Kernel Design for Reduced Interference Distributions

Jechang Jeong, *Member, IEEE,* and William J. Williams, *Senior Member, IEEE*

Abstract—**The mixed time-frequency signal representation has received considerable attention as a powerful tool for analyzing a variety of signals. Typical examples of time-frequency energy distributions are the spectrogram and the Wigner distribution. Despite several interesting properties, however, both of them are known to have critical limitations. For the spectrogram, the time and frequency resolutions cannot be simultaneously optimized. The Wigner distribution exhibits negative values and interference terms, which may lead to misinterpretations regarding the signal spectral contents. To overcome these drawbacks, Choi and Williams introduced a new distribution that has an exponential-type kernel. Although retaining high resolution with suppressed interference terms, the exponential distribution (ED, or Choi–Williams distribution) does not completely satisfy the support properties in time and frequency.**

In this paper, based on desirable distribution properties and associated kernel requirements, we discuss and further define a new class of time-frequency distributions called the reduced interference distribution (RID). Like the ED, the RID suppresses cross terms effectively while retaining high time and frequency resolutions of autoterms. Furthermore, the RID meets most of the desirable kernel requirements, including the time/frequency support properties. A systematic procedure to create RID kernels (or, equivalently, to compute RID's) is proposed. Some aspects and properties of the RID are discussed. It is shown that interpretations in the ambiguity, temporal correlation, spectral correlation, and time-frequency domains are fruitful for conceptualizing the RID. Some experimental results are provided to demonstrate the performance of the RID.

I. Introduction

RECENTLY, the mixed (or joint) time-frequency signal representation has received considerable attention as a powerful tool for analyzing a variety of signals and systems. In particular, if the frequency content is time varying as in nonstationary signals, then this approach is quite attractive. Although either the time domain description, $f(t)$, or its Fourier transform (FT), $F(\omega)$, carries complete information about the signal (assuming one deals with deterministic signals), none of them reveals explicitly the frequency spectrum at a particular time or the time at which a particular frequency component occurs. By mapping a one-dimensional function of time (or frequency) into a two-dimensional function of time and frequency, the mixed time-frequency representation (TFR) localizes the signal energy in both the time and frequency directions.

Traditionally the short-time Fourier transform (STFT)—or its modulus squared, called spectrogram—has played an important role in visualizing the time-varying frequency content of various signals such as human speech [1], [29], [31], [32]. A major drawback inherent in the STFT is that a tradeoff is inevitable between temporal and spectral resolutions [29]. If one uses a longer sliding time window to obtain higher spectral resolution, then the underlying nonstationarity will be smeared out, resulting in lower temporal resolution. Conversely, using a shorter window to achieve better temporal resolution will give lower spectral resolution.

The Wigner distribution (WD) has been employed as an alternative to overcome this shortcoming of the STFT. The WD was first introduced in the context of quantum mechanics [35] and revived for signal analysis by Ville [34]. The WD possesses very high resolution in both time and frequency, and it has many other nice properties as well [8]. Two major drawbacks of the WD are that it is not necessarily nonnegative and its bilinearity produces cross terms (or interferences) between two signal components located at different regions in the time-frequency plane. An excellent discussion on the geometry of interferences has been provided by Hlawatsch and Flandrin [14], [15], [20].

In his pioneering work, Cohen incorporated many TFR's into a general class of time-frequency distributions of TFR's. In his recent work, Cohen has comprehensively discussed many aspects of time-frequency distributions and recent progress in the area [10]. Boashash has compared the performance of several time-frequency distributions in terms of resolution [5]. Desirable properties of a distribution and associated kernel requirements have been extensively investigated by Claasen and Mecklenbrauker [9].

Recently, Choi and Williams introduced a new distribution with an exponential-type kernel [7], which they called the ED. Since most investigators now call this the Choi–Williams distribution, we will defer to this nomenclature for uniformity and clarity. The Choi–Williams distribution overcomes the aforementioned drawbacks of the STFT and the WD, and provides high resolution in time and frequency while suppressing interferences. Unfortunately, however, the Choi–Williams distribution does not completely satisfy the support properties in time and frequency.

In this paper we present a new class of TFR's which

Manuscript received October 23, 1989; revised September 28, 1990. This work was supported in part by grants from the Rackham School of Graduate Studies and the Office of Naval Research, ONR Contracts N00014-89-J-1723 and N00014-90-J-1654.

The authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109.

IEEE Log Number 9104877.

has many nice properties. The rest of this paper is organized as follows. In Section II, we briefly review commonly used TFR's and introduce desirable distribution properties. Particular attention is paid to the interpretation of Cohen's class in the ambiguity, temporal correlation, spectral correlation, and time-frequency domains. In Section III, based on the desirable kernel requirements, we discuss and further define a new class of time-frequency distributions called the reduced interference distribution (RID) [36], [37]. A systematic procedure to create RID kernels (or, equivalently, compute RID's) is proposed. Some aspects and properties of the RID are discussed. In Section IV, we consider design considerations for RID's and compare various selections of the "primitive" window. Section V provides some experimental results demonstrating the performance of the RID.

II. Time-Frequency Representations: Background and Notations

Let $R_f(t, \tau)$ be the instantaneous autocorrelation of a complex signal $f(t)$, defined as

$$R_f(t, \tau) = f(t + \tau/2)f^*(t - \tau/2) \tag{1}$$

where f^* denotes the complex conjugate of f. The Wigner distribution of $f(t)$ is then defined as the Fourier transform (FT) of $R_f(t, \tau)$ with respect to the lag variable τ [35]:

$$W_f(t, \omega) = \int R_f(t, \tau)e^{-j\omega\tau}\, d\tau. \tag{2}$$

(Unless otherwise indicated, the ranges of integrals are from $-\infty$ to ∞ throughout this paper.) Analogous to the WD, but with a different physical meaning, the symmetrical ambiguity function (AF) is defined as the inverse Fourier transform (IFT) of $R_f(t, \tau)$ with respect to the time variable t [39]:

$$A_f(\theta, \tau) = \frac{1}{2\pi}\int R_f(t, \tau)e^{j\theta t}\, dt. \tag{3}$$

Thus, the WD and AF are related by the two-dimensional (2-D) FT:

$$W_f(t, \omega) = \iint A_f(\theta, \tau)e^{-j(t\theta + \omega\tau)}\, d\theta\, d\tau. \tag{4}$$

Let $F(\omega)$ be the FT of $f(t)$, and let $R_F(\omega, \theta)$ be the instantaneous spectral autocorrelation defined as

$$R_F(\omega, \theta) = F(\omega + \theta/2)F^*(\omega - \theta/2). \tag{5}$$

Then it follows that

$$R_F(\omega, \theta) = \iint R_f(t, \tau)e^{-j(\omega\tau + \theta t)}\, dt\, d\tau \tag{6}$$

$$W_F(\omega, t) = \int R_F(\omega, \theta)e^{-j\theta t}\, d\theta = 2\pi W_f(-t, \omega) \tag{7}$$

$$A_F(\tau, \theta) = \frac{1}{2\pi}\int R_F(\omega, \theta)e^{j\omega\tau}\, d\omega = 2\pi A_f(\tau, -\theta). \tag{8}$$

Fig. 1(a) contains a block diagram relating $R_f(t, \tau)$, $A_f(\theta, \tau)$, $W_f(t, \omega)$, $R_F(\omega, \theta)$, $A_F(\tau, \theta)$, and $W_F(\omega, t)$. One can see in Fig. 1(a) that there are basically six domains, all related to one another: the time (t), frequency (ω), temporal correlation (t, τ), spectral correlation (ω, θ), ambiguity (θ, τ), and time-frequency (t, ω) domains. Each domain possesses unique features and is suitable for representing certain types of signals. Note that the procedure $f(t) \rightarrow R_f(t, \tau)$ is irreversible in the sense that $f(t)$ cannot be exactly recovered from $R_f(t, \tau)$. Since $f(t)$ and $f(t)e^{j\alpha}$ produce the identical instantaneous autocorrelation, the reconstructed time waveform can be different from the original signal by a constant phase. Similarly, the signal can be restored from the WD or the AF up to a constant phase [6].

