Copulas: A New Insight Into Positive Time-Frequency Distributions

Manuel Davy and Arnaud Doucet

***Abstract*—In this letter, we establish connections between Cohen–Posch theory of positive time-frequency distributions (TFDs) and copula theory. Both are aimed at designing joint probability distributions with fixed marginals, and we demonstrate that they are formally equivalent. Moreover, we show that copula theory leads to a noniterative method for constructing positive TFDs. Simulations show typical results.**

***Index Terms*—Copulas, imposed marginals, joint distribution, positive time-frequency.**

I. Introduction

POSITIVE Time-Frequency Distributions (TFDs) [1], [2] have the nice property of being positive and having correct marginals. More precisely, let $s(t)$ denote the time-domain signal, and $\mathcal{P}(t,f)$ its positive TFD. Then we have

$$\begin{aligned}
&\mathcal{P}(t,f) \geq 0, \qquad \text{for all } (t,f) \in \mathbb{R}^2\\
&\int_{-\infty}^{+\infty} \mathcal{P}(t,\nu)d\nu = \mathcal{T}(t), \qquad \text{for all } (t) \in \mathbb{R}\\
&\int_{-\infty}^{+\infty} \mathcal{P}(\tau,f)d\tau = \mathcal{F}(f), \qquad \text{for all } (f) \in \mathbb{R}
\end{aligned} \tag{1}$$

where $\mathcal{T}(t) = |s(t)|^2$ is the time marginal, and $\mathcal{F}(f) = |\mathrm{FT}_s(f)|^2$ is the frequency marginal ($\mathrm{FT}_s(f) = \int s(t)\exp(-j2\pi ft)dt$ denotes the Fourier transform of s). For the sake of simplicity, we assume here that $s(t)$ is normalized such that $E_s = \int |s(t)|^2 dt = 1$. Whenever $E_s \neq 1$, the results remain true up to the multiplicative factor E_s. In [1] and [2], it is shown that all distributions such as in (1) can be written as

$$\mathcal{P}(t,f) = \mathcal{T}(t)\mathcal{F}(f)\Omega\left(\mathbf{T}(t), \mathbf{F}(f)\right) \tag{2}$$

where $\mathbf{T}$ (respectively, $\mathbf{F}$) denotes[1] the cumulative distribution related to $\mathcal{T}$ (respectively, $\mathcal{F}$) by $\mathbf{T}(t) = \int_{-\infty}^{t} \mathcal{T}(du)$ [respectively, $\mathbf{F}(f) = \int_{-\infty}^{f} \mathcal{F}(du)$], and Ω is a positive function such that

$$\int_0^1 \Omega(u,v)du = \int_0^1 \Omega(u,v)dv = 1. \tag{3}$$

Positive TFDs are used in a variety of applications such as the analysis and synthesis of multicomponent signals [3], biomedical engineering [4], and speech processing [5]. Moreover, efficient recursive implementations have been proposed [6]–[10], most of which provide minimum cross-entropy (MCE) TFDs, which requires the definition of a prior estimate such as the spectrogram. Aside the work of Cohen and Posch [1], [2], statisticians [11], [12] have addressed the same problem, namely *how to construct a joint probability distribution with imposed marginals?* These researches have led to the concept of *copula*, and we show in this letter that the main result in copula theory, Sklar's theorem, is equivalent to the formulation given in [2]. Each TFD admits a copula that contains all the information about the signal time-frequency dependence and that can be used to measure it. In addition to providing a new insight into time-frequency theory, copulas provide a simple noniterative way to construct positive TFDs with correct marginals. These TFDs are proved to be different from MCE TFDs and are easy to compute in practice.

This letter is organized as follows. In Section II, we recall some basic definitions and properties related to copulas. In Section III, we show that copulas and Cohen–Posch's group of positive TFDs are actually two versions of the same theory. A direct consequence of this connection is presented in Section IV where we compute the copula of a Gaussian chirp, and we propose a new method aimed at computing positive TFDs for any signal. In Section V, we illustrate the interest of copulas in time-frequency analysis by simulations. We give finally some conclusions and research directions in Section VI.

II. Copulas

In this section, we consider *cumulative* distributions rather than noncumulative probability distributions. The elements presented here are taken from [13].

