STOCHASTIC GRADIENT ALGORITHMS IN ACTIVE CONTROL*

Alberto González
Dept. of Comunicaciones, Universidad Politécnica de Valencia,
Camino de Vera s/n,
Valencia 46071, SPAIN
Tel: +34 6 3877300; fax: +34 6 3877309
`e-mail: agonzal@dcom.upv.es`

ABSTRACT

This paper deals with some aspects concerning to the practical implementation of the stochastic gradient algorithms in active control. The control system under study is assumed to be a multichannel feedforward system and it is also assumed that there is not feedback signals from the secondary sources measured at the detection sensors.
Several iterative algorithms were developed in [1] [2] for a frequency domain model of a multichannel active noise control system. Such iterative algorithms related to the p-norm of the error signal vectors were then applied to control pure tones in time domain [3].
When the disturbance signals can be modelled as stationary stochastic processes, a different framework is needed although there exist some analogies with the frequency domain model. This paper reviews the development and implementation techniques of stochastic gradient algorithms for active control under a general point of view, and then focuses on algorithms called minimax type which were studied in the frequency domain in [1] [2] [3].

1 CANCELLATION OF STATIONARY STOCHASTIC PROCESSES IN ACTIVE CONTROL

The model described in [4] and [5] will be used. Fig. 1 shows the block diagram of the active noise control system model. There are M secondary sources, L error sensors and K reference signals. The diagram contains the next blocks: a matrix $\mathbf{W}$ composed by the MxK impulsional responses of the *electric controller* (the adaptive filters); and the system responses matrix $\mathbf{C}$ which is composed by the LxM system responses between each secondary source input and each error sensor output, these system responses are commonly called *the error paths*.

The matrix $\mathbf{W}$ is considered for the moment time invariant and it is implemented as matrix of FIR filters with coefficients w_{mki}. The error paths responses can also be modelled as a matrix $\mathbf{C}$ of FIR filters with coefficients c_{lmj} and a duration of J samples each filter. The output of the j-th error sensor can be written as,

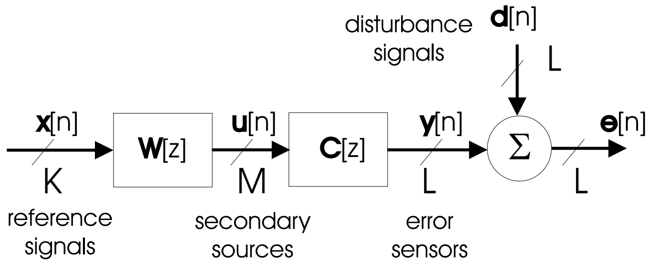

Figure 1: Multichannel pure feedforward ANC system.

$$e_l\left[n\right] = d_l\left[n\right] + \sum_{m=1}^{M}\sum_{j=0}^{J-1} c_{lmj}\, u_m\left[n-j\right] \qquad (1)$$

where $d_l\left[n\right]$ is the noise due to the primary noise measured at the l-th sensor. The input to the m-th secondary source $u_m[n]$, can also be written as,

$$u_m\left[n\right] = \sum_{k=1}^{K}\sum_{i=0}^{I-1} w_{mki}\, x_k\left[n-i\right] \qquad (2)$$

where $x_k\left[n\right]$ is a sampled version of the k-th reference (or primary) signal. Using (2), (1) becomes,

$$e_l\left[n\right] = d_l\left[n\right] + \sum_{k=1}^{K}\sum_{i=0}^{I-1}\sum_{m=1}^{M} w_{mki}\, r_{lmk}\left[n-i\right] \qquad (3)$$

where the filtered reference signals are defined as,

$$r_{lmk}\left[n\right] = \sum_{j=0}^{J-1} c_{lmj}\, x_k\left[n-j\right] \qquad (4)$$

The order of the filtering of the reference signals has be changed in (3) since $\mathbf{W}$ was assumed to be time invariant. Equation (3) illustrates the linear relationship between the error signal and the controller coefficients and can be expressed in matrix form as, [4] [5],

