THE INTRINSIC VALUE OF HFO FEATURES AS A BIOMARKER OF EPILEPTIC ACTIVITY

SV Gliske[1], WC Stacey[1,2], KR Moon[3], AO Hero III[2,3,4]

[1] Department of Neurology, University of Michigan
[2] Department of Biomedical Engineering, University of Michigan
[3] Department of Electrical Engineering and Computer Science, University of Michigan
[4] Department of Statistics, University of Michigan

ICASSP 2016
Shanghai, China
22 March, 2016

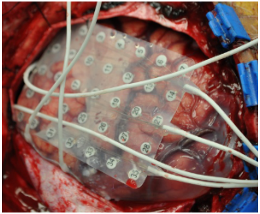
Motivation: Epilepsy

- Fourth most prevalent neurological disease
- Effects about 3% of people worldwide
- Disease characterized by frequent dehabilitating seizures
- Main treatment options:

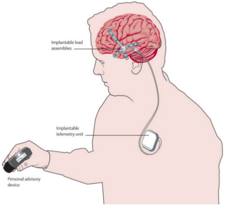
Medications

https://commons.wikimedia.org/wiki/File:VariousPills.jpg

Resective Surgery



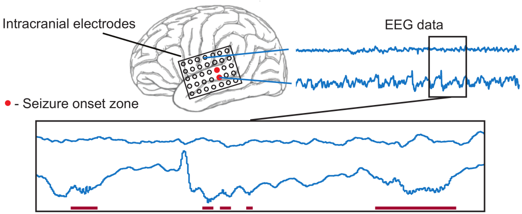
Implanted Devices

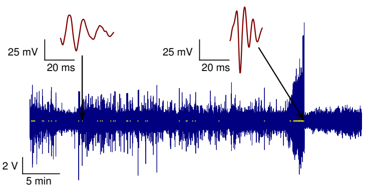


Cook, et al (2014)

- Goal: improve localization of the seizure focus through advanced signal processing to improve surgery outcomes

Signal of Interest: High Frequency Oscillations



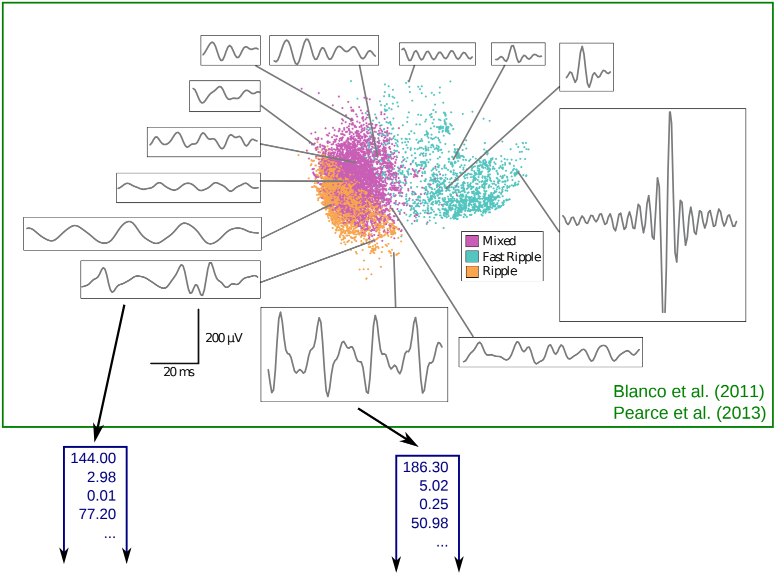


- Strong correlation between HFO rate and the seizure onset zone (SOZ)
  - See, e.g., Gliske et al. (2015)
- However, not all types of HFOs are correlated

	Seizure	not Seizure
Normal Tissue	pHFO	nHFO
Diseased Tissue	pHFO	nHFO & pHFO

- pHFO: pathological HFO
- nHFO: normal HFO

Quantification of HFOs



General Protocol for Classification

Standard Method

1. Feature selection and/or dimensional reduction
2. Selection of classification algorithm
3. Training of the classifier
4. Testing of classifier.

Proposed Method

1. Compute features of the raw data
2. Estimate the topology of the features
3. Compare topology across various confounding factors
4. Reduce dimensionality of the features according to the topology
5. Estimate bounds on Bayes Error
6. Selection of classification algorithm
7. etc.

