A NOVEL ADAPTIVE LEAKAGE FACTOR SCHEME FOR ENHANCEMENT OF A VARIABLE TAP-LENGTH LEARNING ALGORITHM

Leilei Li and Jonathon A. Chambers

Advanced Signal Processing Group, Department of Electronic & Electrical Engineering
Loughborough University, Loughborough LE11 3TU, UK.
(Email: l.li2@lboro.ac.uk, j.a.chambers@lboro.ac.uk)

ABSTRACT

In this paper a new adaptive leakage factor variable tap-length learning algorithm is proposed. Through analysis the converged difference between the segmented mean square error (MSE) of a filter formed from a number of the initial coefficients of an adaptive filter, and the MSE of the full adaptive filter, is confirmed as a function of the tap-length of the adaptive filter to be monotonically non-increasing. This analysis also provides a systematic way to select the key parameters in the fractional tap-length (FT) learning algorithm, first proposed by Gong and Cowan, to ensure convergence to permit calculation of the true tap-length of the unknown system and motivates the need for adaptation in the leakage factor during learning. A new strategy for adaptation of the leakage factor is therefore developed to satisfy these requirements with both small and large initial tap-length. Simulation results are presented which confirm the advantages of the proposed scheme over the original FT scheme.

Index Terms— Adaptive signal processing, Adaptive filters, Variable tap-length, Adaptive leakage factor

1. INTRODUCTION

As a key parameter, tap-length plays an important role in the design of adaptive filters based on the LMS algorithm, which has been utilized in a wide range of applications [1, 2, 3], as a consequence of its simplicity and robustness. Since the concept of variable tap-length in adaptive filters was initially proposed in [4], many related results [5]-[10] have been reported. The algorithm in [4] compares the current MSE of a deficient tap-length adaptive filter to the pre-estimated minimum MSE for a specific tap-length to improve convergence rate. In [5], a variable tap-length algorithm with a precalculated time constant is proposed. However, both algorithms were initial attempts at enhancing the convergence behavior of the MSE in an environment where the tap-length of the system is known. In time-varying scenarios, where the system to be identified has changing length, Riera-Palou et al. presented an algorithm which relies on the concept of partitioned segments, the number of which must be carefully chosen dependent on application [6]. During learning in variable tap-length adaptive filters, when the adaptation noise is low, the "wandering" problem is encountered, that is, the tap-length wanders within a range that is always greater than the optimum tap-length, [7]; an issue we address in this work through careful selection of the leakage factor for the FT algorithm. The algorithms in [8] were presented to make the estimated tap-length converge to the optimum tap-length in the mean. However, both algorithms suffer from slow tap-length convergence in certain scenarios.

Gong and Cowan [9] introduce a low-complexity FT algorithm based on instantaneous errors, which obtains improved convergence properties. In [10], a variable tap-length natural gradient blind equalization algorithm based on the FT algorithm is proposed, which gives a good compromise between steady-state performance and computational complexity. In this paper, an analysis of the converged fractional tap-length function is introduced, which provides the motivation for the novel scheme based on the FT algorithm. To facilitate this analysis, we assume that the final element of the unknown filter is significantly different from zero. In the proposed algorithm, a variable leakage factor based on the squared smoothed error is used to improve the convergence behavior of the fractional tap-length as compared to the original FT algorithm.

The remainder of this paper is organized as follows. The analysis of the fractional tap-length function is described in section 2. The proposed algorithm and its motivation are introduced in section 3. A simulation that confirms the analysis and advantage of the presented algorithm as compared with the original FT algorithm is given in section 4. Section 5 offers conclusions.

