OPTIMAL MULTIPLE WINDOW TIME-FREQUENCY ANALYSIS OF LOCALLY STATIONARY PROCESSES

Maria Hansson and Patrik Wahlberg

Mathematical Statistics, Centre for Mathematical Sciences and Signal Processing Group, dept. of Electroscience
Box 118, SE-221 00 Lund, Sweden
phone: +46 46 222 49 53, email: mh@maths.lth.se, pw@es.lth.se

ABSTRACT

This paper investigates the multiple windows of the mean squared error optimal time-frequency kernel for estimation of the Wigner-Ville spectrum. The kernel is optimal for a certain locally stationary process where the covariance function is determined by two one-dimensional Gaussian functions. The multiple windows are obtained as the eigenvectors of the rotated time-lag estimation kernel. The spectrograms from the different windows are weighted with the eigenvalues and the resulting multiple window spectrogram is an estimate of the optimal smoothed Wigner-Ville spectrum.

1. INTRODUCTION

The area of time-frequency analysis is well covered in the signal processing literature. A lot of work is done concerning deterministic signals disturbed by noise but the area of time-frequency analysis of time-varying stochastic processes remain limited, [1].

The Wigner-Ville spectrum (WVS) of a process exists due to the assumption of harmonizability. It can be estimated from realizations of the process using Cohen's class of time-frequency (TF) representations, which are determined by a TF kernel function. The mean squared error optimal solution to this problem has been obtained by Sayeed and Jones [2]. Instead of calculating the two-dimensional convolution between the optimal kernel and the Wigner-Ville distribution (WVD) of a process realization, the calculations can be simplified with use of multiple windows. The time-lag estimation kernel connected to a optimal TF kernel is rotated and smoothed and the corresponding eigenvectors and eigenvalues are calculated. The estimate of the WVS is given as the weighted sum of the the spectrograms of the data with the different eigenvectors as sliding windows. The weights are the eigenvalues, [3, 4]. The fewer eigenvalues that significantly differ from zero the fewer spectrograms to compute. The phrase multiple windows were originally introduced by Thomson [5] for the case of stationary processes with smooth spectra. For varying spectra however, e.g., spectra with peak and notches, the performance of the Thomson multiple window method degrades due to cross-correlation between spectra, [6]. Other methods have been invented, [7, 8] and multiple windows are also introduced for the case of time-variable processes, [9, 10].

We have studied optimal TF kernels for a class of certain locally stationary processes (LSP), where the covariance function is determined by two one-dimensional Gaussian functions. In [11] a formula was derived, valid for LSPs, for the optimal kernel in the ambiguity domain. In this paper, we present the corresponding eigenvectors and eigenvalues connected to this kernel and study the performance of the multiple window spectrogram estimate.

This project is supported by the Swedish Research Council.

2. OPTIMAL ESTIMATION OF WIGNER-VILLE SPECTRA

The Wigner-Ville spectrum (WVS) of $X(t)$, [1], is defined by

$$W_x(t,\omega) := \int r_x(t+\tau/2, t-\tau/2) e^{-i\omega\tau} d\tau.$$

It exists due to the assumption of harmonizability and can be estimated from process realizations using Cohen's class of time-frequency representations,

$$\hat{W}_x(t,\omega) := W * \Phi(t,\omega) =$$

$$= \frac{1}{2\pi} \int\int W(t-t', \omega-\omega') \Phi(t',\omega') dt' d\omega', \qquad (1)$$

where $\hat{W}_x(t,\omega)$ is the estimate of the WVS, $W(t,\omega)$ is the signal's Wigner-Ville distribution (WVD) and $\Phi(t,\omega)$ is a time-frequency estimation kernel. Sayeed and Jones [2] derived the optimal kernel in the mean square error sense, i.e. minimizing the integrated expected squared error

$$J(\Phi) = \int\int E|\hat{W}_x(t,\omega) - W_x(t,\omega)|^2 dt d\omega. \qquad (2)$$

