EUSIPCO 2013 1569743611

OPTIMAL TIME FREQUENCY ANALYSIS OF MULTIPLE TIME-TRANSLATED LOCALLY STATIONARY PROCESSES

Johan Brynolfsson and Maria Hansson-Sandsten

Lund University,
Mathematical Statistics, Centre for Mathematical Sciences,
Box 118, SE-221 00 Lund, Sweden.
johanb@maths.lth.se, sandsten@maths.lth.se

ABSTRACT

A previously proposed model for non-stationary signals is extended in this contribution. The model consists of multiple time-translated locally stationary processes. The optimal Ambiguity kernel for the process in mean-square-error sense is computed analytically and is used to estimate the time-frequency distribution. The performance of the kernel is compared with other commonly used kernels. Finally the model is applied to electrical signals from the brain (EEG) measured during a concentration task.

***Index Terms*—** Time frequency analysis, Locally stationary process, Optimal Ambiguity kernel, EEG.

1. INTRODUCTION

The paper treats estimation of the Wigner spectrum of Gaussian stochastic processes using the quadratic class of random time-frequency representations. We study the minimum mean square error estimation kernel and restrict to processes that have locally stationary covariances in Silverman's sense, [1, 2, 3]. There are other definitions of locally stationary processes to be found in literature, e.g. [4], but the locally stationary process (LSP) as defined by Silverman has a covariance function which is a multiplication of a covariance function of a stationary process and a time-variable function giving it useful separability features.

The mean square error (MSE) optimal kernel for non-stationary processes was first derived by Sayeed and Jones [5] and in [6] the optimal kernel for a LSP in Silverman's sense in the case of Gaussian covariance and time-variable functions has been derived.

An extension of the LSP model is presented in this paper, making it more useful when studying real life signals. The set of LSPs is parameterized and can be used to model signals ranging from non-stationary to stationary and can thereby be used to represent a wide variety of natural signals, such as speech or electrical brain activity, [7].

Thanks to the Swedish Research Council for funding.

Whereas a LSP consists of one single component we allow the extended model to contain multiple components centered at different times.

The paper is organized as follows. In section 2 we state the definition of a multiple time translated locally stationary process. In section 3 we show the definitions of the Wigner spectrum and the Ambiguity spectrum and how they are related. We derive the optimal Ambiguity kernel in section 4 and compare the performance with other kernel functions in section 5. Finally we apply the model to EEG signals in section 6 and state our conclusions in section 7.

2. MULTIPLE TIME TRANSLATED LOCALLY STATIONARY PROCESSES

Definition 1 *A zero-mean second-order continuous-time random process $X(t)$ is called a* ***Locally Stationary Process*** *(in the wide sense) if its covariance function can be written on the form*

$$r_x(t,\tau) = q\left(t\right) \cdot r\left(\tau\right), \tag{1}$$

where $q\left(t\right)$ may be any positive valued function such that $\int |q(t)|^2 dt < \infty$, (or a positive constant), and $r\left(\tau\right)$ must fulfill the properties of a stationary covariance function [1].

Since this is a unimodal model we want to extend it to contain multiple components which we allow to be centered at different times.

Definition 2 *The* ***Multiple Time Translated Locally Stationary Process*** *(MTTLSP) is defined as a stochastic process with the covariance function*

$$r_x(t,\tau) = \sum_{j=1}^{N} a_j \cdot q_j\left(t\right) \cdot r_j\left(\tau\right), \tag{2}$$

where a_j is a positive valued scaling factor and N is the number of components in the process.

