Research Article
HOC Based Blind Identification of Hydroturbine Shaft Volterra System

Bing Bai[1] and Lixiang Zhang[2]

[1]*School of Civil Engineering and Communication, North China University of Water Resources & Electric Power, Beihuan Road No. 36, Zhengzhou, Henan 450045, China*
[2]*Department of Engineering Mechanics, Kunming University of Science and Technology, South Jingming Road No. 727, Kunming, Yunnan 650500, China*

Correspondence should be addressed to Bing Bai; baibing7777@ncwu.edu.cn

In order to identify the quadratic Volterra system simplified from the hydroturbine shaft system, a blind identification method based on the third-order cumulants and a reversely recursive method are proposed. The input sequence of the system under consideration is an unobservable independent identically distributed (i.i.d.), zero-mean and non-Gaussian stationary signal, and the observed signals are the superposition of the system output signal and Gaussian noise. To calculate the third-order moment of the output signal, a computer loop judgment method is put forward to determine the coefficient. When using optimization method to identify the time domain kernels, we combined the traditional optimization algorithm (direct search method) with genetic algorithm (GA) and constituted the hybrid genetic algorithm (HGA). Finally, according to the prototype observation signal and the time domain kernel parameters obtained from identification, the input signal of the system can be gained recursively. To test the proposed method, three numerical experiments and engineering application have been carried out. The results show that the method is applicable to the blind identification of the hydroturbine shaft system and has strong universality; the input signal obtained by the reversely recursive method can be approximately taken as the random excitation acted on the runner of the hydroturbine shaft system.

1. Introduction

There are lots of nonlinear phenomena in engineering. Knowing of the coherent characteristics is of great importance to engineering safety. Thus exploring an appropriate method to describe an input-output relation defined by the nonlinear system has been a hot research topic. At present, Volterra series is a commonly used technique in researchers, and achievements are involved with plenty of engineering fields. Maheswaran and Khosa [1, 2] use a Volterra series model with wavelet functions to identify the nonlinearity of a natural stream flow and to forecast the water levels in groundwater system, respectively. In forensic investigation, a filter designed with Volterra series can well work when it is necessary to identify the latent fingerprints [3]. A nonlinear phenomenon that is produced by a distorted harmonic audible signal is recognized by estimating sound pressure level with

Volterra series [4]. In addition, the nonlinear structure response of a large container carrier in irregular seaway is calculated with a third-order Volterra model [5]. For a multiple input-multiple output (MIMO) system, a topological assemblage scheme is developed to make synthesis of Volterra system and applied to engineering, such as vibration analysis of wind-excited suspended cable [6].

The applications of Volterra series are mainly demonstrated in two aspects, that is, modeling and identification of an input-output system. Two important properties of Volterra series are (1) output of a Volterra system depends linearly on kernel parameters of this system; (2) nonlinearity of a signal can be represented throughout multidimensional operators on product of samples. Therefore, Volterra model is generally used to describe the input-output relationship of a nonlinear system in a long term. Nonlinear systems, such as system with additive white Gaussian noises [7], structure subjected to
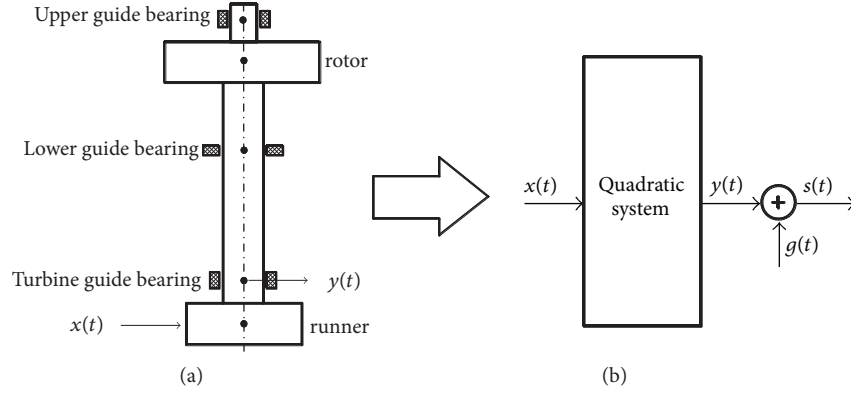
FIGURE 1: Hydroturbine shaft system and quadratic Volterra system.

impact load [8], and turbomachinery [9], can be represented by Volterra series. Systems with multispectral properties are also reported [10, 11].

Peng et al. make a lot of achievements on the modeling, identification, and prediction with Volterra series [12–14]. They discussed lots of problems, for instance, the influence on the force transmissibility of MDOF structures with cubic nonlinear damping device [12], a new method to detect the breathing fatigue crack in nonlinear beam [13] and the analytical expression of calculating the bispectrum with higher order cumulants for nonlinear systems subjected to zero-mean Gaussian excitation [14]. These works highly prove the advantages of Volterra series in describing a nonlinear system and the convenience for derivation of the subsequent formulae.

It is fundamental to use Volterra series method to identify a nonlinear system. Take Duffing oscillator as an example, its "jump" phenomenon can be recognized by the Volterra series method [15], which is expanded to identify a bilinear oscillator [16]. To deal with identification problems more effectively, Volterra series is usually combined with an optimization method [17]. However, the input signal of a nonlinear system is not always known. Thus the blind identification method without the input signal is developed. Le Caillec [18] blindly identified a second-order Wiener model driven by a Gaussian noise with corrupted output data. Ghandchi Tehrani et al. [19] compared the Volterra series method with the harmonic balance method, which proves the ability of blind identification by the Volterra series method.

Actually, when Volterra series is employed to model and blindly identify a nonlinear system, it needs to often meet the situation in which higher order cumulants are used. Antari et al. [20, 21] considered the third-order and fourth-order cumulants to blindly identify a Hammerstein system. However, the expression of higher cumulants sometimes seems to be quite complicated [22]. For a hydroturbine main shaft system in a hydropower station, the nonlinear input-output relation is very complicated. Especially for the input signals, such as the hydraulic excitation on the runner, its working mechanisms are not completely known so far. On the contrary, the corresponding output signals, such as the throw at turbine guide bearing, are always easy to be obtained.

Thus it is a mathematical problem which needs to be solved frequently in hydraulic engineering to identify the whole shaft system with the observed vibration response only.

This paper proposes a novel blind identification method based on higher order cumulants for a nonlinear system described by a hydroturbine shaft Volterra model. A computer loop judgment method is used to determine the coefficients γ_{ijk} when the third-order moment of the input signal is calculated. It greatly simplifies formula and algorithm. Finally, the Volterra kernel parameters of the nonlinear system under consideration are identified by the hybrid genetic algorithm (HGA).

2. Description of System

The shaft system of a certain hydroturbine generating set can be simplified as a quadratic Volterra model, as shown in Figure 1. The input sequence of the system is $\{x(t)\}$, representing the hydraulic excitation on the runner; the output sequence $\{y(t)\}$ denotes the dynamic displacements at turbine guide bearing; the observed signal is $\{s(t)\}$, and Gaussian noise is $\{g(t)\}$.