Definition 1: A copula $\mathbf{C}$ is a function from $[0,1]^2$ to [0,1] such that

1) $\mathbf{C}(u,0) = \mathbf{C}(0,v) = 0$ for all $(u,v) \in [0,1]^2$;
2) $\mathbf{C}(u,1) = u$ and $\mathbf{C}(1,v) = v$ for all $(u,v) \in [0,1]^2$;
3) for all $(u_1,u_2,v_1,v_2) \in [0,1]^4$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$

$$\mathbf{C}(u_2,v_2) - \mathbf{C}(u_1,v_2) - \mathbf{C}(u_2,v_1) + \mathbf{C}(u_1,v_1) \geq 0. \tag{4}$$

Manuscript received April 9, 2002; revised October 8, 2002. The work of M. Davy was supported by the European project MOUMIR. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. See-May Phoon.

M. Davy is with IRCCyN/CNRS, 44321 Nantes Cedex 03, France (e-mail: Manuel.Davy@irccyn.ec-nantes.fr).

A. Doucet is with the University of Cambridge, Department of Engineering, CB2 1PZ, UK (e-mail: ad2@eng.cam.ac.uk).

Digital Object Identifier 10.1109/LSP.2003.811636

[1]Here, distributions are denoted with calligraphic letters (e.g., $\mathcal{X}$), whereas cumulative distributions are denoted with boldface capital letters (e.g., $\mathbf{X}$).

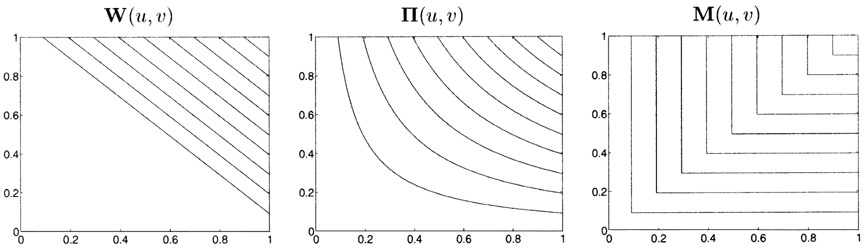

Fig. 1. Three important copulas displayed as contour plots.

Fig. 1 displays the contour plots of three important copulas: the product copula $\mathbf{\Pi}(u,v) = uv$ and the *Fréchet–Hoeffding bounds* $\mathbf{W}$ and $\mathbf{M}$, whose existence is justified by the following theorem:

Theorem 2: Let $\mathbf{C}$ be a copula. Let $\mathbf{W}(u,v) = \max(u+v-1,0)$ and $\mathbf{M}(u,v) = \min(u,v)$. Then, for all $(u,v) \in [0,1]^2$, $\mathbf{W}(u,v) \leq \mathbf{C}(u,v) \leq \mathbf{M}(u,v)$.

We now expose the theorem that justifies the use of copulas in the study of joint distributions with imposed marginals.

Theorem 3 (Sklar's Theorem [11]): Let $\mathbf{P}(x,y)$ be a cumulative distribution with marginals $\mathbf{X}(x)$ and $\mathbf{Y}(y)$. Then there exists a copula $\mathbf{C}$ such that

$$\mathbf{P}(x,y) = \mathbf{C}\left(\mathbf{X}(x), \mathbf{Y}(y)\right). \tag{5}$$

If $\mathbf{X}(x)$ and $\mathbf{Y}(y)$ are continuous, then $\mathbf{C}$ is unique; otherwise $\mathbf{C}$ is uniquely defined on Range[$\mathbf{X}(x)$] $\times$ Range[$\mathbf{Y}(y)$]. Conversely, if $\mathbf{X}(x)$ and $\mathbf{Y}(y)$ are distributions and $\mathbf{C}$ is a copula, then $\mathbf{P}(x,y)$ defined as in (5) is a distribution function with marginals $\mathbf{X}(x)$ and $\mathbf{Y}(y)$.

In practice, this theorem means that, whenever the marginals are fixed, a given probability distribution is uniquely related to a copula and conversely. It also provides a way to construct the copula from a given distribution (see Corollary 5). This requires, however, to define the *quasi inverse* of a marginal distribution $\mathbf{X}$ (or equivalently, $\mathbf{Y}$).

