can be solved by using computationally efficient algorithms. Several examples have been provided to demonstrate the efficacy of our method.

## REFERENCES


[1] P. A. Regalia, S. K. Mitra, and P. P. Vaidyanathan, "The digital all-pass filter: A versatile signal processing building block," *Proc. IEEE*, vol. 76, pp. 19–37, Jan. 1988.

[2] K. Shenoi, M. J. Narasimha, and A. M. Patterson, "On the design of recursive digital filters," *IEEE Trans. Circuit Syst.*, vol. CAS-23, pp. 485–489, Aug. 1976.

[3] F. J. Brophy and A. C. Salazar, "Two design techniques for digital phase networks," *Bell Syst. Tech. J.*, vol. 54, pp. 767–781, Apr. 1975.

[4] A. T. Chottera and G. A. Jullien, "A linear programming approach to recursive filter design with linear phase," *IEEE Trans. Circuits Syst.*, vol. CAS-29, pp. 139–149, Mar. 1982.

[5] Z. Jing, "A new method for digital all-pass filter design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1557–1564, Nov. 1987.

[6] A. G. Deczky, "Synthesis of recursive digital filters using the minimum $p$-norm criterion," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 257–263, Oct. 1972.

[7] ——, "Equiripple and minimax (Chebyshev) approximations for recursive digital filters," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-22, pp. 98–111, Apr. 1974.

[8] M. Ikehara, M. Funaishi, and H. Kuroda, "Design of all-pass networks using Reméz algorithm," in *Proc. IEEE Int. Symp. Circuits and Systems*, May 1991, pp. 364–367.

[9] M. Lang and T. I. Laakso, "Simple and robust method for the design of allpass filters using least-squares phase error criterion," *IEEE Trans. Circuit Syst. II*, vol. 41, pp. 40–48, Jan. 1994.

[10] S.-C. Pei and J.-J. Shyu, "Eigenfilter desing of 1-D and 2-D IIR digital allpass filters," *IEEE Trans. Signal Processing*, vol. 42, pp. 966–968, Apr. 1994.

[11] T. Q. Nguyen, T. I. Laakso, and R. D. Koilpillai, "Eigenfilter approach for the design of allpass filters approximating a given phase response," *IEEE Trans. Signal Processing*, vol. 42, pp. 2257–2263, Sept. 1994.

[12] S. Sunder and V. Ramachandran, "Recursive allpass filter design using least-squares techniques," in *Proc. IEEE Int. Symp. Circuits and Systems*, London, vol. 5, May 1994, pp. 791–794.

[13] G. A. Merchant and T. W. Parks, "Efficient solution of a Toeplitz-plus-Hankel coefficient matrix system of equations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 40–44, Feb. 1982.

[14] I. Gohberg and I. Koltracht, "Efficient algorithm for Toeplitz-plus-Hankel matrices," *Integ. Equations Oper. Theory*, vol. 12, pp. 136–142, Jan. 1989.

[15] G. W. Stewart, *Introduction to Matrix Computation.* New York: Academic, 1973.

[16] S. Sunder and V. Ramachandran, "Design of equiripple digital differentiators and Hilbert transformers using a weighted least-squares technique," *IEEE Trans. Signal Processing*, vol. 42, pp. 2504–2509, Sept. 1994.


# Auto-Term Representation by the Reduced Interference Distributions: A Procedure for Kernel Design

LJubiša Stanković


***Abstract*— An analysis of auto-term presentation using the reduced interference distributions (RID) is done. Comparison with an ideal time-frequency signal representation is taken as a basis for this analysis. The following distributions are considered: Choi-Williams, Zao-Atlas-Marks, Born-Jordan, Sinc, Zhang-Sato, Butterworth, Spectrogram, and the author's recently proposed S-method for time-frequency analysis. Various distributions produce different auto-term shapes. In all cases, the condition for cross-term reduction is contradictory to the condition for high auto-term quality. A procedure for designing a kernel that will produce the desired auto-term shape is demonstrated. An optimal kernel, with respect to the auto-term quality and cross-term suppression, is derived.**


## I. INTRODUCTION

Time-frequency analysis has attracted the attention of many researchers. The main challenge in this area lies in the fact that many fundamental questions are still waiting for viable answers. A whole variety of tools for time-frequency analysis, mainly rendered in the form of energy distributions in the time-frequency plane, has been proposed (Spectrogram, Wigner distribution, Rihaczek distribution, Page distribution, Choi–Williams distribution, etc.; for a complete list and source references, see [1] and [2]). Cohen has shown that all the above distributions are simply special cases of a general class of distributions obtained for a particular choice of an arbitrary function (kernel) [1]. Generally, these distributions belong to the class of quadratic signal transforms. Due to their quadratic nature, they are inevitably accompanied by undesirable effects, manifesting themselves as the cross-terms. The shapes, location, and other cross-term properties have been intensively studied, [3]–[7], [10]–[13], [15]–[17], [20]. A class of distributions having the property of reducing cross-terms is defined as reduced interference distributions (RID's) [10].

However, the related papers devote little attention to the shape of auto-terms that, after all, represent an ultimate goal of any time-frequency analysis. Therefore, the primary motivation for this research was to provide some additional insight into the auto-term shapes for the RID class of distributions.

This correspondence is organized as follows. Section II presents an ideal time-frequency representation of the frequency modulated signals. In Section III, the auto-term function, in the Cohen class of distributions, is defined. Auto-term forms, produced by various distributions, are compared in Section IV. The procedure for kernel design is demonstrated in Section V.

## II. IDEAL TIME-FREQUENCY REPRESENTATION

Generally, for an arbitrary signal, there is not a unique answer to the fundamental question "What should an ideal distribution look like?" However, the answer may be precise, and with a full physical meaning, if one concentrates on a specific class of signals [8], [14]. This is basically the approach that we have pursued in this work.


Manuscript received June 2, 1995; revised November 23, 1995. The associate editor coordinating the review of this paper and approving it for publication was Dr. Boualem Boashash.



