New symmetric decorrelating set-membership NLMS adaptive algorithms for blind speech intelligibility enhancement

Mohamed Djendi¹ · Abdelhak Cheffi¹

Received: 23 August 2019 / Accepted: 5 November 2019 / Published online: 9 November 2019
© Springer Nature Switzerland AG 2019

Abstract

This paper addresses the problem of speech enhancement and acoustic noise reduction by partial and set-membership adaptive algorithms combined with the symmetric decorrelating adaptive (SAD) algorithm structure. In this paper, we propose two new adaptive algorithms based on set-membership principle that improve the original set-membership algorithm behavior in speech enhancement applications. The first proposed algorithm (called Proposed 1), is based on the combination of the SAD algorithm structure with a smart control that uses decorrelating properties between the output and the mixing signal to control and update the SAD adaptive filters. The second proposed algorithm (called Proposed 2) is a modification of Proposed 1 and based on a new regularization relations of the SAD adaptive filters that use a combination between the variance of the mixing and the output signals of the SAD structures. These two proposed algorithms (Proposed 1 and Proposed 2) aim to improve the convergence speed performance and the output signal-to-noise-ratio of the original SAD algorithm when no smart control of the adaptive filters is used. The proposed algorithms have very interesting properties with non-stationary signal like speech when the SAD algorithm is, used alone, fails. The simulation results that are obtained by the comparison between the proposed algorithms (Proposed 1 and Proposed 2) and the original two-channel set-membership NLMS algorithm have shown the best performances of Proposed 1 and Proposed 2 in terms of the following criteria: systems mismatch, segmental SNR and segmental mean square error.

Keywords Speech enhancement · SAD · NLMS · Set-membership · Adaptive algorithm · Convergence speed

Abbreviations

Fs	Sampling frequency
SNR	Signal to noise ratio
LMS	Least-mean-squares
NLMS	Normalized least-mean-squares
AWGN	Additive white Gaussian noise
WGN	White Gaussian noise
USASI	United state of America standard institute
CSP	Convergence speed performance
MSE	Mean square error
BSS	Blind source separation
CD	Cepstral distance
FBSS	Forward BSS
SFT	Short-Fourier-transform

SegMSE	Segmental mean-square-error
SegSNR	Segmental signal-to-noise ratio
SM	System mismatch
PNLMS	Proportionate NLMS
SAD	Symmetric adaptive decorrelating
SMF	Set-membership filtering
AP	Affine projection
SM-AP	Set-membership affine projection
TC-SM-NLMS	Two-channel set membership NLMS

List of symbols

L	Real impulse responses length
k	Time index
dB	Decibel
$s(n)$	Original speech signal

✉ Mohamed Djendi, m_djendi@yahoo.fr; Abdelhak Cheffi, cheffi.abdelhak@yahoo.com | ¹Laboratory of Signal Processing and Imaging (LATSI), University of Blida 1, Route de Soumaa, B.P. 270, 09000 Blida, Algeria.

$b(n)$	Noise
$h_{12}(n), h_{21}(n)$	Cross-coupling impulse responses
$p_1(n), p_2(n)$	Noisy speech signals
$\delta(n)$	Dirac impulse
$w_{12}(n), w_{21}(n)$	Adaptive filters
$u_1(n)$	Estimated speech by forward structure
$u_2(n)$	Estimated noise by forward structure
$H_1(k)$ and $H_2(k)$	Conditions set
μ_1 and μ_2	Step-sizes
N	Adaptive filter length
γ_1 and γ_2	Fixed threshold error
$SNR1, SNR2$	Signal to noise ration in the two inputs
$w_{12}^{opt}(k), w_{21}^{opt}(k)$	Optimal filters
$\mu_{c1}(k), \mu_{c2}(k)$	Control step-sizes of Proposed 1
$\mu_{v1}(k), \mu_{v2}(k)$	Control step-sizes of Proposed 2
β_1, β_2	Smoothing parameters of Proposed 2

1 Introduction

Adaptive filtering principle has been largely used in many area of research in the past decades, and several algorithms based on different types of criterion have been proposed in the literature [1, 2]. Several conventional adaptive techniques assume that the problem to be resolved is based on linear form [3, 4], and this assumption, that is not true in practice, reduces the efficiency of the adaptive approach to accomplish and limits the optimal solutions that can be achieved [5, 6].

The most popular adaptive filtering algorithms like the least mean square (LMS) and the normalized LMS (NLMS) algorithms are robust and have a low computational complexity [7, 8]. The adaptive filters updates of LMS and NLMS algorithms are directly controlled by the input vector [9, 10]. This property makes them very limited in terms of performances with a non-stationary signal like speech, but, in the other hand, very convenient for dispersive-impulse response type systems [11–13]. Several techniques and algorithms based on these principles were proposed to improve the LMS and NLMS algorithms behavior, like convergence speed performance in the transient regime, under similar conditions [14]. In [15, 16], it was proposed the proportionate NLMS (PNLMS) that recovers the convergence speed property of NLMS, nevertheless the PNLMS algorithms are more complexes in comparison with NLMS. The problem's complexity of PNLMS algorithms were resolved by the set-membership filtering (SMF) family [17, 18]. In set-membership algorithm family, the output filtering control parameter is limited by a predetermined threshold [19]. Available methods based on set-membership (SM) approaches exist for linear and non linear models [20–22]. In the set-membership context, the estimated unknown system encloses the set of solutions or a union

of disconnected sub-sets of solutions which are produced by the uncertainty of experimental data and the predefined system error boundaries [23]. However, the set-membership algorithms allow to get fast convergence speed performance as well as low steady-state error. Several techniques of SMF were proposed with low complexity in the literature like the set-membership NLMS (SMNLMS), [24], and the set-membership affine projection (SM-AP) [25] algorithms.

