COMMUN. MATH. SCI.
Vol. 13, No. 3, pp. 673–694

SPARSE TIME FREQUENCY REPRESENTATIONS AND DYNAMICAL SYSTEMS*

THOMAS Y. HOU†, ZUOQIANG SHI‡, AND PEYMAN TAVALLALI§

Dedicated to George Papanicolaou on the occasion of his 70th birthday

Abstract. In this paper, we establish a connection between the recently developed data-driven time-frequency analysis [T.Y. Hou and Z. Shi, Advances in Adaptive Data Analysis, 3, 1–28, 2011], [T.Y. Hou and Z. Shi, Applied and Comput. Harmonic Analysis, 35, 284–308, 2013] and the classical second order differential equations. The main idea of the data-driven time-frequency analysis is to decompose a multiscale signal into the sparsest collection of Intrinsic Mode Functions (IMFs) over the largest possible dictionary via nonlinear optimization. These IMFs are of the form $a(t)\cos(\theta(t))$, where the amplitude $a(t)$ is positive and slowly varying. The non-decreasing phase function $\theta(t)$ is determined by the data and in general depends on the signal in a nonlinear fashion. One of the main results of this paper is that we show that each IMF can be associated with a solution of a second order ordinary differential equation of the form $\ddot{x}+p(x,t)\dot{x}+q(x,t)=0$. Further, we propose a localized variational formulation for this problem and develop an effective l^1-based optimization method to recover $p(x,t)$ and $q(x,t)$ by looking for a sparse representation of p and q in terms of the polynomial basis. Depending on the form of nonlinearity in $p(x,t)$ and $q(x,t)$, we can define the order of nonlinearity for the associated IMF. This generalizes a concept recently introduced by Prof. N. E. Huang et al. [N.E. Huang, M.-T. Lo, Z. Wu, and Xianyao Chen, US Patent filling number 12/241.565, Sept. 2011]. Numerical examples will be provided to illustrate the robustness and stability of the proposed method for data with or without noise. This manuscript should be considered as a proof of concept.

Key words. Sparse time frequency representations; order of nonlinearity; intrinsic mode function; dynamical system.

AMS subject classifications. 37M10, 94A12.

1. Introduction

In many scientific applications such as biology, the underlying physical problem is so complex that we often do not know what is the appropriate governing equation to describe its dynamics. Typically, there are several dominating components that could contribute to the complex phenomena of the underlying physical solution. It is likely that each dominating component can be characterized by a dynamical system. Although we do not know the precise governing equation for these complex phenomena, we can collect a lot of data to characterize the solution of the underlying physical system. A very interesting question to ask is whether or not it is possible to obtain some qualitative understanding of different dominating components from the data that we collect. One of the most important questions is whether the underlying dynamical system is linear or nonlinear. If it is nonlinear, can we quantify the order of nonlinearity of the underlying dynamical system? In this paper, we attempt to provide one possible approach via a recently proposed data-driven time-frequency analysis method [10, 11].

The most commonly used definition of linearity is that the output of a system is linearly dependent on the input. But this definition is not very practical since we may not even know the governing system precisely. It is not easy to define what is input and what is output without knowing the governing system. Another difficulty is that the solution typically consists of several dominating components each of which

*Received: August 12, 2013; accepted (in revised form): February 27, 2014.

†Applied and Comput. Math, MC 9-94, Caltech, Pasadena, CA 91125, USA (hou@cms.caltech.edu).

‡Mathematical Sciences Center, Tsinghua University, Beijing, 100084, P.R. China (zqshi@math.tsinghua.edu.cn).

§Applied and Comput. Math, MC 9-94, Caltech, Pasadena, CA 91125, USA (ptavalla@caltech.edu).

accounts for a different physical mechanism. Some of these mechanisms may be linear and others may be nonlinear. Thus it is not a good idea to work on the entire data directly. We need to first decompose the data into several dominating components and then try to analyze these components separately. How to extract these intrinsic physical components from the data without compromising their hidden physical structure and integrity is highly nontrivial. For the data that we collect from a nonlinear system, such as the stokes wave, the classical Fourier or wavelet analysis would decompose the signal to a collection of fundamental components and harmonics. Each of the components, whether it is a fundamental or harmonic component, looks like a linear signal. A data analysis method based on these linear transformations would suggest that the signal is a superposition of linear components corresponding to a linear system rather than a nonlinear system.

The Empirical Mode Decomposition (EMD) method of Huang et al [12] provides a completely new way to analyze nonlinear and nonstationary signals. The EMD method decomposes a signal into a collection of intrinsic mode functions (IMFs) sequentially. The basic idea behind this approach is the removal of the local median from a signal by using a sifting process and a cubic spline interpolation of local extrema. The EMD method has found many applications, see e.g. [13, 30, 31]. One important property of these IMFs is that they give physically meaningful Hilbert spectral representation. On the other hand, since the EMD method relies on the information of local extrema of a signal, it is unstable to noise perturbation. Recently, an ensemble EMD method (EEMD) was proposed to make it more stable to noise perturbation [29]. Despite of the tremendous success of EMD and EEMD, there is still lack of a theoretical understanding of this method. We remark that the recently developed synchrosqueezed wavelet transform by Daubechies, Lu, and Wu [6] is another attempt to provide a mathematical justification for an EMD like method.

Inspired by EMD/EEMD and the recently developed compressed (compressive) sensing theory [2, 3, 4, 7, 9], Hou and Shi have recently introduced a data-driven time-frequency analysis method [10, 11]. There are two important ingredients of this method. The first one is that the basis that is used to decompose the data is derived from the data rather than determined *a priori*. This explains the name "data-driven" in our method. Finding such nonlinear multiscale basis is an essential ingredient of our method. In some sense, our problem is more difficult than the compressed (compressive) sensing problem in which the basis is assumed to be known *a priori*. The second ingredient is to look for the sparsest decomposition of the signal among the largest possible dictionary consisting of intrinsic mode functions. In our method, we reformulate the problem as a nonlinear optimization and find the basis and the decomposition simultaneously by looking for the sparsest decomposition among all the possible decompositions.

In this paper, we develop a method to quantify the nonlinearity of the IMFs given by the data-driven time-frequency analysis method. The main idea is to establish a connection between the IMFs and the classical second order differential equations. The data-driven time-frequency analysis decomposes a multiscale signal into a sparse collection of IMFs. These IMFs are of the form $a(t)\cos(\theta(t))$, where the amplitude $a(t)$ is positive and slowly varying. The non-decreasing phase function $\theta(t)$ is determined by the data and is in general nonlinear. One of the main results of this paper is that each IMF can be associated with a solution of a second order ordinary differential equation of the form $\ddot{x}+p(x,t)\dot{x}+q(x,t)=f(t)$. We further assume that the coefficients $p(x,t)$, $q(x,t)$ and $f(t)$ are slowly varying with respect to t. Thus, we can freeze these coefficients locally in time and absorb the forcing function into q. This leads to the reduced

autonomous second order ODE, i.e. $\ddot{x}+p(x)\dot{x}+q(x)=0$. Further, we can reformulate the second order ODE in a conservative form: $\ddot{x}+\dot{P}(x)+q(x)=0$, where $\frac{dP(x)}{dx}=p(x)$. We then have the following weak formulation of the equation by integrating by parts:

$$<x,\ddot{\phi}>-<P(x),\dot{\phi}>+<q(x),\phi>=0,$$

where $<\cdot,\cdot>$ is the standard inner product, and ϕ is a smooth test function of compact support. If $p(x)$ and $q(x)$ have a sparse representation in terms of the polynomial basis, then we can represent $P(x)$ and $q(x)$ as follows: $P(x)=\sum_{k=0}^{M}p_k x^{k+1}$, $q(x)=\sum_{k=0}^{M}q_k x^k$ for some integer $M>0$. Then we obtain the following weak formulation:

$$<x,\ddot{\phi}>-\sum_{k=0}^{M}p_k<x^{k+1},\dot{\phi}>+\sum_{k=0}^{M}q_k<x^k,\phi>=0.$$

Based on the above weak formulation, we can design an l^1-based optimization method to solve for p_k and q_k,

$$(p_k,q_k)=\arg\min_{\alpha_k,\beta_k}\gamma\sum_{k=0}^{M}(|\alpha_k|+|\beta_k|)+\sum_{i=1}^{N}\left|\left\langle x,\ddot{\phi}_i\right\rangle-\sum_{k=0}^{M}\alpha_k\left\langle x^{k+1},\dot{\phi}_i\right\rangle+\sum_{k=0}^{M}\beta_k\left\langle x^k,\phi_i\right\rangle\right|^2$$

where ϕ_i's are smooth test functions of compact support and N is the number of the test functions. We will provide some guidance on how to choose these test functions optimally.

