OPTIMAL DETECTION USING BILINEAR TIME-FREQUENCY AND TIME-SCALE REPRESENTATIONS

Akbar M. Sayeed, *Student Member, IEEE,* and Douglas L. Jones, *Member, IEEE*

Coordinated Science Laboratory
University of Illinois at Urbana–Champaign
1308 West Main Street
Urbana, IL 61801*

akbar@csl.uiuc.edu, tel: (217) 244-6384
jones@csl.uiuc.edu, tel: (217) 244-6823

Fax: (217) 244-1642

IEEE Transactions on Signal Processing, pp. 2872–2883, December 1995.

Abstract

Bilinear time-frequency representations (TFRs) and time-scale representations (TSRs) are potentially very useful for detecting a nonstationary signal in the presence of nonstationary noise or interference. As quadratic signal representations, they are promising for situations in which the optimal detector is a quadratic function of the observations. All existing time-frequency formulations of quadratic detection either implement classical optimal detectors equivalently in the time-frequency domain, without fully exploiting the structure of the TFR, or attempt to exploit the nonstationary structure of the signal in an *ad hoc* manner. We identify several important nonstationary composite hypothesis testing scenarios for which TFR/TSR-based detectors provide a "natural" framework; that is, in which TFR/TSR-based detectors are both optimal and exploit the many degrees of freedom available in the TFR/TSR. We also derive explicit expressions for the corresponding optimal TFR/TSR kernels. As practical examples, we show that the proposed TFR/TSR detectors are directly applicable to many important radar/sonar detection problems. Finally, we also derive optimal TFR/TSR-based detectors which exploit only partial information available about the nonstationary structure of the signal.

*This work was supported by the National Science Foundation under Grant No. MIP 90-12747, the Joint Services Electronics Program under Grant No. N00014-90-J-1270, and the Schlumberger Foundation.

1 Introduction

Detection and estimation of signals in the presence of noise is an old and important problem in communications and signal processing. Applications include radar and sonar, and, more recently, automatic fault detection in machinery [1, 2, 3, 4, 5, 6]. Time-frequency representations (TFRs) and time-scale representations (TSRs) have been extensively used for detection in these and other applications, primarily due to the need for dealing with nonstationary signals. Drawing on certain properties of high resolution TFRs, a class of quadratic detectors has also recently been proposed for a variety of nonstationary scenarios [7, 8, 9, 10, 11]. Although such TFR-based detectors try to exploit the nonstationary structure of the signals, they are not necessarily optimal from a detection theoretic viewpoint.

The other major class of time-frequency-based detection schemes is based on Moyal's formula [12] and implements the classical optimal detectors equivalently in the time-frequency domain. For example, time-frequency formulations of optimal linear detectors implement the matched filter via the cross-ambiguity function. Similarly, a time-frequency formulation of Gaussian signal detection in Gaussian noise implements the optimal quadratic detector T by correlating the Wigner distribution W_x of the observed signal x with a two-dimensional time-frequency function S based on the statistics [13]:

$$T(x) = \int\int W_x(t,f)S(t,f)dtdf \ . \tag{1}$$

A similar formulation in terms of the Altes distribution [14] is given in [15], and detection schemes based on the spectrogram are considered in [16]. In [17] and [15, 18], maximum likelihood detectors for detecting noncoherent linear and hyperbolic chirps, respectively, with unknown chirp parameters, are implemented equivalently by integrating certain TFRs along certain curves in the time-frequency plane. However, such time-frequency detectors are merely equivalent realizations of the classical optimal detectors, which do not exploit the information available in the TFR at different time-frequency locations, and are usually more efficiently implemented in other domains. For example, $T(x)$ in (1) is in effect a quadratic form which can be implemented more efficiently (without having to compute W_x) as

$$T(x) = \int\int x^*(t_1)L(t_1,t_2)x(t_2)dt_2dt_1 \ , \quad \text{where} \tag{2}$$

$$L(t_1,t_2) = \int S\left(\frac{t_1+t_2}{2}, f\right) e^{j2\pi f(t_1-t_2)} df \ .$$

Such equivalent time-frequency-based optimal detectors may yield some useful interpretive insights but that may not necessarily justify their time-frequency-based realization.

Thus, in studying the role of TFRs/TSRs as detectors, the main question is to identify detection problems for which time-frequency-based detection is "natural"; that is, the detectors are optimal, the structure of TFRs/TSRs is exploited, and the resultant TFR/TSR-based formulations of the optimal detectors are computationally efficient as well. Since current TFRs/TSRs are either linear or quadratic, clearly TFR/TSR-based detectors can be optimal only in situations in which linear or quadratic detectors are optimal.[1] We focus on bilinear (as opposed to linear) TFRs/TSRs because they can implement a richer class of detectors as evidenced by many important detection scenarios in which the optimal detector is inherently a quadratic function of the observations [19, 20, 21, 22].

[1] If a single TFR/TSR is to be used, this restricts TFR/TSR-based detectors to situations in which the optimal detectors are either purely linear or purely quadratic functions of the observations. General linear-quadratic detectors can be implemented with a combination of linear and bilinear TFRs/TSRs.

A key characteristic of TFRs/TSRs, in the context of classical quadratic detection, is that they can potentially realize a different quadratic function of the observations (a different detector) at each time-frequency/time-scale location. This property of TFRs/TSRs is in fact the fundamental motivation for the use of the narrowband (wideband) cross-ambiguity function for detecting a deterministic signal with unknown delay-doppler(-scale) parameters in white Gaussian noise. In this paper, we identify several important quadratic detection scenarios involving nonstationary signals in which

- the many "degrees of freedom" in a TFR or a TSR can be exploited, and
- TFRs or TSRs, by virtue of their special structure, offer a natural quadratic detection framework instead of being merely inefficient equivalent implementations of the classical optimum detectors.

More specifically, we identify certain composite hypothesis testing [19] situations involving Gaussian signals for which TFRs/TSRs provide a natural, unified detection framework. In particular, we explicitly characterize the TFR/TSR-based detectors for such generalizations of the following well-known quadratic detection problems [19]:

- Detecting a Gaussian signal in white Gaussian noise.
- Detecting a Gaussian signal in arbitrary Gaussian noise (deflection optimal).
- Detecting a Gaussian signal with unknown amplitude (low signal-to-noise ratio assumption) in arbitrary Gaussian noise (locally optimal).

In addition to characterizing such TFR/TSR detection frameworks, we show that they apply to many important radar/sonar detection problems. Moreover, the deflection-based detectors are applicable even if the signal is nonGaussian. In situations in which probabilistic information about the parameters of the composite hypothesis is available, the optimal detector has a complicated nonlinear form. In such cases we derive detectors, based on the maximum *a posteriori* probability estimates of the parameters, which are amenable to TFR/TSR implementation and perform better than the maximum likelihood detectors employed in the case of unknown parameters. Finally, we also show that in certain cases, only partial information about the nonstationary structure of the signal can be exploited to design optimal detectors.

2 Classical Quadratic Detection Scenarios

In this section, we review some classical quadratic detection problems whose composite-hypothesis generalizations will be considered in the next section. Consider the following binary hypotheses in continuous time

$$\begin{aligned} H_0 : x(t) &= n(t) \\ H_1 : x(t) &= as(t) + n(t) \end{aligned} \tag{3}$$

where $t \in T$, the time interval of observation, x is the observed signal, n is a zero-mean complex Gaussian noise with correlation function $R_n(t_1, t_2) = E[n(t_1)n^*(t_2)]$, s is a zero-mean complex Gaussian signal with correlation function $R_s(t_1, t_2)$, and a is a positive parameter. In statistical hypothesis testing, for each observation, x, a real-valued test statistic, $L(x)$, is compared to a threshold to decide whether H_0 or H_1 is true; that is, whether the signal is present or not. For detecting a Gaussian signal in Gaussian noise, the optimal test statistic is a quadratic function of the observations [19, 21]. We assume that both the signal and the noise process are independent of each other and are completely characterized by their correlation functions; that is, $E[n(t_1)n(t_2)] = E[s(t_1)s(t_2)] = 0$ for all $t_1, t_2 \in T$.[2] Although we consider only binary hypotheses, most of the results can be readily extended to multiple hypotheses. We consider three well-known cases, and in all cases we assume that the likelihood ratio (LR) exists [24, 19], which, in particular, implies that detection with zero probability of error is not possible [24].[3]

Case I. $a = 1$, and n is white with its real and imaginary parts independent and of equal power spectral densities so that $R_n(t_1, t_2) = N_0\delta(t_1 - t_2)$, where $\delta(t)$ is the Dirac delta function. The optimum test statistic is given by [19]

$$L_{\rm O}(x) = \frac{1}{N_0}\langle(\mathbf{R}_s(\mathbf{R}_s + N_0\mathbf{I})^{-1}x, x\rangle \ , \tag{4}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product defined as $\langle x, y\rangle = \int x(t)y^*(t)dt$, $\mathbf{I}$ is the identity operator, and $\mathbf{R}_s$ denotes the linear operator defined by the corresponding correlation function R_s as

$$(\mathbf{R}_s x)(t) = \int R_s(t,\tau)x(\tau)d\tau \ . \tag{5}$$

Case II. $a = 1$, and n is arbitrary zero-mean Gaussian noise with correlation function R_n. In this case, the optimum test statistic based on the LR is given by [19]

$$L_{\rm LR}(x) = \langle\mathbf{R}_n^{-1}(\mathbf{R}_s\mathbf{R}_n^{-1} + \mathbf{I})^{-1}\mathbf{R}_s\mathbf{R}_n^{-1}x, x\rangle - \log(\det(\mathbf{R}_s\mathbf{R}_n^{-1} + \mathbf{I})) \ , \tag{6}$$

where $\det(\cdot)$ denotes the determinant of an operator (product of the eigenvalues).

