Multiscale Model. Simul.
Vol. 12, No. 4, pp. 1458–1493
© 2014 Society for Industrial and Applied Mathematics

EXTRACTION OF INTRAWAVE SIGNALS USING THE SPARSE TIME-FREQUENCY REPRESENTATION METHOD*

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

Abstract. Analysis and extraction of strongly frequency modulated signals have been a challenging problem for adaptive data analysis methods, e.g., empirical mode decomposition [N.E. Huang et al., *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 454 (1998), pp. 903–995]. In fact, many of the Newtonian dynamical systems, including conservative mechanical systems, are sources of signals with low to strong levels of frequency modulation. Analysis of such signals is an important issue in system identification problems. In this paper, we present a novel method to accurately extract intrawave signals. This method is a descendant of sparse time-frequency representation methods [T.Y. Hou and Z. Shi, *Appl. Comput. Harmon. Anal.*, 35 (2013), pp. 284–308, T.Y. Hou and Z. Shi, *Adv. Adapt. Data Anal.*, 3 (2011), pp. 1–28]. We will present numerical examples to show the performance of this new algorithm. Theoretical analysis of convergence of the algorithm is also presented as a support for the method. We will show that the algorithm is stable to noise perturbation as well.

Key words. sparse time-frequency representation, instantaneous frequency, intrawave frequency modulation

AMS subject classifications. 94A12, 65T99

DOI. 10.1137/140957767

1. Introduction.

1.1. Adaptive data analysis methods. Signals, in general, are gateways of understanding the world around us. It is not an exaggeration to say that all we know about the materialistic world comes from signals. More specifically, oscillatory signals are prevalent in physical and biomedical recordings and many other fields. Among these signals, many of them have features of nonlinearity and nonstationary. For this kind of signal, in many cases, traditional methods, such as the Fourier transform or the wavelet method [15, 1], are not enough to extract the intrinsic structure beneath the signal, since they use a predetermined basis which may not fit the hidden structure of the signal. In order to analyze and understand these signals, one needs apt adaptive data analysis methods.

Substantial progress has been made by the introduction of the empirical mode decomposition (EMD) method in recent years [11, 12, 13, 14]. EMD is an effective method of decomposing signals into a collection of oscillatory signals with smoothly changing envelopes and frequency of the form

$$x(t) = a(t)\cos\theta(t), \tag{1.1}$$

where $a(t) > 0$, and the phase function $\theta(t)$ is a strictly increasing function of time. We call this entity an intrinsic mode function (IMF). Physically speaking, $\theta(t)$ carries

*Received by the editors February 19, 2014; accepted for publication (in revised form) July 10, 2014; published electronically October 23, 2014. This research was supported in part by AFOSR MURI grant FA9550-09-1-0613, DOE grant DE-FG02-06ER25727, and NSF grants DMS-1159138 and DMS-1318377.
http://www.siam.org/journals/mms/12-4/95776.html
†Department of Applied and Computational Mathematics, MC 9-94, Caltech, Pasadena, CA 91125 (tavallali@gmail.com, hou@cms.caltech.edu).
‡Mathematical Sciences Center, Tsinghua University, Beijing, China, 100084 (zqshi@math.tsinghua.edu.cn). This author's research was supported in part by NSFC grant 11201257.

information about the rate of change of the signal in time. The time derivative of $\theta(t)$ is called the instantaneous frequency (IF). A finite linear combination of a collection of IMFs is called an intrinsic signal (IS),

$$s(t) = \sum_{i=1}^{M} a_i(t) \cos \theta_i(t). \tag{1.2}$$

The drawback of the EMD method is its empirical nature. Consequently, there have been attempts to construct methods like EMD that have a solid mathematical foundation. These methods include, but are not confined to, the optimization-based EMD method [10], sparse time-frequency representation (STFR) methods [8, 7], and synchrosqueezed wavelet transforms [6]. All of these methods can be used to extract IMFs. Among all of these methods, the STFR method preserves many important properties of the original EMD method, especially when it comes to the adaptive nature of the method.

Signals that can be represented by (1.1) and conditions $a(t) > 0$, $\frac{d\theta(t)}{dt} > 0$ cover a large set of signals. Hence, in early versions of the STFR methods, the assumption that the envelope of each IMF is less oscillatory than the oscillatory part $\cos\theta$ was used. More precisely, in the original STFR method, we chose $a(t) \in V(\theta(t))$ and $a(t) > 0$, where

$$V(\theta) = span\left\{1, \cos\left(\frac{\theta}{\lambda}\right), \sin\left(\frac{\theta}{\lambda}\right) \mid \lambda \geqslant 2\right\}. \tag{1.3}$$

This form of dictionary $V(\theta)$ would impose a certain smoothness constraint on the envelope $a(t)$: The envelope is smoother than $\cos\theta$ in the sense that the highest frequency terms in $a(t)$ are $\cos\left(\frac{\theta}{2}\right)$, $\sin\left(\frac{\theta}{2}\right)$. In Fourier terms, if the Fourier transform is performed in the phase, θ, domain (not in the time, t, domain), then the envelope of an IMF is centered around frequency 1 having a support in $\left[\frac{1}{2}, \frac{3}{2}\right]$.

All of the adaptive data analysis methods mentioned above, either empirical or mathematical, perform similarly when it comes to IMFs with smoothly changing IF and envelopes; i.e., signals having envelopes from the standard dictionary $V(\theta)$. However, when it comes to signals with strong intrawave frequency modulation, some of these methods have difficulty in extracting the IMFs accurately, especially when the data are polluted by noise. Intrawave modulation, in many cases, causes the envelope $a(t)$ to be not necessarily smoother than $\cos\theta$. The highest frequency terms in $a(t)$ are no longer $\cos\left(\frac{\theta}{2}\right)$, $\sin\left(\frac{\theta}{2}\right)$. This difficulty can be alleviated somewhat by the introduction of a shape function [19]. However, finding the shape function is a challenge by itself. In what follows, we first introduce the reader to the concept of intrawave signals and then move on to show that a slightly modified version of the STFR method can handle the analysis of intrawave signals.

1.2. Intrawave signals. The main goal of this paper is to extract an intrawave signal in which the frequency has strong intrawave frequency modulation. Generally speaking, intrawave signals are of particular interest. First, these signals are abundant in second order ordinary differential equations, specifically conservative systems. Some of these systems produce oscillatory signals with extremely fast-changing IF near the peak and/or troughs of the solution signal, i.e., signals with sharp peaks and/or troughs. The latter is a good indicator for spotting an intrawave signal by visual inspection. Second, from a theoretical point of view, extraction of intrawave

signals is challenging. To the best of our knowledge, most existing adaptive data analysis methods have difficulty in analyzing such signals with good accuracy, especially when the signal is polluted by noise.

To motivate the difficulty in extracting IMFs from an intrawave signal and also introduce the reader to intrawave signals, we consider the following signal:

$$x(t) = \cos\left(\omega t + \frac{\triangle\,\omega}{p}\sin(pt)\right). \tag{1.4}$$

In (1.4), ω is the carrier frequency, $\triangle\,\omega < \omega$ is the strength of frequency modulation, and p is the frequency of the frequency modulation. This example is inspired by the work on frequency modulated (FM) signals by van der Pol [18]. Taking $\theta(t) = \omega t$, it is possible to express (1.4) in the phase domain $\theta(t)$,

$$x(\theta) = \cos\left(\theta + \frac{\triangle\,\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right). \tag{1.5}$$

Simplifying the latter, we have

$$x(\theta) = \cos\theta\cos\left(\frac{\triangle\,\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right) - \sin\theta\sin\left(\frac{\triangle\,\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right). \tag{1.6}$$

Expanding the terms $\cos\left(\frac{\Delta\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right)$ and $\sin\left(\frac{\Delta\omega}{p}\sin\left(\frac{p}{\omega}\theta\right)\right)$, using the Fourier transform, we get

$$\begin{aligned} x(\theta) = &\left\{J_0\left(\tfrac{\Delta\omega}{p}\right) + 2\textstyle\sum_{k=1}^{\infty} J_{2k}\left(\tfrac{\Delta\omega}{p}\right)\cos\left(2k\tfrac{p}{\omega}\theta\right)\right\}\cos\theta \\ &+\left\{2\textstyle\sum_{k=1}^{\infty} J_{2k-1}\left(\tfrac{\Delta\omega}{p}\right)\sin\left((2k-1)\tfrac{p}{\omega}\theta\right)\right\}\sin\theta. \end{aligned} \tag{1.7}$$

In (1.7), J_k is the the Bessel function of the first kind of order k. Finally, the original signal (1.4) can be expressed in terms of $x(\theta) = a(\theta)\cos\theta + b(\theta)\sin\theta$, where

$$a(\theta) = \left\{J_0\left(\frac{\triangle\,\omega}{p}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega}{p}\right)\cos\left(2k\frac{p}{\omega}\theta\right)\right\}, \tag{1.8}$$

$$b(\theta) = -\left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega}{p}\right)\sin\left((2k-1)\frac{p}{\omega}\theta\right)\right\}. \tag{1.9}$$

Obviously, the envelope of the signal in (1.7) is not smoother than $\cos\left(\frac{\theta}{2}\right)$, $\sin\left(\frac{\theta}{2}\right)$. For a fixed $\frac{\Delta\omega}{p}$, as $k \to \infty$, $J_k\left(\frac{\Delta\omega}{p}\right)$ decreases. However, if $\frac{\Delta\omega}{p}$ is large enough, many of the high-frequency terms in this formula would make a considerable tail in the Fourier domain if the Fourier transform were performed in the phase space θ. This thick tail is the main source of difficulty in the extraction of an intrawave IMF. The specific difficulty varies for different adaptive methods. For example, in STFR methods, an explicit band filter in the phase space is used at each iteration [8]—a short band filter that can solely collect terms in $a(\theta)$ and $b(\theta)$ that are smoother than $\cos\left(\frac{\theta}{2}\right)$, $\sin\left(\frac{\theta}{2}\right)$. This band filter is essentially the numerical counter part of (1.3). The band filter mentioned above cannot extract the terms that construct the tail of the expansion in the Fourier domain if the initial guess of the phase function is given by $\theta(t) = \omega t$. On the other hand, it has been demonstrated in [20] that the ensemble empirical mode decomposition (EEMD) method has properties similar to an

implicit band filter. Although the filter is implicit, it still acts like a short band filter. Except for some special data, EMD or EEMD in general has difficulty in capturing the significant tail of the intrawave in the Fourier domain. In short, by using a short band filter to decompose a wide band signal, the EMD/EEMD and the original STFR methods decompose the signal into subharmonics. In what follows, we investigate this issue in more detail.

We will see, in the next section, that at each iteration of the STFR method, with θ as an initial guess, the algorithm chooses terms that are in $V(\theta)$ to extract the envelope. Hence, having $\theta(t) = \omega t$ in (1.4), the maximum number of terms that can be collected by the STFR method for the envelope $a(\theta)$, based on dictionary (1.3) and (1.7), is $k_{max} = \lfloor \frac{\omega}{4p} \rfloor$. For $b(\theta)$, this value is $k_{max} = \lfloor \frac{\omega}{4p} + \frac{1}{2} \rfloor$. In other words, the approximated envelopes $\tilde{a}(\theta)$, $\tilde{b}(\theta)$, at an iteration having the initial guess as $\theta(t) = \omega t$, would look like

$$\tilde{a}(\theta) = \left\{ J_0\left(\frac{\triangle\,\omega}{p}\right) + 2\sum_{k=1}^{\lfloor \frac{\omega}{4p} \rfloor} J_{2k}\left(\frac{\triangle\,\omega}{p}\right)\cos\left(2k\frac{p}{\omega}\theta\right)\right\}, \tag{1.10}$$

$$\tilde{b}(\theta) = -\left\{ 2\sum_{k=1}^{\lfloor \frac{\omega}{4p}+\frac{1}{2} \rfloor} J_{2k-1}\left(\frac{\triangle\,\omega}{p}\right)\sin\left((2k-1)\frac{p}{\omega}\theta\right)\right\}. \tag{1.11}$$

Obviously, if $\frac{\omega}{4p} \ll 1$, the only term that will be collected by the STFR method is $J_0\left(\frac{\triangle\omega}{p}\right)$. The other terms would all be discarded. The consequence of this incomplete extraction is that the STFR method will break an intrawave signal into many subharmonics.

In order to capture more terms of $a(\theta)$ and $b(\theta)$, in the STFR algorithms, we need to modify the dictionary (1.3), or more precisely the filter, as follows:

$$V(\theta) = \left\{1, \cos\left(\frac{\theta}{\lambda}\right), \sin\left(\frac{\theta}{\lambda}\right) \mid \lambda \geqslant \lambda_0 > 0\right\}. \tag{1.12}$$

Here, λ_0 should be as small as possible such that enough terms in $a(\theta)$ and $b(\theta)$ can be collected. In this case, the maximum number of terms that can be collected by the STFR method, based on dictionary (1.12), is $k_{max} = \lfloor \frac{\omega}{4p\lambda_0} \rfloor$ for $a(\theta)$. For $b(\theta)$, this value is $k_{max} = \lfloor \frac{\omega}{4p\lambda_0} + \frac{1}{2} \rfloor$. Obviously, using the new filter (dictionary) (1.12), the algorithm is now able to collect more terms. Consequently, there will be fewer subharmonics in the residue and a more precise analysis.

The choice of λ_0 depends on the decay rate of $J_k\left(\frac{\triangle\omega}{p}\right)$ and the ratio $\frac{\omega}{p}$. However, if the original signal is polluted with noise, this parameter should be chosen carefully. Now it is time to have a formal definition of an intrawave signal.

Definition 1.1 (intrawave signal). *An IMF $x(t) = a(t)\cos\theta(t)$, having only one extremum between two consequent zeros, is an IS in a phase coordinate θ if the envelope $a > 0$ in the θ-coordinate, and when decomposed by the Fourier transform, has terms like* $\cos\nu\theta$ *and* $\sin\nu\theta$ *for* $\nu \geq \frac{1}{2}$.

This definition clearly covers the sample intrawave that we picked earlier. In (1.4) and (1.7), the envelope of the intrawave in the θ-coordinate system with $\theta = \omega t$ has terms that are more oscillatory compared to $\cos\left(\frac{\theta}{2}\right)$, $\sin\left(\frac{\theta}{2}\right)$. Now we can present a summary of the problem that we want to solve, as well as the road map to the solution.

1.3. Summary. The main difficulty in dealing with intrawave signals comes from the fact that they have a wide band representation in the frequency domain. Hence, any time-frequency method that uses short band filters, either explicitly or implicitly, would not be able to decompose such signals properly. In this paper, we will show that by modifying the envelope dictionary (1.3), in an STFR framework, the intrawave signals can be extracted with reasonable accuracy. The small modification would allow us to treat intrawave signals without major changes in the original STFR algorithm. The convergence of the modified algorithm will be proven for the case of periodic signals, with or without noise.

We remark that by enlarging the filter band width, we require that the IMF components of an IS be well separated in the time-frequency domain. In our technical terms, they must have large-scale separation. In fact, a wide band filter might collect parts of another IMF. This is why we require that they have large-scale separation. However, this limitation can be alleviated by extracting the nonseparable time-frequency intrawave IMFs simultaneously. We will report this method in detail in a future work. In this paper, we will only demonstrate this idea in one of the numerical examples. Finally, we remark that the method of proof in this paper is similar in spirit to that reported in our previous paper [9].

The rest of the paper is organized as follows. In section 2, we present the STFR algorithms as well as their extensions, which can analyze intrawave signals. The theorems shown in section 3 will then explain why such algorithms would work. In section 4, we present several numerical examples to demonstrate the effectiveness of our algorithms and provide support for our convergence theorems. Some concluding remarks are made in section 5. Several technical estimates are deferred to Appendices A, B, and C.

2. Algorithmic analysis. In this section, we present the STFR methods and algorithms. The STFR method was inspired by the EMD/EEMD method [13, 21], recently developed compressive sensing theory [4, 5], and matching pursuit [16]. The main idea in the STFR method is to decompose a signal into the sparsest collection of IMFs. The set of the collection of IMFs constitutes a dictionary.

