

UCLA

UCLA Electronic Theses and Dissertations

Title

Learning characteristic neuronal dynamics of neurological disorders

Permalink

<https://escholarship.org/uc/item/2jq440v6>

Author

Monsoor, Tonmoy

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Learning characteristic neuronal dynamics of neurological disorders

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical and Computer Engineering

by

Tonmoy Monsoor

2024

© Copyright by
Tonmoy Monsoor
2024

ABSTRACT OF THE DISSERTATION

Learning characteristic neuronal dynamics of neurological disorders

by

Tonmoy Monsoor

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Los Angeles, 2024

Professor Vwani P. Roychowdhury, Chair

The understanding of the origins of neurological disorders, such as Epilepsy, Alzheimer's disease (AD), represents one of the most urgent and challenging areas of current scientific enquiry. In USA alone, 20% of the general population fall into one of these categories, thus creating an enormous need for medical intervention. Neurological disorders lead to system-level deficits which can cause disruptions in structural connectivity, functional organization, and information processing across various brain regions. Characterizing these system-level deficits from neuronal dynamics perspective is crucial for developing targeted interventions aimed at restoring normal brain function. Fueled by the rapid advancement in neural recording technologies both at the single and population level, we develop a data driven framework to characterize the neuronal dynamics in neurological disorders that can advance our understanding, diagnosis, and treatment of neurological disorders, ultimately improving patient outcomes and quality of life.

In this thesis, we study one of the most prevalent and debilitating neurological disorders called Epilepsy: approximately 65 million people suffer from it globally. In patients with epilepsy, the normal signaling mechanism in the brain is disrupted by sudden and

synchronized bursts of electrical pulses, leading to recurrent episodes of seizures. Epileptic seizures can be broadly classified into two types: generalized seizures which involve multiple cross-hemisphere epileptic foci and focal seizures where the epileptic focus is localized to a specific brain region. About one-third of patients with focal seizures, cannot be treated with anti-seizure medications and they need to undergo a resective surgery for the removal of Epileptogenic zone (EZ), which is the site of the cortex responsible for generating seizures. In the pre-surgical stage, the patient is placed under intracranial EEG (iEEG) monitoring in the hospital leading to iEEG recordings during actual seizures, referred to as ictal segments. Synchronous electrical signals recorded in the ictal segments have been modeled as network/collective dynamics involving all the channels leading to automated identifications of channels that drive the observed seizures. These channels are referred to as seizure onset zones (SOZs) as they constitute parts of the EZ active during observed seizures. Another correlate of this synchronous activity are short duration oscillatory field potential, known as High Frequency Oscillations (HFO), that are observed at the level of individual electrodes of iEEG. SOZ channels have distinctly higher rates of HFOs during the ictal segment, allowing neurologists to identify SOZs without any explicit network modeling. Once the SOZ has been identified then the surgeons resect the SOZ if possible. However, this approach has led to success in about only 60 – 70% of the patients because there might be parts of EZ that did not participate in generating seizures during the limited observation window; such unobserved parts of EZ are known as potential seizure onset zones (PSOZs). The inability of SOZ to completely encapsulate EZ in many patients, along with the hardships and risks associated with lengthy hospitalization period -often lasting two weeks or more- has prompted the need to find accurate physiological biomarkers of EZ during the interictal period, i.e, the majority of the time when patients do not have seizures.

HFOs observed during ictal periods have also been observed to be present at higher rates in SOZ channels (determined from ictal periods) during interictal periods, leading to the hope that resection of channels with high interictal HFO rates would lead to seizure

freedom. However, the presence of HFOs arising from cognitive processes (physiological HFOs) during interictal periods have diminished the predictive power of interictal HFO rate in the context of surgical outcome prediction. In the first part of the thesis, we develop a weakly supervised deep learning model to filter out the physiological HFOs and thus extract the pathological cluster of HFOs: epileptogenic HFOs (eHFO). In retrospective validation on a patient cohort of 15 patients, the eHFO cluster was found to be a better biomarker of EZ compared to Real HFO cluster (HFO cluster after filtering for artifacts) as it was able to correctly predict the post surgical outcomes of patients with an F1 score of 87% in comparison to Real HFO cluster’s 72%. However, when tested on a much larger patient cohort (159 patients), we found that a significant percentage of patients ($\sim 25\%$) did not have enough HFO detections and as a result the eHFO resection ratio was not able to correctly predict the surgical outcomes of those patients. Therefore, there is a need to look beyond HFOs in the space of potential interictal biomarkers of EZ.

Recently, there is a growing interest in determining whether synchronization effects can be observed in the interictal period and their temporal dynamics can be leveraged to delineate EZ. The problem of studying such effects and using them for better surgical outcome prediction is still open and we address this in the second part of the thesis. In particular, we use Power-Phase coupling amongst channels to construct a sequence of directed weighted networks from interictal segments. We leverage the topological dynamics of the network, both local and global, to train a machine learning model to identify SOZ (ground truth obtained from ictal data) from purely interictal segments. The model identifies the SOZ with over 95% accuracy. One of the hallmarks of the constructed networks is that they occasionally transition into a state of hyper-synchrony with SOZ and PSOZ nodes being the hub of these hyper-synchronous states. We hypothesize that these hyper-synchronous states are ‘mini seizures’ in the interictal phase and our machine learning model is able to identify them and use them for not only accurate SOZ identification but also identify PSOZ. The only way to validate whether our model has truly identified PSOZ and hence EZ is through

surgical outcome prediction. In the third part of thesis, we construct a set of features from SOZ model prediction scores along with the constructed network flow dynamics to propose a network based biomarker of EZ. In retrospective validation on a patient cohort of 159 patients, the network based biomarker was able to correctly predict the post surgical outcomes of patients with an F1 score of 89%. Finally, we develop an integrated framework to exploit the interplay between the constructed network and pathological HFO cluster to propose a novel biomarker of EZ. At the heart of this framework is a regression model which predicts the pathological HFO rate using a mixture of local and global properties of the constructed network. The summary statistics of the R^2 of the regression model along with the previously computed features (SOZ model prediction scores and epileptic network flow dynamics) is proposed as a novel biomarker of EZ. In retrospective validation on a patient cohort of 159 patients, the novel biomarker was able to correctly predict the post surgical outcomes of patients with an F1 score of 98%. A closer inspection into the summary statistics of the regression model for non-seizure free patients reveals a cluster of pathological HFO whose rate cannot be explained by the constructed network thus revealing the presence of a brain region capable of generating seizures outside of the one sampled by iEEG.

The dissertation of Tonmoy Monsoor is approved.

Hiroki Nariai

Xiang 'Anthony' Chen

Jonathan Kao

Lieven Vandenberghe

Vwani P. Roychowdhury, Committee Chair

University of California, Los Angeles

2024

To my beloved Abbu who left us too early

TABLE OF CONTENTS

1	Introduction	1
1.1	Overview	1
1.2	System level deficits in neurological disorders	4
1.2.1	Alzheimer’s Disease (AD)	5
1.2.2	Parkinson’s Disease (PD)	6
1.2.3	Multiple Sclerosis (MS)	8
1.2.4	Epilepsy	8
1.3	Neural recording techniques	9
1.4	Epilepsy: A prevalent and debilitating neurological disorder	12
1.4.1	Types of epileptic seizures and clinical manifestations	14
1.4.2	Treating focal drug resistant epilepsy	16
1.4.3	High frequency oscillations: a putative interictal biomarker of EZ	17
1.4.4	Synchronization network properties: ictal biomarker of EZ	25
1.5	Problem Statement and Contributions	28
1.5.1	Unsupervised learning of spatiotemporal interictal discharges in fo- cal epilepsy: an exterior method for non-negative matrix factorization (eNMF)	29
1.5.2	Purification of detected HFO clusters	31
1.5.3	Novel interictal biomarker of epileptogenic zone	35
1.6	Organization of the rest of the thesis	36
2	Exterior Method for Nonnegative Matrix Factorization	37

2.1	Introduction	37
2.2	Prior work	39
2.3	Exact NMF, Complexity, and SVD	41
2.3.1	Going from SVD to known Exact NMF factors	42
2.3.2	Going from SVD to unknown Exact NMF factors	43
2.4	General NMF: Going from SVD to the Positive Orthant	45
2.5	Exterior NMF Algorithm	46
2.5.1	Optimal orthogonal transformation of unconstrained global optimum	46
2.5.2	Penalty method for attaining feasibility	48
2.5.3	Descent to the local minima	50
2.6	Exterior NMC algorithm	52
2.7	Experiments and Results	53
2.7.1	Datasets and algorithms	53
2.7.2	Reconstruction error and run time comparison	55
2.7.3	Equivalent local minima and relationships among factors obtained by different algorithms	59
2.8	Concluding remarks	66
2.9	Exact NMF solution	72
2.9.1	Updates for W_1 and W_3	74
2.9.2	Updates for W_2 and R_1	75
2.9.3	Updates for R_2 and Λ	76
2.10	Generalized orthogonal transformation	77
2.11	Checking permutation invariance between two factor matrices	78

3	Optimizing detection and deep learning-based classification of pathological high-frequency oscillations in epilepsy	80
3.1	Introduction	80
3.2	Methods	82
3.2.1	Patient Cohort	82
3.2.2	Standard protocol approvals, registrations, and patient consent	83
3.2.3	Patient evaluation	83
3.2.4	Subdural electrode placement	83
3.2.5	Acquisition of three-dimensional (3D) brain surface images	84
3.2.6	Intracranial EEG (iEEG) recordings	84
3.2.7	Automated detection of high frequency oscillations	84
3.2.8	Creation of each HFO set and subsequent training of deep learning-based classification (Fig. 3.1: Study workflow)	85
3.2.9	Evaluating the performance of eHFO detectors in predicting postoperative seizure outcomes	92
3.2.10	Time-frequency plot characteristics of the HFOs in the three HFO sets	92
3.2.11	Distribution of HFO with spike and epileptogenic HFO in three HFO sets	93
3.2.12	Relationship between epileptogenic HFOs and HFOs with spikes . . .	93
3.2.13	Statistical analysis	93
3.3	Results	94
3.3.1	Patients' characteristics	94
3.3.2	Interictal HFO detection and HFOs' classification	94
3.3.3	Spectral characteristics of HFOs in three HFO sets	95

3.3.4	Overlap between eHFOs and HFOs with spikes in three eHFO detectors	95
3.3.5	Characteristic differences of eHFOs vs. non-eHFOs controlled by spike status	96
3.3.6	Distribution of eHFO in three HFO sets	96
3.3.7	Predicting postoperative seizure outcomes using HFO resection ratios	96
3.3.8	Post-hoc sensitivity analysis	97
3.4	Discussion	97
4	Characterizing physiological high-frequency oscillations using deep learning	107
4.1	Introduction	107
4.2	Methods	109
4.2.1	Patient cohort	109
4.2.2	Standard protocol approvals, registrations, and patient consents . . .	109
4.2.3	Patient evaluation	109
4.2.4	Subdural electrode placement	110
4.2.5	Acquisition of three-dimensional (3D) brain surface images	110
4.2.6	iEEG recording	111
4.2.7	Functional cortical mapping	111
4.2.8	Automated detection and classification of HFOs	112
4.2.9	Weakly supervised DL using functional cortical mapping results as labels	112
4.2.10	Model architecture and feature representation of HFOs	113
4.2.11	Model training	114
4.2.12	Performance analysis framework	115

4.2.13	Characterization of ecHFOs	115
4.2.14	Interpretability analysis of ecHFOs	115
4.2.15	Time domain characteristic ecHFO	116
4.2.16	Comparison of resection ratios of HFOs to postoperative seizure out-comes	117
4.2.17	Statistical analysis	117
4.3	Results	118
4.3.1	Clinical information (patient characteristics)	118
4.3.2	Interictal HFO detection	118
4.3.3	Time–frequency plot characteristics of ecHFO and non-ecHFOs . . .	119
4.3.4	Perturbation analyses	119
4.3.5	Feature characterization of ecHFOs vs. non-ecHFOs	120
4.3.6	Inference of channel characteristics	120
4.3.7	Clinical inference using postoperative seizure outcomes	120
4.4	Discussion	121
5	Novel interictal biomarker of epileptogenic zone	131
5.1	Introduction	131
5.2	Materials	133
5.2.1	Patient Cohort	133
5.2.2	Standard protocol approvals, registrations, and patient consent	133
5.2.3	Patient evaluation	133
5.2.4	Subdural electrode placement	134
5.2.5	Acquisition of three-dimensional (3D) brain surface images	134

5.2.6	Intracranial EEG (iEEG) recordings	134
5.3	Dynamic directed weighted interictal networks	135
5.3.1	Frequency domain synchronization network	135
5.4	Seizure onset zone identification using dynamic interictal network properties	140
5.4.1	Node feature vector	140
5.4.2	Node labeling	149
5.4.3	Training and Testing corpus for SOZ identification	150
5.4.4	Results	151
5.4.5	Discussion	152
5.5	Predicting post-surgical outcomes using interictal SOZ classification and net- work dynamics	154
5.5.1	3 dimensional feature vector	155
5.5.2	5 dimensional feature vector	162
5.5.3	Mini Seizures: 7 dimensional feature vector	170
5.6	A Biomarker of EZ agnostic to surgical resection and SOZ labels	176
5.6.1	Analyzing the four clusters	178
6	Concluding remarks and future work	187
	References	190

LIST OF FIGURES

1.1	Vicious cycle of ND [Linazasoro, 2009]: Neurological disorders (ND) which often starts from molecular level changes can lead to neuronal network dysfunction. These network level disruptions can trigger a feedback loop where they can lead to molecular level changes causing further disruptions in neuronal network connectivity	4
1.2	Memory encoding deficit in Alzheimers [Golby et al., 2005]: Alzheimer’s disease (AD), the most common cause of dementia in older adults leads to a deficit in declarative memory with relative sparing of implicit memory. This study uses fMRI to study the functional competency of specific brain regions and the relationship to the behavioral memory deficits in AD. The study involved five early-stage AD patients and five healthy age-matched neurologically normal control subjects (NC) with fMRI at 3 Tesla during intentional memory encoding of scenes.	6
1.3	Structural connectivity in PD [Galantucci et al., 2017]: Affected structural connections in patients with PD and MCI relative to healthy control subjects and patients with Parkinson disease (PD) without mild cognitive impairment (MCI). The principal connected component is shown in red, and the other affected connections not included in the principal connected component are shown in green. Subnetworks show altered structural connectivity (decreased fractional anisotropy or increased mean diffusivity) in patients with PD and MCI relative to healthy control subjects at $P < .01$	7
1.4	EEG node Connectivity in Epilepsy [Bomela et al., 2020]: Networks (graph representation and the heat-map representation of the corresponding adjacency matrix) of the normal state (a) and seizure state (b)	9

1.5	Two Methods of Intracranial EEG: Subdural Electrodes(SDE) and Stereo-EEG (SEEG) While grids and strips of subdural electrodes provide a large coverage over the bare surface of the cerebral cortex, they are often implanted in one hemisphere and do not reach deeper brain structures (e.g., hippocampus or insula). By comparison, depth electrodes can enable bilateral monitoring of superficial and deep cortical structures but only the most superficial and deep contacts will be within the cortical gray matter while the rest of the contacts are placed in the white matter.	11
1.6	Seizure generation and propagation [Osorio et al., 2016]: Intracranial EEG recording obtained from bilateral orthogonally implanted eight-contact depth electrodes; only the signal from the first (deepest), third, and fifth contacts of each depth electrode are shown, with LB1/RB1 situated in the hippocampi and LC1/RC1 in the parahippocampal gyri. Left mesial temporal regional seizure onset in hippocampus and parahippocampal gyrus (LB1 > 3 > 5 and LC1 > 3 > 5), with subsequent propagation (“spread”) to contralateral right mesial temporal region. Note that the contralateral propagation can also be considered as the abrupt onset of an independent seizure in the right temporal lobe because the ictal activity might not be synchronous between the two hemispheres. Seizure prediction signal analysis techniques need to uncover predictive features from the type of signal seen in the left-sided electrodes before the seizure onset (interictal segment) and in the right-sided electrodes before the seizure onset (interictal segment) to differentiate between independent and propagated seizures	13
1.7	Drug resistant epilepsy: patients with epilepsy who do not respond successfully to antiseizure medications. Almost a third of patients with epilepsy suffer from drug-resistant epilepsy	16
1.8	High Frequency Oscillations (HFO): Short-duration oscillatory field potential with spectral content in the 80-500 Hz frequency range, consisting of at least four oscillations that can be distinguished from background.	18

1.9	Workflow of STE HFO detector: Energy function based on root mean squared amplitude is used to compute the energy in short EEG segments. EEG segments with energy above five times the standard deviation (SD) of the mean energy of the entire EEG recording and lasting more than 6 ms were considered as putative HFO events. Putative HFO events are filtered based on some comparison with mean value of rectified signal	21
1.10	Workflow of MNI HFO detector: HFO detections consist of two stages. In the first stage, channels with lots of baseline segments are identified and then HFOs are detected in those channels using an algorithm similar to the STE detector. In the second stage, an iterative algorithm is used to detect HFOs in channels with few baseline segments	22

1.12	Drivers of seizure episodes (SOZ) identification: During this ictal segment, boxed in blue, visually all the red electrodes seem to behave in a very synchronized manner, so doctors mark them as the drivers of this seizure episode. However, since the synchronized activity between the channels happen in such a short time scale so often it is very difficult to identify the channels driving these kind of activity visually. In this example, only a subset of the red electrodes might be the drivers of the synchronized activity. Therefore, network based approaches have also been developed to identify the drivers of such synchronized activity in the ictal segment.	25
1.13	Potential Seizure Onset Zone (PSOZ): Parts of EZ not generating seizures even during 2 week long observation - such regions are called Potential Seizure Onset Zone (PSOZ)	27
1.14	HFO types: Detections by automated detectors can be broadly classified into three categories: i) artifacts, ii) physiological HFO stemming from healthy tissue and iii) pathological HFO stemming from epileptogenic tissue. While there exists well developed algorithms for filtering artifacts, there is a need for developing algorithms to classify into the other two categories.	32
2.1	Distance to the positive orthant (as defined by the sum of the absolute values of the negative entries of the factor matrices post optimization; see Eqs.(2.6) and (2.8) is plotted for synthetic exact datasets with various dimensions and sparsity levels. For lower dimensions, ADMM implementation of (6) generates exact non-negative solutions. For larger dimensions, ADMM implementations for both the general and relaxed version (involving only a single orthogonal matrix R_2 , and setting $R_1 = \Lambda = I$) generate factors that are only close to the positive orthant (PO), thus requiring further processing. . .	44

2.2	Given enough run time, all the algorithms converge to local minima with similar reconstruction error (<i>across all the datasets we used</i>): The above figure illustrates it for the Synthetic dataset with $n = m = 5000, r = 500$ and SNR of $-28dB$. Relative error $= \frac{\ X - U_{NMF} V_{NMF}^T\ _F}{\ X - U_{SVD} V_{SVD}^T\ _F}$ is plotted as a function of run time (secs). <i>eNMF achieves the minimum (i.e. unconstrained global minimum in this case) in 106 seconds, compared to hours of run time for some of the competing algorithms</i> Competitors like AO-ADMM have a fast convergence initially but slow down when approaching the minima, as shown by the long tail.	56
2.3	Exact factorization: eNMF achieves the global minimum for sparsity levels 0.2 through 0.5. Although ALS and HALS achieve the global minimum for sparsity levels 0.3 through 0.5, but they are far away from the global minimum for sparsity level 0.2. For sparsity level 0.1, although HALS and AO-ADMM get similar reconstruction error as ENMF does, they usually have a fast initial convergence but slow down considerably as it approaches the global minimum as shown in Fig. 2.2. Note: We dropped NMF-ADMM from this plot due to its high reconstruction error.	58
2.8	Asymptotic/Tending Equivalence of Factor Matrices: We plotted the percentage of columns in factor matrices returned by eNMF (Algorithm A) that are a scaled permutation of the columns in factor matrices returned by AO-ADMM (Algorithm B) as a function of the reconstruction error difference between the algorithms. As the reconstruction error difference gets smaller, more columns in the factor matrices returned by AO-ADMM converge to a scaled permutation of the columns in the factor matrices returned by eNMF. The percentage of the columns converged approaches 100% as reconstruction error approaches zero in an asymptotic manner showing the tendency of the factor matrices to approach equivalent local minima.	64

2.4	Equal-time constrained NMF reconstruction errors (Synthetic datasets): Relative Reconstruction Error (RE) $\frac{\ X - U_{\text{NMF}} V_{\text{NMF}}^T\ _F}{\ X - U_{\text{SVD}} V_{\text{SVD}}^T\ _F}$ is plotted as a function of the number of latent dimensions (r) for various SNR levels. <i>eNMF achieves the unconstrained global minimum for all latent dimensions and across all the four SNR levels.</i> Among its competitors, only NMF-ADMM attains the global minimum for lower latent dimensions. For higher latent dimensions, the gap in relative error of eNMF and its closest competitor AO-ADMM widens for lower SNR levels (more noise). Note: For a zoomed in picture of the scenario, we have dropped the curves corresponding to the algorithms with a high relative error.	68
2.5	Convergence speed (Synthetic datasets): Algorithm 1 of eNMF directly finds globally optimal solutions : Relative error = $\frac{\ X - U_{\text{NMF}} V_{\text{NMF}}^T\ _F}{\ X - U_{\text{SVD}} V_{\text{SVD}}^T\ _F}$ is plotted as a function of run time (secs) for various SNR levels. <i>The orthogonal part of eNMF (Algorithm 1) finds non-negative factors by itself, thus achieving globally optimal solutions for all SNR levels.</i> Moreover, eNMF converges to the global minimum at the fastest rate (compared to SOTA algorithms) for all SNR levels. As the noise level increases (lower SNR), the gap in the convergence rates of eNMF and AO-ADMM (its closest competitor) widens. Note: In order to get convergence plots where the relative performance of top algorithms can be clearly visualized, we have dropped the convergence curves of slower algorithms.	69
2.6	Equal-error constrained time complexity (Real datasets): Run time is plotted on a log scale as a function of the number of latent dimensions (r). <i>eNMF performs significantly better than in its competitors in terms of time complexity with the margin widening for higher latent dimensions.</i>	70
2.7	Equal-time constrained NMC errors: Relative reconstruction error (RE) $\frac{\ P_{\Omega}(X - UV^T)\ _F}{\ P_{\Omega}X\ _F}$ and Relative prediction error $\frac{\ X_{\text{test}} - UV^T\ _F}{\ X_{\text{test}}\ _F}$ is plotted as a function of the number of latent dimensions (r). <i>eNMF achieves the lowest reconstruction and prediction error for all latent dimensions.</i>	70

3.1	Union eHFO detector Workflow	86
3.2	CNN architecture: The architecture of the model was adapted from Resnet-18. The last layer of the resnet18 was modified to be three fully-connected layers with LeakyReLU, BatchNorm, and 10% dropout in between. The output of the model was fed into a sigmoid function to bound the output between 0 to 1, representing the probability of each task.	89
3.3	Network Input Images: The network input includes three images constructed from a one-second raw EEG segment with a detected HFO in the center (500ms before and after). Left: The time-frequency plot was generated by continuous Gabor Wavelets ranging from 10Hz to 500Hz. Middle: EEG tracing plot was generated on a 2000 x 2000 image by scaling the time-series signal into the 0 to 2000 range. Right: amplitude-coding plot contains the amplitude at every time point; a column of the image represented the signal's actual value rescaled with a color gradient. These three images were resized into 224 x 224 in order to fit into the neural network.	90
3.4	Distribution of HFOs with spikes in three HFO sets: (A) An example of HFO with spike from each HFO set (Intersection, MNI only, and STE only) is shown. The first row shows the original tracing, the second row shows the filtered tracing (80–500 Hz), and the third shows the time–frequency plot. (B) The proportion of HFO with spikes in 15 patients is plotted as box plots and labeled by the HFO set type. Across a patient body of 15 patients, the Intersection set has the highest proportion of HFO with spikes on average, followed by the MNI and STE only set. The proportion of HFOs with spikes was significantly higher in the Intersection, and MNI only set compared to the STE only set (Intersection vs. STE only: mean 0.88 versus 0.36, $p < 0.001$; MNI only vs. STE only: mean 0.87 versus 0.36, $p < 0.01$). HFOs: high-frequency oscillations; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.	101

3.5	Time-frequency plot characteristics of HFOs in three HFO sets:	(A) The time-frequency plots of HFOs, averaged across all patients ($n = 15$), are plotted and labeled based on their set type and four distinct frequency bands. Across four frequency bands, HFOs in the Intersection set have higher amplitude than HFOs in the MNI only set and STE only set. HFOs in the MNI only set have higher amplitude than STE only set in the 10–30 and 30–80 Hz frequency bands.	
	(B) Pixel-wise analysis of time-frequency plot amplitude differences among three HFO sets: We created a custom scalogram to visually represent the frequency ranges where the HFOs belonging to different sets varied. The lighter-colored regions in the figure indicate pixels where the mean amplitude of one HFO set is statistically higher (p-value below 0.005 from a one-tailed t-test) than the other HFO set. From the left plot, it is clear that the HFOs in the Intersection set have higher amplitude than HFOs in the MNI only set in all four frequency bands, with a noticeable difference in the 30–80 Hz and 80–250 Hz frequency bands. Similarly, the right plot shows that HFOs in the MNI only set have higher amplitude than HFOs in the STE only set in the 10–30 Hz and 30–80 Hz frequency bands, with a significant difference in the 10–30 Hz frequency band. HFOs: high-frequency oscillations; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.		102

3.6	Overlap between eHFO and HFO with spike set in three eHFO detectors:	The overlap ratio between eHFO and HFO with spike set in 15 patients is plotted as box plots and labeled by the eHFO detector type. Across the cohort of 15 patients, there was significant overlap between the eHFOs and HFOs with spikes for all three eHFO detectors. The overlap ratio was significantly higher in the MNI eHFO detector when compared to the Union and STE eHFO detectors (MNI vs. Union: mean 0.95 versus 0.79, $p < 0.001$; MNI vs. STE: mean 0.95 versus 0.78, $p < 0.001$) while there was no significant difference in the overlap ratio between Union and STE eHFO detectors (Union vs. STE: mean 0.79 versus 0.78, $p = 0.042$). HFOs: high-frequency oscillations; eHFOs: epileptogenic HFOs; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.	103
3.7	Time-frequency plot characteristics of eHFOs vs. non-eHFOs controlled by spike status:	The time–frequency plots of eHFOs and non-eHFOs (detected with Union HFO detector) averaged across all patients ($n = 15$) were plotted based on spike status (with or without spikes). Across four frequency bands, eHFOs with spikes showed higher amplitude than non-eHFOs with spikes, and eHFOs without spikes showed higher amplitude than non-eHFOs without spikes ($p < 0.0001$ in all comparisons with Welch t-test with unequal variances). HFOs: high-frequency oscillations; eHFOs: epileptogenic HFOs.	104

3.8	Distribution of eHFO in three HFO sets: (A) Proportion of eHFO in 15 patients is plotted as box plots and labeled by the HFO set type. Across the cohort of 15 patients, MNI only set has the highest proportion of eHFO on average, followed by Intersection and STE only set. The proportion of eHFO was significantly higher in the Intersection, and MNI only set when compared to the STE only set (Intersection vs. STE only: mean 0.82 versus 0.45, $p < 0.001$; MNI only vs. STE only: mean 0.85 versus 0.45, $p = 0.03$). (B) Confidence scores of eHFOs in 15 patients were plotted as a histogram and labeled by the HFO set type. The histogram for the Intersection and MNI only set has a skewed distribution strongly biased towards 1 (direction towards eHFO), while the STE only set has a more uniform distribution. HFOs: high-frequency oscillations; eHFOs: epileptogenic HFOs; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.	105
-----	---	-----

3.9 Accuracy of postoperative seizure outcomes prediction using HFO re-

section ratios: We constructed a postoperative seizure outcome prediction model using the resection ratio derived from 90-minute EEG data ($n = 15$) (same approach as in our previous work (Zhang et al. 2022a). Each receiver-operating characteristics (ROC) curve delineates the accuracy of seizure outcome classification of a given model, using the area under the ROC curve (AUC) statistics. A) The HFO resection ratio (detected HFOs filtered for artifacts) obtained from 3 models (Union, MNI, and STE) was used to predict the seizure outcome. Union and MNI detectors have higher AUC values than a model obtained by the STE detector. B) eHFO resection ratio (detected HFOs filtered for artifacts and non-eHFO) obtained from 3 models (Union, MNI, and STE) was used to predict the seizure outcome. Overall significant improvement in prediction performance was seen, with the Union detector model having the highest AUC of 0.89. HFOs: high frequency oscillations; eHFOs: epileptogenic HFOs; MNI: Montreal Neurological Institute (detector); STE: Short-term energy. 106

4.1 **Processing workflow:** (A) the detected HFO (by STE detector) is first filtered by the artifact detector that was developed in our previous study [Zhang et al., 2022b]. The constructed image features along with the channel-wise clinical information, are fed into the convolutional neural network to train the model (ecHFO detector). In the real application of the model (inference), an unseen EEG signal is sent into the pipeline, and the STE detector is applied to detect all HFOs. Then we construct the image feature based on the detected HFO. The same artifacts detector [Zhang et al., 2022b] is applied to reject the artifacts, and then the real HFO is filtered by the ecHFO detector. The detector will output a score from zero to one representing the confidence of each HFO being ecHFO. HFO feature representation and model architecture (B): we captured the time–frequency domain features as well as signal morphology information of the HFO window via three images. The time–frequency plot (scalogram) was generated by continuous Gabor wavelets ranging from 10 Hz to 500 Hz. The EEG tracing plot was generated on a 2000×2000 image by scaling the time-series signal into the 0–2000 range to represent the EEG waveform’s morphology. We constructed another two images from the original EEG signal serving as the additional inputs to the network. The amplitude-coding plot was generated to represent the relative amplitude of the time-series signal: for every time point, the pixel intensity of a column of the image represented the signal’s raw value at that time. These three images were resized into the standard size (224×224), serving as the input to the neural network. We used ResNet 18 with a modified output layer for binary classification. The weights in the convolution layers are frozen and serve as feature extractors in the convolutional neural network. 125

4.2 Characteristics in the time–frequency plot of ecHFOs against non-

ecHFOs: (A) The time–frequency plot characteristics of the eloquent cortex and non-eloquent cortex HFOs for Pt 4, 5, 7, and 8. The yellow-colored regions in the figure stood for the pixels, where the power spectrum of ecHFOs is statistically lower than (p-value below 0.05 from the one-tailed t-test) non-ecHFOs. The figure showed one set of clearly interpretable distinguishing features between ecHFOs and non-ecHFOs: the ecHFOs generally have lower power at lower frequencies during the HFO event (center part along the time axis), panel (B-Right) was generated by taking the average of the individual binary images from each of the 18 patients. It showed the distinguishing features are also significant at the population level. The feature can be assembled as a ‘Bell-Shape’ if we take the region >0.7 in the plot averaging all of the time–frequency plot characteristics (red color). (B-Left). (C) The model’s response to the Bell-shaped perturbation on the time–frequency plot. We provide two examples of perturbation for ecHFO events in Pt 3. Each row presents one example and the first column indicates the original time–frequency plot while the second indicates the perturbed time–frequency plot based on the Bell shape perturbation. The prediction value of the model changed from 0.780 (therefore originally predicting it as non-ecHFO) to 0.299 (thus a change of 0.481), implying that the perturbed HFO would correspond to a non-ecHFO. (D) The change in model confidence in population level. Each column (along the y-axis) is a histogram of the percentage of change in confidence for one distinct patient. It shows the frequency distribution of confidence changes after adding the bell shape perturbation to the time–frequency plot to all classified ecHFOs for the given patient. The change in confidence level is significant, with an average of 0.532 with a 95% confidence interval $[0.528, 0.535]$ noted as the solid red line in the histogram. HFO = high-frequency oscillation; ecHFOs = eloquent cortex HFOs; non-ecHFO = non-eloquent cortex HFOs 126

4.3	The model's responses to injecting a spike-like feature into the amplitude-coding plot: Examples of introducing a downgoing (A) and upgoing (C) spike feature into classified non-ecHFO events. These demonstrate that on the introduction of a spike-like perturbation, the model predicts higher confidence towards non-ecHFOs (B, D). Subfigure (A) shows the amplitude encoding image before perturb, spike template, and the after-perturb (top row), and the corresponding time-series signal with downgoing spike perturbation (bottom row). Similarly, subfigure (C) shows the same information on a different classified ecHFO but with upgoing spike perturbation. For each patient, we compute a histogram for the distribution of the change in confidence (B) for downgoing spike perturbation. The same steps are repeated for upgoing spike perturbation, and the results are shown in (D). The percentage of change in confidence for both up-and-downgoing spike perturbation is significantly greater than zeros, with means downgoing: 0.268 (95% confidence interval [0.263, 0.272]) and upgoing: 0.320 (95% confidence interval [0.316, 0.324]), which are noted as solid red lines in each histogram. HFO = high-frequency oscillation; ecHFOs = eloquent cortex HFOs; non-ecHFO = non-eloquent cortex HFOs.	127
4.4	Traditional feature characterization of ecHFO and Non-ecHFO: The normalized histogram of peak frequency (A), amplitude (B), and duration (C) of predicted ecHFO. (A) The ecHFO generally has a lower peak frequency than the non-ecHFO (p-value < 0.001). (B) The ecHFO generally has a smaller amplitude than the non-ecHFO (p-value < 0.001). (C) The ecHFO generally has the trend of having a longer HFO than the non-ecHFO (p-value = 0.07). However, there are no clear decision boundaries that can be drawn to clearly discriminate between ecHFO and non-ecHFO using each of these traditional features.	128

4.5	Inference of channel characteristics: (A) the ratio of ecHFOs (eHFOs/total HFOs) in different types (physiological: behavior or gamma only; pathological: SZ/AD or spike only; both: both physiological and pathological) of channels from all patients (n = 14) is plotted in box plots. The percentage of ecHFOs was higher in physiological channels than that in pathological channels. (B) The model confidence distribution of each individual eHFOs in channels with physiological, pathological, and both categories are shown. The distribution in both channels is closer to a uniform distribution. . . .	129
4.6	The accuracy of prediction models incorporating HFO resection ratio: We constructed postoperative seizure outcome prediction models using the HFO resection ratio derived from EEG data (n = 14). Each receiver-operating characteristics (ROC) curve delineates the accuracy of seizure outcome classification of a given model, using the area under the ROC curve statistics. (A) Unclassified HFO resection ratio was used as a single classifier. (B) non-ecHFO resection ratio was used as a single classifier. (C) non-ecHFO resection ratio was used as a single classifier. (D, E) A multiple regression model incorporating the resection ratio of HFOs (non-ecHFOs or spk-HFO) and complete removal of the SOZ (yes or no) was used, which demonstrated further improved predictive value of postoperative seizure outcomes. HFO = high-frequency oscillation; non-ecHFO = non-eloquent cortex HFOs; spk-HFO = HFO with spike; SOZ = Seizure onset zone. . . .	130
5.1	Interictal segment [Fernández-Torre, 2010]: Interictal segment (highlighted in green), when the patient is not having seizures, is always available in abundance and much longer in duration. However, as you can see that the synchronization effects in the interictal segments if any are much more subtle	132

5.2	Sequence of synchronization networks: One second long iEEG segment is used to construct one synchronization network. Assuming the interictal segment is T seconds long and non-overlapping one second long sliding window approach is used, we have a sequence of T synchronization networks	135
5.3	Synchronization network workflow: The first step involves selecting a frequency band of interest to filter the iEEG segment and transform it into the frequency domain. Once we have the frequency domain representation, then we construct a synchronization network to capture the interictal interactions between the iEEG channels	136
5.4	Frequency bands: Epilepsy research has focused on specific frequency bands depending on the research question they are trying to answer. We will be studying synchronization in High Gamma, Ripple and Fast Ripple band.	137
5.5	Edge directionality: If the total phase in channel j leads the total phase in channel i then we assign a directionality from i to j and if the total phase in channel j lags the total phase in channel i then we assign a directionality from j to i	139
5.6	SOZ classifier setup: The SOZ classifier takes in as input a node feature constructed from the interictal network property dynamics and outputs a probability of the node being an SOZ	140
5.7	Hypersynchronous state with SOZ node P3 as hub: A snapshot of the synchronization network constructed from a one second long interictal segment with blue nodes representing channels that drove seizures in hospital monitoring, red nodes representing channels that did not drive seizures in hospital monitoring but were resected by the doctors, and green nodes representing channels that neither drove seizures in hospital monitoring nor resected by the doctors. We can observe the emergence of a hypersynchronous state, characterized by a star like network, with a SOZ node P3 as the hub	141

5.8	Hypersynchronous state with SOZ node G53 as hub: Another snapshot of the synchronization network constructed from a different one second long interictal segment with blue nodes representing channels that drove seizures in hospital monitoring, red nodes representing channels that did not drive seizures in hospital monitoring but were resected by the doctors, and green nodes representing channels that neither drove seizures in hospital monitoring nor resected by the doctors. We can observe the emergence of a hypersynchronous state, characterized by a star like network, with a SOZ node G53 as the hub	142
5.9	Hypersynchronous state with Resected Non-SOZ node P1 as hub: Another snapshot of the synchronization network constructed from a different one second long interictal segment with blue nodes representing channels that drove seizures in hospital monitoring, red nodes representing channels that did not drive seizures in hospital monitoring but were resected by the doctors, and green nodes representing channels that neither drove seizures in hospital monitoring nor resected by the doctors. We can observe the emergence of a hypersynchronous state, characterized by a star like network, with a resected Non-SOZ node P1 as the hub	143
5.10	Sharp rise in fiedler eigenvalue during episodes of large scale synchronization: The Fiedler eigenvalue is graphed over time throughout the interictal period. Examination of the plot reveals distinct spikes in the Fiedler value at certain intervals. These intervals align with periods of hypersynchronous interictal network states, thereby confirming the efficacy of the Fiedler value in capturing such states.	144
5.11	Time evolution of out-degree of a node: We have plotted the time evolution of out-degree of a node in the network. From the plot, it can be observed that out-degree occasionally enters into an outlier zone marked in red	147

5.12	Time evolution of eigencentality of a node: We have plotted the time evolution of eigencentality of a node in the network. From the plot, it can be observed that the eigencentality occasionally enters into an outlier zone marked in red	147
5.13	20 dimensional feature vector per node for given frequency band: For a given frequency band, each node is represented by a 5 dimensional feature vector that is evolving over time. We summarize the time evolution of each feature through 4 sufficient statistics thus giving a 20 dimensional feature vector per node per frequency band	148
5.14	Node feature vector construction workflow: We study synchronization effects in three frequency bands: high gamma, ripple and fast ripple. Each frequency band results in a sequence of synchronization networks which in turn gives a 20 dimensional node feature. Concatenating the features across the 3 frequency bands we get a 60 dimensional feature vector per node.	148
5.15	Node labeling: We only consider seizure free patients for training and for those patients we label SOZ nodes as 1 and Preserved Non-SOZ nodes as 0	149
5.16	Training and Testing Corpus: 17 UCLA seizure free patients in the training corpus and the remaining 142 patients in the testing corpus resulting in one random forest model	150
5.17	Training and Testing Corpus for 17 UCLA seizure free patients: For each fold, we have 16 UCLA seizure free patients in the training corpus and the left out UCLA seizure free patient in the testing corpus resulting in one random forest model per fold	151

5.18	SOZ classification results: The SOZ classifier outputs a probability score and we use a threshold of 0.5 to make predictions: $P(i \in SOZ) \geq 0.5 \rightarrow i \in SOZ$. The classifier is able to correctly identify most of the SOZ channels in the test set shown by the high recall of 98%. Few false negatives come from patients where the SOZ labeling by doctors were very generous.	152
5.19	Surgical outcome prediction model: The surgical outcome prediction model takes feature vector of a patient as input and outputs a probability of the patient becoming seizure free which can be then thresholded to predict whether the patient will become seizure free or not	154
5.20	Multicenter patient corpus: A multicenter (UCLA, Wayne State) patient corpus of 159 patients with 113 of them being seizure free ($\sim 70\%$) and 46 of them being non-seizure free ($\sim 30\%$)	156
5.21	Probability of seizure freedom: A box plot illustrating the probability of seizure freedom, as predicted by the 3-dimensional model, stratified by the actual surgical outcomes. Mean seizure freedom probability of seizure free group is slightly higher than the non-seizure free group	157
5.22	10 fold performance metric statistics (3 dimensional model): The table presents the aggregated statistics, including the mean and variance, calculated across 10 folds, for four performance metrics: F1 score, Accuracy, Precision, and Recall. The 3 dimensional model demonstrates proficiency in predicting surgical outcomes, achieving a mean F1 score of 86% and a mean accuracy of 79%. Notably, the model exhibits strong generalization capabilities, as evidenced by minimal variance in performance across folds	158
5.23	Misclassification statistics (3 dimensional model): The 3 dimensional model makes more mistakes in the minority class as shown by the 32 false positives	158

5.24	10 fold performance metric and misclassification statistics comparison (1dim vs 3dim): The 3 dimensional model (highlighted in yellow) outperforms the current 1 dimensional model utilized by doctors (highlighted in green), boasting a significant increase in accuracy by 12%, an improvement in F1 score by 7%, and a reduction of 12 misclassifications	159
5.25	2D visualization of the feature vector for 159 patients: resSOZ ratio is plotted as a function of resection ratio for all 159 patients. The patients are colored based on ground truth surgical outcome: green circles represent seizure free patients and red crosses represent non-seizure free patients. There is a subset of seizure free patients, highlighted with blue circles, that has a low resection ratio with some predicted SOZ channels in the preserved region. These patients are incorrectly predicted as non-seizure free by the 3 dimensional model	161
5.26	Flow analysis pipeline: We partition the set of nodes in the interictal synchronization network into 3 non-overlapping set based on the interictal SOZ classifier predictions and resection information. The flow between these 3 regions are computed by summing the weights of the undirected edges connecting the nodes in these regions	163
5.27	Undirected flow dynamics in a seizure free patient: The flows between the three regions (f_1, f_2, f_3) is plotted as a function of time where each time index corresponds to one interictal synchronization network. On average, there is minimal flow between preserved SOZ and preserved Non-SOZ given the resected SOZ cut	164
5.28	Undirected flow dynamics in a non-seizure free patient: The flows between the three regions (f_1, f_2, f_3) is plotted as a function of time where each time index corresponds to one interictal synchronization network. On average, there is moderately strong flow between preserved SOZ and preserved Non-SOZ given the resected SOZ cut	165

5.29	Probability of seizure freedom (5 dimensional): A box plot illustrating the probability of seizure freedom, as predicted by the 5 dimensional model, stratified by the actual surgical outcomes. Mean seizure freedom probability of seizure free group (0.82) is higher than the non-seizure free group (0.48)	166
5.30	10 fold performance metric and misclassification statistics comparison (3dim vs 5dim): 5 dimensional model (highlighted in pink) performs better than the 3 dimensional model (highlighted in maroon) with a boost in accuracy of 5%, boost in F1 score of 3%, and 14 fewer misclassifications	167
5.31	10 fold performance metric and misclassification statistics comparison (1dim vs 5dim): The 5 dimensional model (highlighted in pink) outperforms the current 1 dimensional model employed by doctors (highlighted in green), demonstrating a remarkable increase in accuracy by 17%, an improvement in F1 score by 10%, and a reduction of 26 misclassifications	168
5.32	2D visualization of the feature vector for 159 patients: resSOZ ratio is plotted as a function of resection ratio for all 159 patients. The patients are colored based on ground truth surgical outcome: green circles represent seizure free patients and red crosses represent non-seizure free patients. Majority of the false positives, boxed in blue, are situated in the far right corner of the plot with a very high resection ratio and all the predicted SOZ channels removed	169

5.33	'Mini Seizures' in interictal period: The Fiedler eigenvalue is plotted over time during the interictal period on the left side of the plot, while the raw waveform amplitude of four channels over the same time period is depicted on the right. Notably, time segments characterized by hypersynchronous activity (sharp rise in fiedler value) are boxed in black. Observing the amplitude plot, it becomes evident that the channels exhibit excessive synchronization in the time domain during these hypersynchronous states, bearing resemblance to that observed during seizures. As a hypothesis, we speculate that these periods of hypersynchronous activity in the interictal period may correspond to "Mini Seizures"	170
5.34	Correlation between interictal network properties and channel level pathological events: The spk-HFO rate of a channel is graphed over time during the interictal period on the left side of the plot, while the eigenvector centrality of the same channel over the same time period is illustrated on the right. Upon examination of both plots, it becomes apparent that the spk-HFO rate of a channel appears to be correlated with eigenvector centrality of that channel	171
5.35	Probability of seizure freedom (7 dimensional): A box plot illustrating the probability of seizure freedom, as predicted by the 7 dimensional model, stratified by the actual surgical outcomes. Mean seizure freedom probability of seizure free group (0.9) is significantly higher than the non-seizure free group (0.27)	174
5.36	10 fold performance metric and misclassification statistics comparison (1dim vs 3dim vs 5dim vs 7dim): The 7 dimensional model (highlighted in blue) significantly outperforms the current 1 dimensional model utilized by doctors (highlighted in green), showcasing a substantial increase in accuracy by 30%, a remarkable improvement in F1 score by 20%, and a reduction of 50 misclassifications	175

5.37	Pattern in the correlation feature space: The ratio of variance to mean $\frac{Var[R_c^2]}{E[R_c^2]}$ is plotted against the mean $E[R_c^2]$ for all 159 patients. Patients are color-coded based on their ground truth surgical outcome: green circles depict seizure free patients, while red crosses represent non-seizure free patients. A distinct separation between the two patient groups is evident in this two-dimensional feature space	176
5.38	k-nearest neighbor clustering in the correlation feature space: In the 2-dimensional correlation feature space, we conducted a k -nearest neighbor clustering with $k = 7$. The clustering algorithm was trained and validated using 10-fold stratified cross-validation, resulting in an impressive mean $F1$ score of 93%.	177
5.39	Four clusters in the correlation feature space: In the 2-dimensional correlation feature space, four distinct clusters emerge. Among them, Cluster 2 and Cluster 4 are primarily composed of seizure free patients, whereas Cluster 1 and Cluster 3 predominantly consist of non-seizure free patients	179
5.40	Box plot of channel level R^2 values of a patient from Cluster 1: For a patient in cluster 1, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. A significant number of channels with spk-HFO detections, have a low R^2 value, suggesting that a sizeable number of detected spk-HFOs remain unexplained by the sampled interictal network	181
5.41	Box plot of channel level R^2 values of a patient from Cluster 4: For a patient in cluster 4, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. Majority of the channels with spk-HFO detections, have a high R^2 value, suggesting that most of the spk-HFOs detected are explained by the sampled interictal network	182

- 5.42 **Box plot of channel level R^2 values of a patient from Cluster 2:** For a patient in cluster 2, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. The majority of channels exhibit zero spk-HFO detections, resulting in zero R^2 values. Among channels with a few spk-HFO detections, their R^2 values are high, suggesting that the spk-HFO rates of these channels are explained by the sampled interictal network property 184
- 5.43 **Box plot of channel level R^2 values of a patient from Cluster 3:** For a patient in cluster 3, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. The majority of channels exhibit zero spk-HFO detections, resulting in zero R^2 values. Among channels with a few spk-HFO detections, their R^2 values are low, suggesting that the spk-HFO rates of these channels remain unexplained by the sampled interictal network property 186

LIST OF TABLES

- 2.1 **Equal-time constrained NMF reconstruction errors (Real datasets):** We tabulated the reconstruction error statistics for 5 runs of each algorithm. From the table, it is evident that eNMF achieves the lowest mean reconstruction error with least variance for all latent dimensions and across both the datasets. **Note:** We didn't tabulate the error statistics for Mult because it has a very high reconstruction error. 57
- 2.2 **Equivalence of Factor Matrices (scaled and rotated versions of each other) for Synthetic datasets** (see Section VII(A) for the construction of such data): We used the algorithm in Appendix B to compute the generalized orthogonal parameters, R_1, R_2 and Λ . We verified that $R_1 \approx I$ by computing $\|R_1 - I\|_F$ for all 4 SNR levels ($\|R_1 - I\|_F \leq 10^{-6}$). We then computed the residual errors (δ_u, δ_v) , where $\delta_u = \|U_{alg} - U_{eNMF} \Lambda R_2^T\|_F$ and $\delta_v = \|V_{alg} - V_{eNMF} \Lambda^{-1} R_2^T\|_F$, and found that $(\delta_u \approx 0, \delta_v \approx 0)$. Thus, the NMF factors obtained from the AO-ADMM algorithm are scaled and rotated versions of those obtained by the eNMF algorithm. Thus, *the feature vectors retain the same pairwise cosine distances, and are equivalent from a clustering perspective.* 64
- 2.4 **eNMF satisfies the KKT conditions for almost all latent dimensions and across both the datasets:** We tabulated the tuple (δ_W, σ_W) , where $\delta_W = \max(\nabla W \odot W)$ and $\sigma_W = \sum_{i,j} (\nabla W \odot W)_{ij}$. As before W is defined as the stacked NMF factors i.e. $W = \begin{bmatrix} U \\ V \end{bmatrix}$. eNMF satisfies the KKT conditions $(\delta_W \approx 0, \sigma_W \approx 0)$ for all datasets and across all latent dimensions. 65

2.3 Numerical Exploration of Permutation Equivalence of Factor matrices: We tabulated the quartet (u_B, v_B, e_d, e_m) . From the top table for the exact factorization scenario, we can see that factor matrices are permutations of each other at the global minimum (for sparsity 0.3, 0.4, 0.5 where global minimum is obtained by several algorithms). From the bottom table for the real dataset scenarios, it can be observed that the factors from the competing algorithms are approaching to a permuted version of the eNMF factors ($u_B \approx 100, v_B \approx 100$) in an asymptotic manner due to the slow convergence of the competing algorithms. As the numerical results show, the factor matrices can be very different even when the algorithms are very close to a local minimum. Thus while reconstruction errors might be similar, the factor matrices might be very different; which might affect performance in downstream tasks. 71

ACKNOWLEDGMENTS

Chapter 3 is reprint of publication at Clinical Neurophysiology (PI: Vwani P. Roychowdhury) that is available under CC-BY 4.0 license (<https://www.sciencedirect.com/journal/clinical-neurophysiology>)

Chapter 4 is reprint of publication at Journal of Neural Engineering (PI: Vwani P. Roychowdhury) that is available under CC-BY 4.0 license (<https://iopscience.iop.org/journal/1741-2552>)

Citations below:

- Tonmoy Monsoor, Yipeng Zhang, Atsuro Daida, Shingo Oana, Qiuqing Lu, Shaun A. Hussain, Aria Fallah, Hiroki Nariai, Vwani Roychowdhury et al. "Optimizing detection and deep learning-based classification of pathological high-frequency oscillations in epilepsy." Clinical Neurophysiology 154 (2023): 129-140.
- Yipeng Zhang, Hoyoung Chung, Jacqueline P. Ngo, Tonmoy Monsoor, Shaun A. Hussain, Joyce H. Matsumoto, Patricia D. Walshaw, Hiroki Nariai, Vwani Roychowdhury et al. "Characterizing physiological high-frequency oscillations using deep learning." Journal of neural engineering 19, no. 6 (2022): 066027.

VITA

- 2010–2014 B.S. (Electrical Engineering and Computer Science), University of Michigan, Ann Arbor
- 2015–2017 M.S. (Electrical and Computer Engineering), University of California, Los Angeles
- 2017–present Ph.D. Candidate (Electrical and Computer Engineering), University of California, Los Angeles

PUBLICATIONS

- [1] **Tonmoy Monsoor**, Yipeng Zhang, Atsuro Daida, Shingo Oana, Qiuqing Lu, Shaun A. Hussain, Aria Fallah, Hiroki Nariai, Vwani Roychowdhury et al. "Optimizing detection and deep learning-based classification of pathological high-frequency oscillations in epilepsy." *Clinical Neurophysiology* 154 (2023): 129-140.
- [2] Yipeng Zhang, Qiuqing Lu, **Tonmoy Monsoor**, Shaun A. Hussain, Joe X. Qiao, Noriko Salamon, Aria Fallah, Hiroki Nariai, Vwani Roychowdhury et al. "Refining epileptogenic high-frequency oscillations using deep learning: a reverse engineering approach." *Brain communications* 4, no. 1 (2022): fcab267.
- [3] Yipeng Zhang, Hoyoung Chung, Jacqueline P. Ngo, **Tonmoy Monsoor**, Shaun A. Hussain, Joyce H. Matsumoto, Patricia D. Walshaw, Hiroki Nariai, Vwani Roychowdhury et al. "Characterizing physiological high-frequency oscillations using deep learning." *Journal of neural engineering* 19, no. 6 (2022): 066027.

- [4] Yipeng Zhang, Lawrence Liu, Yuanyi Ding, Xin Chen, **Tonmoy Monsoor**, Atsuro Daida, Shingo Oana, Hiroki Nariai, Vwani Roychowdhury et al. "PyHFO: lightweight deep learning-powered end-to-end high-frequency oscillations analysis application." *Journal of Neural Engineering* (2023).
- [5] Yipeng Zhang, Zahra M. Aghajan, Matias Ison, Qiujing Lu, Hanlin Tang, Guldamlal Kalender, **Tonmoy Monsoor**, Vwani Roychowdhury et al. "Decoding of human identity by computer vision and neuronal vision." *Scientific reports* 13, no. 1 (2023): 651.
- [6] Linda Comstock, Jianhong He, Pavan Holur, **Tonmoy Monsoor**, Vwani Roychowdhury, and William Speier. "Parallels between auditory and neural processing in TMS-facilitated speech decoding." In *The Night Whites Language Workshop: The Fifth Saint Petersburg Winter Workshop on Experimental Studies of Speech and Language (Night Whites 2019)*, pp. 47-47. 2020.

CHAPTER 1

Introduction

1.1 Overview

Recent advancements in Artificial Intelligence (AI) have led to significant progress in learning efficient models of dynamical systems, revolutionizing our understanding of complex temporal processes. Machine learning techniques, particularly deep learning, have been instrumental in capturing and modeling the intricate dynamics of systems ranging from physics to biology. Deep neural networks (DNN), recurrent neural networks (RNNs), convolutional neural networks (CNNs), and graph neural networks (GNN) have been adapted to infer and predict the behavior of dynamical systems from observational data. These AI-driven approaches enable the extraction of meaningful patterns and dependencies within temporal sequences, facilitating tasks such as time-series prediction, system identification, and control.

Brain is an unparalleled complex system displaying multi-scale memory dynamics ranging from tens of milliseconds to years, with the simultaneous capability to learn new behavior patterns even in single-shot settings. Its complexity arises from the dynamic interplay of billions of neurons, each communicating with one another through electrical and chemical signals. This intricate network gives rise to nonlinear and time-varying behaviors, and its neural activity is finely tuned, relying on the precise coordination of various brain regions to function optimally. However, this delicate balance can be disrupted by neurological disorders, leading to significant alterations in collective brain dynamics.

Neurological disorders often disrupt the connectivity between different brain regions, lead-

ing to aberrant communication patterns. In conditions such as Alzheimer’s disease (AD), for example, there is a progressive loss of connectivity within the default mode network [Grieder et al., 2018], which is crucial for tasks such as memory consolidation and self-referential processing. Similarly, in disorders like autism spectrum disorder (ASD), atypical connectivity patterns between regions involved in social cognition can impair social interaction and communication skills [Abbott et al., 2016]. These disruptions in network connectivity alter the collective dynamics of the brain, impacting its ability to integrate information and perform cognitive functions effectively.

Another aspect of collective brain dynamics pertains to the synchronization of neural oscillations or brain rhythms. These rhythmic activities play a fundamental role in coordinating information processing across different brain regions. However, neurological disorders can lead to abnormalities in brain rhythms [Videnovic and Zee, 2015], affecting synchronization and coherence. For instance, individuals with epilepsy often exhibit abnormal electrical activity characterized by epileptic seizures, disrupting normal brain rhythms and causing transient cognitive impairments. Similarly, in Parkinson’s disease, there is a disruption of beta oscillations in the basal ganglia-thalamocortical circuit [McCarthy et al., 2011], contributing to motor dysfunction. These alterations in brain rhythms not only affect specific functions but also interfere with the overall coordination of neural activity, leading to cognitive and behavioral disturbances.

Collective brain dynamics rely on the efficient processing and integration of information across distributed brain networks. Neurological disorders can impair this process by disrupting neuronal signaling, synaptic transmission, and neurotransmitter systems. In conditions such as schizophrenia, for example, there is evidence of impaired glutamatergic and dopaminergic neurotransmission [Egerton et al., 2020], leading to disturbances in information processing and perception. Similarly, in attention deficit hyperactivity disorder (ADHD), alterations in the dopaminergic system disrupt the regulation of attention and executive functions. These disruptions in information processing undermine the coherence of collective

brain dynamics, manifesting as deficits in perception, cognition, and behavior.

In this thesis, we learn the characteristic neuronal dynamics of one of the most prevalent and debilitating disorders called Epilepsy. It is a neurological disorder where the normal signaling mechanism in the brain is disrupted by sudden and synchronized bursts of electrical pulses, leading to recurrent episodes of seizures. About one-fifth of patients with epileptic seizures cannot be treated with anti-seizure medications and need to undergo a resective surgery to remove the brain regions responsible for generating seizures, known as the epileptogenic zone (EZ). Such surgeries, however, have a success rate of only 60 – 70%, and the need to accurately identify EZ is a pressing problem. As part of a collaboration between Roychowdhury research group and the Pediatric neurology group at UCLA Geffen School of Medicine, we received intracranial EEG (iEEG) recordings of 159 patients with focal drug resistant epilepsy. We use methods from complex network dynamics and machine learning to model iEEG recordings as the output of a networked dynamical system. We extract a mix of global and local topological properties of the time-varying network models to discover novel spatio-temporal biomarkers of EZ from short interictal iEEG recordings; that is, our biomarkers are derived from segments as short as 5 minutes long, sampled from iEEG recordings when the subject has no seizures. In contrast, current practices rely on monitoring patients for two weeks or more, and collecting enough infrequent ictal segments (i.e., iEEG recordings during seizure episodes) to estimate the EZ based on channels/nodes responsible for driving the episodes. One of the primary findings of our approach is that even during the interictal period, the EZ channels are not passive: they still are hubs driving frequent hypersynchronous activity through the network, and thus enabling us to determine EZ accurately even in the absence of any visible seizure episodes. We hypothesize that these hyper-synchronous interictal states are 'mini seizures' that have been anticipated in the literature, and to the best of our knowledge, our method is one of the first to identify and mine them. These novel biomarkers hold great promise to i) accurately identify the EZ and therefore guide neurosurgeons' decisions, and lead to much better postoperative outcomes,

ii) significantly reduce hospitalization time and associated medical risks, and iii) lead to a better understanding of the biochemical origins of seizures and therefore aid in finding effective treatments, including targeted neural stimulation, precise resection and designing drugs to disrupt the hyper synchronous network events.

