ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 14, No. 1, 2019, *pp*.025-034

Error Analysis for Sparse Time-Frequency Decomposition of Non-Integer Period Sampling Signals

Qi Yang [1+]

[1] *Department of Mathematics, Jinan University, Guangzhou, 510632, China*
+ E-mail: yangqi@stu2016.jnu.edu.cn
(Received October 24, 2018, accepted December 21, 2018)

Abstract. In this paper, we review a nonlinear matching pursuit approach (Hou and Shi, 2013), a data-driven time-frequency analysis method, which is looking for the sparsest representation of multiscale data over a dictionary consisting of all intrinsic mode functions (IMFs). In many practical problems, signals are non-integer period sampled. In other words, the time window may not contain exactly an integer number of signal periods. We consider the sparse time-frequency decomposition of non-integer period sampling signals by the nonlinear matching pursuit method and estimate the error. The estimation show that the relative error depends on the separation factor, the frequency ratio, and the number of periods of the IMF.

Keywords: sparse time-frequency decomposition, non-integer period sampling, scale separation

1. Introduction

Data is ubiquitous in our lives, and how to extract useful information from multiscale data has aroused research upsurge in recent years. The Fourier transform is a conventional method for analyzing linear and stationary data by decompose the signal into a linear combination of different frequency components. While the Fourier transform is widely applied in many fields, but it is still has some limitations in processing oscillatory data. When traditional data analysis methods can not provide simple indices with clear physical meanings, new ideas are in need. Time-frequency analysis method represents one-dimensional time signal into time-frequency (TF) domains, which is an important tool for nonstationary and nonlinear signal analysis. These include the short time Fourier transform, the wavalet transform [17], and the Wigner-Ville distribution [15]. The analytic signal (AS) method is another important approach in TF analysis was introduced by Van der Pol [2] and Gabor [7], and it has shown its usefulness in many applications, see, e.g. [11, 16, 1]. The AS method contains the original function and its Hilbert transform. In addition, an analytic signal has only the non-negative frequency components which has more physical meaning. In the past decades, the empirical mode decomposition (EMD) [13] was proposed to study the dynamics hidden inside an oscillatory signal which has offered an effective method for nonstationary data analysis. The EMD algorithm decomposes a signal into several intrinsic mode functions (IMFs) via a sifting process. The EMD method has shown its usefulness in analysing signal components caused by various sources, see, e.g., [5, 14, 26, 27]. But its mathematical foundation is still lacking cannot be ignored, which is due to the empirical nature. Recently, many variations of EMD were proposed, like the ensemble empirical mode decomposition (EEMD) [25], the variational mode decomposition [12], the synchrosqueezed wavelet transforms [10], etc.

Inspired by the EMD method and compressive sensing [6, 8, 9], Hou and Shi proposed a data-driven time-frequency analysis method [21, 22]. The data-driven time-frequency analysis method based on the sparsest representation of multiscale data to decompose the signal into a finite number of IMF with a small residual:

$$f(t)=\sum_{k=1}^{M}a_k(t)\cos\theta_k(t)+r(t),\quad t\in[0,T], \tag{1.1}$$

where each IMF is a signal with amplitude modulation-frequency modulation (AM-FM). The amplitude $a_k(t)>0$, the instantaneous frequency $\omega_k(t)=\theta_k'(t)>0$, and $r(t)$ is a small residual.

The idea of this method is looking for the sparsest representation over the dictionary consisting of all IMFs by solving the following nonlinear optimization problem:

$$\min_{(a_k)_{1\le k\le M},\,(\theta_k)_{1\le k\le M}} \quad M$$
$$\text{subject to} \quad f(t)=\sum_{k=1}^{M} a_k(t)\cos\theta_k(t),\ 0\le t\le T,\ \ a_k\cos\theta_k\in D, \tag{P$_0$}$$

where D is a dictionary which is defined as

$$D=\{a(t)\cos\theta(t):\theta'(t)\ge 0,\ a(t)\in V(\theta)\}, \tag{1.2}$$

$V(\theta)$ is a is the collection of all the function that are less oscillatory than $\cos\theta(t)$. Liu et al analyzed the uniqueness of the optimization problem (P_0) under the assumption of the scale separation by considering a simplified slow evolution chirp model and studying the wavelet transform of each IMF; it is proved that under the assumption of the scale separation and the signal $f(t)$ is well separated, the solution of (P_0) is unique up to an error determined by the scale separation property [3]. Now, we review the definitions of scale separation and well-separated signal.

