Received 10 February 2024, accepted 23 February 2024, date of publication 26 February 2024, date of current version 1 March 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3370471

RESEARCH ARTICLE

Tracking Analysis of Maximum Versoria Criterion Based Adaptive Filter

AZAM KHALILI[1], (Member, IEEE), AMIR RASTEGARNIA[1], (Member, IEEE), ALI FARZAMNIA[2], (Senior Member, IEEE), SAEID SANEI[3], (Senior Member, IEEE), AND THAMER A. H. ALGHAMDI[4,5]
[1]Department of Electrical Engineering, Malayer University, Malayer 65719-95863, Iran
[2]Faculty of Engineering, University Malaysia Sabah, Kota Kinabalu 88400, Malaysia
[3]Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K.
[4]Electrical Engineering Department, Faculty of Engineering, Al-Baha University, Al Bahah 65779, Saudi Arabia
[5]Wolfson Centre for Magnetics, School of Engineering, Cardiff University, CF24 3AA Cardiff, U.K.

Corresponding author: Thamer A. H. Alghamdi (Alghamdit1@cardiff.ac.uk)

ABSTRACT Recently, maximum Versoria criterion-based adaptive algorithms have been introduced as a new solution for robust adaptive filtering. This paper studies the steady-state tracking analysis of an adaptive filter with maximum Versoria criterion (MVC) in a non-stationary (Markov time-varying) system. Our analysis relies on the energy conservation method. Both Gaussian and general non-Gaussian noise are considered, and for both cases, the closed-form expression for steady-state excess mean square error (EMSE) is derived. Regardless of noise type, unlike the stationary environment, the EMSE curves are not increasing functions of step-size parameter. The validity of the theoretical results is justified via simulation.

INDEX TERMS Adaptive filter, non-stationary, performance analysis, tracking, Versoria.

I. INTRODUCTION

Since the invention of adaptive filters by Widrow and Hoff [1], they have been successfully used in many applications such as noise and echo cancellation, signal prediction, channel estimation, and beamforming [2]. Generally, an adaptive filter adjusts its parameters (filter weights) by optimizing a predefined cost function. The cost function quantifies the difference between the desired and actual outputs of the filter. The most commonly used cost function is the mean squared error (MSE), the statistical average of the squared difference between the desired and actual outputs. While MSE-based adaptive filters provide appealing performance for linear models with Gaussian data, they perform poorly when dealing with nonlinear models and non-Gaussian signals.

To improve the robustness against non-Gaussian signals such as impulsive noise, other criteria beyond MSE should be considered. Recently, information-theoretic adaptive filters have been developed to optimize the filter coefficients based on information-theoretic measures, such as mutual information, error entropy [3], [4], [5], [6], correntropy [7], [8], [9], [10], [11] and risk sensitive loss [12], [13], [14], [15]. These filters can effectively capture higher-order moments of data, which allows them to better adapt to the changes in signal conditions and improve the performance of signal processing tasks. For example, in correntropy-based adaptive filters, the filter parameters are adjusted in a way that the correntropy between the desired and estimated outputs is maximized, which, in turn, effectively reduces the impact of non-Gaussian and impulsive noise.

While maximum correntropy focuses on the similarity between the input and output vectors, maximum Versoria [16], [17], [18], [19], [20] takes a different approach. Versoria is a measure of directionality that captures the underlying structure of the data. Maximum Versoria-based adaptive filtering algorithms aim to maximize the Versoria criterion, quantifying the similarity between the desired and estimated output directions. By maximizing the Versoria criterion, these algorithms can extract valuable directional information from the data, even in non-Gaussian noise. This makes maximum Versoria a powerful tool for applications

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamed M. A. Moustafa.

TABLE 1. Symbols and their descriptions.

Symbol	Description
$(\cdot)^{\mathsf{T}}$	Matrix transposition
$\|\cdot\|^2$	Squared Euclidean norm
$\mathsf{Tr}[\cdot]$	Trace of a matrix
$\mathsf{E}[\cdot]$	Statistical expectation
$\mathbf{I}_M$	$M \times M$ identity matrix
$\mathcal{N}(0, \sigma^2)$	Zero-mean Gaussian distribution with variance σ^2

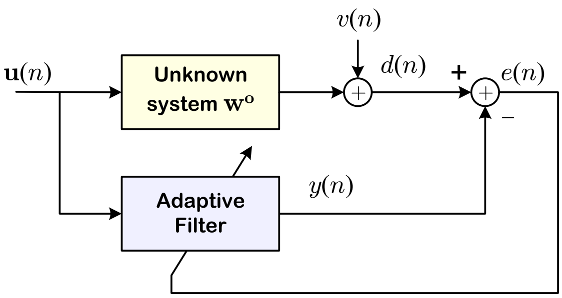

FIGURE 1. An adaptive filter in a system identification setup.

such as beamforming, where the direction of arrival needs to be estimated accurately.

In the context of Versoria-based adaptive filtering algorithms, the tracking analysis provides valuable information about their ability to capture the underlying structure of a non-stationary (Markov time-varying) system [15], [21], [22]. Such non-stationarities arise in many practical applications such as acoustics, communication, control, and power systems. By analyzing the tracking performance, the parameters of an adaptive filter can be adjusted to optimize their performance for specific applications. This paper examines the tracking analysis of adaptive filters with maximum Versoria criterion (MVC). We explore the steady-state behavior of these filters and analyze their performance in both Gaussian and non-Gaussian noise environments. Our analysis is based on the energy conservation method, allowing us to derive closed-form expressions for steady-state excess mean square error (EMSE) for both cases. Furthermore, we investigate the impact of the step-size parameter on the EMSE curves and provide insights into selecting an optimal step-size for different scenarios.

