33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

A Bayesian Framework for Analyzing iEEG Data from a Rat Model of Epilepsy

Sabato Santaniello, *Member, IEEE*, David L. Sherman, *Member, IEEE,* Marek A. Mirski, Nitish V. Thakor, *Fellow, IEEE*, and Sridevi V. Sarma, *Member, IEEE*

Abstract— **The early detection of epileptic seizures requires computing relevant statistics from multivariate data and defining a robust decision strategy as a function of these statistics that accurately detects the transition from the normal to the peri-ictal (problematic) state. We model the afflicted brain as a hidden Markov model (HMM) with two hidden clinical states (normal and peri-ictal). The output of the HMM is a statistic computed from multivariate neural measurements. A Bayesian framework is developed to analyze the *a posteriori* conditional probability of being in peri-ictal state given current and past output measurements. We apply this method to multi-channel intracortical EEGs (iEEGs) from the thalamo-cortical ictal pathway in an epilepsy rat model. We first define the output statistic as the max singular value of a connectivity matrix computed on the EEG channels with spectral techniques Then, we estimate the HMM transition probabilities from this statistic and track the *a posteriori* probability of being in peri-ictal state (the "information state variable"). We show how the information state variable changes as a function of time and we predict a seizure when this variable becomes greater than 0.5. This Bayesian strategy significantly improves over chance level and heuristically-chosen threshold-based predictors.**

I. Introduction

EPILEPTIC seizures in patients can be preceded by early changes in the temporal properties of intracortical EEG (iEEG) signals [1]–[10]. Univariate and bivariate approaches have provided some evidence for such changes. Iasemidis et al. [1][2] showed that the short-term largest Lyapunov exponents of the iEEG recorded in a critical electrode site may significantly decrease ~70 min before the seizure onset. Lehnertz and Elger [3][4] reported that a measure of the EEG complexity can decrease ~12 min before a seizure, Le Van Quyen at al. [5] showed that a measure of the similarity between non overlapping windows of the same EEG signal may significantly modulate ~5 min before a seizure, and Jouny et al. [6] showed that a measure of EEG complexity may increase several seconds before the clinical onset of the seizure. Finally, [7] reported that the accumulated energy (time integrated variance of the power spectrum of the EEG) locally increases in specific electrode sites 50 min before the seizure because of bursts of epileptiform discharges.

Bivariate measures [8]–[10], instead, estimate the phase synchronization between pairs of EEG channels and define phase variables based either on the Hilbert Transform or the Wavelet Transform [11]. Several bivariate measures have been proposed (max linear cross-correlation [8], conditional probability index [9][12], phase difference [10], Shannon entropy index [12], etc.) and it has been shown that such measures may decrease ~80 min before the seizure onset [8].

Based on the variability of these measures, several seizure prediction algorithms were proposed [1][2][9][13]. These algorithms track the measures over time and predict a seizure when such measures pass a heuristically-chosen fixed threshold. However, studies conducted on extensive databases of clinical seizures show that these measures with fixed threshold policies have poor predictive performances and are typically no better than a chance level predictor [9][13]–[16]. Possible explanations can be: (i) univariate and bivariate measures provide a limited description of the peri-ictal activity as they do not capture network effects among multiple sites; (ii) transitions from normal (seizure-free) to peri-ictal state could impact 2nd or higher order statistics, which means that such transitions cannot be captured with fixed threshold-based policies on these measures; (iii) thresholds are heuristic and do not explicitly optimize any performance-related objective function (e.g., minimizing the prediction delay or the probability of false alarms, etc.).

We propose a probabilistic framework for the analysis of multivariate statistics computed from multichannel iEEGs in different clinical states. We characterize the multivariate statistic as the output of a hidden Markov model (HMM) [17] with hidden normal and peri-ictal states. Then, we exploit a Bayesian approach to analyze on line the *a posteriori* probability of being in peri-ictal state given the current and past output measurements [18], and use this conditional probability for predicting a state transition. Our framework computes a power spectrum-based connectivity matrix among all the available iEEG channels and uses the max singular value (σ_{max}) of this matrix as the required time-varying multivariate statistic. The distribution of the values of σ_{max} conditioned on being in normal vs. peri-ictal state is estimated from the data via the Baum-Welch algorithm [19].