Definition 1.1 (scale separation [3]). A function $f(t)=a(t)\cos\theta(t)\,(t\in[0,T])$ is said to satisfy a scale separation property with a separation factor $\varepsilon>0$, if $a(t)$ and $\theta(t)$ satisfy the following conditions:

$$\begin{aligned}&a(t)\in C^1(\mathbb{R}),\ \theta\in C^2(\mathbb{R}),\ \inf_{t\in[0,T]}\theta'(t)>0,\\ &\frac{\sup_{t\in[0,T]}\theta'(t)}{\inf_{t\in[0,T]}\theta'(t)}\le M'<+\infty,\ \left|\frac{a'(t)}{a(t)\theta'(t)}\right|\le\varepsilon,\ \left|\frac{\theta''(t)}{(\theta'(t))^2}\right|\le\varepsilon,\quad \forall t\in[0,T].\end{aligned} \tag{1.3}$$

In the above definition, the condition $\inf\theta'(t)>0\,(t\in[0,T])$ is to ensure that the instantaneous frequency has a physical meaning and the condition $|a'(t)/(a(t)\theta'(t))|\le\varepsilon$ implies the envelope $a(t)$ is smoother than the phase function $\theta(t)$. The variation of the instantaneous frequency $\theta'(t)$ over the time domain is bounded; otherwise, it is more likely to cause the mode mixing. The definition of well separated signal as follows.

Definition 1.2 (well-separated signal [3]). A signal $f:[0,T]\to\mathbb{R}$ is said to be well separated with separation factor ε and frequency ratio β if it can be written as

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

where all $f_k(t)=a_k(t)\cos\theta_k(t)$ satisfy the scale separation property with separation factor ε, $r(t)=O(\varepsilon)$, and their phase function $\theta_k(t)$ satisfy

$$\theta_k'(t)\ge\beta\theta_{k-1}'(t),\qquad \forall t\in[0,T],$$

and $\beta>1, \beta-1=O(1)$.

The optimization problem (P_0) can be seen as a nonlinear version of the L^0 minimization problem which is NP-hard and challenging to solve [21, 22, 23]. Mallat and Zhang introduced an algorithm, called matching pursuit, that builds up a sequence of sparse approximations stepwise and provide the representation of (1.1) [8]. For Gabor dictionaries, S. Qian and D. Chen proposed a similar algorithm [19]. The basis pursuit (BP) is an important optimization principle for decomposing a signal, which was introduced by S. Chen, D. Donoho and M. Saunders [20]. Because the L^1 norm is approximately equivalent to the L^0 norm under certain conditions, the main idea of BP is to find a representation of the signal whose coefficients have the smallest L^1 norm.

It is too difficult to solve the optimization problem (P_0) which is nonlinear and nonconvex. It is like the EMD method, Hou and Shi introduced an algorithm based on matching pursuit to solve problem (P_0) [23, 24]. This nonlinear matching pursuit (NMP) algorithm is stated as follows.

Task. To find a approximate optimizer of the problem (P_0)
Initialization: Set the initial residual $r_0 = f$ and $k = 1$.
Main iteration:
Step 1. Solve the nonlinear least-squares problem as follows:

$$\begin{aligned} &\min_{a_k, \theta_k} \quad \|r_{k-1} - a_k \cos\theta_k\|_{L^2}^2 \\ &\text{subject to} \quad a_k \in V(\theta_k), \theta_k^{'} \geq 0 \end{aligned} \tag{1.4}$$

Step 2. Update the residual

$$r_k = f - \sum_{j=1}^{M} a_j \cos\theta_j. \tag{1.5}$$

Step 3. If $\|r_k\|_{L^2} < \varepsilon_0$, stop. Otherwise, set $k = k+1$ and go back to Step 1.

In the above algorithm, each IMF is given by solving the nonlinear least-squares problem in Step 1. Liu et al gave some estimations of each IMF, and proved each IMF is a local minimizer of (P_0) under the assumptions of scale separation and the signal has integer periods over the time span [3]. A two-level method based on data-driven time-frequency analysis, proposed in [4], to find the sparse time-frequency decomposition. The main idea of the two-level method is run a local algorithm to obtain a approximate instantaneous frequency, then take the result as the initial guess of the global algorithm. In the local algorithm, f is divided into a number of overlapped small time subintervals, and get the local frequency by solving the optimization problem in Step 1 on each time subinterval. Then we can get a pieces of an IMF in each overlapped subinterval, and connect all pieces of the IMF into a global IMF. The main difficulty in the local algorithm is to select a suitable number of periods of IMF over the subintervals.

