

# Adaptive Fourier decomposition-based Dirac type time-frequency distribution

Liming Zhang<sup>a</sup>, Tao Qian<sup>a\*†</sup>, Weixiong Mai<sup>a</sup> and Pei Dang<sup>b</sup>

Communicated by W. Sprößig

The Dirac-type time-frequency distribution (TFD), regarded as *ideal* TFD, has long been desired. It, until the present time, cannot be implemented, due to the fact that there has been no appropriate representation of signals leading to such TFD. Instead, people have been developing other types of TFD, including the Wigner and the windowed Fourier transform types. This paper promotes a practical passage leading to a Dirac-type TFD. Based on the proposed function decomposition method, viz., adaptive Fourier decomposition, we establish a rigorous and practical Dirac-type TFD theory. We do follow the route of analytic signal representation of signals founded and developed by Garbo, Ville, Cohen, Boashash, Picinbono, and others. The difference, however, is that our treatment is theoretically throughout and rigorous. To well illustrate the new theory and the related TFD, we include several examples and experiments of which some are in comparison with the most commonly used TFDs. Copyright © 2016 John Wiley & Sons, Ltd.

**Keywords:** instantaneous frequency; mono-component; multi-component; adaptive Fourier decomposition; time-frequency distribution; the best  $n$ -rational approximation

## 1. Introduction

Instantaneous frequency (IF) and time-frequency distribution (TFD) are important subjects in signal processing [1–11]. The purpose is to provide the representative energy density of a signal in both time and frequency [7]. There, however, have been different approaches and understandings towards frequency, or IF [1]. Together with IF, mono-component (MC), and multi-component are among those most concerned by signal analysts. In the literature, they are described as ‘A signal is said to be a MC, if for this signal, there is only one frequency or a narrow range of frequencies varying as a function of time; and, it is a multi-component if it is not a mono-component’ [1]. In this definition, neither the meaning of ‘frequency’ nor that of ‘narrow range’ is precisely defined. IF, band of IF, and others are not defined because they are thought of as something that objectively exists. According to what is said on MC, although one generally agrees that a single sinusoid or linear combinations of two sinusoids of particular types are MCs, there are, however, no general methods to judge whether a given signal is a MC or not. In the book of Cohen [7], he raises five famous paradoxes in relation to the frequency concept. Boashash, as a pioneer signal analyst in the late 20th century, cited ‘... the definition of IF is controversial, application-related, and empirically assessed’ [1]. This situation has been lasting: Until now, an acceptable theoretical foundation of signal processing has not been established.

In the classical analysis, *frequency* is often referred to the Fourier frequency. When signal analysts concern ‘time-varying frequency’, they still, and often tacitly, refer to Fourier frequency. Such ambiguity could cause conceptual confusion. Like Fourier, there exist different types of frequencies. In wavelet decomposition, for instance, the dilation parameter plays the role of frequency. Integral transformations given by short-time Fourier transform (STFT), Wigner–Ville distribution (WVD), and others are alternatives to series expansions having their own sense of frequency [1]. There are also frequencies defined through empirical mode decomposition [12] and  $\alpha$ -counting frequency [13]. In order to apply the notion of frequency in the analysis of nonstationary signals, it would be necessary to introduce a new and basic family of functions that are nonstationary and of an oscillatory form so that the notion of ‘frequency’ is attainable [1].

Let  $s(t)$  be a complex-valued or real-valued signal with, respectively, the amplitude-phase representation  $s(t) = \rho(t)e^{i\theta(t)}$  or  $s(t) = \rho(t)\cos\theta(t)$ , where  $\rho$  is non-negative and  $\theta$  is real-valued. If the phase derivative  $\theta'(t)$  could be well defined, then a Dirac-type TFD

<sup>a</sup> Faculty of Science and Technology, University of Macau, Macao, China

<sup>b</sup> Department of General Education, Macau University of Science and Technology, Macao, China

\* Correspondence to: Tao Qian, Faculty of Science and Technology, University of Macau, Macao, China.

† E-mail: fsttq@umac.mo

could be given by the following:

$$P(t, \omega) = \rho(t)\delta(\omega - \theta'(t)), \quad (1)$$

where the Dirac  $\delta$  function is defined as  $\delta(0) = 1$  and  $\delta(t) = 0, t \neq 0$ . Existing knowledge on the Dirac-type TFD is referred to, for instance, Cohen [6]. A more general form of the Dirac-type TFD is cited in [2]:

$$P(t, \omega) = A(t, \omega)\delta(\omega - \theta'(t)), \quad (2)$$

where  $A(t, \omega)$  is the time-frequency representation of the amplitude function  $\rho(t)$ . It is quoted in [2] that for MC signals of the form  $\rho(t)e^{i\theta(t)}$ , it would be intuitively satisfying to generate a TFD of the form (2). The same paper further quotes 'Few distributions considered at present provide a TFD of the form given in (2).' Taking the WVD as an example, if a signal is with a large bandwidth time [2] and satisfies  $\theta^{(n)}(t) = 0$  for  $n \geq 3$ , then its WVD is approximately of the form (2). Signals with quadratic phase functions, therefore, satisfy the required conditions [2]. Those restrictive conditions on signals are, of course, neither natural nor practical. Although the Dirac-type TFD is desired, being frequently mentioned and regarded as *ideal*, neither its theory nor its applications have been developed. A general TFD theory and practice should deal with not only signals of linear or quadratic phase, and of smooth phase, such as combination of trigonometric and chirp signals, but also, in general, those of only finite energy, and in particular non-smooth ones.

There exist well-developed formulations in time-energy distribution and frequency-energy distribution, separately. The separated time-energy density and frequency-energy density are not sufficient to present at a time moment which frequencies and how much of each of them are contained in the signal. Time-frequency representations and 'IF estimations' have been studied by many authors, including [1, 2, 5–7, 10, 14–23]. STFT and the WVD are among currently used major techniques for time-frequency analysis for deterministic signals. Our study on adaptive Fourier decomposition (AFD)-based TFD in the present paper is mainly for bringing in a coherent and theoretical foundation of signal analysis. It is not for a thorough comparison between all the existing methods. Nevertheless, we conduct several experiments for illustrations.

A rigorous definition of MCs is recently proposed and developed by Qian *et al.* [24–40]. The new MC definition complies with the idea of the existing MC literature [1], but it is not the same. With rigorous mathematical formulation, the former has the capacity for classifying and representing all types of signals with finite energy into MCs, developing a family of adaptive MC decompositions and relevant topics, as well as being applied in signals analysis and processing. It is a fact that for any multi-component, there exist MC series decompositions that can converge as fast as one desires [31]. To gain certain stability of MC decomposition, we adopt AFD, or, in other terminology, adaptive Takenaka–Malmquist systems [29, 33, 35–37]. Among different kinds of MC decompositions, in this paper, we use a particular one called *n-Best AFD*, being equivalent to *n-Best* rational approximation [36]. Based on *n-Best AFD*, the Dirac-type TFD is realized in this paper.

The theory of AFD-based TFD addresses the main criteria of TFDs, including IF, TFD, marginal distribution, spectrum density, mean of IF, conditional mean of IF, standard deviation of frequency, standard deviation of time, and covariance. The orthogonality of real AFD decomposition and a new uncertainty principle with respect to TFDs of MCs are proved in the Appendix. Through *n-Best AFD*, time-varying IFs are obtained out of the composing MCs. The present paper is a further development of the AFD-based TFD study initially in [41].

The writing plan of this paper is as follow. The theoretical foundation of MC with detailed description and illustrative examples is given in Section 2. Besides serving as a complete summary of the existing theory, it further develops the MC theory through introducing two finer classes of MCs, namely, normal MCs (NMCs) and abnormal MCs (ANMC). The section gives new and non-trivial examples of nonlinear phases of the two kinds. AFD and *n-Best AFD* are briefly introduced in Section 3. Principles of AFD-based Dirac-type TFD are introduced in Section 4. Experiments with comparisons are given in Section 5. The conclusions are drawn in Section 4.

The MC function theory and AFD are available in two parallel contexts, viz., the unit circle and the real line. In this paper, we will concentrate on the unit circle case corresponding to signals living on compact intervals. The two contexts have their individual orthogonal rational systems, but the ideas on adaptive decompositions and Dirac-type TFDs in the two contexts are exactly the same.

## 2. Mono-component

For a complex-valued signal in the quadrature form, the IF is conventionally defined to be the phase derivative [7]. Some significant relations can be deduced based on such defined frequency, which include the following:

$$\langle \omega \rangle_{\hat{s}} = \langle \theta' \rangle_s, \quad (3)$$

where the  $\hat{s}$ -mean uses the measure  $|\hat{s}(\omega)|^2 d\omega$ ,  $\omega$  is the Fourier frequency, and the  $s$ -mean uses the measure  $|s(t)|^2 dt$ , where  $t$  is the time variable (also see the formula after (35)). There are also similar relations for group delay and the induced ones for the derivations [7, 26]. Such definition of IF has certain obstacles and limitations: (1) It is available only for complex piece-wise smooth signals, including quadrature signals with linear and polynomial phases, and (2) various quantities in signal analysis, including means of frequency, and others, which cannot be made well qualified or with the expected significance if negative phase derivatives are present. For real-valued signals such definition of frequency have the extra problem of non-uniqueness of amplitude-phase representations  $s(t) = \rho(t) \cos \theta(t)$ . What is needed is a uniform treatment of all smooth and non-smooth, real and complex signals together. The analytic signal approach proposed by Gabor [42] can well treat the non-uniqueness problem. Let  $s$  be a signal of finite energy; it can be real and complex valued. Then  $s + iHs$  is the analytic signal associated with  $s$ , where  $H$  is the Hilbert transformation in the context. The property of the Hilbert transformation makes the zero spectrum on all negative Fourier frequencies (no matter whether  $s$  is real valued or complex valued)

of  $s + iHs$ . The relation (3) implies the weak non-negative property  $\langle \theta' \rangle_s \geq 0$  for analytic signals. The analytic signal approach has been paid closed attention by renowned signal analysts including Cohen, Boashash, Picibono, and Huang [1–3, 5–7, 12, 43]. In signal analysis, a TFD is desired to be non-negative because it represents the energy. Another desired property is that the conditional mean of the distributional frequency is identical with phase derivative (38). This property is enjoyed by the Dirac-type distribution given by (1). The conditional mean property together with the positivity of TFD implies that  $\theta'(t) \geq 0$ , almost everywhere (a.e.). The analytic signal approach provides not only unique amplitude-phase representation, weak positivity, and further pointwise non-negativity of the phase derivative but also an approach to treat non-smooth signals. Later, we give a detailed account on analytic signals and MCs.

