An Adaptive Pattern Learning Framework to Personalize Online Seizure Prediction

Cao Xiao, Shouyi Wang, *Member, IEEE*, *Member, IEEE*, Wanpracha Art Chaovalitwongse, *Senior Member, IEEE*, Leon Iasemidis, *Member, IEEE*, Stephen Wong

***Abstract*—The sudden and spontaneous occurrence of epileptic seizures can impose a significant burden on patients with epilepsy. If seizure onset can be prospectively predicted, it could greatly improve the life of patients with epilepsy and also open new therapeutic avenues for epilepsy treatment. However, discovering effective predictive patterns from massive brainwave signals is still a challenging problem. The prediction of epileptic seizures is still in its early stage. Most existing studies actually investigated the predictability of seizures offline instead of a truly prospective online prediction, and also the high inter-individual variability was not fully considered in prediction. In this study, we propose a novel adaptive pattern learning framework with a new online feature extraction approach to achieve personalized online prospective seizure prediction. In particular, a two-level online feature extraction approach is applied to monitor intracranial electroencephalogram (EEG) signals and construct a pattern library incrementally. Three prediction rules were developed and evaluated based on the continuously-updated patient-specific pattern library for each patient, including the adaptive probabilistic prediction (APP), adaptive linear-discriminant-analysis-based prediction (ALP), and adaptive Naive Bayes-based prediction (ANBP). The proposed online pattern learning and prediction system achieved impressive prediction results for 10 patients with epilepsy using long-term EEG recordings. The best testing prediction accuracy averaged over the 10 patients were 79%, 78%, and 82% for the APP, ALP, and ANBP prediction scheme, respectively. The experimental results confirmed that the proposed adaptive prediction framework offers a promising practical tool to solve the challenging seizure prediction problem.**

***Index Terms*—time series pattern recognition, online prediction, seizure prediction, adaptive learning, probabilistic decision making**

I. Introduction

Epilepsy is one of the most common neurological disorders, affecting approximately 1% of the world's population [4]. Epileptic seizures generally occur without warning, and the shift between a normal brain state and seizures is often considered an unpredictable phenomenon. The unpredictability of seizures represents a significant source of morbidity in patients with epilepsy. These patients frequently suffer from seizure-related injuries due to a loss of motor control, a loss of consciousness or a delayed reactivity during seizures [17]. The ability to predict the occurrence of impending seizures could significantly improve the life quality and treatment of patients with epilepsy.

A recent comprehensive review of seizure predictability and seizure prediction can be found in [12]. The pioneering efforts to investigate the predictability of seizures were made by Viglione and Walsh in the 1970s [28] and Iasemidis et al. in the 1980s [8]. In particular, Iasemidis et al. [9] noted pre-seizure changes based on analysis of entrainment of brain dynamics. Since then, several studies on seizure predictability have been undertaken utilizing the abundance of computer power and advanced mathematical techniques for biological signals. Lehnertz and Elger [13] showed that the correlation dimension decreases prior to seizures. Le van Quyen et al. [21] reported a reduction in the dynamical similarity index before seizure occurrence. Mormann et al. [15] observed that there was a relative decrease of signal power in the delta band of the EEG up to hours prior to seizure onsets. They also demonstrated statistically significant discrimination between pre-seizure and normal brain states.

Although efforts have been made to investigate the predictability of seizures, only a few investigations have been performed on seizure prediction (prospective analysis). Seizure prediction, what is needed in medical applications, is different from and more difficult to achieve than seizure predictability analysis, as we have argued and shown in the past. One common measure for seizure prediction is sensitivity, which is defined as the number of correct predicted seizures divided by the total number of seizures. A seizure is considered to be correctly predicted if there is at least one warning within its preceding horizon. Another measure is specificity, i.e. false prediction rate, which is defined by the number of false predictions per hour (or unit time). The two measures are widely adopted in more recent studies of seizure prediction.

A more recent investigation in seizure prediction was performed by Schulze-Bonhage's group, a group that in the past was very skeptical with respect to feasibility of seizure prediction schemes [5]. These investigators reported a maximum mean sensitivity of 43.2%, at a false seizure prediction rate of 0.15 per hour. Training on each patient's

Shouyi Wang is with the Department of Industrial & Systems Engineering, and the Integrated Brain Imaging Center, School of Medicine, University of Washington, Seattle, WA 98195, USA. E-mail: shouyi@uw.edu;

Wanpracha Chaovalitwongse is with the Departments of Industrial & Systems Engineering and Radiology, University of Washington, Seattle, WA 98195, USA. E-mail: artchao@uw.edu;

Cao Xiao is with the Department of Industrial & Systems Engineering, University of Washington, Seattle, WA 98195, USA. E-mail: Cao Xiao;

Leon Iasemidis is with the Department of Biomedical Engineering, Louisiana Tech University, Ruston, LA 71272, USA. E-mail: leonidas@latech.edu;

Stephen Wong is with Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick, NJ 08901, USA. E-mail: wongst@umdnj.edu

EEG data was required before a prospective run of the algorithm on a patient. The interval at which a seizure was predicted ahead of time was not clearly reported but appeared to be in the range of 10 to 60 minutes. Although this performance is sensitivity-wise almost half of what we have reported in the past, even with our first-generation seizure prediction algorithms [10], it is an additional piece of independent evidence that seizures not only can be predictable (retrospective analysis) but also predicted (prospective analysis) too. Most recently, two other groups have attempted prospective seizure prediction using pattern recognition methodologies [2], [16]. A training stage per patient was required here too. The reported specificity values varied widely per patient while sensitivity to seizures was very high and appeared encouraging. However, in a closer inspection, both groups used the old Freiburg database [25] that is, a database with fragmented, discontinuous EEG data (one 24 hour interictal period well separated from about 50 min preictal periods for a couple of seizures per patient). Even though this is an intermediate step towards seizure prediction, it may give a false picture of the capabilities of a seizure prediction algorithm, which has to work not only in real time but on line on continuous EEG data. This fact is shown by Freiburg group's recent publication [5] and the poor results they obtained about sensitivity using a fair specificity value when they decided to run their algorithms on long-term continuous EEG data.

In general, most of the current seizure prediction methods involve two steps. First, univariate or multivariate EEG features are extracted from an EEG epoch within a moving window over the entire EEG recording. Then each EEG epoch is classified as either pre-seizure or normal based on an optimized threshold level. Whenever a monitored EEG epoch is classified as pre-seizure, a warning alarm is triggered to indicate that an impending seizure is eminent within a pre-defined prediction horizon. Although some methods have shown good results for selected patients, the reliability and repeatability of the results have been questioned when tested on other EEG datasets. For example, many of the earlier optimistic findings were irreproducible or achieved poor performance in extended EEG datasets [1]. Several aspects of EEG-based seizure prediction are challenging:

- Massive Multichannel EEG data. Many sensors are necessary to monitor the brain activity for long periods. Moreover, continuous EEG data are needed, with sampling rates up to the KHz range. Understanding and decoding multichannel EEG time series, temporally and spatially, is an open research area.
- Variability in pre-seizure patterns. Unlike many biological detection problems with relatively clear patterns to recognize, the pre-seizure EEG patterns are unknown. Given the heterogeneity of epileptogenic regions of the brain, and intracranial electrode placement that is individualized per patient, pre-seizure EEG patterns may vary dramatically across patients.
- High intra-individual variability of epileptic seizures with time [10]. Pre-seizure EEG patterns may vary a lot over time even in the same patient.
- Uncertainty in definition of a prediction horizon. For evaluation of any seizure prediction algorithm, a prediction horizon has to be defined. It provides the time period within which a next seizure may occur upon issue of a warning. Ideally, prediction horizon should be equal to the pre-seizure period. However, the length of pre-seizure period is a priori unknown and is likely variable from one patient to another.

A. Related Work In Adaptive Seizure Prediction

Most current seizure prediction methods are nonadaptive threshold-based approaches. The great inter- and intra-individual variability of epileptic seizures makes it difficult to develop a universal nonadaptive predictor. Manually tuning a threshold level for each individual patient is a subjective procedure and would pose a significant burden on physicians and patients. The inability to benefit a wide spectrum of patients represents a great limitation to current seizure prediction methods. Therefore, there is an urgent need for an automated adaptive framework for epileptic seizure prediction.

Iasemidis et al. [10], [11] and Sackellares et al. [24] developed optimization-based prediction algorithms which, based on dynamical synchronization in the human epileptic brain over time, adaptively selected groups of critical EEG electrode sites to predict impending seizures. Adaptively selecting only EEG channels limited the prediction performance. More recently, Iasemidis' group published similar results, with high sensitivity (85.9%) and specificity (0.18 false positive rate (FPR) per hour), and long warning times prior to seizures (67.6 minutes on average), on prospective seizure prediction in rodents with chronic epilepsy [7]. Rajdev et al. [22] proposed an adaptive seizure prediction algorithm based on linear autoregressive (AR) modeling of intracranial EEG recorded from 4 kainate-treated epileptic rats. Warnings were issued when prediction errors of the model over time exceeded an adaptive threshold. Even though sensitivity to seizures was high, specificity was very low (the online real-time algorithm exhibited 0.08 false warnings per minute, or equivalently 4.8 warnings per hour). In addition, seizures were predicted on the average of only a few seconds prior to seizure onset, which is within the range of seizure 'detection' rather than seizure 'prediction' (e.g. seizures may start in deep brain structures seconds before they manifest themselves on the surface of the brain).

