Volterra series truncation and kernel estimation of nonlinear systems in the frequency domain

B. Zhang [a,b,*], S.A. Billings [a]

[a] Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK
[b] Nuclear AMRC, The University of Sheffield, Advanced Manufacturing Park, Brunel Way, Rotherham S60 5WG, UK

ARTICLE INFO

ABSTRACT

The Volterra series model is a direct generalisation of the linear convolution integral and is capable of displaying the intrinsic features of a nonlinear system in a simple and easy to apply way. Nonlinear system analysis using Volterra series is normally based on the analysis of its frequency-domain kernels and a truncated description. But the estimation of Volterra kernels and the truncation of Volterra series are coupled with each other. In this paper, a novel complex-valued orthogonal least squares algorithm is developed. The new algorithm provides a powerful tool to determine which terms should be included in the Volterra series expansion and to estimate the kernels and thus solves the two problems all together. The estimated results are compared with those determined using the analytical expressions of the kernels to validate the method. To further evaluate the effectiveness of the method, the physical parameters of the system are also extracted from the measured kernels. Simulation studies demonstrates that the new approach not only can truncate the Volterra series expansion and estimate the kernels of a weakly nonlinear system, but also can indicate the applicability of the Volterra series analysis in a severely nonlinear system case.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Volterra series [1] have been used for the modelling and analysis of nonlinear systems in many industries such as marine [2], automotive [3], structural [4], biological [5], and communication systems [6]. The Volterra model is a direct generalisation of the linear convolution integral and provides an intuitive system representation. The multidimensional Fourier transform of the Volterra kernels is a natural extension of the linear frequency response function to the nonlinear case and is often referred to as the Generalised Frequency Response Functions (GFRFs). The GFRFs have received much more research interest over the time-domain Volterra kernels. This is because important nonlinear phenomena such as harmonics, intermodulation and gain expansion/depression can easily be explained by the interactions between different frequency components and orders of these GFRFs [7].

The GFRFs of nonlinear systems can be determined by either a parametric-model-based method or a nonparametric-model-based method [8]. In the parametric approach, a nonlinear parametric model is first identified from the input–output data. The GFRFs are then obtained by mapping the resultant model into the frequency domain using the probing method [9].

---

* Corresponding author at: Nuclear AMRC, The University of Sheffield, Advanced Manufacturing Park, Brunel Way, Rotherham S60 5WG, UK.
   E-mail addresses: bzhang347@gmail.com (B. Zhang), s.billings@sheffield.ac.uk (S.A. Billings).

The nonparametric approach is often referred to as frequency-domain Volterra system identification and is based on the observation that the Volterra model of nonlinear systems is linear in terms of the unknown Volterra kernels, which, in the frequency domain, corresponds to a linear relation between the output frequency response and linear, quadratic, and higher order GFRFs. This linear relationship allows the use of a least squares (LS) approach to solve for the GFRFs. Several researchers [10–12] have used this method to estimate the GFRFs. But they usually made the assumption that it is known a priori that the system under study can be represented by just two or three terms. However, such information is rarely available a priori.

It is well known that the Volterra series cannot represent severely non-linear systems. And even for a weakly nonlinear system, the order of the Volterra series expansion to achieve an approximation accuracy may still be very high. This indicates that the estimation of the GFRFs is related to the truncation of the Volterra series expansion. And because nonlinear system analysis using Volterra series is usually based on a truncated description, the study on the truncation of the Volterra series expansion is important. Although Billings and Lang [13] proposed an algorithm to truncate Volterra series representations, the algorithm makes an assumption that the GFRFs are known a priori or they can be obtained from the time-domain model, which is, however, not practical in many cases.

In this paper, a novel approach utilising a complex-valued orthogonal least squares (OLS) algorithm regularised by an adjustable prediction error sum of squares (APRESS) criterion will be developed for both the truncation of the Volterra series expansion and the estimation of the GFRFs.

2. Volterra modelling of nonlinear systems in the time and frequency domain

The output $y(t)$ of a single input single output (SISO) analytical system can be expressed as a Volterra functional polynomial of the input $u(t)$ to give

$$y(t) = \sum_{n=1}^{\bar{N}} y^{(n)}(t) \tag{1}$$

where $\bar{N}$ is the maximum order of the system nonlinearity and $y^{(n)}(t)$ is the nth-order output of the system, which is given by

$$y^{(n)}(t) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h_n(\tau_1, \cdots, \tau_n) \prod_{i=1}^{n} u(t - \tau_i) d\tau_i, \; n \geq 1 \tag{2}$$

where $h_n(\tau_1, \cdots, \tau_n)$ is a real valued function of $\tau_1, \cdots, \tau_n$ called the nth order impulse response function or Volterra kernel of the system [1]. Volterra generalised the linear convolution concept to deal with nonlinear systems by replacing the single impulse response with a series of multidimensional integration kernels. The nth-order Volterra kernel describes nonlinear interactions among n copies of the input. The multidimensional Fourier transform of the nth-order Volterra kernel yields the nth-order transfer function or generalised frequency response function (GFRF)

$$H_n(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h_n(\tau_1, \cdots, \tau_n) e^{-\bar{j}(\omega_1\tau_1 + \cdots + \omega_n\tau_n)} d\tau_1 \cdots d\tau_n \tag{3}$$

which is a natural extension of the concept of the linear frequency response function to the nonlinear case. In Eq. (3), $\bar{j}$ is the imaginary unit.

The nth-order kernel and the kernel transform are not unique because an interchange of arguments in $h_n(\tau_1, \cdots, \tau_n)$ may give different kernels without affecting the input–output relationships. To ensure that the GFRFs are unique, they are symmetrised to give

$$H_n^{sym}(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) = \frac{1}{n!} \sum_{\substack{all\ permulations \\ of\ \{\omega_1, \cdots, \omega_n\}}} H_n^{asym}(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) \tag{4}$$

Using the concept of GFRF, the general relationship between the input spectrum $U(\bar{j}\omega)$ and the output spectrum $Y(\bar{j}\omega)$ can be obtained as

$$Y(\bar{j}\omega) = \sum_{n=1}^{\bar{N}} \frac{1}{\sqrt{n}(2\pi)^{n-1}} \int_{\omega_1 + \cdots + \omega_n = \omega} H_n(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) \prod_{i=1}^{n} U(\bar{j}\omega_i) d\sigma_\omega \tag{5}$$

where $\int_{\omega_1 + \cdots + \omega_n = \omega} (\cdot) d\sigma_\omega$ denotes the integration of $(\bullet)$ over the n-dimensional hyperplane $\omega_1 + \cdots + \omega_n = \omega$.

When the system is subject to a harmonic input such as

$$u(t) = |A|\cos(\Omega t + \angle A) \tag{6}$$

the output spectrum at the driving frequency can be expressed as [4]

$$
Y(\bar{j}\Omega) = \sum_{n=1,3,\cdots,\bar{N}} \frac{1}{2^n} C\left(n, \left\lfloor \frac{n}{2} \right\rfloor\right) |A|^{2\left\lfloor \frac{n}{2} \right\rfloor} A H_{n,\left\lfloor \frac{n}{2} \right\rfloor}
\tag{7}
$$

where $\left\lfloor \frac{n}{2} \right\rfloor$ denotes the floor function, which gives the largest integer less than or equal to $\frac{n}{2}$, $H_{n,\left\lfloor \frac{n}{2} \right\rfloor}\left(\bar{j}\Omega, \cdots \bar{j}\Omega, -\bar{j}\Omega, \cdots, -\bar{j}\Omega\right)$ is a higher-order GFRF with $n - \left\lfloor \frac{n}{2} \right\rfloor$ arguments of Ω and $\left\lfloor \frac{n}{2} \right\rfloor$ arguments of $-\Omega$, and $C\left(n, \left\lfloor \frac{n}{2} \right\rfloor\right)$ is the number of combinations of $\left\lfloor \frac{n}{2} \right\rfloor$ objects from a set with n objects and given as

$$
C\left(n, \left\lfloor \frac{n}{2} \right\rfloor\right) = \frac{n!}{\left\lfloor \frac{n}{2} \right\rfloor! \left(n - \left\lfloor \frac{n}{2} \right\rfloor\right)!}
\tag{8}
$$

Eq. (7) can also be written as

$$
Y(\bar{j}\Omega) = \sum_{j=1}^{\bar{N}} Y_{2j+1}
\tag{9}
$$

where

$$
\bar{N} = \left\lfloor \frac{\bar{N}}{2} \right\rfloor
\tag{10}
$$

$$
Y_{2j+1} = \frac{1}{2^{2j+1}} C(2j+1, j) |A|^{2j} A H_{2j+1,j}
\tag{11}
$$

is the $(2j+1)$th order output spectrum component.

In Eq. (9), $Y_1 = \frac{1}{2} A H_{1,0}$ is just the output spectrum of the linear system. Thus Eq. (9) clearly demonstrates that how the output energy at the driving frequency contributed by the linear term is modified by the higher-order nonlinear effects to yield the output frequency response $Y(j\Omega)$.

3. Determination of the GFRFs

The concept of GFRF is a natural extension of the concept of the linear frequency response function to the nonlinear case and represents the characteristics of nonlinear systems in a manner which is independent of the inputs. However, GFRFs differ from the frequency response function in linear systems in two aspects. First, the frequency-domain description of a nonlinear system is associated with a sequence of GFRFs instead of only one frequency response function in the linear case. This is because the Volterra series representation of nonlinear systems involves a sequence of Volterra kernels, while GFRFs are defined as the Fourier transform of these kernels. In addition, GFRFs are multi-variable functions even when the underlying system is single-input/single-output. Although these complexities bring about difficulties in the determination of the GFRFs, various computation and estimation methods have been developed.

3.1. Computation of the GFRFs

Given a parametric model of a nonlinear system, there are a number of methods to obtain the GFRFs of the system. Arguably the most direct one is the harmonic probing method of Bedrosian and Rice [6] and Bussgang et al. [14]. In the case of SISO nonlinear systems, the basic idea of the probing method can be introduced as below.