Throughout the paper, we use normal lowercase letters for scalars, bold lowercase letters for column vectors, and bold uppercase letters for matrices. A list of main symbols used throughout the paper is provided in Table 1.

II. MVC ALGORITHM

To begin, without loss of generality, we consider an adaptive filter in a system identification setup, as shown in Fig. 1. The objective is to estimate an unknown weight vector (parameter of parameters) $\mathbf{w}^o \in \mathbb{R}^{M\times 1}$ by minimizing the error between desired signal $d_n \in \mathbb{R}$ and actual output $y(n)$. For a stationary environment (with constant $\mathbf{w}^o$) and linear model assumption, the desired signal d_n at any time instant n is given by

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

where in this model, $\mathbf{u}_n \in \mathbb{R}^{1\times M}$ denote the input data with covariance matrix $\mathbf{R}_\mathbf{u}$. Moreover, $v_n \in \mathbb{R}$ are samples of zero-mean observation (measurement) noise, which are assumed to be independent, identically distributed, and independent of the input signal $\mathbf{u}_n$.

One possible way to estimate $\mathbf{w}^o$ is to use an LMS adaptive filter, which operates by minimizing the mean-square error between the desired signal and the filtered output as follows

$$J_{\mathrm{MSE}}(\mathbf{w}) \triangleq \mathsf{E}\left[e_n^2\right] \tag{2}$$

where e_n is the instantaneous error signal given by

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

The LMS algorithm adjusts the filter coefficients in proportion to the negative gradient of the MSE cost function and solves (2) iteratively as

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

where $\mu > 0$ is a suitably chosen step-size parameter. As mentioned earlier, while the least mean squares (LMS) algorithm is widely used for its simplicity and stability, it may exhibit performance degradation in non-Gaussian environments.

To address these limitations, Versoria has recently been introduced as the cost function to be maximized for adaptive filters, which exploits higher-order data moments with low computational complexity. The cost function in the standard form of the MVC algorithm is defined as

$$\arg\max_{\mathbf{w}} J(\mathbf{w}) \triangleq \mathsf{E}\left[\frac{1}{1+\tau e_n^2}\right] \tag{5}$$

where $\tau > 0$ represents Versoria parameter. Using the steepest ascent method to solve (5) and replacing the statistical expectation with instantaneous mean approximation yields the standard MVC update equation as

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu \frac{\mathbf{u}_n^{\mathsf{T}} e_n}{(1+\tau e_n)^2} \tag{6}$$

Let further denote by $f(e_n)$ a function of the error signal e_n in (6) as

$$f(e_n) \triangleq \frac{e_n}{\left(1+\tau e_n^2\right)^2} \tag{7}$$

Replacing $f(e_n)$ in (6) gives

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mu f(e_n) \mathbf{u}_n^{\mathsf{T}} \tag{8}$$

When $\tau \rightarrow 0$, the MVC algorithm in (6) tends to the LMS algorithm. In [23], steady-state analysis of the MVC algorithm for stationary data has been studied. Although that analysis is required to understand the behavior of adaptive filters with MVC, it does not give any insight into the performance of the MVC algorithm in a non-stationary (time-varying) system. Thus, a tracking analysis that examines the performance of the MVC algorithm for the non-stationary systems is necessary.

III. TRACKING ANALYSIS

To proceed with the analysis, we first introduce the assumptions and definitions used in this paper.

A. ASSUMPTIONS AND DEFINITIONS

In our analysis, a simple Markov process is used to describe the time evolution of the optimal weights $\{\mathbf{w}_n^o\}$, which means that $\mathbf{w}_n^o$ varies according to a random-walk model as follows:

$$\mathbf{w}_n^o = \mathbf{w}_{n-1}^o + \boldsymbol{\theta}_n \tag{9}$$

where $\boldsymbol{\theta}_n$ a zero-mean white noise with covariance matrix $\boldsymbol{\Theta} = \mathsf{E}\left[\boldsymbol{\theta}_n\boldsymbol{\theta}_n^\mathsf{T}\right]$.

For future reference, the following error signals are defined:

$$\tilde{\mathbf{w}}_n \triangleq \mathbf{w}_n^o - \mathbf{w}_n, \quad \text{weight error vector} \tag{10}$$

$$e_{a,n} \triangleq \mathbf{u}_n\tilde{\mathbf{w}}_n, \quad \text{a-priori error signal} \tag{11}$$

Assumption 1.

1) For a priori error signal $e_{a,n}$, we have
   - $\mathsf{E}\left[e_{a,n}\right] = 0$,
   - $g(e_{a,n})$ is independent of v_n for any any given function $g(\cdot)$.
2) At the steady-state i.e. $n \to \infty$, $\|\mathbf{u}_n\|^2$ is independent of e_n.
3) For all $t < n$, $\boldsymbol{\theta}_n$ is independent from $\{d_t\}$ and $\{\mathbf{u}_t\}$ and also $\{\mathbf{w}_0\}$.

We consider the steady-state EMSE as the performance metric, which is defined as

$$\xi = \lim_{n\to\infty} \mathsf{E}\left[e_{a,n}^2\right] \tag{12}$$

Our analysis relies on the energy-conservation approach which shows that certain a-priori and a-posteriori errors maintain an energy balance (known as the variance relation) for all time instants. To derive the variance relation for the MVC algorithm with non-stationary data, we start by taking the Euclidean norm on the update equation in (8) and apply the statistical expectation to the resultant expression and using Assumption 1, we arrive at the variance relation (16). Then, the closed-form expressions for steady-state EMSE are extracted from the variance relation (16) for two different cases for the measurement noise distributions.

