Statistical Tracking Behavior Analysis for the Affine Projection Algorithm Based on Direction Error

Yongfeng Zhi and Wenyan Guo
Research & Development Institute in Shenzhen, Northwestern Polytechnical University, Shenzhen 518057, Guangdong, China

Correspondence should be addressed to Yongfeng Zhi; yongfeng@nwpu.edu.cn

Received 27 January 2021; Revised 19 April 2021; Accepted 26 May 2021; Published 8 June 2021

Academic Editor: Wenqing Wu

Copyright © 2021 Yongfeng Zhi and Wenyan Guo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Under the condition that the step size is less than one, a statistical tracking behavior analysis for the affine projection algorithm based on direction error is discussed. When the unknown true weight vector is modeled by the stochastic walk model, the mean weight error is derived under the four assumptions based on the deterministic recursive equation. Furthermore, the statistical tracking behavior of the steady state is analyzed for the affine projection algorithm based on direction error. Simulation analysis is shown to support the mathematical results.

1. Introduction

Since the normalized least mean square algorithm is computational simplicity, this algorithm is widely put into use by the adaptation algorithm. And the normalized least mean square algorithm is also robust to the length of the finite word effects and ease of implementation in the signal processing. However, the highly colored input signals will cause the normalized least mean square algorithm converge slowly [1]. Compared to the normalized least mean square algorithm, the affine projection (AP) algorithm is a better alternative. The AP algorithm was firstly given by Ozeki and Umeda, and it improves the convergence speed by reusing the input signal [2]. Based on the idea that the successive vectors of the input signal are orthogonal with each other, the best improvement convergence will be obtained; the normalized least mean square algorithm converge slowly [1].

There are lots of work that has been done to analyze the statistical convergence and tracking behavior of the AP and
NLMS-OCF algorithms. Based on the assumption that the input signal is the identically and independent distributed form, the statistical analysis of the class of the AP algorithms was shown, in which the mean weight error (MWE) and the mean-square error (MSE) were shown to study the convergence behavior for the NLMS-OCF algorithm [12, 13]. When we set the step size to be one, the closed-form expression is given for the convergence model of the MWE and MSE, appropriate for autoregressive-moving average (ARMA) input signal models of the AP algorithm, as in [14]. In the adaptive direction of the weight update, when we set each weight error to be zero, the optimal step size for the PAP algorithm is found as

\[ \alpha = \frac{1}{\lambda_1} \]  

The iterated direction of AP is the input vector \( x_n \), which causes the iteration error. Compared with AP algorithms, if the measurement noise is absent, the iteration error of the AP-DE algorithm is caused only by the direction vector \( \phi_n \), which is also the iterated direction of the adaptive filter. Thus, the AP-DE algorithm improves the convergence rate compared with the PAP and AP algorithms.

3. Statistical Properties of the Direction and Input Vectors

In order to study the tracking performance of the AP-DE algorithm, the four assumptions are as follows:

(A1) The successive input vector \( x_n \) is zero mean, identically and independent distributed, and then the covariance matrix can be obtained as [12, 13]

\[ R = E[xx^H] = VAV^H, \]  

where \( A = \text{diag} [\lambda_1, \lambda_2, \ldots, \lambda_N] \) are the eigenvalues of \( R \) and the corresponding eigenvectors \( V = [v_1, v_2, \ldots, v_N] \) are orthonormal to each other, i.e., \( v^H v = 1 \).

(A2) Using assumption A1 and assuming three different independent stochastic variables constitute the direction vector \( x_n \). That is,

\[ x_n = s_n r_n v_n, \]  

where \( s \) is the mean weight error (MWE) and the MSE behaviors are analyzed. When the step size was equal to one, the statistical analysis of convergence behavior for the AP-DE algorithm was given in [19]. When the step size was equal to one, the statistical tracking behavior was given for the AP-DE algorithm in [19]. When the step size was equal to one, the closed-form expression is found as

\[ \alpha = \frac{1}{\lambda_1} \]

