

## IDENTIFYING PRESEIZURE STATE IN INTRACRANIAL EEG DATA USING DIFFUSION KERNELS

DOMINIQUE DUNCAN

101 AKW, 51 Prospect St. New Haven, CT 06511, USA

RONEN TALMON

103 AKW, 51 Prospect St. New Haven, CT 06511, USA

HITTEN P. ZAVERI

716 LLCI, 15 York St. New Haven, CT 06520, USA

RONALD R. COIFMAN

108A AKW, 51 Prospect St. New Haven, CT 06511, USA

**ABSTRACT.** The goal of this study is to identify pre-seizure changes in intracranial EEG (icEEG). A novel approach based on the recently developed diffusion map framework, which is considered to be one of the leading manifold learning methods, is proposed. Diffusion mapping provides dimensionality reduction of the data as well as pattern recognition that can be used to distinguish different states of the patient, for example, interictal and pre-seizure. A new algorithm, which is an extension of diffusion maps, is developed to construct coordinates that generate efficient geometric representations of the complex structures in the icEEG data. In addition, this method is adapted to the icEEG data and enables the extraction of the underlying brain activity.

The algorithm is tested on icEEG data recorded from several electrode contacts from a patient being evaluated for possible epilepsy surgery at the Yale-New Haven Hospital. Numerical results show that the proposed approach provides a distinction between interictal and pre-seizure states.

**1. Introduction.** Approximately 50 million people worldwide suffer from epilepsy, including 2.7 million people in the United States, making it one of the most common neurological disorders [7]. Seizures are successfully controlled in 64% of patients, while the remaining 36% have pharmaco-resistant epilepsy [9]. Epilepsy surgery is an option for some of the patients with pharmaco-resistant epilepsy if their seizures can be localized to a focal area of the brain.

It is important to find a reliable method to predict seizures so that a patient can be warned at least a few minutes prior to the seizure and take effective precautions. Finding an accurate predictor of seizures has become a major focus of research during the last few decades [5]. Seizure prediction algorithms can be used to improve devices to control seizures. The Neuropace device, for example, which is placed within the skull, contains a programmable feature detector that can be employed to stimulate on detection of target features [15]. However, it is still important to

---

2010 *Mathematics Subject Classification.* 68T10, 65F15, 92B99, 92C20, 00A69.

*Key words and phrases.* Intracranial EEG, epilepsy, seizure prediction, diffusion maps, nonlinear independent component analysis.

improve these feature detectors and find ways to predict seizures with sufficient sensitivity and specificity.

In our ongoing research, we have investigated various methods to analyze icEEG data, such as coherence, mutual information, and approximate entropy [4] to analyze resting state icEEG activity. Seizure prediction involves a wide variety of methods, including linear methods, such as autoregressive models, as well as nonlinear methods, such as estimations of the largest Lyapunov exponent and correlation [13]. Other approaches involve the comparison of interictal and preictal epochs [1]. Often the main problem in many seizure prediction studies is that the analysis is performed on a small amount of data from few patients.

In EEG data, we assume that the measurements are controlled by underlying processes that represent brain activity. We would like to recover these underlying brain processes to distinguish different brain states, such as interictal and pre-seizure states. Diffusion maps [3] have been a useful tool in reducing the dimensionality of the data as well as providing a measure for pattern recognition and feature detection. Since diffusion mapping may detect abnormal behavior in the data, it can be used to determine changes in brain states. However, diffusion maps assume access to the underlying process that it aims to reveal. In EEG data, the relationship between samples of the data and the underlying activity may be stochastic, and the data are assumed to be noisy. Hence diffusion mapping is not the most suitable approach to use with our icEEG data. A recently developed algorithm, which is an extension of diffusion maps, may be more applicable in our case [16, 17]. The new algorithm assumes a stochastic mapping between the underlying processes and the measurements; the mapping is inverted and a kernel is used to recover the underlying activity [16]. Thus, the proposed algorithm is more appropriate than diffusion maps for our data.

In this paper, we propose an algorithm that relied on [16] for extracting the underlying brain activity from the icEEG data. The algorithm is an extension of diffusion maps and uses local principal components analysis (PCA). PCA is another dimensionality reduction method. In PCA, the goal is to compute the most meaningful basis to re-express a large and noisy data set. This new basis can reveal hidden patterns and structure in the data as well as remove the noise. An orthogonal linear transformation converts the data to a new coordinate system. The greatest variance in the data is represented by the first coordinate or the first principal component. In EEG, the data may be generated by a nonlinear combination of sources, the underlying processes, so that more sophisticated methods are required to represent the data. An important difference between the proposed algorithm and PCA is the use of nonlinear locality in the extension as opposed to PCA, which retains the linear global information of the data. For the icEEG data, we perform PCA on local regions of data and then integrate the local information using a kernel and obtain a single model. We use a data-driven adapted distance between samples of icEEG recordings to approximate the Euclidean distance between the underlying processes from the noisy EEG.

