

Adaptive Polynomial Filtering using Generalized Variable Step-Size Least Mean p th Power (LMP) Algorithm

Amrita Rai · Amit Kumar Kohli

Received: 6 October 2013 / Revised: 22 May 2014 / Accepted: 22 May 2014 /
Published online: 10 June 2014
© Springer Science+Business Media New York 2014

Abstract This correspondence presents the adaptive polynomial filtering using the generalized variable step-size least mean p th power (GVSS-LMP) algorithm for the nonlinear Volterra system identification, under the α -stable impulsive noise environment. Due to the lack of finite second-order statistics of the impulse noise, we espouse the minimum error dispersion criterion as an appropriate metric for the estimation error, instead of the conventional minimum mean square error criterion. For the convergence of LMP algorithm, the adaptive weights are updated by adjusting $p \geq 1$ in the presence of impulsive noise characterized by $1 < \alpha < 2$. In many practical applications, the autocorrelation matrix of input signal has the larger eigenvalue spread in the case of nonlinear Volterra filter than in the case of linear finite impulse response filter. In such cases, the time-varying step-size is an appropriate option to mitigate the adverse effects of eigenvalue spread on the convergence of LMP adaptive algorithm. In this paper, the GVSS updating criterion is proposed in combination with the LMP algorithm, to identify the slowly time-varying Volterra kernels, under the non-Gaussian α -stable impulsive noise scenario. The simulation results are presented to demonstrate that the proposed GVSS-LMP algorithm is more robust to the impulsive noise in comparison to the conventional techniques, when the input signal is correlated or uncorrelated Gaussian sequence, while keeping $1 < p < \alpha < 2$. It also exhibits flexible design to tackle the slowly time-varying nonlinear system identification problem.

A. Rai
Department of Electronics and Communication Engineering,
R.I.E.T, Faridabad 0121007, Haryana, India
e-mail: subhupari@gmail.com

A. K. Kohli (✉)
Department of Electronics and Communication Engineering, Thapar University,
Patiala 147004, Punjab, India
e-mail: drkohli_jitr@yahoo.co.in

Keywords Volterra filter · Polynomial filters · Time-varying channels · Fixed step-size LMS (FSS-LMS) · Variable step-size LMS (VSS-LMS) · MMSE · Minimum error dispersion (MED)

1 Introduction

In the field of communication engineering, speech processing, image processing, and biomedical engineering, etc., many systems possess certain degrees of nonlinearity, which do not exhibit superposition property. Any such polynomial system [11] is also called Volterra system [9], which is most commonly referred /used paradigm due to its roots in the Taylor series expansion of the nonlinear functions with memory [21]. Therefore, the nonlinear system identification [16] is indispensable to establish a mathematical model for an unknown system through the input–output relationship. The researchers in the field of nonlinear system identification usually consider Volterra, Weiner [17] and Hammerstein [26] models. However, the presented research work will focus on the variable step-size adaptive nonlinear Volterra filtering [12], due to its low computational complexity compared to the variable step-size adaptive Hammerstein filtering [25].

The measurement noise is an inevitable issue in the field of nonlinear system identification, which is generally assumed to be a random process with the finite-order statistics. Under such scenario, the mean square error (MSE) appears as an appropriate metric for the estimation error. However, the impulsive noise [15] with the heavier distribution tail possesses approximately infinite second-order statistics, which connotes non-Gaussian characteristics. It leads to the need of alternate methods for the nonlinear system identification in the presence of impulse noise.

The Gaussian distribution is a special case of α -stable processes with $\alpha = 2$, which is characterized by the finite variance [20]. It is noteworthy that the α -stable processes in the range $1 < \alpha < 2$ are considered to be non-Gaussian with infinite variance. In [24], Stuck has discussed that a finite variance Gaussian model is appropriate over a limited range of data, while an infinite variance model is adequate in terms of matching the observed data over a wider range. Therefore, the impulse noise occurrence may be modeled as non-Gaussian for further analysis, which favors the application of adaptive nonlinear filtering for the noise excision and nonlinear system identification [14]. Under the aforementioned conditions, the cost function based on the minimum error dispersion (MED) outperforms the conventional minimum mean square error (MMSE)-based approach [22]. Moreover, it results in the development of least mean p th power (LMP) adaptation algorithm, in which the cost function is convex with respect to the filter weights for the range $p \geq 1$. However, the performance of LMP algorithm supersedes the conventional LMS algorithm, only when the value of parameter p is close to α for the range $1 < p < \alpha$. But in [28], Weng and Barner have delineated that the large eigenvalue spread of the input signal autocorrelation matrix has been observed in the case of Volterra filtering, which in turn results in the slow convergence speed/rate of the LMP, as well as LMS adaptive algorithms. However, the nonlinear Volterra FSS-LMS filter can encounter divergence in case of the ill-conditioned tap input autocorrelation matrix [25].

The time-varying step-size is one of the tractable solutions to expedite the convergence process in the case of LMP algorithm. Kwong and Johnston have proposed a variable step-size LMS (KVSS-LMS) adaptive algorithm in [10] for the tracking of time-varying first-order Markovian channels, in which the step-size adjustment is controlled by the square of prediction error. Further, Aboulnasr and Mayyas have presented a variable step-size LMS (AVSS-LMS) adaptive algorithm in [1], in which the step-size of the algorithm is adjusted according to the square of time-averaged estimate of the autocorrelation of present/instantaneous estimation error $e(n)$ and the past estimation error $e(n-1)$. In an alternate approach proposed by Ang and Farhang-Boroujeny in [2], the step-size of adaptive filter is changed according to a stochastic gradient adaptive algorithm designed to reduce the squared estimation error at each iteration, which is denoted as SVSS-LMS algorithm. All the aforementioned VSS-LMS algorithms are implemented using the linear filtering perspective.

In this paper, we propose adaptive nonlinear Volterra filtering using the generalized variable step-size least mean p th power (GVSS-LMP) algorithm for the slowly time-varying system identification, in the presence of α -stable impulsive noise. This combination of GVSS and LMP algorithm enhances the convergence rate under the noisy environment. However, it reduces to the various VSS-LMP and VSS-LMS adaptive algorithms under the typical parametric conditions, which signifies its flexibility. This paper is organized as follows. In Sect. 2, we first describe the slowly time-varying nonlinear Volterra system (as shown in Fig. 1) along with the details of α -stable impulse noise characteristics. We next introduce the adaptive nonlinear system identification method based on the MED criterion using the proposed GVSS-LMP algorithm in Sect. 3. Subsequently, the convergence and tracking mode performances of the presented algorithm are compared with KVSS-LMP [10, 28], AVSS-LMP [1, 28] and SVSS-LMS [2] adaptive algorithms in Sect. 4, to manifest its benefits and efficacy on the basis of simulation results. Finally, the concluding remarks and future scope are illustrated in Sect. 5.

