The T-class of time-frequency distributions: time-only kernels with amplitude estimation

Hussain, Zahir; Boashash, Boualem

The T-Class of Time - Frequency Distributions: Time-Only Kernels with Amplitude Estimation

Zahir M. Hussain*, Boualem Boashash°

*School of Electrical and Computer Engineering, RMIT University, 124 Latrobe Street, Melbourne, Victoria 3000, Australia
° School of Engineering Systems, Queensland University of Technology
2 George Street, Brisbane, Queensland 4000, Australia.

Abstract

This paper presents a class of time - frequency distributions (TFDs) characterized by time-lag kernels which are functions of time only. If the parameters of the time-only kernels are properly chosen, their corresponding TFDs, the T-distributions, are more efficient than their two-dimensional counterparts in terms of cross-terms suppression while keeping a high energy concentration (resolution) around the IF law of non-stationary signals. The proposed class is a subclass of Cohen's Class of quadratic TFDs. We have shown that separable time-lag kernels should be lag-independent (or: time-only) for best resolution. In addition, non-parametric amplitude estimation is possible directly from the T-distributions in case of FM signals, a property that is not verified by other TFDs. Two examples of the T-distributions are given and their performance is compared to other TFDs with numerical examples using synthetic and real-life signals.

Key words: Time - frequency analysis, non-parametric amplitude estimation, multicomponent signals, resolution, cross-terms, Gaussian noise.
PACS: 07.50.Qx

1. Introduction

Time-frequency analysis is used to deal with nonstationary signals as it reveals the time-varying characteristics of their spectra and the multicomponent nature of such signals, while the classical Fourier techniques are insufficient in this respect. Joint time - frequency representations (a.k.a time - frequency

---

Email address: zahir.hussain@rmit.edu.au, b.boashash@qut.edu.au (Zahir M. Hussain*, Boualem Boashash°).

URL: http://www.rmit.edu.au, http://www.qut.edu.au (Zahir M. Hussain*, Boualem Boashash°).

distributions, TFDs) are transformations that describe the energy density of the signal simultaneously in time and frequency. Most important TFDs are members of the quadratic (bilinear) class [1]-[4]. The early examples of the quadratic class include the spectrogram and the Wigner-Ville distribution (WVD). This class has originated from the work of Wigner in 1932 [5], formulated with the analytic signal by Ville in 1948 [6], then generalized by the work of Cohen in quantum physics in 1966 [7], hence it is also known as Cohen's Class. However, quadratic TFDs of multicomponent signals suffer from the presence of cross-terms [1]-[11], which can obscure the real features of interest in the signal. The time - frequency resolution (energy concentration) is an important issue for both mono- and multicomponent signals, especially for non-parametric IF estimation [1]. For instance, the Wigner-Ville distribution (WVD) gives optimal resolution for linear monocomponent FM signals but it is suboptimal for non-linear FM [1,8], and it suffers from significant cross-terms for multicomponent signals.

Considerable efforts have been made to define TFDs which reduce the effect of cross-terms while improving the time - frequency resolution (e.g. [9]-[13]). This led to the so-called reduced interference distributions (RIDs), the L-class and signal - dependent optimal time - frequency representations. The general formula for Cohen's Class of quadratic TFDs for a time signal $s(t)$ is given by [1,2]

$$\rho(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\nu, \tau) A(\nu, \tau) e^{-j2\pi(\nu t + \tau f)} d\nu d\tau \qquad (1)$$

where (ν, τ) is the Doppler-lag or the ambiguity domain, $g(\nu, \tau)$ is the Doppler-lag kernel of the distribution $\rho(t, f)$ and $A(\nu, \tau)$ is the signal ambiguity function given by

$$A(\nu, \tau) = \int_{-\infty}^{\infty} s(u + \frac{\tau}{2}) s^*(u - \frac{\tau}{2}) e^{j2\pi\nu u} du \qquad (2)$$

The analytic signal $z(t)$ associated with the signal $s(t)$ is normally used in the above expressions to remove the aliasing effects [1,2]. The properties of a quadratic TFD are completely determined by its kernel. It is well-known in the literature that the kernel should have the shape of a two-dimensional low-pass filter so as to attenuate the cross-terms that exist away from the origin in the ambiguity domain and preserve the auto-terms that concentrate around the origin of this domain [2,14,15]. When a low-pass kernel is utilized, there is always a tradeoff between the resolution and cross-terms reduction.

The problem of non-parametric instantaneous frequency (IF) estimation for multi-component FM signals is an important issue in signal processing [16]. Adaptive techniques for IF estimation for mono- and multicomponent signals have been proposed and analyzed [17–19]. These techniques depend on the amplitude estimation of the signal components. Also amplitude estimation is of general interest in signal processing. Hence, it is of benefit in time - frequency

analysis that a non-parametric amplitude estimation is possible from the TFD. In [19] we have proposed that a TFD can support amplitude estimation for the signal components if its kernel in the time-lag domain, which is the inverse Fourier transform ($\nu \rightarrow t$) of the Doppler-lag kernel $g(\nu, \tau)$, is a function of time only.

In this paper we analyze quadratic TFDs with time-only kernels, hence the name T-distributions (initial ideas on this T - Class of TFDs appeared in [20] and [21]). It is worth noting that the T - Class is a subclass of Cohen's Class of time-frequency distributions. Two important members of the T-class are studied and compared to other efficient TFDs that have similar kernel shape in the time direction. In addition to their simpler structure, it appeared that time-only kernels are more efficient than their two-dimensional counterparts. Amplitude estimation is considered with numerical examples under noise-free and noisy conditions.