## 2. Methods.

### 2.1. Intracranial EEG. <sup>1</sup>

---

<sup>1</sup>Intracranial depth, subdural strip, and subdural grid electrodes were placed as required for the patient. Subdural strip and grid electrode contacts were recessed platinum disk contacts with 2.3 mm exposed surface and 1 cm center-to-center separation (Ad-Tech, Racine, WI, U.S.A.).

Intracranial EEG (icEEG) data were collected from a patient with localization related epilepsy who was undergoing presurgical evaluation at the Yale-New Haven Hospital. A total of 182 electrode contacts were used during the monitoring. The seizure onset area was located on the right occipital lobe. In this paper, we focus on the 3 electrode contacts, which overlaid the seizure onset area. Figure 1 shows the location of the 3 electrode contacts which were studied.

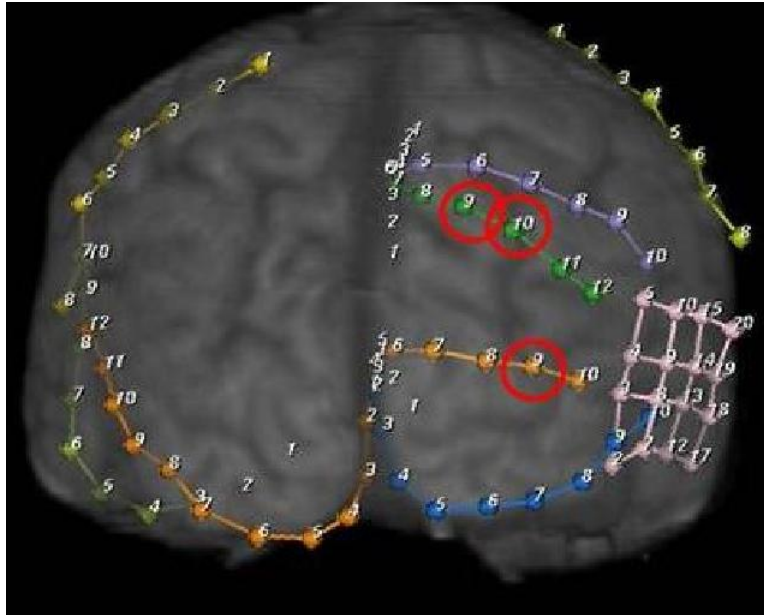


FIGURE 1. The seizure onset area is located in the right occipital lobe. The three electrode contacts overlying the seizure onset area which are used in this study have been circled.

We studied 6 icEEG epochs, each of which corresponds to a seizure experienced by the patient over the course of the icEEG monitoring. Figure 2 shows an example of approximately 16 minutes of icEEG data from one electrode contact that includes a seizure. The seizure start time is marked with a red line. Figure 3 shows the corresponding spectrogram, where the horizontal axis represents time (each point is one time window). We considered the delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-25 Hz), and gamma frequency bands (25-100 Hz).

For each case, we considered 5 minutes of data immediately preceding the seizure and 5 minutes of data from 35 minutes before the seizure, which we considered to be the interictal state. An example is shown in Figure 4. The data were collected

---

Intracranial electrode contacts were located after placement using a computer program (BioImage Suite) developed at Yale University [11] and a procedure which has been described [6]. Briefly, intracranial electrode contacts were located from a postimplantation computed tomography (CT) image. The postimplantation CT was then coregistered first with a postimplantation magnetic resonance (MR) image using a linear coregistration procedure and then with a preimplantation MR image using both a linear and a nonlinear coregistration procedure. This procedure allowed us to visualize intracranial electrode contacts on preimplantation MR images. The icEEG were collected with clinical icEEG acquisition equipment (Natus Medical Inc./Bio-Logic Systems Corp., San Carlos, CA). The signals were sampled at 256 Hz with 16 bit A/D conversion.