*Partially supported by CICYT grant no. TAP97-1128

$$\mathbf{e}[n] = \mathbf{d}[n] + \mathbf{R}[n]\,\mathbf{w} \tag{5}$$

where,

$$\begin{aligned} \mathbf{e}[n] &= [e_1[n], e_2[n], \ldots, e_L[n]]^T \\ \mathbf{d}[n] &= [d_1[n], d_2[n], \ldots, d_L[n]]^T \\ \mathbf{w} &= \left[\mathbf{w}_0^T, \mathbf{w}_1^T, \ldots, \mathbf{w}_{I-1}^T\right]^T \end{aligned} \tag{6}$$

and

$$\mathbf{w}_i = [w_{11i}, w_{12i}, \ldots, w_{21i}, w_{22i}, \ldots, w_{MKi}]^T , \tag{7}$$

$$\mathbf{R}[n] = \begin{bmatrix} \mathbf{r}_1^T[n] & \mathbf{r}_1^T[n-1] & \ldots & \\ \mathbf{r}_2^T[n] & \mathbf{r}_2^T[n-1] & \ldots & \\ . & . & & \\ . & & & \\ . & & & \\ . & & & \mathbf{r}_L^T[n-I+1] \end{bmatrix} \tag{8}$$

with

$$\mathbf{r}_l[n] = [r_{l11}[n], r_{l12}[n], \ldots, r_{l21}[n], r_{l22}[n], \ldots, r_{lMK}[n]]^T . \tag{9}$$

The optimum values of the coefficients vector $\mathbf{w}$ depend on the cost function related to the error signals whose minimisation is wanted. A cost function analogous to the p-norm of the error signal vector, $\mathbf{e}[n]$, can be defined as a first approach to a general cost function, as in [1] [2]. This cost function is chosen as the sum of the p-order moments of the different error signals [6] and it is therefore given by,

$$J_p = E\left\{\|\mathbf{e}[n]\|_p^p\right\} = E\left\{\sum_{l=1}^{L} |e_l[n]|^p\right\} \tag{10}$$

where $0 < p < \infty$.

The existence of optimum vectors or *minimizers* of (10) is assured if J_p is a convex function (it must be also assured that processes $e_l[n]$ has finite p-order moments). A closed expression for the control vector which minimises (10) seems not to exist for the general case, but there exists for $p = 2$. In any case, a iterative expression to minimise J_p based on an steepest descent method can be used to reach this optimum. As the true gradient vector of the cost function is difficult to calculate in practice, an stochastic estimation of the gradient vector can be used instead [7]. The stochastic gradient vector is given by,

$$\mathbf{g}_p[n] = \frac{\partial J_p[n]}{\partial \mathbf{w}} = \sum_{l=1}^{L} |e_l[n]|^{p-2} \mathbf{R}_l^T[n] e_l[n] \tag{11}$$

where $\mathbf{R}_l[n]$ corresponds to the *l-th* row of the filtered reference signals matrix $\mathbf{R}[n]$ and $e_l[n]$ is the signal at the *l-th* sensor.

An iterative algorithm to minimise (10) is built using the stochastic gradient vector definition in (11).

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \alpha \sum_{l=1}^{L} |e_l|^{p-1} \mathbf{R}_l^T[n] \mathrm{sign}\left(e_l[n]\right) \tag{12}$$

where $0 < p < \infty$ and α is called the convergence parameter.

Equation (12) is the vector form of the expression given in [6] for the algorithm called Least Mean p-norm, LMP. The recursion in (12) does not seem very useful in the general case. However there exist some cases where (12) can be useful. For the 1-norm, equation (12) becomes,

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \alpha \sum_{l=1}^{L} \mathbf{R}_l^T[n] \mathrm{sign}\left(e_l[n]\right) \tag{13}$$

which is a general form of the sign LMS algorithm. The single channel version of this algorithm is called Least Mean Absolute Deviation, LMAD, and it can be found in [6]. The sign algorithm has shown to be very useful in real time control systems, since it saves calculations compared to the LMS.

