IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 20 December 2023, accepted 9 January 2024, date of publication 12 January 2024, date of current version 22 January 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3353691

RESEARCH ARTICLE

A Variable Step-Size Regularization-Based Quasi-Newton Adaptive Filter for System Identification

MEHDI BEKRANI AND HADI ZAYYANI, (Member, IEEE)
Faculty of Electrical and Computer Engineering, Qom University of Technology, Qom 3718146645, Iran
Corresponding author: Mehdi Bekrani (bekrani@qut.ac.ir)

ABSTRACT The family of least mean square (LMS) based adaptive filtering algorithms suffers from convergence performance limitation due to the sensitivity of such algorithms to the eigenvalue spread of the input correlation matrix. The quasi-Newton family of adaptive filtering algorithms addresses this limitation, but its performance is restricted by the estimation accuracy of the correlation matrix inverse, especially for highly correlated input signals. Furthermore, the convergence rate and the steady-state performance of both LMS and quasi-Newton families are thoroughly depending on their step-sizes. In this paper, a variable step-size regularized quasi-Newton adaptive algorithm is proposed in the context of system identification. In this algorithm, inspired by the matrix inversion lemma, a modified regularized matrix inverse with a time-varying regularization is computed such that during the convergence, the contribution of matrix inverse in the weight update is reduced, resulting in a more noise-robust algorithm. The paper further provides a convergence analysis of the proposed quasi-Newton algorithm, wherein a variable step-size is proposed to achieve a high initial convergence rate and a low steady-state error in the context of system identification applications.

INDEX TERMS Adaptive filter, quasi-newton algorithm, regularization, system identification, variable step-size.

I. INTRODUCTION

System identification is one of the significant topics of signal processing which is widely used in many areas including acoustic echo cancellation [1], [2], [3], active noise control [4], acoustic channel equalization [5], source separation [6], and wireless multipath channels [7].

So far, numerous adaptive filtering algorithms have been proposed for system identification. The least mean square (LMS) and normalized LMS (NLMS) adaptive algorithms are of the most comment due to their simplicity and low complexity at the expense of a low convergence speed [8], [9]. The affine projection algorithm (APA) is an extension of NLMS which exhibits a faster convergence rate due to a less dependency on the eigenvalue spread of the input correlation matrix [9]. The convergence speed of APA increases with the projection order, however, the computational complexity increases accordingly, and potential numerical instability in matrix inverse computation pose challenges. Although low-complexity variants of APA have been suggested [10] at the cost of convergence degradation, their computational complexity remains high, and performance is sensitive to the spectral variation of the input signal and its statistics.

The quasi-Newton algorithm and its variants represent another class of adaptive algorithms with reduced dependency on eigenvalue spread [11], [12], [13], [14], [15], [16]. The quasi-Newton algorithm shows considerable convergence performance improvements compared to LMS/NLMS for highly correlated input signals [9]. In an ideal scenario where the input correlation matrix is known *a priori*, the quasi-Newton algorithm is insensitive to the eigenvalue spread, achieving the desired response rapidly [9]. In practice, the inverse of the correlation matrix is not priori known and need to be estimated. Various methods,

The associate editor coordinating the review of this manuscript and approving it for publication was Dong Shen.

including recursive approaches based on matrix inversion lemma (MIL) [13], [14], auto-regressive model-based inverse approximation [15], matrix diagonalization [16], utilizing the inverse of the input power spectrum [17], [18], and autocorrelation matrix inverse approximation using Fourier transform [11] and iterative algorithms [19], are employed for this purpose.

To enhance the convergence performance of the quasi-Newton algorithm, diverse methods have been proposed. In [15], a hybrid algorithm combining quasi-Newton and NLMS is proposed to utilize the advantages of these algorithms. This hybrid algorithm acts as quasi-Newton with considerable actuating data and NLMS with the weaker actuating data. In addition, [14] introduces a quasi-Newton algorithm with selective data, determined by output error amplitude. While reducing computational complexity, this approach does not enhance convergence speed.

In addition to the aforementioned contributions, various variable step-size adaptive algorithms, as cited in [8], [20], [21], and [22], and others, have been introduced. These algorithms aim both to improve the convergence rate and reduce the steady-state error in comparison to their fixed step-size counterparts.

In [21], a low-cost variable step-size NLMS algorithm is proposed such that the step-size is dynamically adjusted as a function of the square error. Specifically, when mean squared error (MSE) is high, the step-size is maintained at its maximum, and as the MSE decreases, the step-size is proportionally reduced, contributing to further MSE reduction. Additionally, in [22], a variable step-size fast NLMS algorithm is introduced with the aim of enhancing convergence speed and tracking capability for system identification. The adaptation process begins with a higher convergence speed, and subsequently, the step-size is reduced to achieve a relatively low MSE during stationary steady-state periods. Moreover, the algorithm responds to system variations by automatically adjusting the step-size to its maximum values, thereby enabling effective tracking ability.

In the work presented in [23], variable step-size l_0-norm constraint NLMS algorithms are developed for sparse system identification. The variable step-size schemes are derived by minimizing the forthcoming mean square deviations (MSDs), and the optimal step-size in each iteration is determined through an upper bound of the MSD derivation.

In [24], a variable step-size APA is introduced for the processing of noncircular signals. The variable step-size is determined by minimizing the power of the augmented noise-free *a posteriori* error vector, resulting in accelerated convergence and reduced steady-state misalignment.

In the method developed in [25], a variable step-size gradient descent total least-squares method is proposed to adapt to changes in the signal-to-noise ratio (SNR) for direction of arrival (DOA) estimation. In this approach, the variable step-size strategy is derived by incorporating the instantaneous augmented weight vector and estimated input signal power.

In [26], a variable parameter LMS algorithm based on the generalized maximum correntropy criterion (MCC) is presented for graph signal processing. The generalized MCC exhibits robustness to impulse noise in adaptive filtering. This algorithm introduces a variable recursive step-size strategy based on the average value of the error signal of sampled nodes on the graph, aiming to address the trade-off between fast convergence and small steady-state error.

Taking advantage of MCC, [27] proposes an affine projection MCC with correntropy-induced metric algorithm to counteract the adverse effects of impulsive noise on filter weight updates in sparse systems. In a similar vein, [28] introduces an affine projection sign algorithm (APSA) to mitigate the impact of impulsive noise on weight updates for adaptive filters. This algorithm adopts a variable step-size based on cascaded component filter modules utilizing a weighted MSE to alleviate the effects of impulsive noise on the adaptive filtering system. Furthermore, [29] proposes a variable regularization APSA to combat the impulsive noisy environment, incorporating the time correlation of the input signal and error to adjust the regularization parameter.

It is worth noting that many of the aforementioned time-varying step-size algorithms rely on feedback from the estimated MSE, and consequently, their efficiency for system identification applications, where reducing the MSD error is crucial, may be limited.

In this paper, a modified quasi-Newton adaptive algorithm is proposed. The proposed method incorporates a time-varying regularization in the recursive matrix inverse computation inspired by the matrix inverse lemma. Time-varying regularization is applied to diminish the contribution of matrix inverse in the weight update during the convergence, resulting in a more noise-robust algorithm. Furthermore, based on the convergence analysis of the proposed modified quasi-Newton algorithm, a variable step-size is proposed to provide a high initial convergence rate and a low steady-state mean square weight deviation for application in the system identification.

