

# Improved Sparse Bump Modeling for Electrophysiological Data

François-Benoit Vialatte<sup>1</sup>, Justin Dauwels<sup>1,2</sup>, Jordi Solé-Casals<sup>3</sup>, Monique Maurice<sup>1</sup>, and Andrzej Cichocki<sup>1</sup>

<sup>1</sup> Riken BSI, Wako-Shi, Japan,  
fvialatte@brain.riken.jp

<sup>2</sup> Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>3</sup> University of Vic, Vic, Spain

**Abstract.** Bump modeling is a method used to extract oscillatory bursts in electrophysiological signals, who are most likely to be representative of local synchronies. In this paper we present an improved sparse bump modeling method. The improvements are done in the adaptation method by optimizing the parameters according to the order of their derivatives; and in the window matching method by changing the selection of the initial function. Experimental results, comparing previous method *vs* the improved version, show that the obtained model fits better the signal, hence the result will be much more precise and useful.

## 1 Introduction

Brain rhythms are traditionally described by different oscillatory regimes including theta (4-10 Hz), beta (15-40 Hz) and gamma (40-90 Hz) rhythms. The structural organization and associated functional role of these electroencephalographic (EEG) oscillations are still far from being completely understood. Oscillatory activity can be separated in background and burst pattern activities. The background EEG is constituted by regular waves, whereas bursts are transient and with higher amplitudes. These bursts are organized activities, most likely to be representative of local synchronies. They should consequently play a specific functional role, distinct from background EEG activity. EEG event related synchronization (ERS), and similarly event related desynchronization (ERD), can be interpreted as the reorganization of the spontaneous brain oscillations in response to the stimulus [1, 2].

Once extracted, the activities can be used as features for classification (they can for instance feed neural network classifiers, see *e.g.* [7]) and signal analysis. However, extracting these activities with a reliable and consistent method is not a simple task. Therefore sparse bump modeling, a 2D extension of the 1D bump modeling described in [10], was developed. It was first applied to invasive EEG (local field potentials), recorded from rats olfactory bulb during a go-no go olfactory memory task [16, 15]. Afterwards, it was used to investigate several aspects of brain oscillatory dynamics: scalp EEG data from patients with early

stage of Alzheimer’s disease (AD) [13, 12, 15, 3]; simultaneous time-frequency-space representation using a sonification approach [11]; a synchrony model of time-frequency bursts events (stochastic event synchrony, [4-6]; modeling oscillations of steady state visual event potential epochs [14].

Bump modeling adapts parametric functions, within selected windows, to the wavelet time-frequency representation of a given signal. The present paper seeks to improve the matching algorithm of bump modeling, by the introduction of improved methods of adaptation and window selection.

## 2 Methods

### 2.1 Bump Modeling procedure

Bump modeling follows the following four steps (rationales for this procedure, proofs and technical details are explained in [15]):

1. the signal is first wavelet transformed into time-frequency maps.
2. the time-frequency map is z-score normalized.
3. the map is described by a set of time-frequency window (which could be seen as a low complexity ‘matching’ step if compared against the matching-pursuit method [8])
4. parametric functions are adapted within these windows, in decreasing ordered of energy.

### 2.2 Wavelets

We scale complex Morlet wavelet  $\vartheta$  to compute time-frequency wavelet representations of the signal  $\mathbf{X}$  of dimension  $T$ :

$$C_x(t, s) = \int \mathbf{X}(\tau) \vartheta^* \left( \frac{\tau - t}{s} \right) d\tau \quad (1)$$

where  $s$ , the scaling factor, controls the central frequency  $f$  of the mother wavelet. The modulus of this time-scale representation can therefore be used as a positive time-frequency spectrogram, which we will note  $\mathbf{C}_x$ , a time-frequency matrix of dimension  $T \times F$ .

### 2.3 Z-score

The time-frequency spectrogram is normalized depending on a reference signal. The reference signal  $\mathbf{R}$  of dimension  $U$  is wavelet transformed into a spectrogram  $\mathbf{C}_r$  of dimension  $U \times F$ . The average amplitudes  $\mathbf{M}_f = [\mu_1, \mu_2, \dots, \mu_F]$  and standard deviations of amplitude  $\mathbf{S}_f = [\sigma_1, \sigma_2, \dots, \sigma_F]$  are computed from  $\mathbf{C}_r$  for each of the  $F$  frequencies. The z-scored map  $\mathbf{Z}_x$  is obtained through normalization using these values:

$$\mathbf{Z}_x(f, t) = \frac{\mathbf{C}_x(f, t) - \mu_f}{\sigma_f} \quad (2)$$

## 2.4 z-score offset

We will refer to the positive z-score values as ERS-like components: if the signal is recorded during a stimulation, these oscillatory peaks are the most likely constituent of ERS. If the signal is recorded in pre-stimulus period or in rest condition, these oscillations are representative of organized oscillatory bursts. Z-score returns values in  $\mathbb{R}$ , but bump modeling only accepts values in  $\mathbb{R}^+$  as inputs. In order to model ERS-like oscillations, we reject the negative components of the map with a threshold, the z-score offset  $\phi$ . The thresholded map  $\mathbf{Z}_x^\phi$  is obtained with:

$$\mathbf{Z}_x^\phi = \frac{1}{2} [(\mathbf{Z}_x + \phi) + |(\mathbf{Z}_x + \phi)|] \quad (3)$$

Usual values of  $\phi$  are in the [0-3] range.

We will refer to the negative z-score values as ERD-like components: if the signal is recorded during a stimulation, these oscillatory peaks are the most likely constituent of ERD. If the signal is recorded in pre-stimulus period or in rest condition, these oscillations are representative of a local disorganization of oscillatory activity. These negative z-score values in  $\mathbb{R}^-$  can be extracted into a threshold map  $\mathbf{Z}_x^-$ :

$$\mathbf{Z}_x^- = -\frac{1}{2} [ |(\mathbf{Z}_x)| - (\mathbf{Z}_x) ] \quad (4)$$

## 2.5 Windowing

The map  $\mathbf{Z}_x$  is described by a set of windows  $\omega(s, \tau)$  with  $s \in \mathbf{S}$ ,  $\tau \in \mathbf{T}$  the position on the general time-frequency map respectively in scale and step; each  $\omega$  has the dimensions  $H \times W$  (height and width) determined depending on the time-frequency resolution at the window's central frequency (see [15]).

## 2.6 Parametric functions

We use half ellipsoid functions to model the normalized time-frequency map. The half ellipsoid outskirts are defined as:

$$\Psi(A, h, w, f, t, y, x) = 1 - \frac{(y - f)^2}{h^2} + \frac{(x - t)^2}{w^2} \quad (5)$$

The half ellipsoid bump is obtained with:

$$\xi(A, h, w, f, t, y, x) = \begin{cases} 0 & \text{if } \Psi < \lambda \\ A \cdot \sqrt{\Psi} & \text{else} \end{cases} \quad (6)$$

Hence the adaptation error to be minimized will be:

$$E(A, h, w, f, t, y, x) = \sum_{x=1}^W \sum_{y=1}^H \|\omega_{y,x}(s, \tau) - \xi(A, h, w, f, t, y, x)\|^2 \quad (7)$$

## 2.7 Improved Adaptation

derivatives of half ellipsoid bumps are always null when  $\Psi(A, h, w, f, t, y, x) < \lambda$ . For  $\Psi(A, h, w, f, t, y, x) \geq \lambda$ , they take the following values:

$$\frac{dE}{dA} = 2\sqrt{\Psi} \left( A\sqrt{\Psi} - \omega_{y,x}(s, \tau) \right) = -2\sqrt{E}\sqrt{\Psi} \quad (8)$$

$$\frac{dE}{dh} = \frac{2(y, f)^2 A\sqrt{\Psi}}{h^3} \left( A\sqrt{\Psi} - \omega_{y,x}(s, \tau) \right) = -2\sqrt{E} \frac{(y-f)^2 A}{h^3 \sqrt{\Psi}} \quad (9)$$

$$\frac{dE}{dw} = \frac{2(x-t)^2 A\sqrt{\Psi}}{w^3} \left( A\sqrt{\Psi} - \omega_{y,x}(s, \tau) \right) = -2\sqrt{E} \frac{(x-t)^2 A}{w^3 \sqrt{\Psi}} \quad (10)$$

$$\frac{dE}{df} = \frac{2(y-f)A\sqrt{\Psi}}{h^2} \left( A\sqrt{\Psi} - \omega_{y,x}(s, \tau) \right) = -2\sqrt{E} \frac{(y-f)A}{h^2 \sqrt{\Psi}} \quad (11)$$

$$\frac{dE}{dt} = \frac{2(x-t)A\sqrt{\Psi}}{l^2} \left( A\sqrt{\Psi} - \omega_{y,x}(s, \tau) \right) = -2\sqrt{E} \frac{(x-t)A}{l^2 \sqrt{\Psi}} \quad (12)$$