B. ANALYSIS

By subtracting $\mathbf{w}_n^o$ from both sides of (8) and using the definition of $\mathbf{w}_n^o$ in (9) we have

$$\tilde{\mathbf{w}}_n = \tilde{\mathbf{w}}_{n-1} + \boldsymbol{\theta}_n - \mu\mathbf{u}_n^\mathsf{T} f(e_n) \tag{13}$$

Taking Euclidean norm on both sides of (13) and applying $\mathsf{E}[\cdot]$ on the resultant equation we obtain

$$\begin{aligned}\mathsf{E}\left[\|\tilde{\mathbf{w}}_n\|^2\right] &= \mathsf{E}\left[\|\tilde{\mathbf{w}}_{n-1}\|^2\right] - 2\mu\mathsf{E}\left[\mathbf{u}_n\tilde{\mathbf{w}}_{n-1}f(e_n)\right] \\ &\quad + \mu^2\mathsf{E}\left[\|\mathbf{u}_n\|^2 f^2(e_n)\right] + \mathsf{E}\left[\|\boldsymbol{\theta}_n\|^2\right] \\ &\quad + \underbrace{\mathsf{E}\left[\tilde{\mathbf{w}}_{n-1}^\mathsf{T}\boldsymbol{\theta}_n\right]}_{①} + \underbrace{\mathsf{E}\left[\boldsymbol{\theta}_n^\mathsf{T}\tilde{\mathbf{w}}_{n-1}\right]}_{②} \\ &\quad - 2\mu\underbrace{\mathsf{E}\left[\boldsymbol{\theta}_n^\mathsf{T}\mathbf{u}_n^\mathsf{T} f(e_n)\right]}_{③}\end{aligned} \tag{14}$$

where ③$= 0$ due to Assumption 1. The terms denoted by ① and ② have been calculated in Appendix A as

$$\mathsf{E}\left[\tilde{\mathbf{w}}_{n-1}^\mathsf{T}\boldsymbol{\theta}_n\right] = \mathsf{E}\left[\boldsymbol{\theta}_n^\mathsf{T}\tilde{\mathbf{w}}_{n-1}\right] = 0 \tag{15}$$

Additionally, using $\mathsf{E}\left[\|\boldsymbol{\theta}_n\|^2\right] = \mathsf{E}\left[\mathsf{Tr}\left[\boldsymbol{\theta}_n\boldsymbol{\theta}_n^\mathsf{T}\right]\right] = \mathsf{Tr}[\boldsymbol{\Theta}]$ and replacing (15) in (14) we obtain

$$\begin{aligned}\mathsf{E}\left[\|\tilde{\mathbf{w}}_n\|^2\right] &= \mathsf{E}\left[\|\tilde{\mathbf{w}}_{n-1}\|^2\right] - 2\mu\mathsf{E}\left[e_{a,n}f(e_n)\right] \\ &\quad + \mu^2\mathsf{E}\left[\|\mathbf{u}_n\|^2 f^2(e_n)\right] + \mathsf{Tr}[\boldsymbol{\Theta}]\end{aligned} \tag{16}$$

At the steady-state, where $n \to \infty$ an acceptable approximation is $\mathsf{E}\left[\|\tilde{\mathbf{w}}_n\|^2\right] \approx \mathsf{E}\left[\|\tilde{\mathbf{w}}_{n-1}\|^2\right]$, which in turn simplifies (16) as

$$2\lim_{n\to\infty}\mathsf{E}\left[e_{a,n}f(e_n)\right] = \mu\mathsf{Tr}[\mathbf{R_u}]\lim_{n\to\infty}\mathsf{E}\left[f^2(e_n)\right] + \mu^{-1}\mathsf{Tr}[\boldsymbol{\Theta}] \tag{17}$$

To proceed with the analysis, two different cases for the measurement noise distribution are considered.

C. GAUSSIAN NOISE

In this case, the observation noise v_n is assumed to follow a Gaussian distribution with zero mean and variance of σ_v^2. Then, the following Lemma from Price Theorem [24] can be used to evaluate $\lim_{n\to\infty}\mathsf{E}\left[e_{a,n}f(e_n)\right]$:

Lemma 1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be scalar real-valued zero-mean jointly Gaussian random variables. For any Borel function $h(\cdot)$ we have*

$$\mathsf{E}[\mathcal{X}h(\mathcal{Y})] = \frac{\mathsf{E}[\mathcal{X}\mathcal{Y}]}{\mathsf{E}\left[\mathcal{Y}^2\right]}\mathsf{E}[\mathcal{Y}h(\mathcal{Y})] \tag{18}$$

Setting $\mathcal{X} = e_{a,n}$, $\mathcal{Y} = e_n$ and $h(\cdot) = f(\cdot)$, $\lim_{n\to\infty}\mathsf{E}\left[e_{a,n}f(e_n)\right]$: can be calculated as

$$\begin{aligned}\lim_{n\to\infty}\mathsf{E}\left[e_{a,n}f(e_n)\right] &= \lim_{n\to\infty}\frac{\mathsf{E}\left[e_{a,n}e_n\right]\mathsf{E}[e_nf(e_n)]}{\mathsf{E}\left[e_n^2\right]} \\ &= \frac{\xi}{\sigma_e^2}\lim_{n\to\infty}\mathsf{E}[e_nf(e_n)]\end{aligned} \tag{19}$$

where σ_e^2 denotes the variance of e_n. Note that in (19) we used $e_n = e_{a,n} + v_n$ and Assumption 1. Replacing (19) in (17) yields

$$\xi = \frac{\sigma_e^2\left(\mu\mathsf{Tr}[\mathbf{R_u}]\lim_{n\to\infty}\mathsf{E}\left[f^2(e_n)\right] + \mu^{-1}\mathsf{Tr}[\boldsymbol{\Theta}]\right)}{2\lim_{n\to\infty}\mathsf{E}[e_nf(e_n)]} \tag{20}$$

