Loughborough
University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

Steady-State Performance Analysis of a Variable Tap-Length LMS Algorithm

Yonggang Zhang, *Student Member, IEEE*,　Ning Li,
Jonathon A. Chambers, *Senior Member, IEEE*, and
A. H. Sayed, *Fellow, IEEE*

Abstract—A steady-state performance analysis of the fractional tap-length (FT) variable tap-length least mean square (LMS) algorithm is presented in this correspondence. Based on the analysis, a mathematical formulation for the steady-state tap length is obtained. Some general criteria for parameter selection are also given. The analysis and the associated discussions give insight into the performance of the FT algorithm, which may potentially extend its practical applicability. Simulation results support the theoretical analysis and discussions.

Index Terms—Adaptive filters, steady-state performance analysis, variable tap-length LMS algorithm.

I. INTRODUCTION

The least mean square (LMS) adaptive algorithm has been extensively used as a consequence of its simplicity and robustness [1], [2]. In many applications of the LMS algorithm, the tap length of the adaptive filter is kept fixed. However, in certain situations, the tap length of the optimal filter is unknown or variable. According to the analysis in [3] and [4], the mean square error (MSE) of the adaptive filter is likely to increase if the tap length is undermodeled. To avoid such a situation, a sufficiently large filter tap length is needed. However, the computational cost and the excess mean square error (EMSE) of the LMS algorithm will increase if the tap length is too large; thus, a variable tap-length LMS algorithm is needed to find a proper choice of the tap length.

Several variable tap-length LMS algorithms have been proposed in recent years. Among existing variable tap-length LMS algorithms, some are designed to not only establish a suitable steady-state tap length, but also to speed up the convergence rate [3], [5]. These methods are based on the assumption that the unknown optimal filter has an impulse response sequence with an exponentially decaying envelope, which limits their utility. Other methods are more general and are designed to search for the optimal filter tap length at steady-state [6]–[9]; a summary of these works is given in [10]. As analyzed in [10], the fractional tap-length (FT) algorithm is more robust and has lower computational complexity when compared with other methods. A convex combination structure of the FT algorithm has been proposed

in [11] to establish the optimal tap length in high-noise conditions, in which two filters are updated simultaneously with different parameters, so that the overall filter can obtain both a rapid convergence rate from the fast filter and a smooth curve for the steady-state tap length from the slow filter.

Although the FT method has many advantages in comparison to other methods, and can be potentially used in many applications, no steady-state performance analysis has been given in the literature, due to the nonlinearity between the output error and the tap length. Thus, the choice of the algorithm parameters is difficult. In this correspondence, we provide a steady-state performance analysis of the FT algorithm. At first, the fractional tap-length update equation is rewritten, which gives insight into the fractional tap-length evolution process. Based on this equation, a mathematical formulation of the steady-state tap length is then obtained. Some general guidelines for the parameter choice are also given, which highlight the practical applicability of the FT algorithm. As will be confirmed in the simulations, the guidelines for the parameter choice are reasonable and useful.

The remainder of the correspondence is organized as follows. In Section II, we formulate the FT variable tap-length LMS algorithm. The steady-state performance analysis of the FT algorithm and some guidelines for parameter selection are given in Section III. Simulations are performed in Section IV to illustrate the analysis and discussions. Section V concludes the correspondence.

II. FT VARIABLE TAP-LENGTH LMS ALGORITHM

The FT variable tap-length LMS algorithm is designed to find the optimal filter tap length. In agreement with most approaches used to derive algorithms for adaptive filtering, the design problem is related to the optimization of a certain criterion that is dependent on the tap length. For convenience, we shall formulate the LMS algorithm within a system identification framework, in which the unknown filter $\mathbf{c}_{L_{\mathrm{opt}}}$ has an unknown tap length L_{opt}, which is to be identified. In this model, the desired signal $d(n)$ is represented as

$$d(n) = \mathbf{x}_{L_{\mathrm{opt}}}^T(n)\mathbf{c}_{L_{\mathrm{opt}}} + v(n) \qquad (1)$$

where $\mathbf{x}_{L_{\mathrm{opt}}}(n)$ is the input vector with a tap length of L_{opt}, $v(n)$ is a zero-mean additive noise term uncorrelated with the input, n denotes the discrete time index, and $(\cdot)^T$ denotes the transpose operation. In this correspondence, all quantities are assumed to be real valued.

For convenience of description, we assume that at steady state the tap length of the adaptive filter is a fixed value and denoted by L; $\mathbf{w}_L$ and $\mathbf{x}_L(n)$ are, respectively, the corresponding steady-state adaptive filter vector and input vector. Also, we define the segmented steady-state error as [10]

$$e_M^{(L)}(n) = d(n) - \mathbf{w}_{L,1:M}^T \mathbf{x}_{L,1:M}(n), \text{ as } \quad n \to \infty \qquad (2)$$

where $1 \leq M \leq L$, $\mathbf{w}_{L,1:M}$ and $\mathbf{x}_{L,1:M}(n)$ are, respectively, vectors consisting of the first M coefficients of the steady-state filter vector $\mathbf{w}_L$ and the input vector $\mathbf{x}_L(n)$. The mean square of this segmented steady-state error is defined as $\xi_M^{(L)} = E\{(e_M^{(L)}(n))^2\}$. The underlying basis of the FT method is to find the minimum value of L satisfying [10]:

$$\xi_{L-\Delta}^{(L)} - \xi_L^{(L)} \leq \varepsilon \qquad (3)$$

where Δ is a positive integer less than L and ε is a small positive value determined by the system requirements. The minimum L that satisfies

(3) is then chosen as the optimal tap length. A detailed description of this criterion and another similar criterion can be found in [10].

