International Journal of Innovative Computing, Information and Control
Volume **15**, Number **2**, April **2019** ISSN 1349-4198
pp. **617–628**

RECURSIVE IDENTIFICATION ALGORITHM BASED ON COSINE BASIS FOR RAPID TIME-VARYING SYSTEMS

Xinyu Liu and Lianming Sun

Faculty of Environmental Engineering
The University of Kitakyushu
1-1 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan
z8dcb001@eng.kitakyu-u.ac.jp; sun@kitakyu-u.ac.jp

Received June 2018; revised October 2018

Abstract. *A new recursive identification algorithm using the cosine basis functions has been developed for the time-varying systems. It is demonstrated that through expanding the time-varying parameters within a sliding time window into the virtual even periodic functions, the parameters in the window can explicitly be approximated by the cosine series. Following the orthogonality of the trigonometric functions, the parameter estimation and the efficient update of the data matrices are recursively implemented to decrease the computational complexity. Moreover, a smoothing technique is also considered to reduce the estimation error caused by the noise effects. In contrast with the standard RLS, LMS/NLMS, AP/BOP and the conventional methods based on the basis functions, the proposed algorithm has higher tracking performance even for the high series degree, and can reduce the distortion of the Gibbs effect at the window edges.*
Keywords: System identification, Recursive algorithm, Time-varying, Basis function

1. **Introduction.** Due to component aging, fast variation of the environment, the dynamic characteristics of a physical process vary with time. When the process varies slowly, some existing adaptive algorithms, such as the segmentation approach to separate a time-varying model into several local models [1, 2], the recursive least squares (RLS) with a forgetting factor [3], the least mean square (LMS) or the normalized least mean square (NLMS) algorithms [4, 5], the affine projection algorithm (AP) and block orthogonal projection (BOP) [6], may track the varying dynamics.

Identification of time-varying systems has been successfully applied in many practical applications. For example, the identification of the time-varying system is used to address the terminal control problem [7] in computer engineering and robotic manipulator, the adaptive equalization of rapidly fading communication channels for non-stationary signals [8], the linear parameter varying model in transportation systems such as flight projectile and car steering [9, 10], the identification-based fault diagnosis [11] and time-varying model for effective treatments for certain brain diseases in biomedical engineering [12].

Nevertheless, if the variation is too fast, most of the existing algorithms fail to follow the variation satisfactorily, unless the prior information of the variation is available [13]. For the systems with less prior information, several methods use an explicit approximation of the parameter variation through some orthogonal basis functions such as the trigonometric or Legendre basis. They help to approximate the dynamics at an arbitrary rate when the signals have sufficient excitations, so that they have better tracking performance than other existing methods. Since the trigonometric functions have differentiability and computational stability, the varying parameters are often approximated by the Fourier

DOI: 10.24507/ijicic.15.02.617

series in the conventional basis function methods, where the parameters within a chosen time window are expanded into the virtual periodic functions whose right and left window edges correspond to the newest, the oldest information on the process variation, respectively. Nevertheless, the values of the virtual periodic functions are generally different at the right and left edges. As a result, the discontinuity causes the Gibbs effect in the series expansion and severely degrades the tracking performance at the window edges [14]. On the other hand, when expanding the varying parameters into the virtual even periodic functions that are approximated by the cosine series [15], the virtual functions are continuous at the window edges when the true parameters vary continuously.

On the other hand, it is expected to implement the algorithms recursively for the rapid time-varying systems in many applications, where the recursive computability of the basis functions is an essential issue in the recursive algorithms. Nevertheless, the recursion of cosine series is more complicated than the standard Fourier series when updating the data matrices, and their inversion, recursively, so most of the conventional cosine basis based algorithms work in batches rather than the recursive processing. A recursive identification algorithm with a forgetting factor for time-varying systems is proposed [16]. The forgetting factor makes the time window shift easily since the past data beyond the window decay to zero, and has high tracking performance in rapid varying system. However, it will weaken the orthogonality of the basis functions; as a result, the data matrix often has such large condition number that degrades the tracking performance for the basis series with high degree.

In order to improve the tracking performance, a new recursive identification algorithm based on the trigonometric functions is investigated for the rapid time-varying systems in this paper. In the proposed algorithm, the orthogonality of the trigonometric functions is applied to effectively implementing the recursion of the cosine basis, and some efficient approximation is used to decrease the computational complexity. Moreover, a smoothing technique is also considered to reduce the influence of the noise term. In contrast with the conventional methods based on the Fourier series expansion, the new one holds the orthogonality for the high basis degree, and has less Gibbs effect at the window edges. Consequently, the proposed algorithm has high tracking performance even in the rapid varying processes.

The rest of the paper is organized as follows. In the next section, the main description of the problem considered in this paper is summarized. In Section 3, the preliminaries of the proposed algorithm are illustrated, and then the new recursive identification algorithm is shown in Section 4. Section 5 demonstrates some numerical simulation examples to show the effectiveness of the algorithm. Finally, the conclusion and the future research work are given in Section 6.

2. **Problem Statement.** In many dynamic systems, the characteristics of a physical process are described by impulse response. When the impulse response with a finite length dominates the process characteristics, an FIR model given in (1) can approximate the relation of the process input-output signals:

$$y(k) = h_0^k u(k) + h_1^k u(k-1) + \cdots + h_n^k u(k-n) + e(k) = \boldsymbol{\phi}^T(k)\boldsymbol{h}(k) + e(k), \tag{1}$$

where n is the model order, $u(k)$, $y(k)$ and $e(k)$ are the input, output and noise, respectively, while the superscript and subscript of the parameter h_i^k indicate the lag time i, the normalized sampling instant k, respectively. $\boldsymbol{\phi}(k) = [u(k), \ldots, u(k-n)]^T$, $\boldsymbol{h} = [h_0^k, \ldots, h_n^k]^T$ are the regression and parameter vectors. $u(k)$ is assumed as a wide-sense stationary signal with persistent excitation for system identification, and is independent of the noise $e(k)$.

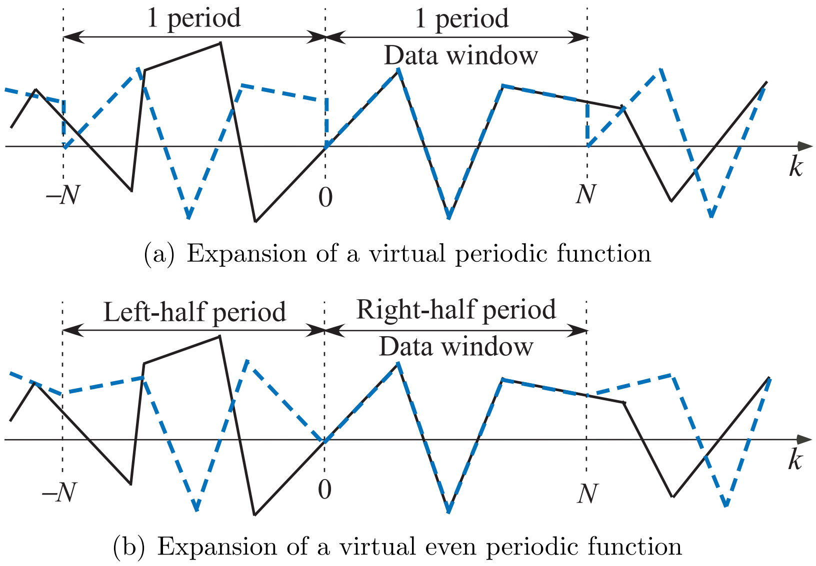
The variation rate of parameters is measured by