When $e(n)$ is Gaussian distributed (per our assumption in this subsection), it is shown in [23] that $\mathsf{E}\left[f^2(e_n)\right]$ is given by

$$\alpha \triangleq \mathsf{E}\left[f^2(e_n)\right] = \phi + \tau\frac{\partial\phi}{\partial\tau} + \frac{\tau^2}{6}\frac{\partial^2\phi}{\partial^2\tau} \tag{21}$$

where

$$\phi \triangleq \mathsf{E}[e_n f(e_n)] \tag{22}$$

Using the moments in (21) and (22) in (20) we have

$$\xi = \frac{\sigma_e^2\left(\mu\mathsf{Tr}[\mathbf{R_u}]\alpha + \mu^{-1}\mathsf{Tr}[\mathbf{\Theta}]\right)}{2\phi} \tag{23}$$

Finally, by defining $\beta \triangleq \alpha/\phi$ and $\mu_\theta \triangleq \phi\mu$ and replacing $\sigma_e^2 = \xi + \sigma_v^2$ in (23) we obtain

$$\xi = \frac{\sigma_v^2\left(\beta\mu\mathsf{Tr}[\mathbf{R_u}] + \mu_\theta^{-1}\mathsf{Tr}[\mathbf{\Theta}]\right)}{2 - \beta\mu\mathsf{Tr}[\mathbf{R_u}] - \mu_\theta^{-1}\mathsf{Tr}[\mathbf{\Theta}]} \tag{24}$$

Remark 1. For stationary data with $\boldsymbol{\theta} = 0$, ξ in (24) changes to

$$\xi = \frac{\sigma_v^2\beta\mu\mathsf{Tr}[\mathbf{R_u}]}{2 - \beta\mu\mathsf{Tr}[\mathbf{R_u}]} \tag{25}$$

which is the theoretical steady-state EMSE of the standard MVC algorithm derived in [23]. Moreover, as $\tau \to 0$, β approaches 1 and ξ in (24) tends to

$$\xi = \xi_{\text{LMS}} = \frac{\sigma_v^2\mu\mathsf{Tr}[\mathbf{R_u}]}{2 - \mu\mathsf{Tr}[\mathbf{R_u}]}, \quad \text{as } \tau \to 0 \tag{26}$$

which is the theoretical steady-state EMSE of LMS adaptive filter [2].

D. NON-GAUSSIAN NOISE

To calculate the required moments in (17) for general non-Gaussian noise, we first need to obtain suitable approximations for $\mathsf{E}\left[e_{a,n}f(e_n)\right]$ and $\mathsf{E}\left[f^2(e_n)\right]$. To this end, we use the Taylor series expansion for $f(e_n)$ as

$$f(e_n) = f(v_n) + f'(v_n)e_{a,n} + \frac{1}{2}f''(v_n)e_{a,n}^2 + o(e_{a,n}^2) \tag{27}$$

Replacing (27) in $\mathsf{E}\left[e_{a,n}f(e_n)\right]$ (in (17)), using Assumption 1 and ignoring higher order terms gives:

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

Similarly, for $\mathsf{E}\left[f^2(e_n)\right]$ in right-hand side of (17) we have:

$$\mathsf{E}\left[f^2(e_n)\right] = \mathsf{E}\left[f^2(v_n)\right] + \mathsf{E}\left[f(v_n)f''(v_n) + \xi\left(f'(v_n)\right)^2\right] \tag{29}$$

By substituting (28) and (29) in (17), the following expression is obtained:

$$\begin{aligned}&2\mathsf{E}\left[f'(v_n)\right]\xi\\ &= \mu\mathsf{Tr}[\mathbf{R_u}]\left(\mathsf{E}\left[f^2(v_n)\right] + \xi\mathsf{E}\left[f(v_n)f''(v_n) + \left(f'(v_n)\right)^2\right]\right)\\ &\quad + \mu^{-1}\mathsf{Tr}[\mathbf{\Theta}]\end{aligned} \tag{30}$$

Solving (30) to obtain ξ yields

$$\xi = \frac{\mu\left(\mathsf{Tr}[\mathbf{R_u}]\,\mathsf{E}\left[f^2(v_n)\right]\right) + \mu^{-1}\mathsf{Tr}[\mathbf{\Theta}]}{2\mathsf{E}[f'(v_n)] - \mu\left(\mathsf{Tr}[\mathbf{R_u}]\,\mathsf{E}\left[f(v_n)f''(v_n) + (f'(v_n))^2\right]\right)} \tag{31}$$

Using (7), $f'(v_n)$ and $f''(v_n)$ are obtained as:

$$f'(v_n) = \frac{(1 - 3\tau v_n^2)}{(1 + \tau v_n^2)^3}, \quad f''(v_n) = \frac{-12\tau v_n(1 - \tau v_n^2)}{(1 + \tau v_n^2)^4} \tag{32}$$

Using (32) in (31) we have

$$\xi = \frac{\mu\left(\mathsf{Tr}[\mathbf{R_u}]\,\mathsf{E}\left[\frac{v_n^2}{(1+\tau v_n^2)^4}\right]\right) + \mu^{-1}\mathsf{Tr}[\mathbf{\Theta}]}{2\mathsf{E}\left[\frac{(1-3\tau v_n^2)}{(1+\tau v_n^2)^3}\right] - \mu\left(\mathsf{Tr}[\mathbf{R_u}]\,\mathsf{E}\left[\frac{1-18\tau v_n^2+21\tau^2 v_n^4}{(1+\tau v_n^2)^6}\right]\right)} \tag{33}$$

