in Time-Frequency Analysis: Concepts and Methods
eds. F. Hlawatsch and F. Auger
London, UK: ISTE and Wiley, 2008
Chap. 10, pp. 279–320

Chapter 10

Time-Frequency Methods for Non-stationary Statistical Signal Processing

Abstract: Time-frequency (TF) methods can be used to analyze and process non-stationary random processes in an efficient and intuitive manner. This chapter presents some of the non-parametric methods in this area. We first discuss two different definitions of a "TF spectrum" for non-stationary processes. For the important subclass of *underspread* processes, it is demonstrated that the various TF spectra are effectively equivalent and approximately satisfy several desirable properties. Methods for estimating TF spectra are presented and studied. Finally, we discuss the use of TF spectra for processing non-stationary random processes of the underspread type. Simple formulations of quasi-optimum estimators and detectors are proposed, which generalize methods for the stationary case (such as, for instance, Wiener-type filters) to the case of underspread non-stationary random processes. These "TF estimators/detectors" have the advantage of allowing an intuitive interpretation and being numerically stable and efficient.

Keywords: underspread non-stationary random processes, statistical signal processing, time-varying spectra, Wigner-Ville spectrum, evolutionary spectrum, time-frequency analysis, non-stationary estimation, Wiener filter, non-stationary detection.

10.1. Introduction

Although most work on time-frequency (TF) concepts and methods is placed within a deterministic framework, the "TF philosophy" is also suited to non-stationary random processes. As long as the random processes considered are stationary, there

Chapter written by Franz HLAWATSCH and Gerald MATZ.

is little reason to use TF methods. In fact, for a stationary process $x(t)$, the *power spectral density*

$$\widehat{r}_x(f) \;=\; \int_{-\infty}^{\infty} r_x(\tau)\, e^{-j2\pi f\tau}\, d\tau\,, \tag{10.1}$$

where $r_x(\tau) = \mathrm{E}\{x(t+\tau)\, x^*(t)\}$ with $\mathrm{E}\{\cdot\}$ denoting the expectation operator, provides a complete description of the second-order statistical properties [PAP 91]. Due to the stationarity of $x(t)$, the spectrum $\widehat{r}_x(f)$ does not depend on time t. On the other hand, when the process $x(t)$ is *non-stationary*, it is quite evident that its spectral properties – and, thus, any relevant description of these properties – must depend on time. Consequently, we will be interested in a time-dependent power spectral representation of the $P_x(t,f)$ type. Such a spectrum can be interpreted as a TF representation of the second-order statistics of $x(t)$.

As we will demonstrate, this situation is closely related to the description of linear systems. As long as a system is time-invariant, the *transfer function*

$$\widehat{h}(f) \;=\; \int_{-\infty}^{\infty} h(\tau)\, e^{-j2\pi f\tau}\, d\tau\,, \tag{10.2}$$

where $h(\tau)$ is the impulse response, provides a complete description of the system's spectral properties [PAP 84]. On the other hand, in the case of a *time-varying* system, the spectral characteristics will depend on time, and thus we are led to consider a transfer function of the $H(t,f)$ type. This function can be interpreted as a TF representation of the system.

At this point, several fundamental questions arise:

– How can $P_x(t,f)$ and $H(t,f)$ be defined? Is the definition unique as in the stationary case?

– Can we use $P_x(t,f)$ and $H(t,f)$ as we use $\widehat{r}_x(f)$ and $\widehat{h}(f)$ in the stationary case? To illustrate what is meant by this question, let us consider, for example, the estimation of a stationary process $s(t)$ from a noisy version $s(t)+n(t)$ by means of a linear filter. The optimum filter according to the criterion of minimum mean-square error (non-causal Wiener filter) is time-invariant with its transfer function given by [PAP 91, VAN 68, POO 88, SCH 91, THE 92, WIE 49]

$$\widehat{h}(f) \;=\; \frac{\widehat{r}_s(f)}{\widehat{r}_s(f)+\widehat{r}_n(f)}\,. \tag{10.3}$$

If, on the other hand, $s(t)$ and $n(t)$ are non-stationary, it is well known that the optimum filter is time-varying [VAN 68, POO 88, SCH 91, THE 92]. Can its transfer function $H(t,f)$ be expressed in a similar manner as in (10.3), i.e., as

$$H(t,f) \;=\; \frac{P_s(t,f)}{P_s(t,f)+P_n(t,f)}\,? \tag{10.4}$$

In this chapter, we will discuss time-varying spectra and their application to the estimation and detection of non-stationary random processes. We will restrict ourselves to *non-parametric* spectra; parametric spectra will be discussed in Chapter 11. To begin with, the preparatory section 10.2 considers the TF representation of time-varying systems, which will provide the fundamentals required for our treatment of time-varying spectra. Some basic principles of non-stationary processes and the important subclass of *underspread* processes are discussed in Section 10.3.

Two broad classes of time-varying spectra are considered in Sections 10.4 and 10.5. We will see that there exist an infinite number of different definitions of a time-varying spectrum. However, we show in Section 10.6 that for an underspread process, all these different spectra provide nearly identical results and satisfy several desirable properties in an approximate manner. The estimation of time-varying spectra from a single process realization is discussed in Section 10.7.

The application of time-varying spectra to the estimation (quasi-optimum filtering) and detection of a non-stationary process is finally studied in Sections 10.8 and 10.9. For underspread processes, it is possible to develop *TF estimators* and *TF detectors* that are nearly optimum as well as numerically stable and efficient and, at the same time, allow an intuitive interpretation in the spirit of (10.4).

10.2. Time-varying systems

Since time-varying systems will occur several times in the course of this chapter, we first develop some of their basic principles. The input-output relation of a time-varying system (operator) $\boldsymbol{H}$ with kernel $h(t,t')$ is

$$y(t) \;=\; (\boldsymbol{H}x)(t) = \int_{-\infty}^{\infty} h(t,t')\,x(t')\,dt'.$$

The *generalized Weyl symbol* of a time-varying system $\boldsymbol{H}$ is defined as [KOZ 92a, KOZ 97a, MAT 98c]

$$L_{\boldsymbol{H}}^{(\alpha)}(t,f) \;:=\; \int_{-\infty}^{\infty} h^{(\alpha)}(t,\tau)\,e^{-j2\pi f\tau}\,d\tau \tag{10.5}$$

with the generalized impulse response

$$h^{(\alpha)}(t,\tau) \;:=\; h\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau\,,\,t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big), \tag{10.6}$$

where $\alpha \in \mathbb{R}$ is a parameter. For the important subclass of "underspread" time-varying systems (see below), $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ can be interpreted and used as a "time-varying transfer function" generalizing the time-invariant transfer function $\widehat{h}(f)$ defined in (10.2). In the case of a time-invariant system, $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ reduces to $\widehat{h}(f)$ and thus is independent of time t. Some special cases are the *Weyl symbol* ($\alpha = 0$), the *Za-*

deh symbol ($\alpha = 1/2$), and the *Kohn-Nirenberg symbol* ($\alpha = -1/2$) [KOZ 97a, FOL 89, JAN 89, KOZ 92b, SHE 94, MAT 98c, ZAD 50, BEL 63, KOH 65]. Due to its symmetric structure, the choice $\alpha = 0$ has some advantages over other choices of α [KOZ 92b, MAT 98c, FOL 89].

A representation that is dual to $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ is given by the *generalized spreading function* [KOZ 92a, KOZ 97a, MAT 98c]

$$D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi) := \int_{-\infty}^{\infty} h^{(\alpha)}(t,\tau)\, e^{-j2\pi\xi t}\, dt\,. \tag{10.7}$$

This representation characterizes the *TF shifts* of the input signal caused by the time-varying system $\boldsymbol{H}$. $D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)$ is a function of the time-shift (delay) variable τ and the frequency (Doppler) shift variable ξ. It can be shown that the magnitude of $D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)$ does not depend on α, so that we can write it as $|D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)| = |D_{\boldsymbol{H}}(\tau,\xi)|$. Moreover, $D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)$ is essentially the two-dimensional Fourier transform of the generalized Weyl symbol defined in (10.5).

The generalized spreading function $D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)$ is the coefficient function in an expansion of $\boldsymbol{H}$ into elementary TF shift operators $\boldsymbol{S}_{\tau,\xi}^{(\alpha)}$ defined by $\big(\boldsymbol{S}_{\tau,\xi}^{(\alpha)}\, x\big)(t) = x(t-\tau)\, e^{j2\pi\xi t} e^{j2\pi(\alpha-1/2)\xi\tau}$ [KOZ 92a, KOZ 97a, FOL 89, MAT 98c, KOZ 97b, SHE 94, BEL 63]. In fact, we have

$$(\boldsymbol{H}x)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)\, \big(\boldsymbol{S}_{\tau,\xi}^{(\alpha)}\, x\big)(t)\, d\tau\, d\xi\,. \tag{10.8}$$

Thus, for a given (τ,ξ), $|D_{\boldsymbol{H}}(\tau,\xi)|$ measures the contribution of the shifted input signal $\big(\boldsymbol{S}_{\tau,\xi}^{(\alpha)}\, x\big)(t) = x(t-\tau)\, e^{j2\pi\xi t} e^{j2\pi(\alpha-1/2)\xi\tau}$ to the output signal $(\boldsymbol{H}x)(t)$. As a consequence, the TF shifts caused by a time-varying system $\boldsymbol{H}$ are globally characterized by the effective support of $|D_{\boldsymbol{H}}(\tau,\xi)|$.

A time-varying system $\boldsymbol{H}$ is called *underspread* if it introduces only small TF shifts; it is called *overspread* otherwise. In view of (10.8), the underspread property amounts to the fact that $|D_{\boldsymbol{H}}(\tau,\xi)|$ is concentrated around the origin of the (τ,ξ) plane [KOZ 97a, MAT 98c, KOZ 97b, MAT 00b]. As a quantitative measure of the concentration of $|D_{\boldsymbol{H}}(\tau,\xi)|$, let us introduce the weighted integral

$$m_{\boldsymbol{H}}^{(\phi)} := \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\tau,\xi)\, |D_{\boldsymbol{H}}(\tau,\xi)|\, d\tau\, d\xi}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |D_{\boldsymbol{H}}(\tau,\xi)|\, d\tau\, d\xi} \;\geq\; 0\,, \tag{10.9}$$

where $\phi(\tau,\xi)$ is a weighting function that penalizes the contributions of $|D_{\boldsymbol{H}}(\tau,\xi)|$ located far from the origin, with $\phi(\tau,\xi) \geq \phi(0,0) = 0$. Important special cases of $m_{\boldsymbol{H}}^{(\phi)}$ are given by the *moments* $m_{\boldsymbol{H}}^{(k,l)}$ (where $k,l \in \mathbb{N}_0$) whose weighting function has the form $\phi(\tau,\xi) = |\tau|^k |\xi|^l$. A system $\boldsymbol{H}$ then is underspread if certain weighted integrals $m_{\boldsymbol{H}}^{(\phi)}$ and/or certain moments $m_{\boldsymbol{H}}^{(k,l)}$ are small.

The underspread property is not equivalent to the property of slow temporal variation (quasi-stationarity property). In fact, slow temporal variation is expressed by a concentration of $|D_{\boldsymbol{H}}(\tau,\xi)|$ with respect to ξ only. In contrast, the underspread property is expressed by a concentration with respect to τ and ξ jointly, where the concentrations with respect to τ and ξ are mutually interchangeable. Thus, a system with slow temporal variation can be overspread if its memory (extension of $|D_{\boldsymbol{H}}(\tau,\xi)|$ with respect to τ) is too long, and conversely a system that does not exhibit slow temporal variation can still be underspread if its memory is short enough.

10.3. Non-stationary processes

The main subject of this chapter is the analysis and processing of non-stationary random processes. The second-order statistical properties of a non-stationary process $x(t)$ are characterized by the *mean function* $m_x(t) := \mathrm{E}\{x(t)\}$ (which will generally be supposed to be zero hereafter) and the *correlation function*

$$R_x(t_1,t_2) \;:=\; \mathrm{E}\big\{x(t_1)\,x^*(t_2)\big\}\,.$$

Sometimes, we will interpret $R_x(t_1,t_2)$ as the kernel of a linear operator $\boldsymbol{R}_x$ that we will call the *correlation operator* of process $x(t)$.

There is another link between non-stationary processes and time-varying systems (operators). Under appropriate conditions, we can represent $x(t)$ as the output of a linear time-varying system $\boldsymbol{H}$ whose input is stationary white noise, denoted $n(t)$, with power spectral density $\widehat{r}_n(f) \equiv 1$ [CRA 61]:

$$x(t) \;=\; (\boldsymbol{H}n)(t) = \int_{-\infty}^{\infty} h(t,t')\,n(t')\,dt'. \tag{10.10}$$

(If $x(t)$ is stationary, $\boldsymbol{H}$ is time-invariant and the power spectral density of $x(t)$ is given by $\widehat{r}_x(f) = |\widehat{h}(f)|^2$.) System $\boldsymbol{H}$ is called an *innovations system* of the process $x(t)$. It is obtained through the factorization $\boldsymbol{H}\boldsymbol{H}^\star = \boldsymbol{R}_x$, where $\boldsymbol{R}_x$ is the correlation operator of $x(t)$ and $\boldsymbol{H}^\star$ is the adjoint of $\boldsymbol{H}$ (i.e., the system with kernel $h^\star(t,t') = h^*(t',t)$). Thus, $\boldsymbol{H}$ is a "square root" of $\boldsymbol{R}_x$. This square root is only unique to within a factor $\boldsymbol{A}$ satisfying $\boldsymbol{A}\boldsymbol{A}^\star = \boldsymbol{I}$: if $\boldsymbol{H}$ is an innovations system of $x(t)$ (thus, $\boldsymbol{H}\boldsymbol{H}^\star = \boldsymbol{R}_x$) and if $\boldsymbol{A}$ satisfies $\boldsymbol{A}\boldsymbol{A}^\star = \boldsymbol{I}$, then $\boldsymbol{H}' = \boldsymbol{H}\boldsymbol{A}$ is another innovations system of $x(t)$.

From $R_x(t_1,t_2)$ (or $\boldsymbol{R}_x$), we can calculate various time-varying spectra that will be discussed in Sections 10.4–10.6. A second-order representation whose interpretation is different from that of a spectrum is the *generalized expected ambiguity function* (GEAF) defined by [KOZ 97a, KOZ 94a]

$$\bar{A}_x^{(\alpha)}(\tau,\xi) \;:=\; \int_{-\infty}^{\infty} r_x^{(\alpha)}(t,\tau)\,e^{-j2\pi\xi t}\,dt \;=\; \mathrm{E}\big\{\big\langle x, \mathbf{S}_{\tau,\xi}^{(\alpha)} x\big\rangle\big\}\,, \tag{10.11}$$

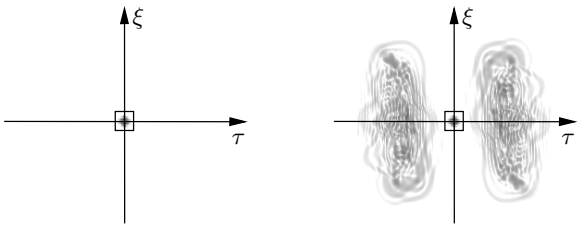

Figure 10.1. *Magnitude of the GEAF of (a) an underspread process and (b) an overspread process. The small squares around the origin have an area of 1 and thus allow an assessment of the underspread or overspread nature. The lowest contour level is at 20 dB below the maximum value of the GEAF,* $\bar{A}_x(0,0)$

with

$$r_x^{(\alpha)}(t,\tau) \;:=\; R_x\!\left(t+\left(\frac{1}{2}-\alpha\right)\tau,\, t-\left(\frac{1}{2}+\alpha\right)\tau\right), \tag{10.12}$$

where $\alpha \in \mathbb{R}$ is a parameter. Comparing (10.11) with (10.7) we note that the GEAF is the generalized spreading function of correlation operator $\boldsymbol{R}_x$, i.e.,

$$\bar{A}_x^{(\alpha)}(\tau,\xi) \;=\; D_{\boldsymbol{R}_x}^{(\alpha)}(\tau,\xi)\,.$$

Thus, $|\bar{A}_x^{(\alpha)}(\tau,\xi)|$ does not depend on α, so that it can be written $|\bar{A}_x^{(\alpha)}(\tau,\xi)| = |\bar{A}_x(\tau,\xi)|$. The interpretation of the GEAF $\bar{A}_x^{(\alpha)}(\tau,\xi)$ is that it characterizes the average correlation of all pairs of TF points separated by τ in time and by ξ in frequency [KOZ 97a, KOZ 94a].

