Enhanced Time-Frequency Representation and Mode Decomposition

Haijian Zhang and Guang Hua

Abstract—Time-frequency representation (TFR) allowing for mode reconstruction plays a significant role in interpreting and analyzing the nonstationary signal constituted of various modes. However, it is difficult for most previous methods to handle signal modes with closely-spaced or spectrally-overlapped instantaneous frequencies (IFs) especially in adverse environments. To address this issue, we propose an enhanced TFR and mode decomposition (ETFR-MD) method, which is particularly adapted to represent and decompose multi-mode signals with close or crossing IFs under low signal-to-noise ratio (SNR) conditions. The emphasis of the proposed ETFR-MD is placed on accurate IF and instantaneous amplitude (IA) extraction of each signal mode based on short-time Fourier transform (STFT). First, we design an initial IF estimation method specifically for the cases involving crossing IFs. Further, a low-complexity mode enhancement scheme is proposed so that enhanced IFs better fit underlying IF laws. Finally, the IA extraction from signal’s STFT coefficients combined with the enhanced IFs enables us to reconstruct each signal mode. In addition, we derive mathematical expressions that reveal optimal window lengths and separation interference of our method. The proposed ETFR-MD is compatible with previous related methods, thus can be regarded as a step toward a more general time-frequency representation and decomposition method. Experimental results confirm the superior performance of the ETFR-MD when compared to a state-of-the-art benchmark.

Index Terms—Time-frequency representation, mode decomposition, multi-mode signals, closely-spaced instantaneous frequencies, spectrally-overlapped instantaneous frequencies.

I. INTRODUCTION

TIME-frequency representation (TFR) of nonstationary signals constituting of different modes has received considerable research attention. More importantly, separation of multi-mode signals, often referred to as mode decomposition, has also been studied for several decades. In practical applications, it is desirable to decompose the signal of interest into its constituent modes, each pertaining to an intrinsic component. Therefore, many signal processing challenges boil down to the problem of signal representation and decomposition, which is still under investigation as well in methodological aspects as in concrete applications. The focus of this paper is put on representing and decomposing multi-mode frequency-modulated (FM) signals with closely-spaced or spectrally-overlapped instantaneous frequencies (IFs), which is a difficult problem yet to be adequately resolved.

A wealth of information could be found in the time-frequency (TF) analysis literature [1]–[4], such as linear and quadratic TFRs [5], polynomial TFRs [6], [7], etc. Most of recent studies developed so far are low-order TFRs, mainly including linear short-time Fourier transform (STFT), linear fractional Fourier transform (FRFT) [8], [9], quadratic pseudo Wigner-Ville distribution (PWVD) and classical Cohen’s class TFRs [10]. Despite high TF resolution of nonlinear TFRs, they generally suffer from cross-term (CT) interference and irreversible TF transform, which cause a serious barrier to mode reconstruction. Prior studies that have noted the importance of mode deconstruction emphasize on linear TFRs due to their low computation cost and invertibility [11].

The theory of linear TFRs has been investigated quite intensively during the past decades. Wavelet transform (WT) [12], which is regarded as a classical linear TFR, employs adaptive window sizes to improve TF resolution compared to linear STFT. Signal representation was afterwards improved by implementing post-processing based on WT or STFT. The reassignment (RS) methods [13], [14] assign the average energy in certain domains to the gravity center of energy distributions to gain IF trajectories. However, a major problem associated with RS related methods is the infeasibility of mode reconstruction. Further, the synchrosqueezing transform (SST) [15] derived on WT [16] or on STFT [17] squeezes TF coefficients into the IF trajectory only in frequency direction instead of in both time and frequency directions for RS, with the aim of reconstructing signal modes. In recent years, advantages to mode representation and reconstruction have resulted in an increasing interest in the development of SST-based methods [18]–[21]. Inspired by the SST, Yu et al. proposed the synchroextracting transform (SET) [22], which employs a newly developed synchroextracting technique to generate a more energy concentrated TFR, and meanwhile the invertible SET allows for mode reconstruction. Thanks to the simplicity and immunity to CTs of linear TFRs, Abdoush et al. [23] developed two linear TF transforms, based on which an adaptive IF estimation method for noisy multi-mode signals was implemented.

Although the above studies work remarkably well, they have to rely on certain separation condition, indicating that the multi-mode signals to be addressed should have non-overlapped and well separated IFs. This requirement leads to the problem of how to deal with multi-mode signals with spatially-adjacent or crossing IF laws especially when the signal-to-noise ratio (SNR) level is exceedingly low, e.g., the illustrative three signals in Fig. 1. Though application of TFRs in the filed of mode decomposition has proliferated in recent years, effort in analyzing complicated signals in strong noise environment is lacking. To implement mode decomposition, accurate IF estimation of each mode is thought of as the key factor. Diverse IF estimation methods have been well explored and developed. The pioneer work of IF estimation can be traced to the works for mono-mode signals [24], and then developed for non-overlapped multi-mode signals [25]–[30] and overlapped multi-mode signals [31]–[34]. While existing

H. Zhang and G. Hua are with the School of Electronic Information, Wuhan University, Wuhan 430072, China.
IF estimation methods have achieved a certain effect, it is not a trivial task to match their ability for complicated signals with crossing IFs. Thus, more works should consider the extension of a general IF estimator to deal with spectrally-overlapped multi-mode signals.

In this paper, we propose an enhanced TFR and mode decomposition (ETFR-MD) method, with the objective of addressing the multi-mode signals consisting of multiple closely-spaced or spectrally-overlapped modes. A major concern of the proposed ETFR-MD is to respectively reconstruct IF and instantaneous-amplitude (IA) of each mode. To achieve this, we firstly design an initial multi-mode IF estimation method based on STFT instead of on SST-based TFRs [30], [35], [36]. To compensate fluctuating or unsmooth IF estimates, an iterative filtering method is then developed to improve IF estimation accuracy [37]. On the other hand, the IA of each mode is extracted from STFT coefficients with a Gaussian window. The combination of enhanced IFs and extracted IAs accomplishes the representation and decomposition of all signal modes. Finally, the proposed ETFR-MD method is verified and justified through experimental study.

The structure of this paper is organized as follows. In Section II, we introduce the signal model and general problem. This is followed by a description of the proposed ETFR-MD, and a detailed presentation of how IF and IA of each mode are recovered in the STFT framework is provided in Section III. We then derive in Section IV the expressions of optimal window lengths and conduct interference analysis, which are necessary for the understanding of our method. Section V presents the experimental results with the emphasis on multi-mode signals consisting of close or crossing IFs. Finally, the ETFR-MD method is summarized in Section VI.

II. SIGNAL MODEL AND GENERAL PROBLEM

Without loss of generality, we consider noisy multi-mode signals with \( I \) time-varying modes:

\[
s[n] = \sum_{i=1}^{I} x_i[n] + \epsilon[n]
\]

where \( I \) is the number of modes, \( \rho_i[n] \) is the IA of the \( i \)-th mode, and \( T \) denotes sampling interval. The IF of the \( i \)-th mode, \( f_i[n] \), is defined by the derivative of the phase of its corresponding analytic signal. The random phase \( \theta_i \) corresponding to the \( i \)-th mode is uniformly distributed on \([0, 2\pi)\), and \( \epsilon[n] \) is the zero-mean additive white noise. The signal model in (1) is quite general to accommodate a variety of practical signals.

Real-world signals are often received with arbitrary number and type of modes, indicating that the IF curves might be closely distributed or even intersecting. Instead of giving a narrowly defined model which does not fit well with practical scenarios, we attempt to relax requirement of the signal defined in (1). For the sake of concentrating on dealing with the complicated cases in Fig. 1 the signal to be handled is assumed to meet the following assumptions: i) the number of signal modes is unknown; ii) the signal might be contaminated by strong noise; iii) IF laws of different modes could be arbitrarily distributed, i.e., sometimes IF intersection occurs.