2. Rationale

We discuss the motivation behind the introduction of time - frequency distributions with time-only kernels. Any quadratic time-frequency distribution (TFD) is completely characterized by its time - lag kernel $G(t, \tau)$ or Doppler - lag kernel $g(\nu, \tau)$ [1]. The time-lag kernel $G(t, \tau)$ of any quadratic TFD is defined as follows

$$G(t, \tau) = \int_{-\infty}^{\infty} g(\nu, \tau) e^{-j2\pi\nu t} d\nu = \mathop{\mathcal{F}}_{\nu \to t} \{g(\nu, \tau)\} \tag{3}$$

where $\mathcal{F}$ is the Fourier transform. The Doppler-lag kernel is usually low-pass, real, and even in t and τ, where we assume these properties to be satisfied in this paper. Using eqs(1), (2), and (3) we can express the time - frequency distribution of the analytic signal $z(t)$ as follows

$$\rho(t, f) = \mathop{\mathcal{F}}_{\tau \to f} [G(t, \tau) \mathop{*}_{(t)} K_z(t, \tau)] \tag{4}$$

where $K_z(t, \tau) = z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2})$ is the instantaneous autocorrelation and $\mathop{*}_{(t)}$ is the convolution in the time domain.

Since the proposed time-only kernels represent a special case of separable time-lag kernels, we will discuss the latter first in the next subsection.

A. Separable Time-Lag Kernels

Suppose we have a time-lag kernel separable along the time and the lag domains as follows

$$G(t, \tau) = \lambda(t)\theta(\tau) \tag{5}$$

where λ and θ are continuous and L^2-integrable functions of time and lag,

respectively. High time - frequency is a main point of interest for any TFD. To test the TFD with the above kernel for resolution, we apply it to a complex sinusoid $z(t) = \exp(j2\pi f_o t)$, where by using eq.(4), we reach the following

$$\rho(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda(t - u)\theta(\tau) \exp\{-j2\pi(f - f_o)\tau\}dud\tau = M\Theta(f - f_o) \quad (6)$$

where

$$M = \int_{-\infty}^{\infty} \lambda(u)du, \quad \Theta(f) = \mathcal{F}_{\tau \to f}\{\theta(\tau)\}. \quad (7)$$

In order to have an ideal (maximum) energy concentration, we should have a Dirac delta function along the IF of the signal as follows

$$\rho(t, f) = \delta(f - f_o) \quad (8)$$

where δ is the Dirac delta function. Using the above relations, this result implies that

$$\theta(\tau) = 1/M \quad (9)$$

hence we have the following kernel

$$G(t, \tau) = \mathcal{G}(t) = \lambda(t)/M, \quad g(\nu, \tau) = g(\nu) = \mathcal{F}_{t \to \nu}^{-1}\{\lambda(t)\}/M \quad (10)$$

where $G(t, \tau)$ should now be a time-only kernel. Hence for best performance in resolution and cross-terms reduction, a separable time-lag kernel should be a time-only one. This is the proposed general formula for all time-only kernels, and the corresponding TFDs will be called accordingly as the T-distributions. Another factor of major importance in designing TFDs is their behavior in reducing cross-terms when analyzing multicomponent signals. Multicomponent signals appear in many applications, e.g., biomedical signals or signals from natural phenomena. To see the performance of the above time-only kernel in cross-terms reduction for multicomponent signals, we consider a signal composed of the sum of two complex sinusoids as follows

$$z(t) = a_1 \exp\{j(2\pi f_1 t + \theta_1)\} + a_2 \exp\{j(2\pi f_2 t + \theta_2)\} \quad (11)$$

where a_1, a_2 are real constants that represent the amplitudes, and θ_1 and θ_2 are phase constants. Complex sinusoids are already in the analytic form. Using eqs. (4), (10), and (11), we have the TFD of the above two-component signal as follows

$$\rho(t, f) = a_1^2 \delta(f - f_1) + a_2^2 \delta(f - f_2)$$
$$+2a_1 a_2 \, g(f_1 - f_2) \, \cos\{2\pi(f_1 - f_2)t + \theta_1 - \theta_2\} \, \delta(f - \frac{f_1 + f_2}{2}) \quad (12)$$

where there is ideal concentration (Dirac delta function) about the auto - terms

4

in the time - frequency plain. Cross-terms appear with a controlling factor $g(f_1 - f_2)$. It is worth noting that the well - known Wigner-Ville distribution (WVD) also utilizes a time - only kernel $G(t, \tau) = \mathcal{G}(t) = \delta(t)$ (whose time-lag counterpart is $g(\nu, \tau) = g(\nu) = 1$), but it has significant oscillatory cross-terms without a controlling factor (they can be larger in amplitude than the auto-terms, causing a confusion in signal analysis). The T - distributions use a low - pass time - only kernel (in the time-lag domain) other than $\delta(t)$; and this will result in controlling the cross-terms by the low-pass function g. In the case of two complex sinusoids above we have the controlling factor $g(f_1 - f_2)$ with cross-terms reduction depending on the shape of the low-pass function g and the frequency separation $f_1 - f_2$. Unlike the Choi - Williams and other two - dimensional kernel distributions which spread the cross-terms in the frequency domain, T - distributions have them more concentrated but attenuated relative to the auto-terms. This would be more evident in practical implementation where the delta functions above are converted by the discrete implementation into finite peaks. Another approach for separable kernels appeared in [15].

Although this derivation of high-resolution and cross-terms reduction does not include noise (interference), we used the result that noise will cause cross-terms with the original information-bearing signal (see [1]), hence, if the TFD is better in noise-free cross-terms reduction, we can conclude that it is better in noisy environments as well.

B. Properties of the T-Distributions

We examine the T-distributions against the following desirable properties of TFD's are explained in [1], [2], and [23]. A newly proposed instantaneous frequency property and the property of amplitude estimation will also be discussed.

