SELECTIVE-PARTIAL-UPDATE PROPORTIONATE NORMALIZED LEAST-MEAN-SQUARES ALGORITHM FOR NETWORK ECHO CANCELLATION

Oğuz Tanrıkulu

Tellabs Media Processing Technologies
175 Middlesex Turnpike
Bedford, MA 01730, USA
E-mail: Oguz.Tanrikulu@tellabs.com

*Kutluyıl Doğançay**

School of Electrical and Information Eng.
University of South Australia
Mawson Lakes, SA 5095, Australia
E-mail: Kutluyil.Dogancay@unisa.edu.au

ABSTRACT

A reduced-complexity, fast converging adaptive filtering algorithm is developed for network echo cancellation applications in sparse echo paths. The new algorithm—selective-partial-update proportionate NLMS—is based on the proportionate NLMS (PNLMS) algorithm and selective partial updating of the adaptive filter coefficients. PNLMS exploits the sparseness of the channel to speed up the initial convergence of the NLMS algorithm by weighting the regressor data proportionately with the estimated magnitude of the channel impulse response. Selective partial updating is essentially a data selection method to reduce the computational complexity. The performance of selective-partial-update PNLMS is compared with PNLMS for echo paths specified in ITU-T Recommendation G.168, and selective-partial-update PNLMS is observed to maintain the convergence speed of PNLMS at a reduced computational complexity.

1. INTRODUCTION

Network echo cancellation problems are characterized by long echo path impulse responses of the order of 128 ms, of which only a small portion (4–22 ms) is actually nonzero. Adaptive echo cancellers are designed to synthesize the full length of the echo path because (i) the location of the nonzero portion of the echo path (the flat delay) is unknown and (ii) the flat delay varies from call to call. The flat delay is due to the long distance between the echo canceller and the hybrid/local-loop circuit.

A recently proposed modification of NLMS—the proportionate normalized least-mean-squares (PNLMS) algorithm [1, 2]—exploits the sparseness of the echo path impulse response to speed up the initial convergence of the conventional NLMS algorithm. PNLMS essentially adapts the nonzero portion of the echo path by weighting the update terms of the adaptive filter with the magnitude of the estimated echo path impulse response. This effectively results in a shorter adaptive filter that converges faster than the full-length adaptive filter. The key to PNLMS is the introduction of weighting for adaptive filter coefficients. A particular choice for the weighting gives rise to PNLMS.

There are some disadvantages to PNLMS, namely, an increase in computational complexity by 50% and the slow convergence of the adaptive filter coefficients after the fast initial convergence. The latter is due to the slow convergence of small coefficients after the convergence of the large coefficients over the "active" portion of the echo path impulse response. It is possible to reduce the increased computational complexity by way of selective partial updating [3] of the adaptive filter coefficients. PNLMS is a coefficient-selective algorithm in that it weights the coefficients to be updated in direct proportion to the magnitude of the coefficient estimates. Selective partial updating introduces a data-selective flavour to PNLMS by selecting the regressor vector entries to be used in the update term. This in turn reduces the number of multiplications in the update term. The slow convergence of PNLMS after the initial fast convergence can be alleviated by providing a transition from PNLMS to NLMS during the adaptation process. This can be achieved by changing one of the PNLMS parameters as new data samples come in. Variants of PNLMS have been proposed to improve the convergence speed, to provide robustness in the face of undetected double-talk, and to reduce the computational complexity [2, 4]. The new algorithm presented in this paper provides further reduction in the computational complexity while maintaining the fast convergence of PNLMS.

The paper is organized as follows. Section 2 derives the PNLMS algorithm with selective partial updates by resorting to a constrained optimization problem incorporating a weighted cost criterion. In Section 3, possible adjustments to the PNLMS parameters are considered to improve

*This work was partially supported by a DITEE Research Grant of the University of South Australia.

0-7803-7402-9/02/$17.00 ©2002 IEEE

Authorized licensed use limited to: University of South Australia. Downloaded on June 9, 2009 at 21:57 from IEEE Xplore. Restrictions apply.

the convergence performance. Section 4 includes simulation examples for comparison of PNLMS and its selective-partial-update variant on sparse network echo paths.

2. PNLMS WITH SELECTIVE PARTIAL UPDATES

This section derives the PNLMS algorithm with selective partial updates from a weighted cost function. The algorithm development is cast into the framework of a constrained optimization problem, which is different to the approach taken in [1].