In this paper, we propose two new adaptive algorithms based on set-membership principle that improve the original set-membership algorithm behavior in speech enhancement applications. The first proposed algorithm (called Proposed 1), is based on the combination of the SAD algorithm structure with a smart control that uses decorrelating properties between the output and the mixing signal to control the SAD filters update. The second proposed algorithm (called Proposed 2) is a modification of Proposed 1 and based on a new regularization relations of the SAD adaptive filters. The two proposed algorithms (Proposed 1 and Proposed 2) aim to improve the convergence speed performance and the output signal-to-noise-ratio of the original SAD algorithm without control system.

This paper is organized as follows: In Sect. 2, the mixing model that we have used in simulation is presented. In Sect. 3, we present the proposed algorithms, i.e. Proposed 1 and Proposed 2. In Sect. 4, the simulation part that contains experiments on three algorithm, i.e. the classical or original TC-SM-NLMS, Proposed 1, and Proposed 2. All the experiments are expressed in terms of three objective criteria as SegSNR, SegMSE, and SM. In Sect. 5, we conclude our work.

2 Two channel mixing model

In Fig. 1, we present the used mixing model in this paper. This model allows to mix two uncorrelated sources that are the speech signal $s(k)$ and the punctual noise $b(k)$ as shown in Fig. 1 [26]. Where $h_{12}(k)$ and $h_{21}(k)$ are the cross-coupling effects. The observed signals $p_1(k)$ and $p_2(k)$ of this model are:

$$p_1(k) = s(k) + b(k) * h_{21}(k) \quad (1)$$

$$p_2(k) = b(k) + s(k) * h_{12}(k) \quad (2)$$

where $*$ represents the linear convolution operator.

3 Proposed TC-SM-NLMS algorithms

In this section, we present the mathematical derivation of two proposed algorithms, i.e.

Fig. 1 The two channel mixing model

- (i) the two channel correlated set membership NLMS (TC-SM-NLMScor) algorithm that we call, in this paper, "Proposed 1".
- (ii) and the two channel correlated set membership regularized NLMS (TC-SM-RNLMScor) that we call, "Proposed 2".

In first, the mixing signals $p_1(k)$ and $p_2(k)$, obtained by Fig. 1, present the inputs of the forward BSS structures given by Fig. 2 and Fig. 3. According to these Figures, the output signals $u_1(k)$ and $u_2(k)$ are estimated as follows:

$$u_1(k) = p_1(k) - p_2(k) * w_{21}(k) \quad (3)$$

$$u_2(k) = p_2(k) - p_1(k) * w_{12}(k) \quad (4)$$

where $w_{12}(k)$ and $w_{21}(k)$ are the two adaptive filters. Inserting Eqs. (1) and (2) into (3) and (4), respectively, and using the optimality assumption for both adaptive filters, i.e. $w_{12}^{opt}(k) = h_{12}(k)$ and $w_{21}^{opt}(k) = h_{21}(k)$, we get the following output relations:

$$u_1(k) = s(k) * [\delta(k) - h_{12}(k) * w_{12}(k)] \quad (5)$$

Fig. 2 The structure of Proposed 1

Fig. 3 The structure of Proposed 2

$$u_2(k) = b(k) * [\delta(k) - h_{21}(k) * w_{21}(k)] \quad (6)$$

Weher $\delta(k)$ is the Kronecker delta. According to relations (5) and (6), we conclude that the outputs $u_1(k)$ and $u_2(k)$ converge respectively to the original signals $s(k)$ and $b(k)$ with distortions that can be corrected by the post-filters [27]. However, The observed distortions are more critical in closely spaced microphone position. In this paper we consider the choice of loosely spaced microphone where the observed distortions could be ignored [28, 31]. Both proposed TC-SM-NLMS algorithms reduce $\|w_{21}(k) - w_{21}(k-1)\|^2$ and $\|w_{12}(k) - w_{12}(k-1)\|^2$, subject to a zero *a posteriori* errors condition. Further condition is to check if the adaptive filters estimates of $w_{21}(k)$ and $w_{12}(k)$ are respectively on the outside conditions sets $H_1(k)$ and $H_2(k)$. We can write:

$$\begin{aligned} H_1(k) &= \{w_{21}(k) \in R^N : |p_1(k) - p_2(k) * w_{21}(k)| \leq \gamma_1\} \\ &= \{w_{21}(k) \in R^N : |u_1(k)| \leq \gamma_1\} \end{aligned} \quad (7)$$

$$\begin{aligned} H_2(k) &= \{w_{12}(k) \in R^N : |p_2(k) - p_1(k) * w_{12}(k)| \leq \gamma_2\} \\ &= \{w_{12}(k) \in R^N : |u_2(k)| \leq \gamma_2\} \end{aligned} \quad (8)$$

where N is the adaptive filter length, and γ_1 and γ_2 are the threshold error constants.