Unless otherwise specified, later, the notation  $s$  represents a real-valued or a complex-valued signal with finite energy in the underlying time interval.

1. If  $s$  is a signal defined on the unit circle or the real line, then

$$As = \frac{1}{2}(s + iHs) \tag{4}$$

is called the analytic signal associated with  $s$ , where  $H$  is Hilbert transformation in the context that in the unit circle case is *circular Hilbert transformation*. We also write  $As$  as  $s^+$ . In the unit circle case,  $s^+$  is conventionally replaced by  $s^+ + \frac{c_0}{2}$ , where  $c_0$  is the 0-th Fourier coefficient of  $s$ . The Fourier expansion of  $s$  and its associated analytic signal are given respectively by the following:

$$s(e^{it}) = \sum_{k=-\infty}^{\infty} c_k e^{ikt}, \quad s^+(e^{it}) = \sum_{k=0}^{\infty} c_k e^{ikt}. \tag{5}$$

A signal  $s$  itself is said to be an *analytic signal* if and only if  $As = s$ . In such case,  $s$  itself has to be complex valued.

2. It is easy to see that  $s$  is an analytic signal if and only if

$$Hs = -i(s - c_0). \tag{6}$$

Furthermore,  $s$  is analytic with finite energy if and only if  $s$  is the (non-tangential) boundary limit (BL) of an analytic function in the unit disk, still denoted by  $s(z)$ , in the Hardy  $H^2$  space defined by the following:

$$H^2 = \left\{ s : \mathbf{D} \rightarrow \mathbf{C} \mid s(z) = \sum_{k=0}^{\infty} c_k z^k, \sum_{k=0}^{\infty} |c_k|^2 < \infty \right\}, \tag{7}$$

where  $\mathbf{D}$  stands for the open unit disk. For the relevant knowledge of the Hardy space and the Hilbert transformation, we refer to [44]. The correspondence between analytic signals of finite energy on the boundary and analytic functions in the Hardy space is an isometry. That gives reasons to use the same notation for the analytic functions and their BLs. We usually denote the BL function by  $s(e^{it}) = \rho(t)e^{i\theta(t)}$ ,  $\rho(t) \geq 0$ , and  $\theta(t)$  being real valued. In this notation,  $\rho(t)$  is a.e. determined, but  $\theta$  not due to the fact that trigonometric functions are  $2\pi$ -periodic.

3. If  $s \in H^2$ , then with the self-explanatory notation  $s(re^{it}) = \rho_r(t)e^{i\theta_r(t)}$ ,  $\rho_r \geq 0$ ,  $\theta_r$  real-valued,  $0 \leq r < 1$ ,  $\rho_r$  is determined, and due to smoothness of analytic functions,  $\theta_r$  is determined up to an integer multiple of  $2\pi$ . By taking derivative with respect to  $t$  to both sides of the aforementioned relation, for  $z = re^{it}$ , we have the following:

$$\frac{d}{dt} \theta_r(t) = \operatorname{Re} \left\{ \frac{zs'(z)}{s(z)} \right\}, \quad \frac{d}{dt} \rho_r(t) = -\rho_r(t) \operatorname{Im} \left\{ \frac{zs'(z)}{s(z)} \right\}. \tag{8}$$

It is a basic property of the functions in the Hardy space that when  $r \rightarrow 1-$ , both  $s(re^{it})$  and  $\rho_r(t)$  have pointwise non-zero limits, a.e., that coincide with  $s(e^{it})$  and  $\rho(t)$ .

4. Let  $s$  be a signal with finite energy, and

$$As(z) = \rho_r(t)e^{i\theta_r(t)}. \tag{9}$$

Then  $\rho_r$  is called analytic amplitude of  $s$  at the level  $r$ , and  $\theta_r$  is analytic phase of  $s$  at the level  $r$ . If  $s$  is real valued, then  $s_r(t) = \rho_r(t) \cos \theta_r(t)$ .

If for a.e.  $t$ , the derivative of the analytic phase function  $\theta_r(t)$  has a non-negative limit as  $r \rightarrow 1-$ ; that is,

$$\lim_{r \rightarrow 1-} \frac{d}{dt} \theta_r(t) \geq 0, \quad \text{a.e.}, \tag{10}$$

then  $s$  is said to be an *MC*. The a.e. non-negative (measurable) limit function is denoted  $\theta'(t)$ , called the *IF (function) of  $s$* . Note that  $\theta'(t)$  should be considered only as a notation but not as a derivative of  $\theta(t)$ ; the latter may not be determined and may not be differentiable.

For a general signal of finite energy, a differentiable phase function  $\theta$  may not exist. However, if the analytic function  $As$  has an analytic continuation across  $e^{it_0}$ , then in a neighborhood of  $e^{it_0}$ , the  $r$ -level analytic phase function  $\theta_r$  and its derivative  $\theta'_r(t)$  both have smooth continuations across the boundary that corresponds to  $r = 1$ . Denote  $\theta_1 = \theta$ , we have the following:

$$\lim_{r \rightarrow 1-} \theta'_r(t) = \theta'(t). \tag{11}$$

In such case, the notation  $\theta'(t)$  coincides with the derivative of the phase function, justifying the notation. In general cases, if the pointwise limit in (10) a.e. exists (but not necessarily a.e. non-negative), then the limit function is called *analytic phase derivative of  $s$*  (but not necessarily IF, which must be non-negative).

If  $s$  is an MC, and there exists  $r_0 < 1$  such that  $\frac{d}{dt}\theta_r(t) \geq 0$  for all  $r \in [r_0, 1)$  and all  $t \in [0, 2\pi]$ , then we say that  $s$  is an NMC. If  $s$  is an MC but not an NMC, then we say that  $s$  is an ANMC. Note that the class of MCs is closed under the multiplication operation but not addition. If an analytic signal  $s$  is not an MC, then we say that it is a *multi-component*. An analytic signal  $s$  is said to be a *pre-MC* if  $e^{it}s(t)$  is an MC.

We note that for a signal of finite energy, its analytic signal  $A_s$ , in general, may not be differentiable. The Hilbert transformation usually reduces the smoothness. In general cases, analytic phase derivative  $\theta'$  should not be assumed to exist in the classical sense. This is why the limit approach from the  $r$ -level analytic function is adopted. In many situations, such as in the inner functions case in the Nevanlinna factorization formulation [44] and the Sobolev spaces case, it can be shown that the non-tangential limits (10) exist [26, 28]. Such rigorous formulation of the MC concept is necessary for theoretical development of functional and signal analysis. Based on this approach, not only a signal analysis theory [24, 28, 29] but also a general frequency-mean and time-mean, deviation, and covariance (versus [7], for instance) theory [26], as well as the theory of all-pass filters and signals of minimum phase, are established [25].

We note that not all analytic signals are MCs. For instance, outer functions in the Nevanlinna factorization formulation are analytic signals but not MCs [24]. MCs form a large pool, including BLs of Möbius transforms [27, 30], BLs of finite and infinite Blaschke products, BLs of singular inner functions [28], BLs of starlike and  $p$ -starlike functions [29], and BLs of modified (weighted) Blaschke products [34, 45, 46]. AFD described in the following section makes use of modified Blaschke products (as given in Example 1), which are particular types of  $p$ -starlike functions.

**Example 1**

A modified Blaschke product of order  $n$  is the product between a Blaschke product of  $n - 1$  zeros and a normalized Szegö (reproducing) kernel of the Hardy space, which reads as follows:

$$s_1(z) = B_n(z) = B_{b_1, \dots, b_n}(z) = e_{b_n}(z) \prod_{l=1}^{n-1} \frac{z - b_l}{1 - \bar{b}_l z}, \tag{12}$$

where  $b_1, \dots, b_n$  are arbitrary complex numbers in the unit disk and  $e_{b_n}$  is the normalized Szegö kernel given by (20). A modified Blaschke product is theoretically a pre-MC but often an NMC. In particular, if one of the  $b_k$ 's is 0, then  $B_n$  is an NMC [33]. For instance, by taking  $(b_1, \dots, b_5) = (0, 0.1455 + 0.6335i, -0.6027 + 0.2434i, -0.2646 - 0.5937i, 0.5218 - 0.3875i)$ , the corresponding  $B_5(z)$  is a four-starlike function and therefore an NMC.

**Example 2**

Let  $s_2(t) = 2 \cos t \cos 2t$ . The associated analytic signal is  $A(s) = 2 \cos t e^{i2t} = e^{i3t} + e^{it} = 2|\cos t|e^{i\theta(t)}$ , where  $\theta(t) = 2t + \phi(t)$  and  $\phi(t) = \sum_{k=-\infty}^{\infty} (-1)^k \chi_{(-\pi/2 + k\pi, \pi/2 + k\pi)}(t)$  is a step function. Because  $\theta_r(t)$  has analytic continuation across all points  $e^{it}$  except the (isolated) discontinuous points of  $\phi(t)$ , the limit  $\theta'_r(t)$  a.e. exists and equals to 2. Therefore,  $s_2(t)$  is an MC. We note, however, that it is not an NMC, but only an ANMC. The plot of the example shows that for any small  $r < 1$ , there always exists some open interval of  $t$  around  $\pi$ , on which the phase derivatives  $\theta'_r(t)$  are negative.

**Example 3**

The signal  $s_3(e^{it}) = (3 + 2 \cos t) \cos 3t$  is the real part of the analytic signal  $A_{s_3}(e^{it}) = e^{i4t} + 3e^{i3t} + e^{i2t}$ , being the BL of the Hardy space function  $A_{s_3}(z) = z^4 + 3z^3 + z^2$ . The analytic amplitude-phase representation of  $s_3$  is with  $\rho(t) = 3 + 2 \cos t > 0$  and  $\theta(t) = 3t$ . The Hardy space function has analytic continuation across the entire unit circle with the phase derivative  $\theta'(t) = 3 > 0$ . Therefore,  $s_3$  is an NMC. This example already shows that an MC can be written as a sum of several other MCs. A signal being an MC or not does not depend on how many pieces of other MCs it can be decomposed into but depends on whether it can be reduced to one single piece of analytic signal  $\rho(t)e^{i\theta(t)}$  with  $\rho \geq 0, \theta' \geq 0$ , a.e.

