Research Article

Optimization of Weighting Factors for Multiple Window Spectrogram of Event-Related Potentials

Maria Hansson-Sandsten (EURASIP Member) and Johan Sandberg (EURASIP Member)

Division of Mathematical Statistics, Centre for Mathematical Sciences, Lund University, Box 118, 221 00 Lund, Sweden

Correspondence should be addressed to Maria Hansson-Sandsten, sandsten@maths.lth.se

This paper concerns the mean square error optimal weighting factors for multiple window spectrogram of different stationary and nonstationary processes. It is well known that the choice of multiple windows is important, but here we show that the weighting of the different multiple window spectrograms in the final average is as important to consider and that the equally averaged spectrogram is not mean square error optimal for non-stationary processes. The cost function for optimization is the normalized mean square error where the normalization factor is the multiple window spectrogram. This means that the unknown weighting factors will be present in the numerator as well as in the denominator. A quasi-Newton algorithm is used for the optimization. The optimization is compared for a number of well-known sets of multiple windows and common weighting factors and the results show that the number and the shape of the windows are important for a small mean square error. Multiple window spectrograms using these optimal weighting factors, from ElectroEncephaloGram data including steady-state visual evoked potentials, are shown as examples.

1. Introduction

Estimation and detection of frequency changes of shorter or longer duration in the ElectroEncephaloGram (EEG) connected to stimuli, for example, evoked or induced potentials are often of great interest. To statistically differ between responses from different types of stimuli, choosing an spectral estimator with small bias and low variance is important.

The idea of multiple windows or multitapering was introduced by Thomson, [1], and in the last decades the Thomson method has been used in many different application areas. It has been shown to outperform the Welch method [2] in terms of leakage, resolution, and variance for a stationary spectrally smooth process, [3]. For nonsmooth spectra, however, the performance of the Thomson method degrades due to cross-correlation between subspectra [4]. Other appropriate choices are then, for example, [5–7]. A comparison of Hermite and Slepian functions (the Thomson method) has shown that in the case of time-varying signals and spectrogram estimation, Hermite functions are a better choice [8].

The choice of windows has been studied in the literature but how to weight the different multiple window spectrograms in the final average has not gained that much attention. In [9], the weighting factors are optimized for the Peak-Matched Multiple Windows, [6]. A criterion is used where normalized bias, variance, and mean square error is optimized for the predefined peaked spectrum. In the nonstationary case, different approaches to approximate a time-varying spectrum with a few windowed spectrograms have been taken, for example, [10–13].

In this paper we compare the Hermite functions, the Thomson windows, the Peak Matched Multiple Windows, and the Welch windows and evaluate the performance with optimal weighting factors for different processes. The cost function for optimization is the normalized mean square error where the normalization factor is the multiple window spectrogram. This means that the unknown weighting factors will be present in the numerator as well as in the
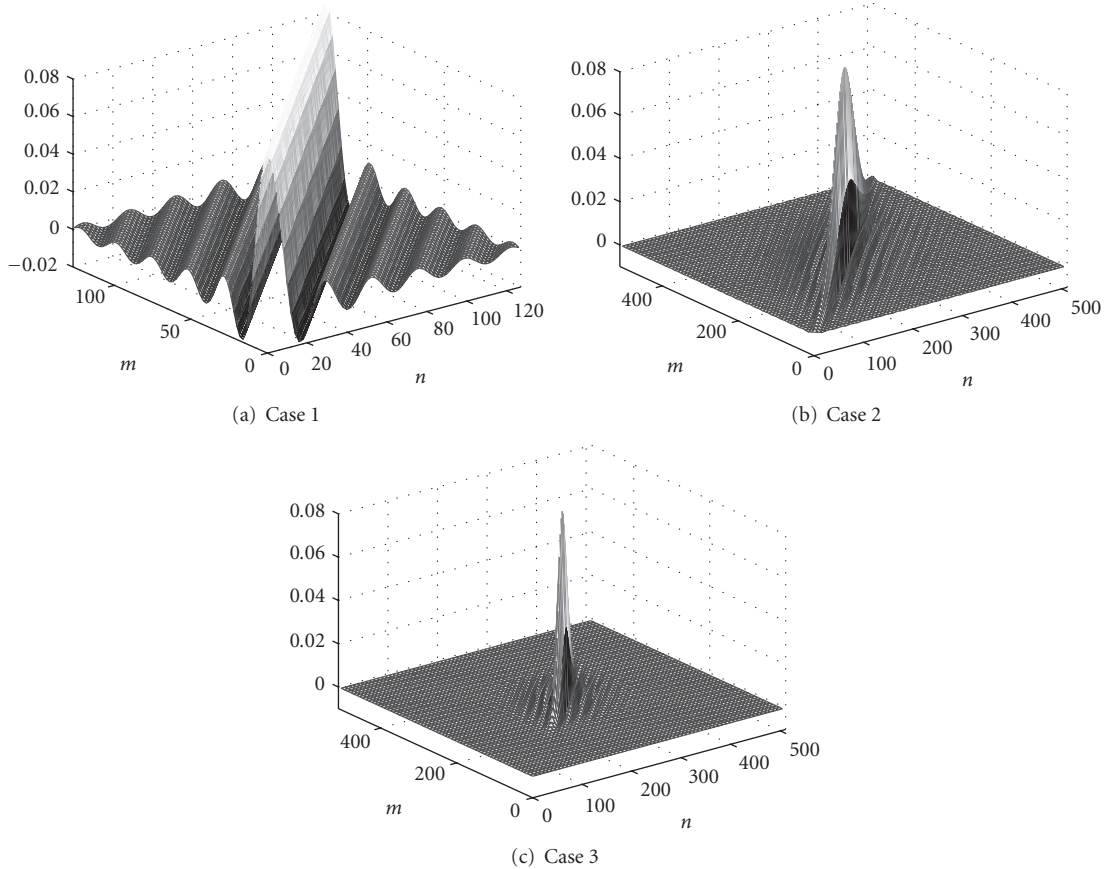
(a) Case 1

(b) Case 2

(c) Case 3

FIGURE 1: The three different test covariance matrices for bandlimited white noise processes: (a) stationary process, (b) long-event nonstationary process, and (c) short-event nonstationary process.

denominator. A quasi-Newton algorithm is used for the optimization. We compare the results from the usual equally weighted multiple window spectrogram as well as an optimal scaling factor-adjusted multiple window spectrogram. Preliminary results have been presented in [14]. A nonstationary process model, which could be appropriate for, for example, induced responses in the EEG, is studied. We illustrate the weighted multiple window spectrogram estimates by showing examples of steady-state visual evoked potential (SSVEP).