3.1 Presentation of Proposed 1

In Proposed 1 (see Fig. 2), we propose a new relations for the step-size updates $\mu_{c1}(k)$ and $\mu_{c2}(k)$ as given by relations (9) and (10) respectively. These new relations are given in the following:

$$\mu_{c1}(k) = \begin{cases} \mu_1 - \frac{\gamma_1}{\|u_2(k)p_1^T(k)\| + \delta} & \text{if } u_2(k) > \gamma_1 \\ 0 & \text{else} \end{cases} \quad (9)$$

$$\mu_{c2}(k) = \begin{cases} \mu_2 - \frac{\gamma_2}{\|u_1(k)\mathbf{p}_2^T(k)\| + \delta} & \text{if } u_1(k) > \gamma_2 \\ 0 & \text{else} \end{cases} \quad (10)$$

From relation (9), the parameter δ is a small constant that is introduced to avoid zero division. We can also see from (9) that the step size $\mu_{c1}(k)$ is sensible to the correlation vector of $\mathbf{p}_1(k)$ and $u_2(k)$, and from relation (10) that the step size $\mu_{c2}(k)$ is sensible to the vector correlation of $\mathbf{p}_2(k)$ and $u_1(k)$. These two relations means that when the two output signals $u_1(k)$ and $u_2(k)$ converge to their optimal solution, i.e. $w_{12}^{opt}(k) = h_{12}(k)$ and $w_{21}^{opt}(k) = h_{21}(k)$, under the condition that the step-size $\mu_{c1}(k)$ takes important values (close to 1) when speech is absent and small values (close to 0) in the other case (speech presence), in such condition, the FBSS structure combined with our proposed algorithm (Proposed 1) provides a recovered speech signal without noise at the output $u_1(k)$. The same conclusions can be generated for the step size $\mu_{c2}(k)$, but as we are only interested in the output $u_1(k)$, we will focus only on this last output.

The new relations of the adaptive filters $\mathbf{w}_{12}(k)$ and $\mathbf{w}_{21}(k)$ are given in the following:

$$\mathbf{w}_{12}(k) = \begin{cases} \mathbf{w}_{12}(k-1) + \mu_{c1}(k) \frac{u_2(k)\mathbf{p}_1(k)}{\|\mathbf{p}_1(k)\|^2 + \delta} & \text{if } |u_2(k)| > \gamma_1 \\ \mathbf{w}_{12}(k-1) & \text{else} \end{cases} \quad (11)$$

$$\mathbf{w}_{21}(k) = \begin{cases} \mathbf{w}_{21}(k-1) + \mu_{c2}(k) \frac{u_1(k)\mathbf{p}_2(k)}{\|\mathbf{p}_2(k)\|^2 + \delta} & \text{if } |u_1(k)| > \gamma_2 \\ \mathbf{w}_{21}(k-1) & \text{else} \end{cases} \quad (12)$$

The summary of the Proposed 1 algorithm is given by Table 1.

3.2 Presentation of Proposed 2

In the second proposed algorithm (Proposed 2, see Fig. 3), we propose new relations for the step-sizes $\mu_{v1}(k)$ and $\mu_{v2}(k)$, they are given as the following:

$$\mu_{v1}(k) = \begin{cases} \mu_1 - \frac{\gamma_1(\beta_1 + \beta_2)}{\beta_1 \|\mathbf{p}_1(k)\|^2 + \beta_2 \|u_2(k)\mathbf{p}_1(k)\|^2 + \delta} & \text{if } u_2(k) > \gamma_1 \\ 0 & \text{else} \end{cases} \quad (13)$$

Table 1 Proposed 1

Initialization: $s(k)$, $b(k)$, γ_1 , γ_2 , $h_{21}(k)$, $h_{12}(k)$, δ

Mixing signals:

$$p_1(k) = s(k) + b(k) * h_{21}(k)$$

$$p_2(k) = b(k) + s(k) * h_{12}(k)$$

For $k = 1, 2, \dots$

Output signals estimation

$$u_1(k) = p_1(k) - \mathbf{w}_{21}^T(k)\mathbf{p}_2(k)$$

$$u_2(k) = p_2(k) - \mathbf{w}_{12}^T(k)\mathbf{p}_1(k)$$

Variable step-size estimation

$$\mu_{c1}(k) = \begin{cases} \mu_1 - \frac{\gamma_1}{\|u_2(k)\mathbf{p}_1^T(k)\| + \delta} & \text{if } u_2(k) > \gamma_1 \\ 0 & \text{else} \end{cases}$$

$$\mu_{c2}(k) = \begin{cases} \mu_2 - \frac{\gamma_2}{\|u_1(k)\mathbf{p}_2^T(k)\| + \delta} & \text{if } u_1(k) > \gamma_2 \\ 0 & \text{else} \end{cases}$$

