Performance analysis of unconstrained partitioned-block frequency-domain adaptive filters in under-modeling scenarios

Zhengqiang Luo[1,2], Ziying Yu[1], Fang Kang[3], Feiran Yang[4*] and Jun Yang[1,2]

*Correspondence:
feiran@mail.ioa.ac.cn

[1] Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China
[2] School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
[3] Department of Intelligent Science, Xi'an Jiaotong Liverpool University, Suzhou 215123, China
[4] State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China

Abstract

The unconstrained partitioned-block frequency-domain adaptive filter (PBFDAF) offers superior computational efficiency over its constrained counterpart. However, the correlation matrix governing the natural modes of the unconstrained PBFDAF is not full rank. Consequently, the mean coefficient behavior of the algorithm depends on the initialization of adaptive coefficients and the Wiener solution is non-unique. To address the above problems, a new theoretical model for the deficient-length unconstrained PBFDAF is proposed by constructing a modified filter weight vector within a system identification framework. Specifically, we analyze the transient and steady-state convergence behavior. Our analysis reveals that modified weight vector is independent of its initialization in the steady state. The deficient-length unconstrained PBFDAF converges to a unique Wiener solution, which does not match the true impulse response of the unknown plant. However, the unconstrained PBFDAF can recover more coefficients of the parameter vector of the unknown system than the constrained PBFDAF in certain cases. Also, the modified filter coefficient yields better mean square deviation (MSD) performance than previously assumed. The presented alternative performance analysis provides new insight into convergence properties of the deficient-length unconstrained PBFDAF. Simulations validate the analysis based on the proposed theoretical model.

Keywords: Adaptive filtering, PBFDAF, Under-modeling, Convergence behavior

1 Introduction

In audio signal processing, adaptive filters are frequently used to model systems with thousands of taps. To reduce the computational burden, frequency-domain adaptive filters (FDAFs), particularly partitioned-block FDAFs (PBFDAFs), are commonly employed [1–7]. By utilizing the fast Fourier transform (FFT) for the calculation of the block convolution and correlation, the FDAFs offer significant computational advantages over time-domain and sub-band adaptive filters [8–11]. However, the FDAF algorithm might not be suitable for certain real-time applications that demand a very low input–output latency. The PBFDAF algorithm can alleviate this limitation by dividing the whole impulse response into shorter and equal-length sub-blocks. The block convolution can be accomplished through a

series of shorter block convolutions [12–14]. The PBFDAF algorithm reduces the algorithm latency, but at the expense of lower convergence speed [15]. In recent years, deep neural networks (DNNs) have been integrated with the FDAF algorithms to enhance their convergence performance [16–18].

The constrained and unconstrained PBFDAFs are two prevalent versions of PBFDAFs [19–21]. The unconstrained PBFDAFs omit the gradient constraint on the time-domain weight vector, and hence, FFT computations are reduced in the weight update of each partition [19, 22]. When a specific adaptive filtering algorithm is developed, it is of importance to study the statistical convergence behaviors [1, 23–29]. Considerable effort has been dedicated to investigate the convergence properties of PBFDAFs [24–29]. More recently, a thorough analysis of the PBFDAF algorithm is conducted and some interesting properties are discovered in [27–29]. Specifically, the transient and steady-state convergence behaviors of the PBFDAFs with 50% overlap in exact-modeling scenarios are described by the theoretical models in [27] and [29]. Subsequently, following a similar approach as in [27], the statistical performance of the PBFDAFs in under-modeling cases is comprehensively investigated in [28]. However, the analysis of the unconstrained PBFDAFs presented in [28] is somewhat difficult to follow due to the singularity of the autocorrelation matrix governing the mean convergence mode. Additionally, the steady-state filter weight vector depends on its initialization, and the mean square deviation (MSD) is influenced by the initial conditions. Consequently, the MSD cannot accurately reflect the true system modeling capability of the unconstrained PBFDAFs

To resolve the aforementioned problems, this paper provides a new theoretical framework for the deficient-length unconstrained PBFDAFs in a system identification scenario. Motivated by methods in [28] and [29], the presented analysis is based on a modified filter weight vector. We firstly perform the mean analysis of the unconstrained PBFDAFs algorithm. Our findings reveal that the steady-state modified weight vector is independent of the initial values and converges to the unique Wiener solution. We also show that the unconstrained PBFDAFs can recover more coefficients of the unknown plant than the constrained version in certain cases, which is not revealed before and is a very nice property for system identification purpose. We then derive closed-form expressions for the steady-state values of MSD and extra mean square error (EMSE). The MSD calculated using the modified filter weight vector is lower than that obtained from the original one, which indicates that the unconstrained PBFDAFs exhibit a stronger system modeling capability than previously assumed. Moreover, the presented statistical analysis of the unconstrained PBFDAFs is much easier to follow compared to our previous work [28]. Finally, simulations support the theoretical results based on the proposed theoretical model.

2 Unconstrained PBFDAFs

In this paper, the system identification is considered as the adaptive filtering task. The input–output relation of the linear time-invariant system is described by

$$d(n) = \mathbf{x}(n)^T \mathbf{w}_o + v(n), \tag{1}$$

Luo *et al. EURASIP Journal on Advances in Signal Processing* (2024) 2024:82

Page 3 of 20

where $\mathbf{x}(n) = [x(n), \ldots, x(n - M + 1)]^T$ is the input signal vector of length M, $\mathbf{w}_o = [w_0, \ldots, w_{M-1}]^T$ is the parameter vector of the system, $d(n)$ is the desired response, $v(n)$ is the zero-mean additive noise and $(\cdot)^T$ denotes transpose operation. To obtain the parameter vector $\mathbf{w}_o$, one can solve the well-known Wiener–Hopf equations. However, this non-iterative solution requires substantial data and computational resources to compute unknown statistical matrices, such as the autocorrelation matrix of the reference signal. In scenarios such as acoustic echo control and acoustic feedback cancelation, it is preferable to estimate the impulse response iteratively using adaptive filtering algorithms due to their computational efficiency and effective tracking capability [1].

We now introduce the unconstrained PBFDAF algorithm, which processes data in blocks rather than individual samples. Once a block of data is collected, the PBFDAF algorithm performs the block convolution and weight updates. The adaptive filter of length $M_1 = PL$ is partitioned into P subfilters, each with a length of L. This paper focuses on the under-modeling scenario, i.e., the length of the parameter vector is longer than that of adaptive filter ($M_1 < M$). The time-domain input signal vector, desired signal vector, output vector and error signal vector for the kth block are defined as follows

$$\mathbf{x}_p(k) = [x((k - p - 1)L), \ldots, x((k - p + 1)L - 1)]^T, \tag{2}$$

$$\mathbf{d}(k) = [d(kL), \ldots, d((k + 1)L - 1)]^T, \tag{3}$$

$$\mathbf{y}(k) = [y(kL), \ldots, y((k + 1)L - 1)]^T, \tag{4}$$

$$\mathbf{e}(k) = [e(kL), \ldots, e((k + 1)L - 1)]^T. \tag{5}$$

The output of the PBFDAFs follows

$$\mathbf{y}(k) = \mathbf{G}_{01}\mathbf{F}^{-1} \sum_{p=0}^{P-1} \boldsymbol{\mathcal{X}}_p(k)\hat{\mathbf{W}}_p(k), \tag{6}$$

where $\boldsymbol{\mathcal{X}}_p(k) = \text{diag}\{\mathbf{F}\mathbf{x}_p(k)\}$ is the input matrix in the frequency domain of dimension $N = 2L$, $\text{diag}(\cdot)$ generates a diagonal matrix, $\mathbf{F}$ is the $N \times N$ FFT matrix and $\hat{\mathbf{W}}_p(k)$ is the frequency-domain weight vector. The window matrix $\mathbf{G}_{01} = [\mathbf{0}_L, \mathbf{I}_L]$ is used to discard the filter output corresponding to circular convolution, where $\mathbf{0}_L$ is a zero matrix of dimension L and $\mathbf{I}_L$ denotes the $L \times L$ identity matrix.

The time-domain error signal is

$$\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{y}(k). \tag{7}$$

Using FFT, the vector $\mathbf{e}(k)$ can be transformed to frequency domain, denoted as $\mathbf{E}(k) = \mathbf{F}[\mathbf{0}_{L \times 1}, \mathbf{e}^T(k)]^T$. Then, the unconstrained PBFDAF algorithm update is given by [1, 12]

$$\hat{\mathbf{W}}_p(k + 1) = \hat{\mathbf{W}}_p(k) + \mu \Lambda^{-1} \boldsymbol{\mathcal{X}}_p^H(k)\mathbf{E}(k), \tag{8}$$

where μ is the step size and $(\cdot)^H$ denotes Hermitian operations. The matrix $\Lambda = E\left[\boldsymbol{\mathcal{X}}_p^H(k)\boldsymbol{\mathcal{X}}_p(k)\right]$ and $\Lambda = \mathbf{I}_N$ represent the normalized and unnormalized PBFDAFs, respectively.