Definition 4: Let $\mathbf{P}$ be a univariate distribution function. A quasi-inverse $\mathbf{P}^{(-1)}$ is such that

$$\mathbf{P}\left(\mathbf{P}^{(-1)}(u)\right) = u, \text{ if } u \in \text{Range}[\mathbf{P}] \tag{6}$$

$$\mathbf{P}^{(-1)}(u) = \inf\left\{x | \mathbf{P}(x) \geq u\right\}, \text{otherwise.} \tag{7}$$

Corollary 5: Let $\mathbf{P}(x,y)$ be a cumulative distribution with marginals $\mathbf{X}(x)$ and $\mathbf{Y}(y)$, and $\mathbf{C}$ the corresponding copula. Then

$$\mathbf{C}(u,v) = \mathbf{P}\left(\mathbf{X}^{(-1)}(u), \mathbf{Y}^{(-1)}(v)\right). \tag{8}$$

These results are useful in practice. On the one hand, given two marginals $\mathbf{X}(x)$ and $\mathbf{Y}(y)$, one can construct a family of distributions whose marginals are $\mathbf{X}(x)$ and $\mathbf{Y}(y)$. On the other hand, given a distribution $\mathbf{P}(x,y) = \Pr(X \leq x, Y \leq y)$, one can use its copula to characterize the dependence of the random variables X and Y. This can be done by using the following measure of dependence:

$$\sigma_{\mathbf{C}} = 12 \int_0^1 \int_0^1 \left|\mathbf{C}(u,v) - uv\right| dudv \tag{9}$$

where $\mathbf{C}(u,v)$ is the copula related to $\mathbf{P}(x,y)$. This measure is maximum $\sigma_{\mathbf{C}} = 1$ whenever $\mathbf{C}(u,v)$ is one of the Fréchet–Hoeffding bounds. In this case, each of X and Y is almost surely a strictly monotone function of the other [13, p. 170]. Another classical characterization of random variables dependence is mutual information $I(X,Y)$. Let $\mathcal{C}$ be such that $\int_{-\infty}^{u}\int_{-\infty}^{v} \mathcal{C}(du', dv') = \mathbf{C}(u,v)$. Then, studying the dependence of (X,Y) using the mutual information $I(X,Y)$ is equivalent to studying the copula entropy $H(\mathcal{C})$, since

$$\begin{aligned} I(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{P}(x,y) \log \frac{\mathcal{P}(x,y)}{\mathcal{X}(x)\mathcal{Y}(y)} dxdy \\ &= -H(\mathcal{C}) \end{aligned} \tag{10}$$

where the copula entropy is defined by $H(\mathcal{C}) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{C}(u,v) \log \mathcal{C}(u,v) dudv$.

III. Redefining Positive Time-Frequency Distributions via Copulas

In order to use copula theory in the context of positive TFDs, we notice that the cumulative distribution $\mathbf{P}(t,f)$ related to the distribution $\mathcal{P}(t,f)$ of (2) is

$$\begin{aligned} \mathbf{P}(t,f) &= \int_{-\infty}^{t}\int_{-\infty}^{f} \mathcal{P}(\tau,\nu) d\tau d\nu \\ &= \mathbf{\Omega}\left(\mathbf{T}(t), \mathbf{F}(f)\right) \end{aligned} \tag{11}$$

where

$$\mathbf{\Omega}(u,v) = \int_{-\infty}^{u}\int_{-\infty}^{v} \Omega(u',v') dudv. \tag{12}$$

It is easy to verity that $\mathbf{\Omega}$ is actually a copula, and from Sklar's theorem, the copula defining a given positive TFD is unique. This result yields a new insight into Cohen–Posch's positive TFDs, by providing a new way of constructing such TFDs: the set of all possible positive TFDs with correct marginals is the set of functions $\mathbf{P}(t,f)$ such that

$$\mathbf{P}(t,f) = \mathbf{C}\left(\mathbf{T}(t), \mathbf{F}(f)\right) \tag{13}$$

where $\mathbf{C}$ is any copula. All results from copula theory can then be applied to time-frequency analysis. For a given signal $s(t)$, however, most $\mathbf{P}(t,f)$, or equivalently, most copulas $\mathbf{C}(u,v)$ are useless, as they do not incorporate information about the actual time-frequency energy location of $s(t)$. In order for $\mathcal{P}(t,f)$

to represent effectively the time-frequency contents of $s(t)$, it is necessary to elaborate signal-dependent copulas.

IV. Time-Frequency Copulas

In this section, we apply Sklar's theorem and its corollary to construct copulas for time-frequency distributions, called time-frequency copulas (TFCs).