Definition 2.1 (IMF dictionary). *The IMF* dictionary $\mathcal{D}$ *is defined as*

$$\mathcal{D} = \left\{ a(t)\cos\theta(t) \mid a(t) \in V(\theta(t)), a(t) > 0, \frac{d\theta(t)}{dt} > 0 \right\}. \tag{2.1}$$

In mathematical terms, finding the sparsest decomposition of a signal can be formulated as the following nonlinear optimization problem:

$$\begin{array}{ll} \text{Minimize} & M \\ \text{subject to} & s(t) = \sum_{i=1}^{M} a_i(t)\cos\theta_i(t), \ a_i(t)\cos\theta_i(t) \in D, \ i = 1, \ldots, M. \end{array} \tag{2.2}$$

This problem is an L_0 minimization. The constraint $s(t) = \sum_{i=1}^{M} a_i(t)\cos\theta_i(t)$ can be relaxed into an inequality if noise is present in the data. Solving this problem is fundamentally difficult. Hence, one should try to find an appropriate approximation of this problem.

There are several ways to approximate the above L_0 minimization problem by a more tractable algorithm; see [8, 7]. The main idea behind all of the simplifications is to decompose the signal into two parts, a mean and a modulated oscillatory part:

$$s = a_0 + a_1\cos\theta. \tag{2.3}$$

Here, θ, a_1, and a_0 are unknown phase function, envelope, and mean, respectively. This sequential decomposition is nothing but a matching pursuit [16] to find one of the IMFs $a_1 \cos\theta$. After the extraction of $a_1 \cos\theta$, the residue (mean) a_0 is treated as a new signal, and the same procedure is repeated until the final residue is small enough.

Here we review the periodic STFR algorithm. Without loss of generality, we assume that the signal is sampled over a uniform grid $t_j = \frac{j}{N}$ for $j = 0, 1, \ldots, N-1$ and even N.

- $k = 1$, $r^{k-1} = s$.
- Step 0: $n = 0$, $\theta_k^n = \theta_0$.
- Step 1: Interpolate r^{k-1} from the t-coordinate into a uniform grid in θ-coordinate. In other words, for $j = 0, \ldots, N-1$,

$$\left(r_{\theta_k^n}^{k-1}\right)_j = \text{Interpolate}\left(t, r^{k-1}, \left(\theta_k^n\right)_j\right). \tag{2.4}$$

- Step 2: Apply the Fourier transform to $r_{\theta_k^n}^{k-1}$ for $\left(\theta_k^n\right)_j = 2\pi L_{\theta_k^n} \frac{j}{N}$:

$$\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_\omega = \frac{1}{N}\sum_{j=1}^{N}\left(r_{\theta_k^n}^{k-1}\right)_j e^{-i\omega\frac{\left(\theta_k^n\right)_j}{L_{\theta_k^n}}},\ \omega = -\frac{N}{2}+1, \ldots, \frac{N}{2}. \tag{2.5}$$

- Step 3: For the Fourier inverse defined as

$$\left(\mathcal{F}_{\theta_k^n}^{-1}\left(\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)\right)\right)_j = \frac{1}{N}\sum_{\omega=-\frac{N}{2}+1}^{\frac{N}{2}} \mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_\omega e^{i\omega\frac{\left(\theta_k^n\right)_j}{L_{\theta_k^n}}},\ j = 0, \ldots, N-1, \tag{2.6}$$

apply a cutoff function χ to $\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)$ to calculate the envelopes

$$a_{\theta_k^n}^k = \mathcal{F}_{\theta_k^n}^{-1}\left[\left(\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega+L_{\theta_k^n}} + \mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega-L_{\theta_k^n}}\right)\chi\left(\frac{\omega}{L_{\theta_k^n}}\right)\right], \tag{2.7}$$

$$b_{\theta_k^n}^k = \mathcal{F}_{\theta_k^n}^{-1}\left[-i\left(\mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega+L_{\theta_k^n}} - \mathcal{F}_{\theta_k^n}\left(r_{\theta_k^n}^{k-1}\right)_{\omega-L_{\theta_k^n}}\right)\chi\left(\frac{\omega}{L_{\theta_k^n}}\right)\right]. \tag{2.8}$$

The cutoff function is

$$\chi(\omega) = \begin{cases} 1, & -\frac{1}{2} < \omega < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

- Step 4: Interpolate the calculated envelopes' background into the t-coordinate for $j = 0, \ldots, N-1$:

$$\left(a_n^k\right)_j = \text{Interpolate}\left(\theta_k^n, a_{\theta_k^n}^k, t_j\right), \tag{2.10}$$

$$\left(b_n^k\right)_j = \text{Interpolate}\left(\theta_k^n, b_{\theta_k^n}^k, t_j\right). \tag{2.11}$$

- Step 5: Update the phase function

$$\triangle\theta' = P_{V_{M_0}}\left(\frac{d}{dt}\arctan\left(\frac{b_n^k}{a_n^k}\right)\right), \triangle\theta(t) = \int_0^t \triangle\theta'(s)\,ds, \tag{2.12}$$

$$\theta_k^{n+1} = \theta_k^n - \beta\triangle\theta. \tag{2.13}$$

Here, $\beta \in [0,1]$ is chosen in a way that the phase function is always strictly increasing:

$$\beta = \max\left\{\alpha \in [0,1] \mid \frac{d}{dt}\left(\theta_k^n - \alpha\triangle\theta\right) > 0\right\}. \tag{2.14}$$

The projection, $P_{V_{M_0}}$, into $V_{M_0} = span\left\{e^{i2\pi kt}, |k| \le M_0, k \in \mathbb{Z}\right\}$ is a smoothing step. M_0 is fixed.

- Step 6: If $\left\|\theta_k^{n+1} - \theta_k^n\right\| \le \epsilon_I$, then $IMF_k = a_n^k \cos\theta_k^n + b_n^k \sin\theta_k^n$, $n \leftarrow 0$, $r^k = r^{k-1} - IMF_k$, $k \leftarrow k+1$. Else, $n \leftarrow n+1$, goto Step 1.
- Step 7: If $\left\|r^k\right\| \le \epsilon_{II}$, stop. Else, goto Step 0.

There is also a similar nonperiodic algorithm. This algorithm is based on solving the discrete form of the following minimization problem:

$$\begin{aligned} &\text{Minimize} \quad \delta\left(\|a(t)\|_1 + \|b(t)\|_1\right) + \|s(t) - A(t)\cos\theta(t)\|_2^2 \\ &\text{subject to} \quad A(t) = \textstyle\int_2^\infty (a(\lambda)\cos\frac{\theta(t)}{\lambda} + b(\lambda)\sin\frac{\theta(t)}{\lambda})d\lambda, \\ &\qquad\qquad \tfrac{d\theta(t)}{dt} > 0. \end{aligned} \tag{2.15}$$

Here, $A(t) = \int_2^\infty (a(\lambda)\cos\frac{\theta(t)}{\lambda} + b(\lambda)\sin\frac{\theta(t)}{\lambda})d\lambda$ means that the envelope $A(t)$ must be represented by terms in $V(\theta)$. The coefficients of that representation are $a(\lambda)$ and $b(\lambda)$.

In order to solve such an optimization problem in the discrete case, we need the following:

$$A(\theta) = \bar{a}^c + \sum_l a_{k_l}^c \cos\frac{\theta}{k_l} + b_{k_l}^c \sin\frac{\theta}{k_l}, \tag{2.16}$$

$$\{k_l\}_{l=1}^m \subset [2,\infty). \tag{2.17}$$

We need the same thing for $B(\theta)$:

$$B(\theta) = \bar{a}^s + \sum_l a_{k_l}^s \cos\frac{\theta}{k_l} + b_{k_l}^s \sin\frac{\theta}{k_l}. \tag{2.18}$$

In fact, $A(\theta)$ and $B(\theta)$ are approximations of the envelope of the IMF. In other words, we assume that the IMF looks like $A(\theta)\cos\theta + B(\theta)\sin\theta$. In this assumption, we set $\theta(0) = 0$. In the discrete domain, for $i = 1, \ldots, N$,

$$\boldsymbol{\theta} = \{\theta_i\}, \ \mathbf{C} = \left\{\cos\tfrac{\theta_i}{k_l}\right\} = \{C_{il}\}, \ \mathbf{S} = \left\{\sin\tfrac{\theta_i}{k_l}\right\} = \{S_{il}\}. \tag{2.19}$$

Now, define vectors $\mathbf{A}$, $\mathbf{B}$:

$$\mathbf{A} = [\mathbf{C}, \mathbf{S}, \mathbf{1}] \begin{bmatrix} \mathbf{a}^c \\ \mathbf{b}^c \\ \bar{\mathbf{a}}^c \end{bmatrix}, \ \mathbf{B} = [\mathbf{C}, \mathbf{S}, \mathbf{1}] \begin{bmatrix} \mathbf{a}^s \\ \mathbf{b}^s \\ \bar{\mathbf{a}}^s \end{bmatrix}. \tag{2.20}$$

We still need a few more matrices to be able to describe the algorithm:

$$\begin{aligned} \Psi &= [\mathbf{C}, \mathbf{S}, \mathbf{1}], & \mathbf{sin} &= \operatorname{diag}(\sin\theta), \\ \mathbf{x} &= \left[\, \mathbf{a}^c \ \ \mathbf{b}^c \ \ \bar{\mathbf{a}}^c \ \ \mathbf{a}^s \ \ \mathbf{b}^s \ \ \bar{\mathbf{a}}^s \,\right]^T, & \mathbf{cos} &= \operatorname{diag}(\cos\theta), \end{aligned} \tag{2.21}$$

$$\mathbf{H} = [\mathbf{cos}\,\Psi, \mathbf{sin}\,\Psi]. \tag{2.22}$$

We are now ready to explain how the nonperiodic STFR algorithm works. Take the sampled signal to be like a column vector $\mathbf{u}$. The STFR algorithm would then be the following.

- $k = 1$, $\mathbf{r}^{k-1} = \mathbf{u}$.
- Step 0: $n = 0$, $\boldsymbol{\theta}_k^{(n)} = \boldsymbol{\theta}_0$.
- Step 1: Find $\mathbf{x}_k^{(n)} = \left(\left[\begin{array}{ccccccc} \mathbf{a}^c & \mathbf{b}^c & \bar{\mathbf{a}}^c & \mathbf{a}^s & \mathbf{b}^s & \bar{\mathbf{a}}^s \end{array}\right]^T\right)_k^{(n)}$ by solving the following convex minimization problem:

$$\text{Minimize } \delta \left\|\mathbf{x}_k^{(n)}\right\|_1 + \left\|\mathbf{s} - \mathbf{H}_k^{(n)}\mathbf{x}_k^{(n)}\right\|_2^2. \tag{2.23}$$

- Step 2: Find

$$\mathbf{A}_k^n = [\mathbf{C}, \mathbf{S}, \mathbf{1}]_k^{(n)} \left(\left[\begin{array}{c} \mathbf{a}^c \\ \mathbf{b}^c \\ \bar{\mathbf{a}}^c \end{array}\right]\right)_k^{(n)}, \quad \mathbf{B}_k^n = [\mathbf{C}, \mathbf{S}, \mathbf{1}]_k^{(n)} \left(\left[\begin{array}{c} \mathbf{a}^s \\ \mathbf{b}^s \\ \bar{\mathbf{a}}^s \end{array}\right]\right)_k^{(n)}. \tag{2.24}$$

- Step 3: Update the phase vector[1]

$$\boldsymbol{\theta}_k^{(n+1)} = \boldsymbol{\theta}_k^{(n)} - \arctan\left(\frac{\mathbf{B}_k^n}{\mathbf{A}_k^n}\right). \tag{2.25}$$

- Step 4: If $\left\|\boldsymbol{\theta}^{(n)} - \boldsymbol{\theta}^{(n-1)}\right\| \leq \epsilon_I$, then $\mathbf{IMF}_k = \mathbf{H}_k^{(n)}\mathbf{x}_k^{(n)}$, $n \leftarrow 0$, $\mathbf{r}^k = \mathbf{r}^{k-1} - \mathbf{IMF}_k$, $k \leftarrow k+1$. Else, $n \leftarrow n+1$, goto Step 1.
- Step 7: If $\left\|\mathbf{r}^k\right\| \leq \epsilon_{II}$, stop. Else, goto Step 0.

One might be curious about the presence of the least-squares l_1-regularized minimization in this algorithm. In fact, the presence of the l_2-regularized–l_1 optimization (2.23) is twofold. First, we assume that the envelope can be approximated by a sparse number of terms, as the original approximation asserts. Second, it acts as a stabilizer of the l_2 norm.

As stated before in detail, the remedy to the failure of time-frequency methods in handling the intrawave signals lies in the fact that the filters of the methods must be modified in a way to capture the thick tail of the intrawave signals in the frequency domain. It is not clear how to achieve this within the framework of the EMD [13] or synchrosqueezed wavelet methods [6, 17]. However, for the STFR methods [8], this is straightforward: The filters or envelope dictionaries must be made wider. This fact manifests itself in the shape of the cutoff function in the periodic algorithm as

$$\chi(\omega) = \begin{cases} 1, & -\frac{1}{\lambda} < \omega < \frac{1}{\lambda}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.26}$$

where λ is a measure of the span of the filter. In this paper, we prove that convergence enhances if $\lambda \to 1^+$. However, our numerical results will show that under intense frequency modulation in the original IMF, $\lambda \to 0^+$ can also be used successfully,

[1] In fact, like before, we have

$$\triangle\theta' = P_{V_{M_0}}\left(\frac{d}{dt}\arctan\left(\frac{\mathbf{B}_k^n}{\mathbf{A}_k^n}\right)\right), \triangle\theta(t) = \int_0^t \triangle\theta'(s)\,ds,$$
$$\theta_k^{n+1} = \theta_k^n - \beta\triangle\theta,$$

and $\beta \in [0,1]$ is chosen in a way that the phase function is always strictly increasing:

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

TABLE 1
Coordinates and symbols.

Domain	Symbol	Analysis	Synthesis
θ	$\mathcal{F}_\theta(.)$	$(\mathcal{F}_\theta(g))_k = \frac{1}{\lvert\theta_{end}-\theta_0\rvert}\int_{\theta_0}^{\theta_{end}} g(\theta)\, e^{\frac{-i2\pi k\theta}{\lvert\theta_{end}-\theta_0\rvert}} d\theta$	$g(\theta) = \sum_{k=-\infty}^{\infty} (\mathcal{F}_\theta(g))_k\, e^{\frac{i2\pi k\theta}{\lvert\theta_{end}-\theta_0\rvert}}$
$t \in [0,1]$	$\mathcal{F}(.)$	$(\mathcal{F}(g))_k = \int_0^1 g(t)\, e^{-i2\pi kt} dt$	$g(t) = \sum_{k=-\infty}^{\infty} (\mathcal{F}(g))_k\, e^{i2\pi kt}$
$\bar{\theta}$	$\hat{()}_{\bar{\theta}}$	$\hat{f}_{\bar{\theta}}(k) = \int_0^1 f(\bar{\theta})\, e^{-i2\pi k\bar{\theta}} d\bar{\theta}$	$f(\bar{\theta}) = \sum_{k=-\infty}^{\infty} \hat{f}_{\bar{\theta}}(k)\, e^{i2\pi k\bar{\theta}}$
t	$P_{V_{M_0}}(.)$	$P_{V_{M_0}}(g) = \sum_{k=-M_0}^{M_0} (\mathcal{F}(g))_k\, e^{i2\pi kt}$	$R_{V_{M_0}}(g) = \sum_{\lvert k\rvert > M_0} (\mathcal{F}(g))_k\, e^{i2\pi kt}$
$k \in \mathbb{Z}$	$\lVert.\rVert_{1,M_0}$	$\lVert z\rVert_{1,M_0} = \sum_{\lvert k\rvert \le M_0} \lvert z(k)\rvert$	

though with extreme caution. For the nonperiodic STFR algorithm, the modification happens where we define the envelope dictionary as

$$\{k_l\}_{l=1}^{m} \subset [\lambda, \infty). \tag{2.27}$$

We remark that the filter parameter λ must be chosen based on the characteristics of the intrawave signal. Ideally, we should choose λ adaptive to the strength of the frequency modulation of different IMFs. This problem will be investigated in our future research. From our experience, the choice of $\lambda = 1^+$ often provides a good practical initial value for λ.[2] For the choice of $\lambda = 1^+$, we can prove the convergence of the periodic algorithm for $\lambda > 1$.

3. Theoretical analysis.

3.1. Convergence analysis. In this section, we prove that, for an intrawave signal, increasing the filter span would reduce the extraction error. We will show that the STFR algorithm will converge to an IMF that is close to the intrawave IMF, but with an error associated with the width (span) of the filter. Before proving our convergence result, we introduce some notation in Table 1.

