Time-Frequency Signal Processing: A Statistical Perspective*

Franz Hlawatsch and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Wien, Austria
phone: +43 1 58801 38916, fax: +43 1 5870583, email: fhlawats@email.tuwien.ac.at
web: http://www.nt.tuwien.ac.at/dspgroup/time.html

Abstract—**Time-frequency methods are capable of analyzing and/or processing nonstationary signals and systems in an intuitively appealing and physically meaningful manner. This tutorial paper presents an overview of some time-frequency methods for the analysis and processing of nonstationary random signals, with emphasis placed on time-varying power spectra and techniques for signal estimation and detection.**

We discuss two major definitions of time-dependent power spectra—the generalized Wigner-Ville spectrum and the generalized evolutionary spectrum—and show their approximate equivalence for *underspread* random processes. Time-dependent power spectra are then applied to nonstationary signal estimation and detection. Specifically, simple expressions and designs of signal estimators (Wiener filters) and signal detectors in the stationary case are extended to underspread nonstationary processes. This results in time-frequency techniques for nonstationary signal estimation and detection which are intuitively meaningful as well as efficient and stable.

I. Introduction

It is an interesting fact that most papers on time-frequency analysis consider deterministic signals whereas a large and important group of applications require signals to be modeled as random processes. As long as these random processes are stationary, a need for time-frequency methods does not arise since the *power spectral density*[1]

$$S_x(f) \;=\; \int_\tau r_x(\tau)\, e^{-j2\pi f\tau}\, d\tau \,, \tag{1}$$

with $r_x(\tau) = \mathrm{E}\{x(t+\tau)\, x^*(t)\}$, provides a complete and unique description of the process' second-order statistics and spectral properties [1]. In particular, due to stationarity the power spectral density does not change with time.

When the process under analysis is *nonstationary*, it is intuitively clear that its spectral properties change with time and, hence, a meaningful representation of these spectral properties must depend on a time variable. Thus, we look for a *time-dependent power spectrum* of the form $S_x(t,f)$, which can be interpreted as a time-frequency representation of the process' second-order statistics.

This situation is somewhat analogous (and, as we will see presently, closely related) to the frequency-domain analysis of linear systems. As long as the system is time-invariant, the *frequency response*

$$H(f) \;=\; \int_\tau h(\tau)\, e^{-j2\pi f\tau}\, d\tau \,, \tag{2}$$

with $h(\tau)$ the system's impulse response, completely and uniquely describes the system's frequency-domain characteristics [2]. However, for a *time-varying* linear system, the frequency-domain characteristics will be time-dependent and we thus must look for a *time-dependent frequency response* of the form $H(t,f)$, which can again be interpreted as a time-frequency representation of the system.

In this tutorial paper, we will discuss time-dependent power spectra and how they can be used for signal estimation and signal detection in nonstationary environments. In Section II, we consider two different approaches to defining a time-dependent power spectrum for nonstationary random processes. Section III discusses the concept of *underspread* processes and shows that the two time-dependent power spectra of Section II become effectively equivalent in the underspread case. The estimation (optimal filtering) and detection of nonstationary random processes will be considered in Sections IV and V, respectively.

II. Time-Dependent Power Spectra

Apart from the *physical spectrum*, which can be interpreted as the expectation of the spectrogram [3–5], there are two fundamentally different approaches to defining a time-dependent power spectrum for a nonstationary random process $x(t)$.

A. Generalized Wigner-Ville Spectrum

The first group of time-dependent spectra, called *generalized Wigner-Ville spectrum* [4–7], is a formally simple extension of the stationary power spectral density $S_x(f)$ in (1):

$$\overline{W}_x^{(\alpha)}(t,f) \;\triangleq\; \int_\tau r_x^{(\alpha)}(t,\tau)\, e^{-j2\pi f\tau}\, d\tau \tag{3}$$

with

$$r_x^{(\alpha)}(t,\tau) \;\triangleq\; r_x\!\left(t+\left(\frac{1}{2}-\alpha\right)\tau,\; t-\left(\frac{1}{2}+\alpha\right)\tau\right), \tag{4}$$

where $r_x(t,t') = \mathrm{E}\{x(t)\, x^*(t')\}$. The parameter $\alpha \in \mathbb{R}$ is arbitrary *a priori*. Special cases are the *Wigner-Ville spectrum* ($\alpha = 0$) [3–8] and the *Rihaczek spectrum* ($\alpha = 1/2$) [4,5,9]. The case $\alpha = 0$ has certain advantages over other choices of α; in particular, $\overline{W}_x^{(0)}(t,f)$ is always real-valued (however, it is not guaranteed to be everywhere nonnegative although it is approximately nonnegative for the practically important underspread processes—see Section III).

*Funding by FWF grant P11904-TEC.

[1] Throughout this paper, $x(t)$ denotes a deterministic signal or a random process and integrals are from $-\infty$ to ∞.

Under appropriate conditions, $\overline{W}_x^{(\alpha)}(t,f)$ can be interpreted as the expected value of the so-called generalized Wigner distribution [10–12] (cf. (19)). For any α, $\overline{W}_x^{(\alpha)}(t,f)$ reduces to the power spectral density $S_x(f)$ in (1) if the process $x(t)$ is (wide-sense) stationary.

B. Generalized Evolutionary Spectrum

The second group of time-dependent spectra, called *generalized evolutionary spectrum* [7,13], is based on an innovations system representation of the nonstationary process $x(t)$ [14]. That is, $x(t)$ is represented as the output of a linear, time-varying system (linear operator) $\mathbf{H}$ whose input is stationary white noise $n(t)$ with normalized power spectral density:

$$x(t) \,=\, (\mathbf{H}n)(t) = \int_{t'} h(t,t')\, n(t')\, dt',$$

with $r_n(t,t') = \delta(t-t')$ and $h(t,t')$ denoting the impulse response (kernel) of $\mathbf{H}$. (In the stationary case, $\mathbf{H}$ is time-invariant and $S_x(f) = |H(f)|^2$.) The innovations system $\mathbf{H}$ is obtained as a solution to the factorization problem $\mathbf{H}\mathbf{H}^+ = \mathbf{R}_x$, where $\mathbf{R}_x$ is the correlation operator[2] of the process $x(t)$ and $\mathbf{H}^+$ is the adjoint of $\mathbf{H}$. Thus, $\mathbf{H}$ is a "square root" of $\mathbf{R}_x$. This square root is unique only up to a factor $\mathbf{A}$ satisfying $\mathbf{A}\mathbf{A}^+ = \mathbf{I}$: if $\mathbf{H}$ is a valid innovations system satisfying $\mathbf{H}\mathbf{H}^+ = \mathbf{R}_x$ and $\mathbf{A}$ satisfies $\mathbf{A}\mathbf{A}^+ = \mathbf{I}$, then it is easily seen that $\mathbf{H}' = \mathbf{H}\mathbf{A}$ is a valid innovations system as well.

A "time-dependent frequency response" of $\mathbf{H}$ that extends the frequency response of time-invariant systems in (2) is given by the *generalized Weyl symbol* [15–17]

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \;\triangleq\; \int_\tau h^{(\alpha)}(t,\tau)\, e^{-j2\pi f\tau}\, d\tau \qquad (5)$$

with

$$h^{(\alpha)}(t,\tau) \;\triangleq\; h\bigg(t+\Big(\frac{1}{2}-\alpha\Big)\tau,\, t-\Big(\frac{1}{2}+\alpha\Big)\tau\bigg), \qquad (6)$$

where $\alpha \in \mathbb{R}$. Special cases are the *Weyl symbol* for $\alpha = 0$ [16–21], Zadeh's time-varying frequency response for $\alpha = 1/2$ [16,17,20,22,23], and the *Kohn-Nirenberg symbol* (equivalently, Bello's *frequency-dependent modulation function*) for $\alpha = -1/2$ [18,23,24]; the case $\alpha = 0$ has again certain advantages over other choices of α [17,18,20]. Note that the generalized Wigner-Ville spectrum in (3) is the generalized Weyl symbol of the correlation operator $\mathbf{R}_x$, i.e.,

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

The *generalized evolutionary spectrum* is now defined as the squared magnitude of the generalized Weyl symbol of the innovations system $\mathbf{H}$ [13]:

$$\mathrm{GES}_x^{(\alpha)}(t,f) \;\triangleq\; \big|L_{\mathbf{H}}^{(\alpha)}(t,f)\big|^2.$$

Note that this definition contains a twofold ambiguity related to the choice of $\alpha \in \mathbb{R}$ and the choice of the innovations system $\mathbf{H}$ for given correlation operator $\mathbf{R}_x$. Special cases are the *evolutionary spectrum* ($\alpha = 1/2$) [25–28] and the *transitory evolutionary spectrum* ($\alpha = -1/2$) [13,29]; furthermore, the *Weyl spectrum* is obtained with $\alpha = 0$ and $\mathbf{H}$ chosen as the positive (semi-)definite square root of $\mathbf{R}_x$ (this choice has certain advantages over other choices of α and $\mathbf{H}$) [13]. Note that $\mathrm{GES}_x^{(\alpha)}(t,f) \geq 0$.

