

OPTIMAL TIME-FREQUENCY KERNELS FOR SPECTRAL ESTIMATION OF LOCALLY STATIONARY PROCESSES

Patrik Wahlberg and Maria Hansson

Signal Processing Group, Dept. Electrosience, Lund University,
Box 118, SE-221 00 Lund, Sweden, email pw@es.lth.se.

ABSTRACT

This paper investigates the mean square error optimal time-frequency kernel for estimation of the Wigner-Ville spectrum of a certain class of nonstationary processes. The class of locally stationary processes have a simplified covariance structure which facilitates analysis. We give a formula for the optimal kernel in the ambiguity domain and conditions that are sufficient for the optimal time-frequency kernel to be a continuous function, decaying at infinity.

1. INTRODUCTION

Generalization of spectrum estimation from stationary to nonstationary stochastic processes is a delicate problem. First of all there is a problem of how to *define* the spectral density of a nonstationary process. It turns out that not all good properties of the stationary spectral density, e.g. non-negativity and marginal properties, can be preserved for nonstationary processes. The Wigner-Ville spectrum (WVS) has however many nice properties and is a generalization of the stationary case [1].

The WVS of a process 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 square error optimal solution to this problem has been obtained by Sayeed and Jones [2].

The optimal kernel is sometimes not an ordinary function but a distribution, containing Dirac measures. If the kernel is a Dirac measure at the origin (a case studied in [2]), no time-frequency averaging is performed, and the optimal solution is to leave the time-frequency representation unaltered.

We study optimal kernels for a class of locally stationary processes (LSP). By this term we mean that the covariance function has a certain structure (as opposed to a large number of researchers who use it to mean "almost stationary" in various senses). The covariance function of a LSP is determined by two one-dimensional functions.

We derive a formula, valid for LSPs, for the optimal kernel in the ambiguity domain. We give conditions on the pair of functions constituting the covariance of a LSP, that are sufficient for the optimal TF kernel to be a continuous function, decaying at infinity. We restrict to circularly symmetric Gaussian processes.

2. OPTIMAL ESTIMATION OF WIGNER-VILLE SPECTRA

The autocovariance function for a zero mean stochastic process $X(t)$ is defined by

$$r_x(t, s) := E\{X(t)\overline{X(s)}\}, \quad (1)$$

and the process is called *harmonizable* [3] if r_x has a Fourier-Stieltjes representation

$$r_x(t, s) = \frac{1}{2\pi} \iint e^{i(t\xi - s\eta)} dR(\xi, \eta) \quad (2)$$

where dR is a measure of bounded variation. The class of stationary processes is included in the class of harmonizable, since in this case dR is diagonal, and a positive measure dR_1 then fulfills $dR(\xi, \eta) = dR_1(\xi)\delta(\xi - \eta)$, according to Bochner's theorem.

The WVS of $X(t)$ [4] is defined by

$$W_E(t, \omega) := \int r_x(t + \tau/2, t - \tau/2) e^{-i\omega\tau} d\tau. \quad (3)$$

It exists due to the assumption of harmonizability. It follows immediately that the WVS reduces to the spectral density function when it is restricted to weakly stationary processes, since $r_x(t, s) = r_1(t - s)$ is a function of $t - s$ only then. The expected ambiguity function of the process is defined by

$$A_E(\theta, \tau) := \int r_x(t + \tau/2, t - \tau/2) e^{-i\theta t} dt, \quad (4)$$

and its Fourier transform is actually W_E ,

$$W_E(t, \omega) = \frac{1}{2\pi} \iint A_E(\theta, \tau) e^{i(\theta t - \tau\omega)} d\theta d\tau =$$

This project is supported by the Swedish Research Council.

$$= \mathcal{F}_1^{-1} \mathcal{F}_2 \{A_E\}, \quad (5)$$

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

$$F(\omega) := (\mathcal{F}f)(\omega) := \int f(t)e^{-i\omega t} dt, \quad (6)$$

$$(\mathcal{F}^{-1}F)(t) := \frac{1}{2\pi} \int F(\omega)e^{i\omega t} d\omega. \quad (7)$$

