Detection, estimation, and classification with spectrograms

Richard A. Altes

ORINCON Corporation, 3366 North Torrey Pines Court, La Jolla, California 92037
(Received 1 March 1979; accepted for publication 20 November 1979)

A locally optimum detector correlates the data spectrogram with a reference spectrogram in order to detect (i) a known signal with unknown delay and Doppler parameters, (ii) a random signal with known covariance function, or (iii) the output of a random, time-varying channel with known scattering function. Spectrogram correlation can also be used for maximum likelihood parameter estimation, e.g., estimation of delay or center frequency of a signal. To estimate an analog input signal from its spectrogram, a modified deconvolution operation can be used together with a predictive noise canceler. If no noise is added to the spectrogram, the mean-square error of this signal estimate is independent of the window function that is used to construct the spectrogram. When estimates of specific signal parameters are obtained directly from the spectrogram, these estimates have mean-square errors that depend upon both signal and window waveforms. Spectrogram correlation can be used for classification as well as for estimation and detection. Parameter estimators and detectors are, in fact, specialized kinds of classifiers.

PACS numbers: 43.60.Gk, 43.60.Cg, 43.85.Kr

INTRODUCTION

Background and goals

A spectrogram provides a convenient signal representation by portraying a waveform as a non-negative function of instantaneous frequency and time. The popularity of spectrogram representations is associated with spectral decomposition of data that has a time–bandwidth product which is too large for an available Fourier transform device. For such data, a sequence of Fourier transforms is obtained by shifting a window function across the data and calculating a Fourier transform for each new window position. The spectrogram retains only the magnitude of each Fourier transform and discards the phase, but we shall see that little information is destroyed by this process.

Although spectrograms can be regarded as a means for estimating the power spectral density of a stationary random process,[1-3] they are often used in practice to portray transient effects, e.g., individual phonemes in speech.[4-6] The ability to represent transient phenomena suggests that spectrograms may contain information about the autocovariance function of a *nonstationary* process. It will be shown that this is indeed the case, and that the autocovariance function of a nonstationary random process can be estimated from the spectrogram of the process. This observation indicates that spectrogram representations and linear prediction coefficients[7] are based upon the same information.

If a spectrogram is a representation of an autocovariance function, then detectors and estimators that use autocovariance functions to describe random signals and noise can use spectrograms instead. Spectrogram descriptions can lead to simple and intuitively straightforward detector and classifier configurations. It will be shown, for example, that very simple spectrogram operations are needed for locally optimum detection of a signal with known covariance matrix in white, Gaussian noise. A simple generalization to non-Gaussian noise distributions is also possible.

Random processes are often created when signals are passed through a randomly time-varying filter. Examples are speech (the output of a vocal tract filter), communication signals that have been passed through a random, time-varying channel, and sonar echoes from moving, range distributed scatterers (e.g., fish schools or surface reverberation). Randomly time-varying filters are often described in terms of a scattering function.[8,9] A relation between scattering functions and spectrograms will be obtained in this paper. A different relation that applies to random, time *invariant* signal transformations will also be discussed.

Overview of the paper

The paper begins with some definitions and properties of spectrograms and related signal representations. One of the most important properties is that an input signal can be reconstructed from a noise-free spectrogram, except for a multiplicative constant with unit magnitude and unknown phase. Another property indicates that only the signal auto-ambiguity function (which can be obtained from a spectrogram by deconvolution) is needed for matched filtering or pulse compression. Reconstruction of the waveform itself is thus unnecessary for detection of a deterministic signal. The effect of noise upon the signal reconstruction process is investigated, and sampling considerations are discussed.

For a random signal, it is shown that an ensemble-average spectrogram is equivalent to a representation of the signal in terms of its covariance function, or in terms of the scattering function of a random channel.

A *locally optimum detector*[10,11] is optimum for very small signal-to-noise ratio (SNR), where optimality is most important, but it may be nonideal for larger SNR, where optimality is not as critical. Locally optimum detectors are determined for (i) a random signal with known covariance function (or the output of a time-varying channel with known scattering function), (ii) a signal that is known except for delay, Doppler shift, and phase, and (iii) a signal that is known except for phase. In all cases it is initially assumed that white,

 0001-4966/80/041232-15$00.80

Gaussian noise has been added to the signals, although the detectors can be adjusted for non-Gaussian noise. A spectrogram representation can be used for all of these problems. The intuitive concept of detecting a signal by summing weighted, envelope detected matched filter responses is justified by the locally optimum solution for case (ii), and a simple implementation of this detector is obtained by using a time window spectrogram with a particular window function.

A measurement problem may involve estimation of an analog time *signal*, or of specific signal *parameters* such as time-of-arrival or center frequency. A minimum mean-square error (MMSE) *signal* estimate can be obtained from a spectrogram, except that an unknown constant will always be added to the signal's phase. Estimates of signal *parameters* can be obtained directly from the spectrogram. Variance bounds that describe the accuracy of different parameter estimates are compared, and their dependence upon signal and window waveforms are assessed.

A maximum likelihood classifier can be configured as a parallel connection of detectors; the largest detector response indicates the most likely signal category. It is argued that detectors and parameter estimators are special cases of such a classifier, and that this concept can lead to an even simpler form of locally optimum detector.

I. DEFINITIONS AND PROPERTIES OF TIME-FREQUENCY ENERGY DISTRIBUTIONS, SPECTROGRAMS, AMBIGUITY FUNCTIONS, AND SCATTERING FUNCTIONS

A given real signal $u_r(t)$ can be used to construct a *complex waveform*[12-14]

$$u(t) = 2^{-1/2}[u_r(t) + j\hat{u}_r(t)], \tag{1}$$

where $\hat{u}_r(t)$ is the Hilbert transform[15] of $u_r(t)$. The waveform $u(t)$ in (1) is the so-called *analytic representation* of $u_r(t)$, and $\mathrm{Re}\{u(t)\} = 2^{-1/2}u_r(t)$. The Fourier transform of $u(t)$ is

$$U(f) = \int_{-\infty}^{\infty} u(t)\exp(-j2\pi ft)\,dt. \tag{2}$$

If $u(t)$ is analytic, $U(f)$ is identically zero for $f < 0$. The factor $2^{-1/2}$ in (1) is included to preserve signal energy.

The envelope of $u_r(t)$ is usually approximated by passing a rectified version of $u_r(t)$ through a low-pass filter or integator. An exact version of the *envelope*, however, is

$$|u(t)| = 2^{-1/2}[u_r^2(t) + \hat{u}_r^2(t)]^{1/2} \tag{3}$$

which is the magnitude of the analytic representation of $u_r(t)$.

The *time–frequency energy density function* of $u(t)$ is defined[16] as

$$e_{uu}(t,f) = u(t)U^*(f)\exp(-j2\pi ft). \tag{4}$$

The *narrow-band ambiguity function* of $u(t)$ is[17]

$$X_{uu}(\tau,\varphi) = \int_{-\infty}^{\infty} u(t)u^*(t+\tau)\exp(-j2\pi\varphi t)\,dt. \tag{5}$$

Similarly, we can define the *cross-energy density function* and *cross-ambiguity function* by

$$e_{uv}(t,f) = u(t)V^*(f)\exp(-j2\pi ft) \tag{6}$$

$$X_{uv}(\tau,\varphi) = \int_{-\infty}^{\infty} u(t)v^*(t+\tau)\exp(-j2\pi\varphi t)\,dt. \tag{7}$$

A spectrogram can be obtained either by bandpass filtering or by Fourier transformation of a time-windowed signal. In the case of bandpass filtering, the *short-time spectral history* or *spectrogram* is[18]

$$S_{uv}(t_1,f_1) = \left|\int_{-\infty}^{\infty} U(f)V(f-f_1)\exp(j2\pi ft_1)\,df\right|^2 \tag{8}$$

$$= \left|\int_{-\infty}^{\infty} u(t_1-t)v(t)\exp(j2\pi f_1 t)\,dt\right|^2, \tag{9}$$

where $u(t)$ and $v(t)$ are analytic. $S_{uv}(t_1,f_1)$ is the squared envelope of the temporal response of a filter with transfer function $V(f-f_1)$, evaluated at time t_1.

The time–frequency energy density function is defined so as to portray the variation of $u(t)$ in the time and frequency domains by means of a three-dimensional plot.[12] The spectrogram is also constructed so as to display this variation. The spectrogram is obtained by passing the signal $u(t)$ through a bank of filters that are all shifted versions of the same basic function $V(f)$.

Historically, the ambiguity function was motivated by the need to detect a radar signal $u(t)$ in Gaussian noise, and to simultaneously estimate its delay τ and Doppler shift ϕ. If $V^*(f) = U^*(f)$, the filter $V^*(f-\phi)$ tests the hypothesis that the transmitted signal spectrum $U(f)$ has been Doppler shifted by ϕ Hz. The radar receiver implements a sequence of such hypotheses by using a bank of filters with ϕ-dependent transfer functions. When $u(t)$ is present at the input, the filter responses as a function of range (time) and Doppler shift (frequency) are represented by the ambiguity function. This interpretation of the ambiguity function explains its close connection to time–frequency signal representations.

Some relevant properties of time–frequency energy density functions, spectrograms, and ambiguity functions are listed in Appendix A. *These properties show that the autoambiguity function of an input signal can be obtained from a noise-free spectrogram by deconvolution.* A correlation (pulse compression) operation can then be implemented by integrating the product of the data autoambiguity function and the conjugated autoambiguity function of a reference signal. Alternatively, one can obtain the input signal itself (apart from a complex constant) from its autoambiguity function.

It is also shown in Appendix A that knowledge of the *covariance function*

$$R_{uu}(t,\tau) = \overline{u(t)u^*(t+\tau)} \tag{10}$$

of a random signal $u(t)$ is equivalent to knowledge of the signal's ensemble-average spectrogram $\overline{S_{uv}(t,f)}$. This

equivalence follows from the fact that $R_{uu}(t,\tau)$ *and* $\overline{S_{uv}(t,f)}$ *can be obtained from one another by linear, reversible transformations.*

A *scattering function* of a time-varying random filter describes the expected dispersion of the input signal in time and frequency.[8,9] A scattering function can be defined in terms of an input–output relation from Appendix A;

$$\overline{S_{zv}(t,f)} = \Gamma_h(t,f) \overset{t,f}{*} S_{uv}(t,f), \tag{11}$$

where $\overline{S_{zv}(t,f)}$ is the ensemble average spectrogram of the filter output $z(t)$, $\Gamma_h(t,f)$ is the filter scattering function ($\overset{t,f}{*}$) denotes two-dimensional convolution in time and frequency, and $S_{uv}(t,f)$ is the spectrogram of the input signal $u(t)$. A time-varying filter smears the spectrogram of the input signal by a two-dimensional convolution operation, and the expected extent of smearing is determined by the scattering function of the filter.

If the random filter is *time invariant*, then

$$\overline{S_{zv}(t,f)} = \overline{e_{hh}(t,f)} \overset{t}{*} e_{uu}(t,f) \overset{t,f}{*} e_{vv}(t,-f), \tag{12}$$

where $\overset{t}{*}$ is a *one*-dimensional convolution in the time variable and where $\overline{e_{hh}(t,f)}$ is the expected time–frequency energy density function of the filter impulse response (Appendix A). Comparison of (11) and (12) indicates that a scattering function is a generalization of a time–frequency energy density function that applies when a filter has a time-varying transfer function.