B.1 Realness:

Since $\mathcal{G}(t)$ is a real and even function of t, we have $g(\nu)$ real and even in ν [22]. This will give a real TFD $\rho(t, f)$.

B.2 Time-Shift and Frequency Shift Invariance:

Since $g(\nu, \tau)$ is independent of time t and frequency f, $\rho(t, f)$ satisfies the time shift and frequency shift invariance properties.

B.3 Frequency Marginal and Group Delay:

A T - distribution $\rho(t, f)$ satisfies the frequency marginal property since, using eqs.(7) and (10), we have $g(0, \tau) = g(0) = 1 \; \forall \tau$. Now the group delay

property, $[\int t\rho(t,f)dt/\int \rho(t,f)dt] = t_g(f)$, is satisfied by the proper choice of g such that $[\partial g(\nu,\tau)/\partial\nu]\mid_{\nu=0} = 0 \,\forall\tau$.

B.4 *Frequency Support:*

The frequency support property is also satisfied since $\int g(\nu,\tau)e^{j2\pi f\tau}d\tau = g(\nu)\delta(f) = 0$ for $|f| > |\nu|/2$.

B.5 *Time Support:*

The time support property entails that $\rho(t,f) = 0$ for $|t| > t_o$ if the signal $s(t) = 0$ for $|t| > t_o$. Hence, we must have $G(t,\tau) = 0$ in the region $|t| > \frac{|\tau|}{2}$ [2,23]. Like the Choi-Williams distribution (CWD) and the Spectrogram, this condition is not exactly satisfied by the T - distributions, but it is approximately true for most of the region $|t| > \frac{|\tau|}{2}$ in the time - lag domain (due to the low-pass shape of $G(t,\tau)$ that substantially decays when $|t| > \frac{|\tau|}{2}$).

B.6 *Reduced Interference and Resolution:*

This property is satisfied by the T - distributions where they had high resolution (delta function for sinusoids) with a cross-terms reduction factor as shown in subsection (2-A) above for the case of two-sinusoids multicomponent signal.

B.7 *Time Marginal and Instantaneous Frequency:*

The instantaneous frequency $f_i(t)$ of a monocomponent signal $s(t)$ is the sinusoidal frequency that instantaneously fits the frequency of the signal, defined mathematically as follows [16]

$$f_i(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{13}$$

where $z(t) = a(t)e^{j\phi(t)}$ is the analytic signal associated with $s(t)$. Traditionally a time - frequency distribution $\rho(t,f)$ is looked upon as analogous to a probability distribution, hence it was imposed that the first moment of $\rho(t,f)$ with respect to f must equal the instantaneous frequency $f_i(t)$, leading to the conditions $g(\nu,0) = 1 \,\forall\nu$ and $\partial g(\nu,\tau)/\partial\tau\mid_{\tau=0} = 0 \,\forall\nu$ [23]. But many important TFDs (the spectrogram, Page, and Rihaczek distributions, for example) do not satisfy these conditions [2]. If the time-frequency distribution does not satisfy one of the marginals, the analogy with a probability distribution will no longer be true, and the traditional definition of the IF property will no longer be valid. Hence, in [20], we postulated the following general IF

property that would be applicable to practical signal analysis: at any time instant t, the time - frequency distribution $\rho(t, f)$ should have an absolute maximum at $f = \frac{1}{2\pi} \frac{d\phi(t)}{dt}$. This is the actual important characteristic needed for IF estimation in practice.

The T - distributions do not satisfy the time marginal, hence do not satisfy the traditional condition for the instantaneous frequency. But at any time t, the T-distributions have absolute maxima at $f = \frac{1}{2\pi} \frac{d\phi(t)}{dt}$ for sinusoids and linear FM signals. This is the basis for IF estimation using TFDs. For non-linear FM signals this IF estimate will be biased, and best IF estimation is achieved in this case by adaptive methods [17–19].

□ *IF Property:* At any time t, the T-distributions have absolute maxima at the frequency $f = \frac{1}{2\pi} \frac{d\phi(t)}{dt}$ for sinusoids and linear FM signals $z(t) = a \, e^{j\phi(t)}$, a being a constant or slowly varying function of time.

Proof: The T-distribution $\rho(t, f)$ of the signal $z(t)$ can be written as follows

$$\rho(t, f) = |a|^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j2\pi f\tau} G(t - u, \tau) \, e^{j[\phi(u+\tau/2)-\phi(u-\tau/2)]} du d\tau$$

$$= |a|^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j2\pi f\tau} \mathcal{G}(t - u) \, e^{j\tau[\phi'(u)+\sum_{\substack{k=3 \\ (k \text{ odd})}}^{\infty} \frac{\tau^{k-1}}{k!2^{k-1}}\phi^{(k)}(u)]} du d\tau$$

where Taylor series expansion has been used for the term $e^{j[\phi(u+\tau/2)-\phi(u-\tau/2)]}$ around the point $\phi(u)$. If we assume relatively small higher-order phase derivatives $\phi^{(k)}(t), k \geq 3$, we reach at the following approximation

$$\rho(t, f) \approx |a|^2 \int_{-\infty}^{\infty} \mathcal{G}(t - u) \, \delta \, [\frac{1}{2\pi}\phi'(u) - f] \, du = |a|^2 \mathcal{G}(t - \psi(f))\psi'(f) \quad (14)$$

where ψ is the inverse of $\frac{1}{2\pi}\phi'$, i.e., $\frac{1}{2\pi}\phi'(\psi(f)) = f$. We assume that $\psi'(f)$ does not have a high peak at any frequency f. Since we assumed that all time-only kernels have low-pass shapes and are even in the variable t, we note that $\mathcal{G}(t - \psi(f))$ should have a peak at $t = \psi(f)$. Hence, the absolute maximum of $\rho(t, f)$ for any time t would be at $\psi(f) = t$, or $f = \frac{1}{2\pi}\phi'(t)$, which is the instantaneous frequency of the FM signal $z(t)$. For non - linear FM signals, the energy peak of $\rho(t, f)$ would be biased from the instantaneous frequency due to the effect of the higher - order derivatives in the term $\sum_{\substack{k=3 \\ (k \text{ odd})}}^{\infty} \frac{\tau^{k-1}}{k!2^{k-1}}\phi^{(k)}(u)$, in which the major factor is $\phi^{(3)}(u)$ [17,18]. Therefore at the instants of rapid change in the IF law, the bias is not negligible and eq.(14) would not be an accurate approximation unless a suitable lag windowing is used.

