Convergence Analysis of a Twin-Reference Complex Least-Mean-Squares Algorithm

Sven Johansson, Sven Nordebo, and Ingvar Claesson

Abstract—In many noise control applications, the noise is dominated by low frequencies and generated by several independent periodic sources. In such situations the tonal noise may be suppressed by using a narrowband *multiple-reference* feedforward controller. The performance characteristics of the control system, e.g., the convergence behavior and noise reduction are directly related to the controller adaptation rate as well as the frequency separation of the tonal components in the noise, i.e., the beat frequency.

This paper treats the convergence performance of a complex least-mean-squares (LMS) algorithm using two reference signals. An analysis of its convergence behavior is presented as well as the results from computer simulations validating the convergence behavior. The convergence of the filter weights and the decrease rate of the squared error (the learning curve) for noise control applications are also discussed.

Index Terms—Active noise control, complex algorithm, convergence performance, feedforward control, LMS algorithm.

I. INTRODUCTION

ACTIVE noise control (ANC) is a highly useful method for reducing noise at low frequencies. This technique is based on destructive interference of sound fields, i.e., the primary sound field from the noise sources and the anti-noise produced by the control system. In many applications, the primary noise is generated by rotating machines and is therefore periodic. In such applications, the use of an adaptive feedforward control system has proved to be a successful means of suppressing the low-frequency noise [1]–[4].

A narrowband feedforward controller utilizes a synchronization signal, e.g., a tachometer signal from the noise source, which provides information about the radiated noise. This signal contains information about the rotational speed of the machine and may therefore be used to generate sinusoidal reference signals that have the same frequency content as the noise components to be suppressed. In many noise control applications, several independent noise sources contribute to the sound field, e.g., in twin-engined propeller aircraft [5]–[8], and vessels and motor boats [9], [10]. In such applications, a synchronization signal from each noise source and a *multiple-reference* control system are usually required to achieve an efficient and substantial noise reduction [11].

Manuscript received December 12, 2000; revised March 18, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Michael S. Brandstein.

S. Johansson and I. Claesson are with the Department of Telecommunications and Signal Processing, Blekinge Institute of Technology, S-372 25 Ronneby, Sweden (e-mail: sven.johansson@bth.se).

S. Nordebo is with the School of Mathematics and Systems Engineering Växjö University, S-351 95 Växjö, Sweden (e-mail: sven.nordebo@msi.vxu.se).

Publisher Item Identifier S 1063-6676(02)06033-9.

Fig. 1. Basic concept of narrowband active noise control using a twin-reference feedforward adaptive controller.

A multiple-reference approach allows the reference signals from each noise source to be processed individually within the controller, thereby enabling individual control of the sound field from each source. The principle of a narrowband twin-reference feedforward control system for ANC is illustrated in Fig. 1. The single-frequency reference signals are filtered individually using an adaptive filter before driving a secondary source, e.g., a loudspeaker. Each filter adjusts the amplitude and the phase of the corresponding reference signal. The loudspeaker produces a secondary sound field of equal amplitude which is 180° out of phase with the primary noise field. An error microphone is used to measure the residual noise achieved, and the output signal from this error sensor is then used to adjust the adaptive filters so that the level of the residual noise is minimized.

Periodic noise sources with a slight difference in the rotational speed induce an acoustic beating which may lead to considerable discomfort. However, comfort could be improved significantly if the level of the noise and beating is significantly reduced, e.g., by using an ANC system. The performance of an ANC system such as convergence rate and noise reduction is a direct result of the running conditions of the engines, i.e., the beat frequency. The running conditions for several engines can generally be divided into three main types:

- engines with exactly the same rotational speed;
- engines which have almost the same rotational speeds;
- engines which have significantly different rotational speeds.

Synchronization of the engines can be achieved either automatically or manually. Most twin-engined propeller aircraft are fitted with a synchrophaser unit to synchronize automatically

weight vector $\mathbf{w}(n)$ in the negative gradient direction of the cost function $J(n)$ [13], [14]. Hence, the algorithm for the updating of the filter weights can be expressed as