3 Statistical analysis

This section starts with a summary of the analysis results for the deficient-length unconstrained PBFDAFs algorithm developed in [28]. Following this, we construct a modified filter weight vector for an alternative performance analysis. We then proceed to analyze the convergence behaviors of the unconstrained PBFDAFs. It should be mentioned that the presented analysis is based on the results in [28], but it provides a more comprehensive understanding of the performance characteristics of the underlying algorithm.

3.1 Analysis of deficient-length unconstrained PBFDAF in [28]

By transforming the frequency-domain weight vector $\hat{\mathbf{W}}_p(k)$ back to time domain, we have

$$\hat{\mathbf{w}}_p(k) = \mathbf{F}^{-1}\hat{\mathbf{W}}_p(k) = [\hat{w}_{p,0}(k), \ldots, \hat{w}_{p,N-1}(k)]^T. \tag{9}$$

In [29], it is demonstrated that the output vector $\mathbf{y}(k)$ can be expressed in terms of the vectors $\hat{\mathbf{w}}_p(k)$ as

$$\hat{\mathbf{y}}(k) = \sum_{p=0}^{P-1} \mathbf{X}_p(k)\hat{\mathbf{w}}_p(k) = \mathbf{X}_a(k)\hat{\mathbf{w}}(k), \tag{10}$$

where $\hat{\mathbf{w}}(k) = [\hat{\mathbf{w}}_0^T(k), \ldots, \hat{\mathbf{w}}_{P-1}^T(k)]^T$ represents the augmented filter weight vector of length NP, $\mathbf{X}_p(k) = \mathbf{G}_{01}\mathbf{F}^{-1}\boldsymbol{\mathcal{X}}_p(k)\mathbf{F} = [\mathbf{x}_{p,0}(k), \ldots, \mathbf{x}_{p,N-1}(k)]$ is the time-domain input matrix of size $L \times N$ and $\mathbf{X}_a(k) = [\mathbf{X}_0(k), \ldots, \mathbf{X}_{P-1}(k)]$ is the augmented input matrix of size $L \times NP$

The parameter vector of the unknown plant can be divided into two parts $\mathbf{w}_{\mathrm{o}} = [\mathbf{w}_*^T, \mathbf{w}_r^T]^T$. The modeling part of the system $\mathbf{w}_* = [\mathbf{w}_0^T, \ldots, \mathbf{w}_{P-1}^T]^T$ corresponds to the first M_1 coefficients of the parameter vector $\mathbf{w}_{\mathrm{o}}$, where $\mathbf{w}_p = [w_{pL}, \ldots, w_{(p+1)L-1}]^T$ is a vector of length L. The under-modeling part $\mathbf{w}_r = [w_{PL}, \ldots, w_{M-1}]^T$ is of length $M_2 = M - M_1$. We define the weight-error vector as

$$\tilde{\mathbf{w}}(k) = \overline{\mathbf{w}} - \hat{\mathbf{w}}(k), \tag{11}$$

where $\overline{\mathbf{w}} = [\overline{\mathbf{w}}_0^T, \ldots, \overline{\mathbf{w}}_{P-1}^T]^T$ is the zero-padded parameter vector of length NP and $\overline{\mathbf{w}}_p = [\mathbf{w}_p^T, \mathbf{0}_{1\times L}]^T$. As shown in [28], the update equation of the weight-error vector is given by

$$\begin{aligned}
\tilde{\mathbf{w}}(k+1) =& \tilde{\mathbf{w}}(k) - \mu\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{X}_a(k)\tilde{\mathbf{w}}(k) \\
&- \mu\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{X}_r(k)\mathbf{w}_r(k) \\
&- \mu\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{v}(k),
\end{aligned} \tag{12}$$

where $\boldsymbol{\theta} = \mathrm{blkdiag}(\boldsymbol{\sigma}, \ldots, \boldsymbol{\sigma})$ is the $NP \times NP$ block diagonal matrix, $\boldsymbol{\sigma} = (\mathbf{F}^{-1}\Lambda\mathbf{F})^{-1}$ and $\mathbf{X}_r(k) = [\mathbf{x}_r(kL), \ldots, \mathbf{x}_r(kL + L - 1)]^T$ is the $L \times M_2$ input matrix with

$\mathbf{x}_r(n) = [x(n - PL), \ldots, x(n - M + 1)]^T$. Using (12), the expectation of the weight-error vector $\tilde{\mathbf{w}}(k)$ is given by

$$
\begin{aligned}
E[\tilde{\mathbf{w}}(k + 1)] =& (\mathbf{I} - \mu E[\mathbf{A}_0(k)]) E[\tilde{\mathbf{w}}(k)] \\
& - E[\mathbf{B}_0(k)] \mathbf{w}_r(k),
\end{aligned}
\tag{13}
$$

where $\mathbf{A}_0(k) = \boldsymbol{\theta} \mathbf{X}_a^T(k) \mathbf{X}_a(k)$, $\mathbf{B}_0(k) = \boldsymbol{\theta} \mathbf{X}_a^T(k) \mathbf{X}_r(k)$ and $E(\cdot)$ is the statistical expectation operator.

The convergence characteristics of the deficient-length unconstrained PBFDAFs are accurately described by the theoretical model proposed in [28]. However, this model has two limitations. First, the analysis of the mean and mean square performance is mathematically complicated. This is because of the correlation matrix $E[\mathbf{A}_0(k)]$ is singular and has zero eigenvalues. Second, because the correlation matrix $E[\mathbf{A}_0(k)]$ is rank-deficient, the filter coefficients are dependent on their initial values in the steady state. In [30], the behaviors of several popular adaptive filtering algorithms are investigated when the covariance matrix of the input signal is singular. A similar convergence behavior of the LMS algorithm is revealed in [30], e.g., the steady-state weights are biased and depend on the weight initialization. However, it should be mentioned that the deficient rank of the input covariance matrix is due to the underlying signal characteristic in [30], while it is caused by the overlap between successive partitions in the PBFDAF algorithm. In [29], we have demonstrated that the bias in the steady-state weights of the unconstrained PBFDAF can be eliminated, allowing PBFDAF to recover the true impulse response in exact-modeling scenarios. Hence, the MSD calculated using the weight-error vector cannot accurately reflect the system modeling capability of the unconstrained PBFDAFs.

3.2 Proposed model

To overcome the above difficulties, we now present an alternative theoretical model, which was originally applied to analyze the performance of the unconstrained PBFDAF in exact-modeling scenarios [29]. In [28], the main difficulty in performance analysis lies in the singularity of correlation matrix $E[\mathbf{A}_0(k)]$ since there are identical columns in the augmented input matrix $\mathbf{X}_a(k)$. Specifically, we have $\mathbf{x}_{p,0}(k) = \mathbf{x}_{p+1,L}(k)$ for $0 \leq p \leq P - 2$. This motivates us to remove the identical columns in the matrix $\mathbf{X}_a(k)$. Subsequently, we construct a new weight vector $\hat{\mathbf{h}}(k)$ to remain the time-domain output vector $\mathbf{y}(k)$ unchanged.

To illustrate our main idea, we partition the adaptive filter of length $M_1 = 2$ into $P = 2$ blocks, i.e., $L = 1$. The output of the filter is calculated by

$$
\mathbf{y}(k) = \begin{bmatrix} x(k) & x(k-1) & x(k-1) & x(k-2) \end{bmatrix} \begin{bmatrix} \hat{w}_{0,0}(k) \\ \hat{w}_{0,1}(k) \\ \hat{w}_{1,0}(k) \\ \hat{w}_{1,1}(k) \end{bmatrix}.
\tag{14}
$$

By removing the same columns of the data matrix and reconstructing the weight vector in (14), the output vector $\mathbf{y}(k)$ can also be expressed by

$$
\mathbf{y}(k) = \begin{bmatrix} x(k) & x(k-1) & x(k-2) \end{bmatrix} \begin{bmatrix} \hat{w}_{0,0}(k) \\ \hat{w}_{0,1}(k) + \hat{w}_{1,0}(k) \\ \hat{w}_{1,1}(k) \end{bmatrix}.
\tag{15}
$$

Since the data matrix in (14) contains the identical columns, its correlation matrix is singular. As suggested in [28], the steady-state values of $\hat{w}_{0,1}(k)$ and $\hat{w}_{1,0}(k)$ depend on their initial values. In contrast, the data matrix in (15) does not contain any identical columns and its correlation matrix is invertible. Hence, the steady-state coefficients $\hat{w}_{0,1}(k) + \hat{w}_{1,0}(k)$ do not depend on its initial values. Our subsequent study shows that this property not only simplifies the convergence behavior analysis but also allows us to focus on the influence of the under-modeling part on the steady-state performance of the algorithm. Finally, based on the proposed model, we demonstrate that the modeling part of the system can be recovered in certain cases

In general, by deleting the columns of $\mathbf{x}_{p,L}(k)$ $0 \leq p \leq P - 2$ in the augmented input matrix $\mathbf{X}_a(k)$, we obtain the modified input matrix $\mathbf{X}_b(k)$ with dimensions $L \times (NP - P + 1)$. The modified input matrix $\mathbf{X}_b(k)$ and input matrix $\mathbf{X}_a(k)$ are related through the following equations

$$\mathbf{X}_b(k) = \mathbf{X}_a(k)\boldsymbol{\beta}. \tag{16}$$

$$\mathbf{X}_a(k) = \mathbf{X}_b(k)\boldsymbol{\alpha}, \tag{17}$$

where the matrices $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are of size $(NP - P + 1) \times NP$ and $NP \times (NP - P + 1)$, respectively. We generate the matrix $\boldsymbol{\beta}$ by removing the $(2p - 1)L$th $(p = 1, \ldots, P - 1)$ columns of the identity matrix $\mathbf{I}_{NP}$. We generate the matrix $\boldsymbol{\alpha}$ by first constructing a matrix $\boldsymbol{\alpha}_0$, where the $(2p - 1)L$th $(p = 1, \ldots, P - 1)$ rows of the identity matrix $\mathbf{I}_{NP}$ are added to the $2pL$th row. Then, by removing $(2p - 1)L$th $(p = 1, \ldots, P - 1)$ rows of the matrix $\boldsymbol{\alpha}_0$, we obtain the matrix $\boldsymbol{\alpha}_0$.