**Example 4**

A more sophisticated example of the same type but with nonlinear phases is given by the BL function of the following:

$$s_4(z) = (3 + h(z))F_4(z), \tag{13}$$

where

$$F_2(z) = \frac{z}{1 - \bar{a}_2 z}, \quad F_4(z) = \frac{z^2}{1 - \bar{a}_4 z} \frac{z - a_2}{1 - \bar{a}_2 z},$$

$a_2$  and  $a_4$  belong to the unit disk, and, with  $z^* = 1/\bar{z}$ ,

$$h(z) = F_2(z) + \overline{F_2(z^*)} = \frac{z}{1 - \bar{a}_2 z} + \frac{1}{z - a_2}. \tag{14}$$

Note that  $F_2$  and  $F_4$  are respectively some orders 2 and 4 modified Blaschke products.

Substituting the expressions of  $F_2$  and  $F_4$  into the definition formula of  $s_4$ , we see that  $s_4$  is in the Hardy space, and therefore, its BL is an analytic signal. Because the poles of the constructing rational functions are all outside the closed unit disk,  $s_4(z)$  has its analytic continuation across the whole unit circle.  $h(z)$  is real valued when  $z = e^{it}$ . Therefore,

$$s_4(e^{it}) = (3 + 2\rho_2(t) \cos \theta_2(t))\rho_4(t)e^{i\theta_4(t)}, \tag{15}$$

where  $F_k(e^{it}) = \rho_k(t)e^{i\theta_k(t)}$ ,  $k = 2, 4$ , with  $|\rho(t)| = |s_4(e^{it})| = (3 + 2\rho_2(t) \cos \theta_2(t))\rho_4(t) > 0$  and the phase derivatives of  $s_4$  coincide with those of the modified Blaschke product  $F_4$ , viz.,  $\theta'_4(t)$ , and thus be positive. Further analysis shows that  $s_4$  is an NMC. Now, because the zero  $a_2$  of  $F_4(z)$  kills the sole pole  $a_2$  of  $h(z)$ ,  $h(z)F_4(z)$  is analytic in the unit disk and even across the boundary. Similarly to the analysis in Example 2, we can show that, restricted to the boundary,

$$h(e^{it})F_4(e^{it}) = 2\rho_2(t) \cos \theta_2(t)e^{i\theta_4(t)} \tag{16}$$

is an ANMC. Thus,  $s_4(e^{it})$  can be expressed as a sum of two MCs, viz.,  $3F_4(e^{it})$  and  $h(e^{it})F_4(e^{it})$ . On the other hand, (13) and (2) imply the following:

$$s_4(z) = 3F_4(z) + F_2(z)F_4(z) + F_2(z)\frac{z}{1 - \bar{a}_4z}. \tag{17}$$

The last factor in the aforementioned formula is a starlike function and hence an NMC. Therefore,  $s_4$  is a sum of three NMCs. This example shows that an MC can be decomposed into sums of MCs in different ways. All MCs involved in the decompositions of  $s_4$ , as well as  $s_4$  itself, are of non-constant amplitude and nonlinear phase. It is common that an MC can be decomposed into finite or infinite sums of some other types of MCs.

### 3. Core adaptive Fourier decomposition and $n$ -Best adaptive Fourier decomposition

Let  $s$  be a real-valued signal. Adopting the notation used in (1) to (4), it is easy to deduce the following:

$$s = 2\text{Res}^+ - c_0. \tag{18}$$

Thus, MC decomposition of  $f$  is reduced to that of  $s^+$ . The AFD, or alternatively core AFD, decomposition of  $f^+$  is given by the following:

$$s^+(e^{it}) = \sum_{k=1}^{\infty} \langle s_k, e_{a_k} \rangle B_k(e^{it}), \tag{19}$$

where  $s_1 = s^+$ ,  $B_k$  are the consecutive modified Blaschke products given by (12) and  $e_{a_k}$ 's denote the  $L^2$ -normalized parameterized reproducing or Szegő kernels given by the following:

$$e_{a_k}(e^{it}) = \frac{\sqrt{1 - |a_k|^2}}{1 - \bar{a}_k e^{it}}. \tag{20}$$

$$s_k(e^{it}) = \frac{s_{k-1}(e^{it}) - \langle s_{k-1}, e_{a_{k-1}} \rangle e_{a_{k-1}}(e^{it})}{\frac{e^{it} - a_{k-1}}{1 - \bar{a}_{k-1} e^{it}}}, k = 1, 2, \dots \tag{21}$$

For each  $k$ , the operator mapping  $s_{k-1}$  to  $s_k$  given by (21) is called the generalized backward shift operator induced by  $a_{k-1}$ . The points

$$a_k = \arg \max \left\{ \sqrt{1 - |b|^2} |s_k(b)| : b \in \text{the unit disk} \right\}. \tag{22}$$

The existence of such  $a_k$  inside the unit disk rests on the maximal selection principle [47] proved for general Hardy space functions [29, 33]. The adaptive TM system  $\{B_n\}$  in (12) is orthonormal. It is not necessary to be a basis but expands the given signal  $s^+$ , and therefore,  $s$ .

In the algorithm, the recursive computation ceases at the smallest  $N$  that makes (23) hold:

$$\|s^+\|^2 - \sum_{k=1}^N (1 - |a_k|^2) |s_k(a_k)|^2 \leq \varepsilon. \tag{23}$$

Based on the relation (18), the corresponding expansion for the originally real-valued signal  $s$  is as follows:

$$\begin{aligned} s &= 2\text{Re} \left\{ \sum_{k=1}^{\infty} \langle s_k, e_{a_k} \rangle B_k \right\} - c_0 \\ &\approx \sum_{k=1}^N \rho_k(t) \cos \theta_k(t) - c_0, \end{aligned} \tag{24}$$

where

$$\rho_k(t)e^{i\theta_k(t)} = 2\langle s_k, e_{a_k} \rangle B_k(e^{it}).$$

Practically, we take the initial value  $a_1 = 0$ , for, in such case, each entry,  $B_k$  in the decomposition series is an NMC, as well as for all  $k$  and all  $t$  [33],

$$0 < \theta'_k(t) < \theta'_{k+1}(t). \tag{25}$$

**Theorem 1**

(Orthogonality of real expansion) Let  $a_1 = 0$ . Then for arbitrary  $a_2, \dots, a_N$ , the real TM expansion defined by  $a_1, a_2, \dots, a_N$  through (24) is an orthogonal expansion.

For a proof, see Appendices.

In core AFD, the selection of  $a_1, \dots, a_n$  is of the one by one manner. The simultaneous selection of all the  $a_k$  for  $k = 1, \dots, n$  to make the left-hand-side of (23) (objective function) be minimized is called  $n$ -Best AFD, and the corresponding function

$$\sum_{k=1}^n \langle s_k, e_{a_k} \rangle B_k(e^{it}) \tag{26}$$

is called an  $n$ -Best Blaschke form. To find an  $n$ -Best Blaschke form is equivalent to find an  $n$ -Best rational approximation, as described later.

Let  $f \in H^2$ . Find a rational function  $\tilde{p}/\tilde{q}$  in  $H^2$ ,  $\tilde{p}$  and  $\tilde{q}$  are co-prime polynomials,  $\deg(\tilde{p}), \deg(\tilde{q}) \leq n$ , such that

$$\|f - \frac{\tilde{p}}{\tilde{q}}\| = \min \|f - \frac{p}{q}\|, \tag{27}$$

where  $p, q$  are polynomials satisfying the same conditions as  $\tilde{p}, \tilde{q}$ . The existence of a solution of this optimization problem was proved more than a half century ago [35, 36, 48, 49], and thereafter, but a practical algorithm to get a solution is an open problem until now. In [35], Qian proposes a method, called cyclic AFD algorithm that offers a *local solution* (called local algorithm). It, however, is practical: When the optimization problem has only one critical point, then the cyclic AFD algorithm converges to the critical point. Cyclic AFD algorithm selects better and better  $n$ -tuples of parameters to form  $n$ -Blaschke forms converging to a solution locally minimizing the objective function given in (23). In general situation, performing cyclic AFD algorithm on a collection of initial  $n$ -tuples with sufficient density, one can obtain all critical points and thus obtain the global  $n$ -Best Blaschke products and thus solutions of the  $n$ -Best rational approximation. Besides cyclic AFD, there exist other local algorithms, including RARL2, being available from INRIA, France [50].

The core AFD algorithm is given in [40]. The cyclic AFD algorithm is given in [35]. Algorithm codes of core AFD and cyclic AFD algorithm are available from the webpage address <http://www.fst.umac.mo/en/staff/fsttq.html>. This paper uses, among the two AFD algorithms there, the one with the main function *find\_poles*.

## 4. Principles of adaptive Fourier decomposition-based time-frequency distributions

From now on, we assume that  $s$  is square integrable with energy 1; that is,  $\|s\|^2 = 1$ . The Dirac-type TFD aspects are discussed for MCs and multi-components in Subsections 4.1 and 4.2, respectively.

### 4.1. Mono-component time-frequency distribution

4.1.1. *Time-frequency distribution.* If  $s$  is an MC, no matter real-valued or complex-valued, the MC TFD of  $s(t) = \rho(t) \cos \theta(t)$  or  $s(t) = \rho(t)e^{i\theta(t)}$ ,  $\rho(t) = |s(t)|$  (TFD of  $s$ ) is defined to be the following:

$$P(t, \omega) = \rho^2(t)\delta(\omega - \theta'(t)). \tag{28}$$

4.1.2. *Marginal distribution.* The  $\omega$ -marginal distribution is as follows:

$$\begin{aligned} P_\omega(t) &= \int_0^\infty P(t, \omega) d\omega \\ &= \int_0^\infty \rho^2(t)\delta(\omega - \theta'(t))d\omega = \rho^2(t). \end{aligned} \tag{29}$$

The  $t$ -marginal distribution is as follows:

$$\begin{aligned} P_t(\omega) &= \int_0^{2\pi} P(t, \omega) dt \\ &= \sum_{t_i \in \theta'^{-1}(\omega)} \rho^2(t_i) / |\theta''(t_i)|, \end{aligned} \tag{30}$$

where

$$\theta'^{-1}(\omega) = \{t_j \mid \theta'(t_j) = \omega\}$$

and the last equal relation (30) is referred to Chapter IV, 4.7, [7].