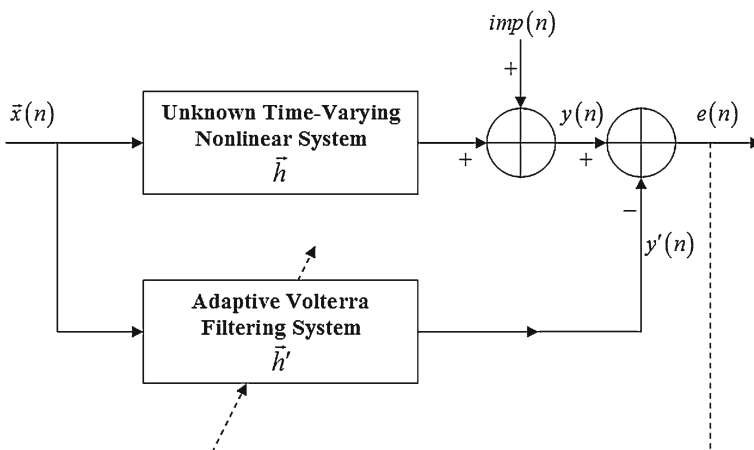


Fig. 1 Nonlinear slowly time-varying system identification configuration

2 Nonlinear System Model in Noisy Environment

2.1 Slowly Time-Varying Volterra System

Among polynomial system models, the Volterra system [21] is the preferred paradigm because its output is nonlinear with respect to the input signals, but it is linear in terms of kernels. Therefore, the adaptive signal processing techniques may be directly extended to the Volterra filtering. In literature, there are many time-varying nonlinear wireless or underwater acoustic communication channels, which need to be tracked or estimated by the nonlinear polynomial adaptive filtering. For identification of these unknown systems, we consider the configuration shown in Fig. 1, in which the underlying system and the adaptive nonlinear Volterra filter are driven by the common input signal vector $\vec{x}(n)$. In the presence of impulse noise, the general input–output relationship of an unknown nonlinear system can be illustrated by a truncated Volterra series as

$$y(n) = h_0 + \sum_{k=1}^K \sum_{m_1=0}^{M-1} \dots \sum_{m_k=0}^{M-1} h_k(n; m_1, \dots, m_k) \prod_{i=1}^k x(n - m_i) + imp(n) \quad (1)$$

Typically, the second-order Volterra series is described by the input–output relationship as

$$y(n) = h_0 + \sum_{m_1=0}^{M-1} h_1(n; m_1)x(n - m_1) + \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} h_2(n; m_1, m_2)x(n - m_1) x(n - m_2) + imp(n), \quad (2)$$

where h_0 is the time-invariant zeroth-order Volterra kernel, h_1, h_2 are the first-order and the second-order Volterra kernels, respectively, M is the memory length, $x(n)$ is the input signal, and $imp(n)$ is the α -stable noise with zero-mean (inevitable disturbance). The complexity of Volterra filter is dependent upon the memory (M). In the general case, the degree of nonlinearity (K) of the Volterra system is usually assumed to be time-invariant [3]. As the Volterra kernels are symmetrical in nature, the value of coefficient $h_k(n; m_1, \dots, m_k)$ is kept unchanged for any of the possible $k!$ permutations of m_1, m_2, \dots, m_k . Hence, these kernels remain time-invariant under the different permutations of its argument.

In the presented work, the values of K and M are considered to be known *a priori*. For the slowly time-varying second-order Volterra system, the input–output relationship is depicted by (1) with $K = 2$. Now, let us consider the $L \times 1$ dimensional expanded filter coefficients vector as

$$\vec{h}(n) = [h_1(n; 0), h_1(n; 1), \dots, h_1(n; M - 1), h_2(n; 0, 0), h_2(n; 0, 1), \dots, \dots, h_2(n; 0, M - 1), h_2(n; 1, 1), \dots, h_2(n; M - 1, M - 1)]^T, \quad (3)$$

where $(.)^T$ is the matrix transpose operator. The $L \times 1$ dimensional expanded input signal vector for the second-order Volterra filter with zero-mean and variance

$\sigma_x^2 = 1/L$ is denoted as

$$\vec{x}(n) = [x(n), x(n - 1), \dots, x(n - M + 1), \dots, x^2(n), x(n)x(n - 1), \dots, \dots, x(n)x(n - M + 1), \dots, x^2(n - 1), \dots, x^2(n - M + 1)]^T \tag{4}$$

Further, we can express Eq. (2) in the vector form as

$$y(n) = \vec{h}^T(n)\vec{x}(n) + imp(n) \tag{5}$$

In the nonlinear system identification, the final goal is to identify the time-varying Volterra kernels $h_k(n; m_1, \dots, m_k)$ in Eq. (1) through measured $y(n)$ and $x(n)$, which follow the Random Walk model [1, 2, 10, 13] given by $\vec{h}(n + 1) = \vec{h}(n) + \vec{w}(n + 1)$; where $\vec{w}(n)$ is the zero-mean white Gaussian process noise vector with variance $\sigma_w^2 = 0.001$ (assumed to be small for the slow time-variations). For the mathematical analysis, the adaptively estimated Volterra kernel vector may be represented by

$$\vec{h}'(n) = [h'_1(n; 0), h'_1(n; 1), \dots, h'_1(n; M - 1), h'_2(n; 0, 0), h'_2(n; 0, 1), \dots, \dots, h'_2(n; 0, M - 1), h'_2(n; 1, 1), \dots, h'_2(n; M - 1, M - 1)]^T \tag{6}$$

Therefore, the estimated received signal is denoted by

$$y'(n) = \vec{h}'^T(n)\vec{x}(n) \tag{7}$$

Hence, the output estimation error in the signal reception is computed by

$$e(n) = y(n) - y'(n) \tag{8}$$

This error signal is fed back to the adaptive filter (self-designing filter), which begins from an initial guess based on the prior knowledge available to the system; and then it converges eventually to the optimal solution in some statistical sense through the successive iterations.