Gradient-based methods can be used to estimate L on the basis of (3). However, the tap length that should be used in the adaptive filter structure must be an integer, and this constrains the adaptation of the tap length. Different approaches have been applied to solve this problem [6]–[10]. In [10], the concept of "pseudo fractional tap length," denoted by $l_f(n)$, is utilized to make instantaneous tap-length adaptation possible. The update of the fractional tap-length is as follows:

$$l_f(n+1) = (l_f(n) - \alpha) - \gamma \left[(e_{L(n)}^{(L(n))})^2 - (e_{L(n)-\Delta}^{(L(n))})^2 \right] \quad (4)$$

where γ is the step size for the tap-length adaptation, and α is a positive leakage parameter [10]. As explained in [10], $l_f(n)$ is no longer constrained to be an integer, and the tap length $L(n+1)$, which will be used in the adaptation of the filter weights in the next iteration, is obtained from the fractional tap length $l_f(n)$ as follows:

$$L(n+1) = \begin{cases} \lfloor l_f(n) \rfloor, & \text{if } |L(n) - l_f(n)| > \delta \\ L(n), & \text{otherwise} \end{cases} \quad (5)$$

where $\lfloor . \rfloor$ is the floor operator, which rounds down the embraced value to the nearest integer and δ is a small integer.

Next, we will give a steady-state performance analysis based on the above formulation, which is the novel contribution in this correspondence. Some general guidelines for the parameter choice will also be provided.

III. STEADY-STATE PERFORMANCE

In the FT algorithm, the filter coefficients are updated as

$$\mathbf{w}_{L(n)}(n+1) = \mathbf{w}_{L(n)}(n) + \mu e_{L(n)}^{(L(n))}(n) \mathbf{x}_{L(n)}(n) \quad (6)$$

where $\mathbf{w}_{L(n)}$ and $\mathbf{x}_{L(n)}$ are, respectively, the $L(n)$-tap adaptive filter vector and the input vector, and μ is the positive step size for the update of the coefficients.

For convenience of analysis, we use a vector $\mathbf{c}_N$ to denote the unknown filter, where N is an integer larger than both the optimal tap length L_{opt} and the maximum value of the variable tap-length sequence $L(n)$, and $\mathbf{c}_N$ is obtained by padding $\mathbf{c}_{L_{\text{opt}}}$ with zeros. This unknown filter vector $\mathbf{c}_N$ can be split into three parts:

$$\begin{pmatrix} \mathbf{c}' \\ \mathbf{c}'' \\ \mathbf{c}''' \end{pmatrix}$$

where $\mathbf{c}'$ is the part modeled by $\mathbf{w}'(n), \mathbf{w}'(n) = \mathbf{w}_{L(n),1:L(n)-\Delta}(n), \mathbf{c}''$ is the part modeled by $\mathbf{w}''(n), \mathbf{w}''(n) = \mathbf{w}_{L(n),L(n)-\Delta+1:L(n)}(n)$ and $\mathbf{c}''$ is the undermodeled part. We let $\mathbf{g}_N(n)$ denote the total coefficient error vector

$$\mathbf{g}_N(n) = \mathbf{c}_N - \mathbf{w}_N(n) \quad (7)$$

where $\mathbf{w}_N(n)$ is obtained by padding $\mathbf{w}_{L(n)}(n)$ with zeros. Therefore, $\mathbf{g}_N(n)$ can be similarly split as

$$\begin{pmatrix} \mathbf{g}'(n) \\ \mathbf{g}''(n) \\ \mathbf{g}'''(n) \end{pmatrix}.$$

The mean square deviation (MSD) between the optimal filter vector and the adaptive filter vector is given by $E\{\|\mathbf{g}_N(n)\|_2^2\}$, where $\|\cdot\|_2^2$ denotes the squared Euclidean distance.

For convenience of description, the input vector $\mathbf{x}_N(n)$ is split similarly to that of $\mathbf{c}_N(n)$ and $\mathbf{g}_N(n)$. With the above notation and substituting (1) and (7) into (2), and padding all the vectors in (1) and (2) with zeros to make their lengths equal to N, we have

$$e_{L(n)}^{(L(n))}(n) = \begin{pmatrix} \mathbf{x}'(n) \\ \mathbf{x}''(n) \\ \mathbf{x}'''(n) \end{pmatrix}^T \begin{pmatrix} \mathbf{g}'(n) \\ \mathbf{g}''(n) \\ \mathbf{c}''' \end{pmatrix} + v(n) \quad (8)$$

and

$$e_{L(n)-\Delta}^{(L(n))}(n) = \begin{pmatrix} \mathbf{x}'(n) \\ \mathbf{x}''(n) \\ \mathbf{x}'''(n) \end{pmatrix}^T \begin{pmatrix} \mathbf{g}'(n) \\ \mathbf{c}'' \\ \mathbf{c}''' \end{pmatrix} + v(n). \quad (9)$$

The term $(e_{L(n)}^{(L(n))})^2 - (e_{L(n)-\Delta}^{(L(n))})^2$, which is the key term in the fractional tap-length update (4), can then be written as

$$\left(e_{L(n)}^{(L(n))} \right)^2 - \left(e_{L(n)-\Delta}^{(L(n))} \right)^2$$
$$= [2v(n) + 2\mathbf{x}'^T(n)\mathbf{g}'(n)$$
$$+ \mathbf{x}''^T(n)\mathbf{g}''(n) + \mathbf{x}''^T(n)\mathbf{c}'' + 2\mathbf{x}'''^T(n)\mathbf{c}''']$$
$$\times [\mathbf{x}''^T(n)\mathbf{g}''(n) - \mathbf{x}''^T(n)\mathbf{c}'']. \quad (10)$$