A non-stationary process $x(t)$ is called *underspread* if all components of $x(t)$ that are sufficiently distant from each other in the TF plane (corresponding to values of τ and/or ξ that are not too small) are effectively uncorrelated; it is called *overspread* in the opposite case [KOZ 94a, KOZ 97a, MAT 00b, MAT 06]. The underspread property is exhibited by many non-stationary processes occurring in applications. In view of the interpretation of the GEAF, this property essentially states that $|\bar{A}_x(\tau,\xi)|$ is concentrated around the origin of the (τ,ξ) plane. The concentration of the GEAF can be characterized by the weighted integral

$$m_x^{(\phi)} \;:=\; \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\tau,\xi)\,|\bar{A}_x(\tau,\xi)|\,d\tau\,d\xi}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\bar{A}_x(\tau,\xi)|\,d\tau\,d\xi} \;\geq\; 0\,, \tag{10.13}$$

where $\phi(\tau,\xi)$ is a weighting function as in (10.9). We will also use the *moments* $m_x^{(k,l)}$ whose weighting function is $\phi(\tau,\xi) = |\tau|^k|\xi|^l$. A process is thus underspread if certain weighted integrals $m_x^{(\phi)}$ and/or certain moments $m_x^{(k,l)}$ are small. Figure 10.1 contrasts the GEAF of an underspread process with that of an overspread process. Some time-varying spectra of these two processes will be presented later in Section 10.6.1 (see Figures 10.3 and 10.4).

The underspread property is not equivalent to the quasi-stationarity property, which is expressed by a concentration of $|\bar{A}_x(\tau,\xi)|$ with respect to the single variable ξ. In fact, the underspread property is expressed by a concentration with respect to τ and ξ jointly, where the concentrations with respect to τ and ξ are mutually interchangeable. Consequently, it is possible that a quasi-stationary process is not underspread if its temporal correlation horizon (extension of $|\bar{A}_x(\tau,\xi)|$ with respect to τ) is too broad, and conversely a process that is not quasi-stationary can be underspread if its temporal correlation horizon is sufficiently small ("quasi-white" process). We note that definitions of limited TF correlation that are somewhat similar conceptually have been proposed and discussed in [MAL 99, MAL 98, SIL 95a, SIL 95b].

The two concepts of underspread *systems* and underspread *processes* are related because $\bar{A}_x^{(\alpha)}(\tau,\xi)$ is the generalized spreading function of $\boldsymbol{R}_x$; thus, a *process* $x(t)$ is underspread if and only if its correlation operator $\boldsymbol{R}_x$ is an underspread *system*. Moreover, there is a relationship between the TF correlation structure of $x(t)$ and the TF shifts caused by the innovations system of $x(t)$. If the innovations system $\boldsymbol{H}$ is underspread, then the correlation operator $\boldsymbol{R}_x = \boldsymbol{H}\boldsymbol{H}^\star$ is also underspread and, consequently, the process $x(t)$ is underspread as well. Conversely, if $x(t)$ is underspread, this does not imply that all innovations systems $\boldsymbol{H}$ are underspread, but there is always a particular $\boldsymbol{H}$ that is underspread.

10.4. TF analysis of non-stationary processes – type I spectra

As we have already observed, any spectrum of a non-stationary process must depend on time t in addition to frequency f. To define a time-varying spectrum within the non-parametric framework studied here, two different approaches are possible. These approaches lead to two classes of spectra that will be termed "type I" and "type II". Type I spectra are the subject of this section, while type II spectra will be introduced and studied in the next section.

10.4.1. *Generalized Wigner-Ville spectrum*

First, we consider an important family of type I time-varying spectra that is known as *generalized Wigner-Ville spectrum* (GWVS) [FLA 89, FLA 97, KOZ 94a, MAT 06]. The GWVS is a simple extension of the power spectral density $\widehat{r}_x(f)$ defined in (10.1):

$$\overline{W}_x^{(\alpha)}(t,f) := \int_{-\infty}^{\infty} r_x^{(\alpha)}(t,\tau)\, e^{-j2\pi f\tau}\, d\tau\,, \tag{10.14}$$

where $r_x^{(\alpha)}(t,\tau)$ with $\alpha \in \mathbb{R}$ has been defined in (10.12). When process $x(t)$ is stationary (in the wide sense), $\overline{W}_x^{(\alpha)}(t,f)$ becomes equal to the power spectral density $\widehat{r}_x(f)$ and, thus, independent of time t.

A comparison with (10.5) shows that the GWVS is the generalized Weyl symbol of the correlation operator $\boldsymbol{R}_x$:

$$\overline{W}_x^{(\alpha)}(t,f) \;=\; L_{\boldsymbol{R}_x}^{(\alpha)}(t,f)\,.$$

It is also the two-dimensional Fourier transform of the GEAF defined in (10.11):

$$\overline{W}_x^{(\alpha)}(t,f) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{A}_x^{(\alpha)}(\tau,\xi)\, e^{-j2\pi(f\tau-t\xi)}\, d\tau\, d\xi\,. \tag{10.15}$$

This relation generalizes the Wiener-Khintchine relation (10.1) to non-stationary processes. Finally, under appropriate conditions [FLA 99], $\overline{W}_x^{(\alpha)}(t,f)$ can be written as the expectation of a TF signal representation that is known under the name of *generalized Wigner-Ville distribution* and defined as [CLA 80, JAN 82, HLA 97]

$$W_x^{(\alpha)}(t,f) \;:=\; \int_{-\infty}^{\infty} x\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau\Big)\, x^*\Big(t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi f\tau}\, d\tau\,. \tag{10.16}$$

Important special cases of the GWVS are the *Wigner-Ville spectrum* ($\alpha = 0$) [MAR 70, MAR 85, FLA 89, KOZ 94a, MAT 06, FLA 97], which has certain advantages due to its symmetric structure, and the *Rihaczek spectrum* ($\alpha = 1/2$) [FLA 89, FLA 97, RIH 68]. The Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$ is always real – although its positivity is not guaranteed [FLA 86b, FLA 99, MAT 00b, MAT 06] – whereas other spectra $\overline{W}_x^{(\alpha)}(t,f)$ are complex-valued in general. However, we will see later on that the problems of interpretation and representation caused by negative or even complex values effectively disappear in the underspread case.

10.4.2. *TF correlations and statistical cross-terms*

The definition of some of the type I spectra can be justified by the effects that the TF correlations of a non-stationary process have on its GWVS. For an underspread process (TF correlations with small horizon), the GEAF is well concentrated around the origin of the (τ,ξ) plane. It then follows from (10.15) that the GWVS is a low-pass function, i.e., a smooth function with slow variations. On the other hand, if the process is *overspread* (TF correlations with large horizon), its GEAF has components located at some distance away from the origin of the (τ,ξ) plane. Thus, it follows from relation (10.15) that the GWVS will contain oscillatory and partly negative components. These components can be interpreted as "statistical cross-terms"; they are indicative of the presence of TF correlations in the process [THO 93, MAT 06, MAT 00b, KOZ 97a].

To illustrate the mechanism and geometry of these statistical cross-terms, we study the basic example of a process with two components $x(t) = x_1(t) + x_2(t)$. The components are given by

$$x_1(t) \;=\; a_1\, x_0(t-t_1)\, e^{j2\pi f_1 t}, \qquad x_2(t) \;=\; a_2\, x_0(t-t_2)\, e^{j2\pi f_2 t},$$

where $x_0(t)$ is a process whose GWVS is assumed to be localized around the origin $(t,f)=(0,0)$ of the TF plane and a_1, a_2 are *random* factors that are uncorrelated with $x_0(t)$. We note that components $x_1(t)$ and $x_2(t)$ are localized around the TF points (t_1,f_1) and (t_2,f_2), respectively, and that they are correlated if and only if a_1 and a_2 are correlated. If $x_1(t)$ and $x_2(t)$ are correlated, we can then say that "the TF points (t_1,f_1) and (t_2,f_2) are correlated".

Let us now consider the Wigner-Ville spectrum (GWVS with $\alpha=0$) $\overline{W}_x^{(0)}(t,f)$ of process $x(t)$. It consists of (i) the two terms

$$\overline{W}_{x_1}^{(0)}(t,f) \;=\; \mathrm{E}\{|a_1|^2\}\,\overline{W}_{x_0}^{(0)}(t-t_1,f-f_1)$$
$$\overline{W}_{x_2}^{(0)}(t,f) \;=\; \mathrm{E}\{|a_2|^2\}\,\overline{W}_{x_0}^{(0)}(t-t_2,f-f_2)$$

which are correctly localized around (t_1,f_1) and (t_2,f_2), respectively, and (ii) a "statistical cross-term" given by

$$\Psi_{12}(t,f) \;=\; c\left(t-\frac{t_1+t_2}{2}, f-\frac{f_1+f_2}{2}\right)$$

with

$$c(t,f) \;:=\; 2|r|\,\overline{W}_{x_0}^{(0)}(t,f)\,\cos\bigl(2\pi\bigl[(f_1-f_2)t-(t_1-t_2)f\bigr]+\varphi\bigr)\,,$$

where $r:=\mathrm{E}\{a_1\,a_2^*\}$ and $\varphi:=2\pi\,(t_1-t_2)\,(f_1-f_2)+\arg\{r\}$. Thus, this cross-term $\Psi_{12}(t,f)$ is localized around the middle point $\left(\frac{t_1+t_2}{2},\frac{f_1+f_2}{2}\right)$ between points (t_1,f_1) and (t_2,f_2); it oscillates, taking on negative as well as positive values. Its amplitude is proportional to $|r|=\bigl|\mathrm{E}\{a_1\,a_2^*\}\bigr|$ and thus constitutes a direct measure of the degree of correlation of the components $x_1(t)$ and $x_2(t)$ or of the TF points (t_1,f_1) and (t_2,f_2). It follows that $\overline{W}_x^{(0)}(t,f)$ will contain large statistical cross-terms if (and only if) process $x(t)$ has strong TF correlations, i.e., if it is *overspread*.

Figure 10.2 gives an example illustrating these results. We can verify, in particular, that the amplitude of the statistical cross-term increases with the degree of TF correlation. We note that the TF geometry of this cross-term is that of the cross-terms of the Wigner-Ville distribution studied in Chapter 5 (this is not surprising, as the Wigner-Ville spectrum is the expectation of the Wigner-Ville distribution). This geometry can be generalized to the case $\alpha\neq 0$ [HLA 97].

10.4.3. *TF smoothing and type I spectra*

While statistical cross-terms are indicative of TF correlations, they very often have the disadvantage of covering and, consequently, masking the other terms of the Wigner-Ville spectrum that characterize more particularly the signal components of the analyzed process. Therefore, in the overspread case where the amplitude of statistical cross-terms is large, it is often desirable to attenuate these terms. Due to their

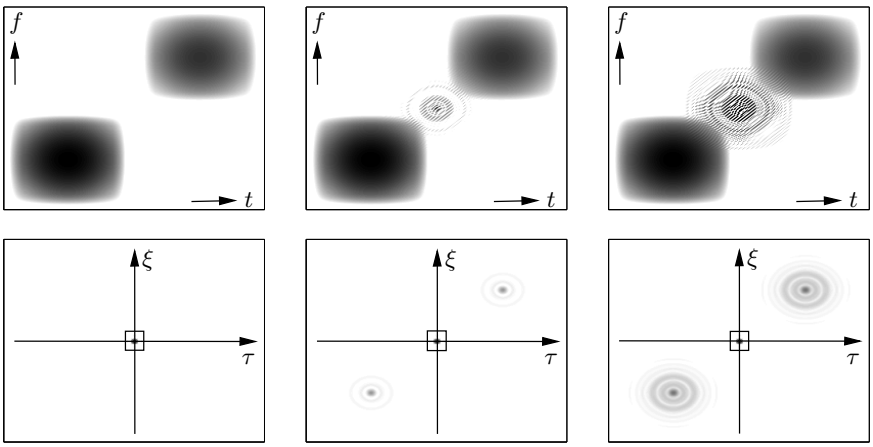

Figure 10.2. *Statistical cross-term in the Wigner-Ville spectrum of a process with two components for (a) zero correlation, (b) medium correlation, (c) maximum correlation of the two components. Above: Wigner-Ville spectrum, below: GEAF magnitude. The small squares around the origin have an area of 1 and thus allow an assessment of the underspread or overspread nature*

oscillatory behavior, cross-terms can be easily attenuated or even suppressed by a *TF smoothing* that is expressed mathematically by a two-dimensional convolution:

$$\begin{aligned} \bar{C}_x(t,f) \;&:=\; \phi_{\text{t-f}}(t,f) \;*_t *_f \; \overline{W}_x^{(0)}(t,f) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{\text{t-f}}(t-t',f-f')\; \overline{W}_x^{(0)}(t',f')\, dt'df', \end{aligned} \tag{10.17}$$

where $\phi_{\text{t-f}}(t,f)$ is a smooth function, of low-pass type.

If we now allow that $\phi_{\text{t-f}}(t,f)$ is an arbitrary function, not necessarily of the low-pass type, then expression (10.17) defines a class of time-varying spectra that we will here refer to as *type I spectra*. This class has an important theoretical interpretation: it consists of all the spectra that depend linearly on $R_x(t_1,t_2)$ and are *covariant to TF shifts* [LOY 68, FLA 97, AMI 92]:

$$\tilde{x}(t) \;=\; x(t-t_0)\, e^{j2\pi f_0 t} \quad \Rightarrow \quad \bar{C}_{\tilde{x}}(t,f) \;=\; \bar{C}_x(t-t_0, f-f_0)\,. \tag{10.18}$$

Under appropriate conditions, the class of type I spectra can also be written as the expectation of *Cohen's class* of TF signal representations [FLA 97] (for Cohen's class, see Chapters 1 and 5 of this book and [COH 95, FLA 99, HLA 92]). The class of type I spectra contains, in particular, all the members of the GWVS family. The GWVS members are an example of the case where $\phi_{\text{t-f}}(t,f)$ is not of the low-pass type and, thus, convolution (10.17) is not a smoothing. Other elements of the class of type I

spectra are the spectra of Page [PAG 52] and Levin [LEV 64] and the "physical spectrum" (expectation of the spectrogram) [MAR 70, FLA 89, FLA 97]. All these particular spectra are obtained by specific choices of the "kernel" $\phi_{\text{t-f}}(t,f)$ in (10.17). We finally note that we can define the same class of type I spectra using "generators" other than $\overline{W}_x^{(0)}(t,f)$ in (10.17), for example $\overline{W}_x^{(\alpha)}(t,f)$ with $\alpha \neq 0$.

10.4.4. *Properties of type I spectra*

All type I spectra satisfy covariance property (10.18). Other desirable properties will be satisfied if the kernel $\phi_{\text{t-f}}(t,f)$ satisfies corresponding constraints. For example, the normalization constraint

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_{\text{t-f}}(t,f)\,dt\,df \;=\; 1 \tag{10.19}$$

is necessary and sufficient for $\bar{C}_x(t,f)$ to preserve the mean energy of the process $x(t)$:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\bar{C}_x(t,f)\,dt\,df \;=\; \bar{E}_x\,,$$

with $\bar{E}_x := \int_{-\infty}^{\infty} R_x(t,t)\,dt = \mathrm{E}\{\|x\|^2\}$.

It is often simpler to formulate such constraints in terms of the two-dimensional Fourier transform $\phi_{\text{d-D}}(\tau,\xi) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_{\text{t-f}}(t,f)\,e^{-j2\pi(\xi t-\tau f)}dt\,df$ of the kernel $\phi_{\text{t-f}}(t,f)$. The kernel $\phi_{\text{d-D}}(\tau,\xi)$ is a function of the time-shift (delay) variable τ and the frequency (Doppler) shift variable ξ. For example, normalization constraint (10.19) is then expressed by $\phi_{\text{d-D}}(0,0) = 1$. Table 10.1 lists some basic properties along with their constraints on $\phi_{\text{d-D}}(\tau,\xi)$. We note, in particular, that the statistical cross-terms in $\bar{C}_x(t,f)$ are attenuated if and only if $\phi_{\text{d-D}}(\tau,\xi)$ is concentrated around the origin, since this corresponds to a smooth kernel $\phi_{\text{t-f}}(t,f)$. In Section 10.6.2, we will see that the properties of Table 10.1 are typically satisfied by type I spectra in an approximate fashion if the analyzed process is underspread.

10.5. TF analysis of non-stationary processes – type II spectra

An alternative to type I spectra is offered by the class of type II spectra, which is based on the innovations representation (10.10). First of all, we will study an important family of type II spectra.