2.2 Symmetric α -Stable Noise Model

An α -stable process can be described by the following characteristic function [19, 22], as it exhibits no closed-form probability density function.

$$\Phi(\Omega) = \exp [j\eta\Omega - \gamma |\Omega|^\alpha \{1 + j\beta \text{sgn}(\Omega)S(\Omega, \alpha)\}] \tag{9}$$

where, $S(\Omega, \alpha) = \begin{cases} \tan(\frac{\alpha\pi}{2}) & \text{for } \alpha \neq 1 \\ \frac{2}{\pi} \log |\Omega| & \text{for } \alpha = 1 \end{cases}$, $0 < \alpha \leq 2$, $-\infty < \eta < +\infty$, $\gamma > 0$,

and $-1 \leq \beta \leq +1$. Thus, a stable distribution is completely determined by four parameters: (1) the location parameter η , (2) the index of skewness β (the distribution

is symmetric about its location parameter η , when $\beta = 0$, therefore, called symmetric α -stable distribution), (3) the scale parameter γ is called dispersion (the parameter $\gamma^{1/\alpha}$ plays a role similar to the standard deviation of the Gaussian distribution), and (4) α is the characteristic exponent. This shape parameter is a measure of the heaviness of the tail of distribution. The processes with small values of α are considered to be impulsive. However, for the large values of α , the observed values of random variable are not far from its central location. Under typical conditions, when the value of $\alpha \rightarrow 2$ and $\beta = 0$, then $\Phi(\Omega) \rightarrow \exp[j\eta\Omega - \gamma|\Omega|^2]$. This relevant stable distribution is Gaussian in nature. However in the presented work, the α -stable random variables do not have finite variance, but these are characterized only by the finite p th-order moments for $p < \alpha$. It is noteworthy fact that all the moments of order less than α do exist and are called the fractional lower order moments (FLOM) [22], which can be derived from its dispersion and the characteristic exponent with zero location parameter as

$$E|X|^p = C(p, \alpha)\gamma^{p/\alpha} \text{ for } 0 < p < \alpha$$

$$\text{where, } C(p, \alpha) = \frac{2^{p+1}\Gamma((p+1)/2)\Gamma(-p/\alpha)}{\alpha\sqrt{\pi}\Gamma(-p/2)} \quad (10)$$

which is dependent on the values of parameters α and p , not on random variable X . In the above equation, the parameter Γ is the gamma function [19,22].

3 GVSS-LMP Algorithm for Nonlinear System Identification

3.1 Least Mean p th Power Adaptive Algorithm

In the field of adaptive signal processing [7], the most popular approaches for the estimation/prediction schemes are based on the MMSE criterion. The corresponding cost function as per the Wiener theory is

$$J_{\text{MMSE}}(\vec{h}') = E \left[\left| y(n) - \vec{h}'^T(n)\vec{x}(n) \right|^2 \right], \quad (11)$$

where $E[*]$ is the ensembled average operator. Using the error signal $e(n)$, the non-mean square error criterion is discussed in the form of least mean forth (LMF) adaptive algorithm in [27], in which the cost function is considered to be $E \left[|e(n)|^{2\bar{K}} \right]$ for $\bar{K} \geq 1$ (only integer values of \bar{K}). In some typical cases, the LMF algorithm with $\bar{K} > 1$ outperforms the conventional LMS algorithm by providing less noise in weights for the same speed of convergence. It has motivated the evolution of LMP algorithm for the noisy situations, in which $1 < p < \alpha < 2$. Particularly, when $\vec{h}'(n) \rightarrow \vec{h}(n)$ in the presence of α -stable noise, the residual $imp(n)$ dominates in the estimation error $e(n)$ in Eq. (8). Therefore, the resulting estimation error may be assumed as approximately α -stable process, such that the FLOM [22] is

$$E \left[|e(n)|^{p-2} \right] = C(p-2, \alpha)\gamma^{((p-2)/\alpha)} = D(p, \alpha, \gamma) \quad (12)$$

Since the variance of α -stable noise is not finite, therefore, we can utilize the MED criterion [24], i.e., the minimization of cost function

$$J_{\text{MED}}(\vec{h}') = E \left[\left| y(n) - \vec{h}'^T(n) \vec{x}(n) \right|^p \right] \tag{13}$$

It is equivalent to the minimization of p th order FLOM. Unfortunately, this cost function $J_{\text{MED}}(\vec{h}')$ does not exhibit closed-form solution. Therefore, the stochastic gradient technique can be utilized as an alternative for the minimization of $J_{\text{MED}}(\vec{h}')$, similar to the LMS adaptive algorithm. The basic idea is to minimize the error dispersion for each successive datum or observation as much as possible. It leads to

$$\vec{h}'(n + 1) = \vec{h}'(n) - \mu(n) \nabla_h J_{\text{MD}}(\vec{h}') \tag{14}$$

Akin to the steepest descent algorithm [7],

$$\nabla_h J_{\text{MD}}(\vec{h}') = \frac{\delta E \left[\left| y(n) - \vec{h}'^T(n) \vec{x}(n) \right|^p \right]}{\delta \vec{h}'} \tag{15}$$

Analogous to the stochastic gradient algorithm [22],

$$\nabla_h J_{\text{MD}}(\vec{h}') \approx \frac{\delta \left[|e(n)|^p \right]}{\delta \vec{h}'} = p |e(n)|^{p-1} \frac{\delta |e(n)|}{\delta \vec{h}'} \tag{16}$$

where,
$$\begin{aligned} |e(n)| &= \text{sgn} \{e(n)\} e(n) \\ &= \text{sgn} \{e(n)\} \left\{ y(n) - \vec{h}'^T(n) \vec{x}(n) \right\} \end{aligned} \tag{17}$$

Hence,
$$\nabla_h J_{\text{MD}}(\vec{h}') \approx -p |e(n)|^{p-1} \text{sgn} \{e(n)\} \vec{x}(n) \tag{18}$$

By substituting (18) in Eq. (14), it can be shown that

$$\vec{h}'(n + 1) = \vec{h}'(n) + \{\mu(n)p\} |e(n)|^{p-1} \text{sgn} \{e(n)\} \vec{x}(n) \tag{19}$$

The simplified version of LMP algorithm can be represented as

$$\vec{h}'(n + 1) = \vec{h}'(n) + \mu'(n) |e(n)|^{p-1} \text{sgn} \{e(n)\} \vec{x}(n), \tag{20}$$

where $\mu'(n) = \mu(n)p$ is the variable step-size (VSS), which plays a critical role in the convergence mode of LMP algorithm [28], for the operating range $1 < p < \alpha < 2$. Although the convergence analysis of LMP algorithm is a tedious problem, yet the convergence range of variable step-size $\mu'(n)$ in (20) can be approximated as

$$0 < \mu'(n) < \frac{2}{D(p, \alpha, \gamma)\lambda_{\text{Max}}} \text{ (loose bounded, [28])} \tag{21}$$

By invoking a better approximation, it can be shown that

$$0 < \mu'(n) < \frac{2}{D(p, \alpha, \gamma)tr(\bar{R}_{xx})} < \frac{2}{3tr(\bar{R}_{xx})} \text{ (tight bounded, [23, 28]),} \quad (22)$$

where $tr(\bar{R}_{xx})$ symbolizes the trace of autocorrelation matrix \bar{R}_{xx} of the input signals. The maximum value of VSS is tightly bounded by the maximum eigenvalue λ_{Max} of the matrix $\bar{R}_{xx} = E[\bar{x}(n)\bar{x}^T(n)]$. The LMS algorithm is a special case of LMP algorithm for $p = 2$ and $\alpha = 2$ in Eq. (20). In the next subsection, we give details about the proposed GVSS criterion to update $\mu'(n)$ in the aforementioned iterative procedure.

