Introduction to Time-Frequency Analysis and Its Biomedical Applications

Dr. Ervin Sejdić, Ph.D.

Innovative Medical Engineering Developments (iMED) lab
Department of Electrical and Computer Engineering
University of Pittsburgh
E-mail: esejdic@ieee.org
Web: www.imedlab.org

Time and frequency representations of signals

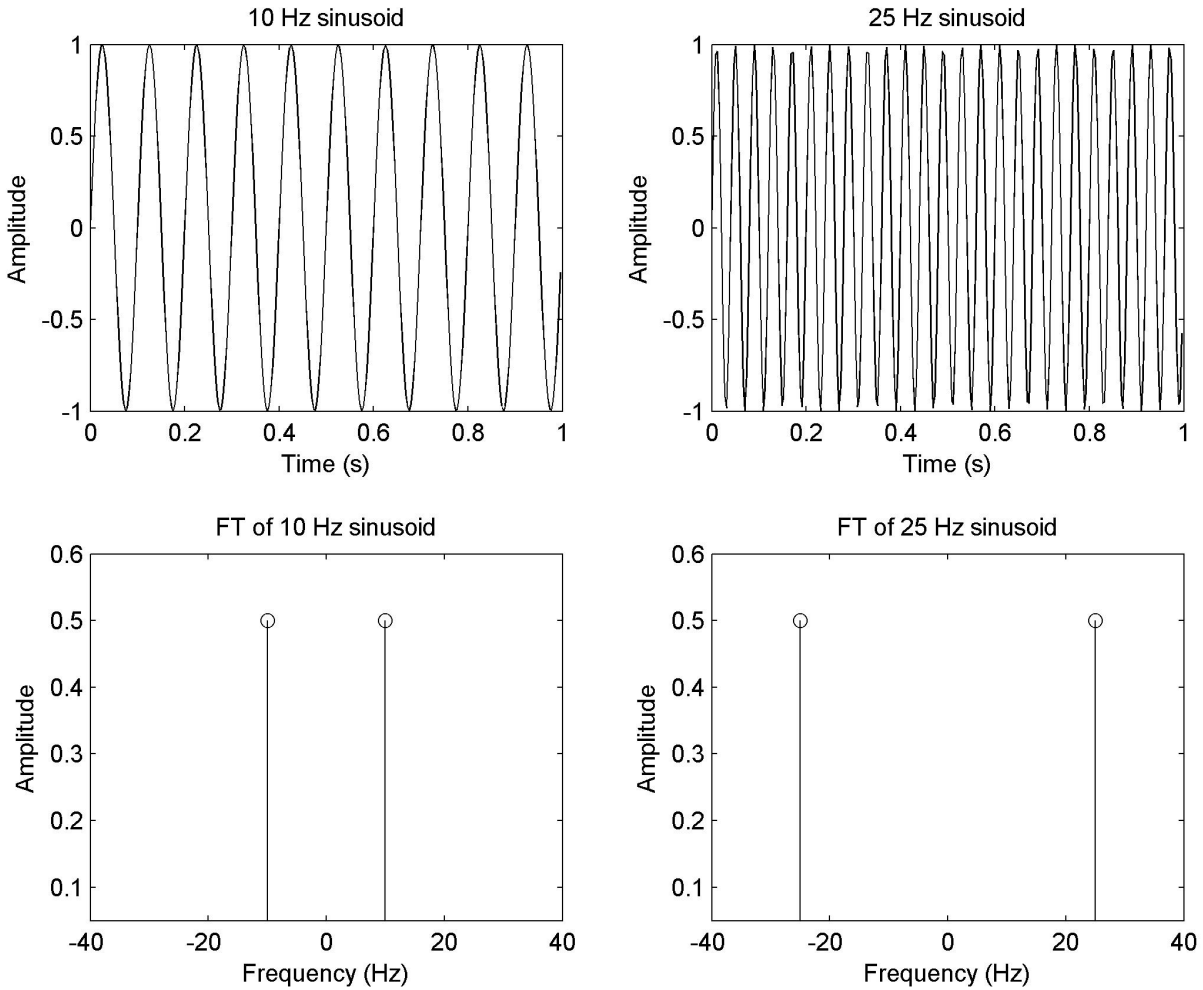


Figure 1: Sample signals.

Houston, we have a problem!

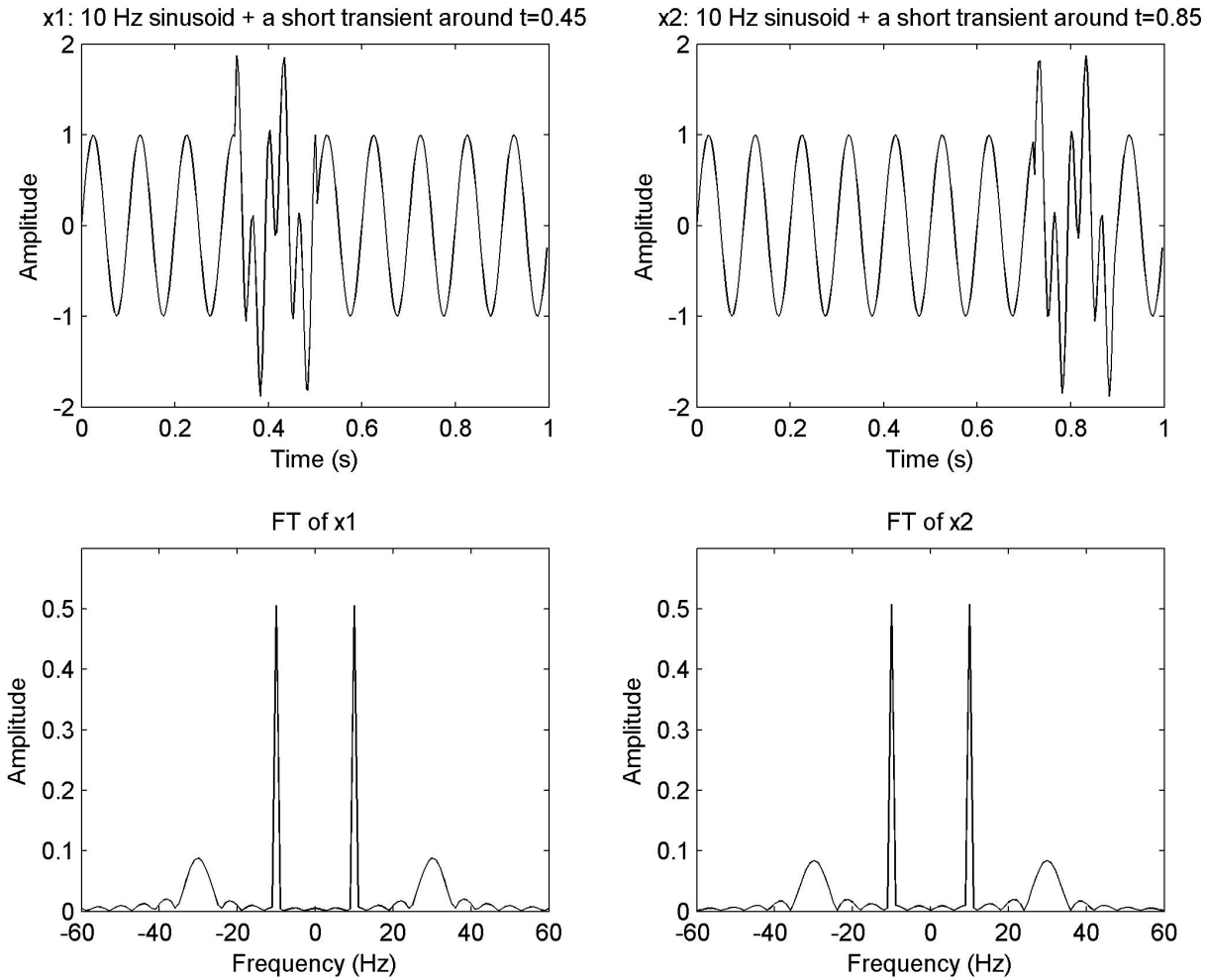


Figure 2: Sample signals.

Ooops...I did it again!

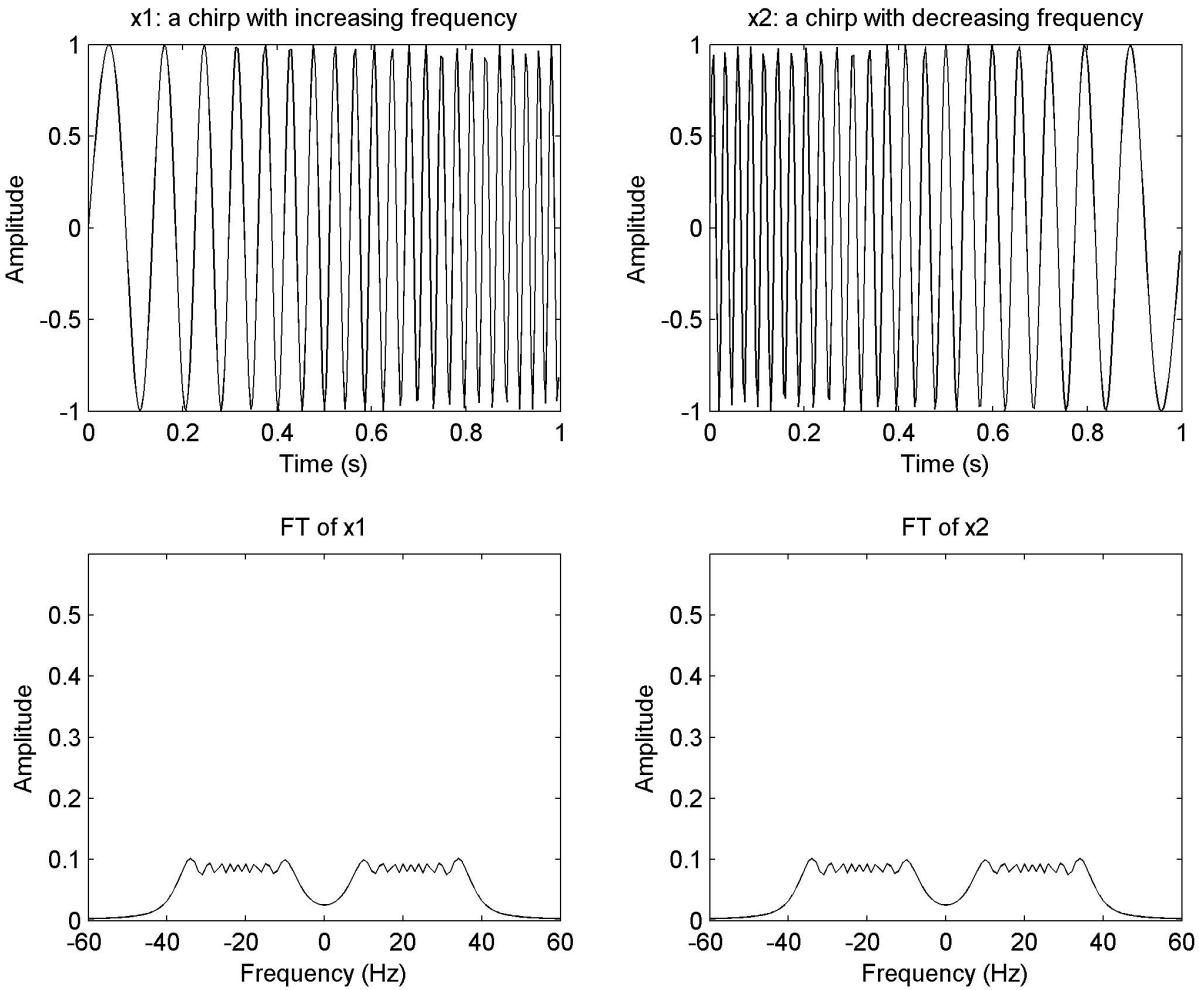


Figure 3: Sample signals.

Why do we need TFA?