For a wide-sense stationary process $x(t)$, the innovations system $\mathbf{H}$ can always be chosen to be time-invariant, in which case the generalized Weyl symbol $L_{\mathbf{H}}^{(\alpha)}(t,f)$ reduces to the system's frequency response $H(f)$. Hence, the generalized evolutionary spectrum here reduces to the power spectral density: $\mathrm{GES}_x^{(\alpha)}(t,f) = |H(f)|^2 = S_x(f)$.

III. Underspread Systems and Processes

For the results obtained with the two classes of time-dependent spectra described above, the time-frequency displacements caused by the innovations system $\mathbf{H}$ and the time-frequency correlation structure of the resulting process $x(t)$ play an important role. A limitation of these time-frequency displacements or time-frequency correlations leads to the important concepts of *underspread* systems or processes, respectively. Broadly speaking, an intuitively pleasing and meaningful interpretation of time-dependent spectra is only possible for underspread processes.

A. Underspread Systems

The time-frequency shifts caused by a linear time-varying system $\mathbf{H}$ are characterized by the *generalized spreading function* [15–17]

$$S_{\mathbf{H}}^{(\alpha)}(\tau,\nu) \;\triangleq\; \int_t h^{(\alpha)}(t,\tau)\, e^{-j2\pi\nu t}\, dt\,, \qquad (7)$$

with $h^{(\alpha)}(t,\tau)$ as defined in (6). It can be shown [15,16] that the magnitude of $S_{\mathbf{H}}^{(\alpha)}(\tau,\nu)$ is independent of α, so that we may write $|S_{\mathbf{H}}^{(\alpha)}(\tau,\nu)| = |S_{\mathbf{H}}(\tau,\nu)|$. Furthermore, $S_{\mathbf{H}}^{(\alpha)}(\tau,\nu)$ is the 2-D Fourier transform of the generalized Weyl symbol in (5).

The generalized spreading function $S_{\mathbf{H}}^{(\alpha)}(\tau,\nu)$ is the coefficient function of an expansion of $\mathbf{H}$ into elementary time-frequency shifts $\mathbf{S}_{\tau,\nu}^{(\alpha)}$, where $\big(\mathbf{S}_{\tau,\nu}^{(\alpha)} x\big)(t) = x(t-\tau)\, e^{j2\pi\nu t} e^{j2\pi(\alpha-1/2)\nu\tau}$ [15–18,20,21,23,30,31]. Hence, for a given (τ,ν), $|S_{\mathbf{H}}(\tau,\nu)|$ indicates how much the time-frequency shifted input signal $\big(\mathbf{S}_{\tau,\nu}^{(\alpha)} x\big)(t) = x(t-\tau)\, e^{j2\pi\nu t} e^{j2\pi(\alpha-1/2)\nu\tau}$ contributes to the output signal. It follows that the time-frequency shifts caused by a linear time-varying system are crudely characterized by the effective support of $|S_{\mathbf{H}}(\tau,\nu)|$.

A system $\mathbf{H}$ is now called *underspread* if $|S_{\mathbf{H}}(\tau,\nu)|$ is concentrated about the origin of the (τ,ν)-plane, so that $\mathbf{H}$ causes only small time-frequency shifts (mathematically precise definitions of underspread systems can be found in [16,17,30,31]). It should be noted that underspread is not equivalent to slowly time-varying, since slow time variation only means a limitation of $|S_{\mathbf{H}}(\tau,\nu)|$ with respect to ν. In contrast, underspread means limitations with respect to both τ and ν; however, the extents of these limitations can be exchanged for one another. Hence, a slowly time-varying system will not be underspread if its memory (extension of $|S_{\mathbf{H}}(\tau,\nu)|$ with respect to τ) is too

[2] The correlation operator $\mathbf{R}_x$ is the positive (semi-)definite linear operator whose kernel equals the correlation $r_x(t,t') = \mathrm{E}\{x(t)\, x^*(t')\}$.

long, whereas a fast time-varying system will be underspread if its memory is sufficiently short.

B. Underspread Processes

Quasi-stationary processes have limited spectral correlation, while quasi-white processes have limited temporal correlation. These two situations are generalized by the concept of *underspread processes.* We first define the *expected generalized ambiguity function* [6,16] of a nonstationary process $x(t)$ as

$$\bar{A}_x^{(\alpha)}(\tau,\nu) \;\triangleq\; \int_t r_x^{(\alpha)}(t,\tau)\, e^{-j2\pi\nu t}\, dt \;=\; \mathrm{E}\big\{\big\langle x, \mathbf{S}_{\tau,\nu}^{(\alpha)} x\big\rangle\big\}\,,$$

with $r_x^{(\alpha)}(t,\tau)$ as in (4). The expected generalized ambiguity function is the generalized spreading function (see (7)) of the correlation operator $\mathbf{R}_x$, i.e.,

$$\bar{A}_x^{(\alpha)}(\tau,\nu) \;=\; S_{\mathbf{R}_x}^{(\alpha)}(\tau,\nu)\,;$$

it is furthermore the 2-D Fourier transform of the generalized Wigner-Ville spectrum in (3). It can be shown [6,16] that the expected generalized ambiguity function $\bar{A}_x^{(\alpha)}(\tau,\nu)$ describes the average correlation of all time-frequency locations separated by τ in time and by ν in frequency.

A nonstationary process $x(t)$ is now called *underspread* if $|\bar{A}_x^{(\alpha)}(\tau,\nu)| \;=\; |\bar{A}_x(\tau,\nu)|$ is concentrated about the origin of the (τ,ν)-plane, so that components of $x(t)$ that are sufficiently separated in the time-frequency plane will be nearly uncorrelated (mathematically precise definitions of underspread processes can be found in [6,7,16]; furthermore, somewhat similar concepts of processes with limited time-frequency correlation have been discussed in [32–35]). This underspread property is satisfied by many nonstationary processes occurring in practice. We emphasize that underspread should not be confused with quasistationarity which only means a limitation of $|\bar{A}_x(\tau,\nu)|$ with respect to ν. In contrast, underspread means limitations with respect to both τ and ν, where again the extents of these limitations can be exchanged for one another. Hence, a quasistationary process will not be underspread if its correlation time (extension of $|\bar{A}_x(\tau,\nu)|$ with respect to τ) is too long, whereas a fast nonstationary process will be underspread if its correlation time is sufficiently short (quasiwhite process).

The concepts of underspread systems and underspread processes are related since $\bar{A}_x^{(\alpha)}(\tau,\nu)$ is the generalized spreading function of $\mathbf{R}_x$, and hence a process $x(t)$ is underspread if and only if its correlation operator $\mathbf{R}_x$ is an underspread system. Furthermore, the time-frequency shifts caused by the innovations system $\mathbf{H}$ are related to the time-frequency correlation structure of the associated process $x(t)$. If the innovations system $\mathbf{H}$ is underspread, the correlation operator $\mathbf{R}_x = \mathbf{H}\mathbf{H}^+$ is an underspread system as well, and hence the process $x(t)$ is itself underspread. Conversely, if $x(t)$ is underspread, then not every innovations system $\mathbf{H}$ is underspread but an underspread $\mathbf{H}$ can always be found.

C. Equivalence of Time-Dependent Spectra

The importance of the underspread property in the context of nonstationary spectral analysis is due to the approximate equivalence of all above-mentioned time-dependent spectra in the case of underspread processes. This equivalence is based on the following two approximations valid for the generalized Weyl symbol of underspread systems (bounds on the associated approximation errors are provided in [16,17,30]):

1. The generalized Weyl symbol of an underspread system $\mathbf{H}$ is approximately independent of α, i.e.,

$$L_{\mathbf{H}}^{(\alpha_1)}(t,f) \;\approx\; L_{\mathbf{H}}^{(\alpha_2)}(t,f)\,. \tag{8}$$

2. For an underspread system $\mathbf{H}$, there is

$$L_{\mathbf{H}\mathbf{H}^+}^{(\alpha)}(t,f) \;\approx\; \big|L_{\mathbf{H}}^{(\alpha)}(t,f)\big|^2. \tag{9}$$

With these approximations, we now obtain the following equivalence results valid for an underspread process $x(t)$:

- With $\overline{W}_x^{(\alpha)}(t,f) = L_{\mathbf{R}_x}^{(\alpha)}(t,f)$, and since for an underspread process $\mathbf{R}_x$ is an underspread system, it follows from (8) that the generalized Wigner-Ville spectrum is approximately independent of α, i.e., [7,16]

$$\overline{W}_x^{(\alpha_1)}(t,f) \;\approx\; \overline{W}_x^{(\alpha_2)}(t,f)\,.$$

- With $\mathrm{GES}_x^{(\alpha)}(t,f) = \big|L_{\mathbf{H}}^{(\alpha)}(t,f)\big|^2$, and since for an underspread process we can always find an underspread innovations system $\mathbf{H}$, it further follows from (8) that the generalized evolutionary spectrum (based on an underspread innovations system $\mathbf{H}$) is approximately independent of α, i.e., [7,13,16]

$$\mathrm{GES}_x^{(\alpha_1)}(t,f) \;\approx\; \mathrm{GES}_x^{(\alpha_2)}(t,f)\,.$$

- With $\mathrm{GES}_x^{(\alpha)}(t,f) = \big|L_{\mathbf{H}}^{(\alpha)}(t,f)\big|^2$ and $\overline{W}_x^{(\alpha)}(t,f) = L_{\mathbf{R}_x}^{(\alpha)}(t,f) = L_{\mathbf{H}\mathbf{H}^+}^{(\alpha)}(t,f)$, and using an underspread innovations system $\mathbf{H}$, it follows from (9) that the generalized evolutionary spectrum is approximately equal to the generalized Wigner-Ville spectrum, i.e., [7,13,16]

$$\mathrm{GES}_x^{(\alpha)}(t,f) \;\approx\; \overline{W}_x^{(\alpha)}(t,f)\,.$$

Since $\mathrm{GES}_x^{(\alpha)}(t,f) \geq 0$, this also shows that the generalized Wigner-Ville spectrum of an underspread process is approximately real-valued and nonnegative.