Substituting (17) into (10), we may express the output vector $\mathbf{y}(k)$ as [29]

$$\mathbf{y}(k) = \mathbf{X}_b(k)\hat{\mathbf{h}}(k), \tag{18}$$

where $\hat{\mathbf{h}}(k) = \boldsymbol{\alpha}\hat{\mathbf{w}}(k) = [\hat{\mathbf{h}}_0^T(k), \ldots, \hat{\mathbf{h}}_{P-1}^T(k)]^T$ is the modified weight vector of length $NP - P + 1$,

$$\hat{\mathbf{h}}_0(k) = [\underbrace{\hat{w}_{0,0}(k), \ldots, \hat{w}_{0,L-1}(k),}_{L}$$
$$\underbrace{\hat{w}_{0,L+1}(k), \ldots, \hat{w}_{0,N-1}(k)}_{L-1}]^T,$$
$$\hat{\mathbf{h}}_p(k) = [\underbrace{\hat{w}_{p,0}(k) + \hat{w}_{p-1,L}(k), \ldots, \hat{w}_{p,L-1}(k),}_{L}$$
$$\underbrace{\hat{w}_{p,L+1}(k), \ldots, \hat{w}_{p,N-1}(k)}_{L-1}]^T, 1 \leq p \leq P - 2$$
$$\hat{\mathbf{h}}_{P-1}(k) = [\underbrace{\hat{w}_{P-1,0}(k) + \hat{w}_{P-2,L}(k), \ldots, \hat{w}_{P-1,L-1}(k),}_{L}$$
$$\underbrace{\hat{w}_{P-1,L}(k), \ldots, \hat{w}_{P-1,N-1}(k)}_{L}]^T.$$

Using the modified matrix $\mathbf{X}_b(k)$, the desired signal vector $\mathbf{d}(k)$ can be expressed as

Luo *et al. EURASIP Journal on Advances in Signal Processing* (2024) 2024:82

Page 7 of 20

$$\mathbf{d}(k) = \mathbf{X}_b(k)\mathbf{h}_o + \mathbf{X}_r(k)\mathbf{w}_r(k) + \mathbf{v}(k), \tag{19}$$

where $\mathbf{h}_o = \boldsymbol{\alpha}\overline{\mathbf{w}} = [\mathbf{h}_0^T, \dots, \mathbf{h}_{P-1}^T]^T$ is the modified augmented weight vector with $\mathbf{h}_p = [\mathbf{w}_p^T, \mathbf{0}_{1\times(L-1)}]^T$ for $0 \leq p \leq P-2$ and $\mathbf{h}_{P-1} = [\mathbf{w}_{P-1}^T, \mathbf{0}_{1\times L}]^T$. The modified weight-error vector is given by

$$\tilde{\mathbf{h}}(k) = \mathbf{h}_o - \hat{\mathbf{h}}(k). \tag{20}$$

Substituting (18) and (19) into (7), we may rewrite the error vector as

$$\mathbf{e}(k) = \mathbf{X}_b(k)\tilde{\mathbf{h}}(k) + \mathbf{X}_r(k)\mathbf{w}_r(k) + \mathbf{v}(k). \tag{21}$$

The connection between the theoretical model in [28] and the proposed model is depicted in Fig. 1. The elements $[\overline{\mathbf{w}}, \mathbf{X}_a(k), \hat{\mathbf{w}}(k)]$ and $[\mathbf{h}_o, \mathbf{X}_b(k), \hat{\mathbf{h}}(k)]$ are the primary components of the model in [28] and that of the proposed model, respectively. The primary components of the two models can be related by the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. As shown in (18) and (19), the two model are equivalent in calculating the output $\mathbf{y}(k)$ and the desired signal vector $\mathbf{d}(k)$. Hence, the mean square convergence behavior of the algorithm can be described by the proposed model.

3.3 Mean convergence behavior

Two assumptions are used for the following derivation. [A1] The noise signal $v(n)$ and the input signal $x(n)$ are independent and both have zero mean. [A2] Independence assumption: The input matrix $\mathbf{X}_b(k)$ and the weight vector $\hat{\mathbf{h}}(k)$ and are mutually independent [1].

Premultiplying the matrix $\boldsymbol{\alpha}$ on both sides of (12) and using (17), we obtain

$$\begin{aligned}
\tilde{\mathbf{h}}(k+1) =& \tilde{\mathbf{h}}(k) - \mu\mathbf{A}(k)\tilde{\mathbf{h}}(k) - \mu\mathbf{B}(k)\mathbf{w}_r \\
& - \mu\boldsymbol{\alpha}\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{v}(k),
\end{aligned} \tag{22}$$

where $\mathbf{A}(k) = \boldsymbol{\alpha}\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{X}_a(k)\boldsymbol{\beta}$ and $\mathbf{B}(k) = \boldsymbol{\alpha}\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{X}_r(k)$. Using the independence assumption [A2], we have

$$\begin{aligned}
E[\tilde{\mathbf{h}}(k+1)] =& [\mathbf{I} - \mu E(\mathbf{A}(k))]E[\tilde{\mathbf{h}}(k)] \\
& - \mu E[\mathbf{B}(k)]\mathbf{w}_r.
\end{aligned} \tag{23}$$

Since the matrix $E[\mathbf{A}(k)]$ does not contain any identical columns, it does not have zero eigenvalues and is invertible. Hence, the steady-state value of the weight-error vector is
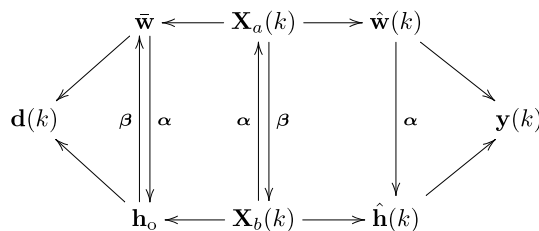

Fig. 1 Connection between the theoretical model in [28] and the proposed model

$$E[\tilde{\mathbf{h}}(\infty)] = -(E[\mathbf{A}(k)])^{-1}E[\mathbf{B}(k)]\mathbf{w}_r. \tag{24}$$

It is apparent that the steady-state weight-error vector $\tilde{\mathbf{h}}(\infty)$ is independent of the initial weight-error vector $\tilde{\mathbf{h}}(0)$, which is quite different from the original weight-error vector $\tilde{\mathbf{w}}(k)$.

We now investigate the Wiener solution of the modified weight-error vector $\tilde{\mathbf{h}}(k)$ and examine whether it converges to the Wiener solution. By replacing the modified weight-error vector $\tilde{\mathbf{h}}(k)$ with a fixed vector $\tilde{\mathbf{h}}$ in (21), we may write the MSE as

$$
\begin{aligned}
J &= \frac{1}{L}E\|\mathbf{e}(k)\|^2 \\
&= \frac{1}{L}(\tilde{\mathbf{h}}^T\mathbf{R}_b\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^T\mathbf{R}_{br}\mathbf{w}_r + \mathbf{w}_r^T\mathbf{R}_{br}^T\tilde{\mathbf{h}}) + \sigma_v^2,
\end{aligned}
\tag{25}
$$

where $\mathbf{R}_b = E[\mathbf{X}_b^T(k)\mathbf{X}_b(k)]$ and $\mathbf{R}_{br} = E[\mathbf{X}_b^T(k)\mathbf{X}_r(k)]$. The Wiener solution of (24) satisfies

$$\frac{\partial J}{\partial\tilde{\mathbf{h}}} = 2\mathbf{R}_b\tilde{\mathbf{h}} + 2\mathbf{R}_{br}\mathbf{w}_r = \mathbf{0}. \tag{26}$$

Since the matrix $\mathbf{R}_b$ is full rank, we can obtain the Wiener solution using (26)

$$\tilde{\mathbf{h}}_{\text{opt}} = -\mathbf{R}_b^{-1}\mathbf{R}_{br}\mathbf{w}_r. \tag{27}$$

Equation (27) indicates that, within our theoretical framework, the Wiener solution is uniquely determined by the under-modeling part and the statistical characteristics of the input signal. In contrast, the theoretical model proposed in [28] results in infinitely many Wiener solutions. In Appendix A, we demonstrated that $E[\tilde{\mathbf{h}}(\infty)] = \tilde{\mathbf{h}}_{\text{opt}}$, which indicates that the modified weight-error vector $\tilde{\mathbf{h}}(k)$ converges to the Wiener solution. Moreover, the correlation matrix $\mathbf{R}_{br}$ is nonzero regardless of the input signal. Consequently, the vector $E[\tilde{\mathbf{h}}(\infty)]$ is not a zero vector and $\hat{\mathbf{h}}(k)$ cannot converge to $\mathbf{h}_o$ in the mean sense.