10.5.1. *Generalized evolutionary spectrum*

To justify the definition of this family, let us recall that in the case of a stationary process, the power spectral density equals the squared magnitude of the transfer function of innovations system $\boldsymbol{H}$ (which is time-invariant): $\widehat{r}_x(f) = |\widehat{h}(f)|^2$. Extend-

Property	*Constraint*
covariance to TF shifts: $\tilde{x}(t) = x(t-t_0)\, e^{j2\pi f_0 t} \;\Rightarrow\; \bar{C}_{\tilde{x}}(t,f) = \bar{C}_x(t-t_0, f-f_0)$	—
real-valued: $\bar{C}_x(t,f) = \bar{C}_x^*(t,f)$	$\phi_{\text{d-D}}(\tau,\xi) = \phi_{\text{d-D}}^*(-\tau,-\xi)$
preservation of mean energy: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{C}_x(t,f)\, dt\, df = \bar{E}_x$	$\phi_{\text{d-D}}(0,0) = 1$
time marginal: $\int_{-\infty}^{\infty} \bar{C}_x(t,f)\, df = R_x(t,t) = \mathrm{E}\{\lvert x(t)\rvert^2\}$	$\phi_{\text{d-D}}(0,\xi) \equiv 1$
frequency marginal: $\int_{-\infty}^{\infty} \bar{C}_x(t,f)\, dt = R_{\hat{x}}(f,f) = \mathrm{E}\{\lvert\hat{x}(f)\rvert^2\}$	$\phi_{\text{d-D}}(\tau,0) \equiv 1$
Moyal-type property: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{C}_x(t,f)\, \bar{C}_y^*(t,f)\, dt\, df$ $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(t_1,t_2)\, R_y^*(t_1,t_2)\, dt_1\, dt_2$	$\lvert\phi_{\text{d-D}}(\tau,\xi)\rvert \equiv 1$
attenuation of statistical cross-terms	$\phi_{\text{d-D}}(\tau,\xi)$ concentrated around $(0,0)$

Table 10.1. *Some properties of type I spectra and the corresponding constraints on* $\phi_{\text{d-D}}(\tau,\xi)$

ing this expression to the non-stationary case, we define the *generalized evolutionary spectrum* (GES) as [MAT 97, MAT 06]

$$G_x^{(\alpha)}(t,f) := \left|L_{\boldsymbol{H}}^{(\alpha)}(t,f)\right|^2,$$

where $\boldsymbol{H}$ is an innovations system of the non-stationary process $x(t)$. In this definition, the generalized Weyl symbol $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ (see (10.5)) takes the place of the transfer function $\widehat{h}(f)$. Evidently, $G_x^{(\alpha)}(t,f)$ is always non-negative. For the special case of a stationary process, we can always choose a stationary $\boldsymbol{H}$; then $L_{\boldsymbol{H}}^{(\alpha)}(t,f) = \widehat{h}(f)$ and the GES becomes equal to the power spectral density: $G_x^{(\alpha)}(t,f) = |\widehat{h}(f)|^2 = \widehat{r}_x(f)$.

The GES $G_x^{(\alpha)}(t,f) = \left|L_{\boldsymbol{H}}^{(\alpha)}(t,f)\right|^2$ depends not only on the parameter $\alpha \in \mathbb{R}$ but also on the choice of the innovations system $\boldsymbol{H}$, which is not unique for a given process $x(t)$ (i.e., for a given $\boldsymbol{R}_x$). This latter dependence may be regarded as a disadvantage of the GES with respect to the GWVS. Furthermore, in contrast to the GWVS, $G_x^{(\alpha)}(t,f)$ is not related to the correlation $r_x(t_1,t_2)$ or $\boldsymbol{R}_x$ by an invertible mapping. Special cases of the GES are the *evolutionary spectrum* ($\alpha = 1/2$) [PRI 65, PRI 81, RIE 93, KAY 94] and the *transitory evolutionary spectrum* ($\alpha = -1/2$) [DET 94, MAT 97]. Finally, the *Weyl spectrum* is obtained by taking $\alpha = 0$

and choosing $\boldsymbol{H}$ as the *positive (semi-)definite* square root of operator $\boldsymbol{R}_x$ (which is unique). This has certain advantages over other choices of α and $\boldsymbol{H}$ [MAT 97].

10.5.2. *TF smoothing and type II spectra*

When the process $x(t)$ is overspread, the innovations system $\boldsymbol{H}$ is necessarily also overspread. Therefore, the generalized spreading function $D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)$ contains components that are located some distance away from the origin of the (τ,ξ) plane. From this it follows that $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$, which is the two-dimensional Fourier transform of $D_{\boldsymbol{H}}^{(\alpha)}(\tau,\xi)$, contains oscillatory components. These components correspond to "statistical cross-terms" in $G_x^{(\alpha)}(t,f) = \left|L_{\boldsymbol{H}}^{(\alpha)}(t,f)\right|^2$. We note, as an important difference from the GWVS, that the statistical cross-terms of the GES are always positive and thus cannot be attenuated by a smoothing of the GES. However, it is possible to smooth $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ before taking the squared magnitude. For the case $\alpha = 0$, we obtain

$$\begin{aligned}\widetilde{G}_x(t,f) &:= \left|\phi_{\text{t-f}}(t,f) \;*_t *_f\; L_{\boldsymbol{H}}^{(0)}(t,f)\right|^2 \\ &= \left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{\text{t-f}}(t-t',f-f')\, L_{\boldsymbol{H}}^{(0)}(t',f')\, dt'df'\right|^2 .\end{aligned} \tag{10.20}$$

If we allow $\phi_{\text{t-f}}(t,f)$ to be an arbitrary function, not necessarily of the low-pass type, this expression defines a class of time-varying spectra that will be called the class of *type II spectra*. All members of the GES family belong to this class (they are, by the way, an example where kernel $\phi_{\text{t-f}}(t,f)$ is not of the low-pass type, and thus the convolution with $\phi_{\text{t-f}}(t,f)$ is not a smoothing). The same class of type II spectra can also be defined by using "generators" other than $L_{\boldsymbol{H}}^{(0)}(t,f)$ in (10.20), for example $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ with $\alpha \neq 0$.

A type II spectrum preserves the mean energy, i.e.,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \widetilde{G}_x(t,f)\, dt\, df \;=\; \bar{E}_x\,,$$

if and only if the two-dimensional Fourier transform of kernel $\phi_{\text{t-f}}(t,f)$, $\phi_{\text{d-D}}(\tau,\xi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{\text{t-f}}(t,f)\, e^{-j2\pi(\xi t-\tau f)} dt\, df$, satisfies $|\phi_{\text{d-D}}(\tau,\xi)| \equiv 1$. Other desirable properties are more difficult to characterize using explicit constraints, since their validity also depends on the choice of innovations system $\boldsymbol{H}$. However, for the important case of an underspread process, many properties will typically be satisfied in an approximate manner (see Section 10.6.2).

10.6. Properties of the spectra of underspread processes

As we have seen in the previous section, there is an infinite variety of type I and II time-varying spectra. In general, these spectra can have very diverse properties, and

they can provide very different results for a given process. This, however, is no longer true when the processes are underspread.

10.6.1. *Approximate equivalences*

We first show that for an underspread process, the spectra tend to provide approximately equivalent results. Thus, in the underspread case, the choice of the "correct" spectrum is not too critical.

10.6.1.1. *Approximate equivalence of type I spectra*

Let $\bar{C}_x^{(1)}(t,f)$ and $\bar{C}_x^{(2)}(t,f)$ be two type I spectra with kernels $\phi_{\text{d-D}}^{(1)}(\tau,\xi)$ and $\phi_{\text{d-D}}^{(2)}(\tau,\xi)$, respectively. It can be shown that the difference between these two spectra is bounded [MAT 06, MAT 00b]:

$$\left|\bar{C}_x^{(1)}(t,f) - \bar{C}_x^{(2)}(t,f)\right| \;\leq\; \|\bar{A}_x\|_1 \, m_x^{(\phi)}, \tag{10.21}$$

where $m_x^{(\phi)}$ is defined as in (10.13) with the weighting function given by $\phi(\tau,\xi) = |\phi_{\text{d-D}}^{(1)}(\tau,\xi) - \phi_{\text{d-D}}^{(2)}(\tau,\xi)|$, and $\|\bar{A}_x\|_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\bar{A}_x(\tau,\xi)|\, d\tau\, d\xi$. Let us assume that $\phi_{\text{d-D}}^{(1)}(0,0) = \phi_{\text{d-D}}^{(2)}(0,0) = 1$ (thus, the spectra preserve the mean energy). Then $\phi(0,0) = 0$, and usually $|\phi(\tau,\xi)|$ will still be small in the immediate vicinity of $(0,0)$. Consequently, for an underspread $x(t)$ (which means that $\bar{A}_x(\tau,\xi)$ is concentrated around $(0,0)$), $m_x^{(\phi)}$ is small and it then follows from the bound (10.21) that the two spectra are roughly equal:

$$\bar{C}_x^{(1)}(t,f) \;\approx\; \bar{C}_x^{(2)}(t,f)\,.$$

As a special case, let us consider the difference between two GWVS with different values of α. From (10.21), we can derive the bound [MAT 06, MAT 00b]

$$\left|\overline{W}_x^{(\alpha_1)}(t,f) - \overline{W}_x^{(\alpha_2)}(t,f)\right| \;\leq\; 2\pi|\alpha_1-\alpha_2|\,\|\bar{A}_x\|_1\, m_x^{(1,1)}.$$

We recall from Section 10.3 that $m_x^{(1,1)}$ uses the weighting function $\phi(\tau,\xi) = |\tau\xi|$. Thus, $m_x^{(1,1)}$ is small if the GEAF $\bar{A}_x^{(\alpha)}(\tau,\xi)$ is concentrated around the τ axis and/or the ξ axis, which is a form of the underspread property; in this case, we will have $\overline{W}_x^{(\alpha_1)}(t,f) \approx \overline{W}_x^{(\alpha_2)}(t,f)$.

10.6.1.2. *Approximate equivalence of type II spectra*

Similar results exist for the case of type II spectra. Let us consider two type II spectra $\widetilde{G}_x^{(1)}(t,f)$ and $\widetilde{G}_x^{(2)}(t,f)$ with respective kernels $\phi_{\text{d-D}}^{(1)}(\tau,\xi)$ and $\phi_{\text{d-D}}^{(2)}(\tau,\xi)$, based on the *same* innovations system $\boldsymbol{H}$. It can be shown [MAT 06, MAT 00b] that the difference between $\widetilde{G}_x^{(1)}(t,f)$ and $\widetilde{G}_x^{(2)}(t,f)$ is bounded as

$$\left|\widetilde{G}_x^{(1)}(t,f) - \widetilde{G}_x^{(2)}(t,f)\right| \;\leq\; 2\,\|D_{\boldsymbol{H}}\|_1^2\, m_{\boldsymbol{H}}^{(\phi)}, \tag{10.22}$$

where the weighting function in $m_{\boldsymbol{H}}^{(\phi)}$ (see (10.9)) is given by $\phi(\tau,\xi) = |\phi_{\text{d-D}}^{(1)}(\tau,\xi) - \phi_{\text{d-D}}^{(2)}(\tau,\xi)|$. Therefore, if $\boldsymbol{H}$ is chosen underspread so that $m_{\boldsymbol{H}}^{(\phi)}$ is small (which is possible if and only if $x(t)$ is underspread), the two spectra are roughly equal:

$$\widetilde{G}_x^{(1)}(t,f) \approx \widetilde{G}_x^{(2)}(t,f)\,.$$

In particular, the following bound on the difference between two GES can be derived from (10.22) [MAT 06, MAT 00b]:

$$\left|G_x^{(\alpha_1)}(t,f) - G_x^{(\alpha_2)}(t,f)\right| \leq 4\pi|\alpha_1-\alpha_2|\,\|D_{\boldsymbol{H}}\|_1^2\, m_{\boldsymbol{H}}^{(1,1)}.$$

Thus, if $m_{\boldsymbol{H}}^{(1,1)}$ is small, we have $G_x^{(\alpha_1)}(t,f) \approx G_x^{(\alpha_2)}(t,f)$.

10.6.1.3. *Approximate equivalence of type I and II spectra*

Slightly more complicated bounds [MAT 06, MAT 00b] show that for an underspread process, even a type I spectrum and a type II spectrum tend to be roughly equal, i.e.,

$$\bar{C}_x(t,f) \approx \widetilde{G}_x(t,f)\,,$$

provided that the innovations system $\boldsymbol{H}$ used in $\widetilde{G}_x(t,f)$ is chosen as an underspread system.

10.6.1.4. *Numerical examples*

Figure 10.3 depicts several type I and type II spectra, with and without smoothing, for the underspread process[1] whose GEAF has been shown in Figure 10.1(a). We note that all these spectra provide essentially equal results, and that they are fairly smooth (non-oscillatory, low-pass type) functions.

In Figure 10.4, these results are contrasted with the results obtained by the same spectra for the overspread process whose GEAF has been shown in Figure 10.1(b). Some of these results are very different. In fact, the GWVS and GES shown in Figures 10.4(a),(b),(d),(e) contain statistical cross-terms of an oscillatory shape. These terms indicate strong correlations between the components "T" and "F" while partially or completely masking the energy-bearing structures ("T" and "F"). In contrast, in the smoothed spectra shown in Figures 10.4(c) and (f), these statistical cross-terms are effectively suppressed, which causes the energy-bearing structures to be better visible. On the other hand, there is no indication of the correlations between the components "T" and "F" and, again due to the smoothing used, the energy-bearing terms are represented with a lower resolution.

1. This process was generated using the TF synthesis technique introduced in [HLA 95].

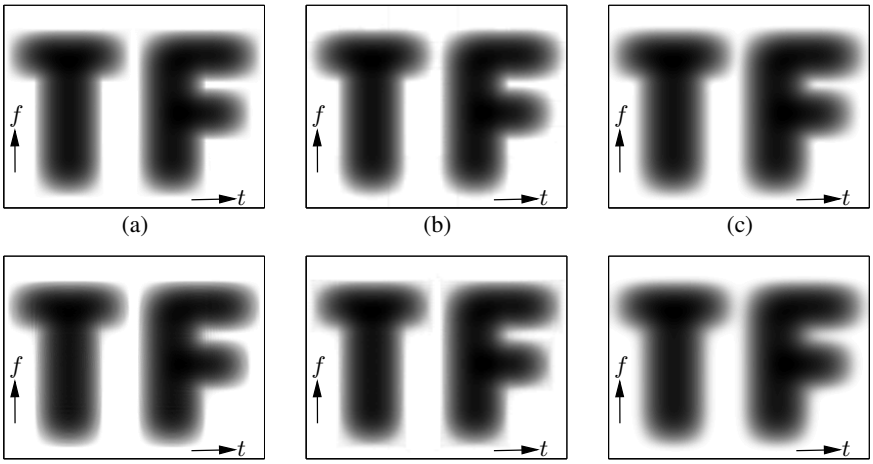

Figure 10.3. *Some type I and type II spectra of an underspread process: (a) Wigner-Ville spectrum* $\overline{W}_x^{(0)}(t,f)$*, (b) real part of the Rihaczek spectrum* $\mathrm{Re}\{\overline{W}_x^{(1/2)}(t,f)\}$*, (c) type I spectrum with smoothing, (d) Weyl spectrum* $G_x^{(0)}(t,f)$*, (e) evolutionary spectrum* $G_x^{(1/2)}(t,f)$ *(also transitory evolutionary spectrum* $G_x^{(-1/2)}(t,f)$ *because the innovations system is positive semi-definite [MAT 97]), (f) type II spectrum with smoothing. The signal's duration is 256 samples*

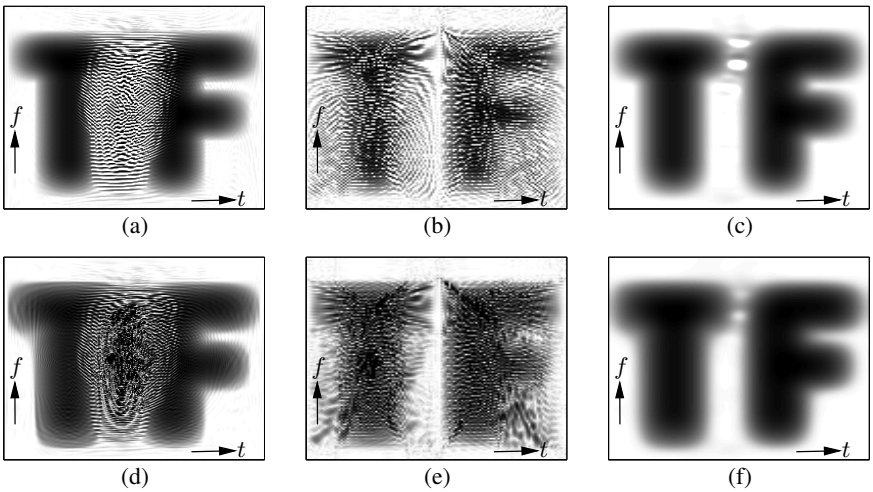

Figure 10.4. *Same spectra as in Figure 10.3, but for an overspread process*

10.6.2. *Approximate properties*

The underspread case is also remarkable because for underspread processes, the various spectra tend to satisfy several desirable properties in an approximate manner, even if they do not satisfy them in the general case.

