Analyzing Spatio-temporal Patterns of Epileptic EEG Signals by Dynamic Mode Decomposition

Jong-Hyeon Seo[1*] and Ichiro Tsuda[2]

[1] Academy of Emerging Science; Chubu University, Japan; hyeonni94@isc.chubu.ac.jp
[2] Academy of Emerging Science; Chubu University, Japan; tsuda@isc.chubu.ac.jp

Abstract: In this work, we apply a dynamic mode decomposition (DMD) to epileptic EEG data to figure out the spatio-temporal patterns in the data. It is found that some patterns with high frequency have direct influence on the epileptic seizure, and instantaneously capture the abnormal patterns of the neuron firing with high precision. We prepare an ictal and interictal EEG data for an epileptic patient, and investigating the dyamics of the DMD modes captured by the proposed algorithm for the windowed EEG signals. From the results we conclude that the DMD modes extracted from the EEG signals can be useful for analyzing and understanding the dynamics of the epileptic EEG data.

Keywords: epileptic seizure; dynamic mode decomposition; EEG; spatio-temporal pattern; pattern recognition, high frequency in signal.

1. Introduction

The dynamic mode decomposition (DMD) originated introduced by Schmid [14] in the fluid dynamics community as a method to decompose complex flows into a simple representation based on spatiotemporal coherent structures [14, 4, 15]. After its introduction, DMD was reframed by Rowley et al. [13] as a numerical technique to approximate the Koopman operator [9], establishing a strong connection to the analysis of nonlinear dynamical systems [2]. The DMD is an equation-free, data-driven method capable of providing an accurate decomposition of a complex system into spatiotemporal coherent structures that may be used for short-time future-state prediction and control [7]. Thus, DMD yields a set of modes along with a linear evolution model. The development of DMD is timely due to the concurrent rise of data science, encompassing a broad range of techniques, from machine learning and statistical regression to computer vision and compressed sensing [7].

Epileptic seizure is often signified by abnormal synchronization in neuronal firing either in a focal region or across several regions in the brain [8]. It is reported that the epileptic

zone may be localized by the abnormal signal pattern of the high-frequency domain so-called persistent high-frequency oscillations (HFO) [6, 10]. However, using conventional transformation algorithms to extract features from the EEG signals does not guarantee to achieve high accuracy [11]. In recent years, newly developed methods based on the DMD have been proposed in order to extract features from EEG data [7, 16]. The DMD estimates the complex frequencies and magnitudes (called "modes") corresponding to the signals. One of the advantage of DMD is that these modes are provided in linearly independent normalized vectors, so that the high-frequency mode pattern can be clearly detected.

In this paper, we present the feature extraction method using DMD modes for recognizing spatial-temporal patterns of the epileptic EEG data between ictal and interictal states. In section 2, we describe the structure of data and explain the process of feature extraction. Finally, in Section 3, we interpret spatial-temporal patterns of the features and the DMD modes through numerical results.

1.1. Background : Dynamic Mode Decomposition

Assume that data is collected from a dynamical system

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, t), \tag{1.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is a vector representing the state of a dynamical system at time t and $\mathbf{f}(\cdot)$ represent the dynmaics. In general, it is impossible to construct a solution to the nonlinear equation (1.1), so numerical solutions are used to approximate the dynamics and predict future states. The DMD procedure constructs the proxy, approximate locally linear dynamical system

$$\frac{d\mathbf{x}}{dt} = \mathcal{A}\mathbf{x} \tag{1.2}$$

with inital condition $\mathbf{x}(0)$ and solution

$$\mathbf{x}(t) = \sum_{k=1}^{n} \phi_k \exp(\omega_k t) b_k = \mathbf{\Phi} \exp(\mathbf{\Omega} t)\mathbf{b},$$

where ϕ_k and ω_k are the eigenvectors and eigenvalues of the matrix $\mathcal{A}$, and the coefficients b_k are the coordinates of $\mathbf{x}(0)$ in the eigenvector basis.