$$\kappa(k) = \sqrt{\frac{\mathrm{tr}\big\{E\{\boldsymbol{\phi}(k)\}\mathrm{cov}\{\boldsymbol{h}(k) - \boldsymbol{h}(k-1)\}\big\}}{\sigma_e^2}}, \tag{2}$$

where σ_e^2 is the variance of the noise $e(k)$. $E\{\boldsymbol{\phi}(k)\}/\sigma_e^2$ is related to the stationary part, and $\mathrm{cov}\{\boldsymbol{h}(k) - \boldsymbol{h}(k-1)\}$ corresponds to the varying part. Empirically, the variation is indicated as a fast one when $\kappa(k) > 0.0005\sqrt{n}$, and the conventional methods may fail to sufficiently track the fast variation.

In the conventional method based on the Fourier series, the parameters h_i^k within a window $[0, N]$ are expanded into the virtual periodic functions, as shown in Figure 1(a) where the period is the same as the window width N. It is noted that the discontinuity at the window edges 0, $\pm N$, $\pm 2N$, $\ldots$ yields distortion of the prediction error caused by Gibbs effect.

In the new algorithm, the even periodic expansion is used, and thus the virtual functions are continuous at the window edges. It implies that the parameter functions within the window $[0, N]$ are regarded as the right-half period of the expanded virtual even functions, as illustrated in Figure 1(b). Therefore, the virtual periodic functions with period $2N$ can be approximated by a finite degree of the cosine series representation with a small bias.



(b) Expansion of a virtual even periodic function

FIGURE 1. Illustration of the virtual expansion of time-varying parameters. Solid line: true parameters; dashed line: virtual expansion.

At the newest sampling instant k, the sliding time window is $[k_0, k]$, which is the right-half period of the expanded virtual even function. Following the convergence theorem of the Fourier series, if the parameters $h_i^{k_0+k_1}$ have at most a finite number of discontinuous points for $0 \le k_1 \le N$, the virtual even periodic functions can be approximated by the cosine series

$$h_i^{k_0+k_1} = h_i^{k-N+1+k_1} \triangleq h_i^k(k_1) \sim \sum_{m=0}^{M} c_{i,m}^k \cos\big(m\omega k_1\big), \tag{3}$$

where $\omega = \frac{\pi}{N}$, M is the degree of the series.

3. **Preliminaries.** Within the window $[k_0, k]$, the input-output relation can be written as follows:

$$\begin{aligned} y(k_0+k_1) &= y(k-N+k_1) \triangleq y^k(k_1) \\ &\approx \sum_{m=0}^{M} c_{0,m}^k \cos\big(m\omega k_1\big) u(k_0+k_1) + \sum_{m=0}^{M} c_{1,m}^k \cos\big(m\omega k_1\big) u(k_0+k_1-1) \\ &\quad + \cdots + \sum_{m=0}^{M} c_{n,m}^k \cos\big(m\omega k_1\big) u(k_0+k_1-n) + e(k_0+k_1) \\ &= \left(\boldsymbol{\phi}_{c,0}^k(k_1)\right)^T \boldsymbol{\theta}_{c,0}^k + \left(\boldsymbol{\phi}_{c,1}^k(k_1)\right)^T \boldsymbol{\theta}_{c,1}^k + \cdots + \left(\boldsymbol{\phi}_{c,M}^k(k_1)\right)^T \boldsymbol{\theta}_{c,M}^k + e(k_0+k_1) \\ &= \left(\boldsymbol{\phi}_{c}^k(k_1)\right)^T \boldsymbol{\theta}_{c}^k + e(k_0+k_1), \end{aligned} \tag{4}$$

where the main regressions are as follows:

$$\begin{aligned} &\boldsymbol{\phi}_c^k(k_1) = \begin{bmatrix} \boldsymbol{\phi}_{c,0}^k(k_1) \\ \boldsymbol{\phi}_{c,1}^k(k_1) \\ \vdots \\ \boldsymbol{\phi}_{c,M}^k(k_1) \end{bmatrix}, \ \boldsymbol{\phi}_s^k(k_1) = \begin{bmatrix} \boldsymbol{\phi}_{s,1}^k(k_1) \\ \vdots \\ \boldsymbol{\phi}_{s,M}^k(k_1) \end{bmatrix}, \\ &\boldsymbol{\phi}_{c,0}^k(k_1) = [u(k_0+k_1), u(k_0+k_1-1), \ldots, u(k_0+k_1-n)]^T, \\ &\boldsymbol{\phi}_{c,m}^k(k_1) = \boldsymbol{W}_{c,m}^{k_1} \boldsymbol{\phi}_{c,0}^k(k_1), \ \boldsymbol{\phi}_{s,m}^k(k_1) = \boldsymbol{W}_{s,m}^{k_1} \boldsymbol{\phi}_{c,0}^k(k_1), \\ &\boldsymbol{W}_{c,m}^{k_1} = \cos(m\omega k_1)\boldsymbol{I}, \ \boldsymbol{\phi}_{s,m}^k(k_1) = \sin(m\omega k_1)\boldsymbol{I}, \end{aligned} \tag{5}$$

while the coefficient vectors are

$$\boldsymbol{\theta}_c^k = \left[\left(\boldsymbol{\theta}_{c,0}^k\right)^T, \left(\boldsymbol{\theta}_{c,1}^k\right)^T, \ldots, \left(\boldsymbol{\theta}_{c,M}^k\right)^T\right]^T, \quad \boldsymbol{\theta}_{c,m}^k = \left[c_{0,m}^k, c_{1,m}^k, \ldots, c_{n,m}^k\right]^T, \tag{6}$$

where $\boldsymbol{I}$ is an identity matrix with the appropriate dimension, $\boldsymbol{W}_{c,m}^{k_1}$ and $\boldsymbol{W}_{s,m}^{k_1}$ are the diagonal matrices whose diagonal elements are $\cos(\omega m k_1)$, $\sin(\omega m k_1)$, respectively.

The approximation in (3) and (4) implies that the variation of model parameters can be approximated by the linear combination of the known basis functions, whereas the estimation of coefficients $\boldsymbol{\theta}_c^k$ becomes an important issue. Therefore, the parameter estimation problem of a time-varying model can be realized by estimating the coefficient vector $\boldsymbol{\theta}_c^k$ instead of the direct estimation of the varying model parameters. $\boldsymbol{\theta}_c^k$ can be estimated by minimizing the criterion function

$$\hat{\boldsymbol{\theta}}_c^k = \arg\min_{\boldsymbol{\theta}_c^k} \sum_{k_1=0}^{N} \left(y^k(k_1) - \left(\boldsymbol{\phi}_c^k(k_1)\right)^T \boldsymbol{\theta}_c^k\right)^2. \tag{7}$$

The minimization problem can be solved by some optimization algorithms such as the least squares algorithm

$$\hat{\boldsymbol{\theta}}_c^k = \left(\sum_{k_1=0}^{N} \boldsymbol{\phi}_c^k(k_1) \left(\boldsymbol{\phi}_c^k(k_1)\right)^T\right)^{-1} \left(\sum_{k_1=0}^{N} \boldsymbol{\phi}_c^k(k_1) y^k(k_1)\right) = \left(\boldsymbol{\Phi}_{cc}^k\right)^{-1} \boldsymbol{\phi}_{cy}^k \triangleq \boldsymbol{P}^k \boldsymbol{\phi}_{cy}^k, \tag{8}$$

where $\boldsymbol{\Phi}_{cc}^k$ and $\boldsymbol{P}^k$ are the correlation matrix and its inverse, $\boldsymbol{\phi}_{cy}^k$ is the correlation vector of the regression and the process output in the window $[k_0, k]$. It indicates that (8) turns the estimation of varying parameter into the estimation problem of the series coefficients.