In this section, we assume that an IS can be represented in the following format:

$$f(t) = f_0(t) + f_1(t)\cos\theta(t) \tag{3.1}$$

for $f_1(t) > 0$, $\theta' > 0$, and $t \in [0,1]$. We assume that the signal is periodic with mean zero. In fact, even if the signal $f(t)$ is periodic with a nonzero mean, we can redefine it by reducing the mean of the signal from the signal itself. Next, we need a lemma that helps us to bound the nth derivative of the mth approximated phase function $\bar{\theta}^m$. This lemma can later be used to bound some of the integrals in the main theorem.

LEMMA 3.1. *If* $\left(\bar{\theta}^m\right)' \in V_{M_0}$, *where* $V_{M_0} = span\left\{e^{i2\pi kt}, |k| \le M_0, k \in \mathbb{Z}\right\}$, *then*

$$\left|\left(\bar{\theta}^m\right)^{(n)}(t)\right| \le (2\pi M_0)^{n-1} \left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1 \tag{3.2}$$

for $n \in \mathbb{N}$.

[2]In this paper, the symbol $\lambda = t^+$ is used extensively. By using this symbol, we mean $\lambda = t + \sigma$ for some positive small σ.

Proof. By definition, we have

$$\begin{aligned}\left|\left(\bar{\theta}^m\right)^{(n)}(t)\right| &= \left|\sum_{|k|\le M_0}\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)^{(n)}\right)\right)_k e^{i2\pi kt}\right| \le \sum_{|k|\le M_0}\left|\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)^{(n)}\right)\right)_k\right| \\ &\le \sum_{|k|\le M_0}\left|\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)^{(n)}\right)\right)_k\right| = \sum_{|k|\le M_0}|i2\pi k|^{n-1}\left|\left(\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right)_k\right| \\ &= (2\pi M_0)^{n-1}\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1. \qquad \square\end{aligned}$$

Using this lemma, we can now present the next lemma to bound integrals like $\left|\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m\right|$ that occur frequently in the main theorem. These integrals would be bounded by the norm of the Fourier transform of the phase correction $\triangle\theta'$. In fact, this bound would help us construct a contraction in the main theorem.

LEMMA 3.2. *If* $\left(\bar{\theta}^m\right)' > 0$, $t \in [0,1]$, $\bar{\theta}^m(0) = 0$, $\bar{\theta}^m(1) = 1$, *and* $\left(\bar{\theta}^m\right)', \triangle\theta' \in V_{M_0}$. *Also, if* $e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}$ *is periodic, then for* $\epsilon \neq 0$, *we have*

$$\left|\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m\right| \le \frac{P_m^n M_0^n}{|\epsilon|^n}\sum_{j=1}^n |\delta|^j (2\pi M_0)^{-j}\left(\|\mathcal{F}(\triangle\theta')\|_1\right)^j. \tag{3.3}$$

Proof. Using integration by parts, we have

$$\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m = \frac{1}{(i\epsilon)^n}\int_0^1\left(\frac{d^n}{d\left(\bar{\theta}^m\right)^n}e^{i\delta\triangle\theta}\right)e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m.$$

Now, using the previous lemma, we have

$$\begin{aligned}\left|\int_0^1 e^{i\delta\triangle\theta}e^{-i\epsilon\bar{\theta}^m}d\bar{\theta}^m\right| &\le \left|\frac{1}{|\epsilon|^n}\int_0^1\left(\frac{d^n}{d\left(\bar{\theta}^m\right)^n}e^{i\delta\triangle\theta}\right)d\bar{\theta}^m\right| \\ &\le \frac{P\left(\frac{\|\mathcal{F}((\bar{\theta}^m)')\|_1}{\min(\bar{\theta}^m)'}, n\right)M_0^n}{\left(\min\left(\bar{\theta}^m\right)'\right)^n|\epsilon|^n}\sum_{j=1}^n|\delta|^j(2\pi M_0)^{-j}\left(\|\mathcal{F}(\triangle\theta')\|_1\right)^j \\ &= \frac{P_m^n M_0^n}{|\epsilon|^n}\sum_{j=1}^n|\delta|^j(2\pi M_0)^{-j}\left(\|\mathcal{F}(\triangle\theta')\|_1\right)^j.\end{aligned}$$

Here, $P(x,n)$ is a polynomial of degree $n-1$ and $P_m^n = P\left(\frac{\|\mathcal{F}((\bar{\theta}^m)')\|_1}{\min(\theta^m)'}, n\right)$. This completes the proof. $\square$

Remark 3.1. Here we present a simple calculation on how to compute the poly-

nomial $P(x,n)$ for small n. For example, for $n=2$, we have

$$\begin{aligned}\left|\frac{d^2}{d\left(\bar{\theta}^m\right)^2}e^{i\delta\triangle\theta}\right| &= \left|i\left(\frac{(\delta\triangle\theta)''}{\left(\left(\bar{\theta}^m\right)'\right)^2}-\frac{(\delta\triangle\theta)'\left(\bar{\theta}^m\right)''}{\left(\left(\bar{\theta}^m\right)'\right)^3}+i\frac{\left((\delta\triangle\theta)'\right)^2}{\left(\left(\bar{\theta}^m\right)'\right)^2}\right)e^{i\delta\triangle\theta}\right| \\ &\le \left|\frac{(\delta\triangle\theta)''}{\left(\left(\bar{\theta}^m\right)'\right)^2}\right|+\left|\frac{(\delta\triangle\theta)'\left(\bar{\theta}^m\right)''}{\left(\left(\bar{\theta}^m\right)'\right)^3}\right|+\left|\frac{\left((\delta\triangle\theta)'\right)^2}{\left(\left(\bar{\theta}^m\right)'\right)^2}\right| \\ &\le \frac{\max\left|(\delta\triangle\theta)''\right|}{\min\left(\left(\bar{\theta}^m\right)'\right)^2}+\frac{\max\left|(\delta\triangle\theta)'\right|\max\left|\left(\bar{\theta}^m\right)''\right|}{\min\left(\left(\bar{\theta}^m\right)'\right)^3}+\frac{\max\left|\left((\delta\triangle\theta)'\right)^2\right|}{\min\left(\left(\bar{\theta}^m\right)'\right)^2} \\ &\le \frac{1}{\min\left(\left(\bar{\theta}^m\right)'\right)^2}\left[\left(1+\frac{\left\|\mathcal{F}\left(\left(\bar{\theta}^m\right)'\right)\right\|_1}{\min\left(\bar{\theta}^m\right)'}\right)2\pi M_0\left\|\mathcal{F}\left((\delta\triangle\theta)'\right)\right\|_1+\left\|\mathcal{F}\left((\delta\triangle\theta)'\right)\right\|_1^2\right],\end{aligned}$$

where we have used $\left(\bar{\theta}^m\right)', \triangle\theta' \in V_{M_0}$. In simple words, we have $P(x,2)=K(x+1)$ for some positive constant K.

Having all these preliminaries, we can now present the convergence theorem. The essence of the algorithm is as follows. We try to construct a contraction iterative scheme on $\left\|\mathcal{F}\left(\theta-\theta^m\right)'\right\|_1$, where θ^m is the approximate value of θ at the mth step. This contraction is built upon the error bounds of the extracted envelopes at each iteration of the algorithm. The notation of this proof follows the notation of the algorithm.

THEOREM 3.1 (convergence theorem). *Assume that the instantaneous frequency in* (3.1) *is M_0-sparse; i.e., $\theta' \in V_{M_0}$. Furthermore, assume that*

$$\left|\hat{f}_{0,\bar{\theta}}(k)\right| \le \frac{C_0}{|k|^p},\ \left|\hat{f}_{1,\bar{\theta}}(k)\right| \le \frac{C_0}{|k|^p}$$

for $C_0>0$ and $p\ge 4$. If the initial guess satisfies $\frac{\|\mathcal{F}(\triangle\theta^0)'\|_1}{2\pi M_0}\le\frac{1}{4}$, then there exists an $\eta_0>0$ such that for $L>\eta_0$, we have

$$\left\|\mathcal{F}\left(\theta-\theta^{m+1}\right)'\right\|_1 \le \Gamma_1\lambda^{2p-2}L^{-p+2}+\tfrac{1}{2}\left\|\mathcal{F}\left(\theta-\theta^m\right)'\right\|_1$$

for $\lambda>1$ and $\Gamma_1>0$.

Proof. We know that if $\triangle\theta^m=\theta-\theta^m$, then $a^m=f_1\cos\Delta\theta^m$, and $b^m=-f_1\sin\Delta\theta^m$. Let $\tilde{a}^m$, $\tilde{b}^m$ be approximate envelope functions. Set the error in envelopes as $\triangle a^m=a-\tilde{a}^m$, and set $\triangle b^m=b-\tilde{b}^m$. Using these, we get $f=f_0+a^m\cos\theta^m+b^m\sin\theta^m$. Let $L_m=\frac{\theta^m(1)-\theta^m(0)}{2\pi}$ and $\bar{\theta}^m=\frac{\theta^m}{2\pi L_m}$. Then, we have $f=f_0+a^m\cos 2\pi L_m\bar{\theta}^m+b^m\sin 2\pi L_m\bar{\theta}^m$. If we take the Fourier transform in $\bar{\theta}^m$ coordinate (see Table 1), we get

$$\hat{f}_{\bar{\theta}^m}(k)=\hat{f}_{0,\bar{\theta}^m}(k)+\frac{1}{2}\left(\hat{a}^m_{\bar{\theta}^m}(k-L_m)+\hat{a}^m_{\bar{\theta}^m}(k+L_m)\right)+\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k-L_m)-\hat{b}^m_{\bar{\theta}^m}(k+L_m)\right).$$

Consequently, one can find

$$\begin{aligned}\hat{a}^m_{\bar{\theta}^m}(k) = {}& \hat{f}_{\bar{\theta}^m}(k+L_m)+\hat{f}_{\bar{\theta}^m}(k-L_m) \\ & -\hat{f}_{0,\bar{\theta}^m}(k+L_m)-\hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ & +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m)-\hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ & +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m)-\hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\end{aligned}$$

and

$$\begin{aligned}\hat{b}^m_{\bar{\theta}^m}(k) = & i\hat{f}_{\bar{\theta}^m}(k+L_m) - i\hat{f}_{\bar{\theta}^m}(k-L_m) \\ & -i\hat{f}_{0,\bar{\theta}^m}(k+L_m) + i\hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ & +\frac{i}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) + \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ & +\frac{1}{2}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) + \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right).\end{aligned}$$

In the periodic STFR algorithm, $\hat{\tilde{a}}^m_{\bar{\theta}^m}$, $\hat{\tilde{b}}^m_{\bar{\theta}^m}$ are approximated as

$$\hat{\tilde{a}}^m_{\bar{\theta}^m}(k) = \begin{cases} \hat{f}_{\bar{\theta}^m}(k+L_m) + \hat{f}_{\bar{\theta}^m}(k-L_m), & -\frac{L_m}{\lambda} \le k \le \frac{L_m}{\lambda}, \\ 0 & \text{otherwise,}\end{cases}$$

$$\hat{\tilde{b}}^m_{\bar{\theta}^m}(k) = \begin{cases} i\left(\hat{f}_{\bar{\theta}^m}(k+L_m) - \hat{f}_{\bar{\theta}^m}(k-L_m)\right), & -\frac{L_m}{\lambda} \le k \le \frac{L_m}{\lambda}, \\ 0 & \text{otherwise.}\end{cases}$$

Here, $1 < \lambda$ defines the width of the filter. Hence, for $\triangle\hat{a}^m_{\bar{\theta}^m}$ and $\triangle\hat{b}^m_{\bar{\theta}^m}$, we have

$$\triangle\hat{a}^m_{\bar{\theta}^m}(k) = \begin{cases} \left\{\begin{array}{l} -\hat{f}_{0,\bar{\theta}^m}(k+L_m) - \hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) - \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) - \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\end{array}\right\}, & |k| \le \frac{L_m}{\lambda}, \\ \hat{a}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda},\end{cases}$$

$$\triangle\hat{b}^m_{\bar{\theta}^m}(k) = \begin{cases} \left\{\begin{array}{l} -i\hat{f}_{0,\bar{\theta}^m}(k+L_m) + i\hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{i}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) + \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ +\frac{1}{2}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) + \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\end{array}\right\}, & |k| \le \frac{L_m}{\lambda}, \\ \hat{b}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda}.\end{cases}$$

The ideal case for updating the phase function is $\frac{d\theta_{new}}{dt} = \frac{d\theta^m}{dt} - \frac{d}{dt}\arctan\frac{\tilde{b}^m}{\tilde{a}^m}$. However, the algorithm works in a way that we must choose $\frac{d\theta^{m+1}}{dt}$ in V_{M_0}. Hence, $\frac{d\theta^{m+1}}{dt} = P_{V_{M_0}}\left(\frac{d\theta_{new}}{dt}\right)$. So, at each step, we force $\frac{d\theta^m}{dt}$ to be in V_{M_0}. In simple words, $\frac{d\theta^m}{dt} \in V_{M_0}$ for all $m \ge 0$, $m \in \mathbb{Z}$. This short analysis tells us that $\frac{d\theta^{m+1}}{dt} = \frac{d\theta^m}{dt} - P_{V_{M_0}}\left(\frac{d}{dt}\arctan\frac{\tilde{b}^m}{\tilde{a}^m}\right)$. Since $\theta \in V_{M_0}$ is sufficiently differentiable, and $\frac{d\theta^m}{dt} \in V_{M_0}$, then $(\theta - \theta^m = \triangle\theta^m) \in V_{M_0}$. Having these in mind, we can find

$$\begin{aligned}\frac{d}{dt}\triangle\theta^{m+1} = & \frac{d}{dt}\left(\theta - \theta^{m+1}\right) = \frac{d\theta}{dt} - \frac{d\theta^m}{dt} + P_{V_{M_0}}\left(\frac{d}{dt}\arctan\frac{\tilde{b}^m}{\tilde{a}^m}\right) \\ = & P_{V_{M_0}}\frac{d}{dt}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right) + R_{V_{M_0}}\left(\frac{d}{dt}\triangle\theta^m\right).\end{aligned}$$

We know $\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}$ is in $\mathcal{C}^1$. For any $g \in \mathcal{C}^1$, we have $P_{V_{M_0}}\frac{d}{dt}(g) = \frac{d}{dt}P_{V_{M_0}}(g)$. Hence, the Fourier transform of the IF error is

$$\begin{aligned}\left(\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right)_k &= \left(\mathcal{F}\left(P_{V_{M_0}}\frac{d}{dt}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right)\right)\right)_k + \left(\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right)_k \\ &= (i2\pi k)\left(\mathcal{F}\left(P_{V_{M_0}}\left(\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right)\right)\right)_k + \left(\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right)_k.\end{aligned}$$

As $(\mathcal{F}(P_{V_{M_0}}(\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m})))_k=0$ for $|k|>M_0$, then

$$\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq (2\pi M_0)\left\|\mathcal{F}\left(P_{V_{M_0}}\left(\arctan\tfrac{\tilde{b}^m}{\tilde{a}^m}-\arctan\tfrac{b^m}{a^m}\right)\right)\right\|_1+\left\|\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right\|_1.$$

Now, we know that for any function $g(t)$, we have

$$\begin{aligned}\left\|\mathcal{F}\left(P_{V_{M_0}}(g)\right)\right\|_1 &= \sum_{k=-M_0}^{M_0}\left|\left(\mathcal{F}(g)\right)_k\right| = \sum_{k=-M_0}^{M_0}\left|\int_0^1 g(t)\,e^{-i2\pi kt}dt\right| \\ &\leq \sum_{k=-M_0}^{M_0}\int_0^1|g(t)|\,dt \leq \sum_{k=-M_0}^{M_0}\|g\|_\infty = (2M_0+1)\|g\|_\infty.\end{aligned}$$

Hence, the l_1 norm of the Fourier of the IF error is

$$\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq (2\pi M_0)(2M_0+1)\left\|\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right\|_\infty+\left\|\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right\|_1.$$