This term can be expanded as

$$\left(e_{L(n)}^{(L(n))} \right)^2 - \left(e_{L(n)-\Delta}^{(L(n))} \right)^2$$
$$= \underbrace{2v(n)\mathbf{x}''^T(n)\mathbf{g}''(n)}_{A} - \underbrace{2v(n)\mathbf{x}''^T(n)\mathbf{c}''}_{B}$$
$$+ \underbrace{2\mathbf{x}'^T(n)\mathbf{g}'(n)\mathbf{x}''^T(n)\mathbf{g}''(n)}_{C}$$
$$- \underbrace{2\mathbf{x}'^T(n)\mathbf{g}'(n)\mathbf{x}''^T(n)\mathbf{c}''}_{D}$$
$$+ \underbrace{[\mathbf{x}''^T(n)\mathbf{g}''(n)]^2}_{E} - \underbrace{[\mathbf{x}''^T(n)\mathbf{c}'']^2}_{F}$$
$$+ \underbrace{2\mathbf{x}'''^T(n)\mathbf{c}'''\mathbf{x}''^T(n)\mathbf{g}''(n)}_{G}$$
$$- \underbrace{2\mathbf{x}'''^T(n)\mathbf{c}'''\mathbf{x}''^T(n)\mathbf{c}''}_{H}. \quad (11)$$

Substituting (11) into (4), the steady-state fractional tap-length update equation can be rewritten as

$$l_f(n+1)$$
$$= l_f(n) - (\alpha + \gamma(A - B + C - D + E - F + G - H)) \quad (12)$$

where terms A, B, C, D, E, F, G, and H are denoted in (11).

Next, a steady-state performance analysis will be given based on this update equation.

A. Steady-State Performance Analysis

Before we perform the steady-state analysis, we shall assume that the system has arrived at steady-state if the quantities $E\{(e_{L(n)}^{(L(n))})^2\}$ and $E\{l_f(n)\}$ tend to constants as $n \to \infty$. To simplify the analysis, we make several further assumptions.

A.1. At steady-state, the tap length will converge, or can be approximately deemed to have converged to a fixed value $L(\infty)$. As will be

shown by the simulations, if all the parameters are set properly, the tap length will slightly fluctuate around a fixed value. Also we assume that at steady state, $L(\infty) > L_{\text{opt}}$. We justify that in most simulations, if the parameter γ is not chosen too small, the steady-state tap length $L(\infty)$ will be always larger than L_{opt}. This overestimate phenomenon is also justified and discussed in [10] and can be seen in the simulations in the next section.

A.2. Both the input signal $x(n)$ and the noise signal $v(n)$ are statistically independent identically distributed (i.i.d.) zero-mean Gaussian white signals with variances σ_x^2 and σ_v^2, respectively.

A.3. The tail elements of the unknown optimal filter vector $\mathbf{c}_{L_{\text{opt}}}$ can be deemed to be drawn from a random white sequence with zero mean and variance σ_c^2. This assumption is used to simplify the analysis. We note that even for a filter with a decaying impulse response structure, the tail elements can be approximately deemed as to have the same variance if the tap length is long enough; thus, this assumption matches the observations in many applications.

A.4. At steady state, the vectors $\mathbf{g}'(n)$ and $\mathbf{g}''(n)$, which are due to the adaptive noise, are independent of $\mathbf{x}_N(n)$. The justification of this assumption is that the updates of $\mathbf{g}'(n)$ and $\mathbf{g}''(n)$ only depend on the past input vectors, and from assumption A.2 we know that $\mathbf{x}_N(n)$ is independent of $\mathbf{x}_N(j)$ if $j \neq n$; thus, $\mathbf{g}'(n)$ and $\mathbf{g}''(n)$ are independent of $\mathbf{x}_N(n)$ [12]. Also, in order to simplify the analysis, we assume that at steady state

$$E\left\{ \begin{pmatrix} \mathbf{g}'(n) \\ \mathbf{g}''(n) \end{pmatrix} \begin{pmatrix} \mathbf{g}'(n) \\ \mathbf{g}''(n) \end{pmatrix}^T \right\} = \sigma_g^2 I \qquad (13)$$

where I is the identity matrix and σ_g^2 is the variance of the elements of $\mathbf{g}'(n)$ and $\mathbf{g}''(n)$.

Also the tap length is constrained to be not less than a lower floor value $L_{\min}$, where $L_{\min} > \Delta$, during its evolution, i.e., if the tap length fluctuates under $L_{\min}$, it will be set to $L_{\min}$. This operation is necessary since $L(n) - \Delta$ is used as a tap length in the FT algorithm, as can be seen in (4), and it should be positive.

Taking expectation of both sides of (12), we have

$$E\{(A - B + C - D + E - F + G - H)\} = -\frac{\alpha}{\gamma}. \qquad (14)$$

Using assumptions A.1 and A.3, we have (15) and (16), shown at the bottom of the page. With (15), it is straightforward to see that the terms G and H in (14) will disappear at steady state. Using assumptions A.1, A.2, A.3, and A.4, we know that the expectations of terms A, B, C, and D will be zero at steady state, and (14) can be written as

$$E\{(A - B + C - D + E - F)\} = \sigma_x^2(E\{\|\mathbf{g}''(n)\|_2^2\} - \|\mathbf{c}''\|_2^2). \qquad (17)$$

It is straightforward to see that if $L(\infty) > L_{\text{opt}} + \Delta$, it will imply that (14) and (17) contradict each other, i.e., together with (16) we know

that the right-hand side (RHS) of (17) is larger than zero if $L(\infty) > L_{\text{opt}} + \Delta$, but the RHS of (14) is a negative value. We therefore conclude that $L(\infty) \leq L_{\text{opt}} + \Delta$, so that by also exploiting assumption A.1, the condition that $L_{\text{opt}} \leq L(\infty) \leq L_{\text{opt}} + \Delta$ will always be used in the following derivations.

