_s	transmitted symbol period;
$p(t)$	transmitter pulse shape.

The transmitted waveform is $s(t) = \Sigma_{l=-\infty}^{\infty} s_l p(t - lT_s)$. The received waveform is $r(t) = g(t) * s(t) + w(t)$, where $w(t)$ is a complex white Gaussian random process with two-sided power spectral density $N_0/2$, and $g(t)$ is the continuous time channel. If the received signal is integrated and dumped at period $T_c = (T_s/N)$, where N is an integer greater than one, the standard discrete-time oversampled communication system model [4], [5] is derived as $y_n(k) \triangleq \int_{kT_s+nT_c}^{kT_s+(n+1)T_c} r(t)\, dt = \Sigma_{l=-\infty}^{\infty} h_n(l) s_{k-l} + \eta_n(k)$, where $\eta_n(k)$ is a sequence of independent and identically distributed (IID) Gaussian random variables with variance $\sigma^2 = (N_0/2)T_c$, and $\{h_n(l)\}_{l=0}^{L-1}$ is the nth subchannel response, whose coefficients are given by $h_n(l) = \int_{-\infty}^{\infty} d\tau\, g(\tau) \int_0^{T_c} dt\, p(t + (lN + n)T_c - \tau)$.

The subchannel response matching algorithm [3] is derived by noting that $h_{n_1}(k) * y_{n_0}(k) = h_{n_0}(k) * y_{n_1}(k)$ in a noiseless system. This suggests that to identify the subchannels $h_{n_0}(k)$ and $h_{n_1}(k)$, an option is to consider subchannel estimates $\hat{h}_{n_0}(k)$ and $\hat{h}_{n_1}(k)$ that minimize the mean squared difference $\mathcal{E}_2(\hat{h}_{n_0}, \hat{h}_{n_1}) \triangleq E|\hat{h}_{n_1}(k) * y_{n_0}(k) - \hat{h}_{n_0}(k) * y_{n_1}(k)|^2$. Clearly, $\mathcal{E}_2 = 0$ is zero when the subchannels are identified correctly. Furthermore, if the subchannels have no common zeroes (which is the well-known identifiability criterion for second-order identification [2]), $\mathcal{E}_2 = 0$ implies $\hat{h}_{n_0}(k) = h_{n_0}(k), \hat{h}_{n_1}(k) = h_{n_1}(k)$ if the trivial solution $\hat{h}_{n_0}(k) = 0, \hat{h}_{n_1}(k) = 0, \forall k$ is excluded. For the general case of $N > 2$, define the vector (of length LN) of concatenated subchannels as $\boldsymbol{h} \triangleq [\mathrm{h}_0(0)h_0(1)\cdots h_0(L-1)h_1(0)h_1(1)\cdots h_1(L-1)\cdots h_{N-1}(0)h_{N-1}(1)\cdots h_{N-1}(L-1)]^T$ and the vector (of length LN) of concatenated subchannel observation vectors as $\boldsymbol{y}(k) \triangleq [\boldsymbol{y}_0^T(k), \boldsymbol{y}_1^T(k), \cdots, \boldsymbol{y}_{N-1}^T(k)]^T$, where $\boldsymbol{y}_n(k) \triangleq [y_n(k)y_n(k-1)\cdots y_n(k-L+1)]^T$. Let $\boldsymbol{x}[n]$ denote the nth element of the vector $\boldsymbol{x}$. The objective function is defined as the sum of all pairwise mean squared differences

$$\mathcal{E}(\hat{\boldsymbol{h}}) = \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} E\left[\left|\sum_{i=0}^{L-1} \hat{\boldsymbol{h}}[mL+i] y_n(k-i) - \sum_{j=0}^{L-1} \hat{\boldsymbol{h}}[nL+j] y_m(k-j)\right|^2\right] \tag{1}$$

which, when minimized over $\hat{\boldsymbol{h}}$, which is the estimate of $\boldsymbol{h}$, gives a solution $\boldsymbol{h}^{\mathrm{opt}}$. A constraint must be added to avoid the trivial solution $\boldsymbol{h}^{\mathrm{opt}} = \boldsymbol{0}$; the constraint employed here is $\|\hat{\boldsymbol{h}}\|^2 = 1$. Note that there is an implicit ambiguity in the problem; multiplying $\boldsymbol{h}$ by a complex constant does not affect the solution $\boldsymbol{h}^{\mathrm{opt}}$.