As aforementioned, the vector $E[\hat{\mathbf{h}}(\infty)]$ does not match $\mathbf{h}_o$ that consists of the first PL elements of the unknown plant. Does it mean that we cannot obtain any useful information about the unknown plant from $\hat{\mathbf{h}}(k)$? The answer is no as we will immediately show.

3.3.1 *White noise input*

We will demonstrate that we can recover (calculate) the first $PL + L$ coefficients of the parameter vector $\mathbf{w}_o$ for white noise input. For $M_2 \geq L$, the under-modeling vector can be divided into two parts $\mathbf{w}_r = [\mathbf{w}_{r1}^T, \mathbf{w}_{r2}^T]^T$, where $\mathbf{w}_{r1} = [w_{PL}, \ldots, w_{PL+L-1}]$ and $\mathbf{w}_{r2} = [w_{PL+L}, \ldots, w_{PL+M_2-1}]$. For $M_2 < L$, we may zero-pad the vector $\mathbf{w}_r$ to obtain $\mathbf{w}_{r0} = [\mathbf{w}_r^T, \mathbf{0}_{1 \times (L-M_2)}]^T$.

For white noise input signal, the matrix $\mathbf{R}_{br}$ can be represented as

$$\mathbf{R}_{br} = \begin{bmatrix} \mathbf{0}_{(NP-P-L+1)\times L} & \mathbf{0}_{(NP-P-L+1)\times(M_2-L)} \\ \mathbf{K} & \mathbf{0}_{L\times(M_2-L)} \end{bmatrix}, \tag{28}$$

where $\mathbf{K} = \mathrm{diag}\{\sigma_x^2[L, L-1, \ldots, 1]^T\}$ is a $L \times L$ diagonal matrix and $\sigma_x^2 = E[x^2(n)]$ is the variance of the input signal. We partition the matrix $\mathbf{R}_b^{-1}$ into four submatrices

$$\mathbf{R}_b^{-1} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_3 & \mathbf{R}_4 \end{bmatrix}, \tag{29}$$

where $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$ and $\mathbf{R}_4$ are the submatrix of dimensions $(NP-P-L+1) \times (NP-P-L+1)$, $(NP-P-L+1) \times L$, $L \times (NP-P-L+1)$ and $L \times L$, respectively. Substituting (28) and (29) into (27), we obtain

$$E[\tilde{\mathbf{h}}(\infty)] = - \begin{bmatrix} \mathbf{R}_2 \mathbf{K} \mathbf{w}_{r1} \\ \mathbf{R}_4 \mathbf{K} \mathbf{w}_{r1} \end{bmatrix}. \tag{30}$$

Note that for white noise input, only the first L entries of vector $\mathbf{w}_r$ contribute to the bias of the steady-state filter weight vector $E[\tilde{\mathbf{h}}(\infty)]$. We split $E[\hat{\mathbf{h}}(\infty)]$ into

$$E[\hat{\mathbf{h}}(\infty)] = \begin{bmatrix} \hat{\mathbf{h}}_1(\infty) \\ \hat{\mathbf{h}}_2(\infty) \end{bmatrix}, \tag{31}$$

where the vectors $\hat{\mathbf{h}}_1(\infty)$ and $\hat{\mathbf{h}}_2(\infty)$ are of length $NP - P - L + 1$ and L, respectively. Using (30) and (31), we have

$$\hat{\mathbf{h}}_2(\infty) = \mathbf{R}_4 \mathbf{K} \mathbf{w}_{r1}. \tag{32}$$

Using (32), We can then obtain

$$\mathbf{w}_{r1} = \mathbf{K}^{-1} \mathbf{R}_4^{-1} \hat{\mathbf{h}}_2(\infty). \tag{33}$$

Substituting (33) into (30), we obtain the vector $\tilde{\mathbf{h}}(\infty)$ and $\mathbf{h}_\mathrm{o}$ can be computed by

$$\mathbf{h}_\mathrm{o} = E[\hat{\mathbf{h}}(\infty)] + E[\tilde{\mathbf{h}}(\infty)]. \tag{34}$$

Since the modeling part $\mathbf{w}_*$ of length PL can be recovered by (34) and the first L coefficient of the under-modeling part $\mathbf{w}_r$ can be recovered by (33), we obtain the first $PL + L$ coefficients of the weight vector $\mathbf{w}_o$ from the weight vector $E[\hat{\mathbf{h}}(\infty)]$, which is referred to as $\check{\mathbf{h}} = [\mathbf{w}_*^T, \mathbf{w}_{r1}^T]^T$.

3.3.2 Correlated input

At this point, we investigate the recovery of the true weight vector $\mathbf{w}_\mathrm{o}$ for correlated input. By permuting the entries of the vector $E[\hat{\mathbf{h}}(\infty)]$, we obtain the vector $\hat{\mathbf{b}} = [\hat{\mathbf{b}}_1^T, \hat{\mathbf{b}}_2^T]$, where

$$\hat{\mathbf{b}}_1 = E[\underbrace{\hat{w}_{0,0}(\infty), \ldots, \hat{w}_{0,L-1}(\infty)}_{L},$$

$$\underbrace{\hat{w}_{1,0}(\infty) + \hat{w}_{0,L}(\infty), \ldots, \hat{w}_{1,L-1}(\infty)}_{L}, \ldots$$

$$\underbrace{\hat{w}_{P-1,0}(\infty) + \hat{w}_{P-2,L+1}(\infty), \ldots, \hat{w}_{P-1,L-1}(\infty)}_{L}]^T,$$

$$\hat{\mathbf{b}}_2 = E[\underbrace{\hat{w}_{0,L+1}(\infty), \ldots, \hat{w}_{0,N-1}(\infty)}_{L-1},$$

$$\underbrace{\hat{w}_{1,L+1}(\infty), \ldots, \hat{w}_{1,N-1}(\infty)}_{L-1}, \ldots$$

$$\underbrace{\hat{w}_{P-1,L}(\infty), \ldots, \hat{w}_{P-1,N-1}(\infty)}_{L}]^T,$$

are of lengths PL and $PL - P + 1$, respectively. The vector $\hat{\mathbf{b}}$ and $E[\hat{\mathbf{h}}(\infty)]$ can be related as

$$\hat{\mathbf{b}} = \mathbf{P}E[\hat{\mathbf{h}}(\infty)], \tag{35}$$

where $\mathbf{P}$ is a $(NP - P + 1) \times (NP - P + 1)$ permutation matrix. Premultiplying both sides of (34) by the permutation matrix $\mathbf{P}$, we have

$$\mathbf{b}_o = \hat{\mathbf{b}} + \tilde{\mathbf{b}}, \tag{36}$$

where $\mathbf{b}_o = \mathbf{P}\mathbf{h}_o = [\mathbf{w}_*^T, \mathbf{0}_{1 \times (PL-P+1)}]^T$ and $\tilde{\mathbf{b}} = \mathbf{P}E[\tilde{\mathbf{h}}(\infty)]$. Using (27), the vector $\tilde{\mathbf{b}}$ can be represented as $\tilde{\mathbf{b}} = \mathbf{Q}\mathbf{w}_r$, where $\mathbf{Q} = -\mathbf{P}\mathbf{R}_b^{-1}\mathbf{R}_{br}$ is of dimension $(NP - P + 1) \times M_2$. We split $\mathbf{Q}$ as $\mathbf{Q} = [\mathbf{Q}_1^T, \mathbf{Q}_2^T]^T$, where the submatrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are of dimensions $PL \times M_2$ and $(PL - P + 1) \times M_2$, respectively. We then rewrite (36) as

$$\begin{bmatrix} \mathbf{w}_* \\ \mathbf{0}_{(PL-P+1) \times 1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{Q}_1 \mathbf{w}_r \\ \mathbf{Q}_2 \mathbf{w}_r \end{bmatrix}. \tag{37}$$

Provided that the matrix $\mathbf{Q}_2$ is of full column rank, $\mathbf{w}_r$ can be obtained by

$$\mathbf{w}_r = -(\mathbf{Q}_2^T \mathbf{Q}_2)^{-1} \mathbf{Q}_2^T \hat{\mathbf{b}}_2. \tag{38}$$

Substituting (38) into (37), the modeling part $\mathbf{w}_*$ can be expressed as

$$\mathbf{w}_* = \hat{\mathbf{b}}_1 + \mathbf{Q}_1 \mathbf{w}_r. \tag{39}$$

Using (38) and (39), the true impulse response $\mathbf{w}_o$ can be recovered.

