Independent University **Bangladesh (IUB)**

IUB Academic Repository

Electrical and Electronics Engineering Article

2016-08-03

Steady-State Tracking Analysis of Adaptive Filter With Maximum Correntropy Criterion

Khalili, Azam

Springer

http://dir.iub.edu.bd:8180/handle/123456789/264

Downloaded from IUB Academic Repository

Circuits Syst Signal Process
DOI 10.1007/s00034-016-0373-9

SHORT PAPER

Steady-State Tracking Analysis of Adaptive Filter With Maximum Correntropy Criterion

Azam Khalili[1] · **Amir Rastegarnia**[1] · **Md Kafiul Islam**[2] · **Tohid Yousefi Rezaii**[3]

Received: 17 December 2015 / Revised: 14 July 2016 / Accepted: 16 July 2016
© Springer Science+Business Media New York 2016

Abstract This letter studies the tracking performance of a stochastic gradient-based adaptive algorithm, namely the maximum correntropy criterion algorithm, where a random walk is used to model the non-stationarity. In our analysis, we use the energy conservation argument to derive expressions for the steady-state excess mean square error (EMSE). We consider two different cases for measurement of noise distribution including the Gaussian noise and general non-Gaussian noise. For the Gaussian case, we derive a fixed-point equation that can be solved numerically to find steady-state EMSE value. For the general non-Gaussian case, we derive an approximate closed-form expression for EMSE. For both cases, unlike the stationary environment, the EMSE curves are not increasing functions of step size parameter. We use this observation to find the optimum step size learning parameter for general non-Gaussian case. The validity of the theoretical results are justified via simulation results.

Keywords Adaptive filter · Correntropy · Tracking · Performance analysis

✉ Azam Khalili
a.khalili@ieee.org

Amir Rastegarnia
a_rastegar@ieee.org

Md Kafiul Islam
kafiul_islam@u.nus.edu

Tohid Yousefi Rezaii
yousefi@tabrizu.ac.ir

1 Department of Electrical Engineering, Malayer University, Malayer 65719-95863, Iran

2 Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583, Singapore

3 Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51664, Iran

Published online: 03 August 2016

1 Introduction

Over the last several years, adaptive filters have been used in wide range of applications [8]. In general, an adaptive filter uses a sequence of input vectors $\mathbf{u}_n \in \mathbb{R}^{1\times M}$ and desired samples $d_n \in \mathbb{R}$, $n = 1, 2, \ldots$ to find the optimal weight vector $\mathbf{w}^o \in \mathbb{R}^{M\times 1}$ that minimizes a cost function. In stationary environment, at every time instant n, d_n is related to the input vector $\mathbf{u}_n$ with a regression model as

$$d_n = \mathbf{u}_n \mathbf{w}^o + v_n \tag{1}$$

where $v_n, n = 1, 2, \ldots$ are samples of the measurement noise signal, which are assumed to be zeromean, independent, identically distributed, and independent of the input signal $\mathbf{u}_n$. So far, numerous adaptive filters have been developed in the literature. However, since its invention by Widrow and Hoff [16], the least mean squares (LMS) algorithm is perhaps the most widely used adaptive filter due to its simplicity, robustness, and ease of implementation. The LMS algorithm has been developed based on the minimum mean square error (MMSE) criterion as the cost function, defined by

$$\mathcal{J}_{\mathrm{MMSE}}(\mathbf{w}) \triangleq \mathbb{E}[e_n^2] \tag{2}$$

where e_n is instantaneous error signal which is given by

$$e_n = d_n - \mathbf{u}_n \mathbf{w} \tag{3}$$

Besides, the LMS algorithm uses the steepest descent method with simple stochastic approximations and provides an iterative solution for (2) as