For linear FM (LFM) signals, the phase derivatives above the third are all zero, i.e., $\phi^{(k)}(t) = 0$ for $k \geq 3$. For the general LFM signal $z(t) = ae^{j2\pi(f_o t + \frac{\beta_o}{2} t^2)}$,

where f_o and β_o are constants, we have the following structure for the T-distributions

$$\rho(t, f) = \frac{1}{\beta_o} |a|^2 \mathcal{G}\{t - \frac{1}{\beta_o}(f - f_o)\} \tag{15}$$

which has an absolute maximum at $f = f_o + \beta_o t$, the instantaneous frequency of the linear FM signal $z(t)$.

The above proof is applicable to any T - distribution $\rho(t, f)$ on condition that it is even in t and that its time - lag kernel is chosen to be of a low-pass shape. The proof is also applicable to slowly varying amplitudes $a(t)$ as per the Bedrosian Theorem [1]. $\square$

B.8 *Amplitude Estimation:*

In many applications it is necessary to extract the amplitude of the FM signal from its TFD. This can be more important when the signal under consideration is multicomponent, where there is no way to get the individual amplitudes except using time-frequency analysis. The T - distributions as explained in eq.(14) above enable amplitude estimation for FM signals using an estimate for ψ from the peak of $\rho(t, f)$. This estimation of ψ is more accurate in the linear parts of TFD. For a single component FM signal of the form $z(t) = a\, e^{j\phi(t)}$, we can estimate its amplitude alternatively as follows [17]

$$\hat{a}^2 = \frac{1}{N} \sum_{n=1}^{N} |z(nT)|^2 \tag{16}$$

For multicomponent FM signals, the task is a bit more complicated. We consider the following signal

$$z(t) = \sum_{q=1}^{M} a_q e^{j\phi_q(t)} \tag{17}$$

where the amplitudes of the individual components $\{a_q\}$ are constant. The individual IF laws are given by [1]

$$f_{i,q}(t) = \frac{1}{2\pi} \frac{d\phi_q}{dt} \quad ; \; q = 1, ..., M. \tag{18}$$

For the above signal, we have the following relationship [19]

$$\sum_{q=1}^{M} |\hat{a}_q|^2 = \frac{1}{N} \sum_{n=1}^{N} |z(nT)|^2 \tag{19}$$

where N is the total number of samples and the hat ˆ refers to the estimated values. In practice it is more accurate in case of multicomponent signals to find

the ratio between the components amplitudes, while the sum of the squares of the component amplitudes can be obtained from the signal in the time domain according to eq.(19). The ratio between the component amplitudes can be estimated. Assume that the ratio of the qth amplitude to the first amplitude is $r_q = |a_q|/|a_1|$, then we have the following estimates

$$|\hat{a}_q|^2 = \hat{r}_q^2 s_o / \{1 + \sum_{i=2}^{M} \hat{r}_i^2\} \tag{20}$$

where $s_o = \sum_{q=1}^{M} |\hat{a}_q|^2$. Using approximations for the T-distributions as per eq.(14), we can estimate the ratio r_q at the peaks (ridges) around the qth and the first components, $P_q(t,f)$ and $P_1(t,f)$, as follows

$$\hat{r}_q^2 = \operatorname*{mean}_{t,f}\{|\frac{P_q(t,f)}{\hat{\psi}_q'(f)}|\}/\operatorname*{mean}_{t,f}\{|\frac{P_1(t,f)}{\hat{\psi}_1'(f)}|\}. \tag{21}$$

We can use the peak trajectory to estimate $\phi_q'(t)$ and $\phi_1'(t)$, from which the quantities $\psi_q'(f)$ and $\psi_1'(f)$ can be estimated. It is better not to include the start and the end parts of the TFD in the estimation using eq.(21). Also the regions of rapid change in the IF law should be excluded as eq.(14) would not be an accurate approximation to the TFD there due to the effect of higher-order phase derivatives.

In the next section we study two members of the T - class.

3. Exponential and Hyperbolic Time - Only Kernels

In this section we present two time-only kernels and compare their performance with other well - known kernels using numerical examples. Amplitude estimation is also considered.

A. *The Exponential Time - Only Kernel*

The Choi - Williams distribution $CW(t,f)$ was a significant step in time-frequency analysis where it opened the way for optimizing resolution with cross-terms reduction [9]. The kernel of the Choi - Williams distribution (CWD) in the Doppler-lag domain is $g(\nu, \tau) = \exp(-4\pi^2 \nu^2 \tau^2 / \sigma)$, with its time-lag counterpart given by

$$G(t, \tau) = \sqrt{\sigma/4\pi\tau^2} e^{-\sigma t^2/4\tau^2} \tag{22}$$

where σ is a real parameter that will control the resolution and the cross-terms reduction. This two-dimensional exponential kernel has shown excellent performance in reducing cross-terms while keeping high resolution, with a compromise between these two requirements decided by the parameter σ.