2. BACKGROUND

In the case of system identification, a reference data measurement $d(i)$ is observed from the model:

$$d(i) = \mathbf{u}_i \mathbf{w}^o + v(i) \tag{1}$$

where $\mathbf{u}_i$ is a $1 \times M$ row input vector, $v(i)$ indicates the zero-mean noise, $\mathbf{w}^o$ denotes an unknown $M \times 1$ column tap vector that we intend to estimate and i denotes the iteration index. All quantities are assumed to be real valued for convenience of development but extension to complex values is straightforward. It is well known in [1, 2, 3] that the LMS algorithm computes $\mathbf{w}_i$ via

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{u}_i^{\mathrm{T}} e(i) \tag{2}$$

where μ is the step-size of the tap weight update, $[\cdot]^{\mathrm{T}}$ denotes the vector transpose operator and $e(i)$ is the difference between the reference data and the output of the adaptive filter, defined as,

$$e(i) = d(i) - \mathbf{u}_i \mathbf{w}_{i-1} \tag{3}$$

When the tap-length of the adaptive filter is unknown, variable tap-length schemes can be employed. But, as mentioned in [8], MSE-based variable tap-length schemes can lead to underestimation of the tap-length in certain applications. In order to overcome this problem, a parameter Δ, which is a positive integer and sufficiently large to avoid sub-optimum tap-lengths, has been introduced in [8]. We suppose that N is the estimated length of the variable tap-length adaptive filter. Thus, the segmented error is constructed, as in [9] by,

$$e_{N-\Delta}^{(N)}(i) = d(i) - \mathbf{u}_i(1:N-\Delta)\mathbf{w}_{i-1}(1:N-\Delta) \tag{4}$$

where $\mathbf{u}_i(1:N-\Delta)$ and $\mathbf{w}_{i-1}(1:N-\Delta)$ are vectors consisting of the first $N-\Delta$ coefficients of $\mathbf{u}_i$ and $\mathbf{w}_{i-1}$ respectively. The steady-state segmented MSE is further defined as $J_{N-\Delta}^{(N)} = E[(e_{N-\Delta}^{(N)}(\infty))^2]$ where $E[\cdot]$ denotes statistical expectation. The cost function to search for the optimum tap-length is described as:

$$\mathbf{min}\{N | J_{N-\Delta}^{(N)} - J_N^{(N)} \leq \varepsilon\} \tag{5}$$

where $\mathbf{min}(\cdot)$ denotes the minimum value and ε is a predetermined small positive value selected according to the requirements of the adaptive filter. For expressing the analysis clearly, when the initial tap-length of the adaptive filter is smaller than the converged tap-length of the adaptive filter, we name it as increasing tap-length (ITL) estimation; when the initial tap-length is larger than the converged tap-length, we name it as decreasing tap-length (DTL) estimation.

3. ANALYSIS OF FRACTIONAL TAP-LENGTH FUNCTION

To simplify the analysis, we make three assumptions:

A1. The input $u(i)$ is a zero-mean stationary white signal with variance σ_u^2 and the input vector $\mathbf{u}_i$ is uncorrelated with $\mathbf{u}_j$ for $i \neq j$.

A2. The background noise $v(i)$ is also a zero-mean stationary white signal with variance σ_v^2 and uncorrelated with $\mathbf{v}_j$ for $i \neq j$ and $u(j)$ for all j.

A3. The final optimum weight vector coefficient is sufficiently different form zero, $w^o(M) \neq 0$.

In the variable tap-length algorithm, the fractional tap-length function is defined as, [9],

$$Lf(i+1) = Lf(i) - \alpha + \beta \cdot \eta_{Ln(i)}(i) \tag{6}$$

where α is the leakage factor, β is the step-size of the fractional tap-length function and $\eta_{Ln(i)}(i)$ is the difference between the segmented MSE and the full MSE, given by,

$$\eta_{Ln(i)}(i) = (e_{Ln(i)-\Delta}^{(Ln(i))})^2 - (e_{Ln(i)}^{(Ln(i))})^2 \tag{7}$$

where $Ln(i)$ denotes the rounded value of $Lf(i)$ and is constrained to be not less than $L\min$, where $L\min > \Delta$, since any tap-length below it cannot be calculated in (7).

