Local Polynomial Modelling of Time-varying Autoregressive Processes and its Application to the Analysis of Event-related Electroencephalogram

Z. G. Zhang, S. C. Chan, and Y. S. Hung

Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
{ zgzhang, scchan, yshung}@eee.hku.hk

Abstract— **This paper proposes a new method for identification of time-varying autoregressive (TVAR) models based on local polynomial modeling (LPM) and applies it to investigate the dynamic spectral information of event-related electroencephalogram (EEG). The proposed method models the TVAR coefficients locally by polynomials and estimates those using least-squares estimation with a kernel having a certain bandwidth. A data-driven variable bandwidth selection method is developed to obtain the optimal bandwidth, which minimizes the mean squared error (MSE). Simulation results show that the LPM-based TVAR identification method outperforms conventional methods for different scenarios. The advantages of the LPM method make it a useful high-resolution time-frequency analysis (TFA) technique for nonstationary biomedical signals like EEG. Experimental results show that the LPM method can reveal more meaningful time-frequency characteristics than wavelet transform.**

I. INTRODUCTION

Time-varying autoregressive (TVAR) models are widely used to describe the dynamics of nonstationary signals, such as speech signals, communication signals, and biomedical signals [1], [2]. A nonstationary signal $s(t)$ can be expressed in the form of a TVAR model as follows:

$$s(t) = \sum_{k=1}^{L} a_k(t)s(t - kT_s) + \varepsilon(t), \qquad (1)$$

where $\boldsymbol{a}(t) = [a_1(t), \ a_2(t), \cdots, a_L(t)]^T$ is the TVAR coefficient vector, L is the order of the model, T_s is the sampling period, and $\varepsilon(t)$ is a zero mean additive white Gaussian process with variance $\sigma^2(t)$. Accurate identification of TVAR models, i.e., estimation of the TVAR coefficients $\boldsymbol{a}(t)$, is crucial to understand the dynamics of signals and to predict future observations. Generally, two classes of TVAR identification methods are common in practice [2], and they are 1) adaptive filtering and Kalman filtering (KF); 2) basis expansion modelling (BEM). Although these methods usually offer efficient implementation and different tradeoffs between performance and complexity, their performances are often dependent on prior information and model parameters [1], [2]. Therefore, the estimation accuracy of these methods is sometimes limited when prior knowledge is vague or parameters are incorrectly specified [2].

In this paper we propose a new TVAR identification method based on a local polynomial modelling (LPM) approach, which does not require explicit information of the coefficient variations. It can achieve a good tracking performance and a better flexibility in a variety of experimental conditions. The proposed LPM method models the TVAR coefficients locally by a set of polynomials and a window or kernel having a certain bandwidth. The estimation of TVAR coefficients is then reduced to the estimation of polynomial coefficients, which can be performed using the least-squares (LS) technique [3]-[7]. Since the TVAR coefficients may vary considerably over time, it is crucial to choose a proper local bandwidth to achieve the best bias-variance tradeoff. For slow-varying coefficients, we would like the bandwidth to be large to reduce the estimation error. For fast-varying coefficients, a small bandwidth is desirable in order to reduce the bias error. A data-driven bandwidth selection technique for LPM is therefore developed, where the bandwidths are determined locally and adaptively by minimizing the approximated mean squared error (MSE), which is the sum of squared bias and variance. Hence, the proposed LPM-based TVAR identification method can obtain good tracking performances for both slowly-varying and fast-varying coefficients by employing variable bandwidths.

Next, we propose to apply the LPM-based TVAR identification method for time-frequency analysis (TFA) of event-related electroencephalogram (EEG). Event-related EEG is the stimulus-induced electrical activities measured on the scalp, and it turns out to be a non-stationary process, because the states of brain regions change rapidly on a time scale of milliseconds [8]. Previous studies showed that event-related EEG consists of a large number of components, such as event-related potentials (ERPs), event-related (de)synchronization (ERS/ERD), and other oscillations and potentials [8], [9]. TFA techniques have been widely adopted in EEG study, because the time-varying spectral components can effectively represent ongoing EEG rhythmic activities, providing useful information for discovering the nature and timing of sensory and cognitive processes.