1.2 System level deficits in neurological disorders

Neurological disorders (ND) often entail system-level deficits, affecting various interconnected networks and functions within the nervous system. From cognitive impairments to motor dysfunction, these deficits manifest across multiple levels, presenting unique challenges in diagnosis, treatment, and management. These system level deficits can trigger a vicious feedback loop causing bio-molecular changes causing further disruptions in neuronal network connectivity. Therefore, key to treating neurological disorders is early detection and intervention which requires the understanding of the collective neuronal dynamics that underlie such disorders.

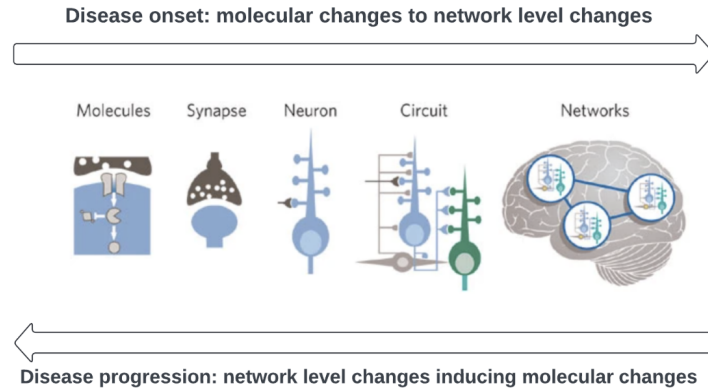


Figure 1.1: **Vicious cycle of ND [Linazasoro, 2009]:** Neurological disorders (ND) which often starts from molecular level changes can lead to neuronal network dysfunction. These network level disruptions can trigger a feedback loop where they can lead to molecular level changes causing further disruptions in neuronal network connectivity

1.2.1 Alzheimer’s Disease (AD)

Alzheimer’s disease (AD), the most common form of dementia, epitomizes the profound disruption of neural networks seen in neurological disorders. Primarily characterized by the accumulation of beta-amyloid plaques and tau protein tangles, AD instigates widespread neuronal death, particularly in regions crucial for memory and cognition, such as the hippocampus and cerebral cortex. However, the repercussions extend far beyond these localized areas.

System-level deficits in AD encompass a spectrum of cognitive impairments, including memory loss, executive dysfunction, and language deficits. These deficits arise from disruptions in large-scale brain networks involved in memory encoding, retrieval, and consolidation, such as the default mode network (DMN) and the salience network. Consequently, individuals with AD experience difficulties in integrating information across brain regions, leading to fragmented cognition and impaired decision-making.

Moreover, AD engenders alterations in neurotransmitter systems, particularly acetylcholine and glutamate, further exacerbating cognitive decline. Dysfunction in cholinergic projections from the basal forebrain disrupts attentional processes, while glutamatergic dysregulation impairs synaptic plasticity and learning mechanisms. Thus, the system-level deficits in AD underscore the intricate interplay between molecular pathology, neural network dysfunction, and cognitive decline.

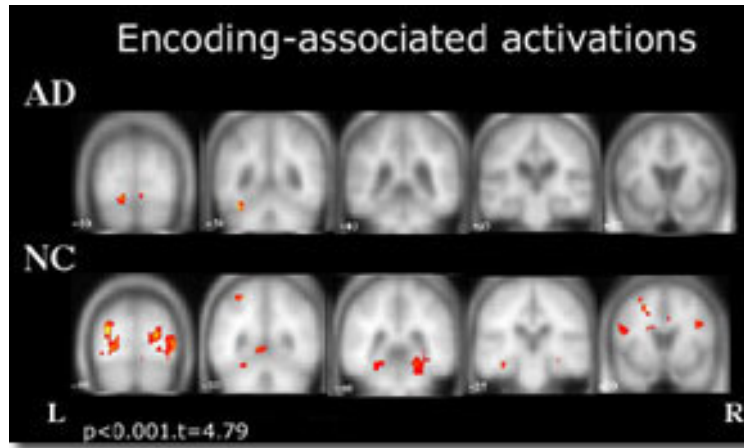


Figure 1.2: **Memory encoding deficit in Alzheimers [Golby et al., 2005]:** Alzheimer’s disease (AD), the most common cause of dementia in older adults leads to a deficit in declarative memory with relative sparing of implicit memory. This study uses fMRI to study the functional competency of specific brain regions and the relationship to the behavioral memory deficits in AD. The study involved five early-stage AD patients and five healthy age-matched neurologically normal control subjects (NC) with fMRI at 3 Tesla during intentional memory encoding of scenes.

1.2.2 Parkinson’s Disease (PD)

Parkinson’s disease (PD), characterized by dopaminergic neurodegeneration in the substantia nigra pars compacta, precipitates a myriad of motor and non-motor symptoms, indicative of widespread neural circuitry disruption. While the basal ganglia-cortical motor circuitry is prominently affected, PD’s impact transcends this motor-centric view, implicating broader network dysfunction.

At the system level, PD entails disruptions in cortico-striato-thalamo-cortical circuits, compromising not only motor control but also cognitive and affective processes. Dopaminergic depletion in the basal ganglia perturbs the delicate balance between direct and indirect pathways, resulting in bradykinesia, rigidity, and tremor. Concurrently, alterations in meso-cortical and mesolimbic dopaminergic projections engender cognitive impairments, such as

executive dysfunction and visuospatial deficits.

Additionally, PD's non-motor symptoms, including depression, anxiety, and sleep disturbances, underscore the involvement of broader neural networks beyond the motor system. Dysfunction in monoaminergic systems, particularly serotonin and noradrenaline, contributes to affective disturbances, while alterations in the hypothalamic-pituitary-adrenal (HPA) axis disrupt circadian rhythms and sleep-wake cycles. Thus, PD exemplifies the systemic nature of neurological disorders, intertwining motor, cognitive, and affective domains through multifaceted network dysfunction.

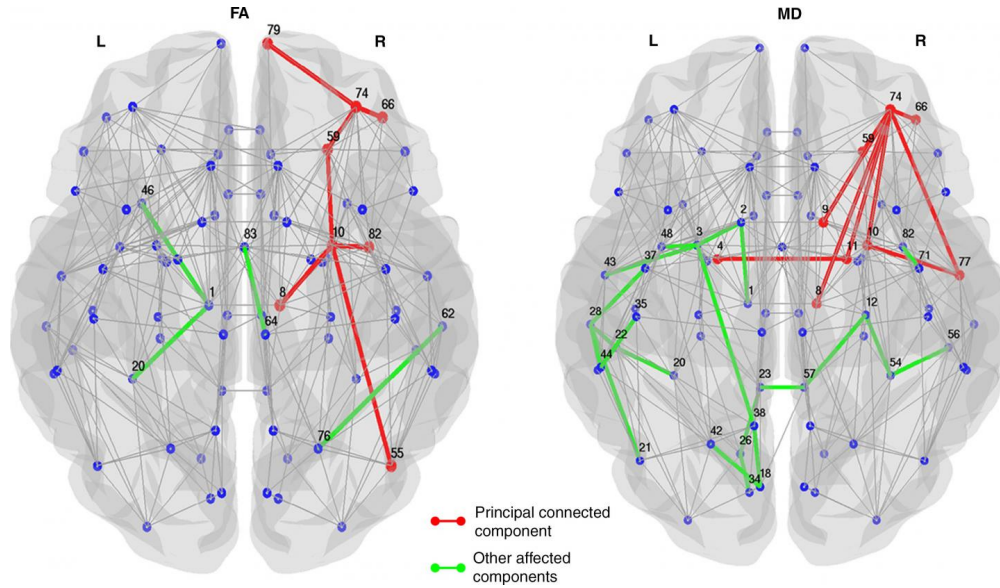


Figure 1.3: **Structural connectivity in PD [Galantucci et al., 2017]:** Affected structural connections in patients with PD and MCI relative to healthy control subjects and patients with Parkinson disease (PD) without mild cognitive impairment (MCI). The principal connected component is shown in red, and the other affected connections not included in the principal connected component are shown in green. Subnetworks show altered structural connectivity (decreased fractional anisotropy or increased mean diffusivity) in patients with PD and MCI relative to healthy control subjects at $P < .01$

1.2.3 Multiple Sclerosis (MS)

Multiple sclerosis (MS), characterized by autoimmune-mediated demyelination and neuroinflammation, epitomizes the heterogeneity and complexity of system-level deficits in neurological disorders. While traditionally viewed as a white matter disease, MS's impact extends to gray matter regions and broader neural networks, precipitating diverse clinical manifestations.

At the system level, MS engenders disruptions in sensorimotor, visual, and cognitive networks, reflective of multifocal demyelination and axonal damage throughout the central nervous system. Sensorimotor deficits, including weakness, spasticity, and ataxia, arise from lesions in corticospinal tracts and cerebellar pathways, compromising motor coordination and control.

Moreover, MS's impact on cognitive function underscores the systemic nature of the disease, transcending focal lesions to affect large-scale brain networks. Demyelination in fronto-subcortical circuits impairs attention, processing speed, and working memory, while alterations in default mode and salience networks contribute to deficits in social cognition and emotional processing. Furthermore, MS's influence on visual pathways, manifested as optic neuritis and visual field defects, underscores the widespread nature of neural network disruption in the disease.

1.2.4 Epilepsy

Epilepsy is a chronic neurological disorder characterized by recurrent seizures resulting from abnormal electrical activity in the brain. While seizures are the hallmark feature of epilepsy, the disorder's impact extends beyond episodic events, implicating broader neural networks and cognitive functions.

At the system level, epilepsy entails disruptions in neural circuits involved in seizure generation, propagation, and termination. Abnormal synchronization of neuronal firing within

cortical and subcortical structures precipitates seizures, disrupting normal brain function and behavior. These disruptions can manifest as alterations in consciousness, motor control, sensation, and cognition, depending on the seizure’s onset and propagation patterns.

Moreover, epilepsy’s systemic effects encompass cognitive impairments and psychiatric comorbidities, reflecting disturbances in large-scale brain networks beyond seizure activity. Chronic epilepsy is associated with deficits in attention, memory, and executive function, stemming from structural and functional alterations in frontal-temporal-parietal networks. Additionally, psychiatric disorders such as depression, anxiety, and psychosis are prevalent in individuals with epilepsy, implicating disruptions in limbic circuits and neurotransmitter systems.

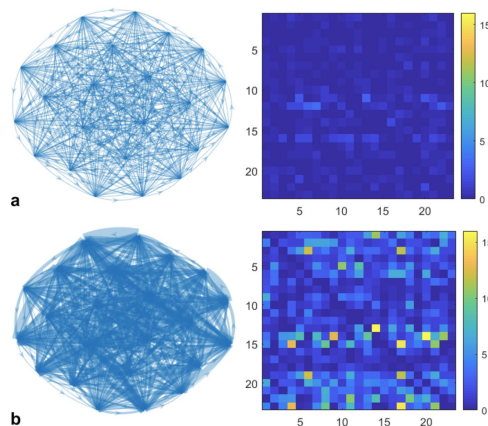


Figure 1.4: **EEG node Connectivity in Epilepsy [Bomela et al., 2020]**: Networks (graph representation and the heat-map representation of the corresponding adjacency matrix) of the normal state (a) and seizure state (b)

1.3 Neural recording techniques

Neurological disorders often arise from aberrant interactions within and between brain regions, leading to disruptions in neural networks. By recording from a large neuronal population, researchers can gain insights into the dynamics of these networks, uncovering patterns

of activity that underlie disease pathology. Whether it's the synchronization of neural oscillations in epilepsy, the dysconnectivity of brain regions in autism spectrum disorder, or the aberrant functional connectivity in Alzheimer's disease, examining the collective activity of neuronal populations provides critical information about network dysfunction and its implications for brain function.

Invasive neural recording techniques involve direct access to the brain tissue, providing detailed insights into neuronal activity at the cellular level. Microelectrode arrays (MEAs) enable researchers to record extracellular action potentials and local field potentials from specific brain regions, facilitating the study of neural circuitry and network dynamics in animal models and human patients undergoing neurosurgery. In conditions such as epilepsy, intracranial EEG (iEEG) monitoring with implanted electrodes helps localize seizure foci and guide surgical interventions for seizure control. Similarly, in movement disorders like Parkinson's disease, deep brain stimulation electrodes not only alleviate symptoms but also provide valuable recordings of neuronal activity, enhancing our understanding of basal ganglia dysfunction.

Advances in recording and implantation of iEEG provide multiple options in the 21st century. This not only includes the choice between subdural electrodes (SDE) and stereo-electroencephalography (SEEG) but also includes the implantation and recordings from microelectrodes. Recordings from SDE and SEEG have both their advantages and disadvantages. Stereo-EEG seems to have a lower rate of complications that are clinically significant, but has limitations in spatial sampling of the cortical surface. Stereo-EEG can sample deeper areas of the brain including deep sulci and hard to reach areas such as the insula.

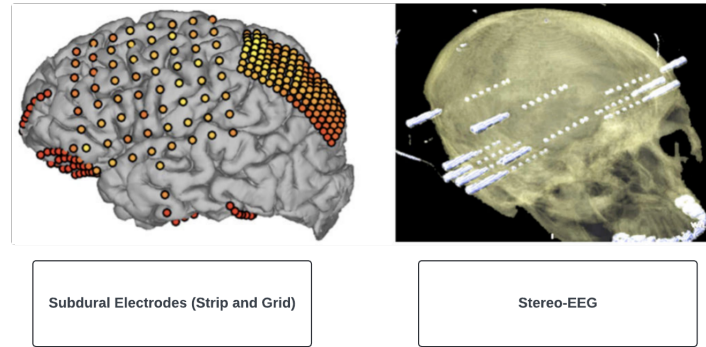


Figure 1.5: **Two Methods of Intracranial EEG: Subdural Electrodes(SDE) and Stereo-EEG (SEEG)** While grids and strips of subdural electrodes provide a large coverage over the bare surface of the cerebral cortex, they are often implanted in one hemisphere and do not reach deeper brain structures (e.g., hippocampus or insula). By comparison, depth electrodes can enable bilateral monitoring of superficial and deep cortical structures but only the most superficial and deep contacts will be within the cortical gray matter while the rest of the contacts are placed in the white matter.

Non-invasive neural recording techniques offer safer and more accessible means of studying brain activity in humans, enabling large-scale investigations and longitudinal studies. Electroencephalography (EEG) measures electrical activity from the scalp, providing insights into brain rhythms, connectivity, and event-related potentials. In epilepsy, EEG serves as a cornerstone for diagnosing and monitoring seizure activity, guiding treatment decisions and evaluating treatment efficacy. Functional magnetic resonance imaging (fMRI) indirectly measures changes in blood flow and oxygenation associated with neural activity, allowing researchers to map brain function and connectivity in various neurological disorders, including Alzheimer’s disease, multiple sclerosis, and stroke.

Advancements in technology have led to the development of novel neural recording techniques with enhanced capabilities and applications. Optogenetics enables precise control of neuronal activity using light-sensitive proteins, offering insights into neural circuit function

and potential targets for therapeutic intervention. Calcium imaging techniques allow researchers to visualize and monitor neuronal activity in real-time, providing valuable insights into cellular dynamics and network activity patterns. These emerging technologies hold promise for unraveling the complexities of neurological disorders and developing targeted therapies tailored to individual patients' neural circuitry.

1.4 Epilepsy: A prevalent and debilitating neurological disorder

In this thesis, we learn the characteristic neuronal dynamics of one of the most prevalent and debilitating disorder called Epilepsy. It is the third most common neurological disorder with a worldwide prevalence of about 1% and is characterized by recurrent episodes of seizures. Epileptic seizures arise from a disturbance in the balance of excitatory and inhibitory neuronal activity in the brain. This imbalance can result from various factors, including genetic mutations, structural brain abnormalities, neurotransmitter imbalances, or alterations in ion channel function. Seizure initiation typically involves the synchronous and excessive firing of neurons, leading to a cascade of abnormal electrical activity that spreads throughout the brain. The precise mechanisms underlying seizure generation and propagation vary among different types of epilepsy and can involve complex interactions within and between brain networks.

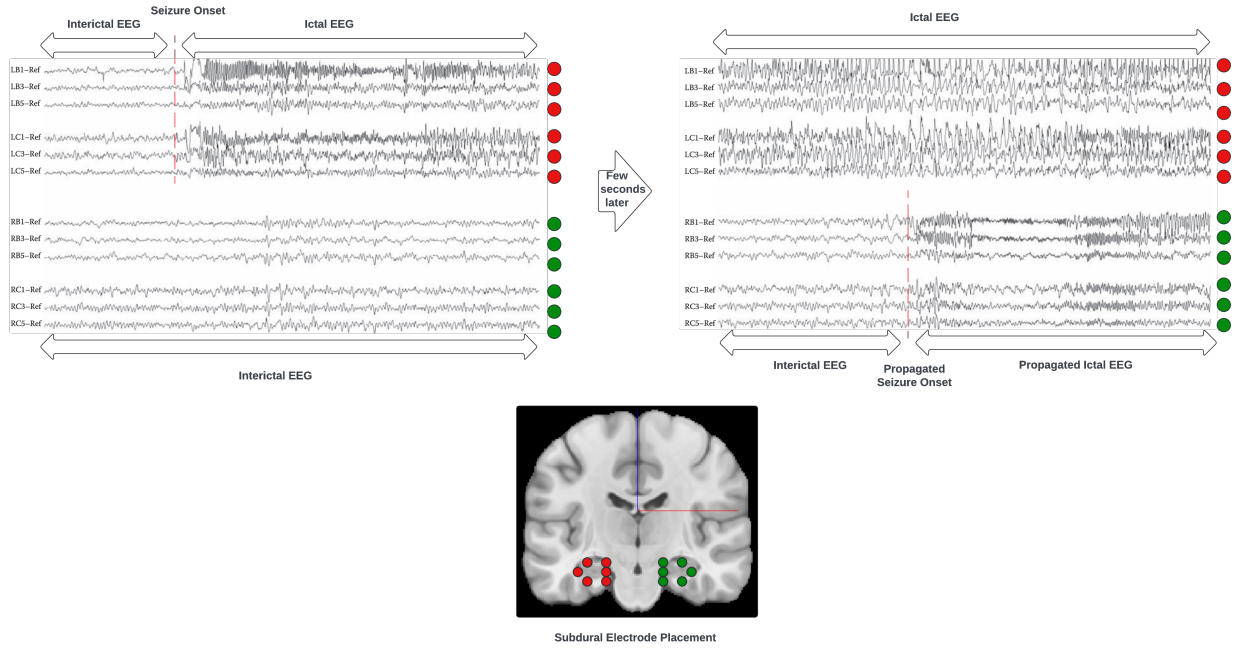


Figure 1.6: **Seizure generation and propagation [Osorio et al., 2016]**: Intracranial EEG recording obtained from bilateral orthogonally implanted eight-contact depth electrodes; only the signal from the first (deepest), third, and fifth contacts of each depth electrode are shown, with LB1/RB1 situated in the hippocampi and LC1/RC1 in the parahippocampal gyri. Left mesial temporal regional seizure onset in hippocampus and parahippocampal gyrus ($LB1 > 3 > 5$ and $LC1 > 3 > 5$), with subsequent propagation (“spread”) to contralateral right mesial temporal region. Note that the contralateral propagation can also be considered as the abrupt onset of an independent seizure in the right temporal lobe because the ictal activity might not be synchronous between the two hemispheres. Seizure prediction signal analysis techniques need to uncover predictive features from the type of signal seen in the left-sided electrodes before the seizure onset (interictal segment) and in the right-sided electrodes before the seizure onset (interictal segment) to differentiate between independent and propagated seizures

1.4.1 Types of epileptic seizures and clinical manifestations

Epileptic seizures manifest in various forms, with focal and generalized seizures representing two primary classifications. Distinguishing between these types is crucial for accurate diagnosis, treatment planning, and providing tailored support to individuals with epilepsy.

1.4.1.1 Focal seizures

Focal seizures originate in a specific area of the brain and may exhibit diverse clinical manifestations, depending on the region affected. These seizures are further categorized into two subtypes:

1. **Simple Focal Seizures:** Characterized by preserved consciousness, simple focal seizures typically involve abnormal sensations or movements localized to a specific body part. Examples include tingling sensations, jerking of limbs, or distortions in sensory perception like changes in taste or smell. While individuals remain aware during these seizures, they may feel alarmed or confused by the unusual sensations
2. **Complex Focal Seizures:** Complex focal seizures are associated with alterations in consciousness and complex, automatic behaviors. During these seizures, individuals may exhibit repetitive, purposeless movements such as lip smacking, picking at clothing, or wandering aimlessly. Despite the outward appearance of purposeful activity, individuals are unaware of their actions and may experience amnesia or confusion afterward

1.4.1.2 Generalized seizures

Generalized seizures involve abnormal electrical activity that affects both hemispheres of the brain simultaneously, resulting in widespread effects on consciousness and motor function. This category encompasses several seizure types, each with distinct clinical features:

1. **Tonic-Clonic (Grand Mal) Seizures:** Among the most dramatic seizure types, tonic-clonic seizures consist of two phases. The tonic phase is characterized by sudden loss of consciousness, stiffening of muscles, and a cry caused by forced expiration. This is followed by the clonic phase, marked by rhythmic, jerking movements of the limbs. Tonic-clonic seizures may result in injuries due to falls, loss of bladder or bowel control, or tongue biting
2. **Absence Seizures:** Commonly observed in children, absence seizures involve brief, sudden lapses in consciousness, typically lasting seconds to half a minute. During these episodes, individuals exhibit a vacant stare, cessation of ongoing activities, and subtle facial movements like blinking or lip smacking. Despite the abrupt interruption in awareness, absence seizures typically resolve spontaneously, often leaving the individual unaware of the episode
3. **Myoclonic Seizures:** Myoclonic seizures manifest as sudden, brief jerks or twitches of muscles, occurring either in isolation or as clusters. These seizures may affect specific muscle groups or involve generalized movements, often resembling startled reactions. Myoclonic seizures can occur spontaneously or be triggered by factors such as sleep deprivation or stress
4. **Atonic Seizures:** Also known as drop attacks, atonic seizures involve sudden loss of muscle tone, leading to abrupt collapse or falls without warning. These seizures may cause individuals to drop objects, slump to the ground, or experience sudden head drops. Atonic seizures pose a risk of injury due to falls and are often associated with a brief period of confusion following the episode

Focal and generalized seizures represent distinct manifestations of epileptic activity, each with unique clinical features and implications for diagnosis and management. In this thesis, we focus on focal seizures.

1.4.2 Treating focal drug resistant epilepsy

Approximately one third of patients with focal epilepsy continue to experience seizures despite treatment with antiepileptic drugs (AEDs) [Beleza, 2009, Ryvlin et al., 2014], known to suffer from focal drug resistant epilepsy. The gold standard for treating focal drug resistant epilepsy is to resect the epileptogenic zone (EZ) which is the site of the cortex responsible for generating seizures [Ryvlin et al., 2014, Lüders et al., 2006]. Various non-invasive techniques, such as surface video-EEG monitoring, neuropsychological assessments, and brain imaging (MRI, PET, ictal SPECT), are currently utilized to identify the EZ and critical cortical regions, facilitating the development of an optimal and safe surgical approach [Sebastiano et al., 2020].

In spite of advancements in diagnostic imaging and electrophysiological techniques, invasive recordings such as intracranial EEG (iEEG) or stereo-EEG (SEEG) are still necessary in a significant portion of patients (25 – 50%) during the presurgical evaluation [Yuan et al., 2014, Cardinale et al., 2019]. This need is particularly pronounced when there are discrepancies between non-invasive test results and when magnetic resonance imaging (MRI) fails to reveal pertinent information. The considerable rate of seizure control failure (30 – 40%) observed one year post-surgery, which tends to increase with longer postoperative follow-up periods, underscores the ongoing challenge in accurately defining the epileptogenic zone (EZ).

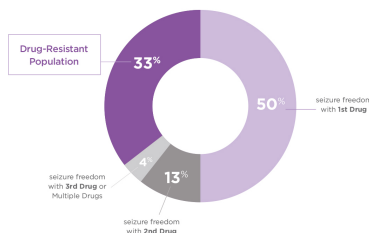


Figure 1.7: **Drug resistant epilepsy:** patients with epilepsy who do not respond successfully to antiseizure medications. Almost a third of patients with epilepsy suffer from drug-resistant epilepsy

The current best practices for EZ identification consists of two main stages:

1. **Invasive recording:** doctors implant electrodes under the skull to record iEEG for around 2 weeks. During the monitoring stage, they try to capture as many seizure episodes as possible where each seizure episode is known as an ictal segment. The recorded time segments for which the patient is not having seizures is known as interictal segment
2. **Analysis of interictal and ictal segments:** use a combination of visual and automated approaches to unearth electrophysiological biomarkers of EZ from ictal and interictal segments

We first review High Frequency Oscillations (HFO), one of the most promising interictal biomarkers of EZ.

1.4.3 High frequency oscillations: a putative interictal biomarker of EZ

High frequency oscillations (HFOs) are short-duration oscillatory field potential with spectral content in the 80-500 Hz frequency range, consisting of at least four oscillations that can be distinguished from background. It was first observed in kainic acid (KA)-induced seizures in rats in late 1990s. Kainic acid, a potent neuroexcitatory substance, is commonly used in experimental models to induce seizures that mimic certain aspects of human epilepsy. One notable study in this regard was "High-frequency oscillations in rat hippocampus during acute epileptic seizures" [Bragin et al., 1999a]. In this pioneering research, the authors investigated the electrographic activity in the hippocampus of rats during KA-induced seizures using iEEG. Bragin and colleagues observed the emergence of distinct high-frequency oscillations (HFOs) in the hippocampus during the acute phase of KA-induced seizures. These HFOs were characterized by their frequency range, which typically exceeded 80 Hz. The study provided evidence that HFOs were closely associated with epileptic activity and occurred concurrently with other seizure-related changes in the EEG, such as spikes and sharp

waves. The findings of this study were instrumental in establishing the presence of HFOs in animal models of epilepsy, particularly in the context of KA-induced seizures in rats. Subsequent research built upon these initial observations, further elucidating the characteristics and significance of HFOs in epileptogenesis, seizure generation, and the identification of epileptogenic zones. Since then, numerous studies have been conducted to explore the mechanistic underpinnings of HFOs, their relationship to epileptogenicity, and their potential utility as biomarkers for epilepsy diagnosis and treatment.

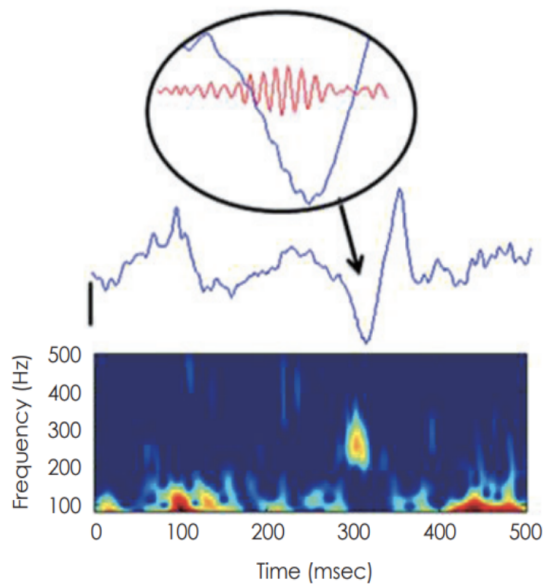


Figure 1.8: **High Frequency Oscillations (HFO)**: Short-duration oscillatory field potential with spectral content in the 80-500 Hz frequency range, consisting of at least four oscillations that can be distinguished from background.

During interictal periods (time segments when the patient is not having seizures), higher rates of HFOs were observed in the seizure onset zone (SOZ) than in other areas [Bragin et al., 1999c, Urrestarazu et al., 2007]. Even though a large proportion of HFOs co-occur with spikes, HFOs can occur also in non-spiking channels or independently from spikes [Jacobs et al., 2008]. The ranking of channels according to rate [Zijlmans et al., 2009] indicated that

HFOs remained confined to the same region during interictal and ictal periods, while spikes presented a wider spread during seizures than interictally [Zijlmans et al., 2011]. Moreover, a postsurgical study showed a correlation between surgical outcome and removal of channels with high HFO rates [Jacobs et al., 2010b]. In summary, interictal HFOs promises to be a potential biomarker of tissue capable of producing seizures i.e epileptogenic zone (EZ).

Visual marking of HFOs provided a good understanding of the relation of HFOs with epilepsy [Zijlmans et al., 2009, Jacobs et al., 2010b, Urrestarazu et al., 2007]. However, visual marking is highly time consuming (it takes about 3 hours to visually mark HFOs in a 10 channel 10 minute recording) and subjectivity is inevitable. Thus, the development of automatic HFO detectors is crucial for the systematic study of HFOs and for their eventual clinical application. The lack of a formal definition makes the detection of HFOs difficult and subjective. Therefore, many groups of researchers have come up with their own rule based methods leading to many automated HFO detectors. In the next section, we review some of the popular automated HFO detectors.

1.4.3.1 Rule based methods for detecting HFOs: automated detectors

Rule based HFO detectors are largely based on the comparison of the signal energy with the EEG epoch that includes the HFO events. The main difference among these detectors is the type of energy function computed on the filtered signal and the post-processing used to prune false detections. The detectors are reviewed in order of the year of original publication.

Staba [Staba et al., 2002] developed a HFO detector based on the energy function defined as the moving average of the root mean square amplitude of a filtered signal. Each channel is first band-pass filtered (80–500 Hz) and then EEG segments with energy above five times the standard deviation (SD) of the mean energy of the entire EEG recording and lasting more than 6 ms are considered as putative HFO events. Putative HFO events less than 10 ms apart are considered as a single HFO event. This detector was developed for microelectrode recordings in humans and rats in temporal regions [Staba et al., 2004a]. A sensitivity of 84%

was reported [Staba et al., 2002]. This detector is referred to as the STE detector.

The HFO detector developed by Gardner [Gardner et al., 2007] is based on the energy function using short-time line length [Esteller et al., 2001]. Each channel is first passed through a spectral equalizer (first order differential filter), then band-pass filtered (80–500 Hz) and the energy threshold is computed as the 95th percentile of the empirical cumulative distribution function for the epoch [Worrell et al., 2008]. A sensitivity of 89.5% was reported. This detector is referred to as the SLL detector.

In the detector developed by Crepon [Crépon et al., 2010] the envelope of the signal was computed with a Hilbert transform. Each channel is first band-pass filtered (80–400 Hz) and events were detected with a threshold five times the SD of the envelope over the whole EEG. A sensitivity of 90.5% was reported. This detector is referred to as the Hilbert detector.

The MNI detector developed by Zelman [Zelmann et al., 2010] is a two-stage algorithm for detecting HFOs. In the first stage, channels with lots of baseline segments are identified and the HFOs in those channels are detected using an algorithm similar to STE detector. In the second stage, the HFOs in channels with few baseline segments are detected using an iterative algorithm. A sensitivity of 96.5% was reported. This detector is referred to as the MNI detector. Although many HFO detectors have been developed, but the two most popular ones to date in terms of usage are STE and MNI.

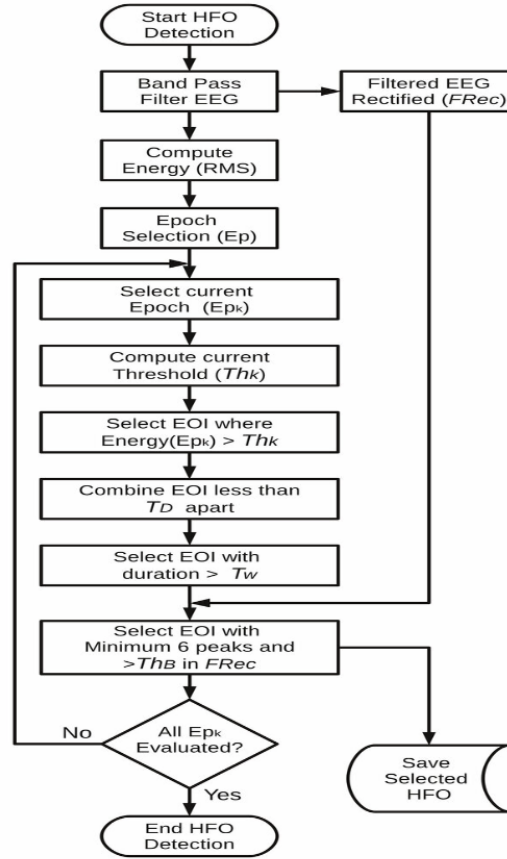


Figure 1.9: **Workflow of STE HFO detector:** Energy function based on root mean squared amplitude is used to compute the energy in short EEG segments. EEG segments with energy above five times the standard deviation (SD) of the mean energy of the entire EEG recording and lasting more than 6 ms were considered as putative HFO events. Putative HFO events are filtered based on some comparison with mean value of rectified signal

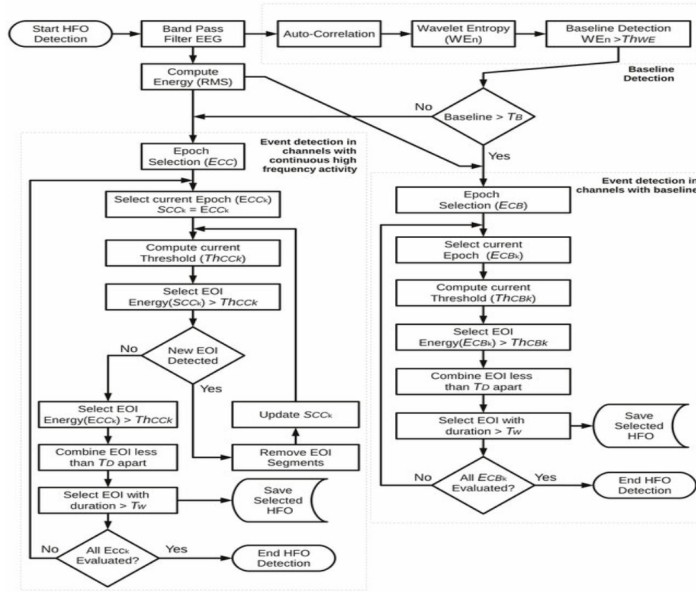


Figure 1.10: **Workflow of MNI HFO detector:** HFO detections consist of two stages. In the first stage, channels with lots of baseline segments are identified and then HFOs are detected in those channels using an algorithm similar to the STE detector. In the second stage, an iterative algorithm is used to detect HFOs in channels with few baseline segments

However, HFOs have also been increasingly recognized for their potential role in cognition, particularly in the context of memory formation and information processing. In the next section, we review the presence of HFOs in various cognitive tasks.

1.4.3.2 Association of HFOs with cognition: physiological HFOs

High frequency oscillations are associated with normal brain function, but also – as mentioned in the preceding discussions – increasingly recognized as potential biomarkers of the epileptogenic brain. Their role in human cognition has been predominantly studied in classical gamma frequencies (30– 100 Hz), which reflect neuronal network coordination involved in attention, learning and memory. High frequency oscillations (HFOs) extend beyond the

limits of gamma band activity and recently have been the focus of animal and human studies of the neurophysiology of cognition (for a review see [Worrell et al., 2012]). In a recent study [Kucewicz et al., 2014], they investigated high frequency oscillations spanning the high gamma (50–125 Hz), ripple (125–250 Hz) and fast ripple (250–500 Hz) frequency bands using intracranial recordings from 12 patients (five males and seven females, age 21–63 years) during memory encoding and recall of a series of affectively charged images. Presentation of the images induced high frequency oscillations in all three studied bands within the primary visual, limbic and higher order cortical regions in a sequence consistent with the visual processing stream. These induced oscillations were detected on individual electrodes localized in the amygdala, hippocampus and specific neocortical areas, revealing discrete oscillations of characteristic frequency, duration and latency from image presentation. Memory encoding and recall significantly modulated the number of induced high gamma, ripple and fast ripple detections in the studied structures, which was greater in the primary sensory areas during the encoding (Wilcoxon rank sum test, $P = 0.002$) and in the higher-order cortical association areas during the recall (Wilcoxon rank sum test, $P = 0.001$) of memorized images. Furthermore, the induced high gamma, ripple and fast ripple responses discriminated the encoded and the affectively charged images. In summary, the results show that high frequency oscillations, spanning a wide range of frequencies, are associated with memory processing and generated along distributed cortical and limbic brain regions.

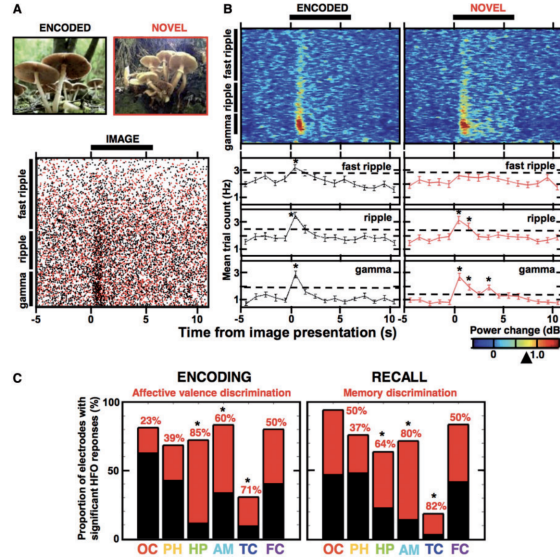


Figure 1.11: **Physiological HFOs [Kucewicz et al., 2014]:** Induced HFOs reveal discriminative responses to the encoded images. (A, top) Examples of two similar pictures one previously seen during encoding (black code) and the other novel (red code); (A, bottom) cumulative scatterplot of all HFO detections in the three frequency bands recorded from an example amygdala electrode in Patient 8 during his memory recall—black dots were detected from trials with encoded images and the red dots from trials with novel images of the same session (black bar indicates the time-course of image presentation). (B, top) Spectrograms summarize trial-averaged power changes in the two trial types from the scatterplot in A, (B, bottom) the mean trial counts for the scatterplot in A are summarized separately for the gamma, ripple and fast ripple detections, aligned to the trial time-course of the spectrograms. Notice significant induction of the HFOs in individual bins as compared to matched average count from the five ‘baseline’ bins, which is not present in the fast ripple response to novel images and was significantly different from the fast ripple response to encoded images. (C) Proportion of all electrodes that showed discriminative HFO responses (as in B) to affectively charged images during encoding (left) and to the encoded images during recall (right) are summarized across the studied structures (the total black bar indicates number of electrodes with significantly induced HFOs and its red part indicates fraction of these electrodes that showed significant differences. Notice that the highest fraction of electrodes with the discriminative HFO responses is reported in the hippocampus, amygdala and the temporal cortex, consistently exceeding 50% (marked with asterisk).

1.4.4 Synchronization network properties: ictal biomarker of EZ

The synchronization effect is prominent in ictal segments and doctors visually mark the iEEG channels driving the ictal segments. However, since the synchronized activity between the channels happen in such a short time scale, it is often very difficult to identify the channels driving seizure activity visually. Therefore, network based approaches have also been developed to identify the drivers of such synchronized activity, known as seizure onset zone (SOZ), in the ictal segment. Once the SOZ has been identified, then it is used to estimate EZ.

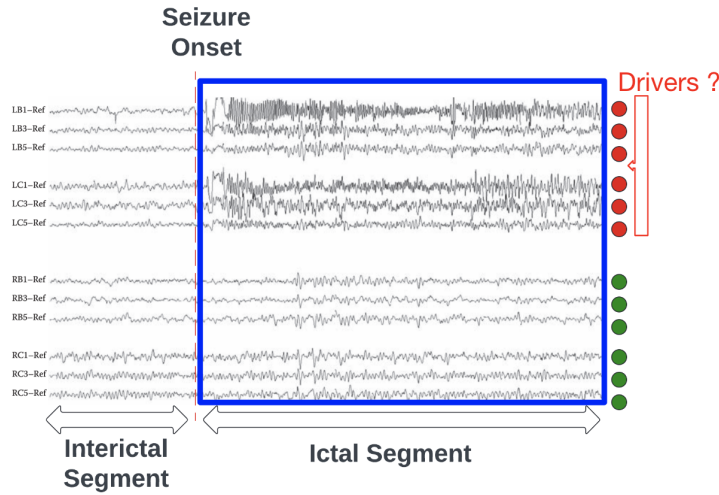


Figure 1.12: **Drivers of seizure episodes (SOZ) identification:** During this ictal segment, boxed in blue, visually all the red electrodes seem to behave in a very synchronized manner, so doctors mark them as the drivers of this seizure episode. However, since the synchronized activity between the channels happen in such a short time scale so often it is very difficult to identify the channels driving these kind of activity visually. In this example, only a subset of the red electrodes might be the drivers of the synchronized activity. Therefore, network based approaches have also been developed to identify the drivers of such synchronized activity in the ictal segment.

1.4.4.1 SOZ identification using ictal synchronization network properties

Graph theory has been widely used to investigate epilepsy. The first indication that network analysis might be helpful in understanding epileptic seizures came from the simulation study of Netoff and others on models of hippocampal networks [Netoff et al., 2004]. After the simulation study, many studies have investigated network alterations in relation to the SOZ using graph theory in patients undergoing surgery for focal drug resistant epilepsy. The procedure to study ictal iEEG using graph theoretic approach usually includes the following steps (i) construct a binary or weighted adjacency matrix of undirected or directed graph from ictal iEEG data (ii) apply a threshold to each element of the adjacency matrix and (iii) analyze the topological properties of the graph.

Researchers have analyzed various topological properties of the ictal network to identify the SOZ. Wilke et al. [Wilke et al., 2011] used directed transfer function (DTF) and betweenness centrality to identify SOZ during ictal iEEG recordings. They found that the betweenness centrality correlated with the location of the resected cortical regions in patients who were seizure free following epilepsy surgery. Furthermore, a better outcome was associated with the resection of brain regions showing the highest betweenness centrality. This finding suggested that critical highly interconnected nodes played a crucial role in seizure onset and spreading. In another paper, van Mierlo et al. [Van Mierlo et al., 2018] studied the temporal changes in effective connectivity during the first 20 seconds of ictal rhythmic iEEG activity in the 3–20 Hz band, in eight patients suffering from focal drug resistant epilepsy who underwent successful resective surgery. In all the patients, the authors found that the iEEG contact showing maximal ictal out-degree was among those visually classified by clinicians as belonging to the SOZ, and it was always included within the resected brain region. Furthermore, the patient specific connectivity patterns were consistent over the majority of seizures, suggesting that most of the seizures originated from the same area and that the same driving hubs play a leading role into the epileptic network. Study carried out by Haneef et al. [Haneef and Chiang, 2014] analyzed different connectivity measures and centrality measures for the

identification of SOZ using ictal iEEG signals. They found that the graph theory centrality measures, betweenness centrality, outdegree, and graph index complexity perform well for the localization of SOZ. Geier et al. [Geier and Lehnertz, 2017] used connectivity measures such as normalized maximum lag correlation function, mean phase coherence, and phase lag index combined with centrality measures such as strength centrality, closeness centrality, betweenness centrality, and eigenvector centrality to localize SOZ. In yet another recent study by Nithin et al. [Nithin et al., 2021], three graph theoretical centrality measures, namely, simple graph energy based centrality, Laplacian centrality, and distance energy centrality, are used to identify the electrodes placed at the SOZ. K-means clustering is used to cluster electrodes into focal and non-focal. The centrality values obtained from the three methods show a significant difference between focal and non-focal electrodes during the seizure phase.

Overall, these studies showed that you can accurately estimate the SOZ by analyzing the topological properties of the ictal network. However, there is an unobserved part of EZ that do not generate seizures during the 2 week long observation known as the Potential Seizure Onset Zone (PSOZ). The current studies based on ictal network properties fail to identify PSOZ and hence leads to inaccurate estimation of EZ.

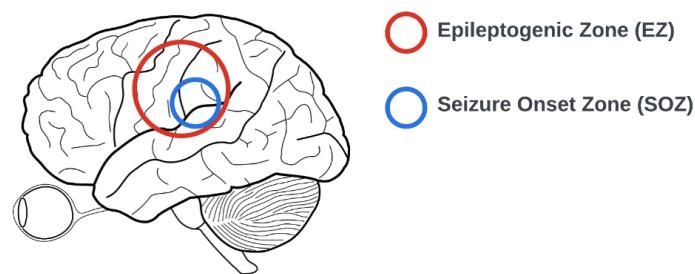


Figure 1.13: **Potential Seizure Onset Zone (PSOZ):** Parts of EZ not generating seizures even during 2 week long observation - such regions are called Potential Seizure Onset Zone (PSOZ)

1.5 Problem Statement and Contributions

As already discussed, about one-third of patients with focal seizures, cannot be treated with anti-seizure medications and they need to undergo a resective surgery for the removal of Epileptogenic zone (EZ), which is the site of the cortex responsible for generating seizures. In the pre-surgical stage, the patient is placed under intracranial EEG (iEEG) monitoring in the hospital leading to iEEG recordings during actual seizures, referred to as ictal segments. Synchronous electrical signals recorded in the ictal segments have been modeled as network/collective dynamics involving all the channels leading to automated identifications of channels that drive the observed seizures. These channels are referred to as seizure onset zones (SOZs) as they constitute parts of the EZ active during observed seizures. Another correlate of this synchronous activity are short duration oscillatory field potential, known as High Frequency Oscillations (HFO), that are observed at the level of individual electrodes of iEEG. SOZ channels have distinctly higher rates of HFOs during the ictal segment, allowing neurologists to identify SOZs without any explicit network modeling. Once the SOZ has been identified then the surgeons resect the SOZ if possible. However, this approach has led to success in about only 60 – 70% of the patients because there might be parts of EZ that did not participate in generating seizures during the limited observation window; such unobserved parts of EZ are known as potential seizure onset zones (PSOZs). The inability of SOZ to completely encapsulate EZ in many patients, along with the hardships and risks associated with lengthy hospitalization period -often lasting two weeks or more- has prompted the need to find accurate physiological biomarkers of EZ during the interictal period, i.e, the majority of the time when patients do not have seizures. HFOs observed during ictal periods have also been observed to be present at higher rates in SOZ channels (determined from ictal periods) during interictal periods, leading to the hope that resection of channels with high interictal HFO rates would lead to seizure freedom. However, the presence of HFOs arising from cognitive processes (physiological HFOs) during interictal periods have diminished the predictive power of interictal HFO rate in the context of surgical outcome prediction. In

the first part of the thesis, we develop a set of weakly supervised deep learning models to separate the cluster of detected HFOs into a cluster of pathological HFO: epileptogenic HFO (eHFO) and a cluster of physiological HFO: eloquent cortex HFO (ecHFO). In retrospective validation on a patient cohort of 15 patients, the eHFO cluster was found to be a better biomarker of EZ compared to Real HFO cluster (HFO cluster after filtering for artifacts) as it was able to correctly predict the post surgical outcomes of patients with an F1 score of 87% in comparison to Real HFO cluster's 72%. However, when tested on a much larger patient cohort (159 patients), we found that a significant percentage of patients ($\sim 25\%$) did not have enough HFO detections and as a result the eHFO resection ratio was not able to correctly predict the surgical outcomes of those patients. Therefore, there is a need to look beyond HFOs in the space of potential interictal biomarkers of EZ. In the second part of the thesis, we develop a synchronization network based approach to model the interactions amongst the channels in the interictal period and leverage the temporal dynamics of the network properties to propose an interictal network biomarker of the epileptogenic zone. In retrospective validation on a patient cohort of 159 patients, the interictal network based biomarker was able to correctly predict the post surgical outcomes of patients with an F1 score of 89%. Finally, in the last part of the thesis, we combine all the biomarkers: rate of pathological HFOs and interictal network properties to propose a novel interictal biomarker of epileptogenic zone. In retrospective validation on a patient cohort of 159 patients, the novel biomarker was able to correctly predict the post surgical outcomes of patients with an F1 score of 98%. In the following paragraphs, we briefly summarize the frameworks that we developed to better estimate the epileptogenic zone.

1.5.1 Unsupervised learning of spatiotemporal interictal discharges in focal epilepsy: an exterior method for non-negative matrix factorization (eNMF)

Interictal epileptiform discharges (IED), like spikes and high frequency oscillations (HFO), are an important biomarker for localization of epileptogenic zone (EZ), especially in patients

suffering from focal drug resistant epilepsy. Although rule based detectors have been developed, but unsupervised learning offers a powerful approach towards automated identification of such pathological IED. Recently, Non-negative matrix factorization (NMF) has been used to automatically detect IED from long iEEG recordings by leveraging spatiotemporal features across all iEEG channels [Baud et al., 2018]. While investigating these results, we found that there are several popular algorithms for NMF and they all tended to converge at different rates and lead to seemingly different factor matrices. Such numerical results prompted us to study the geometry of the underlying optimization problem and numerically establish relationships among the different algorithms. Chapter 2 of this thesis presents fundamental contributions to the field of NMF factorization that evolved as part of our investigations. The NMF problem is to find matrices $U \geq 0$ and $V \geq 0$ of given dimensions so that the product UV^T approximates a data matrix X . Due to its wide applications in machine learning tasks, several algorithms have been developed in the literature for the NMF problem. All these algorithms, however, are interior methods in the sense that they either start with a feasible point or obtain the feasibility of the factors after each iteration. We propose an exterior method that constructs nonnegative factor matrices starting with optimal matrix factors of the unconstrained matrix factorization problem, which are not necessarily nonnegative. A key step is an *algorithm to rotate the optimal matrix factors for the unconstrained problem* (for example, computed using SVD) *to a point closest to the positive orthant* (based on suitable distance measures); simple iterative algorithms are then sufficient *to start at this exterior initial point* and reach a local minimum on the boundary. The proposed algorithm performs better than existing state of the art (SOTA) algorithms in terms of both reconstruction error and time complexity, when tested on multiple real and synthetic data sets. Since the columns of U and V are used as features for downstream tasks, we also empirically investigate if the feature vectors are scaled and orthogonal versions of each other, preserving pair-wise cosine distances and hence, clustering properties. Our empirical results, collected over more than 50 instances of both real and simulated datasets, show that when

our algorithm satisfied the optimality conditions the other algorithms' factor matrices approached solutions that are equivalent; in only one instance a competing algorithm reached a non-equivalent local minima with a higher reconstruction error.

1.5.2 Purification of detected HFO clusters

HFO's detected by automated detectors can be broadly classified into 3 categories:

- Artifacts
- Physiological HFO
- Pathological HFO

Several algorithms have been developed (we also developed one) in the literature to filter out artifacts from detected HFO's. The fact that HFOs are not only pathological in nature but also occur under physiological conditions is a big challenge when assessing the validity of HFOs as a marker for epilepsy. Distinguishing pathological from physiological HFOs might increase the specificity of that marker. This requires defining HFOs to be considered either being physiological or pathological. For instance, continuous high frequency activity in the background EEG has been suggested to reflect physiologic activity distinctive for certain brain regions, such as the hippocampus or the occipital lobe .Concordantly HFOs have been considered to reflect epileptic activity when observed on a flat background, and not when they are embedded in an oscillatory background. Just recently, Liu et al. [Liu et al., 2018] reported on a morphological difference between HFOs obtained in patients with epilepsy and healthy controls, associating stereotypical HFOs with a high degree of waveform similarity to the SOZ of patients and HFOs appearing within random waveforms to functional regions. In addition, HFOs coupled with interictal epileptiform discharges (IEDs), such as spikes, can be considered to reflect epileptic activity as there is a clear association with pathology, and indeed, they have been shown to be more specific for the SOZ than

independent HFO events [Wang et al., 2013]. Furthermore, it has been suggested that the physiologic nature pertains mainly to ripples and that fast ripples mostly represent epileptic activity in these areas [Liu et al., 2018, Jefferys et al., 2012]. In this context, fast ripples might even more exclusively reflect epilepsy-related activity, making them a very promising candidate for clinical use, when present. However, the researchers disagree on the definition of HFOs coupled with spikes and this has led to a lack of inter-rater reliability. Therefore, this approach of classifying HFOs based on morphological differences have not found much clinical acceptance. Hence, there is a need for developing an automated framework for differentiating physiological HFO from pathological HFO and improving localization of EZ.

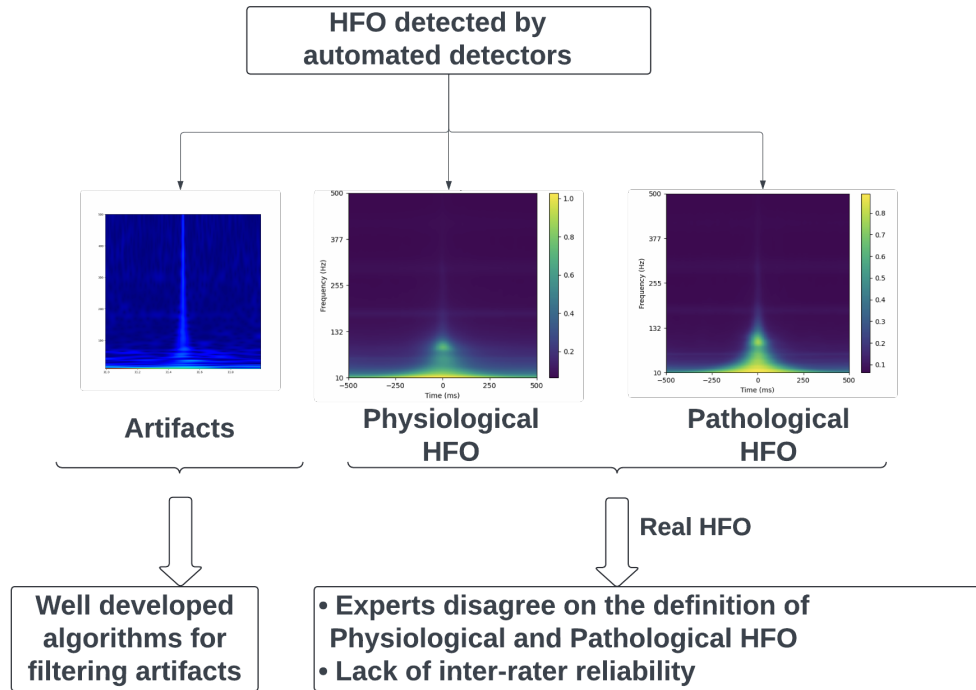


Figure 1.14: **HFO types:** Detections by automated detectors can be broadly classified into three categories: i) artifacts, ii) physiological HFO stemming from healthy tissue and iii) pathological HFO stemming from epileptogenic tissue. While there exists well developed algorithms for filtering artifacts, there is a need for developing algorithms to classify into the other two categories.

1.5.2.1 Clinical evidence guided purification of High frequency oscillations: Discovering Epileptogenic HFOs in the Electroencephalogram (EEG) of Epilepsy Patients

Intracranially-recorded interictal high-frequency oscillations (HFOs) have been proposed as a promising spatial biomarker of the epileptogenic zone. However, visual verification of HFOs is time-consuming and exhibits poor inter-rater reliability. Furthermore, no method is currently available to distinguish HFOs generated from the epileptogenic zone (epileptogenic HFOs: eHFOs) from those generated from other areas (non-epileptogenic HFOs: non-eHFOs). To address these issues, we constructed a deep learning (DL)-based algorithm using HFO events from chronic intracranial electroencephalogram (iEEG) data via subdural grids from 15 children with medication-resistant neocortical epilepsy to: 1) replicate human expert annotation of artifacts and HFOs with or without spikes, and 2) discover eHFOs by designing a novel weakly supervised model (HFOs from the resected brain regions are initially labeled as eHFOs, and those from the preserved brain regions as non-eHFOs). The “purification power” of DL is then used to automatically relabel the HFOs to distill eHFOs. Using 12,958 annotated HFO events from 15 patients, the model achieved 96.3% accuracy on artifact detection (F1 score = 96.8%) and 86.5% accuracy on classifying HFOs with or without spikes (F1 score = 80.8%) using patient-wise cross-validation. Based on the DL-based algorithm trained from 84,602 HFO events from nine patients who achieved seizure-freedom after resection, the majority of such DL-discovered eHFOs were found to be HFOs with spikes (78.6%, $p < 0.001$). While the resection ratio of detected HFOs (number of resected HFOs/number of detected HFOs) did not correlate significantly with post-operative seizure freedom (the area under the curve [AUC]=0.72, $p=0.06$), the resection ratio of eHFOs positively correlated with post-operative seizure freedom (AUC=0.87, $p=0.01$). We discovered that the eHFOs had a higher signal intensity associated with ripple (80-250 Hz) and fast ripple (250-500 Hz) bands at the HFO onset and with a lower frequency band throughout the event time window (the inverted T-shaped), compared to non-eHFOs. We then designed

perturbations on the input of the trained model for non-eHFOs to determine the model’s decision-making logic. The model probability significantly increased towards eHFOs by the artificial introduction of signals in the inverted T-shaped frequency bands (mean probability increase: 0.285, $p < 0.001$), and by the artificial insertion of spike-like signals into the time domain (mean probability increase: 0.452, $p < 0.001$). With this DL-based framework, we reliably replicated HFO classification tasks by human experts. Using a reverse engineering technique, we distinguished eHFOs from others and identified salient features of eHFOs that aligned with current knowledge.

1.5.2.2 Characterizing physiological high-frequency oscillations using deep learning

Intracranially-recorded interictal high-frequency oscillations (HFOs) have been proposed as a promising spatial biomarker of the epileptogenic zone. However, HFOs can also be recorded in the healthy brain regions, which complicates the interpretation of HFOs. In this work we aimed to characterize salient features of physiological HFOs using deep learning (DL). We studied children with neocortical epilepsy who underwent intracranial strip/grid evaluation. Time-series EEG data were transformed into DL training inputs. The eloquent cortex (EC) was defined by functional cortical mapping and used as a DL label. Morphological characteristics of HFOs obtained from EC (eHFOs) were distilled and interpreted through a novel weakly supervised DL model. A total of 63,379 interictal intracranially-recorded HFOs from 18 children were analyzed. The eHFOs had lower amplitude throughout the 80-500 Hz frequency band around the HFO onset and also had a lower signal amplitude in the low frequency band throughout a one-second time window than non-eHFOs, resembling a bell-shaped template in the time-frequency map. A minority of eHFOs were HFOs with spikes (22.9%). Such morphological characteristics were confirmed to influence DL model prediction via perturbation analyses. Using the resection ratio (removed HFOs/detected HFOs) of non-eHFOs, the prediction of postoperative seizure outcomes improved compared

to using uncorrected HFOs (area under the ROC curve of 0.82, increased from 0.76). We characterized salient features of physiological HFOs using a DL algorithm. Our results suggested that this DL-based HFO classification, once trained, might help separate physiological from pathological HFOs, and efficiently guide surgical resection using HFOs.