Relation update estimation

$$\mathbf{w}_{12}(k) = \begin{cases} \mathbf{w}_{12}(k-1) + \mu_{c1}(k) \frac{u_2(k)\mathbf{p}_1(k)}{\|\mathbf{p}_1(k)\|^2 + \delta} & \text{if } |u_2(k)| > \gamma_1 \\ \mathbf{w}_{12}(k-1) & \text{else} \end{cases}$$

$$\mathbf{w}_{21}(k) = \begin{cases} \mathbf{w}_{21}(k-1) + \mu_{c2}(k) \frac{u_1(k)\mathbf{p}_2(k)}{\|\mathbf{p}_2(k)\|^2 + \delta} & \text{if } |u_1(k)| > \gamma_2 \\ \mathbf{w}_{21}(k-1) & \text{else} \end{cases}$$

End.

$$\mu_{v2}(k) = \begin{cases} \mu_2 - \frac{\gamma_2(\beta_1 + \beta_2)}{\beta_1 \|\mathbf{p}_2(k)\|^2 + \beta_2 \|u_1(k)\mathbf{p}_2(k)\|^2 + \delta} & \text{if } |u_1(k)| > \gamma_2 \\ 0 & \text{else} \end{cases} \quad (14)$$

We can well see that the new relations of both step sizes $\mu_{v1}(k)$ and $\mu_{v2}(k)$ are controlled by two parameters β_1 and β_2 . These two constants (that are comprised between 0 and 1) allow mixing the step-size control of $\mu_{v1}(k)$ by $\|\mathbf{p}_1(k)\|^2$ and $\|u_2(k)\mathbf{p}_1(k)\|^2$. In the case of the step-size $\mu_{v2}(k)$, the control is balanced between $\|\mathbf{p}_2(k)\|^2$ and $\|u_1(k)\mathbf{p}_2(k)\|^2$. The controlled balanced parameters are chosen by $\beta_1 + \beta_2 = 1$. This means that the control of β_1 affects directly the control of β_2 , and the variances of the balanced quantities that are controlled by the couple (β_1, β_2) kept unchanged, i.e. there is no amplification of the balanced quantities in the proposed relations (13) and (14). An example of the control of β_1 is given by: $\beta_1 = 1$ if $|u_2(k)| < 10\gamma_1$ or $|u_1(k)| < 10\gamma_2$. The goal of the proposed systems of relations (13) and (14) is the control of the step sizes $\mu_{v1}(k)$ and $\mu_{v2}(k)$ to take opposite values in the range 0 and 1. This means that the proposed control allows to get a couple of values as the following:

$$\mu_{v1}(k) = \begin{cases} \approx 1 \rightarrow \text{Activity speech signal periods} \\ \approx 0 \rightarrow \text{Inactivity speech signal periods} \end{cases} \quad (15)$$

$$\mu_{v2}(k) = \begin{cases} \approx 0 \rightarrow \text{Activity speech signal periods} \\ \approx 1 \rightarrow \text{Inactivity speech signal periods} \end{cases} \quad (16)$$

Under the conditions (15) and (16), we propose to add a further technique to make automatic the adaptive process of both adaptive filters $\mathbf{w}_{12}(k)$ and $\mathbf{w}_{21}(k)$. We propose the new updates relations of both filters as the following:

$$\mathbf{w}_{12}(k) = \begin{cases} \mathbf{w}_{12}(k-1) + \mu_{v1}(k) \frac{u_2(k)\mathbf{p}_1(k)}{\|\mathbf{p}_1(k)\|^2 + \|u_2(k)\|^2 + \delta}; & \text{if } |u_2(k)| > \gamma_1 \\ \mathbf{w}_{12}(k-1) & \text{else} \end{cases} \quad (17)$$

$$\mathbf{w}_{21}(k) = \begin{cases} \mathbf{w}_{21}(k-1) + \mu_{v2}(k) \frac{u_1(k)\mathbf{p}_2(k)}{\|\mathbf{p}_2(k)\|^2 + \|u_1(k)\|^2 + \delta}; & \text{if } |u_1(k)| > \gamma_2 \\ \mathbf{w}_{21}(k-1) & \text{else} \end{cases} \quad (18)$$

The variable step sizes $\mu_{v1}(k)$ and $\mu_{v2}(k)$ are given by relations (13) and (14), respectively. As we can see, we have added to the denominator of the second part of $\mathbf{w}_{12}(k)$ and $\mathbf{w}_{21}(k)$, i.e. relations (17) and (18), the quantities $u_2(k)^2$ and $u_1(k)^2$, respectively. We remind here that the constant δ is introduced to avoid division by zeros in the case of inactivity speech signal. These new modifications give more insurance to be close to the relations (15) and (16) and lead to an automatic adaptive algorithm that is robust and fast in the same time (Table 2).

4 Simulation results of the proposed algorithms (Proposed 1 and 2)

The simulation results of the proposed algorithms are shown in this section. We compare the performance of the proposed algorithms (Proposed 1 and 2) with the classical TC-SM-NLMS algorithm. The comparison is based on the evaluation of stability and convergence speed performances. Two types of noise are used in simulations (i.e. White Gaussian noise WGN, and United States of America Standard Institute (USASI) noise that has a spectrum-like speech signal). The sampling frequency is $F_s = 8$ kHz. The parameter simulations of each algorithm are summarized in Table 3.