3.2 Generalized Variable Step-Size (GVSS) Criterion

The large eigenvalue spread in the case of Volterra filtering necessitates the incorporation of variable step-size, in combination with the LMP adaptive algorithm, for the improved convergence rate. Moreover, the VSS criterion is also beneficial in the tracking of slowly time-varying channels/systems. The VSS must increase or decrease as the mean square error increases or decreases, allowing the adaptive nonlinear filter to track changes in the underlying system and to produce a small steady-state error. It should reduce the tradeoff between misadjustment and the speed of adaptation under the slowly time-varying conditions, due to its innate capability of providing both fast tracking, as well as small misadjustment. Therefore, the generalized variable step-size (GVSS) criterion is proposed to adjust the step-size under the stationary and nonstationary scenarios, which is as follows:

$$\mu'(n) = \bar{\alpha}\mu'(n-1) + \bar{\gamma}J_1(n-1) + \bar{\beta}J_2(n) \quad (23)$$

$$\text{where, } J_1(n) = \sum_{\bar{p}=0}^{\bar{P}} \lambda_1^{\bar{p}} e(n)e(n-\bar{p}) \text{ with } 0 \leq \lambda_1 < 1 \quad (24)$$

$$J_2(n) = \left[\sum_{\bar{q}=1}^{\bar{Q}} \lambda_2^{\bar{q}} e(n-\bar{q}) \bar{x}^T(n-\bar{q}) \right] \bar{x}(n)e(n) \text{ with } 0 \leq \lambda_2 < 1 \quad (25)$$

where $0 < \bar{\alpha} \leq 1$, $0 \leq \bar{\gamma} < 1$, and $0 \leq \bar{\beta} < 1$. The parameter $\bar{\alpha}$ induces the global exponential forgetting to the VSS, the parameter $\bar{\gamma}$ controls the convergence time, as well as the level of misadjustment [10], the parameter $\bar{\beta}$ adjusts the adaptive behavior of the step-size sequence $\mu'(n)$ [13]. However, λ_1 and λ_2 are the local exponential forgetting factors in Eqs. (24) and (25), respectively. For the appropriate convergence, the VSS should be bounded in the range $\mu'_{Min} \leq \mu'(n) \leq \mu'_{Max}$ [1]. The initial step-size is usually taken as μ'_{Max} , which ensures that the MED of algorithm remains bounded. However, μ'_{Min} is chosen to provide the minimum level of tracking ability, which is kept close to the step-size of FSS-LMS algorithm.

Further in special case 1, if the values of parameter are $\bar{P} = 0$ and $\bar{\beta} = 0$ in Eq. (23), then

$$\mu'(n) = \bar{\alpha}\mu'(n - 1) + \bar{\gamma}J_1(n - 1) \tag{26}$$

$$\mu'(n) = \bar{\alpha}\mu'(n - 1) + \bar{\gamma}\underline{e^2(n - 1)} \tag{27}$$

The underlined term in Eq. (27) is similar to the KVSS criterion presented in [10], (as in Eq. (33) of appendix). Next in special case 2, if the parametric values $\bar{P} = 1$ and $\bar{\beta} = 0$ are in Eq. (23), then

$$\mu'(n) = \bar{\alpha}\mu'(n - 1) + \bar{\gamma}e^2(n - 1) + \{\bar{\gamma}\lambda_1\} \underline{\underline{\{e(n - 1)e(n - 2)\}}} \tag{28}$$

The underlined term in Eq. (28) is akin to the AVSS criterion described in [1], (as in Eq. (35) of appendix). Subsequently in special case 3, if $\bar{\alpha} = 1$, $\bar{Q} = 1$, and $\bar{\gamma} = 0$ are in Eq. (23), then

$$\mu'(n) = \mu'(n - 1) + \{\bar{\beta}\lambda_2\} \underline{\underline{\{e(n - 1)\bar{x}^T(n - 1)\bar{x}(n)e(n)\}}} \tag{29}$$

The underlined term in Eq. (29) is analogous to the Mathews’ algorithm proposed in [13]. However, in special case 4, $\bar{\alpha} = 1$ and $\bar{\gamma} = 0$ in Eq. (23) results in

$$\mu'(n) = \mu'(n - 1) + \{\bar{\beta}\lambda_2\} \underline{\underline{\left[\begin{array}{l} e(n - 1)\bar{x}^T(n - 1) + \lambda_2 e(n - 2)\bar{x}^T(n - 2) \\ + \lambda_2^2 e(n - 3)\bar{x}^T(n - 3) + \dots \\ \dots + \lambda_2^{\bar{Q}-1} e(n - \bar{Q})\bar{x}^T(n - \bar{Q}) \end{array} \right] \bar{x}(n)e(n)}} \tag{30}$$

The underlined term in Eq. (30) is similar to the SVSS criterion suggested in [2], (as shown in Eq. (39) of appendix). Therefore, the abovementioned GVSS criterion (23) is incorporated in Eq. (20) to formulate the proposed GVSS-LMP algorithm, which is relatively computationally complex than FSS-LMP, KVSS-LMP, AVSS-LMP, and SVSS-LMS algorithms.

4 Simulation Results

The performance evaluation of the proposed GVSS-LMP algorithm is performed by comparing it with KVSS-LMP, AVSS-LMP, and SVSS-LMS algorithm under the similar conditions, for the nonlinear system identification. The kernels of the unknown system (as shown in Fig. 1) are assumed to follow the Random Walk model for the slow time-variations in the system response (as discussed in Sect. 2.1). As in α -stable noisy environment, the error signal variance could be infinite, therefore, the LMS algorithm based on the MMSE criterion (11) seems to be an inappropriate choice in comparison with the MED criterion $J_{MED}(h')$ (12). However, the value of p in LMP algorithm (20) is kept close to α for excellent results [10] in terms of the transient and steady-state behavior, which are fixed at $\alpha = 1.75$ and $p = 1.6$.

