On the Variable Step-Size NLMS Algorithms

Junghsi Lee, Yu-Hung Teng, and Kuan-Rong Huang

Abstract—**This paper presents a comprehensive comparison of two variable step-size normalized least-mean-square algorithms, one was first introduced by Zipf, Tobias, and Seara in 2010 and was revisited with a stochastic model analyzing its behavior for both transient and steady-state phases in 2018, the other VSS -NLMS algorithm was proposed by Hsu-Chang Huang, and Junghsi Lee in 2012. Considering a system identification problem, some interesting characteristics of two algorithms are verified. Contrary to what Seara *et al*. have claimed in their publication, this paper shows that Lee's VSS-NLMS algorithm outperforms Seara's filter.**

Index Terms—**Acoustic echo cancellation, adaptive filter, normalized least mean-square (NLMS) algorithm, variable step-size NLMS.**

I. Introduction

Adaptive filtering algorithms have been widely employed in many signal processing applications such as channel equalization, active noise control and echo cancellation. The normalized least-mean-square (NLMS) adaptive filter is the most popular approach due to its low computational complexity and robustness [1]. The stability of the conventional NLMS is determined by a fixed step size. This parameter also controls the rate of convergence, speed of tracking ability and the amount of steady-state mean-square error (MSE). In general, the use of a large step size leads to a faster convergence in the early stage but along with a larger steady-state MSE. Conversely, a smaller step size makes convergence rate slower but along with smaller steady-state MSE.

Aiming to solve the conflicting objectives of fast convergence and low excess MSE associated with the basic NLMS, Mandic derived a generalized normalized gradient descent algorithm, which updates the regularization parameter gradient adaptively [2]. In the past two decades, a number of variable step-size NLMS (VSS-NLMS) algorithms have been proposed [2]-[6]. Most VSS-NLMS algorithms require the tuning of several parameters for better performance, which is usually carried out through a annoying trial-and-error process (see [3]-[5] for a detailed discussion). Benesty introduced a relatively tuning-free nonparametric VSS-NLMS algorithm [3]. Seara proposed a nonparametric VSS-NLMS filter in 2010 [4]. Huang and Lee presented a nonparametric algorithm, which employs the MSE and the estimated additive system noise power to control the variable step-size [5].

Manuscript received July 3, 2019; revised July 27, 2019.The authors are with the Department of Electrical Engineering, Yuan-Ze University, Chung-Li, TaoYuan City, 32003, TAIWAN
(email: eejlee@saturn.yzu.edu.tw;eddie830816@gmail.com; s1064711@mail.yzu.edu.tw)

Recently Seara *et al*. presented a comprehensive study of their early work on nonparametric VSS-NLMS algorithm [4] with a stochastic modeling analyzing its behavior for both transient and steady-state phases [6]. Simulations of comparison study showing their filter [4] outperforms VSS-NLMS algorithm developed by Huang and Lee [5] were presented in [6]. In this paper, we revisit these two VSS-NLMS algorithms. Extensive simulation results have demonstrated that Huang's algorithm enjoys a better performance than that of Seara's VSS-NLMS algorithm.

II. Algorithms

A. Basic Model

Let $d(n)$ be the desired response of the adaptive filter

$$d(n)=\mathbf{x}^{\mathrm{T}}(n)\mathbf{h}(n)+v(n)\ =y(n)+v(n) \tag{1}$$

where $y(n)=\mathbf{x}^{\mathrm{T}}(n)\mathbf{h}(n)$, and $\mathbf{h}(n)$ denotes the coefficient. vector of the unknown system with length M,

$$\mathbf{h}(n)=[h_0(n),h_1(n),\ldots,\ h_{M\text{-}1}(n)]^{\mathrm{T}} \tag{2}$$

$\mathbf{x}(n)$ is the input signal vector

$$\mathbf{x}(n)=[x(n),x(n+1),\ldots.,\ x(n-M+1)]^{\mathrm{T}} \tag{3}$$

and $v(n)$ is the system noise that is independent of $\mathbf{x}(n)$.

