SAND2008-4350C

Modeling of Epileptic Seizures using Tensor Analysis

Evrim Acar
Sandia National Labs., Livermore, CA

Rasmus Bro, Copenhagen University, Denmark
Bulent Yener, Rensselaer Polytechnic Institute, Troy, NY
Canan A. Bingol, Yeditepe University, Turkey
Haluk Bingol, Bogazici University, Turkey
Anthony Ritaccio, Albany Medical College, Albany, NY

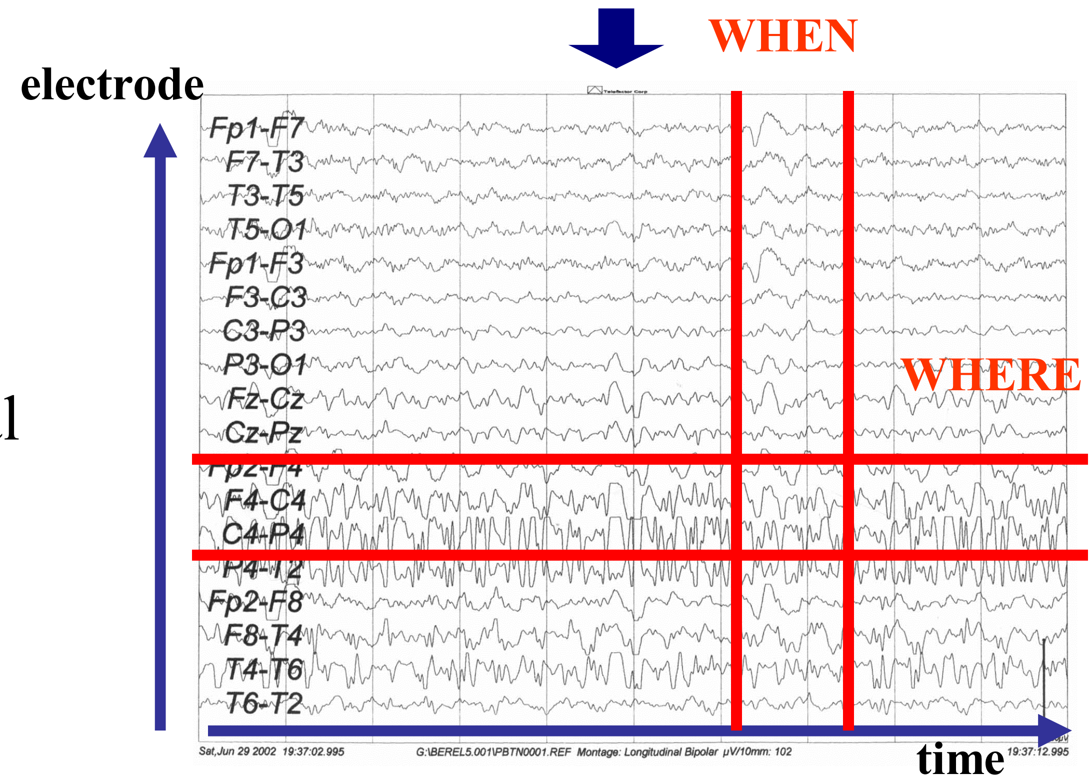
Sandia National Laboratories

Problem Definition

- **Epilepsy:** neurological disorder characterized by recurrent, abrupt seizures.

- WHO: over 50 million people!

- Diagnosis techniques
  – Physical and neurological exams
  – Neuroimaging
  – Video-EEG monitoring



Problem Definition (Cont.)

Motivation:

– Visual analysis is time-consuming

>> saving manpower

– Visual analysis is subjective and error-prone

>> objective/robust analysis

Goal: Developing mathematical models that can capture a seizure structure automatically.