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu \mathbf{u}_n^{\mathrm{T}} e_n \tag{4}$$

where $\mu > 0$ is a suitably chosen step size parameter. Although the MMSE-based adaptive filters work well for Gaussian data, they exhibit performance degradation for nonlinear models and non-Gaussian situations, especially when the data are disturbed by impulsive noise [12]. To address these issues, recently information-theoretic metrics, such as entropy and mutual information, have been introduced as cost functions for adaptive filters. For example, the given algorithms in [5–7] have been developed based on the minimum error entropy (MEE), wherein the filter weights are updated in a way to minimize the entropy of the error signal. The main problem of MEE-based adaptive filters is their high computational complexity. On the other hand, the adaptive filters that rely on maximum correntroy criterion are able to exploit higher-order moments of the data with low complexity as the LMS algorithm [1,2,9,10,13–15,17].

Different aspects of adaptive filters under the maximum correntropy criterion have been studied in the literature. For example, steady-state performance of MCC algorithm has been studied in [3]. In [4], convergence behavior of a fixed-point algorithm under maximum correntropy criterion has been studied. This paper investigates the tracking performance of the MCC algorithm in non-stationary environment where

random walk model is adopted for the optimal parameter variation. In our analysis, we use the energy conservation argument, while the EMSE is considered as performance metric. Two different distributions including the Gaussian and general non-Gaussian measurement noise distributions are considered for measurement noise. For the Gaussian case, we show that EMSE is given by a fixed-point equation, while for the general non-Gaussian case, we can derive an approximate closed-form expression for EMSE. For both cases, unlike the stationary environment, the EMSE curves are not increasing functions of step size parameter. For the general non-Gaussian case, we find the optimum step size parameter which minimizes the EMSE. The validity of the analysis is demonstrated by several computer simulations.

The remainder of this paper is organized as follows. In Sect. 2, we briefly introduce the MCC algorithm. In Sect. 3, tracking analysis of the MCC algorithm is provided. In Sect. 4, we present simulation results to verify our theoretical analysis, and we conclude in Sect. 5.

Notation We adopt small boldface letters for vectors and bold capital letters for matrices.

2 The MCC Algorithm

As we mentioned in the introduction section, the MCC algorithm relies on the correntropy as the cost function. For two random variables X and Y, correntropy is defined as

$$V(X,Y) \triangleq \mathbb{E}[\kappa_\sigma(X-Y)] \tag{5}$$

where $\kappa_\sigma(\cdot,\cdot)$ is a shift-invariant Mercer kernel with the kernel width σ. A popular kernel in correntropy is the Gaussian kernel which is given by

$$\kappa_\sigma(x,y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) \tag{6}$$

To obtain the correntropy from (5), the joint distribution function of (X,Y) is required which is usually unknown. In practice, only finite number of samples $\{x_i, y_i\}$, $i = 1, 2, \ldots, N$ from X and Y are available. Thus, a sample estimator for correntropy can be defined as

$$\hat{V}(X,Y) = \frac{1}{N}\sum_{i=1}^{N}\kappa_\sigma(x_i - y_i) \tag{7}$$

For adaptive filtering, correntropy between the desired signal, d_n, and filter output, $\mathbf{u}_n\mathbf{w}_{n-1}$, is used as the cost function. Using the Gaussian kernel and definition of error e_n, the cost function becomes

$$\mathcal{J}_{\text{corr}}(j) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{N}\sum_{i=j-N+1}^{j}\exp\left(-\frac{e_i^2}{2\sigma^2}\right) \tag{8}$$

The MCC algorithm can be obtained from (8) by applying gradient ascent approach and approximating the sum by the current value $N = 1$ as [3]

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu \exp e_n \left(-\frac{e_n^2}{2\sigma^2} \right) \mathbf{u}_n^{\mathrm{T}} \tag{9}$$

Note that as $\sigma \to \infty$ the MCC algorithm in (9) tends to the LMS algorithm.