We will also restrict the analysis to the class of MTTLSP where the q and r functions are Gaussian on the form

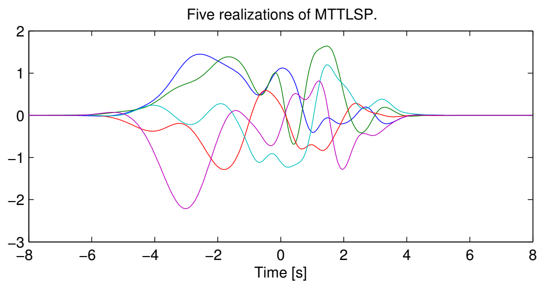

Fig. 1. Realizations of a slow oscillating part centered around $t = -2$ and a faster oscillating part centered around $t = 1$.

$$\begin{cases} q_j(t) = e^{-\frac{1}{2}\left(\frac{t}{f_j}-t_j\right)^2}, \\ r_j(\tau) = e^{-\frac{c_j}{8}\left(\frac{\tau}{f_j}\right)^2}, \end{cases} \quad (3)$$

and where $c_j > 1$ [6]. When $c_j >> 4$, $r_j(\tau)$ decreases quicker than $q_j(\tau)$ and we approach a stationary process as $c_j \to \infty$. The opposite extreme with $c_j \to 1$ give maximal non-stationarity within the model. The parameter f_j is a frequency scaling factor. Realizations of MTTLSP can be seen in figure 1 where the process is sampled at a frequency of 50 Hz and consists of two components. The first with non-stationarity parameter $c_1 = 5$, frequency scaling $f_1 = 1$ and a time-translation of $t_1 = -2$ seconds. The second component has the parameters $c_2 = 15$, $f_2 = 0.8$ and $t_2 = 1$. Both components have unit amplitude $a_1 = a_2 = 1$.

3. WIGNER SPECTRUM

A time-frequency estimate in the quadratic class is written

$$W_X^Q(t,f) = \int\int A_X(\nu,\tau)\phi(\nu,\tau)e^{-i2\pi(\tau f - t\nu)}d\tau d\nu, \quad (4)$$

using the *Ambiguity kernel*, $\phi(\nu,\tau)$, and where the corresponding *Ambiguity spectrum* of a process $X(t)$ with covariance function r_x is defined by

$$\begin{aligned} E\{A_X(\nu,\tau)\} &= E\left\{\int X(t+\tau/2)\overline{X(t-\tau/2)}e^{-i2\pi t\nu}dt\right\} \\ &= \int r_x(t,\tau)e^{-i2\pi t\nu}dt. \end{aligned} \quad (5)$$

Throughout the paper $E\{\cdot\}$ will denote expected value, i is the imaginary unit and $\overline{X(t)}$ will denote the complex conjugate of $X(t)$. The Ambiguity spectrum is related to the *Wigner spectrum*, [8, 5, 9],

$$\begin{aligned} E\{W_X(t,f)\} &= E\left\{\int X(t+\tau/2)\overline{X(t-\tau/2)}e^{-i2\pi f\tau}d\tau\right\} \\ &= \int r_x(t,\tau)e^{-i2\pi f\tau}d\tau. \end{aligned} \quad (6)$$

All integrals are assumed to reach from $-\infty$ to ∞.

4. OPTIMAL AMBIGUITY KERNEL OF A MTTLSP

In this section we will derive the global optimal Ambiguity kernel of MTTLSP, meaning that we don't allow the kernel to depend on time or frequency, only time- and frequency lag. Note that this means that the kernel is invariant to translations of the whole system in both time and frequency. We will also assume the process is circularly symmetric, meaning that $E\{x(t)x(s)\} = 0$. The optimal Ambiguity kernel for a non-stationary process, in the MSE sense $X(t)$, was derived in [5]. Assuming the process is observed noise free, the mean squared error optimal kernel can be formulated as

$$\begin{aligned} \phi_{opt} &= \arg \inf_{\phi(\nu,\tau)} \iint E\{|A_X(\nu,\tau)\phi(\nu,\tau) \\ &\quad - E\{A_X(\nu,\tau)\}|^2\} d\tau d\nu, \end{aligned} \quad (7)$$

where the solution, if it exists, can be stated as

$$\phi_{opt}(\nu,\tau) = \frac{|E\{A_X(\nu,\tau)\}|^2}{E\left\{|A_X(\nu,\tau)|^2\right\}}.$$

