Multi-convex Combination Adaptive Filtering Algorithm Based on Maximum Versoria Criterion (Workshop)

Wenjing Wu, Zhonghua Liang(✉), Yimeng Bai, and Wei Li

School of Information Engineering, Chang'an University,
Xi'an 710064, People's Republic of China
lzhxjd@hotmail.com

Abstract. Aiming at the contradiction between the convergence rate and steady state mean square error of adaptive filter based on Maximum Versoria Criterion (MVC), this paper introduces the multi-convex combination strategy into MVC algorithm, and proposes a multi-convex combination MVC (MCMVC) algorithm. Simulation results show that compared with the existing MVC algorithm, MCMVC algorithm can select the best filter more flexibly under different weight change rates, and thus it has faster convergence speed and stronger tracking ability. Moreover, compared with the existing multi-convex combination maximum correntropy criterion (MCMCC) algorithm, MCMVC algorithm not only ensures the tracking performance, but also has lower exponential computation and steady-state error.

Keywords: Maximum Versoria Criterion (MVC) · Multi-convex combination · Multi-convex combination maximum correntropy criterion (MCMCC) · Steady-state error

1 Introduction

At present, adaptive filters have been widely used in signal processing and machine learning. With the development of linear adaptive filters, many linear adaptive filtering algorithms have been proposed, the most common of which are: least mean square (LMS) algorithm [1], affine projection algorithm (APA) [2] and recursive least squares (RLS) algorithm [3]. Although the above algorithms have the advantage of good performance in Gaussian noise environment, there is still a problem of performance degradation under non-Gaussian noise.

Therefore, many linear adaptive filtering algorithms against non-Gaussian noise have been proposed, such as least mean fourth (LMF) algorithm based on

This work was supported in part by the National Natural Science Foundation of China under Grant 61271262 and Grant 61871314, and in part by the Fundamental Research Funds for the Central Universities, CHD under Grant 300102249303 and Grant 300102249107.

H. Gao et al. (Eds.): ChinaCom 2019, LNICST 313, pp. 297–306, 2020.
https://doi.org/10.1007/978-3-030-41117-6_24

gradient [4], least mean p-power (LMP) algorithm [5] and recursive least p-norm (RLP) algorithm [6]. In recent years, the concepts of entropy, mutual information and correntropy in information theory have been applied in the field of adaptive filtering, among which the maximum correntropy criterion (MCC) algorithm [7] and minimum error entropy (MEE) algorithm [8] have attracted wide attention due to their strong robustness to non-Gaussian environments. The common point of MCC and MEE algorithms is that when the error is outliers, the superposition of weight updating is almost zero, which makes the algorithm resistant to non-Gaussian impulse noise.

However, since the default Gaussian kernel function in the MCC algorithm is not the optimal, Chen et al. proposed generalized maximum correntropy criterion (GMCC) by using generalized Gaussian density function as cost function in [9]. GMCC algorithm can be used in various non-Gaussian noise environments and it includes the original correntropy with a Gaussian kernel as a special case. Later, in [10], Huang et al. proposed maximum versoria criterion (MVC) algorithm in which the versoria function is used as cost function. Compared with GMCC algorithm, MVC algorithm not only avoids the high exponential function, but also has lower steady-state error and stronger resistance to non-Gaussian interference. However, MVC algorithm with invariable step-size has the contradiction between the convergence rate and steady state mean square error.

Therefore, this paper introduces the multi-convex combination strategy into MVC algorithm, and proposes a multi-convex combination adaptive filtering algorithm under non-Gaussian noise, namely MCMVC algorithm. Different from the traditional multi-convex combination strategy [11], the proposed MCMVC algorithm uses MVC to update the mixing factor indirectly, so it has strong robustness against various non-Gaussian noises.

In addition, in order to improve the convergence speed of MCMVC algorithm, the corresponding weight transfer scheme for non-Gaussian noise is presented. Simulation results show that in non-stationary system identification scenarios the proposed algorithm not only has good performance in the presence of non-Gaussian noise, but also has better tracking performance and convergence performance.