We apply our framework to multichannel iEEG signals simultaneously acquired in anterior and posterior thalamus, hippocampus, and cortex in a rodent model of generalized epilepsy [20]. Two male Sprague–Dawley rats were treated with pentylenetetrazol (PTZ) chemoconvulsant to generate short-term (i.e., approximately 10 min after the injection) seizures with selective activation of the anterior thalamus.

S. Santaniello, D. L. Sherman, N. V. Thakor, and S. V. Sarma are with the Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: {ssantan5,sherman,nthakor,ssarma2}@jhu.edu). M. A. Mirski is with the Department of Anesthesiology & Critical Care Medicine, Johns Hopkins University (e-mail: mmirski@jhmi.edu).

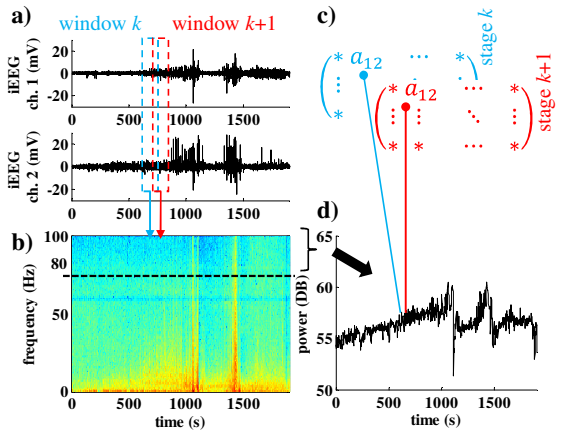

Fig. 1. Multivariate analysis. Consecutive 3 s-long iEEG windows with 0.5 s overlap (a) are used for computing the time-varying cross power spectrogram (b) between any pair of channels. For each window, the cross power in the band [80, 100] Hz is computed (d) and used for filling the elements of the connectivity matrix (c) at the correspondent stage.

Our data set included 12 clinical seizures from 7 recording sessions (session duration: 33.0 ± 5.3 min, mean ± standard error of mean [s.e.m.]). Because of the PTZ, the transition from preictal normal state to peri-ictal state in the iEEGs occurred just ~5 min before the actual onset of every clinical seizure and a few minutes after the injection of PTZ.

With this data, our framework detected the state transition with average lag of 86.3 ± 19.5 s (mean ± s.e.m.), which is significantly lower (p-value $p < 0.05$) than the lag achieved by a chance level and a threshold-based policy, where the threshold is chosen heuristically and applies to σ_{max}.

II. Methods

A. Multivariate Analysis

Multichannel iEEGs sampled at 200 Hz are used. For each pair of channels in each recording session, the cross power spectrogram is computed over consecutive 3 s windows (0.5 s overlap, Fig. 1a). For each window, the power density is computed with the Welch's method [21] (Fig. 1b).

The connectivity matrix is defined as a time-varying matrix $A(k)$, whose (i,j)-th element at stage k is the power stored in the band [80,100] Hz of the cross power spectrum between the i-th and j-th channel in the k-th window (Fig. 1c-d). The max singular value of $A(k)$, $\sigma_{max}(k)$, is computed at each stage k and tracked over time.

B. Hidden Markov Model

We model the afflicted brain as a HMM with two states (Fig. 2). At stage k, the state $x_k \in \{0,1\}$ (0 = normal; 1= peri-ictal). We assume that $x_0 = 0$ and that x_k switches from 0 to 1 at some stage $\overline{T} > 0$, with $\overline{T}$ geometric random variable, and that from $\overline{T}$ onwards $x_k = 1$. The probability of the event $\{\overline{T} = k\}$ is $P(\overline{T} = k) = \rho(1-\rho)^{k-1}$ for $k = 1,2,3,\ldots$, where ρ is the parameter of the geometric distribution and represents the probability of transition from state 0 to state 1 [17].

