Epileptic electroencephalography classification using Embedded Dynamic Mode Decomposition

Jennifer Hellar, Negar Erfanian, and Behnaam Aazhang

Department of Electrical and Computer Engineering
Rice University, Houston, TX 77005

E-mail: jennifer.hellar@rice.edu, ne12@rice.edu, aaz@rice.edu

April 2022

Abstract. Objective: Seizure prediction devices for drug-resistant epileptic patients could lead to improved quality of life and new treatment options, but current approaches to classification of electroencephalography (EEG) segments for early identification of the pre-seizure state typically require many features and complex classifiers. We therefore propose a novel spatio-temporal EEG feature set that significantly aids in separation and easy classification of the interictal and preictal states. Approach: We derive key spectral features from the Embedded Dynamic Mode Decomposition (EmDMD) of the brain state system. This method linearizes the complex spatio-temporal dynamics of the system, describing the dynamics in terms of a spectral basis of modes and eigenvalues. The relative subband spectral power and mean phase locking values of these modes prove to be good indicators of the preictal state that precedes seizure onset. Main results: We analyze the linear separability and classification of preictal and interictal states based on our proposed features using seizure data extracted from the CHB-MIT scalp EEG and Kaggle American Epilepsy Society Seizure Prediction Challenge intracranial EEG databases. With a light-weight support vector machine or random forest classifier trained on these features, we classify the preictal state with a sensitivity of up to 92% and specificity of up to 89%. Significance: The EmDMD-derived features separate the preictal and interictal states, improving classification accuracy and motivating further work to incorporate them into seizure prediction algorithms.

Submitted to: *J. Neural Eng.*

1. Introduction

Epilepsy is a serious neurological disorder characterized by abnormal brain activity that leads to seizures or periods of unusual behavior, sensations and sometimes loss of awareness. Nearly 1% of the world's population suffers from this condition, but close to 30% of epilepsy cases prove to be drug-resistant [1–5]. For those patients, few options remain. Brain resection, a surgery in which the piece of brain believed to be causing the seizures is removed, is not feasible for some patients and is not always completely successful even when performed [6–8]. Therefore, patients with drug-resistant epilepsy can still experience sudden and unexpected seizures that can lead to embarrassment, injury, or even death depending on the location and activity of the patient at the onset e.g. driving a car.

A seizure prediction device that notifies the patient to stop risky activities and get help would significantly mitigate stress and reduce seizure-related injuries. Alternatively, an implantable device that adaptively and preemptively stimulates the brain to prevent seizures could be the next treatment option for drug-resistant epileptic patients [9–11]. The problem of seizure prediction, the capability to anticipate seizure occurrence several minutes beforehand, presents a major roadblock to the development of such devices.

Various seizure prediction techniques have been proposed that process either scalp electroencephalography (sEEG) or intracranial electroencephalography (iEEG) recordings‡. Most focus on deriving and extracting key features from these signals that help to differentiate the interictal (between seizures, or baseline), preictal (short time prior to seizure), ictal (seizure) and post-ictal (after seizure) states of the brain. The seizure prediction problem is then typically reformulated as a classification problem where short time windows of the sEEG or iEEG recordings are labelled as either interictal or preictal and a preictal label is assumed to trigger a "seizure alarm."

For a feasible device implementation, the seizure prediction must be accomplished in near real-time with a minimum number of EEG channels, a small feature set, and a fast classifier. Reducing the number of EEG channels directly lowers the large power consumption of analog to digital converters. It also reduces data dimensionality, and when transformed to a small feature set, reduces overall computational complexity.

A wide variety of time- and frequency-domain features of sEEG/iEEG signals have been analyzed by prior works, including mean phase coherence [12, 13], correlation dimension and entropy [14], Lyapunov exponents [15], spike rate [16], phase correlation [17], power spectral density [18], phase synchronization from empirical mode decomposition [19], attractor states [20], short-time Fourier transform [21], wavelet transform [22, 23], fast Fourier transform [24], and Pearson correlation [25]. Classification has been done using Support Vector Machines (SVMs) [18, 19], Neural Networks [21, 22, 25], and various deep learning approaches [26, 27].

Despite the proliferation of works on the subject, however, EEG classification remains an open research problem due to the many challenges associated with it. Seizure

‡ Throughout this work, "EEG" with no prefix refers to both sEEG and iEEG data.

patterns and even baseline EEG patterns are extremely dependent on the patient, the seizure type, and the EEG electrode placements. EEG signals are in general noisy, subject to biological artifacts, and not wide-sense stationary. Noise can come from interference from nearby machines, electrode displacement, or other sources. Artifacts can stem from patient eye movements, muscle activity, and heartbeats during the recording. Of particular interest to us, EEG recordings contain both temporal and spatial information, but the non-linear dynamics across time and across channels are complex and difficult to model and interpret.

To address these problems, we, first, apply a powerful tool to describe the spatio-temporal dynamics of EEG observations called Embedded Dynamic Mode Decomposition (EmDMD) [28, 29]; and second, propose a novel EmDMD-based feature set that greatly improves the separability and consequently the classification of interictal versus preictal states for seizure prediction. From dynamical system theory, we are familiar with the Koopman operator which describes non-linear system evolution with an infinite-dimensional linear expansion [30, 31]. The eigendecomposition of this Koopman operator gives (1) eigenvalues that capture the temporal growth or decay rates and oscillation frequencies and (2) "modes" that correspond to coherent spatial structures. We numerically approximate the Koopman eigendecomposition with Dynamic Mode Decomposition (DMD) [32–36] applied to the EEG data with time-delay embedding, resulting in modes and eigenvalues that describe the non-linear spatio-temporal dynamics of electrical activity measured from the epileptic brain in a linear setting. The frequencies determined by the eigenvalues of this decomposition allow us to treat it as a spectra similar to those computed by more typical transforms [33].

To illustrate the application of EmDMD, we rely on past works that have shown the usefulness of spectral power and phase synchronization features to distinguish interictal and preictal states [13, 18, 37–39]. Rather than the various transforms previously used to estimate these two features, we redefine them in terms of the power and phase of the spatio-temporal spectra derived from our EmDMD analysis. These novel features, relative subband spectral powers and mean phase locking values, exhibit excellent capability to linearly separate preictal and interictal segments of sEEG and iEEG data. We test EmDMD and validate the derived feature pair for EEG segment classification using the Physionet CHB-MIT sEEG database [40, 41] and the Kaggle American Epilepsy Society Seizure Prediction Challenge iEEG database [27].

The paper is organized as follows. Section 2 summarizes the patient datasets used to evaluate the capabilities of EmDMD. Section 3 describes our methods, including pre-processing, embedded dynamic mode decomposition, spatio-temporal feature extraction, feature selection, and cross-validated classification. In Section 4, we present our results. Section 5 briefly highlights opportunities for future work. And finally, Section 6 summarizes our key contributions and concludes.

2. Datasets

2.1. Data labelling

There is no general consensus in the literature regarding the proper labelling of unlabelled EEG recordings as preictal, interictal, postictal, etc. We therefore follow this labelling rule that roughly corresponds to that used in the Kaggle American Epilepsy Society Seizure Prediction Challenge [27]:

- **Preictal periods**: 1 hour prior to seizure onset,
- **Postictal periods**: 1 hour after seizure end,
- **Interictal periods**: at least 4 hours prior to any seizure onset and 4 hours after seizure end.

To prevent postictal contamination of preictal data, we combine any seizures that occur less than 2 hours apart and consider only the preictal data of the leading seizure.

2.2. CHB-MIT scalp EEG (sEEG) dataset

We test our methods first on the CHB-MIT scalp EEG dataset [40, 41] which contains a total of 23 cases recorded from 22 pediatric patients with intractable seizures at the Children's Hospital Boston. Note that case 21 was obtained 1.5 years after case 01, from the same patient, but all other cases correspond to unique patients. Each case records 23-28 EEG channels sampled at 256Hz with 16-bit resolution. For consistency, we use only the 18 channels present in all recordings[§].

Under the labelling criteria in Section 2.1, we exclude cases 03, 11, 19, and 21 due to insufficient preictal data, as each contains only 1 usable seizure. Similarly, we exclude cases 08, 15, and 24 for insufficient interictal data of less than 2 hours each. In the recording of case 12, the channels measured changed multiple times and the patient experienced a total of 40 seizures in 24 hours, making the consistent selection of channels and extraction of interictal and preictal states especially challenging. For this reason, we do not analyze case 12. The remaining 16 cases are summarized in Table 1.

2.3. Kaggle Challenge intracranial EEG (iEEG) dataset

We additionally evaluate EmDMD on intracranial EEG data from the Kaggle American Epilepsy Society Seizure Prediction Challenge dataset [27]. This dataset includes recordings from five dogs and two human patients. Canine data were recorded from 16 implanted electrodes with a sampling rate of 400 Hz. Patient iEEG data were recorded from 15 depth electrodes (Patient_1) and 24 subdural electrodes (Patient_2) at a sampling rate of 5 kHz.