These equivalence results simplify the spectral analysis of nonstationary processes a great deal: Even though there exist infinitely many different time-dependent spectra (there are the two distinct classes of generalized Wigner-Ville spectrum and generalized evolutionary spectrum, plus the dependence on α within each class), *all these spectra are approximately equivalent for underspread processes.*

This approximate equivalence is demonstrated by **Fig. 1** which compares various time-dependent spectra of an underspread process. It is seen that all spectra are very similar and, furthermore, that they are *smooth* time-frequency functions. In fact, the spectra of underspread processes must be smooth functions since the generalized Wigner-Ville spectrum is the 2-D Fourier transform of the expected generalized ambiguity function (which is concentrated about the origin) and the generalized

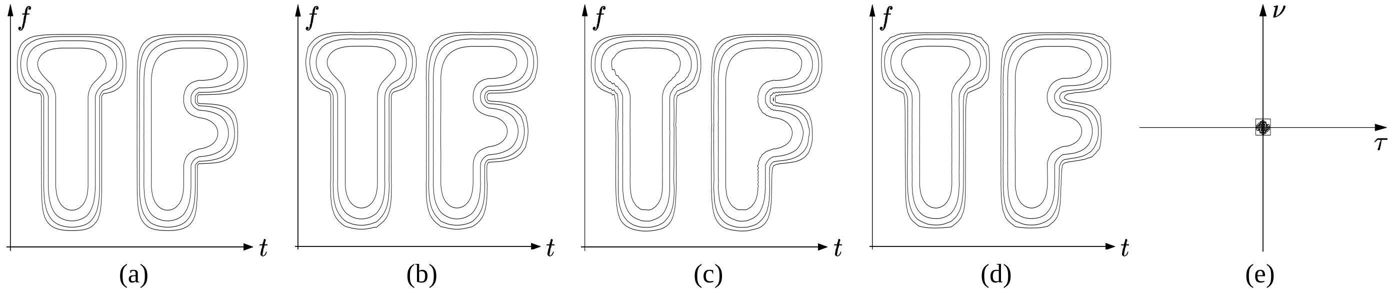


Figure 1. *Time-dependent spectra of an underspread process: (a) Wigner-Ville spectrum, (b) real part of Rihaczek spectrum, (c) Weyl spectrum, (d) (transitory) evolutionary spectrum using a positive semidefinite innovations system (here, the evolutionary spectrum equals the transitory evolutionary spectrum since the positive semidefinite innovations system is used [13]), (e) magnitude of expected ambiguity function (the rectangle shown has area 1 and thus permits to assess the underspread property of the process). The process was generated by means of the time-frequency synthesis technique introduced in [36]. The signal length is 256 samples.*

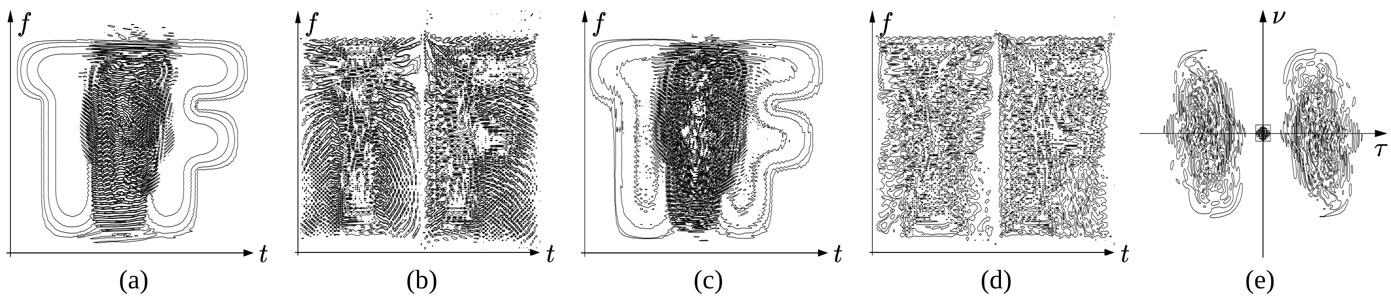


Figure 2. *Time-dependent spectra of an overspread process: (a) Wigner-Ville spectrum, (b) real part of Rihaczek spectrum, (c) Weyl spectrum, (d) (transitory) evolutionary spectrum, (e) magnitude of expected ambiguity function. The overspread characteristic of this process is due to strong statistical correlations between the 'T' and 'F' components. Note that the expected ambiguity function in (e) is widely spread out beyond the rectangle with area 1.*

evolutionary spectrum is similar to the generalized Wigner-Ville spectrum.

In contrast, **Fig. 2** shows that the various spectra yield dramatically different results for an "overspread" process (i.e., a process whose expected generalized ambiguity function is not sufficiently concentrated about the origin of the (τ, ν)-plane, cf. part (e)). Furthermore, the spectra are no longer smooth functions; they contain rapidly oscillating components (cross terms) that can be attributed to the strong statistical correlations existing between widely separated time-frequency points [7,37].

IV. Nonstationary Signal Estimation

In the remainder of this paper, we shall show how time-dependent spectra can be applied to nonstationary signal estimation and detection. We shall use the generalized Wigner-Ville spectrum because it has a mathematically simple structure. However, for underspread processes the generalized evolutionary spectrum is approximately equivalent to the generalized Wigner-Ville spectrum (as explained above), and hence it can be substituted for the generalized Wigner-Ville spectrum in the relevant equations. Other approaches to nonstationary signal estimation are discussed in [34,35,38–41].

The enhancement or estimation of signals corrupted by noise or interference is important in many signal processing applications. In this section, we consider the estimation of a nonstationary random signal $s(t)$ from an observation $x(t) = s(t) + n(t)$, where $n(t)$ is nonstationary noise uncorrelated with $s(t)$, by means of a linear, time-varying system $\mathbf{H}$. Hence, the signal estimate is given by

$$\hat{s}(t) \,=\, (\mathbf{H}x)(t) \,=\, \int_{t'} h(t,t')\, x(t')\, dt'.$$

The system that minimizes the mean-square error is the *time-varying Wiener filter* [42–45]

$$\mathbf{H}_W \,=\, \mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}, \tag{10}$$

where $\mathbf{R}_s$ and $\mathbf{R}_n$ denote the correlation operators of signal and noise, respectively. For stationary random processes, $\mathbf{H}$ simplifies to a time-invariant system whose frequency response is given by [1,42–46]

$$H_W(f) \,=\, \frac{S_s(f)}{S_s(f) + S_n(f)}\,, \tag{11}$$

where $S_s(f)$ and $S_n(f)$ denote the power spectral densities of signal and noise, respectively. This frequency-domain expression contains simple products and reciprocals of functions (instead of products and inverses of operators as in (10)) and hence allows a simple design and interpretation of time-invariant Wiener filters.

A. Time-Frequency Formulation of Time-Varying Wiener Filters

We may ask whether a similarly simple formulation as in the stationary case can be obtained for the time-varying Wiener filter $\mathbf{H}_W$ by introducing in (11) an explicit time dependence, i.e., by substituting for $H_W(f)$ and for $S_s(f), S_n(f)$ suitably defined time-dependent frequency responses and time-dependent power spectra. Indeed, for *jointly underspread*[3] processes $s(t)$ and $n(t)$ it can be shown [47,48] that the time-varying Wiener filter $\mathbf{H}_W$ can be decomposed into an underspread component and an overspread component with the following properties:

- The overspread system component has negligible effect on the system's performance (mean-square error) and can hence be disregarded.

- The underspread part, denoted as $\mathbf{H}_W^u$ in what follows, allows the approximate time-frequency formulation

$$L^{(\alpha)}_{\mathbf{H}_W^u}(t,f) \;\approx\; \frac{\overline{W}_s^{(\alpha)}(t,f)}{\overline{W}_s^{(\alpha)}(t,f) + \overline{W}_n^{(\alpha)}(t,f)}\,, \tag{12}$$

where $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$ denote the generalized Wigner-Ville spectra of signal and noise, respectively.

