

Article

A Neural Network-Assisted Variable Step-Size NLMS Algorithm

Zhipeng Li and Yalan Guo * 

School of Mechanical and Electrical Engineering, Chengdu University of Technology, Chengdu 610059, China; lizhipeng08@cdut.edu.cn

* Correspondence: guoyalan@stu.cdut.edu.cn

Abstract

The traditional normalized least-mean-square (NLMS) algorithm faces an inherent trade-off between convergence rate and steady-state error, and its adaptability is limited in non-stationary environments. This paper proposes a neural network-assisted variable step-size NLMS algorithm (NN-VSS-NLMS). An analytically motivated reference step size is first derived under a zero-mean statistically symmetric signal assumption to characterize the desired step-size trend. Based on this reference, an eight-dimensional feature vector composed of input signal power, error energy, and related statistical descriptors is constructed to describe the instantaneous signal state, and a two-layer fully connected neural network (NN) is introduced as an auxiliary tool to provide data-driven correction to the reference step size. In addition, dynamic modulation, step-size constraints, and smoothing operations are incorporated to regulate the predicted step size and enhance its controllability under time-varying conditions. Through simulations with stationary and non-stationary inputs as well as time-invariant and time-varying systems, the proposed algorithm achieves up to a fourfold improvement in convergence rate and more than 8 dB reduction in steady-state error compared with the classical NLMS algorithm, while maintaining improved tracking ability.

Keywords: variable step-size NLMS; neural network; step-size prediction; statistical symmetry; non-stationary environment

1. Introduction

An adaptive filter plays an important role in many signal processing applications [1–6], such as echo cancellation, line enhancement, channel equalization, system identification, and time delay estimation [7]. Among adaptive algorithms, the least-mean-square (LMS) algorithm [8,9] is one of the most classical adaptive algorithms due to its simple structure and easy implementation, but its convergence rate strongly depends on the input signal power, which leads to a significant performance degradation when the input amplitude varies widely or exhibits strong correlation [10]. To mitigate this problem, the normalized least-mean-square (NLMS) algorithm [11,12] normalizes the input signal power on the basis of LMS, effectively solving the problem that LMS is sensitive to the input amplitude, and improving the adaptability of the algorithm to input signals with different amplitudes to some extent. However, the step size in NLMS is a fixed value that must be manually set, making it difficult to coordinate convergence rate and steady-state error [13]: a large step size leads to fast convergence but may cause oscillation, whereas a small step size provides better steady-state performance but results in slow convergence [14]. From a system perspective, adaptive filtering inherently involves regulating this convergence–stability



Academic Editor: Christos Volos

Received: 1 March 2026

Revised: 2 April 2026

Accepted: 3 April 2026

Published: 12 April 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

trade-off, particularly when signal statistics vary over time. In practical scenarios, signals and environments are often non-stationary, such as time-varying fading in communication channels, dynamic amplitude changes in speech signals, and abrupt variations in noise statistics, and under such conditions, the limitations of a fixed-step-size NLMS algorithm become more pronounced [15].

In recent years, researchers have proposed various variable step-size NLMS (VSS-NLMS) algorithms, enabling the step size to adapt to changes in signal statistics. These algorithms can be broadly divided into three categories. The first category includes step-size adjustment methods driven by error statistics, which directly use error power or noise power to construct step-size update rules. For example, the nonparametric VSS-NLMS (NP-VSS-NLMS) algorithm proposed by Benesty et al. [16] adjusts the step size dynamically according to the ratio between error power and noise power. The VSS-NLMS algorithm proposed by Huang et al. [17] controls the step-size update using the mean-square error (MSE) and the estimated system noise power. The new and effective nonparametric VSS-NLMS algorithm proposed by Wang and Zhang [18] obtains the step size from the square roots of the estimated error power and system noise power. The second category consists of structure-enhanced methods based on statistical inference or optimization theory, which improve step-size adaptability by introducing additional inference mechanisms or optimization frameworks. The EM-NLMS algorithm proposed by Huemmer et al. [19] obtains a variable step size by estimating the state and observation noise variances in real time through the M-step of the EM algorithm. The partial-update NLMS (PU-NLMS) algorithm proposed by Bhotto and Antoniou [20] computes a variable step size by solving a constrained minimization problem and significantly reduces the steady-state misalignment. The step-size optimization NLMS algorithm proposed by Hayashi and Sayed [21] unfolds the NLMS iterations into a differentiable computation graph and trains the step-size parameters directly using an end-to-end optimization framework. The third category adjusts the effective step size from a regularization perspective, where a time-varying regularization term is introduced to control the step-size magnitude equivalently. For example, the generalized squared-error regularized LMS (GSER-LMS) algorithm proposed by Huang and Lee [22] uses the inverse of the weighted squared error as a time-varying regularization parameter to adjust the effective step size dynamically, thereby achieving a balance between fast convergence and steady-state error.

Existing VSS-NLMS algorithms have achieved progress in these aspects, but their step-size adjustment mechanisms are still predominantly based on explicitly designed rules, which are typically derived under stationary or quasi-stationary assumptions. As a result, these methods often exhibit sluggish response or severe step-size oscillations when the signal statistics change abruptly, and their performance degrades significantly in highly non-stationary environments. Moreover, most VSS-NLMS algorithms rely on multiple manually tuned hyperparameters whose optimal values strongly depend on the signal type, signal-to-noise ratio (SNR) level, and system dynamics, thereby limiting their practical applicability and robustness. Although several analytically motivated step-size expressions have been reported, their practical performance is frequently compromised by modeling mismatch, noisy instantaneous estimates, and violated assumptions in real-world scenarios.

More recently, learning-assisted adaptive filtering methods have attracted increasing attention, where neural networks (NNs) or data-driven models are incorporated into classical adaptive filtering frameworks to enhance performance in complex and non-stationary environments [18,21,23]. For instance, NN-augmented adaptive filters have been proposed to learn nonlinear update rules or estimate key parameters such as step size and covariance matrices directly from data [24–26]. In addition, meta-learning and deep-learning-based adaptive filtering frameworks have been explored to improve generalization across vary-

ing signal conditions [27–29]. Despite these advances, most existing learning-assisted approaches either replace the analytical update rule with a fully data-driven model, which reduces interpretability, or rely on loosely coupled combinations of NNs and adaptive filters without a clearly defined analytical reference structure.

These observations indicate that a gap remains between analytically derived VSS methods and learning-assisted approaches. The former provide interpretable structures but are sensitive to modeling assumptions, while the latter improve adaptability but often lack clear analytical grounding. Effectively integrating analytical interpretability with data-driven adaptability remains an open issue.

Motivated by this gap, this work adopts a hybrid learning-assisted framework. Instead of replacing the analytical structure, an explicitly derived reference step size is first constructed based on statistical modeling. From a statistical perspective, the analytical derivation of the reference step size relies on commonly adopted assumptions such as zero-mean Gaussianity and balanced second-order moment structures, which exhibit certain symmetry properties in the signal distributions. Under stationary conditions, these properties facilitate tractable modeling and lead to structured relationships between the input and error signals. However, in practical non-stationary environments, such symmetry conditions may only hold approximately or be partially violated, resulting in biased instantaneous estimates and performance degradation. To address this issue, an NN is introduced as a data-driven correction mechanism to refine the reference step size based on instantaneous signal features and compensate for deviations from the symmetry-based assumptions used in the analytical model, rather than explicitly enforce symmetry. Specifically, the proposed NN-assisted variable step-size NLMS (NN-VSS-NLMS) algorithm constructs an eight-dimensional feature vector to describe the instantaneous features of the signals and uses a two-layer fully connected NN to learn the mapping between the signal features and the required step-size adjustment. Dynamic modulation, steady-state constraint, and smoothing mechanisms are incorporated to further improve the robustness and temporal consistency. The structure of the proposed algorithm is shown in Figure 1. To thoroughly evaluate its performance, the proposed algorithm is compared with classical algorithms under stationary and non-stationary inputs, time-invariant and time-varying systems, and various SNRs.

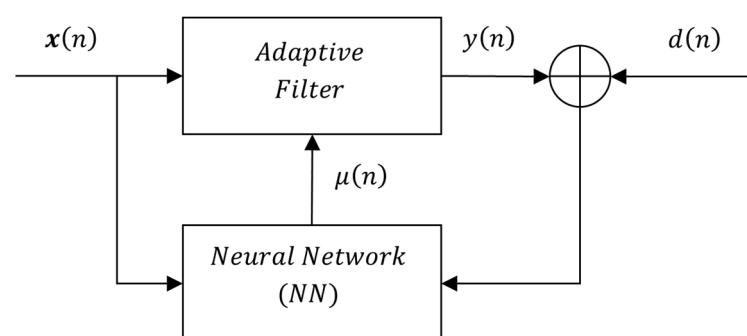


Figure 1. Block diagram of the algorithm.

Compared with existing VSS-NLMS and learning-assisted adaptive filtering methods, the main contributions of this work can be summarized as follows:

1. An analytically motivated reference step-size expression is derived based on statistical modeling, providing an interpretable characterization of the desired step-size behavior;
2. A NN is introduced as a data-driven correction mechanism to refine the analytical step-size model under non-stationary conditions, rather than replacing it;

3. A dynamic modulation and constraint framework is developed to regulate the temporal evolution of the step size and enhance its robustness;
4. The proposed method establishes a hybrid model that integrates analytical interpretability with data-driven adaptability.

2. Algorithm Design

This section first reviews the basic principles of the classical NLMS algorithm and then presents the proposed NN-VSS-NLMS algorithm in a systematic manner, including the theoretical derivation of the step size, the NN-based prediction, the dynamic modulation and steady-state constraint mechanisms, and the convergence analysis.

2.1. Review of the NLMS Algorithm

Let the filter length be L , and the input vector $x(n)$ and the coefficient vector $w(n)$ defined as

$$x(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T. \quad (1)$$

$$w(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)]^T. \quad (2)$$

The filter output is

$$y(n) = w^T(n)x(n). \quad (3)$$

The desired signal is denoted by $d(n)$ and the error signal $e(n)$ is defined as

$$e(n) = d(n) - y(n). \quad (4)$$

The classical LMS algorithm minimizes the MSE through the gradient descent method, and its coefficient update rule is given by

$$w(n+1) = w(n) + \mu e(n)x(n) \quad (5)$$

where μ is the fixed step-size parameter.

To address the instability of coefficient updates caused by variations in the input energy, the NLMS algorithm introduces a normalization term into the update rule

$$w(n+1) = w(n) + \mu \frac{e(n)x(n)}{\|x(n)\|^2 + \delta} \quad (6)$$

where $\|x(n)\|^2 = x^T(n)x(n)$ and δ is a regularization parameter introduced to avoid division by zero.

By normalizing the input power, the NLMS algorithm effectively adjusts the step size, reducing the update magnitude when the input amplitude is large and preserving a meaningful learning rate when the amplitude is small. In non-stationary environments, the input signal energy and noise power vary over time, which makes the choice of step size more critical. A relatively large step size tends to accelerate the convergence rate but increases the steady-state error, whereas a smaller step size improves steady-state performance at the cost of a slower convergence rate. Under such conditions, variable step size schemes are often adopted to enhance overall performance. The goal is to adjust μ dynamically according to the current signal or error condition so that a better compromise between convergence rate and steady-state error can be achieved [30].

