Laguerre-domain adaptive filters

den Brinker, A.C.

Published in:
IEEE Transactions on Signal Processing

DOI:
10.1109/78.285660

Published: 01/01/1994

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
to be essentially independent of the influence of phase error. A similar effect may be seen in the time domain results presented in [6] for a sampling rate four times the input frequency (\( \gamma = 90^\circ \)). Figs. 2 and 3 also depict the maximum stable value of convergence coefficient for time delays of one sample and one cycle. Observe, especially in Fig. 3, that the characteristics of the relationship between the phase error and maximum stable convergence coefficient have changed. In addition, with the time delays, the peaks near the \( \pm 90^\circ \) bound are removed, and so the shape of the curve approximates the cosine function shape of the complex algorithm, as suggested in [6]. However, the curve is still not, in general, symmetric about the \( \phi = 0^\circ \) point, the exception being at a sampling rate four times the input frequency (\( \gamma = 90^\circ \)) used in [6].

Although these convergence characteristics of the time domain-filtered x LMS algorithm with a cancellation loop transfer function phase estimation error are somewhat interesting, they are also somewhat discouraging from the viewpoint of the stated objective of the analysis. In fact, it is almost impossible to provide a more quantitative assessment of the effect of transfer function phase estimation error beyond stating that the tolerable bounds of this error are \( |\phi| < 90^\circ \).

IV. CONCLUSIONS

Errors in the estimation of the cancellation path transfer function for active noise and vibration control systems implementing the filtered-x LMS algorithm will have an influence on the stability of the algorithm. Errors in the estimation of the magnitude of this quantity will alter the maximum stable value of convergence coefficient through an inverse proportional relationship. It can further be said that it is possible for the algorithm to be made stable, provided the error in the estimation of the phase of the transfer function does not exceed \( \pm 90^\circ \). The effect that a phase error has on algorithm stability between these bounds is more difficult to predict, owing to the alteration of the error surface as seen by the algorithm induced by the phase errors. The effect will normally not be symmetric about the \( 0^\circ \) phase error point and may, in fact, cause the stability of the algorithm to increase for some values of error. This is in contrast to the case where the complex algorithm is used, where the maximum stable value of convergence coefficient is simply reduced by a factor proportional to the cosine of the phase error.

REFERENCES

We assume a linear regression model where the filter output $y(n)$ is written as

$$y(n) = \sum_{m=0}^{M} w_m u_m(n).$$

(2)

The signals $u_m(n)$ are internal signals of the adaptive filter $F$, and there are $M+1$ weights indexed 0 to $M$. The signals $u_m(n)$ are derived from $x$ by some linear causal filter operation. In essence the adaptive filter $F$ is a filter bank, not necessarily the usual tapped-delay-line.

Taking the derivatives of this optimization criterion $J(n)$ with respect to the filter weights $w_m$, setting these to zero and using (2) gives

$$\sum_{m=0}^{M} w_m \sum_{k=-\infty}^{n} u_m(k)u_m(n-k) = \sum_{k=-\infty}^{n} r(k)u_m(k)\delta^{n-k}.$$  

(3)

This equation holds for all $m$ in the optimal situation. In (3), we recognize a number of local cross-correlations as introduced in [1], [2]. These can be calculated from the Laguerre spectrum of the windowed signals $\hat{r}(n; k) = r(k)\delta^{n-k}$ and $\hat{u}_m(n; k) = u_m(n)\delta^{n-k}$, where $\delta^k = 1$ and $k \leq n$. Therefore, we write the windowed signals as a Laguerre series

$$\hat{r}(n; k) = r(k)\delta^{n-k} = \sum_{i=0}^{\infty} g_i(n)\phi_i(n; n-k),$$

(4)

$$\hat{u}_m(n; k) = u_m(n)\delta^{n-k} = \sum_{i=0}^{\infty} h_i(n)\phi_i(n; n-k).$$

(5) 