The corresponding nonlinear system is described as

$$y(t) = \sum_{i=0}^{q} \sum_{j=0}^{q} h(i, j) x(t - i) x(t - j), \qquad (1)$$

$$s(t) = y(t) + g(t), \qquad (2)$$

where t stands for the discrete time, $h(i, j)$ stands for the time domain kernels of the quadratic nonlinear system, and q represents the system memory length $(q + 1)$, and i, j are the time delay. The excitation sequence $\{x(t)\}$ of the system is an unobservable, independent identically distributed (i.i.d.), zero-mean and non-Gaussian stationary signal. It is assumed that the following equations hold:

$$\begin{aligned}
\gamma_{2x} &= E\left[x^2(t)\right] \neq 0, \\
\gamma_{3x} &= E\left[x^3(t)\right] \neq 0, \\
\gamma_{4x} &= E\left[x^4(t)\right] \neq 0, \\
\gamma_{6x} &= E\left[x^6(t)\right] \neq 0,
\end{aligned} \qquad (3)$$

where $\gamma_{m,x}$ stands for the mth-order moment at the origin of the input signal. It is proved that, when $\{x(t)\}$ is a zero-mean sequence, the following properties [22] hold:

$$
\begin{aligned}
& E\left\{x\left(t_1\right) \cdots x\left(t_6\right)\right\} \\
& = \begin{cases}
\gamma_{6x} & \text{if } t_1 = t_2 = t_3 = t_4 = t_5 = t_6 \\
\gamma_{4x}\gamma_{2x} & \text{if } t_1 = t_2 \neq t_3 = t_4 = t_5 = t_6 \\
\gamma_{3x}^2 & \text{if } t_1 = t_2 = t_3 \neq t_4 = t_5 = t_6 \\
\gamma_{2x}^3 & \text{if } t_1 = t_2 \neq t_3 = t_4 \neq t_5 = t_6 \\
0 & \text{otherwise.}
\end{cases}
\end{aligned} \tag{4}
$$

Letting $C_{m,y}(\tau_1, \ldots, \tau_{m-1})$ and $C_{m,s}(\tau_1, \ldots, \tau_{m-1})$ be cumulants of the output sequence and the observed sequence, respectively, a relationship between the higher order cumulants (HOC) and time domain kernels $h(i, j)$ of the quadratic nonlinear system is established by calculation. In the present study, we try to establish a relationship between the third-order cumulants of the observed signal and time domain kernels $h(i, j)$. Because the third-order or higher order cumulants of the Gaussian noise are zero, we obtain from (2) the following:

$$
C_{3s}\left(\tau_1, \tau_2\right) = C_{3y}\left(\tau_1, \tau_2\right) + C_{3g}\left(\tau_1, \tau_2\right) = C_{3y}\left(\tau_1, \tau_2\right). \tag{5}
$$

Therefore, the third-order cumulants of the output sequence $\{y(t)\}$ need to only be calculated. For a stationary random sequence $\{y(t)\}$, the third-order cumulants are defined as

$$
\begin{aligned}
& C_{3y}\left(\tau_1, \tau_2\right) \\
& = M_{3y}\left(\tau_1, \tau_2\right) \\
& \quad - M_{1y}\left[M_{2y}\left(\tau_1\right) + M_{2y}\left(\tau_2\right) + M_{2y}\left(\tau_2 - \tau_1\right)\right] \\
& \quad + 2\left(M_{1y}\right)^3,
\end{aligned} \tag{6}
$$

where M_{1y}, $M_{2y}(\tau)$, and $M_{3y}(\tau_1, \tau_2)$ denote the first-order moment, second-order moment, and third-order moment of the sequence $\{y(t)\}$, respectively. The specific expression of M_{1y} is stated as

$$
\begin{aligned}
M_{1y} & = \sum_{i=0}^{q} \sum_{j=0}^{q} h(i, j) E\{x(t-i) x(t-j)\} \\
& = \gamma_{2x} \cdot \sum_{i=0}^{q} h(i, i).
\end{aligned} \tag{7}
$$

It implies that when i, j are not equal, the expectation value is zero; only the diagonal elements $h(i, i)$ are left for the time domain kernels $h(i, j)$. $M_{2y}(\tau)$ is as follows:

$$
\begin{aligned}
M_{2y}(\tau) & = E\{y(t) y(t+\tau)\} \\
& = \sum_{i_1, i_2, j_1, j_2 = 0}^{q} \gamma_{ij} \cdot h(i_1, i_2) h(j_1, j_2),
\end{aligned} \tag{8}
$$

where γ_{ij} is the coefficient, which consists of the second-order moment or the fourth-order moment of the input signal $\{x(t)\}$, and the value depends on the different time delay of the sequence $\{x(t)\}$. Among the four subscripts $-i_1$, $-i_2$, $\tau - j_1$, and $\tau - j_2$, if the value of one subscript (time) is different from the rest of the three, the expectation is zero, and $\gamma_{ij} = 0$. The values of γ_{ij} are yielded as

$$
\gamma_{ij} = \begin{cases}
\gamma_{4x} & \text{if } -i_1 = -i_2 = \tau - j_1 = \tau - j_2 \\
\gamma_{2x}^2 & \text{if } -i_1 = -i_2 \neq \tau - j_1 = \tau - j_2, \text{ or} \\
& \text{if } -i_1 = \tau - j_1 \neq -i_2 = \tau - j_2, \text{ or} \\
& \text{if } -i_1 = \tau - j_2 \neq -i_2 = \tau - j_1 \\
0 & \text{otherwise.}
\end{cases} \tag{9}
$$

In order to calculate the three second-order cumulants in (6), we substitute τ, respectively, by τ_1, τ_2, and $\tau_2 - \tau_1$ in (8). The expression of the last term $M_{3y}(\tau_1, \tau_2)$ in (6) is written as

$$
\begin{aligned}
M_{3y}\left(\tau_1, \tau_2\right) & = E\{y(t) y(t+\tau_1) y(t+\tau_2)\} \\
& = \sum_{i_1, i_2, j_1, j_2, k_1, k_2 = 0}^{q} \gamma_{ijk} \cdot h(i_1, i_2) h(j_1, j_2) h(k_1, k_2)
\end{aligned} \tag{10}
$$

in which coefficient γ_{ijk} is composed of one or several certain order moments of the input signal sequence $\{x(t)\}$. According to the different time delay of the sequence $\{x(t)\}$, the different values are obtained as

$$
\begin{aligned}
& \gamma_{ijk} \\
& = \begin{cases}
\gamma_{6x} & \text{if } -i_1 = -i_2 = \tau_1 - j_1 = \tau_1 - j_2 = \tau_2 - k_1 = \tau_2 - k_2 \\
\gamma_{4x}\gamma_{2x} & \text{if } -i_1 = -i_2 \neq \tau_1 - j_1 = \tau_1 - j_2 = \tau_2 - k_1 = \tau_2 - k_2 \\
\gamma_{3x}^2 & \text{if } -i_1 = -i_2 = \tau_1 - j_1 \neq \tau_1 - j_2 = \tau_2 - k_1 = \tau_2 - k_2 \\
\gamma_{2x}^3 & \text{if } -i_1 = -i_2 \neq \tau_1 - j_1 = \tau_1 - j_2 \neq \tau_2 - k_1 = \tau_2 - k_2 \\
0 & \text{otherwise.}
\end{cases}
\end{aligned} \tag{11}
$$

The judgment of the first line γ_{6x} is simple, and the judgments of the rest $\gamma_{4x}\gamma_{2x}$, γ_{3x}^2, and γ_{2x}^3 are relatively complicated since there exist a variety of different combinations, and all combinations need to compile a programming code to loop for judgment, which is called as a method on the computer loop judgment (details in the Appendix). With this method, it is possible to calculate the second-order moment and third-order moment of the output sequence $\{y(t)\}$ and to have a great superiority in computing.