3 Tracking Analysis of MCC Algorithm

To begin the analysis, we first assume that in a non-stationary environment, the variation in the optimal weight $\mathbf{w}^o$ follows a random walk model as

$$\mathbf{w}_n^o = \mathbf{w}_{n-1}^o + \mathbf{q}_n \tag{10}$$

where $\mathbf{q}_n$ is an i.i.d. vector with positive-definite autocorrelation matrix $\mathbf{Q} = \mathbb{E}[\mathbf{q}\mathbf{q}^{\mathrm{T}}]$ and is independent of $\{\mathbf{u}_i, d_i\}$ for all $i < n$ and also of initial conditions $\{\mathbf{w}_0, \tilde{\mathbf{w}}_0\}$. We consider again the update equation of MCC algorithm with a general function of the error signal e_n as

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu \mathbf{u}_n^{\mathrm{T}} f(e_n) \tag{11}$$

For further reference, we define the weight error vector $\tilde{\mathbf{w}}_n$ and a priori error signal $e_{a,n}$ as follows

$$\tilde{\mathbf{w}}_n \triangleq \mathbf{w}_n^o - \mathbf{w}_n, \quad e_{a,n} \triangleq \mathbf{u}_n \tilde{\mathbf{w}}_n \tag{12}$$

Note that the steady-state excess mean square error is defined in terms of $e_{a,n}$ as

$$\xi = \lim_{n \to \infty} \mathbb{E}\left[e_{a,n}^2 \right] \tag{13}$$

By subtracting $\mathbf{w}_n^o$ from both sides of (11), we get

$$\tilde{\mathbf{w}}_n = \mathbf{w}_n^o - \mathbf{w}_{n-1} - \mu \mathbf{u}_n^{\mathrm{T}} f(e_n) \overset{(a)}{=} \tilde{\mathbf{w}}_{n-1} + \mathbf{q}_n - \mu \mathbf{u}_n^{\mathrm{T}} f(e_n) \tag{14}$$

where (a) follows by replacing $\mathbf{w}_n^o$ from (10). Equating the weighted norm of (14) and taking expectation from the resultant equation we have

$$\begin{aligned} \mathbb{E}\left[\|\tilde{\mathbf{w}}_n\|^2 \right] = \mathbb{E}\left[\|\tilde{\mathbf{w}}_{n-1}\|^2 \right] - 2\mu \mathbb{E}\left[\mathbf{u}_n \tilde{\mathbf{w}}_{n-1} f(e_n) \right] + \mu^2 \mathbb{E}\left[\|\mathbf{u}_n\|^2 f^2(e_n) \right] \\ + \mathbb{E}\left[\|\mathbf{q}_n\|^2 \right] + \underbrace{\mathbb{E}\left[\tilde{\mathbf{w}}_{n-1}^{\mathrm{T}} \mathbf{q}_n \right]}_{①} + \underbrace{\mathbb{E}\left[\mathbf{q}_n^{\mathrm{T}} \tilde{\mathbf{w}}_{n-1} \right]}_{②} - 2\mu \underbrace{\mathbb{E}\left[\mathbf{q}_n^{\mathrm{T}} \mathbf{u}_n^{\mathrm{T}} f(e_n) \right]}_{③} \end{aligned} \tag{15}$$

To evaluate the term ① first note that $\tilde{\mathbf{w}}_{n-1}$ can be rewritten as

$$\tilde{\mathbf{w}}_{n-1} = \mathbf{w}_{n-1}^{o} - \mathbf{w}_{n-1} = \left(\mathbf{w}_{-1}^{o} + \sum_{j=0}^{n-1} \mathbf{q}_j \right) - \mathbf{w}_{n-1}$$