2 Adaptive Filtering Algorithm Based on Maximum Versoria Criterion

Considering the system identification model, the unknown system is modeled as a linear finite length unit impulse response filter. Therefore, the ideal output signal for an unknown system is

$$d(i) = \mathbf{u}(i)^T W_0 + v(i), \tag{1}$$

where $v(i)$ is noise interference, i represents the number of iterations, superscript T represents the vector transpose operation, $\mathbf{u}(i) \in R^m$ represents the input vector of the unknown system, usually defined as: $\mathbf{u}(i) \in [u(i), u(i-1), \ldots, u(i+m-1)]^T$, m represents the filter length, and W_0 represents the unknown weight vector to be estimated.

The system output error is defined as

$$e(i) = d(i) - \mathbf{u}(i)^T W(i-1), \tag{2}$$

where $W(i-1)$ is the weight vector for the $i-1$-th iteration.

The generalized versoria function as a cost function can be expressed as [10]

$$J(W(i-1)) = E[\frac{1}{1+\tau|e(i)|^p}], \tag{3}$$

where E represents the expectation operation, $p > 0$ represents the shape parameter of the generalized versoria function, $\tau = (2\alpha)^{-p}$, and $\alpha > 0$ represents the radius of the circle generated by versoria function. From (3), the gradient of the cost function of the generalized versoria function can be obtained as

$$\nabla J(W(i-1)) = \tau p \frac{1}{(1+\tau|e(i)|^p)^2} |e(i)|^{p-1} sign(e(i))\mathbf{u}(i). \tag{4}$$

According to the random positive gradient principle of the adaptive algorithm, the weight coefficient update formula based on maximum versoria criterion is [10]

$$W(i) = W(i-1) + \eta_1 \frac{1}{(1+\tau|e(i)|^p)^2} |e(i)|^{p-1} sign(e(i))\mathbf{u}(i), \tag{5}$$

where $\eta_1 = \tau p$ is the step-size. When $p = 2$, the weight update formula of the standard MVC algorithm can be obtained

$$W(i) = W(i-1) + \eta_1 \frac{1}{(1+\tau e^2(i))^2} e(i)\mathbf{u}(i). \tag{6}$$

3 Multi-convex Combination Maximum Versoria Criterion Algorithm

According to Eq. (6), the overall weight vector and output of MCMVC algorithm can be expressed as

$$W_{eq}(i) = \sum\nolimits_{k=1}^{L} v_k(i) W_k(i), \tag{7}$$

$$y_{eq}(i) = \sum\nolimits_{k=1}^{L} v_k(i) y_k(i), \tag{8}$$

where $v_k(i)$ represents the mixing factor and satisfies $\sum_{k=1}^{L} v_k(i) = 1$, $y_k(i) = \mathbf{u}(i)^T W_k(i)$, $k = 1, 2 \ldots L$ represents the output of the partial filter, $W_k(i)$ represents the weight of the k-th partial filter, namely

$$W_k(i) = W_k(i-1) + \mu_k \frac{1}{\left(1 + \tau {e_k}^2(i)\right)^2} e_k(i)\mathbf{u}(i), \tag{9}$$

where μ_k represents the step size of the k-th partial filter.

In MCMVC algorithm, the setting of the mixing factor $v_k(i)$ uses a softmax activation function that enhances the stability of the multi-convex combination filter, namely

$$v_k(i) = \frac{\exp(\alpha_k(i))}{\sum_{j=1}^{L} \exp(\alpha_j(i))}, k = 1, 2 \cdots L, \tag{10}$$

where $\alpha_k(i)$ is updated by maximizing the versoria function. The updated expression of improved $\alpha_k(i)$ is as follows

$$\begin{aligned} \alpha_k(i+1) &= \alpha_k(i) + \tfrac{\mu_\alpha}{2\tau} \tfrac{\partial f(e_{eq}(i))}{\partial \alpha_k(i)} \\ &= \alpha_k(i) + \tfrac{\mu_\alpha}{2\tau} \tfrac{\partial f(e_{eq}(i))}{\partial v_k(i)} \tfrac{\partial v_k(i)}{\partial \alpha_k(i)}, \\ &= \alpha_k(i) + \mu_\alpha v_k(i)(y_k(i) - y_{eq}(i)) \tfrac{e_{eq}(i)}{(1+\tau {e_{eq}}^2(i))^2} \end{aligned} \tag{11}$$

where $f(e_{eq}(i) = \frac{1}{1+\tau {e_{eq}}^2(i)})$ is the expression of versoria function, μ_α is the update step of $\alpha_k(i)$, and $\mu_\alpha \gg \mu_1$. In (11), in order to prevent MCMVC algorithm from stopping, limit the range of $\alpha_k(i)$ to $[-\varepsilon, \varepsilon]$, where $\varepsilon = {}^1\!/_2 ln(101 - L)$.

