2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Detecting Epileptic Seizures via Non-Uniform Multivariate Embedding of EEG Signals

Haidong Gu and Chun-An Chou

Abstract— **Efficient real-time detection of epileptic seizures remains a challenging task in clinical practice. In this study, we introduce a new thresholding method to monitor brain activities via a non-uniform multivariate (NUM) embedding of multi-channel electroencephalogram (EEG) signals. Specifically, we present a NUM embedding optimization problem to identify the best embedding parameters. We originate one feature, named non-uniform multivariate multiscale entropy (NUMME), which is extracted from the NUM embedded EEG data. Finally, the extracted feature, compared to an individualized threshold, is used for monitoring and detecting seizure onsets. Experimental results on the real CHB-MIT Scalp EEG database show that our approach achieves a comparable performance to the state-of-art methods. Moreover, it is important to note that we accomplish this without using any sophisticated machine learning algorithms.**

Clinical relevance— **This decision support tool provides a patient-specific measurement of brain complexity for real-time seizure detection at 96% sensitivity rate.**

I. Introduction

Epilepsy is one of most common neurological disorders, which is characterized by recurrent seizures [1]. According to the recent statistics of World Health Organization (WHO), more than 50 million people across different ages suffer from epilepsy in the world [2]. Failure to detect the seizure onsets in time may lead to serious accidents and even threatening the lives of patients. However, up to date, prompt detection of epileptic seizures still remains a challenging task. Scalp electroencephalography (EEG) is a non-invasive technique for real-time recording of neuronal electrical activities. Multi-channel EEG signals in the form of multivariate time series are capable of representing underlying dynamics of the brain system, and thus are widely used for seizure detection. Due to the non-linearity and non-stationarity nature of EEG time series, traditional signal processing techniques such as Fourier transform and wavelet transform are not applicable. Therefore, a variety of nonlinear time series analysis have been implemented to characterize the epileptic patterns of EEG signals [3]–[5].

The state space reconstruction of complex systems by using observed time series is the foundation of nonlinear time series analysis. In past decades, several methods have been proposed to address the reconstruction problem, and time-delay embedding is the most popular approach due to its simplicity and efficiency. The majority of existing embedding schemes are based on uniform univariate (UU) embedding on univariate time series [6]. Nonetheless, there are several limitations regarding the UU embedding. Especially, when it comes to embedded systems inheriting multiple time scales structure, multivariate time series and noise contaminated input data. It is noteworthy that using non-uniform time delays could ensure a successful embedding of systems exhibiting a heterogeneous temporal scales behaviour. In the past decades, some methods including non-uniform time delays with extension to multivariate time series have been introduced [7], [8]. However, we found that their applications are rather limited and not popular compared to the UU embedding. This could be due to the sophistication nature which makes it hard to be implemented and high computational cost for high-dimensional complex times series data. To this end, we introduce a new effective non-uniform multivariate (NUM) embedding algorithm in this work.

To monitor and detect seizure onsets, various informative feature extraction methods have been developed for characterizing epileptic patterns from embedded EEG signals, including recurrence quantification analysis (RQA) and algebraic connectivity analysis [4]. Inspired by multiscale entropy (MSE) analysis [9], [10], we will introduce a new feature extraction technique named non-uniform multivariate multiscale entropy (NUMME) and demonstrate its superiority over existing methods.

Furthermore, in practice, the above mentioned feature extraction methods are usually integrated within a state-of-the-art machine learning framework (e.g., neural networks, support vector machine, logistic regression) [3], [11]–[15]. These elaborate frameworks have been proven to significantly improve the seizure detection accuracy, but the greatest shortcoming is that they are not applicable for real-time seizure detection due to high computational cost.

In this study, we propose a novel computationally efficient framework suitable for real-time seizure detection, which can be applied in automated seizure control system for clinical treatments. We present the empirical results based on the benchmark CHB-MIT Scalp EEG database [16], comparing to the state-of-the-art methods.