In Fig. 4, we give an example of two impulse responses (IRs) that are used in simulations, the IRs are generated according to the model proposed in [29]. The original speech signal and its manual voice activity detector (MVAD) and also a noisy speech signal are given in Fig. 5.

In order to evaluate the convergence speed performance of the proposed algorithms (Proposed 1 and Proposed 2), we have used 3 objective criteria [30].

(i) The system mismatch (SM): it is calculated by the following relation:

$$SM_{dB} = 20 \log_{10} \left(\frac{\mathbf{h}_{21}(k) - \mathbf{w}_{21}(k)}{\mathbf{h}_{21}(k)} \right) \quad (19)$$

where $\mathbf{h}_{21}(k)$ and $\mathbf{w}_{21}(k)$ are respectively the real impulse response and adaptive filters vectors.

(ii) The segmental mean square error (SegMSE): It is evaluated as the following:

$$SegMSE_{dB} = 20 \log_{10} \left(\sum_{k=0}^{M-1} E[u_1(k)] \right) \quad (20)$$

where M is the averaged time frame length of the output signal $u_1(k)$, and E is the expectation operator. We note that this criterion is calculated only in the voice inactivity periods (silence periods).

(iii) The segmental signal to noise ratio (SegSNR): it is given by the following relation:

$$SegSNR_{dB} = \frac{1}{P} \sum_{p=0}^{P-1} \left[10 \log_{10} \left(\frac{\sum_{k=0}^{M-1} |s(k)|^2}{\sum_{k=0}^{M-1} |s(k) - u_1(k)|^2} \right) \right] \quad (21)$$

where $s(k)$ and $u_1(k)$ are the original speech signal and the enhanced one at the output of each algorithm, respectively. The parameter M is the mean averaging value of the output SegSNR criterion. The parameter P is the number of only-speech periods.

Table 2 Proposed 2

Initialization: $s(k), b(k), \gamma_1, \gamma_2, h_{21}(k), h_{12}(k), \delta$
Mixing signals:
$p_1(k) = s(k) + b(k) * h_{21}(k)$ $p_2(k) = b(k) + s(k) * h_{12}(k)$
For $k = 1, 2, \dots$
-Output signals estimation
$u_1(k) = p_1(k) - \mathbf{w}_{21}^T(k) \mathbf{p}_2(k)$ $u_2(k) = p_2(k) - \mathbf{w}_{12}^T(k) \mathbf{p}_1(k)$
-Variable step-size estimation
$\mu_{v1}(k) = \begin{cases} \mu_1 - \frac{\gamma_1(\beta_1 + \beta_2)}{\beta_1 \ \mathbf{p}_1(k)\ ^2 + \beta_2 \ u_2(k) \mathbf{p}_1(k)\ ^2 + \delta} & \text{if } u_2(k) > \gamma_1 \\ 0 & \text{else} \end{cases}$ $\mu_{v2}(k) = \begin{cases} \mu_2 - \frac{\gamma_2(\beta_1 + \beta_2)}{\beta_1 \ \mathbf{p}_2(k)\ ^2 + \beta_2 \ u_1(k) \mathbf{p}_2(k)\ ^2 + \delta} & \text{if } u_1(k) > \gamma_2 \\ 0 & \text{else} \end{cases}$
-Relation update estimation
$\mathbf{w}_{12}(k) = \begin{cases} \mathbf{w}_{12}(k-1) + \mu_{v1}(k) \frac{u_2(k) \mathbf{p}_1(k)}{\ \mathbf{p}_1(k)\ ^2 + \ u_2(k)\ ^2 + \delta} & ; \quad \text{if } u_2(k) > \gamma_1 \\ \mathbf{w}_{12}(k-1) & \text{else} \end{cases}$ $\mathbf{w}_{21}(k) = \begin{cases} \mathbf{w}_{21}(k-1) + \mu_{v2}(k) \frac{u_1(k) \mathbf{p}_2(k)}{\ \mathbf{p}_2(k)\ ^2 + \ u_1(k)\ ^2 + \delta} & ; \quad \text{if } u_1(k) > \gamma_2 \\ \mathbf{w}_{21}(k-1) & \text{else} \end{cases}$
End.

Table 3 Simulation parameters of TC-SM-NLMS and Proposed algorithms (i.e. Proposed 1 and 2)

	TC-SM-NLMS [2]	Proposed 1 (in this paper)	Proposed 2 (in this paper)
Simulation parameters	<i>Step-size</i> $\mu_1 = \mu_2 = 0.2, 0.4 \text{ and } 0.6$ <i>Adaptive filter length</i> $N = 64, 128 \text{ and } 256$ <i>Fixed threshold error</i> $\gamma_1 = \gamma_2 = 20 \text{ and } 80$ Input SNR1 = 3 dB Input SNR2 = 3 dB	<i>Step-size</i> $\mu_1 = \mu_2 = 0.2, 0.4 \text{ and } 0.6$ <i>Adaptive filter length</i> $N = 64, 128 \text{ and } 256$ <i>Fixed threshold error</i> $\gamma_1 = \gamma_2 = 20 \text{ and } 80$ Input SNR1 = 3 dB Input SNR2 = 3 dB	<i>Step-size</i> $\mu_1 = \mu_2 = 0.2, 0.4 \text{ and } 0.6$ <i>Adaptive filter length</i> $N = 64, 128 \text{ and } 256$ <i>Fixed threshold error</i> $\gamma_1 = \gamma_2 = 20 \text{ and } 80$ Input SNR1 = 3 dB Input SNR2 = 3 dB $\beta_1 = \beta_2 = 0.5$