However, for reasons that will be explained in Section 3, we do not use the LR as the test statistic in this case. Instead, we use the deflection criterion for quadratic test statistics, which is a useful alternative optimality criterion [21, 19]. The optimal quadratic test statistic is one which maximizes the deflection

$$H(L) = \frac{(E_1[L] - E_0[L])^2}{\mathrm{Var}_0(L)} \ , \tag{7}$$

where E_i denotes the expectation given that the i-th hypothesis is true, and $\mathrm{Var}_0(L)$ denotes the variance of L under H_0. In this case, the deflection-optimal test statistic is given by [25, 19] [4]

$$L_{\rm H}(x) = \langle\mathbf{R}_n^{-1}\mathbf{R}_s\mathbf{R}_n^{-1}x, x\rangle = \langle\mathbf{R}_s\mathbf{R}_n^{-1}x, \mathbf{R}_n^{-1}x\rangle \ . \tag{8}$$

Case III. Same as Case II except that a is an unknown positive parameter which is assumed to be small enough so that the signal-to-noise ratio (*SNR*) is low. In this case we resort to a locally optimal test to get rid of the unknown parameter a, and the test statistic is given by [19]

$$L_{\rm LO}(x) = \langle\mathbf{R}_s\mathbf{R}_n^{-1}x, \mathbf{R}_n^{-1}x\rangle - \mathrm{Trace}(\mathbf{R}_n^{-1}\mathbf{R}_s) = L_{\rm H}(x) - \mathrm{Trace}(\mathbf{R}_n^{-1}\mathbf{R}_s) \ , \tag{9}$$

where $\mathrm{Trace}(\cdot)$ denotes the trace of an operator (sum of the eigenvalues). Note that the quadratic form in x is identical in both (8) and (9).[5]

---

[2]Such Gaussian processes exist [23] and are sometimes referred to as "circular" Gaussian. Note that for a real process, $E[s(t_1)s(t_2)] = [s(t_1)s^*(t_2)] = R_s(t_1, t_2)$; thus, we don't need this assumption for real processes.

[3]In anticipation of the composite hypotheses introduced in the next section, we include those constant terms in the optimum test statistics which will be affected by the composite hypothesis parameters.

[4]Deflection-optimal detectors can be interpreted as "maximum *SNR*" detectors because deflection is a measure of *SNR* [21, 19].

[5]Note that the effect of the operator $\mathbf{R}_n^{-1}$ in (8) and (9) can be interpreted in terms of a time-varying, "pre-whitening" filter.

3 Composite Hypotheses

As discussed in the Introduction, a key observation in the context of TFR/TSR-based detection is that a bilinear TFR/TSR can realize a different quadratic detector at each time-frequency/time-scale location. In this section, in order to exploit such degrees of freedom in a TFR/TSR detector, we consider composite-hypothesis generalizations of the detection scenarios identified in the previous section. Using fundamental properties of TFRs and TSRs, we characterize the composite hypothesis testing situations, within the proposed framework, for which TFR/TSR-based detectors provide a natural structure. Not surprisingly, the parameters of the composite hypothesis correspond to time-frequency or time-scale shifts in such detection situations.

3.1 Composite hypothesis testing

Instead of (3) we consider the problem in which H_1 is a composite hypothesis; that is,

$$\begin{aligned} H_0 : x(t) &= n(t) \\ H_1 : x(t) &= as(t;\alpha,\beta) + n(t)\ ,\ (\alpha,\beta) \in \mathbb{R}^2 \end{aligned} \tag{10}$$

where $t \in T$, and T, s, n and a are as defined in the previous section. The parameters (α,β) parameterize the composite hypothesis H_1. In the time-frequency context, α and β may represent unknown time and frequency shifts, such as might arise, for example, in a doppler radar scenario. Because of our assumptions, the dependence of s on (α,β) is only through the correlation function, which we denote by $R_s^{(\alpha,\beta)}$. The parameters (α,β) may be deterministic but unknown or random with a known joint probability density function (pdf) $p(\alpha,\beta)$. Optimal detectors for such composite hypothesis testing situations are described next.

Unknown parameters. If the parameters (α,β) are modeled as deterministic but unknown, we use a generalized likelihood ratio test (GLRT) [19] corresponding to the maximum likelihood (ML) estimate of (α,β).[6] The optimal ML detectors in the three cases are given by

$$L_{ML}(x) = \begin{cases} \max_{(\alpha,\beta)} L_{\mathrm{O}}^{(\alpha,\beta)}(x) & \text{Case I} \\ \max_{(\alpha,\beta)} L_{\mathrm{H}}^{(\alpha,\beta)}(x) & \text{Case II (deflection optimal)} \\ \max_{(\alpha,\beta)} L_{\mathrm{LO}}^{(\alpha,\beta)}(x) & \text{Case III (locally optimal)} \end{cases} \quad , \tag{11}$$

where the superscript (α,β) denotes the test statistic corresponding to the signal correlation function $R_s^{(\alpha,\beta)}$. We note that in Case II we use the "maximum deflection" detector, instead of the ML detector, which can also be interpreted as a "maximum *SNR*" detector.

Random parameters. For situations in which the parameters (α,β) are random with joint pdf $p(\alpha,\beta)$, the optimal test statistic is given by

$$T(x) = \int\int LR^{(\alpha,\beta)}(x)p(\alpha,\beta)d\alpha d\beta\ , \tag{12}$$

where $LR^{(\alpha,\beta)}$ is the LR for the binary hypotheses (10), for a given value of the parameters (α,β).[7] However, in general, T in (12) is difficult to compute analytically. Thus, in this case we resort to suboptimal detectors

[6] In a GLRT, an estimate of the parameters (α,β) is first formed, and that estimate is then used in the likelihood ratio for the two hypotheses.

[7] Note that T is the expected value of $LR^{(\alpha,\beta)}$ with respect to the random parameters (α,β).

based on a GLRT (or analogues of it) in which, instead of using ML estimates of (α,β), we exploit the available statistical information about (α,β) and use the maximum *a posteriori* probability (MAP) estimates of (α,β). Thus, we propose the following "MAP GLRT detectors" for the three cases:

$$L_{\mathrm{MAP}}(x) = \begin{cases} \max_{(\alpha,\beta)} \left[L_{\mathrm{O}}^{(\alpha,\beta)}(x) + \log p(\alpha,\beta) \right] - \log p(\hat{\alpha},\hat{\beta}) & \text{Case I} \\ \max_{(\alpha,\beta)} \left[\log\left(L_{\mathrm{H}}^{(\alpha,\beta)}(x) \right) + \log p(\alpha,\beta) \right] - \log p(\hat{\alpha},\hat{\beta}) & \text{Case II (deflection optimal)} \\ \max_{(\alpha,\beta)} \left[\log\left(L_{\mathrm{LO}}^{(\alpha,\beta)}(x) \right) + \log p(\alpha,\beta) \right] - \log p(\hat{\alpha},\hat{\beta}) & \text{Case III (locally optimal)} \end{cases} , \quad (13)$$

$$(\hat{\alpha},\hat{\beta}) = \begin{cases} \arg\max_{(\alpha,\beta)} \left[L_{\mathrm{O}}^{(\alpha,\beta)}(x) + \log p(\alpha,\beta) \right] & \text{Case I} \\ \arg\max_{(\alpha,\beta)} \left[\log\left(L_{\mathrm{H}}^{(\alpha,\beta)}(x) \right) + \log p(\alpha,\beta) \right] & \text{Case II} \\ \arg\max_{(\alpha,\beta)} \left[\log\left(L_{\mathrm{LO}}^{(\alpha,\beta)}(x) \right) + \log p(\alpha,\beta) \right] & \text{Case III} \end{cases} . \quad (14)$$

We note that the proposed form of the "deflection optimal" detector in Case II is based on its similarity to the locally optimal detector as evident from (8) and (9).

3.2 Parameter dependence for TFR/TSR-based detectors

In the previous subsection, we described the optimal quadratic detectors in a number of composite hypothesis testing scenarios involving a parameterized Gaussian signal. From (11) and (13) and the expressions for test statistics in (4), (8) and (9), we note that for a given value of the parameters, (α,β), the component of the test statistics that is a function of the observations is in all cases of the form

$$L^{(\alpha,\beta)}(x) = \langle \mathbf{Q}^{(\alpha,\beta)} x, x \rangle \quad (15)$$

where $\mathbf{Q}^{(\alpha,\beta)}$ is a positive definite operator. The motivation for introducing composite hypotheses was to characterize the situations in which the composite hypothesis parameters, (α,β), naturally correspond to time-frequency or time-scale. We now use fundamental properties of bilinear TFRs and TSRs to characterize the appropriate dependence of the Gaussian signal on the parameters (α,β). First, we review the relevant characterizations of TFRs and TSRs.

Any bilinear TFR from Cohen's class can be expressed as [12]

$$P_x(t,f;\Phi) = \int\int W_x(u,v)\Phi(u-t,v-f)\,du\,dv \ , \quad (16)$$

where W_x is the Wigner distribution (WD) of x, defined as [26, 12]

$$W_x(t,f) = \int x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi f\tau}d\tau \ , \quad (t,f) \in \mathbb{R}^2 \ , \quad (17)$$

and the two-dimensional kernel Φ completely characterizes the TFR $P_x(t,f;\Phi)$. Similarly, any bilinear TSR from the affine class can be expressed as [27]

$$C_x(t,a;\Pi) = \int\int W_x(u,v)\Pi((u-t)/a, av)\,du\,dv \ , \quad (t,a) \in \mathbb{R} \times (0,\infty) \ , \quad (18)$$

where again the kernel Π completely characterizes the TSR $C_x(t,a;\Pi)$. We note that both $P_x(\Phi)$ and $C_x(\Pi)$ are characterized as averaged versions of the WD, the difference being in the nature of the averaging.