II. ADAPTIVE SYSTEM IDENTIFICATION

Figure 1 illustrates the schematic diagram of the adaptive system identification. The input signal to the adaptive filter is assumed a stationary signal represented by the zero-mean Gaussian vector $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^T$ with variance σ_x^2, where the superscript T denotes the transpose operator. The observation noise $\omega(n)$ is assumed to be zero-mean white Gaussian with variance σ_ω^2 and is independent of $\mathbf{x}(n)$. The desired signal $d(n)$, which arises from the output of the unknown system impulse response $\mathbf{h}(n) = [h_0(n), h_1(n), \ldots, h_{N-1}(n)]^T$, is expressed as

$$d(n) = \mathbf{h}^T(n)\mathbf{x}(n) + \omega(n), \tag{1}$$

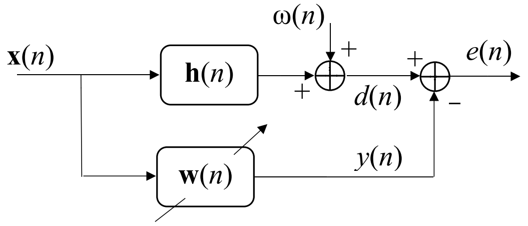

FIGURE 1. **The schematic diagram of the adaptive system identification.**

We assume $\mathbf{h}(n)$ as a time-varying system with standard random walk model as [30]

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mathbf{q}(n) \tag{2}$$

where $\mathbf{q}(n)$ is a white noise vector of length N with zero mean and variance σ_q^2. The vector sequence $\mathbf{q}(n)$ is assumed to be independent of both $\mathbf{x}(n)$ and $\omega(n)$.

The output signal of the adaptive filter is given by

$$y(n) = \mathbf{x}^T(n)\mathbf{w}(n) \tag{3}$$

where $\mathbf{w}(n) = [w_0(n), w_1(n), \ldots, w_{N-1}(n)]^T$ is the weight vector of the adaptive filter of length N. The error signal is derived by subtracting $y(n)$ from $d(n)$ as

$$e(n) = d(n) - y(n). \tag{4}$$

The adaptive filter updates $\mathbf{w}(n)$ to attain an appropriate estimate of $\mathbf{h}(n)$. The weight update equation of the NLMS algorithm is expressed as [9]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu e(n)\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}. \tag{5}$$

where μ is the step-size which controls the rate of convergence. Several extensions of NLMS with time-varying step-sizes have already been introduced. For example, in the set-membership variant of NLMS (SM-NLMS) [31], the step-size is adjusted using

$$\mu(n) = \begin{cases} 1 - \dfrac{t}{|e(n)|}, & |e(n)| > t \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where t is the set-membership error bound. Additionally, in the variable step-size LMS (VSS-LMS) algorithm [20], the step-size is described as

$$\mu(n) = \left[1 - \frac{2}{e^{2\overline{\delta_e}(n)} + 1}\right] / \left[(\overline{\delta_e}(n)\delta_e(n) + \mathbf{x}^T(n)\mathbf{x}(n)\right] \tag{7}$$

where $\delta_e(n) = \delta_e(n-1) + |e(n)|$ and $\overline{\delta_e}(n) = \delta_e(n)/n$.

On the other hand, the weight update equation of the quasi-Newton (Q-Newton) algorithm is given by [9]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\Phi\mathbf{x}(n) \tag{8}$$

where Φ is an estimate of the inverse of the input correlation matrix $\mathbf{R}_x = E\{\mathbf{x}(n)\mathbf{x}^T(n)\}$. A recursive approach for Φ is derived using the matrix inversion lemma (MIL) [9], [16]. According to MIL, the time-varying Φ, denoted by $\Phi(n)$, is derived from the estimated inverse of the input correlation matrix as [14] and [32]

$$\Phi(n) = \frac{1}{1-\lambda}\left\{\Phi(n-1) - \frac{\Phi(n-1)\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\Phi(n-1)}{\mathbf{x}^{\mathrm{T}}(n)\Phi(n-1)\mathbf{x}(n) + \frac{1-\lambda}{\lambda}}\right\}. \tag{9}$$

where λ is a small positive constant that defines the convergence step of inverse calculation. In the fast quasi-Newton (FQ-Newton) algorithm [11], the inverse of the autocorrelation matrix in the weight update equation is replaced by an approximate one, yielding

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu e(n)\mathbf{P}(n)\mathbf{x}(n) \tag{10}$$

where $\mathbf{P}(n)$ is an approximate inversion of $\mathbf{R}_x$ using the inverse power spectrum of the signal corresponding to the truncated autocorrelation sequence [11].

III. THE PROPOSED VARIABLE STEP-SIZE QUASI-NEWTON ADAPTIVE FILTER

In this section, we propose a novel quasi-Newton adaptive algorithm with a variable step-size and a recursive estimate of the regularized matrix inverse. Our contributions include the adjustment of both the step-size and regularization factor based on an analysis of the convergence behaviour of the conventional quasi-Newton adaptive algorithm. The proposed algorithm, initially sets a high step-size and a low regularization factor, gradually adapting them during convergence, in a few steps. This dynamic adjustment results in the step-size decreasing and the regularization factor increasing during the convergence, with a low computational load, and ensures that the algorithm converges to a desired MSD level.

The total objective of the proposed algorithm is to quickly attain a desired low steady-state MSD level in the context of system identification. In the following, we describe the proposed method in detail.

The weight update process of the proposed quasi-Newton adaptive algorithm is similar to that of the adaptive quasi-LMS/Newton algorithm [9], [16], which relies on the inverse of the correlation matrix. However, our proposed method differs in that it employs a time-varying regularized version of the matrix inverse estimate recursively. This modification aims to enhance the initial convergence while diminishing the contribution of the matrix inverse during convergence, leading to a lower steady-state error.

We note that the regularization factor serves to prevent instability in matrix inversion, particularly when dealing with ill-conditioned data matrices that are numerically challenging to invert in practice. This situation commonly arises when handling highly correlated signals. The proposed regularization thus serves a dual purpose: firstly, to mitigate computation errors and construct a noise-robust adaptive algorithm, and secondly, to strike a balance between quasi-Newton and LMS algorithms, achieving a high initial convergence rate along with a low steady-state MSD error.

The weight vector update equation of the proposed variable step-size quasi-Newton (VSS-Q-Newton) algorithm can be expressed as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)e(n)\Phi(n)\mathbf{x}(n) \tag{11}$$

where $\mu(n)$ is a variable step-size and $\Phi(n)$ is a recursive estimate of the regularized matrix inverse such that

$$\Phi(n) = \gamma_n \widehat{\mathbf{R}}^{-1}(n) + (1-\gamma_n)\,\mathbf{I} \tag{12}$$

where $0 \leq \gamma_n \leq 1$ is a time-varying regularization parameter, and $\widehat{\mathbf{R}}^{-1}(n)$ is a time-varying estimate of $\mathbf{R}_x^{-1}$. In extreme cases, with $\gamma_n = 1$, $\Phi(n) = \widehat{\mathbf{R}}^{-1}(n)$, and the proposed algorithm acts as a time-varying quasi-Newton algorithm. Conversely, with $\gamma_n = 0$, $\Phi(n) = \mathbf{I}$, resulting in a time-varying LMS algorithm. It is noteworthy that for $0 \ll \gamma_n < 1$, $\Phi(n)$ serves as a regularized version of the matrix inverse. The derivation of the matrix $\Phi(n)$ is discussed in the following section.