**Improved Adaptation** can be obtained by optimizing the parameters stepwise, with a priority depending on the order of their derivatives. The comparison of the above derivatives is obvious: the term  $-2\sqrt{E}\sqrt{\Psi}$  being always common to all these derivatives, and the other terms being defined in  $\mathbb{R}_+$ , they will all have the same sign. The slope of the adaptation will then be dependant on the multiplicands  $m$  applied to  $-2E\sqrt{\Psi}$ :

1.  $m = 1$  for  $dE/dA$ ;
2.  $m$  is a positive value in  $[0 - A]$  in numerator divided by a variable of order 3 for  $dE/dh$  and  $dE/dw$ ;
3.  $m$  is a positive value in  $[0 - A]$  in numerator divided by a variable of order 2 for  $dE/df$  and  $dE/dt$ .

This would probably be working correctly with properly normalized parameters, however we are here adapting parameters of different ranges:  $A \in [0 - 1]$ , while  $h$  and  $f \in [1 - H/2]$  and  $w$  and  $t \in [1 - W/2]$  with  $H$  and  $W$  usually  $\gg 1$ . Therefore, the multiplicands corresponding to the three above case will be of the order  $m \in O(1)$  in case (1), and  $m \in O(x^{-3})$  in case (2) and  $m \in O(x^{-2})$  in case (3). Practically speaking, it means that the parameters adaptation should be performed stepwise (first  $h$  and  $w$ , then  $f$  and  $t$ , and finally  $A$ ). In the previous implementation of bump modeling [15], we optimized all these parameters simultaneously (first using iterations of first order gradient descent, followed by iterations of the BFGS [9] algorithm). We improve the quality and speed of convergence by performing the following stepwise estimation of these parameters:

1. Update  $h$  and  $w$  until both their derivatives are below a threshold  $t_\Psi$ .
2. Update  $f$  and  $t$  until both their derivative are below a threshold  $t_{pos}$ . If at anytime  $dE/dh$  or  $dE/dw$  becomes above  $t_\Psi$ , go back to 1.

3. Update only  $A$ , until its derivative is below a threshold  $t_A$ . If at anytime  $dE/dh$  or  $dE/dw$  becomes above  $t_\psi$ , go back to 1. If at anytime  $dE/df$  or  $dE/dt$  becomes above  $t_{pos}$ , go back to 2.

The adaptation is still performed using the BFGS [9] algorithm. We used  $t_\psi = 10^{-5}$ .

## 2.8 Improved window matching

The first step of bump modeling consist in the selection of a window within the boundaries of which the function is estimated. In the previous version of bump modeling [15], the best candidate window  $\Omega$  was selected as:

$$\Omega = \arg \max_{s,\tau} \left[ \sum_y \sum_x \omega_{y,x}(s,\tau) \right] \quad (13)$$

Because these windows are used to determine the initial condition of the function adaptation, finding the best suitable window is primordial. The new optimized method is more related to matching pursuit [8], in that we will match the window content with the initial bump function  $\xi_{s,\tau}(y,x) = \xi(A_{s,\tau}, h_{s,\tau}, l_{s,\tau}, f_{s,\tau}, t_{s,\tau}, y, x)$  with  $f_{s,\tau} = h_{s,\tau} = H/2$ ,  $x_{s,\tau} = l_{s,\tau} = L/2$ , and  $A_{s,\tau}$  the highest peak in the window:

$$A_{s,\tau} = \max_{x,y} [\omega_{y,x}(s,\tau)] \quad (14)$$

We compute the matrix  $\Xi_{s,\tau}$  of values of this initial function for each time-frequency points of the window  $\omega(s,\tau)$ . The best window is thus matched as:

$$\Omega = \arg \max_{s,\tau} \left[ \frac{\Xi_{s,\tau} : \omega_{s,\tau}}{\|\Xi_{s,\tau}\|_F} \right] \quad (15)$$

Where  $:$  denotes the Frobenius inner product, and  $\|\cdot\|_F$  indicates the Frobenius norm. This generalize the matching criterion of matching pursuit methods to 2D data. For bump modeling high amplitude has priority on best fit. Therefore, contrary to matching pursuit, the product is not normalized by the norm of  $\omega_{s,\tau}$ .

## 2.9 EEG Data

We demonstrate the improved method with data collected from 23 patients with Alzheimer's disease. This is the same data that was modeled with bumps in [12]. EEG recording was done in an awake resting state with eyes closed, under vigilance control. Ag/AgCl electrodes (disks of diameter 8 mm) were placed on 21 sites according to 10/20 international system, with the reference electrode on the right ear-lobe. EEG was recorded with Biotop 6R12 (NEC Sanei, Tokyo, Japan) using sampling frequency of 200Hz.