– Seizure recognition **(when)**

– Seizure localization **(where)**

Overview

- **Part I: Tensor Basics**
- **Part II: Epileptic Seizure Recognition**
  - Related Work (e.g., feature extraction from epileptic EEG)
  - Methodology:
    - Construction of a third-order **Epilepsy Feature Tensor**
    - Seizure recognition using Multilinear PLS
- **Part III: Discussions**
  - Future Research Directions
  - Other studies on epileptic EEG signals

Sandia National Laboratories

Overview

- **Part I: Tensor Basics**
- **Part II: Epileptic Seizure Recognition**
  - Related Work (e.g., feature extraction from epileptic EEG)
  - Methodology:
    - Construction of a third-order **Epilepsy Feature Tensor**
    - Seizure recognition using Multilinear PLS
- **Part III: Discussions**
  - Future Research Directions
  - Other studies on epileptic EEG signals

Why tensors?

Two-way data	Higher-order data

- **Social networks:**

  <user, keyword> <user, keyword, time>

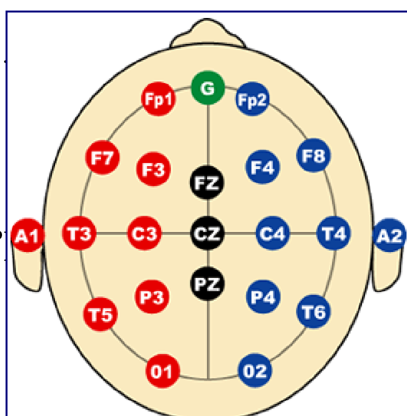
- **Text mining:**

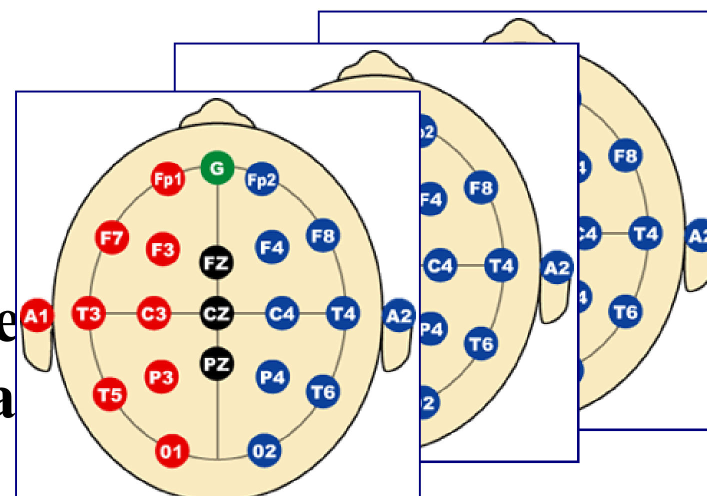
  <document, term> < document, term, author>

- **Face recognition:**

  <person, pixel> < person, pixel, viewpoint>

- **Neuroscience:**





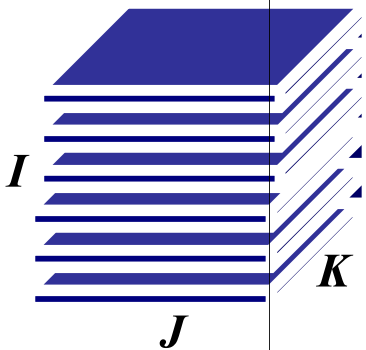
Matr ugh to represe t

he data from a

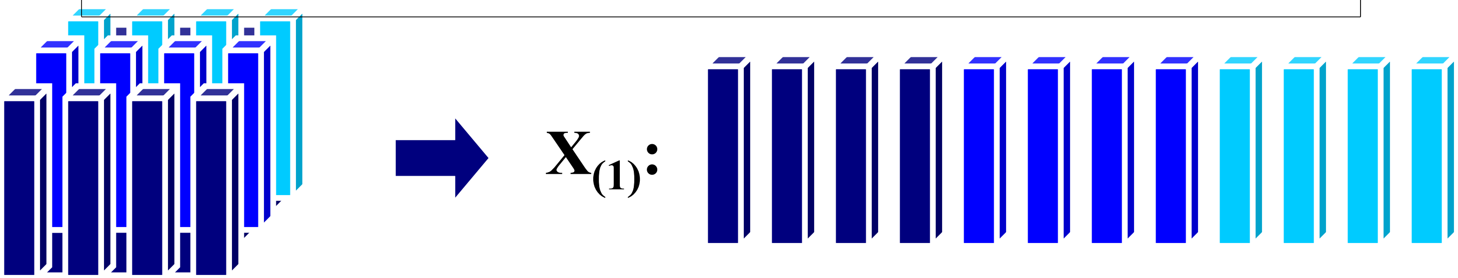
What is a tensor?