A. DERIVATION OF THE REGULARIZED MATRIX INVERSE

In this section, we propose a recursive approach for $\Phi(n)$ inspired by the MIL. To achieve this, we modify (9) in a manner that eliminates the dependency of the initial convergence on the correlation of inputs. Subsequently, after the initial convergence, $\Phi(n)$ contains more regularization amounts and hence, approaches the identity matrix to achieve a low steady-state MSD. The rationale behind this approach is to mitigate estimation errors in the weight update process after the initial convergence, ultimately achieving a lower steady-state MSD error than what is possible with a conventional quasi-Newton algorithm. With this objective in mind, the computation of $\Phi(n)$ is performed using

$$\begin{cases} \Phi(n) = \dfrac{1}{1-\lambda}\Big\{\widetilde{\Phi}(n-1) - \dfrac{\widetilde{\Phi}(n-1)\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\widetilde{\Phi}(n-1)}{\mathbf{x}^{\mathrm{T}}(n)\widetilde{\Phi}(n-1)\mathbf{x}(n) + \frac{1-\lambda}{\lambda}}\Big\}, \\ \widetilde{\Phi}(n) = \gamma_n\Phi(n) + (1-\gamma_n)\,\mathbf{I}, \end{cases} \tag{13}$$

where the coefficient γ_n is determined such that $\Phi(n)$ is initially a regularized estimate of $\mathbf{R}_x^{-1}$, and through varying regularization, it gradually approaches the identity matrix. Notably, when $\gamma_n = 1$, (13) is equivalent to (9), and $\Phi(n)$ is therefore an estimate of $\mathbf{R}_x^{-1}$. Moreover, to achieve $\Phi(n \gg 1) \approxeq \mathbf{I}$, one should set $\gamma_n \to 0$ for $n \gg 1$.

This implies that the initial value of γ_n is set to one and it approaches zero for large values of n, i.e., after initial convergence. In order to consider the variation of γ_n proportional to the rate of convergence, its value can be determined as a function of the number of iterations required for the convergence, r_n, given by

$$\gamma_n = \begin{cases} 1 - (\frac{n}{r_n})^{\alpha}, & n < r_n \\ 0, & n \geq r_n \end{cases} \tag{14}$$

where r_n will be derived in the next section for two cases when $r_n = r =$ constant and when r_n is time-varying, and $\alpha \geq 1$ is a constant that indicates the decay rate for γ_n.

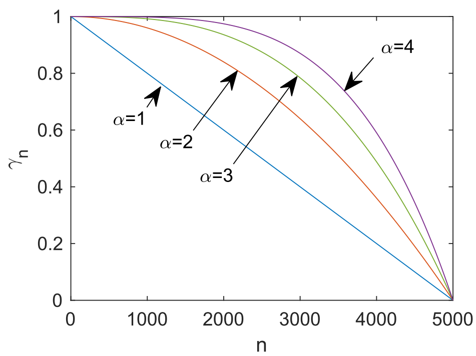

FIGURE 2. Typical variations of γ_n for various amounts of α with $r = 5000$.

Figure 2 shows typical variations of γ_n for various values of $\alpha = 1, 2, 3, 4$ when $r = 5000$ is considered. As observed, the decay rate depends on the value of α such that higher values of α result in slower variations of γ_n initially and faster variations as n approaches r. This signifies that initially, the contribution of $\widehat{\mathbf{R}}^{-1}(n)$ to $\Phi(n)$ is reduced more gradually for higher values of α. We empirically choose $\alpha = 2$ to achieve a desirable convergence performance in the simulations.

B. DERIVATION OF THE STEP-SIZE

To determine the variable step-size $\mu(n)$, we first provide an analysis of the MSD convergence behaviour of the Q-Newton algorithm for Gaussian inputs.

1) STOCHASTIC ANALYSIS OF THE Q-NEWTON CONVERGENCE BEHAVIOUR

Substituting (2) and (1) into (4) gives

$$\begin{aligned} e(n) &= \mathbf{x}^T(n)\mathbf{h}(n) - \mathbf{x}^T(n)\mathbf{w}(n) + \omega(n) \\ &= -\mathbf{x}^T(n)\mathbf{v}(n) + \omega(n) \end{aligned} \tag{15}$$

where $\mathbf{v}(n) = \mathbf{w}(n) - \mathbf{h}(n)$. Subtracting $\mathbf{h}(n)$ from both sides of (8) and using (2) and (15) yields

$$\begin{aligned} \mathbf{v}(n+1) &= \mathbf{v}(n) + \mu e(n)\Phi\mathbf{x}(n) - \mathbf{q}(n) \\ &= \mathbf{v}(n) + \mu\big(-\mathbf{x}^T(n)\mathbf{v}(n) + \omega(n)\big)\Phi\mathbf{x}(n) - \mathbf{q}(n) \\ &= \mathbf{v}(n) - \mu\Phi\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n) \\ &\quad + \mu\omega(n)\Phi\mathbf{x}(n) - \mathbf{q}(n) \end{aligned} \tag{16}$$

As a result,

$$\mathbf{v}(n+1) = \Big[I - \mu\Phi\mathbf{x}(n)\mathbf{x}^T(n)\Big]\mathbf{v}(n) + \mu\omega(n)\Phi\mathbf{x}(n) - \mathbf{q}(n). \tag{17}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. Assuming that $\mathbf{x}(n)$ is statistically independent from $\mathbf{v}(n)$, post-multiplying (17) by its transpose and taking its average results in

$$\begin{aligned} &\mathrm{E}\left\{\mathbf{v}(n+1)\mathbf{v}^T(n+1)\right\} \\ &\quad = \mathrm{E}\left\{\mathbf{v}(n)\mathbf{v}^T(n)\right\} \\ &\quad + \mu^2\Phi\mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\right\}\Phi \end{aligned}$$

$$- \mu \mathrm{E}\left\{\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\right\}\Phi$$
$$- \mu\Phi \mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\right\}$$
$$+ \mu^2 \mathrm{E}\left\{\omega^2(n)\Phi\mathbf{x}(n)\mathbf{x}^T(n)\Phi\right\} + \sigma_q^2\mathbf{I} \quad (18)$$

We note that

$$\mathrm{E}\left\{\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\right\} = 2R_xK(n)R_x + \mathrm{Tr}\left[R_xK(n)\right]R_x \quad (19)$$

where $K(n) = \mathrm{E}\left\{\mathbf{v}(n)\mathbf{v}^T(n)\right\}$. The above equality holds for zero-mean Gaussian inputs $x(n)$ [33] (Eq. (10.4.26)). Consequently,

$$\begin{aligned} K(n+1) = K(n) &+ \mu^2\Phi\left\{2R_xK(n)R_x + \mathrm{Tr}\left[R_xK(n)\right]R_x\right\}\Phi \\ &- \mu\left[K(n)R_x(n)\Phi + \Phi R_x(n)K(n)\right] \\ &+ \mu^2\sigma_\omega^2\Phi R_x(n)\Phi + \sigma_q^2\mathbf{I} \end{aligned} \quad (20)$$