For simplicity, the adaptive filter is assumed to have the same structure as that of the unknown system. Denoting its coefficient vector at iteration *n as* $\mathbf{w}(n)$, the a priori estimation error is evaluated as

$$e(n)=d(n)-\mathbf{x}^{\mathrm{T}}(n)\mathbf{w}(n) \tag{4}$$

The key issue in any VSS-NLMS algorithm is the means by which to vary the step size. The general weight update equation of VSS-NLMS algorithms is

$$\mathbf{w}(n+1)=\mathbf{w}(n)+\frac{\mu(n)}{\|\mathbf{x}(n)\|^2+\varepsilon}\mathbf{x}(n)e(n) \tag{5}$$

where ε is a positive regularization parameter, the variable step size $\mu(n)$ is bounded by zero and two to guarantee the stability, so the key point of VSS-NLMS algorithm is how to choose suitable step size value.

B. VSS-NLMS-H Algorithm

We briefly review Huang and Lee work in [5] and denote it as VSS-NLMS-H algorithm. This filter employs the mean-square error and the estimated system noise power to control the step-size update as follows.

$$\mu(n) = \alpha\mu(n-1) + (1-\alpha)\,\frac{\hat{\sigma}_e^2(n)}{\beta\hat{\sigma}_v^2(n)} \qquad (6)$$

where β is a positive parameter, the estimated MSE $\hat{\sigma}_e^2(n)$ and the system noise power $\hat{\sigma}_v^2(n)$ can be obtained recursively

$$\hat{\sigma}_e^2(n) = \alpha\hat{\sigma}_e^2(n-1) + (1-\alpha)e^2(n) \qquad (7)$$

$$\hat{\sigma}_v^2(n) = \hat{\sigma}_e^2(n) - \frac{1}{\hat{\sigma}_x^2(n)}\hat{\boldsymbol{r}}_{ex}^T(n)\hat{\boldsymbol{r}}_{ex}(n) \qquad (8)$$

In (8), $\hat{\boldsymbol{r}}_{ex}^T(n)$ denotes the cross-correlation between $\boldsymbol{x}(n)$ and $\boldsymbol{e}(n)$, *and* $\hat{\sigma}_x^2(n)$ is the input signal power, that can be estimated as

$$\hat{\sigma}_x^2(n) = \alpha\hat{\sigma}_x^2(n-1) + (1-\alpha)x^2(n) \qquad (9)$$

$$\hat{\boldsymbol{r}}_{ex}(n) = \alpha\hat{\boldsymbol{r}}_{ex}(n-1) + (1-\alpha)\mathbf{x}(n)e(n) \qquad (10)$$

The variable step size $\mu(n)$ value depends on a statistic $\zeta(n)$

$$\begin{cases} \mu(n) = \alpha\,\mu(n-1) + (1-\alpha)\dfrac{\hat{\sigma}_e^2(n)}{\beta\,\hat{\sigma}_v^2(n)} & ,\ \zeta(n) < \zeta_{th} \\ \mu(n) = 1 & ,\ \zeta(n) > \zeta_{th} \end{cases} \qquad (11)$$

where ζ(th) is a small positive quantity. The statistic ζ(n) is defined to be

$$\zeta(n) = \frac{|\,\hat{r}_{de}(n) - \hat{\sigma}_e^2(n)|}{|\,\hat{\sigma}_d^2(n) - \hat{r}_{de}(n)| + c} \qquad (12)$$

where $\hat{\boldsymbol{r}}_{de}$(n) is an estimate of E{d(n)e(n)}.