The goal of the proposed ETFR-MD method is to respectively enhance and recover IF and IA information of each mode, i.e., \( \rho_i[n] \) and \( f_i[n] \), \( i = 1, \ldots, I \).

III. ENHANCED TIME-FREQUENCY REPRESENTATION AND MODE DECOMPOSITION

The basic process of the proposed ETFR-MD method is depicted in Fig. 2 wherein the noisy signal \( s[n] \) defined in (1) serves as the input, and the output of ETFR-MD could be either enhanced TFR or decomposed signal modes (in highlighted round boxes). Note that our emphasis is placed on the enhancement of IF estimation (the gray area in the figure), by which the fluctuating property of initial IF estimate can be mitigated. An additional benefit of IF enhancement is that the IF information of each mode can be individually recovered with almost negligible interference from other modes. Afterwards, the IA information of each mode is extracted from Gaussian window based STFT, which offers a simple solution to the IA recovery.
In the following subsections, we first introduce an effective IF estimation method to solve complex situations involving close and crossing IFs. Second, the initial IF estimation is refined by implementing an iterative filtering method, so that the IFs of all signal modes are enhanced. Last, the enhanced IFs and the extracted IAs derived from STFT are combined for the purpose of mode representation and decomposition.

A. Initial IF Estimation

In the literature, few works have been reported to deal with multi-mode signals with overlapped IFs. To solve this problem, we extend the work of Djurović and Stanković [27] originally designed for non-overlapped signal modes, and propose a smooth-and-curvature constraint based IF estimation method for both non-overlapping and overlapping cases, aiming at finding a nonparametric STFT-based solution without requiring prior knowledge of signal's IF laws. To be specific, three penalty functions are considered, and Fig. 3 exhibits an illustrative example to explain the three IF constraints in the presence of crossing IFs and low SNR.

1) Penalty for IF Energy: Current IF estimation methods can be classified as phase differencing, signal modeling, phase modeling and TFR based methods, which can be called categorized into parametric and nonparametric ones. Parametric methods are complicated and time-consuming in the case of distinct IF laws, furthermore, they are only suitable for IF estimation when the signal model is known. In most situations, the IF exhibits unforeseen variations and the parametric model is unknown. Recently, some nonparametric methods were proposed, among which a common way is to track IF curves by detecting the maxima positions of all time-slices in STFT plane, which has been demonstrated to be useful for IF tracking.

2) Penalty for IF Continuity: In the case of low SNRs, the fluctuation in IF estimates becomes significant, and spurious estimates (outliers) far from true IFs frequently occur. Therefore, the continuous property is incorporated into the IF detection to punish improbable outliers and generate continuous frequency estimates. We combine the TF energy maxima detection with minimization of IF variations between consecutive IF samples, and the basic idea comes from the problem of edge-following in digital image processing. The main criteria are: IF should pass through as many as possible points of STFT with the highest magnitudes, while IF variation between two consecutive points should not be too fast. To relieve computational burden, the IF estimation method can be realized recursively as an instance of the Viterbi algorithm.

3) Penalty for IF Ridge Curvature: Nevertheless, only considering the energy and continuity constraints will not work well for multi-mode signals with IF intersections. To tackle this issue, our method not only considers continuous property between consecutive samples, but also minimizes the effect of large curvature of estimated IF segments (e.g., penalty 3 in Fig. 3) by looking at the overall IF trajectory structure without prior knowledge of IF laws. This extended method enables a smooth IF detection for multi-mode signals with non-overlapping or overlapping IF laws, and is also valid for mono-mode signals.

According to the above analysis, the essential idea of our method is to find an optimal path from the beginning time-slice to the ended time-slice in STFT plane. This optimal path corresponds to the expected IF estimate which minimizes the following expression:

$$\text{IF}[n] = \arg \min_{p[n] \in \mathbf{P}} \left\{ \sum_{n=0}^{N-1} \mathcal{P}_1 \left( \text{STFT}_x[n, p[n]] \right) + \sum_{n=0}^{N-2} \mathcal{P}_2 \left( p[n], p[n+1] \right) + \sum_{s=1}^{S-1} \mathcal{P}_3 \left( g[s], g[s+1] \right) \right\},$$

where \(n \in [0, N-1]\) denotes the time index, and \(N\) is the number of time-slices, \(p[n]\) denotes all the possible paths belong to a set \(\mathbf{P}\). Assuming every path can be approximated with a series of linear segments (red segments in Fig. 3), and \(S\) is the number of IF segments. \(g[s]\) denotes the gradient of the \(s\)-th segment. \(\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3\) correspond to the three penalty functions, which are mathematically defined as follows:

- \(\mathcal{P}_1\): \(j \in [1, N_{\text{in}}]\), and \(N_{\text{in}}\) is the number of fre-
Algorithm 1: Initial IF Estimation.

Input: STFTs; \( c_1, c_2, \Delta_1, \Delta_2, \delta, \varepsilon_0 \);

1 Step 1: Applying the optimal expression in (2) to find the first optimal path, i.e., \( \text{IF}^{(i)}[n] \), \( i = 0 \), which corresponds to the strongest signal mode;

2 Step 2: Generating a new STFT by setting zero value in the TF region around the detected optimal path by Step 1, i.e., \( \text{IF}^{(i)}[n] - \delta, \text{IF}^{(i)}[n] + \delta \leftarrow 0 \);

3 Step 3: If the energy in the remaining STFT is less than a predefined threshold value \( \varepsilon_0 \), terminate the procedure. Otherwise, repeat Step 1 and Step 2 to obtain the next optimal path \( \text{IF}^{(i)}[n] \), \( i = 1, \ldots, I \).

Output: \( \hat{f}_i[n] \leftarrow \text{IF}^{(i)}[n] \), \( i = 1, \ldots, I \).

Since the STFTs are sorted by a nonincreasing sequence: \( \text{STFT}_s[n, f_1] \geq \text{STFT}_s[n, f_2] \geq \ldots \geq \text{STFT}_s[n, f_m] \), this implies that the stronger the value of \( \text{STFT}_s[n, f_j] \), a smaller value is assigned to the penalty function \( P_1 \).

- \( P_2 \): In traditional TF maxima detection methods, each IF point at a specific time-slice is independently detected from other time-slices, which omits the dependency among consecutive IF points. Since the IF law of an FM mode should be a smooth curve, we merge the maxima detection with the smooth constraint describing the transition rule of adjacent instants, which is mathematically quantified by the penalty parameter \( c_1 \) and the tolerable frequency change \( \Delta_1 \).

- \( P_3 \): This penalty function is mainly used to avoid estimation errors caused by IF intersections. As shown in Fig. 5, we calculate the curvature of every two adjacent IF segments, and ensure correct IF estimation of signal modes by penalizing the combination of linear segments with large curvature. The penalty parameter \( c_2 \) and the tolerable curvature change \( \Delta_2 \) are determined according to the specific application at hand.

The procedure of the initial IF estimation method for multimode signals is summarized in Algorithm 1. The determination of various parameters and the reasons for their choice will be discussed briefly in Section V.

B. Enhanced IF Estimation

Despite the effectiveness of Algorithm 1 for complicated signals, the resultant IF curves are probably unsmooth or fluctuating especially in strong noise environments. Inaccurate IF estimation is the main factor causing mode reconstruction error, i.e., random IF deviation from true IF trajectories would have a serious effect for recovering underlying signal modes, which motivates us to further enhance the initial IF information obtained in Section III.A.

In this subsection, the theoretical foundation, under which an iterative IF enhancement method is designed, is described. To begin with we provide a brief background on the sinusoidal time-frequency distribution [38], [39], based on which a kernel phase averaging is explored for signal filtering and separation with low computational complexity. Detailed analysis on optimal filtering window and separation interference will be discussed in later section.