II. Related Work

For the state space reconstruction, an embedding vector should satisfy two properties: (1) $Relevance$: it is able to predict the dynamics of the system, in other words, it is relevant to the future states of the system and (2) $Nonredundancy$: its components are as independent as

The authors are with the Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: gu.hai@northeastern.edu; ch.chou@northeastern.edu).

This work is supported by Burroughs Wellcome Fund (Award No.1019976) and Tufts CTSI NIH Clinical and Transnational Science Award (UL1TR002544).

possible to each other. In previous work, several NUM embedding algorithms were developed [7], [8] to optimize the embedding with a trade-off between $Nonredundancy$ and $Relevance$. However, there are two major limitations in their work. Firstly, they put the same weights on $Nonredundancy$ and $Relevance$, whereas in practice we are more interested in whether the reconstructed vector can well explain the underlying dynamics. Hence, we will assign more weights to $Relevance$ in our model. Secondly, for the $Relevance$ objective, they only consider the performance of predicting future state in one step forward, which is not enough to disclose the relevance property. Instead, we will extend the prediction horizon to several time steps ahead.

Quantifying the complexity of signals (e.g., EEG) is a prerequisite for understanding the mechanisms of signal generating systems (e.g., human brain). In the literature, several methods have been developed to measure the complexity of physical and biological time series [9], [10]. The multiscale entropy (MSE) analysis is a univariate technique and only capable of measuring the complexity of single time series [9]. The recently proposed multivariate multiscale entropy (MMSE) method generalized the MSE scheme for multivariate time series [10]. However, both of MSE and MMSE are based on the uniform embedding, and obviously, there are limitations on such embedding like mentioned before. To bridge the gap, we herein introduce the NUMME criterion based on our proposed NUM embedding.

III. METHODOLOGY

In this section, we present a framework for real-time seizure detection. First of all, we introduce an automated NUM embedding method based on a combinatorial optimization problem. Secondly, we propose a feature, i.e., NUMME, for seizure detection by adapting the concept from MSE analysis [9], [10]. The NUMME feature is extracted from the NUM embedded EEG data. Finally, we show a simple thresholding approach applied on the extracted NUMME feature for monitoring seizure onsets.

A. Non-Uniform Multivariate (NUM) Embedding

Given a multivariate time series $\{x_{n,t}\}$, where channel $n = 1 \ldots N$ and time point $t = 1 \ldots T$, the NUM embedding vector at time t is written as

$$\boldsymbol{x}_t = \left(x_{1,t+\tau_{11}}, \ldots, x_{1,t+\tau_{1m_1}}, \ldots, x_{N,t+\tau_{Nm_N}} \right), \tag{1}$$

where τ_{ij} is the j-th lag (time delay) for the channel i which is selected by solving an optimization problem shown later.

In our NUM embedding algorithm, we choose to use mutual information to evaluate $Nonredundancy$ and $Relevance$. The mutual information is a fundamental notion in information theory. It measures the mutual dependence between two random variables or time series.

With the maximum embedding lags L_n for each time series $n = 1, \ldots, N$, we have a set of $\sum_{n=1}^{N} L_n$ candidate lags. In our experiment, we chose $L_n = 2$. The main goal is to find an optimal combination of these candidates to build the embedding vector. This is a combinatorial optimization problem, and it is computationally intractable to investigate all the possible combinations. Thus, we propose an iterative scheme shown below to handle it. We choose one candidate lag in each iteration and progressively expand the embedding vector until meeting a stopping criterion.

Start with an initial empty embedding vector $\boldsymbol{e}_0$. At iteration i, note that we already have a reconstructed vector $\boldsymbol{e}_{i-1} = (x_1^*, x_2^*, \ldots, x_{i-1}^*)$ of dimension $i-1$, and meanwhile the feasible set of candidate lags shrinks to $\mathcal{S}_i = \{x_{1,t}, \ldots, x_{1,t+L_1}, x_{2,t}, \ldots, x_{N,t+L_N}\} \setminus \{x_1^*, \ldots, x_{i-1}^*\}$. Then, we choose an optimal lag x_i^* to add into the embedding vector by solving the following optimization problem (2).