- The time or the frequency domain descriptions of a signal alone cannot provide comprehensive information for feature extraction and classification.
- The time domain lacks the frequency description of the signals.
- The Fourier transform of the signal cannot depict how the spectral content of the signal changes with time.
- The time variable is introduced in the Fourier based analysis in order to provide a proper description of the spectral content changes as a function of time

STFT of signals from Figure 2

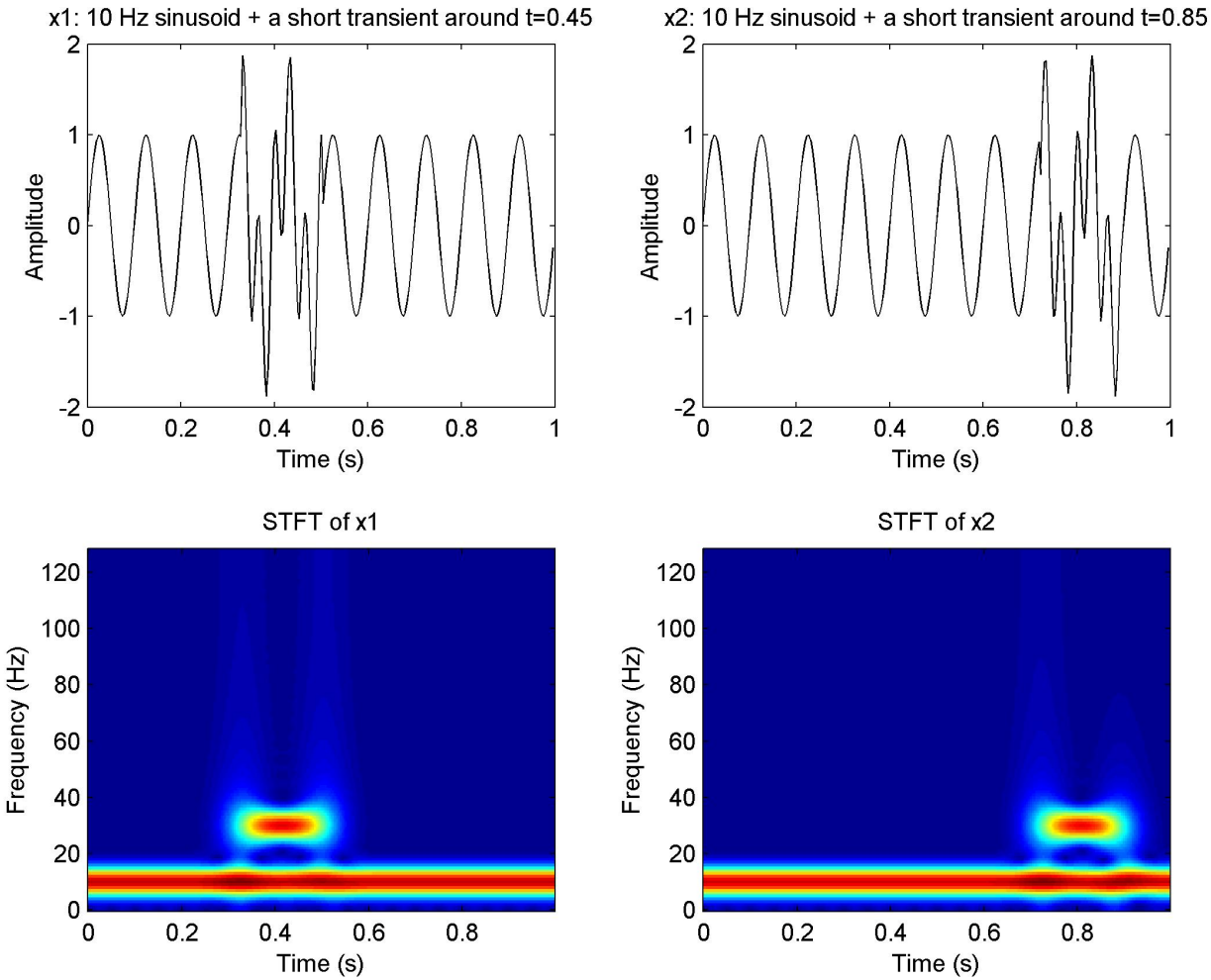


Figure 4: TFR of sample signals.

STFT of signals from Figure 3

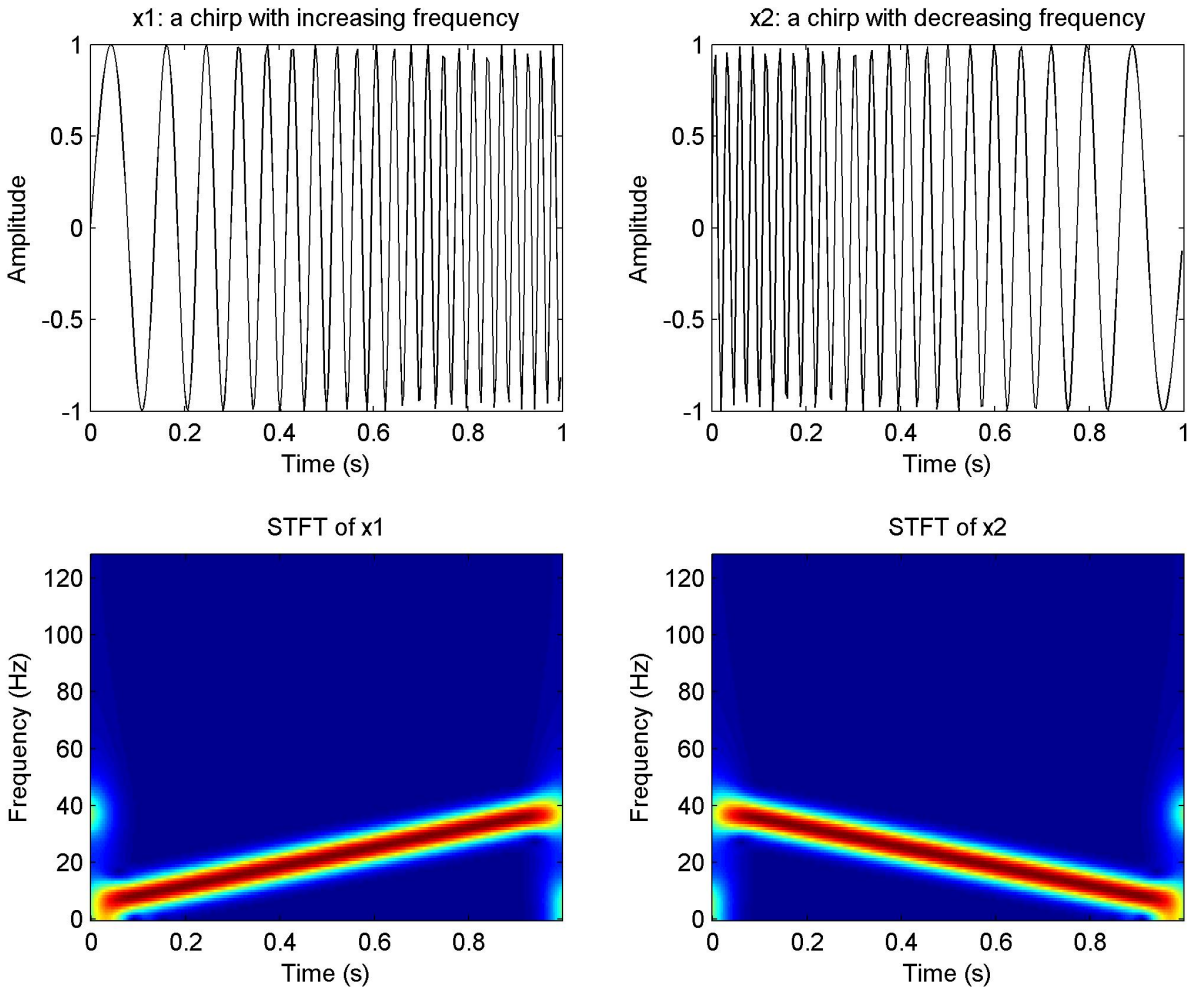


Figure 5: TFR of sample signals.

TFA

- The basic goal of the TFA is to determine the energy concentration along the frequency axis at a given time instant, i.e., to search for joint time-frequency representation of the signal.
- In an ideal case, the time-frequency transform would provide direct information about the frequency components occurring at any given time by combining the local information of an "instantaneous frequency spectrum"with the global information of the temporal behaviour of the signal.
- In reality, the Heisenberg uncertainty principle prohibits the existence of windows/kernels with arbitrarily small duration and arbitrarily small bandwidth.
- Resolution tradeoff: improving the time resolution (by using a short window) results in a loss of frequency resolution, and vice versa.

Example

- The main goal of the TFA of a signal is to determine the energy distribution along the frequency axis at each time instant.
- 3 sample signals: $x_1(t)$ - a signal with four short transients; $x_2(t)$ - a linear chirp; and $x_3(t)$ - a signal with sinusoidally modulated frequency.
- The TF domain representations of the signals are obtained by four different TFRs: STFT, S-transform, S-method, and Wigner distribution (WD).

Example

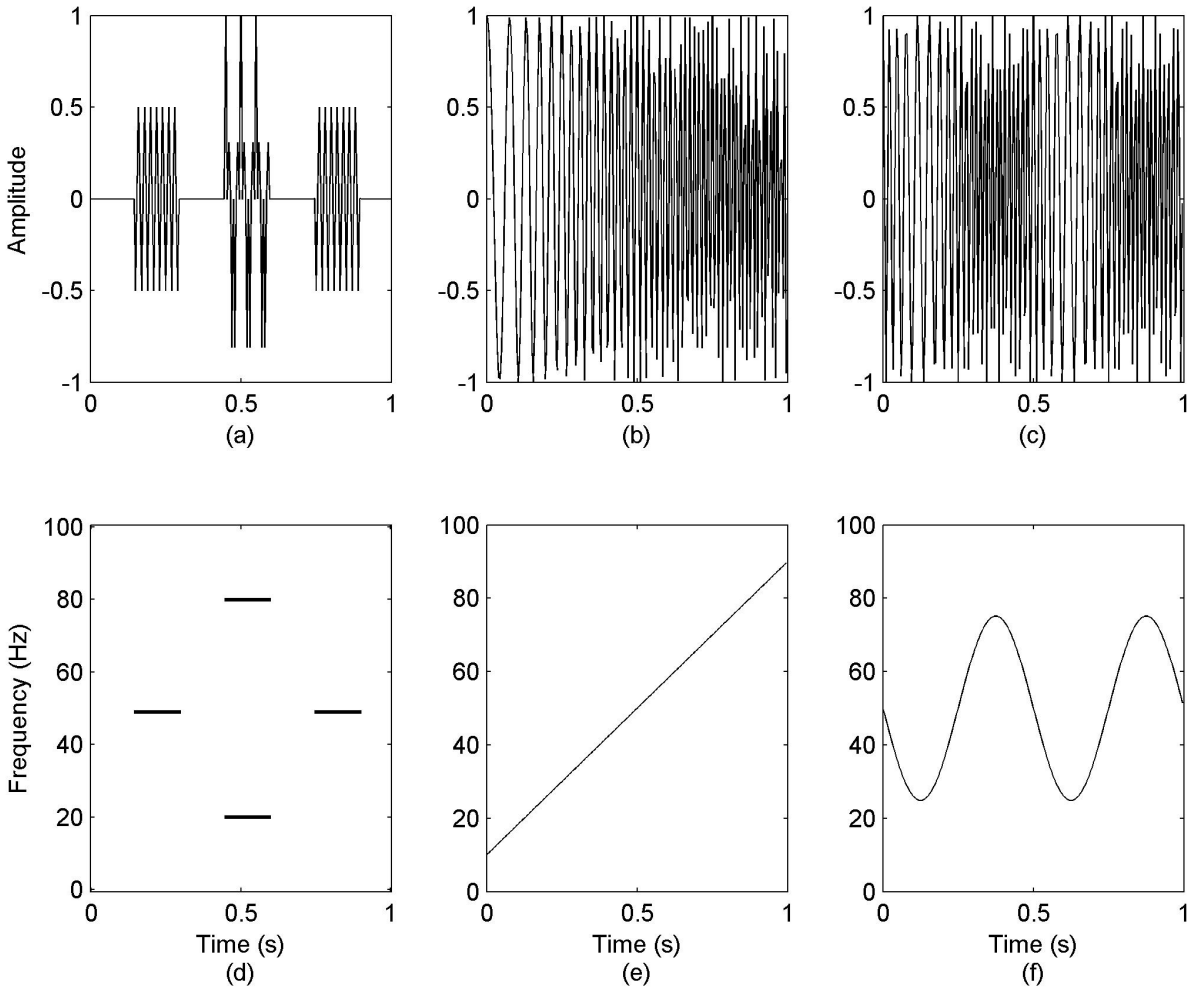


Figure 6: Ideal TFRs of sample signals.

