# Data-Driven Time-Frequency Analysis of Multiscale Data via Nonlinear Optimization

**Thomas Y. Hou**

**Applied and Comput. Mathematics,**
**Caltech**

**Joint work with Dr. Zuoqiang Shi and Peyman Tavallali**

**IPAM Workshop on Adaptive Data Analysis and Sparsity, Jan. 29, 2013**

# Sparse representation of multiscale data

- In many real world applications, we often deal with signals with ever-changing frequency, chirp signals, speech, etc.
- It is advantageous to use an adaptive sparse representation of these signals that preserve the intrinsic physical properties.
- This calls for the need to develop local and adaptive data analysis methods for nonlinear and nonstationary data.
- Currently, most data analysis methods use pre-determined bases (e.g. Fourier, wavelets, or polynomial bases).
- Inspired by the EMD method and compressed sensing, we introduce a new data-driven time-frequency analysis method based on nonlinear optimization.

# Instantaneous Frequency

The Analytic Signal method by Van der Pol (1946) and Gabor (1946) was one of the commonly used methods in defining instantaneous frequency.

- Given a signal $x(t)$, they define $y(t) = H(x)(t)$ ($H$ is the Hilbert transform) and an analytic signal $z(t)$

  $$z(t) = x(t) + iy(t) = a(t)e^{i\phi(t)},$$

  where $a(t) = \sqrt{x^2(t) + y^2(t)}$ and $\phi(t) = \tan^{-1} \frac{y(t)}{x(t)}$.
- Instantaneous Frequency is defined as $\omega(t) = \frac{d}{dt}\phi(t)$.
- This definition could be problemic if we do not remove the local median from the signal. Let $x_1(t) = c_0 + x(t)$. Then we have

  $$\phi_1(t) = \tan^{-1} \frac{H(c_0 + x(t))}{c_0 + x(t)} = \tan^{-1} \frac{H(x(t))}{c_0 + x(t)} \neq \tan^{-1} \frac{H(x(t))}{x(t)} = \phi(t).$$

# Empirical Mode Decomposition–continued

- In EMD, one first subtracts the local median (or detrending) $c_1(t)$ from the signal $x(t)$ so that the residual $x(t) - c_1(t)$ has zero mean and oscillates around zero.

- Then, the instantaneous frequency is defined through the Analytic Signal method using the Hilbert Transform:

$$z_1(t) = c_1(t) + iH(c_1)(t) = a_1(t)exp(i\theta_1(t)).$$

This gives

$$c_1(t) = a_1(t)\cos(\theta_1(t)),$$

and the instantaneous frequency is defined as $\omega = \frac{d\theta_1(t)}{dt}$.

- Define an Instantaneous Fourier Mode (IFM) as $\cos(\theta(t)$ with $\theta$ chosen adaptively to the signal so that we obtain a **sparsest representation of the signal**:

$$f(t) = a_0(t) + a_1(t)\cos(\theta_1(t)),$$

where $a_0$ is the local median of $f(t)$ and $a_1(t)$ the amplitude of the envelope, which are assumed to be smoother than $\cos(\theta_1(t))$.

- The Instantaneous Frequency is defined as $\omega(t) = \frac{d}{dt}\theta_1(t)$.
- We can expand $a_0(t)$ to generate a sparse decomposition of $f(t)$:

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

with the **smallest possible** $M$ among all possible IFMs $\{\cos(\theta_k(t))\}$, where $r_M(t)$ is the residual which is either monotone or has at most one extremum.

# Building the largest dictionary for IFMs

**Definition**

(Instantaneous Fourier Mode) $g(t),\ t \in [a,b]$ is a Instantaneous Fourier Mode (IFM for short), if there is a function $\theta(t)$ such that $g(t) = \cos\theta(t)$, where $\theta(t)$ satisfies the following conditions:

1. $\theta(t)$ is continuous;
2. $\theta(t)$ is monotonously increasing;
3. $\theta(t)$ is piecewise differentiable.

The instantaneous frequency is defined as

$$\omega(t) = \frac{d}{dt}\theta(t), \quad \theta(t) = \arccos(g(t)). \tag{1}$$

- The concept of IFM generalizes the traditional concept of frequency in a substantial way.

# General Framework

Now, the problem we need to solve is

$$
\begin{aligned}
&\textit{Minimize} && M \\
&\textit{Subject to} && \sum_{k=1}^{M} a_k(t)\cos(\theta_k(t)) = f(t) \\
& && \cos(\theta_k(t)) \textit{ are IFMs}, \\
& && a_k(t) \textit{ is smoother than } \cos\theta_k.
\end{aligned}
\tag{2}
$$

This is a very difficult optimization problem. Notice that the smoother a signal is, the fewer IMFs it would contain. Based on this observation, we propose two nonlinear optimization methods to decompose the signal sequentially. The first method is based on nonlinear **Basis Pursuit** ($L^1$-minimization), and the second on nonlinear **Matching Pursuit** (Greedy Algorithms).

# Nonlinear Basis Pursuit using 3rd order Total Variation

To measure the smoothness of the envelope function and the median, we use the 3rd order total variation norm and solve the following nonlinear optimization problem to obtain a sparsest decomposition of a given signal $f$,

$$\begin{aligned} (P) \quad \text{minimize} \quad & TV^3(a_0(t)) + TV^3(a_1(t)), \\ \text{subject to:} \quad & a_0(t) + a_1(t)\phi(t) = f(t) \\ & \phi(t) = \cos(\theta(t)) \text{ is an IFM.} \end{aligned} \tag{3}$$

where the 3rd order total variation is defined as follows:

$$TV^3(g) = \int_a^b |g^{(4)}(t)| dt \tag{4}$$

where $g^{(4)}(t)$ is the 4th derivative. $TV^3$-based optimization tends to favor cubic splines, which give properties similar to those of EMD.

# Optimization-Projection Iteration

- $\phi^0 = \Phi[f]$, $\phi^0 = \cos\theta^0, \quad \psi^0 = \sin\theta^0.$

Step 1: Get $a_0^n,\ a_1^n,\ b_1^n$ by solving following linear optimization problem:

$$\begin{aligned} \text{minimize} \qquad & TV^3(a_0^n) + TV^3(a_1^n) + \lambda^{n-1} TV(b_1^n), \\ \text{Subject to} \ : \qquad & a_0^n + a_1^n \phi^{n-1}(t) + b_1^n \psi^{n-1}(t) = f(t). \end{aligned}$$

Step 2: Update the IFM:

$$\theta^n = \theta^{n-1} - \arctan\left(\frac{b_1^n}{a_1^n}\right), \quad \phi^n = \cos\theta^n, \quad \psi^n = \sin\theta^n.$$

Step 3: If $\phi^n$ is not an IFM

$$\phi^n = \Phi[f - a_0^n].$$

Step 4: If $\|b_1\|_2 < \epsilon$, stop. Otherwise, go to Step 1.