Since θ' is sparse in V_{M_0}, then the last term vanishes. In fact, we previously showed that $\frac{d\theta^m}{dt}\in V_{M_0}$, and hence

$$\left(\mathcal{F}\left(\triangle\theta^m\right)'\right)_k=\left(\mathcal{F}\left(\theta-\theta^m\right)'\right)_k=\begin{cases}\left(\mathcal{F}(\theta')\right)_k-\left(\mathcal{F}\left(\left(\theta^m\right)'\right)\right)_k, & |k|\leq M_0,\\ \left(\mathcal{F}(\theta')\right)_k=0, & |k|>M_0,\end{cases}$$

so

$$R_{V_{M_0}}\left(\triangle\theta^m\right)'=\sum_{|k|>M_0}\left(\mathcal{F}\left(\triangle\theta^m\right)'\right)_k e^{i2\pi kt}=\sum_{|k|>M_0}\left(\mathcal{F}(\theta')\right)_k e^{i2\pi kt}=0,$$

and then $\left(\mathcal{F}\left(R_{V_{M_0}}\left(\triangle\theta^m\right)'\right)\right)_k=0$. Finally, we have the following bound on the IF error:

$$\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq (2\pi M_0)(2M_0+1)\left\|\arctan\tfrac{\tilde{b}^m}{\tilde{a}^m}-\arctan\tfrac{b^m}{a^m}\right\|_\infty.$$

The last term on the right-hand side of this bound can be simplified further. If we use the fact that $x^2+xy\leq\frac{x^2}{2}-y^2$ for real x and y, we have

$$\begin{aligned}\left|\arctan\frac{\tilde{b}^m}{\tilde{a}^m}-\arctan\frac{b^m}{a^m}\right| &= \left|\arctan\left(\frac{\tilde{b}^m a^m-\tilde{a}^m b^m}{a^m\tilde{a}^m+b^m\tilde{b}^m}\right)\right| \leq \left|\frac{\tilde{b}^m a^m-\tilde{a}^m b^m}{a^m\tilde{a}^m+b^m\tilde{b}^m}\right| \\ &= \left|\frac{\left(a^m+\triangle a^m\right)\triangle b^m-\left(b^m+\triangle b^m\right)\triangle a^m}{\left(a^m\right)^2+\left(b^m\right)^2+\left(\triangle a^m\right)a^m+\left(\triangle b^m\right)b^m}\right| \\ &\leq \frac{\left(|a^m|+|\triangle a^m|\right)|\triangle b^m|+\left(|b^m|+|\triangle b^m|\right)|\triangle a^m|}{\frac{(a^m)^2+(b^m)^2}{2}-\left(\left(\triangle a^m\right)^2+\left(\triangle b^m\right)^2\right)} \\ &\leq D\left(|\triangle a^m|+|\triangle b^m|\right)\end{aligned}$$

for $D=\max\left(\frac{f_1+|\triangle a^m|}{\frac{f_1^2}{2}-((\triangle a^m)^2+(\triangle b^m)^2)},\frac{f_1+|\triangle b^m|}{\frac{f_1^2}{2}-((\triangle a^m)^2+(\triangle b^m)^2)}\right)$. In the latter, we have taken into account that $f_1>0$. Now, consider the fact that

$$|\triangle a^m(t)|=|\triangle a^m(\theta)|=\left|\triangle a^m\left(\bar{\theta}^m\right)\right|=\left|\sum_{k=-\infty}^{\infty}\triangle\hat{a}^m_{\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m}\right|\leq\sum_{k=-\infty}^{\infty}\left|\triangle\hat{a}^m_{\bar{\theta}^m}(k)\right|=\left\|\triangle\hat{a}^m_{\bar{\theta}^m}\right\|_1,$$

and as in general $\hat{g}_{\bar{\theta}}(-k) = \overline{\hat{g}_{\bar{\theta}}(k)}$ for any g, then $|\hat{g}_{\bar{\theta}}(-k)| = |\hat{g}_{\bar{\theta}}(k)|$. Consequently, the bound on the envelopes errors can be expressed as

$$|\triangle a^m| \leq \left\|\triangle \hat{a}^m_{\bar{\theta}^m}\right\|_1 \leq 2 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq k \leq \left(1+\frac{1}{\lambda}\right)L_m} \left|\hat{f}_{0,\bar{\theta}^m}(k)\right| + \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq k \leq \left(2+\frac{1}{\lambda}\right)L_m} \left(\left|\hat{a}^m_{\bar{\theta}^m}(k)\right| + \left|\hat{b}^m_{\bar{\theta}^m}(k)\right|\right) + \sum_{|k| > \frac{L_m}{\lambda}} \left|\hat{a}^m_{\bar{\theta}^m}(k)\right|$$

and

$$|\triangle b^m| \leq \left\|\triangle \hat{b}^m_{\bar{\theta}^m}\right\|_1 \leq 2 \sum_{\left(1-\frac{1}{\lambda}\right)L_m \leq k \leq \left(1+\frac{1}{\lambda}\right)L_m} \left|\hat{f}_{0,\bar{\theta}^m}(k)\right| + \sum_{\left(2-\frac{1}{\lambda}\right)L_m \leq k \leq \left(2+\frac{1}{\lambda}\right)L_m} \left(\left|\hat{a}^m_{\bar{\theta}^m}(k)\right| + \left|\hat{b}^m_{\bar{\theta}^m}(k)\right|\right) + \sum_{|k| > \frac{L_m}{\lambda}} \left|\hat{b}^m_{\bar{\theta}^m}(k)\right|.$$

We will soon use approximations on these terms. Before that, we recall that according to one of our assumptions, we have the following: "The observation is periodic with mean zero." Hence, $\hat{f}_{0,\bar{\theta}^m}(0) = 0$. Furthermore, as $e^{i2\pi k\bar{\theta}} e^{-i2\pi\omega\bar{\theta}^m}$ is periodic, so is $e^{i2\pi\bar{\theta}^m\left(k\frac{L_m}{L}-\omega\right)} e^{ik\frac{\triangle\theta^m}{L}}$, and then we have the following estimate using these facts and the previous lemma:

(3.4)
$$\begin{aligned}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \;\leq\; & 2C_0\left(\tfrac{\alpha\lambda}{|\omega|}\right)^{p-1} \\ & +2C_0P_m^n\left(\tfrac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\left(\sum_{j=1}^{p-2}\left(\tfrac{\gamma}{L}\right)^j\tfrac{\pi^2}{3}+2\sum_{j=p}^{n}\left|\tfrac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\tfrac{\gamma}{L}\right)^j+2\left(\tfrac{\gamma}{L}\right)^{p-1}\left(1+\tfrac{|\omega|}{\alpha\lambda}\right)\right).\end{aligned}$$

Here, $\gamma = \frac{\|\mathcal{F}(\triangle\theta^m)'\|_1}{2\pi M_0}$. The proof of this inequality can be found in Appendix A. Now it is time to find the estimate on $\hat{a}^m_{\bar{\theta}^m}(k)$. The approach is essentially the same as what we just had above. The reader can find the approximation procedure in Appendix B.

$$\begin{aligned}|\hat{a}^m_{\bar{\theta}^m}(\omega)| \leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + \left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^{n}\gamma^j + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\frac{2p}{p-1}\sum_{j=1}^{n}2^j\gamma^j \\ & + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \\ & \times\left(\sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3}+\sum_{j=p}^{n}2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j+2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).\end{aligned} \tag{3.5}$$

There is a similar bound for $\hat{b}^m_{\bar{\theta}^m}(\omega)$ as well:

$$\begin{aligned}\left|\hat{b}^m_{\bar{\theta}^m}(\omega)\right| \leq & 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + \left|\hat{f}_{1,\bar{\theta}}(0)\right| P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^{n}\gamma^j + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\frac{2p}{p-1}\sum_{j=1}^{n}2^j\gamma^j \\ & + C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \\ & \times\left(\sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3}+\sum_{j=p}^{n}2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j+2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).\end{aligned}$$

Now we approximate the terms in the inequality for $|\triangle a^m|$. One can find the details of this approximation in Appendix C. We have a similar inequality for $|\triangle b^m|$. These can be expressed in the following way:

$$|\triangle a^m| \leq C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma\,, \tag{3.6}$$

$$|\triangle b^m| \leq C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma\,, \tag{3.7}$$

where C_1 and C_2 depend on C_0, M_0, ε, α, P_m^n, n, p. Before moving forward, we need to show that parameters like α and P_m^n are uniformly bounded. We start with α. We have

$$\begin{aligned}|1-\alpha| &= \left|1-\frac{L_m}{L}\right| = \left|1-\frac{\frac{\theta^m(1)-\theta^m(0)}{2\pi}}{\frac{\bar\theta^m(1)-\bar\theta^m(0)}{2\pi}}\right| = \left|\frac{\Delta\theta^m(1)-\Delta\theta^m(0)}{2\pi L}\right| \\ &\leq \left\|\frac{(\Delta\theta^m)'}{2\pi L}\right\|_\infty \leq \left\|\frac{\mathcal{F}(\Delta\theta^m)'}{2\pi L}\right\|_1 = \frac{\gamma M_0}{L} \leq \frac{M_0}{4L} \leq \frac{1}{8}.\end{aligned} \tag{3.8}$$

Here, we have used the fact that $\gamma \leq \frac{1}{4}$. At the last part of the proof, we will show that this condition remains intact for all iterations. The last inequality in (3.8) can be true for the condition $L \geq 2M_0$. Hence, $\frac{7}{8} \leq \alpha \leq \frac{9}{8}$. In order to prove the boundedness of P_m^n at every step, we need to find bounds on $\|\mathcal{F}((\bar\theta^m)')\|_1$ and $\min\left(\bar\theta^m\right)'$. If we take the condition $\min\left(\bar\theta'\right) \geq \frac{M_0}{2L}$, we get $\bar\theta' \geq \frac{M_0}{2L} > \frac{M_0}{4L} \geq \frac{\gamma M_0}{L} = \|\frac{\mathcal{F}(\Delta\theta^m)'}{2\pi L}\|_1 \geq |\frac{(\Delta\theta^m)'}{2\pi L}|$. Using this, we have

$$\begin{aligned}\left|\left(\bar\theta^m\right)'\right| &= \left|\frac{(\theta^m)'}{2\pi L_m}\right| = \left|\frac{(\theta-\Delta\theta^m)'}{2\pi\alpha L}\right| \geq \frac{8}{9}\left|\frac{(\theta-\Delta\theta^m)'}{2\pi L}\right| \\ &= \frac{8}{9}\left|\bar\theta' - \frac{(\Delta\theta^m)'}{2\pi L}\right| \geq \frac{8}{9}\left(\bar\theta' - \frac{(\Delta\theta^m)'}{2\pi L}\right) \geq \frac{8}{9}\left(\bar\theta' - \frac{M_0}{4L}\right) \\ &\geq \frac{8}{9}\left(\bar\theta' - \frac{M_0}{4L}\right) \geq \frac{8}{9}\left(\bar\theta' - \frac{\bar\theta'}{2}\right) \geq \frac{4}{9}\min\left(\bar\theta'\right).\end{aligned}$$

We also have

$$\begin{aligned}\left\|\mathcal{F}\left(\left(\bar\theta^m\right)'\right)\right\|_1 &= \frac{1}{\alpha}\left\|\mathcal{F}\left(\frac{(\theta-\Delta\theta^m)'}{2\pi L}\right)\right\|_1 = \frac{1}{\alpha}\left\|\mathcal{F}\left(\bar\theta'\right) - \frac{\mathcal{F}\left((\Delta\theta^m)'\right)}{2\pi L}\right\|_1 \\ &\leq \frac{8}{7}\left(\left\|\mathcal{F}\left(\bar\theta'\right)\right\|_1 + \frac{\left\|\mathcal{F}\left((\Delta\theta^m)'\right)\right\|_1}{2\pi L}\right) = \frac{8}{7}\left(\left\|\mathcal{F}\left(\bar\theta'\right)\right\|_1 + \frac{M_0\gamma}{L}\right) \\ &\leq \frac{8}{7}\left(\left\|\mathcal{F}\left(\bar\theta'\right)\right\|_1 + \frac{M_0}{4L}\right).\end{aligned}$$

These two estimates pave the way to rigorously prove that P_m^n is bounded at every single step. Now it is time for a bound on D. In fact, if we bound D uniformly over all steps, we can set a contraction. Assuming that $1 < \lambda < \lambda_0$, and since $f_1 > 0$, taking the condition

$$\frac{C_1\lambda_0^{2p-2}}{L^{-p+1}} + \frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1$$

would result in[3]

$$\begin{aligned}|\triangle a^m| \leq & C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma \\ \leq & \frac{C_1\lambda_0^{2p-2}}{L^{p-2}} + \frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1.\end{aligned}$$

Similarly, we have the same bound for $|\triangle b^m|$, and then

$$\begin{aligned}D = & \max\left(\frac{f_1 + |\triangle a^m|}{\frac{f_1^2}{2} - \left((\triangle a^m)^2 + (\triangle b^m)^2\right)}, \frac{f_1 + |\triangle b^m|}{\frac{f_1^2}{2} - \left((\triangle a^m)^2 + (\triangle b^m)^2\right)}\right) \\ \leq & \frac{\max f_1 + \frac{\sqrt{2}}{4}\min f_1}{\min\left(\frac{f_1^2}{2} - \left((\triangle a^m)^2 + (\triangle b^m)^2\right)\right)} \leq \frac{\max f_1 + \frac{\sqrt{2}}{4}\min f_1}{\min\frac{f_1^2}{2} - \frac{1}{4}\min f_1^2} \\ = & \frac{4\max f_1 + \sqrt{2}\min f_1}{\min f_1^2} = E_0.\end{aligned}$$

The latter shows that D is also bounded. So, we get

$$\begin{aligned}\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq & (2\pi M_0)(2M_0+1)\left\|\arctan\frac{\tilde{b}^m}{\tilde{a}^m} - \arctan\frac{b^m}{a^m}\right\|_\infty \\ \leq & (2\pi M_0)(2M_0+1)D\left(|\triangle a^m| + |\triangle b^m|\right) \\ \leq & E_0(4\pi M_0)(2M_0+1)C_1\lambda^{2p-2}L^{-p+2} \\ & + E_0(4\pi M_0)(2M_0+1)C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma.\end{aligned}$$

The last inequality is nothing but

$$\left\|\mathcal{F}\left(\triangle\theta^{m+1}\right)'\right\|_1 \leq \Gamma_1\lambda^{2p-2}L^{-p+2} + \Gamma_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\left\|\mathcal{F}\left(\triangle\theta^m\right)'\right\|_1.$$

If we have the condition that $\Gamma_2\lambda_0^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)} \leq \frac{1}{2}$, we have the contraction that we were looking for. Before finishing, we need to state the following: There exists $\eta_0 > 0$ such that $L > \eta_0$; then all the conditions

$$\begin{aligned}L \geq 2M_0, \quad \min\left(\bar{\theta}'\right) \geq \frac{M_0}{2L}, \\ \frac{C_1\lambda_0^{2p-2}}{L^{p-2}} + \frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1, \\ \Gamma_2\lambda_0^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)} \leq \frac{1}{2}\end{aligned}$$

would be satisfied. As a result, we have

$$\left\|\mathcal{F}\left(\theta - \theta^{m+1}\right)'\right\|_1 \leq \Gamma_1\lambda^{2p-2}L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}\left(\theta - \theta^m\right)'\right\|_1.$$

Having the latter, the condition $\gamma \leq \frac{1}{4}$ would remain intact for all iterations. In other words, when there is a contraction on $\|\mathcal{F}(\theta - \theta^{m+1})'\|_1$, this term would remain

[3]Remember that if f_1 has a zero crossing, this condition would never be satisfied. Hence, the rest of the proof will not be valid.

bounded. Hence, if $\frac{\|\mathcal{F}(\triangle\theta^0)'\|_1}{2\pi M_0} \leq \frac{1}{4}$ for the first iteration, it will remain bounded by $\frac{1}{4}$ for all iterations. This completes the proof. ∎

Remark 3.2. In our proof, the convergence of the method depends mainly on the filter parameter λ and the initial guess of the phase function. In many examples that we have tested, the convergence of our method does not depend on the filter parameter λ in a sensitive manner. The convergence is observed in a much wider range of parameter than what is required by our convergence analysis. Actually, we found that the choice of λ plays a critical role in obtaining good accuracy. In many cases, with an inappropriate choice of λ, the method still converges, but the accuracy would be low.

Remark 3.3. If f_0 has only high-frequency components, there would be no interference between f_0 and $f_1 \cos\theta$. In this case, we can prove the convergence of our method for $\lambda > 0$.