Defining $\xi(n) = E\left[\mathbf{v}^T(n)\mathbf{v}(n)\right] = \mathrm{Tr}\left[K(n)\right]$ as the MSD criterion, we derive

$$\begin{aligned} \xi(n+1) = \xi(n) &+ \mu^2\{2\mathrm{Tr}\left[\Phi R_xK_{vv}(n)R_x\Phi\right] \\ &+ \mathrm{Tr}\left[R_xK_{vv}(n)\right]\mathrm{Tr}\left[\Phi R_x\Phi\right]\} \\ &- 2\mu\mathrm{Tr}\left[\Phi R_xK_{vv}(n)\right] + \mu^2\sigma_\omega^2\mathrm{Tr}\left[\Phi R_x\Phi\right] + \sigma_q^2N \end{aligned} \quad (21)$$

To proceed further, after some manipulations, we can rewrite (21) as

$$\xi(n+1) = a\xi(n) + b \quad (22)$$

where

$$a \approx 1 + \mu^2(2\widetilde{\sigma}_x^2 + N\sigma_x^2\widehat{\sigma}_x^2) - 2\mu\breve{\sigma}_x^2 \quad (23)$$

is the decay rate and $b \approx \mu^2N\sigma_\omega^2\widehat{\sigma}_x^2 + \sigma_q^2N$ in which, $\sigma_x^2 = \mathrm{Tr}[R_x]/N$, $\widetilde{\sigma}_x^2 = \mathrm{Tr}[\mathbf{R}_x\Phi^2R_x]/N$, $\widehat{\sigma}_x^2 = \mathrm{Tr}[\Phi^2R_x]/N$, and $\breve{\sigma}_x^2 = \mathrm{Tr}[\Phi R_x]/N$. Solving the recursive equation (22), we achieve

$$\xi(n) = a^n\xi(0) + b\sum_{k=0}^{n-1}a^k, \quad n \geq 1 \quad (24)$$

Without loss of generality, assuming $\mathbf{h}^T(n)\mathbf{h}(n) = 1$ and the initial value of the weight vector $\mathbf{w}(0) = \mathbf{0}$, we obtain $\xi(0) = 1$, and hence,

$$\xi(n) = a^n + b\sum_{k=0}^{n-1}a^k, \quad n \geq 1 \quad (25)$$

For the convergence of Q-Newton, from (25), we should have $0 < a < 1$. As a result,

$$\mu(2\widetilde{\sigma}_x^2 + N\sigma_x^2\widehat{\sigma}_x^2) - 2\breve{\sigma}_x^2 < 0 \quad (26)$$

and we satisfy the condition for convergence of (25), establishing upper and lower bounds for the step-size as

$$0 < \mu < \frac{2\breve{\sigma}_x^2}{2\widetilde{\sigma}_x^2 + N\sigma_x^2\widehat{\sigma}_x^2} \quad (27)$$

Solving the recursive equation (25) for steady-state, ξ_{ss}, we obtain

$$\xi_{ss} = \lim_{n\to\infty}\xi(n) = \frac{b}{1-a} \approx \frac{\mu^2N\sigma_\omega^2\widehat{\sigma}_x^2 + \sigma_q^2N}{2\mu\breve{\sigma}_x^2 - \mu^2(2\widetilde{\sigma}_x^2 + N\sigma_x^2\widehat{\sigma}_x^2)} \quad (28)$$

from which we derive

$$\mu_{ss} = \frac{\xi_{ss} + \sqrt{\xi_{ss}^2 - C\sigma_q^2N/\breve{\sigma}_x^2}}{C\breve{\sigma}_x^2} \quad (29)$$

where μ_{ss} denotes the step-size μ that the MSD level $\xi(n)$ approaches ξ_{ss}, and

$$C = \frac{1}{\breve{\sigma}_x^2}\left\{N\sigma_\omega^2\widehat{\sigma}_x^2 + \xi_{ss}\left(2\widetilde{\sigma}_x^2 + N\sigma_x^2\widehat{\sigma}_x^2\right)\right\}. \quad (30)$$

As already defined, r is the number of iterations required for the adaptive algorithm to reach its steady state, which can be computed using

$$r = \frac{\log_{10}\left[\frac{\xi_{ss}}{\xi_{init}}\right]}{\log_{10}\left[a\right]}, \quad (31)$$

where ξ_{init} is the initial value of $\xi(n)$ at $n = 0$. Assuming without loss of generality, $\mathbf{w}(0) = 0$ and $\mathbf{h}^T(n)\mathbf{h}(n) = 1$, then $\xi_{init} = 1$.

2) GENERALIZATION OF THE ANALYSIS TO A Q-NEWTON ALGORITHM WITH TIME-VARYING MATRIX

In the case where Φ is *time-varying*, C and as a result, μ_{ss}, a, and r are time-varying as well. Denoting the corresponding time-varying parameters by $\Phi(n)$, C_n, $\mu_{ss}(n)$, $a(n)$, and r_n, respectively, we modify (29) as

$$\mu_{ss}(n) = \frac{\xi_{ss} + \sqrt{\xi_{ss}^2 + C(n)\sigma_q^2N}}{C(n)} \quad (32)$$

where

$$C(n) = \sigma_\omega^2\mathrm{Tr}\left[\Phi(n)\right] + \xi_{ss}\left(2 + \frac{\mathrm{Tr}\left[R_x\right]\mathrm{Tr}\left[\Phi(n)\right]}{N}\right). \quad (33)$$

In addition, the decay rate can be modified as

$$a(n) \approx 1 + \mu^2(n)(2\widetilde{\sigma}_x^2(n) + N\sigma_x^2\widehat{\sigma}_x^2(n)) - 2\mu(n)\breve{\sigma}_x^2(n) \quad (34)$$

where

$$\widetilde{\sigma}_x^2(n) = \mathrm{Tr}[\mathbf{R}_x\Phi^2(n)R_x]/N,$$
$$\widehat{\sigma}_x^2(n) = \mathrm{Tr}[\Phi^2(n)R_x]/N,$$
$$\breve{\sigma}_x^2(n) = \mathrm{Tr}[\Phi(n)R_x]/N.$$

Moreover, r_n is defined as the number of iterations required for the adaptive algorithm to reach its steady state with a time-varying decay rate $a(n)$, which can be computed using

$$r_n = \frac{\log_{10}\left[\frac{\xi_{ss}}{\xi_{init}}\right]}{\log_{10}\left[a(n)\right]}. \quad (35)$$

As can be seen, r_n is updated at each iteration n until $\Phi(n)$ reaches its steady state.

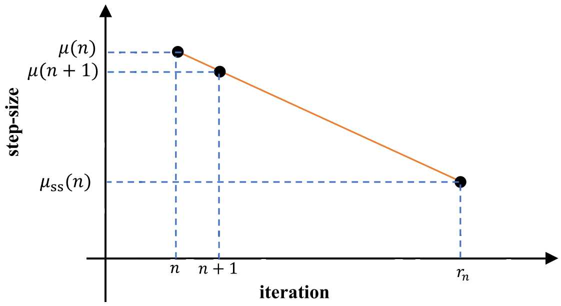

FIGURE 3. Derivation of $\mu(n+1)$ versus $\mu(n)$ using the line equation.

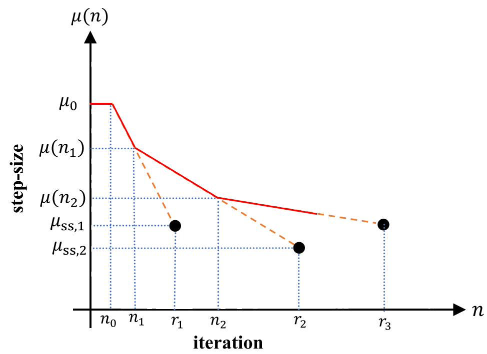

FIGURE 4. Variation of $\mu(n)$ in steps.