Example

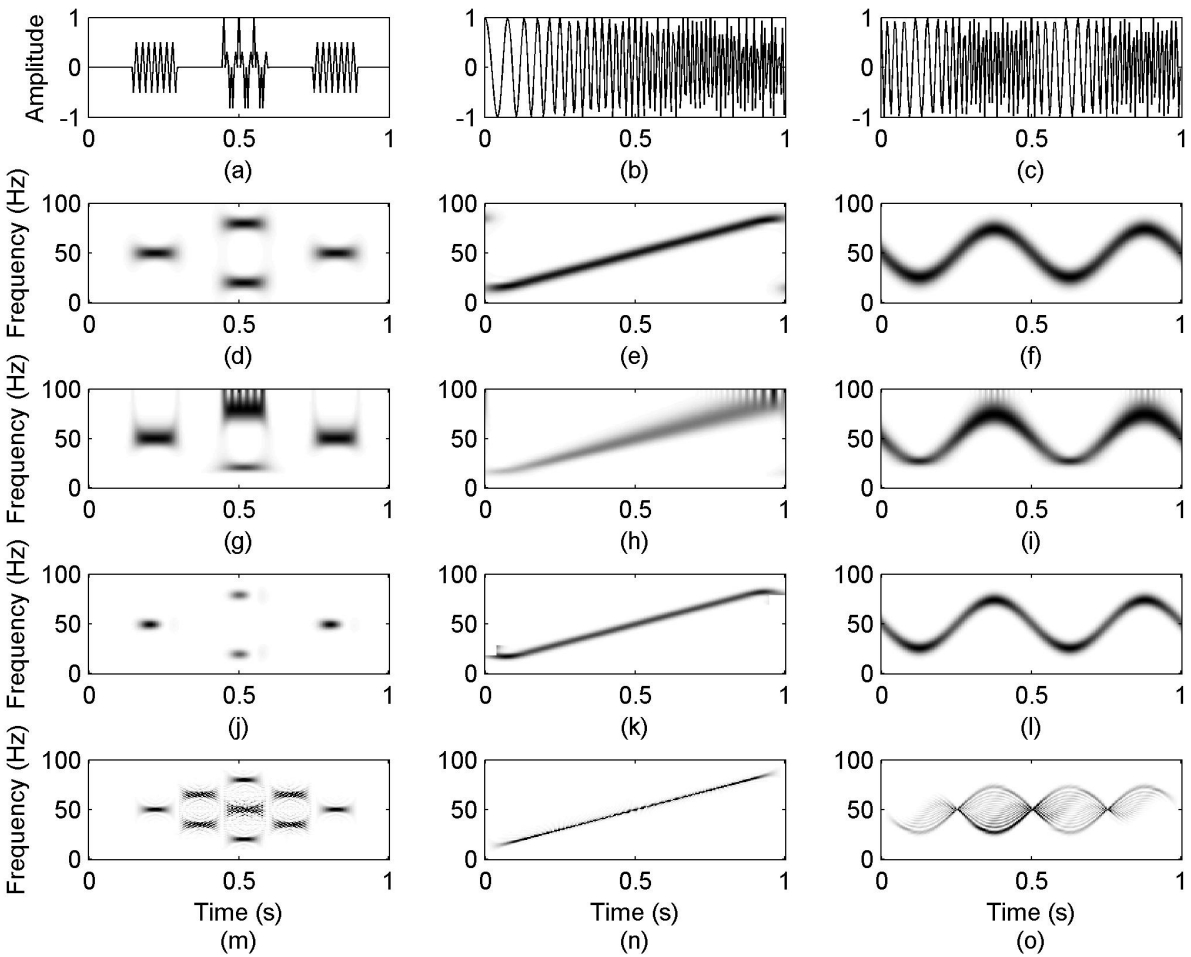


Figure 7: TFRs of sample signals.

Signal decomposition based TFRs

- A signal is represented by TF functions derived from translating, modulating and scaling a basis function having a definite time and frequency localization.
- Mathematically

$$TF_x(t,\omega) = \int_{-\infty}^{+\infty} x(\tau)\phi^*_{t,\omega}(\tau)d\tau = \langle x, \phi_{t,\omega}\rangle \tag{1}$$

  where $\phi_{t,\omega}$ represents the basis functions (also called the TF atoms) and * represents the complex conjugate.
- The basis functions are assumed to be square integrable, $\phi_{t,\omega} \in \mathbf{L}^2(\mathbb{R})$, i.e., they have finite energy
- STFT, wavelets, and matching pursuit algorithms are typical examples in this category.

Typical transforms

- short-time Fourier transform

$$\phi_{t,\omega}(\tau) = h(\tau - t)\exp(-j\omega\tau) \tag{2}$$

where $h(\cdot)$ is a window function;

- wavelet transform:

$$\phi_{t,\omega}(\tau) = \sqrt{s}h(s(\tau - t)) \tag{3}$$

where $h(\cdot)$ is a wavelet function and s is the scale;

- multiresolution Fourier transform (MFT):

$$\phi_{t,\omega}(\tau) = \sqrt{s}h(s(\tau - t))\exp(-j\omega\tau) \tag{4}$$

where $h(\cdot)$ is a window function and s is the scale similar to one used in the wavelet analysis;

Typical transforms

- S-transform:

$$\phi_{t,\omega}(\tau) = h(\tau - t, \sigma(\omega)) \exp(-j\omega\tau) \tag{5}$$

where $h(\cdot)$ is a Gaussian window function and $\sigma(\omega)$ is the standard deviation of the Gaussian window;

- short-time harmonic transform (STHRT):

$$\phi_{t,\omega}(\tau) = h(t - \tau)\varphi_u^{(1)}(\tau)\exp(-j\omega\varphi_u(\tau)) \tag{6}$$

where $\varphi_u(\tau)$, known as the unit phase function, is the phase function of the fundamental divided by its nominal IF and $\varphi_u^{(1)}(\tau)$ is the first-order derivative of $\varphi_u(\tau)$;

- short-time Hartley transform (STHT):

$$\phi_{t,\omega}(\tau) = h(t - \tau)\text{cas}(\omega\tau) \tag{7}$$

where $\text{cas}(\cdot) = \cos(\cdot) + \sin(\cdot)$.

Comparison of STFT and S-transform

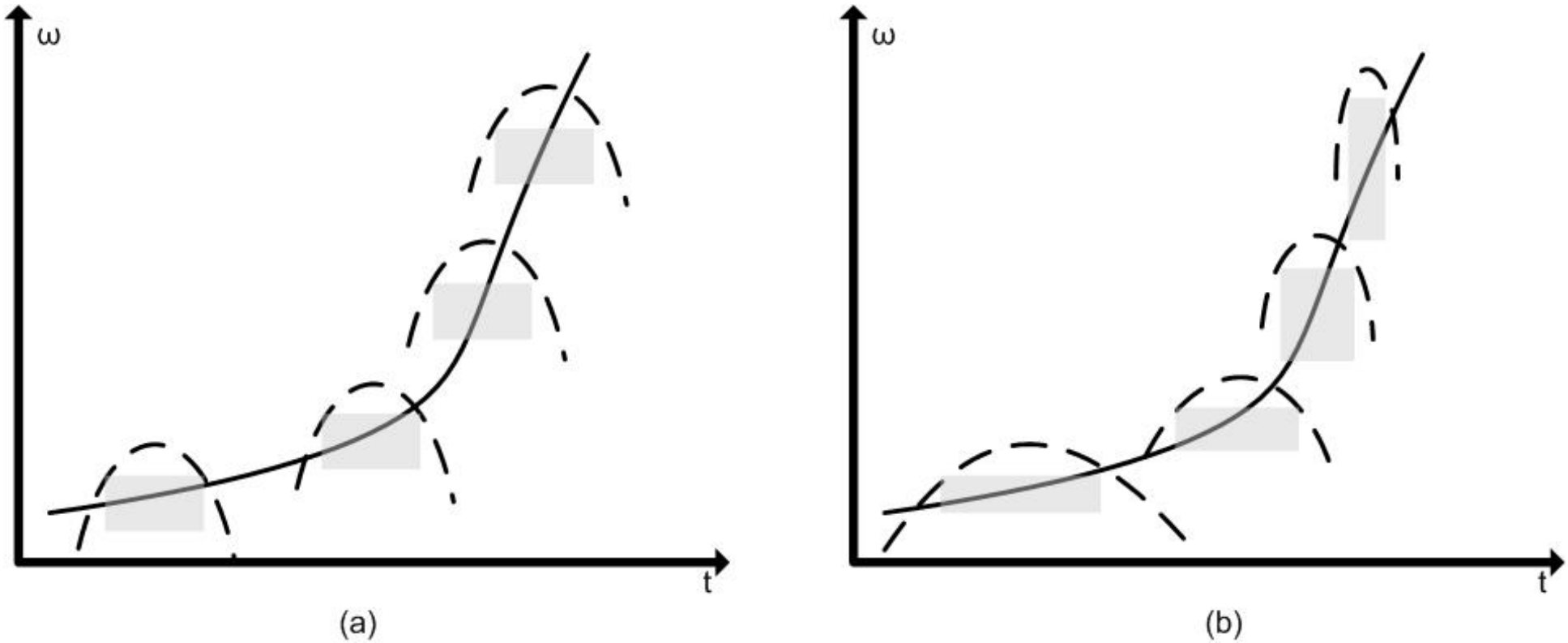


Figure 8: Comparison of STFT and S-transform.

Example

- A hyperbolic FM signal, $x(t) = \exp(j20\pi \ln(11|t| + 1))$, is used to examine the effects of a variable window width over a constant window. The signal is analyzed with STFT and the S-transform.

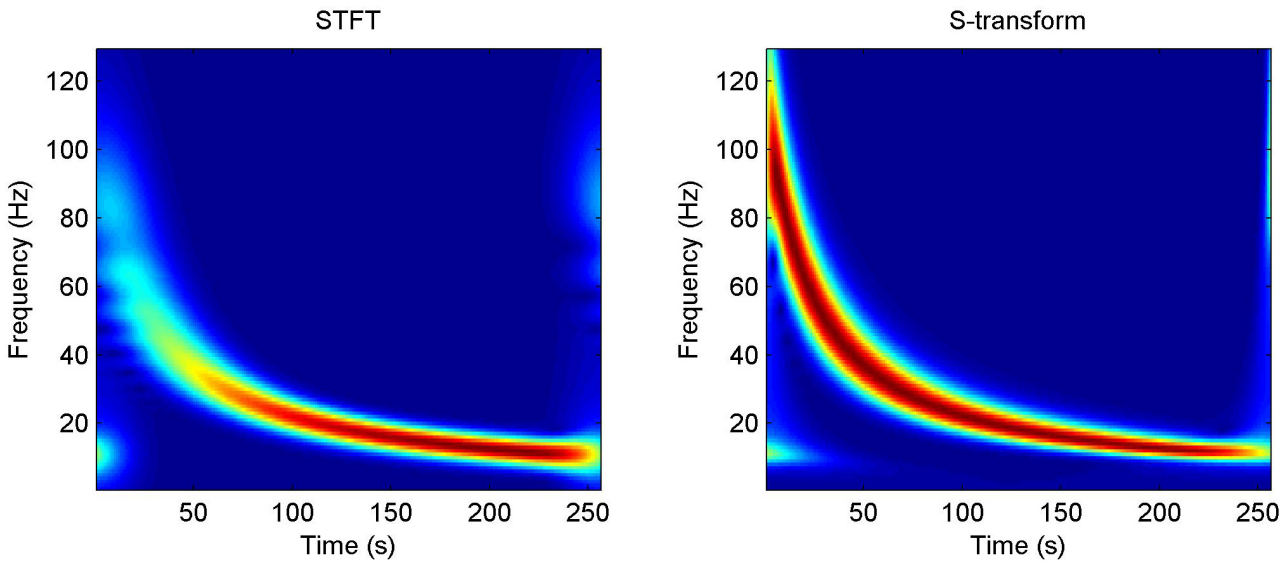


Figure 9: TFR of sample signals.