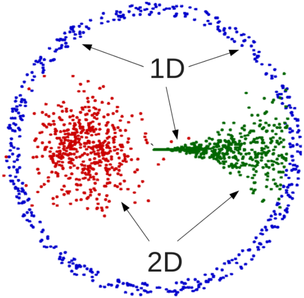
Context of Contributions

- Main contribution: connecting the various methods
- Other contributions:

Previous State	New Contributions
kNN local intrinsic dimension	application to neural data
Angular distance	application to comparing intrinsic dimension
Greedy-LDA	application to HFOs
Generalized Grassman distance	application and modified eq. for Chordal distance
Henze-Penrose Divergence and Bayes Error Estimates	application to neural data
f-Divergence computation	application to neural data

Local Intrinsic Dimension

- Local intrinsic dimension is the local dimension of the submanifold at a given location
- Estimates are provided by a nonlinear k-Nearest Neighbor algorithm (Carter et al., 2010).

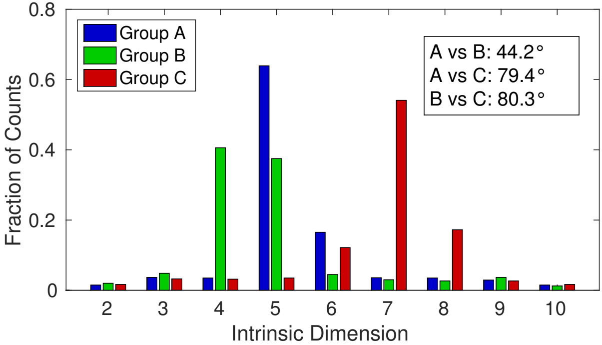


- The basis of the algorithm is a least squares minimization between
  - Total k-NN graph edge length $L_{\gamma,k}(X_n) = \sum_{i=1}^{n} \sum_{y \in \mathcal{N}_{k,i}} D^{\gamma}(y, x_i)$
    - X_n: matrix of n samples; $\mathcal{N}_{k,i}$: k-NN neighborhood of x_i; γ: free parameter
  - Asymptotic functional form of $L_{\gamma,k}(X_n)$, $cn^{1-\gamma/m} + \epsilon_n$.
    - c: constant based on distributions; m: intrinsic dimension; ϵ_n: noise term
- Algorithm is applied to local subsets of data to get local estimates
- To average out ϵ_n, the algorithm bootstraps multiple subblocks of data.

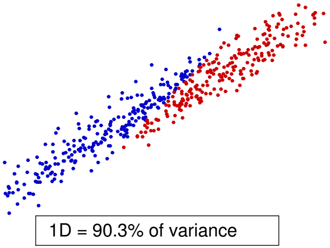
Comparison of Intrinsic Dimension Multisets

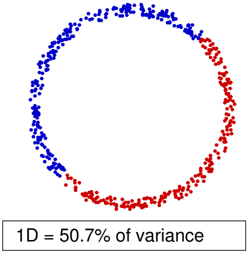
- Address confounding factors by comparing multisets of intrinsic dimension
- Chosen method: angular distance θ_I (Ochiai, 1957; Barkman, 1958)
  - Let A and B be two multisets of integers
  - Let the multiplicity functions be denoted 1_A and 1_B
  - Let $N = \max(A \cup B)$, $n_i = 1_A(i)/|A|$, and $m_i = 1_B(i)/|B|$ $\forall i \leq N$.
  - Angular distance is the Euclidean angle between $\boldsymbol{n}$ and $\boldsymbol{m}$ in $\mathbb{R}^N$

Toy model



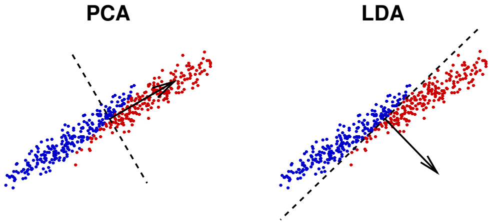
Global linearity of Feature Manifolds





- To assess global linearity, the non-linear local intrinsic dimension is compared with global linear method (PCA)
- Comparison quantified by determining fraction of variance accounted for by mean intrinsic dimension.