Sampling of the spectrogram is discussed in Appendix B. Two different sampling densities apply, depending upon how the spectrogram is to be used. For parsimonious representation of the spectrogram itself, samples should be spaced

$$\Delta t = (2B_v)^{-1} \text{ s} \tag{13}$$

apart in the time direction and

$$\Delta f = (2T_v)^{-1} \text{ Hz} \tag{14}$$

apart in the frequency direction, where B_v and T_v are the approximate bandwidth and duration of $v(t)$, the filter impulse response or window function that is used to construct the spectrogram.

If deconvolution is to be applied in order to reconstruct a signal or its ambiguity function, spectrogram samples should be at most

$$\Delta t = [2 \max(B_u, B_v)]^{-1} \text{ s} \tag{15}$$

apart in time and at most

$$\Delta f = [2 \max(T_u, T_v)]^{-1} \text{ Hz} \tag{16}$$

apart in frequency, where B_u and T_u are the approximate bandwidth and duration of the signal.

II. DETECTOR CONFIGURATIONS

A locally optimum detector[10,11] performs optimally for low signal-to-noise ratio (SNR) where optimality is critical, but its performance may be suboptimum for larger SNR, where optimality is less important. In Appendix C, locally optimum detectors are obtained for three different problems. Each of these detectors implements a statistical test to distinguish between two hypotheses.

For case 1, the hypotheses are

H_1: The data $z(t)$ = [a random signal $u(t)$ with known covariance function] + white, Gaussian noise (WGN).

H_0: $z(t)$ = WGN alone.

From properties 16 and 17 in Appendix A, this problem is equivalent to detecting the output of a randomly time-varying filter, where the scattering function of the filter is specified. The locally optimum detector for this problem compares the statistic

$$l[z(t)] = \sum_{m=1}^{M} \sum_{n=1}^{N} z_{mn} \overline{s_{mn}} \tag{17}$$

with a threshold, where $\{z_{mn}\}$ are statistically independent samples of the data spectrogram, i.e.,

$$z_{mn} = S_{zv}(t_m, f_n), \quad m = 1, \ldots, M; n = 1, \ldots, N \tag{18}$$

and

$$\overline{s_{mn}} = \overline{S_{uv}(t_m, f_n)}, \quad m = 1, \ldots, M; n = 1, \ldots, N \tag{19}$$

are samples of the ensemble average spectrogram of the random noise-free signal, $u(t)$. From property 16 in Appendix A, $\{\overline{s_{mn}}\}$ can be obtained from the known covariance function of the random signal. Alternatively, $\{\overline{s_{mn}}\}$ can be obtained from the scattering function of a time-varying target or channel (property 17). For random, time invariant targets or channels, $\{\overline{s_{mn}}\}$ can be obtained from the expected energy density function of the random impulse response (property 18).

The results of Poor and Thomas[19] indicate that the locally optimum detector can be reconfigured for non-Gaussian noise by using a nonlinearity other than a square law device for determining the samples $\{z_{mn}\}$ in (17).

Equations (17)–(19) describe a *spectrogram correlation process*, where samples of the data spectrogram are correlated with corresponding samples of a reference spectrogram.

For case 2, the hypotheses are

H_1: $z(t) = \{$a signal $u(t-\tau) \exp[j(2\pi\phi t + \lambda)]$ with random phase λ and with unknown delay and Doppler parameters (τ, ϕ) which are described by a prior pdf $p(\tau, \phi)\}$ + WGN.

H_0: $z(t)$ = WGN alone.

In this case, the signal $u(t)$ itself is assumed to be known exactly. From Appendix C, the locally optimum detector for this case compares the statistic

$$l[z(t)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\tau, \phi) |X_{zu}(\tau, \phi)|^2 \, d\tau \, d\phi \tag{20}$$

with a threshold. According to property 4 in Appendix A, the magnitude-squared cross-ambiguity function $|X_{zu}(\tau, \phi)|^2$ can be interpreted as a spectrogram that is generated by a bank of filters with transfer functions $U^*(f-\phi)$, where $U(f)$ is the Fourier transform of the transmitted signal. *The statistic in* (20) *is thus a form of spectrogram correlator.* The spectrogram that is generated by a bank of matched filters (i.e., the mag-

nitude-squared cross-ambiguity function) is correlated with the probability density function $p(\tau, \phi)$ which describes prior knowledge of delay and Doppler shift.

A sampled version of $|X_{zu}(\tau, \phi)|^2$ can be efficiently generated with a fast Fourier transform (FFT) algorithm. Data is multiplied by a sequence of time-shifted window functions $w(t-t_1)=u^*(t-t_1)$, and the product is passed through a discrete Fourier transform. From (A24), we see that the magnitude-squared samples of this discrete Fourier transform are samples of the time window spectrogram $S^T_{zu*}(t_1, f_1)$, and from (7), $S^T_{zu*}(t_1, f_1)=|X_{zu}(-t_1, f_1)|^2$. A time window spectrogram can thus be made equivalent to a bank of matched filters. The time window spectrogram is advantageous for digital implementations because it is easy to synthesize with FFT algorithms, and because the window function can easily be changed to accommodate new signal waveforms.

For case 3, the hypotheses are

$H_1: z(t)=$(a signal that is known exactly except for a random, uniformly distributed phase parameter) +WGN.

$H_0: z(t)=$WGN alone.

The locally optimum detector for an analytic signal $u(t)\exp(j\lambda)$, where λ is unknown, compares the statistic

$$l[z(t)]=\left|\int_{-\infty}^{\infty} z(t)u^*(t)\,dt\right|^2 \tag{21}$$

with a threshold.

The statistic in (21) can be implemented if only the spectrogram of the data is given. From property 8 in Appendix A,

$$X_{uu}(\tau, \phi)=F\{S_{uu}(t,f); \tau, \phi\}/X_{vv}(-\tau, \phi)\,, \tag{22}$$

$$X_{zz}(\tau, \phi)=F\{S_{zz}(t,f); \tau, \phi\}/X_{vv}(-\tau, \phi)\,, \tag{23}$$

where $F\{\cdot\}$ denotes a two-dimensional Fourier transform as in (A3). Equations (22) and (23) describe a deconvolution operation that can be generalized to include noisy spectrograms. This generalization will be discussed further on.

From property 14 in Appendix A, X_{uu} and X_{zz} in (22) and (23) can be used to implement the statistic in (21), since

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X_{zz}(\tau, \phi)X^*_{uu}(\tau, \phi)\,d\tau\,d\phi=\left|\int_{-\infty}^{\infty} z(t)u^*(t)\,dt\right|^2. \tag{24}$$

The detection statistic in (24) can be interpreted as *a modified form of spectrogram correlation process*. From (22) and (23), we see that (24) involves cross correlation of two *filtered* spectrograms.

A more laborious alternative is to reconstruct $z(t)\exp(j\lambda_1)$ from X_{zz} by means of property 9 in Appendix A (Siebert's separation theorem). A constant phase λ_1 is an artifact of the reconstruction process. The reconstructed data can then be correlated with $u^*(t)$ as in (21), and one obtains

$$\left|\int_{-\infty}^{\infty}[z(t)\exp(j\lambda_1)]u^*(t)\,dt\right|^2=\left|\int_{-\infty}^{\infty} z(t)u^*(t)\,dt\right|^2. \tag{25}$$

Equation (25) demonstrates that the unknown, constant phase parameter λ_1 is eliminated by envelope detecting the correlator response, and pulse compression[20] can be implemented even though λ_1 is unknown.

III. ESTIMATION OF A SIGNAL FROM ITS SPECTROGRAM

A. Generalized deconvolution

The problem is to estimate a signal from its noise-corrupted spectrogram. From property 9 in Appendix A, a simpler version of the problem is to estimate $e_{uu}(t,f)$ or $X_{uu}(\tau, \phi)$, the signal's energy density function or its ambiguity function. The spectrogram can be corrupted by noise that is external to the receiver and also by noise that is internal to the receiver.

Figure 1 shows the formulation of the problem and the assumed form of its solution. The input consists of a signal $u(t)$ plus a sample function of noise $n(t)$, and the input is represented by its energy density function or by its ambiguity function, which is the Fourier transform of its energy density function (property 3). According to property 8 in Appendix A, the Fourier transform of the data spectrogram is obtained by forming the product $X_{u+n,u+n}(\tau, \phi)X_{vv}(-\tau, \phi)$, where $X_{vv}(\tau, \phi)$ is the ambiguity function of the filter impulse response $v(t)$. The external noise $n(t)$ is characterized by the function

$$N_e(\tau, \phi)=X_{nn}(\tau, \phi)+X_{un}(\tau, \phi)+X_{nu}(\tau, \phi) \tag{26}$$

and

$$X_{u+n,u+n}(\tau, \phi)=X_{uu}(\tau, \phi)+N_e(\tau, \phi)\,. \tag{27}$$

In addition to the external noise $n(t)$, Fig. 1 shows a second, independent noise source $N_i(\tau, \phi)$ that is added to the Fourier transform of the data spectrogram. $N_i(\tau, \phi)$ represents a sample function of the internal noise, which can be associated, for example, with quantization errors in the representation of the spectrogram.

An estimating filter $H(\tau, \phi)$ is used to form a minimum mean-square error (MMSE) estimate of the input function $X_{uu}(\tau, \phi)$.