3. Hybrid Genetic Algorithm

3.1. Consistent Estimate. For a known observed signal $\{s(t)\}$, we need to have an expectation operation. Because the number of the observed signals is limited, it is difficult to precisely calculate the expectation required. So, an approximating

method is used to calculate this value by consistent estimate; namely,

$$
\begin{aligned}
\widehat{C}_{3s}\left(\tau_1, \tau_2\right) \\
= \widehat{M}_{3s}\left(\tau_1, \tau_2\right) \\
- \widehat{M}_{1s}\left[\widehat{M}_{2s}\left(\tau_1\right) + \widehat{M}_{2s}\left(\tau_2\right) + \widehat{M}_{2s}\left(\tau_2 - \tau_1\right)\right] \\
+ 2\left(\widehat{M}_{1s}\right)^3,
\end{aligned}
\tag{12}
$$

$$
\widehat{M}_{3s}\left(\tau_1, \tau_2\right) = \frac{1}{N}\sum_{n=1}^{N} s\left(n\right) s\left(n + \tau_1\right) s\left(n + \tau_2\right),
\tag{13}
$$

$$
\widehat{M}_{2s}\left(\tau\right) = \widehat{R}_s\left(\tau\right) = \frac{1}{N}\sum_{n=1}^{N} s\left(n\right) s\left(n + \tau\right),
\tag{14}
$$

$$
\widehat{M}_{1s} = \widehat{\mu}_s = \frac{1}{N}\sum_{n=1}^{N} s\left(n\right).
\tag{15}
$$

In order to complete the consistent estimate of the third-order cumulants in (12), we replace τ with τ_1, τ_2, and $\tau_2 - \tau_1$ in (14), respectively.

3.2. Algorithm.
The hybrid genetic algorithm (HGA) is an intelligent optimization algorithm which combines the simple genetic algorithm (SGA) with the traditional optimization methods.

As we know, for a kind of optimization algorithm, the global search ability of the genetic algorithm (GA) is very strong. However, when the search process is close to a local optimal point, its local search ability starts to become poor. That is a shortcoming of the simple genetic algorithm. On the contrary, the traditional direct search method of discrete variables is lack of global search ability, although the local search ability is very strong. Therefore, the authors make a hybrid genetic algorithm by combining the traditional direct search method with the genetic algorithm. The hybrid genetic algorithm (HGA) improves the ability of the global and local searches in calculation.

The basic steps of HGA are sentenced in the process of GA. The local search is carried out in a certain probability p_h (for example 5%) to find out a local optimal point as soon as possible. Then it goes back to the genetic operation for the global search to complete the global search. The flow chart of HGA is as shown in Figure 2.

3.3. Blind Identification of Time Domain Kernels.
Through the above derivation, we calculate the consistent estimates of the third-order cumulants of the observed signal $\{s(t)\}$ by using (12) and (13)–(15). On the other hand, we obtain a relationship between time-domain kernels $h(i, j)$ of the nonlinear system and the third-order cumulants by using (6) and (7)–(11). Thus, letting (6) equal (12), we obtain an equation as

$$
\begin{aligned}
\widehat{C}_{3s}\left(\tau_1, \tau_2\right) \\
= M_{3y}\left(\tau_1, \tau_2\right)
\end{aligned}
$$

Figure 2: Flow chart of hybrid GA.

$$
\begin{aligned}
- M_{1y}\left[M_{2y}\left(\tau_1\right) + M_{2y}\left(\tau_2\right) + M_{2y}\left(\tau_2 - \tau_1\right)\right] \\
+ 2\left(M_{1y}\right)^3.
\end{aligned}
\tag{16}
$$
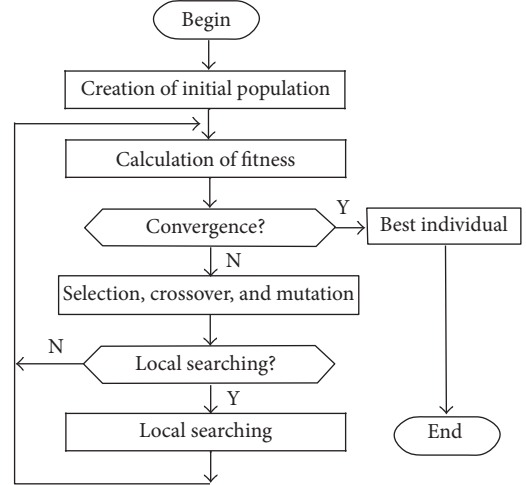
Obviously, this is a cubic nonlinear equation with unknown $h(i, j)$. By solving this equation, $h(i, j)$ is obtained. It is noted that because the output signal of the system is only used in calculation, namely, no input signal $\{x(t)\}$ is needed, so, this identification method is called blind identification.

In actual application, the third-order cumulants of $(q+1)^2$ have to be computed; that is, $C_{3y}(\tau_1, \tau_2)$, $0 \leq \tau_1, \tau_2 \leq q$. Since it is not easy to directly solve a cubic nonlinear system, the above-mentioned HGA is proposed to solve (16). The objective function (the fitness value of an individual) of the optimization problem is as follows:

$$
\min F\left(\mathbf{h}\right) = \sum_{\tau_1=0}^{q}\sum_{\tau_2=\tau_1}^{q}\left[C_{3y}\left(\tau_1, \tau_2\right) - \widehat{C}_{3s}\left(\tau_1, \tau_2\right)\right]^2,
\tag{17}
$$

where $\mathbf{h}$ is the unknown column matrix composed of the elements in time domain kernels $h(i, j)$. Due to the symmetry properties of the third-order cumulants, namely, $C_{3y}(\tau_1, \tau_2) = C_{3y}(\tau_2, \tau_1)$, the loop variable τ_2 in (17) increases from τ_1 to q.

For HGA, when the Gaussian noise is zero-mean i.i.d. signal, the following constraint condition can be used:

$$
\left|M_{1y} - \widehat{M}_{1s}\right| \leq \varepsilon,
\tag{18}
$$

where ε denotes the admissible error of the constraint condition. Comparing (7) with (15), it is seen that the introduction of the constraint (18) can guarantee the small deviation of the diagonal elements $h(i, i)$ in kernels $h(i, j)$ to be identified.