The sign algorithm has been used in some ANC applications, as an example, in [8] the single error version of this algorithm is used to build a computational efficient implementation of the waveform synthesis method. This approach can be also found in [9]. In any case, the physical sense of the minimisation of J_1 needs to be explained to validate this cost function in active control applications.

When $p = 2$ equation (12) becomes,

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \alpha \sum_{l=1}^{L} \mathbf{R}_l^T[n] e_l[n] \tag{14}$$

which is, as it was expected, the same iterative expression as the definition of the multiple error filtered-X LMS, MELMS, [10], which has been broadly used in the field of active control.

There can be found some algorithms which minimise the cost function defined in (10) with other finite values of p [11] [12] but all of them have been formulated for the single channel case and not for the general case as we do here. The application described in [12] uses a single channel-single error version of the iterative algorithm in (12) to carry out the active control of impulsive noise in a duct, with $0 < p < 2$.

2 MINIMAX TYPE STOCHASTIC GRADIENT ALGORITHMS

The cost function in (10) cannot be defined when the p-order moments of the processes $e_l[n]$ does not exist. Therefore, the limiting case, $p \to \infty$, which was carried out in the frequency domain model to develop the minimax algorithm [1], has in general no sense with (10).

The point is that the sum of the ∞-order moment of the error signal is not what is really wanted to minimise. A minimax type algorithm in active control is intended to balance the acoustic field after control, therefore it is needed to define which measure of the acoustic field is wanted to balance and then apply a minimax strategy of minimisation using this measure. A cost function definition which takes into account the last discussion is given below,

$$J_\infty(q) = \lim_{p\to\infty} \sqrt[p]{\sum_{1\le l\le L} \left(E\{\,|e_l[n]|^q\,\}\right)^p} = \max_{1\le l\le L} E\{|e_l[n]|^q\} = E\{|e_b[n]|^q\} \qquad (15)$$

where $q > 0$ is a parameter which can be selected to change the error signals measure. This measure is given in (15) by the error signals q-order moments.

Subscript b selects between the error signals, $1 \le l \le L$, that one with larger q-order moment. It is important to note that the value of subscript b depends on the current value of the control vector $\mathbf{w}$, a change of this vector could imply a change of the error signal with larger q-moment and so of the value of b.

It is useful to consider the instantaneous value of (15). This objective function, instantaneous or stochastic, consists in the instantaneous maximum value of the error signals (without sign) raised to the *q-th* power,

$$J_\infty[n,q] = \max_{1\le l\le L} |e_l[n]|^q = |e_b[n]|^q \qquad (16)$$

This function chooses the sample of the error signal of maximum absolute value at each iteration (any of them is chosen in case there are several equal values).

The stochastic gradient is calculated as in previous cases and is given as below (the values of q can be restricted to even numbers to avoid problems with the signs in the derivatives),

$$\mathbf{g}_\infty[n,q] = q|e_b[n]|^{q-2}\mathbf{R}_b^T[n]e_b[n] \qquad (17)$$

The gradient in (17) is valid for a given value of the control vector $\mathbf{w}$, any change of this vector will change the value of the gradient.

It has really practical sense the case with $q = 2$ since the 2-order moments of the error signals are easily related to physical quantities whose mean square value has practical sense. Using the stochastic gradient an iterative algorithm can be found as,

$$\mathbf{w}[n+1] = \mathbf{w}[n] - \alpha\mathbf{R}_b^T[n]e_b[n] \qquad (18)$$

The algorithm in (18) will be called Least Maximum Mean Squares algorithm, LMMS. Subscript b denotes the error signal with maximum mean squared value for a given value of the control vector. Different versions of this algorithm can be implemented in practice, depending on the mean squared values estimation. However, any of them should lead to the same optimum value (in case they converge). In the same way, the control vector update in (18) can be carried out each block of samples instead of each single sample, so defining a block LMMS.

The main idea of the LMMS algorithm lays on the reasoning given above. Even though the LMMS implementations using the stochastic gradient lead to the same iterative algorithm, different mean squares estimations lead to different versions of the algorithm. These different versions do not change neither the update equation nor the optimum value but they can have different convergence properties. Therefore it can exist some versions with poor convergence properties or even with no convergence at all.