Assume that $X(t)$ is a MTTLSP giving,

$$\begin{aligned} |E\{A_X(\nu,\tau)\}|^2 &= \left|\int r_x(t,\tau)e^{-i2\pi\nu t}dt\right|^2 \\ &= \left|\int\left(\sum_{j=1}^N a_j q_j(t) r_j(\tau)\right)e^{-i2\pi\nu t}dt\right|^2 \\ &= \left|\sum_{j=1}^N a_j Q_j(\nu) r_j(\tau)\right|^2, \end{aligned} \quad (8)$$

where $Q_j(\nu) := \mathcal{F}_{t\to\nu}(q(t))$ and $\mathcal{F}$ represents the Fourier transform. Now we go to the more laborious denominator.

$$\begin{aligned} E\left\{|A_X(\nu,\tau)|^2\right\} &= E\left\{\left|\int X(t+\frac{\tau}{2})\overline{X(t-\frac{\tau}{2})}e^{-i2\pi\nu t}dt\right|^2\right\}. \\ &= E\left\{\int X(t_1+\frac{\tau}{2})\overline{X(t_1-\frac{\tau}{2})}e^{-i2\pi\nu t_1}dt_1\right. \\ &\quad \left.\cdot\overline{\int X(t_2+\frac{\tau}{2})\overline{X(t_2-\frac{\tau}{2})}e^{-i2\pi\nu t_2}dt_2}\right\} \\ &= E\left\{\iint X(t_1+\frac{\tau}{2})\overline{X(t_1-\frac{\tau}{2})X(t_2+\frac{\tau}{2})}\right. \cdot \\ &\quad \left.\cdot X(t_2-\frac{\tau}{2})e^{-i2\pi\nu(t_1-t_2)}dt_1dt_2\right\}. \end{aligned} \quad (9)$$

As X is a zero mean Gaussian stochastic process, we can make use of Isserlis' theorem [10] which states that

$$\begin{aligned} E\{x_1x_2x_3x_4\} &= E\{x_1x_2\}E\{x_3x_4\} \\ &\quad + E\{x_1x_3\}E\{x_2x_4\} + E\{x_1x_4\}E\{x_2x_3\}. \end{aligned}$$

Using this we can represent

$$E\left\{|A_X(\nu,\tau)|^2\right\} = B(\nu,\tau) + C(\nu,\tau) + D(\nu,\tau), \quad (10)$$

and derive the different terms separately.

Beginning with $B(\nu,\tau)$,

$$B(\nu,\tau) = \iint E\left\{X(t_1+\frac{\tau}{2})\overline{X(t_1-\frac{\tau}{2})}\right\}\cdot$$

$$\cdot E\left\{\overline{X(t_2+\frac{\tau}{2})}X(t_2-\frac{\tau}{2})\right\}e^{-i2\pi\nu(t_1-t_2)}dt_1dt_2$$
$$= \int r_x(t_1,\tau)e^{-i2\pi\nu t_1}dt_1\cdot\overline{\int r_x(t_2,\tau)e^{-i2\pi\nu t_2}dt_2}$$

$$= \left|\sum_{j=1}^{N} a_jQ_j(\nu)r_j(\tau)\right|^2. \tag{11}$$

Continue then with $C(\nu,\tau)$.

$$\begin{aligned} C(\nu,\tau) &= \iint E\left\{X(t_1+\frac{\tau}{2})\overline{X(t_2+\frac{\tau}{2})}\right\}\cdot \\ &\quad\cdot E\left\{\overline{X(t_1-\frac{\tau}{2})}X(t_2-\frac{\tau}{2})\right\}e^{-i2\pi\nu(t_1-t_2)}dt_1dt_2 \\ &= \iint rx\left(\frac{t_1+t_2+\tau}{2},t_1-t_2\right)\cdot \\ &\quad\overline{\cdot rx\left(\frac{t_1+t_2-\tau}{2},t_1-t_2\right)}e^{-i2\pi\nu(t_1-t_2)}dt_1dt_2. \end{aligned} \tag{12}$$