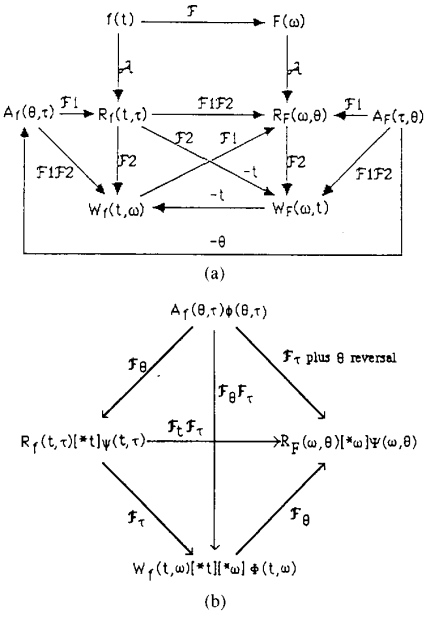

Fig. 1 (a) Relationships between the time function, its Fourier transform, instantaneous autocorrelation, spectral autocorrelation, ambiguity function, and Wigner distribution. Variable descriptions are t: time, τ: time lag, ω: frequency, θ: frequency lag. Operator descriptions are $\mathcal{F}$: Fourier transform, $\mathcal{F}_1$: Fourier transform w.r.t. the first variable, $\mathcal{F}_2$: Fourier transform w.r.t. the second variable, $\mathcal{A}$: autocorrelation. (b) Cohen's class in four different domains.

The WD possesses very high resolution in both time and frequency, and it has many other nice properties as well [8], [22]. Through a series of papers published in the early 1980's [8], [9], Claasen and Mecklenbrauker rekindled interest in this area. Their discussion on properties and aspects of the WD provided valuable guidance in the usage of the WD.

Two major drawbacks of the WD are that it is not necessarily nonnegative and it produces cross terms (i.e., interferences) between signal components located in different regions in the time-frequency plane. These two phenomena may cause difficulties in interpreting the WD as an energy distribution over time and frequency. Sometimes the nonnegativity and cross terms are closely related, since the cross terms in the time-frequency domain are, in general, fluctuating and produce a plethora of ne-

gativity. Thus, in many cases, suppression of cross terms accompanies reduction of negative values in magnitude.

Remark: Although it is desirable to suppress the cross terms to obtain better interpretability regarding the time-varying spectrum, the cross terms carry important information regarding the relationship between signal components. For example, the cross-terms, after being converted into the ambiguity domain form, provide information regarding the moving target in radar/sonar applications.

In order to suppress the possible negative values and interferences in the WD, smoothing in both the time and frequency directions is often carried out:

$$W_f'(t, \omega) = \frac{1}{4\pi^2} \int\int W_f(u, \xi)\Phi(t - u, \omega - \xi)\, du\, d\xi \tag{9}$$

where $\Phi(t, \omega)$ is a 2-D smoothing function. By the 2-D convolution theorem, we have

$$\begin{aligned} W_f'(t, \omega) &= \int\int A_f(\theta, \tau)\phi(\theta, \tau) e^{-j(\theta t + \tau\omega)}\, d\tau\, d\theta \qquad (10)\\ &= \frac{1}{2\pi} \int\int\int e^{j(-\theta t - \tau\omega + \theta u)}\, \phi(\theta, \tau) f(u + \tau/2) \\ &\quad \cdot f^*(u - \tau 2)\, du\, d\tau\, d\theta \\ &\triangleq C_f(t, \omega; \phi) \qquad (11) \end{aligned}$$

where $\phi(\theta, \tau)$ is the 2-D IFT of $\Phi(t, \omega)$. The collection of $C_f(t, \omega; \phi)$ is referred to as Cohen's class of time-frequency distributions with kernel $\phi(\theta, \tau)$ [11]. In general, the kernel $\phi(\theta, \tau)$ can depend on t, ω, and the signal $f(t)$ [10]. In this work, however, we consider shift-invariant cases, in which the kernels do not depend on t or ω (see Table I). In (11), different kernels produce different distributions. For example, $\phi(\theta, \tau) = 1$, $e^{j\theta\tau/2}$, $e^{j\theta|\tau|/2}$, and $e^{-\theta^2\tau^2/\sigma}$ correspond to the Wigner, Rihaczek, Page, and Choi-Williams distributions, respectively [10].

$C_f(t, \omega; \phi)$ can also be expressed as the FT of the generalized autocorrelation function [7]:

$$C_f(t, \omega; \phi) = \int R_f'(t, \tau) e^{-j\omega\tau}\, d\tau \tag{12}$$

where the generalized autocorrelation is defined as

$$R_f'(t, \tau) = \frac{1}{2\pi} \int R_f(u, \tau)\psi(t - u, \tau)\, du \tag{13}$$

with the autocorrelation domain kernel

$$\psi(t, \tau) = \int \phi(\theta, \tau) e^{-j\theta t}\, d\theta. \tag{14}$$

$C_f(t, \omega; \phi)$ can also be obtained from the IFT of the generalized spectral correlation $R_F'(\omega, \theta)$:

$$C_f(t, \omega; \phi) = \frac{1}{2\pi} \int R_F'(\omega, \theta) e^{j\theta t}\, d\theta \tag{15}$$

TABLE I
DISTRIBUTION PROPERTIES AND ASSOCIATED KERNEL REQUIREMENTS

P0. nonnegativity: $C_f(t, \omega; \phi) \geq 0\ \forall t, \omega$.
Q0. $\phi(\theta, \tau)$ is the ambiguity function of some function $w(t)$.
P1. realness: $C_f(t, \omega; \phi) \in R$.
Q1. $\phi(\theta, \tau) = \phi^*(-\theta, -\tau)$.
P2. time shift: $g(t) = f(t - t_o) \Rightarrow C_g(t, \omega; \phi) = C_f(t - t_o, \omega; \phi)$
Q2. $\phi(\theta, \tau)$ does not depend on t.
P3. frequency shift: $g(t) = f(t)e^{j\omega_0 t} \Rightarrow C_g(t, \omega; \phi) = C_f(t, \omega - \omega_0; \phi)$.
Q3. $\phi(\theta, \tau)$ does not depend on ω.
P4. time marginal: $1/2\pi \int W_f(t, \omega)\, d\omega = f(t) f^*(t)$.
Q4. $\phi(\theta, 0) = 1\ \forall\theta$.
P5. frequency marginal: $\int C_f(t, \omega; \phi)\, dt = F(\omega) F^*(\omega)$.
Q5. $\phi(0, \tau) = 1\ \forall\tau$.
P6. instantaneous frequency: $[\int \omega C_f(t,\omega; \phi)\, d\omega / \int C_f(t, \omega; \phi)\, d\omega] = \omega_i(t)$.
Q6. Q4 and $[\partial\phi(\theta, \tau)/\partial\tau]|_{\tau=0} = 0\ \forall\theta$.
P7. group delay: $[\int t C_f(t, \omega; \phi)\, dt / \int C_f(t, \omega; \phi)\, dt] = t_g(\omega)$.
Q7. Q5 and $[\partial\phi(\theta, \tau)/\partial\theta]|_{\theta=0} = 0\ \forall\tau$.
P8. time support: $f(t) = 0$ for $|t| > t_c \Rightarrow C_f(t, \omega; \phi) = 0$ for $|t| > t_c$.
Q8. $\psi(t, \tau) \triangleq \int \phi(\theta, \tau) e^{-j\theta t}\, d\theta = 0$ for $|\tau| < 2|t|$.
P9. frequency support: $F(\omega) = 0$ for $|\omega| > \omega_c \Rightarrow C_f(t, \omega; \phi) = 0$ for $|\omega| > \omega_c$.
Q9. $\int \phi(\theta, \tau) e^{j\omega\tau}\, d\tau = 0$ for $|\theta| < 2|\omega|$.
P10. Reduced interference.
Q10. $\phi(\theta, \tau)$ is a 2-D low-pass filter type.