In the ambiguity domain the Fourier transform converts the convolution (1) into a multiplication

$$\hat{A}_x(\theta,\tau) = A(\theta,\tau) \cdot \phi(\theta,\tau),$$

where the expected ambiguity function of the process is defined by

$$A_x(\theta,\tau) := \int r_x(t+\tau/2, t-\tau/2) e^{-i\theta t} dt.$$

Its Fourier transform is actually W_x,

$$W_x(t,\omega) = \frac{1}{2\pi} \int\int A_x(\theta,\tau) e^{i(\theta t - \tau\omega)} d\theta d\tau =$$

$$= \mathcal{F}_1^{-1} \mathcal{F}_2 \{A_x\},$$

where $\mathcal{F}_{1,2}$ denotes Fourier transform in the first and second variable. It is denoted by capital letters and defined by

$$F(\omega) := (\mathcal{F} f)(\omega) := \int f(t) e^{-i\omega t} dt.$$

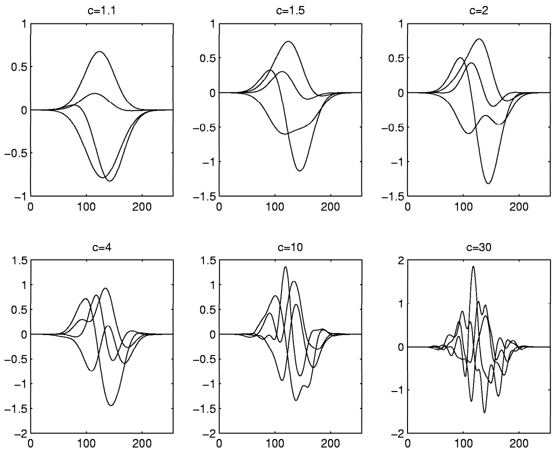

Figure 1: Four different realizations of the Gaussian locally stationary process for different values of c.

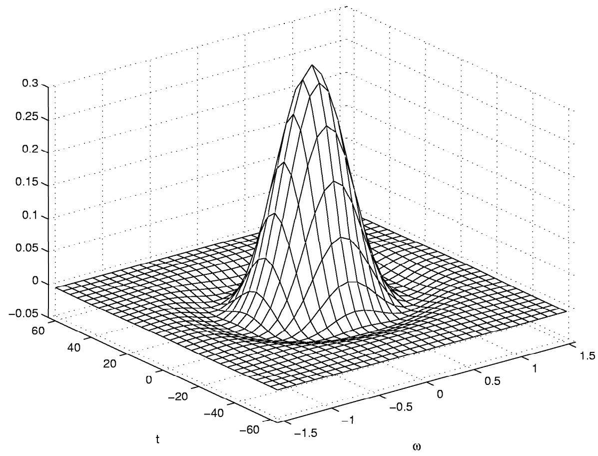

Figure 2: The TF kernel for $c = 30$.

The optimal kernel was deduced to be,

$$\phi_{opt}(\theta,\tau) = \frac{|A_x(\theta,\tau)|^2}{E[|A(\theta,\tau)|^2]}, \tag{3}$$

in the ambiguity domain, [11]. The time-frequency kernel is computed from the ambiguity domain kernel by a Fourier transformation

$$\Phi_{opt}(t,\omega) = \frac{1}{2\pi}\iint \phi_{opt}(\theta,\tau)e^{i(\theta t-\tau\omega)}d\theta d\tau. \tag{4}$$

3. OPTIMAL KERNEL OF LOCALLY STATIONARY PROCESSES

A locally stationary process (LSP) [12] has, per definition, a covariance function determined by two functions q, r and has the form

$$r_x(t,s) = q\left(\frac{t+s}{2}\right)\cdot r(t-s). \tag{5}$$

It can be shown that q can be taken to be non-negative, and r is non-negative definite (i.e. the covariance of a stationary process) [1, 12]. The normalization $r(0) = 1$ is used without loss of generality (any other constant can be incorporated into m). Such processes do exist [12].