The window shifts forward to $[k_0+1, k+1]$ at the next instant $k+1$. Correspondingly, the updated regression in the new time window is

$$\boldsymbol{\phi}_{c,0}^{k+1}(k_1) = \left[u(k_0+k_1+1), u(k_0+k_1), \ldots, u(k_0+k_1+1-n)\right]^T = \boldsymbol{\phi}_{c,0}^k(k_1+1), \tag{9}$$

and following $\boldsymbol{W}_{c,m}^{k_1} = \boldsymbol{W}_{c,m}^{1}\boldsymbol{W}_{c,m}^{k_1+1} + \boldsymbol{W}_{s,m}^{1}\boldsymbol{W}_{s,m}^{k_1+1}$, $\boldsymbol{\phi}_{c,m}^{k+1}(k_1)$ can be expressed by

$$\boldsymbol{\phi}_{c,m}^{k+1}(k_1) = \boldsymbol{W}_{c,m}^{k_1}\boldsymbol{\phi}_{c,0}^{k+1}(k_1) = \begin{bmatrix} \boldsymbol{W}_{c,m}^{1} & \boldsymbol{W}_{s,m}^{1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}_{c,m}^{k}(k_1+1) \\ \boldsymbol{\phi}_{s,m}^{k}(k_1+1) \end{bmatrix}. \tag{10}$$

It illustrates that besides the time-shift term $\boldsymbol{W}_{c,m}^{1}\boldsymbol{\phi}_{c,m}^{k}(k_1+1)$, an extra term $\boldsymbol{W}_{s,m}^{1}\boldsymbol{\phi}_{s,m}^{k}$ (k_1+1) also appears in (10). Then, the correlation matrix $\boldsymbol{\Phi}_{cc}^{k+1}$ is updated by

$$\begin{aligned} \boldsymbol{\Phi}_{cc}^{k+1} &= \sum_{k_1=0}^{N} \boldsymbol{\phi}_{c}^{k+1}(k_1)\left(\boldsymbol{\phi}_{c}^{k+1}(k_1)\right)^T \\ &= \begin{bmatrix} \boldsymbol{W}_c & \boldsymbol{W}_s \end{bmatrix} \left(\begin{bmatrix} \boldsymbol{\phi}_{c}^{k}(N+1) \\ \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}_{c}^{k}(N+1) & \boldsymbol{0} \end{bmatrix} \right. \\ &\quad \left. - \begin{bmatrix} \boldsymbol{\phi}_{c}^{k}(0) \\ \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\phi}_{c}^{k}(0) & \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\Phi}_{cc}^{k} & \boldsymbol{\Phi}_{cs}^{k} \\ \boldsymbol{\Phi}_{sc}^{k} & \boldsymbol{\Phi}_{ss}^{k} \end{bmatrix} \right) \begin{bmatrix} \boldsymbol{W}_c \\ \boldsymbol{W}_s \end{bmatrix}, \end{aligned} \tag{11}$$

where $\boldsymbol{W}_c$ and $\boldsymbol{W}_s$ are the diagonal matrices with the diagonal blocks $\boldsymbol{W}_{c,m}^{1}$, $\boldsymbol{W}_{s,m}^{1}$, $m = 0, 1, \ldots, M$, respectively.

From (11), it is deduced that $\boldsymbol{\Phi}_{cc}^{k+1}$ consists of the innovative term, the term beyond the new window and the transition term. Let the inverse of $\boldsymbol{\Phi}_{cc}^{k+1}$ be denoted as $\boldsymbol{P}^{k+1}$, the coefficient vector can be obtained by $\hat{\boldsymbol{\theta}}_{c}^{k+1} = \boldsymbol{P}^{k+1}\boldsymbol{\phi}_{cy}^{k+1}$, where the computation of $\boldsymbol{P}^{k+1}$ suffers from the heavy computational load due to extra terms such as $\boldsymbol{\Phi}_{cs}^{k}$, $\boldsymbol{\Phi}_{sc}^{k}$ and $\boldsymbol{\Phi}_{ss}^{k}$ in (11). Therefore, a novel recursive algorithm is proposed to solve this problem.

4. New Recursive Identification Algorithm.

The mild assumptions in the proposed algorithm are summarized as follows: the input signal $u(k)$ has quasi-stationarity and ergodicity, and is independent of the noise $e(k)$; the window width $N > (n+1)(M+1)$; the varying parameters have at most a finite number of discontinuous points in the window. Then, following the expansion theorem of cosine series, the approximation in (3) is guaranteed for the parameter estimation if these assumptions hold.

4.1. Properties of data matrices.

In order to illustrate the update of the matrices in the recursive algorithm, their properties are investigated. Let the matrices $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$ be denoted as

$$\begin{aligned} &\boldsymbol{\Phi}_1 = \boldsymbol{\phi}_{c}^{k}(N+1)\left(\boldsymbol{\phi}_{c}^{k}(N+1)\right)^T - \boldsymbol{\phi}_{c}^{k}(0)\left(\boldsymbol{\phi}_{c}^{k}(0)\right)^T + \boldsymbol{\Phi}_{cc}^{k} = \boldsymbol{\Psi}^{k+1} + \boldsymbol{\Phi}_{cc}^{k}, \\ &\boldsymbol{\Phi}_2 = \boldsymbol{W}_{cs}^{-1}\boldsymbol{\Phi}_{sc}^{k} + \boldsymbol{\Phi}_{cs}^{k}\boldsymbol{W}_{cs}^{-1} + \boldsymbol{W}_{cs}^{-1}\boldsymbol{\Phi}_{ss}^{k}\boldsymbol{W}_{cs}^{-1}, \\ &\boldsymbol{\psi}^{k+1} = \begin{bmatrix} \boldsymbol{\phi}_{c}^{k}(N+1), & \boldsymbol{\phi}_{c}^{k}(0) \end{bmatrix}, \ \bar{\boldsymbol{\psi}}^{k+1} = \begin{bmatrix} \boldsymbol{\phi}_{c}^{k}(N+1), & -\boldsymbol{\phi}_{c}^{k}(0) \end{bmatrix}, \\ &\bar{\boldsymbol{y}}^{k+1} = \begin{bmatrix} y^k(N+1), & -y^k(0) \end{bmatrix}^T, \end{aligned} \tag{12}$$

where $\boldsymbol{W}_{cs}^{-1} = \boldsymbol{W}_{c}^{-1}\boldsymbol{W}_s$, then $\boldsymbol{\Phi}_{cc}^{k+1}$ in (11) can be expressed by

$$\boldsymbol{\Phi}_{cc}^{k+1} = \boldsymbol{W}_c\big(\boldsymbol{\Phi}_1 + \boldsymbol{\Phi}_2\big)\boldsymbol{W}_c. \tag{13}$$

It is seen that the extra matrices in $\boldsymbol{\Phi}_2$ make the inverse of $\boldsymbol{\Phi}_{cc}^{k+1}$ be very complicated. The matrix inverse will be simplified by using the properties of data matrices in the new recursive algorithm.