# Numerical Results

**Example 1:**

$$a_0 = \sin 2\pi t + 0.5\cos 4\pi t + 0.2 + \cos 6\pi t, \tag{5}$$

$$a_1 = 2 + \cos(2\pi t) + 0.5\cos(4\pi t) + 0.3\sin(6\pi t), \tag{6}$$

$$\theta = 5\sin(2\pi t) + 0.4\cos(4\pi t) + 20\pi(t+0.5)^2 + 2, \tag{7}$$

$$f = a_0 + a_1\cos\theta \tag{8}$$

# Numerical Results

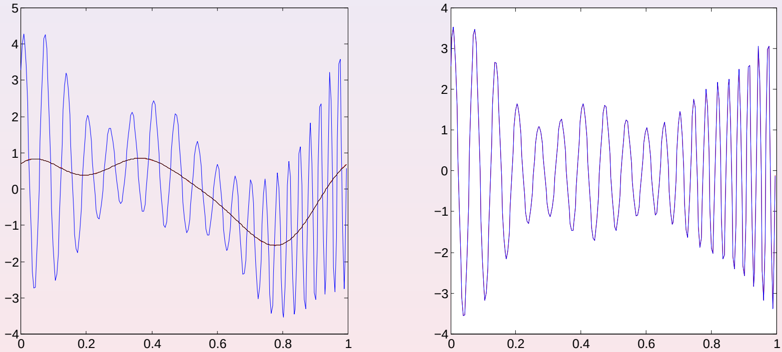



Figure: Original data and local median, Example 1.

# Numerical Results

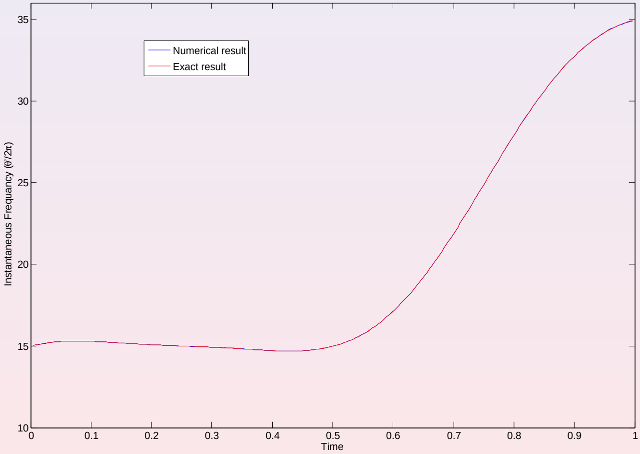



Figure: Instantaneous frequency, Example 1.

# Numerical Results

**Example 2:**

$$
\begin{aligned}
a_0 &= \frac{1}{1.1+\cos 2\pi t}, && (9)\\
a_1 &= 2+\cos(2\pi t)+0.5\cos(4\pi t)+0.3\sin(6\pi t), && (10)\\
\theta_1 &= \sin(2\pi t)+0.4\cos(4\pi t)+30\pi t+2, && (11)\\
\theta_2 &= \sin(2\pi t)+0.4\cos(4\pi t)+60\pi t+2, && (12)\\
f &= a_0+a_1\cos\theta_1+a_1\cos\theta_2 && (13)
\end{aligned}
$$

# Numerical Results

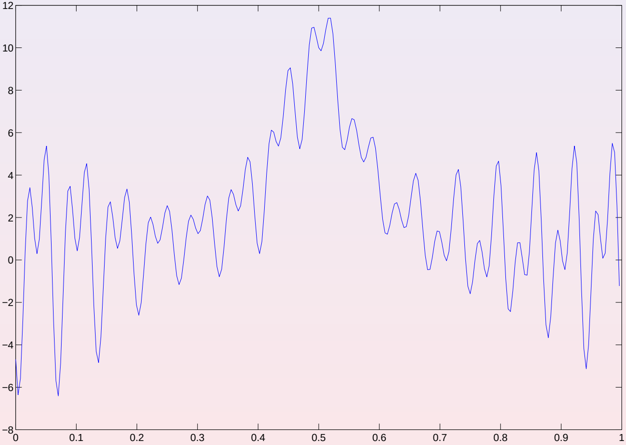



Figure: Original data and local median, Example 2.

# Numerical Results

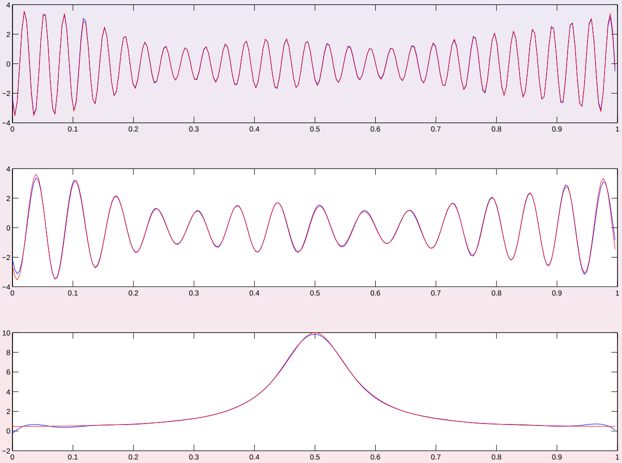



Figure: two IMFs and the trend. Blue: Numerical resutls; Red: Exact results.

# Length of the day

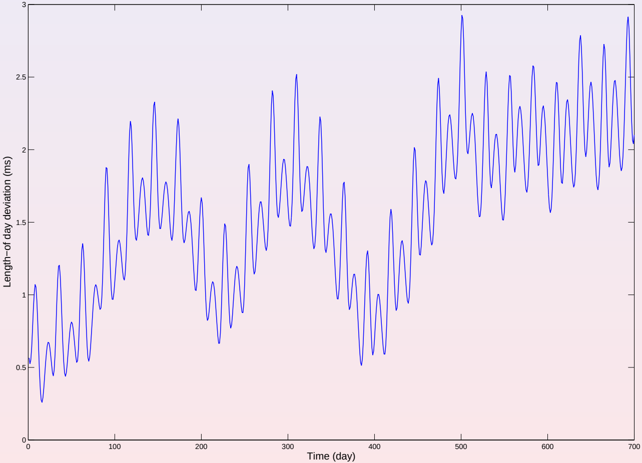



Figure: Length of the day.