The continuous dynamics in (1.2) can be described by analogous discrete-time system sampled every Δt in time such as

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k, \quad k = 0, 1, 2, \ldots, m$$

where $\mathbf{A} = \exp(\mathcal{A}\Delta t)$ and m is the number of the states. The solution to this system is simply expressed by

$$\mathbf{x}_{k+1} = \sum_{j=1}^{n} \phi_j \lambda_j^k b_j = \mathbf{\Phi}\mathbf{\Lambda}^k \mathbf{b} \tag{1.3}$$

where λ_j and ϕ_j are the eigenvalues and eigenvectors of the discrete-time map $\mathbf{A}$ respectively, and $\mathbf{b}$ is the coefficient of the initial contion $\mathbf{x}_0\,(= \mathbf{x}(0))$ in the eigenvector basis so that

$\mathbf{x}_0 = \mathbf{\Phi b}$. The final goal of the DMD algorithm is to produce a low-rank eigendecomposition (1.3) of the matrix $\mathbf{A}$ that optimally fits $\mathbf{x}_k$ for $k = 0, 1, \ldots, m$ in a least-sqaure sense so that

$$\|\mathbf{x}_{k+1} - \mathbf{A}\mathbf{x}_k\|_2 \tag{1.4}$$

is minimized for $k = 0, 1, \ldots, m-1$.

Let $X = [\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_m] \in \mathbb{R}^{n\times(m+1)}$ be the data matrix, then to minimize the approximation error (1.4), we arrange the matrix into two data matrices

$$\begin{aligned} X_1 &= [\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_{m-1}] \in \mathbb{R}^{n\times m}, \\ X_2 &= [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m] \in \mathbb{R}^{n\times m}. \end{aligned} \tag{1.5}$$

Then the locally linear approximation (1.2) can be written by

$$X_2 \approx \mathbf{A}X_1$$

and the best-fit $\mathbf{A}$ matrix is given by

$$\mathbf{A} = X_2 X_1^{\dagger}, \tag{1.6}$$

where $\dagger$ denotes the Moore-Penrose pseudoinverse. The solution (1.6) minimize the error

$$\|X_2 - \mathbf{A}X_1\|_F,$$

where $\|\cdot\|_F$ is the Frobenius norm [12] given by

$$\|M\|_F = \sqrt{\sum_{j=1}^{p}\sum_{k=1}^{q} M_{jk}^2}$$

for $M \in \mathbb{R}^{p\times q}$. Thus, the DMD of the pair (X_1, X_2) is the eigendecomposition of the matrix $\mathbf{A}$. However, in practice, if the state dimension n is large, the matrix $\mathbf{A}$ may be intractable to analyze directly. Instead, we compute a low-rank approximation $\tilde{\mathbf{A}}$ by the following algorithm [15]:

(1) Compute the reduced and appropriately truncated SVD [1] of the data matrix X_1 in (1.5)

$$X_1 \approx U_r \Sigma_r V_r^*, \tag{1.7}$$

where the columns of $U_r \in \mathbb{R}^{n\times r}$ and $V_r \in \mathbb{R}^{m\times r}$ are orthonormal eigenvectors of $X_1X_1^T$ and $X_1^TX_1$, respectively, the diagonal entries of $\Sigma_r \in \mathbb{R}^{r\times r}$ are the square roots of the non-negative eigenvalues of both $X_1X_1^T$ and $X_1^TX_1$, and $r(< \min n, m)$ refers to the reduced rank of the approximated matrix given in (1.7). It is notable that the columns of U_r are called the left singular vectors of X_1.