Thus, current adaptive seizure prediction approaches are generally based on an adaptively-optimized set of EEG channels [10], [11], [24] or an adaptive threshold [22]. In principle, these approaches employed the prediction settings optimized from one or several occurred seizures in the past to predict the next seizure. An intrinsic assumption is that the next seizure is similar to the most recent ones. However, this is not always true in reality. Therefore, it is desirable to develop a prediction system that is capable of

accumulating knowledge of predictive EEG patterns over time instead of holding only 'short memories' from the past.

B. Personalized Seizure Prediction

Given our experience with seizure prediction, we conjecture that a promising approach should possess intelligent learning ability and autonomously adapt to individual patient's EEG patterns. In this study, we make an important progress towards this direction and formulate the seizure prediction problem as an online adaptive pattern recognition and learning problem. In particular, we provide a new online feature extraction and prediction methodology for nonstationary multivariate time series data, and apply it to construct an adaptive framework for online monitoring and prediction of seizure onset. The proposed online learning and prediction framework combines probabilistic theory, adaptive learning theory and new feature extraction techniques to solve the challenging online seizure prediction problem. It can efficiently process massive non-stationary EEG data, and summarize millions of complex time series patterns in a concise feature space at a very low computational cost. We show that, with adaptive pattern learning capabilities, the proposed online prediction framework has a great potential to realize accurate personalized seizure prediction for each individual patient.

The rest of the paper is organized as follows. In section II, the online monitoring and prediction approaches are presented, including feature extraction, feature selection, three adaptive prediction schemes, and the evaluation metrics of prediction performance. The experimental results are presented in Section III, and the conclusions of the study is in Section IV.

II. Materials and Methods

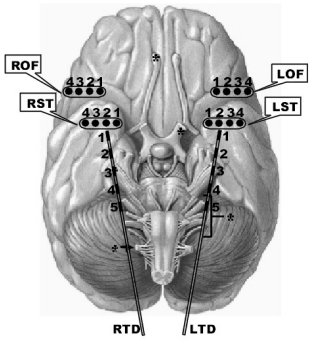
A. Data Collection

We used a dataset containing long-term continuous intracranial EEG recordings from ten epileptic patients. The EEG was recorded with a standard 26 electrode montage (see electrode placement in Figure 1). Recording durations ranged from 3 to 13 days per patient. Expert epileptologists annotated the EEG recordings of the patients to determine the number of seizures, their onset and offset points. The characteristics of the ten patients and the EEG data statistics are outlined in Table I.

B. Online Learning and Prediction Framework

In this study, we employ an adaptive online learning and prediction framework to discover hidden predictive patterns for epileptic seizures. The flowchart of the online prediction scheme is shown in Figure 2. The proposed online prediction framework has the following significant components:

- employed a deterministic chaotic measure Lyapunov exponent to achieve dimensionally reduction.



Subdural electrode strips are placed over:
left orbitofrontal cortex (LOF)
right orbitofrontal cortex (ROF)
left subtemporal cortex (LST)
right subtemporal cortex (RST)

Depth electrodes are placed in:
left hippocampus (LTD)
right hippocampus (RTD)

Fig. 1: The interior transverse view of the brain [18] and the placement of the 26 EEG electrodes.

- developed a two-level time series monitoring and feature extraction framework to characterize multivariate EEG patterns.
- constructed an efficient online pattern library by introducing pattern clusters of original feature vectors in discrete feature space.
- proposed an adaptive probabilistic online prediction rule using the pattern occurrence time and frequency information in pre-seizure and normal periods stored in the pattern library. Two other prediction rules were also experimented and compared using a popular binary classification technique and Naive Bayesian theory.

These key components of the online monitoring and prediction framework are discussed in detail in the following subsections.

C. Two-Level Feature Extraction

As shown in Figure 2, we employ a two-level sliding window approach to monitor and extract features from multichannel EEG signals. *The First-Level Feature Extraction:* is performed on raw EEG signals directly. Since EEG signals are highly nonstationary and seemingly chaotic, there has been an increasing interest in analyzing EEG signals in the context of chaos theory [23]. Several commonly used chaotic measures include largest Lyapunov exponent [10], correlation dimension [26], Hurst exponent [3] and entropy [20]. Among these measures, the Lyapunov exponent has been shown to be useful in characterizing the stability of the brain [27]. We have developed an estimation algorithm called the short-term largest Lyapunov exponent (STL_{max}) to quantify EEG dynamics [8], [10]. We also employed this measure in the current study. For each channel of EEG time series, we convert each 10-second non-overlapping EEG signal into one STL_{max} number. For a raw EEG epoch with 26 channels, 29 first-level features are extracted in the first-level sliding window. The averaged STL_{max} values of each channel contributed 26 univariate features. And we also extracted three bivariate features: averaged pairwise Euclidean distance, T-statistic, and Pearson correlation over all pairs of channels. The second-level feature extraction is

TABLE I: The characteristics of the ten patients and the EEG data statistics

Patient	Gender	Age (years)	EEG Duration (days)	Number of Seizures	Seizure Rate (per hour)	Seizure Type	Onset Region
1	F	45	3.55	7	0.082	CP, SC	RH
2	M	60	11.98	7	0.024	CP, GTC, SC	RH, RF
3	F	41	8.55	22	0.104	CP	RH
4	M	19	13.13	17	0.054	CP, SC	RH
5	M	33	12.24	18	0.061	CP, SC	RH
6	M	38	3.18	9	0.118	CP, SC	RH
7	M	44	6.09	23	0.157	CP, SC	LH, RH
8	M	29	6.07	19	0.130	CP, SC	RH
9	F	37	11.53	20	0.061	CP, SC	LH, RH
10	M	37	9.88	12	0.051	CP, GTC	LH, RH
total	-	-	86.20	154	-	-	-

Onset region: LH, left hippacampal; RH, right hippacampal. Seizure types: CP, complex partial; SC, subclinical; GTC, generalized tonic/clonic.

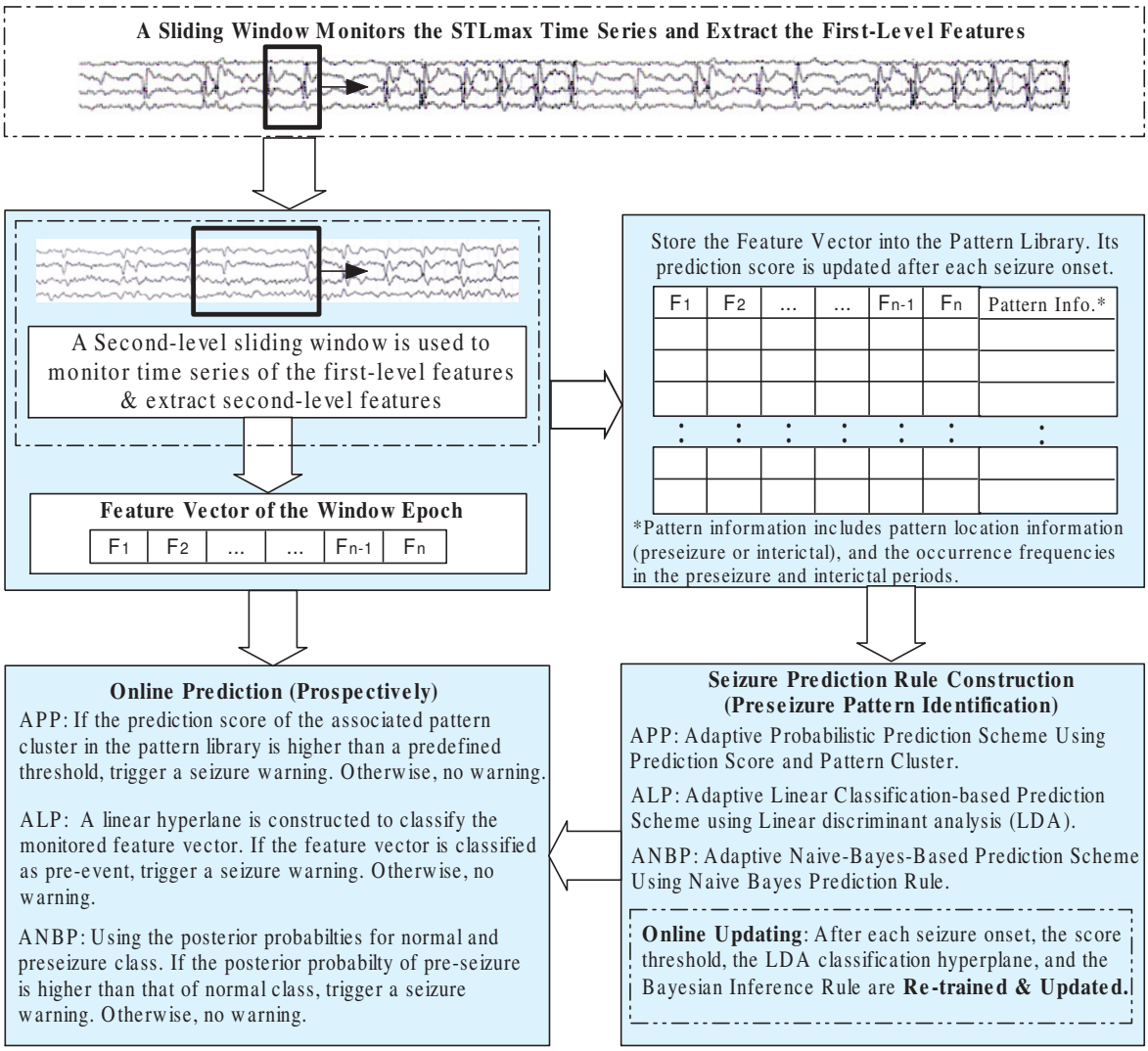


Fig. 2: Flowchart of the online learning and prediction framework for personalized seizure prediction.

to monitor and characterize the temporal evolutions of the first-level features.