### 3 Results and discussion

The purpose of the present paper was to describe an improved method for sparse time-frequency bump modeling, which addresses efficiently the problem of detecting automatically reproducible timefrequency burst events in electro-physiological data. We extracted ERS-like and ERD-like components from the EEG signals using the improved modeling (improved adaptation and/or window matching), as compared with the previous method. The improved bump modeling allowed a more reliable extraction of time-frequency oscillatory bursts events as compared to the method described in [15]. We observed that:

1. The original algorithm already fits nicely the parametric functions to the map.
2. The improved window matching mainly changes the adaptation order, but sometimes selects windows with better location than the previous method. The difference with the previous method (b *vs.* c) is subtle but significant.
3. The improved adaptation fits better the parametric function's shapes to the signal (b *vs.* d).
4. With both improved methods, the representation becomes even more precise (b *vs.* e).

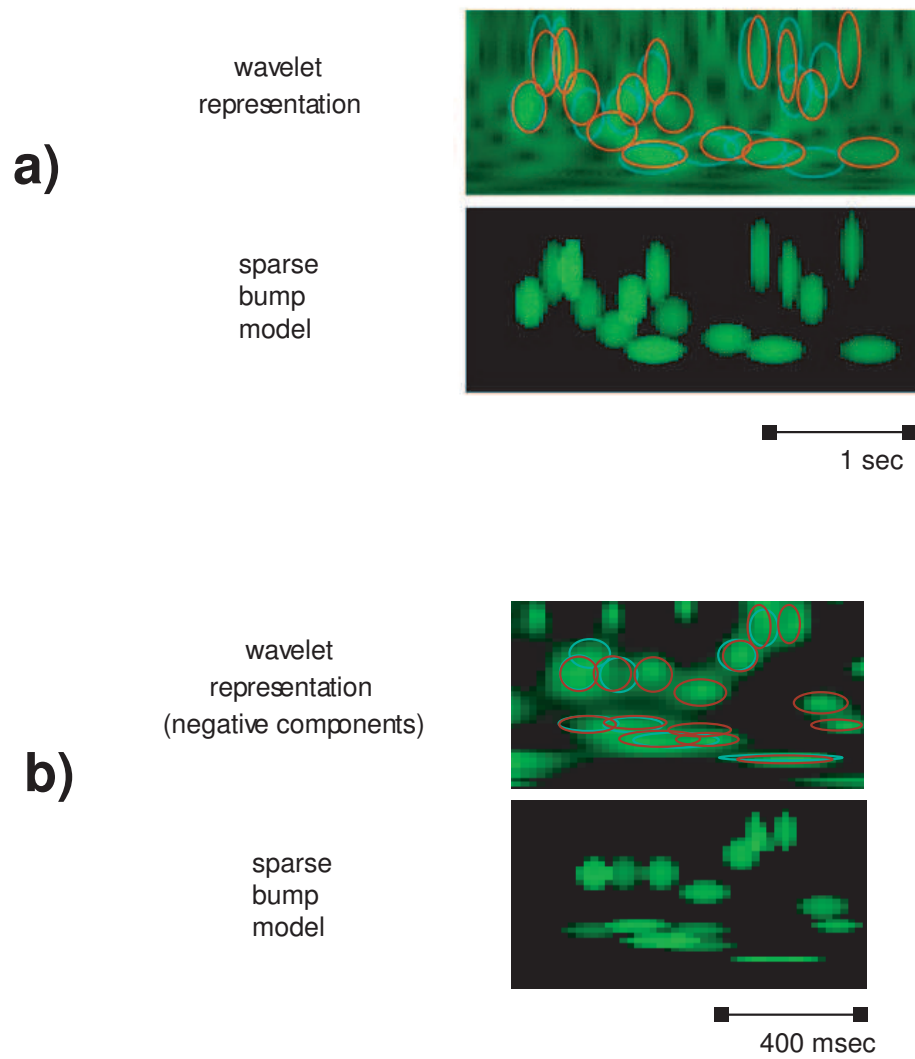
In order to illustrate this difference, we show on Fig.1 an example of time-frequency modeling with a comparison of the previous and new model (ERS-like components, with  $\phi = 3$ , and ERD-like components).

### 4 Acknowledgments

Jordi Solé-Casals was supported by the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya, and by the Spanish grant TEC2007-61535/TCM. Monique Maurice is supported by the JSPS fellowship P-08811. Justin Dauwels is supported by the Henri-Benedictus fellowship from the King Baudouin Foundation and the BAEF.