To guarantee that $\mathbf{Q}_2$ is of full column rank, it should have $M_2 \leq PL - P + 1$. However, this is not a sufficient condition. For instance, for a first-order autoregressive (AR) process as input, the matrix $\mathbf{Q}_2$ has rank at most L, and hence, the true impulse response cannot be recovered in this case.

The results in Sects. 3.3.1 and 3.3.2 are of particularly interest and show we can indeed obtain useful information about the true system impulse response using the unconstrained PBFDAFs. Specifically, the unconstrained PBFDAFs can recover more coefficients of the unknown plant than the constrained version in certain cases. Also, we can use (32) to

determine whether the unconstrained PBFDAF is sufficient length or deficient length. In other words, if $\hat{\mathbf{h}}_2(\infty)$ is a zero vector (or very close to a zero vector), it indicates that $\mathbf{w}_r$ is a zero vector and we are using a sufficient-length adaptive algorithm.

3.4 Mean square convergence behavior

Using (21) and invoking the independence assumption, the EMSE produced by the algorithm is

$$
\begin{aligned}
\xi_{ex}(k) = &\frac{1}{L}\mathrm{tr}(E[\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)]\mathbf{R}_b \\
&+ E[\tilde{\mathbf{h}}(k)\mathbf{w}_r^T]\mathbf{R}_{br}^T \\
&+ E[\mathbf{w}_r\tilde{\mathbf{h}}^T(k)]\mathbf{R}_{br} \\
&+ E[\mathbf{w}_r\mathbf{w}_r^T]\mathbf{R}_{rr}),
\end{aligned}
\tag{40}
$$

where $\mathbf{R}_{rr} = E[\mathbf{X}_r^T(k)\mathbf{X}_r(k)]$. The last three terms of (40) are introduced by the under-modeling part $\mathbf{w}_r$. To continue with our analysis, it is required to calculate the covariance matrix $E[\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)]$. Using (22), we have

$$
\begin{aligned}
E[\tilde{\mathbf{h}}(k+1)\tilde{\mathbf{h}}^T(k+1)] = &E[\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)] \\
&- \mu E[\mathbf{A}(k)\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)] \\
&- \mu E[\mathbf{B}(k)\mathbf{w}_r\tilde{\mathbf{h}}^T(k)] \\
&- \mu E[\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)\mathbf{A}^T(k)] \\
&- \mu E[\tilde{\mathbf{h}}(k)\mathbf{w}_r^T\mathbf{B}^T(k)] \\
&+ \mu^2 E[\mathbf{A}(k)\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)\mathbf{A}^T(k)] \\
&+ \mu^2 E[\mathbf{B}(k)\mathbf{w}_r\tilde{\mathbf{h}}^T(k)\mathbf{A}^T(k)] \\
&+ \mu^2 E[\mathbf{A}(k)\tilde{\mathbf{h}}(k)\mathbf{w}_r^T\mathbf{B}^T(k)] \\
&+ \mu^2 E[\mathbf{B}(k)\mathbf{w}_r\mathbf{w}_r^T\mathbf{B}^T(k)] \\
&+ \mu^2 \sigma_v^2 E[\boldsymbol{\alpha}\boldsymbol{\theta}\mathbf{X}_a^T(k)\mathbf{X}_a(k)\boldsymbol{\theta}^T\boldsymbol{\alpha}^T].
\end{aligned}
\tag{41}
$$

To facilitate the derivation of the recursion of the covariance matrix, we introduce the vectorization operator $\mathrm{vec}(\cdot)$, which stacks the input matrix into one column. Applying the operator $\mathrm{vec}(\cdot)$ to the matrix product $\mathbf{MQN}$, we have [31]

$$
\mathrm{vec}(\mathbf{MQN}) = (\mathbf{N}^T \otimes \mathbf{M})\mathrm{vec}(\mathbf{Q}),
\tag{42}
$$

where $\otimes$ denotes the Kronecker product.

We define the vectorization of the covariance matrix as $\mathbf{z}(k) = \mathrm{vec}(\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k))$. Using (41) and (42), $\mathbf{z}(k)$ satisfies the difference equation

$$
\mathbf{z}(k+1) = \mathbf{H}\mathbf{z}(k) + \boldsymbol{\theta}(k),
\tag{43}
$$

where the matrices $\mathbf{H}$ and $\boldsymbol{\theta}(k)$ are

$$
\mathbf{H} = \mathbf{I} - \mu\mathbf{D} + \mu^2\mathbf{J},
\tag{44}
$$

$$\boldsymbol{\theta}(k) = \sum_{i=1}^{4} \boldsymbol{\theta}_i(k). \tag{45}$$

The matrices $\mathbf{D}$ and $\mathbf{J}$ in (44) are given by $\mathbf{D} = E[\mathbf{A}(k)]\otimes\mathbf{I} + \mathbf{I}\otimes E[\mathbf{A}(k)]$ and $\mathbf{J} = E[\mathbf{A}(k)\otimes\mathbf{A}(k)]$. The vectors $\boldsymbol{\theta}_i(k)$ in (45) are given by

$$\begin{aligned}\boldsymbol{\theta}_1(k) =& (-\mu E[\mathbf{B}(k)]\otimes\mathbf{I} + \mu^2 E[\mathbf{B}(k)\otimes\mathbf{A}(k)]) \\ & \times \mathrm{vec}(\tilde{\mathbf{h}}(k)\mathbf{w}_r^T), \end{aligned} \tag{46}$$

$$\begin{aligned}\boldsymbol{\theta}_2(k) =& (-\mu\mathbf{I}\otimes E[\mathbf{B}(k)] + \mu^2 E[\mathbf{A}(k)\otimes\mathbf{B}(k)]) \\ & \times \mathrm{vec}(\mathbf{w}_r\tilde{\mathbf{h}}^T(k)), \end{aligned} \tag{47}$$

$$\boldsymbol{\theta}_3(k) = \mu^2 E[\mathbf{B}(k)\otimes\mathbf{B}(k)]\times\mathrm{vec}(\mathbf{w}_r\mathbf{w}_r^T), \tag{48}$$

$$\boldsymbol{\theta}_4(k) = \mu^2\sigma_v^2\mathrm{vec}[E(\boldsymbol{\alpha}\boldsymbol{\theta}\bar{\mathbf{X}}_a^T(k)\bar{\mathbf{X}}_a(k)\boldsymbol{\theta}^T\boldsymbol{\alpha}^T)], \tag{49}$$

where σ_v^2 is the additive noise variance. We then obtain the MSD learning curve

$$\delta_m(k) = \mathrm{tr}(E[\tilde{\mathbf{h}}(k)\tilde{\mathbf{h}}^T(k)]) = \mathrm{tr}(\mathrm{vec}^{-1}[\mathbf{z}(k)]). \tag{50}$$

Since the control matrix $\mathbf{H}$ is invertible, the steady-state value of the vector $\mathbf{z}(k)$ is

$$\mathbf{z}(\infty) = (\mathbf{I} - \mathbf{H})^{-1}\boldsymbol{\theta}(\infty). \tag{51}$$

Using (51), the steady-state value of the EMSE is given by

$$\begin{aligned}\xi_{ex}(\infty) =& \frac{1}{L}\mathrm{tr}(\mathrm{vec}^{-1}[\mathbf{z}(\infty)]\mathbf{R}_b \\ & + E[\tilde{\mathbf{h}}(\infty)\mathbf{w}_r^T]\mathbf{R}_{br}^T \\ & + E[\mathbf{w}_r\tilde{\mathbf{h}}^T(\infty)]\mathbf{R}_{br} \\ & + E[\mathbf{w}_r\mathbf{w}_r^T]\mathbf{R}_{rr}). \end{aligned} \tag{52}$$

In [28], the MSD is evaluated using the original weight vector $\hat{\mathbf{w}}(k)$ which depends on the initialization $\hat{\mathbf{w}}(0)$. Here, we evaluate the MSD using the modified weight vector $\hat{\mathbf{h}}(k)$ that is independent of the initialization $\hat{\mathbf{h}}(0)$. Also, we observe from (51) and (52) that the steady-state MSD and EMSE are influenced by the degree of under-modeling and the system noise level.