So we have

$$\mathbb{E}\left[\tilde{\mathbf{w}}_{n-1}^{\mathrm{T}} \mathbf{q}_n\right] = \underbrace{\mathbb{E}\left[\left(\mathbf{w}_{-1}^{o} + \sum_{j=0}^{n-1} \mathbf{q}_j \right)^{\mathrm{T}} \mathbf{q}_n \right]}_{=0} - \underbrace{\mathbb{E}\left[\mathbf{w}_{n-1}^{*} \mathbf{q}_n\right]}_{=0} = 0 \tag{16}$$

where for the first terms in (16), we used the assumption that $\mathbf{q}_n$ is independent of all $\mathbf{q}_k$ for $k < n$ and of initial value $\mathbf{w}_0^o$. Moreover, as $\mathbf{w}_{n-1}$ depends on data $\{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}, d_0, d_1, \ldots, d_{n-1}\}$ and all are independent of $\mathbf{q}_n$ we can conclude that the second term equals zero. Similarly we have ② = ③ = 0. Finally, using $\mathbb{E}\left[\|\mathbf{q}_n\|^2\right] = \mathbb{E}\left[\mathrm{Tr}\left[\mathbf{q}_n \mathbf{q}_n^{\mathrm{T}}\right]\right] = \mathrm{Tr}\left[\mathbf{Q}\right]$ we obtain the following energy conservation relation

$$\mathbb{E}\left[\|\tilde{\mathbf{w}}_n\|^2\right] = \mathbb{E}\left[\|\tilde{\mathbf{w}}_{n-1}\|^2\right] - 2\mu\mathbb{E}\left[e_{a,n} f(e_n)\right] + \mu^2 \mathbb{E}\left[\|\mathbf{u}_n\|^2 f^2(e_n)\right] + \mathrm{Tr}\left[\mathbf{Q}\right] \tag{17}$$

To derive ξ we consider the following assumptions.

Assumption 1 The a priori error $e_{a,n}$ is zero mean and independent of the measurement noise v_n.

Assumption 2 The filter is long enough such that $e_{a,n}$ is Gaussian, $\|\mathbf{u}_n\|^2$ and is asymptotically uncorrelated with $f^2(e_n)$.

Note that Assumption 2 enables us to rewrite the third term in the right-hand side of (17) as

$$\lim_{n\to\infty} \mathbb{E}\left[\|\mathbf{u}_n\|^2 f^2(e_n)\right] = \mathrm{Tr}\left[\mathbf{R}_u\right] \lim_{n\to\infty} \mathbb{E}\left[f^2(e_n)\right] \tag{18}$$

As in this paper our aim is to evaluate the steady-state tracking performance of MCC algorithm, at the steady-state we have

$$\lim_{n\to\infty} \mathbb{E}\left[\|\tilde{\mathbf{w}}_n\|^2\right] = \lim_{n\to\infty} \mathbb{E}\left[\|\tilde{\mathbf{w}}_{n-1}\|^2\right] \tag{19}$$

So we can simplify (18) at the steady-state as

$$2 \lim_{n\to\infty} \mathbb{E}\left[e_{a,n} f(e_n)\right] = \mu \mathrm{Tr}\left[\mathbf{R}_u\right] \lim_{n\to\infty} \mathbb{E}\left[f^2(e_n)\right] + \mu^{-1} \mathrm{Tr}\left[\mathbf{Q}\right] \tag{20}$$

In the following analysis, we consider two different cases for measurement noise distribution.

3.1 Gaussian Noise

In this case, we assume that measurement noise v_n has zero-mean Gaussian distribution with variance σ_v^2. Then, we can evaluate $\lim_{n\to\infty} \mathbb{E}\left[e_{a,n} f(e_n)\right]$ using the following result from the Price theorem [11]

Lemma 1 *Let x_1 and x_2 be scalar real-valued zero-mean jointly Gaussian random variables and assume functions h and g so that $h(x_1, x_2) = x_1 g(x_2)$. Then, using the Price theorem, the following equality holds*

$$\mathbb{E}\left[h(x_1, x_2)\right] = \mathbb{E}[x_1 x_2]\mathbb{E}\left[\frac{\mathrm{d}g}{\mathrm{d}x_2}\right] \tag{21}$$

Now, we can evaluate $\lim_{n\to\infty} \mathbb{E}\left[e_{a,n} f(e_n)\right]$ with $x_1 = e_{a,n}$ and $x_2 = e_n = e_{a,n} + v_n$ as