Furthermore, there are important facts about this theorem on which we want to elaborate. We begin with the nonuniqueness issue. It is possible that an IS has multiple representations.[4] This theorem states merely that the algorithm converges to an IMF in one of these representations. The theorem does not mention which one the algorithm converges to. The theorem that we observed, in this part, says that if in a representation we have $|\hat{f}_{0,\bar{\theta}}(k)| \leq \frac{C_0}{|k|^p}$, $|\hat{f}_{1,\bar{\theta}}(k)| \leq \frac{C_0}{|k|^p}$, then increasing the width of the filter will reduce the error in extraction. More specifically, if in a representation we have a wide band signal (like an intrawave signal), we are more likely to capture it by widening the filter width. This approach helps us to find a sparser representation compared to the case in which one uses a normal envelope dictionary (1.3). From the algorithmic point of view, the initial guess θ_0 and the parameter λ define the representation to which we converge. In practice, if the IMFs that constitute the signal are separated enough in the time-frequency domain, the reduction of the parameter λ is always beneficial. Let us illustrate this idea in the following example.

Example 3.1. Again consider generic intrawave IMF $x(t) = \cos(\omega t + \frac{\Delta\omega}{p}\sin(pt))$. We can have many representations for this signal. The first representation would be

$$x(t) = \cos\left(\omega t + \frac{\triangle\,\omega}{p}\sin(pt)\right) \tag{3.9}$$

for $\bar{\theta} = \omega t + \frac{\Delta\omega}{p}\sin(pt)$. The second representation would be

$$\begin{aligned} x(\theta) = &\left\{J_0\left(\frac{\triangle\,\omega}{p}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega}{p}\right)\cos(2kpt)\right\}\cos(\omega t) \\ &- \left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega}{p}\right)\sin((2k-1)pt)\right\}\sin(\omega t) \end{aligned}$$

for $\bar{\theta} = \omega t$. So, with a constant initial guess for θ_0 close to $\bar{\theta} = \omega t$, the second representation is the one that is seen by the algorithm at the first iteration. Hence, if a short band filter (large value of λ) is used, only the first harmonic is extracted by the algorithm. In order to capture more terms in the envelope of the second representation, we need a smaller value of λ.

[4] This is the topic of our future research. We intend to quantify this phenomenon.

3.2. Recovery of signals polluted by noise. The next theorem shows that if the observation is polluted by noise, under mild presence of noise, the algorithm still converges. Take

$$f(t) = f_0(t) + f_1(t)\cos\theta(t) + \Im, \tag{3.10}$$

where $\Im$ is a periodic perturbation to the original signal (3.1).

THEOREM 3.2 (convergence theorem in the presence of noise). *Assume that the instantaneous frequency in* (3.10) *is M_0-sparse; i.e., $\theta' \in V_{M_0}$. Furthermore, assume that*

$$\left|\hat{f}_{0,\bar{\theta}}(k)\right| \le \frac{C_0}{|k|^p}, \left|\hat{f}_{1,\bar{\theta}}(k)\right| \le \frac{C_0}{|k|^p}$$

for $C_0 > 0$ and $p \ge 4$. If the initial guess satisfies $\frac{\|\mathcal{F}(\triangle\theta^0)'\|_1}{2\pi M_0} \le \frac{1}{4}$, then there exists an $\eta_0 > 0$ such that for $L > \eta_0$ and $\|\Im\|_\infty \le \epsilon_0$ (ϵ_0 sufficiently small), we have

$$\left\|\mathcal{F}\left(\theta - \theta^{m+1}\right)'\right\|_1 \le \Upsilon_0(L,\lambda)\|\Im(t)\|_\infty + \Gamma_1\lambda^{2p-2}L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}\left(\theta - \theta^m\right)'\right\|_1$$

for $\lambda > 1$, $\Gamma_1 > 0$, and $\Upsilon_0(L,\lambda)$.

Proof. The proof of this theorem is essentially the same as the proof of the convergence theorem in the absence of noise. There is only one minor change. Using the same trend of proof as before, we have

$$\triangle\hat{a}^m_{\bar{\theta}^m}(k) = \begin{cases} \Big\{-\hat{\Im}_{\bar{\theta}^m}(k+L_m) - \hat{\Im}_{\bar{\theta}^m}(k-L_m) \\ -\hat{f}_{0,\bar{\theta}^m}(k+L_m) - \hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) - \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) - \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\Big\}, & |k| \le \frac{L_m}{\lambda}, \\ \hat{a}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda}, \end{cases}$$

$$\triangle\hat{b}^m_{\bar{\theta}^m}(k) = \begin{cases} \Big\{-i\hat{\Im}_{\bar{\theta}^m}(k+L_m) + i\hat{\Im}_{\bar{\theta}^m}(k-L_m) \\ -i\hat{f}_{0,\bar{\theta}^m}(k+L_m) + i\hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{i}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) + \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ +\frac{1}{2}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) + \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\Big\}, & |k| \le \frac{L_m}{\lambda}, \\ \hat{b}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda}. \end{cases}$$

Here, in simple words, we have $\triangle a^m(t) = \Im^m_{a,trunc}(t) + \triangle a^m_{old}(t)$, where $\Im^m_{a,trunc}\left(\bar{\theta}^m\right)$ is the truncated part of $\Im$ in a:

$$\Im^m_{a,trunc}\left(\bar{\theta}^m\right) = \sum_{k=-\infty}^{\infty}\hat{\Im}^m_{a,trunc,\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m} = \sum_{|k|\le\frac{L_m}{\lambda}}\left(-\hat{\Im}_{\bar{\theta}^m}(k+L_m) - \hat{\Im}_{\bar{\theta}^m}(k-L_m)\right)e^{i2\pi k\bar{\theta}^m}.$$

Also, $\triangle a^m_{old}(t)$, in Fourier domain, is defined as

$$\triangle\hat{a}^m_{old,\bar{\theta}^m}(k) = \begin{cases} \Big\{-\hat{f}_{0,\bar{\theta}^m}(k+L_m) - \hat{f}_{0,\bar{\theta}^m}(k-L_m) \\ +\frac{1}{2}\left(-\hat{a}^m_{\bar{\theta}^m}(k+2L_m) - \hat{a}^m_{\bar{\theta}^m}(k-2L_m)\right) \\ +\frac{1}{2i}\left(\hat{b}^m_{\bar{\theta}^m}(k+2L_m) - \hat{b}^m_{\bar{\theta}^m}(k-2L_m)\right)\Big\}, & |k| \le \frac{L_m}{\lambda}, \\ \hat{a}^m_{\bar{\theta}^m}(k), & |k| > \frac{L_m}{\lambda}. \end{cases}$$

The latter is telling us that we need only find a bound on $\Im^m_{trunc}(t)$ and then use the results of the former theorem to finalize the proof. One can simplify $\Im^m_{a,trunc}\left(\bar{\theta}^m\right)$, using

$$\chi_\lambda(k) = \begin{cases} 1, & |k| \leq \frac{L_m}{\lambda}, \\ 0, & |k| > \frac{L_m}{\lambda}, \end{cases}$$

as

$$\begin{aligned}\Im^m_{a,trunc}\left(\bar{\theta}^m\right) = &\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k+L_m) - \hat{\Im}_{\bar{\theta}^m}(k-L_m)\right)\chi_\lambda(k)\, e^{i2\pi k\bar{\theta}^m} \\ =&e^{-i2\pi L_m\bar{\theta}^m}\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k)\right)\chi_\lambda(k-L_m)\, e^{i2\pi k\bar{\theta}^m} \\ &+ e^{i2\pi L_m\bar{\theta}^m}\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k)\right)\chi_\lambda(k+L_m)\, e^{i2\pi k\bar{\theta}^m}.\end{aligned}$$

Here, we need to pay attention to $\sum_{k=-\infty}^{\infty}\hat{\Im}_{\bar{\theta}^m}(k)\,\chi_\lambda(k+L_m)\, e^{i2\pi k\bar{\theta}^m}$ for further simplifications. Rename $\hat{\Im}_{\bar{\theta}^m}(k)\,\chi_\lambda(k+L_m)$ as $\hat{P}_{1,\bar{\theta}^m}(k)$. Hence, we have

$$P_1\left(\bar{\theta}^m\right) = \int_0^1 \Im(\tau)\,\check{\chi}_{1,\lambda}\left(\bar{\theta}^m - \tau\right) d\tau,$$

where

$$\check{\chi}_{1,\lambda}\left(\bar{\theta}^m\right) = \sum_{k=-\infty}^{\infty}\chi_\lambda(k+L_m)\, e^{i2\pi k\bar{\theta}^m} = \sum_{L_m\left(-\frac{1}{\lambda}-1\right)\leq k\leq L_m\left(\frac{1}{\lambda}-1\right)} e^{i2\pi k\bar{\theta}^m}.$$

Also, in $\sum_{k=-\infty}^{\infty}(-\hat{\Im}_{\bar{\theta}^m}(k))\chi_\lambda(k-L_m)e^{i2\pi k\bar{\theta}^m}$, rename $(-\hat{\Im}_{\bar{\theta}^m}(k))\chi_\lambda(k-L_m)$ as $\hat{P}_{2,\bar{\theta}^m}(k)$. Hence, we get

$$P_2\left(\bar{\theta}^m\right) = \int_0^1 \Im(\tau)\,\check{\chi}_{2,\lambda}\left(\bar{\theta}^m - \tau\right) d\tau,$$

where

$$\check{\chi}_{2,\lambda}\left(\bar{\theta}^m\right) = \sum_{k=-\infty}^{\infty}\chi_\lambda(k-L_m)\, e^{i2\pi k\bar{\theta}^m} = \sum_{L_m\left(-\frac{1}{\lambda}+1\right)\leq k\leq L_m\left(\frac{1}{\lambda}+1\right)} e^{i2\pi k\bar{\theta}^m}.$$

As a consequence, one can find

$$\begin{aligned}
\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| &\leq \left|e^{-i2\pi L_m\bar{\theta}^m}\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k)\right)\chi_\lambda(k-L_m)\,e^{i2\pi k\bar{\theta}^m}\right| \\
&\quad+\left|e^{-i2\pi L_m\bar{\theta}^m}\sum_{k=-\infty}^{\infty}\left(-\hat{\Im}_{\bar{\theta}^m}(k)\right)\chi_\lambda(k-L_m)\,e^{i2\pi k\bar{\theta}^m}\right| \\
&=\left|\sum_{k=-\infty}^{\infty}\hat{P}_{2,\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m}\right|+\left|\sum_{k=-\infty}^{\infty}\hat{P}_{1,\bar{\theta}^m}(k)\,e^{i2\pi k\bar{\theta}^m}\right| \\
&=\left|\int_0^1\Im(\tau)\,\check{\chi}_{2,\lambda}\left(\bar{\theta}^m-\tau\right)d\tau\right|+\left|\int_0^1\Im(\tau)\,\check{\chi}_{1,\lambda}\left(\bar{\theta}^m-\tau\right)d\tau\right| \\
&\leq\|\Im(t)\|_\infty\left(\int_0^1\left|\check{\chi}_{2,\lambda}\left(\bar{\theta}^m-\tau\right)\right|d\tau+\int_0^1\left|\check{\chi}_{1,\lambda}\left(\bar{\theta}^m-\tau\right)\right|d\tau\right).
\end{aligned}$$

Since $\check{\chi}_{2,\lambda}\left(\bar{\theta}^m-\tau\right)$ and $\check{\chi}_{1,\lambda}\left(\bar{\theta}^m-\tau\right)$ are periodic, we have

$$\int_0^1\left|\check{\chi}_{i,\lambda}\left(\bar{\theta}^m-\tau\right)\right|d\tau=-\int_{\bar{\theta}^m}^{\bar{\theta}^m-1}\left|\check{\chi}_{i,\lambda}(\tau)\right|d\tau=\int_{\bar{\theta}^m-1}^{\bar{\theta}^m}\left|\check{\chi}_{i,\lambda}(\tau)\right|d\tau=\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\check{\chi}_{i,\lambda}(\tau)\right|d\tau$$

for $i=1,2$. So, we get

$$\begin{aligned}
\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| &\leq\|\Im(t)\|_\infty\left(\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\check{\chi}_{1,\lambda}(\tau)\right|d\tau+\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\check{\chi}_{2,\lambda}(\tau)\right|d\tau\right) \\
&\leq\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}} \\
&\quad\times\left(\left|\sum_{L_m\left(-\frac{1}{\lambda}-1\right)\leq k\leq L_m\left(\frac{1}{\lambda}-1\right)}e^{i2\pi k\tau}\right|+\left|\sum_{L_m\left(-\frac{1}{\lambda}+1\right)\leq k\leq L_m\left(\frac{1}{\lambda}+1\right)}e^{i2\pi k\tau}\right|\right)d\tau.
\end{aligned}$$

Now, using the identity $e^{i2\pi kt}=\frac{\pi t}{\sin\pi t}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}e^{i2\pi\omega t}d\omega$, we get

$$\begin{aligned}
\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right| &\leq\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\sum_{L_m\left(-\frac{1}{\lambda}-1\right)\leq k\leq L_m\left(\frac{1}{\lambda}-1\right)}\frac{\pi t}{\sin\pi t}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}e^{i2\pi\omega t}d\omega\right|dt \\
&\quad+\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\sum_{L_m\left(-\frac{1}{\lambda}+1\right)\leq k\leq L_m\left(\frac{1}{\lambda}+1\right)}\frac{\pi t}{\sin\pi t}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}}e^{i2\pi\omega t}d\omega\right|dt \\
&\leq\frac{\pi}{2}\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\frac{e^{i2\pi\left(L_m\left(\frac{1}{\lambda}-1\right)+\frac{1}{2}\right)t}-e^{i2\pi\left(L_m\left(-\frac{1}{\lambda}-1\right)-\frac{1}{2}\right)t}}{2\pi it}\right|dt \\
&\quad+\frac{\pi}{2}\|\Im(t)\|_\infty\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\frac{e^{i2\pi\left(L_m\left(\frac{1}{\lambda}+1\right)+\frac{1}{2}\right)t}-e^{i2\pi\left(L_m\left(-\frac{1}{\lambda}+1\right)-\frac{1}{2}\right)t}}{2\pi it}\right|dt.
\end{aligned}$$

In the last inequality, we have used the fact that $\max_{-\frac{1}{2}\leq t\leq\frac{1}{2}}\left(\frac{\pi t}{\sin\pi t}\right)\leq\frac{\pi}{2}$. From the previous theorem, we know that $\frac{L_m}{L}\leq\frac{9}{8}$. Hence, the above bound would be

$$\left|\Im^m_{a,trunc}\left(\bar{\theta}^m\right)\right|\leq\Upsilon(L,\lambda)\,\|\Im(t)\|_\infty$$

for some constant $\Upsilon(L,\lambda)$. The same thing is true for $|\Im^m_{b,trunc}(\bar{\theta}^m)|$:

$$\left|\Im^m_{b,trunc}\left(\bar{\theta}^m\right)\right| \leq \Upsilon(L,\lambda)\left\|\Im(t)\right\|_\infty .$$

Hence, we get

$$\begin{aligned}|\triangle a^m| &\leq \Upsilon(L,\lambda)\left\|\Im(t)\right\|_\infty + C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma ,\\ |\triangle b^m| &\leq \Upsilon(L,\lambda)\left\|\Im(t)\right\|_\infty + C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma .\end{aligned}$$

Now, if $\frac{\|\mathcal{F}(\triangle\theta^m)'\|_1}{2\pi M_0} = \gamma \leq \frac{1}{4}$, one can find $\eta_0 > 0$ such that $L > \eta_0$ in a way that all the following conditions are satisfied for sufficiently small ϵ_0 (in $\left\|\Im(t)\right\|_\infty \leq \epsilon_0$):

$$\begin{gathered}L \geq 2M_0,\\ \min\left(\bar{\theta}'\right) \geq \frac{M_0}{2L},\\ \Upsilon(L,\lambda)\left\|\Im(t)\right\|_\infty + \frac{C_1\lambda_0^{2p-2}}{L^{-p+1}} + \frac{C_2\lambda_0^{2n}}{4L^{-\max(-n,-2p+1,-n-p+2,-2p)}} \leq \frac{\sqrt{2}}{4}\min f_1,\\ \Gamma_2\lambda_0^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)} \leq \frac{1}{2};\end{gathered}$$

then we have, for some real positive constant $\Upsilon_0(L,\lambda)$,

$$\left\|\mathcal{F}\left(\theta-\theta^{m+1}\right)'\right\|_1 \leq \Upsilon_0(L,\lambda)\left\|\Im(t)\right\|_\infty + \Gamma_1\lambda^{2p-2}L^{-p+2} + \tfrac{1}{2}\left\|\mathcal{F}\left(\theta-\theta^m\right)'\right\|_1 . \qquad \square$$

4. Numerical results. In this part, we present examples in which we show the credibility of the STFR method in extracting intrawave IMFs with high accuracy. Needless to say, these examples will show how the proposed method is capable of extracting intrawave IMFs when other methods like EMD/EEMD would perform weakly. This weak performance is even intensified when the data is polluted by noise.