In a manner similar to [3] and [5], in order to speed up the convergence rate of the FT variable tap-length LMS algorithm, the step size is made variable rather than fixed, according to the range of μ described in [3]:

$$\mu(n) = \mu'/((L(n) + 2)\sigma_x^2) \qquad (18)$$

where μ' is a constant. With this variable step size, the term $E\{\|\mathbf{g}''(n)\|_2^2\}$ can be derived as in (9) in Appendix I. Substituting (16) and (9) into (17) yields

$$E\{A - B + C - D + E - F\}$$
$$\approx \sigma_x^2 \left[\frac{\Delta \mu' \sigma_v^2}{(2 - \mu')L_{\text{opt}}\sigma_x^2} - (L_{\text{opt}} + \Delta - L(\infty))\sigma_c^2 \right]. \qquad (19)$$

Utilizing (14) in (19), we have

$$-\frac{\alpha}{\gamma \sigma_x^2} \approx \frac{\Delta \mu' \sigma_v^2}{(2 - \mu')L_{\text{opt}}\sigma_x^2} - (L_{\text{opt}} + \Delta - L(\infty))\sigma_c^2. \qquad (20)$$

From (20), we obtain

$$L(\infty) \approx L_{\text{opt}} + \Delta - \frac{\alpha}{\gamma \sigma_x^2 \sigma_c^2} - \frac{\Delta \mu' \sigma_v^2}{(2 - \mu')L_{\text{opt}}\sigma_x^2\sigma_c^2}. \qquad (21)$$

This equation gives a mathematical formulation of the steady-state tap length. Since the steady-state tap-length value given in (21) will seldom be an integer, in practice the steady-state tap length will fluctuate around this value. This can be clearly seen in later simulations. The steady-state MSE can then be easily obtained with the steady-state tap length given in (21) by using the analysis results provided in [3]. In practice, the last two terms of the RHS of (21) will be small, and the steady-state tap length will be close to the value $L_{\text{opt}} + \Delta$, as will be shown in the simulations in next section.

To avoid the undermodeling situation, the parameters should be chosen to make $L(\infty) > L_{\text{opt}}$ and obtain a small fluctuation of the steady-state tap length. Next, we will give some guidelines for the parameters choice.

B. Guidelines for the Parameter Choice

In this section, we give some general guidelines for parameter choice. To choose the parameters properly, we need estimations of the optimal tap length, the input variance, the noise variance, and the desired system MSE. The availability of these estimations will be application dependent and therefore outside of the scope of this study.

$$\mathbf{c}''' = \mathbf{0} \qquad (15)$$

and

$$\|\mathbf{c}''\|_2^2 \approx \begin{cases} (L_{\text{opt}} + \Delta - L(\infty))\sigma_c^2, & \text{if } L_{\text{opt}} < L(\infty) \leq L_{\text{opt}} + \Delta \\ 0, & \text{if } L(\infty) > L_{\text{opt}} + \Delta \end{cases} \qquad (16)$$

With these estimated values, the parameters used in the FT algorithm can be determined as follows.

1. The parameter δ in (5) is not a crucial parameter, since it is just used to obtain an integer value of the tap length for the coefficients adaptation and can be easily set to a small integer.

2. The choice of the parameter μ' can be determined according to the system MSE requirement. Similar to the step-size choice of the LMS algorithm, μ' can be a large value in low-noise conditions to obtain faster convergence of the MSE and should be small in high-noise conditions to avoid a large MSE.

3. The parameter Δ should be as large as possible to obtain a fast convergence rate of the tap length, but also much smaller than the estimate of L_{opt}, so that the steady-state tap length formulated in (21) will not be significantly biased from L_{opt}. For example, $\Delta \approx 0.1 L_{\mathrm{opt}}$ will be a good choice for a wide range of optimal tap lengths.

4. The leakage parameter α should not be too large, so that it will not influence the initial tap-length convergence rate too much. The parameter α should not be too small either, so that once the tap length is overestimated, α can make the tap length converge close to the steady-state value as soon as possible. For example, $\alpha = 0.0001$ is not a good choice, since it means that after 10 000 iterations, the leakage parameter α will reduce the tap length by one tap, which is usually too slow. Generally, values between 0.001 and 0.01 are good choices for α.

5. The parameter γ is the step-size parameter that controls the adaptation process of the variable tap length. Similar to the step size in the LMS algorithm, a large parameter γ will speed up the convergence rate of the tap length, but will result in a large fluctuation of the steady-state tap length. On the other hand, a small parameter γ can obtain a small fluctuation of the steady-state tap length but lead to a slow convergence rate. Thus, γ provides a tradeoff between the convergence rate of the tap length and the steady-state tap-length variance. The choice of this parameter is important in the FT algorithm. A detailed discussion for the choice of this parameter is as follows.

At first, to avoid undermodeling the optimal tap length, the steady-state tap length of the FT method $L(\infty)$ should not be less than L_{opt}. Considering the fluctuation of the steady-state tap length, the parameter γ should be set properly so that $L(\infty) > L_{\mathrm{opt}} + \kappa\delta$, where κ is a small positive integer and can be chosen according to the system requirement of the fluctuation of the steady-state tap length. For example, $\kappa = 2$ is a reasonable choice. Substituting (21) into the inequality $L(\infty) > L_{\mathrm{opt}} + \kappa\delta$, we can obtain the lower bound value γ_l:

$$\gamma_l = \frac{\alpha}{(\Delta - \kappa\delta)\sigma_x^2\sigma_c^2 - \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}}}. \tag{22}$$