The method described above provides a new way to interpret the hidden intrinsic information contained in the extracted IMF. Depending on the local form of nonlinearity in $p(x,t)$ and $q(x,t)$, we can define the order of nonlinearity for each associated IMF. Moreover, we also recover accurately the coefficients for the nonlinear terms in p and q. This generalizes a similar concept recently introduced by Prof. N. E. Huang et. al. [14]. Numerical examples will be provided to illustrate the robustness and stability of the proposed method.

The organization of the paper is as follows. In Section 2, we give a brief review of the data-driven time-frequency analysis. Section 3 is devoted to the connection between IMFs and second order ODEs. We will illustrate through some examples that solutions of many linear and nonlinear second order ODEs have solutions that are essentially IMFs. In Section 4, we introduce two numerical methods to extract the coefficients of the second order ODE from a given IMF. Based on the order of nonlinearity of the extracted coefficients, we introduce the order of nonlinearity for each IMF. This is called *nonlinear order analysis*. In Section 5, we demonstrate the effectiveness of the proposed method by a number of numerical examples. Some concluding remarks are made in Section 6.

2. A brief review of the data-drive time-frequency analysis

The data-driven time-frequency analysis method is based on finding the sparsest decomposition of a signal by solving a nonlinear optimization problem. First, we need to construct a large dictionary that can be used to obtain a sparse decomposition of the signal. In our method, the dictionary is chosen to be:

$$\mathcal{D}=\{a\cos\theta:a,\theta' \text{ is smoother than } \cos\theta,\forall t\in\mathbb{R},\theta'(t)\geq 0\}. \tag{2.1}$$

Let $V(\theta,\lambda)$ be the collection of all the functions that are smoother than $\cos\theta(t)$. In general, it is most effective to construct $V(\theta,\lambda)$ as an overcomplete Fourier basis given

below:

$$V(\theta,\lambda)=\text{span}\left\{1,\left(\cos\left(\frac{k\theta}{2L_\theta}\right)\right)_{1\le k\le 2\lambda L_\theta},\left(\sin\left(\frac{k\theta}{2L_\theta}\right)\right)_{1\le k\le 2\lambda L_\theta}\right\}, \tag{2.2}$$

where $L_\theta=\lfloor\frac{\theta(T)-\theta(0)}{2\pi}\rfloor$, $\lfloor\mu\rfloor$ is the largest integer less than μ, and $\lambda\le 1/2$ is a parameter to control the smoothness of $V(\theta,\lambda)$. The dictionary $\mathcal{D}$ then becomes:

$$\mathcal{D}=\{a\cos\theta:\ a\in V(\theta,\lambda),\ \theta'\in V(\theta,\lambda),\ \text{and}\ \theta'(t)\ge 0\ \forall t\in\mathbb{R}\}. \tag{2.3}$$

Each element of the dictionary $\mathcal{D}$ is an IMF with inter-wave frequency modulation. By an IMF with inter-wave frequency modulation, we mean that both the amplitude $a(t)$ and the instantaneous frequency $\theta'(t)$ are less oscillatory than $\cos\theta(t)$. In the case when the instantaneous frequency $\theta'(t)$ is as oscillatory as $\cos\theta(t)$ or more oscillatory than $\cos\theta(t)$, we say that this IMF has intra-wave modulation. The IMFs with intra-wave frequency modulation are not included in this dictionary. We will consider the IMFs with intra-wave frequency modulation in the next section. By saying that a function f is less oscillatory than another function g, we mean that f contains fewer high frequency modes than those of g or the high frequency mods of f decay much faster than those of g.

Since the dictionary $\mathcal{D}$ is highly redundant, the decomposition over this dictionary is not unique. We need a criterion to select the "best" one among all possible decompositions. We assume that the data we consider have an intrinsic sparse structure in the time-frequency plane in some nonlinear and nonstationary basis. However, we do not know this basis *a priori* and we need to derive (or learn) this basis from the data. Based on this consideration, we adopt sparsity as our criterion to choose the best decomposition. This criterion yields the following nonlinear optimization problem:

$$\begin{aligned} P_\delta:\ &\underset{(a_k)_{1\le k\le M},(\theta_k)_{1\le k\le M}}{\text{Minimize}} && M \\ &\text{Subject to:} && \begin{cases}\|f-\sum_{k=0}^{M}a_k\cos\theta_k\|_{l^2}\le\delta,\\ a_k\cos\theta_k\in\mathcal{D},\quad k=0,\cdots,M,\end{cases}\end{aligned} \tag{2.4}$$

where δ depends on the noise level of the signal.

REMARK 2.1. The above decomposition is based on the assumption that the signal has sparse representation over the dictionary. We believe that this assumption is reasonable in many cases. First of all, the dictionary we use is very large. Many of the oftenly used bases, such as Fourier basis and Wavelet basis, are included in our dictionary. This provides a higher probability that the signal is sparse over the dictionary. On the other hand, many studies also reveal that many interesting physical signals do have this property. We refer to [12, 13, 30] for more details.

The above optimization problem can be seen as a nonlinear l^0 minimization problem. Thanks to the recent developments of compressed sensing, two types of methods have been developed to study this problem. Since we have infinitely many elements in the basis (in fact uncountably many), we could not generalize basis pursuit directly to solve our problem. On the other hand, matching pursuit can be generalized. However, straightforward generalization of matching pursuit to our nonlinear optimization problem could be ill-conditioned and would introduce severe interference among different IMFs. In order to develop a stable nonlinear optimization method and remove the interference, we add an l^1 term to regularize the nonlinear least squares problem. This

gives rise to the following algorithm based on an l^1 regularized nonlinear least squares. We begin with $r_0 = f,\ k = 0$.

Step 1: Solve the following l^1 regularized nonlinear least-square problem (P_2):

$$P_2: \quad (a_k, \theta_k) \in \underset{a,\theta}{\text{Argmin}} \quad \gamma \|\widehat{a}\|_{l^1} + \|r_{k-1} - a\cos\theta\|_{l^2}^2 \\ \text{Subject to: } a \in V(\theta, \lambda), \quad \theta' \geq 0, \forall t \in \mathbb{R}, \tag{2.5}$$

where $\gamma > 0$ is a regularization parameter and $\widehat{a}$ is the representation of a in the overcomplete Fourier basis.

Step 2: Update the residual $r_k = f - \sum_{j=1}^{k} a_j \cos\theta_j$.

Step 3: If $\|r_k\|_{l^2} < \epsilon_0$, stop. Otherwise, set $k = k+1$ and go to Step 1.

If signals are periodic, we can use the standard Fourier basis to construct $V(\theta, \lambda)$ instead of the overcomplete Fourier basis. The l^1 regularization term is not needed (i.e. we can set $\gamma = 0$) since the standard Fourier basis elements are orthogonal to each other. For data with poor samples (i.e. the number of samples is not sufficient to resolve the signal) or for data with poor scale separation, we would still require the l^1 regularization even for periodic data.

One of the main difficulties in solving our l^1 regularized nonlinear least squares problem is that the objective functional is non-convex since the basis is not known *a priori*. We need to find the basis and the decomposition simultaneously. In [11], a Gauss-Newton type method was proposed to solve the l^1 regularized nonlinear least squares problem.

2.1. Numerical method for IMFs with intra-wave frequency modulation. The data-driven time-frequency analysis method described in the previous section is applicable to those signals whose IMFs have only inter-wave modulation but do not have intra-wave frequency modulation. As we will see in Section 3, the IMF with inter-wave frequency modulation is typically associated with a linear second order ODE, while the IMF with intra-wave frequency modulation is associated with a nonlinear second order ODE. In order to analyze the nature of nonlinearity in a signal, we must consider those IMFs with intra-wave frequency modulation. In this section, we describe a modified data-driven time-frequency analysis method that is capable of decomposing signals which contain IMFs with intra-wave frequency modulation.

For a signal that contains IMFs with intra-wave frequency modulation, they still have a sparse decomposition:

$$f(t) = \sum_{k=1}^{M} a_k \cos\theta_k, \tag{2.6}$$

where a_k are smooth amplitude functions. An important difference for data with intra-wave modulation is that their instantaneous frequencies, θ_k', are no longer in $V(\theta_k, \lambda)$. Typically, the phase function has the form $\theta_k = \phi_k + \epsilon\cos(\omega_k \phi_k)$, where ϕ_k is a smooth function, $\epsilon > 0$ is a small number, and ω_k is a positive integer.