According to the properties of quasi-stationarity and orthogonality, $\|\boldsymbol{\Phi}_1\| \gg \|\boldsymbol{\Phi}_2\|$ holds for $N \gg M$, and the entries of $\boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2$ are much smaller than 1. Consequently, the following inversion can be approximated by

$$\left(\boldsymbol{I} + \boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2\right)^{-1} = \boldsymbol{I} - \boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2 + \left(\boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2\right)^2 - \cdots \approx \boldsymbol{I} - \boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2, \tag{14}$$

and it yields the approximation of inverse of $\boldsymbol{P}^{k+1}$, i.e., the inverse of $\boldsymbol{\Phi}_{cc}^{k+1}$ as follows

$$\boldsymbol{P}^{k+1} = \left(\boldsymbol{\Phi}_{cc}^{k+1}\right)^{-1} = \boldsymbol{W}_c^{-1}\left(\boldsymbol{\Phi}_1 + \boldsymbol{\Phi}_2\right)^{-1}\boldsymbol{W}_c^{-1} \approx \boldsymbol{W}_c^{-1}\left(\boldsymbol{I} - \boldsymbol{\Phi}_1^{-1}\boldsymbol{\Phi}_2\right)\boldsymbol{\Phi}_1^{-1}\boldsymbol{W}_c^{-1}, \tag{15}$$

where $\boldsymbol{\Phi}_1^{-1}$ can be updated following matrix inversion lemma

$$\boldsymbol{\Phi}_1^{-1} = \left(\boldsymbol{I} - \boldsymbol{g}^{k+1}\left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T\right)\boldsymbol{P}^k, \tag{16}$$

whereas $\boldsymbol{g}^{k+1}$ is a gain vector given by

$$\boldsymbol{g}^{k+1} = \boldsymbol{P}^k\boldsymbol{\psi}^{k+1}\left(\boldsymbol{I}_2 + \left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T\boldsymbol{P}^k\boldsymbol{\psi}^{k+1}\right)^{-1}. \tag{17}$$

In (17), $\boldsymbol{I}_2$ is a (2×2) identity matrix, so the calculation of $\left(\boldsymbol{I}_2 + \left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T\boldsymbol{P}^k\boldsymbol{\psi}^{k+1}\right)^{-1}$ is very easy in the recursive algorithm.

Similarly as $\boldsymbol{\Phi}_{cc}^{k+1}$, the correlation vectors $\boldsymbol{\phi}_{cy}^{k+1}$ and $\boldsymbol{\phi}_{sy}^{k+1}$ can be updated by

$$\begin{aligned}\boldsymbol{\phi}_{cy}^{k+1} &= \boldsymbol{W}_c\left(\boldsymbol{\psi}^{k+1}\bar{\boldsymbol{y}}^{k+1} + \boldsymbol{\phi}_{cy}^k + \boldsymbol{W}_{cs}^{-1}\boldsymbol{\phi}_{sy}^k\right),\\ \boldsymbol{\phi}_{sy}^{k+1} &= -\boldsymbol{W}_s\left(\boldsymbol{\psi}^{k+1}\bar{\boldsymbol{y}}^{k+1} + \boldsymbol{\phi}_{cy}^k\right) + \boldsymbol{W}_c\boldsymbol{\phi}_{sy}^k.\end{aligned} \tag{18}$$

From (18), the extra term $\boldsymbol{\phi}_{sy}^k$ can be expressed by the past data

$$\boldsymbol{\phi}_{sy}^k = -\boldsymbol{W}_s\left(\boldsymbol{\psi}^k\bar{\boldsymbol{y}}^k + \boldsymbol{\phi}_{cy}^{k-1}\right) + \boldsymbol{W}_c\boldsymbol{\phi}_{sy}^{k-1}, \tag{19}$$

Since the matrices $\boldsymbol{W}_c$, $\boldsymbol{W}_s$ and $\boldsymbol{W}_{cs}^{-1}$ are diagonal, $\boldsymbol{\phi}_{sy}^k$ can be rewritten as

$$\boldsymbol{\phi}_{sy}^k = \boldsymbol{W}_c\left(\boldsymbol{I} + \boldsymbol{W}_{cs}^{-2}\right)\boldsymbol{\phi}_{sy}^{k-1} - \boldsymbol{W}_{cs}^{-1}\boldsymbol{\phi}_{cy}^k. \tag{20}$$

4.2. **Update of parameter estimation.** Now substitute the approximated formulae of $\boldsymbol{P}^{k+1}$ and $\boldsymbol{\phi}_{cy}^{k+1}$ to deduce the recursive estimation $\hat{\boldsymbol{\theta}}_c^{k+1} = \boldsymbol{P}^{k+1}\boldsymbol{\phi}_c^{k+1}$ in the new window $[k_0+1, k+1]$, where $\boldsymbol{\phi}_{cy}^{k+1}$ in (18) is split into two parts: $\boldsymbol{W}_c\left(\boldsymbol{\psi}^{k+1}\bar{\boldsymbol{y}}^{k+1} + \boldsymbol{\phi}_{cy}^k\right)$ and $\boldsymbol{W}_s\boldsymbol{\phi}_{sy}^k$.

Multiplying $\boldsymbol{\Phi}_1^{-1}\boldsymbol{W}_c^{-1}$ by the first part of $\boldsymbol{\phi}_{cy}^{k+1}$ yields that

$$\boldsymbol{\Phi}_1^{-1}\boldsymbol{W}_c^{-1}\boldsymbol{W}_c\left(\boldsymbol{\psi}^{k+1}\bar{\boldsymbol{y}}^{k+1} + \boldsymbol{\phi}_{cy}^k\right) = \left(\boldsymbol{\psi}^{k+1}\left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T + \boldsymbol{\Phi}_{cc}^k\right)^{-1}\left(\boldsymbol{\psi}^{k+1}\bar{\boldsymbol{y}}^{k+1} + \boldsymbol{\phi}_{cy}^k\right). \tag{21}$$

Similarly as the standard recursive formula in [17], (21) can be compactly rewritten as

$$\boldsymbol{P}^k\boldsymbol{\phi}_{cy}^k + \boldsymbol{g}^{k+1}\varepsilon^{k+1} = \hat{\boldsymbol{\theta}}_c^k + \boldsymbol{g}^{k+1}\varepsilon^{k+1}, \tag{22}$$

where the prediction error ε^{k+1} is defined by

$$\varepsilon^{k+1} = \bar{\boldsymbol{y}}^{k+1} - \left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T\hat{\boldsymbol{\theta}}_c^k. \tag{23}$$