1) Sinusoidal Time-Frequency Distribution (STFD): The STFD is particularly designed for a sinusoidal frequency modulated (SFM) signal \( z[n] \) with a sinusoidal IF \( f[n] = \rho \cos(2 \pi f n T + \theta) \), defined as below:

\[
z[n] = \exp \left\{ j 2 \pi T \sum_{m=0}^{n} \rho \cos (2 \pi f n T + \theta) \right\} = \exp \left\{ j \frac{p}{f} \sin (2 \pi f n T + \theta) \right\}.
\]

where \( f \) means a constant modulation frequency. The computation of STFD is based on a kernel function adapted to the SFM signal \( z[n] \) [38]:

\[
\text{STFD}_z[n, v] = \mathcal{F} \{ \mathcal{K}_z[n, l] \},
\]

where \( n \) and \( v \) denote time index and frequency index, \( \mathcal{F} \) denotes the Fourier transform, and \( \mathcal{K}_z[n, l] \) is the kernel function:

\[
\mathcal{K}_z[n, l] = \prod_{p=1}^{P} \left\{ z[n + l + a_p b_p \rho \pi n - l - a_p b_p \rho \pi] \right\},
\]

where * denotes the conjugate operator. It is expected that the STFD can optimally concentrate the signal energy along its sinusoidal IF, i.e.,

\[
\mathcal{F} \{ \mathcal{K}_z[n, l] \} = \delta (v - \rho \cos(2 \pi f n T + \theta)),
\]

which is equivalent to:

\[
\Theta \{ \mathcal{K}_z[n, l] \} = 2 \pi f T \cos(2 \pi f n T + \theta),
\]

where \( \delta (\cdot) \) denotes the Dirac delta function, and \( \Theta \{ \cdot \} \) denotes the phase operator. One of the solutions to (7) can be obtained by using simple trigonometric relations when \( P = 2 \), which gives

\[
\begin{align*}
[a_1 = 0, \quad a_2 = \frac{f_T}{2f}] ;
\end{align*}
\]

\( b_1 = \pi f T \sin(2 \pi f T), \quad b_2 = \pi f T \cos(2 \pi f T) \);

where \( [\cdot] \) denotes the rounding operator, and \( f_s = \frac{1}{f_T} \) is the sampling rate. Therefore, one of the possible kernel functions for STFD in [4] and [5] is derived as below:

\[
\mathcal{K}_z[n, l] = \left\{ z[n + l] z^* [n - l] \right\}^{\pi f T \sin (2 \pi f T)} \left\{ z[n + l + \left[ \frac{f_s}{4f} \right] z^* [n - l - \left[ \frac{f_s}{4f} \right]] \right\}^{\pi f T \cos (2 \pi f T)}.
\]

It is noted that the kernel \( \mathcal{K}_z[n, l] \) of STFD is closely dependent on the modulation frequency \( f \), and the exponent of \( \mathcal{K}_z[n, l] \) is a function of \( f \) and the time lag \( l \), indicating that the parameter \( f \) should be known in advance or estimated to compute STFD. Details of the above derivations could be found in [38] and are not presented here for brevity. This technique has also been applied to enhance electric network signal, a mono-mode signal, for forensic analysis [39].
we consider a more complicated multi-mode scenario.

2) Signal Encoding: The purpose of introducing STFD is to explore its property for enhancing initial IF estimation of the input \( s[n] \) in Fig.2. By comparing the signal \( s[n] \) in (1) and the signal \( z[n] \) adapted to STFD in (3), we attempt to transform the sinusoidal modulated waveform \( s[n] \) into an approximate SFM signal by encoding \( s[n] \) into the IF of a unit-amplitude analytic signal:

\[
z_s[n] = \exp \left\{ j2\pi T_\mu \sum_{m=0}^{n} s[m] \right\}
\]

\[
= \exp \left\{ j2\pi T_\mu \sum_{m=0}^{n} x_i[m] \right\} \exp \left\{ j2\pi T_\mu \sum_{m=0}^{n} \epsilon[m] \right\}
\]

\[
= z_s[n] z_e[n],
\]  

(10)

where \( \mu \) denotes the modulation index, \( z_s[n] = \exp \left\{ j2\pi T_\mu \sum_{m=0}^{n} \epsilon[m] \right\} \) denotes the encoded noise, which becomes multiplicative relationship with the encoded clean signal \( z_s[n] = \exp \left\{ j2\pi T_\mu \sum_{m=0}^{n} x_i[m] \right\} \).

By applying the STFD in (4) on the encoded signal \( z_s[n] \), the underlying signal \( x[n] = \sum_{i=1}^{l} x_i[n] \) is obtained by estimating the IF of the STFD on \( z_s[n] \), which could be realized by detecting the peaks of STFD \( z_s[n, v] \) at all the timeslices. However, it is cumbersome for this purpose since in practical applications: i) the implementation of STFD \( z_s[n, v] \) is expensive due to the discrete Fourier transform (DFT) on the nonlinear kernel function; ii) even though the underlying signal \( x[n] \) is recovered, it is still hard to decompose it into individual modes. Therefore, a low complexity method allowing for mode separation is designed, as discussed in the next part.

3) Kernel Phase Averaging: This method, referred to as kernel phase averaging (KPA), has two attractive features. Instead of finding the STFD followed by peak search, the IF estimation of \( z_s[n] \) can be realized by a simple phase operation on kernel function without implementing the complex STFD; More importantly, the underlying signal modes \( \{ x_i[n] \}_{i=1}^{l} \) can be individually separated by providing initial IFs of the signal \( s[n] \), which gives rise to enhanced IF of each mode. We then present the following theorems of the KPA method as well as their proofs.

**Theorem 1** For a noisy mono-mode signal:

\[
s[n] = x[n] + \epsilon[n] = \rho \cos\left(2\pi f[n]T + \theta\right) + \epsilon[n], \quad n = 0, \ldots, N-1,
\]

the kernel phase averaging on STFD’s kernel of \( z_s[n] \) theoretically equals to the underlying clean signal \( x[n] \), provided that the IF of \( s[n] \) approximates to a constant within a certain range:

\[
E\left\{ K_{z_s}[n, l] \right\} \approx \rho \cos\left(2\pi f[n]T + \theta\right) = x[n],
\]  

(11)

where \( E\{ \cdot \} \) denotes the expectation operator, and \( K_{z_s} \) denotes the kernel function of STFD on \( z_s[n] \). \( f[n] \) is the IF at the \( n \)-th time-slice, which is regarded as a constant within \([n-L/2, n+L/2]\), where \( L \) is the window length of the KPA method.