In MCMVC algorithm, this paper proposes a weight transfer scheme suitable for non-Gaussian noise environment, which accelerates the convergence performance of the combined filter. The weight transfer scheme transfers a part of the weight coefficient of the combined filter to the filter which is worse than the combined filter. The weight of the improved k-th partial filter can be expressed as

$$W_k(i+1) = \beta W_k(i) + \mu_i \frac{1}{\left(1 + \tau {e_k}^2(i)\right)^2} e_k(i)\mathbf{u}(i) + (1-\beta) W_{eq}(i), \tag{12}$$

where β is the smoothing factor. The use condition of Eq. (12) is that the combined filter is obviously superior to the partial filter, and its judgment is based on the versoria estimator of the filter. Therefore, this paper defines the versoria estimator of partial filters and combined filters as follows

$$ver(e_k(i)) = 0.9 ver(e_k(i-1)) + 0.1 \frac{1}{1 + \tau {e_k}^2(i)}, \tag{13}$$

$$ver(e_{eq}(i)) = 0.9 ver(e_{eq}(i-1)) + 0.1 \frac{1}{1 + \tau {e_{eq}}^2(i)}. \tag{14}$$

When $ver(e_{eq}(i))/ver(e_k(i)) \geq \gamma_k$ and $\gamma_k > 1$, the weight coefficient transfer can be performed using Eq. (12). Through a lot of experiments, when γ_k and β take 2 and 0.8 respectively, the algorithm achieves the best transfer effect. Then, the implementation process of the MCMVC algorithm is shown in Algorithm 1.

Algorithm 1. Implementation Process of MCMVC Algorithm

Initialization:

Parameter: μ_α, ε, β, τ, L, γ_k, μ_k, $k = 1, 2, \cdots, L$.

Initialization: $\alpha_k(0) = 0$, $v_k(0) = 1/L$, $W_k(0) = 0$, $k = 1, 2, \cdots, L$, $ver(e_k(0)) = 0$, $ver(e_{eq}(0)) = 0$.

Computation:

while $i \geq 1$ do

(1) Compute partial filter output: $y_k(i) = {W_k}^T(i)\mathbf{u}(i)$, $k = 1, 2 \cdots L$;

(2) Compute partial filter error: $e_k(i) = d(i) - y_k(i)$, $k = 1, 2 \cdots L$;

(3) Compute combined filter output: $y_{eq}(i) = \sum\limits_{k=1}^{L} v_i(i) y_i(i)$;

(4) Compute combined filter error: $e_{eq}(i) = d(i) - y_{eq}(i)$;

(5) Compute versoria estimator of the filter:

$ver(e_k(i)) = 0.9 ver(e_k(i-1)) + 0.1 \frac{1}{1+\tau {e_k}^2(i)}$, $k = 1, 2 \cdots L$;

$ver(e_{eq}(i)) = 0.9 ver(e_{eq}(i-1)) + 0.1 \frac{1}{1+\tau {e_{eq}}^2(i)}$;

(6) Update partial filter weight vector:

if $\gamma_k \leq ver(e_{eq}(i))/ver(e_k(i))$

$W_k(i+1) = \beta W_k(i) + \mu_i \frac{1}{(1+\tau {e_k}^2(i))^2} e_k(i)\mathbf{u}(i) + (1-\beta) W_{eq}(i)$;

else

$W_k(i+1) = W_k(i) + \mu_k \frac{1}{(1+\tau {e_k}^2(i))^2} e_k(i)\mathbf{u}(i)$;

(7) Update mixing factor:

$\alpha_k(i+1) = \alpha_k(i) + \mu_\alpha v_k(i)(y_k(i) - y_{eq}(i)) \frac{e_{eq}(i)}{(1+\tau {e_{eq}}^2(i))^2} |_{-\varepsilon}^{\varepsilon}$;

$v_k(i+1) = \frac{\exp(\alpha_k(i+1))}{\sum_{j=1}^{L} \exp(\alpha_j(i+1))}$, $k = 1, 2 \cdots L$;

(8) Update combined filter weight vector: $W_{eq}(i+1) = \sum_{k=1}^{L} v_k(i+1) W_k(i+1)$.