where

$$R_F'(\omega, \theta) = \frac{1}{2\pi} \int R_F(\xi, \theta)\Psi(\omega - \xi, \theta)\, d\xi \tag{16}$$

with the spectral correlation domain kernel

$$\Psi(\omega, \theta) = \int \phi(-\theta, \tau) e^{-j\tau\omega}\, d\tau. \tag{17}$$

Thus the generalized time-frequency distribution can be viewed in four ways [9], [21]:

1) The 2-D convolution of the WD with the time-frequency domain kernel.

2) The 2-D FT of the product between the ambiguity function and the kernel. This product is referred to as the generalized ambiguity function [12].

3) The 1-D FT (with respect to the lag variable) of the convolution (in the time direction) between the instantaneous autocorrelation and the temporal correlation domain kernel.

4) The 1-D IFT (with respect to the Doppler-shift variable) of the convolution (in the frequency direction) between the instantaneous spectral autocorrelation and the spectral correlation domain kernel.

These interpretations are very important when one wishes to design a kernel or implement a time-frequency distribution. In general, the first two relationships are useful for the kernel design, while the last two are often adopted for the practical implementation. Fig. 1(b) shows the relationships among these four interpretations. We will rely on these four interpretations throughout this paper.

Desirable distribution properties and associated kernel requirements (sufficient conditions) are summarized in Table I. P0-P9 are adapted from the work of Claasen and Mecklenbrauker [8], while the reduced interference property P10 has been introduced by Williams and Jeong [36],

[37] based on [7]. Instead of P10, one may wish to take into account other properties, such as regularity (related to signal recovery) and unitarity (related to Moyal's formula) [9], [21]. In this paper, however, we will stress the importance of the interference reduction property and we attempt to combine this property with P1–P9. Unfortunately, it is known that no kernel satisfies all of the requirements simultaneously [8]. In particular, P0 (nonnegativity) often conflicts with some other desirable properties, as is the case with the spectrogram. No TFR behaves "nicely" for all purposes. Each TFR, with its own advantages and disadvantages, is appropriate only for certain types of signals and applications. Therefore, the choice of a TFR relies on signal characteristics and applications [2], [3], [5].

The STFT of $f(t)$, using a sliding window $h(t)$, is defined as

$$T_f(t, \omega) = \int f(u) h^*(t - u) e^{-j\omega u} \, du \tag{18}$$

$$= \frac{1}{2\pi} \int F(\xi) H^*(\omega - \xi) e^{j(\xi - \omega)t} \, d\xi \tag{19}$$

where $F(\omega)$ and $H(\omega)$ are the FT's of $f(t)$ and $h(t)$, respectively. Conceptually, the STFT is based on a series of Fourier transforms applied to the signal segment chosen by the sliding window at each instant. The window length determines the tradeoff between time and frequency resolutions. Characteristics of several windows have been extensively investigated by Harris [18] and Nuttall [30]. The STFT can be viewed as a generalized version of Gabor's signal expression [4], [16], [19]. Properties of the STFT have been extensively investigated in [29].

The spectrogram is the modulus squared of the STFT:

$$S_f(t, \omega) = |T_f(t, \omega)|^2. \tag{20}$$

It can be easily seen that the spectrogram is a member of Cohen's class [9], [21], with the kernels in (t, τ), (θ, τ), (ω, θ), and (t, ω) domains being

$$\psi_s(t, \tau) = 2\pi R_h(t, \tau) \tag{21}$$

$$\phi_s(\theta, \tau) = 2\pi A_h(\theta, \tau) \tag{22}$$

$$\Psi_s(\omega, \theta) = R_H(\omega, \theta) \tag{23}$$

$$\Phi_s(t, \omega) = 2\pi W_h(t, \omega) \tag{24}$$

respectively. Thus the spectrogram is a 2-D convolution of two WD's in the time-frequency domain (or, equivalently, a 2-D multiplication of two AF's in the ambiguity domain) [9]:

$$S_f(t, \omega) = \frac{1}{2\pi} \iint W_f(u, \xi) W_h(t - u, \omega - \xi) \, du \, d\xi \tag{25}$$

$$= 2\pi \iint A_f(\theta, \tau) A_h(\theta, \tau) e^{-j(\theta t + \tau\omega)} \, d\tau \, d\theta. \tag{26}$$

Altes has derived many interesting properties of the spectrogram [1].

Remark: The STFT is a linear signal decomposition and there are no cross terms between signal components. However, the spectrogram is a bilinear signal energy distribution due to the magnitude squaring operation (the integral over the time-frequency plane gives the signal energy if $\|h(t)\|^2 = 1$). Thus, like the WD, the spectrogram has cross terms. In other words, the cross terms which do not exist in the STFT may appear visible in the spectrogram through the magnitude-squaring process. In many applications, the cross terms in the spectrogram are not conspicuous simply because they are inherently filtered out by a low-pass filter defined by the AF of the window [14], [23].

The spectrogram is nonnegative everywhere and this is a desirable property especially when one wishes to interpret the spectrogram as the signal energy distribution in the time-frequency plane. It does not, however, meet other useful properties such as P4–P9 in Table I. Moreover, it requires a tradeoff between time and frequency resolutions [29], [32].

Another class of TFR's, which has attracted a considerable attention over the past five years, is the wavelet transform (WT). The WT decomposes the signal into a family of functions that are "dilate-and-translated" versions of an analyzing wavelet function [17]. The WT is basically a time-scale (or time-dilation) representation, but it can also be regarded as a TFR since the dilation parameter is related to frequency in an inversely proportional manner. Daubechies has provided an excellent discussion on many aspects of the WT compared to the STFT [13]. In contrast to the fixed time-frequency resolution of the STFT, the WT provides good spectral and poor temporal resolutions at low frequencies, and good temporal and poor spectral resolutions at high frequencies. This feature makes the WT especially suitable for analyzing a variety of acoustic signals and images [27], [28].