Ten minute segments of preictal and interictal recordings were extracted by the organizers such that, for each lead seizure, six preictal segments were extracted from 65

§ CHB-MIT channels used: FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, FZ-CZ, CZ-PZ.

Table 1. CHB-MIT sEEG Dataset Summary. Seizure types: Simple Partial (SP), Complex Partial (CP), or Generalized Tonic-Clonic (GTC).

Case ID	Seizure type	Seizure origin	Seizures (used)	Interictal hours	Preictal hours
01	SP, CP	Frontal	5	14.3	4.9
02	SP, CP, GTC	Temporal	2	23.9	2
04	SP, CP, GTC	Temporal	3	131.4	3
05	CP, GTC	Frontal	4	14.4	4
06	CP, GTC	Temporal/Occipital	7	24.6	5.5
07	SP, CP, GTC	Temporal	3	51.3	2
09	CP, GTC	Temporal/Occipital	3	48.6	2
10	SP, CP, GTC	Temporal	6	21.9	6
13	SP, CP, GTC	Temporal/Occipital	3	12.6	3
14	CP, GTC	Frontal/Temporal	3	2.2	3
16	SP, CP, GTC	Temporal	2	5.6	2
17	SP, CP, GTC	Temporal	2	7.1	1.6
18	SP, CP	Frontal	2	25.0	2
20	SP, CP, GTC	Temporal/Parietal	2	10.8	2
22	-	Temporal	3	12.9	3
23	-	Temporal	2	14.2	2
Total			52	420.7	48

Table 2. Kaggle iEEG Dataset Summary.

Case ID	Channels	Seizures	Interictal hours	Preictal hours
Patient_1	15	3	8.3	3
Patient_2	24	3	7.0	3
Dog_1	16	4	78.7	3.9
Dog_2	16	7	81.9	6.9
Dog_3	16	12	236.0	4
Dog_4	16	17	131.8	15.9
Dog_5	15	5	73.8	4.9
Total		50	617.5	41.6

min to 5 min before seizure onset. Interictal segments were randomly selected at least 1 week from any seizure for the canine data and at least 4 hours from any seizure for the patient data.

The human patient data of the iEEG dataset was originally sampled at 5kHz. To reduce computational complexity, we downsample this data by a factor of 10 to an effective sampling rate of 500Hz. Downsampling is performed by averaging every 10 samples and storing the result. In addition, due to the large class imbalance and to speed up processing, we use only the first 40 hours of interictal data for each canine case. This dataset is summarized in Table 2.

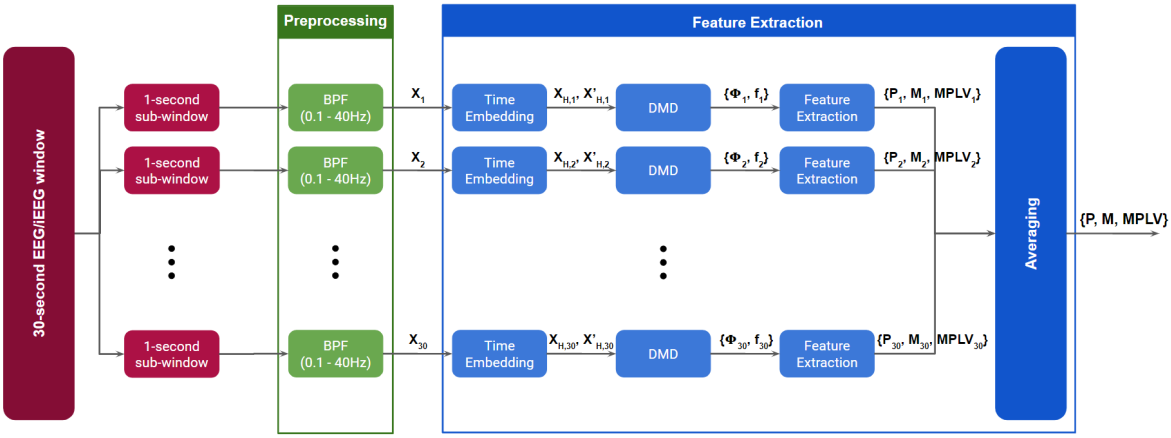


Figure 1. Diagram of proposed methodology for EmDMD preprocessing and feature extraction.

3. Methods

At the highest level, our methodology is divided into the following stages: preprocessing, EmDMD spectra estimation, feature extraction, feature selection, and cross-validated classification. The pipeline for the first three stages is summarized in Fig. 1. In the following sections, we review the theory of dynamic mode decomposition and describe in detail the implementation of each stage.

3.1. Preprocessing

There exists no consensus in the literature on the optimal window size to analyze EEG data, and choices range from 1 second [19] up to 5 minutes [17]. To be conservative, we process all sEEG and iEEG data in sub-windows of $w = 1$ second and average the resulting features within each non-overlapping window of $W = 30$ seconds. Prior to any feature extraction, each 30-second window of data is filtered with a 2nd-order digital Butterworth bandpass filter from 0.1 to 40Hz with padding of 150 samples.

3.2. EmDMD

As mentioned previously, the Koopman operator describes non-linear system evolution with an infinite-dimensional linear expansion. We quickly review this operator and its significance before detailing the EmDMD methodology for estimating the Koopman spectra of EEG data. The resulting EmDMD spectra forms the basis of our proposed novel feature set for the classification of preictal versus interictal segments.

3.2.1. Linearizing a nonlinear system with the Koopman operator Let us consider the electrical activity of the brain to be a dynamical system evolving on a finite-dimensional manifold $\mathcal{M}$ such that the vector $\mathbf{s}_k \in \mathcal{M}$ represents the complete state of the system

at time step k and is in turn some non-linear function of the previous state,

$$\mathbf{s}_k = \mathbf{f}(\mathbf{s}_{k-1}).$$

We know that the linear Koopman operator $\mathcal{U}$ acts on scalar functions $g : \mathcal{M} \to \mathbb{R}$ such that

$$\mathcal{U}g(\mathbf{s}_{k-1}) = g(\mathbf{f}(\mathbf{s}_{k-1})) = g(\mathbf{s}_k), \tag{1}$$

and can be decomposed into an infinite number of eigenvalues $\lambda_j \in \mathbb{C}$ and eigenfunctions $\theta_j(\mathbf{s}) : \mathcal{M} \to \mathbb{R}$ for $j \in \mathbb{N}_0$, the set of non-negative integers [30, 31].

Since we can never measure the complete state of the brain's electrical activity, we instead have a set of n measured spatial values that are referred to as sEEG or iEEG channels. We represent these observations at a given timestep k as a vector $\mathbf{x}_k \in \mathbb{R}^n$. A summary of these and other parameters can be found in Table 3. We formalize the n-dimensional EEG recording as a vector-valued observable $\mathbf{g} : \mathcal{M} \to \mathbb{R}^n$, and can write

$$\mathbf{x}_k = \mathbf{g}(\mathbf{s}_k) = \sum_{i=0}^{\infty} \lambda_i^k \theta_i(\mathbf{s}_0)\phi_i, \tag{2}$$

where *Koopman modes* $\phi_i \in \mathbb{C}^n$ have growth rate and frequency determined by the magnitude and phase of the corresponding eigenvalues λ_i [32]. Koopman eigenfunctions θ_i are evaluated on the initial state of the system $\mathbf{s}_0$, making the observed state $\mathbf{x}_k$ at timestep k a linear function of the complete state $\mathbf{s}_0$ at timestep 0.

The Koopman operator is therefore a linear transformation that describes the evolution of observables in the state space of a dynamic system. That evolution is fully characterized by the set of all Koopman modes and eigenvalues $\{\phi_i, \lambda_i\}$ for $i = 0, 1, 2, ...$, called the Koopman spectra. Intuitively, the Koopman modes are strongly related to the eigenvectors of linear systems and the discrete Fourier transform (DFT) of periodic systems [33]. In the latter case, the phases of the Koopman eigenvalues are proportional to the frequencies of the DFT.

Table 3. EmDMD and dataset parameters.

Parameter	**Description**	**CHB-MIT**	**Patient-iEEG**	**Dog-iEEG**
n	channels	18	15, 24	16, 15
f_s	sampling frequency	256 Hz	500 Hz	400 Hz
W	window size	30 sec	30 sec	30 sec
w	sub-window size	1 sec	1 sec	1 sec
r	singular values kept	100	100	100
m	sub-window size (in samples)	$m = f_s w$		
h	Hankel embedding parameter	$\lceil 3m/n \rceil$		
k	sample/time-step index	$k \in \{1, 2, ..., m\}$		
$\mathbf{x}_k$	observations at timestep k	$\mathbf{x}_k \in \mathbb{R}^n$		

3.2.2. Approximating Koopman spectra with Dynamic Mode Decomposition Since the Koopman operator exists in infinite-dimensional space, we approximate its spectral modes and eigenvalues in a lower-dimensional space using DMD, a data-driven algorithm that been successfully used in a number of applications [28, 34].