The signal's Wigner-Ville distribution (WVD) is defined by the stochastic integral

$$W(t, \omega) := \int X(t + \tau/2) \overline{X(t - \tau/2)} e^{-i\omega\tau} d\tau, \quad (8)$$

and $E\{W(t, \omega)\} = W_E(t, \omega)$ holds under certain conditions on the process' fourth order moments [2] which we assume fulfilled. The ambiguity process is defined by

$$A(\theta, \tau) := \int X(t + \tau/2) \overline{X(t - \tau/2)} e^{-i\theta t} dt. \quad (9)$$

The WVS can be estimated from process realizations using Cohen's class of time-frequency representations,

$$\begin{aligned} \widehat{W}_E(t, \omega) &:= W * \Phi(t, \omega) = \\ &= \frac{1}{2\pi} \iint W(t - t', \omega - \omega') \Phi(t', \omega') dt' d\omega', \end{aligned} \quad (10)$$

where Φ 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

$$J(\Phi) = \iint E|\widehat{W}_E(t, \omega) - W_E(t, \omega)|^2 dt d\omega. \quad (11)$$

In the ambiguity domain the Fourier transform converts the convolution (10) into a multiplication

$$\widehat{A}_E(\theta, \tau) = A(\theta, \tau) \cdot \phi(\theta, \tau), \quad (12)$$

and the optimal kernel was deduced to be, in the ambiguity domain,

$$\phi_{opt}(\theta, \tau) = \frac{|A_E(\theta, \tau)|^2}{E|A(\theta, \tau)|^2}. \quad (13)$$

The time-frequency kernel is computed from the ambiguity domain kernel by a Fourier transformation

$$\Phi_{opt} = \mathcal{F}_1^{-1} \mathcal{F}_2 \{\phi_{opt}\}. \quad (14)$$

3. LOCALLY STATIONARY PROCESSES

A locally stationary process (LSP) [5] has, per definition, a covariance function determined by two functions q , r and has the form

$$r_x(t, s) = q\left(\frac{t+s}{2}\right) \cdot r(t-s). \quad (15)$$

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) [4, 5]. The normalization $r(0) = 1$ is used without loss of generality (any other constant can be incorporated into m). Such processes do exist [5]. For instance, if r is the covariance of any stationary process and $q(\frac{t+s}{2})$ is positive definite (a so called *exponentially convex covariance* [5, 6]). Then q is a real analytic function which is the Laplace-Stieltjes transform of a non-negative non-decreasing function F ,

$$q(t) = \int e^{xt} dF(x). \quad (16)$$

It is exponentially growing either at $+\infty$ or at $-\infty$, which means that the LSP is not harmonizable. But there also exists LSPs where the function q is bounded, and where $q(\frac{t+s}{2})$ is not a covariance (it is not non-negative definite).

Lemma 1. *If $r(t-s)$ and $q(\frac{t+s}{2})$ are non-negative definite functions, and if*

$$q(t)e^{-2at^2} = \int e^{it\theta} dQ(\theta) \quad (17)$$

where dQ is a bounded measure, then $q_1(t) = q(t)e^{-2at^2}$ and $r_1(t) = r(t)e^{-(a/2)t^2}$ define a covariance $r_x(t, s) = q_1(\frac{t+s}{2}) \cdot r_1(t-s)$ of a harmonizable LSP.

Example 2. Let $q(t) \equiv 1$, $r(t) = e^{-(b/2)t^2} \Rightarrow q_1(t) = e^{-2at^2}$, $r_1(t) = e^{-\frac{a+b}{2}t^2}$. Then r_1 and q_1 define an LSP for any $b \geq 0$.