2.2. Theoretical Derivation of the Reference Step Size

Conventional VSS-NLMS algorithms typically rely on empirically designed step-size rules, whose performance degrades under non-stationary conditions due to their

dependence on simplified statistical assumptions. To overcome this limitation, this paper first derives an analytically motivated reference step size, which is subsequently refined using an NN based on instantaneous signal features.

The derivation of the reference step size follows a combination of analytical transformations and commonly adopted modeling approximations in adaptive filtering. The initial steps are based on algebraic manipulations of the NLMS update and the MSE cost function, while subsequent simplifications are introduced to obtain a tractable expression under practical conditions. These simplifications are guided by standard assumptions, including local stationarity within a short-time window, slow variation in the optimal system, approximate projection of the coefficient error onto the input subspace, and weak Gaussianity of the involved random variables. Together, these considerations lead to a reference step-size expression that remains analytically interpretable while being suitable for practical implementation.

In adaptive filtering, the objective is to drive the error $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$ toward zero with a fast convergence rate while preserving stability. Accordingly, the MSE at the next time instant is defined as the cost function:

$$J(n+1) = E[e(n+1)^2] \quad (7)$$

where $E[\cdot]$ denotes the expectation operator.

The error signal at time $n+1$ is defined as

$$e(n+1) = d(n+1) - \mathbf{w}^T(n+1)\mathbf{x}(n+1). \quad (8)$$

Substituting Equation(6) into Equation (8) and rearranging yields

$$e(n+1) = d(n+1) - \mathbf{w}^T(n)\mathbf{x}(n+1) - \mu \frac{e(n)\mathbf{x}^T(n)\mathbf{x}(n+1)}{\|\mathbf{x}(n)\|^2 + \delta}. \quad (9)$$

To simplify the derivation, the two intermediate variables, the residual signal $r(n+1)$ and the correlation coefficient α_n , are defined as:

$$r(n+1) = d(n+1) - \mathbf{w}^T(n)\mathbf{x}(n+1). \quad (10)$$

$$\alpha_n = \frac{\mathbf{x}^T(n)\mathbf{x}(n+1)}{\|\mathbf{x}(n)\|^2 + \delta}. \quad (11)$$

Thus, Equation (9) can be rewritten as

$$e(n+1) = r(n+1) - \mu e(n)\alpha_n. \quad (12)$$

Substituting Equation (12) into Equation (7), taking the derivative with respect to $\mu(n)$, and setting it to zero yields:

$$\frac{\partial J(n+1)}{\partial \mu(n)} = \frac{\partial E[(r(n+1) - \mu e(n)\alpha_n)^2]}{\partial \mu(n)} = 0. \quad (13)$$

This yields the theoretical reference step size as

$$\mu(n) = \frac{E[e(n)\alpha_n r(n+1)]}{E[e(n)^2 \alpha_n^2]}. \quad (14)$$

The expectation operator is difficult to estimate in real time, especially under non-stationary environments. Therefore, the expectation is approximated using a short-term

smoothing operator $\mathbb{E}_s[\cdot] = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k (\cdot)(n - k)$, where $\lambda \in (0, 1)$. By assuming that the signal is approximately locally stationary within the smoothing window (meaning that its statistics vary more slowly than the window length), Equation (14) is modified as

$$\mu(n) = \frac{\mathbb{E}_s[e(n)\alpha_n r(n+1)]}{\mathbb{E}_s[e(n)^2 \alpha_n^2]}. \quad (15)$$

To further simplify the computation of the reference step size and improve its practical usefulness, it is necessary to expand the residual signal in a mathematical form. Assume that the desired signal satisfies

$$d(n) = \mathbf{w}_0^T(n) \mathbf{x}(n) + v(n) \quad (16)$$

where $\mathbf{w}_0(n)$ is the optimal coefficient vector that minimizes the MSE, and $v(n)$ is zero-mean additive noise uncorrelated with the input signal. Define the coefficient error vector [31] as

$$\tilde{\mathbf{w}}(n) = \mathbf{w}_0(n) - \mathbf{w}(n). \quad (17)$$

Substituting $d(n+1)$ into the residual signal in Equation (10), expanding it and approximating $\mathbf{w}_0(n+1) \approx \mathbf{w}_0(n)$ under the assumption of the slow variation in the optimal system, yields

$$\begin{aligned} r(n+1) &= d(n+1) - \mathbf{w}^T(n) \mathbf{x}(n+1) \\ &= \mathbf{w}_0^T(n+1) \mathbf{x}(n+1) + v(n+1) - \mathbf{w}_0^T(n) \mathbf{x}(n+1) + \tilde{\mathbf{w}}^T(n) \mathbf{x}(n+1) \\ &\approx \tilde{\mathbf{w}}^T(n) \mathbf{x}(n+1) + v(n+1). \end{aligned} \quad (18)$$

Since the noise $v(n+1)$ is uncorrelated with $e(n)$ and α_n , it follows that $\mathbb{E}_s[e(n)\alpha_n v(n+1)] = 0$, so the numerator of Equation (15) can be simplified as

$$\mathbb{E}_s[e(n)\alpha_n r(n+1)] = \mathbb{E}_s[e(n)\alpha_n \tilde{\mathbf{w}}^T(n) \mathbf{x}(n+1)]. \quad (19)$$

Furthermore, approximate $\tilde{\mathbf{w}}(n)$ as the projection of the input $\mathbf{x}(n)$, that is $\tilde{\mathbf{w}}(n) = \gamma(n) \mathbf{x}(n)$, which yields

$$e(n) = d(n) - y(n) = \tilde{\mathbf{w}}^T(n) \mathbf{x}(n) = \gamma(n) \|\mathbf{x}(n)\|^2. \quad (20)$$

Thus, the instantaneous estimate of $\gamma(n)$ becomes $\gamma(n) = \frac{e(n)}{\|\mathbf{x}(n)\|^2}$, which yields

$$\tilde{\mathbf{w}}(n) = \frac{e(n) \mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}. \quad (21)$$

If the input signal is assumed to be stationary, the numerator and denominator are approximately canceled, and the step size tends to a constant value (close to 1). However, in non-stationary environments, this approximation is clearly idealized, because the true projection of the coefficient error onto the input subspace is not always strictly equal to the single-sample estimate $\frac{e(n) \mathbf{x}(n)}{\|\mathbf{x}(n)\|^2}$. A more accurate estimate should come from short-term least-squares or generalized linear estimation [32], which integrates information from several past samples and therefore includes correction factors related to the covariance matrix as well as residual terms. To characterize the signal dynamics, it is necessary to further perform a subspace decomposition of $\tilde{\mathbf{w}}(n)$.

If two vectors lie in this subspace, then any linear combination of them also lies in the subspace. Considering that a single-sample estimate cannot capture covariance information [32], a smoothing correction is introduced:

$$\tilde{w}(n) = A(n) \frac{e(n)x(n)}{\|x(n)\|^2} + B(n) \frac{\zeta(n)}{\|x(n)\|^2} \quad (22)$$

where $\zeta(n)$ denotes the residual subspace component that cannot be described solely by $x(n)$ and $e(n)$. This step introduces an approximate projection of the coefficient error vector onto the input signal subspace. The subsequent decomposition further accounts for residual components that cannot be captured by single-sample estimation.

Substituting Equation (22) into Equation (10) (neglecting noise) yields

$$r(n+1) \approx A(n) \frac{e(n)x(n)x(n+1)}{\|x(n)\|^2} + B(n) \frac{\zeta^T(n)x(n+1)}{\|x(n)\|^2}. \quad (23)$$

At this point, the numerator of Equation (15) becomes

$$\begin{aligned} & \mathbb{E}_s[e(n)\alpha_n r(n+1)] \\ &= \mathbb{E}_s \left[e(n)\alpha_n A(n) \frac{e(n)x(n)x(n+1)}{\|x(n)\|^2} \right] + \mathbb{E}_s \left[e(n)\alpha_n B(n) \frac{\zeta^T(n)x(n+1)}{\|x(n)\|^2} \right]. \end{aligned} \quad (24)$$

Approximating $x(n+1)$ by $x(n)$ as $x(n+1) = px(n) + q$ (which holds within the short-time analysis window of speech or stationary inputs, and although high-transient regions may introduce errors, these will be suppressed by the subsequent smoothing process). Substituting this relation into Equation (24) reveals that the numerator and denominator contain terms of the form $\mathbb{E}_s[e^2(n)(\dots)]$, which involve fourth-order or higher-order moments such as second- and fourth-order terms in $\|x(n)\|^2$. To simplify the high-order moments involved, a weak Gaussian approximation is adopted, allowing fourth-order terms to be expressed in terms of second-order moments via Isserlis' theorem [33]. This approach relies on the symmetry of even-order moments for Gaussian variables and the key assumption that the involved variables (or their linear combinations) can be approximated as zero-mean Gaussian variables with symmetric probability density functions. For any such variables u_1, u_2, u_3, u_4 , this yields the decomposition:

$$\mathbb{E}_s[u_1 u_2 u_3 u_4] = \mathbb{E}_s[u_1 u_2] \mathbb{E}_s[u_3 u_4] + \mathbb{E}_s[u_1 u_3] \mathbb{E}_s[u_2 u_4] + \mathbb{E}_s[u_1 u_4] \mathbb{E}_s[u_2 u_3]. \quad (25)$$

In this way, the symmetry of the underlying distributions provides a tractable mechanism to reduce higher-order statistical dependencies into lower-order terms, which facilitates the derivation of an analytically interpretable step-size expression. Without this symmetry-related simplification, the analytical derivation would involve intractable higher-order moment terms and would be difficult to express in a closed form. It should be noted that this symmetry-based simplification is approximate and mainly holds under locally stationary conditions. Applying this decomposition idea to the fourth-order mixed terms in the numerator and denominator of Equation (15) yields

$$\mathbb{E}_s[e(n)\alpha_n r(n+1)] \approx c_1 \mathbb{E}_s[e^2(n) \|x(n)\|^2] + c_2 \mathbb{E}_s[e^2(n)] + \dots \quad (26)$$

$$\begin{aligned} & \mathbb{E}_s[e(n)^2 \alpha_n^2] \\ & \approx d_1 \mathbb{E}_s[e^2(n)(\|x(n)\|^2)^2] + d_2 \mathbb{E}_s[e^2(n) \|x(n)\|^2] + d_3 \mathbb{E}_s[e^2(n)] + \dots \end{aligned} \quad (27)$$

where c_i and d_i are slowly varying coefficients.

When the input energy $\|x(n)\|^2$ exhibits a significant dynamic range (that is, its order is higher than the constant terms), the dominant denominator term becomes $d_1 \mathbb{E}_s[e^2(n) \|x(n)\|^4]$ because α_n brings higher-order input energy, while the dominant numerator term becomes $c_1 \mathbb{E}_s[e^2(n) \|x(n)\|^2]$. In this case, taking the ratio of dominant terms yields

$$\mu(n) \approx \frac{c_1 \mathbb{E}_s[e(n)^2 \|x(n)\|^2]}{d_1 \mathbb{E}_s[e(n)^2 (\|x(n)\|^2)^2]} = \frac{c_1}{d_1} \cdot \frac{1}{\mathbb{E}_s[\|x(n)\|^2]}. \quad (28)$$

In this step, only the dominant terms are retained, while higher-order and less significant components are neglected, which leads to a simplified yet interpretable form.