The input signal to the underlying unknown system (as shown in Fig. 1) may be correlated or uncorrelated Gaussian sequence \bar{x} . The white Gaussian input is quite

apposite for the identification of kernels in the Volterra system because it has adequate spectral representatives and sufficient amplitude variations [6]. Moreover, the Volterra series can be expressed as G-functionals [12], which form an orthogonal set when the input is white Gaussian. However, the non-identical-independent-distributed input signals lead to the large eigenvalue spread of the autocorrelation matrix \vec{R}_{xx} (particularly in the case of Volterra filters), which in turn results in the slow convergence [7]. The value of minimum step-size is $\mu'_{\text{Min}} = 0.0008$ and the maximum bounded value of step-size μ'_{Max} is set by the Eq. (22). The signal-to-noise-ratio (SNR) is defined as the input signal variance to the dispersion of the α -stable noise, i.e., $\text{SNR} = \sigma_x^2/\gamma$, which is kept 15 dB for all the simulations [28]. The Volterra kernel mean square estimation error (performance appraisal factor) is calculated by the following formula

$$\hat{J}(n) = E \left[|h(n) - h'(n)|^2 \right] \quad (31)$$

As per the Monte-Carlo simulations, the performance of adaptive algorithms is compared on the basis of measured performance appraisal factor as

$$\hat{J}(n) = \sum_{j=1}^{2500} \left[\frac{|h(n, j) - h'(n, j)|^2}{2500} \right] \quad (32)$$

Example 1 We consider the second-order Volterra filter with $K = 2$ and $M = 3$ in the first simulation setup, with the uncorrelated Gaussian white input sequence. Similar to the methodology opted in [1], the parameter values of the adaptive algorithms are selected to produce a comparable level of misadjustment. The values of parameters are $\lambda_1 = 0.8$, $\lambda_2 = 0.5$, $\bar{\alpha} = 0.97$, $\bar{\gamma} = 15 \times 10^{-5}$, $\bar{P} = \bar{Q} = 2$, $\bar{\alpha}_A = 0.97$, $\bar{\alpha}_W = 0.8$, $\bar{\rho}_W = 15 \times 10^{-5}$. The value of parameter $\bar{\beta}$ is varied as $\bar{\beta} = 0.00003$, GVSS-LMP1, $\bar{\beta} = 0.00005$, GVSS-LMP2, and $\bar{\beta} = 0.00015$, GVSS-LMP3. It is apparent from the simulation results depicted in Fig. 2 that the performance of GVSS-LMP algorithm improves as the value of $\bar{\beta}$ increases. The performance of $\bar{\beta} = 0.00015$, GVSS-LMP3 is approximately 7 dB better than AVSS-LMP algorithm in the tracking mode, and this proposed algorithm converges at the higher rate than other conventional algorithms.

Further, the value of $\bar{\beta} = 0.00005$ is fixed under the similar conditions. However, the values of \bar{P} and \bar{Q} are varied as $\bar{P} = \bar{Q} = 1$, $\bar{P} = \bar{Q} = 2$, $\bar{P} = \bar{Q} = 3$ in the GVSS-LMP algorithm. The simulation results demonstrated in Fig. 3 evidenced that $\bar{P} = \bar{Q} = 2$ is the suitable preference for GVSS-LMP algorithm, which also restricts its computational complexity.

Subsequently, the values of $\bar{\beta} = 0.00005$ and $\bar{P} = \bar{Q} = 2$ are fixed under the similar conditions. However, the values of $\lambda_2 = 0.3, 0.5, 0.7$ are varied in the proposed GVSS-LMS algorithm. It may be inferred from the simulation results in Fig. 4 that the performance of presented algorithm can be improved by increasing the value of λ_2 . However, for $\lambda_2 > 0.75$, the observed performance advantage is marginal.

Example 2 Now, we consider the third-order Volterra filter with $K = 3$ and $M = 3$ in this simulation setup with the uncorrelated Gaussian white input sequence. As

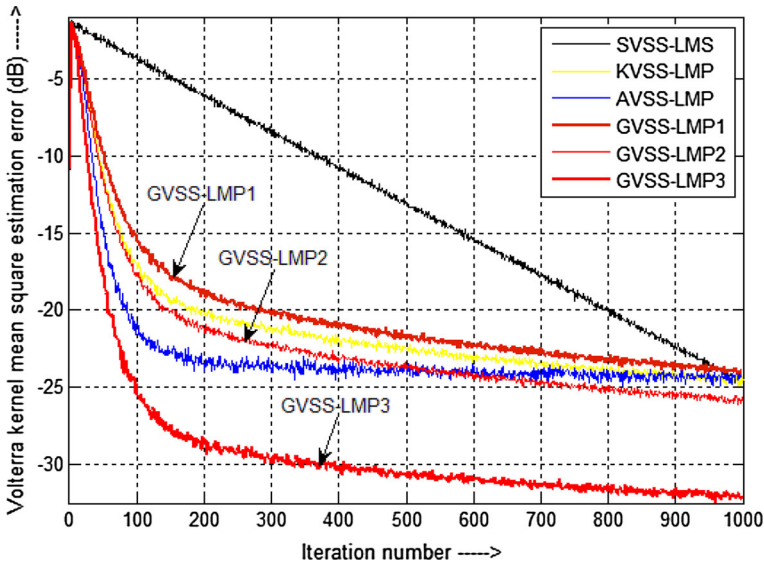


Fig. 2 Comparison of GVSS-LMP algorithm with conventional algorithms under varying value of $\bar{\beta}$ for the second-order Volterra filter

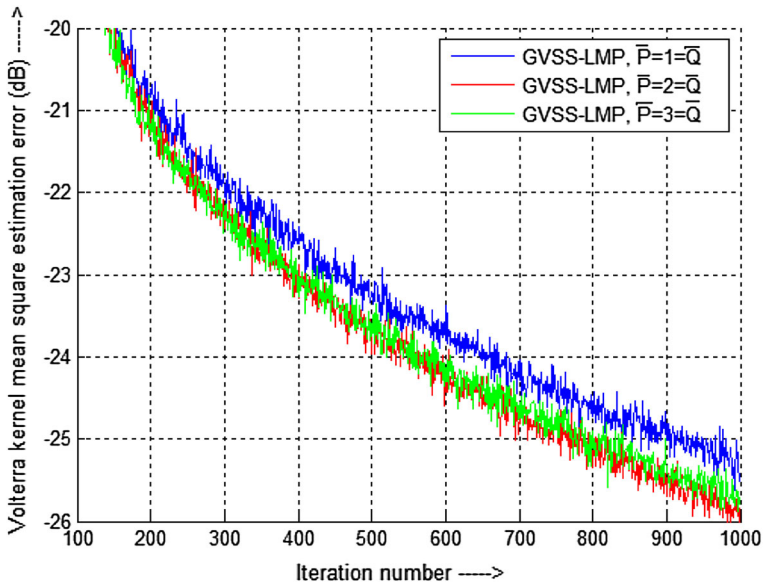


Fig. 3 Effects of the variation in the value of \bar{P} and \bar{Q} on GVSS-LMP algorithm

the number of filter weights increases in this case [1,7], the parameter values need to be changed to maintain the value of GVSS within limits. The values of parameters are $\lambda_1 = 0.98$, $\lambda_2 = 0.5$, $\bar{\alpha} = 0.91$, $\bar{\beta} = \bar{\gamma} = 0.000025$, $\bar{P} = \bar{Q} = 2$, $\bar{\alpha}_A = 0.98$, $\bar{\alpha}_W = 0.8$, $\bar{\rho}_W = 15 \times 10^{-9}$. The results in Fig. 5 manifest that the