1.5.3 Novel interictal biomarker of epileptogenic zone

Recently, there is a growing interest in determining whether synchronization effects can be observed in the interictal period and their temporal dynamics can be leveraged to delineate EZ. The problem of studying such effects and using them for better surgical outcome prediction is still open and we address it in this part of the thesis. In particular, we use Power-Phase coupling amongst channels to construct a sequence of directed weighted networks from interictal segments. We leverage the topological dynamics of the network, both local and global, to train a machine learning model to identify SOZ (ground truth obtained from ictal data) from purely interictal segments. The model identifies the SOZ with over 95% accuracy. One of the hallmarks of the constructed networks is that they occasionally transition into a state of hyper-synchrony with SOZ and PSOZ nodes being the hub of these hyper-synchronous states. We hypothesize that these hyper-synchronous states are 'mini seizures' in the interictal phase and our machine learning model is able to identify them and use them for not only accurate SOZ identification but also identify PSOZ. The only way to validate whether our model has truly identified PSOZ and hence EZ is through surgical outcome prediction. We construct a set of features from SOZ model prediction scores along with the constructed network flow dynamics to propose a network based biomarker of EZ. In retrospective validation on a patient cohort of 159 patients, the network based biomarker was able to correctly predict the post surgical outcomes of patients with an F1 score of 89%. Finally, we develop an integrated framework to exploit the interplay between the constructed network and pathological HFO cluster to propose a novel biomarker of EZ. At the heart of this framework is a regression model which predicts the pathological HFO rate using a mix-

ture of local and global properties of the constructed network. The summary statistics of the R^2 of the regression model along with the previously computed features (SOZ model prediction scores and network flow dynamics) is proposed as a novel biomarker of EZ. In retrospective validation on a patient cohort of 159 patients, the novel biomarker was able to correctly predict the post surgical outcomes of patients with an F1 score of 98%. A closer inspection into the summary statistics of the regression model for non-seizure free patients reveals a cluster of pathological HFO whose rate cannot be explained by the constructed network thus revealing the presence of a brain region capable of generating seizures outside of the one sampled by iEEG.

1.6 Organization of the rest of the thesis

The rest of the thesis is organized as follows:

- In chapter 2, we present an efficient algorithm for Non-negative matrix factorization (NMF).
- In chapter 3, we propose a weakly supervised deep learning method which uses clinical clues to refine noisy clustering in discovered HFOs to extract the cluster of epileptogenic HFO (eHFO)
- In chapter 4, we propose a weakly supervised method which uses clinical clues to refine noisy clustering in discovered HFOs to extract the cluster of eloquent cortex HFO (ecHFO)
- In chapter 5, we construct a dynamic directed weighted network from interictal iEEG and exploit the dynamics of the network topology to accurately identify EZ from short interictal segments
- Finally, in Chapter 6 we summarize our contributions and discuss several avenues to extend the work reported in this thesis

CHAPTER 2

Exterior Method for Nonnegative Matrix Factorization

2.1 Introduction

Interictal epileptiform discharges (IEDs) are a hallmark of the epileptic brain, and their localization is a critical complement to definitive delineation of the epileptogenic zone (EZ) for optimal treatment of focal drug resistant epilepsy. IEDs include spikes (or sharp waves) and high-frequency oscillations (HFOs) that can coexist or occur independently and are thought to collectively reflect synchronous activity in a hyper-connected epileptic network. To date, their identification in intracranial EEG (iEEG) is based on visual recognition of 3 distinctive features: (i) a sudden increase in slope (spikes), (ii) a transient increase in HFOs, and (iii) propagation to neighboring brain areas. Manual interpretation becomes limited and impractical when data increases as a factor of time of recording and number of channels in the setting of chronic high-density intracranial implantation for presurgical workup (>50 electrodes, weeks long). Computer based iEEG labeling has been investigated for decades and is available with some commercial EEG packages, but often rely on rule-based detection (eg, amplitude, slope) and, consequently, the sensitivity and specificity trade off is usually poor. The issue of variable IEDs morphology was recently addressed by multi-template matching, clustering algorithms, and adjustable wavelet analyses, but was not extended to the question of patient-specific networks of spike propagation. Recently, a group of researchers at UCSF [Baud et al., 2018] developed a machine-learning algorithm, based on non-negative matrix factorization (NMF), that leverages spatiotemporal features across all channels and all time points of 24-h iEEG recordings. They hypothesized that

NMF could recognize epileptic events from normal physiological data in multichannel EEG. The rationale being that IEDs could be schematically represented as discrete activation of recurring patterns of channel involvement. They found that their algorithm is capable of personalizing to each patient, synthesizing the data into a single anatomic map without human supervision, and detecting and quantifying innumerable IEDs.

Non-negative matrix factorization (NMF) is the problem of finding good approximations of a given non-negative matrix as a low-rank product of two non-negative matrices. This linear dimensionality reduction technique has been used very successfully for a large variety of machine learning and data mining tasks, including hyperspectral imaging [Gillis, 2017], topic modeling [Arora et al., 2012b], audio denoising [Schmidt et al., 2007], and video structuring [Essid and Févotte, 2013, Berry et al., 2007], communication complexity [Lee et al., 2009], probability [Carlini and Rapallo, 2010], and computational geometry [Gillis and Glineur, 2012b].

A standard optimization framework for computing the nonnegative factors uses the Frobenius norm, and a bi-convex formulation is given by

$$\underset{U \geq 0, V \geq 0}{\text{minimize}} \quad f(U, V) = \frac{1}{2} \|X - UV^T\|_F^2, \quad (2.1)$$

where $X \geq 0$. There exist other variants of the NMF depending on the distance measure used to quantify the quality of the approximation. For example, in [Lee and Seung, 2001] both the Frobenius norm (as in (2.1)) and the generalized Kullback-Leibler divergence measure are proposed to measure approximation error, in [Févotte et al., 2009] the author utilizes Itakura-Saito divergence for music analysis and in [Zhang et al., 2008] total variation norm is designed to extract image patterns. In yet another variation (see Section 2.6), the approximation error minimization in (2.1) is done only over a subset of the entries of the data matrix X . In such a case it is referred to as the *matrix completion factorization problem*, and we address this problem as well in Section 2.6.

In addition to using different cost functions to capture the approximation errors, a number

of works have added further constraints and regularization terms to the objective function to suit the needs of a wide variety of applications. For example, to mitigate the impact of outliers [Wei et al., 2022] uses optimizable weight for each attribute of each data point to emphasize their importance while downplaying the influence of potential outliers or noisy data point; [Chen and Li, 2022] introduces a new entropy loss to minimize the entropy of the residue distribution while still allowing for some larger error samples; [Zhang et al., 2019] employs NMF for zero shot learning; [Jia et al., 2019] utilizes NMF with a pair of similarity and dissimilarity regularizers to obtain more discriminative representations for clustering tasks; [Che et al., 2021] employs Recurrent Neural Networks (RNNs) to optimize for sparse factors with an l_0 norm regularizer. To further incorporate the inherent relationships present in the data into the factorization process, [Meng et al., 2019] adds additional regularization term for latent vectors based on a similarity matrix derived from the data; [Liu et al., 2016] introduces a large cone constraint for the left factor; [Yang et al., 2016] utilizes a data-based smoothness factor to achieve a balance between the decomposition error and the sparsity of factors; [Guan et al., 2012] provides an online algorithm for approximating factors with streaming data.

2.2 Prior work

In this chapter we address the primary challenge in solving the optimization problem underlying NMF: ensuring the non-negativity of the factors U and V . While, as mentioned above, regularization terms are indeed added to modify the objective function in (2.1), the algorithms used are almost always variants of the algorithms developed for (2.1). For example, [Wei et al., 2022], [Chen and Li, 2022], mentioned above, use multiplicative update algorithms introduced in the original paper on NMF [Lee and Seung, 2001]. We thus focus on developing a novel algorithm for solving (2.1) and study properties of the corresponding set of local minima (i.e., the factor matrices) obtained by different algorithms. The algorithm

developed here can be directly adapted for the applications mentioned above and we leave such comparative studies as future work.

Several algorithms have been developed to find good local minima of (2.1) and they can be divided into three general classes: multiplicative update algorithms, gradient descent algorithms, and alternating least squares algorithms. The prototypical multiplicative algorithm (Mult) originated with Lee and Seung [Lee and Seung, 2001] and the algorithm is initialized with non-negative factor matrices and they remain non-negative throughout the iterations. However, it is notoriously slow to converge and requires many more iterations than alternatives such as the gradient descent and alternating least squares algorithms. To overcome such shortcomings accelerated versions of the multiplicative algorithm have been proposed [Gillis and Glineur, 2012a].

NMF algorithms of the second class are based on gradient descent methods (Grad-Mult) [Lin, 2007a]. These algorithms always take a step in the direction of the negative gradient, the direction of steepest descent. In order to ensure non-negativity of the factor matrices after the descent step, a projection step is employed by these algorithms.

NMF algorithms of the third class are based on alternating optimization framework. This framework exploits the fact that, while the optimization problem of (2.1) is not convex in both U and V , it is convex in either U or V . Therefore, these algorithms update the factor matrices in an alternating fashion by solving simpler subproblem. The subproblem that arises for each factor is non-negative linear least squares (NNLS). Various algorithms have been developed to solve NNLS and we list the most popular ones below:

- ALS [Lin, 2007b]: Solve unconstrained least squares followed by a simple projection to the non-negative orthant
- AO-ADMM [Huang et al., 2016]: Alternating direction method of multipliers (ADMM) is used to solve NNLS. The algorithm uses cholesky decomposition of the mixing matrix $U^T U$ for an efficient implementation

- Active set [Kim and Park, 2011]: This algorithm searches iteratively for optimal set of variables to optimize over by exchanging a variable between active and passive sets
- Hierarchical ALS (HALS) [Cichocki and Phan, 2009]: Each column of left factor matrix (U) or each row of the right factor matrix (V) are updated by solving an unconstrained least squares and then projecting them to the non-negative orthant.

The three classes of NMF algorithms have one thing in common: *these iterative algorithms either start with a feasible point* or obtain the feasibility of the iterates after each iteration.

2.3 Exact NMF, Complexity, and SVD

The complexity of a special case of the optimization problem in (2.1) where one asks whether there is an *exact* non-negative factorization, has been studied extensively in the literature. More precisely, the *Exact NMF* problem takes as input a matrix $X \geq 0 \in \mathbb{R}^{n \times m}$ with rank equal to r , $r \geq 1$, and asks whether there is a pair of matrices (U, V) , where $U \geq 0 \in \mathbb{R}^{n \times r}$ and $V \geq 0 \in \mathbb{R}^{m \times r}$, such that $X = UV^T$. Vavasis [Vavasis, 2009] and Arora [Arora et al., 2012a] have shown that this problem is NP hard and there is no $(nm)^{o(r)}$ time algorithm for *Exact NMF*, unless $P = NP$. For some special cases of the Exact NMF problem, namely the separable case, first Donoho [Donoho and Stodden, 2004] and then [Arora et al., 2012a] have developed polynomial time algorithms in n , m , and r to compute the exact factorization if one exists. Since the exact factorization scenario is a special case of the general *NMF* problem, it implies that the *NMF* problem (2.1) is also NP-hard, and we cannot find any provably efficient algorithm unless $P = NP$. In order to develop a framework for solving (2.1) that is more efficient than the ones in literature as reviewed in Section 2.2, one will have to judiciously exploit any geometry or structure inherent in the landscape of the underlying optimization problem. We next provide one such geometrical framework.

2.3.1 Going from SVD to known Exact NMF factors

Let $U_1, U_2 \in \mathbb{R}^{n \times r}$ and $V_1, V_2 \in \mathbb{R}^{m \times r}$ be matrices with full column rank and $U_1 V_1^T = U_2 V_2^T$. Then, there exists $A \in \mathbb{R}^{r \times r}$ such that $U_2 = U_1 A$ and $V_2 = V_1 A^{-T}$ [Vavasis, 2009] [Stewart, 2001].

In the context of Exact-NMF, let

$$X = U_0 V_0^T,$$

where $X \geq 0 \in \mathbb{R}^{n \times m}$ is a non-negative data matrix and $U_0 \geq 0 \in \mathbb{R}^{n \times r}$, $V_0 \geq 0 \in \mathbb{R}^{m \times r}$ are the Exact-NMF factors. Also, let X admit the following SVD

$$X = U \Sigma V^T,$$

and we construct the scaled SVD factor matrices as

$$\begin{aligned} U^* &= U \Sigma^{\frac{1}{2}} \\ V^* &= V \Sigma^{\frac{1}{2}}. \end{aligned}$$

Since,

$$U_0 V_0^T = U^* V^{*T}$$

and all the matrices have full column rank so by the result stated above there exists A such that

$$\begin{aligned} U_0 &= U^* A \\ V_0 &= V^* A^{-T}. \end{aligned} \tag{2.2}$$

If we can solve (2.2) for A then we have a way to go from the SVD factors to known Exact-NMF factors. This exercise of going from SVD to known Exact NMF factors will motivate the optimization problem for the inverse case, where we need to solve for an A that transforms SVD factors to unknown Exact-NMF factors.

Let the SVD of A be given by,

$$A = R_1 \Lambda R_2^T,$$

where $R_1, R_2 \in \mathbb{R}^{r \times r}$ are orthogonal matrices and Λ is a diagonal scaling matrix. Substituting the SVD of A in (2.2) we get,

$$\begin{aligned} U_0 &= U^* R_1 \Lambda R_2^T \\ V_0 &= V^* R_1 \Lambda^{-1} R_2^T. \end{aligned} \tag{2.3}$$

Substituting the SVD of U^* in (2.3) we get

$$U_0 = U \Sigma^{\frac{1}{2}} R_1 \Lambda R_2^T \tag{2.4}$$

Multiplying both sides of (2.4) to the left by $\Sigma^{-\frac{1}{2}} U^T$ we get

$$\Sigma^{-\frac{1}{2}} U^T U_0 = R_1 \Lambda R_2^T = A \tag{2.5}$$

Hence, we get A from the SVD of $\Sigma^{-\frac{1}{2}} U^T U_0$.

2.3.2 Going from SVD to unknown Exact NMF factors

To solve the Exact NMF problem, we can first compute the scaled SVD factors, (U^*, V^*) , of a nonnegative X , and then solve for $R_1, R_2 \in \mathbb{R}^{r \times r}$ ($R_1 R_1^T = I = R_2 R_2^T$), and diagonal scaling matrix Λ , such that $U^* R_1 \Lambda R_2 \geq 0$, and $V^* R_1 \Lambda^{-1} R_2 \geq 0$. If X has an exact NMF solution, then this problem is feasible. One can define the following non-convex optimization problem to solve for the orthogonal and scaling matrices:

$$\begin{aligned} &\underset{R_1, R_2, \Lambda}{\text{minimize}} \quad \sum_{i=1}^n \sum_{j=1}^r h((U^* R_1 \Lambda R_2)_{ij}) + \sum_{i=1}^m \sum_{j=1}^r h((V^* R_1 \Lambda^{-1} R_2)_{ij}) \\ &\text{subject to} \quad R_1^T R_1 = I, \quad R_2^T R_2 = I, \quad \Lambda \geq 0 \end{aligned} \tag{2.6}$$

where $h(q) = \max\{0, -q\}$. Clearly the optimization problem in (2.6) has a minimum value $= 0$ if and only if X admits an Exact solution. This, however, is an NP Hard problem, and our ADMM implementations for solving (2.6) (a generalized version of the eNMF algorithm presented in Section 2.5) did yield exact non-negative solutions in reasonable time, only when X has low dimensions. For example for synthetic datasets (for data generation procedure read Section 2.7.1) with dimensions $n = 50, m = 40, r = 10, s = 0.1$ and $n = 100, m = 80, r = 20, s = 0.1$ (where s is the sparsity factor) the ADMM implementation yielded almost exact non-negative solutions. However, for an X when it's dimensions n, m are large, the ADMM implementations yielded factors that are only close to the positive orthant, **which are still external infeasible points** (see Fig. 2.1 for details) This computational intractability of finding optimal R_1, Λ, R_2 to move the SVD solutions to the feasible region (even when such solutions exist), prompted a more approximate approach, as described next in the context of the general NMF problem.

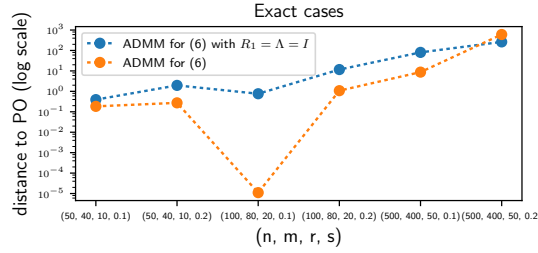


Figure 2.1: Distance to the positive orthant (as defined by the sum of the absolute values of the negative entries of the factor matrices post optimization; see Eqs.(2.6) and (2.8) is plotted for synthetic exact datasets with various dimensions and sparsity levels. For lower dimensions, ADMM implementation of (6) generates exact non-negative solutions. For larger dimensions, ADMM implementations for both the general and relaxed version (involving only a single orthogonal matrix R_2 , and setting $R_1 = \Lambda = I$) generate factors that are only close to the positive orthant (PO), thus requiring further processing.

2.4 General NMF: Going from SVD to the Positive Orthant

The unconstrained optimization version of the problem in (2.1) has a well known global optimum (the minimum value, however, is no longer necessarily 0 as in the case of Exact NMF), given by the truncated singular value decomposition (SVD) of the data matrix X , (U^*, V^*) , as defined in the previous section. In general, the optimal value of (2.1) is bounded below by the optimal value of the unconstrained problem. Therefore, solving (2.6) exactly would still generate factor matrices that are outside the positive orthant and we will have to find a way to move to the positive orthant from this close exterior point. Hence, *to make the problem of getting an approximate solution to (2.6) more tractable, we restrict ourselves, to the case where $R_1 = I$, and $\Lambda = I$.* From Fig. 2.1 it can be observed that this relaxation produces factor matrices that are comparable (in terms of distance to the positive orthant (PO))to the ones obtained by solving the general problem in (2.6). In the relaxed version, we search for a closest point to the positive orthant (as defined by the sum of the absolute values of the negative elements of the point) from the set:

$$\mathcal{X}^* = \{(U^*R, V^*R) \mid R^T R = I, R \in \mathbb{R}^{r \times r}\}. \quad (2.7)$$

Next, to explain further why such a close point in $\mathcal{X}^*$ is a “good” point to start at, we consider the following scenario: If we do a descent on $f(U, V)$ starting from any local minimum of the *NMF* problem, then we will converge to a pair of factor matrices (U, V) and *this external unconstrained global minimum is in some sense “closest” to the boundary with respect to the local minimum* we started at. This external globally-optimal (U, V) is in the same level-set as (U^*, V^*) and as shown in [Zhu et al., 2021] this level set of globally optimal solutions is the set $\mathcal{X}^*$ (modulo any scaling of the (U, V)). In light of this geometric intuition, *we can “reverse” the above descent process: start with an element of $\mathcal{X}^*$ that is close to the boundary* (based on suitable distance measures) *and then use simple exterior methods to find a local minimum of the NMF problem.* We next present such an algorithm for the *NMF*.

2.5 Exterior NMF Algorithm

The proposed algorithm consists of three blocks:

- **Optimal orthogonal transformation of unconstrained global optimum (2.5.1):** find an orthogonal transformation that moves the unconstrained global optimum close to the positive orthant. For several datasets, the transformed factors are non-negative and thereby achieving globally optimal solutions. The next two blocks are for cases for which the transformed factors are infeasible
- **Penalty method for attaining feasibility (2.5.2):** attain feasibility of the transformed unconstrained global optimum using projected block coordinate descent
- **Descent to the local minima (2.5.3):** descend the feasible factors to a local minima of the NMF problem using hierarchical alternating least squares

In the following sections, we describe each block in detail.

2.5.1 Optimal orthogonal transformation of unconstrained global optimum

An orthogonal transformation that minimizes the sum of the negative entries in $U^\star R$ and $V^\star R$ can be found by solving the optimization problem in (2.6) with $R_1 = \Lambda = I$. Although the problem is non-convex, we can use ADMM as a heuristic method to solve it [Boyd et al., 2011, Wang et al., 2019, Hong et al., 2016, Gao et al., 2020]. In order to aid the ADMM formulation, we introduce the following notation

$$W = \begin{bmatrix} U^\star \\ V^\star \end{bmatrix} \in \mathbb{R}^{(n+m) \times r}, Z \in \mathbb{R}^{(n+m) \times r},$$

where Z is the splitting variable. Then, with the above notation we can rewrite the opti-

mization problem in standard ADMM form given by

$$\begin{aligned}
& \underset{Z, R}{\text{minimize}} && \sum_{i=1}^{n+m} \sum_{j=1}^r h(Z_{ij}) \\
& \text{subject to} && Z - WR = 0 \\
& && R^T R = I.
\end{aligned} \tag{2.8}$$

The ADMM for solving (2.8) is given in Algorithm 1.

Algorithm 1 orthogonal

Input: U, V, R^0, Y^0, ρ

Stack U, V vertically to get W

Repeat until convergence:

$$\begin{aligned}
Z^{(k+1)} &\leftarrow \arg \min_Z \left(\sum_{i=1}^{n+m} \sum_{j=1}^r h(Z_{ij}) + \right. \\
&\quad \left. \frac{\rho}{2} \|Z - WR^{(k)} + \frac{1}{\rho} Y^{(k)}\|_F^2 \right) \\
R^{(k+1)} &\leftarrow \arg \min_{R^T R = I} \left(\frac{\rho}{2} \|Z^{(k+1)} - WR + \frac{1}{\rho} Y^{(k)}\|_F^2 \right) \\
Y^{(k+1)} &\leftarrow Y^{(k)} + \rho(Z^{(k+1)} - WR^{(k+1)}) \\
k &\leftarrow k + 1
\end{aligned}$$

Output: UR, VR

2.5.1.1 Updates for Z

We can update each entry of Z independently. The updates can be obtained by solving a simple quadratic problem:

$$Z_{ij}^{(k+1)} = \begin{cases} b_{ij} & b_{ij} > 0 \\ 0 & -\frac{1}{\rho} \leq b_{ij} \leq 0 \\ b_{ij} + \frac{1}{\rho} & b_{ij} < -\frac{1}{\rho}, \end{cases} \tag{2.9}$$

where $b_{ij} = [WR^{(k)} - \frac{1}{\rho} Y^{(k)}]_{ij}$.

2.5.1.2 Updates for R

We can update R by solving a non-convex problem known as Orthogonal Procrustes Problem [Gower et al., 2004]. If we define $B = Z^{(k+1)} + \frac{1}{\rho}Y^{(k)}$, then the update for R is given by

$$R^{(k+1)} = CD^T, \quad (2.10)$$

where C and D are obtained from the SVD $W^T B = C\Sigma D^T$.

2.5.2 Penalty method for attaining feasibility

The rotated point returned by Algorithm 1 is expected to be close to the positive orthant but not necessarily inside it (unless there is an NMF solution which is also globally optimal for the unconstrained problem; see Fig. 2.4 for an example). We use an exterior penalty method to attain feasibility of the NMF factors starting from the rotated point:

$$\min_{U,V} \frac{1}{2} \|X - UV^T\|_F^2 + \delta_u \sum_{i,j} h(U_{ij}) + \delta_v \sum_{i,j} h(V_{ij}),$$

where $h(q) = \max\{0, -q\}$, and δ_u, δ_v are the penalty parameters. We solve the optimization problem by iteratively updating U and V using gradient descent.

2.5.2.1 Updating U

Assuming δ_u is large, the partial derivatives of the cost function for negative entries of U are approximately $-\delta_u$, and the negative entries can be updated as,

$$U_{ij} = U_{ij} + \rho_u \delta_u, \quad \forall (i, j) \in I_- , \quad (2.11)$$

where I_- denotes the set of negative entries and ρ_u is a fixed step size.

Updating the non-negative entries of U in any given row does not affect the gradient of the cost function for non-negative entries of U in other rows and we can leverage this to update non-negative entries in the rows of U independently. The updates for all the nonnegative

entries in the i^{th} row can be computed as

$$U_{ij} = (U_{ij} - d_i^*(\nabla_U f)_{ij})_+, \quad \forall (i, j) \in I_+, \quad (2.12)$$

where I_+ denotes the set of non-negative coordinates, $\nabla_U f = (UV^T - X)V$, and the projection operator, $(x)_+ = \max\{0, x\}$, retains their nonnegativity, and d_i^* is the optimal step size. In order to find the optimal step size, we need to formulate an optimization problem. To aid the formulation, we introduce the error matrix at $(k+1)$ iteration as

$$E^{(k+1)} = X - U^{(k+1)}V^{(k+1)T}.$$

Then with the above definitions, we can define the optimization problem as $\min. L = \|E^{(k+1)}(i, :)\|^2$, or

$$\min_{d_i} \|X(i, :) - [U^{(k)}(i, :) - d_i\{M_{U_+} \circ \nabla_U f\}(i, :)]V^{(k)T}\|^2 \quad (2.13)$$

where M_{U_+} is the binary vector mask over the negative entries of U . The problem given by (2.13) is an unconstrained quadratic problem and has a global minimum. The global minimum can be computed by setting the gradient with respect to d_i to zero

$$\nabla_{d_i} L = \sum_{j=1}^n \{2E^{(k)}(i, j)a_j + 2a_j^2 d_i\}$$

where $a_j = \{M_{U_+} \circ \nabla_U f\}(i, :)V^{(k)T}(:, j)$. Solving $\nabla_{d_i} L = 0$, we get

$$d_i^* = \frac{-\sum_{j=1}^n E^{(k)}(i, j)a_j}{\sum_{j=1}^n a_j^2}$$

The expression for d_i^* can be simplified further to give

$$d_i^* = \frac{\|\{M_{U_+} \circ \nabla_U f\}(i, :)\|^2}{\|\{M_{U_+} \circ \nabla_U f\}(i, :)V^{(k)T}\|^2} \quad (2.14)$$

2.5.2.2 Updating V

We follow a similar routine to update V . First we update the negative entries of V as,

$$V_{ij} = V_{ij} + \rho_v \delta_v, \quad \forall (i, j) \in I_- , \quad (2.15)$$

The non-negative entries in the rows of V are updated independently. The updates for all the non-negative entries in the i^{th} row can be computed as

$$V_{ij} = (V_{ij} - t_i^*(\nabla_V f)_{ij})_+, \quad \forall (i, j) \in I_+, \quad (2.16)$$

where t_i^* is the optimal step size, and $\nabla_V f = (UV^T - X)^T U$. The optimal step size computation is similar as before and is given by

$$t_i^* = \frac{\|\{M_{V_+} \circ \nabla_V f\}(i, :)\|^2}{\|\{M_{V_+} \circ \nabla_V f\}(i, :)^T U^{(k+1)}\|^2} \quad (2.17)$$

The pseudocode for updating the nonnegative entries is given below:

Algorithm 2 Projected Block Coordinate Descent (PBCD)

- 1: **Input:** $X, U, V, \epsilon, \text{max_iter}, M_{U_+}, M_E$
 - 2: $\nabla_{U_{\text{init}}} f \leftarrow M_{U_+} \circ ((M_E \circ (UV^T - X))V)$
 - 3: $n \leftarrow$ number of rows in U
 - 4: **Repeat until** $\text{projected_grad} < \epsilon \times \|\nabla_{U_{\text{init}}} f\|_F$ **or** max_iter :
 - 5: **for** $i = 1..n$
 - 6: $d_i^* \leftarrow$ Using Eq. 2.14
 - 7: $U_{i,:} \leftarrow$ Using Eq. 2.12
 - 8: $\text{projected_grad} \leftarrow \|M_{U_+} \circ ((M_E \circ (UV^T - X))V)\|_F$
 - 9: **Output:** U
-

We refer to this attaining feasibility stage as the ascent stage because it leads to a reconstruction error which is higher than the one attained by the rotated SVD factors.

2.5.3 Descent to the local minima

The ascent stage described in Section 2.5.2 returns factors that are already very close to a local minima (as evaluated by the KKT conditions) and any standard descent algorithm will move the feasible factors to a local minima of the NMF problem. The PBCD algorithm

outlined in the previous section is sufficient to move the feasible factors to a local minima in a reasonable time complexity but we use hierarchical alternating least squares (HALS) for the descent stage. The descent stage terminates once the factor matrices U and V converge to a local minima. The convergence to a local minima is verified by checking for the KKT optimality conditions of (2.1) given below:

$$\begin{aligned}\nabla W, W &\geq 0 \\ \nabla W \odot W &= 0\end{aligned}\tag{2.18}$$

where $W = \begin{bmatrix} U \\ V \end{bmatrix}$. The pseudocode for eNMF is provided below with three previously introduced main blocks:

Algorithm 3 eNMF

- 1: **Input:** $X, r, \rho, \epsilon, \text{max_iter}$
 - 2: $U^*, V^* \leftarrow SVD(X, r)$
 - 3: $(U^*R, V^*R) \leftarrow \text{orthogonal}(U^*, V^*, R^0, Y^0, \rho)$
 - 4: $(U, V) \leftarrow (U^*R, V^*R)$
 - 5: **Repeat until U,V non negative:**
 - 6: Update negative elements in U using Eq. 2.11
 - 7: Compute mask M_{U+}
 - 8: $U \leftarrow PBCD(X, U, V, \epsilon, \text{max_iter}, M_{U+}, \mathbf{1})$
 - 9: Update negative elements in V using Eq. 2.15
 - 10: Compute mask M_{V+}
 - 11: $V \leftarrow PBCD(X^T, V, U, \epsilon, \text{max_iter}, M_{V+}, \mathbf{1})$
 - 12: $U, V \leftarrow HALS(X, U, V)$
 - 13: **Output:** U, V
 - 14: **Note:** $\mathbf{1}$ is the matrix of all 1's
-

2.6 Exterior NMC algorithm

The eNMF algorithm is extended to the case where the data matrix X has missing entries.

We compute a local minimum of the unconstrained problem

$$\min_{U,V} \hat{f}(U,V) = \frac{1}{2} \|M_E \circ (X - UV^T)\|_F^2,$$

where M_E is the binary mask over the unknown entries of X , using an alternating least squares algorithm (softImpute-ALS) proposed in [Hastie et al., 2015]. Once we obtain a local minimum, then we follow the same sequence (i.e. orthogonal of (U, V) closest to the positive orthant, followed by an exterior penalty method) to obtain a solution to the nonnegative matrix completion problem. The gradient computations in the exterior penalty method are updated accordingly: $\nabla_U \hat{f} = (M_E \circ (UV^T - X))V$, $\nabla_V \hat{f} = (M_E^T \circ (UV^T - X)^T)U$. The pseudocode for eNMC is given below:

Algorithm 4 eNMC

- 1: **Input:** $X, r, M_E, \epsilon, \text{max_iter}$
 - 2: $U_0, V_0 \leftarrow \text{softImpute}(X, r, M_E)$
 - 3: $(U_0 R, V_0 R) \leftarrow \text{orthogonal}(U_0, V_0, R^0, Y^0, \rho)$
 - 4: $(U, V) \leftarrow (U_0 R, V_0 R)$
 - 5: **Repeat until convergence:**
 - 6: Update negative elements in U using Eq. 2.11
 - 7: Compute mask M_{U+}
 - 8: $U \leftarrow \text{PBCD}(X, U, V, \epsilon, \text{max_iter}, M_{U+}, M_E)$
 - 9: Update negative elements in V using Eq. 2.15
 - 10: Compute mask M_{V+}
 - 11: $V \leftarrow \text{PBCD}(X^T, V, U, \epsilon, \text{max_iter}, M_{V+}, M_E^T)$
 - 12: **Output:** U, V
-

2.7 Experiments and Results

In this section, we run experiments on synthetic and real datasets and demonstrate the superiority of the proposed algorithm in terms of reconstruction error and time complexity.

2.7.1 Datasets and algorithms

Exact factorization dataset: Although we developed an algorithm for the *NMF* problem, we would also like to evaluate the performance of eNMF when applied to the *exact NMF* problem. For this evaluation process, we created a synthetic dataset that admits an exact factorization. The dataset was created by following the procedure: **i)** generate the entries of the ground truth factor matrices U and V by sampling randomly from $U(0, 1)$, **ii)** a binary mask with sparsity s is applied to the ground truth factors and **iii)** generate the data matrix using $X = UV^T$.

Synthetic dataset (*Dense and High Dimensional X*): It is standard in the NMF literature to use randomly generated matrices to compare the relative performance of algorithms. We generated X as follows: **i)** generate the entries of two matrices $W \in \mathbb{R}^{5000 \times 1000}$, and $H \in \mathbb{R}^{1000 \times 5000}$ by sampling randomly from $U(0, 1)$, **ii)** generate the entries of $N \in \mathbb{R}^{5000 \times 5000}$ by sampling randomly from $\mathcal{N}(0, \tau^2)$, where τ controls the signal to noise ratio (SNR) level, and **iii)** generate the data matrix using $X = WH + N$. The dense and high dimensional X allows one to compare both the scalability and accuracy of different algorithms at various SNR levels.

Verb dataset (*Highly Sparse and Moderate Dimensional X*): This dataset is comprised of entity-verb relationships extracted from a corpus of text data [Tangherlini et al., 2016], where $X(i, j)$ is the number of times an entity i appears as the object of verb j ; $X \in \mathbb{R}^{282 \times 1528}$ and has 97.3% sparsity.

Face dataset (*Highly skewed and High Dimensional X*): The data matrix with shape 32256×64 contains the pixel values of a person’s image under various levels of illumination

[Lee et al., 2005]. Part based decomposition of images is widely used in image processing and sparse factor matrices help to extract the various parts of an image like nose, eyes, etc. This dataset is useful for comparing the trade offs between reconstruction error and sparsity of the factor matrices obtained from different algorithms.

MovieLens dataset (*A popular dataset with 95.5% Missing entries in X*): This dataset contains 1 million ratings by 6040 users on 3706 movies [Harper and Konstan, 2016]. The rating matrix, $X \in \mathbb{R}^{6040 \times 3706}$, has 95.5% missing (user, rating) entries. The quality of recommendation systems are often benchmarked based on their performance on this dataset.

Six well-known *NMF* algorithms were used for performance comparison: accelerated multiplicative algorithm (Mult) [Gillis and Glineur, 2012a], gradient based algorithm (Grad-Mult) [Lin, 2007a], alternating non-negative least squares (ALS) [Lin, 2007b], alternating optimization with ADMM (AO-ADMM) [Huang et al., 2016], hierarchical alternating least squares (HALS) [Cichocki and Phan, 2009], and alternating direction method of multipliers (NMF_ADMM) [Hajinezhad et al., 2016]. However, in [Hajinezhad et al., 2016] a *nested ADMM* approach (where subproblems are themselves solved using ADMM) was used and it had difficulty converging. To alleviate this, we introduced a formulation

$$\begin{aligned} & \underset{U, V, A, B}{\text{minimize}} \quad \|X - AB^T\|_F^2 + I_+(A) + I_+(B) \\ & \text{subject to} \quad U = A, \quad V = B, \end{aligned} \tag{2.19}$$

where I_+ is the indicator function for element-wise nonnegative matrix, and solved (2.19) using ADMM [Boyd et al., 2011] by alternately minimizing the blocks (U, B) and (V, A) . The subproblems in this ADMM formulation turned out to be simple non-negative least squares problems, which can be solved efficiently. This NMF_ADMM algorithm performed much better than the one proposed in [Hajinezhad et al., 2016].

The performance of eNMC algorithm was benchmarked against three widely used NMC algorithms: multiplicative updates (Mult) [Lin and Boutros, 2018], sequential coordinate descent (SCD) [Lin and Boutros, 2018], and alternating direction method (ADM) [Xu et al.,

2012].

2.7.2 Reconstruction error and run time comparison

We observed an interesting phenomenon: if we run the competing algorithms for a long enough time (often for days) then they achieve a similar reconstruction error as eNMF (Fig. 2.2). Therefore, for a meaningful comparison of the relative performance of the algorithms we perform two experiments: (i) **Comparison of equal-time constrained reconstruction errors** (see Fig. 2.4, Table 2.1): For any dataset and choice of r , we first run eNMF until it reaches a local minimum (verified by checking for the KKT conditions, see Table 2.4), and obtain the resulting reconstruction error and note the run time. Then we run the competing algorithms for a time that is at least as long as eNMF's (hence, *they have a better chance of lowering their reconstruction error*) by interrupting their running iterations right after the execution time exceeds the run time of eNMF. (ii) **Comparison of equal-error constrained time complexity** (see Fig. 2.5, 2.6): We set the reconstruction error to that of eNMF, and then run the rest of competing algorithms until they reach eNMF's reconstruction error.

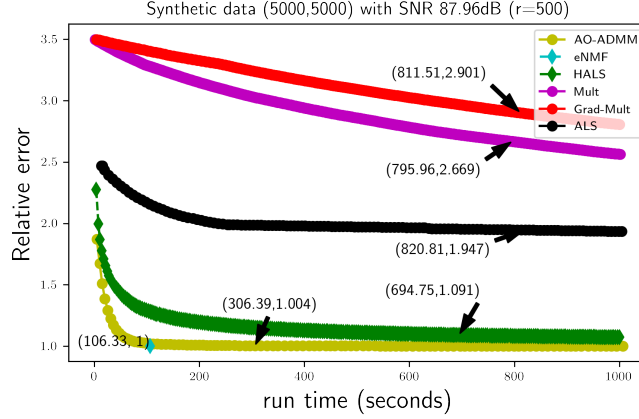


Figure 2.2: **Given enough run time, all the algorithms converge to local minima with similar reconstruction error** (across all the datasets we used): The above figure illustrates it for the Synthetic dataset with $n = m = 5000, r = 500$ and SNR of $-28dB$. Relative error = $\frac{\|X - U_{\text{NMF}}V_{\text{NMF}}^T\|_F}{\|X - U_{\text{SVD}}V_{\text{SVD}}^T\|_F}$ is plotted as a function of run time (secs). *eNMF achieves the minimum (i.e. unconstrained global minimum in this case) in 106 seconds, compared to hours of run time for some of the competing algorithms* Competitors like AO-ADMM have a fast convergence initially but slow down when approaching the minima, as shown by the long tail.

2.7.2.1 Performance of eNMF in exact factorization scenarios

The performance of eNMF, in terms of reconstruction error, was benchmarked against the six existing algorithms listed previously for various sparsity levels. As before, we do an equal-time constrained experiment by letting the competing algorithms run for at least as long as eNMF. Fig. 2.3 plots the reconstruction error as a function of the sparsity level s . It is clear that eNMF outperforms its competitors by converging to the exact factors for sparsity levels 0.2 through 0.5 and achieving the lowest reconstruction error for sparsity level 0.1.

Reconstruction error statistics: mean $\pm$ standard deviation

Dataset	r	eNMF	HALS	NMF-ADMM	Grad-Mult	ALS	AO-ADMM
Face	5	18708.23 $\pm$.00	18708.23 $\pm$.00	18708.23 $\pm$.00	19510.79 $\pm$ 540.99	18708.23 $\pm$.00	18708.23 $\pm$.00
	10	12557.48 $\pm$.00	12583.28 $\pm$ 12.90	17317.94 $\pm$ 2519.46	14051.48 $\pm$ 259.70	12605.34 $\pm$ 54.59	12580.3 $\pm$ 0
	15	9796.13 $\pm$.10	9856.23 $\pm$ 50.33	18037.35 $\pm$ 5816.46	11529.95 $\pm$ 217.71	9954.2 $\pm$ 27.72	9823.4 $\pm$.00
	20	7916.49 $\pm$.01	7976.74 $\pm$ 37.70	29791.24 $\pm$ 7203.81	10330.96 $\pm$ 159.80	7975.07 $\pm$ 33.86	7960.5 $\pm$.00
	25	6371.27 $\pm$ 17.93	6374.11 $\pm$ 18.67	27566.53 $\pm$ 7361.19	8912.64 $\pm$ 543.97	6465.98 $\pm$ 35.98	6372.65 $\pm$.00
Verb	10	14.39 $\pm$.00	14.39 $\pm$.00	14.49 $\pm$.01	14.41 $\pm$.00	14.4 $\pm$.01	14.39 $\pm$.00
	20	13.58 $\pm$.00	13.6 $\pm$.01	13.98 $\pm$.03	13.61 $\pm$.00	13.6 $\pm$.00	13.59 $\pm$.00
	40	12.32 $\pm$.00	12.33 $\pm$.01	14.31 $\pm$.32	12.39 $\pm$.01	12.33 $\pm$.01	12.34 $\pm$.00
	80	10.47 $\pm$.00	10.56 $\pm$.03	17.68 $\pm$ 1.63	10.68 $\pm$.01	10.49 $\pm$.01	10.53 $\pm$.00
	100	9.68 $\pm$.00	9.84 $\pm$ 0.10	18.48 $\pm$ 1.31	9.95 $\pm$.01	9.71 $\pm$.01	9.77 $\pm$.00
Audio	10	9665.11 $\pm$.00	9915.33 $\pm$.17	10192.72 $\pm$ 833.40	51637.65 $\pm$ 88.65	9912.53 $\pm$.06	9810.46 $\pm$.00
	20	9440.42 $\pm$.00	9619.71 $\pm$ 6.89	11981.93 $\pm$ 1521.42	16147.03 $\pm$ 193.48	9498.46 $\pm$ 13.40	9486.61 $\pm$.00
	40	9036.95 $\pm$ 0.02	9292.87 $\pm$ 2.50	13681.51 $\pm$ 5186.36	14186.23 $\pm$ 25.31	9115.48 $\pm$ 3.11	9082.16 $\pm$.00
	80	8342.99 $\pm$ 0.15	8570.29 $\pm$ 4.09	39708.22 $\pm$ 13032.29	18584.87 $\pm$ 131.23	9490.27 $\pm$ 136.78	8402.73 $\pm$.00
	100	8034.89 $\pm$ 0.53	8515.05 $\pm$ 17.47	47585.91 $\pm$ 14363.41	16032.56 $\pm$ 33.48	9538.46 $\pm$ 43.90	8085.34 $\pm$.00

Table 2.1: **Equal-time constrained NMF reconstruction errors (Real datasets):** We tabulated the reconstruction error statistics for 5 runs of each algorithm. From the table, it is evident that eNMF achieves the lowest mean reconstruction error with least variance for all latent dimensions and across both the datasets. **Note:** We didn't tabulate the error statistics for Mult because it has a very high reconstruction error.

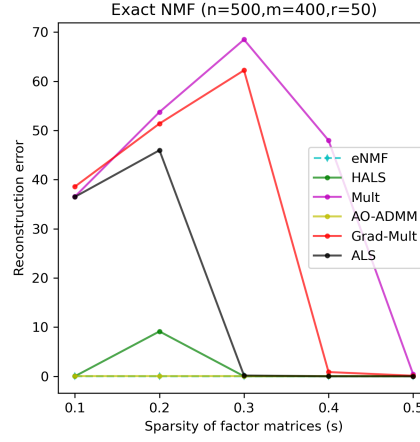


Figure 2.3: **Exact factorization:** eNMF achieves the global minimum for sparsity levels 0.2 through 0.5. Although ALS and HALS achieve the global minimum for sparsity levels 0.3 through 0.5, but they are far away from the global minimum for sparsity level 0.2. For sparsity level 0.1, although HALS and AO-ADMM get similar reconstruction error as ENMF does, they usually have a fast initial convergence but slow down considerably as it approaches the global minimum as shown in Fig. 2.2. **Note:** We dropped NMF-ADMM from this plot due to its high reconstruction error.

2.7.2.2 Performance of eNMF on synthetic datasets

The performance of eNMF, in terms of reconstruction error and run time, was benchmarked against the six existing algorithms on synthetic dataset with four SNR levels. Fig. 2.4 plots the equal-time constrained relative reconstruction error as a function of the latent dimension for various SNR levels. From the plot, it is evident that eNMF achieves the unconstrained global minimum for all latent dimensions and across all the four SNR levels. For higher latent dimensions, the gap in relative error of eNMF and its closest competitor AO-ADMM widens for lower SNR levels (more noise). Fig. 2.5 plots the relative error as a function of run time for various SNR levels. eNMF converges to the global minimum at the fastest rate for all SNR levels and as the noise level increases the gap in the convergence rates of eNMF and its closest competitor widens.

2.7.2.3 Performance of eNMF on real datasets

Besides the synthetic datasets, the performance of eNMF, in terms of reconstruction error and run time, was benchmarked against the six existing algorithms on additional two datasets from real-world scenarios: **Verb** and **Face**. Table 2.1 reports the reconstruction error statistics (over 5 runs) as a function of the latent dimension. We can observe that eNMF achieves the lowest mean reconstruction error with least variance for all latent dimensions and across both datasets. Fig. 2.6 plots the equal-error constrained time complexity as a function of the latent dimensions. Similar to the synthetic case, eNMF attains the local minima fastest for all latent dimensions with wider gap between the convergence rates of eNMF and its closest competitor for higher latent dimensions.

2.7.2.4 Performance of eNMC on movielens dataset

NMC is used to predict the ratings that an user would give on an item. The predicted ratings are then used to make recommendations to the user. Thus, the accuracy of the predicted ratings plays a crucial role in recommendation engines. We evaluated the performance of eNMC in terms of the prediction error: we trained the eNMC algorithm using 80% ratings of the movielens dataset and then used the factors to predict 20% ratings of the movielens dataset. eNMC has a significant advantage over its competitors as shown in Fig. 2.7 where it obtains the lowest prediction error across all latent dimensions. Its superior generalization ability can be explained by the fact that its factors capture the structure in the training data the best considering the lowest relative error for all latent dimensions.

2.7.3 Equivalent local minima and relationships among factors obtained by different algorithms

The primary application of NMF in ML is semantic dimensionality reduction: The factors U and V are used as dimensionality-reduced features in downstream tasks, and they have been

shown to retain key semantic structures present in the higher dimensional original datasets such as images. A natural question to ask is: *Do different algorithms lead to geometrically different factors? If so, then the downstream algorithms' performance would depend on the algorithm used for NMF and also the particular random seeds used in different runs of any given algorithm.* To study this vital question we introduce the notion of equivalent local minima and empirically study relationships amongst factors obtained by different algorithms.

Clearly, if two algorithms reach local minima with different objective values, then the factors are different and non-equivalent. While it is widely believed that the NMF problem admits local minima that have different reconstruction errors, the existence of such distinct local minima for real datasets is not well-known (see Table 2.3 for an example we found).

Let A and B be two NMF algorithms that reach their respective local minima with factor matrices (U_A, V_A) and (U_B, V_B) . Furthermore, let us assume that they reach the same reconstruction error: $U_A V_A^T = U_B V_B^T$. Then, as shown in Vavasis [Vavasis, 2009, Stewart, 2001], there must exist a matrix $Z \in \mathbb{R}^{r \times r}$ such that $U_B = U_A Z$ and $V_B = V_A Z^{-T}$. Thus by considering an SVD of Z , $Z = R_1 \Lambda R_2^T$, we find that there exist orthogonal matrices $R_1, R_2 \in \mathbb{R}^{r \times r}$ ($R_1 R_1^T = I = R_2 R_2^T$) and diagonal scaling matrix Λ such that

$$U_B = U_A R_1 \Lambda R_2^T, V_B = V_A R_1 \Lambda^{-1} R_2^T. \quad (2.20)$$

In the later sections, we describe an algorithm to compute $Z = R_1 \Lambda R_2^T$, given $U_A V_A^T = U_B V_B^T$.

In this general situation, however, the feature vectors are arbitrary linear transformations of each other and do not preserve geometrical properties, such as pairwise distances or pairwise cosine similarity. Therefore, we define *algorithms A and B as reaching equivalent local minima if there exist a diagonal scaling matrix Σ and an orthonormal orthogonal matrix R ($R^T R = I$) such that*
$$\begin{bmatrix} U_A \\ V_A \end{bmatrix} = \begin{bmatrix} U_B \Sigma R \\ V_B \Sigma^{-1} R \end{bmatrix}.$$
 That is, $R_1 = I$ in Eq. 2.20, and (U_B, V_B) are scaled and rotated versions of (U_A, V_A) , thus preserving pairwise cosine similarity and thus, clustering properties.

Later section provides an algorithm to check whether such an equivalence is satisfied. We applied this algorithm to determine whether for the case of synthetic data (see Section 2.7.1 for the construction of such data) –where eNMF achieves globally optimal objective value, and AO-ADMM converges to a value very close to this optimum - the factor matrices computed by the two algorithms are equivalent. As shown in Table 2.2, the factor matrices are indeed scaled and rotated versions of each other, with very small residual errors. The AO-ADMM algorithm only asymptotically approaches globally optimum reconstruction error, accounting for the observed residual errors. Thus the non-negative factors generated by both eNMF and asymptotically by AO-ADMM (which takes longer time to converge than eNMF) preserve the same pair-wise feature distance metrics (in this case cosine similarity) as obtained by performing unconstrained optimization using SVD.

The case of synthetic data is an interesting special case, where the globally optimal manifold defined in Eqn. (2.7) intersects with the positive orthant, thus creating a manifold of scaled and orthogonalally equivalent global minima in the positive orthant. *This empirical result supports the intuition behind eNMF and provides evidence of how eNMF exploits the geometry of the optimization landscape to get to better solutions than SOTA competitors.* Even for real datasets (Audio), we have the globally optimal manifold intersecting with the positive orthant.

In most cases, however, the globally-optimal unconstrained manifold in Eqn. 2.7 does not intersect with the positive orthant and the local and global minima for the NMF problem lie on the boundaries of the positive orthant (i.e., with several coordinates set to zero). These minima are reached by algorithms as they iteratively approach the unconstrained optimal manifolds that lie outside of the positive orthant. Hence, it is more appropriate to look for equivalence where the orthogonal matrix R reduces to the case of a permutation matrix P . We determine that the factor matrices U_A and U_B are scaled permutations of each other by computing the cosine similarity between columns from U_A and U_B (see later sections for detailed computation procedure), and determining the underlying permutation matrix if it

exists. Once we have the permutation matrix, mapping U_A to U_B , we check whether the same permutation matrix P maps V_A to V_B . If so, then the existence of a P is numerically established.

The synthetic exact-NMF datasets provide us with a sandbox to determine such permutation equivalences, since the globally optimal factors are known (i.e. those used to construct datasets) and are sparse by construction, i.e., they lie on the boundaries of the positive orthant. In fact (see Fig. 2.3), the competing algorithms converged to global minima (i.e. attained a reconstruction error of zero for various sparsity levels) providing us with the opportunity to explore how the factors obtained by the different algorithms are related. In the top table of Table 2.3, we tracked the percentage of the columns of the factor matrices that have converged to a unique column of ground truth factors as a function of sparsity levels for the exact factorization scenario. It can be observed that when the algorithms converge to the global minimum then the factor matrices obtained from different algorithms are permuted versions of the ground truth factors.

The top table in Table 2.3 also provides an important observation regarding when an *algorithm trends towards a minimum that is equivalent to a ground truth minimum or a minimum that is achieved by another algorithm*. For example, for the sparsity level of 0.1 we find that for the eNMF algorithm, 49 (47) columns out of 50 of its factor matrix U (V) are permuted and scaled versions of the ground truth factor matrices. That is, if (U^*, V^*) is the globally optimum factor matrix pair, then there is a permutation of the 50 columns of U and V obtained by the eNMF such that for a subset of 45 columns we have: $\cos(U_i, U_{P(i)}^*) = 1 - \epsilon$ and $\cos(V_i, V_{P(i)}^*) = 1 - \epsilon$, where $\epsilon \ll 1$; in our computations we picked $\epsilon = 0.05$. In fact, if the permutation mapping is computed prior to when eNMF reaches its minimum value (in this case, 8.7, instead of 0) then a fewer number of columns of factor matrices would have been permutations of the ground truth factors. The AO-ADMM and HALS algorithms that reach a reconstruction error of 8.9 have a fewer number of converged columns. For sparsity level 0.2, both AO-ADMM and HALS are further away and hence, have a fewer number

of converged columns in their factor matrices than eNMF, which reaches global equivalent minimum and has perfect reconstruction.

This is a general phenomenon that we observed for other datasets as well, where the globally optimum factor matrices are not known. Suppose algorithm A reaches a local minimum (in almost all the datasets we tried, the eNMF algorithm reaches a local minimum as determined by KKT conditions; see Eqn 2.18) then it is often the case that the landscape around the minimum is extremely flat along most directions, causing an algorithm B to only asymptotically approach an equivalent local minimum. In such a case, we expect a large fraction of the columns of U_B and V_B to be scaled permutations of the columns in U_A and V_A . Moreover, as B is run for longer iterations, this fraction is expected to increase until the algorithm literally stalls away from the minimum. We find that most of these algorithms do not progress, even after they run for tens of thousands of iterations longer, and neither do they satisfy the KKT conditions. **In such cases, we say that B is tending towards an equivalent (TE) local minimum as A .**(Fig. 2.8)

r	SNR	eNMF vs AO-ADMM: (δ_u, δ_v)
100	28.45	$(10^{-10}, 10^{-10})$
	18.40	$(10^{-10}, 10^{-10})$
	12.44	$(10^{-10}, 10^{-10})$
	10.02	$(10^{-10}, 10^{-9})$

Table 2.2: **Equivalence of Factor Matrices (scaled and rotated versions of each other) for Synthetic datasets** (see Section VII(A) for the construction of such data): We used the algorithm in Appendix B to compute the generalized orthogonal parameters, R_1, R_2 and Λ . We verified that $R_1 \approx I$ by computing $\|R_1 - I\|_F$ for all 4 SNR levels ($\|R_1 - I\|_F \leq 10^{-6}$). We then computed the residual errors (δ_u, δ_v) , where $\delta_u = \|U_{alg} - U_{eNMF} \Lambda R_2^T\|_F$ and $\delta_v = \|V_{alg} - V_{eNMF} \Lambda^{-1} R_2^T\|_F$, and found that $(\delta_u \approx 0, \delta_v \approx 0)$. Thus, the NMF factors obtained from the AO-ADMM algorithm are scaled and rotated versions of those obtained by the eNMF algorithm. Thus, *the feature vectors retain the same pairwise cosine distances, and are equivalent from a clustering perspective.*

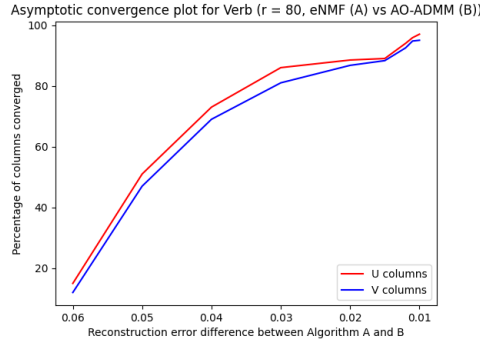


Figure 2.8: **Asymptotic/Tending Equivalence of Factor Matrices:** We plotted the percentage of columns in factor matrices returned by eNMF (Algorithm A) that are a scaled permutation of the columns in factor matrices returned by AO-ADMM (Algorithm B) as a function of the reconstruction error difference between the algorithms. As the reconstruction error difference gets smaller, more columns in the factor matrices returned by AO-ADMM converge to a scaled permutation of the columns in the factor matrices returned by eNMF. The percentage of the columns converged approaches 100% as reconstruction error approaches zero in an asymptotic manner showing the tendency of the factor matrices to approach equivalent local minima.

KKT optimality conditions [Eqn 2.18]: $(\max(\nabla W \odot W), \sum_{i,j}(\nabla W \odot W)_{ij})$

Dataset	r	eNMF	HALS	AO-ADMM	ALS
Face	5	(0, 0)	(0.06, 5.8)	(0.087, 5.2)	(0, 4.7)
	10	(0, 0)	(0.01, 1.1)	(0, 0)	(0, 1.1)
	15	(0, 0)	(3, 220)	(0, 0)	(0.06, 158)
Verb	10	(0, 0)	(0, 0)	(0, 0)	(0, 0)
	20	(0, 0)	(0, 0)	(0, 0)	(0, 0)
	40	(0, 0)	(0, 0.2)	(0, 0)	(0, 0)
	80	(0, 0)	(0, 0)	(0, 0)	(0, 0)
	100	(0, 0)	(0, 0.2)	(0, 0)	(0.8, 42)

Table 2.4: **eNMF satisfies the KKT conditions for almost all latent dimensions and across both the datasets:** We tabulated the tuple (δ_W, σ_W) , where $\delta_W = \max(\nabla W \odot W)$ and $\sigma_W = \sum_{i,j}(\nabla W \odot W)_{ij}$. As before W is defined as the stacked NMF factors i.e. $W = \begin{bmatrix} U \\ V \end{bmatrix}$. eNMF satisfies the KKT conditions ($\delta_W \approx 0, \sigma_W \approx 0$) for all datasets and across all latent dimensions.

Having verified the observation on the synthetic dataset, we wanted to validate it on real datasets. We checked for permutations between the factors obtained from equal-error constrained experiments on verb and face dataset. That is, we obtain an eNMF solution, and then let the other algorithms run until the error either becomes the same as that of eNMF or their reconstruction error doesn't improve for a long enough time. As listed in Table 2.4, the KKT conditions show that our eNMF algorithm reaches a local minimum for all data sets and almost all latent dimensions (for Face $r = 20, 25$ we are very close to a local minima). For permutation checking purpose, we fix the algorithm A to be eNMF and then compare its factors with the factors obtained from the competing algorithms. In Table 2.3 we compared the percentage of the columns of the factor matrices that have converged to a unique column of the eNMF factors along with the error ratio as a function of the latent

dimension for the Verb and Face dataset. From the table , we can make the following claim: *the factors from the NMF algorithms converge to a pair of factor matrices that are unique modulo permutation and scaling only at the local minima (error ratio = 1), and even a small deviation from the minima leads to a large reduction in the number of columns that have converged to a unique column of the eNMF..*

Only in one instance (see Table 2.3, Face ($r = 10$)) we found that the AO-ADMM algorithm converged to a non-equivalent minimum with higher reconstruction error. That is, the factor matrices do not trend towards that of eNMF algorithm, and it satisfies the KKT conditions.