# Length of the day

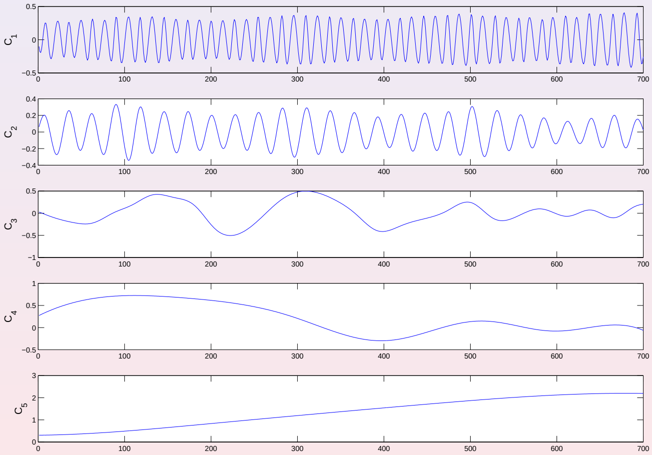



Figure: The IMFs decomposed by our method.

# Length of the day

The oscillations in different IMFs are dominated by the tides of both lunar and solar origin.
$C_1$: semi-monthly tides;
$C_2$: monthly tides;
$C_3$: semi-annual cycle;
$C_4$: annual cycle.

# Global Surface Temperature

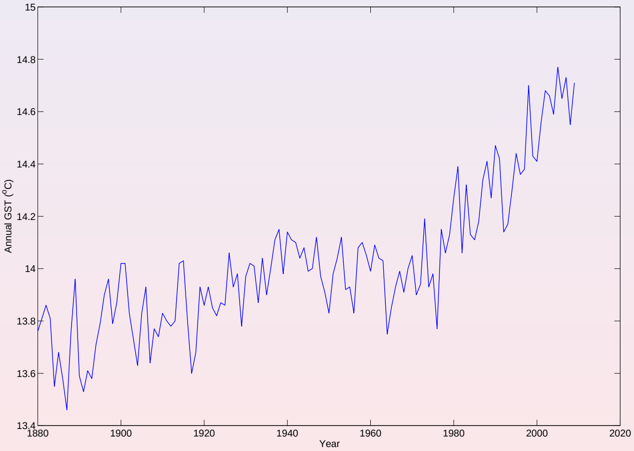



Figure: Annual global surface temperature from 1880 to 2009.

# Global Surface Temperature

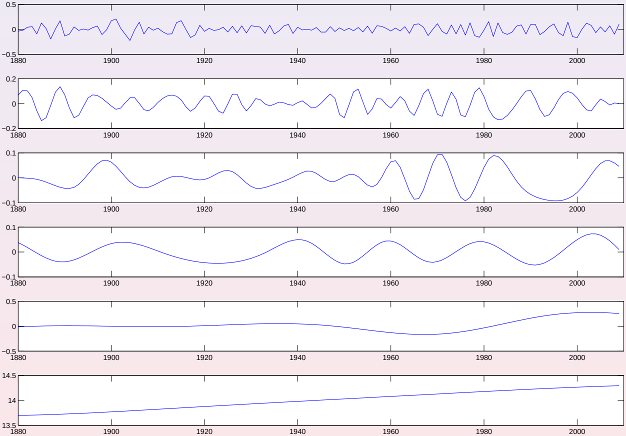



Figure: IMFs.

# Global Surface Temperature

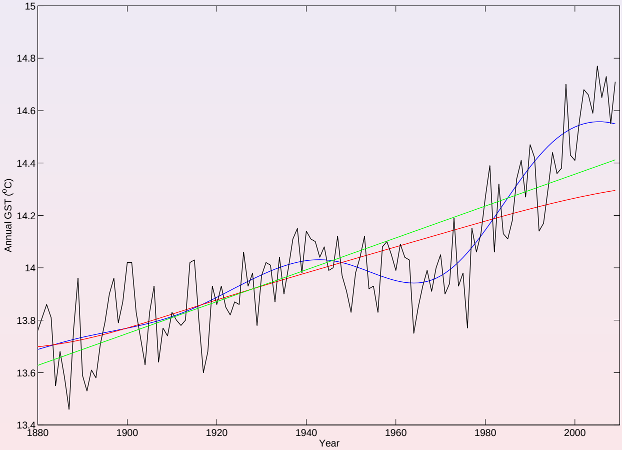



Figure: Adaptive and linear trends. Black: original data; Blue: multidecadal trend; Red: overall adaptive trend; Green: linear trend.

# Exact Recovery of Nonlinear Sparse Data

If we assume that $a_0$, $a_1$ and $\theta$ are sparse under the Fourier or Wavelet basis, then we can design a nonlinear $l^1$ minimization of the coefficients of $a_0$, $a_1$ and $\theta$. This gives rise to the following nonlinear $l^1$ optimization algorithm:

$$
\begin{array}{lll}
(P_n^w) & \text{minimize} & |\boldsymbol{\omega}_{a_0} \widehat{a}_0^{n+1}|_1 + |\boldsymbol{\omega}_{a_1} \widehat{a}_1^{n+1}|_1 + |\boldsymbol{\omega}_\theta \widehat{\delta_\theta}^{n+1}|_1, \\
 & \text{Subject to:} & a_0^{n+1} + a_1^{n+1} \cos\theta^n - a_1^n \delta_\theta^{n+1} \sin\theta^n = f.
\end{array}
\tag{14}
$$

# Exact Recovery of Nonlinear Sparse Data

**Theorem**

*Let* $T = \mathrm{supp}\left(\widehat{a}_0, \widehat{a}_1, \widehat{\theta} - \widehat{\theta}^n\right)$. $\mu_i = \max_{j \in T,\, j \neq i} \mu_{ij}$, *where* $\mu_{ij} = \frac{|<A_i^n, A_j^n>|}{\|A_i^n\|_2 \|A_j^n\|_2}$. $\mu_0 = \max_{i \in T} \mu_i$, *and the weight is choosing according to* $\omega_i = \max\left\{\frac{\mu_i}{\lambda}, 1\right\}$, $\lambda$ *is a constant. If the support* $T$ *satisfies that*

$$K_0 = |T| \leq \frac{1}{2\max\{\mu_0, \lambda\}} \tag{15}$$

*then*

$$\|e^{n+1}\|_1 \leq C\|e^n\|_1^2 \tag{16}$$

*where* $e^n = (a_0 - a_0^n, a_1 - a_1^n, \theta - \theta^n)$ *and* $C$ *is a constant.*