Change of variables $t_a = \frac{t_1+t_2+\tau}{2}$ and $t_b = t_1 - t_2$ gives that the integral now can be expressed as

$$\begin{aligned} C(\nu,\tau) &= \iint r_x(t_a,t_b)\overline{r_x(t_a-\tau,t_b)}e^{-i2\pi\nu t_b}dt_adt_b \\ &= \iint\left(\sum_{j=1}^{N}a_jq_j\left(t_a\right)r_j(t_b)\right)\cdot \\ &\quad\cdot\overline{\left(\sum_{k=1}^{N}a_kq_k\left(t_a-\tau\right)r_k(t_b)\right)}e^{-i2\pi\nu t_b}dt_adt_b \\ &= \iint\left(\sum_{j=1}^{N}\sum_{k=1}^{N}a_jq_j\left(t_a\right)r_j(t_b)\cdot\right. \\ &\quad\left.\left(\overline{a_kq_k\left(t_a-\tau\right)r_k(t_b)}\right)e^{-i2\pi\nu t_b}dt_adt_b\right) \\ &= \sum_{j=1}^{N}\sum_{k=1}^{N}a_j\overline{a_k}\int q_j\left(t_a\right)\overline{q_k\left(t_a-\tau\right)}dt_a\cdot \\ &\quad\cdot\int r_j(t_b)\overline{r_k(t_b)}e^{-i2\pi\nu t_b}dt_b \\ &= \sum_{j=1}^{N}\sum_{k=1}^{N}a_j\overline{a_k}(q_j*\tilde{q_k})(\tau)\cdot\mathcal{F}_{t\to\nu}(r_j\overline{r_k}) \end{aligned} \tag{13}$$

where the notation $\tilde{q_k}(\tau) = q_k(-\tau)$ is used. Assuming the process X is circularly symmetric, part $D(\nu,\tau)$ is equal to zero as it contains no complex conjugation.

Denoting $EA_X = \left|\sum_{j=1}^{N}a_jQ_j(\nu)r_j(\tau)\right|^2$, the optimal Ambiguity kernel can so far be found as

$$\phi_{opt}(\nu,\tau) = \frac{EA_X}{EA_X + \sum_{j,k}a_j\overline{a_k}(q_j*\tilde{q_k})(\tau)\cdot\mathcal{F}_{t\to\nu}(r_j\overline{r_k})}. \tag{14}$$

Find yet unknown parts:

$$\begin{aligned} Q_j(\nu) &= \int_{-\infty}^{\infty}q_j(t)e^{-i2\pi\nu t}dt \\ &= \int_{-\infty}^{\infty}e^{-\frac{1}{2}\left(\frac{t}{f_j}-t_j\right)^2}e^{-i2\pi\nu t}dt \\ &= e^{-i2\pi f_jt_j\nu}\sqrt{2\pi f_j^2}e^{-2(\pi f_j\nu)^2}. \end{aligned} \tag{15}$$

Assuming that f_j and c_j are real-valued $\forall j$, we have that

$$\begin{aligned} (q_j*\tilde{q_k})(\tau) &= \int q_j\left(t\right)\overline{q_k\left(t-\tau\right)}dt, \\ &= \int e^{-\frac{1}{2}\left(\frac{t}{f_j}-t_j\right)^2}e^{-\frac{1}{2}\left(\frac{t-\tau}{f_k}-t_k\right)^2}dt, \\ &= \sqrt{2\pi}\frac{e^{-\frac{1}{2}\frac{(\tau+t_kf_k-t_jf_j)^2}{f_k^2+f_j^2}}}{\sqrt{\frac{f_j^2+f_k^2}{f_j^2f_k^2}}}. \end{aligned} \tag{16}$$