It was shown by Rugh [15] that for nonlinear systems which are described by the Volterra model , and excited by a combination of harmonic exponentials

$$
u(t) = \sum_{i=1}^{K} e^{\bar{j}\omega_i t}, \ 1 \leq K \leq N
\tag{12}
$$

the output response can be expressed as

$$
\begin{aligned}
y(t) &= \sum_{n=1}^{N} \sum_{i_1=1}^{K} \cdots \sum_{i_n=1}^{K} H_n\left(\bar{j}\omega_{i_1}, \cdots, \bar{j}\omega_{i_n}\right) e^{\bar{j}\left(\omega_{i_1} + \cdots + \omega_{i_n}\right)t} \\
&= \sum_{n=1}^{N} \sum_{\substack{\sum_{i=1}^{K} m_i = n, m_i \geq 0}} G_{m_1 \cdots m_K}\left(\bar{j}\omega_1, \cdots, \bar{j}\omega_K\right) e^{\bar{j}t \sum_{i=1}^{K} m_i \omega_i}
\end{aligned}
\tag{13}
$$

where

$$G_{m_1 \cdots m_K}(\bar{j}\omega_1, \cdots, \bar{j}\omega_K) = \frac{n!}{m_1! \cdots m_K!} H_n\left(\underbrace{\bar{j}\omega_1, \cdots, \bar{j}\omega_1}_{m_1}, \cdots, \underbrace{\bar{j}\omega_K, \cdots, \bar{j}\omega_K}_{m_K} \right) \tag{14}$$

In most cases, Eq. (14) will contain repeated frequency arguments. In the special case where $K = n$, however, all the frequency components are distinct and namely $m_i=1$, $i = 1, \cdots, K$. Therefore,

$$G_{m_1 \cdots m_K}(\bar{j}\omega_1, \cdots, \bar{j}\omega_K) = n! H_n(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) \tag{15}$$

Considering Eq. (15), Eq. (13) can then be written as

$$y(t) = n! H_n(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) e^{\bar{j} t \sum_{i=1}^{n} \omega_i} + (terms\ with\ repeated\ frequencies) + (terms\ from\ lower\ orders) \tag{16}$$

For nonlinear systems which have a parametric model with parameter vector θ,

$$y(t) = f_0(t, \theta, y(t), u(t)) \tag{17}$$

and which can also be described by the Volterra model and , substituting Eqs. (12) and (16) into Eq. (17) for $y(t)$ and $u(t)$, and extracting the coefficient of $exp(\bar{j} t \sum_{i=1}^{n} \omega_i)$ from the resulting expression produces an equation from which the GFRF $H_n(\bar{j}\omega_1, \cdots \bar{j}\omega_n)$ can be obtained.

By using the aforementioned probing method, the GFRFs of the generalized higher-order Duffing oscillator model

$$\sum_{l=1}^{L} \sum_{\alpha=0}^{A} c(l, \alpha)(D^\alpha y)^l = u \tag{18}$$

where $D=d/dt$ denotes the differential operator, α is the order of the derivative, l is the order of the exponential and $c(l, \alpha)$ are the model coefficients, were derived as follows[16]

$$H_n(\bar{j}\omega_1, \cdots, \bar{j}\omega_n) \triangleq - \frac{\sum_{l=2}^{L} \sum_{\alpha=0}^{A} l! c(l, \alpha) \left\{ \sum_{p=1}^{|S_n^l|} S_n^l[p] \right\}}{n! \sum_{\alpha=0}^{A} \left[c(1, \alpha) \left(\bar{j} \sum_{i=1}^{n} \omega_i \right)^\alpha \right]} \qquad (n \geq L) \tag{19}$$

where S_n^l, the Stirling set of the second kind, denotes the set whose elements cover all the partitions of a set $(1,2, \cdots n)$ into l blocks, $S_n^l[p]$ denotes the pth element of S_n^l, and $|S_n^l|$, the Stirling number of the second kind, is the cardinality of the Stirling set of the second kind, and

$$\sum_{p=1}^{|S_n^l|} S_n^l[p] \triangleq \sum_{(r;l,n)} \sum_{N}' \left[r_1!\left(\bar{j}\omega_1 + \cdots + \bar{j}\omega_{r_1}\right)^\alpha H_{r_1}\left(\bar{j}\omega_1, \cdots, \bar{j}\omega_{r_1}\right) r_2!\left(\bar{j}\omega_{r_1+1} + \cdots + \bar{j}\omega_{r_1+r_2}\right)^\alpha \right.$$
$$\left. \times H_{r_2}\left(j\omega_{r_1+1}, \cdots, j\omega_{r_1+r_2}\right) \cdots r_l!\left(j\omega_\mu + \cdots + j\omega_n\right)^\alpha H_{r_l}\left(j\omega_\mu, \cdots, j\omega_n\right) \right] \tag{20}$$

In Eq. (20), $\mu = r_1 + r_2 + \cdots r_{l-1}+1 = n - r_l+1$ and $(r;l, n)$ beneath the leftmost $\sum$ denotes summation taken over those partitions of n which have l parts such that

$$r_1 + r_2 + \cdots + r_l = n, \quad r_1 \leq r_2 \leq \cdots \leq r_l \tag{21}$$

The second summation $\sum_N'$ in Eq. (20) extends over the N symmetric products. The number of terms in $\sum_N'$ is

$$N = \frac{n!}{r_1! r_2! \cdots r_l! w_1! w_2! \cdots w_k!} \tag{22}$$

where w_1 is the number of equal r's in the first run of equalities in the arrangement $r_1 \leq r_2 \leq \cdots \leq r_l$, w_2 the number in the second run, and so on. When the r's are unequal, the w's do not appear.

The generalized higher-order Duffing oscillator model represents a wide class of nonlinear systems frequently encountered in engineering. Specially, when $A = 2$, $c(1,2)=1$ and $c(l, 2)=0$ for $l \geq 2$, Eq. (18) becomes

$$\ddot{y} + \sum_{l=1}^{L} c(l, 1)\dot{y}^l + \sum_{l=1}^{L} c(l, 0)y^l = u \tag{23}$$

which represents the generalized Duffing oscillator model, in which $f_1(\bullet) = \sum_{l=1}^{L} c(l, 1)(\bullet)^l$ is the nonlinear damping polynomial function and $f_0(\bullet) = \sum_{l=1}^{L} c(l, 0)(\bullet)^l$ is the nonlinear stiffness polynomial function.

The probing method can also be extended to the single input multiple output nonlinear systems. If the system is of a

single input and two outputs and can be described by the following parametric model

$$
\begin{cases}
y_1(t) = f_1\big(t, \theta, y_1(t), y_2(t), u_1(t)\big) \\
y_2(t) = f_2\big(t, \theta, y_1(t), y_2(t), u_1(t)\big)
\end{cases}
\tag{24}
$$

Eq. (16) can be written as

$$
y(t) = n! H_{nj_1}(\bar{j}\,\omega_1, \cdots, \bar{j}\,\omega_n) e^{\bar{j}\, t \sum_{i=1}^{n} \omega_i} + \big(terms\ \ with\ repeated\ frequencies\big) + \big(terms\ from\ lower\ orders\big), \quad j_1 = 1, 2
\tag{25}
$$

Then substituting $u_1(t) = \sum_{i=1}^{n} e^{\bar{j}\,\omega_i t}$, and $y_1(t)$ and $y_2(t)$ expressed by Eq. (25) into Eq. (24), and extracting the coefficient of $exp\big(\bar{j}\, t \sum_{i=1}^{n} \omega_i\big)$ from the resulting expressions produces two coupled equations from which the GFRF matrix $\big[H_{n1}(\bar{j}\,\omega_1, \cdots \bar{j}\,\omega_n), H_{n2}(\bar{j}\,\omega_1, \cdots \bar{j}\,\omega_n) \big]$ can be obtained.

3.2. Estimation of the GFRFs

Eq. (7) shows that $Y(\bar{j}\,\Omega)$ is a function of the excitation amplitude A. Therefore, if one measures $Y(\bar{j}\,\Omega)$ for various excitation amplitudes and neglects higher-order terms, one can estimate the GFRFs [17–19]. For example, if one measures $Y_i(\bar{j}\,\Omega), i = 1, 2, \cdots, N$ for N different excitation amplitudes A_i respectively, and considers the first $\bar{N}$ terms on the right-hand side of Eq. (7), one can write the following equation,

$$
Y_i(\bar{j}\,\Omega) = \sum_{j=1}^{\bar{N}} \theta_j \phi_{ij} + \varepsilon_i
\tag{26}
$$

where

$$
\theta_j = H_{2j+1,j}
\tag{27}
$$

$$
\phi_{ij} = C(2j + 1, j)\frac{A_i^{2j+1}}{2^{2j+1}}
\tag{28}
$$

and $\varepsilon_i, i = 1, 2, \cdots, N$, is the model residual. The relationship between $\bar{N}$ and the maximum order of the system nonlinearity $\bar{\bar{N}}$ is given by Eq. (10).

Eq. (26) can also be written in the matrix form as

$$
\mathbf{Y} = \mathbf{\Phi}\mathbf{\Theta} + \mathbf{\Xi}
\tag{29}
$$

where

$$
\mathbf{Y} = \big[\, Y_1(\bar{j}\,\Omega), Y_2(\bar{j}\,\Omega), \cdots, Y_N(\bar{j}\,\Omega) \big]^T
\tag{30}
$$

$$
\mathbf{\Phi} = \big[\, \boldsymbol{\varphi_1}, \ \boldsymbol{\varphi_2}, \ \cdots, \ \boldsymbol{\varphi_{\bar{N}}} \big]
\tag{31}
$$

$$
\mathbf{\Theta} = \big[\, H_{1,0}, H_{3,1}, \cdots, H_{2\bar{N}+1,\bar{N}} \big]^T
\tag{32}
$$

$$
\mathbf{\Xi} = \big[\, \varepsilon_1, \ \cdots, \ \varepsilon_N \big]^T
\tag{33}
$$

The residual vector $\mathbf{\Xi}$ is assumed to be of zero mean and uncorrelated with $\boldsymbol{\varphi_j}, j = 1, 2, \cdots, \bar{N}$ and

$$
\boldsymbol{\varphi_j} = \big[\, \phi_{1j}, \ \cdots, \ \phi_{Nj} \big]^T
\tag{34}
$$

where $\phi_{ij}, j = 1, 2, \cdots, \bar{N}, i = 1, 2, \cdots, N$, is given by Eq. (28).

The solution of Eq. (29) can be obtained by the LS algorithm as

$$
\hat{\mathbf{\Theta}} = \big(\mathbf{\Phi}^H \mathbf{\Phi}\big)^{-1} \mathbf{\Phi}^H \mathbf{Y}
\tag{35}
$$

where $\mathbf{\Phi}^H = \bar{\mathbf{\Phi}}^T$ which is the conjugate transpose of $\mathbf{\Phi}$, $\mathbf{\Phi}^T$ denotes the transpose, and $\bar{\mathbf{\Phi}}$ denotes the matrix with complex conjugated entries.

