Performance analysis of the deficient length NSAF algorithm and a variable step size method for improving its performance

Yi Yu^{1,2}, Haiquan Zhao^{*1,2}

Abstract: In all presented analyses of the normalized subband adaptive filter (NSAF) algorithm, there is a common assumption that the length of the adaptive filter is equal to that of the unknown system. In many practices, however, the adaptive filter usually works in an under-modeling situation. Namely, the length of the adaptive filter is less than that of the unknown system. Therefore, **for this case, the existing analysis results for the NSAF algorithm are not applicable.** In this paper, we analyze the performance of the deficient length NSAF algorithm based on some reasonable assumptions and approximations. More precisely, the expressions that characterize the transient-state and steady-state mean-square behavior of the algorithm are presented. Simulation results in various scenarios support our theoretical expressions. **In addition, based on our analyses, a variable step size NSAF algorithm suitable for the under-modeling case is developed, which improves the performance.**

Keywords: Normalized subband adaptive filter algorithm, deficient length adaptive filter, transient and steady-state analyses, colored input, variable step size.

1. Introduction

As a subfield of modern signal processing theory, adaptive filtering algorithms play a very important role in some practical applications such as system identification, active noise control, beamforming, channel equalization, and echo cancellation [1]-[3]. Among the existing algorithms, the normalized least mean square (NLMS) is well-known, due to its simplicity and robustness to the power of input signal. Furthermore, to obtain both fast convergence rate and low steady-state error, many variable step size NLMS algorithms were proposed [4]-[6]. Nevertheless, these algorithms will suffer from slow convergence when the input signals are **colored, also called the correlated signals.**

To deal with this problem, in a recent decade, the subband adaptive filter (SAF) has received significant attention. In subband adaptive filtering, the colored input signal is divided into almost mutually exclusive multiple subband signals and

¹ Key Laboratory of Magnetic Suspension Technology and Maglev Vehicle, Ministry of Education, Chengdu, 610031, China.

² School of Electrical Engineering, Southwest Jiaotong University, Chengdu, 610031, China.

e-mail: yuyi_xyuan@163.com (Y. Yu), hqzhao_swjtu@126.com (H. Zhao).

* Corresponding author

then critically decimated so that each decimated subband signal is approximately white, thereby improving the convergence performance [3]. In contrast to the conventional subband structure, the multiband structure of SAF has no aliasing and band edge effects, and thus it is more effective [3], [7]. Based on this multiband-structure, Lee and Gan proposed the normalized SAF (NSAF) algorithm by means of the principle of least perturbation [7]. This algorithm exhibits faster convergence rate than the NLMS algorithm for the colored input signals, due mainly to the fact that it inherits the decorrelating property of SAF [8]. Apart from the above property, the NSAF algorithm has almost the same computational complexity as the NLMS algorithm for applications of long adaptive filter. As a matter of fact, the NSAF algorithm will reduce to the NLMS algorithm when number of subbands is one. Following this algorithm, to balance a tradeoff of the NSAF algorithm between the fast convergence rate and low final estimation error which is controlled by the fixed step size and/or regularization constant, several variable step size [9]-[12] and/or variable regularization constant [13]-[15] versions were developed successively.

The performance analysis is a very important research topic for adaptive filtering algorithm, because it is very beneficial to predict the behavior of a specific adaptive filtering algorithm and to provide some guidelines to further improve the filter performance [16]-[26]. In [16] and [17], the steady-state mean-square error (MSE) results of the NSAF algorithm for the fixed step size and regularization constant were analyzed, respectively. However, in system identification and channel estimation scenarios, analyzing the mean-square-deviation (MSD) of the algorithm is more reasonable than the MSE of that, because the goal in these applications is to identify the system impulse response. In [18], Yin and Mehr studied the MSD behavior (including the transient-state and steady-state) of the NSAF algorithm based on the paraunitary assumption of the analysis filter bank. Moreover, in this analysis, three methods were used to solve some hyperelliptic integrals, i.e., the Lobatto quadrature, chi-square, and partial fraction methods, at the expense of high computational complexity especially for a long filter. The previous theoretical analyses in [16]-[18] obtain good agreement with the simulated results, but they require the assumption that the length of the adaptive filter is equal to that of the unknown impulse response. However, in many practical applications such as acoustic echo cancellation [4], [5], the length of unknown impulse response is unknown and very large, which would yield under-modeling situation. Under-modeling here means that, the length of the adaptive filter is less than that of the unknown impulse response, which is also called the deficient length. Consequently, the previous theoretical results on the sufficient length NSAF algorithm do not necessarily apply to the deficient length situation. For such scenarios, the performance of many algorithms has been studied in the literature such as the LMS [19], [20], the frequency-domain block LMS (FBLMS) [21], and the distributed LMS [22], [23]. To the best of our knowledge, however, there are no available studies to accurately evaluate the performance of the deficient length NSAF algorithm. To this end, benefited from [19]-[23], we will focus on analyzing the performance of the deficient length NSAF algorithm under some reasonable assumptions and

approximations, involving the transient-state and steady-state behaviors. Following the previous analysis processes, we also derive a variable step size NSAF algorithm which provides good performance in under-modeling situation. The proposed theoretical results and algorithm are also supported by extensive simulations.

The remainder of this paper is organized as follows. In the next section, we briefly review the NSAF algorithm. In Section 3, the performance of the deficient length NSAF algorithm is analyzed in mean and mean-square senses. In Section 4, simulations are performed to verify the theoretical analysis. Section 5 develops the variable step size NSAF algorithm for under-modeling. Finally, conclusions are presented in Section 6.

Notation: for the ease of reference, the main notations used in this paper are as follows: $(\cdot)^T$ denotes the transpose of a vector or matrix; $E\{\cdot\}$ denotes the mathematical expectation of a random variable; $\|\cdot\|$ stands for the l_2 -norm of a vector; $\mathbf{0}_{M \times L}$ is a $M \times L$ zero matrix; and $\mathbf{I}_L$ is the identity matrix of size $L \times L$.

2. The standard NSAF algorithm

Consider the desired signal $d(n)$ that arises from the model

$$d(n) = \mathbf{u}_{Lopt}^T(n) \mathbf{w}_{Lopt} + \eta(n), \quad (1)$$

where $\mathbf{w}_{Lopt} = [w_{1,o}, w_{2,o}, \dots, w_{L,o}]^T$ is the unknown L -dimensional vector to be estimated with an adaptive filter, $\mathbf{u}_{Lopt}(n) = [u(n), u(n-1), \dots, u(n-L+1)]^T$ is the input signal vector, and $h(n)$ is the system noise.