4.1.3. *Total energy.* Fubini's theorem implies the following:

$$\int \int P(t, \omega) dt d\omega = \int_0^{2\pi} P_\omega(t) dt = \int_0^\infty P_t(\omega) d\omega = 1, \quad (31)$$

where

$$\begin{aligned} \int_0^{2\pi} P_\omega(t) dt &= \int_0^{2\pi} \rho^2(t) dt = \|s\|^2 = 1 = \\ &= \int_0^\infty P_t(\omega) d\omega = \int_0^\infty \sum_{t_i \in \theta'^{-1}(\omega)} \rho^2(t_i) / |\theta''(t_i)| d\omega. \end{aligned}$$

4.1.4. *Spectrum density (frequency density).* Define the IF spectrum density by the following:

$$\hat{S}(\omega) := P_t^{1/2}(\omega) = \sqrt{\sum_{t_i \in \theta'^{-1}(\omega)} \rho^2(t_i) / |\theta''(t_i)|}. \quad (32)$$

The Fourier series spectrum density is known to be the following:

$$\hat{s}(n) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} s(t) e^{-int} dt. \quad (33)$$

By (31), we have the Plancherel theorems:

$$\|\hat{s}\|^2 = \|s\|^2 = \|\hat{S}\|^2, \quad (34)$$

where the first equal relation is classical and the second is a new Plancherel theorem for IF.

4.1.5. *Mean of instantaneous frequency and mean of Fourier frequency.* We have the following remarkable relations between the mean of Fourier frequency, the mean of the IF, and the mean of the TFD:

$$\langle \omega \rangle_{\hat{S}} = \langle \theta' \rangle_s = \langle \omega \rangle_P, \quad (35)$$

where for point measure  $dn(\omega)$  at the non-negative integers

$$\langle \omega \rangle_{\hat{S}} = \int_0^\infty \omega |\hat{S}(\omega)|^2 dn(\omega), \quad \langle \theta' \rangle_s = \int_0^{2\pi} \theta'(t) |s(t)|^2 dt,$$

and

$$\begin{aligned} \langle \omega \rangle_P &= \int_0^\infty \int_0^{2\pi} \omega \rho^2(t) \delta(\omega - \theta'(t)) d\omega dt \\ &= \int_0^\infty \omega |\hat{S}(\omega)|^2 d\omega. \end{aligned} \quad (36)$$

The last relation shows that  $\langle \omega \rangle_P$  can also be denoted  $\langle \omega \rangle_{\hat{S}}$ . The first equal relation in (35) is referred to (1.91), §1.5, Ch. 1, [7], and also to [26], while it takes into account that Fourier transforms of MCs are supported in  $[0, \infty)$ .

The second equal relation is new, proved through the following:

$$\begin{aligned} \int_0^\infty \omega |\hat{S}(\omega)|^2 d\omega &= \int_0^\infty \omega \sum_{t_i \in \theta'^{-1}(\omega)} \rho^2(t_i) / |\theta''(t_i)| d\omega \\ &= \int_0^\infty \omega \int_0^{2\pi} \rho^2(t) \delta(\omega - \theta'(t)) dt d\omega \\ &= \int_0^{2\pi} \rho^2(t) dt \int_0^\infty \omega \delta(\omega - \theta'(t)) d\omega \\ &= \int_0^{2\pi} \theta'(t) \rho^2(t) dt. \end{aligned} \quad (37)$$

4.1.6. *Conditional means of frequency in time-frequency distribution.* The important feature of IF is that the conditional mean of TFD at the time moment  $t$  is IF itself:

$$\langle \omega \rangle_{P,t} = \frac{1}{P_\omega(t)} \int_0^\infty \omega P(t, \omega) d\omega = \theta'(t), \quad (38)$$

being consistent with the idea of MTFD for MCs. The relation is proved by invoking the property of the Dirac distribution and the value of the marginal distribution  $P_\omega(t)$ . It can be easily shown for any positive integer  $k$ ; there hold

$$\langle \omega^k \rangle_{P,t} = \theta'^k(t) = \langle \omega \rangle_{P,t}^k. \quad (39)$$

By defining

$$\langle \omega^k \rangle_P = \int_0^{2\pi} \langle \omega^k \rangle_{P,t} P_\omega(t) dt, \quad (40)$$

we have

$$\langle \omega^k \rangle_P = \int_0^{2\pi} \theta'^k(t) |s(t)|^2 dt = \langle \theta'^k \rangle_s. \quad (41)$$

This is a generalization of the second equality in (35). The first equality of (35) cannot be such generalized. For Fourier frequency, the relation

$$\langle \omega^k \rangle_s = \langle \theta'^k \rangle_s \quad (42)$$

holds only for  $k = 1$  (Chapter I, 1.5., (1.90), (1.98), [7]).

4.1.7. *Standard deviations.* The conditional standard deviation of TFD at the time moment  $t$  is as follows:

$$\begin{aligned} \sigma_{P,\omega/t}^2 &= \frac{1}{P_\omega(t)} \int_0^\infty (\omega - \langle \omega \rangle_{P,t})^2 P(t, \omega) d\omega \\ &= \langle \omega^2 \rangle_{P,t} - \langle \omega \rangle_{P,t}^2 \\ &= \theta'^2(t) - \theta'^2(t) = 0. \end{aligned} \quad (43)$$

The last relation implies a strong uncertainty principle for the conditional standard deviation for MCs (see later). Thanks to this relation and the general formula (4.71) in §4, Ch.4, [7], that is,

$$\sigma_{P,\omega}^2 = \int_0^{2\pi} \sigma_{P,\omega/t}^2 P_\omega(t) dt + \int_0^{2\pi} (\langle \omega \rangle_{P,t} - \langle \omega \rangle_P)^2 P_\omega(t) dt, \quad (44)$$

we have

$$\begin{aligned} \sigma_{P,\omega}^2 &= \int_0^{2\pi} (\theta'(t) - \langle \theta' \rangle_s)^2 |s(t)|^2 dt \\ &= \langle \theta' \rangle_s^2 - \langle \theta'^2 \rangle_s = \sigma_{IF}^2. \end{aligned} \quad (45)$$

4.1.8. *Conditional mean of time in time-frequency distribution.* Define

$$\langle t^l \rangle_{P,\omega} = \frac{1}{P_t(\omega)} \int_0^{2\pi} t^l P(t, \omega) dt, \quad l = 1, 2, \dots \quad (46)$$

It can be directly verified

$$\begin{aligned} \langle t^l \rangle_{P,\omega} &= \frac{1}{P_t(\omega)} \int_0^{2\pi} t^l P(t, \omega) dt \\ &= \frac{1}{\int_0^{2\pi} \rho^2(t) \delta(\omega - \theta'(t)) dt} \int_0^{2\pi} t^l \rho^2(t) \delta(\omega - \theta'(t)) dt \\ &= \frac{\sum_{t_i \in \theta'^{-1}(\omega)} t_i^l \rho^2(t_i) / |\theta''(t_i)|}{\sum_{t_i \in \theta'^{-1}(\omega)} \rho^2(t_i) / |\theta''(t_i)|}. \end{aligned} \quad (47)$$

Further defining

$$\langle t^l \rangle_P = \int_0^\infty \langle t^l \rangle_{P,\omega} P_t(\omega) d\omega, \quad (48)$$

through a direct computation, we have

$$\langle t^l \rangle_P = \langle t^l \rangle_s. \quad (49)$$

4.1.9. Standard deviation of time.

$$\sigma_{p,t/\omega}^2 = \langle t^2 \rangle_{p,\omega} - \langle t \rangle_{p,\omega}^2, \tag{50}$$

and

$$\begin{aligned} \sigma_{p,t}^2 &= \langle t^2 \rangle_p - \langle t \rangle_p^2 \\ &= \int_0^\infty \sum_{t_i \in \theta'^{-1}(\omega)} t_i^2 \rho^2(t_i) / |\theta''(t_i)| d\omega \\ &\quad - \left( \int_0^\infty \sum_{t_i \in \theta'^{-1}(\omega)} t_i \rho^2(t_i) / |\theta''(t_i)| d\omega \right)^2. \end{aligned} \tag{51}$$

4.1.10. Covariance. The mean of  $t\omega$  for the TFD is as follows:

$$\begin{aligned} \langle t\omega \rangle_p &= \int_0^\infty \int_0^{2\pi} t\omega \rho^2(t) \delta(\omega - \theta'(t)) dt d\omega \\ &= \int_0^{2\pi} t \theta'(t) \rho^2(t) dt. \end{aligned} \tag{52}$$

The covariance for TFD is as follows:

$$\begin{aligned} \text{Cov}_{t,\omega}^p &\triangleq \langle t\omega \rangle_p - \langle t \rangle_p \langle \omega \rangle_p \\ &= \int_0^{2\pi} t \theta'(t) \rho^2(t) dt - \langle t \rangle \langle \omega \rangle \\ &\triangleq \text{Cov}_{t,\omega}^s, \end{aligned} \tag{53}$$

where  $\text{Cov}_{t,\omega}^s$  is referred as covariance of the signal [7].

4.1.11. Uncertainty principle for time-frequency distribution. For an MC signal  $s = \rho e^{i\theta(t)}$ , we have the following uncertainty principle for the TFD:

Theorem 2

$$\sigma_{p,\omega}^2 \sigma_{p,t}^2 \geq [\text{COV}_{t,\omega}^s]^2 \geq [\text{Cov}_{t,\omega}^p]^2, \tag{54}$$

where

$$\text{COV}_{t,\omega}^s = \int_0^{2\pi} |(t - \langle t \rangle_p)(\theta'(t) - \langle \theta'(t) \rangle)| \rho^2(t) dt \tag{55}$$

is called *absolute covariance* of  $s$ .

See Appendices for a proof. We remark that it is due to the relation (43) that the usual additive constant 1/4 is not involved in (54), being compared with the classical uncertainty principle. This means that the observation errors for MCs can be significantly reduced.

4.2. Multi-components time-frequency distribution

Assume that  $s$  is a multi-component with unit energy. Performing an orthogonal MC decomposition for  $s$ , we obtain the following:

$$s = \sum_k s_k, \tag{56}$$

where each  $s_k$  is an MC. A realization of orthogonal MC decomposition is  $n$ -Best AFD in which each  $s_k$  is an MC of the form  $\rho_k(t)e^{i\theta_k(t)}$  or  $\rho_k(t) \cos \theta_k(t)$ . We note that if a signal is a finite combination of sinusoid functions of integer-frequencies (a finite Fourier series), then  $n$ -Best AFD for large enough  $n$  can reproduce the expansion.