$k \leq n$, where the Laguerre functions are given by ($Z^{-1}$ denotes the inverse $z$-transformation)

$$\phi_i(n; z) = \frac{1}{Z^{-1}}\left\{\sqrt{1 - \xi^2 z(1 - z)'}\left(\frac{z}{z-1}\right)^{i+1}\right\}.$$  

The Laguerre coefficients can be determined by a convolution $g_i(n) = r(n) * d_i(n)$ and $h_i(n) = u_m(n) * d_i(n)$, where $d_i$ is the $i$th decomposition filter

$$d_i(k) = \phi_i(n; k)\xi^k = \frac{Z^{-1}\left\{\sqrt{1 - \xi^2 z(1 - z)'}(z-\theta)^{i+1}\right\}}{Z^{-1}\{D_i(z)\}.$$

(7)

Substituting (4) and (5) in (3) yields [2]

$$\sum_{m=0}^{M} w_m \left\{\sum_{i=0}^{\infty} h_i(n)h_i(n)\right\} = \sum_{i=0}^{\infty} g_i(n)h_i(n).$$

(8)

In matrix notation, we have $HH^T w = Hg$, where $w = [w_0, w_1, \ldots, w_M]^T$, a vector containing the filter weights $g = [g_0, g_1, \ldots, g_M]^T$, the Laguerre spectrum of $\hat{r}$, and $H$ is a matrix containing rowwise the Laguerre spectra of the windowed internal filter signals $u_m (0 \leq m \leq M)$

$$H = \begin{bmatrix}
o_h 0_h & 0_h & 0_h & \cdots \\
o_h & 1_h & 1_h & 1_h \\
\vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots \\
0_h & 0_h & 0_h & 0_h \\
1_h & 1_h & 1_h & 1_h \\
\vdots & \vdots & \vdots & \vdots \\
M_h & M_h & M_h & M_h \\
\end{bmatrix}.$$  

(9)

The relation $HH^T w = Hg$ is the normal equation for the coefficients of the adaptive filter.

The derived normal equation is equal to a deterministic normal equation for a tapped delay line (cf. [3]). The matrix $HH^T$ contains the coefficients $c_0$ (see [1] and [2]) of the local cross-correlation functions of $w_m$ and $u_m$, whereas the vector $Hg$ contains the coefficients $c_0$ of the local cross-correlation functions of $u_m$ and $r$.

---

### III. THE FILTER BANK

So far, our analysis and signal descriptions only gave rise to more computational complexity since all signals had to be windowed and transformed to the Laguerre domain. However, we still have complete freedom in the choice of the filter bank. We will make use of this freedom by taking the transfer functions $F_m(z)$ such that both the filtering of $x$ and the decomposition of the signals $u_m$ are performed within the adaptive filter.

To derive the Laguerre coefficient $m_h(n)$ the input signal $x(n)$ has to be filtered by $F_m(z)D_m(z)$. Consider the following choice for $F_m(z)$. We take the ratio of successive filters $F_m$ and $F_{m-1}$ equal to the ratio of successive decomposition filters, so

$$F_m(z) = \left[\frac{\sqrt{\theta(1-z)}}{z-\theta}\right]^m F_0(z)$$

(10)

where $F_0(z)$ is an arbitrary system function. We assume an infinite number of filters, $M = \infty$. Consequently, we have $F_m(z)D_m(z) = [F_{m+1}(z) + \sqrt{\theta}F_{m+1}(z)]/\sqrt{\theta}$, and thus

$$m_{h}(n) = \frac{1}{\sqrt{1-\theta}}[u_{m+1}(n) + \sqrt{\theta}u_{m+1}(n)].$$

(11)

Equation (11) states that the $i$th Laguerre coefficient of the signal $u_m$ is equal to two consecutive signals in the filter bank itself. Since $m_h$ is dependent on the sum $m+i$ and not on $m$ and $i$ separately, the matrix $H$ assumes the form of a Hankel matrix, and therefore, $H^T = H$.