First, we define an EEG sample sub-window of $w = 1$ second over which to estimate the Koopman spectra. Assuming a sample rate of f_s Hz, we have $m = f_s w$ samples (timesteps) per window. We gather these m snapshots into two $n \times (m-1)$ matrices,

$$\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad ... \quad \mathbf{x}_{m-1}], \qquad \mathbf{X}' = [\mathbf{x}_2 \quad \mathbf{x}_3 \quad ... \quad \mathbf{x}_m].$$

We then approximate the relationship between $\mathbf{X}$ and the time-shifted $\mathbf{X}'$ with a linear operator $\mathbf{A}$ such that

$$\mathbf{X}' = \mathbf{A}\mathbf{X}. \tag{3}$$

Here, $\mathbf{A}$ approximates the Koopman operator, so that the eigendecomposition of $\mathbf{A}$ approximates the Koopman spectra that we are seeking [33–35]. The dynamic mode decomposition of the pair of matrices $\mathbf{X}$, $\mathbf{X}'$ is therefore the eigendecomposition of $\mathbf{A}$, which can be computed with the psuedoinverse of $\mathbf{X}$ obtained by its singular value decomposition (SVD).

The SVD of $\mathbf{X}$ produces at most n non-zero singular values, resulting in at most n DMD modes and eigenvalues. In the case of EEG data, the number of samples per window $m >> n$, the number of spatial channels, so n DMD modes and eigenvalues are insufficient to accurately approximate the complex dynamics over m timesteps. Therefore, we instead construct the following augmented versions of the data matrices with time-delay embedding, called Hankel matrices [28, 36],

$$\mathbf{X_H} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{m-h} \\ \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_{m-h+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_h & \mathbf{x}_{h+1} & \cdots & \mathbf{x}_{m-1} \end{bmatrix}, \quad \mathbf{X'_H} = \begin{bmatrix} \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_{m-h+1} \\ \mathbf{x}_3 & \mathbf{x}_4 & \cdots & \mathbf{x}_{m-h+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{h+1} & \mathbf{x}_{h+1} & \cdots & \mathbf{x}_m \end{bmatrix}$$

such that $\mathbf{X_H}$, $\mathbf{X'_H} \in \mathbb{R}^{nh \times (m-h)}$. The number of "channels" is now nh, and we can obtain up to nh spatial DMD modes by analysis of

$$\mathbf{X'_H} = \mathcal{A}\mathbf{X_H}. \tag{4}$$

This embedding of the data into the higher-dimensional "Hankel space" is controlled by the Hankel embedding parameter h and results in an embedded operator $\mathcal{A}$ that describes how the EEG signals evolve in that space. In general, it is recommended to choose the embedding parameter h such that $hn > 2m$ [28], therefore we use

$$h = \left\lceil \frac{3m}{n} \right\rceil.$$

To compute the eigendecomposition of $\mathcal{A}$ and obtain our DMD modes and eigenvalues, we first evaluate the SVD of $\mathbf{X_H}$,

$$\mathbf{X_H} = \mathbf{W}\mathbf{\Sigma}\mathbf{U}^{\mathsf{T}},$$

obtaining left singular vectors $\mathbf{W} \in \mathbb{R}^{nh\times(m-h)}$, singular values on the diagonal of $\mathbf{\Sigma} \in \mathbb{R}^{(m-h)\times(m-h)}$, and right singular vectors $\mathbf{U} \in \mathbb{R}^{(m-h)\times(m-h)}$. Re-arranging Eq. (4), we have

$$\mathcal{A} = \mathbf{X'_H}\mathbf{X}_\mathbf{H}^\dagger = \mathbf{X'_H}(\mathbf{W\Sigma U}^\mathsf{T})^\dagger = \mathbf{X'_H U\Sigma}^{-1}\mathbf{W}^\mathsf{T} \tag{5}$$

where † denotes the pseudoinverse of a matrix.

To reduce computational complexity, we assume that a lower-dimensional structure exists within the embedded EEG data and truncate the SVD by keeping only the $r = 100$ most significant singular values and corresponding vectors so that

$$\mathbf{X_H} \approx \tilde{\mathbf{X}}_\mathbf{H} = \mathbf{W}_r\mathbf{\Sigma}_r\mathbf{U}_r^\mathsf{T}, \tag{6}$$

where $\mathbf{W}_r \in \mathbb{R}^{nh\times r}$, $\mathbf{\Sigma}_r \in \mathbb{R}^{r\times r}$, and $\mathbf{U}_r \in \mathbb{R}^{(m-h)\times r}$. Here, the truncation parameter r is chosen such that the highest frequencies present in the final DMD spectra are above 60 Hz; we experimentally verify this on 10 minute preictal and interictal segments selected per patient. We map operator $\mathcal{A}$ to the same lower-dimensional space with

$$\mathcal{A}_r = \mathbf{W}_r^\mathsf{T}\mathcal{A}\mathbf{W}_r \in \mathbb{R}^{r\times r}. \tag{7}$$

Now, we have from Eq. (5), (6) and (7) that

$$\mathcal{A}_r = \mathbf{W}_r^\mathsf{T}\mathbf{X'_H}(\mathbf{U}_r\mathbf{\Sigma}_r^{-1}\mathbf{W}_r^\mathsf{T})\mathbf{W}_r = \mathbf{W}_r^\mathsf{T}\mathbf{X'_H}\mathbf{U}_r\mathbf{\Sigma}_r^{-1}. \tag{8}$$

From this, we compute the eigendecomposition of $\mathcal{A}_r$,

$$\mathcal{A}_r\mathbf{V} = \mathbf{V\Lambda} \tag{9}$$

obtaining eigenvectors $\mathbf{V} \in \mathbb{C}^{r\times r}$ and eigenvalues on the diagonal of $\mathbf{\Lambda} \in \mathbb{C}^{r\times r}$ that allow us to approximate the corresponding eigenvectors and eigenvalues of the higher-dimensional operator $\mathcal{A}$,

$$\mathcal{A}\mathbf{\Psi} \approx \mathbf{\Psi\Lambda} \tag{10}$$

where eigenvectors $\mathbf{\Psi} \approx \mathbf{W}_r\mathbf{V} \in \mathbb{C}^{nh\times r}$ are the EmDMD modes [35].

The time-delay embedding of EmDMD results in h very similar "stacks" (groups of n rows) within the DMD modes in $\mathbf{\Psi}$, so we truncate all but the first n rows to get an approximation for the eigenvectors $\mathbf{\Phi} \in \mathbb{C}^{n\times r}$ of the original operator $\mathbf{A}$ that approximates the evolution of the EEG signals [28]. The infinite-dimensional Koopman operator that describes the non-linear dynamics of the system is therefore now approximated with the linear operator $\mathbf{A}$ in Eq. (3) which can be written

$$\mathbf{X'} \approx \mathbf{AX} = \mathbf{\Phi\Lambda\Phi}^\dagger\mathbf{X}, \tag{11}$$

and we have the modes $\mathbf{\Phi} \in \mathbb{C}^{n\times r}$ and eigenvalues $\mathbf{\Lambda} \in \mathbb{C}^{r\times r}$ that characterize the spatio-temporal impact of that operator. The magnitude and phase of elements in mode $\phi_i \in \mathbb{C}^n$ correspond to spatially coherent voltage and relative phases across the n electrodes. The real part of each eigenvalue $\lambda_i \in \mathbb{C}$ associated with ϕ_i corresponds to a temporal growth or decay rate of the frequency given by the imaginary part.

3.2.3. Defining the EmDMD spectra Recall that for an arbitrary complex number $c = a+jb$ where we have the imaginary number $j = \sqrt{-1}$, the magnitude $|c| = \sqrt{a^2 + b^2}$, and the phase $\theta = \arctan(\frac{b}{a})$, so that $c = a + jb = |c|e^{j\theta}$ in polar form. We know from [34, 36] that an eigenvalue $\lambda_i \in \mathbf{\Lambda}$ is directly related to the Discrete Fourier Transform frequency associated with mode ϕ_i by

$$\lambda_i = |\lambda_i|e^{j\theta_i} \propto e^{j2\pi f_i \Delta t}. \tag{12}$$

For a short time window m and small number r of the most stable singular values, we can assume that EEG frequency components are stable over those few seconds and so the growth or decay rate $|\lambda_i| \approx 1$, therefore

$$\theta_i = \arctan\left(\frac{\mathrm{Im}\left[\lambda_i\right]}{\mathrm{Re}\left[\lambda_i\right]}\right) = 2\pi f_i \Delta t, \tag{13}$$

where the time delay between consecutive samples $\Delta t = f_s^{-1}$, and we can compute Fourier frequencies for each mode in $\mathbf{\Phi}$ as

$$f_i = \frac{\theta_i}{2\pi\Delta t}. \tag{14}$$

Since our EEG signal is real, this results in symmetric positive and negative frequencies, therefore we analyze only the modes corresponding to positive frequencies. Additionally, since [29] showed excellent results on analysis of DMD frequencies below 60 Hz, we similarly discard any modes above that threshold, giving us a final EmDMD mode matrix $\mathbf{\Phi} \in \mathbb{R}^{n\times p}$. Here, p is the final number of EmDMD modes used for feature extraction and is determined dynamically in real time according to the frequencies present in the DMD analysis.