- A tensor is a higher-order generalization of a matrix.

Example: A third-order tensor $\underline{X} \in R^{I \times J \times K}$



- Matricization (unfolding/flattening)
  - rearranging a tensor as a matrix.
  - The mode-n matricization of a higher-order dataset, e.g., $\underline{X}$, denoted by $X_{(n)}$ unfolds the data in the n^{th} mode.



Tensor times a matrix

- The mode-n product of a tensor $\underline{X} \in \mathbb{R}^{I_1 \times I_2 \ldots \times I_n \times \ldots I_N}$ with a matrix $M \in \mathbb{R}^{P \times I_n}$ is denoted by

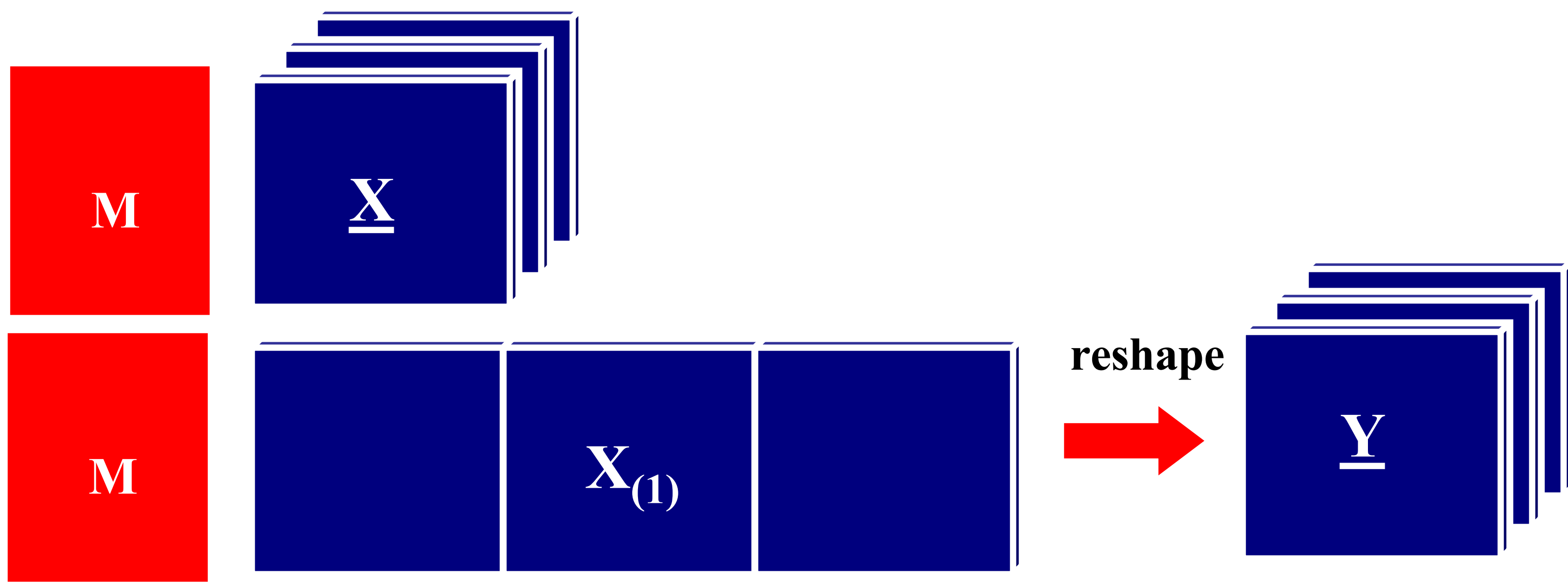
$$\underline{Y} = \underline{X} \times_n M$$

where $\underline{Y} \in \mathbb{R}^{I_1 \times I_2 \ldots I_{n-1} \times P \times I_{n+1} \ldots I_N}$

Example:

$\underline{X} \times_1 M =$

$=$



Overview

- Part I: Tensor Basics
- **Part II: Epileptic Seizure Recognition**
  - Related Work (e.g., feature extraction from epileptic EEG)
  - Methodology:
    - Construction of a third-order **Epilepsy Feature Tensor**
    - Seizure recognition using Multilinear PLS
- Part III: Discussions
  - Future Research Directions
  - Other studies on epileptic EEG signals

Sandia National Laboratories

Related Work

Common Approach: Divide an EEG recording into time epochs and extract features that can distinguish between seizures and other brain dynamics.