The Second-Level Feature Extraction: given a time series of a first-level feature, we first applied a piecewise linear approximation algorithm to partition the time series into piecewise linear segments using its key-turning points. We have developed a reliable and efficient algorithm for piecewise linear segmentation of time series data, called two-stage-top-down (TSTD) approach. A more detailed discussion of this algorithm can be found in [29]. After piecewise linear segmentation of a time series $X = (x_1, x_2, ..., x_n)$, its key-turning points become prominent (shown with black dots in Figure 3). There are six linear segments to describe the original time series. Three segments (a, c, e) have increasing trends, and the other three segments (b, d, f) have decreasing trends. Then, the extracted increasing

and decreasing trends characterize the temporal fluctuation pattern in a time series. The following four second-level features are proposed to capture the temporal fluctuation of first-level feature time series.

- Feature 1: accumulated vertical decrease in the segmented piecewise linear time series, which is calculated as
$$F_1 = H(a) + H(c) + H(e), \quad (1)$$
where the function $H(.)$ means the vertical distance from the starting point to the ending point of a sub-segment in the segmented time series.
- Feature 2: accumulated vertical increase in the segmented piecewise linear time series, which is calculated as
$$F_2 = H(b) + H(d) + H(f), \quad (2)$$
- Feature 3: percentage of the decreasing sub-segments in the time series, which is calculated as
$$F_3 = T(a + c + e)/T(X), \quad (3)$$
where $T(.)$ is the horizontal distance from the starting point to the ending point of a sub-segment.
- Feature 4: range of the time series, which is calculated as
$$F_4 = max(X) - min(X), \quad (4)$$
where $max(X)$ and $min(X)$ means the maximum and minimum values of the time series, respectively.

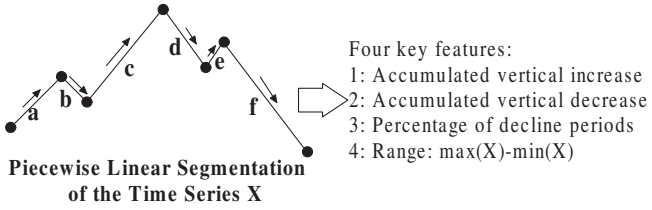


Fig. 3: Four skeleton-point-based features are employed to represent the temporal fluctuation of a time series.

D. Feature Selection

For each EEG epoch monitored online, we have 29 first-level features, and each first-level feature has 4 second-level features to describe temporal variations of the first-level features. Thus, each multivariate EEG epoch has $29 \times 4 = 116$ features after the second-level feature extraction. Since not all of these features are informative to seizures and also to achieve dimensionality reduction, we employed the Pudil's floating search [19] to select which temporal features of which first-level features have the discrimination power to separate pre-seizure from normal epochs. Pudil's floating search provides the possibility of trading significant reduction of search time for often negligible decrease of the classification accuracy. The criterion for feature selection was the nearest-neighbor leave-one-out classification performance. The selected optimal feature subset has the highest leave-one-out classification accuracy. In our experiment, we selected eight most important features from the 106 candidates from training dataset for online pattern monitoring and prediction.

E. Pattern Library Construction in Discrete Space

The two-level sliding window monitors EEG signals and convert EEG epoch at each step of sliding window into a 8-dimensional feature vector. A pattern library was constructed by storing each 8-dimensional feature vector at each step of sliding window. Instead of using raw feature vector, we represent each pattern vector in a pattern cluster formulation in discrete feature space. *Pattern Cluster In Discrete Feature Space:* we partitioned each feature space into a number of non-overlap intervals. Feature vectors that fall into the same feature bins are considered as similar feature patterns and form a pattern cluster. Using the concept of pattern cluster, one can represent millions or billions of feature vectors by a largely reduced number of pattern clusters. As shown in Figure 4, we partitioned each feature space into nonoverlapping intervals, two similar pattern vectors can be associated to and represented by the same pattern cluster in the discrete feature space. The pattern cluster representation achieves dimensionality reduction and allows a very efficient storage, visualization, and computational analysis. More importantly, it enable us to perform probabilistic analysis to pattern clusters.

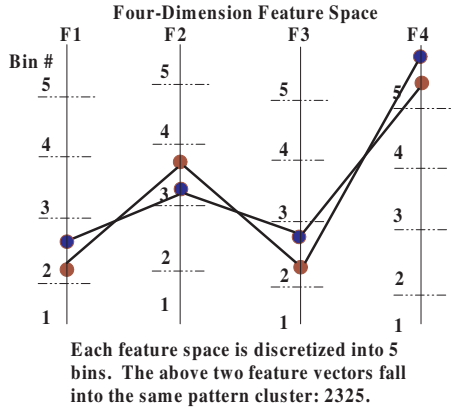


Fig. 4: Demonstration of the pattern cluster concept. The two similar feature vectors can be associated to and represented by the same pattern cluster.

F. Adaptive Probabilistic Prediction Scheme

In the probabilistic prediction scheme, each feature vector is represented by its corresponding pattern cluster in discrete feature space and stored into the pattern library. In addition, as shown in Figure 2, the pattern occurrence frequency and occurrence time related to seizure onset is also stored in the pattern library. Using these information, we propose an adaptive probabilistic prediction framework to discover hidden predictive pattern clusters to seizure onset.

Prediction Score of A Pattern Cluster: given a pattern cluster, indexed as the kth cluster in the pattern library, its prediction score S_k is defined as follows:

$$S_k = \frac{N_{pre}/N_{tot}}{R_{pre}} \times \frac{N_{pre}^{dist}}{N_{evt}}, \quad (5)$$

where N_{pre} is the number of occurrences of the pattern cluster in all monitored pre-seizure periods; and N_{pre}^{dist} is the number of pre-seizure periods in which the pattern cluster appears at least once; N_{tot} is the total number of occurrences of the pattern cluster; and N_{evt} is the total number of seizures that have occurred. For example, if two seizures occurred, and a pattern cluster occurred three times in the first pre-seizure period, and twice in the normal periods, but did not show up in the second pre-seizure period, then $N_{pre} = 3$, $N_{pre}^{dist} = 1$, $N_{tot} = 5$, and $N_{evt} = 2$. Finally, R_{pre} is the ratio of pre-seizure periods over normal periods defined as follows:

$$R_{pre} = \frac{T_{pre}}{T_{tot} - T_{pre}} = \frac{N_{evt} \times T_{hrzn}}{T_{tot} - N_{evt} \times T_{hrzn}}, \quad (6)$$

where T_{pre} is the total length of monitored pre-seizure periods, T_{tot} is the total length of monitored EEG time series, and T_{hrzn} is the length of prediction horizon.

The predictive score proposed in formula 5 indicates how strong a pattern cluster is associated with seizure onset. In particular, the first term of formula 5 is to evaluate if the pattern cluster occurs in pre-seizure periods at a random level. If the pattern is pure random in both pre-seizure and normal periods, then we have $E(N_{pre}/N_{tot}) = E(R_{pre})$. If the pattern occurs more frequently in pre-seizure periods than the normal periods, we have $E(N_{pre}/N_{tot}) > E(R_{pre})$. The higher the ratio value, the more likely the pattern cluster is associated with seizure onset. If the pattern occurs less frequently in pre-seizure periods than the normal periods, we have $E(N_{pre}/N_{tot}) < E(R_{pre})$. The second term of formula 5 is to evaluate if a pattern cluster occurs in many pre-seizure periods. We expect that an ideal candidate of predictive pattern should appear in most pre-seizure periods, not only in a few ones. That is $N_{pre}^{dist}/N_{evt} \approx 1$.

Formula 5 estimates the likelihood of a pattern cluster to appear in pre-seizure periods. The higher the prediction score, the higher probability the pattern cluster appears in pre-seizure periods, and thus the more probable it is to predict seizures.

Prediction Score-Based Prediction Rule: the pattern library continuously collects pattern clusters online and updates their prediction scores according to formula 5. An adaptive threshold strategy is proposed to discriminate pre-seizure and normal pattern clusters online. We introduced a score threshold S^* which is defined as the value that **retrospectively** maximizes the prediction accuracy in the monitored historical time series. The threshold S^* is updated after each occurrence of a seizure. Thus we call this prediction scheme as adaptive probabilistic prediction (APP) defined by:

$$prediction = \begin{cases} preseizure, & \text{if } S_k \geq S^*, \\ normal, & \text{otherwise.} \end{cases}$$

G. Prediction Rule Construction in Continuous Space

In addition to the probabilistic prediction rule in discrete feature space, we also constructed and tested two prediction rules based on the pattern library using original feature vectors in continuous feature space. In this case, if a pattern appears in pre-seizure periods, it is labeled as a pre-seizure pattern; if it is in normal periods, it is labeled as a normal pattern. To discriminate pre-seizure patterns and normal patterns, we employed two popular classification techniques including the Fisher's Linear Discriminant Analysis (LDA) and the Bayesian decision-making theory. The two prediction schemes will be presented in the following.