In engineering implementations, the short-time averages are often replaced by instantaneous estimates, which yields the empirical instantaneous-driven form (scaled and regularized):

$$\mu(n) \propto \frac{e^2(n)}{\|x(n)\|^2}. \quad (29)$$

Directly using $\mu(n) = \frac{e^2(n)}{\|x(n)\|^2}$ would cause the step size to become excessively large when the input energy is low and may even lead to numerical divergence. To improve robustness, this form is modified as follows. The denominator is regularized by replacing $\|x(n)\|^2$ with $\|x(n)\|^2 + \varepsilon$ to avoid division by zero. A joint normalization term $\frac{e^2(n)}{\|x(n)\|^2 + e^2(n) + \varepsilon}$ is applied to balance the contributions of the error and the input. The square root of the result is taken to smooth the dynamic range and stabilize the step-size variation. This yields the practical reference step size:

$$\mu_{ref}(n) = \sqrt{\frac{e^2(n)}{\|x(n)\|^2 + e^2(n) + \varepsilon}}. \quad (30)$$

When the error energy $e^2(n)$ is large, and the input energy is small, $\mu_{ref}(n)$ approaches 1, which tends to promote faster convergence. When the error energy becomes stable or the input energy increases, $\mu_{ref}(n)$ is compressed, which helps maintain steady-state performance.

It should be noted that the above derivation relies on standard simplifying assumptions commonly adopted in adaptive filtering analysis. As a result, the obtained step-size expression (Equation (30)) is not claimed to be globally optimal in a strict theoretical sense. Instead, it serves as a reference form that captures the general dependence of the step size on the error energy and the input signal power.

In particular, the derived step size can be regarded as a reasonable reference choice under the considered assumptions:

1. The input signal statistics are approximately stationary or slowly time-varying within a short time window;
2. The error signal can be locally approximated as a zero-mean Gaussian process;
3. The weight perturbations between successive iterations are sufficiently small.

From a practical perspective, the analytically derived $\mu_{ref}(n)$ can therefore be interpreted as an approximate solution obtained by retaining the dominant terms in the instantaneous MSE recursion under locally stationary assumptions. It provides a structured reference that reflects the desired step-size trend while remaining sensitive to instantaneous estimation noise and modeling mismatch.

It should be emphasized that $\mu_{ref}(n)$ is not intended to be the final step size used in the adaptive update. When used alone, it reflects the desired dependency of the step size on the error energy and the input signal power but may suffer from performance degradation under strong non-stationarity and noisy instantaneous estimates. Nevertheless, $\mu_{ref}(n)$ is

bounded within the interval $(0, 1]$, which is consistent with the stability condition of the NLMS algorithm.

Outside these conditions, the analytical expression alone may not provide sufficient robustness, which motivates the use of a learning-based predictor in the proposed framework.

2.3. NN Prediction, Dynamic Modulation and Steady-State Constraint Mechanism

Theoretically, Equation (30) can adjust the step size automatically according to the magnitude of the error. In practical environments, the expected performance is not always achieved. This is mainly because the expression in Equation (30) is derived under specific signal models and instantaneous statistical conditions, and its effectiveness depends on assumptions on the input distribution and error characteristics. As a result, the formula shows limited generalization capability, and changes in signal conditions, especially under strongly non-stationary scenarios, may lead to abnormal step-size behavior.

To address this issue, this paper introduces an NN as an auxiliary tool for step-size adjustment, aiming to improve the adaptability of the analytically derived step-size under practical conditions. The NN takes an eight-dimensional feature vector as input. These features include the input signal power, error magnitude and energy, normalized error terms, and energy ratio-based quantities. They capture the instantaneous deviation, local energy, and short-term statistical behavior of the filter, and describe the filter state and signal characteristics at different convergence stages:

$$f(n) = \left[\frac{x_2(n)}{L}, |e(n)|, e^2(n), \bar{e}^2(n), \frac{\bar{x}_2(n)}{L}, \frac{e(n)}{\sqrt{x_2(n) + \varepsilon}}, \bar{\phi}_1(n), \bar{\phi}_2(n) \right] \quad (31)$$

where $\frac{x_2(n)}{L} = \frac{\|x(n)\|^2}{L}$ denotes the average input signal power, and the normalization by L ensures scale consistent with respect to the filter length. $|e(n)|$ and $e^2(n)$ characterize the error magnitude and energy; $\bar{e}^2(n)$ and $\bar{x}_2(n)$ is the smoothed estimate of $e^2(n)$ and $x_2(n)$; $\frac{e(n)}{\sqrt{x_2(n) + \varepsilon}}$ provides a normalized signed error, which reflects both the magnitude and direction of the error relative to the input; $\bar{\phi}_1(n)$ and $\bar{\phi}_2(n)$ are the smoothed estimates of $\phi_1(n) = \frac{e^2(n)}{x_2(n) + \varepsilon}$ and $\phi_2(n) = \frac{e^2(n)}{e^2(n) + x_2(n) + \varepsilon}$, capturing the relative contribution of the error energy with respect to the input and total energy, respectively; ε is a small constant preventing division by zero.

This work employs a two-layer fully connected NN. The input layer size matches the feature dimension; the two hidden layers contain 64 and 24 neurons, respectively. ReLU activation [34] is used to enhance nonlinear mapping capability, and light dropout is applied to reduce overfitting. The output layer uses a Sigmoid function to constrain the predicted step size to the interval $(0, 1)$, which helps constrain the output within a bounded range. During the training phase, the analytically derived reference step size $\mu_{ref}(n)$ is directly used as the supervision signal. In this sense, $\mu_{ref}(n)$ can be viewed as a practical surrogate target, while the NN aims to learn a smoother and more robust mapping from signal features to step-size adjustment under real signal fluctuations. However, it should be emphasized that the $\mu_{ref}(n)$ is computed from instantaneous signal observations under practical conditions, and therefore inherently contains statistical fluctuations and modeling inaccuracies. From this perspective, the NN is not intended to replicate the analytical expression itself, but to learn the underlying nonlinear relationship between signal features and the desired step-size behavior in practical scenarios. The discrepancies between the analytical model and its practical realization arise from various simplifying assumptions, such as local stationarity, projection approximations, and Gaussian moment approximations, which are difficult to model explicitly in closed form. As a result, directly refining the reference step size using analytical methods would require complex high-

order statistical modeling, which is generally difficult to implement in real-time adaptive filtering. The NN provides a practical alternative by learning this relationship from data through its nonlinear mapping capability. In this sense, the NN is expected to capture the discrepancy between the idealized analytical model and its practical realization, which can be interpreted as a form of data-driven correction to the analytically motivated step-size structure. In the inference stage, the step size predicted by the NN is further processed through dynamic modulation, constraint, and smoothing operations to enhance stability and temporal consistency. These mechanisms operate on top of the NN output and are designed to regulate its temporal behavior, whereas the NN itself focuses on capturing the instantaneous nonlinear mapping between signal features and step-size adjustment. Based on this learning objective, the practical training process of the network is described as follows. Training samples with different SNRs are constructed, and the network parameters are updated by minimizing the loss between the NN-predicted step-size $\mu_{pred}(n)$ and $\mu_{ref}(n)$, enabling the NN to learn the step-size adjustment patterns under different noise conditions and different convergence stages. During the inference phase, the network generates the predicted step size $\mu_{pred}(n)$ based on the instantaneous signal features. This data-driven correction refines the reference step size and forms a nonlinear step-size adjustment mechanism that is expected to improve robustness and generalization under varying conditions.

The network is trained using the Adam optimizer with an initial learning rate of 10^{-3} . The mini-batch size is set to 64 and the training process runs for 40 epochs. To improve generalization ability, a dropout rate of 0.1 is applied in the hidden layers and an L_2 regularization term with coefficient 10^{-4} is adopted. Early stopping is employed based on the validation loss with a patience of six validation checks to prevent overfitting.

However, the NN generates the step size mainly based on instantaneous signal features, focusing on the mapping at the current time instant rather than explicitly reflecting the temporal evolution of the step size. In rapidly time-varying scenarios, the predicted step size may exhibit instantaneous fluctuations. Without further regulation, this behavior may lead to excessively aggressive filter updates or degraded steady-state performance. Therefore, on top of the NN-predicted step size, this work introduces a dynamic modulation and stabilization mechanism to enhance the robustness of the algorithm.

Specifically, an exponential-decay amplification factor is employed as a dynamic modulation term to accelerate early-stage convergence and gradually restore stability. The amplification factor is computed as follows:

$$\text{boostFactor}(n) = \gamma \cdot e^{-\frac{\beta \cdot n}{N}} \quad (32)$$

where γ is the initial amplification coefficient, β is a hyperparameter controlling the decay rate, n is the current iteration index, and N is the total number of iterations. At this time, the step size update is defined as

$$\mu_{\text{boost}}(n) = \mu_{\text{pred}}(n) \cdot (1 + \text{boostFactor}(n)) \quad (33)$$

where $\mu_{\text{pred}}(n)$ denotes the step-size predicted by the NN.

The step-size is limited to the interval $[\mu_{\min}, \mu_{\max}]$ to avoid instability caused by excessively large or small values, which leads to

$$\mu_{\text{constr}}(n) = \min(\max(\mu_{\text{boost}}(n), \mu_{\min}), \mu_{\max}) \quad (34)$$

where $\mu_{\max}$ is the maximum step-size and $\mu_{\min}$ is the minimum step-size.

Furthermore, the parameter $maxStepRatio$ is used to restrict the change in the step size with respect to its previous value, yielding

$$\mu_{ratio}(n) = \min(\mu_{constr}(n), \mu_{prev} \cdot maxStepRatio) \quad (35)$$

$$\mu_{ratio}(n) = \max\left(\mu_{ratio}(n), \frac{\mu_{prev}}{maxStepRatio}\right) \quad (36)$$

where μ_{prev} denotes the step-size at the previous time instant.

A smoothing operation is then applied to obtain the final step-size used in the filter update:

$$\mu_{final}(n) = \alpha \mu_{ratio}(n) + (1 - \alpha) \mu_{prev} \quad (37)$$

where $\alpha \in (0, 1)$ is the smoothing coefficient.

Through these operations, the step size predicted by the NN is converted into a step size used in the filter update. The flowchart of the NN-VSS-NLMS algorithm is shown in Figure 2.

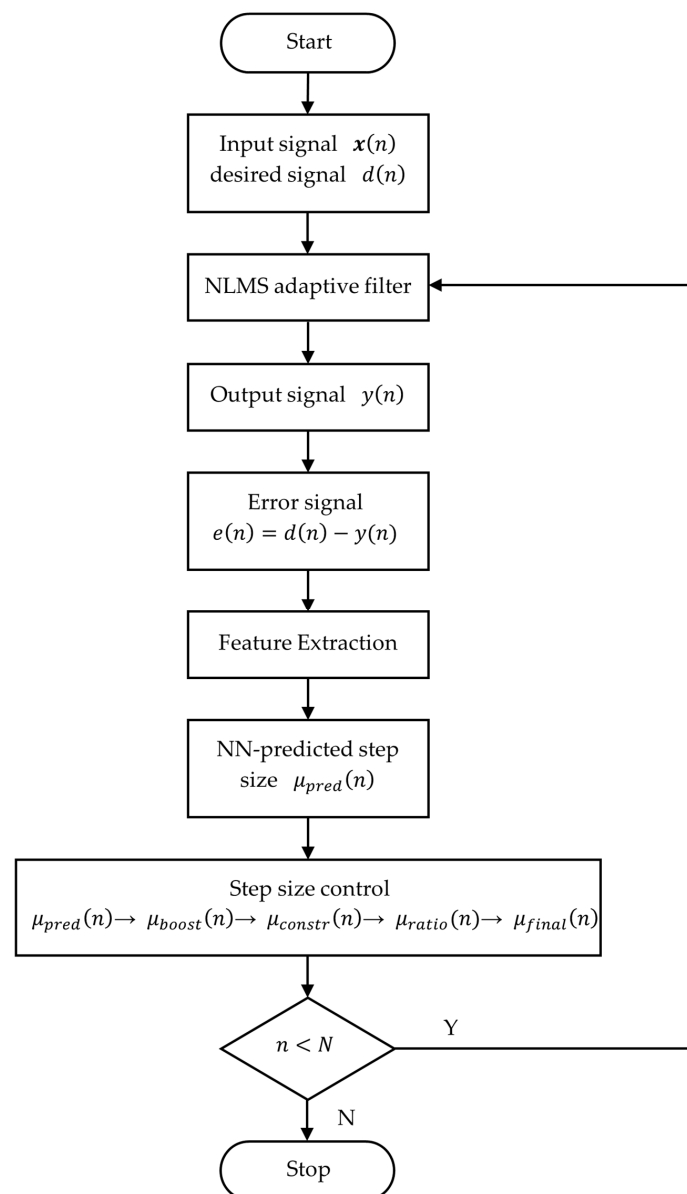


Figure 2. Flowchart of the NN-VSS-NLMS algorithm.