An essential difficulty for this type of data is that the instantaneous frequency, θ_k', is as oscillatory as or even more oscillatory than $\cos\theta_k$. In the method proposed in the previous section, we assume that a_k and $\dot{\theta}_k$ are less oscillatory than $\cos\theta_k$. We use this property to construct the dictionary $\mathcal{D}$. In the case when an IMF has strong intra-wave modulation, θ_k' is as oscillatory as $\cos\theta_k$. Thus the method described in the previous section would not be able to give a good approximation of θ_k. To overcome

this difficulty, we introduce a shape function, s_k, to replace the cosine function. The idea is to absorb the high frequency intra-wave modulation into the shape function s_k. This will ensure that θ_k' is still less oscillatory than $s_k(\theta_k)$. This idea was proposed by Dr. H.-T. Wu in [28], but he did not provide an efficient algorithm to compute such shape function.

Note that s_k is not known *a priori* and is adapted to the signal. We need to learn s_k from the physical signal. This consideration naturally motivates us to modify the construction of the dictionary as follows:

$$\mathcal{M}=\{a_k s_k(\theta_k):\quad a_k,\theta_k'\in V(\theta_k,\lambda),\, s_k \text{ is } 2\pi\text{-period function}\}, \tag{2.7}$$

where $V(\theta,\lambda)$ is defined in (2.2) and s_k is an unknown 2π-periodic 'shape function' and is adapted to the signal. If we choose s_k to be the cosine function, then the new dictionary $\mathcal{M}$ is reduced to the dictionary $\mathcal{D}$ that we used previously, i.e. $\mathcal{M}=\mathcal{D}$.

We also use "sparsity" as the criterion to select the decomposition over the redundant dictionary $\mathcal{M}$. This would give us the following optimization problem:

$$\begin{aligned} &\underset{(s_k)_{1\le k\le M},(a_k)_{1\le k\le M},(\theta_k)_{1\le k\le M}}{\text{Minimize}} && M \\ &\text{Subject to:} && \begin{cases} \|f-\sum_{k=0}^{M} a_k\cdot s_k(\theta_k)\|_{l^2}\le\delta, \\ a_k\cdot s_k(\theta_k)\in\mathcal{M}, \quad k=0,\cdots,M, \end{cases} \end{aligned} \tag{2.8}$$

where δ depends on the noise level of the signal.

The above optimization problem is much more complicated than (2.4), since the shape function s_k is also unknown instead of being determined *a priori* as in (2.4). In order to simplify this problem, we further assume that the non-zero Fourier coefficients of s_k are confined to a finite number of low frequency modes, i.e. for each s_k, there exists $N_k\in\mathbb{N}$, such that

$$s_k(t)\in \operatorname{span}\left\{e^{ijt},\ j=-N_k,\cdots,N_k\right\}. \tag{2.9}$$

We further assume that we know how to obtain an estimate for N_k by some method. We call this the low-frequency confinement property of s_k. Based on this property of s_k, we can represent s_k by its Fourier series,

$$s_k(t)=\sum_{j=-N_k}^{N_k} c_{k,j}e^{ijt}. \tag{2.10}$$

For a given θ_k, we can use this representation and apply the singular value decomposition (SVD) to recover the Fourier coefficients $c_{k,j}$ of each s_k. This enables us to obtain the shape function s_k. Once we get an approximation of the shape function s_k, we can use s_k to update θ_k. This process continues until it converges. The detail of this method will appear in a subsequent paper. In this paper, we will focus on using this generalized data analysis method to perform nonlinearity analysis of multiscale data whose IMFs have intra-wave modulation.

3. IMFs and second order ODEs

One of the main objectives of this paper is to establish a connection between an IMF that we decompose from a multiscale signal and a second order ODE. Moreover, we propose an effective method to find such second order ODE and study the order of nonlinearity of the associated ODE. For a given IMF of the form $x(t)=a(t)\cos\theta(t)$, it

is not difficult to show that it satisfies the following second order ordinary differential equation:

$$\ddot{x}+\left(-\frac{\ddot{\theta}}{\dot{\theta}}-2\frac{\dot{a}}{a}\right)\dot{x}+\left(\dot{\theta}^2+\frac{\dot{a}\ddot{\theta}}{a\dot{\theta}}+2\left(\frac{\dot{a}}{a}\right)^2-\frac{\ddot{a}}{a}\right)x=0. \tag{3.1}$$

Let $p(t)=\left(-\frac{\ddot{\theta}}{\dot{\theta}}-2\frac{\dot{a}}{a}\right)$ and $q(t)=\left(\dot{\theta}^2+\frac{\dot{a}\ddot{\theta}}{a\dot{\theta}}+2\left(\frac{\dot{a}}{a}\right)^2-\frac{\ddot{a}}{a}\right)$, then we get a second order ODE

$$\ddot{x}+p(t)\dot{x}+q(t)x=0. \tag{3.2}$$

Note that $p(t)$ and $q(t)$ in general depend on $x(t)$. Thus the above ODE may be nonlinear in general. This formal connection does not give much information about the nature of the ODE. We will perform further analysis to reveal the nature of the associated ODE depending on the regularity of the amplitude, $a(t)$, and the instantaneous frequency, $\theta'(t)$, of a given IMF, $a(t)\cos(\theta(t))$.

3.1. Connection between linear second order ODEs and IMFs. Many second order linear differential equations with smooth coefficients have solutions that have the form of an IMF, i.e. $x=a(t)\cos\theta(t)$. Moreover, the corresponding amplitude $a(t)$ and the instantaneous frequency $\dot{\theta}(t)$ are smoother than $\cos\theta(t)$. To see this, we consider the following linear second order ODE:

$$\ddot{x}+b(t)\dot{x}+c(t)x=0. \tag{3.3}$$

It can be also rewritten in the following form:

$$\ddot{v}+Q(t)v=0, \tag{3.4}$$

where

$$v=e^{\frac{1}{2}\int_0^t b(\xi)d\xi}x, \quad Q(t)=c(t)-\frac{1}{4}b^2(t)-\frac{1}{2}\dot{b}(t). \tag{3.5}$$

Assume that $Q(t)>0$ and $Q(t)\gg 1$. Using the WKB method [1], we can get the asymptotic approximation of $v(t)$,

$$v(t)\sim c_1\cos\left(\int_0^t\sqrt{Q(\xi)}d\xi\right)+c_2\sin\left(\int_0^t\sqrt{Q(\xi)}d\xi\right). \tag{3.6}$$

In terms of the original variables, the solution of (3.3) has the form:

$$x(t)\sim e^{-\frac{1}{2}\int_0^t b(\xi)d\xi}\left(c_1\cos\left(\int_0^t\sqrt{Q(\xi)}d\xi\right)+c_2\sin\left(\int_0^t\sqrt{Q(\xi)}d\xi\right)\right), \tag{3.7}$$

which is essentially an IMF without intra-wave frequency modulation in which both the amplitude and the instantaneous frequency are smoother than $\cos\theta(t)$ due to the smoothness of b and Q.

On the other hand, for those IMFs $a(t)\cos\theta(t)$ that do not have intra-wave frequency modulation (meaning that both $a(t)$ and $\dot{\theta}(t)$ are smoother than $\cos(\theta(t))$), it is easy to see that the coefficients q and p given in (3.2) are smooth functions with respect to t. This seems to suggest that there is a close connection between oscillatory solutions of a linear second order ODE with smooth coefficients and IMFs without intra-wave frequency modulation.

3.2. IMFs with intra-wave frequency modulation and nonlinear ODEs. For IMFs with intra-wave frequency modulation, the situation is much more complicated. In this case, the coefficients $p(t)$ and $q(t)$ that appear in Equation (3.2) are no longer smooth since $\dot{\theta}(t)$ is not smooth. As we will demonstrate later, intra-wave frequency modulation is usually associated with a solution of a nonlinear ODE.