A special case occurs by taking $F_0(z) = \sqrt{\theta}N^\theta(z - \theta)$. The adaptive filter then contains a filter bank equal to the decomposition filter bank. Our goal is to be able to describe, or at least approximate, arbitrary linear causal operators by the proposed adaptive filter. The question is whether this is possible.

First, consider the $i$th system function $F_i(z)$. For $F_i(z) = D_i(z)$, it can be easily verified that $F_i(z)$ can be written as a finite sum of $\Phi_j$

$$F_i(z) = \frac{1}{\sqrt{1+\theta}}\left(\frac{\sqrt{\theta}}{1+\theta}\right)^i \sum_{j=0}^{i} \left(\begin{array}{c} i \\
-j \end{array}\right) (-1)^{i-j} \Phi_j(\theta; z),$$

(12)

where $\Phi_j$ is the $j$-transform of $\Phi_j$.

A Laguerre filter $L_M(\theta; z)$ of order $M+1$ ($M = 0, 1, \ldots$) is defined by

$$L_M(\theta; z) = \sum_{i=0}^{M} \hat{w}_i \Phi_i(\theta; z)$$

(13)

where $\hat{w}_i$ are unspecified constants. From (12), we infer that a limited number $M+1$ of partial system functions $F_i(z)$ and arbitrary weights $w_m$ encompasses the same set of functions as a Laguerre filter of order $M+1$. The Laguerre filter for $M = \infty$ describes the system functions of all square-summable time-invariant causal impulse responses [3], and consequently, so does the proposed adaptive filter for this special choice of $F_0$. Any other choice for $F_0$ can conceptually be split into a cascade of two filters with one of them having a system function $D_0$. The other filter can then be considered as prefiltering the input signal $x$. 

---

[1] and [2] refer to the original sources of the equations and statements used in the text.
Equation for the bias proposed adaptive filter. For this situation three statistical properties are derived. Linearly combined in an output signal banks have input signal filter bank is combined with variable weights and constitutes the length M. Note that (15) states that we do not need to perform a prefiltering operation on $A$. As.

From (18) and $\Phi(n) = \Phi(n)H(n)$, we can construct the updating equation for the bias $b(n) = b(n-1) + b(n)\alpha(n)$. (20)

Substitution of (17) in (20) and using $k_p(n) = P(n)\Phi(n)$ and (14) gives the following time-varying difference equation for $b(n)$:

$$b(n) = \theta P(n)H(n-1)H^T(n-1) b(n-1) = P(n)\Phi(n) r_2(n).$$

This equation is linear in $r_2(n)$ provided that $r_2$ and $x$ are statistically independent. Since $\Phi(n)$, $H(n-1)H^T(n-1)$, and $P(n)$ are proportional, proportional to the square, and inversely proportional to the square of the amplitude of $x(n)$, respectively, the bias is inversely proportional to the amplitude of the input signal $x(n)$.

Suppose now that we start at time $n = 1$ with $b(0) = b_0$ and matrix $H(0)H^T(0)$. (The initial matrix $H(0)H^T(0)$ should be taken nonsingular. In RLS algorithms, one commonly takes $H(0)H^T(0) = \delta I$, where $I$ is the identity matrix, and $\delta$ is some small positive constant [4].) Solving the difference equation (21) gives

$$b(n) = \theta^n P(n)H(0)H^{T}(0)b_0 + P(n) \sum_{k=1}^{n} \theta^{n-k} \Phi(k) r_2(k).$$

Suppose now that $x(n)$ is a deterministic signal and thus that $y(k)$ and $P(n)$ are deterministic quantities. Suppose furthermore that $r_2$ is a stochastic signal derived from a zero-mean process. Consequently, the bias $b(n)$ is a stochastic signal and its expectation $E[b(n)]$ is

$$E[b(n)] = \theta^n P(n)H(0)H^{T}(0)b_0.$$ (22)

The foregoing equation expresses that if $P(n)$ increases slower than exponentially, the expectation $E[b(n)]$ tends to zero for $n \rightarrow \infty$. In fact, (23) expresses the identifiability of the unknown process given a certain input signal $x(n)$. One can always construct signals $x(n)$ such that $\Phi(n)$ does not tend to $\Phi(n)$ for $n \rightarrow \infty$.