(2) Define a low-rank approximation $\tilde{\mathbf{A}}$ of $\mathbf{A}$ in (1.6)

$$\tilde{\mathbf{A}} = U_r^* \mathbf{A} U_r = U_r^* X_2 V_r \Sigma_r^{-1}. \tag{1.8}$$

(3) Compute the eigendecomposition of $\tilde{\mathbf{A}}$ in (1.8)

$$\tilde{\mathbf{A}}W = W\mathbf{\Lambda}_r, \quad \mathbf{\Lambda}_r = \operatorname{diag}(\lambda_1, \lambda_{2,} \cdots, \lambda_r) \tag{1.9}$$

where the columns of $W \in \mathbb{C}^{r\times r}$ and the diagonal entries of $\mathbf{\Lambda}_r \in \mathbb{C}^{r\times r}$ are the eigenvectors and the eigenvalues of $\tilde{\mathbf{A}}$ respectively.

(4) Since the eigenvalues in (1.9) is also the eigenvalues of $\mathbf{A}$, the DMD mode (i.e. the eigenvector of $\mathbf{A}$) corresponding to the DMD eigenvalue λ_i in (1.9) is given by

$$\tilde{\phi}_i \equiv U_r w_i,$$

where w_i is the i^{th} column of W in (1.9).

Each snapshot $\mathbf{x}_k$ is approximated by $\mathbf{x}_k \approx \mathbf{A}\mathbf{x}_{k-1}$, and so, this algorithm allows us to express a local approximtation of a windowed signal as a composition of a coupled spatio-temporal model

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_0 = \mathbf{\Phi}_r \mathbf{\Lambda}_r^k \mathbf{c}, \tag{1.10}$$

where $\mathbf{\Phi}_r \in \mathbb{C}^{n\times r}$ consists of the columns of $\tilde{\phi}_i$ and $\mathbf{c}$ is the coefficient of the initial condition $\mathbf{x}_0$ satisfying $\mathbf{x}_0 = \mathbf{\Phi}_r \mathbf{c}$ [7]. The phase of eigenvalues can be converted to frequency (Hz) by

$$f_i = \frac{\Im(\log(\lambda_i)/\Delta t)}{2\pi}, \tag{1.11}$$

where $\Im(\cdot)$ is the imaginary part of a complex number. f_i represents the frequency of oscillation of mode $\tilde{\phi}_i$ in units of cycles per second.

The spatial resolution for neurological signals is usually less than the temporal resolution, e.g. we have 49 electrodes sampling at 2000Hz. Therefore the standard DMD algorithm represented above must be modified to capture the dynamics of the neurologocial activity properly. The modification is adopted to augment the data matrix X by stacking h number of observation vectors $\mathbf{x}_k$ such that the number of rows in X becomes at least twice the number of columns. The augmented matrix for the data matrix X with stacking number h is represented by

$$X_{\text{aug}} = \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \cdots & \mathbf{x}_{m-h+1} \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{m-h+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_h & \mathbf{x}_{h+1} & \cdots & \mathbf{x}_m \end{bmatrix} \in \mathbb{R}^{nh\times(m-h+1)}. \tag{1.12}$$

More details are in [3]. Then the DMD mode amplitude is defined by

$$P_i = 2|c_i|/\sqrt{h} \tag{1.13}$$

where c_i is i^{th} entry of $\mathbf{c}$ given in (1.10). Meanwhile, the DMD mode power is defined by $\left\|\tilde{\phi}_i\right\|_2^2$ [3, 16].

Figure 1.1 illustrates the variation in spectral information and DMD modes corresponding to signals from different data regions of the EEG recordings such as interictal and ictal states. The signals used in this work are obtained from 49 channel with one-second duration and they are resampled by a sampling rate of 500Hz with 50 stacking number defined in 1.12 , hence, the data matrix X contains 2450 columns and 450 rows (the original sample has 500 columns and 49 rows.). In Figure 1.1 we see that while the amplitudes at high frequencies are relatively