A. TFC of a Simple Signal

In order to obtain a better comprehension of copulas, we compute the TFC for an important class of signals. Consider a chirp with Gaussian envelope [6]

$$s(t) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} \exp\left[-\frac{\alpha}{2}t^2 + j\frac{\beta}{2}t^2 + j2\pi f_0 t\right]. \tag{14}$$

Its cumulative marginal distributions are

$$\mathbf{T}(t) = Q(\sqrt{\alpha}t) \tag{15}$$

$$\mathbf{F}(f) = Q\left(\sqrt{\frac{\alpha^2+\beta^2}{\alpha}}2\pi(f-f_0)\right) \tag{16}$$

where $Q(x) = \int_{-\infty}^{x} \exp(-u^2)du$. The Wigner–Ville distribution of $s(t)$ [denoted $\mathcal{P}_{WV}(t,f)$] is positive,[2] and it has the correct marginals. $\mathcal{P}_{WV}(t,f)$ is also the MCE TFD of $s(t)$ [6] with the uniform distribution as prior. The cumulative distribution is

$$\mathbf{P}_{WV}(t,f) = \frac{\sqrt{\alpha}}{\pi}\int_{-\infty}^{t} e^{-\alpha\tau^2} Q\left(\frac{2\pi(f-f_0)}{\sqrt{\alpha}} - \frac{\beta}{\sqrt{\alpha}}\tau\right)d\tau. \tag{17}$$

We can then compute the corresponding TFC in some special cases

$$\mathbf{C} = \mathbf{\Pi} \quad \text{if} \quad \beta = 0 \tag{18}$$

$$\mathbf{C} \underset{\alpha\to 0}{\longrightarrow} \mathbf{M} \quad \text{if} \quad \beta > 0 \tag{19}$$

$$\mathbf{C} \underset{\alpha\to 0}{\longrightarrow} \mathbf{W} \quad \text{if} \quad \beta < 0. \tag{20}$$

Equations (19) and (20) are obtained using $Q((a-b\tau)/\sqrt{\alpha}) \longrightarrow \Gamma((a-b\tau)/\sqrt{\alpha})$, $(a,b) \in \mathbb{R}^2$, as α tends to zero (where Γ is the unit step), i.e., as the time support of $s(t)$ tends to be infinite.

The interpretation of this result is the following: there is no time-frequency dependence when the signal is stationary ($\beta = 0$); the dependence is "maximum" whenever the chirp is not stationary ($\beta < 0$ or $\beta > 0$). Depending on the sign of the chirp slope, the copula reaches one or the other Fréchet–Hoeffding bounds. However, recall that in such cases, one of the variables (e.g., f) is almost surely a strictly monotone function of the other (e.g., t). This explains our result and extends it other strictly monotonic time-frequency functions than chirps.

B. Construction of TFCs

In the general case, the above computation is generally not feasible. Moreover, the Wigner–Ville distribution is often nonpositive; thus there is no underlying copula. It is, however, necessary to find a TFC that captures most time-frequency dependences of the signal at hand.

[2]This signal is the only one for which the Wigner–Ville distribution is positive [6].

Let $\mathcal{S}(t,f)$ be a "satisfactory" TFR of $s(t)$ (i.e., a positive TFD that represents correctly the time-frequency contents of $s(t)$) that however does not satisfy the marginal requirement. Its cumulative marginals are denoted $\widetilde{\mathbf{T}}(t)$ and $\widetilde{\mathbf{F}}(f)$. Then, construct the copula $\mathbf{C}^{\mathcal{S}}$ of $\mathbf{S}(t,f)$ (using Corollary 5) such that

$$\mathbf{S}(t,f) = \mathbf{C}^{\mathcal{S}}\left(\widetilde{\mathbf{T}}(t), \widetilde{\mathbf{F}}(f)\right). \tag{21}$$

A still more satisfactory cumulative positive TFD (satisfying the marginals) of $s(t)$ is then

$$\mathbf{P}(t,f) = \mathbf{C}^{\mathcal{S}}\left(\mathbf{T}(t), \mathbf{F}(f)\right) \tag{22}$$

where $\mathbf{T}(t)$ and $\mathbf{F}(f)$ are the correct marginals. Typically, $\mathcal{S}(t,f)$ can be the spectrogram of $s(t)$ computed with the windowing function h.

C. Discussion

This technique has many advantages. First, it is noniterative as opposed to the optimization techniques in [6], [8], and [9]. Second, the information about the time-frequency dependence is made free from "marginal effects"; in other words, the copula itself can be used to analyze the signal, measure its time-frequency dependence, or characterize a class of signals. Third, in the case of spectrogram-based TFCs, the choice of the window h has little influence on the shape of the copula (provided the window length is reasonable), as will be shown in Section V. However, similar to [6] where the choice of the prior distribution had an influence on the resulting TFD, different choices of h lead to slightly different positive TFDs.