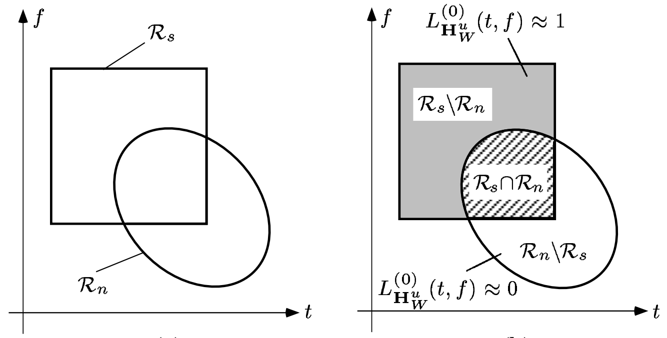
The time-frequency formulation in (12) provides the looked-for extension of the frequency-domain formulation (11) to the nonstationary (underspread) case. For $\alpha = 0$ (recall that $\overline{W}_{\cdot}^{(0)}(t,f)$ is real-valued), (12) allows a simple and intuitively appealing time-frequency interpretation of (the underspread component of) the time-varying Wiener filter (see **Fig. 3**). Let $\mathcal{R}_s$ and $\mathcal{R}_n$ denote the effective support regions of $\overline{W}_s^{(0)}(t,f)$ and $\overline{W}_n^{(0)}(t,f)$, respectively. Regarding the action of the time-varying Wiener filter, the following three time-frequency regions can be distinguished:

- *Pass region.* In the "signal only" time-frequency region $\mathcal{R}_s \backslash \mathcal{R}_n$, i.e., in the time-frequency region where only signal energy is present, there is $L^{(0)}_{\mathbf{H}_W^u}(t,f) \approx 1$. Thus, $\mathbf{H}_W^u$ passes all "noise-free" observation components without attenuation or distortion.

- *Stop region.* In the "noise only" time-frequency region $\mathcal{R}_n \backslash \mathcal{R}_s$ where only noise energy is present, there is $L^{(0)}_{\mathbf{H}_W^u}(t,f) \approx 0$, i.e., $\mathbf{H}_W^u$ suppresses all observation components in time-frequency regions where there is no signal.

- *Transition region.* In the "signal plus noise" time-frequency region $\mathcal{R}_s \cap \mathcal{R}_n$ where both signal energy and noise energy are present, $L^{(0)}_{\mathbf{H}_W^u}(t,f)$ assumes values approximately between 0 and 1. Here, $\mathbf{H}_W^u$ performs a time-frequency varying attenuation that depends on the relative signal and noise energy at the respective time-frequency point. In particular, for equal signal and noise energy, i.e., time-frequency points (t,f) with $\overline{W}_s^{(0)}(t,f) = \overline{W}_n^{(0)}(t,f)$, there is $L^{(0)}_{\mathbf{H}_W^u}(t,f) \approx 1/2$.

[3]The processes $s(t)$ and $n(t)$ are said to be *jointly* underspread if their expected generalized ambiguity functions, $\bar{A}_s^{(\alpha)}(\tau,\nu)$ and $\bar{A}_n^{(\alpha)}(\tau,\nu)$, are concentrated within the *same* region about the origin of the (τ,ν)-plane. For example, a quasistationary process and a quasiwhite process may be individually underspread but not jointly underspread.



(a) (b)

Figure 3. *Time-frequency interpretation of the time-varying Wiener filter $\mathbf{H}_W$ for jointly underspread signal and noise processes: (a) Effective time-frequency support regions of signal and noise, (b) time-frequency pass, stop, and transition regions of the time-varying Wiener filter.*

B. Time-Frequency Design of Time-Varying Wiener Filters

The time-frequency formulation (12) suggests a simple *time-frequency design* of nonstationary signal estimators. Let us define the "time-frequency pseudo-Wiener filter" $\widetilde{\mathbf{H}}_W$ by setting its generalized Weyl symbol equal to the right-hand side of (12) [47,48]:

$$L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}(t,f) \;\triangleq\; \frac{\overline{W}_s^{(\alpha)}(t,f)}{\overline{W}_s^{(\alpha)}(t,f) + \overline{W}_n^{(\alpha)}(t,f)}\,. \tag{13}$$

For jointly underspread processes $s(t)$, $n(t)$ where (12) is a good approximation, combination of (12) and (13) yields $L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}(t,f) \approx L^{(\alpha)}_{\mathbf{H}_W^u}(t,f)$, and hence the time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ is a close approximation to (the underspread part of) the optimal Wiener filter $\mathbf{H}_W$; furthermore, $\widetilde{\mathbf{H}}_W$ will then be nearly independent of the value of α used in (13). For processes that are not jointly underspread, however, $\widetilde{\mathbf{H}}_W$ must be expected to perform poorly.

While the time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ is *designed* in the time-frequency domain, the actual calculation of the signal estimate can be performed directly in the time domain according to

$$\hat{s}(t) \;=\; \big(\widetilde{\mathbf{H}}_W\, x\big)(t) \;=\; \int_{t'} \tilde{h}_W(t,t')\, x(t')\, dt',$$

where $\tilde{h}_W(t,t')$, the impulse response of $\widetilde{\mathbf{H}}_W$, can be obtained from $L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}(t,f)$ as [15–17]

$$\tilde{h}_W(t,t') = \int_f L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}\bigg(\Big(\frac{1}{2}+\alpha\Big)t + \Big(\frac{1}{2}-\alpha\Big)t', f\bigg)\, e^{j2\pi f(t-t')}\, df\,. \tag{14}$$

An efficient implementation of the time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ that is based on the multiwindow short-time Fourier transform is discussed in [48,49].

Compared to the Wiener filter $\mathbf{H}_W$, the time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ possesses two practical advantages:

• *Modified a priori knowledge.* The calculation (design) of $\mathbf{H}_W$ requires knowledge of the correlation operators $\mathbf{R}_s$ and $\mathbf{R}_n$ (cf. (10)), whereas the design of $\widetilde{\mathbf{H}}_W$ requires knowledge of the generalized Wigner-Ville spectra $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$ (cf. (13)). Although correlation operators and generalized Wigner-Ville spectra are mathematically equivalent due to the one-to-one mapping (3), the generalized Wigner-Ville spectra are much easier and more intuitive to handle than the correlation operators or correlation functions. For example, an approximate or partial knowledge of the Wigner-Ville spectra will often suffice for a reasonable filter design. This fact is especially important for practical applications where the *a priori* knowledge has to be estimated from the available data [49].

• *Reduced computation.* The calculation (design) of $\mathbf{H}_W$ requires a computationally intensive and potentially unstable operator inversion (or, in a discrete-time setting, a matrix inversion). In the time-frequency design (13), this operator inversion is replaced by simple and easily controllable pointwise divisions of functions. Assuming discrete-time signals of length N, the computational cost of the design of $\mathbf{H}_W$ grows with N^3, whereas that of $\widetilde{\mathbf{H}}_W$ (using divisions and FFTs) grows only with $N^2 \log_2 N$.

C. Robust Variations

If the actual correlations $\mathbf{R}_s$ and $\mathbf{R}_n$ deviate from the nominal correlations for which the Wiener filter $\mathbf{H}_W$ was designed, the filter's performance may degrade significantly. This sensitivity of the performance of the Wiener filter (and also of the time-frequency pseudo-Wiener filter) to variations of the second-order statistics motivates the use of *minimax robust Wiener filters* that optimize the worst-case performance within specified uncertainty classes of operating conditions.

Recently, the minimax robust time-*in*variant Wiener filters introduced in [50–53] for stationary processes were extended to the nonstationary case, and time-frequency designs of robust time-varying Wiener filters were proposed [54,55]. Particularly simple and intuitively appealing results were obtained for so-called *p-point uncertainty classes*. Let $\{\mathcal{R}_i\}_{i=1,2,\ldots,K}$ be a partition of the time-frequency plane, i.e., $\bigcup_{i=1}^K \mathcal{R}_i = \mathbb{R}^2$ and $\mathcal{R}_i \cap \mathcal{R}_j = \emptyset$ for $i \neq j$. Extending the stationary case definition in [51,53], p-point uncertainty classes can be defined for Wigner-Ville spectra as [54]

$$\mathcal{S} = \left\{ \overline{W}_s^{(0)}(t,f) : \iint_{\mathcal{R}_i} \overline{W}_s^{(0)}(t,f)\, dt\, df = s_i\,, \;\; i = 1,...,K \right\}$$

$$\mathcal{N} = \left\{ \overline{W}_n^{(0)}(t,f) : \iint_{\mathcal{R}_i} \overline{W}_n^{(0)}(t,f)\, dt\, df = n_i\,, \;\; i = 1,...,K \right\},$$

i.e., as the sets which contain all Wigner-Ville spectra $\overline{W}_s^{(0)}(t,f)$ and $\overline{W}_n^{(0)}(t,f)$ that have prescribed energies $s_i \geq 0$ and $n_i \geq 0$ in prescribed time-frequency regions $\mathcal{R}_i$.