Differently from a traditional Markov chain, the states of an HMM are inaccessible or "hidden". However, output observations, z_k, are available and depend probabilistically on the states. One can think of z_k as a "noisy" observation of x_k. We assume that, for any k, $z_k = \sigma_{max}(k)$. For each recorded session of multichannel iEEGs, we estimated off-line the probability mass function $q_x(z) = P(z_k=z \mid x_k=x)$ for $x \in \{0,1\}$ and any value $z > 0$ by running the Baum-Welch algorithm [19] on training data (~50% of the available observations).

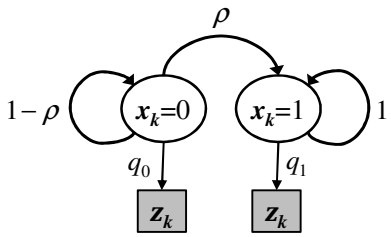

Fig.2. HMM schematic with $z_k = \sigma_{max}(k)$.

C. Bayesian Evolution Model and Estimation Policy

Because the state x_k is inaccessible, we define the new "information state variable" $\pi_k = P(\overline{T} \le k \mid z_0,\ldots, z_k) = P(x_k = 1 \mid z_0,\ldots, z_k)$ which is the Bayesian *a posteriori* probability of being in state 1 at stage k given the observations up to and including stage k. The evolution of π_k is given by [18][22]:

$$\pi_0 = 0$$

$$\pi_{k+1} = \frac{L(z_{k+1})\left[\pi_k + (1-\pi_k)\rho\right]}{(1-\pi_k)(1-\rho) + L(z_{k+1})\left[\pi_k + (1-\pi_k)\rho\right]} = f(\pi_k, z_{k+1}) \quad (1)$$

where we used $L(z_{k+1}) = q_1(z_{k+1})/q_0(z_{k+1})$ [18][22] and ρ is the parameter of the geometric distribution in section II-B.

By using the Bayesian framework, an estimation T_S of the time of state transition is given by:

$$T_S = \min\{\, k > 0 \mid \pi_k > 0.5 \,\} \quad (2)$$

i.e., we decide that a change has occurred when the *a posteriori* probability of a change exceeds 50% [18].

D. Evaluation of the Detection Policy

Since each recorded session in our data set includes at least one seizure (see Section II-E), we evaluated the detection performances of the policy (2) by measuring the absolute distance $|T_S - \overline{T}|$ between the estimated and actual change time for each seizure event. For each event, we marked the peri-ictal interval by running the Viterbi's algorithm [19] on the sequence of observations z_k, k = 1,2,3,…, (Fig. 3a-b) with the hidden states and the probability mass functions $q_x(z)$, $x \in \{0,1\}$ as defined in section II-B (Fig. 3c). The actual change time $\overline{T}$ was set at the beginning of such marked interval and maximizes the distance between the probability distributions of the z_k inside vs. outside the interval [19].

We compared the policy (2) with the chance level (CL) predictor $T_S^{CL} = E[\overline{T}] = 1/\rho$, where ρ is defined in section II-B and $E[\cdot]$ is the expected value.

We also compared (2) with the heuristic threshold-based (HT) policy:

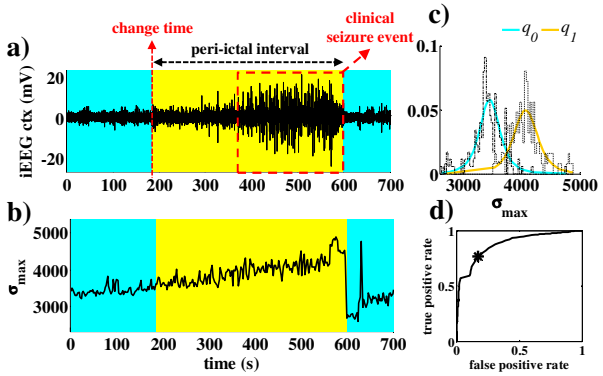

Fig. 3. Change times. a-b) The peri-ictal interval (yellow) of the seizure (a) is extracted by running the Viterbi's algorithm on the observations $z_k = \sigma_{max}(k)$, k=1,2,3,…. in (b). c) Probability functions q_0, q_1 of z_k in state 0 (preictal normal) and 1 (peri-ictal), respectively, for the same seizure. Histograms of the actual σ_{max} are overlapped. d) Mean ROC curve. The asterisk denotes the point of the curve that corresponds to the max distance between true and false positive rates.