TABLE 1: Identified results [E1: diagonal element $h(i, i)$].

	True values	SNR = 0 dB	SNR = 20 dB	SNR = 40 dB	SNR = ∞
		Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
$h(0,0)$	1.00	0.9437 ± 0.1250	0.9877 ± 0.0913	0.9930 ± 0.0857	0.9962 ± 0.0685
$h(1,1)$	−0.30	−0.2460 ± 0.1368	−0.3018 ± 0.1124	−0.2997 ± 0.1138	−0.3010 ± 0.1145
$h(2,2)$	−0.85	−0.7958 ± 0.1394	−0.8328 ± 0.1003	−0.8377 ± 0.1001	−0.8390 ± 0.0872
$h(3,3)$	0.66	0.6053 ± 0.1356	0.6557 ± 0.0981	0.6540 ± 0.0990	0.6534 ± 0.0953
F		$0.751054E - 01$	$0.316731E - 02$	$0.314600E - 02$	$0.357726E - 02$

TABLE 2: Identified results [E1: full element $h(i, j)$].

	True values	SNR = 0 dB	SNR = 20 dB	SNR = 40 dB	SNR = ∞
		Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
$h(0,0)$	1.00	0.9568 ± 0.0631	0.9745 ± 0.0575	0.9783 ± 0.0610	0.9731 ± 0.0501
$h(0,1)$	0	−0.0161 ± 0.0768	−0.0162 ± 0.0712	0.0001 ± 0.0712	−0.0071 ± 0.0717
$h(0,2)$	0	0.0054 ± 0.0658	−0.0038 ± 0.0602	−0.0045 ± 0.0645	−0.0083 ± 0.0624
$h(0,3)$	0	−0.0147 ± 0.0426	−0.0061 ± 0.0501	−0.0162 ± 0.0517	−0.0186 ± 0.0533
$h(1,1)$	−0.30	−0.2752 ± 0.0687	−0.2928 ± 0.0706	−0.2935 ± 0.0733	−0.2948 ± 0.0668
$h(1,2)$	0	−0.0484 ± 0.0785	−0.0212 ± 0.0693	−0.0405 ± 0.0600	−0.0294 ± 0.0635
$h(1,3)$	0	0.0012 ± 0.0676	−0.0012 ± 0.0579	−0.0063 ± 0.0597	−0.0047 ± 0.0570
$h(2,2)$	−0.85	−0.8289 ± 0.0668	−0.8324 ± 0.0678	−0.8348 ± 0.0713	−0.8328 ± 0.0598
$h(2,3)$	0	−0.0096 ± 0.0782	−0.0191 ± 0.0627	−0.0143 ± 0.0579	−0.0044 ± 0.0632
$h(3,3)$	0.66	0.6448 ± 0.0604	0.6543 ± 0.0627	0.6528 ± 0.0639	0.6534 ± 0.0516
F		$0.693044E - 01$	$0.201256E - 02$	$0.184063E - 02$	$0.259258E - 02$

4. Numerical Experiment

To examine the proposed method, two quadratic nonlinear Volterra systems, E1 by (19) and E2 by (20), are herein considered as

$$y(n) = x^2(n) - 0.30x^2(n-1) - 0.85x^2(n-2)$$
$$+ 0.66x^3(n-3), \tag{19}$$

$$y(n) = x^2(n) - 0.3x(n)x(n-1) + 0.5x(n)x(n-2)$$
$$- 0.26x(n)x(n-3) - 0.30x^2(n-1)$$
$$+ 0.36x(n-1)x(n-2)$$
$$- 0.28x(n-1)x(n-3) - 0.85x^2(n-2)$$
$$- 0.2x(n-2)x(n-3) + 0.66x^2(n-3). \tag{20}$$

In (19) and (20), n is used to replace time t, to denote discrete time.

Because (19) is of completely squared terms, the coefficients of each term are the true solutions. Therefore, the true solutions of E1 are

$$h(i, i) = \begin{bmatrix} 1.00 & -0.30 & -0.85 & 0.66 \end{bmatrix}. \tag{21}$$

In (20), the coefficient of a squared term is the true solution; however, half of the coefficient of a cross term is the true solution. Hence the true values of E2 are

$$h(i, j) = \begin{bmatrix} 1.00 & -0.15 & 0.25 & -0.13 \\ & -0.30 & 0.18 & -0.14 \\ & & -0.85 & -0.10 \\ \text{sym} & & & 0.66 \end{bmatrix}. \tag{22}$$

The input signal $x_0(n)$ is generated by RAND command in MATLAB. Their lengths are 1024 and 8192, and each has 50 groups. Then the generated uniform white noises are scaled as independent identically distributed (i.i.d.) signal with variance 1. Assuming that the memory length of the model is $M = q + 1 = 4$, the simulation is under the condition of 50 Monte-Carlo runs. At the same time, the different values of the signal-to-noise ratio (SNR) defined by (23) are considered [20]:

$$\text{SNR} = 10 \cdot \log_{10}\left[\frac{E\left(y^2(t)\right)}{E\left(g^2(t)\right)}\right]. \tag{23}$$

Only the diagonal elements exist in the time domain kernels $h(i, i)$ of E1; in other words, the system is a simplified Volterra Model: Hammerstein Model. The identification results are listed in Table 1 for four cases of SNR = 0, 20, 40, ∞. At the same time, to assess the universality of the method, the time domain $h(i, i)$ is identified by treating it as a full matrix with element $h(i, j)$. The results are listed in Table 2. The time

TABLE 3: Identified results [E2: full element $h(i, j)$].

	True values	SNR = 0 dB Mean ± SD	SNR = 20 dB Mean ± SD	SNR = 40 dB Mean ± SD	SNR = ∞ Mean ± SD
$h(0, 0)$	1.00	0.9793 ± 0.0689	1.0118 ± 0.0523	1.0138 ± 0.0546	1.0191 ± 0.0507
$h(0, 1)$	−0.15	−0.1713 ± 0.0945	−0.1395 ± 0.0593	−0.1364 ± 0.0543	−0.1319 ± 0.0518
$h(0, 2)$	0.25	0.2185 ± 0.1070	0.2273 ± 0.0789	0.2324 ± 0.0768	0.2282 ± 0.0773
$h(0, 3)$	−0.13	−0.0719 ± 0.0953	−0.0890 ± 0.0604	−0.0907 ± 0.0609	−0.0972 ± 0.0564
$h(1, 1)$	−0.30	−0.3062 ± 0.0699	−0.3024 ± 0.0604	−0.3071 ± 0.0620	−0.3167 ± 0.0599
$h(1, 2)$	0.18	0.2247 ± 0.0913	0.1723 ± 0.0654	0.1765 ± 0.0580	0.1812 ± 0.0529
$h(1, 3)$	−0.14	−0.0892 ± 0.1079	−0.1223 ± 0.0693	−0.1246 ± 0.0704	−0.1183 ± 0.0685
$h(2, 2)$	−0.85	−0.8162 ± 0.0606	−0.8569 ± 0.0660	−0.8494 ± 0.0662	−0.8467 ± 0.0603
$h(2, 3)$	−0.10	−0.0878 ± 0.0777	−0.1008 ± 0.0520	−0.1014 ± 0.0518	−0.1001 ± 0.0513
$h(3, 3)$	0.66	0.6423 ± 0.0722	0.6586 ± 0.0707	0.6558 ± 0.0635	0.6567 ± 0.0611
F		$0.107841E + 00$	$0.351261E - 02$	$0.340253E - 02$	$0.384778E - 02$

domain kernels $h(i, j)$ of E2 are originally full matrix with the elements, and the E2's results are listed in Table 3.

