Robustness Analysis of the Adaptive Periodic Noise Canceller Applied to Resonance Cancellation

J. Timoney
Department of Computer Science,
NUI Maynooth,
Maynooth,
Co. Kildare,
Ireland

J.B. Foley
Department of Electronic and Electrical Engineering,
Trinity College,
Dublin 2,
Ireland

Abstract

In this paper the criteria determining robustness of a LMS-driven Adaptive Periodic Noise Canceller (APNC) when applied to the cancellation of coloured interference signals are investigated. The upper bound on the algorithm stepsize is the crucial quantity for determining robustness and here relevant expressions for this upper bound are developed, followed by experimental evaluation.

1. Introduction

In applications where cancellation of a time-varying interference, such as an acoustic resonance, is required, adaptive filtering is the most commonly proposed solution as the unknown nature of the interference demands a time-variant filter structure to ensure continuous tracking and cancellation. As the interference may exhibit rapid fluctuations in magnitude, an adaptive filter with robust performance properties is essential to prevent disturbance of the adaption process by these fluctuations. Specifically, the LMS-driven Adaptive Periodic Noise Canceller (APNC) has been shown experimentally to be a successful solution to resonance cancellation [1] [2]. However, although the requirement of algorithm stability is necessary for proper algorithm operation, it may not be sufficient to guarantee a practically useful level of performance at all times. Fortunately, the concept of robustness provides an alternative framework within which it is possible to establish a more meaningful criterion for algorithm performance. A theoretical proof of the robustness of the LMS algorithm in the sense of it being H^∞ -optimal was given in [3]. Here, the value of the algorithm stepsize was shown to be the crucial parameter for ensuring algorithm robustness. Following from this, an empirical determination of the necessary conditions for robust performance of the APNC in a one-dimensional open-loop feedback system with a white Gaussian input was given in [4], and for a one-dimensional closed-loop feedback system under similar input conditions was given in [5]. Limiting the input properties to be Gaussian, this work demonstrated how an empirical justification of APNC robustness could be achieved. This paper aims to extend these analyses to a two-dimensional open-loop APNC configuration whose input consists of an acoustic resonance and an information signal. The increased dimensionality allows arbitrary placement of the resonance in terms of frequency and magnitude. Additionally, the aim is to supersede the empirical approach in favour of a more rigorous derivation of a suitable expression for the maximum allowable stepsize to ensure robust performance.

2. Background

Little work has gone towards creating a methodology under which the performance of interference cancelling algorithms can be deemed to be robust. However, significant advances have been made in the field of estimation theory as many of the classical methods were not always directly applicable in cases where there was a high level of uncertainty regarding the statistics and distributions of the underlying signals being input to the system under examination. One particular approach known as H^∞ estimation theory arose from

developments in robust control theory [6]. It was founded on the premise that for a non-robust estimation procedure small disturbances corrupting the measurement combined with modelling errors may lead to large estimation errors. Therefore, any approach to robust estimation requires a measure of the largeness and smallness of the signals involved, and specifically within the H^∞ approach the focus was on the energy of these quantities [7]. The main idea was to come up with estimators that minimise (or in the suboptimal case, bound) the maximum energy gain from any disturbances to the estimation errors. Thus, it is guaranteed that if the disturbances are small (in energy) then, no matter what those disturbances are, the estimation errors will be as small as possible (in energy). The robustness of H^∞ estimators, with respect to disturbance variation, follow from the fact that they protect against the estimators worst-case performance and make no assumption on the statistics or distributions of the disturbance signals. In other words, the maximum energy gain is minimised over all possible disturbances. The primary drawback of the H^∞ approach is that as it makes no assumptions about the disturbances and has to accommodate for all conceivable ones, it may be over-conservative.

Although it has long been known experimentally that the Least-Mean Square (LMS) algorithm was robust, a good example being its superior tracking capability over the RLS algorithm in the presence of non-stationary inputs [8], this observed robustness was not justified theoretically until [3] established its formal link to H^∞ theory. This paper set out to prove the LMS algorithm was robust in the sense of it being H^∞ optimal and the primary result was that the value of the algorithm stepsize was shown to be the crucial parameter for ensuring algorithm robustness. Although the interference cancelling problem was not examined within this work, the principles given exposed a procedure which could be adapted and applied.

