Quickest detection of drug-resistant seizures: An optimal control approach

Sabato Santaniello [a], Samuel P. Burns [a], Alexandra J. Golby [b], Jedediah M. Singer [c], William S. Anderson [d], Sridevi V. Sarma [a,*]

[a] Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA
[b] Department of Neurosurgery and Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA
[c] Department of Ophthalmology and Neurology, Children's Hospital, Boston, MA, USA
[d] Department of Neurosurgery, Johns Hopkins Hospital, Baltimore, MD, USA

ARTICLE INFO

ABSTRACT

Epilepsy affects 50 million people worldwide, and seizures in 30% of the cases remain drug resistant. This has increased interest in responsive neurostimulation, which is most effective when administered during seizure onset. We propose a novel framework for seizure onset detection that involves (i) constructing statistics from multichannel intracranial EEG (iEEG) to distinguish nonictal versus ictal states; (ii) modeling the dynamics of these statistics in each state and the state transitions; you can remove this word if there is no room. (iii) developing an optimal control-based "quickest detection" (QD) strategy to estimate the transition times from nonictal to ictal states from sequential iEEG measurements. The QD strategy minimizes a cost function of detection delay and false positive probability. The solution is a threshold that non-monotonically decreases over time and avoids responding to rare events that normally trigger false positives. We applied QD to four drug resistant epileptic patients (168 hour continuous recordings, 26–44 electrodes, 33 seizures) and achieved 100% sensitivity with low false positive rates (0.16 false positive/hour).

This article is part of a Supplemental Special Issue entitled The Future of Automated Seizure Detection and Prediction.

1. Introduction

Automatic online seizure detection (AOSD) in intractable epilepsy has generated great interest in the last 20 years and is a fundamental step toward the development of neurostimulation-based responsive antiepileptic therapies [1–3]. Pioneering works in the late 1970s and early 1980s by Gotman et al. [4,5] showed that seizures can be automatically separated from interictal activity, and since then, several approaches to AOSD have been proposed by exploiting either scalp or intracranial EEG recordings, single or multichannel analysis, linear or nonlinear features.

Osorio et al. [6–9] used a wavelet-based decomposition of selected intracranial EEG recordings (iEEGs) to (i) separate the seizure-related component from the background noise, (ii) track the ratio between these components in the time–frequency domain, and (iii) detect a seizure when such a ratio crosses a fixed threshold for a sufficiently long time. Parameters of the detection method (e.g., threshold,

duration of the suprathreshold condition, etc.) can be either fixed [6] or adaptive [7,8]. Fixed threshold-based approaches were also proposed in [10–12], where the threshold was applied to linear spectral features of the iEEGs.

Gotman et al. [13–15] proposed a probabilistic framework for seizure detection using both scalp EEGs [15] and iEEGs [13,14]. In this framework, amplitude and energy measures in multiple frequency bands are computed for each channel via wavelet decomposition and the corresponding sampled probability distribution function is estimated. Then, the probability of a seizure is conditioned to the value of such measures and estimated via Bayes' rule. A patient-specific threshold is finally applied on this conditional probability of seizure to decide, for each channel, whether a seizure is likely or not, and a seizure is detected when that threshold is passed in a sufficient number of channels.

In all the studies mentioned above, the AOSD was solved by (i) introducing a relevant statistic that is computed from the available measurements and that captures changes in brain activity at the seizure onset, and (ii) a rule that, based on this statistic, determines whether a seizure has occurred or not.

More recently, AOSD has been implemented using sophisticated classification tools. In particular, EEG channels (either scalp or intracortical) have been processed individually to extract multiple univariate or bivariate features in time, frequency, or wavelet domain [16–32]. Then,

* Corresponding author at: Institute for Computational Medicine, Johns Hopkins University, Hackerman Hall 316c, Baltimore, MD 21218–2686, USA. Fax: +1 410 516 5294.
E-mail address: ssarma2@jhu.edu (S.V. Sarma).

for each channel, the available features have been combined into vectors and classified via support vector machines (SVM) [21,23,25,29], principal component analysis (PCA) [27,28], artificial neural networks (ANN) [16,18–20,24,30–32], fuzzy logics [22], or pattern recognition tools [17]. Finally, decisions made for different channels have been combined or ranked to ultimately determine whether a seizure has occurred or not. As a variation to this paradigm, [26,33] merged features extracted from different channels into one vector and applied the classification rule to this vector.

Although many of the aforementioned methods have sensitivity well above 90%, results generally show lower specificity (i.e., higher number of false positives) when applied to test data, and a comparative study by Lee et al. [34] reported performance varying from patient to patient. This may not be surprising considering that (i) several parameters in these methods are patient specific and tuned heuristically; and (ii) these methods do not explicitly minimize relevant detection performance criteria (e.g., delay or probability of false positives). The latter drawback reflects the important fact that current detection paradigms develop algorithms first which then define and limit the performances. We believe that performance specifications should be stated up front in a cost function to be minimized (e.g., delay length, false positive rate, etc.) which then defines the algorithm.

We propose a novel computational framework for AOSD that involves (a) constructing network-based statistics from multichannel neural data to distinguish nonictal from ictal states; (b) modeling the evolution of such statistics in nonictal and ictal states and their transition probabilities; and (c) developing an optimal model-based strategy that detects the transitions from nonictal to ictal states by using sequential multichannel measurements. The combination of (a) and (b) results in a *dynamic detector,* which, unlike a standard classifier, evolves over time based on current and past measurements, thus automatically adapting to brain dynamics.

In our formulation, AOSD is posed as a change-point detection problem and solved by minimizing a cost function of the average distance between actual and estimated seizure onset, the probability of false positives and the probability of delayed detection. This formulation is a variation of the "Quickest Detection" (QD) [35,36], whose theory we extended to allow for different cost functions and, more importantly, for dependent measurements, which is more applicable to neural data [37].

We applied our framework to four subjects with drug-resistant seizures (168 hours [h] of continuous recordings from 26–44 iEEG electrodes, 33 seizures), which resulted in 100% sensitivity on both training and validation data, low false-positive rate (min: 0.16 false positive/hour [FP/h]; max: 2.95 FP/h; mean: 1.39 FP/h), and an average detection delay of 9.6 seconds. Compared with a standard Bayesian and a heuristic threshold detector (i.e., non-optimal policies for which no explicit cost function is minimized), our approach showed lower false-positive rates and higher sensitivity.

2. Methods

2.1. Clinical data

Four subjects with intractable epilepsy were surgically implanted with subdural grid and strip electrodes (26–75 channels, Ad-Tech® Medical Instrument Corp., Racine, WI, USA) for approximately 1 week before surgical resection of the focal region and monitored by clinicians for seizures and interictal epileptic activity. Electrodes are 4 mm diameter platinum contacts embedded in a silicone sheet with 2.3 mm exposed. Data were digitized and stored using an XLTEK® EMU128FS system (Natus Medical Incorporated, San Carlos, CA, USA) with 250 to 500 Hz sampling frequency. Table 1 and Fig. 1 report patient-specific information, number of electrodes included in this study, and electrode position.

Table 1

Seizure origin, number of channels, duration of recordings, sampling frequency, and frequency band used for estimation of the connectivity matrix.

Patient	Sex	Seizure origin	Types of seizures	# iEEG channels	Hours of recordings	Sampling frequency (Hz)	Frequency band for σ_1
1	Male	T	CP	34	40	500	13–30
2	Female	T	TC	28	47	500	4–7
3	Male	F	CP	44	47	250	13–30
4	Male	O	SP	26	34	500	13–30

Note. For each patient, the frequency band was chosen by maximizing the distance between ictal and nonictal generalized linear models (GLM) parameters (training data only).

Board-certified electroencephalographers (up to three) marked, by consensus, the unequivocal electrographic onset (UEO) [38] of each seizure and the period between seizure onset and termination. The time of seizure onset was indicated by a variety of stereotypical electrographic features, which could include, but were not limited to, the onset of fast rhythmic activity, an isolated spike or spike–wave complex followed by rhythmic activity, or an electrodecremental response. These features were typically present in one to a few channels at ictal onset. At all times, concurrent changes in the patient's behavior were sought from the video segment of a video-EEG recording. UEOs were used as the "gold standard" for evaluating the performance of the detection algorithm. Electrode recordings (iEEG in the following) included in this study were made available to the authors with the written consent of the patients, in accordance with the protocol approved by the institutional review boards at Brigham and Women's Hospital and Children's Hospital, Boston, MA, USA. Recordings included in this study cover a period of two consecutive days per patient (days with seizures).

2.2. Multichannel analysis

Recent studies have introduced schemes that analyze all the available electrode channels simultaneously [39–48]. In these schemes, each electrode is treated as a node in a graph, and any two nodes are considered *connected* (i.e., an edge exists between them) if the activities at these sites are dependent. The connectivity (topology) of the graph can then be described by a matrix. Statistics computed from this matrix (which is referred to as "connectivity" or "adjacency" matrix [49]) can show if the topology changes significantly from nonictal to ictal states or vice versa, and significant changes in these statistics can be used to detect a seizure onset.

In order to capture potential linear dependencies between all of the recording sites, we computed the connectivity matrix, A, as the cross-power in a chosen frequency band (theta, alpha, etc.) between the available iEEG channels. That is, for each pair of channels (i, j) the corresponding element of the connectivity matrix, A_{ij}, was

$$A_{ij} \triangleq \int_{lb}^{ub} P_{ij}(\omega)d\omega \qquad (1)$$

where $P_{ij}(\omega)$ is the cross-power spectral density of channels i and j at frequency ω [50]. The frequency band $[lb, ub]$ in (1) was patient specific and chosen among theta (4–7 Hz), alpha (8–13 Hz), and beta (14–30 Hz) bands (see Table 1 for the selection criteria) in our data set. Finally, the matrix A was computed over a 5-second-long sliding window (1-second [s] slide), resulting in a sequence of matrices $\{A(k)\}$, one for each time second k.