1) Binary Classification Based Prediction Scheme: For a pattern library with labeled feature vectors, we employed a popular binary classification technique LDA to construct a hyperplane to discriminate pre-seizure and normal patterns. Fisher's LDA aims to find an optimal projection by minimizing the intraclass variance and maximizing the distance between the two classes simultaneously [6]. Mathematically, the optimal projection $\omega^* \in R^{n\times 1}$ can be obtained by solving the following optimization problem:

$$\omega^* = argmax_\omega \frac{\omega^T S_b \omega}{\omega^T S_\omega \omega}, \quad (7)$$

where ω is the direction of the hyperplane that is used to eparate the two data sets. S_b and S_ω are the interclass and intraclass covariance matrices, respectively. Once ω^* is obtained, the optimal decision boundary of LDA can be represented by ${\omega^*}^T Y + b = 0$. The bias term b is defined by $b = -{\omega^*}^T(m_1 + m_2)/2$.

LDA-based Online Prediction Rule: an optimal LDA hyperplane was trained by the pattern library with feature vectors of preseizure and normal. If monitored feature vector is X_k, then the LDA-based prediction rule is defined by:

$$prediction = \begin{cases} preseizure, & \text{if } {\omega^*}^T X_k + b > 0, \\ normal, & \text{if } {\omega^*}^T X_k + b \leq 0. \end{cases}$$

The LDA hyperplane is re-trained after each seizure onset using the latest updated pattern library. We call this prediction scheme an adaptive LDA-based prediction (ALP).

2) Naive Bayesian Based Prediction Scheme: Bayesian decision theory is one of the most widely used statistical approaches in many data mining problems. Naive Bayesian rule provides us with another tool to explore the pattern library. For any monitored feature vector X_k, we have two classes normal and pre-seizure, denoted as C_1 and C_2 respectively. By Bayes theorem, the posterior probabilities for the two classes are calculated by

$$P(C_i \| X_k) = \frac{P(X_k \| C_i) P(C_i)}{P(X_k)}, i = 1, 2 \quad (8)$$

where $P(C_i)$ is called prior probability of class C_i, which can be estimated by the portion of feature vectors that are labeled as C_i in the pattern library, $P(X_k \| C_i)$ is the likelihood probability that the feature vector X_k belongs to class C_i, $P(X_k)$ is the probability of witnessing X_k overall regardless of its class, and $P(C_i \| X_k)$ is the posterior probability that X_k belongs to C_i when X_k is observed. A Bayesian prediction rule can be constructed determining the pattern class by maximizing the posterior probability. As it

is computationally expensive to compute the conditional probability for a multivariate feature vector X_k, Naive Bayes rule assumes that each features are independent. With the independence assumption, the conditional distribution over the class variable C_i is then expressed as follows:

$$P(C_i \| X_k = (F_1, \ldots, F_n,)) = \frac{1}{Z} P(C_i) \prod_{k=1}^{n} P(F_j \| C_i) \tag{9}$$

where $i = 1, 2$ and $j = 1, \ldots, n$, representing two classes and n feature dimensions; $Z = P(X_k)$ is a constant for all classes, and it is a scaling factor only dependent on feature vector X_k. One can easily estimate the independent probability distributions $P(F_j \| C_i)$, $j = 1, \ldots, n$ from the training data set (pattern library). In this study, each feature was assumed to follow Gaussian distribution.

Naive-Bayes-Based Prediction Rule: for each monitored feature vector X_k, $P(C_i)P(X_k \| C_i)$ is evaluated for both classes: pre-seizure C_1 and normal C_2. The Naive-Bayes-based prediction rule is then defined by:

$$prediction = \begin{cases} preseizure, & \text{if } P(C_1 \| X_k) > P(C_2 \| X_k), \\ normal, & \text{if } P(C_1 \| X_k) \leq P(C_2 \| X_k). \end{cases}$$

The Bayesian decision rule is re-trained after each seizure onset using the latest updated pattern library. We name this prediction scheme by adaptive Naive Bayesian-based prediction (ANBP).

H. Seizure Detection

The proposed adaptive prediction approaches have to work with a seizure detection algorithm for real-time detection of seizure onset and subsequent characterization of a period in the EEG as pre-seizure or normal. There have been a number of automated online seizure detection algorithms embedded in clinical EEG systems and our proposed prediction approaches can be readily integrated with them. Seizure detection was beyond the scope of this paper; all herein reported results were produced assuming an ideal seizure detection algorithm that runs in parallel with our seizure prediction algorithms and gives its input to the seizure prediction algorithms reliably and without delay upon a seizure occurrence.

I. Evaluation of Prediction Performance

To evaluate our prediction model and enable comparison with other models, we employed the measure specificity and sensitivity. We defined sensitivity as the number of correctly predicted seizures divided by the total number of seizures, denoted as sen_{blk} in this study. For prediction specificity, most current studies calculated a false prediction rate, which is defined by the number of false predictions per hour (or unit time). However, false prediction rate does not provide enough information to infer the effect of prediction horizon on the prediction performance. For example, a patient has to wait until the end of prediction horizon to determine if a warning is false. Given the same false prediction rate, an algorithm with a 5-hour prediction horizon will give a patient a much longer false awaiting time than the one with a 10-minute prediction horizon. Also the sensitivity of using a 5-hour prediction horizon is potentially much higher than that of using a 10-minute prediction horizon. This creates great confusion when evaluating the true prediction performance of a prediction system in real life. To overcome this issue, Mormann et al. [14] suggested to measure prediction specificity by the portion of time during the normal period that is not in false awaiting time. We herein employed this time-based specificity measure, denoted as spe_{blk}. The sen_{blk} and spe_{blk} estimation is schematically shown in Figure 5. To evaluate the overall performance by a single metric, we also define the overall prediction accuracy (PA) as the mean of sen_{blk} and spe_{blk}, that is $PA = (sen_{blk} + spe_{blk})/2$.

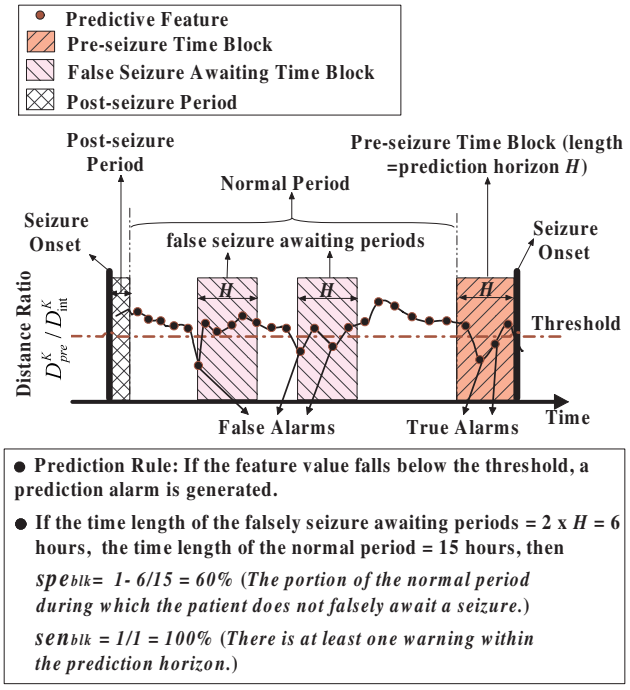


Fig. 5: A demonstration of the time block based sensitivity (sen_{blk}), specificity (spe_{blk}), false alarms, and false seizure awaiting periods in a generic seizure prediction procedure. The black dot line indicates critical predictive features. If it drops below the threshold, a prediction alarm is generated.

J. Training and Testing

For each patient, the EEG recordings were divided into training and testing datasets. The training datasets were the EEG recordings that contained the first half of seizures. It is used to select the most discriminative features of pre-seizure and normal patterns and obtain the best parameter settings of the prediction system, including prediction horizon (H), sliding window size (L_{mw}), and moving-step length (L_{step}). For each patient, the prediction framework ***prospectively*** with best parameter settings trained from the training dataset was evaluated by the testing EEG recordings that contained the second half of seizures of the patient. The pattern library and the prediction rules were

adaptively updated during the monitoring process of the testing EEG recordings.

III. Results

A. Computational Settings

The proposed online monitoring and prediction framework with three prediction rules was tested on EEG recordings from 10 patients with epilepsy. The complete parameter settings of the prediction system are summarized in Table II. We evaluated three choices of prediction horizons, seven choices of window length, and seven choices of step length of the sliding window. In the feature selection step, we selected eight most important features using the Pudil's floating search approach from the training data set (the first half of EEG recordings of each patient). For the pattern cluster formulation, we discretized each feature space into six equal bins. For an 8-dimensional feature space, the maximum number of possible pattern clusters in the pattern library is $6^8 = 1679616$.