The *minimax robust time-frequency Wiener filter* $\widetilde{\mathbf{H}}_R$ is now defined as the linear, time-varying system whose Weyl symbol minimizes a time-frequency expression of the mean-square error for the worst-case choice of $\overline{W}_s^{(0)}(t,f) \in \mathcal{S}$ and $\overline{W}_n^{(0)}(t,f) \in \mathcal{N}$ [55]. The Weyl symbol of this robust time-frequency Wiener filter is obtained as [54,55]

$$L_{\widetilde{\mathbf{H}}_R}^{(0)}(t,f) \,=\, \sum_{i=1}^{K} \frac{s_i}{s_i + n_i}\, I_{\mathcal{R}_i}(t,f)\,,$$

where $I_{\mathcal{R}_i}(t,f)$ is the indicator function of $\mathcal{R}_i$ (i.e., $I_{\mathcal{R}_i}(t,f) = 1$ for $(t,f) \in \mathcal{R}_i$ and 0 otherwise). Note that $L_{\widetilde{\mathbf{H}}_R}^{(0)}(t,f)$ is piecewise constant, expressing constant time-frequency weighting in a given time-frequency region $\mathcal{R}_i$. It can be shown [54,55] that the performance of the robust time-frequency Wiener filter $\widetilde{\mathbf{H}}_R$ is approximately independent of the actual operating conditions as long as they are within $\mathcal{S}, \mathcal{N}$. An intuitive and computationally efficient approximate time-frequency implementation of $\widetilde{\mathbf{H}}_R$ in terms of the multi-window Gabor transform [56,57] exists if the partition $\{\mathcal{R}_i\}$ corresponds to a uniform rectangular tiling of the time-frequency plane [55].

D. Simulation Results

Fig. 4 shows the Wigner-Ville spectra and expected ambiguity functions (with $\alpha = 0$) of signal and noise processes $s(t)$ and $n(t)$ as well as the Weyl symbols of the resulting Wiener filter $\mathbf{H}_W$, its underspread part $\mathbf{H}_W^u$, and the time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$. From the expected ambiguity functions in parts (c) and (d), it is seen that the processes $s(t)$ and $n(t)$ are jointly underspread. From the Weyl symbols in parts (e)–(g), the time-frequency pass, stop, and transition regions (cf. Fig. 3) of the filters $\mathbf{H}_W$, $\mathbf{H}_W^u$, and $\widetilde{\mathbf{H}}_W$ are easily recognized. It is verified that the Weyl symbol of $\widetilde{\mathbf{H}}_W$ closely approximates that of $\mathbf{H}_W^u$. The mean SNR improvement achieved was 6.14 dB for the Wiener filter $\mathbf{H}_W$, 6.10 dB for its underspread part $\mathbf{H}_W^u$, and 6.11 dB for the time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$. Hence, the performance of the time-frequency pseudo-Wiener filter is seen to be very close to that of the Wiener filter.

To illustrate the application and performance of the robust time-frequency Wiener filter described in Subsection IV-C, we defined p-point uncertainty classes $\mathcal{S}$ and $\mathcal{N}$ based on $K = 4$ rectangular time-frequency regions $\mathcal{R}_i$. The regional energies s_i and n_i used for the definition of $\mathcal{S}$ and $\mathcal{N}$ were derived from the nominal Wigner-Ville spectra $\overline{W}_s^{(0)}(t,f)$ and $\overline{W}_n^{(0)}(t,f)$ shown in Fig. 4(a), (b) according to $s_i = \iint_{\mathcal{R}_i} \overline{W}_s^{(0)}(t,f)\, dt\, df$ and $n_i = \iint_{\mathcal{R}_i} \overline{W}_n^{(0)}(t,f)\, dt\, df$. The Weyl symbol of the robust time-frequency Wiener filter $\widetilde{\mathbf{H}}_R$ obtained for these uncertainty classes is shown in Fig. 4(h). The rectangular time-frequency regions $\mathcal{R}_i$ underlying the uncertainty classes are clearly visible. **Fig. 5** compares the performance (output SNR vs. input SNR) of the ordinary Wiener filter $\mathbf{H}_W$ designed for the nominal Wigner-Ville spectra, the robust time-frequency Wiener filter $\widetilde{\mathbf{H}}_R$, and a trivial filter $\mathbf{H}_T$ that suppresses (passes) all signals in the case of negative (positive) input SNR. It is seen that at nominal operating conditions $\mathbf{H}_W$ performs only slightly better than $\widetilde{\mathbf{H}}_R$ but at its worst-case operating conditions $\mathbf{H}_W$ performs much worse than $\widetilde{\mathbf{H}}_R$ or even $\mathbf{H}_T$. Hence, in this example, the robust time-frequency Wiener filter $\widetilde{\mathbf{H}}_R$ achieves a drastic performance improvement over $\mathbf{H}_W$ at worst-case operating conditions with only a slight performance loss at nominal operating conditions.

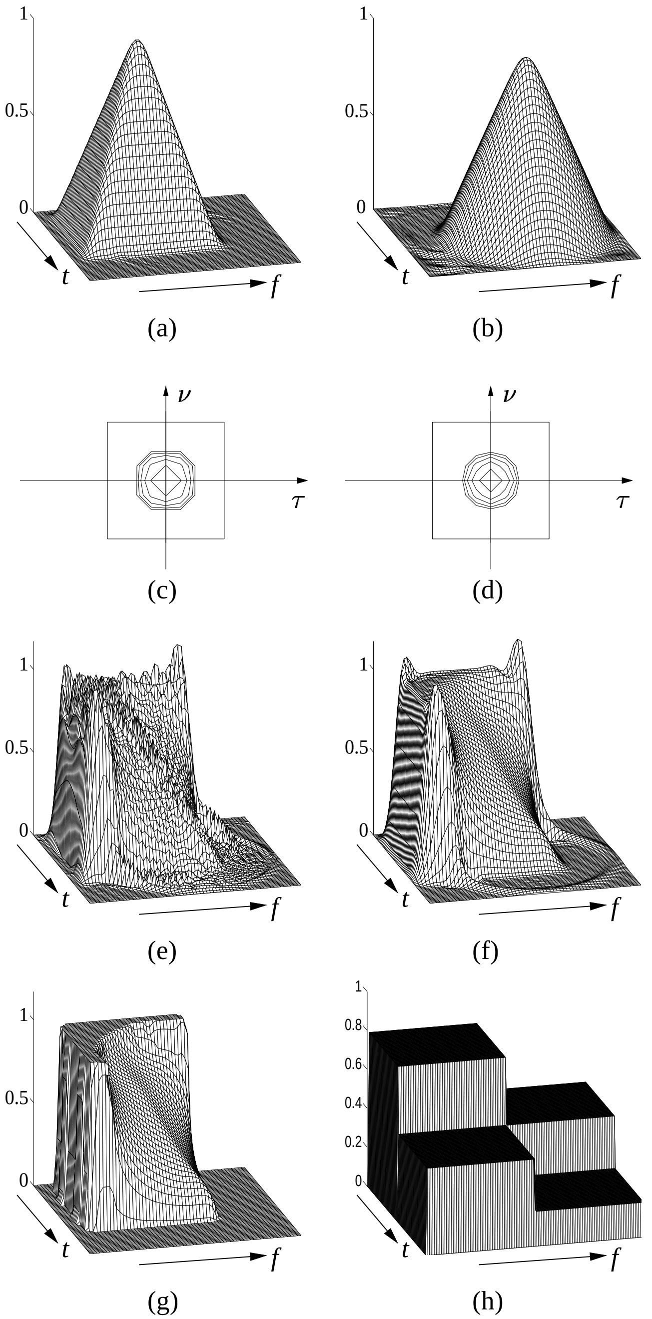


Figure 4. *Time-frequency representations of signal and noise statistics and of various Wiener-type filters: (a) Wigner-Ville spectrum of $s(t)$, (b) Wigner-Ville spectrum of $n(t)$, (c) magnitude of expected ambiguity function of $s(t)$, (d) magnitude of expected ambiguity function of $n(t)$, (e) Weyl symbol of Wiener filter $\mathbf{H}_W$, (f) Weyl symbol of underspread part $\mathbf{H}_W^u$ of $\mathbf{H}_W$, (g) Weyl symbol of time-frequency pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$, (h) Weyl symbol of robust time-frequency Wiener filter $\widetilde{\mathbf{H}}_R$. The rectangles in parts (c) and (d) have area 1 and thus permit to assess the underspread property of $s(t)$ and $n(t)$. The processes $s(t)$ and $n(t)$ were generated using the time-frequency synthesis technique introduced in [36]. The signal length is 128 samples.*

V. Nonstationary Signal Detection

Next, we discuss time-frequency methods for nonstationary signal detection. We shall again use the generalized Wigner-Ville spectrum as time-dependent power spectrum, while keeping in mind that, in the underspread case considered, the generalized evolutionary spectrum can approximately be substituted in the relevant equations.