The limitations of the WD and the spectrogram described motivate new time-frequency representations that overcome the drawbacks. The Choi–Williams distribution is one such attempt [7]. It has an exponential-type kernel

$$\phi(\theta, \tau) = e^{-\theta^2 \tau^2 / \sigma} \tag{27}$$

where the parameter σ trades off autoterm resolution for cross-term suppression or vice versa. The Choi–Williams distribution proves to be quite effective in suppressing the interferences while retaining high resolution. Its performance has been compared with those of the spectrogram and the WD in a variety of environments [36], [37]. Unfortunately, however, in a strict sense, this distribution violates the support properties P8 and P9 in Table I. Nontheless, the performance of the Choi–Williams distribution has pointed the way to the design of desirable time-frequency distributions and the ambiguity domain interpretation of its performance suggests a new methodology for obtaining such distributions.

III. A New Class of Time-Frequency Distributions

Incorporating the idea of interference reduction in the Choi-Williams distribution (P10 in Table I) with other desirable properties (P0–P9 in Table I), a new class of time-frequency distributions, called reduced interference distributions, has been introduced [36], [37]. While not satisfying P0, the RID does satisfy P1–P10 and provides high resolution in time and frequency. (The spectrogram, on the other hand, satisfies P0, but does not satisfy other nice properties). To meet Q1–Q10, the RID kernel should be a cross-shaped low-pass filter, satisfying [36], [37]

$$|\phi(\theta, \tau)| \ll 1 \qquad \text{for } |\theta\tau| \gg 0. \tag{28}$$

Exploiting several dual properties in Table I, we propose the following procedure to design a RID kernel.

Step 1: Design a primitive real-valued function $h(t)$ that satisfies the following:

R1: $h(t)$ has unit area, i.e., $\int h(t)\,dt = 1$.
R2: $h(t)$ is a symmetrical function of time, i.e., $h(-t) = h(t)$.
R3: $h(t)$ is time-limited on $[-1/2, 1/2]$, i.e., $h(t) = 0$ for $|t| > 1/2$.
R4: $h(t)$ tapers smoothly towards both ends so that its frequency response has little high frequency content. That is, $|H(\theta)| \ll 1$ for $|\theta| \gg 0$, where $H(\theta)$ is the FT of $h(t)$.

Step 2: Take the FT of $h(t)$, i.e.,

$$H(\theta) = \int h(t) e^{-j\theta t}\,dt. \tag{29}$$

Step 3: Replace θ by $\theta\tau$ in $H(\theta)$:

$$\phi(\theta, \tau) = H(\theta\tau). \tag{30}$$

We discuss each step in more detail in the following.

In step 1, R1–R3 are related to Q1–Q9 in Table I.

Since $\phi(\theta, \tau)$ is a function of the product of θ and τ, it is straightforward to show that R1 implies Q4 and Q5. Thus, the marginals are preserved in time and frequency.

R2 produces a real $H(\theta)$, which in turn implies Q1. R2 also implies Q6 and Q7 under the condition that $dH(\theta)/d\theta$ exists.

The implication of R3 is a little more complicated, but it does imply Q8 and Q9. Using the scaling property of the FT, Q8 is satisfied as follows:

$$\psi(t, \tau) = \int \phi(\theta, \tau) e^{-j\theta t}\,d\theta \tag{31}$$

$$= \frac{2\pi}{|\tau|} h\left(\frac{-t}{\tau}\right) \tag{32}$$

$$= 0 \text{ if } |\tau| < 2|t|. \tag{33}$$

By symmetry, Q9 is also satisfied.

Q2 and Q3 are automatically satisfied in the above design procedure since ϕ does not depend on t or ω.

R4 plays the role of suppressing interference (P10). In most cases, the autoterms are located near the origin in the ambiguity domain, while the cross terms occur far away from the origin [14], [20]. Therefore, a low-pass filter-type kernel imposed by R4 can effectively reduce the interferences while retaining the resolution of the autoterms. It can be easily seen that R4 is equivalent to the requirement Q10. The requirements R1–R4 on $h(t)$ are summarized in Table II along with associated kernel requirements.

Steps 2 and 3 utilize the scaling property of the FT and the dual properties between Q4–Q5, Q6–Q7, and Q8–Q9. As a result, the RID kernel is a type of product kernel which was introduced in [10], i.e., it is a function of $\theta\tau$.

From R1–R4, the RID will have a cross-shaped kernel. Figs. 2(a)–(c) show the kernels of the WD, spectrogram, and RID in the ambiguity domain.

Using a primitive function $h(t)$ designed according to the requirements R1–R4, the RID has the following integral expression:

$$\text{RID}_f(t, \omega; h) = \int\int \frac{1}{|\tau|} h\left(\frac{u-t}{\tau}\right) f(u + \tau/2) \cdot f^*(u - \tau/2) e^{-j\tau\omega}\,du\,d\tau. \tag{34}$$

For the RID computation, the generalized autocorrelation function

$$R_f'(t, \tau; h) = \int \frac{1}{|\tau|} h\left(\frac{u-t}{\tau}\right) f(u + \tau/2) \cdot f^*(u - \tau/2)\,du \tag{35}$$

is computed for all τ at time instant t; then the FT

$$\text{RID}_f(t, \omega; h) = \int R_f'(t, \tau; h) e^{-j\tau\omega}\,d\tau \tag{36}$$

is carried out.

Comparing the above expressions with Cohen's class in Section II, it can be easily seen that the RID has the following kernels in the (t, τ), (ω, θ), (θ, τ), and (t, ω) domains:

$$\psi_{\text{RID}}(t, \tau) = \frac{2\pi}{|\tau|} h(t/\tau) \tag{37}$$

$$\Psi_{\text{RID}}(\omega, \theta) = \frac{4\pi^2}{|\theta|} h(\omega/\theta) \tag{38}$$

$$\phi_{\text{RID}}(\theta, \tau) = H(\theta\tau) \tag{39}$$

$$\Phi_{\text{RID}}(t, \omega) = 2\pi \int \frac{1}{|\tau|} h(t/\tau) e^{-j\omega\tau}\,d\tau. \tag{40}$$

Note that the kernel in the (ω, θ) domain has the same shape as that in the (t, τ) domain. The RID kernel is cone shaped (or, bowtie shaped) in both the (t, τ) and (ω, θ) domains (Fig. 2(d)–(e)). In the time-frequency domain, the RID kernel has a somewhat "tricky" shape (Fig. 2(f)).

Hlawatsch has investigated many interesting properties

TABLE II
REQUIREMENTS ON $h(t)$ AND THEIR COUNTERPARTS IN TABLE I

R1. unit area: $\int h(t)\,dt = 1$	Q4, Q5
R2. symmetrical: $h(-t) = h(t)$	Q1, Q6, Q7
R3. time-limited: $h(t) = 0\ \forall \lvert t\rvert > \frac{1}{2}$	Q8, Q9
R4. low-pass type: $\lvert H(\theta)\rvert \ll 1$ for $\lvert\theta\rvert \gg 0$	Q10

Note: In this kernel design procedure, Q2 and Q3 are always satisfied, while Q0 is not satisfied.