Currently the most popular TFA method for EEG is the wavelet transform (WT) [10]. The superiority of WT is that it addresses the time-frequency resolution tradeoff problem by applying a short window at high frequency and a long window at low frequency. However, WT has a degraded frequency resolution for high-frequency components and a degraded time resolution for low-frequency components. Therefore, WT may not able to reveal useful components in the whole time-frequency plane and cannot acquire accurate temporal information on rapid (maybe event-related) changes of EEG. Another type of TFA methods (parametric methods) describe an EEG signal by a parametric model, and the time-varying power spectral density (PSD) of the signal can be obtained from the estimated model coefficients, hence providing a time-frequency distribution (TFD) of the EEG. Parametric methods generally have a high frequency

resolution than nonparametric methods like WT and Lomb periodogram [11]-[13]. In EEG study, TVAR models and KF are respectively the most popular parametric model and identification method. However, KF is sensitive to model parameters so that it often leads to a long tracking lag or a large estimation variance (time-frequency resolution tradeoff) [11]. The proposed LPM-based TVAR identification method is well-suited for TFA of EEG because its bandwidth (time scale) is adaptively selected so that the TFD can achieve high time-resolution without any loss of frequency resolution. EEG signals recorded in a visual oddball paradigm are tested and the experimental results show that the LPM method can reveal more meaningful time-frequency components than WT.

The rest of the paper is organized as follows. In Section II, the LPM for TVAR models is introduced. The adaptive bandwidth selection method for the LPM-based TVAR identification is developed in Section III. The proposed method is further extended to TFA of EEG in Section IV. Experimental results are presented in Section V. Finally, conclusions are drawn in Section VI.

II. LPM FOR IDENTIFICATION OF TVAR MODEL

In the TVAR model of (1), the k-th coefficient $a_k(t)$ can be modelled locally at $t=t_0$ as a p-th order polynomial [3]-[6]:

$$a_k(t) \approx \sum_{j=0}^{p} \frac{1}{j!}\alpha_k^{(j)}(t_0)(t-t_0)^j , \qquad (2)$$

where $\alpha_k^{(j)}(t_0)$ are the polynomial coefficients. These polynomial coefficients can be estimated locally by maximum likelihood (ML) estimation. Since the additive noise is zero mean and white Gaussian distributed, the ML estimation is equivalent to minimizing a locally weighted LS criterion as follows:

$$\min_{\beta} \sum_{i=1}^{n}[s_i - \sum_{k=1}^{L}\sum_{j=0}^{p}\beta_k^{(j)}(t_0)(t_i-t_0)^j s(t_i-kT_s)]^2 K_h(t_i-t_0) , \qquad (3)$$

where $s_i = s(t_i)$, $\beta_k^{(j)}(t_0) = \frac{1}{j!}a_k^{(j)}(t_0)$, and $K_h(t_i-t_0) = \frac{1}{h}K(\frac{1}{h}(t_i-t_0))$ is a kernel which controls the number of neighboring samples around t_0 used to estimate $\beta_k^{(j)}(t_0)$. The kernel $K_h(\cdot)$ is obtained by scaling a basis kernel $K(\cdot)$ using a bandwidth of h.

Next, we rewrite (3) more compactly in a matrix form as

$$\min_{\beta}(y - X\beta(t_0))^T W(y - X\beta(t_0)) , \qquad (4)$$

where $y = [s_1, s_2, \cdots, s_n]^T$, $W = diag\{K_h(t_i - t_0)\}$,

$$X = \begin{pmatrix} x^T(t_1) & (t_1-t_0)x^T(t_1) & \cdots & (t_1-t_0)^p x^T(t_1) \\ x^T(t_2) & (t_2-t_0)x^T(t_2) & \cdots & (t_2-t_0)^p x^T(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x^T(t_n) & (t_n-t_0)x^T(t_n) & \cdots & (t_n-t_0)^p x^T(t_n) \end{pmatrix}$$ with

$x(t_i) = [s(t_i-T_s), \cdots, s(t_i-LT_s)]^T$, and $\beta(t_0) = \{[\beta^{(p)}(t_0)]^T, \cdots, [\beta^{(p)}(t_0)]^T\}^T$ with $\beta^{(j)}(t_0) = [\beta_1^{(j)}(t_0), \cdots, \beta_L^{(j)}(t_0)]^T$. The LS solution to (4) is:

$$\hat{\beta}(t_0) = (X^T W X)^{-1} X^T W y , \qquad (5)$$

and finally the k-th TVAR coefficient at time instant t_0 is obtained as $\hat{a}_k(t_0) = \hat{\alpha}_k^{(0)}(t_0) = \hat{\beta}_k^{(0)}(t_0)$.