2.4. Convergence Analysis

For the conventional NLMS algorithm, mean-square stability is guaranteed when the step size satisfies $0 < \mu < 2$.

In the proposed method, the reference step size is constructed within a predefined interval $[\mu_{min}, \mu_{max}]$, where $0 < \mu_{min} < \mu_{max} < 2$. Through subsequent modulation and constraint operations, the step size used in the coefficient update remains within this interval.

Therefore, the effective step size involved in the adaptive update is consistent with the stability range typically associated with the NLMS algorithm, which supports stable operation under general conditions.

To further analyze the behavior of the proposed NN-VSS-NLMS algorithm under non-stationary conditions, the evolution of the coefficient-error vector $\tilde{\mathbf{w}}(n)$ is considered. The update equation is given by

$$\tilde{\mathbf{w}}(n+1) = \tilde{\mathbf{w}}(n) - \mu(n) \frac{e(n)\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2 + \delta}. \quad (38)$$

Taking the expectation of the norm squared of Equation (38) yields

$$\begin{aligned} & E[\|\tilde{\mathbf{w}}(n+1)\|^2] \\ &= E[\|\tilde{\mathbf{w}}(n)\|^2] - 2E\left[\mu(n) \frac{e(n)\tilde{\mathbf{w}}^T(n)\mathbf{x}(n)}{\|\mathbf{x}(n)\|^2 + \delta}\right] + E\left[\mu^2(n) \frac{e^2(n)\|\mathbf{x}(n)\|^2}{(\|\mathbf{x}(n)\|^2 + \delta)^2}\right]. \end{aligned} \quad (39)$$

Since error signal $e(n) = \tilde{\mathbf{w}}^T(n)\mathbf{x}(n) + v(n)$, where the noise $v(n)$ is uncorrelated with both $\tilde{\mathbf{w}}(n)$ and $\mathbf{x}(n)$, it follows that

$$E[e(n)\tilde{\mathbf{w}}^T(n)\mathbf{x}(n)] = E[\|\tilde{\mathbf{w}}(n)\mathbf{x}(n)\|^2] \geq 0. \quad (40)$$

Under commonly adopted assumptions in adaptive filtering, the first-order term in Equation (39) generally dominates, while the higher-order term associated with $\mu^2(n)$ has a relatively smaller influence. In this case, a strict theoretical proof of monotonic convergence is difficult to establish, while the above analysis still reflects the overall convergence behavior.

However, it should be noted that the step size $\mu(n)$ is jointly determined by the analytical model and the data-driven prediction, and thus varies with the instantaneous signal conditions. In this case, a strict theoretical proof of monotonic convergence is difficult to establish, while the above analysis still reflects the overall convergence behavior.

2.5. Computational Complexity Analysis

This subsection analyzes the computational complexity of the considered algorithms in terms of the number of operations per iteration. Specifically, the total computational cost is divided into two components: operations proportional to the filter length L and constant operations independent of L .

For the NLMS algorithm, the dominant computations per iteration include the output calculation, input energy estimation, and weight update. These operations result in approximately $3L + 1$ arithmetic operations, leading to an overall computational complexity of $O(L)$.

The NP-VSS-NLMS and VSS-NLMS algorithms introduce additional recursive estimations and step-size adaptation mechanisms. These operations incur a limited number of extra scalar computations, while preserving the overall $O(L)$ complexity.

For the proposed NN-VSS-NLMS algorithm, the adaptive filtering update follows the same structure as NLMS, and thus retains the $O(L)$ complexity with respect to L . The additional computational cost arises from feature extraction, NN inference, and step-size control. The feature extraction process involves only a small number of arithmetic operations (approximately 14 operations) and remains independent of L . The step-size control includes dynamic boosting, bounding constraints, and temporal smoothing. These operations mainly consist of multiplications, comparisons, and exponential evaluations, introducing only a small constant computational overhead (approximately 10 operations). The computational cost of the NN depends on its architecture. For the adopted 8–64–24–1 fully connected structure, the forward-pass computation requires approximately $8 \times 64 + 64 \times 24 + 24 \times 1 \approx 2100$ arithmetic operations per iteration. Therefore, the additional complexity of the proposed algorithm is primarily dominated by the NN inference, while the feature extraction and step-size control introduce only minor constant overhead. Consequently, the overall computational complexity remains $O(L)$.

A summary of the computational complexity of the considered algorithms is provided in Table 1.

Table 1. Computational complexity comparison of the considered algorithms.

Algorithm	Order	Operations Dependent on L	Fixed Overhead
NLMS	$O(L)$	$3L + 1$	1
NP-VSS-NLMS	$O(L)$	$3L + 1$	≈ 10
VSS-NLMS	$O(L)$	$4L + 1$	≈ 20
NN-VSS-NLMS	$O(L)$	$3L + 1$	≈ 2100

3. Algorithm Simulation and Analysis

To validate the performance of the proposed NN-VSS-NLMS algorithm, this paper conducts comprehensive simulations on the MATLAB R2024a platform under both stationary and non-stationary input conditions.

For a meaningful and balanced comparison, several representative algorithms are selected as benchmarks. The NP-VSS-NLMS algorithm adjusts the step size according to the relationship between error power and noise power, resulting in a direct and responsive adaptation mechanism. The VSS-NLMS algorithm proposed by Huang et al. employs recursive estimation of signal and noise statistics, leading to a more structured step-size adaptation process. The NLMS algorithm is included as a baseline due to its simplicity and well-established performance characteristics. These algorithms represent different approaches to step-size adaptation and together provide a comprehensive basis for performance evaluation. In contrast, the proposed NN-VSS-NLMS algorithm integrates an analytically derived reference step size with NN-assisted refinement, providing a more flexible and controllable step-size adaptation mechanism.

3.1. Input Signals

Two representative signals are used for the experiments in stationary environments. The first one is white noise that follows a standard normal distribution with zero mean and unit variance [35], which is employed to evaluate the algorithm's performance under uncorrelated inputs. The second one is an AR(1) signal generated by $x(n) = ax(n-1) + v(n)$, where a is the autoregressive coefficient and $v(n)$ is white noise. This signal is used to assess the algorithm's behavior when dealing with colored inputs.

For the experiments in non-stationary environments, speech signals are selected from the LibriSpeech corpus. Specifically, a total of 1000 speech recordings from different

speakers are used for training, while 100 recordings from another set of speakers are used for testing to ensure speaker independence between the training and test sets. All recordings are resampled to 8 kHz. To ensure consistency in the adaptive filtering process, each speech signal is truncated to a fixed length of $N = 2000$ samples. This fixed-length segmentation ensures uniform input conditions and facilitates stable feature extraction and training. Speech signals inherently exhibit strong non-stationary characteristics and temporal variations, making them suitable for evaluating the algorithm's performance under realistic and time-varying conditions.

3.2. Algorithm Parameter Configuration

The parameter settings of all algorithms are summarized in Table 2. The parameters of the compared algorithms are selected according to the recommended settings reported in their original publications and are used directly in the experiments without further retuning in order to ensure a fair and unbiased comparison. The hyperparameters in the proposed algorithm are determined through a combination of theoretical analysis and empirical tuning. Specifically, the step-size bounds (μ_{min} and μ_{max}) are selected according to the stability conditions of NLMS to ensure convergence. The remaining parameters (e.g., γ , β , $maxStepRatio$, and α) are tuned through repeated experiments under different signal types (white noise, AR(1), and speech) and various SNR conditions. The final parameter configuration is chosen to achieve a favorable trade-off between convergence rate and steady-state error. In the experiments, the order of the adaptive filter is set to $L = 8$ and $L = 32$ for separate tests, and the sample length per iteration is $N = 2000$. This allows the evaluation of the proposed algorithm under both low-order and higher-order filter conditions.

Table 2. Simulation parameters of our NN-VSS-NLMS and other comparing algorithms.

Algorithm:	NLMS
Parameters:	$\mu_{NLMS} = 0.01$
Benesty's NP-VSS-NLMS [16]	
Algorithm:	$\mu = \begin{cases} 1 - \frac{\hat{\sigma}_v}{\zeta + \hat{\sigma}_e(n)}, & \text{if } \hat{\sigma}_e(n) \geq \hat{\sigma}_v \\ 0, & \text{otherwise} \end{cases}$
	$\hat{\sigma}_e^2(n) = \kappa \hat{\sigma}_e^2(n-1) + (1 - \kappa)e^2(n)$
Parameters:	$\zeta = \frac{\hat{\sigma}_v}{1000}$, $\delta = 10^{-3}$, $\kappa = 0.95$
Huang's VSS-NLMS [17]	
Algorithm:	$\mu = \begin{cases} \alpha \mu(n-1) + (1 - \alpha) \frac{\hat{\sigma}_e^2(n)}{\beta \hat{\sigma}_v^2(n)}, & \text{if } \zeta(n) < \zeta_{th} \\ 1, & \text{otherwise} \end{cases}$
Parameters:	$\alpha = 0.998$, $\beta = 30$, $\varepsilon = 0.1$, $\mu_{min} = 10^{-5}$, $\mu_{max} = 1$, $\mu(0) = 1$
Proposed NN-VSS-NLMS	
Algorithm:	$\mu_{boost}(n) = \mu_{pred}(n) \cdot (1 + \gamma \cdot e^{-\frac{\beta \cdot n}{N}})$ $\mu_{constr}(n) = \min(\max(\mu_{boost}(n), \mu_{min}), \mu_{max})$ $\mu_{ratio}(n) = \max(\min(\mu_{constr}(n), \mu_{prev} \cdot maxStepRatio), \frac{\mu_{prev}}{maxStepRatio})$ $\mu_{final}(n) = \alpha \mu_{ratio}(n) + (1 - \alpha) \mu_{prev}$
Parameters:	$\gamma = 6.8$, $\beta = 12$, $\mu_{min} = 10^{-4}$, $\mu_{max} = 1$, $maxStepRatio = 2$, $\alpha = 0.9$

During the NN training stage, the proposed algorithm uses the Adam optimizer to update the network parameters, with an initial learning rate of 0.001. The SNR of the training data is randomly selected from 0, 5, 10, 15, and 20 dB, enabling the network to learn the step-size adaptation behavior under different noise conditions. The target labels for supervised learning are generated based on the reference step size derived in the previous section. The constructed dataset is randomly shuffled and divided into training and

validation subsets. Approximately 10% of the samples are used for validation. This split ensures sufficient training data while enabling monitoring of generalization performance.