Example

- The harmonic signal, $x(t) = \exp(j2\pi(10t+5t^2))+\exp(j2\pi(20t+5t^2))+\exp(j2\pi(30t+5t^2))$, is used to examine the STHRT and the STFT.

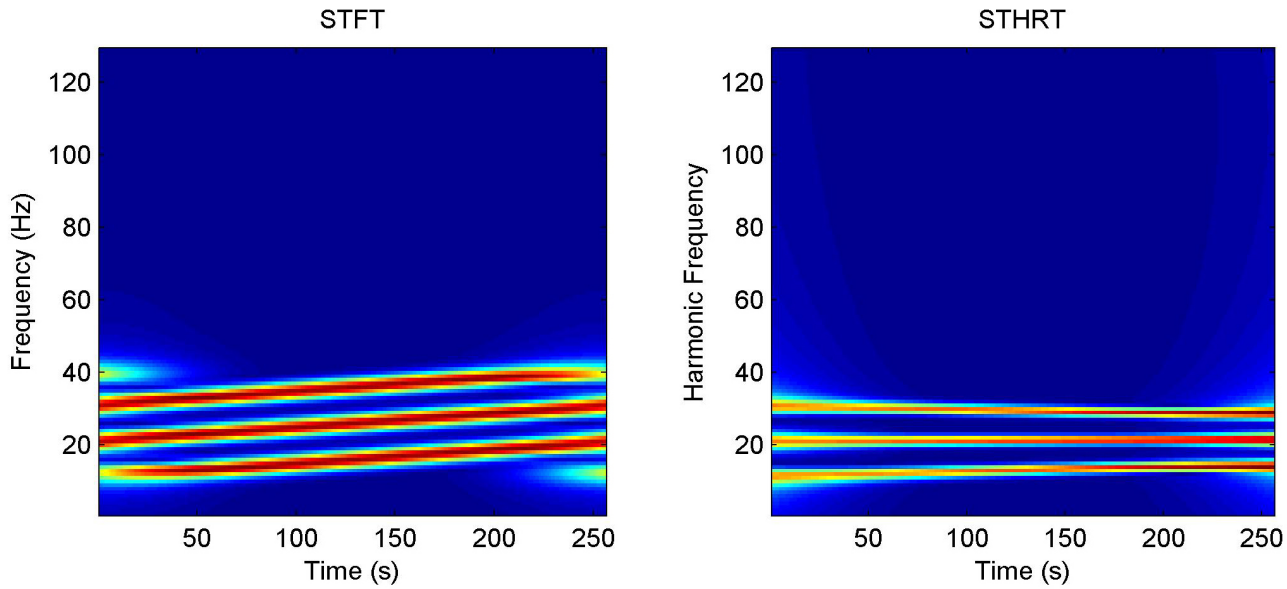


Figure 10: TFR of sample signals.

Properties

Method	Advantages	Disadvantages
STFT	Very simple for implementation.	Constant window width limits time-frequency resolution.
Wavelet analysis	Variable resolution.	Does not maintain the absolute phase of the signal components. A scale to frequency conversion is dependent on a mother wavelet.

Properties

Method	Advantages	Disadvantages
MFT	Variable resolution. Absolute phase of each component is maintained.	Complex requirements for the window function. Choice of scale might require oversampling.
S-transform	Variable resolution. Absolute phase of each component is maintained.	Single window function.
STHRT	Good energy concentration obtained for the harmonic signals.	$\varphi_u(\tau)$ has to be known or precisely estimated.
STHT	Easy for hardware implementation.	Same disadvantages as STFT.

Avoiding shortcomings

- A generalized S-transform is introduced to allow greater control over the window function. This generalization also allows nonsymmetric windows to be used, which were also introduced in the literature.
- The solution to the problem of the constant window width associated with the STHT is proposed in the form of a Hartley S-transform. The Hartley S-transform introduces a variable window width framework for the Hartley analysis.

TFD

- The properties of the representation are reflected by simple constraints on the kernel that produces the TFR with prescribed, desirable properties.
- Mathematically,

$$TFD_x(t,\omega) = \frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\left(u+\frac{1}{2}\tau\right)x^*\left(u-\frac{1}{2}\tau\right) \phi(\theta,\tau)e^{-j\theta t-j\tau\omega+j\theta u}dud\tau d\theta \quad (8)$$

  where $\phi(\theta,\tau)$ is a two-dimensional kernel function, determining the specific representation in this category, and hence, the properties of the representation.
- Wigner distribution, Choi-Williams distribution, and spectrogram are some of the methods commonly used for obtaining the TFDs.

TFD

- Cross terms and inner interferences lead to the ambiguous representation of a signal in the TF domain.

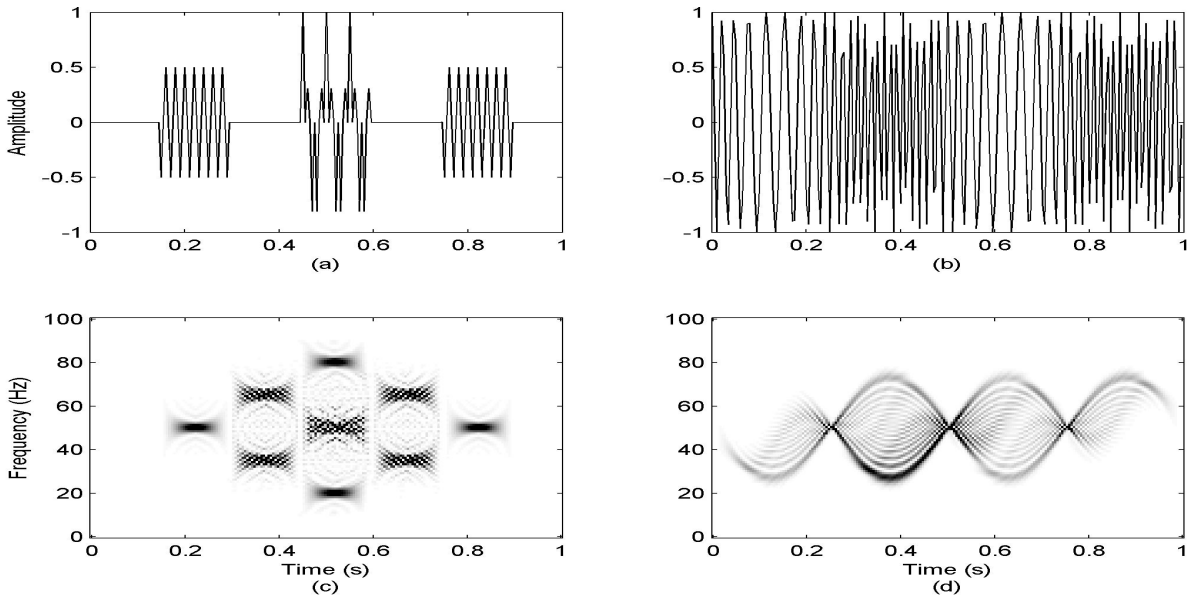


Figure 11: TFR of sample signals.

- When cross terms and inner interferences are minimized, these transforms can produce very high resolution representations.

TFD

- The cross terms can be reduced or eliminated by introducing a kernel function $\phi(\theta, \tau)$.
- Let's rewrite Cohen's class of the TFRs in terms of the ambiguity function, $A(\theta, \tau)$, which is defined as:

$$A(\theta, \tau) = \int_{-\infty}^{+\infty} x\left(u + \frac{1}{2}\tau\right) x^*\left(u - \frac{1}{2}\tau\right) e^{j\theta u} du \tag{9}$$

and the Cohen's class can then be rewritten as

$$TFD_x(t, \omega) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(\theta, \tau)\phi(\theta, \tau) e^{-j\theta t - j\tau\omega} d\tau d\theta. \tag{10}$$

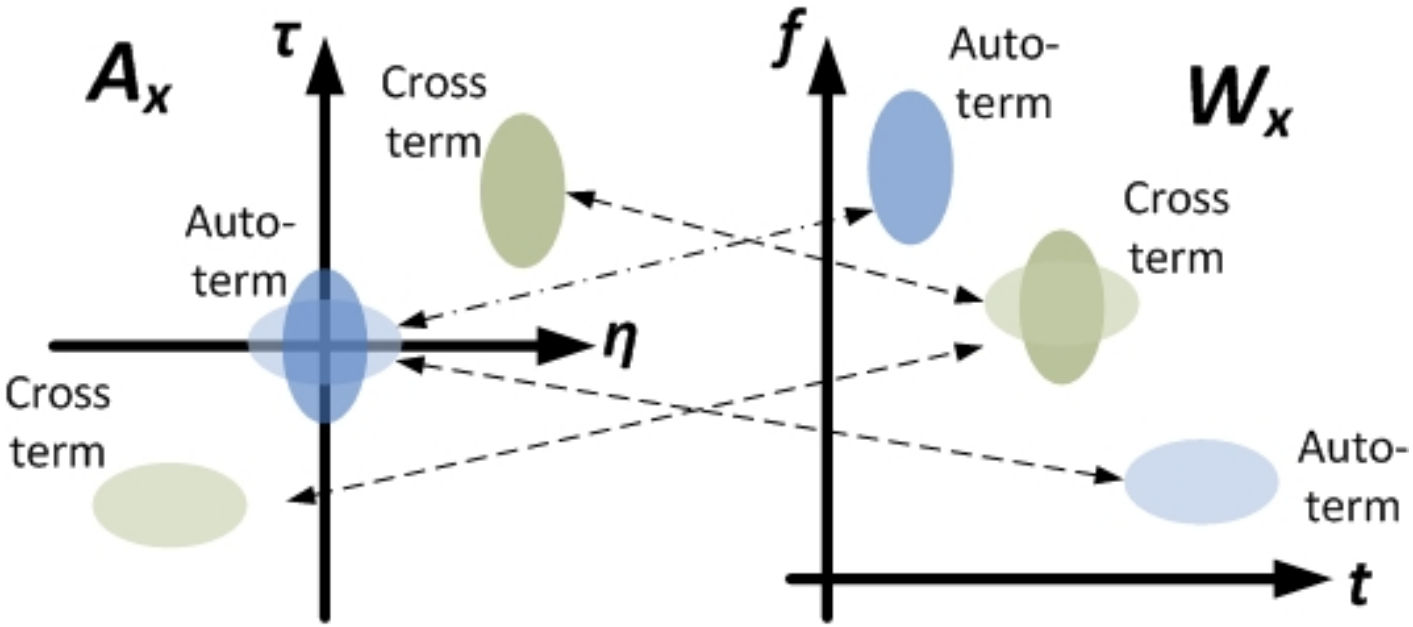


Figure 12: Relationship between AF and WD.

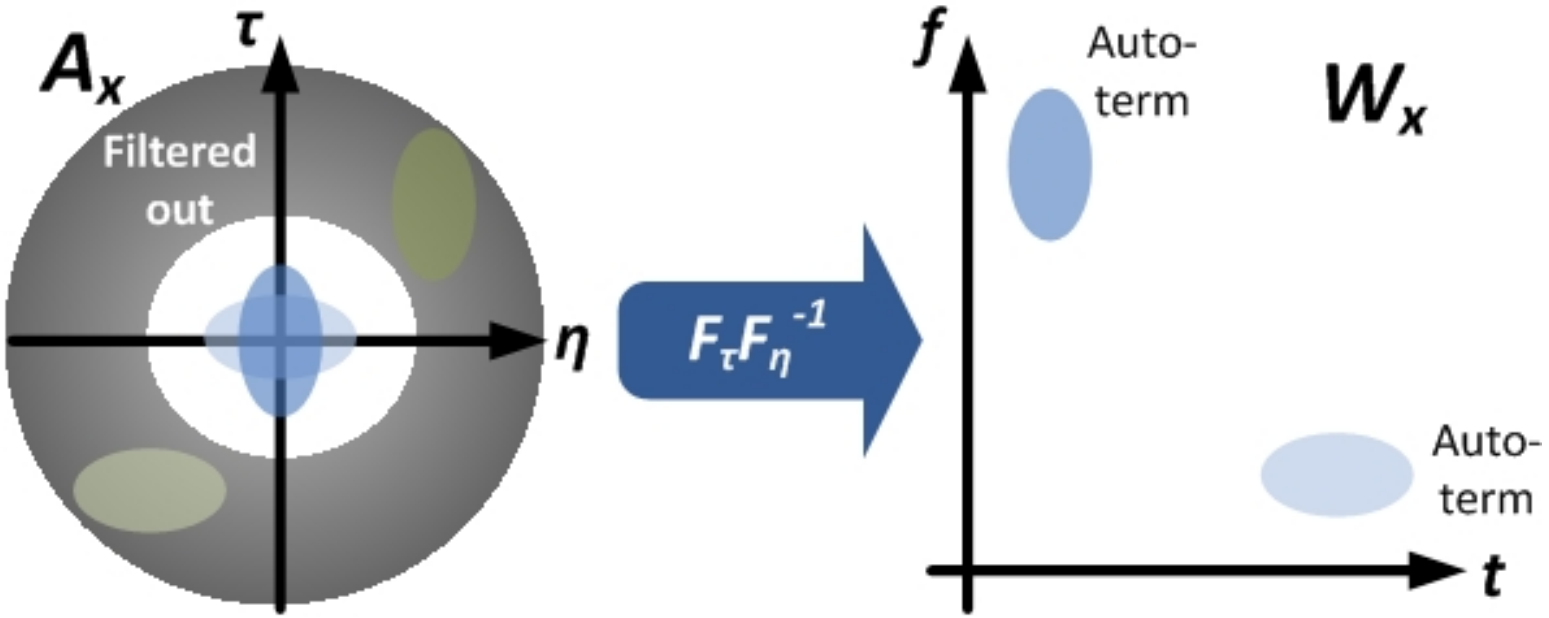


Figure 13: Relationship between AF and WD.