2.8 Concluding remarks

We developed a general framework for the *NMF* problem that performs significantly better than the existing ones: From Fig. 2.4 and Fig. 2.5, it is evident that, eNMF’s performance is significantly better than the existing algorithms on both the reconstruction error and run time metrics with the performance gap widening for lower SNR levels (more noise). Also, from Fig. 2.5 it can be observed that eNMF’s closest competitors (AO-ADMM and HALS) have a very slow convergence as they approach the local minimum, depicted by its very long tail. In the context of real datasets, eNMF also achieves the lowest reconstruction error as shown in Table 2.1. The slower convergence of baseline algorithms when approaching the local minima gives eNMF a significant advantage in terms of run time metrics as shown in Fig. 2.6.

We also performed empirical studies on the factor matrices that are computed by different algorithms for the same NMF problem. For examples where ground-truth factors are known, such as synthetic data for exact-NMF problems, we found that eNMF can reach almost zero reconstruction errors for most settings and yield factor matrices whose columns are permuted and scaled versions of the ground-truth factor columns. The factor matrices,

however, only partially converge to ground-truth factors – i.e. only a subset of the columns are permuted versions of the ground-truth factors– when the reconstruction error is not zero. For general NMF problems we found that the factor matrices are sensitive to the reconstruction errors achieved by an algorithm, and even a slight increase in reconstruction error can change the factors significantly. Since the factors are used as features for downstream tasks, our empirical results show that the algorithms used to perform NMF can lead to varying performance in classification tasks downstream. In particular, our numerical results indicate that to assure stable factor matrices it is important to reach local minima with low reconstruction errors. In this regard, eNMF seems to be the algorithm of choice, given its faster convergence and lower reconstruction errors.

These extensive numerical results imply that numerical evaluations of the relative performance of NMF algorithms should consider both measures: the run-time complexity required for achieving the same reconstruction-errors, and the reconstruction errors achieved for the same run-time. Moreover, as an algorithm approaches a local minimum, even a small difference in the reconstruction error could yield factor matrices that are geometrically non-equivalent, and thus potentially affect performance of downstream tasks.

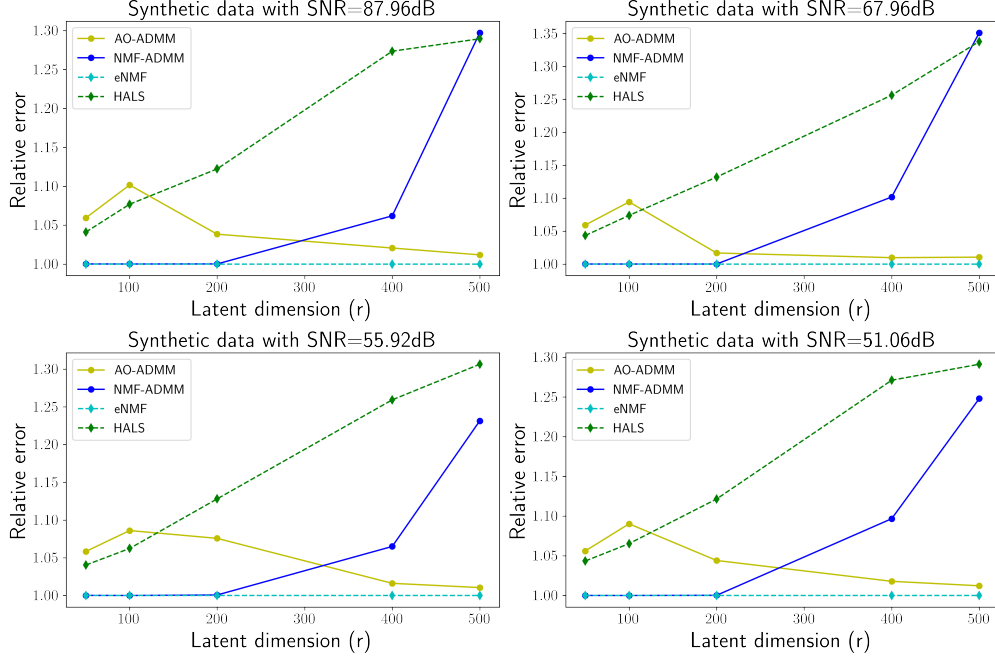


Figure 2.4: **Equal-time constrained NMF reconstruction errors (Synthetic datasets):** Relative Reconstruction Error (RE) $\frac{\|X - U_{\text{NMF}} V_{\text{NMF}}^T\|_F}{\|X - U_{\text{SVD}} V_{\text{SVD}}^T\|_F}$ is plotted as a function of the number of latent dimensions (r) for various SNR levels. *eNMF* achieves the unconstrained global minimum for all latent dimensions and across all the four SNR levels. Among its competitors, only NMF-ADMM attains the global minimum for lower latent dimensions. For higher latent dimensions, the gap in relative error of eNMF and its closest competitor AO-ADMM widens for lower SNR levels (more noise). **Note:** For a zoomed in picture of the scenario, we have dropped the curves corresponding to the algorithms with a high relative error.

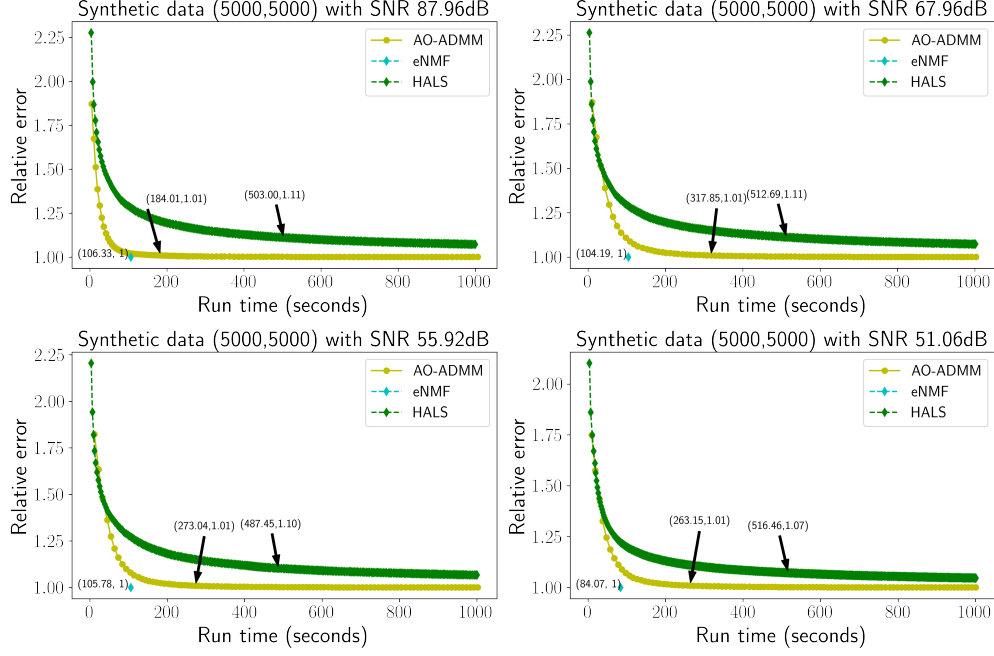


Figure 2.5: **Convergence speed (Synthetic datasets): Algorithm 1 of eNMF directly finds globally optimal solutions** : Relative error $= \frac{\|X - U_{\text{NMF}} V_{\text{NMF}}^T\|_F}{\|X - U_{\text{SVD}} V_{\text{SVD}}^T\|_F}$ is plotted as a function of run time (secs) for various SNR levels. *The orthogonal part of eNMF (Algorithm 1) finds non-negative factors by itself, thus achieving globally optimal solutions for all SNR levels.* Moreover, eNMF converges to the global minimum at the fastest rate (compared to SOTA algorithms) for all SNR levels. As the noise level increases (lower SNR), the gap in the convergence rates of eNMF and AO-ADMM (its closest competitor) widens. **Note:** In order to get convergence plots where the relative performance of top algorithms can be clearly visualized, we have dropped the convergence curves of slower algorithms.

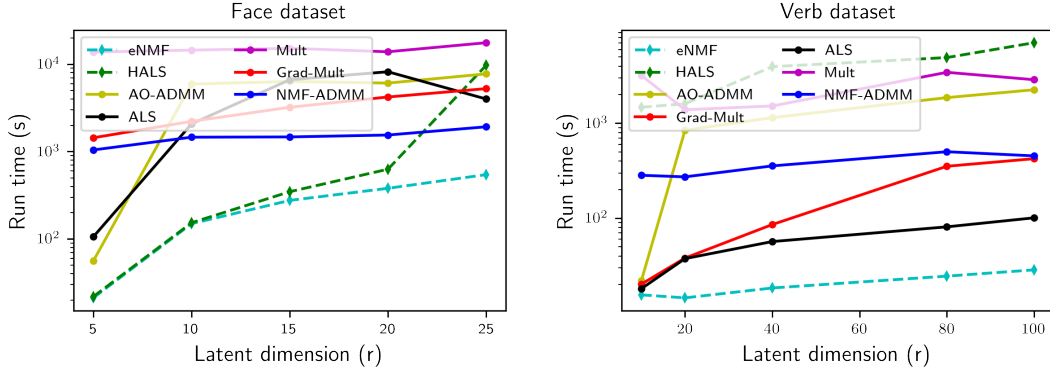


Figure 2.6: **Equal-error constrained time complexity (Real datasets)**: Run time is plotted on a log scale as a function of the number of latent dimensions (r). *eNMF performs significantly better than in its competitors in terms of time complexity with the margin widening for higher latent dimensions.*

Reconstruction error comparison for various latent dimensions (MovieLens dataset)

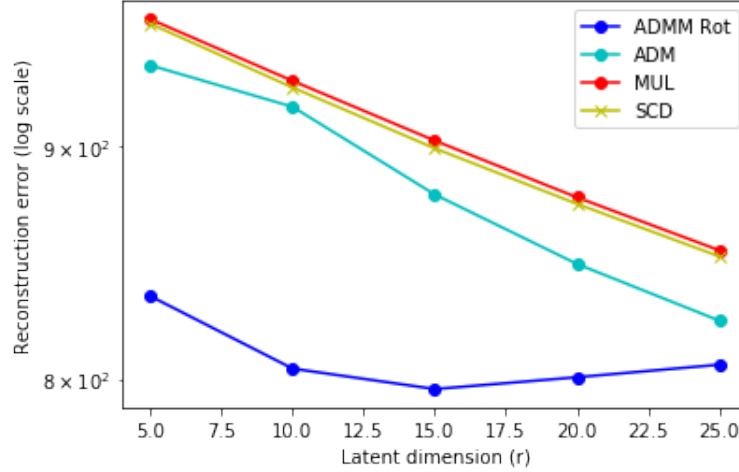


Figure 2.7: **Equal-time constrained NMC errors**: Relative reconstruction error (RE) $\frac{\|P_{\Omega}(X-UV^T)\|_F}{\|P_{\Omega}X\|_F}$ and Relative prediction error $\frac{\|X_{test}-UV^T\|_F}{\|X_{test}\|_F}$ is plotted as a function of the number of latent dimensions (r). *eNMF achieves the lowest reconstruction and prediction error for all latent dimensions.*

Permutation results: (% of converged U columns, % of converged V columns, reconstruction error difference, equivalence of local minima (**E**: equivalent local minima, **TE**: tending towards equivalent local minima, **NE**: non-equivalent local minima))

Sparsity	eNMF vs gt	HALS vs gt	AO-ADMM vs gt	ALS vs gt	Mult vs gt	Grad-Mult vs gt
0.1	(98,95,8.7,TE)	(96,93,8.9,TE)	(93,92,8.9,TE)	(10,7,36.5,TE)	(8,6,36.5,TE)	(5,2,38.6,TE)
0.2	(100,100,0,E)	(29,26,31,TE)	(35,33,31,TE)	(18,16,46,TE)	(0,0,53.7,TE)	(0,0,51.4,TE)
0.3	(100,100,0,E)	(100,100,0,E)	(100,100,0,E)	(25,20,58,TE)	(19,15,68.5,TE)	(23,22.5,62.2,TE)
0.4	(100,100,0,E)	(100,100,0,E)	(100,100,0,E)	(100,100,0,E)	(91,89,48,TE)	(7,6,115,TE)
0.5	(100,100,0,E)	(100,100,0,E)	(100,100,0,E)	(100,100,0,E)	(100,100,0.4,E)	(8,5,113,TE)

Dataset	r	HALS vs eNMF	AO-ADMM vs eNMF	ALS vs eNMF
Face	5	(100,100,0,E)	(100,100,0,E)	(100,100,0,E)
	10	(100,100,0,E)	(18,15,32.3,NE)	(100,100,0,E)
	15	(75,59,0.25,TE)	(100,100,0,E)	(75,59,0.25,TE)
Verb	10	(91,86,0.01,TE)	(100,100,0,E)	(93,89,0.0096,TE)
	20	(97,97,0.004,TE)	(83,82,0.01,TE)	(75,73,0.016,TE)
	40	(93,92,0.0004,TE)	(85,84.75,0.01,TE)	(86,85.5,0.009,TE)
	80	(75.25,73.25,0.069,TE)	(94.5,92.25,0.01,TE)	(87,85.75,0.02,TE)
	100	(98,96.65,0.00003,TE)	(86,83,0.01,TE)	(75,71,2.79,TE)

Table 2.3: **Numerical Exploration of Permutation Equivalence of Factor matrices:** We tabulated the quartet (u_B, v_B, e_d, e_m) . From the top table for the exact factorization scenario, we can see that factor matrices are permutations of each other at the global minimum (for sparsity 0.3, 0.4, 0.5 where global minimum is obtained by several algorithms). From the bottom table for the real dataset scenarios, it can be observed that the factors from the competing algorithms are approaching to a permuted version of the eNMF factors ($u_B \approx 100, v_B \approx 100$) in an asymptotic manner due to the slow convergence of the competing algorithms. As the numerical results show, the factor matrices can be very different even when the algorithms are very close to a local minimum. Thus while reconstruction errors might be similar, the factor matrices might be very different; which might affect performance in downstream tasks.

2.9 Exact NMF solution

$$\begin{aligned}
& \underset{R_1, R_2, \Lambda}{\text{minimize}} && \sum_{i=1}^n \sum_{j=1}^r h((U^\star R_1 \Lambda R_2)_{ij}) + \\
& && \sum_{i=1}^m \sum_{j=1}^r h((V^\star R_1 \Lambda^{-1} R_2)_{ij}) \\
& \text{subject to} && R_1^T R_1 = I, R_2^T R_2 = I, \Lambda \geq 0
\end{aligned} \tag{2.21}$$

where $h(q) = \max\{0, -q\}$. We will solve (2.21) using ADMM. In order to aid the ADMM formulation, we introduce the following notation:

$$W_{svd} = \begin{bmatrix} U^\star \\ V^\star \end{bmatrix} \in \mathbb{R}^{(m+n) \times r}, W_1, W_2, W_3$$

where W_1, W_2, W_3 are the splitting variables. Then, with the above notation we have the following equivalent optimization problem

$$\begin{aligned}
& \underset{W_1, W_2, W_3, R_1, R_2}{\text{minimize}} && \sum_{i=1}^{n+m} \sum_{j=1}^r h([W_3]_{ij}) \\
& \text{subject to} && W_1 = \begin{bmatrix} W_{11} \\ W_{12} \end{bmatrix} = \begin{bmatrix} U^\star \\ V^\star \end{bmatrix} R_1 = W_{svd} R_1 \\
& && W_2 = \begin{bmatrix} W_{21} \\ W_{22} \end{bmatrix} = \begin{bmatrix} W_{11} \Lambda \\ W_{12} \Lambda^{-1} \end{bmatrix} \\
& && W_3 = W_2 R_2 \\
& && R_1^T R_1 = I \\
& && R_2^T R_2 = I
\end{aligned} \tag{2.22}$$

The ADMM for (2.22) is given by Algorithm (5)

Algorithm 5 ADMM for (2.22)

Input: U^*, V^*

Stack U^*, V^* vertically to get W_{svd}

Repeat until convergence:

$$(W_1^{(k+1)}, W_3^{(k+1)}) \leftarrow \arg \min_{W_1, W_3} f_1(W_1, W_3)$$

$$(W_2^{(k+1)}, R_1^{(k+1)}) \leftarrow \arg \min_{W_2, R_1} f_2(W_2, R_1)$$

$$(R_2^{(k+1)}, \Lambda^{(k+1)}) \leftarrow \arg \min_{R_2, \Lambda} f_3(R_2, \Lambda)$$

$$M^{(k+1)} \leftarrow M^{(k)} + W_1^{(k+1)} - W_{svd} R_1^{(k+1)}$$

$$\begin{bmatrix} N_1 \\ N_2 \end{bmatrix}^{(k+1)} \leftarrow \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}^{(k)} + \begin{bmatrix} W_{21} \\ W_{22} \end{bmatrix}^{(k+1)} - \begin{bmatrix} W_{11} \Lambda \\ W_{12} \Lambda^{-1} \end{bmatrix}^{(k+1)}$$

$$P^{k+1} \leftarrow P^k + W_3^{(k+1)} - W_2^{(k+1)} R_2^{(k+1)}$$

$$k \leftarrow k + 1$$

Output: R_1, R_2, Λ

In algorithm (5), we have three subproblems with the following cost functions:

$$\begin{aligned}
f_1(W_1, W_3) &= \sum_{i=1}^n \sum_{j=1}^r h([W_3]_{ij}) + \frac{\rho}{2} \|W_1 - W_{svd} R_1^{(k)} + \\
&\quad M^{(k)}\|^2 + \frac{\gamma}{2} \|W_{21}^{(k)} - W_{11} \Lambda^{(k)} + N_1^{(k)}\|^2 + \\
&\quad \frac{\gamma}{2} \|W_{22}^{(k)} - W_{12} \Lambda^{(k)-1} + N_2^{(k)}\|^2 + \\
&\quad \frac{t}{2} \|W_3 - W_2^{(k)} R_2^{(k)} + P^{(k)}\|^2 \\
f_2(W_2, R_1) &= \frac{\rho}{2} \|W_1^{(k+1)} - W_{svd} R_1 + M^{(k)}\|^2 + \\
&\quad \frac{\gamma}{2} \|W_{21} - W_{11}^{(k+1)} \Lambda^{(k)} + N_1^{(k)}\|^2 + \\
&\quad \frac{\gamma}{2} \|W_{22} - W_{12}^{(k+1)} \Lambda^{(k)-1} + N_2^{(k)}\|^2 + \\
&\quad \frac{t}{2} \|W_3^{(k+1)} - W_2 R_2^{(k)} + P^{(k)}\|^2 \\
f_3(R_2, \Lambda) &= \frac{\gamma}{2} \|W_{21}^{(k+1)} - W_{11}^{(k+1)} \Lambda + N_1^{(k)}\|^2 + \\
&\quad \frac{\gamma}{2} \|W_{22}^{(k+1)} - W_{12}^{(k+1)} \Lambda^{-1} + N_2^{(k)}\|^2 + \\
&\quad \frac{t}{2} \|W_3^{(k+1)} - W_2^{(k+1)} R_2 + P^{(k)}\|^2
\end{aligned}$$

2.9.1 Updates for W_1 and W_3

$f_1(W_1, W_3)$ is separable in W_1 and W_3 , and hence they can be updated in parallel by solving 2 subproblems in parallel.

2.9.1.1 Updates for W_3

We can update each entry of W_3 independently. The updates can be obtained by solving a simple quadratic problem

$$[W_3]_{ij}^{(k+1)} = \arg \min_{[W_3]_{ij}} h([W_3]_{ij}) + \frac{t}{2} ([W_3]_{ij} - [W_2^{(k)} R_2^{(k)} - P^{(k)}]_{ij})^2.$$

The updates are

$$[W_3]_{ij}^{(k+1)} = \begin{cases} b_{ij} & b_{ij} > 0 \\ 0 & -\frac{1}{t} \leq b_{ij} \leq 0 \\ b_{ij} + \frac{1}{t} & b_{ij} < -\frac{1}{t} \end{cases}$$

where $b_{ij} = [W_2^{(k)} R_2^{(k)} - P^{(k)}]_{ij}$.

2.9.1.2 Updates for W_1

We can update the blocks W_{11} and W_{12} in parallel by solving independent subproblems.

$$\begin{aligned} W_{11}^{(k+1)} &= \arg \min_{W_{11}} \frac{\rho}{2} \|W_{11} - U^* R_1^{(k)} + M_1^{(k)}\|^2 + \\ &\quad \frac{\gamma}{2} \|W_{11} \Lambda^{(k)} - W_{21}^{(k)} - N_1^{(k)}\|^2 \\ W_{12}^{(k+1)} &= \arg \min_{W_{12}} \frac{\rho}{2} \|W_{12} - V^* R_1^{(k)} + M_2^{(k)}\|^2 + \\ &\quad \frac{\gamma}{2} \|W_{12} \Lambda^{(k)} - W_{22}^{(k)} - N_2^{(k)}\|^2 \end{aligned}$$

The subproblems are separable in the elements of W_{11} and W_{12} and can be solved to give the following updates:

$$\begin{aligned} [W_{11}]_{ij}^{(k+1)} &= \frac{\rho[U^* R_1^{(k)} - M_1^{(k)}]_{ij} + \gamma \lambda_j [W_{21}^{(k)} + N_1^{(k)}]_{ij}}{\rho + \gamma \lambda_j^2} \\ [W_{12}]_{ij}^{(k+1)} &= \frac{\rho[V^* R_1^{(k)} - M_2^{(k)}]_{ij} + \frac{\gamma}{\lambda_j} [W_{22}^{(k)} + N_2^{(k)}]_{ij}}{\rho + \frac{\gamma}{\lambda_j^2}} \end{aligned}$$

where λ_j is the j^{th} diagonal element of Λ .

2.9.2 Updates for W_2 and R_1

$f_2(W_2, R_1)$ is separable in W_2 and R_1 , and hence they can be updated in parallel by solving 2 subproblems in parallel.

2.9.2.1 Updates for R_1

We can update R_1 by solving a non-convex problem known as Orthogonal Procrustes Problem. If we define $B = W_1^{(k+1)} + M^{(k)}$, then the update for R_1 is given by

$$R_1^{(k+1)} = CD^T, \quad (2.23)$$

where C and D are obtained from the SVD $W_{svd}^T B = C\Sigma D^T$.

2.9.2.2 Updates for W_2

We can update the blocks W_{21} and W_{22} in parallel by solving independent subproblems:

$$\begin{aligned} W_{21}^{(k+1)} &= \arg \min_{W_{21}} \frac{\gamma}{2} \|W_{21} - W_{11}^{(k+1)} \Lambda^{(k)} + N_1^{(k)}\|^2 + \\ &\quad \frac{t}{2} \|W_{31}^{(k+1)} - W_{21} R_2^{(k)} + P_1^{(k)}\|^2 \\ W_{22}^{(k+1)} &= \arg \min_{W_{21}} \frac{\gamma}{2} \|W_{22} - W_{12}^{(k+1)} \Lambda^{(k)-1} + N_2^{(k)}\|^2 + \\ &\quad \frac{t}{2} \|W_{32}^{(k+1)} - W_{22} R_2^{(k)} + P_2^{(k)}\|^2 \end{aligned}$$

The subproblems can be solved to give the following updates:

$$\begin{aligned} W_{21}^{(k+1)} &= \frac{\gamma W_{11}^{(k+1)} \Lambda^{(k)} - \gamma N_1^{(k)} + t(W_{31}^{(k+1)} + P_1^{(k)}) R_2^{(k)T}}{\gamma + t} \\ W_{22}^{(k+1)} &= \frac{\gamma W_{12}^{(k+1)} \Lambda^{(k)-1} - \gamma N_2^{(k)} + t(W_{32}^{(k+1)} + P_2^{(k)}) R_2^{(k)T}}{\gamma + t} \end{aligned}$$

2.9.3 Updates for R_2 and Λ

$f_3(R_2, \Lambda)$ is separable in R_2 and Λ , and hence they can be updated in parallel by solving 2 subproblems in parallel.

2.9.3.1 Updates for R_2

We can update R_2 by solving a non-convex problem known as Orthogonal Procrustes Problem. If we define $B = W_3^{(k+1)} + P^{(k)}$, then the update for R_2 is given by

$$R_2^{(k+1)} = CD^T,$$

where C and D are obtained from the SVD $W_2^{(k+1)T} B = C\Sigma D^T$.

2.9.3.2 Updates for Λ

Since we assume $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, so we can update Λ by solving for the λ_i 's in parallel.

$$\begin{aligned} \lambda_i^{(k+1)} = \arg \min_{\lambda_i} & \frac{\gamma}{2} \|W_{21}^{(k+1)}(:, i) - \lambda_i W_{11}^{(k+1)}(:, i) + N_1^{(k+1)}(:, i)\|^2 \\ & + \frac{\gamma}{2} \|W_{22}^{(k+1)}(:, i) - \frac{W_{12}^{(k+1)}(:, i)}{\lambda_i} + N_2^{(k+1)}(:, i)\|^2 \end{aligned}$$

The above is an unconstrained optimization problem and can be solved by setting the gradient with respect to λ_i to zero. It gives rise to the quartic equation,

$$\begin{aligned} & \lambda_i^4 W_{11}^{(k+1)}(:, i)^T W_{11}^{(k+1)}(:, i) - \lambda_i^3 (W_{21}^{(k+1)}(:, i) + \\ & N_1^{(k+1)}(:, i))^T W_{11}^{(k+1)}(:, i) + \lambda_i (W_{22}^{(k+1)}(:, i) + \\ & N_2^{(k+1)}(:, i))^T W_{12}^{(k+1)}(:, i) - W_{12}^{(k+1)}(:, i)^T W_{12}^{(k+1)}(:, i) = 0 \end{aligned}$$

which can be solved efficiently for λ_i .

2.10 Generalized orthogonal transformation

Given eNMF and another algorithm that reaches the same reconstruction error, i.e., $U_{alg} V_{alg}^T = U_{eNMF} V_{eNMF}^T$. We want to find R_1 , R_2 , and Λ such that the matrix factors from different

algorithms are related by the following generalized orthogonal transformation

$$U_{alg} = U_{eNMF} R_1 \Lambda R_2^T$$

$$V_{alg} = V_{eNMF} R_1 \Lambda^{-1} R_2^T$$

and such a transformation always exists [Vavasis, 2009] [Stewart, 2001]. Following a similar approach as in Section (2.3.1), we can find R_1, R_2 and Λ by computing the SVD of $Q\Gamma^{-1}P^T U_{alg}$ where:

$$SVD(U_{eNMF}) = P\Gamma Q^T$$

2.11 Checking permutation invariance between two factor matrices

Let A and B be two NMF algorithms returning factor matrices (U_A, V_A) and (U_B, V_B) respectively. We identify the subset of columns in U_B that are scaled permutations of the columns in U_A using the cosine-similarity measure. Assuming U_A and U_B have r columns, we seek to find the subset

$$A_{candidates} \subseteq \{1, 2, 3, \dots, r\}$$

$$B_{candidates} \subseteq \{1, 2, 3, \dots, r\}$$

such that

$$\cos(U_A^{(i)}, U_B^{(P_i)}) = 1 \tag{2.24}$$

where $i \in A_{candidates}$ and $P_i \in B_{candidates}$. If Eqn 2.24 is satisfied with equality, then we are done with identifying the subset of columns in U_A that have mapped to columns in U_B . However, due to rounding error and numerical instability often is the case where Eqn 2.24 is not satisfied with equality. To circumvent this, we relax the criterion in Eqn 2.24 to fall within some bound

$$\cos(U_A^{(i)}, U_B^{(P_i)}) \geq 1 - \epsilon$$

where ϵ is a relaxation parameter with value close to zero (for our experiments we used $\epsilon = 0.05$). In addition to the relaxed criterion, we also perform some additional cross cosine-similarity computations to validate the identification of the subset.

Suppose we construct the following square submatrices

$$\begin{aligned}[M]_{ij} &= \cos(U_{Asub}[:, i], U_{Asub}[:, j]) \\ [N]_{ij} &= \cos(U_{Bsub}[:, i], U_{Bsub}[:, j]) \\ [MN]_{ij} &= \cos(U_{Asub}[:, i], U_{Bsub}[:, j]),\end{aligned}$$

where

$$\begin{aligned}U_{Asub} &= U_A[:, A_{candidates}] \\ U_{Bsub} &= U_B[:, B_{candidates}]\end{aligned}$$

. We can think of M and N as the self-similarity matrices and MN as the cross-similarity matrix. We sort the entries in each column of the above matrices in descending order and perform the following steps for each column of M :

- Compare the entries in the column of M with the entries of the corresponding column in N and MN rounded to the nearest one decimal place.
- If more than 90% of the entries in the column of M match with the entries of column N and MN , then we keep the corresponding index in $A_{candidates}$ and $B_{candidates}$. Else, we delete the corresponding index in $A_{candidates}$ and $B_{candidates}$.

After running the above steps through all the columns of M , we get the pruned arrays $A_{candidates}$ and $B_{candidates}$ which contains the index of columns in U_A and U_B that are permutations of each other based on cosine-similarity.

CHAPTER 3

Optimizing detection and deep learning-based classification of pathological high-frequency oscillations in epilepsy

3.1 Introduction

Complete resection of the epileptogenic zone (EZ: areas of the brain responsible for generating seizures) is necessary to achieve freedom from seizures in drug-resistant focal epilepsy [Rosenow and Lüders, 2001]. Discovering a biomarker that can accurately map the spatial extent of EZ will be significant. Research on epilepsy, involving both humans and animals, has proposed that intracranially recorded interictal high-frequency oscillations (HFOs) on EEG is a promising spatial neurophysiological biomarker of EZ [Bragin et al., 1999c, Bragin et al., 1999b, Worrell et al., 2008, Zijlmans et al., 2009]. There have been many studies in the past showing the correlation between postoperative seizure freedom and the resection of brain regions generating HFOs [Jacobs et al., 2010a, Wu et al., 2010, Akiyama et al., 2011, Nariai et al., 2011, van’t Klooster et al., 2015]. However, a recent multi-institutional prospective study was unable to reproduce the findings [Jacobs et al., 2018], and a clinical trial that was recently published could not demonstrate that HFO-guided resection during electrocorticography (ECoG) resulted in better seizure outcomes than standard spike-guided resection [Zweiphenning et al., 2022a].

At this stage, although HFOs show promise, it is unclear whether it can be used to guide

epilepsy surgery. Some technical issues with HFO analysis include the following: (1) The standard detection methods of HFOs have not been established [Remakanthakurup Sindhu et al., 2020, Wong et al., 2021]. Currently, there are various methods available for HFO detection, including the STE [Staba et al., 2002], Short Line Length (SLL) [Gardner et al., 2007], and Hilbert transform [Crépon et al., 2010] methodologies. The recent and widely used method is the MNI detector [Zelmann et al., 2009], in which the HFO detection incorporates information from previously detected baselines to simulate expert HFO analysis to enhance its detection sensitivity. However, although the STE detector is simple and commonly used, it is still unclear whether this simpler method is less effective in detecting HFOs than the more sophisticated MNI detector. (2) The presence of physiological HFOs complicates the interpretation of the results [Frauscher et al., 2017, Frauscher et al., 2018]. Currently, no methods are available to detect pathological HFOs, while rejecting physiological HFOs and artifacts. Visual analysis by a human expert is commonly performed, but it is time-consuming and has limitations in inter-rater reliability [Spring et al., 2017, Nariai et al., 2018]. Although it works well for artifact rejection, human experts cannot confidently annotate each event as pathological or physiological except for recognizing HFOs with or without spikes. Efforts to selectively detect pathological HFOs have been attempted, including automated detection of spikes along with HFOs [Weiss et al., 2016, Guth et al., 2021] and crossfrequency coupling analysis [Nonoda et al., 2016, Motoi et al., 2019, Kuroda et al., 2021]. However, the evidence of its efficacy in differentiating pathological and physiological HFOs is still limited. Furthermore, it remains uncertain whether the MNI detector is more sensitive than the simple STE-based detection in detecting pathological HFOs.

In recent work, we developed a novel weakly-supervised DL algorithm, using channel resection status as DL labels, to characterize the pathological HFOs generated by EZ [Zhang et al., 2022b]. In another related work, we utilized a similar framework, with functional cortical mapping results as DL labels, to characterize the physiological HFOs generated by the eloquent cortex (EC) [Zhang et al., 2022a]. However, we applied the DL-based classifiers

after we detected HFOs using the STE detector. It is unknown whether we had increased sensitivity in detecting pathological or physiological HFOs before applying the classifiers with the MNI detector or a combination of both.

The study presented in this chapter aimed to provide insights into what HFO detection methods should be used to maximize the sensitivity of pathological HFOs by meticulously characterizing HFOs detected by the STE and MNI detectors. We also engineered the Union detector, which encompassed events detected by either STE or MNI detectors to further enhance the sensitivity of the detection. Then, the detected HFOs from each detector were evaluated for their association with spikes and time–frequency characteristics. Finally, our previously developed novel DL-based classifier was applied to determine what detection methods would provide the most sensitive detection of pathological HFOs, and what HFO detection method would provide the most accurate postoperative seizure outcomes.

3.2 Methods

3.2.1 Patient Cohort

This study employed a retrospective cohort design. We consecutively recruited children (under the age of 21) who were suffering from medically refractory epilepsy—characterized by monthly or more frequent seizures and the inefficacy of more than three anti-seizure medications—from the Pediatric Epilepsy Program at UCLA. These children had intracranial electrodes implanted for epilepsy surgery planning and subsequent cortical resection between August 2016 and August 2018. Those who underwent a diagnostic stereo-EEG evaluation, which wasn’t intended for resective surgery, were excluded from the study.

3.2.2 Standard protocol approvals, registrations, and patient consent

The UCLA’s institutional review board approved the involvement of human subjects and dispensed with the necessity for written informed consent. All the tests were regarded as clinically relevant for patient care, and all the EEG data used in this retrospective study was anonymized before it was extracted and analyzed. This study was not classified as a clinical trial and was not registered in any public registry.

3.2.3 Patient evaluation

All children referred during the study period, who were suffering from medically refractory epilepsy, underwent a systematic pre-surgery evaluation. This included inpatient video-EEG monitoring, high-resolution brain magnetic resonance imaging (MRI) of 3.0 Tesla, and 18 fluoro-deoxyglucose positron emission tomography (FDG-PET), coupled with MRI-PET co-registration. The boundaries and scope of the resections were primarily based on the seizure onset zone (SOZ), which is clinically identified as the areas initially showing continuous rhythmic waveforms at the beginning of typical seizures. In certain instances, to avoid causing unacceptable neurological damage, the seizure onset zones were not fully resected.

3.2.4 Subdural electrode placement

Surgical implantation was performed for macroelectrodes, which included platinum grid electrodes with an intercontact distance of 10 mm, and depth electrodes, made of platinum with a 5 mm intercontact distance. The total number of electrode contacts for each patient ranged from 40 to 128, with a median of 96 contacts. The positioning of these intracranial electrodes was primarily guided by the results of scalp video-EEG recording and neuroimaging studies. All electrode plates were sewn to neighboring plates, the dura mater’s edge, or both, in order to minimize the movement of subdural electrodes post-placement.

3.2.5 Acquisition of three-dimensional (3D) brain surface images

We collected high-resolution 3D magnetization-prepared rapid acquisition with gradient echo (MPRAGE) T1-weighted images of the whole head preoperatively. Utilizing post-implant computed tomography (CT) images, we created a FreeSurfer-based 3D surface image with the locations of the electrodes directly identified on the brain surface [Zhang et al., 2022b]. Furthermore, digital photographs were captured intraoperatively before the dura was closed to improve the spatial precision of electrode placement on the 3D brain surface. When the area was re-exposed for resective surgery, we visually verified that the electrodes had remained in place, matching the digital photo taken during the electrode implantation surgery.

3.2.6 Intracranial EEG (iEEG) recordings

The intracranial EEG (iEEG) data was recorded with a digital sampling frequency of 2,000 Hz using the Nihon Kohden Systems (Neurofax 1100A, Irvine, California, USA). For each participant, a consistent 90-minute EEG segment was chosen, ensuring it was at least two hours before or after any seizures, prior to tapering of anti-seizure medications, and before cortical stimulation mapping. This mapping usually took place two days post-implant. All the iEEG data used in the study were extracted from the clinical EEG recording.

3.2.7 Automated detection of high frequency oscillations

For the HFO analysis, a customized average reference was applied, which involved excluding electrodes that exhibited significant artifacts [Nariai et al., 2019]. Two sets of candidate interictal HFOs were identified using the STE detector [Staba et al., 2002] and the MNI detector [Zelmann et al., 2009]. HFOs were defined as spontaneous EEG patterns in the frequency range of 80– 500 Hz, consisting of at least four oscillations that can be distinguished from the background. The STE detector [Staba et al., 2002] uses a sliding window approach to detect segments in the EEG signal as candidate interictal HFOs. The root mean

square (RMS) energy of the segment under consideration is compared with a global energy threshold that is computed from the sequence of the RMS values of the entire EEG signal. In contrast, the MNI detector [Zelmann et al., 2009] takes a more localized approach and detects HFO segments by comparing them with the surrounding baseline activity. The MNI detector consists of a baseline detector block and an energy-based event detector block. The baseline detector block identifies the baseline segments in the EEG, and the event detector block identifies the HFO segments using the energy threshold computed using the baseline segments. The detected baselines characterizes the background activity and plays an important role in channels where a large number of interictal epileptiform discharges (IED) are present. From this point onwards, we will refer to the set of candidate interictal HFOs detected by the STE and the MNI detector as STE set and MNI set, respectively. This study used the Matlab code and their default detector settings on RippleLab software [Navarrete et al., 2016].

3.2.8 Creation of each HFO set and subsequent training of deep learning-based classification (Fig. 3.1: Study workflow)

3.2.8.1 Jaccard similarity between two detected HFO events (Fig. 3.1(A))

Assume there are two detected HFO events, represented by a tuple of start and end times, (s_1, e_1) and (s_2, e_2) . Without loss of generality, we can assume that $s_1 < s_2$. Then, we define the Jaccard similarity between the two detected HFO events $d(s_1, s_2, e_1, e_2)$ in the following manner (see the study workflow):

- If $e_2 < e_1$, the detected HFO_2 is a subset of the detected HFO_1 , then we set the Jaccard similarity $d(s_1, s_2, e_1, e_2) = 1$
- If $s_2 > e_1$, the two detected HFOs are disjoint, then we set the Jaccard similarity $d(s_1, s_2, e_1, e_2) = 0$

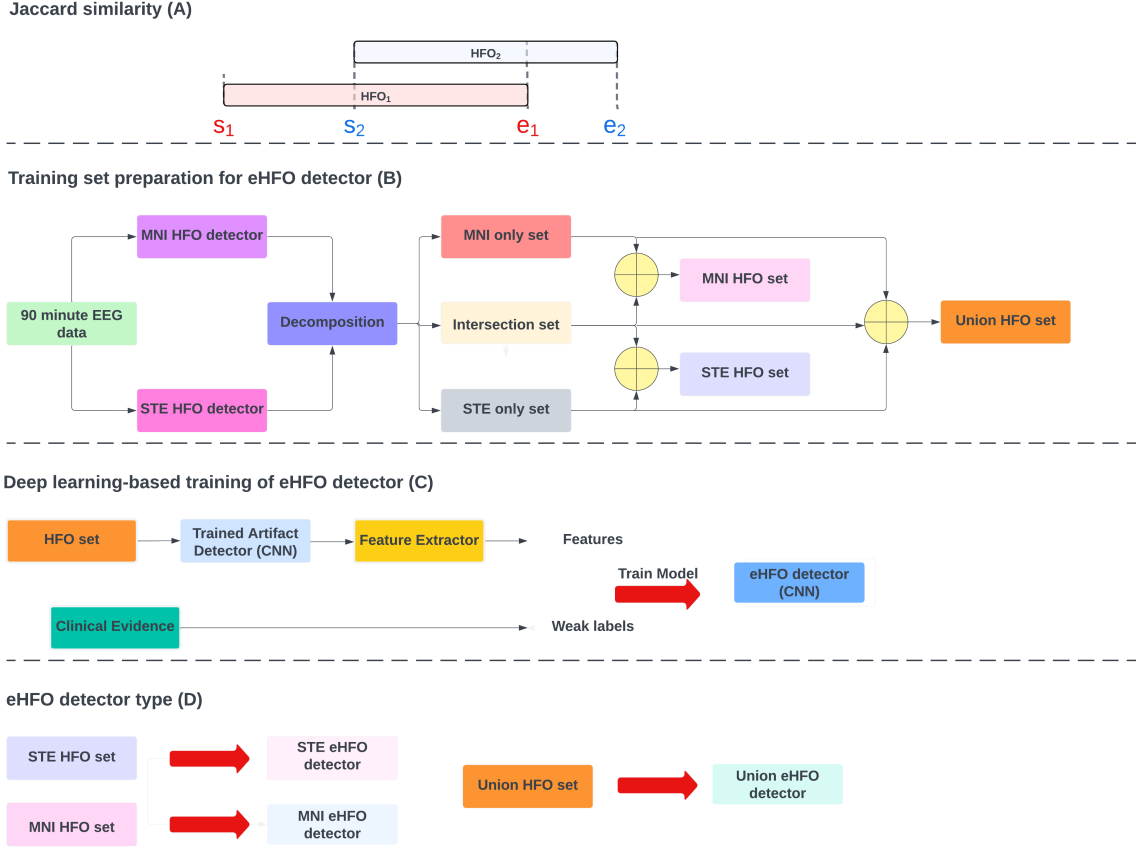


Figure 3.1: **Union eHFO detector Workflow**

- If $e_2 > e_1$, and $e_1 > s_2$, the two detected HFOs have partial overlap, and then we set the Jaccard similarity $d(s_1, s_2, e_1, e_2) = \frac{(e_1 - s_2)}{(e_2 - s_2)}$

3.2.8.2 Decomposition of STE and MNI set into three disjoint sets (Fig. 3.1(B))

We used the Jaccard similarity metric to decompose the STE and MNI set into three disjoint sets of HFOs: the set of HFOs that are detected by both the STE and MNI detector (Intersection set), the set of HFOs that are detected by the MNI detector but missed by the STE detector (MNI only set) and the set of HFOs that are detected by STE detector but missed by the MNI detector (STE only set). We populated the three disjoint sets using an

algorithm described below:

- Step 1: For each HFO in the MNI set (s_{MNI}, e_{MNI}) , we compute its Jaccard similarity to all the HFOs detected by the STE detector and sort them in descending order
- Step 2: If there is an HFO in the STE set, (s_{STE}, e_{STE}) , with a Jaccard similarity greater than 0.95, then we place both the HFOs (s_{MNI}, e_{MNI}) and (s_{STE}, e_{STE}) in the Intersection set. If there is no HFO in the STE set with a Jaccard similarity greater than 0.95, then we place the HFO (s_{MNI}, e_{MNI}) in the MNI only set
- Step 3: Keep repeating Steps 1 and 2 until all the HFOs in the MNI set have been placed in either the Intersection or the MNI only set
- Step 4: The HFOs in the STE set that have not been placed in the Intersection set are used to populate the STE only set

The empirical distribution of Jaccard similarity scores was bimodal with two peaks at 0 and 1, prompting us to use a strict threshold of 0.95 for our greedy algorithm.

3.2.8.3 Training epileptogenic HFO (eHFO) detector (Fig. 3.1(C/D))

We introduced the class of eHFOs, which aimed to capture HFOs generated from the epileptogenic tissues of the brain [Zhang et al., 2022b]. It is obtained by training a DL-based detector for the HFOs that are distinctively present in the SOZ channels and other resection channels in post-operative seizure free patients. Since it is a data-driven procedure, the exact DL model is a function of the type of HFO detector. A defining characteristic of such data-driven eHFOs — irrespectively of the type of the HFOs used — is high amplitude in time–frequency plots across frequency bands at HFO onset (the inverted T-shaped template) [Hoogteijling and Zijlmans, 2022].

HFO set: The HFOs obtained by running the 90-minute intracranial EEG data through the STE and MNI detector were passed through the decomposition block to obtain the three

disjoint HFO sets: MNI only set, Intersection set, and the STE only set. We constructed the Union HFO set by taking the Union of the three disjoint sets (MNI only + Intersection + STE only). Similarly, we constructed the MNI HFO set by taking the Union of Intersection and MNI only sets (Intersection + MNI only).

Artifact detector: A human expert (HN: board certified in clinical neurophysiology and epilepsy, with experience in HFO analysis) classified candidate HFOs (c-HFOs) in the HFO set of each patient’s 10-minute EEG segments into two classes: artifacts and real HFOs. Artifacts are false-positive events, including ringing (filtering of sharp transients), as well as muscle and background fluctuations. Another expert with similar qualifications (SH) independently scored c-HFOs from two representative patients, and inter-rater reliability was examined using Cohen’s kappa statistics. Each c-HFO was represented by a one-second window, with the c-HFO was located at the center (0 ms), and including 500ms of EEG signal before and after. The time-frequency plot (scalogram) was generated by continuous Gabor Wavelets ranging from 10Hz to 500Hz. The scalogram was resized into the standard size (224 x 224) and repeated 3 times to serve as the input to the 3 channels of the Convolutional Neural Network (CNN).

The neural network architecture for the artifact detector was adapted from ResNet18 with a modification in the last few layers to accommodate the binary classification tasks. Specifically, the last layer of the Resnet18 was modified to be three fully connected layers with LeakyReLU, BatchNorm, and 10% dropout in between. The output of the model was fed into a sigmoid function to bound the output between 0 to 1, representing the probability of each task.

Feature extractor: After filtering for artifacts, we extracted three features from each real HFO in the set as our prior studies [Zhang et al., 2022a, Zhang et al., 2022b]: i) a time-frequency plot (scalogram) generated using continuous Gabor wavelet transforms ranging from 10-500 Hz, ii) an EEG-tracing plot generated on a 2000×2000 image by scaling the time-series signal into the 0–2000 range to represent the EEG waveform’s morphology,

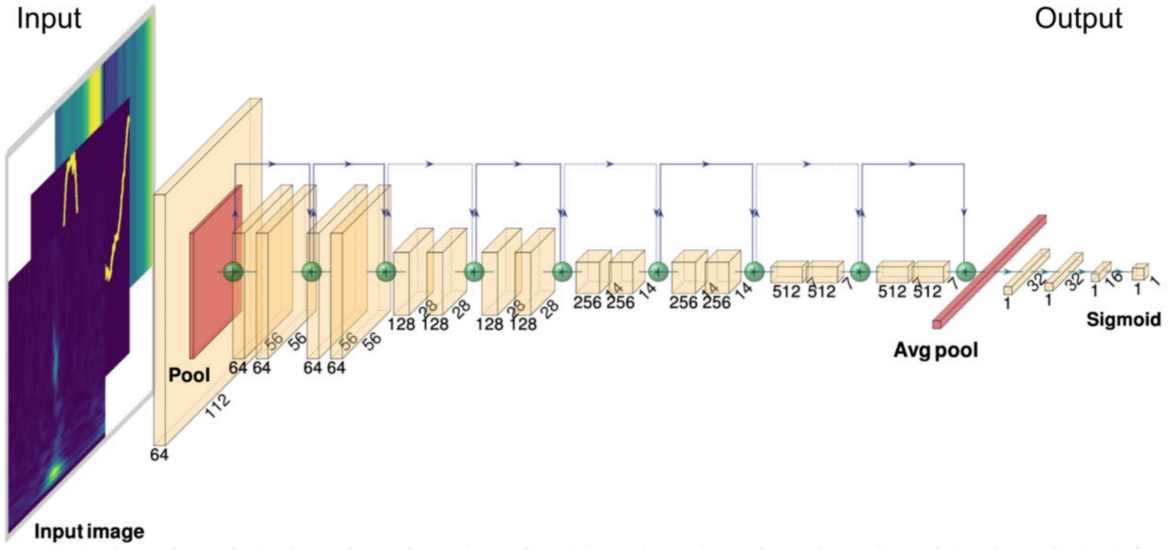


Figure 3.2: **CNN architecture:** The architecture of the model was adapted from Resnet-18. The last layer of the resnet18 was modified to be three fully-connected layers with LeakyReLU, BatchNorm, and 10% dropout in between. The output of the model was fed into a sigmoid function to bound the output between 0 to 1, representing the probability of each task.

iii) an amplitude-coding plot generated to represent the relative amplitude of the time series signal: for every time point, the pixel intensity of a column of the image represented the signal’s raw value at that time. These three images were resized into the standard size (224×224), serving as the input to the neural network.

Weak labeling based on clinical evidence:

We generated a weakly labeled training set with the following assumptions. For those patients that became seizure-free after resection (9 patients), we assumed that all epileptogenic tissue was contained in the resection. Similarly, the preserved regions in post-surgery seizure free patients could not contain epileptogenic tissue, leading us to assume that the HFOs in the preserved regions are non-epileptogenic. We thus labeled all of the HFOs from

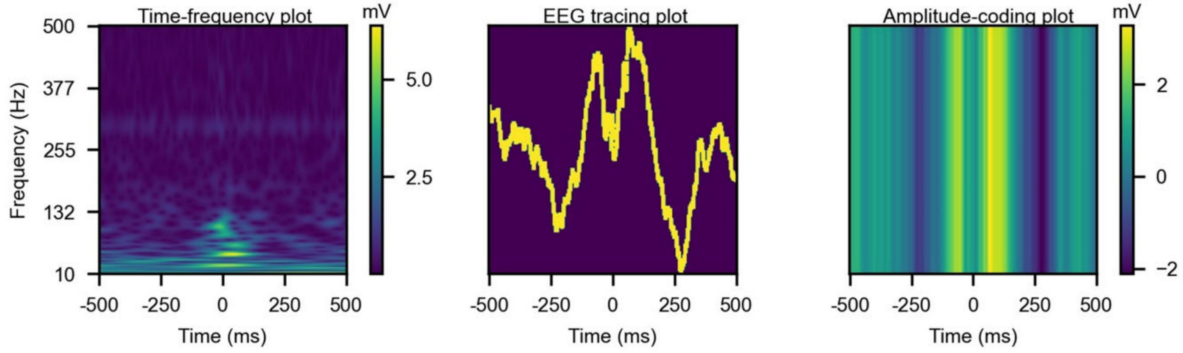


Figure 3.3: **Network Input Images:** The network input includes three images constructed from a one-second raw EEG segment with a detected HFO in the center (500ms before and after). Left: The time-frequency plot was generated by continuous Gabor Wavelets ranging from 10Hz to 500Hz. Middle: EEG tracing plot was generated on a 2000 x 2000 image by scaling the time-series signal into the 0 to 2000 range. Right: amplitude-coding plot contains the amplitude at every time point; a column of the image represented the signal’s actual value rescaled with a color gradient. These three images were resized into 224 x 224 in order to fit into the neural network.

resected channels as 1 and all HFOs from preserved channels as 0. By doing such weak supervision, we introduced potential errors in the data labels. Specifically, (1) the epileptogenic zone may generate non-eHFOs in addition to eHFOs; (2) some tissue was resected based on anatomical location (e.g. the channel was located within the same gyrus as the SOZ) and was likely not epileptogenic, leading to false-positive labels; and (3) in the non-resected region, while the majority of the HFOs were indeed non-epileptogenic, a few HFOs that are morphologically similar to epileptogenic HFOs could have been introduced by phenomena such as propagation. To address (1) and (3), we believed the automatic geometric purification property of the neural network could denoise the errors introduced by this mislabeling since the clinical outcomes constrain the number of mislabels to be few. To address (2), we introduced a weight term in the network’s loss function to reduce the noise introduced by

the false positives.

eHFO detector training:

For each input with a label of y and prediction of x , we introduced a weight term, w , in the Binary Cross Entropy (BCE) loss:

$$\mathcal{L} = -w * [y \log(x) + (1 - y) \log(1 - x)].$$

If $y = 0$ (the HFO is from a preserved channel), then $w = 1$. If $y = 1$ (the HFO is from a resected channel) and the HFO is from a SOZ channel, then also $w = 1$. However, if $y = 1$ and the HFO is not from a SOZ channel, we used a value of $w = 0.5$, which was found to be optimal for our dataset. We used the same CNN architecture as the one in Figure 3.2, except the CNN’s convolution layers were frozen to only act as feature extractors. Each of the three different eHFO detectors was trained by varying the training set: Union eHFO detector trained with Union HFO set, MNI eHFO detector trained with MNI set, and STE eHFO detector trained with STE set.

A patient-wise elimination training strategy was adopted to validate the findings of the detectors and generate the receiver operating characteristics curve. In this training strategy, we trained ten models in total: one model each for the seizure-free patients (9 seizure-free patients) and one model for the non-seizure-free patients. The model for a seizure-free patient was trained by excluding the iEEG data of that seizure-free patient from the training set. For example, we trained a model for a particular patient by constructing the training set from iEEG data of the remaining 8 seizure-free patients. The model for the non-seizure-free patients was trained by constructing the training set from iEEG data of the 9 seizure-free patients. In the inference stage, we used the patient-wise models to predict the outcome of the seizure-free patients and the non-seizure-free model to predict the outcome of the non-seizure-free patients. Therefore, we are using an independent cohort of EEGs to validate the findings of the detectors, as there is no overlap between the training and test set.

3.2.9 Evaluating the performance of eHFO detectors in predicting postoperative seizure outcomes

The trained eHFO detectors were used to classify the HFOs in the Union HFO set into epileptogenic HFOs (eHFO) and nonepileptogenic HFOs (non-eHFO). The ratio of the rates of eHFO in the resected region to the preserved region was used as a feature to train a logistic regression model for predicting the postoperative seizure outcomes of 15 patients. By varying the classification threshold on the ratio, a receiver operating characteristics curve (ROC) was generated. The area under the ROC curve was computed and used as a metric to evaluate the performance of the eHFO detectors in predicting the postoperative seizure outcome (seizure freedom at 24 months after surgery) of patients.

3.2.10 Time-frequency plot characteristics of the HFOs in the three HFO sets

We computed the mean time–frequency scalogram of HFOs in the three sets (Intersection, MNI only, STE only) by averaging the power across all the scalograms of HFOs in the sets. The mean scalograms were then compared across the three sets in four distinct frequency bands (10–30 Hz, 30–80 Hz, 80–250 Hz, 250– 500 Hz) by performing a t-test between the mean scalograms. We also created a customized scalogram, using the same technique as in our previous work [Zhang et al., 2022b], to visually represent the frequency ranges in which the scalograms of the HFOs belonging to different sets varied in. For every pixel (x, y) in the 224×224 image, we created two sets of data points: $S_{int}(x, y)$ and $S_{mni}(x, y)$, where $S_{int}(x, y)$ consists of the intensity values of the scalogram of the HFOs in the Intersection set and $S_{mni}(x, y)$ consists of the intensity values of the scalogram of the HFOs in the MNI only set. Then we performed one-tailed t-tests to determine whether a random variable $A(x, y)$, whose samples are given by $S_{int}(x, y)$, is greater than that of a random variable $B(x, y)$ whose samples are given by $S_{mni}(x, y)$. If this hypothesis is found to be true, with a p-value less than 0.001, we set the pixel value $I(x, y) = 1$; otherwise, $I(x, y) = 0$. Similarly,

by following the above procedure, we compared the time–frequency scalogram of HFOs in the MNI only, and the STE only sets with two sets of data points: $S_{mni}(x, y)$ and $S_{ste}(x, y)$.

3.2.11 Distribution of HFO with spike and epileptogenic HFO in three HFO sets

We passed the three HFO sets (Intersection, MNI only, STE only) through the trained spike detector [Zhang et al., 2022b] to classify them into HFOs with spikes (spk-HFO) and HFOs without spikes (non-spk-HFO) categories. Similarly, we passed the three HFO sets through the trained union eHFO detector to classify them into eHFO and non-eHFO categories. We compared the distribution of HFO with spike and eHFO across the three HFO sets.

3.2.12 Relationship between epileptogenic HFOs and HFOs with spikes

We passed the Union HFO set through the three trained eHFO detectors (Union, MNI, and STE) to classify the HFOs into eHFO and non-eHFO categories. We also passed the Union HFO set through the trained spike detector to classify the HFOs into HFOs with spikes (spk-HFO) and HFOs without spikes (non-spk-HFO) categories. For each eHFO detector type, we computed the overlap between the eHFO and spk-HFO set and then compared the overlap across the three eHFO detectors.

3.2.13 Statistical analysis

We used Python (version 3.7.3; Python Software Foundation, USA) for above mentioned statistical calculations. PyTorch (version 1.6.0; Facebook’s AI Research lab) was used to develop the deep neural network. Quantitative measures are described by medians with interquartile or mean with standard deviations. Comparisons between groups were performed using chi-square for comparing two distributions and Student’s t-test for quantitative measures (in means with standard deviations). All comparisons were two-sided, and significant

results were considered at $p < 0.05$ unless stated otherwise. The particular statistical tests executed for each experiment are detailed in their respective sections.

3.3 Results

3.3.1 Patients' characteristics

There were 15 patients (7 males, 8 females) enrolled during the study period. The median age at surgery was 13 years (range: 3–20 years). The median electrocorticography monitoring duration was four days (range: 2–12 days), and the median number of seizures captured during the monitoring was 7 (range: 1–35). All 15 patients underwent resection immediately following intracranial monitoring and provided postoperative seizure outcomes at 24 months (9 of 15 became seizure-free).

3.3.2 Interictal HFO detection and HFOs' classification

A total of 105,622 HFOs were detected with the STE detector and 47,143 with the MNI detector in 90-minute EEG data from the 15 patients. Of those, 29,668 were in the Intersection set. After artifact rejection, there were 84,602 HFOs detected with the STE detector, 34,965 with the MNI detector and 21,248 of them belonged to the Intersection set. We classified the HFOs in the three sets (Intersection, MNI only, and STE only) into HFOs with spikes and HFOs without spikes. Across 15 patients, on average, 88% of HFOs in the Intersection set were HFOs with spikes, 87% of HFOs in the MNI only set were HFOs with spikes, and 36% of HFOs in the STE only set were HFOs with spikes. The percentage of HFO with spikes was significantly higher in the Intersection and the MNI only set when compared to the STE only set (Intersection vs. STE only: mean 0.88 versus 0.36, $p < 0.001$; MNI only vs. STE only: mean 0.87 versus 0.36, $p < 0.01$) (Fig. 3.4).

3.3.3 Spectral characteristics of HFOs in three HFO sets

The analysis of the mean time–frequency map averaged across the 15 patients demonstrated that HFOs in the Intersection set showed higher amplitude in all the frequency bands (fast ripple: 250–500 Hz; ripple: 80–250 Hz; gamma: 30–80; and beta: 10–30 Hz) than HFOs in the MNI only or STE only set ($p < 0.0001$ for all comparisons). However, comparisons of the amplitude between the MNI only and the STE only HFOs did not show differences in fast ripple, ripple, and gamma bands ($p = 0.90, 0.99, \text{ and } 0.99$, respectively). In the beta band, HFOs in the MNI only set have more amplitude than the ones in the STE only set ($p < 0.001$) (Fig. 3.5A). With more detailed pixel-by-pixel analysis, we demonstrated that the HFOs in the Intersection set had higher amplitude than HFOs in the MNI only set in all four frequency bands, with a noticeable difference in the 30–80 Hz and 80–250 Hz frequency bands. Similarly, HFOs in the MNI only set have higher amplitude than HFOs in the STE only set in the 10–30 Hz and 30–80 Hz frequency bands, with a significant difference in the 10–30 Hz frequency band (Fig. 3.5B).