In the early stage $\hat{\sigma}_e^2(n)$ is generally big due to the system mismatch, so the adaptive filter uses a large $\mu(n)$. When the algorithm starts to converge, $\hat{\sigma}_e^2(n)$ becomes smaller, and $\mu(n)$ get smaller. When the adaptive filter converges to the optimum solution, $\hat{\sigma}_e^2(\infty)$ is pretty close to $\hat{\sigma}_v^2(\infty)$ resulting in a constant step size, $\mu(n) \approx 1/\beta$. Note that if system noise becomes larger suddenly, $\mu(n)$ tends to decrease.

C. VSS-NLMS-S Algorithm

This paper denotes the algorithm proposed by Seara et al. in [4] as VSS-NLMS-S. The variable step size is updated as

$$\mu(n) = \frac{p^2(n)}{q^2(n)} \qquad (13)$$

where $p(n)$ is the error correlation between $e(n)$ and $e(n$-1)

$$p(n) = \beta p(n-1) + (1-\beta)e(n)e(n-1) \qquad (14)$$

and $q(n)$ is the smoothed squared error signal

$$q(n) = \beta q(n-1) + (1-\beta)e^2(n) \qquad (15)$$

It should be noted this β is a forgetting factor close to 0.99 and is different from the one used in (6).

In the early stage $e(n)$ is strongly correlated with $e(n$-1), implying that $p^2(n)$ is approximately equal to $q^2(n)$. Therefore, the step size $\mu(n)$ is likely close to 1, speeding up the convergence process. When the adaptive filter converges to nearly optimum, leading to a weak correlation between $e(n)$ and $e(n$-1), as a result $p^2(n)$ becomes smaller than $q^2(n)$ and step size $\mu(n)$ tends to be small thus reducing the steady-state error. So $\mu(n)$ varies dynamically between 1 and 0 (very close to zero) during the adaptation process.

III. SIMULATION RESULTS

In this section, we conduct simulation comparison study for VSS-NLMS-S and VSS-NLMS-H.

A. Example 1

The adaptive filter is used to identify a system. A zero-mean Gaussian input signal x(n) with variance $\sigma_x^2 = 1$ is used, which is obtained from

$$x(n) = -a_1x(n-1) - a_2x(n-2) + v(n) \qquad (16)$$

where a_1 and a_2 denote the AR (2) coefficients, and $v(n)$ is a white Gaussian noise with variance.

$$\sigma_v^2 = \left(\frac{1-a_1}{1+a_2}\right)\left[(1+a_2)^2 - a_1^2\right] \qquad (17)$$

Two values of signal-to-noise (SNR), 20dB, 40dB were considered. The impulse response of System-1 with length M=128 are first taken from the sinc function

$$h(n) = [sinc(0), sinc(1/\mathrm{M}), \cdot\cdot, sinc(\mathrm{M}-1/\mathrm{M})]^T \qquad (18)$$

as shown in Fig. 1, then being normalized, *i.e.*, h^T(n) h(n)=1 so that the simulation scenario is the same as that in [6]. The parameter settings are tabulated in Table I. The impulse response changes signs at 50000th sample. The excess MSE (EMSE) curves (in decibels) of averages of 200 independent runs are presented in Fig. 2.

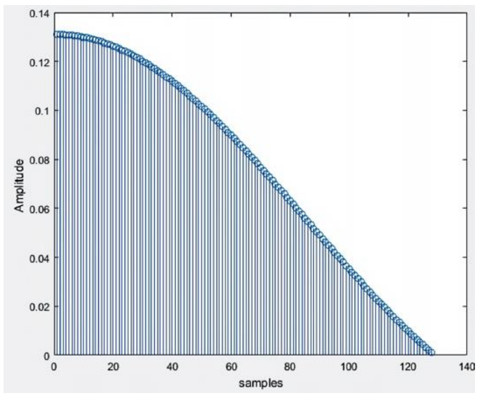

Fig. 1. Impulse response of System-1.