When $q(\tau) = e^{-\tau^2/2}$ is a fix Gaussian function and $r(\tau) = e^{-\frac{c}{4}\tau^2/2}$ is a variable Gaussian function, $r_x(t,s)$ is a covariance if and only if $c \geq 1$, [11]. The limit case $c = 1$ results in a deterministic Gaussian function with stochastic amplitude. Examples of the processes is depicted in Figure 1 for different values of c.

The expected ambiguity function of a locally stationary process is separable (rank one),

$$A_x(\theta,\tau) = Q(\theta)r(\tau).$$

Restricting to circularly symmetric Gaussian distributed processes, the denominator of (3) reduces to

$$E[|A(\theta,\tau)|^2] = |A_x(\theta,\tau)|^2 + (\mathcal{F}_1^{-1}\mathcal{F}_2\{|A_x|^2\})(\theta,\tau).$$

For the case of LSPs this gives $E[|A(\theta,\tau)|^2] =$

$$= |Q(\theta)|^2\cdot|r(\tau)|^2 + (\mathcal{F}\,|r|^2)(\theta)\cdot(\mathcal{F}^{-1}|Q|^2)(\tau).$$

The optimal kernel (3) is thus $\phi_{opt}(\theta,\tau) =$

$$\begin{aligned} &= \frac{|Q(\theta)|^2|r(\tau)|^2}{|Q(\theta)|^2|r(\tau)|^2+(\mathcal{F}\,|r|^2)(\theta)(\mathcal{F}^{-1}|Q|^2)(\tau)} \\ &= \frac{1}{1+c^{-1/2}e^{(1-\frac{1}{c})\theta^2+\frac{c-1}{4}\tau^2}}. \end{aligned} \tag{6}$$

With use of (4) the kernel in the TF domain is determined. An example of $\Phi_{opt}(t,\omega)$ is depicted in Figure 2. For smaller values of c the TF kernel becomes more narrow with a larger amplitude. The limit case $c = 1$ results in $\phi_{opt}(\theta,\tau) \equiv \frac{1}{2} \Leftrightarrow \Phi_{opt}(t,\omega) = \frac{1}{2}\delta(t,\omega)$.

4. MULTIPLE WINDOW TIME-FREQUENCY ANALYSIS

The optimal multiple windows for a time-variable process are obtained as the eigenvectors of a rotated time-lag estimation kernel, [3, 4]. The time-lag estimation kernel is calculated as the inverse Fourier transform of (6) in the first variable,

$$R_{opt}(t,\tau) = \mathcal{F}_1^{-1}\phi_{opt}(\theta,\tau).$$

For implementation and calculation of the multiple windows, the optimal kernel is sampled,

$$\phi^d_{opt}(k_1,l_2) = \phi_{opt}\left(\frac{2\pi(k_1-K-1)}{2K}M, \frac{l_2-L-1}{M}\right),$$

for $k_1 = 1\ldots 2K$ and $l_2 = 1\ldots 2L$ and M is a scaling variable to sample a proper area of the kernel. Using the inverse discrete Fourier transform in the first variable gives the discrete time-lag estimation kernel $R^d_{opt}(l_1,l_2)$, $l_1 = 1\ldots 2L$ and $l_2 = 1\ldots 2L$. The rotation is made via the transformation, $Rrot^d_{opt}(n_1,n_2)$

$$= R^d_{opt}\left(\frac{|(n_1-1)+(n_2-1)|}{2}+1, \frac{|(n_1-1)-(n_2-1)|}{2}+1\right) \tag{7}$$

when both n_1 and n_2 are odd or when they both are even. When n_1 is odd and n_2 is even or the opposite case we approximate the covariance function as,