# Numerical Results

$$a_0 = \sin 2\pi t + 0.5\cos 4\pi t + 0.2 + \cos 6\pi t, \tag{17}$$

$$a_1 = 2 + \cos(2\pi t) + 0.5\cos(4\pi t) + 0.3\sin(6\pi t), \tag{18}$$

$$\theta = \sin(2\pi t) + 0.4\cos(4\pi t) + 20\pi t + 2, \tag{19}$$

$$f = a_0 + a_1 \cos\theta \tag{20}$$

# Numerical Results

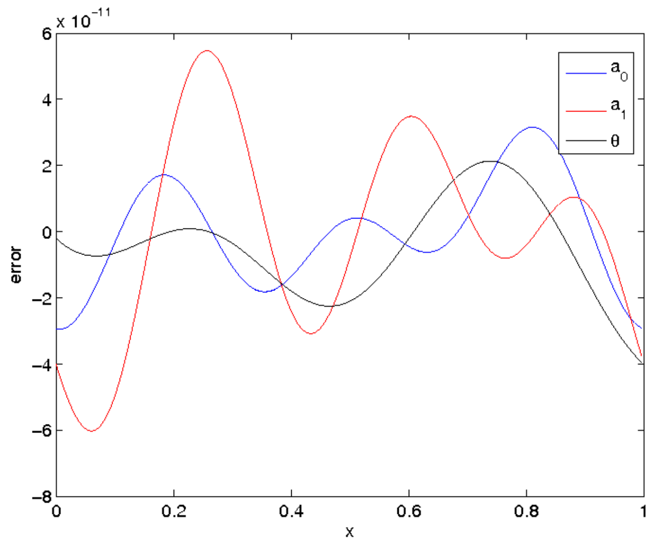



Figure: error of $a_0, a_1$ and $\theta$.

# Optimization based on nonlinear Matching Pursuit

Let $V(\theta,\lambda)$ $(\lambda \le 1/2)$ be an overcomplete Fourier basis given below:

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

where $L_\theta = \lfloor \frac{\theta(1)-\theta(0)}{2\pi} \rfloor$. We start with $r_0 = f$ and an initial guess for $\theta_0$.
**Step 1**: Solve the following $l^1$-regularized nonlinear least-square problem:

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

where $\gamma > 0$ is a 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.

# An $l^1$ regularized nonlinear least-sqaure solver

**Step 1**: Solve for $(a_k^{n+1}, b_k^{n+1})$ from the following $l^1$ regularized least-square problem:

$$\mathrm{Min}_{(a,b)} \;\; \gamma(\|\widehat{a}\|_{l^1} + \|\widehat{b}\|_{l^1}) + \|r_{k-1} - a\cos\theta_k^n - b\sin\theta_k^n\|_{l^2}^2$$
$$\text{Subject to: } \; a \in V(\theta_k^n, \lambda), \; b \in V(\theta_k^n, \lambda),$$

where $\widehat{a}, \widehat{b}$ are the representations of $a, b$ in the $V(\theta_k^n, \lambda)$ space.
**Step 2**: Update $\theta_k^n$:

$$\Delta\theta' = P_{V(\theta^n;\,\eta)}\left(\frac{d}{dt}\left(\arctan\left(\frac{b_k^{n+1}}{a_k^{n+1}}\right)\right)\right), \qquad \theta_k^{n+1} = \theta_k^n - \beta\Delta\theta,$$

where $\beta \in [0,1]$ is chosen to make sure that $\theta_k^{n+1}$ is monotonically increasing, $P_{V(\theta_k^n;\,\eta)}$ is the projection operator to the space $V(\theta_k^n;\eta)$.
**Step 3**: If $\|\theta_k^{n+1} - \theta_k^n\|_2 > \epsilon_0$, set $n = n+1$ and go to Step 1. Otherwise, go to Step 4.
**Step 4**: If $\eta \geq \lambda$, stop. Otherwise, set $\eta = \eta + \Delta\eta$ and go to Step 1.

# A fast algorithm based on FFT for periodic data

For periodic data, we can use a standard Fourier basis to construct the $V(\theta, \lambda)$ space instead of the overcomplete Fourier basis:

$$V_p(\theta, \lambda) = \text{span}\left\{1, \left(\cos\left(\frac{k\theta}{L_\theta}\right)\right)_{1\leq k\leq \lambda L_\theta}, \left(\sin\left(\frac{k\theta}{L_\theta}\right)\right)_{1\leq k\leq \lambda L_\theta}\right\}.$$

Since the standard Fourier basis is an orthogonal basis, the $l^1$ regularized term is not necessary in our nonlinear optimization.

$$\min_{a,b} \quad \|r_k - a\cos\theta_k^n - b\sin\theta_k^n\|_{l^2}^2.$$

$$\text{Subject to} \quad a,\ b \in V_p(\theta_k^n, \lambda).$$

# Gauss-Newton type iteration

- $\theta_k^0 = \theta_0, \ r^0 = f$.

Step 1: Solve the following linear least-square problem:

$$\begin{aligned} &\text{Minimize} \quad \| r^{k-1} - a_k^{n+1}(t)\cos\theta_k^n(t) - b_k^{n+1}(t)\sin\theta_k^n(t)\|_2^2 \\ &\text{Subject to} \quad a_k^{n+1}(t), \ b_k^{n+1}(t) \in V(\theta_k^n). \end{aligned}$$

Step 2: Update $\theta_k^n$,

$$\theta_k^{n+1} = \theta_k^n - \lambda \arctan\left(\frac{b_k^{n+1}}{a_k^{n+1}}\right), \qquad (21)$$

where $\lambda \in [0,1]$ is chosen to make sure that $\theta_k^{n+1}$ is a monotonely increasing function.

$$\lambda = \max\left\{\alpha \in [0,1] : \frac{d}{dt}\left(\theta_k^n - \alpha \arctan\left(\frac{b_k^{n+1}}{a_k^{n+1}}\right)\right) \geq 0\right\}. \qquad (22)$$

Step 3: If $\|\theta_k^{n+1} - \theta_k^n\|_2 < \epsilon_0$, stop. Otherwise, go to Step 1.

# Fast algorithm for periodic data

The least square problem can be solved approximately by FFT in $\theta$ space by using the fact that $\cos\theta$ and $\sin\theta$ are single Fourier basis in $\theta$ space:

Step 1: Interpolate $r_{k-1}$ from $\{t_i\}_{i=1}^N$ in the physical space to a uniform mesh in the $\theta_k^n$-coordinate to get $r_{\theta_k^n}$ and compute the Fourier transform $\widehat{r}_{\theta_k^n}$:

$$r_{\theta_k^n, j} = \text{Interpolate } \left(r_{k-1}, \theta_{k,j}^n\right), \tag{23}$$

where $\theta_{k,j}^n$, $j = 0, \cdots, N-1$ are uniformly distributed in the $\theta_k^n$-coordinate,i.e. $\theta_{k,j}^n = 2\pi L_{\theta_k^n}\, j/N$. And the Fourier transform of $r_{\theta_k^n}$ is given as follows

$$\widehat{r}_{\theta_k^n}(\omega) = \frac{1}{N}\sum_{j=1}^{N} r_{\theta_k^n, j} e^{-i2\pi\omega\overline{\theta}_{k,j}^n}, \quad \omega = -N/2+1, \cdots, N/2, \tag{24}$$

where $\overline{\theta}_{k,j}^n = \frac{\theta_{k,j}^n - \theta_{k,0}^n}{2\pi L_{\theta_k^n}}$.