Consider a conservative system $\ddot{x}=F(x)$, where $F(x)=-\frac{dU(x)}{dx}$ for some smooth function U. The total energy of the system is $E=\frac{1}{2}\dot{x}^2+U(x)$. Assume $E>U(x)$ for all values of x in $D=[x_0,x_1]$, except the end points where $E=U(x_0)=U(x_1)$. It is obvious that $\dot{x}=0$ only at x_0 and x_1. Take $a=\frac{x_0+x_1}{2}$, $b=\frac{-x_0+x_1}{2}$. Consequently, the range of $\frac{x(t)-a}{b}$ lies within $[-1,1]$. We define

$$\begin{aligned}\theta(t) &= \arccos\left(\tfrac{x(t)-a}{b}\right), \qquad for \quad \dot{x}<0,\\ \theta(t) &= \arccos\left(\tfrac{-x(t)+a}{b}\right)+\pi, \ for \quad \dot{x}>0.\end{aligned}$$

As a result, we have

$$\begin{aligned}\dot{\theta}(t) &= \frac{-\dot{x}(t)}{b\sqrt{1-\left(\frac{x(t)-a}{b}\right)^2}}, \quad for \quad \dot{x}<0,\\ \dot{\theta}(t) &= \frac{\dot{x}(t)}{b\sqrt{1-\left(\frac{x(t)-a}{b}\right)^2}}, \quad for \quad \dot{x}>0,\end{aligned}$$

and $\dot{\theta}>0$ if $\frac{x(t)-a}{b}\neq\pm1$. Now, if $\frac{x(t)-a}{b}\to-1$, then $\dot{\theta}(t)\to\sqrt{\frac{\ddot{x}}{b}}$. Remember that $\ddot{x}>0$ as $\frac{x(t)-a}{b}\to-1$. Similarly, we can show that as $\frac{x(t)-a}{b}\to1$, we have $\ddot{x}<0$ and $\dot{\theta}(t)\to\sqrt{\frac{-\ddot{x}}{b}}$. Therefore, the solution of $\ddot{x}=F(x)$ can be represented as $x(t)=a+b\cos\theta(t)$, where a,b are constants and $\theta(t)\in C^1(t)$, $\dot{\theta}(t)>0$. The period T of the oscillation, using $\dot{\theta}$, can be defined as a real positive number such that $\int_0^T\dot{\theta}dt=2\pi$.

To illustrate this point further, we consider the solution of the Duffing equation. The undamped Duffing equation has the form $\ddot{x}+x+x^3=0$. The energy E of the system is $E=\frac{\dot{x}^2}{2}+\frac{x^2}{2}+\frac{x^4}{4}$. Obviously, the potential energy is $U(x)=\frac{x^2}{2}+\frac{x^4}{4}$. Now, assume that the solution varies within the interval $[-A,A]$. Due to symmetry, we look for a solution of the form $x(t)=A\cos\theta(t)$. Substituting this into the energy equation gives

$$A^2\cos^2\theta+A^2\dot{\theta}^2\sin^2\theta+\frac{1}{2}A^4\cos^4\theta=A^2+\frac{A^4}{2},$$

which can be further simplified as

$$\dot{\theta}^2=\frac{A^2}{2}\left(1+\cos^2\theta\right)+1.$$

The right hand side of this equation is strictly positive. Since $\dot{\theta}>0$, we obtain

$$\dot{\theta}=\sqrt{\frac{A^2}{2}\left(1+\cos^2\theta\right)+1}.$$

This shows that the solution of the Duffing equation is an IMF with intra-wave frequency modulation. We can see that the peaks and troughs of the signal coincide with the maximum of the instantaneous frequency $\dot{\theta}$.

4. Nonlinear order analysis

In this section, we propose a new method to analyze the order of nonlinearity of the IMFs that we decompose from a multiscale signal. We will present an effective optimization method to construct a second order ODE for each IMF. Moreover, based on the order of the nonlinearity of the coefficients associated with the second order ODE, we define the order of nonlinearity for each IMF.

To begin with, we consider the second order ODE of the following type:

$$\ddot{x}+p(x,t)\dot{x}+q(x,t)=f(t), \tag{4.1}$$

where $p(x,t)$, $q(x,t)$, and $f(t)$ are slowly varying with respect to t. For example, in case of the Duffing equation, we have $p(x)=0$, $q(x)=x+x^3$.

Based on this assumption, we can freeze $p(x,t)$, $q(x,t)$, and $f(t)$ locally in time over a local time interval (a few periods) since they vary slowly in time. Thus we can replace the above ODE by the corresponding autonomous ODE over this local time interval and absorb f into q (meaning that we can set $f=0$):

$$\ddot{x}+p(x)\dot{x}+q(x)=0. \tag{4.2}$$

This approximation reduces the level of difficulty significantly.

4.1. A strong formulation. In order to determine the autonomous ODE locally, we propose to use polynomials to approximate $p(x)$ and $q(x)$,

$$p(x)=\sum_{k=0}^{M}p_k x^k, \quad q(x)=\sum_{k=0}^{M}q_k x^k, \tag{4.3}$$

where M is the order of polynomials which is given *a priori*, and p_k, q_k are unknown coefficients.

One way to get the coefficients p_k, q_k is to substitute (4.3) to (4.2). This leads to

$$\ddot{x}+\sum_{k=0}^{M}p_k(x^k)\dot{x}+\sum_{k=0}^{M}q_k x^k=0. \tag{4.4}$$

Then p_k, q_k can be obtained by using a least squares method,

$$(p_k,q_k)=\arg\min_{\alpha_k,\beta_k}\|\ddot{x}+\sum_{k=0}^{M}\alpha_k(x^k)\dot{x}+\sum_{k=0}^{M}\beta_k x^k\|_2. \tag{4.5}$$

To study the order of nonlinearity, we are most interested in the highest order terms in p and q. Further, we assume that the coefficients p_k and q_k are sparse. Due to the strong correlation between x and $\dot{x}$, the direct least squares proposed in (4.5) would be unstable to noise perturbation. In order to stabilize this optimization algorithm, we add an l^1 term to regularize the least squares and look for the sparsest representation,

$$(p_k,q_k)=\arg\min_{\alpha_k,\beta_k}\gamma\sum_{k=0}^{M}(|\alpha_k|+|\beta_k|)+\|\ddot{x}+\sum_{k=0}^{M}\alpha_k(x^k)\dot{x}+\sum_{k=0}^{M}\beta_k x^k\|_2^2, \tag{4.6}$$

where γ is a parameter to control the sparsity of the coefficients. In order to capture the leading order term, γ is chosen to be $O(1)$. In the following examples, γ is chosen to be 2.

In the method described above, we need to compute $\ddot{x}$ and $\dot{x}$. This tends to amplify the error introduced in our approximation of the IMF, x. Next, we will introduce another method based on the weak formulation of the second order ODE.

4.2. A weak formulation. In this section, we will introduce an l^1-based optimization based on a weak formulation for the second order ODE. Let $P(x)$ be the primitive function of $p(x)$, i.e. $\dot{P}(x)=p(x)\dot{x}$. Then the ODE can be rewritten in a conservation form:

$$\ddot{x}+\dot{P}(x)+q(x)=0. \tag{4.7}$$

Suppose the span of time of the signal that we want to study is $[0,T]$. For any test function $\phi\in C_0^2[0,T]$ satisfying $\dot{\phi}(0)=\dot{\phi}(T)=0$, we have the following weak formulation of the equation by performing integration by parts:

$$<x,\ddot{\phi}>-<P(x),\dot{\phi}>+<q(x),\phi>=0, \tag{4.8}$$

where $<\cdot,\cdot>$ is the standard inner product.

If $p(x),q(x)$ can be approximated by polynomials as what we have done in (4.3), then $P(x)$ and $q(x)$ can be expanded in terms of the polynomial basis:

$$P(x)=\sum_{k=0}^{M}p_k x^{k+1},\quad q(x)=\sum_{k=0}^{M}q_k x^k. \tag{4.9}$$

Then we get

$$<x,\ddot{\phi}>-\sum_{k=0}^{M}p_k<x^{k+1},\dot{\phi}>+\sum_{k=0}^{M}q_k<x^k,\phi>=0. \tag{4.10}$$

Using this formulation, we can design the following optimization problem to solve for p_k and q_k,

$$(p_k,q_k)=\arg\min_{\alpha_k,\beta_k}\gamma\sum_{k=0}^{M}(|\alpha_k|+|\beta_k|)+\sum_{i=1}^{N}\left|\left\langle x,\ddot{\phi}_i\right\rangle-\sum_{k=0}^{M}\alpha_k\left\langle x^{k+1},\dot{\phi}_i\right\rangle+\sum_{k=0}^{M}\beta_k\left\langle x^k,\phi_i\right\rangle\right|^2 \tag{4.11}$$

where N is the number of the test functions that we use. In our computations, we choose $N=2M$ to make sure that we have enough measurements to determine the coefficients. The test functions that we use are given below:

$$\phi_i(t)=\begin{cases}\frac{1}{2}(1+\cos(\pi(t-t_i)/\lambda)), & -\lambda<t-t_i<\lambda,\\ 0, & \text{otherwise},\end{cases}\qquad i=1,\cdots,N, \tag{4.12}$$

where t_i's $(i=1,\cdots,N)$ are the centers of the test functions and the parameter λ determines their support. In order to enhance stability, we should make the support of the test functions as large as possible by choosing a large λ. On the other hand, if the support of ϕ is too large, we cannot get the high frequency information of the signal, which is essential in capturing the nonlinearity of the signal. Thus, we should determine λ based on the balance between stability and resolution. The strategy that we use is that to make λ as large as possible without compromising the resolution. In our computations, λ is chosen to be 1/5 of the local period (or wavelength) of the signal. After λ is determined, we choose t_i, $i=1,\cdots,N$ to be uniformly distributed over $[\lambda,T-\lambda]$, where $[0,T]$ is the time span of the signal.