The TVAR model can be viewed as a special case of a time-varying linear regression model, where $s(t)$, $s(t-kT_s)$, and $a(t)$ are respectively the observations, explanatory variables and regression coefficients. In [7], LPM has been proposed for estimating the regression coefficients of the time-varying linear regression model, and the asymptotic properties of bias and variance have been derived. In the content of TVAR model, the asymptotic bias and variance of $\hat{\beta}_k^{(j)}(t_0)$ are respectively expressed as:

$$Bias(\hat{\beta}_k^{(j)}(t_0)) = \frac{\widetilde{S}_{b,k}^{(j)}\beta_k^{(p+1)}(t_0)j!h^{p+1-j}}{(p+1)!s(t_0-kT_s)} , \qquad (6)$$

$$Var(\hat{\beta}_k^{(j)}(t_0)) = \frac{\widetilde{S}_{v,k}^{(j)}j!^2 \sigma^2(t_0)}{nh^{2j+1}s^2(t_0-kT_s)} , \qquad (7)$$

where $\widetilde{S}_{b,k}^{(j)}$ and $\widetilde{S}_{v,k}^{(j)}$ are two quantities determined by the basis kernel $K(\cdot)$ and indices j and k, and $\beta_k^{(p+1)}(t_0)$ is the $(p+1)$-th derivative of $\beta_k(t_0)$. The derivation and details of (6) and (7) can be found in [7]. It can be seen from (6) and (7) that, when h increases, the bias will increase while the variance will decrease. Hence, there exists a locally optimal bandwidth $h^{opt}(k,t_0)$ for estimating $\beta_k^{(j)}(t_0)$, and $h^{opt}(k,t_0)$ should minimize the MSE. As some of the quantities in (6) and (7) are difficult to be calculated directly, the optimal bandwidth is difficult to be estimated accurately. We next introduce an empirical method to select the optimal bandwidth from a set of possible bandwidths.

III. ADAPTIVE BANDWIDTH SELECTION FOR LPM

In [7], we have developed an effective bandwidth selection method for time-varying linear regression models. This method is also applicable to the TVAR models with slight modifications. Given a finite set of bandwidth parameters in a geometric grid, say,

$$H = \{h_j \mid h_j = h_0(h_a)^j, \ j = 0, \cdots, J-1\} , \qquad (8)$$

where $h_a > 1$ is a step factor, $h_0 > 0$ is a base bandwidth, and J is the number of bandwidths, the empirical bandwidth selection method in [7] approximates the bias, variance, and MSE values of each bandwidth and determines an optimal bandwidth as the one that minimizes the approximated MSE.

Firstly, the bias of $\hat{\beta}(t_0)$ can be estimated using a Taylor's expansion with an order $p+p_{ex}$:

$$\hat{b}_\beta(t_0) = (X^T W X)^{-1} X^T W \tau , \qquad (9)$$

where τ is an $n \times 1$ vector with $\sum_{\varsigma=1}^{p_{ex}}\sum_{k=1}^{L}\beta_k^{(p+\varsigma)}(t_0)(t_i-t_0)^{p+\varsigma} s(t_0-kT_s)$ as its i-th element. $\beta_k^{(p+\varsigma)}(t_0)$ can be estimated by fitting a polynomial of degree $p+p_{ex}$, where p_{ex} is an excess order [3]. However, the $(p+p_{ex})$-th order LPM still requires an initial bandwidth parameter, which is called the pilot bandwidth h^* and its selection will be discussed later.