Second, the parameter γ should not be too large to avoid a large fluctuation of the steady-state tap length. The update process for the steady-state fractional tap length is formulated in (12). The fluctuation of the steady-state fractional tap length is brought about by the fluctuation of the term $\alpha + \gamma(A - B + C - D + E - F + G - H)$. It is straightforward to see in (12) that large variance of the term $\alpha + \gamma(A - B + C - D + E - F + G - H)$, which is denoted as σ_f^2, will result in large fluctuations of the steady-state fractional tap length. To avoid such a situation, a simple and intuitive approach is to make the standard deviation σ_f much smaller than the parameter δ in (5), so that the probability of the steady-state fractional tap length fluctuating outside

the range $(L(\infty) - \delta, L(\infty) + \delta)$ can be very small. A simple criterion to satisfy such a requirement is

$$\sigma_f < \rho\delta \tag{23}$$

where ρ is a small positive value and can be decided according to the system requirement of the fluctuation of the steady-state tap length. The derivation of the variance σ_f^2 is given in (51) in Appendix II. Using (51) in (23) and after rearrangement, we obtain the upper bound value γ_u:

$$\gamma_u = \frac{-\frac{K_2\alpha}{K_3} - \sqrt{\frac{K_2^2\alpha^2}{K_3^2} + \frac{4(\rho^2\delta^2 + \alpha^2)}{K_3}}}{2} \tag{24}$$

where K_2 and K_3 are, respectively, formulated in (54) and (55).

The parameter ρ should be chosen so that the possibility of the tap-length fluctuating under L_{opt} is nearly zero. In general, for high noise condition, this parameter should be chosen small, and for low noise condition it can be chosen larger. Examples for the choice of ρ can be seen in the simulations in the next section. With the lower bound value given in (22) and the upper bound value given in (24), the parameter γ can then be easily chosen. According to the motivation of the upper bound value γ_u we know that values close to this value will be good choices to avoid a large fluctuation of the steady-state tap length while retaining as quick as possible convergence rate; thus, in practice γ should be chosen close to γ_u and larger than γ_l.

Since in practice, all the parameters $\sigma_x^2, \sigma_v^2, \sigma_c^2$, especially the parameter L_{opt} are unknown, approximate estimations of these parameters can be used in the calculations. Next, we will perform several simulations to confirm the above analysis and discussions.

IV. SIMULATION RESULTS

In this section, we will perform two simulations to support the analysis and discussions in the previous section. In the first simulation, a low-noise condition is used while a high-noise environment is utilized in the second simulation.

A. Low-Noise Case: SNR = 20 dB

The setup of this simulation is as follows. The impulse response sequence of the unknown filter is a white Gaussian sequence with zero mean and variance 0.01. The tap length L_{opt} is set to 200. The input signal is another white Gaussian sequence with zero mean and unit variance. The noise signal is a zero mean uncorrelated random Gaussian sequence and scaled to make the SNR 20 dB. According to the parameter choice guidelines in Section III-B, the parameter δ is set to 2. The step size μ' is set to 0.5. The leakage parameter α is set to 0.005, and Δ is set to 20. To obtain the lower bound value of γ, we set the parameter $\kappa = 2$, and using the above parameter settings in (22), we have $\gamma_l = 0.0314$. Similarly, we have $\gamma_u = 17.954$ by using $\rho = 0.5$ in (24). To compare the performance with different values of γ, we use $\gamma = 0.1\gamma_u, \gamma = \gamma_u$ and $\gamma = 10\gamma_u$ in the simulation, respectively. Note that all these sets of γ are larger than the lower bound value γ_l. The evolution curves of the tap length with different parameter γ values are shown in Fig. 1. The evolution curves of the EMSE with different parameter γ values are shown in Fig. 2.

It is clear to see from Fig. 1 that $\gamma = \gamma_u$ provides a good tradeoff between the convergence rate of the tap length and the steady-state tap-length variance. The algorithm with a parameter $\gamma = 0.1\gamma_u$ gives a very smooth curve of the steady-state tap length, but the convergence rate of both the tap length and the EMSE is very slow. The algorithm with a parameter $\gamma = 10\gamma_u$ provides a quick convergence rate, but the
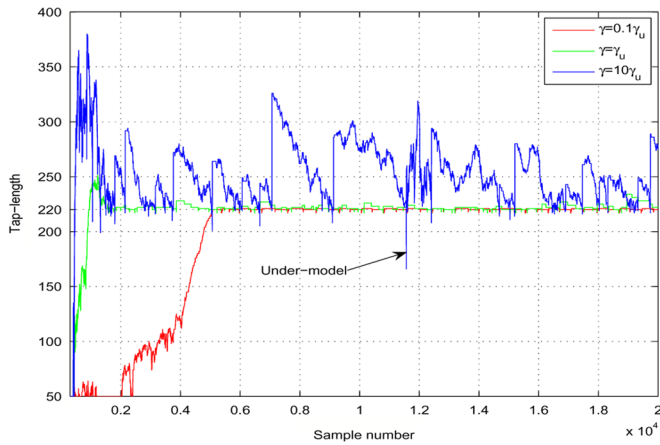
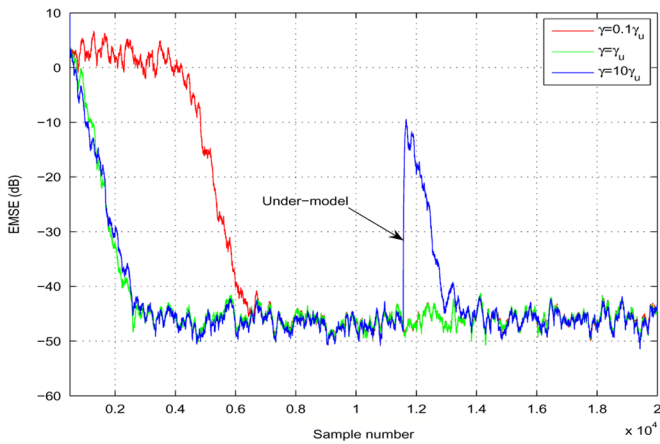Fig. 1. Evolution curves of the tap length with different step sizes under a low-noise condition, SNR = 20 dB.
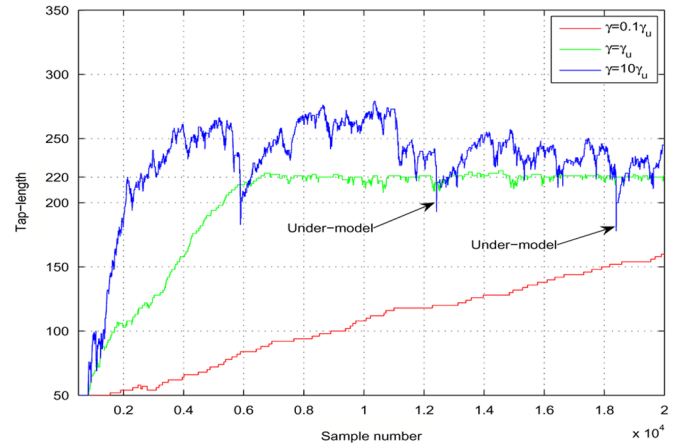