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[e_{a,n} f(e_n)\right] &= \lim_{n\to\infty} \mathbb{E}\left[e_{a,n} f(e_{a,n} + v_n)\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[e_{a,n}^2\right] \mathbb{E}\left[f'(e_n)\right] \\
&= \frac{1}{\sqrt{2\pi}\sigma_e} \int_{-\infty}^{\infty} \left(1 - \frac{e_n^2}{\sigma^2}\right) \exp\left(-\frac{e_n^2}{2\sigma^2}\right) \exp\left(-\frac{e_i^2}{2\sigma_e^2}\right) \mathrm{d}e_n \\
&= \frac{1}{\sqrt{2\pi}\sigma_e} \int_{-\infty}^{\infty} \left(1 - \frac{e_n^2}{\sigma^2}\right) \exp\left(-\frac{e_n^2}{2\sigma_{total}^2}\right) \mathrm{d}e_n \\
&= \frac{\sigma^3}{(\xi + \sigma_v^2 + \sigma^2)^{3/2}}
\end{aligned} \tag{22}$$

with $\sigma_{\text{total}}^2 = \frac{\sigma_e^2 \sigma^2}{2\sigma_e^2 + \sigma^2}$. Similarly, we can evaluate $\mathbb{E}[f^2(e_n)]$ as

$$\mathbb{E}\left[f^2(e_n)\right] = \frac{1}{\sqrt{2\pi}\sigma_e} \int_{-\infty}^{\infty} e_n^2 \exp\left(-\frac{e_n^2}{2\sigma_{\text{total}}^2}\right) \mathrm{d}e_n = \frac{\sigma^3(\xi + \sigma_v^2)}{(2\xi + 2\sigma_v^2 + \sigma^2)^{3/2}} \tag{23}$$

Replacing (22) and (23) in (20) gives

$$\frac{2\xi}{(\xi + \sigma_v^2 + \sigma^2)^{3/2}} = \frac{\mu \mathrm{Tr}\left[\mathbf{R}_u\right](\xi + \sigma_v^2)}{(2\xi + 2\sigma_v^2 + \sigma^2)^{3/2}} + \frac{1}{\sigma^3} \mu^{-1} \mathrm{Tr}\left[\mathbf{Q}\right] \tag{24}$$

It must be noted that although the steady-state EMSE satisfies the above equation, a closed-form expression for EMSE cannot be extracted from (24) as it is not an explicit function of step size. However, we can find ξ numerically by solving the following fixed-point equation

$$\xi = \frac{\mu \mathrm{Tr}\left[\mathbf{R}_u\right]}{2} \frac{(\xi + \sigma_v^2)(\xi + \sigma_v^2 + \sigma^2)^{3/2}}{(2\xi + 2\sigma_v^2 + \sigma^2)^{3/2}} + \frac{\mu^{-1} \mathrm{Tr}\left[\mathbf{Q}\right](\xi + \sigma_v^2 + \sigma^2)^{3/2}}{2\sigma^3} \tag{25}$$

Remark 1 As the kernel size $\sigma \to \infty$, the EMSE value ξ given by (25) tends to the EMSE expression of the LMS algorithm, i.e.,

$$\lim_{\sigma\to\infty} \xi = \frac{\mu \mathrm{Tr}\,[\mathbf{R}_u]\,\sigma_v^2 + \mu^{-1}\mathrm{Tr}\,[\mathbf{Q}]}{2 - \mu \mathrm{Tr}\,[\mathbf{R}_u]} = \xi_{\mathrm{LMS}} \tag{26}$$

3.2 Non-Gaussian Noise

To derive the theoretical expression for general non-Gaussian noise data, we consider again the steady-state relation (20). Similar to the Gaussian noise case, we again need to evaluate $\mathbb{E}[e_{a,n} f(e_n)]$ and $\mathbb{E}[f^2(e_n)]$. For the first moment, we have[1]

$$\mathbb{E}\left[e_{a,n} f(e_n)\right] \approx \mathbb{E}\left[e_{a,n}(f(v_n) + f'(v_n)e_{a,n})\right] \approx \xi \mathbb{E}[f'(v_n)] \tag{27}$$