Secondly, suppose local homoscedasticity, the estimation covariance of $\hat{\beta}(t_0)$ can be estimated as:

$$\hat{V}_\beta(t_0) = (X^T W X)^{-1} X^T W^2 X (X^T W X)^{-1} \sigma^2(t_0) . \qquad (10)$$

The noise variance $\sigma^2(t_0)$ can be obtained as the normalized weighted residual sum of squares [3]:

$$\hat{\sigma}^2(t_0) = \frac{\sum_{i=1}^{n}(s_i - \hat{s}_i)^2 K_{h^*}(t_i-t_0)}{tr\{W^* - W^* X^*(X^{*T}W^* X^*)^{-1} X^{*T}W^*\}} , \qquad (11)$$

where X^* and W^* are respectively the design matrix and weighting matrix in the $(p+p_{ex})$-th order LPM using the pilot bandwidth h^*. Consequently, the MSE of $\hat{\beta}(t_0)$ is

$$MSE(\hat{\beta}(t_0)) = \hat{b}_\beta(t_0)\hat{b}_\beta^T + tr(\hat{V}_\beta(t_0)) , \qquad (12)$$

and the optimal bandwidth is determined as the bandwidth having the minimum MSE.

To select a pilot bandwidth h^*, the intersection of confidence interval (ICI) method can be used [5]-[7]. The ICI method selects bandwidth $h_{ICI}(k,t)$ for each k-th coefficient at each time instant t by calculating the variance of (10) using different bandwidths in H. Because the TVAR coefficients at

each time should be estimated as a whole (using one bandwidth), not individually (using different bandwidths for different k), $h_{ICI}(k,t)$ will be combined (generally, averaged) to produce the pilot bandwidth $h^*(t)$. The algorithm of ICI is omitted to save space, and more details can be found in [5]-[7].

IV. LPM-BASED TVAR FOR EEG

We now apply the LPM-based TVAR identification method for TFA of event-related EEG. By modelling an EEG signal by the TVAR model of (1), we can obtain its time-varying PSD from the estimated coefficients and noise variance as:

$$P(t,f) = \hat{\sigma}^2(t) \left/ \left| 1 - \sum_{k=1}^{L} \hat{a}_k(t) e^{-j2\pi fk} \right|^2 \right. . \quad (13)$$

We now discuss the parameter selection when applying the LPM method in our EEG study. Following the literature [3]-[7], the Epanechnikov kernel is employed due to its low complexity. The effective size, i.e., the time interval, of an Epanechnikov kernel is $2h$. Next, to determine the bandwidth set H of (8), h_0 should be selected as the minimum bandwidth that makes (5) solvable (overdetermined). Because a $(p+p_{ex})$-th order LPM is used to approximate MSE, the number of samples included in the kernel should be larger than $(p+p_{ex}+1)L$ (the dimension of $\hat{\boldsymbol{\beta}}(t)$). If the data are uniformly distributed at a sampling frequency of f_s, it is required that

$$h_0 \geq (p + p_{ex} + 1)L / (2f_s) . \quad (14)$$

Moreover, to achieve a good tradeoff between performance and complexity, in our study, the largest bandwidth of H is selected to make the largest kernel cover at least 1/8 of the data set, and the number of bandwidths, J, is set as 4. As for the selection of polynomial order p and p_{ex}, since the major advantage of the adaptive bandwidth selection is its fine time-resolution, p and p_{ex} should be as small as possible to fulfill the condition of (14). Hence, we set $p=0$ and $p_{ex}=1$ in this study. The proposed bandwidth setting gave satisfactory results in our experiments.

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. Simulated TVAR Model

We first used a second-order TVAR model to illustrate the good performance of the LPM method. The total number of samples is 250 and the additive noise is zero mean with a variance of 0.01. The coefficient a_1 jumps from -1.9 to -1.0 at time 125, as shown in Figure 1, while a_2 equals a constant -1.0. For comparison, the KF and BEM methods are also tested. In KF, the state transition matrix is chosen as identity matrices to produce a random-walk KF. As mentioned in [1], [11], the ratio between the variances of the state noise and the observation noise, R_σ, determines the tracking performance of KF. Here, two values (0.01, 1) of R_σ are tested in KF. In BEM, Legendre polynomial functions are employed and the number of basis functions is 5.