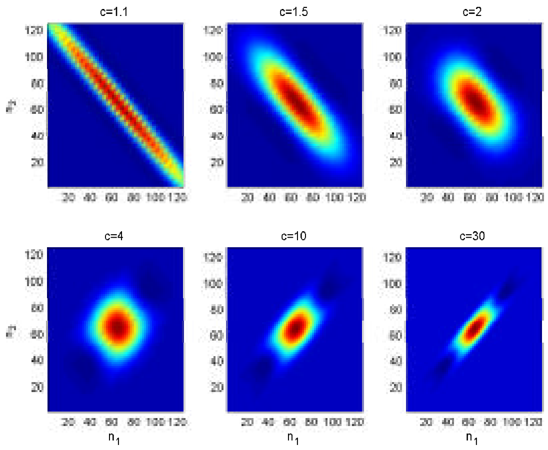

Figure 3: The rotated time-lag estimation kernel for different values of c.

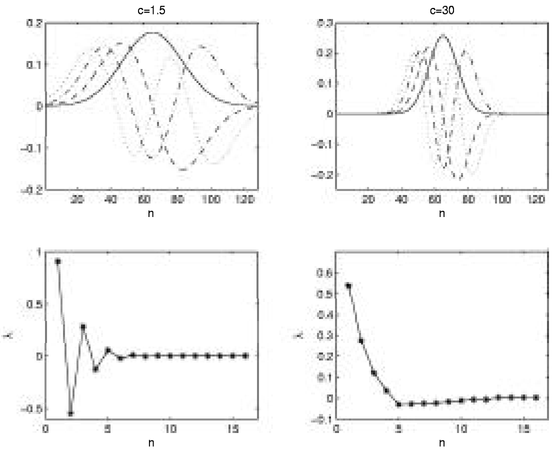

Figure 4: Examples of eigenvectors (upper figures) and eigenvalues (lower figures) for two different values of c.

$$Rrot^d_{opt}(n_1,n_2) = \frac{1}{4}Rrot^d_{opt}(n_1-1,n_2) + \frac{1}{4}Rrot^d_{opt}(n_1+1,n_2)$$

$$+\frac{1}{4}Rrot^d_{opt}(n_1,n_2-1) + \frac{1}{4}Rrot^d_{opt}(n_1,n_2+1)$$

except at the edges of the matrix where just three values can be averaged. In figure 3 $Rrot^d_{opt}(n_1,n_2)$ is depicted for different values of c.

Computing the eigenvectors and eigenvalues of the time-lag estimation kernel and calculating the sum of eigenvalue-weighted WVD of the different windows will give the optimal discrete TF-kernel, $\Phi^d_{opt}(l_1,k_2)$, [3, 4]. If we use each eigenvector as a sliding window and compute the spectrograms of data, these spectrograms weighted with the eigenvalues and summed will give an mean squared error optimal estimate of the WVS of the sequence.

In figure 4 the four first eigenvectors (upper figures) and the largest eigenvalues (lower figures) for two different values of c are shown. In the sampling process the parameters are $L = 64$, $K = 256$ and $M = 10$. The eigenvectors which are of length $2L = 128$ are very similar to hermite functions. The eigenvalues for $c = 1.5$ are alternating between positive and negative values where for $c = 30$ the largest eigenvalues are positive. The number of eigenvalues that differ significantly from zero is the number of spectrograms that must be calculated and averaged. In figure 5, the six first eigenvalues are depicted as functions of the parameter c. For small values of c there are several eigenvalues that differ significantly from zero. For larger c, a few eigenvalues, typically 3 or 4 differ significantly from zero.