The author is with Elektrotehnicki fakultet, University of Montenegro, Podgorica, Montenegro, Yugoslavia.



Publisher Item Identifier S 1053-587X(96)03967-0.

The bandwidth relation for any signal of the form $x(t) = A(t)e^{j\phi(t)}$ is given by [8], [18]

$$B^2 = \int_{-\infty}^{\infty} \left(\frac{A'(t)}{A(t)}\right)^2 A^2(t)dt + \int_{-\infty}^{\infty} \left(\phi'(t) - \langle\omega\rangle\right)^2 A^2(t)dt. \quad (1)$$

If the mean frequency is equal to the instantaneous frequency ($\langle\omega\rangle = \omega_i(t) = \phi'(t)$), then the instantaneous bandwidth is $\sigma^2_{f/t} = (A'(t)/A(t))^2$, [1], [8]. Details on this definition may be found in [8]. It is apparent that the ideal distribution for signals with $|A'(t)/A(t)| << 1$, i.e., $\sigma^2_{f/t} \to 0$ , should be in the form

$$I(\omega, t) = 2\pi A^2(t)\delta(\omega - \phi'(t)). \quad (2)$$

This distribution is ideally concentrated, along the $\omega$ axis, at the instantaneous frequency $\phi'(t)$. It is also ideally concentrated along the $t$ axis at the group delay $t = (\phi'(\omega))^{-1} = \varphi(\omega)$ (an inverse function of the instantaneous frequency $\phi'(t)$ under the conditions described in [14] is the group delay function $\varphi(\omega)$). The same expression for the ideal distribution may be obtained from the analysis of $x(t)$ using the stationary phase method. This analysis is done in [13] (see also the Appendix). Similar forms of the ideal distribution are proposed and analyzed in [9] and [14].

If signal $x(t)$ is multicomponent, i.e., $x(t) = \sum_{m=1}^{M} x_m(t)$ [6], [9], [14], then it is very difficult to define how the ideal distribution should look. We will assume that it should be equal to a sum of the ideal distributions of each individual signal's component. Note that the ideal distribution, in this case, is not an energetic one with respect to $x(t)$, but it is energetic with respect to each component separately.

## III. Auto-Terms in the Cohen Class of Distributions

As stated in the previous section, a specific class of signals will be defined for the auto-term shape analysis. In order to define such a class of signals, consider the Cohen class of distributions, bearing in mind that all shift covariant time-frequency distributions belong to this class [1], [2], [13]:

$$\begin{aligned} CD(\omega, t) = \tfrac{1}{2\pi} \iiint_{-\infty}^{\infty} & c(\theta, \tau)x(u + \tau/2) \\ & x^*(u - \tau/2)e^{-j\theta t - j\omega\tau + j\theta u} du d\theta d\tau \end{aligned} \quad (3)$$

where $c(\theta, \tau)$ is an arbitrary kernel function. Ideal distribution (2) may be easily translated into form (3) (taking an inverse 2-D Fourier transform of $I(\omega, t)$ and then its 2-D Fourier transform) as

$$I(\omega, t) = \frac{A^2}{2\pi} \iiint_{-\infty}^{\infty} e^{j\phi'(u)\tau} e^{-j\theta t - j\omega\tau + j\theta u} du d\theta d\tau \quad (4)$$

where $A(t)$ is treated as a constant $A$. Comparing (3) and (4) while having in mind the uniqueness of the Fourier transform, we get that signal $x(t) = Ae^{j\phi(t)}$ has the distribution equal to the ideal one (see (2)) iff

$$c(\theta, \tau)e^{j\phi(u+\tau/2) - j\phi(u-\tau/2)} = e^{j\phi'(u)\tau}.$$

Expanding $\phi(u \pm \tau/2)$ into a Taylor series around $u$, up to the second-order term, we get

$$c(\theta, \tau) = e^{-j\frac{\phi^{(3)}(u+\tau_1) + \phi^{(3)}(u-\tau_2)}{3!}\left(\frac{\tau}{2}\right)^3}$$

where $\tau_1$ and $\tau_2$ are variables ranging from 0 to $\tau/2$. From the last equation, one may conclude that for any signal $x(t)$, there exists its own kernel such that the Cohen distribution is equal to the ideal one [9]. Here, we will restrict the analysis to the case of signal-independent kernels. With this assumption (which is of practical importance), we get that the ideal distribution may be obtained only if $\phi^{(3)}(u) \equiv 0$, i.e., $c(\theta, \tau) \equiv 1$. This is the Wigner distribution kernel. The previous requirement ($\phi^{(3)}(u) \equiv 0$) is met only if the signal is linear frequency modulated:

$$x(t) = A(t)e^{j(at^2/2 + bt)} \quad (5)$$

where $A(t)$ should be treated as a constant within the considered time interval.

This class of signals satisfies some other important properties as well:

1) Signals (5) have simple mathematical form.
2) At the same time, these signals are not so simple that two or more distributions from the Cohen class have the same auto-term.
3) There exists one distribution having the ideal time-frequency representation defined by (2).
4) Such a class of signals is of great practical importance.
5) The Fourier transform of $x(t)$ belongs to the same class of signals in the $\omega$ domain $FT\{Ae^{j(at^2/2+bt)}\} = A\sqrt{j2\pi/a}\; e^{-j(\omega-b)^2/(2a)}$ so that definitions from one domain ($t$ or $\omega$) may be directly applied to the other one ($\omega$ or $t$).
6) In addition, theoretically, a very important delta pulse $\delta(t)$ may be written as a limit of (5) (details on this form of the delta pulse may be found in [18]):

$$\delta(t) = \lim_{a \to \infty} \sqrt{\frac{a}{2\pi j}} e^{jat^2/2}. \quad (6)$$

This way, the delta pulse may be formally treated as a frequency-modulated signal of form (5) with $A = \sqrt{\frac{a}{2\pi j}}$.Its ideal representation, corresponding to (2), is $I(\omega, t) = \lim_{a\to\infty} 2\pi \frac{a}{2\pi}\delta(\omega - at) = \delta(t)$.