Our EmDMD analysis therefore results in a "spectra" $\{\mathbf{\Phi}, \mathbf{f}\}$ where each frequency element of $\mathbf{f} \in \mathbb{R}^p$ corresponds to a spatio-temporal mode of $\mathbf{\Phi}$. Examples of this spectra, evaluated for a 1-second window of interictal data from case 01 of the CHB-MIT sEEG dataset and for a 1-second window of preictal data from Dog 1 of the Kaggle iEEG dataset, are shown in Fig. 2.

3.3. Spatio-temporal feature extraction

Conceptually, when evaluating windows of sEEG/iEEG samples such as those shown in Fig. 3, we are interested in extracting key distinguishing features to linearly separate and classify them as interictal versus preictal. EEG spectral power and phase synchronization features have respectively been identified as key features for preictal versus interictal classification and therefore seizure prediction in a variety of prior works [12, 13, 17, 18, 37–39, 42]. These have in the past mostly been computed by empirical mode decomposition, Fast Fourier Transform (FFT), Hilbert transform, or wavelet transform techniques. More recently, the first DMD-based subband power and spectral moment EEG features have been defined with good preliminary classification results [29]. Building on this, we now propose two new spatio-temporal feature types based on the EmDMD spectra: *relative subband powers* and *channel-wise mean phase locking values*.

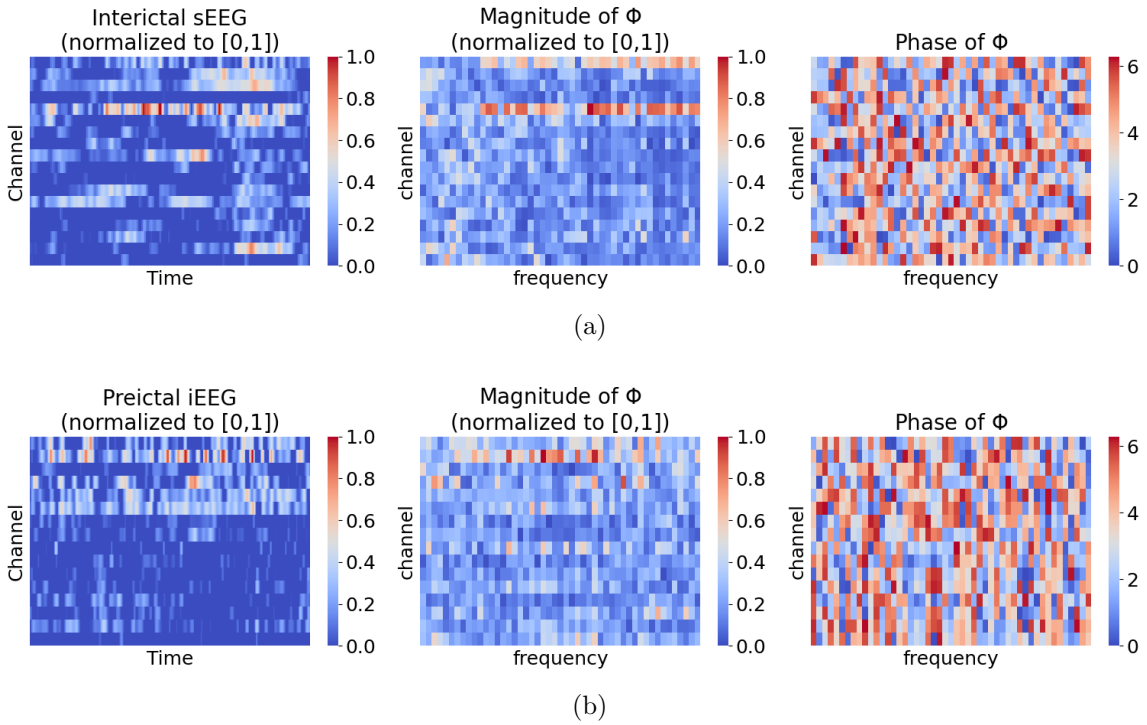


Figure 2. For a 1-second window of interictal sEEG data (a) and of preictal iEEG data (b), the EmDMD spectra plotted respectively as complex magnitude and phase versus frequency captures spatio-temporal activity.

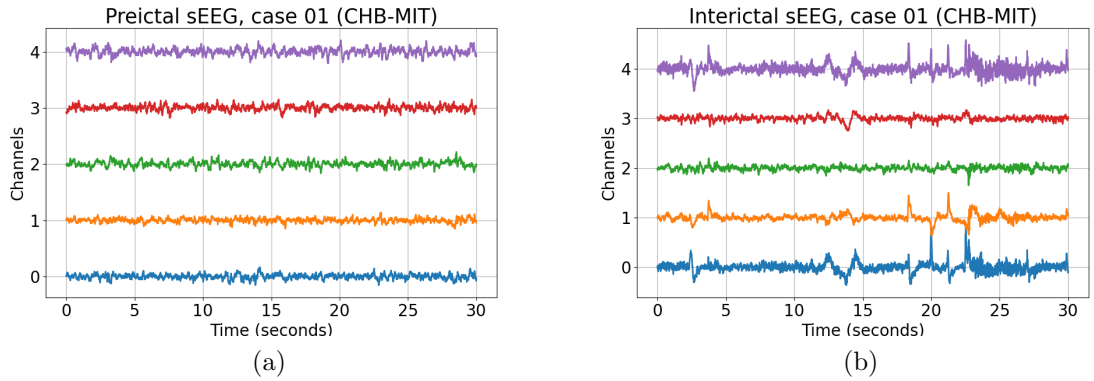
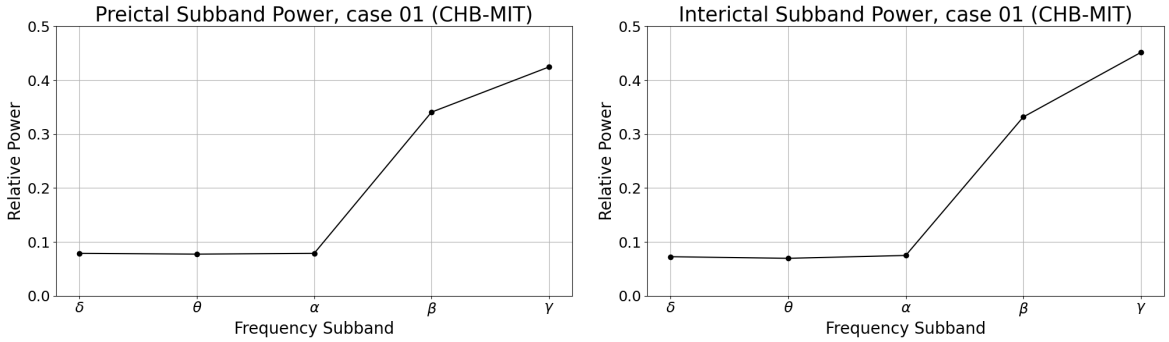


Figure 3. For case 01 of the CHB-MIT scalp EEG dataset, we plot a few channels of the 30-second preictal window ending 5 minutes prior to the first seizure (a), as well as a 30-second interictal window occurring 5 minutes into the first interictal section (b).

3.3.1. Relative subband power features First, we compute the power of each EmDMD mode by taking the squared l^2-norm of each column of $\mathbf{\Phi}$ such that

$$|\mathbf{\Phi}| = [|\phi_1|^2 \quad |\phi_2|^2 \quad \cdots \quad |\phi_p|^2] \in \mathbb{R}^p.$$

This gives us a familiar power spectra $\{|\mathbf{\Phi}|, \mathbf{f}\}$ that describes the distribution of power at various frequencies in the 0-60 Hz range [28]. From this, we compute the subband powers for the δ (0-4 Hz), θ (4-8 Hz), α (8-12 Hz), β (12-30 Hz), and γ (30-60 Hz)



(a) (b)

Figure 4. For 30-second windows of preictal and interictal sEEG data, we find a slight decrease in higher frequency subband powers in the preictal (a) versus interictal (b) states.

frequency bands by summing the individual mode powers corresponding to the respective bands [29]. To obtain our novel *relative subband power feature set*, we additionally normalize the resulting subband powers by dividing out the total power across all 5 bands,

$$P = [P_\delta \quad P_\theta \quad P_\alpha \quad P_\beta \quad P_\gamma] \in \mathbb{R}^5.$$

These features coarsely estimate the relative distribution of EmDMD spectral power over short sub-windows of sEEG/iEEG recordings. For the sEEG segments in Fig. 3, we show the corresponding relative subband powers in Fig. 4. In this case, we find a slight decrease in the higher frequency subband powers in the preictal window compared to the interictal window.