Based on this work, we propose a time - frequency distribution $T_e(t, f)$ with exponential time-only kernel as follows

$$G(t, \tau) = \mathcal{G}_e^\eta(t) = \sqrt{\eta/\pi} e^{-\eta t^2} \tag{23}$$

where η is a real parameter and $\sqrt{\eta/\pi}$ is a normalization factor as it appears in eqs.(7-10). Using an argument similar to that in the Appendix, it can be shown that when η approaches ∞, the exponential T - distribution $T_e(t, f)$ will approach the Wigner - Ville distribution (WVD).

Now we examine the performance of this kernel in resolution (energy concentration) and cross-terms reduction using various examples and compare the performance of the corresponding exponential T - distribution (ETD) with the Choi - Williams distribution.

Example 1: *The Sum of Two Complex Sinusoids*

The exponential T - distribution $T_e(t, f)$ of the signal $z(t)$ given in eq.(11) can be written as follows

$$T_e(t, f) = a_1^2 \, \delta(f - f_1) + a_2^2 \, \delta(f - f_2) + 2a_1 a_2 \Lambda_\eta(t) \, \delta(f - \frac{f_1 + f_2}{2}) \tag{24}$$

where
$$\Lambda_\eta(t) = e^{-\frac{\pi^2 (f_1 - f_2)^2}{\eta}} \cos\left(2\pi(f_1 - f_2)t + \theta_1 - \theta_2 \right)$$

If η is small, the cross-terms can be reduced so that they are negligible compared with the auto - terms. If $\eta \to \infty$, we have $\exp\{-\pi^2(f_1 - f_2)^2/\eta\}$ approaches 1 and $T_e(t, f)$ approaches $W(t, f)$. The relative reduction in cross - terms is clearer in the above formula than the formula of CWD [9]. To see the efficiency of this cross - terms reduction compared to the two - dimensional exponential time - lag kernel in the practical implementation, we consider next a more complicated signal and compare $T_e(t, f)$ with $CW(t, f)$.

Example 2: *Multicomponent Linear FM Signals*

For linear FM signals, we compare the performance of $T_e(t, f)$ with $CW(t, f)$ using the sum of two linear FM signals

$$z(nT) = e^{j2\pi[0.1(nT)+0.0008(nT)^2]} + e^{j2\pi[0.3(nT)+0.0008(nT)^2]}$$

Figure 1 shows the discrete versions of $T_e(t, f)$ with parameter $\eta = 0.01$ and the Choi-Williams distribution $CW(t, f)$ with parameter $\sigma = 20$ for the above signal, with total length $N = 64$ and sampling interval $T = 1$ at the discrete time instant $n = t/T = 30$. Note that for both TFD's there is a compromise between the two requirements: resolution and cross - terms reduction.
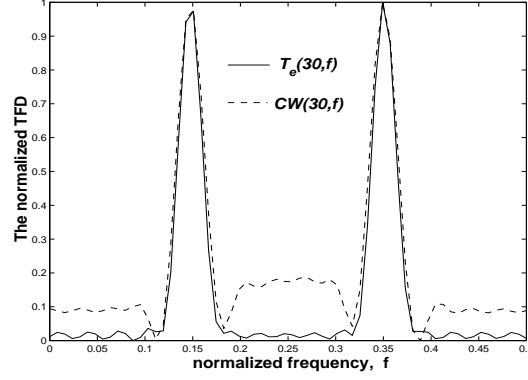
Fig. 1. Performance comparison between the discrete versions of $T_e(t, f)$ for $\eta = 0.01$ and $CW(t, f)$ for $\sigma = 20$ using a two - component linear FM signal with $N = 64$ and sampling instant $T = 1$ at the time instant $t = 30$. Due to the inherent compromise between high resolution and cross-terms suppression, changing these parameters will improve one of these two requirements at the expense of the other.

Changing the parameter of the TFD results in improving one of the above requirements at the expense of the other.

Example 3: *Multicomponent Real - Life Signals*

In this example we compare the one-dimensional exponential kernel with the Choi - Williams two - dimensional exponential kernel using a non - linear FM real-life signal, a bat signal. As shown in Figure 2, The cross - terms in $CW(t, f)$ can obscure the weak component (far right). In this comparison we used $\sigma = 30$ for $CW(t, f)$ and $\eta = 0.05$ for $T_e(t, f)$, with total signal length $N = 400$ and sampling interval $T = 1$.

Example 4: *A Comparison with WVD: Linear FM Signals*

It is well - known that the Wigner - Ville distribution $W(t, f)$ is ideal for a mono-component linear FM signal since it gives a delta function around the IF law of such a signal. The proposed exponential T - distribution $T_e(t, f)$ has a performance comparable to that of WVD for linear FM signals as shown in Figure 3 using the signal $z(nT) = e^{j2\pi[0.1(nT)+0.0008(nT)^2]}$ with total signal length $N = 64$, $T = 1$, $\eta = 0.01$ at the discrete time instant $n = t/T = 30$.