The Cohen distribution of signal $x(t)$, given by (5), is

$$CD(\omega, t) = A^2 \int_{-\infty}^{\infty} c(-a\tau, \tau)e^{j(at+b-\omega)\tau} d\tau \quad (7)$$

$$\begin{aligned} CD(\omega, t) &= A^2 \mathbf{C}(\omega - at - b) \quad \text{with} \\ \mathbf{C}(\omega) &= FT\{c(-a\tau, \tau)\}. \end{aligned} \quad (8)$$

The auto-term shape is determined by function $\mathbf{C}(\omega)$ which will be referred to as the *auto-term function.*[1]

It is evident that a member of the Cohen class (considering only signal independent kernels) having the ideal auto-term shape is the one with $c(-a\tau, \tau) = 1$ for any $\tau$ and $a$. This is precisely the Wigner distribution.[2] However, as it is widely known, this distribution has very emphatic cross-term effects since its kernel is not of a lowpass 2-D filter type [10]. Any other distribution will have auto-terms that are more or less distorted when compared with the ideal representation (2).

In order to provide a basis for the analysis that follows, we will also review some results on the locations of the auto-terms and cross-terms in the $(\theta, \tau)$ plane. Consider a simple two-component signal $x(t) = x_1(t) + x_2(t)$. Suppose that the following practical assumption can be made: The signal's components may be treated as being time and frequency limited, i.e., $x_{1,2}(t) = 0$ for $|t - t_{1,2}| > T/2$ and $X_{1,2}(\omega) = FT\{x_{1,2}(t)\} = 0$ for $|\omega - \omega_{1,2}| > W/2$. From (3), it follows that the auto-terms are located in the $\tau$ direction within the interval $|\tau| < T$, whereas the cross-terms lie in the region

[1] Although we have defined $\mathbf{C}(\omega)$ as a function of $\omega$, we may also write $CD(\omega, t) = A^2\mathbf{C}(a(\frac{\omega-b}{a} - t))$, where $(\omega - b)/a$ is a group delay of signal (5). Thus, the auto-term function may be written as a function of $t$ in the form $\mathbf{C}(at)$.

[2] A signal-dependent kernel should have $c(\theta, \tau) = 1$ only along the line $\theta = -a\tau$ for a given $a$ in the $(\theta, \tau)$ plane. Since the ambiguity function is equal to zero outside this line, $c(\theta, \tau)$ may take any values there. Refer to [19], where the distribution (with signal dependent kernel) having this property is analyzed.

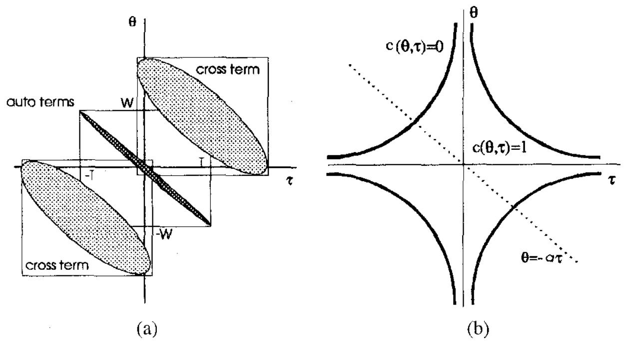


Fig. 1. (a) Auto-term and cross-term illustration in $(\theta, \tau)$ domain; (b) reduced interference distribution kernel support (Sinc distribution).

$|\tau - t_1 + t_2| < T$. The same results may be obtained from the Cohen distribution definition in the frequency domain [1], [2]. The auto-terms are located in the region $|\theta| < W$, whereas the cross-terms are dislocated from the origin $|\theta - \omega_1 + \omega_2| < W$. A graphical representation of the previous consideration is given in Fig. 1(a) [2]. A distribution, with reduced cross-terms should have the kernel of a lowpass filter type [10]. At the same time, the marginal properties[3] are satisfied if $c(\theta, 0) = c(0, \tau) \equiv 1$. This means that a distribution that satisfies marginal properties and belongs to the RID class should have a kernel support, as illustrated in Fig. 1(b). It is apparent that the condition for a good cross-term reduction is that the region of kernel support is as narrow as possible around the origin and the coordinate axes $\theta$ and $\tau$. However, the auto-term quality is higher if this region is wider; see (8). Those are contradictory requirements that are differently compromised in various distributions. Since the cross-terms are studied in detail in the cited references, we will not pursue their analytical forms in this paper. We will instead focus our attention on the auto-terms only.

## IV. Auto-Terms in the RID's

According to (8), one is able to derive the auto-term function for any distribution from the Cohen class. Results for some typical distributions are presented in Table I and Figs. 2 and 3. The following distributions are considered: Pseudo Wigner distribution (PWD) [1]–[3], Sinc distribution [1], [2], Choi-Williams distribution (CWD) [5], Zhao–Atlas–Marks distribution (ZAMD) [11], Born–Jordan distribution (BJD) [1], [2], Zhang–Sato distribution (ZSD) [20], Butterworth distribution (BD) [21], smoothed pseudo Wigner distribution [1], [2], Spectrogram [1], [2], and the S-method [12], [13], [15]–[17].

Here, we will provide some additional explanations for the results contained in Table I (presenting auto-term functions and their widths). Auto-term functions for the ZAMD and the ZS distributions, as well as for the Spectrogram, are approximations obtained using the stationary phase method (see the Appendix). Functions $K(\omega), C(\omega)$ and $S(\omega)$ (in Table I) are the Fresnel function and its real and imaginary parts, respectively.[4] The ambiguity function $A_{ww}(\theta, \tau)$ [1], [2] is defined by

$$A_{ww}(\theta, \tau) = \int_{-\infty}^{\infty} w(t + \tau/2) w^*(t - \tau/2) e^{-j\theta t} dt$$

[3] The marginal properties are satisfied if 1) the integral of a distribution over frequency is equal to the signal power $|x(t)|^2$, and 2) the integral of a distribution over time is equal to the spectral energy density $|X(\omega)|^2$.