**Proof of Theorem 1.** The exponent of the kernel function in (9) is dependent on the time lag \( l \), which allows for a direct estimation of the underlying signal \( x[n] \) by performing an averaging operation on the phase of the kernel in (9) by multiplying the term \( \frac{1}{2\pi T_\mu} \):

\[
\Theta\{ K_{z_s}[n, l] | f[n] \} \approx \frac{2}{2\pi T_\mu} \left\{ \frac{1}{n} \sum_{i=1}^{n} \left\{ z_s[n - i + l, \epsilon[n]] \right\} \right\},
\]

(12)

where \( l \in [-L/2, L/2] \). By encoding \( s[n] \) into \( z_s[n] \) according to (10), and then bringing \( z_s[n] \) into (12), the expectation operation allows for an unbiased signal recovery of \( x[n] \) from the noisy signal \( s[n] \), derived as below:

\[
E\left\{ \Theta\{ K_{z_s}[n, l] | f[n] \} \right\} = E\left\{ \Theta\{ K_{z_s}[n, l] \cdot K_{z_s}[n, l] \} \right\}
\]

\[
= E\left\{ \Theta\{ K_{z_s}[n, l] \} \right\} + E\left\{ \Theta\{ K_{z_s}[n, l] \} \right\},
\]  

(13)

where

\[
K_{z_s}[n, l] = \exp \left\{ j\pi \left(2\pi f[n]T \sum_{n+l}^{n} x[\lambda] + \cos(2\pi f[n]T) \sum_{n-l}^{n} \epsilon[\lambda] \right) \right\};
\]

\[
K_{z_s}[n, l] = \exp \left\{ j\pi \left(2\pi f[n]T \sum_{n+l}^{n} \epsilon[\lambda] + \cos(2\pi f[n]T) \sum_{n-l}^{n} x[\lambda] \right) \right\}.
\]

Under the assumption that \( \epsilon(t) \) is zero-mean additive noise and \( x[n] = \rho \cos\left(2\pi f[n]T + \theta\right) \), we have \( E\{ \Theta\{ K_{z_s}[n, l] \} \} \approx 0 \), therefore

\[
E\left\{ \Theta\{ K_{z_s}[n, l] \} \right\} \approx E\left\{ \Theta\{ K_{z_s}[n, l] \} \right\}
\]

\[
= A \left\{ \sin(2\pi f[n]T) \sum_{n-l}^{n+l} x[\lambda] + \cos(2\pi f[n]T) \sum_{n-l}^{n+l} \epsilon[\lambda] \right\}
\]

\[
= \rho \cos(2\pi f[n]T + \theta) = x[n],
\]  

(14)

where \( A = \pi T_\mu \). Note that the phase term \( \Theta\{ K_{z_s}[n, l] \} \) on \( z_s[n] \) keeps constant over the time lag \( l \), which gives a precise estimation of the clean signal \( x[n] \). In the presence of noise, \( \Theta\{ K_{z_s}[n, l] \} \) on \( z_s[n] \) changes with the variation of \( l \), under this circumstance, the underlying signal \( x[n] \) can be recovered by taking the mean operation on the fluctuating phase term \( \Theta\{ K_{z_s}[n, l] \} \) over the time lag \( l \). The proof of Theorem 1 is completed.

**Theorem 2** For a noisy multi-mode signal with \( I \) modes:

\[
s[n] = \sum_{i=1}^{I} x_i[n] + \epsilon[n] = \sum_{i=1}^{I} \rho_i \cos\left(2\pi f_i[n]T + \theta_i\right) + \epsilon[n], \quad n = 0, \ldots, N-1,
\]

the kernel phase averaging on STFD’s kernel of \( z_s[n] \) theoretically equals to the underlying clean mode \( x_i[n] \), provided that i) the IF of the \( i \)-th mode inserted into the kernel function approximates to a constant within a certain range; ii) a separation condition in (27) should be satisfied. Under the above conditions, we have

\[
\hat{x}_i[n] = E\left\{ \Theta\{ K_{z_s}[n, l] \} | e[n] \right\}
\]

by
where \( f_i[n] \) is the IF of the \( i \)-th mode at the \( n \)-th time-slice, which is regarded as a constant within \( n - L/2, n + L/2 \). Consequently, the clean multi-mode signal \( x[n] \) can be recovered through:

\[
\hat{x}[n] = \sum_{i=1}^{I} \rho_i \cos \left( 2\pi f_i[n]nT + \theta_i \right) \approx \sum_{i=1}^{I} x_i[n]. \tag{16}
\]

**Remark.** According to (15) and (16), it is worth stressing that the \( i \)-th signal mode, \( x_i[n] \), can be successfully reconstructed if its IF, \( f_i[n] \), is known or estimated. Intuitively, the reconstructed \( \hat{x}_i[n] \) is inevitably influenced by other signal modes, i.e., \( \sum_{j \neq i} x_j[n] \), especially when the IF laws \( \{f_i[n]\}_{i=1, \ldots, I} \) are closely-spaced or crossing. The proof of Theorem 2 will be provided in Section IV.A.

4) Iterative IF Refinement: Each signal mode can be filtered by the operation in (15), and meanwhile the noise variance is greatly mitigated. Thus, the initial IF estimation in Section III.A is further enhanced based on STFT of estimated \( \{\hat{x}_i[n]\}_{i=1, \ldots, I} \) by detecting STFT peaks across each time index [24]:

\[
\hat{f}_i[n] = \arg \max_{v} \left\{ \text{STFT}_{\hat{x}_i[n]}[n, v] \right\} / N_{\text{ft}}, \quad i = 1, \ldots, I \tag{17}
\]

where \( N_{\text{ft}} \) is the DFT length of STFT computation, and \( \hat{x}_i[n] \) denotes the analyzed version of \( \hat{x}_i[n] \).

It has been demonstrated that one iteration would be sufficient for a desirable enhancement performance. However, the enhanced signal modes using initial IFs can offer an updated IF estimate, which means that the KPA method can be implemented in an iterative manner, i.e., (15) in Theorem 2 is repeatedly conducted by updating IF information until a satisfied performance is achieved or a prescribed number of iterations is reached.

C. Combining IF and IA for Mode Decomposition

After obtaining the IF information of each mode, the subsequent task is the IA recovery, so that the mode reconstruction can be realized by combining extracted IFs and IAs. Despite the effectiveness on IF enhancement using the KPA method, the IA of enhanced \( \hat{x}[n] \) in (16) might be severely distorted. It would become more difficult to recover the IA as the SNR level decreases. In our study, the distortion of mode’s IA can be compensated by taking into account the information in linear STFT, which is a commonly-used linear TFR with a complex-valued TF description that involves magnitude contribution, thus enabling IA reconstruction.

On the one hand, the Gaussian window based STFT is explored for reconstructing the IA of each mode:

\[
\text{STFT}_s[n, v] = \sum_{l=0}^{N-1} s'[l]g[l - n] \exp \left\{ -j2\pi lv / N \right\}, \tag{18}
\]

where \( s'[n] \) denotes the analyzed version of \( s[n] \), and \( g[n] \) is the real Gaussian window: \( g[n] = \sigma^{-1} \exp\{-\pi n^2 / \sigma^2\} \), where \( \sigma \) controls the Gaussian window size.

On the other hand, according to the enhanced IFs \( \{\hat{f}_i[n]\}_{i=1, \ldots, I} \) in (17), the synchroextracting operator (SEO) of our proposed ETFR-MD is defined as below:

\[
\text{SEO}_f[n, v] = \sum_{i=1}^{I} \delta (v - \hat{f}_i[n]TN_{\text{ft}}). \tag{19}
\]

Based on the idea of the SET [22], the proposed ETFR associated with STFT is formulated using the SEO in (19), expressed as:

\[
\text{ETFR}_s[n, v] = \text{STFT}_s[n, v] \odot \text{SEO}_f[n, v], \tag{20}
\]

where \( \odot \) denotes the element-wise product operator. Note that the Gaussian window based STFT achieves the maximum values along IF laws, thus the extracted IAs through (20) will have the best performance of noise robustness.