Remark 2. As can be seen from (24) and (33), unlike the stationary case [23], the obtained expressions for EMSE for both Gaussian and non-Gaussian cases are not monotonic functions of step-size parameter. This means that reducing step-size value does not necessarily result in a smaller steady-state EMSE value. So, selecting an appropriate step-size parameter is crucial to achieving acceptable performance. For example, the given expression for EMSE in (33) can be written as

$$\xi = \frac{\mu a_1 + \mu^{-1}a_2}{b_1 - \mu b_2} \tag{34}$$

with

$$a_1 = \mathsf{Tr}[\mathbf{R_u}]\,\mathsf{E}\left[\frac{v_n^2}{(1 + \tau v_n^2)^4}\right] \tag{35a}$$

$$a_2 = \mathsf{Tr}[\mathbf{\Theta}] \tag{35b}$$

$$b_1 = 2\mathsf{E}\left[\frac{(1 - 3\tau v_n^2)}{(1 + \tau v_n^2)^3}\right] \tag{35c}$$

$$b_2 = \mathsf{Tr}[\mathbf{R_u}]\,\mathsf{E}\left[\frac{1 - 18\tau v_n^2 + 21\tau^2 v_n^4}{(1 + \tau v_n^2)^6}\right] \tag{35d}$$

Setting $\partial\xi/\partial\mu = 0$ yields

$$a_1 b_1\mu^2 + 2a_2 b_2\mu - a_2 b_1 = 0 \tag{36}$$

The positive root of (36) is the optimum step-size parameter that achieves the minimum value of EMSE. It should be noted that the optimal step-size parameter requires information (statistical moments), which in practical applications is not available in advance. One possible way is to use instantaneous approximations for needed statistical moments in the solution of (36) and use variable step-size μ_n to achieve a sub-optimum solution.

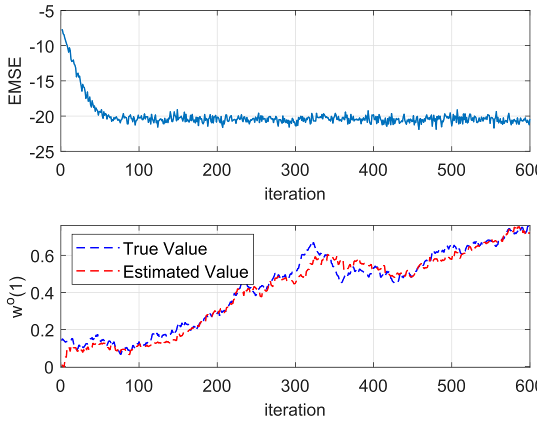

FIGURE 2. **The EMSE learning curve (top) and estimation performance (bottom) for Gaussian noise.**

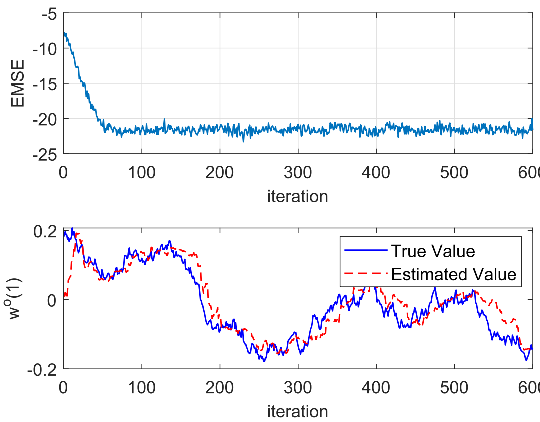

FIGURE 3. **The EMSE learning curve (top) and estimation performance (bottom) for uniform noise.**

IV. SIMULATION RESULTS

In this section, simulation results are presented to verify the theoretical analysis. To this end, we consider the system identification setup, as shown in Fig. 1. For unknown system we set $M = 6$ and initial value $\mathbf{w}_0^o = (1/M)\mathbb{1}_M$. For input vectors we assume that $\mathsf{E}[\mathbf{u}_n\mathbf{u}_t] = 0$ when $n \neq t$, and $\{\mathbf{u}_n\}$ are generated from a Gaussian process with covariance matrix $\mathbf{R}_\mathbf{u} = \mathbf{I}_M$. For random-walk model (9) with $\mathbf{\Theta} = \sigma_\theta^2 \mathbf{I}_M$. Three different distributions for the measurement noise v_n are considered, including:

- $v_n \sim \mathcal{N}(0, 0.01)$, i.e. zero-mean Gaussian noise with variance $\sigma_v^2 = 0.01$.
- uniform noise is distributed over $[0, +1]$.
- $v_n \sim (1-p)\mathcal{N}(0, 0.01) + p\mathcal{N}(0, 10)$, i.e., impulsive noise condition with white Gaussian background noise, and impulse probability of occurrence $p = 0.1$.

The tracking performance of the MVC algorithm is shown in Figs. 2-4 for $\mu = 0.08$, $\sigma_\theta^2 = 10^{-4}$ and $\tau = 1$. These figures depict the EMSE learning curves for true unknown parameter and its estimation for different measurement noises.

From these figures, it is clear that for a suitably chosen step-size parameter, the MVC algorithm can track the variation of the unknown parameter.