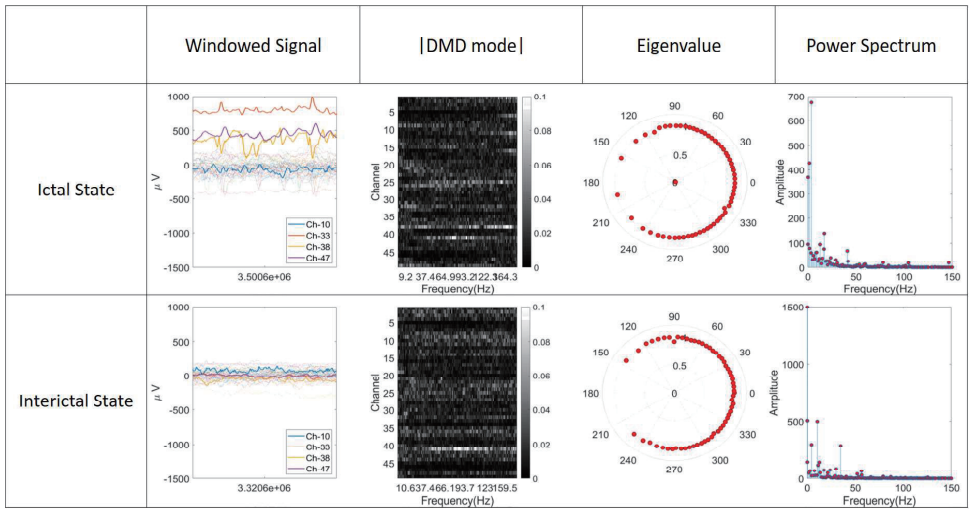

FIGURE 1.1. Examples of the patterns of DMD modes and power spectrums for ictal signals and interictal signals.

much smaller than the amplitudes at other frequencies, the patterns of high-frequency of the DMD mode are more significant than those of other frequencies. Beside high frequency DMD mode patterns are relatively significant than low. Since our purpose is to detect anormal spatial-temporal patterns of high frequency from electrode signals, we ignore these biased spectral information to focus on DMD mode patterns itselves in this study. Therefore, in order to detect abnormal spatial-temporal patterns of high frequencies from electrode ictal and interictal signals, we ignore the biased spectral information to focus on DMD mode patterns themselves in this study.

2. Material and Method

In this paper, a kind of EEG data, called ECoG, is used. The ECoG recordings are sampled at intervals of 0.2 second, and they are segmented with one second duration. Each windowed signal is decomposed by the DMD into a dynamic mode, and the columns of the mode are sorted according to their frequencies and then vectorized. Those vectors are projected to the approximated left singular vector space so that the features are extracted.

2.1. Dataset

The ECoG recording of an epileptic patient was provided by Ikeda laboratory at Kyoto University. The recording is measured through 49 channels and composed of only two types of signals: an ictal signal and an interictal signal. Since the sampling rate of the original data provided is very large as 2000Hz, it has been reduced to 500Hz to match the rate used in most of the previous related studies.

2.2. Feature Extraction

Each DMD mode is used for the feature extraction after being vectorized and aligned. The left singular vectors are obtained by decomposing the matrix constructed with the modes. Then, the features are extracted by projecting the computed modes to the truncated left singular vector space which is prepared in the previous states. The feature to be used in the proposed method is a vectorized and aligned DMD mode. The dimension of the feature is reduced by a linear transformation mapped to the space of truncated left singular vectors, which come from a row-wise stacked feature matrix of modes. The detail of feature extraction is explained below.

For simplicity of notation without loss of generality we define the signal set from windowed regions of EEG recordings as follows

$$\mathcal{W} = \{X^{(j)} \in \mathbb{R}^{n\times(m+1)} | \, j = 1, 2, \cdots, J\} \tag{2.1}$$

where n and $(m+1)$ are the number of channels and that of snapshots within one second durations at 0.2 second intervals, respectively. Note that J is the number of the windowed signals in the set.