End

4 Simulation

In this section, the convergence and tracking performance of MCMVC algorithm are analyzed and compared with MVC and MCMCCC algorithms [12] via simulations of non-stationary linear system identification. In the simulation, the length of the identification system and the length of the adaptive filter are both set to 32; the input signal is a Gaussian white noise sequence with zero mean and variance 1. For noise $v(i)$, this paper uses mixed noise model to simulate non-Gaussian noise with impulse noise, which can be expressed as

$$v(i) = (1 - \varsigma(i))A(i) + \varsigma(i)B(i), \tag{15}$$

where $\varsigma(i)$ is 0.04, $B(i)$ is a Gaussian noise with zero mean and variance 15, and four different noises are considered for $A(i)$ as follows: (1) Gaussian noise with zero mean and variance 0.25; (2) Uniform noise distributed in $\{-0.5, 0.5\}$; (3) Laplacian noise with zero mean and variance 1; (4) Binary distribution noise distributed in $\{-0.5, 0.5\}$ with equal-probability, that is $p(x = 0.5) = p(x = 0.5) = 0.5$.

This paper introduces different weight vector change rates through the random walk model. The random walk model can be expressed as

$$W_0(i+1) = W_0(i) + q(i), \tag{16}$$

where the initial value of ideal weight $W_0(i)$ is generated randomly within the range of $[-1,1]$, $q(i)$ represents the random zero mean vector with independent and identical distribution, its positive definite auto-correlation matrix is $Q = E[q(i)q^T(i)]$. $Tr(Q)$ is the measurement of the velocity of weight vector. In addition, $q(i)$ is considered to be an independent gaussian distribution.

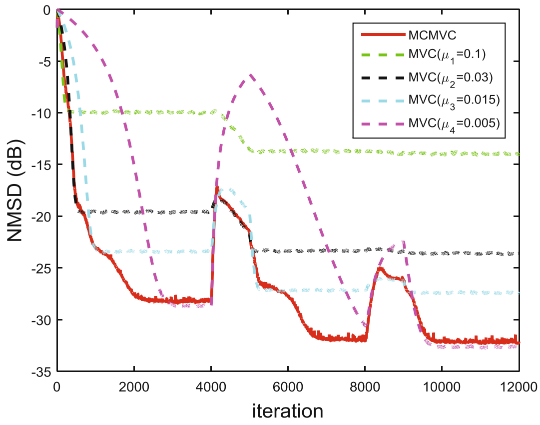

Fig. 1. Performance of different algorithms when $A(i)$ is Uniform noise

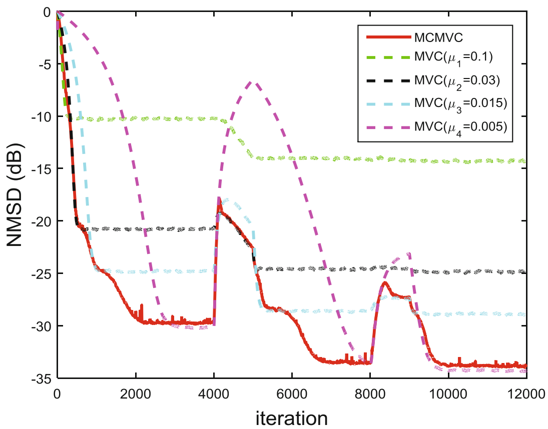

Fig. 2. Performance of different algorithms when $A(i)$ is Gaussian noise

Figures 1, 2, 3 and 4 show the tracking performance comparison between MCMVC algorithm and MVC algorithm with corresponding step size under four non-Gaussian noise environments. The number of combined filters in MCMVC algorithm is $L = 4$, the step size parameters are $\mu_1 = 0.1$, $\mu_1 = 0.03$, $\mu_1 = 0.015$ and $\mu_1 = 0.005$, the parameter μ_α is 5, and the parameter τ is 0.12. In addition, when the number of iterations are $4000 \leq i \leq 5000$ and $8000 \leq i \leq 9000$, weight vectors are added to $Tr(Q_1) = 10^{-6}$ and $Tr(Q_1) = 10^{-7}$ respectively.