5. CALCULATION OF MEAN SQUARED ERROR

The bias and variance of a process can be computed if the covariance matrix is known. We calculate the time-variable covariance matrix of the locally stationary process as,

$$\begin{aligned} R_x(n_1,n_2) &= q\left(\frac{n_1-2L-1}{2M}+\frac{n_2-2L-1}{2M}\right)\cdot \\ & \quad r\left(\frac{n_1-2L-1}{M}-\frac{n_2-2L-1}{M}\right), \end{aligned} \qquad (8)$$

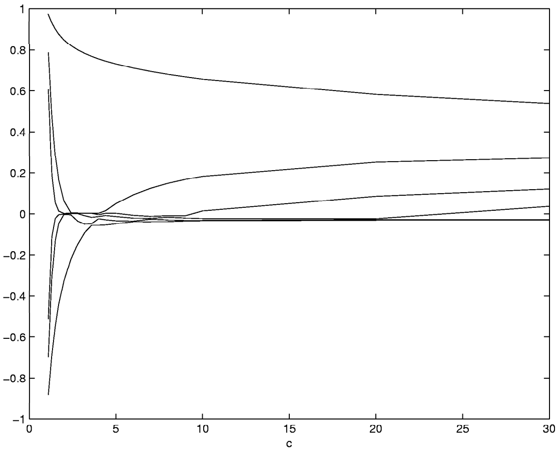

Figure 5: The six first eigenvalues as a function of c.

where $n_1 = 1\ldots 4L$ and $n_2 = 1\ldots 4L$. Bias of the estimated WVS is then defined as

$$\text{Bias } \hat{W}_x(l_1,k_2) = E[\hat{W}_x(l_1,k_2)] - W_x(l_1,k_2),$$

where $W_x(l_1,k_2)$ is the WVS of the process. The expected value of the estimate of the WVS is calculated to be

$$E[\hat{W}_x(l_1,k_2)] = \sum_{i=1}^{I} \lambda_i \mathbf{h}_i^T \mathbf{\Phi}^H(k_2)\mathbf{R}_x^{l_1}\mathbf{\Phi}(k_2)\mathbf{h}_i,$$

where the values of the matrix $\mathbf{R}_x^{l_1}$ is the $2L\times 2L$ sub-matrix from the diagonal of (8), $\mathbf{R}_x^{l_1} = R_x(l_1+L : l_1+2L-1, l_1+L : l_1+2L-1)$ for $l_1 = 1\ldots 2L$. The windows $\mathbf{h}_i$ are the eigenvectors and λ_i are the eigenvalues of the rotated optimal time-lag estimation kernel and $\mathbf{\Phi}(k_2) = \text{diag}[1, e^{-j2\pi\frac{k_2}{2K}},\ldots,e^{-j2\pi(2K-1)\frac{k_2}{2K}}]$ is the Fourier transform matrix. The variance of the WVS estimate is given by all combinations of the different subspectra covariances,

$$\text{Variance } \hat{W}_x(l_1,k_2) = \sum_{j=1}^{I}\sum_{i=1}^{I} \lambda_j\lambda_i \text{cov}(\hat{W}_i(l_1,k_2)\hat{W}_j(l_1,k_2)),$$

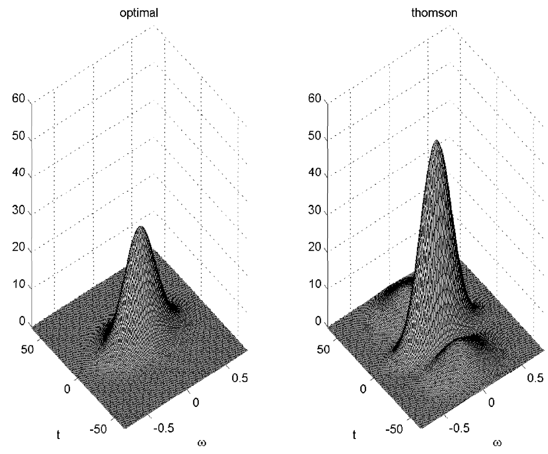

Figure 6: The mean squared error mse $\hat{W}_x(l_1,k_2) = E|\hat{W}_x(l_1,k_2) - W_x(l_1,k_2)|^2$, for the optimal windows and Thomson windows.