The paper is organized as follows. Section 2 presents the optimized weighting factors and in Section 3 the evaluation for different stationary and nonstationary processes is presented. In Section 4 examples of estimation of SSVEP are shown. Section 5 concludes the paper.

2. Optimization of Weighting Factors

The Multiple Window Spectrogram $\hat{S}_x(n,k)$ of the zero mean real-valued random process $x(n)$, $n = 0, \ldots, N_0 - 1$ is defined by

$$\hat{S}_x(n,k) = \sum_{j=1}^{I} \alpha_j \left| \sum_{m=0}^{N-1} x(m+nL)h_j(m)e^{-i2\pi(k/K)m} \right|^2, \quad (1)$$

for $k = 0 \cdots K - 1$ and $0 \le n \le (N_0 - N)/L$, where the assumption is that the data is stationary for the N samples $x(nL) \cdots x(N - 1 + nL)$. Equation (1) is a weighted sum of spectrograms obtained by using the data windows $\mathbf{h}_j = [h_j(0) \cdots h_j(N-1)]^T$, and the weighting factors α_j, $j = 1 \cdots I$. The parameter L is the step size and K the number of values in the DFT.

With only one window, $I = 1$, the spectrogram has too large variance to be useful in the analysis of a stochastic process, as the variance is approximately the squared Wigner spectrum $S_x(n,k)^2$.

2.1. Mean Square Error Optimization. The mean square error (MSE) is a natural choice of optimization since it includes both variance and squared bias. Optimizing the MSE for a model where the power varies with time, which might be the case for nonstationary processes, focuses too much on high-power parts of the process. To avoid this, the optimization can be done normalizing with the true Wigner spectrum at each time and frequency value. However, this might give a strange result if the Wigner estimate is biased. Therefore, we consider the normalized MSE (nMSE) where the expectation of the multiple window spectrogram is used for the normalization at each time and frequency value.
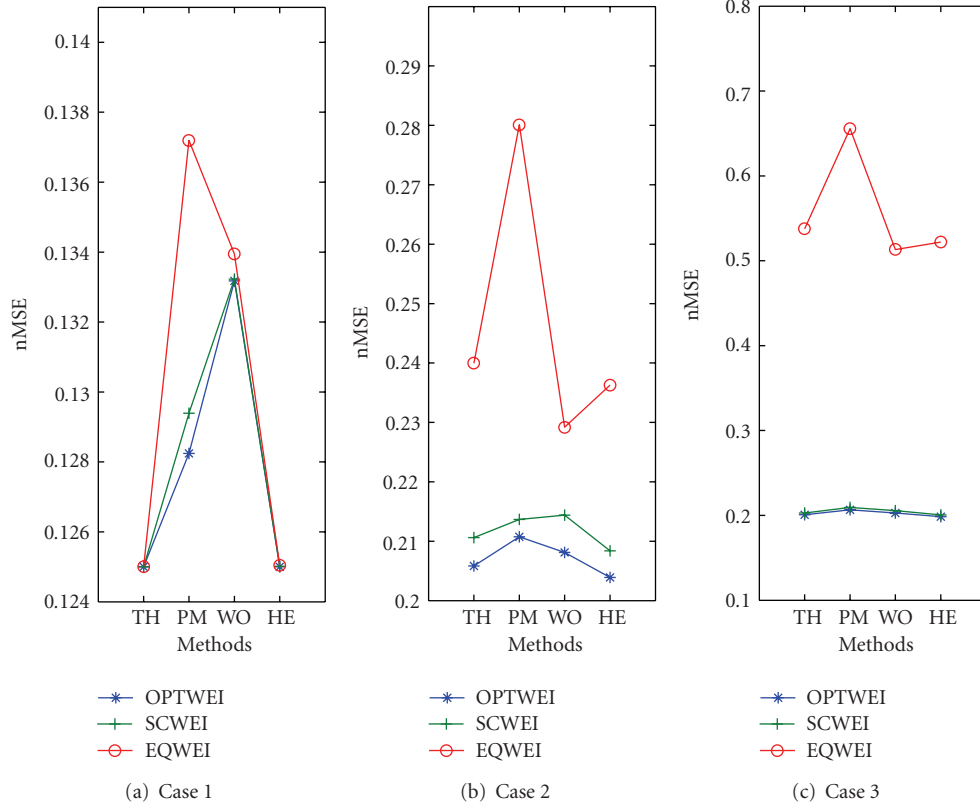
FIGURE 2: The normalized mean square error for the three bandlimited white noise processes with EQWEI (circles), SCWEI (pluses), and OPTWEI (stars) for different window sets, Thomson multiple windows (TH), Peak Matched multiple windows (PM), Welch method (WO), and Hermite functions (HE): (a) stationary process, (b) long-event nonstationary process, and (c) short-event nonstationary process.

The nMSE, which is computed in the time interval $[-T \cdot L \cdots T \cdot L]$ and in the frequency interval $[-M/K \cdots M/K]$, is the average of a number of $(2M+1 \times 2T+1)$ time-frequency values, giving the cost function:

$$\xi = \frac{1}{(2M+1)(2T+1)} \sum_{n=-T}^{T} \sum_{k=-M}^{M} \frac{\varepsilon(n,k)}{E^2\left[\hat{S}_x(n,k)\right]}, \quad (2)$$

where the MSE for each time and frequency value is defined as

$$\varepsilon(n,k) = \text{Variance}\left[\hat{S}_x(n,k)\right] + \text{Bias}^2\left[\hat{S}_x(n,k)\right]. \quad (3)$$