Using Weyl correspondence [28, 29, 30], the test statistic in (15) can be expressed as[8]

$$L^{(\alpha,\beta)}(x) = \langle \mathbf{Q}^{(\alpha,\beta)} x, x \rangle = \int\int W_x(u,v) \mathrm{WS}_{Q^{(\alpha,\beta)}}(u,v) du dv \tag{19}$$

where $\mathrm{WS}_{Q^{(\alpha,\beta)}}$ is the Weyl symbol [28, 29, 30] of the operator $\mathbf{Q}^{(\alpha,\beta)}$ (or its kernel $Q^{(\alpha,\beta)}$) defined as

$$\mathrm{WS}_{Q^{(\alpha,\beta)}}(u,v) = \int Q^{(\alpha,\beta)}(u+\tau/2, u-\tau/2) e^{-j2\pi v\tau} d\tau \ . \tag{20}$$

Comparing (19) and (16), we note that by using a time-frequency-varying kernel, any test statistic $L^{(\alpha,\beta)}$, corresponding to any arbitrary dependence of $\mathbf{Q}^{(\alpha,\beta)}$ on (α,β), can be implemented via Cohen's class as

$$L^{(\alpha,\beta)}(x) = P_x(\alpha,\beta;\Phi^{(\alpha,\beta)}) \ , \quad (\alpha,\beta) \in \mathbb{R}^2 \ , \tag{21}$$

by identifying the parameters (α,β) with (t,f), where the kernel $\Phi^{(\alpha,\beta)}$ is given by

$$\Phi^{(\alpha,\beta)}(u,v) = \mathrm{WS}_{Q^{(\alpha,\beta)}}(u+\alpha, v+\beta) \ . \tag{22}$$

Clearly, with such an arbitrary dependence of the kernel on the parameters, the TFR-based realization of $L^{(\alpha,\beta)}$ in (21) is no different from an arbitrary unstructured array of quadratic detectors. Moreover, with such an arbitrarily time-frequency-varying kernel, the resulting TFR is not necessarily a valid (in terms of time-frequency localization properties) TFR.

Thus, for a natural and efficient TFR-based realization of the test statistic $L^{(\alpha,\beta)}$, the kernel should not vary with the (α,β) parameters. In Appendix A, it is shown that $\Phi^{(\alpha,\beta)}$ defined in (22) is independent of (α,β) if and only if $Q^{(\alpha,\beta)}$ has the following dependence on the parameters $(\alpha,\beta) \in \mathbb{R}^2$:

$$Q^{(\alpha,\beta)}(t_1,t_2) = Q^{(0,0)}(t_1-\alpha, t_2-\alpha) e^{j2\pi\beta(t_1-t_2)} \ , \quad (t_1,t_2) \in \mathbb{R}^2 \ . \tag{23}$$

That is, the correlation function $Q^{(\alpha,\beta)}$ is a time- and frequency-shifted version of a fixed correlation function $Q^{(0,0)}$. Thus, for a natural TFR-based detection framework, the composite hypothesis parameters (α,β) must correspond to time- and frequency-shifts, respectively.

Similarly, it can be shown that for a natural TSR-based realization (using a fixed kernel Π) of the test statistics $L^{(\alpha,\beta)}$, a necessary and sufficient condition on the dependence of $Q^{(\alpha,\beta)}$ on $(\alpha,\beta) \in \mathbb{R} \times (0,\infty)$ is

$$Q^{(\alpha,\beta)}(t_1,t_2) = \beta Q^{(0,1)}(\beta(t_1-\alpha), \beta(t_2-\alpha)) \ , \quad (t_1,t_2) \in \mathbb{R}^2 \ . \tag{24}$$

That is, for TSR-based detectors, the parameters (α,β) must correspond to time-shifts and scalings, respectively.

3.3 Signal models for TFR/TSR-based detectors

It can be shown that the condition (23) on the operator $\mathbf{Q}^{(\alpha,\beta)}$, in conjunction with the expressions for the detectors in the three cases, translates into the following Gaussian signal model in (10) which characterizes the composite hypothesis detection situations for which TFR-based detectors provide a natural framework.

Case A. $(\alpha,\beta) = (\tau,\nu) \in T \times \mathbb{R}$ and

$$R_s^{(\tau,\nu)}(t_1,t_2) = R_{TF}(t_1-\tau, t_2-\tau) e^{j2\pi\nu t_1} e^{-j2\pi\nu t_2} \ , \quad (t_1,t_2) \in T \times T \ , \tag{25}$$

[8] We note that Weyl correspondence can be defined in terms of any unitary [31] TFR. For our purposes, any such TFR which is covariant to scale changes as well will suffice. We choose WD simply out of convenience; from a detection point of view, the particular choice of such a representation is immaterial.

for some correlation function R_{TF} whose effective support is small compared with $T \times T$.[9] The subscript TF reflects that Case A corresponds to TFR implementation. Note that (25) is equivalent to $s(t;\tau,\nu) = s_{(\tau,\nu)}(t-\tau)e^{j2\pi\nu t}$ in (10), where, for each (τ,ν), $s_{(\tau,\nu)}$ is any Gaussian signal with correlation function R_{TF}; that is, for each (τ,ν), $s(t;\tau,\nu)$ is a time-frequency shifted version of some Gaussian signal with correlation function R_{TF}.

Similarly, using (24), we deduce that TSR-based detectors provide a natural framework for composite hypothesis testing situations characterized by the following signal model in (10).

Case B. $(\alpha,\beta) = (\tau,c) \in T \times (0,\infty)$ and

$$R_s^{(\tau,c)}(t_1,t_2) = cR_{TS}(c(t_1-\tau),c(t_2-\tau)) \ , \quad (t_1,t_2) \in T \times T \ , \tag{26}$$

for some correlation function R_{TS} (with small effective support compared with $T \times T$), where the subscript TS is in anticipation of TSR implementation of Case B. Again, (26) is equivalent to $s(t;\tau,c) = \sqrt{c}s_{(\tau,c)}(c(t-\tau))$ for any Gaussian signal $s_{(\tau,c)}$ with correlation function R_{TS}. That is, for each (τ,c), $s(t;\tau,c)$ is a time-shifted and scaled version of some Gaussian signal with correlation R_{TS}.

We note that the conditions (23) and (24) when applied in Case II to the test statistic (6), based on the LR, do not yield the simple signal models above. In other words, for the signal models in Cases A and B above, the test statistic $L_{\mathrm{LR}}^{(\alpha,\beta)}$ cannot be implemented using TFRs/TSRs with fixed kernels. This is the reason for using the deflection-optimal test statistic $L_{\mathrm{H}}^{(\alpha,\beta)}$ in Case II, which is amenable to efficient TFR/TSR-based realization.

Thus, we have characterized composite-hypothesis generalizations of detecting a Gaussian signal in Gaussian noise for which bilinear TFRs/TSRs provide a natural, unified detection framework. In particular, we have focused on generalizations of the three Gaussian detection scenarios described in Section 2. We will see in Section 5 that the signal models in Cases A and B can be identified with narrowband and wideband radar, respectively, by using appropriate scatterer models.

4 TFR/TSR-Based Realization of the Detectors

In this section, we derive explicit expressions for the kernels of TFR/TSR-based detectors for the various composite hypothesis testing situations characterized in the previous section. As we will see, such "natural" TFR/TSR-based formulations not only exploit the information at different time-frequency/time-scale locations, but are also *computationally efficient* implementations of the optimal detectors. We also express the optimal TFR/TSR detectors in terms of the eigenexpansion of the signal correlation function because such descriptions yield useful insight into the structure of the TFR/TSR detectors which is exploited in Section 6 to derive detectors based on partial signal information. We start by characterizing the eigenfunctions of the families (25) and (26).

4.1 Eigenfunctions of $R_s^{(\tau,\nu)}$ and $R_s^{(\tau,c)}$

First consider the family of correlation functions, $\{R_s^{(\tau,\nu)}\}$, defined in (25) in terms of the correlation function R_{TF}. We assume that R_{TF} is continuous on $T \times T$ and has finite support which is small compared to the

[9] The real underlying assumption is that the Weyl symbol of R_{TF}, $\mathrm{WS}_{R_{TF}}(t,f)$, has effectively finite support, with the temporal support small compared to the observation interval T. Note that this is not a restrictive assumption in practice if we are dealing with transient, nonstationary signals. The same assumption applies to R_{TS} in Case B.

length of the observation interval;[10] that is,

$$R_{TF}(t_1,t_2) = 0 \quad \text{for } |t_1| > T_1 \text{ or } |t_2| > T_1 \ , \tag{27}$$

for some T_1 such that $2T_1 \ll |T|$, where $|T|$ denotes the length of the observation interval T. By Mercer's theorem [32], R_{TF} admits the eigenexpansion

$$R_{TF}(t_1,t_2) = \sum_k \lambda_k u_k(t_1) u_k^*(t_2) \ , \tag{28}$$

where the u_k's are the eigenfunctions of R_{TF} and the λ_k's are the corresponding eigenvalues. The eigenfunctions form an orthonormal set; that is, they satisfy $\langle u_k, u_l\rangle = \delta_{kl}$, where δ_{kl} is the Kronecker delta function. From (25) we see that the support of $R_s^{(\tau,\nu)}$ is confined to $[-T_1+\tau, T_1+\tau] \times [-T_1+\tau, T_1+\tau]$ and $R_s^{(\tau,\nu)}$ admits the expansion

$$R_s^{(\tau,\nu)}(t_1,t_2) = \sum_k \lambda_k u_k^{(\tau,\nu)}(t_1) u_k^{(\tau,\nu)*}(t_2) \ , \quad \text{where} \tag{29}$$

$$u_k^{(\tau,\nu)}(t) = e^{j2\pi\nu t} u_k(t-\tau) \ . \tag{30}$$

It can be easily verified that (29) is indeed the eigenexpansion of $R_s^{(\tau,\nu)}$ with the $u_k^{(\tau,\nu)}$'s as the eigenfunctions and the λ_k's as the corresponding eigenvalues.