2.1 H^∞ approach to LMS Robustness

First it is necessary to begin with the definition of the H^∞ norm of a transfer operator:

Definition - The H^∞-norm: Consider a linear time-invariant system that maps an input sequence $\{u_i\}$ to an output sequence $\{y_i\}$. The energy gain γ^2 between the input and output is

$$\gamma^2 = \frac{\|y\|_2^2}{\|u\|_2^2} \tag{1}$$

where $\|u\|_2^2$ denotes the squared h_2-norm of the real causal sequence $\{u_i\}$ (i.e. $\|u\|_2^2 = \sum_{k=0}^{\infty} u_k u_k$ and $\|y\|_2^2 = \sum_{k=0}^{\infty} y_k y_k$).

Denoting T as the transfer operator that maps the input vector sequence $\{u_i\}$ to the output sequence $\{y_i\}$,

$$y_i = T u_i \tag{2}$$

The H^∞-norm of T is defined as the maximum value for γ over all possible square-summable causal input sequences h_2,

$$\|T\|_\infty = \sup_{u \neq 0, u \in h_2} \gamma \tag{3}$$

or

$$\|T\|_\infty = \sup_{u \neq 0, u \in h_2} \frac{\|y\|_2}{\|u\|_2} \tag{4}$$

and is thus directly related to the maximum energy gain from the input to the output [3].
In the context of adaptive filtering, which belongs to the finite horizon case above, robust performance implies that that the adaptive solution minimises the maximum energy gain from

the disturbances to the estimation errors [3]. A measure of the energy gain is found by taking the H^∞- norm of the transfer operator T that maps these disturbances to the estimation errors. As the adaptive filter is a time-varying system, this measurement must be taken with each iteration of the algorithm. Furthermore, T is a function of the weight estimation strategy i.e. $T(\mathbf{W})$, which in turn is a function of all previous filter inputs and outputs.

$$T(\mathbf{W}) = f\big(e(0),\ldots,e(n); X(0),\ldots,X(n)\big) \tag{5}$$

where $e(n)$ is the filter output and $X(n)$ is the filter input vector at time n.

In the case of the LMS algorithm, denoting the a priori error $\varepsilon(n)$ and the input disturbances collectively as $v(n)$, the idea is to find a suitable value of algorithm stepsize so that the LMS algorithm will be a H^∞-optimal strategy that minimises $\|T(\mathbf{W})\|_\infty$ and will also obtain the resulting,

$$\gamma_{\text{opt}}^2 = \inf_f \|T(\mathbf{W})\|_\infty^2 = \inf_f \sup_{w,v \in h_2} \frac{\|\varepsilon\|_2^2}{\mu^{-1}\big|\mathbf{W}^* - \mathbf{W}_0\big| + \|v\|_2^2} \tag{6}$$

where $\mathbf{W}^*$ is the optimum Wiener weight value, $\mathbf{W}_0$ is the initial weight vector and μ is the algorithm stepsize [3].

Examining (6), it can be seen that the transfer operator is a measure of the "amplification" of the noise given the estimate. Furthermore, the ratio depends on the choice of input disturbance. To remove this dependence it is important to consider the largest energy gain over all possible disturbance sequences, $v(n)$. This is equivalent to finding the H^∞- norm of the transfer operator $\|T(\mathbf{W})\|_\infty$.

For the LMS algorithm to exhibit robust performance the worst case disturbance energy gain must be less than or equal to 1, i.e. no amplification, and for it to be H^∞-optimal [3]

$$\gamma_{\text{opt}} = \inf_f \|T(\mathbf{W})\|_\infty \leq 1 \tag{7}$$

For the LMS algorithm H^∞-optimality is determined by the value of stepsize μ, so by choosing the value of the stepsize of the algorithm to lie below a defined upper bound, the criterion of optimality given in (7) will be satisfied. The proof of H^∞-optimal of the LMS algorithm given in [3] for the system identification problem relies on the filter input being persistently exciting, i.e.

$$\lim_{N \to \infty} \sum_{n=0}^{N} X^2(n) = \infty \tag{8}$$

If this is not the case, the solution will be sub-optimal.