It is seen from Tables 1–3 that the proposed method is suitable to well identify the quadratic nonlinear system under consideration. Table 1 shows that, when SNR = 0 dB, the relative errors of the four values of the time domain kernels $h(i, i)$ are basically within 10%, and the standard deviation is between 0.12 and 0.14. With the increase of the signal-to-noise ratio, the relative errors are gradually reduced to less than 2%, and the standard deviation is to 0.06–0.11; the objective function value (sum of the square of the third-order cumulants errors) is also reduced. The identification error becomes large with the small SNR; this is due to the fact that there exists certain error between Gaussian noise signal generated by the numerical method and the real Gaussian distribution, which leads a direct effect of the identification results.

In Table 2, the same sequence to the signal data as Table 1 is used. The identified mean values of the nondiagonal elements are generally under 0.02 (the real value is 0), and the standard deviation is below 0.08. This is very important; when the blind identification method is used to identify a quadratic system, we do not know the length of the memory, or whether the time domain kernels $h(i, j)$ have only the diagonal elements. By using different memory length to identify an identical system, we obtained very similar results, which show that this method has strong universality.

5. Blind Identification of Quadratic Volterra System of Number 4 Hydroturbine Shaft in a Certain Hydropower Station

The number 4 hydroturbine generating set shaft system in a certain hydropower station, as shown in Figure 1(a), can be simplified as a second-order Volterra system (shown in Figure 1(b)). The observed data of the lateral throw $y_2(n)$ at the turbine guide bearing in the experiment is shown in Figure 3 (after filtration). These observed signals are taken as the output signals $y(n)$, and this engineering problem is identified according to the above proposed method. Note that, for the problem in which the input signals are unknown,
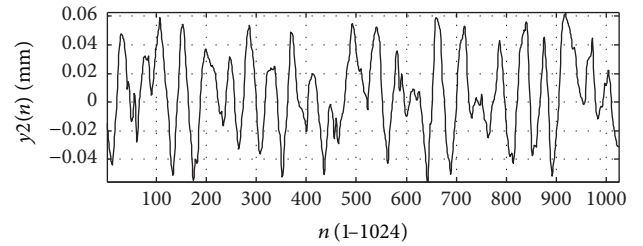

FIGURE 3: Transient waveshape of the lateral throw at turbine guide bearing (filtered).

there is a precondition here: the input must be zero-mean, i.i.d., stationary signals, and their statistics are known.

According to the 50 groups of the zero-mean, i.i.d. stationary signals generated by MATLAB, the statistics of the input signals are chosen as follows:

$$\gamma_{2x} = E\left[x^2(n)\right] = 1.000,$$
$$\gamma_{3x} = E\left[x^3(n)\right] = 2.645 \times 10^{-2},$$
$$\gamma_{4x} = E\left[x^4(n)\right] = 1.800,$$
$$\gamma_{6x} = E\left[x^6(n)\right] = 3.900.$$

(24)

The memory length of the system is taken as $M_2 = q_2 + 1 = 6$, so the time domain kernels parameters (optimization variables) to be identified are

$$h(i, j) = \begin{bmatrix} h(0,0) & h(0,1) & h(0,2) & h(0,3) & h(0,4) & h(0,5) \\ & h(1,1) & h(1,2) & h(1,3) & h(1,4) & h(1,5) \\ & & h(2,2) & h(2,3) & h(2,4) & h(2,5) \\ & & & h(3,3) & h(3,4) & h(3,5) \\ & & & & h(4,4) & h(4,5) \\ & & & & & h(5,5) \end{bmatrix}.$$
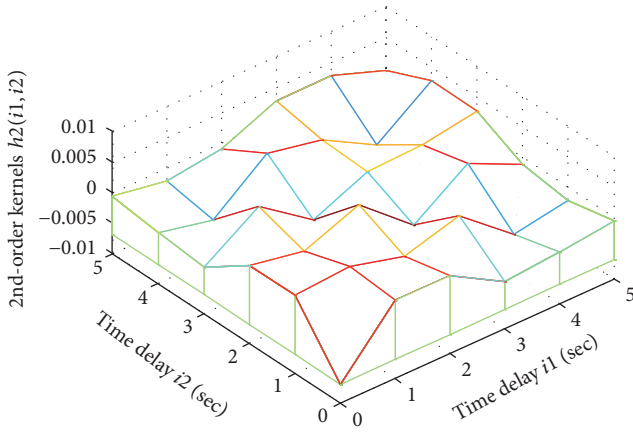
(25)

FIGURE 4: The second-order time domain kernels of the shaft system in overspeed test.

The system becomes an optimization problem with 21 design variables. When using HGA, we set the population as 50, crossover probability 50%, mutation probability 5%, and local search probability 5%. The lower and upper bounds of the variables are $[-100, 100]$, and the increment of the discrete variables is 0.01. The interval of each variable is narrowed down after some test run, and the increment of the discrete variables is ultimately decreased to 0.00001. The number of the evolution generations is one million.

After identification, the time domain kernels parameters $h(i, j)$ of the Volterra system for this shaft system are obtained, as shown in Figure 4. The frequency domain kernels parameters (Generalized Frequency Response Function, GFRF) can be obtained by two-dimensional Fourier Transform, as shown in Figure 5.

So far, by use of the blind identification method proposed in this paper, the blind identification of the system was successfully carried out, based on the observed throw signals at the turbine guide bearing.

6. Reversely Recursive Solution to Hydraulic Excitation

6.1. Reversely Recursive Algorithm. The hydraulic excitation acted on the hydraulic turbine runner is very important in the design and calculation of shaft system in a hydroelectric generating set but difficult to be observed. A reversely recursive algorithm is proposed in the paper, which can be used to approximately find out the input signal of the system based on the following two sets of data: (1) the observed output signal $y_0(n)$ of the system; (2) the system time domain kernels $h(i, j)$ obtained through blind identification.

Expanding the second-order discretized time domain Volterra system (1), the specific expression of $y(n)$ can be obtained. When $n = 0$, from the following expression:

$$y(0) = h(0, 0) x(0) x(0).$$ (26)

We can get

$$x(0) = \pm \sqrt{\frac{y_0(0)}{h(0, 0)}},$$ (27)

where $y_0(0)$ is the observed output signal of the system. $x(0)$ thus obtained has two roots.