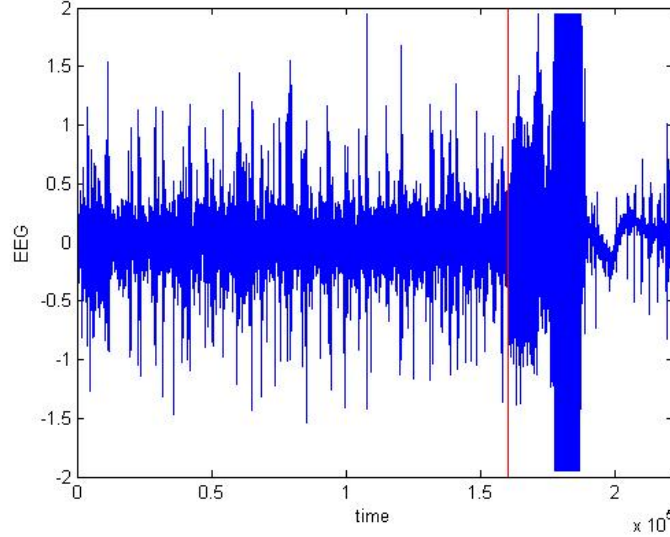


FIGURE 2. The icEEG from one electrode contact for approximately 16 minutes including one seizure where the seizure start time is marked with a red line. The horizontal axis represents time (in samples).

at 256 samples per second, so for each case, we analyzed 10 minutes or 153,600 samples. The horizontal axis represents time (in samples). From the raw icEEG data for this patient, it is typically unclear to determine when a seizure began until the actual start time of the seizure.

**2.2. Data features and metric.** Let  $y(n)$  be the icEEG data from a single electrode contact where  $n$  is the time index; we describe the following procedure for data from merely one electrode contact for simplicity. Vectors and matrices will be denoted in lowercase and uppercase bold, respectively. Applying the Short Time Fourier Transform (STFT) on  $y(n)$  yields

$$y_{m,k} = \sum_n y(n) \tilde{\phi}_{m,k}(n), \quad (1)$$

where

$$\tilde{\phi}_{m,k}(n) = \tilde{\phi}(n - mL) e^{j \frac{2\pi}{N} k(n - mL)}, \quad (2)$$

and  $m$  is the time frame index,  $k$  is the frequency band index,  $\tilde{\phi}(n)$  is a Hanning window of length  $N$ , and  $L$  is the discrete-time shift [12]. The Hanning window is defined as

$$\tilde{\phi}(n) = \frac{1}{2} \left( 1 - \cos \left( \frac{2\pi n}{N-1} \right) \right). \quad (3)$$

The data are split into segments overlapped by 50%, or  $L = N/2$ , each of which are windowed and then Fourier transformed to produce an estimate of the short-term frequency content of the signal. Let

$$S_y(m, n) = |y_{m,n}|^2 \quad (4)$$

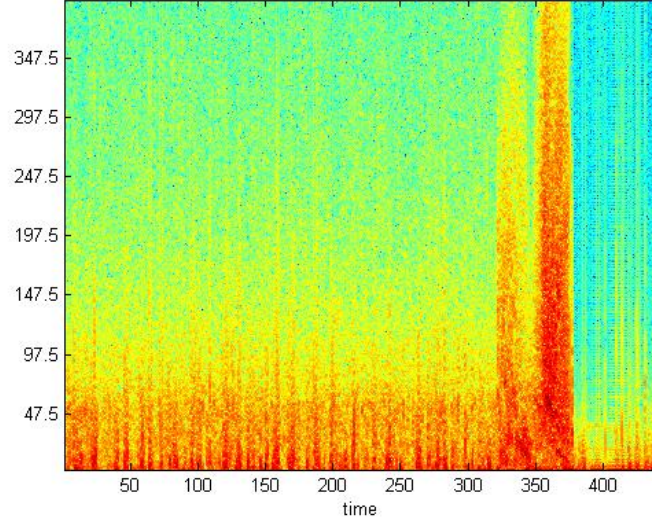


FIGURE 3. The spectrogram of the above segment of data (approximately 16 minutes) that includes a seizure, showing the delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-25 Hz), and gamma frequency bands (25-100 Hz). The vertical axis represents frequency ranging from 0-100 Hz, and the horizontal axis represents time (in analysis windows).

be the squared amplitude of the STFT. We collect a few frequency bands for each time frame into vectors defined as

$$\mathbf{s}_y(m) = [\dots, S_y(m, n), \dots]^T, n \in \mathcal{K}, \quad (5)$$

where  $\mathcal{K}$  is the collection of empirically relevant frequency bands. The STFT is commonly used in time series analysis because it allows us to see the changes of the spectral composition over time. In EEG, it may enable us to identify differences in the data closer to seizures as opposed to interictal states. We call  $\{\mathbf{s}_y(m)\}$  the feature vectors.