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{\mu}{2} \nabla J(n) \quad (7)$$

where $\nabla J(n)$ denotes the gradient vector of the cost function and μ is the step-size parameter that controls the convergence rate and the stability of the algorithm. The factor of two is included to simplify the final form of the algorithm. The relationship between the gradient vector and the derivative of the cost function with respect to the complex filter weights is given by [13, App. B]

$$\nabla J(n) = 2 \frac{\partial J(n)}{\partial \mathbf{w}^*(n)} \quad (8)$$

where

$$\frac{\partial J(n)}{\partial \mathbf{w}^*(n)} = \mathbf{x}(n)d^*(n) + \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}(n) = \mathbf{x}(n)e^*(n). \quad (9)$$

If (8) and (9) are now substituted into the filter weight update (7), the complex LMS adaptive algorithm is given by [12], [13]

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \mathbf{x}(n)e^*(n). \quad (10)$$

A. Convergence Analysis

The convergence investigation is based on a deterministic analysis. Hence, no assumptions are made about the independence between the input vector $\mathbf{x}(n)$ and the weight vector $\mathbf{w}(n)$. On the other hand, such assumptions are often used in the analysis of adaptive algorithms [13], [14].

Consider the derivative of the cost function with respect to the complex filter weights in (9). By setting the differential of this function with respect to the weight vector to zero, $(\partial J(n)/\partial \mathbf{w}^*(n)) = \mathbf{0}$, the following expression is obtained:

$$\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{w}_o = -\mathbf{x}(n)d^*(n). \quad (11)$$

Here, $\mathbf{w}_o$ denotes the optimum filter weight vector for which the quadratic cost function attains its minimum value $J_{\min}$. Note that the matrix $\mathbf{x}(n)\mathbf{x}^H(n)$ does not have full rank, so that $\mathbf{w}_o$ is nonunique, and for this reason we may choose $\mathbf{w}_o$ as

$$\mathbf{w}_o = -[\mathbf{x}(n)\mathbf{x}^H(n)]^+ \mathbf{x}(n)d^*(n) \quad (12)$$

where $[\cdot]^+$ denotes the pseudoinverse [17].

Now we define the weight-error vector $\mathbf{v}(n)$ as the difference between the adaptive weights and the optimal solution

$$\mathbf{v}(n) = \mathbf{w}(n) - \mathbf{w}_o. \quad (13)$$

By substituting (4) into (10) and using the expressions in (11) and (13), the recursive relation of the adaptive weights can be rewritten as

$$\mathbf{v}(n+1) = [\mathbf{I} - \mu \mathbf{x}(n)\mathbf{x}^H(n)]\mathbf{v}(n) \quad (14)$$

where $\mathbf{I}$ is the identity matrix. To simplify this equation, a new time variable matrix $\mathbf{A}(n)$ is introduced:

$$\mathbf{A}(n) = \mathbf{I} - \mu \mathbf{x}(n)\mathbf{x}^H(n). \quad (15)$$

Equation (14) can now be written as

$$\mathbf{v}(n+1) = \mathbf{A}(n)\mathbf{v}(n). \quad (16)$$

Equation (16) represents a first-order difference equation, the solution of which is

$$\mathbf{v}(n) = \left[\prod_{i=1}^n \mathbf{A}(n-i) \right] \mathbf{v}(0) \quad (17)$$

where $\mathbf{v}(0)$ denotes the initial weight-error vector. The matrix $\mathbf{A}(n)$ in (15) can be expressed in expanded form as

$$\mathbf{A}(n) = \begin{pmatrix} 1 - \mu & -\mu e^{-j\Delta\omega n} \\ -\mu e^{j\Delta\omega n} & 1 - \mu \end{pmatrix} \quad (18)$$

where

$$\Delta\omega = \omega_2 - \omega_1 = 2\pi \frac{\Delta f}{f_s} \quad (19)$$

and $\Delta f = f_2 - f_1$. Accordingly, the matrix $\mathbf{A}(n)$ depends both on the step-size μ and the frequency separation Δf between the reference signals $x_1(n)$ and $x_2(n)$.

By decomposing the Hermitian matrix $\mathbf{A}(n)$ [17], the matrix can be written as

$$\mathbf{A}(n) = \mathbf{Q}(n)\mathbf{\Gamma}\mathbf{Q}^H(n) \quad (20)$$

where $\mathbf{Q}(n)$ is a unitary matrix with columns containing the eigenvectors of $\mathbf{A}(n)$

$$\mathbf{Q}(n) = \frac{1}{\sqrt{2}} \begin{pmatrix} -e^{-j\Delta\omega n} & e^{-j\Delta\omega n} \\ 1 & 1 \end{pmatrix} \quad (21)$$

and $\mathbf{\Gamma}$ is a diagonal matrix whose elements consist of eigenvalues of $\mathbf{A}(n)$

$$\mathbf{\Gamma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 - 2\mu \end{pmatrix}. \quad (22)$$

Observe that the eigenvalues are determined by the step-size μ , while the eigenvectors are determined by the frequency separation $\Delta\omega$ and the time index n , resulting in a rotation of the eigenvectors where the rotation speed is determined by the frequency separation.