The LS-based parameter estimation approach needs to make an assumption that the output frequency response in Eq. (7) can be truncated by $\bar{N}$ terms while $\bar{N}$ is a sufficiently large number. However, many of these candidate model terms may be redundant. The inclusion of redundant model terms often makes the model become oversensitive to the training data and is also likely to make the information matrix $\mathbf{\Phi}^H \mathbf{\Phi}$ ill-conditioned which may result in biased parameter estimates. Therefore, a truncated Volterra series expansion must be determined prior to the estimation of the GFRFs. On the other hand, Volterra

series analysis is based on a truncated description and a finite Volterra series is required in practical nonlinear system analysis. To solve these problems all together, the OLS algorithm can be used. The OLS method provides a powerful tool to select the significant model terms, determine the optimal number of model terms, and then estimate the model parameters and has already been widely applied in the identification of nonlinear systems. But because both the output frequency response and the GFRFs are complex, the complex-valued OLS algorithm is required. Several complex-valued OLS algorithms [20,21] were proposed but in forms different from the widely used real-valued algorithm. However, the OLS algorithm in itself is complex-valued. In this paper, the conventional algorithm was revisited. A unique form of the complex-valued and real-valued algorithm was presented. This can avoid confusions and help ease of use of the OLS algorithm.

4. Complex-valued orthogonal least squares algorithm

Since the $N \times \bar{N}$ $(\bar{N} \leq N)$ measured matrix $\mathbf{\Phi}$ has full column rank, it can be uniquely decomposed as

$$\mathbf{\Phi} = \mathbf{QR} \tag{36}$$

where $\mathbf{Q}$ is an $N \times \bar{N}$ unitary matrix and $\mathbf{R}$ is an $\bar{N} \times \bar{N}$ upper triangular matrix with positive diagonal elements $r_{11}, r_{22}, \cdots, r_{\bar{N}\bar{N}}$. Denote $\mathbf{D}=\text{diag}[\, r_{11}, r_{22}, \cdots, r_{\bar{N}\bar{N}} \,]$ and then Eq. (36) can be rewritten as

$$\mathbf{\Phi} = \mathbf{WA} \tag{37}$$

where $\mathbf{A}=\mathbf{D}^{-1}\mathbf{R}$ is an $\bar{N} \times \bar{N}$ upper triangular matrix with unit diagonal elements, that is,

$$\mathbf{A} = \begin{bmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1\bar{N}} \\ 0 & 1 & a_{23} & \cdots & a_{2\bar{N}} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & a_{\bar{N}-1\bar{N}} \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{38}$$

and $\mathbf{W} = \mathbf{QD}$ is an $N \times \bar{N}$ matrix with orthogonal columns $\mathbf{w_j}$, $j=1, 2,\cdots, \bar{N}$ such that

$$\mathbf{W}^{\mathrm{H}}\mathbf{W} = \mathbf{D}^2 = \mathbf{\Lambda} = \text{diag}[\, \lambda_1, \quad \lambda_2, \quad \cdots, \quad \lambda_{\bar{N}} \,] \tag{39}$$

where

$$\lambda_j = \langle \mathbf{w_j}, \ \mathbf{w_j} \rangle, \qquad j = 1, 2, \cdots, \bar{N} \tag{40}$$

and the symbol $\langle \bullet, \bullet \rangle$ denotes the inner product of two vectors.

Note that for two complex vectors $\boldsymbol{\alpha}=[\alpha_1,\cdots,\alpha_K]^{\mathrm{T}}$, $\boldsymbol{\beta}=[\beta_1,\cdots,\beta_K]^{\mathrm{T}}$,

$$\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = \boldsymbol{\beta}^{\mathrm{H}}\boldsymbol{\alpha} = (\bar{\boldsymbol{\beta}})^{\mathrm{T}}\boldsymbol{\alpha} \tag{41}$$

$$\langle \boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle = \boldsymbol{\alpha}^{\mathrm{H}}\boldsymbol{\alpha} = \|\boldsymbol{\alpha}\|_2^2 = \sum_{k=1}^{K} |\alpha_i|^2 \tag{42}$$

Substituting Eq. (37) into Eq. (29) gives

$$\mathbf{Y} = \mathbf{WA\Theta} + \mathbf{\Xi} \tag{43}$$

Denote

$$\mathbf{A\Theta} = \mathbf{g} \tag{44}$$

and then Eq. (43) can be expressed as

$$\mathbf{Y} = \mathbf{Wg} + \mathbf{\Xi} \tag{45}$$

or

$$\mathbf{Y} = \sum_{j=1}^{\bar{N}} g_j \mathbf{w_j} + \mathbf{\Xi} \tag{46}$$

which is an auxiliary model equivalent to Eq. (29) and the space spanned by the orthogonal basis vectors $\mathbf{w}_1, \mathbf{w}_2,\cdots,\mathbf{w}_{\bar{N}}$ is the same as that spanned by the original model basis $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2,\cdots,\boldsymbol{\varphi}_{\bar{N}}$.

By using the LS algorithm, the auxiliary parameter vector $\mathbf{g}$ can be solved from Eq. (45),

$$\mathbf{g} = \left(\mathbf{W}^{\mathrm{H}}\mathbf{W} \right)^{-1}\mathbf{W}^{\mathrm{H}}\mathbf{Y} \tag{47}$$

Substituting Eq. (39) into Eq. (47) gives

$$\mathbf{g} = \Lambda^{-1}\mathbf{W}^{H}\mathbf{Y}$$

(48)

or

$$g_j = \frac{\langle \mathbf{Y}, \mathbf{w_j} \rangle}{\langle \mathbf{w_j}, \mathbf{w_j} \rangle}, \quad j = 1, 2, \cdots, \bar{N}$$

(49)

Several orthogonalization procedures including classical Gram-Schmidt, modified Gram-Schmidt and Householder transformation[22] can be used to implement the orthogonal decomposition of the measured matrix $\boldsymbol{\Phi}$. Then after obtaining the auxiliary parameter vector $\mathbf{g}$ by Eq. (48), the parameter vector $\boldsymbol{\Theta}$ can be easily solved from Eq. (44) by using backward substitutions. However, our objective is not just to estimate the parameters, but also to detect which terms are significant and should be included within the model. This can be achieved by computing the error reduction ratio(ERR) described below.

Suppose that $Y_i(j\Omega)$, , $i=1, 2, \cdots, N$ is the output after its mean has been removed. Since $\boldsymbol{\Xi}$ is uncorrelated with $\boldsymbol{\varphi_i}$, $i=1, 2, \cdots, \bar{N}$, the variance of $Y_i(j\Omega)$ can be expressed as

$$\begin{aligned}
\sigma_Y^2 &= \frac{1}{N}\mathbf{Y}^{H}\mathbf{Y} \\
&= \frac{1}{N}\left[\mathbf{g}^{H}\mathbf{W}^{H}\mathbf{W}\mathbf{g} + 2(\mathbf{W}\mathbf{g})^{H}\boldsymbol{\Xi} + \boldsymbol{\Xi}^{H}\boldsymbol{\Xi} \right] \\
&= \frac{1}{N}\left[\mathbf{g}^{H}\Lambda\mathbf{g} + 2(\boldsymbol{\Phi}\boldsymbol{\Theta})^{H}\boldsymbol{\Xi} + \boldsymbol{\Xi}^{H}\boldsymbol{\Xi} \right] \\
&= \frac{1}{N}\left[\sum_{j=1}^{\bar{N}} |g_j|^2 \lambda_j + 2\boldsymbol{\Theta}^{H}\boldsymbol{\Phi}^{H}\boldsymbol{\Xi} + \boldsymbol{\Xi}^{H}\boldsymbol{\Xi} \right] \\
&= \frac{1}{N}\sum_{j=1}^{\bar{N}} |g_j|^2 \mathbf{w_j}^{H}\mathbf{w_j} + \frac{1}{N}\boldsymbol{\Xi}^{H}\boldsymbol{\Xi}
\end{aligned}$$

(50)

where the first part $\left(\sum_{j=1}^{\bar{N}} |g_j|^2 \mathbf{w_j}^{H}\mathbf{w_j} \right)/N$, which can be explained by the involved terms, is the desired output variance while the second part represents the unexplained variance. Thus $|g_j|^2 \mathbf{w_j}^{H}\mathbf{w_j}/N$ is the increment to the explained desired output variance brought by the jth term $\mathbf{w_j}$ and the jth error reduction ratio introduced by $\mathbf{w_j}(\boldsymbol{\Xi}^{H}\boldsymbol{\Xi})/N$ can be defined as

$$ERR_j = \frac{|g_j|^2 \mathbf{w_j}^{H}\mathbf{w_j}}{\mathbf{Y}^{H}\mathbf{Y}} \times 100\%, \quad j = 1, 2, \cdots, \bar{N}$$

(51)

Substituting Eq. (49) into Eq. (51) yields

$$ERR_j = \frac{|\langle \mathbf{Y}, \mathbf{w_j} \rangle|^2}{\langle \mathbf{Y}, \mathbf{Y} \rangle \langle \mathbf{w_j}, \mathbf{w_j} \rangle} \times 100\%, \quad j = 1, 2, \cdots, \bar{N}$$

(52)

which is also called the squared correlation coefficient between $\mathbf{Y}$ and $\mathbf{w_j}$.

From Eq. (50), the residual sum of squares $\|\boldsymbol{\Xi_{\bar{N}}}\|^2 = \left\| \mathbf{Y} - \hat{\mathbf{Y}} \right\|^2$, where $\hat{\mathbf{Y}}$ is the model prediction produced by the associated $\bar{N}$ terms model, can also be obtained,

$$\|\boldsymbol{\Xi_{\bar{N}}}\|^2 = \langle \mathbf{Y}, \mathbf{Y} \rangle - \sum_{j=1}^{\bar{N}} \frac{|\langle \mathbf{Y}, \mathbf{w_j} \rangle|^2}{\langle \mathbf{w_j}, \mathbf{w_j} \rangle}$$

(53)

while the residual vector $\boldsymbol{\Xi_{\bar{N}}}$ can be expressed from Eq. (46) as

$$\boldsymbol{\Xi_{\bar{N}}} = \mathbf{Y} - \sum_{j=1}^{\bar{N}} \frac{\langle \mathbf{Y}, \mathbf{w_j} \rangle}{\langle \mathbf{w_j}, \mathbf{w_j} \rangle}\mathbf{w_j}$$

(54)

Note that Eqs. (50)–(53) have been extended to the complex-valued case.
Dividing both sides of Eq. (50) by $\mathbf{Y}^{T}\mathbf{Y}/N$ gives

$$1 - \sum_{j=1}^{\bar{N}} ERR_j = \frac{\left(\mathbf{\Xi}^{\mathrm{H}}\mathbf{\Xi}\right)/N}{\left(\mathbf{Y}^{\mathrm{H}}\mathbf{Y}\right)/N} = \frac{\sigma_\varepsilon^2}{\sigma_Y^2} \tag{55}$$

which clearly indicates that the larger the ERR value associated with a particular term is, the more reduction in the residual variance will be produced if this term is included in the model. Thus the ERR provides a simple but effective means to detect which term is significant and should be selected. Notice that a term which is introduced at an early stage will have a larger ERR than that would be obtained if it were reordered to enter as a candidate term at a later stage. To overcome the order dependency of ERR, the terms can be selected in a forward stepwise manner. The detailed orthogonalization, for example, using the classical Gram-Schmidt algorithm, and terms selection procedure is described as follows.