2.2.1. Singular value decomposition

It has been reported that the brain enters a more organized, lower-complexity state prior to a seizure [51,52]. As the brain becomes more organized and nodes become more connected, the *rank* (number of linearly independent rows or columns) of the
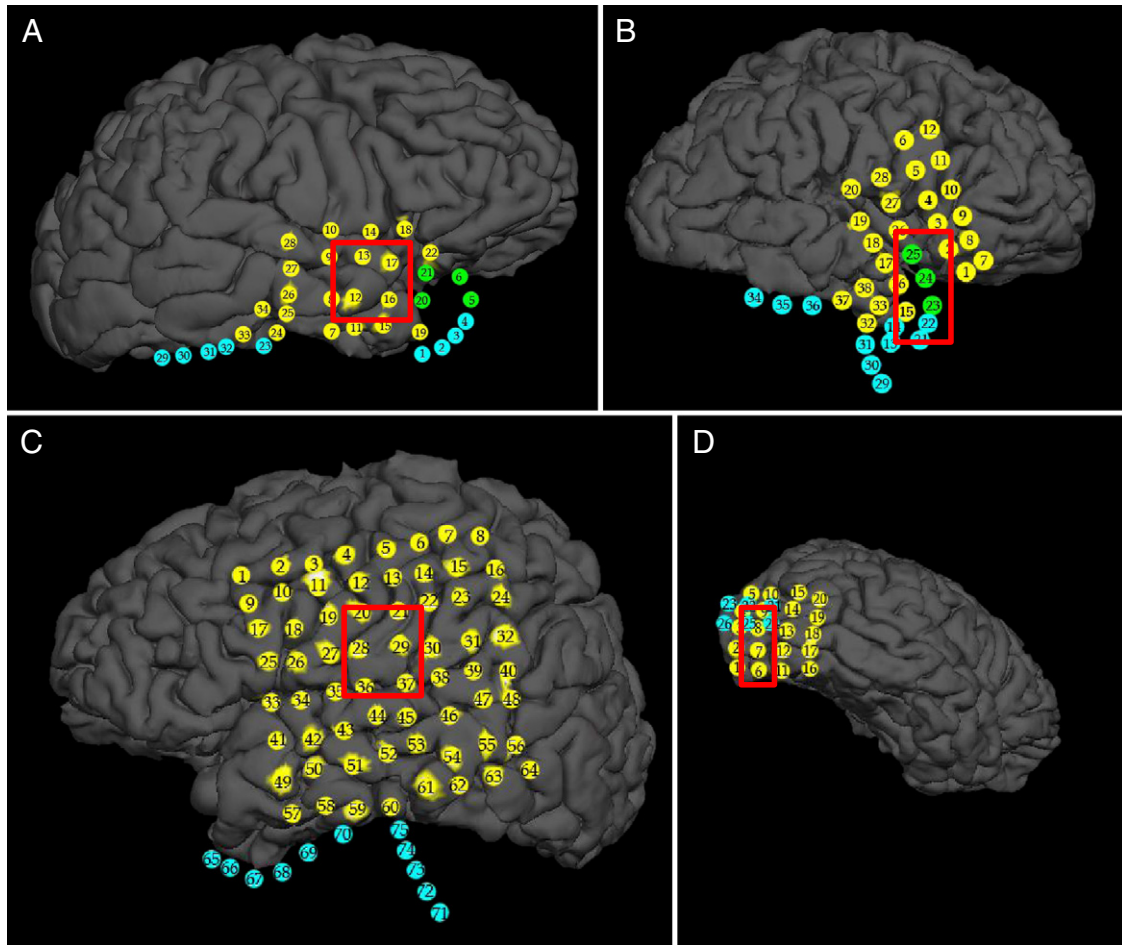
Fig. 1. Three-dimensional reconstructions of the brain for patients 1 (A), 2 (B), 3 (C), and 4 (D) with the electrode grids superimposed. Red-squared boxes denote the hand-annotated seizure foci. In (A) to (D), yellow circles are electrodes visible in their proper location on the cortex; green circles are electrodes that would be visible where the underlying cortex rendered properly; cyan circles are electrodes occluded by cortex (rendered or not) because they wrap around underneath or behind the temporal lobe. In patient 2, recordings from electrodes 1, 4, 6, 23, and 33–37 were not included in this study. In patient 3, recordings from electrode 34–64 were not included in this study. Electrodes not included in this study showed negligible modulation during the ictal periods.

connectivity matrix drops. The singular value decomposition (SVD) of a matrix highlights the rank of a matrix and also generates a weighted set of vectors that span the range space and null space of the matrix [53]. We therefore used SVD to measure the time-varying complexity of the brain by tracking the rank and its associated subspaces as a means to characterize nonictal versus ictal states.

The SVD of a generic $m \times n$ connectivity matrix A is defined as

$$A = USV^* = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_m] \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ \hline & & & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1^* \\ \boldsymbol{v}_2^* \\ \vdots \\ \boldsymbol{v}_n^* \end{bmatrix} \quad (2)$$

where U is an $m \times m$ unitary ($UU^* = I$) matrix whose columns, $\boldsymbol{u}_i$, are the eigenvectors of the matrix AA^*, V is an $n \times n$ unitary matrix whose columns, $\boldsymbol{v}_i$, are the eigenvectors of the matrix A^*A, and * denotes the complex conjugate transpose operator. S is an $m \times n$ matrix whose first r diagonal entries $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ are the nonzero singular values of A, with r being the rank of A [53]. The first r columns of U span the column space of A and the first r rows of V span the row space of A. When $m = n$ and A is symmetric, the SVD reduces to the conventional eigenvalue decomposition, where the singular values are the square of the eigenvalues of A, $U = V^{-1}$, and the columns of U and V are the eigenvectors of A [53].

An example is shown in Fig. 2. Here, two three-node graphs are analyzed. In Fig. 2A all the nodes have similar weak connections (strength ≤ 1). The SVD of the corresponding connectivity matrix, A, reveals that the matrix of this graph has full rank (three non-zeros and comparable singular values). More physically, full rank here indicates that the activity in the three nodes spans a three-dimensional space, or has 3 degrees of freedom.

If one of the connections is strengthened, as is the case between nodes 1 and 2 (Fig. 2B), one of the singular values of the corresponding connectivity matrix, B, becomes small in comparison to the other two, indicating that the rank of matrix has approximately dropped to 2. This means that with the addition of one strong connection, the activity of the graph collapses to two dimensions and become more "ordered". The singular vectors of graphs in Figs. 2A and B are plotted in Fig. 2C and indicate that the dominant direction of the vectors has rotated and the average amplification of the connectivity matrix has increased when the strong connection is added. We investigate the first (i.e., maximum) singular value, $\sigma_1(k)$, from each connectivity matrix $A(k)$ computed in (1) and apply QD on this statistic.

2.3. Quickest detection of seizure onsets

2.3.1. Hidden Markov model estimation

For any given patient, we assume that the maximum singular value computed at each second (observations), $z_k \triangleq \sigma_1(k)$, $k = 1, 2, \ldots$, is
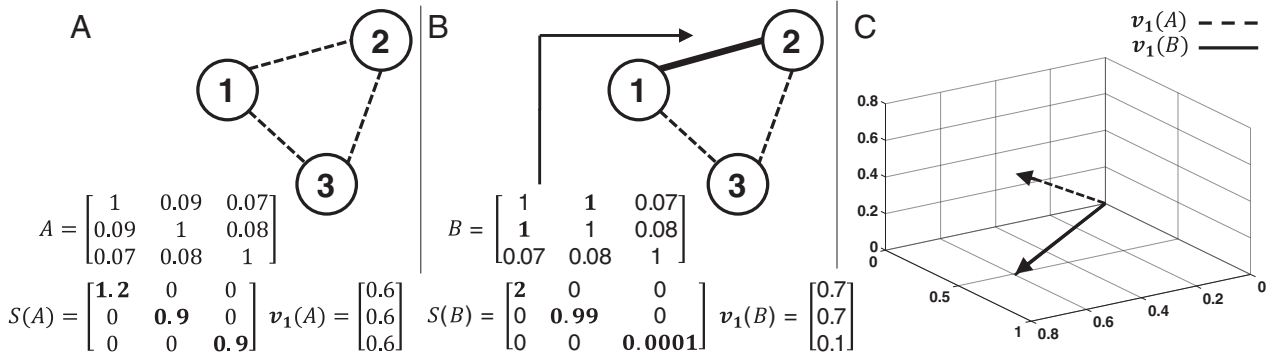
Fig. 2. (A) Three nodes in a network are loosely connected and the correspondent connectivity matrix A is close to identity with full rank. (B) Two of three nodes are strongly connected and the connectivity matrix B is close to losing rank. (C) Average direction and amplification of the first singular vectors corresponding to matrices A and B.

generated by a hidden Markov model (HMM) [54]. At each stage $k \geq 0$, the brain is in one of m states, that is, $x_k \in \{0, 1, ..., m-1\}$, which follows a Markov chain [54],

$$\Pr(x_{k+1} = j | x_k = i, x_{k-1}, ..., x_0) = \Pr(x_{k+1} = j | x_k = i) \triangleq p_{ij},$$

$$\sum_{h=1}^{m-1} p_{ih} = 1 \tag{3}$$

$$p_0 + p_1 + \cdots + p_{m-1} = 1,$$

for all $i, j = 0, 1, 2, ..., m-1$, where $p_i \triangleq \Pr(x_0 = i)$, $i = 0, 1, ..., m-1$, is the probability of starting in state i. For a fixed state i, the observations z_k are generated according to a known history-dependent probability law $q_i(z|H_k) \triangleq \Pr(z_k = z | x_k = i, H_k)$, where $H_k \triangleq \{z_0, z_1, ..., z_{k-1}\}$ denotes the sequence of past observations. Note that the dependency of z_k on previous observations is introduced to account for temporal dependencies that exist in neural data. The HMM is uniquely defined by the triple $\{\mathcal{P}, \Sigma, q\}$, with $\mathcal{P} \triangleq [p_0 p_1 \cdots p_{m-1}], \Sigma_{ij} \triangleq p_{ij}, i, j = 0, 1, ..., m-1$, and $q \triangleq [q_0 ... q_{m-1}]$. See Fig. 3A for a schematic of a $m = 2$ state HMM. For our QD framework, we fitted a $m = 2$ state HMM on each patient, with state $x = 0$ and $x = 1$ denoting the nonictal and ictal conditions, respectively (Fig. 3B). The ictal state begins and ends with the unequivocal ictal onset and offset determined by clinicians. Early ictal or preictal conditions were subsumed in the nonictal state as they may not exist in all patients.