Fig. 1 shows the multiband-structure diagram of the SAF [7] which has been used to derive the NSAF family, where N denotes number of subbands. The desired signal $d(n)$ and input data $u(n)$ are partitioned into multiple subband signals $d_i(n)$ and $u_i(n)$ through the analysis filter bank $\{H_i(z), i = 0, 1, \dots, N-1\}$, respectively. The subband signals $y_{i,D}(k)$ and $d_{i,D}(k)$ are obtained by critically decimating $y_i(n)$ and $d_i(n)$. Here, n and k are used to indicate the original sequences and the decimated sequences, respectively. It is easy to know that the subband error signals $e_{i,D}(k)$ for $i = 0, 1, \dots, N-1$ are expressed as

$$\begin{aligned} e_{i,D}(k) &= d_{i,D}(k) - y_{i,D}(k) \\ &= d_{i,D}(k) - \mathbf{u}_i^T(k) \mathbf{w}(k) \end{aligned} \quad (2)$$

where $\mathbf{w}(k) = [w_1(k), w_2(k), \dots, w_M(k)]^T$ denotes the tap-weight vector of the adaptive filter with length M , $\mathbf{u}_i(k) = [u_i(kN), u_i(kN-1), \dots, u_i(kN-M+1)]^T$, and $d_{i,D}(k) = d_i(kN)$.

In [7], the standard NSAF algorithm for updating the tap-weight vector is expressed as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \sum_{i=0}^{N-1} \frac{e_{i,D}(k) \mathbf{u}_i(k)}{\|\mathbf{u}_i(k)\|^2} \quad (3)$$

where μ is the step-size. In practical applications, the length of the unknown system might be very large, which would result in under-modeling situation, i.e., $M < L$ [4], [5]. In this case, therefore, we will study the performance of the NSAF algorithm.

Fig. 1. Multiband-structure diagram of SAF

3. Performance analyses

3.1 Some assumptions and preliminaries

In order to convenience of analysis, the following assumptions are necessary.

Assumption 1: The system additive noise $\eta(n)$ is a white Gaussian process with zero-mean and variance σ_η^2 .

Assumption 2: The cosine modulated analysis filter bank for partitioning the input signal $u(n)$ and the observed output $d(n)$ is assumed to be paraunitary [16], [18], [24], [26]. Based on this assumption, we have

$$d_{i,D}(k) = \mathbf{u}_{Lopt,i}^T(k) \mathbf{w}_{Lopt} + \eta_i(k) \quad (4)$$

where the subband noises $\eta_i(k)$ for $i = 0, 1, \dots, N-1$ are also zero-mean white but variances $\sigma_{\eta_i}^2 = \sigma_\eta^2 / N$ which is obtained from $\eta(n)$ through the analysis filter bank.

Assumption 3: The tap-weight vector $\mathbf{w}(k)$, the subband input vector $\mathbf{u}_i(k)$, and the subband noise $\eta_i(k)$ are statistically independent. This type of assumption is commonly done in the context of adaptive filtering algorithms, see [1], [16]-[26].

Assumption 4: Due to the inherent decorrelating property of SAF, each decimated subband input signal can be assumed to be white, i.e., $E\{\mathbf{u}_i(k)\mathbf{u}_i^T(k)\} \approx \mathbf{I}_M \sigma_{u_i}^2$ and $E\{\|\mathbf{u}_i(k)\|^2\} \approx M \sigma_{u_i}^2$ [17], [24], where the variance $\sigma_{u_i}^2$ is usually obtained by the ensemble average. Also, if the number of subbands N is larger, this assumption will be more reasonable. The reason behind this assumption has been explained in Sections 3.2.1 and 4.1.2 in [3] in detail.

Assumption 5: When adaptive filter is long, i.e., $M \gg 1$, it can assume that the fluctuation of $\|\mathbf{u}_i(k)\|^2$ from one iteration to the next is small enough [9], [17], [24].

For deficient-length estimation, $\mathbf{w}_{\text{Lopt}}$ and $\mathbf{u}_{\text{Lopt},i}(k)$ can be split into two parts, respectively, i.e.,

$$\mathbf{w}_{\text{Lopt}} = \begin{bmatrix} \mathbf{w}_{\text{Mopt}} \\ \mathbf{w}_{\text{Lopt-M}} \end{bmatrix}, \quad \mathbf{u}_{\text{Lopt},i}(k) = \begin{bmatrix} \mathbf{u}_i(k) \\ \bar{\mathbf{u}}_i(k) \end{bmatrix} \quad (5)$$

where $\mathbf{w}_{\text{Mopt}}$ and $\mathbf{u}_i(k)$ are the first M entries of $\mathbf{w}_{\text{Lopt}}$ and $\mathbf{u}_{\text{Lopt},i}(k)$, respectively. In other words, $\mathbf{w}_{\text{Mopt}}$ is the part of $\mathbf{w}_{\text{Lopt}}$ that is estimated by $\mathbf{w}(k)$ at each iteration. With (4) and (5), (2) is rewritten as

$$\begin{aligned} e_{i,D}(k) &= \mathbf{u}_{\text{Lopt},i}^T(k) \mathbf{w}_{\text{Lopt}} + \eta_i(k) - \mathbf{u}_{\text{Lopt},i}^T(k) \begin{bmatrix} \mathbf{w}(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} \\ &= \mathbf{u}_{\text{Lopt},i}^T(k) \begin{bmatrix} \mathbf{w}_{\text{Mopt}} - \mathbf{w}(k) \\ \mathbf{w}_{\text{Lopt-M}} \end{bmatrix} + \eta_i(k) \\ &= \mathbf{u}_{\text{Lopt},i}^T(k) \begin{bmatrix} \tilde{\mathbf{w}}(k) \\ \mathbf{w}_{\text{Lopt-M}} \end{bmatrix} + \eta_i(k) \\ &= \mathbf{u}_{\text{Lopt},i}^T(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k) + \eta_i(k) \end{aligned} \quad (6)$$

where the $L \times 1$ error vector $\tilde{\mathbf{w}}_{\text{Lopt}}(k)$ indicates the difference between the estimated weight vector $\mathbf{w}(k)$ and the unknown vector $\mathbf{w}_{\text{Lopt}}$, i.e., $\tilde{\mathbf{w}}_{\text{Lopt}}(k) = [\tilde{\mathbf{w}}^T(k), \mathbf{w}_{\text{Lopt-M}}^T]^T$, where $\tilde{\mathbf{w}}(k) = \mathbf{w}_{\text{Mopt}} - \mathbf{w}(k)$ is a measure of difference in the fact that the first M entries of $\mathbf{w}_{\text{Lopt}}$ are estimated by $\mathbf{w}(k)$. Substituting (6) into (3), we have