# Fast algorithm for periodic data

Step 2: Apply a cutoff function to the Fourier Transform of $r_{\theta_k^n}$ to compute $a$ and $b$ on the mesh of the $\theta_k^n$-coordinate, denoted by $a_{\theta_k^n}$ and $b_{\theta_k^n}$:

$$
\begin{aligned}
a_{\theta_k^n} &= \mathcal{F}^{-1}\left[\left(\widehat{r}_{\theta_k^n}\left(\omega + L_{\theta_k^n}\right) + \widehat{r}_{\theta_k^n}\left(\omega - L_{\theta_k^n}\right)\right) \cdot \chi_\lambda\left(\omega / L_{\theta_k^n}\right)\right], \\
b_{\theta_k^n} &= \mathcal{F}^{-1}\left[i \cdot \left(\widehat{r}_{\theta_k^n}\left(\omega + L_{\theta_k^n}\right) - \widehat{r}_{\theta_k^n}\left(\omega - L_{\theta_k^n}\right)\right) \cdot \chi_\lambda\left(\omega / L_{\theta_k^n}\right)\right].
\end{aligned}
$$

$\mathcal{F}^{-1}$ is the inverse Fourier transform defined in the $\theta_k^n$ coordinate:

$$
\mathcal{F}^{-1}\left(\widehat{r}_{\theta_k^n}\right) = \frac{1}{N} \sum_{\omega=-N/2+1}^{N/2} \widehat{r}_{\theta_k^n} e^{i 2\pi\omega \overline{\theta}_{k,j}^n}, \quad j = 0, \cdots, N-1.
$$

Step 3: Interpolate $a_{\theta_k^n}$ and $b_{\theta_k^n}$ from the uniform mesh $\{\theta_{k,j}^n\}_{j=1}^N$ in the $\theta_k^n$-coordinate back to the physical grid points $\{t_i\}_{i=1}^N$:

$$
\begin{aligned}
a(t_i) &= \text{Interpolate}\left(a_{\theta_k^n}, t_i\right), \quad i = 0, \cdots, N-1, \\
b(t_i) &= \text{Interpolate}\left(b_{\theta_k^n}, t_i\right), \quad i = 0, \cdots, N-1,.
\end{aligned}
$$

# Fast algorithm for periodic data

Step 4: Update $\theta^n$ in the $t$-coordinate:

$$\Delta\theta' = P_{V(\theta;\,\eta)}\left(\frac{d}{dt}\left(\arctan\left(\frac{b_k^{n+1}}{a_k^{n+1}}\right)\right)\right), \quad \Delta\theta(t) = \int_0^t \Delta\theta'(s)ds,$$
$$\theta_k^{n+1} = \theta_k^n - \beta\Delta\theta,$$

where $\beta \in [0,1]$ is chosen to make sure that $\theta_k^{n+1}$ is monotonically increasing:

$$\beta = \max\left\{\alpha \in [0,1] : \frac{d}{dt}\left(\theta_k^n - \alpha\Delta\theta\right) \geq 0\right\}.$$

and $P_{V_p(\theta;\,\eta)}$ is the projection operator to the space $V_p(\theta;\,\eta)$.

Step 5: If $\|\theta_k^{n+1} - \theta_k^n\|_2 < \epsilon_0$, go to step 6. Otherwise, set $n = n+1$ and go to Step 1.

Step 6: If $\eta \geq \lambda$, stop. Otherwise, set $\eta = \eta + \Delta\eta$ and go to step 1.

**Signal I:**

$$f(t) = \cos(60\pi t + 15\sin(2\pi t)) \tag{25}$$

where $X(t)$ is a white noise with zero mean and variance $\sigma^2 = 1$.

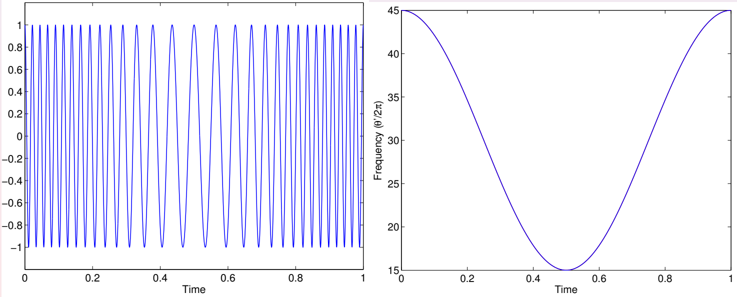



Figure: Original data without noise and its instantaneous frequency.

# Numerical example: periodic data

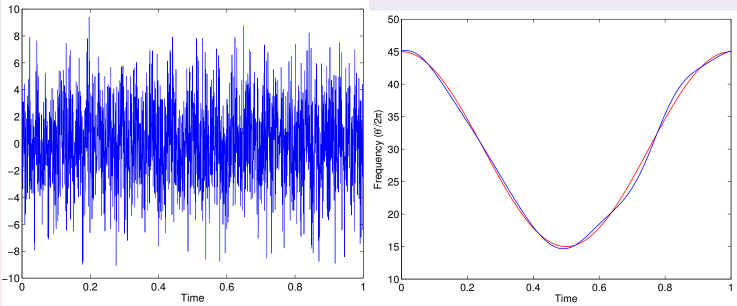



Figure: Noised signal $f(t) + 3X(t)$ and its instantaneous frequency, , corresponding SNR is $-12.55$ dB. The Signal-Noise Ratio (SNR, measured in dB) is defined by $\mathrm{SNR[dB]} = 10\log_{10}\left(\frac{\mathrm{var} f}{\sigma^2}\right)$.

# Numerical example: periodic data

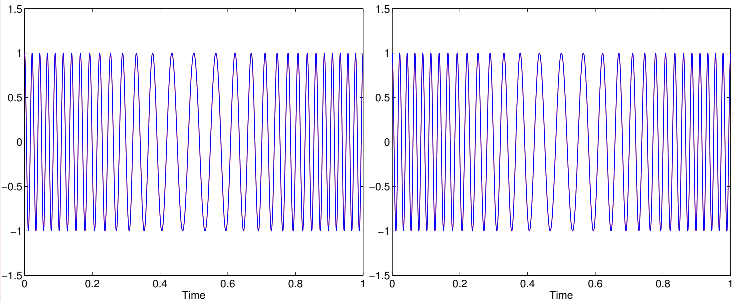



Figure: The IMFs extracted by our method and EMD method for $f(t)$ without noise, where $f(t)$ is signal I.

