Identification and frequency domain analysis of non-stationary and nonlinear systems using time-varying NARMAX models

He, F., Wei, H. L. & Billings, S. A.

Original citation & hyperlink:
He, F, Wei, HL & Billings, SA 2015, 'Identification and frequency domain analysis of non-stationary and nonlinear systems using time-varying NARMAX models' International Journal of Systems Science, vol. 46, no. 11, pp. 2087-2100.
https://dx.doi.org/10.1080/00207721.2013.860202

DOI 10.1080/00207721.2013.860202
ISSN 0020-7721
ESSN 1464-5319
Publisher: Taylor and Francis

Identification and Frequency Domain Analysis of Nonstationary and Nonlinear Systems Using Time-Varying NARMAX Models

Fei He, Hua-Liang Wei, and Stephen A. Billings

Department of Automatic Control and Systems Engineering, The University of Sheffield, Sheffield, S1 3JD, U.K.

Corresponding author: Stephen A. Billings. Email: s.billings@sheffield.ac.uk

Identification and Frequency Domain Analysis of Nonstationary and Nonlinear Systems Using Time-Varying NARMAX Models

This paper introduces a new approach for nonlinear and nonstationary (time-varying) system identification based on time-varying NARMAX (TV-NARMAX) models. The challenging model structure selection and parameter tracking problems are solved by combining a multi-wavelet basis function expansion of the time-varying parameters with an orthogonal least squares algorithm. Numerical examples demonstrate that the proposed approach can track rapid time-varying effects in nonlinear systems more accurately than the standard recursive algorithms. Based on the identified time domain model, a new frequency domain analysis approach is introduced based on a time-varying generalized frequency response function (TV-GFRF) concept, which enables the analysis of nonlinear nonstationary systems in the frequency domain. Features in the TV-GFRFs which depend on the TV-NARMAX model structure and time-varying parameters are investigated. It is shown that the high dimensional frequency features can be visualized in a low dimensional time-frequency space.

Keywords: Generalized frequency response functions; nonlinear and nonstationary systems; system identification; time varying systems; wavelet basis functions

1. Introduction

Many processes are inherently nonstationary, including a large number of physical, physiological, and biochemical systems (Fitzgerald, Smith, Walden, and Young 2000). Modeling and identification of nonstationary processes have been widely studied for linear systems. Traditional techniques for identifying linear time-varying (LTV) systems are primarily based on adaptive recursive methods, for example recursive least squares (RLS), least mean squares (LMS), and the Kalman filter (Ljung and Soderstrom 1983; Ljung and Gunnarsson 1990; Bermudez and Bershad 1996). More recently, approaches based on expanding the time-varying (TV) coefficients using a finite sequence of basis functions, such as Legendre, Fourier or wavelet bases, have been proposed (Zou, Wang, and Chon 2003; Niedzwiecki 1988; Tsatsanis and Giannakis 1993; Niedzwiecki and Klaput 2002; Zheng, Lin, and Tay 2001; Li, Wei, and Billings 2011) mainly for polynomial LTV models (e.g. TV-ARMA, and TV-ARX models). This approach reduces the initial TV problem to a time invariant model identification problem and provides significant improvements on the tracking of rapid changes in TV coefficients. For such LTV systems, providing the correct model structure can be identified and tracked in time, the models can be mapped into the frequency domain (Zou and Chon 2004; Ball, Gohberg, and Kaashoek 1995) which is important for the model analysis and especially for real applications, e.g. geophysical, electrophysiological signals (EEG, EMG, etc) (Billings 2013, Wei, Billings, and Liu 2010; Li, Wei, Billings, and Sarrigiannis 2011).

However, many real systems are nonlinear. If a process is nonlinear and the transition is relatively slow and smooth, LTV models can be used to approximate the 'true' nonlinear process and track the nonstationary effects. Such approaches are widely used and can often provide good approximations in real-time implementations, e.g. model predictive control, and online adaptive control (Peng, Nakano, and Shioya 2007; Chowdhury 2000). Nevertheless, when analyzing the characteristics of a system in the frequency domain, such a LTV approximation can never reproduce nonlinear effects such as harmonics, intermodulations and energy transfer which can only be produced from a nonlinear model. Such features are critically important when the frequency behavior of a system is crucial to the analysis, such as in nonlinear vibrations, nonlinear communications, and neurophysiological signals. A key aim therefore of this study is to be able to identify and track nonstationarity in nonlinear systems and directly map the time-varying nonlinear behavior into the frequency domain.

A wide class of nonlinear time-invariant (NTI) systems can be represented by Volterra series models (Schetzen 1980; Marmarelis 1981) or Nonlinear Auto-Regressive Moving Average with eXogenous variable (NARMAX) models (Billings, 2013, Leontaritis and Billings 1985). Polynomial NARMAX models have been demonstrated to be a computationally efficient approach that can well capture nonlinear effects for a diverse range of nonlinear continuous and discrete time systems (Zheng and Billings 1999; Billings and Li 2000). However, limited studies exist on nonlinear and nonstationary system identification. Until recently, approaches based on time-varying Volterra series combining artificial neural networks (Iatrou, Berger, and Marmarelis 1999) or principal dynamic modes (Zhong, Jan, Ju, and Chon 2007) as well as time variable parameters (TVP) estimation based dynamic harmonic regression (DHR) (Young 2011) have been proposed, however the model structure selection problem is still an unsolved issue and the frequency domain analysis based on Volterra models can be computationally costly, while the DHR method only provide pseudo-spectrum analysis for a small class of nonlinear systems. In this study, a new procedure is proposed for the identification of nonlinear and nonstationary dynamic systems based on the time-varying NARX (TV-NARX) model. The basis function expansion strategy proposed for LTV identification is extended to more general nonlinear time-varying cases. The TV parameters in TV-NARX models are expanded using multi-wavelet basis functions, as a result the TV-NARX model is transformed into an expanded time invariant model, and the challenging TV model selection and parameter estimation problem can then be solved by using the orthogonal least squares (OLS) algorithm (Chen, Billings, and Luo 1989; Billings, Chen, and Korenberg 1989). This identification approach is also extended to more general time-varying NARMAX models by introducing a modified extended least squares (ELS) algorithm.

After the TV-NARX model has been identified in the time domain, the model is mapped into the frequency domain so that the nonlinear and nonstationary effects can be characterized and analyzed. Conventionally, the frequency domain representation of a NTI system can be obtained by employing the multi-dimensional Fourier transformation of the Volterra series expansion of the system. This yields the so-called generalized frequency response functions (GFRFs), which can be estimated using non-parametric approaches (Boyd, Tang, and Chua 1983; Chua and Liao 1989) or parametric approaches based on nonlinear differential equations or the NARMAX models (Billings and Tsang 1989a; Peyton Jones and Billings 1989; Billings and Peyton Jones 1990). The latter approach is computationally more attractive and the analytical expressions of the GFRFs up to any order can be obtained. However, the frequency domain representation of a nonlinear time-varying system has not yet been discussed in the literature. In this paper, the concept of time-varying GFRFs (TV-GFRFs) is proposed, which allows the computation and analysis of time-dependent GFRFs of nonlinear and nonstationary systems. More importantly, 'features' in the TV-GFRFs which functionally depend on the TV-NARX model structure and TV parameters can be analyzed systematically. It is shown that the gain of the TV-GFRFs up to any order can be transformed and visualized in a two-dimensional time-frequency space. The results obtained in this work greatly facilitate the understanding and practical analysis of the frequency characteristics of nonlinear and nonstationary systems.

The paper is organized as follows. Section II presents the time domain identification procedure based on TV-NARX models. In section III, the frequency domain analysis of the identified TV-NARX models is proposed, including the novel characteristic analysis of GFRFs and the visualization of TV-GFRFs. Two simulation examples are provided in section IV, and conclusions are given in section V.

2. Identification of Time-Varying Nonlinear Systems

A wide range of NTI systems can be represented using NARMAX models (Leontaritis and Billings 1985):

$$y(t) = f\left(y(t-1),\ldots,y(t-n_y),u(t-1),\ldots u(t-n_u),e(t-1),\ldots e(t-n_e)\right)+e(t) \tag{1}$$

where f is a nonlinear function, $u(t)$ and $y(t)$ are the sampled input and output sequences, with maximum lags n_u and n_y, respectively; The stochastic variable $e(t)$ is assumed to be to be independent, bounded and uncorrelated with the input $u(t)$. For many practical problems, the nonlinear function $f(\cdot)$ is generally unknown. The most commonly used approach to approximate the unknown function $f(\cdot)$ is to employ a polynomial NARMAX expression (Chen and Billings 1989). The NARX model, a subset of NARMAX model without considering the moving average noise terms in (1), can be further expressed in a linear-in-the-parameters form:

$$y(t) = \varphi^T(t)\theta + e(t) \tag{2}$$

where $\varphi(t)$ is the regressor vector, containing monomials of lagged input and output terms; θ is the parameter vector; $e(t)$ here is simply a zero mean *i.i.d.* measurement noise sequence. Because the NARX model is the core part required for the frequency domain analysis of a nonlinear system (Peyton Jones and Billings 1989), the modeling and identification procedure of time-varying NARX model is firstly introduced in Section 2.1 and later extended to the more general time-varying NARMAX cases in Section 2.2.