The objective (1) can be rewritten in an alternate form. Define the L by L matrices $R_{mn} \triangleq E[\boldsymbol{y}_m(k)\boldsymbol{y}_n^H(k)]$. Then, $\mathcal{E}(\hat{\boldsymbol{h}}) = \hat{\boldsymbol{h}}^T S \hat{\boldsymbol{h}}^*$, where $\hat{\boldsymbol{h}}^*$ is the complex conjugate (no transpose) of $\hat{\boldsymbol{h}}$, $S = [I_N \otimes \Sigma_{n=0}^{N-1} R_{nn}] - V$, I_N is an N by N identity matrix, $\otimes$ is the Kronecker product, and

$$V = \begin{bmatrix} R_{00} & R_{10} & \cdots & R_{(N-1)0} \\ R_{01} & R_{11} & & \vdots \\ \vdots & & \ddots & \\ R_{0(N-1)} & \cdots & & R_{(N-1)(N-1)} \end{bmatrix}.$$

The matrix S is Hermitian symmetric and positive semidefinite, and if the subchannels have no common zeroes, there will be a *unique* minimum eigenvalue of S, where the channel $\boldsymbol{h}$ is the conjugate of its associated eigenvector [2].

III. Adaptive Algorithms

Historically, algorithms for iteratively solving for the eigenvector corresponding to the minimum eigenvalue of a positive semidefinite matrix have fallen into three classes:

- stochastic gradient [12], [13];
- conjugate gradient [14];
- Newton's method [15].

Attention here is restricted to the stochastic gradient and conjugate gradient algorithms. Although finding the eigenvector corresponding to the minimum eigenvalue of a matrix corresponds closely to Pisarenko's harmonic retrieval method [12] and direction-of-arrival estimation, S is not expressible as the expectation of an outer product of the observed vector unless $N = 2$; hence, the simple data updates of [12] do not apply. The algorithms presented here update $\hat{\boldsymbol{h}}_{k-1}$, which is the estimate of $\boldsymbol{h}^{\text{opt}}$ at iteration $k-1$ based on an additional symbol period of data $y_0(k), y_1(k), \cdots, y_{N-1}(k)$.

The stochastic gradient update equation for $\hat{\boldsymbol{h}}_k$ is given by $\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_{k-1} - \mu \nabla e_k(\hat{\boldsymbol{h}}_{k-1})$, where μ is the step size to be chosen, ∇ is the gradient operator, and the error function $e_k(\hat{\boldsymbol{h}}_{k-1})$ employed here is the empirical estimate of the Rayleigh quotient of S^* (i.e., the Rayleigh quotient with the positive semidefinite empirical matrix S_k, which is identical to S derived above except with the expectations removed from (1), replacing the ensemble average matrix S). Substituting in appropriate values and absorbing a factor of 2 into μ, the stochastic gradient update is obtained as

$$\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_{k-1} - \mu \frac{(\|\hat{\boldsymbol{h}}_{k-1}\|^2 S_k^* \hat{\boldsymbol{h}}_{k-1} - (\hat{\boldsymbol{h}}_{k-1}^T S_k \hat{\boldsymbol{h}}_{k-1}^*)\hat{\boldsymbol{h}}_{k-1})}{\|\hat{\boldsymbol{h}}_{k-1}\|^4}. \quad (2)$$

Performing a normalization based on the convergence analysis (see Appendix A) yields the normalized stochastic gradient update

$$\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_{k-1} - \mu \frac{(\|\hat{\boldsymbol{h}}_{k-1}\|^2 S_k^* \hat{\boldsymbol{h}}_{k-1} - (\hat{\boldsymbol{h}}_{k-1}^T S_k \hat{\boldsymbol{h}}_{k-1}^*)\hat{\boldsymbol{h}}_{k-1})}{\|\hat{\boldsymbol{h}}_{k-1}\|^2}. \quad (3)$$