At the first step, consider all the possible $\boldsymbol{\varphi_j}$, $j=1, 2, \cdots, \bar{N}$ as candidates for $\mathbf{w_1}$, and for $j=1, 2, \cdots, \bar{N}$, compute

$$\left\| \begin{aligned} &\mathbf{w_1^{(j)}} = \boldsymbol{\varphi_j}; \\ &ERR_1^{(j)} = \frac{\left|\left\langle \mathbf{Y}, \ \mathbf{w_1^{(j)}} \right\rangle\right|^2}{\left\langle \mathbf{Y}, \ \mathbf{Y} \right\rangle \left\langle \mathbf{w_1^{(j)}}, \ \mathbf{w_1^{(j)}} \right\rangle} \times 100\%. \end{aligned} \right. \tag{56}$$

Find the maximum of $ERR_1^{(j)}$, say $ERR_1^{(j_1)} = \max\left\{ ERR_1^{(j)}, 1 \le j \le \bar{N} \right\}$. Then the first term to be included in the model is $\boldsymbol{\varphi_{j_1}}$. $\mathbf{w_1} = \mathbf{w_1^{(j_1)}} = \boldsymbol{\varphi_{j_1}}$ is then selected as the first column of $\mathbf{W}$ together with the first element of the auxiliary parameter vector $\mathbf{g}$, $g_1 = \langle \mathbf{Y}, \mathbf{w_1} \rangle / \langle \mathbf{w_1}, \mathbf{w_1} \rangle$, the error reduction ratio produced by the first term, $ERR_1 = ERR_1^{(j_1)}$, and the associated sum-squared-error $\|\mathbf{\Xi_1}\|^2 = \langle \mathbf{Y}, \mathbf{Y} \rangle - |\langle \mathbf{Y}, \mathbf{w_1} \rangle|^2 / \langle \mathbf{w_1}, \mathbf{w_1} \rangle$. As defined in Eq. (38), the first column of $\mathbf{A}$, $a_{11}=1$.

At the kth step where $k \ge 2$, all the $\boldsymbol{\varphi_j}$, $j=1, 2, \cdots, \bar{N}, j \notin \left\{ j_1, \cdots, j_{k-1} \right\}$ are considered as possible candidates for $\mathbf{w_k}$, and for $j=1, 2, \cdots, \bar{N}, j \notin \left\{ j_1, \cdots, j_{k-1} \right\}$, calculate

$$\left\| \begin{aligned} &\mathbf{w_k^{(j)}} = \boldsymbol{\varphi_j} - \sum_{p=1}^{k-1} \frac{\left\langle \boldsymbol{\varphi_j}, \ \mathbf{w_p} \right\rangle}{\left\langle \mathbf{w_p}, \ \mathbf{w_p} \right\rangle} \mathbf{w_p}; \\ &ERR_k^{(j)} = \frac{\left|\left\langle \mathbf{Y}, \ \mathbf{w_k^{(j)}} \right\rangle\right|^2}{\left\langle \mathbf{Y}, \ \mathbf{Y} \right\rangle \left\langle \mathbf{w_k^{(j)}}, \ \mathbf{w_k^{(j)}} \right\rangle} \times 100\%. \end{aligned} \right. \tag{57}$$

Find the maximum of $ERR_k^{(j)}$, say $ERR_k^{(j_k)} = \max\left\{ ERR_k^{(j)}, 1 \le j \le \bar{N}, j \ne j_1, \cdots, j \ne j_{k-1} \right\}$. Then the kth term to be included in the model is $\boldsymbol{\varphi_{j_k}}$ while the kth column of $\mathbf{W}$, $\mathbf{w_k} = \mathbf{w_k^{(j_k)}}$, the kth element of the auxiliary parameter vector $\mathbf{g}$, $g_k = \langle \mathbf{Y}, \mathbf{w_k} \rangle / \langle \mathbf{w_k}, \mathbf{w_k} \rangle$, the kth error reduction ratio $ERR_k = ERR_k^{(j_k)}$, and the kth sum-squared-error $\|\mathbf{\Xi_k}\|^2 = \langle \mathbf{Y}, \mathbf{Y} \rangle - \sum_{j=1}^{k} |\langle \mathbf{Y}, \mathbf{w_j} \rangle|^2 / \langle \mathbf{w_j}, \mathbf{w_j} \rangle$. The elements of the kth column of $\mathbf{A}$ are computed by

$$a_{pk} = \begin{cases} \dfrac{\left\langle \boldsymbol{\varphi_{j_k}}, \ \mathbf{w_p} \right\rangle}{\left\langle \mathbf{w_p}, \ \mathbf{w_p} \right\rangle}, & p = 1, \cdots, k - 1 \\ 1, & p = k \end{cases} \tag{58}$$

According to Eq. (55), the procedure can be terminated at the $\bar{\bar{N}}$th step ($\bar{\bar{N}} \le \bar{N}$) when

$$1 - \sum_{j=1}^{\bar{\bar{N}}} ERR_j < \rho, \qquad 0 < \rho < 1 \tag{59}$$

where ρ is a chosen error tolerance and in practice, can actually be learnt during the selection procedure.

The criterion (59) concerns only the performance of the model (variance of residuals). Because a more accurate performance is often achieved at the expense of using a more complex model, a trade-off between the performance and complexity of the model is often desired. A number of model selection criteria that provide a compromise between the performance and the number of parameters have been introduced and incorporated into the OLS algorithm over the past few decades. Despite the differences amongst these model selection criteria, they are asymptotically equivalent under general conditions [23]. In this paper, the adjustable prediction error sum of squares (APRESS) [24] is employed to solve the model length determination problem,

$$APRESS(n) = c(n)MSE(n) \tag{60}$$

where

$$c(n) = \left(\frac{1}{1 - \alpha n/N} \right)^2 \tag{61}$$

with $\alpha \geq 1$, is the complexity cost function and

$$MSE(n) = \frac{\|\Xi_n\|^2}{N} \tag{62}$$

is the mean squared error corresponding to the model performance.

The model selection procedure is terminated at the $\hat{N}$th step when

$$APRESS\left(\hat{N} \right) = \min_{1 \leq n \leq \bar{N}} \left[APRESS(n) \right] \tag{63}$$

Practically a distinct turning point of the APRESS statistic versus the model length can be easily found, especially when computed by using several adjustable parameters α, and this can then be used to determine the model length.

The final model is thus the linear combination of the $\hat{N}$ significant terms $\varphi_{j_1}, \cdots, \varphi_{j_{\hat{N}}}$ selected from the $\bar{N}$ candidate terms $\varphi_1, \cdots, \varphi_{\bar{N}}$,

$$Y(t) = \sum_{k=1}^{\hat{N}} \hat{\theta}_k \phi_{j_k}(t) + \varepsilon(t) \tag{64}$$

where the parameter $\hat{\Theta} = \left[\hat{\theta}_1, \cdots, \hat{\theta}_{\hat{N}} \right]^T$ can easily be computed from Eq. (44) by using backward substitutions,

$$\left\| \begin{array}{l} \hat{\theta}_{\hat{N}} = g_{j_{\hat{N}}}; \\ \hat{\theta}_k = g_{j_k} - \sum_{p=k+1}^{\hat{N}} a_{kp} \hat{\theta}_p \quad for \ k = \hat{N} - 1, \hat{N} - 2, \cdots, 1. \end{array} \right. \tag{65}$$

It should be pointed out that in practice the mean of the output does not need to be removed because adding a constant to the denominator of the ERR will not affect the result of the maximization in this selection procedure. Because of the orthogonal property, this procedure is very efficient and leads to a parsimonious model. Moreover, any numerical ill-conditioning can be avoided by eliminating $\mathbf{w_k}$ if $\mathbf{w_k^H w_k}$ is less than a predetermined threshold. Similar selection procedures can also be derived using the modified Gram-Schmidt algorithm and Householder transformation algorithm. Simulation studies will be conducted in the next section to apply the developed OLS algorithm to the truncation of the Volterra series expansion and the estimation of the GFRFs of a typical nonlinear system.

5. Simulation study

Consider a single degree of freedom (SDOF) system shown in Fig. 1. This represents a mass supported on a linear spring $k_1(\bullet)$ in parallel with a nonlinear damper with a cubic polynomial characteristic $a_1(\bullet) + a_3(\bullet)^3$ where a_3 represents the system nonlinearity. The mass is subjected to a harmonic excitation force of amplitude F_m and frequency Ω, and the output is the force $F_T(t)$ transmitted to the system of interest via the mass–spring–damper element. This is a quite simple model but can represent a wide range of engineering systems such as automotive suspensions, machinery mounts, and base isolators of buildings.

Denote

$$y_F(t) = F_T(t) \tag{66}$$

$$y_d(t) = x(t) \tag{67}$$

and

$$u(t) = F_m \cos(\Omega t) \tag{68}$$

The equilibrium equation for the system in Fig. 1 and the transmitted force at the support can be expressed as

$$m\ddot{y}_d(t) + k_1 y_d(t) + a_1 \dot{y}_d(t) + a_3 \dot{y}_d^3(t) = u(t) \tag{69}$$
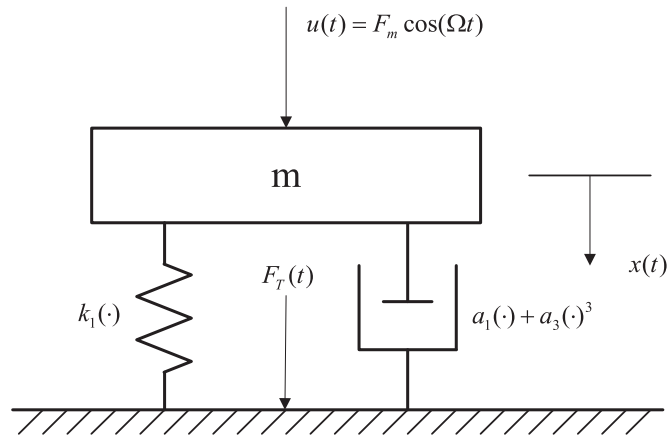
Fig. 1. A SDOF mass–spring–damper system.

$$y_F(t) = k_1 y_d(t) + a_1 \dot{y}_d(t) + a_3 \dot{y}_d^3(t) \tag{70}$$

This is a single input two output model but the transmitted force is the primary concern. Take the system linear characteristic and input parameters as follows:

$m = 240$kg, $k_1 = 16000$ N m^{-1}, $a_1 = 29.6$ s N m^{-1}, $F_m = 10$N, $\Omega = 8.1$ rad^{-1}.