3) DERIVATION OF THE VARIABLE STEP-SIZE

According to the previous derivation, the step-size μ can be considered as a time-varying quantity that initiates from an initial value μ_0 and follows the variation of $\mu_{ss}(n)$. Consequently, we introduce a recursive approach for $\mu(n)$ expressed as

$$\mu(n+1) = \begin{cases} \mu_0, & n = 1 \\ \dfrac{r_n - n - 1}{r_n - n}\mu(n) + \dfrac{1}{r_n - n}\mu_{ss}(n), & 1 < n < r_n \\ \mu_{ss}(n), & n \geq r_n \end{cases} \tag{36}$$

Figure 3 visually illustrates the computation of $\mu(n+1)$ at iteration $n+1$, utilizing the line equation constructed from the known values of $\mu(n)$, r_n, and $\mu_{ss}(n)$ for $1 < n < r_n$.

As can be seen from (36), once the desired MSD level is attained at $n = r_n$, the value of $\mu(n)$ becomes $\mu_{ss}(n)$, which can be constant if $\Phi(n)$ achieves its steady-state value. It is worth noting that, since $a(n)$ in (34) is a function of $\mu(n)$, r_n in (35) also becomes a function of $\mu(n)$. In essence, r_1 is initially computed using μ_0, and then r_n and $\mu(n)$ subsequently update each other in a loop.

4) DERIVATION OF A LOW-COMPLEXITY VARIABLE STEP-SIZE

It is notable that the computations of $\mu_{ss}(n)$ in (32) and r_n in (35) in each iteration require a substantial computational load. The total number of multiplications, additions, and divisions to compute $\mu_{ss}(n)$ and r_n per iteration are $1.5N^3 + 14$, $1.5N^2(N-1) + 4N + 3$, and 8, respectively.

To address this, we can establish a low-complexity version of the step-size that starts from the initial value and varies based on the updated values of $\mu_{ss}(n)$ and r_n in a few steps. Assuming n_i, $i = 1, \ldots, K$ as the values of n at K steps, we consider r_i and $\mu_{ss,i}$, respectively, as the values of r_n and $\mu_{ss}(n)$ at n_i. Additionally, at the final step, $n_K = r_K$. In general, a higher value of K provides a faster convergence to ξ_{ss} at the cost of increased computational load. To this end, we adjust the time-varying step-size $\mu(n)$ as follows

$$\mu(n) = \begin{cases} \mu_0, & n < n_0 \\ \dfrac{r_{i+1} - n}{r_{i+1} - n_i}\mu(n_i) + \dfrac{n - n_i}{r_{i+1} - n_i}\mu_{ss,i+1}, & n_i \leq n < n_{i+1}, \\ \mu_{ss,K}, & n \geq n_K \end{cases} \tag{37}$$

where $i = 1, \ldots, K-1$. Figure 4 illustrates the variation of $\mu(n)$ (with the red line) in steps. As observed, $\mu(n)$ is updated with a new decay rate in each step. As highlighted in (37), after achieving the desired MSD level at $n = n_K$, the value of $\mu(n)$ becomes $\mu_{ss,K}$, which is constant with respect to n.

TABLE 1. The step-size reset algorithm [34].

Initialization	$Z_{\text{old}} = 0$, $\mathbf{M} = \text{diag}(1, \ldots, 1, 0, \ldots, 0)$
At each iteration n	
$\mathbf{u} = \text{sort}\left[\frac{\lvert e(n)\rvert}{\lVert\mathbf{x}(n)\rVert}, \ldots, \frac{\lvert e(n-V_T+1)\rvert}{\lVert\mathbf{x}(n-V_T+1)\rVert}\right]$,	
if mod $(n, V_T) = 0$,	$Z_{\text{new}} = \frac{\mathbf{u}^T\mathbf{M}\mathbf{u}}{V_T - V_D}$.
if $\frac{Z_{\text{new}} - Z_{\text{old}}}{\mu(n-1)} > \epsilon$,	$\mu(n) = \mu_0$, $\gamma_n = 1$,
else,	$\mu(n) =$ update equation (37), $\gamma_n =$ update equation (14).
$Z_{\text{old}} = Z_{\text{new}}$	

5) RESET ALGORITHM FOR A SYSTEM SUDDEN CHANGE

In situations where there is an abrupt change in the unknown system, it is necessary to recalibrate the step-size $\mu(n)$ and regularization parameter γ_n to their initial values in order to maintain the tracking performance of the adaptive algorithm. In pursuit of this objective, we incorporate the step-size reset algorithm, proposed in [34] into our proposed algorithm. Table 1 outlines the step-size reset algorithm in accordance to our propose algorithm.

In Table 1, V_T and V_D (where $V_D < V_T$) represent positive integers. The notation mod (p, q) denotes the remainder of the division between the integers p and q. In addition, $\mathbf{M}$ is a diagonal matrix with $V_T - V_D$ ones in the diagonal. The operator sort(.) denotes the ascending sort operation, and ϵ represents a threshold value. The step-size reset algorithm involves $V_T - V_D$ multiplications, $V_T - V_D + 1$ additions, and $V_T + 2$ divisions.

C. COMPLEXITY ANALYSIS

In this section, we evaluate the computational complexity of the proposed method and compare it with some other algorithms. To ensure a fair comparison, we specifically focus on second-order algorithms which use matrix inverse. Furthermore, we leverage the symmetric property of the correlation matrix in our analysis.

The MIL expressed in (9) requires $1.5N^2 + N + 1$ multiplications, $1.5N^2$ additions, and 3 divisions. In comparison, the regularized MIL, as defined in (13) and (14), requires $2N^2 + N + 3$ multiplications, $1.5N^2 + N + 2$ additions, and 4 divisions.

The weight update equation for Q-Newton and FQ-Newton in (8) and (10) involves $N^2 + N + 1$ multiplications and N^2 additions. In contrast, for VSS-Q-Newton, the required computations associated with (11) and (37) include $N^2 + N + 3$ multiplications and $N^2 + 2$ additions.

In the APA weight update equation, considering L as the projection order, the computational complexity of the matrix inverse is of the order $O(L^3)$. Additionally, the number of multiplications, additions, and divisions in the weight update are $NL^2 + (2N + 1)L$, $(N + 1)L^2 + (2N - 1)L$, and 0, respectively.

The computations required for generating the output and error signals are identical for Q-Newton, FQ-Newton, and VSS-Q-Newton, involving N multiplications and N additions. In contrast, for APA, these computations entail N^2 multiplications and N^2 additions.

Table 2 presents a comparative analysis of computational complexity, considering the number of required multiplications/divisions and additions per iteration. It is important to highlight that the computations for r_i and $\mu_{ss,i}$ to calculate $\mu(n)$ in (37) are performed only in a very limited number of transients (K), which can be considered negligible in comparison to the iterations needed to attain the desired MSD level. The computational requirements for these operations are indicated as $O(K)$ in Table 2. Additionally, it is worth noting that the complexity of the step-size reset algorithm, as discussed in the previous section, is minimal compared to other adaptation operations and has been omitted from the table for the sake of clarity in comparison.