$$\begin{aligned} \mathcal{F}(r_j\overline{r_k})(\nu) &= \int r_j(t)\overline{r_k(t)}e^{-i2\pi\nu t}dt \\ &= \int e^{-\frac{c_j}{8}\left(\frac{t}{f_j}\right)^2}\overline{e^{-\frac{c_k}{8}\left(\frac{t}{f_k}\right)^2}}e^{-i2\pi\nu t}dt \\ &= \int e^{-\frac{t^2}{8}\left(\frac{c_j}{f_j^2}+\frac{c_k}{f_k^2}\right)}e^{-i2\pi\nu t}dt \\ &= \sqrt{\frac{8\pi}{\left(\frac{c_j}{f_j^2}+\frac{c_k}{f_k^2}\right)}}e^{-\frac{8\pi^2\nu^2}{\left(\frac{c_j}{f_j^2}+\frac{c_k}{f_k^2}\right)}}. \end{aligned} \tag{17}$$

Plugging this result into equation 14 we have arrived to the analytic expression of the MSE optimal Ambiguity kernel of a MTTLSP process.

5. COMPARISON WITH OTHER KERNELS

The MSE is calculated for a set of commonly used kernels to compare the performance and show that the MTTLSP kernel in fact performs better than existing methods. To make a fair comparison we optimize all methods over their respecting parameters. The methods we choose to compare with are the Wigner spectrum (i.e. using no kernel), the Choi-Williams kernel $\phi_{CW} = exp(-\alpha\tau^2\nu^2)$ [11], the Hanning window (corresponding to the short-time Fourier transform) and the LSP optimal kernel (for one single component). When using the Hanning window we transform it into Ambiguity kernel functions [12]. The MSE is then calculated for each discrete (ν, τ) and the expected mean of the MSE (MMSE) is then calculated over the time span of the process and the Nyquist frequency. Assume that we have an observed process with N_{obs} samples with a sampling frequency F_s and the discrete Fourier transform is calculated using zero padding up to N_{FFT} samples. The MMSE will hence be calculated as

$$MMSE = \frac{1}{N_{obs}} \cdot \frac{1}{N_{FFT}} \sum_{\tau} \sum_{\nu} E\left\{ |A_X(\nu,\tau)\phi(\nu,\tau) - EA_X(\nu,\tau)|^2 \right\}, \quad (18)$$

using the analytic expressions for the expected ambiguity functions derived in equations (5)-(17). The kernels are evaluated on MTTLSP with two components where the stationarity parameters are $c_1 = 7$ and $c_2 = 3$, the scaling frequencies $f_1 = f_2 = 1$ Hz, the time translations $t_1 = -1.3$ and $t_2 = 1.7$ seconds and both components had maximum instant variance equal to one. The sampling frequency was 30 Hz. Results of the evaluations are presented in table 1.

The Wigner-Ville estimate generated the highest MMSE which is not surprising as there will be significant cross terms caused by the multiple components.

The Choi-Williams kernel was optimized over the parameter α giving a minimum MMSE for $\alpha = 685$, meaning we are approaching the Wigner-Wille spectrum. The reason for the high value of α is due to the shape of the Choi-Williams kernel where, when used on MTTLSP processes, as a higher cross-term reduction renders also renders an auto-term reduction.

The Hanning window was optimized over window length. The best MMSE was found for a 10 sample long window.