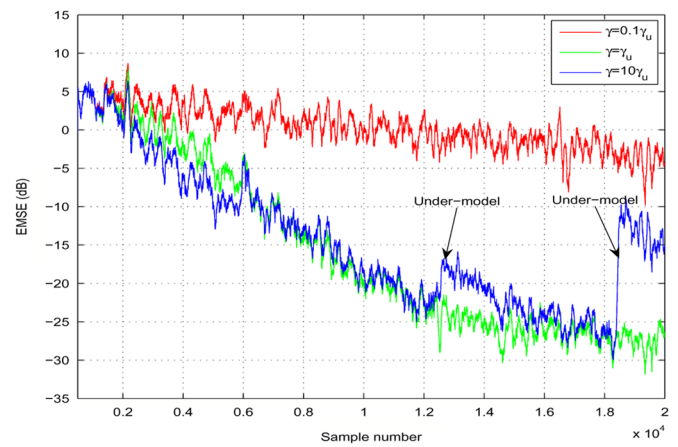Fig. 3. Evolution curves of the tap length with different step sizes under a high-noise condition, SNR = 0 dB.

Fig. 2. Evolution curves of the EMSE with different step sizes under a low-noise condition, SNR = 20 dB.

Fig. 4. Evolution curves of the EMSE with different step sizes under a high-noise condition, SNR = 0 dB.

tap length fluctuates greatly. When the tap length is undermodeled, i.e., $L(n) < L_{\rm opt}$, the EMSE will increase, as can be seen in Fig. 2.

Substituting all the parameter sets into (21) we obtain the theoretical values of the steady-state tap length $L(\infty \,|\, \gamma = 0.1\gamma_u) = 219.65$, $L(\infty \,|\, \gamma = \gamma_u) = 219.90$ and $L(\infty \,|\, \gamma = 10\gamma_u) = 219.93$, which are all close to $L_{\rm opt} + \Delta$. It can be seen from Fig. 1 that for the parameter sets $\gamma = \gamma_u$ and $\gamma = 0.1\gamma_u$, the simulation results of the steady-state tap length match the theoretical values very well: the steady-state tap length fluctuates around the theoretical value $L(\infty)$, which confirms (21).

B. High-Noise Case: SNR = 0 dB

In this simulation, a high-noise environment is used. The setup for this simulation is as follows. The unknown filter is the same as that in the previous simulation, which is a filter with an impulse response sequence drawn from a white Gaussian sequence with zero mean and a variance of 0.01, and a tap length of 200. The input signal is another white Gaussian sequence with zero mean and unit variance. The noise signal is a zero mean uncorrelated random Gaussian sequence, and scaled to make the SNR 0 dB. According to the parameter choice guidelines in Section III-B, the parameter δ is set 2. The step size μ' is set as 0.05 to obtain a small EMSE. The leakage parameter α is set to 0.005. Δ is set to 20.

Similar to the first simulation, in order to obtain the lower bound value of γ, we set the parameter $\kappa = 2$, and using the above parameter

settings in (22), we have $\gamma_l = 0.0323$. Similarly, we have $\gamma_u = 1.0255$ by using $\rho = 0.2$ in (24). To compare the performance with different values of γ, we use $\gamma = 0.1\gamma_u$, $\gamma = \gamma_u$ and $\gamma = 10\gamma_u$ in the simulation, respectively. The evolution curves of the tap length with different parameter γ values are shown in Fig. 3. The evolution curves of the EMSE with different parameter γ values are shown in Fig. 4.

Again from Fig. 3 we can see that $\gamma = \gamma_u$ provides a good tradeoff between the convergence rate of the tap length and the steady-state tap-length variance. The convergence rate of both the tap length and EMSE with parameter $\gamma = 0.1\gamma_u$ is too slow for the algorithm. The algorithm with parameter $\gamma = 10\gamma_u$ provides a quick convergence rate of the tap length, but the fluctuation of the steady-state tap length is very large. Once the tap length fluctuates under $L_{\rm opt}$, EMSE will increase, as can be seen in Fig. 4.

Substituting all the parameter sets into (21), we obtain the theoretical values of the steady-state tap length $L(\infty|\gamma = 0.1\gamma_u) = 214.6$, $L(\infty|\gamma = \gamma_u) = 219.0$ and $L(\infty|\gamma = 10\gamma_u) = 219.4$, which are all close to $L_{\rm opt} + \Delta$. For $\gamma = \gamma_u$, the simulation results of the steady-state tap length match the theoretical values very well, which confirms (21).