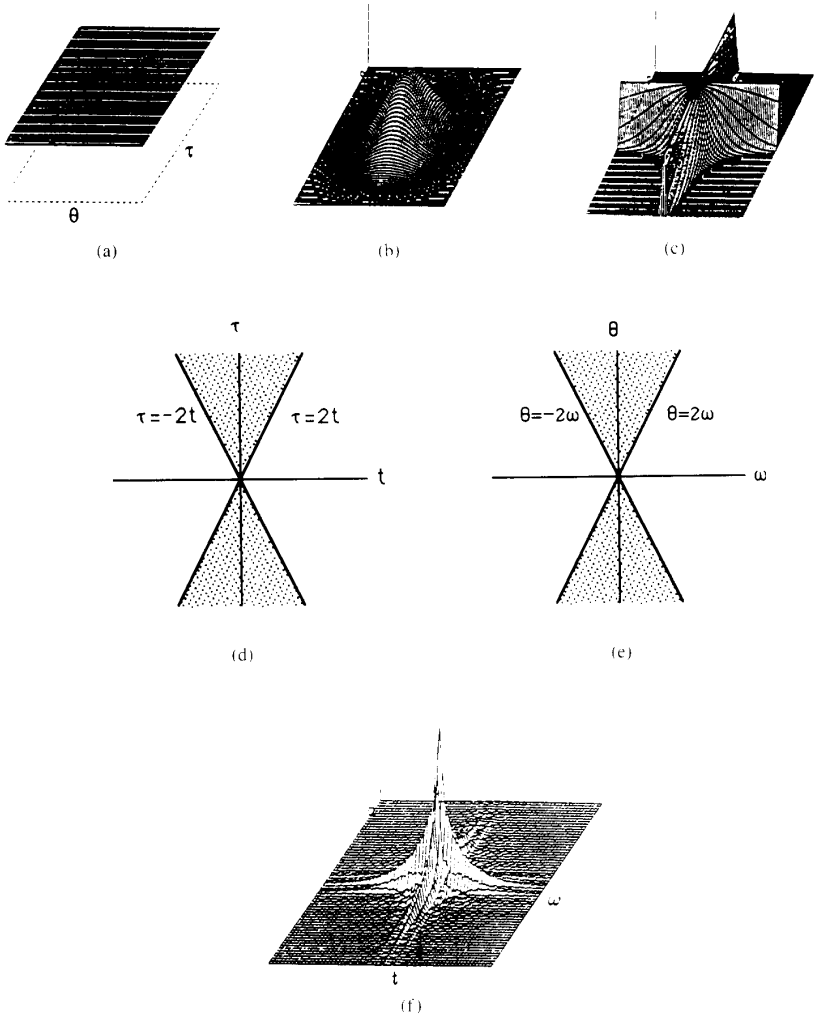

Fig. 2. Ambiguity domain kernels of (a) WD, (b) spectrogram, (c) RID. Nonzero support of the RID kernel in the (d) temporal correlation domain (e) spectral correlation domain. (f) The RID kernel in the time-frequency domain.

of product-type kernels, including the properties of scale invariance, regularity, and unitarity as well as P1–P9 in Table I [21]; P10 (reduced interference property) was not considered, however, in this work. As is clear in the above RID construction procedure, the RID emphasizes P10 and has a cross-shaped kernel. The WD, for instance, satisfies P1–P9 and the properties of scale invariance, regularity, and unitarity, but it does not satisfy P10. Hence it is not a member of the RID. In fact, the RID automatically has the scale-invariance property due to its product-type kernel; the RID may or may not satisfy the regularity property; and the RID does not have the unitarity property due to the incompatibility between this property and P10.

The RID is quite different in many respects from the "cone-shaped kernel" which has been recently proposed [40]. The cone-shaped kernel has the following (t, τ) do-

main expression:

$$\psi(t, \tau) = \begin{cases} g(\tau), & |t| < |\tau|/2 \\ 0, & \text{otherwise.} \end{cases} \quad (41)$$

Like the RID, this kernel has the properties of realness, shift invariance, and time support listed in Table I. Unlike the RID, however, this kernel does not satisfy some of the other desirable properties in Table I. Nevertheless, this kernel provides good resolution in some applications [40].

The RID is not dedicated to a certain type of signals. The key idea underlying the RID is to develop TFR's that satisfy many desirable properties (see Table III). Of course, one can find signals that will not be effectively handled by the RID; for example, a chirp. Since the RID has a cross-shaped kernel in the ambiguity domain, if the AF of the chirp falls on a 45° diagonal line, then it will not intersect well with the RID kernel, resulting in low time/frequency resolution. As for cross-term suppression in the RID, the cross terms are, in general, reduced in height but spread over a larger time-frequency area. This is also true for the cross terms in the spectrogram. Recall that the cross terms in the RID (or spectrogram) are the 2-D convolution between the cross terms in the WD and the time-frequency domain kernel of the RID (or spectrogram). In particular, if a cross term is located on the θ or the τ axis, it will not be suppressed in the RID. Therefore, kernels should be examined carefully in terms of the signals at hand and kernel design should be optimized depending on particular needs. In other words, one may wish to relinquish some desirable properties of TFR's in order to meet a specific goal. Except for those extreme cases aforementioned, the RID seems to perform quite well in a number of real-world situations.

For practical implementation of the RID, the generalized discrete-time TFR [5] may be used with proper constraints on the discrete-kernel. In this case, it is well known that in order to avoid aliasing, the signal should be sampled at no less than twice the Nyquist rate. An alternative approach, which can avoid the aliasing problem, can be found in [24], and this new approach has been applied to the tracking of the continuous-time instantaneous frequency [26].

IV. Design Considerations for RID

One easy way of designing a primitive function $h(t)$ is to use the well-known windows, e.g., the Hann window, and their frequency characteristics. Various windows and their frequency domain characteristics have been extensively investigated in [18], [30]. Two important remarks on selecting $h(t)$ directly are 1) in contrast to conventional one-dimensional spectral analysis, we need to consider the 2-D characteristics of $H(\theta\tau)$, which has a cross shape as mentioned before, and 2) since the time-frequency analysis kernel is multiplicative in the ambiguity domain, $H(\theta\tau)$ should be interpreted as a filter rather than as a convolution operator which is the interpretation it is given by most spectral analysis methods. For instance, although the rectangular window provides a very high resolution in conventional spectral analysis, it yields a low resolution kernel in the RID design since it has a very narrow bandwidth.

Another way of designing $h(t)$ is to employ the window technique from FIR filter design. One can first specify the characteristics of $\phi(\theta, \tau) = H(\theta\tau)$ (which is usually a 2-D cross-shaped low-pass filter), replace $\theta\tau$ by θ in $H(\theta\tau)$, and perform the inverse Fourier transform to produce $h(t)$. Finally, $h(t)$ is multiplied by a window $w(t)$ which is time limited on $[-1/2, 1/2]$. This procedure will enable one to determine the passband of the kernel directly in the ambiguity domain, as well as to satisfy all of R1–R4.

In the above kernel design procedure, the waveform of $h(t)$ determines the autoterm resolution and the cross-term reduction of the distribution. As might be expected, there is a tradeoff between the autoterm resolution and the interference suppression. If $h(t)$ is an impulse-like function (as in the WD), then the resultant distribution suffers from strong interference while maintaining high auto-term resolution. On the other hand, if $h(t)$ is a smooth function with little high frequency content, then the resultant distribution effectively suppresses the interference terms at the expense of the autoterm resolution. The impulse, rectangular, and Gaussian functions will produce the Wigner, Born–Jordan, and Choi–Williams kernels, respectively. Various selections for $h(t)$ are examined in the following. Some typical selections are compared in Table III.

Example 1: $h(t) = \delta(t)$: This choice of $h(t)$ results in the WD, for which $\phi(\theta, \tau) = 1$. This satisfies R1–R3, but not R4. Hence, the WD satisfies properties P1–P9 only. It offers the best resolution of the autoterms at the expense of strong interference between signal components.