3. Method

The aim here is to investigate robust interference cancellation, specifically, the removal of coloured interference from a desired information signal by an Adaptive Periodic Noise Canceller (APNC), a variant on the Adaptive Line Enhancer of Widrow [9]. Following from this, Figure 1 below shows the APNC connected in an open-loop configuration. The intention is that the filter will cancel the resonance at its input, leaving only the information signal $s(n)$ at its output. In this case, the information signal is assumed to have Gaussian properties so the filter decorrelation delay can set to unity; however this does not restrict the generality of the analysis.
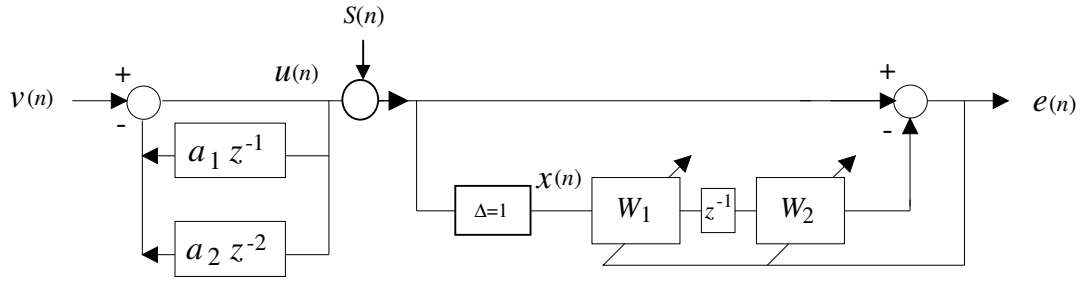
Fig. 1 Block diagram of APNC Configuration

The resonance is represented by $u(n)$ which is the output of a two-dimensional AR process. The investigation is focused on robust performance of the APNC when the filter input $x(n)$ is considered as the disturbance.

The filter output is given by

$$e(n) = s(n) + u(n) - \mathbf{W}_n X^T(n) \qquad (9)$$

where $\mathbf{W}_n$ is the filter weight vector

The filter input can be denoted as

$$X(n) = [s(n-1) + u(n-1) \quad s(n-2) + u(n-2)]^T \qquad (10)$$

giving

$$e(n) = x(n+1) - \mathbf{W}_n X^T(n) \qquad (11)$$

Ideally, the filter will remove the autoregressive signal and only the information will be present at the output [2], i.e.

$$e(n) \approx s(n) \qquad (12)$$

The best possible performance of the APNC will be given by the optimum Wiener solution whose filter weight vector can be denoted as $\mathbf{W}^*$. In practice, once the algorithm stepsize is within a certain bound, the LMS solution will approach but not exactly reach the optimum solution and will therefore produce a misadjustment error vector which is given by

$$\boldsymbol{\varepsilon}(n) = (\mathbf{W}_n - \mathbf{W}^*) X^T(n) \qquad (13)$$

where

$$\boldsymbol{\varepsilon}(n) = \begin{bmatrix} \varepsilon_1(n) \\ \varepsilon_2(n) \end{bmatrix}$$

Substituting $\boldsymbol{\varepsilon}(n)$ into (11) above gives

$$e(n) = x(n+1) - \boldsymbol{\varepsilon}(n) - \mathbf{W}^* X^T(n) \qquad (14)$$

Writing the LMS algorithm,

$$\mathbf{W}_{n+1} = \mathbf{W}_n + 2\mu e(n) X(n) \qquad (15)$$

Then, subtracting $\mathbf{W}^*$ from both sides of (15) gives the LMS algorithm update equation in terms of the weight misadjustment

$$\mathbf{B}_{n+1} = \mathbf{B}_n - 2\mu e(n)X(n) \qquad (16)$$

The goal is to ensure robustness of the APNC against fluctuations in the filter input $X(n)$. Thus, based on (6), the criterion for robustness can be cast as ensuring that the energy of residual error $\varepsilon(n)$ is upper bounded by the energy of the disturbances and the initial uncertainty [3], i.e.