To start with, let us show how PNLMS can be derived from the solution of a constrained minimization problem as in the case of NLMS [5]. The output $y(k)$ of a network echo canceller at time k is given by

$$y(k) = \boldsymbol{w}^T(k)\boldsymbol{x}(k)$$

where $\boldsymbol{w}(k) = [w_0(k), w_1(k), \ldots, w_{N-1}(k)]^T$ is the $N \times 1$ filter coefficient vector and $\boldsymbol{x}(k) = [x(k), x(k-1), \ldots, x(k-N+1)]^T$ is the $N \times 1$ regressor vector of the far-end signal that excites the echo path. Introducing a weighting in the constrained optimization problem that leads to the NLMS algorithm yields

$$\min_{\boldsymbol{w}(k+1)} (\boldsymbol{w}(k+1) - \boldsymbol{w}(k))^T \boldsymbol{G}^{-1}(k)(\boldsymbol{w}(k+1) - \boldsymbol{w}(k)) \quad \text{subject to} \quad \boldsymbol{w}^T(k+1)\boldsymbol{x}(k) = d(k) \tag{1}$$

where $\boldsymbol{G}(k) = \text{diag}\{g_0(k), g_1(k), \ldots, g_{N-1}(k)\}$ is the diagonal coefficient weighting matrix and $d(k)$ is the desired filter response which is the near-end signal (i.e., echo plus the near-end noise). The resulting PNLMS algorithm is given by the recursion

$$\boldsymbol{w}(k+1) = \boldsymbol{w}(k) + \mu \frac{\boldsymbol{G}(k)\boldsymbol{x}(k)e(k)}{\boldsymbol{x}^T(k)\boldsymbol{G}(k)\boldsymbol{x}(k)} \tag{2}$$

where μ is a stepsize introduced to control the convergence speed and $e(k)$ is the error(residual echo) signal defined by

$$e(k) = d(k) - y(k).$$

Up to this point, we have not specified how the diagonal weighting matrix $\boldsymbol{G}(k)$ should be formed. Indeed, this information is not essential for the purposes of algorithm derivation. While in general $\boldsymbol{G}(k)$ can be chosen as any reasonable diagonal matrix with nonnegative entries, for PNLMS the diagonal entries of $\boldsymbol{G}(k)$ are chosen to approximate the magnitude of the estimated echo path impulse response. This results in an uneven distribution of the available energy in the regressor vector over the adaptive filter coefficients. If we set $\boldsymbol{G}(k) = \boldsymbol{I}$, PNLMS reduces to NLMS.

The PNLMS algorithm in (2) is somewhat different to the PNLMS recursion obtained in [1] where the normalization factor in the denominator of the update term is replaced by $\boldsymbol{x}^T(k)\boldsymbol{x}(k)$. In practice, to avoid division by zero (or a very small number), the normalization factor should be modified to $\boldsymbol{x}^T(k)\boldsymbol{G}(k)\boldsymbol{x}(k) + \delta$ where δ is a small positive constant.

For PNLMS, the entries of the diagonal weighting matrix $\boldsymbol{G}(k)$ are calculated at every iteration according to [2]

$$l_\infty(k) = \max\{\delta_p, |w_0(k)|, \ldots, |w_{N-1}(k)|\} \tag{3a}$$

$$\gamma_i(k) = \max\{\rho l_\infty(k), |w_i(k)|\}, \quad 0 \le i \le N-1 \tag{3b}$$

$$g_i(k) = \frac{\gamma_i(k)}{\sum_{j=0}^{N-1} \gamma_j(k)}, \quad 0 \le i \le N-1 \tag{3c}$$

where δ_p and ρ are the PNLMS parameters that effect small-signal regularization. The parameter δ_p prevents the algorithm from misbehaving when the adaptive filter coefficients $\boldsymbol{w}(k)$ are very small as at initialization, and ρ prevents individual filter coefficients from freezing (i.e., being never adapted again) when their magnitude is much smaller than the largest coefficient l_∞ [2]. Typical values for the PNLMS parameters are $\delta_p = 0.01$ and $\rho = 5/N$. If $\rho \ge 1$, $g_i(k) = 1/N$, i.e., $\boldsymbol{G}(k) = \frac{1}{N}\boldsymbol{I}$, and PNLMS behaves like NLMS.