By using the probing method described in Section 3.1, the analytical expressions of the GFRFs for the transmitted force have already been derived by Zhang et al. [4] and were given in Appendix A. Simulation studies will be conducted in this section for systems (69) and (70) subject to the harmonic input to evaluate the nth order output spectrum component Y_{2j+1} for the following three nonlinear damping characteristics using the new complex-valued OLS algorithm,

(i) $a_3 = 100$ s^3 N m^{-3}
(ii) $a_3 = 200$ s^3 N m^{-3}
(iii) $a_3 = 500$ s^3 N m^{-3}

The estimated results will then be compared with those determined using the analytical expressions of the GFRFs to demonstrate the effectiveness of the new method. Alternatively, by using the approach in Appendix B, the physical parameters of the system can be extracted from the measured GFRFs and then be compared with their true values to validate the method.

5.1. For the condition when $a_3 = 100$ s^3 N m^{-3}

In this case, the system was excited using the excitation amplitudes from 1 N to 10 N with an increment of 0.3 N. Solving Eqs. (69) and (70) yielded the system time-domain output $y_F(t)$. The output spectra $Y_F(\bar{j}\Omega)$ was then obtained by performing an FFT operation on the steady state output $y_F(t)$. Note that the FFT calculation of a harmonic signal should be conducted using a sampled sequence over a time interval which is a multiple of the signal period [25]. The output spectra together with the corresponding excitation amplitudes can then form Eq. (29). And the complex-valued OLS algorithm developed in Section 4 was then used to truncate the Volterra series expansion and to estimate the GFRFs.

The model length $\bar{N}$ in Eq. (29) was assumed to be the same as the number of measurements N. The initial model thus involved a total of 31 candidate model terms. The complex-valued OLS algorithm was then employed to select and rank the significant model terms. By setting the adjustable parameter $\alpha = 0, 0.2, \cdots, 0.8$, the APRESS statistic versus the model length over the measured data, were calculated and shown in Fig. 2, where the bottom line with circles, corresponding to $\alpha = 0$, indicates the mean-squared-errors. It can be seen from Fig. 2 that there is a turning point at the abscissa 6 for various values of the adjustable parameter α. Model structure detection is quite critical in system identification. To validate the effectiveness of the APRESS statistic, the Bayesian information criterion (BIC) statistic versus the model length was computed,

$$BIC(n) = \frac{\bar{N} + n[\ln(\bar{N}) - 1]}{\bar{N} - n} MSE(n) \tag{71}$$

The results are shown in Fig. 3 indicating the model length is also 6 where BIC arrives at the minimum value. The indexes of the first six model terms selected and ranked in order of the significance by the complex-valued OLS algorithm, together with the coefficient of each term and its corresponding ERR, are shown in Table 1. Table 1 indicates that the system can be represented by a Volterra series with maximum order 11. Reapplying the complex-valued OLS algorithm to the model represented by these terms over the measured data gave the estimation of the GFRFs which were substituted into Eq. (11) to
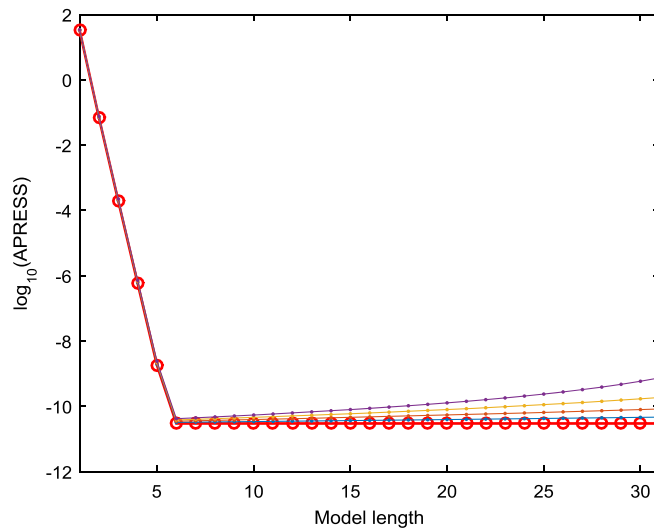
Fig. 2. The APRESS statistic versus the model length when $a_3 = 100 \text{ s}^3 \text{ N m}^{-3}$: the lines from bottom to the top correspond to $\alpha = 0, 0.2, \cdots, 0.8$. The bottom line with circles, corresponding to $\alpha = 0$, indicates the mean-squared-errors (MSE).
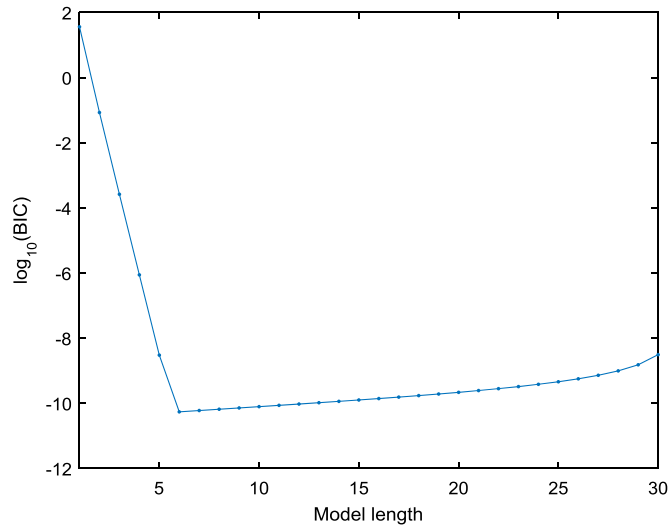

Fig. 3. The BIC statistic versus the model length when $a_3 = 100 \text{ s}^3 \text{ N m}^{-3}$.

Table 1
The model terms selected and ranked in order of the significance by the complex-valued OLS algorithm together with the coefficient of each term and its corresponding ERR.

Index	1	2	3	4	5	6
Terms	$H_{1,0}$	$H_{3,1}$	$H_{5,2}$	$H_{7,3}$	$H_{9,4}$	$H_{11,5}$
ERR (%)	99.95	0.0444	9.34e−5	2.61e−7	7.95e−10	2.45e−12

obtain the output spectrum components Y_{2j+1}. These results were then compared with the analytical results determined using the GFRFs expressions in Appendix A and were shown in Fig. 4. Clearly, the complex-valued OLS algorithm correctly truncated the Volterra series expansion of the system and correctly estimated the GFRFs.

The aforementioned procedure was repeated for the output $y_d(t)$. The only difference is that the GFRFs for the excitation frequency 10 rad s^{-1}, in addition to 8.1 rad s^{-1} in the transmitted force case, was also identified. It should be pointed out that as indicated by the OLS algorithm, the system excited at the non-resonant frequency 10 rad s^{-1} should be represented by a Volterra series with maximum order 3 which is significantly smaller than the maximum order 11 for the resonant frequency 8.1 rad s^{-1}. This is actually a fundamental behaviour of nonlinear systems. It is well known that the dynamic characteristics of a linear system are independent of the input and are described by only the frequency response function. But for a nonlinear system, despite that the first, second, and higher order GFRFs are still independent of the input, the
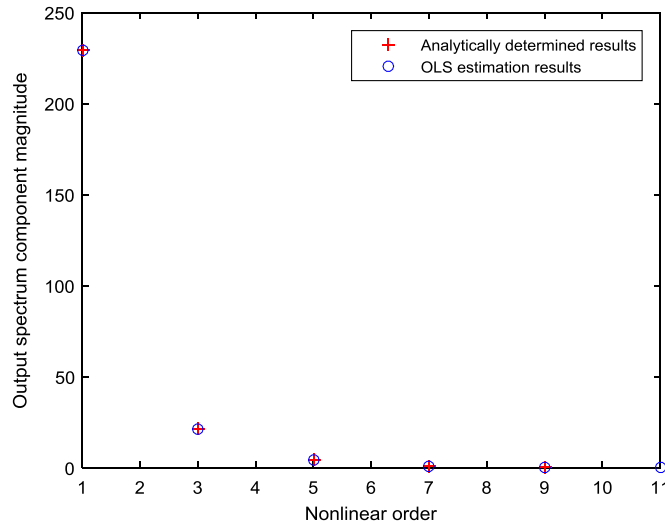
Fig. 4. The magnitudes of output spectrum components Y_{2j+1} at the driving frequency when a_3=100 s^3 N m^{-3}.

dynamic characteristics depend on not only the GFRFs but also the input applied to it. Although a nonlinear system subject to similar inputs such as harmonic excitations with similar frequencies may be represented by a Volterra series of the same maximum order, a system subject to different inputs are normally represented by Volterra series expansions of different orders. And a good practice is to apply the complex-valued OLS algorithm to them separately to determine the appropriate order. After obtaining the GFRFs for the excitation frequencies 8.1 rad s^{-1} and 10 rad s^{-1}, the method in Appendix B was then used to estimate the system parameters to further verify the method. The results are listed in Table 2 and agree with the true values very well, which confirms the effectiveness of the proposed method. Notice that in Table 2, the relative error is defined by $\left| x_{true} - \hat{x} \right| / x_{true} \times 100\%$, where x_{true} and $\hat{x}$ are the true value and estimated value of the parameter x, respectively.

5.2. For the condition when a_3=200 s^3 N m^{-3}

In this case, the system was also excited using the excitation amplitudes from 1 N to 10 N with an increment of 0.3 N. The excitation together with the FFT transformation of the solution of Eqs. (69) and (70) formed Eq. (29). As a starting point, the model length was assumed to be the same as the amount of measurements and the initial model involved 31 candidate model terms. The significant terms were then selected and ranked by the complex-valued OLS algorithm. After setting the adjustable parameter α=0, 0. 2, ⋯,0.8, the APRESS statistic versus the model length over the estimation data, were calculated and shown in Fig. 5, where the bottom line with circles, corresponding to α=0, indicates the mean-squared-errors. It can be seen from Fig. 5 that there is a turning point at the abscissa 8 for various values of the adjustable parameter α. The indexes of the first eight significant model terms, together with the coefficient of each term and its corresponding ERR, are shown in Table 3. Reapplying the complex-valued OLS algorithm to the model represented by these terms produced the estimation of the GFRFs. The output spectrum components Y_{2j+1} were then computed and compared with the corresponding analytical results determined using the GFRFs expressions in Appendix A, which is shown in Fig. 6. And again, the complex-valued OLS algorithm correctly truncated the Volterra series expansion of the system and correctly estimated the GFRFs. The procedure was repeated for the output $y_d(t)$ with the excitation frequencies 8.1 rad s^{-1}8.1rads^{-1} and 10 rad s^{-1}. The system parameters were then estimated using the method in Appendix B and are provided in Table 4. And once again, the system parameters were accurately estimated, which further validate the effectiveness of the method.