3.3.2. Channel-wise mean phase locking value features For an arbitrary window of m samples, the phase locking value (PLV) measures phase synchronization between two signals x and y at a particular frequency f as

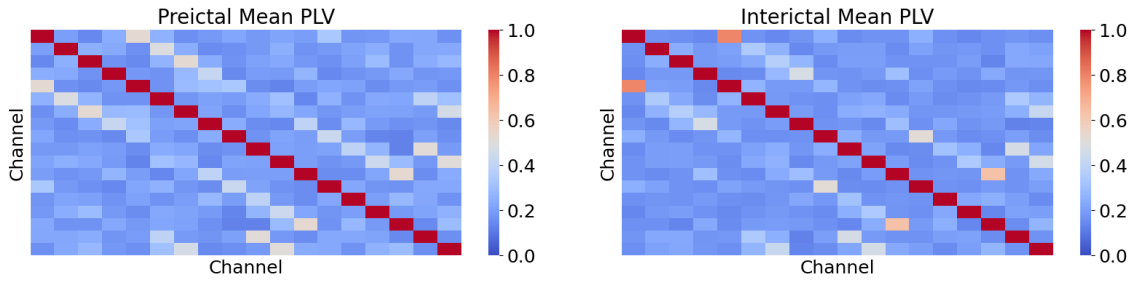
$$PLV(f) = \frac{1}{m}\left|\sum_{k=1}^{m} e^{i\Theta_{x,y}(k,f)}\right|, \tag{15}$$

where $\Theta_{x,y}(k,f)$ is the instantaneous phase difference between x and y at timestep k and frequency f [19]. We adapt this concept to our EmDMD spectra as follows.

First, we obtain a phase matrix $\angle\mathbf{\Phi} \in \mathbb{R}^{n\times p}$ by computing the element-wise phases of $\mathbf{\Phi}$. Then, we define the phase difference between two channels x and y for frequency $f \in \mathbf{f}$ as

$$\Theta_{x,y}(f) = \angle\mathbf{\Phi}_x(f) - \angle\mathbf{\Phi}_y(f),$$

where $\angle\mathbf{\Phi}_x(f)$ denotes the xth row of $\angle\mathbf{\Phi}$ and the column (mode) corresponding to frequency f. Now for every unique pair of channels (x, y), we compute a mean phase



(a) (b)

Figure 5. For 30-second windows of preictal and interictal sEEG data from case 01 (CHB-MIT), we find a decrease in MPLVs for many channel pairs in the preictal (a) versus interictal (b) features.

locking value, averaging over modes (frequencies) rather than time,

$$MPLV_{x,y} = \frac{1}{p}\left|\sum_{f\in\mathbf{f}} e^{i\Theta_{x,y}(f)}\right|, \tag{16}$$

to obtain our *mean phase locking feature set*,

$$MPLV \in \mathbb{R}^{\frac{n(n-1)}{2}}.$$

Intuitively, each value in this feature set measures the average phase synchronization between a pair of channels with respect to the EmDMD spectra. Considering again the sEEG segments in Fig. 3, we show the corresponding MPLV's in Fig. 5 and find a decrease for many channel pairs in the preictal window.

3.3.3. Spectral moment features In addition to the above novel EmDMD feature sets, we also compute the first three DMD spectral moments [29], where the jth moment is given by

$$M^j = \sum_{i=1}^{p} f_i^j |\phi_i|, \tag{17}$$

obtaining the *spectral moment feature set*

$$M = [M^1 \quad M^2 \quad M^3] \in \mathbb{R}^3.$$

3.3.4. Feature smoothing We evaluate EEG recordings and extract the above EmDMD-based features in sub-windows of $w = 1$ second. To get smoothly varying and stable features, we average all extracted features within each non-overlapping window of $W = 30$ seconds. For an n-channel analysis, our final feature set, $\mathbf{F} = [P \quad M \quad MPLV]$ therefore contains $q = 8 + n(n-1)/2$ features per 30-second segment of sEEG/iEEG data.

3.4. Feature standardization and selection

Our EmDMD features differ significantly in the scale of values present, which can negatively impact learning and classification performance. Therefore, for any training set of features extracted from segments of sEEG/iEEG data, we first estimate the mean and variance per feature. Then we standardize each feature of both the training set and corresponding test set by subtracting the appropriate mean and dividing out the variance.

Additionally, we know that for relatively small datasets, classification based on many input features can become challenging as training will tend to overfit the large number of parameters. Automated feature selection based on learned "importance" values is therefore a common technique to improve performance [43, 44]. To reduce the size of our EmDMD feature set prior to training a classifier, we first fit a logistic regression model on the training data. This model generates a set of coefficients corresponding to the features so that their weighted linear combination estimates the log-odds probability of a positive label. The coefficients therefore form a relative "importance" measure of the input features for classification of the training set.

To select the most useful features, we simply keep all features whose regression coefficient is greater than or equal to the mean regression coefficient. Conceptually, we can view this process as the selection of the most informative frequency bands, spectral moments, and channel pairs for classification of preictal versus interictal sEEG/iEEG segments.

3.5. Cross-validated classification

To measure the potential utility of these EmDMD features for seizure prediction, we train and test two binary classifiers, a linear support vector machine (SVM) and a Random Forest (RF) classifier [45–47], to identify interictal versus preictal segments based only on our reduced EmDMD feature sets. The RF classifier is initialized with 100 estimators. Note that we do not consider ictal, post-ictal, or other states as these are essentially irrelevant to seizure prediction. Additionally, the seizure detection (ictal classification) problem is much easier and many works already have excellent results [48, 49].

Seizure types and locations and EEG features can vary widely from patient to patient. We expect that a seizure prediction device would likely incorporate some level of patient-specific training and optimization, and we therefore evaluate our methodology in a patient-specific manner by training and testing on only a single patient at a time.

3.5.1. Data selection A key stage in developing a seizure prediction algorithm is the definition of the prediction window. Assuming that the positive classification of an EEG segment results in a "seizure alarm," we define the seizure prediction horizon (SPH) as the minimum time period after an alarm before the seizure is expected to occur. During this period, the patient would be alerted and ideally take precautionary or preventative

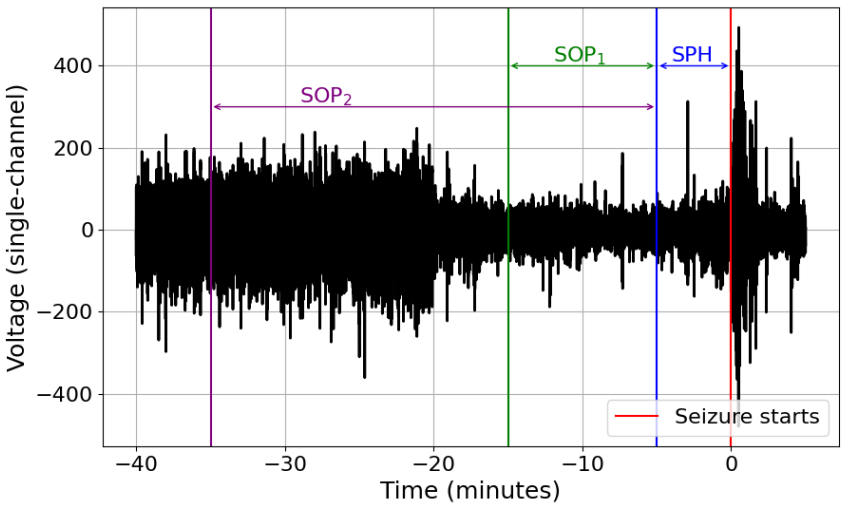


Figure 6. We extract and classify sEEG/iEEG preictal segments according to the seizure prediction horizon (SPH, blue) and two values for the seizure occurrence period: $\text{SOP}_1 = 10$ min (green) and $\text{SOP}_2 = 30$ min (purple).

actions. Following this, we define the seizure occurrence period (SOP) as the interval during which the seizure is expected to occur. For a correct prediction, seizure onset must be after the SPH and within the SOP.