Our objective is to reduce computational complexity of the PNLMS algorithm by updating a subset of the filter coefficients at every iteration. Partition the regressor vector $\boldsymbol{x}(k)$ and the coefficient vector $\boldsymbol{w}(k)$ into M blocks of length $L = N/M$ where L is an integer:

$$\boldsymbol{x}(k) = \begin{bmatrix} \boldsymbol{x}_1^T(k) & \boldsymbol{x}_2^T(k) & \cdots & \boldsymbol{x}_M^T(k) \end{bmatrix}^T$$

$$\boldsymbol{w}(k) = \begin{bmatrix} \boldsymbol{w}_1^T(k) & \boldsymbol{w}_2^T(k) & \cdots & \boldsymbol{w}_M^T(k) \end{bmatrix}^T.$$

Let $\boldsymbol{G}_i(k)$ denote the corresponding $L \times L$ block of $\boldsymbol{G}(k)$. Here the coefficient vector blocks $\boldsymbol{w}_1(k), \ldots, \boldsymbol{w}_M(k)$ represent candidate subsets of $\boldsymbol{w}(k)$ that can be updated at time instant k. For a single-block update, the constrained weighted minimization problem can be written as

$$\min_{1 \le i \le M} \min_{\boldsymbol{w}_i(k+1)} \|\boldsymbol{G}_i^{-1/2}(k)(\boldsymbol{w}_i(k+1) - \boldsymbol{w}_i(k))\|_2^2 \quad \text{subject to} \quad \boldsymbol{w}^T(k+1)\boldsymbol{x}(k) = d(k). \tag{4}$$

The solution of this optimization problem will find the block for which the coefficient update is minimal in the squared Euclidean norm sense while satisfying the constraint that $\boldsymbol{w}^T(k+1)\boldsymbol{x}(k)$ should be identical to the desired response $d(k)$.

We will first consider the minimization problem for a given block. If i is given and fixed, then (4) reduces to

$$\min_{\boldsymbol{w}_i(k+1)} \|\boldsymbol{G}_i^{-1/2}(k)(\boldsymbol{w}_i(k+1) - \boldsymbol{w}_i(k))\|_2^2 \quad \text{subject to} \quad \boldsymbol{w}^T(k+1)\boldsymbol{x}(k) = d(k) \tag{5}$$

Authorized licensed use limited to: University of South Australia. Downloaded on June 9, 2009 at 21:57 from IEEE Xplore. Restrictions apply.

which can be solved in a similar way to NLMS by using the method of Lagrange multipliers [5]. The cost function to be minimized is

$$J_i(k) = \|\boldsymbol{G}_i^{-1/2}(k)(\boldsymbol{w}_i(k+1) - \boldsymbol{w}_i(k))\|_2^2 + \lambda\big(d(k) - \boldsymbol{w}^T(k+1)\boldsymbol{x}(k)\big)$$

where λ is a Lagrange multiplier. Setting $\partial J_i(k)/\partial \boldsymbol{w}_i(k+1) = \boldsymbol{0}$ and $\partial J_i(k)/\partial\lambda = 0$, we get

$$\boldsymbol{w}_i(k+1) - \boldsymbol{w}_i(k) - \frac{\lambda}{2}\boldsymbol{G}_i(k)\boldsymbol{x}_i(k) = \boldsymbol{0} \quad (6a)$$

$$d(k) - \big(\boldsymbol{w}^T(k+1)\boldsymbol{x}(k)\big) = 0. \quad (6b)$$

Substituting (6a) into (6b), we get

$$\frac{\lambda}{2} = \frac{e(k)}{\boldsymbol{x}_i^T(k)\boldsymbol{G}_i(k)\boldsymbol{x}_i(k)}. \quad (7)$$

After substituting (7) into (6a) and introducing a small positive stepsize μ, we obtain the following recursion

$$\boldsymbol{w}_i(k+1) = \boldsymbol{w}_i(k) + \mu\frac{\boldsymbol{G}_i(k)\boldsymbol{x}_i(k)e(k)}{\boldsymbol{x}_i^T(k)\boldsymbol{G}_i(k)\boldsymbol{x}_i(k)}. \quad (8)$$

According to (4), the selection of the block to be updated should be made by determining the block with the smallest weighted squared-Euclidean-norm update. Thus, using (8), the block to be updated at every k is given by

$$i = \arg\min_{1\le j\le M} \|\boldsymbol{G}_j^{-1/2}(k)(\boldsymbol{w}_j(k+1) - \boldsymbol{w}_j(k))\|_2^2 \quad (9a)$$

$$= \arg\min_{1\le j\le M} \frac{e^2(k)}{\boldsymbol{x}_j^T(k)\boldsymbol{G}_j(k)\boldsymbol{x}_j(k)} \quad (9b)$$

$$= \arg\max_{1\le j\le M} \boldsymbol{x}_j^T(k)\boldsymbol{G}_j(k)\boldsymbol{x}_j(k). \quad (9c)$$

Combining (8) and (9c) yields a selective-single-block-update PNLMS algorithm.