The mean-square error is

$$\text{MSE}=E\left\{\int\int |X_{uu}-[(X_{uu}+N_e)X_{vv}+N_i]H|^2\,d\tau\,d\phi\right\}. \tag{28}$$

A variational method can be used to find $H(\tau, \phi)$ such that MSE is minimized. The resulting function $H_0(\tau, \phi)$ is

$$H_0(\tau, \phi)=\frac{\overline{|X_{uu}(\tau, \phi)|^2}X^*_{vv}(-\tau, \phi)}{|X_{vv}(-\tau, \phi)|^2[|X_{uu}(\tau, \phi)|^2+P_e(\tau, \phi)]+P_i(\tau, \phi)}, \tag{29}$$

where

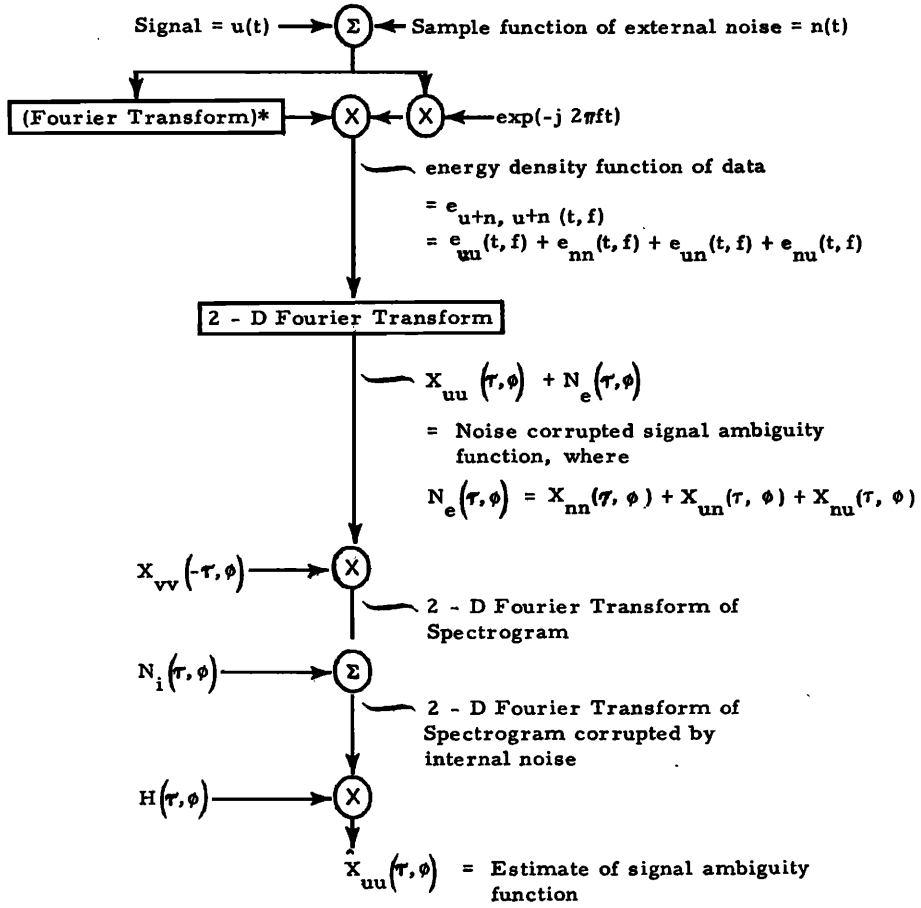

FIG. 1. An estimation problem that is encountered when a signal or its auto-ambiguity function is reconstructed from a spectrogram.

$$P_e(\tau,\phi)=\overline{|N_e(\tau,\phi)|^2},$$

$$P_i(\tau,\phi)=\overline{|N_i(\tau,\phi)|^2},$$

and $\overline{|X_{uu}(\tau,\phi)|^2}$ is an expected value of $|X_{uu}(\tau,\phi)|^2$ that is based upon accumulated knowledge of the signals that are encountered in practice.

In order to implement a locally optimum detector for a signal with random phase λ, it is only necessary to estimate $X_{ss}(\tau,\phi)$ from the data spectrogram, where

$$z(t)=u(t)\exp(j\lambda)+n(t).$$

For this estimation problem, a sample function of the external noise is part of the data that is to be reconstructed, and $N_e(\tau,\phi)=0$. A MMSE estimate $\hat{X}_{ss}(\tau,\phi)$ of the data ambiguity function is obtained with a filter

$$H_0(\tau,\phi)=\frac{\overline{|X_{zz}(\tau,\phi)|^2}X^*_{vv}(-\tau,\phi)}{\overline{|X_{zz}(\tau,\phi)|^2}|X_{vv}(-\tau,\phi)|^2+P_i(\tau,\phi)}, \quad (30)$$

where $\overline{|X_{zz}(\tau,\phi)|^2}$ is an expected magnitude-squared data ambiguity function that is based upon past observations of the data.

If $P_i(\tau,\phi)=0$ in (30), we obtain the deconvolution process in (22) and (23), but we also encounter realizability problems if $|X_{vv}(-\tau,\phi)|$ has any zeros. For a small but nonzero value of $P_i(\tau,\phi)$, realizability problems are avoided.

A generalized deconvolution filter as in (30), together with the operations in property 9 of Appendix A, yield a data estimate

$$\hat{z}(t)=z(t)\exp(j\lambda)+(\text{reconstruction error}),$$

with Fourier transform

$$\hat{Z}(f)=Z(f)\exp(j\lambda)+(\text{reconstruction error}), \quad (31)$$

where λ is an unknown constant.

If the response of *one* of the filters (say the mth) can be observed *before* envelope detection, then we can measure $Z(f)V(f-f_m)$, and the complex frequency sample $Z(f_m)$ at the center of the filter's pass band can be obtained. Substituting this frequency sample into (31) and neglecting reconstruction errors, we have

$$\hat{Z}(f_m)/Z(f_m)=\exp(j\lambda), \quad (32)$$

and we can solve for the unknown, constant phase parameter λ. By observing the phase, as well as the amplitude, at the output of a *single filter*, we can deduce the value of the parameter λ that is added to the phase of the reconstructed spectrum $\hat{Z}(f)$. Knowledge of λ allows us to *completely* determine $z(t)$ from its spectrogram. It is only required that $|Z(f_m)|$ not be zero in (32), i.e., a phase measurement should be made at the output of a filter that lies within the pass band of the data. Since the data usually consists of a sample function of broadband noise along with a signal, the condition $|Z(f_m)|\neq 0$ is not very restrictive.

B. Estimator performance for waveform estimation via deconvolution

Substituting $H_0(\tau, \phi)$ from (29) back into (28), we obtain

$$\mathrm{MSE}_0 = \int\int \overline{|X_{uu}|^2}\left(1 - \frac{\overline{|X_{uu}|^2}}{|X_{uu}|^2 + P_e + (P_t/|X_{vv}|^2)}\right) d\tau\, d\phi\,. \tag{33}$$

If both P_e and P_t are zero, (29) is identical to (A13), $e_{uu}(t,f)$ is obtained by deconvolution, and $\mathrm{MSE}_0 = 0$. If P_t is zero but P_e is not, MSE_0 in (33) is independent of the filter impulse response $v(t)$. *For no internal noise, any convenient filter function can be used, provided that $H_0(\tau, \phi)$ is bounded.* This result is surprising, since it implies that the time and frequency resolution of the reconstructed signal energy density function $e_{uu}(t,f)$ is independent of the bandwidth of the analyzing filters! If there is no internal noise, spectrograms that are obtained with octave-wide filters yield the same mean-square error as spectrograms that are obtained with very narrow-band filters, even for narrow-band signals! Without frequency interpolation, this result holds only if (16) is satisfied, and T_u will be large for narrow-band signals. Although wide-band filters can be used, their center frequencies must be closely spaced, and the resulting transfer functions will overlap.

C. Estimation by prediction-subtraction

Spectrogram samples for a particular time-window position can be partially predicted by using prior knowledge of spectrogram behavior, along with spectrogram samples from previous time windows. The difference between spectral energy samples for a given time-window position and predicted spectral energy samples (based upon samples from previous window positions) can be used to form spectrogram innovations.[21] The innovations can be used to sequentially estimate or "track" the spectrograms (or autoambiguity functions, or covariance functions) of signal and noise. If only the noise is tracked, a correct noise prediction produces an innovation that subtracts the noise contribution from the data spectrogram. The result is a sequential version of the spectral noise subtractor suggested by Boll.[22] If the noise prediction is computed from a linear combination of past spectrogram samples, the estimator is similar to the adaptive noise canceler of Widrow *et al.*[23] After noise is reduced by the prediction–subtraction process, the signal and its ambiguity function can be reconstructed by deconvolution.

IV. ESTIMATION OF SIGNAL PARAMETERS

A. Maximum likelihood estimation with matched filters followed by envelope detectors

The parameter estimation problem involves a known signal that depends upon unknown parameters. For example, $u(t+\tau)$ specifies the dependence of a known signal $u(t)$ upon a delay parameter τ, and $u(t)\exp(j2\pi\phi t)$ describes a known signal $u(t)$ that has been frequency shifted by a Doppler parameter ϕ. For unknown delay τ, frequency shift ϕ, and phase shift λ, the data $z(t)$ can be written

$$z(t) = u(t+\tau)\exp[j(2\pi\phi t + \lambda)] + n(t)\,, \tag{34}$$

where $n(t)$ is white, Gaussian noise. A vector of data samples is denoted $\boldsymbol{\zeta}$.

Maximum likelihood (ML) estimates of delay, τ and Doppler shift, ϕ are obtained by finding the values of τ_H and ϕ_H that maximize the likelihood ratio $\Lambda(\boldsymbol{\zeta}\,|\,\tau_H, \phi_H)$. When a nuisance parameter λ has been added to the phase of an analytic signal as in (34), the likelihood ratio is (Appendix C)

$$\Lambda(\boldsymbol{\zeta}\,|\,\tau_H, \phi_H) = e^{-E_u/2\sigma^2} I_0(|X_{zu}(\tau_H, \phi_H)|/\sigma^2)\,. \tag{35}$$

Since the modified Bessel function $I_0(\cdot)$ is a monotone increasing function of its argument, maximum likelihood estimates are the values of τ_H and ϕ_H that maximize

$$|X_{zu}(\tau_H, \phi_H)|^2 = \left|\int_{-\infty}^{\infty} z(t)u^*(t+\tau_H)e^{-j2\pi\phi_H t}\,dt\right|^2, \tag{36}$$

where $z(t)$ is the observed data (signal plus noise). The ML estimator can be implemented by passing the data through a bank of matched filters with transfer functions $U^*(f-\phi_H)$. The envelope detected outputs of these filters are analyzed in order to find a maximum.

The asymptotic variance of a ML estimate of τ or ϕ, for many independent observations of the data, can be obtained from the Cramér–Rao (CR) bound.[24,25] It can be shown that, for *small* SNR,

$$\begin{aligned} \Delta\tau &\sim \mathrm{SNR}^{-1} B_u^{-1} \\ \Delta\phi &\sim \mathrm{SNR}^{-1} T_u^{-1} \end{aligned} \tag{37}$$

where $\Delta\tau$ and $\Delta\phi$ are the standard deviations of ML delay and Doppler (or center frequency) estimates, respectively.[26]

Bounds for *large* SNR have been worked out by Kelly[27] and Kelly, Reed, and Root.[28] The resulting standard deviations are

$$\begin{aligned} \Delta\tau &\sim \mathrm{SNR}^{-1/2} B_u^{-1} \\ \Delta\phi &\sim \mathrm{SNR}^{-1/2} T_u^{-1}\,. \end{aligned} \tag{38}$$

For a bank of matched filters with envelope detected outputs, both large SNR and small SNR results indicate that

$$\begin{aligned} \Delta\tau &\sim B_u^{-1} \\ \Delta\phi &\sim T_u^{-1}\,. \end{aligned} \tag{39}$$

B. Parameter estimation with a spectrogram correlator

From (C5), (C6), and (C9) in Appendix C, the log-likelihood ratio for a spectrogram representation of $z(t)$ in (34) is

$$\ln\Lambda(\mathbf{z}\,|\,\mathbf{s}) \approx \sum_{m=1}^{M}\sum_{n=1}^{N}\left[-s_{mn}/\sigma^2 + z_{mn}s_{mn}/\sigma^4\right], \tag{40}$$

where the approximation holds for very small SNR. In (40), $\mathbf{s}$ is an array of spectrogram samples corresponding to the noise-free data, and $\mathbf{z} = \{z_{mn}\}$ is an array of

samples from the data spectrogram, $S_{zz}(t,f)$.