For the second part of $\boldsymbol{\phi}_{cy}^{k+1}$, substituting (19) into the multiplication of $\boldsymbol{P}^{k+1}$ and $\boldsymbol{W}_s\boldsymbol{\phi}_{sy}^k$ yields that

$$\begin{aligned}\boldsymbol{\Phi}_1^{-1}\boldsymbol{W}_c^{-1}\boldsymbol{W}_s\boldsymbol{\phi}_{sy}^k \approx{}& \left(\boldsymbol{I} - \boldsymbol{g}^{k+1}\left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T\right)\left(\boldsymbol{I} + \boldsymbol{W}_{cs}^{-2}\right)\bar{\boldsymbol{\theta}}_s^k\\ &- \left(\boldsymbol{I} - \boldsymbol{g}^{k+1}\left(\bar{\boldsymbol{\psi}}^{k+1}\right)^T\right)\boldsymbol{W}_{cs}^{-2}\bar{\boldsymbol{\theta}}_c^k.\end{aligned} \tag{24}$$

Now, we complement the rest terms in the update of $\boldsymbol{P}^{k+1}$ and $\hat{\boldsymbol{\theta}}_c^{k+1}$. Denote the following gain matrices to simplify the recursive formulae

$$\boldsymbol{\Omega}^k = \boldsymbol{P}^k \left(\boldsymbol{W}_{cs}^{-1} \boldsymbol{\phi}_{cs}^k + \left(\boldsymbol{\phi}_{cs}^k \right)^T \boldsymbol{W}_{cs}^{-1} + \boldsymbol{W}_{cs}^{-1} \boldsymbol{\phi}_{ss}^k \boldsymbol{W}_{cs}^{-1} \right),$$

$$\boldsymbol{G}^{k+1} = \boldsymbol{W}_c^{-1} \left(\boldsymbol{I} - \left(\boldsymbol{I} - \boldsymbol{g}^{k+1} \left(\bar{\boldsymbol{\psi}}^{k+1} \right)^T \right) \boldsymbol{\Omega}^k \right), \ \boldsymbol{G}_s^{k+1} = \boldsymbol{G}^{k+1} \left(\boldsymbol{I} - \boldsymbol{g}^{k+1} \left(\bar{\boldsymbol{\psi}}^{k+1} \right)^T \right).$$

Then, by combining (22) with (24), the new parameter vector and the inverse of the correlation matrix can be concluded as follows:

$$\hat{\boldsymbol{\theta}}_c^{k+1} = \boldsymbol{P}^{k+1} \boldsymbol{\phi}_{cy}^{k+1} = \boldsymbol{G}^{k+1} \left(\hat{\boldsymbol{\theta}}_c^k + \boldsymbol{g}^{k+1} \varepsilon^{k+1} \right) + \hat{\boldsymbol{\theta}}_s^{k+1}, \tag{25}$$

$$\boldsymbol{P}^{k+1} = \boldsymbol{G}^{k+1} \boldsymbol{\Phi}_1^{-1} \boldsymbol{W}_c^{-1} = \boldsymbol{G}^{k+1} \left(\boldsymbol{I} - \boldsymbol{g}^{k+1} \left(\bar{\boldsymbol{\psi}}^k \right)^T \right) \boldsymbol{P}^k \boldsymbol{W}_c^{-1}. \tag{26}$$

The estimate in (25) is composed of 2 parts: the first part projects the term $\hat{\boldsymbol{\theta}}_c^k + \boldsymbol{g}^{k+1} \varepsilon^{k+1}$ onto the cosine basis in the new window, while the second part $\hat{\boldsymbol{\theta}}_s^{k+1}$ corresponds to transition effect on the sliding window with respect to the extra terms $\boldsymbol{\phi}_s^k$, $\boldsymbol{\phi}_{sy}^k$ appeared in the update of $\boldsymbol{\phi}_c^{k+1}$ and $\boldsymbol{\phi}_{cy}^{k+1}$.

Moreover, the correlation matrix and vector are updated as follows:

$$\begin{bmatrix} \boldsymbol{\Phi}_{cc}^{k+1} & \boldsymbol{\Phi}_{cs}^{k+1} \\ \boldsymbol{\Phi}_{sc}^{k+1} & \boldsymbol{\Phi}_{ss}^{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_c & \boldsymbol{W}_s \\ -\boldsymbol{W}_s & \boldsymbol{W}_c \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{cc}^k + \boldsymbol{\Psi}^{k+1} & \boldsymbol{\Phi}_{cs}^k \\ \boldsymbol{\Phi}_{sc}^k & \boldsymbol{\Phi}_{ss}^k \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_c & -\boldsymbol{W}_s \\ \boldsymbol{W}_s & \boldsymbol{W}_c \end{bmatrix}. \tag{27}$$

4.3. **Simplification of recursive computation.** It is seen that in (25) the matrix size $(M+1)(n+1)$ makes the computation of matrix multiplication complicated for the high order model approximated by the high degree cosine series. In order to reduce the computational complexity, the implementation of the parameters and matrices' update is divided into $(M+1)$ sub-blocks for each of the basis function $\cos(m\omega k_1)$, with respect to the orthogonality of the trigonometric functions. Following the structure of $\boldsymbol{\phi}_{c,m}^k(k_1)$ and $\boldsymbol{\phi}_c^k(k_1)$ defined in (5), it is seen that $\boldsymbol{\Phi}_{cc}^k$ is composed of the following sub-blocks

$$\boldsymbol{\Phi}_{cc,m_1m_2}^k = \sum_{k_1=0}^{N} \boldsymbol{\phi}_{c,m_1}^k(k_1) \left(\boldsymbol{\phi}_{c,m_2}^k(k_1) \right)^T. \tag{28}$$

Similar definitions are given for the blocks of $\boldsymbol{\Phi}_{cs,m_1m_2}^k$, $\boldsymbol{\Phi}_{sc,m_1m_2}^k$ and $\boldsymbol{\Phi}_{ss,m_1m_2}^k$. For the simplicity of notation, the blocks for $m_1 = m_2 = m$ are abbreviated as $\boldsymbol{\Phi}_{cc,m}^k$, $\boldsymbol{\Phi}_{ss,m}^k$.

The updates are just given by replacing the counterparts with the sub-blocks corresponding to mth degree of the basis function. For example, the gain vector $\boldsymbol{g}_m^{k+1}$ is calculated as follows:

$$\boldsymbol{g}_m^{k+1} = \boldsymbol{P}_m^k \, \boldsymbol{\psi}_m^{k+1} \left(\boldsymbol{I}_2 + \left(\bar{\boldsymbol{\psi}}_m^{k+1} \right)^T \boldsymbol{P}_m^k \boldsymbol{\psi}_m^{k+1} \right)^{-1}. \tag{29}$$

Furthermore, the matrices $\boldsymbol{G}_m^{k+1}$, $\boldsymbol{G}_{s,m}^{k+1}$ are given by

$$\boldsymbol{G}_m^{k+1} = \boldsymbol{W}_{c,m}^{-1} \left(\boldsymbol{I} - \left(\boldsymbol{I} - \boldsymbol{g}_m^{k+1} \left(\bar{\boldsymbol{\phi}}_m^{k+1} \right)^T \right) \boldsymbol{\Omega}_m^k \right), \tag{30}$$

$$\boldsymbol{G}_{s,m}^{k+1} = \boldsymbol{G}_m^{k+1} \left(\boldsymbol{I} - \boldsymbol{g}_m^{k+1} \left(\bar{\boldsymbol{\phi}}_m^{k+1} \right)^T \right). \tag{31}$$