The convergence performance of the algorithm is related to the matrix product of $\mathbf{A}(n)$. By substituting (20) into (17) the matrix product can be expressed as

$$\begin{aligned} \prod_{i=1}^n \mathbf{A}(n-i) &= \mathbf{A}(n-1)\mathbf{A}(n-2) \dots \mathbf{A}(0) \\ &= \mathbf{Q}(n-1)\mathbf{\Gamma}\mathbf{Q}^H(n-1)\mathbf{Q}(n-2) \\ &\quad \times \mathbf{\Gamma}\mathbf{Q}^H(n-2)\mathbf{Q}(n-3) \dots \mathbf{Q}(0)\mathbf{\Gamma}\mathbf{Q}^H(0). \end{aligned} \quad (23)$$

By defining a new matrix

$$\begin{aligned} \mathbf{B} &= \mathbf{\Gamma}\mathbf{Q}^H(m)\mathbf{Q}(m-1) \\ &= \frac{1}{2} \begin{pmatrix} 1 + e^{j\Delta\omega} & 1 - e^{j\Delta\omega} \\ (1 - 2\mu)(1 - e^{j\Delta\omega}) & (1 - 2\mu)(1 + e^{j\Delta\omega}) \end{pmatrix} \end{aligned} \quad (24)$$

respect to the weight error vector $\mathbf{v}(n)$ can be obtained by substituting $\mathbf{w}(n) = \mathbf{w}_o + \mathbf{v}(n)$ into (6). Now the learning curve can be rewritten as

$$J(n) = J_{\min} + \mathbf{v}^H(n)\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{v}(n) + e_o(n)\mathbf{x}^H(n)\mathbf{v}(n) + \mathbf{v}^H(n)\mathbf{x}(n)e_o^*(n). \quad (35)$$

where

$$e_o(n) = d(n) + \mathbf{w}_o^H \mathbf{x}(n) \quad (36)$$

and

$$J_{\min} = |e_o(n)|^2. \quad (37)$$

Here, $e_o(n)$ is the error signal obtained when $\mathbf{w}(n) = \mathbf{w}_o$ and $J_{\min}$ denotes the minimum squared error.

The difference between the instantaneous cost function $J(n)$ and $J_{\min}$ is defined as the instantaneous excess squared error $J_{\text{ex}}(n)$

$$J_{\text{ex}}(n) = J(n) - J_{\min}. \quad (38)$$

Accordingly, an explicit expression of $J_{\text{ex}}(n)$ with respect to the weight-error vector $\mathbf{v}(n)$ can now be written as

$$J_{\text{ex}}(n) = \mathbf{v}^H(n)\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{v}(n) + e_o(n)\mathbf{x}^H(n)\mathbf{v}(n) + \mathbf{v}^H(n)\mathbf{x}(n)e_o^*(n). \quad (39)$$

By studying the learning curve in (35) and the weight-error vector in (31) it is obvious that the decrease rate of the learning curve is related to the eigenvalues given in (30), where eigenvalue λ_i is associated with the i th mode of the learning curve. Accordingly, the convergence performance of the algorithm is strongly influenced by the *step-size* μ and the *frequency separation* between the reference signals $\Delta\omega$. The *sampling rate* also affects the performance, since $\Delta\omega = 2\pi(\Delta f/f_s)$. The smaller the absolute value of λ_i , the faster the decay of the corresponding mode, and the closer to unity the absolute value of λ_i is, the slower the decay of the mode will be. Thus, with proper eigenvalues the algorithm is convergent, resulting in decreasing filter weights errors

$$\lim_{n \rightarrow \infty} \mathbf{v}(n) = \mathbf{0} \quad (40)$$

and in decreasing excess squared error

$$\lim_{n \rightarrow \infty} J_{\text{ex}}(n) = 0. \quad (41)$$

The fastest convergence is obtained when the two eigenvalues are equal $\lambda_1 = \lambda_2$. By setting the square root in (30) to zero the value of μ for which the eigenvalues are equal can be calculated. Accordingly, the optimum positive step-size μ_{opt} for a given frequency separation Δf and sampling rate f_s can be calculated from

$$\mu_{\text{opt}} = -\tan^2\left(\frac{\Delta\omega}{2}\right) + \sqrt{\tan^2\left(\frac{\Delta\omega}{2}\right)\left(1 + \tan^2\left(\frac{\Delta\omega}{2}\right)\right)}. \quad (42)$$

Fig. 4. Eigenvalue plot for $\Delta f = 0.5$ Hz and $\mu = 0.1 - 0.9$.