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) + \mu \sum_{i=0}^{N-1} \frac{\mathbf{u}_i(k) \mathbf{u}_{\text{Lopt},i}^T(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k)}{\|\mathbf{u}_i(k)\|^2} \\ &\quad + \mu \sum_{i=0}^{N-1} \frac{\eta_i(k) \mathbf{u}_i(k)}{\|\mathbf{u}_i(k)\|^2}. \end{aligned} \quad (7)$$

3.2 Mean analysis

Subtracting (7) from $\mathbf{w}_{\text{Mopt}}$ and then taking the mathematical expectation under the condition of assumptions 3 and 5, we get (see Appendix A for derivation)

$$E\{\tilde{\mathbf{w}}(k+1)\} = \left(\mathbf{I}_M - \mu \sum_{i=0}^{N-1} \frac{E\{\mathbf{u}_i(k)\mathbf{u}_i^T(k)\}}{E\{\|\mathbf{u}_i(k)\|^2\}} \right) E\{\tilde{\mathbf{w}}(k)\} - \mu \sum_{i=0}^{N-1} \frac{E\{\mathbf{u}_i(k)\bar{\mathbf{u}}_i^T(k)\}}{E\{\|\mathbf{u}_i(k)\|^2\}} \mathbf{w}_{L_{\text{opt}}-M} \quad (8)$$

Since the last term of (8) is a constant in statistical sense, to guarantee the mean convergence of the algorithm, the maximal eigenvalue of the matrix $(\mathbf{I}_M - \mu\Phi)$ in the first term at the right side of (8) must be less than one [25], [26], where

$\Phi = \sum_{i=0}^{N-1} E\{\mathbf{u}_i(k)\mathbf{u}_i^T(k)\} / E\{\|\mathbf{u}_i(k)\|^2\}$, which yields the following inequality

$$0 < \mu < \frac{2}{\lambda_{\max}(\Phi)}, \quad (9)$$

which has the same form as the mean stability of the NSAF algorithm with sufficient length (i.e., $M = L$) [26]¹.

Based on assumption 4, we have $\Phi \approx (N/M)\mathbf{I}_M$ so that the mean convergence condition (9) can be approximated as

$$0 < \mu < \frac{2M}{N}. \quad (10)$$

Let us assume that the algorithm has converged to the steady-state under the condition of (9), i.e., $E\{\tilde{\mathbf{w}}(k+1)\} = E\{\tilde{\mathbf{w}}(k)\}$

for $k \rightarrow \infty$, the steady-state solution of the deficient length NSAF algorithm is obtained from (8) as²

$$E\{\mathbf{w}(\infty)\} = \mathbf{w}_{M_{\text{opt}}} + \Phi^{-1} \sum_{i=0}^{N-1} \frac{E\{\mathbf{u}_i(\infty)\bar{\mathbf{u}}_i^T(\infty)\}}{E\{\|\mathbf{u}_i(\infty)\|^2\}} \mathbf{w}_{L_{\text{opt}}-M} \quad (11)$$

According to assumption 4, we know $E\{\mathbf{u}_i(\infty)\bar{\mathbf{u}}_i^T(\infty)\} \approx \mathbf{0}_{M \times (M-L)}$, and this approximation is more reasonable for a larger number of subbands N . As a result, (11) becomes $E\{\mathbf{w}(\infty)\} \approx \mathbf{w}_{M_{\text{opt}}}$, which signifies that the deficient length NSAF algorithm is an approximately unbiased estimation in the mean for the first M entries of the unknown vector $\mathbf{w}_{L_{\text{opt}}}$.

3.3 Mean-square analysis

In the following, we concentrate on studying the MSD behavior of the deficient length NSAF algorithm. The MSD is defined as

$$\text{MSD}(k) = E\{\|\tilde{\mathbf{w}}_{L_{\text{opt}}}(k)\|^2\}. \quad (12)$$

¹ When the affine projection order is equal to one, the equation (15) in [26] is the convergence condition of the sufficient length NSAF in the mean.

² Note that the second term at the right side of (11) is zero in sufficient length (i.e., $M = L$) case.

To this end, padding the vectors with length M in (7) by $L-M$ zeros results in vectors with length L

$$\begin{aligned} \begin{bmatrix} \mathbf{w}(k+1) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} &= \begin{bmatrix} \mathbf{w}(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} + \mu \sum_{i=0}^{N-1} \begin{bmatrix} \mathbf{u}_i(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} \frac{\eta_i(k)}{\|\mathbf{u}_i(k)\|^2} \\ &\quad + \mu \sum_{i=0}^{N-1} \begin{bmatrix} \mathbf{u}_i(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} \frac{\mathbf{u}_{\text{Lopt}, i}^T(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k)}{\|\mathbf{u}_i(k)\|^2}. \end{aligned} \quad (13)$$

Subtracting both sides of (13) from $\mathbf{w}_{\text{Lopt}}$, we have

$$\tilde{\mathbf{w}}_{\text{Lopt}}(k+1) = \mathbf{A}(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k) - \mu \sum_{i=0}^{N-1} \begin{bmatrix} \mathbf{u}_i(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} \frac{\eta_i(k)}{\|\mathbf{u}_i(k)\|^2} \quad (14)$$

where

$$\mathbf{A}(k) = \mathbf{I}_L - \mu \sum_{i=0}^{N-1} \begin{bmatrix} \mathbf{u}_i(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} \frac{\mathbf{u}_{\text{Lopt}, i}^T(k)}{\|\mathbf{u}_i(k)\|^2}. \quad (15)$$

By taking the squared l_2 -norm of both sides of (14), we found the following relation

$$\begin{aligned} \|\tilde{\mathbf{w}}_{\text{Lopt}}(k+1)\|^2 &= \tilde{\mathbf{w}}_{\text{Lopt}}^T(k) \mathbf{A}^T(k) \mathbf{A}(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k) \\ &\quad - 2\mu \tilde{\mathbf{w}}_{\text{Lopt}}^T(k) \mathbf{A}^T(k) \sum_{i=0}^{N-1} \begin{bmatrix} \mathbf{u}_i(k) \\ \mathbf{0}_{(L-M) \times 1} \end{bmatrix} \frac{\eta_i(k)}{\|\mathbf{u}_i(k)\|^2} \\ &\quad + \mu^2 \sum_{i=0}^{N-1} \frac{\eta_i^2(k)}{\|\mathbf{u}_i(k)\|^2}. \end{aligned} \quad (16)$$