Example 2: $h(t) = \delta(t + 1/2)$: This produces the Rihaczek distribution, for which $\phi(\theta, \tau) = e^{j\theta\tau/2}$. Here R1 and R3 are satisfied, but not R2 or R4. Hence, only properties P2–P5 and P7–P8 are satisfied.

Example 3: $h(t) = \delta(t + \alpha)$: This yields the generalized Wigner distribution (GWD), for which $\phi(\theta, \tau) = e^{j\alpha\theta\tau}$. This was defined in [15] to provide a geometric interpretation of the interference mechanism in the Wigner and Rihaczek distributions; $\alpha = 0$ and $\alpha = 1/2$ correspond to the Wigner and Rihaczek distributions, respectively. If $|\alpha| > 1/2$ then R4 is violated and the resultant distribution satisfies P2–P5 only.

Example 4: $h(t) = [\delta(t - 1/2) + \delta(t + 1/2)/2]$: This produces the real part of the Rihaczek distribution, for which $\phi(\theta, \tau) = \cos \theta\tau/2$. With this $h(t)$, R1–R3 are met, but not R4. Hence, only properties P1–P9 are satisfied.

Example 5: $h(t) = [\delta(t - \alpha) + \delta(t + \alpha)/2]$: This is a generalization of example 4. If $|\alpha| \leq 1/2$ then R1–R3 are met and properties P1–P9 are satisfied. Otherwise, only R1, R2 are met and properties P1–P7 are satisfied.

Example 6: $h(t) = \text{rect}(t)$, where rect $(t) = 1$ on $[-1/2, 1/2]$, $= 0$ otherwise: This corresponds to the Born–Jordan distribution [11]. It satisfies all of R1–R4,

TABLE III
COMPARISONS AMONG SEVERAL TIME-FREQUENCY DISTRIBUTIONS

Distribution	Time Function $h(t)$	Kernel $\phi(\theta, \tau)$	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Wigner	$\delta(t)$	1		x	x	x	x	x	x	x	x	x	
Rihaczek	$\delta(t + 1/2)$	$e^{j\theta\tau/2}$			x	x	x	x			x	x	
GWD $w/\lvert\alpha\rvert > 0.5$	$\delta(t + \alpha)$	$e^{j\alpha\theta\tau}$			x	x	x	x					
Re {Rihaczec}	$\frac{\delta(t - 1/2) + \delta(t + 1/2)}{2}$	$\cos(\theta\tau/2)$		x	x	x	x	x	x	x	x	x	
Choi-Williams	$\frac{1}{\sqrt{2\pi\alpha}} \exp\left(\frac{-t^2}{2\alpha^2}\right)$	$e^{-\alpha^2\theta^2\tau^2/2}$		x	x	x	x	x	x	x			x
Sinc	$\frac{\sin(\alpha t/2)}{\pi t}$	rect $(\theta\tau/\alpha)$		x	x	x	x	x	x	x			x
Spectrogram	n/a	$A_w(\theta, \tau)$ of a window $w(t)$	x	x	x	x							x
Born-Jordan*	rect (t)	$\varphi(\theta\tau) \triangleq \frac{\sin(\theta\tau/2)}{\theta\tau/2}$		x	x	x	x	x	x	x	x	x	x
G-Hamming*	$[1 + \beta \cos 2\pi t]$ rect (t)	$\varphi(\theta\tau) + \frac{\beta}{2}[\varphi(\theta\tau + \pi) + \varphi(\theta\tau - \pi)]$		x	x	x	x	x	x	x	x	x	x
Truncated-CW*	$\frac{1}{\sqrt{2\pi\alpha}} \exp\left(\frac{-t^2}{2\alpha^2}\right) w(t)$	$e^{-\alpha^2\theta^2/2} * W(\theta)\lvert_{\theta=\theta\tau}$		x	x	x	x	x	x	x	x	x	x
Truncated-Sinc*	$\frac{\sin(\alpha t/2)}{\pi t} w(t)$	rect $(\theta/\alpha) * W(\theta)\lvert_{\theta=\theta\tau}$		x	x	x	x	x	x	x	x	x	x
Triangular*	$2 - 4\lvert t\rvert$, $\lvert t\rvert < 0.5$	$\varphi^2(\theta\tau)$		x	x	x	x	x	x	x	x	x	x

*belongs to the RID.

and properties P1–P10 are satisfied. Since $h(t)$ is flat on $[-1/2, 1/2]$, the resultant $H(\theta\tau)$ has a very narrow bandwidth. Hence, the Born–Jordan kernel provides very good interference reduction at the expense of autoterm resolution.

Example 7: $h(t) = 1/\alpha$ rect (t/α): This produces the generalized Born–Jordan distribution (GBJ). If $|\alpha| > 1$ then it violates R4; hence P8, P9 are not satisfied. A larger value of α will lead to larger degree of smoothing, or equivalently, lower resolution in time and frequency with improved interference suppression.

Example 8: $h(t) = [1 + \beta \cos 2\pi t]$ rect (t): This will give a generalized-Hamming-type kernel. β controls the amount of tapering in $h(t)$, or, equivalently, the shape of the resultant 2-D kernel in the ambiguity domain. $\beta = 0$, 0.92, and 1 correspond to the GBJ, Hamming-type, and Hanning-type kernel, respectively.

Example 9: $h(t) = 1/\alpha\,[1 + \beta \cos 2\pi t/\alpha]$ rect (t/α): This is an extension of example 8. α plays the same role as in example 7.

Example 10: $h(t) = (1/\sqrt{2\pi\alpha}) \exp(-t^2/2\alpha^2)$: This will produce the Choi–Williams distribution [7]. Only R1, R2, and R4 will be met, resulting in properties P1–P7 and P10 being satisfied. α trades off the autoterm resolution and interference reduction. If α is sufficiently small then R3 is approximately satisfied.

Example 11: $h(t) = (1/\sqrt{2\pi\alpha}) \exp(-t^2/2\alpha^2) w(t)$: This is a truncated version of example 10 such that R3 is also met. The window $w(t)$ is time limited within $[-1/2, 1/2]$, symmetrical, and scaled appropriately so that the resulting function $h(t)$ satisfies R3, R2, and R1, respectively. Hence, the resultant distribution has all of P1–P10.

Example 12: $h(t) = \sin(\alpha t/2)/\pi t$: This yields a sharp transition between the passband and stopband of the kernel in the abmiguity domain since the Fourier transform of $h(t)$ is a rectangular waveform. It does not, however, meet R3; hence, P8, P9 are not satisfied. The passband in the ambiguity domain is determined by the value of α.

Example 13: $h(t) = [\sin(\alpha t/2)/\pi t] w(t)$: This is a truncated version of example 12, where $w(t)$ plays the same role as in example 10. This procedure to generate $h(t)$ is analogous to the well-known FIR low-pass filter design technique using windows. It now meets all of R1–R4 and the resultant distribution satisfies all of the properties P1–P10.

Example 14: $h(t) = 2 - 4|t|$ for $|t| < 0.05$, $= 0$ otherwise: This triangular window meets all of R1–R4 and has properties P1–P10.

V. EXPERIMENTAL RESULTS

In order to compare the performances of three typical time-frequency distributions (WD, spectrogram, and RID), several computer experiments have been conducted using both computer-generated signals and real-world signals. In this section, an example of a three-tone signal is shown which clearly reveals the differences in performance among those three distributions. Other examples have also been reported [36]–[38] or are in preparation for publication.