In this paper, we consider the sparse time-frequency decomposition of non-integer period sampling signals. For any given non-integer period sampling signal $f(t)$ we decompose f over the dictionary D_ε which can be written as

$$D_\varepsilon := \{a(\cdot)\cos\theta(\cdot) : (a,\theta) \in U_\varepsilon\}, \tag{1.6}$$

where

$$U_\varepsilon := \left\{(a,\theta) : a > 0, \theta' > 0; \left|\frac{a'}{a\theta'}\right| \leq \varepsilon, \left|\frac{\theta''}{[\theta']^2}\right| \leq \varepsilon, \frac{\sup_{t\in[0,T]}\theta'(t)}{\inf_{t\in[0,T]}\theta'(t)} \leq M'\right\}. \tag{1.7}$$

Now, we consider the following problem

$$\begin{aligned} &\text{Minimize} \quad p(a,\theta) := \|f(t) - a(t)\cos\theta(t)\|_{L^2}^2 \\ &\text{subject to} \quad (a,\theta) \in U_{\varepsilon.} \end{aligned} \tag{P_2}$$

We can find a sparse representation of the non-integer period sampling signal f by solving the optimization problem (P_2). By the proof of [3, Theorem 3.1], we have the following result:

Lemma 1.1. Let $f(t)$ be a function satisfying the scale property with separation factor ε and frequency ratio β as defined in Definition 1.2:

$$f(t) = \sum_{k=1}^{M} a_k(t)\cos\theta_k(t) + r(t), \quad a_k \cos\theta_k \in U_\varepsilon, \quad a_k = O(1), r = O(\varepsilon).$$

Suppose that $l = \{1, \ldots, M\}$. If

$$p(a,\theta) \leq p(a_l, \theta_l), \tag{1.8}$$

where $p(a,\theta)$ is given in (P_2), then we have

$$\frac{\|a\cos\theta - a_l\cos\theta_l\|_{L^2}}{\|a_l\cos\theta_l\|_{L^2}} \leq \delta_2 + \sqrt{\delta_2^2 + 2(\delta_1 + \delta_2)}, \tag{1.9}$$

where

$$\delta_1 = \sum_{k\neq l} \mu_{k,l} \frac{\|a_k \cos\theta_k\|_{L^2}}{\|a_l \cos\theta_l\|_{L^2}},\ \mu_{k,l} = \frac{|\langle a_l \cos\theta_l, a_k \cos\theta_k\rangle|}{\|a_l \cos\theta_l\|_{L^2} \|a_k \cos\theta_k\|_{L^2}}, \tag{1.10}$$

and

$$\delta_2 = \sum_{k\neq l} \gamma_{k,l} \frac{\|a_k \cos\theta_k\|_{L^2}}{\|a_l \cos\theta_l\|_{L^2}} + \frac{\|r(t)\|_{L^2}}{\|a_l \cos\theta_l\|_{L^2}},\ \gamma_{k,l} = \frac{|\langle a \cos\theta, a_k \cos\theta_k\rangle|}{\|a \cos\theta\|_{L^2} \|a_k \cos\theta_k\|_{L^2}}. \tag{1.11}$$

2. Discussion on problem (P2) with non-integer period sampling signal

In the rest of this paper all the functions that we discuss are supposed to belong to $L^2[0,T]$. For convenience, we write the L^2 norm $\|\cdot\|_{L^2}$ as $\|\cdot\|$ for short. Let $f(t)$ be represented in the following form

$$f(t) = \sum_{k=1}^{M} a_k(t)\cos\theta_k(t) + r(t),\ \forall t\in[0,T]. \tag{2.1}$$

where $a_k(t)\cos\theta_k(t)(k=1,\ldots,M)$ are the IMFs. In this paper, we always suppose that $\eta_1 \ge 2$, where η_1 is the number of periods of $\cos\theta_1(t)$ in the interval $[0,T]$. In many practical problems, signals usually have dozens of periods.

Our main result is summarized in Theorem 2.1.

Theorem 2.1. Let $f(t)$ be a function satisfying the scale property with separation factor ε and frequency ratio β as defined in Definition 1.2:

$$f(t) = \sum_{k=1}^{M} a_k(t)\cos\theta_k(t),\ \ (a_k,\theta_k)\in U_\varepsilon,\ \ a_k = O(1),\ \ \forall t\in[0,T]. \tag{2.2}$$

Suppose that there exist $(a,\theta)\in U_\varepsilon$, $\alpha\in[1,d)$, and $l\in\{1,\ldots,M\}$ such that

$$\alpha^{-1}\theta_l'(t) \le \theta'(t) \le \alpha\theta_l'(t), \qquad \forall t\in[0,T]. \tag{2.3}$$

If

$$p(a,\theta) \le p(a_l,\theta_l), \tag{2.4}$$

where $p(a,\theta)$ is given in (P2), then we have

$$\frac{\|a\cos\theta - a_l\cos\theta_l\|_{L^2}}{\|a_l\cos\theta_l\|_{L^2}} \le O\left(\sqrt{(\eta_l\beta)^{-1}+\varepsilon}\right). \tag{2.5}$$

The proof of Theorem 2.1 relies on some estimates, which we demonstrate in the following lemmas.