Taking the mathematical expectation of both sides of (16) and applying assumptions 2, 3, and 5, (16) becomes

$$\begin{aligned} \text{MSD}(k+1) &= E \left\{ \tilde{\mathbf{w}}_{\text{Lopt}}^T(k) \mathbf{A}^T(k) \mathbf{A}(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k) \right\} \\ &\quad + \mu^2 \sum_{i=0}^{N-1} \frac{\sigma_{\eta_i}^2}{E \left\{ \|\mathbf{u}_i(k)\|^2 \right\}}. \end{aligned} \quad (17)$$

To further proceed, we need to evaluate the first term of the right side of (17). Using assumption 3, this term can be represented as

$$\begin{aligned} E \left\{ \tilde{\mathbf{w}}_{\text{Lopt}}^T(k) \mathbf{A}^T(k) \mathbf{A}(k) \tilde{\mathbf{w}}_{\text{Lopt}}(k) \right\} &= \\ E \left\{ \tilde{\mathbf{w}}_{\text{Lopt}}^T(k) E \left\{ \mathbf{A}^T(k) \mathbf{A}(k) \right\} \tilde{\mathbf{w}}_{\text{Lopt}}(k) \right\}. \end{aligned} \quad (18)$$

Evidently, the calculation of $E \left\{ \mathbf{A}^T(k) \mathbf{A}(k) \right\}$ is necessary to obtain the MSD behavior of the algorithm. As shown in the

Appendix B, the following relation holds true, i.e.,

$$E \left\{ \mathbf{A}^T(k) \mathbf{A}(k) \right\} = \begin{bmatrix} \beta \mathbf{I}_M & \mathbf{0}_{M \times (L-M)} \\ \mathbf{0}_{(L-M) \times M} & \alpha \mathbf{I}_{L-M} \end{bmatrix} \quad (19)$$

where

$$\beta = 1 - 2\mu \frac{N}{M} + \mu^2 \frac{N}{M} \quad (20)$$

$$\alpha = 1 + \mu^2 \frac{N}{M} . \quad (21)$$

According to the definition of $\tilde{\mathbf{w}}_{\text{Lopt}}(k)$ in (5), we have

$$\|\tilde{\mathbf{w}}_{\text{Lopt}}(k)\|^2 = \|\tilde{\mathbf{w}}(k)\|^2 + \|\mathbf{w}_{\text{Lopt}-M}\|^2 . \quad (22)$$

Substituting (19) and (22) into (18), and after some manipulations, (18) can be changed as

$$\begin{aligned} E\left\{\tilde{\mathbf{w}}_{\text{Lopt}}^T(k)\mathbf{A}^T(k)\mathbf{A}(k)\tilde{\mathbf{w}}_{\text{Lopt}}(k)\right\} = \\ \beta \text{MSD}(k) + (\alpha - \beta) \|\mathbf{w}_{\text{Lopt}-M}\|^2 . \end{aligned} \quad (23)$$

By combining (17) and (23) yields

$$\begin{aligned} \text{MSD}(k+1) = \beta \text{MSD}(k) + \\ 2\mu \frac{N}{M} \|\mathbf{w}_{\text{Lopt}-M}\|^2 + \mu^2 \sum_{i=0}^{N-1} \frac{\sigma_{\eta_i}^2}{M \sigma_{u_i}^2} . \end{aligned} \quad (24)$$

The equation (24) describes the MSD learning curve of the deficient length NSAF algorithm. To ensure that the deficient length NSAF algorithm is convergent in the mean-square sense, the range of the step size can be obtained by imposing $|b| < 1$, which results in

$$0 < m < 2 . \quad (25)$$

It is worth to note that the result (25) is also the mean-square convergence condition of the sufficient length NSAF algorithm [7]. Let us rewrite (20) as $\beta = 1 + \hbar N$, where $\hbar = (\mu - 2)\mu/M$. According to (25) we know $\hbar < 0$, thus β will decrease as N increases. It follows that the convergence rate of the algorithm is faster for a larger N . If the algorithm with the step size in the range of (25) has gone into the steady-state, we have $\text{MSD}(k+1) = \text{MSD}(k) = \text{MSD}(\infty)$. Hence, the steady-state MSD of the algorithm can be obtained from (24) as

$$\text{MSD}(\infty) = \frac{2}{(2-\mu)} \|\mathbf{w}_{\text{Lopt}-M}\|^2 + \frac{\mu}{(2-\mu)N} \sum_{i=0}^{N-1} \frac{\sigma_{\eta_i}^2}{\sigma_{u_i}^2} . \quad (26)$$

So far, we have completed the mean-square analysis of the NSAF algorithm in the under-modeling case including the transient-state and steady-state results.

As shown in (26), the steady-state MSD of the deficient length NSAF algorithm consists of two components. Compared to

the steady-state MSD of the sufficient length NSAF (i.e., $M = L$), the first term on the right side of (26) is indeed an additional error when the NSAF algorithm works in the deficient length scenario (i.e., $M < L$). This term consists of all the omitting coefficients of the unknown vector $\mathbf{w}_{L_{\text{opt}}}$ in the adaptation process, i.e., $\mathbf{w}_{L_{\text{opt}}-M}$. Apparently, as much the selected length M of filter is less than L , this additional error will be larger. It is also observed that as μ approaches to zero, the steady-state MSD of the NSAF in the deficient length case will tend to a constant $\|\mathbf{w}_{L_{\text{opt}}-M}\|^2$, unlike in the sufficient length case the steady-state MSD tends to zero.

4. Simulations for theoretical analyses

Simulations are presented to evaluate the theoretical results of the deficient length NSAF algorithm in system identification setup. The unknown vector $\mathbf{w}_{L_{\text{opt}}}$ is randomly generated from the uniform distribution $[-1, 1]$ with $\mathbf{w}_{L_{\text{opt}}}^T \mathbf{w}_{L_{\text{opt}}} = 1$, unless otherwise specified. The tap-length M of the NSAF algorithm is always less than L . The input signal $u(n)$ is generated by the following autoregressive (AR) processes [18]: $u(n) = a_1 u(n-1) + z(n)$ and $u(n) = a_1 u(n-1) + a_2 u(n-2) + z(n)$ which respectively are called the AR(1) input and the AR(2) input, where $z(n)$ is a zero mean white Gaussian process, and a_1 and a_2 are the AR coefficients. The white Gaussian noise $h(n)$ is added to the output-end of the unknown system to give a certain signal signal-to-noise rate (SNR). The SNR is defined as $\text{SNR} = 10 \log_{10} \left(E \left\{ \left(\mathbf{u}^T(n) \mathbf{w}_{L_{\text{opt}}} \right)^2 \right\} / \sigma_{\eta}^2 \right)$. The cosine modulated filter banks are used for the subband structure; and to maintain 60 dB stopband attenuation, the length of the prototype filter is 32 and 64, respectively, for $N = 4$ and 8 subbands [7]. The MSD defined by $10 \log_{10} (\text{MSD}(k))$, (dB), is used to evaluate the previous analyses. Simulation results are averaged over 30 independent runs.