Example 5: *Amplitude Estimation*

We investigate the amplitude ratio estimate for a multicomponent FM signal under noise - free and noisy conditions using the component peaks (ridges) of $T_e(t, f)$. Two-component FM signal with non - linear IF laws is considered as follows
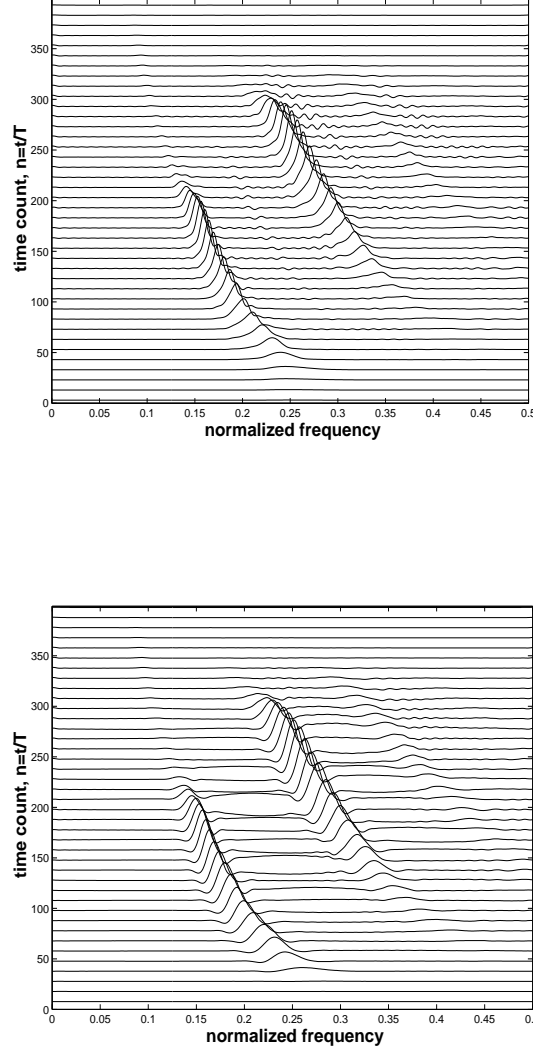
11

Fig. 2. Performance comparison between $T_e(t, f)$ for $\eta = 0.05$ (above) and $CW(t, f)$ for $\sigma = 30$ (below) using a bat signal. The total signal length is $N = 400$ and the sampling interval is $T = 1$.

$$z(nT) = a_1 e^{j2\pi[47 + 2.5\sinh^{-1}(10(nT - 0.3))]} + a_2 e^{j2\pi[30 + 2.5\sinh^{-1}(20(nT - 0.8))]} + \epsilon(t)$$

where $\epsilon(t)$ is a complex-valued white Gaussian noise with independent identically distributed (i.i.d.) real and imaginary parts with total variance σ_ϵ^2. If $r = a_1/a_2$ and $\hat{r}$ is the estimated value of the ratio r as per eq.(21), then the estimation error can be found as follows

$$e_r = |(\hat{r} - r)/r|. \tag{25}$$

The estimation error for different values of r and signal - to - noise ratios is shown in Figure 4. We used the exponential T - distribution $T_e(t, f)$ with a
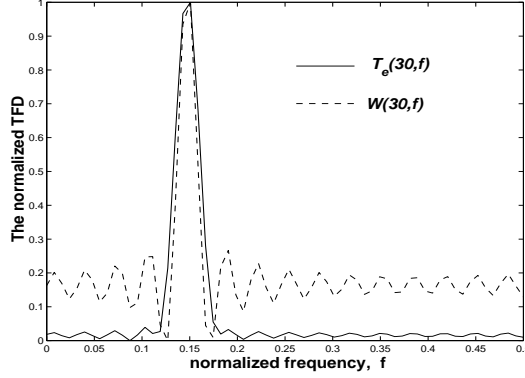
Fig. 3. Performance comparison between $T_e(t,f)$ for $\eta = 0.01$ and $W(t,f)$ using a mono-component linear FM signal at the time instant $t = 30$. Total signal length is $N = 64$ and the sampling interval is $T = 1$. Ripples appear due to finite signal length, which is equivalent to time and lag windowing. However, $T_e(t,f)$ performs better than $W(t,f)$ in reducing these ripples, although the latter is slightly better in resolution.

total number of samples $N = 128$, $T = 1$, and control parameter $\eta = 0.01$. For the amplitudes ratio estimate we used the following formula:

$$\hat{r}^2 = |\max_{(t,f)\in L_1} \{P_1(t,f)\beta_1\} / \max_{(t,f)\in L_2} \{P_2(t,f)\beta_2\}| \tag{26}$$

where we used the linear regions L_1 and L_2 of the IF laws, with β_1 and β_2 representing the approximate slopes of the IF laws as functions of time in the linear regions L_1 and L_2 as per eq.(15).

It is worth noting that less overall root - mean - squared error over all ratios is obtained for higher values of the signal - to - noise ratio (SNR). In the above example, the RMS values of error for noise - free, SNR=30 dB, and SNR=15 dB are 0.0357, 0.0363, and 0.0411, respectively. There is a non-zero error even in the noiseless case due to windowing and discretization of the TFD (calculating its value using windowing and finite summations). Another source of error may emerge when there is no linear part in the IF law of the signal, where in this case the bias in the estimate of $\psi'(f)$ would be large as the effect of higher-order derivatives of the IF law would be non-negligible.