The variance is

$$\text{Variance}\left[\hat{S}_x(n,k)\right] = \sum_{j=1}^{I} \sum_{g=1}^{I} \alpha_j \alpha_g \, \text{cov}\left[\hat{S}_j(n,k), \hat{S}_g(n,k)\right]$$

$$\approx \sum_{j=1}^{I} \sum_{g=1}^{I} \alpha_j \alpha_g \left|\mathbf{h}_j^T \mathbf{\Phi}^H(k) \mathbf{R}_X^n \mathbf{\Phi}(k) \mathbf{h}_g\right|^2, \quad (4)$$

where the covariance matrix $R_X^n = E[\mathbf{x}\mathbf{x}^T]$ with $\mathbf{x} = [x(nL) \cdots x(nL + N - 1)]^T$ and $\mathbf{\Phi}(k) = \text{diag}[1 \, e^{-i2\pi k/K} \, \cdots \, e^{-i2\pi(N-1)k/K}]$ and the superscript H denotes conjugate transpose, according to [4]. Reduction

of the variance is established if the correlation between the windowed periodogram (subspectra),

$$S_j(n,k) = \left|\sum_{m=0}^{N-1} x(m + n\,L) h_j(m) e^{-i2\pi(k/K)m}\right|^2, \quad (5)$$

from the windows $\mathbf{h}_j$ and h_g, $j \neq g$, is small for all frequency values k.

The bias is

$$\text{Bias}\left[\hat{S}_x(n,k)\right] = E\left[\hat{S}_x(n,k)\right] - S_x(n,k)$$

$$= \sum_{j=1}^{I} \alpha_j \mathbf{h}_j^T \mathbf{\Phi}^H(k) \mathbf{R}_X^n \mathbf{\Phi}(k) \mathbf{h}_j - S_x(n,k), \quad (6)$$

where $S_x(n,k)$ is the known Wigner spectrum of the model. The optimization cost function of (2) includes the expressions of (4) and (6) where $\mathbf{h}_j$, $j = 1, \ldots, I$ are known windows and $\mathbf{R}_X^n$ is the time-variable nonstationary covariance matrix. The unknown variables are α_j, $j = 1, \ldots, I$ which appear both in the numerator and the denominator of (2). The minimization of the criterion is therefore done iteratively with a quasi-Newton algorithm [15]. The criterion and its derivative are used in the algorithm. The algorithm is described in [9]. Using these weights in the multiple window spectrogram is referred to as optimal weights (OPTWEI).
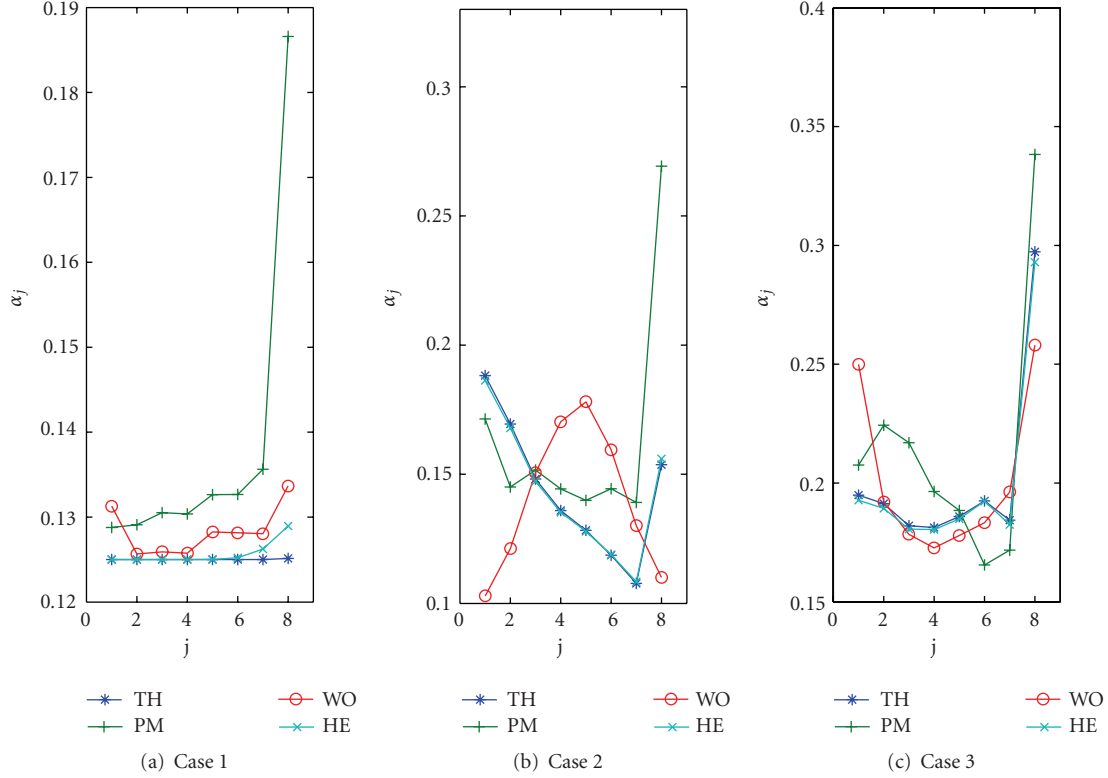
FIGURE 3: The optimized weighting factors for the bandlimited white noise process when the different multiple window sets are applied to the different cases, Thomson windows (stars), Peak Matched multiple windows (pluses), Welch method (circles), Hermite functions (crosses): (a) stationary process, (b) long-event nonstationary process, and (c) short-event nonstationary process.

2.2. Averaging and Scale Optimization. Usually, the spectrograms from different windows are equally weighted and averaged in the final estimate, that is,

$$\alpha_j = \frac{1}{I}, \quad j = 1 \cdots I. \tag{7}$$

Using equal weights according to (7) is referred to as equal weights (EQWEI).

The mean square error could be optimized according to the nMSE criterion, using equal weights scaled with a constant factor, that is,

$$\alpha_j^c = \frac{c}{I}, \quad j = 1 \cdots I, \tag{8}$$

where a closed form expression for the factor c is found from

$$c = \frac{\sum_{n=-T}^{T} \sum_{k=-M}^{M} S_x^2(n,k)/E^2\left[\hat{S}_x(n,k)\right]}{\sum_{n=-T}^{T} \sum_{k=-M}^{M} S_x(n,k)/E\left[\hat{S}_x(n,k)\right]}. \tag{9}$$

The weighting factors α_j^c are referred to as scaled weights (SCWEI).