After the initial convergence, assuming close proximity to the optimal adaptive filter length and small misadjustment, the tap-length $Ln(i)$ should be designed to vary within $[M+\Delta-1,\ M+\Delta]$, from which the true tap-length of the unknown system M can be found. When $Ln(i)$ equals to $M+\Delta-1$, the fractional tap-length function should be increased towards $M+\Delta$, namely $\eta_{M+\Delta-1}(i) > \frac{\alpha}{\beta}$. On the other hand, when $Ln(i) = M+\Delta$, the fractional tap-length should be decreased, namely $\eta_{M+\Delta-1}(i) < \frac{\alpha}{\beta}$. Therefore, the problem becomes how to select the parameters α and β in equation (6) to satisfy the above requirements. In order to make the appropriate selection of $\frac{\alpha}{\beta}$, let us evaluate the performance of $\eta_{M+\Delta-1}(i)$, given by,

$$\eta_{M+\Delta-1}(i) = (e_{M-1}^{(M+\Delta-1)})^2 - (e_{M+\Delta-1}^{(M+\Delta-1)})^2 \tag{8}$$

Taking statistical expectation of both sides, we obtain

$$E[\eta_{M+\Delta-1}(i)] = E[(e_{M-1}^{(M+\Delta-1)})^2] - E[(e_{M+\Delta-1}^{(M+\Delta-1)})^2] \tag{9}$$

Due to the assumption *A2*, this becomes

$$E[\eta_{M+\Delta-1}(i)] = A + \sigma_u^2 w^o(M)^2 - B \tag{10}$$

where $w^o(M)$ is the final coefficient of the tap weight in the unknown system, B equals to $E[(\mathbf{u}_i(1:M+\Delta-1)[\mathbf{w}^o(1:M+\Delta-1) - \mathbf{w}_i(1:M+\Delta-1)])^2]$ and A is constructed as,

$$\begin{aligned} A = {} & E[(\mathbf{u}_i(1:M-1)[\mathbf{w}^o(1:M-1) - \mathbf{w}_{i-1}(1:M-1)])^2] \\ & -2w^o(M)E[u_i(M)\mathbf{u}_i(1:M-1)\mathbf{w}_{i-1}(1:M-1)] \\ & +2w^o(M)E[u_i(M)\mathbf{u}_i(1:M-1)]\mathbf{w}^o(1:M-1) \end{aligned} \tag{11}$$

According to the assumption *A1*, $u_i(M)$ is uncorrelated with $\mathbf{u}_i(1:M-1)$ and $\mathbf{w}_{i-1}(1:M-1)$. Thus, the result of the last two items on the right side of equation (11) is zero. In addition, as $i \rightarrow \infty$, $J_{M-1,\ excess}^{(M+\Delta-1)}$ and $J_{M+\Delta-1,\ excess}^{(M+\Delta-1)}$ are used to indicate the converged excess mean square errors (EMSE). Therefore, equation (10) becomes

$$\begin{aligned} E[\eta_{M+\Delta-1}(i)] \quad = \quad & J_{M-1,\ excess}^{(M+\Delta-1)} - J_{M+\Delta-1,\ excess}^{(M+\Delta-1)} \\ & +\sigma_u^2 w^o(M)^2, \quad as \quad i \rightarrow \infty \end{aligned} \tag{12}$$

After the same manipulations as in (9)-(12), the converged $E[\eta_{M+\Delta}(i)]$ is also obtained as,

$$\begin{aligned} E[\eta_{M+\Delta}(i)] &= J^{(M+\Delta)}_{M,\ excess} - J^{(M+\Delta)}_{M+\Delta,\ excess} \\ &= C, \qquad\qquad as \quad i \to \infty \end{aligned} \tag{13}$$

where C is a small negative value and given by $-(\mathbf{u}_i(M+1 : M+\Delta)\mathbf{w}_i(M+1 : M+\Delta))^2$. Therefore, given that the leakage parameter is positive, $E[\eta_{M+\Delta-1}(\infty)]$ should be bigger than zero, namely $w^o(M)^2 > J^{(M+\Delta-1)}_{M-1,\ excess} - J^{(M+\Delta-1)}_{M+\Delta-1,\ excess}$ as required in *A3*. Only when we select the ratio of the optimum parameters $\frac{\alpha}{\beta}$ bigger than zero and smaller than the value of $E[\eta_{M+\Delta-1}(\infty)]$, can the steady-state tap-length of the FT algorithm be used to calculate the true tap-length.