Comparing the results of Sections 5.1 and 5.2 indicates that the stronger the nonlinearity is, the more terms in the Volterra series expansion are necessary to represent the system. A natural question is what happens if the nonlinearity is

Table 2
The parameters of the cubic damper with a_3 = 100 s^3 N m^{-3} estimated by using the measured GFRFs of the displacement output.

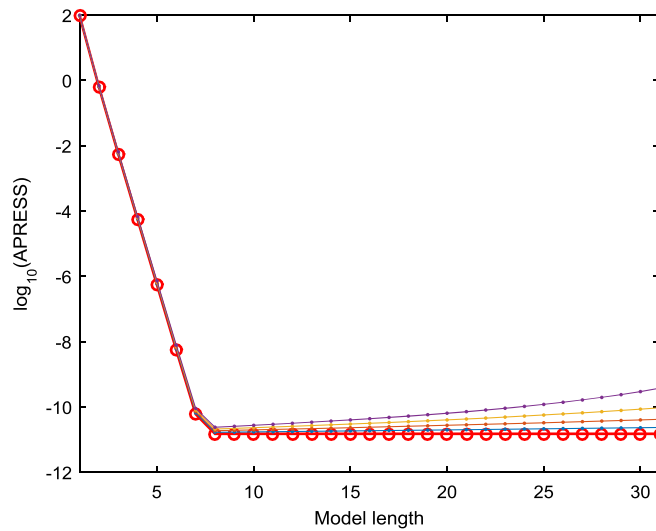| Parameter	True value	Estimated value	Relative error (%)
m (kg)	240	242.19	0.91
$a_1 \left(s\, N\, m^{-1} \right)$	29.6	29.82	0.75
$k_1 \left(N\, m^{-1} \right)$	16,000	16,142.2	0.89
$a_3 \left(s^3\, N\, m^{-3} \right)$	100	101.05	1.05

Fig. 5. The APRESS statistic versus the model length when $a_3 = 200 \, s^3 \, N \, m^{-3}$: the lines from bottom to the top correspond to $\alpha = 0, 0. \, 2, \cdots, 0. \, 8$. The bottom line with circles, corresponding to $\alpha = 0$, indicates the mean-squared-errors (MSE).

Table 3
The model terms selected and ranked in order of the significance by the complex-valued OLS algorithm together with the coefficient of each term and its corresponding ERR.

Index	1	2	3	4	5	6	7	8
Terms	$H_{1,0}$	$H_{3,1}$	$H_{5,2}$	$H_{7,3}$	$H_{9,4}$	$H_{11,5}$	$H_{13,6}$	$H_{15,7}$
ERR (%)	99.86	0.135	$8.88e-4$	$7.87e-6$	$7.75e-8$	$7.85e-10$	$7.88e-12$	$7.59e-14$

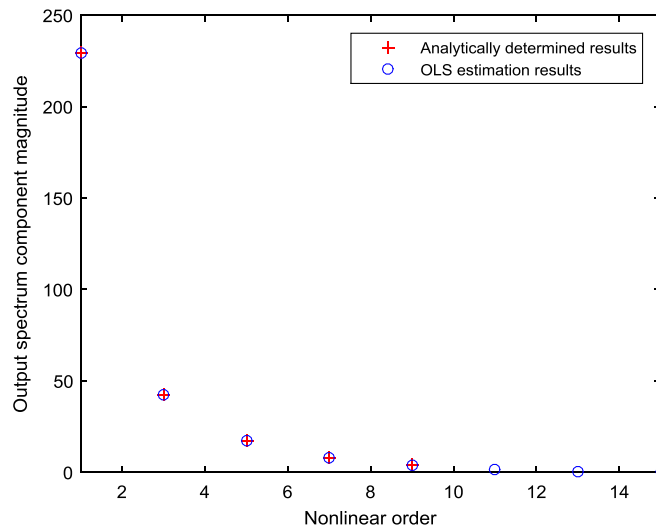**Fig. 6.** The magnitudes of output spectrum components Y_{2j+1} at the driving frequency when $a_3 = 200 s^3 \, N \, m^{-3}$.

very strong and the Volterra series diverges. Is the complex-valued OLS algorithm able to predict this non-convergent behaviour? This will be investigated in the next section.

5.3. For the condition when $a_3 = 500 \, s^3 \, N \, m^{-3}$

The excitations were exactly the same as those in Sections 5.1 and 5.2. And the complex-valued OLS algorithm was employed to select and rank the 31 candidate model terms. The APRESS statistic versus the model length with the adjustable parameter $\alpha = 0, 0. \, 2, \cdots, 0.8$ were then calculated and shown in Fig. 7. There is an obvious turning point at the abscissa 10 even with the bottom line with circles which corresponds to $\alpha = 0$ and represents the mean-squared-errors. The indexes of the first ten significant model terms, together with the coefficient of each term and its corresponding ERR, are

Table 4
The parameters of the cubic damper with $a_3 = 200 \, s^3 \, N \, m^{-3}$ estimated by using the measured GFRFs of the displacement output.

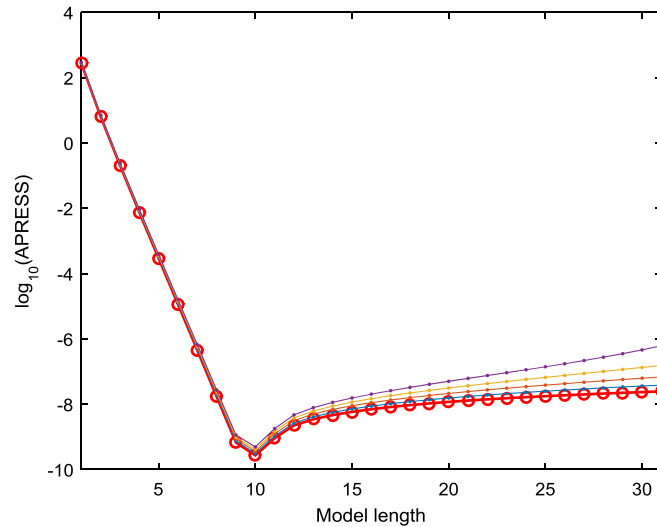| Parameter	True value	Estimated value	Relative error (%)
$m \, (\text{kg})$	240	243.33	1.39
$a_1 \, (s \, N \, m^{-1})$	29.6	29.87	0.91
$k_1 \, (N \, m^{-1})$	16,000	16,217.04	1.36
$a_3 \, (s^3 \, N \, m^{-3})$	200	193.94	3.03

Fig. 7. The APRESS statistic versus the model length when $\boldsymbol{a}_3 = 500 \, s^3 \, N \, m^{-3}$: the lines from bottom to the top correspond to $\alpha = 0, 0.2, \cdots, 0.8$. The bottom line with circles, corresponding to $\alpha = 0$, indicates the mean-squared-errors (MSE).

Table 5
The model terms selected and ranked in order of the significance by the complex-valued OLS algorithm together with the coefficient of each term and its corresponding ERR.

Index	1	2	3	4	5	6	7	8	9	10
Terms	$H_{1,0}$	$H_{3,1}$	$H_{5,2}$	$H_{7,3}$	$H_{9,4}$	$H_{11,5}$	$H_{13,6}$	$H_{15,7}$	$H_{17,8}$	$H_{19,9}$
ERR (%)	99.54	0.454	0.0105	3.35e−4	1.22e−05	4.71e−07	1.85e−08	7.22e−10	2.76e−11	1.36e−12

shown in Table 5. The GFRFs were estimated by reapplying the complex-valued OLS algorithm to the system model re-presented by these terms. The output spectrum components Y_{2j+1} were then computed and compared with the corresponding analytical results determined using the GFRFs expressions in Appendix A, which is shown in Fig. 8. It can be observed from the analytical results that the Volterra series expansion diverges as the amplitudes of the output spectrum components increase with the nonlinear order. But even in this situation, the complex-valued OLS algorithm correctly estimated the GFRFs up to the ninth order, which allows it to be used as an indicator of the convergence of the Volterra series expansion. The identification procedure was also repeated for the output $y_d(t)$ with the excitation frequencies 8.1 rad s^{-1} and 10 rad s^{-1}. The system parameters were again estimated using the method in Appendix B and are provided in Table 6 which indicates that they were accurately estimated. Lee [18] also reported that the GFRFs could be correctly estimated even when an excitation amplitude is outside of the convergence range. But note that only a convergent Volterra series is meaningful. If the numerically estimated results show the Volterra series expansion of a system diverges, it means the system nonlinearity is too strong and the Volterra series analysis shouldn't be pursued further.

6. Conclusions

A complex-valued OLS algorithm in the same form of the conventional OLS algorithm was developed and applied into the truncation of the Volterra series expansion and the estimation of the GFRFs of nonlinear systems. The estimated GFRFs
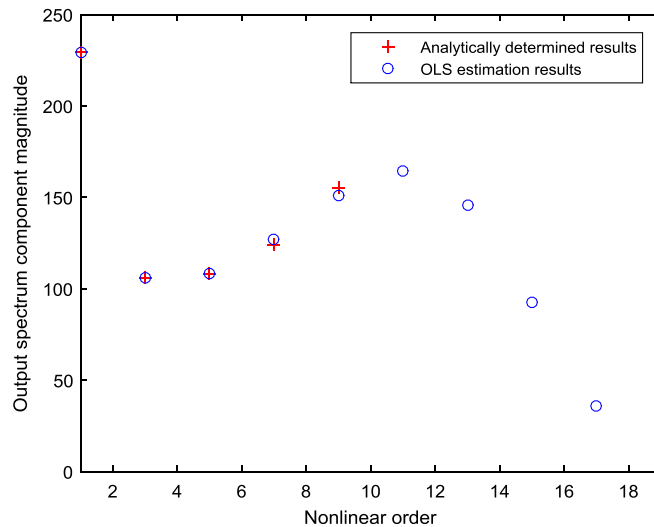
Fig. 8. The magnitudes of output spectrum components Y_{2j+1} at the driving frequency when $a_3 = 500$ s^3 N m^{-3}.

Table 6
The parameters of the cubic damper with $\boldsymbol{a}_3 = 500$ s^3 N m^{-3} estimated by using the measured GFRFs of the displacement output.

Parameter	True value	Estimated value	Relative error (%)
m (kg)	240	244.13	1.72
a_1 (s N m^{-1})	29.6	29.92	1.07
k_1 (N m^{-1})	16,000	16,269.29	1.68
a_3 (s^3 N m^{-3})	500	518.15	3.63

were then compared with the analytical results determined using the probing method to evaluate the effectiveness of the new method. The system parameters were also extracted from the estimated GFRFs and compared with their true values to further validate the method. Simulation studies demonstrate that the complex-valued OLS algorithm can correctly truncate the Volterra series expansion and correctly estimate the Volterra kernels of a weakly nonlinear system. And for a severely nonlinear system, the complex-valued OLS algorithm can correctly estimate the first few higher-order GFRFs and predict the non-applicability of the Volterra series analysis. In this paper, the approach is demonstrated by using a SDOF system subject to a sinusoidal excitation. For a harmonically excited system, only the diagonal values of the GFRFs are necessary for the analysis of the system and thus were estimated. But the main idea in this paper can easily be extended to more general cases such as systems under multi-tone excitations so as to obtain the complete multidimensional GFRFs' points and also multiple-degree-of-freedom (MDOF) nonlinear systems. These will be studied in later publications.