Let us suppose that we wish to update B blocks out of M at every iteration. Let $\mathcal{I}_B = \{i_1, i_2, \ldots, i_B\}$ denote a B-subset (subset with B members) of the set $\{1, 2, \ldots, M\}$. Following the same line of development as above for updating B blocks results in the *selective-partial-update PNLMS (SPU-PNLMS) algorithm* given by

$$\boldsymbol{w}_{\mathcal{I}_B}(k+1) = \boldsymbol{w}_{\mathcal{I}_B}(k) + \mu\frac{\boldsymbol{G}_{\mathcal{I}_B}(k)\boldsymbol{x}_{\mathcal{I}_B}(k)e(k)}{\boldsymbol{x}_{\mathcal{I}_B}^T(k)\boldsymbol{G}_{\mathcal{I}_B}(k)\boldsymbol{x}_{\mathcal{I}_B}(k)},$$

$$\mathcal{I}_B = \{i : \boldsymbol{x}_i^T(k)\boldsymbol{G}_i(k)\boldsymbol{x}_i(k) \text{ is one of the } B \text{ largest among } \boldsymbol{x}_1^T(k)\boldsymbol{G}_1(k)\boldsymbol{x}_1(k), \ldots, \boldsymbol{x}_M^T(k)\boldsymbol{G}_M(k)\boldsymbol{x}_M(k)\} \quad (10)$$

where $\boldsymbol{x}_{\mathcal{I}_B}(k) = [\boldsymbol{x}_{i_1}^T(k), \ldots, \boldsymbol{x}_{i_B}^T(k)]^T$ is the augmented vector containing the $\boldsymbol{x}(k)$ blocks with indices in $\mathcal{I}_B$. Other augmented matrices and vectors are defined similarly.

The overhead introduced by the selection of coefficients using selective partial updating can be alleviated by employing fast ranking algorithms and strobe-down methods [1].

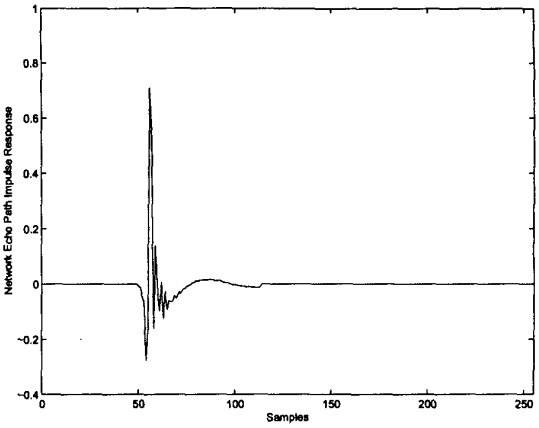

Fig. 1. A network echo path impulse response specified in ITU-T Recommendation G.168.

3. PARAMETER ADJUSTMENTS

The parameter ρ appears to control the location of the "break point" in the algorithm convergence where the rapid initial convergence comes to an end. It also determines the extent to which the proportionate update characteristic of PNLMS is effective. As ρ decreases, the initial convergence rate becomes faster [1]). However decreasing ρ beyond $1/N$ does not increase the initial convergence speed very much. In fact, too small ρ slows down the convergence speed after the break point.

The convergence of PNLMS is faster than NLMS before the break point and slower than NLMS after the break point. Ideally, we wish to have a convergence speed that is at least as fast as that of NLMS after the break point. These observations give us a hint as to how the PNLMS parameters should be adjusted to attain an ideal convergence behaviour.

We propose to redefine ρ so that a transition from PNLMS to NLMS occurs around the break point. The most general approach to doing this is to replace ρ with a time-varying parameter $\rho(k)$, whose definition will be dependent on the break point β. An example $\rho(k)$ is

$$\rho(k) = \begin{cases} 5/N, & k \le \beta \\ 1, & k > \beta. \end{cases} \quad (11)$$

The location of β can be determined from the echo return loss enhancement (ERLE) measurements and the link between ERLE and $\rho(k)$ [1].

4. SIMULATIONS

Figs 1 and 2 show the application of PNLMS to a sparse network echo path specified in [6] and compare the conver-

Authorized licensed use limited to: University of South Australia. Downloaded on June 9, 2009 at 21:57 from IEEE Xplore. Restrictions apply.