Similarly, for the second moment we have

$$\begin{aligned}\mathbb{E}\left[f^2(e_n)\right] &\approx \mathbb{E}\left[(f(v_n) + f'(v_n)e_{a,n} + \frac{1}{2}f''(v_n)e_{a,n}^2)^2\right] \\ &\approx \mathbb{E}\left[f^2(v_n)\right] + \xi \mathbb{E}\left[f(v_n)f''(v_n) + (f'(v_n))^2\right]\end{aligned} \tag{28}$$

The required terms $f'(v_n)$ and $f''(v_n)$ are given by

$$f'(v_n) = \left(1 - \frac{v_n^2}{\sigma^2}\right)\exp\left(\frac{-v_n^2}{2\sigma^2}\right), \quad f''(v_n) = \left(\frac{v_n^3}{v_n^4} - \frac{3v_n}{\sigma^2}\right)\exp\left(\frac{-v_n^2}{2\sigma^2}\right) \tag{29}$$

Replacing (27) and (28) in (20) result in the desired EMSE expression for general noise distribution as follows

$$\xi = \frac{\mu \mathrm{Tr}\,[\mathbf{R}_u]\,\mathbb{E}\left[v_n^2 \exp\left(\frac{-v_n^2}{\sigma^2}\right)\right] + \mu^{-1}\mathrm{Tr}\,[\mathbf{Q}]}{2\mathbb{E}\left[\left(1 - \frac{v_n^2}{\sigma^2}\right)\exp\left(\frac{-v_n^2}{2\sigma^2}\right)\right] - \mu \mathrm{Tr}\,[\mathbf{R}_u]\,\mathbb{E}\left[\left(1 + \frac{2v_n^4}{\sigma^4} - \frac{5v_n^2}{\sigma^2}\right)\exp\left(\frac{-v_n^2}{\sigma^2}\right)\right]} \tag{30}$$

Remark 2 The given expression for EMSE in (30) is not an increasing monotonic function of μ. This can be easily verified by witting it as

$$\xi = \frac{\mu\mathcal{A} + \mu^{-1}\mathcal{B}}{\mathcal{C} - \mu\mathcal{D}} \tag{31}$$

[1] We use the Taylor expansion of $f(e_n)$ as $f(e_n) = f(v_n) + f'(v_n)e_{a,n} + \frac{1}{2}f''(v_n)e_{a,n}^2 + o(e_{a,n}^2)$.

with

$$\mathcal{A} = \mathrm{Tr}\left[\mathbf{R}_u\right]\mathbb{E}\left[v_n^2 \exp\left(\frac{-v_n^2}{\sigma^2}\right)\right] \tag{32a}$$

$$\mathcal{B} = \mathrm{Tr}\left[\mathbf{Q}\right] \tag{32b}$$

$$\mathcal{C} = 2\mathbb{E}\left[\left(1 - \frac{v_n^2}{\sigma^2}\right)\exp\left(\frac{-v_n^2}{2\sigma^2}\right)\right] \tag{32c}$$

$$\mathcal{D} = \mathrm{Tr}\left[\mathbf{R}_u\right]\mathbb{E}\left[\left(1 + \frac{2v_n^4}{\sigma^4} - \frac{5v_n^2}{\sigma^2}\right)\exp\left(\frac{-v_n^2}{\sigma^2}\right)\right] \tag{32d}$$

By setting the first derivative of the above equation to zero (i.e., $\frac{\mathrm{d}\xi}{\mathrm{d}\mu} = 0$), we obtain the following equation

$$\mathcal{A}\mathcal{C}\mu^2 + 2\mathcal{B}\mathcal{D} - \mathcal{B}\mathcal{C} = 0 \tag{33}$$

The optimum step size for which the ξ takes its minimum is the positive root of (33).