Example 1. In this example, we study the effect of the filter width on one intrawave IMF with constant envelope. Consider the mild intrawave signal

$$x(t) = \cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right).$$

The IF of this signal is obviously $8\pi + 4\sin(4\pi t)$. The strength of frequency modulation is 4, and the frequency of the frequency modulation is 4π. Compared to the carrier frequency, 8π, the modulated frequency is not negligible. Hence, we expect to get incomplete and inaccurate extraction if a short band filter is used. In Figure 1, the extraction is performed by a normal short band filter, with parameter $\lambda = 2$. As can be seen, the IMF extraction error is relatively large, as expected. On the other hand, in Figure 2, a wider filter with parameter $\lambda = 1^+$ is used. As can be perceived from the figures, increasing the length of the filter enhances the extraction drastically. The maximum error is now reduced by 10^{-2}. Mentioned before, a short band filter extracts only one of the harmonics of the intrawave signal. It essentially acts like a one-mode Fourier transform. However, when making the filter wider, more subharmonics are collected as a whole. Consequently, the result looks much better.

Finally, there is an observation that we want to mention regarding Figure 1. We recall that when using a short band filter, the total number of terms that can be collected by the algorithm is $k^a_{max} = \lfloor\frac{\omega}{4p}\rfloor$ and $k^b_{max} = \lfloor\frac{\omega}{4p} + \frac{1}{2}\rfloor$ in the best case for

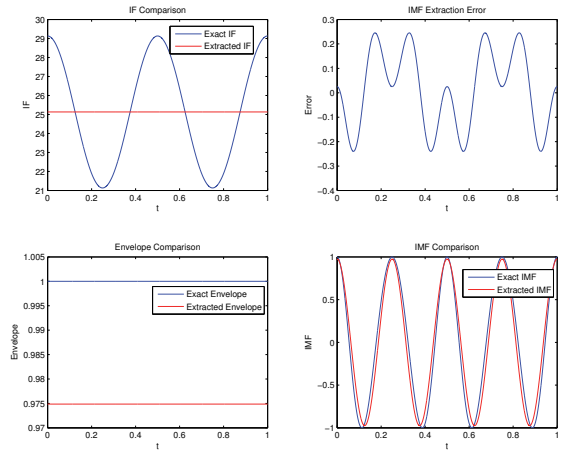

FIG. 1. *Mild intrawave signal vs. short band filter.*

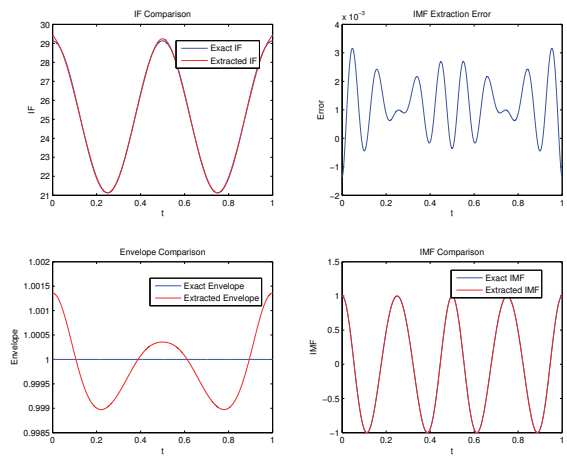

FIG. 2. *Mild intrawave signal vs. wide filter.*

a good constant initial guess. In our case, $k^a_{max} = 0$, $k^b_{max} = 1$. Hence, the main contribution comes from $J_0\left(\frac{\triangle\omega}{p}\right)$, which is 0.9748, in this example. This observation can be seen perfectly well in Figure 1. The plot in the bottom left corner of the figure shows that the extracted envelope is roughly 0.975. Also, the extracted IMF in the bottom right corner of the same figure shows that there is a tiny phase shift, which we can certainly say is coming from $k^b_{max} = 1$.

Example 2. Here, we again take the same intrawave signal from the previous example and mix it with a high-frequency signal

$$v(t) = (1 + 0.3\cos(8\pi t))\left(\cos\left(40\pi t + \frac{5}{2}\sin(2\pi t)\right)\right).$$

This high-frequency signal has small intrawave modulation as well. However, it is not comparable to the frequency modulation of $\cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right)$. The final

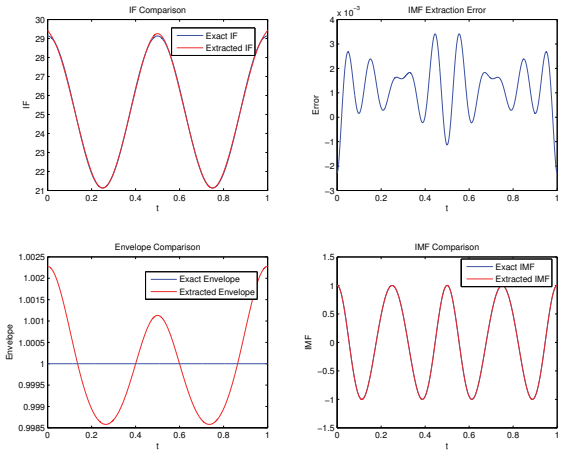

FIG. 3. *Intrawave part of the mixed signal.*

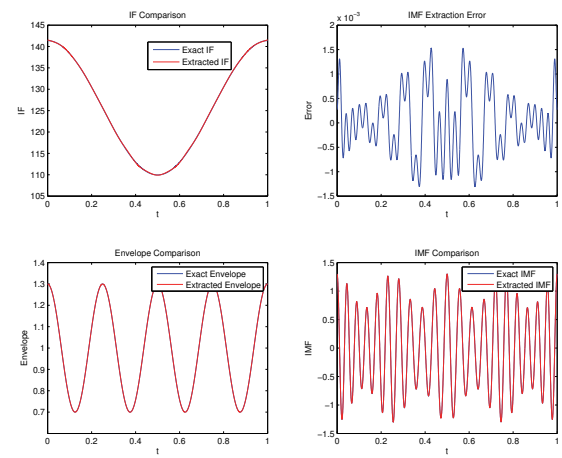

FIG. 4. *High-frequency part of the mixed signal.*

form of the signal is

$$x(t) = \cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right) + (1 + 0.3\cos(8\pi t))\left(\cos\left(40\pi t + \frac{5}{2}\sin(2\pi t)\right)\right).$$

To extract the intrawave part, $\cos\left(8\pi t + \frac{1}{\pi}\sin(4\pi t)\right)$, we use a filter with parameter $\lambda = 1^+$, and for the other part, we use a normal filter with parameter of $\lambda = 2$. Figure 3 shows the result of the extraction for the intrawave part of the signal. If a short band filter were used, this extraction would not be possible. Figure 4 shows the result of the extraction for the high-frequency part. The maximum error in extraction is less than 5×10^{-3}.

There is a subtle point here that we need to emphasize again. As was seen before, if a wide band filter is used, then the IMFs constituting the signal must have a larger scale separation. In other words, in the frequency domain, they should be fairly apart from each other. In a case wherein the signals do not have large-scale separation, they

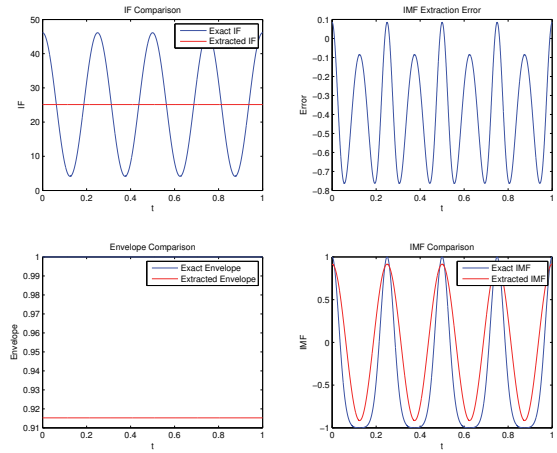

FIG. 5. *Intense intrawave extraction failure with a short band filter.*

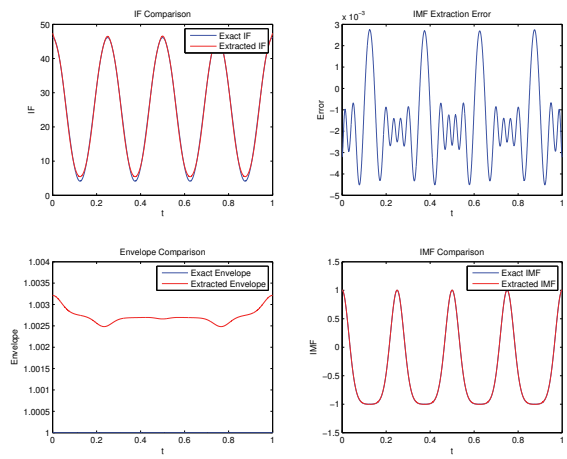

FIG. 6. *Intense intrawave extraction with a wide band filter.*

must be extracted simultaneously. We will demonstrate this in a later example.

Example 3. In this example, we consider a case in which the intrawave modulation is intense. The signal has very sharp peaks and flat troughs. The mathematical form of the signal is

$$x(t) = \cos\left(8\pi t + \frac{21}{8\pi t}\sin(8\pi t)\right).$$

Here, even if the filter parameter is taken to be $\lambda = 1^+$, the results are not yet acceptable; see Figure 5. In order to overcome this issue, we decrease the filter parameter to $\lambda = 0.3$. As seen in Figure 6, the extraction of this intrawave signal with strong frequency modulation is now possible. As seen in the bottom left corner of the figure, the envelope is not recovered exactly; however, the final extraction is acceptable.

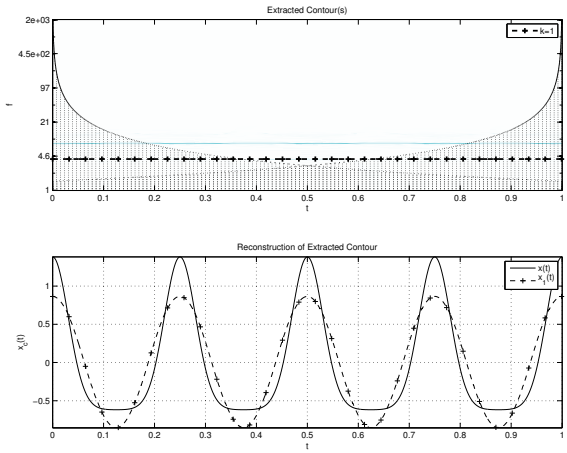

FIG. 7. *Synchrosqueezed wavelet comparison. Top: The frequency spectrum shows that the synchrosqueezed method detects two major frequency trends. Bottom: The first IMF extracted using this analysis is like the first dominant subharmonic.*

The synchrosqueezed wavelet transform (SWT) [6, 17] is another effective adaptive data analysis method for analyzing signals with frequency modulation. Naturally, we would like to apply the SWT to analyze the above signal and compare with the result that we obtained using our method. As we can see from Figure 7, the SWT method tends to decompose the intrawave signal into a number of harmonics. We used a publicly available MATLAB code for the SWT to perform our decomposition. This code can be found in [3]. There are a number of parameters in this code, including the choice of wavelets being used. In our implementation, the Morlet wavelet was used. We realize that the performance of the SWT depends on the choice of the wavelet and the ridge detection method. The parameters that we used in the above comparison may not be the optimal ones, and there may be a more optimal ridge detection method that one can use to achieve better results. The parameters that we presented in the SWT code provided in [3] were the best ones that we could find after many trials and errors. It is very likely that there exists a better choice of wavelet and edge detection methods to be used to obtain a more optimal result than the one presented here.

We continue this example by considering an intense intrawave signal with a changing envelope. See Figure 8. As before, we take $\lambda = 0.3$. Again, the envelope extraction is not accurate; however, the final extraction of the signal is reliable.

We again compare our results with other methods. This time, we use the EMD method. Using EMD, the method would not be able to decompose the intrawave in one piece. The result of this extraction can be seen in Figure 9.

As perceived from this example, our method is more general compared to other adaptive methods. Next, we consider the noise stability of our method.

Example 4. In this example, we investigate the effect of noise on the extraction of an intrawave signal. We take an intrawave signal with strong frequency modulation of the form

$$x(t) = \cos\left(8\pi t + \frac{21}{8\pi t}\sin(8\pi t)\right) + 0.1\mathcal{N}(0,1).$$

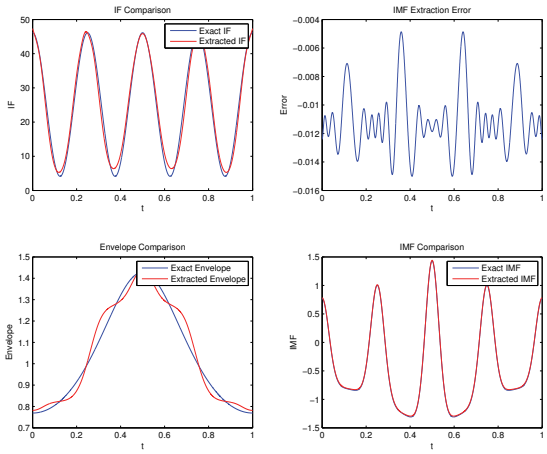

FIG. 8. *Intense intrawave with a nonconstant envelope.*

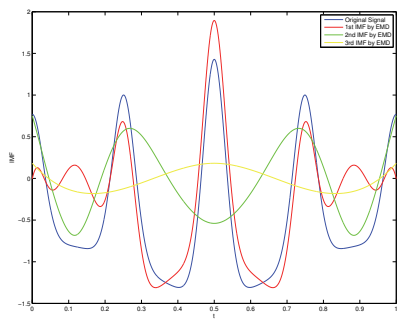

FIG. 9. *EMD extraction result for intrawave signal with a changing envelope. The method decomposes the IS into many IMFs.*

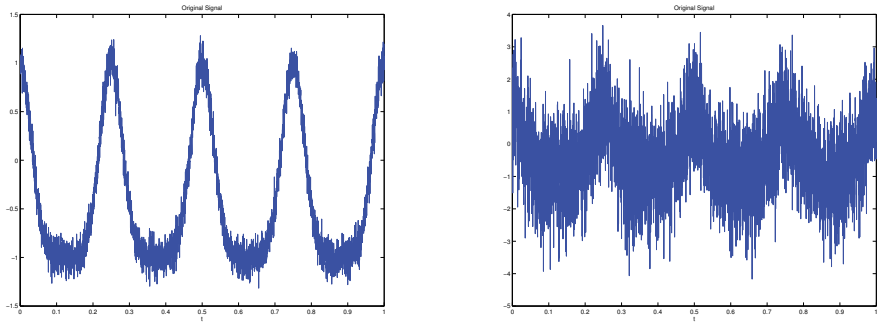

FIG. 10. *Original mildly noisy signal (left) and heavy noisy signal (right).*

Here $\mathcal{N}(0,1)$ is a normal Gaussian random variable. The original noisy signal can be seen in Figure 10. Taking $\lambda = 0.3$, the result of the extraction can be seen in Figure 11. Despite minor error in the extraction of the envelope and the instantaneous

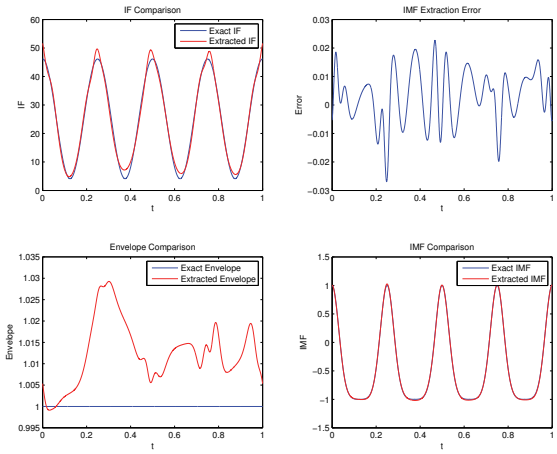

FIG. 11. *Extraction result of the mildly noisy signal.*

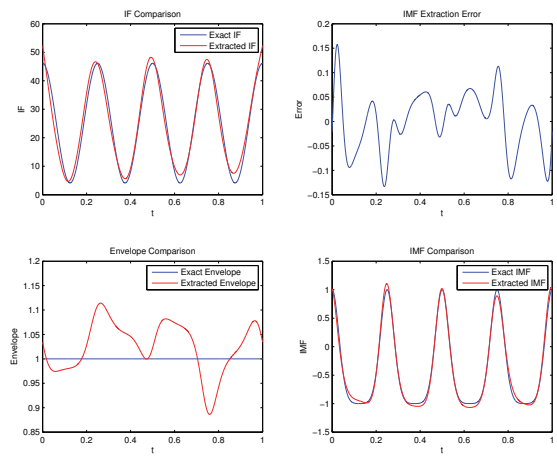

FIG. 12. *Extraction result of the intensely noisy signal.*

frequency, the final extraction is faithful. In this case, the noise perturbation was not large compared to the energy of the signal.