In the inference stage, all algorithm comparisons are conducted under a fixed SNR of 20 dB. The test set contains 100 speech samples from 30 different speakers to ensure diversity and independence. For each experimental setting, the reported results are obtained by averaging over 100 independent trials. Specifically, the MSE curves are first computed for each test sample and then averaged across all samples to produce the final performance curves.

3.3. Performance Metrics

This study adopts the MSE as the primary performance metric, based on which both the convergence rate and the steady-state error are evaluated. The MSE is defined as

$$MSE = 10\log_{10}(E\{|e(n)|^2\}). \quad (41)$$

The convergence rate is defined as the average decrease in the MSE per iteration (unit: dB/iteration), computed as

$$R = \frac{MSE_{start} - MSE_{stable}}{N_{conv}} \quad (42)$$

where MSE_{start} denotes the average MSE over the first 50 iterations, MSE_{stable} denotes the average MSE during the steady-state phase, and N_{conv} is the number of iterations from the beginning until the algorithm enters steady state.

The steady-state error is defined as the average value of the MSE after the MSE curve enters the steady-state region. The steady-state region is determined as follows: when the fluctuation amplitude of the MSE curve, estimated using a sliding standard deviation with a window length of 50, falls below a threshold and remains below this value for at least 200 iterations, the algorithm is considered to have entered steady state. If this condition is not satisfied throughout the entire process, the last 20% of the iterations are regarded as the steady-state interval. The threshold is set according to the type of input signal, where 0.3 is used for stationary inputs and 1.0 is used for non-stationary inputs.

To further evaluate the statistical stability of the algorithms, the standard deviation of the steady-state MSE is also considered. For each algorithm, multiple independent trials are conducted under the same experimental settings, resulting in a set of steady-state MSE values.

Specifically, let MSE_k denote the steady-state MSE obtained from the k -th trial. The mean and standard deviation are computed as

$$\mu_{MSE} = \frac{1}{K} \sum_{k=1}^K MSE_k \quad (43)$$

$$\sigma_{MSE} = \sqrt{\frac{1}{K} \sum_{k=1}^K (MSE_k - \mu_{MSE})^2} \quad (44)$$

where K is the number of trials. The MSE values are averaged in the linear domain and then converted to the dB scale for reporting.

The standard deviation reflects the variability of the algorithm performance across different trials. A smaller value indicates more stable and consistent behavior.

To evaluate the tracking ability of the algorithm under time-varying systems, the true FIR filter is multiplied by a triangular gain function (with amplitude varying from 1 to 3.5)

within the iteration interval $[600, 800]$. This constructs a time-varying system for testing the algorithm under non-stationary conditions (as shown in Figure 3).

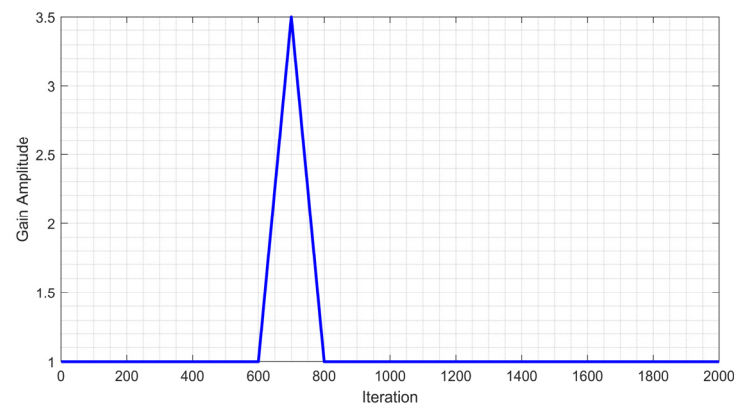


Figure 3. Gain function of the time-varying system.

3.4. Performance Analysis of Step-Size Forms

To clarify the individual roles of different step-size components in the proposed NN-VSS-NLMS algorithm, a structured ablation study is conducted with the adaptive filter order set to $L = 32$. Seven step-size variants are evaluated under identical conditions, including the reference step size $\mu_{ref}(n)$, the NN prediction $\mu_{pred}(n)$, the boosted step size $\mu_{boost}(n)$, the amplitude-constrained step size $\mu_{constr}(n)$, the ratio-constrained step size $\mu_{ratio}(n)$, the final smoothed step size $\mu_{final}(n)$, and the step size $\mu_{noNN}(n)$ obtained by applying boosting, constraint, and smoothing directly to the $\mu_{ref}(n)$ without introducing the NN. For white noise input, the MSE curves and step-size curves are shown in Figures 4 and 5, respectively. For speech input, the MSE curves and step-size curves are shown in Figures 6 and 7, respectively.

The $\mu_{ref}(n)$ provides a baseline evolution with relatively small magnitude, leading to a more conservative adaptation process. However, under non-stationary speech conditions, $\mu_{ref}(n)$ exhibits noticeable fluctuations in the steady-state phase, suggesting limited robustness to rapidly varying signal statistics.

With the introduction of NN, $\mu_{pred}(n)$ exhibits larger values, leading to a more responsive adaptation process. Compared with $\mu_{ref}(n)$, it achieves a faster reduction in MSE after the initial stage.

It should be noted that $\mu_{boost}(n)$ is introduced as an intermediate modulation signal in the step-size generation process, aiming to enhance the NN output during the early adaptation stage. Since this step-size form does not yet incorporate sufficient stability-oriented constraints for NLMS adaptation, it does not constitute an independent step-size strategy, and its corresponding MSE performance is therefore not taken as a primary evaluation metric.

After introducing constraints, $\mu_{constr}(n)$ and $\mu_{ratio}(n)$ exhibit very similar step-size trajectories, indicating that they produce comparable step-size ranges. Their corresponding MSE curves almost overlap, suggesting that these mechanisms primarily regulate the step-size behavior rather than directly reducing the error level.

Finally, the $\mu_{final}(n)$ achieves favorable performance under both stationary and non-stationary inputs, exhibiting a fast convergence rate and a low steady-state error. These results indicate that the analytically motivated reference step size alone is insufficient under non-stationary conditions, and that the learning-assisted correction and subsequent regulation are necessary to maintain stable and robust adaptation.

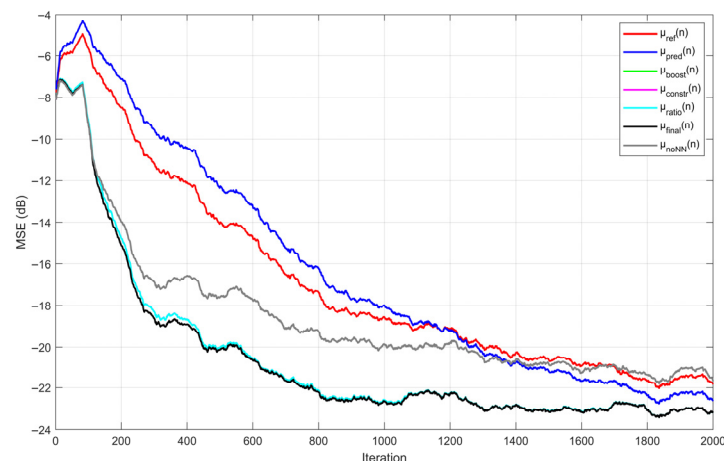


Figure 4. MSE curves of the ablation study with different step-size forms with white noise input.

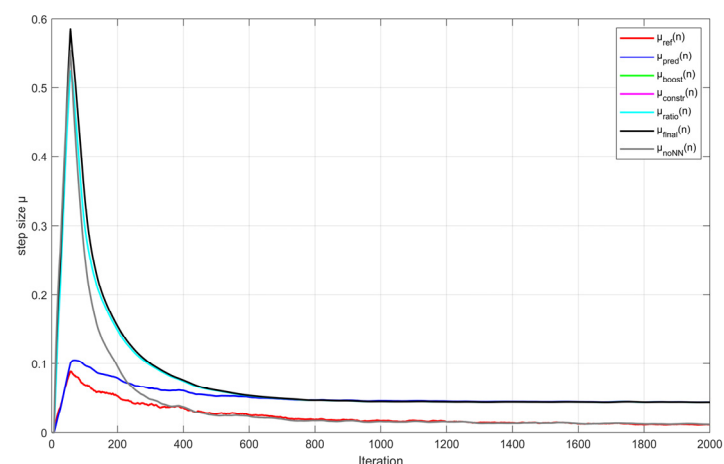


Figure 5. Step-size curves of the ablation study with different step-size forms with white noise input.

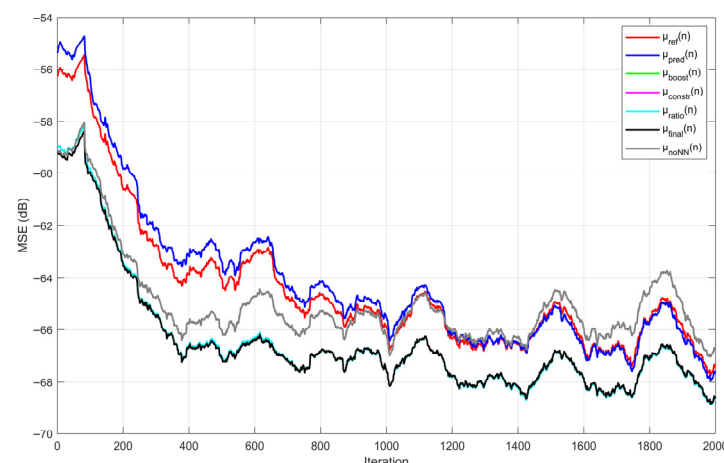


Figure 6. MSE curves of the ablation study with different step-size forms with speech input.

It is worth noting that the NN and the subsequent heuristic mechanisms play different roles in the proposed framework. The NN focuses on learning the instantaneous nonlinear mapping between signal features and step-size adjustment, while the dynamic modulation and constraint mechanisms regulate the temporal evolution and stability of the step size. It can be observed from the figures that $\mu_{final}(n)$ exhibits superior overall performance to $\mu_{noNN}(n)$. Therefore, the improvement is mainly attributed to the enhanced step-size adaptation capability introduced by the NN, which enables a more accurate characteriza-

tion of the desired step-size behavior under non-stationary conditions. The subsequent heuristic mechanisms further refine this adaptation by improving stability and preventing undesirable fluctuations.

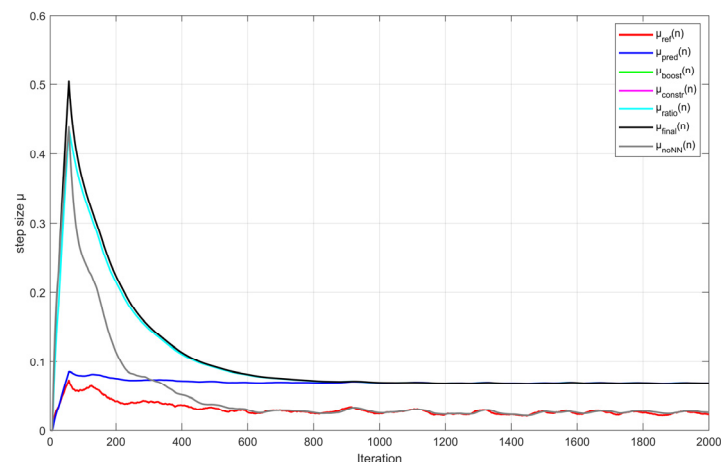


Figure 7. Step-size curves of the ablation study with different step-size forms with speech input.