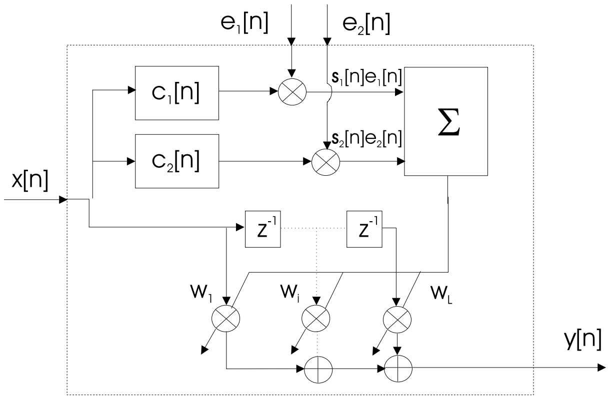

Figure 2: ANC system scheme for two microphones and the *MELMS* algorithm.

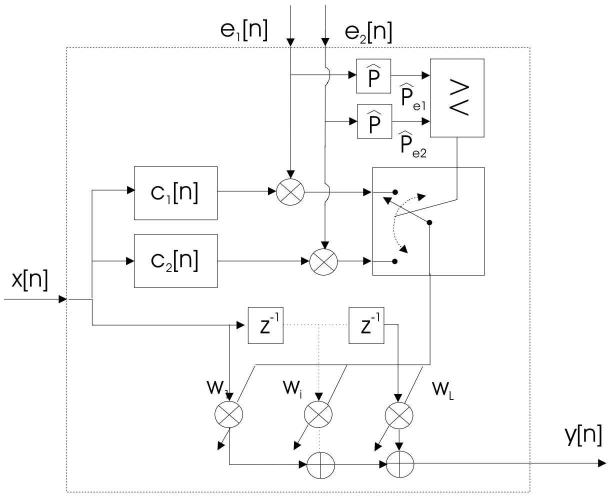

Figure 3: ANC system scheme for two microphones and the LMMS algorithm.

3 A PRACTICAL EXAMPLE

Two iterative algorithms described in the last section were used to cancel a broadband gaussian noise signal whose spectrum was in the range from 10 to 150 Hz in a pure feedforward active control system working in a duct [3]. The system was composed by one primary source, one secondary source and two error sensors. The number of adaptive filter taps was 145. The next expressions apply in this case: equation (14) for the MELMS and equation (18) for the LMMS.

The block diagrams that describe these algorithms are given by figures 2 and 3.

The LMMS algorithm is shown in figure 3. The taps update depends on the comparison between the error signals powers. The error signal whose estimated mean power, $\hat{P}_{e1}$ and $\hat{P}_{e2}$, was larger at instant n is used in the corresponding algorithm iteration. The mean power of the error signal was calculated in this example by means of an exponential window estimator which is defined as,

$$\hat{P}_e\,[n+1] = \alpha \hat{P}_e\,[n] + (1-\alpha) e^2\,[n]\,, \qquad (19)$$

where the parameter α is called the forgetting factor and it is typically chosen $0.91 \leq \alpha \leq 0.94$. The exponential window was used for its low computation requirements and it was validated in the laboratory experiments.

Figure 4 illustrates the cancellation results when either the MELMS or the LMMS algorithms are used. Both algorithms lead to the same final error signals as it was expected [3]. It was shown in this particular application that the LMMS algorithm saves computations and achieves in general a faster convergence than the MELMS algorithm.