A ML estimate of a parameter ρ can be obtained by writing

$$\mathbf{s}(\rho)=\{s_{mn}(\rho);\, m=1,\ldots,M;\, n=1,\ldots,N\}, \tag{41}$$

where $s_{mn}(\rho)$ is a sample of the spectrogram of a signal $u(t,\rho)$. If ρ is the delay parameter τ, for example, then $s_{mn}(\rho)$ is a sample of the spectrogram of $u(t+\tau)$. The ML estimate of ρ is then the hypothesized parameter ρ_H that maximizes

$$\ln\Lambda[\mathbf{z}\,|\,\mathbf{s}(\rho_H)]\approx\sum_{m=1}^{M}\sum_{n=1}^{N}[-s_{mn}(\rho_H)/\sigma^2+z_{mn}s_{mn}(\rho_H)/\sigma^4]. \tag{42}$$

From property 6 in Appendix A, $\sum_{m,n}s_{mn}(\rho_H)$ is proportional to E_uE_v. The filter energy E_v does not vary with ρ_H, and neither does the signal energy E_u when ρ_H is delay or frequency shift. The first term in (42) can therefore be ignored, and *the ML estimator for small SNR is a spectrogram correlator*, i.e., a device that correlates the data spectrogram $\{z_{mn}\}$ with a sequence of hypothesized noise-free signal spectrograms, $\{s_{mn}(\rho_H)\}$.

For a large number of observations, the variance of a ML estimate approaches the Cramér–Rao (CR) lower bound.[24,25] A ML estimate for low SNR data is implemented by spectrogram correlation. The asymptotic accuracy of parameter estimation by means of a spectrogram correlator is therefore given by a CR bound that is computed for low SNR data. For an estimate $\hat{\rho}$ of a parameter ρ, the CR bound is

$$\mathrm{Var}(\rho-\hat{\rho})\geq\{-E[(\partial^2/\partial\rho^2)\ln\Lambda[\mathbf{z}\,|\,\mathbf{s}(\rho)]]\}^{-1}, \tag{43}$$

evaluated at the true value of ρ.

The asymptotic variances of ML spectrogram estimates of delay and center frequency for small SNR can be calculated from (43).[26] For a broadband pulse with $T_u\leq T_v$ and $B_u\geq B_v$,

$$\Delta_\tau\sim[B_uB_v]^{-1/2}. \tag{44}$$

For a narrow-band sinusoid with $B_u\leq B_v$ and $T_u\geq T_v$,

$$\Delta f\sim[T_uT_v]^{-1/2}. \tag{45}$$

These standard deviations agree with $\Delta\tau$ and Δf for matched filters followed by envelope detectors. It has already been remarked that a bank of matched filters followed by envelope detectors creates a special form of spectrogram, with $v(t)=u^*(-t)$, as in property 4 of Appendix A. If $v(t)=u^*(-t)$, then $B_v=B_u$ and $T_v=T_u$. In this case, (44) and (45) become identical with (39).

Since the standard deviations in (39) are lower bounds for those in (44) and (45), we conclude that the best estimates of delay and frequency shift for a known signal are obtained when spectrogram filters or window functions are matched to the signal.

V. CLASSIFICATION OF SIGNALS FROM THEIR SPECTROGRAMS

A classifier generally tests more than two hypotheses. Letting H_0 be the hypothesis that only noise is present, K different signal types with additive noise can be represented by $\{H_k\}_{k=1}^K$. If the kth hypothesis has a corresponding noise-free spectrogram $\mathbf{s}^{(k)}$, then the most likely hypothesis corresponds to

$$\max_{1\leq k\leq K}[\ln\Lambda(\mathbf{z}\,|\,\mathbf{s}^{(k)})],$$

where $\ln\Lambda(\mathbf{z}\,|\,\mathbf{s})$ is given by (40) when SNR is small. The most likely hypothesis is chosen by a maximum likelihood classifier, which minimizes probability of misclassification.[29] Comparison of

$$\max_{1\leq k\leq K}[\ln\Lambda(\mathbf{z}\,|\,\mathbf{s}^{(k)})]$$

with a threshold determines whether H_0 is actually the most likely hypothesis. For energy-normalized reference signals, (A8) and (40) imply that a parallel connection of spectrogram correlators is a locally optimum, maximum likelihood classifier for a spectrogram representation of the data. The maximum correlator response is compared with a threshold.

The locally optimum classifier configuration suggests the existence of a strong interconnection between detectors, estimators, and classifiers. The *classifier* is a parallel connection of *detectors*, where each detector searches for a different waveform, signal covariance function, or target/channel scattering function. A *detector* is a two-hypothesis *classifier*, where one of the hypotheses is H_0. The estimator computes $\ln\Lambda[\mathbf{z}\,|\,\mathbf{s}(\rho_H)]$ for an unknown, hypothesized signal parameter ρ_H. K different parameter hypotheses correspond to K different spectrograms, i.e., $\{\mathbf{s}(\rho_H^{(k)})\}_{k=1}^K\equiv\{\mathbf{s}^{(k)}\}_{k=1}^K$. The maximum likelihood *estimator* is then a special kind of *classifier*.

Estimators and detectors are thus specific examples of signal classifiers. This relation implies that simplified detector and estimator configurations may be obtained by using appropriate feature sets from the pattern recognition literature.[30-32] An example of a simplified spectrogram correlator is given by Altes and Faust.[33]

The relation between parameter estimators and classifiers also implies that a classifier can be evaluated in terms of a generalized ambiguity function.[34] Classifier ambiguity peaks are associated with large correlation between time-shifted reference spectrograms.

VI. SUMMARY AND CONCLUSION

Properties of spectrograms and related functions indicate that a signal can be reconstructed from its spectrogram, except that the reconstructed signal is multiplied by a factor with unit magnitude and unknown, constant phase. If the zero crossing information at the output of any one filter can be observed, however, then the input data can be completely reconstructed, provided that the data (signal plus noise) has energy in the frequency band of the filter that takes the phase measurement.

Information about the covariance function of a random process is preserved in a spectrogram representation, and the covariance function can be obtained from an ensemble-average spectrogram. This observation implies that optimum detection of random signals with

known covariance can be accomplished with a spectrogram representation of the data. A locally optimum spectrogram detector for random signals in Gaussian noise is a spectrogram correlator, and this detector can easily be adjusted for non-Gaussian noise backgrounds.

Spectrogram correlation can be used to detect signals that have been passed through a time-varying channel or reflected from a time-varying target, if the scattering function of the channel or target is known. Spectrogram correlation is also applicable to locally optimum detection of signals that have been passed through a random, time *invariant* filter, provided the expected energy density function of the filter impulse response is known.

Locally optimum detection of signals that are known except for Doppler shift (or center frequency), time of arrival, and phase, can be accomplished with a bank of matched filters. A weighted sum of envelope detected filter outputs is formed, where the weights are determined in accordance with prior information about expected time of arrival and center frequency. The resulting receiver is a special form of spectrogram correlator, and it can again be easily adjusted to accommodate non-Gaussian noise.

Since a bank of matched filters is a special form of time-window spectrogram, there is a simple digital implementation for this receiver configuration. Input data is multiplied by a shifted window function that is identical to the conjugated reference signal, and the magnitude-squared discrete Fourier transform of the product is computed.

A spectrogram representation can be used to detect signals that are known except for phase. Although reconstructed data can be used for this detection operation, it is easier to simply reconstruct the data ambiguity function by deconvolution, and to integrate the product of the data ambiguity function and the conjugated ambiguity function of the reference signal. The resulting process is again a spectrogram correlator, where filtered (deconvolved) spectrograms are used.

A signal can be estimated from its spectrogram when noise has been added to the signal (external noise) and to the spectrogram (internal noise). A minimum mean-square error (MMSE) estimate of the signal ambiguity function is obtained by using a modified deconvolution filter. An estimate of the signal (except for a constant phase parameter) can then be obtained from the estimated ambiguity function. The deconvolution process can be augmented by using prediction–subtraction to cancel predictable spectral energy due to noise or nonstationary interference.

If internal noise power is very small, the mean-square error in the reconstructed signal ambiguity function is insensitive to the filter impulse response or window function that is used to construct the spectrogram. This result implies that wide filter bandwidths are acceptable even for reconstruction of narrow-band signals. For wide filter bandwidths, the required sampling rates for signal reconstruction will lead to considerable filter overlap.

Signal parameter estimates such as time of arrival or center frequency can be obtained by using a bank of matched filters or a spectrogram correlator. For a matched filter implementation, the minimum standard deviations of maximum likelihood delay and frequency estimates are $\Delta_\tau \sim B_u^{-1}$ and $\Delta f \sim T_u^{-1}$, respectively, where B_u and T_u are the approximate bandwidth and duration of the signal. For a spectrogram that is constructed from filters with bandwidth B_v and duration T_v, $\Delta_\tau \sim (B_v B_u)^{-1/2}$ and $\Delta f \sim (T_v T_u)^{-1/2}$ for a broadband pulse ($B_u \geq B_v$) and a narrow-band sinusoid ($T_u \geq T_v$), respectively. Although the error in the reconstructed waveform (and in parameters that are obtained from this waveform) is independent of the window function for zero internal noise, the error in parameters that are obtained directly from a spectrogram is window-dependent.

The most interesting result from a detection theory viewpoint is that a spectrogram correlator is a locally optimum detector that can take its place alongside the better-known Karhunen–Loève and estimator–correlator configurations for detection of random signals.[35] The spectrogram correlator is in fact an attractive alternative to the better-known configurations because there is no eigenfunction equation to solve, the detector is easily implemented by using a fast Fourier transform algorithm, and it can easily be adjusted for non-Gaussian noise by using the results of Poor and Thomas.[19]

The most interesting result from a signal theory viewpoint is that very little information about a signal is lost in a spectrogram representation. In fact, if zero crossing information is obtained from only one filter, the signal can be completely reconstructed from its spectrogram. The reconstruction method may be useful for speech synthesis from vocoders that transmit speech spectrograms,[6] although less elaborate reconstruction methods can still yield intelligible speech.[22]

Spectrograms have traditionally been constructed with window functions that have small time–bandwidth product. A bank of matched filters for a large time–bandwidth product signal, however, can easily be synthesized with a time window spectrogram, if the complex, analytic window function is a conjugated version of the signal. The most accurate delay and Doppler estimates are also obtained in this case. Properties 11 and 12 in Appendix A seem to indicate some additional advantages that may accrue from using coded or modulated windows that have thumbtack ambiguity functions.

Detectors and parameter estimators are special kinds of classifiers. Simplified spectrogram correlators for detection and estimation can then be obtained by using appropriate feature sets. Ambiguity concepts can be used to analyze classifiers as well as estimators.

APPENDIX A: PROPERTIES OF SPECTROGRAMS AND RELATED FUNCTIONS

Some properties of time–frequency energy density functions, spectrograms, and ambiguity functions are listed below. Where limits on integrals are omitted, the limits are $(-\infty, \infty)$.