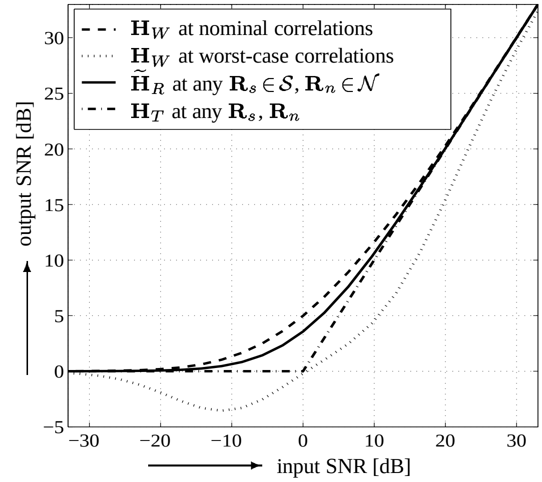


Figure 5. *Performance of the ordinary Wiener filter $\mathbf{H}_W$, the robust time-frequency Wiener filter $\widetilde{\mathbf{H}}_R$, and the trivial filter $\mathbf{H}_T$. The input SNR was varied by scaling $\mathbf{R}_s$.*

We consider the detection of a nonstationary, Gaussian random signal $s(t)$ from a noise-contaminated observation $x(t)$. The hypotheses are

$$H_0 \;:\; x(t) \;=\; n(t)$$
$$H_1 \;:\; x(t) \;=\; s(t) + n(t)\,,$$

where $n(t)$ is nonstationary, Gaussian noise uncorrelated with $s(t)$. The optimal detector (likelihood ratio detector) [42–44] forms a quadratic form of the observed signal $x(t)$,

$$\Lambda(x) \;=\; \langle \mathbf{H}_l\, x, x\rangle \;=\; \int_t \int_{t'} h_l(t,t')\, x(t')\, x^*(t)\, dt\, dt', \qquad (15)$$

with the operator $\mathbf{H}_l$ given by

$$\mathbf{H}_l \;=\; \mathbf{R}_n^{-1} - (\mathbf{R}_s + \mathbf{R}_n)^{-1} \;=\; \mathbf{R}_n^{-1}\mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}, \quad (16)$$

or $\mathbf{H}_l = \mathbf{R}_n^{-1}\mathbf{H}_W$, where $\mathbf{H}_W = \mathbf{R}_s(\mathbf{R}_s + \mathbf{R}_n)^{-1}$ is the time-varying Wiener filter considered in Section IV. The test statistic $\Lambda(x)$ is then compared to a threshold to decide whether H_0 or H_1 is in force. For stationary processes, $\Lambda(x)$ can be expressed in the frequency domain as

$$\Lambda(x) \;=\; \int_f \frac{S_s(f)}{S_n(f)\,[S_s(f) + S_n(f)]}\, |X(f)|^2\, df, \qquad (17)$$

where $X(f)$ is the Fourier transform of the observation $x(t)$ and $S_s(f)$ and $S_n(f)$ are the power spectral densities of $s(t)$ and $n(t)$, respectively. This frequency-domain expression involves simple products and reciprocals of functions (instead of products and inverses of operators as in (16)) and hence allows a simple interpretation and design of optimal detectors in the stationary case.

A. Time-Frequency Formulation of Nonstationary Detectors

It is known [58–60] that the quadratic test statistic in (15) can be rewritten as an inner product in the time-frequency domain,

$$\Lambda(x) \;=\; \langle L^{(\alpha)}_{\mathbf{H}_l}, W^{(\alpha)}_x \rangle \;=\; \int_t \int_f L^{(\alpha)}_{\mathbf{H}_l}(t,f)\, W^{(\alpha)*}_x(t,f)\, dt\, df. \tag{18}$$

Here,

$$W^{(\alpha)}_x(t,f) \triangleq \int_\tau x\left(t+\left(\frac{1}{2}-\alpha\right)\tau\right) x^*\left(t-\left(\frac{1}{2}+\alpha\right)\tau\right) e^{-j2\pi f\tau}\, d\tau \tag{19}$$

is the *generalized Wigner distribution* [10–12] of the observed signal $x(t)$. Thus, $\Lambda(x)$ can be interpreted as a weighted integral of $W^{(\alpha)*}_x(t,f)$, where the time-frequency weighting function is the generalized Weyl symbol of the operator $\mathbf{H}_l$.

In analogy to Subsection IV-A, a simplified approximate time-frequency formulation of $\Lambda(x)$ exists for *jointly underspread* processes $s(t)$ and $n(t)$. Here, the operator $\mathbf{H}_l$ can be decomposed into an overspread component whose effect is negligible and an underspread component, denoted as $\mathbf{H}^u_l$, whose generalized Weyl symbol can be approximated as [60]

$$L^{(\alpha)}_{\mathbf{H}^u_l}(t,f) \;\approx\; \frac{\overline{W}^{(\alpha)}_s(t,f)}{\overline{W}^{(\alpha)}_n(t,f)\left[\overline{W}^{(\alpha)}_s(t,f)+\overline{W}^{(\alpha)}_n(t,f)\right]}\,. \tag{20}$$

Inserting (20) in (18), we obtain the following approximate time-frequency formulation of our test statistic,

$$\Lambda(x) \;\approx\; \int_t \int_f \frac{\overline{W}^{(\alpha)}_s(t,f)\, W^{(\alpha)*}_x(t,f)}{\overline{W}^{(\alpha)}_n(t,f)\left[\overline{W}^{(\alpha)}_s(t,f)+\overline{W}^{(\alpha)}_n(t,f)\right]}\, dt\, df.$$

This extends the frequency-domain expression (17) to the nonstationary (underspread) case. For $\alpha = 0$ (note that $\overline{W}^{(0)}_{\cdot}(t,f)$ and $W^{(0)}_x(t,f)$ are real-valued), the above approximation allows a simple and intuitively appealing time-frequency interpretation that is analogous to the one given in Subsection IV-A in the context of the approximation (12). In essence, the test statistic $\Lambda(x)$ picks up energy of the observation $x(t)$ in time-frequency regions where there is large mean signal energy but little mean noise energy, and tends to reject observation components in time-frequency regions where there is little mean signal energy and large mean noise energy.

B. Time-Frequency Design of Nonstationary Detectors

The time-frequency formulation (20) suggests a simple *time-frequency design* of nonstationary detectors. In analogy to Subsection IV-B, we define the system $\widetilde{\mathbf{H}}_l$ by setting its generalized Weyl symbol equal to the right-hand side of (20) [60]:

$$L^{(\alpha)}_{\widetilde{\mathbf{H}}_l}(t,f) \;\triangleq\; \frac{\overline{W}^{(\alpha)}_s(t,f)}{\overline{W}^{(\alpha)}_n(t,f)\left[\overline{W}^{(\alpha)}_s(t,f)+\overline{W}^{(\alpha)}_n(t,f)\right]}\,. \tag{21}$$

Inserting in (18) yields the time-frequency designed test statistic

$$\tilde{\Lambda}(x) \;=\; \int_t \int_f \frac{\overline{W}^{(\alpha)}_s(t,f)\, W^{(\alpha)*}_x(t,f)}{\overline{W}^{(\alpha)}_n(t,f)\left[\overline{W}^{(\alpha)}_s(t,f)+\overline{W}^{(\alpha)}_n(t,f)\right]}\, dt\, df.$$

For jointly underspread processes $s(t)$, $n(t)$ where (20) is a good approximation, combination of (20) and (21) yields $L^{(\alpha)}_{\widetilde{\mathbf{H}}_l}(t,f) \approx L^{(\alpha)}_{\mathbf{H}^u_l}(t,f)$. Thus, $\widetilde{\mathbf{H}}_l$ is a close approximation to (the underspread part of) the optimal operator $\mathbf{H}_l$; furthermore, $\widetilde{\mathbf{H}}_l$ will then be nearly independent of the value of α used in (21). Hence, for jointly underspread processes the performance of the time-frequency designed detector $\tilde{\Lambda}(x)$ will be similar to that of the optimal detector $\Lambda(x)$ and approximately independent of α. For processes that are not jointly underspread, however, $\tilde{\Lambda}(x)$ must be expected to perform poorly.

While the detector $\tilde{\Lambda}(x)$ is designed in the time-frequency domain, it can be implemented directly in the time domain according to (cf. (15))

$$\tilde{\Lambda}(x) \;=\; \langle \widetilde{\mathbf{H}}_l x, x\rangle \;=\; \int_t \int_{t'} \tilde{h}_l(t,t')\, x(t')\, x^*(t)\, dt\, dt',$$

where $\tilde{h}_l(t,t')$, the impulse response of $\widetilde{\mathbf{H}}_l$, can be obtained from $L^{(\alpha)}_{\widetilde{\mathbf{H}}_l}(t,f)$ by an inverse Weyl transformation (cf. (14)). Efficient implementations of the time-frequency detector $\tilde{\Lambda}(x)$ that are based on the multiwindow short-time Fourier transform or the multiwindow spectrogram are discussed in [61].

Compared to the optimal detector $\Lambda(x)$, the time-frequency designed detector $\tilde{\Lambda}(x)$ is practically advantageous because the statistical *a priori* knowledge used in its design is formulated in the intuitively accessible time-frequency domain, and because its design is computationally less intensive and numerically more stable (since operator inversions are replaced by pointwise divisions of functions). These advantages are analogous to the advantages of the time-frequency pseudo-Wiener filter discussed in Subsection IV-B.

C. Simulation Results

Fig. 6 compares the performance of the optimal likelihood ratio detector $\Lambda(x)$ with that of the time-frequency designed detector $\tilde{\Lambda}(x)$ for jointly underspread signal and noise processes. It is seen that the time-frequency designed detector closely approximates the optimal detector.