As expected the MTTLSP kernel has the lowest MMSE closely followed by the LSP kernel. The reason for the quite small difference in MMSE is largely explained by that we only evaluate on a dual component process. To show that MTTLSP is still an improvement over the original one-component LSP model we calculate the MMSE of a 10 component process with stationarity parameters, $c_j's$, between 2 and 11; scaling frequencies, $f_j's$, between 0.1 and 1.9; time translations, $t_j's$ between -2 and 2 seconds and scaling factors $a_j's$ between 0.1 and 1.9. In this case the MMSE of the optimal LSP Ambiguity kernel gave a MMSE of $81.7 \cdot 10^{-3}$ whereas the MMSE of the optimal MTTLSP kernel was $69.6 \cdot 10^{-3}$; roughly 17% better.

Method	MMSE
MTTLSP	$4.2 \cdot 10^{-3}$
Wigner-Ville	$42.7 \cdot 10^{-3}$
Choi-Williams	$7.1 \cdot 10^{-3}$
Hanning	$8.5 \cdot 10^{-3}$
LSP	$4.3 \cdot 10^{-3}$

Table 1. Performance of kernels on a MTTLSP process

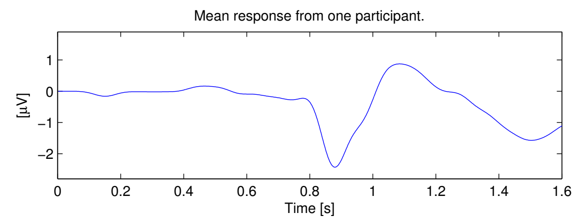

Fig. 2. Mean response from one of the participants in the trial.

6. EEG ANALYSIS USING MTTLSP KERNELS

The set of locally stationary processes have previously been proposed as a model of EEG signals [13]. During an psychological experiment participants were asked to quickly identify letters appearing on a screen. Simultaneously their brain activity was recorded using EEG electrodes. A mean response from one of the participants can be seen in figure 2. The brain responses are assumed to be realizations of the same MTTLSP every repetition of a certain trigger. If we have M recordings and denote the $m : th$ brain response $x_m(t)$ we can estimate the instantaneous covariance as

$$\hat{r}_x(t,\tau) = \frac{1}{M} \sum_{m=1}^{M} x_m(t + \tau/2)\bar{x}_m(t - \tau/2).$$

The estimated covariance matrix can be seen in figure 3. In the top right corner of the same figure one can see the diagonal of the matrix where time-lag is identically zero. Assuming the EEG signal is a MTTLSP, one then observes the sum of time-translated Gaussian functions ($q_j(t)$ in eq. (3)) as $\tau = 0$ gives that $r_j(\tau) = 1$. The lower three plots in the figure show the anti-diagonal across three main peaks ($r_j(\tau)$ in eq. (3)). The similar principle holds for the anti-diagonals, where $q_j(t) = 1$ for each respective component and one therefore only observe $r_j(\tau)$. We then fitted a model consisting of three components which can be seen as red dashed lines in the plot. The estimated scaling frequencies for the three components were $f_1 = 15.7$, $f_2 = 37.2$ and $f_3 = 49.4$ and the stationarity parameters were estimated to $c_1 = 1$, $c_2 = 1.16$ and $c_3 = 1.18$.

Note that, in calculating the optimal kernel, we assumed the signal to be circularly symmetric. As the real valued EEG signal does not fulfills this, we use the Hilbert transform to get the corresponding complex-valued analytic signal which may be considered approximately circularly symmetric.

$$x_a(t) = x(t) + i\,[\mathcal{H}x]\,(t), \quad (19)$$

where $\mathcal{H}$ is the Hilbert transform. We then estimated the

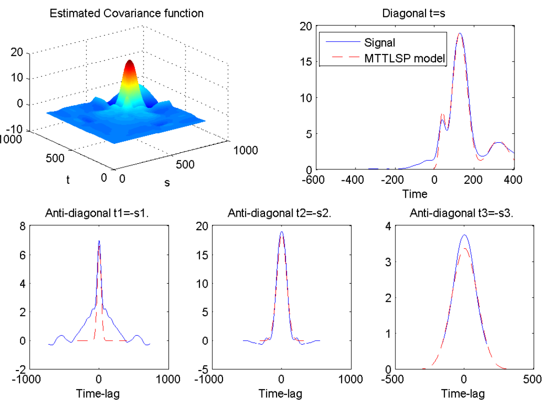

Fig. 3. The estimated covariance function of the EEG signal can be seen in the upper left corner. In the upper right corner is the diagonal of the matrix in blue and the MTTLSP model in dashed red. The three lower plots show the anti-diagonals across each of the three components in blue and the model in dashed red.