Because we begin monitoring a patient in the nonictal state 0, we set $\mathcal{P} = [1 \ 0]$. We also assume that the state transition probability matrix is

$$\Sigma = \begin{bmatrix} 1-\rho & \rho \\ 0 & 1 \end{bmatrix} \tag{4}$$

where ρ was estimated from training data via maximum likelihood estimation [55]. The output probability law $q_x(z|H_k)$, $x = 0, 1$ was computed by combining generalized linear models (GLMs) [56] and maximum likelihood estimation. Observations were first quantized and mapped to integer nonnegative numbers in order to have a discrete observation domain, i.e., $\mathcal{Q}([z_k]) \triangleq n_k$ with $n_k \in \mathcal{Z}_0^+$, for all k. Then, we assumed that the probability mass function of the sequence n_k follows the Poisson law,

$$q_x(z|H_k) \cong \Pr(n_k = \mathcal{Q}([z]) | H_k, x) \triangleq e^{-\lambda_{x,k}} \frac{\lambda_{x,k}^{\mathcal{Q}([z])}}{\mathcal{Q}([z])!} \tag{5}$$

where the instantaneous rate $\lambda_{x,k}$ depends on the current state x, stage k, and history H_k. Finally, we assumed $\lambda_{x,k}$ is given by the GLM

$$\log \lambda_{x,k} = \alpha_x + \sum_{j=1}^{L} \beta_{x,j} n_{k-j} \tag{6}$$

where $\Theta_x \triangleq \{\alpha_x, \beta_{x, 1}, ..., \beta_{x, L}\}$ is a vector of parameters to be fitted on the data via maximum likelihood estimation. The number $L = 15$ of past observations to be used in (6) determines the number of parameters in Θ_x and was chosen by minimizing the Akaike's information criterion (AIC) [57] over a set of candidate models. For each patient, the parameter vector Θ_x was estimated separately for state $x = 0$ and $x = 1$ on training data. Such data included 3 h of continuous interictal recordings well before seizure (min 3 h, max 12 h before the seizure) and 1 (patients 1–3) or 2 (patient 4) hand-annotated ictal periods.

2.3.2. HMM state evolution and QD policy

Because the state of a HMM is hidden, we introduce the Bayesian *information state variable* $\pi_k \triangleq \Pr(x_k = 1 | z_k, H_k)$ [37,58] in order to estimate how likely the transition from the nonictal to ictal state is at each stage k. Note that π_k is the *a posteriori* probability of being in state 1 at stage k and depends on the observations up to and including stage k. The evolution equation of π_k is recursive and given by

$$\pi_{k+1} = \frac{L_{k+1}[\pi_k + (1-\pi_k)p_{01}]}{(1-\pi_k)(1-p_{01}) + L_{k+1}[\pi_k + (1-\pi_k)p_{01}]} \triangleq \Phi_k(\pi_k, z_{k+1}, H_{k+1}) \tag{7}$$

with initial condition

$$\pi_0 = P(x_0 = 1 | z_0) = \frac{q_1(z_0)(1-p_0)}{q_0(z_0)p_0 + q_1(z_0)(1-p_0)}$$

where $q_x(z_0)$ is the probability of observing z_0 in state x at time $k = 0$, $p_{01} = \rho$ because of (4), and

$$L_k \triangleq \frac{q_1(z_k | H_k)}{q_0(z_k |, H_k)}$$

is the likelihood ratio. See Appendix for details.

The Quickest Detection (QD) problem is an *online* decision problem, where at each stage k we test the hypothesis $\mathcal{H} \triangleq \{a \text{ seizure onset has occurred}\}$ conditioned on the observations (H_k, z_k). We introduce the decision variable $u_k \in \{0, 1\}$, where $u_k = 0$ ($u_k = 1$) denotes that the hypothesis $\mathcal{H}$ is rejected (accepted) at stage k. In this way,

$$\pi_{k+1} = f_k(\pi_k, z_{k+1}, H_{k+1}, u_k) \triangleq \begin{cases} \Phi_k(\pi_k, z_{k+1}, H_{k+1}) & u_k = 0 \\ terminate \ \& \ restart & u_k = 1 \end{cases} \tag{8}$$

where the "*terminate & restart*" state implies that we restart the detection algorithm after a seizure has been detected. With this setup, the detection problem boils down to *deciding when* to switch from $u_k = 0$ to $u_k = 1$ and claiming that a seizure has occurred.
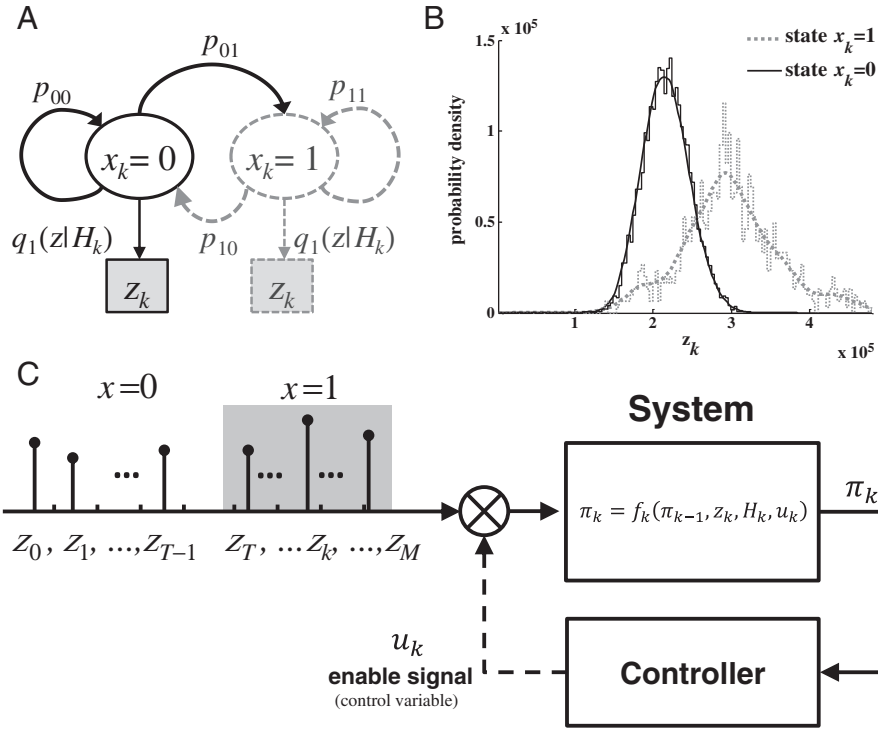
Fig. 3. (A) Schematic of an $m=2$ state HMM. (B) Sampling probability distribution function of the observations z_k in state $x_k=0$ (nonictal) and $x_k=1$ (ictal). Data collected from patient 2. (C) Seizure onset detection formulated as a feedback control problem.

We designed a decision strategy that minimizes the following cost function, which weighs the average detection delay and the probability of a false positive:

$$\mathcal{J}_0 \triangleq (1-\gamma)E_{T_{|T_{QD}<T}}\{T-T_{QD}\} + \gamma\, E_{T_{|T_{QD}>T}}\left\{(T_{QD}-T)^2\right\} \tag{9}$$

where T and T_{QD} denote the actual and estimated seizure onsets, respectively. T is a random variable whose distribution is defined by the HMM transition probabilities, i.e., $P(T=k)=(1-\rho)^{k-1}\rho$. It is important to note that a time-varying HMM would determine that $P(T=k)=\prod_{i=1}^{k-1}(1-p_{01}(i))p_{01}(k)$.

$E_{T_{|T_{QD}<T}}\{\cdot\}$ and $E_{T_{|T_{QD}>T}}\{\cdot\}$ denote the expected value of the distance between T_{QD} and T in case of false positive ($T_{QD}<T$) and delayed detection ($T_{QD}>T$), respectively. Finally, the parameter $\gamma\in[0,1]$ allows one to trade off false-positive and delayed detection, while the expected value $E_T\{\cdot\}$ accounts for the average temporal distance between actual and estimated seizure onset. Note that the distance $T_{QD}-T$ is squared in (9) for $T_{QD}>T$ as we intend to penalize more situations where $T_{QD}\gg T$, which account for very late detections.

To construct our QD policy, we constructed the cost (9) as a function of the information state π_k. To do so, we first defined a cost-per-stage, $G_k(\pi_k,u_k)$, that penalizes both rejecting $\mathcal{H}$ after the state transition has occurred (i.e., delay, $k\geq T$) and accepting $\mathcal{H}$ before the actual state transition has occurred (i.e., false positive, $k<T$), and is 0 otherwise:

$$G_k(\pi_k,u_k) \triangleq \begin{cases} \gamma E_{T|k>T}k-T\}\cdot\pi_k & u_k=0 \\ (1-\gamma)E_{T|k<T}T-k\}\cdot(1-\pi_k) & u_k=1 \\ 0 & otherwise \end{cases}$$

Then, the decision deals with choosing the stage $T_{QD}>0$ such that the policy ($u_1=0, u_2=0, ..., u_{T_{QD}-1}=0, u_{T_{QD}}=1$) minimizes the overall cost

$$\mathcal{J}_0 \triangleq E_{z_0,...,z_M}\left[G_M(\pi_M) + \sum_{i=0}^{M-1} G_i(\pi_i,u_i)\right] \tag{10}$$

where $G_M(\pi_M)$ is the final stage cost for rejecting hypothesis $\mathcal{H}$. M is the horizon over which a seizure must be detected, and we set M for each patient to be the value of the average inter-seizure interval. It is fairly straightforward to show that (10) is equivalent to (9) [59]. One can interpret the minimization of (10) with respect to the variable u_k given the evolution model (8), as an optimal feedback control problem where u_k is the control variable (Fig. 3C). As a control problem, this formulation can be solved recursively via dynamic programming [59], and leads to the optimal QD policy

$$T_{QD} = min\{0<k<M|\pi_k>F_k(\pi_k,z_k,H_k)\}$$

where $F_k(\pi_k,z_k,H_k)$ is an adaptive threshold that depends on the current observation, history, and information state variable. There is no simple closed form expression for $F_k(\cdot)$ which is computed recursively and decreases over time non-monotonically as discussed below. Details on the derivation of $F_k(\cdot)$ can be obtained in [37].