$$\begin{aligned} \max_{x_i^* \in \mathcal{S}_i} \ & \lambda \left\{ \frac{1}{N} \sum_{n=1}^{N} \left(\frac{1}{P_n} \sum_{p=1}^{P_n} \boldsymbol{I}\left(x_{n,t+L_n+p};\ x_i^*\right) \right) \right\} \\ & + (1-\lambda) \left\{ - \sum_{j=1}^{i-1} \boldsymbol{I}\left(x_j^*;\ x_i^*\right) \right\}, \end{aligned} \tag{2}$$

where $\boldsymbol{I}(\cdot;\cdot)$ represents mutual information, N is the number of channels, P_n indicates the maximum prediction horizon for the n-th channel, and λ is the trade-off parameter. We chose $P_n = 2$ and $\lambda = 0.9$ in our experiment.

The first term weighted by λ in model (2) represents the $Relevance$ objective, while the second term in model (2) resembles the $Nonredundancy$ objective. We observe that the first term in (2) is large as the future states $x_{n,t+L_n+p}$ are dependent on the new added lag x_i^*, while the second term in (2) is large when the candidate x_i^* is independent to the already selected components x_j^* in the embedding vector $\boldsymbol{e}_{i-1}$. Therefore, our model (2) can generate an optimal embedding vector that maximizes the satisfaction of two favorable properties for phase space reconstruction. Actually, it gives the best trade-off between $Relevance$ and $Nonredundancy$.

To terminate the above iterative algorithm, we choose the false nearest neighbor (FNN) method [6] as the stopping criterion. The progressive building of embedding vector ends when the percentage of false nearest neighbors going from $\boldsymbol{e}_{i-1}$ to $\boldsymbol{e}_i$ at iteration i is smaller than a properly chosen threshold (e.g., 0.01 in the experiment). The final embedding vector is then $\boldsymbol{e}_{i-1}$.

B. Non-Uniform Multivariate Multiscale Entropy (NUMME)

Given a N-variate time series $\{x_{n,t}\}$, where $n = 1, \ldots, N$ represents the channel index and $t = 1 \ldots T$ indicates the time point, we can evaluate its NUMME measure of complexity in the following three steps.

1) Set the time scale range. For each scale factor s in the range, we construct a non-overlapping coarse-grained time series $y_{n,i}^s = \frac{1}{s} \sum_{t=(i-1)s+1}^{is} x_{n,t}$, where $i = 1, \ldots, \lfloor \frac{T}{s} \rfloor$ and $n = 1, \ldots, N$.
2) Compute non-uniform multivariate sample entropy (NUMSampEn) for each coarse-grained multivariate time series $\{y_{n,i}^s\}$.
3) Perform cumulative summation of NUMSampEn over the temporal scale range to obtain the NUMME measure for each scale factor s.

Algorithm 1 Non-Uniform Multivariate Sample Entropy

Set the threshold $r = 1$ as the standard deviation of the standardized data.
Construct $(T-L)$ NUM embedding vectors $\boldsymbol{x}_M(i) \in \mathbb{R}^M$, where $i = 1 \ldots T-L$ and $L = \max(\boldsymbol{\tau}) + 1$.
for $i = 1 \ldots T-L$ **do**
  $B_i^M(r) \leftarrow \frac{1}{T-L-1} \sum_{j=1, j \neq i}^{T-L} \mathbb{1}\left(d\left(\boldsymbol{x}_M(i), \boldsymbol{x}_M(j)\right) \leq r\right)$
end for
$B^M(r) \leftarrow \frac{1}{T-L} \sum_{i=1}^{T-L} B_i^M(r)$
Construct a total of $p(T-L)$ vectors $\boldsymbol{x}_{M+1}^n(i) \in \mathbb{R}^{M+1}$ by separately increasing the embedding dimension of $\boldsymbol{x}_M(i)$ from m_n to m_n+1 (i.e., add a new lag $\tau_{n,m_n}+1$) for each channel $n \in \mathcal{S} = \{n|1 \leq n \leq N, m_n > 0\}$, where $p = \sum_{n=1}^{N} \mathbb{1}\left(m_n > 0\right)$.
for $n \in \mathcal{S}$ **do**
  for $i = 1 \ldots T-L$ **do**
    $B_i^{M+1}(r) \leftarrow \frac{1}{T-L-1} \sum_{j=1, j \neq i}^{T-L} \mathbb{1}\left(d\left(\boldsymbol{x}_{M+1}^n(i), \boldsymbol{x}_{M+1}^n(j)\right) \leq r\right)$
  end for
  $A_n^M(r) \leftarrow \frac{1}{T-L} \sum_{i=1}^{T-L} B_i^{M+1}(r)$
end for
$A^M(r) \leftarrow \frac{1}{p} \sum_{n \in \mathcal{S}} A_n^M(r)$
$\text{NUMSampEn}\left(\boldsymbol{m}, \boldsymbol{\tau}, r, T\right) \leftarrow -\log \frac{A^M(r)}{B^M(r)}$