$$T_S^{HT} = \min\{ k > 0 \mid \sigma_{max}(k) > H^* \}. \quad (3)$$

H^* was the same for every recording session and was set as follows: we first computed, for each seizure, the receiver operator characteristic (ROC) curve [2] of the observations $\sigma_{max}(k)$, i.e., we linearly varied the threshold H to span the set of values of $\sigma_{max}(k)$, and, for each H, we computed the true and false positive rates as the fraction of samples in the peri-ictal and preictal normal interval that are larger than H, respectively [2]. Then, for each H, we averaged the true and positive rates over the available seizures and constructed the mean ROC curve (Fig. 3d). Note that each point of the mean ROC curve corresponds to a specific value of H. Finally, we chose H^* as the threshold at which the max distance between true and false positive rates of the mean ROC curve is achieved. This choice of H^* aimed at keeping both the probability of the event $\{T_S < \overline{T}\}$ and the average delay low.

E. Experimental Set Up

Details about the experimental set up are in [20].The Johns Hopkins Medical Institutional Review Board Committee for laboratory investigation approved the research protocol.

Male Sprague-Dawley rats (250-300 g) were implanted with four skull screw EEG electrodes placed bifrontally and posteriorly behind bregma. Bipolar insulated steel electrodes (0.125-mm diameter, 2-mm tip separation) were placed on cortex (CTX), in anterior (2 electrodes, one per side) and posterior thalamus. A fifth depth electrode was placed in hippocampus (Fig 4). Animals were allowed to recover for 2 days with *ad lib* food and water. Then, they were implanted with a jugular venous catheter and recovered for a minimum of 1 h. In each animal, baseline EEG was recorded for at least 60 s prior to the infusion of PTZ (100 mg ml^{-1}, Sigma Chemical, St. Louis, MO), administered at 5.5 mg $kg^{-1}min^{-1}$. Behavior and EEGs were both continuously monitored and the extent of seizure was noted according to the modified clinical Racine scale [20]. Animals passed through all stages of the Racine scale. iEEGs from the implanted electrodes were acquired in 7 nonconsecutive sessions (min and max

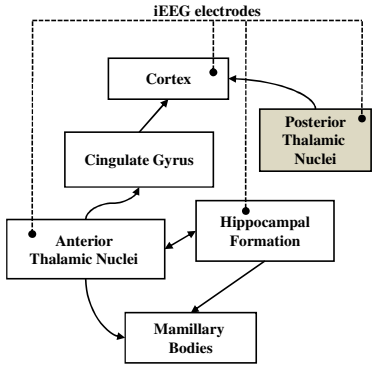

Fig. 4. Simplified schematic of the connections from mammillary bodies (MB) to the anterior thalamus (AN) via the mammillothalamic tract. AN, MB, and hippocampus belong to the "circuit of Papez". AN has primary rostral connections to cingulate gyrus and, ultimately, to cortex. The posterior thalamus is assumed to be unaffiliated with activity during the epileptic seizure and used as a reference site.

session duration: 12 and 58 min, respectively; average: 33.0 ± 5.3 min, mean ± s.e.m.) and 12 seizure events were noted.

Analog EEGs were amplified with a Grass 8D-10 eight-channel portable polygraph with internal 0.3 Hz high-pass and 70 Hz low-pass filter cutoffs. A 60 Hz notch was also employed. Analog EEGs were collected using a 7-channel FM data recorder (TEAC MR-30) and digitized by CODAS (DATAQ Instruments Inc., Akron, OH) with sampling rate of 1000 Hz. Then, data was downsampled offline at 200 Hz.

III. Results

Parameter ρ in (1) was estimated by fitting (maximum likelihood estimation) a geometric distribution on the actual change times $\overline{T}$. Results are in Fig. 3, 5-6.

The two-state HMM fitted on the sequential observations $\sigma_{max}(k)$ clearly isolates the problematic state (yellow portion in Fig. 3b) well before the onset of the clinical seizure, while the raw iEEGs do not show significant modulation before the seizure onset (Fig.3a). The connectivity matrix combines information from all the available channels, captures the network interactions among the electrode sites, and therefore modulates during the entire transition from normal preictal to peri-ictal state (Fig. 3b). Such modulation is reflected in the probability function q_x of σ_{max}, which is fitted on actual observations and is different in state x=0 vs. x=1 (Fig. 3c).