A conjugate gradient algorithm [14] is given by the update $\tilde{\boldsymbol{h}}_{k+1} = \hat{\boldsymbol{h}}_k + \mu_k \boldsymbol{p}_k$ followed by the normalization $\hat{\boldsymbol{h}}_{k+1} = (\tilde{\boldsymbol{h}}_{k+1}/\|\tilde{\boldsymbol{h}}_{k+1}\|)$, where μ_k is chosen to minimize the Rayleigh quotient in direction $\boldsymbol{p}_k$ and is given by $\mu_k = (-B + \sqrt{B^2 - 4CD}/2D)$ with $D \triangleq P_b(k)P_c(k) - P_a(k)P_d(k), B \triangleq P_b(k) - e_{k+1}(\hat{\boldsymbol{h}}_k)P_d(k), C \triangleq P_a(k) - e_{k+1}(\hat{\boldsymbol{h}}_k)P_c(k)$

$$P_a(k) \triangleq \frac{\boldsymbol{p}_k^H S_{k+1}^* \hat{\boldsymbol{h}}_k}{\|\hat{\boldsymbol{h}}_k\|^2} \quad P_b(k) \triangleq \frac{\boldsymbol{p}_k^H S_{k+1}^* \boldsymbol{p}_k}{\|\hat{\boldsymbol{h}}_k\|^2}$$

$$P_c(k) \triangleq \frac{\boldsymbol{p}_k^H \hat{\boldsymbol{h}}_k}{\|\hat{\boldsymbol{h}}_k\|^2} \quad P_d(k) \triangleq \frac{\boldsymbol{p}_k^H \boldsymbol{p}_k}{\|\hat{\boldsymbol{h}}_k\|^2}.$$

TABLE I
Zeroes of the "Good" and "Bad" Channels

Subchannel	*Good*	*Bad*
0	0.353 ± 0.353 j	0.636 ±0.636j
	0.259 ± 0.150 j	0.692 ± 0.400j
1	-0.104 ± 0.590 j	-0.121 ± 0.689j
	1.691 ± 0.615 j	1.122 ± 0.410j
2	0.069 ± 0.393 j	0.148 ± 0.837j
	-1.221 ± 1.455 j	1.034 ± 0.376j
3	0.752 ± 2.067 j	0.393 ± 1.081j
	-0.800 ± 1.385 j	-0.525 ± 0.909j

The new search direction $\boldsymbol{p}_k$ is selected by $\boldsymbol{p}_k = -\nabla e_{k+1}(\hat{\boldsymbol{h}}_k) + b_k \boldsymbol{p}_{k-1}$. In Chen *et al.* [14], b_k is chosen such that $\boldsymbol{p}_k^H S_k^* \boldsymbol{p}_{k-1} = 0$, which yields

$$b_k = \frac{\nabla^H e_{k+1}(\hat{\boldsymbol{h}}_k) S_k^* \boldsymbol{p}_{k-1}}{\boldsymbol{p}_{k-1}^H S_k^* \boldsymbol{p}_{k-1}}.$$

The stochastic gradient and conjugate gradient algorithms require $O(L^2N^2)$ operations as written due to matrix-vector multiplications, where S_k is the matrix in each case. To see how this can be simplified, consider the multiplication $S_k \hat{\boldsymbol{h}}_{k-1}^*$ required by the stochastic gradient algorithm. Let $\mathcal{E}_k(\hat{\boldsymbol{h}}_{k-1}) = \hat{\boldsymbol{h}}_{k-1}^T S_k \hat{\boldsymbol{h}}_{k-1}^*$ be the empirical estimate of $\mathcal{E}(\hat{\boldsymbol{h}}_{k-1})$, and note that $\nabla \mathcal{E}_k(\hat{\boldsymbol{h}}_{k-1}) = 2S_k^* \hat{\boldsymbol{h}}_{k-1}$. Element $mL + l$ of $S_k \hat{\boldsymbol{h}}_{k-1}^*$ can be obtained as

$$\begin{aligned} &\frac{\partial \mathcal{E}_k(\hat{\boldsymbol{h}}_{k-1})}{\partial \hat{\boldsymbol{h}}_{k-1}[mL+l]} \\ &\quad = 2 \sum_{n=0, n\neq m}^{N-1} \left(\sum_{i=0}^{L-1} \hat{\boldsymbol{h}}_{k-1}[mL+i] y_n(k-i) y_n^*(k-l) \right. \\ &\qquad \left. - \sum_{j=0}^{L-1} \hat{\boldsymbol{h}}_{k-1}[nL+j] y_m(k-j) y_n^*(k-l) \right) \\ &\quad = 2 \left(\sum_{i=0}^{L-1} \hat{\boldsymbol{h}}_{k-1}[mL+i] \sum_{n=0}^{N-1} y_n(k-i) y_n^*(k-l) \right. \\ &\qquad \left. - \sum_{j=0}^{L-1} y_m(k-j) \sum_{n=0}^{N-1} \hat{\boldsymbol{h}}_{k-1}[nL+j] y_n^*(k-l) \right). \end{aligned} \quad (4)$$