$$\frac{\displaystyle\sum_{n=1}^{\infty}\varepsilon_1^2(n)}{(2\mu)^{-1}B_{11}^2 + \displaystyle\sum_{n=1}^{\infty}x^2(n)} \leq 1 \qquad (17a)$$

$$\frac{\displaystyle\sum_{n=1}^{\infty}\varepsilon_2^2(n)}{(2\mu)^{-1}B_{21}^2 + \displaystyle\sum_{n=1}^{\infty}x^2(n-1)} \leq 1 \qquad (17b)$$

Denote the transfer operators that map the input disturbance vectors $\left\{(2\mu^{-1/2})B_{11}, \{x(n)\}_{n=1}^{\infty}\right\}$ and $\left\{(2\mu^{-1/2})B_{21}, \{x(n-1)\}_{n=1}^{\infty}\right\}$, where $(2\mu^{-1/2})B_{11}$ and $(2\mu^{-1/2})B_{21}$ are the (weighted) energies of the respective weight errors due to the initial guess, to the residual output error $\{\varepsilon(n)\}_{n=1}^{\infty}$ as $D_{1n}(\mu)$ and $D_{2n}(\mu)$. Then, for robustness it is sufficient to ensure that for each iteration of the algorithm the H^{∞}-norms of the transfer operators $D_{1n}(\mu)$ and $D_{2n}(\mu)$ are less then one [7], i.e.

$$\inf_{n}\left\|D_{1n}(\mu)\right\|_{\infty} \leq 1 \qquad (18a)$$

and

$$\inf_{n}\left\|D_{2n}(\mu)\right\|_{\infty} \leq 1 \qquad (18b)$$

To determine an expression for a suitable upper bound on the algorithm stepsize that will ensure algorithm robustness against the disturbances, first multiply (16) by $(2\mu)^{-1/2}$

$$(2\mu)^{-1/2}\mathbf{B}_{n+1} = (2\mu)^{-1/2}(\mathbf{B}_n - 2\mu e(n)X(n)) \qquad (19)$$

Given the disturbance vector $X(n)$, square (19) and then subtract the square on the disturbance from both sides, i.e.

$$(2\mu)^{-1}\mathbf{B}_{n+1}^2 - X^2(n) = (2\mu)^{-1}(\mathbf{B}_n + 2\mu e(n)X(n))^2 - X^2(n) \qquad (20)$$

This becomes

$$(2\mu)^{-1}\mathbf{B}_{n+1}^2 - X^2(n) = (2\mu)^{-1}(\mathbf{B}_n^2 + 4\mu\mathbf{B}_n e(n)X(n) + 4\mu^2 e^2(n)X^2(n)) - X^2(n) \qquad (21)$$

giving

$$(2\mu)^{-1}\mathbf{B}_{n+1}^2 - X^2(n) = (2\mu)^{-1}\mathbf{B}_n^2 + 2C_n\,e(n)X(n) + 2\mu\,e^2(n)X^2(n) - X^2(n) \quad (22)$$

Then, substitute $x(n+1) - \boldsymbol{\varepsilon}(n) - \mathbf{W}^* X^T(n)$ for $e(n)$ in the second term on the RHS of (22)

$$(2\mu)^{-1}\mathbf{B}_{n+1}^2 - X^2(n) = (2\mu)^{-1}\mathbf{B}_n^2 + 2\boldsymbol{\varepsilon}(n)\big(x(n+1) - \boldsymbol{\varepsilon}(n) - \mathbf{W}^* X^T(n)\big) - \big(1 - 2\mu e^2(n)\big)X^2(n)$$

(23)

leading to

$$(2\mu)^{-1}\mathbf{B}_{n+1}^2 + \boldsymbol{\varepsilon}^2(n) = (2\mu)^{-1}\mathbf{B}_n^2 + X^2(n) + 2\boldsymbol{\varepsilon}(n)x(n+1) - \boldsymbol{\varepsilon}^2(n) - 2\mathbf{W}^* X^T(n)\boldsymbol{\varepsilon}(n)$$
$$- \big(1 - 2\mu e^2(n)\big)X^2(n)$$

(24)

For robustness, the stepsize must be chosen so that the sum of the last three terms on the RHS of (24) is zero or smaller, i.e.