Appendix A. Analytical expressions of the GFRFs for the transmitted force

The expressions of the GFRFs for the transmitted force were derived by Zhang et al. [4] and will simply be quoted here.

$$H_{nF}(\bar{j}\,\omega_1, \cdots, \bar{j}\,\omega_n) = \begin{cases} \dfrac{\bar{j}\,\omega_1 a_1 + k_1}{\beta(\bar{j}\,\omega_1)}, & n = 1 \\[4mm] -m\left(\bar{j}\,\displaystyle\sum_{i=1}^{n}\omega_i\right)^2 H_{nd}(\bar{j}\,\omega_1, \cdots, \bar{j}\,\omega_n), & n = 2, 3, \cdots \end{cases} \tag{72}$$

and

$$H_{nd}(\bar{j}\,\omega_1, \cdots, \bar{j}\,\omega_n) = \begin{cases} \dfrac{1}{\beta(\bar{j}\,\omega_1)}, & n = 1 \\[4mm] -\dfrac{3!\,a_3 \sum_{p=1}^{|S_n^3|} S_n^3[p]}{n!\,\beta\!\left(\bar{j}\,\sum_{i=1}^{n}\omega_i\right)}, & n = 3, 5, \cdots \\[4mm] 0, & n = 2, 4, \cdots \end{cases} \tag{73}$$

where

$$\beta\!\left(\bar{j}\,\sum_{i=1}^{n}\omega_i\right) = m\left(\bar{j}\,\sum_{i=1}^{n}\omega_i\right)^2 + a_1\left(\bar{j}\,\sum_{i=1}^{n}\omega_i\right) + k_1 \tag{74}$$

and

$$\sum_{p=1}^{|S_n^3|} S_n^3[p] = M^{(3)} \triangleq \sum_{(r;3,n)} \sideset{}{'}\sum_{N} r_1!\left(\bar{j}\,\omega_1 + \cdots + \bar{j}\,\omega_{r_1}\right) H_{r_1 d}(\bar{j}\,\omega_1, \cdots, \bar{j}\,\omega_{r_1}) r_2!\left(\bar{j}\,\omega_{r_1+1} + \cdots + \bar{j}\,\omega_{r_1+r_2}\right)$$
$$H_{r_2 d}(\bar{j}\,\omega_{r_1+1}, \cdots, \bar{j}\,\omega_{r_1+r_2}) r_3!\left(\bar{j}\,\omega_{r_1+r_2+1} + \cdots + \bar{j}\,\omega_n\right) H_{r_3 d}(\bar{j}\,\omega_{r_1+r_2+1}, \cdots, \bar{j}\,\omega_n) \tag{75}$$

Specifically,

$$H_{3d}(\bar{j}\,\omega_1, \bar{j}\,\omega_2, \bar{j}\,\omega_3) \triangleq -\frac{3!\,a_3 M_{111}^{(3)}}{3!\,\beta\!\left(\bar{j}\,\sum_{i=1}^{3}\omega_i\right)} \triangleq -\frac{a_3(1)(2)(3)}{\beta\!\left(\bar{j}\,\sum_{i=1}^{3}\omega_i\right)} \tag{76}$$

where

$$(1)(2)(3) \triangleq \left(\bar{j}\,\omega_1\right) H_{1d}(\bar{j}\,\omega_1)\left(\bar{j}\,\omega_2\right) H_{1d}(\bar{j}\,\omega_2)\left(\bar{j}\,\omega_3\right) H_{1d}(\bar{j}\,\omega_3) \tag{77}$$

And in the case $n = 5$,

$$H_{5d}(\bar{j}\,\omega_1, \bar{j}\,\omega_2, \bar{j}\,\omega_3, \bar{j}\,\omega_4, \bar{j}\,\omega_5) \triangleq -\frac{a_3 M_{113}^{(3)}}{20\beta\!\left(\bar{j}\,\sum_{i=1}^{5}\omega_i\right)} \tag{78}$$

Here

$$M_{113}^{(3)} = (5)(4)(321) + (5)(3)(421) + (5)(2)(341) + (5)(1)(324) + (4)(3)(521) +$$
$$(4)(2)(351) + (4)(1)(325) + (3)(2)(541) + (3)(1)(524) + (2)(1)(354) \tag{79}$$

where

$$(5)(4)(321) \triangleq \left(\bar{j}\,\omega_5\right) H_{1d}(\bar{j}\,\omega_5)\left(\bar{j}\,\omega_4\right) H_{1d}(\bar{j}\,\omega_4) 3!\left(\bar{j}\,\omega_3 + \bar{j}\,\omega_2 + \bar{j}\,\omega_1\right) H_{3d}(\bar{j}\,\omega_3, \bar{j}\,\omega_2, \bar{j}\,\omega_1) \tag{80}$$

and so on.
In the case $n = 7$,

$$H_{7d}(\bar{j}\,\omega_1, \bar{j}\,\omega_2, \bar{j}\,\omega_3, \bar{j}\,\omega_4, \bar{j}\,\omega_5, \bar{j}\,\omega_6, \bar{j}\,\omega_7) \triangleq -\frac{a_3\left(M_{115}^{(3)} + M_{331}^{(3)}\right)}{840\beta\!\left(\bar{j}\,\sum_{i=1}^{7}\omega_i\right)} \tag{81}$$

Here

$$
\begin{aligned}
M_{115}^{(3)} =\ & (7)(6)(54321) + (7)(5)(64321) + (7)(4)(56321) + (7)(3)(54621) + (7)(2)(54361) \\
& + (7)(1)(54326) + (6)(5)(74321) + (6)(4)(57321) + (6)(3)(54721) + (6)(2)(54371) \\
& + (6)(1)(54327) + (5)(4)(76321) + (5)(3)(74621) + (5)(2)(74361) + (5)(1)(74326) \\
& + (4)(3)(57621) + (4)(2)(57361) + (4)(1)(57326) + (3)(2)(54761) + (3)(1)(54726) \\
& + (2)(1)(54376)
\end{aligned}
\tag{82}
$$

and

$$
\begin{aligned}
M_{331}^{(3)} =\ & (765)(432)(1) + (765)(431)(2) + (765)(421)(3) + (765)(321)(4) + (764)(532)(1) \\
& + (764)(531)(2) + (764)(521)(3) + (764)(321)(5) + (763)(452)(1) + (763)(451)(2) \\
& + (763)(421)(5) + (763)(521)(4) + (762)(435)(1) + (762)(431)(5) + (762)(451)(3) \\
& + (762)(351)(4) + (761)(432)(5) + (761)(435)(2) + (761)(425)(3) + (761)(325)(4) \\
& + (754)(632)(1) + (754)(631)(2) + (754)(621)(3) + (754)(321)(6) + (753)(462)(1) \\
& + (753)(461)(2) + (753)(421)(6) + (753)(621)(4) + (752)(436)(1) + (752)(431)(6) \\
& + (752)(461)(3) + (752)(361)(4) + (751)(432)(6) + (751)(436)(2) + (751)(426)(3) \\
& + \ \ (751)(326)(4) + (743)(652)(1) + (743)(651)(2) + (743)(621)(5) + (743)(521)(6) \\
& + (742)(635)(1) + (742)(631)(5) + (742)(651)(3) + (742)(351)(6) + (741)(632)(5) \\
& + (741)(635)(2) + (741)(625)(3) + (741)(325)(6) + (732)(465)(1) + (732)(461)(5) \\
& + (732)(451)(6) + (732)(651)(4) + (731)(462)(5) + (731)(465)(2) + (731)(425)(6) \\
& + (731)(625)(4) + (721)(436)(5) + (721)(435)(6) + (721)(465)(3) + (721)(365)(4) \\
& + (654)(321)(7) + (653)(421)(7) + (652)(431)(7) + (651)(432)(7) + (643)(521)(7) \\
& + (642)(351)(7) + (641)(325)(7) + (632)(451)(7) + (631)(425)(7) + (621)(435)(7)
\end{aligned}
\tag{83}
$$

where

$$
(7)(6)(54321) \triangleq (\bar{j}\omega_7)H_{1d}(\bar{j}\omega_7)(\bar{j}\omega_6)H_{1d}(\bar{j}\omega_6)5!(\bar{j}\omega_5 + \bar{j}\omega_4 + \bar{j}\omega_3 + \bar{j}\omega_2 + \bar{j}\omega_1)H_{5d}(\bar{j}\omega_1, \bar{j}\omega_2, \bar{j}\omega_3, \bar{j}\omega_4, \bar{j}\omega_5)
\tag{84}
$$

$$
(764)(531)(2) \triangleq 3!(\bar{j}\omega_7 + \bar{j}\omega_6 + \bar{j}\omega_4)H_{3d}(\bar{j}\omega_7, \bar{j}\omega_6, \bar{j}\omega_4)3!(\bar{j}\omega_5 + \bar{j}\omega_3 + \bar{j}\omega_1)H_{3d}(\bar{j}\omega_5, \bar{j}\omega_3, \bar{j}\omega_1)(\bar{j}\omega_2)H_{1d}(\bar{j}\omega_2)
\tag{85}
$$

and so on.

Note that the first five terms of $M_{331}^{(3)}$ were missing in [4] due to typographical errors but are added in here.

In the case $n = 9$,

$$
H_{9d}(\bar{j}\omega_1, \bar{j}\omega_2, \bar{j}\omega_3, \bar{j}\omega_4, \bar{j}\omega_5, \bar{j}\omega_6, \bar{j}\omega_7, \bar{j}\omega_8, \bar{j}\omega_9) \triangleq -\frac{3!a_3\left(M_{117}^{(3)} + M_{135}^{(3)} + M_{333}^{(3)}\right)}{9!\beta\left(\bar{j}\sum_{i=1}^{9}\omega_i\right)}
\tag{86}
$$

where $M_{117}^{(3)}$, $M_{135}^{(3)}$, and $M_{333}^{(3)}$ have 36, 504 and 280 terms respectively. These expressions are omitted here due to space limitations.

Appendix B. Estimation of system parameters using the GFRFs of the displacement

Once the GFRFs of the displacement are measured, the linear and nonlinear parameters of the system can then be estimated.