The computation of the right side of (4) proceeds as follows: The sums over n must be done for each l and i (or j), thus leading to $O(L^2N)$ operations; given these sums, $S_k \hat{\boldsymbol{h}}_{k-1}^* = \nabla \mathcal{E}_k(\hat{\boldsymbol{h}}_{k-1})$ can be found in $O(LN)$ inner products of vectors of length L. Thus, the update (3) can be done in $O(L^2N)$ flops. Since $N \gg L$ for highly oversampled systems, this amounts to a large savings in computational effort for the algorithm. By reversing the summations in (4), the update can be performed in $O(LN^2)$ flops.

IV. Numerical Results

The normalized mean squared error at iteration k is defined by $\text{NMSE}_k = \|(\hat{\boldsymbol{h}}_k/\|\hat{\boldsymbol{h}}_k\|) - (\boldsymbol{h}/\|\boldsymbol{h}\|)\|^2$. Binary phase-shift keyed (BPSK) modulation over two channels with zeroes as in Table I and in Fig. 1 is considered. For each channel, $N = 4$ and $L = 5$, which yields $\mu = 0.14$ from the convergence analysis in Appendix A. The "standard" stochastic gradient algorithm (2), "normalized" stochastic gradient algorithm (3), conjugate gradient algorithm, and the off-line SRM algorithm from [3] were run on channels "Good" and "Bad" at signal-to-noise ratios of 20 and 10 dB. For the stochastic gradient algorithms, the best gain factors from among $\mu = 0.06, \mu = 0.14,$ and $\mu = 0.30$ (the gain factor the convergence analysis

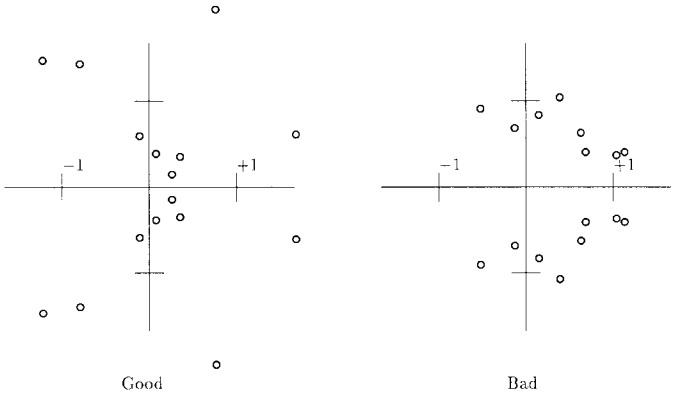

Fig. 1. Zeroes of the "Good" and "Bad" channels in the z plane.

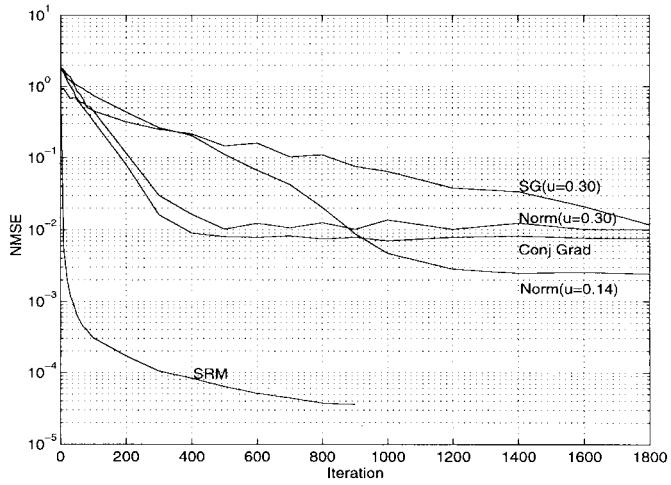

Fig. 2. Comparison of the SRM-based algorithms for the "Good" channel, SNR $= 20$ dB. SG: Standard stochastic gradient algorithm. Norm: Normalized stochastic gradient algorithm. SRM: Off-line SRM algorithm of [3], where S_k is obtained at time k by averaging all of the data through time k to estimate ensemble averages. Conj Grad: Conjugate gradient algorithm. The results are averaged over 200 sample trials.

suggests and one on either side) were selected. The results are shown in Figs. 2–5.