4.1 Transient-state behavior

In this section, we will examine the effectiveness of the theoretical transient model (24) for the deficient length NSAF algorithm in different scenarios given by the SNR, m , N , input signal, L and M . The length of the unknown vector is $L = 512$, except for Figs. 10-13. To further quantify the effectiveness of (24), the standard deviation (SD) between simulation and theory values of MSD is also computed, which is defined as $\text{SD}(k) = \sqrt{\frac{1}{30} \sum_{j=1}^{30} (\text{MSD}_j(k) - \text{MSD}_{th}(k))^2}$ [26], where $\text{MSD}_j(k)$ is the result of the j -th independent run, and $\text{MSD}_{th}(k)$ is the theory value from (24).

First of all, Fig. 2 shows a comparison of the theoretical and simulated MSD curves of the deficient length NSAF algorithm

with different lengths M and subbands N . The SNR is 30 dB, and the step size is chosen as $m = 0.2$. As one can see, the theoretical results are very close to the simulated results. In Fig. 2, there is also a slight discrepancy in the transient stage due to the use of some assumptions for convenient for analysis. By observing Fig. 2 (a) and (b), we can find that as compared to $N = 4$, this discrepancy in the case of $N = 8$ is smaller. This is because that the decimated subband input signals are closer to white signal for a larger number of subbands, as explained in assumption 4. This fact can also be verified from Fig. 3 which shows the SD curves corresponding to Fig. 2, since the SD for $N = 4$ is greater than the SD for $N = 8$. Figs. 4 and 5 show the MSD and SD curves of the deficient length NSAF algorithm in the case of 20 dB, respectively. The conclusions obtained from Figs. 4 and 5 are consistent with those of Figs. 2 and 3. In addition, the convergence rate of the deficient length NSAF does not depend on the SNR.

(a)

(b)

Fig. 2. The MSD curves of the deficient length NSAF algorithm in the case of SNR = 30 dB. (a) $N = 4$. (b) $N = 8$. [AR(1) input with $a_1=0.8$].

Fig. 3. The SD curves of corresponding to Fig. 2.

(a)

(b)

Fig. 4. The MSD curves of the deficient length NSAF algorithm in the case of SNR = 20 dB. (a) $N = 4$. (b) $N = 8$. [AR(1) input with $a_1=0.8$].

Fig. 5. The SD curves of corresponding to Fig. 4.

Then, Fig. 6 compares the theoretical and simulated MSDs of the deficient length NSAF algorithm using different step sizes $m = 0.1, 0.5$ and 1 . And, the corresponding SD results are depicted in Fig. 7. One can see that the theory values calculated by (24) match well with simulation values. Also, in Figs. 6 and 7, a large step size (e.g., $m = 1$) leads to fast convergence rate but meanwhile brings about large MSD or SD in the steady-state; conversely, a small step size (e.g., $m = 0.1$) provides low steady-state MSD or SD but slows the convergence rate. In other words, as the sufficient length (i.e., $M = L$) NSAF algorithm, the deficient length NSAF algorithm also exists a tradeoff problem between the convergence rate and steady-state error for the choice of the step size. To overcome this problem, designing variable step size method is a good choice, as shown in Section 5.

Fig. 6. The MSD curves of the deficient length NSAF algorithm for different step sizes. (a) $M = 500$. (b) $M = 480$. [AR(1) input with $a_1=0.8$, SNR = 30 dB, $N = 8$].

Fig. 7. The SD curves of corresponding to Fig. 6.

Figs. 8 and 9 depict the MSD and SD results of the deficient length NSAF for the input signals with different correlation levels, respectively. The correlation level of the input signal is characterized by the eigenvalue spread χ of its autocorrelation matrix $\mathbf{R}$, defined as $\chi = \lambda_{\max} / \lambda_{\min}$ [3], where $\lambda_{\max}$ and $\lambda_{\min}$ denote the maximum and minimum eigenvalues of $\mathbf{R}$. According to this definition, we have $\chi = 1$ for white input, $\chi = 35$ for AR(1) input with $a_1 = 0.7$, and $\chi = 135$ for AR(2) input with $a_1 = -0.1$ and $a_2 = -0.8$ [18]. Note that a larger value of χ means that the correlation of the input signal is stronger. As can be seen from these two figures, the difference between the theory values and the simulation values will become obvious when the input signal is highly correlated (e.g., $\chi = 135$). However, based on the assumption 4,

this difference when dealing with the strong correlation input signal can be mitigated by increasing number of subbands N . Overall, the theoretical results show a good match with the simulations.

Fig. 8. The MSD curves of the deficient length NSAF algorithm for the input signals with different correlation levels. [$m = 0.2$, $M = 480$, $\text{SNR} = 30$ dB, $N = 8$].

Fig. 9. The SD curves of corresponding to Fig. 8.

We examine the relation (24) for the deficient length NSAF algorithm in a low-order scenario, where $L = 64$. Correspondingly, the MSD and SD results are plotted in Figs. 10 and 11, respectively. Also, considering a higher-order case, where the unknown vector is a realistic acoustic echo path with length of $L = 1024$ [9] which is significantly larger than M . And, Figs. 12 and 13 shows the MSD and SD results of the deficient length NSAF algorithm. In these two extreme cases, the theory values are still in agreement with simulations values. Notice that in Figs. 3, 5, 7, 11 and 13, there is a common fact that for a larger M (when L is the same), the SD in the steady-state is smaller. This result is reasonable, since the MSD will be

smaller as M is closer to L .

Fig. 10. The MSD curves of the deficient length NSAF algorithm in the case of a low-order system. [AR(1) input with $a_1=0.8$, $m=0.2$, SNR = 30 dB, and

$N=8$].

Fig. 11 The SD curves of corresponding to Fig. 10.

Fig. 12. The MSD curves of the deficient length NSAF algorithm in the case of a higher-order system. [AR(1) input with $a_1=0.8$, $m= 0.2$, SNR = 30 dB, and $N = 8$].

Fig. 13 The SD curves of corresponding to Fig. 12.

4.2 Steady-state behavior

Finally, Figs. 14 and 15 illustrate the steady-state MSDs curves of the deficient length NSAF in various simulation environments. Here, the length of the unknown vector is also chosen as $L=512$. The theoretical results are directly calculated using (26), and the simulation results are obtained by averaging more than 1000 instantaneous MSDs in the steady-state.