Kernel functions

- Born-Jordan distribution:

$$\phi(\theta,\tau) = \frac{\sin(\theta\tau/2)}{\theta\tau/2}. \tag{11}$$

- Choi-Williams distribution:

$$\phi(\theta,\tau) = \exp\left(-\frac{\theta^2\tau^2}{\sigma^2}\right) \tag{12}$$

where σ is a scaling factor.

- Zhang-Sato distribution:

$$\phi(\theta,\tau) = \exp\left(-\frac{\theta^2\tau^2}{\sigma^2}\right)\cos(2\pi\beta\tau) \tag{13}$$

where σ and β are parameters. For $\beta = 0$ a Choi-Williams distribution is obtained, since σ is defined in the same manner as for the Choi-Williams distribution.

Kernel functions

- Radial Butterworth distribution:

$$\phi(\theta,\tau) = \frac{1}{1+\left(\frac{\theta^2+\tau^2}{r_o}\right)^M} \tag{14}$$

with constraints $r_o \neq 0$ and $M \in \mathbb{Z}^+$.

- Bessel distribution:

$$\phi(\theta,\tau) = \frac{J_1(2\pi\alpha\theta\tau)}{\pi\alpha\theta\tau} \tag{15}$$

where J_1 is the first kind Bessel function of order one and $\alpha > 0$ is a scaling factor.

- Generalized exponential distribution:

$$\phi(\theta,\tau) = \exp\left(-\left(\frac{\theta}{\theta_1}\right)^{2N}\left(\frac{\tau}{\tau_1}\right)^{2M}\right) \tag{16}$$

where N, M are positive integers, and θ_1, τ_1 are chosen such that $\phi(\theta_1,\tau_1) = \exp(-1)$.

Kernel functions

- S-method:

$$\phi(\theta,\tau) = P\left(-\frac{\theta}{2}\right) *_\theta \int_{-\infty}^{\infty} w\left(u+\frac{\tau}{2}\right) w^*\left(u-\frac{\tau}{2}\right) \exp(-j\theta u)du \tag{17}$$

where $*_\theta$ represents a convolution in θ, $P(\theta)$ is a smoothing function and $w(t)$ is a window function used for the STFT.

- Hyperbolic distribution:

$$\phi(\theta,\tau) = \frac{1}{\cosh(\beta\theta\tau)} \tag{18}$$

where β is a parameter to control the exponential terms of the hyperbolic function.

Kernel functions

- It is important to mention that all the kernels presented above, except the kernel for the Born-Jordan distribution, contain one or more adjustable parameters.
- This implies that for a given kernel the parameter(s) can be chosen such that the resulting kernel produces a representation similar to a representation obtained by some other kernel with the same number of parameters.
- Having the opportunity to "fine tune" the kernel generally represents an advantage, since the kernel can be optimized to achieve maximal reduction of the cross term effects.

Example

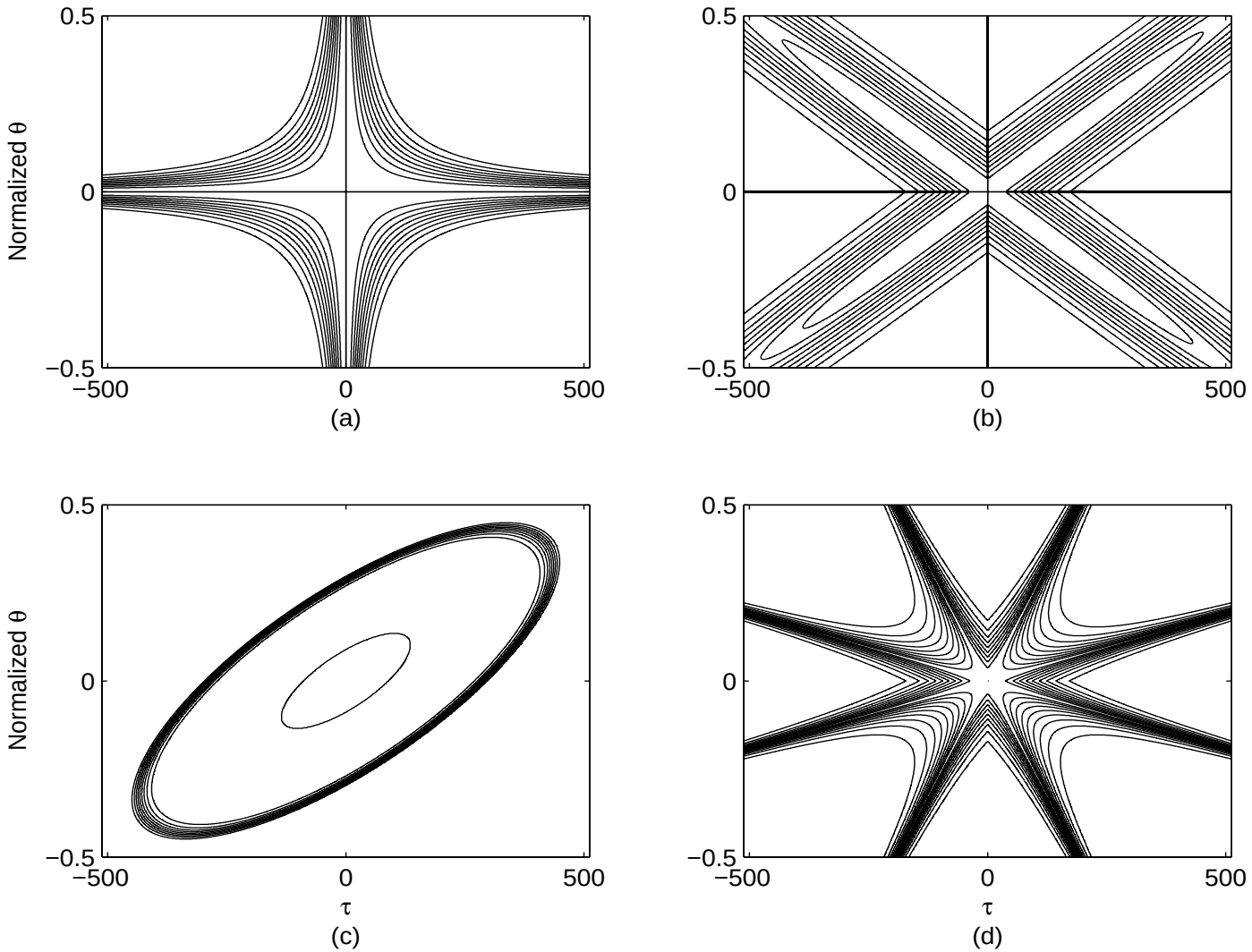


Figure 14: Sample kernels.

Rotated TFRs

- The rotation of the TF plane has been introduced to improve energy concentration for signals whose components are not aligned with either the time or the frequency axis.

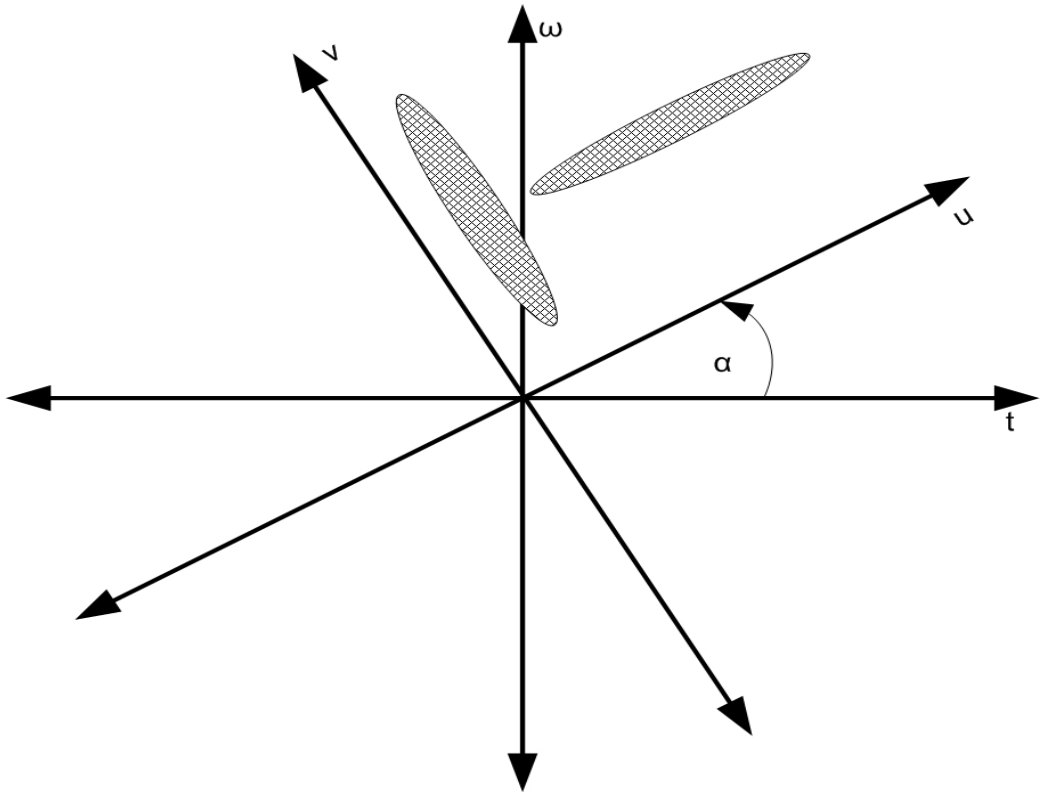


Figure 15: Rotated TFRs.