Then $\hat{\boldsymbol{\theta}}_{s,m}^{k+1}$ and $\hat{\boldsymbol{\theta}}_{c,m}^{k+1}$ can be given by

$$\varepsilon_m^{k+1} = \bar{\boldsymbol{y}}_m^{k+1} - \left(\bar{\boldsymbol{\psi}}_m^{k+1}\right)^T \hat{\boldsymbol{\theta}}_{c,m}^k, \tag{32}$$

$$\hat{\boldsymbol{\theta}}_{s,m}^{k+1} = \boldsymbol{G}_{s,m}^{k+1} \left(\left(\boldsymbol{I} + \boldsymbol{W}_{cs,m}^{-2}\right) \bar{\boldsymbol{\theta}}_{s,m}^k - \boldsymbol{W}_{cs,m}^{-2} \bar{\boldsymbol{\theta}}_{c,m}^k\right), \tag{33}$$

$$\hat{\boldsymbol{\theta}}_{c,m}^{k+1} = \boldsymbol{G}_m^{k+1} \left(\bar{\boldsymbol{\theta}}_{c,m}^k + \boldsymbol{g}_m^{k+1} \varepsilon_m^{k+1}\right) + \bar{\boldsymbol{\theta}}_{s,m}^{k+1}. \tag{34}$$

Moreover, the correlation matrix and its inverse are updated as follows:

$$\boldsymbol{P}_m^{k+1} = \boldsymbol{G}_m^{k+1} \left(\boldsymbol{P}_m^k - \boldsymbol{g}_m^{k+1} \left(\bar{\boldsymbol{\psi}}_m^{k+1}\right)^T \boldsymbol{P}_m^k\right) \boldsymbol{W}_{c,m}^{-1}, \tag{35}$$

$$\begin{bmatrix} \boldsymbol{\Phi}_{cc,m}^{k+1} & \boldsymbol{\Phi}_{cs,m}^{k+1} \\ \boldsymbol{\Phi}_{sc,m}^{k+1} & \boldsymbol{\Phi}_{ss,m}^{k+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_{c,m} & \boldsymbol{W}_{s,m} \\ -\boldsymbol{W}_{s,m} & \boldsymbol{W}_{c,m} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{cc,m}^k + \boldsymbol{\Psi}_m^{k+1} & \boldsymbol{\Phi}_{cs,m}^k \\ \boldsymbol{\Phi}_{sc,m}^k & \boldsymbol{\Phi}_{ss,m}^k \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_{c,m} & -\boldsymbol{W}_{s,m} \\ \boldsymbol{W}_{s,m} & \boldsymbol{W}_{c,m} \end{bmatrix}. \tag{36}$$

It is obvious that $\boldsymbol{W}_{c,0}$ is an identity matrix, and all the elements of $\boldsymbol{W}_{s,0}$, $\boldsymbol{\phi}_{s,0}(k_1)$ are zero, then $\boldsymbol{\Omega}_0^k = 0$, $\boldsymbol{\theta}_{s,0}^k = \boldsymbol{\theta}_{s,0}^{k+1} = 0$. Therefore, the update for $m = 0$ is almost the same as the standard RLS algorithm. The simplified algorithm can be regarded as an extension of RLS algorithm into the high degree of the basis series to track the rapid variations.

4.4. **Smoothing of parameter estimates.** A smoothing technique is applied in the proposed algorithms to improve the resistance against the noise. Notice that the $(k+1)$th recursion of estimation $\hat{\boldsymbol{\theta}}_c^{k+1}$ yields the parameter estimates of $\hat{h}_i^{k+1}$, $\hat{h}_i^k$, $\hat{h}_i^{k-1}$, $\ldots$ within the window $[k_0 + 1, k + 1]$, where $\hat{h}_i^k$, $\hat{h}_i^{k-1}$, $\ldots$ are overlapped with the window $[k_0, k]$. Consequently, the simple smoothing way is by using the overlapped parameters to smooth the estimated parameters at k, $k - 1, \ldots$. For example, the parameters of $\hat{h}_i^{k-n_0}$ can be smoothed by using the cosine expansion with $\hat{\boldsymbol{\theta}}_c^{k-n_0+1}, \ldots, \hat{\boldsymbol{\theta}}_c^{k+1}$ at point $k - n_0$ as follows:

$$\hat{h}_i^{k-n_0+1} = \frac{1}{n_0+1} \sum_{k_1=1}^{n_0+1} \left(\sum_{m=0}^{M} \hat{c}_{i,m}^{k-n_0+k_1} \cos\left(m\omega(N+1-k_1)\right) \right). \tag{37}$$

Moreover, by using some methods to estimate the jump points [18], the Gibbs effect inside the window can be further mitigated at the discontinuous points, and can improve the approximation accuracy of the cosine series.

5. **Numerical Examples.** A time-varying digital communication channel is considered in the numerical examples. Assume that the true channel model is described by (1) where $u(k)$ and $y(k)$ are the transmitted training signal, received signal, respectively, $e(k)$ is a white additive noise that is independent of $u(k)$, the parameter h_i^k varies with the instant k, and the variation rate in (2) is $\kappa(k) > 0.1 >> 0.0005\sqrt{n}$, so the processes are the fast time-varying ones and their parameter estimation is not an easy task.

5.1. **Simulation example 1.** Let the model order $n = 5$ and the time varying parameters of the channel model be the time functions. The following are the examples of h_0^k and h_2^k given by

$$h_0^k = \begin{cases} -1, & \text{for } k \in [1, 1204], \text{ or } [3072, 3372], \\ 0.35 \sin\left(\dfrac{k\pi}{150} + \dfrac{\pi}{6}\right), & \text{for } k \in [2561, 2880], \\ 1 + 0.35 \sin\left(\dfrac{k\pi}{150} + \dfrac{\pi}{6}\right), & \text{otherwise,} \end{cases}$$

$$h_2^k = \begin{cases} -0.75, & \text{for } k \in [1, 1204], \text{ or } [3072, 3372], \\ -\frac{1}{4}\left(1 - \sin\left(\frac{k\pi}{100} + \frac{\pi}{3}\right)\right), & \text{for } k \in [2561, 2880], \\ 0.5 + 0.25\sin\left(\frac{k\pi}{100} + \frac{\pi}{3}\right), & \text{otherwise}. \end{cases} \tag{38}$$

As shown in Figure 2, it is seen that there are significant jump points at $k = 2561, 2880$, 3072, 3372. Assume that the signal to noise ratio (SNR) is 10dB.