We implement our classification methodology with an SPH of 5 minutes and an SOP of 10 min and 30 min in mind. This is a short enough SOP to minimize patient anxiety while waiting for possible seizure onset and a large enough SPH to allow the patient at a minimum to halt dangerous activities. The "prediction window" of preictal data that we use for training and testing therefore consists of the sEEG/iEEG segments from 15 to 5 min and from 35 to 5 min prior to each leading seizure onset. These definitions of SPH and SOP are visualized on an EEG segment in Fig. 6.

For each patient, we extract the EmDMD features for these preictal segments and for a randomly selected subset of interictal data. To balance the classes, we choose the same amount of interictal data as preictal.

3.5.2. Cross-validation Once the preictal and interictal features have been obtained, we combine, shuffle, and split each patient's data using 5-fold cross-validation (CV). For each iteration, we utilize 4 folds (80% of the data) for training and the remaining fold (20% of the data) for testing. The training data is used for standardization and feature selection as described in Section 3.4. The modified training data is then used to train the linear SVM or RF classifier. Performance metrics are measured on the fitted classifier's prediction of the test set and then averaged across the 5 CV iterations.

3.5.3. Performance metrics As performance and calibration measures, for each dataset, we report the metrics defined in Table 4, averaged across the CV iterations,

Table 4. Performance and calibration metrics. The calibration curve plots the predicted versus observed cases for the test set. TP ≡ True Positive, TN ≡ True Negative, FP ≡ False Positive, FN ≡ False Negative.

Metric, Abbreviation	**Formula**	**Interpretation**	**Best**
Area Under Receiver Operating Characteristic Curve, AUROC	-	overall ability to discriminate	1.0
F1 Score, F1	-	harmonic mean of precision and recall	1.0
Kappa Score, KAP	-	Cohen's coefficient of agreement	1.0
Sensitivity, SEN	$\frac{TP}{TP+FN}$	ability to detect preictal segments	1.0
Specificity, SPE	$\frac{TN}{TN+FP}$	ability to detect interictal segments	1.0
Calibration slope, CAS	-	regression slope	1.0
Calibration intercept, CAI	-	regression intercept	0.0
Calibration r^2, CAR	-	regression fit	1.0

as well as the average number of 30-second sEEG/iEEG segments utilized and average number of features selected across the patients in the dataset. Individual patient results are reported in the Supplementary Document.

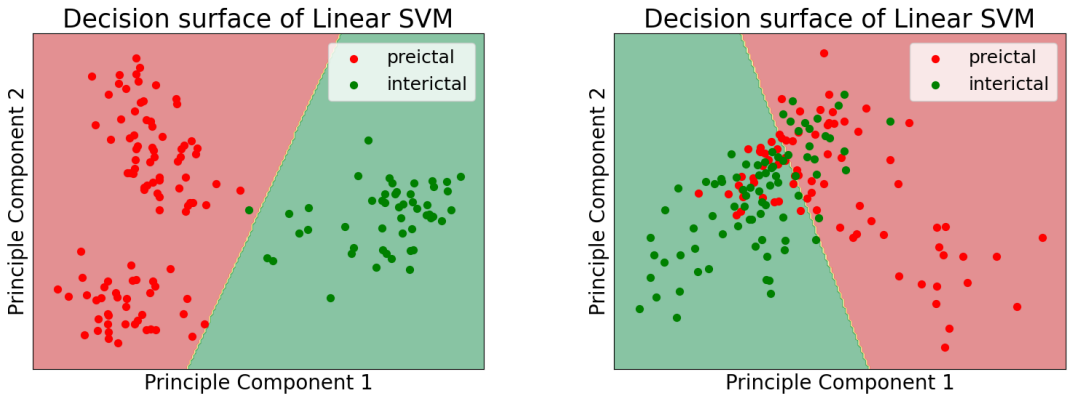
3.6. Visualization of linear separability

Beyond classification results, we also visualize the linear separation induced by our features and the decision boundary learned for a few patients. For case 01 of the CHB-MIT dataset, we use the preictal segments from 15 to 5 min prior to the first five seizures and the interictal segments from 50-60 min into the first five interictal sections.

For one cross-validation split, we apply the feature scaling and feature selection described previously on the training set. Then, to reduce dimensionality so that we can view the linear SVM decision boundary, we additionally apply Principle Components Analysis (PCA) to the training features, reducing the final feature set size to 2. These are used to train the linear SVM, and the same projection is applied to the test set of features before prediction.

Each preictal or interictal 30-second window (training and test) is then plotted as a function of the two PCA-flattened EmDMD features as shown in Fig. 7a. Here, the PCA-flattened features result in a clear separation of the classes, and the linear SVM decision boundary splits them accordingly.

Executing the same process on the first four seizures and interictal sections of Dog 1 from the Kaggle iEEG dataset with the result shown in Fig. 7b, we see that although the classes are relatively well clustered based on the PCA-flattened features, they are not fully separated in just two dimensions.



(a) (b)

Figure 7. For 30-second windows of preictal (red) and 30-second windows of interictal (green) sEEG data taken from case 01 of the CHB-MIT dataset (a) and iEEG data from Dog_1 of the Kaggle dataset (b), our PCA-flattened EmDMD-based features effectively separate the two classes for classification based on the linear SVM decision boundary. In the case of the iEEG dataset, two dimensions is insufficient to fully separate the classes.

Table 5. Mean classification results for the CHB-MIT sEEG dataset, SPH = 5 min.

	Segments (per case)	**Features (per case)**	**AUROC**	**F1**	**KAP**	**SEN**	**SPE**	**SOP**
Linear SVM	122.5	66.8	0.893	0.891	0.780	0.901	0.885	10 min
Random Forest	122.5	66.8	0.889	0.886	0.770	0.899	0.879	10 min
Linear SVM	360.5	65.5	0.902	0.902	0.803	0.909	0.895	30 min
Random Forest	360.5	65.5	0.913	0.914	0.824	0.928	0.897	30 min

4. Results

4.1. CHB-MIT scalp EEG classification

The linear SVM and Random Forest classification results on the scalp EEG dataset, averaged across all cases, are given in Table 5 for an SOP of 10 min and 30 min respectively. Individual patient results are provided in the Supplementary Document. Average calibration metrics for both classifiers are given in Table 6. Overall, both classifiers exhibit similar performance, and we find that our EmDMD features can distinguish preictal versus interictal 30-second segments for classification with an average AUROC $> 91\%$, sensitivity $> 92\%$, and specificity $> 89\%$.

Table 7‖ compares our results versus those of recent works published utilizing the CHB-MIT scalp EEG dataset. Of these, [21, 22, 25] rely on convolutional neural

‖ [18, 20–22] report average false positive rates (FPR) of 0.16/h, 0.423/h, 0.05/h, and 0.147/h respectively. We compute SPE = 1 - FPR.

Table 6. Mean calibration results for the CHB-MIT sEEG dataset.

	SOP	**CAS**	**CAI**	**CAR**
Linear SVM	10 min	1.240	-0.160	0.746
Random Forest	10 min	1.367	-0.188	0.809
Linear SVM	30 min	1.346	-0.228	0.801
Random Forest	30 min	1.365	-0.193	0.845

Table 7. Comparison of classification results for the CHB-MIT sEEG dataset. PCC = Pearson Correlation Coefficient. STFT = Short-Time Fourier Transform. CNN = Convolutional Neural Network. MEMD = Multi-variate Empirical Mode Decomposition. PSD = Power Spectral Density.

Method	**Patients**	**Features**	**Classifier**	**AUROC**	**SEN**	**SPE**	**SPH**	**SOP**
Zhang et al. [18]	17	PSD	SVM	-	0.987	0.950	0 min	60 min
Cho et al. [19]	21	MEMD	SVM	0.877	0.828	0.830	0 min	5 min
Chu et al. [20]	13	Attractor state	Thres.	-	0.821	0.577	30 sec	85 min
Truong et al. [21]	13	STFT	CNN	-	0.812	0.84	5 min	30 min
Khan et al. [22]	15	Wavelets	CNN	0.866	0.878	0.853	0 min	10 min
Zhang et al. [25]	19	PCC	CNN	-	0.929	0.870	0 min	15 min
Usman et al. [26]	23	Many	Ensemble	-	0.963	0.957	-	-
Proposed	16	EmDMD	SVM	0.893	0.901	0.885	5 min	10 min
Proposed	16	EmDMD	RF	0.913	0.928	0.897	5 min	30 min

networks (CNNs) that are far more computationally intensive than our linear SVM or random forest classifiers. Similarly, [26] utilizes a large feature set and an ensemble learning approach that is computationally heavy, and fails to clearly report their preictal/interictal labelling critera and SPH/SOP parameters. This key information makes valid comparisons impossible. [18, 19, 22, 25] all implicitly or explicitly define an SPH of 0 min, which is not clinically useful for seizure prediction.

4.2. Kaggle intracranial EEG classification

The linear SVM and Random Forest classification results on the iEEG dataset, averaged across all cases, are given in Table 8 for an SOP of 10 min and 30 min respectively. Individual patient results are provided in the Supplementary Document, and average calibration metrics for both classifiers are given in Table 9.