2.1 Identification of a TV-NARX Model

When modeling a time-varying system, the parameter vector θ in a NARX model (2) is replaced with a time-varying parameter vector $\theta(t)$ to obtain a polynomial time-varying NARX (TV-NARX) model.

$$\begin{aligned} y(t) &= \sum_{n=1}^{M} y_n(t) + e(t) \\ y_n(t) &= \sum_{p=0}^{n} \sum_{k_1,k_{p+q}=1}^{K} c_{p,q}(k_1,\ldots,k_{p+q},t) \times \prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} u(t-k_i) \\ &= \varphi^T(t)\theta(t) \end{aligned} \tag{3}$$

where $y_n(t)$ is the nth-order output of the system and M is the order of the nonlinearity, with $p+q=n$, $k_i = 1,\ldots,K$, $\sum_{k_1,k_{p+q}=1}^{K} \equiv \sum_{k_1=1}^{K} \cdots \sum_{k_{p+q}=1}^{K}$. The TV parameter vector is $\theta(t) = [c_{0,1}(1,t),\ldots,c_{0,1}(K,t),c_{1,0}(1,t),\ldots,$ $c_{p,q}(K,\ldots,K,t)]^T$. The model (3) is linear-in-the-parameters, however, it consists of a large number of candidate terms, which increases dramatically as the order (M), the number of input and output terms (q and p) and the corresponding maximum lags (K) of the model increases. The linear TV-ARX model is a special case of the TV-NARX model which only includes the linear terms of the polynomial expression in (3). A general procedure for identifying a nonlinear TV system is presented in following paragraphs, including selection of the appropriate model structure for the TV-NARX model, and parameter estimation.

When modeling a time-varying nonlinear system using the TV-NARX model (3), the classical identification procedure can still be applied in which the model structure is determined first and the TV parameters are then estimated. However, the standard model subset selection algorithms, such as forward OLS, and principal component analysis approximation (Kaipio and Karjalainen 1997), cannot be applied directly for a TV system. These approaches were originally proposed and are only applicable for time invariant systems, and cannot distinguish parametric time-varying effects from extra linear or nonlinear regression terms.

To identify a TV-NARX model, two strategies can be considered. One strategy is to approximate the time-varying parameters using basis functions (e.g. wavelets, Fourier bases) expansion, so that the

TV-NARX is transformed into an expanded time invariant regression model, and the subset selection (e.g. OLS) algorithm can then be applied to the extended model to select the significant basis functions which can then be used to recover the relevant TV parameters in the original TV model. The other strategy is to first divide the input-output time series into several sub-intervals, by treating each segment as a time invariant model (e.g. NARX) and selecting the corresponding model structure using, for example, forward OLS, finally a 'common' model structure can be determined by comparing the selected model terms from different intervals. Based on the selected model structure, a recursive algorithm (e.g. RLS, LMS) can be used to estimate the TV parameters.

The first strategy is theoretically elegant as the forward regression OLS can be used to determine the model structure of the expanded time invariant model, and therefore ensure the correctness of the original TV model structure thanks to the basis function transformation. Also the algorithm can easily track the rapid changes in the TV parameters due to the properties of the basis functions, for example wavelets and the introduction of multi-wavelet decompositions. In this study, a novel identification procedure is proposed based on the first strategy, while the second strategy can be used as a complement to pre-process the data and to reduce the candidate model set especially for large scale problems. The detailed algorithm is introduced in the following paragraphs.

Firstly, each TV parameter in (3) is expanded using multi-wavelet basis functions:

$$c_{p,q}(k_1,\ldots,k_{p+q},t)=\sum_{m}\sum_{l\in\Gamma_m}\alpha_{p,q,l}^{m}(k_1,\ldots,k_{p+q})\phi_l^m\left(\frac{t}{N}\right) \tag{4}$$

where $\phi_l^m(\cdot)$ are wavelet basis functions, with the shift indices $l\in\Gamma_m$, $\Gamma_m=\{k:-m\le l\le 2^j-1\}$, and wavelet scale *j*. The superscript *m* denotes the order of the wavelet basis functions. $\alpha_{i,l}^{m}(\cdot)$ are the corresponding expanded basis function parameters which are time invariant. *N* is the number of measurement data. There are various types of wavelet basis functions which can be combined with different orders in order to track both rapid and slow variations of TV parameters. For example, cardinal B-splines are an important class of basis function and can be used for multi-resolution wavelet decompositions (Ahuja, Lertrattanapanich, and Bose 2005). The B-spline function of m^{th}-order can be computed recursively according to (de Boor 1978; Chui and Wang 1992) and several low order B-spline functions are provided in (Wei and Billings 2006). Then the $\phi_l^m(\cdot)$ can be expressed from the m^{th}-order B-spline (B_m) with the scale j and shift indices l:

$$\phi_{j,l}^m(x) = 2^{j/2} B_m(2^j x - l) \tag{5}$$

since the function variable $x = t/N$ is normalized within [0,1] and $-m \le l \le 2^j - 1$. The higher the value j, the more basis functions are used and thus the resolution improves, however, this will also introduce more parameters and increase the computational cost. Normally j is selected to be 3 or a larger number according to the criteria given in (Wei and Billings 2002), and the wavelets are usually selected from $\{\phi_l^m : m = 2,3,4,5\}$.

By substituting (4) into (3), the TV-NARX model is expanded as

$$\begin{aligned} y(t) &= \sum_{n=1}^{M}\sum_{p=0}^{n}\sum_{k_1,k_{p+q}=1}^{K}\sum_{m}\sum_{l\in\Gamma_m} \alpha_{p,q,l}^m(k_1,\ldots,k_{p+q}) \times \left(\phi_l^m\left(\frac{t}{N}\right) \times \prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} u(t-k_i) \right) + e(t) \\ &= \psi^T(t)\eta + e(t) \end{aligned} \tag{6}$$

where $\psi(t)$ is the expanded regressor vector at time t, and $\eta = [\alpha_{0,1,l}^m, \ldots, \alpha_{1,0,l}^m, \ldots, \alpha_{p,q,l}^m, \ldots]^T$ is the corresponding expanded time invariant parameter vector.

The original TV-NARX model is now transformed into a time invariant regression model which is also linear-in-the-parameters. This enables the parameter estimation problem to be solved within the least squares framework. However, the number of candidate regression terms $\psi(t)$ can be very large if the number of wavelet basis functions, the maximum lags and/or the order of the TV-NARX model is large. Hence, detecting the correct structure and reducing the number of terms in the expanded model becomes a crucial step in the identification of the original nonlinear time-varying problem. The forward regression OLS algorithm is an efficient approach to deal with such model term selection problems (Chen, Billings, and Luo 1989; Billings, Chen, and Korenberg 1989). By applying the forward regression OLS to (6), an appropriate parsimonious model structure can be determined and the relevant coefficients, $\alpha_{p,q,l}^{m}(\cdot)$, can be estimated. By re-substituting those estimated parameters into (4), the original TV parameters $c_i(t)$ can then be recovered.

For a nonstationary process, it is very difficult and possibly impossible to validate the TV model using standard approaches such as splitting the data into training and testing sets, due to the time-varying nature of the model. Also the whiteness or auto-correlation test in the residuals is insufficient for a nonlinear system (Billings and Voon 1983). In order to balance the model complexity (in terms of the number of basis functions) and the value of MSE, a model order determination technique, such as the generalized cross-validation (GCV) criterion (Golub, Heath, and Wahha 1979) or Akaike information criterion (AIC) (Akaike 1974), can be employed to select an appropriate number of wavelet basis functions to be included in the parsimonious model.

2.2 The Stochastic TV-NARMAX Model Identification

To ensure unbiased estimation and accommodate the stochastic perturbations or additive colored (process or measurement) noise in the systems, time-varying moving average (MA) noise terms may need to be further included into the TV-NARX model (3) to form a polynomial TV-NARMAX model:

$$\begin{aligned} y(t) &= \sum_{n=1}^{M} y_n(t) + e(t) \\ y_n(t) &= \sum_{p=0}^{n} \sum_{k_1, k_{p+q}=1}^{K} c_{p,q}(k_1, \ldots, k_{p+q}, t) \times \prod_{i=1}^{p} y(t-k_i) \prod_{i=p+1}^{p+q} u(t-k_i) + \sum_{k_1=1}^{K} \cdots \sum_{k_n=k_{n-1}}^{K} b_{k_1,\cdots,k_n}(t) e(t-k_1) \cdots e(t-k_n) \end{aligned} \tag{7}$$

where sequence $e(t)$ is assumed to be independent, bounded and uncorrelated with the input $u(t)$. (7) can be represented as $y(t) = \varphi^T(t)\theta(t) + \xi^T(t)\beta(t)$, where $\xi(t)$ is the stochastic prediction error vector and $\beta(t)$ is the corresponding parameter vector. Because the stochastic variable $e(t)$ cannot be measured directly, the unobserved noise sequence has to be estimated during identification using a predication error method - known as extended least squares (ELS) algorithm, which has been successfully applied for time invariant NARMAX model identifications (Billings and Zhu 1991). Here, a modified ELS algorithm is proposed and incorporated with the identification procedure discussed in Section 2.1, to solve the identification of TV-NARMAX model (7) as follows:

Step 1: Set $s = 0$. Fit a TV-NARX model (3) to the measurement data, and using the identification procedure proposed in Section 2.1 to determine the TV-NARX model structure and estimate model parameters.