We assume that the high dimensional features  $\mathbf{s}_y(m)$  are controlled by the low dimensional processes that represent brain activity. The factors that describe a preseizure state are considered to be a collection of  $d$  stochastic processes  $\boldsymbol{\theta}_m = (\theta_m^1, \dots, \theta_m^d)$ , where  $m$  is the time frame index, such that each of these processes satisfies a stochastic differential equation that can be described as

$$\theta_{m+1}^i = \theta_m^i + a^i(\theta_m^i) + w_m^i, \quad (6)$$

where  $i = 1, \dots, d$ . It is assumed that the drift terms  $(a^1, \dots, a^d)$  are unknown and  $(w_m^1, \dots, w_m^d)$  are  $d$  independent white Gaussian noises.  $\mathbf{s}_y(m)$  is a random variable whose statistics can be expressed as a nonlinear stochastic function of the underlying factors. Thus, the goal is to recover these underlying factors, which represent the preseizure state, from the features,  $\mathbf{s}_y(m)$ .

We propose the following method to reduce the dimension of the data while keeping the information we need to distinguish preseizure from interictal states.

We compare feature vectors for the 5 minutes of data by calculating the STFT of the windows of data. Combining the vectors,  $\mathbf{s}_y(m)$ , and plotting the outcome shows that the feature vectors (Figure 5) do not exhibit significant variation between interictal and pre-seizure state for the 6 separate seizure episodes. The lower frequency bands are most often used for detecting different states in patients using EEG; we consider the frequency bands up to 100 Hz for our analysis. These are used as our feature vectors for each point in time (each window of almost 4 seconds).

Given a feature vector  $\mathbf{s}_y(m)$ , we compute the local covariance matrix in a time interval of length  $J$ :

$$\Sigma_m = \frac{1}{J} \sum_{m'=m-J+1}^m (\mathbf{s}_y(m') - \boldsymbol{\mu}_m)(\mathbf{s}_y(m') - \boldsymbol{\mu}_m)^T, \quad (7)$$

where  $\boldsymbol{\mu}_m$  is the empirical local mean of the feature vectors in the interval.

We define a nonsymmetric distance using the covariance matrices  $\Sigma_m$ s:

$$a_{\Sigma}^2(m, m') = (\mathbf{s}_y(m) - \mathbf{s}_y(m'))^T \Sigma_m^{-1} (\mathbf{s}_y(m) - \mathbf{s}_y(m')) \quad (8)$$

and a symmetric distance

$$d_{\Sigma}^2(m, m') = (1/2)(a_{\Sigma}^2(m, m') + a_{\Sigma}^2(m', m)). \quad (9)$$

The distance in (9) is termed the Mahalanobis distance, and it is shown in [14] that it approximates the Euclidean distance between the underlying factors. This distance allows us in Section 2.3 to recover the underlying factors in the icEEG data via eigendecomposition of an appropriate Laplace operator (kernel).

**2.3. Kernel and embedding.** Consider a collection  $\{\mathbf{s}_y(m)\}$  of  $N_s$  feature vectors. Let  $\mathcal{R}$  be a subset of reference feature vectors of size  $\bar{N}_s$  and  $\bar{N}_s < N_s$ .  $\mathcal{R}$  is obtained by selecting an arbitrary subset of measurements to enable a more efficient computation of the data.

A kernel is used to compare the underlying factors, whose  $(m, m')$ th element is

$$W_{\mathcal{R}}^{m, m'} = \exp \left\{ -\frac{d_{\Sigma}^2(m, m')}{\epsilon} \right\} \quad (10)$$

where  $\epsilon$  is the kernel scale set according to the Mahalanobis distance defined above. This kernel,  $\mathbf{W}_{\mathcal{R}}$ , depends on the reference set and defines the local geometries of the graph [3].