ISBN: 978-988-14048-7-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

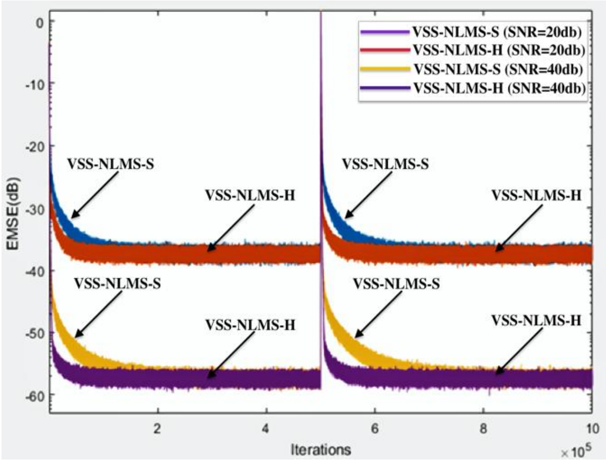

Fig. 2. EMSE curves of the VSS-NLMS-S algorithm and VSS-NLMS-H algorithm. System-1is the unknown system.

As seen in the Fig. 2, both algorithms maintain an equal steady-state MSE while the VSS-NLMS-H algorithm obviously performs with faster convergence rate and better tracking capability for time-varying system than that of the VSS-NLMS-S algorithm.

TABLE I
THE PARAMETER SETTING OF EXAMPLE 1

Input signal	AR
AR coefficients	a_1= -0.6、a_2= 0.85
Unknown system-1	Sinc function
Algorithm	Parameter setting
VSS-NLMS-S	β=0.975
VSS-NLMS-H	α=0.998，β=30，ζ_{th} = 0.35

B. Example 2

We evaluate these two VSS-NLMS algorithms in the same scenario as the previous experiment using a random unknown system model, denoted as System-2, as shown in Fig. 3. The EMSE curves (in decibels) of averages of 200 independent runs are presented in Fig. 4.

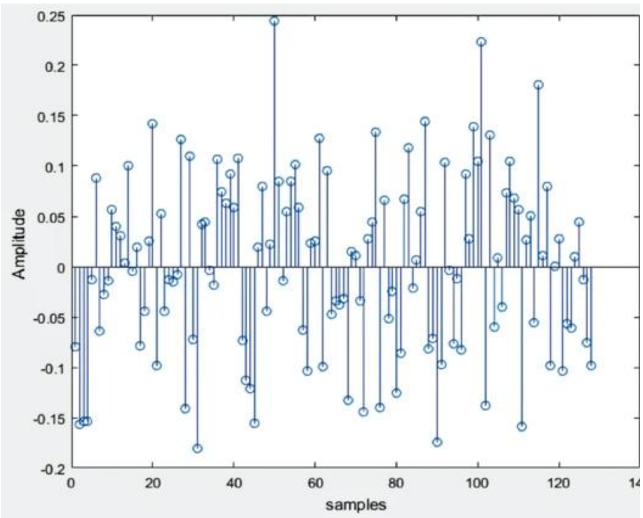

Fig. 3. Impulse response of System-2.

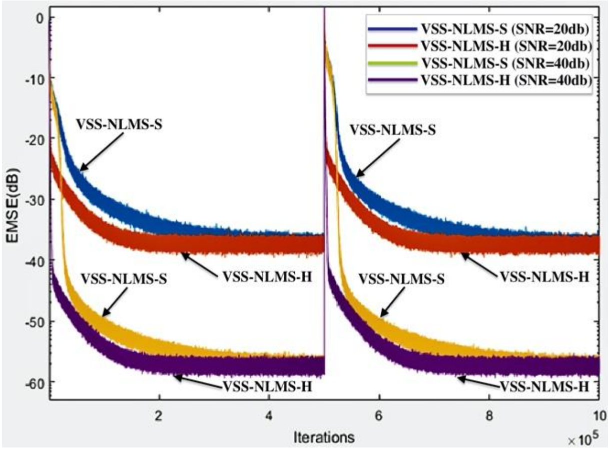

Fig. 4. EMSE curves of the VSS-NLMS-S algorithm and VSS-NLMS-H algorithm. System -2 is the unknown system.