3.3.4 Overlap between eHFOs and HFOs with spikes in three eHFO detectors

We classified the HFOs in the Union HFO set (Intersection + MNI only + STE only) into HFOs with spikes and HFOs without spikes. We also classified the HFOs in the Union HFO set into eHFOs and non-eHFOs using the three types of eHFO detectors (Union, MNI, and STE). Across the 15 patients, there was an overlap ratio of 0.79, 0.95, and 0.78 between HFOs with spikes and eHFOs with the Union, MNI, and STE eHFO detectors, respectively. The overlap percentage was significantly higher in the MNI HFOs than in the Union and STE HFOs (MNI vs. Union: mean 0.95 vs. 0.79, $p < 0.001$; MNI vs. STE: mean 0.95 vs. 0.78, $p < 0.001$). There were no significant differences in the overlap percentage when comparing the Union and STE HFOs (Union vs. STE: mean 0.79 vs. 0.78, $p = 0.042$, which became insignificant after Bonferroni correction)(Fig. 3.6).

3.3.5 Characteristic differences of eHFOs vs. non-eHFOs controlled by spike status

We investigated the differences in the time–frequency characteristics between eHFOs and non-eHFOs controlled by spike status. Across four frequency bands, eHFOs with spikes showed higher power than non-eHFOs with spikes, and eHFOs without spikes showed higher power than non-eHFOs without spikes ($p < 0.0001$ in all comparisons with Welch t-test with unequal variances) (Fig. 3.7).

3.3.6 Distribution of eHFO in three HFO sets

We classified the HFOs in the three sets (Intersection, MNI only, and STE only) into eHFOs and non-eHFOs using the Union eHFO detector. Across 15 patients, on average, 82%, 85%, and 45% of HFOs in the Intersection set, MNI only set, and STE only set were eHFOs, respectively. The percentages of eHFOs were significantly higher in the Intersection, and MNI only set when compared to the STE only set (Intersection vs. STE only: mean 82% versus 45%, $p < 0.001$; MNI only vs. STE only: mean 85% versus 45%, $p = 0.03$). We plotted the distribution of the confidence scores (the probability of each HFO being eHFO) returned by the Union eHFO detector for the three sets. The distribution for the Intersection and the MNI only set was skewed towards 1 (higher confidence towards eHFO), while the STE only set showed a more balanced distribution (Fig. 3.8).

3.3.7 Predicting postoperative seizure outcomes using HFO resection ratios

We used the HFO and eHFO resection ratio (ratio = total number of resected HFOs/total number of detected HFOs) as a classifier to predict the postoperative seizure outcomes. Three types of HFO detectors (Union, MNI, and STE) were used for computing the ratios. For both cases (HFO and eHFO resection ratios), the Union HFO detector achieved the best performance in predicting the postoperative seizure outcome (AUC of 0.78 and 0.89),

followed by the MNI HFO detector (AUC of 0.76 and 0.85), and then the STE eHFO detector (AUC of 0.72 and 0.81) (Fig. 3.9).

3.3.8 Post-hoc sensitivity analysis

We conducted a post-hoc sensitivity analysis to prove that the superior performance of the Union eHFO detector was not due to biased HFO samples. We trained the STE and the Union eHFO detector using the same number of training samples ($n = 84602$). In training the STE eHFO detector, the training set only consisted of HFO samples from the STE set. In training the Union eHFO detector, the training set consisted of an equal number of HFO samples from the STE and MNI set. The trained models were then used to predict the patient outcomes and generate receiver operating characteristics (ROC) curves. We plotted the area under the ROC curve (AUC) as a function of the number of training samples used to train the detectors. From the plot, it can be observed that there is a significant gap between the curves showing that even for the same number of training samples union eHFO detector performs better in terms of predicting the seizure outcomes.

3.4 Discussion

Although HFOs are widely investigated as one of the spatial biomarkers of the EZ, the definition of what constitutes an HFO event is computational and algorithm-dependent [Frauscher et al., 2017, Remakanthakurup Sindhu et al., 2020]. There have been some efforts at comparing the signal characteristics of the two widely used detectors, namely the MNI and the STE detectors; for example, the study [Zelmann et al., 2010] that introduced the MNI detector, showed it could detect HFO events in EEG data with noisy background (based on the baseline adjustment principle) that were often missed by the STE detector [Zelmann et al., 2012]. However, this does not answer whether the MNI detector—which has more stringent settings—also misses the detection of certain physiological and pathological HFOs that could

have clinical impact. Thus, a comprehensive study comparing the signal and morphological differences of STE and MNI HFOs, especially in the context of pathological HFOs, is needed. In the study presented in this chapter, we first verified that, even at the level of individual subjects, each detector produced unique HFO events not detected by the other (referred to as MNI only and STE only HFO sets, respectively) while sharing an intersection set of HFOs (Intersection-set HFOs). We further showed that the Intersection set HFOs, which were detected by both detectors, contained most pathological features demonstrated by high association rates with spikes and their signal characteristics being consistent with eHFOs (high amplitude in time– frequency plots across frequency bands at HFO onset) as reported in our previous study [Zhang et al., 2022b]. The MNI HFOs were more biased towards HFOs with spikes and eHFOs and the STE HFOs were more biased towards HFOs without spikes and non-eHFOs. However, both the MNI only and STE only HFO sets contained significant number of pathological HFOs (for example, the STE only HFO set has 36% of total spk-HFO that are thus missed by the MNI detector). Thus the Union eHFO detector that processes all the HFOs by either the MNI or the STE detector could potentially better capture the characteristics of HFOs that are distinctively present in EZ. We then created three different eHFO detectors namely Union eHFO, STE eHFO, and MNI eHFO. Indeed, we demonstrated that (i) using any eHFO resection ratio outperformed the unclassified HFO resection ratio in any HFO detection method in post operative seizure outcome prediction; (ii) the Union eHFO detector outperformed all other detectors in postoperative seizure outcome prediction.

Our findings are clinically meaningful. We presented the data to show that neither the MNI nor the STE detector is sensitive enough to evaluate pathological HFOs. When we use HFO evaluation in clinical context, we might consider using several types of detection methods, as proposed in our study, to enhance the sensitivity of the detection. In addition, not all detected HFOs are pathological. In a recent clinical trial and observational study, issues with the feasibility of utilizing HFOs in clinical settings were elucidated [Jacobs et al., 2018, Zweiphenning et al., 2022a]. These included the time constraints of expert-based visual

analysis in detecting and classifying HFOs (artifacts, pathological HFOs, and physiological HFOs). Automated and reproducible detection and classification of HFOs will be crucial to use HFOs in clinical practice. In another work which we will discuss in the next chapter, data from cortical stimulation mapping were used to characterize HFOs with origin in the eloquent cortex and hence, physiological in origin [Zhang et al., 2022a]. If HFOs can be automatically classified into such HFOs that are expressed in the eloquent cortex, then one can start localizing the eloquent cortex from HFO detections, which would have a significant clinical impact. Since the STE detector can detect many more of such physiological HFOs than the MNI detector, it further underscores the need for using both detectors and then designing DL-assisted classifiers to automatically extract clinically relevant information.

Given that discovering the characteristics of pathological HFOs is an active area of research, we now discuss the relationship of such work with our findings. For example, HFOs with higher amplitude are hypothesized to be characteristics of pathological HFOs [Charupanit et al., 2020]. Similarly, HFOs with spikes are believed to be correlated with the pathological nature of HFOs [Weiss et al., 2016]. Our study also substantiated such findings as we observed that the Intersection-set HFOs had higher amplitude in all the frequency bands and a high spike association rate. We further investigated characterizations of pathological HFOs using a DL based data-driven approach as introduced in our recent work [Zhang et al., 2022b], where no a priori feature engineering is required to train the algorithms. In particular, data from patients who became seizure-free post-surgery was used to characterize HFOs into eHFOs and non-eHFOs. Such eHFOs were shown to share known characteristics of pathological HFOs [Zhang et al., 2022b]. Thus, data obtained from post operative seizure free patients along with DL can be used to automatically distill manually defined HFOs to provide clinically useful localization information about EZ, and thus obtain automatic characterizations of pathological HFOs. Future studies utilizing larger datasets and more sophisticated DL-architectures may further characterize pathological and physiological HFOs [Hoogteijling and Zijlmans, 2022].

There are several limitations to our study. We only tested the STE and MNI detection methods of HFO analysis. We plan to investigate other methodologies, such as short line length (SLL) [Gardner et al., 2007] and Hilbert detections [Crépon et al., 2010]. The iEEG data was recorded by placing intracranial strips and grid electrodes. Although we expect the analysis to hold in a clinical setting where the data is recorded using stereotactic EEG (SEEG), the analysis and the deep learning model need to be validated on a dataset of sEEG recordings. Another potential limitation of our study is the diversity of the dataset in terms of the number of patients, age group of the patients, state of the patients (we analyzed sleep segment only), and the duration of the EEG recordings. We used EEG data from non-REM sleep, as pathological HFOs were more commonly detected during this state [von Ellenrieder et al., 2017]. Awake EEG data can be affected by many factors, such as subjects’ motion, thus we did not prepare such a dataset. Since the morphology of the detections might be affected by the age and vigilance state of the patients, the transferability of the analysis and the model needs to be verified in a larger cohort of patients with a diversified age range, vigilance state, and epilepsy etiology.

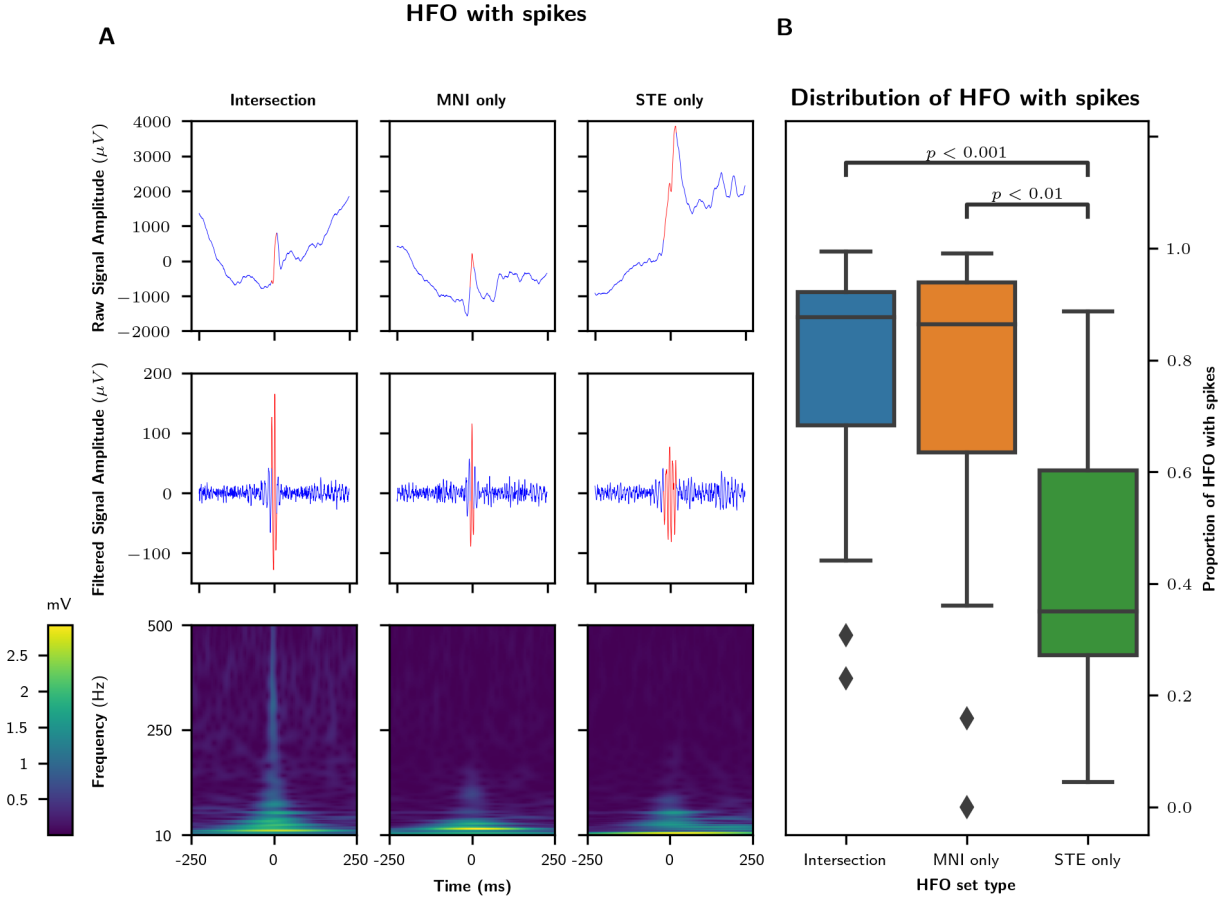


Figure 3.4: **Distribution of HFOs with spikes in three HFO sets:** (A) An example of HFO with spike from each HFO set (Intersection, MNI only, and STE only) is shown. The first row shows the original tracing, the second row shows the filtered tracing (80–500 Hz), and the third shows the time–frequency plot. (B) The proportion of HFO with spikes in 15 patients is plotted as box plots and labeled by the HFO set type. Across a patient body of 15 patients, the Intersection set has the highest proportion of HFO with spikes on average, followed by the MNI and STE only set. The proportion of HFOs with spikes was significantly higher in the Intersection, and MNI only set compared to the STE only set (Intersection vs. STE only: mean 0.88 versus 0.36, $p < 0.001$; MNI only vs. STE only: mean 0.87 versus 0.36, $p < 0.01$). HFOs: high-frequency oscillations; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.

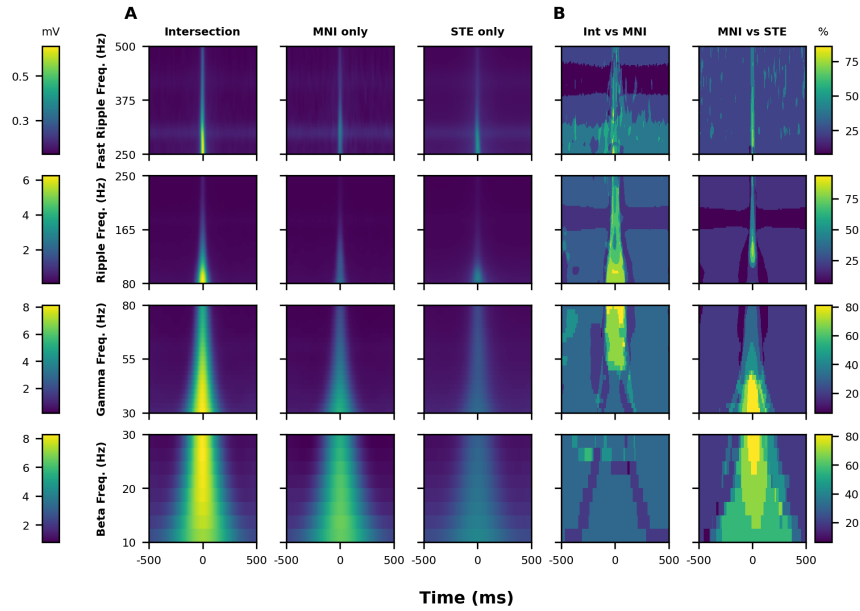


Figure 3.5: Time-frequency plot characteristics of HFOs in three HFO sets: (A) The time-frequency plots of HFOs, averaged across all patients ($n = 15$), are plotted and labeled based on their set type and four distinct frequency bands. Across four frequency bands, HFOs in the Intersection set have higher amplitude than HFOs in the MNI only set and STE only set. HFOs in the MNI only set have higher amplitude than STE only set in the 10–30 and 30–80 Hz frequency bands. (B) Pixel-wise analysis of time-frequency plot amplitude differences among three HFO sets: We created a custom scalogram to visually represent the frequency ranges where the HFOs belonging to different sets varied. The lighter-colored regions in the figure indicate pixels where the mean amplitude of one HFO set is statistically higher (p-value below 0.005 from a one-tailed t-test) than the other HFO set. From the left plot, it is clear that the HFOs in the Intersection set have higher amplitude than HFOs in the MNI only set in all four frequency bands, with a noticeable difference in the 30–80 Hz and 80–250 Hz frequency bands. Similarly, the right plot shows that HFOs in the MNI only set have higher amplitude than HFOs in the STE only set in the 10–30 Hz and 30–80 Hz frequency bands, with a significant difference in the 10–30 Hz frequency band. HFOs: high-frequency oscillations; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.

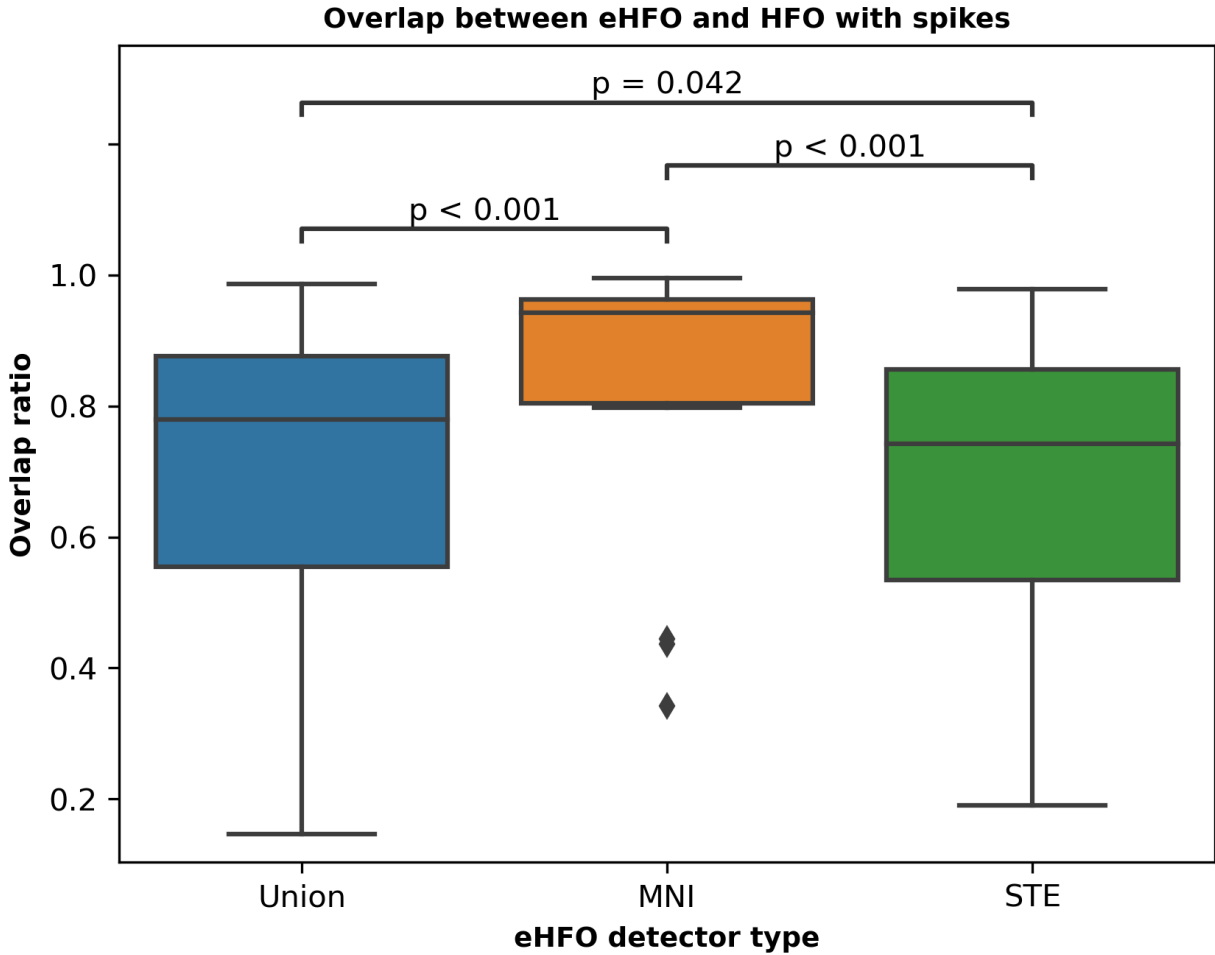


Figure 3.6: **Overlap between eHFO and HFO with spike set in three eHFO detectors:** The overlap ratio between eHFO and HFO with spike set in 15 patients is plotted as box plots and labeled by the eHFO detector type. Across the cohort of 15 patients, there was significant overlap between the eHFOs and HFOs with spikes for all three eHFO detectors. The overlap ratio was significantly higher in the MNI eHFO detector when compared to the Union and STE eHFO detectors (MNI vs. Union: mean 0.95 versus 0.79, $p < 0.001$; MNI vs. STE: mean 0.95 versus 0.78, $p < 0.001$) while there was no significant difference in the overlap ratio between Union and STE eHFO detectors (Union vs. STE: mean 0.79 versus 0.78, $p = 0.042$). HFOs: high-frequency oscillations; eHFOs: epileptogenic HFOs; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.

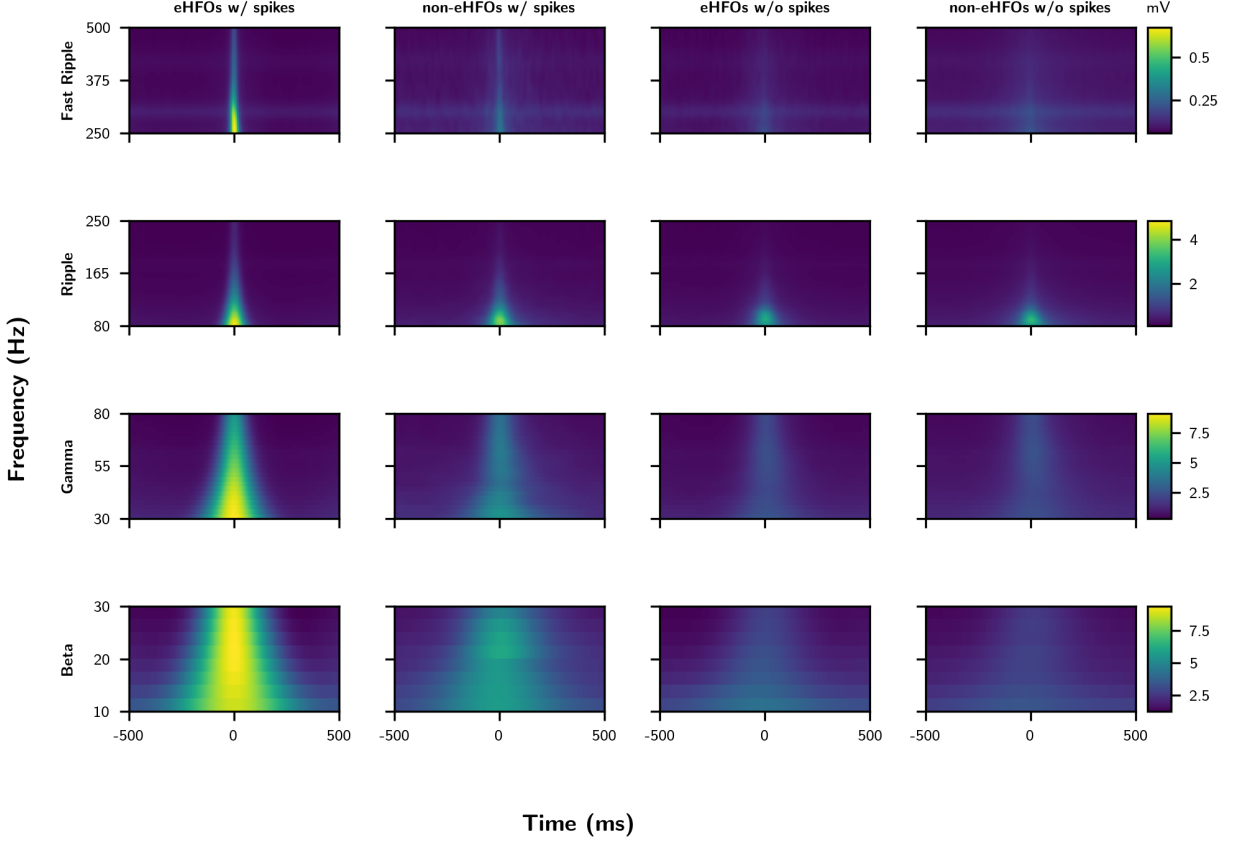


Figure 3.7: **Time-frequency plot characteristics of eHFOs vs. non-eHFOs controlled by spike status:** The time–frequency plots of eHFOs and non-eHFOs (detected with Union HFO detector) averaged across all patients ($n = 15$) were plotted based on spike status (with or without spikes). Across four frequency bands, eHFOs with spikes showed higher amplitude than non-eHFOs with spikes, and eHFOs without spikes showed higher amplitude than non-eHFOs without spikes ($p < 0.0001$ in all comparisons with Welch t-test with unequal variances). HFOs: high-frequency oscillations; eHFOs: epileptogenic HFOs.

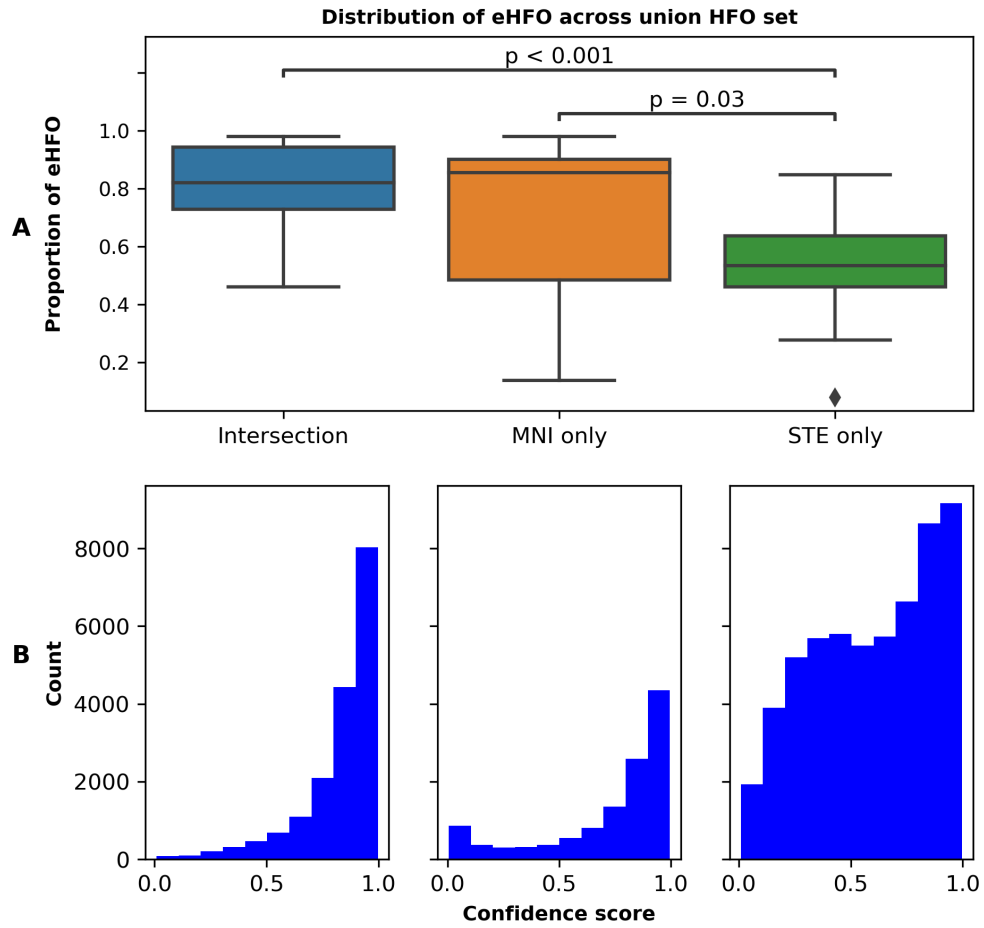


Figure 3.8: **Distribution of eHFO in three HFO sets:** (A) Proportion of eHFO in 15 patients is plotted as box plots and labeled by the HFO set type. Across the cohort of 15 patients, MNI only set has the highest proportion of eHFO on average, followed by Intersection and STE only set. The proportion of eHFO was significantly higher in the Intersection, and MNI only set when compared to the STE only set (Intersection vs. STE only: mean 0.82 versus 0.45, $p < 0.001$; MNI only vs. STE only: mean 0.85 versus 0.45, $p = 0.03$). (B) Confidence scores of eHFOs in 15 patients were plotted as a histogram and labeled by the HFO set type. The histogram for the Intersection and MNI only set has a skewed distribution strongly biased towards 1 (direction towards eHFO), while the STE only set has a more uniform distribution. HFOs: high-frequency oscillations; eHFOs: epileptogenic HFOs; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.

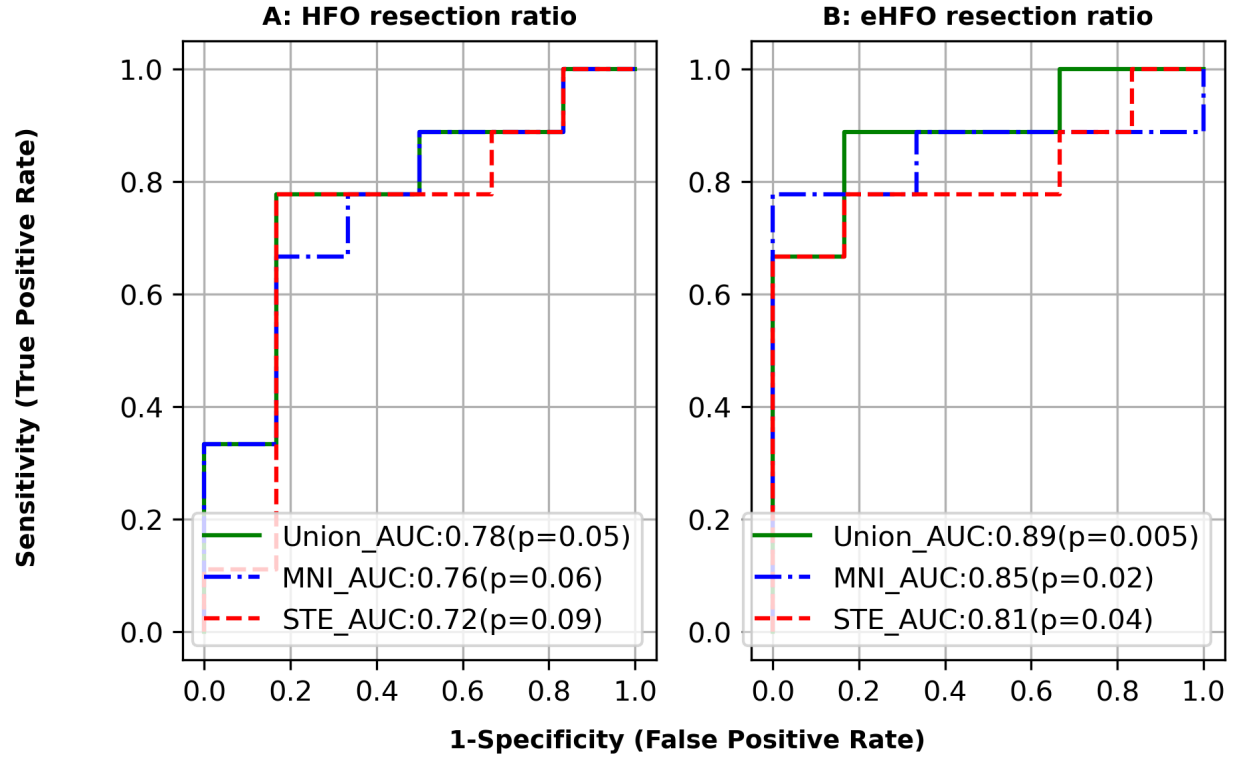


Figure 3.9: Accuracy of postoperative seizure outcomes prediction using HFO resection ratios: We constructed a postoperative seizure outcome prediction model using the resection ratio derived from 90-minute EEG data ($n = 15$) (same approach as in our previous work (Zhang et al. 2022a)). Each receiver-operating characteristics (ROC) curve delineates the accuracy of seizure outcome classification of a given model, using the area under the ROC curve (AUC) statistics. A) The HFO resection ratio (detected HFOs filtered for artifacts) obtained from 3 models (Union, MNI, and STE) was used to predict the seizure outcome. Union and MNI detectors have higher AUC values than a model obtained by the STE detector. B) eHFO resection ratio (detected HFOs filtered for artifacts and non-eHFO) obtained from 3 models (Union, MNI, and STE) was used to predict the seizure outcome. Overall significant improvement in prediction performance was seen, with the Union detector model having the highest AUC of 0.89. HFOs: high frequency oscillations; eHFOs: epileptogenic HFOs; MNI: Montreal Neurological Institute (detector); STE: Short-term energy.

CHAPTER 4

Characterizing physiological high-frequency oscillations using deep learning

4.1 Introduction

Intracranially-recorded interictal high-frequency oscillations (HFOs) have been proposed as a promising spatial biomarker of the epileptogenic zone (EZ) [Bragin et al., 1999a, Gotman, 2010]. Animal and human studies have demonstrated the association between HFOs and brain tissue capable of generating seizures [Bragin et al., 1999b, Staba et al., 2004b, Jirsch et al., 2006, Worrell et al., 2004]. Several retrospective studies have linked favorable post-surgical seizure outcomes to the resection of cortical sites showing interictal HFOs [Jacobs et al., 2010a, Wu et al., 2010, Akiyama et al., 2011, van’t Klooster et al., 2015, Fedele et al., 2016].

However, despite the potential, HFOs can also be recorded in healthy brain regions, which complicates the interpretation of HFOs when one attempts to guide resection using HFOs. Several recent studies, including a large multicenter prospective study, failed to correlate the removal of HFO-generating brain regions with postoperative seizure freedom; some patients became seizure-free despite part of the brain regions generating HFOs being preserved [Jacobs et al., 2018, Roehri et al., 2018]. An ongoing clinical trial utilizing HFOs in electrocorticography to guide resection has excluded enrollment of occipital lobe epilepsy, due to abundant physiological HFOs in the visual areas [van’t Klooster et al., 2015]. The current impasse is that there are no methods to separate pathological from physiological

HFOs.

Within the field of artificial intelligence, machine learning can bridge statistics and computer science to develop algorithms to complete tasks by exposure to meaningful clinical data without explicit instruction. Indeed, machine learning has been successfully applied to the problem of classifying HFOs based on a priori manual engineering of event-wise features, which includes: linear discriminant analysis [Jrad et al., 2015], support vector machines [Amiri et al., 2016, Jrad et al., 2016], decision trees [Chaibi et al., 2014], and clustering [Liu et al., 2018]. More recently, the deep learning (DL) framework has been adopted, which directly works with raw data (avoiding any a priori feature engineering) and yields better performance in the field of neuroimaging [Lundervold and Lundervold, 2019]. Leveraging DL’s revolutionary success in the field of computer vision using convolutional neural networks (CNNs), prior studies explored the use of CNNs in electroencephalogram (EEG) analysis, especially converting one-dimensional EEG signals into a two dimensional image for CNNs input [Jing et al., 2020, Zhao et al., 2020, Zuo et al., 2019]. The previous DL approaches conducted the HFO classification in a supervised manner, requiring human annotated labels which constrains the spectrum of usage of their methods, especially the needs of human expert labeling. In the context of medical image analysis, recent work has shown that optimized model architectures and loss functions could mitigate data labeling errors, thus making the DL framework even more applicable [Karimi et al., 2020]. Our recent work demonstrated that using the channel resection status as DL labels, a novel weakly-supervised DL algorithm characterized HFOs generated by the EZ [Zhang et al., 2022b].

The study presented in this chapter aimed to characterize physiological HFOs, using functional cortical mapping results as DL training labels. We leveraged results from functional cortical mapping by stimulation and gamma-related language mapping in children with neocortical medication-resistant epilepsy who underwent intracranial EEG (iEEG) monitoring with a grid/strip approach. We characterized salient features of physiological HFOs represented by the eloquent cortices (echFOs) through the DL algorithm. Also, we investigated

if removing cortical regions with HFOs excluding ecHFOs (to mitigate the effect of physiological HFOs) correlated with postoperative seizure freedom better than using uncorrected HFOs alone.

4.2 Methods

4.2.1 Patient cohort

This was a retrospective cohort study. Children (below age 21) with medically refractory epilepsy (typically with monthly or greater seizure frequency and failure of more than three first-line anti-seizure medications) who had intracranial electrodes implanted for the planning of epilepsy surgery with anticipated cortical resection with the Pediatric Epilepsy Program at University of California, Los Angeles (UCLA) were consecutively recruited between August 2016 and August 2018. Diagnostic stereo-EEG evaluation (not intended for resective surgery) was excluded.

4.2.2 Standard protocol approvals, registrations, and patient consents

The institutional review board at UCLA approved the use of human subjects and waived the need for written informed consent. All testing was deemed clinically relevant for patient care, and also all the retrospective EEG data used for this study were de-identified before data extraction and analysis. This study was not a clinical trial, and it was not registered in any public registry

4.2.3 Patient evaluation

All children with medically refractory epilepsy referred during the study period underwent a standardized presurgical evaluation, which—at a minimum—consisted of inpatient video-EEG monitoring, high resolution (3.0 T) brain magnetic resonance imaging (MRI),

and 18 fluoro-deoxyglucose positron emission tomography (FDG-PET), with MRI-PET co-registration [Nariai et al., 2019]. The margins and extent of resections were determined mainly based on seizure onset zone (SOZ), clinically defined as regions initially exhibiting sustained rhythmic waveforms at the onset of habitual seizures. In some cases, the SOZs were incompletely resected to prevent an unacceptable neurological deficit.

4.2.4 Subdural electrode placement

Macroelectrodes, including platinum grid electrodes (10 mm intercontact distance) and depth electrodes (platinum, 5 mm intercontact distance), were surgically implanted. The total number of electrode contacts in each subject ranged from 40 to 128 (median 96 contacts). The placement of intracranial electrodes was mainly guided by the results of scalp video-EEG recording and neuroimaging studies. All electrode plates were stitched to adjacent plates, the edge of the dura mater, or both, to minimize the movement of subdural electrodes after placement.

4.2.5 Acquisition of three-dimensional (3D) brain surface images

We obtained preoperative high-resolution 3D magnetization-prepared rapid acquisition with gradient echo T1-weighted image of the entire head. A FreeSurfer-based 3D surface image was created with the location of electrodes directly defined on the brain surface, using post-implant computed tomography images [Nakai et al., 2017]. In addition, intraoperative pictures were taken with a digital camera before dural closure to enhance the spatial accuracy of electrode localization on the 3D brain surface. Upon re-exposure for resective surgery, we visually confirmed that the electrodes had not migrated compared to the digital photo obtained during the electrode implantation surgery.

4.2.6 iEEG recording

iEEG recording was obtained using Nihon Kohden Systems (Neurofax 1100 A, Irvine, California, USA). The study recording was acquired with a digital sampling frequency of 2000 Hz, which defaults to a proprietary Nihon Kohden setting of a low frequency filter of 0.016 Hz and a high frequency filter of 600 Hz at the time of acquisition. For each subject, a continuous 90 min EEG segment were selected at least two hours before or after seizures, before anti-seizure medication tapering, and before cortical stimulation mapping, which typically occurred two days after the implant. All the study iEEG data were part of the clinical EEG recording.

4.2.7 Functional cortical mapping

Cortical stimulation was performed as part of clinical management to define regions of the eloquent cortices (EC) to help guide resections. A pulse train of repetitive electrical stimuli was delivered to neighboring electrode pairs, with a stimulus frequency of 50 Hz, pulse duration of 300 μ s, and train duration ranging up to 5 s (sensorimotor or visual mapping) or 7 s (language mapping). Stimulus intensity ranged from 1 mA to 13 mA. Seven out of 19 patients also underwent auditory and picture-naming tasks to supplement cortical stimulation mapping. In short, patients were instructed to overtly verbalize an answer to a given auditory question or name a picture. Each EEG trial data was transformed into the time–frequency domain using complex demodulation via BESA software (BESA GmbH, Germany). The iEEG signal at each channel was assigned an amplitude (a measure proportional to the square root of power) as a function of time and frequency (in steps of 10 ms and 5 Hz). The time–frequency transform was obtained by multiplication of the time domain signal with a complex exponential, followed by a band-pass filter iEEG traces were aligned to: (a) stimulus (question) onset; (b) stimulus offset; and (c) response (answer) onset. We determined whether the degree of such gamma-augmentation reached significance using

studentized bootstrap statistics followed by Simes’ correction [Brown et al., 2008, Koga et al., 2011]. Sites surviving correction showing significant gamma-augmentation spanning (a) at least 20 Hz in width and (b) at least 20 ms in duration were defined as ‘language-related gamma sites’ [Kojima et al., 2013]. Further methodological details of the gamma mapping are described in the prior studies [Kambara et al., 2018].

4.2.8 Automated detection and classification of HFOs

A customized average reference was used for the HFO analysis, with the removal of electrodes containing significant artifacts [Nariai et al., 2019, Kuroda et al., 2021]. Candidate interictal HFOs were identified by an automated short term energy detector (STE) [Staba et al., 2002, Navarrete et al., 2016]. This detector considers HFOs as oscillatory events with at least six peaks and a center frequency occurring between 80 and 500 Hz. The root mean square threshold was set at five standard deviations (SD), and the peak threshold was set at three SD. The HFO events are segments of EEG signals with durations ranging from 60 to 200 ms [Navarrete et al., 2016]. The candidate HFOs were evaluated by our previously validated algorithm (recall = 98.0%, precision = 96.1%, F1 score = 96.8% against human experts validation) to reject artifacts for further analyses [Zhang et al., 2022b]. For simplicity, we use HFO in the later paragraphs to represent the HFO after the artifact rejection.

4.2.9 Weakly supervised DL using functional cortical mapping results as labels

The general workflow of the DL training and inference was shown in the flowchart (Figure 4.1(A)). The channel-wise annotations include behavior (positive behavioral changes with cortical stimulation mapping or language-related gamma sites), seizure (SZ)/after discharge (AD) (channels with seizures or after discharges with cortical stimulation mapping), spike (cortical sites showing spontaneous interictal spikes), and none (no behavioral changes with mapping, SZ/AD, and spikes). We hypothesize that a morphologically distinct class of

HFOs generated by EC, representative as physiological HFOs (defined as ecHFOs). We expect a large percentage of the HFOs in the behavioral channels are ecHFOs. However, such channels could also have HFOs that are morphologically distinct from ecHFOs, since each channel also picks up neuronal activities generated by other physiological or pathological processes. We refer to such complementary HFOs as non-ecHFOs. Since the physiological processes generating ecHFOs are also present in different regions of the brain, for example, the unstimulated brain region, we expect non-behavioral channels to also have ecHFOs. However, as there is no ground truth annotation for an ecHFO event, we relied on channel-level annotations acquired from clinical experiments. Specifically, positive labels are assigned to HFOs from channels that only have behavior responses, excluding spike and SZ/AD channels, and negative labels are assigned to all other HFOs. When an HFO is from a channel with both behavior response and pathological (SZ/AD or Spike) response, we excluded them from the training set, as the distribution of ecHFOs and non-ecHFOs might be comparable. Clearly, such labeling has errors. However, if ecHFOs have distinct morphological signatures, then we could utilize the generalization power of the neural network to distill the ecHFO from this weak supervision.

4.2.10 Model architecture and feature representation of HFOs

We adopt the same CNN architecture as the prior study (Chapter 3) because it has demonstrated excellent performance in classifying HFOs [Zhang et al., 2022b]. A one-second window represented each HFO, centered on the HFO (0 ms). While the HFO duration is generally shorter than 200 ms, it does not necessarily mean that the discovered features representing the HFO’s characteristics are fully describable within that window. For example, spikes associated with HFOs provide distinct attributes in the time–frequency window, extending before and after the HFO detection point. Therefore, we used the longer window size to provide a more comprehensive description of an HFO event, which encompasses HFO itself and surrounding EEG waveforms [Zhang et al., 2022b]. Then three images are constructed

to capture the time–frequency domain features as well as signal morphology information of the HFO window. The time–frequency plot (scalogram) was generated by continuous Gabor Wavelets ranging from 10 Hz to 500 Hz. The EEG tracing plot was generated on a 2000×2000 image by scaling the timeseries signal into the 0–2000 range to represent the EEG waveform’s morphology. The amplitude-coding plot was generated to represent the relative amplitude of the time-series signal: for every time point, the pixel intensity of a column of the image represented the signal’s raw value at that time. No frequency filtering was used to generate the EEG tracing and amplitude coding plots. Then these three images were resized into the standard size (224×224), serving as the input to the neural network (Figure 4.1(B)). Regarding the time–frequency plot, the resizing operation is just binning the frequency bands, and the time–frequency plot still keeps the frequency information of the HFO signal. The EEG tracing plot captures the general EEG waveform of the HFO event, and the amplitude coding plot captures the HFO’s relative amplitude information. Thus, resizing the image does not lose HFO’s frequency information. This operation provides comprehensive HFO information as inputs to the neural network.

4.2.11 Model training

We used 90 min of data from each patient for training. Since our task is binary classification, we use binary cross-entropy as the loss function, $L = -[y\log(x) + (1 - y)\log(1 - x)]$, where y is the label (1 for ecHFO, 0 for non-ecHFO [ecHFOs were subtracted from total HFOs]), and x is the model output ecHFO probability. In training, we adopt stratified sampling to balance the data distribution in different labels. The Adam optimizer was adopted with a learning rate of 0.0001. The training was conducted using 25 epochs (training iterations), and validation loss was plotted with respect to the number of epochs completed. For selecting the best model across epochs, we picked the model corresponding to the balanced global minima over 25 iterations, i.e. the highest balanced accuracy (the averaged recall for each class). In order to fully explore the performance of the model, five-fold cross-validation was

conducted using the pooled data across the full patient cohort. For each fold, 20% of the dataset was selected as the test set, 70% was selected as the training set, and the remaining 10% was used for validation.

4.2.12 Performance analysis framework

Since the model was evaluated using five-fold cross-validation, there were five distinct models trained by different folds of data. For the downstream analysis, if an HFO is from a channel with both behavior response and SZ/AD response, the classification of these HFOs is determined by the mean of the probabilities from five models. Otherwise, the classification of the HFO is determined by the model where this specific HFO is in the test set.

4.2.13 Characterization of ecHFOs

We adopted the same procedure as the prior study [Zhang et al., 2022b] to determine which characterization in the time–frequency scalogram of ecHFOs differed from that of non-ecHFOs, that is, conducting a t-test across all pixels in the time–frequency scalogram (Figure 4.2(A)). The null hypothesis is that the values within the scalogram of ecHFO are not lower than those of non-ecHFO. We summarized the results by taking the average of the difference discovered in each patient to prove the characterization in the time–frequency plot generalizes at the population level (Figure 4.2(B)). In addition, to examine the influence of age, we divided the patient cohort into two groups (young patients [$age < 10$] and old patients [$age \geq 10$]) and repeated the same process.

4.2.14 Interpretability analysis of ecHFOs

Perturbation on time-frequency plot: we adopted a similar perturbation analysis as prior work [Zhang et al., 2022b] based on the characterization of ecHFO in the time–frequency plot (Figure 4.2(C)). We measured the change in probabilities as the ratio ($P_{original} -$

$P_{perturb})/P_{original}$ where $P_{original}$ is the original model confidence toward the event being ecHFO, and $P_{perturb}$ is the model confidence with perturbed input for the same event being ecHFO. We summarized the change in the ratio via the histogram for each patient across all the predicted ecHFOs (Figure 4.2(D)). A one-tailed t-test was also performed on the change of that output probability score to ensure that the change was significant and generalized well at the population level. Perturbation on amplitude coding plot: we adopted a data-driven approach to discover the effect of introducing a spike in the amplitude coding plot (Figure 4.3). As in our prior work [Zhang et al., 2022b], we built a detector to capture HFO with a spike (spk-HFO) trained on expert annotations (86.5% accuracy with an F1 score of 80.8%) [Zhang et al., 2022b]. The spk-HFO detector was applied to our data and predicted whether each HFO event was a spk-HFO or not. For all detected spk-HFOs in each patient, we separated these HFOs into upward spikes and downward spikes. The upward (downward) spikes are defined as spk-HFO events, with the average amplitude in the center 20 ms duration being greater (smaller) than the average amplitude in the peripheral 250 ms. We then created upgoing and downgoing spike templates by taking the average of all tracings. The corresponding amplitude coding plots of the upgoing and downgoing spike templates are built for perturbing the amplitude coding plots for each patient (Figures 4.3(A) and (C)). This perturbation was performed by adding the spike template to the original amplitude coding plot, and we measured the change in model confidence. We summarized the results (ratio change) via the histogram for each patient across all the predicted ecHFOs (Figures 4.3(B) and (D)). A one-tailed t-test was performed on the change of output probability to evaluate whether the change was significant and generalized well at the population level.

4.2.15 Time domain characteristic ecHFO

We reported three time-domain characteristics of the predicted ecHFOs and non-ecHFOs, which are amplitude, duration, and max frequency (Figure 4.4). The duration is defined as the time duration of the HFO predicted by the STE HFO detector. The amplitude is

defined as the maximum absolute value within the HFO interval, and the max frequency is defined as the largest frequency component within a window (0.2 s before and after) around the center of the HFO. We plotted the normalized histogram to compare the distribution of these three characteristics.

4.2.16 Comparison of resection ratios of HFOs to postoperative seizure outcomes

We estimated the probability of each patient’s surgical success (for the 14 patients who underwent resective surgery with known seizure outcomes at 24 months) based on the resection ratio of HFOs (number of resected HFOs/number of detected HFOs) as a classifier. We constructed the receiver operating characteristic (ROC) curve and calculated the area under the curve (AUC) values in the resection ratio of unclassified HFOs, ecHFOs, and HFOs with spikes (spk-HFOs). The classification of HFOs with spikes was performed using our prior algorithm [Zhang et al., 2022b]. Determination of the channel resection status (resected vs. preserved) was determined based on intraoperative pictures (pre-and post-resection) and also on post-resection brain MRI, based on discussion among a clinical neurophysiologist (HN), neurosurgeon (AF), and radiologist. A multiple logistic regression model incorporating the resection ratio of classified HFOs (non-ecHFOs and spk-HFOs) and complete resection of SOZ was also created. The surgical outcomes were determined 24 months after resection as either seizure-free or not seizure-free.

4.2.17 Statistical analysis

Above mentioned statistical calculations were carried out using Python (version 3.7.3; Python Software Foundation, USA). The deep neural network was developed using PyTorch (version 1.6.0; Facebook’s AI Research lab). Quantitative measures are described by medians with interquartile or means with SD. Comparisons between groups were performed using

chi-square for comparing two distributions and Student’s t-test for quantitative measures (in means with SD). All comparisons were two-sided and significant results were considered at $p < 0.05$ unless stated otherwise. Specific statistical tests performed for each experiment were described in each section.

4.3 Results

4.3.1 Clinical information (patient characteristics)

There were 19 patients (nine males, ten females) enrolled during the study period. The median age at surgery was 14 years (range: 3–20 years). The median electrocorticography monitoring duration was 4 d (range: 2–14 d), and the median number of seizures captured during the monitoring was 8 (IQ range: 4–25). There were 15 patients who underwent resection, and 14 patients provided postoperative seizure outcomes at 24 months (9 of 14 became seizure-free). Due to unknown postoperative seizure outcomes, patient 10 was removed from further HFO analyses.

4.3.2 Interictal HFO detection

A total of 63379 HFO events were detected (median 2258.5 events per patient) in 90 min EEG data from the 18 patients (patient 10 was removed for subsequent HFO analysis due to lack of seizure outcomes). There were 15111 HFOs recorded from the stimulated channels (48268 HFOs from non-stimulated channels). There were 7173 HFOs obtained from behavior channels (positive behavioral responses from cortical stimulation mapping and/or gamma mapping), and ‘behavior only’ channels (behavior channels without spontaneous spikes or SZ/AD) exhibited 3822 HFOs. There were 5317 HFOs obtained from ‘none’ channels (no spontaneous spikes and no behavioral responses with mapping). After the DL training, we noted that 22.9% (4547/19816) of the ecHFOs were HFOs with spikes, and

77.1% (30647/43563) of non-ecHFOs were HFOs with spikes ($p < 0.0001$, a chi-square test).

4.3.3 Time–frequency plot characteristics of ecHFO and non-ecHFOs

The analysis of the time–frequency plots from the selected individual patient demonstrated that ecHFOs had lower amplitude throughout the frequency band, including both ripples (80–250 Hz) and fast ripples (250–500 Hz) around the center point (0 ms, where HFOs were detected) than non-ecHFOs (Figure 4.2(A)). We inspected the averaged overall characterization and selected regions greater than 0.7 to define the salient characteristic difference between ecHFOs and non-ecHFOs. The selected region resembled a bell-shaped (Figure 4.2(B)). We observed the similar characteristic template in the time–frequency plot regardless of whether a young or old cohort was used for the training.

4.3.4 Perturbation analyses

By utilizing the bell-shaped template found in the time–frequency map, we observed that the bell-shaped perturbation on the time–frequency plot significantly increased the model prediction probability towards non-ecHFOs (mean increase ratio was 0.532 [95% confidence interval: 0.528–0.535], $p < 0.001$) (Figures 4.2(C) and (D)). Furthermore, we analyzed the effect of introducing a spike-like shape in the amplitude-coding plot. By introducing a downgoing or an upgoing spike in an ecHFOs event, the model confidence increased towards a non-ecHFO event (Figures 4.3(A) and (C)). On the population level, the prevalent probability increase in ecHFO events among all patients (Figures 4.3(B) and (D)) demonstrated the non-trivial model response by introducing a spike in the time domain (mean ratio increase 0.268 for a downgoing spike introduction [95% confidence interval: 0.263–0.272], and 0.320 for an upgoing spike introduction [95% confidence interval: 0.316–0.324], both with $p < 0.001$).

4.3.5 Feature characterization of ecHFOs vs. non-ecHFOs

We plotted the histogram of the peak frequency, amplitude, and duration of both ecHFO and non-ecHFO, respectively (Figure 4.4). The ecHFOs exhibited a smaller amplitude (p-value < 0.001), a lower max frequency (p-value < 0.001) and a trend of shorter HFO length (p-value $= 0.07$) than non-ecHFO. However, there was no decision boundary to clearly discriminate between ecHFO and non-ecHFO using each of these traditional features.

4.3.6 Inference of channel characteristics

After the model was trained, it assigned one prediction label for all of the HFOs in each channel, and we plotted the ecHFO ratio (proportion of ecHFOs divided by the number of all HFOs) and the model confidence distribution towards ecHFOs in different channels (Figure 4.5). In channels with a physiological response (only behavior responses with functional mapping), the ecHFO ratio was high, and confidence scores were skewed towards ecHFOs. Channels with a pathological response (spike and SZ/AD with stimulation) showed a lower ecHFO ratio, and confidence scores were skewed towards non-ecHFOs. Channels with no behavioral responses showed the confidence scores skewed towards non-ecHFOs. Notably, channels with both properties, in which both behavioral responses and spikes or/and SZ/AD, model confidence showed a uniform distribution, suggesting the presence of both ecHFOs and non-ecHFOs.

4.3.7 Clinical inference using postoperative seizure outcomes

We created the ROC curves using the HFO resection ratio to predict postoperative seizure freedom at 24 months ($n = 14$) (Figure 4.6). Using the resection ratio of HFOs showed acceptable prediction performance (AUC = 0.76; $p < 0.001$). The resection ratio of non-ecHFOs and spk-HFOs exhibited higher AUC values of 0.82 and 0.87 than unclassified HFOs, respectively ($p < 0.001$ for both). The performance was further augmented by using a

multiple regression model incorporating the resection ratio of such classified HFOs (non-ecHFOs and spk-HFOs) and the complete removal of SOZ (AUC = 0.933 for both classifiers, $p = 0.047$ and 0.043 , respectively).

4.4 Discussion

HFOs with similar frequency ranges emerge despite considerably different mechanisms [Jefferys et al., 2012, Zijlmans et al., 2012]. A traditional hypothesis-driven approach to separate physiological HFOs from pathological HFOs poses challenges because numerous yet-to-be-identified features must be considered. Simple engineering features of HFOs, including amplitude, frequency, and duration, do not appear to successfully separate pathological from physiological HFOs [Matsumoto et al., 2013, Cimbalka et al., 2018]. Visual classification of HFOs with or without spikes along with artifact removal (such as ringing) is commonly performed because HFOs with spike-wave discharges are considered representative of pathological HFOs [Weiss et al., 2016]. However, this task is time consuming and exhibits poor inter-rater reliability among human experts [Spring et al., 2017, Nariai et al., 2018]. Fast ripples (250–500 Hz) might more specifically localize EZs than ripples do (80–250 Hz), but their detection rate is much lower than ripples [Roehri et al., 2018]. Correcting the HFO detection rate with region-specific normative values seems a reasonable approach [Kuroda et al., 2021, Frauscher et al., 2018, Guragain et al., 2018, Zweiphenning et al., 2022b], but this does not determine each HFO event as either pathological or physiological.

In the study presented in this chapter, we leveraged the robust clinical definition of the EC with cortical stimulation and gamma-related language mapping to identify cortical areas generating physiological HFOs (represented by ecHFOs). Our approach using a DL-based algorithm is distinct from other approaches. Based on our hypothesis that physiological HFOs look morphologically different from pathological HFOs, leveraging DL’s ability to analyze imaging input (transformed from EEG time-series data) seemed logical. By training

a DL model with labels based on functional cortical mapping results, we then investigated the salient features of the physiological HFOs through the DL algorithm. The novel findings include that we found a bell-shaped template in the time–frequency plot as the discriminating feature of ecHFOs. More specifically, ecHFOs would have lower signal amplitude at the center of the HFO onset across the frequencies (including both ripple and fast ripple band) and lower signal amplitude in the low frequency band throughout the time window than non-ecHFOs (representing pathological HFOs). The ecHFOs generally had a lower peak frequency, a smaller amplitude, and tended to have a longer duration than non-ecHFOs. However, these time-domain characteristics showed significant overlaps between ecHFOs and non-ecHFOs, and none of such simple engineering features could clearly separate them, consistent with the previous studies [Matsumoto et al., 2013, Cimbalnik et al., 2018]. Once the DL model was trained, we verified specific morphological features of ecHFOs via perturbation analysis. Insertion of spike templates significantly decreased ecHFO probability, which confirmed the traditional knowledge that HFOs with spikes are most likely pathological HFOs.