Step 2: Compute the one-step-ahead prediction errors from the identified TV-NARX model as:

$$\begin{aligned} e^{(s)}(t) &= y(t) - \hat{y}(t) = y(t) - \hat{f}(y(t-1),\dots,u(t-1),\dots) \\ &= y(t) - \varphi^T(t)\hat{\theta}^{(s)}(t) \end{aligned} \tag{8}$$

Step 3: Set $s = s + 1$, and update the prediction error vector $\xi^{(s)}(t) = \xi^{(s-1)}(t)$. When $s = 1$, a candidate prediction error set structure $\xi^{(s)}(t) = \left[e^{(s-1)}(t-k_1),\dots,e^{(s-1)}(t-k_1)e^{(s-1)}(t-k_2),\dots\right]$ is firstly arbitrarily selected and this prediction error model structure will be iteratively selected and updated as s increases. Now a complete TV-NARMAX model (7) is employed and the identification procedure is similar to the approach proposed in Section 2.1. More specifically, both TV-NARX and prediction error (TV-MA) model parameters, $\theta^{(s)}(t)$ and $\beta^{(s)}(t)$, are now expanded using B-spline basis functions (5). By using forward regression OLS approach on the expanded model, the prediction error model structure is therefore selected (the TV-NARX model structure is selected in step $s = 0$ and assumed fixed here) and all the time varying parameters in TV-NARMAX model are estimated. The model prediction error can now be updated as

$$\begin{aligned} e^{(s)}(t) &= y(t) - \hat{y}(t) = y(t) - \hat{f}(y(t-1),\dots,u(t-1),\dots,e^{(s-1)}(t-1),\dots) \\ &= y(t) - \varphi^T(t)\hat{\theta}^{(s)}(t) - \xi^{(s)T}(t)\hat{\beta}^{(s)}(t) \end{aligned} \tag{9}$$

This step is repeated until the termination conditions in step 4 are satisfied.

Step 4: Termination tests. The iteration in Step 3 can be terminated when one of the following two convergence tests is satisfied

$$\frac{\left\|\theta^{(s)} - \theta^{(s-1)}\right\|}{\left\|\theta^{(s)}\right\|} \le \delta_1 \qquad \text{or} \qquad \frac{1}{N}\left\|e^{(s)} - e^{(s-1)}\right\| \le \delta_2 \tag{10}$$

where δ_1 and δ_2 are small tolerance values, and $\|\cdot\|$ denotes l_2-norm.

3. Frequency-Domain Analysis of Time-Varying Nonlinear Systems

Compared to time domain identification, the frequency domain analysis of nonlinear time-varying systems is an even greater challenge that does not appear to have been studied in the literature. As an extension of linear FRF, the frequency domain analysis of NTI systems is mainly based on the generalised frequency response function (GFRF) theory. The *n*-th order GFRFs are defined as the multiple Fourier transform of the *n*-th order Volterra kernel:

$$H_n(f_1, f_2, \ldots, f_n) = \int_0^{\infty} \cdots \int_0^{\infty} h_n(\tau_1, \tau_2, \ldots, \tau_n) e^{-j2\pi(f_1\tau_1 + \cdots + f_n\tau_n)} d\tau_1 \ldots \tau_n \tag{11}$$

The concept of GFRFs, introduced based on the Volterra model (Schetzen 1980), can be extended to the NARMAX model case (Billings and Tsang 1989a; Peyton Jones and Billings 1989; Billings and Peyton Jones 1990). The general expression of the *n*th-order GFRF for a time invariant NARX model is given in (Peyton Jones and Billings 1989). It is important to note that the *n*th-order GFRF is a function of the NARX model coefficients which are constants, while the structure of the GFRF expression purely depends on the NARX model structure. As only parametric TV effects are taken into account in this work without considering the time variation in the model structure, previous results obtained for NARX models can be extended to TV-NARX cases, while an extra time dimension relating to the time '*t*' needs to be introduced to accommodate the parametric TV effects and then the TV parameters can be regarded as 'constants' in a local region without affecting the model structure and thus the formulations of the GFRFs. Hence, by replacing the time invariant parameters with TV parameters and including the time dimension *t*, the *n*th-order time-varying GFRFs (TV-GFRFs) with respect to a TV-NARX model (3) can be proposed as

$$H_n(f_1, \ldots, f_n, t) = \frac{H_{n_u}(f_1, \ldots, f_n, t) + H_{n_{uy}}(f_1, \ldots, f_n, t) + H_{n_y}(f_1, \ldots, f_n, t)}{1 - \sum_{k_1=1}^{K} c_{1,0}(k_1, t) e^{(-j2\pi(f_1 + \ldots + f_n)k_1/f_s)}} \tag{12}$$

where the contributions of the pure input, output and cross-product non-linearities, H_{n_u}, H_{n_y} and $H_{n_{uy}}$, are defined in the Appendix; f_s is the sampling frequency. For example, the 1st and 2nd-order TV-GFRFs are explicitly given as:

$$H_1(f,t) = \frac{\sum_{k_1=1}^{K} c_{0,1}(k_1,t)e^{-j2\pi k_1 f/f_s}}{1-\sum_{k_1=1}^{K} c_{1,0}(k_1,t)e^{-j2\pi k_1 f/f_s}} \tag{13}$$

$$\begin{aligned} H_2(f_1, f_2, t) &= H_{2_{uy}}(\cdot) + H_{2_y}(\cdot) + H_{2_u}(\cdot) \\ &= \frac{1}{1-\sum_{k_1=1}^{K} c_{1,0}(k_1,t)e^{-j2\pi k_1 (f_1+f_2)/f_s}} \times \left(\sum_{k_1=1}\sum_{k_2=1} c_{1,1}(k_1,k_2,t)H_1(f_1,t)e^{-j2\pi(f_1k_1+f_2k_2)/f_s} \right. \\ &\quad \left. + \sum_{k_1=1}\sum_{k_2=1} c_{2,0}(k_1,k_2,t)H_1(f_1,t)H_1(f_2,t)e^{-j2\pi(f_1k_1+f_2k_2)/f_s} + \sum_{k_1=1}\sum_{k_2=1} c_{0,2}(k_1,k_2,t)e^{-j2\pi(f_1k_1+f_2k_2)/f_s} \right) \end{aligned} \tag{14}$$

where $c_{1,0}(k_1)$ are the parameters of the linear output terms; $c_{0,2}(k_1,k_2)$, $c_{2,0}(k_1,k_2)$ and $c_{1,1}(k_1,k_2)$ are the parameters of the bilinear input, output and cross-product terms, i.e. $u(t-k_1)u(t-k_2)$, $y(t-k_1)y(t-k_2)$, $u(t-k_1)y(t-k_2)$. For a time invariant system (ignoring the time index *t*, or assuming $t=t_k$), the 2nd-order GFRF (14) can be visualized in a 3-dimensional (3D) space, while higher order GFRFs can be computed recursively from the lower order GFRFs according to (12).

3.1 How the Frequency Domain Features Depend on NARX Model Structure and Parameters

When analyzing the linear frequency response or the 1st-order GFRF of a nonlinear system, the main interest is often to identify the ‘peaks’ in the frequency response, as it will indicate at which frequency ranges the output response would be stronger. When analyzing a higher order GFRF of a nonlinear system, we are interested in not only revealing the ‘peaks’ but more importantly in identifying the ‘ridges’ where the gain of the GFRF reach their maxima. The locations of the ‘ridges’ indicate the frequency combinations in the input excitation that produce strong intermodulation, harmonic, or significant d.c. shift effects in the output (Billings and Tsang 1989b; 1990; Yue, Billings, and Lang 2005a).

In order to analyze and visualize a time-varying nonlinear system in the frequency domain with corresponding TV-GFRFs, an important question that needs to be answered is whether the features and especially the directions and locations of the 'ridges' in a GFRF (i.e. TV-GFRF at a specific time) functionally depend on the structure of the TV-NARX model, or on the parameter values, or both.

By analyzing the general expressions of the 2nd- and higher order TV-GFRFs in (14) and (12), and according to the characteristic analysis of GFRFs given by (Yue, Billings, and Lang 2005b; Jing, Lang, Billings, and Tomlinson 2006), it can be revealed that the features in the GFRFs in general depend on both the structure and the parameter values of the model. However, this study will demonstrate that the directions of the 'ridges' are mainly dependent on the NARX model structure, and are rarely affected by changes in one or a few parameters, while the positions and magnitudes of the 'ridges' depend on the model parameters. This result is summarized in Proposition 1.