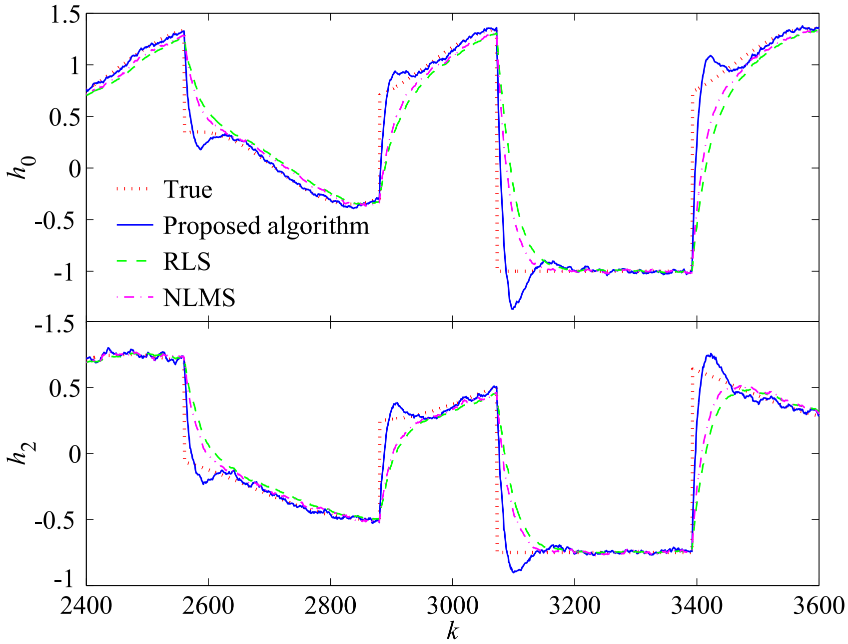


FIGURE 2. Estimation of time-varying channel

Generally, the more rapid variations arise in the time window, the higher series degree is necessary. On the other hand, the approximation with high series degree is easily influenced by the noise, and a long time window is often required to reduce the noise influence. Therefore, both the parameter variation velocity and the noise level should be considered when choosing the time window length N and series degree M. In the simulation, the window width is chosen as $N = 1024$, and the series degree $M = 3$. The mean values of the estimated parameters $\hat{h}_{0,k}$ and $\hat{h}_{2,k}$ for 50 simulation runs are plotted in Figure 2. As a comparison, the results of RLS with the forgetting factor 0.95, and NLMS with the updating step size 0.15 are also shown in the figure. It illustrates that the proposed algorithm tracks the variation more promptly than RLS and NLMS, especially at the sharp jump points.

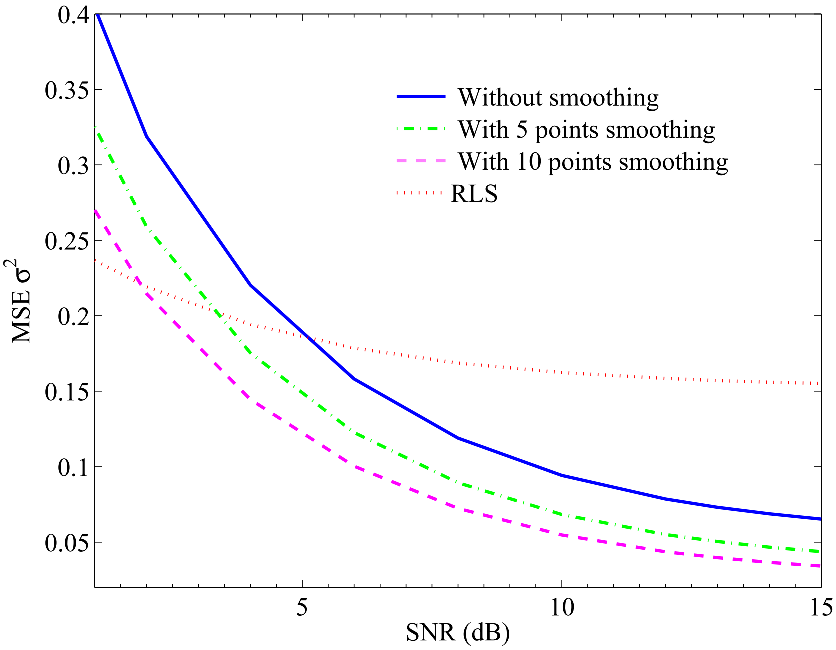
Define the mean square error (MSE) σ^2 of the estimated parameters as follows:

$$\sigma^2 = \frac{1}{LK} \sum_{l=1}^{L} \sum_{k_1=1}^{K} \left(\sum_{i=0}^{n} \left(\hat{h}_i^{k_0+k_1} - h_i^{k_0+k_1} \right)^2 \right), \tag{39}$$

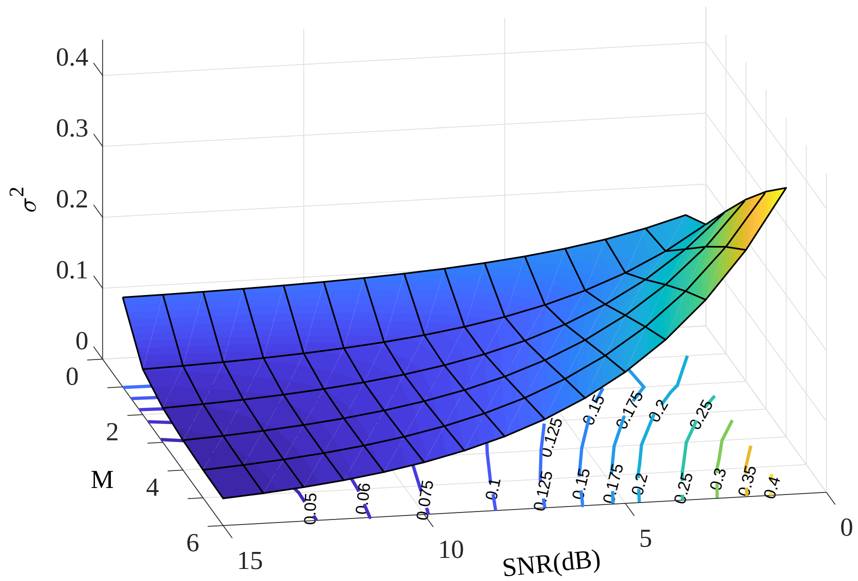
where L and K are the number of simulation runs, the number of recursions, respectively. The values of σ^2 for $k \in [2400, 3600]$ in the proposed algorithm, the standard RLS and NLMS algorithms are 0.0844, 0.1584, 0.1475, respectively. It is seen that the proposed algorithm has a smaller mean square error than the other two methods under the same simulation conditions, especially around the rapid varying points.

5.2. **Simulation example 2.** Let SNR be from 0dB to 15dB. The other simulation conditions are the same as example 1. The mean squares errors σ^2 obtained by estimating with 5 points, 10 points in 50 simulation runs are shown in Figure 3(a). It is shown that the MSE can be decreased by applying the smoothing technique.

5.3. **Simulation example 3.** Let the series degree M be chosen from 1 to 6. The values of σ^2 under various noise environments are shown in Figure 3(b). It is seen that σ^2 decreases with increasing M under the low noise environment, whereas it becomes large for the high degree M under the strong noise environment due to the fluctuation in the



(a) Smoothing effect in parameter estimation



(b) Estimation errors vs. series degree M and SNR

FIGURE 3. Performance of parameter estimation

estimation of $\boldsymbol{\theta}_{c,m}^{k+1}$. These results imply that the optimal identification performance can be obtained through selecting an appropriate M with respect to the noise level and the variation velocity of the model parameters at the present instant, which may be detected by combining the algorithm with some detection methods for the rapid variations.

