Design of an Optimal Piece-Wise Spline Wigner-Ville Distribution for TFD Performance Evaluation and Comparison

Mohammad Al-Sa'd, *Member, IEEE*, Boualem Boashash, *Fellow, IEEE*, and Moncef Gabbouj, *Fellow, IEEE*

***Abstract*—This paper proposes a new performance evaluation process for time-frequency distributions (TFD) by designing a reference optimal TFD and novel accuracy and resolution measures. The motivation comes from the need for a TFD performance evaluation method that is objective, capable of quantifying the TFD accuracy and resolution, can determine the performance difference among different TFDs, and suitable for signals with an arbitrary number of components, instantaneous frequency and amplitude. We formulate the proposed optimal TFD, namely the piece-wise spline Wigner-Ville distribution (PW-WVD), by decomposing a standard non-stationary signal model using piece-wise linear frequency modulated (LFM) basis and by exploiting the Wigner-Ville distribution optimality for LFM signals. We compare the designed PW-WVD to conventional optimal TFDs and show that the former is more suitable to serve as a reference for TFD performance evaluation. Using the PW-WVD we derive TFD accuracy and resolution measures, compare them to conventional approaches, and analyze their sensitivity to form a TFD selection criteria. We evaluate the accuracy and resolution of twelve different TFDs and develop precise TFD selection strategies with or without prior information on the signal parameters. Results indicate that the compact kernel distribution is the best performing TFD given no prior information on the signal parameters and different TFDs must be selected upon the availability of prior information.**

***Index Terms*—Multi-component signals, performance evaluation, piece-wise LFM signals, time-frequency distribution, time-frequency measures, time-frequency representation.**

I. Introduction

A TIME-FREQUENCY representation (TFR), obtained by a time-frequency distribution (TFD), describes the spectral contents of a signal through time [1]. It is an adequate representation to analyze and process non-stationary signals where their spectral information varies with time [1]. Time-frequency (TF) signal processing techniques have shown superior results, when compared to time only or frequency only methods, in various applications such as anomaly detection and classification [2]–[5], localization [6], [7], frequency estimation [8]–[10], signal modelling and synthesis [11], [12], and parameter estimation [13]–[15]. The accuracy and resolution of a TFD define its performance [1]. High accuracy ensures minimal artifacts (cross-terms), while high resolution enhances the ability to resolve closely spaced components [1]. Visual assessment and TF measures are viable tools to evaluate the TFD performance [16]. Nonetheless, visual assessment is subjective and TF measures cannot define the performance difference among different TFDs. Therefore, we propose a new TFD performance evaluation process by designing a reference optimal TFD and new accuracy and resolution measures.

The objectives of this paper are: (1) designing an optimal TFD for multi-component non-stationary signals with non-linear instantaneous frequency (IF) and amplitude (IA); (2) and adopting the proposed distribution as a reference to compare and evaluate the performance of TFDs.

The main contributions of the paper are as follows: (1) designing an optimal piece-wise spline Wigner-Ville distribution (PW-WVD); (2) deriving TFD accuracy and resolution measures using the proposed distribution; (3) evaluating the performance of various TFDs using the proposed measures; (4) and providing TFD selection strategies with or without prior information on the signal parameters.

The Wigner-Ville distribution (WVD) yields an optimal TFR for mono-component linear frequency modulated (LFM) signals [1]. However, it is known to suffer from cross-terms when analyzing multi-component or non-linear frequency modulated signals [17]. High performing TFDs are attained by filtering the WVD using a TF kernel that achieves the best trade-off between accuracy and resolution [1]. Various studies adopt TFDs to improve the performance of conventional signal processing methods in different areas such as biomedicine [18], [19], radar [20], [21] and telecommunication [22], [23]. Currently, the TFD performance is optimized and evaluated using TF measures such as ratio of norms [24], energy concentration [25], entropy [26], and application-specific criterion [27]. Energy concentration is used in [28] to optimize the Spectrogram based on the short-time fractional Fourier transform. The window parameters and the fractional angle are adjusted to maximize

Manuscript received September 26, 2020; revised March 15, 2021 and June 1, 2021; accepted June 9, 2021. Date of publication June 15, 2021; date of current version July 27, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lorenzo Galleani. The work was supported in part by NSF IUCRC CVDI AMALIA project funded by Business Finland, and in part by Stroke-Data project funded by Haltian and the Office of the Provost. *(Corresponding author: Mohammad Al-Sa'd.)*

Mohammad Al-Sa'd and Moncef Gabbouj are with the Faculty of Information Technology and Communication Sciences, Tampere University, 33720 Tampere, Finland (e-mail: mohammad.al-sad@tuni.fi; moncef.gabbouj@tuni.fi).

Boualem Boashash is with the Center for Clinical Research, The University of Queensland, Brisbane, QLD 4072, Australia (e-mail: b.boashash@uq.edu.au).

This article has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the authors.

Digital Object Identifier 10.1109/TSP.2021.3089291

concentration. Moreover, a normalized version of this measure is used in [25] to automatically optimize the performance of various TFDs. Rényi entropy is also used to optimize the performance of TFDs by minimizing their complexity [29]. In [24], various Rényi entropy and energy concentration measures are used to analyze and quantify performance. In the same way, application-specific criterion such as classification accuracy and estimation error are common to evaluate the performance of TFDs [30], [31].

A TF measure is a proxy that is correlated with the performance of a TFD. It is useful to compare the performance of different TFDs, but it cannot define the performance difference among those TFDs. The normalized instantaneous resolution (NIR) aims to mitigate the TF measures pitfalls [32], [33]. However, it demands accurate detection for the signal components and a maximum of two components at any time instant with approximately equal amplitudes [34]. The Reinhold measure extends the NIR by relaxing the equal amplitudes restriction [34]. It employs an adaptive detection algorithm to identify the signal components regardless of their IF complexity and amplitude differences. Nonetheless, it is still limited to signals with maximum of two components at any time instant. Consequently, the NIR and Reinhold measures are not suitable when analyzing signals with arbitrary number of components. The utility of optimal distributions as reference for comparison, allows direct evaluation of the TFD accuracy and resolution without relying on proxy measures. Besides, reference distributions are suited to determine the performance difference among different TFDs even when analyzing different signals. Currently, reference TFDs are derived using ideal distributions [28], [35], WVD-based approaches [36], parameterized TFDs [37], or by a Matching Pursuit distribution (MPD) [38], [39]. Ideal TFDs are comprised of knife edges along the signal IF law. Despite that, they are not suitable as a reference because their TF tiling approach results in discontinuities. Moreover, the WVD-based reference distributions are created by summing the WVD of each signal component; hence, their utility is limited to multi-component LFM signals due to the rise of cross-terms. On the other hand, parameterized TFDs are optimal for a specific class of signals. Therefore, they do not generalize to signals with arbitrary number of components and IF/IA complexities. In contrast, the MPD can theoretically serve as a reference distribution, given a dictionary with sufficient diversity [40]. However, its capability is hindered by computational complexity and by the input signal variability [41]. As a result, the MPD is not suitable to serve as a reference distribution when analyzing signals with arbitrary number of components and IF/IA complexities [42]. Driven by the WVD optimality for mono-component LFM signals [1], the theory of linear B-splines [43], [44], and the piece-wise signal model in [45], we design an optimal TFD for multi-component non-stationary signals with non-linear IF/IA laws. Furthermore, we utilize the proposed distribution as a reference to compare and evaluate the performance of various TFDs. We derive novel TFD accuracy and resolution measures and develop precise TFD selection strategies.

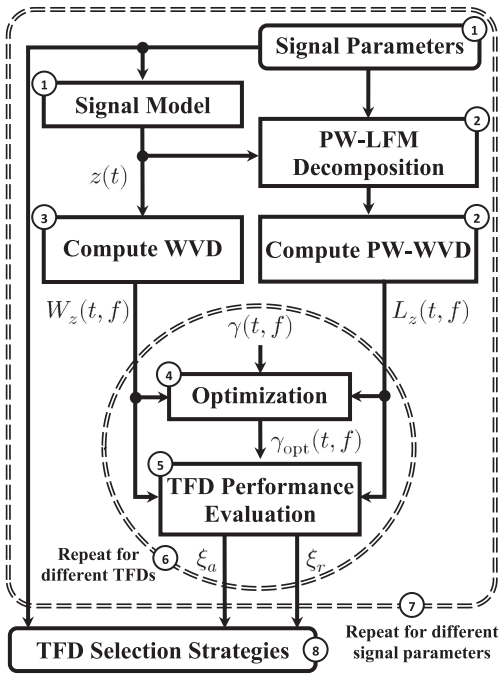


Fig. 1. Flowchart of the proposed TFD performance evaluation process using the optimal PW-WVD. The process steps are summarized as follows: (1) given a set of parameters, generate a multi-component non-stationary signal $z(t)$, (2) compute the optimal PW-WVD $L_z(t,f)$, (3) compute the WVD $W_z(t,f)$, (4) optimize the kernel $\gamma(t,f)$, (5) evaluate the TFD performance in terms of accuracy and resolution, (6) repeat steps 4-5 for different TFDs, (7) repeat the entire process for different signal parameters, and (8) develop TFD selection strategies based on the performance evaluations.

The rest of the paper is organized as follows: Section II formulates the proposed PW-WVD and presents the methodology to evaluate the TFD accuracy and resolution. Afterwards, we present the TFD performance evaluation results and the developed TFD selection strategies in Section III. Finally, Section IV concludes the work.

II. Methodology

The proposed TFD performance evaluation process using the optimal PW-WVD is depicted in Fig. 1 and summarized as follows:

1) Given a set of parameters (number of components, IF, IA, and time support), generate a multi-component signal $z(t)$ using a standard non-stationary signal model.
2) Decompose $z(t)$ using piece-wise LFMs (PW-LFM) and compute its optimal PW-WVD $L_z(t,f)$.
3) Compute the signal WVD $W_z(t,f)$.
4) Given a kernel $\gamma(t,f)$, optimize its parameters by minimizing the difference between $L_z(t,f)$ and the TFD.
5) Using the optimized TFD kernel $\gamma_{\text{opt}}(t,f)$, $W_z(t,f)$ and $L_z(t,f)$, evaluate the TFD performance using the proposed accuracy ξ_a and resolution ξ_r measures.
6) Repeat steps 4-5 for different TFD kernels and aggregate the performance evaluations.

7) Repeat steps 1-6 for different signal parameters and aggregate all performance evaluations.
8) Develop TFD selection strategies when prior information on the signal parameters is available and when it is not.

The remaining subsections detail each step in the proposed TFD performance evaluation process.

A. Signal Model

Multi-component non-stationary signals are comprised of several mono-components and form the general case. A multi-component finite non-stationary signal $z(t)$ is defined as:

$$z(t) = \sum_{p=1}^{P} a_p(t) \exp\left(j2\pi \int_{-\infty}^{\infty} f_p(t)\,\mathrm{d}t\right) \Pi_p(t)\,, \quad (1)$$

$$\Pi_p(t) = \begin{cases} 1 : t_{p,i} \le t \le t_{p,f} \\ 0 : \text{otherwise} \end{cases}, \quad (2)$$

where P is the number of components in $z(t)$, $a_p(t)$ and $f_p(t)$ are the IA and IF of the p*th* component, respectively, $j = \sqrt{-1}$, and $\Pi_p(t)$ is a rectangular window that defines the p*th* component time support from $t_{p,i}$ to $t_{p,f}$.