However, in the FT algorithm $Ln(i)$ may converge to a value within $(M,\ M+\Delta]$ when the parameters α and β are not chosen appropriately. Thus, the analysis for this case is developed. The expectation of the MSE difference, $E[\eta_L]$ is given by,

$$E[\eta_L(i)] = E[(e^{(L)}_{L-\Delta})^2] - E[(e^{(L)}_{L})^2] \tag{14}$$

where $M < L \leq M + \Delta$. Applying the same manipulations as in the earlier analysis, we obtain

$$\begin{aligned} E[\eta_L(\infty)] = J^{(L)}_{L-\Delta,\ excess} - J^{(L)}_{L,\ excess} \\ +(M+\Delta-L)||\mathbf{w}^o(L-\Delta+1 : M)||^2 \end{aligned} \tag{15}$$

where $\|\cdot\|^2$ denotes the squared Euclidean norm. The EMSE of the LMS algorithm in [3] is formulated as

$$J_{LMS,\ excess} = \frac{\mu L \sigma_u^2 \sigma_v^2}{2 - \mu L \sigma_u^2} \tag{16}$$

According to the stability condition of the LMS algorithm (see [1]), the value of $\mu L \sigma_u^2$ is chosen significantly smaller than 2. As a result, the EMSE of the LMS algorithm (16) becomes

$$J_{LMS,\ excess} \approx \frac{\mu L \sigma_u^2 \sigma_v^2}{2} \tag{17}$$

In order to faciliate the analysis, we generally suppose that with a small μ the difference between the segmented EMSE and the full length EMSE is trivial compared to $(M + \Delta - L)\sigma_u^2||\mathbf{w}^o(L-\Delta+1 : M)||^2$. Thus, equation (15) is approximated as

$$E[\eta_L(\infty)] \approx (M+\Delta-L)\sigma_u^2||\mathbf{w}^o(L-\Delta+1 : M)||^2 \tag{18}$$

From the above analysis, it is clear that $E[\eta_L(\infty)]$ is a monotonic non-increasing function with respect to the tap-length L, which is within the range $(M,\ M+\Delta]$. From the above analysis, when we choose $\frac{\alpha}{\beta}$ close to zero the converged $Ln(i)$ can be utilized to calculate the true tap-length of the unknown system by subtracting Δ. However, when $\frac{\alpha}{\beta}$ is chosen too small, as observed in [7] the "wandering" problem occurs in the DTL estimation, which results in a very slow convergence rate for the fractional tap-length of the adaptive filter. Careful choice of $\frac{\alpha}{\beta}$ is therefore crucial for successful operation of the FT algorithm, and this motivates the work in section 4.

4. PROPOSED NOVEL ALGORITHM

In this section, we present a novel algorithm with an adaptive leakage factor based on the FT algorithm [9]. Recalling (13), the value of $E[\eta_{M+\Delta}(i)]$ is negative when $Ln(i) = M + \Delta$. Thus, we need the variable leakage factor to be close to zero at the steady-state stage so that $Lf(i)$ is not reduced too much in (6) and the true tap-length M of the unknown system can be calculated, and big enough during the learning stage to avoid the "wandering" problem. To proceed, the smoothed error is obtained from

$$\hat{e}(n) = (1-\rho)\sum_{i=1}^{n} \rho^{n-i} e^{(Ln(i))}_{Ln(i)}(i) \tag{19}$$

where the initial smoothed error is chosen as $\hat{e}(0) = 1$ and ρ $(0 \ll \rho < 1)$ is a forgetting factor that governs the time averaging window to reduce the effect of the distant past and adapt to the current statistics. With the assumption *A2*, the expected performance of the squared smoothed error can then be simplified as,

$$E[(\hat{e}(i))^2] \approx \frac{(1-\rho)}{(1+\rho)}\sigma_v^2, \qquad as \quad i \to \infty \tag{20}$$

where we use the assumption that at steady state the error signal $e^{(Ln(i))}_{Ln(i)}(i)$ is approximately equal to the noise signal $v(i)$.