[4] The Fresnel function is defined by $K(\omega) = \sqrt{\frac{2}{\pi}} \int_0^{\sqrt{\omega}} e^{ju^2} du$.

whereas $P(\theta) = FT\{p(t)\}$ is a window-function in the S-method:

$$SM(t, \omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} P(\theta) STFT(t, \omega + \theta) STFT^*(t, \omega - \theta) d\theta$$

where $STFT(t, \omega) = FT_\tau\{x(t+\tau)w(\tau)\}$. The values of $\alpha$, $\sigma$, and $\theta_1\tau_1$ are chosen such that the widths of distribution kernels' main lobes, in the $\theta\tau$ direction, are the same. For positive kernels, we assumed that their widths are defined by the points where the kernels were attenuated $e \approx 2.718$ times, whereas the width of the Butterworth distribution kernel is defined by the well-known 3 [dB] value. The width of window $w^2(\tau/2)$ is denoted by $T$. The width of $P(\theta)$, in the S-method, is assumed to be one tenth of $\theta_{\max}$, thus producing approximately the same kernel width, along $\theta$, for numerical data given in Fig. 2. This width of $P(\theta)$ is twice (five times) less in Fig. 2. (Fig. 3) than the theoretical one that would produce the same auto-term as the pseudo Wigner distribution does [13], [15]–[17]. However, this window narrowing resulted not in a wider auto-term but in small side lobe appearances, see Figs. 2(g) and 3(g) (note that the S-method may be understood as a deconvolution of the STFT's). In the smoothed pseudo-Wigner distribution, we assumed the same width in the $\theta$ (and $\tau$) directions as in the S-method $\beta(\theta_{\max}/10)^2/2 = 1$. The auto-term widths (Table I) are defined following the same logic as in case of kernels widths. Some additional details on the auto-term forms and their derivations may be found in [23].

It is evident from this analysis (which is summarized in Table I and in Figs. 2 and 3) that when comparing various reduce interference distributions, not only a cross-term reduction but the auto-term shape, as well, should be taken to be an important comparison parameter. The results presented in Table I and Figs. 2 and 3 are, in our opinion, so apparent that there is no need for their further description or discussion.

## V. A Procedure for RID Kernel Design

On the basis of (8), one may construct a distribution with the desired auto-term shape in the following way: If $\mathbf{C}(\omega)$ is a given auto-term function for the linear frequency modulated signal, then the product kernel $c(\theta, \tau) = c(\theta\tau)$, that will produce this auto-term form is defined by

$$\begin{gathered} c(\theta\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{C}(\omega) e^{j\omega g} d\omega \\ \text{at} \quad -ag^2 = \theta\tau \quad \text{or} \quad |g| = \sqrt{|\theta\tau/a|} \end{gathered} \tag{9}$$

where $a$ is the instantaneous frequency coefficient $x(t) = e^{jat^2/2+bt}$.

*Example:* Let us determine the kernel function that will, for $a = 1$, produce the Hanning auto-term function $\mathbf{C}(\omega) = \frac{k}{2}[1 + \cos(\omega\pi)]$ for $|\omega| < 1$ and $\mathbf{C}(\omega) = 0$ elsewhere (coefficient $k$ will follow from the condition that $c(0) = 1$).

According to (9), we get

$$c(\theta, \tau) = c(\theta\tau) = \frac{\pi^2 \sin(\sqrt{|\theta\tau|})}{\sqrt{|\theta\tau|}(\pi^2 - |\theta\tau|)}. \tag{10}$$

This kernel decreases in the $(\theta, \tau)$ plane as $1/|\theta\tau|^{3/2}$. Thus, kernel (10) will have better cross-term reduction than the ones decreasing as $1/(\theta\tau)$, which were presented in Section IV.

In the analysis performed in Section III, we have seen that the width of $c(-\tau, \tau)$ (i.e., the width of $c(-a\tau, \tau)$) should be as small as possible in order to have high cross-term suppression. However, at the same time, the width of auto-term function $\mathbf{C}(\omega) = FT\{c(-a\tau, \tau)\}$ should be small (i.e., $c(-a\tau, \tau)$) wide) in order to produce a concentrated and sharp distribution in the time-frequency plane. Product $m$ of the width of $\mathbf{C}(\omega)$ (which is denoted by $D$) and the width of $c(-\tau, \tau)$ (which is denoted by $d$) is constant for a given kernel. It satisfies the uncertainty principle relation $Dd = m \succeq 1/2$,

TABLE I
AUTO-TERM FUNCTION IN THE REDUCED INTERFERENCE DISTRIBUTIONS. (*) IN THIS COLUMN, THE VALUES $\beta = \gamma = 1$, $\alpha/2 = \sigma = \theta_1\tau_1 = 2\pi$, AS WELL AS THE HANNING WINDOW $w^2(\tau/2)$ OF THE WIDTH $T = 28.2$, ARE ASSUMED. FUNCTIONS ARE FOR $a > 0$; IF $a < 0$, THEN THE SAME FORMS ARE VALID WITH $|a|$