Pattern classification

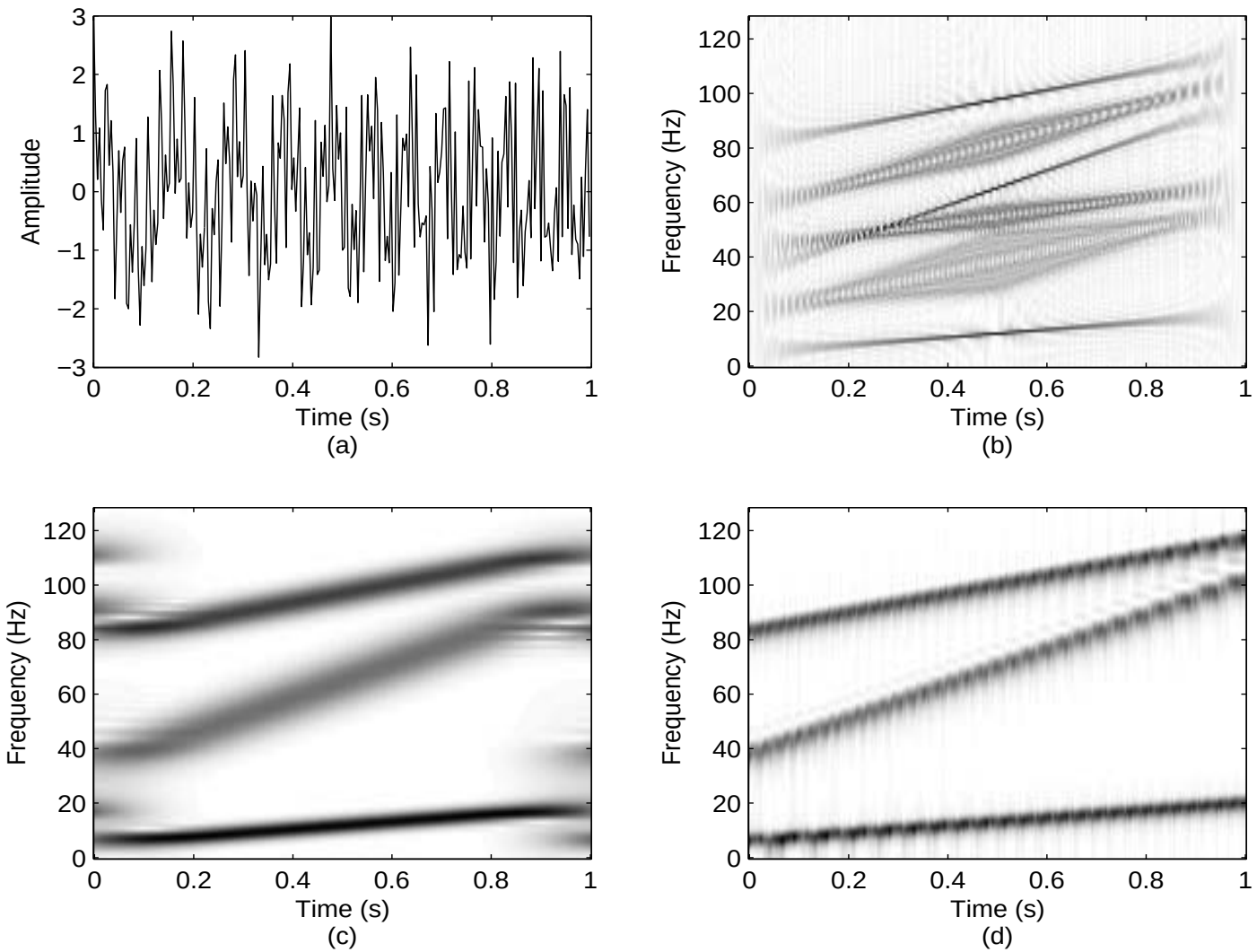


Figure 16: Rotated TFRs.

Rotated TFRs

The TFA based on the rotation of the TF plane can be achieved in several ways:

- Fractional Fourier transform (FRFT):

$$F_\alpha(u) = \begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j(u^2/2)\cot\alpha} \int_{-\infty}^{+\infty} x(t) e^{j(t^2/2)\cot\alpha - jut\csc\alpha} dt \\ \text{if } \alpha \text{ is not multiple of } \pi \\ x(t) \\ \text{if } \alpha \text{ is a multiple of } 2\pi \\ x(-t) \\ \text{if } \alpha + \pi \text{ is a multiple of } 2\pi \end{cases} \tag{19}$$

The standard Fourier transform is a special case of the FRFT with a rotation angle $\alpha = \pi/2$.

Rotated TFRs

- Local polynomial Fourier transform (LPFT):

$$LPFT_x(t,\overline{\omega}) = \int_{-\infty}^{+\infty} x(t+\tau)w(\tau)$$
$$\exp\left(-j\omega_1\tau - j\omega_2\tau^2/2 - ... - j\omega_M\tau^M/M\right)d\tau \qquad (20)$$

where $\overline{\omega} = (\omega_1, \omega_2, ..., \omega_M)$. The LPFT enables one to estimate both the time-varying frequency and its derivatives. The technique is based on fitting a local polynomial approximation of the frequency which implements a high-order nonparametric regression.

Rotated TFRs

- Radon-Wigner distribution (RWD):

$$RWD(r,\vartheta) = \mathcal{R}\left[WV_x(t,\omega)\right]$$
$$= \int WV_x(t,\omega_o + mt)dt\Big|_{m=-1/\tan(\vartheta);\omega_o=r/\sin(\vartheta)} \tag{21}$$

  where $\mathcal{R}\left[f(x,y)\right] = \int f(r\cos\vartheta - s\sin\vartheta; r\sin\vartheta + s\cos\vartheta)ds$ and r and s represent x and y axes rotated counterclockwise by an angle ϑ.

- Even though the RWD is considered a tool for the rotation of the TF plane at a certain angle, the RWD was developed primarily for detection and classification of multicomponent linear FM signals in noise.

Relations between rotated TFRs

- It is shown that the Radon-Wigner distribution is the squared modulus of the fractional Fourier transform:

$$RW\left[x(t)\right] = \left|FRFT\left[x(t)\right]\right|^2. \tag{22}$$

- To establish the relationship between the FRFT and the LPFT, the FRFT can be written as:

$$F_\alpha(u) = \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j(u^2/2)\cot\alpha} \int_{-\infty}^{+\infty} x_w(\tau) e^{j(\tau^2/2)\cot\alpha - ju\tau\csc\alpha} d\tau \tag{23}$$

where $x_w(\tau) = x(t+\tau)w(\tau)$. For $M = 2$, $\omega_1 = u\csc\alpha$, and $\omega_2 = \cot\alpha$ in (20), equation (23) can be expressed in terms of the LPFT as:

$$F_\alpha(u) = \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j(u^2/2)\cot\alpha} LPFT_x(t, \omega_1, \omega_2). \tag{24}$$

Properties of rotated TFRs

Approach	Advantages	Disadvantages
FRFT	Allows representation of a signal on the orthonormal basis formed by chirps.	$\cot(\alpha)$ can take enormous values and oversampling may be needed to satisfy the sampling theorem.
LPFT	Provides generalization of the FRFT to any order of the polynomial IF.	A drawback of the LPFT is the increase in dimensionality, i.e., an increase of the calculation complexity.
RWD	Excellent for establishing the direction of the linear FM modulated signal in the ambiguity plane.	Not suitable for long data records, and the segmentation of such records is needed. Analyzed in depth only for the WD.

Signal dependent TFRs

- The feature extractors described in the previous sections deal with several concepts regarding the improvement of energy concentration: reducing the effects of spectral leakage; diminishing the effects of cross terms; and aligning the axis of analysis with the principal axis of the signal components.
- Can a single feature extractor be optimal for all signals?
- Unfortunately not, since a major drawback of all fixed mappings is that, for each mapping, the resulting TFR is satisfactory only for a limited class of signals.
- Thus, the enhanced concentration in the TF domain is desirable for a variety of classes of signals.

Signal dependent TFRs

- Concentrated components generally overlap or interfere with other nearby components as little as possible, and yield a "sharp" representation.
- The maximal concentration also implies that components are confined as closely as possible to their proper support in the TF domain.
- Signal dependent TFRs are generally nonlinear and nonquadratic due to the nature of the computation process.
- They are based on:
  - concentration measures;
  - reassignment methods;
  - signal optimized kernels/windows.

Concentration measures

- The concentration measure approach examines the effects of certain parameter variations on the energy concentration of the signal in the TF domain.
- The parameter value yielding the highest energy concentration is chosen for the signal dependent TFR.
- The development of the concentration measure can be divided into two groups based either on the distribution norms or on the entropy of the distributions.

Concentration measures

- Jones and Parks proposed a measure based on the STFT for signal concentration that allows the fully automated determination of the optimal basis parameters:

$$CM = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |STFT(t,\omega)|^4 \, dtd\omega}{\left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |STFT(t,\omega)|^2 \, dtd\omega \right)^2}. \tag{25}$$

- CM in (25) favours those components with higher concentration.
- For multicomponent signals, a local measure is required to determine the concentration of the dominant component at each location in the TF domain.
- Eqn. (25) can be turned into a local concentration measure by multiplying the squared magnitude of the short-time Fourier transform by a "localization weighting function."

Concentration measures

- A solution to the problem in the Jones-Parks measure is proposed by Stanković.
- His CM does not discriminate low concentrated components with respect to the highly concentrated ones within the same distribution, and it is given by:

$$CM = \left(\sum_{k=1}^{N} \sum_{n=1}^{N} |TFR_x(n,k)|^{1/p} \right)^{p} \tag{26}$$

where $TFR_x(n,k)$ is a discrete version of any of the TFRs.

Concentration measures

- Williams *et al* considered how the information measures could be used to provide information on TFDs.
- The Shannon information measure is appropriate only for positive TFRs.
- The Rényi measure conforms closely to the visually based notion of complexity when inspecting TFRs and can be used for other TFRs.
- For Cohen's class of the TFRs, the Shannon information measure is given as

$$H(TF_x(t,\omega)) = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} TF_x(t,\omega)\log_2 TF_x(t,\omega)dtd\omega \quad (27)$$

and the Rényi measure as

$$R_\alpha(TF_x(t,\omega)) = -\frac{1}{1-\alpha}\log_2\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} TF_x(t,\omega)dtd\omega \quad (28)$$

where $\alpha > 0$, and the Shannon entropy is recovered as the limit of R_α, as $\alpha \to 1$.

Example

- A signal consisting of a sinusoidally FM and linear FM components, $x(t) = \exp(j20\pi t + j30\pi t^2) + \exp(j5\pi \cos(4\pi t) + j150\pi t)$.

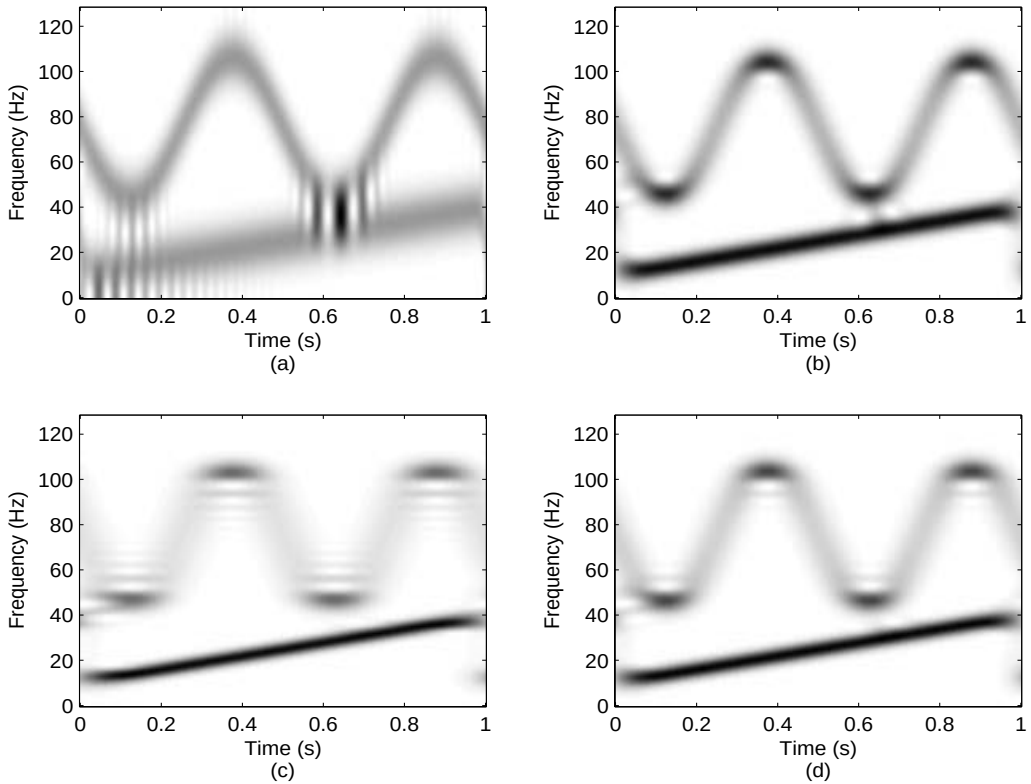


Figure 17: TFRs based on concentration measures.