After obtaining the enhanced ETFRs \( \text{ETFR}_s[n, v] \), the \( i \)-th signal mode can be reconstructed by integrating \( \text{STFT}_s[n, v] \) in the IF trajectory of the corresponding mode, i.e., using only the STFT coefficients on the enhanced IF ridge of each mode:

\[
\hat{x}_i[n] = \text{Re} \left\{ \text{ETFR}_s \left[n, \left| \hat{f}_i[n]TN_{\text{ft}} \right| \right] \right\} = \text{Re} \left\{ \text{STFT}_s[n, v] \cdot \delta (v - \hat{f}_i[n]TN_{\text{ft}}) \right\}. \tag{21}
\]

where \( i = 1, \ldots, I \) and \( \text{Re}\{\cdot\} \) means the real part of a complex number. It also enables the signal enhancement in a straightforward manner, i.e., \( \hat{x}[n] = \sum_{i=1}^{I} \hat{x}_i[n] \), which can be alternatively derived by:

\[
\hat{x}[n] = \text{Re} \left\{ \sum_{i=1}^{I} \text{ETFR}_s \left[n, \left| \hat{f}_i[n]TN_{\text{ft}} \right| \right] \right\}. \tag{22}
\]

**Remark.** Signal energy of an ideal TFR will be only concentrated on true IF trajectories. Nevertheless, the utilization of STFT will inevitably give rise to biased IF estimation due to the non-linearity of some modes, which further leads to mode reconstruction error. Regarding the use of STFT for both initial IF estimation and IA extraction, one of the key issues is to determine an appropriate choice of the Gaussian window size in (15). In order to minimize the impact on IF estimation and mode decomposition, we will choose empirical values via experiments like in [20]–[22].

Returning to the problem posed at Section II, we now state that the goal to recover IF and IA information of each mode is achieved, and the proposed ETFR-MD method is summarized in Algorithm 2, where \( K \) denotes the number of iterations for the KPA method. We next provide the theoretical analysis of the proposed method in Section IV, and validate the theoretical analysis by experimental results in Section V.

IV. PERFORMANCE ANALYSIS

In this section, we first analyze that under what conditions Theorem 2 for mode enhancement holds. In this regard, a closed-form expression as a function of frequency difference
is derived, which indicates the negligible mutual interference between different modes, as long as the frequency difference is below a certain limit. In addition, the optimal window length $L$ of the KPA method in (12) used for IF enhancement is discussed, which is useful in analyzing how to adjust an appropriate window length according to the property of different modes.

A. Interference Analysis of Mode Separation

The closed-form expression of mutual interference for mode separation via the KPA Method is firstly derived, and then an illustrative example is provided to verify the theoretical expression.

Recall Theorem 2 in Section III.B, when the received signal of interest has multiple modes, the KPA method can be used for mode separation according to (13). In the following, we quantitatively analyze the cross-interference between signal modes by using KPA, and derive a condition under which the mode separation in (15) can be well carried out. In order to avoid bias, we assume a multi-mode signal $s[n]$, each mode of which has approximately constant IF within the window range of the KPA method. According to (10), we obtain the encoded multi-mode signal $z_s[n] = \exp\left\{j2\pi nT \sum_{m=0}^{n} \sum_{i=1}^{I} f_i[n] n T + \theta_i\right\}$. To prove Theorem 2, it is equivalent to proving the validity of the following expression:

$$\mathbb{E}\left\{\Theta\left\{K_z[n,l]\right\} \mid \Theta(z_s[n])\right\} \approx x_i[n], \quad i = 1, \ldots, I.$$  \hspace{1cm} (23)

Proof of Theorem 2. In a straightforward manner, the $i$-th mode in (10) is calculated as below:

$$\mathbb{E}\left\{\Theta\left\{K_z[n,l]\right\} \mid \Theta(z_s[n])\right\} = \mathbb{E}\left\{\Theta(K_z[n,l])\right\} = \mathbb{E}\left\{\Theta\left(\sum_{i=1}^{I} b_i f_i[n] \left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right)\right\}$$

$$= \sum_{i' \neq i}^I \mathbb{E}\left\{\rho_{i'} \cos\left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right\} + \mathbb{E}\left\{\rho_i \cos\left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right\}$$

$$+ \sum_{i' \neq i}^I \mathbb{E}\left\{b_{i'} f_{i'}[n] \rho_{i'} \cos\left(2\pi f_{i'}[n] T(l + a_{i'}))\right)\sin(2\pi f_{i'}[n] T(l + a_{i'}))\right\}$$

$$+ \sum_{i' \neq i}^I \mathbb{E}\left\{b_i f_i[n] \rho_i \cos\left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right\}$$

$$= \sum_{i' \neq i}^I \mathbb{E}\left\{\rho_{i'} \cos\left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right\} + \mathbb{E}\left\{\rho_i \cos\left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right\}$$

$$+ \sum_{i' \neq i}^I \mathbb{E}\left\{b_{i'} f_{i'}[n] \rho_{i'} \cos\left(2\pi f_{i'}[n] T(l + a_{i'}))\right)\sin(2\pi f_{i'}[n] T(l + a_{i'}))\right\}$$

$$+ \sum_{i' \neq i}^I \mathbb{E}\left\{b_i f_i[n] \rho_i \cos\left(2\pi f_i[n] n T + \theta_i\right)\sin(2\pi f_i[n] T(l + a_i))\right\}$$

where $\varepsilon_1$ is a small threshold value which denotes the allowed interference level. Under this condition, the expression (16) is
Absolute Amplitude Error (AAE)

\[ \text{Averaged AAE} \]

\[ 10 \] Hz, and drops sharply to zero as long as the window length is indefinitely. This naturally leads to the question of the optimal window length. The mode’s initial IF estimation \( f_{\Delta}[n] \) represents the two-mode case when \( I = 2 \), and there actually exist \( C_f^2 \) sub-optimal window lengths for more general cases when \( f > 2 \). It will become difficult to weigh an optimal value of \( L \) from multiple sub-optimal values according to (28).

Alternatively, an optimal window length can be determined by assessing the trade-off between signal bias and noise variance. It is seen from Fig. 4 that the window length significantly affects the accuracy of IF enhancement. On the one hand, the bias for IF estimation needs to be controlled by using a limited window length. On the other hand, a large window length is expected to reduce the noise variance as well as the separation interference especially in low SNR situations.

Note that our previous analysis assumes that the IF of each mode, \( f[n] \), is known and kept at a constant value within the analyzed window. However, more practical signals with varying IFs and estimated IF errors should be taken into account for further theoretical analysis. Let us consider a signal \( x[n] \) whose IF is approximately regarded as a linear frequency-modulated (LFM) segment within the range \([-L/2, L/2]\), as any nonstationary signal can be viewed as the combination of a group of LFM segments. Defining an LFM signal segment \( x[n] = \cos(2\pi T \sum_{m=0}^{n} f[m]) \) with the IF \( f[n] = f_0 + r_0 n T \), where \( f_0 \) and \( r_0 \) denote initial frequency and chirp rate of the LFM segment, respectively. Without specially referring to any particular IF estimator, let us assume the IF estimation error, \( \Delta[n] = f[n] - \hat{f}[n] \), which is modeled as a zero-mean Gaussian noise process with a variance \( \sigma_z^2 \).

The signal’s nonstationarity and inaccurate IF estimation will lead to a deterministic bias for IF enhancement, thus we derive the bias expression of the KPA method by considering time-varying IF with estimation error. According to (10), the encoded LFM signal segment becomes:

\[ z_x[n] = \exp \left\{ j \frac{\sin(2\pi f_0 n T + \pi r_0 n^2 T^2)}{f_0 + r_0 n T} \right\}. \]
\[ \hat{x}[n] = \mathbb{E} \left\{ \Theta \left\{ K_{n_l}[n,l] \right\} f[n] \right\} \]

\[ = \sum_{p=1}^{2} b_p \left\{ \sin \left( 2\pi T f_0 (n + l + a_p) + \pi T^2 r_0 (n + l + a_p)^2 \right) \right\} x_f(n) + \frac{\sin \left( 2\pi T f_0 (n - l - a_p) + \pi T^2 r_0 (n - l - a_p)^2 \right)}{f_0 + r_0 T (n - l - a_p)} \]

\[ = \mathbb{E} \left\{ \left( f_0 + r_0 n T \right) f[n] \sin \left( 2\pi T f_0 + \pi T^2 n^2 + l^2 \right) \right\} \]

\[ + \left( f_0 + r_0 n T - r_0 T - r_0 T \right) \left( f_0 + r_0 n T - r_0 T \right) \left( f_0 + r_0 n T - r_0 T \right) \]

The enhanced signal using the KPA method is derived as in (30), which is presented on the top of the current page, where \( a_1 = 0, a_2 = \frac{1}{4 f_0 n T}, b_1 = \bar{f}[n] \sin (2\pi f_0 T)/2, b_2 = \frac{f}{n T} \cos (2\pi f_0 n T)/2, \) and \( \bar{f}[n] = f_0 + r_0 n T + \Delta[n]. \)

To obtain simplified bias expressions from (30), we next consider the bias analysis in two different cases: constant \( f[n] \) with the impact of estimated IF error as well as time-varying \( f[n] \) with perfect IF estimation. The mathematical bias expressions for these two cases are introduced as follows.