10.6.2.1. *Approximate properties of type I spectra*

Let $\bar{C}_x(t,f)$ be a type I spectrum with kernel $\phi_{\text{d-D}}(\tau,\xi)$. We decompose $\bar{C}_x(t,f)$ as

$$\bar{C}_x(t,f) \;=\; \bar{C}_x^{+}(t,f) \;+\; \bar{C}_x^{-}(t,f) \;+\; j\,\mathrm{Im}\big\{\bar{C}_x(t,f)\big\}\,,$$

where $\bar{C}_x^{\pm}(t,f) := \frac{1}{2}\big[\mathrm{Re}\big\{\bar{C}_x(t,f)\big\} \pm \big|\mathrm{Re}\big\{\bar{C}_x(t,f)\big\}\big|\big]$ is the positive real part ($+$ sign) or the negative real part ($-$ sign) of $\bar{C}_x(t,f)$. It can then be shown [MAT 06, MAT 00b] that the imaginary part of $\bar{C}_x(t,f)$ is bounded as

$$\big|\mathrm{Im}\{\bar{C}_x(t,f)\}\big| \;\leq\; \frac{1}{2}\,\|\bar{A}_x\|_1\, m_x^{(\phi)},$$

where $\phi(\tau,\xi) = |\phi_{\text{d-D}}(\tau,\xi) - \phi^{*}_{\text{d-D}}(-\tau,-\xi)|$. Thus, for $x(t)$ underspread such that $m_x^{(\phi)}$ is small, $\bar{C}_x(t,f)$ will be almost real. Furthermore, the negative real part of $\bar{C}_x(t,f)$ is bounded as[2] [MAT 06, MAT 00b]

$$\big|\bar{C}_x^{-}(t,f)\big| \;\leq\; \|\bar{A}_x\|_1 \inf_{\mathbf{C}\geq\mathbf{0}} \big\{m_x^{(\phi_{\mathbf{C}})}\big\}\,,$$

with $\phi_{\mathbf{C}}(\tau,\xi) = \big|\phi_{\text{d-D}}(\tau,\xi) - D^{(0)}_{\mathbf{C}}(\tau,\xi)\big|$. Consequently, if $x(t)$ is underspread such that $\inf_{\mathbf{C}\geq\mathbf{0}}\big\{m_x^{(\phi_{\mathbf{C}})}\big\}$ is small, $\bar{C}_x^{-}(t,f)$ will be almost zero. These results show that type I spectra of underspread processes are approximately real and positive:

$$\bar{C}_x(t,f) \;\approx\; \bar{C}_x^{+}(t,f)\,.$$

Other properties that are typically satisfied in an approximate manner in the underspread case are the *marginal properties* (see Table 10.1)

$$\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,df \;\approx\; R_x(t,t) \;=\; \mathrm{E}\big\{|x(t)|^2\big\}\,,$$

$$\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,dt \;\approx\; R_{\widehat{x}}(f,f) \;=\; \mathrm{E}\big\{|\widehat{x}(f)|^2\big\}\,.$$

In fact, it can be shown that [MAT 06, MAT 00b]

$$\left|\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,df - R_x(t,t)\right| \;\leq\; \lambda_x \quad \text{and} \quad \left|\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,dt - R_{\widehat{x}}(f,f)\right| \;\leq\; \theta_x\,,$$

2. The notation $\mathbf{C} \geq \mathbf{0}$ means that operator $\mathbf{C}$ is positive semi-definite.

where the bounds λ_x and θ_x are degenerate forms of $m_x^{(\phi)}$ defined as $\lambda_x := \int_{-\infty}^{\infty} |1 - \phi_{\text{d-D}}(0,\xi)|\,|\bar{A}_x(0,\xi)|\,d\xi$ and $\theta_x := \int_{-\infty}^{\infty} |1 - \phi_{\text{d-D}}(\tau,0)|\,|\bar{A}_x(\tau,0)|\,d\tau$. Thus, for an underspread process where λ_x and θ_x are small, the marginal properties will be approximately satisfied. Finally, other bounds indicate that, typically, type I spectra of two processes that are jointly underspread[3] approximately satisfy a *Moyal-type property* [MAT 06, MAT 00b]:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,\bar{C}_y^*(t,f)\,dt\,df \;\approx\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(t_1,t_2)\,R_y^*(t_1,t_2)\,dt_1 dt_2\,.$$

10.6.2.2. *Approximate properties of type II spectra*

For type II spectra, it is generally more difficult to develop bounds on the various approximation errors; in addition, these bounds depend on the innovations system $\boldsymbol{H}$ [MAT 06, MAT 00b]. However, for an underspread $\boldsymbol{H}$ (which presupposes an underspread process), also type II spectra tend to approximately satisfy our desirable properties. This can be explained quite simply by the fact that for an underspread $\boldsymbol{H}$, type II spectra are roughly equal to type I spectra (see Section 10.6.1.3).

10.7. Estimation of time-varying spectra

Sometimes we need to estimate[4] a time-varying type I or II spectrum from a single realization of process $x(t)$. Such an estimation can prove difficult since the non-stationarity of process $x(t)$ does not allow us to perform an extended temporal smoothing (averaging). However, in the underspread case, the spectra are typically smooth themselves; therefore, a TF smoothing can be employed to reduce the variance of the estimate without incurring a high bias.

10.7.1. *A class of estimators*

Most time-varying spectral estimators proposed so far [MAR 85, FLA 97, FLA 89, FLA 99, KOZ 94b, SAY 95b] are TF signal representations of Cohen's class [COH 95, FLA 99, HLA 92]. This class consists of all TF signal representations with the following two properties: (i) quadratic dependence on the signal analyzed and (ii) covariance to TF shifts (see Chapters 1 and 5). Cohen's class is the "deterministic counterpart" of the class of type I spectra, and we will thus concentrate on the estimation of these

3. Two non-stationary processes $x(t)$ and $y(t)$ are called *jointly underspread* if their GEAFs $\bar{A}_x^{(\alpha)}(\tau,\xi)$ and $\bar{A}_y^{(\alpha)}(\tau,\xi)$ are concentrated in the *same* region $\mathcal{S}$ of area $|\mathcal{S}| \ll 1$ around the origin of the (τ,ξ) plane. For example, a quasi-stationary process and a quasi-white process may be underspread individually but not jointly.

4. We note at this point that any estimator of a time-varying spectrum reversibly related to the correlation $R_x(t_1,t_2)$ directly leads to a correlation estimator.

spectra. However, these estimators can also be used for estimating type II spectra, due to the approximate equivalence of type I and type II spectra in the underspread case.

Let us consider a type I spectrum $\bar{C}_x(t,f)$ with kernel $\phi_{\text{t-f}}(t,f)$. A single realization of the process $x(t)$ is observed. Estimators of $\bar{C}_x(t,f)$ belonging to Cohen's class are given by the expression

$$\begin{aligned}\hat{C}_x(t,f) &:= \hat{\phi}_{\text{t-f}}(t,f) \;*_t *_f \; W_x^{(0)}(t,f) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\phi}_{\text{t-f}}(t-t', f-f')\; W_x^{(0)}(t',f')\, dt'df',\end{aligned} \tag{10.23}$$

where $W_x^{(0)}(t,f)$ is the Wigner-Ville distribution of the realization $x(t)$ (see (10.16)) and $\hat{\phi}_{\text{t-f}}(t,f)$ is a kernel that is generally different from the kernel $\phi_{\text{t-f}}(t,f)$ defining $\bar{C}_x(t,f)$. Let $\hat{\boldsymbol{C}}$ be the linear operator such that $\hat{\phi}_{\text{t-f}}(t,f) = L^{(0)}_{\hat{\boldsymbol{C}}}(-t,-f)$. We can then write the estimator as

$$\hat{C}_x(t,f) \;=\; \big\langle \hat{\boldsymbol{C}}_{t,f}\, x, x\big\rangle \qquad \text{with } \; \hat{\boldsymbol{C}}_{t,f} := \boldsymbol{S}^{(\alpha)}_{t,f}\hat{\boldsymbol{C}}\boldsymbol{S}^{(\alpha)\star}_{t,f},$$

where $\hat{\boldsymbol{C}}_{t,f}$ does not depend on α (recall that $\boldsymbol{S}^{(\alpha)}_{t,f}$ is defined by $\big(\boldsymbol{S}^{(\alpha)}_{t,f}\, x\big)(t') = x(t'-t)\, e^{j2\pi ft'} e^{j2\pi(\alpha-1/2)ft}$).

10.7.2. *Bias-variance analysis*

If $x(t)$ now indicates the random process, the estimator $\hat{C}_x(t,f)$ is also random. For an analysis of the statistical behavior (bias and variance) of $\hat{C}_x(t,f)$ that is not too complicated, we suppose that operator $\hat{\boldsymbol{C}}$ is normal (i.e., $\hat{\boldsymbol{C}}\hat{\boldsymbol{C}}^\star = \hat{\boldsymbol{C}}^\star\hat{\boldsymbol{C}}$; the set of estimators generated by normal operators includes, in particular, all real-valued estimators) and normalized as $\mathrm{tr}\{\hat{\boldsymbol{C}}\} = 1$ where $\mathrm{tr}\{\hat{\boldsymbol{C}}\}$ denotes the trace of $\hat{\boldsymbol{C}}$ (this normalization is equivalent to $\hat{\phi}_{\text{d-D}}(0,0) = 1$, which means that $\hat{C}_x(t,f)$ preserves the energy E_x).

It can then be shown [KOZ 94b] that the *bias* of $\hat{C}_x(t,f)$, $B(t,f) := \mathrm{E}\big\{\hat{C}_x(t,f) - \bar{C}_x(t,f)\big\} = \mathrm{E}\big\{\hat{C}_x(t,f)\big\} - \bar{C}_x(t,f)$, is given by

$$B(t,f) \;=\; \big[\hat{\phi}_{\text{t-f}}(t,f) - \phi_{\text{t-f}}(t,f)\big] \;*_t *_f \; \overline{W}_x^{(0)}(t,f)\,,$$

where $\overline{W}_x^{(0)}(t,f)$ is the Wigner-Ville spectrum of the process $x(t)$ (see (10.14)). If $\hat{\phi}_{\text{t-f}}(t,f) \equiv \phi_{\text{t-f}}(t,f)$ or, equivalently, $\hat{\phi}_{\text{d-D}}(\tau,\xi) \equiv \phi_{\text{d-D}}(\tau,\xi)$, $B(t,f)$ is zero for all (t,f) and thus the estimator $\hat{C}_x(t,f)$ is unbiased. On the other hand, if $\hat{\phi}_{\text{d-D}}(\tau,\xi) \neq \phi_{\text{d-D}}(\tau,\xi)$, $B(t,f)$ can be large in those TF regions where $\overline{W}_x^{(0)}(t,f)$ is large. For the squared norm of the bias, $B^2 := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |B(t,f)|^2\, dt\, df$, the following expression can be shown:

$$B^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \big|\hat{\phi}_{\text{d-D}}(\tau,\xi) - \phi_{\text{d-D}}(\tau,\xi)\big|^2\, |\bar{A}_x(\tau,\xi)|^2\, d\tau\, d\xi\,. \tag{10.24}$$

Therefore, B^2 will be small if $\hat{\phi}_{\text{d-D}}(\tau,\xi) \approx \phi_{\text{d-D}}(\tau,\xi)$ on the effective support of $|\bar{A}_x(\tau,\xi)|$. Consequently, to have $B^2 \approx 0$ in the case of an underspread process, it suffices that $\hat{\phi}_{\text{d-D}}(\tau,\xi) \approx \phi_{\text{d-D}}(\tau,\xi)$ for (τ,ξ) close to the origin.

If the process $x(t)$ is Gaussian, the *variance* of estimator $\hat{C}_x(t,f)$, $V^2(t,f) := \mathrm{E}\{\left|\hat{C}_x(t,f) - \mathrm{E}\{\hat{C}_x(t,f)\}\right|^2\}$, can be expressed as [KOZ 94b]

$$V^2(t,f) \;=\; \mathrm{tr}\{\hat{\boldsymbol{C}}_{t,f}\boldsymbol{R}_x\hat{\boldsymbol{C}}^{\star}_{t,f}\boldsymbol{R}_x\}\,.$$

It can then be shown that $V^2(t,f)$ is large in those TF regions where $\overline{W}_x^{(0)}(t,f)$ is large. For the integral $V^2 := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V^2(t,f)\,dt\,df$, we obtain

$$V^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |D_{\hat{\boldsymbol{C}}}(\tau,\xi)|^2\,|\bar{A}_x(\tau',\xi')|^2\,e^{j2\pi(\tau\xi'-\tau'\xi)}d\tau\,d\xi\,d\tau' d\xi'.$$

If $\hat{\boldsymbol{C}}$ and $\boldsymbol{R}_x$ are jointly underspread[5], $|D_{\hat{\boldsymbol{C}}}(\tau,\xi)|^2$ and $|\bar{A}_x(\tau,\xi)|^2 = |D_{\boldsymbol{R}_x}(\tau,\xi)|^2$ are effectively zero outside a common small region around the origin of the (τ,ξ) plane. We can thus set $e^{j2\pi(\tau\xi'-\tau'\xi)} \approx 1$ in the above integral. This yields

$$V^2 \;\approx\; \|\boldsymbol{R}_x\|^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\hat{\phi}_{\text{d-D}}(\tau,\xi)|^2\,d\tau\,d\xi\,,$$

where $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\bar{A}_x(\tau,\xi)|^2\,d\tau\,d\xi \;=\; \|\boldsymbol{R}_x\|^2$ and $|D_{\hat{\boldsymbol{C}}}(\tau,\xi)|^2 \;=\; |\hat{\phi}_{\text{d-D}}(-\tau,-\xi)|^2$ have been used. Thus, V^2 will be small if the integral of $|\hat{\phi}_{\text{d-D}}(\tau,\xi)|^2$ is small. With the normalization $\hat{\phi}_{\text{d-D}}(0,0)=1$, this essentially amounts to the condition that the effective support of $|\hat{\phi}_{\text{d-D}}(\tau,\xi)|$ around the origin is small or, equivalently, that $\hat{\phi}_{\text{t-f}}(t,f)$ is a smooth function. In this case, the convolution (10.23) corresponds to a *smoothing* of $W_x^{(0)}(t,f)$. However, a small effective support of $|\hat{\phi}_{\text{d-D}}(\tau,\xi)|$ around the origin may cause $\left|\hat{\phi}_{\text{d-D}}(\tau,\xi) - \phi_{\text{d-D}}(\tau,\xi)\right|^2$ to be large for (τ,ξ) outside this effective support. According to (10.24), this may entail a large bias B^2 – unless the process is underspread. In the underspread case, $|\bar{A}_x(\tau,\xi)|^2$ in (10.24) is strongly concentrated around the origin, and thus B^2 will not be affected by the behavior of $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ outside a neighborhood of the origin.

These results demonstrate the existence of a *bias-variance tradeoff*. For a stronger smoothing of $W_x^{(0)}(t,f)$ in (10.23), the variance of estimator $\hat{C}_x(t,f)$ decreases, but the bias of $\hat{C}_x(t,f)$ may increase; this increase, however, will not be large in the underspread case where the spectrum $\bar{C}_x(t,f)$ is smooth.

---

5. Two linear operators $\boldsymbol{A}$ and $\boldsymbol{B}$ are called *jointly underspread* if their generalized spreading functions $D_{\boldsymbol{A}}^{(\alpha)}(\tau,\xi)$ and $D_{\boldsymbol{B}}^{(\alpha)}(\tau,\xi)$ are concentrated in the *same* region $\mathcal{S}$ of area $|\mathcal{S}| \ll 1$ around the origin of the (τ,ξ) plane.

10.7.3. *Designing an estimator*

The bias-variance analysis presented in the previous section leads us to the following conclusion: when the process $x(t)$ is underspread, it is advantageous to choose $\hat{\phi}_{\text{d-D}}(\tau,\xi) \approx \phi_{\text{d-D}}(\tau,\xi)$ on the effective support of $|\bar{A}_x(\tau,\xi)|$ and $\hat{\phi}_{\text{d-D}}(\tau,\xi) \approx 0$ otherwise. In fact, it has been shown in [KOZ 94b] that when the support $\mathcal{S}$ of $|\bar{A}_x(\tau,\xi)|$ is compact, the choice

$$\hat{\phi}_{\text{d-D}}(\tau,\xi) \;=\; \begin{cases} \phi_{\text{d-D}}(\tau,\xi)\,, & (\tau,\xi) \in \mathcal{S} \\ 0\,, & (\tau,\xi) \notin \mathcal{S} \end{cases}$$

yields the unbiased estimator with minimal variance (i.e., $B^2 = 0$ and, moreover, V^2 is minimum among all estimators satisfying $B^2 = 0$). Unfortunately, the assumption that $|\bar{A}_x(\tau,\xi)|$ has compact support is rarely satisfied in practice; in addition, it is possible that an even smaller mean-square error of $\hat{C}_x(t,f)$ can be achieved by allowing a non-zero bias in favor of a smaller variance[6].