To obtain NUMSampEn in Step 2), recall from the NUM embedding vector (1), let us define $\boldsymbol{m} = (m_1, m_2, \ldots, m_N) \in \mathbb{R}^N$ as the embedding dimension vector and $M = \sum_{n=1}^{N} m_n$ as the dimensionality of the NUM embedding vector. Also, we denote $\boldsymbol{\tau} = (\tau_{1,1}, \ldots, \tau_{1,m_1}, \tau_{2,1}, \ldots, \tau_{N,m_N}) \in \mathbb{R}^M$ as the embedding lag vector. Henceforth, we rewrite $\boldsymbol{x}_t$ in (1) as $\boldsymbol{x}_M(t) \in \mathbb{R}^M$. Then the method for calculating NUMSampEn of a N-variate time series $\{x_{n,t}\}_{t=1}^{T}$, $n = 1, \ldots, N$ is presented in Algorithm 1, which represents a natural extension of uniform univariate sample entropy (SampEn) [17] and uniform multivariate sample entropy (MSampEn) [10]. Note that we use $\mathbb{1}(\cdot)$ as the indicator function (returns 1 if condition true, otherwise 0), and $d(\cdot)$ as the maximum norm distance function.

C. Control-Chart-based Thresholding for Seizure Detection

The feature values extracted from each time window of multivariate EEG signals are concatenated in time order to form a new univariate time series of feature. Sharp fluctuations are expected in the feature time series, as we use a short time window of one second. Hence, we implement a three-point moving average on it for smoothing purpose. Then, both upper threshold ($UT = \hat{\mu} + 2\hat{\sigma}$) and lower threshold ($LT = \hat{\mu} - 2\hat{\sigma}$) are applied on the smoothed univariate feature time series for detecting abnormal events of seizure onset, where $\hat{\mu}$ and $\hat{\sigma}$ are mean and standard deviation estimates, respectively, of 180 time windows (i.e., 3 minutes) located at the beginning of the corresponding feature time series. The seizure onset alarm is triggered when 2 of 3 points rise above UT or fall below LT on the same side.

IV. Experimental Results

We test the proposed seizure detection framework on the benchmark CHB-MIT Scalp EEG Database [16]. This public database consists of multi-channel EEG recordings of 23 epileptic patients from Children's Hospital Boston. Each subject's EEG signals were recorded continuously and divided into several sessions with length of one hour. The sampling rate of 256 Hz and around 23 electrodes, which constitute the EEG channels, were used for the recording. In EEG recordings of all 129 sessions of 23 subjects, there were 182 epileptic seizures marked by clinician experts.

The whole procedure of our framework is depicted as follows. Firstly, we preprocess the multivariate EEG data by a band-pass filter of $1-50$ Hz to reject artifacts, and then standardize each channel to mean of 0 and standard deviation of 1. Secondly, we perform the NUM embedding on the multivariate EEG data for each subject. Next, we slice the NUM embedded EEG time series into non-overlapping time windows of one second length, namely 256 time points due to the 256 Hz sampling rate. Then, we extract features (e.g., NUMME) from each time window and concatenate in time order to build a new univariate time series of feature. Finally, we apply the constant thresholding method on the obtained feature time series for real-time detection of epileptic seizures.