The different probability distribution of σ_{max} in state 0 vs. 1 also influences the evolution of the variable π_k in (1). As in Fig. 5, π_k is generally low in state 0 and begins to increase at the transition to state 1. The dynamics of π_k depends on the ratio $L(\cdot)$ between the functions q_1 and q_0 in (1), and is generally fast (a few seconds are required to reach the steady state value π_k = 1) provided that the probability distribution of σ_{max} is significantly different in state 0 vs. 1.

Based on the dynamics of π_k, we proposed a Bayesian estimator (BE) in (2) to predict the transition from interictal to peri-ictal state. Cumulative results in Fig. 6 indicate that the BE performed better than the CL and HT predictors. The average absolute distance $|T_S - \overline{T}|$ (Fig. 6a), delay (Fig. 6b) and anticipation (Fig. 6c) with BE were lower than with CL

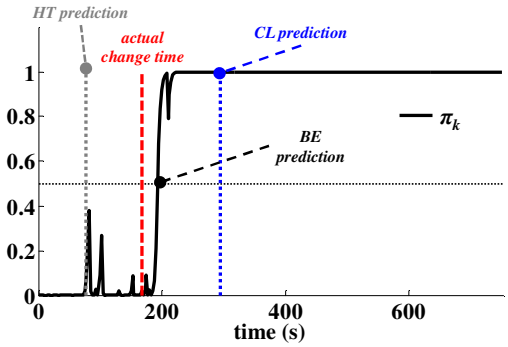

Fig. 5. Evolution of the information state variable π_k for a seizure event. Actual change time and estimation given by BE, CL, and HT predictor are marked.

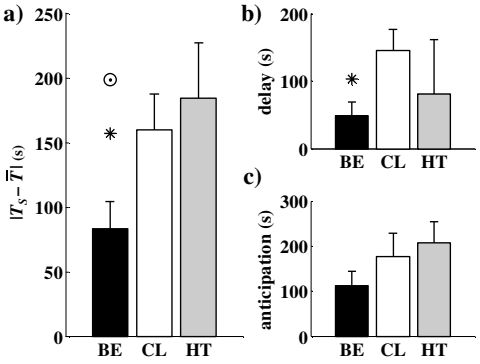

Fig. 6. Detection policies. a) Average absolute distance between the actual change time $\bar{T}$ and the estimation T_S achieved with the Bayesian estimator (2) (BE), the chance level predictor (CL), and the heuristic threshold-based policy (3) (HT). b), c) show results separately for delays (i.e., events $\{T_S > \bar{T}\}$ only) and anticipations (i.e., events $\{T_S < \bar{T}\}$ only), respectively. Bars are mean + s.e.m. Asterisks and circles indicate significant difference ($p < 0.05$) BE vs. CL and BE vs. HT, respectively.

and HT, and the difference was statistically significant (t-test, $p<0.05$) for absolute distances (BE vs. CL and BE vs. HT) and delays (BE vs. CL only). The difference between delays with BE and HT, instead, was not significant because $T_S^{HT} > \bar{T}$ only in 2 out of 12 seizures, which is due to the large variance of σ_{max} in state 0 (Fig. 3c). Although the t-test was not passed in case of $\{T_S < \bar{T}\}$, the average anticipation was remarkably less with BE than with CL (66s) or HT (96s) (Fig. 6c). Also, the absolute distance $|T_S - \bar{T}|$ with BE was lower than with CL and HT in 11 out of 12 and 10 out of 12 seizures, respectively, which means that the improvements achieved with the BE policy over the CL and HT predictor were independent from the specific seizure event.

ACKNOWLEDGMENT

Dr. Sarma was supported by the Burrough's Wellcome Fund CASI and L'Oreal USA FWIS. Dr. Sherman was supported by NIH NS035528.