LSPs can be seen approximately as stationary processes that have a time varying power described by q . The Cauchy-Schwartz inequality

$$|r_x(t, s)|^2 = |E\{X(t)\overline{X(s)}\}|^2 \leq E|X(t)|^2 E|X(s)|^2 \quad (18)$$

is necessary for a function r_x to be the covariance of any stochastic process. For an LSP it is the inequality relation between q and r

$$(q(t))^2 |r(\tau)|^2 \leq q(t + \tau/2)q(t - \tau/2), \quad \forall t, \tau. \quad (19)$$

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

$$A_E(\theta, \tau) = Q(\theta)r(\tau). \quad (20)$$

Restricting to circularly symmetric Gaussian distributed processes, the denominator of (13) reduces to

$$E|A(\theta, \tau)|^2 = |A_E(\theta, \tau)|^2 + (\mathcal{F}_1^{-1} \mathcal{F}_2\{|A_E|^2\})(\tau, \theta). \quad (21)$$

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

$$= |Q(\theta)|^2 \cdot |r(\tau)|^2 + (\mathcal{F}|r|^2)(\theta) \cdot (\mathcal{F}^{-1}|Q|^2)(\tau). \quad (22)$$

The optimal kernel (13) is thus $\phi_{opt}(\theta, \tau) =$

$$= \frac{|Q(\theta)|^2 |r(\tau)|^2}{|Q(\theta)|^2 |r(\tau)|^2 + (\mathcal{F}|r|^2)(\theta) (\mathcal{F}^{-1}|Q|^2)(\tau)} = \quad (23)$$

$$= \frac{|Q(\theta)|^2 |r(\tau)|^2}{|Q(\theta)|^2 |r(\tau)|^2 + (\mathcal{F}|r|^2)(\theta) (q * \tilde{q})(\tau)}, \quad (24)$$

where $\tilde{q}(\tau) := q(-\tau)$. It can be shown easily that the functions $\mathcal{F}|r|^2$ and $\mathcal{F}^{-1}|Q|^2$ are non-negative, which gives $0 \leq \phi_{opt}(\theta, \tau) \leq 1$. A special case of LSPs is weakly stationary processes, $q(\tau) \equiv c > 0$. But $q * \tilde{q}$ is not well defined then, so this case can not be treated with formula (24). From (14) and (24) we can directly make the following observations concerning the kernels $\phi_{opt}(\theta, \tau)$ and $\Phi_{opt}(t, \omega)$.

- The τ variable support $\text{supp}_\tau(\phi_{opt}) = \text{supp}(r)$. If $\text{supp}(r) = K = \text{compact}$ then $\Phi(t, \omega)$ is analytic in the ω variable.
- The θ variable support $\text{supp}_\theta(\phi_{opt}) = \text{supp}(Q)$. If $\text{supp}(Q) = K = \text{compact}$ then $\Phi(t, \omega)$ is analytic in the t variable.
- If $\text{supp}(q * \tilde{q}) \subset \text{supp}(r)$, then $\phi_{opt}(\theta, \tau) = 1$, $\tau \in \text{supp}(r) \setminus \text{supp}(q * \tilde{q})$. In particular, if $\text{supp}(q * \tilde{q}) \subset \text{supp}(r) = \mathbf{R}$ then $\Phi(t, \omega)$ will contain Dirac measure in the ω variable.
- If $\text{supp}(\mathcal{F}|r|^2) \subset \text{supp}(Q)$, then $\phi_{opt}(\theta, \tau) = 1$, $\theta \in \text{supp}(Q) \setminus \text{supp}(\mathcal{F}|r|^2)$. In particular, if $\text{supp}(\mathcal{F}|r|^2) \subset \text{supp}(Q) = \mathbf{R}$ then $\Phi(t, \omega)$ will contain Dirac measure in the t variable.

4. LSPS WITH CONTINUOUS OPTIMAL TF KERNELS

Here we give sufficient conditions for the time-frequency kernel $\Phi_{opt}(t, \omega)$ for a LSP to be a continuous function. Sayeed and Jones gave a characterization of when $\phi_{opt}(\theta, \tau) \equiv c$ is a constant, i.e. when $\Phi_{opt}(t, \omega)$ is a Dirac measure at the origin. This means that no averaging is performed in neither the time nor frequency domain by the estimator. The following (very restrictive) condition is necessary and sufficient for $\phi_{opt}(\theta, \tau) \equiv c$ to hold. $X(t) = X_0 \cdot u(t)$

where X_0 is an arbitrary random variable with $|X_0| = \text{const}$ a.s., and $u(t)$ is a deterministic function [2].