4.2.1. Explicit formula of instantaneous frequency. The instantaneous amplitude  $\rho_k(t)$  is explicitly given by the following:

$$\rho_k(t) = |\langle s_k, e_{a_k} \rangle B_k(e^{it})| = |\langle s_k, e_{a_k} \rangle e_{a_k}(e^{it})|.$$

Calculation gives

$$\theta'_k(t) = \frac{|a_k| \cos(t - \theta_{a_k}) - |a_k|^2}{1 - 2|a_k| \cos(t - \theta_{a_k}) + |a_k|^2} + \sum_{l=1}^{k-1} \frac{1 - |a_l|^2}{1 - 2|a_l| \cos(t - \theta_{a_l}) + |a_l|^2}, \tag{57}$$

where  $a_k = |a_k|e^{i\theta_{a_k}}$ .

4.2.2. *Time-frequency distribution of multi-components.* Define the TFD of a multi-component  $s$  by the following:

$$P(t, \omega) = \sum_k P_k(t, \omega), \quad (58)$$

where  $P_k$  is the Dirac-type TFD of the composing MC  $s_k$ . Denote

$$\tilde{s} = (s_1, \dots, s_k, \dots), \quad (59)$$

$$\tilde{\hat{s}} = (\hat{s}_1, \dots, \hat{s}_k, \dots), \quad (60)$$

$$|\tilde{s}|^2 = (|s_1|^2, \dots, |s_k|^2, \dots), \quad (61)$$

$$\tilde{\theta}' = (\theta'_1, \dots, \theta'_k, \dots), \quad (62)$$

$$\tilde{\theta}'^l = (\theta'^l_1, \dots, \theta'^l_k, \dots), \quad l = 1, 2, \dots, \quad (63)$$

$$\tilde{\hat{S}} = (\hat{S}_1, \dots, \hat{S}_k, \dots). \quad (64)$$

4.2.3. *Marginal distribution, total energy, and spectrum density.* For a multi-component TFD, there holds

$$P_\omega(t) = \int_0^\infty \sum_k P_k(t, \omega) d\omega = \sum_k |s_k(t)|^2 = |\tilde{s}|^2. \quad (65)$$

The  $t$ -marginal distribution is as follows:

$$\begin{aligned} P_t(\omega) &= \int_0^{2\pi} \sum_k P_k(t, \omega) dt \\ &= \sum_k \sum_{t_{k,i} \in \theta'^{-1}_k(\omega)} \rho_k^2(t_{k,i}) / |\theta'^l_k(t_{k,i})| \\ &= |\tilde{\hat{S}}(\omega)|^2. \end{aligned} \quad (66)$$

The last quantity is called multi-component spectrum density.

The analog of (34) (the Plancherel theorems) is as follows:

$$\|\tilde{\hat{S}}\|^2 = \|\tilde{s}\|^2 = \|\tilde{\hat{S}}(\omega)\|^2 = 1. \quad (67)$$

4.2.4. *Mean of instantaneous frequency and mean of Fourier frequency.* The analogy of (35) is as follows:

$$\begin{aligned} \int_0^\infty \omega |\tilde{\hat{S}}(\omega)|^2 d\omega &= \int_0^{2\pi} \langle \tilde{\theta}'(t), |\tilde{s}|^2(t) \rangle dt \\ &= \int_0^\infty \omega |\tilde{\hat{S}}(\omega)|^2 d\omega, \end{aligned} \quad (68)$$

which can be proved by using (35). The second part of the equal relations in (35) can be extended to the following:

$$\int_0^{2\pi} \langle \tilde{\theta}'^l(t), |\tilde{s}|^2(t) \rangle dt = \int_0^\infty \omega^l |\tilde{\hat{S}}(\omega)|^2 d\omega, \quad l = 1, 2, \dots, \quad (69)$$

while the first part cannot. See the remarks for MCs given in Subsection 4.1.5.

4.2.5. *Conditional mean of instantaneous frequency.* The instantaneous moments for IF are as follows:

$$\begin{aligned} \langle \omega^l \rangle_{P,t} &= \frac{1}{P_\omega(t)} \int_0^\infty \omega^l P(t, \omega) d\omega \\ &= \frac{\sum_k (\theta'^l_k)'(t) |s_k(t)|^2}{\sum_k |s_k(t)|^2}, \quad l = 1, 2, \dots \end{aligned} \quad (70)$$

When  $l = 1$ , we obtain the mean of IF at time moment  $t$ :

$$\bar{\theta}'(t) = \frac{\sum_{k=1}^N \theta'_k(t) \rho_k^2(t)}{\sum_{k=1}^N \rho_k^2(t)}. \quad (71)$$

$\bar{\theta}'(t)$  is time varying that can be treated as a candidate for IF of the original  $f$ . The mean of IF  $\bar{\theta}'(t)$  can be compared with the mean of the Fourier frequency  $\langle l \rangle = \sum |c_l|^2 / \sum |c_l|^2$ , which is a constant.

The global mean of the frequency is as follows:

$$\langle \omega \rangle_P = \int_0^{2\pi} \langle \omega \rangle_{P,t} P_\omega(t) dt = \int_0^{2\pi} \sum_k \theta'_k(t) |s_k(t)|^2 dt. \tag{72}$$

This is the quantity treated in the equality chain (68).

4.2.6. *Standard deviation of frequency.* The conditional standard deviation of IF at the time moment  $t$  for a multi-component is as follows:

$$\begin{aligned} \sigma_{P,\omega/t}^2 &= \langle \omega^2 \rangle_{P,t} - \langle \omega \rangle_{P,t}^2 \\ &= \frac{\sum_k (\theta'_k(t))^2 |s_k(t)|^2}{\sum_k |s_k(t)|^2} - \left( \frac{\sum_k \theta'_k(t) |s_k(t)|^2}{\sum_k |s_k(t)|^2} \right)^2. \end{aligned} \tag{73}$$

Therefore,

$$\begin{aligned} \sigma_{P,\omega}^2 &= \langle \omega^2 \rangle_P - \langle \omega \rangle_P^2 \\ &= \int_0^{2\pi} \sum_k (\theta'_k(t))^2 |s_k(t)|^2 dt - \\ &\quad - \left( \int_0^{2\pi} \sum_k \theta'_k(t) |s_k(t)|^2 dt \right)^2. \end{aligned} \tag{74}$$

4.2.7. *Standard deviation of time.* The conditional moments of time are as follows:

$$\begin{aligned} \langle t^l \rangle_{P,\omega} &= \frac{1}{P_t(\omega)} \int_0^{2\pi} t^l P(t, \omega) dt \\ &= \frac{\sum_k \sum_{t_{k,i} \in \theta_k'^{-1}(\omega)} t_{k,i}^l \rho_k^2(t_{k,i}) / |\theta_k''(t_{k,i})|}{\sum_k \sum_{t_{k,i} \in \theta_k'^{-1}(\omega)} \rho_k^2(t_{k,i}) / |\theta_k''(t_{k,i})|}, \\ & \quad l = 1, 2, \dots \end{aligned} \tag{75}$$

The standard deviation for  $t$  is as follows:

$$\begin{aligned} \sigma_{P,t}^2 &= \langle t^2 \rangle_P - \langle t \rangle_P^2 \\ &= \int_0^\infty \sum_k \sum_{t_{k,i} \in \theta_k'^{-1}(\omega)} t_{k,i}^2 \rho_k^2(t_{k,i}) / |\theta_k''(t_{k,i})| d\omega \\ &\quad - \left( \int_0^\infty \sum_k \sum_{t_{k,i} \in \theta_k'^{-1}(\omega)} t_{k,i} \rho_k^2(t_{k,i}) / |\theta_k''(t_{k,i})| d\omega \right)^2. \end{aligned} \tag{76}$$

## 5. Experiment results

There are four experiments conducted in this paper. Experiment 1 shows that an MC can be expressed either as a single component or as a sum of other MCs. Experiment 2 demonstrates that our proposed  $n$ -Best AFD-based TFD has higher resolution in a specifically designed function by Daubechies. Experiment 3 illustrates that our proposed approach works well to at least certain practical signals. Experiment 4 demonstrates that our proposed  $n$ -Best AFD-based TFD is robust to Gaussian white noise.

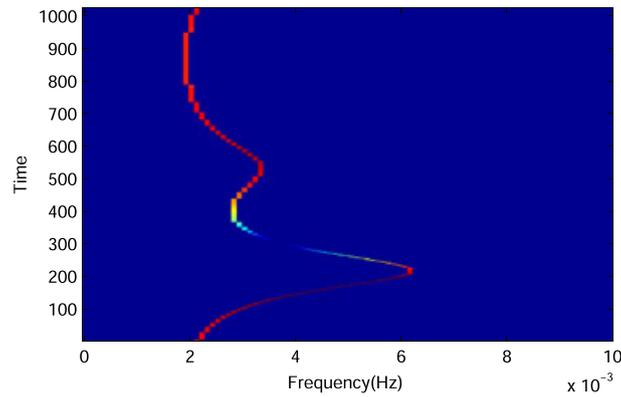
### 5.1. Experiment 1

Graphical illustrations of Example 4 in Section 2 are provided in Experiment 1 with the following parameters:

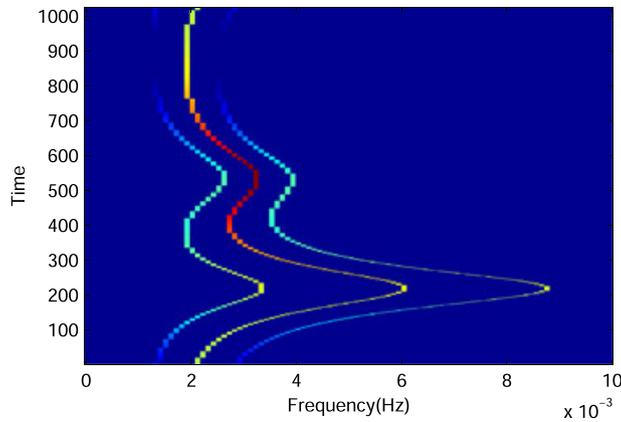
$$a_2 = 0.14495 + 0.62206i,$$

$$a_4 = -0.51325 - 0.075967i.$$

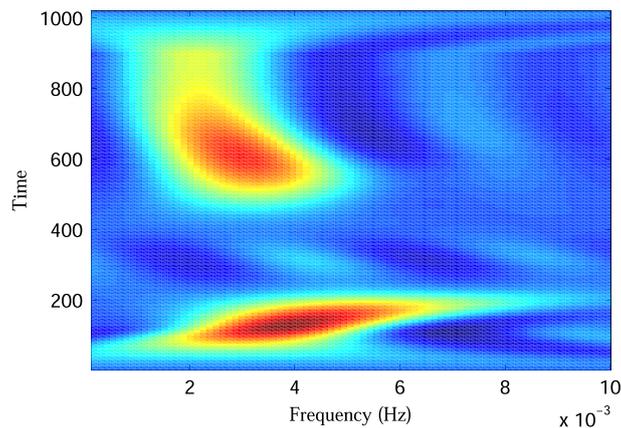
It is an MC. Figure 1 gives an MTFD as a single component. Figure 2 shows that an MTFD can be given by a sum of three NMCs with energy concentration in a restricted area. Figures 3, 4, and 5 show respectively the TFDs by WVD, smoothed WVD, and polynomial WVD. Its WVD, smoothed WVD, and polynomial WVD are plotted by TFSA 6.0 [51] with the parameters 255 of Laq window length in WVD and smoothed WVD and 512 of data window length in polynomial WVD (order 4), respectively. In all the distributions, FFT length is 4096, and frequency zoom max is 0.02. All the other parameters are chosen by default. The parameters of window length and FFT length are optimally selected in the ranges 63–512 and 255–4096, respectively. There is no parameter selection issue for  $n$ -Best AFD-based TFD.