One simulation is presented in Figure 1. It can be seen that, BEM gives a smooth estimation, but rapid changes are smoothed out at the same time. The performance of KF heavily depends on the selection of R_σ. A large R_σ enables the KF to track fast-varying coefficients while a small R_σ helps to reduce the variance of slowly-varying coefficients. In LPM, a small bandwidth can detect fast change of coefficients, but it may lead to large estimation variations, and a large bandwidth can obtain smooth estimates when coefficients vary slowly, but it cannot accurately estimate fast-varying coefficients. LPM with variable bandwidth can obtain satisfactory results for the whole data because of the adaptive variable bandwidths used. The mean squared deviation (MSD) values,

$$MSD = 10\log_{10}\left\{ \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{L}[a_k(t_i) - \hat{a}_k(t_i)]^2 \right\},$$ which are obtained by averaging from 100 Monte-Carlo simulations, are given in notation of Figure 1, and we can see that the LPM with variable bandwidth method has the minimum MSD value.
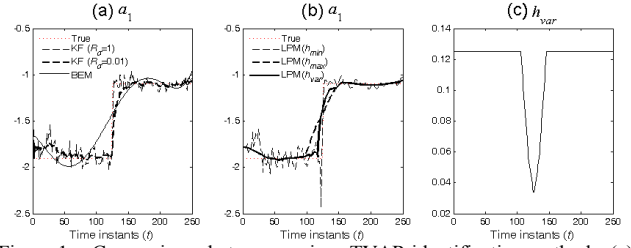

Figure 1. Comparisons between various TVAR identification methods: (a) KF with different R_σ and BEM, (b) LPM with a constant small kernel (h_{min}) and a constant large kernel (h_{max}), and adaptive variable bandwidths (h_{var}), (d) the estimated adaptive bandwidths h_{var}. The MSD values from 100 Monte-Carlo simulations are, KF (R_σ=1): -16.9 dB; KF (R_σ=0.01): -16.5 dB; BEM: -15.5 dB; LPM (h_{min}): -17.9 dB; LPM (h_{max}): -16.8dB; LPM (h_{var}): -20.2dB.

B. TFA of Simulated Signals

We next compare various TFA methods, including short-time Fourier transform (STFT), WT and TVAR-based PSD with different identification methods, using a chirp signal with an instantaneous frequency $w(t)$ shown in Figure 2 as red lines. The number of samples is 250 and the sampling frequency is 250Hz. An additive Gaussian white noise with a SNR of 10 dB is added. Hann windows with sizes 20 and 100 are used in STFT, while the Morlet wavelet is used in WT. The order of TVAR models is selected as 4 according to the Akaike's information criterion. Parameters for TVAR identification methods are the same as those in previous simulation. One representative simulation is presented in Figure 2. The MSD values of estimated instantaneous frequency $\hat{w}(t)$, $MSD = 10\log_{10}\left\{ \frac{1}{n}\sum_{i=1}^{n}[w(t_i) - \hat{w}(t_i)]^2 \right\}$, are obtained by averaging from 100 Monte-Carlo simulations and are listed in notation of Figure 2. It can be seen that the TFD obtained using the LPM with variable bandwidth has better time-frequency resolution than the other methods tested and it can attain fine frequency resolution in a wide frequency range.
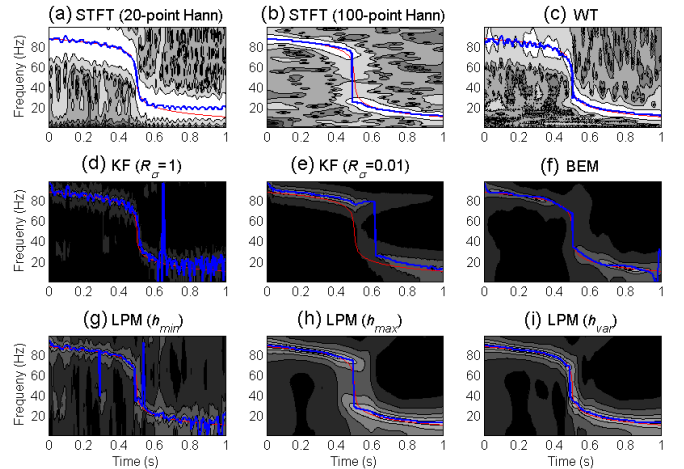