| Distribution | Kernel $c(\theta,\tau)$ | Auto-term function $\mathbf{C}(\omega) = FT\{c(-a\tau,\tau)\}$ | Auto-term width | Auto-term width (*) | Delta pulse signal |
|---|---|---|---|---|---|
| Pseudo Wigner distribution | $w^2(\tau/2)$ | $FT\{w^2(\tau/2)\} = \mathbf{W}(\omega)$ | $8\pi/T$ | 0.891 | $\delta(t)$ |
| Sinc distribution | $rect(\theta\tau/\alpha)$ | $2\frac{\sin(\omega\sqrt{\alpha/(2a)}}{\omega}$ | $2\pi\sqrt{2a/\alpha}$ | $2.51\sqrt{a}$ | $\delta(t)$ |
| Choi-Williams distribution | $e^{-\theta^2\tau^2/\sigma^2}$ | $\int_{-\infty}^{\infty} e^{-a^2\tau^4/\sigma^2} e^{-j\omega\tau} d\tau$ | $2.8\sqrt{2\pi a/\sigma}$ | $2.8\sqrt{a}$ | $\delta(t)$ |
| ZAM distribution | $\lvert\tau\rvert \frac{\sin(\theta\tau/2)}{\theta\tau/2} w(\tau)$ | $\frac{2\pi}{a}\left[\frac{1}{2} - \left\lvert K(\frac{\omega^2}{2a})\right\rvert^2\right] w(\frac{\omega}{a})$ | $2.56\sqrt{a}$ | $2.56\sqrt{a}$ | 0 |
| Born-Jordan distribution | $\frac{\sin(\theta\tau/2)}{\theta\tau/2}$ | $\frac{2\pi}{a}\omega\left[S(\frac{\omega^2}{2a}) - C(\frac{\omega^2}{2a})\right] + +\sqrt{\frac{8\pi}{a}}\sin(\frac{\omega^2}{2a} + \frac{\pi}{4})$ | $3.3\sqrt{a}$ | $3.3\sqrt{a}$ | $\delta(t)$ |
| Zhang-Sato distribution | $e^{-\theta^2\tau^2/\sigma^2}\cos\frac{\theta\tau}{2}$ | $\sqrt{\frac{2\pi}{a}} e^{-\frac{\omega^4}{a^2\sigma^2}} \cos(\frac{\omega^2}{2a} - \frac{\pi}{4})$ | $4.3\sqrt{a}$ | $4.3\sqrt{a}$ | $\delta(t)$ |
| Butterworth distribution | $\frac{1}{1+\left(\frac{\theta\tau}{\theta_1\tau_1}\right)^{2N}}$ | $2\frac{\sin(\omega\sqrt{\theta_1\tau_1/a}}{\omega}$ for large $4N$ | $6.24\sqrt{a/(\theta_1\tau_1)}$ | $2.49\sqrt{a}$ | $\delta(t)$ |
| S-method | $\frac{1}{2\pi}P(\frac{-\theta}{2}) *_\theta A_{ww}(\theta,\tau)$ | $\mathbf{W}(\omega)$ | $8\pi/T$ | 0.891 | $2p(2t)w^2(t)$ |
| Smoothed pseudo Wigner distribution Gaussian | $e^{-\beta\theta^2/2-\gamma\tau^2/2}$ | $\sqrt{\frac{2\pi}{a^2\beta+\gamma}} e^{-\frac{\omega^2}{2(a^2\beta+\gamma)}}$ | $2.82\sqrt{a^2\beta+\gamma}$ | $2.82\sqrt{a^2+1}$ | $\sqrt{\frac{2\pi}{\beta}} e^{-\frac{t^2}{2\beta}}$ |
| Spectrogram | $A_{ww}(\theta,\tau)$ | $\frac{2\pi}{a} w^2(\omega/a)$ | $aT$ | $28.2a$ | $w^2(t)$ |
| Optimal kernel distribution | $e^{-\lvert\theta\tau\rvert/\sigma}$ | $\sqrt{\frac{\pi\sigma}{a}} e^{-\sigma\omega^2/(4a)}$ | $1.6\sqrt{2\pi a/\sigma}$ | $1.6\sqrt{a}$ | $\delta(t)$ |

[18], [22]. Thus, if one fixes the value of $D$ (which is the auto-term width), then the remaining value $d$ (being the measure of cross-terms suppression) will be minimal if $Dd$ is minimal, i.e., equal to $1/2$. The same is valid if one fixes $d$. A kernel defined by $Dd = 1/2$ (optimal in the described way) is presented by the following theorem.

*Theorem:* The product kernel $c(\theta,\tau) = c(\theta\tau)$, which has the property that the product of its width $d$ along line $\theta = -a\tau$ and width $D$ of auto-term function $C(\omega)$, is minimal, is defined by

$$c(\theta\tau) = e^{-|\theta\tau|/\sigma}. \tag{11}$$

*Proof:* This kernel directly follows from Gabor's uncertainty relation [18], [22]:

$$D^2 d^2 = \left(\frac{1}{2\pi E}\int_{-\infty}^{\infty} \omega^2 |\mathbf{C}(\omega)|^2 d\omega\right)\left(\frac{1}{E}\int_{-\infty}^{\infty} \tau^2 |c(-\tau,\tau)|^2 d\tau\right) \succeq 1/4.$$

The minimal value $1/4$ is achieved for $c'(-\tau,\tau) = k\tau c(-\tau,\tau)$. The finite energy solution of the above differential equation is the normal function $c(-\tau,\tau) = Be^{-k\tau^2}$. It follows that the optimal kernel with respect to this criterion is

$$c(-\tau,\tau) = c(-\tau^2) = e^{-\tau^2/\sigma}$$

where the values $B = 1$ (unbiased energy condition), as well as $k = 1/\sigma$, are taken. The kernel function, which corresponds to this equation, according to (9), is defined by (11). Q.E.D.

The kernel defined by (11) could be treated as the one belonging to the generalized Choi–Williams distributions [2] (in a wide sense) with $N = 1/2$. The auto-term function is of the form

$$\mathbf{C}(\omega) = \sqrt{\frac{\pi\sigma}{a}} e^{-\sigma\omega^2/(4|a|)}. \tag{12}$$

It is positive for all $a$ and $\sigma$. Assuming $\sigma = 2\pi$ and $a = 1$, as in Section IV, we get that the auto-term function is attenuated $e$ times at $\omega = 0.80$. For other $a$ and $\sigma$, the auto-term width is $W = 1.6\sqrt{2\pi|a|/\sigma}$.

The optimal kernel distribution (11) in the case of multicomponent signal $x(t) = e^{-t^2/4}\cos(6t)$ at instant $t = 0$ when the cross-term assumes its maximal value is shown in Fig. 4 along with the Wigner, Sinc, and ZAM distributions. Note that the S-method would produce the same auto-terms as in Figs. 2(g) and 3(g), without cross-term [12], [13], [15]–[17].