Property 1[18]

$$S_{uv}(t_1,f_1)=e_{uu}(t,f)\overset{t,f}{*}e_{vv}(t,-f), \tag{A1}$$

where $\overset{t,f}{(*)}$ denotes two-dimensional convolution, i.e.,

$$S_{uv}(t_1,f_1)=\int\int e_{uu}(t,f)e_{vv}(t_1-t,f-f_1)\,dt\,df.$$

Property 2[18]

$$|X_{uv}(-\tau,\phi)|^2=e_{uu}(t,f)\overset{t,f}{*}e^*_{vv}(-t,-f)$$
$$=\int\int e_{uu}(t,f)e^*_{vv}(t-\tau,f-\phi)\,dt\,df. \tag{A2}$$

Property 3[36,16]

$$F\{e_{uv}(t,f);\tau,\phi\}=X_{uv}(\tau,\phi), \tag{A3}$$

where $F\{e_{uv}(t,f);\tau,\phi\}$ denotes a two-dimensional Fourier transform that maps $e_{uv}(t,f)$ onto the τ,ϕ plane, i.e.,

$$\int\int e_{uv}(t,f)e^{-j2\pi(\phi t+\tau f)}\,dt\,df=X_{uv}(\tau,\phi).$$

Property 4

From (6)–(9),

$$|X_{uv}(t_1,f_1)|^2=S_{u(t)v^*(-t)}(-t_1,f_1) \tag{A4}$$

$$S_{uv}(t_1,f_1)=|X_{u(t)v^*(-t)}(-t_1,f_1)|^2. \tag{A5}$$

Since many properties of the magnitude-squared ambiguity function have been derived,[36,20] the above identity leads to similar properties of the spectrogram. Properties 5 to 7, for example, are shared by $S_{uv}(t,f)$ and $|X_{uv}(t,f)|^2$.

Property 5

Smeared energy density spectra and time envelopes can be obtained from marginals of the spectrogram, i.e.,

$$\int S_{uv}(t_1,f_1)\,dt_1=\int |U(f)|^2|V(f-f_1)|^2\,df \tag{A6}$$

$$\int S_{uv}(t_1,f_1)\,df_1=\int |u(t)|^2|v(t_1-t)|^2\,dt. \tag{A7}$$

These relations are verified by substitution of (8) into the left-hand side of (A6), and by substituting (9) into the left-hand side of (A7).

Property 6

$$\int\int S_{uv}(t_1,f_1)\,dt_1\,df_1=E_uE_v, \tag{A8}$$

where

$$E_u\equiv\int |U(f)|^2\,df=\int |u(t)|^2\,dt$$

$$E_v\equiv\int |V(f)|^2\,df=\int |v(t)|^2\,dt. \tag{A9}$$

This property states that the volume under the spectrogram depends strictly upon the energy of the signal and filter functions. The volume is independent of the form of these functions. To verify (A8), integrate (A6) with respect to f_1, or integrate (A7) with respect to t_1.

Property 7

$$S_{uv}(t_1,f_1)\le E_uE_v \tag{A10}$$

with equality if and only if there exist values of t_1 and f_1 such that

$$V(f-f_1)\exp(j2\pi ft_1)=U^*(f). \tag{A11}$$

The magnitude of the spectrogram never exceeds the product E_uE_v, and the upper bound is only obtained if one of the filters that is used to form the spectrogram is matched to the signal. Equation (A10) is obtained by applying the Schwarz inequality to $S_{uv}(t_1,f_1)$, as defined by (8).

Property 8

$$F\{S_{uv}(t,f);\tau,\phi\}=X_{uu}(\tau,\phi)X_{vv}(-\tau,\phi) \tag{A12}$$

or

$$X_{uu}(\tau,\phi)=F\{S_{uv}(t,f);\tau,\phi\}/X_{vv}(-\tau,\phi). \tag{A13}$$

Equations (A12) and (A13) follow directly from (A1) and (A3). Equation (A13) implies that, in the absence of noise, $X_{uu}(\tau,\phi)$ can be obtained from a spectrogram, if the filter function $v(t)$ is known. The significance of this result is enhanced by the following property and by property 14.

Property 9[37,38]

$$X_{u_1u_1}(\tau,\phi)=X_{u_2u_2}(\tau,\phi)$$

if, and only if,

$$u_1(t)=u_2(t)\exp(j\lambda),$$

where λ is real. This observation was originally derived from the separation equation

$$\int X_{uu}(\tau-t,\phi)\exp(j2\pi t\phi)\,d\phi=u(t)u^*(\tau). \tag{A14}$$

The result can also be derived from property 3 by taking $F\{X_{uu}(\tau,\phi);t,f\}=e_{uu}(t,f)$, and by forming

$$\int e_{uu}(t,f)\exp[j2\pi f(t-\tau)]\,df=u(t)u^*(\tau). \tag{A15}$$

In either case, $|u(\tau_0)|$ is obtained by setting $t=\tau=\tau_0$, where τ_0 is a constant. Then $u(t)[u^*(\tau_0)/|u(\tau_0)|]$ is the desired waveform $u(t)$, multiplied by a complex constant with unity magnitude.

It follows that one can obtain $u(t)\exp(j\lambda)$ from $X_{uu}(\tau,\phi)$ or $e_{uu}(t,f)$, where λ is a real, unknown constant.

Property 10[36]

The counterpart of property 8 for the magnitude-squared ambiguity function is

$$F\{|X_{uu}(-\tau,\phi)|^2;t,f\}=X_{uu}(t,f)X_{vv}^*(t,f), \tag{A16}$$

i.e.,

$$\int\int |X_{uu}(-\tau,\phi)|^2 e^{-j2\pi(f\tau+t\phi)}\,d\tau\,d\phi = X_{uu}(t,f)X_{vv}^*(t,f).$$

In particular,

$$F\{|X_{uu}(-\tau,\phi)|^2;t,f\}=|X_{uu}(t,f)|^2,$$

i.e., the magnitude-squared autoambiguity function is its own Fourier transform. Equation (A2) can be obtained directly from (A3) and (A16) by using the fact that multiplication of two functions is the same as convolution of their Fourier transforms.

Property 11

Multiplying both sides of (A12) by $X_{vv}^*(-\tau,\phi)$, we have

$$F\{S_{uv}(t,f);\tau,\phi\}X_{vv}^*(-\tau,\phi)=X_{uu}(\tau,\phi)|X_{vv}(-\tau,\phi)|^2.$$

Using property 10 and (A3),

$$S_{uv}(t,f)\overset{t,f}{*}e_{vv}^*(-t,f)=e_{uu}(t,f)\overset{t,f}{*}|X_{vv}(t,f)|^2. \tag{A17}$$

The above relation can also be obtained from (A1) and (A2), and this alternate derivation leads to a generalized version of (A17):

$$S_{u_1v_1}(t,f)\overset{t,f}{*}e_{v_2v_2}^*(-t,f)=e_{u_1u_1}(t,f)\overset{t,f}{*}|X_{v_1v_2}(t,f)|^2$$

or

$$\int\int S_{u_1v_1}(t,f)e_{v_2v_2}^*(t-\tau,\phi-f)\,dt\,df = \int\int e_{u_1u_1}(t,f)|X_{v_1v_2}(\tau-t,\phi-f)|^2\,dt\,df. \tag{A18}$$

These results suggest that the energy density function of a signal can be estimated from the spectrogram of the signal without the deconvolution operation in (A13). From (A17), $e_{uu}(t,f)$ can be approximated without a deconvolution process if $|X_{vv}(t,f)|^2$ resembles a thumbtack. The time–bandwidth product of $v(t)$ must be large in order to obtain a thumbtack ambiguity function, and the operation in (A17) is a two-dimensional pulse compression process.

Property 12

Two different short-time spectral histories can be compared by means of the cross-correlation operation The above sequence of equalities made use of (A1) and (A2). Equation (A19) is similar to a result for wide-band *auto*ambiguity functions that was published by Flaska,[39] with credit to C. E. Persons.

$$\int\int S_{u_1v_1}(t,f)S_{u_2v_2}(t+\tau,f+\phi)\,dt\,df = S_{u_1v_1}^*(-t,-f)\overset{t,f}{*}S_{u_2v_2}(t,f)=e_{u_1u_1}^*(-t,-f)\overset{t,f}{*}e_{v_1v_1}^*(-t,f)\overset{t,f}{*}e_{u_2u_2}(t,f)\overset{t,f}{*}e_{v_2v_2}(t,-f)$$

$$=|X_{u_2u_1}(-t,f)|^2\overset{t,f}{*}|X_{v_2v_1}(t,-f)|^2=\int\int |X_{u_2u_1}(t,f)|^2|X_{v_2v_1}(\tau+t,f-\phi)|^2\,dt\,df. \tag{A19}$$

If $|X_{v_2v_1}(t,f)|^2$ is a thumbtack, then two-dimensional cross correlation of the two spectrograms in (A19) is approximately equivalent to forming the magnitude-squared cross-ambiguity function of $u_1(t)$ and $u_2(t)$. If $u_1(t)$ is input data and $u_2(t)$ is a reference signal, (C22) indicates that the cross-ambiguity function is part of a locally optimum detection statistic for a signal with unknown delay and Doppler shift, and with random phase. Spectrogram cross correlation as in (A19) can also be used to estimate delay and Doppler parameters from spectrograms, as indicated by (42).

Property 13

Another way to compare two different signals is to integrate the product of their time–frequency energy density functions. By using (6) and Parseval's theorem, it is easy to prove that

$$\int\int e_{u_1v_1}(t,f)e_{u_2v_2}^*(t,f)\,dt\,df = \int u_1(t)u_2^*(t)\,dt\int v_1^*(t)v_2(t)\,dt. \tag{A20}$$

An important special case of (A20) is obtained by letting $v_1(t)=u_1(t)$ and $v_2(t)=u_2(t)=u_1(t-\tau)\exp(j2\pi\phi t)$. In this case, (A20) becomes

$$\int\int e_{u_1u_1}(t,f)e_{u_2u_2}^*(t,f)\,dt\,df=|X_{u_1u_1}(\tau,\phi)|^2. \tag{A21}$$

Property 14

By using property 3 and applying Parseval's theorem, (A20) can be written

$$\int\int X_{u_1v_1}(\tau,\phi)X_{u_2v_2}^*(\tau,\phi)\,d\tau\,d\phi=\int u_1(t)u_2^*(t)\,dt\int v_1^*(t)v_2(t)\,dt. \tag{A22}$$

The above equation can also be verified by direct substitution of (7) into the left-hand side. If $v_1(t)=u_1(t)$ and $v_2(t)=u_2(t)=u_1(t-\tau_1)\exp(j2\pi\phi_1 t)$, we have

$$\int\int X_{u_1u_1}(\tau,\phi)X_{u_2u_2}^*(\tau,\phi)\,d\tau\,d\phi=|X_{u_1u_1}(\tau_1,\phi_1)|^2.$$

Similarly, if $v_1(t)=u_1(t)$ and $v_2(t)=u_2(t)$ in (A22), we have

$$\int\int X_{u_1u_1}(\tau,\phi)X_{u_2u_2}^*(\tau,\phi)\,d\tau\,d\phi=\left|\int u_1(t)u_2^*(t)\,dt\right|^2. \tag{A23}$$

The significance of (A23) is that two signals can be correlated by operating upon their autoambiguity functions, rather than upon the signals themselves. Equation (A23) can also be obtained from Stutt's Fourier transform equation, (A16).