B. Piece-Wise LFM Decomposition

A continuous function $h \in C\,[0,1]$ can be approximated, given a cardinal finite set of points in $\{t_k\} \subset [0,1]$ and the values $\{h(t_k)\} \forall t_k \in \{t_k\}$, without explicitly knowing its true parametric form [46]. One can formulate a linear B-spline approximation for h by joining the adjacent points $(t_k, h(t_k))$ and $(t_{k+1}, h(t_{k+1}))$ with lines. This approximation becomes a decomposition for h when $\{t_k\}$ is the infinite cardinal set of points between 0 and 1 [46].

The IF of the signal p*th* component, described in Eq. (1), can be decomposed using linear B-splines as follows [44]:

$$f_p(t) = \sum_{q=1}^{\infty} B_{p,q}(t)\,, \quad (3)$$

$$B_{p,q}(t) = \begin{cases} \lambda_{p,q}\,t + \varepsilon_{p,q} & : t_{p,q} \le t < t_{p,q+1} \\ 0 & : \text{otherwise} \end{cases}, \quad (4)$$

where $B_{p,q}(t)$ is the q*th* linear B-spline for $f_p(t)$, $t_{p,i} = t_{p,1} < t_{p,2} < \cdots < t_{p,\infty} = t_{p,f}$ is the infinite cardinal knot sequence, $T_p = t_{p,q+1} - t_{p,q}$ is the infinitesimal support of $B_{p,q}(t)$, $\lambda_{p,q} = (f_p(t_{p,q+1}) - f_p(t_{p,q}))/T_p$ is the slope of $B_{p,q}(t)$, and $\varepsilon_{p,q} = f_p(t_{p,q}) - \lambda_{p,q}\,t_{p,q}$ is its intercept.

Using Eq. (3), the signal definition in Eq. (1) becomes:

$$z(t) = \sum_{p=1}^{P} a_p(t) \exp\left(j2\pi \int_{-\infty}^{\infty} \sum_{q=1}^{\infty} B_{p,q}(t)\,\mathrm{d}t\right) \Pi_p(t)\,. \quad (5)$$

Since $[t_{p,q}, t_{p,q+1}) \cap [t_{p,q+1}, t_{p,q+2}) = \varnothing$ and $B_{p,q}(t) \neq 0$ if and only if $t \in [t_{p,q}, t_{p,q+1})$, Eq. (5) can be expressed in the following alternative form:

$$z(t) = \sum_{p=1}^{P} a_p(t) \underbrace{\left(\sum_{q=1}^{\infty} \psi_{p,q}(t)\right)}_{\text{PW-LFM decomposition}} \Pi_p(t)\,. \quad (6)$$

$$\psi_{p,q}(t) = \begin{cases} e^{\left(j2\pi \int_{-\infty}^{\infty} \lambda_{p,q}\,t + \varepsilon_{p,q}\,\mathrm{d}t\right)} & : t_{p,q} \le t < t_{p,q+1} \\ 0 & : \text{otherwise} \end{cases}. \quad (7)$$

By comparing Eq. (1) to Eq. (6), one observes that the term $\sum_{q=1}^{\infty} \psi_{p,q}(t)$ is a PW-LFM decomposition of the p*th* complex exponential in Eq. (1). This decomposition allows us to exploit the WVD optimality for LFM signals to design the proposed optimal PW-WVD.

C. The Piece-Wise Spline Wigner-Ville Distribution

Given an analytic signal $z(t)$, the WVD is defined as [1]:

$$W_z(t,f) = \underset{\tau \to f}{\mathcal{F}} \{K_z(t,\tau)\}\,, \quad (8)$$

$$K_z(t,\tau) = z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right), \quad (9)$$

where $K_z(t,\tau)$ is the instantaneous autocorrelation function (IAF) of $z(t)$, $W_z(t,f)$ is the WVD, $z^*(t)$ is the complex conjugate of $z(t)$, and $\underset{\tau \to f}{\mathcal{F}}$ denotes the Fourier transform from lag τ to frequency f. The WVD is known to suffer from cross-terms[1]. Nonetheless, it is optimal for mono-component LFM signals. This optimality allows us to formulate the PW-WVD using the signal PW-LFM decomposed model.

1) Formulation: Using Eq. (6), the IAF of $z(t)$ auto-terms is expressed as:

$$K_z(t,\tau) = \sum_{p=1}^{P} K_{a_p}(t,\tau)\, K_{\psi_p}(t,\tau)\, K_{\Pi_p}(t,\tau)\,. \quad (10)$$

$$K_{a_p}(t,\tau) = a_p\left(t + \frac{\tau}{2}\right) a_p\left(t - \frac{\tau}{2}\right). \quad (11)$$

$$K_{\psi_p}(t,\tau) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \psi_{p,i}\left(t + \frac{\tau}{2}\right) \psi_{p,j}^*\left(t - \frac{\tau}{2}\right). \quad (12)$$

$$K_{\Pi_p}(t,\tau) = \Pi_p\left(t + \frac{\tau}{2}\right) \Pi_p\left(t - \frac{\tau}{2}\right). \quad (13)$$

Since, by definition, $\psi_{p,i}(t) \cap \psi_{p,j}(t) = \varnothing$ if $i \neq j$, the product terms in Eq. (12) describe the signal auto-terms only when $i = j$. Therefore, $K_{\psi_p}(t,\tau)$ reduces to:

$$K_{\psi_p}(t,\tau) = \sum_{q=1}^{\infty} \psi_{p,q}\left(t + \frac{\tau}{2}\right) \psi_{p,q}^*\left(t - \frac{\tau}{2}\right). \quad (14)$$

Using Eqs. (8) and (10), the PW-WVD is defined as:

$$L_z(t,f) = \sum_{p=1}^{P} W_{a_p}(t,f) \underset{f}{*} \left(\sum_{q=1}^{\infty} W_{\psi_{p,q}}(t,f)\right) \underset{f}{*} W_{\Pi_p}(t,f)\,, \quad (15)$$

[1]The WVD cross-terms are comprised of outer and inner-terms [1]. Outer-terms emerge when $z(t)$ is a multi-component signal, while inner-terms evolve for non-linear frequency modulated signals [17]. Appendix A presents a formulation for the WVD inner-terms using the binomial theorem.

where $L_z(t,f)$ is the PW-WVD of $z(t)$, $W_{a_p}(t,f)$ is the WVD of $a_p(t)$, $W_{\Pi_p}(t,f)$ is the WVD of $\Pi_p(t)$, and $W_{\psi_{p,q}}(t,f)$ is the WVD of the PW-LFM in $\psi_{p,q}(t)$, i.e.:

$$W_{\psi_{p,q}}(t,f)=\begin{cases}\delta\left(f-\lambda_{p,q}\,t-\varepsilon_{p,q}\right) & : t_{p,q}\leq t<t_{p,q+1}\\ 0 & : \text{otherwise}\end{cases}. \quad (16)$$

The PW-WVD is optimal given the signal parameters $(f(t),a(t),\Pi(t),P)$. Besides, it provides an approach to separate[2] the WVD cross-terms $\mathcal{X}_z(t,f)$ from its auto-terms, i.e.:

$$\mathcal{X}_z(t,f)=W_z(t,f)-L_z(t,f)\,. \quad (17)$$

Fig. 2 illustrates the PW-WVD, WVD, and the WVD separated cross-terms for an example 2-component non-stationary signal. In addition, Appendix B discusses the PW-WVD optimality and its properties, while Appendix C presents the discrete PW-WVD formulation.

In practice when analyzing measured signals, utilizing the PW-WVD requires estimation of the signal parameters. Therefore, the PW-WVD accuracy and resolution solely depend on the estimation quality that is acceptable to the user. Nonetheless, in this work, we compute the optimized PW-WVD using the input signal true parameters to serve as a reference distribution in the proposed TFD performance evaluation process.

2) Comparison With the Ideal and Matching Pursuit Distributions: The ideal TFD of a finite multi-component non-stationary signal is defined as [1, Sec. 2.1.4.1]:

$$I_z(t,f)=\sum_{p=1}^{P} W_{a_p}(t,f)\underset{f}{*}\delta\left(f-f_p(t)\right)\underset{f}{*}W_{\Pi_p}(t,f)\,. \quad (18)$$

By comparing Eq. (18) to the proposed PW-WVD in Eq. (15), one observes that the two approaches construct the signal TFR differently. In Eq. (18), the signal is decomposed using independent sinusoid basis. This corresponds to tiling the TF domain uniformly along the frequency axis; see Fig. 3(a). In other words, the ideal TFD is thought of as an accumulation of infinitely many independent frequency representations. In contrast, the PW-WVD uses PW-LFM basis. Consequently, the TF domain tiling adapts to the IF orientation at each time instant; see Fig. 3(b). As a result, the designed PW-WVD is thought of as the sum of infinitely many optimal WVDs. In practice, when analyzing discrete signals, the number of independent frequency representations in the ideal TFD and the number of optimal WVDs in the PW-WVD are finite. This leads to a major difference between the two distributions. On one hand, the ideal TFD suffers from discontinuities because each frequency representation (or time-slice) is independent from its neighbors. As a result, the ideal TFD fails to link the transition between consecutive TF samples; see the outlined green region in Fig. 3(a). On the other hand, the proposed PW-WVD eliminates such discontinuities by linearizing all TF transitions; it links every pair of samples with a PW-LFM that is aligned with the IF orientation, see Fig. 3(b).

[2]The expression in Eq. (17) can be extended to separate the contribution of inner and outer-terms using the formulation presented in Appendix A.

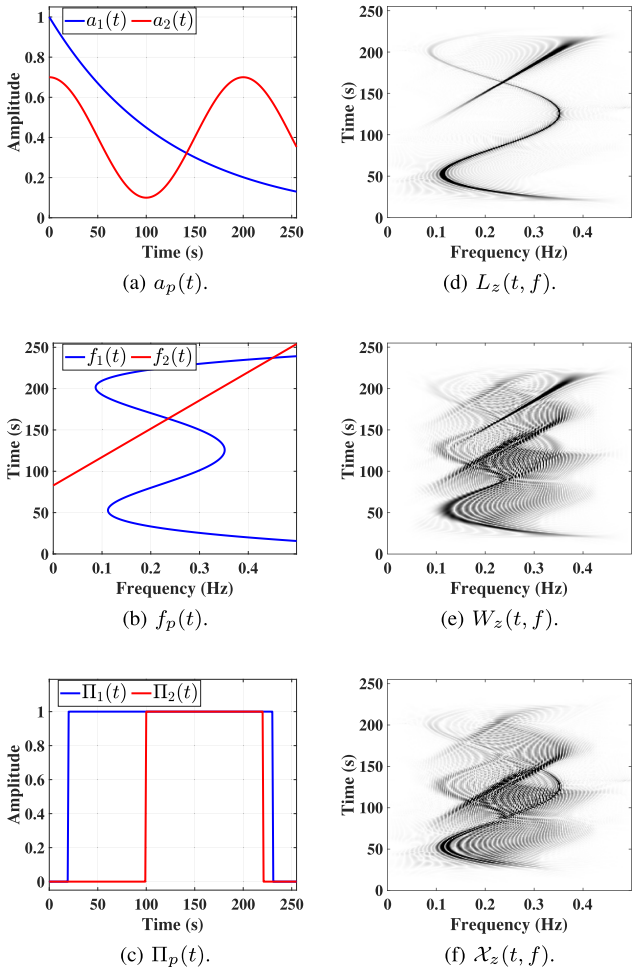
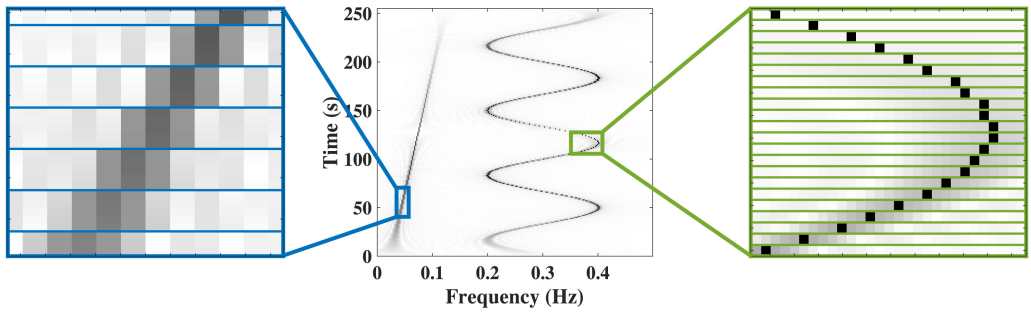
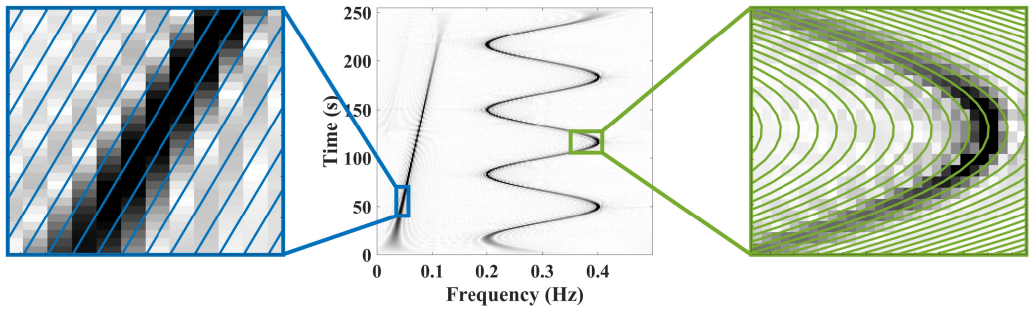