Lemma 2.1. Let $a(t),\theta(t)$ satisfy (1.3). Then

$$\frac{1}{2}\|a(t)\|^2\left[1-3\varepsilon-\frac{1}{2\eta}\right] \le \|a(t)\cos\theta(t)\|^2 \le \frac{1}{2}\|a(t)\|^2\left[1+3\varepsilon+\frac{1}{2\eta}\right], \tag{2.6}$$

where

$$\eta = \frac{\theta(T)-\theta(0)}{2\pi}. \tag{2.7}$$

The proof of lemma 2.1 can be found in Appendix A, and it is easy to show that, for any $k\in\{1,\ldots,M\}$, we have

$$\left|\|a_k(t)\cos\theta_k(t)\|^2 - \frac{1}{2}\|a_k(t)\|^2\right| < \frac{1}{2}\|a_k(t)\|^2\cdot\left[3\varepsilon+\frac{1}{2\eta_1\beta^{k-1}}\right]. \tag{2.8}$$

Lemma 2.2. Let $(a,\theta),(\hat{a},\hat{\theta})$ both satisfy (1.3), and there exist $d>1$ such that

$$\theta'(t) < d\hat{\theta}'(t), \qquad \forall t\in[0,T]. \tag{2.9}$$

Then

$$\left|\left\langle a(t)\cos\theta(t), \hat{a}(t)\cos\hat{\theta}(t)\right\rangle\right| < \int_0^T a(t)\hat{a}(t)dt \cdot \left[\frac{\eta}{\eta^2 - \hat{\eta}^2} + 2\varepsilon\left(1 + \frac{1}{(1-d^{-1})^2}\right)\right], \tag{2.10}$$

where

$$\eta = \frac{\theta(T) - \theta(0)}{2\pi}, \qquad \hat{\eta} = \frac{\hat{\theta}(T) - \hat{\theta}(0)}{2\pi}. \tag{2.11}$$

The proof of lemma 2.2 can be found in Appendix B.

2.1. Proof of Theorem 2.1

Using the above estimates, we are ready to prove Theorem 2.1. Now, we give the proof of Theorem 2.1.

Proof of Theorem 2.1. For any $1 \le k \ne l \le M$, if $l > k$, we have

$$\theta_l^{'}(t) < \beta^{l-k}\theta_k^{'}(t), \qquad \forall t \in [0,T], \tag{2.12}$$

and the following inequality by applying Lemma 2.2,

$$\left|\left\langle a_k(t)\cos\theta_k(t), a_l(t)\cos\theta_l(t)\right\rangle\right| < \int_0^T a_k(t)a_l(t) \cdot \left[\frac{\eta_l}{\eta_l^2 - \hat{\eta}_k^2} + 2\varepsilon\left(1 + \frac{1}{(1-\beta^{-|l-k|})^2}\right)\right], \tag{2.13}$$

Since

$$\frac{\eta_l}{\eta_l^2 - \eta_k^2} = \frac{1}{1 - \left(\frac{\eta_k}{\eta_l}\right)^2} \cdot \frac{1}{\eta_l}, \qquad \frac{\eta_k}{\eta_l} \le \frac{1}{\beta^{|l-k|}}, \tag{2.14}$$

we have

$$\left|\left\langle a_k(t)\cos\theta_k(t), a_l(t)\cos\theta_l(t)\right\rangle\right| < \mu_1(k,l)\int_0^T a_k(t)a_l(t)dt, \tag{2.15}$$

where

$$\mu_1(k,l) := \frac{1}{(1-\beta^{-2|l-k|})\eta_l} + 2\varepsilon\left(1 + \frac{1}{(1-\beta^{-|l-k|^2})}\right). \tag{2.16}$$