4.1 System mismatch (SM) criterion evaluation

We have evaluated the SM criterion for TC-SM-NLMS, Proposed 1 and Proposed 2 algorithms. We focus only on the output $u_1(k)$ and the adaptive filter $w_{21}(k)$. The SM criterion is calculated according to relation (19). All simulation parameters are summarized in Table 3. For each simulation, we have selected the adaptive filter length, i.e. $N = 64, 128$ and 256 . In Fig. 6, the original speech and the white Gaussian noise (WGN) are used as the input of the system. In Fig. 7, we use the same speech signal with USASI noise as a second source. According to Figs. 6 and 7, we observe that the three algorithms converge toward the

optimal solution with different noise types and adaptive filter lengths. We clearly show also the fast convergence of the two proposed algorithms in comparison with the classical TC-SM-NLMS. These experiments prove that the Proposed 1 is more efficient than the classical TC-SM-NLMS algorithms in terms of convergence speed to the optimal solution, and the Proposed 2 behaves more faster than the two other ones. This means that the proposed modification that incorporate the correlation parameters in the variance normalization has improved the behavior of Proposed 2.

This good performance still the same even with important adaptive filter lengths. The good behavior of Proposed

Fig. 4 A sample of impulse responses ($h_{12}(k)$ in the left, $h_{21}(k)$ in the right), generated by the model of [29]

Fig. 5 Original speech signal and its MVAD (left), and noisy signal (right)

Fig. 6 SM evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2 for speech signal and white Gaussian noise (WGN) at source. $\mu_1 = \mu_2 = 0.2$, $\gamma_1 = \gamma_2 = 20$ and adaptive filter length $N = 64, 128, 256$ (from left to right)

Fig. 7 SM evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2 for speech signal and USASI noise at source. $\mu_1 = \mu_2 = 0.4$, $\gamma_1 = \gamma_2 = 80$ and adaptive filter length $N = 64, 128, 256$ (from left to right)

2 is managed by the new formulas of the step-size that accelerates the algorithm when the input signal variance is low (in the speech absence periods) and to slow down it when the variance is high (in the presence of speech plus noise), this smart automatic mechanism lead to an ANC system with only noise at the reference signal and this is exactly why Proposed 2 has the best performance.

4.2 Segmental Mean Square Error (SegMSE) criterion evaluation

The SegMSE is evaluated for TC-SM-NLMS, Proposed 1 and Proposed 2 algorithms and the obtained results are reported in this section. The MSE evaluation is done in the silence periods of the speech signal after the denoising process, we focus only on the output $u_1(k)$. The segmental MSE criterion is computed by (20). We keep all simulation parameters given in Sect. 4.1 unchanged. In each experiment, we change the adaptive filter length, i.e. $N=64, 128$ and 256. We have done two experiments according to the source signals. First, we use the speech signal shown in Fig. 5 and WGN as the source signal and noise, respectively, in the mixing model. The obtained results for different adaptive filters length are presented in Fig. 8. In the second experiment, the noise source is USASI, and the results are reported on Fig. 9. We confirm again that the output speech signal obtained by Proposed 2 is more intelligible than the other results.

Basing on Figs. 8 and 9, we confirm that the three algorithms converge to the optimal solution. However, we see clearly the better performance of Proposed 1 and 2 in comparison with the TC-SM-NLMS algorithms in term of convergence speed. Also, we confirm that the introduced modification on Proposed 2 algorithm has

improved the MSE values of Proposed 1, which means that Proposed 2 algorithm is more efficient than the other algorithms and allows to quickly converge to the smallest MSE values in the permanent regime. This allows low distortion of the output speech signal (more Intelligible speech signal) and more noise reduction at the output result.

4.3 Segmental signal to noise ratio (SegSNR) criterion evaluation

We have evaluated the SegSNR criterion of the following algorithms, TC-SM-NLMS, Proposed 1 and Proposed 2. We have used relation (21) to evaluate the SegSNR criterion. All simulation parameters are summarized in Table 3. The obtained results of three experiments, with three adaptive filter length, i.e. $N=64, 128$ and 256 are presented in this section. In all experiments, we have used two punctual noise types, i.e. white Gaussian in the first experiment, and USASI noises in the second one, in both experiments the source signal is a speech signal which is phonetically balanced and given by Fig. 5. The obtained results of both experiments are reported on Figs. 10 and 11, respectively.

From the obtained results, we clearly show the supremacy of both proposed algorithms (Proposed 1 and 2) in the transient phase in comparison with the classical partial TC-SM-NLMS algorithm. Also, we have observed the best performance of the second proposed (Proposed 2) algorithm in all cases ($N=64, 128$ and 256 and with both types noises, i.e. white Gaussian in Fig. 10 or USASI in Fig. 11). Also, we confirm that the processing speech signal obtained by Proposed 2 is more intelligible than the other results obtained with other algorithms.