3.5. Performance Analysis Under Stationary Inputs

Figures 8 and 9 presents the performance of the proposed NN-VSS-NLMS algorithm compared with NLMS, NP-VSS-NLMS, and VSS-NLMS under white-noise input, for filter orders of 8 and 32 respectively. The corresponding quantitative results are provided in Table 3.

Table 3. Performance metrics of different algorithms with white noise input in a time-invariant system at SNR = 20 dB.

Filter Order L	Metric	NLMS	NP-VSS-NLMS	VSS-NLMS	NN-VSS-NLMS
8	Convergence Rate (dB/iteration)	0.010839	0.053242	0.049901	0.054135
	Steady-state Error (dB)	−15.495	−23.655	−23.426	−23.752
	Standard Deviations (dB)	0.40742	0.41015	0.34955	0.35424
32	Convergence Rate (dB/iteration)	\	0.044393	0.042922	0.044432
	Steady-state Error (dB)	\	−22.672	−22.587	−22.779
	Standard Deviations (dB)	0.30585	0.34954	0.33223	0.33409

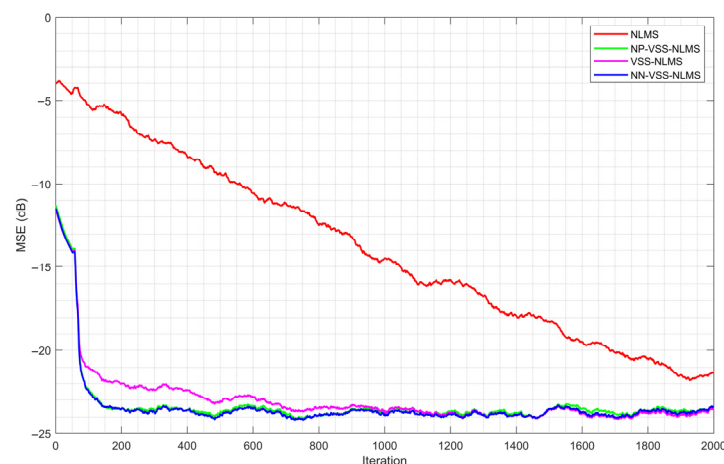


Figure 8. MSE curves of different algorithms with white noise input in a time-invariant system at SNR = 20 dB and $L = 8$.

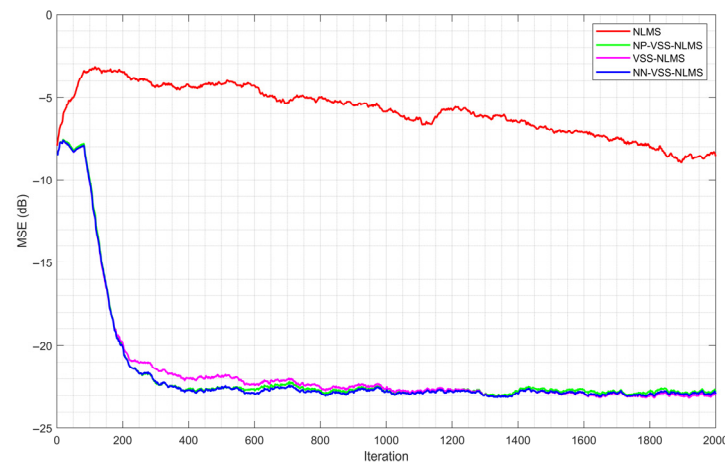


Figure 9. MSE curves of different algorithms with white noise input in a time-invariant system at SNR = 20 dB and $L = 32$.

For $L = 8$, the proposed algorithm achieves a higher convergence rate than NP-VSS-NLMS and VSS-NLMS, indicating faster adaptation during the initial stage. In the steady-state phase, it attains a lower steady-state error than NLMS and remains slightly below NP-VSS-NLMS and VSS-NLMS, reflecting a moderate improvement in steady-state accuracy. In addition, its standard deviation is close to that of VSS-NLMS and lower than that of NP-VSS-NLMS, indicating that the steady-state performance is achieved without a noticeable increase in fluctuation level.

For $L = 32$, the overall convergence rate decreases due to the increased system order, and the differences among the variable step-size algorithms become smaller. Under this condition, NN-VSS-NLMS still yields a faster convergence rate, although the margin over NP-VSS-NLMS becomes limited, while remaining slightly faster than VSS-NLMS. In the steady-state phase, it continues to produce a slightly lower steady-state error, and its standard deviation remains within a similar range to the other algorithms, indicating that the fluctuation level does not increase with filter order.

Figures 10 and 11 show the performance of the proposed NN-VSS-NLMS algorithm and the NLMS, NP-VSS-NLMS, and VSS-NLMS algorithms under AR(1) inputs with four correlation levels ($a = 0.3, 0.5, 0.7, 0.9$) for filter orders of 8 and 32, respectively. The corresponding quantitative results are provided in Table 4.

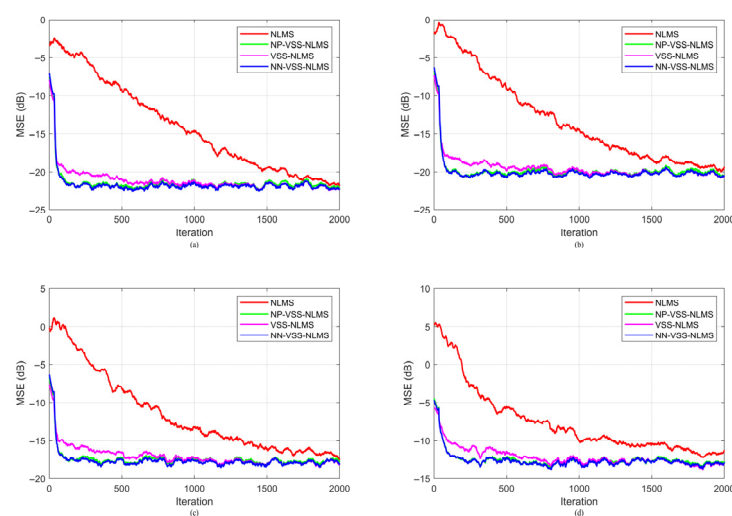


Figure 10. MSE curves of different algorithms with AR(1) input in a time-invariant system at SNR = 20 dB and $L = 8$ under different correlation levels: (a) $a = 0.3$, (b) $a = 0.5$, (c) $a = 0.7$, (d) $a = 0.9$.

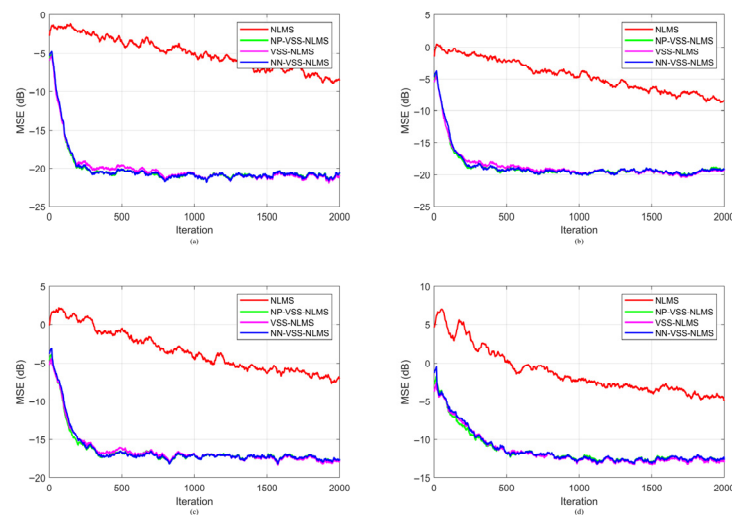


Figure 11. MSE curves of different algorithms with AR(1) input in a time-invariant system at SNR = 20 dB and $L = 32$ under different correlation levels: (a) $a = 0.3$, (b) $a = 0.5$, (c) $a = 0.7$, (d) $a = 0.9$.

Table 4. Performance metrics of different algorithms with AR(1) input in a time-invariant system at SNR = 20 dB for different correlation coefficients ($a = 0.3, 0.5, 0.7, 0.9$).

Filter Order L	Correlation Level a	Metric	NLMS	NP-VSS-NLMS	VSS-NLMS	NN-VSS-NLMS
8	0.3	Convergence Rate (dB/iteration)	0.01245	0.062827	0.064369	0.068016
		Steady-state Error (dB)	−17.641	−21.758	−21.512	−21.955
		Standard Deviations (dB)	0.29186	0.34086	0.31145	0.31162
	0.5	Convergence Rate (dB/iteration)	0.01437	0.067417	0.057339	0.06983
		Steady-state Error (dB)	−15.642	−20.043	−19.811	−20.252
		Standard Deviations (dB)	0.3216	0.36726	0.33839	0.34103
	0.7	Convergence Rate (dB/iteration)	0.014548	0.054301	0.04367	0.057732
		Steady-state Error (dB)	−15.391	−17.682	−17.404	−17.873
		Standard Deviations (dB)	0.43746	0.41572	0.38412	0.38363
	0.9	Convergence Rate (dB/iteration)	0.019465	0.037923	0.012667	0.037146
		Steady-state Error (dB)	−10.086	−12.798	−12.687	−12.903
		Standard Deviations (dB)	0.51497	0.55338	0.52526	0.53198

Table 4. Cont.

Filter Order L	Correlation Level α	Metric	NLMS	NP-VSS-NLMS	VSS-NLMS	NN-VSS-NLMS
32	0.3	Convergence Rate (dB/iteration)	\	0.053511	0.050999	0.054507
		Steady-state Error (dB)	\	−20.845	−20.746	−20.853
		Standard Deviations (dB)	0.37233	0.3827	0.34918	0.36255
	0.5	Convergence Rate (dB/iteration)	\	0.04374	0.043561	0.045154
		Steady-state Error (dB)	\	−19.461	−19.355	−19.449
		Standard Deviations (dB)	0.33843	0.34599	0.31988	0.32531
	0.7	Convergence Rate (dB/iteration)	\	0.029781	0.029003	0.031949
		Steady-state Error (dB)	\	−17.195	−17.161	−17.204
		Standard Deviations (dB)	0.40094	0.37122	0.36554	0.36818
	0.9	Convergence Rate (dB/iteration)	\	0.02369	0.021877	0.026133
		Steady-state Error (dB)	\	−12.236	−12.233	−12.256
		Standard Deviations (dB)	0.41295	0.48468	0.44401	0.45335

For $L = 8$, NN-VSS-NLMS maintains a relatively higher convergence rate across most correlation levels. As the input correlation increases, the convergence rates of all algorithms decrease, and the differences among the variable step-size algorithms become smaller. In the steady-state phase, NN-VSS-NLMS consistently achieves a lower steady-state error than the other algorithms.

For $L = 32$, a similar overall behavior is observed. NN-VSS-NLMS maintains a higher convergence rate than NP-VSS-NLMS and VSS-NLMS across different correlation levels. As the correlation increases, the margin in convergence rate gradually decreases. In the steady-state phase, the performance of the variable step-size algorithms becomes very close. NN-VSS-NLMS generally maintains a slightly lower steady-state error across different correlation levels. The standard deviations of the variable step-size algorithms remain within similar ranges under all conditions, indicating comparable fluctuation levels.