Consequently, in VSS-Q-Newton, the complexity is slightly elevated compared to other second-order algorithms, owing to the additional computations involved in determining $\mu(n)$ and $\Phi(n)$.

TABLE 2. Computational complexity of different second-order algorithms.

Algorithm	Multiplications/Divisions	Additions
APA	$N^2 + N(L^2 + 2L) + L + O(L^3)$	$N^2 + N(L^2 + 2L) + L(L-1) + O(L^3)$
Q-Newton	$2.5N^2 + 3N + 4$	$2.5N^2 + N$
FQ-Newton	$N^2 + 4N + 1$	$2N^2 + N$
VSS-Q-Newton	$3N^2 + 3N + 12 + O(K)$	$2.5N^2 + 2N + 4 + O(K)$

D. EFFICIENT IMPLEMENTATION FOR HIGH SYSTEM LENGTHS

In applications such as active noise control or acoustic echo cancellation, where the system length is substantial, challenges in terms of computational cost and even, the impracticality of the adaptive algorithm can arise. To address this issue, we can implement the proposed variable step-size approach along with the variable regularization utilizing a fast quasi-Newton algorithm. This modification proves beneficial when dealing with longer unknown systems, as the size of the matrix inverse becomes considerably smaller than N, making it suitable for such applications [15], [35].

Further complexity reduction can be achieved by employing a block fast LMS/Newton algorithm, facilitating a block update within the LMS part of the algorithm [36]. The resulting algorithms exhibit a significant reduction in arithmetic complexity. We consider extensions of the proposed approach for future developments in our ongoing research endeavors.

IV. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed VSS-Q-Newton algorithm and compare its convergence performance with NLMS, SM-NLMS, Q-Newton, APA, VSS-LMS algorithm [20], and F-Q-Newton [11] in the context of system identification.

For the evaluation of algorithms, we utilize the normalized MSD (NMSD) criterion, defined as

$$\text{NMSD (dB)} = 10 \log_{10} \left\{ \frac{\|\mathbf{h}(n) - \mathbf{w}(n)\|^2}{\|\mathbf{h}(n)\|^2} \right\}. \quad (38)$$

In the following simulations, an independent white Gaussian noise signal is added to the output signal $y(n)$ with SNR= 30 dB. The unknown impulse response $\mathbf{h}(n)$ is considered time-invariant with a length of $N = 64$. The adaptive weights are initialized to zero for all algorithms. The simulation results for NMSD are obtained by averaging across 20 independent trials.

In the first simulation, we consider the input signal $x(n)$ as a zero-mean white Gaussian signal with a variance of $\sigma_x^2 = 1$. For the proposed VSS-Q-Newton algorithm, $K = 10$ and $\xi_{ss} = -55$ dB are considered. The values of the step-size for NLMS, SM-NLMS, and Q-Newton algorithms are chosen as $\mu = 0.015$. Moreover, the value of the step-size for APA is chosen as $\mu = 0.003$, and its projection order is $L = 7$.

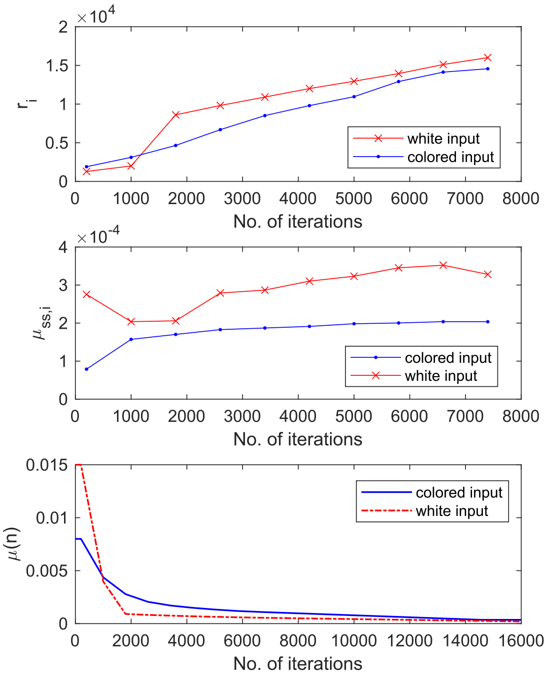

FIGURE 5. Variations of r_i, $\mu(n)$, and $\mu_{ss,i}$ for typical white and colored Gaussian input signals.

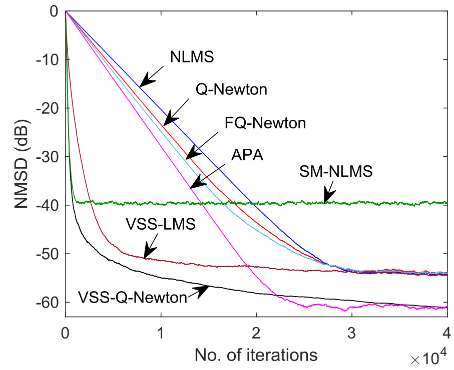

FIGURE 6. Comparison of NMSD of various algorithms for white input.

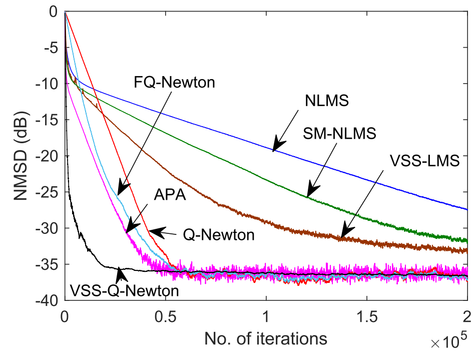

FIGURE 7. Comparison of NMSD of various algorithms for colored input.

The variation of r_i, $\mu(n)$, and $\mu_{\text{ss,i}}$ for the above white Gaussian input signal are depicted in Fig. 5. As can be seen, $r_1 = 1283$ and gradually increases to eventually reach $r_{10} = 16010$ after 10 steps. Additionally, $\mu_{\text{ss,i}}$ varies from 2×10^{-4} to 3.5×10^{-4}. The value of $\mu(n)$ reduces from $\mu_0 = 0.015$ to 0.00021.

Figure 6 compares the NMSD convergence behaviour of the proposed algorithm with its counterparts. As can be seen, the VSS-Q-Newton achieves an NMSD less than -50 dB after 3500 iterations, while APA, VSS-LMS, FQ-Newton, Q-Newton, and NLMS reach that NMSD level after 6000, 18000, 24300, 24900, and 25700 iterations, respectively. The SM-NLMS achieves a high initial convergence rate, however, its steady-state NMSD is restricted to -40 dB. In addition, the steady-state NMSD of the proposed VSS-Q-Newton achieves -61 dB, while for other algorithms, except SM-NLMS and APA, it achieves -54 dB.

It is noteworthy that the convergence performance of FQ-Newton, Q-Newton, and NLMS is similar, as the Newton-based algorithms produce a near-identity matrix for the correlation matrix inverse. As a result, the strength of the proposed VSS-Q-Newton lies in its variable step-size.

Next, we evaluate the NMSD behaviour of adaptive algorithms using a correlated input signal created by passing a white Gaussian signal through an FIR filter with coefficients [0.3574, 0.9, 0.3574] to generate a signal with a pseudo-speech spectrum [37].