Fig. 2. The auto-terms and cross-terms of an example 2-component non-stationary signal described by the signal model in Eq. (1). The signal IA, IF, and time-support functions are demonstrated in (a-c), respectively. The PW-WVD shows the signal auto-terms in (d). The signal WVD and its separated cross-terms are shown in (e) and (f), respectively. The signal first component is characterized by an exponentially decaying IA in the form: $\exp(-0.008t)$, a 4*th* order polynomial IF that passes through the TF points $\{(20,0.4),(80,0.2),(130,0.35),(180,0.15),(230,0.3)\}$, and a time support between 20 and 230 seconds. The signal second component is characterized by a sinusoidal IA in the form: $0.4+0.3\cos(0.01\pi t)$, a linear IF that passes through $\{(100,0.05),(220,0.4)\}$, and a time support between 100 and 220 seconds. The simulation parameters are: $N=256$ samples, $f_s=1$ Hz and $M=256$ samples.

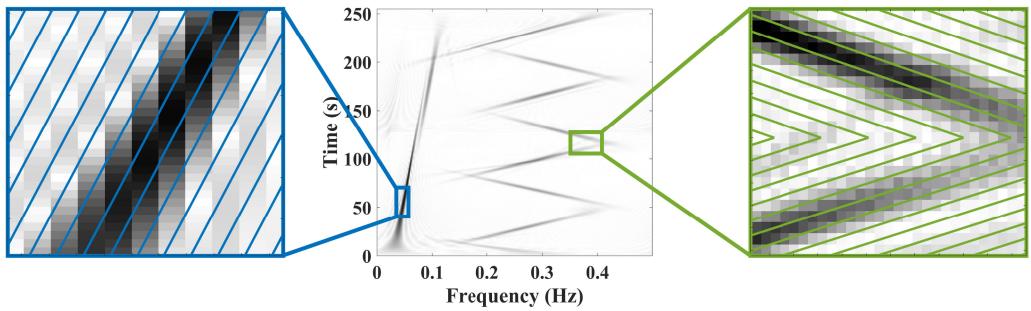
The MPD is an iterative cross-terms free TFD [38], [39]. Nonetheless, its performance is hindered by computational complexity and is influenced by the signal variability [42]. The MPD requires a dictionary with diverse TF atoms to match a wide range of signal patterns. In this work, we utilize a chirplet dictionary that holds approximately 6 million different atoms; see Appendix D for more details. By comparing the MPD in Fig. 3(c) to the PW-WVD in Fig. 3(b), one notes that the two TFDs perform similarly across local structures that are members of the chirplet dictionary, e.g. the LFM region outlined in blue. However, in regions with non-linear structure, as the one outlined in green, the MPD fails to concentrate the signal energy around its IF because the pattern is not included in



(a) The ideal TFD $I_z(t, f)$.



(b) The optimal PW-WVD $L_z(t, f)$.



(c) The MPD with chirplet atoms $M_z(t, f)$.

Fig. 3. Comparison between the ideal TFD, MPD, and the proposed optimal PW-WVD using an example 2-component non-stationary signal. The signal first component is characterized by a unit constant IA, a sinusoidal IF in the form: $0.3 - 0.1\sin(0.03\pi\, t)$, and a time support between 0 and 255 seconds. The signal second component is characterized by a decaying IA in the form: $\exp(-0.0037\, t)$, a linear IF in the form: $0.025 + (0.1/255)\, t$, and a time support between 0 and 255 seconds. Two example regions of interest, outlined by blue and green rectangles in all subfigures, are enlarged for better visualization and interpretation. The first region in blue (on the left) shows a local linear pattern, while the second region in green (on the right) depicts a local non-linear structure. The TF domain tiling of each TFD is demonstrated on the regions of interest by solid lines. Note that the number of tiling lines is adjusted to ease visualization. The simulation parameters are: $N = 256$ samples, $f_s = 1$ Hz and $M = 256$ samples.

its dictionary. Increasing the MPD dictionary size to account for the signal variability boosts its performance, but it also escalates the complexity, i.e. an exact decomposition becomes an NP-hard problem [41]. Besides, according to the uncertainty principle, using many short-time atoms to match the signal temporal shape lowers the TFD resolution [47]. In conclusion, the designed PW-WVD is more suitable to serve as a reference distribution to compare and evaluate the TFDs accuracy and resolution.

D. TFD Performance Optimization and Evaluation

High performing TFDs are attained by filtering the WVD using a TF kernel $\gamma(t, f)$ as follows:

$$\rho_z(t, f) = \gamma(t, f) \underset{(t,f)}{**} W_z(t, f)\,, \tag{19}$$

where $\rho_z(t, f)$ is a smoothed TFD and $\underset{(t,f)}{**}$ denotes the convolution operation in time and frequency [1]. The TF kernel $\gamma(t, f)$ suppresses the WVD cross-terms, but it also smears the signal auto-terms. Hence, $\gamma(t, f)$ must achieve the best trade-off between accuracy and resolution.

1) Optimization: Using the proposed PW-WVD, we optimize the TFD performance by minimizing the sum square error (SSE) between $\rho_z(t, f)$ and $L_z(t, f)$, i.e.:

$$\min_{\gamma(t,f)} \left\{ \iint \left[L_z(t, f) - \rho_z(t, f)\right]^2 \mathrm{d}t\, \mathrm{d}f \right\}. \tag{20}$$

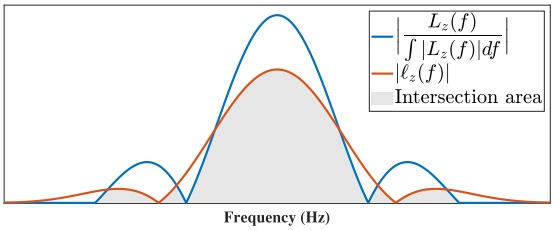


Fig. 4. One dimensional visual interpretation for the TFD resolution measure ξ_r in Eq. (23). $L_z(f)$ and $\ell_z(f)$ are the original and smeared auto-terms, respectively, and $\xi_r = 0.7080$ in this example.

Minimizing the SSE maximizes the TFD accuracy and resolution and yields an optimal TF kernel $\gamma_{\text{opt}}(t,f)$, i.e.:

$$\gamma_{\text{opt}}(t,f) = \arg\min_{\gamma(t,f)} \left\{ \iint \left[L_z(t,f) - \rho_z(t,f)\right]^2 \, \mathrm{d}t \, \mathrm{d}f \right\}. \quad (21)$$

In this work, we use the Bayesian optimization algorithm in MATLAB to optimize the performance of the TFDs listed in Table II. The optimization is executed with random initial kernel parameters, for 200 iterations, and by using the expected improvement plus acquisition function [48]. Note that optimization of the SPWVD kernel uses the PWVD optimal parameters as initial solution.

2) Evaluation: Using Eqs. (17) and (21), we quantify the TFD accuracy by measuring the relative amount of unfiltered cross-terms, i.e.:

$$\xi_a = 1 - \iint \left| \gamma_{\text{opt}}(t,f) \underset{(t,f)}{**} \frac{\mathcal{X}_z(t,f)}{\iint |\mathcal{X}_z(t,f)| \, \mathrm{d}t \, \mathrm{d}f} \right| \mathrm{d}t \, \mathrm{d}f \,, \quad (22)$$

where ξ_a is the TFD accuracy which is 0 when cross-terms are not altered and 1 when they are completely removed. Likewise, we quantify the TFD resolution by measuring the relative intersection area between the PW-WVD original auto-terms and their smeared counterparts, i.e.:

$$\xi_r = \iint \min \left\{ \left| \frac{L_z(t,f)}{\iint |L_z(t,f)| \mathrm{d}t \, \mathrm{d}f} \right|, |\ell_z(t,f)| \right\} \mathrm{d}t \, \mathrm{d}f \,, \quad (23)$$

$$\ell_z(t,f) = \gamma_{\text{opt}}(t,f) \underset{(t,f)}{**} \frac{L_z(t,f)}{\iint |L_z(t,f)| \mathrm{d}t \, \mathrm{d}f} \,, \quad (24)$$

where ξ_r is the TFD resolution which is 0 when auto-terms are completely smeared and 1 when they are not altered. Fig. 4 presents a one-dimensional interpretation for ξ_r.

In this work, we use ξ_a, ξ_r, and their average ξ to evaluate the performance of the 12 TFDs formulated in Appendix D. Apart from the proposed measures, we have considered the error, divergence, and correlation between the PW-WVD and the TFD under analysis. We found that the proposed measures are more suitable because they yield values that are: bounded between 0 and 1, have physical meaning and interpretation, measure actual correspondence not similarity, aligned with the definitions of TFD accuracy and resolution, and finally, they decouple the TFD performance into accuracy and resolution measures. In spite of that, performance evaluation is application dependent; therefore, alternative measures are relevant in other applications.