1) Bias analysis for constant \( f[n] \) with \( \Delta[n] \neq 0 \): In this case, the estimated IF is \( \hat{f}[n] = f_0 + \Delta[n] \), i.e., \( r_0 = 0 \). The bias of the KPA method is thus derived from (30):

\[ \mathcal{B}_1[n] = x[n] - \mathbb{E} \left\{ \hat{x}[n] \right\} \]

\[ = x[n] - x[n] \cdot \mathbb{E} \left\{ \frac{\hat{f}[n]}{f_0} \sin \left( 2\pi f_0 n T \right) \right\} \]

\[ + \cos \left( 2\pi f_0 n T \right) \sin \left( 2\pi f_0 n T \right) \]

where \( \mathbb{E} \) denotes the mean computation over \( l \). When the error variance \( \sigma_e^2 \) is small, \( \mathcal{B}_1[n] \) approximates to:

\[ \mathcal{B}_1[n] \approx x[n] \cdot \mathbb{E} \left\{ \frac{1 - \hat{f}[n]}{f_0} \right\} \]

which shows that the value of the bias depends on \( \Delta[n] \), \( l \) and \( x[n] \). The maximum deviation occurs at a peak or a valley of the sinusoidal waveform, whereas the deviation is zero when \( x[n] = 0 \), meaning that the effect of the inaccurate IF estimation is to reduce the signal amplitude and slightly affect the signal frequency. This indicates the KPA method for the purpose of spectral restoration can tolerate certain amount of IF estimate error.

2) Bias analysis for time-varying \( f[n] \) with \( \Delta[n] = 0 \): In the similar manner as in (31), the bias expression for time-varying \( f[n] \) with \( r_0 \neq 0 \) is simplified into:

\[ \mathcal{B}_2[n] = x[n] - \mathbb{E} \left\{ \mathcal{D}_1 \sin^2 \left( 2\pi l f_0 + r_0 n T \right) \right\} \]

\[ + \mathcal{D}_2 \cos^2 \left( 2\pi l f_0 + r_0 n T \right) \]

where

\[ \mathcal{D}_1 = \frac{f^2[n]}{\left( f[n] + r_0 T \right) \left( f[n] - r_0 T \right)} \]

\[ \mathcal{D}_2 = \frac{f^2[n]}{\left( f[n] + r_0 T + r_0 \frac{f}{n T} \right) \left( f[n] - r_0 T - r_0 \frac{f}{n T} \right)} \]

When the value of \( \frac{f}{n T} \) is small, the bias expression in (33) degrades into:

\[ \mathcal{B}_2[n] \approx \cos \left( 2\pi f_0 n T + \pi r_0 n^2 T^2 \right) \]

\[ - \mathbb{E} \left\{ \mathcal{D}_1 \cos \left( 2\pi f_0 n T + \pi r_0 n^2 T^2 + \pi r_0 l^2 T^2 \right) \right\} \]

which equals to zero when \( r_0 = 0 \). The expressions in (32) and (34) show that the most significant impact for IF enhancement is the chirp rate \( r_0 \) in \( \mathcal{B}_2[n] \) rather than IF estimation error \( \Delta \) in \( \mathcal{B}_1[n] \). Fig. 5 presents the mean absolute biases (MABs) calculated by (33) and (34) as the window length \( L \) increases with varying \( r_0 \in \{ 1, 2, 5, 10, 20, 50, 100 \} \). It is seen that the approximate expression (34) well fits the exact expression (33) under the condition \( f_0 = 50 \text{ Hz} \) and \( f_s = 1024 \text{ Hz} \).

Since the KPA method is more concerned with the effect of bias on frequency information rather than amplitude information, the optimal window size for IF enhancement can be chosen by limiting the bias quantity in (34). Specifically, the deterministic bias due to the signal’s nonstationarity can be minimized by constraining the maximum of the elementary part \( \pi r\text{ol}^2 T^2 \) in (34):

\[ \max_{l} \left\{ \frac{\pi r\text{ol}^2 T^2}{f_s^2} \right\} < \varepsilon_2, \quad l \in \left[ -L, \frac{L}{2}, \frac{L}{2} \right] \]

where \( \varepsilon_2 \) is a constant coefficient used to control the quantity of MAB. As a result, the criterion of choosing the window length \( L \) is derived as a function of \( r_0, \varepsilon_2, \) and \( f_s \):

\[ L \leq 2 f_s \sqrt{\frac{\varepsilon_2}{\pi r_0}} \]

which implies that the window size should be reduced when a strict MAB (a small value of \( \varepsilon_2 \)) is required. Furthermore, as the value of \( r_0 \) decreases, large window sizes are allowed. This is reasonable because no window or an infinite window length can be used when \( r_0 = 0 \). Consequently, the window...
Mean Absolute Bias (MAB)

The computational complexity for the KPA method is

compared to that in (4) by avoiding DFT computation and peak

detecting time-slice peaks of STFD in (4). From the viewpoint

performing the phase averaging operation in (16) instead of

for IF enhancement is that the enhanced signal is obtained by

IA extraction via STFT. The main merit of the KPA method

rithm 1

by increasing the sampling frequency

shown in Fig. 5. Besides, the window length can be enlarged

the window size changes slowly with the variation of

0

Window Length (samples)

10-5
10-4
10-3
10-2
10-1
100
Mean Absolute Bias (MAB)               
(33)
(34)
(33)
(34)
50
20
10
r0=100
5
 2
1

Figure 5. Mean absolute biases (MABs) computed from (33) and (34) versus
different window length L with ρ = 1, f₀ = 50 Hz, fₛ = 1024 Hz, N = 1024,
and r₀ ∈ {1, 2, 5, 10, 20, 50, 100}.

remark. The above biases have been analyzed based on
LFM signal model in (29), but the derived theoretical result in
(36) is also suitable for other nonstationary signals with
arbitrary frequency variation. The reason lies in that every
segment of a nonstationary signal within a limited lag window
can be approximately viewed as an LFM waveform according
to the Weierstrass approximation theorem.

C. Complexity Analysis of the Proposed ETFR-MD

The overall computational complexity of the proposed
ETFR-MD should include the initial IF estimation in Algorithm 1, the IF enhancement using the KPA method, and the
IA extraction via STFT. The main merit of the KPA method
for IF enhancement is that the enhanced signal is obtained by
performing the phase averaging operation in (16) instead of
detecting time-slice peaks of STFD in (4). From the viewpoint
of implementation, the operation in (16) is much less complex
compared to that in (4) by avoiding DFT computation and peak
search process. Therefore, the mathematical description of the
computational complexity for the KPA method is O(KIN),
where K is the number of iterations, I is the number of signal
modes, and N denotes the number of signal samples. The
proposed KPA method not only has a linear complexity, but
also is compatible with parallel computation. In addition, the
IAs of signal modes are recovered through the computation of
STFT in [18], which has a complexity O(N log N).