**Figure 1.** The mono-component in Example 4. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



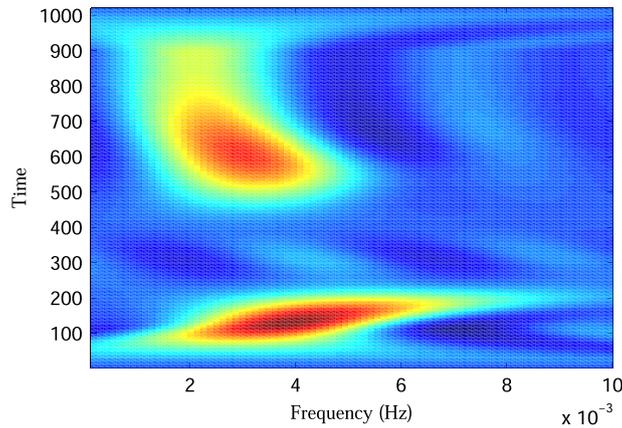
**Figure 2.** The mono-component in Example 4 can be decomposed into a sum of three normal mono-components. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



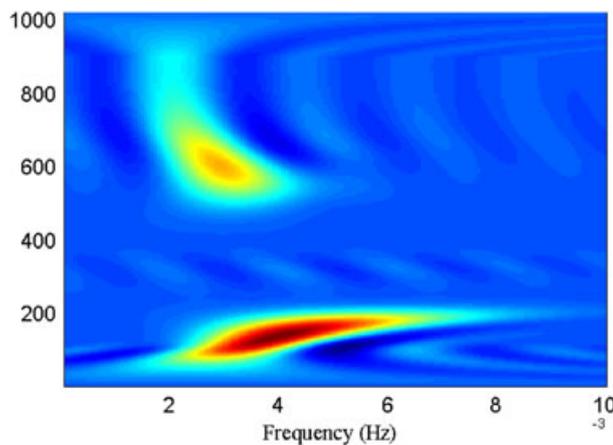
**Figure 3.** The mono-component in Example 4 under Wigner-Ville distribution. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Time-frequency distributions of the Wigner type suffer from the cross-term problem. It is commonly agreed that applied to MCs, however, the cross-term problem does not show in Wigner-type TFDs. By nature, the proposed AFD-based TFDs, including the n-Best AFD-based TFD, do not have cross-term problem at all, no matter being applied to MCs or multi-components. Figures 1–5 offer examples for comparison between various TFDs of the WVD type and that for n-Best AFD applied to the MC given in Example 4. It shows that all being free of cross-terms.

In all the TFDs, the colors indicate the energy densities represented by the square of the amplitude values. The latter three figures show that the WVD methods have similar frequency density concentrations as the MTFDs. The MTFDs, however, are of much clearer resolutions.



**Figure 4.** The mono-component in Example 4 under smoothed Wigner–Ville distribution. [Colour figure can be viewed at wileyonlinelibrary.com]



**Figure 5.** The mono-component in Example 4 under polynomial Wigner–Ville distribution. [Colour figure can be viewed at wileyonlinelibrary.com]

### 5.2. Experiment 2

In Daubechies’s book [16], she designs the signal

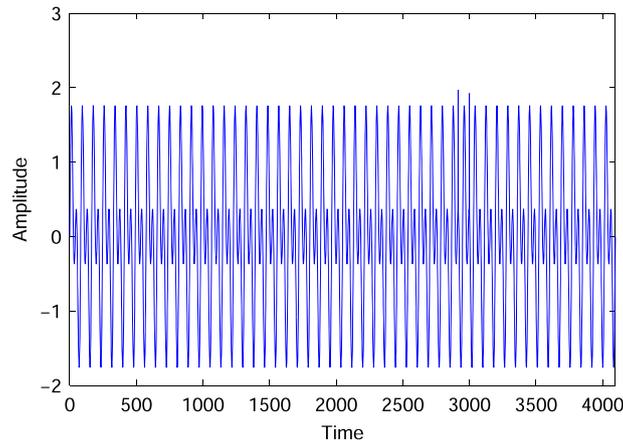
$$f(t) = \sin(2\pi\nu_1 t) + \sin(2\pi\nu_2 t) + \gamma[\delta(t - t_1) + \delta(t - t_2)]. \quad (77)$$

This signal is used to demonstrate that it is hard to find suitable window widths in STFT to do accurate time-frequency analysis. This example is selected in this paper to show that by using the proposed AFD-based TFD, we can obtain a result of high resolution. The original signal is shown in Figure 6. In our experiment, the following parameters are chosen:  $2\pi\nu_1 = 50$ ,  $2\pi\nu_2 = 100$ ,  $t_1 = 2916$ ,  $t_2 = 3000$ , and with 4096 samples. The original signal (77) is not an MC. It is a multi-component and decomposed by AFD. The AFD-based TFD, smoothed WVD, and polynomial WVD are illustrated in Figures 7, 8, and 9, respectively. Smoothed WVD can usually reduce or remove the cross-term effects [11]. It is used in our experiments for comparison of multi-component signals.

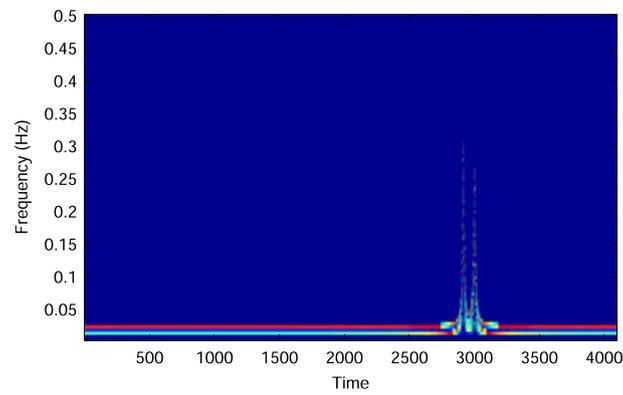
Smoothed WVD and polynomial WVD are plotted by TFSA 6.0 [51]. In smoothed WVD, Laq window length is 255. In polynomial WVD (order 4), data window length is 512. FFT length is 2048. All the other parameters are chosen by default. The parameters of window length and FFT length are selected in the ranges 63–511 and 256–4096, respectively. The parameter values chosen give rise to promising resolutions. The AFD-based TFD shows higher resolution in both the major frequencies and the two pulses, and there is no cross-term problem in our proposed approach.

### 5.3. Experiment 3

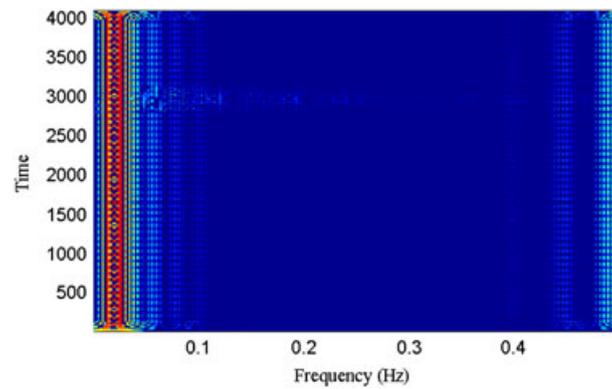
The data of Experiment 3 come from Whale data of the TFSA 6.0 package under demo signals [51]. The data contain 7000 points and were collected at a sample rate of 8 kHz. The first 512 data points are chosen to produce  $n$ -Best AFD-based TFD, smoothed WVD, and polynomial WVD. The latter two WVDs are plotted by TFSA 6.0. The parameters selected in TFSA 6.0 include 255 of Laq window length in smoothed WVD and 256 of data window length in polynomial WVD (order 4), respectively. In both the distribution, FFT length is set to be 1024. All the other parameters are chosen by default. The parameters of window length and FFT length are selected in the ranges 63–512 and 255–4096, respectively. The parameter values chosen result in optimal resolutions. In  $n$ -Best AFD,  $n$  is chosen to be 200, and no further parameter selections are involved.