The variation of r_i and $\mu(n)$ for the aforementioned colored Gaussian input signal is depicted in Fig. 5. It is observed that r_i is initially equal to $r_1 = 1900$ and gradually increases to ultimately reach $r_{10} = 14600$ after 10 steps. In addition, $\mu_{\text{ss,i}}$ varies from 0.8×10^{-4} to 2×10^{-4}. As a result, the value of $\mu(n)$ decreases from its initial value of $\mu_0 = 0.008$ to reach 0.00036 after 16000 iterations, as illustrated in Fig. 5.

Figure 7 compares the NMSD convergence behaviour of the proposed algorithm with its counterparts for the specified colored input. The values of step-size for NLMS, SM-NLMS, Q-Newton algorithms are set to $\mu = 0.015$. Additionally, for APA, the value of step-size and projection order are chosen as $\mu = 0.003$ and $L = 7$, respectively. These values are chosen such that these algorithms converge to the same steady-state NMSD as the VSS-Q-Newton. As can be seen, the VSS-Q-Newton algorithm achieves the NMSD less than -30 dB after 6000 iterations while APA, FQ-Newton, Q-Newton, VSS-LMS, SM-NLMS, and NLMS reach that NMSD level after 29000, 35000, 40000, 105000, 164000, and 230000 iterations, respectively. Consequently, the proposed algorithm exhibits a significantly higher convergence rate compared to its counterparts.

In order to evaluate the performance of the proposed VSS-Q-Newton algorithm using a speech input signal, we conducted the following simulation. The sampling frequency is considered to be 8000 Hz. The speech signal sample, as shown in Fig. 8, was taken from the TIMIT database [38], uttering: *Thus technical efficiency is achieved at the expense of actual experience*. The step-sizes and

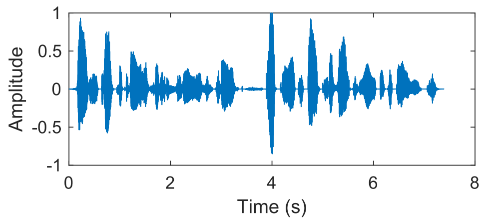

FIGURE 8. Speech input sample from the TIMIT database for the evaluation of algorithms.

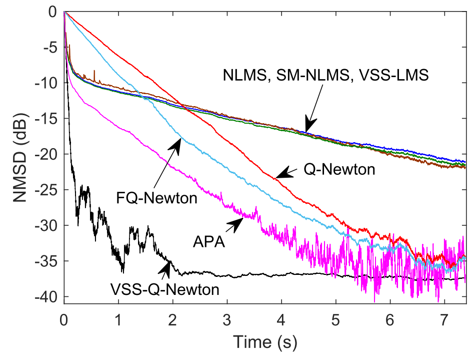

FIGURE 9. Comparison of NMSD of various algorithms for speech input.

other parameters are the same as those used in the previous simulation.

Figure 9 compares the NMSD convergence behaviour of the proposed algorithm with its counterparts for the mentioned speech signal. As can be seen, the proposed VSS-Q-Newton method reaches its steady state after 2 seconds, while it takes 5 seconds for APA, 7 seconds for FQ-Newton and Q-Newton, and 15 seconds for VSS-LMS, SM-NLMS, and NLMS algorithms. Consequently, the proposed algorithm outperforms other algorithms for non-stationary input signals.

In the reminder, we evaluate the NMSD performance of the proposed VSS-Q-Newton algorithm when an abrupt change occurs in the system impulse response. To achieve this, we consider the previous simulation configuration, wherein the constant $\mathbf{h}(n)$ is considered in the initial 30000 iterations and subsequently, an abrupt change is introduced to $\mathbf{h}(n)$, wherein it transitions to a new random WGN sequence of the same length. The input signal employed is the stationary colored signal utilized in the preceding simulation. Similar to [34], the parameters of the reset algorithm is assigned as $V_T = 2N$, $V_D = 0.75V_T$, and $\epsilon = 25$.

Figure 10 compares the NMSD performance of VSS-Q-Newton with Q-Newton, which has a fixed step-size. The curves are plotted using a single trial. At the moment of abrupt change, the NMSD error values reach around 1.5 dB and then begin to decrease to converge again. As can be seen, after the abrupt system change, the proposed algorithm reaches its steady state after 10000 iterations, while the Q-Newton reaches the steady state after 50000 iterations. As a result, the

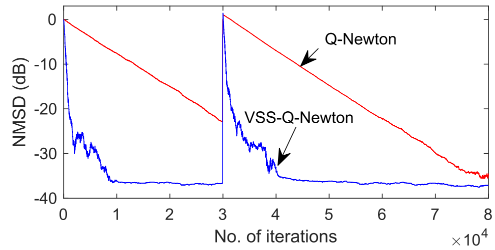

FIGURE 10. Comparison of NMSD of VSS-Q-Newton and Q-Newton when an abrupt impulse response change occurs at n = 30,000.

proposed VSS-Q-Newton algorithm outperforms its counterparts, demonstrating superior convergence performance.

V. CONCLUSION

This study introduced a novel variable step-size regularization-based quasi-Newton adaptive algorithm designed for system identification applications. The step-size adjustment procedure was derived through a statistical analysis of the mean squared deviation (MSD) convergence behavior of the quasi-Newton algorithm. Exploiting this analysis, the step-size was dynamically tuned by incorporating regularization in the recursive matrix inverse estimation and aligning it with the desired MSD. The proposed variable step-size algorithm demonstrated a high convergence rate for both white/colored stationary and non-stationary speech input signals. Simulation results verified the superior performance of the proposed algorithm in terms of MSD convergence behavior when compared to several competing algorithms.

REFERENCES

[1] M. Bekrani, A. W. H. Khong, and M. Lotfizad, "A clipping-based selective-tap adaptive filtering approach to stereophonic acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 6, pp. 1826–1836, Aug. 2011.

[2] M. Bekrani and M. Lotfizad, "Clipped LMS/RLS adaptive algorithms: Analytical evaluation and performance comparison with low-complexity counterparts," *Circuits, Syst., Signal Process.*, vol. 34, no. 5, pp. 1655–1682, May 2015.

[3] M. Bekrani, A. W. H. Khong, and M. Lotfizad, "Neural network based adaptive echo cancellation for stereophonic teleconferencing application," in *Proc. IEEE Int. Conf. Multimedia Expo*, Singapore, Jul. 2010, pp. 1172–1177.

[4] S. Chen, F. Gu, C. Liang, H. Meng, K. Wu, and Z. Zhou, "Review on active noise control technology for α-stable distribution impulsive noise," *Circuits, Syst., Signal Process.*, vol. 41, no. 2, pp. 956–993, Feb. 2022.

[5] M. Bekrani and A. W. H. Khong, "A robust MINT equalization algorithm based on near-common zero concept," in *Proc. Iranian Conf. Electr. Eng. (ICEE)*, Tehran, Iran, May 2017, pp. 1685–1690.

[6] V. Zarzoso and A. K. Nandi, "Adaptive blind source separation for virtually any source probability density function," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 477–488, Feb. 2000.

[7] E. Vlachos, A. S. Lalos, and K. Berberidis, "Stochastic gradient pursuit for adaptive equalization of sparse multipath channels," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 413–423, Sep. 2012.

[8] M. Wu, L. Tan, and N. Xiong, "Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications," *Inf. Sci.*, vol. 329, pp. 800–818, Feb. 2016.