B. Random Prediction Models

There has been a debate on how well prospective algorithms can predict seizures based on EEG analysis. Before any clinical application, it is necessary to evaluate if the designed prediction model is indeed able to perform better than a chance model. Therefore, we compared the performance of the proposed adaptive prediction model with two random prediction schemes that issue warning irrespectively of recorded EEG signals and depend only on seizure occurrences. They are periodic prediction schemes and Poisson prediction schemes. The periodic prediction scheme gives warnings at a fixed time interval T. The Poisson prediction scheme issues warnings according to an exponential distributed random time interval with a fixed mean λ. We performed the periodic prediction scheme and the Poisson prediction scheme for each patient. The values of λ and T were determined according to the average length of inter-seizure intervals for each patient as shown in Table I. For example, for patient 1, the averaged inter-seizure interval is 12.17 hours, we set $\lambda = T = 12.17$ hours. Clearly, this is the best value setting of T and λ the one could a-priori assume for such schemes.

C. Prediction Performance

Table III summarizes the training and testing prediction performances of the three prediction schemes in terms of sen_{blk} and spe_{blk}. All three online prediction schemes generated very promising prediction results. In particular, for the APP scheme, the averaged testing prediction accuracies of prediction horizons of 30, 90, 150 minutes were 79%, 65%, and 69% respectively. The averaged testing prediction accuracies for the three prediction horizons for ALP scheme were 78%, 70%, and 67% respectively; and for ANBP scheme, 82%, 79%, and 72% respectively. Overall, the ANBP scheme achieved the best testing prediction accuracy of 82% using the prediction horizon of 30 minutes. Figure 6, Figure 7, and Figure 8 provide demonstrations of the three prediction schemes for one patient with their best training parameter settings. One can see clearly how the three prediction schemes work in real-time online prediction of epileptic seizures. f

We notice that all three prediction schemes achieved their best testing accuracies using the prediction horizon of 30 minutes. With this 30-minute horizon, the average testing sen_{blk} and spe_{blk} of the APP scheme were 73% and 84%; of the ALP scheme 80% and 75%; and of the ANBP scheme 71% and 93% respectively. In addition, the corresponding average prediction times were 12.95 minutes, 14.26 minutes, and 9.30 minutes for the APP, ALP, and ANBP schemes respectively. The short prediction horizon of 30 minutes and the short prediction times at the level of 10 minitues provide a good time resolution for impending seizures, and also avoids long awaiting times once a false prediction occurs. Compared to previous approaches using longer prediction horizons [10], [11], [24], this improves the time resolution of seizure prediction. Also the high sensitivity and specificity values strongly indicate that the designed **prospective** prediction schemes are effective in learning predictive patterns online through our adaptive-updating learning strategy. For completeness, the prediction performances of the two random prediction schemes we tested are summarized in Table IV. The averaged prediction accuracies of the Poisson and periodic prediction scheme were both around 50% for all settings of prediction horizon. The averaged testing prediction accuracies of APP, ALP, and ANBP are about 60% higher than those of the two naive predictors. Thus, the proposed approaches perform much better than chance level. The experimental results are significan, especially since most of the current approaches are still striving to work better than a chance level [5].

D. Effectiveness of Adaptive Updating

The three prediction rules were adaptively updated in online. To evaluate the effectiveness of the online updating strategy, we experimented with different length of updating periods. That is, we update the decision-making rules of APP, ALP and ANBP for a portion of the entire monitoring time, stop updating at some time point, and keep the decision-making rules unchanged in the remaining part of online monitoring and prediction process. Figure 9 (a), (b), and (c) show the accuracies of the three prediction schemes with respect to using different portions for adaptive-updating. Point 0 on the x-axis in our plot indicates that the prediction rule (score threshold, LDA hyperplane, or Bayesian inference rule) keeps the initial-trained form unchanged throughout the prediction process without any updating. Point 0 on x-axis of each plot indicates that the prediction rule (score threshold, LDA hyperplane, or Bayesian inference rule) keeps the initial form (training dataset) unchanged throughout the prediction process without any updating at seizures. Point 1 on x-axis indicates that the prediction rule was updated throughout the prediction process. The blue and red dashed lines plot

TABLE II: Computational settings of the online prediction framework for epileptic seizure prediction.

Parameter Setting	Setting Choices
Prediction Horizon	30, 90, 150 minutes
1st-level sliding window (monitor raw time series)	window size: 10 minutes moving step length: 1 minute
2nd-level sliding window (monitor feature time series)	window size: 15, 30, 60, 90, 120, 150, 180 minutes moving step length: 1, 3, 6, 9, 12, 15, 18 minutes
Online Prediction Scheme	1. Adaptive-Threshold-Based Prediction Scheme (APP) 2. Adaptive LDA-Based Prediction Scheme (ALP) 3. Adaptive Naive Bayesian-Based Prediction Scheme (ANBP)
Feature Selection Method	Pudil's floating search based on 1-Nearest Neighbour leave-one-out classification performance.
1st-level features	1-26: Lyapunov exponents of 26 channels of raw EEG 27: averaged pair-wise Euclidean distances 28: averaged pairwise T-statistics 29: averaged pairwise correlations.
2nd-level features (temporal pattern feature)	1. accumulated vertical increase 2. accumulated vertical decrease 3. percentage of decline periods 4. amplitude range

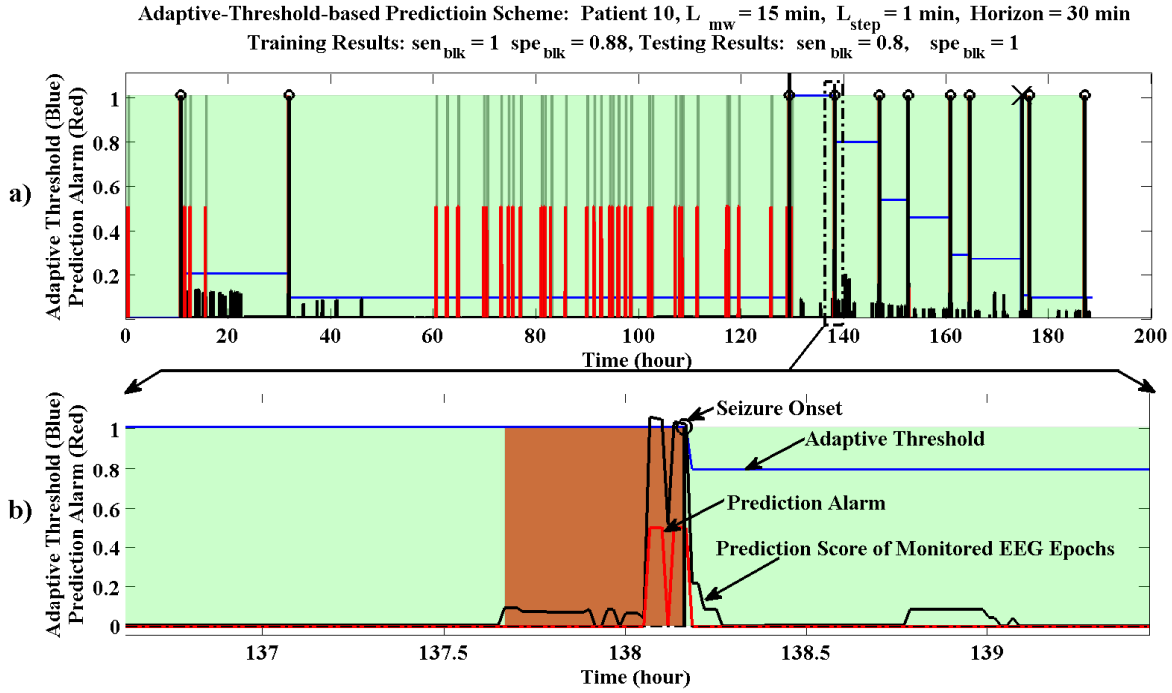


Fig. 6: a) Demonstration of the APP prediction scheme for patient 10. The vertical black lines that span the graph indicate seizure occurrences. On top of each seizure line, a mark 'O' indicates a seizure was correctly predicted and a mark 'X' indicates the seizure was miss-predicted. The piecewise horizontal blue line represents the values of the adaptive score-threshold, which is updated after each seizure occurrence. The black line is the prediction score of the monitored EEG epochs over time. If it is higher than the score-threshold (the blue line), a prediction alarm is generated. The red vertical lines represent prediction alarms. The vertical green lines indicates the false-awaiting periods due to false alarms. b) A magnified snapshot of prediction of one seizure. The brown-colored area indicates the pre-seizure period. The adaptive threshold, the prediction score, and the prediction warning are annotated in the graph.

the averaged sen_{blk} and spe_{blk}, respectively. The red solid line shows the averaged prediction accuracies (PAs) over the 10 patients.

We observe trends of increase in the overall prediction accuracy for all three prediction schemes with inclusion of updating, especially in the beginning of the monitoring process. These increasing trends indicate the strong learning capability of the proposed prediction schemes. Without any prior pattern knowledge, the proposed online learning and prediction framework is indeed effective in

TABLE III: The training and testing performance characteristics of the APP, ALP and ANBP prediction frameworks for three prediction horizons, respectively. The average prediction time is denoted by T_{pred} in the table.