REMARK 4.1. The choice of λ depends on the regularity of the signal that we want to study. If the signal is nearly singular, we should choose a small λ to make sure that the information of the signal can be well captured by the test functions.

REMARK 4.2. If the test functions $\phi_i(t)$ are chosen to be the classical piecewise linear finite element basis, then the weak formulation is equivalent to the strong formulation if we approximate $\ddot{x}$ and $\dot{x}$ by a second order central difference approximation.

Based on the coefficients that we recover from the signal, we can define two indices associated with each IMF to characterize the nonlinearity of this IMF.

DEFINITION 4.1. (Order of Nonlinearity.) *The order of nonlinearity of an IMF is defined to be the following two indices*

$$I_1 = \max\{k : p_k \neq 0, k = 0, \cdots, M\}, \quad I_2 = \max\{k : q_k \neq 0, k = 0, \cdots, M\}. \tag{4.13}$$

From the above definition, we can see that the order of nonlinearity of the signal corresponds to the highest order of the nonlinear terms. The case of $I_1 = 0$ and $I_2 = 1$ corresponds to a linear ODE. When $I_1 > 0$ or $I_2 > 1$, the IMF becomes nonlinear. The larger the index is, the more nonlinear the IMF becomes. We not only quantify the order of nonlinearity of the IMF, we can also recover the coefficients associated with the leading order nonlinear terms. This information is very helpful in quantifying how nonlinear an IMF is and may have an important implication in engineering and biomedical applications.

REMARK 4.3. In the above definition, we have only used the highest order. Other information arising from the ODE, such as the amplitude of the coefficients, will also play roles in the nonlinearity of the signal. It is part of our future research to develop more comprehensive definition of the nonlinearity of the signal by using all the information carried out by the ODE itself.

In practical computations, the signal may be polluted by noise or measurement errors. As a result, our recovery of the coefficients will be influenced by these errors. To alleviate this side effect, we set up a small threshold ν_0 to enforce sparsity of the coefficients by keeping only those coefficients that are larger than ν_0. This leads to the following modified definition of the order of nonlinearity:

$$I_1 = \max\{k : |p_k| > \nu_0, k = 0, \cdots, M\}, \quad I_2 = \max\{k : |q_k| > \nu_0, k = 0, \cdots, M\}. \tag{4.14}$$

In the computations to be presented in the next section, we set $\nu_0 = 0.05$.

The method proposed in this paper is still not very robust to large noise perturbations. The reason is twofold. First, in our method, the second order derivative of an IMF is involved, which would amplify the perturbation of the IMF polluted by noise. Second, the definition of the order of nonlinearity is sensitive to perturbation. In short, the noise perturbation changes the dynamics of the system. Hence, the presence of so much noise would completely tarnish the dynamics of the signal. As we mentioned before, in the future work, we will study how to combine all the information of the coefficients and orders of the ODE to improve the definition of the order of nonlinearity.

The method based on the l^1 regularized least squares performs very well in identifying those nonlinear terms with large coefficients. On the other hand, the l^1 regularization also compromises the accuracy of the coefficients at the expense of producing a sparse representation of the signal. In order to recover the coefficients accurately, we propose the following procedure to improve the accuracy.

First, we identify the dominant coefficients,

$$\Gamma_1 = \{k : |p_k| > \nu_1, k = 0, \cdots, M\}, \quad \Gamma_2 = \{k : |q_k| > \nu_1, k = 0, \cdots, M\}. \tag{4.15}$$

In our computations, ν_1 is chosen to be 0.05.

Secondly, we solve a least squares problem without l^1 regularization to obtain more accurate coefficients for these dominant terms, for $k_2 \in \Gamma_2,\ k_1 \in \Gamma_1$,

$$(p_{k_2}, q_{k_1}) = \arg \min_{\alpha_{k_2}, \beta_{k_1}} \sum_{i=1}^{N} \left| \left\langle x, \ddot{\phi}_i \right\rangle - \sum_{k_2 \in \Gamma_2} \alpha_{k_2} \left\langle x^{k_2+1}, \dot{\phi}_i \right\rangle + \sum_{k_1 \in \Gamma_1} \beta_{k_1} \left\langle x^{k_1}, \phi_i \right\rangle \right|^2. \quad (4.16)$$

REMARK 4.4. If the signal is free of noise and accurate, the above refinement procedure does help to get more accurate coefficients. But when the signal is polluted with noise, the IMF that we extract from the signal is not very accurate. In this case, the error of the coefficients is still relatively large even with the above refinement procedure.

Before we end this section, we summarize the discussion by giving the following algorithm. We first partition the entire physical domain into a number of subdomains and localize the signal locally by multiplying a smooth cut-off function. Then we apply the above optimization algorithm to the localized signal to extract the local order of nonlinearity of the signal.

An l^1-l^2 Refinement Algorithm.

- Calculate the phase function $\theta(t)$ of the signal. Choose K points $t_j,\ j = 1, \cdots, K$ such that the time variation of P and q is well resolved by the local resolution $(t_{j+1} - t_j)$.
- For $j = 1 : K$
- Extract the signal around the point t_j,

$$f_j(t) = f(t)\chi(\theta(t) - \theta(t_j)),$$

where $\chi(t)$ is a cutoff function. In our computations, it is chosen to be

$$\chi(t) = \begin{cases} \frac{1}{2}(1 + \cos(t/\mu)), & -\mu\pi < t < \mu\pi, \\ 0, & \text{otherwise.} \end{cases}$$

μ is a parameter to control the width of the cutoff function. In this paper, we choose $\mu = 3$, which means that for each point, we localize the signal within 3 periods to perform the order of nonlinearity analysis.

- Extract the IMF c_j for $f_j(t)$ using the algorithm in Section 2.
- Solve the optimization problem (4.11) with $x = c_j$ to get the coefficients of the polynomials, $P_j(x)$ and $q_j(x)$.
- (optional) Apply the refinement procedure to update the coefficients.
- End
- Calculate the order of nonlinearity of the signal according to (4.14).

5. Numerical results

In this section, we will show several numerical results to demonstrate the performance of our nonlinearity analysis method proposed previously. We first apply our method to study the order of nonlinearity from the signal generated from the solution of the Van der Pol equation.

Example 1: Consider the Van der Pol Equation

$$\ddot{x} + \left(x^2 - 1\right)\dot{x} + x = 0.$$

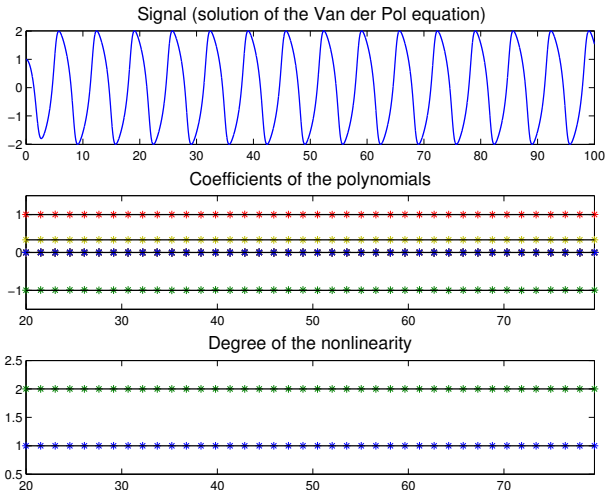

FIG. 5.1. *Top: The solution of the Van der Pol equation; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

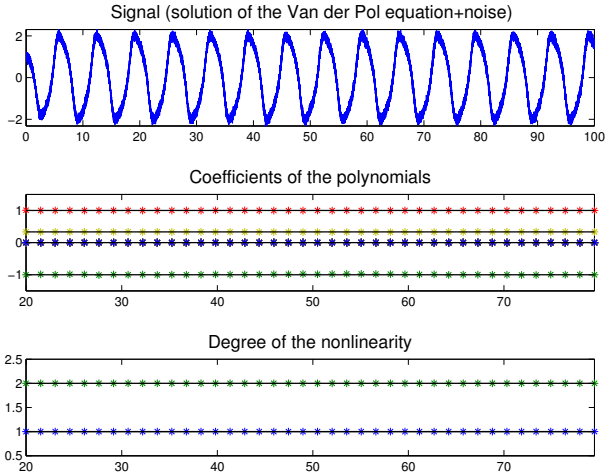

FIG. 5.2. *Top: The solution of the Van der Pol equation with noise* $0.1X(t)$*, where* $X(t)$ *is the white noise with standard deviation* $\sigma^2 = 1$*; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

The equation is solved from $t = 0$ to $t = 100$ with the initial condition $x(0) = 1$, $\dot{x}(0) = 0$. Figure 5.1 shows the original signal and the extracted coefficients and nonlinearity at different times. We choose $M = 10$ in our computations. With this choice, there are 22 coefficients in total, and only three of them are not zero. They correspond to $p_1 = -1$, $p_2 = 1/3$, and $q_1 = 1$ respectively. As shown in figure 5.1, we can get almost exact recovery of all the coefficients. When the signal is polluted by noise, our method

can still give reasonably accurate results, see figure 5.2.