As can be seen in the Fig. 4, both algorithms exhibit slower convergence rate. However, VSS-NLMS-H algorithm clearly outperforms VSS-NLMS-S algorithm for a not-so-smooth impulse response system.

C. Example 3

Simulation results in Example 2 indicates that both VSS-NLMS algorithms displayed deteriorated convergence performance in identifying a rough-shape system. We conduct another experiment by adding zero-mean white Gaussian noise with 3 different variances (0.031, 0.1, and 0.316) into System-1 and observe the filter performance for different rough-shape-level systems. The parameter settings are tabulated in Table III. A standard NLMS algorithm is also employed for comparison purpose. The input signal is a white Gaussian process with zero mean and unit variance. The performance index shown in Figs. 5-8 are system distance (SD) curves and MSE curves, which are ensemble averages of 200 independent runs.

TABLE II
THE PARAMETER SETTING OF EXAMPLE 3

Input signal	White Gaussian random process		
Unknown System-1	system-1		
	system-1	σ^2 = 0.031	AWGN is added
	system-1	σ^2 = 0.1	AWGN is added
	system-1	σ^2 = 0.316	AWGN is added
Algorithm	Parameter setting		
NLMS	μ=1		
VSS-NLMS-S	β=0.99		
VSS-NLMS-H	α=0.998，β=30，ζ_{th} =0.35		

ISBN: 978-988-14048-7-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

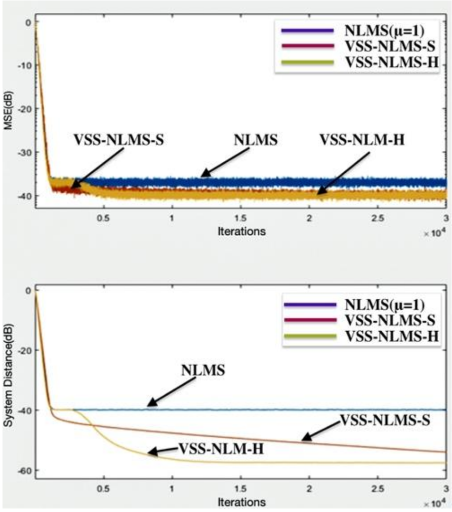

Fig. 5. MSE curves and SD curves of the NLMS algorithm (step size is 1) VSS-NLMS-S algorithm and VSS-NLMS-H algorithm. System-1 is the unknown system.

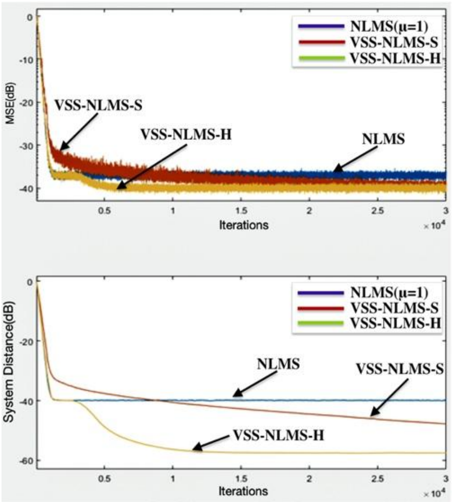

Fig. 6. MSE curves and SD curves of the NLMS algorithm (step size is 1) VSS-NLMS-S algorithm and VSS-NLMS-H algorithm. System-1 with AWGN σ^2=0.031 is the unknown system.