The AP-DE algorithm updates the weight vector by the adaptive filter [6], as follows:

\[ e_n = d_n - \sum_{k=1}^{m} \tilde{a}_{nk} d_{n-k} - w_n^H \phi_n, \]  

\[ w_{n+1} = w_n + \mu \cdot \frac{e_n^*}{\sqrt{\tilde{e}_n^* \tilde{e}_n}} \phi_n, \]  

where \( H \) denotes a transposed matrix or vector and the direction vector \( \phi_n \) of the input signal is estimated based on

\[ \phi_n = x_n - X_{n-1,m} \tilde{a}_n, \]  

And from the least-squares formulation, the vector \( \tilde{a}_n = [a_{\land n,1} \ a_{\land n,2} \ \cdots \ a_{\land n,m}]^T \) is found as

\[ \tilde{a}_n = (X_{n-1,m}^H X_{n-1,m}^{-1} X_{n-1,m}^H x_n. \]  

(A3) Based on assumption A2, make the following assumption. The direction vectors \( \phi_n \) of the successive input signal are independent on the weight vector \( w_n \), and they are the Gaussian stochastic vectors with variance \( \sigma^2_\phi \) and zero mean. Therefore, the direction vectors are also three independent stochastic variables [20], i.e.,
where the vector $\varphi_n = \tilde{s}_n \tilde{\varphi}_n \tilde{v}_n$, (7)

$$P(\tilde{s}_n = \pm 1) = \frac{1}{2},$$

$$\tilde{r}_n \sim ||\varphi_n||,$$

$$P(\tilde{v}_n = v_i) = \frac{1}{N}, \quad 1 \leq i \leq N,$$

where the stochastic variable $\tilde{r}_n \sim ||\varphi_n||$ means that $\tilde{r}_n$ has the same distribution as the direction vectors of the input signal, and then, the square of the stochastic variable is denoted by $r^2 = N \sigma^2_{\varphi}$.

(A4) Based on assumption A3, make assumption about the vector $\omega_n$. The desired output signal $d_n$ is given by the mathematical model as follows:

$$d_n = w_n^oH x_n + \epsilon_n,$$ (9)

where the measurement noise $\epsilon_n$ is identically distributed, independent, and stationary with variance $\sigma^2_{\epsilon}$ and zero mean and the time-variant vectors $w_n^o$ are the system model parameters. The unknown weight vector $w_n^*$ is modeled by the stochastic walk model given by

$$w_n^* = w_{n-1}^* + \omega_{n-1},$$ (10)

where the vector $\omega_{n-1}$ is an independent, stationary zero mean vector process with the variance $\sigma^2_{\omega}$ and it is also the three independent stochastic variables that are identically and independent distributed. That is,

$$\omega_n = \tilde{s}_n \tilde{\omega}_n \tilde{v}_n,$$ (11)

$$P(\tilde{s}_n = \pm 1) = \frac{1}{2},$$

$$\tilde{r}_n \sim ||\omega_n||,$$

$$P(\tilde{v}_n = v_i) = \frac{1}{N}, \quad 1 \leq i \leq N,$$

where the independent stochastic variable $\tilde{r}_n \sim ||\omega_n||$ means that it has the same distribution as the vector $\omega_n$ and the square of the independent stochastic variable is denoted by $r^2 = N \sigma^2_{\omega}$.

Therefore, based on (3a) and (9), the iteration error $e_n$ of the adaptive filtering can be obtained as

$$e_n = w_n^oH x_n - \sum_{k=1}^{m} \tilde{a}_{nk} w_n^o H x_{n-k} - w_n^o \varphi_n + \epsilon_n - \sum_{k=1}^{m} \tilde{a}_{nk} \epsilon_{n-k},$$ (13)

From (10) and (13), we have

$$e_n = w_n^oH x_n - \sum_{k=1}^{m} \tilde{a}_{nk} w_n^o H x_{n-k} - w_n^o \varphi_n + \epsilon_n - \sum_{k=1}^{m} \tilde{a}_{nk} \epsilon_{n-k},$$

$$- w_n^o \varphi_n + \epsilon_n - \sum_{k=1}^{m} \tilde{a}_{nk} \epsilon_{n-k},$$ (14)

From (3c) and (14), it yields

$$e_n = w_n^o \varphi_n + \tilde{a}_{nk} (\sum k \omega_{n-j}^H) x_{n-k} + \epsilon_n - \sum_{k=1}^{m} \tilde{a}_{nk} \epsilon_{n-k},$$ (15)

where

$$\tilde{w}_n = w_n^o - w_n.$$ (16)