B. *The Hyperbolic Time - Only Kernel*

The hyperbolic time - only kernel was proposed in [19,20] as follows

$$G(t,\tau) = \mathcal{G}_h^\alpha(t) = k_\alpha / \cosh^{2\alpha}(t) \tag{27}$$

where α is a real positive number and k_α is a normalization factor given by $k_\alpha = \int_{-\infty}^{\infty} 1/\cosh^{2\alpha}(t) = \Gamma(2\alpha)/2^{2\alpha-1}\Gamma^2(\alpha)$ (Γ stands for the gamma function). However, the T - distribution associated with this kernel, the hyperbolic T -
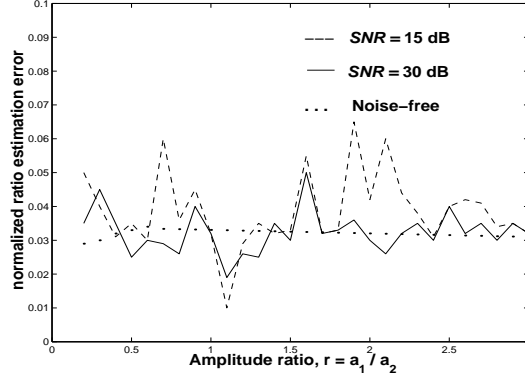
Fig. 4. The amplitudes ratio estimation error for a two - component non - linear FM signal with different values of the signal - to - noise ratio (SNR). The distribution $T_e(t, f)$ is used with a total number of samples $N = 128$, $T = 1$, and $\eta = 0.01$
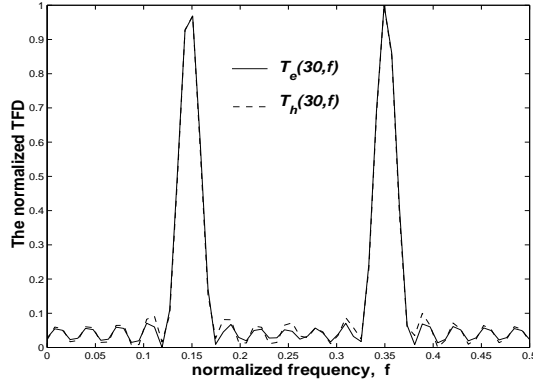

Fig. 5. Performance comparison between the ETD $T_e(t, f)$ for $\eta = 0.015$ and the HTD $T_h(t, f)$ for $\alpha = 0.1$ using a two-component linear FM signal at the time instant $t = 30$. The total signal length is $N = 64$ and the sampling interval is $T = 1$. The parameters are arranged such that both TFDs have the same resolution. Slight reduction in cross-terms and ripples is obtained by $T_e(t, f)$.

distribution (HTD) $T_h(t, f)$, has a performance slightly inferior to that of $T_e(t, f)$ discussed in the previous subsection as shown in Figure 5. As $\alpha \to \infty$, we have $\mathcal{G}_h^\alpha(t, \tau) \to \delta(t)$ and $T_h(t, f)$ approaches the Wigner-Ville distribution $W(t, f)$ (see Appendix).

Example 6: *Instantaneous Frequency Estimation*

We investigate the performance of the above T - distributions in IF estimation of FM signals and compare that to the performance of WVD and CWD. A complex linear FM signal was used as follows $z(t) = a\, e^{j\phi(t)}$, $\phi(t) = 2\pi(f_o t + \beta t^2/2)$, with $a = 1$, $f_o = 20$ Hz, and $\beta = 0.5$. The signal length was $N = 512$ samples, and the sampling frequency was $f_s = N$ Hz. The exponential T - distribution (ETD) was used with parameter $\eta = 0.01$, the hyperbolic T -

14

distribution (HTD) was used with parameter $\alpha = 0.1$, while the CWD was used with parameter $\sigma = 19$. The instantaneous frequency was calculated using

$$\widehat{f}_i(t) = \arg[\max_f \rho(t, f)] \quad ; \; 0 \le f \le f_s/2 \tag{28}$$

at the mid of the signal, i.e., at the time instant $n = N/2$ (where the TFD is most accurate), then an estimate was taken as the average of these calculations over 1000 Monte Carlo simulations. Figure 6 shows that the above two T - distributions (which have similar performance represented by the solid curve) surpass WVD and CWD in robustness; they give the minimum estimation variance that is close to the Cramer - Rao Bound, especially at low SNRs, where the performance curves of both WVD and CWD deteriorate significantly.
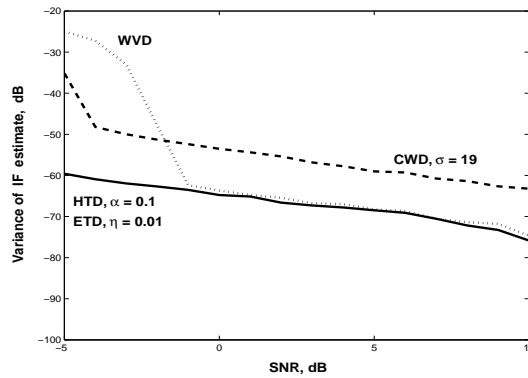

Fig. 6. Instantaneous frequency estimation of a linear FM signal using different TFDs. The T - distributions (solid curve) surpass WVD and CWD in robustness where they give the minimum variance, especially at low SNRs.

5. Conclusions

In this paper we have presented a sub-class of Cohen's Class of time - frequency distributions (TFDs) that has interesting properties. The kernels of the proposed class are time-only in the continuous time-lag domain. We have analyzed separable time - lag kernels and proved that for best performance, these kernels should be lag-independent. If their parameters are properly chosen, time-only kernels (and their corresponding TFDs: the T - distributions) are efficient in optimizing the two requirements: cross-terms reduction and high energy concentration around the IF law of the signal signal. Numerical examples have shown that time-only kernels are more efficient than their two-dimensional counterparts in this respect. Two distributions of this class are studied and numerical comparisons are made using synthetic and real - life signals. In addition to their superior performance in resolution and cross-terms reduction, non-parametric amplitude estimation is possible directly from the T - distributions in case of FM signals, a merit that does not exist in other TFDs. The performance of the amplitude estimator is studied for a two - com-

ponent non - linear FM signal under noise - free condition and in the presence of additive Gaussian noise.