Overall, under both white noise and AR(1) stationary inputs, the proposed algorithm demonstrates better performance, validating its capability of adaptively adjusting the step-size in stationary environments.

3.6. Performance Analysis Under Non-Stationary Inputs

Non-stationary inputs represent typical conditions in real-world applications, where the algorithm must maintain convergence and tracking ability when the system or the input statistics undergo abrupt changes. Figures 12–15 present the performance of the proposed NN-VSS-NLMS algorithm, NLMS, NP-VSS-NLMS, and VSS-NLMS under speech inputs for both time-invariant and time-varying systems with filter orders of 8 and 32. Specifically, Figures 12 and 13 correspond to the time-invariant system with $L = 8$ and $L = 32$, respectively, while Figures 14 and 15 illustrate the corresponding results for the

time-varying system. The corresponding quantitative results for the time-invariant system are provided in Table 5.

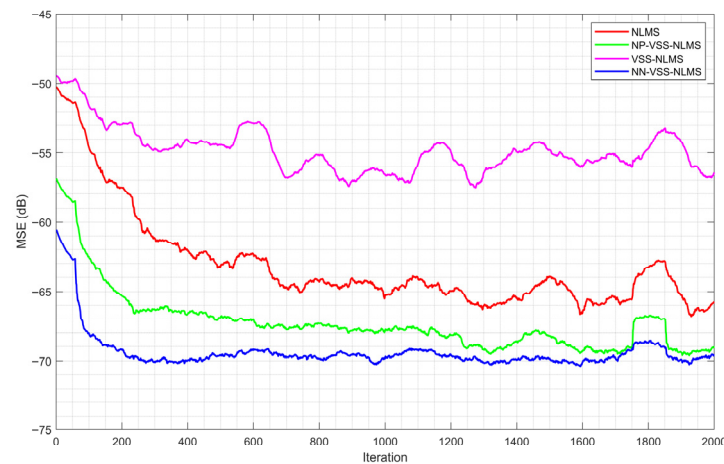


Figure 12. MSE curves of different algorithms with speech input in a time-invariant system at SNR = 20 dB and $L = 8$.

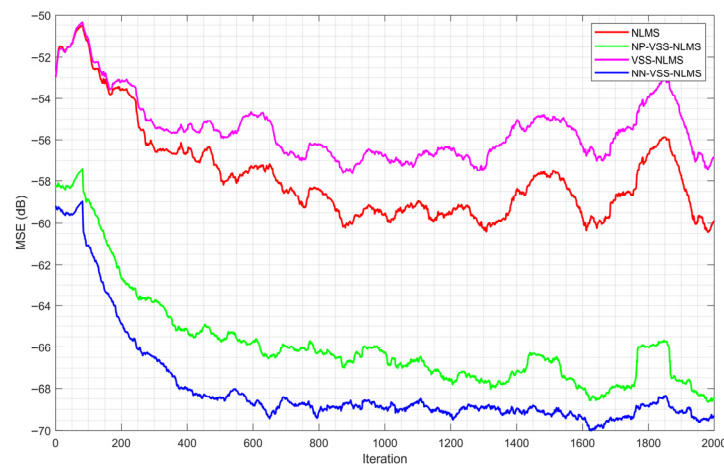


Figure 13. MSE curves of different algorithms with speech input in a time-invariant system at SNR = 20 dB and $L = 32$.

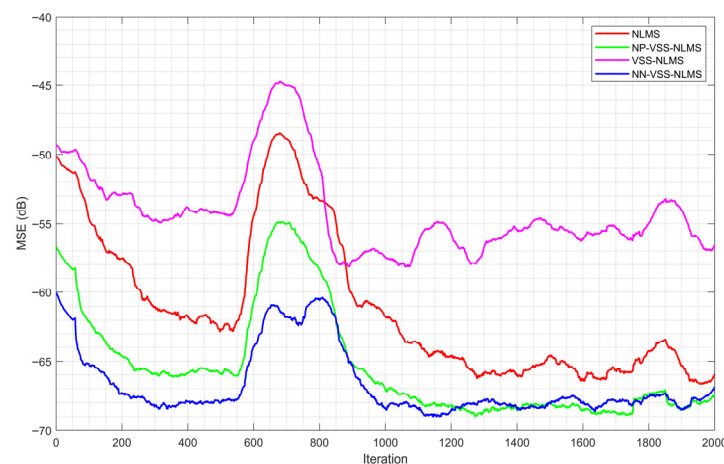


Figure 14. MSE curves of different algorithms with speech input in a time-varying system at SNR = 20 dB and $L = 8$.

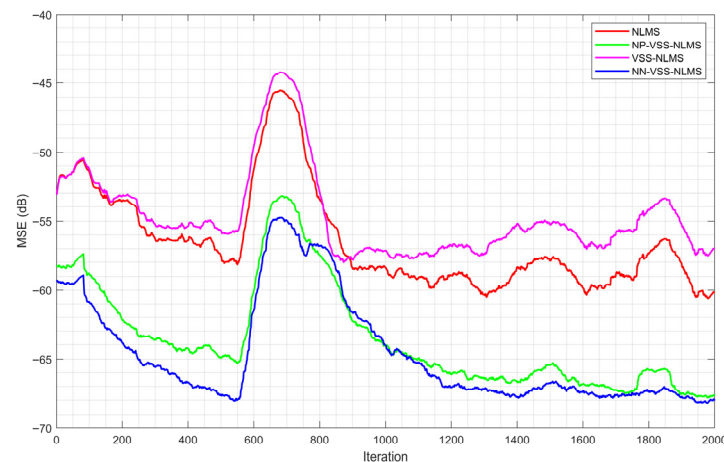


Figure 15. MSE curves of different algorithms with speech input in a time-varying system at SNR = 20 dB and $L = 32$.

Table 5. Performance metrics of different algorithms with speech input in a time-invariant system at SNR = 20 dB.

Filter Order L	Metric	NLMS	NP-VSS-NLMS	VSS-NLMS	NN-VSS-NLMS
8	Convergence Rate (dB/iteration)	0.043339	0.049149	0.018884	0.057867
	Steady-state Error (dB)	−63.886	−67.766	−55.087	−69.673
	Standard Deviations (dB)	0.21326	0.21728	0.14934	0.32904
32	Convergence Rate (dB/iteration)	0.0086559	0.021447	0.0063509	0.027948
	Steady-state Error (dB)	−58.303	−66.491	−55.852	−68.682
	Standard Deviations (dB)	0.22064	0.26039	0.17306	0.25474

For the time-invariant system, the corresponding quantitative results are provided in Table 5. When $L = 8$, the proposed algorithm exhibits a higher convergence rate (0.057867 dB/iteration) than the other algorithms, indicating a faster adaptation process under speech input. In the steady-state phase, it achieves a lower steady-state error (−69.673 dB), reflecting improved estimation accuracy. It is also observed that the standard deviation of NN-VSS-NLMS is relatively larger under this condition. This behavior can be associated with the more responsive step-size adjustment, which facilitates faster convergence while introducing slightly increased fluctuations around the steady state. However, the magnitude of this increase remains limited compared to the gain in steady-state accuracy.

When the filter order increases to $L = 32$, the convergence rate decreases due to the higher system dimensionality, while NN-VSS-NLMS still maintains a relatively higher convergence rate (0.027948 dB/iteration) among the compared algorithms. In steady state, it continues to achieve a lower steady-state error (−68.682 dB). In this case, the standard deviation becomes closer to those of NP-VSS-NLMS and VSS-NLMS, indicating that the fluctuation level is moderated as the filter order increases.

In time-varying systems (i.e., non-stationary environments), the tracking performance of all algorithms is evaluated using a variable filter, whose evolution pattern is shown in Figure 3 [36]. The main challenge in this scenario is that the system changes occur during periods when the step-size tends to be small, while VSS-NLMS algorithms must increase the step-size promptly to respond to system variations [36]. For $L = 8$, it can be observed that the proposed NN-VSS-NLMS algorithm responds effectively to the system change, with the MSE increasing during the abrupt transition and then decreasing as adaptation

resumes. Compared with NLMS, the proposed algorithm maintains a lower MSE level throughout the transition process. In comparison with NP-VSS-NLMS and VSS-NLMS, the overall MSE trajectory after the abrupt change remains at a relatively lower level, indicating improved tracking performance.

For $L = 32$, a similar trend is observed, while the adaptation process becomes slower due to the increased system order. During and after the abrupt-change interval, NN-VSS-NLMS maintains a relatively lower MSE level compared with the other algorithms. After the transition, its steady-state error remains lower, while the fluctuation range is comparable to that of NP-VSS-NLMS and VSS-NLMS.

3.7. Performance Analysis Under Different SNR Conditions

To evaluate the robustness of the algorithms under different noise conditions, the proposed algorithm is compared with NLMS, NP-VSS-NLMS, and VSS-NLMS over an SNR range from 0 to 50 dB, as illustrated in Figures 16 and 17 for filter orders of 8 and 32, respectively. When the SNR is 0 dB, the signal power is comparable to the noise power, and the step-size prediction of the NN becomes more sensitive to noise disturbances. As a result, the MSE curves exhibit relatively larger fluctuations, and the proposed algorithm performs slightly worse than NP-VSS-NLMS. As the SNR increases to 10–50 dB, the influence of noise is gradually reduced, leading to more stable step-size prediction. For $L = 8$, the proposed algorithm exhibits a higher convergence rate and achieves a lower steady-state error compared with the other algorithms across most SNR conditions. For $L = 32$, a similar tendency can be observed. As the SNR increases, all algorithms show improved convergence behavior and reduced steady-state error. NN-VSS-NLMS maintains a relatively faster convergence rate and a lower steady-state error, while the differences among the variable step-size algorithms become less pronounced compared with the case of $L = 8$.

Although its performance is affected at very low SNR, the proposed algorithm exhibits clear advantages in medium and high SNR scenarios, which supports its applicability in practical acoustic signal processing tasks.

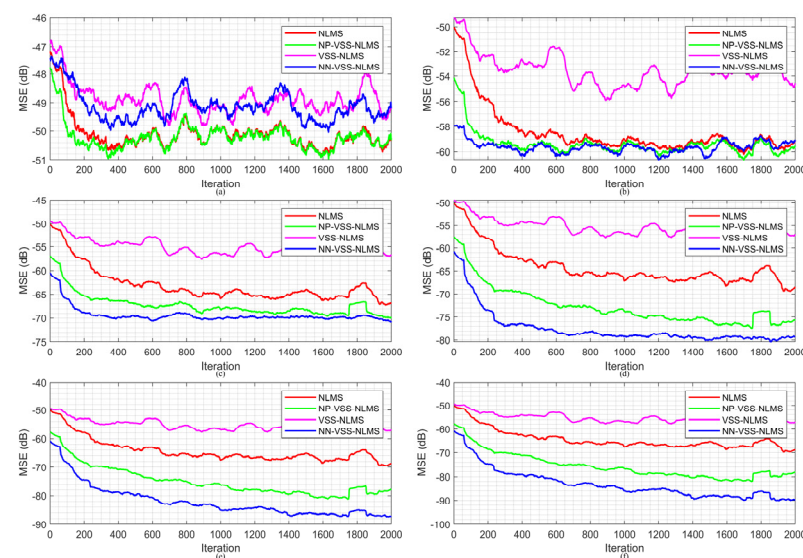


Figure 16. MSE curves of different algorithms under various SNR conditions with speech input in a time-invariant system at $L = 8$: (a) $SNR = 0$ dB, (b) $SNR = 10$ dB, (c) $SNR = 20$ dB, (d) $SNR = 30$ dB, (e) $SNR = 40$ dB, (f) $SNR = 50$ dB.