Next, we take

$$x(t) = \cos\left(8\pi t + \frac{21}{8\pi t}\sin(8\pi t)\right) + \mathcal{N}(0,1).$$

This time, the noise is stronger than the previous case. The original signal is shown in Figure 10. As can be seen from the figure, the signal is buried in noise. The result of the extraction using $\lambda = 0.8$ is shown in Figure 12. Again, the result of the extraction, even in the presence of intense noise perturbation, is acceptable. There are errors in the extracted envelope and IF; however, the final extracted IMF is in a good shape.

We would like to point out that if we had taken smaller values of λ, the algorithm would have diverged due to the strong presence of noise. In fact, this divergence was predicted by the convergence theorem on the presence of noise.

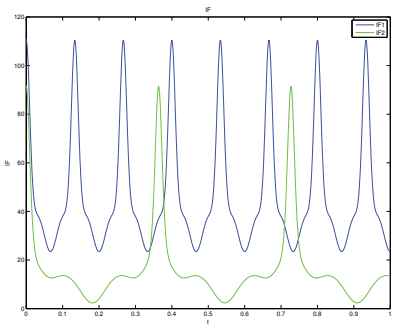

FIG. 13. *IF of both IMFs.*

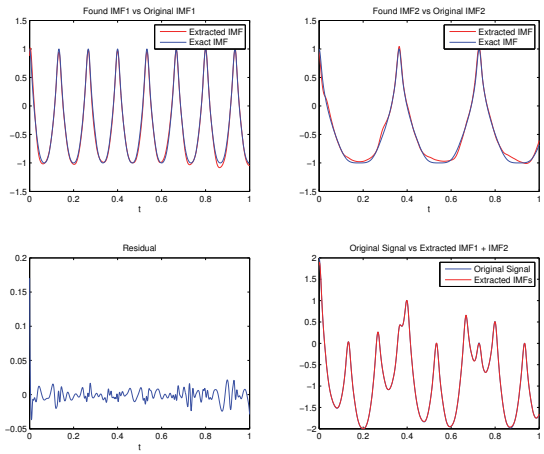

FIG. 14. *Results of the extraction of mixed intrawaves.*

Example 5. This example is part of our ongoing future research. Here, we investigate the extraction of two intrawave signals with strong frequency modulation that do not have well-separated scales. The signal is of the form $f(t) = \cos\theta_1 + \cos\theta_2$. The IFs are shown in Figure 13. As can be seen from the figure, the IMFs have intense intrawave characteristics. At the same time, the mode mixture is apparent in the signal. The signals are not separate in the time-frequency domain. The extraction of such IMFs is challenging. None of the sparse time frequency methods can extract them successfully, to the best of our knowledge.

To overcome this difficulty, we extract them simultaneously using the same algorithm presented in this paper. The results are shown in Figure 14. As can be seen, our method, using a wide band filter combined with the simultaneous extraction, would result in a fairly successful extraction. The error of the extraction is acceptable, as no other method is capable of such performance. The only a priori knowledge that we have used is that there are two IMFs in the original signal. However, this assumption can be dropped. We will present the results of this kind of extraction in more detail in a future paper.

It is not hard to show formally why the last example extraction works. Assume

that the signal can be written as

$$x(t) = \cos\left(\omega_1 t + \frac{\triangle\,\omega_1}{p_1}\sin(p_1 t)\right) + \cos\left(\omega_2 t + \frac{\triangle\,\omega_2}{p_2}\sin(p_2 t)\right). \tag{4.1}$$

Next, consider that we intend to extract both IMFs simultaneously. Take the initial guesses as $\theta_1(t) = \omega_1 t$ and $\theta_2(t) = \omega_2 t$. Using the expansion in (1.7), we can write (4.1) as

$$\begin{aligned} x(\theta_1,\theta_2) = &\left\{J_0\left(\frac{\triangle\,\omega_1}{p_1}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega_1}{p_1}\right)\cos\left(2k\frac{p_1}{\omega_1}\theta_1\right)\right\}\cos\theta_1 \\ &- \left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega_1}{p_1}\right)\sin\left((2k-1)\frac{p_1}{\omega_1}\theta_1\right)\right\}\sin\theta_1 \\ &+ \left\{J_0\left(\frac{\triangle\,\omega_2}{p_2}\right) + 2\sum_{k=1}^{\infty} J_{2k}\left(\frac{\triangle\,\omega_2}{p_2}\right)\cos\left(2k\frac{p_2}{\omega_2}\theta_2\right)\right\}\cos\theta_2 \\ &- \left\{2\sum_{k=1}^{\infty} J_{2k-1}\left(\frac{\triangle\,\omega_2}{p_2}\right)\sin\left((2k-1)\frac{p_2}{\omega_2}\theta_2\right)\right\}\sin\theta_2. \end{aligned} \tag{4.2}$$

The sparse representation of (4.2) suggests the use of definition (1.12) for both θ_1 and θ_2 in a singular form (i.e., with a small λ_0). In a future work, we will analyze this method in more detail.

4.1. Uniqueness issues. One of the most important issues here is the uniqueness of extraction. The theorems, proven in this paper, solely explain that we have a convergent algorithm. However, the uniqueness issue is not addressed. In fact, the uniqueness is a difficult theoretical problem. In this section, we explain insights about how we can handle this topic in implementation and numerical analysis.

Here, we reduce uniqueness into sparsity. In fact, any signal can have many representations.[5] In our STFR methodology, we prefer a representation in which there is the least number of IMFs extracted from the signal. In other words, a sparse representation is the unique representation in STFR numerical terminology. One might ask, What if two different extractions would result in the same number of IMFs? Then, which one is better? The answer is that both of them are acceptable. However, we pick the one in which the components have the smoothest envelopes compared to the IFs.

The most important fact is that the result of a numerical implementation of STFR must be sparse. In other words, we pick the set of IMFs as the constituent blocks of a signal if we cannot find another set that contains fewer IMFs. A preferred extraction is the one that has fewer IMF components. In fact, as we observed, in case of intrawave signals, widening the filter width would help to find the smallest number of IMFs in an extraction.

One should keep in mind that if there is any mode mixture present (the IFs are not separate in the time-frequency domain), it is highly possible to have a nonsparse extraction. In order to tackle this problem, one should find an estimate of the number of mode mixtures and then try to extract the mode mixed IMFs simultaneously. The latter would impose sparsity. The last example of this paper showed this method.

[5] This is a topic of our future research.

Quantitatively, having two extractions with the same number of IMFs, the smoothness measure (SM) can be one of the following:

$$SM_{TV} = \frac{TV\left(\tilde{a}\left(t\right)\right)}{TV\left(\tilde{\dot{\theta}}\left(t\right)\right)},$$

$$SM_{max-local} = \max\left(\frac{\tilde{\dot{a}}\left(t\right)}{\tilde{\dot{\theta}}\left(t\right)}\right).$$

In these definitions, $\tilde{a}\left(t\right)$ is the extracted envelope, $\tilde{\dot{\theta}}\left(t\right)$ is the extracted IF, and TV is the total variation. In case of many IMFs, one can sum up the SMs of all extracted IMFs to judge the quality of extraction.

As we mentioned previously, the question of uniqueness is a difficult one. Our understanding of this subject is still in a very preliminary stage. We will continue to study this question in our future research.

5. Concluding remarks and future work. In this paper, we presented a modified version of the STFR method capable of extracting intrawave signals. Traditionally, the intrawave signals have been challenging problems to analyze. So far, almost all adaptive signal processing methods have difficulty in extracting intrawave signals properly. In this paper, we proposed expanding the length of the band filters used in STFR methods to successfully extract intrawave signals. It was shown theoretically that using a wide filter would reduce the extraction error. We also showed that the algorithm is convergent and stable to noise. Numerical examples demonstrated the effectiveness of the modified STFR method. Our numerical study also showed that our modified STFR method offers superior performance compared with other well-known time-frequency methods, such as the EMD and synchrosqueezed wavelet methods. Currently, these other methods have difficulty in extracting the intrawave signals properly. Our method has a strong mathematical foundation. It offers an effective alternative as an adaptive data analysis method.

We also stated that using a wide band filter in any STFR algorithm can cause some problems if the IMFs are not well separated in the time-frequency domain. This shortcoming was alleviated by extracting mixed (nonseparated) IMFs simultaneously. The last example in this paper clearly showed the merit of simultaneous extraction and a wide band filter in the extraction of mixed intrawave IMFs. However, the latter is still a topic of future work. We intend to show the robustness of this method analytically and also examine the stability of this method in the presence of noise perturbation.

Finally, in our future work, we will also address the nonuniqueness of the IMF representations. In fact, we believe this nonuniqueness is as fundamental as the uncertainty principle. We believe that looking for the sparsest decomposition within some appropriately chosen dictionary would be the most effective way to address the uniqueness issue. We hope to address this issue in a future paper in depth.

Appendix A. Approximating $|\hat{f}_{0,\bar{\theta}^m}(\omega)|$. In this appendix, we will derive a bound for $|\hat{f}_{0,\bar{\theta}^m}(\omega)|$, which is given in (3.4). In the following calculations, we will

bound some of the summations by integrals. Specifically, we have

$$\begin{aligned}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| &= \left|\int_0^1 f_0\left(\bar{\theta}^m\right) e^{-i2\pi\omega\bar{\theta}^m} d\bar{\theta}^m\right| = \left|\int_0^1 f_0\left(\bar{\theta}^m\left(\bar{\theta}\right)\right) e^{-i2\pi\omega\bar{\theta}^m} d\bar{\theta}^m\right| \\ &= \left|\sum_{k\neq 0}\hat{f}_{0,\bar{\theta}}(k)\int_0^1 e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m\right| \\ &\leq \sum_{|\alpha k|>\frac{|\omega|}{\lambda}}\left|\hat{f}_{0,\bar{\theta}}(k)\right| + \sum_{0<|\alpha k|\leq\frac{|\omega|}{\lambda}}\left|\hat{f}_{0,\bar{\theta}}(k)\right|\left|\int_0^1 e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m\right|.\end{aligned}$$

Since we have $|\hat{f}_{0,\bar{\theta}}(k)| \leq |k|^{-p}$ for $k \neq 0$, and using the first two lemmas in section 3, we get

$$\begin{aligned}&\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \\ \leq& C_0\sum_{|k|>\frac{|\omega|}{\alpha\lambda}}|k|^{-p} + C_0\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}|k|^{-p}\left|\int_0^1 e^{i2\pi\bar{\theta}^m(\alpha k-\omega)} e^{ik\frac{\triangle\theta^m}{L}} d\bar{\theta}^m\right| \\ \leq& 2C_0(\alpha\lambda)^{p-1}|\omega|^{-p+1} + 2C_0P_m^n\left(\frac{M_0\lambda}{2\pi}\right)^n|\varepsilon\omega|^{-n}\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}\sum_{j=1}^n|k|^{j-p}\left(\frac{\left\|\mathcal{F}\left(\triangle\theta^m\right)'\right\|_1}{2\pi M_0 L}\right)^j \\ \leq& 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + 2C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \\ &\times\left(\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}\left(\sum_{j=1}^{p-2}|k|^{j-p}\left(\frac{\gamma}{L}\right)^j + \sum_{j=p}^n|k|^{j-p}\left(\frac{\gamma}{L}\right)^j + |k|^{-1}\left(\frac{\gamma}{L}\right)^{p-1}\right)\right) \\ \leq& 2C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + 2C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \\ &\times\left(\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3} + 2\sum_{j=p}^n\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j + 2\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).\end{aligned}$$

In the inequalities above, we have also used the fact that if $0 < k \leq \frac{\omega}{\alpha\lambda}$, then $\frac{1}{\omega-\alpha k} < \frac{\lambda}{\lambda-1}\frac{1}{\omega} < \frac{\lambda}{\varepsilon}\frac{1}{\omega}$ provided $0 < \varepsilon \ll 1$ and $\lambda \geq 1+\varepsilon$. Also we called $\frac{\|\mathcal{F}(\triangle\theta^m)'\|_1}{2\pi M_0} = \gamma$. We recall that in these calculations $n > p$. This finalized the derivation of the bound on (3.4).

Appendix B. Approximating $\left|\hat{a}^{m}_{\bar{\theta}^m}(\omega)\right|$. This approximation (see (3.5)) is essentially the same as the one we derived in the previous part, except that we also use Jensen's inequality. In the main theorem of this section, we extensively use this lemma. So, we present the following lemma to introduce the inequality.

LEMMA B.1 (Jensen's inequality). *If φ is a convex function, and $a_i \in \mathbb{R}^+$ for $i \in \mathbb{N}$, $i > 1$, then*

$$\varphi\left(\frac{\sum a_i x_i}{\sum a_i}\right) \leq \frac{\sum a_i\varphi(x_i)}{\sum a_i}. \tag{B.1}$$

A proof of this lemma can be found in [2]. A consequence of this lemma is that, for a positive integer j, we have

$$2^{-j}\left|1+\frac{k}{L}\right|^{j}=\left|\frac{1}{2}1+\frac{1}{2}\frac{k}{L}\right|^{j}\leq\frac{1}{2}\left(|1|^{j}+\left|\frac{k}{L}\right|^{j}\right). \tag{B.2}$$

This bound is used several times in the next approximations. As mentioned above, we also bound summations by integrals. So, for $\left|\hat{a}^{m}_{\bar{\theta}^{m}}(\omega)\right|$ we have

$$\begin{aligned}\left|\hat{a}^{m}_{\bar{\theta}^{m}}(\omega)\right|=&\left|\int_{0}^{1}a^{m}\left(\bar{\theta}^{m}\right)e^{-i2\pi\omega\bar{\theta}^{m}}d\bar{\theta}^{m}\right|=\left|\int_{0}^{1}a^{m}\left(\bar{\theta}^{m}\left(\bar{\theta}\right)\right)e^{-i2\pi\omega\bar{\theta}^{m}}d\bar{\theta}^{m}\right|\\
=&\frac{1}{2}\left|\int_{0}^{1}\left(e^{i\Delta\theta^{m}}+e^{-i\Delta\theta^{m}}\right)\sum_{k=-\infty}^{\infty}\hat{f}_{1,\bar{\theta}}(k)\,e^{i2\pi\bar{\theta}^{m}(\alpha k-\omega)}e^{ik\frac{\triangle\theta^{m}}{L}}d\bar{\theta}^{m}\right|\\
\leq&\frac{1}{2}\left|\int_{0}^{1}\sum_{k=-\infty}^{\infty}\hat{f}_{1,\bar{\theta}}(k)\,e^{i2\pi\bar{\theta}^{m}(\alpha k-\omega)}e^{i\triangle\theta^{m}\left(\frac{k}{L}+1\right)}d\bar{\theta}^{m}\right|\\
&+\frac{1}{2}\left|\int_{0}^{1}\sum_{k=-\infty}^{\infty}\hat{f}_{1,\bar{\theta}}(k)\,e^{i2\pi\bar{\theta}^{m}(\alpha k-\omega)}e^{i\triangle\theta^{m}\left(\frac{k}{L}-1\right)}d\bar{\theta}^{m}\right|.\end{aligned}$$

Now, we break up the the summation $\sum_{k=-\infty}^{\infty}$ into summations over $k=0$, $|\alpha k|>\frac{|\omega|}{\lambda}$, and $0<|\alpha k|\leq\frac{|\omega|}{\lambda}$ and also use the first and second lemmas in section 3 to get

$$\begin{aligned}\left|\hat{a}^{m}_{\bar{\theta}^{m}}(\omega)\right|\leq&2C_{0}\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_{m}^{n}\left(\frac{M_{0}}{2\pi|\omega|}\right)^{n}\sum_{j=1}^{n}\gamma^{j}\\
&+C_{0}P_{m}^{n}\left(\frac{M_{0}\lambda}{2\pi|\varepsilon\omega|}\right)^{n}\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}|k|^{-p}\sum_{j=1}^{n}\left(\left|1+\frac{k}{L}\right|^{j}+\left|1-\frac{k}{L}\right|^{j}\right)\gamma^{j}\\
\leq&2C_{0}\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_{m}^{n}\left(\frac{M_{0}}{2\pi|\omega|}\right)^{n}\sum_{j=1}^{n}\gamma^{j}\\
&+C_{0}P_{m}^{n}\left(\frac{M_{0}\lambda}{2\pi|\varepsilon\omega|}\right)^{n}\sum_{0<|k|\leq\frac{|\omega|}{\alpha\lambda}}|k|^{-p}\sum_{j=1}^{n}2^{j}\left(1+\left|\frac{k}{L}\right|^{j}\right)\gamma^{j}\\
\leq&2C_{0}\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_{m}^{n}\left(\frac{M_{0}}{2\pi|\omega|}\right)^{n}\sum_{j=1}^{n}\gamma^{j}\\
&+C_{0}P_{m}^{n}\left(\frac{M_{0}\lambda}{2\pi|\varepsilon\omega|}\right)^{n}\frac{2p}{p-1}\sum_{j=1}^{n}2^{j}\gamma^{j}\\
&+C_{0}P_{m}^{n}\left(\frac{M_{0}\lambda}{2\pi|\varepsilon\omega|}\right)^{n}\\
&\times\left(\sum_{j=1}^{p-2}2^{j}\left(\frac{\gamma}{L}\right)^{j}\frac{\pi^{2}}{3}+\sum_{j=p}^{n}2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^{j}+2^{p}\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).\end{aligned}$$

The latter completes the derivation of the bound on $\left|\hat{a}^{m}_{\bar{\theta}^{m}}(\omega)\right|$; see (3.5).