# Numerical example: periodic data

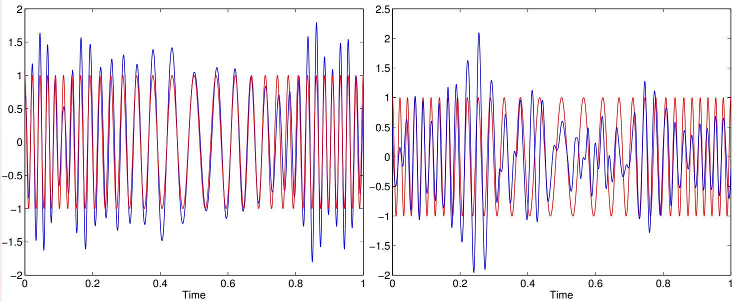



Figure: The IMFs extracted by our method and EEMD method for $f(t) + 3X(t)$, where $f(t)$ is signal I.

# Numerical example: periodic data

$$\begin{aligned} f(t) &= \frac{1}{1.5+\cos(2\pi t)}\cos(60\pi t + 15\sin(2\pi t)) \\ &+ \frac{1}{1.5+\sin(2\pi t)}\cos(160\pi t + \sin(16\pi t)) \end{aligned} \tag{26}$$

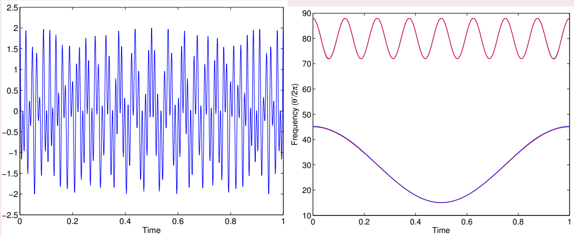
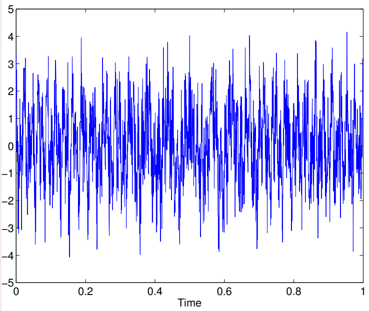



Figure: Upper row: left: signal defined in (26) without noise; right: Instantaneous frequencies; red: exact frequencies; blue: numerical results.

# Numerical example: periodic data



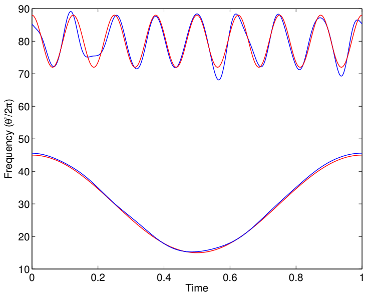




Figure: Upper row: left: signal defined in (26) with Gaussian noise $X(t)$; right: Instantaneous frequencies; red: exact frequencies; blue: numerical results.

# Numerical example: non-periodic data

**Signal II:**

$$\theta_1 = 20\pi(t+1)^2 + 1, \quad \theta_2 = 161.4\pi t + 4(1-t)^2 \sin(16\pi t),$$
$$f(t) = \frac{1}{1.5 + \sin(1.5\pi t)} + (2t+1)\cos\theta_1 + (2-t)^2 \cos\theta_2. \quad (27)$$

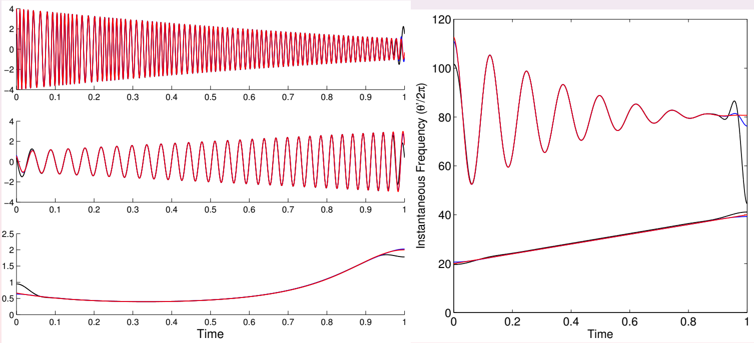
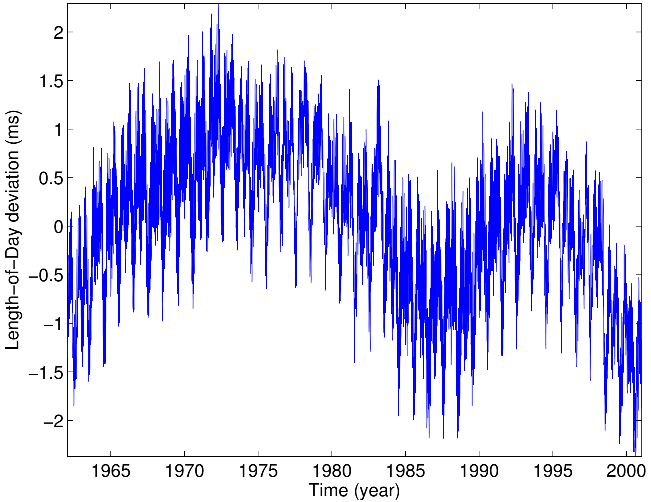



Figure: IMF (left) and Instantaneous frequency (right) of the signal obtained from different methods. Red: exact; Blue: $l^1$ regularized least square; Black: Fourier transform with mirror reflection.

# Length of the day

The method based on FFT is very effective in computation, so we can handle data with large size, for example the length of the day data, from 20 January 1962 to 6 January 2001, for a total of 14,232 days.



# Length of the day

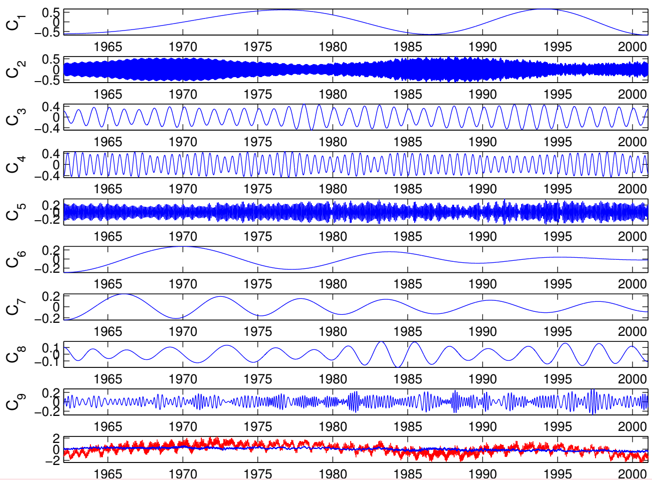



Figure: The IMFs decomposed by our method based on FFT.