Property 15[18]

The time-window spectrogram, $S_{uv}^T(t_1,f_1)$, is obtained by multiplying the data $u(t)$ by a delayed window func-

tion $w(t-t_1)$ and by taking the Fourier transform of the resulting product, i.e.,

$$S^T_{uw}(t_1,f_1)=\left|\int u(t)w(t-t_1)\exp(-j2\pi f_1 t)\,dt\right|^2. \quad \text{(A24)}$$

By changing variables in (9), the frequency-window (bank-of-filters) spectrogram can be written

$$S_{uv}(t_1,f_1)=\left|\int_{-\infty}^{\infty} u(t)v(t_1-t)\exp(-j2\pi f_1 t)\,dt\right|^2. \quad \text{(A25)}$$

Comparison of (A24) and (A25) indicates that

$$S^T_{uw}(t_1,f_1)=S_{uv}(t_1,f_1)$$

if

$$w(t)=v(-t), \quad \text{(A26)}$$

where $v(t)$ is the impulse response of the basic filter function that is used to construct the frequency-window spectrogram.

Property 16

Let $\overline{S_{uv}(t,f)}$ denote an ensemble average spectrogram. Knowledge of $\overline{S_{uv}(t,f)}$ is equivalent to knowledge of the autocorrelation function $\overline{u(t)u^*(t+\tau)}$, where $u(t)$ is a random process.

This statement is proven by using (A12):

$$F\{\overline{S_{uv}(t,f)};\tau,\phi\}=\overline{X_{uu}(\tau,\phi)}X_{vv}(-\tau,\phi), \quad \text{(A27)}$$

where, from (5),

$$\overline{X_{uu}(\tau,\phi)}=\int_{-\infty}^{\infty}\overline{u(t)u^*(t+\tau)}\exp(-j2\pi\phi t)\,dt. \quad \text{(A28)}$$

If the autocovariance function of the random process $u(t)$,

$$R_{uu}(t,\tau)=\overline{u(t)u^*(t+\tau)} \quad \text{(A29)}$$

is specified, then the expected spectrogram $\overline{S_{uv}(t,f)}$ can be determined by using (A27) and (A28). Similarly, if $\overline{S_{uv}(t,f)}$ is specified, then (A27) gives $\overline{X_{uu}(\tau,\phi)}$, and from (A28),

$$R_{uu}(t,\tau)=\int_{-\infty}^{\infty}\overline{X_{uu}(\tau,\phi)}\exp(j2\pi\phi t)\,d\phi. \quad \text{(A30)}$$

Property 17

Let $z(t)$ be the output of a time-varying random channel with input $u(t)$. Then the expected spectrogram of the channel output is

$$\overline{S_{zv}(t,f)}=\Gamma_h(t,f)\overset{t,f}{*}S_{uv}(t,f), \quad \text{(A31)}$$

where $\Gamma_h(t,f)$ is the channel scattering function.

To prove (A31), we start with the input–output relation for a wide-sense stationary random channel[8,9]:

$$\psi_{zz}(\tau,\phi)=R_H(\tau,\phi)\psi_{uu}(\tau,\phi), \quad \text{(A32)}$$

where

$$R_H(\tau,\phi)=\overline{H^*(t,f)H(t+\tau,f+\phi)},$$

for a channel with time-varying transfer function $H(t,f)$,

$$\psi_{zz}(\tau,\phi)=\int\overline{z^*(t-\tau/2)z(t+\tau/2)}e^{-j2\pi\phi t}dt,$$
$$\psi_{uu}(\tau,\phi)=\int u^*(t-\tau/2)u(t+\tau/2)e^{-j2\pi\phi t}dt. \quad \text{(A33)}$$

The symmetrized ambiguity functions in (A33) can be related to the asymmetric version in (5) by a change of variable;

$$\psi_{uu}(\tau,\phi)=X_{uu}(-\tau,\phi)\exp(j\pi\phi\tau). \quad \text{(A34)}$$

Multiplying both sides of (A32) by $\exp(-j\pi\phi\tau)$, we have

$$\overline{X_{zz}(-\tau,\phi)}=R_H(\tau,\phi)X_{uu}(-\tau,\phi). \quad \text{(A35)}$$

Using the two-dimensional Fourier transform as defined in property 3, along with (A3), we have

$$\overline{e_{zz}(t,-f)}=F^{-1}\{R_H(\tau,\phi);t,f\}\overset{t,f}{*}e_{uu}(t,-f). \quad \text{(A36)}$$

The scattering function $\Gamma_h(t,f)$ is defined by Bello[8] to be

$$\Gamma_h(t,f)=\int\int R_H(\tau,\phi)\exp[j2\pi(t\phi-f\tau)]\,d\tau\,d\phi. \quad \text{(A37)}$$

It follows that

$$\Gamma_h(t,-f)=F^{-1}\{R_H(\tau,\phi);t,f\}, \quad \text{(A38)}$$

and

$$\overline{e_{zz}(t,f)}=\Gamma_h(t,f)\overset{t,f}{*}e_{uu}(t,f). \quad \text{(A39)}$$

Convolving both sides of (A39) with $e_{vv}(t,-f)$ and using (A1), we obtain (A31), which is the desired result. Other versions of (A31), written in terms of magnitude-squared ambiguity functions, are found in Stewart and Westerfield,[40] Gaarder,[41] and Rihaczek.[16]

Property 18

Let $z(t)$ be the output of a time *invariant* random filter with input $u(t)$ and impulse response $h(t)$. Then

$$\overline{S_{zv}(t,f)}=\int\overline{e_{hh}(t,\alpha)}\overset{t}{*}e_{uu}(t,\alpha)\overset{t}{*}e_{vv}(t,\alpha-f)\,d\alpha, \quad \text{(A40)}$$

where $\overset{t}{*}$ is a *one*-dimensional convolution in the time variable only.

To prove (A40), we start with

$$z(t)=u(t)\overset{t}{*}h(t)=\int u(x)h(t-x)\,dx \quad \text{(A41)}$$

and we compute

$$\overline{e_{zz}(t,f)}=\overline{\left(\int u(x)h(t-x)\,dx\right)[U(f)H(f)]^*\exp(-j2\pi ft)}$$
$$=\int u(x)U^*(f)e^{-j2\pi xf}\overline{[h(t-x)H^*(f)e^{-j2\pi(t-x)f}]}\,dx$$
$$=e_{uu}(t,f)\overset{t}{*}\overline{e_{hh}(t,f)}. \quad \text{(A42)}$$

Convolving both sides of (A42) with $e_{vv}(t,-f)$ and using (A1), we obtain (A40).

APPENDIX B: SAMPLING THE SPECTROGRAM

1. Minimum sampling rates in time and frequency directions

Let each filter $V(f-f_1)$ of a frequency window spectrogram have one-sided bandwidth B_v. The amplitude and phase of the filter output can then each be sampled every B_v^{-1} s.[17] The squared envelope of the filter output

has twice the bandwidth of the envelope itself, and it must therefore be sampled every $(2B_v)^{-1}$ s. Spectrogram samples should thus be

$$\Delta t = (2B_v)^{-1} \text{ s} \tag{B1}$$

apart in the time direction.

A similar argument can be used to determine the minimum sampling rate for a constant-t_1 profile of $S_{ww}(t_1, f_1)$. It must be assumed, however, that the filter impulse response $v(t)$ is time limited as well as band limited, and this situation cannot occur. An alternate assumption is that a function is adequately specified if it can be reconstructed to within a given small error ϵ, and that, for any filter of interest, there exist sampling densities in time and frequency such that ϵ is not exceeded. In other words, we assume the existence of numbers B_v and T_v such that $v(t)$ is adequately specified by samples that are $1/(2B_v)$ apart and $V(f)$ is adequately specified by samples that are $1/(2T_v)$ apart. B_v has been called the bandwidth of $v(t)$, and T_v will be called the duration of $v(t)$.

Frequency domain sampling rate is determined from the duration of the corresponding time function. In (9), shifted input data is multiplied by $v(t)$, and the maximum time width of the windowed data is T_v. The squared envelope of the Fourier transform of the windowed data can then be specified with samples that are

$$\Delta f = (2T_v)^{-1} \text{ Hz} \tag{B2}$$

apart in the frequency direction.

2. Sampling rates for two-dimensional deconvolution

Property 1 in Appendix A states that the spectrogram can be formed by convolution of $e_{uu}(t,f)$ with $e_{vv}(t,f)$. From the definition of the energy density function (4), $e_{uu}(t,f)$ and $e_{vv}(t,f)$ can both be adequately represented by using a sampling rate in time that is at least $2\max(B_w, B_v)$ samples/s and a sampling rate in frequency that is at least $2\max(T_w, T_v)$ samples/Hz. These sampling rates are therefore required for two-dimensional digital convolution of $e_{uu}(t,f)$ and $e_{vv}(t,f)$. The output of the convolution process will be sampled at the same rates as above, and these rates will often be much larger than the minimum required sampling rates for the spectrogram ($2B_v$ samples/s and $2T_v$ samples/Hz). High sampling rates are also required if $S_{ww}(t,f)$ is to be subjected to a digital deconvolution operation in order to separate $e_{uu}(t,f)$ from $e_{vv}(t,f)$. The estimate of $e_{uu}(t,f)$ after deconvolution may be undersampled unless the spectrogram is sampled at $2\max(B_w, B_v)$ samples/s and $2\max(T_w, T_v)$ samples/Hz.

One way to obtain the relatively high sampling rates that may be required for deconvolution is to interpolate between the more sparse samples that are required for efficient representation of the spectrogram. Another method is to construct a spectrogram that is deliberately oversampled. Physically, a given frequency sample is obtained by using a given value of f_1 in (8), i.e., by using a filter with the specified center frequency. If frequency interpolation is not to be used, the filter center frequencies should be separated by at most

$$\Delta f_1 = [2\max(T_w, T_v)]^{-1} \text{ Hz}. \tag{B3}$$

If $T_u \gg T_v$, then Δf_1 will usually be much smaller than B_v, the bandwidth of the spectrogram filter. *The filter transfer functions will therefore overlap* if frequency interpolation is not to be used. The required degree of overlap increases as the signal duration T_u becomes longer.

A wide-band noise process (e.g., thermal noise) is usually added to the signal, $u(t)$. In this case, a maximum signal bandwidth B_u and maximum signal duration T_u must be explicitly implemented by the receiver, since the data (signal plus noise) is neither band limited nor time limited. In order to reconstruct $e_{uu}(t,f)$ from spectrogram samples, the receiver must introduce a window function or gate in the time domain. Most transducers experience a gradual decrease in sensitivity at sufficiently high frequencies, but for sampling purposes it may be advantageous to deliberately introduce a more rapid high-frequency cutoff as well as a time window.

3. Sampling rates for statistically uncorrelated sample values

Suppose that the data consists of a time-gated sample function of $u(t)+n(t)$, where $u(t)$ is a random process and $n(t)$ is zero mean, white noise. The time gate is T_u seconds long, and $n(t)$ is statistically independent of $u(t)$.