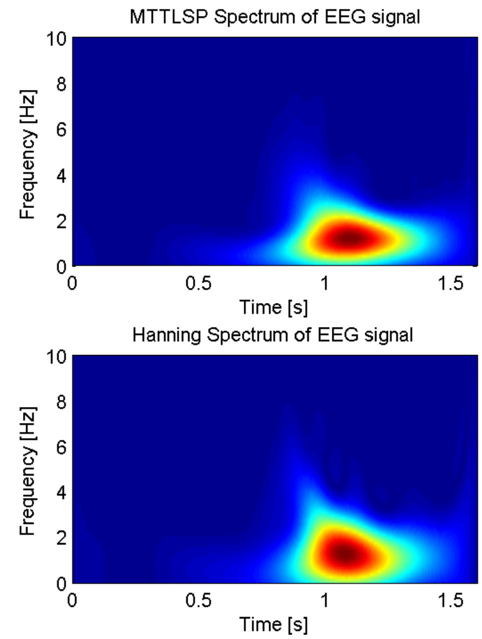

Fig. 4. Spectral estimation of EEG signal using the MTTLSP kernel in the upper figure and the estimation using a Hanning kernel in the lower figure.

Wigner spectrum using the MTTLSP kernel with the parameters fitted to the data. The resulting spectral estimate can be seen in figure 4 together with an spectrum estimated using a Hanning kernel. Note that center of mass is located at a higher frequency in the Hanning spectrum due to too much smoothing.

7. CONCLUSIONS

An extended model of locally stationary processes were presented and the MSE optimal Ambiguity kernel was derived. It was also shown that when estimating the frequency content of the process the said MTTLSP Ambiguity kernel had significantly better performance. By giving an applied example we want to show the usefulness of this new model.

8. REFERENCES

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

[2] R. A. Silverman, "A matching theorem for locally stationary random processes," *Comm. Pure and Appl. Math.*, vol. 12, pp. 373–383, 1959.

[3] P. Wahlberg and P. J. Schreier, "Gabor discretization of the Weyl product for modulation spaces and filtering of nonstationary stochastic processes," *Appl. Comput. Harmon. Anal.*, vol. 26, pp. 97–120, 2009.

[4] R. Dahlhaus, "Fitting time series models to nonstationary processes," *The Annals of Stat.*, vol. 25, no. 1 pp. 1–37, 1997.

[5] A. M. Sayeed and D. L. Jones, "Optimal kernels for nonstationary spectral estimation," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 478–491, Feb. 1995.

[6] P. Wahlberg and M. Hansson, "Kernels and multiple windows for estimation of the Wigner-Ville spectrum of Gaussian locally stationary processes," *IEEE Trans. on Signal Processing*, vol. 55, no. 1, pp. 73–87, January 2007.

[7] M. Hansson-Sandsten, "Optimal estimation of the time-varying spectrum of a class of locally stationary processes using Hermite functions," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, 2011, Article ID 980805.

[8] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.

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

[10] G. Lindgren, *Stationary Stochastic Processes, Theory and Applications*, CRC Press, 2013.

[11] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multi-component signals using exponential kernels," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, pp. 862–871, June 1989.

[12] L. Cohen, *Time-Frequency Analysis*, Prentice-Hall, 1995.

[13] M. Hansson-Sandsten and J. Sandberg, "Optimization of weighting factors for multiple window Spectrogram of event related potentials," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, 2011,