Now consider the family, $\{R_s^{(\tau,c)}\}$, defined in (26) in terms of R_{TS}. Again, we assume that the support of R_{TS} is confined to $[-T_1,T_1]\times[-T_1,T_1]$ where $2T_1 \ll |T|$. Then, the support of $R_s^{(\tau,c)}$ is contained in $[-T_1/c+\tau, T_1/c+\tau]\times[-T_1/c+\tau, T_1/c+\tau]$, and if the eigenexpansion of R_{TS} is

$$R_{TS}(t_1,t_2) = \sum_k \mu_k v_k(t_1) v_k^*(t_2) \ , \tag{31}$$

then $R_s^{(\tau,c)}$ admits the eigenexpansion

$$R_s^{(\tau,c)}(t_1,t_2) = \sum_k \mu_k v_k^{(\tau,c)}(t_1) v^{(\tau,c)*}(t_2) \ , \quad \text{where} \tag{32}$$

$$v_k^{(\tau,c)}(t) = \sqrt{c}\, v_k(c(t-\tau)) \ . \tag{33}$$

4.2 TFR/TSR detectors

From (11) and (13) we note that in order to characterize the TFR/TSR detectors in *all* the situations that we have considered, we only need to represent the test statistics $L_{\mathrm{O}}^{(\alpha,\beta)}$, $L_{\mathrm{H}}^{(\alpha,\beta)}$ and $L_{\mathrm{LO}}^{(\alpha,\beta)}$ in terms of TFRs for the signal correlation family (25) in Case A, and in terms of TSRs for the family (26) in Case B. Furthermore, all the test statistics have the quadratic form (15).

Weyl correspondence (or, equivalently, Moyal's formula [12, 31]) coupled with the covariance properties of the Weyl symbol (or, equivalently, the WD) provides the crucial link between TFRs/TSRs and quadratic detectors.[11] Weyl correspondence equates the quadratic form of an operator $\mathbf{L}$ with the inner product between the WD of the signal and and Weyl symbol of the operator as in (19) and repeated here for convenience:

$$\langle \mathbf{L}x, x\rangle = \int\!\!\int \mathrm{WS}_L(t,f) W_x(t,f) \ , \tag{34}$$

[10] See footnote 9.

[11] The restriction of Weyl correspondence to rank-1 operators, $\mathbf{L}x = \langle x, y\rangle y$, yields Moyal's formula [12, 31], and using the eigenexpansion of operators we can derive (34) from Moyal's formula. We have used Weyl correspondence because it yields a more streamlined presentation.

where WS_L is the Weyl symbol of the operator $\mathbf{L}$ as defined in (20). The relevant covariance properties of the Weyl symbol [28, 33] are captured by the relations

$$\mathrm{WS}_{R_s^{(\tau,\nu)}}(t,f) = \mathrm{WS}_{R_{TF}}(t-\tau, f-\nu) \ , \tag{35}$$

$$\mathrm{WS}_{R_s^{(\tau,c)}}(t,f) = \mathrm{WS}_{R_{TS}}(c(t-\tau), f/c) \ , \tag{36}$$

for the correlation families (25) and (26), which are generalizations of the familiar covariance relations for the WD [31]. We note that the Weyl symbol of the correlation function of a process is identical to the Wigner-Ville spectrum [34, 35] of the process.

For the composite hypothesis testing situations characterized by the family $\{R_s^{(\tau,\nu)}\}$ defined in (25), we can equivalently express the test statistics (8) and (9) using (34) and (35) as

$$\begin{aligned} L_{\mathrm{H}}^{(\tau,\nu)}(x) &= \langle \mathbf{R}_s^{(\tau,\nu)}\mathbf{R}_n^{-1}x, \mathbf{R}_n^{-1}x\rangle = \int\int \mathrm{WS}_{R_{TF}}(t-\tau, f-\nu) W_{\mathbf{R}_n^{-1}x}(t,f)dtdf \\ &= \sum_k \lambda_k |\langle \mathbf{R}_n^{-1}x, u_k^{(\tau,\nu)}\rangle|^2 \ , \quad \text{and} \end{aligned} \tag{37}$$

$$\begin{aligned} L_{\mathrm{LO}}^{(\tau,\nu)}(x) &= L_{\mathrm{H}}^{(\tau,\nu)}(x) - \mathrm{Trace}(\mathbf{R}_n^{-1}\mathbf{R}_s^{(\tau,\nu)}) = L_{\mathrm{H}}^{(\tau,\nu)}(x) - \int\int \mathrm{WS}_{R_{TF}}(t-\tau, f-\nu)\mathrm{WS}_{R_n^{-1}}(t,f)dtdf \\ &= L_{\mathrm{H}}^{(\tau,\nu)}(x) - \sum_k \lambda_k \langle u_k^{(\tau,\nu)}, \mathbf{R}_n^{-1}u_k^{(\tau,\nu)}\rangle \ , \end{aligned} \tag{38}$$

where in (38) we have used the fact that

$$\mathrm{Trace}(\mathbf{L}_1\mathbf{L}_2) = \int\int \mathrm{WS}_{L_1}(t,f)\mathrm{WS}_{L_2}(t,f)dtdf \tag{39}$$

which follows from (34) using the eigenexpansion of $\mathbf{L}_2$. Similarly, the test statistic (4) can be expressed as

$$L_{\mathrm{O}}^{(\tau,\nu)}(x) = \frac{1}{N_0}\sum_k \frac{\lambda_k}{\lambda_k + N_0}|\langle x, u_k^{(\tau,\nu)}\rangle|^2 = \int\int \mathrm{WS}_{\hat{R}_{TF}}(t-\tau, f-\nu) W_x(t,f)dtdf \ , \tag{40}$$

where $\mathrm{WS}_{\hat{R}_{TF}}$ is the Weyl symbol of the correlation function $\hat{R}_{TF}$ given by

$$\hat{R}_{TF}(t_1,t_2) = \frac{1}{N_0}\sum_k \frac{\lambda_k}{\lambda_k + N_0} u_k(t_1)u_k^*(t_2) \ . \tag{41}$$

By comparing the forms of (37) and (40) with (16), the TFR implementation of the test statistics is immediate; for example, $L_{\mathrm{O}}^{(\tau,\nu)}$ in (40) corresponds to a TFR with kernel $\Phi = \mathrm{WS}_{\hat{R}_{TF}}$. The following proposition summarizes the results by explicitly characterizing the TFR-based detectors.

Proposition A. *In the composite hypothesis testing problem (10), if the dependence of the Gaussian signal on the parameters $(\alpha,\beta) = (\tau,\nu)$ is characterized by the family of correlation functions, $\{R_s^{(\tau,\nu)}\}$, defined in (25), then the test statistics for both the ML and MAP GLRT detectors identified in Section 3 can be implemented using bilinear TFRs as*

$$L_A(x) = \begin{cases} \max_{(\tau,\nu)}\left[P_y(\tau,\nu;\Phi)\right] & \textit{ML detectors; Cases I and II} \\ \max_{(\tau,\nu)}\left[P_y(\tau,\nu;\Phi) - F_A(\tau,\nu)\right] & \textit{ML detectors; Case III} \\ \max_{(\tau,\nu)}\left[P_y(\tau,\nu;\Phi) + \log p(\tau,\nu)\right] - \log p(\hat{\tau},\hat{\nu}) & \textit{MAP detectors; Case I} \\ \max_{(\tau,\nu)}\left[\log\{P_y(\tau,\nu;\Phi)\} + \log p(\tau,\nu)\right] - \log p(\hat{\tau},\hat{\nu}) & \textit{MAP detectors; Case II} \\ \max_{(\tau,\nu)}\left[\log\{P_y(\tau,\nu;\Phi) - F_A(\tau,\nu)\} + \log p(\tau,\nu)\right] - \log p(\hat{\tau},\hat{\nu}) & \textit{MAP detectors; Cases III} \end{cases} \tag{42}$$

where

$$y = \begin{cases} x & \text{Case I} \\ \mathbf{R}_n^{-1}x & \text{Cases II and III} \end{cases} , \tag{43}$$

$$F_A(\tau,\nu) = P_{R_n^{-1}}(\tau,\nu;\Phi = \mathrm{WS}_{R_{TF}}) = \iint \mathrm{WS}_{R_n^{-1}}(t,f)\mathrm{WS}_{R_{TF}}(t-\tau,f-\nu)dtdf , \tag{44}$$

$$(\hat{\tau},\hat{\nu}) = \begin{cases} \arg\max_{(\tau,\nu)}\left[P_y(\tau,\nu;\Phi) + \log p(\tau,\nu)\right] & \textit{MAP detectors; Case I} \\ \arg\max_{(\tau,\nu)}\left[\log\left\{P_y(\tau,\nu;\Phi)\right\} + \log p(\tau,\nu)\right] & \textit{MAP detectors; Case II} \\ \arg\max_{(\tau,\nu)}\left[\log\left\{P_y(\tau,\nu;\Phi) - F_A(\tau,\nu)\right\} + \log p(\tau,\nu)\right] & \textit{MAP detectors; Case III} \end{cases} , \tag{45}$$

and the kernel Φ *characterizing the TFR* $P_y(\Phi)$ *can be expressed as*

$$\Phi(t,f) = \begin{cases} \mathrm{WS}_{\hat{R}_{TF}}(t,f) & \text{Case I} \\ \mathrm{WS}_{R_{TF}}(t,f) & \text{Cases II and III} \end{cases} , \tag{46}$$

where

$$\mathrm{WS}_{\hat{R}_{TF}}(t,f) = \frac{1}{N_0}\sum_k \frac{\lambda_k}{\lambda_k + N_0} W_{u_k}(t,f) , \tag{47}$$

and

$$\mathrm{WS}_{R_{TF}}(t,f) = \int R_{TF}(t+\tau/2, t-\tau/2)e^{-j2\pi f\tau}d\tau = \sum_k \lambda_k W_{u_k}(t,f) . \tag{48}$$

Similarly, in the other composite hypothesis testing situations characterized by the family $\{R_s^{(\tau,c)}\}$ defined in (26), by using (34) and (36), we can equivalently realize the test statistics $L_{\mathrm{O}}^{(\tau,c)}$, $L_{\mathrm{H}}^{(\tau,c)}$ and $L_{\mathrm{LO}}^{(\tau,c)}$ via TSRs, and the structure of such detectors is summarized in the next proposition.