		APP								ALP								ANBP							
		Setting		Training			Testing			Setting		Training			Testing			Setting		Training			Testing		
Horizion	Patient	L_{mw} (min)	L_{step} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	L_{mw} (min)	L_{step} (min)	sen_{blk}	spe_{blk}	T_{pred}	sen_{blk}	spe_{blk}	T_{pred} (min)	L_{mw} (min)	L_{step} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)
30 min.	1	30	1	0.67	0.99	3.50	0.67	0.72	15.50	30	9	1.00	0.81	20.33	0.50	0.69	26.00	15	9	0.67	0.99	5.00	0.33	0.85	29.00
	2	30	1	1.00	0.85	16.00	0.67	0.94	19.50	30	15	1.00	0.86	9.00	1.00	0.76	13.00	30	15	1.00	0.97	4.00	0.67	0.92	2.00
	3	30	6	0.90	0.96	10.33	0.67	0.92	8.67	30	9	1.00	0.90	10.50	0.89	0.81	12.63	15	3	0.90	0.95	6.89	0.89	0.93	5.50
	4	30	1	1.00	0.89	15.50	0.71	0.82	8.00	15	12	1.00	0.82	17.75	0.86	0.81	6.33	60	6	0.88	0.95	14.71	0.57	0.95	4.25
	5	60	3	0.71	0.90	20.20	0.67	0.78	26.50	60	1	1.00	0.84	18.57	0.83	0.66	23.60	30	15	0.43	0.98	13.67	0.67	0.97	9.00
	6	120	1	0.75	0.89	4.33	1.00	0.73	11.00	120	12	0.75	0.69	19.00	1.00	0.80	15.00	30	1	0.25	1.00	1.00	1.00	0.90	8.75
	7	15	6	0.75	0.75	19.33	0.86	0.69	18.33	15	6	1.00	0.73	15.63	0.86	0.62	21.33	15	1	0.38	0.97	11.33	0.57	0.87	17.50
	8	90	9	0.71	0.55	11.40	0.17	0.79	4.00	90	15	0.86	0.60	16.33	0.33	0.78	0.00	15	1	0.86	0.96	11.17	0.67	0.91	6.50
	9	30	3	0.67	0.99	5.00	1.00	0.90	13.78	15	12	1.00	0.90	8.56	1.00	0.84	14.67	30	6	0.89	0.98	4.00	0.67	0.96	7.33
	10	15	1	1.00	0.88	8.50	0.80	1.00	4.25	30	15	1.00	0.93	10.50	0.60	0.79	10.00	30	6	1.00	0.96	7.50	1.00	0.89	3.20
	Ave.			0.81	0.87	11.41	0.73	0.84	12.95	Ave.		0.96	0.83	14.62	0.80	0.75	14.26	Ave.		0.74	0.97	7.93	0.71	0.93	9.30
	PA			**0.84**			**0.79**			PA		**0.90**			**0.78**			PA		**0.85**			**0.82**		
90 min	1	15	1	0.67	0.90	4.50	0.67	0.59	66.50	90	18	1.00	0.50	59.33	1.00	0.56	80.00	30	6	0.67	0.91	45.50	0.33	0.76	41.00
	2	30	1	1.00	0.70	34.33	0.67	0.89	52.00	150	15	0.67	0.80	79.50	1.00	0.73	73.00	15	18	1.00	0.84	23.00	1.00	0.88	37.00
	3	60	1	0.70	0.89	48.57	0.78	0.69	19.00	30	15	1.00	0.74	39.00	0.89	0.36	46.75	15	6	1.00	0.94	17.80	0.89	0.92	15.88
	4	30	1	1.00	0.78	33.38	0.43	0.65	58.67	60	18	1.00	0.70	44.75	1.00	0.48	48.43	90	3	1.00	0.85	36.50	0.86	0.91	25.67
	5	150	18	0.83	0.77	75.60	0.50	0.78	62.67	90	18	0.86	0.71	49.67	0.83	0.58	65.40	15	3	0.71	0.87	41.60	0.83	0.84	28.80
	6	60	1	0.75	0.71	74.33	0.50	0.46	52.50	90	12	0.75	0.67	35.67	0.75	0.66	57.33	150	1	1.00	0.36	53.50	0.75	0.78	7.33
	7	90	9	0.63	0.72	34.20	0.57	0.60	52.75	15	12	1.00	0.50	58.63	0.86	0.36	59.00	15	1	0.75	0.77	33.50	0.71	0.60	30.80
	8	90	6	0.86	0.23	48.33	0.50	0.60	59.33	150	15	1.00	0.41	66.71	0.50	0.47	29.33	120	15	0.71	0.77	50.60	0.50	0.77	29.33
	9	60	12	0.89	0.78	37.38	0.56	0.74	16.20	120	6	1.00	0.69	55.11	0.89	0.50	55.13	30	9	0.89	0.94	15.63	0.78	0.88	17.57
	10	15	1	1.00	0.70	8.50	0.60	1.00	30.00	60	12	1.00	0.57	39.50	1.00	0.44	65.60	60	6	1.00	0.92	5.50	0.80	0.77	23.25
	Ave.			0.82	0.72	39.91	0.58	0.73	46.96	Ave.		0.96	0.63	52.79	0.86	0.53	58.00	Ave.		0.87	0.86	32.31	0.76	0.82	25.66
	PA			0.77			0.65			PA		0.79			0.70			PA		0.87			0.79		
150 min	1	15	1	1.00	0.26	90.00	0.67	0.69	147.50	60	18	1.00	0.28	131.33	1.00	0.39	128.00	15	9	1.00	0.69	84.33	0.67	0.54	93.00
	2	150	9	0.67	0.89	144.00	1.00	0.75	124.00	150	18	0.67	0.92	126.00	1.00	0.59	124.00	15	15	1.00	0.89	59.00	1.00	0.75	128.00
	3	120	1	0.90	0.74	52.89	1.00	1.00	75.33	120	12	1.00	0.73	92.20	0.89	0.31	123.00	15	9	1.00	0.86	56.80	1.00	0.96	69.89
	4	15	1	0.88	0.75	45.86	1.00	0.52	103.14	120	15	1.00	0.50	110.75	1.00	0.23	90.29	90	3	1.00	0.86	32.38	0.86	0.86	37.17
	5	150	12	0.67	0.69	78.50	0.67	0.57	121.00	30	18	1.00	0.48	76.71	1.00	0.41	120.83	15	3	1.00	0.70	58.14	0.83	0.79	72.60
	6	60	1	1.00	0.81	94.00	0.75	0.33	103.00	60	6	1.00	0.50	100.25	1.00	0.36	147.00	60	15	1.00	0.19	135.50	0.75	0.53	142.00
	7	150	15	0.88	0.42	89.43	1.00	0.49	116.00	15	18	1.00	0.51	107.13	1.00	0.30	79.14	15	15	0.88	0.64	98.57	0.57	0.50	88.00
	8	150	15	0.86	0.28	117.83	0.50	0.31	49.33	120	12	1.00	0.45	112.57	0.83	0.49	67.60	120	18	0.86	0.39	129.33	0.17	0.53	145.00
	9	90	12	0.67	0.62	93.50	0.56	0.46	119.20	90	18	1.00	0.62	84.89	0.89	0.41	101.63	15	15	1.00	0.63	78.22	0.89	0.66	57.38
	10	30	1	0.83	0.82	32.80	0.80	0.89	28.00	90	18	1.00	0.24	137.50	0.60	0.55	122.33	30	6	1.00	0.85	7.50	0.80	0.63	26.25
	Ave.			0.83	0.64	83.88	0.79	0.58	98.65	Ave.		0.99	0.47	107.93	0.91	0.42	110.38	Ave.		0.96	0.72	73.98	0.76	0.69	85.93
	PA			0.74			0.69			PA		0.73			0.67			PA		0.84			0.72		

TABLE IV: The prediction performances of two random prediction schemes (periodic and Poisson) are also reported. The prediction periods of the periodic and Poisson schemes for each patient are equal to the averaged length of inter-seizure intervals of the patient.

	30 Minutes						90 Minutes						150 Minutes					
	Poisson			Periodic			Poisson			Periodic			Poisson			Periodic		
Patient	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)	sen_{blk}	spe_{blk}	T_{pred} (min)
1	0.03	0.95	13.82	0.00	0.95	5.00	0.08	0.88	52.63	0.02	0.88	78.88	0.23	0.80	100.32	0.12	0.80	110.79
2	0.01	0.98	16.67	0.02	0.99	15.38	0.04	0.95	38.94	0.02	0.96	24.39	0.06	0.93	62.21	0.03	0.93	24.16
3	0.05	0.95	14.87	0.01	0.95	19.42	0.14	0.86	45.03	0.13	0.87	44.98	0.22	0.77	73.82	0.16	0.78	58.51
4	0.02	0.98	16.94	0.01	0.97	17.13	0.07	0.93	42.07	0.08	0.93	41.30	0.10	0.88	69.77	0.12	0.89	67.89
5	0.01	0.97	16.74	0.03	0.97	16.47	0.07	0.92	53.25	0.07	0.93	63.63	0.11	0.87	83.02	0.14	0.88	81.39
6	0.09	0.96	16.23	0.06	0.95	18.88	0.22	0.86	37.64	0.15	0.87	48.17	0.28	0.76	48.83	0.17	0.76	54.81
7	0.04	0.95	14.51	0.04	0.95	14.39	0.12	0.87	49.18	0.11	0.89	41.38	0.24	0.81	87.05	0.24	0.84	85.83
8	0.11	0.95	13.82	0.07	0.95	16.71	0.23	0.86	36.10	0.18	0.86	46.19	0.33	0.76	62.71	0.25	0.77	59.13
9	0.05	0.97	14.92	0.04	0.97	14.13	0.14	0.91	40.26	0.10	0.91	35.99	0.20	0.84	61.13	0.11	0.84	47.51
10	0.02	0.98	15.44	0.02	0.97	14.74	0.07	0.93	53.95	0.06	0.93	29.06	0.15	0.89	78.35	0.06	0.88	37.05
Ave.	0.04	0.96	15.40	0.03	0.96	15.22	0.12	0.90	44.91	0.09	0.90	45.40	0.19	0.83	72.72	0.14	0.84	62.71
PA	0.50			0.50			0.51			0.50			0.51			0.49		