$$(2\mu)^{-1}\mathbf{B}_{n+1}^2 + \boldsymbol{\varepsilon}^2(n) \le (2\mu)^{-1}\mathbf{B}_n^2 + X^2(n) \quad (25)$$

if

$$2\boldsymbol{\varepsilon}(n)x(n+1) - \boldsymbol{\varepsilon}^2(n) - 2\mathbf{W}^* X^T(n)\boldsymbol{\varepsilon}(n) - \big(1 - 2\mu e^2(n)\big)X^2(n) \le 0 \quad (26)$$

Rearranging (26) results in the following upper bounds for the stepsize μ

$$\mu_1 \le \frac{\varepsilon_1^2(n) - 2\varepsilon_1(n)x(n+1) + x^2(n) + 2W_1^* x(n)\varepsilon_1(n)}{2e^2(n)x^2(n)} \quad (27\text{a})$$

and

$$\mu_2 \le \frac{\varepsilon_2^2(n) - 2\varepsilon_2(n)x(n+1) + x^2(n-1) + 2W_2^* x(n-1)\varepsilon_2(n)}{2e^2(n)x^2(n-1)} \quad (27\text{b})$$

Taking (27a) as an example, to eliminate the dependence on $x^2(n)$ in the above, substitute for $\varepsilon_1(n)$

$$\mu_1 \le \frac{B_{1n}^2 x^2(n) + x^2(n) + 2W^* B_{1n} x^2(n) - 2B_{1n}x(n)x(n+1)}{2e^2(n)x^2(n)} \quad (28)$$

leaving

$$\mu_1 \le \frac{\left(B_{1n}^2 + 2W_1^* B_{1n} - 2B_{1n}\,x(n+1)\big/x(n) + 1\right)}{2e^2(n)} \quad (29)$$

Examining (29), if the algorithm is stable and convergent and has a zero initial weight vector, the maximum value that $e(n)$ should reach is equivalent the maximum of the sum of the input signals, i.e.

$$e_{\max}^2 = \sup_{n\ne0}\big(s(n) + u(n)\big)^2 \quad (30)$$

Then, evaluating

$$B_{1n}^2 + 2W_1^* B_{1n} - 2B_{1n} \frac{x(n+1)}{x(n)} \qquad (31)$$

gives

$$-W_1^{*2} + 2W_1^* \frac{x(n+1)}{x(n)} \qquad (32)$$

If it is assumed that, as the algorithm is convergent, the largest value of B_{1n} will be $-W_1^*$. Also, to ensure that the value of (32) and therefore the nominator of (28) are small, take the minimum of the absolute value of the ratio $\frac{x(n+1)}{x(n)}$, i.e.

$$D_{1\min} = -W_1^{*2} + 2W_1^* \min\left|\frac{x(n+1)}{x(n)}\right| \qquad (33)$$

Then, the following upper bound can be presented for the algorithm stepsize to ensure robustness

$$\mu_1 \le \frac{1 + D_{1\min}}{2\max(s(n)+u(n))^2} \qquad (34a)$$

Examining (33) and (34), the upper bound on the stepsize for robustness is a function of both the filter input and the optimum weight value.
Applying a similar analysis, (27b) becomes

$$\mu_2 \le \frac{1 + D_{2\min}}{2\max(s(n)+u(n))^2} \qquad (34b)$$

where

$$D_{2\min} = -W_2^{*2} + 2W_2^* \min\left|\frac{x(n+1)}{x(n-1)}\right| \qquad (35)$$