Example 2: The second example is the Duffing equation

$$\ddot{x}+x+x^3=0,$$

with initial conditions $x(0)=1$ and $\dot{x}(0)=0$. The solution is also solved from $t=0$ to $t=100$. Figure 5.3 shows the original signal and recovery of the coefficients and order of nonlinearity. Again, we use $M=10$ in our computations. In this case, there are actually two coefficients that are not zero, $q_1=1,\ q_3=1$.

When the signal does not have noise, the recovery is very good for both of the coefficients and the order of nonlinearity, see figure 5.3. But when the signal is polluted by noise, the results for the Duffing equation are not as good as those for the Van der Pol equation, see figure 5.3. The reason is that the solution of the Duffing equation is closer to the linear sinusoidal wave with $q_3=0$. A small perturbation would introduce a large perturbation to the coefficients. Nevertheless, even in this case, our method can still give the correct order of nonlinearity, see figure 5.4.

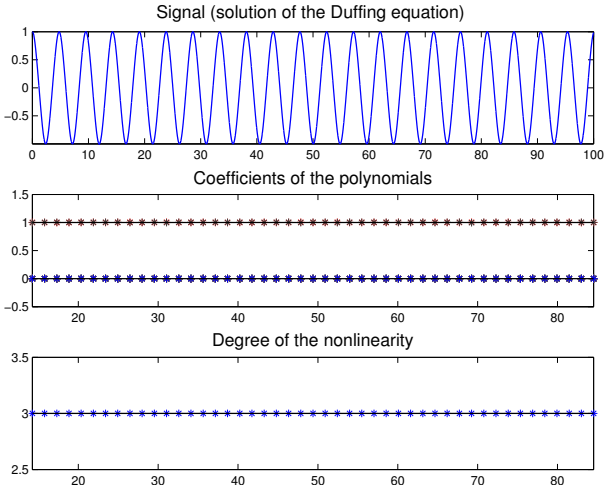

FIG. 5.3. *Top: The solution of the Duffing equation; Middle: Coefficients* (q_k,p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

Example 3: The equations in the previous two examples are both autonomous. For this kind of equations, the coefficients can be extracted globally, since it does not change over the whole time span. In order to demonstrate the locality of our method, we consider an equation which is not autonomous:

$$\ddot{x}+a(t)(x^2-1)\dot{x}+(1-a(t))x^3+x=0, \tag{5.1}$$

where $a(t)=\frac{1}{2}\left(1-\frac{t-100}{\sqrt{(t-100)^2+400}}\right)$. The initial condition is that $\dot{x}(0)=0,\ x(0)=1$, and the equation is solved over $t\in[0,200]$.

As we can see, this equation is essentially of the Van der Pol type when t is small $(t<100)$. As t increases, the equation changes to the Duffing type equation gradually. This equation has to be analyzed locally. A global approach would predict the wrong

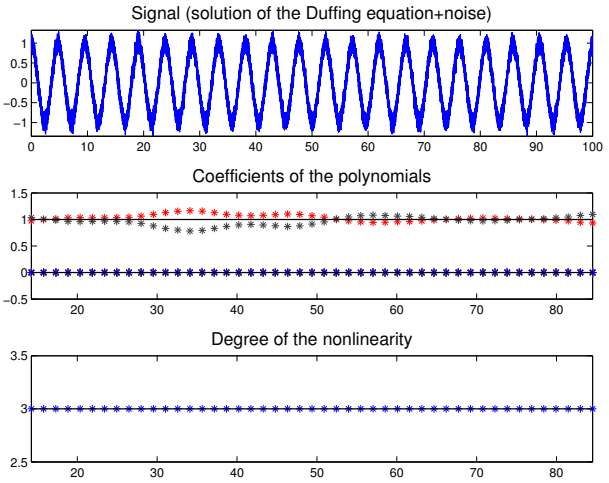

FIG. 5.4. *Top: The solution of the Duffing equation with noise* $0.1X(t)$*, where* $X(t)$ *is the white noise with standard deviation* $\sigma^2 = 1$*; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

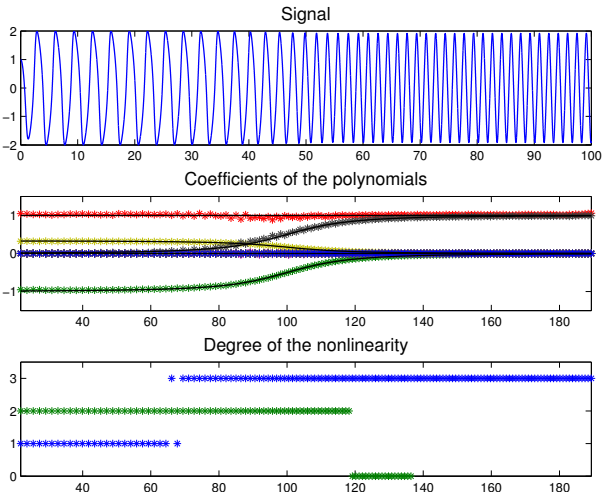

FIG. 5.5. *Top: The solution of the equation given in* (5.1)*; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

order of nonlinearity. We first present our results in figure 5.5 when the solution is free of noise. Our method can capture the time variation of the coefficients very accurately. Even if the solution is polluted with noise, the results are still with reasonable accuracy, figure 5.6. The error of the coefficients is relatively large when $t > 100$. The reason is that in this region, the equation is qualitatively of Duffing type, and the Duffing equation is more sensitive to noise than the Van der Pol equation, as we pointed out in the previous example.

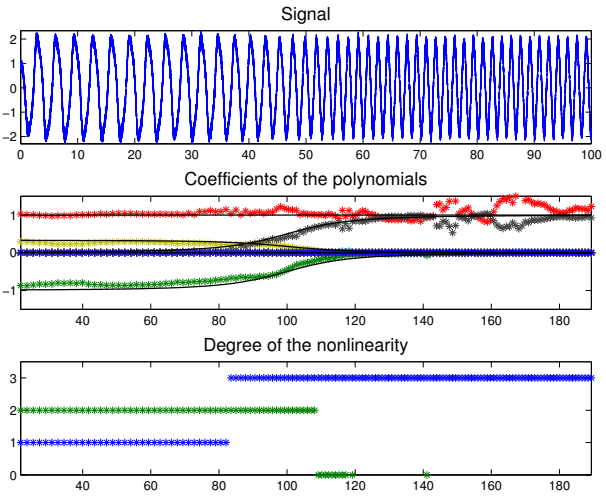

FIG. 5.6. *Top: The solution of the equation given in* (5.1) *with noise* $0.1X(t)$*, where* $X(t)$ *is the white noise with standard deviation* $\sigma^2 = 1$*; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

Example 4: In this example, we consider a more challenging equation, the coefficients have a sharp change instead of a smooth transition as in Example 3. The equation is given as follows:

$$\ddot{x} + \frac{1}{2}(1 - \mathrm{sgn}(t-100))\dot{x} + \frac{1}{2}(1 + \mathrm{sgn}(t-100))x^3 + x = 0, \tag{5.2}$$

where $\mathrm{sgn}(\cdot)$ is the sign function. This equation has a sharp transition from the Van der Pol equation to the Duffing equation at point $t = 100$.

When applying our method to analyze the solution of this equation, it is not hard to imagine that there would be some problem near the transition point, since we require that the coefficients be constants over a few periods of the signal. This assumption is not satisfied near the transition point.

We present the results in figure 5.7. It is not surprising that the error near $t = 100$ is very large, but in the region away from the transition point, our method still gives a reasonably accurate recovery. Due to the poor accuracy near the transition point, our method cannot locate the transition point accurately. But the good news is that our method does tell us that the nonlinearity of the signal changes from the Van der Pol type to the Duffing type, although it cannot give the precise location of the transition point. When the signal is polluted by noise, the performance of our method is qualitatively the same, see figure 5.8.