Proposition 1: The directions of the 'ridges' in the *n*th-order GFRF (gain) mainly depend on the NARX model structure. There always exist 'ridges' along the $f_1+f_2+\cdots+f_n=C_i$ direction, and often with extra 'ridges' along the 'off-diagonal' directions of subspaces in the *n*-dimensional frequency domain, i.e. $f_j=C_i$, $f_j+f_{j+1}=C_i$,…, $f_1+\cdots+f_{j+n-2}=C_i$ with $j\in\{1,\cdots,n\}$ depending on the model nonlinear part structure. The positions of the above mentioned 'ridges' only depend on the parameters of the linear output terms, i.e.

$$C_i = \arg\min_{f_1+f_2+\cdots+f_n}\left|1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1(f_1+f_2+\cdots+f_n)/f_s}\right|$$

while the magnitudes of the 'ridges' can depend on all the parameters of the NARX model.

Proof: The proof begins with the 2nd-order GFRF, and the results are generalized to higher-order GFRFs later. As discussed in (Yue, Billings, and Lang 2005b), the factors in the denominator jointly determine the maxima of the *n*th-order GFRF. Since the 'ridges' are a subset of the maxima, only the denominator is required to be analyzed. By substituting the 1st-order GFRF expression (13) into the 2nd-order GFRF (14), the denominator of the GFRF becomes

$$\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1 f_1/f_s}\right)\times\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1 f_2/f_s}\right)\times\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1 (f_1+f_2)/f_s}\right) \tag{15}$$

The denominator in (15) is the product of three parts; the first term is introduced from $H_1(f_1,t)$, the second term is introduced from $H_1(f_2,t)$, and the last term is the default denominator in (14). At a specific time and frequencies when the gain of any of these parts reaches a minimum, the overall gain of the GFRF will reach a large value. A subset of these 'large values' will appear as a 'ridge' in the GFRF gain plot.

For some models with specific structures, only the last part in the denominator (15) exists, if i) there are no linear input terms (i.e. $u(t-k_1)$), or ii) there exist linear input terms in the linear part, but there exist only pure nonlinear input terms without any other 2nd-order nonlinear terms. For the above two cases, the 'ridges' in the 2nd-order GFRF are always along the direction defined by $f_1 + f_2 = C_i$, where the constants C_i determine the position of the ridges when the first part of (15) reaches a minimum, that is

$$C_i = \arg\min_{f_1+f_2}\left|1-\sum_{k_1=1}^{K} c_{1,0}(k_1,t)e^{-j2\pi k_1(f_1+f_2)/f_s}\right| \tag{16}$$

It is obvious that the positions of the ridges only depend on the parameters of the linear output terms, since C_i only functionally depends on $c_{1,0}(k_1,t)$ in (16) when the model structure (i.e. k_1) is fixed. It is also straightforward to demonstrate that if there is only one linear output delay term (i.e. K=1) in the model, $C_i = 0$. Otherwise, if there are more than one linear output terms, C_i can be non-zero.

Moreover, if there are linear input terms in the model, and either cross-product or pure output nonlinear terms exist, then the second or both the second and third parts in (15) would be present together with the first part. Hence extra ridges that are perpendicular to either or both the frequency axis, i.e. $f_1 = C_i$ and/or $f_2 = C_i$, can be observed additional to the ridges along the direction $f_1 + f_2 = C_i$. Again, the positions of those ridges also depend on the parameters of the linear output model.

The above results can be generalized to the higher-order GFRFs. By substituting (26)-(28) into (12), the denominator of the nth-order GFRF can be expressed as

$$\prod_{i=1}^{n}\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1 f_i/f_s}\right)\prod_{i=1}^{n-1}\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1(f_i+f_{i+1})/f_s}\right)\times\cdots\times\left(1-\sum_{k_1=1}^{K}c_{1,0}(k_1,t)e^{-j2\pi k_1(f_1+\cdots+f_n)/f_s}\right) \tag{17}$$

As long as the n^{th}-order GFRF exists for a given nonlinear system, the last part in (17) always exists, while other parts may appear depending on the NARX model structure. This indicates that there always exists a ridge in the hyperspace along $f_1+f_2+\cdots+f_n=C_i$, and sometimes with extra ridges within one or a few of the sub-spaces, for example $f_j=C_i$, $f_j+f_{j+1}=C_i$, …, $f_j+\cdots+f_{j+n-2}=C_i,\ j\in\{1,\cdots,n\}$, which again depends on the NARX model structure. The positions of these ridges can be determined in a similar way as (16). The magnitudes of the ridges are affected by both the factors in the denominator and numerator of the GFRF, and hence the magnitudes depend on all the model parameters up to n^{th}-order. □

The locations and magnitudes of ridges in an n^{th}-order GFRF can indicate the transfer of energy from input spectral components to the output components at their summation, known as intermodulation effects, and determine the strength of effects. Additionally, when $f_1=f_2=\cdots=f_n$, peaks in the reduced sub-plane, $H_n(f_1,f_1,\ldots,f_1)$, would indicate whether the n^{th} harmonics will have significant effects on the output.

To illustrate the statements in Proposition 1, an example is given here where only the 2nd-order GFRF will be discussed. The simplified NARX model from (Billings and Tsang 1990) is given as

$$y(k)=1.604y(k-1)-0.9493y(k-2)+0.06187u(k-1)-0.01373y(k-1)y(k-1) \tag{18}$$

Since a pure output nonlinear term exists in the model (18), all the three parts in the denominator (15) of the 2nd-order will exist. Hence, the ridges in the GFRF in all three directions can be observed as in the left plot of Figure 1. According to (16), the positions of the ridges can be determined as $f_1+f_2=\pm0.5$, $f_1=\pm0.5$ and $f_2=\pm0.5$. If the NARX model structure is modified by replacing the pure output nonlinear term (i.e. $y(k\text{-}1)y(k\text{-}1)$) with a cross-product term (i.e. $y(k\text{-}1)u(k\text{-}1)$) without changing the parameter value, there will be ridges along two directions, $f_1+f_2=\pm0.5$, $f_2=\pm0.5$, as shown in Figure 1 central plot. If the nonlinear part is replaced with a pure input nonlinear term (i.e. $u(k\text{-}1)u(k\text{-}1)$), there are no ridges along the directions of the axes. The only ridges are along $f_1 + f_2 = \pm 0.5$. If any parameters in the linear output terms in (18) vary, only the positions of the ridges in Figure 1 would change without affecting their directions, and the variation of any other parameters only affects the magnitudes of the ridges, which clearly support the theoretical analysis above.

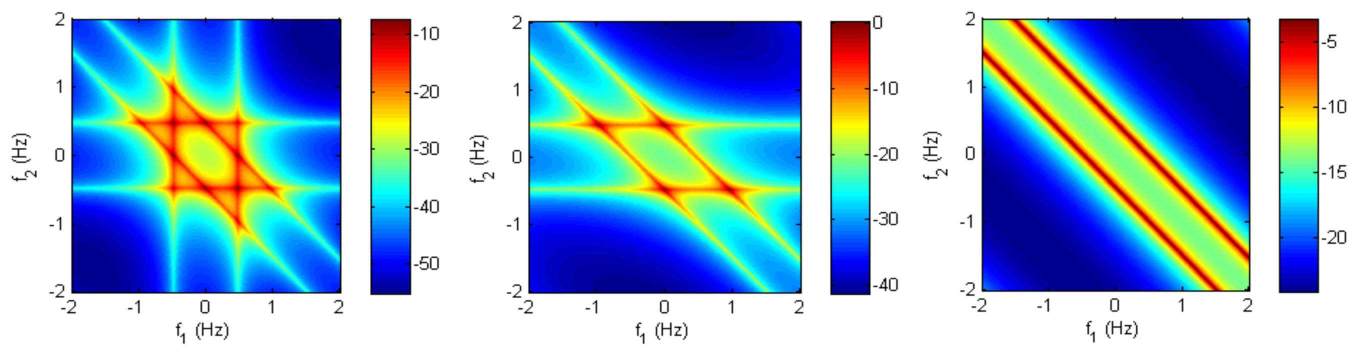

Figure 1. The contours of the 2nd-order GFRF gain plots with respect to the pure output (left), cross-product (central), or pure input (right) nonlinear terms in the NARX model (18).