REFERENCES

[1] L. D. Iasemidis, P. M. Pardalos, J. C. Sackellares, and D. S. Shiau, "Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures," *J. Comb. Optim.*, vol. 5, pp. 9-26, 2001.

[2] J. C. Sackellares, D. S. Shiau, J. C. Principe, M. C. K. Yang, L. K. Dance, W. Suharitdamrong, W. Chaovalitwongse, P. M. Pardalos, and L. D. Iasemidis, "Predictability analysis for an automated seizure prediction algorithm," *J. Clin. Neurophysiol.*, vol. 23, pp. 509-520, 2006.

[3] C. E. Elger, and K. Lehnertz, "Seizure prediction by non-linear time series analysis of brain electrical activity," *Eur. J. Neurosci.*, vol. 10, pp. 786-789, 1998.

[4] K. Lehnertz, and C. E. Elger, "Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity," *Phys. Rev. Lett.*, vol. 80, pp. 5019-5023, 1998.

[5] M. Le Van Quyen, C. Adam, J. Martinerie, M. Baulac, S. Clémenceau, and F. Varela, "Spatio-temporal characterizations of non-linear changes in intracranial activities prior to human temporal lobe seizures," *Eur. J. Neurosci.*, vol. 12, pp. 2124-2134, 2000.

[6] C. C. Jouny, G. K. Bergey, and P. J. Franaszczuk, "Partial seizures are associated with early increases in signal complexity," *Clin. Neurophysiol.*, vol. 121, pp. 7-13, 2010.

[7] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, R. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter, and G. Vachtsevanos, "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients," *Neuron*, vol. 30, pp. 51-64, 2001.

[8] F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, "Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients," *Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys.*, vol. 67, pp. 021912, 2003.

[9] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *Clin. Neurophysiol.*, vol. 116, pp. 569-587, 2005.

[10] B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, A. Schulze-Bonhage, and J. Timmer, "Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction," *Chaos*, vol. 16, pp. 013108, 2006.

[11] M. Le Van Quyen, J. Foucher, J. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F. J. Varela, "Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony," *J. Neurosci. Methods*, vol. 111, pp. 83-98, 2001.

[12] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. S. Pikovsky, J. Volkmann, A. Schnitzler, and H. J. Freund, "Detection of n:m phase locking from noisy data: application to magnetoencephalography," *Phys. Rev. Lett.*, vol. 81, pp. 3291-3294, 1998.

[13] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. Schulze-Bonhage, and J. Timmer, "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic," *Physica D*, vol. 194, pp. 357-368, 2004.

[14] P. E. McSharry, L. A. Smith, and l. Tarassenko, "Prediction of epileptic seizures: are nonlinear methods relevant?," *Nat. Med.*, vol. 9, pp. 241-242, 2003.

[15] Y. C. Lai, M. A. Harrison, M. G. Frei, and I. Osorio, "Controlled test for predictive power of Lyapunov exponents: their inability to predict epileptic seizures," *Chaos*, vol. 14, pp. 630-642, 2004.

[16] M. Winterhalder, T. Maiwald, H. U. Voss, R. Aschenbrenner-Scheibe, J. Timmer, and A. Schulze-Bonhage, "The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods," *Epilepsy Behav.*, vol. 4, pp. 318-325, 2003.

[17] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models. Estimation and Control*. New York, NY: Springer, 1995.

[18] M. Basseville, and I. V. Nikiforov, *Detection of Abrupt Changes. Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[19] R. Durbin, S. Eddy, A. Krogh and G. Mitchison, *Biological Sequence Analysis*. Cambridge, UK: Cambridge University Press, 1998.

[20] D. L. Sherman, C. B. Patel, N. Zhang, L. A. Rossell, Y. C. Tsai, N. V. Thakor, and M. A. Mirski, "Sinusoidal modeling of ictal activity along a thalamus-to-cortex seizure pathway I: new coherence approaches," *Ann. Biomed Eng.*, vol. 32, pp. 1252-1264, 2004.

[21] 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 Trans. Audio Electroacoust.*, vol. AU-15, pp. 70-73, 1967.

[22] H. V. Poor, "Quickest detection with exponential penalty for delay," *Ann. Statist.*, vol. 26, pp. 2179–2205, 1998.