### 4. Behavior of Mean Weight Error

In order to study the tracking performance of (3b), according to (15), we have

$$w_{n+1} = w_n + \mu \frac{\varphi_n \varphi_n^H}{\varphi_n^H \varphi_n} \tilde{w}_n + \mu \frac{\varphi_n}{\varphi_n^H \varphi_n} \sum_{k=1}^{m} \tilde{a}_{nk} x_{n-k}^H (\sum_{j=1}^{k} \omega_{n-j})$$

$$+ \mu \frac{\varphi_n}{\varphi_n^H \varphi_n} (\epsilon_n - \sum_{k=1}^{m} \tilde{a}_{nk} \epsilon_{n-k}) + \omega_n.$$ (17)

Combining (10), (16), and (17), the adaptation equation is shown as

$$\tilde{w}_{n+1} = \left( I - \mu \frac{\varphi_n \varphi_n^H}{\varphi_n^H \varphi_n} \right) \tilde{w}_n - \mu \frac{\varphi_n}{\varphi_n^H \varphi_n} \sum_{k=1}^{m} \tilde{a}_{nk} x_{n-k}^H (\sum_{j=1}^{k} \omega_{n-j})$$

$$+ \mu \frac{\varphi_n}{\varphi_n^H \varphi_n} (\epsilon_n - \sum_{k=1}^{m} \tilde{a}_{nk} \epsilon_{n-k}) + \omega_n.$$ (18)

Since the parameter $\epsilon_n$, and the stochastic vector $\omega_n$ are both white noise with zero mean, we can take expectation on both sides of (18), and the last three terms in (18) becomes zero, so they can be obtained as

$$E[\tilde{w}_{n+1}] = E \left[ \left( I - \mu \frac{\varphi_n \varphi_n^H}{\varphi_n^H \varphi_n} \right) \tilde{w}_n \right].$$ (19)

Based on the vectors $\{v_1, v_2, \cdots, v_N\}$ which are orthonormal to each other, the representation of $E[\tilde{w}_n]$ can be defined as the vector $\rho_n$. That is,

$$\rho_n = \mathbf{V}^H E[\tilde{w}_n].$$ (20)
Therefore,
\[
\rho_{n,i} = \mathbf{u}_i^H E[\hat{w}_n] = E[\mathbf{u}_i^H \hat{w}_n].
\] (21)

Using this result, according to (7), premultiplication on both sides of (19) by \(\mathbf{u}_j^H\), we have
\[
\rho_{n+1,j} = E[\mathbf{u}_j^H (1 - \hat{\mathbf{w}}_{n}^H \mathbf{w}_{n}) \hat{w}_n].
\] (22)

For the orthonormality of the vector \(\mathbf{u}_j\). That is,
\[
\mathbf{u}_j^H \mathbf{u}_j^H = \begin{cases} \mathbf{u}_j^H, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}
\] (23)

Based on (23), (22) becomes
\[
\rho_{n+1,i} = \left(1 - \frac{\mu}{\mathbb{N}}\right) \rho_{n,i}.
\] (24)

5. Statistical MSE Behavior

Under assumption A3 where the direction vectors \(\mathbf{q}_n\) of the input signal are independent on the vector \(\mathbf{w}_n\), according to (18), the covariance of the weight error is proposed as

\[
\text{cov} (\hat{w}_{n+1}) = E \left(1 - \hat{\mathbf{w}}_{n+1}^H \mathbf{w}_{n} \right) \text{cov} (\hat{w}_n) \left(1 - \hat{\mathbf{w}}_{n}^H \mathbf{w}_{n} \right) E[\hat{w}_n^H \hat{w}_n] + \mu^2 \text{tr} \left(\sum_{k=1}^{m} E[\hat{w}_n^H \hat{w}_n^H \hat{w}_n] E[\mathbf{x}_{n-k}^H \mathbf{x}_{n-k}] E\left[\sum_{j=1}^{k} \mathbf{w}_{n-j} \omega_{n-j}^H\right]\right) E \left[\frac{\mathbf{q}_n^H \mathbf{q}_n}{(\mathbf{q}_n^H \mathbf{q}_n)^2}\right]
\]