The power spectral density of the data is the signal power spectral density plus that of the noise (a constant). If the data is passed through a filter with one-sided bandwidth B_v, then uncorrelated sample values are obtained when complex samples of the output process are separated by[17,20,42]

$$\Delta t \geq [\min(B_w, B_v)]^{-1} \text{ s}. \tag{B4}$$

Similarly, uncorrelated complex sample values in the frequency direction are obtained when

$$\Delta f \geq [\min(T_w, T_v)]^{-1} \text{ Hz}. \tag{B5}$$

If the signal and noise processes are both Gaussian, the sample values are independent as well as uncorrelated. Squared independent samples (i.e., real-valued spectrogram samples) are also independent. If the samples are separated as in (B4) and (B5), and if the signal and noise processes are Gaussian, then spectrogram sample values are statistically independent.

APPENDIX C: DETECTOR CONFIGURATIONS

1. Random signal with known covariance

We seek a statistical test to distinguish between hypotheses H_0 and H_1, given a sample function of the data, $z(t)$. The hypotheses are

$$H_0 : z(t) = n(t),$$

$$H_1 : z(t) = u(t) + n(t),$$

where $n(t)$ is a sample function of zero-mean white, Gaussian noise (WGN) and $u(t)$ is a random process with known covariance function

$$R_{uu}(t, \tau) = \overline{u(t)u^*(t+\tau)}.$$

Since the signal and noise processes are assumed to be independent, we have

$$R_{zz}(t, \tau) = R_{uu}(t, \tau) + R_{nn}(t, \tau).$$

From property 16 in Appendix A, we know that the spectrogram of $z(t)$ contains all the information about $\overline{z(t)z^*(t+\tau)}$ that is contained in the original data. Without loss of generality, we can therefore compute the data spectrogram. A likelihood ratio test can then be constructed from the probability density functions (PDF's) of spectrogram samples under H_0 and H_1.

For any random process $x(t)$ that has a Hilbert transform $\hat{x}(t)$,

$$\overline{x(t)\hat{x}(t)} = 0, \tag{C1}$$

i.e., $x(t)$ and $\hat{x}(t)$ are uncorrelated.[43]

For Gaussian input noise, the real and imaginary parts of each filter output are therefore statistically independent and Gaussian. This statement is also true for *non*-Gaussian input noise, provided that the linear filtering operation of the spectrogram filter $V(f-f_j)$ causes non-Gaussian signals to become Gaussian. This tendency toward Gaussianity is most pronounced for narrow-band filters[44] and for wide-band filters with large time–bandwidth product.[45]

Let x_{ij} denote the response $x(t)$ of the jth filter at time i, and let y_{ij} be a sample at time i of the Hilbert transform of the jth filter response. The envelope of the filter response is formed as in (3), and we are interested in the PDF of the *squared* envelope, which is a sample of the spectrogram:

$$S_{ss}(t_i, f_j) = z_{ij} = (x_{ij}/\sqrt{2})^2 + (y_{ij}/\sqrt{2})^2. \tag{C2}$$

The random variables $x = x_{ij}/\sqrt{2}$ and $y = y_{ij}/\sqrt{2}$ are assumed to be Gaussian and independent, with joint PDF

$$p(x, y \mid s_{ij}) = (\pi\sigma^2)^{-1} \exp\{-[(x-\eta_1)^2 + (y-\eta_2)^2]/\sigma^2\}, \tag{C3}$$

where the noise-free spectrogram sample is

$$s_{ij} = \eta_1^2 + \eta_2^2. \tag{C4}$$

To find the PDF of $S_{ss}(t_i, f_j)$ in (C3), we can use the results of Papoulis,[46] pp. 196 and 129;

$$p(z_{ij} \mid s_{ij}, H_1) = \sigma^{-2} \exp[-(z_{ij} + s_{ij})/\sigma^2] \times I_0[2(z_{ij}s_{ij})^{1/2}/\sigma^2] \tag{C5}$$

and

$$p(z_{ij} \mid H_0) = \sigma^{-2} \exp(-z_{ij}/\sigma^2). \tag{C6}$$

In (C5) $I_0[\cdot]$ is a modified Bessel function of zero order.

Equation (C5) is the same as the PDF obtained by McGill[47] from the Rayleigh–Rice distribution. Rice's derivation however, is based upon the assumption that the filter responses are narrow band.[48] *Equation* (C1) *has allowed us to discard the narrow-band assumption.*

Equations (C5) and (C6) specify the conditional PDF for one sample of the spectrogram. Assuming that samples are sufficiently far apart, as in (B4) and (B5), to be statistically independent, we can construct the conditional PDF of an entire spectrogram by forming the products of the PDF's of the individual, independent samples. If the signal $u(t)$ is amplified by a factor A, i.e., if the signal is written $Au(t)$, then s_{ij} in (C5) is replaced by $A^2 s_{ij}$. An SNR parameter θ can then be defined as

$$\theta = A^2/\sigma^2. \tag{C7}$$

In terms of θ, the likelihood ratio is

$$\Lambda_\theta(z) = \int \cdots \int \Lambda(\mathbf{z} \mid \mathbf{s}) p(\mathbf{s})\, d\mathbf{s} = \int \cdots \int \exp\Big(-\sum_{i,j} \theta s_{ij}\Big) \times \prod_{i,j} I_0[2(\theta s_{ij} z_{ij}/\sigma^2)^{1/2}] p(\mathbf{s})\, d\mathbf{s}, \tag{C8}$$

where[49]

$$I_0[2(\theta s_{ij} z_{ij}/\sigma^2)^{1/2}] = 1 + \theta s_{ij} z_{ij}/\sigma^2 + (\theta s_{ij} z_{ij}/\sigma^2)^2/4 + \cdots. \tag{C9}$$

A locally optimum detector for very small SNR can be obtained from (C8) and (C9). The locally optimum detection statistic is determined by calculating $(\partial/\partial\theta) \ln\Lambda_\theta(z)$ at $\theta = 0$.[10,11] Since $\Lambda_\theta(z)$ evaluated at $\theta = 0$ is the ratio of two identical noise-only PDF's, we have

$$\left.\frac{\partial \ln\Lambda_\theta(z)}{\partial\theta}\right|_{\theta=0} = \left.\frac{(\partial/\partial_\theta)\Lambda_\theta(z)}{\Lambda_\theta(z)}\right|_{\theta=0} = \left.\frac{\partial\Lambda_\theta(z)}{\partial\theta}\right|_{\theta=0}. \tag{C10}$$

Substituting (C8) and (C9) into (C10), we obtain

$$\left.\frac{\partial\Lambda_\theta(z)}{\partial\theta}\right|_{\theta=0} = \int \cdots \int \sum_{i,j} (-s_{ij} + s_{ij} z_{ij}/\sigma^2) p(\mathbf{s})\, d\mathbf{s} = \sum_{i=1}^{M}\sum_{j=1}^{N} (-\bar{s}_{ij} + \bar{s}_{ij} z_{ij}/\sigma^2), \tag{C11}$$

where

$$\bar{s}_{ij} = \int s_{ij} p(s_{ij})\, ds_{ij}. \tag{C12}$$

The locally optimum detector compares the quantity in (C11) with a threshold. If the signal $u(t)$ and the filter function $v(t)$ are energy normalized, then from (A8) the term

$$\sum_{i=1}^{M}\sum_{j=1}^{N} \bar{s}_{ij} \propto \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{S_{ss}(t,f)}\, dt\, df = E_u E_v \tag{C13}$$

is the same for all situations, and this term can be incorporated into the threshold. The threshold can also be multiplied by σ^2, and the locally optimum test for distinguishing between signal plus noise (H_1) or noise alone (H_0) is

$$\sum_{i=1}^{M}\sum_{j=1}^{N} \bar{s}_{ij} z_{ij} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{C14}$$

where γ is a threshold.

The locally optimum detector correlates sampled versions of the expected noise-free signal spectrogram $S_{uu}(t,f)$ and the data spectrogram, $S_{zu}(t,f)$.

A generalization of the above formulation for non-Gaussian noise has recently been published.[19] In the Poor–Thomas detector, the square-law operations in (C2) are replaced by nonlinearities that are determined by the PDF of the noise at the filter outputs. From the results of Papoulis[44] and Altes,[45] we expect that these nonlinearities will often not be much different from the square-law devices which were used for Gaussian noise.

If the signal process is not Gaussian, then samples that are separated as in (B4) and (B5) are uncorrelated but not necessarily independent. Poor and Thomas have shown, however, that the locally optimum detector configuration depends only upon the covariance between samples. The spectrogram correlator configuration is therefore unaffected when the signal process is non-Gaussian.

2. Known signal with unknown delay and center frequency, and random, unpredictable phase shift

In this case, the hypotheses are

$$H_0: z(t) = n(t),$$

$$H_1: z(t) = u(t+\tau)\exp[j(2\pi\varphi t + \lambda)] + n(t),$$

where $n(t)$ is white, Gaussian noise (zero mean, variance σ^2) and $u(t)$ is a known, nonrandom, analytic time function. The function $u(t+\tau)\exp(j2\pi\varphi t)$ is a time- and frequency-shifted version of $u(t)$, and the parameters τ and φ are specified only by their joint *a priori* PDF $p(\tau,\varphi)$. The phase shift λ is considered to be a nuisance parameter that changes from pulse to pulse, so that sequential estimation of λ as in Roberts[50] and Hodgkiss[51] cannot be used. The phase λ is assumed to be statistically independent of (τ,φ) and to be uniformly distributed between 0 and 2π radians.

Let ζ denote a set of M statistically independent time samples of the data, $z(t)$. The likelihood ratio is

$$\Lambda(\zeta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{2\pi} \Lambda(\zeta\,|\,\tau,\varphi,\lambda)p(\lambda)p(\tau,\varphi)\,d\lambda\,d\tau\,d\varphi, \tag{C15}$$

where, for Gaussian noise,

$$\Lambda(\zeta\,|\,\tau,\varphi,\lambda) = \exp\{[-E_u + 2\,\mathrm{Re}[\chi_{zu}(\tau,\varphi)e^{j\lambda}]]/2\sigma^2\} \tag{C16}$$

and

$$E_u = \sum_{i=1}^{M} |u(t_i+\tau)|^2 \equiv \sum_{i=1}^{M} |u(t_i)|^2, \tag{C17}$$

$$\chi_{zu}(\tau,\varphi) = \sum_{i=1}^{M} z(t_i)u^*(t_i+\tau)\exp(-j2\pi\varphi t_i), \tag{C18}$$

$$\mathrm{Re}[\chi_{zu}(\tau,\varphi)e^{j\lambda}] = |\chi_{zu}(\tau,\varphi)|\cos[\lambda+\alpha(\tau,\varphi)], \tag{C19}$$

where $\alpha(\tau,\varphi)$ is the phase of the complex function $\chi_{zu}(\tau,\varphi)$. Performing the λ integration in (C15), we have

$$\begin{aligned}\Lambda(\zeta\,|\,\tau,\varphi) &= \int_0^{2\pi}\Lambda(\zeta\,|\,\tau,\varphi,\lambda)p(\lambda)\,d\lambda \\ &= e^{-E_u/2\sigma^2}(2\pi)^{-1}\int_0^{2\pi} e^{[|\chi_{zu}(\tau,\varphi)|/\sigma^2]\cos[\lambda+\alpha(\tau,\varphi)]}\,d\lambda \\ &= e^{-E_u/2\sigma^2} I_0[|\chi_{zu}(\tau,\varphi)|/\sigma^2],\end{aligned} \tag{C20}$$

where $I_0(\cdot)$ is a modified Bessel function.