To evaluate the performance of our method, we use three common metrics for seizure detection tasks defined below.

1) *Sensitivity*: Ratio of truly detected seizure states to total seizure states.
2) *False Alarm Rate (FAR)*: Ratio of falsely detected seizure states to total normal states.
3) *Latency*: Time lag in seconds between the start of seizure onset and the earliest detection.

We compare our proposed NUMME feature to a selection of RQA features (e.g., LMAX, TND, LAM, TT) and algebraic connectivity feature (λ_2) [4] which are widely used for seizure detection in the literature. In addition, we also compare the effect of our introduced NUM embedding to well-established uniform multivariate (UM) embedding on the seizure detection performance. Furthermore, we make the comparison between our seizure detection framework with the state-of-art approaches which usually include sophisticated machine learning techniques and numerous complex features [3], [5], [11]–[15], [18].

The overall comparison results with respect to the performance metrics averaged over all subjects are summarized in Table I. The monitor chart of selected feature signals for real-time seizure detection is displayed in Figure 1.

V. Discussions

In Table I, we find the NUMME feature consistently outperforms other features by yielding the highest *Sensitivity*, the lowest *False Alarm Rate* and the shortest *Latency*, under both NUM embedding and UM embedding scenarios. We also notice that, comparing to the UM embedding, although the NUM embedding significantly reduces the embedding dimension, it increases *Sensitivity* by 1% and lowers *Latency* by around 1 second of the NUMME feature. Moreover, the proposed information theoretic approach can parallel the state-of-art methods in the literature. It is important to note that our framework can achieve the comparable performance without using any elaborate machine learning algorithms and

TABLE I

PERFORMANCE SUMMARY

Feature	Sensitivity	FAR	Latency
NUM embedding (dim = 13)			
LMAX	85%	22%	11.03
TND	88%	20%	11.86
LAM	85%	22%	13.10
TT	88%	21%	16.95
λ_2	87%	16%	9.91
NUMME	**96%**	**15%**	**6.68**
UM embedding (dim = 23)			
LMAX	85%	23%	10.78
TND	89%	20%	10.48
LAM	85%	22%	11.14
TT	87%	22%	11.98
λ_2	86%	16%	10.61
NUMME	**95%**	**14%**	**7.24**
Other Studies			
Shoeb et al. [11]	96%	5/hr	4.60
Ahammad et al. [12]	98%	14%	1.76
Thodoroff et al. [13]	85%	0.8/hr	n/a
Zabihi et al. [3]	88%	7%	n/a
Samiee et al. [14]	70%	0.4/hr	n/a
Bhattacharyya et al. [15]	97%	1%	n/a
Khanmohammadi et al. [18]	96%	0.1%	4.21
Bomela et al. [5]	79%	0.02/hr	4.44

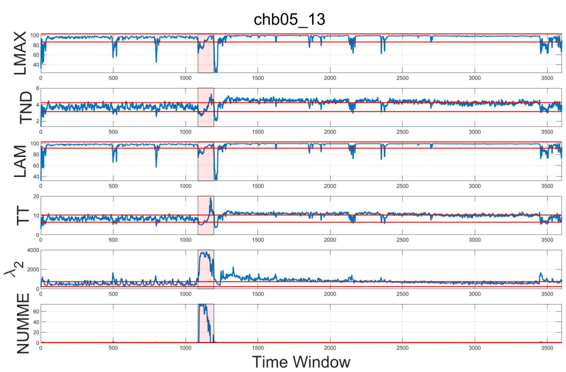

Fig. 1. The real-time monitor chart for Session 13 of Subject 5. The highlighted segments (in red shadows) mark the seizure onset periods. The horizontal red lines represent constant upper and lower thresholds.

without additional complicated features. According to Figure 1, the NUMME feature is more sensitive than RQA and λ_2 features for detecting seizure onsets. Figure 1 also implies that the NUMME measure yields more stable monitor than the rest for seizure detection. On the contrary, RQA measures exhibit relatively high variation. Furthermore, potential preictal events are observed in some cases. In summary, our findings support the claim that our proposed framework is promising for real-time seizure detection in clinical practice.