Let $X^{(j)} = \left[\mathbf{x}_0^{(j)}, \mathbf{x}_1^{(j)}, \cdots, \mathbf{x}_m^{(j)}\right] \in \mathcal{W}$ be a j-th data matrix of the signal set in (2.1). Then we denote the augmented matrix $X_{\text{aug}}^{(j)}$ defined in (1.12). For each $X_{\text{aug}}^{(j)}$ we calculate $\mathbf{\Phi}^{(j)} \in \mathbb{R}^{nh\times r}$, $f^{(j)}$ and $P^{(j)}$ defined in (1.10), (1.11) and (1.13) respectively. Here for the purpose of extracting optimal features, the low-rank approximation size r is fixed (in this paper $r = 200$). The discussion for choosing an optimal threshold on the singular values can be found in [5]. After that we get the aligned modes by frequency as follows

$$\mathbf{\Phi}_{\text{align}}^{(j)} = \left[\phi_{\sigma(1)}^{(j)}, \phi_{\sigma(2)}^{(j)}, \cdots, \phi_{\sigma(r)}^{(j)}\right] \tag{2.2}$$

where σ is a permutation of $\{1, 2, \cdots, r\}$ satisfying

$$f_{\sigma(1)}^{(j)} \geq f_{\sigma(2)}^{(j)} \geq \cdots \geq f_{\sigma(r)}^{(j)}.$$

In order to reduce the computational cost in featuring process without loss of the information one can extract submatrix by truncating the rows by

$$\mathbf{\Phi}_{\text{trunc}}^{(j)} = \mathbf{\Phi}_{\text{align}}^{(j)}[1 : n, 1 : r] \tag{2.3}$$

where $Z[\mathbf{u}, \mathbf{v}]$ denotes a submatrix of Z indexed by sequences $\mathbf{u}$ and $\mathbf{v}$, and $i : k$ implies the index sequences $\{i, i+1, i+2, \cdots, k\}$ satisfying $i \leq k$ respectively. Note that in this paper the size of each $\mathbf{\Phi}_{\text{trunc}}^{(j)}$ is given by 49×200 for $j = 1, 2, \ldots, J$. The remarkable point is that most of entries of $\mathbf{\Phi}_{\text{trunc}}^{(j)}$ are close to zero (or almost sparse matrices) and can be clustered by small groups. So we suppose that dynamics of every signals from the EEG recording can be explained by a few entry of DMD modes. It means that we can apply low rank reduction schemes to the modes such as singular value decomposition (or principle component analysis).

From the windowed signals, the corresponding truncated DMD modes in (2.3) are gathered to form libraries. Construction of the libraries is performed as follows

$$\mathcal{L} = \left[\begin{array}{cccc} \tilde{\phi}_1 & \tilde{\phi}_2 & \cdots & \tilde{\phi}_J \end{array}\right] \in \mathbb{R}^{nr\times J}, \tag{2.4}$$

where

$$\tilde{\phi}_j = \mathbf{vec}\left(\mathbf{\Phi}^{(j)}_{\text{trunc}}\right) \in \mathbb{R}^{nr}.$$

$\mathbf{\Phi}^{(j)}_{\text{trunc}}$ is a algined DMD mode matrix in (2.3) for the j-th signal of $\mathcal{W}$ in (2.1), and the $\mathbf{vec}\,(M)$ represents a column vector obtained by rearranging $M = \begin{bmatrix} m_1 & m_2 & \cdots & m_k \end{bmatrix} \in \mathbb{C}^{p\times q}$with vertically stacking the column vectors m_i of ther matrix below in the order of the index, that is,

$$\mathbf{vec}(M) = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_k \end{bmatrix} \in \mathbb{C}^{pq\times 1}.$$

Here we only consider the absolute values of DMD modes. Thus the features deternmined in d-dimensional left singular vector space is given by the transformation of each column of $|\mathcal{L}|$ onto the first d left singular vectors of $|\mathcal{L}|$ such as

$$|\mathcal{L}| = U_{\mathcal{L}}\Sigma_{\mathcal{L}}V^*_{\mathcal{L}} \quad \text{and} \quad F_{\mathcal{L}} = U_d^T|\mathcal{L}| \in \mathbb{R}^{d\times J} \tag{2.5}$$

where $U_{\mathcal{L}} \in \mathbb{R}^{n\times n}$ and $V_{\mathcal{L}} \in \mathbb{R}^{J\times J}$ are matrices orthogonal to each other, $\Sigma_{\mathcal{L}} \in \mathbb{R}^{n\times J}$ is a non-negative diagonal matrix, U_d^T is the transpose of the first d columns of $U_{\mathcal{L}}$, and columns of $F_{\mathcal{L}} \in \mathbb{R}^{d\times J}$ are regarded as the features of the signal set, respectively.