The distribution described by (11) satisfies the marginal properties since $c(0) = 1$. It satisfies the following conditions as well: realness, time and frequency shift, time-frequency scaling, etc., but there is a question: Are the unbiased instantaneous frequency and group delay properties satisfied? The mean instantaneous frequency of a

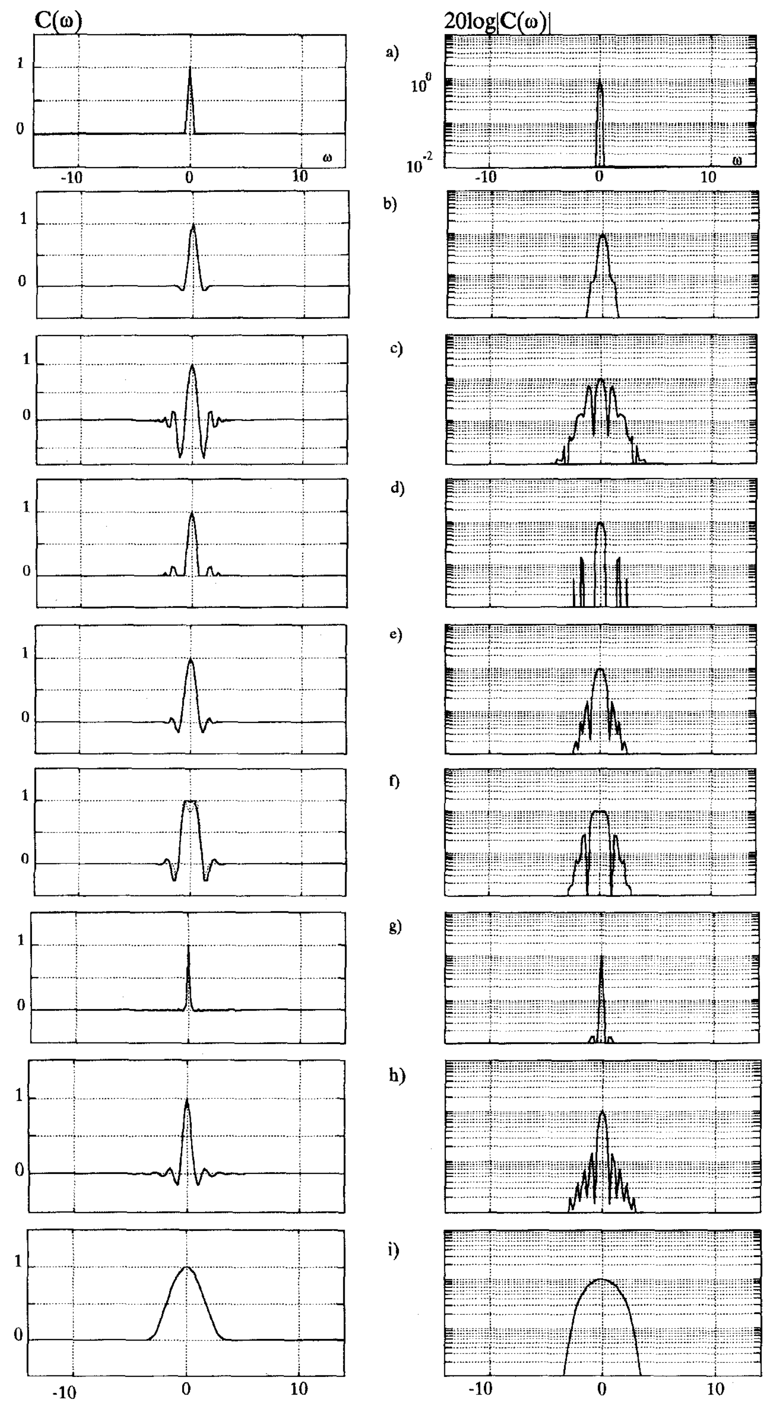


Fig. 2. Normalized auto-term function for $a = 0.25$ : (a) Pseudo Wigner distribution; (b) Choi–Williams distribution; (c) ZAM distribution; (d) nonnegative ZAM distribution; (e) Born–Jordan distribution; (f) ZS distribution; (g) S-method; (h) Butterworth distribution $N = 4$; (i) Spectrogram. The Hanning window is used. In all examples, the region $\tau \in [-14.1, 14.1]$, $\theta \in [-14.1, 14.1]$ is considered. Note that the same figures will be obtained if the time and frequency axis are scaled by $Q$, i.e., $\tau \rightarrow \tau/Q$, $\theta \rightarrow \theta Q$ and $a \rightarrow aQ^2$, where $Q$ is any real number not equal to 0.

distribution belonging to the Cohen class is [1], [8]

$$\langle\omega\rangle = \phi'(t) + 2\frac{A'(t)}{A(t)}\frac{c'(0)}{c(0)}.$$

It follows that $\langle\omega\rangle = \phi'(t)$ if $c'(0) = 0$ or $c'(0)$ is a constant, and $A'(t)/A(t) \rightarrow 0$. In other words, if the amplitude variations are negligible compared with the phase variations (whose measure is $\phi'(t)$), then $\langle\omega\rangle = \phi'(t)$ even if $c'(0)$ is a nonzero constant. If the amplitude variations are of order of the phase variations, then there is a question: "What does $\phi'(t)$ represent in that case" or "Can it be considered to be the instantaneous frequency at all?" (For example, if we have a signal $x(t) = A_1 e^{j\phi_1(t)} + A_2 e^{j\phi_2(t)} = A(t)e^{j\phi(t)}$, where the amplitude variation $A'(t)$ is of order of the phase variation $\phi'(t)$, may we treat the value of $\phi'(t)$ as the instantaneous frequency at all