Fig. 8 MSE evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2 for two source signals: the speech signal and the white Gaussian noise (WGN). The parameters simulations are: $\mu_1 = \mu_2 = 0.2$, $\gamma_1 = \gamma_2 = 20$; the adaptive filter lengths are $N = 64, 128, 256$ (from left to right)

Fig. 9 MSE evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2 for two source signals: the speech signal and the USASI noise. The parameters simulations are: $\mu_1 = \mu_2 = 0.4$, $\gamma_1 = \gamma_2 = 80$; and adaptive filter lengths are: $N = 64, 128, 256$ (from left to right)

4.4 Performance quantification of the proposed algorithms

In this section, we have evaluated the convergence speed performance (CSP) of the proposed algorithms (i.e. Proposed 1 and Proposed 2) in comparison with the classical TC-SM-NLMS algorithm. We have quantified the CSP in term of number of iteration to achieve the steady state regime. Two experiments were done with the three algorithms, the input signal is USASI noise in the first experiment, and a speech signal in the second one. The input SNR in both experiments is 3 dB. The maximal step-size value of each algorithm is 0.6. We have

evaluated the CSP in term of the threshold values which are selected to be equal to 20, 30, and 50 (best values of the threshold and they are selected experimentally). We have evaluated the percentage of iteration performed by each algorithm before to achieve convergence toward the optimal solutions. The obtained results are given by the Figs. 12 and 13.

According to the obtained results of Figs. 12 and 13, we can easily see the best performance of Proposed 2 in comparison with the other algorithms. The same remark is noted when we use a speech signal or USASI noise as an input. We have also noted that the threshold parameters must be experimentally selected and when this items is

Fig. 10 SegSNR criterion evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2 for two source signals: the speech signal and the white Gaussian noise (WGN). The parameters simulations of each

algorithm are: $\mu_1 = \mu_2 = 0.2$, $\gamma_1 = \gamma_2 = 20$. The adaptive filter lengths are $N = 64, 128, 256$ (from left to right)

Fig. 11 SegSNR criterion evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2 for two source signals: the speech signal and the USASI noise. The parameters simulations of each algorithm

are: $\mu_1 = \mu_2 = 0.4$, $\gamma_1 = \gamma_2 = 80$. The adaptive filter lengths are $N = 64, 128, 256$ (from left to right)

Fig. 12 Percentage of performed iteration evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2. The source signal is USASI noise. The parameters simulations of each algorithm are: $\mu_1 = \mu_2 = 0.4$, $\gamma_1 = \gamma_2 = 20, 30, 40, 50$, and the adaptive filter length is $N = 256$

Fig. 13 Percentage of performed iteration evaluation of TC-SM-NLMS, Proposed 1 and Proposed 2. The source signal is speech. The parameters simulations of each algorithm are: $\mu_1 = \mu_2 = 0.4$, $\gamma_1 = \gamma_2 = 20, 30, 40, 50$, and the adaptive filter length is $N = 256$

selected around 50, the convergence speed performance of Proposed 2 get its good values in both experiments.

5 Conclusion

In this paper, two new algorithms based set-membership principle are proposed. Both of them combine the advantages of the SAD algorithm and regularization techniques based on the cross-correlation between the output signals of the SAD structure and the mixing ones. Three criteria were used to evaluate the performances of the proposed algorithms in comparison with the original TC-SM-SAD algorithm, i.e. The system mismatch (SM), the output Seg-SNR, and SegMSE. All the criteria have shown the good behavior of the proposed algorithms (Proposed 1 and Proposed 2) even when non-stationary signals like speech are used in the input. Intensives experiments were done on both algorithms and show, in terms of these criteria, the efficiency of Proposed 2 in comparison with Proposed 1 and the original TC-SM-SAD algorithms. The proposed combination between the variance of the output signals and the mixing signal and its used as a regularization factor have shown efficiency in improving the behavior of Proposed 2 algorithm. In conclusion, we can say that we get a less distortion at the output speech signal when small values of the TC-SM-SAD algorithm step-sizes are chosen, nevertheless, the convergence speed will be degraded. The new Proposed 1 and Proposed 2 algorithms allow to improve in the same time the convergence speed performance and the distortion with low step-sizes values. We have also noted that the intelligibility property of the processed speech signal has been improved in all the cases especially with the Proposed 2 algorithm. This is why we recommend the use of these two proposed algorithms in speech enhancement and acoustic noise reduction applications.

Acknowledgements This study was carried out without funding. Authors are grateful to the Blida University, Algeria, and Professor Abderreak Guessoum, Professor at Blida University, Algeria, and Director of Signal Processing and imaging Laboratory (LATSI) for providing the infrastructure to achieve this work.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that they have no conflict of interest.