Proposition B. *In the composite hypothesis testing problem (10), if the dependence of the Gaussian signal on the parameters* $(\alpha,\beta) = (\tau,c)$ *is characterized by the family of correlation functions,* $\{R_s^{(\tau,c)}\}$*, defined in (26), then the test statistics for both the ML and MAP GLRT detectors identified in Sectionr 3 can be implemented using bilinear TSRs as*

$$L_B(x) = \begin{cases} \max_{(\tau,c)}\left[C_y(\tau,1/c;\Pi)\right] & \textit{ML detectors; Cases I and II} \\ \max_{(\tau,c)}\left[C_y(\tau,1/c;\Pi) - F_B(\tau,c)\right] & \textit{ML detectors; Case III} \\ \max_{(\tau,c)}\left[C_y(\tau,1/c;\Pi) + \log p(\tau,c)\right] - \log p(\hat{\tau},\hat{c}) & \textit{MAP detect ors; Case I} \\ \max_{(\tau,c)}\left[\log\left\{C_y(\tau,1/c;\Pi)\right\} + \log p(\tau,c)\right] - \log p(\hat{\tau},\hat{c}) & \textit{MAP detectors; Case II} \\ \max_{(\tau,c)}\left[\log\left\{C_y(\tau,1/c;\Pi) - F_B(\tau,c)\right\} + \log p(\tau,c)\right] - \log p(\hat{\tau},\hat{c}) & \textit{MAP detectors; Case III} \end{cases} \tag{49}$$

where

$$y = \begin{cases} x & \text{Case I} \\ \mathbf{R}_n^{-1}x & \text{Cases II and III} \end{cases} , \tag{50}$$

$$F_B(\tau,c) = C_{R_n^{-1}}(\tau,1/c;\Pi = \mathrm{WS}_{R_{TS}}) = \iint \mathrm{WS}_{R_n^{-1}}(t,f)\mathrm{WS}_{R_{TS}}((t-\tau)c, f/c)dtdf , \tag{51}$$

$$(\hat{\tau},\hat{c}) = \begin{cases} \arg\max_{(\tau,c)}\left[C_y(\tau,c;\Pi) + \log p(\tau,c)\right] & \textit{MAP detectors; Case I} \\ \arg\max_{(\tau,\nu)}\left[\log\left\{C_y(\tau,c;\Pi)\right\} + \log p(\tau,c)\right] & \textit{MAP detectors; Case II} \\ \arg\max_{(\tau,\nu)}\left[\log\left\{C_y(\tau,c;\Pi) - F_B(\tau,c)\right\} + \log p(\tau,c)\right] & \textit{MAP detectors; Case III} \end{cases} , \tag{52}$$

and the kernel Π *characterizing the TSR* $C_y(\Pi)$ *can be expressed as*

$$\Pi(t,f) = \begin{cases} \mathrm{WS}_{\hat{R}_{TS}}(t,f) & \text{Case I} \\ \mathrm{WS}_{R_{TS}}(t,f) & \text{Cases II and III} \end{cases} , \tag{53}$$

where

$$\mathrm{WS}_{\hat{R}_{TS}}(t,f) = \frac{1}{N_0} \sum_k \frac{\mu_k}{\mu_k + N_0} W_{v_k}(t,f) , \tag{54}$$

and

$$\mathrm{WS}_{R_{TS}}(t,f) = \int R_{TS}(t+\tau/2, t-\tau/2) e^{-j2\pi f \tau} d\tau = \sum_k \mu_k W_{v_k}(t,f) . \tag{55}$$

Remark. We note in passing that in the low *SNR* case with random parameters, in addition to the MAP GLRT detector, we can also implement the locally optimal detector based on the actual LR. Given the family $\{R_s^{(\alpha,\beta)}\}$ and the joint pdf $p(\alpha,\beta)$, the locally optimum test statistic based on the LR is given by

$$L_{\mathrm{LO,LR}}(x) = \int\int \langle \mathbf{R}_s^{(\alpha,\beta)} \mathbf{R}_n^{-1} x, \mathbf{R}_n^{-1} x \rangle p(\alpha,\beta) d\alpha d\beta = \int\int L_{\mathrm{LO}}^{(\alpha,\beta)}(x) p(\alpha,\beta) d\alpha d\beta , \tag{56}$$

It can be easily verified, using the previous discussion, that the detectors corresponding to the families $\{R_s^{(\tau,\nu)}\}$ and $\{R_s^{(\tau,c)}\}$ admit the following TFR/TSR implementation:

$$L_{\mathrm{LO,LR}}^{A}(x) = \int\int p(\tau,\nu) P_{\mathbf{R}_n^{-1}x}(\tau,\nu;\Phi = \mathrm{WS}_{R_{TF}}) d\tau d\nu \quad \text{and} \tag{57}$$

$$L_{\mathrm{LO,LR}}^{B}(x) = \int\int p(\tau,c) C_{\mathbf{R}_n^{-1}x}(\tau,1/c;\Pi = \mathrm{WS}_{R_{TS}}) d\tau dc . \tag{58}$$

This is a generalization of locally optimal detection of a deterministic signal with random amplitude, time- and frequency-shifts as discussed in [16, 13].

4.3 Discussion

Propositions A and B, in conjunction with the discussion in Section 3, state the main results of the paper. In Section 3, we characterized certain composite hypothesis testing situations, corresponding to detecting a parameterized Gaussian signal in Gaussian noise, for which TFRs/TSRs provide a natural, unified detection framework. Propositions A and B explicitly characterize the corresponding TFR and TSR detectors in the three detection cases on which we have focused. Essentially, TFR-based detectors are characteristic of situations in which the Gaussian signal to be detected has unknown or random time-frequency shifts. On the other hand, the TSR-based detection framework applies to situations in which the signal has unknown or random time-shifts and scalings. The information contained in the TFR/TSR at different time-frequency/time-scale locations is amply exploited because at each point $(t,f) = (\tau,\nu)$ or $(t,a) = (\tau,1/c)$, the optimum detector corresponding to the value of the parameters (τ,ν) or (τ,c) is realized by the TFR or TSR, respectively. Moreover, as evident from (42) and (49), the location (τ,ν) or (τ,c) at which the optimal TFR/TSR-based detector is realized is exactly the ML or MAP estimate of the corresponding signal parameters. Furthermore, owing to the fact that the TFR/TSR-based detectors use a *fixed* kernel, the TFR/TSR-based realization of the optimal detectors is computationally efficient as well; any efficient algorithm for TFR/TSR implementation with a given kernel can be used. The computational efficiency of TFR/TSR-based detectors can alternatively be exploited by implementing them as a bank of spectrograms/scalograms, as shown in Section 6.

We finally make a few comments about estimation of necessary statistics if they are not available *a priori*. Evidently, the knowledge of any one of the pairs (R_n, R_s), (R_n, R_{n+s}) or (R_s, R_{n+s}) is sufficient to characterize all the detectors. Moreover, in order to characterize the complete family $\{R_s^{(\tau,\nu)}\}$ or $\{R_s^{(\tau,c)}\}$, we only need to estimate one of the pairs for only a *single* value of the parameters (τ, ν) or (τ, c). The reason is that once we know, say, $R_s^{(\tau,\nu)}$ for a particular value (τ_o, ν_o), we can determine R_{TF} from (25) which characterizes the entire family $\{R_s^{(\tau,\nu)}\}$.

5 Signal Correlation for Narrowband and Wideband Radar

In Section 3 we derived two families of signal correlation functions, $\{R_s^{(\tau,\nu)}\}$ and $\{R_s^{(\tau,c)}\}$, which facilitated efficient TFR/TSR implementation of the corresponding optimal detectors. In this section we show that many radar/sonar detection problems can be naturally modeled with such signal families. In particular, the family $\{R_s^{(\tau,\nu)}\}$ corresponds to the narrowband signal model and $\{R_s^{(\tau,c)}\}$ corresponds to the wideband model.

5.1 Narrowband model

Let the transmitted radar pulse be given by (the real part of)

$$p(t) = p_o(t) e^{j2\pi f_c t} \ , \tag{59}$$

where p_o is the baseband pulse which is time-limited to $[0, T_p]$ and f_c is the carrier frequency. According to the narrowband signal model, the received pulse from a point scatterer is given by

$$r^{(\tau,\nu)}(t) = a p(t-\tau) e^{j2\pi\nu t} \ , \tag{60}$$

where a is some complex constant, τ is the delay, and ν is the doppler shift. An arbitrary scatterer can be modeled as a continuum of scatterers, and thus the return is given by

$$r^{(\tau_o,\nu_o)}(t) = \int_{S_T}\int_{S_F} p(t-\tau_o-\tau) e^{j2\pi(\nu+\nu_o)t} \rho_{NB}(\tau,\nu) d\tau d\nu \tag{61}$$

where ρ_{NB} is the (complex) reflectivity of the scatterer, $S_T = [-\Delta_T, +\Delta_T]$ and $S_F = [-\Delta_F, +\Delta_F]$ for some $\Delta_T > 0$ and $\Delta_F > 0$. τ_o and ν_o are the *nominal* range and doppler of the scatterer, and Δ_T and Δ_F denote the range- and doppler-spread of the scatterer. For example, in the case of an aircraft, τ_o and ν_o may reflect the range and doppler of the center of the fuselage. The model (61) for an arbitrary scatterer is based on the observation that the reflectivities of the same scatterer, corresponding to the nominal values (τ_o, ν_o) and $(\tau_o + \Delta_\tau, \nu + \Delta_\nu)$ of range-doppler, must be related by $\rho_{NB}^{(\tau_o+\Delta_\tau,\nu_o+\Delta_\nu)}(\tau,\nu) = \rho_{NB}^{(\tau_o,\nu_o)}(\tau-\Delta_\tau, \nu-\Delta_\nu) = \rho_{NB}(\tau-\Delta_\tau-\tau_o, \nu-\Delta_\nu-\nu_o)$.[12]

In many situations, the reflectivity ρ_{NB} may be best modeled as being random. For example, the reflectivities corresponding to the many different orientations of an aircraft may be best modeled as different realizations of a random process. Thus, we assume that $\{\rho_{NB}(\tau,\nu) : \tau \in S_T, \nu \in S_F\}$ is a zero-mean, second-order random process. Moreover, we make the wide-sense stationary (WSS) scatterer assumption [36], which implies that the correlation function of ρ_{NB} is given by[13]

$$E[\rho_{NB}(\tau,\nu)\rho_{NB}^*(\tau',\nu')] = M_{NB}(\tau,\tau';\nu)\delta(\nu-\nu') \ , \tag{62}$$

[12] Note that $\rho_{NB} \equiv \rho_{NB}^{(0,0)}$.