If $u(t)$ is multiplied by a real parameter A, then $|\chi_{zu}(\tau,\varphi)|$ is also multiplied by A. Defining a signal-to-noise ratio parameter θ as in (C7), and using the series expansion for $I_0(\cdot)$ as in (C9), we have

$$\begin{aligned}\Lambda_\theta(\xi) = e^{-E_u/2\sigma^2}\int\int \{1 &+ [\theta|\chi_{zu}(\tau,\varphi)|^2/4\sigma^2] \\ &+ [\theta^2|\chi_{zu}(\tau,\varphi)|^4/64\sigma^2] + \cdots\}p(\tau,\varphi)\,d\tau\,d\varphi,\end{aligned} \tag{C21}$$

where unlabeled limits of integration are $(-\infty,\infty)$. The locally optimum test for this detection problem is then

$$\left.\frac{\partial\Lambda_\theta(\zeta)}{\partial\theta}\right|_{\theta=0} = \frac{e^{-E_u/2\sigma^2}}{4\sigma^2}\int\int |\chi_{zu}(\tau,\varphi)|^2 p(\tau,\varphi)\,d\tau\,d\varphi \underset{H_1}{\overset{H_1}{\gtrless}} \gamma, \tag{C22}$$

where γ is a threshold and $\int\int |\chi_{zu}(\tau,\varphi)|^2 p(\tau,\varphi)\,d\tau\,d\varphi$ is a correlator that operates upon $|\chi_{zu}(\tau,\varphi)|^2$.

A sampled version of the function $|\chi_{zu}(\tau,\varphi)|^2$ can be generated by a bank of filters with transfer functions $U^*(f-\varphi_j)$, $j=1,2,\ldots,N$. Each of these filters is followed by a square-law envelope detector. The resulting configuration is identical to the spectrogram correlator in case 1, except that the reference spectrogram $S_{uu}(t,f)$ is replaced by the *a priori* PDF $p(\tau,\varphi)$. Property 4 in Appendix A indicates that the magnitude-squared ambiguity function is a specialized form of spectrogram.

The results of Poor and Thomas[19] again indicate that the test in (C22) can be easily reconfigured for non-Gaussian noise by changing the square-law envelope detectors into different zero memory nonlinearities.

3. Known signal with random, unpredictable phase shift

In this case,

$$H_0: z(t) = n(t),$$

$$H_1: z(t) = u(t)\exp(j\lambda) + n(t),$$

where $u(t)$ is known, $n(t)$ is WGN with zero mean and variance σ^2, and λ is a nuisance parameter that changes from pulse to pulse and is uniformly distributed between 0 and 2π radians.

This problem is just a special version of case 2. The problems are the same if

$$p(\tau,\varphi) = \delta(\tau)\delta(\varphi) \tag{C23}$$

in (C22). The locally optimum detector is then

$$\left.\frac{\partial \Lambda_\theta(\zeta)}{\partial \theta}\right|_{\theta=0} = \frac{e^{-E_u/2\sigma^2}}{4\sigma^2} |\chi_{zu}(0,0)|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{C24}$$

where

$$|\chi_{zu}(0,0)|^2 \propto \left| \int_{-\infty}^{\infty} z(t)u^*(t)\,dt \right|^2 \tag{C25}$$

= the envelope detected response of a matched filter or correlator that operates on the data, $z(t)$.

[1]R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra* (Dover, New York, 1958).

[2]M. R. Schroeder and B. S. Atal, "Generalized short-time power spectra and autocorrelation functions," J. Acoust. Soc. Am. 34, 1679–1683 (1962).

[3]G. Gambardella, "A contribution to the theory of short-time spectral analysis with non-uniform bandwidth filters," IEEE Trans. Circuit Theory CT-18, 455–460 (1971).

[4]R. A. Cole, A. I. Rudnicky, D. R. Reddy, and V. Zue, "Performance of an expert spectrogram reader. I." J. Acoust. Soc. Am. Suppl. 1 64, S178(A) (1978).

[5]A. I. Rudnicky, R. A. Cole, and V. Zue, "Performance of an expert spectrogram reader. II." J. Acoust. Soc. Am. Suppl. 1 64, S178(A) (1978).

[6]J. L. Flanagan, *Speech Analysis, Synthesis, and Perception* (Springer–Verlag, Berlin, 1965), pp. 140–163, 323–327.

[7]J. Makhoul, "Linear prediction: A tutorial review," Proc. IEEE 63, 561–580 (1975).

[8]P. A. Bello, "Characterization of randomly time-variant linear channels," IEEE Trans. Commun. Syst. CS-11, 360–393 (1963).

[9]R. F. Daly, "Signal design for efficient detection in dispersive channels," IEEE Trans. Theory IT-16, 206–213 (1970).

[10]J. Capon, "On the asymptotic efficiency of locally optimum detectors," IRE Trans. Inf. Theory IT-7, 67–71 (1961).

[11]D. Middleton, "Canonically optimum threshold detection," IEEE Trans. Inf. Theory IT-12, 230–243 (1966).

[12]D. Gabor, "Theory of Communication," JIEE (London) 93, 429–457 (1946).

[13]A. W. Rihaczek, *Principles of High-Resolution Radar* (McGraw–Hill, New York, 1969), pp. 10–40, 331–349.

[14]L. E. Franks, *Signal Theory* (Prentice–Hall, Englewood Cliffs, 1969), pp. 78–80.

[15]A. Erdélyi, *Tables of Integral Transforms* (McGraw–Hill, New York, 1954), Vol. II., p. 243.

[16]A. W. Rihaczek, "Signal energy distributions in time and frequency," IEEE Trans. Inf. Theory IT-14, 369–374 (1968).

[17]P. M. Woodward, *Probability and Information Theory with Applications to Radar* (Pergamon, Oxford, 1964), pp. 115–125.

[18]M. H. Ackroyd, "Short-time spectra and time-frequency energy distributions," J. Acoust. Soc. Am. 50, 1229–1231 (1971).

[19]H. V. Poor and J. B. Thomas, "Locally optimum detection of discrete-time stochastic signals in non-Gaussian noise," J. Acoust. Soc. Am. 63, 75–80 (1978).

[20]C. E. Cook and M. Bernfeld, *Radar Signals* (Academic, New York, 1967), pp. 88, 160–164.

[21]T. Kailath, "The innovations approach to detection and estimation theory," Proc. IEEE 58, 680–695 (1970).

[22]S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust. Speech, Signal Process. ASSP-27, 113–120 (1979).

[23]B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," Proc. IEEE 63, 1692–1716 (1975).

[24]H. Cramér, *Mathematical Methods of Statistics* (Princeton U. P., Princeton, 1945).

[25]H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York, 1968), pp. 19–86.

[26]R. A. Altes, "Spectrogram Correlation," Vol. 2, in *Further development and new concepts for bionic sonar*, by R. A. Altes and W. J. Faust, Report OC-R-78-A004-1, Orincon Corp., Oct. 1978. Available as Report TR-404, NOSC Hawaii Lab., P. O. Box 997, Kailua, HI, 96734 (1978).

[27]E. J. Kelly, "The radar measurement of range, velocity, and acceleration," IRE Trans. Mil. Electronics MIL-5, 51–57 (1961).

[28]E. J. Kelly, I. S. Reed, and W. L. Root, "The detection of radar signals in noise, Parts I and II," J. Soc. Ind. Appl. Math. 8, 309–341, 481–507 (1960).

[29]N. Abramson and D. Braverman, "Learning to recognize patterns in a random environment," IEEE Trans. Inf. Theory IT-8, 58–63 (1962).

[30]Y. T. Chien and K. S. Fu, "Selection and ordering of feature observations in a pattern recognition system," Inf. Control 12, 395–414 (1968).

[31]D. H. Foley and J. W. Sammon, "An optimal set of discriminant vectors," IEEE Trans. Comput. C-24, 281–289 (1975).

[32]T. T. Kadota and L. A. Shepp, "On the best finite set of linear observables for discriminating two Gaussian signals," IEEE Trans. Inf. Theory IT-13, 278–284 (1967).

[33]R. A. Altes and W. J. Faust, "A unified method of broadband acoustic characterization for fish schools, channels, and range-extended targets," Oceans '79 Conf. Rec. IEEE No. 79-CH1478-7 OEC, 92–99 (1979).

[34]R. A. Altes, "Target position estimation in radar and sonar, and generalized ambiguity analysis for maximum likelihood parameter estimation," Proc. IEEE 67, 920–930 (1979).

[35]H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part III* (Wiley, New York, 1971), pp. 8–32.

[36]C. A. Stutt, "Some results on real-part/imaginary-part and magnitude-phase relations in ambiguity functions," IEEE Trans. Inf. Theory IT-10, 321–327 (1964).

[37]W. M. Siebert, "Studies of Woodward's uncertainty function," Quarterly Progress Report, MIT, 90–94 (1958).

[38]E. L. Titlebaum and N. DeClaris, "Linear transformations of the ambiguity function," IEEE Trans. Inf. Theory IT-12, 120–125 (1966).

[39]M. D. Flaska, "Cross correlation of short-time spectral histories," J. Acoust. Soc. Am. 59, 381–388 (1976).

[40]J. L. Stewart and E. C. Westerfield, "A theory of active sonar detection," Proc. IRE 47, 872–881 (1959).

[41]N. T. Gaarder, "Scattering function estimation," IEEE Trans. Inf. Theory IT-14, 684–693 (1968).

[42]D. Slepian, "Estimation of signal parameters in the presence of noise," IRE Trans. Inf. Theory IT-1, 68–69 (1954).

[43]A. Papoulis, *Systems and Transforms with Applications in Optics* (McGraw–Hill, New York, 1968), p. 174.

[44]A. Papoulis, "Narrowband systems and Gaussianity," IEEE Trans. Inf. Theory IT-18, 20–27 (1972).

[45]R. A. Altes, "Wide-band systems and Gaussianity," IEEE Trans. Inf. Theory IT-21, 679–682 (1975).

[46]A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw–Hill, New York, 1965), pp. 129–196.

[47]W. J. McGill, "Variations on Marrill's detection formula," J. Acoust. Soc. Am. 43, 70–73 (1968).

[48]S. O. Rice, "Mathematical analysis of random noise," in *Noise and Stochastic Processes*, edited by N. Wax (Dover, New York, 1954).

[49]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, NBS AMS 55 (US Government Printing Office, Washington, DC, 1964), p. 375.

[50]R. A. Roberts, "On the detection of a signal known except for phase," IEEE Trans. Inf. Theory IT-11, 76–82 (1965).

[51]W. S. Hodgkiss, "Detection of LPM signals with estimation of their velocity and time of arrival," J. Acoust. Soc. Am. 64, 177–180 (1978).