We present here a simple heuristic method for designing an estimator $\hat{C}_x(t,f)$ with the special structure of a multi-window spectrogram, which allows an efficient implementation. We assume that the *effective* support of $\bar{A}_x(\tau,\xi)$ – but not necessarily the detailed form of $\bar{A}_x(\tau,\xi)$ – is known. The kernel of the operator $\hat{\boldsymbol{C}}$ defining $\hat{C}_x(t,f)$ is chosen

$$(\hat{\boldsymbol{C}})(t,t') \;=\; \frac{1}{K}\sum_{k=1}^{K} c_k(t)\, c_k^*(t')\,,$$

with orthonormal functions $c_k(t)$, $k = 1,\ldots,K$. We can observe that the functions $c_k(t)$ are the eigenfunctions of operator $\hat{\boldsymbol{C}}$, with the eigenvalues given by $1/K$. The kernels of estimator $\hat{C}_x(t,f)$ follow as

$$\hat{\phi}_{\text{t-f}}(t,f) \;=\; \frac{1}{K}\sum_{k=1}^{K} W_{c_k}^{(0)}(-t,-f)\,, \qquad \hat{\phi}_{\text{d-D}}(\tau,\xi) \;=\; \frac{1}{K}\sum_{k=1}^{K} A_{c_k}^{(0)*}(\tau,\xi)\,.$$

With (10.23), the estimator is then obtained as

$$\begin{aligned} \hat{C}_x(t,f) \;&=\; \frac{1}{K}\sum_{k=1}^{K}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{c_k}^{(0)}(t'-t,f'-f)\, W_x^{(0)}(t',f')\,dt'df' \\ &=\; \frac{1}{K}\sum_{k=1}^{K} S_x^{c_k}(t,f)\,, \end{aligned} \tag{10.25}$$

6. It should be noted at this point that a direct minimization of the mean-square error of $\hat{C}_x(t,f)$ is not feasible. In fact, the result would depend on the second-order statistics of $x(t)$ [SAY 95b]. However, these statistics are unknown (had they been known, we would not have to estimate $\bar{C}_x(t,f)$!).

where $S_x^{c_k}(t,f) := \left| \int_{-\infty}^{\infty} x(t')\, c_k^*(t'-t)\, e^{-j2\pi f t'}\, dt' \right|^2$ is the spectrogram of $x(t)$ using $c_k(t)$ as analysis window (see Chapters 1 and 5). The TF representation (10.25) is known as a *multi-window spectrogram* [CUN 94]; it generalizes the stationary multi-window techniques proposed in [THO 82].

With this special structure, the design of the estimator reduces to the choice of the order K and the orthonormal windows $c_k(t)$. According to our discussion at the beginning of this section, we have to adapt the effective support of $\hat{\phi}_{\text{d-D}}(\tau,\xi) = \frac{1}{K}\sum_{k=1}^{K} A_{c_k}^{(0)*}(\tau,\xi)$ to the effective support of $\bar{A}_x(\tau,\xi)$, hereafter denoted by $\mathcal{S}$. To this end, we propose to choose the windows as

$$c_k(t) \;=\; \sqrt{a}\, h_k(at) \quad \text{or} \quad c_k(t) \;=\; \sqrt{a}\, p_k(at)\,, \qquad k = 1, \ldots, K\,,$$

where the $h_k(t)$ are the K first Hermite functions [FOL 89, HLA 98], the $p_k(t)$ are the K first prolate spheroidal functions [FLA 99, HLA 98], and $a > 0$ is a compression/dilation factor. Hermite functions are particularly appropriate when $\mathcal{S}$ has an elliptical form, whereas prolate spheroidal functions are preferable for an $\mathcal{S}$ with a rectangular form.

Parameter a allows us to adapt the ratio of the extensions of $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ in the τ and ξ directions to that of $\bar{A}_x(\tau,\xi)$. Such an adaptation is obtained by setting

$$a \;=\; \sqrt{\frac{\Delta t}{\Delta f}\,\frac{\xi_{\max}}{\tau_{\max}}}\,, \tag{10.26}$$

where Δt and Δf are respectively the effective duration and effective bandwidth of $h_1(t)$ (or of $p_1(t)$), and $\tau_{\max}$ and $\xi_{\max}$ are respectively the effective time and frequency correlation horizons of process $x(t)$.

Finally, the choice of the order K can be based on the observation that the area of the effective support of the kernel $\hat{\phi}_{\text{t-f}}(t,f) = \frac{1}{K}\sum_{k=1}^{K} W_{c_k}^{(0)}(-t,-f)$ is approximately equal to K [HLA 98]. Consequently, the area of the effective support of $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ is approximately given by $1/K$. Now $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ must be selected such that the area of its effective support is equal to $|\mathcal{S}|$, the area of $\mathcal{S}$. Thus, we obtain the following rule for the choice of K:

$$K \;=\; \text{round}\left\{\frac{1}{|\mathcal{S}|}\right\}. \tag{10.27}$$

Therefore, K will be larger for a process that is "more underspread".

10.7.4. *Numerical results*

We illustrate the application of our method to the estimation of the Wigner-Ville spectrum of an underspread process. This process was generated using the TF syn-

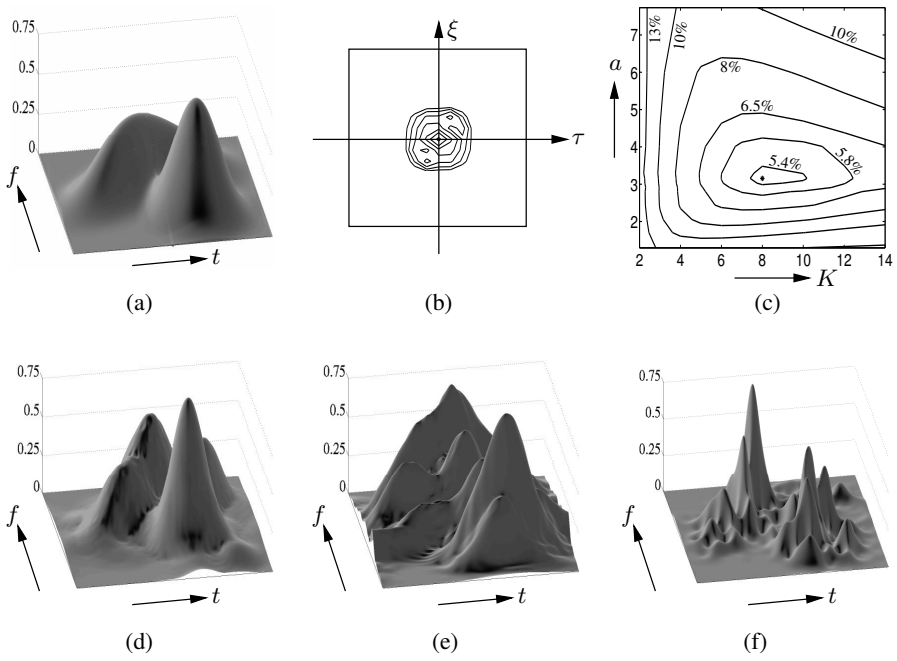

Figure 10.5. *Estimation of the Wigner-Ville spectrum of an underspread process $x(t)$ by means of the estimator (10.25) using Hermite functions: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$, (b) GEAF magnitude $|\bar{A}_x(\tau,\xi)|$ (the square around the origin has an area of 1 and thus allows an assessment of the underspread nature of $x(t)$), (c) mean-square error of the estimator for various values of the parameters K and a (the asterisk indicates the position of the minimum error), (d) result of the quasi-optimum estimator with $K = 9$, $a = 3.76$ (these parameters were obtained from the rules (10.26) and (10.27)), (e) result of the estimator with $K = 9$, $a = 6$ (duration too large), (f) result of the estimator with $K = 1$, $a = 3.76$ (order too small)*

thesis technique introduced in [HLA 95]. The Wigner-Ville spectrum and the GEAF of the process are represented in Figures 10.5(a) and (b); the area of the GEAF's effective support $\mathcal{S}$ has been estimated as $|\mathcal{S}| = 0.107$. Due to the nearly elliptical form of $\mathcal{S}$, the estimator was based on the Hermite functions $h_k(t)$. Figure 10.5(c) shows the (normalized) mean-square error of the estimator for different values of order K and compression/dilation factor a. The minimum of the mean-square error is 5.3%; it is obtained for $K = 8$ and $a = 3.15$. Our heuristic rules (10.26) and (10.27) yield parameters $K = 9$ and $a = 3.76$, corresponding to a mean-square error of 5.6%. Thus, the performance loss relative to the optimum parameters is small. The result of our quasi-optimum estimator obtained for a realization of the process is illustrated in Figure 10.5(d), while results obtained for the same realization with estimators designed using "incorrect" values of K and a are shown in Figures 10.5(e) and (f).

10.8. Estimation of non-stationary processes

Estimating non-stationary signals contaminated by noise or other interferences has considerable importance in many practical applications. Therefore, the subject of this section will be the application of time-varying spectra to the estimation of underspread non-stationary processes. We will follow the approach introduced in [HLA 00], using the GWVS because of its simple mathematical structure. However, we recall from Section 10.6.1 that for an underspread process, the GES and other spectra are approximately equivalent to the GWVS; thus, these spectra can effectively be substituted for the GWVS in the equations below. We also note that other TF approaches to estimating non-stationary processes are discussed in [ABD 69, SIL 95a, SIL 95b, KHA 97, SAY 95c, LAN 97].

More specifically, we consider the estimation of a centered non-stationary random signal $s(t)$ from an observed signal $x(t) = s(t) + n(t)$, where $n(t)$ is a centered non-stationary "noise" process that is uncorrelated with $s(t)$. The correlation operators $\boldsymbol{R}_s$ and $\boldsymbol{R}_n$ are initially assumed known. The estimation of $s(t)$ is performed using a linear time-varying system $\boldsymbol{H}$, i.e.,

$$\hat{s}(t) \;=\; (\boldsymbol{H}x)(t) \;=\; \int_{-\infty}^{\infty} h(t,t')\, x(t')\, dt'. \tag{10.28}$$

The non-causal system $\boldsymbol{H}$ minimizing the mean-square estimation error $\mathrm{E}\{|\hat{s}(t) - s(t)|^2\}$ is the *time-varying Wiener filter* given by [VAN 68, POO 88, SCH 91, THE 92]

$$\boldsymbol{H}_{\mathrm{W}} \;=\; \boldsymbol{R}_s(\boldsymbol{R}_s + \boldsymbol{R}_n)^{-1}. \tag{10.29}$$

As we mentioned in Section 10.1, for $s(t)$ and $n(t)$ stationary, $\boldsymbol{H}_{\mathrm{W}}$ is a time-invariant system whose transfer function is given by a simple expression involving the power spectral densities of $s(t)$ and $n(t)$ [PAP 91, VAN 68, POO 88, SCH 91, THE 92, WIE 49]:

$$\widehat{h}_{\mathrm{W}}(f) \;=\; \frac{\widehat{r}_s(f)}{\widehat{r}_s(f) + \widehat{r}_n(f)}\,. \tag{10.30}$$

This frequency-domain expression allows an easy design and interpretation of the time-invariant Wiener filter, since it uses a simple division of functions instead of products and inverses of operators as in (10.29).

In Section 10.1, we wondered whether in the non-stationary/time-varying case there is an expression similar to (10.30), of comparable simplicity – of course admitting that the transfer function $\widehat{h}_{\mathrm{W}}(f)$ and the spectral densities $\widehat{r}_s(f)$ and $\widehat{r}_n(f)$ are replaced by non-stationary definitions that depend on time. We will now present an answer to this question.

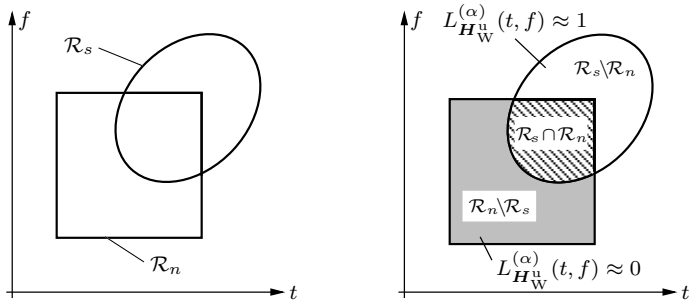

Figure 10.6. *TF interpretation of the time-varying Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ for jointly underspread signal and noise processes: (a) Effective TF support regions of signal and noise, (b) TF pass, stop, and transition regions of $\boldsymbol{H}_{\mathrm{W}}$*

10.8.1. *TF formulation of the optimum filter*

In fact, a positive answer exists for *jointly underspread* processes $s(t)$ and $n(t)$ (see footnote on page 296). It has been shown [HLA 00] that the time-varying Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ can be decomposed as $\boldsymbol{H}_{\mathrm{W}} = \boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}} + \boldsymbol{H}^{\mathrm{o}}_{\mathrm{W}}$, where the components $\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}$ and $\boldsymbol{H}^{\mathrm{o}}_{\mathrm{W}}$ are characterized as follows:

– $\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}$ is an underspread system that admits an approximate TF formulation involving the GWVS of $s(t)$ and $n(t)$:

$$L^{(\alpha)}_{\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}}(t,f) \;\approx\; \frac{\overline{W}_s^{(\alpha)}(t,f)}{\overline{W}_s^{(\alpha)}(t,f) + \overline{W}_n^{(\alpha)}(t,f)}\;; \tag{10.31}$$

– $\boldsymbol{H}^{\mathrm{o}}_{\mathrm{W}}$ is an overspread system that has only a small effect on the performance (mean-square error) and can thus be ignored.

The TF formulation (10.31) is the desired extension of (10.30) to the underspread non-stationary case. It allows a simple and intuitive interpretation of the time-varying Wiener filter that is illustrated in Figure 10.6. Let $\mathcal{R}_s$ and $\mathcal{R}_n$ be the effective support regions of $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$, respectively. Regarding the action of the time-varying Wiener filter, three TF regions can be distinguished:

– *pass region.* In the region $\mathcal{R}_s \backslash \mathcal{R}_n$ where there is only the signal, expression (10.31) yields $L^{(\alpha)}_{\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}}(t,f) \approx 1$. This shows that TF components of the observed signal $x(t)$ that are not contaminated by noise are preserved without modification by $\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}$;

– *stop region.* In the region $\mathcal{R}_n \backslash \mathcal{R}_s$ where there is only noise, (10.31) gives $L^{(\alpha)}_{\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}}(t,f) \approx 0$. Therefore, TF components of $x(t)$ that contain no component of signal $s(t)$ are completely suppressed by $\boldsymbol{H}^{\mathrm{u}}_{\mathrm{W}}$;

– *transition region.* In the region $\mathcal{R}_s \cap \mathcal{R}_n$ that contains both signal and noise, $L^{(\alpha)}_{\boldsymbol{H}^{\rm u}_{\rm W}}(t,f)$ is approximately[7] between 0 and 1. In this region, $\boldsymbol{H}^{\rm u}_{\rm W}$ performs an attenuation that varies with time and frequency according to the "TF signal-to-noise ratio" $\overline{W}^{(\alpha)}_s(t,f)/\overline{W}^{(\alpha)}_n(t,f)$. For instance, at those points (t,f) where signal and noise are equally strong, i.e., $\overline{W}^{(\alpha)}_s(t,f) = \overline{W}^{(\alpha)}_n(t,f)$ or $\overline{W}^{(\alpha)}_s(t,f)/\overline{W}^{(\alpha)}_n(t,f) = 1$, expression (10.31) yields $L^{(\alpha)}_{\boldsymbol{H}^{\rm u}_{\rm W}}(t,f) \approx 1/2$.

10.8.2. *TF design of a quasi-optimum filter*

The TF formulation (10.31) suggests a *TF design* of non-stationary estimators. Let us define the "pseudo-Wiener filter" $\widetilde{\boldsymbol{H}}_{\rm W}$ by identifying its generalized Weyl symbol with the right-hand side of (10.31) [HLA 00]:

$$L^{(\alpha)}_{\widetilde{\boldsymbol{H}}_{\rm W}}(t,f) := \frac{\overline{W}^{(\alpha)}_s(t,f)}{\overline{W}^{(\alpha)}_s(t,f) + \overline{W}^{(\alpha)}_n(t,f)}\,. \tag{10.32}$$

Note that $\widetilde{\boldsymbol{H}}_{\rm W}$ depends on the value of the parameter α chosen in (10.32). For jointly underspread processes $s(t)$ and $n(t)$, the approximation (10.31) is valid; with (10.32), we obtain $L^{(\alpha)}_{\widetilde{\boldsymbol{H}}_{\rm W}}(t,f) \approx L^{(\alpha)}_{\boldsymbol{H}^{\rm u}_{\rm W}}(t,f)$. This shows that in the underspread case, the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\rm W}$ is a good approximation of the Wiener filter $\boldsymbol{H}_{\rm W}$ (to be more precise, of its underspread part $\boldsymbol{H}^{\rm u}_{\rm W}$, but this effectively amounts to the same thing). From this, we can also conclude that $\widetilde{\boldsymbol{H}}_{\rm W}$ is approximately independent of the parameter α chosen in (10.32) (cf. the results of Section 10.6.1.1). On the other hand, for processes $s(t)$, $n(t)$ that are not jointly underspread, we must expect that $\widetilde{\boldsymbol{H}}_{\rm W}$ is very different from $\boldsymbol{H}_{\rm W}$ and that its performance is significantly poorer.