2.3.3. Significance and performance tests

For each patient, we compared the QD policy with a classic Bayesian estimator (BE) [58], which is widely used in the field of change-point detection [35,58]. We also compared the QD policy versus a threshold-based detector (HT), where the threshold is chosen heuristically. The formula for the estimated seizure onset with each of these predictors is

$$BE : T_{BE} \triangleq min\{k>0|\pi_k>0.5\}$$

$$HT : T_{HT} \triangleq min\left\{k>0\left|z_k>\bar{h}\right.\right\}$$

where the threshold $\bar{h}$ is fixed. In particular, we chose $\bar{h} \triangleq \mu_z + 3\sigma_z$, where μ_z and σ_z are the mean and standard deviation (SD) of z_k over the 3-h nonictal training data, respectively.

For each detection policy, we measured the delay between each estimated seizure onset time and the correspondent UEO, the number of true positives (TP), false positives (FP), and false negatives (FN).
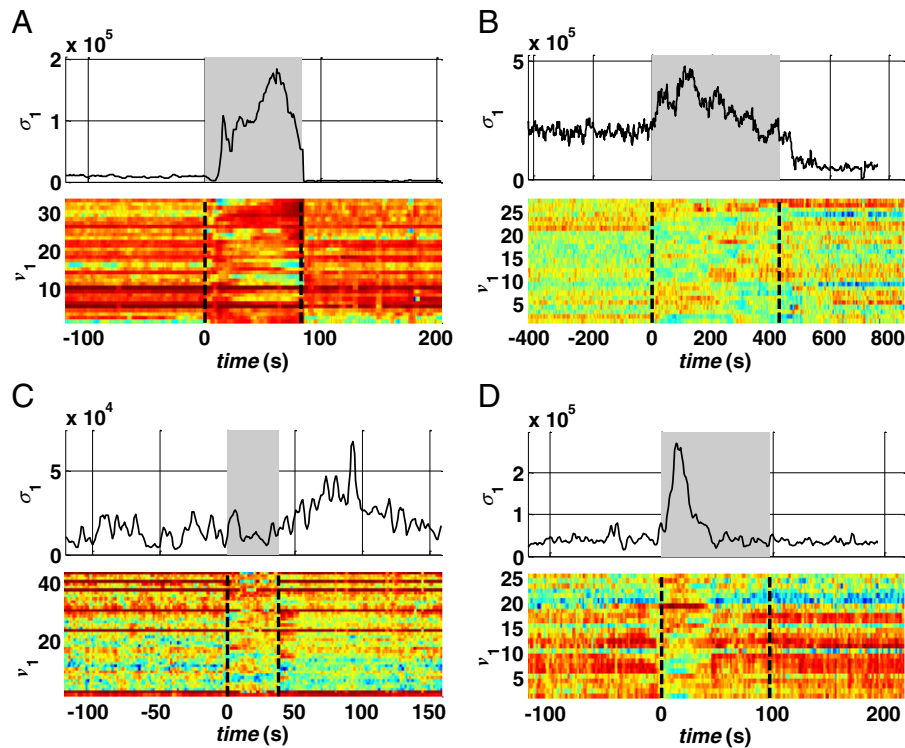
Fig. 4. Sequence of first singular values σ_1 and correspondent first singular vectors v_1 around an ictal period (gray background) in patient 1 (seizure s_1, A), 2 (s_1, B), 3 (s_3, C), and 4 (s_3, D).

Every detection was classified as TP or FP if an UEO occurred within 20 s from the detection time or not. This time was chosen to be comparable to [14]. Each UEO that was not detected was classified as FN. Results are summarized in Tables 2 and Table 3.

3. Results

3.1. Multichannel analysis

The maximum singular value σ_1 estimated at each stage k from the connectivity matrix (1) is plotted in Fig. 4 (one ictal period per patient). The sequence of σ_1 had a consistent pattern across the patients, with large values in nonictal state (pre- and post-seizure). The corresponding singular vector $\boldsymbol{v}_1$ shows a leading direction before the seizure onset, which depends on both the specific patient and the location of the foci. For example, components 8–10 in patient 1 (Fig. 4A) and 1–4 in patient 3 (Fig. 4C) correspond to the focal area (Figs. 1A, C) and have significantly higher values than the other components of $\boldsymbol{v}_1$ before the seizure.

During a seizure, however, σ_1 rapidly increases compared with the nonictal background activity in the previous minutes, reaches a local maximum approximately half of the ictal period (gray boxes, Figs. 4A–D), and then slowly decreases to smaller, nonictal values. The change in the dynamics of σ_1 was observed almost at the beginning of the hand-annotated seizure onset while the return to the nonictal condition was usually slower and may last longer than the hand-annotated end time of the seizure. Interestingly, after every seizure, σ_1 decreased below the average value achieved before the seizure and, then, increased to the preictal values with a long drift (at least 2 hours, data not reported), which may be consistent with the definition of a postictal state given in [51,52]. The stereotypical dynamics of σ_1 were associated with a sudden change in the direction of the singular vector $\boldsymbol{v}_1$: the components with the largest values during the nonictal period decreased during the seizure while the remaining components increased. As a consequence, the distance between the largest and smallest components consistently decreased during a seizure (more than 30%), indicating a rotation of $\boldsymbol{v}_1$ toward a new direction which varies with the specific seizure.

Table 2
Performance analysis.

	Patient	h of recordings	# of seizures	QD					BE					HT				
				Specificity			Sensitivity		Specificity			Sensitivity		Specificity			Sensitivity	
				FP	FPR (FP/h)	FN	TP	% total	FP	FPR (FP/h)	FN	TP	% total	FP	FPR (FP/h)	FN	TP	% total
Training	1	3	1	0	**0**	0	1	100	0	0	0	1	100	8	2.67	0	1	100
	2	3	1	0	**0**	0	1	100	0	0	0	1	100	8	2.67	0	1	100
	3	3	1	3	**1**	0	1	100	3	1	0	1	100	8	2.67	1	0	0
	4	3	2	0	**0**	0	2	100	0	0	0	2	100	0	0	0	2	100
Validation	1	37	1	6	**0.16**	0	1	100	6	0.16	0	1	100	42	1.14	0	1	100
	2	44	2	7	**0.16**	0	2	100	11	0.25	0	2	100	13	0.30	0	2	100
	3	44	3	130	**2.95**	0	3	100	320	7.27	0	3	100	75	1.71	3	0	0
	4	31	22	71	**2.29**	0	22	100	138	4.45	0	22	100	277	8.94	0	22	100

Note. FP, false positive; TP, true positive; FN, false negative; FPR, false positive rate.
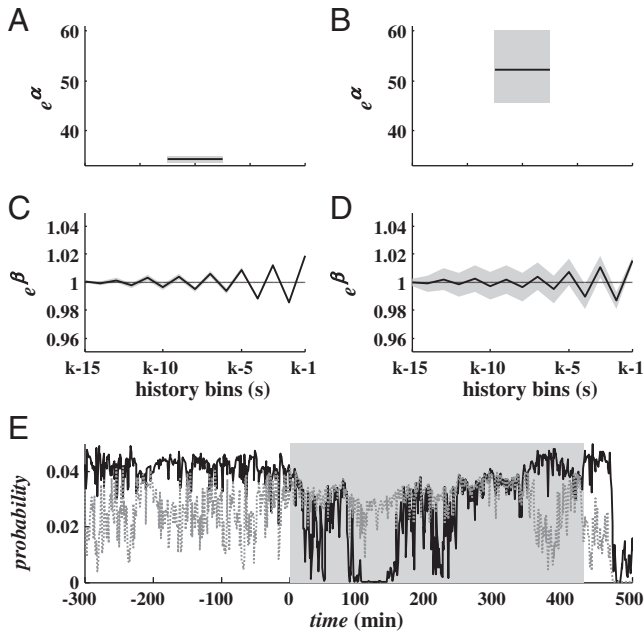
Fig. 5. (A–D) Maximum likelihood estimate of parameters Θ_x (black line) in (6) with 95% confidence bounds (gray background) for nonictal (A, C) and ictal (B, D) state. (E) History-dependent estimate of the probability $q_0(z_k|H_k)$ (black line) and $q_1(z_k|H_k)$ (gray dotted line) around a specific hand-annotated seizure (gray background). Plots refer to patient 2. Probabilities in (E) refer to seizure s_3 (validation data).

Modulo a scaling factor, the dynamics of $\sigma 1$ and $\mathbf{v}_1$ were similar in patients 1, 2, and 4 (Figs. 4A, B, D), independently of the type/origin of the seizure and the connectivity matrix (σ_1 and $\mathbf{v}_1$ were computed from connectivity matrices of cross-power in different frequency bands, see Table 1). These dynamics were less clear in patient 3, where σ_1 showed slow oscillations independently of the seizure occurrence (Fig. 4C, top). However, at the seizure onset, the value of σ_1 first decreased, then rapidly rose to a local maximum, and finally drifted toward the baseline value, as did occur in the other patients. This indicates that, modulo a scaling factor, the same dynamics occurred in patient 3 as well as patients 1, 2, and 4. Interestingly, in Fig. 4C (top), σ_1 achieves a peak value ~50 s after termination of the seizure. However, this peak is shorter than the pattern during the previous seizure, is not followed by a recovery phase, and is not associated with a rotation of $\mathbf{v}_1$. The leading components of $\mathbf{v}_1$, indeed, do not change when this second peak occurs and are the same as in the preictal phase (Fig. 4C, bottom). These facts indicate that, although the absolute value of σ_1 may be higher after the seizure, the specific pattern is different from the ictal one.