3. Results

3.1. Bandlimited White Noise Process. The evaluation is done for different stationary and nonstationary processes.

The bandlimited white noise process with the covariance function

$$r_w(n-m) = B\frac{\sin(\pi B(n-m))}{\pi B(n-m)} \tag{10}$$

generates a Toeplitz covariance matrix $R_{\text{stat}}(n,m)$, which is shown in Figure 1(a) for $B = (8+3)/128 \approx 0.0859$, (Case 1). The locally stationary process approach [16, 17], where the covariance function of a nonstationary process is defined by

$$r_{ns}(n,m) = r_w(n-m)e^{-((n-m)/2F_{sc})^2}e^{-((n+m)/F_{sc})^2}, \tag{11}$$

gives a time-variable bandlimited spectrum where the time-variable power of the bandlimited white-noise process changes with a Gaussian envelope. Two examples are seen in Figure 1(b) as a long-event nonstationary process ($F_{sc} = 120$) (Case 2) and in Figure 1(c) as a short-event nonstationary process ($F_{sc} = 50$) (Case 3).

The weighting factors are optimized using four different sets of multiple windows. The Thomson multiple windows (TH) [1] give uncorrelated subspectra and thereby low variance for a stationary white noise process and the window functions $\mathbf{h}_j$ are given by the eigenvectors of the ($N \times N$) Toeplitz covariance matrix with elements given by (10) with

$$B = \frac{I+3}{N}, \tag{12}$$

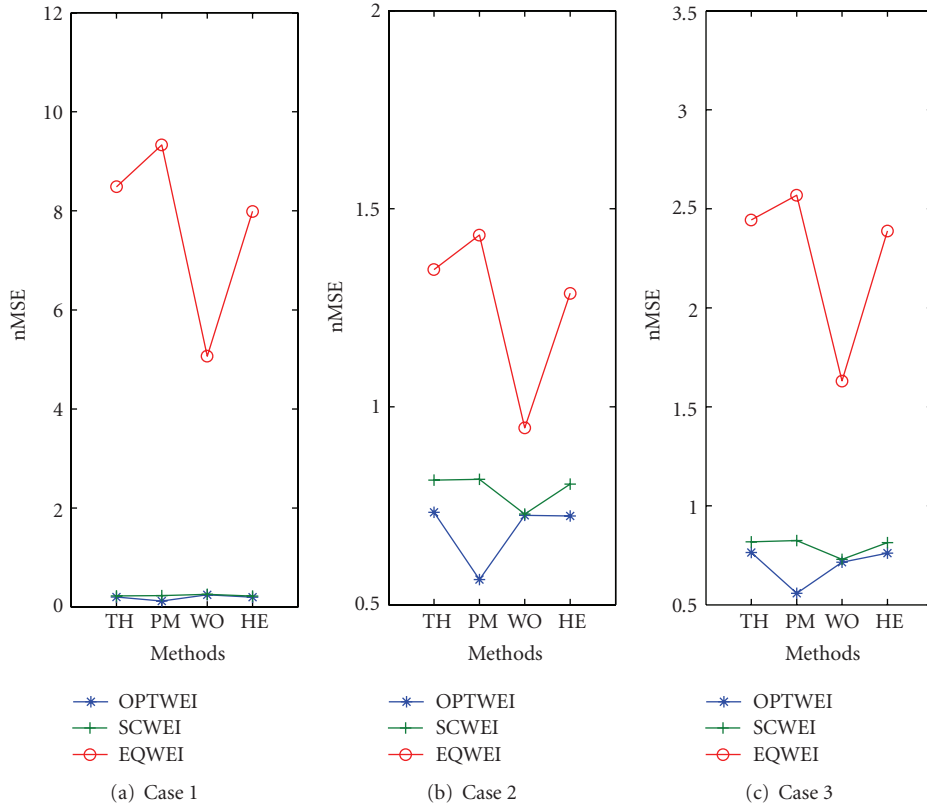where I is the number of multiple windows in the set.

FIGURE 4: The normalized mean square error for the three bandlimited peaked spectrum processes with EQWEI (circles), SCWEI (pluses), and OPTWEI (stars) for different window sets, Thomson multiple windows (TH), Peak-Matched multiple windows (PM), Welch method (WO) and Hermite functions (HE): (a) stationary process, (b) long-event nonstationary process, and (c) short-event nonstationary process.

The Peak-Matched (PM) multiple windows [6] are designed to give small correlation between subspectra when the spectrum of the stationary process includes peaks and notches. The windows are given by the solution of the generalized eigenvalue problem where the number of windows satisfies (12). Other parameters to be defined are the peak height chosen as $C = 20$ dB and the sidelobe suppression chosen as $K = 30$ dB [6]. The number of windows is I and is related to the bandwidth B and window length N as in (12).

The Welch method (WO) [2] utilizes time-shifted equal windows. In this paper we use a Hanning window of appropriate length so that the number of windows, I, is fitted into the total window length N with 50% overlap.

A set of Hermite functions (HE) is computed as

$$h_1(t) = e^{-t^2/2},$$

$$h_2(t) = 2te^{-t^2/2}, \qquad (13)$$

$$h_j(t) = 2th_{j-1}(t) - 2(j-2)h_{j-2}(t), \quad j = 3 \cdots I,$$

with $t = n/F_H$ for $n = -N/2 \cdots N/2 - 1$. The parameter F_H is chosen so that the first Hermite function is approximately equal to the first Slepian function of the Thomson method in each case (similar approach as in [8]).

The number of windows is chosen as $I = 8$ for all different methods and the window lengths are in all cases $N = 128$ giving $B \approx 0.0859$ for the Thomson and Peak-Matched multiple windows. For Case 1 (stationary process), the nMSE is computed and optimized only for the frequency, $f = 0$, that is, $M = 0$ and $T = 0$. For the nonstationary cases we choose, $M = 8$ and $T \cdot L = 192$ with $T = 8$, (17×17 values) for Case 2 and $T \cdot L = 32$ with $T = 8$, (17×17 values) for Case 3. These choices include the whole covariance matrix in each case and give a balance between different time and frequency values in the average.