Despite the low complexity of IF and IA recovery in the
ETFR-MD, its practical implementation requires an initial
IF estimation by carrying out state-of-the-art methods or the
proposed method in Algorithm 1, which implies an additional
increase in the computational cost. Nevertheless, it does not
significantly increase the overall complexity. Experimental
results in the next section will demonstrate that the gain of
enhanced representation and decomposition performance well
offsets the additional complexity on initial IF estimation.

V. Experimental Evaluation

In this section, we demonstrate the benefits of the proposed
ETFR-MD method in dealing with multi-mode signals with
close or crossing IFs in terms of IF enhancement, enhanced
TFR, and mode decomposition:

• IF enhancement: Experimental results of the proposed
KPA method are firstly given to illustrate its effectiveness
in enhancing initial IF estimates.

• Enhanced TFR: The enhanced time-frequency representations
by combining extracted IF and IA information are compared with the SST [14] and SET [22] methods.

• Mode decomposition: The proposed ETFR-MD for mode
decomposition is evaluated via a comparison against SST [14], SST4 [20], MSST [21], and SET [22], which are state-of-the-art methods in TF representation and mode
reconstruction. The decomposed modes are assessed via
mean square error (MSE) and Output SNR metrics [14].

In our study, the three typical multi-mode FM signals with
close or crossing IFs in Fig. 1 are selected for the computer simulation. The initial IF estimation of these multi-mode signals
is implemented using Algorithm 1. Note that an alternative
IF estimation method, e.g., ridge detection related ones, can
be used when non-overlapping FM modes are involved. Since
the number and type of signal modes are unknown, we set
empirical parameters for the entire multi-mode signal, thus
forcing a compromise in the parameter setting for signals with
different modes. Specifically, the experimental parameters in
Algorithm 1 are empirically selected: c₁ = 50, c₂ = 5000,
Δ₁ = 3, Δ₂ = 0.3, δ = 30. The number of iterations in
the KPA method is set to K = 3, moreover, we suggest the
window length L = 32, which is chosen so that the bias of
the enhanced modes in (36) is approximately equal to the
magnitude of 10⁻². It is assumed that the noise is additive
white Gaussian noise (AWGN), and we also assume that all
the signal modes exist throughout the time domain.

A. Enhanced IF Estimation

Firstly, we consider the multi-mode signal whose five IF
laws are non-overlapping but spatially close in TF domain.
As shown in the first column of Fig. 6, two quadratic TFRs,
I.e., PWVD and B-distribution (BD) [40], are also simulated
for comparison with STFT at SNR = 5 dB. The remaining
columns of Fig. 6 show the corresponding TFRs of the five
enhanced modes using the KPA method in [15]. It is worth
noting that the PWVD generating strong CTs significantly
impacts the initial IF estimation, and fails to implement the
KPA method. In contrast, the BD has both high TF resolution and
maximum CT reduction, thus obtaining accurate IF estimate. Although BD exhibits more concentrated TFR than STFT, we can still implement the initial IF estimation from the simpler STFT. It is obvious from the second and third rows of Fig. 6 that the IF ridge of each mode can be clearly recognized, based on which the IF enhancement of each mode becomes possible.

Secondly, to test initial IF estimation and IF enhancement using the KPA method, two multi-mode signals with intersected IF laws are considered. It is seen from the first column of Fig. 7 that the initial IFs based on STFT could be well detected despite the interference of crossing TF points, and then the IF of each mode is further enhanced via (17). As illustrated in the second and third columns of Fig. 7, the enhanced IFs are more smooth and accurate compared with initial ones. The comparative results show the superiority of using the KPA method on IF enhancement even for noisy signals at a low SNR level.

B. Enhanced Time-Frequency Representation

Next, we present the enhanced TFRs by a comparative study with the classical SST and currently reported SET to verify
the robustness of the proposed ETFR-MD in [20]. The first columns of Fig. 8 and Fig. 9 show the TFR behaviors of different methods on the three multi-mode signals when SNR = 10 dB. It is observed that the proposed ETFR-MD provides the clearest and most concentrated TFRs, while somewhat more expensive since initial IF estimation is required. For the methods of comparison, their parameters are adjusted to result in the most visually appealing TFRs, but these TFRs exhibit frequency and amplitude variations under low SNR conditions and complex situations with close or crossing IFs.

C. Time-Frequency Mode Decomposition

Finally, we demonstrate the performance improvement of the proposed ETFR-MD method that contributes to mode decomposition. Fig. 8 and 9 respectively show the mode decomposition results of the above three multi-mode signals.
(b) Two crossing weakly-modulated FM modes.

(c) Two crossing highly-modulated FM modes.

Output SNR (dB)

-20 -10 0 10 20 30

Input SNR (dB)

-5 0 5 10 15 20 25 30

Output SNR (dB)

-5 0 5 10 15 20 25 30

Input SNR (dB)

-5 0 5 10 15 20 25 30

Using the SST [14], the SET [22], and our ETFR-MD in [21] when SNR = 10 dB. The MSEs of estimated signal modes calculated with ideal modes are indicated on the top of each figure. Note that both the SST and SET based mode decomposition requires an initial IF estimation of each mode, which is realized by the ridge detection method from SST and SET, respectively. Experimental results in Fig. 8 and 9 reveal the benefit of the ETFR-MD using our proposed Algorithm 1 for multi-mode FM signals with close and crossing IFs.

Considering that more quantitative measures would usefully supplement and extend the previous qualitative analysis. In addition to SST and SET, we also compare our ETFR-MD with the SST4 [20] and the MSST [21], two of the state-of-the-art methods, in terms of the Output SNR metric. For SST-related methods, a larger parameter d means more accurate mode reconstruction, thus d = 10 is selected for a fair comparison. Fig. 10 presents the Output SNRs of reconstructed signal modes versus different input SNRs using four existing methods and our ETFR-MD. It is observed from Fig. 10 (a) that the ETFR-MD outperforms other methods in most cases for the non-overlapping multi-mode signal. While the MSST method achieves the best performance at high SNRs, it is conducted in an iterative process (6 iterations in our experiments). Fig. 10 (b) and (c) show that our ETFR-MD has attained much larger performance gain for spectrally-overlapping multi-mode signals, which largely benefits from robust IF estimation and enhancement of the KPA method.

VI. CONCLUSION

To address TF representation and mode reconstruction of complicated FM signals with close or crossing IFs, this paper proposes an STFT-based ETFR-MD method considering the widespread use of STFT and its invertibility. Moreover, an efficient method is designed for initial IF estimation of multi-mode signals. The essence of the proposed ETFR-MD method lies in the separate IF and IA reconstruction of each mode, which are respectively accomplished by the KPA method and STFT coefficients. Experimental results have verified our theoretical analysis and confirmed the central importance of IF enhancement. Our method does not rely on the parametric assumption of signal’s IF model, thus extending the benefits of TF representation and decomposition to more general signals of arbitrary mode combination. However, the performance of the ETFR-MD is limited by the IF estimation in STFT domain, the estimation error of which becomes more significant as the signal’s nonstationarity increases. Although our IF estimation method exceeds the cost of existing ridge detection methods, the benefits of representation and mode reconstruction may often outweigh the additional cost. To alleviate the computational cost without influencing the performance, an alternative low-complexity IF estimation method can be applied instead for non-overlapping cases. The proposed ETFR-MD method is expected to empower the development of TF representation and mode decomposition technologies in practical applications where close or crossing signal modes are involved.

ACKNOWLEDGMENT

We would like to thank the researchers for kindly sharing the source codes in their works [13], [20]–[22], which are publicly available.