According to (10.32), the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\rm W}$ is *designed* in the TF plane. However, the calculation of the estimated signal $\hat{s}(t)$ can be carried out in the time domain using the relation (see (10.28))

$$\hat{s}(t) = \big(\widetilde{\boldsymbol{H}}_{\rm W}\, x\big)(t) = \int_{-\infty}^{\infty} \tilde{h}_{\rm W}(t,t')\, x(t')\, dt'.$$

The impulse response $\tilde{h}_{\rm W}(t,t')$ of $\widetilde{\boldsymbol{H}}_{\rm W}$ is obtained from $L^{(\alpha)}_{\widetilde{\boldsymbol{H}}_{\rm W}}(t,f)$ by inversion of (10.5), (10.6):

$$\tilde{h}_{\rm W}(t,t') = \int_{-\infty}^{\infty} L^{(\alpha)}_{\widetilde{\boldsymbol{H}}_{\rm W}}\!\left(\Big(\frac{1}{2}+\alpha\Big)t + \Big(\frac{1}{2}-\alpha\Big)t', f\right) e^{j2\pi f(t-t')}\, df\,. \tag{10.33}$$

7. Since $\boldsymbol{H}^{\rm u}_{\rm W}$ is an underspread system, $L^{(\alpha)}_{\boldsymbol{H}^{\rm u}_{\rm W}}(t,f)$ is approximately real and even for $\alpha \neq 0$ [MAT 00b].

Alternatively, an efficient implementation of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ can be based on the multi-window version of the short-time Fourier transform or the Gabor transform [KOZ 96, HLA 00, MAT 02a].

The pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ has two practical advantages over the Wiener filter $\boldsymbol{H}_{\mathrm{W}}$:

– the *a priori* knowledge required for designing $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ consists of the GWVS $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$. These functions are clearly more intuitive and easier to handle than the correlation operators $\boldsymbol{R}_s$ and $\boldsymbol{R}_n$ arising in the design of $\boldsymbol{H}_{\mathrm{W}}$ according to (10.29); this is true despite the fact that the GWVS and correlation operator are mathematically equivalent. This advantage can be particularly important when the statistics have to be estimated from observed signals (see Section 10.7);

– the calculation of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ according to (10.32) is numerically less expensive and more stable than the calculation of $\boldsymbol{H}_{\mathrm{W}}$ according to (10.29) since it uses a simple division of functions rather than an inversion of operators (or matrices in a discrete-time implementation).

We finally note that robust versions (in the *minimax* sense) of the time-varying Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ and the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ have been proposed in [MAT 98a, MAT 99a, MAT 00a, MAT 01].

10.8.3. *Numerical results*

Figure 10.7 shows the Wigner-Ville spectra and GEAFs of signal and noise processes that are jointly underspread to a certain extent because the GEAFs are effectively contained in a common region with area < 1. For these processes, Figure 10.8 shows the Weyl symbols of the Wiener filter $\boldsymbol{H}_{\mathrm{W}}$, of its underspread part $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$, and of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$. The TF pass, stop and transition regions of these filters are easily recognizable (cf. Figure 10.6). Furthermore, it can be verified that the Weyl symbol of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is a good approximation to the Weyl symbol of $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$. The improvement of the signal-to-noise ratio (SNR) achieved by the filters $\boldsymbol{H}_{\mathrm{W}}$ and $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is finally illustrated in Figure 10.9, which represents the output SNR as a function of the input SNR. We see that for an input SNR of 0 dB, the SNR improvement is more than 6 dB for both filters. The performance of both filters is nearly identical for all values of the input SNR. This shows that our TF design method produces quasi-optimum estimators.

Our second example concerns the application of an adaptive "on-line" variant of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ to the denoising of a speech signal contaminated by stationary white noise. The filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ has been implemented in an approximate manner by a multi-window Gabor filter [MAT 02a]. This implementation is very efficient – especially for signals of long duration – and it allows on-line estimation of the statistical *a priori* knowledge during the filter operation. Figure 10.10 shows the noise-free sig-

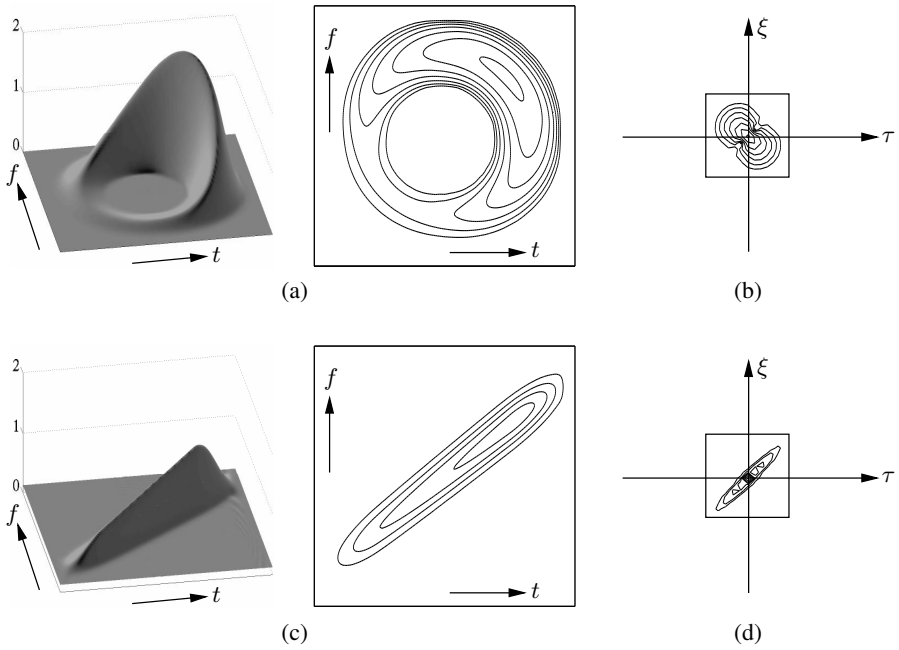

Figure 10.7. *TF representations of the statistics of the signal $s(t)$ and the noise $n(t)$: (a) Wigner-Ville spectrum of $s(t)$, (b) GEAF magnitude of $s(t)$, (c) Wigner-Ville spectrum of $n(t)$, (d) GEAF magnitude of $n(t)$. In (b) and (d), the squares around the origin have an area of 1 and thus allow an assessment of the jointly underspread nature of $s(t)$ and $n(t)$. The duration of all signals is 128 samples*

nal, the noisy signal (input signal of the filter) and the denoised signal (output signal of the filter, estimation result). The SNR is increased by 6.05 dB.

10.9. Detection of non-stationary processes

Just as their estimation, the detection and classification of non-stationary signals are important tasks in signal processing. In this section, following the approach introduced in [MAT 96, MAT 00b], we will study the application of time-varying spectra to this task. We note that other TF approaches to the detection and classification of non-stationary signals and processes are discussed in Chapter 13 as well as in [FLA 86a, FLA 88b, FLA 88a, MAT 96, MAT 98b, MAT 99b, MAT 03, SAY 95a, SAY 96, HEI 95, RIC 98, VIN 94, KAY 85, KUM 84, MAR 97, DAV 98, DAV 01, LEM 94].

The processes are assumed Gaussian and jointly underspread. We will again use the GWVS, with the understanding that, for jointly underspread processes, the GES

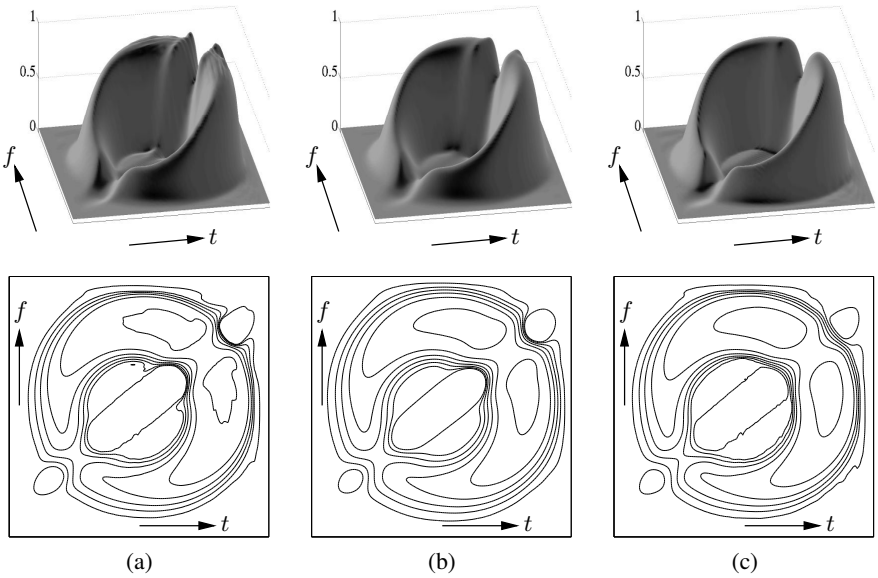

Figure 10.8. *TF representation of the various Wiener-type filters for the processes of Figure 10.7: (a) Weyl symbol of the Wiener filter $\boldsymbol{H}_{\mathrm{W}}$, (b) Weyl symbol of the underspread part $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$ of $\boldsymbol{H}_{\mathrm{W}}$, (c) Weyl symbol of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$*

and other spectra can be substituted for the GWVS in the following equations. The problem studied in this section is, more specifically, the discrimination between two centered non-stationary Gaussian random signals $x_0(t)$ and $x_1(t)$. Thus, we have the hypotheses

$$H_0 : \; x(t) = x_0(t) \qquad \text{versus} \qquad H_1 : \; x(t) = x_1(t)\,.$$

It is known that the optimum detector (likelihood ratio detector) [VAN 68, POO 88, SCH 91] calculates a quadratic form of the observed signal $x(t)$

$$\Lambda(x) \;=\; \langle \boldsymbol{H}_{\mathrm{L}}\, x, x\rangle \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_{\mathrm{L}}(t,t')\, x(t')\, x^*(t)\, dt\, dt'. \tag{10.34}$$

Here, the operator (time-varying system) $\boldsymbol{H}_{\mathrm{L}}$ is given by

$$\boldsymbol{H}_{\mathrm{L}} \;=\; \boldsymbol{R}_{x_0}^{-1} - \boldsymbol{R}_{x_1}^{-1} \;=\; \boldsymbol{R}_{x_0}^{-1}(\boldsymbol{R}_{x_1} - \boldsymbol{R}_{x_0})\,\boldsymbol{R}_{x_1}^{-1}, \tag{10.35}$$

where $\boldsymbol{R}_{x_0}$ and $\boldsymbol{R}_{x_1}$ denote the correlation operators of $x_0(t)$ and $x_1(t)$, respectively. The test statistic $\Lambda(x)$ is then compared to a threshold to decide whether H_0 or H_1 is in force.

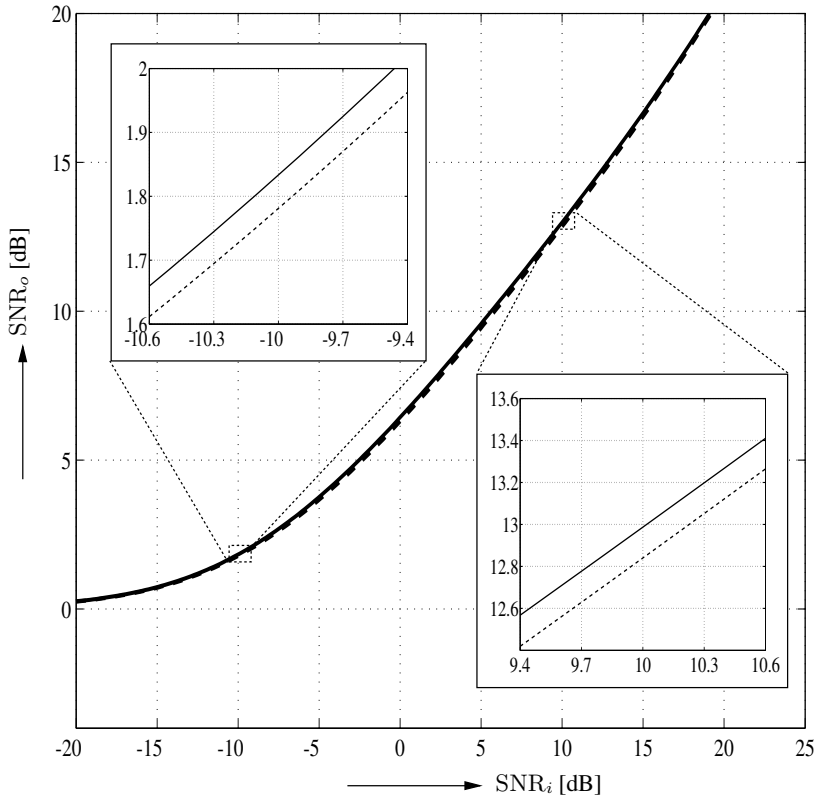

Figure 10.9. *Output SNR versus input SNR for $\boldsymbol{H}_{\mathrm{W}}$ (solid line) and for $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ (dashed line)*

For stationary processes, $\boldsymbol{H}_{\mathrm{L}}$ is a time-invariant system with transfer function

$$\widehat{h}_{\mathrm{L}}(f) \;=\; \frac{\widehat{r}_{x_1}(f) - \widehat{r}_{x_0}(f)}{\widehat{r}_{x_0}(f)\,\widehat{r}_{x_1}(f)}\;.$$

Consequently, the optimum test statistic (10.34) can be written as

$$\Lambda(x) \;=\; \int_{-\infty}^{\infty} \widehat{h}_{\mathrm{L}}(f)\,|\widehat{x}(f)|^2\,df \;=\; \int_{-\infty}^{\infty} \frac{\widehat{r}_{x_1}(f) - \widehat{r}_{x_0}(f)}{\widehat{r}_{x_0}(f)\,\widehat{r}_{x_1}(f)}\,|\widehat{x}(f)|^2\,df\,. \qquad (10.36)$$

This expression involves simple products and reciprocals of functions, instead of products and inverses of operators as in (10.35). We will now extend this formulation to the underspread non-stationary case.

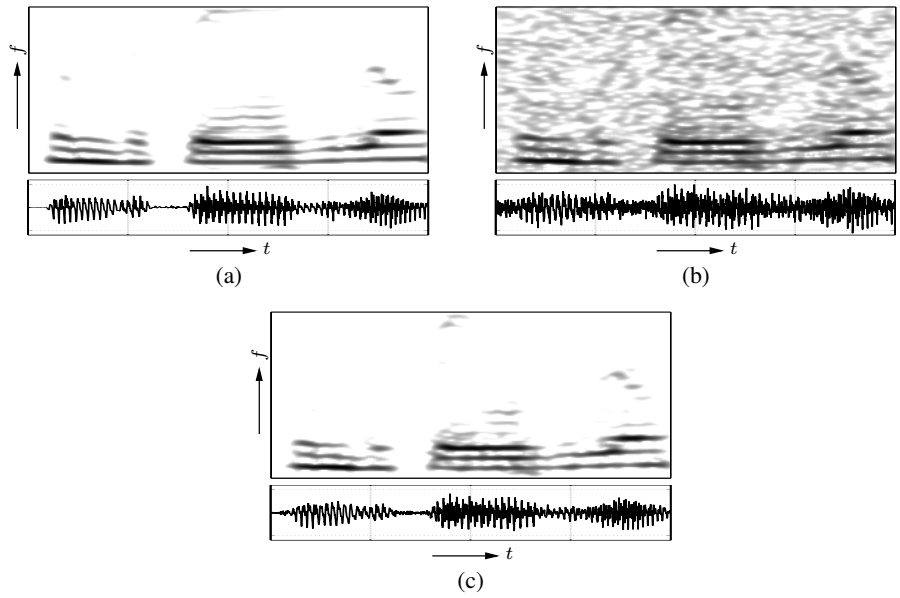

Figure 10.10. *Denoising of a speech signal using an adaptive "on-line" version of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ implemented by a multi-window Gabor filter with on-line estimation of the statistical* a priori *knowledge: (a) Noise-free speech signal, (b) noisy signal (SNR_{input} = 3 dB), (c) denoised/estimated signal (SNR_{output} = 9.05 dB). The TF representations shown are smoothed Wigner-Ville distributions. The signal duration is 4,096 samples*

10.9.1. *TF formulation of the optimum detector*

It is well known [FLA 88a, SAY 95a, MAT 96] that the quadratic test statistic (10.34) can be expressed as an inner product in the TF domain:

$$\Lambda(x) \;=\; \left\langle L^{(\alpha)}_{\boldsymbol{H}_{\mathrm{L}}}, W^{(\alpha)}_{x} \right\rangle \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L^{(\alpha)}_{\boldsymbol{H}_{\mathrm{L}}}(t,f)\, W^{(\alpha)*}_{x}(t,f)\, dt\, df\;, \qquad (10.37)$$

for all $\alpha \in \mathbb{R}$. In this expression, $W^{(\alpha)}_{x}(t,f)$ is the generalized Wigner-Ville distribution of signal $x(t)$ whose definition has been given in (10.16). Thus, $\Lambda(x)$ can be interpreted as a weighted integral of $W^{(\alpha)*}_{x}(t,f)$, where the weighting function is the generalized Weyl symbol of operator $\boldsymbol{H}_{\mathrm{L}}$.