The analysis in this section indicates that the directions of the 'ridges' mainly depend on the NARX model structure, and that the parameters of the linear output terms determine their positions. However, there can be exceptions. This is because the overall gain of an nth-order TV-GFRF at specific times and frequencies is essentially a ratio between the gain of the numerator with that of the denominator of the GFRF as in (12). Hence, the overall direction of a 'ridge' in a GFRF is actually a trade-off between the two ridge directions that are implicitly contained in the numerator and denominator. In most cases, the ridge contained in the denominator dominates the overall direction. However, for some special cases when the gradient (changing rate) of the ridge in the denominator is extremely small, the ridge in the numerator can dominate. Such exceptions can only happen when all the parameters of the linear output terms are very small in magnitude, for example $c_{1,0}(k_i, t) \approx 0,\ i = 1, \ldots, K$, as a result the gain with respect to the denominator of a nth-order GFRF would always be close to 1 (as in (17)) with a small gradient. In such cases, the varying of any parameters that change the ratio can affect the directions of the ridges. Nevertheless, such situations rarely occur in practice, as the parameter scales of the linear terms are usually larger or at least similar with those of the nonlinear terms.

3.2 Visualization of TV-GFRFs

When studying a time-varying nonlinear system in frequency, the *n*th-order TV-GFRF would be in an *n*+2 dimensional space (e.g. the 2nd-order TV-GFRF would be in a 4D space), which is in general difficult to visualize and analyze. However, as analyzed in Section 3.1 that only several 'ridges' along upto very few directions are often observed in the GFRF of a nonlinear system, and the directions of the ridges depend on the (TV)-NARX model structure which can be determined from Proposition 1. An important approach to analysis TV-GFRFs can then be developed by averaging all the elements in a GFRF along those specific 'ridge' directions at each sampling time, and the variations of the magnitudes and positions of one or several 'ridges' can therefore be 'tracked' and visualized as the time-varying parameters change. Again considering the example in (18) with pure input nonlinearity, the only 'ridges' in the 2nd-order GFRF are along the $f_1 + f_2 = C_i$ direction. Then the gain of TV-GFRF can be averaged along this direction as

$$\left|H_2^{f_1+f_2}(f,t)\right| = \int_{f_1+f_2=f} \frac{1}{N_f}\left|H_2(f_1, f_2, t)\right| df \tag{19}$$

with $H_2(f_1, f_2, t)$ given in (14), and the N_f denotes the number of samples along the 'ridge' direction at each frequency f, i.e. $f = f_1 + f_2$. The phase of the 2nd-order TV-GFRF can be averaged in a similar way.

Such averaging strategy can be easily extended to the visualization of higher-order TV-GFRFs. The n^{th}-order TV-GFRF can be averaged along a specific 'ridge' direction

$$\begin{aligned} \left|H_n^{f_1+f_2+\cdots+f_j}(f,t)\right| &= \int_{f_1+f_2+\cdots+f_j=f} \frac{1}{N_f}\left|H_n(f_1,\ldots, f_n,t)\right| df \\ \phi\left(H_n^{f_1+f_2+\cdots+f_j}(f,t)\right) &= \int_{f_1+f_2+\cdots+f_j=f} \frac{1}{N_f}\phi\left(H_n(f_1,\ldots, f_n,t)\right) df \end{aligned} \tag{20}$$

where $\phi(\cdot)$ denotes phase. When the subscript $j=n$, it represents an averaging along the main ridge direction $f_1+f_2+\cdots+f_n=C_i$. When $j<n$, (20) represents averaging along the other ridges in the $j+1$ dimensional subspace, and with the rest (i.e. n-j) frequency dimensions fixed.

4. Simulation Examples

Two examples are studied in this section. One is a simulated nonlinear time-varying system with measurement data generated from a known TV-NARX or TV-NARMAX system. It is demonstrated that the proposed algorithm can accurately identify the true system structure and estimate the time-varying and time invariant parameters. Also based on the new frequency domain analysis approach, the identified TV-NARX model is mapped to the TV-GFRFs to be analyzed and visualized. The second example is based on a simulated model of a true nonlinear time-varying damping system, which is a 2nd-order continuous-time differential equation with 3rd-order nonlinearity. It is demonstrated that the proposed TV-NARX approach can estimate and recover the time-varying parameters of the continuous-time nonlinear and nonstationary system, and the corresponding frequency domain analysis can then be investigated based on the identified discrete-time TV-NARX model.

4.1 A TV-NAR(MA)X Example

The measurement data were generated from the following TV-NARX system:

$$y(k)=c_{1,0}(1,k)y(k-1)+c_{1,0}(2,k)y(k-2)+c_{0,1}(1,k)u(k-1)+c_{0,2}(2,2,k)u(k-2)^2+e(k) \tag{21}$$

where the time-varying and time invariant parameters are given as:

$$c_{1,0}(1,k)=\begin{cases}1, & 0\le k\Delta t<0.2s\\ 0.4, & 0.2\le k\Delta t<0.4s\\ 0.8, & 0.4\le k\Delta t\le 1s\end{cases},\quad c_{1,0}(2,k)=-0.3,$$

$$c_{0,1}(1,k)=0.1,\quad c_{0,2}(2,2,k)=\begin{cases}0.2, & 0\le k\Delta t<0.6s\\ 0.5, & 0.6\le k\Delta t<0.8s\\ 0.3, & 0.8\le k\Delta t\le 1s\end{cases}$$

The input $u(k)$ was a Gaussian random sequence with sampling time Δt = 0.0025s (f_s =400Hz), and variance 1. $e(k)$ was the additive Gaussian *i.i.d.* noise, with variance σ^2 = 0.0004, to give a signal-to-noise ratio (SNR) of around 16dB. The set of initial candidate model terms was selected as {$y(k$-1), $y(k$-2), $u(k$-1), $u(k$-2), $u(k$-1)^2, $u(k$-1)$u(k$-2), $u(k$-2)^2} with corresponding parameters $c_{1,0}(1,k)$, $c_{1,0}(2,k)$, $c_{0,1}(1,k)$, $c_{0,1}(2,k)$, $c_{0,2}(1,1,k)$, $c_{0,2}(1,2,k)$, $c_{0,2}(2,2,k)$. B-spline basis functions were then used to expand the time-varying parameters and transform them into constant kernel parameters. The B-spline basis functions were selected from $\{\phi_l^m : m = 3,4,5\}$, with scale index j=4. Then the OLS algorithm was applied to select the appropriate model structure from the expanded candidate terms and estimate the parameters only based on the available input and output data. After selection and estimation, the original time-varying parameters can be re-constructed and are displayed in the Figure 2 (right plot). For comparison, the parameter estimates using RLS (with forgetting factor λ=0.95) is also provided in Figure 2 (left plot).

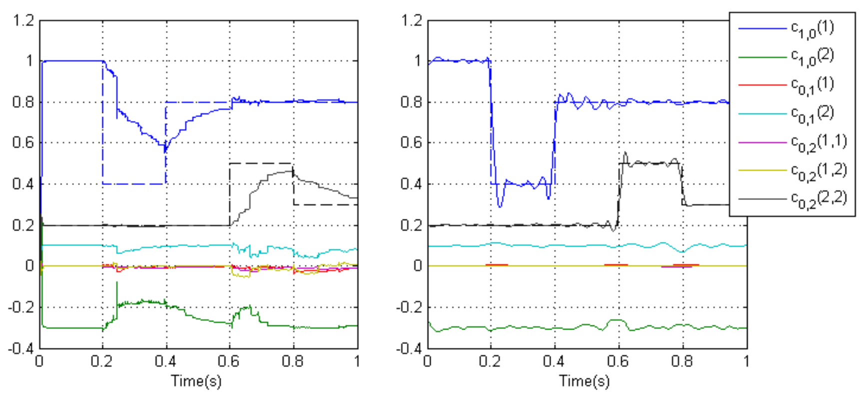

Figure 2. Parameters estimates of candidate model terms using RLS (left) and wavelet with OLS approach (right), with true TV parameter values are shown as dashed lines.

Figure 2 demonstrates that standard RLS estimation fails to track the rapid changes in the time-varying parameters, e.g. $c_{1,0}(1)$, $c_{0,2}(2,2)$. In contrast, the proposed wavelet with OLS approach can accurately estimate and track the time-varying parameters, estimate time invariant parameters, and set all the other parameters of the ‘redundant’ candidate terms to 0. The estimation results have been tested under different SNRs, and the proposed algorithm can effectively identify and track the TV parameters even with an SNR of 13dB. The advantages of using the wavelet expansion based approach compared to other classical recursive approaches, e.g. LMS, Kalman filer, have been demonstrated in the literature (e.g. Tsatsanis and Giannakis 1993; Li, Wei, and Billings 2011) for linear systems which show excellent rapid tracking behavior with smaller estimation variances. The proposed approach can therefore successfully identify the model structure for a discrete-time nonlinear and nonstationary system, and accurately estimate the time-varying and time invariant parameters. To test a more realistic and general noise situation, the data generated from the following TV-NARMAX model is used as a new input-output measurement set,

$$y(k)=c_{1,0}(1,k)y(k-1)+c_{1,0}(2,k)y(k-2)+c_{0,1}(1,k)u(k-1)+c_{0,2}(2,2,k)u(k-2)^2+e(k)+b_1(k)e(k-1) \quad (22)$$

where the deterministic part of the TV-NARMAX model is the same as TV-NARX model (21) but with an extra moving average term, and $b_1(k)=0.5\sin(2\pi k\Delta t)$. The initial candidate input and output model terms are set to be the same as the TV-NARX case and with extra moving average candidate set $\{e(k\text{-}1), e(k\text{-}2), e(k\text{-}1)^2, e(k\text{-}1)e(k\text{-}2), e(k\text{-}2)^2\}$. By applying the proposed identification algorithm in Section 2.2, and after 6 iterations in step 3 the algorithm terminated with the correct moving average terms selected. The parameter estimates of the deterministic part of the model are given in Figure 3.