Our findings have significant clinical implications. Using the resection ratio (removed HFOs/ detected HFOs) of non-ecHFOs (likely representing pathological HFOs), we demonstrated that the prediction of postoperative seizure outcomes significantly improved compared to using the uncorrected HFO resection ratio (AUC of 0.82, increased from 0.76). Combined with the current clinical standard of complete resection of SOZ, the AUC further improved to 0.93. The above findings suggest that the classification of HFOs (physiological vs. pathological) can potentially improve the clinical utility of HFOs in guiding resection, and this can still be combined with our current clinical standard of capturing habitual seizures to decide the resection margin. A similar improvement in AUC was observed using spk-HFOs (AUC of 0.87, increased from 0.76), which seemed even better than non-ecHFOs. This result suggests that spk-HFOs may better represent pathological HFOs than non-ecHFO. However, HFOs with spikes need to be evaluated by human experts, which exhibits concerns about inter-rater reliability [Nariai et al., 2018]. There is also no objective definition of spikes. So,

analyzing ‘HFOs with spikes’ may not be as straightforward as it sounds. At least, the determination of ecHFOs can be made by objective functional mapping results. We are planning more studies to investigate the additional clinical utility of ecHFOs, such as identifying EC by analyzing ecHFOs, without the need for stimulation mapping.

A recent study demonstrated that a cortical location could produce different types of HFOs, such as both physiological and pathological HFOs [Liu et al., 2018]. We probed this point by investigating cortical sites with both EC characteristics (behavioral responses with stimulation mapping or/and gamma activity with language tasks) and pathological properties (spontaneous spikes or/and AD/SZ). The ecHFO ratio was between cortical sites with only behavioral responses and that with spikes or/and SZ/AD only. The confidence scores of HFO-type prediction are diffusely distributed from ecHFOs to non-ecHFOs, without clear bimodal distribution. If such ‘both’ channels produce morphologically distinct physiological and pathological HFOs, one would expect to see a clear bimodal distribution. Our findings suggest that in cortical sites with both pathological and physiological properties, the morphologies of HFOs are heterogeneous. This may be a limitation of using macroelectrodes since multiple local generators might influence each HFO morphology.

There are several limitations in the study presented in this chapter. We have analyzed only 18 patients to define physiological HFOs. Also, postoperative outcomes were correlated in only 14 patients. Although we analyzed an extended EEG dataset (90 min from each subject) to maximize the number of HFOs for analysis, we will need to analyze more patients to have a definitive conclusion. We demonstrated the discovered ecHFOs’ characteristics were consistent even trained by different age groups. Still, a robust conclusion should be drawn from analyses of a larger number of subjects to examine the effect of age. Although we assumed that HFOs from the healthy brain regions look morphologically similar, such characteristics may be different in the different brain regions. With more subjects, we will be able to characterize the difference in ecHFOs generated in functionally distinct cortical areas, such as the somatosensory, language, and visual cortex. In addition, we analyzed only sleep

EEG to maximize the number of HFOs for analysis. States of consciousness (including awake and different sleep stages) and vigilance levels may affect HFO morphology [von Ellenrieder et al., 2017]. Lastly, with an increasing number of diagnostic stereotactic EEG (SEEG) studies, it is of interest to investigate how HFO morphologies via SEEG would look different compared to strip/grid-sampled HFOs.

In summary, we proposed to use a weakly supervised DL algorithm to characterize physiological HFOs using robust clinical outcomes (functional cortical mapping results) as training labels. Although future work to include a larger number of subjects will be needed, this method provided salient morphological features of physiological HFOs. Our results also suggested that this DL-based HFO classification, once trained, might help separate pathological from physiological HFOs, and efficiently guide surgical resection using HFOs.

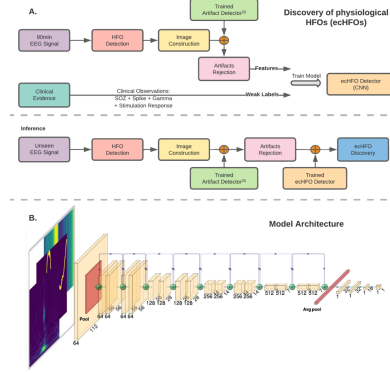


Figure 4.1: **Processing workflow:** (A) the detected HFO (by STE detector) is first filtered by the artifact detector that was developed in our previous study [Zhang et al., 2022b]. The constructed image features along with the channel-wise clinical information, are fed into the convolutional neural network to train the model (ecHFO detector). In the real application of the model (inference), an unseen EEG signal is sent into the pipeline, and the STE detector is applied to detect all HFOs. Then we construct the image feature based on the detected HFO. The same artifacts detector [Zhang et al., 2022b] is applied to reject the artifacts, and then the real HFO is filtered by the ecHFO detector. The detector will output a score from zero to one representing the confidence of each HFO being ecHFO. HFO feature representation and model architecture (B): we captured the time–frequency domain features as well as signal morphology information of the HFO window via three images. The time–frequency plot (scalogram) was generated by continuous Gabor wavelets ranging from 10 Hz to 500 Hz. The EEG tracing plot was generated on a 2000×2000 image by scaling the time-series signal into the 0–2000 range to represent the EEG waveform’s morphology. We constructed another two images from the original EEG signal serving as the additional inputs to the network. The amplitude-coding plot was generated to represent the relative amplitude of the time-series signal: for every time point, the pixel intensity of a column of the image represented the signal’s raw value at that time. These three images were resized into the standard size (224×224), serving as the input to the neural network. We used ResNet 18 with a modified output layer for binary classification. The weights in the convolution layers are frozen and serve as feature extractors in the convolutional neural network.

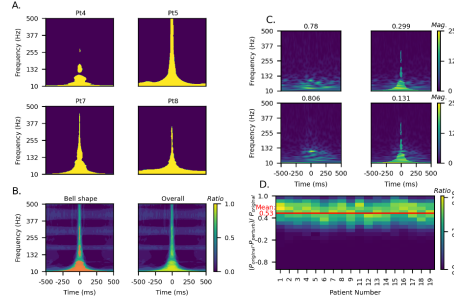


Figure 4.2: **Characteristics in the time–frequency plot of ecHFOs against non-ecHFOs:** (A) The time–frequency plot characteristics of the eloquent cortex and non-eloquent cortex HFOs for Pt 4, 5, 7, and 8. The yellow-colored regions in the figure stood for the pixels, where the power spectrum of ecHFOs is statistically lower than (p-value below 0.05 from the one-tailed t-test) non-ecHFOs. The figure showed one set of clearly interpretable distinguishing features between ecHFOs and non-ecHFOs: the ecHFOs generally have lower power at lower frequencies during the HFO event (center part along the time axis), panel (B-Right) was generated by taking the average of the individual binary images from each of the 18 patients. It showed the distinguishing features are also significant at the population level. The feature can be assembled as a ‘Bell-Shape’ if we take the region >0.7 in the plot averaging all of the time–frequency plot characteristics (red color). (B-Left). (C) The model’s response to the Bell-shaped perturbation on the time–frequency plot. We provide two examples of perturbation for ecHFO events in Pt 3. Each row presents one example and the first column indicates the original time–frequency plot while the second indicates the perturbed time–frequency plot based on the Bell shape perturbation. The prediction value of the model changed from 0.780 (therefore originally predicting it as non-ecHFO) to 0.299 (thus a change of 0.481), implying that the perturbed HFO would correspond to a non-ecHFO. (D) The change in model confidence in population level. Each column (along the y-axis) is a histogram of the percentage of change in confidence for one distinct patient. It shows the frequency distribution of confidence changes after adding the bell shape perturbation to the time–frequency plot to all classified ecHFOs for the given patient. The change in confidence level is significant, with an average of 0.532 with a 95% confidence interval $[0.528, 0.535]$ noted as the solid red line in the histogram. HFO = high-frequency oscillation; ecHFOs = eloquent cortex HFOs; non-ecHFO = non-eloquent cortex HFOs

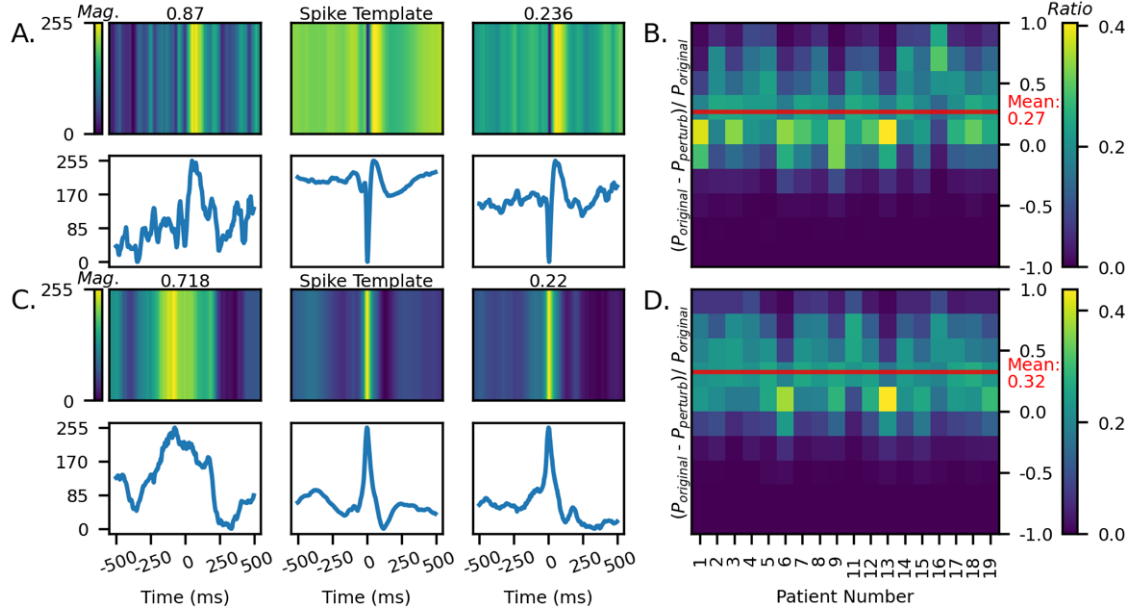


Figure 4.3: **The model's responses to injecting a spike-like feature into the amplitude-coding plot:** Examples of introducing a downgoing (A) and upgoing (C) spike feature into classified non-ecHFO events. These demonstrate that on the introduction of a spike-like perturbation, the model predicts higher confidence towards non-ecHFOs (B, D). Subfigure (A) shows the amplitude encoding image before perturb, spike template, and the after-perturb (top row), and the corresponding time-series signal with downgoing spike perturbation (bottom row). Similarly, subfigure (C) shows the same information on a different classified ecHFO but with upgoing spike perturbation. For each patient, we compute a histogram for the distribution of the change in confidence (B) for downgoing spike perturbation. The same steps are repeated for upgoing spike perturbation, and the results are shown in (D). The percentage of change in confidence for both up-and-downgoing spike perturbation is significantly greater than zeros, with means downgoing: 0.268 (95% confidence interval [0.263, 0.272]) and upgoing: 0.320 (95% confidence interval [0.316, 0.324]), which are noted as solid red lines in each histogram. HFO = high-frequency oscillation; ecHFOs = eloquent cortex HFOs; non-ecHFO = non-eloquent cortex HFOs.

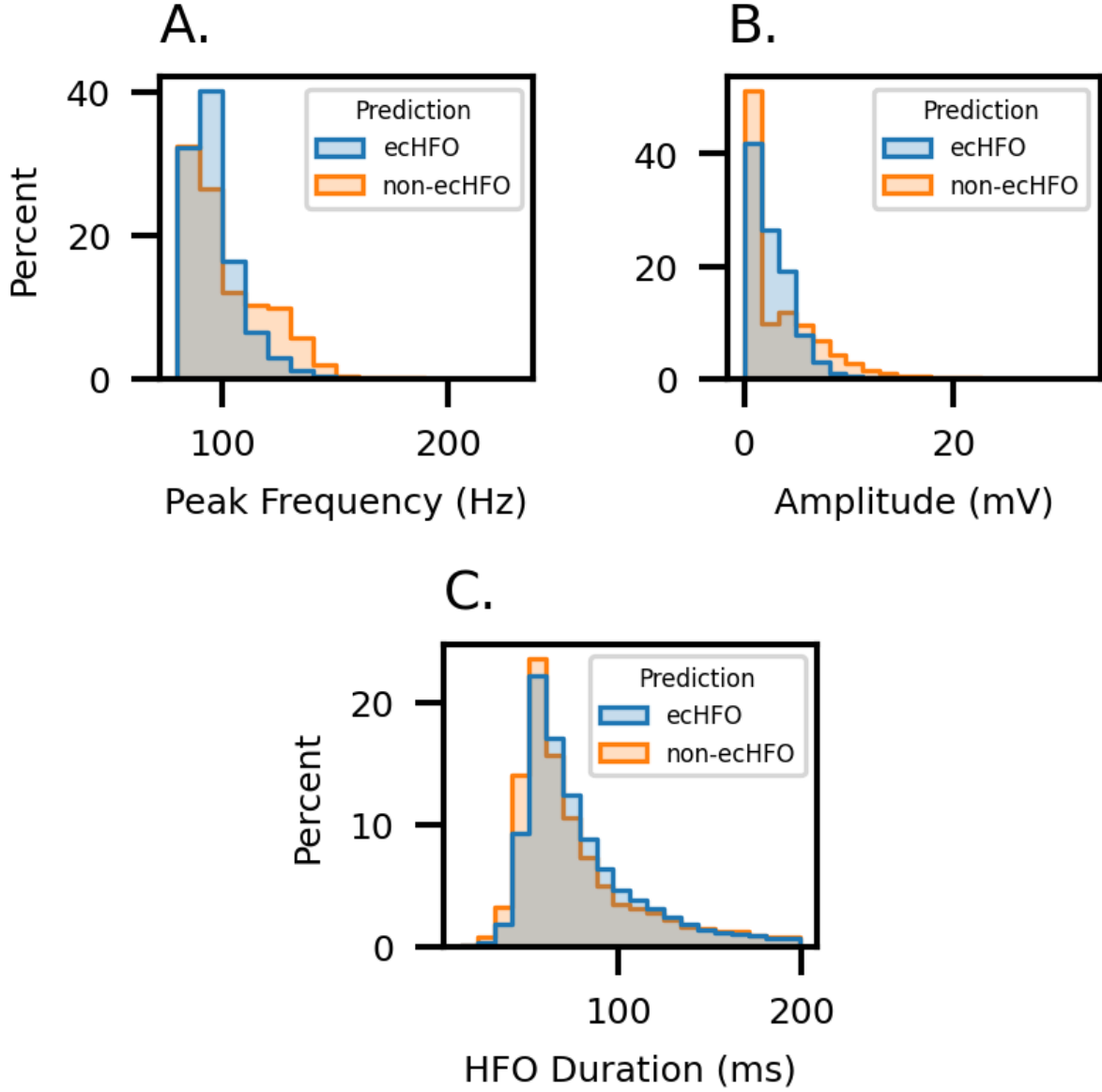


Figure 4.4: **Traditional feature characterization of ecHFO and Non-ecHFO:** The normalized histogram of peak frequency (A), amplitude (B), and duration (C) of predicted ecHFO. (A) The ecHFO generally has a lower peak frequency than the non-ecHFO (p-value < 0.001). (B) The ecHFO generally has a smaller amplitude than the non-ecHFO (p-value < 0.001). (C) The ecHFO generally has the trend of having a longer HFO than the non-ecHFO (p-value = 0.07). However, there are no clear decision boundaries that can be drawn to clearly discriminate between ecHFO and non-ecHFO using each of these traditional features.

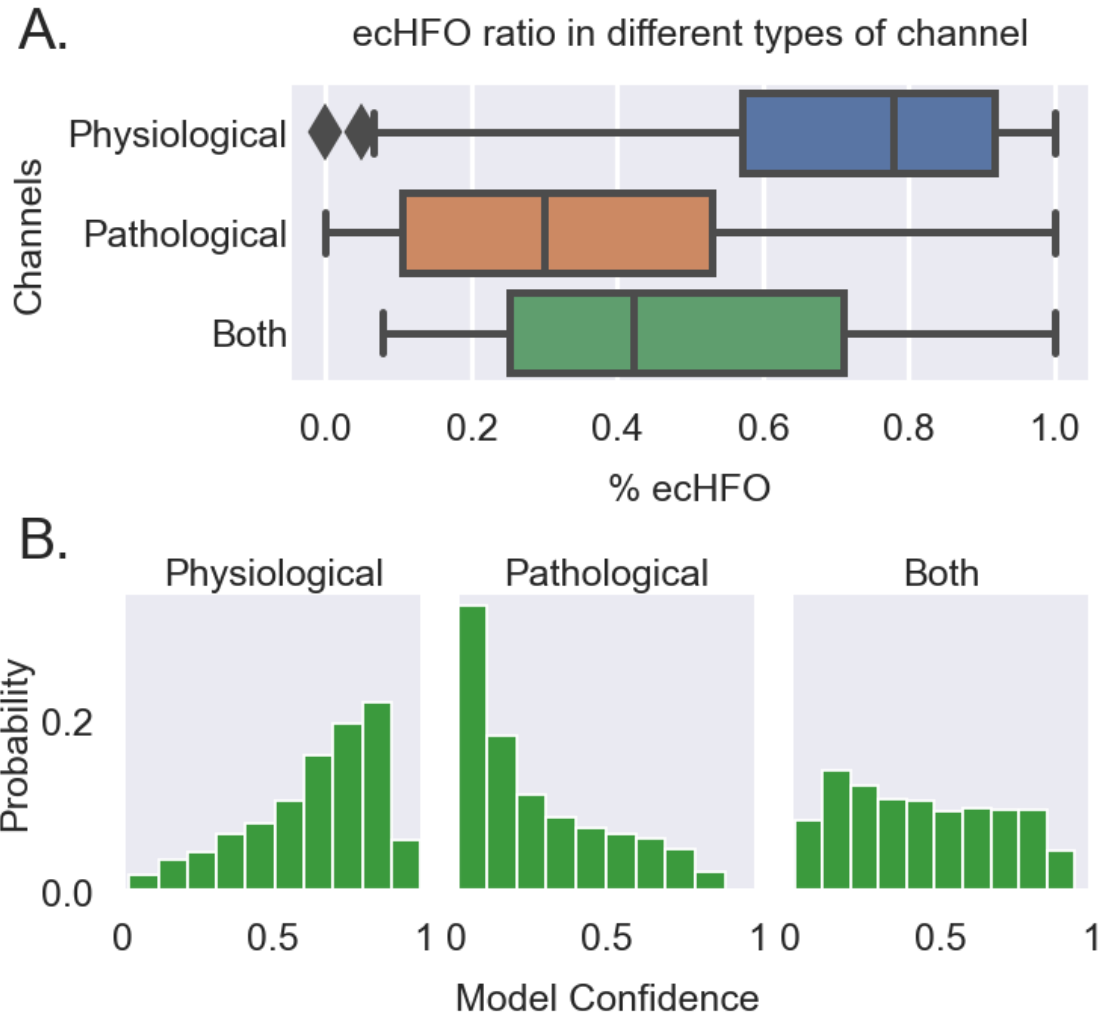


Figure 4.5: **Inference of channel characteristics:** (A) the ratio of ecHFOs (ecHFOs/total HFOs) in different types (physiological: behavior or gamma only; pathological: SZ/AD or spike only; both: both physiological and pathological) of channels from all patients ($n = 14$) is plotted in box plots. The percentage of ecHFOs was higher in physiological channels than that in pathological channels. (B) The model confidence distribution of each individual ecHFOs in channels with physiological, pathological, and both categories are shown. The distribution in both channels is closer to a uniform distribution.

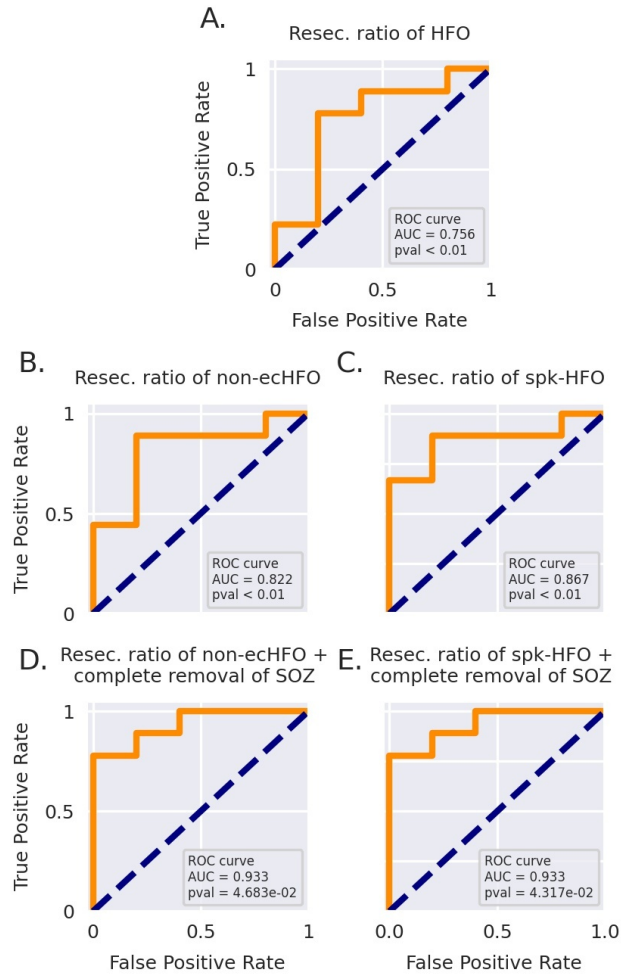


Figure 4.6: The accuracy of prediction models incorporating HFO resection ratio:

We constructed postoperative seizure outcome prediction models using the HFO resection ratio derived from EEG data ($n = 14$). Each receiver-operating characteristics (ROC) curve delineates the accuracy of seizure outcome classification of a given model, using the area under the ROC curve statistics. (A) Unclassified HFO resection ratio was used as a single classifier. (B) non-ecHFO resection ratio was used as a single classifier. (C) non-ecHFO resection ratio was used as a single classifier. (D, E) A multiple regression model incorporating the resection ratio of HFOs (non-ecHFOs or spk-HFO) and complete removal of the SOZ (yes or no) was used, which demonstrated further improved predictive value of postoperative seizure outcomes. HFO = high-frequency oscillation; non-ecHFO = non-eloquent cortex HFOs; spk-HFO = HFO with spike; SOZ = Seizure onset zone.

CHAPTER 5

Novel interictal biomarker of epileptogenic zone

5.1 Introduction

In Section 1.4.2, we highlighted the pivotal role of accurately identifying and resecting the epileptogenic zone (EZ) in treating focal drug-resistant epilepsy. Recent years have witnessed considerable efforts towards developing advanced signal analysis techniques aimed at improving the localization of the EZ. As focal epilepsy is increasingly recognized as a "network disease," the field of network neuroscience has provided promising insights for presurgical evaluation. However, despite these advancements, several persistent barriers impede the translation of research findings into clinical practice [Sinha et al., 2022].

Most of the research to date has focused on examining network properties during ictal segments, which are very short in duration and occur infrequently during the 2 week long hospital monitoring. Therefore, if we can develop a network based approach to unearth biomarkers of EZ from segments when the patient is not having seizures, known as interictal segments, then it has potential to significantly shorten invasive procedures and reduce the necessity of recording numerous spontaneous seizures (refer to [Lagarde et al., 2022] for an extensive review). Although interictal segments are much longer in duration and always available in abundance, the synchronization effect if any are much more subtle in this segment. Therefore, a dynamic network modeling approach is required to discover interictal biomarkers of EZ.

In this chapter, we model the interictal interactions amongst the iEEG/SEEG channels

using a dynamic synchronization network and leverage the temporal dynamics of the network properties, both local and global, to accurately identify the channels that drove all the seizure episodes during the 2 week long hospital monitoring stage (SOZ) and the channels that did not lead to seizures in hospital monitoring but can lead to seizures (PSOZ) if given a long enough observation window ($>> 2$ weeks). The ability of our network model to accurately estimate the EZ ($SOZ \cup PSOZ$) from interictal segments with duration as short as 5 minutes holds great promise of cutting down invasive hospital monitoring time significantly, which is fraught with risk of infection.

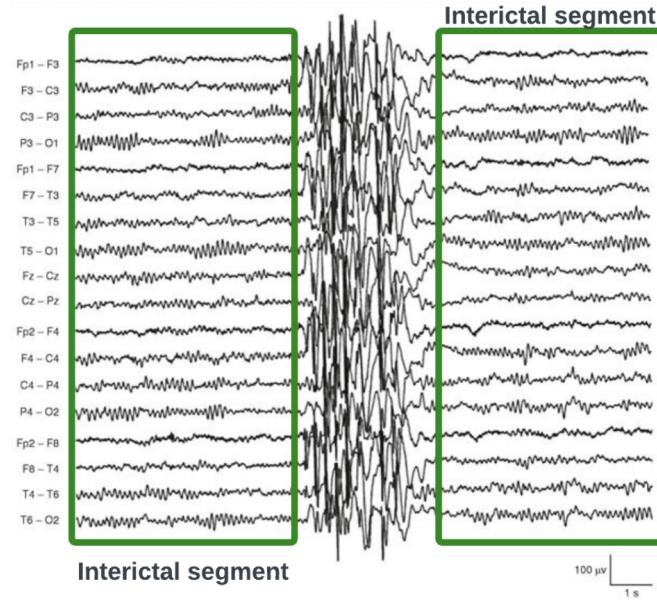


Figure 5.1: **Interictal segment [Fernández-Torre, 2010]:** Interictal segment (highlighted in green), when the patient is not having seizures, is always available in abundance and much longer in duration. However, as you can see that the synchronization effects in the interictal segments if any are much more subtle

5.2 Materials

5.2.1 Patient Cohort

This study employed a retrospective cohort design. We consecutively recruited 159 children (under the age of 30) who were suffering from medically refractory epilepsy—characterized by monthly or more frequent seizures and the inefficacy of more than three anti-seizure medications—from the Pediatric Epilepsy Program at UCLA and Wayne State. These children had intracranial electrodes implanted for epilepsy surgery planning and subsequent cortical resection between August 2016 and August 2018.

5.2.2 Standard protocol approvals, registrations, and patient consent

The UCLA’s institutional review board approved the involvement of human subjects and dispensed with the necessity for written informed consent. All the tests were regarded as clinically relevant for patient care, and all the EEG data used in this retrospective study was anonymized before it was extracted and analyzed. This study was not classified as a clinical trial and was not registered in any public registry.

5.2.3 Patient evaluation

All children referred during the study period, who were suffering from medically refractory epilepsy, underwent a systematic pre-surgery evaluation. This included inpatient video-EEG monitoring, high-resolution brain magnetic resonance imaging (MRI) of 3.0 Tesla, and 18 fluoro-deoxyglucose positron emission tomography (FDG-PET), coupled with MRI-PET co-registration. The boundaries and scope of the resections were primarily based on the seizure onset zone (SOZ), which is clinically identified as the areas initially showing continuous rhythmic waveforms at the beginning of typical seizures. In certain instances, to avoid causing unacceptable neurological damage, the seizure onset zones were not fully resected.

5.2.4 Subdural electrode placement

Surgical implantation was performed for macroelectrodes, which included platinum grid electrodes with an intercontact distance of 10 mm, and depth electrodes, made of platinum with a 5 mm intercontact distance. The total number of electrode contacts for each patient ranged from 40 to 128, with a median of 96 contacts. The positioning of these intracranial electrodes was primarily guided by the results of scalp video-EEG recording and neuroimaging studies. All electrode plates were sewn to neighboring plates, the dura mater’s edge, or both, in order to minimize the movement of subdural electrodes post-placement.

5.2.5 Acquisition of three-dimensional (3D) brain surface images

We collected high-resolution 3D magnetization-prepared rapid acquisition with gradient echo (MPRAGE) T1-weighted images of the whole head preoperatively. Utilizing post-implant computed tomography (CT) images, we created a FreeSurfer-based 3D surface image with the locations of the electrodes directly identified on the brain surface [Zhang et al., 2022b]. Furthermore, digital photographs were captured intraoperatively before the dura was closed to improve the spatial precision of electrode placement on the 3D brain surface. When the area was re-exposed for resective surgery, we visually verified that the electrodes had remained in place, matching the digital photo taken during the electrode implantation surgery.

5.2.6 Intracranial EEG (iEEG) recordings

The intracranial EEG (iEEG) data was recorded with a digital sampling frequency of 2,000 Hz using the Nihon Kohden Systems (Neurofax 1100A, Irvine, California, USA). For each participant, a consistent 90-minute EEG segment was chosen, ensuring it was at least two hours before or after any seizures, prior to tapering of anti-seizure medications, and before cortical stimulation mapping. This mapping usually took place two days post-implant. All the iEEG data used in the study were extracted from the clinical EEG recording.

5.3 Dynamic directed weighted interictal networks

In order to capture the temporal dynamics of the interictal interactions amongst the channels, we split the interictal iEEG data into one second long non-overlapping data segments. For each iEEG segment, we construct a frequency domain synchronization network. Assuming the interictal segment is T seconds long, we will have T such synchronization networks. We describe the construction of one such network in the following section.

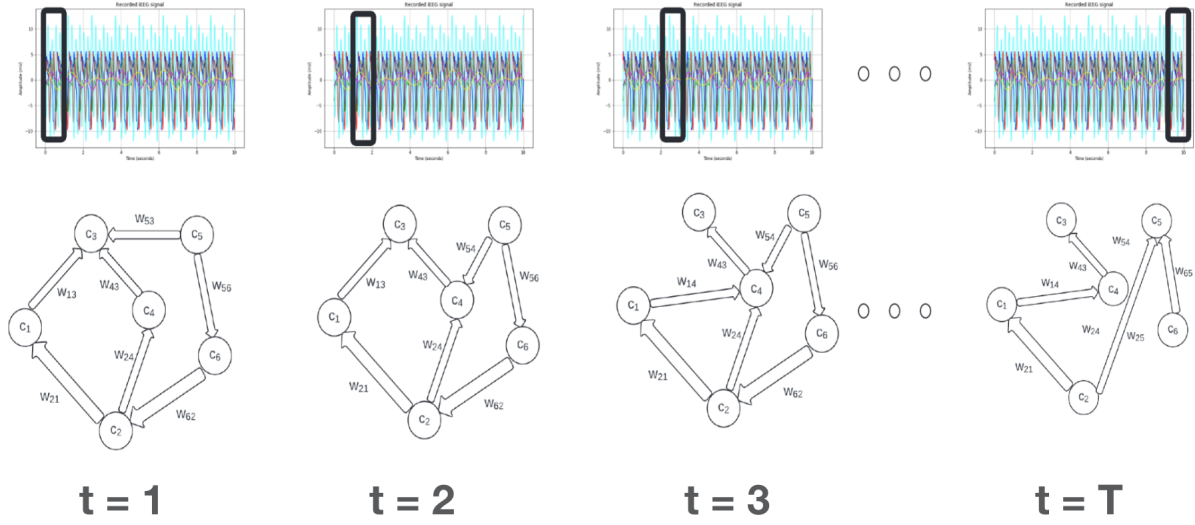


Figure 5.2: **Sequence of synchronization networks:** One second long iEEG segment is used to construct one synchronization network. Assuming the interictal segment is T seconds long and non-overlapping one second long sliding window approach is used, we have a sequence of T synchronization networks

5.3.1 Frequency domain synchronization network

We model the interictal interactions amongst the iEEG channels in the frequency domain. Below is a flow diagram showing the steps involved in constructing the frequency domain synchronization networks: i) select a specific frequency band to filter the intracranial

EEG(iEEG) segment and transform the filtered signal into its frequency domain representation and ii) subsequently, utilize the frequency domain representation to build a synchronization network to capture the interictal interactions amongst the channels. In the following sections, we describe each step in detail

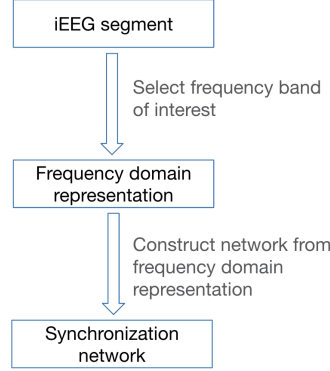


Figure 5.3: **Synchronization network workflow:** The first step involves selecting a frequency band of interest to filter the iEEG segment and transform it into the frequency domain. Once we have the frequency domain representation, then we construct a synchronization network to capture the interictal interactions between the iEEG channels

5.3.1.1 Frequency bands of interest in Epilepsy

The analysis of frequency bands within intracranial EEG (iEEG) data has provided valuable insights into the pathophysiology of epilepsy, aiding in diagnosis, treatment, and research endeavors. For example, increased delta power may indicate abnormal neuronal synchronization, especially during epileptic seizures originating from deep brain structures like the thalamus while alterations in alpha rhythms, such as alpha power suppression, have been observed in the preictal and ictal states, suggesting their involvement in seizure generation and propagation. Similarly, ripple band and fast ripple band activity has been hypothesized to be a biomarker of EZ. Although, the framework that we are presenting is general and can be used for any frequency bands of interest we will be studying synchronization in High

Gamma, Ripple and Fast Ripple band.

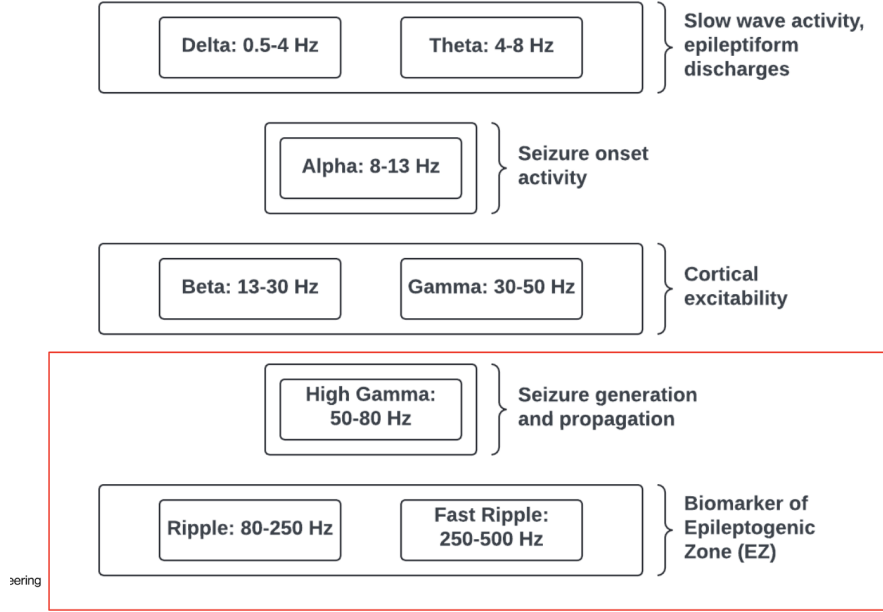


Figure 5.4: **Frequency bands:** Epilepsy research has focused on specific frequency bands depending on the research question they are trying to answer. We will be studying synchronization in High Gamma, Ripple and Fast Ripple band.

5.3.1.2 Network construction

The network consist of N nodes where N is the number of channels in iEEG. For each channel, we compute it's power and phase spectrum using Fast Fourier Transform (FFT). Since the goal of the network is to capture synchronous activity between channels, so we capture the synchronous activity between channels i and j through phase coupling coefficient σ_{ij}^ϕ , where

$$\sigma_{ij}^\phi = -\sigma_{ji}^\phi = \frac{1}{K} \sum_{m=1}^K \sin(\phi_{j,m} - \phi_{i,m}) \quad (5.1)$$

and K is the number of discrete frequency bins in the frequency band of interest. However, we want to capture interactions during episodes of excessive synchronization so we scale

the phase coupling coefficient by the total combined power in channels i and j , to get the synchronization coefficient between channels i and j

$$d_{ij} = \sigma_{ij}^{\mathcal{P}} \cdot \sigma_{ij}^{\phi}, \quad (5.2)$$

where

$$\mathcal{P}_i = \sum_{m=1}^K |P_{i,m}|^2 \quad (5.3)$$

$$\mathcal{P}_j = \sum_{m=1}^K |P_{j,m}|^2 \quad (5.4)$$

$$\sigma_{ij}^{\mathcal{P}} = \mathcal{P}_i + \mathcal{P}_j. \quad (5.5)$$

Hence, high total power along with large phase coupling coefficient will lead to large synchronization coefficient between nodes i and j . Having computed the synchronization coefficient, the next step is to map it into edge weights. We want the mapping to satisfy the following two criterion:

- Edge weights should be bounded between 0 and 1 as it will facilitate network property computations
- Edge weights should be a monotonically increasing function of the absolute value of the synchronization coefficient because high synchronization coefficient should map to a large edge weight

The following mapping from synchronization coefficient to edge weight satisfies both the criterion

$$w_{ij} = 1 - e^{-|d_{ij}|} \quad (5.6)$$

To emphasize substantial synchronization, we retain only the top 10 percent (ranked in terms of edge weights) of edges and discard the rest by setting their edge weights to zero. The final step is to assign directionality to the edges. We use the total phase method to assign the

directionality. The total phase in node j is given by

$$\eta_j^\phi = \sum_{m=1}^K \phi_{j,m} \quad (5.7)$$

where K is the number of discrete frequency bins in the given frequency band. Analogously, the total phase in node i is given by

$$\eta_i^\phi = \sum_{m=1}^K \phi_{i,m} \quad (5.8)$$

where K is the number of discrete frequency bins in the given frequency band. If there is a greater total phase in node j than node i , that is

$$\eta_j^\phi > \eta_i^\phi \quad (5.9)$$

then the weighted edge points from node i to node j ($i \rightarrow j$). If there is a greater total phase in node i than node j , that is

$$\eta_i^\phi > \eta_j^\phi \quad (5.10)$$

then the weighted edge points from node j to node i ($j \rightarrow i$).



Figure 5.5: **Edge directionality:** If the total phase in channel j leads the total phase in channel i then we assign a directionality from i to j and if the total phase in channel j lags the total phase in channel i then we assign a directionality from j to i

5.4 Seizure onset zone identification using dynamic interictal network properties

Recall, that our goal is to model the interictal interactions amongst the iEEG/SEEG channels using a dynamic synchronization network and leverage the temporal dynamics of the network properties, both local and global, to accurately identify the EZ ($SOZ \cup PSOZ$). Specifically, we want to build a classifier that takes in a node feature vector, constructed from the interictal network property dynamics, and outputs the probability of the node being a part of the seizure onset zone (SOZ).



Figure 5.6: **SOZ classifier setup:** The SOZ classifier takes in as input a node feature constructed from the interictal network property dynamics and outputs a probability of the node being an SOZ

In the next section, we explain how we construct the node feature vector.

5.4.1 Node feature vector

We can observe from the interictal network snapshots shown in Figures 5.7, 5.8, 5.9 that they enter into a state of large scale synchronization where a star like network evolves with either a SOZ node or a resected Non-SOZ node being the hub. Since epileptic seizures are believed to be caused by large scale synchronization effects, so these hypersynchronous states and the nodes driving these states are of particular importance to us for SOZ identification.

A network feature which has been used to study synchronization in networks [Bertrand and Moonen, 2013] is Fiedler eigenvalue which is the second smallest eigenvalue of the laplacian matrix L , where

$$L = D - A,$$

D is the degree matrix and A is the adjacency matrix. In our context of interictal networks, we also observe a sharp rise in fiedler value during episodes of large scale synchronization as shown in Figure 5.10

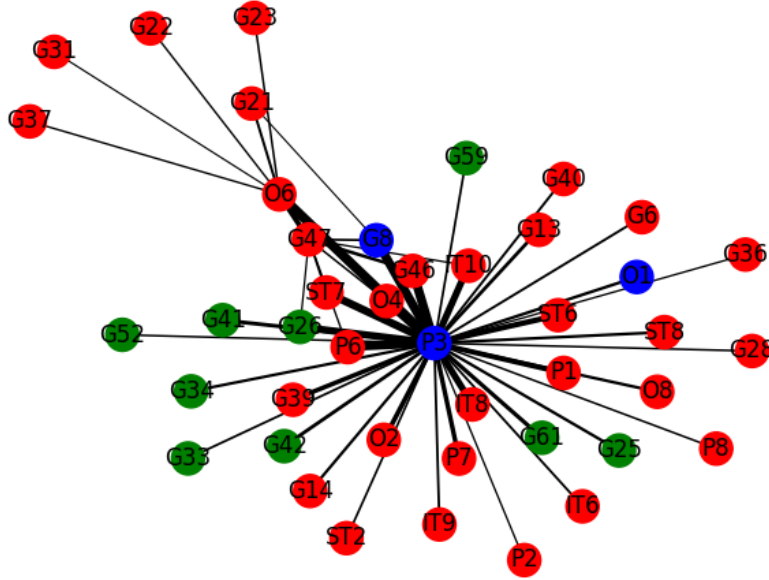


Figure 5.7: **Hypersynchronous state with SOZ node P3 as hub:** A snapshot of the synchronization network constructed from a one second long interictal segment with blue nodes representing channels that drove seizures in hospital monitoring, red nodes representing channels that did not drive seizures in hospital monitoring but were resected by the doctors, and green nodes representing channels that neither drove seizures in hospital monitoring nor resected by the doctors. We can observe the emergence of a hypersynchronous state, characterized by a star like network, with a SOZ node P3 as the hub

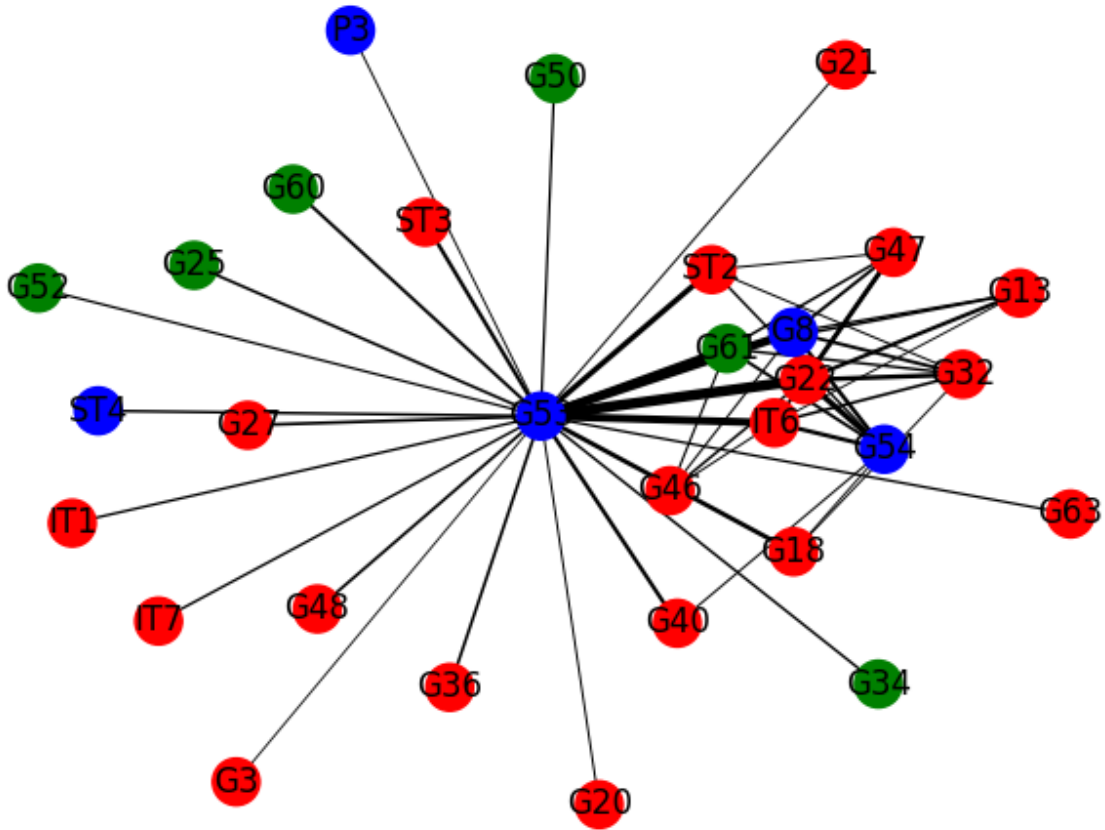


Figure 5.8: **Hypersynchronous state with SOZ node G53 as hub:** Another snapshot of the synchronization network constructed from a different one second long interictal segment with blue nodes representing channels that drove seizures in hospital monitoring, red nodes representing channels that did not drive seizures in hospital monitoring but were resected by the doctors, and green nodes representing channels that neither drove seizures in hospital monitoring nor resected by the doctors. We can observe the emergence of a hypersynchronous state, characterized by a star like network, with a SOZ node G53 as the hub

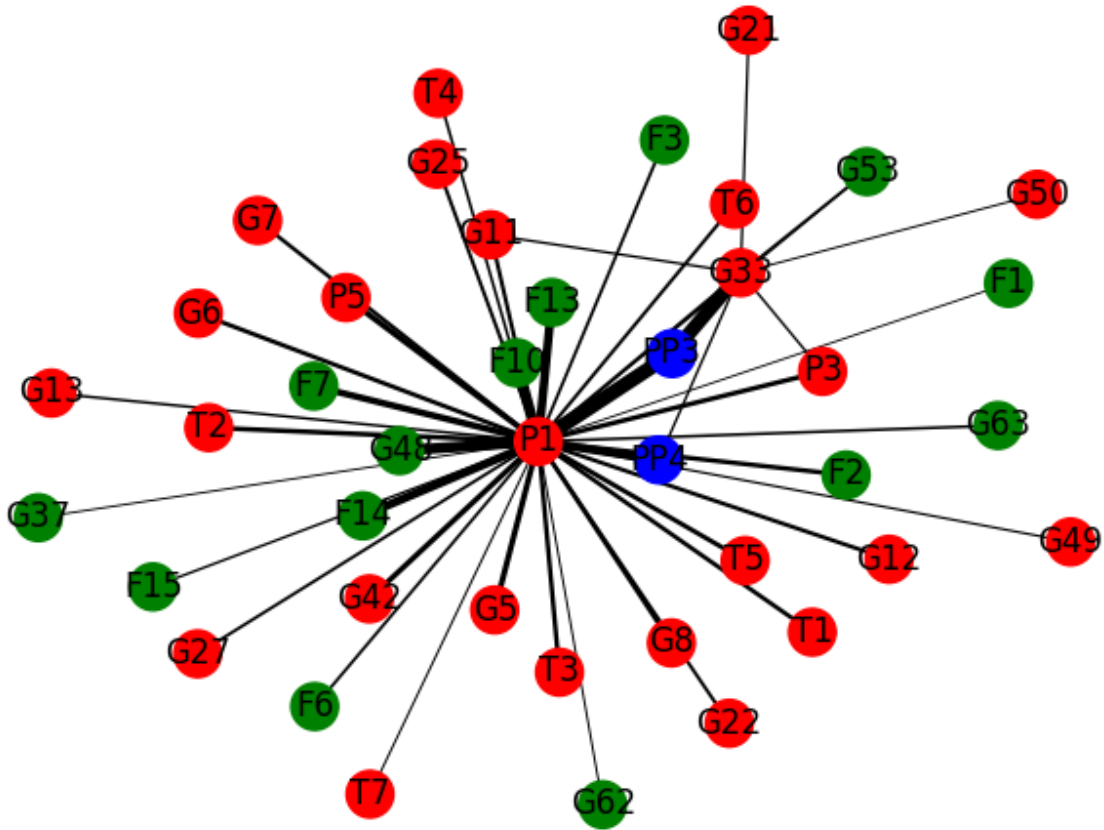


Figure 5.9: **Hypersynchronous state with Resected Non-SOZ node P1 as hub:** Another snapshot of the synchronization network constructed from a different one second long interictal segment with blue nodes representing channels that drove seizures in hospital monitoring, red nodes representing channels that did not drive seizures in hospital monitoring but were resected by the doctors, and green nodes representing channels that neither drove seizures in hospital monitoring nor resected by the doctors. We can observe the emergence of a hypersynchronous state, characterized by a star like network, with a resected Non-SOZ node P1 as the hub

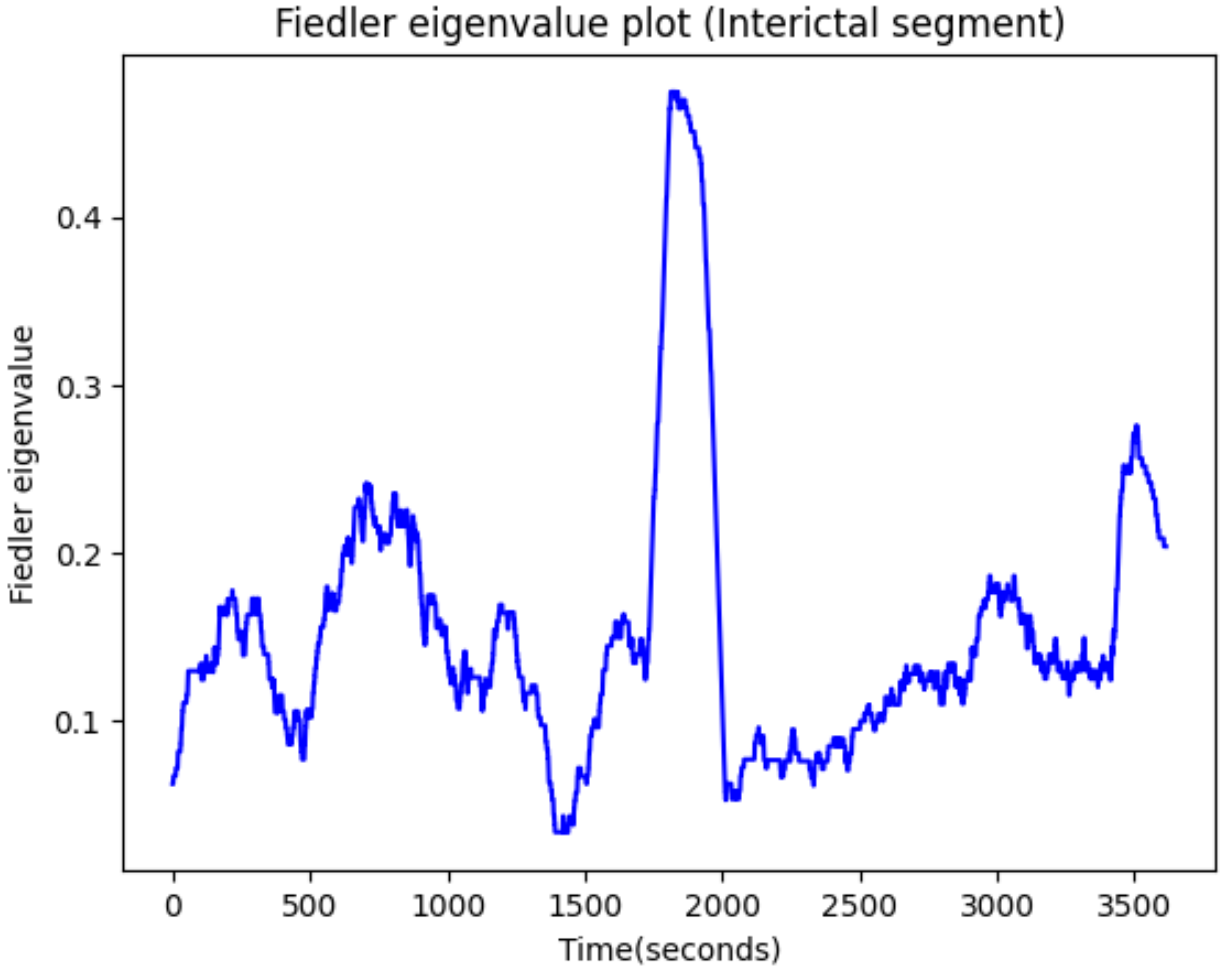


Figure 5.10: **Sharp rise in fiedler eigenvalue during episodes of large scale synchronization:** The Fiedler eigenvalue is graphed over time throughout the interictal period. Examination of the plot reveals distinct spikes in the Fiedler value at certain intervals. These intervals align with periods of hypersynchronous interictal network states, thereby confirming the efficacy of the Fiedler value in capturing such states.

The most basic node feature to consider for capturing hub like phenomenon is in degree and out degree. Also, we want to put particular emphasis on the node degrees during these episode of large scale synchronization so we also add the scaled degrees as node features, in degree and out degree scaled by fiedler eigenvalue. This scaling will amplify the node degrees

during episodes of such large scale synchronization.

Another feature to consider at a node level is Eigencentality because it has been shown in epileptic network research that eigencentality of a node is correlated to the rate of pathological events generated by that node during seizure episodes [Zimmermann et al., 2024]. It measures a node's importance not just by the number of connections it has, but also by the importance of the nodes it is connected to. Nodes with high centrality are not only well connected but are also connected to other important nodes in the network. It is computed by solving the set of consistency equations,

$$\lambda c_i = \sum_{j \neq i} a_{ij} c_j, \quad \forall i \in N,$$

where c_i and c_j are the eigencentralities of nodes i and j respectively. In matrix form, the equation is

$$\lambda c = A^T c \tag{5.11}$$

where c is the eigenvector centrality vector and A is the adjacency matrix. By Perron-Frobenius theorem, c is well-defined for any connected network and λ is the largest eigenvalue of A^T with corresponding eigenvector c .

In summary, for one interictal network, each node in the network is represented by a 5 dimensional feature vector

$$f(v_i) = \begin{bmatrix} in(v_i) \\ out(v_i) \\ f_{in}(v_i) \\ f_{out}(v_i) \\ ei(v_i) \end{bmatrix}$$

Since, we have a sequence of interictal networks so each node in the network is represented by a 5 dimensional time series representing the time evolution of this node across the feature space. Therefore, we need to summarize the 5 dimensional time series to get a feature vector

for a node. To gain some insight into what type of summary statistics we should look into, we have plotted the time evolution of out degree (Figure 5.11) and eigencentrality of a node (Figure 5.12) in the network. From the plots, it can be observed that the node features occasionally enter into an outlier zone (marked in red). Therefore, we need higher order statistics, in addition to mean and variance, to capture this kind of bursty phenomenon of node features in time. The two higher order summary statistics that we use are kurtosis and skewness. Kurtosis quantifies the peakedness or flatness of a probability distribution compared to a normal distribution. High kurtosis indicates that the distribution has heavy tails (presence of outliers) and a sharp peak, while low kurtosis indicates a flatter distribution with lighter tails (absence of outliers). Skewness measures the asymmetry of the distribution : positive skewness indicates that the right tail of the distribution is longer or fatter than the left tail (presence of samples with extremely high values), while negative skewness indicates the opposite (presence of samples with extremely low values).

$$Kurt[X] = E\left[\frac{(X - \mu)^4}{\sigma^4}\right]$$

$$Skew[X] = E\left[\frac{(X - \mu)^3}{\sigma^3}\right]$$

Putting everything together, each node is represented by a 5 dimensional feature vector that is evolving over time. We summarize the time evolution of each feature through 4 sufficient statistics thus giving a 20 dimensional feature vector per node (Figure 5.13).

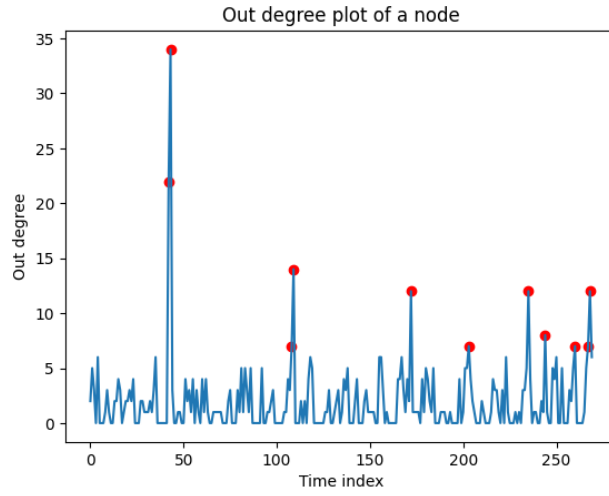


Figure 5.11: **Time evolution of out-degree of a node:** We have plotted the time evolution of out-degree of a node in the network. From the plot, it can be observed that out-degree occasionally enters into an outlier zone marked in red

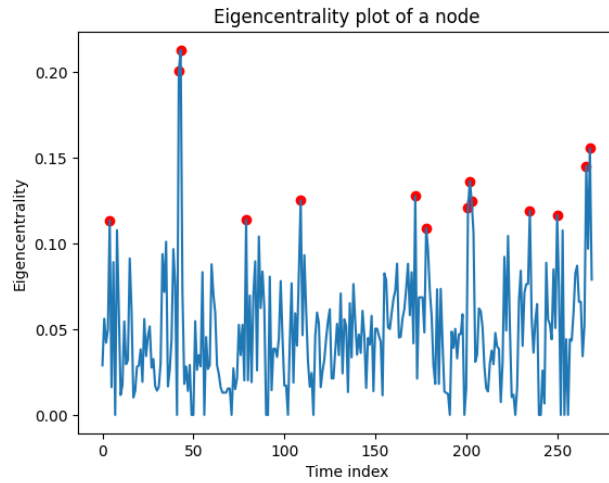


Figure 5.12: **Time evolution of eigencentrality of a node:** We have plotted the time evolution of eigencentrality of a node in the network. From the plot, it can be observed that the eigencentrality occasionally enters into an outlier zone marked in red

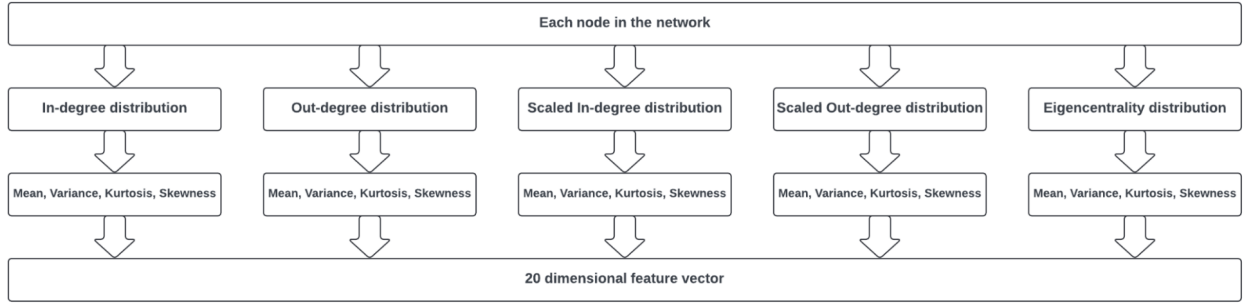


Figure 5.13: **20 dimensional feature vector per node for given frequency band:** For a given frequency band, each node is represented by a 5 dimensional feature vector that is evolving over time. We summarize the time evolution of each feature through 4 sufficient statistics thus giving a 20 dimensional feature vector per node per frequency band

As mentioned earlier, we study synchronization effects in three frequency bands: high gamma, ripple and fast ripple. Each frequency band results in a sequence of synchronization networks which in turn gives a 20 dimensional node feature. Concatenating the features across the 3 frequency bands we get a 60 dimensional feature vector per node (Figure 5.14).