$\mathbf{W}_{\mathcal{R}}$  is normalized by a diagonal density matrix,

$$\mathbf{D}^{m, m} = \sum_{m' \in \mathcal{R}} \mathbf{W}_{\mathcal{R}}^{m, m'}, \quad (11)$$

which enables us to view the sampling as uniform. This normalized matrix is denoted by

$$\tilde{\mathbf{W}}_{\mathcal{R}} = D^{-1/2} \mathbf{W}_{\mathcal{R}} D^{1/2}. \quad (12)$$

We can further normalize according to [8]. In order to obtain the representation of the training set, we perform an eigendecomposition to handle the nonuniform sampling of the data and acquire the eigenvalues,  $\boldsymbol{\lambda}_j$ , and eigenvectors,  $\boldsymbol{\varphi}_j$ . The eigenvalues of the diagonal matrix are all non-negative and sorted in decreasing order. The Laplace-Beltrami operator is given by  $\mathbf{I} - \tilde{\mathbf{W}}_{\mathcal{R}}$ , which shares the same eigenvectors as  $\tilde{\mathbf{W}}_{\mathcal{R}}$ .

In order to incorporate samples from the entire set, we form an  $N_s \times \bar{N}_s$  non-symmetric affinity matrix  $\mathbf{A}$  whose  $(m, m')$ th element is defined as

$$A^{mm'} = \exp\left(-\frac{a_{\Sigma}^2(m, m')}{\epsilon}\right), \quad (13)$$

where  $\epsilon > 0$  is the kernel scale,  $m' \in \mathcal{R}$ , and  $m = 1, \dots, N_s$ .

We define the diagonal matrix  $\tilde{\mathbf{D}}$  and normalize  $\mathbf{A}$ , similar to the previous normalization, to form

$$\tilde{\mathbf{A}} = \tilde{\mathbf{D}}^{-1} \mathbf{A}. \quad (14)$$

Next we define a kernel on the entire set:

$$\mathbf{W} = \tilde{\mathbf{A}} \tilde{\mathbf{A}}^T. \quad (15)$$

Then we extend the eigenvectors of  $\mathbf{W}_{\mathcal{R}}$  and use the diagonal density to get the eigenvalues of  $\mathbf{W}$  without computing the eigendecomposition of  $\mathbf{W}$ . It can be shown by [8] that

$$\tilde{\mathbf{W}}_{\mathcal{R}} = \tilde{\mathbf{A}}^T \tilde{\mathbf{A}}. \quad (16)$$

The eigenvalues and eigenvectors of  $\tilde{\mathbf{W}}_{\mathcal{R}}$  are determined, and it is known that  $\mathbf{W}$  converges to a diffusion operator [8, 17].

Proposition 3.1 from [8] states that the eigenvectors  $\psi_j$  of  $\mathbf{W}$  corresponding to nonzero eigenvalues  $\lambda_j > 0$  are

$$\psi_j = \frac{1}{\lambda_j^{1/2}} \mathbf{A} \varphi_j. \quad (17)$$

Using  $\mathbf{W}_{\mathcal{R}}$  based on our reference set is useful, because it is smaller than the size of the kernel based on the whole data set, and using the entire data set would be computationally intensive. In practice, it is difficult to handle the icEEG data, because it is so large, so using a small matrix based on the reference set is efficient. In the algorithm, we use  $\mathbf{W}_{\mathcal{R}}$  to denote the first stage with the reference feature vectors as opposed to  $\mathbf{W}$  for the second stage. Initially, we use the reference feature vectors to perform our analysis using the kernel based on the Mahalanobis distance (9) and apply an eigendecomposition. Then we use (17) to acquire a representation of the entire set and finally define a mapping that allows us to recover the underlying brain activity in the icEEG data.

Based on the eigendecomposition, we define an  $\ell$ -dimensional embedding of each feature vector as the diffusion mapping:

$$\mathbf{s}_y(m) \mapsto [\lambda_1 \psi_1(m), \lambda_2 \psi_2(m), \dots, \lambda_{\ell} \psi_{\ell}(m)]^T, \quad (18)$$

and  $\ell$  is related to the intrinsic dimensionality of the data for each  $m = 1, \dots, N_s$ .

The eigenvectors of  $\mathbf{W}$  corresponding to the largest eigenvalues provide a parameterization of the features. Thus, we reduce the data from a large dimensional space of noisy observations to a small dimensional space, which captures the features of pre-seizure activity. In our case, we set  $\ell = 3$ , so our embedding is in a 3-dimensional space. Our empirical example approximates heat flow on a manifold; we have high dimensional complex data, and we want to understand the intrinsic coordinates of the manifold. The proposed algorithm is useful for distinguishing different states (interictal and pre-seizure) in the icEEG data.