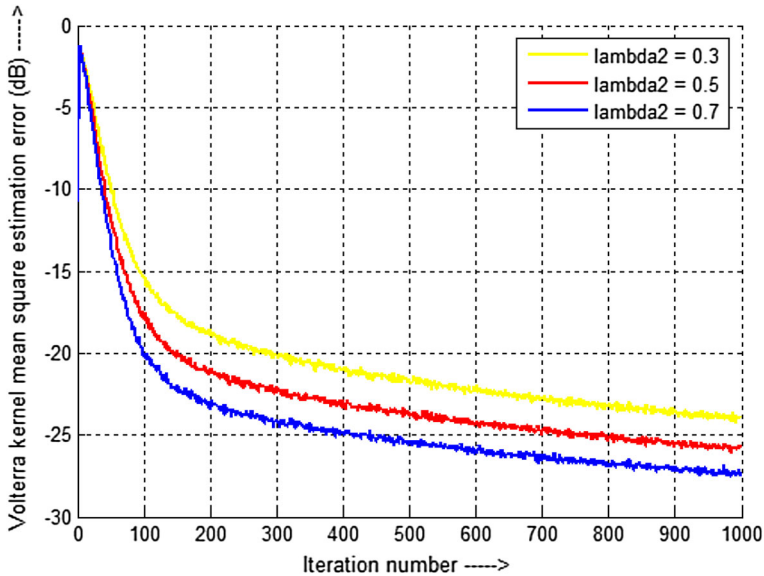


Fig. 4 Effects of variation in the value of λ_2 on GVSS-LMS algorithm.

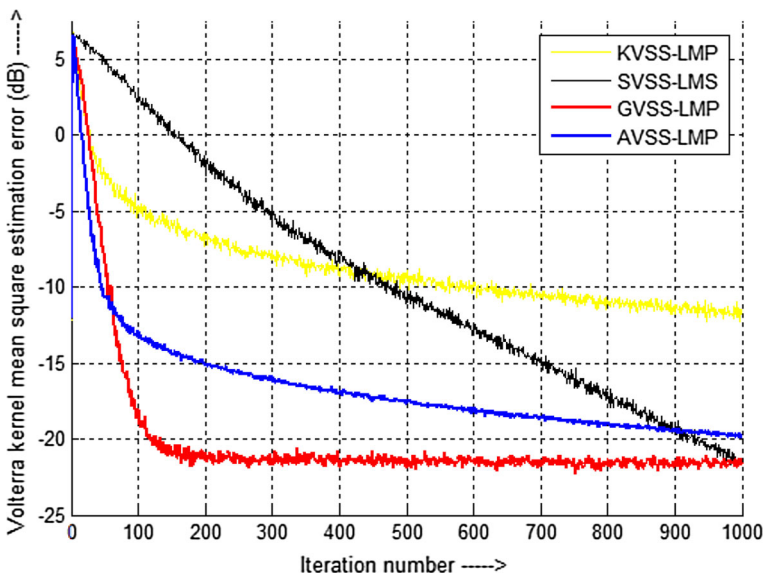


Fig. 5 Comparison of GVSS-LMP algorithm with conventional algorithms for the third-order Volterra filter

performance advantage of GVSS-LMP algorithm is approximately 3 dB better than the AVSS-LMP algorithm in the tracking mode. However, the convergence rate of both algorithms is approximately same in the initial phase. But, the significant performance degradation is observed in the case of KVSS-LMP algorithm.

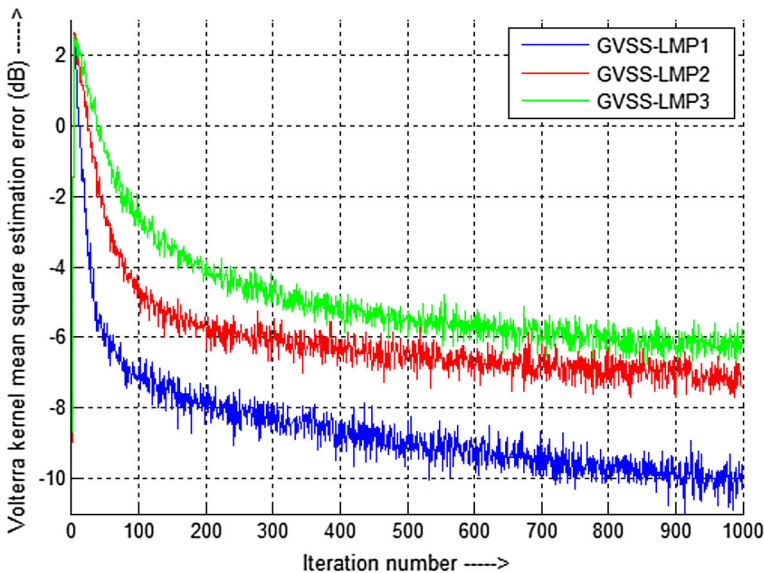


Fig. 6 Effects of variation in the value of $\tilde{\beta}$ on GVSS-LMP algorithm for the correlated input signal

Example 3 Next, we consider the second-order Volterra filter with $K = 2$ and $M = 3$ in this simulation setup, when the unknown system is excited by a correlated input signal as $\vec{x}(n) = 0.9\vec{x}(n-1) + \vec{v}_x(n)$; where $\vec{v}_x(n)$ is a zero-mean, uncorrelated Gaussian noise of unity variance. This type of input signals results in the flattened elliptical contours, which usually cause difficulties in the convergence of stochastic gradient adaptive algorithms. The values of parameters are $\lambda_1 = 0.8$, $\lambda_2 = 0.5$, $\bar{\alpha} = 0.97$, $\bar{\gamma} = 15 \times 10^{-6}$, $\bar{P} = \bar{Q} = 2$, $\bar{\alpha}_A = 0.97$, $\bar{\alpha}_W = 0.8$, $\bar{\rho}_W = 15 \times 10^{-5}$. The value of parameter $\tilde{\beta}$ is varied as $\tilde{\beta} = 15 \times 10^{-6}$, GVSS-LMP1, $\tilde{\beta} = 15 \times 10^{-7}$, GVSS-LMP2, and $\tilde{\beta} = 15 \times 10^{-8}$, GVSS-LMP3. It is observed from the results in Fig. 6 that the performance of GVSS-LMP algorithm improves as the value of $\tilde{\beta}$ increases, but the overall performance degradation is noticed for all the algorithms. We now fix the value of parameter $\tilde{\beta} = 15 \times 10^{-6}$ for the simulation results in Fig. 7, which indicate that the proposed GVSS-LMP algorithm still outperforms the conventional algorithms. The variable step-size controls the problem of eigenvalue spread, and consequently leads to the enhanced convergence rate in the presence of impulse noise and correlated input signal.