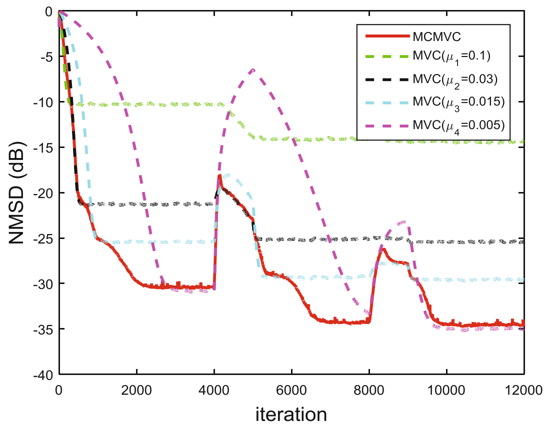

Fig. 3. Performance of different algorithms when $A(i)$ is Laplacian noise

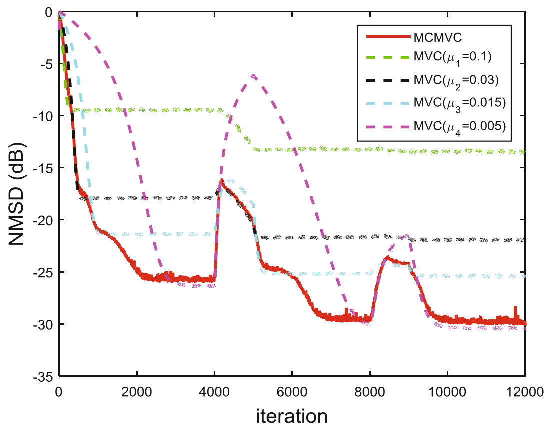

Fig. 4. Performance of different algorithms when $A(i)$ is Binary noise

According to Fig. 5, when the noise $A(i)$ is uniform noise, the change process of the four mixing factors of MCMVC algorithm can be seen. At different weight change rates, MCMVC algorithm can adaptively select partial adaptive filters with optimal performance, so that the performance of the algorithm shows the performance of the optimal filter. For example, after the weight vector $Tr(Q_1)$

is added, MCMVC algorithm first selects the adaptive filter with the optimal performance step size μ_2, then selects the adaptive filter with the optimal performance step size μ_3 to play the main role, finally, when the stable state is reached, the adaptive filter with the optimal performance step size μ_4 plays the main role.

Figure 6 shows performance comparison of weight transfer method for MCMVC algorithm when $A(i)$ is uniform noise. It can be seen from the graph that in the convergence stage and the convergence stage after the weight vector changes, the MCMVC algorithm with $\beta = 1$ has faster tracking speed than the MCMVC algorithm with $\beta = 0.8$. The weight transfer method transfers part of the weight of the combined filter to the filter with worse performance than the combined filter, so the convergence speed of the multi-convex combined filter is further improved.