[13] Note that the WSS scatterer assumption does not imply that the returned signal is a (wide-sense) stationary process.

for some M_{NB}, known as the *scattering function.* For a zero-mean WSS scatterer, the returned signal $r^{(\tau_o,\nu_o)}$ is zero-mean and its correlation function is given by

$$R_r^{(\tau_o,\nu_o)}(t_1,t_2) = R_{NB}(t_1-\tau_o, t_2-\tau_o)e^{+j2\pi\nu_o t_1}e^{-j2\pi\nu_o t_2} \ , \quad \text{where} \tag{63}$$

$$R_{NB}(t_1,t_2) = \int_{S_T}\int_{S_T}\int_{S_F} p(t_1-\tau)e^{j2\pi\nu t_1}p^*(t_2-\tau')e^{-j2\pi\nu t_2}M_{NB}(\tau,\tau';\nu)d\tau d\tau' d\nu \ . \tag{64}$$

Thus, we see that the family of correlation functions $\{R_r^{(\tau_o,\nu_o)}\}$ in (63), corresponding to different nominal values of range and doppler, is identical in form to the family $\{R_s^{(\tau,\nu)}\}$ defined in (25).

5.2 Wideband model

Let the transmitted pulse, p, be as defined in (59). Using the wideband signal model, the received pulse from a point scatterer is given by

$$r^{(\tau,c)}(t) = a\sqrt{c}p(c(t-\tau)) \ , \tag{65}$$

where a is some complex constant, τ is the delay, and c is the scaling proportional to the relative velocity between the radar and the scatterer. For an arbitrary scatterer, modeled as a continuum of point scatterers, the return is given by

$$r^{(\tau_o,c_o)}(t) = \int_{S_T}\int_{S_C} \sqrt{cc_o}p(c(c_o(t-\tau_o)-\tau))\rho_{WB}(\tau,c)d\tau dc \ , \tag{66}$$

where $S_C = [1/\Delta_C, \Delta_C]$ for some $\Delta_C \geq 1$, and τ_o and c_o are the *nominal* range and *scale* of the scatterer. The above wideband model for an arbitrary scatterer is derived from the natural assumption that the reflectivities of the same scatterer, corresponding to the nominal values (τ_o, c_o) and $(\tau_o/\Delta_c + \Delta_\tau, c_o\Delta_c)$ of delay-scale, must be related by $\rho_{WB}^{(\tau_o/\Delta_c+\Delta_\tau,c_o\Delta_c)}(\tau,c) = \rho_{WB}^{(\tau_o,c_o)}(\Delta_c(\tau-\Delta_\tau), c/\Delta_c)$, which is equivalent to the relation $\rho_{WB}^{(\tau_o,c_o)}(\tau,c) = \rho_{WB}(c_o(\tau-\tau_o), c/c_o)$ used in (66).[14] Again, modeling the scatterer reflectivity ρ_{WB} as being random, the correlation function of the received pulse can be expressed as

$$R_r^{(\tau_o,c_o)}(t_1,t_2) = c_o R_{WB}(c_o(t_1-\tau_o), c_o(t_2-\tau_o)) \ , \quad \text{where} \tag{67}$$

$$R_{WB}(t_1,t_2) = \int_{S_T}\int_{S_T}\int_{S_C}\int_{S_C} \sqrt{cc'}p(c(t_1-\tau))p^*(c'(t_2-\tau'))M_{WB}(\tau,\tau';c,c')d\tau d\tau' dc dc' \ , \tag{68}$$

and $M_{WB}(\tau,\tau';c,c') = E[\rho_{WB}(\tau,c)\rho_{WB}^*(\tau',c')]$ is the wideband scattering function. Thus, we note that the family of correlation functions $\{R_r^{(\tau_o,c_o)}\}$, corresponding to different values of nominal range and scale, is of the same form as the family $\{R_s^{(\tau,c)}\}$ defined in (26).

We just note in passing that narrowband and wideband radar are not the only examples of systems in which correlation functions of the form (25) and (26) arise. In fact, any randomly time-varying channel or linear system characterized by parameterized families of input-output relations of the form (61) and (66), with $p(t)$ as the input and $r(t)$ as the output, will have output correlation functions characterized as (25) or (26). For example, (61) may characterize a random communication channel with an unknown time-delay and/or frequency-offset. With noisy output measurements, the detection of a signal in the output of such systems can be accomplished using the same TFR/TSR detectors as discussed in Section 4.

[14]Note that $\rho_{WB} \equiv \rho_{WB}^{(0,1)}$.

6 Detection Schemes with Partial Signal Information

Although the TFR/TSR-based detectors derived in Section 4 do exploit the information available at different time-frequency/time-scale locations, any information about the nonstationary structure of the signal is not explicitly exploited. The reason is that the nonstationarity of the signal is implicit in the structure of the correlation function; the eigenfunctions completely characterize the nonstationarity. In this section we show how partial information about the signal eigenfunctions only can be exploited to design optimal detectors.[15]

Availability of such partial knowledge of signal eigenfunctions is conceivable in some applications. For example, in radar/sonar scenarios, if we assume that the point scatterers for different values of (τ, ν) and (τ, c) are uncorrelated[16], then (64) and (68) seem to suggest that the eigenfunctions of the correlation function of the received signal may be strongly dependent on the transmitted pulse $p(t)$;[17] thus, different targets may roughly correspond to different eigenvalues if the different time-frequency/time-scale translates of the pulse are roughly orthogonal. In other situations, knowledge about the nonstationary structure of the signal may be used to choose an appropriate basis to model the eigenfunctions of the signal correlation function R_{TF} or R_{TS} characterizing the families $\{R_s^{(\tau,\nu)}\}$ or $\{R_s^{(\tau,c)}\}$, respectively. For example, faults in rotating machinery typically exhibit transient, nonstationary signatures for which an appropriately chosen wavelet basis might be a good model [1, 2]. In such situations, ML estimates of the eigenvalues may be used for the various TFR/TSR detectors.

In order to characterize the TFR/TSR detectors based on ML estimates of signal eigenvalues, it is very convenient to interpret the detectors as a weighted sum of a bank of spectrograms or scalograms, respectively. The short-time Fourier transform (STFT) of a signal x is defined as [12, 31]

$$\mathrm{STFT}_x(t, f; h) = \int x(u) h^*(u - t) e^{-j2\pi f u} du \ , \tag{69}$$

where h is the analysis window, and the wavelet transform (WT) of x is defined as [31, 27]

$$\mathrm{WT}_x(t, a; g) = \int x(u) |a|^{-1/2} g^* \left(\frac{u - t}{a} \right) du \ , \tag{70}$$

where g is called the analysis or mother wavelet. By comparing (69) with (40), (37) and (38), it can be easily verified that

$$\begin{aligned}
L_{\mathrm{O}}^{(\tau,\nu)}(x) &= \frac{1}{N_0} \sum_k \frac{\lambda_k}{\lambda_k + N_0} |\mathrm{STFT}_x(\tau, \nu; u_k)|^2 + \sum_k \log\left(\frac{N_0}{\lambda_k + N_0} \right) \ , & (71)\\
L_{\mathrm{H}}^{(\tau,\nu)}(x) &= \sum_k \lambda_k |\mathrm{STFT}_{\mathbf{R}_n^{-1} x}(\tau, \nu; u_k)|^2 \ , \text{ and} & (72)\\
L_{\mathrm{LO}}^{(\tau,\nu)}(x) &= L_{\mathrm{H}}^{(\tau,\nu)}(x) - \sum_k \lambda_k \langle u_k^{(\tau,\nu)}, \mathbf{R}_n^{-1} u_k^{(\tau,\nu)} \rangle \ , & (73)
\end{aligned}$$

where in (71) we have included a constant term which depends on the signal eigenvalues. Similarly, the TSR-based test statistics can be expressed as

$$L_{\mathrm{O}}^{(\tau,c)}(x) = \frac{1}{N_0} \sum_k \frac{\mu_k}{\mu_k + N_0} |\mathrm{WT}_x(\tau, 1/c; v_k)|^2 + \sum_k \log\left(\frac{N_0}{\mu_k + N_0} \right) \ , \tag{74}$$

[15]It is worth noting that determining the eigendecomposition for any one member of each of families $\{R_s^{(\tau,\nu)}\}$ and $\{R_s^{(\tau,c)}\}$ characterizes the eigendecompositions for the entire family.

[16]That is, $M_{NB}(\tau, \tau'; \nu) = M_{NB}^o(\tau, \nu)\delta(\tau - \tau')$ in (64) and $M_{WB}(\tau, \tau'; c, c') = M_{WB}^o(\tau, c)\delta(\tau - \tau')\delta(c - c')$ in (68). The assumption in the narrowband case is equivalent to the "WSS uncorrelated scatterer" assumption in [36].