Figure 2. Comparisons between various TFA methods. The red lines are true instantaneous frequency, and the blue lines are estimated instantaneous frequency (the frequency with the maximum PSD value at each time point). The MSDs of instantaneous frequency estimates from 100 Monte-Carlo simulations are respectively: (a) 11.7 dB; (b) 13.7 dB; (c) 10.2 dB; (d) 19.3 dB; (e) 25.2 dB; (f) 11.2 dB; (g) 16.4 dB; (h): 10.8 dB; (i): 9.5 dB.

C. Real Event-related EEG

A set of visual oddball event-related EEG was collected from the Electrical Geodesic 128-channel EEG system, with vertex referenced. The standard stimulus is a visual character 'O' with 80% occurrence, while the target stimulus is a visual character 'X' with 20% occurrence. Inter-stimulus interval was 1500ms and the duration of one trial was 1000 ms (100 ms before and 900 ms after stimulus onset). Twelve subjects were recruited and they were required to press the correct button immediately after he/she recognized the stimulus. A total of 250 stimuli were presented, and the sampling rate was 250 Hz. Automatic artifact correction for EEG by blind source separation was performed to attenuate the artifacts [14]. The behavioral data, such as onset time, subject response time and subject response accuracy, were also recorded. The order of TVAR models is selected as 12 by the Akaike's information criterion, which is consistent with previous study [11].

Figure 3 (a) shows the conventional WT- and LPM-based TFDs of an averaged target-response EEG (averaged over 50 trials; recorded at O2) of one subject. It can be seen that, compared with WT, the LPM-based TFD can achieve a higher frequency resolution for high-frequency components and a higher time resolution for low-frequency components. The power peaks of low-frequency components A and B in alpha and theta bands can be seen clearly in LPM-based TFD, and they should be due to the P300 and other early ERP components [8], [9]. A 60-Hz component C starting from 0ms is pronounced in the LPM-based TFD, which should be the steady-state visual evoked potential (SSVEP) in response to computer monitor flicker [15]. In addition, we can see clearly a beta-band oscillation (17-20Hz, component D) after 700ms, which should be the visual attention waiting for the next stimulus [9], [16]. For single-trial EEG, LPM can also achieve a meaningful result. Figure 3 (b) illustrates the TFDs of a single-trial target EEG signal when the subject pressed the wrong button. We can see that, except the 60-Hz SSVEP, P300, and post-response beta oscillation, the component E around 600ms and theta band is distinct. This component should be linked to the error-related negativity (ERN), which occurs when the subject knows his/her error immediately after the response [9]. On the other hand, the ERN pattern in WT is very vague due to the coarse resolution of WT in low frequency bands.

Due to page limitation, more results (different subject and different electrodes) are not presented here. It should be noted that LPM is a powerful method to discover fine time-frequency components of EEG. It facilitates the further analysis, such as statistical analysis, source location, etc., so as to interpret more thoroughly the disclosed time-frequency components and reveal the underlying physiological and psychological mechanisms.

VI. CONCLUSION

A novel LPM approach and a data-driven variable bandwidth selection scheme for identification of TVAR processes were presented. Simulation results showed that the performance of the LPM with adaptive bandwidth selection method is superior to conventional TVAR identification methods in a variety of testing conditions. The proposed method was further applied to TFA of event-related EEG, and experimental results showed that the LPM method can reveal more clearly time-frequency components than WT.
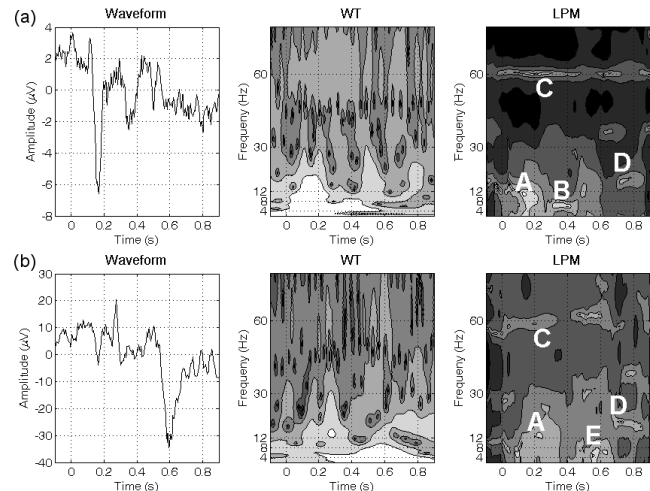
ACKNOWLEDGMENT