3. Result

Figure 3.1 and Figure 3.2 shows spatio-temporal patterns of the DMD modes and the features captured by the proposed algorithm for the windowed signals of the EEG recordings respectively.

In Figure 3.1, the DMD modes of the channel 33 and 47 capture the HFO phenomenon at the front of the ictal onset. Besides that of channel 38 derives significant values in the ictal area. Figure 3.2 exhibits the dynamics of the features generated by projection to the left singular vector space of the library $\mathcal{L}$ in 2.4. The graph shows the feature components associated with the high-frequency patterns of the ictal and interictal signals. The components corresponding the third and fifth left singular vectors indicate the patterns of correlated channels and frequencies at front of the ictal onset. Besides, that of the second left singular vectors shows corelated channels and frequencies in the ictal area.

Acknowledgments

This work was partially supported by Grant-in-Aid for JST Strategic Basic Research Programs (An Exploration of the Principle of Emerging Interactions in Spatio-temporal Diversity, CREST Grant Number JPMJCR17A4) and Scientific Research on Innovative Areas (Non-linear Neuro-oscillology: Towards Integrative Understanding of Human Nature, KAKENHI, 15H05878) from the Ministry of Education, Culture, Sports, Science and Technology, Japan. This work was partially supported by the Research Institute for Mathematical Sciences, an International Joint Usage/Research Center located in Kyoto University.