# Convergence analysis

## Theorem

*For the data $f = f_0 + f_1 \cos\theta$, assume that the instantaneous frequency $\theta'$ is $M_0$-sparse over the Fourier basis in physical space,*
$\theta' \in V_{M_0} = span\left\{e^{i2k\pi t/T}, k = -M_0, \cdots, 1, \cdots, M_0\right\}$,
*mean $f_0$ and envelop $f_1$ are $M_1$-sparse over the Fourier basis in $\theta$-space,*

$$\widehat{f}_{0,\theta}(k) = \widehat{f}_{1,\theta}(k) = 0, \quad \forall |k| > M_1.$$

*If the initial guess of $\theta$ satisfies*

$$\|\mathcal{F}\left(\left(\theta^0 - \theta\right)'\right)\|_1 \le \pi M_0/2, \tag{28}$$

*then there exist $\eta_0 > 0$ such that*

$$\|\mathcal{F}\left(\left(\theta^{m+1} - \theta\right)'\right)\|_1 \le \frac{1}{2}\left\|\mathcal{F}\left((\theta^m - \theta)'\right)\right\|_1, \tag{29}$$

*provided $L \ge \eta_0$, where $\eta_0$ is a constant determined by $M_0$, $M_1$, where $L = \frac{\theta(T)-\theta(0)}{2\pi}$.*

# Numerical Validation

$$\theta = 20\pi t + 2\cos 2\pi t + 2\sin 4\pi t, \quad \overline{\theta} = \theta/10$$
$$a_0 = 2 + \cos\overline{\theta} + 2\sin 2\overline{\theta} + \cos 3\overline{\theta}, \quad a_1 = 3 + \cos\overline{\theta} + \sin 3\overline{\theta}$$
$$f = a_0 + a_1 \cos\theta.$$

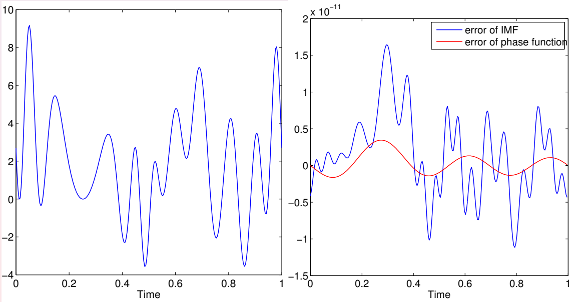



Figure: Original data and errors of IMF and phase function.

# Data with sparse samples

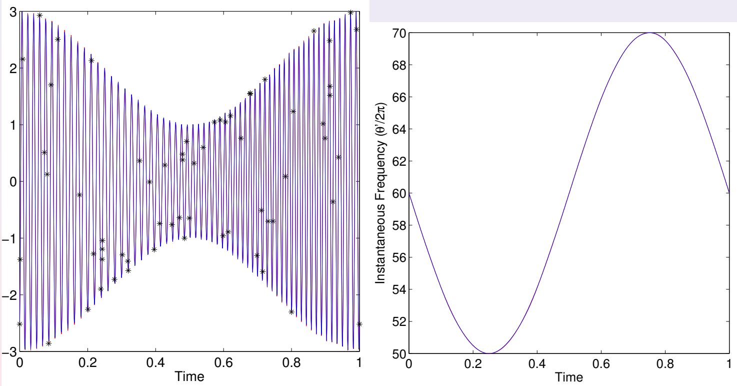



Figure: Left: original samples, red: exact; blue: recovered; '*' represent the sample points. Right: instantaneous frequency, red: exact; blue: numerical.

In this example, the number of samples is 64, the number of periods is 60. There is approximately one sample point in one period of the signal. The location $t_i$ is chosen randomly in $[0, 1]$.

**Theorem**

*If the samples points are selected at random, under the same assumption in Theorem 3, we have there exist $\eta_0 > 0$, $\eta_1 > 0$, such that with overwhelming probability*

$$\|\mathcal{F}\left(\left(\theta^{m+1}-\theta\right)'\right)\|_1 \leq \frac{1}{2}\left\|\mathcal{F}\left(\left(\theta^{m}-\theta\right)'\right)\right\|_1, \tag{30}$$

*provided $L \geq \eta_0$ and $N_s \geq \eta_1 \max(\overline{\theta})' \left(\log N_b\right)^6$, $N_s$ is the number of samples, $N_b$ is the number of basis.*

# Data with poor scale-separation

Suppose, we have a set of data $f$ which has following sparse decomposition over dictionary $\mathcal{D}$:

$$f(t) = \sum_{k=1}^{M} a_k \cos\theta_k, \quad a_k \cos\theta_k \in \mathcal{D}, \tag{31}$$

where $\mathcal{D}$ is defined as following

$$\mathcal{D} = \left\{ a(t)\cos\theta(t) : \; \theta'(t) \geq 0, \; a(t) \in V(\theta) \right\}, \tag{32}$$

But now the instantaneous frequencies $\theta_k'(t)$ are not well separated, so $f(t)$ does not satisfy the scale-separation condition.

It is known that for the data consist of components with close frequencies, matching pursuit with Gabor dictionary may not get the sparse decomposition. Our method is based on the matching pursuit, so it may not get the right sparsest decomposition neither.

# Data with poor scale-separation

$$f(t) = \cos(20\pi t + 40\pi t^2 + \sin(2\pi t)) + \cos(40\pi t) \tag{33}$$

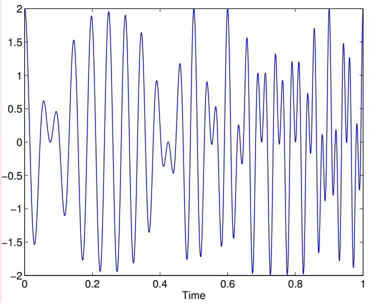



Figure: Original data.

# Data with poor scale-separation

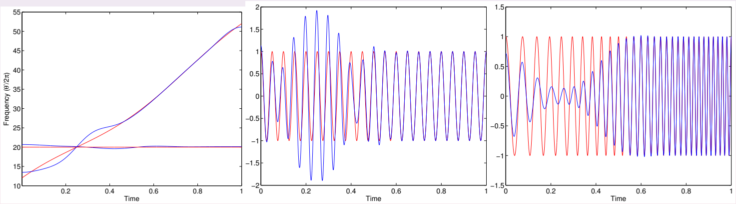



Figure: Left: Instantaneous frequencies; red: exact frequencies; blue: numerical results. Middle and Right: IMFs extracted by previous method .