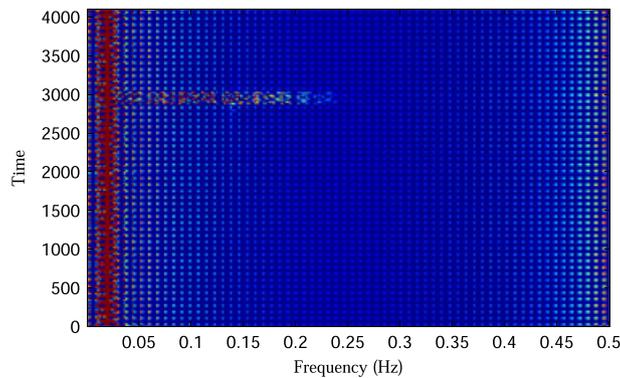
**Figure 6.** The original signal of Experiment 2. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



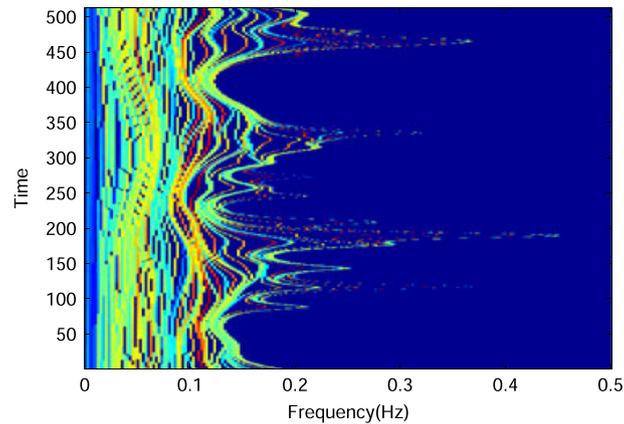
**Figure 7.** The  $n$ -Best adaptive Fourier decomposition-based time-frequency distribution of Experiment 2. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



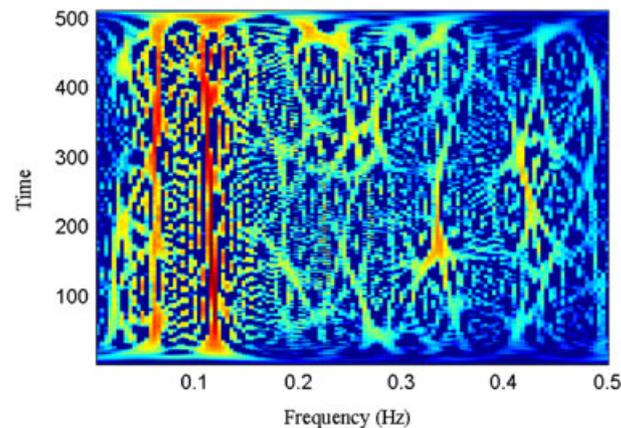
**Figure 8.** The smoothed Wigner–Ville distribution of Experiment 2. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



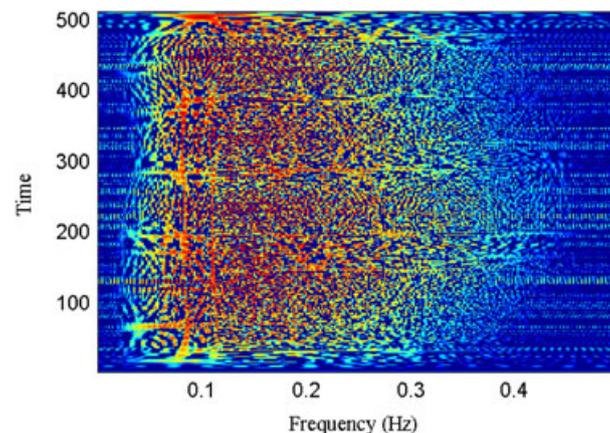
**Figure 9.** The polynomial Wigner–Ville distribution of Experiment 2. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**Figure 10.** The  $n$ -Best adaptive Fourier decomposition-based time-frequency distribution of Experiment 3. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



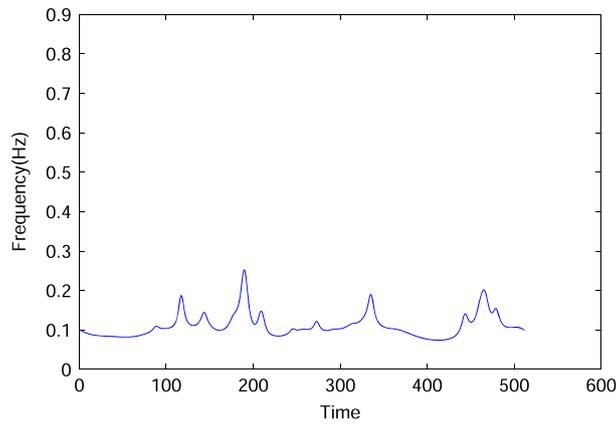
**Figure 11.** The smoothed Wigner–Ville distribution of Experiment 3. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



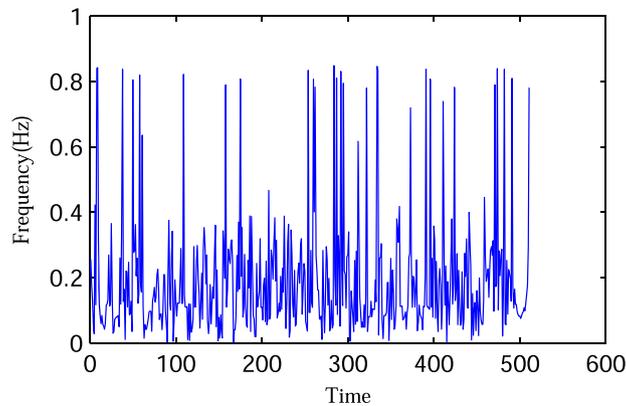
**Figure 12.** The polynomial Wigner–Ville distribution of Experiment 3. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Figures 10, 11, and 12 give respectively multi-components time-frequency distribution ( $n$ -Best AFD-based TFD for multi-components), smoothed WVD, and polynomial WVD, of the data. The latter two are similar but still different, while the first one is apparently a different kind. They represent frequency distributions in their respective ways. The characteristic of the multi-components time-frequency distribution is that the frequency curves do not overlap that illustrate the frequency levels from lower to higher with high resolution.

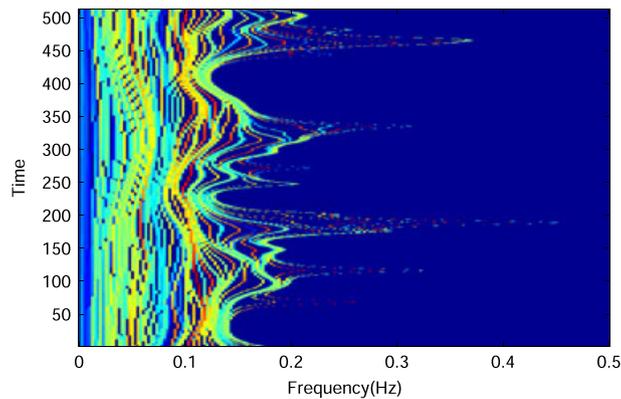
The  $n$ -Best AFD-based mean IF is illustrated in Figure 13. Figure 14 shows the result of the peak of polynomial WVD-based IF estimation. It is plotted by TFSA 6.0 [51] with the parameters 127 of Laq window length and 128 of FFT length. All the other parameters are chosen by default. The experiment results show that AFD-based IF estimation provides reasonable IF values.



**Figure 13.** The  $n$ -Best adaptive Fourier decomposition-based mean instantaneous frequency of Experiment 3. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**Figure 14.** The peak of polynomial Wigner-Ville distribution-based instantaneous frequency of Experiment 3. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**Figure 15.** The  $n$ -Best adaptive Fourier decomposition-based time-frequency distribution of Experiment 3 with Gaussian white noise. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

#### 5.4. Experiment 4

Experiment 4 is conducted for the noise influence analysis to the proposed TFD. The Gaussian white noise is added to the signal treated in Experiment 3 with mean 0 and standard deviation 0.01. In the MC TFD, the noise does not show influence. The  $n$ -Best AFD-based TFD with Gaussian white noise is shown in Figure 15.

## 6. Conclusion

Based on the well-developed AFD theory and, specifically, the  $n$ -Best AFD, a Dirac-type TFD is proposed. It is a practical approach for ideal TFD. It works for all types of signals with finite energy. Based on the orthogonal AFD, the proposed TFD has most commonly desired properties for a TFD, including the following:

1. The TFD is non-negative.
2. The composing frequency values are all non-negative.
3. Marginal or conditional distribution at time  $t$  is identical with the energy density of the signal.
4. For MCs, the time  $t$ -conditional mean of the TFD is identical with the frequency value at the time  $t$ .
5. With the proposed spectrum density, we have a new Plancherel theorem generalizing the classical one.
6. A mean of IF is well defined being a reasonable candidate of the overall IF of the signal.
7. We prove a new uncertainty principle that shows that for MC, the observation errors are significantly reduced and a lower bound is determined by a quantity called absolute covariance.
8. In both the MC and multi-component cases, the one and higher order deviations for the time  $t$  and for the frequency are all given by explicit formulas.
9. The first moments of Fourier frequency are identical with that of IF and that of the TFD frequency. For higher moments, only the latter two are identical. This shows that the TFD frequency is coherent with IF.

Compared with WVD, smoothed WVD, and polynomial WVD, the AFD-based Dirac-type TFD gives better resolution providing detailed and precise information in terms of its composing frequencies. We note that resolutions of the proposed method are robust under interference of zero mean noises because of the integration and the analytic transformations. Our examples on noise-corrupted signals demonstrate this fact.

## Appendix A. Proof of orthogonality of real expansion

*Proof*

Because a TM system is orthogonal, it suffices to show that, under the assumption  $a_1 = 0$ , the real parts of any two orthogonal complex entries in series expansion (26) are also orthogonal. We first assert that any analytic signal has a representation of the form  $x + iHx$ , where  $x$  is a real-valued signal of finite energy. This is in contrast with the conventional  $s^+ = 1/2(s + iHs) + c_0/2$ . In fact, for a real-valued signal  $s$ , we have the following:

$$s^+ = 1/2(s + iHs) + c_0/2 = \frac{s + c_0/2}{2} + iH\left(\frac{s + c_0/2}{2}\right), \tag{A1}$$

verifying the assertion. Let  $B_k$  and  $B_l$  be two different modified Blaschke products of (26),  $l < k$ . Based on the aforementioned assertion, we let  $B_l = x_l + iy_l$ ,  $B_k = x_k + iy_k$ , where  $x_l, y_l, x_k, y_k$  are real-valued functions and  $y_i = Hx_i, i = l, k$ . If  $l = 1$ , then  $B_l = 1$ , and  $\langle 1, B_k \rangle = 0$ , implying  $\langle 1, x_k \rangle = 0$ . Now, assume  $1 < l < k$ . It is easy to prove that the Hilbert transformation satisfies

$$H^* = -H, \quad H^2 = -I + c_0, \tag{A2}$$

where the second relation is a consequence of (6). It is known that

$$\langle B_k, B_l \rangle = 0. \tag{A3}$$

Considering the real part of the left-hand side of (A3), we have the following:

$$\begin{aligned} 0 &= \langle x_k, x_l \rangle + \langle y_k, y_l \rangle \\ &= \langle x_k, x_l \rangle + \langle Hx_k, Hx_l \rangle \\ &= \langle x_k, x_l \rangle + \langle x_k, -H^2x_l \rangle \\ &= \langle x_k, x_l \rangle + \langle x_k, x_l - d_l \rangle \\ &= 2\langle x_1, x_2 \rangle, \end{aligned} \tag{A4}$$

where we used the orthogonality  $\langle 1, x_k \rangle = 0$ ,  $d_l$  is the 0th-Fourier coefficient of  $x_l$ . We thus conclude the orthogonality  $\langle x_1, x_2 \rangle = 0$ .  $\square$