Analysis of the performance of multiple features on **single channel** EEG data

Time Epochs
i=1,2,…I

Features
j=1,2,…J

Analysis of the performance of a **single feature** on multi-channel EEG data

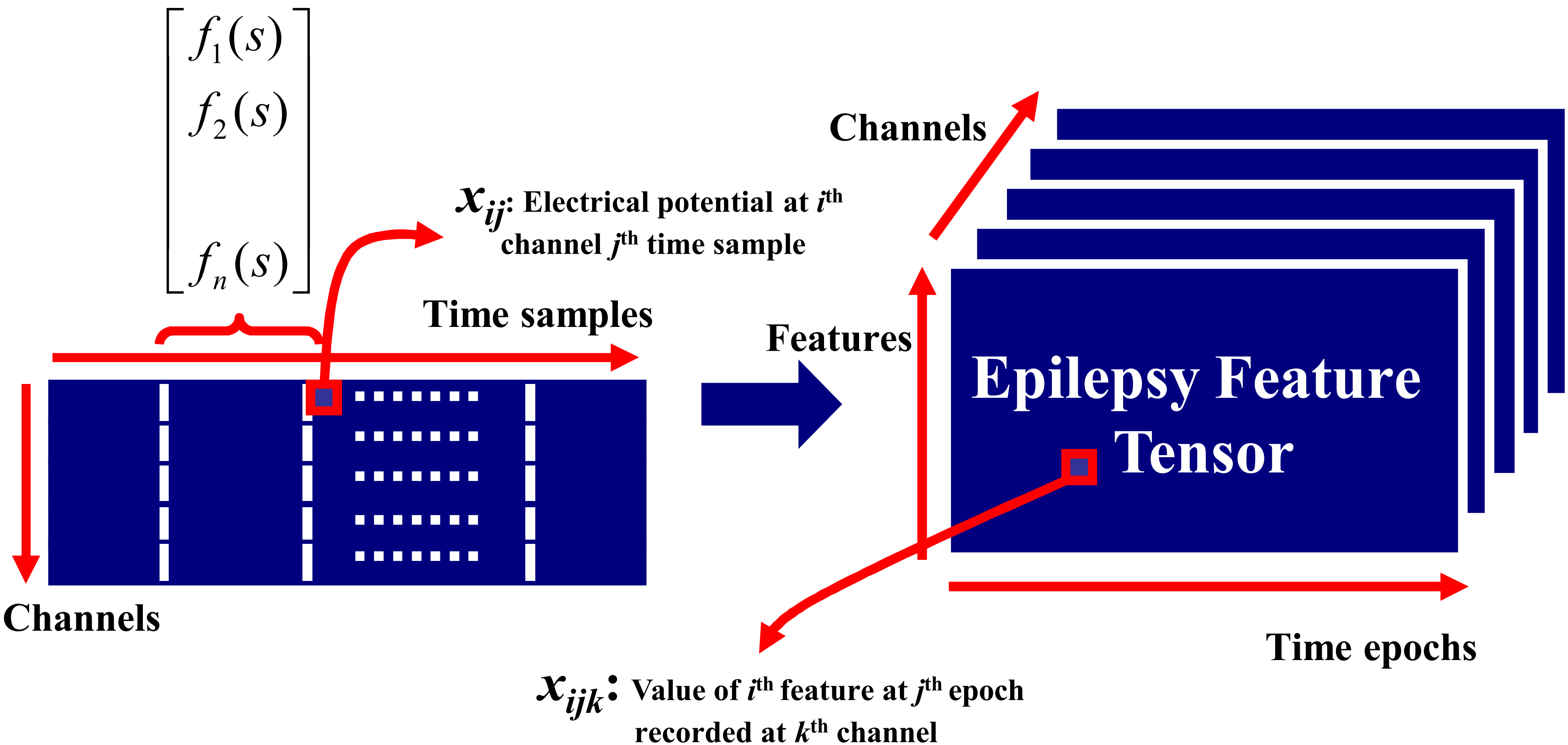
Time Epochs
i=1,2,…I

Channels
k=1,2,…K

Our Approach: Analysis of **multiple features** from different domains on **multi-channel** EEG data