An important question is that of comparison with MCE TFDs. For a given spectrogram, the TFD computed with our method is visually very close to the MCE TFD obtained by choosing as prior the same spectrogram (see the simulations displayed in Section V). However, the closed-form calculations in Section IV-A show that the MCE TFD copula of a chirp (namely $\mathbf{W}$ or $\mathbf{M}$) is not the same as the copula of the distribution chosen as prior (namely $\mathbf{\Pi}$ for the uniform prior), which shows that MCE TFDs are different from our TFDs.

Note finally that the theory can be extended to discrete TFDs, since copulas are defined for any cumulative distribution (continuous or discrete). Given a discrete starting TFD, one can construct a discrete positive TFD with correct marginals using the technique presented above.

Remark: In a recent publication [10], a class of isentropic positive TFDs is presented. The entropy of a given positive TFD $\mathcal{P}(t,f)$ is $H(\mathcal{P}) = H(\mathcal{T}) + H(\mathcal{F}) + H(\mathcal{C})$; thus, the choice of a class of isentropic copulas leads to isentropic TFDs (for given marginals). The copula corresponding to the parameterization function proposed in [10] is

$$\mathbf{C}(u,v) = uv + \frac{\epsilon}{4\pi^2 nm}\left[\cos 2\pi(mv-\Delta) + \cos 2\pi(nu-\Delta) - \cos 2\pi(nu+mv-\Delta) - \cos 2\pi\Delta\right] \tag{23}$$

where $\epsilon \in \{-1,1\}$, $\Delta \in [0,2\pi]$ and $(m,n) \in \mathbb{Z}^2$. For the special case $m = n = 1$, $\mathbf{C}(v,u)$ is symmetric, i.e., $\mathbf{C}(v,u) = \mathbf{C}(u,v)$. Note that symmetric copulas are not adapted to most signals, as shown in the following simulations.

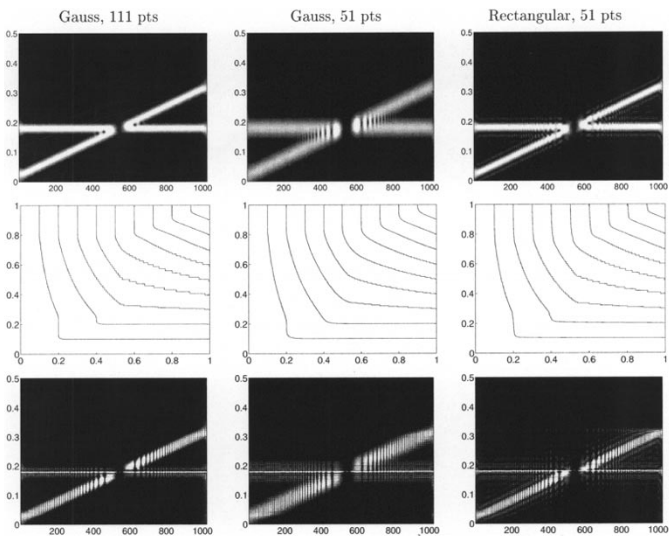

Fig. 2. (Top) Spectrograms of the data computed with three different windows. (Middle) Copulas computed from the above spectrograms. (Bottom) Positive TFDs with correct marginals computed with the copulas above.

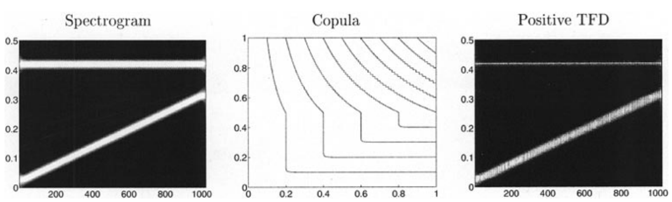

Fig. 3. Spectrogram, copula, and positive TFD with correct marginals of another signal (Gaussian window, 111 points).

V. Simulations

In this section, we display simulation results. The 1024-samples discrete signal $s(t)$ processed is composed of a chirp and a stationary tone. Its spectrogram is displayed Fig. 2. Three copulas are computed from three different spectrograms of $s(t)$ (see Fig. 2, top row). The shape of the copula is not much affected by the choice of the window h. The copulas of Fig. 2, top row, are used to compute the positive TFDs as explained in the previous section. The bottom row of Fig. 2 displays the results. As can be seen, the choice of the spectrogram window used to compute the copula influences the positive TFD in the same way as it influences the spectrogram in terms of time-frequency resolution. The positive TFD resolution is, however, better than the spectrogram resolution.