V. Summary and Conclusions

The need for channel identification in highly oversampled communication systems motivates the search for efficient blind channel identification algorithms when the oversampling rate is high. Stochastic gradient and conjugate gradient algorithms have been presented that are low complexity (only linear in the oversampling rate). Numerical results suggest that there is quite a bit of performance loss between the off-line SRM algorithm and the data-adaptive implementations of it. Furthermore, it appears that the gain of the conjugate gradient algorithms over the stochastic gradient algorithms is not as prominent as was shown in Pisarenko's harmonic retrieval [16]; however, the conjugate gradient algorithm showed markedly faster initial convergence in general.

Appendix
Stochastic Gradient Algorithm Convergence

Since all of the eigenvectors except those of the minimum and maximum eigenvalues are saddle points of the error surface, the stochastic gradient algorithm will leave these eigenvectors [13]; thus,

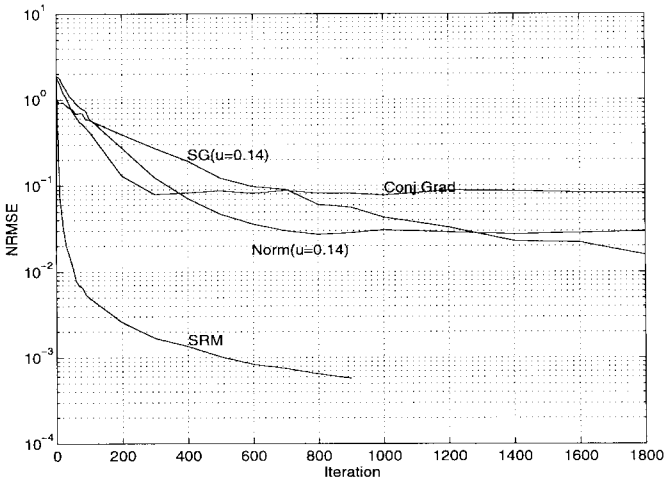

Fig. 3. Comparison of the SRM-based algorithms for the "Good" channel, SNR $= 10$ dB. SG: Standard stochastic gradient algorithm. Norm: Normalized stochastic gradient algorithm. SRM: Off-line SRM algorithm of [3], where S_k is obtained at time k by averaging all of the data through time k to estimate ensemble averages. Conj Grad: Conjugate gradient algorithm. The results are averaged over 200 sample trials.

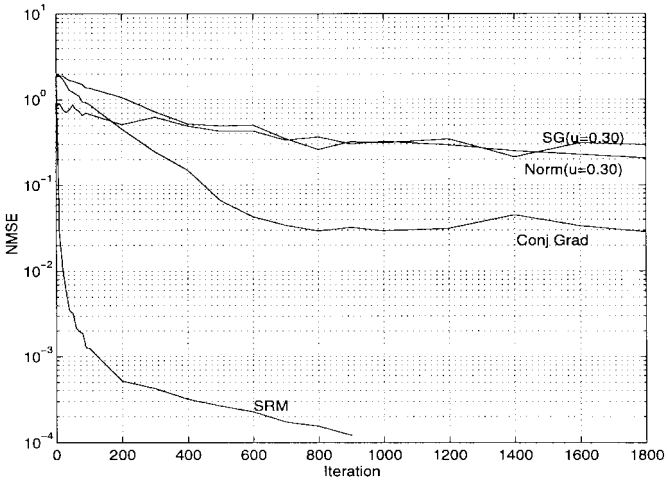

Fig. 4. Comparison of the SRM-based algorithms for the "Bad" channel, SNR $= 20$ dB. SG: Standard stochastic gradient algorithm. Norm: Normalized stochastic gradient algorithm. SRM: Off-line SRM algorithm of [3], where S_k is obtained at time k by averaging all of the data through time k to estimate ensemble averages. Conj Grad: Conjugate gradient algorithm. The results are averaged over 200 sample trials.