The proposed adaptive leakage factor algorithm is formulated as follows:

$$\hat{e}(i+1) = \rho\hat{e}(i) + (1-\rho)e^{(Ln(i+1))}_{Ln(i+1)}(i+1) \tag{21}$$

$$\bar{\alpha}(i+1) = \frac{(\hat{e}(i+1))^2}{(\hat{e}(i+1))^2 + \delta} \tag{22}$$

$$\bar{\alpha}(i+1) = \mathbf{min}(\bar{\alpha}(i+1),\ \alpha_{\max}) \tag{23}$$

$$Lf(i+1) = Lf(i) - \bar{\alpha}(i+1) + \beta \cdot \eta_{Ln(i)}(i) \tag{24}$$

where δ indicates a positive constant and is related to σ_v^2. According to (22), the converged variable leakage factor approximately equals to $\frac{(1-\rho)\sigma_v^2}{(1+\rho)\delta}$, when we choose $\delta \gg (\hat{e}(i+1))^2$ as $i \to \infty$, which is verified in section 5. The proper selection of ρ and δ enables the variable leakage factor to converge to the desired value.

5. SIMULATION

In this section, we show results of the computer simulation which compare the performances of the original FT algorithm and the novel algorithm. Two unknown system $\mathbf{W}_1$ and $\mathbf{W}_2$ are introduced and the coefficients of both filters are chosen from a zero-mean uniformly distributed random sequence and the tap-lengths are 20 and 10 respectively, to satisfy *A3*, $w^o(M)^2$ is always bigger than 0.01. The input signal $u(i)$ is a zero-mean white-noise Gaussian sequence with $\sigma_u^2 = 1$ and

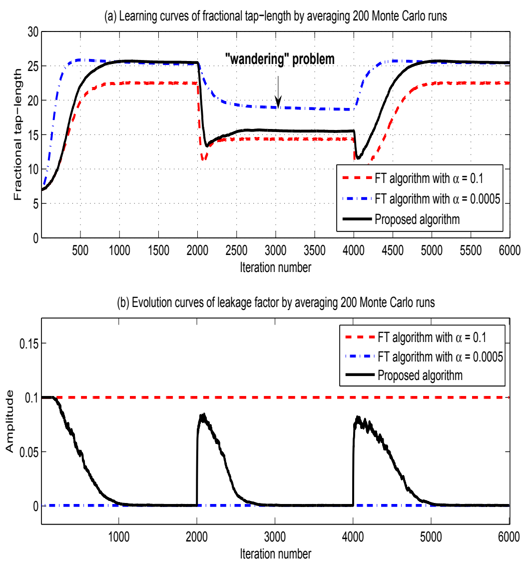

Fig. 1. Learning curves of the FT algorithm and the proposed algorithm by averaging 200 Monte Carlo runs in 30dB SNR: fractional tap-length (a), leakage factor (b).

the background noise is also a white-noise Gaussian sequence with zero mean. We select the parameter $\Delta = 6$ and the initial tap-length $Ln(0)$ equal to the minimum tap-length $L\min$ chosen as $\Delta + 1 = 7$. In this example, the step-size of the adaptive filter is chosen as $\mu = 0.005$, which ensures the stability of LMS (see [3]) and guarantees $w^o(M)^2 > J^{(M+\Delta-1)}_{M-1,\ excess} - J^{(M+\Delta-1)}_{M+\Delta-1,\ excess}$ (see Section 3), and the smoothing factor of error is chosen as $\rho = 0.99$, which results in the converged smoothed MSE of the novel algorithm having approximately the value of $0.005\sigma_v^2$, as in equation (20).