REFERENCES

[1] L. Stankovic, M. Dakovic, and T. Thayaparan, “Time-Frequency Signal Analysis With Applications,” Artech House, 2013.
[2] B. Boashash, “Time-Frequency Signal Analysis and Processing (Second Edition),” Academic Press, 2016.
[3] P. Flandrin, “Explorations in Time-Frequency Analysis,” Cambridge University Press, 2018.
[4] E. Sejdic, I. Djurović, and J. Jiang, “Time–frequency feature representation using energy concentration: An overview of recent advances,” Digital Signal Processing, vol. 19, no. 1, pp. 153–183, 2009.
[5] F. Hlawatsch and G. F. Boudreaux-Bartels, “Polynomial Time–frequency Transforms,” IEEE Trans. Signal Processing, vol. 40, no. 5, pp. 134–141, 1992.
[6] B. Boashash and B. Ristic, “Polyanalytic Polynomial Time–Frequency Distributions and Time–Varying Higher Order Spectra: Application to the Analysis of Multicomponent FM Signals and to the Treatment of Multiplicative Noise,” IEEE Trans. Signal Process., vol. 47, no. 1, pp. 1–13, 1999.
[7] G. Bi and Y. Wei, “Split-Radix Algorithms for Arbitrary Order of Polynomial Time Frequency Transforms,” IEEE Trans. Signal Process., vol. 55, no. 1, pp. 134–141, 2007.
[8] S. Liu, T. Shan, R. Tao, Y. D. Zhang, G. Zhang, F. Zhang, and Y. Wang, “Sparse Discrete Fractional Fourier Transform and Its Applications,” IEEE Trans. Signal Process., vol. 62, no. 24, pp. 6582–6595, 2014.
[9] J. Shi, J. Zheng, X. Liu, W. Xiang, and Q. Zhang, “Novel Short-Time Fractional Fourier Transform: Theory, Implementation, and Applications,” IEEE Trans. Signal Process., vol. 68, pp. 3280–3295, 2020.
[10] L. Cohen, “Time-frequency distributions—a review,” Proceedings of the IEEE, vol. 77, no. 7, pp. 941–981, 1989.
[11] D. Griffin and J. Lim, “Signal estimation from modified short-time Fourier transform,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 32, no. 2, pp. 236–243, 1984.
[12] I. Daubechies, “The Wavelet Transform, Time-frequency Localization and Signal Analysis,” IEEE Trans. Inf. Theory, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
[13] F. Auger and P. Flandrin, “Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method,” IEEE Trans. Signal Process., vol. 43, no. 5, pp. 1068–1089, May 1995.
[14] F. Auger, P. Flandrin, Y. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H. Wu, “Time-frequency reassignment and synchrosqueezing: An overview,” IEEE Signal Processing Magazine, vol. 30, no. 6, pp. 32–41, 2013.
[15] T. Oberlin, S. Meignen, and V. Perrier, “The Fourier-based synchrosqueezing transform,” in 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 315–319.
[16] I. Daubechies, J. Lu, and H.-T. Wu, “Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool,” Applied and Computational Harmonic Analysis, vol. 30, no. 2, pp. 243–261, 2011.
[17] T. Thakur and H.-T. Wu, “Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples,” SIAM Journal on Mathematical Analysis, vol. 43, no. 5, pp. 2078–2095, 2011.
[18] S. Wang, X. Chen, G. Cai, B. Chen, X. Li, and Z. He, “Matching Demodulation Transform and SynchroSqueezing in Time-Frequency Analysis,” IEEE Trans. Signal Process., vol. 62, no. 1, pp. 69–84, 2014.
[19] T. Oberlin, S. Meignen, and V. Perrier, “Second-Order Synchrosqueezing Transform or Invertible Reassignment? Towards Ideal Time-Frequency Representations,” IEEE Trans. Signal Process., vol. 63, no. 5, pp. 1335–1344, 2015.
[20] D. Pham and S. Meignen, “High-order synchrosqueezing transform for multicomponent signals analysis—with an application to gravitational-wave signal,” IEEE Trans. Signal Process., vol. 65, no. 12, pp. 3168–3178, 2017.
[21] G. Yu, Z. Wang, and P. Zhao, “Multisynchrosqueezing transform,” IEEE Transactions on Industrial Electronics, vol. 66, no. 7, pp. 5441–5455, 2019.
[22] G. Yu, M. Yu, and C. Xu, “Synchroextracting Transform,” IEEE Trans. Industrial Electronics, vol. 64, no. 10, pp. 8042–8054, 2017.
[23] Abdoush, Y. and Pojani, G. and Corazza, G. E., “Adaptive Instantaneous Frequency Estimation of Multicomponent Signals Based on Linear Time–Frequency Transforms,” IEEE Trans. Signal Process., vol. 67, no. 12, pp. 3100–3112, 2019.
[24] B. Boashash, “Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications,” Proceedings of the IEEE, vol. 80, no. 4, pp. 540–568, 1992.
[25] R. A. Carmona, W. L. Hwang, and B. Torresani, “Characterization of signals by the ridges of their wavelet transforms,” IEEE Trans. Signal Process., vol. 45, no. 10, pp. 2586–2590, 1997.
[26] ———, “Multiridge detection and time-frequency reconstruction,” IEEE Trans. Signal Process., vol. 47, no. 2, pp. 480–492, 1999.
[27] I. Djurović and L. Stanković, “An algorithm for the wigner distribution based instantaneous frequency estimation in a high noise environment,” Signal Processing, vol. 84, no. 3, pp. 631–643, 2004.
[28] Z. M. Hussain and B. Boashash, “Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions,” IEEE Trans. Signal Process., vol. 50, no. 8, pp. 1866–1876, 2002.
[29] D. Iatsenko, P. McClintock, and A. Stefanovska, “Extraction of instantaneous frequencies from ridges in time–frequency representations of signals,” Signal Processing, vol. 125, pp. 290–303, 2016.
[30] Q. Zhu, M. Chen, C.-W. Wong, and M. Wu, “Adaptive multi-trace carving for robust frequency tracking in forensic applications,” arXiv:2005.06686, 2020.
[31] L. Rankine, M. Mesbah, and B. Boashash, “IF estimation for multicomponent signals using image processing techniques in the time–frequency domain,” Signal Processing, vol. 87, no. 6, pp. 1234 – 1250, 2007.
[32] H. Zhang, G. Bi, W. Yang, S. G. Razul, and C. M. S. See, “IF estimation of FM signals based on time-frequency image,” IEEE Transactions on Vehicular Technology and Electronic Systems, vol. 51, no. 1, pp. 326–343, 2015.
[33] N. A. Khan and S. Ali, “A robust and efficient instantaneous frequency estimator of multi-component signals with intersecting time-frequency signatures,” Signal Processing, vol. 177, p. 107728, 2020.
[34] X. Zhu, H. Yang, Z. Zhang, J. Gao, and N. Liu, “Frequency-chirprate reassignment,” Digital Signal Processing, vol. 104, pp. 102783, 2020.
[35] I. Djurović and L. Stanković, “Time–frequency representation based on the reassigned S-method,” Signal Processing, vol. 77, no. 1, pp. 115–120, 1999.
[36] D. Iatsenko, P. V. McClintock, and A. Stefanovska, “Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms,” Digital Signal Processing, vol. 42, pp. 1–26, 2015.
[37] H. Zhang, G. Bi, S. G. Razul, and C. M. S. See, “Robust time-varying filtering and separation of some nonstationary signals in low SNR environments,” Signal Processing, vol. 106, pp. 141–158, 2015.
[38] V. Valeau, J.-C. Valière, and C. Mellet, “Instantaneous frequency tracking of a sinusoidally frequency-modulated signal with low modulation index: application to laser measurements in acoustics,” Signal Processing, vol. 84, no. 7, pp. 1147–1165, 2004.
[39] G. Hua and H. Zhang, “ENF Signal Enhancement in Audio Recordings,” IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1868–1878, 2020.
[40] B. Barkat and B. Boashash, “A High-Resolution Quadratic Time-Frequency Distribution for Multicomponent Signals Analysis,” IEEE Trans. Signal Process., vol. 49, no. 10, pp. 2232–2239, Oct. 2001.