It follows that

$$\begin{aligned}
\mu_{k,l} &= \frac{\left|\left\langle a_l\cos\theta_l, a_k\cos\theta_k\right\rangle\right|}{\left|\left\langle a_l, a_k\right\rangle\right|} \cdot \frac{\left|\left\langle a_l, a_k\right\rangle\right|}{\left\|a_l\cos\theta_l\right\|\left\|a_k\cos\theta_k\right\|} \\
&\le \frac{\left|\left\langle a_l\cos\theta_l, a_k\cos\theta_k\right\rangle\right|}{\left|\left\langle a_l, a_k\right\rangle\right|} \cdot \frac{\left\|a_k\right\|}{\left\|a_k\cos\theta_k\right\|} \cdot \frac{\left\|a_l\right\|}{\left\|a_l\cos\theta_l\right\|} \\
&< 2\left(1 - 3\varepsilon - \frac{1}{2\eta_k}\right)^{-\frac{1}{2}} \cdot \left(1 - 3\varepsilon - \frac{1}{2\eta_l}\right)^{-\frac{1}{2}} \cdot \left[\frac{1}{(1-\beta^{-2|l-k|})\eta_l} + 2\varepsilon\left(1 + \frac{1}{(1-\beta^{-|l-k|^2})}\right)\right] \\
&< 2\left(1 - 3\varepsilon - \frac{1}{2\eta_1}\right)^{-\frac{1}{2}} \cdot \left[\frac{1}{(1-\beta^{-2|l-k|})\eta_l} + 2\varepsilon\left(1 + \frac{1}{(1-\beta^{-|l-k|^2})}\right)\right] \\
&= O\left(\frac{1}{\eta_l\beta} + \varepsilon\right).
\end{aligned} \tag{2.17}$$

On the other hand, we have

$$\theta'(t) \ge \alpha^{-1}\beta^{l-k}\theta_k^{'}(t), \qquad \forall t \in [0,T] \tag{2.18}$$

by (2.3) and (2.12). Similarly as the estimation of $\mu_{k,l}$, we have

$$\begin{aligned}\gamma_{k,l} &< 2\left(1-3\varepsilon-\frac{1}{2\eta_k}\right)^{-\frac{1}{2}}\cdot\left(1-3\varepsilon-\frac{1}{2\eta}\right)^{-\frac{1}{2}} \\ &\quad\cdot\left[\frac{1}{(1-\alpha^2\beta^{-2|l-k|})\alpha^{-1}\eta_l}+2\varepsilon\left(1+\frac{1}{(1-\alpha\beta^{-|l-k|^2})}\right)\right] \\ &= O\left(\frac{1}{\eta_l\beta}+\varepsilon\right).\end{aligned} \tag{2.19}$$

where

$$\eta=\frac{\theta(T)-\theta(0)}{2\pi}. \tag{2.20}$$

If $l<k$, we have

$$\begin{aligned}&\theta_k^{'}(t)\ge\beta^{k-l}\theta_l^{'}(t),\\ &\theta_k^{'}(t)\ge\alpha^{-1}\beta^{k-l}\theta'(t),\end{aligned}\qquad \forall t\in[0,T], \tag{2.21}$$

and

$$\eta\ge\alpha^{-1}\eta_l. \tag{2.22}$$

So we can also obtain that

$$\begin{aligned}\mu_{k,l} &< 2\left(1-3\varepsilon-\frac{1}{2\eta_1}\right)^{-1}\cdot\left[\frac{1}{(1-\beta^{-2|k-l|})\eta_l\beta^{k-l}}+2\varepsilon\left(1+\frac{1}{(1-\beta^{-|l-k|^2})}\right)\right] \\ &= O\left(\frac{1}{\eta_l\beta}+\varepsilon\right),\end{aligned} \tag{2.23}$$

and

$$\begin{aligned}\gamma_{k,l} &< 2\left(1-3\varepsilon-\frac{1}{2\eta_1}\right)^{-\frac{1}{2}}\cdot\left(1-3\varepsilon-\frac{\alpha}{2\eta_l}\right)^{-\frac{1}{2}} \\ &\quad\cdot\left[\frac{1}{(1-\alpha^2\beta^{-2|k-l|})\eta_l\beta^{k-l}}+2\varepsilon\left(1+\frac{1}{(1-\alpha\beta^{-|k-l|^2})}\right)\right] \\ &= O\left(\frac{1}{\eta_l\beta}+\varepsilon\right).\end{aligned} \tag{2.24}$$

It follows from (1.9), (1.10) and (1.11) that

$$\frac{\|a\cos\theta-a_l\cos\theta_l\|}{\|a_l\cos\theta_l\|}\le O\left(\sqrt{(\eta_l\beta)^{-1}+\varepsilon}\right). \tag{2.25}$$

This completes the proof.