Greedy Linear Discriminant Analysis



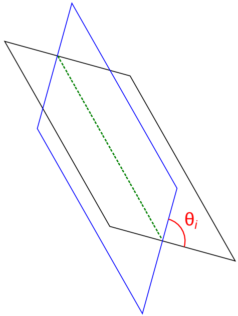
- Often preferable to choose basis based on class separation (LDA) than total variance (PCA)
- Standard Fisher's Linear Discriminant Analysis (LDA) (Fisher, 1936)
  - Direction of best separability is given by

$$\boldsymbol{w} \propto (\Sigma_A + \Sigma_B)^{-1} (\boldsymbol{\mu}_A - \boldsymbol{\mu}_B) .$$

- Greedy LDA (Wang et al., 2010)
  - Apply Fisher's LDA, project out resultant dimension, repeat
  - Set of all $\boldsymbol{w}$ form basis of selected subspace

Comparing Similarity of Subspaces

- To assess if manifold globally linear, we need to compare the subspaces selected by PCA and/or greedy-LDA
- Binary comparison of subspaces accomplished using generalized Grassman-chordal distance
  - Previous work incorporated affine translations with unequal dimension (Ye & Lim, 2014)
  - We additionally included a factor of $1/k$ and converted the quantity back to an angle
  - For k principle angles $\{\theta_i\}_{i=1}^{k}$, the generalized Grassman-chordal distance θ_C is



$$\theta_C = \arcsin\left(\left(\frac{1}{k}\sum_{i=1}^{k}\sin^2\theta_i\right)^{1/2}\right).$$

Bayes Error Estimates

- Bounds on the Bayes error provide an expected range of classification performance
  - Measure of separability between classes in the feature space
  - Benchmark for classification
- Bounds on the estimated Bayes error P_e^* can be estimated in any dimension N using the Henze-Penroze divergence (Moon et al., 2015; Berisha et al., 2015)

$$\frac{1}{2} - \frac{1}{2}\sqrt{\widetilde{D}_{q_1}(p_1, p_2)} \leq P_e^* \leq \frac{1}{2} + \frac{1}{2}\widetilde{D}_{q_1}(p_1, p_2).$$

$$\widetilde{D}_{q_1}(p_1, p_2) = \int d^N \boldsymbol{x} \frac{(q_1 p_1(\boldsymbol{x}) - q_2 p_2(\boldsymbol{x}))^2}{(q_1 p_1(\boldsymbol{x}) + q_2 p_2(\boldsymbol{x}))},$$

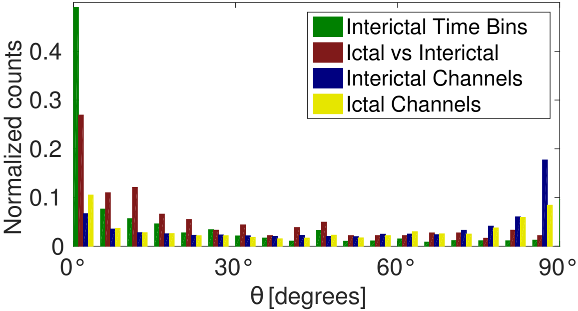
- q_i: normalized prior i
- p_i: pdf i

- Henze-Penrose divergence computed using non-parametric approach of Moon et al. (2014a; 2014b), which achieves the parametric convergence rate.