The pair of functions q, r should define a non-negative definite function $r_x(t, s)$ in order to be the covariance. The two following conditions are sufficient for $\Phi_{opt} \in C_0 =$ continuous functions vanishing at infinity.

Theorem 3. If

$$\frac{|Q|^2}{\mathcal{F}|r|^2} \in L^1(\mathbf{R}), \quad (25)$$

then $\Phi_{opt}(t, \omega) \in C_0(t)$, i.e. it is a continuous function, vanishing at infinity, in the t variable. If

$$\frac{|r|^2}{q * \tilde{q}} \in L^1(\mathbf{R}). \quad (26)$$

then $\Phi_{opt}(t, \omega) \in C_0(\omega)$. i.e. it is a continuous function, vanishing at infinity, in the ω variable.

Example 4. Let $q(t) \equiv 1 \Rightarrow q_1(t) = e^{-2at^2}$, $r_1(t) = e^{-(a/2)t^2} r(t)$. Then $q_1 * \tilde{q}_1(t) = 2\sqrt{\pi/a} e^{-at^2}$, and $\frac{|r_1|^2}{q_1 * \tilde{q}_1} = \frac{1}{2}\sqrt{a/\pi} |r(t)|^2$. According to (26), $\Phi_{opt}(t, \omega) \in C_0(\omega)$ if $r \in L^2(\mathbf{R})$.

Example 5. Let $q(t) \equiv 1 \Rightarrow q_1(t) = e^{-2at^2}$, $r(t) = \cos(\theta_0 t) \Rightarrow r_1(t) = e^{-(a/2)t^2} \cos(\theta_0 t)$. Then $r \notin L^2(\mathbf{R})$, so $\Phi_{opt}(t, \omega) \in C_0(\omega)$ is not guaranteed by (26). The optimal kernel (24) is

$$\phi_{opt}(\theta, \tau) = \frac{\cos^2(\theta_0 \tau)}{\cos^2(\theta_0 \tau) + 4e^{-\theta_0^2/(4a)} \cosh(\frac{\theta \theta_0}{2a})}. \quad (27)$$

For fixed θ , ϕ_{opt} is oscillating as a function of τ . Therefore $\Phi_{opt}(t, \omega)$ will contain Dirac measure as a function of ω . From

$$\begin{aligned} \frac{|Q_1|^2}{\mathcal{F}|r_1|^2} &= \frac{\sqrt{\pi/a}}{1 + \frac{1}{4}e^{-\theta_0^2/a} \cosh(\theta \theta_0/a)} \\ &\sim \frac{1}{\cosh(\frac{\theta \theta_0}{2a})} \in L^1(\mathbf{R}), \end{aligned} \quad (28)$$

we see by theorem 3 that $\Phi_{opt} \in C_0(t)$.

5. LSPS WITH GAUSSIAN CONSTITUENT FUNCTIONS

Here we study the case when both q and r are Gaussian functions. Assume q is the fix Gaussian $q(\tau) = e^{-\tau^2/2}$ and $r(\tau) = e^{-\frac{c}{4}\tau^2/2}$. $r_x(t, s)$ is a covariance if and only if $c \geq 1$. The two conditions (25), (26) of Theorem 1 are fulfilled if and only if $c > 1$. The limit case $c = 1$ results in

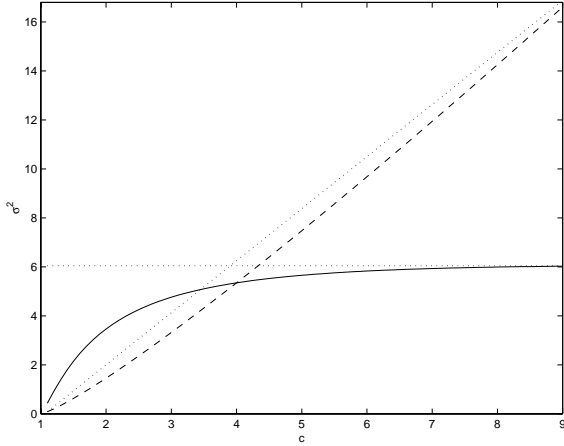


Figure 1. Variances σ_t^2 (solid), σ_ω^2 (dashed), and asymptotes (dotted).