Reassignment method

- The reassignment method creates a modified version of a representation by moving its values away from where they are computed to produce a better localization of the signal components.
- In order to perform such an operation, for each point in the TF plane, one calculates the center of gravity for the signal energy such as:

$$\widehat{t}(t,\omega) = t - \frac{\int\int uTFR(t-u,\omega-\Omega)dud\Omega}{\int\int TFR(t-u,\omega-\Omega)dud\Omega} \quad (29)$$

$$\widehat{\omega}(t,\omega) = \omega - \frac{\int\int \Omega TFR(t-u,\omega-\Omega)dud\Omega}{\int\int TFR(t-u,\omega-\Omega)dud\Omega}. \quad (30)$$

- Given these centers of gravities, the reassigned TFR is obtained by

$$RTFR(t,\omega) = \int\int TFR(\tau,\upsilon)\delta(t-\widehat{t}(\tau,\upsilon))\delta(\omega-\widehat{\omega}(\tau,\upsilon))d\tau d\upsilon \quad (31)$$

  where $\delta(t)$ is a Dirac function.
- The technique is highly sensitive to noise.
- The reassignment method is also computationally expensive.

Signal dependent kernels and windows

- The first two approaches to signal dependent TFRs are based upon the fact that an optimized representation is found for each new signal.
- Another stream of research in this area is based on the development of the signal dependent kernels/windows for a class of signals through an optimization design procedure.
- The initial research has been conducted for so-called radially Gaussian distributions.
- The problem of finding the optimized kernel boils down to finding the optimal $\sigma(\psi)$ for radially Gaussian functions for the given signal:

$$\max_{\Phi} \int_0^{2\pi} \int_0^{\infty} |A(r,\psi)\Phi(r,\psi)|^2 \, r dr d\psi \tag{32}$$

with a constraint that the energy of $\Phi(r,\psi)$ must be finite, where $r = \sqrt{\theta^2 + \tau^2}$ and $A(r,\psi)$ is the ambiguity function of the signal in the polar coordinates.

Properties of the approaches for obtaining signal dependent TFRs

Approach	Advantages	Disadvantages
Concentration measure	Usually easy to implement. Good energy concentration can be obtained.	It has to be calculated for each signal.
Reassignment methods	Excellent energy concentration can be obtained.	Computationally expensive. Sensitive to noise.
Signal optimized kernels/windows	It does not need recalculation for every signal, but it is rather based on class of signals.	Needs careful implementation when working with signals in noisy environment.

Pattern classification

- TFR-based classification methods are preferred because TFRs have discriminant capabilities for signals belonging to different signal classes.

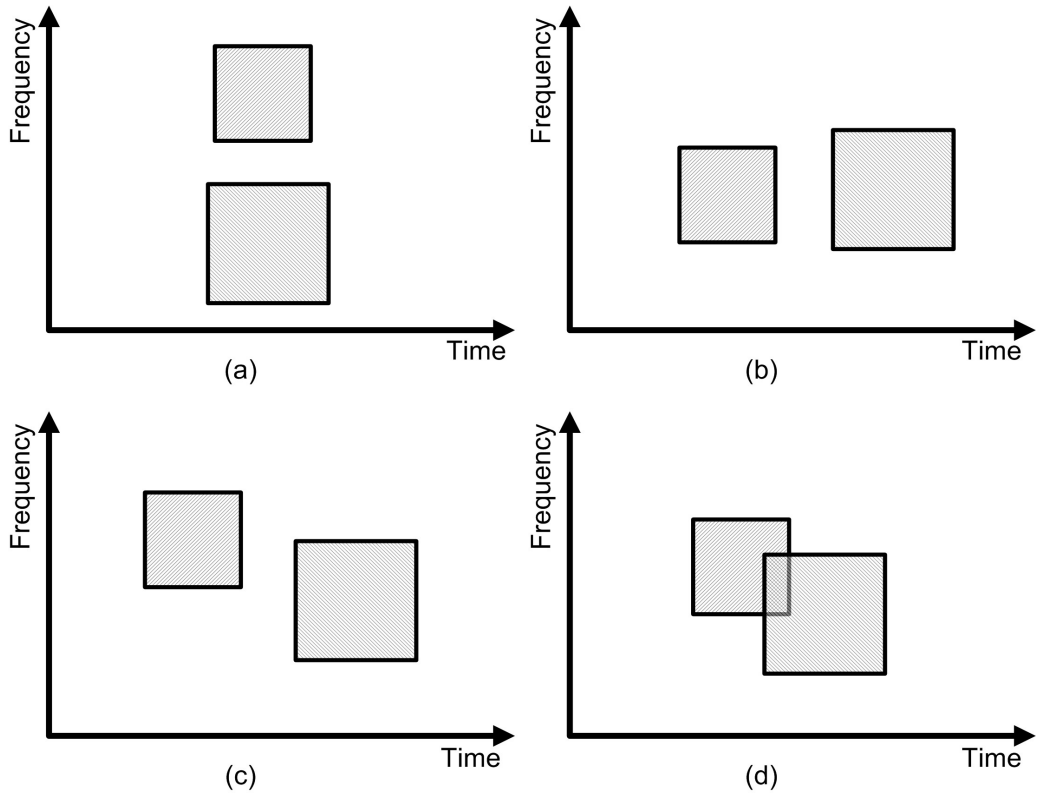


Figure 18: Sample patters in TF Domain.

Pattern classification

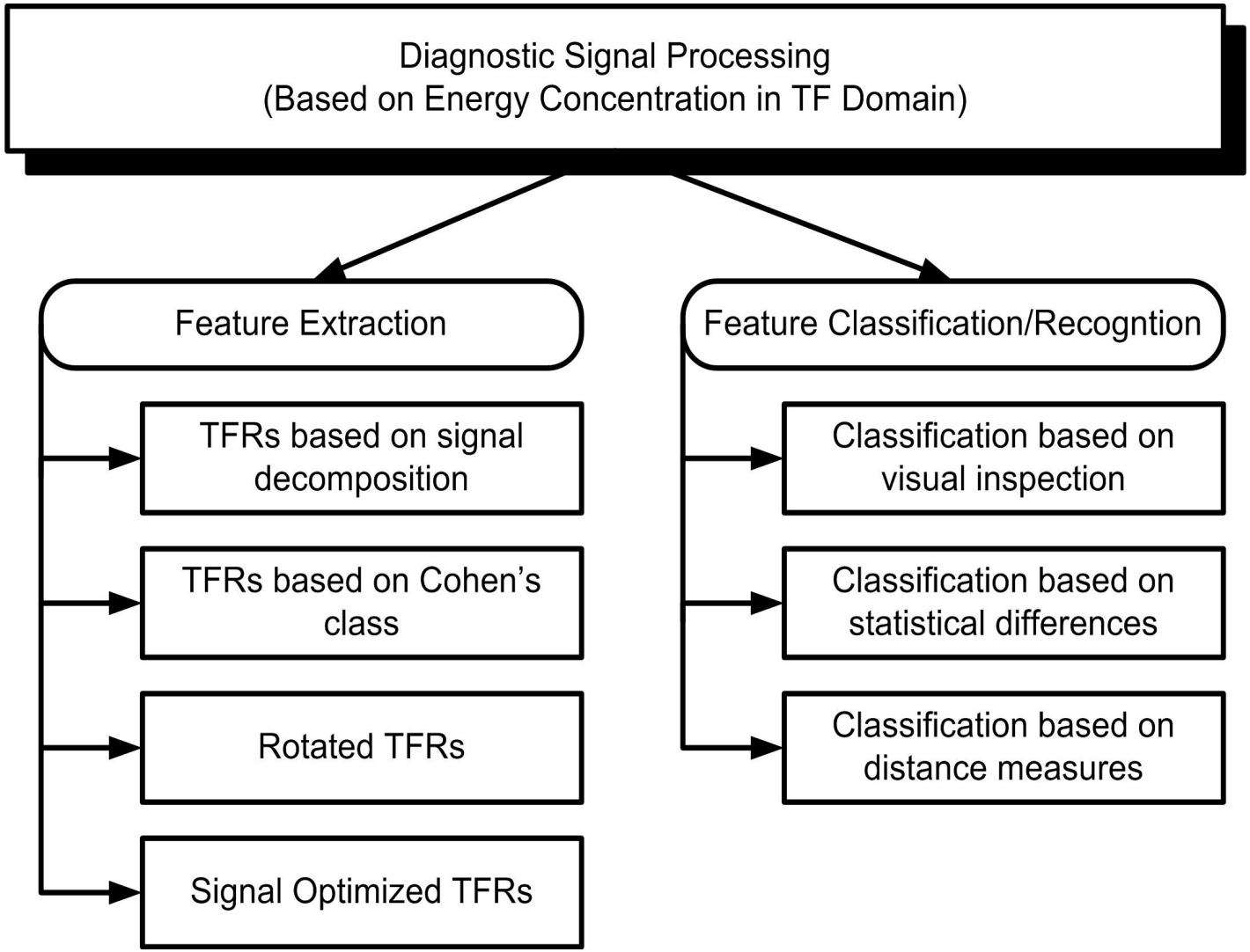


Figure 19: Pattern classification in TF domain.

IF estimation

- The IF can be estimated as a first moment of the TFR:

$$\omega(t) = \frac{\int_{-\infty}^{+\infty} \omega\, TF_x(t,\omega) d\omega}{\int_{-\infty}^{+\infty} TF_x(t,\omega) d\omega} \tag{33}$$

  or based on the position of the maximum value of the energy concentration in the TF domain as

$$\omega(t) = \arg\max_{\omega} \left[|\, TF_x(t,\omega)| \right]. \tag{34}$$

- The first moment provides an unbiased estimate of the IF of a signal.
- The presence of additive noise leads to the serious degradation of the first moment estimate.
- It may have a high statistical variance even at high values of input SNR.
- The first moment estimate is not affected by the multiplicative noise.

IF estimation

- The maximum value estimate is greatly affected by the multiplicative noise when the power spectral density of the noise has a maximum at a frequency other than DC.
- The maximum value estimate is hence used for the signals contaminated with the additive noise.
- It is based on the detection of a distribution maxima positions.
- This estimate is also prone to some estimation errors and the sources of estimation error are:
  - bias;
  - random deviation of the maxima within the auto-term caused by a small noise;
  - large random deviations due to false maxima detection outside the auto-term caused by a high noise.

IF estimation

- Consider the problem of the problem of the IF estimation from the discrete-time observations

$$x(nT) = m(nT) + \epsilon(nT) \tag{35}$$

with

$$m(t) = A\exp(j\phi(t)) \tag{36}$$

where n is integer, T is a sampling interval and $\epsilon(nT)$ complex-valued white Gaussian noise with i.i.d. real and imaginary parts.

- $\Re(\epsilon(nT))$ and $\Im(\epsilon(nT)) \sim N(0, \sigma_\epsilon^2/2)$ and the total variance of the noise is equal to σ_ϵ^2.
- The IF is the first derivative of the phase $(\omega(t) = \phi^{'}(t))$
- the IF estimate is a solution of the optimization problem

$$\widehat{\omega}_h(t) = \arg\left[\max_{\omega \in Q_\omega} TF_x(\omega, t)\right] \tag{37}$$

IF estimation

- The estimation error, at a time-instant t, is defined as

$$\Delta\widehat{\omega}(t) = \omega(t) - \widehat{\omega}_h(t) \tag{38}$$

- $\Delta\widehat{\omega}(t)$ is a random variable and is characterized by its bias and variance.
- The stationary point of $TF_x(t,\omega)$ is determined by the zero value of the derivative $TF_x(t,\omega)$, which is given as

$$\frac{\partial TF_x(t,\omega)}{\partial\omega} = 0 \tag{39}$$

IF estimation

- The linearization of $\partial TF_x(t,\omega)/\partial\omega = 0$ with respect to the small
  1) estimation error $\Delta\widehat{\omega}_h(t)$
  2) residual of the phase deviation $\Delta\phi$;
  3) noise ϵ;

  gives

$$\frac{\partial TF_x(t,\omega)}{\partial\omega}|_0 + \frac{\partial^2 TF_x(t,\omega)}{\partial\omega^2}|_0\Delta\widehat{\omega}(t) + \frac{\partial TF_x(t,\omega)}{\partial\omega}|_{0\delta_{\Delta\phi}} + \frac{\partial TF_x(t,\omega)}{\partial\omega}|_{0\delta_\epsilon} = 0 \quad (40)$$

  where $|_0$ mean that the corresponding derivatives are calculated at the point $\omega = \phi^{'}(t)$, $\Delta\phi = 0$ and $\epsilon(nT) = 0$.