### References

1. E. Başar. *EEG-brain dynamics: Relation between EEG and brain evoked potentials*. Elsevier, Amsterdam, 1980.
2. E. Başar, T. Demilrap, M. Schürmann, C. Başar-Eroglu, and A. Ademoglu. Oscillatory brain dynamics, wavelet analysis, and cognition. *Brain and Language*, 66:146–183, 1999.
3. M Buscema, P Rossini, C Babiloni, and E Grossi. The ifast model, a novel parallel nonlinear eeg analysis technique, distinguishes mild cognitive impairment and alzheimer's disease patients with high degree of accuracy. *Artificial Intelligence In Medicine*, 40:127–141, 2007.
4. J Dauwels, F Vialatte, and A Cichocki. A novel measure for synchrony and its application to neural signals. In *Proceedings of the 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'07): 15-20 April 2007, Honolulu, USA*, volume IV, pages 1165–1168, 2007.



**Fig. 1.** Improved bump modeling, as compared to classical bump modeling. a) ERS-like components ( $\phi = 3$ ), b) ERD-like components. Top: Wavelet transform, with superimposed positions of the previous model (blue) and the improved model (red). Bottom: improved bump model.

5. J Dauwels, F Vialatte, and A Cichocki. On synchrony measures for the detection of alzheimer's disease based on eeg. In *LNCS 4984, Proceedings of the 14th International Conference on Neural Information Processing (ICONIP'07): 13-16 November 2007; Kita Kyushu, Japan*, pages 112–125, 2008.
6. J Dauwels, F Vialatte, T M Rutkowski, and A Cichocki. Measuring neural synchrony by message passing. In *Advances in Neural Information Processing Systems, NIPS 2007: 6-9 December 2007; Vancouver, Canada*, 2008.
7. N. Kasabov. *EEvolving connectionist systems: The knowledge engineering approach*. Springer, 2007.
8. SG Mallat and Zhang Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 12:3397–3415, 1993.
9. W H Press, B P Flannery, S A Teukolsky, and Vetterling W T. *Numerical recipes in C: The art of scientific computing*. Cambridge University Press, New York, 2002.
10. Dubois R, P Maison-Blanche, B Quenet, and G Dreyfus. Automatic eeg wave extraction in long-term recordings using gaussian mesa function models and non-linear probability estimators. *Comput Methods Programs Biomed*, 88(3):217–233, 2007.
11. F Vialatte and A Cichocki. Sparse bump sonification: a new tool for multichannel eeg diagnosis of mental disorders; application to the detection of the early stage of alzheimer's disease. In *LNCS 4234, Proceedings of the 13th International Conference on Neural Information Processing (ICONIP'06): 3-6 October 2006, Hong Kong, China*, pages 92–101, 2006.
12. F Vialatte, A Cichocki, G Dreyfus, T Musha, T Rutkowski, and R Gervais. Blind source separation and sparse bump modelling of time frequency representation of eeg signals: New tools for early detection of alzheimer's disease. In *Proceedings of the IEEE Workshop on Machine Learning for Signal Processing 2005 (MLSP'05): 28-30 September 2005, Mystic CT, USA*.
13. F Vialatte, A Cichocki, G Dreyfus, T Musha, S L Shishkin, and R Gervais. Early detection of alzheimer's disease by blind source separation, time frequency representation, and bump modeling of eeg signals (invited presentation). In *LNCS 3696, Proceedings of the International Conference on Artificial Neural Networks 2005 (ICANN'05): 11-15 September 2005, Warsaw, Poland*, pages 683–692, 2005.
14. F Vialatte, J Dauwels, T M Rutkowski, and A Cichocki. Oscillatory event synchrony during steady state visual evoked potentials. In Springer, editor, *Advances in Cognitive Neurodynamics, Proceedings of the First International Conference on Cognitive Neurodynamics (ICCN'07): 17-21 November 2007; Shanghai, China*, 2008.
15. F Vialatte, C Martin, R Dubois, J Haddad, B Quenet, R Gervais, and Dreyfus G. A machine learning approach to the analysis of time-frequency maps, and its application to neural dynamics. *Neural Networks*, 20:194–209, 2007.
16. F Vialatte, C Martin, N Ravel, B Quenet, G Dreyfus, and R Gervais. Oscillatory activity, behaviour and memory, new approaches for lfp signal analysis. In *Acta Neurobiologiae Experimentalis 2003, Vol. 63., Proceedings of the 35th annual general meeting of the European Brain and Behaviour Neuroscience Society (EBBS'03): 17-20 September 2003, Barcelona, Spain*, 2003.