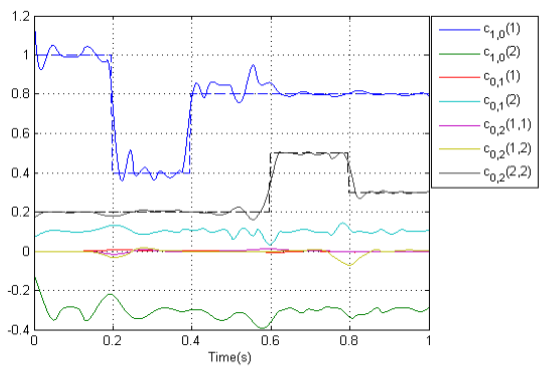

Figure 3. Parameters estimates of TV-NARMAX model terms using proposed ELS approach, with true TV parameter values are shown as dashed lines.

Figure 3 demonstrates that the by using the proposed algorithm the time-varying parameters in the TV-NARMAX model can still be relatively well identified and tracked in time. However, because the time-varying moving average noise terms enter into the output of the system iteratively as in (22) and such effect will affect the estimation accuracy, both parameter estimates of time-varying and time invariant parameters in the TV-NARMAX model are not as accurate as the TV-NARX case in some time intervals as given in Figure 2. The following frequency domain analysis results are based on the identification results of the TV-NARX system in Figure 2.

Based on the estimated time domain TV-NARX model, the 2nd-order GFRF at two different sampling time points was computed from (14). The gain plots are illustrated in Figure 4.

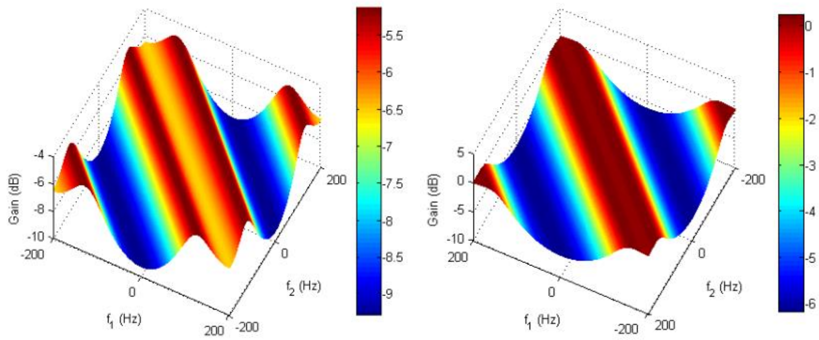

Figure 4. The gain of 2nd-order GFRF (i.e. $H_2(f_1,f_2,t)$) at different sampling times: $k\Delta t$ = 0.3s (left) and $k\Delta t$ = 0.7s (right)

As shown in Figure 4, the ridges of the 2nd-order GFRF are always along the $f_1 + f_2 = C_i$ direction, although with different positions (i.e. C_i = ±80 and ±40) and magnitudes. This indicates there would be a significant intermodulation effect on the output response whenever frequencies of the input excitation sum to C_i. Based on the proposed approach in Section III, the time dependent 1st and the averaged 2nd order TV-GFRFs can be computed according to (13) and (19), and the contours of corresponding gain plots are illustrated in Figure 5.

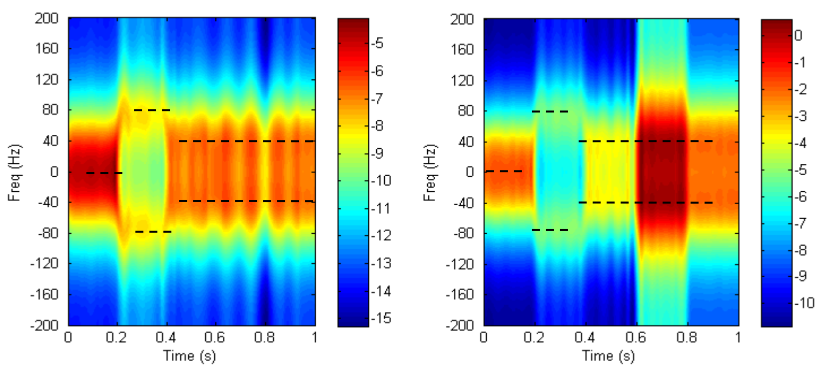

Figure 5. The contours of the 1st-order TV-GFRF (H_1) (left) and the averaged 2nd-order TV-GFRF (H_2) (right) gain plots based on the estimated TV-NARX model. Ridges are marked with black dashed lines.

Figure 5 presents the complete time-frequency characteristics of the identified time-varying system, and clearly demonstrates how the features in TV-GFRFs rely on the time-varying parameters. As expected from the theoretical analysis, the positions in the 1st-order TV-GFRF purely depend on the parameters of the linear output terms (i.e. $c_{1,0}(1,t)$ and $c_{1,0}(2,t)$) while the magnitudes depend on both the linear input and output terms. Here, $c_{1,0}(1,t)$ is the only time-varying parameter within the linear terms, so the positions and magnitudes of the ridges in the 1st-order TV-GFRF change only as this parameter varies, as in the left plot in Figure 5. For the 2nd-order TV-GFRF, the ridge positions are still only affected by the parameters of the linear output terms. Nevertheless, the magnitudes of the ridges are affected by the time-varying parameters of both the linear and nonlinear terms as shown in the right plot in Figure 5. Such visualization of TV-GFRFs greatly facilitates the analysis of time-varying nonlinear systems in practice. For instance, with the positions and magnitudes of the ridges known, it is easy to tell when there are significant intermodulation effects on the output response if the sum of frequency components in input excitation is close to a specific ridge and how strong the effects are, without actually computing the output spectrum from the real experiments or simulations.

4.2 Time-Varying Nonlinear Damping Oscillator

Time-varying and nonlinear damping systems have been observed in various fields of engineering, such as the rain-wind induced vibrations to the dynamics of cable-stayed bridges (van der Burgh and Hartono 2004), and the negative aeroelastic damping in a wind turbine (Murtagh and Basu 2007). The modeling and identification of such systems are important for many practical purposes, including response prediction, real-time monitoring and diagnosis, load identification and control. A general representation of time-varying nonlinear autonomous damping systems is provided in (Jang, Baek, Kim, and Moon 2011), by further including the input signals *u*, a modified representation is given as follows

$$m\ddot{y}(t)+ky(t)=\lambda(t)f(\dot{y}(t))+r(y(t))+u(t) \tag{23}$$

where m and k are constants, $r(\cdot)$ and $f(\cdot)$ denote general nonlinear functions of the displacement y and the velocity $\dot{y}$, $\lambda(t)$ is the time-varying damping coefficient. For illustration, a simple model is discussed here with $f(\dot{y}) = \dot{y}$ and $r(y) = 0.1y^3$. The time-varying damping coefficient is set to $\lambda(t) = 2(1+0.5\sin(t))$, and the input excitation is given by $u(t) = A\sin(\omega_1 t) + A\sin(\omega_2 t)$ with a sampling time $\Delta t = 0.02$s, where $A = 2$, $\omega_1 = 2\pi$, and $\omega_2 = 1.6\pi$. This continuous time-varying system can be modeled and identified using a discrete-time TV-NARX model based on a numerical integration scheme (e.g. Euler, or Runge-Kutta method). Here, by using the Euler method, i.e. $\dot{y} \approx [y(t+\Delta t) - y(t)]/\Delta t$ and $\ddot{y} \approx [\dot{y}(t+\Delta t) - \dot{y}(t)]/\Delta t$, the original differential equation model (23) can be approximated as

$$y(k\Delta t) = (2-\lambda(t)\Delta t)y((k-1)\Delta t) - (1-\lambda(t)\Delta t + \Delta t^2)(y(k-2)\Delta t) - 0.1\Delta t^2(y(k-2)^3\Delta t) - \Delta t^2(u(k-2)\Delta t) \tag{24}$$

When the sampling time Δt is relatively small, (24) can provide an accurate approximation to (23). The above TV-NARX model was two time-varying parameters in the linear output terms, $c_{1,0}(1,t) = 2-\lambda(t)\Delta t$ and $c_{1,0}(2,t) = 1-\lambda(t)\Delta t+\Delta t^2$. After the TV-NARX model (24) is identified using the procedure in Section II, the time-varying damping coefficient in (23) can be estimated from

$$\hat{\lambda}(t) = \frac{2-\hat{c}_{1,0}(1,t)}{\Delta t} \quad \text{or} \quad \hat{\lambda}(t) = \frac{1+\hat{c}_{1,0}(2,t)+\Delta t^2}{\Delta t} \tag{25}$$

The estimation results of the time-varying damping coefficient $\lambda(t)$ using the proposed multi-wavelet expansion approach and the standard RLS approach are illustrated in Figure 6. It is shown that parameter estimates can follow the true time-varying coefficient very well using the new algorithm, while RLS fails to track the true parametric variation.