**3. Results.** Figure 5 shows the spectrograms for the two segments of interictal and pre-seizure icEEG data. The horizontal axis represents time frames, and the vertical axis represents frequency ranging from 0-100 Hz. The spectrogram is computed by using the Short Time Fourier Transform (STFT) in (4) with a window length of 1,000 points overlapping by 50%. For EEG data, lower frequency bands reveal the most information about the state of a patient [2]. In our analysis, we considered the delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-25 Hz), and gamma (25-100 Hz) frequency bands.

After combining the resulting spectrograms for the two segments (five minutes immediately preceding the seizure and five minutes from 35 minutes before the seizure), we apply our proposed algorithm. Figure 6 depicts a 3D scatter plot of the embedding obtained using the three leading eigenvectors from (18). Each point in the figure represents a time frame defined in (5). The colors represent time for the 10 minutes of data, with the red points corresponding to the data directly before the seizure. We find that this method shows a clear separation of the pre-seizure points from the interictal state points, which tend to be scattered around the origin.

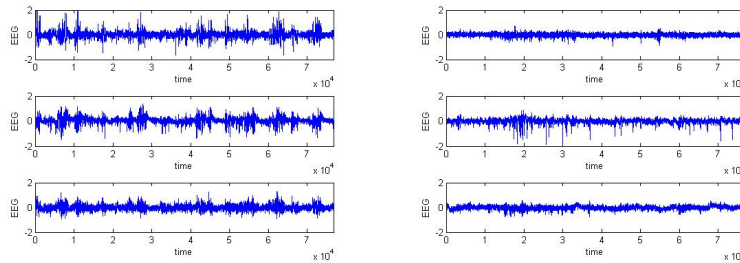


FIGURE 4. The raw icEEG data recorded from the 3 electrode contacts in the seizure onset area for the interictal (left) and pre-seizure (right) periods; the horizontal axis represents time (in samples).

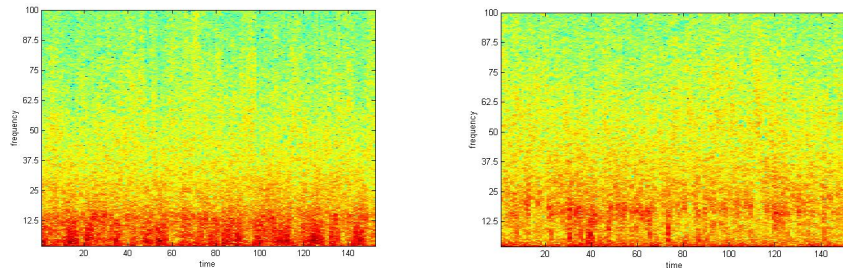


FIGURE 5. The spectrograms from the interictal (left) and pre-seizure (right) periods; the vertical axis represents frequency ranging from 0-100 Hz, and the horizontal axis represents time (in analysis windows).



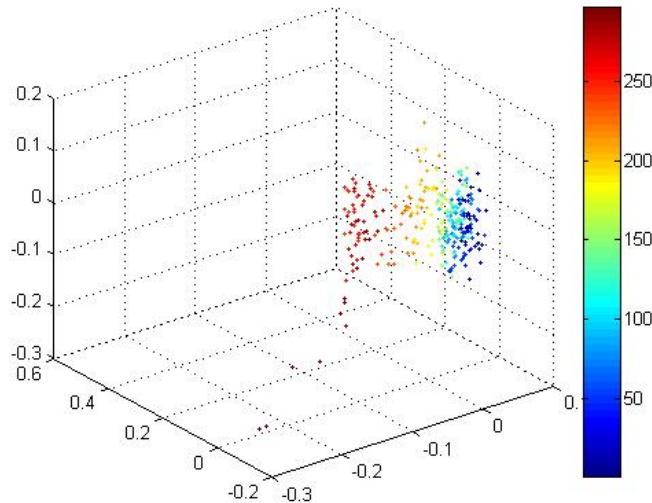


FIGURE 6. The embedding using 3 eigenvectors; the colors represent time, with the red points closest to seizure onset.

From the three dimensional plot, it is quite evident that there is a clear separation between the points in the preseizure state and the points in the interictal state. The red points on the graph are the points closest in time to the seizure onset, and they are separated from the rest of the cloud of points that are near the origin. For the data analysis of this patient, we see that for each of the times prior to the seizure occurrences, there is no apparent difference between the interictal states and the preseizure states in the raw data or in the spectrograms of the data. On the other hand, there is a clear separation between the two states from the embeddings that arise from an application of our method. This experiment was repeated for all six seizures that the patient experienced while being monitored, and we obtained similar and consistent results.