μ is chosen such that the algorithm will converge to the correct solution when it is near the minimizing eigenvector. A first-order convergence analysis is performed, which is equivalent to making the assumption that S can be estimated without error. The error after iteration k of the algorithm is given by the Rayleigh quotient $e(\hat{\boldsymbol{h}}_k) = (\hat{\boldsymbol{h}}_k^T S \hat{\boldsymbol{h}}_k^*/\|\hat{\boldsymbol{h}}_k\|^2) = (\hat{\boldsymbol{h}}_k^T S_Z \hat{\boldsymbol{h}}_k^*/\|\hat{\boldsymbol{h}}_k\|^2) + \lambda_0$, where S_Z is the matrix S for a noiseless system $(w(t) = 0)$, and $\lambda_0 = (N-1)\sigma^2$ is the unique smallest eigenvalue of S. Dropping the irreducible error λ_0, substituting in the appropriate quantities, and preserving terms that are linear or constant in $e(\hat{\boldsymbol{h}}_k)$ (i.e., assuming $e(\hat{\boldsymbol{h}}_k) \ll 1$ near the solution), we can show [10] that e_k^i, which is the squared error in the direction of the ith eigenvector of S_Z at step k, is given by

$$e_k^i \le e_{k-1}^i \left(1 - \frac{\mu\lambda_i}{\|\hat{\boldsymbol{h}}_{k-1}\|^2}\right)^2 \tag{5}$$

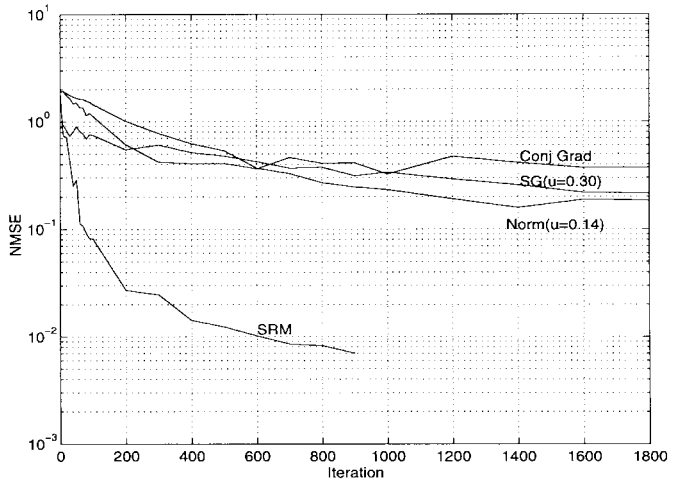

Fig. 5. Comparison of the SRM-based algorithms for the "Bad" channel, SNR $= 10$ dB. SG: Standard stochastic gradient algorithm. Norm: Normalized stochastic gradient algorithm. SRM: Off-line SRM algorithm of [3], where S_k is obtained at time k by averaging all of the data through time k to estimate ensemble averages. Conj Grad: Conjugate gradient algorithm. The results are averaged over 200 sample trials.

where λ_i is the eigenvalue corresponding to the ith eigenvector of S_Z. If the initial guess is chosen on the unit circle (i.e., $\|\hat{\boldsymbol{h}}_0\| = 1$), the fact that the magnitude of $\hat{\boldsymbol{h}}_k$ is nondecreasing and (5) can be used to show [10] that convergence occurs in all of the modes if $\mu < (2/\lambda_{\max})$, where $\lambda_{\max}$ is the largest eigenvalue of S_Z. Since $\lambda_{\max}$ is difficult to obtain, the conservative rule $\mu \leq (2/\mathrm{tr}(S_Z))$ is used instead, where $\mathrm{tr}(S_Z) = L(N-1)\ \Sigma_{i=0}^{N-1}\ E|y_i(k)|^2 - LN(N-1)\sigma^2$, which can be readily estimated if the signal-to-noise ratio is known.

Equation (5) suggests defining a normalized algorithm to eliminate the $\|\hat{\boldsymbol{h}}_{k-1}\|^2$ dependence in the convergence condition. The normalized algorithm is obtained by using a variable gain factor $\mu\|\hat{\boldsymbol{h}}_{k-1}\|^2$ at the kth step. Employing this variable gain factor yields

$$\hat{\boldsymbol{h}}_k = \hat{\boldsymbol{h}}_{k-1} - \mu \frac{(\|\hat{\boldsymbol{h}}_{k-1}\|^2 S_k^* \hat{\boldsymbol{h}}_{k-1} - (\hat{\boldsymbol{h}}_{k-1}^T S_k \hat{\boldsymbol{h}}_{k-1}^*)\hat{\boldsymbol{h}}_{k-1})}{\|\hat{\boldsymbol{h}}_{k-1}\|^2} \quad (6)$$

which is (3) of the main text. The convergence analysis of this normalized stochastic gradient update follows similarly and implies convergence if $\mu < (2/\lambda_{\max})$.