\[
+ \mu^2 E \left[\sum_{k=1}^{m} \sum_{k'=1}^{m} \mathbf{w}_{n-k}^H \mathbf{w}_{n-k'} \mathbf{x}_{n-k}^H \mathbf{x}_{n-k'} \mathbf{q}_n^H \mathbf{q}_n + \mathbf{w}_{n-k}^H \mathbf{w}_{n-k} \mathbf{x}_{n-k}^H \mathbf{x}_{n-k} \mathbf{q}_n^H \mathbf{q}_n\right] + E[\mathbf{w}_n^H \mathbf{w}_n].
\] (26)
Since $\varepsilon_n$ is identically distributed, independent with variance $\sigma^2_\varepsilon$ and zero mean, under assumptions A2, A3, and A4, using (5), (7), and (11), (26) becomes as follows:

$$\begin{align*}
cov (\mathbf{w}_{n+1}) & = E \left[ (1 - \mu \mathbf{v}_n \hat{\mathbf{w}}_n^H) \cdot \mathbf{w}_{n+1}^H \right] \\
& = \mu^2 \frac{\sigma^2_\varepsilon}{T_n} \sum_{k=1}^{m} k E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] E \left[ \mathbf{v}_n \mathbf{v}_n^H \right] \\
& + \frac{\mu^2 R_n}{T_n^2} \left( 1 + \sum_{k=1}^{m} E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] \right) E \left[ \mathbf{v}_n \mathbf{v}_n^H \right] \\
& + \frac{\mu^2 R_n^2}{T_n^3} E \left[ \tilde{v}_n \tilde{v}_n^H \right].
\end{align*}$$

(27)

For the covariance matrix, we define the diagonal elements as $\hat{\lambda}_{n+1,j}$. So we have

$$\hat{\lambda}_{n+1,j} \equiv \left[ \mathbf{V}^H \cdot \text{cov} (\mathbf{w}_{n+1}) \cdot \mathbf{V} \right]_{jj} \equiv \mathbf{v}_j^H \cdot \text{cov} (\mathbf{w}_{n+1}) \cdot \mathbf{v}_j.$$

(28)

We premultiply and postmultiply on both sides of (27) by $\mathbf{v}_j^H$ and $\mathbf{v}_j$, respectively; under assumption A3, using (29), we can obtain

$$E \left[ \mathbf{v}_j^H (1 - \mu \mathbf{v}_n \hat{\mathbf{w}}_n^H) \cdot (1 - \mu \mathbf{v}_n \hat{\mathbf{w}}_n^H) \cdot \mathbf{v}_j \right] = \left( 1 + \frac{\mu^2 - 2\mu}{N} \right) \mathbf{v}_j \cdot \mathbf{A} \cdot \mathbf{v}_j,$$

(29)

where $\mathbf{A}$ is a $N \times N$ dimension matrix; we can obtain

$$\hat{\lambda}_{n+1,j} = \left( 1 + \frac{\mu^2 - 2\mu}{N} \right) \hat{\lambda}_{n,j} + \mu^2 \frac{\sigma^2_\varepsilon}{T_n} \sum_{k=1}^{m} k E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] E \left[ \mathbf{v}_j^H \mathbf{v}_j \right]$$

$$+ \frac{\mu^2 R_n}{T_n^2} \left( 1 + \sum_{k=1}^{m} E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] \right) E \left[ \mathbf{v}_j^H \mathbf{v}_j \right] + \frac{\mu^2 R_n^2}{T_n^3} E \left[ \tilde{v}_n \tilde{v}_n^H \right].$$

(30)

Based on assumptions A3 and A4, we can obtain that the vectors $\tilde{v}_n$ and $\tilde{v}_j$ both have the probability of $1/N$ to be $\mathbf{v}_j$, respectively. From (30), we can obtain

$$\hat{\lambda}_{n+1,j} = \left( 1 + \frac{\mu^2 - 2\mu}{N} \right) \hat{\lambda}_{n,j} + \mu^2 \frac{\sigma^2_\varepsilon}{N T_n^2} \sum_{k=1}^{m} k E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right]$$

$$+ \frac{\mu^2 R_n}{N T_n^2} \left( 1 + \sum_{k=1}^{m} E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] \right) + \frac{\mu^2 R_n^2}{N T_n^3}.$$