The nMSE for Case 1 is shown in Figure 2(a), for the different multiple window sets, where the nMSE from EQWEI is shown with circles, the SCWEI with pluses, and OPTWEI with stars. The Thomson windows and Hermite functions are optimal for the stationary bandlimited white-noise process using the EQWEI and thereby the optimization of the weighting factors (SCWEI and OPTWEI) does not give any improvement of the nMSE. The Peak-Matched multiple windows do not give a small error using EQWEI, but with SCWEI and also OPTWEI, the nMSE decreases. The overall smallest error, however, is given by the Thomson and Hermite multiple windows as expected, as these two sets are optimal for a stationary bandlimited process. The Peak-Matched multiple windows and the Welch method are not
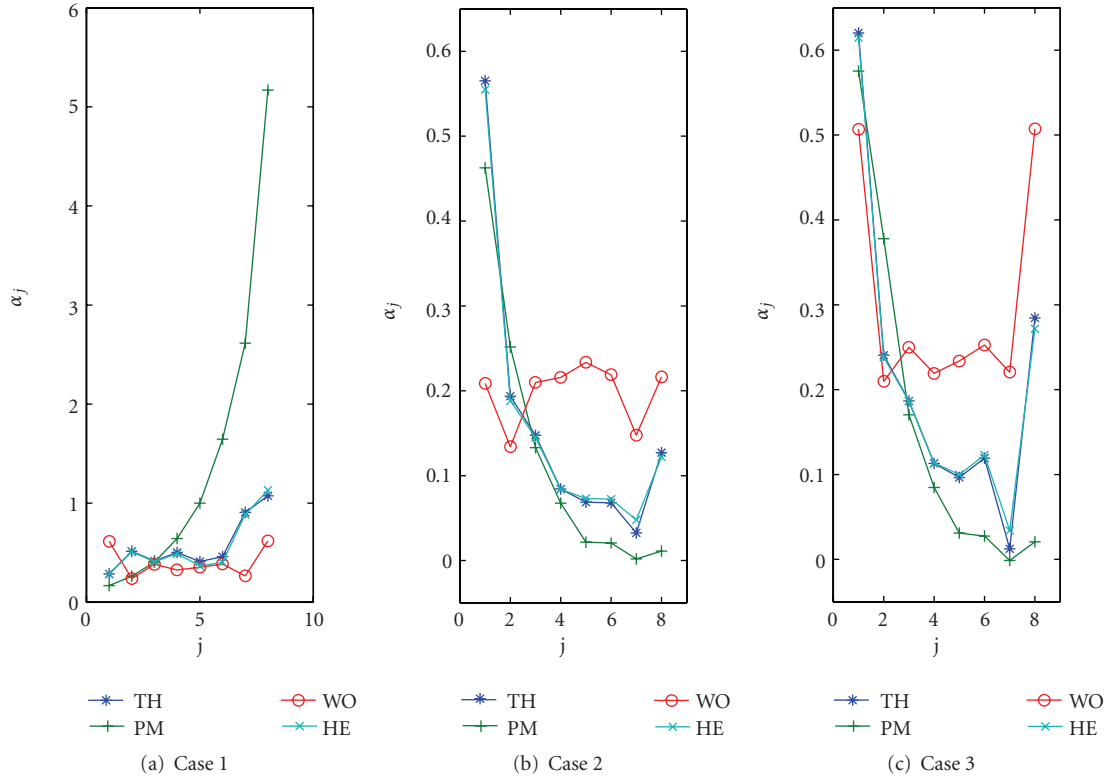
FIGURE 5: The optimized weighting factors for the bandlimited peaked spectrum process when the different multiple window sets are applied to the different cases, and Thomson windows (stars), Peak Matched multiple windows (pluses), Welch method (circles), and Hermite functions (crosses): (a) stationary process, (b) long-event nonstationary process, and (c) short-event nonstationary process.

able to reach the same nMSE even when the weighting factors are optimized.

In Case 2, Figure 2(b), the results from the long-event nonstationary process show the importance of using SCWEI and OPTWEI compared to the EQWEI in the nonstationary case. The difference of these two sets of weights is, however, not that large. It could also be noted that the Hermite functions perform slightly better than the Thomson multiple windows, which is in concordance with the study of nonstationary processes in [8]. In Case 3, Figure 2(c), using EQWEI on the short-event nonstationary process gives a very large error. Using SCWEI and OPTWEI gives a much lower nMSE.

The weighting factors α_j for OPTWEI are depicted in Figure 3 for the different window sets and the different cases. For the stationary process, the optimal weighting factors for the Thomson multiple windows (stars) are equally given by $\alpha_j = 1/8 \approx 0.125$. This is almost also the case for the Hermite functions (crosses), where the Peak-Matched multiple windows as well as the Welch method give more irregular weighting factors. Overall, however, the optimal weighting factors result almost in equally averaged spectra in all cases which coincides with theory for stationary processes. Of more interest is the long-event nonstationary process in Figure 3(b), where now both the Thomson and the Hermite windows give weights where more power is given to the spectrogram from the first window function with decreasing power to the following ones. Similar appearance is seen

for the Welch method, where we should remember that all the windows have the same frequency shape but have their power centered at different time points. Most of the power is laid on the resulting spectrograms of the middle windows which intuitively seems quite natural. For the more short-event nonstationary process, the resulting weighting factors have a different behavior, see Figure 3(c), where now the multiple windows located at end points of the time interval for the Welch method are given most power. This shows the importance of considering the weighting factors in estimation procedure. However, for the bandlimited white noise process, we should remember that using SCWEI in all cases gave almost as small error as OPTWEI.

As the optimization is made using a quasi-Newton algorithm, we cannot be sure of convergence to the global minimum. To verify, we optimize the weighting factors in all cases using 100 different initial sets of weighting factors. The set of initial values is randomly picked from a rectangle distribution with values between zero and one and the resulting sum is normalized to one. For all three bandlimited white noise processes and for all sets of windows, the optimization converged to the same minimum error for all the 100 cases, based on equal sets of weighting factors.