[17]Note that under the assumption of uncorrelated scatterers (see footnote 16), (64) and (68) reduce to integral expansions for $R_s^{(\tau,\nu)}$ and $R_s^{(\tau,c)}$, respectively, in terms of outer products of time-frequency translates, and time-shifted and scaled versions, respectively, of the transmitted pulse $p(t)$; that is, they resemble eigenexpansions.

$$L_{\mathrm{H}}^{(\tau,c)}(x) = \sum_k \mu_k |\mathrm{WT}_{\mathbf{R}_n^{-1}x}(\tau, 1/c; v_k)|^2 \ , \text{ and} \tag{75}$$

$$L_{\mathrm{LO}}^{(\tau,c)}(x) = L_{\mathrm{H}}^{(\tau,c)}(x) - \sum_k \mu_k \langle v_k^{(\tau,c)}, \mathbf{R}_n^{-1} v_k^{(\tau,c)} \rangle \ . \tag{76}$$

It is shown in Appendix B that in Case I, considering the family of correlation functions $\{R_s^{(\tau,\nu)}\}$, the ML estimates of the eigenvalues of $R_s^{(\tau,\nu)}$, for each value of (τ,ν), are given by

$$\hat{\lambda}_k^{(\tau,\nu)}(x) = \max\left\{0, |\langle u_k^{(\tau,\nu)}, x\rangle|^2 - N_0\right\} = \max\left\{0, |\mathrm{STFT}_x(\tau,\nu;u_k)|^2 - N_0\right\} \ . \tag{77}$$

Similarly, considering the family $\{R_s^{(\tau,c)}\}$ for Case I, the ML estimates of the eigenvalues of $R_s^{(\tau,c)}$, for each value of (τ, c), are

$$\hat{\mu}_k^{(\tau,c)}(x) = \max\left\{0, |\langle v_k^{(\tau,c)}, x\rangle|^2 - N_0\right\} = \max\left\{0, |\mathrm{WT}_x(\tau,1/c;v_k)|^2 - N_0\right\} \ . \tag{78}$$

The above eigenvalue estimates can then be substituted in (71) and (74) to realize the TFR/TSR detectors for Case I, as described in Propositions A and B. An interpretation of the ML estimates (77)/(78) is that for each value of $(\tau,\nu)/(\tau,c)$, only those STFT/WT banks should be included in the expansion (71)/(74) whose output power is greater than the noise power. This makes intuitive sense because in the absence of knowledge about signal power in the different "coordinates" (projections onto eigenfunctions), only those coordinates should be used in making a decision in which the measured power exceeds the noise power.

In Cases II and III, because of the nature of the detectors, we have to impose an energy constraint on the signal:

$$\mathrm{Trace}(R_s^{(\tau,\nu)}) = \mathrm{Trace}(R_{TF}) = \sum_k \lambda_k \leq d_{TF} \quad \text{and} \tag{79}$$

$$\mathrm{Trace}(R_s^{(\tau,c)}) = \mathrm{Trace}(R_{TS}) = \sum_k \mu_k \leq d_{TS} \tag{80}$$

for some bounds, d_{TF} and d_{TS}, on signal energy. Similar to the derivation for Case I in Appendix B, it can be shown that for Case II, subject to the above energy constraints, the "maximum *SNR*" estimates of the eigenvalues of $R_s^{(\tau,\nu)}$ and $R_s^{(\tau,c)}$ are given by

$$\tilde{\lambda}_k^{(\tau,\nu)}(x) = \begin{cases} d_{TF} & \text{if } \ k = \arg\max_i \left\{|\langle u_i^{(\tau,\nu)}, \mathbf{R}_n^{-1}x\rangle| = |\mathrm{STFT}_{\mathbf{R}_n^{-1}x}(\tau,\nu;u_i)|\right\} \\ 0 & \text{else} \end{cases} \ , \text{ and} \tag{81}$$

$$\tilde{\mu}_k^{(\tau,c)}(x) = \begin{cases} d_{TS} & \text{if } \ k = \arg\max_i \left\{|\langle v_i^{(\tau,c)}, \mathbf{R}_n^{-1}x\rangle| = |\mathrm{WT}_{\mathbf{R}_n^{-1}x}(\tau,1/c;v_i)|\right\} \\ 0 & \text{else} \end{cases} \ . \tag{82}$$

It should be noted that only a single eigenvalue estimate is nonzero in (81) and (82); if more than one eigenvectors result in the maximum projection, the eigenvalue corresponding to any one of them may be chosen to be the nonzero one. Similarly, for Case III, subject to the energy constraints, the "maximum (local) likelihood" estimates of the eigenvalues are given by

$$\breve{\lambda}_k^{(\tau,\nu)}(x) = \begin{cases} d_{TF} & \text{if } \ k = \arg\max_i \left\{A_i = |\mathrm{STFT}_{\mathbf{R}_n^{-1}x}(\tau,\nu;u_i)|^2 - \langle u_i^{(\tau,\nu)}, \mathbf{R}_n^{-1}u_i^{(\tau,\nu)}\rangle : A_i > 0\right\} \\ 0 & \text{else} \end{cases} \tag{83}$$

$$\breve{\mu}_k^{(\tau,c)}(x) = \begin{cases} d_{TS} & \text{if } \ k = \arg\max_i \left\{B_i = |\mathrm{WT}_{\mathbf{R}_n^{-1}x}(\tau,1/c;v_i)|^2 - \langle v_i^{(\tau,c)}, \mathbf{R}_n^{-1}v_i^{(\tau,c)}\rangle : B_i > 0\right\} \\ 0 & \text{else} \end{cases} \ . \tag{84}$$

The estimates in (81), (82), (83) and (84) can then be substituted in (72), (75), (73) and (76), respectively, to implement the TFR/TSR detectors for Cases II and III as described in Propositions A and B.

The eigenvalue estimates for the deflection-optimal and locally optimal test statistics have a very interesting interpretation: only the largest projection (STFT/WT bank corresponding to largest output magnitude), after subtracting the corresponding noise term in the locally optimum case, should be included in the expansions (72), (75), (73) and (76). Thus, for each value of (τ, ν) or (τ, c), effectively, the output of a "matched filter" (corresponding to the largest projection) is used in the computation of the test statistic. This is consistent with the fact that for a given signal energy, a rank-1 signal maximizes the deflection [22].[18]

7 Conclusions

The need to detect transients and for detection in nonstationary environments has led to extensive use of bilinear time-frequency representations in such applications. However, current usage has been limited to merely equivalent time-frequency implementations of classical quadratic detectors and to *ad hoc* approaches with no assurance that time-frequency representations are even appropriate, let alone optimal, for such problems.

Recognizing that bilinear TFRs/TSRs are quadratic in the data, we used classical quadratic detection theory to identify the role of TFRs/TSRs as detectors. We characterized certain hypothesis testing situations in Section 3, involving Gaussian signals, in which time-frequency-based detection is both optimal and results in efficient realizations of the detection statistics via a TFR or TSR. Propositions A and B in Section 4 explicitly characterize the corresponding TFR/TSR detectors in those situations. In general, the scenarios in which time-frequency-based detection is optimal are composite hypothesis tests involving time and frequency shifts or time-shifts and scalings as unknown or random parameters; the TFR/TSR performs a search over these parameters, implementing an optimal quadratic detector at each point. While detectors for other situations can often be described equivalently in terms of TFRs (e.g., the LR-based detector in Case II), we believe there is little reason to do so.

In certain situations (random parameters with known pdf), the form of the truly optimal detector is complicated and an explicit description of it as a function of the observations is usually not possible. In such cases, we propose a MAP GLRT detector which, while suboptimal, is easy to implement and should perform better than the ML detector.

As shown in Section 5, standard narrowband and wideband radar/sonar models fall within the class of problems for which TFR/TSR-based detection is optimal. In addition to radar/sonar, the proposed TFR/TSR-based detectors can be directly applied to other situations that involve a random linear channel (or time-varying system) with unknown or random time and frequency shifts or time-shifts and scalings.

The proposed bilinear TFR/TSR-based detectors are intimately related to their linear counterparts implemented via STFTs/WTs, respectively. The bilinear TFR/TSR detectors can be interpreted as a weighted sum of a bank of spectrograms/scalograms, with the number of filters in the bank being equal to the rank of the signal correlation function. Thus, for rank-1 signals, the bilinear detectors essentially reduce to their corresponding linear counterparts for detecting a deterministic signal with unknown phase, time-frequency, or time-scale parameters. This yields a unified subspace-based detection interpretation: a weighted projection of the observed signal onto the subspace spanned by eigenfunctions of the signal correlation function yields the optimal test statistic. This subspace-based formulation yields alternative forms for the optimal detectors which are naturally suited for efficient implementation via spectrograms/scalograms. In many situations, even though only partial information is available about the nonstationary structure of

[18] By a rank-1 signal we mean that its correlation function is rank-1 (one term in the eigenexpansion).

the signal, optimal detectors can be designed using the subspace-based formulation, as shown in Section 6.

In this paper, with quadratic detection in mind, we only considered variants of the classical problem of detecting a Gaussian signal in Gaussian noise. However, even for arbitrary signal and noise statistics, optimal quadratic detectors can be derived based on the deflection criterion [21]. In fact, the optimal deflection-based detector for an arbitrary second-order signal in Gaussian noise is exactly the detector in Case II [19]. Thus, it is conceivable that TFR/TSR detectors may be useful in such situations as well. The main issue would be to identify meaningful scenarios in which the parameters of the detection problem naturally correspond to time-frequency or time-scale. A second possible generalization stems from the fact that both the classes of signal correlation functions that we derived in this paper resulted from certain unitary transformations of random signals. It is conceivable that some applications may be characterized by some other unitary signal transformations [37, 38, 39]. In such cases the optimal detectors may be naturally characterized by joint distributions of the variables defined by the unitary transformation [38, 40, 41, 39].