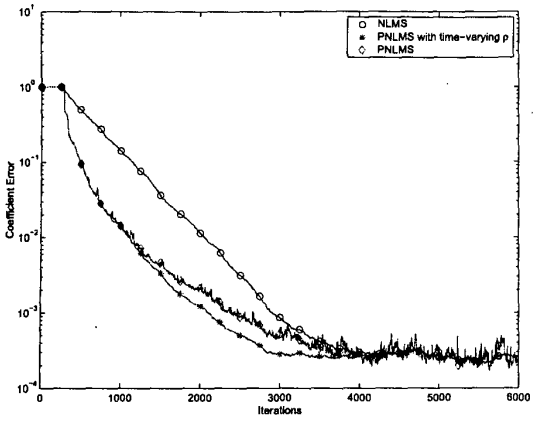

Fig. 2. Convergence curves for NLMS, PNLMS and PNLMS with time-varying ρ ($N = 256$, $\mu = 0.4$, $\rho = 5/N$, $\beta = 800$, SNR=30 dB).

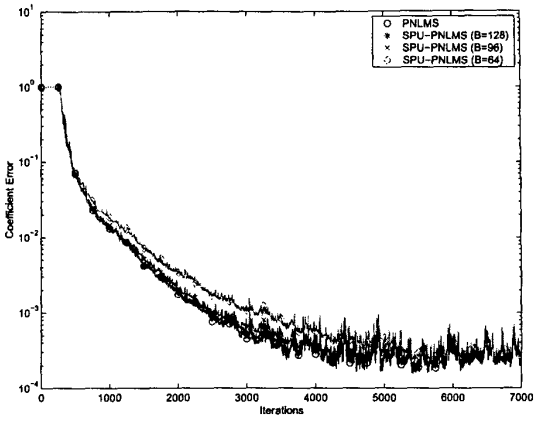

Fig. 3. Convergence curves for PNLMS and SPU-PNLMS.

gence of NLMS with standard PNLMS, as well as PNLMS with time varying ρ as in (11) ($\beta = 800$). The PNLMS algorithm using (11) attains a faster convergence speed than both NLMS and standard PNLMS by about 1000 iterations. In fact, its convergence rate approximates the faster of the two algorithms.

The convergence speeds of PNLMS and its reduced-complexity variant SPU-PNLMS have been simulated for the sparse network echo path in Fig. 1. The echo excitation signal $x(k)$ is zero-mean, unit-variance white Gaussian noise. The near-end noise is also zero-mean white Gaussian. The signal-to-noise ratio (SNR) for the echo signal is 30 dB. Fig. 3 shows the convergence curves for PNLMS and SPU-PNLMS for different parameter selections. The simulation parameters were $N = 256$, $M = 256$ ($L = 1$), $\mu = 0.4$. SPU-PNLMS was simulated for three values of B, viz., 128, 96 and 64, corresponding to $1/2$, $3/8$ and $1/4$ of the filter coefficients being updated at every iteration, respectively. The convergence speed of SPU-PNLMS is almost identical to that of PNLMS for values of B larger than the length of the active network echo path segment, which is 64 in this case.

5. CONCLUSION

A reduced-complexity selective-partial-update PNLMS algorithm has been developed for identification of discrete-time sparse systems. The chief application of this type of adaptive filtering algorithms is in adaptive network echo cancellation. The algorithm is derived from the solution of a constrained weighted optimization problem. A time-varying parameter approach has been proposed to remedy the slow convergence of PNLMS after the rapid initial convergence. SPU-PNLMS was also shown to be capable of maintaining the same convergence speed as PNLMS at a reduced complexity.

6. REFERENCES

[1] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaption in echo cancelers," *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 5, pp. 508–518, September 2000.

[2] S. L. Gay, "An efficient, fast converging adaptive filter for network echo cancellation," in *Proc. ASILOMAR*, November 1998, pp. 394–398.

[3] K. Doğançay and O. Tanrıkulu, "Adaptive filtering algorithms with selective partial updates," *IEEE Trans. on Circuits and Systems II*, vol. 48, no. 8, August 2001, in press.

[4] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*, Springer-Verlag, Berlin, Germany, 2001.

[5] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction, and Control*, Prentice Hall, New Jersey, 1984.

[6] International Telecommunication Union, *Digital Network Echo Cancellers*, Recommendation ITU-T G.168, April 2000.

Authorized licensed use limited to: University of South Australia. Downloaded on June 9, 2009 at 21:57 from IEEE Xplore. Restrictions apply.