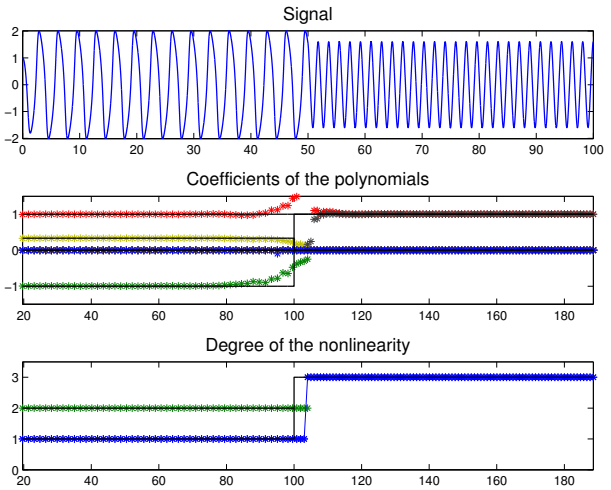

FIG. 5.7. *Top: The solution of the equation given in* (5.2)*; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

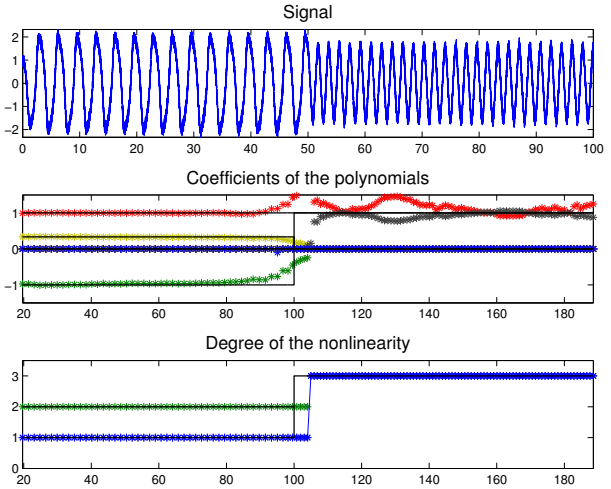

FIG. 5.8. *Top: The solution of the equation given in* (5.2) *with noise* $0.1X(t)$*, where* $X(t)$ *is the white noise with standard deviation* $\sigma^2 = 1$*; Middle: Coefficients* (q_k, p_k) *recovered by our method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

In order to improve the accuracy in the region near the transition point, we combine the idea of the ENO method in computing shock waves in fluid dynamics [16] with the method that we proposed earlier. This gives rise to the following algorithm.

- Calculate the phase function $\theta(t)$ of the signal. Choose K points t_j, $j = 1, \cdots, K$ such that the time variation of P and q is well resolved by the local resolution $(t_{j+1} - t_j)$.
- For $j = 1 : K$

S1: Extract the signal centered around the point t_j and also extract the signal to the left and to the right of t_j,

$$\begin{aligned} f_j^c(t) &= f(t)\chi^c(\theta(t)-\theta(t_j)), \\ f_j^l(t) &= f(t)\chi^l(\theta(t)-\theta(t_j)), \\ f_j^r(t) &= f(t)\chi^r(\theta(t)-\theta(t_j)), \end{aligned}$$

where $\chi^c(t), \chi^l(t), \chi^r(t)$ are cutoff functions

$$\chi^c(t) = \begin{cases} \frac{1}{2}(1+\cos(t/\mu)), & -\mu\pi < t < \mu\pi, \\ 0, & \text{otherwise}, \end{cases}$$

$$\chi^l(t) = \begin{cases} \frac{1}{2}(1+\cos(t/\mu+\pi)), & -2\mu\pi < t < 0, \\ 0, & \text{otherwise}, \end{cases}$$

$$\chi^r(t) = \begin{cases} \frac{1}{2}(1+\cos(t/\mu-\pi)), & 0 < t < 2\mu\pi, \\ 0, & \text{otherwise}. \end{cases}$$

As before, we choose $\mu = 3$.

S2: Extract the IMFs c_j^c, c_j^l, c_j^r for $f_j^c(t), f_j^l(t), f_j^r(t)$ respectively.

S3: Pick up the IMF c_j^* such that the residual $\|c_j^\alpha - f_j^\alpha\|_2$ is minimized over the choices $\alpha = c, l, r$, i.e.

$$c_j^* = \arg\min_{\alpha\in\{c,l,r\}} \|c_j^\alpha - f_j^\alpha\|_2.$$

S4: Solve the optimization problem (4.11) with $x = c_j^*$ to get the coefficients of the polynomials, $P_j(x)$ and $q_j(x)$.

S5: (optional) Apply the refinement procedure to update the coefficients.

- End
- Calculate the order of nonlinearity of the signal according to (4.14).

Figure 5.9 gives the performance of the above modified algorithm. The result is much better than the one obtained earlier. The coefficients are now accurate over the whole time span of the signal. The location of the transition point is also captured accurately. Even when the signal is polluted with noise, this method is still capable of approximating the order of nonlinearity and the transition point accurately as shown in figure 5.10.

Example 5: The signal $f(t)$ we consider in this last example consists of several components,

$$f(t) = s(t) + \cos(16\pi t/200) + t/100 + 0.1X(t), \quad t \in [0, 200], \tag{5.3}$$

where $s(t)$ is the solution of the Van der Pol equation with the initial condition $\dot{x}(0) = 0$, $x(0) = 2$ and $X(t)$ is the white noise with standard deviation $\sigma^2 = 1$.

For this kind of signal, we have to decompose it to several IMFs first and apply the nonlinearity analysis to each IMF to obtain their order of nonlinearity. Figure 5.11 gives the signal and two IMFs that we decompose from the signal. In figure 5.12 and figure 5.13, we present the results of the nonlinearity analysis for each IMF. As we can see that for this signal, the performance of our method is still reasonably good.

These examples show that our data-driven time-frequency analysis can be used to detect the type of nonlinearity (or at least its leading order of nonlinearity). A future goal is to combine this method with statistical study to make the nonlinear system identification algorithm more accurate and more stable.

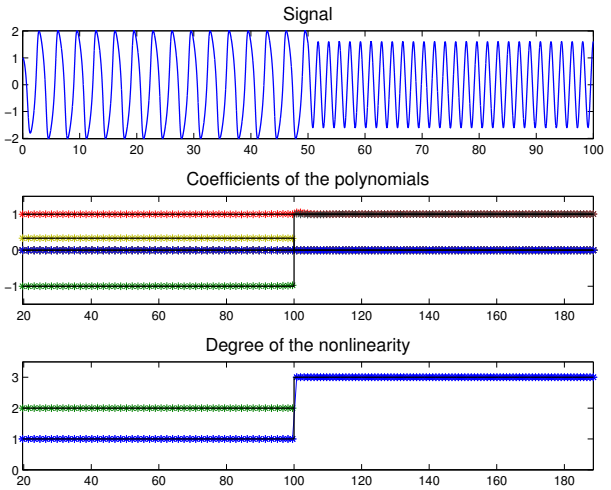

FIG. 5.9. *Top: The solution of the equation given in* (5.2)*; Middle: Coefficients* (q_k, p_k) *recovered by our method together with the trick in ENO method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

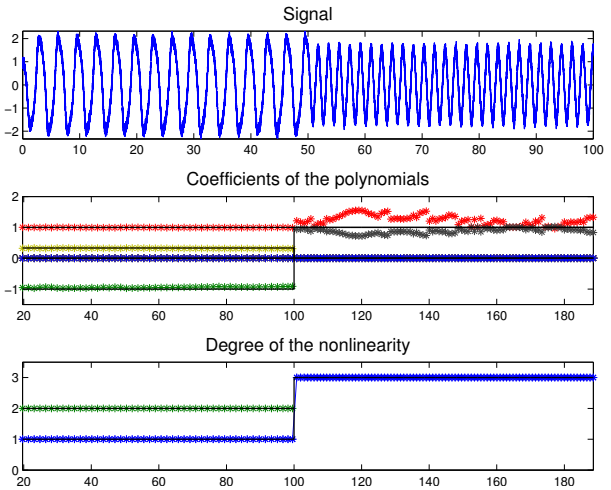

FIG. 5.10. *Top: The solution of the equation given in* (5.2) *with noise* $0.1X(t)$*, where* $X(t)$ *is the white noise with standard deviation* $\sigma^2 = 1$*; Middle: Coefficients* (q_k, p_k) *recovered by our method together with the ENO type method, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

6. Concluding remarks

In this paper, we have shown that many of the IMFs can be analyzed from the point of view of dynamical systems. This explains to some extent why adaptive methods such EMD or our data-driven time-frequency analysis method provide a natural way to analyze such signals. By establishing a connection between each IMF and a second order ODE, we can use the information of the associated second order ODE to obtain further information about the IMF that we extract, including the order of nonlinearity