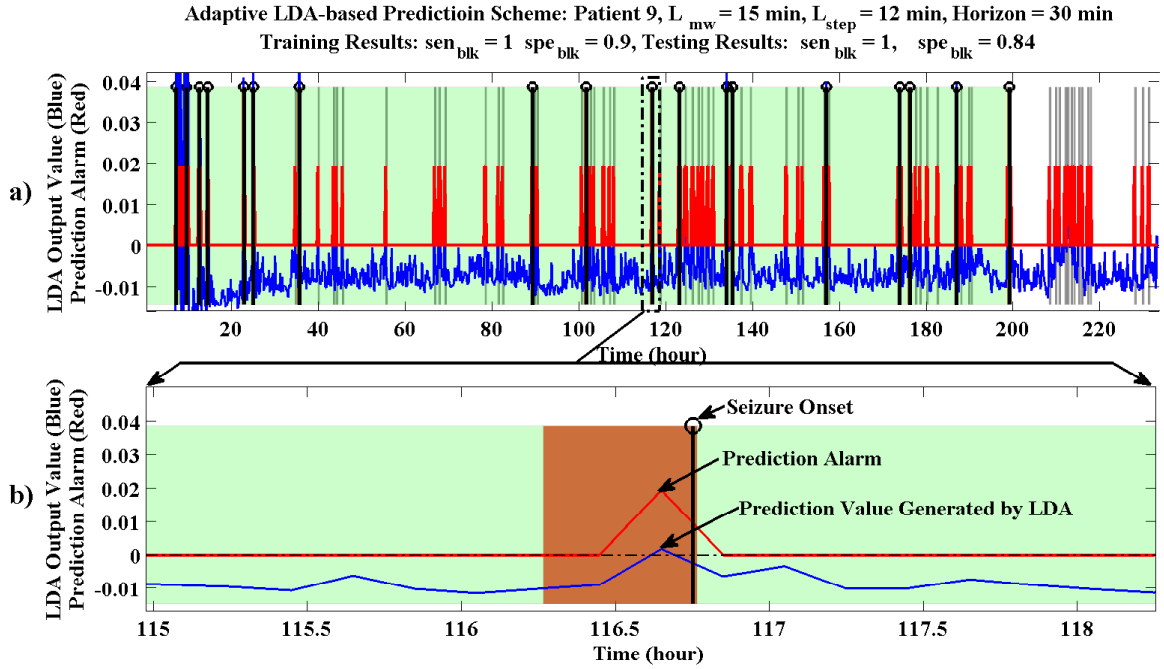


Fig. 7: a) A demonstration of the ALP prediction scheme for patient 9. The vertical black lines that span the graph indicate seizure occurrences. On top of each seizure line, the mark 'O' indicates a seizure was correctly predicted and the mark 'X' indicates the seizure was miss-predicted. The blue line in the graph represents the prediction values generated by the LDA classifier over time. An EEG epoch is classified as 'pre-seizure' if a LDA-prediction value is higher than 0. The red vertical line represents prediction alarms over time. b) A magnified snapshot with one seizure. The brown-colored area indicates the pre-seizure period. The LDA-predicted values and a prediction alarm are annotated in the graph.

learning predictive patterns online for each individual patient, and construct a personalized adaptive prediction rule for each patient autonomously. This significant feature of the proposed prediction system makes it convenient to be embedded in existing EEG recording systems, and provide personalized seizure prediction.

IV. Conclusions

This study investigated the challenging problem of epileptic seizure prediction. We introduced a new online feature extraction approach to characterize patterns of massive multivariate EEG data. We also proposed a new personalized online pattern learning framework by constructing an online pattern library for each individual patient. Three adaptive prediction rules were proposed and tested to perform online prospective seizure prediction from long-term EEG recordings of 10 patients with epilepsy. The APP scheme employed a probabilistic prediction rule to discriminate pre-seizure and normal patterns in discrete feature space. The ALP scheme employed the binary classification technique LDA to construct a linear decision boundary to classify pre-seizure and normal patterns. And the ANBP scheme employed the Bayesian decision theory to determine normal and pre-seizure patterns by posterior probabilities calculated from the pattern library. The experimental outcomes were promising compared to current seizure prediction approaches that mostly are offline, non-adaptive, non-prospective, and non-personalized. The rela-

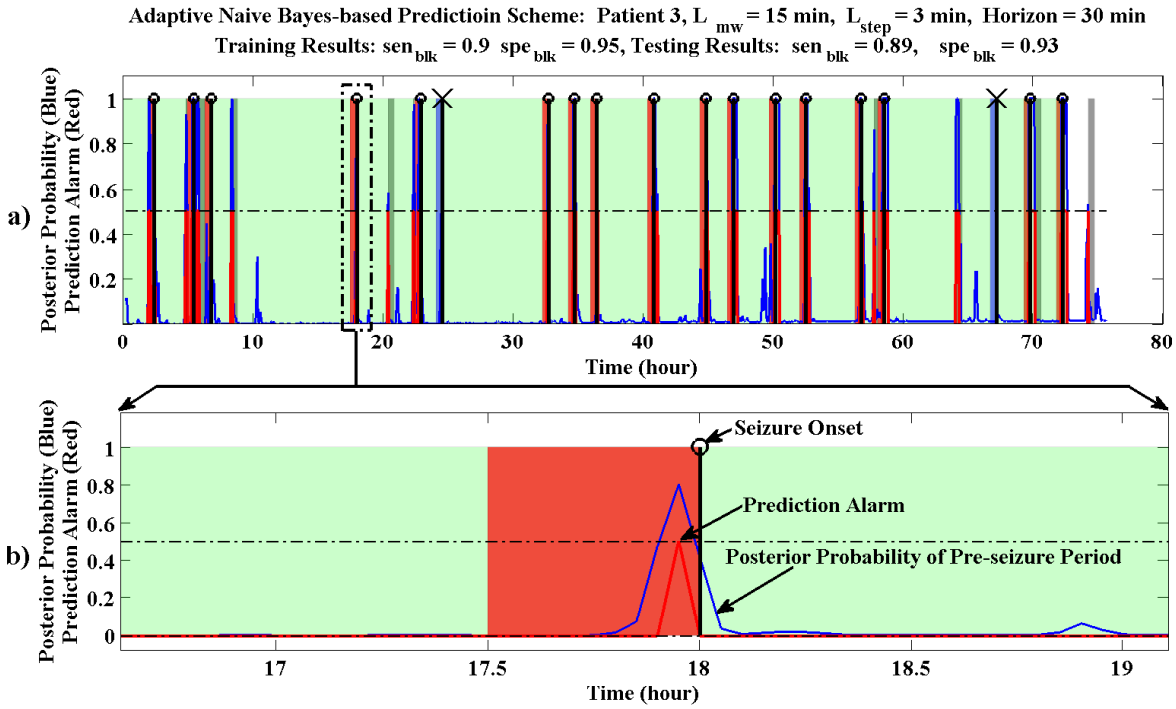
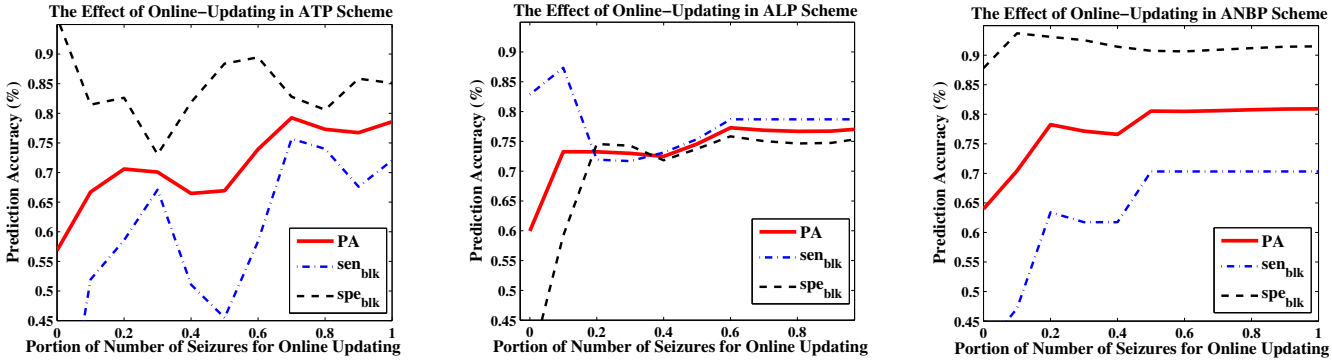


Fig. 8: a) A demonstration of the ANBP prediction scheme for patient 3. The vertical black lines that span the graph indicate seizure occurrences. On top of each seizure line, a mark 'O' indicates a seizure was correctly predicted and a mark 'X' indicates the seizure was miss-predicted. The blue line represents the posterior probability of pre-seizure for a monitored pattern. The horizontal dot-dashed line is the probability threshold of 0.5. If a posterior probability is higher than 0.5, it indicates an EEG epoch is more likely to be in 'pre-seizure' period, and thus a prediction alarm is generated. The Naive Bayesian prediction rule is updated after each seizure onset. The red line represents prediction alarms over time. b) A magnified snapshot of prediction for several hours with one seizure onset. The brown-colored area indicates the pre-seizure period. The LDA, the prediction score, and prediction warning are annotated in the graph.