Figures 1(a) and 1(b) show the algorithms in a time varying scenario, where $\mathbf{w}^o = \mathbf{W}_1$ for $i \leq 2000$ or $i > 4000$, and $\mathbf{w}^o = \mathbf{W}_2$ for $2000 < i \leq 4000$, when the SNR is 30dB. And in both algorithms, the step-size β is set as 1. The leakage factors of the FT algorithm are chosen as 0.1 and 0.0005 respectively. For the proposed algorithm, $\alpha_{\max}$ is chosen as 0.1 and $\delta = 0.01$ is used to control the adaptive leakage factor. According to equations (20) and (22), the converged leakage factor can be estimated and the value is approximately 0.0005. As expected, Figure 1(a) shows that the proposed scheme not only avoids "wandering" in high value areas, which the FT algorithm with $\alpha = 0.0005$ encounters during the period $[2001,\ 4000]$, but also obtains an improvement that its fractional tap-length better calculate the true tap-length as compared to the FT algorithm with $\alpha = 0.1$ for example during periods $[1000,\ 2000]$, $[2500,\ 4000]$ and $[5000,\ 6000]$. In addition, it gives a good compromise for the fractional tap-length convergence rate in both the ITL estimation and the DTL estimation. The learning curve in Figure 1(b) shows the fluctuation of the variable leakage factor in the time varying scenario which confirms the tracking ability of the variable leakage factor. Therefore, we find that when the value of the leakage factor of the novel scheme for the true tap-length is located in the range $[0.0005,\ 0.1)$, the proposed algorithm can be utilized to search for the true tap-length without the "wandering" problem.

6. CONCLUSIONS

We have presented a novel adaptive leakage factor variable tap-length learning algorithm based on the FT algorithm together with an analysis of the converged difference between the segmented MSE of a filter formed from a number of initial coefficients of an adaptive filter, and the MSE of the full adaptive filter, which motivates the novel scheme. The simulation results have confirmed the advantages of the presented algorithm over the original FT algorithm in terms of the performance behavior of the fractional tap-length to ensure convergence to permit calculation of the true tap-length in both the ITL estimation and DTL estimation cases. Although in this paper we limited our attention to the uncorrelated white-noise input, we may expect further improvements by extending it to general coloured input.

7. REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.

[2] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, John Wiley & Sons, Inc., Hoboken, NJ, 1998.

[3] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, Inc., Hoboken, NJ, 2003.

[4] Z. Pritzker and A. Feuer, "Variable length stochastic gradient algorithm," *IEEE Trans. Signal Process*, vol. 39, no. 4, pp. 977–1001, Apr. 1991.

[5] Y. K. Won, R. H. Park, J. H. Park, and B. U. Lee, "Variable LMS algorithm using the time constant concept," *IEEE Trans. Consumer Electron*, vol. 40, no. 3, pp. 655–661, 1994.

[6] F. Riera-Palou, J. M. Norsa, and D. J. M. Cruickshank, "Linear equalizsers with dynamic and automatic length selection," *Electron. Lett.*, vol. 37, no. 25, 2001.

[7] Y. Gu, K. Tang, H. Cui, and W. Du, "LMS algorithm with gradient descent filter length," *IEEE Signal Process. Lett.*, vol. 11, no. 3, 2004.

[8] Y. Gong and C. F. N. Cowan, "Structure adaptation of linear MMSE adaptive filters," *Proc. Inst. Elect. Eng.-Vision, Image, Signal Process.*, vol. 151, no. 4, pp. 271–277, Aug. 2004.

[9] Y. Gong and C. F. N. Cowan, "An LMS style variable tap-length algorithm for structure adaptation," *IEEE Tran. Signal Processing*, vol. 53, no. 7, pp. 2400–2407, 2005.

[10] Y. Zhang and J. A. Chambers, "A variable tap-length natural gradient blind deconvolution/equalization algorithm," *Electron. Lett.*, vol. 43, no. 14, 2007.