Overall, the clear modulation of σ_1 at the seizure onset and its specific pattern during the ictal state may capture the overall change in complexity of the brain. This is possible because σ_1 exploits information recorded from multiple sites simultaneously and combines linear dependencies between all the possible pairs of electrodes in the beta (patients 1, 3, 4) and theta (patient 2) frequency bands.

3.2. HMM estimation

Figs. 3B and 5 show results for the HMM estimation. Although the mean value and the variance of σ_1 were different in the ictal versus nonictal states, the sampling probability distribution functions had some overlap (Fig. 3B), which means that several of the same values of σ_1 were likely to be achieved in both ictal and nonictal states.

In order to better characterize the distribution of σ_1 in each state, we introduced a history-dependent representation of the

Table 3
Detection delay on validation data.

Patient	Seizure	Duration (s)	Delay (s)		
			QD	BE	HT
1	s_1	101	15	14	12
	s_2	180	37	35	10
2	s_3	428	25	31	19
	mean		**31**	**33**	**14.5**
	s_2	74	13	–5	*na*
	s_3	38	–7	–10	*na*
3	s_4	36	21	17	*na*
	mean		**9**	**0.67**	***na***
	mean abs. value		**13.67**	**10.67**	***na***
	median		**13**	**–5**	***na***
	s_3	97	1	–1	–2
	s_4	46	14	12	2
	s_5	45	–7	–10	–11
	s_6	60	–4	–7	–8
	s_7	16	9	7	0
	s_8	37	–8	–11	–11
	s_9	35	10	7	6
	s_{10}	33	20	5	4
	s_{11}	41	9	–1	4
	s_{12}	38	11	9	4
	s_{13}	38	12	7	5
4	s_{14}	31	14	11	6
	s_{15}	30	12	9	5
	s_{16}	94	–4	–14	–15
	s_{17}	55	1	–1	–2
	s_{18}	21	7	0	–1
	s_{19}	51	6	–1	–1
	s_{20}	29	3	0	0
	s_{21}	49	8	3	–1
	s_{22}	45	13	11	8
	s_{23}	50	23	20	2
	s_{24}	25	15	12	7
	mean		**7.5**	**3.05**	**0.045**
	mean abs. value		**9.59**	**7.23**	**4.77**
	median		**9**	**4**	**1**
	standard deviation		**8.32**	**8.56**	**6.30**
	mean		**9.6**	**5.32**	**1.68**
	mean abs. value		**11.75**	**9.68**	**5.84**
TOT	**median**		**10.5**	**6**	**2**
	standard deviation		**10.5**	**11.78**	**7.55**

Note. Positive and negative values denote delay and anticipation, respectively.
[a] Mean of the absolute values of delays and anticipations.

probability distribution of σ_1 (Fig. 5). At each stage k, the model (5)–(6) modulates the probability that the observation $\sigma_1(k)$ has been emitted while in the ictal or nonictal state based on the values of σ_1 in the last 15 s. The dynamics of the instantaneous rate $\lambda_{x,k}$ are captured by parameters Θ_x in (5) and (6), as Θ_x were estimated from actual observations in both states. Parameters in Figs. 5A–D indicate that the transition from nonictal to ictal state is characterized by (i) a significant increase in parameter α_x (Figs. 5A, B), which is a scaling factor and accounts for the average value of σ_1 in state x, and (ii) larger 95% confidence bounds for parameters $\beta_{x,\ j}$ (Figs. 5C, D), which accounts for larger differences between consecutive observations of σ_1.

The different model parameters Θ_x resulted in distinct functions q_0, q_1, which (i) varied the probability of any given observation σ_1 at each stage k depending on the past observations, and (ii) had opposite dynamics in the ictal and nonictal states (Fig. 5E). In particular, for the computed sequence of σ_1 in each patient, $q_1(\cdot)$ was consistently larger than $q_0(\cdot)$ during the ictal periods, but decreased during the nonictal periods and was almost zero during the postictal phase, while $q_0(\cdot)$ was generally high (and almost always larger than $q_1(\cdot)$) during the nonictal periods. In each patient, q_0 and q_1 were almost zero after every seizure independently of the model parameters, thus suggesting that the postictal period is characterized by a resetting of the brain activity, as argued in [51,52].

3.3. Quickest detection policy

Tables 2 and 3 and Figs. 6 and 7 report the results for the QD policy versus the Bayesian estimator (BE) and the heuristic threshold-based detector (HT) described under Methods.

The QD policy was derived by penalizing the probability of false positives more than the delay (i.e., $\gamma < 0.5$ in the cost function (9)). In this way, QD guaranteed 100% sensitivity (i.e., all the seizures were correctly detected) on both training and validation data (Table 2) and delays were generally small (average across patients: 9.6 ± 10.5 s, mean $\pm$ SD). Detection occurred earlier than the UEO (anticipation) in 5 of 28 seizures on validation data (Table 3), with a lag ranging from 4 to 7 s. The number of false positives, instead, varied with the patient (Table 2) and determined an average FPR value of 1.39 ± 1.45 FP/h across four patients.

A comparison between QD, BE, and HT indicates that the non-optimal policies BE and HT can achieve shorter delays, but the number of false positives rapidly increases, which may make these methods unsuitable for clinical application. While the average delay was 5.32 ± 11.78 s and 1.68 ± 7.55 s for BE and HT, respectively, the FPR was 3.0 ± 3.46 and 3.0 ± 3.99 FP/h, respectively. Also, the performance of the HT was influenced by the noise in the sequence of σ_1 values, resulting in 0% sensitivity in one patient (i.e., no seizure was correctly detected) and an average value of 82.5% across the whole population.

These results depend on the performance goals imposed by the cost (9). By increasing the penalty for detection delay in (9) (i.e.,

increasing γ), the QD was able to reduce the delay to the values achieved via BE (16 s, Fig. 6B, black dash-dot line), while keeping a lower number of FPs (7 vs. 11). By decreasing γ, instead, we achieved higher robustness to early modulations of π_k, which can be due to abrupt spikes in the sequence of σ_1, and were able to decrease the number of FPs and detect a seizure with less anticipation than the other methods (Figs. 7A, B).

A sensitivity analysis of the QD policy to variations in the parameter γ is reported in Fig. 8. The policy was implemented for several values of γ in [0, 1], and for each value, the average delay and FPR per patient were estimated on the validation data. Results indicate that delays are quite insensitive to modulations in γ in patients 1 and 2, and low FPR values (<0.17 FP/h) can be achieved while keeping the delay to the minimum value. However, in patient 4, delays were generally small (<10 seconds) for FPRs below 2 FP/h, and further reductions in delay rapidly increase the FPR. Finally, performances for patient 3 was similar to that for BE and HT in terms of FPRs, presumably because of the less significant modulation of σ_1 in the ictal versus nonictal state (Fig. 4).

4. Discussion

In this study, we propose a multichannel statistic that measures linear dependencies among all recorded sites of the epileptic brain simultaneously by combining power spectrum analysis and matrix theory. This statistic (the maximum singular value σ_1 of the cross-
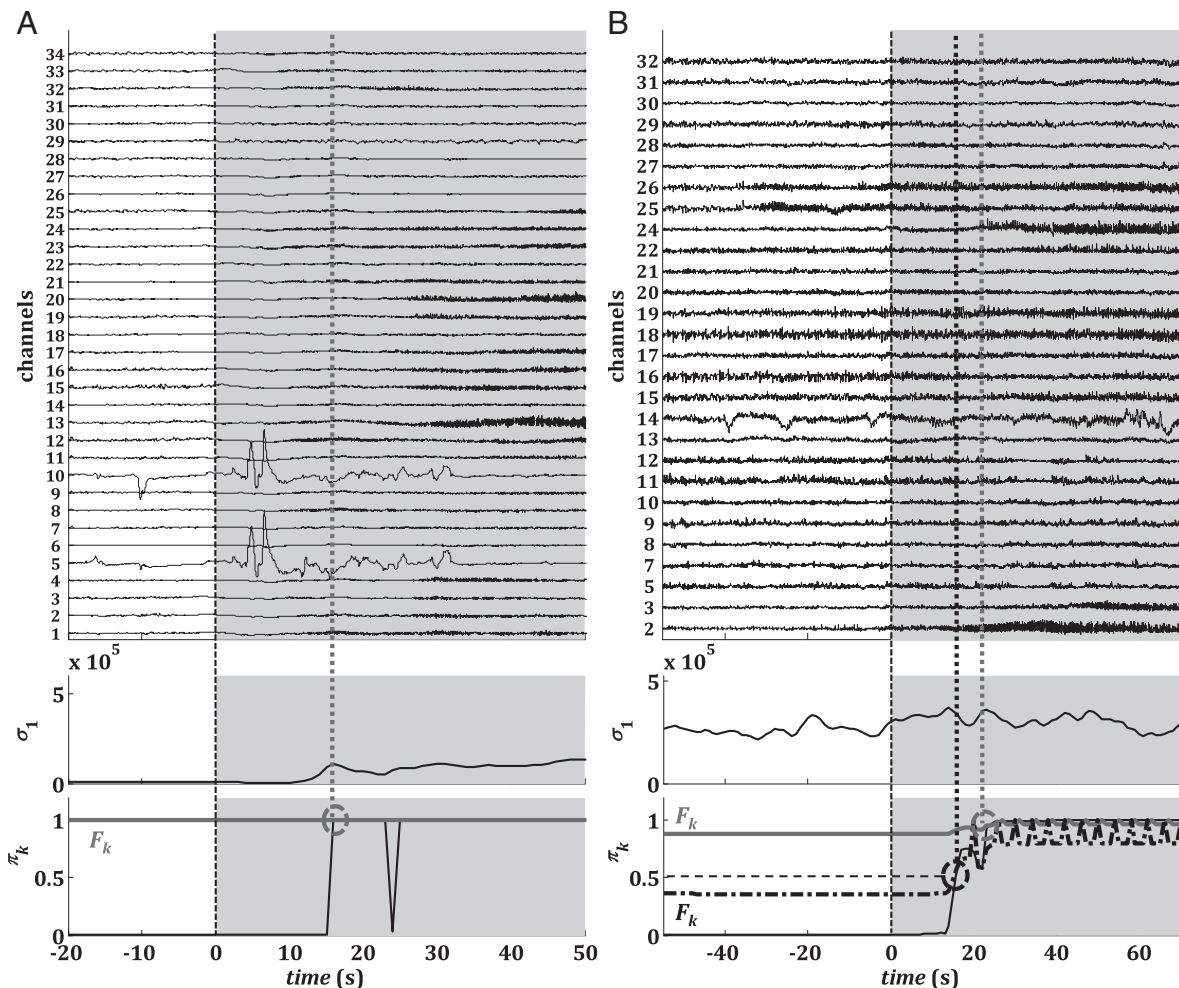