3.2. Bandlimited Peaked Spectrum Process. Instead of using a bandlimited white noise process, the stationary covariance function in (10) is replaced with the covariance function

(a) Ex. 1, Hanning

(b) Ex. 1 EQWEI/SCWEI

(c) Ex. 1, OPTWEI

(d) Ex. 2, Hanning
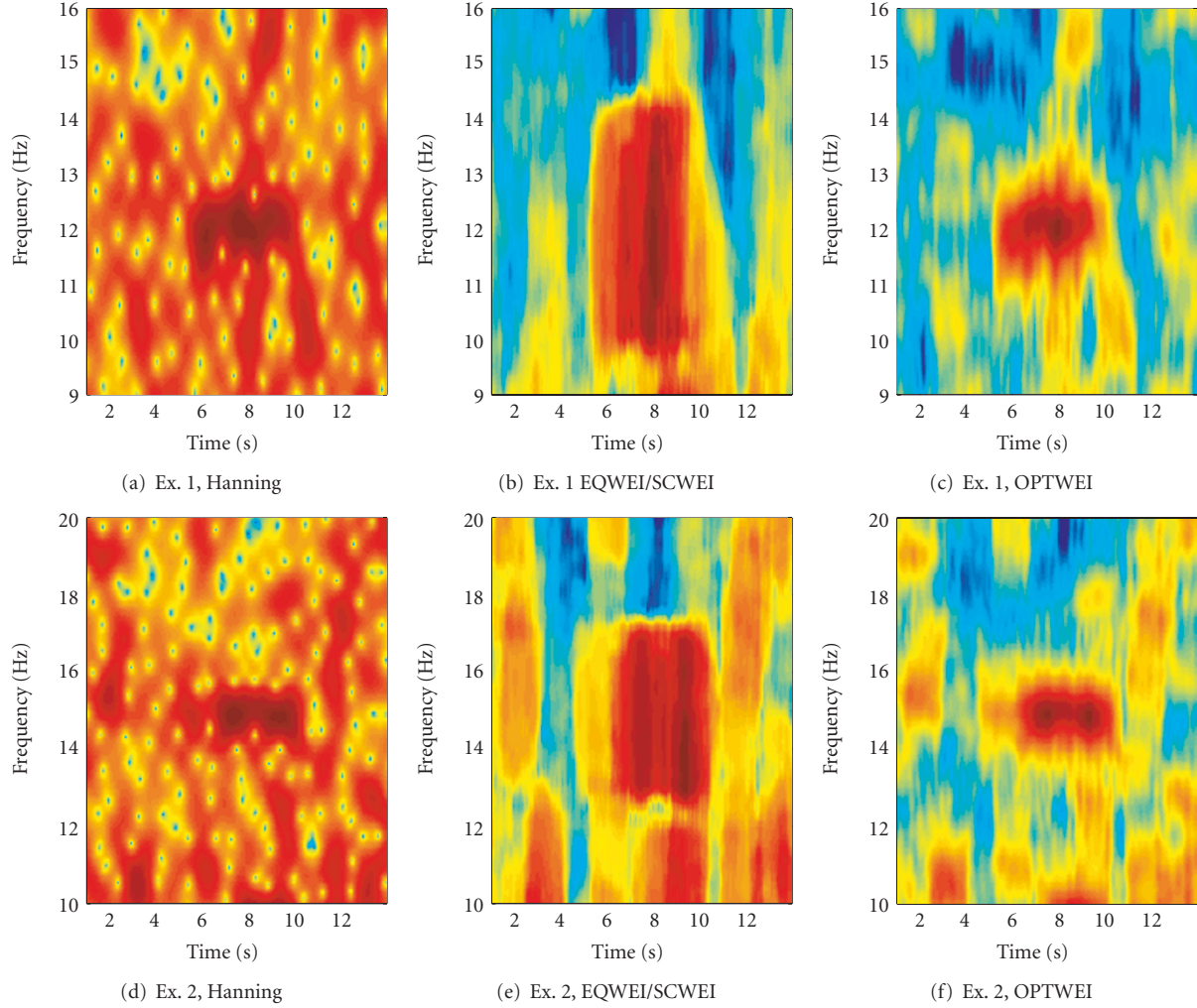
(e) Ex. 2, EQWEI/SCWEI

(f) Ex. 2, OPTWEI

FIGURE 6: Logarithmic spectrogram of SSVEP of long-event nonstationary character, estimated using a single window spectrogram with a Hanning window and the Peak-Matched multiple windows with EQWEI/SCWEI and with OPTWEI from Figure 5(b). Horizontal and vertical lines indicate flickering light of (a), (b), (c) 12 Hz, 5–10 s, (d), (e),(f) 15 Hz, 5–10 s.

$r_x(n - m)$ of a bandlimited peaked spectrum according to [6]:

$$S_x(f) = \begin{cases} e^{-2C|f|/10B\log_{10}(e)}, & |f| \leq \dfrac{B}{2}, \\ 0, & |f| > \dfrac{B}{2}. \end{cases} \quad (14)$$

In (14), $S_x(f)$ is a peaked spectrum with $S_x(0) = 1$, $B = 11/128 \approx 0.0859$, and $S_x(B/2) = -C$ dB, where $C = 20$ dB. The nonstationary processes are found from (11) with $r_w(n - m)$ replaced with $r_x(n - m)$.

The results from these processes are presented in Figure 4. For the stationary-peaked spectrum process, Figure 4(a), the EQWEI of the Welch method happens to give the smallest error. Using SCWEI we can lower the nMSE for all methods but using OPTWEI combined with Peak Matched multiple windows gives the smallest nMSE of all methods, which is concordance with [6, 9], where these windows and optimized weighting factors are shown to be optimal for this process. In Case 2, Figure 4(b), for the long-event nonstationary process, the benefit of using windows with properties suitable for the process becomes visible as the smallest nMSE is given for the Peak-Matched multiple windows combined with OPTWEI. In this case, the SCWEI is far from giving the same result. In Case 3, we also see a similar result for the short-event nonstationary process.