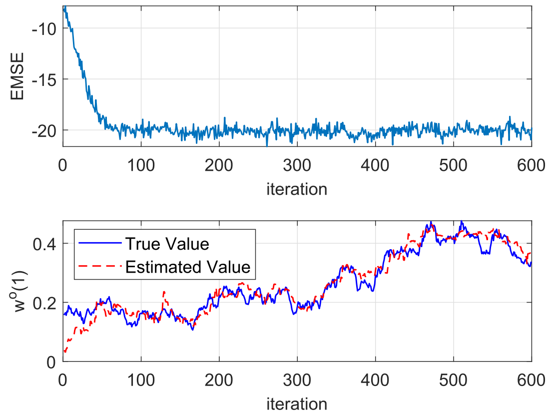

FIGURE 4. **The EMSE learning curve (top) and estimation performance (bottom) for impulsive noise.**

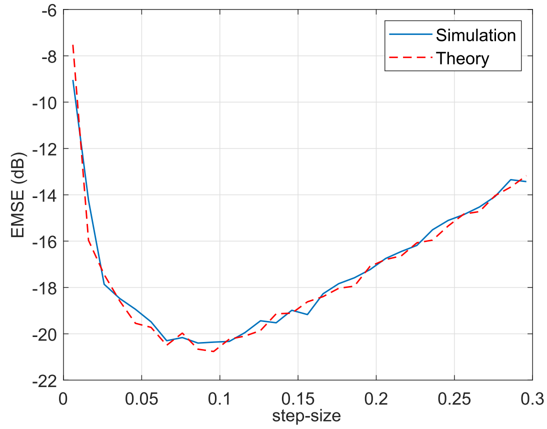

FIGURE 5. **Theoretical and simulated EMSE values for Gaussian measurement noise.**

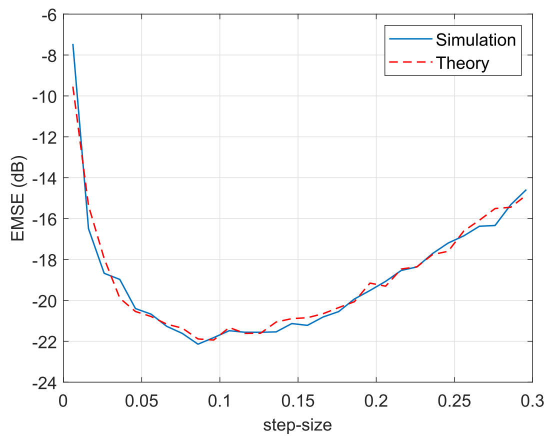

FIGURE 6. **Theoretical and simulated EMSE values for uniform measurement noise.**

To verify the accuracy of the derived expression for EMSE in (24) and (33), we compare the theoretical EMSE values with those of the simulated ones for different values of the step-size parameter. To this end, the theoretical EMSE values are first generated using expressions in (24) and (33). Derivation of a closed-form expression for every required

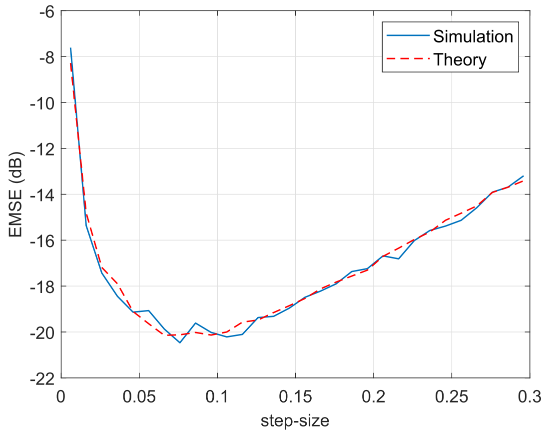

FIGURE 7. Theoretical and simulated EMSE values for impulsive measurement noise.

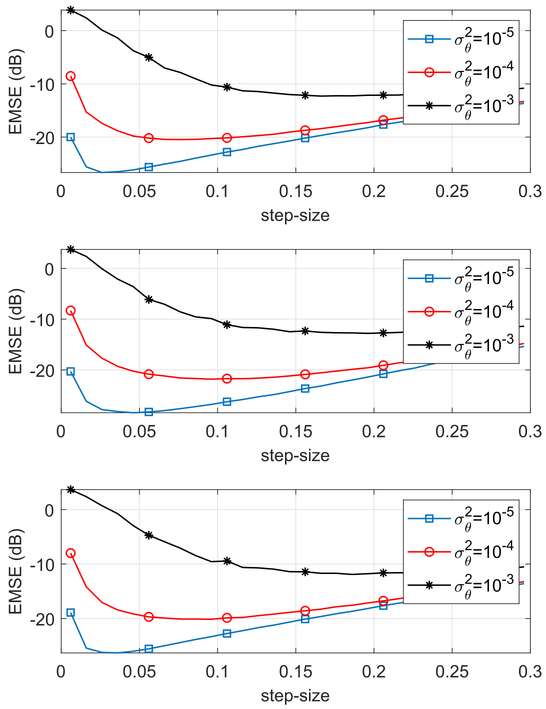

FIGURE 8. Steady-state EMSE values for different values of σ_θ^2 and different noise distributions.

moment in (24) and (33) is not mathematically tractable. So, to obtain an accurate estimate for each required moment in (24) and (33), the ensemble averages of over 5000 trails have been replaced with statistical expectations. The MVC algorithm is carried out for 2000 iterations to reach its steady and then averages the last 100 samples to obtain simulated values. Figs. 5-7 show the theoretical and simulated EMSE values regarding the step-size parameter for different measurement noises.