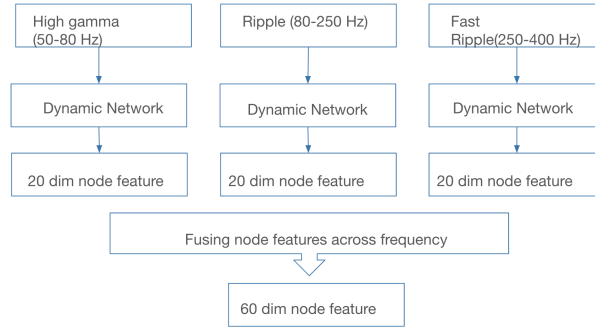


Figure 5.14: **Node feature vector construction workflow:** We study synchronization effects in three frequency bands: high gamma, ripple and fast ripple. Each frequency band results in a sequence of synchronization networks which in turn gives a 20 dimensional node feature. Concatenating the features across the 3 frequency bands we get a 60 dimensional feature vector per node.

5.4.2 Node labeling

We have described the node feature vector construction in the previous section. The only step left is to assign node labels before we can train and test our SOZ classification model. We assign a label of 1 to all the SOZ nodes. For the rest of the nodes we can assign a label of 0 but that will make the training set noisy for the following reasons. In seizure free patients there might be channels in the resected zone that might have the capacity to generate seizures. From Figure 5.9 we can see one such node (Resected Non-SOZ) being the hub of a hypersynchronous state and most likely it is PSOZ. Therefore, we drop resected Non-SOZ nodes from training. In non-seizure free patients, there might be channels in both the resected and preserved zone that might have the capacity to generate seizures since the surgical outcome was poor. Therefore, we drop non seizure free patients from training. Putting everything together, we only consider seizure free patients for training and for those patients we label SOZ nodes as 1 and Preserved Non-SOZ nodes as 0.

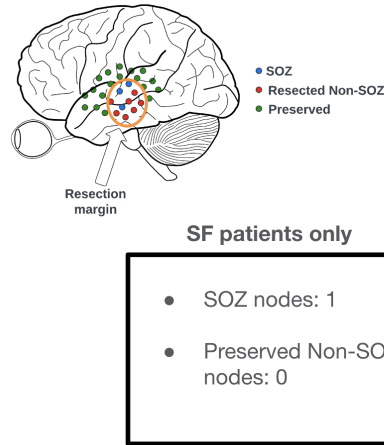


Figure 5.15: **Node labeling:** We only consider seizure free patients for training and for those patients we label SOZ nodes as 1 and Preserved Non-SOZ nodes as 0

5.4.3 Training and Testing corpus for SOZ identification

We constructed the training corpus by selecting seizure free patients with long interictal recording. Hence, 17 UCLA seizure free patients, each with an interictal recording of duration 90 minutes formed the training corpus and the remaining 142 patients formed the testing corpus. We trained one random forest SOZ classifier model on the channels of 17 patients and used it to infer the status of the channels of the remaining 142 patients (Figure 5.16). However, we also want to infer the status of the channels of 17 patients. A patient-wise elimination training strategy was adopted to validate the findings of the SOZ classifier for these 17 patients. In this training strategy, we trained 17 models in total: one random forest model each for the 17 seizure free patients (Figure 5.17). Therefore, we are using an independent cohort of channels to validate the findings of the detectors, as there is no overlap between the training and test set.

Random Forest is a machine learning algorithm that constructs a multitude of decision trees at training time and outputs the mode of the classes of the individual trees. Each tree is trained on a random subset of the training data and also uses a random subset of features for splitting at each node, thereby reducing overfitting and increasing robustness and accuracy.

Institution	Average length	Seizure Free	Not Seizure Free	Age
UCLA	90 minutes	17	7	3-20 years
Wayne State	5 minutes	96	39	4-30 years

- 17 patients in the training corpus
- Leads to one random forest model
- 142 patients in the testing corpus

Figure 5.16: **Training and Testing Corpus:** 17 UCLA seizure free patients in the training corpus and the remaining 142 patients in the testing corpus resulting in one random forest model

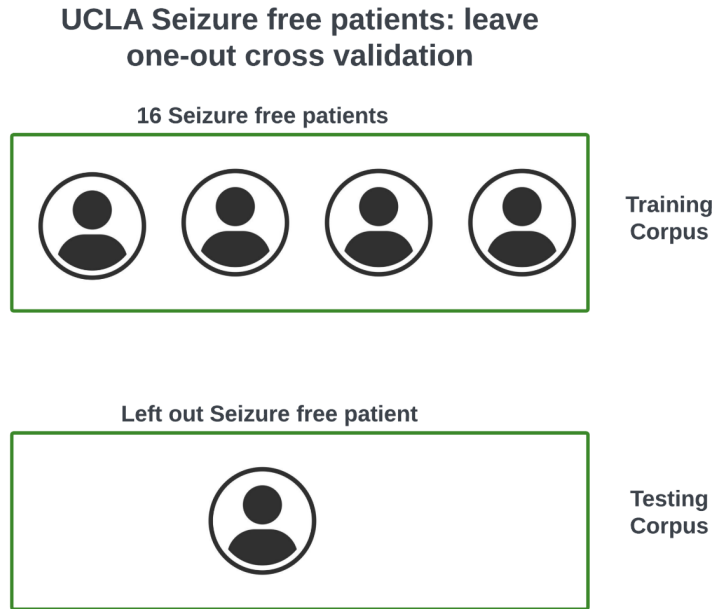


Figure 5.17: **Training and Testing Corpus for 17 UCLA seizure free patients:** For each fold, we have 16 UCLA seizure free patients in the training corpus and the left out UCLA seizure free patient in the testing corpus resulting in one random forest model per fold

5.4.4 Results

Recall, from our SOZ classifier setup shown in Figure 5.6, that it outputs a probability of the node being part of SOZ. We can threshold the probability scores to get the predicted node labels and compare them with the ground truth labels. In Figure 5.18, we tabulate the performance metrics when the trained SOZ classifier was used to predict the status of all the channels except the resected Non-SOZ ones. Since the ground truth label distribution is highly biased with only around 10% of the channels having a ground truth label of SOZ so recall is a more important metric than accuracy. From the figure, we can observe that the classifier is able to correctly identify most of the SOZ channels in the test set shown by

the high recall of 98%. The few false negatives (i.e. ground truth SOZ channels predicted as Non-SOZ), comes from patients where the SOZ labeling by doctors was generous. For example, most of the false negatives came from patients DL and GM where each of them had more than 30% of the channels labeled as SOZ.

When the same trained classifier was used to predict the status of the resected Non-SOZ channels, then around 18% of them got predicted as SOZ with a high probability. This result is along expected lines because from Figure 5.9 we know that these types of channels also emerge as hubs of hypersynchronous states like the SOZ channels.

Testing set	F1	Accuracy	Recall	Precision
No Resected Non-SOZ channels	98.7%	99.6%	98.1%	99.3%

Figure 5.18: **SOZ classification results:** The SOZ classifier outputs a probability score and we use a threshold of 0.5 to make predictions: $P(i \in SOZ) \geq 0.5 \rightarrow i \in SOZ$. The classifier is able to correctly identify most of the SOZ channels in the test set shown by the high recall of 98%. Few false negatives come from patients where the SOZ labeling by doctors were very generous.

5.4.5 Discussion

We have modeled the interictal interactions amongst the iEEG channels using a sequence of synchronization networks and leveraged the temporal dynamics of the network properties to accurately identify the channels that drove all the seizure episodes (SOZ) during the 2 week long hospital monitoring with over 98% accuracy. One of the hallmarks of the constructed networks is that they occasionally transition into a state of hypersynchrony with SOZ nodes being the hub of these hypersynchronous states and our model is able to identify them and

exploit them for accurate SOZ identification. Therefore, we have built a model that can identify the drivers of almost all the seizure episodes that happened during the 2 week long hospital monitoring from interictal segments with duration as short as 5 minutes. This result can significantly shorten the invasive hospital monitoring time, which is fraught with risk of infection.

The same model also identifies another set of channels (around 18%) of having the capacity to generate seizures and we hypothesize that these channels might be PSOZ. The only way to validate our hypothesis is through the evaluation of a surgical outcome prediction model, which we build in the following section.

5.5 Predicting post-surgical outcomes using interictal SOZ classification and network dynamics

In the last section, we discovered a set of channels that got predicted as SOZ with a high probability but did not drive seizure episodes during the 2 week long hospital monitoring. These channels might constitute the unobserved/inactive part of EZ known as Potential Seizure Onset Zone (PSOZ). Since there is no diagnostic modality to know whether the entire EZ ($\text{SOZ} \cup \text{PSOZ}$) has been identified and resected, so the only way to know is through surgical outcome evaluation: if the patient becomes seizure free post surgery then the EZ has been resected otherwise not. Therefore, we want to build a model to predict the surgical outcome of patients. The first and the most important step in this model building is to design a good set of features that truly encapsulates all the necessary information about EZ. In the following sections, we will be reasoning our way to develop such a set of features.

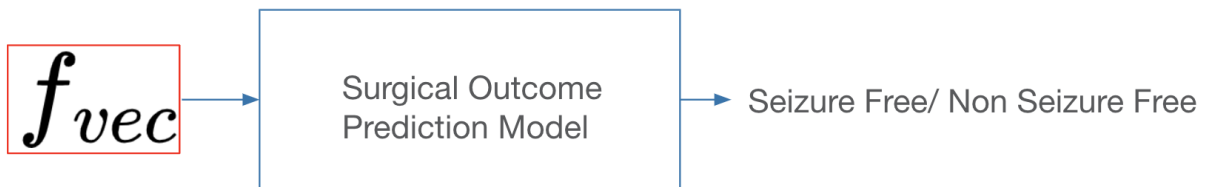


Figure 5.19: **Surgical outcome prediction model:** The surgical outcome prediction model takes feature vector of a patient as input and outputs a probability of the patient becoming seizure free which can be then thresholded to predict whether the patient will become seizure free or not

5.5.1 3 dimensional feature vector

Complete resection of SOZ has been correlated with the patient gaining seizure freedom after surgery. Therefore, the first feature we consider is the extent of SOZ resection. It is captured by the categorical variable, $catSOZ$, defined as follows

$$catSOZ = \begin{cases} 1, \text{ all channels predicted as SOZ by interictal SOZ classifier is resected} \\ 0, \text{ otherwise.} \end{cases}$$

Since our interictal SOZ classifier returns a probability score for each channel being in SOZ, so we can also use the scores to construct a set of features rather than the hard predicted labels as in the $catSOZ$ construction. We construct the second feature by computing what fraction of the total probability score is captured by the resected SOZ nodes. It is captured by the variable, resSOZ ratio, defined as follows

$$\text{resSOZ ratio} = \frac{\sum_{(i \in \text{res}, P(i \in SOZ) \geq 0.5)} P(i \in SOZ)}{\sum_{i \in \text{all}} P(i \in SOZ)} \quad (5.12)$$

In equation 5.12, the numerator captures the total probability scores of all the channels that were resected and predicted as SOZ by our interictal SOZ classifier and the denominator captures the total probability scores of all the channels. Similarly, we construct the third feature by computing what fraction of the total probability score is captured by the resected nodes. It is captured by the variable, resection ratio, defined as follows

$$\text{resection ratio} = \frac{\sum_{(i \in \text{res}, P(i \in SOZ) \geq 0.5)} P(i \in SOZ) + \sum_{(i \in \text{res}, P(i \in SOZ) < 0.5)} P(i \in SOZ)}{\sum_{i \in \text{all}} P(i \in SOZ)} \quad (5.13)$$

In equation 5.12, the numerator captures the total probability scores of all the channels that were resected and the denominator captures the total probability scores of all the channels. Concatenating the 3 constructed features, we get our 3 dimensional feature vector

$$f_{vec} = \begin{bmatrix} catSOZ \\ \text{resSOZ ratio} \\ \text{resection ratio} \end{bmatrix}$$

5.5.1.1 Training and Testing corpus for surgical outcome prediction

We train and validate the performance of the 3 dimensional surgical outcome prediction model on the same multicenter patient corpus of 159 patients that we introduced earlier (Figure 5.20). Since the patient corpus is unbalanced with $\sim 70\%$ patients being seizure free and $\sim 30\%$ patients being non-seizure free, so we use 10 fold stratified cross validation to evaluate the performance and generalization power of our model. In 10 fold stratified cross validation, we split the corpus of 159 patients into 10 patients groups of roughly equal size where each group has around 70% seizure free and around 30% non-seizure free patients. Then for each fold, we train the model using 9 patient groups and then test the model on the left out patient group.

Institution	Average length	Seizure Free	Not Seizure Free	Age
UCLA	90 minutes	17	7	3-20 years
Wayne State	5 minutes	96	39	4-30 years

Figure 5.20: **Multicenter patient corpus:** A multicenter (UCLA, Wayne State) patient corpus of 159 patients with 113 of them being seizure free ($\sim 70\%$) and 46 of them being non-seizure free ($\sim 30\%$)

5.5.1.2 Surgical outcome prediction results using 3 dimensional feature vector

Recall, from our surgical outcome prediction model setup shown in Figure 5.19, that it outputs a probability of the patient being seizure free. Based on the stratified cross validation approach described in the previous section, each patient gets to be in the test set once and hence gets predicted a probability of seizure freedom by the trained model. In Figure 5.21, we have plotted the predicted probabilities categorized by the ground truth surgical outcomes. From the plot, it can be observed that the mean predicted probability of seizure freedom is slightly higher for the seizure free group compared to the non-seizure free group.

Predicted probability of success categorized by surgical outcomes (3 dimensional)

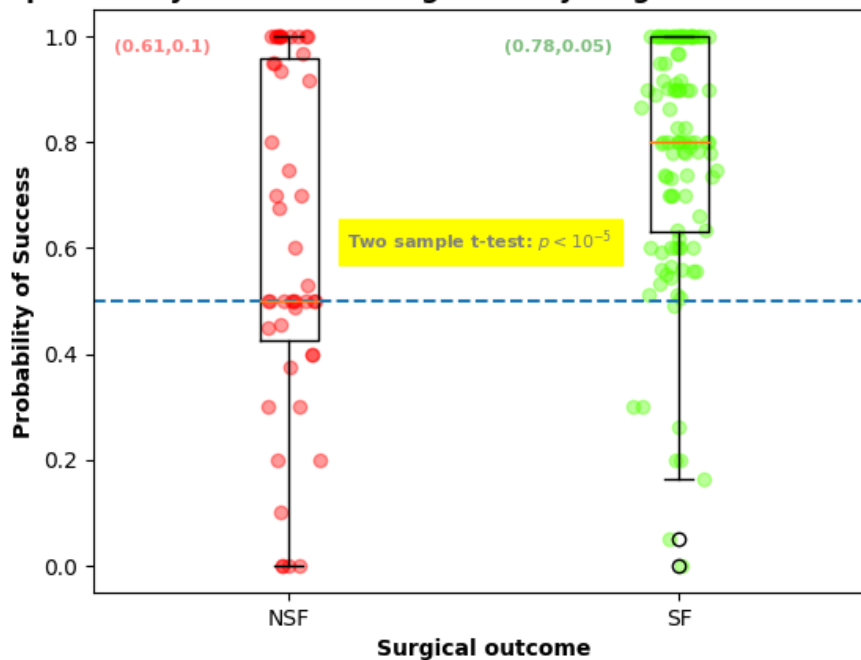


Figure 5.21: **Probability of seizure freedom:** A box plot illustrating the probability of seizure freedom, as predicted by the 3-dimensional model, stratified by the actual surgical outcomes. Mean seizure freedom probability of seizure free group is slightly higher than the non-seizure free group

We can threshold the probability scores to get the predicted surgical outcome labels and

compare them with the ground truth surgical outcome labels. In Figure 5.22, we have tabulated the 10 fold performance metric statistics. Since the labels are unbalanced, so F1 score is more relevant than accuracy. The 3 dimensional model is able to predict the surgical outcomes of patients with a mean F1 score of 86% across the 10 folds and also generalizes well as shown by the small variance in performance across folds. Going into the misclassification statistics, shown in Figure 5.23, we observe that the model makes more mistakes in the minority class, that is the non-seizure free patients, as shown by the 32 false positives.

Performance Metric	(Mean, Variance)
F1	(0.86,0.003)
Accuracy	(0.79,0.007)
Precision	(0.82,0.008)
Recall	(0.92,0.002)

Figure 5.22: **10 fold performance metric statistics (3 dimensional model):** The table presents the aggregated statistics, including the mean and variance, calculated across 10 folds, for four performance metrics: F1 score, Accuracy, Precision, and Recall. The 3 dimensional model demonstrates proficiency in predicting surgical outcomes, achieving a mean F1 score of 86% and a mean accuracy of 79%. Notably, the model exhibits strong generalization capabilities, as evidenced by minimal variance in performance across folds

False Negatives (SF predicted as NSF)	False Positives (NSF predicted as SF)
9	32

Figure 5.23: **Misclassification statistics (3 dimensional model):** The 3 dimensional model makes more mistakes in the minority class as shown by the 32 false positives

5.5.1.3 Surgical outcome prediction results using doctor's SOZ labels

The gold standard for doctors is to identify the SOZ channels by analyzing all the ictal segments captured in two weeks of hospital monitoring and then surgically resect them with the hope that the patient will become seizure free. We can evaluate the efficacy of the doctors' process in our context of predicting the post surgical outcomes by defining a categorical variable, *docSOZ*, which is defined as follows

$$docSOZ = \begin{cases} 1, & \text{all channels labeled as SOZ by doctors are resected} \\ 0, & \text{otherwise.} \end{cases}$$

Using the same 10 fold stratified cross validation approach described in Section 5.5.1.1, we train and test the 1 dimensional model currently used by the doctors. In Figure 5.24, we have tabulated the 10 fold performance statistics for both the 1 dimensional and the 3 dimensional model. We observe that the 3 dimensional model highlighted in yellow performs better than the one currently used by the doctors which is highlighted in green, with an accuracy boost of more than 12%, an F1 score boost of more than 7%, and 12 fewer misclassifications.

Feature Vector Dimension	Mean F1	Mean Accuracy	Mean Precision	Mean Recall	False Negatives	False Positives
1(<u>docSOZ</u>)	78.6%	66.7%	71.2%	88.4%	13	40
3	86%	79%	82%	92%	9	32

Figure 5.24: **10 fold performance metric and misclassification statistics comparison (1dim vs 3dim):** The 3 dimensional model (highlighted in yellow) outperforms the current 1 dimensional model utilized by doctors (highlighted in green), boasting a significant increase in accuracy by 12%, an improvement in F1 score by 7%, and a reduction of 12 misclassifications

5.5.1.4 Discussion

We have constructed a 3 dimensional feature vector from the interictal SOZ classifier model scores to predict the surgical outcomes of patients. It performs better than the one currently used by doctors and is therefore a better biomarker of EZ than the docSOZ based biomarker used by the doctors currently. We believe that the performance boost is mainly due to two reasons: 1) our interictal SOZ classifier extracts a portion of the unobserved part of EZ which we call PSOZ and 2) the use of the probability scores to construct the features rather than just using the hard predicted labels of the SOZ classifier.

However, our current 3 dimensional model still makes incorrect predictions on 41 patients. To get a better sense of why our current model make incorrect predictions on these patients, we have plotted the resSOZ ratio as a function of resection ratio for all the 159 patients in Figure 5.25. From the plot, it can be observed that there is a subset of seizure free patients (marked with blue circles) which has a low resection ratio and $\text{catSOZ} = 0$ meaning that some of the predicted SOZ were not resected by the doctors. Hence, our current 3 dimensional feature vector (catSOZ , resSOZ ratio, resection ratio) is not able to explain the outcome of these patients because of the low feature values and we suspect that there might be some interesting patterns embedded in the network flows. In order to extract such patterns and augment them into the feature vector, we develop a framework for investigating the network flow dynamics in the next section.

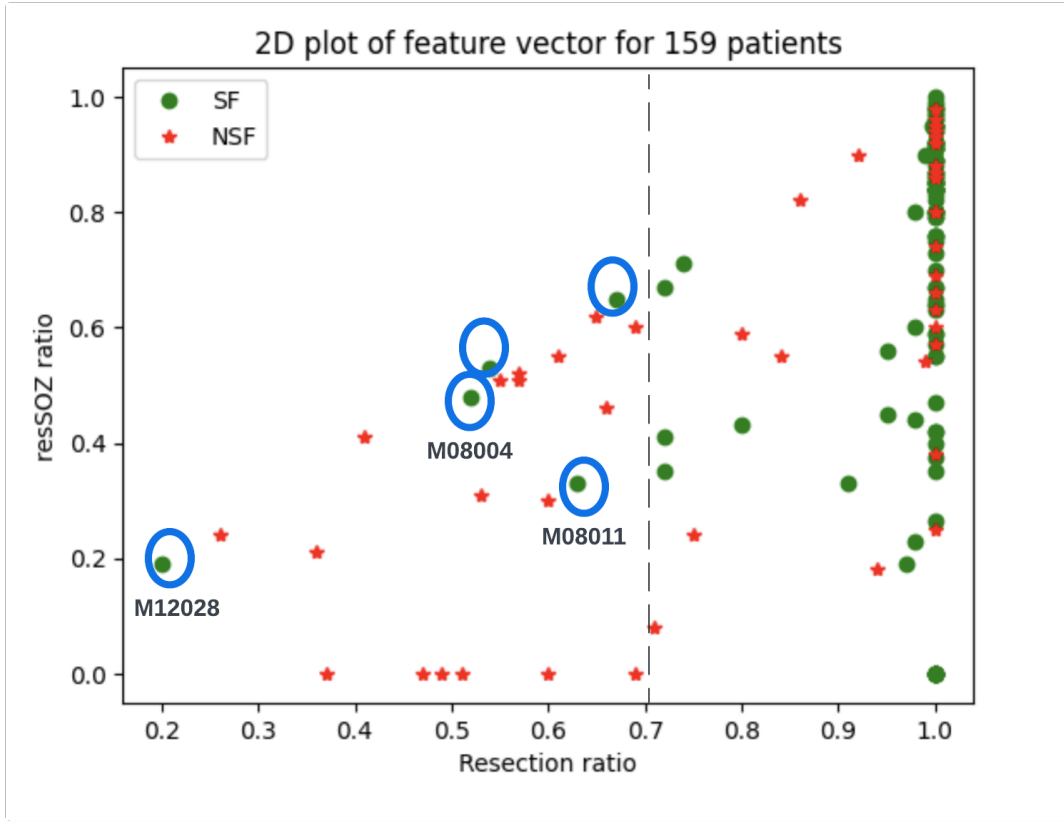


Figure 5.25: **2D visualization of the feature vector for 159 patients:** resSOZ ratio is plotted as a function of resection ratio for all 159 patients. The patients are colored based on ground truth surgical outcome: green circles represent seizure free patients and red crosses represent non-seizure free patients. There is a subset of seizure free patients, highlighted with blue circles, that has a low resection ratio with some predicted SOZ channels in the preserved region. These patients are incorrectly predicted as non-seizure free by the 3 dimensional model

5.5.2 5 dimensional feature vector

In Figure 5.25, we found presence of a set of patients who became seizure free even though the entire predicted SOZ was not resected. This evidence led us to investigate the flow dynamics in the constructed interictal synchronization networks. In this framework, we partition the interictal network into 3 non-overlapping regions based on the interictal SOZ classifier predictions and the resection information: predicted SOZ channels that are resected (resected SOZ), predicted SOZ channels that are not resected (preserved SOZ) and predicted Non-SOZ channels that are not resected (preserved Non-SOZ). Based on this partitioning, we compute the flow between these regions by summing the weights of all the edges connecting the nodes in these regions (Figure 5.26):

f_1 := Flow between preserved SOZ and preserved Non-SOZ

f_2 := Flow between resected SOZ and preserved Non-SOZ

f_3 := Flow between preserved SOZ and resected SOZ,

where

$$\begin{aligned} f_1 &= \sum_{i \in \text{preserved SOZ}, j \in \text{preserved Non-SOZ}} w_{ij} \\ f_2 &= \sum_{k \in \text{resected SOZ}, j \in \text{preserved Non-SOZ}} w_{kj} \\ f_3 &= \sum_{i \in \text{preserved SOZ}, k \in \text{resected SOZ}} w_{ik}. \end{aligned}$$

Note that such a partitioning might not always exist, when there is no preserved SOZ, in such cases we set $f_1 = f_2 = f_3 = 0$.

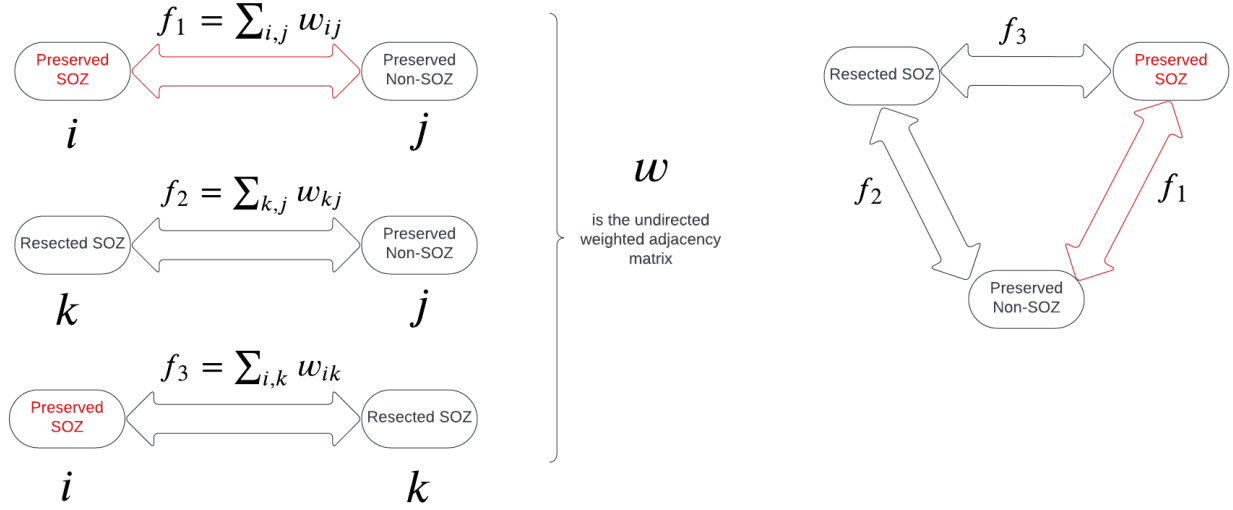


Figure 5.26: **Flow analysis pipeline:** We partition the set of nodes in the interictal synchronization network into 3 non-overlapping set based on the interictal SOZ classifier predictions and resection information. The flow between these 3 regions are computed by summing the weights of the undirected edges connecting the nodes in these regions

In Figure 5.27, we have plotted f_1, f_2, f_3 as a function of time for a seizure free patient. From the plot, it can be observed that given the resected soz cut, on average there is minimal flow between the preserved SOZ and the preserved Non-SOZ, showing that preserved SOZ cannot trigger large scale synchronization events. Similarly, in Figure 5.28, we have plotted f_1, f_2, f_3 as a function of time for a non-seizure free patient. In this case, given the resected SOZ cut, there is moderately strong flow between preserved SOZ and preserved Non-SOZ, showing that preserved SOZ channels can trigger large scale synchronization events. Although, we have just plotted the flow evolution for only one seizure free and one non-seizure free patient but such patterns hold for more seizure and non-seizure free patients where such a partitioning exists. Based on these observations, we construct the additional features Flow 1 and Flow 2 which are defined as follows

$$\text{Flow 1} = \frac{E[f_1]}{E[f_2]}, \text{Flow 2} = \frac{E[f_3]}{E[f_2]}.$$

In terms of Flow 1 and Flow 2,

Seizure free patients := Flow 1 << Flow 2

Non-Seizure free patients := Flow 1 $\approx$ Flow 2

Augmenting the previously constructed 3 dimensional feature vector with the flow features, we get our 5 dimensional feature vector

$$f_{vec} = \left[catSOZ, resSOZ \text{ ratio}, resection \text{ ratio}, Flow 1, Flow 2 \right]^T$$

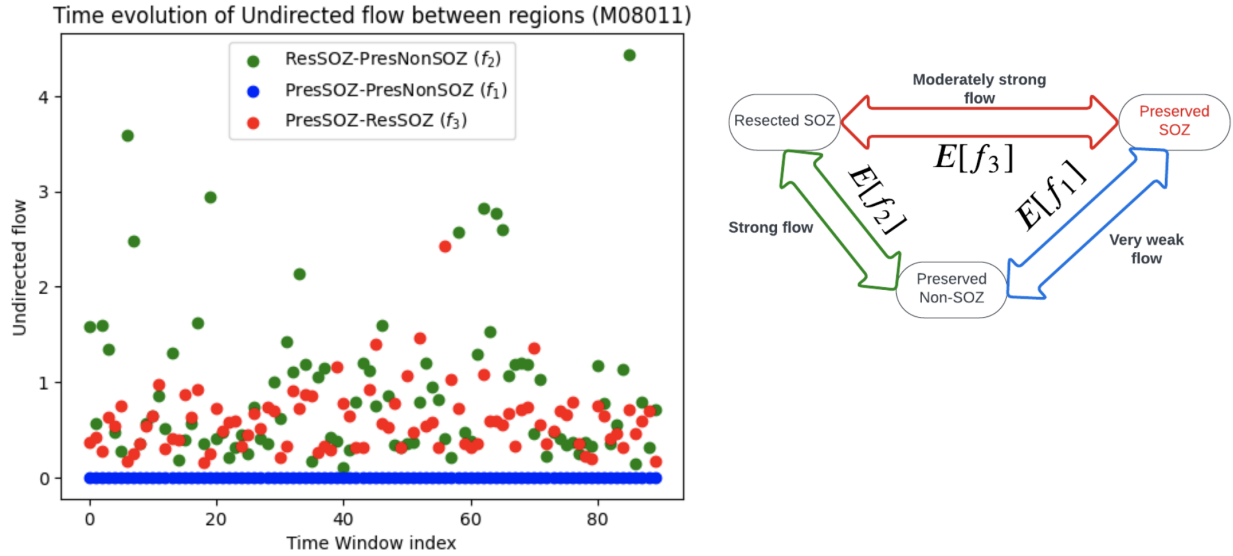


Figure 5.27: **Undirected flow dynamics in a seizure free patient:** The flows between the three regions (f_1, f_2, f_3) is plotted as a function of time where each time index corresponds to one interictal synchronization network. On average, there is minimal flow between preserved SOZ and preserved Non-SOZ given the resected SOZ cut

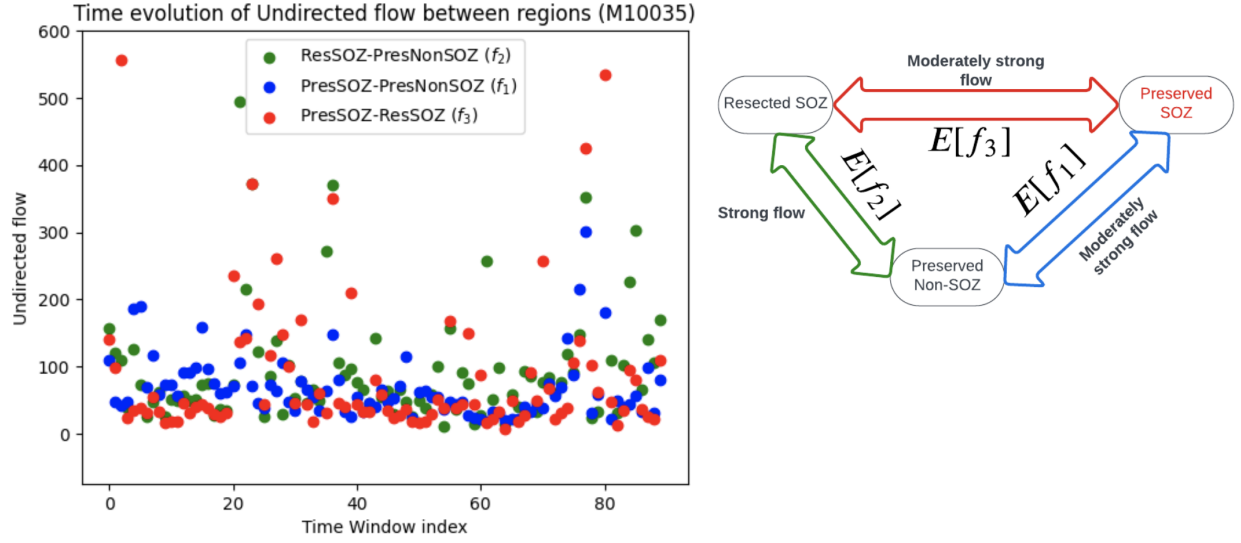


Figure 5.28: **Undirected flow dynamics in a non-seizure free patient:** The flows between the three regions (f_1, f_2, f_3) is plotted as a function of time where each time index corresponds to one interictal synchronization network. On average, there is moderately strong flow between preserved SOZ and preserved Non-SOZ given the resected SOZ cut

We train and validate the performance of the 5 dimensional surgical outcome prediction model using the same 10 fold stratified cross validation approach described in Section 5.5.1.1.

5.5.2.1 Surgical outcome prediction results using 5 dimensional feature vector

Recall, from our surgical outcome prediction model setup shown in Figure 5.19, that it outputs a probability of the patient being seizure free. Based on the stratified cross validation approach described in Section 5.5.1.1, each patient gets to be in the test set once and hence gets predicted a probability of seizure freedom by the trained model. In Figure 5.29, we have plotted the predicted probabilities categorized by the ground truth surgical outcomes. From the plot, it can be observed that the mean predicted probability of seizure freedom is higher for the seizure free group compared to the non-seizure free group showing the efficacy of the 5 dimensional feature vector in predicting surgical outcomes.

Predicted probability of success categorized by surgical outcomes (5 dimensional)

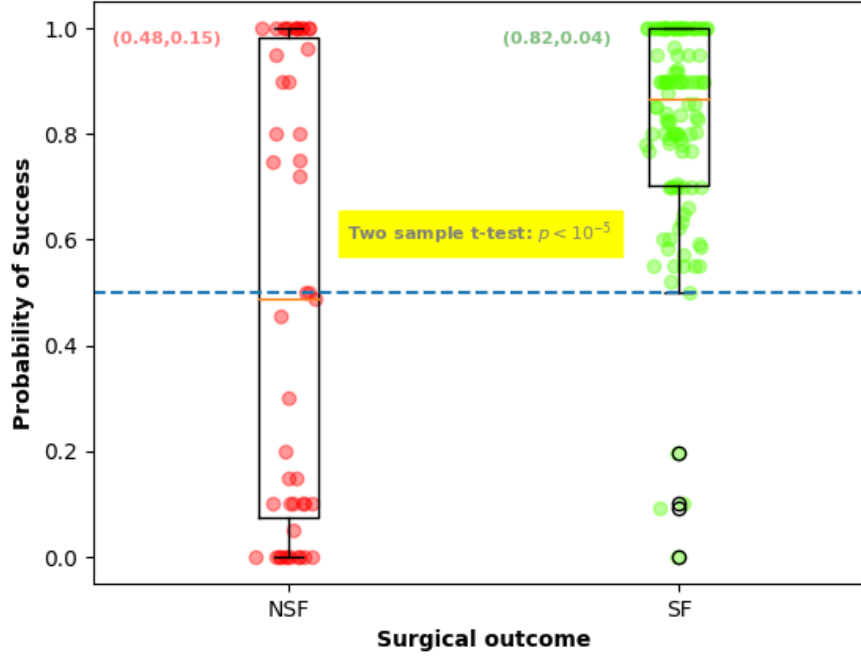


Figure 5.29: **Probability of seizure freedom (5 dimensional)**: A box plot illustrating the probability of seizure freedom, as predicted by the 5 dimensional model, stratified by the actual surgical outcomes. Mean seizure freedom probability of seizure free group (0.82) is higher than the non-seizure free group (0.48)

We can threshold the probability scores to get the predicted surgical outcome labels and compare them with the ground truth surgical outcome labels. In Figure 5.30, we have tabulated the 10 fold performance metric and misclassification statistics for both the 3 dimensional and 5 dimensional model. The 5 dimensional model which incorporates the network flow dynamics performs better than the 3 dimensional model with a boost in accuracy of 5%, boost in F1 score of 3% and 14 fewer misclassifications. By incorporating the flow dynamics, the model not only corrects the 5 circled Seizure free patients in Fig 5.25, but also corrects 9 Non-Seizure free patients

Feature Vector Dimension	Mean F1	Mean Accuracy	Mean Precision	Mean Recall	False Negatives	False Positives
3	86%	79%	82%	92%	9	32
5	89%	84%	84%	95%	4	23

Figure 5.30: **10 fold performance metric and misclassification statistics comparison (3dim vs 5dim)**: 5 dimensional model (highlighted in pink) performs better than the 3 dimensional model (highlighted in maroon) with a boost in accuracy of 5%, boost in F1 score of 3%, and 14 fewer misclassifications

5.5.2.2 Discussion

We have constructed a 5 dimensional feature vector from the interictal SOZ classifier model scores and the interictal network flow dynamics to predict the surgical outcomes of patients. It performs significantly better than the one currently used by doctors (Fig 5.31, accuracy boost of 17%, F1 score boost of 10%, 26 fewer misclassifications). Therefore, we have developed an interictal network based method that can predict the post surgical outcome of patients better and hence is a better biomarker of epileptogenic zone than the SOZ based biomarker used by doctors currently. Also, with the interictal network flow dynamics framework we have provided evidence that complete resection of the SOZ is not a necessary condition for seizure freedom and sometimes by just resecting the right set of SOZ nodes one can disrupt the flow in epileptic network and hence achieve seizure freedom. More importantly this framework will allow us to come up with algorithms that can guide the doctors to design minimum resection margins for seizure freedom.

Feature Vector Dimension	Mean F1	Mean Accuracy	Mean Precision	Mean Recall	False Negatives	False Positives
1 (docSOZ)	78.6%	66.7%	71.2%	88.4%	13	40
3	86%	79%	82%	92%	9	32
5	89%	84%	84%	95%	4	23

Figure 5.31: **10 fold performance metric and misclassification statistics comparison (1dim vs 5dim)**: The 5 dimensional model (highlighted in pink) outperforms the current 1 dimensional model employed by doctors (highlighted in green), demonstrating a remarkable increase in accuracy by 17%, an improvement in F1 score by 10%, and a reduction of 26 misclassifications

However, our current 5 dimensional model still makes incorrect predictions on 27 patients and the bulk of these mistakes are false positives. To get a better sense of why our current model make incorrect predictions on these patients, we have plotted the resSOZ ratio as a function of resection ratio for all the 159 patients in Figure 5.32. Almost all of the 23 false positives are located in the far right corner boxed in blue. These patients have a very high resection ratio (> 0.9) and have all the predicted SOZ removed (flow based partitioning doesn't exist) so the current set of features is not able to explain their outcomes. Therefore, we need to look for some other sources of features that are not necessarily confined to the space of interictal networks. In the next section, we develop such a set of features that combines the interictal network features along with the existing biomarkers of EZ.

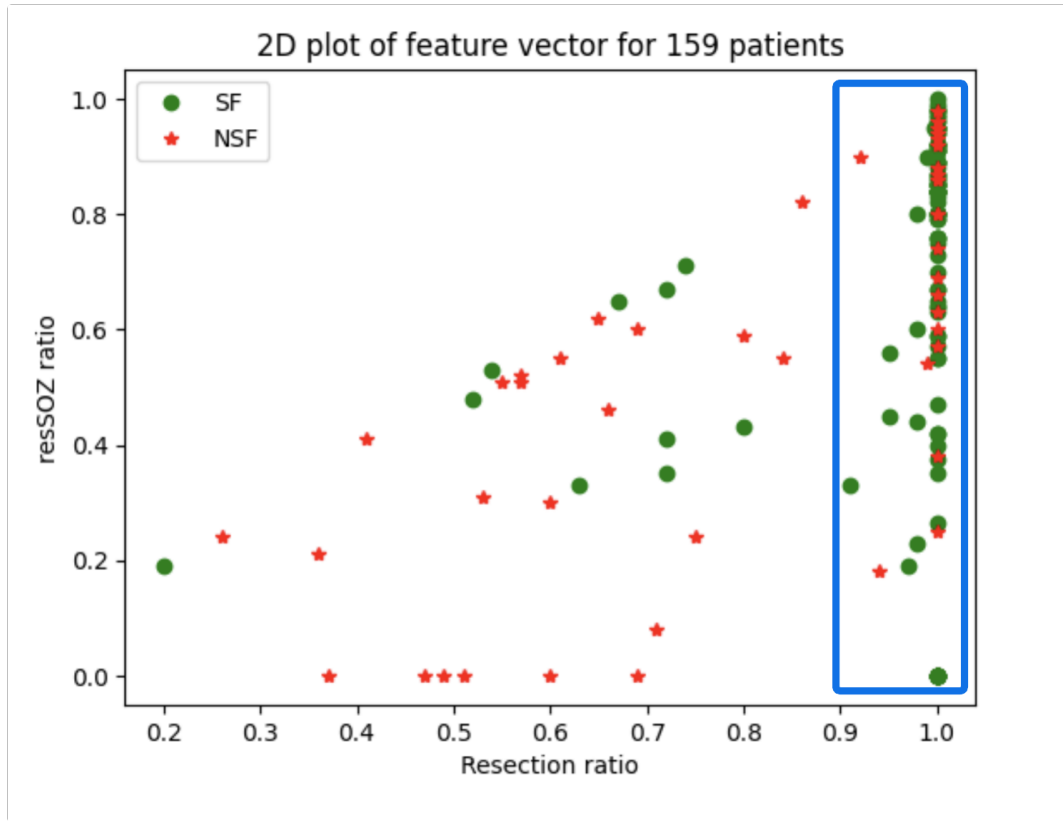


Figure 5.32: **2D visualization of the feature vector for 159 patients:** resSOZ ratio is plotted as a function of resection ratio for all 159 patients. The patients are colored based on ground truth surgical outcome: green circles represent seizure free patients and red crosses represent non-seizure free patients. Majority of the false positives, boxed in blue, are situated in the far right corner of the plot with a very high resection ratio and all the predicted SOZ channels removed

5.5.3 Mini Seizures: 7 dimensional feature vector

One of the key observations that led to the accurate identification of EZ from interictal segments was the emergence of hypersynchronous states in the interictal networks. From Figure 5.33, we can observe excessive synchronization in the time domain as well during such episodes of hypersynchronization in the interictal period. The observed excessive time domain synchronization in the interictal period looks similar to the ones observed during seizure episodes except it's less pronounced in terms of amplitude. We hypothesize that these hypersynchronous states are like 'Mini Seizures' happening in the interictal period with the SOZ/PSOZ nodes driving them and our model is able to mine them for accurate EZ identification.

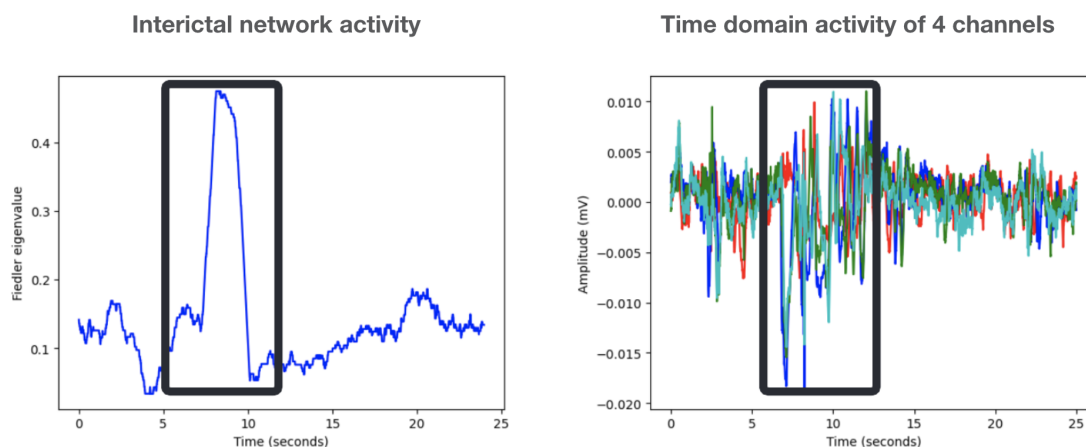


Figure 5.33: **'Mini Seizures' in interictal period:** The Fiedler eigenvalue is plotted over time during the interictal period on the left side of the plot, while the raw waveform amplitude of four channels over the same time period is depicted on the right. Notably, time segments characterized by hypersynchronous activity (sharp rise in fiedler value) are boxed in black. Observing the amplitude plot, it becomes evident that the channels exhibit excessive synchronization in the time domain during these hypersynchronous states, bearing resemblance to that observed during seizures. As a hypothesis, we speculate that these periods of hypersynchronous activity in the interictal period may correspond to "Mini Seizures"

In Chapter 3, we crafted a weakly supervised deep learning framework aimed at enhancing the specificity of the channel level interictal biomarker known as HFO by isolating the pathological HFO cluster. Our findings indicated that the pathological HFO cluster serves as a superior biomarker of the epileptogenic zone (EZ) compared to the general HFO cluster.

In Section 5.5.2, we introduced an interictal network property based biomarker of the epileptogenic zone (EZ), which demonstrated superior performance compared to the currently utilized EZ biomarker based on the seizure onset zone (SOZ). To gain insight into the potential relationship between these network level and channel level biomarkers, we generated a plot depicting the time evolution of pathological HFO (spk-HFO) rate of a channel (Figure 5.34, left) and time evolution of its eigenvector centrality (Figure 5.34, right). The plot suggests a correlation between the eigenvector centrality, representing an interictal network property, and the spk-HFO rate, a channel level interictal biomarker. This observation suggests that channel level pathological events like spk-HFO may be influenced by interictal network effects, implying that we could potentially predict certain statistics of these events using interictal network properties. In the next section, we develop a regression based framework to explore the correlation between spk-HFO and interictal network properties.

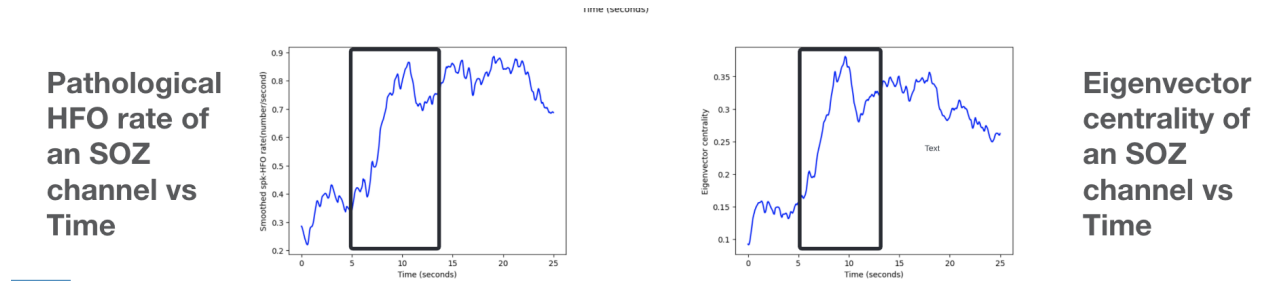


Figure 5.34: **Correlation between interictal network properties and channel level pathological events:** The spk-HFO rate of a channel is graphed over time during the interictal period on the left side of the plot, while the eigenvector centrality of the same channel over the same time period is illustrated on the right. Upon examination of both plots, it becomes apparent that the spk-HFO rate of a channel appears to be correlated with eigenvector centrality of that channel

5.5.3.1 Regression model

We build a regression model based framework to capture the correlation between the interictal network properties and spk-HFO rate. Specifically, we predict the smoothed spk-HFO rate of a channel c at time step t using three properties of channel c in the interictal network at time step t : in degree of channel c , out degree of channel c and eigenvector centrality of channel c .

$$sp_c(t) = a_1 \times i_c(t) + a_2 \times o_c(t) + a_3 \times ei_c(t).$$

At a channel level, we quantify the correlation using R^2 which measures how much of the variations in spk-HFO rate of a channel c can be explained by the interictal network properties.

$$R_c^2 = 1 - \frac{\sum_{t=0}^T (sp_c(t) - \hat{sp}_c(t))^2}{\sum_{t=0}^T (sp_c(t) - E[sp_c])^2},$$

where

- $sp_c(t)$: true smoothed spk-HFO rate of channel c at time step t
- $\hat{sp}_c(t)$: predicted smoothed spk-HFO rate of channel c at time step t
- $E[sp_c]$: Average smoothed spk-HFO rate of channel c
- R_c^2 : R^2 score of channel c

A R^2 score close to 1 implies that the independent variables (interictal network properties) explain most of the variations in the dependent variable (smoothed spk-HFO rate). An iEEG recording from a patient will have many channels, so to get a patient level feature we summarize the R^2 values across channels using mean and the ratio of variance to mean

$$\begin{aligned} \text{feature 6} &= E[R_c^2] \\ \text{feature 7} &= \frac{Var[R_c^2]}{E[R_c^2]}. \end{aligned}$$

Augmenting the previously constructed 5 dimensional feature vector with the R^2 statistics based features, we get our 7 dimensional feature vector

$$f_{vec} = \begin{bmatrix} catSOZ \\ resSOZ \text{ ratio} \\ resection \text{ ratio} \\ Flow \ 1 \\ Flow \ 2 \\ E[R_c^2] \\ \frac{Var[R_c^2]}{E[R_c^2]} \end{bmatrix}$$

We train and validate the performance of the 7 dimensional surgical outcome prediction model using the same 10 fold stratified cross validation approach described in Section 5.5.1.1.

5.5.3.2 Surgical outcome prediction results using 7 dimensional feature vector

Recall, from our surgical outcome prediction model setup shown in Figure 5.19, that it outputs a probability of the patient being seizure free. Based on the stratified cross validation approach described in Section 5.5.1.1, each patient gets to be in the test set once and hence gets predicted a probability of seizure freedom by the trained model. In Figure 5.35, we have plotted the predicted probabilities categorized by the ground truth surgical outcomes. From the plot, it can be observed that the mean predicted probability of seizure freedom is significantly higher for the seizure free group compared to the non-seizure free group showing the efficacy of the 7 dimensional feature vector in predicting surgical outcomes.

Predicted probability of success categorized by surgical outcomes (7 dimensional)

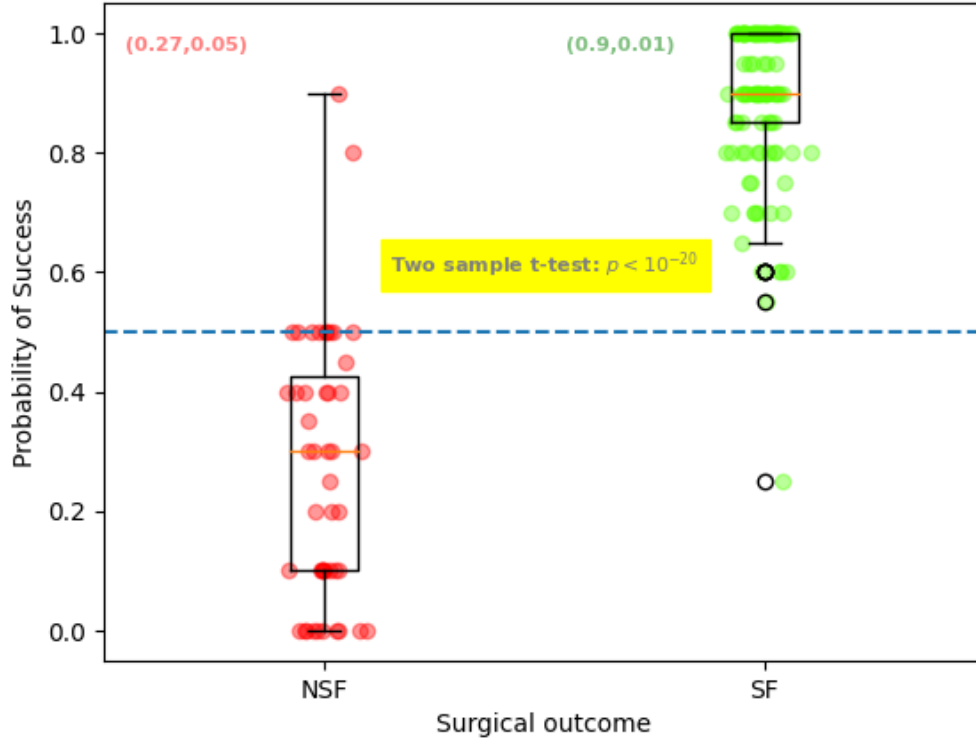


Figure 5.35: **Probability of seizure freedom (7 dimensional)**: A box plot illustrating the probability of seizure freedom, as predicted by the 7 dimensional model, stratified by the actual surgical outcomes. Mean seizure freedom probability of seizure free group (0.9) is significantly higher than the non-seizure free group (0.27)

We can threshold the probability scores to get the predicted surgical outcome labels and compare them with the ground truth surgical outcome labels. In Figure 5.36, we have tabulated the 10 fold performance metric and misclassification statistics for all the models (1 dimensional, 3 dimensional, 5 dimensional, 7 dimensional). The 7 dimensional model which combines both the interictal network biomarker and channel level pathological event based biomarker performs significantly better than the SOZ based biomarker used by doctors currently with a boost in accuracy of 30%, boost in F1 score of 20% and 50 fewer misclassifications. Also by combining the biomarkers, we have reduced the number of false positives

from 23 to only 2.

Feature Vector Dimension	Mean F1	Mean Accuracy	Mean Precision	Mean Recall	False Negatives	False Positives
1(docSOZ)	78.6%	66.7%	71.2%	88.4%	13	40
3	86%	79%	82%	92%	9	32
5	89%	84%	84%	95%	4	23
7	98.7%	98%	98%	99%	1	2

Figure 5.36: **10 fold performance metric and misclassification statistics comparison (1dim vs 3dim vs 5dim vs 7dim):** The 7 dimensional model (highlighted in blue) significantly outperforms the current 1 dimensional model utilized by doctors (highlighted in green), showcasing a substantial increase in accuracy by 30%, a remarkable improvement in F1 score by 20%, and a reduction of 50 misclassifications

5.5.3.3 Discussion

We have developed a set of features that uses the interictal network properties and exploits the interplay between the network properties and spk-HFO to predict the post surgical outcome of patients better. The significant performance boost in retrospective validation makes the set of features a novel biomarker of epileptogenic zone. Also, the remarkable reduction in invasive monitoring time required to compute the set of features makes this biomarker more appealing as it can considerably cut down the risk and costs associated with hospital stays.

5.6 A Biomarker of EZ agnostic to surgical resection and SOZ labels

In the previous section, we observed that adding the features capturing the correlation between interictal network properties and spk-HFO rate significantly boosted the performance of the surgical outcome prediction model. Therefore, there might be some interesting pattern embedded in the correlation feature space that's driving such good performance. In Figure 5.37, we have plotted $\frac{Var[R_c^2]}{E[R_c^2]}$ as a function of $E[R_c^2]$ for all the 159 patients. Four clusters evolve in this 2 dimensional feature space with two of them predominantly containing seizure free patients and the other two containing non seizure free patients.

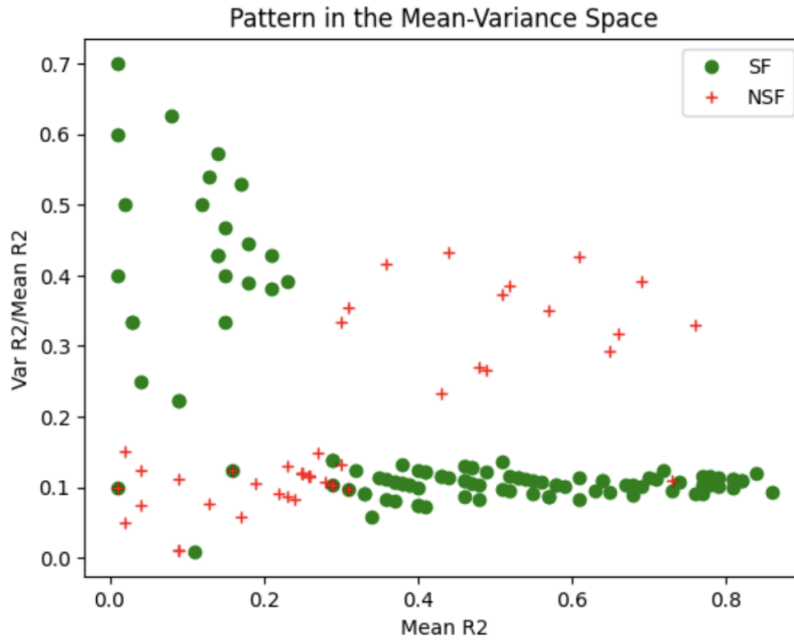


Figure 5.37: **Pattern in the correlation feature space:** The ratio of variance to mean $\frac{Var[R_c^2]}{E[R_c^2]}$ is plotted against the mean $E[R_c^2]$ for all 159 patients. Patients are color-coded based on their ground truth surgical outcome: green circles depict seizure free patients, while red crosses represent non-seizure free patients. A distinct separation between the two patient groups is evident in this two-dimensional feature space

Performing a k -nearest neighbor clustering with 10 fold stratified cross validation on the correlation feature space we get a mean $F1$ score of 93%. Therefore, there exists an interesting structure in the correlation space that's enabling the 7 dimensional model to have a near perfect (Fig 5.36) surgical outcome prediction power

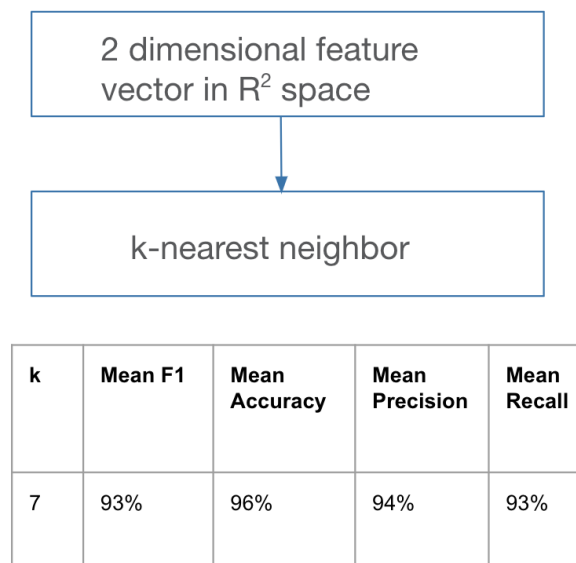


Figure 5.38: **k -nearest neighbor clustering in the correlation feature space:** In the 2-dimensional correlation feature space, we conducted a k -nearest neighbor clustering with $k = 7$. The clustering algorithm was trained and validated using 10-fold stratified cross-validation, resulting in an impressive mean $F1$ score of 93%

One interesting thing to note is that these two features were constructed by exploiting the correlation between the interictal network properties and the spk-HFO rate and is completely agnostic to surgical resection and SOZ labels. Although, we have unearthed a minimal set of features that can guide surgical resections and help achieve better outcomes in a completely unsupervised manner but there is more information encoded in the four clusters that we observed in Figure 5.37. In the next section, we analyze the four clusters and pinpoint some findings that can help the doctors sample the brain regions for optimal surgical resection.