4 Simulation Results

In this section, we provide the simulation results in order to verify the theoretical analysis. To this end, we consider a system identification setup which involves determining the coefficients of an unknown filter with length $M = 10$. The input vector $\mathbf{u}_n$ is generated from a Gaussian process with covariance matrix $\mathbf{R} = \mathbf{I}$. For the non-stationary environment, we assume a random walk model with $\mathbf{Q} = 10^{-4}\mathbf{I}$ with initial vector $\mathbf{w}_0 = \mathbf{0}$. We use a Gaussian kernel with size $\sigma = 2$. The steady-state EMSE curves are generated by performing the MCC algorithm for 10,000 iterations and then averaging the last 200 samples.

For the Gaussian case, we assume that measurement noise in (1) is zero-mean Gaussian noise with variance $\sigma_v^2 = 0.1$. The steady-state curve for Gaussian case is shown in Fig. 1. We can observe that the simulation result matches well with the theoretical expression given by the Eq. (25). We can also see that the EMSE curve is not monotonic increasing function of step size parameter. It is worth noting that similar to LMS algorithm, the requirement of step size for convergence rate, time-varying tracking accuracy and convergence precision (steady-state performance) is contradictory. When μ is too small, convergence rate is slow and the MCC algorithm cannot track the optimal weight variations, which, in turn, results in large steady-state EMSE. Increasing μ improves convergence rate and tracking accuracy and reduces the steady-state EMSE. Finally, as μ increases further, oscillation occurs during the convergence and the steady-state performance deteriorates again.

Note further that for this case an optimum step size value cannot be obtained from (25) as it is not an explicit function of step size.

For the Gaussian case, we consider two different distributions for measurement noise including (1) uniform distribution, where the uniform noise is distributed over $[-1 \ +1]$, and (2) exponential distribution with mean parameter $\lambda = 2$. The other simulation parameters remain unchanged. Figure 2 shows the EMSE curves for both

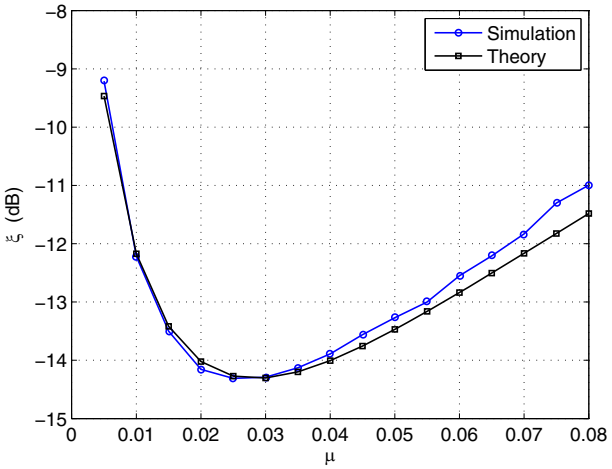


Fig. 1 Comparing theoretical and simulation EMSE for Gaussian noise: uniform distribution (*left*), exponential distribution (*right*)

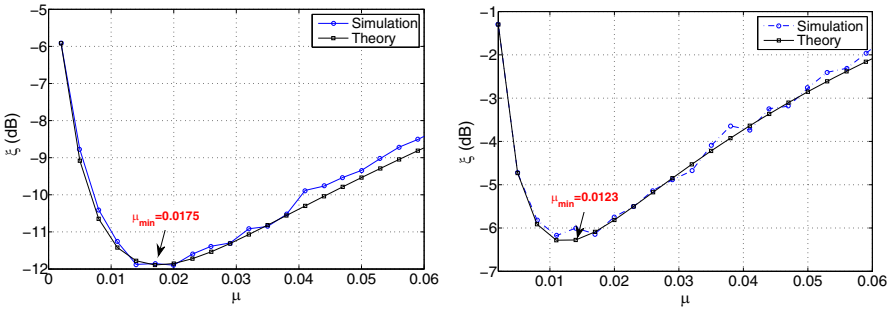


Fig. 2 Comparing theoretical and simulation EMSE for non-Gaussian noise

of the non-Gaussian noise distributions. As it is seen from Fig. 2, the EMSE obtained from the simulation nicely fits the theoretical result obtained earlier. Moreover, for both cases the optimum values given by (33) are very close to the optimum values given by simulations.