The simulation results are in close agreement with those by the theoretical analysis, confirming the accuracy of the derived expressions for the EMSE in (24) and (33). Moreover, regardless of measurement noise distribution, the EMSE

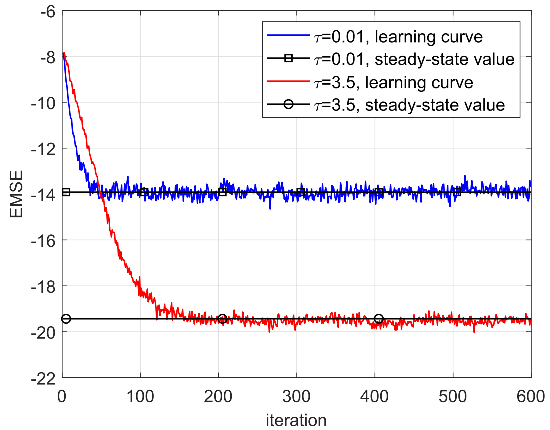

FIGURE 9. The EMSE learning curve for different values of τ for impulsive noise.

curve is not a monotonic increasing function of the step-size parameter. This result emphasizes the importance of selecting an appropriate step-size value to achieve optimal tracking and convergence in non-stationary scenarios.

The effect of $\boldsymbol{\theta}$ on the steady-state tracking performance of the MVC algorithm has been considered in Fig. 8, where the steady-state EMSE values for different values of σ_θ^2 and different noise distributions are plotted. From Fig. 8 it is seen that for all noise distribution, when σ_θ^2 (as a source of noise) increases, the tracking ability of the MVC algorithm decreases, which in turn, increases the steady-state value.

Fig. 9 shows the effect of τ on the learning rate and steady-state tracking performance of the MVC algorithm for impulsive noise distribution have been considered. From Fig. 9 we can see that a larger value of τ yields a lower steady-state misalignment but a slower convergence rate, and vice versa.

V. CONCLUSION

In conclusion, the tracking analysis of adaptive filter with maximum Versoria criterion provides valuable insights into its performance in non-stationary environments. The analysis, based on the energy conservation approach, allows us to understand the behavior of these filters and optimize their performance by selecting an appropriate step-size. The derived theoretical expressions for the steady-state EMSE provide a framework for evaluating the performance of adaptive filters in both Gaussian and non-Gaussian noise environments. The simulation results validate the analysis and highlight the importance of considering the noise characteristics in optimizing the filter's performance.

APPENDIX A CALCULATION OF THE TERMS IN (14)

To calculate ①, it should be noted that $\tilde{\mathbf{w}}_{n-1}$ can be rewritten as

$$\tilde{\mathbf{w}}_{n-1} = \mathbf{w}^o_{n-1} - \mathbf{w}_{n-1} = \left(\mathbf{w}^o_{-1} + \sum_{j=0}^{n-1} \boldsymbol{\theta}_j\right) - \mathbf{w}_{n-1} \quad \text{(A1)}$$

So we have

$$\mathsf{E}\left[\tilde{\mathbf{w}}_{n-1}^{\mathsf{T}}\boldsymbol{\theta}_n\right]=\mathsf{E}\left[\left(\mathbf{w}_{-1}^{o}+\sum_{j=0}^{n-1}\boldsymbol{\theta}_j\right)^{\mathsf{T}}\boldsymbol{\theta}_n\right]-\mathsf{E}[\mathbf{w}_{n-1}\boldsymbol{\theta}_n] \quad (A2)$$

Note that, $\mathsf{E}[\boldsymbol{\theta}_n\boldsymbol{\theta}_t]=0$ for $t<n$ so we have

$$\mathsf{E}\left[\left(\mathbf{w}_{-1}^{o}+\sum_{j=0}^{n-1}\boldsymbol{\theta}_j\right)^{\mathsf{T}}\boldsymbol{\theta}_n\right]=0 \quad (A3)$$

Moreover, $\mathbf{w}_{n-1}$ is independent of $\boldsymbol{\theta}_n$, which means that $\mathsf{E}[\mathbf{w}_{n-1}\boldsymbol{\theta}_n]=0$. Similarly, we have ② $=0$.