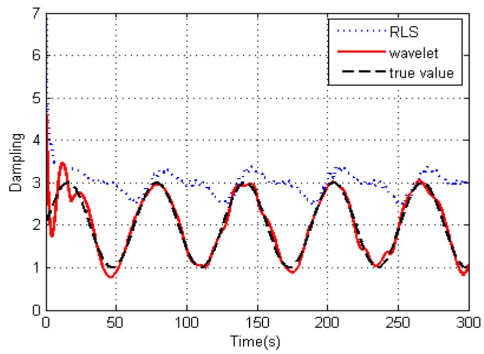

Figure 6. The estimation of the time-varying damping coefficient $\lambda(t)$ based on the multi-wavelet basis function expansion (solid red) and the RLS (dot blue).

Figure 6 also shows that the time-varying damping parameter oscillates between values of 1 and 3. The contours of the 3rd-order GFRF gain plots with the damping coefficient at those two extreme values (i.e. $\lambda(t)$=1 and $\lambda(t)$=3) are given in Figure 7. Since the existence of nonlinear output terms (i.e. $y(k\text{-}2)^3$) in the system (24), the ridges along $f_1+f_2+f_3=C_i$, $f_1(f_3)=C_i$, $f_2=C_i$ directions can be observed in both cases (ridges along $f_1(f_3)+f_2=C_i$, $f_1+f_3=C_i$, are overlapped with $f_1+f_2+f_3=C_i$, $f_1(f_3)=C_i$, since $f_1=f_3$), although with different positions (i.e. C_i=0.35 and 0).

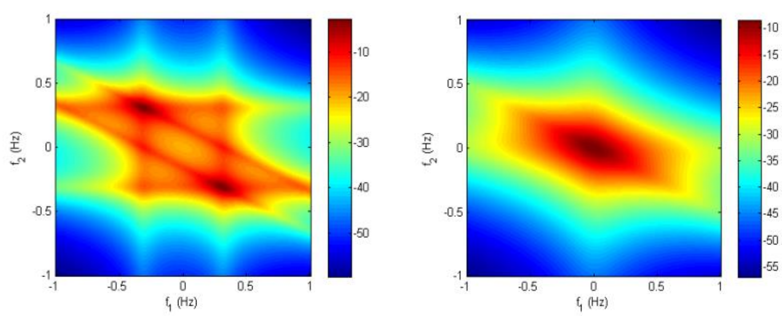

Figure 7. The contours of the 3rd-order GFRF gain plots (i.e. $H_3(f_1,f_2, f_3,t)$, $f_1= f_3$) when the damping coefficient: $\lambda(t)$=1 (left) and $\lambda(t)$=3 (right).

The complete time-dependent frequency domain information in the 1st and averaged 3rd-order TV-GFRFs is illustrated in Figure 8. Since the time-varying damping coefficient $\lambda(t)$ only affects the parameters of the linear output terms in the TV-NARX model (24), both the 1st and averaged 3rd-order TV-GFRFs present very similar contours in the gain figures although with different magnitudes. The right plot in Figure 8 demonstrates how the ridges positions in the 3rd-order GFRF (along the $f_1+f_2+f_3$ direction) vary as a result of the periodic nonstationarity in the damping coefficient. At some time point, only a single ridge along $f_1+f_2+f_3=0$ is observed, but gradually it evolves to two ridges which later converge to one ridge again. This indicates the time-varying damping coefficient significantly affects the frequency range at which the intermodulation effects can be introduced and the strength of the effects, although there are no time-varying effects in the nonlinear part of the system. This result provides a complete time-frequency view of the system under investigation that summarizes the key features in the sampled GFRFs at all time, which cannot be achieved by existing standard techniques.

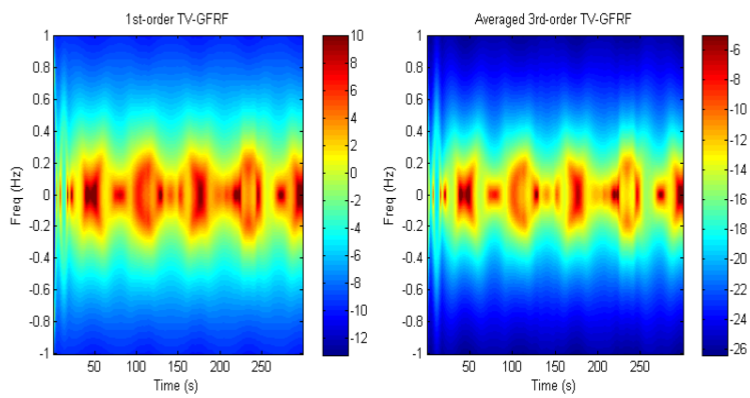

Figure 8. The contours of the 1st-order TV-GFRF (H_1) gain plot (left) and the averaged 3^{nd}-order TV-GFRF (H_2) gain plot along the $f_1+f_2+f_3$ direction (right) based on the estimated TV-NARX model.

5. Conclusions

The identification of nonlinear and nonstationary systems has long been a challenging task, and the frequency domain analysis of such systems has not been investigated systematically. In this study, a complete identification procedure has been proposed based on the multi-wavelet basis function expansion of the TV-NARX model parameters and an OLS algorithm. The system nonstationarity is automatically detected, and the system nonlinearity is identified as part of the model selection process. The numerical examples indicate the proposed approach can estimate and track fast changing time-varying parameters as well as any time invariant model parameters, which are more accurate than traditional RLS algorithm. Furthermore, based on the identified time domain model, a novel frequency domain analysis approach is presented based on a TV-GFRF concept. How features in the TV-GFRFs depend on the TV-NARX model structure and time-varying parameters have been theoretically quantified. The results enable the analysis of nonlinear nonstationary systems in the frequency domain, where the high dimensional TV-GFRFs can be visualized in a low dimensional time-frequency space.

Appendix

The contributions of the pure input, output and cross-product non-linearities in (12) are given as

$$
\begin{aligned}
H_{n_u}(f_1,\ldots,f_n,t) &= \sum_{k_1,k_n=1}^{K} c_{0,n}(k_1,\ldots,k_n,t)e^{-j2\pi(f_1k_1+\cdots+f_nk_n)/f_s} \\
H_{n_{uy}}(f_1,\ldots,f_n,t) &= \sum_{q=1}^{n-1}\sum_{p=1}^{n-q}\sum_{k_1,k_{p+q}=1}^{K} c_{p,q}(k_1,\ldots,k_{p+q},t) \\
&\quad \times H_{n-q,p}(f_1,\ldots,f_{n-q},t)e^{-j2\pi(f_{n-q+1}k_{n-q+1}+\cdots+f_{p+q}k_{p+q})/f_s} \\
H_{n_y}(f_1,\ldots,f_n,t) &= \sum_{p=2}^{n}\sum_{k_1,k_n=1}^{K} c_{0,n}(k_1,\ldots,k_n,t)H_{n,p}(f_1,\ldots,f_n,t)
\end{aligned}
\tag{26}
$$

The contribution of the *p*th order non-linearity in $y(t)$ to the *n*th order TV-GFRF, $H_{n,p}(\cdot)$, can be recursively computed according to (Peyton Jones and Billings 1989) as

$$
H_{n,p}(\cdot) = \sum_{i=1}^{n-p+1} H_i(f_1,\ldots,f_i,t)H_{n-i,p-1}(f_{i+1},\ldots,f_n,t)e^{-j2\pi(f_1+\ldots+f_i)k_p/f_s} \tag{27}
$$

The above recursion finishes with p=1, where the $H_{n,1}(f_1,\ldots,f_n,t)$ is defined as

$$
H_{n,1}(f_1,\ldots,f_n,t) = H_n(f_1,\ldots,f_n,t)e^{-j2\pi(f_1+\ldots+f_i)k_1/f_s} \tag{28}
$$

Acknowledgements

The authors would like to acknowledge the support of the Engineering and Physical Sciences Research Council (EPSRC), U.K. and the European Research Council (ERC).

References

Ahuja, N., Lertrattanapanich, S., and Bose, N.K. (2005), ‘Properties determining choice of mother wavelet’, *IEE Proceedings Vision, Image and Signal Processing*, 152, 659-663.

Akaike, H. (1974), ‘A new look at the statistical model identification’, *IEEE Transactions on Automatic Control*, 19, 716-723.

Ball, J.A., Gohberg, I., and Kaashoek, M.A. (1995), ‘A frequency response function for linear, time-varying systems’, *Mathematics of Control, Signals, and Systems*, 8, 334-351.