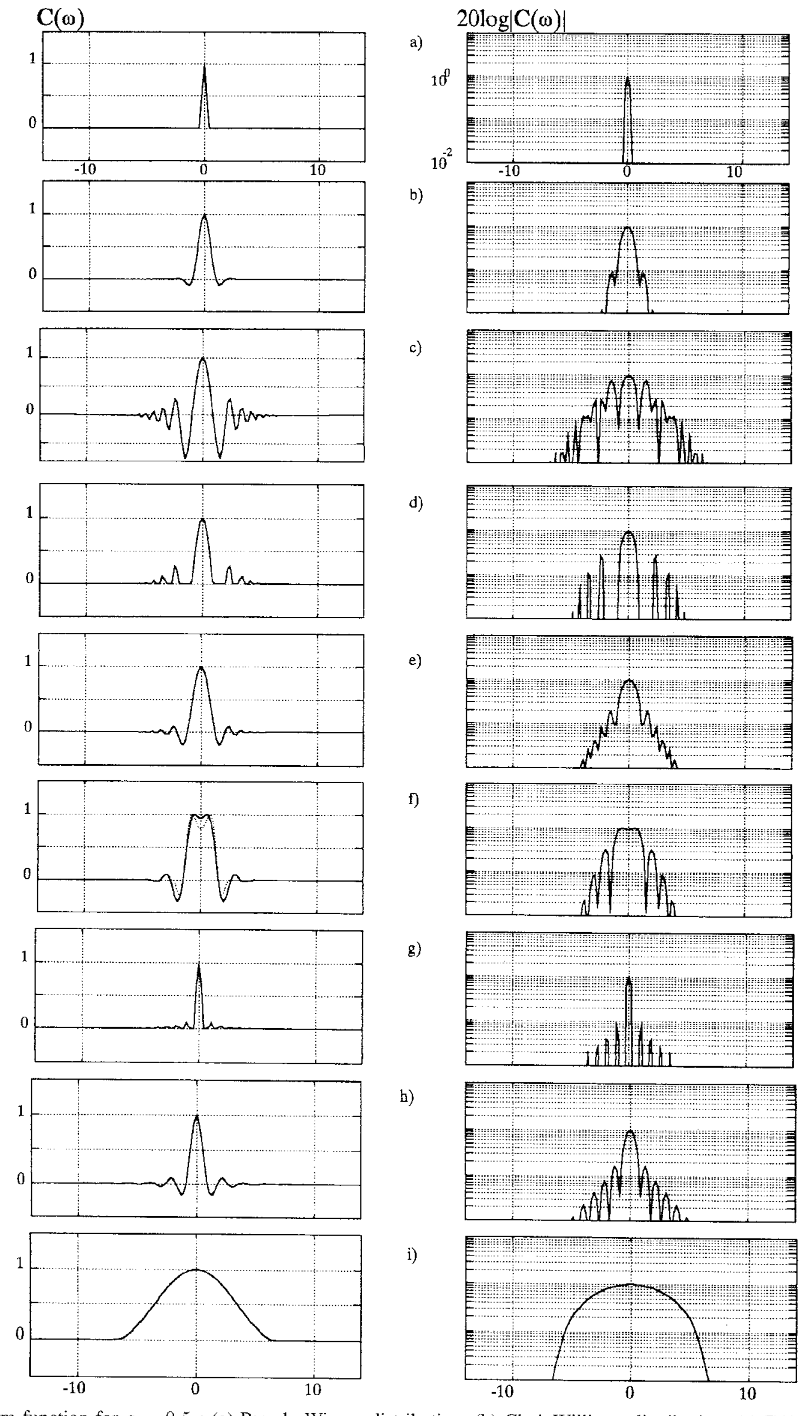


Fig. 3. Normalized auto-term function for $a = 0.5$ : (a) Pseudo Wigner distribution; (b) Choi–Williams distribution; (c) ZAM distribution; (d) nonnegative ZAM distribution; (e) Born–Jordan distribution; (f) ZS distribution; (g) nonnegative S-method; (h) Butterworth distribution $N = 4$; (i) Spectrogram. The Hanning window is used. In all examples, the region $\tau \in [-14.1, 14.1]$, $\theta \in [-14.1, 14.1]$ is considered. Note that the same figures will be obtained if the time and frequency axis are scaled by $Q$, i.e., $\tau \to \tau/Q$, $\theta \to \theta Q$ and $a \to aQ^2$, where $Q$ is any real number not equal to 0.

[8], [14]?) If the unbiased instantaneous frequency and group delay conditions are to be satisfied anyway, then the optimal kernel may by modified:

$$c(\theta\tau) = e^{-|\theta\tau|/\sigma}(1 + \frac{|\theta\tau|}{\sigma}). \quad (13)$$

This kernel has the values $c(0) = 1$, $c'(0+) = c'(0-) = 0$. Therefore, it satisfies the marginal conditions.

In the end, we will mention that the minimization of $Dd$ is just one of possible criteria to define an optimal kernel. Other criteria will produce different optimal kernel functions.

## VI. Conclusion

The analysis of auto term representation in the RID class of distribution is presented. It is shown that various reduced interference

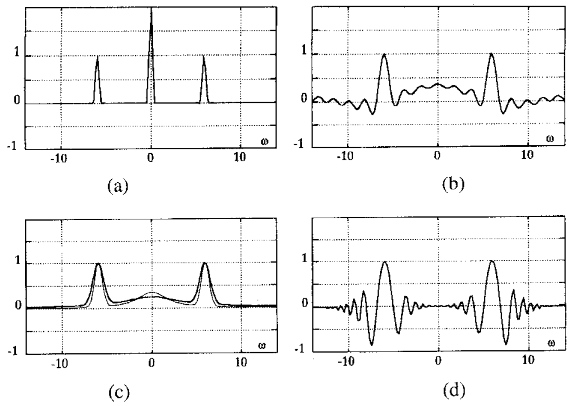

Fig. 4. Time-frequency representation of a multicomponent signal at instant $t = 0$, where the cross-term reaches its maximal value. Normalized values of (a) Pseudo Wigner distribution; (b) Sinc distribution ($\alpha/2 = 2\pi$); (c) optimal kernel distribution (thin line $\sigma = 2\pi$ and the thick line $\sigma = \pi$); (d) Zao–Atlas–Marks distribution.

distributions produces different auto-term functions. The procedure for a kernel design is given.