Fig. 14 plots the steady-state MSDs of the deficient length NSAF algorithm versus deficient length M , under the condition of different N and SNR. The value of the step size is $m= 0.2$. Clearly, there is a good match between theory and simulation results. One can see from Fig. 14, as the length of the adaptive filter becomes more and more deficient, i.e., M is increasingly less than L , the steady-state MSD will become larger and larger. This conclusion can be observed directly from (26). In

addition, the steady-state MSD of the deficient length NSAF algorithm is not dependent of number of subbands N .

In Fig. 15, we compare the steady-state MSDs of the deficient length NSAF algorithm versus different step sizes. The step-size m varies from 0.01 to 1.99. As the step-size approaches the stability bound at $m = 2$, the steady-state MSD of the deficient length NSAF algorithm increases tremendously. Besides, when the step size is close to zero, the steady-state MSD of the deficient length NSAF algorithm will tend to a constant but that of the sufficient length NSAF algorithm (i.e., $M = 512$) will be close to zero. This result is consistent with the theory relation (26).

(a)

(b)

Fig. 14. The steady-state MSDs of the deficient length NSAF algorithm for deficient lengths M . (a) SNR = 30 dB. (b) SNR = 20 dB. [AR(1) input with $a_1=0.8$ and $m=0.2$].

(a)

(b)

Fig. 15. The steady-state MSDs of the deficient length NSAF algorithm for different step sizes m . (a) SNR = 30 dB. (b) SNR = 20 dB. [AR(1) input with $a_1=0.8$ and $N=8$].

5. Variable step size algorithm for under-modeling

5.1 The derivation of algorithm

By assigning an individual time-varying step size $\mu_i(k)$ for each subband, (3) is changed as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \sum_{i=0}^{N-1} \mu_i(k) \frac{e_{i,D}(k) \mathbf{u}_i(k)}{\|\mathbf{u}_i(k)\|^2}. \quad (27)$$

Then, imposing the manipulations used for obtaining (24) on (27), we can obtain

$$\begin{aligned} \text{MSD}(k+1) = & \left(1 - \frac{2}{M} \sum_{i=0}^{N-1} \mu_i(k) + \frac{1}{M} \sum_{i=0}^{N-1} \mu_i^2(k) \right) \text{MSD}(k) + \\ & \frac{2}{M} \left\| \mathbf{w}_{\text{Lopt-M}} \right\|^2 \sum_{i=0}^{N-1} \mu_i(k) + \sum_{i=0}^{N-1} \mu_i^2(k) \frac{\sigma_{\eta_i}^2}{M \sigma_{u_i}^2} \end{aligned} \quad (28)$$

Setting the derivative of (28) with respect to $\mu_i(k)$ for $i=0, 1, \dots, N-1$ to be zero, we get the following relation:

$$\mu_i(k) = \frac{\text{MSD}(k) \sigma_{u_i}^2 - \left\| \mathbf{w}_{\text{Lopt-M}} \right\|^2 \sigma_{u_i}^2}{\text{MSD}(k) \sigma_{u_i}^2 + \sigma_{\eta_i}^2} \quad (29)$$

Substituting (12) and (22) into (29), (29) becomes

$$\mu_i(k) = \frac{E\{e_{a,i}^2(k)\}}{E\{e_{a,i}^2(k)\} + \sigma_{Ae}^2} \quad (30)$$

where $E\{e_{a,i}^2(k)\} = E\{\|\tilde{\mathbf{w}}(k)\|^2\} \sigma_{u_i}^2$ denotes the power of the disturbance-free subband error $e_{a,i}(k) = \mathbf{u}_i^T(k) \tilde{\mathbf{w}}(k)$, and

$\sigma_{Ae}^2 = \left\| \mathbf{w}_{\text{Lopt-M}} \right\|^2 \sigma_{u_i}^2 + \sigma_{\eta_i}^2$ stands for the contaminated noise power. In general, the power $E\{e_{a,i}^2(k)\}$ can be estimated in a recursive way as [4], [11]:

$$\hat{\sigma}_{e_{a,i}}^2(k) = \theta \hat{\sigma}_{e_{a,i}}^2(k-1) + (1-\theta) e_{a,i}^2(k) \quad (31)$$

where θ is a weighting factor chosen as $\theta = 1 - N/(\kappa M)$ with $\kappa > 1$. According to assumptions 3 and 4, and using (5), we have $E\{e_{i,D}^2(k)\} = E\{e_{a,i}^2(k)\} + \sigma_{Ae}^2$. Hence, similar to [11], [28], the disturbance-free subband error $e_{a,i}(k)$ in (31) can be recovered from the subband error signal $e_{i,D}(k)$ by using the *noniterative shrinkage* method, as follows:

$$e_{a,i}(k) = \text{sign}(e_{i,D}(k)) \max(|e_{i,D}(k)| - f_i, 0) \quad (32)$$

where $\text{sign}(\cdot)$ is the sign function, and f_i is the threshold parameter which is chosen by $f_i = \sqrt{\pi \sigma_{Ae}^2}$ with $\pi \geq 1$.

Obviously, to implement the propose VSS scheme (30), σ_{Ae}^2 appeared in (30) and (32) must be determined. Under assumptions 3 and 4, we also have $\sigma_{Ae}^2 = E\{d_{i,D}^2(k)\} - E\{d_{M,i}^2(k)\}$, where $d_{M,i}(k) = \mathbf{u}_i^T(k) \mathbf{w}_{\text{Mopt}}$. Also, let us suppose that the adaptive filter coefficients have converged to a certain degree, it can be considered that $E\{d_{M,i}^2(k)\} \cong E\{y_{i,D}^2(k)\}$ [4], [29]. Therefore, σ_{Ae}^2 can be estimated as:

$$\hat{\sigma}_{Ae}^2(k) \cong \left| \hat{\sigma}_{d_{i,D}}^2(k) - \hat{\sigma}_{y_{i,D}}^2(k) \right| \quad (33)$$

where $\hat{\sigma}_{d_{i,D}}^2(k)$ and $\hat{\sigma}_{y_{i,D}}^2(k)$ represent the estimates of $E\{d_{i,D}^2(k)\}$ and $E\{y_{i,D}^2(k)\}$, respectively, and are computed using the same way as (31), i.e.,

$$\hat{\sigma}_{d_{i,D}}^2(k) = \theta \hat{\sigma}_{d_{i,D}}^2(k-1) + (1-\theta) d_{i,D}^2(k), \quad (34)$$

$$\hat{\sigma}_{y_{i,D}}^2(k) = \theta \hat{\sigma}_{y_{i,D}}^2(k-1) + (1-\theta) y_{i,D}^2(k). \quad (35)$$

Considering the fact that the approximation (33) may be not effective at the starting stages of the adaptation, therefore, the step sizes $\mu_i(k)$ for $i=0, 1, \dots, N-1$ of the proposed variable step size NSAF for under-modeling (VSS-NSAF-UM) algorithm are expressed as

$$\mu_i(k) = \begin{cases} 1, & \text{for } k \leq M \\ \frac{\hat{\sigma}_{e_{a,i}}^2(k)}{\hat{\sigma}_{e_{a,i}}^2(k) + \hat{\sigma}_{Ae}^2(k)}, & \text{for } k > M \end{cases}. \quad (36)$$

Remark 1: The proposed VSS-NSAF-UM algorithm is similar at first glance to the new VSS-NSAF (NVSS-NSAF) algorithm presented in [11]. However, the NVSS-NSAF algorithm was derived based on an exact modeling case so that it is not suitable for the under-modeling case. Comparing with the NVSS-NSAF algorithm, the VSS-NSAF-UM algorithm has the following advantages: 1) it is an extension of the NVSS-NSAF in the under-modeling situation; it does not require knowing the power of the system noise in advance, since all the parameters in (36) have been estimated.