## Appendix B. Proof of uncertainty principle for mono-components

$$\begin{aligned} &\left[ \int_0^{2\pi} |(t - \langle t \rangle_\rho)(\theta'(t) - \langle \theta'(t) \rangle)| \rho^2(t) dt \right]^2 \\ &\leq \int_0^{2\pi} (t - \langle t \rangle_\rho)^2 \rho^2(t) dt \int_0^{2\pi} (\theta'(t) - \langle \theta'(t) \rangle)^2 \rho^2(t) dt \\ &= \int_0^{2\pi} \int_0^\infty (t - \langle t \rangle_\rho)^2 \rho^2(t) \delta(\omega - \theta'(t)) dt d\omega \sigma_{P,\omega}^2 \\ &= \sigma_{P,t}^2 \sigma_{P,\omega}^2. \end{aligned} \tag{B1}$$

On the other hand,

$$\begin{aligned} \text{Cov}_{t,\omega}^p &= \text{Cov}_{t,\omega}^s = \int_0^{2\pi} t\theta'(t)\rho^2(t)dt - \langle t \rangle \langle \omega \rangle \\ &= \int_0^{2\pi} (t - \langle t \rangle_p)(\theta'(t) - \langle \theta'(t) \rangle)\rho^2(t)dt \\ &\leq \text{COV}_{t,\omega}^s. \end{aligned} \tag{B2}$$

## Acknowledgements

This work was supported by Macao FDCT 098/2012/A3 and Multi-Year Research Grant of the University of Macau no. MYRG116(Y1-L3)-FST13-QT and MYRG2014-00009-FST.

## References

- Boashash B. Estimating and interpreting the instantaneous frequency of a signal – Part 1: fundamentals. *Proceedings of The IEEE* 1992; **80**(4):520–538.
- Boashash B, Jones G, O’Shea P. Instantaneous frequency of signals: concepts, estimation techniques and applications. *Proceedings of SPIE - Advanced Algorithms and Architectures for Signal Processing IV*. 1989; **1152**:382–400.
- Boashash B, O’Shea P, Arnold MJ. Algorithms for instantaneous frequency estimation: a comparative study. *Proceedings of SPIE - Advanced Algorithms and Architectures for Signal Processing IV*. 1990; **1348**:126–148.
- Boashash B, Ristic B. 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. *Signal Processing* 1998; **67**(1):1–23.
- Cohen L, Lee C. Instantaneous frequency, its standard deviation and multi-component signals. *Proceeding SPIE - Advanced Algorithms and Architectures for Signal Processing III*. 1988; **975**:186–208.
- Cohen L. Time-frequency distributions – a review. *Proceedings of The IEEE* 1989; **77**(7):941–981.
- Cohen L. *Time-frequency analysis*. Prentice Hall: New Jersey, 1995.
- Kay SM, Marple Jr SL. Spectrum analysis – a modern perspective. *Proceedings of The IEEE* 1981; **69**(11):1380–1419.
- Lovell BC, Williamson RC, Boashash B. The relationship between instantaneous frequency and time-frequency representations. *IEEE Transactions on Signal Processing* 1993; **41**(3):1458–1461.
- Orovic I, Orlandic M, Stankovic S, Uskovi Z. A virtual instrument for time-frequency analysis of signals with highly nonstationary instantaneous frequency. *IEEE Transactions on Instrumentation and Measurement* 2011; **60**(3):791–803.
- Stankovic L. A method for time-frequency analysis. *IEEE Transactions on Signal Processing* 1994; **42**(1):225–229.
- Huang N, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N, Tung CC, Liu HH. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A* 1998; **454**:903–995.
- Zhang L, Liu N, Yu P. A novel instantaneous frequency algorithm and its application in stock index movement prediction. *IEEE Selected Topics in Signal Processing* 2012; **6**(4):311–318.
- Boashash B, Ristic B. 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. *Signal Processing* 1998; **67**:1–23.
- Claassen TACM, Mecklenbrauker WFG. The Wigner distribution – a tool for time-frequency signal analysis, Part III – relations with other time-frequency signal transformations. *Philips Journal of Research* 1980; **35**:276–350.
- Daubechies I. *Ten Lectures on Wavelet*. Society for Industrial and Applied Mathematics: Philadelphia, Pennsylvania, 1992.
- Fink LM. Relations between the spectrum and instantaneous frequency of a signal. *Problemy Peredachi Informatsii* 1966; **2**(4):26–38.
- Mandel L. Interpretation of instantaneous frequencies. *American Journal of Physics* 1974; **42**:840–846.
- Qian S, Chen D. Joint time-frequency analysis. *IEEE Signal Processing Magazine* 1999; **16**(2):52–67.
- Stankovic S, Stankovic L. Introducing time-frequency distribution with a ‘complex-time’ argument. *Electronics Letters* 1996; **32**(14):1265–1267.
- Stankovic L. Time–frequency distributions with complex argument. *IEEE Transactions on Signal Processing* 2002; **50**(3):475–486.
- Sun MG, Scabassi RJ. Discrete-time instantaneous frequency and its computation. *IEEE Transactions on Signal Processing* 1993; **41**(5):1867–1880.
- Xia XG, Cohen L. On analytic signals with nonnegative instantaneous frequencies. *Proceedings of the ICASSAP*. 1999; **3**:1329–1332.
- Qian T. Mono-components for decomposition of signals. *Mathematical Methods in the Applied Sciences* 2006; **29**(10):1187–1198.
- Dang P, Qian T. Analytic phase derivatives, all-pass filters and signals of minimum phase. *IEEE Transactions on Signal Processing* 2011; **59**(10):4708–4718.
- Dang P, Qian T, You Z. Hardy–Sobolev spaces decomposition and applications in signal analysis. *Journal of Fourier Analysis and Applications* 2011; **17**(1):36–64.
- Qian T. Analytic signals and harmonic measures. *Journal of Mathematical Analysis and Applications* 2006; **314**:526–536.
- Qian T. Boundary derivatives of the phases of inner and outer functions and applications. *Mathematical Methods in the Applied Sciences* 2009; **32**:253–263.
- Qian T. Intrinsic mono-components decomposition of functions: an advance of Fourier theory. *Mathematical Methods in the Applied Sciences* 2010; **33**:880–891.
- Qian T, Chen Q, Li L. Analytic unit quadrature signals with non-linear phase. *Physica D: Nonlinear Phenomena* 2005; **303**:80–87.
- Qian T, Ho I, Leong I, Wang Y. Adaptive decomposition of functions into pieces of non-negative instantaneous frequencies. *International Journal of Wavelets, Multiresolution and Information Processing* 2010; **8**(5):813–833.
- Qian T, Sproessig W, Wang J. Adaptive Fourier decomposition of functions in quaternionic Hardy spaces. *Mathematical Methods in the Applied Sciences* 2012; **35**(1):43–64.
- Qian T, Wang Y. Adaptive Fourier series – a variation of greedy algorithm. *Advances in Computational Mathematics* 2011; **34**(3):279–293.
- Qian T, Wang R, Xu Y, Zhang H. Orthonormal bases with nonlinear phases. *Advances in Computational Mathematics* 2010; **33**(1):75–95.

35. Qian T. Cyclic AFD algorithm for best rational approximation. *Mathematical Methods in the Applied Sciences* 2014; **37**(6):846–859.
36. Qian T, Wegert E. Optimal approximation by Blaschke forms. *Complex Variables and Elliptic Equations* 2013; **58**(1):123–133.
37. Walsh JL. *Interpolation and Approximation by Rational Functions in the Complex Domain*. American Mathematical Society (Fourth Edition): Providence, Rhode Island, 1965.
38. Qian T. Characterization of boundary values of functions in Hardy spaces with applications in signal analysis. *Journal of Integral Equations and Applications* 2005; **17**(2):159–198.
39. Qian T, Zhang L, Li H. Mono-components in signal decomposition. *International Journal of Wavelets. Multiresolution and Information Processing*. 2008; **6**(3):353–374.
40. Qian T, Zhang L, Li Z. Algorithm of adaptive Fourier transform. *IEEE Transactions on Signal Processing* 2011; **59**(12):5899–5906.
41. Dang P. Time-frequency analysis based on mono-components, *Ph.D Thesis*, Univrsity of Macau, 2010.
42. Gabor D. Theory of communication. *Journal of the IEE* 1946; **93**(26):429–441.
43. Picinbono B. On instantaneous amplitude and phase of signals. *IEEE Transactions on Signal Processing* 1997; **45**(3):552–560.
44. Garnett JB. *Bounded Analytic Functions*. Academic Press, INC.: San Diego, CA 92101, 1981.
45. Tan L, Shen L, Yang L. Rational orthogonal bases satisfying Bedrosian identity. *Advances in Computational Mathematics* 2010; **33**(3):285–303.
46. Yu B, Zhang H. The Bedrosian identity and homogeneous semi-convolution equations. *Journal of Integral Equations and Applications*. 2008; **20**(4):527–568.
47. Mallat S, Zhang Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing* 1993; **14**(12):3397–3415.
48. Mi W, Qian T. Frequency domain identification: an algorithm based on adaptive rational orthogonal system. *Automatica* 2012; **48**(6):1154–1162.
49. Mi W, Qian T, Wan F. A fast adaptive model reduction method based on Takenaka–Malmquist systems. *Systems & Control Letters* 2012; **61**(1):223–230.
50. Baratchart L, Cardell M, Olivi M. Identification and rational  $L^2$  approximation: a gradient algorithm. *Aotomatica* 1991; **27**(2):413–418.
51. Boashash B. *TFSA 6.0, Time-frequency Signal Analysis Toolbox*. Signal Processing Research Concentration, UQ Centre for Clinical Research, The University of Queensland: Australia, 2011. Available from: <http://time-frequency.net/tf/> (Access date: 15/10/2016).