## APPENDIX

The stationary phase method states that if $|A'(t)| \ll |\phi'(t)|$, then [18]

$$X(\omega) = \int_{-\infty}^{\infty} A(t)e^{j\phi(t)}e^{-j\omega t}dt \cong e^{j\phi(t_0)}e^{-j\omega t_0}A(t_0)\sqrt{\frac{2\pi j}{\phi^{(2)}(t_0)}}$$

with $\phi'(t_0) = \omega$, and $\phi^{(2)}(t_0) \neq 0$.

If we have a product $A(t)w(t)$ that satisfies the stationary phase method condition, and if the instantaneous frequency is linear, i.e., $\phi'(t_0) = at_0$, then

$$X_w(\omega) = \int_{-\infty}^{\infty} A(t)w(t)e^{j\phi(t)}e^{-j\omega t}dt \cong X(\omega)w(t_0)$$

with $w(t_0) = w(\omega/a)$.

This relation may be applied in order to get the auto-term function in the spectrogram of signal (5):

$$\begin{aligned}|STFT(\omega,t)|^2 =& \left|\int_{-\infty}^{\infty} Ae^{ja(t+\tau)^2/2}w(\tau)e^{-j\omega\tau}d\tau\right|^2\\ \cong& \frac{2\pi A^2}{a}w^2\left(\frac{\omega - at}{a}\right)\end{aligned}$$

or in the Zao–Atlas–Marks distribution:

$$\begin{aligned}\mathbf{C}(\omega) =& \int_{-\infty}^{\infty} |\tau|\frac{\sin(a\tau^2/2)}{a\tau^2/2}w(\tau)e^{-j\omega\tau}d\tau\\ =& \frac{1}{a}\left[\frac{1}{2} - \left|K(\frac{\omega^2}{2a})\right|^2\right] *_\omega W(\omega)\\ \cong& \frac{2\pi}{a}\left[\frac{1}{2} - \left|K(\frac{\omega^2}{2a})\right|^2\right] w(\omega/a).\end{aligned}$$

The approximations is obtained for large $a\tau_{\max}^2/2$, practically meaning that inside window $w(\omega)$ ($|\tau| < \tau_{\max}$), there exist few quasisemiperiods of function $\sin(a\tau^2/2)$ , i.e., this function has few zeros inside $|\tau| < \tau_{\max}$). In the same way, we get the auto-terms in the ZS distribution. Agreement of these auto-terms with the numerically obtained ones (according to (8)) is almost complete.

## REFERENCES


[1] L. Cohen, "Time-frequency distribution; A review" *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, July 1989.

[2] F. Hlawatsch and G.F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Mag.*, pp. 21–67, Apr. 1992.

[3] S. Kadambe and G.F. Boudreaux-Bartels, "A comparison of the existence of 'cross terms' in the Wigner distribution and the squared magnitude of the Wavelet transform and Short-time Fourier transform," *IEEE Trans. Signal Processing*, vol. 40, no. 10, pp. 2498–2517, Oct. 1992.

[4] J. Jeong and W. J. Williams, "Mechanism of the cross-terms in spectrograms," *IEEE Trans. Signal Processing*, vol. 40, no. 10, pp. 2608–2613, Oct. 1992.

[5] H. Choi and W. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 862–871, June 1989.

[6] P. Flandrin, "Some features of time-frequency representation of multicomponent signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1984, pp. 41B.4.1.–4.1.

[7] F. Hlawatsch, "Interference terms in the Wigner distribution," in *Proc. Digital Signal Processing 84*, V. Cappellini *et al.*, Eds., 1984, pp. 363-367.

[8] L. Cohen and C. Lee, "Instantaneous bandwidth," in *Time-Frequency Signal Analysis*, B. Boashash Ed. Melbourne, Australia: Longman Cheshire, 1992.

[9] L. Cohen, "Distributions concentrated along the instantaneous frequency," *Adv. Signal Processing Algorithms, Archit. Implement.*, vol. 1348, pp. 149–157, 1992.

[10] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, no. 2, pp. 402–412, Feb. 1992.

[11] S. Oh and R. J. Marks, II, "Some properties of generalized time-frequency representation with cone-shaped kernels," *IEEE Trans. Signal Processing*, vol. 40, no. 7, pp. 1735–1745, July 1992.

[12] LJ. Stanković, "A method for time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 225–229, Jan. 1994.

[13] ——, "An analysis of some time-frequency and time-scale distributions," *Annales des Telecommunications*, no. 9/10, pp. 505–517, Sept./Oct. 1994.

[14] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal Part 1—Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 519–538, Apr. 1992.

[15] LJ. Stanković, "A method for improved energy concentration in the time-frequency signal analysis using the L-Wigner distribution," *IEEE Trans. Signal Processing*, vol. 43, no. 5, pp. 1262–1268, May 1995.

[16] ——, "A multi-time definition of the Wigner higher order distribution; L-Wigner distribution," *IEEE Signal Processing Lett.*, vol. 1, no. 7, pp. 106–109, July 1994.

[17] S. Stanković, LJ. Stanković, and Z. Uskoković, "On the local frequency, group shift and cross-terms in the multidimensional time-frequency distributions; A method for multidimensional time–frequency analysis," *IEEE Trans. Signal Processing*, vol. 43, no. 7, pp. 1719–1725, July 1995.

[18] A. Papoulis, *Signal Analysis*. New York: McGraw Hill, 1977.

[19] B. Ristic and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Trans. Signal Processing*, vol. 41, no. 5, pp. 1996–2008, May 1993.

[20] B. Zhang and S. Sato, "A time-frequency distribution of Cohen's class with a compound kernel and its application to speech signal processing," *IEEE Trans. Signal Processing*, vol. 42, no. 1, pp. 54-64, Jan. 1994.

[21] D. Wu and J. Morris, "Time-frequency representations using a radial Butterworth kernel," in *Proc. IEEE Symp. TFTSA*, Philadelphia, PA, Oct. 1994, pp. 60-63.

[22] D. Gabor, "Theory of communications," *J. Inst. Elec. Eng.*, vol. 93, Nov. 1946.

[23] LJ. Stanković, S. Stanković, and Z. Uskoković, Eds., "Time-frequency signal analysis," research monograph, Epsilon-Montenegropublic, 1994.