We now show that as the step size μ converges to zero, the steady-state modified MSD does not decrease to zero. Substituting (44) and (45) into (51), we get

$$\begin{aligned}\mathbf{z}(\infty) &= (-\mu\mathbf{D} + \mu^2\mathbf{J})^{-1}\sum_{i=1}^{4}\boldsymbol{\theta}_i(k) \\ &= \mu^{-1}(-\mathbf{D} + \mu\mathbf{J})^{-1}\sum_{i=1}^{4}\boldsymbol{\theta}_i(k). \end{aligned} \tag{53}$$

In Appendix B, we prove that as the step size μ converges to zero, the vector $\mathbf{z}(\infty)$ converges to

$$
\begin{aligned}
\mathbf{z}(\infty)_{\mu \to 0} = - \ & \mathbf{D}^{-1}(E[\mathbf{B}(k) \otimes \mathbf{I}]) \times \text{vec}(\tilde{\mathbf{h}}(k)\mathbf{w}_r^T) \\
& + (\mathbf{I} \otimes E[\mathbf{B}(k)]) \times \text{vec}(\mathbf{w}_r \tilde{\mathbf{h}}^T(k)).
\end{aligned}
\tag{54}
$$

As shown in (54), the MSD will not decrease to zero due to the influence of under-modeling part $\mathbf{w}_r$. Substituting (52) into (49) and (53), we have the minimum value $\delta_{m,min}(\infty)$ and $\xi_{ex,min}(\infty)$ for the MSD and EMSE, respectively.

The aim of this paper is not to replace the model in [28] with the presented one, but the presented model can be used as useful supplement to the model in [28], as will be explained. The mean weight behavior $E[\hat{\mathbf{w}}(k)]$ of the deficient-length unconstrained PBFDAF has been described in [28], while the presented model describes mean the weight behavior $E[\hat{\mathbf{h}}(k)]$ instead. The model in [28] and the presented model are indeed equivalent in describing the MSE performance of the deficient-length unconstrained PBFDAF. However, the steady-state MSD calculated using the vector $\hat{\mathbf{h}}(k)$ is smaller than that obtained using the vector $\hat{\mathbf{w}}(k)$. This is because the MSD calculated with $\hat{\mathbf{h}}(k)$ is independent of its initial condition $\hat{\mathbf{h}}(0)$. Consequently, the unconstrained PBFDAFs exhibit better modeling capability than previously assumed. Also, in the under-modeling scenarios, the steady-state MSD and EMSE have nonzero lower bounds regardless of the step size. In contrast, for the unconstrained PBFDAFs with sufficient length, both MSD and EMSE tend to zero as the step size is decreased [27]. A comparison of the theoretical model in [28] and the proposed theoretical is summarized in Table 1.

4 Simulation results

The proposed theoretical model is verified via Monte Carlo simulations (over 1000 independent runs). We consider a system identification scenario where the system to be modeled has an impulse response of length $M = 15$. The adaptive filter, with a total length of $M_1 = 8$, is partitioned into two subfilters $P = 2$, each of length $L = 4$. The reference signal $x(n)$ is either white uniform noise or an AR(15) process generated by applying the filter $H(z) = 1/(1 + 0.3z^{-1} + 0.3z^{-7} - 0.4z^{-15})$ to white Gaussian noise. The white noise $v(n)$ is scaled to achieve a signal-to-noise ratio (SNR) of 10 dB.

We study the mean convergence behavior for correlated inputs. Figure 2 shows the transient behavior of the mean modified weight-error vector $E[\tilde{\mathbf{h}}(k)]$ with a step size of 0.03. For clarity, the learning curves of the first six entries of the weight-error

Table 1 Comparison of the theoretical model in [28] and the proposed theoretical model

|	Mean convergence behavior	Mean square convergence behavior
Theoretical model in [28]	● Steady-state solution depends on the initial coefficients	● Large steady-state MSD
Proposed model	● Steady-state solution does not depend on the initial coefficients	● Smaller steady-state MSD
	● For white noise input, $PL + L$ coefficients of the impulse response can be recovered	
	● For correlated input, a maximum of $NP - P + 1$ coefficients of the impulse response can be recovered	
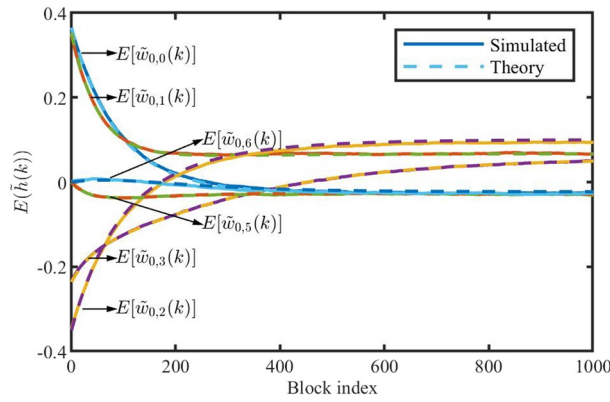
Fig. 2 Transient behavior of the modified weight-error vector $\tilde{\mathbf{h}}(k)$
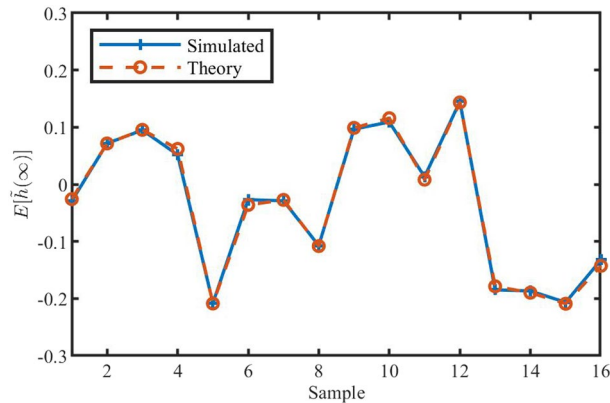

Fig. 3 Steady-state behavior of the modified weight-error vector $\tilde{\mathbf{h}}(k)$

vector $E[\tilde{\mathbf{h}}(\infty)]$ are presented. As shown in Fig. 1, the experimental results closely match the predicted learning curves calculated by (23). Figure 3 presents the steady-state modified weight-error vector $E[\tilde{\mathbf{h}}(\infty)]$. The theoretical predictions, calculated via (24), correspond well with the experimental results, and the mean weight-error vector $E[\tilde{\mathbf{h}}(\infty)]$ is nonzero. This implies the modified weight vector $\tilde{\mathbf{h}}(k)$ converges to a biased solution, i.e., the deficient-length unconstrained PBFDAFs cannot converge to $\mathbf{h}_o$ that consists of the first PL coefficients of unknown plant.

Figure 4 shows the recovery of the weight vector $\check{\mathbf{h}}$ for white input. Figure 4a illustrates the steady-state weight vector $\hat{\mathbf{b}}_1$ corresponding to the samples that lack a recovery weight vector. The vector $\hat{\mathbf{b}}_1$ is obtained from (37). Due to the influence of under-modeling part, the steady-state weight vector $\hat{\mathbf{b}}_1$ does not match the modeling part $\mathbf{w}_*$ for white noise input. Figure 4b presents the vector $\check{\mathbf{h}}$ obtained from (33) and (34). It is apparent that the first $PL + L$ coefficients of the true weight vector $\mathbf{w}_o$ can be recovered for white noise input.

We demonstrate that the unknown parameter vector can be recovered when using the AR(15) process as input. Figure 5a suggests that the weight vector $\hat{\mathbf{b}}_1$ and the steady-state weight vector $E[\hat{\mathbf{w}}_{cn}(\infty)]$ of the constrained PBFDAFs do not match the

modeling part $\mathbf{w}_*$. The results of the constrained PBFDAFs are obtained from [28]. Figure 5b shows that the vector $\mathbf{w}_o$ calculated by (38) and (39) is consistent with the impulse response. Figures 4 and 5 confirm that more coefficients of the unknown plant can be recovered using the unconstrained PBFDAFs compared to the constrained PBFDAFs, which is not revealed in the previous works. This is because the steady-state weight vector $E[\hat{\mathbf{h}}(\infty)]$ is independent on the initialization of the filter coefficient and we can extract part of steady-state weight error from the vector $E[\hat{\mathbf{h}}(\infty)]$.

We examine the mean square convergence behavior for correlated inputs. Figure 6 the learning curves of the MSD and EMSE. The theoretical learning curves of the MSD and EMSE are obtained from (50) and (51), respectively. The used step sizes are 0.03 and 0.1. As shown in Fig. 6, there is a good agreement between theoretically derived learning curve and the experimental learning curve. We also repeat this experiment using SNR = 30 dB and obtain the very similar learning curves as that in Fig. 6, which is because the MSD and EMSE are dominated by the under-model part and the system noise does not affect the steady-state MSD and EMSE significantly in this case.