When $n = 1$,

$$y(1) = h(0, 0) x(1) x(1) + h(0, 1) x(1) x(0)$$
$$+ h(1, 0) x(0) x(1) + h(1, 1) x(0) x(0)$$ (28)

in which $x(0)$ is known and $x(1)$ is unknown. Equation (28) becomes a quadratic equation of $x(1)$:

$$h(0, 0) x(1) x(1)$$
$$+ [h(0, 1) x(0) + h(1, 0) x(0)] x(1)$$ (29)
$$+ h(1, 1) x(0) x(0) - y(1) = 0.$$

Set $a = h(0, 0)$, $b = h(0, 1)x(0) + h(1, 0)x(0) = 2h(0, 1)x(0)$, and $c = h(1, 1)x(0)x(0) - y_0(1)$; then

$$a \cdot x(1) x(1) + b \cdot x(1) + c = 0.$$ (30)

So

$$x(1) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$ (31)

Because $x(0)$ has two roots, substituting, respectively, the two roots into (31), we can obtain the four roots of $x(1)$. The question now is which roots should $x(0)$ and $x(1)$ take? The answer is obvious. First, calculate $y(0)$ and $y(1)$ through (26) and (28), then calculate the errors of $y(0) - y_0(0)$ and $y(1) - y_0(1)$, and finally compare the errors. The roots with minimum error are the true solution. So we take

$$F = [y(1) - y_0(1)]^2 + [y(0) - y_0(0)]^2$$ (32)

as the optimal objective function and find out the minimum value $F_{\min}$ from several values of objective function F, and the corresponding $x(0)$ and $x(1)$ are taken as the true solution (input signal).

When the previous values of $x(n - 2)$ and the $x(\cdot)$ have been confirmed, the general formulae for calculating $x(n)$ are as follows:

$$y(n) = h(0, 0) x(n - 0) x(n - 0)$$
$$+ h(0, 1) x(n - 0) x(n - 1) + \cdots$$
$$+ h(0, q) x(n - 0) x(n - q)$$
$$+ h(1, 0) x(n - 1) x(n - 0)$$
$$+ h(1, 1) x(n - 1) x(n - 1) + \cdots$$ (33)
$$+ h(1, q) x(n - 1) x(n - q) + \cdots$$
$$+ h(q, 0) x(n - q) x(n - 0)$$
$$+ h(q, 1) x(n - q) x(n - 1) + \cdots$$
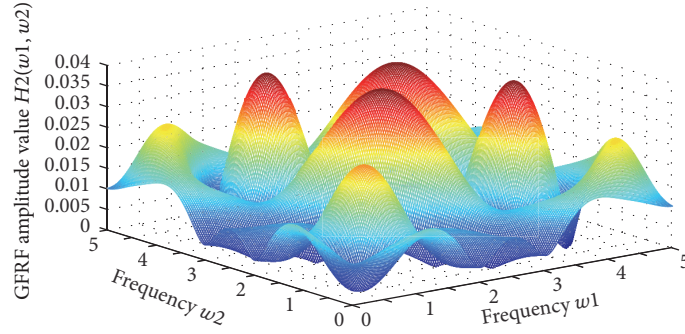$$+ h(q, q) x(n - q) x(n - q).$$

FIGURE 5: The frequency domain kernels of the shaft system in overspeed test.

Let $a = h(0, 0)$, $b = \sum_{i=1}^{q} 2 \cdot h(0, i) x(n - i)$, and $c = [\sum_{i=1}^{q} \sum_{j=1}^{q} h(i, j) x(n - i) x(n - j)] - y_0(n)$; thus

$$x(n) = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{34}$$

Because $x(n - 1)$ has two values, we can obtain the four roots of $x(n)$. In order to determine $x(n - 1)$ and $x(n)$, we still use optimization method. The objective function for optimization is

$$F = \sum_{i=0}^{q} [y(n - i) - y_0(n - i)]^2. \tag{35}$$

Finding the minimum value $F_{\min}$, the corresponding $x(n - 1)$ and $x(n)$ can be taken as the input signal. Thus through recursive calculation, we can find out $x(0) \sim x(N)$; here $N + 1$ is the sampling length. The above method is the reversely recursive algorithm.

6.2. Numerical Experiment E2. For convenience, we still use the experimental data E2 in Section 4. According to the above method, the calculated input data $x(n)$ can be obtained, as shown in Figure 6, where $xx(n)$ is the calculated result and $x0(n)$ is the original input signal.

To illustrate the correctness and feasibility of the calculation results, substituting in (1) the calculated results $xx(n)$ and the original input signal $x0(n)$, respectively, we can find out the output signal $y(n)$ and the original output signal $y0(n)$ and their differences:

$$dy(n) = |y(n) - y_0(n)| \tag{36}$$

and draw in Figure 7. At the same time, the relative error of output signal energy is calculated by using the following formula:

$$R_E = \frac{\sum y^2(n) - \sum y_0^2(n)}{\sum y_0^2(n)} \times 100\%. \tag{37}$$

From Figures 6 and 7 we can see that, by the reversely recursive method proposed in this paper, we can approximately find out the input signal of the system based on the system output signal. The calculated input signal $xx(n)$ is slightly
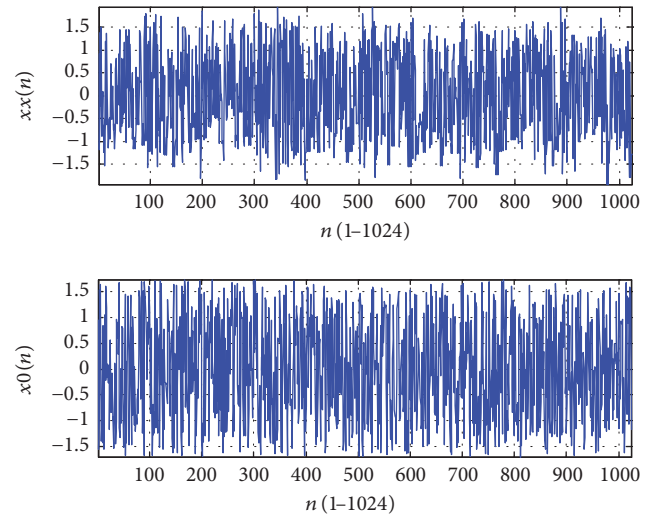

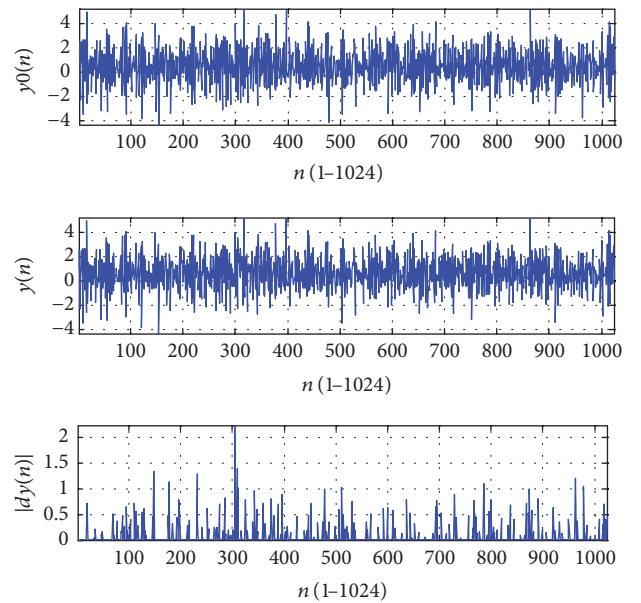
FIGURE 6: E2 numerical experiment results (input signals).