In the previous example, the noise contained a strong white component and hence the operators $\mathbf{R}_n$ and $\mathbf{R}_s + \mathbf{R}_n$ (that have to be inverted for calculating $\mathbf{H}_l$ according to (16)) were nonsingular. In practice, this need not be the case. Our next example considers the application of the optimal detector $\Lambda(x)$ and the time-frequency designed detector $\tilde{\Lambda}(x)$ to the detection of knocking combustions in car engines (see [61] for background and details). The hypotheses are

$$H_0 \;:\; x(t) \;=\; x_0(t) \;=\; s_0(t) + n(t)$$
$$H_1 \;:\; x(t) \;=\; x_1(t) \;=\; s_1(t) + n(t)\,,$$

where $s_0(t)$ and $s_1(t)$ denote respectively the non-knocking and knocking signal and $n(t)$ is stationary white Gaussian noise. The format of these hypotheses is different from that of our previous hypotheses; however, the optimal detector $\Lambda(x)$ and

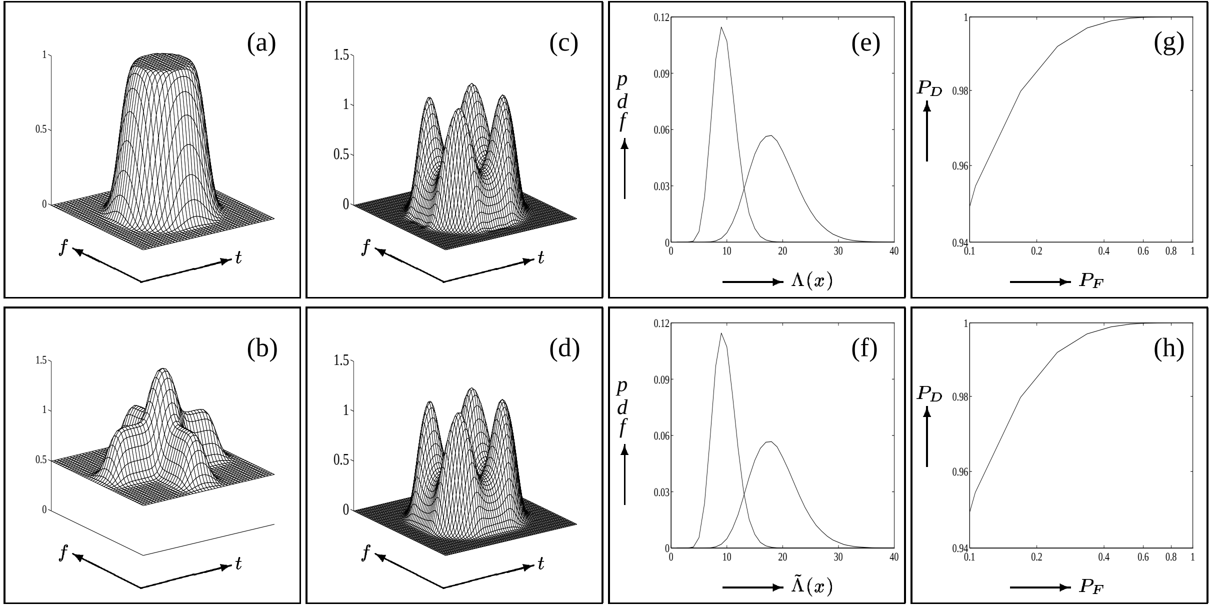


Figure 6. *Comparison of optimal detector $\Lambda(x)$ and time-frequency designed detector $\tilde{\Lambda}(x)$: (a) Wigner-Ville spectrum of $s(t)$, (b) Wigner-Ville spectrum of $n(t)$, (c) Weyl symbol of optimal operator $\mathbf{H}_l$, (d) Weyl symbol of time-frequency designed operator $\tilde{\mathbf{H}}_l$, (e) conditional probability density functions (pdf's) of optimal test statistic $\Lambda(x)$ under either hypothesis, (f) conditional pdf's of time-frequency designed test statistic $\tilde{\Lambda}(x)$, (g) receiver operator characteristics (ROC) [43,44] of optimal detector $\Lambda(x)$, and (h) ROC of time-frequency designed detector $\tilde{\Lambda}(x)$. The performance results in (e)–(h) were obtained by Monte Carlo simulation. The signal length is 128 samples.*

the time-frequency designed detector $\tilde{\Lambda}(x)$ can be extended to this more general type of hypotheses [61]. In this example, the second-order statistics (correlations or Wigner-Ville spectra) were estimated from a set of labeled training data, the optimal and time-frequency designed detectors were constructed using these estimated second-order statistics, and the performance of the detectors was evaluated by applying them to a different set of labeled data. The calculation of the optimal detector was made difficult by the poor conditioning of the estimated correlation operators of $x_0(t)$ and $x_1(t)$ (hence, pseudo-inverses were used for implementing the necessary inversions). In contrast, the design of $\tilde{\Lambda}(x)$ using the estimated Wigner-Ville spectra merely involves divisions of functions that are easily stabilized. **Fig. 7** shows the estimated Wigner-Ville spectra of $x_0(t)$ and $x_1(t)$ as well as the resulting time-frequency weighting functions $L^{(0)}_{\mathbf{H}_l}(t,f)$ for the optimal detector $\Lambda(x)$ and $L^{(0)}_{\tilde{\mathbf{H}}_l}(t,f)$ for the time-frequency designed detector $\tilde{\Lambda}(x)$ (see (18)). The receiver operating characteristics of the two detectors are compared in **Fig. 8**. It is seen that, due to its more stable design, the time-frequency designed detector performs significantly better than the theoretically optimal detector.

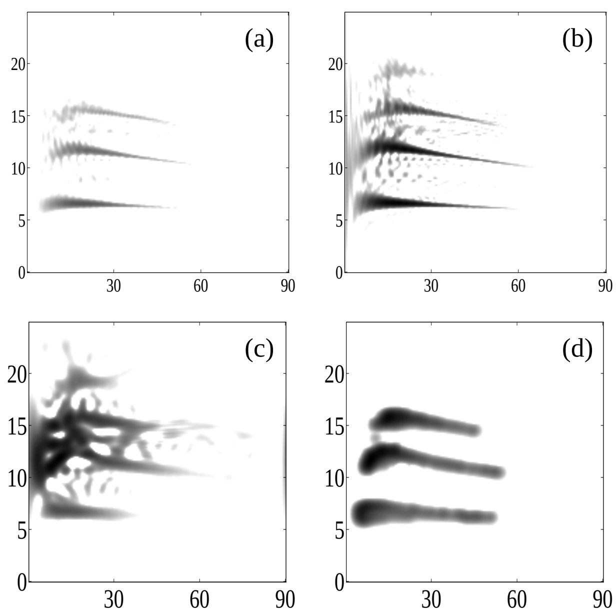


Figure 7. *Estimated Wigner-Ville spectra of the observed signal and time-frequency weighting functions of the two detectors: (a) Estimated Wigner-Ville spectrum of $x_0(t)$ calculated from non-knocking training data, (b) estimated Wigner-Ville spectrum of $x_1(t)$ calculated from knocking training data, (c) time-frequency weighting function of the optimal detector $\Lambda(x)$, (d) time-frequency weighting function of the time-frequency designed detector $\tilde{\Lambda}(x)$. Horizontal axis: crank angle (in degrees) which is proportional to time, vertical axis: frequency (in kHz). The signal length is 186 samples.*

VI. Conclusion

We have shown that the time-frequency domain allows an extension of the spectral representation of random processes and the frequency-domain formulation of statistical signal processing techniques to the nonstationary case. However, it is important to be aware that this extension is possible only if the processes involved are underspread. In this paper, we have emphasized techniques for signal estimation and signal detec-

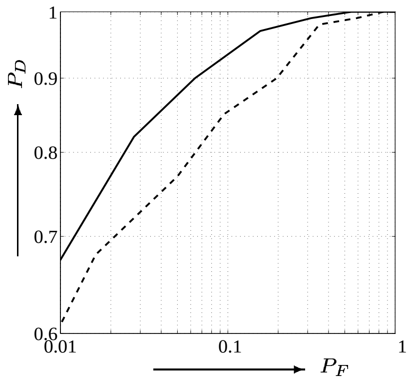


Figure 8. *ROCs of the optimal detector $\Lambda(x)$ (dashed line) and the time-frequency designed detector $\tilde{\Lambda}(x)$ (solid line).*

tion. An important aspect that has not been considered is the problem of estimating the various time-dependent spectra [4,5,8,27,28,62,63]. Furthermore, related results that have not been mentioned are time-frequency techniques for the detection of deterministic signals in nonstationary noise [60,64], for deflection-optimal detection [59,60], for subspace-based estimation and detection [48,61,65], and for minimax robust signal estimation based on uncertainty models other than the p-point model [55]. Beyond that, we emphasize that the general approach—formally replacing the power spectral density by a suitably defined time-dependent power spectrum—is applicable to other fields of statistical signal processing as well.

ACKNOWLEDGMENTS

We are grateful to S. Carstens-Behrens, M. Wagner, and J. F. Böhme for stimulating and illuminating discussions and for providing us with the car engine data used in our last simulation example.

REFERENCES

[1] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 3rd ed., 1991.

[2] A. Papoulis, *Signal Analysis*. Singapore: McGraw-Hill, 1984.

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

[4] P. Flandrin, "Time-dependent spectra for nonstationary stochastic processes," in *Time and Frequency Representation of Signals and Systems* (G. Longo and B. Picinbono, eds.), pp. 69–124, Wien: Springer, 1989.

[5] P. Flandrin and W. Martin, "The Wigner-Ville spectrum of nonstationary random signals," in *The Wigner Distribution — Theory and Applications in Signal Processing* (W. Mecklenbräuker and F. Hlawatsch, eds.), pp. 211–267, Amsterdam (The Netherlands): Elsevier, 1997.