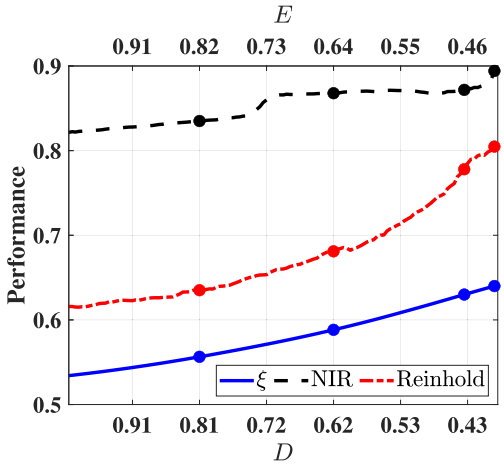


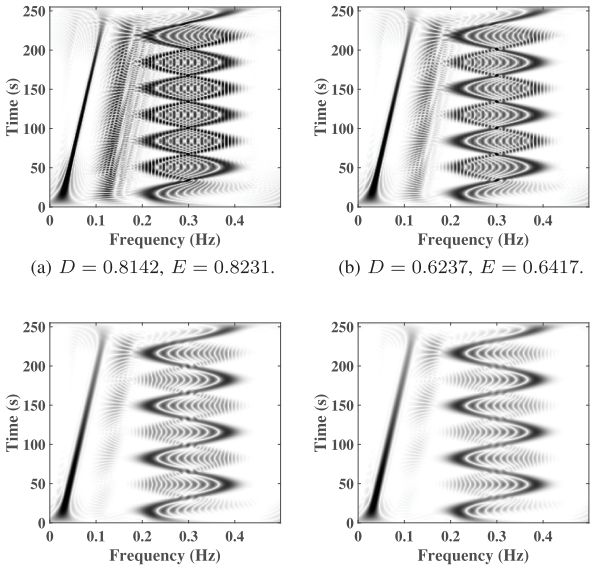
Fig. 5. Comparison between the proposed average TFD performance measure ξ and the NIR and Reinhold measures using the CKD of the example signal shown in Fig. 3. The four increasing performance levels are illustrated by filled circles and their corresponding CKDs are plotting in Fig. 6.

For instance, error is suitable to evaluate component separation algorithms when tested using the designed PW-WVD.

3) Comparison With the NIR and Reinhold Measures: The proposed average TFD performance measure ξ is compared with the NIR and Reinhold measures using the example signal in Fig. 3 and the CKD. First, we optimize the CKD kernel parameters (c, D, E) using the Bayesian optimization algorithm to minimize the cost function in Eq. (20). After that, we fix the CKD shape parameter to its optimal value, i.e. $c = 2.995$, and compute the performance measures for a range of (D, E) values starting from (1,1) and ending at the global optimal solution, i.e. (0.3951,0.424); see Fig. 5. In addition, we visualize the TFD performance progression by computing the CKD at four increasing performance levels; see Fig. 6.

By examining Fig. 5, one observes that ξ shows a similar but more robust trend when compared to the NIR and Reinhold measures. For instance, the NIR overestimates the TFD performance such that it yields high values even for TFDs that are abundant in cross-terms; see Fig. 6(a). On the other hand, the Reinhold measure extends the NIR and reduces its overestimation error. However, as its predecessor, it is still limited to signals with maximum of two components at any time instant. In contrast, ξ is independent of the signal number of components and generates bounded values that correspond to the actual TFD performance; it does not overestimate. Consequently, the proposed average measure ξ is more appropriate to quantify the TFD accuracy and resolution.

4) TFD Selection Criteria: By inspecting the CKDs in Fig. 6(c) and Fig. 6(d), one deduces that a 0.01 drop in ξ results in slightly noticeable accuracy and resolution degradation. Therefore, in this work, we select a set of best performing TFDs according to the following criteria; if two TFDs exhibit a performance difference less/more than 0.01, they are considered equivalent/different in terms of performance.



(a) $D = 0.8142$, $E = 0.8231$. (b) $D = 0.6237$, $E = 0.6417$.

(c) $D = 0.4379$, $E = 0.4648$. (d) $D = 0.3951$, $E = 0.424$.

Fig. 6. The CKD at four increasing performance levels are shown in (a)-(d) in ascending order along with their (D, E) kernel parameters. Note that the CKD shape parameter c is fixed to its optimal value, i.e. $c = 2.995$. The TFD performance measures in each sub-figure are as follows: (a) $\xi = 0.556$, NIR $= 0.835$, Reinhold $= 0.635$, (b) $\xi = 0.588$, NIR $= 0.868$, Reinhold $= 0.681$, (c) $\xi = 0.63$, NIR $= 0.872$, Reinhold $= 0.778$, and lastly, (d) $\xi = 0.64$, NIR $= 0.894$, Reinhold $= 0.805$. The simulation parameters are: $N = 256$ samples, $f_s = 1$ Hz and $M = 256$ samples.

III. Results and Discussion

A. Database

The proposed TFD performance evaluation process, depicted in Fig. 1, requires a set of test signals that are capable of exploring the TFDs advantages and limitations. In this work, we generate 1000 multi-component non-stationary signals following the model in Eq. (1). The test signals are sampled at 1 Hz and characterized by random number of components between 1 and 4, polynomial IF laws with random orders[3] between 1 and 3, random constant IA between 0.5 and 1, and random time support within 0 and 255. The signal random time support is constrained with minimum and maximum IF curve lengths to produce signals with realistic durations.

B. Performance Evaluation

The performance of each utilized TFD is optimized for each test signal and evaluated using the proposed measures. The performance evaluation results for each TFD are averaged across all test signals and summarized in Fig. 7. According to the TFD selection criteria in Section II-D4, the averaged performance trend ξ in Fig. 7 indicates that the CKD and SPWVD form the set of best performing TFDs. Closer inspection of the averaged accuracy and resolution show that the CKD and SPWVD jointly maximize their TFD accuracy and resolution.

[3]An IF polynomial of order R is generated such that when regressed by a polynomial of order $R-1$, the adjusted r-square value is below 0.95. This is done to validate the generated IF complexity.

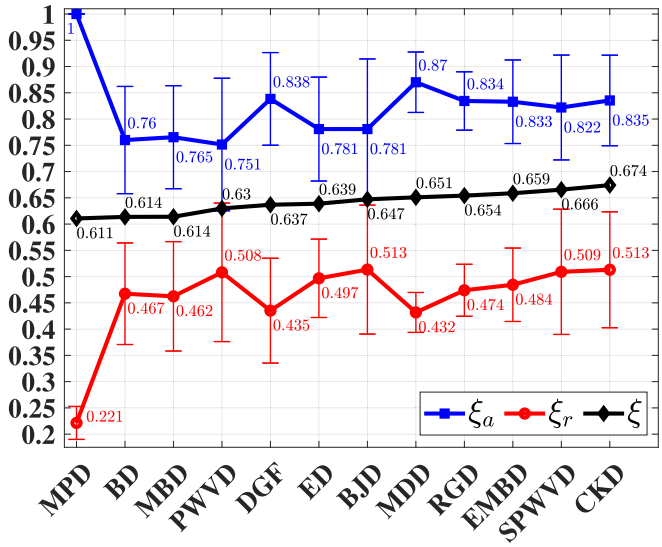


Fig. 7. The TFD performance evaluation results using the proposed accuracy and resolution measures. ξ_a is the TFD accuracy, ξ_r is the TFD resolution, and ξ is their average. Error bars indicate the 1-sigma confidence interval. TFDs on the x-axis are sorted according to their averaged performance.

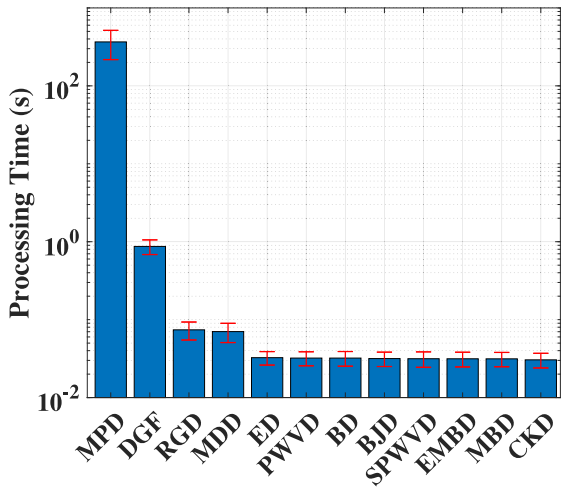


Fig. 8. The TFD computational complexity measured in terms of averaged processing time. Error bars indicate the 2-sigma confidence interval. TFDs on the x-axis are sorted according to their averaged processing time.

Therefore, their utility is preferred when the TFD accuracy and resolution are equally important. Additionally, by examining the averaged accuracy, one observes that the MDD and DGF are highly accurate but with low resolution. By investigating the MDD and DGF kernel formulations in Table II, we suspect that their low resolution is due to the MDD constant minor axis support and the DGF excessive directional filtering. As a result, these limitations prevent them from achieving highest overall performance. Moreover, the MPD with chirplet dictionary shows the lowest averaged performance, because according to the uncertainty principle, selecting many short-time atoms to match the IF non-linearity leads to low resolution.

TABLE I
THE TFD PERFORMANCE EVALUATION RESULTS WITH RESPECT TO THE NUMBER OF SIGNAL COMPONENTS P AND THE IF POLYNOMIAL ORDERS R. RESULTS ARE SUMMARIZED USING THE PROPOSED AVERAGED PERFORMANCE MEASURE ξ AND ROUNDED TO THREE DECIMAL PLACES TO ACCOUNT FOR SENSITIVITY. FOLLOWING THE TFD SELECTION CRITERIA IN SECTION II-D4, BEST PERFORMING TFDS ARE IN BOLD AND SHADED IN GRAY

P	R	MPD	BD	MBD	PWVD	DGF	ED	BJD	MDD	RGD	EMBD	SPWVD	CKD
1	(1)	0.609	0.866	0.932	**1**	0.841	0.857	**1**	0.747	0.816	0.873	**0.998**	**0.997**
	(2)	0.596	0.585	0.583	0.681	0.638	0.633	0.624	0.616	0.638	0.648	0.66	**0.697**
	(3)	0.601	0.57	0.573	**0.674**	0.625	0.625	0.618	0.611	0.632	0.637	0.67	**0.684**
2	(1,1)	0.615	0.608	0.582	0.588	0.609	0.591	0.578	0.653	**0.667**	0.627	0.656	**0.662**
	(1,2)	0.611	0.605	0.598	0.639	0.641	0.627	0.621	0.639	0.648	0.648	0.663	**0.678**
	(1,3)	0.604	0.593	0.59	0.61	0.615	0.613	0.617	0.635	0.647	0.634	0.648	**0.659**
	(2,2)	0.603	0.598	0.597	0.64	0.649	0.629	0.634	0.641	0.638	0.651	0.663	**0.674**
	(2,3)	0.604	0.593	0.592	0.632	0.642	0.625	0.631	0.64	0.642	0.647	**0.657**	**0.666**
	(3,3)	0.606	0.592	0.591	0.631	0.638	0.627	0.635	0.64	0.644	0.648	0.653	**0.668**
3	(1,1,1)	0.618	0.606	0.606	0.577	0.605	0.61	0.614	**0.671**	**0.671**	0.638	0.643	0.658
	(1,1,2)	0.613	0.614	0.614	0.591	0.614	0.623	0.628	**0.651**	**0.65**	0.647	0.648	**0.659**
	(1,1,3)	0.614	0.608	0.607	0.615	0.631	0.629	0.637	0.646	0.647	0.648	**0.649**	**0.659**
	(1,2,2)	0.61	0.615	0.614	0.61	0.63	0.629	0.635	0.644	0.647	0.651	**0.656**	**0.662**
	(1,2,3)	0.609	0.607	0.606	0.615	0.63	0.631	0.638	0.648	0.644	**0.652**	**0.654**	**0.662**
	(1,3,3)	0.609	0.6	0.599	0.613	0.621	0.631	0.636	0.639	**0.645**	**0.646**	**0.645**	**0.655**
	(2,2,2)	0.61	0.611	0.609	0.617	0.648	0.635	0.644	0.649	0.648	**0.656**	**0.658**	**0.666**
	(2,2,3)	0.608	0.605	0.604	0.623	0.647	0.635	0.643	0.65	0.647	**0.657**	**0.656**	**0.662**
	(2,3,3)	0.607	0.598	0.597	0.618	0.627	0.632	0.637	0.642	0.643	**0.651**	**0.65**	**0.658**
	(3,3,3)	0.607	0.596	0.595	0.619	0.639	0.639	0.642	**0.648**	0.645	**0.654**	**0.651**	**0.657**
4	(1,1,1,1)	0.621	0.622	0.622	0.584	0.591	0.617	0.627	**0.66**	**0.656**	0.643	0.643	**0.652**
	(1,1,1,2)	0.616	0.617	0.615	0.608	0.625	0.636	0.644	0.652	0.654	**0.655**	**0.657**	**0.665**
	(1,1,1,3)	0.62	0.614	0.614	0.608	0.589	0.637	**0.646**	**0.646**	0.645	**0.65**	**0.649**	**0.656**
	(1,1,2,2)	0.616	0.614	0.614	0.615	0.641	0.636	0.644	**0.654**	0.651	**0.655**	**0.656**	**0.662**
	(1,1,2,3)	0.613	0.622	0.621	0.615	0.626	0.64	0.648	0.65	0.65	**0.658**	**0.658**	**0.662**
	(1,1,3,3)	0.617	0.616	0.615	0.618	0.636	0.643	0.649	0.651	0.652	**0.661**	**0.66**	**0.665**
	(1,2,2,2)	0.614	0.616	0.614	0.624	**0.667**	0.643	0.652	0.657	0.656	**0.664**	**0.664**	**0.669**
	(1,2,2,3)	0.61	0.612	0.612	0.618	0.646	0.644	0.645	**0.657**	0.652	**0.661**	**0.659**	**0.664**
	(1,2,3,3)	0.617	0.607	0.606	0.618	0.605	0.645	0.646	**0.652**	0.651	**0.661**	**0.658**	**0.662**
	(1,3,3,3)	0.612	0.608	0.607	0.62	0.637	0.644	0.649	0.653	**0.656**	**0.664**	**0.659**	**0.666**
	(2,2,2,2)	0.611	0.62	0.619	0.617	0.638	0.641	0.649	**0.66**	0.652	**0.662**	**0.661**	**0.668**
	(2,2,2,3)	0.609	0.611	0.61	0.621	0.649	0.642	0.647	**0.656**	0.648	**0.661**	**0.659**	**0.661**
	(2,2,3,3)	0.611	0.609	0.609	0.618	0.636	0.645	0.649	**0.66**	0.653	**0.663**	**0.659**	**0.665**
	(2,3,3,3)	0.612	0.609	0.608	0.618	0.64	0.645	0.649	**0.659**	**0.653**	**0.662**	**0.657**	**0.661**
	(3,3,3,3)	0.609	0.604	0.603	0.622	0.637	0.647	0.65	**0.658**	**0.656**	**0.664**	**0.66**	**0.666**