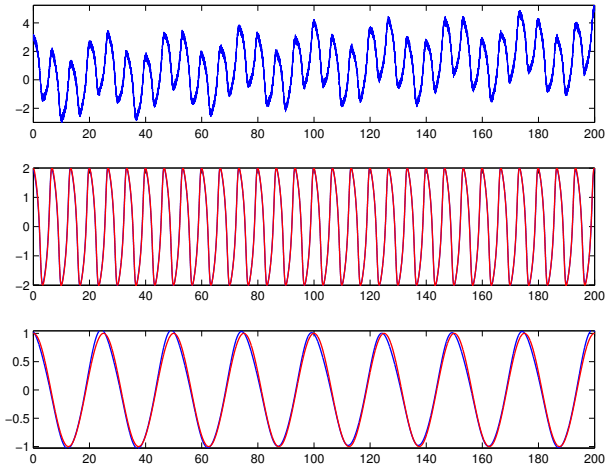

FIG. 5.11. *Top: The signal consists of the solution of the Van der Pol equation and a cosine function and a linear trend and noise* $0.1X$*; Middle: The IMF extracted from the signal corresponding to the solution of the Van der Pol equation, blue: numerical result; red: exact solution; Bottom: The IMF extracted from the signal corresponding to the cosine function, blue: numerical result; red: exact solution.*

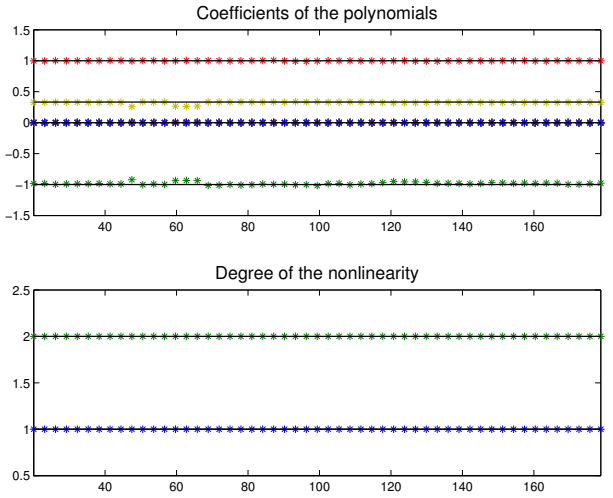

FIG. 5.12. *Top: Coefficients* (q_k, p_k) *recovered by our method for the first IMF in figure 5.11, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

and their energy levels. This information can be also used to provide a quantitative and qualitative description of the extracted IMFs of a multiscale signal. This may prove to be very useful in a number of engineering or biomedical applications. A possible future direction is to use statistical methods to do system identification and detect whether the system is linear or nonlinear.

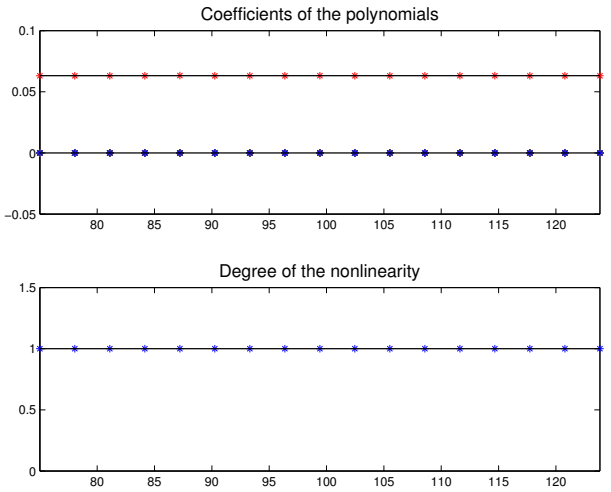

FIG. 5.13. *Top: Coefficients* (q_k, p_k) *recovered by our method for the second IMF in figure 5.11, star points* $*$ *represent the numerical results, black line is the exact one; Bottom: Nonlinearity of the signal according to the recovered coefficients, star points* $*$ *represent the numerical results, black line is the exact one.*

Acknowledgments. We would like to thank Professor Norden E. Huang for a number of stimulating discussions on the topic of the order of nonlinearity. This work was supported by NSF FRG Grant DMS-1159138, an AFOSR MURI Grant FA9550-09-1-0613 and a DOE grant DE-FG02-06ER25727. The research of Dr. T.Y. Hou was partially supported by NSF Grant DMS-1318377. The research of Dr. Z. Shi was also in part supported by a NSFC Grant 11201257.

REFERENCES

[1] C.M. Bender and S.A. Orszag, *Advanced mathematical methods for scientists and engineers*, Springer, 54–68, 1999.
[2] A.M. Bruckstein, D.L. Donoho, and M. Elad, *From sparse solutions of systems of equations to sparse modeling of signals and images*, SIAM Rev., 51, 34–81, 2009.
[3] E.J. Candès, J. Romberg, and T. Tao, *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52, 489–509, 2006.
[4] E.J. Candès and T. Tao, *Near optimal signal recovery from random projections: universal encoding strategies?* IEEE Trans. on Information Theory, 52(12), 5406–5425, 2006.
[5] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series on Applied Mathematics, SIAM Publications, 61, 1992.
[6] I. Daubechies, J. Lu, and H. Wu, *Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool*, Appl. Comp. Harmonic Anal., 30, 243–261, 2011.
[7] D.L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, 52, 1289–1306, 2006.
[8] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.
[9] R. Gribonval and M. Nielsen, *Sparse representations in unions of bases*, IEEE Trans. Inform. Theory, 49(12), 3320–3325, 2003.
[10] T.Y. Hou and Z. Shi, *Adaptive Data Analysis via Sparse Time-Frequency Representation*, Adv. Adapt. Anal., 3, 1–28, 2011.
[11] T.Y. Hou and Z. Shi, *Data-driven Time-Frequency Analysis*, Appl. Comput. Harmon. Anal., 35, 284–308, 2013.
[12] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proc. R. Soc. Lond. A, 454, 903–995, 1998.

[13] N.E. Huang and Z. Wu, *A review on Hilbert-Huang Transform: the method and its applications on geophysical studies*, Rev. Geophys., 46, 2008. RG2006, doi:10.1029/2007RG000228.
[14] N.E. Huang, M.T. Lo, Z. Wu, and Xianyao Chen, *Method for quantifying and modeling degree of nonlinearity, combined nonlinearity and nonstantionarity*, US Patent filling number 12/241.565, Sept. 2011.
[15] D. L. Jones and T. W. Parks, *A High Resolution Data-Adaptive Time-Frequency Representation*, IEEE Trans. Acoust. Speech Signal Process, 38, 2127–2135, 1990.
[16] R.J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhauser-Verlag Publ., 1992.
[17] P.J. Loughlin and B. Tracer, *On the amplitude- and frequency-modulation decomposition of signals*, J. Acoust. Soc. Am., 100, 1594–1601, 1996.
[18] B.C. Lovell, R.C. Williamson, and B. Boashash, *The relationship between instantaneous frequency and time-frequency representations*, IEEE Trans. Signal Process, 41, 1458–1461, 1993.
[19] S. Mallat and Z. Zhang, *Matching pursuit with time-frequency dictionaries*, IEEE Trans. Signal Process, 41, 3397–3415, 1993.
[20] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, Academic Press, 2009.
[21] W.K. Meville, *Wave modulation and breakdown*, J. Fluid Mech., 128, 489–506, 1983.
[22] S. Olhede and A.T. Walden, *The Hilbert spectrum via wavelet projections*, Proc. Roy. Soc. London A, 460, 955–975, 2004.
[23] B. Picinbono, *On instantaneous amplitude and phase signals*, IEEE Trans. Signal Process, 45, 552–560, 1997.
[24] S. Qian and D. Chen, *Joint Time-Frequency Analysis: Methods and Applications*, Prentice Hall, 1996.
[25] S.O. Rice, *Mathematical analysis of random noise*, Bell Syst. Tech. J., 23, 282–332, 1944.
[26] J. Shekel, *Instantaneous frequency*, Proc. IRE, 41, 548–548, 1953.
[27] B. Van der Pol, *The fundamental principles of frequency modulation*, Proc. IEE, 93, 153–158, 1946.
[28] H. Wu, *Instantaneous frequency and wave shape functions (I)*, Appl. Comput. Harmon. Anal. 35, 181–199, 2013.
[29] Z. Wu and N.E. Huang, *Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method*, Advances in Adaptive Data Analysis, 1, 1–41, 2009.
[30] Z. Wu, N.E. Huang, S.R. Long, and C.K. Peng, *On the trend, detrending, and variability of nonlinear and nonstationary time series*, PNAS. 104(38), 14889–14894, 2007.
[31] Z. Wu, N. E Huang, and X. Chen, *The multi-dimensional Ensemble Empirical Mode Decomposition method*, Adv. Adapt. Data Anal., 1(3), 339–372, 2009.