In general, we would expect the classification of iEEG data to be easier than that of sEEG data, since the iEEG electrodes are often less noisy. Contrary to expectation, therefore, we find that our EmDMD features are much less successful at separating the preictal and interictal states in iEEG data, achieving on average over the whole dataset an AUROC $> 83\%$, sensitivity $> 83\%$, and specificity $> 83\%$.

When we restrict the data to only the first 4 seizures (if available) and the first

Table 8. Mean classification results for the Kaggle iEEG dataset, SPH = 5 min.

	Segments (per case)	**Features (per case)**	**AUROC**	**F1**	**KAP**	**SEN**	**SPE**	**SOP**
Linear SVM	291.4	59.7	0.805	0.801	0.607	0.806	0.804	10 min
Random Forest	291.4	59.7	0.836	0.832	0.667	0.839	0.832	10 min
Linear SVM	857.7	55.7	0.788	0.787	0.576	0.791	0.785	30 min
Random Forest	857.7	55.7	0.832	0.831	0.663	0.839	0.825	30 min

Table 9. Mean calibration results for the Kaggle iEEG dataset.

	SOP	**CAS**	**CAI**	**CAR**
Linear SVM	15 min	1.135	-0.119	0.727
Random Forest	15 min	1.337	-0.173	0.790
Linear SVM	30 min	1.227	-0.189	0.823
Random Forest	30 min	1.293	-0.150	0.877

Table 10. Average classification results for the Kaggle iEEG dataset, SPH = 5 min, considering only the first 4 seizures (if available) and 4 hours of interictal data per case.

	Segments (per case)	**Features (per case)**	**AUROC**	**F1**	**KAP**	**SEN**	**SPE**	**SOP**
Linear SVM	148.6	58.3	0.892	0.889	0.779	0.884	0.899	10 min
Random Forest	148.6	58.3	0.929	0.923	0.856	0.901	0.957	10 min
Linear SVM	445.7	57.5	0.889	0.887	0.777	0.886	0.893	30 min
Random Forest	445.7	57.5	0.915	0.910	0.828	0.889	0.942	30 min

4 hours of interictal data per case, we obtain the far better classification results in Table 10, achieving an AUROC $> 92\%$, sensitivity $> 90\%$, and specificity $> 95\%$. Considering the significant difference in performance on this subset of the data versus the whole dataset, we can conclude that while our method is suitable for application to iEEG data, some aspect of the larger dataset is reducing overall performance.

For the Kaggle dataset, the location and type of seizures is unknown, therefore the feature selection stage which mostly chooses key MPLVs of channel pairs may be learning the optimal features to classify seizures in one or two brain locations and then failing to detect others. Alternatively, the overall dataset may contain unknown noisy or dropout sections of iEEG recording that require additional preprocessing beyond our basic bandpass filtering approach.

Table 11 compares our results versus those of the original Kaggle competition leaders [27] and recent works published utilizing the Kaggle iEEG dataset. Of these, [21,23,24] again rely on CNNs that are computationally intensive, and most others utilize

Table 11. Comparison of classification results for the Kaggle iEEG dataset. STFT: Short-Time Fourier Transform, CNN: Convolutional Neural Network, FFT: Fast Fourier Transform, CWT: Complex Wavelet Transform.

Method	**Features**	**Classifier**	**AUROC**	**SEN**	**SPE**	**SPH**	**SOP**
M. Hills [27]	Many	SVM	0.862	-	-	5 min	60 min
QMSDP [27]	Many	Ensemble	0.859	-	-	5 min	60 min
Birchwood [27]	Log spectral power, covariance	SVM	0.839	-	-	5 min	60 min
Truong et al. [21]	STFT	CNN		0.75	0.79	5 min	30 min
Hussein et al. [23]	CWT	CNN	0.928	0.885	0.856	5 min	60 min
Liu et al. [24]	PCA, FFT	CNN	0.84	0.83	0.82	5 min	60 min
Usman et al. [26]	Many	Ensemble		0.942	0.958	-	-
Proposed	EmDMD	RF	0.836	0.839	0.832	5 min	10 min
Proposed	EmDMD	RF	0.832	0.839	0.825	5 min	30 min

large feature sets or ensemble learning approaches that also are computationally heavy. In addition, these works classified 10-min segments of iEEG data, often thresholding or taking a consensus of labels over smaller windows to improve the sensitivity and reduce false positives, while our simulations directly classify short 30-sec windows. Even so, our light-weight linear SVM and random forest classifiers trained only on EmDMD-derived features are able to achieve comparable performance to several of the prior works.

5. Discussion

We have examined the capability of our EmDMD-derived spatio-temporal features to distinguish preictal from interictal segments in a patient-specific manner with excellent results. The robustness of these features to noise and artifacts remains to be fully tested and understood. We have not implemented a framework to separate segments of noisy versus clean EEG recording for examining the impact of noise or implemented advanced pre-processing techniques to reduce noise.

We have observed that iEEG classification is less effective, possibly due to variety in seizure locations and types or else due to some unknown noise or artifacts in the large Kaggle dataset. Further work is needed to understand and characterize these results and to investigate performance on other iEEG datasets.

The extension of EmDMD to incorporate the multi-resolution characteristics of wavelets [50] has potential to improve the time and space localization of extracted features, perhaps helping to identify and isolate seizure onset zone (SOZ) features for better performance in processing the EEG data.

The implementation of this method is very suitable for incorporation into true seizure prediction algorithms executing in real time. Once a classifier is trained with these and likely other relevant features, each new 30-sec window of EEG data would be filtered, the features computed, and a prediction made. If necessary, a consensus of

several consecutive windows would potentially improve practical performance, resulting in an overall seizure prediction every few minutes. Further testing would then be required to estimate how far in advance a seizure is typically predicted in the clinical scenario and performance metrics computed on a per seizure basis rather than per segment as in this work.

6. Conclusion

We have proposed two novel spatio-temporal features of sEEG/iEEG data derived from the EmDMD of the dynamic brain system for EEG segment classification and ultimately seizure prediction. EmDMD linearizes the complex dynamic non-linear evolution of the brain state so that we can extract spectral relative subband power and mean phase locking value features across channels (space) and frequency (time). Compared to neural network learning methods, these features are easily computed and interpretable with respect to the underlying brain state. These features enable us to linearly separate and classify interictal versus preictal segments of EEG recordings. To test this separability, we have extracted interictal and preictal segments from multiple patients in the CHB-MIT sEEG and Kaggle Seizure Prediction iEEG public databases to perform preictal/interictal binary classification in a patient-specific manner, ultimately achieving high sensitivity and specificity. These results motivate future work to analyze the robustness and multiresolution capabilities of these EmDMD-derived features as well as their application in a true seizure prediction algorithm.

Acknowledgments

This work was funded by National Science Foundation (NSF) grants 1406447 and 1533688.

References

[1] Patrick Kwan and Martin J Brodie. Early identification of refractory epilepsy. New England Journal of Medicine, 342(5):314–319, 2000.

[2] MJ Brodie, SD Shorvon, R Canger, P Halasz, S Johannessen, P Thompson, HG Wieser, and P Wolf. Commission on european affairs: appropriate standards of epilepsy care across europe. Epilepsia, 38(11):1245–1250, 1997.

[3] Alessandra Del Felice, Ettore Beghi, Giovanni Boero, Angela La Neve, Graziella Bogliun, Alessia De Palo, and Luigi M Specchio. Early versus late remission in a cohort of patients with newly diagnosed epilepsy. Epilepsia, 51(1):37–42, 2010.

[4] Dieter Schmidt and Lennart Gram. Monotherapy versus polytherapy in epilepsy. CNS drugs, 3(3):194–208, 1995.

[5] JWAS Sander. Some aspects of prognosis in the epilepsies: a review. Epilepsia, 34(6):1007–1016, 1993.

[6] Samuel Wiebe, Warren T Blume, John P Girvin, and Michael Eliasziw. A randomized, controlled trial of surgery for temporal-lobe epilepsy. New England Journal of Medicine, 345(5):311–318, 2001.

[7] José F Téllez-Zenteno, Raj Dhar, and Samuel Wiebe. Long-term seizure outcomes following epilepsy surgery: a systematic review and meta-analysis. Brain, 128(5):1188–1198, 2005.

[8] Barbara C Jobst and Gregory D Cascino. Resective epilepsy surgery for drug-resistant focal epilepsy: a review. Jama, 313(3):285–293, 2015.

[9] William H Theodore and Robert S Fisher. Brain stimulation for epilepsy. The Lancet Neurology, 3(2):111–118, 2004.

[10] Tobias Loddenkemper, Andrew Pan, Silvia Neme, Kenneth B Baker, Ali R Rezai, Dudley S Dinner, Erwin B Montgomery Jr, and Hans O Lüders. Deep brain stimulation in epilepsy. Journal of Clinical Neurophysiology, 18(6):514–532, 2001.