5.2 Verification of algorithm

Fig. 16 compares the performance of the proposed VSS-NSAF-UM algorithm with that of the NSAF with two step sizes and NVSS-NSAF algorithms in the under-modeling case. In this example, a realistic acoustic echo path $\mathbf{w}_{Lopt}$ with $L = 1024$ length needs to be identified, and the length of adaptive filter is 512. Also, an abrupt change for the echo path occurs in the middle of input samples by changing $\mathbf{w}_{Lopt}$ to $-\mathbf{w}_{Lopt}$. As can be seen, the proposed VSS-NSAF-UM is more practical than the NVSS-NSAF algorithm, due to its superiority performance for the case of under-modeling.

(a)

(b)

Fig. 16. The MSD curves of various NSAF-type algorithms for under-modeling case. (a) AR(1) input with $a_1=0.8$. (b) AR(2) input with $a_1 = -0.1$ and

$a_2 = -0.8$. [SNR = 30 dB and $N=8$]. Parameters setting of algorithms: $\kappa = 4$ and $\lambda = 4$ for the NVSS-NSAF [11]; $\kappa = 4$ and $\pi = 3$ for the

VSS-NSAF-UM.

6. Conclusion

In this study, we carried out a rather detailed mean and mean-square performance evaluation of the deficient length NSAF algorithm for the colored input. Under the assumptions 1-5, we derived a theoretical model that quantifies the evolution behavior of the MSD for the algorithm. We also gave a closed-form expression for predicting the steady-state MSD of this algorithm. Several simulation results have proved the validity of our analyses.

Some important results in our analyses were summarized. Compared to the sufficient length case, the deficient length

NSAF algorithm has the same mean-square stability condition and conflict between the convergence rate and steady-state error caused by the step size. Unlike the sufficient length case, as the step size tends toward zero, the steady-state MSD of the deficient length NSAF algorithm will approach to a constant rather than zero. And, this constant is determined by the coefficients of the unknown system vector that are not identified by adaptive filter.

In addition, the proposed VSS-NSAF-UM algorithm has been shown to have better performance than its counterparts in the under-modeling case in terms of convergence rate and steady-state MSD.

Acknowledgement

This work was partially supported by National Science Foundation of P.R. China (Grant: 61271340, 61571374 and 61433011).

Appendix A: Derivation of (8)

Subtracting (7) from $\mathbf{w}_{Mopt}$ and applying the definitions of $\tilde{\mathbf{w}}_{Lopt}(k)$ and $\mathbf{u}_{Lopt,i}(k)$, we obtain

$$\begin{aligned} \tilde{\mathbf{w}}(k+1) = & \tilde{\mathbf{w}}(k) - \mu \sum_{i=0}^{N-1} \frac{\mathbf{u}_i(k) \mathbf{u}_i^T(k)}{\|\mathbf{u}_i(k)\|^2} \tilde{\mathbf{w}}(k) \\ & - \mu \sum_{i=0}^{N-1} \frac{\mathbf{u}_i(k) \bar{\mathbf{u}}_i^T(k)}{\|\mathbf{u}_i(k)\|^2} \mathbf{w}_{Lopt-M} - \mu \sum_{i=0}^{N-1} \frac{\eta_i(k) \mathbf{u}_i(k)}{\|\mathbf{u}_i(k)\|^2} \end{aligned} \quad (A1)$$

Taking the expectation of both sides of (A1) and using assumption 3, (A1) becomes

$$\begin{aligned} E\{\tilde{\mathbf{w}}(k+1)\} = & E\{\tilde{\mathbf{w}}(k)\} - \mu \sum_{i=0}^{N-1} E\left\{ \frac{\mathbf{u}_i(k) \mathbf{u}_i^T(k)}{\|\mathbf{u}_i(k)\|^2} \right\} E\{\tilde{\mathbf{w}}(k)\} \\ & - \mu \sum_{i=0}^{N-1} E\left\{ \frac{\mathbf{u}_i(k) \bar{\mathbf{u}}_i^T(k)}{\|\mathbf{u}_i(k)\|^2} \right\} \mathbf{w}_{Lopt-M} \end{aligned} \quad (A2)$$

Then, based on assumption 5, (8) can be obtained from (A2).

Appendix B:

Proof of (19)

In this section, we will show how to obtain (19). Postmutiplying (15) by its transpose $\mathbf{A}^T(k)$ gives

$$\begin{aligned} \mathbf{A}^T(k) \mathbf{A}(k) = & \mathbf{I}_L - 2\mu \sum_{i=0}^{N-1} \frac{\mathbf{u}_{Lopt,i}(k) \left[\mathbf{u}_i^T(k) \quad \mathbf{0}_{1 \times (L-M)} \right]}{\|\mathbf{u}_i(k)\|^2} \\ & + \mu^2 \sum_{i=0}^{N-1} \frac{\mathbf{u}_{Lopt,i}(k) \mathbf{u}_{Lopt,i}^T(k)}{\|\mathbf{u}_i(k)\|^2} \end{aligned} \quad (B1)$$

where we employed the fact that the input signals of different subbands are mutually orthogonal, i.e., $\mathbf{u}_i^T(k)\mathbf{u}_j(k) \approx 0$, $i \neq j$, which has been reported in the derivation of the NSAF algorithm [7].