A simplification of the TF formulation (10.37) can be obtained if the processes $x_0(t)$ and $x_1(t)$ are jointly underspread. In that case, similarly to Section 10.8.1, the operator $\boldsymbol{H}_{\mathrm{L}}$ can be decomposed as $\boldsymbol{H}_{\mathrm{L}} = \boldsymbol{H}^{\mathrm{u}}_{\mathrm{L}} + \boldsymbol{H}^{\mathrm{o}}_{\mathrm{L}}$, where the components $\boldsymbol{H}^{\mathrm{u}}_{\mathrm{L}}$ and $\boldsymbol{H}^{\mathrm{o}}_{\mathrm{L}}$ have the following properties [MAT 00b]:

– $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}$ is an underspread system that admits the approximate TF formulation

$$L_{\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}}^{(\alpha)}(t,f) \;\approx\; \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\;; \tag{10.38}$$

– $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{o}}$ is an overspread system that has only a small effect on the performance of the detector and can thus be ignored.

Inserting (10.38) into (10.37) while omitting the component $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{o}}$ (i.e., replacing $L_{\boldsymbol{H}_{\mathrm{L}}}^{(\alpha)}(t,f)$ by $L_{\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}}^{(\alpha)}(t,f)$), we obtain the following approximation to our optimum test statistic:

$$\Lambda(x) \;\approx\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\, W_x^{(\alpha)*}(t,f)\,dt\,df\,. \tag{10.39}$$

This TF formulation extends the pure frequency formulation (10.36) to the non-stationary (underspread) case. It allows a very simple and intuitive interpretation of $\Lambda(x)$ in the spirit of the interpretation given in Section 10.8.1 (see Figure 10.6). In fact, apart from the "normalization" by $\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)$, $\Lambda(x)$ collects more (less) energy of the observation $x(t)$ in TF regions where the GWVS $\overline{W}_{x_0}^{(\alpha)}(t,f)$ and $\overline{W}_{x_1}^{(\alpha)}(t,f)$ are more (less) different. This is evidently reasonable, given that our task is to *discriminate* between the two processes $x_0(t)$ and $x_1(t)$.

10.9.2. *TF design of a quasi-optimum detector*

From the TF formulation (10.38), we can easily obtain a *TF design* of non-stationary detectors. Similarly to Section 10.8.2, we define the system $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ by identifying its generalized Weyl symbol with the right-hand side of (10.38) [MAT 96, MAT 00b]:

$$L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}(t,f) \;:=\; \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\,. \tag{10.40}$$

We note that $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ depends on the value of parameter α chosen in (10.40). Substituting $L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}(t,f)$ for $L_{\boldsymbol{H}_{\mathrm{L}}}^{(\alpha)}(t,f)$ in (10.37), we obtain the *TF test statistic*[8]

$$\tilde{\Lambda}(x) \;:=\; \big\langle L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}, W_x^{(\alpha)} \big\rangle \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\, W_x^{(\alpha)*}(t,f)\,dt\,df. \tag{10.41}$$

8. To be able to compare this test statistic to a threshold, we propose using the real part $\mathrm{Re}\{\tilde{\Lambda}(x)\}$. In fact, the imaginary part of $\tilde{\Lambda}(x)$ is small because of approximation (10.39).

For jointly underspread processes $x_0(t)$ and $x_1(t)$, the approximations (10.38) and (10.39) yield $L^{(\alpha)}_{\widetilde{\boldsymbol{H}}_{\rm L}}(t,f) \approx L^{(\alpha)}_{\boldsymbol{H}^{\rm u}_{\rm L}}(t,f)$ and $\tilde{\Lambda}(x) \approx \Lambda(x)$. Therefore, the TF detector $\tilde{\Lambda}(x)$ will be a good approximation to the optimum detector $\Lambda(x)$ and, furthermore, $\tilde{\Lambda}(x)$ will be approximately real and approximately independent of the value of α used in (10.41). On the other hand, if the processes $x_0(t)$ and $x_1(t)$ are not jointly underspread, then $\tilde{\Lambda}(x)$ can be very different from $\Lambda(x)$ and its performance can be significantly poorer.

Although the detector $\tilde{\Lambda}(x)$ is designed in the TF domain, it can also be implemented in the time domain. This is performed using the expression (see (10.34))

$$\tilde{\Lambda}(x) \;=\; \langle \widetilde{\boldsymbol{H}}_{\rm L}\, x, x\rangle \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{h}_{\rm L}(t,t')\, x(t')\, x^*(t)\, dt\, dt',$$

where the impulse response $\tilde{h}_{\rm L}(t,t')$ is obtained from $L^{(\alpha)}_{\widetilde{\boldsymbol{H}}_{\rm L}}(t,f)$ using the inverse generalized Weyl transformation (see (10.33)). An alternative implementation is based on the multi-window version of the short-time Fourier transform or the Gabor transform [MAT 98b, MAT 02a].

Compared to the optimum detector $\Lambda(x)$, the detector $\tilde{\Lambda}(x)$ has two practical advantages. First, the *a priori* knowledge needed to construct the operator $\widetilde{\boldsymbol{H}}_{\rm L}$ consists of the GWVS $\overline{W}^{(\alpha)}_{x_0}(t,f)$ and $\overline{W}^{(\alpha)}_{x_1}(t,f)$ – two TF functions that are much more intuitive and easier to handle than the correlation operators $\boldsymbol{R}_{x_0}$ and $\boldsymbol{R}_{x_1}$ used in the design of $\boldsymbol{H}_{\rm L}$. Second, the calculation of $\widetilde{\boldsymbol{H}}_{\rm L}$ according to (10.40) is numerically less expensive and more stable than the calculation of $\boldsymbol{H}_{\rm L}$ according to (10.35) since it employs a simple division of functions instead of an inversion of operators.

10.9.3. *Numerical results*

Our first example concerns the detection of a signal contaminated by noise, which corresponds to the special case given by $x_0(t) = n(t)$ and $x_1(t) = s(t) + n(t)$ where $s(t)$ and $n(t)$ are uncorrelated. Signal $s(t)$ and noise $n(t)$ are assumed non-stationary, centered, Gaussian and jointly underspread. More specifically, we will consider the processes $s(t)$ and $n(t)$ whose Wigner-Ville spectra and GEAFs have been shown in Figure 10.7. For these processes, Figure 10.11 compares the Weyl symbols of the optimum operator $\boldsymbol{H}_{\rm L}$, its underspread part $\boldsymbol{H}^{\rm u}_{\rm L}$, and the operator $\widetilde{\boldsymbol{H}}_{\rm L}$ designed in the TF domain. We see that these operators are very similar, and we may thus expect the performance of detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$ to be similar too. This similarity is clearly demonstrated by the conditional probability densities, receiver operating characteristics (ROC) [POO 88, SCH 91], and power curves [POO 88, SCH 91] of $\Lambda(x)$ and $\tilde{\Lambda}(x)$ that are shown in Figure 10.12.

In this example, the noise contained a white component (see Figure 10.7(c)); thus, the inversion of the operators $\boldsymbol{R}_{x_0} = \boldsymbol{R}_n$ and $\boldsymbol{R}_{x_1} = \boldsymbol{R}_s + \boldsymbol{R}_n$ (which is required

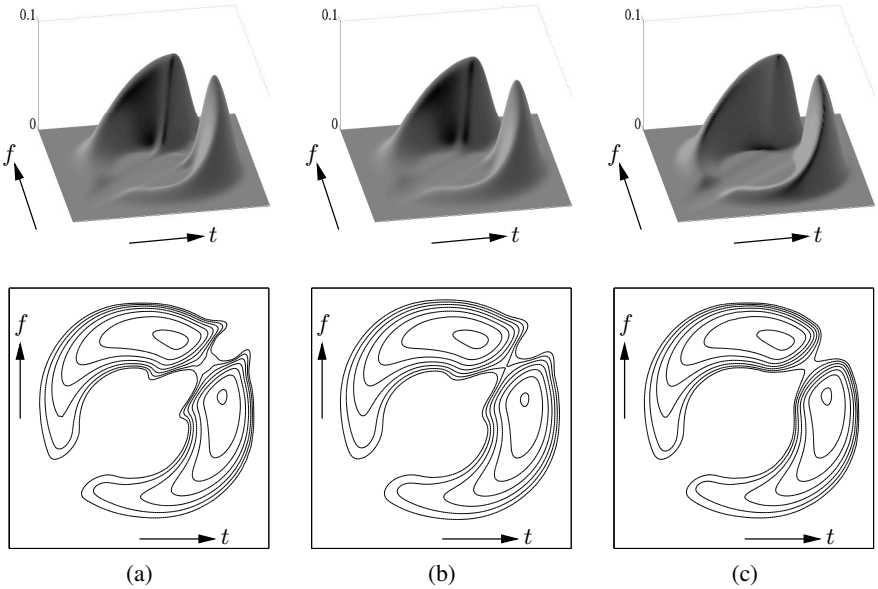

Figure 10.11. *Weyl symbol of the various detection operators for the synthetic processes $s(t)$ and $n(t)$ (see Figure 10.7): (a) Optimum operator $\boldsymbol{H}_{\mathrm{L}}$, (b) its underspread part $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}$, (c) operator $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ designed in the TF domain*

in order to calculate $\boldsymbol{H}_{\mathrm{L}}$ according to (10.35)) did not cause any problem. The next example demonstrates that in practice a less lucky situation may be encountered, and that the TF detector $\tilde{\Lambda}(x)$ then allows a stabilization of the calculation which can result in an improved performance.

We consider the application of the optimum detector $\Lambda(x)$ and of the TF detector $\tilde{\Lambda}(x)$ to the detection of knock in internal combustion engines (see [MAT 98b, MAT 99b, KÖN 94] for background and details). The processes $x_0(t)$ and $x_1(t)$ correspond to pressure signals without and with knock, respectively. The correlation operators and Wigner-Ville spectra of $x_0(t)$ and $x_1(t)$ were estimated from a set of labeled training signals. The estimated Wigner-Ville spectra are shown in Figures 10.13(a) and (b). The operators $\boldsymbol{H}_{\mathrm{L}}$ and $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ underlying detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$ were calculated from these estimated statistics, and then the performance of these detectors was evaluated by using another set of labeled signals.

The Weyl symbols $L^{(0)}_{\boldsymbol{H}_{\mathrm{L}}}(t,f)$ and $L^{(0)}_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}(t,f)$ of operators $\boldsymbol{H}_{\mathrm{L}}$ and $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ are shown in Figures 10.13(c) and (d); it is seen that these operators are quite different. In view of the fact that $L^{(0)}_{\boldsymbol{H}_{\mathrm{L}}}(t,f)$ and $L^{(0)}_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}(t,f)$ are the TF weighting functions of detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$ (see (10.37) and (10.41)), it is then not surprising that the performance

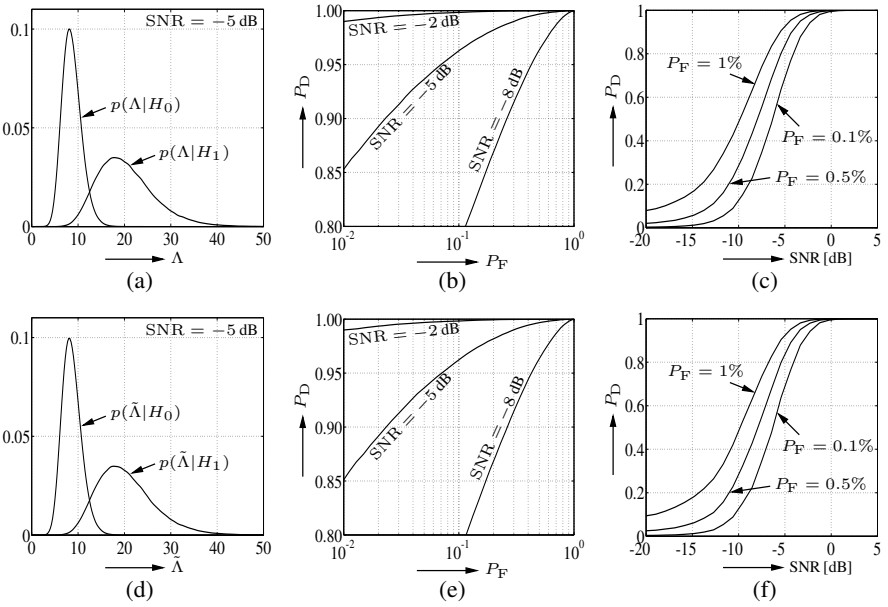

Figure 10.12. *Comparison of the performance of the optimum detector $\Lambda(x)$ and of the "TF detector" $\tilde{\Lambda}(x)$ for the synthetic processes (see Figures 10.7 and 10.11): (a) Conditional probability densities of $\Lambda(x)$ under the two hypotheses, (b) ROC of $\Lambda(x)$, (c) power curves of $\Lambda(x)$, (d) conditional probability densities of $\tilde{\Lambda}(x)$ under the two hypotheses, (e) ROC of $\tilde{\Lambda}(x)$, (f) power curves of $\tilde{\Lambda}(x)$ (SNR = signal-to-noise ratio, P_D = probability of detection, P_F = probability of false alarm)*

of these detectors is very different, too. In fact, the ROC curves shown in Figure 10.13(e) demonstrate that the performance of the TF detector $\tilde{\Lambda}(x)$ is significantly better than that of the detector $\Lambda(x)$, in spite of the theoretical optimality of the latter. This perhaps unexpected result is explained by numerical difficulties encountered in the calculation of operator $\boldsymbol{H}_{\mathrm{L}}$ – difficulties that were caused by the bad conditioning of the estimated operators $\boldsymbol{R}_{x_0}$ and $\boldsymbol{R}_{x_1}$ (these operators are actually matrices in a discrete-time implementation). Despite the use of pseudo-inverses, the inversion of these matrices could not be stabilized sufficiently. In contrast, the calculation of $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ from the estimated Wigner-Ville spectra merely involves a simple division of functions (see (10.40)), which can be stabilized more easily.

10.10. Conclusion

Time-varying spectra and, more fundamentally, the time-frequency (TF) domain allow a natural extension of the spectral representation of random processes from the stationary to the non-stationary case. As we have demonstrated in this chapter, it is

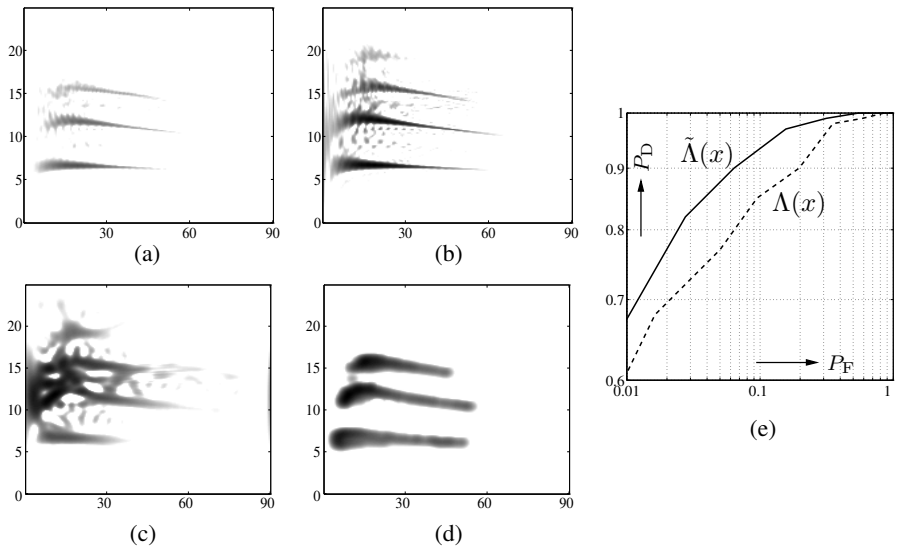

Figure 10.13. *Comparison of the optimum detector $\Lambda(x)$ and of the "TF detector" $\tilde{\Lambda}(x)$ for real processes $x_0(t)$ and $x_1(t)$ (knock detection): (a) Estimated Wigner-Ville spectrum of $x_0(t)$, (b) estimated Wigner-Ville spectrum of $x_1(t)$, (c) Weyl symbol of the operator $\boldsymbol{H}_{\mathrm{L}}$, (d) Weyl symbol of the operator $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$, (e) ROC curves of the detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$. In (a)–(d), the horizontal axis represents the crank angle (proportional to time) and the vertical axis represents frequency (in kHz). The signal duration is 186 samples*

then possible to extend classical spectral techniques for statistical signal processing (estimation and detection) to non-stationary processes, provided that the latter possess an "underspread" property. This property, which is often satisfied in practical applications, means that the correlations of the processes have a small TF horizon.

Our study of statistical signal processing has been centered on the TF formulation and TF design of the Wiener filter (estimator) and of the likelihood ratio detector. However, the fundamental approach – substituting a time-varying spectrum for the traditional power spectrum – can also be used to extend other concepts and methods of stationary statistical signal processing to the underspread non-stationary case.