[6] W. Kozek, F. Hlawatsch, H. Kirchauer, and U. Trautwein, "Correlative time-frequency analysis and classification of nonstationary random processes," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Philadelphia, PA), pp. 417–420, Oct. 1994.

[7] G. Matz and F. Hlawatsch, "Time-varying spectra for underspread and overspread nonstationary processes," in *Proc. 32nd Asilomar Conf. Signals, Systems, Computers*, (Pacific Grove, CA), Nov. 1998. invited paper.

[8] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 1461–1470, Dec. 1985.

[9] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Inf. Theory*, vol. 14, pp. 369–374, May 1968.

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

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

[12] F. Hlawatsch and P. Flandrin, "The interference structure of the Wigner distribution and related time-frequency signal representations," in *The Wigner Distribution — Theory and Applications in Signal Processing* (W. Mecklenbräuker and F. Hlawatsch, eds.), pp. 59–133, Amsterdam, The Netherlands: Elsevier, 1997.

[13] G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes," *IEEE Trans. Signal Processing*, vol. 45, pp. 1520–1534, June 1997.

[14] H. Cramér, "On some classes of nonstationary stochastic processes," in *Proc. 4th Berkeley Symp. on Math. Stat. and Prob.*, pp. 57–78, Univ. Calif. Press, 1961.

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

[16] W. Kozek, *Matched Weyl-Heisenberg Expansions of Nonstationary Environments*. PhD thesis, Vienna University of Technology, March 1997.

[17] G. Matz and F. Hlawatsch, "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, pp. 4041–4071, Aug. 1998.

[18] G. B. Folland, *Harmonic Analysis in Phase Space*, vol. 122 of *Annals of Mathematics Studies*. Princeton (NJ): Princeton University Press, 1989.

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

[20] W. Kozek, "Time-frequency signal processing based on the Wigner-Weyl framework," *Signal Processing*, vol. 29, pp. 77–92, Oct. 1992.

[21] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 318–331, Feb. 1994.

[22] L. A. Zadeh, "Frequency analysis of variable networks," *Proc. of IRE*, vol. 76, pp. 291–299, March 1950.

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

[24] J. J. Kohn and L. Nirenberg, "An algebra of pseudo-differential operators," *Comm. Pure Appl. Math.*, vol. 18, pp. 269–305, 1965.

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

[26] M. B. Priestley, *Spectral Analysis and Time Series – Part II*. London: Academic Press, 1981.

[27] K. Riedel, "Optimal data-based kernel estimation of evolutionary spectra," *IEEE Trans. Signal Processing*, vol. 41, pp. 2439–2447, July 1993.

[28] A. S. Kayhan, A. El-Jaroudi, and L. F. Chaparro, "Evolutionary periodogram for nonstationary signals," *IEEE Trans. Signal Processing*, vol. 42, pp. 1527–1536, June 1994.

[29] C. S. Detka and A. El-Jaroudi, "The transitory evolutionary spectrum," in *Proc. IEEE ICASSP-94*, (Adelaide, Australia), pp. 289–292, April 1994.

[30] W. Kozek, "On the transfer function calculus for underspread LTV channels," *IEEE Trans. Signal Processing*, vol. 45, pp. 219–223, Jan. 1997.

[31] W. Kozek, "Adaptation of Weyl-Heisenberg frames to underspread environments," in *Gabor Analysis and Algorithms: Theory and Applications* (H. G. Feichtinger and T. Strohmer, eds.), ch. 10, pp. 323–352, Boston (MA): Birkhäuser, 1998.

[32] S. G. Mallat, G. Papanicolaou, and Z. Zhang, "Adaptive covariance estimation of locally stationary processes," *Annals of Stat.*, vol. 26, Feb. 1998.

[33] S. G. Mallat, *A Wavelet Tour of Signal Processing*. San Diego: Academic Press, 1998.

[34] J. A. Sills and E. W. Kamen, "Wiener filtering of nonstationary signals based on spectral density functions," in *Proc. 34th IEEE Conf. Decision and Control*, (Kobe, Japan), pp. 2521–2526, Dec. 1995.

[35] J. A. Sills, *Nonstationary Signal Modeling, Filtering, and Parameterization*. PhD thesis, Georgia Institute of Technology, Atlanta, March 1995.

[36] F. Hlawatsch and W. Kozek, "Second-order time-frequency synthesis of nonstationary random processes," *IEEE Trans. Inf. Theory*, vol. 41, pp. 255–267, Jan. 1995.

[37] R. Thomä, J. Steffens, and U. Trautwein, "Statistical cross-terms in quadratic time-frequency distributions," in *Proc. Int. Conf. on DSP*, (Nicosia, Cyprus), July 1993.

[38] N. A. Abdrabbo and M. B. Priestley, "Filtering non-stationary signals," *J. Roy. Stat. Soc. Ser. B*, vol. 31, pp. 150–159, 1969.

[39] H. A. Khan and L. F. Chaparro, "Nonstationary Wiener filtering based on evolutionary spectral theory," in *Proc. IEEE ICASSP-97*, (Munich, Germany), pp. 3677–3680, May 1997.

[40] A. M. Sayeed, P. Lander, and D. L. Jones, "Improved time-frequency filtering of signal-averaged electrocardiograms," *J. of Electrocardiology*, vol. 28, pp. 53–58, 1995.

[41] P. Lander and E. J. Berbari, "Time-frequency plane Wiener filtering of the high-resolution ECG: Background and time-frequency representations," *IEEE Trans. Biomedical Engineering*, vol. 44, pp. 247–255, April 1997.

[42] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. New York: Wiley, 1968.

[43] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 1988.

[44] L. L. Scharf, *Statistical Signal Processing*. Reading (MA): Addison Wesley, 1991.

[45] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs (NJ): Prentice Hall, 1992.

[46] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. Cambridge (MA): MIT Press, 1949.

[47] H. Kirchauer, F. Hlawatsch, and W. Kozek, "Time-frequency formulation and design of nonstationary Wiener filters," in *Proc. IEEE ICASSP-95*, (Detroit, MI), pp. 1549–1552, May 1995.

[48] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation," *submitted to IEEE Trans. Signal Processing*, 1998.

[49] W. Kozek, H. G. Feichtinger, and J. Scharinger, "Matched multiwindow methods for the estimation and filtering of nonstationary processes," in *Proc. IEEE ISCAS-96*, (Atlanta, GA), pp. 509–512, May 1996.

[50] S. A. Kassam and T. L. Lim, "Robust Wiener filters," *J. Franklin Inst.*, vol. 304, pp. 171–185, Oct./Nov. 1977.

[51] L. J. Cimini and S. A. Kassam, "Robust and quantized Wiener filters for p-point spectral classes," in *Proc. Conf. Information Sciences and Systems*, (Princeton, NJ), pp. 314–318, March 1980.

[52] H. V. Poor, "On robust Wiener filtering," *IEEE Trans. Autom. Contr.*, vol. 25, pp. 531–536, June 1980.

[53] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, March 1985.

[54] G. Matz and F. Hlawatsch, "Robust time-varying Wiener filters: Theory and time-frequency formulation," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Pittsburgh, PA), pp. 401–404, Oct. 1998.

[55] G. Matz and F. Hlawatsch, "Minimax robust time-frequency filters for nonstationary signal estimation." Submitted to IEEE ICASSP-99, (Phoenix, AZ), Mar. 1999.

[56] M. Zibulski and Y. Y. Zeevi, "Analysis of multiwindow Gabor-type schemes by frame methods," *Applied and Computational Harmonic Analysis*, vol. 4, pp. 188–221, April 1997.

[57] Y. Y. Zeevi, M. Zibulski, and M. Porat, "Multi-window Gabor schemes in signal and image representations," in *Gabor Analysis and Algorithms: Theory and Applications* (H. G. Feichtinger and T. Strohmer, eds.), ch. 12, pp. 381–407, Boston (MA): Birkhäuser, 1998.

[58] P. Flandrin, "A time-frequency formulation of optimum detection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 9, pp. 1377–1384, 1988.

[59] A. M. Sayeed and D. L. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Trans. Signal Processing*, vol. 43, pp. 2872–2883, Dec. 1995.

[60] G. Matz and F. Hlawatsch, "Time-frequency formulation and design of optimal detectors," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Paris, France), pp. 213–216, June 1996.

[61] G. Matz and F. Hlawatsch, "Time-frequency methods for signal detection with application to the detection of knock in car engines," in *Proc. IEEE-SP Workshop on Statistical Signal and Array Proc.*, (Portland, OR), pp. 196–199, Sept. 1998.

[62] W. Kozek and K. Riedel, "Quadratic time-varying spectral estimation for underspread processes," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Philadelphia, PA), pp. 460–463, Oct. 1994.

[63] A. S. Kayhan, L. F. Chaparro, and A. El-Jaroudi, "Wold-Cramer evolutionary spectral estimators," in *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, (Victoria, Canada), pp. 115–118, Oct. 1992.

[64] B. V. K. Kumar and C. W. Carroll, "Performance of Wigner distribution function based detection methods," *Opt. Eng.*, vol. 23, pp. 732–737, Nov. 1984.

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