Figure 3. Comparisons between morlet WT and LPM-based PSD estimation: (a) averaged target EEG, and (b) single-trial target EEG with a wrong response.

REFERENCES

[1] M. Niedzwiecki, *Identification of Time-varying Processes*, Chechester: John Wiley, 2000.

[2] T. Hsiao, "Identification of time-varying autoregressive systems using maximum a posteriori estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3497-3509, Aug. 2008.

[3] J. Fan and I. Gijbels, *Local Polynomial Modelling and Its Applications*, London: Chapman and Hall, 1996.

[4] W. Zhang and S. Y. Lee, "Variable bandwidth selection in varying-coefficient models," *J. Multivar. Anal.*, pp. 74, no. 1, pp. 116-134, 2000.

[5] V. Katkovnik, "A new method for varying adaptive bandwidth selection," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2567-2571, Sept. 1999.

[6] Z. G. Zhang, S. C. Chan, K. L. Ho, and K. C. Ho, "On bandwidth selection in local polynomial regression analysis and its application to multi-resolution analysis of non-uniform data," *J. Signal Process. Syst.*, vol. 52, no. 3, pp. 263-280, Sept. 2008.

[7] S. C. Chan and Z. G. Zhang, "Local polynomial modeling and bandwidth selection for time-varying linear models," in *Proc. International Conferences on Information, Communications and Signal Processing (ICICS 2009)*, Macau, China, 7-10 Dec., 2009.

[8] E. Niedermeyer and F. H. Lopes da Silva, *Electroencephalography: Basic Principles, Clinical Applications, And Related Fields*, 5th ed., Philadelphia: Lippincott Williams & Wilkins, 2005.

[9] S. Makeig, A. Delorme, M. Westerfield, T. P. Jung, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Electroencephalographic brain dynamics following manually responded visual targets," *PLoS Biol.*, vol. 2, no. 6, pp. 747-762, June 2004.

[10] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc. IEEE*, vol. 84, pp. 626-638, Apr. 1996.

[11] M. Arnold, H. Miltner, H.Witte, R. Bauer, and C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 553–562, May 1998.

[12] M. P. Tarvainen, J. K. Hiltunen, P. O. Ranta-aho, and P. A. Karjalainen, "Estimation of nonstationary EEG with Kalman smoother approach: an application to event-related synchronization (ERS)," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 3, pp. 516-524, March 2004.

[13] Z. G. Zhang, X. L. Cai, S. C. Chan, Y. Hu, L. Hu, and C. Q. Chang, "Time-frequency coherence analysis of multi-channel event-related potential using adaptive windowed Lomb periodogram," in *Proc. IEEE/EMBS Conference on Neural Engineering (NER 2009)*, Antalya, Turkey, 29 Apr.-2 May, 2009, pp. 657-660.

[14] K. H. Ting, P. C. W. Fung, C. Chang, and F. H. Y. Chan, "Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only," *Med. Eng. Phys.*, vol. 28, pp. 780-794, 2006.

[15] E. Lyskov, V. Ponomarev, M. Sandström, K. H. Mild, and S. Medvedev, "Steady-state visual evoked potentials to computer monitor flicker," *Int. J. Psychophysiol.*, vol. 28, no. 3, pp. 285-290, May 1998.

[16] A. Wróbel, "Beta activity: a carrier for visual attention," *Acta. Neurobiol. Exp.*, vol. 60, no.2 , pp. 247-260, 2000.