The different weighting factors are depicted in Figure 5, and for the stationary case, we see in Figure 5(a) the characteristic weighting for the PM given by $1/\lambda_i$, where λ_i is the eigenvalues from the solution of the eigenvalue problem giving the peak-matched multiple windows optimizing at the frequency $M = 0$; see [9]. Of more interest is the nonstationary processes of Cases 2 and 3. The optimal weightings of Peak-Matched and Thomson multiple windows are similar and they all give most power to the spectrogram from the first window and decreasing power to the following spectrograms. It is also worth notifying that this power increases for the short-event process of Case 3; see Figure 5(c). Most important, however, is how the weighting changes between

(a) Ex. 3, Hanning

(b) Ex. 3, EQWEI/SCWEI

(c) Ex. 3, OPTWEI

(d) Ex. 4, Hanning
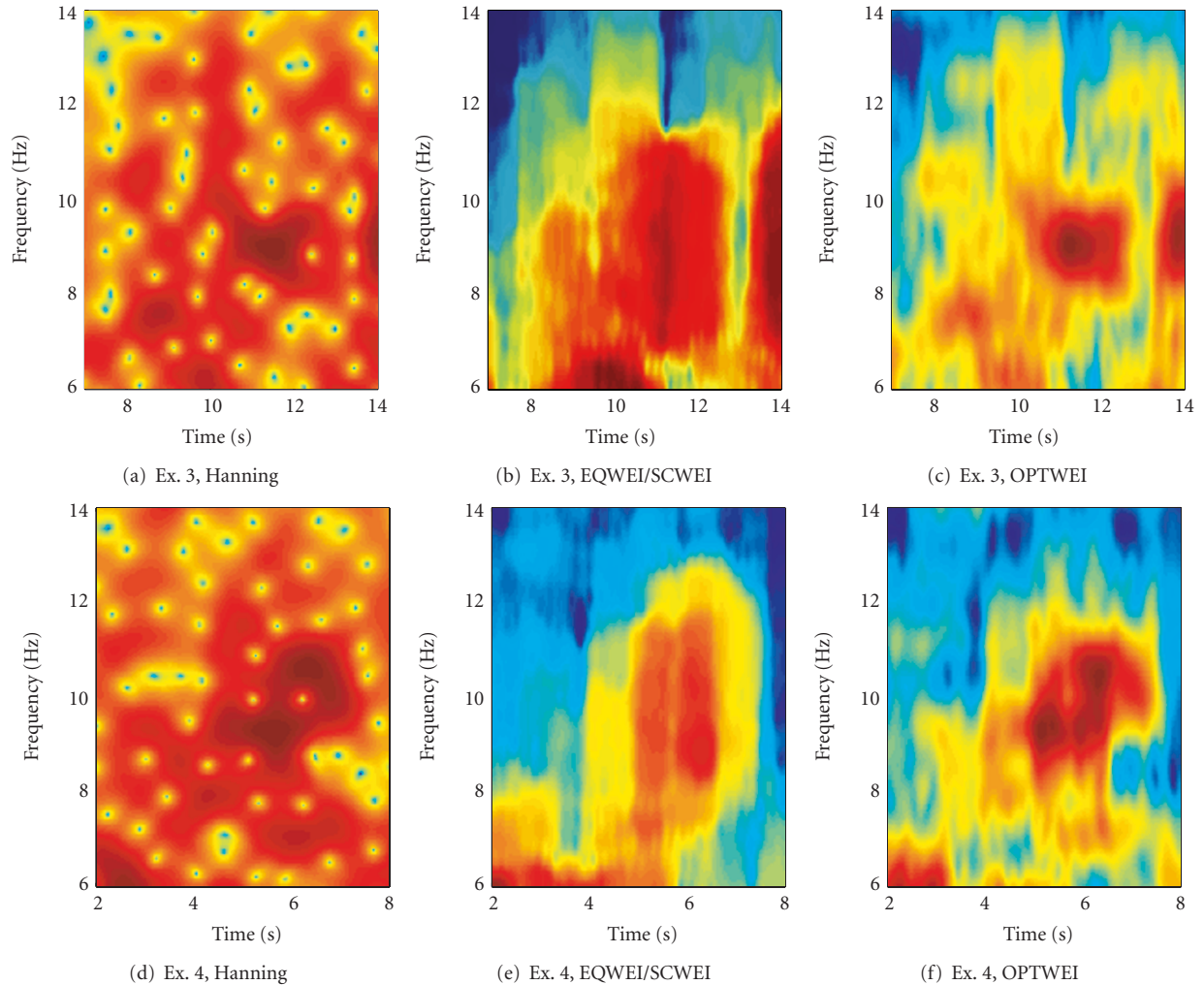
(e) Ex. 4, EQWEI/SCWEI

(f) Ex. 4, OPTWEI

FIGURE 7: Logarithmic spectrogram of SSVEP of short-event nonstationary character, estimated using a single window spectrogram with a Hanning window and the Peak-Matched multiple windows with EQWEI/SCWEI and with OPTWEI from **Figure** 5(b). Horizontal and vertical lines indicate flickering light of (a), (b), (c) 9 Hz, 10.4–12.2 s, (d), (e), (f) 9 Hz, 4.7–5.7 s.

the peaked spectrum process and the bandlimited white noise process and also how the weighting changes with the non-stationarity of the process.

The convergence of the optimization of the weighting factors in the three cases of bandlimited peaked spectrum processes is also investigated using the same randomly picked initial values as for the bandlimited white noise process and all different window sets. The results show that a minimum of 90% (usually around 95%) of the initial values converge to the global minimum giving the true optimal weighting factors. In the cases where the algorithm did not converge, the final error and the weighting factors were very far away from the true values and the divergence was easily discovered.

4. Real Data Examples

To show the performance for real-data, sampled ElectroEncephaloGram data (EEG) were studied, where a flickering light (Grass Photic stimulator Model PS22C) was introduced

at different time points. The light stimulation lasted approximately 1 s or 5 s. For a repetitive periodic visual stimulus a steady-state visual evoked potential (SSVEP) arises in the EEG. We assume the short stimulation (~1 s) to introduce a short-event nonstationary process and the long stimulation (~5 s) to introduce a long-event nonstationary process in the measured EEG. The subject was supine with closed eyes on a bed in a silent laboratory where ambient light was dimmed. The flickering light, with set frequency and time interval, was flashed at the subject from a distance of approximately 1 m. Data were recorded using a Neuroscan system with a digital amplifier (SYNAMP 5080, Neuro Scan, Inc.). Amplifier band-pass settings were 0.3 and 50 Hz. The sample rate was 256 Hz which was downsampled to a sample rate of $F_s = 64$ Hz in Matlab. In all examples channel PZ is chosen.