(31)

According to (15), we have the statistical MSE of tracking performance for the AP-DE algorithm as

$$E[\varepsilon_n^2] = \sigma^2_\varepsilon \sum_{k=1}^{N} \hat{\lambda}_{n+k} + \frac{R_n^2}{N} \sigma^2_\varepsilon + \sigma^2_\varepsilon \left( 1 + \sum_{k=1}^{m} E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] \right).$$

(34)

6. Steady-State Behavior

Assuming convergence, as $n \rightarrow \infty$, (31) can be transformed into

$$\lambda_{ss} = \frac{1}{2\mu - \mu^2} \left[ \mu^2 \frac{\sigma^2_\varepsilon}{T_n^2} \sum_{k=1}^{m} k E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] + \frac{\mu^2 R_n^2}{T_n^3} \left( 1 + \sum_{k=1}^{m} E \left[ \tilde{a}_{nk} \tilde{a}_{nk}^* \right] \right) \right].$$

(35)
Combining with (34) and (35), based on $r^2 = N^2\sigma^2_\phi$, the mean square error of the steady-state tracking performance can be written as

$$E[e^* e^*_n] = \frac{1}{2\mu - \mu^2} \left\{ \mu^2 r^2_n \sigma^2_w \sum_{k=1}^{m} k E[\hat{a}_{n,k} \hat{a}^*_n] + \mu^2 \sigma^2_r \left( 1 + \sum_{k=1}^{m} E[\hat{a}_{n,k} \hat{a}^*_n] \right) + r^2_n \sigma^2_w \sum_{k=1}^{m} E[\hat{a}_{n,k} \hat{a}^*_n] + \sigma^2_r \left( 1 + \sum_{k=1}^{m} E[\hat{a}_{n,k} \hat{a}^*_n] \right) \right\}.$$

(36)

7. Simulation

We give the statistical MSE of the tracking performance learning curves from the simulations in this section, and the derived models are given by (34) and (36). The initial system true weight $w_0$ is stochastic produced by 32 taps. The length of the adaptive filtering is given with the same as the derived model. The weight of the initial estimated values is given as $w_0 = 0$. The 100 independent simulations are averaged for each experiment, and each step size is equal to 0.1. The variance of the measurement noise and the vector $\omega_0$ are set to be $\sigma^2_\epsilon = 0.01$ and $2.5 \times 10^{-5}$, respectively.

Case 1. Consider that the input signal model defined by $x_n = -0.95x_{n-1} + z_n$ is an AR (1) model and the parameter $z_n$ is the white Gaussian noise with the zero mean. The parameter $m$ is set to be one. The statistical MSE for tracking behavior predicted by the model given by (34) and (36) are shown in Figure 1, together with the simulation results. We observe that the derived models are almost the same simulation results in this case.

Case 2. When the parameter $z_n$ is assumed to be the white Gaussian noise with the zero mean, the input signal is an ARMA (2, 1) model defined by $x_n = (0.7 - 0.3i)x_{n-1} - (0.4 - 0.6i)x_{n-2} + z_n + 0.5z_{n-1}$. The parameter is set to be $m = 3$. The mean square error of the tracking behavior predicted by the derived model given with (34) and (36) is shown in Figure 2. We find that the derived model cooperates with the simulation results well for the input ARMA (2, 1) model.

8. Conclusion

Under the condition that the unknown true weight vector is given by the stochastic walk model, a statistical tracking model for the AP-DE algorithm is analyzed. We give four assumptions that show the properties for both the input
and direction vectors. A statistical tracking model of the MWE and the MSE is derived based on these four assumptions. A prediction tracking model of the statistical steady-state MSE of the AP-DE algorithm is also proposed. According to the derived models in this paper, the simulation results for the input given by AR (1) and ARMA (2, 1) show the affine projection algorithm based on direction error, the better statistical tracking behavior.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Science, Technology and Innovation Commission of Shenzhen Municipality (Grant nos. U20B2040 and 61671379).