FIGURE 7: The comparison of the calculated output signals to the original observed signals.
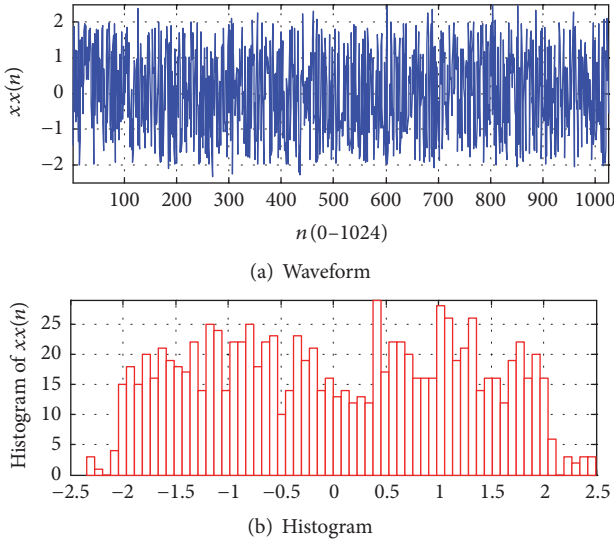
(a) Waveform

(b) Histogram

FIGURE 8: The input signals calculated by recursive method and their histogram.

(a) $y(n)$ waveform

(b) Errors compared with the observed signal

FIGURE 9: The output signals $y(n)$ calculated by $xx(n)$ and their errors.

different from the original input signal $x0(n)$, which is mainly caused by the quadratic nonlinearity of the system. However, the errors $dy(n)$ between the system output signal $y(n)$ calculated through the input signal $xx(n)$ and the original output signal $y0(n)$ are smaller (Figure 7). Only a few signal values have errors, and most output signal values are the same as the original value. The relative error of the output signal energy by (37) is only 3.96%. Therefore, the input data $xx(n)$ thus calculated can be approximately used as the system input signal.

6.3. Reversely Recursive Solution to Hydraulic Excitation on Turbine. The hydraulic excitation (input signal) $xx(n)$ on the turbine runner and its distribution can be calculated recursively using the proposed method in Section 6.1, as shown in Figure 8.

Since there is no real input signals that can be compared, we cannot tell the difference between the input signal $xx(n)$ obtained with recursive method and the real input signals. However, we can compare the calculated output signal $y(n)$ with the real output signal $y_0(n)$. Figure 9 shows the calculated output signal $y(n)$ and the absolute errors with observed output signal $y_0(n)$. From the figure we can see that the error is not very big. Besides, the relative error of the output signal energy was calculated, whose value is 8.95%. This shows that the method is feasible; that is to say, the input signal $xx(n)$ is acceptable and can be approximately taken as the hydraulic excitation on the turbine runner of a hydroelectric generating set shaft system.

7. Conclusions

A blind identification method based on the third-order cumulants is proposed for a quadratic Volterra nonlinear system. Through three numerical experiments and engineering application, the following conclusions are drawn.
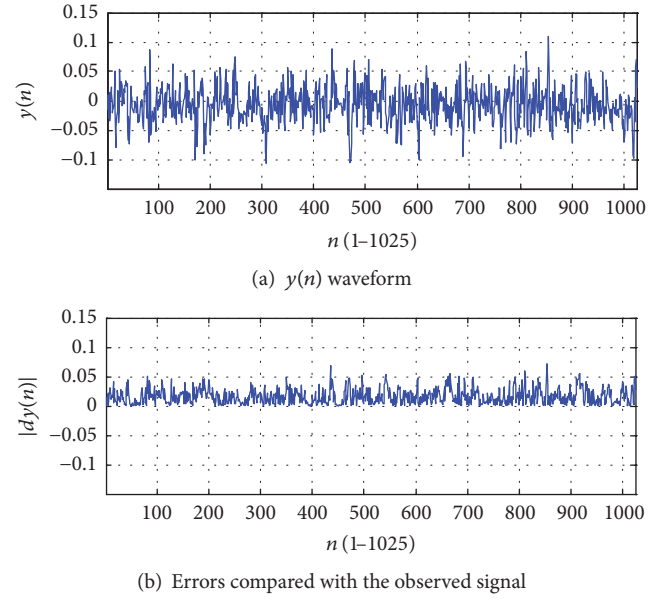
(1) Two different types of the memory length are used to identify the same sets of the signal data sequence of hydroturbine main shaft, and the error between the two groups of the time domain kernels $h(i, j)$ is small. It shows that the proposed method is applicable to the blind identification of the hydroturbine main shaft system and has strong universality.

(2) The proposed computer loop judgment method to determine the coefficient γ_{ijk} for calculating the third-order moment of the output signal greatly simplifies the formulae derivation and compiling of the programming code and is suitable for calculating the fourth- or higher order cumulants.

(3) With the gradual increase of the signal-to-noise ratio, the relative error of the identified time domain kernels decreases, the standard deviation is reduced, and the objective function value is also reduced.

(4) According to the observed signal of the shaft system of a certain hydroturbine generating set, the time domain kernels of the Volterra system can be blindly identified by the method proposed in this paper. Furthermore, according to these time domain kernels and the observed signal, the input signal of the system can be recursively calculated. The results show that the computed input signal can be approximately taken as the stochastic excitation acted on the turbine runner, and the error is not big (relative error of the output signal energy is acceptable). The proposed method is of enlightening significance for the form of hydraulic action.

Appendix

Loop Judgment Method for Determining Coefficient Value

In the expressions of the third-order moment (10), suppose the six subscripts of input sequence $\{x(t)\}$ are, respectively, $t_1 = -i_1, t_2 = -i_2, t_3 = \tau_1 - j_1, t_4 = \tau_1 - j_2, t_5 = \tau_2 - k_1$, and $t_6 = \tau_2 - k_2$; then the coefficient values γ_{ijk} can be determined by the loop judgment method from computer.

Step 1. If $t_1 = t_2 = t_3 = t_4 = t_5 = t_6$, obviously $\gamma_{ijk} = \gamma_{6x}$. If not satisfied, go to Step 2.