Patient Population and Data Description

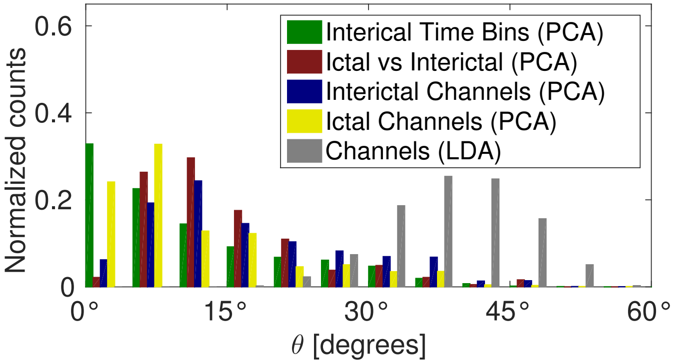
- 17 adult patients from two centers
  - 100,000 channel-hours of recordings with 5 kHz sampling rate
  - >1.6 million HFOs computed using qHFO algorithm (Gliske et al., 2015)
  - 33 features per HFO
    - Duration, peak power, peak frequency, mean Teager-Kaiser energy, various spectral properties, etc.
- To address confounding factors of time, space and brain state, we stratify the data four different ways:
  - Stratify interictal ($>$ 30 min from nearest seizure) HFOs by channels and by 30 minute time windows
  - Stratify HFOs by channels and by ictal or interictal
  - Stratify interictal HFOs by channel
  - Stratify ictal HFOs by channel
- All binary comparisons per patient (per channel) are considered

Variation of Intrinsic Dimension: Results



- Rule of thumb is 0-30° is fairly similar; 30-60° is intermediate; 60-90° is quite dissimilar.
- Temporal variability of intrinsic dimension during interictal times is quite small
- Ictal versus interictal times are also fairly consistent
- However, in some cases HFO features vary significantly from channel to channel
  - One cannot simply aggregate across channels

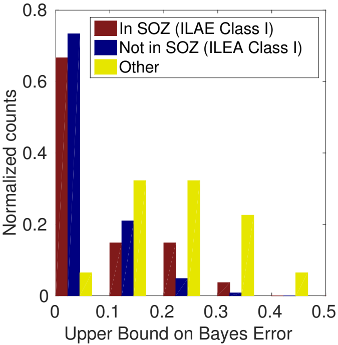
Comparison between Subspaces: Results



- Same rule of thumb is 0-30°: fairly similar; 30-60°: intermediate; 60-90°: quite dissimilar.
- All PCA subspaces are quite similar, but Greedy-LDA are not (40-50°).
  - Again, important variations across channels are observed

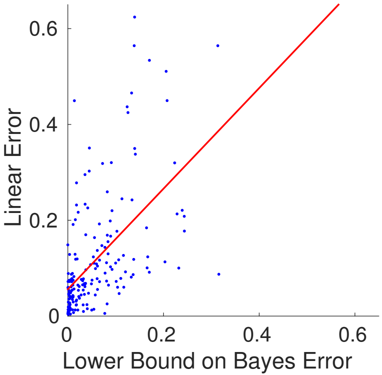
Results of Bayes Error Estimate

- The classification problem: label interictal HFOs as “ictal-like” or “interictal-like”
- Data should be more separable in healthy tissue



- Ictal and interictal HFOs are observed to be fairly distinguishable in ILAE Class I patients
- Poor distinguishability may be a new biomarker for poor surgery outcome

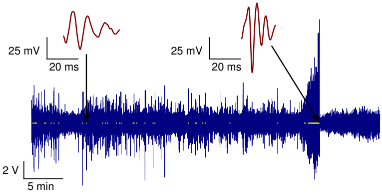
Comparison of Bayes Error Estimate



- We also compared the lower bound with a simple box (greedy LDA) classifier
- Regression line had
  - Offset of 0.06 (0.04–0.08 at 95% C.L.)
  - Slope of 1.05 (0.82–1.28 at 95% C.L.)
- The simple classifier is performing fairly well, given the input data

Conclusions and Outlook

- HFO features vary significantly from channel to channel
  - Variation present in both the intrinsic dimension and the best separating subspace
  - One cannot simply aggregate across channels
- Ictal and interictal HFOs are distinct
  - Promising avenue to identifying pathological HFOs
- Patients where in whom ictal and interictal HFOs are not distinct are likely to have poor surgery outcome.
- These general methods for feature analysis are widely applicable to many large neural data sets
  - We thus propose a standard protocol to prepare neural data for classification



Acknowledgments

- We gratefully acknowledge funding from
  - Doris Duke Foundation (WCS)
  - US National Institute of Health (NIH) grant K08-NS069783-05 (WCS)
  - NIH Big Data to Knowledge Mentored Training Grant K01-ES026839-01A1 (SVG)
  - US National Science Foundation (NSF) grant CCF-1217880 (KRM,AOH)
  - NSF Graduate Research Fellowship, grant F031543 (KRM)