[9] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, 2nd ed. Hoboken, NJ, USA: Wiley, 2013. [Online]. Available: https://www.wiley.com/en-dk

[10] M. Bekrani, R. Bibak, and M. Lotfizad, "Improved clipped affine projection adaptive algorithm," *IET Signal Process.*, vol. 13, no. 1, pp. 103–111, Feb. 2019.
[11] M. S. Salman, O. Kukrer, and A. Hocanin, "A fast quasi-Newton adaptive algorithm based on approximate inversion of the autocorrelation matrix," *IEEE Access*, vol. 8, pp. 47877–47887, 2020.
[12] M. Wu, N. Xiong, A. V. Vasilakos, V. C. M. Leung, and C. L. P. Chen, "RNN-K: A reinforced Newton method for consensus-based distributed optimization and control over multiagent systems," *IEEE Trans. Cybern.*, vol. 52, no. 5, pp. 4012–4026, May 2022.
[13] M. S. Ahmad, O. Kukrer, and A. Hocanin, "Recursive inverse adaptive filtering algorithm," *Digit. Signal Process.*, vol. 21, no. 4, pp. 491–496, Jul. 2011.
[14] C. G. Tsinos and P. S. R. Diniz, "Data-selective LMS-Newton and LMS-quasi-Newton algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 4848–4852.
[15] W. Y. Yuan, Y. Zhou, Z. Y. Huang, and H. Q. Liu, "A VAD-based switch fast LMS/Newton algorithm for acoustic echo cancellation," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 967–970.
[16] M. Bekrani, M. Lotfizad, and A. W. H. Khong, "An efficient quasi LMS/Newton adaptive algorithm for stereophonic acoustic echo cancellation," in *Proc. IEEE Asia–Pacific Conf. Circuits Syst.*, Dec. 2010, pp. 684–687.
[17] M. S. Salman, O. Kukrer, and A. Hocanin, "A fast implementation of quasi-Newton LMS algorithm using FFT," in *Proc. 2nd Int. Conf. Digit. Inf. Commun. Technol. Appl. (DICTAP)*, May 2012, pp. 510–513.
[18] M. A. Poletti and P. D. Teal, "A superfast Toeplitz matrix inversion method for single- and multi-channel inverse filters and its application to room equalization," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 3144–3157, 2021.
[19] A. E. Zorkun, M. A. Salas-Natera, and R. Martínez Rodriguez-Osorio, "Improved iterative inverse matrix approximation algorithm for zero forcing precoding in large antenna arrays," *IEEE Access*, vol. 10, pp. 100964–100975, 2022.
[20] L. Li and X. Zhao, "Variable step-size LMS algorithm based on hyperbolic tangent function," *Circuits, Syst., Signal Process.*, vol. 42, no. 7, pp. 4415–4431, Feb. 2023.
[21] D. G. Tiglea, R. Candido, and M. T. M. Silva, "A variable step size adaptive algorithm with simple parameter selection," *IEEE Signal Process. Lett.*, vol. 29, pp. 1774–1778, 2022.
[22] I. Gueraini, A. Benallal, and A. Tedjani, "New variable step-size fast NLMS algorithm for non-stationary systems," *Signal, Image Video Process.*, vol. 17, no. 6, pp. 3099–3107, Apr. 2023.
[23] M. Lee, T. Park, and P. Park, "Variable step-size ℓ_0-norm constraint NLMS algorithms based on novel mean square deviation analyses," *IEEE Trans. Signal Process.*, vol. 70, pp. 5926–5939, 2022.
[24] L. Shi, H. Zhao, Y. Zakharov, B. Chen, and Y. Yang, "Variable step-size widely linear complex-valued affine projection algorithm and performance analysis," *IEEE Trans. Signal Process.*, vol. 68, pp. 5940–5953, 2020.
[25] H. Zhao, W. Luo, Y. Liu, and C. Wang, "A variable step size gradient-descent TLS algorithm for efficient DOA estimation," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 12, pp. 5144–5148, Dec. 2022.
[26] H. Zhao, W. Xiang, and S. Lv, "A variable parameter LMS algorithm based on generalized maximum correntropy criterion for graph signal processing," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 9, pp. 140–151, 2023.
[27] H. Zhao, B. Liu, and P. Song, "Variable step-size affine projection maximum correntropy criterion adaptive filter with correntropy induced metric for sparse system identification," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 11, pp. 2782–2786, Nov. 2020.
[28] Y.-R. Chien, S. S.-D. Xu, and D.-Y. Ho, "Combined boosted variable step-size affine projection sign algorithm for environments with impulsive noise," *Digit. Signal Process.*, vol. 140, Aug. 2023, Art. no. 104110.
[29] Y.-R. Chien, "Variable regularization affine projection sign algorithm in impulsive noisy environment," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. 102, no. 5, pp. 725–728, May 2019.
[30] A. H. Sayed, *Adaptive Filtering*. New York, NY, USA: Wiley, 2008. [Online]. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/9780470374122
[31] S. Zhang and J. Zhang, "Set-membership NLMS algorithm with robust error bound," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 61, no. 7, pp. 536–540, Jul. 2014.
[32] S. Haykin, *Adaptive Filter Theory*, 5th ed. London, U.K.: Pearson, 2014. [Online]. Available: https://www.pearson.com/
[33] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing*. Norwood, MA, USA: Artech, 2005. [Online]. Available: https://ieeexplore.ieee.org/document/9107189
[34] J. Yoo, J. Shin, and P. Park, "Variable step-size affine projection sign algorithm," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 61, no. 4, pp. 274–278, Apr. 2014.
[35] H. I. K. Rao and B. Farhang-Boroujeny, "Fast LMS/Newton algorithms for stereophonic acoustic echo cancelation," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 2919–2930, Aug. 2009.
[36] Y. Zhou, S. C. Chan, and K. L. Ho, "A new block-exact fast LMS/Newton adaptive filtering algorithm," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 374–380, Jan. 2006.
[37] S. Attallah, "The wavelet transform-domain LMS adaptive filter with partial subband-coefficient updating," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 1, pp. 8–12, Jan. 2006.
[38] *TIMIT Acoustic-Phonetic Continuous Speech Corpus*. Accessed: Jun. 2018. [Online]. Available: https://catalog.ldc.upenn.edu/ldc93s1

MEHDI BEKRANI received the B.Sc. degree in electrical engineering from the Ferdowsi University of Mashhad, Mashhad, Iran, in 2002, and the M.Sc. and Ph.D. degrees in electrical engineering from Tarbiat Modares University, Tehran, Iran, in 2004 and 2010, respectively. From 2010 to 2012, he was a Research Fellow with Nanyang Technological University, Singapore. Since 2012, he has been an Assistant Professor with the Faculty of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran. His current research interests include adaptive signal processing and its applications.

HADI ZAYYANI (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering and in the communication field from the Sharif University of Technology, in 2000, 2003, and 2010, respectively. Since 2012, he has been an Assistant Professor with the Qom University of Technology (QUT), where he has been an Associate Professor, since 2019. He is selected as top two percent scientists from Stanford University, in 2021 and 2022. He has published around 90 articles which around half of them are in IEEE venue. He is also an Editor of *Iranian Journal of Science and Technology* and a reviewer of high impact IEEE and Elsevier journals. His research interests include statistical signal processing, sparse signal processing, adaptive filter theory, radar signal processing, graph signal processing, and distributed signal processing.