4. Results

Simulations were carried out for the system shown in Fig. 1 to investigate the necessary conditions to ensure robust performance of the APNC. The input to the all-pole filter section of Figure 1 was a 200-point Gaussian white noise sequence and the value of the all-pole filter coefficients were varied over the range 0.1 to 1 in steps of 0.1 to examine any possible relationship between noise colouration and robustness. The most difficult test for system robustness is when the filter poles are located on the unit circle as self-sustaining oscillations of increasing magnitude then appear at the filter output, and consequently the coupling between $D_{1n}(\mu)$ and $D_{2n}(\mu)$ is maximised. The information signal $s(n)$ was chosen to be Gaussian. The value of stepsize selected was as given by (34a). The H^∞-norms were calculated by finding the maximum singular values of $D_{1n}(\mu)$ and $D_{2n}(\mu)$ at each time instant, and are shown in Figure 2 and 3.
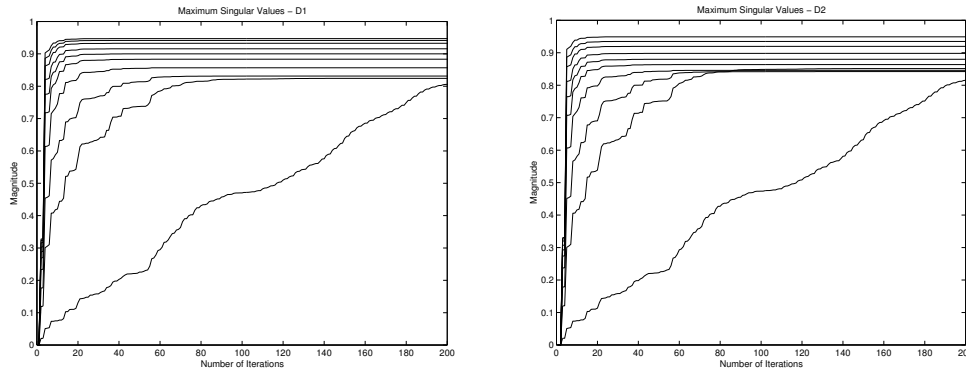
Fig. 2 Maximum Singular Values of $D_{1n}(\mu)$ Fig. 3 Maximum Singular Values of $D_{2n}(\mu)$

From Figs. 2 and 3, it is clear that the maximum singular values of $D_{1n}(\mu)$ and $D_{2n}(\mu)$ are less than unity in all cases demonstrating that robust performance of the APNC is guaranteed within the stepsize bound given by (34a). Examination of the results shows that for the larger values of the all-pole filter parameters the maximum singular values of $D_{1n}(\mu)$ and $D_{2n}(\mu)$ exhibit a slower rate of growth towards unity. This can be attributed to the relatively slower convergence of the LMS algorithm in these cases.

5. Conclusion

This paper has presented expressions for an upper bound on the stepsize of the LMS algorithm to ensure robustness of the APNC in an open-loop interference cancellation configuration. The expressions derived in both cases were similar and showed a dependence on both the magnitude of the filter input and the optimum weight value. The benefits of the robustness criterion are that it provides a practically useful qualitative measure of algorithm performance over the more vague requirement, in performance terms, of stability and it helps to quantify the notion that it is better to be conservative in choosing an algorithm stepsize. As regards future work, it would be interesting to extend the analyses for over- and under-determined systems, i.e. when the number of filter weights is mismatched with the underlying parameters of the interference, and for algorithms other than the LMS.

6. References

[1] J.B. Wright and J.B. Foley, "Adaptive periodic noise cancellation for the control of acoustic howling," in *Signal Processing V: Theories and Applications*, L.Torres, E. Masgrau, and M.A. Lagunas (eds.), Elsevier Science Publishers, 1990, pp. 1979-1982.

[2] J. Timoney and J.B. Foley, "Performance analysis of the adaptive periodic noise canceller with a second-order autoregressive input," Poster presentation *Euraco Tutorial Workshop*, Portugal, May 1996.

[3] B. Hassibi, A.H. Sayed and T. Kailath, "H^{∞} optimality of the LMS algorithm," *IEEE Trans. Acoust. Speech Sig. Proc.*, vol. ASSP-44, no. 2, Feb. 1996, pp. 267-280.

[4] J. Timoney, and B. Foley, "Robust performance of the adaptive periodic noise canceller applied to echo control," *Proc. IWAENC'97,* London, June 1997, pp. 25-28.

[5] J. Timoney, and B. Foley, "Robust performance of the adaptive periodic noise canceller in a closed loop system," *Proc. Eusipco'98,* Rhodes, Sept. 1998, pp. 1177-1180.

[6] H. Kwakernaak, "Robust control and H^{∞}-optimisation - Tutorial paper," *Automatica.*, vol. 29, no. 2, 1993, pp. 255-273.

[7] B. Hassibi, '*Indefinite metric spaces in estimation, control and adaptive filtering*,' PhD. Thesis, Stanford University, California, Aug.1996.

[8] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall International, London, 1996.

[9] B. Widrow et al, *Adaptive signal processing*, Prentice-Hall, London, 1985.