To obtain both a fast convergence rate and a small steady-state EMSE for high noise condition, the convex combination approach can be considered, in which two filters are updated simultaneously with different parameters $\gamma = 10\gamma_u$ and $\gamma = \gamma_u$, so that the overall filter can obtain

both a rapid convergence rate from the fast filter and a smooth curve for the steady-state tap length from the slow filter [11].

V. Conclusion

A steady-state performance analysis for the FT variable tap-length LMS algorithm is given in this correspondence. In this analysis, a mathematical formulation of the steady-state tap-length value is obtained, and some general guidelines for the parameter choice are also given, which aid the practical applicability of the algorithm. Simulation results show that the analysis is correct and the parameter choice guidelines are very useful. The analysis and discussion results can be widely used in the applications of the FT algorithm.

Appendix I

According to the analysis in [3], we know that with a fixed tap length, the steady-state value of $E\{\|\mathbf{g}_N(n)\|_2^2\}$ can be expressed as

$$E\{\|\mathbf{g}_N(n)\|_2^2\} = \frac{(s-r)\|\mathbf{c}'''\|_2^2 + t}{1-r} \tag{25}$$

where

$$r = 1 - 2\mu\sigma_x^2 + (L(\infty) + 2)\mu^2\sigma_x^4 \tag{26}$$

$$s = 1 + L(\infty)\mu^2\sigma_x^4 \tag{27}$$

and

$$t = L(\infty)\mu^2\sigma_x^2\sigma_v^2. \tag{28}$$

Moreover, the MSD can be divided into three parts [3]:

$$E\{\|\mathbf{g}_N(n)\|_2^2\} = E\{\|\mathbf{g}'(n)\|_2^2\} + E\{\|\mathbf{g}''(n)\|_2^2\} + \|\mathbf{c}'''\|_2^2. \tag{29}$$

Substituting (25) into (29), we have

$$E\{\|\mathbf{g}'(n)\|_2^2\} + E\{\|\mathbf{g}''(n)\|_2^2\} = \frac{(s-1)\|\mathbf{c}'''\|_2^2 + t}{1-r}. \tag{30}$$

Substituting (15), (26), (27), and (28) into (30), and with the approximation $L(\infty) \approx L(\infty) + 2$, we have

$$E\{\|\mathbf{g}'(n)\|_2^2\} + E\{\|\mathbf{g}''(n)\|_2^2\} \approx \frac{\mu'\sigma_v^2}{(2-\mu')\sigma_x^2}. \tag{31}$$

With assumption A.4, we have $E\{\|\mathbf{g}'(n)\|_2^2\} = (L(\infty) - \Delta)\sigma_g^2$ and $E\{\|\mathbf{g}''(n)\|_2^2\} = \Delta\sigma_g^2$, together with (31), we have

$$E\{\|\mathbf{g}''(n)\|_2^2\} \approx \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L(\infty)\sigma_x^2}. \tag{32}$$

From the previous discussion, we know $L(\infty) \leq L_{\text{opt}} + \Delta$, and in practice $\Delta \ll L_{\text{opt}}$; thus, if $L(\infty) > L_{\text{opt}}$, $L(\infty)$ will be very close to L_{opt}. Equation (32) can then be approximately written as

$$E\{\|\mathbf{g}''(n)\|_2^2\} \approx \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\text{opt}}\sigma_x^2}. \tag{33}$$

Appendix II

Note that all the following derivation is based on the condition $L_{\text{opt}} \leq L(\infty) \leq L_{\text{opt}} + \Delta$, and the terms G and H are equal to 0 at steady state. The variance of the term $\alpha + \gamma(A - B + C - D + E - F + G - H)$ is

$$\sigma_f^2 = E\{(\alpha + \gamma(A - B + C - D + E - F)) \\ -E\{\alpha + \gamma(A - B + C - D + E - F)\}\}^2. \tag{34}$$

From (14), we have

$$E\{\alpha + \gamma(A - B + C - D + E - F)\} = 0. \tag{35}$$

Substituting (14) and (35) into (34), and using assumptions A.2, A.3, A.4, and the mathematical formulation of terms A, B, C, D, E and F, it is straightforward to obtain that

$$\sigma_f^2 = \gamma^2(\sigma_A^2 + \sigma_B^2 + \sigma_C^2 + \sigma_D^2 + \sigma_E^2 + \sigma_F^2 - 2E\{EF\}) - \alpha^2. \tag{36}$$

With assumptions A.2 and A.4 and using (33) in the mathematical formulation of term A, we can obtain the variance of term A as

$$\sigma_A^2 = \frac{4\Delta\mu'\sigma_v^4}{(2-\mu')L_{\text{opt}}}. \tag{37}$$

Similarly, with assumptions A.2 and A.3, and using (16) in the mathematical formulation of term B, we obtain the variance of term B as

$$\sigma_B^2 \approx 4\sigma_v^2(L_{\text{opt}} + \Delta - L(\infty))\sigma_c^2\sigma_x^2. \tag{38}$$

Substituting (21) into (38), we have

$$\sigma_B^2 \approx 4\left(\frac{\alpha\sigma_v^2}{\gamma} + \frac{\Delta\mu'\sigma_v^4}{(2-\mu')L_{\text{opt}}}\right). \tag{39}$$

Using assumptions A.2 and A.4 in the mathematical formulation of term C, we can obtain the variance of term C as

$$\sigma_C^2 \approx 4\Delta(L(\infty) - \Delta)\sigma_g^4\sigma_x^4. \tag{40}$$

With A.4 and using (33), we have

$$\sigma_g^2 \approx \frac{\mu'\sigma_v^2}{(2-\mu')L_{\text{opt}}\sigma_x^2}. \tag{41}$$