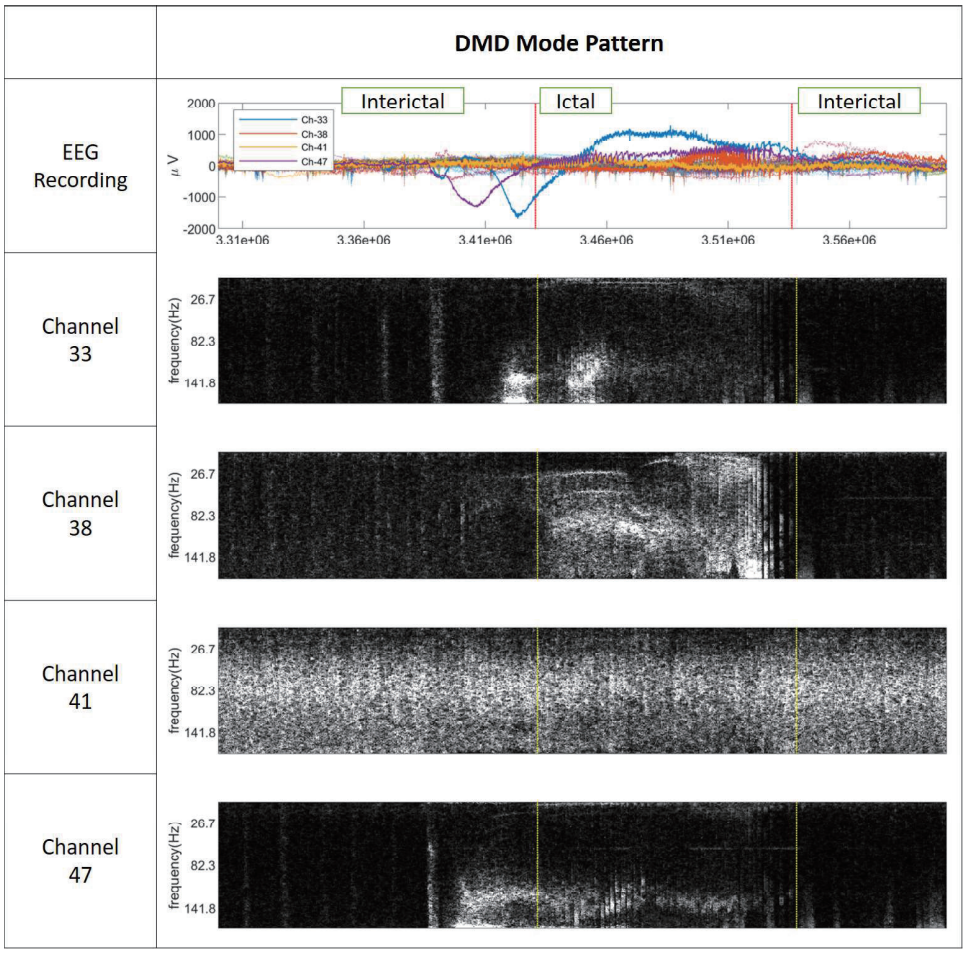

FIGURE 3.1. Spatiotemporal patterns of the DMD modes for EEG data. The first row is the sample of the orignal EEG recordings containing ictal and interictal states. The second through fifth graphs represent the dynamics of the DMD mode values of the associated channels, respectiviely.

References

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical System*, SIAM: Philadelphia, U.S.A., 2005.

[2] Z. Bai, E Kaiser, J. L. Proctor, J. N. Kutz and S. L. Brunton, *Dynamic mode decomposition for compressive system identification*, AIAA Journal, 2019, 1˜14.

[3] L. A. Brunton, L.A., J. G. Johnson, J. G. Ojemann and J. N. Kutz, *Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition,* J. Neurosci. Methods. 2016, 258, 1˜15.