We illustrate the performance of the methods with examples of four different data sets. Example 1 is given from flickering light of 12 Hz and example 2 is given from flickering light of 15 Hz. For both these examples

the flickering lasts between time points 5 and 10 s and we assume these responses to be two long-event nonstationary processes. The third and fourth examples are given from flickering of 9 Hz between time points 10.4 and 12.2 s and time points 4.7 and 5.7 s, respectively. These two examples are assumed to be responses of short-event nonstationary processes.

We assume that we can model the different long-event nonstationary SSVEPs of Examples 1 and 2 as bandlimited peaked spectrum processes; see [18]. Logarithmic spectrograms are depicted in Figure 6, where we compare the spectrograms using the single Hanning window with different weightings of the Peak Matched multiple windows using OPTWEI from Figure 5(b) and EQWEI/SCWEI. In all cases, the window length is $N = 128$. Note that the spectrogram using EQWEI is equal to SCWEI as the difference is only a gain factor and the coloring is adjusted between the minimum and maximum value of each plot. We should also remember that the bandwidth of this estimator is $B = F_s(I + 3)/N = 5.5$ Hz, which is also clearly seen in the examples in Figures 6(b) and 6(e). The spread of the power caused by the large frequency bandwidth makes it difficult to know where the actual response frequency is located. Equal weighting of the multiple window spectrograms is not appropriate for data where it is important to locate the maximum power at a certain frequency. In the time-scale, however, we see that the resulting responses show up in the time interval where they should be located according to the stimuli given. The single window Hanning spectrograms are well resolved in frequency but the variance is however too large to be reliable, which is seen in Figures 6(a) and 6(d).

In Figure 7, we compare the spectrogram estimates using the single Hanning window with different weightings of the Peak-Matched multiple windows using OPTWEI from Figure 5(c) and EQWEI/SCWEI. The single Hanning spectrogram in Figures 7(a) and 7(d) is difficult to interpret and the spectrograms using EQWEI/SCWEI give a too wide estimate in frequency; see Figures 7(b) and 7(e). Using OPTWEI, the short-event nonstationary processes in Figures 7(c) and 7(f) are located correctly in the time interval as well as at the appropriate frequency. The last case however, Figure 7(f), has a large amount of power outside the time interval of stimuli, around 6-7 s. This is explained by the fact the stimulus sequence also activated the person and thereby also the alpha activity raised; see [18].

Even better is to actually estimate a model covariance function, using, for example, many trials from the same experiment for a robust estimate. From the properties of this modeled covariance function an appropriate set of multiple windows can be chosen and the weighting factors could be nMSE optimized to estimate the single stimulus response.

5. Conclusion

We compare the Hermite functions, the Thomson windows, the Peak-Matched Multiple Windows, and the Welch windows and compute the performance with optimal weighting factors for different stationary and nonstationary processes.

The cost function for optimization is the normalized mean square error where the normalization factor is the multiple window spectrogram. This means that the unknown weighting factors will be present in the numerator as well as in the denominator. A quasi-Newton algorithm is used for the optimization. The results show that the weighting factors, as well as the shape of the windows, are important factors for a small error. It is also shown that a scaling optimization of the usual averaging could give almost as small mean square error as an optimization of the individual weighting factors in case of a smooth spectrum. For a peaked spectrum, a significant reduction of the normalized mean square error is achieved using individual optimization of the weights.

References

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

[2] P. D. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[3] T. P. Bronez, "On the performance advantage of multitaper spectral analysis," *IEEE Transactions on Signal Processing*, vol. 40, no. 12, pp. 2941–2946, 1992.

[4] A. T. Walden, E. McCoy, and D. B. Percival, "Variance of multitaper spectrum estimates for real Gaussian processes," *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 479–482, 1994.

[5] K. S. Riedel and A. Sidorenko, "Minimum bias multiple taper spectral estimation," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 188–195, 1995.

[6] M. Hansson and G. Salomonsson, "A multiple window method for estimation of peaked spectra," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 778–781, 1997.

[7] B. Farhang-Boroujeny, "Prolate filters for nonadaptive multitaper spectral estimators with high spectral dynamic range," *IEEE Signal Processing Letters*, vol. 15, pp. 457–460, 2008.

[8] M. Bayram and R. G. Baraniuk, "Multiple window time-frequency analysis," in *Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 173–176, June 1996.

[9] M. Hansson, "Optimized weighted averaging of peak matched multiple window spectrum estimates," *IEEE Transactions on Signal Processing*, vol. 47, no. 4, pp. 1141–1146, 1999.

[10] F. Çakrak and P. J. Loughlin, "Multiple window time-varying spectral analysis," *IEEE Transactions on Signal Processing*, vol. 49, no. 2, pp. 448–453, 2001.

[11] J. Xiao and P. Flandrin, "Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2851–2860, 2007.

[12] W. J. Williams and S. Aviyente, "Spectrogram decompositions of time-frequency distributions," in *Proceedings of the 6th International, Symposium on Signal Processing and its Applications (ISSPA '01)*, pp. 587–590, 2001.

[13] L. L. Scharf and B. Friedlander, "Toeplitz and Hankel kernels for estimating time-varying spectra of discrete-time random

processes," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 179–189, 2001.

[14] M. Hansson-Sandsten and J. Sandberg, "Optimization of weighting factors for multiple window time-frequency analysis," in *Proceedings of the European Signal Processing Conference (EUSIPCO '09)*, Glasgow, UK, 2009.

[15] R. Fletcher, *Practical Methods for Optimization*, John Wiley & Sons, New York, NY, USA, 1987.

[16] R. A. Silverman, "Locally stationary random processes," *IRE Transactions on Information Theory*, vol. 3, pp. 182–187, 1957.

[17] P. Wahlberg and M. Hansson, "Kernels and multiple windows for estimation of the Wigner-Ville spectrum of Gaussian locally stationary processes," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 73–84, 2007.

[18] M. Hansson and M. Lindgren, "Multiple-window spectrogram of peaks due to transients in the electroencephalogram," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, pp. 284–293, 2001.