References

[1] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind equalization," in *Conf. Rec. Twenty-Seventh Asilomar Conf. Signals, Syst., Comput.*, 1993, pp. 751–755.

[2] L. Baccalá and S. Roy, "Time-domain blind channel identification algorithms," in *Proc. 1994 Conf. Inform. Sci. Syst.*, vol. 2, pp. 863–867.

[3] S. Schell and D. Smith, "Improved performance of blind equalization using prior knowledge of transmitter filter," in *Conf. Rec. 1994 IEEE Mil. Commun. Conf.*, vol. 1, pp. 128–132.

[4] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[5] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[6] L. Baccalá and S. Roy, "A new blind time-domain channel identification method based on cyclostationarity," *IEEE Signal Processing Lett.*, vol. 1, pp. 89–91, June 1994.

[7] H. Zeng and L. Tong, "Blind channel estimation using the second-order statistics: Algorithms," *IEEE Trans. Signal Processing*, vol. 45, pp. 1919–1930, Aug. 1997.

[8] M. Tsatsanis and G. Giannakis, "Multirate filter banks for code-division multiple access systems," *Proc. 1995 Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, pp. 1484–1487.

[9] ——, "Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework," *IEEE Trans. Signal Processing*, vol. 44, pp. 3044–3055, Dec. 1996.

[10] D. Goeckel, A. Hero, III, and W. Stark, "Blind channel identification for direct-sequence spread-spectrum systems," in *Conf. Rec. 1995 IEEE Mil. Commun. Conf.*, pp. 368–372.

[11] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol. 21, pp. 125–136, Apr. 1996.

[12] P. Thompson, "An adaptive spectral analysis technique for unbiased frequency estimation in the presence of white noise," in *Conf. Rec. Thirteenth Asilomar Conf. Circuits, Syst., Comput.*, 1979, pp. 529–533.

[13] M. Larimore and R. Calvert, "Convergence studies of Thompson's unbiased adaptive spectral estimator," in *Conf. Rec. Fourteenth Asilomar Conf. Circuits, Syst., Comput.*, 1980, pp. 258–262.

[14] H. Chen, T. Sarkar, S. Dianat, and J. Brulé, "Adaptive spectral estimation by the conjugate gradient method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 272–284, Apr. 1986.

[15] G. Mathew, S. Dasgupta, and V. Reddy, "Improved Newton-type algorithm for adaptive implementation of Pisarenko's harmonic retrieval method and its convergence analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 434–437, Feb. 1994.

[16] X. Yang, T. Sarkar, and E. Arvas, "A survey of conjugate gradient algorithms for solution of extreme eigen-problems of a symmetric matrix," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1550–1555, Oct. 1989.

Cascaded Power Symmetric IIR Filter Banks and Continuity Constrained Adaptive Algorithms for Acoustic Echo Cancellation in Subbands

Oğuz Tanrıkulu and Anthony G. Constantinides

***Abstract*— The problem of aliasing in subband acoustic echo cancellation (AEC) is addressed. Filter banks with implicit notch filtering are derived from cascaded power symmetric–infinite impulse response (CPS-IIR) filters. It is shown that adaptive filters used with these filter banks must be coupled via continuity constraints to reduce the aliasing in the residual echo. A continuity constrained NLMS algorithm is therefore proposed and evaluated.**

***Index Terms*—Adaptive filters, echo suppression, IIR digital filters.**

I. Introduction

Teleconferencing systems and hands-free mobile terminals use acoustic echo cancellation (AEC) for high-quality full-duplex speech communication [1]. AEC in subbands is an effective way of reducing the computational complexity [2], [3]. However, the performance is

Manuscript received January 27, 1997; revised January 27, 1998. This work was supported in part by the U.K., EPSRC Research Project GR/K48617 "Multi-rate Echo Cancellation." The associate editor coordinating the review of this paper and approving it for publication was Prof. P. P. Vaidyanathan.

O. Tanrıkulu is with DSP Software Engineering, Inc., Bedford, MA 01730 USA.

A. G. Constantinides is with the Signal Processing and Digital Section, Department of Electrical and Electronic Engineering, Imperial College of Science, Technology, and Medicine, London, U.K.

Publisher Item Identifier S 1053-587X(98)05232-5.