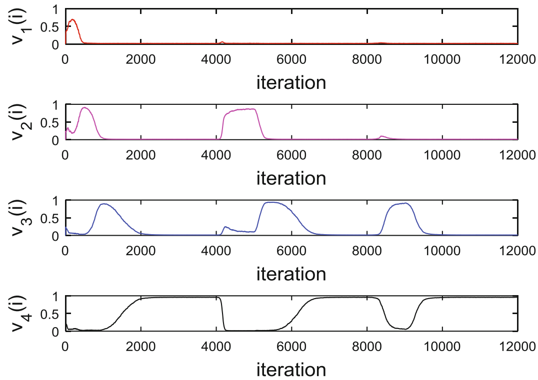

Fig. 5. Four mixing factors of MCMVC algorithm when $A(i)$ is Uniform noise

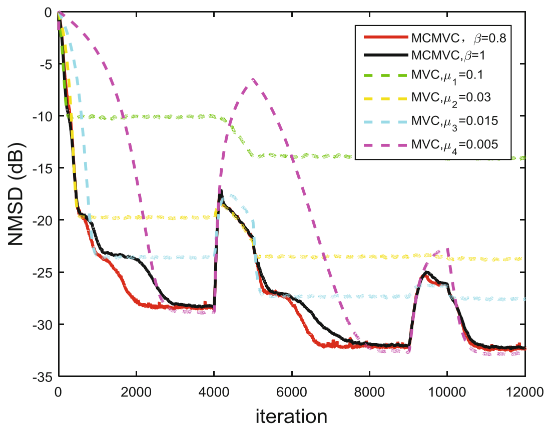

Fig. 6. Performance comparison of weight transfer method for MCMVC algorithm

Figure 7 shows the performance comparison between MCMCC algorithm and MCMVC algorithm when $A(i)$ is uniform noise. Among them, the step parameters of MCMCC algorithm are $\mu_1 = 0.1$, $\mu_2 = 0.03$, $\mu_3 = 0.015$ and $\mu_4 = 0.005$. The step parameters of MCMVC algorithm are $\mu_1 = 0.03$, $\mu_2 = 0.01$, $\mu_3 = 0.005$ and $\mu_4 = 0.001$. In all the algorithms, parameter μ_α is taken as 5. In addition, when the number of iterations is $8000 \leq k \leq 9000$, the weight vector $Tr(Q_1) = 10^{-6}$ is added to the weight change rate. As can be seen from Fig. 7, MCMVC algorithm is composed of MVC algorithms with low steady-state error and no exponential operation. Therefore, compared with MCMCC, MCMVC algorithm has lower steady-state error and exponential computation.

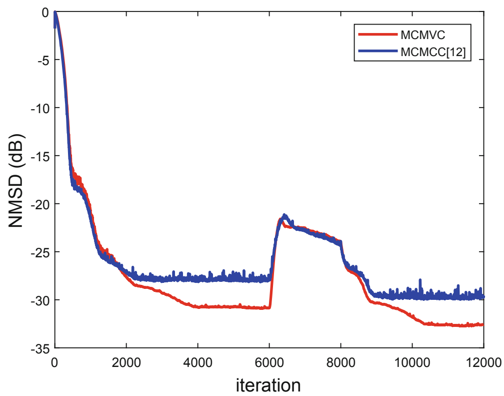

Fig. 7. Performance comparison between MCMVC and MCMCC algorithms when A(i) is Uniform noise

5 Conclusions

In this paper, the multi-convex combination strategy is combined with MVC algorithm, and MCMVC algorithm is proposed to overcome the contradiction between the convergence speed and the steady-state error of MVC algorithm. Simulation results show that compared with MVC algorithm, MCMVC algorithm can adaptively select the optimal performance filter under different weight change rates, and therefore has faster convergence speed and stronger tracking ability. At the same time, compared with MCMCC algorithm, MCMVC algorithm has lower exponential computation and steady-state error while guaranteeing tracking performance.

References

1. Sayed, A.H.: Adaptive Filters, pp. 139–209. Wiley, Hoboken (2008)
2. Ozeki, K., Umeda, T.: An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. Electron. Commun. Jpn. **67**(5), 19–27 (1984)
3. Gao, Y.: An adaptive filtering algorithm based on least square criterion. J. Guangzhou Univ. 32–34 (2001)
4. Walach, E., Widrow, B.: The least mean fourth (LMF) adaptive algorithm and its family. IEEE Trans. Inf. Theory **30**(2), 275–283 (1984)
5. Pei, S.C., Tseng, C.C.: Least mean p-power error criterion for adaptive FIR filter. IEEE J. Sel. Areas Commun. **12**(9), 1540–1547 (1994)
6. Zha, D.F., Qiu, T.S.: Adaptive generalized recursive least p-norm filtering algorithm based on minimum dispersion criterion. J. Electron. Inf. Technol. 54–58 (2007)
7. Singh, A., Principe, J.C.: Using correntropy as a cost function in linear adaptive filters. In: International Joint Conference on Neural Networks (IJCNN), pp. 2950–2955 (2009)
8. Principe, J.C.: Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives. Springer, New York (2010). https://doi.org/10.1007/978-1-4419-1570-2
9. Chen, B., Xing, L., Zhao, H., et al.: Generalized correntropy for robust adaptive filtering. IEEE Trans. Sig. Process. **64**(13), 3376–3387 (2016)
10. Huang, F., Zhang, J., Zhang, S.: Maximum versoria criterion-based robust adaptive filtering algorithm. IEEE Trans. Circuits Syst. II Express Briefs **64**(10), 1252–1256 (2017)
11. Arenas-Garcia, J., Gomez-Verdejo, V., Figueiras-Vidal, A.R.: New algorithms for improved adaptive convex combination of LMS transversal filters. IEEE Trans. Instrum. Meas. **54**(6), 2239–2249 (2005)
12. Wu, W., Liang, Z., Luo, Q., Li, W.: Tracking performance of improved convex combination adaptive filter based on maximum correntropy criterion. In: Liu, X., Cheng, D., Jinfeng, L. (eds.) ChinaCom 2018. LNICST, vol. 262, pp. 184–193. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-06161-6_18