Acknowledgement

Appendix

We here prove that when α approaches infinity, the hyperbolic time-only kernel $\mathcal{G}_h^\alpha(t)$ approaches $\delta(t)$. Similar reasoning can be applied to the exponential time - only kernel $\mathcal{G}_e^\eta(t)$. It is more convenient here to use the following formula [24]

$$\Gamma(2\alpha) = 2^{2\alpha-1}\Gamma(\alpha)\Gamma(\alpha + \frac{1}{2})/\sqrt{\pi}$$

Hence, the hyperbolic time-lag kernel can be written as

$$\mathcal{G}_h^\alpha(t) = \Gamma(\alpha + \frac{1}{2})/\sqrt{\pi}\Gamma(\alpha)\cosh^{2\alpha}(t)$$

The Stirling's asymptotic formula states that [25]

$$\Gamma(x) \to \sqrt{2\pi}\, x^{x+\frac{1}{2}} e^{-x+\frac{\beta}{12x}} \quad \text{as} \ \ x \to \infty, \quad 0 < \beta < 1$$

Using this formula and taking the natural logarithm of $\mathcal{G}_h^\alpha(t)$ we have

$$\log \mathcal{G}_h^\alpha(t) \to \alpha(\frac{\log \alpha}{2\alpha} + 2\log\frac{2}{e^t + e^{-t}}) \quad \text{as} \ \ \alpha \to \infty.$$

Noting that $\lim\limits_{\alpha\to\infty} \dfrac{\log \alpha}{\alpha} = 0$ (L'Hopital's rule), we have

$$\lim_{\alpha\to\infty} \log \mathcal{G}_h^\alpha(t) = \begin{cases} -\infty & t \neq 0 \\ \infty & t = 0. \end{cases}$$

Hence, when $\alpha \to \infty$, $\mathcal{G}_h^\alpha(t)$ is zero everywhere except at $t = 0$ where it is infinite. Since we have $\mathcal{G}_h^\alpha(-t) = \mathcal{G}_h^\alpha(t) \ \forall \alpha$ and

$$\int_{-\infty}^{\infty} \mathcal{G}_h^\alpha(t)\, dt = 1 \ \ \forall \alpha$$

then we can say that

$$\mathcal{G}_h^\alpha(t) \to \delta(t) \quad \text{as} \ \ \alpha \to \infty.$$

16

References

[1] Boashash B. Time-frequency signal analysis. In: Advances in Spectrum Estimation (S.Haykin: editor). Prentice-Hall, NJ, 1991.

[2] Boashash B (editor). Time-frequency Signal Analysis. Longman Cheshire, Melbourne, Australia, 1992

[3] L. Cohen, "Time-frequency distributions - a review" *Proc. IEEE*, vol. 77, no. 7, pp. 941-981, July 1989.

[4] L. Cohen, *Time-Frequency Analysis*, Prentice Hall PTR, Englewood Cliffs, New Jersey, 1995.

[5] E. P. Wigner, "On the quantum correction for thermody- namic equilibrium," Phys. Rev., vol. 40, pp. 749-759, 1932.

[6] J. Ville, "Theorie et applications de la notion de signal ana- lytique," Cables et Transmission, vol. 2A, pp. 61-74, 1948.

[7] L. Cohen, "Generalized phase-space distribution functions" *J. Math. Phys.*, vol. 7, pp. 781-786, 1966.

[8] Stankovic LJ, Katkovnic V. The Wigner distribution of noisy signals with adaptive time-frequency varying window. IEEE Trans Signal Processing 1996; 47 (4): 1099-1108

[9] Choi H, Williams W. Improved time-frequency representation of multi-component signals using exponential kernels. IEEE Trans Signal Processing 1989; 37 (6): 862-871

[10] Jeong J, Williams W. Kernel design for reduced interference distributions. IEEE Trans Signal Processing 1992; 40 (2) 402-412

[11] Amin MG, Carroll JF. Time-frequency kernel design via point and derivative constraints. Proc IEEE ICASSP, Adelaide, 1994; 4: 309-312

[12] Stankovic LJ. A method for improved distribution concentration in the time-frequency analysis of multi-component signals using the L-Wigner distribution. IEEE Trans Signal Processing 1995; 43 (5) 1262-1269

[13] Jones DL, Baraniuk RG. An adaptive optimal-kernel time-frequency representation. IEEE Trans Signal Processing 1996; 44 (9) 2352-2356

[14] Qian S, Chen D. Joint time-frequency analysis: methods and applications. Upper Saddle River, N.J.: PTR Prentice Hall, 1996

[15] Flandrin P. Time-frequency/time-scale analysis. Academic Press, San Diego, 1999.

[16] Boashash B. Estimating and interpreting the instantaneous frequency of a signal. Proc IEEE 1992; 8 (4) 520-568

17

[17] Katkovnik V, Stankovic LJ. Instantaneous frequency estimation using the Wigner distribution with varying and data driven window length. IEEE Trans Signal Processing 1998; 46 (9) 2315-2325

[18] Hussain ZM, Boashash B. Adaptive instantaneous frequency estimation of multi-component FM signals. Proc IEEE ICASSP, Istanbul, 2000; 2: 657-660

[19] Hussain ZM, Boashash B. Multi-component IF estimation. Proc IEEE SSAP 2000; 559-563

[20] Hussain ZM, Boashash B. Adaptive instantaneous Ffrequency estimation of multicomponent FM signals using quadratic timeŰfrequency distributions. IEEE Trans Signal Processing 2002; 50 (8) 1866-1876

[21] Hussain ZM, Boashash B. Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals. Proc IEEE ISSPA 2001; 339-342

[22] Oppenheim AV, Willsky AS, Nawab SH. Signals and Systems. Prentice Hall, New Jersey, 1997

[23] Claasen TCM, Mecklenbrauker WFG. The Wigner distribution - A tool for time-frequency signal analysis - Part III: Relations with other time-frequency signal transformations. Philips J Res 1980; 35 (6) 372-389

[24] Abramowitz M, Stegun IA (editors). Handbook of Mathematical Functions. U.S. Govt. Printing Office, Washington, 1972

[25] Pierpont J. Lectures on the Theory of Functions of Real Variables. Dover Publications, New York, 1912