STEP1: Epilepsy Feature Tensor

- Construction of an Epilepsy Feature Tensor from multi-channel EEG



Sandia National Laboratories

Features from Multiple Domains

Time domain

Let s ={s(1), s(2),…s(N)} be the time sequence for a particular epoch of length N.

- Activity: $f_1(s) = \sigma_s^2$
- Mobility: $f_2(s) = \frac{\sigma_{s'}}{\sigma_s}$
- Complexity

$$f_3(s) = \frac{(\sigma_{s''} / \sigma_{s'})}{(\sigma_{s'} / \sigma_s)}$$

- Mean Absolute Slope: f_4(s)

Frequency domain

Take the first difference and then compute the Fourier coefficients to construct the amplitude spectrum

- Median Frequency: f_5(s)

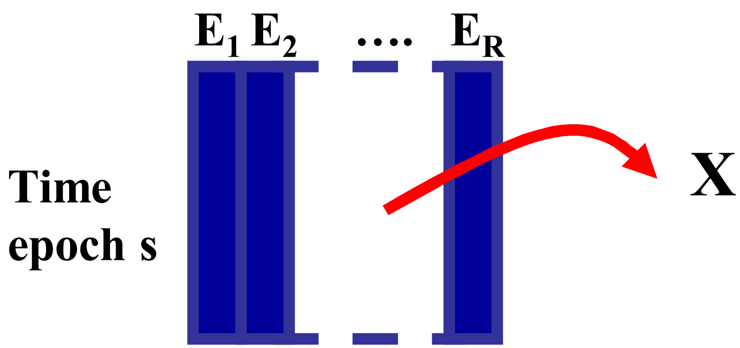
Compute the energy spread across different EEG bands, i.e., δ (0.5-3.5Hz), θ (3.5-7.5Hz), α (7.5-12.5Hz), β(12.5-30Hz) and γ(>30Hz)

- Spectral Entropy: f_6(s)
- Relative Energy in each freq. band:

f_7(s)- f_{10}(s)

More features

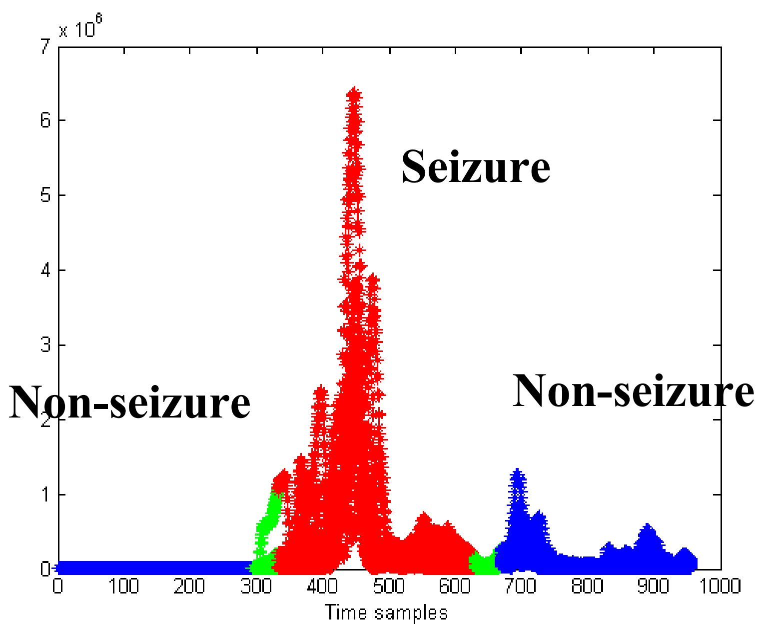
- To quantify the synchronization among channels, especially among neighboring channels as the seizure starts.