5.4. **Simulation example 4.** Consider the channel model has long lag time of $n = 60$. The window width and the series degree are chosen as $N = 3072$, $M = 30$, respectively. The value of σ^2 is 0.1237, and the estimates' mean values of $\hat{h}_{0,k}$ and $\hat{h}_{2,k}$ are plotted in Figure 4. As a comparison, the parameters are also estimated by AP, BOP, RLS and NLMS, and the values of σ^2 are 0.2498, 0.2748, 0.2751, 0.2778, respectively, which are more than 2 times of that in the proposed algorithm. Though the orthogonal projection of input data can improve the convergence performance when the input signal is colored, the AP and BOP methods also have lower convergence rate than the proposed algorithm since they do not use the explicit projection of the parameter variation. Moreover, the estimation error σ^2 of the algorithm [16] where the forgetting factor is chosen as 0.96 is 0.1934, which is larger than that of the proposed algorithm since the noise influences become relatively large for high degree M.

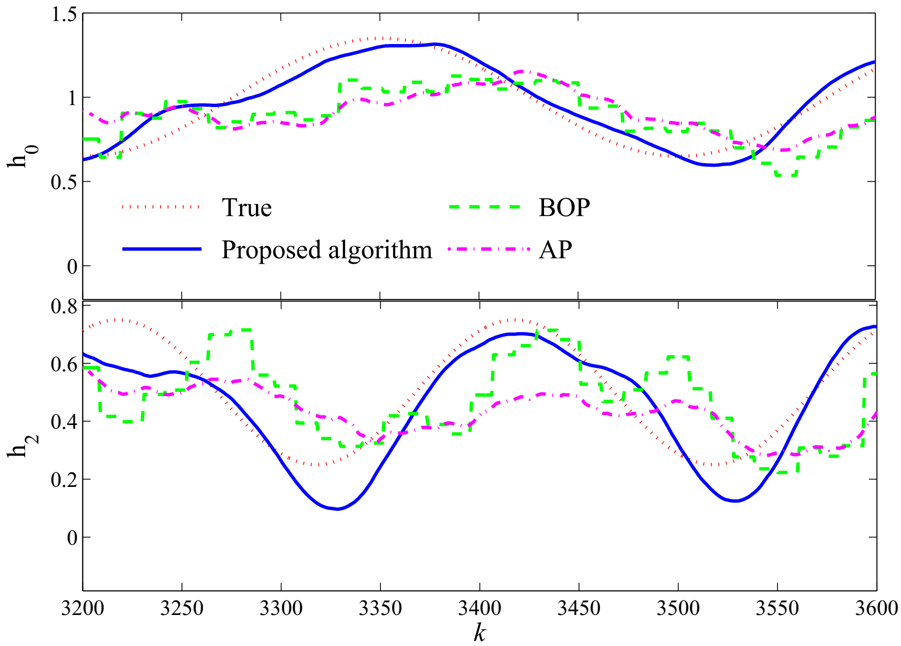


FIGURE 4. Estimation of h_0 and h_2 in example 4

6. **Conclusions.** The recursive identification algorithm based on the trigonometric functions has been developed for the linear time-varying systems. When the parameters of the process model have at most finite discontinuous points in the data window, they can be approximated by the cosine series through virtually expanding them into the even periodic functions, and then the parameter estimation can be obtained through estimating the coefficients of the cosine series. By making use of the orthogonality of the basis functions, the recursive identification algorithm has been proposed where the recursive computability is guaranteed. The simulation results demonstrate that the proposed algorithm has a higher convergence rate than the conventional methods. The extension of the algorithm under the strong noise environment will be investigated in our future work.

REFERENCES

[1] M. Basseville and I. Nikiforov, *Detecting of Abrupt Changes – Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[2] Q. H. Zhang and L. Ljung, From structurally independent local LTI models to LPV model, *Automatica*, vol.84, pp.232-235, 2017.

[3] S. V. Vaerenbergh, I. Santamaría and M. Lázaro-Gredilla, Estimation of the forgetting factor in kernel recursive least squares, *IEEE International Workshop on Machine Learning for Signal Processing*, vol.50, no.6, pp.1-6, 2012.

[4] Z. H. Zhang and D. J. Zhang, New variable step size LMS adaptive filtering algorithm and its performance analysis, *Systems Engineering & Electronics*, vol.31, no.9, pp.2238-2241, 2009.

[5] N. Li, Y. Zhang, Y. Hao and J. A. Chambers, A new variable step-size NLMS algorithm designed for applications with exponential decay impulse responses, *Signal Processing*, vol.88, no.9, pp.2346-2349, 2008.

[6] K. Ikeda, Convergence analysis of block orthogonal projection and affine projection algorithms, *Signal Processing*, vol.82, pp.491-496, 2002.

[7] W. B. J. Hakvoort, R. G. K. M. Aarts, J. V. Dijk and J. B. Jonker, A computationally efficient algorithm of iterative learning control for discrete-time linear time-varying systems, *Automatica*, vol.45, no.12, pp.2925-2929, 2009.

[8] M. K. Tsatsanis and G. B. Giannakis, Modelling and equalization of rapidly fading channels, *International Journal of Adaptive Control and Signal Processing*, vol.10, nos.2-3, pp.159-176, 2010.

[9] G. Rödönyi, P. Körös, P. Dániel, A. Soumelidis, Z. Szabo and J. Bokor, Structure selection and LPV model identification of a car steering dynamics, *The 18th IFAC Symposium on System Identification*, pp.1086-1091, 2018.

[10] D. Machala, S. Dobre, M. Albisser, F. Collin and M. Gilson, Quasi-LPV modeling of a projectile's behaviour in flight, *The 18th IFAC Symposium on System Identification*, pp.1080-1085, 2018.

[11] S. Simani and R. J. Patton, Fault diagnosis of an industrial gas turbine prototype using a system identification approach, *Control Engineering Practice*, vol.16, no.7, pp.769-786, 2008.

[12] H. Wang, L. Bai, J. Xu and W. Fei, EEG recognition through time-varying vector autoregressive model, *The 12th Int. Conf. on Fuzzy Systems and Knowledge Discovery*, pp.292-296, 2015.

[13] M. Niedzwiecki, *Identification of Time-varying Processes*, John Wiley & Sons Ltd., 2000.

[14] S. A. Broughton and K. Bryan, Discrete Fourier analysis and wavelets, *Journal of Applied Statistics*, vol.37, 2008.

[15] M. D. Spiridonakos and S. D. Fassois, Non-stationary random vibration modelling and analysis via functional series time-dependent ARMA (FS-TARMA) models – A critical survey, *Mechanical Systems and Signal Processing*, vol.47, nos.1-2, pp.175-224, 2014.

[16] L. Sun and L. Li, New recursive identification algorithm for time-varying systems, *ICCAS-SICE*, pp.4917-4922, 2009.

[17] L. Ljung, *System Identification – Theory for the User*, 2nd Edition, Prentice Hall, Englewood Cliffs NJ, 1999.

[18] L. Blackmore, S. Rajamanoharan and B. C. Williams, Active estimation for jump Markov linear systems, *IEEE Trans. Automat. Contr.*, vol.53, no.10, pp.2223-2236, 2008.