$\phi_{opt}(\theta, \tau) \equiv \frac{1}{2} \Leftrightarrow \Phi_{opt}(t, \omega) = \frac{1}{2}\delta(t, \omega)$. The kernel (24) is for arbitrary c

$$\phi_{opt}(\theta, \tau) = \frac{1}{1 + c^{-1/2}e^{(1-\frac{1}{c})\theta^2 + \frac{c-1}{4}\tau^2}}. \quad (29)$$

We would like to study the influence of the parameter c , i.e. the quotient of the variances between q and r , on the TF kernel Φ_{opt} . The function ϕ_{opt} is almost rank-one for small $c \geq 1$, and of slowly increasing rank for larger c (ratio of largest eigenvalue to the sum of eigenvalues is ≥ 0.9 for $1 \leq c \leq 100$). Of this reason, it is reasonable to approximate ϕ_{opt} by a rank-one function. We compute the variances of these functions, i.e. the variances of f and g which fulfill

$$\arg \min_{f,g} \|\phi_{opt}(\theta, \tau) - f(\theta)g(\tau)\|_2. \quad (30)$$

The variances are denoted by σ_θ^2 and σ_τ^2 , respectively. After Fourier transformation $\phi_{opt} \rightarrow \Phi_{opt}$, $\theta \rightarrow t$, $\tau \rightarrow \omega$, the variance of Φ_{opt} will approximately (exactly if f, g were Gaussians) be $\sigma_t^2 = 4\sigma_\theta^{-2}$ in the t direction and $\sigma_\omega^2 = 4\sigma_\tau^{-2}$ in the ω direction. Figure 1 shows σ_t^2 and σ_ω^2 as a function of c in the interval $1 \leq c \leq 9$.

From the figure we conclude that σ_ω^2 has a close to linear increase. The increase is natural, since increasing c corresponds to r approaching the covariance of a white process, implying that optimal estimation involves increased smoothing in the ω direction. It can also be seen that σ_t^2 grows slowly for large c and seems to have an upper bound. This can be expected since the function q , which can be seen as representing the process power as a function of time, has a fixed variance = 2.

6. CONCLUSIONS

The covariance function of a locally stationary processes is determined by two functions q and r . Restricted to circularly symmetric Gaussian distributed processes, we have obtained (i) a formula for LSPs for the optimal spectral estimation kernel in the ambiguity domain, expressed in q and r , (ii) sufficient conditions on q and r to give a continuous TF kernel, vanishing at infinity, and (iii) various examples of combinations of q and r leading to distinct types of optimal kernels. Also, we have studied the case of Gaussian functions q and r numerically. It turned out that the resulting kernel is close to rank-one.

7. REFERENCES

- [1] W. Kozek, *Matched Weyl-Heisenberg expansions of nonstationary environments*, Ph.D. thesis, Tech. Univ. Vienna, 1996.
- [2] A. M. Sayeed and D. L. Jones, "Optimal kernels for nonstationary spectral estimation," *IEEE Trans Sig Proc*, vol. 43, pp. 478–491, 1995.
- [3] M Loève, *Probability theory*, Van Nostrand, 1963.
- [4] P. Flandrin, *Time-frequency/Time-scale analysis*, Academic Press, 1999.
- [5] R. A. Silverman, "Locally stationary random processes," *IRE Trans. Info. Theory*, vol. 3, 1957.
- [6] M. Loève, "Fonctions aléatoires a décomposition orthogonale exponentielle," *Rev. Sci.*, vol. 84, pp. 159–162, 1946.