For a well-separated non-integer period sampling signal, we know from estimate (2.24) that, the larger the frequency ratio or the number of periods of the IMF, the smaller the error of the solution of optimization problem (P_2). However, when an IMF has large number of periods in the given time domain, its instantaneous frequency may change greatly so that the inequality $\sup\theta'(t)/\inf\theta'(t)\le M'$ $(t\in[0,T])$ may not be satisfied. In this case, mode mixing might happen (for example, (3.47) in [3]). So in practice, we may divide the whole time domain into small intervals so that η_l in each interval is suitable, not too small or too big.

3. Concluding remarks.

In this paper, we considered the sparse time-frequency decomposition of non-integer period sampling signals. For the non-integer period sampling signals, we analyze the error of looking for a sparse representation by the nonlinear matching pursuit method. By the theoretical proof, we show that the relative error depends on the separation factor, the frequency ratio, and the number of periods of the IMF.

4. Acknowledgements.

This work was supported by the Fundamental Research Funds for the Central Universities (Grant No. 21617417).

Appendix A. Proof of Lemma 2.1.

To proof Lemma 2.1, we need the following estimation.

Lemma A.1. Let $\varepsilon \in (0,1)$, and $g(t) \in C^1[c,d]$ be a positive function. Suppose

$$\left|\frac{g'(t)}{g(t)}\right| \le \varepsilon \qquad \forall t \in [c,d]. \tag{A.1}$$

Then we have

$$\left|\int_c^d g(t)\cos t dt\right| \le \left(2\varepsilon + \frac{1}{\eta}\right)\int_c^d g(t)dt, \tag{A.2}$$

where

$$\eta := \frac{d-c}{2\pi}. \tag{A.3}$$

Proof. First we have that

$$\int_c^d g(t)\cos t dt = \left[\int_c^{c+2\pi} + \int_{c+2\pi}^{c+4\pi} + \cdots + \int_{c+(2[\eta]-1)\pi}^{c+2[\eta]\pi} + \int_{c+2[\eta]\pi}^{d}\right] g(t)\cos t dt, \tag{A.4}$$

Without loss of generality, suppose that

$$\int_{c+2[\eta]\pi}^{d} |g(t)\cos t| dt = \min\left\{\int_{c+2l\pi}^{d-2[\eta]+2l\pi} |g(t)\cos t| dt : 0 \le l \le [\eta]\right\}. \tag{A.5}$$

Thus

$$\begin{aligned}\int_{c+2[\eta]\pi}^{d} |g(t)\cos t| dt &\le \frac{1}{[\eta]+1}\sum_{l=0}^{[\eta]}\left[\int_{c+2l\pi}^{d-2[\eta]\pi+2l\pi} |g(t)\cos t| dt\right] \\ &= \frac{1}{[\eta]+1}\int_c^d g(t)\,|\cos t|\,dt \\ &\le \frac{1}{\eta}\int_c^d g(t)dt.\end{aligned} \tag{A.6}$$

On the other hand, for any integer $m \in [1,\eta]$,

$$
\begin{aligned}
&\left|\int_{c+2(m-1)\pi}^{c+2m\pi} g(t)\cos t dt\right| \\
&=\left|\int_{c+2(m-1)\pi}^{c+(2m-1)\pi} g(t)\cos t dt+\int_{c+(2m-1)\pi}^{c+2m\pi} g(t)\cos t dt\right| \\
&=\left|\int_{c+2(m-1)\pi}^{c+(2m-1)\pi} g(t)\cos t dt+\int_{c+2(m-1)\pi}^{c+(2m-1)\pi} g(t+\pi)\cos(t+\pi) dt\right| \\
&=\left|\int_{c+2(m-1)\pi}^{c+(2m-1)\pi} \left|g(t)-g(t+\pi)\right|\cos t dt\right| \\
&\le \int_{c+2(m-1)\pi}^{c+(2m-1)\pi} \left|\cos t\right| dt\int_{t}^{t+\pi}\left|g'(s)\right| ds \\
&\le \int_{c+2(m-1)\pi}^{c+(2m-1)\pi} \left|\cos t\right| dt\int_{c+2(m-1)\pi}^{c+2m\pi}\left|g'(s)\right| ds \\
&= 2\int_{c+2(m-1)\pi}^{c+2m\pi}\left|g'(s)\right| ds \\
&\le 2\varepsilon\int_{c+2(m-1)\pi}^{c+2m\pi} g(s) ds.
\end{aligned} \tag{A.7}
$$

Thus

$$
\left|\left[\int_{c}^{c+2\pi}+\int_{c+2\pi}^{c+4\pi}+\cdots+\int_{c+2([\eta]-1)\pi}^{c+2[\eta]\pi}\right] g(t)\cos t dt\right| \le 2\varepsilon\int_{c}^{c+2[\eta]\pi} g(t)dt. \tag{A.8}
$$

This and (A.6) imply (A.2).