Fig. 6. QD on validation data. The electrographic onset (dashed vertical black line), the correspondent QD estimation (circles), and threshold F_k (gray and dash-dot black lines) for patients 1 (A) and 2 (B). For patient 2, two QD estimations are computed for different values of parameter γ (black circle $\gamma = 0.52$; gray circle $\gamma = 0.01$). Plots refer to seizures s_1 (patient 1) and s_3 (patient 2), respectively. In (B) (bottom plot) the horizontal dashed black line denotes the threshold for the BE detector.
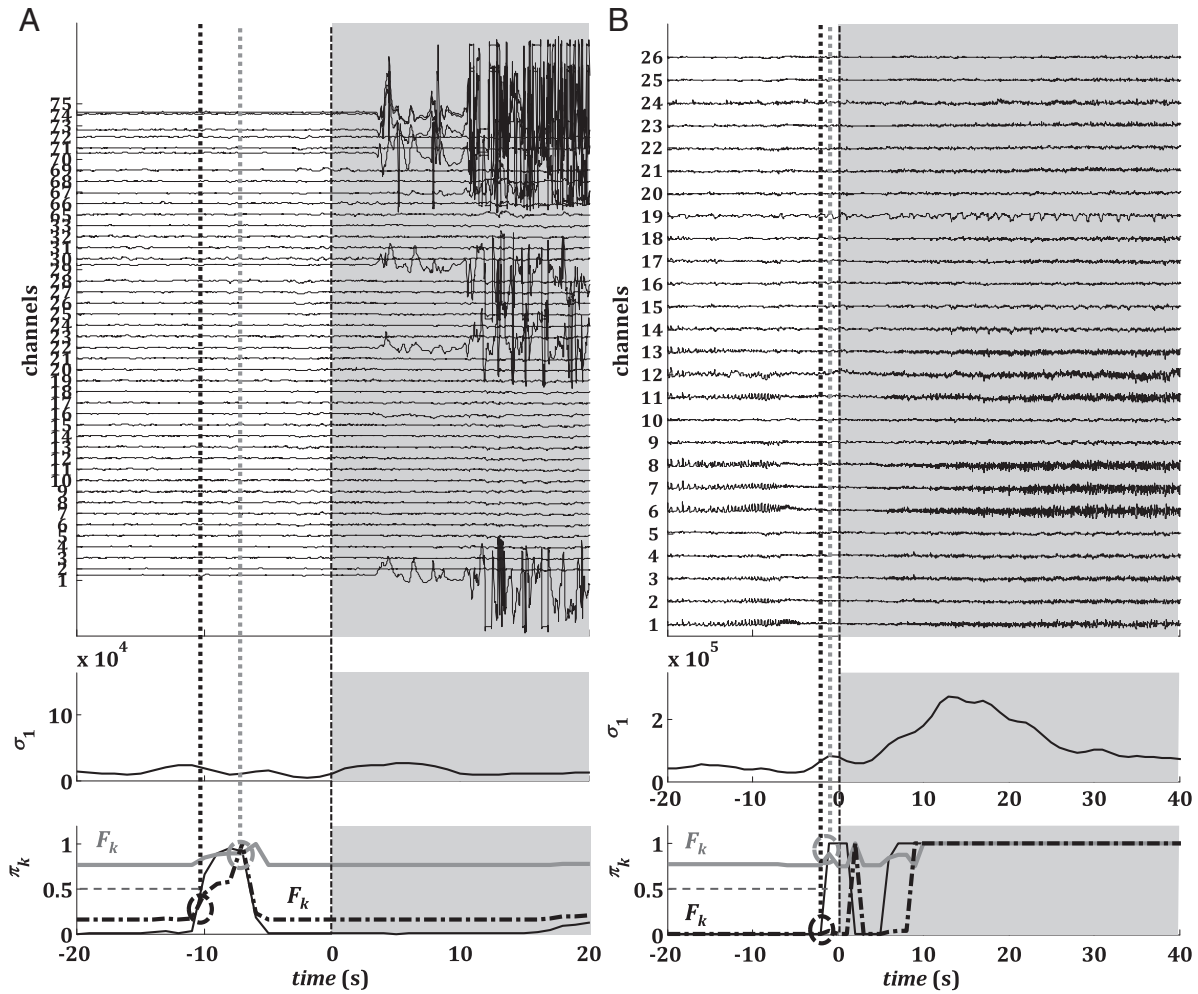
Fig. 7. QD on validation data. The electrographic onset (dashed vertical black line), the correspondent QD estimation (circles), and threshold F_k (gray and dash-dot black lines) for patients 3 (A) and 4 (B). For each patient, two QD estimations are computed (gray and dash-dot black) for different values of parameter γ. Plots refer to seizures s_3 (patient 3) and s_3 (patient 4), respectively. In (A) and (B) (bottom plot) the horizontal dashed black line denotes the threshold for the BE detector.

power-based connectivity matrix) summarizes the change in topology that occurs in the brain network at the seizure onset and shows significantly different dynamics in ictal versus nonictal periods.

Using this statistic, we developed a computational framework for the AOSD. This framework combines Bayesian estimation (we use the a posteriori probability π_k) with optimal control (we minimize the cost function $\mathcal{J}_0$ in (10)) and provides a threshold-based detection policy, where the threshold varies with time based on the dynamics of σ_1 and π_k.

4.1. Multichannel versus single-channel statistics

Several statistics, both single channel and multichannel, have been computed from iEEG signals in the last 20 years to capture changes occurring in the brain network at seizure onset [6–33]. Although these statistics show some modulation between different nonictal and ictal states, they have a few drawbacks: (i) The statistics are usually computed on single channels [6–9,11,28,30–33] or a small subset of channels [10,29] from the focal area, which means that the foci
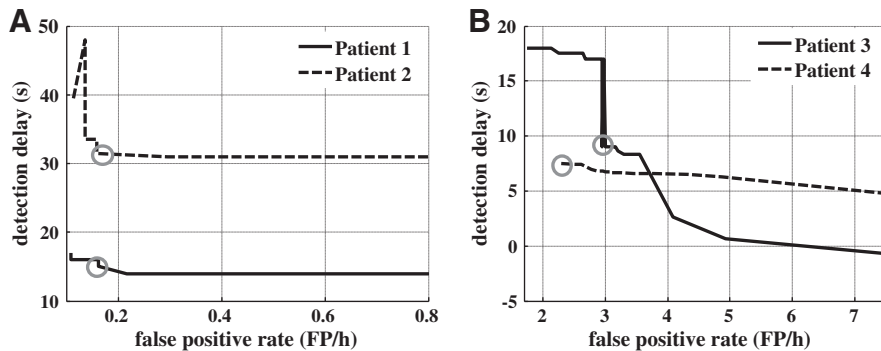

Fig. 8. Sensitivity analysis. The average detection delay and false-positive rate for patients 1 and 2 (A), 3 and 4 (B) are computed for γ ranging from 0.01 to 0.999 on the validation data. Gray circles denote the average performance achieved with the value of γ used in Tables 2 and Table 3.

must be known a priori with reasonable accuracy (i.e., localization and detection are correlated problems). This is less stringent with the multichannel statistics, where the only requirement is that the electrode grid is large enough to include the focal areas. (ii) The nonlinear multichannel statistics usually outperform linear single-channel and two-channel measures [60], but require larger amounts of data and computation. (iii) The modulation of these statistics around the seizure onset may vary with the subject or during the sleep/wake cycle [61], resulting in less predictable patterns.

These limitations can be (at least) in part addressed by increasing the number of combined channels and computing simple measures off of large enough matrices (on the order of hundreds of channels). Such a combination can provide more information about the brain network (albeit still incomplete) and exploits both spatial and temporal features, while the computation of the maximum singular value σ_1 captures the overall degree of (linear) dependency among different brain sites and the sampled network topology. These dynamics were similar in all the patients in our data set, with clear differences during ictal and nonictal periods (Fig. 4) and a postictal resetting pattern that may last several minutes to a few hours.

These results positively affected the dynamics of the Bayesian a posteriori probability of an ictal state (π_k). This probability evolves recursively with (7) and is used as a marker of the actual brain condition in the QD policy. Depending on the evolution of the maximum singular value, this probability clearly separates the ictal and nonictal periods, contributing to the success of the QD policy (Fig. 5). As indicated in Figs. 6 and 7, this a posteriori probability selectively increases at the seizure onset and remains high during the ictal period, before finally decreasing to zero. π_k rises to 1 (i.e., its maximum value) quickly at the seizure onset and usually has minor modulations outside the ictal period (Figs. 6 and 7). For this reason, the detection delay with QD is comparable to that achieved via BE, although the FPR is significantly lower in each patient with QD. The FPR is larger in BE because of the occurrence of abrupt spikes in the sequence of π_k that trigger false positives. These abrupt spikes in π_k may be caused by fluctuations or noise in the measure of the maximum singular value σ_1. In this case, an adaptive threshold allows increasing the performance by selectively avoiding these peaks (see below).

4.2. Quickest detection policy

The proposed QD policy combines several well-known topics in abrupt change-point detection theory (e.g., quickest detection [35,36], bayesian estimation with loss function [58], dynamic programming [59], HMM [54]), generalizes the framework to the case of history-dependent sequential observations (which may be relevant for neural data and iEEG recordings [37], etc.), generalizes the framework to the case of history-dependent sequential observations (which may be relevant for neural data and iEEG recordings [37]), and results in an adaptive, unsupervised, threshold-based detection strategy that can potentially be implemented online.