Taking the expectation of the equation above and using assumption 5, we get

$$\begin{aligned} E\{\mathbf{A}^T(k)\mathbf{A}(k)\} &= \mathbf{I}_L - 2\mu \sum_{i=0}^{N-1} \frac{E\left\{\mathbf{u}_{Lopt,i}(k) \begin{bmatrix} \mathbf{u}_i^T(k) & \mathbf{0}_{1 \times (L-M)} \end{bmatrix}\right\}}{E\{\|\mathbf{u}_i(k)\|^2\}} \\ &\quad + \mu^2 \sum_{i=0}^{N-1} \frac{E\left\{\mathbf{u}_{Lopt,i}(k)\mathbf{u}_{Lopt,i}^T(k)\right\}}{E\{\|\mathbf{u}_i(k)\|^2\}} \end{aligned} \quad (\text{B2})$$

From assumption 4, we obtain

$$E\left\{\mathbf{u}_{Lopt,i}(k) \begin{bmatrix} \mathbf{u}_i^T(k) & \mathbf{0}_{1 \times (L-M)} \end{bmatrix}\right\} = \begin{bmatrix} \sigma_{u_i}^2 \mathbf{I}_M & \mathbf{0}_{1 \times (L-M)} \\ \mathbf{0}_{1 \times (L-M)} & \mathbf{0}_{L-M} \end{bmatrix} \quad (\text{B3})$$

$$E\left\{\mathbf{u}_{Lopt,i}(k)\mathbf{u}_{Lopt,i}^T(k)\right\} = \sigma_{u_i}^2 \mathbf{I}_L. \quad (\text{B4})$$

Substituting (B3) and (B4) into (B2), and again using assumption 4 (i.e., $E\{\|\mathbf{u}_i(k)\|^2\} \approx M\sigma_{u_i}^2$), (19) can be proofed.

References

- [1] S. Haykin, Adaptive Filter Theory, 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002.
- [2] J. Benesty and Y. Huang, Adaptive signal processing—applications to real-world problems, Berlin, Germany: Springer-Verlag, 2003.
- [3] K. A. Lee, W. S. Gan, and S. M. Kuo, Subband adaptive filtering: theory and implementation, Hoboken, NJ: Wiley, 2009.
- [4] C. Paleologu, S. Ciochină, and J. Benesty, “Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation,” IEEE Signal Process. Lett., vol. 15, pp. 5–8, 2008.
- [5] C. Paleologu, S. Ciochina, and J. Benesty, “Double-talk robust VSS-NLMS algorithm for under-modeling acoustic echo cancellation,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Las Vegas, NV, 2008, pp. 245–248.
- [6] S. Ciochină, C. Paleologu, and J. Benesty, “An optimized NLMS algorithm for system identification,” Signal Process., vol. 118, pp. 115–121, 2016.
- [7] K. A. Lee and W. S. Gan, “Improving convergence of the NLMS algorithm using constrained subband updates,” IEEE Signal Process. Lett., vol. 11, no. 9, pp. 736–739, 2004.
- [8] K. A. Lee and W. S. Gan, “Inherent decorrelating and least perturbation properties of the normalized subband adaptive filter,” IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4475–4480, 2006.

- [9] J. Ni and F. Li, "A variable step-size matrix normalized subband adaptive filter," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 6, pp. 1290–1299, 2010.
- [10] J. W. Shin, N. W. Kong, and P. G. Park, "Normalized subband adaptive filter with variable step size," *Electron. Lett.*, vol. 48, no. 4, pp. 204–206, 2012.
- [11] Y. Yu, H. Zhao, and B. Chen, "A new normalized subband adaptive filter algorithm with individual variable step sizes," *Circuits, Systems, and Signal Process.*, vol. 35, no. 4, pp. 1407–1418, 2016.
- [12] Y. Yu, H. Zhao, Z. He, and B. Chen, "A robust band-dependent variable step size NSAF algorithm against impulsive noises," *Signal Process.*, vol. 119, pp. 203–208, 2016.
- [13] J. Ni and F. Li, "A variable regularization matrix normalized subband adaptive filter," *IEEE Signal Process. Lett.*, vol. 16, no. 2, pp. 105–108, 2009.
- [14] J. Ni and F. Li, "Normalised subband adaptive filter with new variable regularization matrix," *Electron. Lett.*, vol. 46, no. 18, pp. 1293–1295, 2010.
- [15] J. J. Jeong, K. Koo, G. Koo, and S. W. Kim, "Variable regularization for normalized subband adaptive filter," *Signal Process.*, vol. 104, pp. 432–436, 2014.
- [16] K. A. Lee, W. S. Gan, and S. M. Kuo, "Mean-square performance analysis of the normalized subband adaptive filter," in *Proc. 40th Asilomar Conf. Signals, Syst., Comput.*, 2006, pp. 248–252.
- [17] J. Ni and X. Chen, "Steady-state mean-square error analysis of regularized normalized subband adaptive filters," *Signal Process.*, vol. 93, pp. 2648–2652, 2013.
- [18] W. Yin and A. S. Mehr, "Stochastic analysis of the normalized subband adaptive filter algorithm," *IEEE Trans. Circuits Syst. I: Reg. Pap.*, vol. 58, no. 5, pp. 1020–1033, 2011.
- [19] K. Mayyas, "Performance analysis of the deficient length LMS adaptive algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2727–2734, 2005.
- [20] Y. Gu, K. Tang, H. Cui, and W. Du, "Convergence analysis of a deficient-length LMS filter and optimal-length sequence to model exponential decay impulse response," *IEEE Signal Process. Lett.*, vol. 10, no. 1, pp. 4–7, 2003.
- [21] M. Wu, J. Yang, Y. Xu, and X. Qiu, "Steady-state solution of the deficient length constrained FBLMS algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6681–6687, 2012.
- [22] G. Azarnia and M. A. Tinati, "Steady-state analysis of the deficient length incremental LMS adaptive networks," *Circuits, Systems, and Signal Process.*, vol. 34, no. 9, pp. 2893–2910, 2015.

- [23] G. Azarnia and M. A. Tinati, "Steady-state analysis of the deficient length incremental LMS adaptive networks with noisy links," *AEU-Int. J. Electron. Commun.*, vol. 69, no. 1, pp. 153-162, 2015.
- [24] Y. Yu, H. Zhao, and B. Chen, "Steady-state mean-square-deviation analysis of the sign subband adaptive filter algorithm," *Signal Process.*, vol. 120, pp. 36-42, 2016.
- [25] H. C. Shin, A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 90-102, 2004.
- [26] F. Yang, M. Wu, P. Ji, Z. Kuang, and J. Yang, "Transient and steady-state analyses of the improved multiband-structured subband adaptive filter algorithm," *IET Signal Process.*, vol. 9, no. 8, pp. 596-604, 2015.
- [27] T. T. Soong, *Fundamentals of Probability and Statistics for Engineers*, John Wiley, 2004.
- [28] M. Z. A. Bhotto, A. Antoniou, "A family of shrinkage adaptive filtering algorithms," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1689-1697, 2013.
- [29] C. Paleologu, J. Benesty, and S. Ciochina, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp. 1466-1478, 2008.