[11] Graham V Goddard. Development of epileptic seizures through brain stimulation at low intensity. Nature, 214(5092):1020–1021, 1967.

[12] Florian Mormann, Klaus Lehnertz, Peter David, and Christian E Elger. Mean phase coherence as a measure for phase synchronization and its application to the eeg of epilepsy patients. Physica D: Nonlinear Phenomena, 144(3-4):358–369, 2000.

[13] Yang Zheng, Gang Wang, Kuo Li, Gang Bao, and Jue Wang. Epileptic seizure prediction using phase synchronization based on bivariate empirical mode decomposition. Clinical Neurophysiology, 125(6):1104–1111, 2014.

[14] Ardalan Aarabi and Bin He. A rule-based seizure prediction method for focal neocortical epilepsy. Clinical Neurophysiology, 123(6):1111–1122, 2012.

[15] N Arunkumar, K Ramkumar, S Hema, A Nithya, Poornima Prakash, and V Kirthika. Fuzzy lyapunov exponent based onset detection of the epileptic seizures. In 2013 IEEE Conference on Information & Communication Technologies, pages 701–706. IEEE, 2013.

[16] Shufang Li, Weidong Zhou, Qi Yuan, and Yinxia Liu. Seizure prediction using spike rate of intracranial eeg. IEEE transactions on neural systems and rehabilitation engineering, 21(6):880–886, 2013.

[17] Mohammad Zavid Parvez and Manoranjan Paul. Epileptic seizure prediction by exploiting spatiotemporal relationship of eeg signals using phase correlation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24(1):158–168, 2015.

[18] Zisheng Zhang and Keshab K Parhi. Low-complexity seizure prediction from ieeg/seeg using spectral power and ratios of spectral power. IEEE transactions on biomedical circuits and systems, 10(3):693–706, 2015.

[19] Dongrae Cho, Beomjun Min, Jongin Kim, and Boreom Lee. Eeg-based prediction of epileptic seizures using phase synchronization elicited from noise-assisted multivariate empirical mode decomposition. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25(8):1309–1318, 2016.

[20] Hyunho Chu, Chun Kee Chung, Woorim Jeong, and Kwang-Hyun Cho. Predicting epileptic seizures from scalp eeg based on attractor state analysis. Computer methods and programs in biomedicine, 143:75–87, 2017.

[21] Nhan Duy Truong, Anh Duy Nguyen, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Samuel Ippolito, and Omid Kavehei. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. Neural Networks, 105:104–111, 2018.

[22] Haidar Khan, Lara Marcuse, Madeline Fields, Kalina Swann, and Bülent Yener. Focal onset seizure prediction using convolutional networks. IEEE Transactions on Biomedical Engineering, 65(9):2109–2118, 2018.

[23] Ramy Hussein, Soojin Lee, Rabab Ward, and Martin J McKeown. Semi-dilated convolutional neural networks for epileptic seizure prediction. Neural Networks, 139:212–222, 2021.

[24] Chien-Liang Liu, Bin Xiao, Wen-Hoar Hsaio, and Vincent S. Tseng. Epileptic seizure prediction with multi-view convolutional neural networks. IEEE Access, 7:170352–170361, 2019.

[25] Shasha Zhang, Dan Chen, Rajiv Ranjan, Hengjin Ke, Yunbo Tang, and Albert Y Zomaya. A lightweight solution to epileptic seizure prediction based on eeg synchronization measurement. The Journal of Supercomputing, 77(4):3914–3932, 2021.

[26] Syed Muhammad Usman, Shehzad Khalid, and Sadaf Bashir. A deep learning based ensemble learning method for epileptic seizure prediction. Computers in Biology and Medicine, 136:104710, 2021.

[27] Benjamin H. Brinkmann, Joost Wagenaar, Drew Abbot, Phillip Adkins, Simone C. Bosshard, Min Chen, Quang M. Tieng, Jialune He, F. J. Muñoz-Almaraz, Paloma Botella-Rocamora, Juan Pardo, Francisco Zamora-Martinez, Michael Hills, Wei Wu, Iryna Korshunova, Will Cukierski, Charles Vite, Edward E. Patterson, Brian Litt, and Gregory A. Worrell. Crowdsourcing reproducible seizure forecasting in human and canine epilepsy. Brain, 139(6):1713–1722, 03 2016.

[28] Bingni W Brunton, Lise A Johnson, Jeffrey G Ojemann, and J Nathan Kutz. Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition. Journal of neuroscience methods, 258:1–15, 2016.

[29] Ozlem Karabiber Cura and Aydin Akan. Analysis of epileptic eeg signals by using dynamic mode decomposition and spectrum. Biocybernetics and Biomedical Engineering, 41(1):28–44, 2021.

[30] Bernard O Koopman. Hamiltonian systems and transformation in hilbert space. Proceedings of the national academy of sciences of the united states of america, 17(5):315, 1931.

[31] BO Koopman and J v Neumann. Dynamical systems of continuous spectra. Proceedings of the National Academy of Sciences of the United States of America, 18(3):255, 1932.

[32] Igor Mezić. Spectral properties of dynamical systems, model reduction and decompositions. Nonlinear Dynamics, 41(1):309–325, 2005.

[33] Clarence W Rowley, Igor Mezić, Shervin Bagheri, Philipp Schlatter, and Dan S Henningson. Spectral analysis of nonlinear flows. Journal of fluid mechanics, 641:115–127, 2009.

[34] Peter J Schmid. Dynamic mode decomposition of numerical and experimental data. Journal of fluid mechanics, 656:5–28, 2010.

[35] Jonathan H Tu, Clarence W Rowley, Dirk M Luchtenburg, Steven L Brunton, and J N Kutz. On dynamic mode decomposition: Theory and applications. Journal of Computational Dynamics, 1(2):391–421, 2014.

[36] Hassan Arbabi and Igor Mezic. Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the koopman operator. SIAM Journal on Applied Dynamical Systems, 16(4):2096–2126, 2017.

[37] Florian Mormann, Thomas Kreuz, Ralph G Andrzejak, Peter David, Klaus Lehnertz, and Christian E Elger. Epileptic seizures are preceded by a decrease in synchronization. Epilepsy research, 53(3):173–185, 2003.

[38] Piotr Mirowski, Deepak Madhavan, Yann LeCun, and Ruben Kuzniecky. Classification of patterns of eeg synchronization for seizure prediction. Clinical neurophysiology, 120(11):1927–1940, 2009.

[39] Mojtaba Bandarabadi, César A Teixeira, Jalil Rasekhi, and António Dourado. Epileptic seizure prediction using relative spectral power features. Clinical Neurophysiology, 126(2):237–248, 2015.

[40] Ali H Shoeb and John V Guttag. Application of machine learning to epileptic seizure detection. In ICML, 2010.

[41] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. circulation, 101(23):e215–e220, 2000.

[42] Yun Park, Lan Luo, Keshab K Parhi, and Theoden Netoff. Seizure prediction with spectral power of eeg using cost-sensitive support vector machines. Epilepsia, 52(10):1761–1770, 2011.

[43] S.I. Dimitriadis, Dimitris Liparas, and Magda N. Tsolaki. Random forest feature selection, fusion and ensemble strategy: Combining multiple morphological mri measures to discriminate among healhy elderly, mci, cmci and alzheimer's disease patients: From the alzheimer's disease neuroimaging initiative (adni) database. Journal of Neuroscience Methods, 302:14–23, 2018.

[44] Beatriz Remeseiro and Veronica Bolon-Canedo. A review of feature selection methods in medical

applications. Computers in Biology and Medicine, 112:103375, 2019.

[45] Christopher JC Burges. A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery, 2(2):121–167, 1998.

[46] Thea Radüntz, Jon Scouten, Olaf Hochmuth, and Beate Meffert. Automated eeg artifact elimination by applying machine learning algorithms to ica-based features. Journal of neural engineering, 14(4):046004, 2017.

[47] Alexander Craik, Yongtian He, and Jose L Contreras-Vidal. Deep learning for electroencephalogram (eeg) classification tasks: a review. Journal of neural engineering, 16(3):031001, 2019.

[48] Rishi Raj Sharma, Piyush Varshney, Ram Bilas Pachori, and Santosh Kumar Vishvakarma. Automated system for epileptic eeg detection using iterative filtering. IEEE Sensors Letters, 2(4):1–4, 2018.

[49] Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, Xiaodi Huang, and Nasir Hussain. A review of epileptic seizure detection using machine learning classifiers. Brain informatics, 7:1–18, 2020.

[50] J Nathan Kutz, Xing Fu, and Steven L Brunton. Multiresolution dynamic mode decomposition. SIAM Journal on Applied Dynamical Systems, 15(2):713–735, 2016.