IF estimation

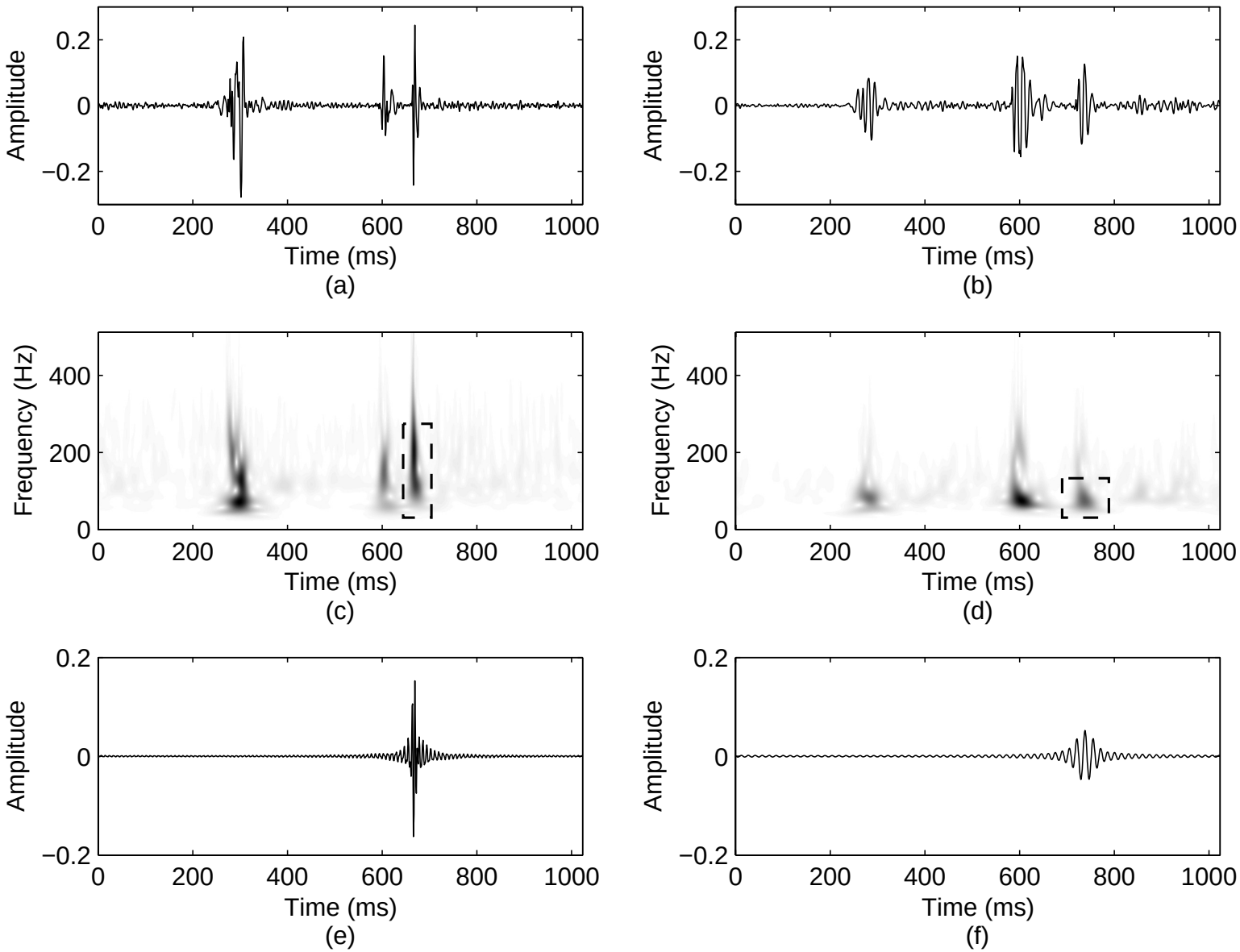
- The estimation bias is

$$E\left(\Delta\widehat{\omega}_h(t)\right) = \frac{-\frac{\partial TF_x(t,\omega)}{\partial\omega}|_0 - \frac{\partial TF_x(t,\omega)}{\partial\omega}|_{0\delta_{\Delta\phi}} - E\left\{\frac{\partial TF_x(t,\omega)}{\partial\omega}|_{0\delta_\epsilon}\right\}}{\frac{\partial^2 TF_x(t,\omega)}{\partial\omega^2}} \tag{41}$$

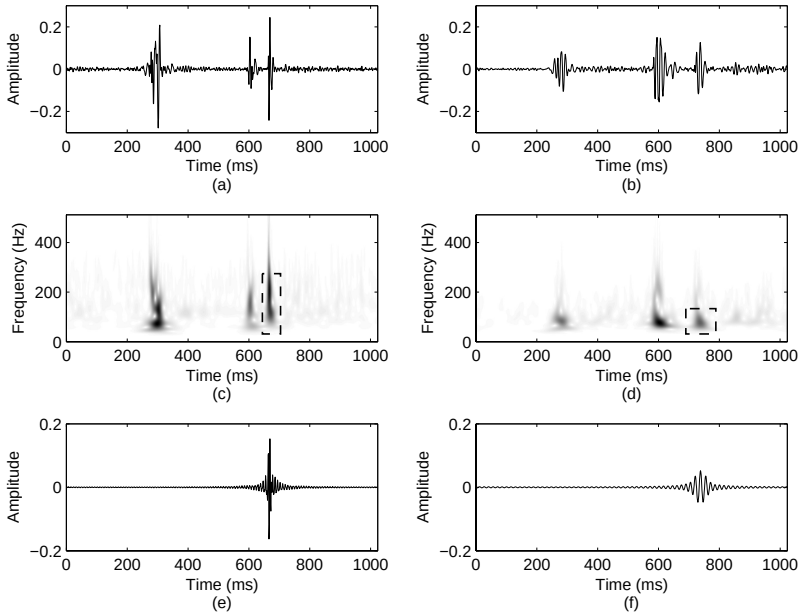
and the variance is of the following form

$$var\left(\Delta\widehat{\omega}_h(t)\right) = \frac{E\left\{\left(\frac{\partial TF_x(t,\omega)}{\partial\omega}|_{0\delta_\epsilon}\right)^2\right\} - E\left\{\frac{\partial TF_x(t,\omega)}{\partial\omega}|_{0\delta_\epsilon}\right\}}{\frac{\partial^2 TF_x(t,\omega)}{\partial\omega^2}} \tag{42}$$

Selective Regional Correlation (SRC)

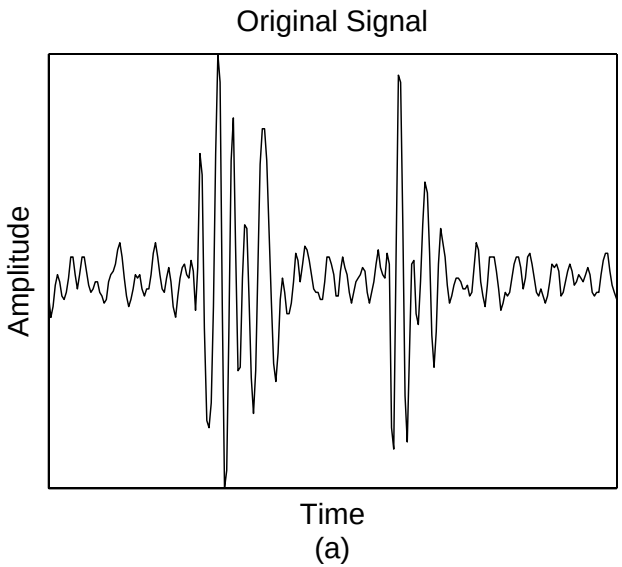


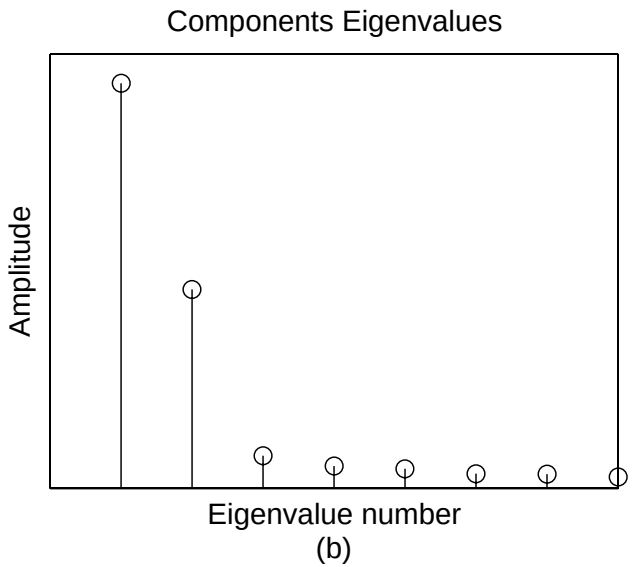
Selective Regional Correlation (SRC)

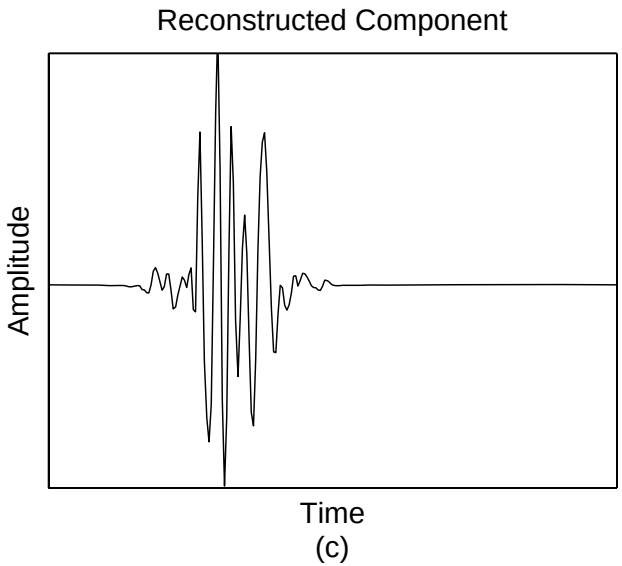


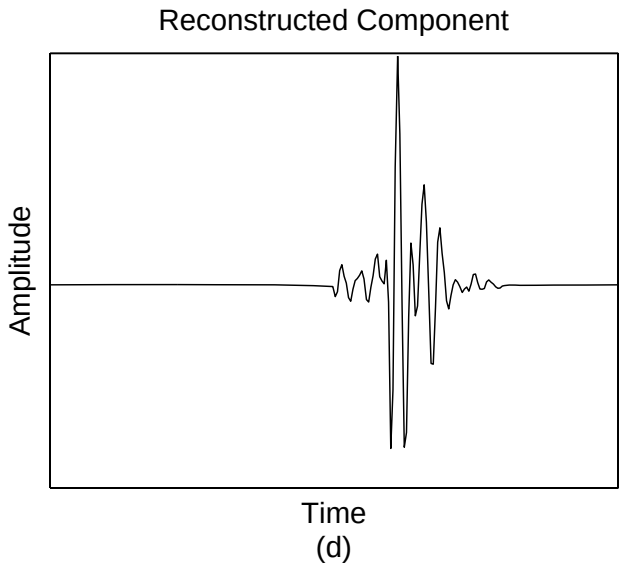
Method		ρ_M	ρ_{NM}	φ	EP
SRC	STFT	0.5533	0.2413	0.3120	10.00%
	CWT	0.6731	0.5248	0.1483	16.67%
	S-transform	0.4905	0.1661	0.3244	6.670%
General Correlation		0.3477	0.3352	0.0126	56.67%

Segmentation of heart sounds









Further reading

- L. Cohen, "Time-frequency distribution - a review," *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941-981, Jul. 1989.
- O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14-38, Oct. 1991.
- F. Hlawatsch and G. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21-67, Apr. 1992.
- N. Hess-Nielsen and M. V. Wickerhauser, "Wavelets and time-frequency analysis," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 523-540, Apr. 1996.
- S. Qian and D. Chen, "Joint time-frequency analysis," *IEEE Signal Processing Magazine*, vol. 16, no. 2, pp. 52-67, Mar. 1999.
- E. Sejdić, I. Djurović, J. Jiang, "Time-frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153-183, January 2009.