C. Computational Complexity

The complexity of each optimized TFD is measured in terms of processing time. The assessment is conducted on a workstation equipped with 2 Intel® Xeon® E5-2697V2 x64-based processors, 192 GB of memory, and MATLAB R2020b. The TFD computational complexity is estimated by Monte-Carlo simulations where the processing time of each optimized TFD is judged by generating the TFR of the 1000 test signals and repeating the process for 10 times. Fig. 8 demonstrates the averaged processing time of each optimized TFD. The results indicate that the MPD has the highest computational complexity followed by the DGF. In addition, it shows that the RGD and MDD share similar complexity and that the rest of TFDs are comparable in terms of processing time. From the TFD performance evaluations in Fig. 7, it becomes apparent that by far the greatest complexity goes to the TFDs that perform worst. These findings highlight a demand to take into account both the TFD accuracy and resolution when designing computationally expensive TFDs.

D. TFD Selection Strategies

The performance evaluation results presented in Fig. 7 assume no prior information on the signal parameters. Nevertheless, some TFDs are designed to perform best given a set of signal parameters. For this reason, we extend the former evaluations to introduce precise TFD selection strategies when some prior information on the signal parameters is available. Table I summarizes the averaged performance measure ξ for each TFD according to the number of signal components P and the IF polynomial orders R. Following the TFD selection criteria in Section II-D4, the set of best performing TFDs are identified by a gray shading. By analyzing the set consistency with respect to

the signal parameters, one can draw the following TFD selection strategies:

1) **CKD for non-stationary signals with one or two components**: the CKD kernel has the ability to reduce to an impulse when no filtering is needed, e.g. in case of mono-component LFM signals. Besides, it offers flexibility in terms of shape adaptation, directionality, and support compactness to suppress cross-terms with minimum effect on the auto-terms. Therefore, we recommend selecting the CKD when analyzing signals with one or two components regardless of their IF complexities.
2) **RGD for multi-component LFM signals**: the RGD and MDD kernels are directional filters that are suited for LFM signals. Nonetheless, the constant minor axis support of the MDD kernel leads to low resolution. Similarly, the upper limit on the RGD kernel parameter ψ prevents it from reducing to an impulse when analyzing mono-component LFM signals. Nevertheless, closer inspection of the results in Table I reveal that the RGD drawback becomes insignificant when $P > 1$. Consequently, we recommend selecting the RGD when analyzing multi-component LFM signals. It is worth noting that the MDD outperforms the RGD when we relax the restriction on its minor axis support. However, in this work, we confine our scope to evaluate the conventional MDD.
3) **CKD for non-stationary signals with three components and non-linear IF laws**: the CKD high performance extends to include signals with three components and non-linear IF laws. By analyzing the results in Table I for $P = 3$, we observe that the CKD yields the most consistent high performance when compared to other TFDs. As a result, we recommend selecting the CKD when analyzing signals with three components and non-linear IF laws.
4) **CKD, SPWVD, or EMBD for non-stationary signals with high number of components and non-linear IF laws**: the CKD, SPWVD, and EMBD kernels become equivalent, in terms of TFD accuracy and resolution, when analyzing signals with high number of components and IF complexities. Note that results in Fig. 7 identified the SPWVD as one of the best performing TFDs and disregarded the EMBD utility when no information on the signal parameters is available. However, it is apparent from Table I that the SPWVD and EMBD are equivalent and their utility is only justified when $P > 3$. Hence, we recommend selecting the CKD, SPWVD, or EMBD when analyzing signals with high number of components and non-linear IF laws.
5) **ED, BJD, PWVD, BD, MBD, DGF, and MPD are not recommended**: the ED and BJD kernels are non-directional filters; they smooth the WVD uniformly in all directions. Moreover, the PWVD, BD, and MBD kernels are mono-directional filters; they smooth the WVD along the time or frequency axis. For these reasons, none of the former kernels can maximize the TFD accuracy and resolution due to their limited directionality. On the other hand, the DGF is a time-dependent directional kernel that aims to maximize the TFD energy concentration. Nevertheless, its excessive filtering reduces resolution. Additionally, the MPD with chirplet dictionary selects many short-time atoms to match the IF non-linear regions; hence, it also reduces resolution. For these reasons, in addition to their computational complexities, the DGF and MPD are not recommended.

IV. Conclusions

The performance of a TFD is defined by its accuracy and resolution. High accuracy ensures minimal cross-terms, while high resolution enhances the ability to resolve the auto-terms. Conventionally, the TFD performance is assessed visually or quantified using TF measures. Nevertheless, visual assessment is subjective and TF measures cannot define the performance difference among different TFDs. This study presented a TFD performance evaluation process by designing a reference optimal TFD and deriving accuracy and resolution measures. The TFD design started by decomposing a standard non-stationary signal model using PW-LFM basis. Afterwards, we exploited the WVD optimality for LFM signals to formulate the proposed optimal PW-WVD. We adopted the proposed distribution to derive novel TFD accuracy and resolution measures and analyzed their sensitivity to form a TFD selection criteria. In addition, we compared the PW-WVD and the proposed measures to conventional approaches. Moreover, we evaluated the performance of 12 different TFDs using these measures and developed precise TFD selection strategies with or without prior information on the signal parameters. In summary, the study findings indicated that the CKD yields highest accuracy and resolution with minimum computational complexity given no prior information on the signal parameters. Additionally, we recommended the RGD for multi-component LFM signals.

The implications of the findings of this paper are as follows: (1) the designed optimal PW-WVD is a viable tool to evaluate and compare the performance of various TF signal processing techniques, e.g. source separation, IF estimation, TFD optimization, denoising, filtering, and many others; (2) the proposed TFD accuracy and resolution measures can be used as design criteria to formulate new high-performing TFDs; (3) TFD selection is no longer convoluted as one can follow the proposed strategies to maximize accuracy and resolution; and finally (4) machine learning techniques, e.g. neural networks, are suited to design new TFDs due to the availability of the proposed optimal TFD as true labels. The designed PW-WVD is limited to signals that can be decomposed using PW-LFMs. The TFR of naturally disperse functions, e.g. Gaussian signals, cannot be described using the proposed TFD due to their inherent ambiguity. Future research directions could involve (1) extending the PW-WVD by using spread directional basis for the signal decomposition, (2) finding an analytical solution of the Ambiguity function using the PW-WVD, (3) investigating the feasibility of different cost functions for TFD optimization using the proposed TFD accuracy and resolution measures, (4) forming a matching pursuit distribution with PW-LFM atoms, and (5) improving TFD design techniques by incorporating the proposed optimal TFD with machine learning tools.

The software package that is needed to reproduce the results of this paper along with the developed PW-WVD are provided as supplementary material and can downloaded from [49].

APPENDIX A
FORMULATING THE WVD INNER-TERMS USING THE BINOMIAL THEOREM

Using the signal model in Eq. (1), let us assume a mono-component ($P = 1$) non-stationary signal $z(t)$ with a polynomial IF of order R, i.e.:

$$f(t) = \sum_{r=0}^{R} \eta_r \, t^{\,r} \,, \tag{A.1}$$

where η_r is the IF r*th* polynomial coefficient. Using Eq. (9), the signal IAF is expressed as follows:

$$K_z(t,\tau) = K_a(t,\tau) \exp\left(j2\pi\,\varphi(t,\tau)\right) K_{\Pi}(t,\tau) \,, \tag{A.2}$$

$$\varphi(t,\tau) = \sum_{r=0}^{R} \left(\frac{\eta_r}{r+1}\right) \left[\left(t+\frac{\tau}{2}\right)^{r+1} - \left(t-\frac{\tau}{2}\right)^{r+1}\right] , \tag{A.3}$$

where $K_a(t,\tau) = a(t+\tau/2)a(t-\tau/2)$ and $K_{\Pi}(t,\tau) = \Pi(t+\tau/2)\Pi(t-\tau/2)$.