According to Eq. (73), the linear GFRF of the displacement can be expressed as,

$$
H_{1d}(j\omega) = \frac{1}{m(\bar{j}\omega)^2 + a_1(\bar{j}\omega) + k_1}
\tag{87}
$$

Eq. (87) can be rewritten by

$$
\begin{bmatrix}
\operatorname{Re}\left\{-\omega^2 H_{1d}(\bar{j}\omega)\right\} & \operatorname{Re}\left\{\bar{j}\omega H_{1d}(\bar{j}\omega)\right\} & \operatorname{Re}\left\{H_{1d}(\bar{j}\omega)\right\} \\
\operatorname{Im}\left\{-\omega^2 H_{1d}(\bar{j}\omega)\right\} & \operatorname{Im}\left\{\bar{j}\omega H_{1d}(\bar{j}\omega)\right\} & \operatorname{Im}\left\{H_{1d}(\bar{j}\omega)\right\}
\end{bmatrix}
\begin{bmatrix}
m \\
a_1 \\
k_1
\end{bmatrix}
=
\begin{bmatrix}
1 \\
0
\end{bmatrix}
\tag{88}
$$

where Re[•] and Im[•] denote the real and imaginary part of a complex number respectively. If the linear GFRFs $H_{1d}(\bar{j}\,\Omega_i)$, $i = 1,2, \cdots, N$ for N different excitation frequencies Ω_i are measured by the new method proposed in this paper, Eq. (88) can be expressed in the following matrix form,

$$\mathbf{H}\Theta = \mathbf{I} \tag{89}$$

where

$$\mathbf{H} = \begin{bmatrix} \mathrm{Re}\left\{-\Omega_1^2 H_{1d}(\bar{j}\,\Omega_1)\right\} & \mathrm{Re}\left\{\bar{j}\,\Omega_1 H_{1d}(\bar{j}\,\Omega_1)\right\} & \mathrm{Re}\left\{H_{1d}(\bar{j}\,\Omega_1)\right\} \\ \mathrm{Im}\left\{-\Omega_1^2 H_{1d}(\bar{j}\,\Omega_1)\right\} & \mathrm{Im}\left\{\bar{j}\,\Omega_1 H_{1d}(\bar{j}\,\Omega_1)\right\} & \mathrm{Im}\left\{H_{1d}(\bar{j}\,\Omega_1)\right\} \\ \vdots & \vdots & \vdots \\ \mathrm{Re}\left\{-\Omega_N^2 H_{1d}(\bar{j}\,\Omega_N)\right\} & \mathrm{Re}\left\{\bar{j}\,\Omega_N H_{1d}(\bar{j}\,\Omega_N)\right\} & \mathrm{Re}\left\{H_{1d}(\bar{j}\,\Omega_N)\right\} \\ \mathrm{Im}\left\{-\Omega_N^2 H_{1d}(\bar{j}\,\Omega_N)\right\} & \mathrm{Im}\left\{\bar{j}\,\Omega_N H_{1d}(\bar{j}\,\Omega_N)\right\} & \mathrm{Im}\left\{H_{1d}(\bar{j}\,\Omega_N)\right\} \end{bmatrix} \tag{90}$$

$$\Theta = \begin{bmatrix} m, & a_1, & k_1 \end{bmatrix}^T \tag{91}$$

$$\mathbf{I} = \begin{bmatrix} 1, & 0, & \cdots, & 1, & 0 \end{bmatrix}^T \tag{92}$$

The linear parameters of the system $\hat{m}, \hat{a}_1, \hat{k}_1$ can be obtained from Eq. (89) by the LS algorithm,

$$\hat{\Theta} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{I} \tag{93}$$

According to Eq. (76), the third-order GFRF of the displacement can be expressed as,

$$H_{3d}(\bar{j}\,\omega, \bar{j}\,\omega, -\bar{j}\,\omega) = -\bar{j}\,\omega^3 H_{1d}^3(\bar{j}\,\omega)\hat{H}_{1d}(-\bar{j}\,\omega)a_3 \tag{94}$$

where $\hat{H}_{1d}(-\bar{j}\,\omega)$ can be obtained by Eq. (87) and the linear parameters are given by Eq. (93).

As $H_{3d}(\bar{j}\,\Omega_i), i = 1,2, \cdots, N$, has already been measured by the new method, Eq. (94) can be written in the following matrix form,

$$\mathbf{Q}a_3 = \bar{\mathbf{H}} \tag{95}$$

where

$$\mathbf{Q} = \left[\mathrm{Re}\left\{-\bar{j}\,\Omega_1^3 H_{1d}^3(\bar{j}\,\Omega_1)\hat{H}_{1d}(-\bar{j}\,\Omega_1)\right\}, \quad \mathrm{Im}\left\{-\bar{j}\,\Omega_1^3 H_{1d}^3(\bar{j}\,\Omega_1)\hat{H}_{1d}(-\bar{j}\,\Omega_1)\right\}, \quad \cdots, \right.$$
$$\left. \mathrm{Re}\left\{-\bar{j}\,\Omega_N^3 H_{1d}^3(\bar{j}\,\Omega_N)\hat{H}_{1d}(-\bar{j}\,\Omega_N)\right\}, \quad \mathrm{Im}\left\{-\bar{j}\,\Omega_N^3 H_{1d}^3(\bar{j}\,\Omega_N)\hat{H}_{1d}(-\bar{j}\,\Omega_N)\right\} \right]^T \tag{96}$$

$$\bar{\mathbf{H}} = \left[\mathrm{Re}\left\{H_{3d}(\bar{j}\,\Omega_1, \bar{j}\,\Omega_1, -\bar{j}\,\Omega_1)\right\}, \quad \mathrm{Im}\left\{H_{3d}(\bar{j}\,\Omega_1, \bar{j}\,\Omega_1, -\bar{j}\,\Omega_1)\right\}, \quad \cdots, \right.$$
$$\left. \mathrm{Re}\left\{H_{3d}(\bar{j}\,\Omega_N, \bar{j}\,\Omega_N, -\bar{j}\,\Omega_N)\right\}, \quad \mathrm{Im}\left\{H_{3d}(\bar{j}\,\Omega_N, \bar{j}\,\Omega_N, -\bar{j}\,\Omega_N)\right\} \right]^T \tag{97}$$

The nonlinear parameter a_3 can then be obtained by,

$$\hat{a}_3 = \left(\mathbf{Q}^T\mathbf{Q}\right)^{-1}\mathbf{Q}^T\bar{\mathbf{H}} \tag{98}$$

References

[1] V. Volterra, Theory of Functionals and of Integral and Integro-Differential Equations, Dover Publications, New York, 1959 (Unabridged republication of the first English translation).
[2] L. Adegeest, Third-order Volterra modelling of ship responses based on regular wave results, Twenty-First Symposium on Naval Hydrodynamics, National Academy of Sciences, Washington D.C., (1997) 189-204.
[3] S. Cafferty, G.R. Tomlinson, Characterization of automotive dampers using higher order frequency response functions, Proc. Inst. Mech. Eng. D – J. Automob. 211 (1997) 181–203.
[4] B. Zhang, S.A. Billings, Z.Q. Lang, G.R. Tomlinson, A novel nonlinear approach to suppress resonant vibrations, J. Sound Vib. 317 (2008) 918–936.
[5] V.Z. Marmarelis, Nonlinear Dynamic Modelling of Physiological Systems, Wiley, Hoboken, New Jersey, 2004.
[6] E. Bedrosian, S.O. Rice, The output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and Gaussian inputs, Proc. IEEE 59 (1971) 1688–1707.
[7] S.A. Billings, K.M. Tsang, Spectral-analysis for non-linear systems. 2. Interpretation of non-linear frequency-response functions, Mech. Syst. Signal Process. 3 (1989) 341–359.

[8] S.A. Billings, Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains, Wiley, Chichester, West Sussex, United Kingdom, 2013.

[9] S.A. Billings, K.M. Tsang, Spectral-analysis for non-linear systems. 1. Parametric non-linear spectral-analysis, Mech. Syst. Signal Process. 3 (1989) 319–339.

[10] K. Kim, E.J. Powers, A digital method of modeling quadratically nonlinear-systems with a general random input, IEEE Trans. Acoust. Speech 36 (1988) 1758–1769.

[11] S.W. Nam, E.J. Powers, Application of higher-order spectral-analysis to cubically nonlinear-system identification, IEEE Trans. Signal Process. 42 (1994) 1746–1765.

[12] L.M. Li, S.A. Billings, Estimation of generalized frequency response functions for quadratically and cubically nonlinear systems, J. Sound Vib. 330 (2011) 461–470.

[13] S.A. Billings, Z.Q. Lang, Truncation of nonlinear system expansions in the frequency domain, Int. J. Control 68 (1997) 1019–1042.

[14] J.J. Bussgang, L. Ehrman, J.W. Graham, Analysis of nonlinear-systems with multiple inputs, Proc. IEEE 62 (1974) 1088–1119.

[15] W.J. Rugh, Nonlinear System Theory: the Volterra/Wiener Approach, Johns Hopkins University Press, Baltimore, 1981.

[16] B. Zhang, S.A. Billings, Z.Q. Lang, G.R. Tomlinson, Analytical description of the frequency response function of the generalized higher order duffing oscillator model, IEEE Trans. Circuits-I 56 (2009) 224–232.

[17] S. Boyd, Y.S. Tang, L.O. Chua, Measuring Volterra kernels, IEEE Trans. Circuits Syst. 30 (1983) 571–577.

[18] G.M. Lee, Estimation of non-linear system parameters using higher-order frequency response functions, Mech. Syst. Signal Process. 11 (1997) 219–228.

[19] D.M. Storer, G.R. Tomlinson, Recent developments in the measurement and interpretation of higher-order transfer-functions from nonlinear structures, Mech. Syst. Signal Process. 7 (1993) 173–189.

[20] K.M. Tsang, S.A. Billings, Orthogonal estimation algorithm for complex number-systems, Int. J. Syst. Sci. 23 (1992) 1011–1018.

[21] S. Chen, S. Mclaughlin, B. Mulgrew, Complex-valued radial basis function network. 1. Network architecture and learning algorithms, Signal Process 35 (1994) 19–31.

[22] G.A.F. Seber, Linear Regression Analysis, John Wiley & Sons, New York, 1977.

[23] X. Hong, R.J. Mitchell, S. Chen, C.J. Harris, K. Li, G.W. Irwin, Model selection approaches for non-linear system identification: a review, Int. J. Syst. Sci. 39 (2008) 925–946.

[24] S.A. Billings, H.L. Wei, An adaptive orthogonal search algorithm for model subset selection and non-linear system identification, Int. J. Control 81 (2008) 714–724.

[25] Z.Q. Lang, S.A. Billings, F. Xie, G.R. Tomlinson, Accurate computation of output frequency responses of nonlinear systems, Proc. SPIE 5503 (2004) 131–142.