Figure 7 depicts the variation of the MSD and EMSE across a range of step sizes from 0.003 to 0.62. The stability bound of the step size is obtained by the method introduced in [27]. The MSD and EMSE of the constrained version are calculated via (77) and (78) in [28], respectively. Specifically, the unconstrained PBFDAF shows a higher MSD but a lower EMSE than the constrained PBFDAF, which is consistent with the results in [28].
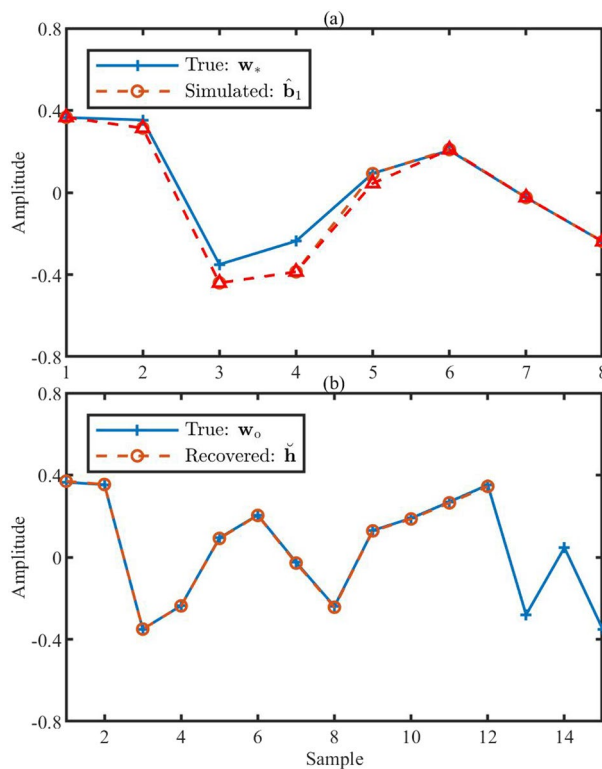

Fig. 4 Recovery of the weight vector $\check{\mathbf{h}}$ for white inputs. **a** Steady-state weight vector $\hat{\mathbf{b}}_1$. **b** Recovered weight vector $\check{\mathbf{h}}$
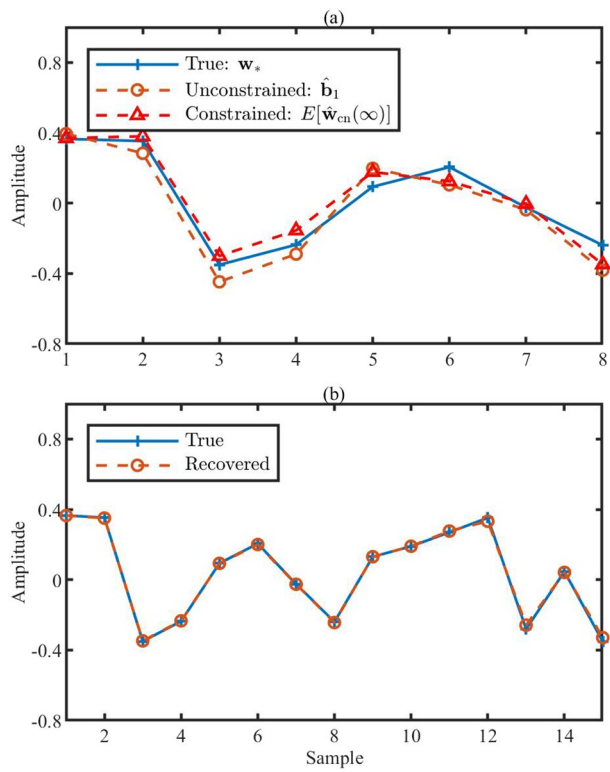
Luo *et al. EURASIP Journal on Advances in Signal Processing*   (2024) 2024:82

Page 16 of 20

Fig. 5 Recovery of the weight vector $\check{\mathbf{h}}$ for correlated inputs. **a** Steady-state weight vector $\hat{\mathbf{b}}_1$. **b** Recovered weight vector $\mathbf{w}_o$
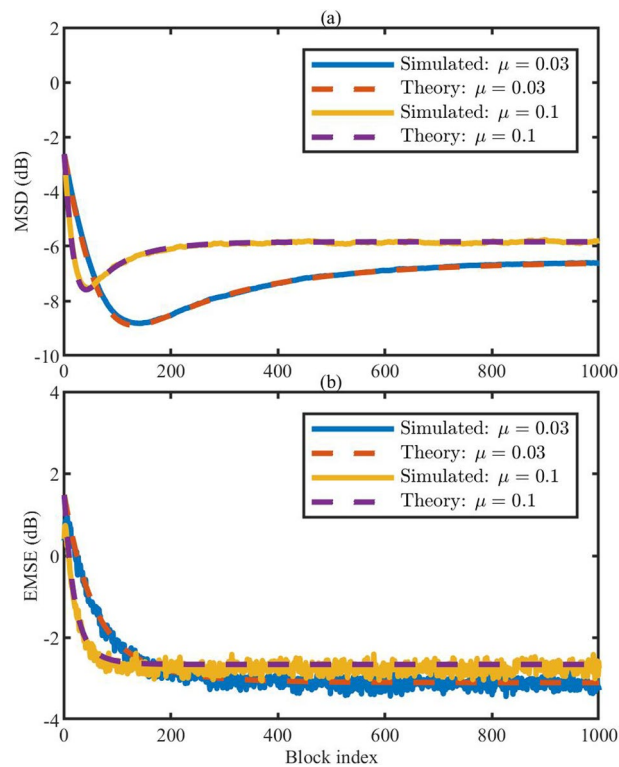

Fig. 6 Learning curves of the deficient-length unconstrained PBFDAFs. **a** MSD. **b** EMSE

Luo *et al. EURASIP Journal on Advances in Signal Processing* (2024) 2024:82

Page 17 of 20

Figure 7a illustrates that the steady-state MSD derived from the modified weight vector $\tilde{\mathbf{h}}(k)$ is lower than that obtained from $\tilde{\mathbf{w}}(k)$. This reduction in MSD is attributed to the steady-state MSD, and $\delta_m(\infty)$ does not depend on the initial conditions of the adaptive coefficients. In exact-modeling scenarios, the steady-state MSD and EMSE of the unconstrained PBFDAFs decrease as the step size is reduced [29]. However, in under-modeling scenarios, the steady-state MSD and EMSE remain almost unchanged when the step size is very small due to the under-modeling part $\mathbf{w}_r$. Additionally, the values $\delta_{m,min}(\infty)$ and $\xi_{ex,min}(\infty)$ accurately represent the lower bounds for the steady-state MSD and EMSE, respectively.

5 Conclusion

This paper provided a new statistical performance analysis of the deficient-length unconstrained PBFDAFs. Our findings indicate that the theoretical analysis is significantly simplified and more comprehensible compared to our previous work in [28] due to the utilization of the new weight vector $\hat{\mathbf{h}}(k)$. The developed theoretical model reveals new properties of unconstrained PBFDAFs in under-modeling scenarios and provides insights into the convergence behavior of the unconstrained PBFDAFs. For instance, the first $PL + L$ coefficients of the unknown plant can be recovered using the unconstrained PBFDAF for white noise input, whereas the constrained PBFDAF of the same order can only identify the first PL entries. In certain cases, the true impulse response can be recovered using the unconstrained PBFDAF for correlated input. A comparison
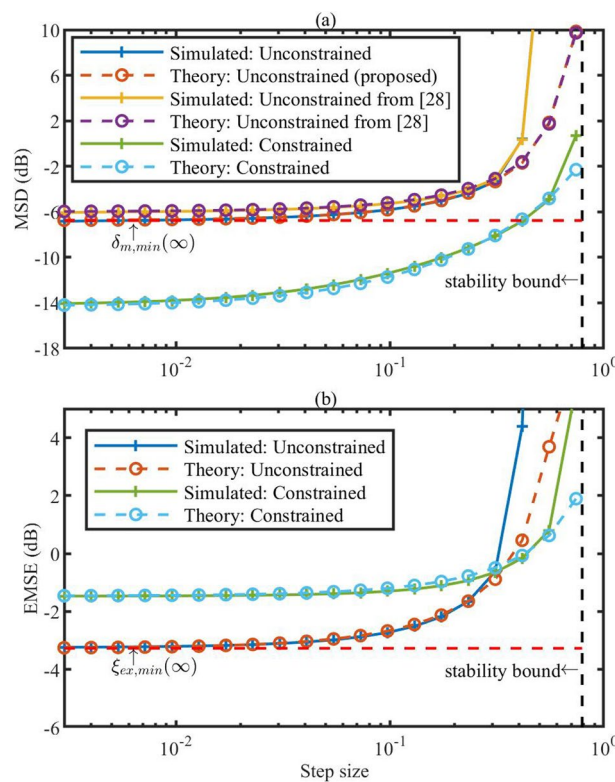

Fig. 7 Steady-state MSD and EMSE as a function of step size. **a** MSD. **b** EMSE

of steady-state MSD shows that the modified weight vector yields a lower MSD than the original one, which indicates that the unconstrained PBFDAFs exhibit a stronger system modeling capability than previously assumed. The theoretical model matches the simulation results well.