According to the binomial theorem, one can expand any non-negative power of $a+b$ into a sum of the form:

$$(a+b)^c = a^{\,c} \sum_{i=0}^{c} \binom{c}{i} \left(\frac{b}{a}\right)^{i} . \tag{A.4}$$

Using Eq. (A.4), one can express Eq. (A.3) as follows:

$$\left(t+\frac{\tau}{2}\right)^{r+1} = t^{\,r+1} \sum_{i=0}^{r+1} \binom{r+1}{i} \left(\frac{\tau}{2\,t}\right)^{i} . \tag{A.5}$$

$$\left(t-\frac{\tau}{2}\right)^{r+1} = t^{\,r+1} \sum_{i=0}^{r+1} \binom{r+1}{i} \left(\frac{-\tau}{2\,t}\right)^{i} . \tag{A.6}$$

$$\varphi(t,\tau) = \sum_{r=0}^{R} \left(\frac{\eta_r \, t^{\,r+1}}{r+1}\right) \underbrace{\sum_{i=0}^{r+1} \binom{r+1}{i} \left(\frac{\tau}{2\,t}\right)^{i} \kappa_i}_{\mathfrak{B}(t,\tau)} . \tag{A.7}$$

$$\kappa_i = \begin{cases} 0 & : i \text{ is even} \\ 2 & : i \text{ is odd} \end{cases} . \tag{A.8}$$

Let us expand the binomial sum $\mathfrak{B}(t,\tau)$ in Eq. (A.7) as:

$$\mathfrak{B}(t,\tau) = (r+1)\left(\frac{\tau}{t}\right) + \sum_{i=2}^{r+1} \binom{r+1}{i} \left(\frac{\tau}{2\,t}\right)^{i} \kappa_i \,. \tag{A.9}$$

Substituting this expansion into Eq. (A.7) yields:

$$\varphi(t,\tau) = \sum_{r=0}^{R} \left[\eta_r \, t^{\,r}\tau + \left(\frac{\eta_r \, t^{\,r+1}}{r+1}\right) \sum_{i=2}^{r+1} \binom{r+1}{i} \left(\frac{\tau}{2\,t}\right)^{i} \kappa_i\right]$$

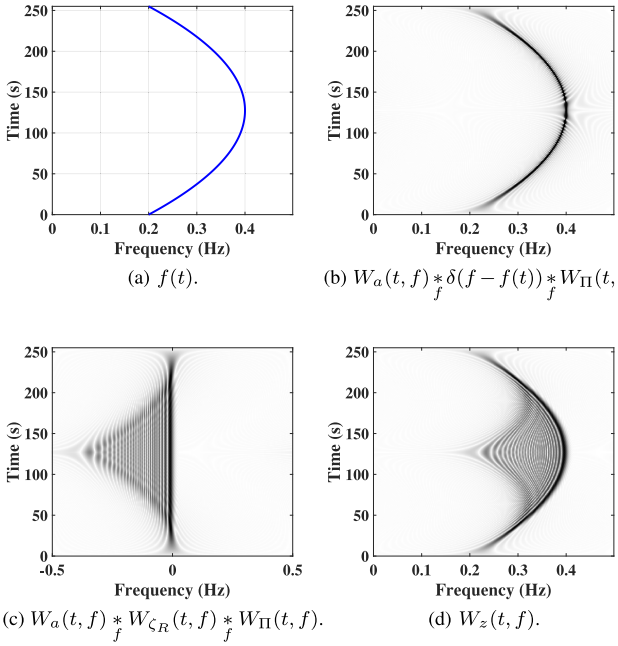


Fig. 9. The WVD auto-terms and inner-terms for an example mono-component non-stationary signal. The signal is characterized by a unit constant IA, a 2*nd* order polynomial IF that passes through the TF points $\{(0,0.2),(127,0.4),(255,0.2)\}$, and a time support between 0 and 255 seconds. (a) shows the signal IF, (b) illustrates the WVD auto-terms, (c) presents the WVD inner-terms, and (d) shows the WVD of the signal computed by Eq. (A.13). The simulation parameters are: $N = 256$ samples, $f_s = 1$ Hz and $M = 256$ samples.

$$= \tau \underbrace{\sum_{r=0}^{R} \eta_r \, t^{\,r}}_{f(t)} + \underbrace{\sum_{r=2}^{R} \sum_{i=2}^{r+1} \binom{r+1}{i} \left(\frac{\eta_r \, t^{\,r+1}}{r+1}\right) \left(\frac{\tau}{2\,t}\right)^{i} \kappa_i}_{\zeta_R(t,\tau)}$$

$$= \tau f(t) + \zeta_R(t,\tau), \tag{A.10}$$

where $f(t)$ is the signal IF and $\zeta_R(t,\tau)$ is a recursive residual term coming from the bilinear nature of the IAF operator, i.e.:

$$\zeta_R(t,\tau) = \zeta_{R-1}(t,\tau) + \sum_{i=2}^{R+1} \binom{R+1}{i} \left(\frac{\eta_R \, t^{\,R+1}}{R+1}\right) \left(\frac{\tau}{2\,t}\right)^{i} \kappa_i \,. \tag{A.11}$$

Using the expression in Eq. (A.10), Eq. (A.2) becomes:

$$K_z(t,\tau) = K_a(t,\tau)\, e^{j2\pi\tau f(t)} e^{j2\pi\zeta_R(t,\tau)} K_{\Pi}(t,\tau) \,. \tag{A.12}$$

Finally, applying the WVD definition in Eq. (8) yields:

$$W_z(t,f) = W_a(t,f) \underset{f}{*} \delta\left(f-f(t)\right) \underset{f}{*} W_{\zeta_R}(t,f) \underset{f}{*} W_{\Pi}(t,f) \,, \tag{A.13}$$

where $\delta(f-f(t))$ holds the WVD auto-terms, $W_{\zeta_R}(t,f)$ holds the WVD inner-terms, $W_a(t,f) = \underset{\tau\to f}{\mathcal{F}}\{K_a(t,\tau)\}$, $W_{\Pi}(t,f) = \underset{\tau\to f}{\mathcal{F}}\{K_{\Pi}(t,\tau)\}$, and $\underset{f}{*}$ denotes the convolution operation in frequency.

By examining Eq. (A.13), one observes that the WVD is optimal if and only if $W_{\zeta_R}(t,f) = \delta(f)$ which corresponds to $\zeta_R(t,\tau) = 0$. Using Eq. (A.11), we see that this is only

possible when $R = 0$ or $R = 1$; $z(t)$ is an LFM signal. Hence, the WVD shows no inner-terms for LFM signals. Additionally, the recursive non-linear definition of $\zeta_R(t,\tau)$ indicates that the complexity of $W_{\zeta_R}(t,f)$ escalates for higher IF polynomial orders. Fig. 9 demonstrates a numerical verification for Eq. (A.13) by presenting the WVD auto-terms and inner-terms for a non-stationary signal with quadratic IF law.

Appendix B
The PW-WVD Optimality and Properties

The PW-WVD definition in Eq. (15) is optimal if and only if $\sum_{q=1}^{\infty} W_{\psi_{p,q}}(t,f) = \delta(f - f_p(t))$ holds $\forall p \in [1,2,\ldots,P]$. In the following we prove this statement.

Let us express the q*th* PW-LFM frequency rate $\lambda_{p,q}$, starting frequency $\varepsilon_{p,q}$, and time support when $T_p \to 0$:

$$\lambda_{p,q} = \lim_{T_p \to 0} \left(\frac{f_p\left(t_{p,q} + T_p\right) - f_p\left(t_{p,q}\right)}{T_p} \right) = f_p'\left(t_{p,q}\right) ,$$

$$\varepsilon_{p,q} = \lim_{T_p \to 0} \left(f_p\left(t_{p,q}\right) - \lambda_{p,q}\, t_{p,q} \right) = f_p(t_{p,q}) - f_p'(t_{p,q})\, t_{p,q} ,$$

$$\lim_{T_p \to 0} \left(t_{p,q} \le t < t_{p,q} + T_p \right) \to t = t_{p,q} ,$$

where $f_p'(t_{p,q}) = \dfrac{\mathrm{d}f_p(t)}{\mathrm{d}t}|_{t=t_{p,q}}$. Using these expressions, the definition of $W_{\psi_{p,q}}(t,f)$ in Eq. (16) becomes:

$$\begin{aligned} & W_{\psi_{p,q}}(t,f) \\ &= \begin{cases} \delta\left(f - f_p'(t_{p,q})\, t - f_p(t_{p,q}) + f_p'(t_{p,q})\, t_{p,q}\right) & : t = t_{p,q} \\ 0 & : \text{otherwise} \end{cases} \\ &= \delta\left(f - \cancel{f_p'(t_{p,q})\, t_{p,q}} - f_p(t_{p,q}) + \cancel{f_p'(t_{p,q})\, t_{p,q}}\right) \\ &= \delta\left(f - f_p(t_{p,q})\right) . \end{aligned}$$

This allows us to express the infinite sum in Eq. (15) as:

$$\sum_{q=1}^{\infty} W_{\psi_{p,q}}(t,f) = \delta\left(f - f_p(t_{p,q})\right) \ \forall t_{p,q} \in [t_{p,i}, t_{p,f}] .$$

Since $f_p(t)$ is only defined within its temporal support $[t_{p,i}, t_{p,f}]$, $f_p(t_{p,q})\ \forall t_{p,q} \in [t_{p,i}, t_{p,f}] \equiv f_p(t)$; hence:

$$\sum_{q=1}^{\infty} W_{\psi_{p,q}}(t,f) = \delta\left(f - f_p(t)\right) .$$

This proves the PW-WVD optimality for each signal component p separately. Nevertheless, since the multi-component PW-WVD is constructed through a linear sum, it is also optimal by linearity, i.e.:

$$\sum_{q=1}^{\infty} W_{\psi_{p,q}}(t,f) = \delta\left(f - f_p(t)\right) \ \forall p \in [1,2,\ldots,P] .$$

Apart from its optimality, the PW-WVD inherent the following set of WVD properties: realness, time covariance, frequency covariance, (time marginal)*, (frequency marginal)*, (global energy)*, (instantaneous frequency)*, (spectral delay)*, time support, frequency support, convolution invariance, modulation invariance, (invertibility)*, (Moyal's formula)*, and Ambiguity function. Note that properties marked with * are only valid for mono-component signals, because the PW-WVD is cross-terms free. The supplementary material includes proofs for these properties.

Appendix C
The Discrete PW-WVD

Let $z[n]$ be a discrete multi-component non-stationary signal sampled at f_s such that $z[n] \triangleq z((n-1)/f_s)$ where $n \in [1,N]$ is the time sample index and N is the total number of time samples in $z[n]$. The discrete PW-LFM decomposed model is expressed as follows:

$$z[n] = \sum_{p=1}^{P} a_p[n] \left(\sum_{q=1}^{Q_p} \psi_{p,q}[n] \right) \Pi_p[n] , \tag{C.1}$$

$$\Pi_p[n] = \begin{cases} 1 & : n \in \left[n_{p,i}, n_{p,f}\right] \\ 0 & : \text{otherwise} \end{cases} , \tag{C.2}$$

$$\psi_{p,q}[n] = \begin{cases} e^{\,j2\pi \left(\frac{\lambda_{p,q}}{2}\, n^2 + \varepsilon_{p,q}\, n + c \right)} & : n \in \left[n_{p,q-1}, n_{p,q}\right] \\ 0 & : \text{otherwise} \end{cases} , \tag{C.3}$$

where $a_p[n]$, $f_p[n]$ and $\Pi_p[n]$ are the p*th* component discrete IA, IF and sample support window, respectively, $\psi_{p,q}[n]$ is the q*th* discrete PW-LFM for $f_p[n]$, $Q_p = n_{p,f} - n_{p,i}$, $n_{p,i} \ge 1$, $n_{p,f} \le N$, $n_{p,q} = n_{p,i} + q$, $\lambda_{p,q} = f_p[n_{p,q}] - f_p[n_{p,q-1}]$, $\varepsilon_{p,q} = f_p[n_{p,q-1}] - \lambda_{p,q} n_{p,q}$, and c is an integration constant that ensures continuity.