III. COMPUTER EXPERIMENT

In order to evaluate the performance of the complex twin-reference algorithm, computer simulations were performed using different frequency separations Δf between the reference signals and different step-size parameters μ . The frequency of reference signal $x_1(n)$ see (Fig. 2) was fixed at 100 Hz, while the reference signal $x_2(n)$ was assigned a frequency in the range of 100–200 Hz ($0 < \Delta f \leq 100$ Hz). The sampling rate was $f_s = 1$ kHz. The desired signal $d(n)$ —the noise signal—consisted of two sinusoidal signals with exactly the same frequencies as the two reference signals and additive white noise $w(n)$ with the power of $E[w^2(n)] = \sigma_w^2 = 10^{-10}$ (−100 dB), $d(n) = H(\omega_1)e^{j\omega_1 n} + H(\omega_2)e^{j\omega_2 n} + w(n)$. Small frequency separations between the reference signals correspond to two sources running at almost the same rotational speeds, while large separations correspond to sources running at completely different rotational speeds, or harmonic frequencies generated by a single source. Throughout the simulations, the initial filter weight vector $\mathbf{w}(0)$ was zero.

IV. RESULTS

For small frequency separations, Δf , the eigenvalues of the matrix $\mathbf{B}$ are approximately real valued $\lambda_1 \approx (1 - 2\mu)$ and $\lambda_2 \approx 1$; see (30). Thus, one of the eigenvalues is very closely related to the step-size, while the other is approximately unity for all step-size values. An eigenvalue plot is shown in Fig. 4 for the frequency separation $\Delta f = 0.5$ Hz and the step-size in the range $\mu = 0.1 - 0.9$. The corresponding learning curves, (35), of the twin-reference algorithm are shown in Fig. 5, and the normalized weight error versus time, $|w_i(n) - w_{o,i}|^2 / |w_{o,i}|^2$ $i = 1, 2$, for some of the step-size parameters are shown in Fig. 6. The adaptation is turned on after ten samples.

Note that when the magnitude of both eigenvalues is less than unity, the rate of convergence increases as the magnitude decreases. As is clearly shown, small magnitude eigenvalues result in a rapidly decaying initial mode, and the closer to zero the value is, the faster the mode decays, reaching its maximum for $\lambda_1 = 0$ ($\mu = 0.5$). For positive real parts of the eigenvalues

Fig. 9. The normalized weight errors versus time for $\Delta f = 100$ Hz. Dashed lines: $|v_1(n)|/|w_{0,1}|$ and solid lines: $|v_2(n)|/|w_{0,2}|$.

Fig. 10. Change of the absolute value of the eigenvalues versus frequency separation between the two reference frequencies.

when the two eigenvalues are equal $\lambda_1 = \lambda_2$, i.e., when the two eigenvalues reach each other on the arc, see Fig. 7. The convergence performance using the calculated optimum step-size $\mu_{\text{opt}} = 0.2361$ for $\Delta f = 100$ Hz is shown in Fig. 8. However, by further increasing the step-size, the magnitude of one of the eigenvalues approaches unity, which slows down the convergence speed, and the rippled behavior disappears as a result of different magnitudes of eigenvalues. An additional increase in the step-size causes convergence behavior with initial decreasing oscillations (underdamped), which then change to monotonic slowly decreasing modes or steady-state levels. In cases of large Δf the $J_{\text{ex}}(n)$ increases with an increasing step-size [13], [14]. On the other hand, by using small step-size values ($\mu < 0.1$) to reduce the excess square error, a very slow rate of convergence and rippled convergence behavior is obtained as a result of eigenvalues with magnitudes quite close to unity. This is clearly shown in Fig. 10, which illustrates the

Fig. 11. Attenuation obtained after 1000 samples for different Δf and μ . Note that the attenuation limit is set at 100 dB.