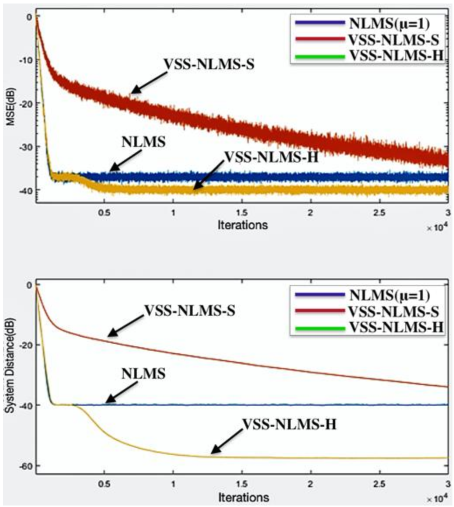

Fig. 7. MSE curves and SD curves of the NLMS algorithm (step size is 1) VSS-NLMS-S algorithm and VSS-NLMS-H algorithm. System-1 with AWGN σ^2=0.1 is the unknown system.

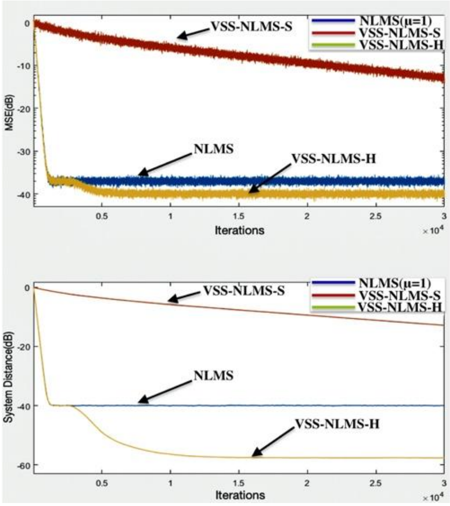

Fig. 8. MSE curves and SD curves of the NLMS algorithm (step size is 1) VSS-NLMS-S algorithm and VSS-NLMS-H algorithm. System-1 with AWGN σ^2=0.316 is the unknown system.

Fig. 5 shows the log of MSE curves (above) and log of SD curves (bottom) of NLMS and two VSS-NLMS algorithms for the smooth system-1. VSS-NLMS-S algorithm seems to have slight convergence advantage. Fig. 6 shows that for the a-bit-rough-shape system, the VSS-NLMS-S converges even a bit slower than the basic NLMS. Figs. 7 and 8 show more surprisingly results that the VSS-NLMS-S performs the worst the whole process. And it is obvious that the VSS-NLMS-H has the best performance and has performed very consistent for all different types of systems.

IV. Conclusions

This paper presented a comprehensive comparison of two VSS-NLMS algorithms. Contrary to what Seara *et al*. have claimed in their recent publication, Lee's VSS-NLMS algorithm has shown to perform with fast convergence rate, good tracking, and low misadjustment. Simulation results demonstrate that it outperforms Seara's filter in all types of systems.

References

[1] S. Haykin, *Adaptive Filter Theory*, 5th ed., Prentice-Hall, Upper Saddle River, NJ, 2014.
[2] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Process. Lett.* vol. 11, no. 2, pp. 115-118, Feb. 2004.
[3] J. Benesty, H. Rey, L. R. Vega, S. Tressens, "A nonparametric VSS NLMS algorithm," *IEEE Signal Process. Lett.* vol. 13, no. 10, pp. 581-584, Oct. 2006.
[4] J. G. F. Zipf, O. J. Tobias, R. Seara, "Non-parametric VSS-NLMS algorithm with control parameter based on the error correlation," *Proceedings of the IEEE International Telecommunications* Symposium (ITS), pp. 1–5, Sep. 2010.
[5] H. C. Huang and J. Lee, "A new variable step-size NLMS algorithm and its performance analysis," *IEEE Trans. Signal Process*., vol. 60, no. 4, pp. 2055-2060, Apr. 2012.
[6] E. V. Kuhn, J. G. F. Zipf, R. Seara, "On the stochastic modeling of a VSS-NLMS algorithm with high immunity against measurement noise," *Signal Process*., vol. 147, pp. 120-132, 2018.

ISBN: 978-988-14048-7-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)