The discrete PW-WVD is defined as follows:

$$L_z[n,m] = \sum_{p=1}^{P} W_{a_p}[n,m] \underset{m}{*} \sum_{q=1}^{Q_p} W_{\psi_{p,q}}[n,m] \underset{m}{*} W_{\Pi_p}[n,m] , \tag{C.4}$$

where $m \in [1,M]$ is the frequency sample index, M is the total number of frequency samples, $L_z[n,m]$ is the discrete PW-WVD of $z[n]$, $W_{a_p}[n,m]$ and $W_{\Pi_p}[n,m]$ are the WVDs of $a_p[n]$ and $\Pi_p[n]$, respectively, and $W_{\psi_{p,q}}[n,m]$ is the WVD of the q*th* PW-LFM $\psi_{p,q}[n]$, i.e.:

$$W_{\psi_{p,q}}[n,m] = \begin{cases} \delta\left(m - \lambda_{p,q}\, n - \varepsilon_{p,q}\right) & : n \in [n_{p,q-1}, n_{p,q}] \\ 0 & : \text{otherwise} \end{cases} . \tag{C.5}$$

Appendix D
Formulation for Selected Time-Frequency Distributions

This study evaluates the accuracy and resolution of the following TFDs: Matching Pursuit distribution (MPD), Born-Jordan distribution (BJD), exponential distribution (ED), pseudo WVD (PWVD), smoothed-pseudo WVD (SPWVD), B-distribution (BD), modified BD (MBD), extended MBD (EMBD), compact kernel distribution (CKD), multi-directional distribution

TABLE II
TFD KERNEL FORMULATIONS AND CONTROL PARAMETERS [1]. L_w AND L_h ARE THE LENGTHS OF THE WINDOWS $w(\tau)$ AND $h(\nu)$ RESPECTIVELY, $\Omega = \{$RECTANGULAR, HANNING, HAMMING, BARTLETT$\}$ IS THE SET OF WINDOW FUNCTIONS, AND N IS THE TOTAL NUMBER OF TIME SAMPLES

TFD	$g(\nu,\tau)$	**Parameters**
BJD	$\operatorname{sinc}(2\nu\tau\alpha)$	$\alpha \in [0,1]$
ED	$\exp\left(\frac{-\nu^2\tau^2}{2\pi\sigma}\right)$	$\sigma \in [0,1]$
PWVD	$w(\tau)$	$w(\tau) \in \Omega$, $L_w \in [1,3,\cdots,N]$
SPWVD	$w(\tau)\,h(\nu)$	$w(\tau) \in \Omega$, $L_w \in [1,3,\cdots,N]$ $h(t) \in \Omega$, $L_h \in [1,3,\cdots,N]$
BD	$\frac{\lvert\tau\rvert^{\beta}\lvert\Gamma(\beta+j\pi\nu)\rvert^2}{2^{1-2\beta}\,\Gamma(2\beta)}$	$\beta \in [0,1]$
MBD	$\frac{\lvert\Gamma(\beta+j\pi\nu)\rvert^2}{\Gamma^2(\beta)}$	$\beta \in [0,1]$
EMBD	$\frac{\lvert\Gamma(\beta+j\pi\nu)\rvert^2}{\Gamma^2(\beta)}\,\frac{\lvert\Gamma(\alpha+j\pi\tau)\rvert^2}{\Gamma^2(\alpha)}$	$\alpha \in [0,1]$, $\beta \in [0,1]$
CKD	$\exp\left(2c+\frac{c\,D^2}{\nu^2-D^2}+\frac{c\,E^2}{\tau^2-E^2}\right) : \lvert\nu\rvert < D\,,\ \lvert\tau\rvert < E$	$c \in [0,3]$, $D \in [0,1]$, $E \in [0,1]$
MDD	$\frac{\exp(2c)}{\#\boldsymbol{\theta}}\sum_{i=1}^{\#\boldsymbol{\theta}}\exp\left(\frac{c\,D^2}{x_i^2-D^2}+\frac{c\,E_i^2}{y_i^2-E_i^2}\right) : \lvert x_i\rvert < D,\ \lvert y_i\rvert < E_i$	$x_i = \nu\cos(\theta_i) - \tau\sin(\theta_i)$ $y_i = \nu\sin(\theta_i) + \tau\cos(\theta_i)$ $c \in [0,3]$, $D = 0.01$, $\eta \in [0,1]$
RGD	$\max\iint\left\lvert\exp\left(-\frac{\nu^2+\tau^2}{2\sigma^2(\boldsymbol{\theta})}\right)A_z(\nu,\tau)\right\rvert^2 \mathrm{d}\nu\,\mathrm{d}\tau : \frac{1}{2\pi}\int_0^{\pi}\sigma^2(\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{\theta} \le \psi$	$\psi \in [1,20]$
DGF	$\underset{\tau\leftarrow f}{\mathcal{F}}{}^{-1}\left\{\underset{t\to\nu}{\mathcal{F}}\left\{\frac{\mathrm{d}^2}{\mathrm{d}f_{\boldsymbol{\theta}_t}^2}\left\{e^{-a^2\left(t\cos(\boldsymbol{\theta}_t)+f\sin(\boldsymbol{\theta}_t)\right)^2-b^2\left(f\cos(\boldsymbol{\theta}_t)-t\sin(\boldsymbol{\theta}_t)\right)^2}\right\}\right\}\right\}$	$a \in [0,15]$, $b \in [0,15]$

(MDD), radial Gaussian distribution (RGD), and directional Gaussian filter (DGF).

A. The Matching Pursuit Distribution

The MP approach decomposes an input signal $z(t)$ via a dictionary that contains a set of elementary TF atoms [38]. It projects $z(t)$ over the TF atoms and then selects those that match its local structure in an adaptive greedy fashion [39]. The MP decomposition is expressed as follows [38]:

$$z(t) = \sum_{k=0}^{K-1} c_k\,\mathcal{G}_{\Lambda_k}(t) + \varepsilon_K z(t)\,, \tag{D.1}$$

where $\mathcal{G}_{\Lambda_k}(t)$ is the TF atom that belongs to the dictionary, Λ_k is the set of atom parameters, $c_k = \int \varepsilon_k z(t)\,\mathcal{G}^*_{\Lambda_k}(t)\,\mathrm{d}t$ is the atom coefficient, K is the total number of decompositions (or atoms), and $\varepsilon_K z(t)$ is the MP residual after K decompositions where $\varepsilon_0 z(t) = z(t)$.

The MPD is computed by applying the WVD definition in Eq. (8) to the selected atoms, i.e. [39]:

$$M_z(t,f) = \sum_{k=0}^{K-1} \lvert c_k\rvert^2\,W_{\mathcal{G}_{\Lambda_k}}(t,f)\,, \tag{D.2}$$

where $M_z(t,f)$ is the MPD and $W_{\mathcal{G}_{\Lambda_k}}(t,f)$ is the WVD of $\mathcal{G}_{\Lambda_k}(t)$. In this work, we populate the MP dictionary with chirplet atoms which are expressed as [50], [51]:

$$\mathcal{G}_{\Lambda_k}(t) = \frac{1}{\sqrt{s_k}}\,G\left(\frac{t-\mu_k}{s_k}\right)e^{\left(j\omega_k^s(t-\mu_k)+j\omega_k^r(t-\mu_k)^2\right)}\,, \tag{D.3}$$

where $G(t) = 2^{1/4}\,e^{-\pi t^2}$ is a Gaussian function, $s_k \in [2^{10}, 2^9, \ldots, 2^0]$ controls the temporal width of the atom, $\mu_k \in [0, 1, \ldots, 127] \times T/127$ is the time shift, T is the input signal duration, $\omega_k^s \in [0, 1, \ldots, 63] \times \pi f_s/63$ is the chirp start frequency, f_s is the signal sampling frequency, and $\omega_k^r \in [0 \times 2 - 63, 1 \times 2 - 63, \ldots, 63 \times 2 - 63] \times \pi f_s/3150$ is the chirp rate. Note that Gaussian atoms are special case of this expression when $\omega_k^r = 0$.

Using the signal parameters summarized in Section III-A, the MP dictionary holds approximately 6 million atoms. Moreover, the signal decomposition in Eq. (D.1) is performed using the Orthogonal Matching Pursuit algorithm in MATLAB [52]. The decomposition is terminated when reaching a 1% minimum L_2 relative error or when $K = T$.

B. Quadratic TFDs

The TFD definition in Eq. (19) can be expressed by means of multiplication in an alternative domain, called ambiguity domain or Doppler-lag domain, as follows [1]:

$$\rho_z(t,f) = \underset{t\leftarrow\nu}{\mathcal{F}}{}^{-1}\left\{\underset{\tau\to f}{\mathcal{F}}\left\{g(\nu,\tau)\,A_z(\nu,\tau)\right\}\right\}, \tag{D.4}$$

where $\underset{t\leftarrow\nu}{\mathcal{F}}{}^{-1}$ denotes the inverse Fourier transform from Doppler to time, $g(\nu,\tau)$ is the Doppler-lag kernel, and $A_z(\nu,\tau)$ is the ambiguity function. Table II lists the Doppler-lag kernel formulations for the quadratic TFDs used in this study alongside their control parameters.

ACKNOWLEDGMENT

The authors would like to thank Dr. Isabella Reinhold and Prof. Maria Sandsten for sharing the implementation of their work.

REFERENCES

[1] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing*, *2nd ed*. Oxford, U.K.: Academic Press, 2016.

[2] Y. Li, W. Cui, M. Luo, K. Li, and L. Wang, "Epileptic seizure detection based on time-frequency images of EEG signals using gaussian mixture model and gray level co-occurrence matrix features," *Int. J. Neural Syst.*, vol. 28, no. 07, 2018, Art no. 1850003.

[3] M. Sharma, D. Goyal, P. V. Achuth, and U. R. Acharya, "An accurate sleep stages classification system using a new class of optimally time-frequency localized three-band wavelet filter bank," *Comput. Biol. Med.*, vol. 98, pp. 58–75, 2018.

[4] B. Boashash and S. Ouelha, "Designing high-resolution time-frequency and time-scale distributions for the analysis and classification of non-stationary signals: A tutorial review with a comparison of features performance," *Digit. Signal Process.*, vol. 77, pp. 120–152, 2018.

[5] Z.-K. Gao, Q. Cai, Y.-X. Yang, N. Dong, and S.-S. Zhang, "Visibility graph from adaptive optimal kernel time-frequency representation for classification of epileptiform EEG," *Int. J. Neural Syst.*, vol. 27, no. 04, 2017, Art no. 1750005.

[6] S. Ouelha, A. Aïssa-El-Bey, and B. Boashash, "Improving DOA estimation algorithms using high-resolution quadratic time-frequency distributions," *IEEE Trans. Signal Process.*, vol. 65, no. 19, pp. 5179–5190, Oct. 2017.

[7] K. Cui, X. Chen, J. Huang, Q. Wang, and N. Yuan, "DOA estimation of LFM sources based on time-frequency interferometer in UCA," *AEU - Int. J. Electron. Commun.*, vol. 104, pp. 76–85, 2019.

[8] Y. Abdoush, G. Pojani, and G. E. Corazza, "Adaptive instantaneous frequency estimation of multicomponent signals based on linear time-frequency transforms," *IEEE Trans. Signal Process.*, vol. 67, no. 12, pp. 3100–3112, Jun. 2019.

[9] M. Brajović, V. Popović-Bugarin, I. Djurović, and S. Djukanović, "Post-processing of time-frequency representations in instantaneous frequency estimation based on ant colony optimization," *Signal Process.*, vol. 138, pp. 195–210, 2017.

[10] V. S. Amin, Y. D. Zhang, and B. Himed, "Improved instantaneous frequency estimation of multi-component FM signals," in *Proc. IEEE Radar Conf.*, Apr. 2019, pp. 1–6.

[11] R. Anderson and M. Sandsten, "Stochastic modelling and optimal spectral estimation of EEG signals," in *Proc. EMBEC & NBC 2017*, H. Eskola, O. Väisänen, J. Viik, and J. Hyttinen, Eds. Singapore: Springer, 2018, pp. 908–911.