A three-tone signal has been constructed over a 512-point sample interval as follows:

$$f(n) = \begin{cases} e^{j\omega_1 n}, & 32 \le n \le 192 \\ e^{j\omega_2 n}, & 193 \le n \le 319 \\ e^{j\omega_3 n}, & 320 \le n \le 480 \\ 0, & \text{otherwise} \end{cases} \tag{42}$$

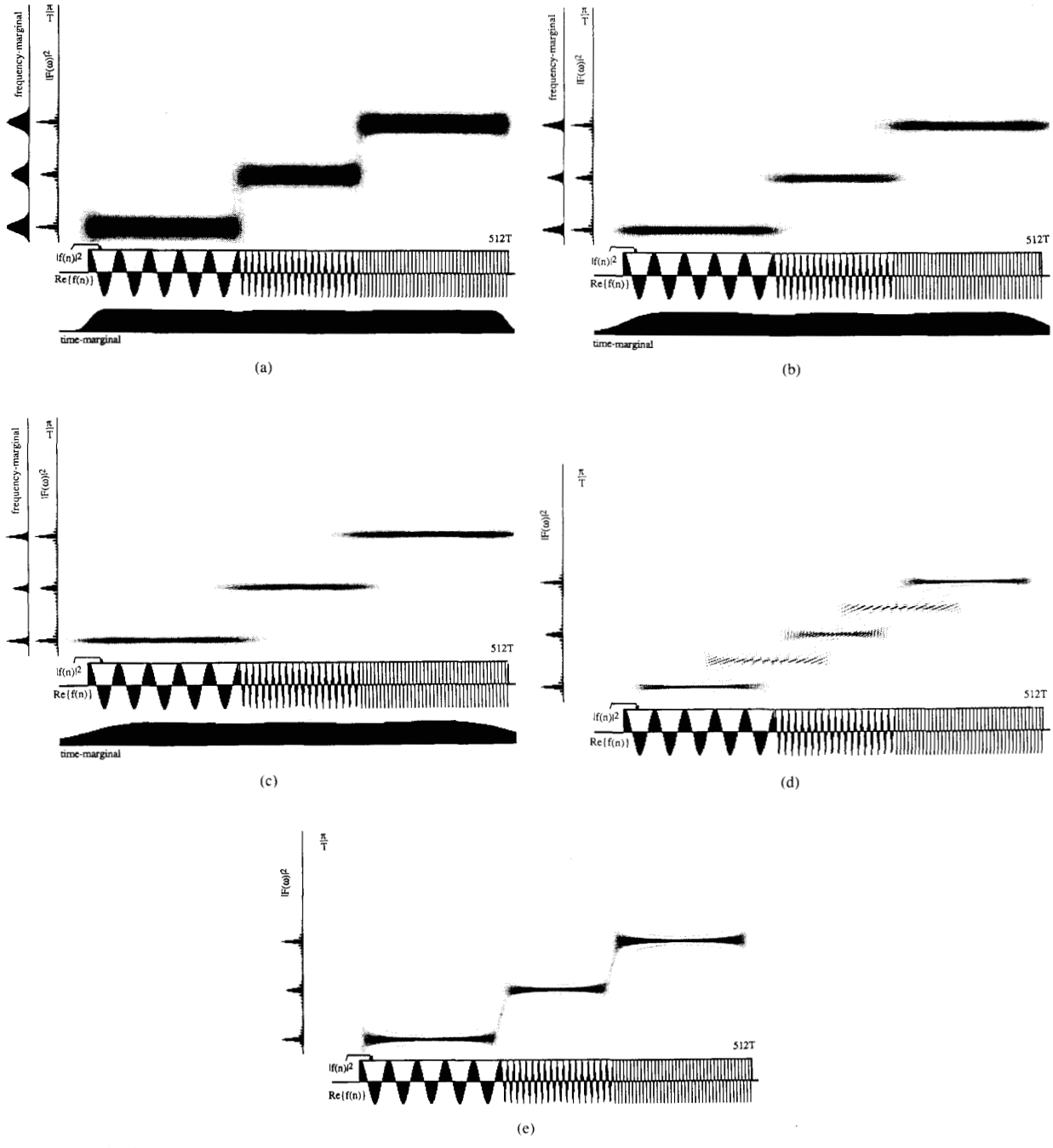

Fig. 3. Time-frequency distributions for a three-tone signal. (a) Spectrogram with a 64-point window. (b) Spectrogram with a 160-point window. (c) Spectrogram with a 256-point window. (d) WD. (e) RID.

where $\omega_1 = 2\pi \cdot 16/512$, $\omega_2 = 2\pi \cdot 72/512$, and $\omega_3 = 2\pi \cdot 128/512$. Obviously, the instantaneous power $|f(n)|^2 = 1$ for $32 \leq n \leq 480$, and $= 0$, otherwise.

Figs. 3(a)–(c) are the spectrograms for $f(n)$ using 64-, 160-, and 256-point Hann (or Hanning) windows, respectively. They clearly illustrate the tradeoff between time and frequency resolutions when using the spectrogram. One can easily see that the 64-point window provides a fairly good time resolution, but does a poor job in representing the frequency spectra. The frequency spectrum at each instant is spread considerably beyond the true frequency support. It is well known that the time (frequency)

marginal of the spectrogram is identical to the convolution between the instantaneous power (energy spectral density) of the signal and that of the window used [10]. The marginals in time and frequency have been calculated from the spectrogram, and they have been superimposed in Figs. 3(a)–(c) for comparison with the true instantaneous power and energy spectral density. In Fig. 3(a) the time marginal shows a small amount of spread, while the frequency marginal shows a great amount of spread. In contrast, the 256-point window represents the frequency spectra faithfully at the expense of time resolution. This can also be verified by looking at the marginals in time and frequency. In Fig. 3(c), unlike the 64-point case, the time marginal is considerably spread out, while the frequency marginal is relatively less spread. Being placed between these two extremes, the 160-point window (Fig. 3(b)) offers an intemediate level of performance in terms of time and frequency resolutions, but still is not a satisfactory representation. Figs. 3(d), (e) are the WD and the RID for the three-tone signal, respectively. For the RID, a truncated Gaussian function (with $\alpha = 1/6$) is chosen as the primitive function $h(t)$ (see example 11 in the previous section). As explained earlier, both the WD and the RID satisfy the marginals and the support properties. The WD (Fig. 3(d)), however, suffers from cross terms between tones, which appear as displaced "shadow pattern" companions to the true autoterm components. In particular, the interference between the first and the third tones disrupts the second autoterm component, resulting in potential misinterpretation regarding the signal spectrum. The RID (Fig. 3(e)) reduces the cross terms fairly effectively, showing only small hints of cross terms. It still has very good resolution in time and frequency, offering excellent interpretability as a time-varying frequency spectrum.

VI. Conclusions

A new class of time-frequency distributions has been introduced, which has many desirable properties. It has been shown that this new distribution provides a high-resolution, easy-to-interpret localization of the signal energy spectrum in the time-frequency plane. A systematic procedure to design the new kernels has been proposed as well. Some aspects and properties of the new distributions, including its limitations, have been discussed. Interpretations in the ambiguity, temporal correlation, spectral correlation, and time-frequency domains have been presented for better insight into the process of kernel design.

Acknowledgment

The authors would like to thank the referees for their constructive suggestions and comments. In particular, they thank one of the referees for pointing out [21]. Some results of [21] have been cited in this revision, and this have greatly enhanced the rigor and clarity of this paper.

References

[1] R. A. Altes, "Detection, estimation, and classification with spectrograms," *J. Acoust. Soc. Amer.*, vol. 67, no. 4, pp. 1232–1246, Apr. 1980.