Whatever h, the copula shape is in some way a mixture between the product copula $\mathbf{\Pi}$ (related to the tone) and $\mathbf{M}$ (related to the increasing chirp). On Fig. 3, the spectrogram, copula, and positive TFD of another signal are displayed. The copula shape is different compared to the previous case: the "weights" of $\mathbf{\Pi}$ and $\mathbf{M}$ in the mixture characterize the relative location of the spectral components.

Fig. 4 displays the MCE TFD and the copula-based TFD computed from the same spectrogram for a 2048-sample real loudspeaker test signal [14]. As can be seen, the two images are very similar. The computational complexity is $139 \cdot 10^6$ flops with our method, and $42 \cdot 10^6$ flops per iteration with the iterative MCE TFD computation. In other words, the computational cost of the copula-based technique is lower as far as more than four iterations are necessary for the MCE TFD, which is often true in practice for a similar accuracy.

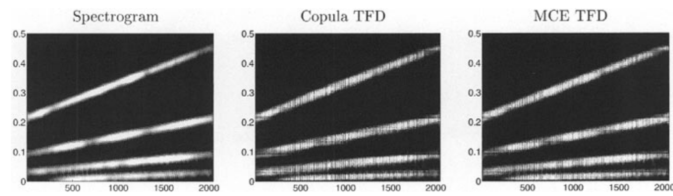

Fig. 4. Spectrogram, copula-based positive TFD, and minimum cross entropy TFD computed for the same signal. Both the MCE TFD and the copula-based TFD were computed from the same spectrogram (left image), with Gaussian window (111 points).

VI. Conclusion

In this letter, we have established connections between Cohen–Posch's positive TFDs and copulas. As a direct application, we have introduced a new noniterative technique aimed at computing positive TFDs with correct marginals. Other possible applications including signal classification/detection using copulas, and time-frequency dependence analysis will be investigated soon.

Acknowledgment

The authors would like to thank M. K. Pitt and C. Doncarli for helpful discussions.

References

[1] L. Cohen and Y. I. Zaparovanny, *J. Math. Phys.*, pp. 794–796, 1980.
[2] L. Cohen and T. Posh, "Positive time-frequency distribution functions," *IEEE Trans. Acoust., Speech Signal Processing*, vol. ASSP-33, pp. 31–38, Feb. 1985.
[3] A. Francos and M. Porat, "Analysis and synthesis of multicomponent signals using positive time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 493–504, Feb. 1999.
[4] P. Bonato, M.-S. S. Cheng, J. Gonzalez-Cueto, A. Leardini, J. O'Connor, and S. H. Roy, "Assessment of dynamic conditions can provide information about compensatory, muscle function in ACL patients," *IEEE Eng. Med. Biol.*, pp. 133–143, Nov./Dec. 2001.
[5] J. Pitton, L. Atlas, and P. Loughlin, "Applications of positive time-frequency distributions to speech processing," *IEEE Trans. Speech Audio Processing*, vol. 2, pp. 554–566, Oct. 1994.
[6] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "Construction of positive time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, pp. 2697–2705, Oct. 1994.
[7] J. R. Fonollosa, "Positive time-frequency distributions based on joint marginals constraints," *IEEE Trans. Signal Processing*, vol. 44, pp. 2086–2091, Aug. 1996.
[8] D. Groutage, "A fast algorithm for computing minimum cross-entropy positive time frequency distributions," *IEEE Trans. Signal Processing*, vol. 45, pp. 1954–1970, Aug. 1997.
[9] M. K. Emresoy and P. J. Loughlin, "Weighted least squares implementation of Cohen-Posh time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 46, pp. 753–757, Mar. 1998.
[10] L. Knockaert, "A class of positive isentropic time-frequency distributions," *IEEE Signal Processing Lett.*, vol. 9, pp. 22–25, Jan. 2002.
[11] A. Sklar, "Fonctions de répartiton ă n dimensions et leurs marges," *Pub. Inst. Stat. Univ. Paris*, vol. 8, pp. 229–231, 1959.
[12] A. Finch and A. Groblicki, "Bivariate probability densities with given margins," *Found. Phys.*, vol. 14, no. 6, pp. 549–552, 1984.
[13] *An Introduction to Copulas*, ser. Lecture Notes in Statistics, Springer, New York, 1999.
[14] M. Davy and C. Doncarli, "A new nonstationary test procedure for improved loudspeaker fault detection," *J. Audio Eng. Soc.*, vol. 50, no. 6, June 2002.