Appendix C. Estimates on $|\triangle a^{m}|$, $|\triangle b^{m}|$. In this appendix, we derive the

bound on $|\triangle a^m|$ (see (3.6)). We first check the term involving $|\hat{f}_{0,\bar{\theta}^m}(\omega)|$. We have

$$
\begin{aligned}
&2\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \\
\le&4\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+4\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}C_0P_m^n\left(\frac{M_0\lambda}{2\pi\,|\varepsilon\omega|}\right)^n\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3} \\
&+8\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}C_0P_m^n\left(\frac{M_0\lambda}{2\pi\,|\varepsilon\omega|}\right)^n\left(\sum_{j=p}^{n}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j+\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right)\right).
\end{aligned}
$$

Here, we can use a trick to bound the summations with integrals. When $\left(1-\frac{1}{\lambda}\right)L_m\le\omega\le\left(1+\frac{1}{\lambda}\right)L_m$, we know that ω will at least start from 1. The latter is due to the fact that $\lambda>1$, and hence $\left(1-\frac{1}{\lambda}\right)L_m>0$. So, we can always find a fixed $\varepsilon>0$ such that $0<|1-\frac{1}{1-\varepsilon}|L_m<\min\left(\left(1-\frac{1}{\lambda}\right)L_m,1\right)$. We can take this ε to be the same as the one we used before. So, the whole inequality, using the fact that $L_m=\alpha L$, becomes

$$
\begin{aligned}
&2\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \\
\le&\frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2}+\frac{4\pi^2C_0P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\alpha^{-n+1}}{3\,(n-1)}L^{-n+1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^j \\
&+8C_0\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^nP_m^n\alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}L^{-p+2}\frac{1}{p+2}\sum_{j=p}^{n}\left(\frac{\gamma}{L}\right)^j \\
&+8C_0P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\left(\frac{\gamma}{L}\right)^{p-1}\left(\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+1}}{n-1}+\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+2}}{\alpha\lambda\,(n-2)}\right).
\end{aligned}
$$

Here, we use the assumption that $\gamma\le\frac{1}{4}$. This assumption would remain intact throughout the steps for large enough L. In fact, the condition $\gamma\le\frac{1}{4}$ would remain intact for all iterations. In other words, when there is a contraction on $\|\mathcal{F}(\theta-\theta^{m+1})'\|_1$, this term would remain bounded. Hence, if $\frac{\|\mathcal{F}(\triangle\theta^0)'\|_1}{2\pi M_0}\le\frac{1}{4}$ for the first iteration, it will remain bounded by $\frac{1}{4}$ for all iterations. Hence, using this, we have $\sum_{j=1}^{p-2}\left(\frac{\gamma}{L}\right)^j\le\left(\frac{\gamma}{L}\right)\sum_{j=1}^{\infty}\left(\frac{\gamma}{L}\right)^{j-1}\le\frac{4\gamma}{3L}$. For the second sum, we use the same trick, namely $\sum_{j=p}^{n}\left(\frac{\gamma}{L}\right)^j\le\left(\frac{\gamma}{L}\right)^p\sum_{j=1}^{\infty}\left(\frac{\gamma}{L}\right)^{j-1}\le(4L)^{-p}\frac{4\gamma}{3L}$. So, we get

$$
\begin{aligned}
&2\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \\
\le&\frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2}+\frac{4\pi^2C_0P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\alpha^{-n+1}}{3\,(n-1)}L^{-n+1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}\frac{4\gamma}{3L} \\
&+8C_0\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^nP_m^n\alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}L^{-p+2}\frac{1}{p+2}(4L)^{-p}\frac{4\gamma}{3L} \\
&+8C_0P_m^n\left(\frac{M_0\lambda}{2\pi\varepsilon}\right)^n\left(\frac{\gamma}{L}\right)^{p-1}\left(\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+1}}{n-1}+\frac{\left(\alpha L\left(\frac{\varepsilon}{1-\varepsilon}\right)\right)^{-n+2}}{\alpha\lambda\,(n-2)}\right).
\end{aligned}
$$

Simplifying further, we get

$$
\begin{aligned}
&2\sum_{(1-\frac{1}{\lambda})L_m\le\omega\le(1+\frac{1}{\lambda})L_m}\left|\hat{f}_{0,\bar{\theta}^m}(\omega)\right| \\
&\le \frac{4C_0\alpha}{p-2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-p+2}\lambda^{p-1}L^{-p+2}+\frac{16}{9}\pi^2C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n+1}}{n-1}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}L^{-n}\lambda^n\gamma \\
&\quad+\frac{8}{3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+2}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}\frac{4^{-p+1}}{p+2}L^{-2p+1}\lambda^n\gamma \\
&\quad+8C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n+1}\left(\frac{\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+1}}{n-1}L^{-n-p+1}\lambda^n\gamma+\frac{\left(\frac{\varepsilon}{1-\varepsilon}\right)^{-n+2}}{n-2}L^{-n-p+2}\lambda^{n-1}\gamma\right).
\end{aligned}
$$

Now, it is time to check the term involving $\left|\hat{a}^m_{\bar{\theta}^m}(k)\right|+\left|\hat{b}^m_{\bar{\theta}^m}(k)\right|$. We have

$$
\begin{aligned}
&\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}\left(\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right|+\left|\hat{b}^m_{\bar{\theta}^m}(\omega)\right|\right) \\
&\le\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}\left(4C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1}+2\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi|\omega|}\right)^n\sum_{j=1}^n\gamma^j\right) \\
&\quad+2\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\frac{2p}{p-1}\sum_{j=1}^n2^j\gamma^j \\
&\quad+2\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\sum_{j=1}^{p-2}2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3} \\
&\quad+2\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n\sum_{j=p}^n2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j \\
&\quad+2\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}C_0P_m^n\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right).
\end{aligned}
$$

In this case, due to the presence of the term $\left(2-\frac{1}{\lambda}\right)L_m\le\omega\le\left(2+\frac{1}{\lambda}\right)L_m$, the substitution of the sum with an integral is much easier. Since $\lambda>1$, we have $1<2-\frac{1}{\lambda}<2$. As a result, $\alpha L=L_m<\left(2-\frac{1}{\lambda}\right)L_m\le\omega$. Hence, for any $\zeta>1$ we have

$$
\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}\frac{1}{\omega^\zeta}\le\sum_{\alpha L<\omega}\frac{1}{\omega^\zeta}=\frac{1}{(\alpha L)^\zeta}\sum_{1<\frac{\omega}{\alpha L}}\frac{1}{\left(\frac{\omega}{\alpha L}\right)^\zeta}=\frac{1}{(\alpha L)^\zeta}\sum_{1<k}\frac{1}{k^\zeta}\le\frac{1}{(\alpha L)^\zeta}\int_1^\infty\frac{1}{k^\zeta}=\frac{(\alpha L)^{-\zeta}}{\zeta-1}.
$$

Consequently, we get

$$
\begin{aligned}
&\sum_{(2-\frac{1}{\lambda})L_m\le\omega\le(2+\frac{1}{\lambda})L_m}\left(\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right|+\left|\hat{b}^m_{\bar{\theta}^m}(\omega)\right|\right) \\
&\le\frac{4C_0}{p-2}\lambda^{p-1}L^{-p+1}+\frac{8}{3}\left|\hat{f}_{1,\bar{\theta}}(0)\right|P_m^n\left(\frac{M_0}{2\pi}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\gamma+\frac{16p}{p-1}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\lambda^n\gamma \\
&+\frac{8\pi^2}{3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-1}\lambda^n\gamma+2^{-p+4}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n}L^{-2p}\lambda^n\gamma \\
&+2^{-p+3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-p+1}\lambda^n\gamma+2^{-p+3}C_0P_m^n\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{(n-2)\lambda}L^{-n-p+2}\lambda^n\gamma.
\end{aligned}
$$

Finally, it is time to find the bound on the term that involves $\left|\hat{a}^m_{\bar{\theta}^m}(k)\right|$. This term appears in the bound of $|\triangle a^m|$. We have

$$\begin{aligned}\sum_{|\omega|>\frac{L_m}{\lambda}}\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right| \leq & 2\sum_{|\omega|>\frac{L_m}{\lambda}} C_0\left(\frac{\alpha\lambda}{|\omega|}\right)^{p-1} + \sum_{|\omega|>\frac{L_m}{\lambda}}\left|\hat{f}_{1,\bar{\theta}}(0)\right| P^n_m\left(\frac{M_0}{2\pi|\omega|}\right)^n \sum_{j=1}^n \gamma^j \\ & + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P^n_m\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \frac{2p}{p-1}\sum_{j=1}^n 2^j\gamma^j \\ & + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P^n_m\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \sum_{j=1}^{p-2} 2^j\left(\frac{\gamma}{L}\right)^j\frac{\pi^2}{3} \\ & + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P^n_m\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n \sum_{j=p}^{n} 2^{j+1}\left|\frac{\omega}{\alpha\lambda}\right|^{j-p+1}\left(\frac{\gamma}{L}\right)^j \\ & + \sum_{|\omega|>\frac{L_m}{\lambda}} C_0 P^n_m\left(\frac{M_0\lambda}{2\pi|\varepsilon\omega|}\right)^n 2^p\left(\frac{\gamma}{L}\right)^{p-1}\left(1+\frac{|\omega|}{\alpha\lambda}\right).\end{aligned}$$

Again we have for any $\zeta > 1$

$$\sum_{|\omega|>\frac{L_m}{\lambda}}\frac{1}{|\omega|^\zeta} = 2\left(\frac{\lambda}{\alpha L}\right)^\zeta \sum_{\frac{\lambda\omega}{\alpha L}>1}\frac{1}{\left(\frac{\lambda\omega}{\alpha L}\right)^\zeta} = 2\left(\frac{\lambda}{\alpha L}\right)^\zeta \sum_{1<k}\frac{1}{k^\zeta} \leq 2\left(\frac{\lambda}{\alpha L}\right)^\zeta \int_1^\infty \frac{1}{k^\zeta} = 2\left(\frac{\lambda}{\alpha L}\right)^\zeta \frac{1}{\zeta-1}.$$

Hence, we can again bound the summations with integrals. So, we get

$$\begin{aligned}\sum_{|\omega|>\frac{L_m}{\lambda}}\left|\hat{a}^m_{\bar{\theta}^m}(\omega)\right| \leq & \frac{4C_0}{p-2}\lambda^{2p-2}L^{-p+1} + \frac{8}{3}\left|\hat{f}_{1,\bar{\theta}}(0)\right| P^n_m\left(\frac{M_0}{2\pi}\right)^n \frac{\alpha^{-n}}{n-1}L^{-n}\lambda^n\gamma \\ & + \frac{16p}{p-1}C_0P^n_m\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n}\lambda^{2n}\gamma \\ & + \frac{8\pi^2}{3}C_0P^n_m\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-1}\lambda^{2n}\gamma \\ & + 2^{-p+5}C_0P^n_m\left(\frac{M_0}{2\pi\varepsilon}\right)^n\alpha^{-n}L^{-2p-1}\lambda^{2n-2}\gamma \\ & + 2^{-p+5}C_0P^n_m\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{n-1}L^{-n-p+1}\lambda^{2n}\gamma \\ & + 2^{-p+5}C_0P^n_m\left(\frac{M_0}{2\pi\varepsilon}\right)^n\frac{\alpha^{-n}}{(n-2)\lambda}L^{-n-p+2}\lambda^{2n-2}\gamma.\end{aligned}$$

In all of the previous estimates, each term can be bounded by either $\lambda^{2p-2}L^{-p+2}$ or $\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma$. Thus, we prove that

$$|\triangle a^m| \;\leq\; C_1\lambda^{2p-2}L^{-p+2} + C_2\lambda^{2n}L^{\max(-n,-2p+1,-n-p+2,-2p)}\gamma\,.$$

This completes the derivation of the bound for (3.6). This bound can be used, as it is, for the term that includes $\sum_{|k|>\frac{L_m}{\lambda}}|\hat{b}^m_{\bar{\theta}^m}(k)|$ in (3.7).

REFERENCES

[1] C. Blatter, *Wavelets: A Primer*, Universities Press, New Delhi, 2003.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[3] E. Brevdo, *Synchrosqueezing Wavelet Transform*, https://web.math.princeton.edu/~ebrevdo/synsq/, 2012.

[4] E.J. Candes and T. Tao, *Near-optimal signal recovery from random projections: Universal encoding strategies?*, IEEE Trans. Inform. Theory, 52 (2006), pp. 5406–5425.

[5] S.S. Chen, D.L. Donoho, and M.A. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comput., 20 (1998), pp. 33–61.

[6] I. Daubechies, J. Lu, and H. Wu, *Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 243–261.

[7] T.Y. Hou and Z. Shi, *Adaptive data analysis via sparse time-frequency representation*, Adv. Adapt. Data Anal., 3 (2011), pp. 1–28.

[8] T.Y. Hou and Z. Shi, *Data-driven time-frequency analysis*, Appl. Comput. Harmon. Anal., 35 (2013), pp. 284–308.

[9] T.Y. Hou, Z. Shi, and P. Tavallali, *Convergence of a data-driven time-frequency analysis method*, Appl. Comput. Harmon. Anal., 37 (2014), pp. 235–270.

[10] B. Huang and A. Kunoth, *An optimization based empirical mode decomposition scheme*, J. Comput. Appl. Math., 240 (2013), pp. 174–183.

[11] C. Huang, L. Yang, and Y. Wang, *Convergence of a convolution-filtering-based algorithm for empirical mode decomposition*, Adv. Adapt. Data Anal., 1 (2009), pp. 561–571.

[12] N.E. Huang and S.S. Shen, *Hilbert-Huang Transform and Its Applications*, Interdiscip. Math. Sci. 5, World Scientific, Hackensack, NJ, 2005.

[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), pp. 903–995.

[14] N.E. Huang, Z. Wu, S.R. Long, K.C. Arnold, X. Chen, and K. Blank, *On instantaneous frequency*, Adv. Adapt. Data Anal., 1 (2009), pp. 177–229.

[15] S.G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1999.

[16] S.G. Mallat and Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Trans. Signal Process., 41 (1993), pp. 3397–3415.

[17] G. Thakur, E. Brevdo, N.S. Fučkar, and H. Wu, *The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications*, Signal Process., 93 (2013), pp. 1079–1094.

[18] B. van der Pol, *Frequency modulation*, Proc. Inst. Radio Eng., 18 (1930), pp. 1194–1205.

[19] H. Wu, *Instantaneous frequency and wave shape functions* (I), Appl. Comput. Harmon. Anal., 35 (2013), pp. 181–199.

[20] Z. Wu and N.E. Huang, *A study of the characteristics of white noise using the empirical mode decomposition method*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci., 460 (2004), pp. 1597–1611.

[21] Z. Wu and N.E. Huang, *Ensemble empirical mode decomposition: A noise-assisted data analysis method*, Adv. Adapt. Data Anal., 1 (2009), pp. 1–41.