Bermudez, J.C.M., and Bershad, N.I. (1996), ‘Transient and tracking performance analysis of the quantized LMS algorithm for time-varying system identification’, *IEEE Transactions on Signal Processing*, 44, 1990–1997.

Billings S A Nonlinear System Identification: Narmax methods in the time, frequency, and spatio-temporal domains, Wiley Sept 2013

Billings, S.A., and Tsang, K.M. (1989a), ‘Spectral analysis for non-linear systems, part I: parametric nonlinear spectral analysis’, *Journal of Mechanical Systems and Signal Processing*, 3, 319-339.

Billings, S.A., and Tsang, K.M. (1989b), ‘Spectral analysis for non-linear systems, part II: Interpretation of non-linear frequency response functions’, *Journal of Mechanical Systems and Signal Processing*, 3, 341-359.

Billings, S.A., Chen, S., and Korenberg, M.J. (1989), ‘Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator’, *International Journal of Control*, 49, 2157-2189.

Billings, S.A., and Tsang, K.M. (1990), ‘Spectral analysis for non-linear systems, part III: case study examples’, *Journal of Mechanical Systems and Signal Processing*, 4, 3-21.

Billings, S.A., and Peyton Jones, J.C. (1990), ‘Mapping nonlinear integro-differential equations into the frequency domain’, *International Journal of Control*, 52, 863-879.

Billings, S.A., and Li, L. (2000), ‘Reconstruction of linear and nonlinear continuous time system models using the kernel invariance algorithm’, *Journal of Sound and Vibration*, 233, 877-896.

Billings, S.A., and Voon, W.S.F. (1983), 'Structure detection and model validity tests in the identification of nonlinear systems', *Proc.IEE, Pt.D*, 130, 193-199.

Billings, S.A., and Zhu, Q.M. (1991), ‘Rational model identification using an extended least-squares algorithm’, *International Journal of Control*, 54:3, 529-546.

Boyd, S., Tang, Y.S., and Chua, L.O. (1983), ‘Measuring Volterra kernels’, *IEEE Transactions on Circuits and Sys*tems, 30, 571-577.

Chen, S., and Billings, S.A. (1989), ‘Representations of non-linear systems: the NARMAX model’, *International Journal of Control*, 49, 1012-32.

Chen, S., Billings, S.A., and Luo, W. (1989), ‘Orthogonal least squares methods and their application to nonlinear system identification’, *International Journal of Control,* 50, 1873-1896.

Chowdhury, F.N. (2000), ‘Input-output modelling of nonlinear systems with time-varying linear models’, *IEEE Transactions on Automatic Control*, 45, 1355-1358.

Chua, L.O., and Liao, Y. (1989), ‘Measuring Volterra kernels (II)’, *Int. J. Circuit Theory and Application*, 17, 151-190.

Chui, C.K., and Wang, J.Z. (1992), ‘On compactly supported spline wavelets and a duality principle’, *Transactions of the American Mathematical Society*, 330, 903-915.

de Boor, C. (1978), *A Practical Guide to Spline*, Springer Verlag.

Fitzgerald, W., Smith, R., Walden, A., and Young, P. (2000), *Nonlinear and Nonstationary Signal Processing*, Cambridge University Press, Cambridge, UK.

Golub, G.H., Heath, M., and Wahha, G. (1979), ‘Generalized cross-validation as a method for choosing a good ridge parameter’, *Technometrics*, 21, 215-223.

Iatrou, M., Berger, T.W., Marmarelis, V.Z. (1999), ‘Modeling of nonlinear nonstationary dynamic systems with a novel class of artificial neural networks’, *IEEE Transactions on Neural Networks*, 10, 327-339.

Jang, T.S., Baek, H., Kim, M.C., and Moon, B.Y. (2011), ‘A new method for detecting the time-varying nonlinear damping in nonlinear oscillation systems: nonparametric identification’, *Mathematical Problems in Engineering*, 2011, 1-12.

Jing, X.J., Lang, Z.Q., Billings, S.A., and Tomlinson, G.R. (2006), ‘The parametric characteristic of frequency response functions for nonlinear systems’, *International Journal of Control*, 79, 1552-1564.

Kaipio, J.P., and Karjalainen, P.A. (1997), ‘Estimation of event-related synchronization changes by a new TVAR method’, *IEEE Transactions on Biomedical Engineering*, 44, 649-656.

Leontaritis, I.J., and Billings, S.A. (1985), ‘Input-output parametric models for nonlinear systems, part I: deterministic nonlinear systems’, *International Journal of Control*, 41, 303-328.

Li, Y., Wei, H.L., and Billings, S.A. (2011), ‘Identification of time-varying systems using multi-wavelet basis functions’, *IEEE Transactions on Control System Technology*, 19, 656-663.

Li, Y., Wei, H.L., Billings, S.A., and Sarrigiannis, P.G. (2011), ‘Time-varying model identification for time-frequency feature extraction from EEG data’, *Journal of Neuroscience Methods*, 196, 151-158.

Ljung, L., and Soderstrom, T., (1983), *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA.

Ljung, L., and Gunnarsson, S., (1990), ‘Adaptation and tracking in system identification – a survey’, *Automatica*, 26, 7-21.

Marmarelis, V.Z. (1981), ‘Practicable identification of nonstationary nonlinear systems’, *IEE Proceedings D: Control Theory and Applications*, 5, 211–214.

Murtagh, P.J., and Basu, B. (2007), ‘Identification of equivalent modal damping for a wind turbine at standstill using Fourier and wavelet analysis’, *Journal of Multi-body Dynamics*, 221, 1464–4193.

Niedzwiecki, M. (1988), ‘Functional series modeling approach to identification of nonstationary stochastic systems’, *IEEE Transactions on Automatic Control*, 33, 955-961.

Niedzwiecki, M., and Klaput, T. (2002), ‘Fast recursive basis function estimators for identification of time-varying processes’, *IEEE Transactions on Signal Processing*, 50, 1925-1934.

Peng, H., Nakano, K., and Shioya, H. (2007), ‘Nonlinear predictive control using neural nets-based local linearization ARX model-stability and industrial application’, *IEEE Transactions on Control System Technology*, 15, 130-143.

Peyton Jones, J.C., and Billings, S.A. (1989), ‘A recursive algorithm for computing the frequency response of a class of nonlinear difference equation models’, *International Journal of Control*, 50, 1925–1940.

Schetzen, M. (1980), *The Volterra and Wiener Theories of Nonlinear Systems*, Chichester: John Wiley.

Tsatsanis, M.K., and Giannakis, G.B. (1993), ‘Time-varying system identification and model validation using wavelets’, *IEEE Transactions on Signal Processing*, 41, 3512-3523.

van der Burgh, A.H.P., and Hartono, (2004), ‘Rain-wind-induced vibrations of a simple oscillator’, *International Journal of Nonlinear Mechanics*, 39, 93-100.

Wei, H.L., and Billings, S.A. (2002), ‘Identification of time-varying systems using multiresolution wavelet models’, *International Journal of Systems Science*, 33, 1217-1228.

Wei, H.L., and Billings, S.A. (2006), ‘An efficient nonlinear cardinal B-spline model for high tide forecasts at the Venice lagoon’, *Nonlinear Processes in Geophysics*, 13, 577-584.

Wei, H.L., Billings, S.A., and Liu, J. (2010), ‘Time-varying parametric modeling and time-dependent spectral characterization with applications to EEG signals using multiwavelets’, *International Journal of Modelling, Identification and Control*, 9, 215-224.

Young, P.C., (2011), *Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner*, Springer-Verlag, Berlin.

Yue, R., Billings, S.A., and Lang, Z.Q. (2005a), ‘An investigation into the characteristics of non-linear frequency response functions. Part 1: Understanding the higher dimensional frequency spaces’, *International Journal of Control*, 78, 1031-1044.

Yue, R., Billings, S.A., and Lang, Z.Q. (2005b), ‘An investigation into the characteristics of non-linear frequency response functions. Part 2: New analysis methods based on symbolic expansions and graphical techniques’, *International Journal of Control*, 78, 1130-1149.

Zheng, G.L., and Billings, S.A. (1999), ‘Qualitative validation and generalisation in nonlinear systems identification’, *International Journal of Control*, 72, 1592-1608.

Zheng, Y., Lin, Z., and Tay, D.B.H. (2001), ‘Time-varying parametric system multiresolution identification by wavelets’, *International Journal of Systems Science*, 32, 775-793.

Zhong, Y., Jan, K.M., Ju, K.H., and Chon, K.H. (2007), ‘Representation of time-varying nonlinear systems with time-varying principal dynamic modes’, *IEEE Transactions on Biomedical Engineering*, 54, 1983-1992.

Zou, R., Wang, H., and Chon, K.H. (2003), ‘A robust time-varying identification algorithm using basis functions’, *Annals of Biomedical Engineering*, 31, 840-853.

Zou, R., and Chon, K.H. (2004), ‘Robust algorithm for estimation of time-varying transfer functions’, *IEEE Transactions on Biomedical Engineering*, 51, 219-228.