$$f_{11}(s,i) = \sum_{j \in NEIGH_i} \left|(X^T X)_{ij}\right|$$

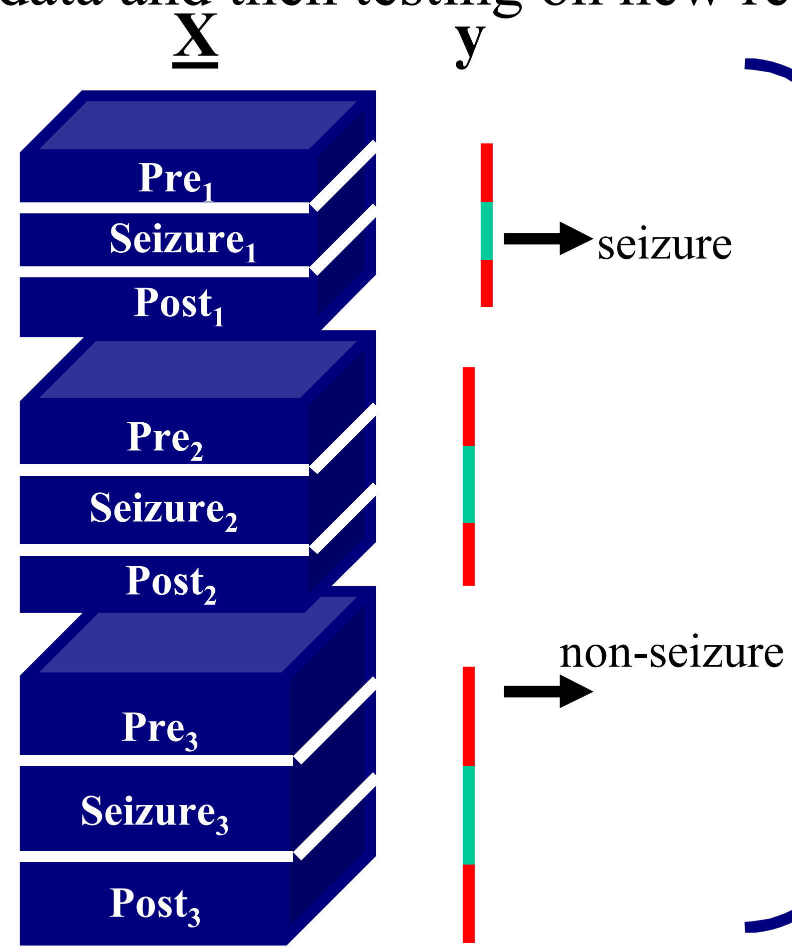
where $NEIGH_i$ contains the neighbors of electrode ***i***.

<u>Example:</u>



STEP2: Seizure Recognition

- Building a mathematical model based on Multilinear PLS on available data and then testing on new recordings

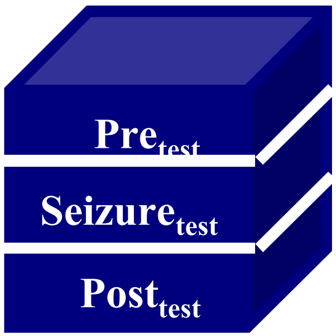


Training Set

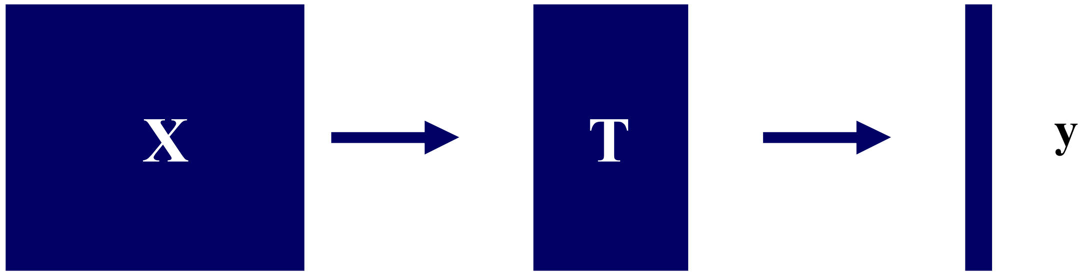
- Build a model based on N-PLS using the training set **X** and the labels **y**.