On contrary to the case of uncorrelated input signal, it may be inferred from the results presented in Fig. 6 and Fig. 7 that the gradient-misadjustment [30] is relatively more in the case of correlated input signal. However, the convergence of GVSS-LMP algorithm is strictly dependent on the appropriate parameter tuning/setting in (23), while keeping the value of GVSS below μ'_{Max} (22). Akin to the VSS-LMS algorithms [4, 23], the GVSS-LMP algorithm is found to be sensitive to noise disturbances in the low signal-to-noise-ratio (SNR) environment.

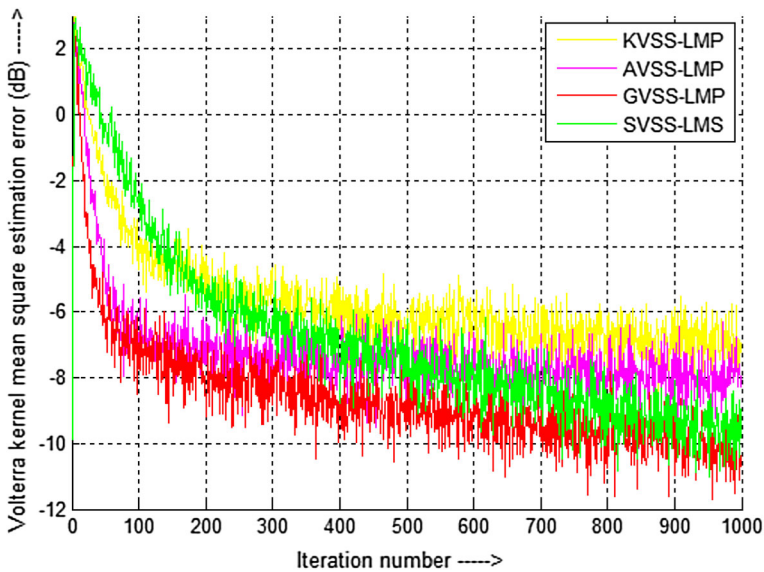


Fig. 7 Comparison of GVSS-LMP algorithm with conventional algorithms for the second-order Volterra filter with correlated input signal

5 Concluding Remarks

This paper presents a generalized variable step-size least mean p th power (LMP) adaptive algorithm for the α -stable noisy environment, which is based on the MED criterion. This algorithm is implemented to identify the unknown time-varying nonlinear systems using the Volterra filtering approach. However, the MMSE criterion is found to be a special case of MED approach. The GVSS-LMP algorithm exploits the knowledge about the previous step-size, the error autocorrelation values, the value of parameter α , the crosscorrelation between error sequence and input sequence. For excellent results, the value of parameter p is kept close to α in the range $1 < p < \alpha < 2$.

It is apparent from the simulation results that the GVSS-LMP algorithm supercedes the KVSS-LMP, AVSS-LMP, and SVSS-LMS algorithms in the convergence, as well as tracking mode, when the input signal is either correlated or uncorrelated Gaussian process. The proposed algorithm also controls the adverse effects of eigenvalue spread of the input signal autocorrelation matrix, by the GVSS criterion to track the time-varying Volterra kernels. The outperforming GVSS-LMP algorithm may find applications in the systems disturbed due to the presence of non-Gaussian impulsive measurement noise, where the conventional FSS-LMS algorithm fails to perform well. Moreover, the different LMP algorithms with $p \neq 2$ and LMS algorithms with $p = 2$ can be derived from the GVSS-LMP algorithm by adjusting the parameters according to the requirements. Future work includes the application of proposed adaptive nonlinear Volterra filtering technique in the emerging fields of bio-signal processing, biomedical engineering [29], nonlinearly amplified digital as well as analog communication signal processing [17] and equalization of nonlinear communication channels [18].

Appendix

The literature [8, 30] of fixed step-size LMS (FSS-LMS) algorithm reflects a tradeoff between the misadjustment and speed of adaptation, which depicts that a small step-size produces small misadjustment, but at the cost of longer convergence time. Under time-varying environment, the optimum value of the step-size in FSS-LMS algorithm strikes a balance between the amount of lag noise and gradient noise [5]. However, the optimum value of step-size can not be determined *a priori* due to the unknown channel parameters. Therefore, in KVSS-LMS algorithm [10], the variable step-size (VSS) is attuned using

$$\mu'(n) = \bar{\alpha}\mu'(n-1) + \bar{\gamma}\underline{e^2(n-1)} \quad (33)$$

In this KVSS-LMS algorithm, a large prediction error causes the step-size to increase in order to provide fast tracking, while a small prediction error leads to reduction in the step-size to yield small misadjustment. The step-size increases or decreases as the MSE increases or decreases, allowing the adaptive filter to track changes in the time-varying system, as well as to produce a small steady-state error. It also reduces sensitivity of the misadjustment to the level of nonstationarity. This approach is heuristically sound and has resulted in several ad hoc techniques, where the selection of convergence parameters is based on the magnitude of estimation error, polarity of the successive samples of the estimation error, measurement of the crosscorrelation of the estimation error with input data. However, the VSS-LMS algorithms are found to be sensitive to noise disturbances [4, 23] in the low signal-to-noise-ratio (SNR) environment because the step-size update of these algorithms are directly obtained from the instantaneous error that is contaminated by the disturbance noise.

Further in AVSS-LMS algorithm [1], the VSS is controlled using

$$\mu'(n) = \bar{\alpha}\mu'(n-1) + \bar{\gamma}\theta^2(n-1) \quad (34)$$

$$\theta(n-1) = \bar{\alpha}_A\theta(n-2) + (1 - \bar{\alpha}_A)\underline{\{e(n-1)e(n-2)\}} \quad (35)$$

Here, the error autocorrelation is usually a fine measure of the proximity to the optimum, which rejects the effect of uncorrelated noise sequence on the step-size update. In the early stages of adaptation, the error autocorrelation estimate is large, resulting in a large step-size. However, the small error autocorrelation leads to a small step-size under the optimum conditions. It results in effective adjustment of the step-size, while sustaining the immunity against independent noise disturbance, for the flexible control of misadjustment. The AVSS-LMS algorithm [1] shows substantial convergence rate improvement over the KVSS-LMS algorithm [10] and FSS-LMS algorithm [30] under the stationary environment for the low SNR, as well as the high SNR values. However, the performance of AVSS-LMS algorithm is comparable to the FSS-LMS and KVSS-LMS adaptive algorithms under the nonstationary conditions.