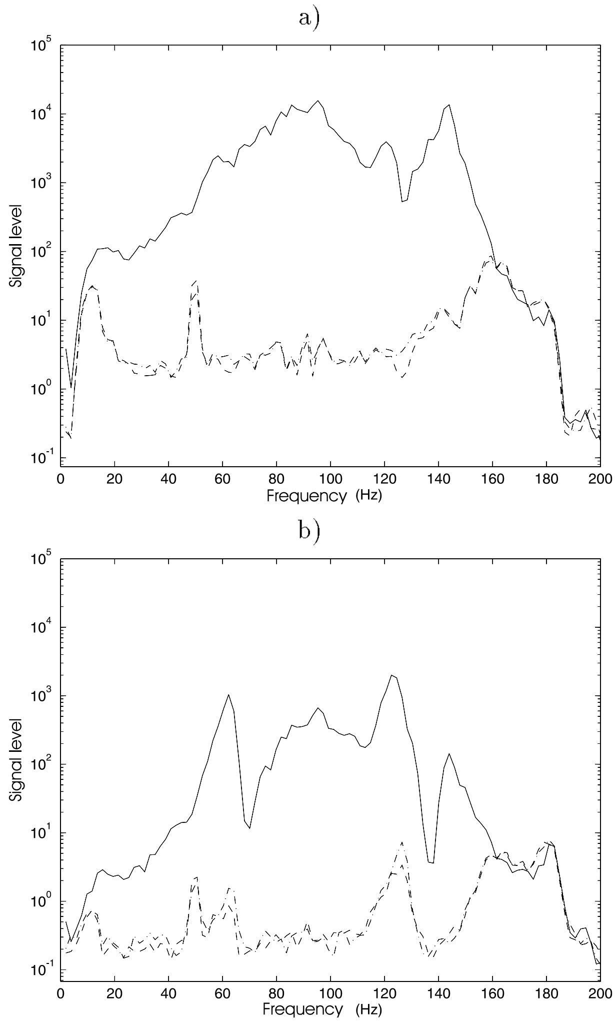

Figure 4: Power spectral density of the signal measured at the error sensors: before the ANC system operation (solid line), after the ANC system operation using the MELMS (dashed line) and after the ANC system operation using the LMMS (pointed and dashed line). a) First sensor signal and b) second sensor signal.

References

[1] A. Gonzalez, A. Albiol and S.J. Elliott, "Minimisation of the maximum error signal in active control," *IEEE Transactions on Acous. Speech and Signal Processing*, vol. 6, no. 3, pp. 1-14, May 1998.

[2] A. Gonzalez, S.J. Elliott and A. Albiol, "Minimisation of the maximum error signal in active control," *International Conference on, Acoustics, Speech and Signal Processing, ICASSP 97*, vol 1, pp. 387-390, April 1997.

[3] A. Gonzalez, S.J. Elliott, E. Cervantes and A. Albiol, "A practical approach for the minimisation of the maximum error signal in an active control system,"*1997 International Symposium on Active Control of Sound and Vibration, ACTIVE 97*, Budapest, Hungary, August 1997.

[4] P.A. Nelson and S.J. Elliott, *Active Control of Sound.* Academic Press, New York, 1992.

[5] S.J. Elliott and T.J. Sutton, "Performance of feedforward and feedback systems for active control," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 4, no. 3, pp. 214-223, May 1996.

[6] M. Shao and C.L. Nikias, "Signal processing with fractional lower order moments: stable processes and their applications". *Proc. IEEE*, vol. 81, no. 7, pp. 986-1010, July 1993.

[7] S. Haykin, *Adaptive Filter Theory.* Prentice-Hall, 3 Ed., 1996.

[8] G.B.B. Chaplin and R.A. Smith, "Waveform synthesis-the Essex solution to repetitive noise and vibration," *Proceedings of Inter-noise*, pp. 399-402, 1983.

[9] S.M. Kuo and D.R. Morgan, *Active Noise Control Systems.* John Wiley & Sons, 1996.

[10] S.J. Elliott, I.M. Stothers and P.A. Nelson, "A multiple error LMS algorithm and its application to the active control of sound and vibration". *IEEE Trans. Acoust. Speech Signal Processing*, vol. 35, no. 10, pp. 1423-1434, October 1987.

[11] E. Walach and B. Widrow, "The least mean fourth (LMF) algorithm and its family," *IEEE Trans. Information Theory*, vol. IT-30, pp. 275-283, 1984.

[12] R. Leahy, Z. Zhou and Y. Hsu, "Adaptive filtering of stable processes for active attenuation of impulsive noise," *International Conference on, Acoustics, Speech and Signal Processing, ICASSP 95*, pp. 2983-2986, April 1995.