References

1. Sayed AH (2003) Fundamentals of adaptive filtering. Wiley, New York
2. Rupp M (1998) A family of adaptive filter algorithms with decorrelating properties. *IEEE Trans Signal Process* 46(3):771–775
3. Haykin S (2002) Adaptive filter theory, 4th edn. Prentice-Hall, Upper Saddle River
4. Diniz PSR (2008) Data-selective adaptive filtering, adaptive filtering. Springer, Berlin
5. Van Gerven S, Van Compernelle D (1995) Signal separation by symmetric adaptive decorrelation: stability, convergence, and uniqueness. *IEEE Trans Signal Process* 43(7):1602–1612
6. Al-Kindi MJ, Dunlop J (1989) Improved adaptive noise cancellation in the presence of signal leakage on the noise reference channel. *Sig Process* 17(3):241–250
7. Djendi M (2015) New efficient adaptive fast transversal filtering FTF-type algorithms for mono and stereophonic acoustic echo cancellation. *Int J Adapt Control Signal Process* 29(3):273–301
8. Sayoud A, Djendi M, Medahi S, Guessoum A (2018) A dual fast NLMS adaptive filtering algorithm for blind speech quality enhancement. *Appl Acoust* 135:101–110
9. Loizou PC (2017) Speech enhancement: theory and practice, 2nd edn. CRC Press, Taylor & Francis Group, Boca Raton
10. Djendi M (2017) A new two-microphone Gauss–Seidel pseudo affine projection algorithm for speech quality enhancement. *Int J Adapt Control Signal Process* 31:1162–1183
11. Huang J, Zhang X, Zhang Y, Zou X, Zeng L (2014) Speech denoising via lowrank and sparse matrix decomposition. *ETRI J* 36(1):167–170
12. Deshpande A, Grant SL (2005) A new multi-algorithm approach to sparse system adaptation. In: *Proceedings of the 13th, European signal processing conference (EUSIPCO'05)*, Antalya, Turkey, September 2005
13. Martin RK, Sethares WA, Williamson RC, Johnson CR Jr (2002) Exploiting sparsity in adaptive filters. *IEEE Trans Signal Process* 50(8):1883–1894
14. Djendi M, Scalart P (2012) Double pseudo affine projection algorithm for speech enhancement and acoustic noise reduction. In: *Proceedings of IEEE. EUSIPCO, Romania, Bucharest*, vol 1, pp 2080–2084
15. Duttweiler DL (2000) Proportionate normalized least-mean squares adaptation in echo cancelers. *IEEE Trans Speech Audio Process* 8(5):508–518
16. Deng H, Doroslovacki M (2006) Proportionate adaptive algorithms for network echo cancellation. *IEEE Trans Signal Process* 54(5):1794–1803
17. Gollamudi S, Nagaraj S, Kapoor S, Huang Y-F (1998) Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size. *IEEE Signal Process Lett* 5(5):111–114
18. Werner S, Diniz PSR (2001) Set-membership affine projection algorithm. *IEEE Signal Process Lett* 8(8):231–235
19. Diniz PSR, Werner S (2003) Set-membership binormalized data-reusing LMS algorithms. *IEEE Trans Signal Process* 51(1):124–134
20. Cheffi A, Djendi M, Guessoum A (2018) New two channel set-membership partial-update NLMS algorithms for acoustic noise reduction. In: *Applied acoustic*, 2018, pp 322–332
21. Yazdanpanah H, Lima MVS, Diniz PSR (2017) On the robustness of set-membership adaptive filtering algorithms. *EURASIP J Adv Signal Process* 72:1–12
22. Jin Z, Li Y, Liu J (2018) An improved set-membership proportionate adaptive algorithm for a block-sparse system. *J Symmetry MDPI Symmetry* 10(75):1–15
23. Mao WL (2017) Robust set-membership filtering techniques on GPS sensor jamming mitigation. *IEEE Sens J* 17(6):1810–1818
24. Zhang S (2014) Zhang J (2014) Set-membership NLMS algorithm with robust error bound. *IEEE Trans Circ Syst II Express Br* 61(7):536–540

25. Werner S, Apolinario JA, Diniz PSR (2007) Set-membership proportionate affine projection algorithms. *EURASIP J Audio Speech Music Process* 2007, 34242, pp 1–10
26. Van Gervan S, Van Compernelle D (1992) Feed forward and feedback in symmetric adaptive noise canceller: stability analysis in a simplified case. In: *European signal processing conference*, Brussels, Belgium, pp 1081–1084
27. Zinser RL, Mirchandani G, Evans JB (1985) Some experimental and theoretical results using a new adaptive filter structure for noise cancellation in the presence of crosstalk. In: *IEEE international conference on acoustic, speech and signal processing*, Mars 1985, pp 1253–1256
28. Djendi M, Scalart P, Gilloire A (2007) New frequency domain post-filters for noise cancellation using two closely spaced microphones. In: *15th European signal processing conference (EUSIPCO)*, Poland, September 3–7, 2007
29. Djendi M, Scalart P, Gilloire A (2006) Noise cancellation using two closely spaced microphones: experimental study with a specific model and two adaptive algorithms. In: *IEEE international conference on acoustic, speech and signal processing*, Toulouse, France, May 2006
30. Hu Y, Loizou PC (2008) Evaluation of objective quality measures for speech enhancement. *IEEE Trans Audio Speech Lang Process ASLP* 16(1):229–238
31. Henni R, Djendi M, Djebbari M (2019) A new efficient two-channel fast transversal adaptive filtering algorithm for blind speech enhancement and acoustic noise reduction. *Comput Electr Eng* 73:349–368. <https://doi.org/10.1016/j.compeleceng.2018.12.009>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.