[12] M. F. Al-Sa'd and B. Boashash, "Design and implementation of a multi-sensor newborn EEG seizure and background model with inter-channel field characterization," *Digit. Signal Process.*, vol. 90, pp. 71–99, 2019.

[13] D. Fourer, F. Auger, and G. Peeters, "Local AM/FM parameters estimation: Application to sinusoidal modeling and blind audio source separation," *IEEE Signal Process. Lett.*, vol. 25, no. 10, pp. 1600–1604, Oct. 2018.

[14] P. Wang, P. V. Orlik, K. Sadamoto, W. Tsujita, and F. Gini, "Parameter estimation of hybrid sinusoidal FM-Polynomial phase signal," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 66–70, Jan. 2017.

[15] A. Serbes, "On the estimation of LFM signal parameters: Analytical formulation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 2, pp. 848–860, Apr. 2018.

[16] B. Boashash and S. Ouelha, "Automatic signal abnormality detection using time-frequency features and machine learning: A newborn EEG seizure case study," *Knowl.-Based Syst.*, vol. 106, pp. 38–50, 2016.

[17] R. B. Pachori and A. Nishad, "Cross-terms reduction in the wigner-ville distribution using tunable-q wavelet transform," *Signal Process.*, vol. 120, pp. 288–304, 2016.

[18] R. Alazrai, R. Homoud, H. Alwanni, and M. I. Daoud, "EEG-Based emotion recognition using quadratic time-frequency distribution," *Sensors*, vol. 18, no. 8, 2018, Art no. 2739.

[19] J. Escrivá Muñoz, Y. Pan, S. Ge, E. W. Jensen, and M. Vallverdú, "Novel characterization method of impedance cardiography signals using time-frequency distributions," *Med. Biol. Eng. Comput.*, vol. 56, no. 10, pp. 1757–1770, Oct. 2018.

[20] S. Shao, A. Liu, C. Yu, H. Yang, Y. Li, and B. Li, "Spatial time-frequency distribution of cross term-based direction-of-arrival estimation for weak non-stationary signal," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–12, Oct. 2019.

[21] B. Jokanović and M. Amin, "Fall detection using deep learning in range-doppler radars," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 1, pp. 180–189, Feb. 2018.

[22] P. Wang, E. Cetin, A. G. Dempster, Y. Wang, and S. Wu, "Time frequency and statistical inference based interference detection technique for GNSS receivers," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 6, pp. 2865–2876, Dec. 2017.

[23] P. Wang, E. Cetin, A. G. Dempster, Y. Wang, and S. Wu, "Improved characterization of GNSS jammers using short-term time-frequency Rényi entropy," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1918–1930, Aug. 2018.

[24] M. Abed and A. Belouchrani, "Performance analysis and computational cost evaluation of high-resolution time-frequency distributions derived from compact support time-lag kernels," *Digit. Signal Process.*, vol. 78, pp. 1–19, 2018.

[25] M. A. Awal and B. Boashash, "An automatic fast optimization of quadratic time-frequency distribution using the hybrid genetic algorithm," *Signal Process.*, vol. 131, pp. 134–142, 2017.

[26] N. Saulig, Ž. Milanović, and C. Ioana, "A local entropy-based algorithm for information content extraction from time-frequency distributions of noisy signals," *Digit. Signal Process.*, vol. 70, pp. 155–165, 2017.

[27] A. F. Hussein, S. J. Hashim, A. F. A. Aziz, F. Z. Rokhani, and W. A. W. Adnan, "Performance evaluation of time-frequency distributions for ECG signal analysis," *J. Med. Syst.*, vol. 42, no. 1, pp. 1–16, Nov. 2017.

[28] M. A. Awal, S. Ouelha, S. Dong, and B. Boashash, "A robust high-resolution time-frequency representation based on the local optimization of the short-time fractional fourier transform," *Digit. Signal Process.*, vol. 70, pp. 125–144, 2017.

[29] N. Saulig, I. Orović, and V. Sucic, "Optimization of quadratic time-frequency distributions using the local rényi entropy information," *Signal Process.*, vol. 129, pp. 17–24, 2016.

[30] B. Boashash, H. Barki, and S. Ouelha, "Performance evaluation of time-frequency image feature sets for improved classification and analysis of non-stationary signals: Application to newborn EEG seizure detection," *Knowl.-Based Syst.*, vol. 132, pp. 188–203, 2017.

[31] Z. Zhang, C. Wang, C. Gan, S. Sun, and M. Wang, "Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 3, pp. 469–478, Sep. 2019.

[32] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1253–1263, May 2003.

[33] B. Boashash and S. Ouelha, "An improved design of high-resolution quadratic time-frequency distributions for the analysis of nonstationary multicomponent signals using directional compact kernels," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2701–2713, May 2017.

[34] I. Reinhold and M. Sandsten, "Optimal time-frequency distributions using a novel signal adaptive method for automatic component detection," *Signal Process.*, vol. 133, pp. 250–259, 2017.

[35] L. Stanković, I. Djurović, S. Stanković, M. Simeunović, S. Djukanović, and M. Daković, "Instantaneous frequency in time-frequency analysis: Enhanced concepts and performance of estimation algorithms," *Digit. Signal Process.*, vol. 35, pp. 1–13, 2014.

[36] S. Pikula and P. Beneš, "Comparison of measures of time-frequency distribution optimization," in *Proc. 8th Int. Congr. Ultra Modern Telecommun. Control Syst. Workshops*, Oct. 2016, pp. 314–319.

[37] Y. Yang, Z. Peng, W. Zhang, and G. Meng, "Parameterised time-frequency analysis methods and their engineering applications: A review of recent advances," *Mech. Syst. Signal Process.*, vol. 119, pp. 182–221, 2019.

[38] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[39] S. Ghofrani and D. C. McLernnon, "Comparison of the cross deleted wigner representation and the matching pursuit distribution (via adaptive signal decomposition)," in *Image and Signal Processing*, A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, Eds. Berlin, Heidelberg: Springer, 2008, pp. 516–526.

[40] F. Jaillet and B. Torrésani, "Time-frequency jigsaw puzzle: Adaptive multiwindow and multilayered gabor expansions," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 5, no. 2, pp. 293–315, 2007.

[41] G. M. Davis, S. G. Mallat, and Z. Zhang, "Adaptive time-frequency decompositions," *Opt. Eng.*, vol. 33, no. 7, pp. 2183–2191, 1994.

[42] G. Davis, S. Mallat, and Z. Zhang, "Adaptive time-frequency approximations with matching pursuits," in *Wavelets: Theory, Algorithms, and Applications, ser. Wavelet Analysis and Its Applications*, C. K. Chui, L. Montefusco, and L. Puccio, Eds. San Diego, CA, USA: Academic Press, 1994, vol. 5, pp. 271–293.

[43] G. Chen, J. Chen, and G. M. Dong, "Chirplet Wigner-Ville distribution for time-frequency representation and its application," *Mech. Syst. Signal Process.*, vol. 4 1, no. 1, pp. 1–13, 2013.
[44] X. Shen, Q.-Q. Li, G. Wu, and J. Zhu, "Decomposition of LiDAR waveforms by B-spline-based modeling," *ISPRS J. Photogrammetry Remote Sens.*, vol. 128, pp. 182–191, 2017.
[45] B. Boashash, B. K. Jawad, and S. Ouelha, "Refining the ambiguity domain characteristics of non-stationary signals for improved time-frequency analysis: Test case of multidirectional and multicomponent piecewise LFM and HFM signals," *Digit. Signal Process.*, vol. 83, pp. 367–382, 2018.
[46] K. Höllig and J. Hörner, *Approximation and Modeling With B-Splines*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 2013.
[47] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.
[48] A. D. Bull, "Convergence rates of efficient global optimization algorithms," *J. Mach. Learn. Res.*, vol. 12, no. 88, pp. 2879–2904, 2011.
[49] M. Al-Sa'd, B. Boashash, and M. Gabbouj, "The piece-wise spline Wigner-Ville distribution (PW-WVD) MATLAB package V1.0," *GitHub*, 2021. [Online]. Available: https://github.com/Al-Sad/PW-WVD
[50] S. Ghofrani, D. C. McLernon, and A. Ayatollahi, "Conditional spectral moments in matching pursuit based on the chirplet elementary function," *Digit. Signal Process.*, vol. 18, no. 5, pp. 694–708, 2008.
[51] S. Ghofrani, D. C. McLernon, and A. Ayatollahi, "Comparing Gaussian and chirplet dictionaries for time-frequency analysis using matching pursuit decomposition," in *Proc. 3rd IEEE Int. Symp. Signal Process. Inf. Technol.*, Dec. 2003, pp. 713–716.
[52] Y. Lv, J. Luo, and C. Yi, "Enhanced orthogonal matching pursuit algorithm and its application in mechanical equipment fault diagnosis," *Shock Vib.*, vol. 2017, 2017, pp. 1–13.

Mohammad Al-Sa'd (Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Qatar University, Doha, Qatar, in 2012 and 2016, respectively. He specialized in signal processing and graduated with honors under Professor Boualem Boashash's supervision. He was a Research Assistant with Qatar University, and he is currently a Doctoral Researcher with the Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland. His research interests include time-frequency signal theory, machine learning, EEG analysis and processing, information flow and theory, signal modeling, and optimization. He is a technical reviewer for several journals, including the IEEE TRANSACTIONS ON SIGNAL PROCESSING, *Digital Signal Processing*, *Signal Processing*, *Biomedical Signal Processing and Control*, and IEEE ACCESS.

Boualem Boashash (Fellow, IEEE) is a Scholar, Professor, and Senior Academic with experience in five leading Universities in France and Australia. He is currently an Emeritus Professor with The University of Queensland, Brisbane, QLD, Australia. He has authored or coauthored more than 500 technical publications, including more than 150 journal publications covering engineering, applied mathematics, and biomedicine. He pioneered the field of time-frequency signal processing for which he authored or coauthored the most comprehensive book and most powerful software package. Among many initiatives, he founded ISSPA, a leading conference since 1985. On leave from The University of Queensland, he took an appointment in the UAE, as the Dean of engineering, then moved to Qatar University, Doha, Qatar, as an Associate Dean, then a Research Professor until 2017 when he returned to The University of Queensland. At the time of publication, his work had been cited about 20000 times. His current interests include time-frequency signal theory, machine learning, DSP for EEG abnormalities detection, newborn brain networks analysis and the relevance of spirituality and ethics in advancing scientific research for the harmonious development of Society.

Moncef Gabbouj (Fellow, IEEE) received the B.S. degree in electrical engineering from Oklahoma State University, Stillwater, OK, USA, in 1985, and the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1986 and 1989, respectively. He is currently a Professor of signal processing with the Department of Computing Sciences, Tampere University, Tampere, Finland. From 2011 to 2015, he was an Academy of Finland Professor. His research interests include big data analytics, multimedia content-based analysis, indexing and retrieval, artificial intelligence, machine learning, pattern recognition, nonlinear signal and image processing and analysis, voice conversion, and video processing and coding. He is a Member of the Academia Europaea and the Finnish Academy of Science and Letters. He is the past Chairman of the IEEE CAS TC on DSP and the Committee Member of the IEEE Fourier Award for Signal Processing. He was an Associate Editor and the Guest Editor of many IEEE, and international journals and a Distinguished Lecturer for the IEEE CASS. He is the Finland Site Director of the NSF IUCRC funded Center for Visual and Decision Informatics (CVDI) and leads the Artificial Intelligence Research Task Force of the Ministry of Economic Affairs and Employment funded Research Alliance on Autonomous Systems (RAAS).