Substituting (41) into (40), and with the approximation $L_{\mathrm{opt}} \approx L(\infty) - \Delta$, we have

$$\sigma_C^2 \approx \frac{4\Delta\mu'^2\sigma_v^4}{(2-\mu')^2 L_{\mathrm{opt}}}. \tag{42}$$

Similarly, with assumptions A.2, A.3, and A.4, and using (16), (21), and (41) in the mathematical formulation of term D, and with the approximation $L_{\mathrm{opt}} \approx L(\infty) - \Delta$, we can obtain the variance of term D as

$$\sigma_D^2 \approx \frac{4\mu'\sigma_v^2}{(2-\mu')}\left(\frac{\alpha}{\gamma} + \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}}\right). \tag{43}$$

Since the term $\mathbf{x}''^T(n)\mathbf{g}''(n)$ can be approximately deemed as a sum of Δ i.i.d. random variables, from the central limit theory we know that the probability density function (pdf) of this term will be very close to a Gaussian distribution with a zero mean; thu,s term E can be approximately deemed as chi-squared distributed with one degree of freedom. With assumptions A.2 and A.4, and using (9) we have the mean value of term E

$$m_E \approx \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}}. \tag{44}$$

Thus, the variance of term E is

$$\sigma_E^2 = 2m_E \approx 2\frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}}. \tag{45}$$

Similarly, the term $\mathbf{x}''^T(n)\mathbf{c}''$ can be approximately deemed as Gaussian distributed; thus, term F can be approximately deemed as chi-squared distributed with one degree of freedom. With assumptions A.3 and A.4, and using (16) and (21), we have the mean value of term F

$$m_F \approx \frac{\alpha}{\gamma} + \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}}. \tag{46}$$

Thus, the variance for term F is

$$\sigma_F^2 = 2m_F = 2\left(\frac{\alpha}{\gamma} + \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}}\right). \tag{47}$$

With assumptions A.3 and A.4, the term $E\{EF\}$ can be expanded as

$$\begin{aligned} E\{EF\} &= E\{(\mathbf{x}''^T(n)\mathbf{g}''(n))^2(\mathbf{x}''^T(n)\mathbf{c}'')^2\} \\ &= E\{\mathbf{x}''^T(n)\mathbf{g}''(n)\mathbf{g}''^T(n)\mathbf{x}''(n)\mathbf{x}''^T(n)\mathbf{c}''\mathbf{c}''^T\mathbf{x}(n)\} \\ &\approx \sigma_g^2\sigma_c^2 E\{\mathbf{x}''^T(n)\mathbf{x}''(n)\mathbf{x}''^T(n)\mathbf{x}''(n)\} \end{aligned} \tag{48}$$

by approximating the instantaneous term $\mathbf{c}''\mathbf{c}''^T$ with its statistical averaging value.

The input signal is an i.i.d Gaussian sequence, thus $E\{x^4\} = 3\sigma_x^2$. From (48) it is straightforward to obtain

$$E\{EF\} = \sigma_g^2\sigma_c^2(\Delta(\Delta-1)\sigma_x^4 + 3\Delta\sigma_x^2). \tag{49}$$

Substituting (41) into (49), we have

$$E\{EF\} = \frac{\Delta\mu'\sigma_v^2\sigma_c^2((\Delta-1)\sigma_x^2 + 3)}{(2-\mu')L_{\mathrm{opt}}}. \tag{50}$$

Substituting (37), (39), (42), (43), (45), (47), and (50) into (36), we have

$$\sigma_f^2 = \gamma^2\left(K_2\frac{\alpha}{\gamma} + K_3\right) - \alpha^2 \tag{51}$$

where

$$K_1 = \frac{\Delta\mu'\sigma_v^2}{(2-\mu')L_{\mathrm{opt}}} \tag{52}$$

and

$$K_2 = 2 + 4\sigma_v^2 + \frac{4L_{\mathrm{opt}}K_1}{\Delta} \tag{53}$$

and

$$K_3 = 2K_1K_2 - 2K_1\sigma_c^2\left((\Delta-1)\sigma_x^2 + 3\right). \tag{54}$$

REFERENCES

[1] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. New York: Wiley, 1998.

[2] A. H. Sayed, *Fundamentals of Adaptive Filtering*. New York: Wiley, 2003.

[3] Y. Gu, K. Tang, H. Cui, and W. Du, "Convergence analysis of a deficient-length LMS filter and optimal-length sequence to model exponential decay impulse response," *IEEE Signal Process. Lett.*, vol. 10, no. 1, pp. 4–7, Jan. 2003.

[4] K. Mayyas, "Performance analysis of the deficient length LMS adaptive algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2727–2734, Aug. 2005.

[5] Y. Zhang, J. A. Chambers, S. Sanei, P. Kendrick, and T. J. Cox, A new variable tap-length LMS algorithm to model an exponential decay impulse response vol. 14, no. 4, pp. 263–266, Apr. 2007.

[6] F. Riero-Palou, J. M. Noras, and D. G. M. Cruickshank, "Linear equalisers with dynamic and automatic length selection," *Electron. Lett.*, vol. 37, no. 25, pp. 1553–1554, Dec. 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, pp. 305–307, Mar. 2004.

[8] Y. Gong and C. F. N. Cowan, "A novel variable tap-length algorithm for linear adaptive filters," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Jan. 2004, pp. 825–828.

[9] 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.

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

[11] Y. Zhang and J. A. Chambers, "Convex combination of adaptive filters for variable tap-length LMS algorithm," *IEEE Signal Process. Lett.*, vol. 13, no. 10, pp. 628–631, Oct. 2006.

[12] M. Tarrab and A. Feuer, "Convergence and performance analysis of the normalized LMS algorithm with uncorrelated Gaussian data," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 680–691, Jul. 1988.