Differently from current AOSD paradigms, our QD policy constructs the threshold by explicitly minimizing a cost function of the desired detection performance goals. Many current detectors track the modulation of a statistic computed from iEEG recordings over time, and then set a threshold on the statistic to estimate the ictal onset time [6,10–15]. The choice of the threshold is *supervised* and usually dependent on the fluctuations of the statistic, the specific patient, or the electrode position, and may require long training sessions to be more accurate. In our approach, instead, an *unsupervised* adaptive threshold falls out of the methodology and adapts according to the cost function, the model of the multichannel statistic (HMM), and current and past measurements.

More recent detection schemes that exploit classifiers (e.g., [16–33]), based on either ANN, SVM, or PCA, also generate unsupervised criteria that separate the feature space (statistic space) into dominant

ictal and nonictal regions. Although optimization methods are used for separating ictal and nonictal data, the clustering procedure exploits the distance (in a specific high-dimensional geometric feature space) between the data, but does not encompass penalties for specific performance goals (e.g., minimize FPR).

In summary, the detection approaches proposed thus far follow a "bottom-up" flow, i.e., they determine criteria that most likely separate ictal and nonictal data, and then apply such criteria on sequential neural measurements and evaluate detection performance. Our QD policy, instead, follows a "top-down" approach, i.e., it requires a cost function that explicitly accounts for different performance goals (e.g., low probability of false positives, low distance between actual and detected seizure onset time, low probability of late detection, etc.) and then defines the detection strategy to minimize the cost function. Depending on the specific application (e.g., online detection of a clinical seizure, offline investigation of the electrographic onset of a known clinical seizure, etc.), the QD can be tuned to achieve a specific goal and reach the required level of sensitivity and specificity (Fig. 8).

Our QD policy also exploits a model-based control approach to seizure detection. Here, the QD formulation was derived for a two-state HMM representation of brain activity in ictal versus nonictal state. HMMs have been successfully used in several fields in statistics and engineering (for an overview, see [54]), and recently introduced for modeling neural data in sensitive tasks and neural prosthetics (e.g., [62,63]). In the case of seizure detection, the transition from a nonictal to an ictal state can be suitably treated as a hidden state transition, where the hidden state is the actual brain condition, which is unknown and partially sampled with electrodes. According to this interpretation, an epileptic brain can sequentially transition into $m > 2$ states, depending on the actual physiological conditions (e.g., the patient is awake or sleeping), the type of epilepsy, or the type of seizure occurring. The number of HMM states is generally patient specific and may vary with the available data. However, the framework outlined in (1)–(10) is general and can be extended to the case of HMMs with $m > 2$ states (e.g., imagine a separate m-state HMM for the nonictal periods and an n-state HMM for the ictal periods).

It is interesting that for AOSD, a minimal HMM with two states appears to be enough and led to low FPRs (<0.2 FP/h) in two of the four patients in our data set. This may stem from the clear patterned dynamics of the maximum singular values in ictal versus nonictal periods, which is captured by the GLM structure in (5) and (6) (Fig. 5). GLM and maximum likelihood methods have been widely used before in the analysis and simulation of neuronal spike train analysis for several types of neural disorders [55,64,65] and provide a flexible framework for both stationary and nonstationary analyses. In our case, the GLM parameters are able to accurately capture changes that occur in the maximum singular value as soon as the seizure starts, and require a minimal set of training data to be estimated (only one seizure and 3 hours of nonictal data) in both conditions.

We are aware, however, that in the remaining two patients the FPR was quite high (>2 FP/h). In this case, it is possible that the training data was not enough for modeling additional (slower) rhythms in the maximum singular value during the interictal state. Such additional dynamics presumably depend on the location of the focal region, which is not temporal in Patient 3-4, and the specific type of seizures. Indeed, while Patient 1 and 2 had distinct and large complex partial and tonic-clonic seizures, respectively, Patient 3-4 had short complex partial seizures with minor clinical evidence and simple partial seizures, respectively (Table 1). Another possible reason could be the spectral selectivity of our model, i.e., the fact that we consider only one frequency band per patient. Finally, a possible reason could be the limited number of states in our HMM, which could not be enough for these patients. These reasons led to a lack of model accuracy in patient 3 and 4, which presumably caused an increased fluctuation of the state variable π_k with frequent erroneous peaks and ultimately decreased the QD performance. Better results could be achieved by improving the

HMM–GLM model framework with further information about the type of epilepsy and patient behavior.

Another aspect of the QD results in Tables 2 and 3 that needs to be addressed is the excessive detection delay (>10 s) for a few seizures, which would not be useful for clinical application. Also, we note that the delay with QD was longer than with the HT and BE detector for several seizures, although the FPR was significantly lower.

The results of QD versus those of BE and HT are a consequence of the specific cost function that we defined, which penalizes the probability of false positives and very late detections more severely than detection delays in the order of 10–20 s. Our choice was motivated by the limited duration of the available recordings (less than 2 days per patient) and definitively resulted in a conservative policy, i.e., the threshold $F_k(\cdot) \gg 0$, and, therefore, a seizure was detected only when the state variable $\pi_k \gg 0$ (Figs. 6B, 7).

The issue of excessive delay may also be due to the spectral selectivity of the adopted connectivity. This can be a limitation as the seizures are characterized by several features (e.g., fast rhythmic activity, electrodecremental activity, isolated spikes, etc.), which may occur across multiple frequency bands. In this case, observations in a single frequency band would lead the multichannel statistic to late modulation and therefore a delayed detection.

A possible solution to this issue could be combining multichannel statistics across several frequency bands by combining maximum singular values of connectivity matrices computed in different frequency bands. Another solution to be explored is the use of nonlinear functions of the detection delay $T_{QD} - T$ in the cost function (9). In the current formulation, the penalty for delay in (9) grows quadratically with delay. We could allow this penalty to grow exponentially in delay ($e^{(T_{QD} - T)}$), and as long as the function is a non-decreasing function of the delay, the QD method will hold.

Future work entails validating our preliminary findings on larger data sets and reducing the detection delays so that they may be actionable for clinical intervention. We will also make direct comparisons with current AOSD approaches on the same set of statistics so that we may understand the degree of performance improvement we achieve with the QD methodology.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgments

Appendix. Recursion of π_k

Let us find the recursion equation for π_k as a function of π_{k-1} in a $m = 2$ HMM. We start with stage 0:

$$\pi_0 = \Pr(x_0 = 1 | z_0) = \Pr(x_0 = 1, z_0)/\Pr(z_0)$$

$$= p_1 q_1(z_0)/(p_1 q_1(z_0) + (1 - p_1)q_0(z_0)).$$

$$\pi_1 = \Pr(T \leq 1 | z_0, z_1) = \Pr(T \leq 1, z_0, z_1)/\Pr(z_0, z_1)$$

$$= \Pr(T \leq 0, z_0, z_1 \; or \; T = 1, z_0, z_1)/\Pr(z_0, z_1)$$

$$= \Pr(T \leq 0, z_0, z_1) + \frac{\Pr(T = 1, z_0, z_1)}{\Pr(z_0, z_1)}$$

$$= \frac{\Pr(z_0, z_1 | T \leq 0)\Pr(T \leq 0) + \Pr(z_0, z_1 | T = 1)\Pr(T = 1)}{\Pr(z_0, z_1)}$$

$$= \Pr(z_1 | T \leq 0, z_0)\Pr(z_0 | T \leq 0)/\Pr(z_1 | z_0)$$

$$+ \Pr(z_1 | T = 1, z_0)\Pr(z_0 | T = 1)\Pr(T = 1)/\Pr(z_1 | z_0)$$

$$= \frac{\Pr(T \leq 0 | z_0)\Pr(z_1 | z_0, T \leq 0) + \Pr(T = 1 | z_0)\Pr(z_1 | z_0, T = 1)}{\Pr(z_1 | z_0)\Pr(z_0)}$$

$$= \frac{\pi_0 q_1(z_1 | z_0) + (1 - \pi_0)p_{01}q_1(z_1 | z_0)}{\pi_0 q_1(z_1 | z_0) + (1 - \pi_0)p_{01}q_1(z_1 | z_0) + (1 - \pi_0)(1 - p_{01})q_0(z_1 | z_0)}.$$

Define the likelihood ratio

$$L_1 \triangleq \frac{q_1(z_1 | z_0)}{q_0(z_1 | z_0)},$$

then

$$\pi_1 = \frac{[L_1[\pi]_0 + (1 - \pi_0)p_{01}]}{L_1[\pi_0 + (1 - \pi_0)p_{01}] + (1 - \pi_0)(1 - p_{01})}$$

Similarly, it is straightforward to show that for

$$L_k \triangleq \frac{q_1(z_k | H_k)}{q_0(z_k | H_k)},$$

where $H_k = \{z_0, z_1, ..., z_{k-1}\}$, then

$$\pi_k = \frac{[L_k[\pi]_{k-1} + (1 - \pi_{k-1})p_{01}]}{L_k[\pi_{k-1} + (1 - \pi_{k-1})p_{01}] + (1 - \pi_{k-1})(1 - p_{01})} \triangleq \Phi(\pi_{k-1}, z_k, H_k).$$