REFERENCES

[1] B. Widrow and M. E. Hoff, "Adaptive switching circuits," in *Proc. IRE WESCON Conv. Rec.*, New York, NY, USA, 1960, pp. 96–104.
[2] A. H. Sayed, *Adaptive Filters*. Hoboken, NJ, USA: Wiley, 2008.
[3] Z. Wu, S. Peng, W. Ma, B. Chen, and J. Principe, "Minimum error entropy algorithms with sparsity penalty constraints," *Entropy*, vol. 17, no. 5, pp. 3419–3437, May 2015.
[4] G. Wang, X. Yang, L. Wu, Z. Fu, X. Ma, Y. He, and B. Peng, "A kernel recursive minimum error entropy adaptive filter," *Signal Process.*, vol. 193, Apr. 2022, Art. no. 108410.
[5] G. Wang, B. Peng, Z. Feng, X. Yang, J. Deng, and N. Wang, "Adaptive filtering based on recursive minimum error entropy criterion," *Signal Process.*, vol. 179, Feb. 2021, Art. no. 107836.
[6] G. Qian, J. Liu, C. Qiu, H. Ho-Ching Iu, J. Qian, and S. Wang, "Minimum total complex error entropy for adaptive filter," *Expert Syst. Appl.*, vol. 237, Mar. 2024, Art. no. 121522.
[7] A. Singh and J. C. Principe, "Using correntropy as a cost function in linear adaptive filters," in *Proc. Int. Joint Conf. Neural Netw.*, 2009, pp. 2950–2955.
[8] S. Peng, B. Chen, L. Sun, W. Ser, and Z. Lin, "Constrained maximum correntropy adaptive filtering," *Signal Process.*, vol. 140, pp. 116–126, Nov. 2017.
[9] X. Zhang, K. Li, Z. Wu, Y. Fu, H. Zhao, and B. Chen, "Convex regularized recursive maximum correntropy algorithm," *Signal Process.*, vol. 129, pp. 12–16, Dec. 2016.
[10] H. Zhao, Y. Gao, and Y. Zhu, "Robust subband adaptive filter algorithms-based mixture correntropy and application to acoustic echo cancellation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 31, pp. 1223–1233, 2023.
[11] H. Zhao and Z. Cao, "Robust generalized maximum Blake–Zisserman total correntropy adaptive filter for generalized Gaussian noise and noisy input," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 11, pp. 6757–6765, May 2023.
[12] B. Chen and R. Wang, "Risk-sensitive loss in kernel space for robust adaptive filtering," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 921–925.
[13] G. Qian, F. Dong, and S. Wang, "Robust constrained minimum mixture kernel risk-sensitive loss algorithm for adaptive filtering," *Digit. Signal Process.*, vol. 107, Dec. 2020, Art. no. 102859.
[14] W. Wang, H. Zhao, L. Lu, and Y. Yu, "Robust nonlinear adaptive filter based on kernel risk-sensitive loss for bilinear forms," *Circuits, Syst., Signal Process.*, vol. 38, no. 4, pp. 1876–1888, Apr. 2019.
[15] A. Rastegarnia, P. Malekian, A. Khalili, W. M. Bazzi, and S. Sanei, "Tracking analysis of minimum kernel risk-sensitive loss algorithm under general non-Gaussian noise," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 7, pp. 1262–1266, Jul. 2019.
[16] F. Huang, J. Zhang, and S. Zhang, "Maximum versoria criterion-based robust adaptive filtering algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 10, pp. 1252–1256, Oct. 2017.
[17] S. Radhika, F. Albu, and A. Chandrasekar, "Proportionate maximum versoria criterion-based adaptive algorithm for sparse system identification," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 3, pp. 1902–1906, Mar. 2022.
[18] S. S. Bhattacharjee, D. Ray, and N. V. George, "Adaptive modified versoria zero attraction least mean square algorithms," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 12, pp. 3602–3606, Dec. 2020.
[19] A. Sharma and B. S. Rajpurohit, "Maximum versoria criteria based adaptive filter algorithm for power quality intensification," in *Proc. IEEE 9th Power India Int. Conf. (PIICON)*, Feb. 2020, pp. 1–5.
[20] S. Jain and S. Majhi, "Zero-attracting kernel maximum versoria criterion algorithm for nonlinear sparse system identification," *IEEE Signal Process. Lett.*, vol. 29, pp. 1546–1550, 2022.
[21] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.
[22] A. Khalili, A. Rastegarnia, M. K. Islam, and T. Y. Rezaii, "Steady-state tracking analysis of adaptive filter with maximum correntropy criterion," *Circuits, Syst., Signal Process.*, vol. 36, no. 4, pp. 1725–1734, Apr. 2017.
[23] S. Radhika, F. Albu, and A. Chandrasekar, "Steady state mean square analysis of standard maximum versoria criterion based adaptive algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 4, pp. 1547–1551, Apr. 2021.
[24] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IEEE Trans. Inf. Theory*, vol. IT-4, no. 2, pp. 69–72, Jun. 1958.

AZAM KHALILI (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2011. In 2011, she joined the Department of Electrical Engineering, Malayer University, as an Assistant Professor. Her current research interests include theory and methods for adaptive filtering and distributed adaptive estimation.

AMIR RASTEGARNIA (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2011. In 2011, he joined the Department of Electrical Engineering, Malayer University, as an Assistant Professor. His current research interests include theory and methods for adaptive filtering, distributed optimization, and signal processing for biomedical engineering.

ALI FARZAMNIA (Senior Member, IEEE) received the Ph.D. degree in electrical engineering telecommunication engineering from Universiti Teknologi Malaysia (UTM), in 2014. He has been appointed as a Senior Lecturer (an Assistant Professor) with the Electrical and Electronic Engineering Program, Faculty of Engineering, University Malaysia Sabah (UMS), since 2014. He has secured several research grants with numerous dedicated collaborative research partners. His research interests include wireless communications, signal processing, network coding, information theory, and bio-medical signal processing. He is a member of IET. He is a Chartered Engineer (C.Eng.), U.K.

SAEID SANEI (Senior Member, IEEE) received the Ph.D. degree from Imperial College London, U.K. Currently, he is a Visiting Professor with the Department of Electrical and Electronic Engineering, Imperial College London. His current research interests include the application of adaptive and nonlinear signal processing, subspace analysis, and tensor factorization to EEG, speech, and medical images. He has published five monograms, several edited books and book chapters, and approximately 430 peer reviewed articles. He has served as an Associate Editor for IEEE Signal Processing Letters, *IEEE Signal Processing Magazine*, and *Journal of Computational Intelligence and Neuroscience*. He has also served as a member for IEEE MLSP and SPTM Committees.

THAMER A. H. ALGHAMDI received the B.Sc. degree from Al-Baha University, Al Bahah, Saudi Arabia, in 2012, the M.Sc. degree from Northumbria University, Newcastle, U.K., in 2016, and the Ph.D. degree from Cardiff University, Cardiff, U.K., in 2023. Then, he was a Power Distribution Engineer for the Saudi Electricity Company (SEC), until 2013. He was a Lecturer Assistant with Al-Baha University, from 2016 to 2018, where he is currently an Assistant Professor in electrical power engineering. His main research interests include power systems, power quality, the integration of renewables, and AI applications in electrical engineering.