5 Conclusions

In this paper, we studied the tracking analysis of MCC algorithm, when the optimum weight vector varies according to a random walk model. Our analysis, which relies on the energy conservation approach revealed that, independent of the measurement noise model, the EMSE curve is not an increasing function of step size parameter. Therefore, in non-stationary environments it is vital to select an appropriate step size to achieve an acceptable performance. The simulation results were in good agreement with the analysis.

References

1. W. Bazzi, A. Rastegarnia, A. Khalili, A robust diffusion adaptive network based on the maximum correntropy criterion, in *Computer Communication and Networks (ICCCN), 2015 24th International Conference on*, pp. 1–4 (2015)
2. B. Chen, J. Wang, H. Zhao, N. Zheng, J. Principe, Convergence of a fixed-point algorithm under maximum correntropy criterion. Signal Process. Lett. IEEE **22**(10), 1723–1727 (2015)
3. B. Chen, L. Xing, J. Liang, N. Zheng, J. Principe, Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion. Signal Process. Lett. IEEE **21**(7), 880–884 (2014)
4. B. Chen, J. Wang, H. Zhao, N. Zheng, J. Principe, Convergence of a fixed-point algorithm under maximum correntropy criterion. Signal Process. Lett. IEEE **22**(10), 1723–1727 (2015)
5. D. Erdogmus, Information theoretic learning: Renyi's entropy and its applications to adaptive system training. Ph.D. dissertation, University of Florida (2002)
6. D. Erdogmus, J.C. Principe, Generalized information potential criterion for adaptive system training. IEEE Trans. Neural Netw. **13**(5), 1035–1044 (2002)
7. S. Han, A family of minimum Renyi's error entropy algorithm for information processing. Ph.D. dissertation, University of Florida (2007)
8. S. Haykin, *Adaptive Filter Theory*, 3rd edn. (Prentice-Hall, Upper Saddle River, NJ, 1996)
9. A. Khalili, A. Rastegarnia, S. Sanei, Robust frequency estimation in three-phase power systems using correntropy-based adaptive filter. Sci. Meas. Technol. IET **9**(8), 928–935 (2015)
10. W. Liu, P.P. Pokharel, J.C. Principe, Correntropy: properties, and applications in non-gaussian signal processing. IEEE Trans. Signal Process. **55**(11), 5286–5298 (2007)
11. R. Price, A useful theorem for nonlinear devices having gaussian inputs. Inf. Theory, IRE Trans. **4**(2), 69–72 (1958)
12. J.C. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives* (Springer, Berlin, 2010)
13. L. Shi, Y. Lin, Convex combination of adaptive filters under the maximum correntropy criterion in impulsive interference. Signal Process. Lett. IEEE **21**(11), 1385–1388 (2014)
14. A. Singh, J. Principe, Using correntropy as a cost function in linear adaptive filters, in *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pp. 2950–2955 (2009)
15. R. Wang, B. Chen, N. Zheng, J. Principe, A variable step-size adaptive algorithm under maximum correntropy criterion, in *Neural Networks (IJCNN), 2015 International Joint Conference on*, pp. 1–5 (2015)
16. B. Widrow, M. Hoff, Adaptive switching circuits. IRE WESCON Conv. Rec. **4**, 96–104 (1960)
17. S. Zhao, B. Chen, J.C. Principe, Kernel adaptive filtering with maximum correntropy criterion, in *International Joint Conference on Neural Networks (IJCNN)*, pp. 2012–2017 (2011)

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and rea