Appendix A

We need to show that the kernel $\Phi^{(\alpha,\beta)}$ defined in (22) is independent of (α,β) if and only if the condition (23) holds. We have

$$\Phi^{(\alpha,\beta)}(u,v) = \Phi^{(\alpha',\beta')}(u,v) \text{ for all } (u,v,\alpha,\beta,\alpha',\beta') \in \mathbb{R}^6$$

$$\Leftrightarrow \mathrm{WS}_{Q^{(\alpha,\beta)}}(u+\alpha, v+\beta) = \mathrm{WS}_{Q^{(0,0)}}(u,v) \text{ for all } (u,v,\alpha,\beta) \in \mathbb{R}^4 \text{ (using (22))}$$

$$\Leftrightarrow \int \left[Q^{(\alpha,\beta)}\left(u+\alpha+\frac{\tau}{2}, u+\alpha-\frac{\tau}{2}\right) e^{-j2\pi\beta\tau} - Q^{(0,0)}\left(u+\frac{\tau}{2}, u-\frac{\tau}{2}\right)\right] e^{-j2\pi v\tau} d\tau = 0 \text{ for all } (u,v,\alpha,\beta) \in \mathbb{R}^4$$

$$\Leftrightarrow Q^{(\alpha,\beta)}\left(u+\frac{\tau}{2}, u-\frac{\tau}{2}\right) = Q^{(0,0)}\left(u+\frac{\tau}{2}-\alpha, u-\frac{\tau}{2}-\alpha\right) e^{+j2\pi\beta\tau} \text{ for al } (u,\tau,\alpha,\beta) \in \mathbb{R}^4$$

which is equivalent to (23) via the coordinate transformation $t_1 = u + \tau/2$, $t_2 = u - \tau/2$.

Appendix B

We derive the ML estimates for the eigenvalues of the test statistic in Case I in Section 2. We assume that the signal correlation function R_s has a finite rank N:[19]

$$R_s(t_1, t_2) = \sum_{k=1}^{N} \lambda_k \varphi_k(t_1) \varphi^*(t_2) \ . \tag{85}$$

The log likelihood function in Case I is given by [19]

$$\log L(x) = \sum_{k=1}^{N} \log\left(\frac{N_0}{\lambda_k + N_0}\right) + \frac{1}{N_0} \sum_{k=1}^{N} \frac{\lambda_k}{\lambda_k + N_0} |\langle x, \varphi_k \rangle|^2 \ , \tag{86}$$

Given the observed signal, x, we are interested in finding the ML estimates, $\hat{\lambda}_k(x)$'s, of the signal eigenvalues. For a given x, let us denote the log likelihood function (86) by $f(\lambda)$, a function of the eigenvalues $\lambda = (\lambda_1, \lambda_2, \cdots \lambda_N)$. Our objective is to maximize $f(\lambda)$ subject to the constraints that $\lambda_i \geq 0$, for all i. The first-order necessary (Kuhn-Tucker) conditions for a solution are [42]

$$\begin{aligned} -\frac{\partial f(\lambda)}{\partial \lambda_i} - \gamma_i &= -\frac{|\langle x, \varphi_i \rangle|^2 - (N_0 + \lambda_i)}{(\lambda_i + N_0)^2} - \gamma_i = 0 \ , \\ \gamma_i \lambda_i &= 0 \ , \quad i = 1, 2 \cdots N \end{aligned} \tag{87}$$

[19]Even if R_s has infinite rank, our analysis could be construed as a finite rank approximation to the solution.

where $\gamma_i \geq 0$ are the Lagrange multipliers. It can be easily inferred from (87) that the ML estimates of the eigenvalues are given by

$$\hat{\lambda}_k(x) = \max\left\{|\langle\varphi_k, x\rangle|^2 - N_0, 0\right\} \quad , k = 1, 2 \cdots N \ . \tag{88}$$

Acknowledgement

The authors would like to thank the reviewers for their useful comments and suggestions that resulted in a more streamlined presentation.

References

[1] M. J. Dowling, "Application of non-stationary analysis to machinery monitoring", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '93*, 1993, pp. I–59–I–62.

[2] R. A. Rohrbaugh, "Application of time-frequency analysis to machinery condition assessment", in *Proc. 27th Asilomar Conference on Signals, Systems, and Computers*, November 1993.

[3] L. P. Heck, "Signal processing research in automatic tool wear monitoring", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '93*, 1993, pp. I–55–I–58.

[4] T. Brotherton, T. Pollard, and D. Jones, "Applications of time-frequency and time-scale representations to fault detection and classification", in *Proc. IEEE Int'l Symp. on Time-Frequency and Time-Scale Analysis*, 1992, pp. 95–98.

[5] K. C. Chou and L. P. Heck, "A multiscale stochastic modeling approach to the monitoring of mechanical systems", in *Proc. IEEE Int'l Symp. on Time-Frequency and Time-Scale Analysis*, 1994, pp. 25–27.

[6] J. E. Lopez, R. R. Tenney, and J. C. Deckert, "Fault detection and identification using real-time wavelet feature extraction", in *Proc. IEEE Int'l Symp. on Time-Frequency and Time-Scale Analysis*, 1994, pp. 217–220.

[7] L. E. Atlas, J. Fang, P. Loughlin, and W. Music, "Resolution advantages of quadratic signal processing", in *Advanced Signal-Processing Algorithms, Architectures and Implementation II*, F. T. Luk, Ed., 1991, vol. Proc. SPIE 1566, pp. 134–143.

[8] L. E. Atlas and J. Fang, "Quadratic detectors for general nonlinear analysis of speech", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '92*, 1992, pp. II–9—II–12).

[9] J. Fang and L. E. Atlas, "Quadratic detectors for noise-immune detection of modulated signals", in *Proc. 27th Asilomar Conference on Signals, Systems, and Computers*, November 1993.

[10] J. Fang and L. E. Atlas, "Quadratic detectors for energy estimation", *Submitted to IEEE Trans. Signal Processing*, September 1993.

[11] J. Fang, L. Atlas, and G. Bernard, "Advantages of cascaded quadratic detectors for analysis of manufacturing sensor data", in *Proc. IEEE Int'l Symposium on Time-Frequency and Time-Scale Analysis*, Victoria, Canada, 1992, pp. 345–348.

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

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

[14] R. A. Altes, "Wide-band proportional-bandwidth Wigner-Ville analysis", *IEEE Trans. Acoust., Speech Signal Processing*, vol. 38, pp. 1005–1012, June 1990.

[15] A. Papandreou, S. M. Kay, and G. F. Boudreaux-Bartels, "The use of hyperbolic time-frequency representations for optimum detection and parameter estimation of hyperbolic chirps", in *Proc. IEEE Int'l Symp. on Time-Frequency and Time-Scale Analysis*, 1994, pp. 369–372.

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

[17] S. Kay and G. F. Boudreaux-Bartels, "On the optimality of Wigner distribution on detection", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '85*, Tampa, FL, 1985, pp. 27.2.1–27.2.4.

[18] P. Flandrin and P. Gonçalvès, "From wavelets to time-scale energy distributions", in *Recent Advances in Wavelet Analysis*, L. L. Schumaker and G. Webb, Eds., Academic Press, 1994, pp. 309–334.

[19] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 1988.

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

[21] B. Picinbono and P. Duvaut, "Optimal linear-quadratic systems for detection and estimation", *IEEE Trans. Inform. Theory*, vol. 43, pp. 304–311, March 1988.

[22] C. R. Baker, "Optimum quadratic detection of a random vector in Gaussian noise", *IEEE Trans. Comm. Tech.*, vol. COM-14, no. 6, pp. 802–805, December 1966.

[23] K. S. Miller, *Complex Stochastic Processes*, Addison-Wesley, 1974.

[24] E. Parzen, "Probability density functionals and reproducing kernel Hilbert spaces", in *Time Series Analysis*, John Wiley & Sons, Inc., New York, 1963, (M. Rosenblatt, Ed.), pp. 155–169.

[25] V. V. Veeravalli and H. V. Poor, "Quadratic detection of signals with drifting phase", *J. Acoust. Soc. Am.*, vol. 89, no. 2, pp. 811–819, February 1991.

[26] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution — a tool for time-frequency signal analysis — part I: Continuous-time signals", *Philips Journal of Research*, vol. 35, no. 3, pp. 217–250, 1980.

[27] O. Rioul and P. Flandrin, "Time-scale distributions: A general class extending the wavelet transform", *IEEE Trans. Signal Processing*, vol. 40, pp. 1746–1757, May 1992.

[28] G. B. Folland, *Harmonic Analysis in Phase Space*, Princeton University Press, Princeton, NJ, 1989.

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

[30] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis", *IEEE Trans. Signal Processing*, vol. 42, no. 2, pp. 318–332, Feb. 1994.

[31] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations", *IEEE Signal Processing Magazine*, April 1992.

[32] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice Hall, New Jersey, 1986.

[33] R. G. Shenoy and T. W. Parks, "Wide-band ambiguity functions and affine Wigner distributions", *Signal Processing*, vol. 41, no. 3, pp. 339–363, 1995.

[34] W. Martin, "Time-frequency analysis of random signals", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '82*, 1982, pp. 1325–1328.

[35] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of non-stationary processes", *IEEE Trans. Acoust., Speech and Signal Processing*, vol. 33, no. 6, pp. 1461–1470, December 1985.

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

[37] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing", *IEEE Trans. Signal Processing*, vol. 43, no. 10, pp. 2269–2282, October 1995.

[38] A. M. Sayeed and D. L. Jones, "Integral transforms covariant to unitary operators and their implications for joint signal representations", *IEEE Trans. Signal Processing*, vol. 44, pp. 1365–1377, June 1996.

[39] F. Hlawatsch and H. Bölcskei, "Unified theory of displacement-covariant time-frequency analysis", in *Proc. IEEE Int'l Symp. on Time-Frequency and Time-Scale Analysis*, 1994, pp. 524–527.

[40] R. G. Baraniuk, "Beyond time-frequency analysis: energy densities in one and many dimensions", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '94*, 1994, vol. 3, pp. 357–360.

[41] A. M. Sayeed and D. L. Jones, "On the equivalence of generalized joint signal representations", in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Proc. — ICASSP '95*, 1995, pp. 1533–1536.

[42] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison-Wesley, Reading, Massachusetts, second edition, 1989.