# Data with poor scale-separation

For this kind of signals, we have to decompose these components simultaneously, since they have strong correlation. Following this idea, the signal should be decomposed by solving following optimization problem:

$$\min \|f(t) - \sum_{k=1}^{K} a_k \cos\theta_k\|_2, \quad s.t. \quad a_k \cos\theta_k \in \mathcal{D}. \tag{34}$$

Here, we assume that $K$ is known.

This problem is much more difficult to solve than the original one, since the different components may have strong correlation.

Based on the idea of $l_1$ regularized least square, we have developed a method to solve above optimization problem.

Here we give an example to demonstrate that this new method has capability to deal with the signal without scale separation.

# Data with poor scale-separation

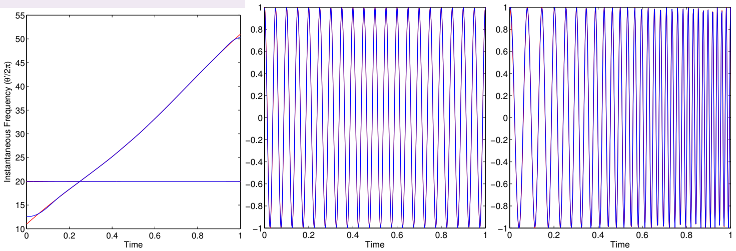



Figure: Left: Instantaneous frequencies; Middle and right: IMFs extracted by extracting two IMFs together. red: exact results; blue: numerical results.

# Data with Intra-wave frequency modulation

In the previous approach, we want to decompose the signal to the form

$$f(t) = \sum_{k=1}^{M} a_k(t) \cos \theta_k(t). \tag{35}$$

For some applications, decompositions of above form are too restrictive. For example, the decomposition of following ECG data is not sparse at all.

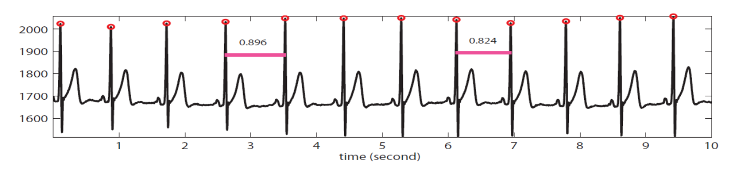



Figure: Typical ECG data

# Data with Intra-wave frequency modulation

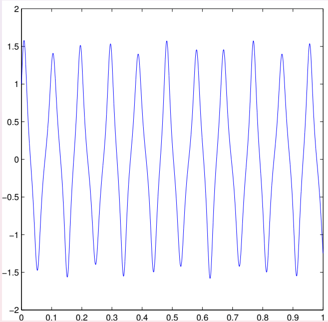
Inspired by the work of Prof. Daubechies and Wu, we decompose the signal to the following form

$$f(t) = \sum_{k=1}^{M} a_k(t) s_k\left(\theta_k(t)\right). \tag{36}$$

where $s_k(\cdot)$ is any $2\pi$-period function, which is called the shape function or the oscillatory pattern. Based on the method looking for the

instantaneous frequency, we have developed an algorithm to find the shape function $s_k$ and decompose the signal.

# Data with intra-wave frequency modulation: Duffing equation



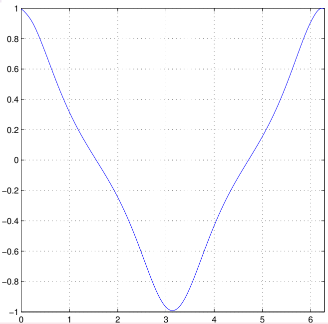




Figure: Left: The solution of the Duffing equation; Right: The shape function obtain by our method.

# Nonlinearity Analysis

Once the shape function has been obtained, we can conduct nonlinearity analysis to identify its nonlinearity.

The basic idea is that for each shape function (or IMF), assume that there exists an ODE such that the shape function (or IMF) is the solution of this ODE with proper initial conditions. More repcisely, we assume that this ODE is of second order and has following form:

$$\ddot{x} + a(x,t)\dot{x} + b(x,t) = 0 \tag{37}$$

and $a(x,t)$, $b(x,t)$ are slowly varying with respect to $t$.

$a(x,t)$ and $b(x,t)$ can be recovered by solving an optimization problem under the assumption that they are sparse over some basis.

# Nonlinearity Analysis

The Duffing equation is a nonlinear ODE which has the following form:

$$\ddot{u} + u + u^3 = 0 \tag{38}$$

$u(0) = 1, \dot{u}(0) = 0.$

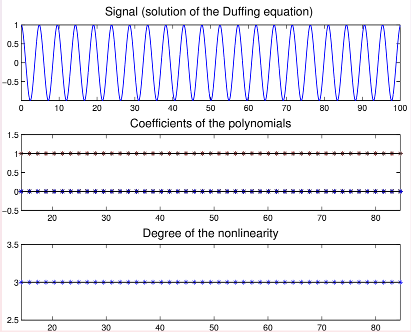



Figure: Top: the solution of the Duffing equation; Middle: Coefficients of polynomials; Bottom: Degrees of nonlinearity.

# Nonlinearity Analysis

Consider a ODE with varying coefficients,

$$\ddot{u} + a(t)(u^2 - 1)\dot{u} + (1 - a(t))u^3 + u = 0 \tag{39}$$

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

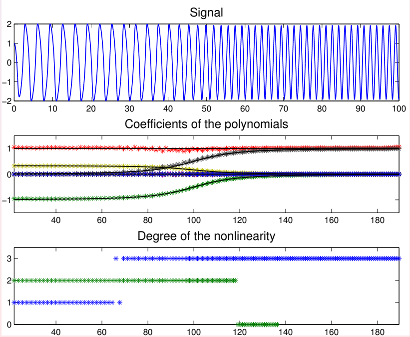



Figure:

# Concluding Remarks

- We introduce a sparse time-frequency analysis method to study Trend and Instantaneous Frequency of multiscale signals.
- By combining ideas from EMD and compressed sensing, we develop effective nonlinear optimization methods which give a sparsest representation of multiscale signals, which preserves some intrinsic physical properties of the original signal.
- This method gives a rigorous definition of Instantatneous Frequency and can be considered as a nonlinear version to Compressed Sensing and a mathematical foundation for the EMD method.
- Convergence analysis has been performed under some scale separation assumption on the mutiscale data.
- Applications to biomedical problems, climate data, and geophysical problems are under investigation.

## Reference:


- T. Y. Hou and Z. Shi, Adaptive Data Analysis via Sparse Time-Frequency Representation, *Advances in Adaptive Data Analysis*, **3**, pp. 1-28, 2011.
- T. Y. Hou and Z. Shi, Data-driven Time-Frequency Analysis, to appear in ACHA.


# Thank you!