Given a proper input signal such that the unknown process is identifiable, the weights $\Phi$ will fluctuate around the optimal solution $\Phi$ after a first learning period. The difference between the actual and optimal solution is again denoted as $b(n)$ and is called the weight-error vector. We calculate the weight-error correlation matrix $E[b(n)b^T(n)]$.

We assume that the adaptive filter is working in its steady state and that the signals $x(n)$ and $r_2(n)$ extend from $n = -\infty$ to $+\infty$. We assume that $x(n)$ is a deterministic signal and that $r_2(n)$ is a wide-sense stationary, zero-mean process with variance $\sigma^2$. Repeated application of (21) gives

$$E[b(n)b^T(n)] = \sigma^2 P(n) \left[ \sum_{k=-\infty}^{n} \theta^{n-k} \Phi(k) \Phi^T(k) \right] P(n).$$ (24)

The covariance matrix is proportional to the variance of $r_2(n)$. Since $b(n)$ is inversely proportional to the amplitude of the input signal $x$, the covariance matrix is inversely proportional to the square of this amplitude.

As a next item, we consider the mean-squared innovation $J_2(n)$ defined as $J_2(n) = -E[(\alpha \Phi)^2]$. We again consider the steady state situation, and we assume that $r_2$ is a wide-sense stationary, zero-mean, white-noise process with variance $\sigma^2$. Since $r_2$ is assumed to be a white-noise process, $r_2(n)$ and $b(n-1)$ are statistically independent. By using $\alpha(n) = r_2(n) - b^T(n-1)\Phi(n)$ and (24) we find

$$J_2(n) = \sigma^2 + \sigma^2 E[b^T(n)P(n)]$$

$$= \left[ \sum_{k=-\infty}^{n} \theta^{n-k} \Phi(k) \Phi^T(k) \right] P(n).$$ (25)

From the foregoing relation, we conclude that the mean-squared innovation is proportional to the variance of $r_2$ and independent of the amplitude of the input signal.

We have simulated the adaptive filter in the situation shown in Fig. 2. The signals $x(n)$ and $r_2(n)$ were derived from two
The normalized mean-squared innovation $L^2$ are equal (Section IV), the differences between these not actually use the information provided by the local analysis in the present study. As a consequence of this and the fact that the model described as Laguerre series.

The second degree of freedom is the parameter $\theta$, i.e., a perfect performance of the algorithm. The conclusion is that there are two degrees of freedom in the Laguerre domain adaptive filter. In the case that the unknown process fluctuating parameters yields larger fluctuations in the innovation. For $\theta \to 1$, i.e., operating with a long memory, we have $J'(\infty) \to \sigma^2$, i.e., the method based on the Remez exchange algorithm. Design of differentiators based on minimization of the relative mean-square error is also carried out. Finally, our method is extended to the design of frequency selective differentiators.

Fig. 3. Normalized mean-squared innovation $J'(\infty)/\sigma^2$ as a function of the number of weights in the adaptive filter for $\theta = 0.8, 0.9,$ and 0.95 (indicated by crosses, circles and stars, respectively). The data are calculated from the average over 200 independent runs of the adaptive mechanism.

The essence a scale factor determining how long the memory of the filter is can already suppress these parts of the input signal by prefiltering the input data $x$ and subsequently using a truncated Laguerre filter. In this way, one can shape the spectrum of the input signal on the basis of a priori information in order to optimally exploit the properties of a truncated LDAF with a minimal number of coefficients, e.g., if it is known that we are approximating a process with a high-frequency falloff, one can already suppress these parts of the input signal by $F_0$.