To verify that the separation did not occur from merely the time separation of the two data segments, two separate time segments of interictal state data were compared, as seen in Figure 7. In the previous case, we use one segment of preseizure data compared to one segment of interictal data from 30 minutes before that preseizure segment. We compare that segment of interictal data to another segment of interictal data from 30 minutes prior to it and also computed the spectrograms using (4), as seen in Figure 8. It was seen that there was no discernible separation of these two interictal state data segments using (18) in the embedding in Figure 9. An analysis of other interictal state and preseizure data for this patient showed the same type of separation as shown in Figure 6, whereas other embeddings using two separate interictal periods showed no separation.

**4. Discussion.** Based on this initial study, we have found our algorithm to be a potentially useful method for characterizing the preseizure state in intracranial EEG recordings from the seizure onset area. Our algorithm may identify the underlying processes that represent the brain activity. In our analysis, in which we examined

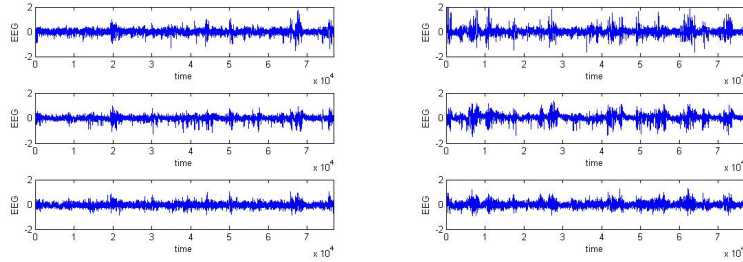


FIGURE 7. The raw icEEG data recorded from the 3 electrode contacts in the seizure onset area for two interictal periods 30 minutes apart from each other. The horizontal axis represents time (in samples).

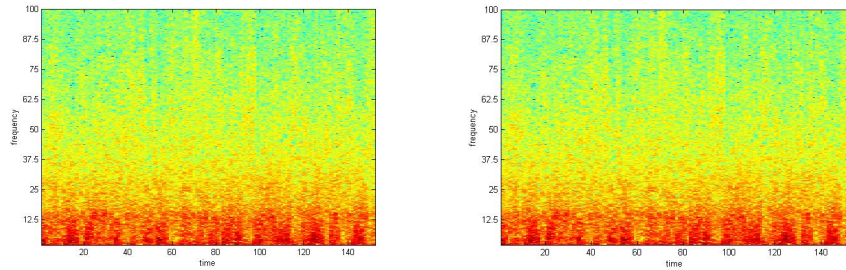


FIGURE 8. The spectrograms from two interictal periods; the vertical axis represents frequency ranging from 0-100 Hz, and the horizontal axis represents time (in analysis windows).

the pre-seizure icEEG data from a patient who experienced 6 seizures, we found that our method resulted in an embedding that showed a distinct separation between interictal and pre-seizure states. Using the same method while comparing two separate interictal segments that were also 30 minutes apart resulted in no separation in the embedding. Thus it appears that the method distinguished a pre-seizure state from an interictal state. While this analysis of the data is preliminary, it is planned to develop this approach to construct a method for predicting seizures. Our goal, relying on these results, is to define a threshold that separates the interictal state from the pre-seizure state in a wide range of cases and develop an automatic method to predict seizures.

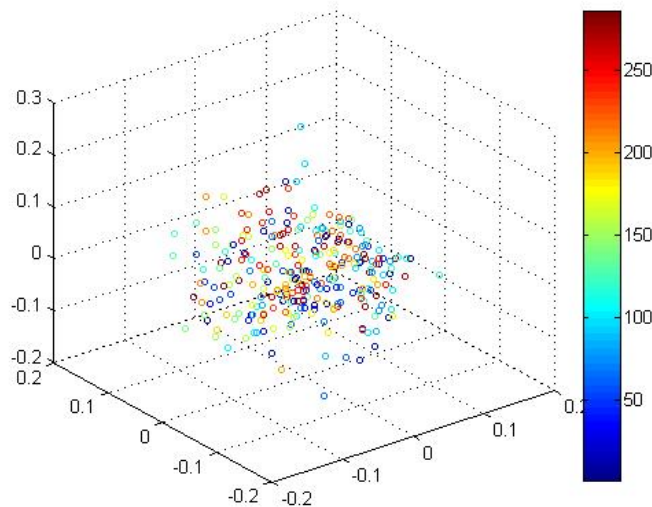


FIGURE 9. The embedding of the two interictal state segments of data using 3 eigenvectors; the colors represent time.