where

$$\operatorname{cov}(\hat{W}_i(l_1,k_2)\hat{W}_j(l_1,k_2)) = |\mathbf{h}_i^T \mathbf{\Phi}^H(k_2)\mathbf{R}_x^{l_1}\mathbf{\Phi}(k_2)\mathbf{h}_j|^2.$$

The calculated variance and bias could be combined to the mean squared error as

$$\text{mse}\,\hat{W}_x(l_1,k_2) = \text{Variance}\,\hat{W}_x(l_1,k_2) + (\text{Bias}\,\hat{W}_x(l_1,k_2))^2.$$

The windows and weighting factors optimal for the LSP process with $c = 30$ are calculated and utilized to estimate the mean squared error of the estimate of the WVS of this process. The window length is $2L = 128$ and for comparison the mean squared error is also calculated when using the Thomson windows for the case where the number of windows are $I = 8$ and the resolution $B = 0.08$. The result are shown in Figure 6.

6. CONCLUSIONS

The covariance function of a locally stationary process is determined by two Gaussian functions q and r. The optimal kernel for this process can be calculated. An optimal multiple window estimate of the WVS with few windows can be computed as the sum of weighted spectrograms. The windows and weighting factors are calculated as the eigenvectors and eigenvalues of the rotated optimal time-lag estimation kernel. We study the eigenvectors and eigenvalues of this kernel and calculate the mean squared error of the estimate of the WVS for the optimal windows and the Thomson windows.

REFERENCES

[1] P. Flandrin, *Time-frequency/Time-scale analysis*, Academic Press, 1999.

[2] A. M. Sayeed and D. L. Jones, "Optimal kernels for nonstationary spectral estimation," *IEEE Trans. on Sig. Proc.*, vol. 43, pp. 478–491, Feb. 1995.

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

[4] M. G. Amin, "Spectral Decomposition of Time-Frequency Distribution Kernels", *IEEE Trans. on Sig. Proc.*, vol. 42, no. 5, pp. 1156–1165, May 1994.

[5] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. of the IEEE*, vol. 70, no. 9, pp. 1055–1096, Sept 1982.

[6] A. T. Walden, E. McCoy, and D. B. Percival, "The variance of multitaper spectrum estimates for real gaussian processes,"*IEEE Trans. on Sig. Proc.*, vol. 42, no. 2, pp. 479–482, Feb 1994.

[7] K. S. Riedel and A. Sidorenko, "Minimum bias multiple taper spectral estimation", *IEEE Trans. on Sig. Proc.*, vol. 43, no. 1, pp. 188–195, Jan 1995.

[8] M. Hansson and G. Salomonsson, "A multiple window method for estimation of peaked spectra," *IEEE Trans. on Sig. Proc.*, vol. 45, no. 3, pp. 778–781, March 1997.

[9] L. L. Scharf and B. Friedlander, "Toeplitz and Hankel Kernels for Estimating Time-Varying Spectra of Discrete-Time Random Processes", *IEEE Trans. on Sig. Proc.*, vol. 49, no. 1, pp. 179–189, Jan. 2001.

[10] F. Çakrak and P. J. Loughlin, "Multiple Window Time-Varying Spectral Analysis ", *IEEE Trans. on Sig. Proc.*, vol. 49, no. 2, pp. 448–453, Feb. 2001.

[11] P. Wahlberg and M. Hansson, "Optimal time-frequency kernels for spectral estimation of locally stationary processes," in *Proc. IEEE Workshop on Statistical Signal Processing*, St. Louis, USA, 2003.

[12] R. A. Silverman, "Locally stationary random processes," *IRE Trans. Info. Theory.*, vol. 3, pp. 182–187, Sept. 1957.