Step 2. If $t_1 = t_2 = t_3 \neq t_4 = t_5 = t_6$, then $\gamma_{ijk} = \gamma_{3x}^2$. In this condition, there are totally 10 kinds of combinations. These combinations can be expressed (only use subscripts) as 123 and 456; 124 and 356; 125 and 346; 126 and 345; 134 and 256; 135 and 246; 136 and 245; 145 and 236; 146 and 235; 156 and 234. If not satisfied, go to Step 3.

Step 3. if $t_1 = t_2 \neq t_3 = t_4 \neq t_5 = t_6$, then $\gamma_{ijk} = \gamma_{2x}^3$; Furthermore, if $t_1 = t_2 \neq t_3 = t_4 = t_5 = t_6$, then $\gamma_{ijk} = \gamma_{4x}\gamma_{2x}$.

First, divide the six time t_i into two groups: $2 + 4$. For example, when $t_1 = t_2$, then to judge whether t_3, t_4, t_5, and t_6 are equal, there are three possibilities.

(1) All four times t_3, t_4, t_5, and t_6 equal each other; that is, $t_3 = t_4 = t_5 = t_6$; then $\gamma_{ijk} = \gamma_{4x}\gamma_{2x}$.

(2) Two times are equal; say $t_3 = t_4$ and $t_5 = t_6$, and in the two groups there is one group equal to t_1; then $\gamma_{ijk} = \gamma_{4x}\gamma_{2x}$.

(3) Two times are equal; say $t_3 = t_4, t_5 = t_6$, and in the two groups there is no one group equal to t_1; then $\gamma_{ijk} = \gamma_{2x}^3$.

This grouping can be distinguished by the following five kinds of combinations; that is, the 2 in $2 + 4$ are $t_1 = t_2, t_1 = t_3, t_1 = t_4, t_1 = t_5$, and $t_1 = t_6$, and the rest is 4. In each combination, the previous comparison and judgment should be repeated.

If the previous conditions are not satisfied, go to Step 4.

Step 4. If Steps 1, 2, and 3 are not met, $\gamma_{ijk} = 0$.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] R. Maheswaran and R. Khosa, "Wavelet-Volterra coupled model for monthly stream flow forecasting," *Journal of Hydrology*, vol. 450-451, pp. 320–335, 2012.

[2] R. Maheswaran and R. Khosa, "Long term forecasting of groundwater levels with evidence of non-stationary and nonlinear characteristics," *Computers & Geosciences*, vol. 52, pp. 422–436, 2013.

[3] V. S. Hari, V. P. Jagathy Raj, and R. Gopikakumari, "Unsharp masking using quadratic filter for the enhancement of fingerprints in noisy background," *Pattern Recognition*, vol. 46, no. 12, pp. 3198–3207, 2013.

[4] W. Ji and W.-S. Gan, "Identification of a parametric loudspeaker system using an adaptive Volterra filter," *Applied Acoustics*, vol. 73, no. 12, pp. 1251–1262, 2012.

[5] Y. Kim, J.-H. Kim, and Y. Kim, "Time series prediction of nonlinear ship structural responses in irregular seaways using a third-order Volterra model," *Journal of Fluids and Structures*, vol. 49, pp. 322–337, 2014.

[6] L. Carassale and A. Kareem, "Synthesis of multi-variate Volterra systems by a topological assemblage scheme," *Probabilistic Engineering Mechanics*, vol. 37, pp. 109–122, 2014.

[7] Z. Sigrist, E. Grivel, and B. Alcoverro, "Estimating second-order Volterra system parameters from noisy measurements based on an LMS variant or an errors-in-variables method," *Signal Processing*, vol. 92, no. 4, pp. 1010–1020, 2012.

[8] J. Liu, Y. Zhang, and B. Yun, "A new method for predicting nonlinear structural vibrations induced by ground impact loading," *Journal of Sound and Vibration*, vol. 331, no. 9, pp. 2129–2140, 2012.

[9] H. Tang, Y. H. Liao, J. Y. Cao, and H. Xie, "Fault diagnosis approach based on Volterra models," *Mechanical Systems and Signal Processing*, vol. 24, no. 4, pp. 1099–1113, 2010.

[10] J. M. Nichols and K. D. Murphy, "Modeling and detection of delamination in a composite beam: a polyspectral approach," *Mechanical Systems and Signal Processing*, vol. 24, no. 2, pp. 365–378, 2010.

[11] J. M. Nichols and C. C. Olson, "Optimal bispectral detection of weak, quadratic nonlinearities in structural systems," *Journal of Sound and Vibration*, vol. 329, no. 8, pp. 1165–1176, 2010.

[12] Z. K. Peng, Z. Q. Lang, L. Zhao, S. A. Billings, G. R. Tomlinson, and P. F. Guo, "The force transmissibility of MDOF structures with a non-linear viscous damping device," *International Journal of Non-Linear Mechanics*, vol. 46, no. 10, pp. 1305–1314, 2011.

[13] C. M. Cheng, Z. K. Peng, X. J. Dong, W. M. Zhang, and G. Meng, "A novel damage detection approach by using Volterra kernel functions based analysis," *Journal of the Franklin Institute*, vol. 352, no. 8, pp. 3098–3112, 2015.

[14] Z. K. Peng, W. M. Zhang, B. T. Yang, G. Meng, and F. L. Chu, "The parametric characteristic of bispectrum for nonlinear systems subjected to Gaussian input," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 456–470, 2013.

[15] L. M. Li and S. A. Billings, "Analysis of nonlinear oscillators using Volterra series in the frequency domain," *Journal of Sound and Vibration*, vol. 330, no. 2, pp. 337–355, 2011.

[16] A. Chatterjee, "Identification and parameter estimation of a bilinear oscillator using Volterra series with harmonic probing," *International Journal of Non-Linear Mechanics*, vol. 45, no. 1, pp. 12–20, 2010.

[17] W.-D. Chang, "Volterra filter modeling of nonlinear discrete-time system using improved particle swarm optimization," *Digital Signal Processing*, vol. 22, no. 6, pp. 1056–1062, 2012.

[18] J.-M. Le Caillec, "Spectral inversion of second order Volterra models based on the blind identification of Wiener models," *Signal Processing*, vol. 91, no. 11, pp. 2541–2555, 2011.

[19] M. Ghandchi Tehrani, L. Wilmshurst, and S. J. Elliott, "Receptance method for active vibration control of a nonlinear system," *Journal of Sound and Vibration*, vol. 332, no. 19, pp. 4440–4449, 2013.

[20] J. Antari, A. Zeroual, and S. Safi, "Blind identification of nonlinear quadratic systems using higher order cumulants and non-Gaussian input signals," *International Journal of General Systems*, vol. 37, no. 5, pp. 585–602, 2008.

[21] J. Antari, S. Chabaa, R. Iqdour, A. Zeroual, and S. Safi, "Identification of quadratic systems using higher order cumulants and neural networks: application to model the delay of video-packets transmission," *Applied Soft Computing*, vol. 11, no. 1, pp. 1–10, 2011.

[22] H.-Z. Tan and T. W. S. Chow, "Blind identification of quadratic nonlinear models using neural networks with higher order cumulants," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 3, pp. 687–696, 2000.