5.6.1 Analyzing the four clusters

Clustering the patients in the 2 dimensional correlation space, we get the following four clusters:

- **Cluster 1:** contains non-seizure free patients only and most patients in this cluster have a high mean R_c^2 ($E[R_c^2] > 0.3$) and a high variance in R_c^2 ($\frac{Var[R_c^2]}{E[R_c^2]} > 0.2$)
- **Cluster 4:** contains mostly seizure free patients and majority of the patients in this cluster have a high mean R_c^2 ($E[R_c^2] > 0.3$) and a low variance in R_c^2 ($\frac{Var[R_c^2]}{E[R_c^2]} < 0.2$)
- **Cluster 2:** contains seizure free patients only and majority of the patients in this cluster have a low mean R_c^2 ($E[R_c^2] < 0.3$) and a high variance in R_c^2 ($\frac{Var[R_c^2]}{E[R_c^2]} > 0.2$)
- **Cluster 3:** contains mostly non-seizure free patients and majority of the patients in this cluster have a low mean R_c^2 ($E[R_c^2] < 0.3$) and a low variance in R_c^2 ($\frac{Var[R_c^2]}{E[R_c^2]} < 0.2$)

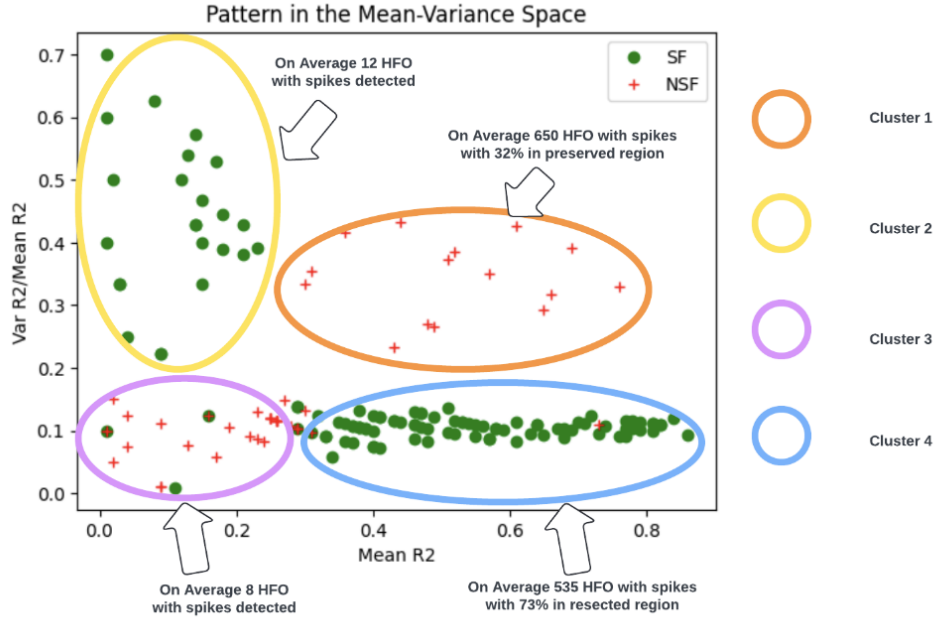


Figure 5.39: **Four clusters in the correlation feature space:** In the 2-dimensional correlation feature space, four distinct clusters emerge. Among them, Cluster 2 and Cluster 4 are primarily composed of seizure free patients, whereas Cluster 1 and Cluster 3 predominantly consist of non-seizure free patients

5.6.1.1 Cluster 1: high mean high variance cluster of non-seizure free patients

To gain insights into why patients in Cluster 1 continue to experience seizures despite surgery, we examined the correlation between biomarkers, specifically interictal network properties and spk-HFO rate. We illustrated this correlation by plotting the channel level R^2 values for a representative patient from Cluster 1 (see Figure 5.40). The plot reveals a substantial number of channels with low R^2 values indicating poor explanatory power of interictal network properties for the variability in spk-HFO rate within those channels.

Given that R^2 quantifies the degree to which interictal network properties can account for

the variance in spk-HFO rate, these findings suggest that the constructed interictal network, based on sampled iEEG electrode data, falls short in elucidating the underlying pathological events recorded in certain channels. This discrepancy hints at the presence of an additional epileptic network that remains unexplored by the current electrode placement, potentially responsible for generating these pathological events.

While the interictal network biomarker proposed in Section 5.5.2, determined by the interictal SOZ classifier prediction scores and network flow dynamics, adequately accounts for the surgical outcomes of patients in cluster 1—predicted as non-seizure free due to predicted SOZ channels in the preserved region with adequate flow to the rest of the preserved channels—upon closer examination, it appears these patients were unlikely to attain seizure freedom even if all predicted SOZ channels were resected. This is likely attributable to the presence of an unsampled epileptic network.

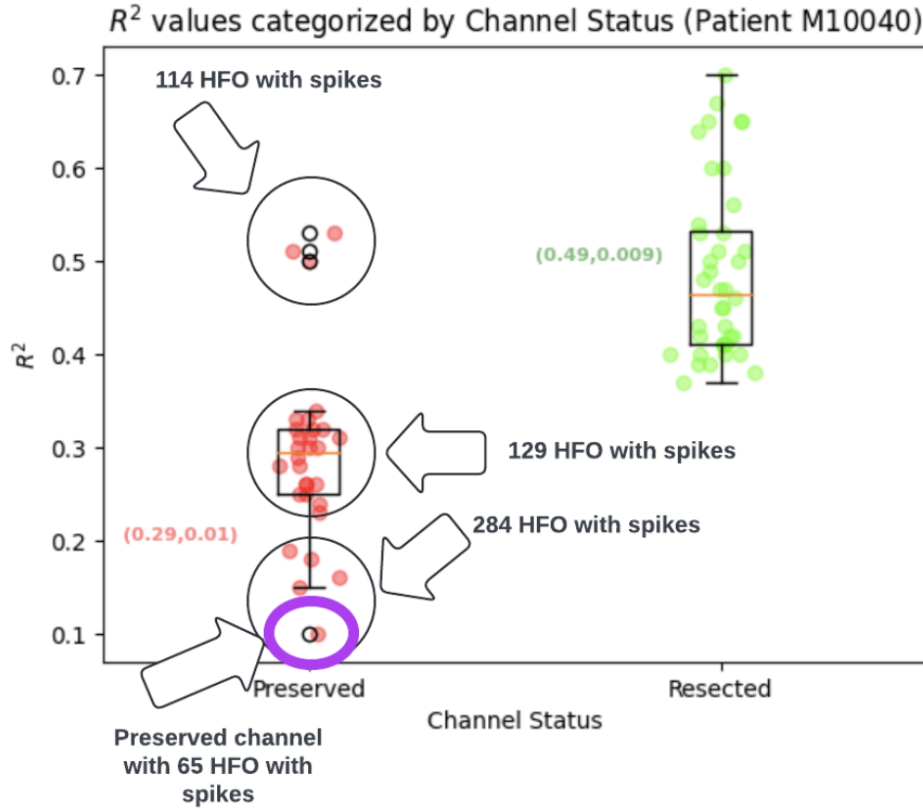


Figure 5.40: **Box plot of channel level R^2 values of a patient from Cluster 1:** For a patient in cluster 1, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. A significant number of channels with spk-HFO detections, have a low R^2 value, suggesting that a sizeable number of detected spk-HFOs remain unexplained by the sampled interictal network

5.6.1.2 Cluster 4: high mean low variance cluster of seizure free patients

To uncover the reasons behind the seizure free status of patients in Cluster 4 following surgery, we delved into the correlation between biomarkers, specifically examining interictal network properties and spk-HFO rate. Demonstrating this correlation, we plotted channel

level R^2 values for a representative patient from Cluster 4 (refer to Figure 5.41). The plot reveals predominantly high R^2 values across most channels, resulting in a high mean and low variance in R_c^2 overall. Given that nearly all spk-HFO occurrences are encompassed by the sampled network, it is highly likely that this precludes the presence of an unsampled epileptic network. Thus, through precise identification and surgical resection of the sampled epileptic network, one can achieve seizure freedom.

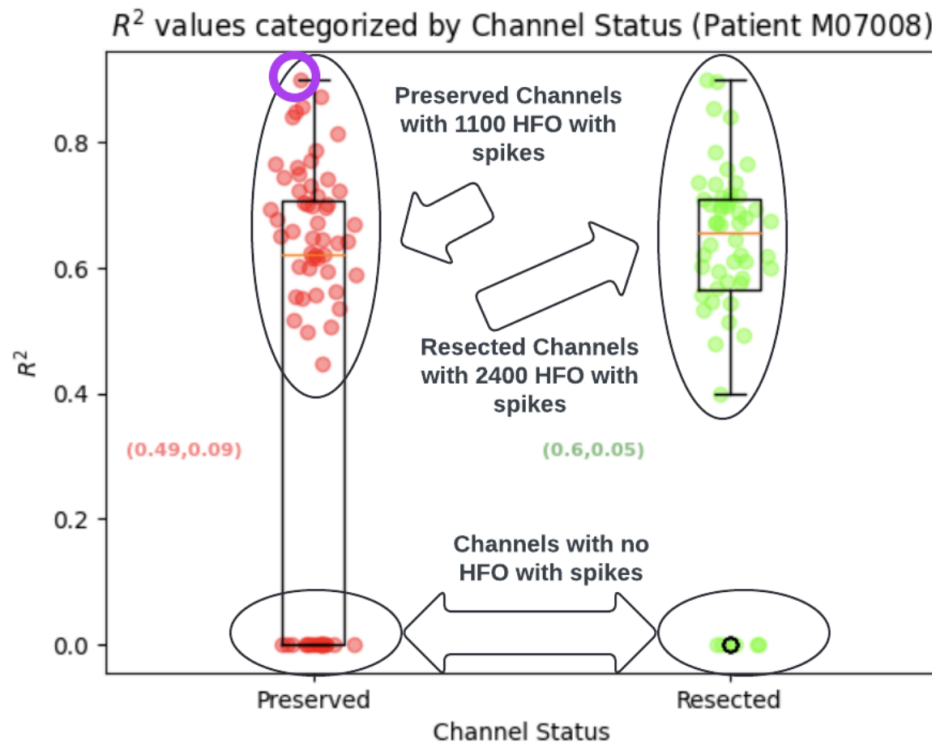


Figure 5.41: **Box plot of channel level R^2 values of a patient from Cluster 4:** For a patient in cluster 4, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. Majority of the channels with spk-HFO detections, have a high R^2 value, suggesting that most of the spk-HFOs detected are explained by the sampled interictal network

5.6.1.3 Cluster 2: low mean high variance cluster of seizure free patients

To understand why patients in Cluster 2 remain seizure-free post-surgery, we closely examined the relationship between two important biomarkers: interictal network properties and spk-HFO rate. This correlation was illustrated by plotting the channel level R^2 values for a representative patient from Cluster 2 (see Figure 5.42). Due to the limited number of detected spk-HFO in patients from these clusters, many channels display zero R^2 values, contributing to the overall low mean R_c^2 .

However, among the channels where some spk-HFO were detected, they exhibited high R^2 values, leading to a substantial variance in R_c^2 overall. Since almost all spk-HFO occurrences are accounted for by the sampled network, it's highly probable that this negates the existence of an unsampled epileptic network. Therefore, by accurately identifying and resecting the sampled epileptic network one can achieve seizure freedom.

tive patient from Cluster 3 (see Figure 5.43). Given the low number of detected spk-HFO in patients from these clusters, most channels exhibit zero R^2 values contributing to the overall low mean R_c^2 .

Nevertheless, among the channels where some spk-HFO were detected, their low R^2 values suggest the presence of an additional epileptic network that remains unexplored by the current electrode placement. This unseen network could potentially be responsible for generating these isolated pathological events, even after surgery.

The 5 dimensional model (Section 5.5.2), integrating interictal SOZ classifier prediction scores and network flow dynamics, erroneously labeled patients in Cluster 3 as seizure free due to the removal of all predicted SOZ channels. However, a deeper analysis, incorporating the correlation between interictal network properties and spk-HFO rate, in the 7-dimensional model (Section 5.5.3), reveals an unsampled epileptic network. This refinement significantly improves outcome prediction precision, reducing false positives from 23 to 2 (Figure 5.36).

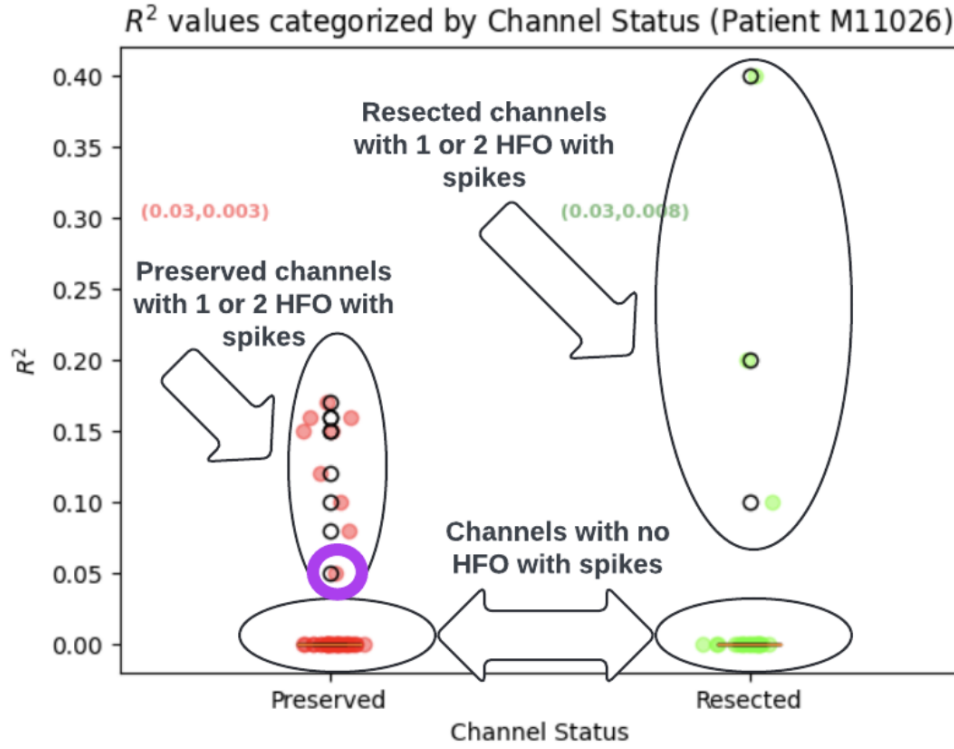


Figure 5.43: **Box plot of channel level R^2 values of a patient from Cluster 3:** For a patient in cluster 3, channel level R^2 values categorized by channel resection status are plotted, with green points indicating resected channels and red points representing preserved channels. The majority of channels exhibit zero spk-HFO detections, resulting in zero R^2 values. Among channels with a few spk-HFO detections, their R^2 values are low, suggesting that the spk-HFO rates of these channels remain unexplained by the sampled interictal network property

CHAPTER 6

Concluding remarks and future work

Our interictal synchronization network based approach to EZ identification has uncovered a significant revelation: even during the interictal period, all the channels in EZ remain active and serve as pivotal hubs, driving frequent hypersynchronous activity throughout the network. This discovery empowers us to accurately pinpoint the EZ even in the absence of observable seizure episodes. We propose that these episodes of hypersynchronous interictal activity resemble 'mini seizures,' a phenomenon previously discussed in the literature. To our knowledge, our method stands as one of the pioneering approaches to detect and analyze mini seizures in the interictal period. The transition from 'mini seizures' to full-blown epileptic seizures bears intriguing similarities to earthquakes, where numerous minor tremors precede major ones. This intriguing aspect of epilepsy prompts further exploration in future research endeavors. One plausible hypothesis regarding transitions from mini seizures to actual seizures could be based on the so-called kindling effect observed in experimental epilepsy models [Bertram, 2007]. In this phenomenon, repeated subthreshold stimulation gradually decreases the seizure initiation threshold. Over time, a series of mini seizures may cumulatively sensitize the brain, rendering it more prone to subsequent seizure activity. This sensitization process likely contributes to the progression from smaller-scale to larger-scale seizure events.

The correlation between pathological high-frequency oscillations (HFOs) and interictal network properties, as observed in our experiments leading up to the development of the 7 dimensional feature vector (Section 5.5.3), hints at an intricate relationship between the

epileptic network and channel level pathological events. In forthcoming studies, our aim is to inquire into the underlying mechanism driving the emergence of these channel level pathological events from network synchronization. One plausible theory for this mechanism could be rooted in emergent properties. The emergence of pathological HFOs from network synchronization may be influenced by various emergent properties of the neural network, such as oscillatory resonance or network excitability thresholds [Memmesheimer, 2010]. These emergent properties possess the capability to mold the spatiotemporal dynamics of synchronized activity and play a pivotal role in the generation of HFOs. Further exploration of these mechanisms promises to shed light on the intricate workings of epileptic networks and pave the way for more effective therapeutic interventions.

In conclusion, this thesis has demonstrated our capability to accurately estimate EZ from interictal segments, even with durations as short as 5 minutes. Building upon this foundation, the development of a real time platform holds tremendous promise for clinical applications. Such a platform could be integrated into existing EEG monitoring systems to continuously analyze patient data in real time. By providing instantaneous feedback to clinicians, this platform could aid in making critical decisions regarding resection margins during epilepsy surgery. The potential benefits of a real time platform are manifold. First and foremost, it could significantly reduce the time required for surgical planning, allowing doctors to promptly identify and delineate the epileptogenic zone during preoperative evaluations thus cutting down the invasive monitoring time which is fraught with risk of infection. Second, as observed in our retrospective studies, patients who were not seizure free after surgery—especially patients for whom all SOZ and PSOZ channels were resected—had HFOs with spike that could not be predicted by the epileptic network. Such a scenario would occur if the electrode placement did not cover the entire EZ, leaving parts of the epileptic network unobserved, which in turn could generate HFOs with spike that cannot be predicted by the observed network. If this observation is proven to be true in clinical trials then our method could provide feedback to clinicians about a critical diagnostic step: they could add new

electrodes to cover the entirety of EZ and hence obtain better surgical outcomes. Third, in addition to guiding electrode placement, our platform could lead to more streamlined surgical procedures and ultimately improve patient outcomes. For example, we observed that a number of patients became seizure free even when only a subset of the SOZ channels were resected and others were left intact in the preserved region. We showed that such cases corresponded to an effective technique of disrupting the epileptic network without complete resection: the preserved SOZ channels did not have enough interaction strength with the rest of the preserved brain regions to cause seizures. This observation naturally leads to the following question: Given the interictal networks, can one design an optimal strategy that would resect a minimum volume of brain tissue while protecting the eloquent cortices, such that the preserved regions do not have enough connections to lead to seizures. An automated algorithm could analyze the interictal synchronization networks and generate multiple such resection strategies to assist neurosurgeons.

REFERENCES

- [Abbott et al., 2016] Abbott, A. E., Nair, A., Keown, C. L., Datko, M., Jahedi, A., Fishman, I., and Müller, R.-A. (2016). Patterns of atypical functional connectivity and behavioral links in autism differ between default, salience, and executive networks. *Cerebral cortex*, 26(10):4034–4045.
- [Akiyama et al., 2011] Akiyama, T., McCoy, B., Go, C. Y., Ochi, A., Elliott, I. M., Akiyama, M., Donner, E. J., Weiss, S. K., Snead III, O. C., Rutka, J. T., et al. (2011). Focal resection of fast ripples on extraoperative intracranial eeg improves seizure outcome in pediatric epilepsy. *Epilepsia*, 52(10):1802–1811.
- [Amiri et al., 2016] Amiri, M., Lina, J.-M., Pizzo, F., and Gotman, J. (2016). High frequency oscillations and spikes: separating real hfos from false oscillations. *Clinical Neurophysiology*, 127(1):187–196.
- [Arora et al., 2012a] Arora, S., Ge, R., Kannan, R., and Moitra, A. (2012a). Computing a Nonnegative Matrix Factorization–Provably. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 145–162. ACM.
- [Arora et al., 2012b] Arora, S., Ge, R., and Moitra, A. (2012b). Learning Topic Models–Going Beyond SVD. In *Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on*, pages 1–10. IEEE.
- [Baud et al., 2018] Baud, M. O., Kleen, J. K., Anumanchipalli, G. K., Hamilton, L. S., Tan, Y.-L., Knowlton, R., and Chang, E. F. (2018). Unsupervised learning of spatiotemporal interictal discharges in focal epilepsy. *Neurosurgery*, 83(4):683–691.
- [Beleza, 2009] Beleza, P. (2009). Refractory epilepsy: a clinically oriented review. *European neurology*, 62(2):65–71.
- [Berry et al., 2007] Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., and Plemmons, R. J. (2007). Algorithms and applications for approximate nonnegative matrix factorization. *Computational statistics & data analysis*, 52(1):155–173.
- [Bertram, 2007] Bertram, E. (2007). The relevance of kindling for human epilepsy. *Epilepsia*, 48:65–74.
- [Bertrand and Moonen, 2013] Bertrand, A. and Moonen, M. (2013). Distributed computation of the fiedler vector with application to topology inference in ad hoc networks. *Signal Processing*, 93(5):1106–1117.
- [Bomela et al., 2020] Bomela, W., Wang, S., Chou, C.-A., and Li, J.-S. (2020). Real-time inference and detection of disruptive eeg networks for epileptic seizures. *Scientific Reports*, 10(1):8653.

- [Boyd et al., 2011] Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al. (2011). Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122.
- [Bragin et al., 1999a] Bragin, A., Engel Jr, J., Wilson, C. L., Fried, I., and Buzsáki, G. (1999a). High-frequency oscillations in human brain. *Hippocampus*, 9(2):137–142.
- [Bragin et al., 1999b] Bragin, A., Engel Jr, J., Wilson, C. L., Fried, I., and Mathern, G. W. (1999b). Hippocampal and entorhinal cortex high-frequency oscillations (100–500 hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures. *Epilepsia*, 40(2):127–137.
- [Bragin et al., 1999c] Bragin, A., Engel Jr, J., Wilson, C. L., Vizingin, E., and Mathern, G. W. (1999c). Electrophysiologic analysis of a chronic seizure model after unilateral hippocampal ka injection. *Epilepsia*, 40(9):1210–1221.
- [Brown et al., 2008] Brown, E. C., Rothermel, R., Nishida, M., Juhász, C., Muzik, O., Hoechstetter, K., Sood, S., Chugani, H. T., and Asano, E. (2008). In vivo animation of auditory-language-induced gamma-oscillations in children with intractable focal epilepsy. *Neuroimage*, 41(3):1120–1131.
- [Cardinale et al., 2019] Cardinale, F., Rizzi, M., Vignati, E., Cossu, M., Castana, L., d’Orio, P., Revay, M., Costanza, M. D., Tassi, L., Mai, R., et al. (2019). Stereoelectroencephalography: retrospective analysis of 742 procedures in a single centre. *Brain*, 142(9):2688–2704.
- [Carlini and Rapallo, 2010] Carlini, E. and Rapallo, F. (2010). Probability matrices, non-negative rank, and parameterization of mixture models. *Linear Algebra and its Applications*, 433(2):424–432.
- [Chaibi et al., 2014] Chaibi, S., Lajnef, T., Samet, M., Jerbi, K., and Kachouri, A. (2014). Detection of high frequency oscillations (hfos) in the 80–500 hz range in epilepsy recordings using decision tree analysis. In *International image processing, applications and systems conference*, pages 1–6. IEEE.
- [Charupanit et al., 2020] Charupanit, K., Sen-Gupta, I., Lin, J. J., and Lopour, B. A. (2020). Amplitude of high frequency oscillations as a biomarker of the seizure onset zone. *Clinical Neurophysiology*, 131(11):2542–2550.
- [Che et al., 2021] Che, H., Wang, J., and Cichocki, A. (2021). Bicriteria sparse nonnegative matrix factorization via two-timescale duplex neurodynamic optimization. *IEEE Transactions on Neural Networks and Learning Systems*.
- [Chen and Li, 2022] Chen, M. and Li, X. (2022). Entropy minimizing matrix factorization. *IEEE Transactions on Neural Networks and Learning Systems*.

- [Cichocki and Phan, 2009] Cichocki, A. and Phan, A.-H. (2009). Fast local algorithms for large scale nonnegative matrix and tensor factorizations. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 92(3):708–721.
- [Cimbalnik et al., 2018] Cimbalnik, J., Brinkmann, B., Kremen, V., Jurak, P., Berry, B., Gompel, J. V., Stead, M., and Worrell, G. (2018). Physiological and pathological high frequency oscillations in focal epilepsy. *Annals of clinical and translational neurology*, 5(9):1062–1076.
- [Crépon et al., 2010] Crépon, B., Navarro, V., Hasboun, D., Clemenceau, S., Martinerie, J., Baulac, M., Adam, C., and Le Van Quyen, M. (2010). Mapping interictal oscillations greater than 200 hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*, 133(1):33–45.
- [Donoho and Stodden, 2004] Donoho, D. and Stodden, V. (2004). When Does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts? In *Advances in neural information processing systems*, pages 1141–1148.
- [Egerton et al., 2020] Egerton, A., Grace, A. A., Stone, J., Bossong, M. G., Sand, M., and McGuire, P. (2020). Glutamate in schizophrenia: Neurodevelopmental perspectives and drug development. *Schizophrenia research*, 223:59–70.
- [Essid and Févotte, 2013] Essid, S. and Févotte, C. (2013). Smooth Nonnegative Matrix Factorization for Unsupervised Audiovisual Document Structuring. *IEEE Transactions on Multimedia*, 15(2):415–425.
- [Esteller et al., 2001] Esteller, M., Corn, P. G., Baylin, S. B., and Herman, J. G. (2001). A gene hypermethylation profile of human cancer. *Cancer research*, 61(8):3225–3229.
- [Fedele et al., 2016] Fedele, T., van’t Klooster, M., Burnos, S., Zweiphenning, W., van Klink, N., Leijten, F., Zijlmans, M., and Sarnthein, J. (2016). Automatic detection of high frequency oscillations during epilepsy surgery predicts seizure outcome. *Clinical neurophysiology*, 127(9):3066–3074.
- [Fernández-Torre, 2010] Fernández-Torre, J. L. (2010). 101 interictal eeg.
- [Févotte et al., 2009] Févotte, C., Bertin, N., and Durrieu, J.-L. (2009). Nonnegative Matrix Factorization with the Itakura-Saito Divergence: With Application to Music Analysis. *Neural computation*, 21(3):793–830.
- [Frauscher et al., 2017] Frauscher, B., Bartolomei, F., Kobayashi, K., Cimbalnik, J., van ‘t Klooster, M. A., Rampp, S., Otsubo, H., Höller, Y., Wu, J. Y., Asano, E., et al. (2017). High-frequency oscillations: the state of clinical research. *Epilepsia*, 58(8):1316–1329.

- [Frauscher et al., 2018] Frauscher, B., von Ellenrieder, N., Zermann, R., Rogers, C., Nguyen, D. K., Kahane, P., Dubeau, F., and Gotman, J. (2018). High-frequency oscillations in the normal human brain. *Annals of neurology*, 84(3):374–385.
- [Galantucci et al., 2017] Galantucci, S., Agosta, F., Stefanova, E., Basaia, S., Van Den Heuvel, M. P., Stojković, T., Canu, E., Stanković, I., Spica, V., Copetti, M., et al. (2017). Structural brain connectome and cognitive impairment in parkinson disease. *Radiology*, 283(2):515–525.
- [Gao et al., 2020] Gao, W., Goldfarb, D., and Curtis, F. E. (2020). Admm for multiaffine constrained optimization. *Optimization Methods and Software*, 35(2):257–303.
- [Gardner et al., 2007] Gardner, A. B., Worrell, G. A., Marsh, E., Dlugos, D., and Litt, B. (2007). Human and automated detection of high-frequency oscillations in clinical intracranial eeg recordings. *Clinical neurophysiology*, 118(5):1134–1143.
- [Geier and Lehnertz, 2017] Geier, C. and Lehnertz, K. (2017). Which brain regions are important for seizure dynamics in epileptic networks? influence of link identification and eeg recording montage on node centralities. *International Journal of Neural Systems*, 27(01):1650033.
- [Gillis, 2017] Gillis, N. (2017). Introduction to nonnegative matrix factorization. *arXiv preprint arXiv:1703.00663*.
- [Gillis and Glineur, 2012a] Gillis, N. and Glineur, F. (2012a). Accelerated multiplicative updates and hierarchical als algorithms for nonnegative matrix factorization. *Neural computation*, 24(4):1085–1105.
- [Gillis and Glineur, 2012b] Gillis, N. and Glineur, F. (2012b). On the geometric interpretation of the nonnegative rank. *Linear Algebra and its Applications*, 437(11):2685–2712.
- [Golby et al., 2005] Golby, A., Silverberg, G., Race, E., Gabrieli, S., O’Shea, J., Knierim, K., Stebbins, G., and Gabrieli, J. (2005). Memory encoding in alzheimer’s disease: an fmri study of explicit and implicit memory. *Brain*, 128(4):773–787.
- [Gotman, 2010] Gotman, J. (2010). High frequency oscillations: the new eeg frontier? *Epilepsia*, 51(Suppl 1):63.
- [Gower et al., 2004] Gower, J. C., Dijksterhuis, G. B., et al. (2004). *Procrustes problems*, volume 30. Oxford University Press on Demand.
- [Grieder et al., 2018] Grieder, M., Wang, D. J., Dierks, T., Wahlund, L.-O., and Jann, K. (2018). Default mode network complexity and cognitive decline in mild alzheimer’s disease. *Frontiers in neuroscience*, 12:388987.

- [Guan et al., 2012] Guan, N., Tao, D., Luo, Z., and Yuan, B. (2012). Online nonnegative matrix factorization with robust stochastic approximation. *IEEE Transactions on Neural Networks and Learning Systems*, 23(7):1087–1099.
- [Guragain et al., 2018] Guragain, H., Cimbalka, J., Stead, M., Groppe, D. M., Berry, B. M., Kremen, V., Kenney-Jung, D., Britton, J., Worrell, G. A., and Brinkmann, B. H. (2018). Spatial variation in high-frequency oscillation rates and amplitudes in intracranial eeg. *Neurology*, 90(8):e639–e646.
- [Guth et al., 2021] Guth, T. A., Kunz, L., Brandt, A., Dümpelmann, M., Klotz, K. A., Reinacher, P. C., Schulze-Bonhage, A., Jacobs, J., and Schönberger, J. (2021). Interictal spikes with and without high-frequency oscillation have different single-neuron correlates. *Brain*, 144(10):3078–3088.
- [Hajinezhad et al., 2016] Hajinezhad, D., Chang, T.-H., Wang, X., Shi, Q., and Hong, M. (2016). Nonnegative matrix factorization using admm: Algorithm and convergence analysis. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 4742–4746. IEEE.
- [Haneef and Chiang, 2014] Haneef, Z. and Chiang, S. (2014). Clinical correlates of graph theory findings in temporal lobe epilepsy. *Seizure*, 23(10):809–818.
- [Harper and Konstan, 2016] Harper, F. M. and Konstan, J. A. (2016). The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)*, 5(4):19.
- [Hastie et al., 2015] Hastie, T., Mazumder, R., Lee, J. D., and Zadeh, R. (2015). Matrix Completion and Low-Rank SVD via Fast Alternating Least Squares. *The Journal of Machine Learning Research*, 16(1):3367–3402.
- [Hong et al., 2016] Hong, M., Luo, Z.-Q., and Razaviyayn, M. (2016). Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems. *SIAM Journal on Optimization*, 26(1):337–364.
- [Hoogteijling and Zijlmans, 2022] Hoogteijling, S. and Zijlmans, M. (2022). Deep learning for epileptogenic zone delineation from the invasive eeg: challenges and lookouts. *Brain Communications*, 4(1):fcab307.
- [Huang et al., 2016] Huang, K., Sidiropoulos, N. D., and Liavas, A. P. (2016). A flexible and efficient algorithmic framework for constrained matrix and tensor factorization. *IEEE Transactions on Signal Processing*, 64(19):5052–5065.
- [Jacobs et al., 2018] Jacobs, J., Wu, J. Y., Perucca, P., Zelmann, R., Mader, M., Dubeau, F., Mathern, G. W., Schulze-Bonhage, A., and Gotman, J. (2018). Removing high-frequency oscillations: A prospective multicenter study on seizure outcome. *Neurology*, 91(11):e1040–e1052.

- [Jacobs et al., 2010a] Jacobs, J., Zijlmans, M., Zelmann, R., Chatillon, C.-É., Hall, J., Olivier, A., Dubeau, F., and Gotman, J. (2010a). High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 67(2):209–220.
- [Jacobs et al., 2010b] Jacobs, V. R., Lamb, L. L., and Philipp, R. A. (2010b). Professional noticing of children’s mathematical thinking. *Journal for research in mathematics education*, 41(2):169–202.
- [Jacobs et al., 2008] Jacobs, Z., Wintle, A. G., Roberts, R. G., and Duller, G. A. (2008). Equivalent dose distributions from single grains of quartz at sibudu, south africa: context, causes and consequences for optical dating of archaeological deposits. *Journal of Archaeological Science*, 35(7):1808–1820.
- [Jefferys et al., 2012] Jefferys, J. G., De La Prida, L. M., Wendling, F., Bragin, A., Avoli, M., Timofeev, I., and da Silva, F. H. L. (2012). Mechanisms of physiological and epileptic hfo generation. *Progress in neurobiology*, 98(3):250–264.
- [Jia et al., 2019] Jia, Y., Kwong, S., Hou, J., and Wu, W. (2019). Semi-supervised non-negative matrix factorization with dissimilarity and similarity regularization. *IEEE transactions on neural networks and learning systems*, 31(7):2510–2521.
- [Jing et al., 2020] Jing, J., Sun, H., Kim, J. A., Herlopian, A., Karakis, I., Ng, M., Halford, J. J., Maus, D., Chan, F., Dolatshahi, M., et al. (2020). Development of expert-level automated detection of epileptiform discharges during electroencephalogram interpretation. *JAMA neurology*, 77(1):103–108.
- [Jirsch et al., 2006] Jirsch, J. D., Urrestarazu, E., LeVan, P., Olivier, A., Dubeau, F., and Gotman, J. (2006). High-frequency oscillations during human focal seizures. *Brain*, 129(6):1593–1608.
- [Jrad et al., 2016] Jrad, N., Kachenoura, A., Merlet, I., Bartolomei, F., Nica, A., Biraben, A., and Wendling, F. (2016). Automatic detection and classification of high-frequency oscillations in depth-eeg signals. *IEEE Transactions on Biomedical Engineering*, 64(9):2230–2240.
- [Jrad et al., 2015] Jrad, N., Kachenoura, A., Merlet, I., Nica, A., Bénar, C. G., and Wendling, F. (2015). Classification of high frequency oscillations in epileptic intracerebral eeg. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 574–577. IEEE.
- [Kambara et al., 2018] Kambara, T., Sood, S., Alqatan, Z., Klingert, C., Ratnam, D., Hayakawa, A., Nakai, Y., Luat, A. F., Agarwal, R., Rothermel, R., et al. (2018). Presurgical language mapping using event-related high-gamma activity: The detroit procedure. *Clinical Neurophysiology*, 129(1):145–154.

- [Karimi et al., 2020] Karimi, D., Dou, H., Warfield, S. K., and Gholipour, A. (2020). Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis. *Medical image analysis*, 65:101759.
- [Kim and Park, 2011] Kim, J. and Park, H. (2011). Fast nonnegative matrix factorization: An active-set-like method and comparisons. *SIAM Journal on Scientific Computing*, 33(6):3261–3281.
- [Koga et al., 2011] Koga, S., Rothermel, R., Juhász, C., Nagasawa, T., Sood, S., and Asano, E. (2011). Electrocoricographic correlates of cognitive control in a stroop task—intracranial recording in epileptic patients. *Human brain mapping*, 32(10):1580–1591.
- [Kojima et al., 2013] Kojima, K., Brown, E. C., Rothermel, R., Carlson, A., Fuerst, D., Matsuzaki, N., Shah, A., Atkinson, M., Basha, M., Mittal, S., et al. (2013). Clinical significance and developmental changes of auditory-language-related gamma activity. *Clinical Neurophysiology*, 124(5):857–869.
- [Kucewicz et al., 2014] Kucewicz, M. T., Cimbalka, J., Matsumoto, J. Y., Brinkmann, B. H., Bower, M. R., Vasoli, V., Sulc, V., Meyer, F., Marsh, W., Stead, S., et al. (2014). High frequency oscillations are associated with cognitive processing in human recognition memory. *Brain*, 137(8):2231–2244.
- [Kuroda et al., 2021] Kuroda, N., Sonoda, M., Miyakoshi, M., Nariai, H., Jeong, J.-W., Motoi, H., Luat, A. F., Sood, S., and Asano, E. (2021). Objective interictal electrophysiology biomarkers optimize prediction of epilepsy surgery outcome. *Brain communications*, 3(2):fcab042.
- [Lagarde et al., 2022] Lagarde, S., Bénar, C.-G., Wendling, F., and Bartolomei, F. (2022). Interictal functional connectivity in focal refractory epilepsies investigated by intracranial eeg. *Brain connectivity*, 12(10):850–869.
- [Lee and Seung, 2001] Lee, D. D. and Seung, H. S. (2001). Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, pages 556–562.
- [Lee et al., 2005] Lee, K., Ho, J., and Kriegman, D. (2005). Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 27(5):684–698.
- [Lee et al., 2009] Lee, T., Shraibman, A., et al. (2009). Lower bounds in communication complexity. *Foundations and Trends® in Theoretical Computer Science*, 3(4):263–399.
- [Lin, 2007a] Lin, C.-J. (2007a). On the convergence of multiplicative update algorithms for nonnegative matrix factorization. *IEEE Transactions on Neural Networks*, 18(6):1589–1596.

- [Lin, 2007b] Lin, C.-J. (2007b). Projected gradient methods for nonnegative matrix factorization. *Neural computation*, 19(10):2756–2779.
- [Lin and Boutros, 2018] Lin, X. and Boutros, P. C. (2018). Fast Nonnegative Matrix Factorization and Applications to Pattern Extraction, Deconvolution and Imputation. *bioRxiv*, page 321802.
- [Linazasoro, 2009] Linazasoro, G. (2009). A global view of parkinson’s disease pathogenesis: Implications for natural history and neuroprotection. *Parkinsonism & related disorders*, 15(6):401–405.
- [Liu et al., 2018] Liu, S., Gurses, C., Sha, Z., Quach, M. M., Sencer, A., Bebek, N., Curry, D. J., Prabhu, S., Tummala, S., Henry, T. R., et al. (2018). Stereotyped high-frequency oscillations discriminate seizure onset zones and critical functional cortex in focal epilepsy. *Brain*, 141(3):713–730.
- [Liu et al., 2016] Liu, T., Gong, M., and Tao, D. (2016). Large-cone nonnegative matrix factorization. *IEEE transactions on neural networks and learning systems*, 28(9):2129–2142.
- [Lüders et al., 2006] Lüders, H. O., Najm, I., Nair, D., Widdess-Walsh, P., and Bingman, W. (2006). The epileptogenic zone: general principles. *Epileptic disorders*, 8:S1–S9.
- [Lundervold and Lundervold, 2019] Lundervold, A. S. and Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, 29(2):102–127.
- [Matsumoto et al., 2013] Matsumoto, A., Brinkmann, B. H., Matthew Stead, S., Matsumoto, J., Kucewicz, M. T., Marsh, W. R., Meyer, F., and Worrell, G. (2013). Pathological and physiological high-frequency oscillations in focal human epilepsy. *Journal of neurophysiology*, 110(8):1958–1964.
- [McCarthy et al., 2011] McCarthy, M., Moore-Kochlacs, C., Gu, X., Boyden, E., Han, X., and Kopell, N. (2011). Striatal origin of the pathologic beta oscillations in parkinson’s disease. *Proceedings of the national academy of sciences*, 108(28):11620–11625.
- [Memmesheimer, 2010] Memmesheimer, R.-M. (2010). Quantitative prediction of intermittent high-frequency oscillations in neural networks with supralinear dendritic interactions. *Proceedings of the National Academy of Sciences*, 107(24):11092–11097.
- [Meng et al., 2019] Meng, Y., Shang, R., Shang, F., Jiao, L., Yang, S., and Stolkin, R. (2019). Semi-supervised graph regularized deep nmf with bi-orthogonal constraints for data representation. *IEEE transactions on neural networks and learning systems*, 31(9):3245–3258.

- [Motoi et al., 2019] Motoi, H., Jeong, J.-W., Juhász, C., Miyakoshi, M., Nakai, Y., Sugiura, A., Luat, A. F., Sood, S., and Asano, E. (2019). Quantitative analysis of intracranial electrocorticography signals using the concept of statistical parametric mapping. *Scientific reports*, 9(1):17385.
- [Nakai et al., 2017] Nakai, Y., Jeong, J.-w., Brown, E. C., Rothermel, R., Kojima, K., Kambara, T., Shah, A., Mittal, S., Sood, S., and Asano, E. (2017). Three-and four-dimensional mapping of speech and language in patients with epilepsy. *Brain*, 140(5):1351–1370.
- [Nariai et al., 2019] Nariai, H., Hussain, S. A., Bernardo, D., Fallah, A., Murata, K. K., Nguyen, J. C., Rajaraman, R. R., Rao, L. M., Matsumoto, J. H., Lerner, J. T., et al. (2019). Prospective observational study: Fast ripple localization delineates the epileptogenic zone. *Clinical neurophysiology*, 130(11):2144–2152.
- [Nariai et al., 2011] Nariai, H., Nagasawa, T., Juhász, C., Sood, S., Chugani, H. T., and Asano, E. (2011). Statistical mapping of ictal high-frequency oscillations in epileptic spasms. *Epilepsia*, 52(1):63–74.
- [Nariai et al., 2018] Nariai, H., Wu, J. Y., Bernardo, D., Fallah, A., Sankar, R., and Hussain, S. A. (2018). Interrater reliability in visual identification of interictal high-frequency oscillations on electrocorticography and scalp eeg. *Epilepsia open*, 3:127–132.
- [Navarrete et al., 2016] Navarrete, M., Alvarado-Rojas, C., Le Van Quyen, M., and Valderama, M. (2016). Ripplelab: A comprehensive application for the detection, analysis and classification of high frequency oscillations in electroencephalographic signals. *PloS one*, 11(6):e0158276.
- [Netoff et al., 2004] Netoff, T. I., Clewley, R., Arno, S., Keck, T., and White, J. A. (2004). Epilepsy in small-world networks. *Journal of neuroscience*, 24(37):8075–8083.
- [Nithin et al., 2021] Nithin, G., Sathidevi, P., and Ameer, P. (2021). Graph energy based centrality measures to detect epileptogenic focal invasive eeg electrodes. *Seizure*, 85:127–137.
- [Nonoda et al., 2016] Nonoda, Y., Miyakoshi, M., Ojeda, A., Makeig, S., Juhász, C., Sood, S., and Asano, E. (2016). Interictal high-frequency oscillations generated by seizure onset and eloquent areas may be differentially coupled with different slow waves. *Clinical Neurophysiology*, 127(6):2489–2499.
- [Osorio et al., 2016] Osorio, I., Zaveri, H. P., Frei, M. G., and Arthurs, S. (2016). *Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics*. CRC press.
- [Remakanthakurup Sindhu et al., 2020] Remakanthakurup Sindhu, K., Staba, R., and Lopour, B. A. (2020). Trends in the use of automated algorithms for the detection of high-frequency oscillations associated with human epilepsy. *Epilepsia*, 61(8):1553–1569.

- [Roehri et al., 2018] Roehri, N., Pizzo, F., Lagarde, S., Lambert, I., Nica, A., McGonigal, A., Giusiano, B., Bartolomei, F., and Bénar, C.-G. (2018). High-frequency oscillations are not better biomarkers of epileptogenic tissues than spikes. *Annals of neurology*, 83(1):84–97.
- [Rosenow and Lüders, 2001] Rosenow, F. and Lüders, H. (2001). Presurgical evaluation of epilepsy. *Brain*, 124(9):1683–1700.
- [Ryvlin et al., 2014] Ryvlin, P., Cross, J. H., and Rheims, S. (2014). Epilepsy surgery in children and adults. *The Lancet Neurology*, 13(11):1114–1126.
- [Schmidt et al., 2007] Schmidt, M. N., Larsen, J., and Hsiao, F.-T. (2007). Wind noise reduction using non-negative sparse coding. In *Machine Learning for Signal Processing, 2007 IEEE Workshop on*, pages 431–436. IEEE.
- [Sebastiano et al., 2020] Sebastiano, D. R., Tassi, L., Duran, D., Visani, E., Gozzo, F., Cardinale, F., Nobili, L., Del Sole, A., Rubino, A., Dotta, S., et al. (2020). Identifying the epileptogenic zone by four non-invasive imaging techniques versus stereo-eeeg in mri-negative pre-surgery epilepsy patients. *Clinical Neurophysiology*, 131(8):1815–1823.
- [Sinha et al., 2022] Sinha, N., Johnson, G. W., Davis, K. A., and Englot, D. J. (2022). Integrating network neuroscience into epilepsy care: Progress, barriers, and next steps. *Epilepsy Currents*, 22(5):272–278.
- [Spring et al., 2017] Spring, A. M., Pittman, D. J., Aghakhani, Y., Jirsch, J., Pillay, N., Bello-Espinosa, L. E., Josephson, C., and Federico, P. (2017). Interrater reliability of visually evaluated high frequency oscillations. *Clinical Neurophysiology*, 128(3):433–441.
- [Staba et al., 2004a] Staba, R. J., Bergmann, P. C., and Barth, D. S. (2004a). Dissociation of slow waves and fast oscillations above 200 hz during gaba application in rat somatosensory cortex. *The Journal of Physiology*, 561(1):205–214.
- [Staba et al., 2002] Staba, R. J., Wilson, C. L., Bragin, A., Fried, I., and Engel Jr, J. (2002). Quantitative analysis of high-frequency oscillations (80–500 hz) recorded in human epileptic hippocampus and entorhinal cortex. *Journal of neurophysiology*, 88(4):1743–1752.
- [Staba et al., 2004b] Staba, R. J., Wilson, C. L., Bragin, A., Jhung, D., Fried, I., and Engel Jr, J. (2004b). High-frequency oscillations recorded in human medial temporal lobe during sleep. *Annals of neurology*, 56(1):108–115.
- [Stewart, 2001] Stewart, G. W. (2001). *Matrix Algorithms: Volume II: Eigensystems*. SIAM.
- [Tangherlini et al., 2016] Tangherlini, T. R., Roychowdhury, V., Glenn, B., Crespi, C. M., Bandari, R., Wadia, A., Falahi, M., Ebrahimzadeh, E., and Bastani, R. (2016). “Mommy Blogs” and the vaccination exemption narrative: results from a machine-learning approach for story aggregation on parenting social media sites. *JMIR public health and surveillance*, 2(2).

- [Urrestarazu et al., 2007] Urrestarazu, E., Chander, R., Dubeau, F., and Gotman, J. (2007). Interictal high-frequency oscillations (100–500 hz) in the intracerebral eeg of epileptic patients. *Brain*, 130(9):2354–2366.
- [Van Mierlo et al., 2018] Van Mierlo, P., Lie, O., Staljanssens, W., Coito, A., and Vulliëmoz, S. (2018). Influence of time-series normalization, number of nodes, connectivity and graph measure selection on seizure-onset zone localization from intracranial eeg. *Brain topography*, 31:753–766.
- [van’t Klooster et al., 2015] van’t Klooster, M. A., Van Klink, N. E., Leijten, F. S., Zelmann, R., Gebbink, T. A., Gosselaar, P. H., Braun, K. P., Huiskamp, G. J., and Zijlmans, M. (2015). Residual fast ripples in the intraoperative corticogram predict epilepsy surgery outcome. *Neurology*, 85(2):120–128.
- [van’t Klooster et al., 2015] van’t Klooster, M. A., Leijten, F. S., Huiskamp, G., Ronner, H. E., Baayen, J. C., van Rijen, P. C., Eijkemans, M. J., Braun, K. P., Zijlmans, M., and Group, H. S. (2015). High frequency oscillations in the intra-operative ecog to guide epilepsy surgery (“the hfo trial”): study protocol for a randomized controlled trial. *Trials*, 16:1–10.
- [Vavasis, 2009] Vavasis, S. A. (2009). On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377.
- [Videnovic and Zee, 2015] Videnovic, A. and Zee, P. C. (2015). Consequences of circadian disruption on neurologic health. *Sleep medicine clinics*, 10(4):469–480.
- [von Ellenrieder et al., 2017] von Ellenrieder, N., Dubeau, F., Gotman, J., and Frauscher, B. (2017). Physiological and pathological high-frequency oscillations have distinct sleep-homeostatic properties. *NeuroImage: Clinical*, 14:566–573.
- [Wang et al., 2013] Wang, S., Wang, I. Z., Bulacio, J. C., Mosher, J. C., Gonzalez-Martinez, J., Alexopoulos, A. V., Najm, I. M., and So, N. K. (2013). Ripple classification helps to localize the seizure-onset zone in neocortical epilepsy. *Epilepsia*, 54(2):370–376.
- [Wang et al., 2019] Wang, Y., Yin, W., and Zeng, J. (2019). Global convergence of admm in nonconvex nonsmooth optimization. *Journal of Scientific Computing*, 78:29–63.
- [Wei et al., 2022] Wei, J., Tong, C., Wu, B., He, Q., Qi, S., Yao, Y., and Teng, Y. (2022). An entropy weighted nonnegative matrix factorization algorithm for feature representation. *IEEE Transactions on Neural Networks and Learning Systems*.
- [Weiss et al., 2016] Weiss, S. A., Orosz, I., Salamon, N., Moy, S., Wei, L., Van’t Klooster, M. A., Knight, R. T., Harper, R. M., Bragin, A., Fried, I., et al. (2016). Ripples on spikes show increased phase-amplitude coupling in mesial temporal lobe epilepsy seizure-onset zones. *Epilepsia*, 57(11):1916–1930.

- [Wilke et al., 2011] Wilke, C., Worrell, G., and He, B. (2011). Graph analysis of epileptogenic networks in human partial epilepsy. *Epilepsia*, 52(1):84–93.
- [Wong et al., 2021] Wong, S. M., Arski, O. N., Workewych, A. M., Donner, E., Ochi, A., Otsubo, H., Snead III, O. C., and Ibrahim, G. M. (2021). Detection of high-frequency oscillations in electroencephalography: A scoping review and an adaptable open-source framework. *Seizure*, 84:23–33.
- [Worrell et al., 2008] Worrell, G. A., Gardner, A. B., Stead, S. M., Hu, S., Goerss, S., Cascino, G. J., Meyer, F. B., Marsh, R., and Litt, B. (2008). High-frequency oscillations in human temporal lobe: simultaneous microwire and clinical macroelectrode recordings. *Brain*, 131(4):928–937.
- [Worrell et al., 2012] Worrell, G. A., Jerbi, K., Kobayashi, K., Lina, J.-M., Zelmann, R., and Le Van Quyen, M. (2012). Recording and analysis techniques for high-frequency oscillations. *Progress in neurobiology*, 98(3):265–278.
- [Worrell et al., 2004] Worrell, G. A., Parish, L., Cranstoun, S. D., Jonas, R., Baltuch, G., and Litt, B. (2004). High-frequency oscillations and seizure generation in neocortical epilepsy. *Brain*, 127(7):1496–1506.
- [Wu et al., 2010] Wu, J., Sankar, R., Lerner, J., Matsumoto, J., Vinters, H., and Mathern, G. (2010). Removing interictal fast ripples on electrocorticography linked with seizure freedom in children. *Neurology*, 75(19):1686–1694.
- [Xu et al., 2012] Xu, Y., Yin, W., Wen, Z., and Zhang, Y. (2012). An alternating direction algorithm for matrix completion with nonnegative factors. *Frontiers of Mathematics in China*, 7(2):365–384.
- [Yang et al., 2016] Yang, Z., Xiang, Y., Xie, K., and Lai, Y. (2016). Adaptive method for nonsmooth nonnegative matrix factorization. *IEEE transactions on neural networks and learning systems*, 28(4):948–960.
- [Yuan et al., 2014] Yuan, S., Zhou, W., Yuan, Q., Zhang, Y., and Meng, Q. (2014). Automatic seizure detection using diffusion distance and blda in intracranial eeg. *Epilepsy & Behavior*, 31:339–345.
- [Zelmann et al., 2010] Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Chander, R., and Gotman, J. (2010). Automatic detector of high frequency oscillations for human recordings with macroelectrodes. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 2329–2333. IEEE.
- [Zelmann et al., 2012] Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Dubeau, F., and Gotman, J. (2012). A comparison between detectors of high frequency oscillations. *Clinical Neurophysiology*, 123(1):106–116.

- [Zelmann et al., 2009] Zelmann, R., Zijlmans, M., Jacobs, J., Châtillon, C.-E., and Gotman, J. (2009). Improving the identification of high frequency oscillations. *Clinical Neurophysiology*, 120(8):1457–1464.
- [Zhang et al., 2019] Zhang, H., Mao, H., Long, Y., Yang, W., and Shao, L. (2019). A probabilistic zero-shot learning method via latent nonnegative prototype synthesis of unseen classes. *IEEE transactions on neural networks and learning systems*, 31(7):2361–2375.
- [Zhang et al., 2008] Zhang, T., Fang, B., Liu, W., Tang, Y. Y., He, G., and Wen, J. (2008). Total variation norm-based nonnegative matrix factorization for identifying discriminant representation of image patterns. *Neurocomputing*, 71(10-12):1824–1831.
- [Zhang et al., 2022a] Zhang, Y., Chung, H., Ngo, J. P., Monsoor, T., Hussain, S. A., Matsumoto, J. H., Walshaw, P. D., Fallah, A., Sim, M. S., Asano, E., et al. (2022a). Characterizing physiological high-frequency oscillations using deep learning. *Journal of neural engineering*, 19(6):066027.
- [Zhang et al., 2022b] Zhang, Y., Lu, Q., Monsoor, T., Hussain, S. A., Qiao, J. X., Salamon, N., Fallah, A., Sim, M. S., Asano, E., Sankar, R., et al. (2022b). Refining epileptogenic high-frequency oscillations using deep learning: a reverse engineering approach. *Brain communications*, 4(1):fcab267.
- [Zhao et al., 2020] Zhao, B., Hu, W., Zhang, C., Wang, X., Wang, Y., Liu, C., Mo, J., Yang, X., Sang, L., Ma, Y., et al. (2020). Integrated automatic detection, classification and imaging of high frequency oscillations with stereoelectroencephalography. *Frontiers in Neuroscience*, 14:546.
- [Zhu et al., 2021] Zhu, Z., Li, Q., Tang, G., and Wakin, M. B. (2021). The global optimization geometry of low-rank matrix optimization. *IEEE Transactions on Information Theory*, 67(2):1308–1331.
- [Zijlmans et al., 2011] Zijlmans, M., Jacobs, J., Kahn, Y. U., Zelmann, R., Dubeau, F., and Gotman, J. (2011). Ictal and interictal high frequency oscillations in patients with focal epilepsy. *Clinical Neurophysiology*, 122(4):664–671.
- [Zijlmans et al., 2009] Zijlmans, M., Jacobs, J., Zelmann, R., Dubeau, F., and Gotman, J. (2009). High-frequency oscillations mirror disease activity in patients with epilepsy. *Neurology*, 72(11):979–986.
- [Zijlmans et al., 2012] Zijlmans, M., Jiruska, P., Zelmann, R., Leijten, F. S., Jefferys, J. G., and Gotman, J. (2012). High-frequency oscillations as a new biomarker in epilepsy. *Annals of neurology*, 71(2):169–178.
- [Zimmermann et al., 2024] Zimmermann, M. L., Breedt, L. C., Centeno, E. G., Reijneveld, J. C., Santos, F. A., Stam, C. J., van Lingen, M. R., Schoonheim, M. M., Hillebrand, A.,

- and Douw, L. (2024). The relationship between pathological brain activity and functional network connectivity in glioma patients. *Journal of Neuro-oncology*, pages 1–11.
- [Zuo et al., 2019] Zuo, R., Wei, J., Li, X., Li, C., Zhao, C., Ren, Z., Liang, Y., Geng, X., Jiang, C., Yang, X., et al. (2019). Automated detection of high-frequency oscillations in epilepsy based on a convolutional neural network. *Frontiers in computational neuroscience*, 13:6.
- [Zweiphenning et al., 2022a] Zweiphenning, W., van’t Klooster, M. A., van Klink, N. E., Leijten, F. S., Ferrier, C. H., Gebbink, T., Huiskamp, G., van Zandvoort, M. J., van Schooneveld, M. M., Bourez, M., et al. (2022a). Intraoperative electrocorticography using high-frequency oscillations or spikes to tailor epilepsy surgery in the netherlands (the hfo trial): a randomised, single-blind, adaptive non-inferiority trial. *The Lancet Neurology*, 21(11):982–993.
- [Zweiphenning et al., 2022b] Zweiphenning, W. J., von Ellenrieder, N., Dubeau, F., Martineau, L., Minotti, L., Hall, J. A., Chabardes, S., Dudley, R., Kahane, P., Gotman, J., et al. (2022b). Correcting for physiological ripples improves epileptic focus identification and outcome prediction. *Epilepsia*, 63(2):483–496.