[2] M. G. Amin, "Time and lag window selection in the Wigner–Ville distribution," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1987, pp. 1529–1532.

[3] J. C. Andrieux, M. R. Feix, G. Mourgues, P. Bertland, B. Izrar, and V. T. Nguyen, "Optimum smoothing of the Wigner–Ville distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 764–769, 1987.

[4] M. J. Bastiaans, "A sampling theorem for the complex spectrograms, and Gabor's expansion of a signal in Gaussian elementary signals," *Opt. Eng.*, vol. 20, no. 4, pp. 594–598, Aug. 1981.

[5] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing*, vol. 1, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[6] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using the Wigner distribution synthesis techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 442–451, 1986.

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

[8] T. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution—A tool for time-frequency signal analysis—Part I: Continuous-time signals," *Philips J. Res.*, vol. 35, pp. 217–250, 1980.

[9] T. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution—A tool for time-frequency signal analysis—Part III: Relations with other time-frequency signal transformations," *Philips J. Res.*, vol. 35, pp. 372–389, 1980.

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

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

[12] L. Cohen and T. E. Posch, "Generalized ambiguity function," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1985, pp. 1033–1036.

[13] I. Daubechies, "The wavelet transform, time-frequency localization, and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, no. 5, pp. 961–1005, 1990.

[14] P. Flandrin, "Some features of time-frequency representations of multicomponent signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1984, pp. 41B.4.1–4.4.

[15] P. Flandrin and F. Hlawatsch, "Signal representations geometry and catastrophies in the time-frequency plane," in *Mathematics in Signal Processing*, T. Durrani *et al.*, Eds. Oxford: Clarendon, 1987, pp. 3–14.

[16] D. Gabor, "Theory of communication," *J. Inst. Elec. Eng.*, vol. 93, no. III, pp. 429–457, 1946.

[17] A. Grossman and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM J. Math. Anal.*, vol. 15, pp. 723–736, 1984.

[18] F. J. Harris, "On the use of windows for harmonic analysis with discrete Fourier transform," *Proc. IEEE*, vol. 66, pp. 51–83, Jan. 1978.

[19] C. W. Helstrom, "An expansion of a signal in Gaussian elementary signals," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 81–82, 1966.

[20] F. Hlawatsch, "Interference terms in the Wigner distribution," in *Digital Signal Processing-84*, V. Cappellini and A. Constantinides, Eds. North-Holland, 1984, pp. 363–367.

[21] F. Hlawatsch, "A study of bilinear time-frequency signal representations, with applications to time-frequency signal synthesis," doctoral dissertation, Technische Universitat Wien, Vienna, Austria, May 1988.

[22] C. P. Janse and J. M. Kaizer, "Time-frequency distributions of loudspeakers: The application of the Wigner distribution," *J. Audio Eng. Soc.*, vol. 31, pp. 198–223, 1983.

[23] J. Jeong and W. J. Williams, "On the cross-terms in spectrograms," *IEEE Int. Symp. Circuits Syst.*, 1990, pp. 1565–1568.
[24] J. Jeong and W. J. Williams, "Alias-free generalized discrete-time time-frequency distributions," *IEEE Trans. Signal Processing*, to be published.
[25] J. Jeong and W. J. Williams, "Variable-windowed spectrograms: Connecting Cohen's and the wavelet transform," in *IEEE ASSP Workshop Spectrum Estimation Modeling*, 1990, pp. 270–279.
[26] J. Jeong, G. C. Cunningham, and W. J. Williams, "Instantaneous frequency and kernel requirements for discrete time-frequency distributions," *Proc. SPIE Int. Soc. Opt. Eng.*, 1990.
[27] R. Kroland-Martinet, J. Morlet, and A. Grossmann, "Analysis of sound patterns through wavelet transforms," *Int. J. Pattern Recog. Art. Intell.*, vol. 1, pp. 273–301, 1987.
[28] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.
[29] S.H. Nawab and T. F. Quatieri, "Short-time Fourier transform," in *Advanced Topics in Signal Processing*, J. S. Lim and A. V. Oppenheim, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1988, pp. 289–337.
[30] A. H. Nuttall, "Some windows with very good sidelobe behavior," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 84–91, Feb. 1981.
[31] M. R. Portnoff, "Representation of digital signals and systems based on the short-time Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 55–69, Feb. 1980.
[32] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ:Prentice-Hall, 1978.
[33] W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 369–374, 1968.
[34] J. Ville, "Theorie et applications de la notion de signal analytique," *Cables Transmiss.*, vol. 20A, pp. 61–74, 1948.
[35] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, pp. 749–759, 1932.
[36] W. J. Williams and J. Jeong, "New time-frequency distributions: Theory and applications," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 2, May 1989, pp. 1243–1247.
[37] W.J. Williams and J. Jeong, "New time-frequency distributions for the analysis of multicomponent signals," *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 1152, pp. 483–495, 1989.
[38] W. J. Williams and J. Jeong, "Reduced interference time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash, Ed. Melbourne, Australia: Longman Cheshire, 1991.
[39] P. M. Woodward, *Probability and Information Theory with Application to Radar*. London: Pergamon, 1953.
[40] Y. Zhao, L. Atlas, and R. Marks, "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no.7, pp. 1084–1091, July 1990.

Jechang Jeong (M'91) received the B.S.E. degree in electronic engineering from Seoul National University in 1980, the M.S.E. degree in electrical engineering from the Korean Advanced Institute of Science and Technology in 1982, and the Ph.D. degree in electrical engineering from the University of Michigan, Ann Arbor, in 1990.

From 1982 to 1986, he was with the Korean Broadcasting System, where he helped develop teletext systems. Since 1990 he has been with Samsung Electronics Inc., Suwon, Korea, where he is currently a Research Manager. His research interests include digital signal processing, time-varying spectral analysis, digital communication, and fast algorithms/VLSI implementation. He has written about 20 published technical papers.

Dr. Jeong, during his study at the University of Michigan, received the Rackham Research Partnership Grant in 1987, and the Rackham Predoctoral Fellowship in 1989. He was also nominated for the Distinguished Dissertation Award in 1990.

William J. Williams (S'63–M'64–SM'73) was born in Ohio. He received the B.E.E. degree from Ohio State University, Columbus, in 1958, the M.S. degree in physiology from the University of Michigan, Ann Arbor, in 1966, and the M.S. and Ph.D. degrees in electrical engineering from the University of Iowa, Iowa City, in 1961 and 1963, respectively.

From 1958 to 1960 he was a Research Engineer at Battelle Memorial Institute, Columbus, OH. He was a Senior Engineer Specialist in communication systems with Emerson Electric from 1963 to 1964. He joined the Faculty of the (now) Electrical Engineering and Computer Science Department and was appointed Professor in 1974. During a sabbatical leave in 1974 he was a Visiting Scientist in Physiology at Johns Hopkins University. He has been active in the bioengineering program at the University of Michigan for a number of years. His interests include the theory and application of signal processing and communication techniques, especially, but not exclusively, to biological problems. He has a particular interest in time-frequency distributions and their applications in nonstationary signal analysis.

Dr. Williams was the Guest Editor for a special issue of the PROCEEDINGS OF THE IEEE on biological signal processing and analysis which appeared in May 1977. He is a member of Tau Beta Pi, Eta Kappa Nu, Sigma Xi, and the Society for Neuroscience.