#### REFERENCES

- [1] R. G. Andrzejak, F. Mormann, R. Kreuz, C. Rieke, A. Kraskov, C. E. Elger and K. Lehnertz, *Testing the null hypothesis of the nonexistence of a preseizure state*, Phys. Rev. E, **67** (2003), 10901.
- [2] J. Britz, D. Van De Ville and C. M. Michel, *BOLD correlates of EEG topography reveal rapid resting-state network dynamics*, NeuroImage, **52** (2010), 1162–1170.
- [3] R. R. Coifman and S. Lafon, *Diffusion maps*, Appl. Comp. Harm. Anal., **21** (2006), 5–30.
- [4] D. Duncan, R. B. Duckrow, R. R. Coifman and H. P. Zaveri, *Intracranial EEG Evaluation of a Resting State Network*, Challenges of Modern Technology, **1** (2010), 27–29.
- [5] M. G. Frei, H. P. Zaveri, S. Arthurs, G. K. Bergey, C. C. Jouny, K. Lehnertz, J. Gotman, I. Osorio, T. I. Netoff, W. J. Freeman, J. Jefferys, G. Worrell, M. Le Van Quyen, S. J. Schiff and F. Mormann, *Controversies in epilepsy: Debates held during the fourth international workshop on seizure prediction*, Epilepsy and Behavior, **19** (2010), 4–16.
- [6] I. I. Goncharova, H. P. Zaveri, R. B. Duckrow, E. J. Novotny and S. S. Spencer, *Spatial distribution of intracranially recorded spikes in medial and lateral temporal epilepsies*, Epilepsia, **50** (2009), 2575–85.
- [7] W. A. Hauser, J. F. Annegers and W. A. Rocca, *Descriptive epidemiology of epilepsy: Contributions of population-based studies from Rochester, Minnesota*, Mayo. Clin. Proc., **71** (1996), 576–586.
- [8] D. Kushnir, A. Haddad and R. R. Coifman, *Anisotropic diffusion on sub-manifolds with application to earth structure classification*, Appl. Comp. Harm. Anal., **32** (2012), 280–294.
- [9] P. Kwan and M. J. Brodie, *Early identification of refractory epilepsy*, N. Engl. J. Med., **342** (2000), 314–319.
- [10] F. Mormann, R. G. Andrzejak, C. E. Elger and K. Lehnertz, *Seizure prediction: the long and winding road*, Brain, **130** (2007), 314–333.
- [11] X. Papademetris, A. P. Jackowski, R. T. Schultz, L. H. Staib and J. S. Duncan, *Integrated intensity and point-feature non-rigid registration*, in “Medical Image Computing and Computer-Assisted Intervention” (Eds. C. Barillot, D. Haynor and P. Hellier) Saint-Malo: Springer, (2004), 763–770.
- [12] M. R. Portnoff, *Time-frequency representation of digital signals and systems based on short-time Fourier analysis*, IEEE Trans. Signal Process, **ASSP-28**, 55–69.

- [13] A. Schulze-Bonhage, H. Feldwisch-Drentrup and M. Ihle, *Epilepsy and behavior*, Elsevier, **22** (2011), S88–S93.
- [14] A. Singer and R. R. Coifman, *Non-linear independent component analysis with diffusion maps*, Appl. Comp. Harm. Anal., **25** (2008), 226–239.
- [15] E. Susman, *Brain stimulation reduces seizures in refractory adult epilepsy*, Neurology Today, **9** (2009), 22–23.
- [16] R. Talmon and R. R. Coifman, *Differential stochastic sensing: intrinsic modeling of random time series with applications to nonlinear tracking*, submitted to PNAS, (2012).
- [17] R. Talmon, D. Kushnir, R. R. Coifman, I. Cohen and S. Gannot, *Parametrization of linear systems using diffusion kernels*, IEEE Trans. Signal Process, **60** (2012).

Received July 05, 2012; Accepted December 22, 2012.

*E-mail address:* [dominique.duncan@yale.edu](mailto:dominique.duncan@yale.edu)

*E-mail address:* [ronen.talmon@yale.edu](mailto:ronen.talmon@yale.edu)

*E-mail address:* [hitten.zaveri@yale.edu](mailto:hitten.zaveri@yale.edu)

*E-mail address:* [coifman@math.yale.edu](mailto:coifman@math.yale.edu)