Now, we can give the proof of Lemma 2,1.

Proof of Lemma 2.1. First, we know

$$
\begin{aligned}
\left\|a(t)\cos\theta(t)\right\|^2 &= \int_0^T a^2(t)[\cos\theta(t)]^2 dt = \frac{1}{2}\int_0^T a^2(t)[\cos 2\theta(t)+1]dt \\
&= \frac{1}{2}\left[\left\|a(t)\right\|^2+\int_0^T a^2(t)\cos 2\theta(t) dt\right].
\end{aligned} \tag{A.9}
$$

Let $s=2\theta(t)$. Then we obtain

$$
\int_0^T a^2(t)\cos 2\theta(t)dt = \frac{1}{2}\int_{2\theta(0)}^{2\theta(T)} g(s)\cos s ds, \tag{A.10}
$$

where $t(s):=\theta^{-1}(s/2)$ and (noticing that $t'(s)=1/(2\theta'(t(s)))$)

$$
g(s)=\frac{a^2(t(s))}{\theta'(t(s))}. \tag{A.11}
$$

So the derivative of g is

$$
g'(s)=\frac{a(t(s))a'(t(s))-a^2(t(s))\cdot\theta''(t(s))t'(s)}{\left[\theta'(t(s))\right]^2}. \tag{A.12}
$$

Then we obatin

$$
\left|\frac{g'(s)}{g(s)}\right| = \left|\frac{a'(t(s))}{a(t(s))\theta'(t(s))}-\frac{\theta''(t(s))}{2\left[\theta'(t(s))\right]^2}\right| \le \frac{3}{2}\varepsilon, \qquad \forall s\in[2\theta(0),2\theta(T)]. \tag{A.13}
$$

Using Lemma A.1, we get

$$
\begin{aligned}
\left|\int_0^T a^2(t)\cos 2\theta(t)dt\right| &< \frac{1}{2}\left(3\varepsilon+\frac{1}{2\eta}\right)\int_{2\theta(0)}^{2\theta(T)} g(s)ds. \\
&= \left(3\varepsilon+\frac{1}{2\eta}\right)\cdot\left\|a(t)\right\|^2.
\end{aligned} \tag{A.14}
$$

Thus (2.6) follows by combining (A.9) and (A.14).

Appendix B. Proof of Lemma 2.2.

Proof of Lemma 2.2. We have the following equality:

$$\begin{aligned}&\left\langle a(t)\cos\theta(t), \hat{a}(t)\cos\hat{\theta}(t)\right\rangle \\ &= \frac{1}{2}\left[\int_0^T a(t)\hat{a}(t)\cos(\theta(t)-\hat{\theta}(t))dt + \int_0^T a(t)\hat{a}(t)\cos(\theta(t)+\hat{\theta}(t))dt\right] \\ &= \frac{1}{2}[I_1+I_2].\end{aligned} \tag{B.1}$$

For the integral I_1, let $s=\theta(t)-\hat{\theta}(t)$. Then

$$I_1 = \int_{\theta(0)-\hat{\theta}(0)}^{\theta(T)-\hat{\theta}(T)} g(s)\cos s ds, \tag{B.2}$$

where $t(s)=(\theta-\hat{\theta})^{-1}(s)$ and (noting that $t'(s)=1/(\theta'(t(s))-\hat{\theta}'(t(s)))$)

$$g(s) = \frac{a(t(s))\hat{a}(t(s))}{\theta'(t(s))-\hat{\theta}'(t(s))}. \tag{B.3}$$

Thus, we have

$$g'(s) = \frac{(\theta'-\hat{\theta}')(a'\hat{a}+a\hat{a}')-a\hat{a}(\theta''-\hat{\theta}'')}{\left(\theta'(t(s))-\hat{\theta}'(t(s))\right)^3}, \tag{B.4}$$

and

$$\left|\frac{g'(s)}{g(s)}\right| \le \frac{2\varepsilon}{(1-d^{-1})^2}, \qquad \forall s\in[\theta(0)-\hat{\theta}(0), \theta(T)-\hat{\theta}(T)]. \tag{B.5}$$

Using Lemma A.1, we get

$$|I_1| \le \left(\frac{4\varepsilon}{(1-d^{-1})^2}+\frac{1}{\eta-\hat{\eta}}\right)\int_0^T a(t)\hat{a}(t)dt. \tag{B.6}$$

Similarly, we can obtain that

$$|I_2| \le \left(4\varepsilon+\frac{1}{\eta+\hat{\eta}}\right)\int_0^T a(t)\hat{a}(t)dt. \tag{B.7}$$

Combine (B.1), (B.6) and (B.7), we complete the proof.