We finally note that other TF methods for non-stationary estimation and detection have been proposed: concerning estimation of time-varying spectra [MAR 85, FLA 97, FLA 89, FLA 99, KOZ 94b, SAY 95b, RIE 93, KAY 94, KAY 92], estimation of non-stationary processes [ABD 69, SIL 95a, SIL 95b, KHA 97, SAY 95c, LAN 97], detection and classification of deterministic signals or non-stationary random signals in non-stationary noise [KAY 85, KUM 84, FLA 86a, FLA 88a, MAR 97, SAY 95a, SAY 96, MAT 96, HEI 95, FLA 88b, DAV 98, DAV 01, RIC 98, LEM 94, VIN 94], estimation and detection in subspaces [HLA 98, HLA 00, MAT 98b,

MAT 99b, MAT 02b], and robust estimation and detection using a minimax approach [MAT 98a, MAT 00a, MAT 99a, MAT 03]. A training-based TF methodology for classification will be presented in Chapter 13.

10.11. Acknowledgements

We would like to thank S. Carstens-Behrens, M. Wagner and J. F. Böhme for instructive discussions and for having placed at our disposal the engine signals used in Section 10.9.3 (courtesy of Aral-Forschung, Bochum). We would also like to thank F. Auger for reviewing this chapter and suggesting numerous improvements – grammatical, stylistic and others.

10.12. Bibliography

[ABD 69] ABDRABBO N. A., PRIESTLEY M. B., "Filtering non-stationary signals", *J. Roy. Stat. Soc. Ser. B*, vol. 31, pp. 150–159, 1969.

[AMI 92] AMIN M., "Time-frequency spectrum analysis and estimation for non-stationary random processes", BOASHASH B. (Ed.), *Advances in Spectrum Estimation*, pp. 208–232, Longman Cheshire, Melbourne, Australia, 1992.

[BEL 63] BELLO P. A., "Characterization of randomly time-variant linear channels", *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.

[CLA 80] CLAASEN T. A. C. M., MECKLENBRÄUKER W. F. G., "The Wigner distribution – A tool for time-frequency signal analysis; Parts I–III", *Philips J. Res.*, vol. 35, pp. 217–250, 276–300, and 372–389, 1980.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[CRA 61] CRAMÉR H., "On some classes of nonstationary stochastic processes", *Proc. 4th Berkeley Sympos. Math. Stat. Prob.*, University of California Press, pp. 57–78, 1961.

[CUN 94] CUNNINGHAM G. S., WILLIAMS W. J., "Kernel decomposition of time-frequency distributions", *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1425–1442, 1994.

[DAV 98] DAVY M., DONCARLI C., "Optimal kernels of time-frequency representations for signal classification", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Pittsburgh, PA, pp. 581–584, Oct. 1998.

[DAV 01] DAVY M., DONCARLI C., BOUDREAUX-BARTELS G. F., "Improved optimization of time-frequency-based signal classifiers", *IEEE Signal Process. Letters*, vol. 8, no. 2, pp. 52–57, Feb. 2001.

[DET 94] DETKA C. S., EL-JAROUDI A., "The transitory evolutionary spectrum", *Proc. IEEE ICASSP-94*, Adelaide, Australia, pp. 289–292, Apr. 1994.

[FLA 86a] FLANDRIN P., "On detection-estimation procedures in the time-frequency plane", *Proc. IEEE ICASSP-86*, Tokyo, Japan, pp. 2331–2334, 1986.

[FLA 86b] FLANDRIN P., "On the positivity of the Wigner-Ville spectrum", *Signal Process.*, vol. 11, no. 2, pp. 187–189, 1986.

[FLA 88a] FLANDRIN P., "A time-frequency formulation of optimum detection", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1377–1384, 1988.

[FLA 88b] FLANDRIN P., "Time-frequency receivers for locally optimum detection", *Proc. IEEE ICASSP-88*, New York, pp. 2725–2728, 1988.

[FLA 89] FLANDRIN P., "Time-dependent spectra for nonstationary stochastic processes", LONGO G., PICINBONO B. (Eds.), *Time and Frequency Representation of Signals and Systems*, pp. 69–124, Springer, Vienna, Austria, 1989.

[FLA 97] FLANDRIN P., MARTIN W., "The Wigner-Ville spectrum of nonstationary random signals", MECKLENBRÄUKER W., HLAWATSCH F. (Eds.), *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 211–267, Elsevier, Amsterdam, The Netherlands, 1997.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[FOL 89] FOLLAND G. B., *Harmonic Analysis in Phase Space*, vol. 122 of *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, 1989.

[HEI 95] HEITZ C., "Optimum time-frequency representations for the classification and detection of signals", *Applied Signal Processing*, vol. 3, pp. 124–143, 1995.

[HLA 92] HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Magazine*, vol. 9, no. 2, pp. 21–67, Apr. 1992.

[HLA 95] HLAWATSCH F., KOZEK W., "Second-order time-frequency synthesis of nonstationary random processes", *IEEE Trans. Inform. Theory*, vol. 41, no. 1, pp. 255–267, Jan. 1995.

[HLA 97] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations", MECKLENBRÄUKER W., HLAWATSCH F. (Eds.), *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam, The Netherlands, 1997.

[HLA 98] HLAWATSCH F., *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation*, Kluwer, Boston, MA, 1998.

[HLA 00] HLAWATSCH F., MATZ G., KIRCHAUER H., KOZEK W., "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation", *IEEE Trans. Signal Process.*, vol. 48, pp. 1417–1432, May 2000.

[JAN 82] JANSSEN A. J. E. M., "On the locus and spread of pseudo-density functions in the time-frequency plane", *Philips J. Res.*, vol. 37, no. 3, pp. 79–110, 1982.

[JAN 89] JANSSEN A. J. E. M., "Wigner weight functions and Weyl symbols of non-negative definite linear operators", *Philips J. Res.*, vol. 44, pp. 7–42, 1989.

[KAY 85] KAY S. M., BOUDREAUX-BARTELS G. F., "On the optimality of the Wigner distribution for detection", *Proc. IEEE ICASSP-85*, Tampa, FL, pp. 1017–1020, Mar. 1985.

[KAY 92] KAYHAN A. S., CHAPARRO L. F., EL-JAROUDI A., "Wold-Cramer evolutionary spectral estimators", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Victoria, Canada, pp. 115–118, Oct. 1992.

[KAY 94] KAYHAN A. S., EL-JAROUDI A., CHAPARRO L. F., "Evolutionary periodogram for nonstationary signals", *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1527–1536, Jun. 1994.

[KHA 97] KHAN H. A., CHAPARRO L. F., "Nonstationary Wiener filtering based on evolutionary spectral theory", *Proc. IEEE ICASSP-97*, Munich, Germany, pp. 3677–3680, May 1997.

[KOH 65] KOHN J. J., NIRENBERG L., "An algebra of pseudo-differential operators", *Comm. Pure Appl. Math.*, vol. 18, pp. 269–305, 1965.

[KÖN 94] KÖNIG D., BÖHME J. F., "Application of cyclostationary and time-frequency signal analysis to car engine diagnosis", *Proc. IEEE ICASSP-94*, Adelaide, Australia, pp. 149–152, Apr. 1994.

[KOZ 92a] KOZEK W., "On the generalized Weyl correspondence and its application to time-frequency analysis of linear time-varying systems", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Victoria, Canada, pp. 167–170, Oct. 1992.

[KOZ 92b] KOZEK W., "Time-frequency signal processing based on the Wigner-Weyl framework", *Signal Process.*, vol. 29, no. 1, pp. 77–92, Oct. 1992.

[KOZ 94a] KOZEK W., HLAWATSCH F., KIRCHAUER H., TRAUTWEIN U., "Correlative time-frequency analysis and classification of nonstationary random processes", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 417–420, Oct. 1994.

[KOZ 94b] KOZEK W., RIEDEL K., "Quadratic time-varying spectral estimation for underspread processes", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 460–463, Oct. 1994.

[KOZ 96] KOZEK W., FEICHTINGER H. G., SCHARINGER J., "Matched multiwindow methods for the estimation and filtering of nonstationary processes", *Proc. IEEE ISCAS-96*, Atlanta, GA, pp. 509–512, May 1996.

[KOZ 97a] KOZEK W., Matched Weyl-Heisenberg Expansions of Nonstationary Environments, Doctoral dissertation, Vienna University of Technology, Mar. 1997.

[KOZ 97b] KOZEK W., "On the transfer function calculus for underspread LTV channels", *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 219–223, Jan. 1997.

[KUM 84] KUMAR B. V. K., CARROLL C. W., "Performance of Wigner distribution function based detection methods", *Opt. Eng.*, vol. 23, no. 6, pp. 732–737, Nov. 1984.

[LAN 97] LANDER P., BERBARI E. J., "Time-frequency plane Wiener filtering of the high-resolution ECG: Background and time-frequency representations", *IEEE Trans. Biomed. Eng.*, vol. 44, no. 4, pp. 247–255, Apr. 1997.

[LEM 94] LEMOINE O., ICART S., BERENGUER C., "Detection of transient signals using time-frequency distributions", *Proc. EUSIPCO-94*, Edinburgh, UK, pp. 272–275, Sep. 1994.

[LEV 64] LEVIN M. J., "Instantaneous spectra and ambiguity functions", *IEEE Trans. Inform. Theory*, vol. 10, pp. 95–97, 1964.

[LOY 68] LOYNES R. M., "On the concept of the spectrum for non-stationary processes", *J. Roy. Stat. Soc. Ser. B*, vol. 30, no. 1, pp. 1–30, 1968.

[MAL 99] MALLAT S. G., *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, San Diego, CA, 1999.

[MAL 98] MALLAT S. G., PAPANICOLAOU G., ZHANG Z., "Adaptive covariance estimation of locally stationary processes", *Ann. Stat.*, vol. 26, no. 1, pp. 1–47, Feb. 1998.

[MAR 70] MARK W. D., "Spectral analysis of the convolution and filtering of non-stationary stochastic processes", *J. Sound Vib.*, vol. 11, no. 1, pp. 19–63, 1970.

[MAR 85] MARTIN W., FLANDRIN P., "Wigner-Ville spectral analysis of nonstationary processes", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1461–1470, Dec. 1985.

[MAR 97] MARINOVICH N. M., "The singular value decomposition of the Wigner distribution and its applications", MECKLENBRÄUKER W., HLAWATSCH F. (Eds.), *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 319–373, Elsevier, Amsterdam, The Netherlands, 1997.

[MAT 96] MATZ G., HLAWATSCH F., "Time-frequency formulation and design of optimal detectors", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 213–216, Jun. 1996.

[MAT 97] MATZ G., HLAWATSCH F., KOZEK W., "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes", *IEEE Trans. Signal Process.*, vol. 45, no. 6, pp. 1520–1534, Jun. 1997.

[MAT 98a] MATZ G., HLAWATSCH F., "Robust time-varying Wiener filters: Theory and time-frequency formulation", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Pittsburgh, PA, pp. 401–404, Oct. 1998.

[MAT 98b] MATZ G., HLAWATSCH F., "Time-frequency methods for signal detection with application to the detection of knock in car engines", *Proc. IEEE-SP Workshop Statist. Signal Array Process.*, Portland, OR, pp. 196–199, Sep. 1998.

[MAT 98c] MATZ G., HLAWATSCH F., "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory", *J. Math. Phys., Special Issue on Wavelet and Time-Frequency Analysis*, vol. 39, no. 8, pp. 4041–4071, Aug. 1998.

[MAT 99a] MATZ G., HLAWATSCH F., "Minimax robust time-frequency filters for nonstationary signal estimation", *Proc. IEEE ICASSP-99*, Phoenix, AZ, pp. 1333–1336, Mar. 1999.

[MAT 99b] MATZ G., HLAWATSCH F., "Time-frequency subspace detectors and application to knock detection", *Int. J. Electron. Commun. (AEÜ)*, vol. 53, no. 6, pp. 379–385, 1999.

[MAT 00a] MATZ G., HLAWATSCH F., "Minimax robust nonstationary signal estimation based on a p-point uncertainty model", *J. Franklin Inst., Special Issue on Time-Frequency Signal Analysis and Its Applications*, vol. 337, pp. 403–419, Jul. 2000.

[MAT 00b] MATZ G., A time-frequency calculus for time-varying systems and nonstationary processes with applications, Doctoral dissertation, Vienna University of Technology, Nov. 2000.

[MAT 01] MATZ G., HLAWATSCH F., RAIDL A., "Signal-adaptive robust time-varying Wiener filters: Best subspace selection and statistical analysis", *Proc. IEEE ICASSP-01*, Salt Lake City, UT, pp. 3945–3948, May 2001.

[MAT 02a] MATZ G., HLAWATSCH F., "Linear time-frequency filters: Online algorithms and applications", PAPANDREOU-SUPPAPPOLA A. (Ed.), *Applications in Time-Frequency Signal Processing*, pp. 205–271, CRC Press, Boca Raton, FL, 2002.

[MAT 02b] MATZ G., HLAWATSCH F., "Time-frequency projection filters: Online implementation, subspace tracking, and application to interference suppression", *Proc. IEEE ICASSP-02*, Orlando, FL, pp. 1213–1216, May 2002.

[MAT 03] MATZ G., RAIDL A., "Minimax robust nonstationary signal detection for p-point uncertainty classes", *Proc. IEEE ICASSP-03*, Hong Kong, pp. 641–644, Apr. 2003.

[MAT 06] MATZ G., HLAWATSCH F., "Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations", *IEEE Trans. Information Theory*, vol. 52, no. 3, pp. 1067–1086, Mar. 2006.

[PAG 52] PAGE C. H., "Instantaneous power spectra", *J. Appl. Phys.*, vol. 23, no. 1, pp. 103–106, Jan. 1952.

[PAP 84] PAPOULIS A., *Signal Analysis*, McGraw-Hill, New York, 1984.

[PAP 91] PAPOULIS A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 3rd edition, 1991.

[POO 88] POOR H. V., *An Introduction to Signal Detection and Estimation*, Springer, New York, 1988.

[PRI 65] PRIESTLEY M. B., "Evolutionary spectra and non-stationary processes", *J. Roy. Stat. Soc. Ser. B*, vol. 27, no. 2, pp. 204–237, 1965.

[PRI 81] PRIESTLEY M. B., *Spectral Analysis and Time Series – Part II*, Academic Press, London, 1981.

[RIC 98] RICHARD C., LENGELLE R., "Two algorithms for designing optimal reduced-bias data-driven time-frequency detectors", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Pittsburgh, PA, pp. 601–604, Oct. 1998.

[RIE 93] RIEDEL K., "Optimal data-based kernel estimation of evolutionary spectra", *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2439–2447, Jul. 1993.

[RIH 68] RIHACZEK A. W., "Signal energy distribution in time and frequency", *IEEE Trans. Inform. Theory*, vol. 14, no. 3, pp. 369–374, May 1968.

[SAY 95a] SAYEED A. M., JONES D. L., "Optimal detection using bilinear time-frequency and time-scale representations", *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2872–2883, Dec. 1995.

[SAY 95b] SAYEED A. M., JONES D. L., "Optimal kernels for nonstationary spectral estimation", *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 478–491, Feb. 1995.

[SAY 95c] SAYEED A. M., LANDER P., JONES D. L., "Improved time-frequency filtering of signal-averaged electrocardiograms", *J. Electrocard.*, vol. 28, pp. 53–58, 1995.

[SAY 96] SAYEED A. M., JONES D. L., "Optimal reduced-rank time-frequency/time-scale detectors", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 209–212, Jun. 1996.

[SCH 91] SCHARF L. L., *Statistical Signal Processing*, Addison Wesley, Reading, MA, 1991.

[SHE 94] SHENOY R. G., PARKS T. W., "The Weyl correspondence and time-frequency analysis", *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 318–331, Feb. 1994.

[SIL 95a] SILLS J. A., KAMEN E. W., "Wiener filtering of nonstationary signals based on spectral density functions", *Proc. 34th IEEE Conf. Decision Control*, Kobe, Japan, pp. 2521–2526, Dec. 1995.

[SIL 95b] SILLS J. A., Nonstationary signal modeling, filtering, and parameterization, PhD thesis, Georgia Institute of Technology, Atlanta, GA, Mar. 1995.

[THE 92] THERRIEN C. W., *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1992.

[THO 82] THOMSON D. J., "Spectrum estimation and harmonic analysis", *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, 1982.

[THO 93] THOMÄ R., STEFFENS J., TRAUTWEIN U., "Statistical cross-terms in quadratic time-frequency distributions", *Proc. Int. Conf. DSP*, Nicosia, Cyprus, Jul. 1993.

[VAN 68] VAN TREES H. L., *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*, Wiley, New York, 1968.

[VIN 94] VINCENT I., DONCARLI C., LE CARPENTIER E., "Non stationary signals classification using time-frequency distributions", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 233–236, Oct. 1994.

[WIE 49] WIENER N., *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, MIT Press, Cambridge, MA, 1949.

[ZAD 50] ZADEH L. A., "Frequency analysis of variable networks", *Proc. IRE*, vol. 76, pp. 291–299, Mar. 1950.