But in SVSS-LMS algorithm [2], the VSS is adjusted using the following recursive relation by adjusting the control parameters $\bar{\rho}_w$ and $\bar{\alpha}_w$.

$$\mu'(n) = \mu'(n-1) + \bar{\rho}_w \bar{\psi}^T(n) \bar{x}(n) e(n) \quad (36)$$

$$\bar{\psi}(n) = \bar{\alpha}_w \bar{\psi}(n-1) + e(n-1) \bar{x}(n-1) \quad (37)$$

The above equation can be rewritten in the expanded form as

$$\bar{\psi}^T\{n\} = \bar{\alpha}_w^{\bar{Q}} \bar{\psi}^T\{n - \bar{Q}\} + \bar{\alpha}_w^{\bar{Q}-1} e(n - \bar{Q}) \bar{x}^T(n - \bar{Q}) + \dots + \bar{\alpha}_w e(n - 2) \bar{x}^T(n - 2) + e(n - 1) \bar{x}^T(n - 1) \quad (38)$$

For $0 \leq \bar{\alpha}_w < 1$ and $\bar{Q} \rightarrow \text{high value}$, the Eq. (38) can be approximated as

$$\bar{\psi}^T\{n\} \approx \frac{\bar{\alpha}_w^{\bar{Q}-1} e(n - \bar{Q}) \bar{x}^T(n - \bar{Q}) + \dots + \bar{\alpha}_w^2 e(n - 3) \bar{x}^T(n - 3) + \bar{\alpha}_w e(n - 2) \bar{x}^T(n - 2) + e(n - 1) \bar{x}^T(n - 1)}{\dots} \quad (39)$$

This algorithm [2] outperforms the Mathews' algorithm [13], when both are set to track the random walk channel under the similar conditions.

References

1. T. Aboulnasr, K. Mayyas, A robust variable step-size LMS-type algorithm: analysis and simulations. *IEEE Trans. Signal Process.* **45**(3), 631–639 (1997)
2. W.P. Ang, B. Farhang-Boroujeny, A new class of gradient adaptive step-size LMS algorithms. *IEEE Trans. Signal Process.* **49**(4), 805–810 (2001)
3. A. Benveniste, Design of adaptive algorithms for the tracking of time varying systems. *Int. J. Adapt. Control Signal Process.* **1**(1), 3–29 (1987)
4. J.B. Evans, P. Xue, B. Liu, Analysis and implementation of variable step size adaptive algorithms. *IEEE Trans. Signal Process.* **41**(8), 2517–2535 (1993)
5. B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications* (Wiley, New York, 1998)
6. A. Feuer, E. Weinstein, Convergence analysis of LMS filters with uncorrelated Gaussian data. *IEEE Trans. Acoust. Speech Signal Process.* **33**(1), 222–230 (1985)
7. S. Haykin, *Adaptive Filter Theory*, 4th edn. (Prentice Hall, New Jersey, 2002)
8. A.K. Kohli, D.K. Mehra, Tracking of time-varying channels using two-step LMS-type adaptive algorithm. *IEEE Trans. Signal Process.* **54**(7), 2606–2615 (2006)
9. A.K. Kohli, A. Rai, Numeric variable forgetting factor RLS algorithm for second-order Volterra filtering. *Circuits Syst. Signal Process.* **32**(1), 223–232 (2013)
10. R.H. Kwong, E.W. Johnston, A variable step size LMS algorithm. *IEEE Trans. Signal Process.* **40**(7), 1633–1642 (1992)
11. V.J. Mathews, Adaptive polynomial filters. *IEEE Signal Process. Mag.* **8**(3), 10–26 (1991)
12. V.J. Mathews, G.L. Sicuranza, *Polynomial Signal Processing* (Wiley, New York, 2000)
13. V.J. Mathews, Z. Xie, Stochastic gradient adaptive filters with gradient adaptive step size. *IEEE Trans. Signal Process.* **41**(6), 2075–2087 (1993)
14. C.J. Masreliez, R.D. Martin, Robust Bayesian estimation for the linear model and robustifying the Kalman filter. *IEEE Trans. Autom. Control.* **22**(3), 361–372 (1977)
15. P. Mertz, Model of impulsive noise for data transmission. *IRE Trans. Commun. Syst.* **9**(2), 130–137 (1961)
16. R.D. Nowak, Nonlinear system identification. *Circuits Syst. Signal Process.* **21**(1), 109–122 (2002)

17. T. Ogunfunmi, *Adaptive Nonlinear System Identification: Volterra and Wiener Model Approaches* (Springer, Berlin, 2007)
18. T. Ogunfunmi, T. Drullinger, Equalization of non-linear channels using a Volterra-based non-linear adaptive filter, in *Proceedings of IEEE International Midwest Symposium on Circuits and Systems, Seoul, South Korea* (2011), pp. 1–4.
19. A. Papoulis, *Probability Random Variables and Stochastic Processes* (McGraw-Hill, New York, 1984)
20. G. Samorodnitsky, M.S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variances* (Chapman and Hall, New York, 1994)
21. M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems* (Wiley, New York, 1980)
22. M. Shao, C.L. Nikias, Signal processing with fractional lower order moments: stable processes and their applications. *Proc. IEEE*. **81**(7), 986–1010 (1993)
23. J. Soo, K.K. Pang, A multi step size (MSS) frequency domain adaptive filter. *IEEE Trans. Signal Process.* **39**(1), 115–121 (1991)
24. B.W. Stuck, Minimum error dispersion linear filtering of scalar symmetric stable processes. *IEEE Trans. Autom. Control.* **23**(3), 507–509 (1978)
25. I.J. Umoh, T. Ogunfunmi, An adaptive nonlinear filter for system identification. *EURASIP J. Adv. Signal Process.* **2009**, 1–7 (2009)
26. I.J. Umoh, T. Ogunfunmi, An affine projection-based algorithm for identification of nonlinear Hammerstein systems. *Signal Process.* **90**(6), 2020–2030 (2010)
27. E. Walach, B. Widrow, The least mean fourth (LMF) adaptive algorithm and its family. *IEEE Trans. Inf. Theory* **30**(2), 275–283 (1984)
28. B. Weng, K.E. Barner, *IEEE Trans. Signal Process.* **53**, 2588 (2005)
29. D.T. Westwick, R.E. Kearney, *Identification of Nonlinear Physiological Systems* (Wiley, New York, 2003)
30. B. Widrow, J.M. McCool, M.G. Larimore, C.R. Johnson, Stationary and nonstationary learning characteristics of LMS adaptive filter. *Proc. IEEE*. **64**(8), 1151–1162 (1976)