Appendix A Proof of $E[\tilde{\mathbf{h}}(\infty)] = \tilde{\mathbf{h}}_{\mathbf{opt}}$

Using (17) and the definition of $E[\mathbf{A}(k)]$ and $E[\mathbf{B}(k)]$, we have

$$E[\mathbf{A}(k)] = \boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T E[\mathbf{X}_b^T(k)\mathbf{X}_b(k)] = \boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T \mathbf{R}_b, \tag{A1}$$

$$E[\mathbf{B}(k)] = \boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T E[\mathbf{X}_b^T(k)\mathbf{X}_r(k)] = \boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T \mathbf{R}_{br}. \tag{A2}$$

Since the matrices $\mathbf{R}_b$ and $E[\mathbf{A}(k)]$ are invertible, the matrix $\boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T$ is invertible. We have

$$\begin{aligned} E[\tilde{\mathbf{h}}(\infty)] &= -(\boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T \mathbf{R}_b)^{-1}\boldsymbol{\alpha}\boldsymbol{\theta}\boldsymbol{\alpha}^T \mathbf{R}_{br}\mathbf{w}_r \\ &= -\mathbf{R}_b^{-1}\mathbf{R}_{br}\mathbf{w}_r \\ &= \tilde{\mathbf{h}}_{\text{opt}}. \end{aligned} \tag{A3}$$

Appendix B Detailed derivation of $\mathbf{z}(\infty)_{\mu \to 0}$ in (54)

Using (51), we have

$$\begin{aligned} \mathbf{z}(\infty) = {} & (\mathbf{D} - \mu\mathbf{J})^{-1} \\ & \times \{(-E[\mathbf{B}(k)]\otimes\mathbf{I} + \mu E[\mathbf{B}(k)\otimes\mathbf{A}(k)])\times\text{vec}[\tilde{\mathbf{h}}(k)\mathbf{w}_r^T] \\ & + (-\mathbf{I}\otimes E[\mathbf{B}(k)] + \mu E[\mathbf{A}(k)\otimes\mathbf{B}(k)])\times\text{vec}[\mathbf{w}_r\tilde{\mathbf{h}}^T(k)] \\ & + \mu E[\mathbf{B}(k)\otimes\mathbf{B}(k)]\times\text{vec}(\mathbf{w}_r\mathbf{w}_r^T) \\ & + \mu\sigma_v^2\text{vec}[E(\boldsymbol{\alpha}\boldsymbol{\theta}\bar{\mathbf{X}}_a^T(k)\bar{\mathbf{X}}_a(k)\boldsymbol{\theta}^T\boldsymbol{\alpha}^T)]\}. \end{aligned} \tag{B4}$$

Since the matrix $\mathbf{D}$ is invertible, we have

$$\begin{aligned} \lim_{\mu \to \infty} \mathbf{z}(\infty) = {} & -\mathbf{D}^{-1}\{(E[\mathbf{B}(k)]\otimes\mathbf{I})\times\text{vec}[\tilde{\mathbf{h}}(k)\mathbf{w}_r^T] \\ & + (\mathbf{I}\otimes E[\mathbf{B}(k)])\times\text{vec}[\mathbf{w}_r\tilde{\mathbf{h}}^T(k)]\} \\ & = \mathbf{z}(\infty)_{\mu \to 0}. \end{aligned} \tag{B5}$$

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Conflict of interests
The authors declare that they have no conflict of interest.

References

1. B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications* (Wiley, Hoboken, 2013)
2. M. Schneider, W. Kellermann, The generalized frequency-domain adaptive filtering algorithm as an approximation of the block recursive least-squares algorithm. EURASIP J. Adv. Signal Process. **2016**, 6 (2016)
3. J.M.P. Borrallo, M.G. Otero, On the implementation of a partitioned block frequency-domain adaptive filter (PBFDAF) for long acoustic echo cancellation. Signal Process. **27**(3), 301–315 (1992)
4. Z.M. Saric, I.I. Papp, D.D. Kukolj, I. Velikic, G. Velikic, Partitioned block frequency domain acoustic echo canceller with fast multiple iterations. Digital Signal Process. **27**, 119–128 (2014)
5. F. Kuech, E. Mabande, G. Enzner, State-space architecture of the partitioned-block-based acoustic echo controller, in Proceedings of IEEE International Conference on Acoustics, Speech, Signal Processing, pp. 1309–1313 (2014)
6. G. Bernardi, T. van Waterschoot, J. Wouters, M. Moonen, Adaptive feedback cancellation using a partitioned-block frequency-domain Kalman filter approach with PEM-based signal prewhitening. IEEE/ACM Trans. Audio Speech Lang. Process. **25**(9), 1784–1798 (2017)
7. D. Shi, W.-S. Gan, B. Lam, X. Shen, Comb-partitioned frequency-domain constraint adaptive algorithm for active noise control. Signal Process. **188**, 108222 (2021)
8. P. Sommen, P. Gerwen, H.J. Kotmans, A. Janssen, Convergence analysis of a frequency-domain adaptive filter with exponential power averaging and generalized window function. IEEE Trans. Circuits Syst. **34**(7), 788–798 (1987)
9. J. Lu, X. Qiu, H. Zou, A modified frequency-domain block LMS algorithm with guaranteed optimal steady-state performance. Signal Process. **104**, 27–32 (2014)
10. M.A. Jung, S. Elshamy, T. Fingscheidt, An automotive wideband stereo acoustic echo canceler using frequency-domain adaptive filtering. in Proceedings of IEEE European Signal Processing Conference, pp. 452–1456 (2014)
11. F. Yang, J. Yang, Mean-square performance of the modified frequency-domain block LMS algorithm. Signal Process. **163**, 18–25 (2019)
12. J.-S. Soo, K. Pang, Multidelay block frequency domain adaptive filter, IEEE Trans. Acoust., Speech, Signal Process., **38**(2), 373–376 (1990)
13. J.J. Shynk, Frequency-domain and multirate adaptive filtering. IEEE Signal Process. Mag., **9**(1), 14–37 (1992)
14. G.P.M. Egelmeers, P.C.W. Sommen, A new method for efficient convolution in frequency domain by nonuniform partitioning for adaptive filtering. IEEE Trans. Signal Process. **44**(12), 3123–3129 (1996)
15. B. Farhang-Boroujeny, Analysis and efficient implementation of partitioned block LMS adaptive filters. IEEE Trans. Signal Process. **44**(11), 2865–2868 (1996)
16. J.K. Wu, J. Casebeer, N.J. Bryan, P. Smaragdis, Meta-learning for adaptive filters with higher-order frequency dependencies, in Proceedings of International Workshop on Acoustics Signal Enhancement (2022)
17. T. Haubner, A. Brendel, W. Kellermann, End-to-end deep learning-based adaptation control for frequency-domain adaptive system identification, in Proceedings of IEEE International Conference on Acoustics, Speech Signal Processing, pp. 766–770 (2022)
18. J. Casebeer, N.J. Bryan, P. Smaragdis, Meta-AF: meta-learning for adaptive filters. IEEE/ACM Trans. Audio Speech Lang. Process. **31**, 355–370 (2023)
19. D. Mansour, A.H. Gray, Unconstrained frequency-domain adaptive filter, IEEE Trans. Acoust. Speech Signal Process., **30**(5), 726–734 (1982)
20. R.M.M. Derkx, G.R.M. Egelmeers, P.C.W. Sommen, New constraining method for partitioned block frequency-domain adaptive filters. IEEE Trans. Signal Process. **50**(9), 2177–2186 (2002)
21. M.L. Valero, E. Mabande, E.A.P. Habets, An alternative complexity reduction method for partitioned-block frequency-domain adaptive filters. IEEE Signal Process. Lett. **23**(5), 66–672 (2016)
22. X. Zhang, Y. Xia, C. Li, L. Yang, D.P. Mandic, Analysis of the unconstrained frequency-domain block LMS for second-order noncircular inputs. IEEE Trans. Signal Process. **67**(159), 3970–3984 (2019)
23. I.J. Umoh, T. Ogunfunmi, An adaptive nonlinear filter for system identification. EURASIP J. Adv. Signal Process. **2009**, 859698 (2009)
24. E. Moulines, O.A. Amrane, Y. Grenier, The generalized multidelay adaptive filter: structure and convergence analysis. IEEE Trans. Signal Process. **43**(1), 14–28 (1995)
25. K.S. Chan, B. Farhang-Boroujeny, Analysis of the partitioned frequency-domain block LMS (PFBLMS) algorithm. IEEE Trans. Signal Process. **49**(9), 1860–1874 (2001)
26. D. Comminiello, A. Nezamdoust, S. Scardapane, M. Scarpiniti, A. Hussain, A. Uncini, A new class of efficient adaptive filters for online nonlinear modeling. IEEE Trans. Syst. Man Cybern. Syst. **53**(3), 1384–1396 (2023)
27. F. Yang, G. Enzner, J. Yang, On the convergence behavior of partitioned-block frequency-domain adaptive filters. IEEE Trans. Signal Process. **69**, 4906–4920 (2021)

28. F. Yang, Analysis of deficient-length partitioned-block frequency-domain adaptive filters. IEEE/ACM Trans. Audio Speech Lang. Process. **30**, 456–467 (2022)
29. F. Yang, Analysis of unconstrained partitioned-block frequency-domain adaptive filters. IEEE Signal Process. Lett. **29**, 2377–2381 (2022)
30. E. Eweda, Convergence analysis of adaptive filtering algorithms with singular data covariance matrix. IEEE Trans. Signal Process. **49**(2), 334–343 (2001)
31. D.S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas* (Princeton University Press, Princeton, 2009)

Publisher's Note

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