VI. CONCLUSIONS

In this study, we presented a novel information theoretic framework for real-time epileptic seizure detection. Firstly, an effective NUM embedding algorithm was proposed for reconstructing state space of the complex dynamic brain system. Then, we originated the NUMME to quantify structural complexity of a multivariate time series (e.g., EEG signals). We found the complexity measure (i.e., NUMME) of embedded EEG data can successfully characterize epileptic patterns. Hence, this work connects the complexity property of brain system with the epileptic seizures, which in turn inspires future research on relationship between brain system complexity and neurological disorders (e.g., epilepsy).

Our framework achieves a high sensitivity of 96% on the benchmark CHB-MIT Scalp EEG Database [16], which parallels the state-of-art approaches in the literature. While, we have to emphasize that our framework achieves comparable performance without using any sophisticated machine learning algorithms and without additional complicated features. Furthermore, interpretability and low computational cost are two significant advantages of our framework which makes it suitable for real-time applications.

REFERENCES

[1] "The revised definition of epilepsy — epilepsy foundation," https://www.epilepsy.com/article/2014/4/revised-definition-epilepsy.

[2] "Epilepsy," https://www.who.int/health-topics/epilepsy#tab=tab_1.

[3] M. Zabihi, S. Kiranyaz, A. B. Rad, A. K. Katsaggelos, M. Gabbouj, and T. Ince, "Analysis of high-dimensional phase space via poincaré section for patient-specific seizure detection," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 386–398, 2015.

[4] M. Fan and C.-A. Chou, "Detecting abnormal pattern of epileptic seizures via temporal synchronization of eeg signals," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 601–608, 2018.

[5] W. Bomela, S. Wang, C.-A. Chou, and J.-S. Li, "Real-time inference and detection of disruptive eeg networks for epileptic seizures," *Scientific Reports*, vol. 10, no. 1, pp. 1–10, 2020.

[6] M. B. Kennel, R. Brown, and H. D. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical review A*, vol. 45, no. 6, p. 3403, 1992.

[7] I. Vlachos and D. Kugiumtzis, "State space reconstruction from multiple time series," in *Topics on Chaotic Systems: Selected Papers from Chaos 2008 International Conference*. World Scientific, 2009, pp. 378–387.

[8] K. H. Kraemer, G. Datseris, J. Kurths, I. Z. Kiss, J. L. Ocampo-Espindola, and N. Marwan, "A unified and automated approach to attractor reconstruction," *New Journal of Physics*, 2021.

[9] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical review letters*, vol. 89, no. 6, p. 068102, 2002.

[10] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy analysis," *IEEE signal processing letters*, vol. 19, no. 2, pp. 91–94, 2011.

[11] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 975–982.

[12] N. Ahammad, T. Fathima, and P. Joseph, "Detection of epileptic seizure event and onset using eeg," *BioMed research international*, vol. 2014, 2014.

[13] P. Thodoroff, J. Pineau, and A. Lim, "Learning robust features using deep learning for automatic seizure detection," in *Machine learning for healthcare conference*. PMLR, 2016, pp. 178–190.

[14] K. Samiee, P. Kovács, and M. Gabbouj, "Epileptic seizure detection in long-term eeg records using sparse rational decomposition and local gabor binary patterns feature extraction," *Knowledge-Based Systems*, vol. 118, pp. 228–240, 2017.

[15] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient-specific eeg seizure detection using empirical wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2003–2015, 2017.

[16] "Chb-mit scalp eeg database," https://physionet.org/content/chbmit/1.0.0/.

[17] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, 2000.

[18] S. Khanmohammadi and C.-A. Chou, "Adaptive seizure onset detection framework using a hybrid pca–csp approach," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp. 154–160, 2017.