Test Set

- Predict the labels of new recordings.



Partial Least Squares (PLS)



Let $\mathrm{X} \in \mathrm{R}^{I \times J}$ and $\mathrm{y} \in \mathrm{R}^{I}$ be the independent and dependent (response) variables, respectively

- Map X to a low-dimensional space and regress onto y

$$\mathrm{y} = \mathrm{Tb} + \mathrm{e}$$

- Determine the columns of $\mathrm{T} \in \mathrm{R}^{I \times K}$ (where $\mathrm{K} < \mathrm{J}$), such that

$$\max_{\mathrm{w}} \quad \mathrm{cov}(\mathrm{Xw}, \mathrm{y})$$

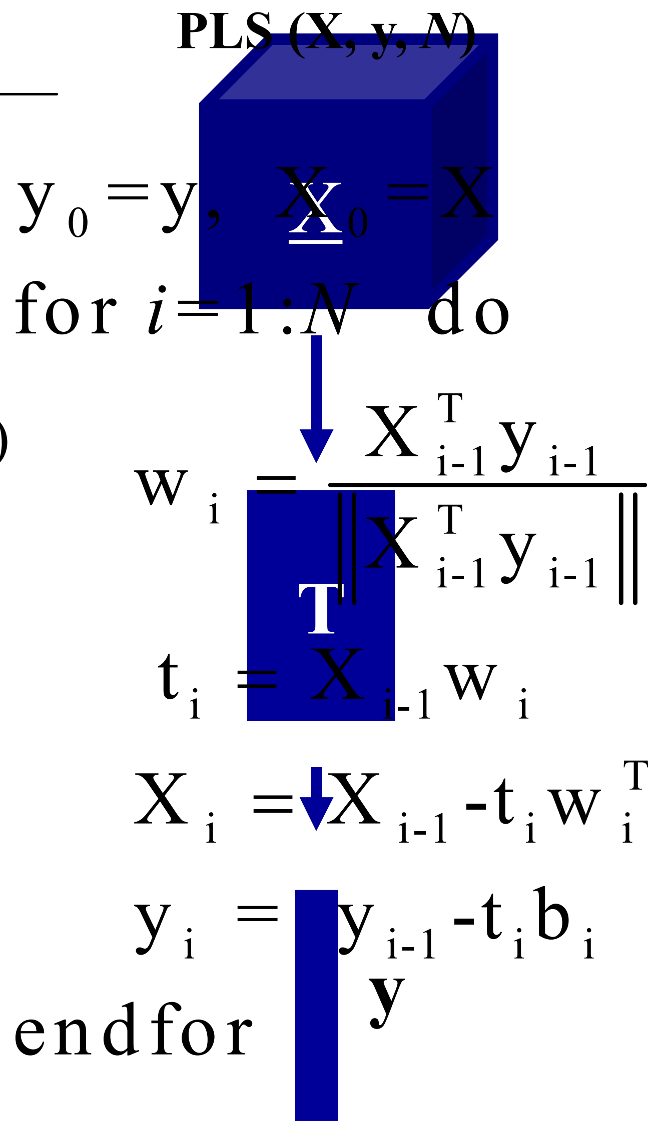
where **t**= **Xw** is a column of matrix **T**.

Multilinear PLS (N-PLS) (by Bro'96)