(a) (b) (c)

Fig. 9: The effectiveness of the adaptive online-updating strategy for the three prediction schemes. In each subplot, the horizontal axis indicates the portion of seizures the prediction rule was updated. The point 0 means that only the initial trained prediction rule was employed, and it was unchanged throughout the prediction process. The point 1 means that the prediction rule was updated in the entire experimental time period. The red line shows the overall prediction accuracies, the black dashed line shows the spe_{blk}, and the blue dash-dotted line shows the sen_{blk}. An increasing trend of prediction accuracy can be clearly observed in each subplot. This observation confirms that the proposed prediction framework has a strong incremental learning capability and is capable of improving prediction performance online as it learns more from an individual patient over time.

tive high prediction accuracies achieved in this study reflect the effectiveness of the proposed online feature extraction and prediction framework.

For the ALP and ANBP schemes, a potential drawback

is that the constructed pattern library in continuously may increase rapidly in size over time. Accordingly, the computational load to re-train the LDA classifier or the Naive-Bayes Posterior Probabilities could increase significantly over time. Another drawback for the ALP scheme is that its prediction performance may be seriously deteriorated by monitoring noises and outliers. On the other hand, the probabilistic APP scheme constructs a pattern library based on pattern clustering. So, the resulting advantage of APP scheme is that the size of the pattern library is relatively small due to feature space discretization. Thus, the pattern-cluster library may only require a small memory space. In our experiments the total number of stored pattern clusters was at a level of one thousand, and the number of pre-seizure pattern clusters was at a level of one hundred. Another significant technical advantage of the APP scheme over the other two prediction schemes is that it is not sensitive to signals noises and outliers for online prediction. A warning is triggered only if the monitored pattern cluster is an already identified pre-seizure pattern cluster in the pattern library. All other monitored patterns (including any noise patterns and outliers) cannot generate prediction alarms. This property of the proposed probabilistic APP prediction is attractive to reduce false alarms in real-life clinical applications.

This study confirmed the hypothesis that it is possible to prospectively predict impending seizures. The proposed adaptive learning framework is capable of self-adjusting decision boundaries autonomously based on an incremental online-updated pattern library. Moreover, with online pattern learning ability, the proposed online adaptive prediction system has a great potential to further improve the prediction performance if more EEG data are available for each patient. The proposed adaptive learning approach is a pilot framework that can be potentially benefit a wide range of patients with epilepsy. The long-term goal of this research is to design an intelligent machine-learning interface to achieve accurate personalized seizure prediction. The developed prediction system could eventually be incorporated in closed-loop devices that deliver pre-seizure targeted therapy to the brain to avert seizure occurrence [12].

REFERENCES

[1] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. Voss, J. Timmer, and A. Schulze-Bonhage. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. *Brain*, 126:2616–2626, 2003.

[2] L. Chisci, A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio, and F. Fuggetta. Real-time epileptic seizure prediction using AR models and support vector machines. *IEEE Transactions on Biomedical Engineering*, 57(5):1124–1132, 2010.

[3] S. Dangel, P. Meier, H. Moser, S. Plibersek, and Y. Shen. Time series analysis of sleep EEG. *Computer assisted Physics*, pages 93–95, 1999.

[4] J. Engel and T. Pedley. *Epilepsy: A Comprehensive Textbook*. Lippincott Williams & Wilkins, Philadelphia, PA, 1997.

[5] H. Feldwisch-Drentrup, B. Schelter, M. Jachan, J. Nawrath, J. Timmer, and A. Schulze-Bonhage. Joining the benefits: combining epileptic seizure prediction methods. *Epilepsia*, 51(8):1598–1606, 2010.

[6] K. Fukunaga. *Statistical Pattern Recognition, seconde edition*. Academic Press, 1990.

[7] L. Good, S. Sabesan, S. Marsh, K. Tsakalis, D. Treiman, and L. Iasemidis. Nonlinear dynamics of seizure prediction in a rodent model of epilepsy. *Nonlinear Dynamics Psychol Life Science*, 14(4):411–434, 2010.

[8] L. Iasemidis. *On the dynamics of the human brain in temporal lobe epilepsy*. PhD thesis, University of Michigan, Ann Arbor, 1991.

[9] L. Iasemidis, J. Principe, J. Czaplewski, R. Gilman, S. Roper, and J. Sackellares. Spatiotemporal transition to epileptic seizures: A nonlinear dynamical analysis of scalp and intracranial EEG recordings. In *Spatiotemporal Models in Biological and Artificial Systems*, pages 81–88. Amsterdam: IOS Press, 1997.

[10] L. Iasemidis, D. Shiau, W. Chaovalitwongse, J. Sackellares, P. Pardalos, J. Principe, P. Carney, A. Prasad, B. Veeramani, and K. Tsakalis. Adaptive epileptic seizure prediction system. *IEEE Transactions on Biomedical Engineering*, 50(5):616–627, 2003.

[11] L. Iasemidis, D. Shiau, P. Pardalos, W. Chaovalitwongse, K. Narayanana, A. Prasada, K. Tsakalis, P. Carney, and J. Sackellares. Long-term prospective online real-time seizure prediction. *Clinical Neurophysiology*, 116:532–544, 2005.

[12] L. D. Iasemidis. Seizure prediction and its applications. *Neurosurgery Clinics of North America*, 22(4):489–506, 2011.

[13] K. Lehnertz and C. Elger. Can epileptic seizures be predicted? evidence from nonlinear time series analysis of brain electrical activity. *Physics Review Letters*, 80:5019–5022, 1998.

[14] F. Mormann, R. Andrzejak, C. Elger, and K. Lehnertz. Seizure prediction: The long and winding road. *Brain*, 130(2):314–333, 2007.

[15] F. Mormann, T. Kreuz, C. Rieke, R. Andrzejak, A. Kraskov, P. David, C. Elger, and K. Lehnertz. On the predictability of epileptic seizures. *Journal of Clinical Neurophysiology*, 116(3):569–587, 2006.

[16] Y. Park, L. Luo, K. Parhi, and T. Netoff. Seizure prediction with spectral power of eeg using cost-sensitive support vector machines. *Epilepsia*, 52(10):1761–1770, 2011.

[17] H. Persson, K. Alberts, B. Farahmand, and T. Tomson. Risk of extremity fractures in adult outpatients with epilepsy. *Epilepsia*, 43(7):768–772, 2002.

[18] H. Potter. Anatomy of the brain. http://faculty.ucc.edu/biology-potter/TheBrain/, 2006.

[19] P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, 1994.

[20] R. Quiroga, J. Arnhold, K. Lehnertz, and P. Grassberger. Kulback-leibler and renormalized entropies: applications to electroencephalograms of epilepsy patients. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 62:8380–8386, 2000.

[21] M. L. V. Quyen, V. Navarro, M. Baulac, B. Renault, and J. Martinerie. Anticipation of epileptic seizures from standard EEG recordings. *The Lancet*, 361(9361):970–971, 2003.

[22] P. Rajdev, M. Ward, J. Rickus, R. Worth, and P. Irazoqui. Real-time seizure prediction from local field potentials using an adaptive Wiener algorithm. *Computers in biology and medicine*, 40(1):97–108, 2010.

[23] P. Rapp, T. Bashore, J. Martinerie, A. Albano, I. Zimmerman, and A. Mess. Dynamics of brain electrical activity. *Brain Topography*, 2:99–118, 1989.

[24] J. Sackellares, D. Shiau, J. Principe, M. Yang, L. Dance, W. Suharitdamrong, W. Chaovalitwongse, P. Pardalos, and L. Iasemidis. Predictibility analysis for an automated seizure prediction algorithm. *Journal of Clinical Neurophysiology*, 23(6):509–520, 2006.

[25] B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, J. Timmer, and A. Schulze-Bonhage. Do false predictions of seizures depend on the state of vigilance? a report from two seizure prediction methods and proposed remedies. *Epilepsia*, 47(12):2058–2070, 2006.

[26] C. Silva, I. Pimentel, A. Andrade, J. Foreid, and E. Ducla-Soares. Correlation dimension maps of EEG from epileptic absences. *Brain Topography*, 11:201–209, 1999.

[27] J. Vastano and E. Kostelich. Comparison of algorithms for determining Lyapunov exponents from experimental data. In *International conference on dimensions and entropies in chaotic systems*, pages 100–107, Pecos River, NM, USA, 1985.

[28] S. Viglione and G. Walsh. Epileptic seizure prediction. *Electroencephalography and Clinical Neurophysiology*, 39:435–436, 1975.

[29] S. Wang. *Online Monitoring and Prediction of Complex Time Series Events From Nonstationary Time Series Data*. PhD thesis, Rutgers University, 2012.