References

[1] S. Haykin, Adaptive Filter Theory, Prentice-Hall, Englewood Cliffs, NJ, 1991.
[2] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," Electronics and Communications in Japan (Part I: Communications), vol. 67, no. 5, pp. 19–27, 1984.
[3] S. G. Sankaran and A. A. L. Beex, "Normalized LMS algorithm with orthogonal correction factors," in Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers (Cat. No.97CB36136), Pacific Grove, CA, USA, 1997.
[4] M. Rupp, "A family of adaptive filter algorithms with decorrelating properties," IEEE Transactions on Signal Processing, vol. 46, no. 3, pp. 771–775, 1998.
[5] F. Bouteille, P. Scalart, and M. Corazza, "Pseudo affine projection algorithm new solution for adaptive identification," Proceedings Eurospeech, vol. 1, pp. 427–430, 1999.
[6] Y. F. Zhi, R. Li, and H. X. Li, "A new affine projection algorithm and its statistical behavior," Chinese Journal of Electronics, vol. 22, no. 3, pp. 537–542, 2013.
[7] S. G. Sankaran and A. A. L. Beex, "Fast generalized affine projection algorithm," International Journal of Adaptive Control and Signal Processing, vol. 14, no. 6, pp. 623–641, 2000.
[8] L. Lu, G. Zhu, X. Yang, K. Zhou, Z. Liu, and W. Wu, "Affine projection algorithm-based high-order error power for partial discharge denoising in power cables," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 4, pp. 1821–1832, 2020.
[9] J. Guo, F. Yang, and J. Yang, "Mean-square performance of the modified filtered-x affine projection algorithm," Circuits Systems and Signal Processing, vol. 39, no. 8, pp. 4243–4257, 2020.
[10] Vasundhara, "Sparsity aware affine-projection-like filtering integrated with robust set membership and M-estimate approach for acoustic feedback cancellation in hearing aids," Applied Acoustics, vol. 175, no. 2, article 107778, 2021.
[11] P. Song, H. Zhao, X. Zeng, W. Quan, and L. Zhao, "Robust time-varying parameter proportionate affine-projection-like algorithm for sparse system identification," Circuits Systems and Signal Processing, vol. 1, pp. 1–22, 2021.
[12] S. G. Sankaran and A. A. L. Beex, "Convergence behavior of affine projection algorithms," IEEE Transactions on Signal Processing, vol. 48, no. 4, pp. 1086–1096, 2000.
[13] T. K. Paul and T. Ogunfunmi, "On the convergence behavior of the affine projection algorithm for adaptive filters," IEEE Transactions on Circuits and Systems, vol. 58, no. 8, pp. 1813–1826, 2011.
[14] Y. F. Zhi, J. L. Li, J. Zhang, and Z. Wang, "Statistical convergence behavior of affine projection algorithms," Applied Mathematics and Computation, vol. 270, no. 1, pp. 511–526, 2015.
[15] Y. F. Zhi, F. F. Shang, J. Zhang, and Z. Wang, "Optimal step-size of pseudo affine projection algorithm," Applied Mathematics and Computation, vol. 273, no. 15, pp. 82–88, 2016.
[16] S. J. M. de Almeida, J. C. M. Bermudez, N. J. Bershad, and M. H. Costa, "A statistical analysis of the affine projection algorithm for unity step size and autoregressive inputs," IEEE Transactions on Circuits and Systems, vol. 52, no. 7, pp. 1394–1405, 2005.
[17] S. I. M. de Almeida, J. C. M. Bermudez, N. J. Bershad, and M. H. Costa, "A stochastic model for the convergence behavior of the affine projection algorithm for Gaussian inputs," in 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP ’03), pp. 313–316, Hong Kong, China, 2003.
[18] S. J. M. de Almeida, J. C. M. Bermudez, and N. J. Bershad, "A stochastic model for a pseudo affine projection algorithm," IEEE Transactions on Signal Processing, vol. 57, no. 1, pp. 107–118, 2009.
[19] Y. F. Zhi, X. Zheng, and R. Li, "On the convergence behavior of affine projection algorithm with direction error," Asian Journal of Control, vol. 16, no. 2, pp. 530–538, 2014.
[20] Y. F. Zhi, Y. Y. Yang, X. Zheng, J. Zhang, and Z. Wang, "Statistical tracking behavior of affine projection algorithm for unity step size," Applied Mathematics and Computation, vol. 283, no. 20, pp. 22–28, 2016.