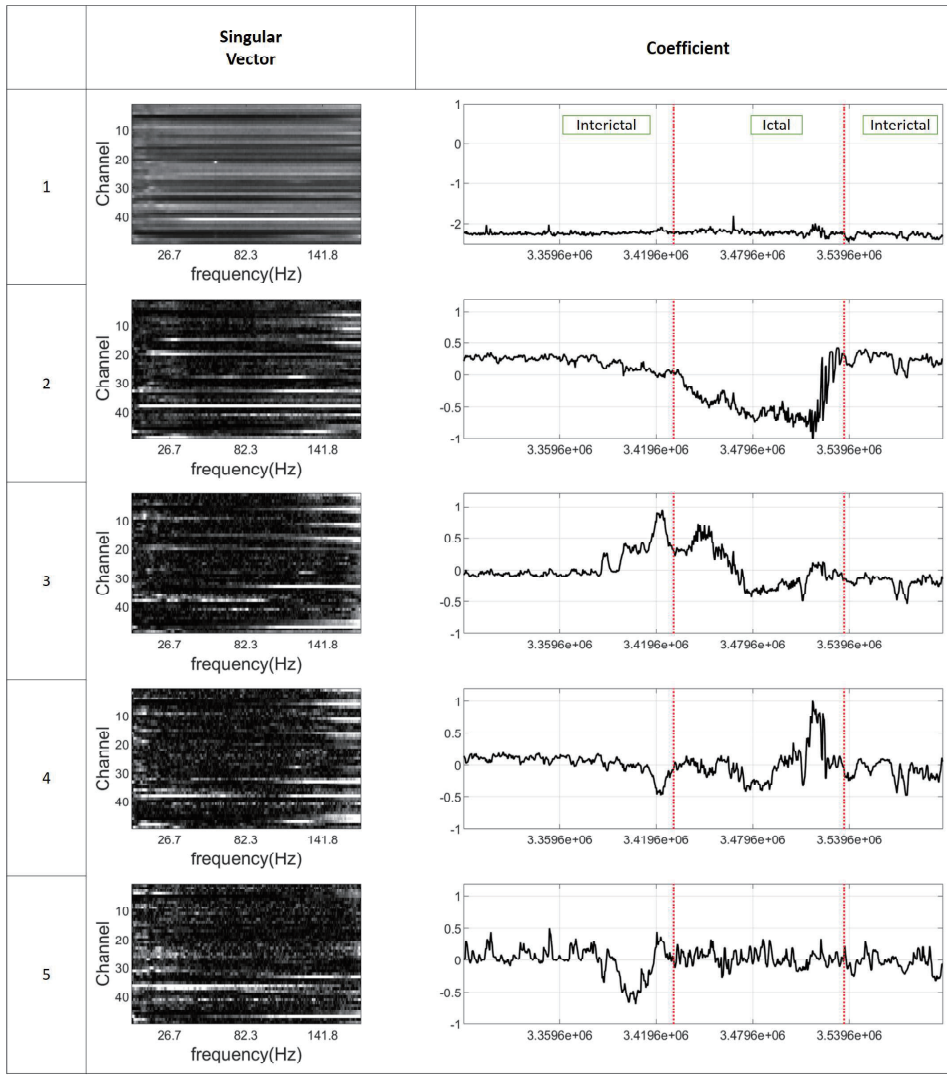

FIGURE 3.2. Spatiotemporal patterns of the features for EEG data. The first column and the second of each row shows the pattern of the left singular vector and the dynamics of the feature corresponding to $U_{\mathcal{L}}$ and $F_{\mathcal{L}}$ in 2.5, respectively.

[4] K. K. Chen, J. H. Tu, and C. W. Rowley, *Variants of dynamic mode decomposition: Boundary condition, Koopman, and Fourier analysis*, J. Nonlinear Sci.,2012, 22, 887˜915.

[5] M. Gavish, and D. L. Donoho, *The optimal hard threshold for singular values is* $4/\sqrt{3}$, IEEE T. Inform. Theory, 2014, 60(8), 5040˜5053.

[6] T. Inoue, M. Inouchi, M. Matsuhashi, R. Matsumoto, T. Hitomi, M. Daifu-Kobayashi, K. Kobayashi, M. Nakatani, K. Kanazawa, A. Shimotake, T. Kikuchi, K. Yoshida, T. Kunieda, S. Miyamoto, R. Takahashi and A. Ikeda, *Interictal Slow and High-Frequency Oscillations: Is it an Epileptic Slow or Red Slow?,* J. Clin. Neurophysiol. 2019, 36, 166˜170.

[7] J. N. Kutz, S. L. Brunton, B. W. Brunton and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, SIAM: Philadelphia, U.S.A., 2016.

[8] K. K. Majumdar, *Automatic seizure detection in ECoG by differential operator and windowed variance*, IEEE Trans. Neural. Syst. Rehabil. Eng., 19(4), 2011.

[9] I. Mezić, *Spectral properties of dynamical systems, model reduction and decompositions*, Nonlinear Dynamics, 2005, 41, 309˜325.

[10] J. G. Ochoa and W. G. Rusyniak, *Description of Ictal HFO Mapping in Patients with Both Temporal and Extratemporal Seizure Focus*, Neurol. Res. Int., 2016, 5, 1˜4.

[11] M. Z. Parvez and M. Paul, *Classification of Ictal and Interictal EEG signals. Proceedings of the 10th IASTED Conference on Biomedical Engineering*, Innsbruck, Austria, 13˜15 Feb. 2013; ACTA Press: Alberta, Canada, 2013, 113˜141.

[12] G. Golub and C. van Loan, *Matrix Computations*, JHU Press, Fourth edition, 2013.

[13] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, *Spectral anlaysis of nonlinear flows*, J. Fluid Mech., 2009, 641, 115˜127.

[14] P. J. Schmid, *Dynamic mode decomposition of numerical and experimental data.* J. Fluid Mech. 2010, 656, 5˜28.

[15] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton and J. N. Kutz, *On dynamic mode decomposition: theory and applications.*, J. Comput. Dyn. 2014, 1, 391˜421.

[16] M. S. J. Solaija, S. Saleem, K. Khurshid, S. A. Hassan and A. M. Kamboh, *Dynamic Mode Decomposition Based Epileptic Seizure Detection from Scalp EEG*, IEEE Access, 2018, 6, 38683˜38692.