References

[1] Jiruska P, Powell AD, Deans JK, Jefferys JG. Effects of direct brain stimulation depend on seizure dynamics. Epilepsia 2010;51:S93–7.
[2] Jobst BC, Darcey TM, Thadani VM, Roberts DW. Brain stimulation for the treatment of epilepsy. Epilepsia 2010;51:S88–92.
[3] Morrell M. Brain stimulation for epilepsy: can scheduled or responsive neurostimulation stop seizures? Curr Opin Neurol 2006;19:164–8.
[4] Gotman J, Gloor P. Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. Electroencephalogr Clin Neurophysiol 1976;41: 513–29.
[5] Gotman J. Automatic recognition of epileptic seizures in the EEG. Electroencephalogr Clin Neurophysiol 1982;54:530–40.
[6] Osorio I, Frei MG, Wilkinson SB. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. Epilepsia 1998;39:615–27.
[7] Osorio I, Frei MG, Giftakis J, et al. Performance reassessment of a real-time seizure-detection algorithm on long ECoG series. Epilepsia 2002;43:1522–35.
[8] Haas SM, Frei M, Osorio I. Strategies for adapting automated seizure detection algorithms. Med Eng Phys 2007;29:895–909.
[9] Osorio I, Frei MG. Real-time detection, quantification, warning, and control of epileptic seizures: the foundations for a scientific epileptology. Epilepsy Behav 2009;16:391–6.
[10] Jerger KK, Netoff TI, Francis JT, et al. Early seizure detection. J Clin Neurophysiol 2001;18:259–68.
[11] Esteller R, Echauz J, D'Alessandro M, et al. Continuous energy variation during the seizure cycle: towards an on-line accumulated energy. Clin Neurophysiol 2005;116:517–26.
[12] Gardner AB, Worrell GA, Marsh E, Dlugos D, Litt B. Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings. Clin Neurophysiol 2007;118:1134–43.
[13] Khan YU, Gotman J. Wavelet based automatic seizure detection in intracerebral electroencephalogram. Clin Neurophysiol 2003;114:898–908.
[14] Grewal S, Gotman J. An automatic warning system for epileptic seizures recorded on intracerebral EEGs. Clin Neurophysiol 2005;116:2460–72.
[15] Saab ME, Gotman J. A system to detect the onset of epileptic seizures in scalp EEG. Clin Neurophysiol 2005;116:427–42.
[16] Alkan A, Koklukaya E, Subasi A. Automatic seizure detection in EEG using logistic regression and artificial neural network. J Neurosci Methods 2005;148:167–76.
[17] Van Putten MJ, Kind T, Visser F, Lagerburg V. Detecting temporal lobe seizures from scalp EEG recordings: a comparison of various features. Clin Neurophysiol 2005;116:2480–9.
[18] Wilson SB. A neural network method for automatic and incremental learning applied to patient-dependent seizure detection. Clin Neurophysiol 2005;116:1785–95.
[19] Wilson SB. Algorithm architectures for patient dependent seizure detection. Clin Neurophysiol 2006;117:1204–16.

[20] Srinivasan V, Eswaran C, Sriraam N. Approximate entropy-based epileptic EEG detection using artificial neural networks. IEEE Trans Inf Technol Biomed 2007;11:288–95.

[21] Chan AM, Sun FT, Boto EH, Wingeier BM. Automated seizure onset detection for accurate onset time determination in intracranial EEG. Clin Neurophysiol 2008;119:2687–96.

[22] Aarabi A, Fazel-Rezai R, Aghakhani Y. A fuzzy rule-based system for epileptic seizure detection in intracranial EEG. Clin Neurophysiol 2009;120:1648–57.

[23] Tito M, Cabrerizo M, Ayala M, Jayakar P, Adjouadi M. Seizure detection: an assessment of time- and frequency-based features in a unified two-dimensional decisional space using nonlinear decision functions. J Clin Neurophysiol 2009;26:381–91.

[24] Minasyan GR, Chatten JB, Chatten MJ, Harner RN. Patient-specific early seizure detection from scalp electroencephalogram. J Clin Neurophysiol 2010;27:163–78.

[25] Temko A, Thomas E, Marnane W, Lightbody G, Boylan G. EEG-based neonatal seizure detection with support vector machines. Clin Neurophysiol 2011;122:464–73.

[26] Shoeb A, Edwards H, Connolly J, Bourgeois B, Treves ST, Guttag J. Patient-specific seizure onset detection. Epilepsy Behav 2004;5:483–98.

[27] Ghosh-Dastidar S, Adeli H, Dadmehr N. Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE Trans Biomed Eng 2007;54:1545–51.

[28] Ghosh-Dastidar S, Adeli H, Dadmehr N. Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. IEEE Trans Biomed Eng 2008;55:512–8.

[29] Meier R, Dittrich H, Schulze-Bonhage A, Aertsen A. Detecting epileptic seizures in long-term human EEG: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns. J Clin Neurophysiol 2008;25:119–31.

[30] Abibullaev B, Kim MS, Seo HD. Seizure detection in temporal lobe epileptic EEGs using the best basis wavelet functions. J Med Syst 2010;34:755–65.

[31] Guo L, Rivero D, Dorado J, Rabuñal JR, Pazos A. Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. J Neurosci Methods 2010;191:101–9.

[32] Guo L, Rivero D, Pazos A. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. J Neurosci Methods 2010;193:156–63.

[33] Shoeb A, Pang T, Guttag J, Schachter S. Non-invasive computerized system for automatically initiating vagus nerve stimulation following patient-specific detection of seizures or epileptiform discharges. Int J Neural Syst 2009;19:157–72.

[34] Lee HC, van Drongelen W, McGee AB, Frim DM, Kohrman MH. Comparison of seizure detection algorithms in continuously monitored pediatric patients. J Clin Neurophysiol 2007;24:137–46.

[35] Shiryaev AN. On optimum methods in quickest detection problems. Theory Probab Appl 1963;8:22–46.

[36] Poor HV, Hadjiliadis O. Quickest detection. Cambridge, UK: Cambridge Univ. Press; 2008.

[37] Sarma SV, Santaniello S. Quickest detection of state-transition in point processes: application to neuronal activity. Proceedings, 18th IFAC World Conference; Milan, Italy, 18; August 29–September 2 2011. p. 10021–6.

[38] Risinger MV, Engel JJ, Van Ness PC, Henry TR, Crandall PH. Ictal localization of temporal seizures with scalp-sphenoidal recordings. Neurology 1989;39:1288–93.

[39] Bernhardt BC, Chen Z, He Y, Evans AC, Bernasconi N. Graph-theoretical analysis reveals disrupted small-world organization of cortical thickness correlation networks in temporal lobe epilepsy. Cereb Cortex 2011;21:2147–57.

[40] Müller M, Baier G, Galka A, Stephani U, Muhle H. Detection and characterization of changes of the correlation structure in multivariate time series. Phys Rev E 2005;71:046116.

[41] Rummel C, Baier G, Müller M. The influence of static correlations on multivariate correlation analysis of the EEG. J Neurosci Methods 2007;166:138–57.

[42] Rummel C, Müller M, Baier G, Amor F, Schindler K. Analyzing spatio-temporal patterns of genuine cross-correlations. J Neurosci Methods 2010;191:94–100.

[43] Rummel C, Abela E, Müller M, et al. Uniform approach to linear and nonlinear interrelation patterns in multivariate time series. Phys Rev E 2011;83:066215.

[44] Schindler KA, Leung H, Elger CE, Lehnertz K. Assessing seizure dynamics by analyzing the correlation structure of multichannel intracranial EEG. Brain 2007;130:65–77.

[45] Schindler KA, Bialonski S, Horstmann MT, Elger CE, Lehnertz K. Evolving functional network properties and synchronizability during human epileptic seizures. Chaos 2008;18:033119.

[46] Schindler KA, Amor F, Gast H, et al. Peri-ictal correlation dynamics of high-frequency (80–200 Hz) intracranial EEG. Epilepsy Res 2010;89:72–81.

[47] Wilke C, Worrell G, He B. Graph analysis of epileptogenic networks in human partial epilepsy. Epilepsia 2011;52:84–93.

[48] Kramer MA, Kolaczyk ED, Kirsh HE. Emergent network topology at seizure onset in humans. Epilepsy Res 2008;79:173–86.

[49] Newman MEJ. Networks: an introduction. Cambridge, UK: Oxford Univ. Press; 2010.

[50] Oppenheim AV, Schafer RW. Discrete-time signal processing. Upper Saddle River, NJ: Prentice–Hall; 1999.

[51] Iasemidis LD, Shiau D-S, Sackellares JC, Pardalos PM, Prasad A. Dynamical resetting of the human brain at epileptic seizures: application of nonlinear dynamics and global optimization techniques. IEEE Trans Biomed Eng 2004;51:493–506.

[52] Liu CC, Pardalos PM, Chaovalitwongse WA, et al. Quantitative complexity analysis in multi-channel intracranial EEG recordings form epilepsy brains. J Comb Optim 2008;15:276–86.

[53] Golub HG, Van Loan CF. Matrix computations. 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press; 1996.

[54] Elliott RJ, Aggoun L, Moore JB. Hidden Markov models: estimation and control. New York, NY: Springer; 1995.

[55] Brown EN, Barbieri R, Eden UT, Frank LM. Likelihood methods for neural data analysis. In: Feng J, editor. Computational neuroscience: a comprehensive approach. London, UK: CRC Press; 2003 pp. 253–86.

[56] McCullagh P, Nelder JA. Generalized linear models. 2nd ed. Boca Raton, FL: CRC Press; 1990.

[57] Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control 1974;19:716–23.

[58] Berger JO. Statistical decision theory and Bayesian analysis. 2nd ed. New York, NY: Springer; 1985.

[59] Bertsekas DP. Dynamic programming and optimal control. 3rd ed. Belmont, MA: Athena Scientific; 2005.

[60] Li D, Zhou W, Drury I, Savit R. Linear and nonlinear measures and seizure anticipation in temporal lobe epilepsy. J Comput Neurosci 2003;15:335–45.

[61] Allen CN. Circadian rhythms, diet, and neuronal excitability. Epilepsia 2008;49 (Suppl. 8):124–6.

[62] Miller P, Katz DB. Stochastic transitions between neural states in taste processing and decision-making. J Neurosci 2010;30:2559–70.

[63] Kemere C, Santhanam G, Yu BM, et al. Detecting neural-state transitions using hidden Markov models for motor cortical prostheses. J Neurophysiol 2008;100:2441–52.

[64] Coleman TP, Sarma SV. A computationally efficient method for nonparametric modeling neural spiking activity with point processes. Neural Comp 2010;22:2002–30.

[65] Sarma SV, Eden UT, Cheng ML, Williams ZM, Hu R, Eskandar EN, Brown EN. Using point process models to compare neuronal activity in subthalamic nucleus of Parkinson's patients and a healthy primate. IEEE Trans Biomed Eng 2010;57:1297–305.