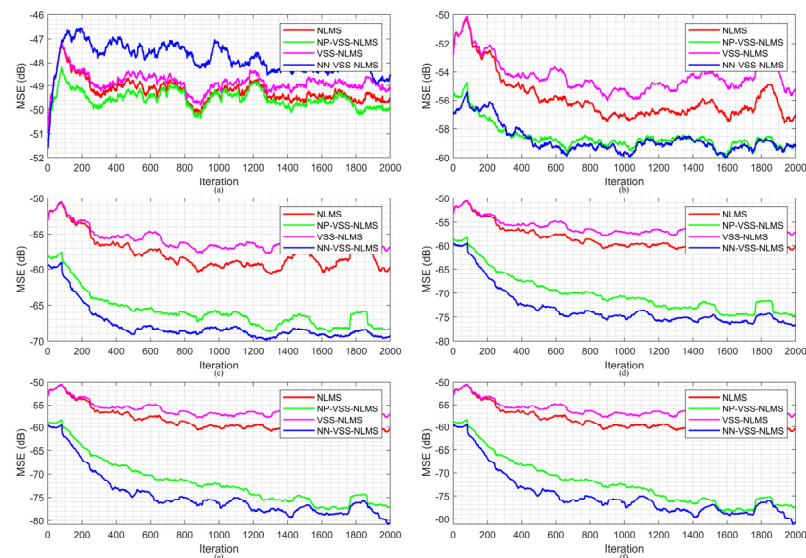


Figure 17. MSE curves of different algorithms under various SNR conditions with speech input in a time-invariant system at $L = 32$: (a) $SNR = 0$ dB, (b) $SNR = 10$ dB, (c) $SNR = 20$ dB, (d) $SNR = 30$ dB, (e) $SNR = 40$ dB, (f) $SNR = 50$ dB.

4. Conclusions

The proposed NN-VSS-NLMS algorithm improves adaptive filtering performance through learning-assisted step-size adjustment and demonstrates enhanced robustness under different signal conditions. By combining an analytically motivated reference step size with NN-based auxiliary adjustment, the algorithm achieves a fast convergence rate, reduced steady-state error, and strong tracking ability under stationary inputs, non-stationary inputs, and time-varying systems, thereby extending the applicability of traditional VSS-NLMS schemes to more challenging scenarios. The algorithm can be applied to practical signal-processing scenarios such as acoustic echo cancellation, active noise control, and system identification in environments with time-varying characteristics. Despite these advantages, several limitations should be acknowledged. First, the incorporation of an NN for step-size prediction introduces additional computational complexity, which may limit its applicability in large-scale or real-time scenarios. Second, the effectiveness of the NN-based prediction depends on the quality of the extracted features, and its performance may degrade under very low SNR conditions or severely corrupted sensor data. Third, although the algorithm exhibits good tracking ability in moderately time-varying systems, its adaptability to highly dynamic environments with abrupt changes may still be limited. Future work will focus on improving robustness under low SNR conditions by refining feature extraction and step-size control while reducing the network complexity to improve computational efficiency for real-time operation. Further evaluation in more realistic acoustic scenarios, including speech signals affected by nonlinear distortions or background interference, is expected to provide a more comprehensive understanding of the algorithm's performance.

Author Contributions: Conceptualization, Z.L. and Y.G.; methodology, Z.L. and Y.G.; software, Z.L. and Y.G.; validation, Z.L. and Y.G.; formal analysis, Z.L. and Y.G.; investigation, Z.L. and Y.G.; data curation, Z.L. and Y.G.; writing—original draft preparation, Z.L. and Y.G.; writing—review and editing, Z.L. and Y.G.; supervision, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

LMS	least-mean-square
NLMS	normalized least-mean-square
NN	neural network
SNR	signal-to-noise ratio
MSE	mean-square error

References

- Widrow, B.; Stearns, S.D. *Adaptive Signal Processing*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1985.
- Haykin, S. *Adaptive Filter Theory*, 4th ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 2002.
- Benesty, J.; Huang, Y. *Adaptive Signal Processing: Applications to Real-World Problems*; Springer: Berlin/Heidelberg, Germany, 2003.
- Huang, Y.; Benesty, J.; Chen, J. *Acoustic MIMO Signal Processing*; Springer: Boston, MA, USA, 2006.
- Ma, T.; Chen, J.; Chen, W.; Peng, Z. Unsupervised robust adaptive filtering against impulsive noise. *J. Syst. Eng. Electron.* **2012**, *23*, 32–39. [\[CrossRef\]](#)
- Li, Z.F.; Li, D.; Xu, X.L.; Zhang, J.Q. New normalized LMS adaptive filter with a variable regularization factor. *J. Syst. Eng. Electron.* **2019**, *30*, 259–269. [\[CrossRef\]](#)
- Wu, C.Y.; Li, W.D. A new variable step-size LMS algorithm based on t-distribution transformation. *J. Shenyang Univ. Technol.* **2021**, *40*, 82–87. (In Chinese)
- Widrow, B.; Hoff, M.E. Adaptive switching circuits. In *Proceedings of the IRE WESCON Convention Record*; IEEE: Piscataway, NJ, USA, 1960; Volume 4, pp. 96–104.
- Widrow, B.; McCool, J.; Ball, M. The complex LMS algorithm. *Proc. IEEE* **1975**, *63*, 719–720. [\[CrossRef\]](#)
- Lopes, P.A.C. Analysis of the LMS and NLMS algorithms using the misalignment norm. *Signal Image Video Process.* **2023**, *17*, 3623–3628. [\[CrossRef\]](#)
- Albert, A.E.; Gardner, L.S. *Stochastic Approximation and Nonlinear Regression*, 1st ed.; MIT Press: Cambridge, MA, USA, 1967.
- Nagumo, J.I.; Noda, A. A learning method for system identification. *IEEE Trans. Autom. Control* **1967**, *12*, 282–287. [\[CrossRef\]](#)
- Zhang, R.; Shi, G.C.; Liu, B.T.; Cheng, Y.R. Low-complexity variable step-size sign algorithm based on the Weibull distribution function. *Telecommun. Sci.* **2018**, *34*, 87–96. (In Chinese)
- Valin, J.-M.; Collings, I.B. Interference-normalized least mean square algorithm. *IEEE Signal Process. Lett.* **2007**, *14*, 988–991. [\[CrossRef\]](#)
- Rusu, A.G.; Paleologu, C.; Benesty, J.; Ciochină, S. A variable step size normalized least-mean-square algorithm based on data reuse. *Algorithms* **2022**, *15*, 111.
- Becker, A.C.; Kuhn, E.V.; Matsuo, M.V.; Benesty, J. On the NP-VSS-NLMS algorithm: Model, design guidelines, and numerical results. *Circuits Syst. Signal Process.* **2024**, *43*, 2409–2427. [\[CrossRef\]](#)
- Huang, H.-C.; Lee, J. A new variable step-size NLMS algorithm and its performance analysis. *IEEE Trans. Signal Process.* **2012**, *60*, 2055–2060. [\[CrossRef\]](#)
- Benyahia, A.; Bendoumia, R.; Hassani, I. New advanced deep learning variable step-size adaptive feed-forward algorithm for two-sensor acoustic noise reduction. *Trait. Signal* **2025**, *42*, 2453–2471. [\[CrossRef\]](#)
- Huemmer, C.; Maas, R.; Kellermann, W. The NLMS algorithm with time-variant optimum stepsize derived from a Bayesian network perspective. *IEEE Signal Process. Lett.* **2015**, *22*, 1874–1878. [\[CrossRef\]](#)
- Bhotto, M.Z.A.; Antoniou, A. A new partial-update NLMS adaptive-filtering algorithm. In *Proceedings of the IEEE 27th Canadian Conference on Electrical and Computer Engineering*, Toronto, ON, Canada, 4–7 May 2014; pp. 1–5.
- Hayashi, K.; Shiohara, K.; Sasaki, T. Differentiable programming based step size optimization for LMS and NLMS algorithms. In *Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2019*, Lanzhou, China, 18–21 November 2019; pp. 1721–1727.
- Huang, H.-C.; Lee, J. A new square-error-regularized NLMS. In *Proceedings of the Advances in Electrical and Electronics Engineering—IAENG Special Edition of the World Congress on Engineering and Computer Science 2008*, San Francisco, CA, USA, 22–24 October 2008; pp. 30–34.

23. Zhang, A.Y.; Gong, N.S. A Novel Variable Step Size LMS Algorithm Based on Neural Network. In Proceedings of the International Conference on Intelligent Systems and Knowledge Engineering (ISKE 2007), Chengdu, China, 15–16 October 2007. [\[CrossRef\]](#)
24. Wu, B.; Zhou, F. Application of neural network adaptive filter method to simultaneous detection of polymetallic ions based on ultraviolet-visible spectroscopy. *Front. Chem.* **2024**, *12*, 1409527. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Zhang, Y.; Yu, M.; Zhang, H.; Yu, D.; Wang, D.L. NeuralKalman: A learnable Kalman filter for acoustic echo cancellation. *arXiv* **2023**, arXiv:2301.12363. [\[CrossRef\]](#)
26. Li, K.; Yu, Y.; He, H.; Yu, T.; de Lamare, R. Novel normalized subband adaptive filtering algorithms with weights-dependent variable step-size. *Digit. Signal Process.* **2025**, *158*, 104945. [\[CrossRef\]](#)
27. Wang, Q.; Wang, G.; Liang, C. Deep neural network-driven adaptive filtering. *arXiv* **2025**, arXiv:2508.04258. [\[CrossRef\]](#)
28. Esmaeilbeig, Z.; Soltanian, M. Deep learning meets adaptive filtering: A Stein’s unbiased risk estimator approach. In *Proceedings of the 2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
29. Feng, P.; So, H.C. Meta-learning-based delayless subband adaptive filter using complex self-attention for active noise control. *Neurocomputing* **2025**, *650*, 130637. [\[CrossRef\]](#)
30. Shi, J.H.; Luo, Y.Y.; Liu, W.Y.; Mei, J.H.; Wang, W.K. A modified variable step size LMS adaptive filtering algorithm. *Comput. Appl. Softw.* **2013**, *30*, 184–186. (In Chinese)
31. Wang, W.H.; Zhang, H.M. A new and effective nonparametric variable step-size normalized least-mean-square algorithm and its performance analysis. *Signal Process.* **2023**, *210*, 109060. [\[CrossRef\]](#)
32. Baraniuk, R.G.; Cevher, V.; Wakin, M.B. Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective. *Proc. IEEE* **2010**, *98*, 959–971. [\[CrossRef\]](#)
33. Isserlis, L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **1918**, *12*, 134–139. [\[CrossRef\]](#)
34. Xie, X.; Liu, X.L.; Wang, J.X.; Guo, S.; Li, K.Q. RF tag quantity estimation method based on multilayer perceptron. *Chin. J. Comput.* **2023**, *46*, 499–511. (In Chinese)
35. Qing, C.J.; Du, Y.H.; Ye, Q.; Yang, N.; Zhang, M.T. Enhanced ELM superimposed CSI feedback with CSI estimation errors. *Comput. Sci.* **2022**, *49*, 632–638. (In Chinese)
36. Gueraini, I.; Benallal, A.; Tedjani, A. New variable step-size fast NLMS algorithm for non-stationary systems. *Signal Image Video Process.* **2023**, *17*, 3099–3106. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.