Fig. 12. Learning curves for $\Delta f = 10$ Hz and $\mu = 0.1 - 0.9$. Dashed lines: $0.1 \leq \mu \leq 0.4$, dashdotted line: $\mu = 0.5$, and solid lines: $0.6 \leq \mu \leq 0.9$.

change in the magnitude of the eigenvalues for different frequency separations and step-size parameters.

A summary of the noise attenuation obtained after 1000 sample iterations (1 s) versus Δf and using different μ is shown in Fig. 11. High noise reduction is obtained for frequency separations up to a few hertz using large step-size parameters, while the reverse is true for large separations. In the transition region at approximately $\Delta f = 10$ Hz, the system exhibits degraded performance and significantly reduced attenuation. The convergence performance for $\Delta f = 10$ Hz is shown in Fig. 12. Here a slowly decaying mode is obtained for $\mu < 0.5$ due to the fact that the magnitude of the eigenvalue close to unity has been reduced; compare the results presented in Fig. 5. However, if the frequency separation Δf is further increased, slowly decaying modes are also obtained for $\mu > 0.5$; consider the convergence performance presented in Fig. 8.

Sometimes the noise is composed of harmonic components generated from sources running at the same or almost identical

- [14] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [15] D. J. Rossetti, M. R. Jolly, and S. C. Southward, "Control effort weighting in feedforward adaptive control systems," *J. Acoust. Soc. Amer.*, vol. 99, no. 5, pp. 2955–2964, May 1996.
- [16] S. J. Elliott, C. C. Boucher, and P. A. Nelson, "The behavior of a multiple channel active control system," *IEEE Trans. Signal Processing*, vol. 40, pp. 1041–1052, May 1992.
- [17] G. Strang, *Linear Algebra and Its Applications*. San Diego, CA: Harcourt Brace Jovanovich, 1986.
- [18] S. Johansson, "Active control of propeller-induced noise in aircraft—algorithms and methods," Ph.D. dissertation, Dept. Telecommun. Signal Process., Blekinge Inst. Technol., Ronneby, Sweden, Dec. 2000.
- [19] S. Johansson, I. Claesson, S. Nordebo, and P. Sjösten, "Evaluation of multiple reference active noise control algorithms on Dornier 328 aircraft data," *IEEE Trans. Speech Audio Processing*, vol. 7, pp. 473–477, July 1999.
- [20] I. Claesson, S. Nordholm, and P. Eriksson, "Noise canceling convergence rates for the LMS algorithm," *Mech. Syst. Signal Process.*, vol. 5, no. 5, pp. 375–388, 1991.

Sven Johansson was born in Sweden in 1968. He received the M.Sc. degree in electrical engineering in 1994 from the University of Lund, Lund, Sweden. In 2000, he received the Ph.D. degree in applied signal processing from the Blekinge Institute of Technology, Ronneby, Sweden, with the dissertation "Active control of propeller induced noise in aircraft-algorithms and methods."

He has been working in the active noise and vibration control group at Blekinge Institute of Technology since 1995, and, since 2001, he has been a

Senior Lecturer. His research interests are in active noise and vibration control and adaptive signal processing.

Sven Nordebo received the M.Sc. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1989, and the Ph.D. degree in electrical engineering from Luleå University of Technology, Luleå, Sweden, in 1995.

From 1996 to 2001, he was a Senior Lecturer with the Blekinge Institute of Technology, Ronneby, Sweden. During 1997–1998, he was with the Australian Telecommunications Research Institute (ATRI), Curtin University of Technology, Perth, Australia, working as a Postdoctoral Research

Fellow. He was appointed Docent in signal processing in 1999. Since 2002, he has been Professor of Signal Processing with the School of Mathematics and Systems Engineering, Växjö University, Växjö, Sweden. His research interests are in array signal processing, wireless communications, adaptive antennas, smart antennas, antennas and propagation, microwave engineering, adaptive signal processing, acoustic signal processing, detection and estimation, digital filter design, and optimization.

Ingvar Claesson was born in Sweden in 1957. He received the M.Sc. degree in 1980 and the Ph.D. degree in 1986, both in electrical engineering from the University of Lund, Lund, Sweden.

Since 1990, he has been building the Telecommunication and Signal Processing Department at the new Blekinge Institute of Technology, where he is Head of research and Chair in applied signal processing. His interests are in acoustic signal processing, filter design, adaptive filtering, and noise cancellation.