References

[1] A. Venkitaraman, S. Chatterjee, P. Händel, On Hilbert Transform, Analytic Signal, and Modulation Analysis for Signals over Graphs, CoRR, abs/1611.05269, 2016.
[2] B. Van der Pol, The fundamental principles of frequency modulation, Proc. IEE, 93 (1946), 153-158.
[3] C. G. Liu, Z. Q. Shi, T. Y. Hou. On the uniqueness of sparse time-frequency representation of multiscale data, Multiscale Model. Simul. 13(3) (2015), 790 – 811.
[4] C. G. Liu, Z. Q. Shi, T. Y. Hou. A two-level method for sparse time-frequency representa- tion of multiscale data, Science China Mathematics. 60(10) (2017), 1377-1752.
[5] D. A. Cummings, R. A. Irizarry, N. E. Huang, T. P. Endy, A. Nisalak, K. Ungchusak, and D. S. Burke, Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand, Nature 427 (2004), 344-347.
[6] D. L. Donoho. Compressed Sensing. IEEE Transactions on Information Theory, 52(4) (2006), 1289-1306.
[7] D. Gabor, Theory of communication, J. IEE. 93 (1946), 426-457.
[8] E. J. Cand ‘es, J. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inform. Theory, 52 (2006), 489 – 509.
[9] E. J. Cand ‘es and T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies, IEEE Trans. Inform. Theory, 52 (2006), 5406 – 5425.
[10] I. Daubechies, J. Lu, and H. Wu, Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool, Appl. Comput. Harmonic Anal., 30 (2011), 243 – 261.
[11] J. Justice, Analytic signal processing in music computation. IEEE Transactions on Acoustics, Speech, and Signal Processing. 27(6) (1979), 670 – 684.
[12] K. Dragomiretskiy and D. Zosso, Variational mode decomposition, IEEE Trans. Signal Process. ,62 (2014), 531 – 544.

[13] 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, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci., 454 (1998), 903 – 995.
[14] 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. A 454 (1998), 903-995.
[15] P. Flandrin, Time-Frequency/Time-Scale Analysis, Academic Press, San Diego, (1999).
[16] Potamianos, Alexandros and Petros Maragos. A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation. Signal Processing 37 (1994), 95-120.
[17] S. Mallat, A Wavelet Tour of Signal Processing: The Sparse Way, 3rd ed., Elsevier/Academic Press, Amsterdam, (2009).
[18] S. Mallat and Z. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Trans. Signal Process, 41 (1993), 3397-3415.
[19] S. Qian and D. Chen. Signal representation using adaptive normalized Gaussian functions. Signal Processing, 36 (1994), 1-11.
[20] S. Chen, D. Donoho and M. Saunders, Atomic decomposition by basis pursuit, SIAM J. Sci. Comput., 20 (1998), 33-61.
[21] T. Y. Hou and Z. Shi Adaptive data analysis via sparse time-frequency representation. Adv Adapt Data Anal, 3 (2011), 1 – 28
[22] T. Y. Hou and Z. Shi, Data-driven time-frequency analysis, Appl. Comput. Harmon. Anal., 35 (2013), 284 – 308.
[23] T. Y. Hou, Z. Shi, and P. Tavallali, Convergence of a data-driven time-frequency analysis method, Appl. Comput. Harmon. Anal., 37 (2014), 235 – 270.
[24] T. Y. Hou and Z. Shi, Sparse time-frequency representation of nonlinear and nonstationary data, Science China Mathematics. 56(12) (2013), 2489-2506.
[25] Z. Wu and N. E. Huang, Ensemble empirical mode decomposition: A noise assisted data analysis method, Adv. Adapt. Data Anal., 1 (2008), 1 – 41.
[26] Z. Wu, N. E. Huang, X. Chen. The multi-dimensional ensemble empirical mode decomposition method. Adv Adapt Data Anal, 1 (2009), 339 – 372.
[27] Z. Wu, N. E. Huang, S. R. Long, et al. On the trend, detrending, and variability of nonlinear and nonstationary time series. Proc Natl Acad Sci USA, 104 (2007), 14889 – 14894.