The second degree of freedom is the parameter $\theta$, which is in essence a scale factor determining how long the memory of the filter is taken. Although each stable linear causal time-invariant operator can be approximated by a Laguerre filter, the discount factor $\theta$ must be carefully chosen in order to obtain a minimal number of meaningful coefficients in the adaptive filter. In the case that the unknown process has a broadband spectrum, one can resort to multiscale adaptive filtering [12].

Current research includes further software simulations of the LDAF. Hardware implementation of the proposed adaptive filter is considered as well.

VI. DISCUSSION

We developed an adaptive filter starting by considering windowed versions of the input and reference signals. Windowing of the data is required since in adaptive filtering it is assumed that the correlation between input and reference signal is (slowly) time varying. An exponential window was chosen, and consequently, the signals were described as Laguerre series.

Our starting point differs from the usual approach in system identification using Laguerre functions [5]-[12]. There one commonly starts by a description of the model by a Laguerre filter, whereas we started by transforming the windowed signals to the Laguerre domain. This provides a local analysis of the input signal. We did not actually use the information provided by the local analysis in the present study. As a consequence of this and the fact that the model sets of the introduced adaptive filter and the more commonly used Laguerre filter are equal (Section IV), the differences between these two approaches should be sought in aspects of implementation.

There are two degrees of freedom in the Laguerre domain adaptive filter. The first one is the initial filtering stage $F_0$. Taking $F_0(z) = z/\sqrt{1 - \theta}(z - \theta)$ is equal to taking a truncated Laguerre filter. Other choices for $F_0$ can be interpreted as prefiltering the input data $x$ and subsequently using a truncated Laguerre filter. In this way, one can shape the spectrum of the input signal on the basis of a priori information in order to optimally exploit the properties of a truncated LDAF with a minimal number of coefficients, e.g., if it is known that we are approximating a process with a high-frequency falloff, one can already suppress these parts of the input signal by $F_0$.

The second degree of freedom is the parameter $\theta$, which is in essence a scale factor determining how long the memory of the filter is taken. Although each stable linear causal time-invariant operator can be approximated by a Laguerre filter, the discount factor $\theta$ must be carefully chosen in order to obtain a minimal number of meaningful coefficients in the adaptive filter. In the case that the unknown process has a broadband spectrum, one can resort to multiscale adaptive filtering [12].

Current research includes further software simulations of the LDAF. Hardware implementation of the proposed adaptive filter is considered as well.

REFERENCES


Leonie—A method is described that can be used to design non-recursive linear-phase higher order differentiators that can perform differentiation over any frequency range. The method is based on formulating the absolute mean-square error between the amplitude responses of the practical and ideal differentiator as a quadratic function. The coefficients of the differentiators are obtained by solving a set of linear equations. This method leads to a lower mean-square error and is computationally more efficient than both the eigenfilter method and the method based on the Remes exchange algorithm. Design of differentiators based on minimization of the relative mean-square error is also carried out. Finally, our method is extended to the design of frequency selective higher order differentiators.

Abstract—A method is described that can be used to design non-recursive linear-phase higher order differentiators that can perform differentiation over any frequency range. The method is based on formulating the absolute mean-square error between the amplitude responses of the practical and ideal differentiator as a quadratic function. The coefficients of the differentiators are obtained by solving a set of linear equations. This method leads to a lower mean-square error and is computationally more efficient than both the eigenfilter method and the method based on the Remes exchange algorithm. Design of differentiators based on minimization of the relative mean-square error is also carried out. Finally, our method is extended to the design of frequency selective higher order differentiators.

Manuscript received April 24, 1992; revised March 17, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. Tamal Bose.

S. Sunder is with the Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada.

R. P. Ramachandran is with the Caip Center, Department of Electrical Engineering, Rutgers University, Piscataway, NJ 08855.

IEEE Log Number 9215275.