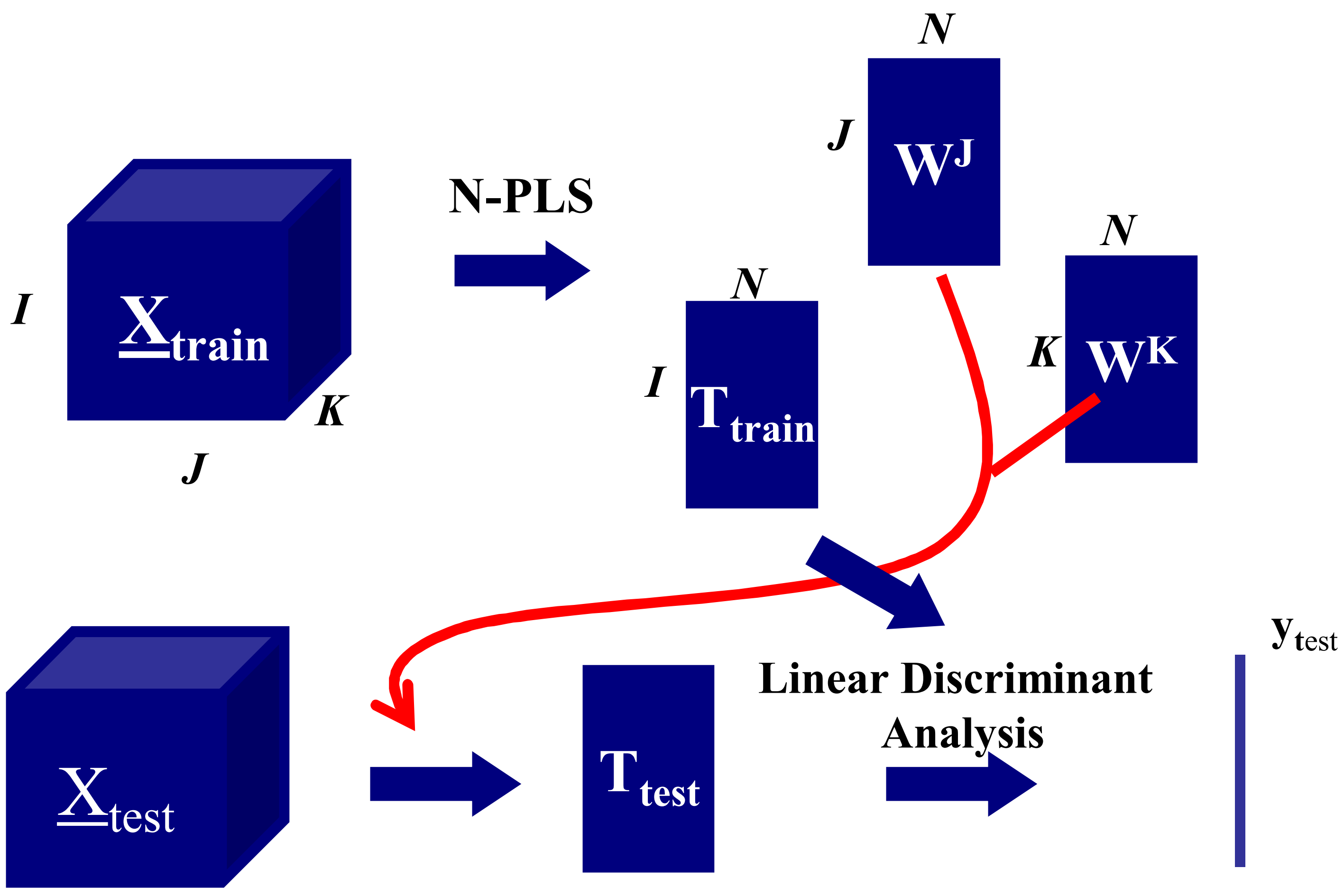
Multilinear PLS $(\underline{X}, y, N)$

1. $y_0 = y, X_0 = X_{(1)}$
2. for $i = 1$ to N do
3. $\quad z = y_{i-1}^T X_{i-1}$
4. $\quad Z = USV^T$, where $Z(m,n) = z(m + J(n-1))$
5. $\quad w^J = U(:,1), \quad w^K = V(:,1)$

   $\quad W^J(:,i) = w^J, W^K(:,i) = w^K$
6. $\quad T(:,i) = X_{i-1}(w^K \otimes w^J)$
7. $\quad X_i = X_{i-1} - T(:,i)(w^K \otimes w^J)^T$
8. $\quad b_i = (T^T T)^{-1} T^T y_{i-1} = T^+ y_{i-1}$
9. $\quad y_i = y_{i-1} - Tb_i$
10. endfor



To predict y_{test}: Combine N-PLS and LDA



Feature Selection: Generalization of VIP to 3-way

- **Goal:** Eliminate the features that are not very relevant with the classification of non-seizure and seizure epochs.
- **Approach:**

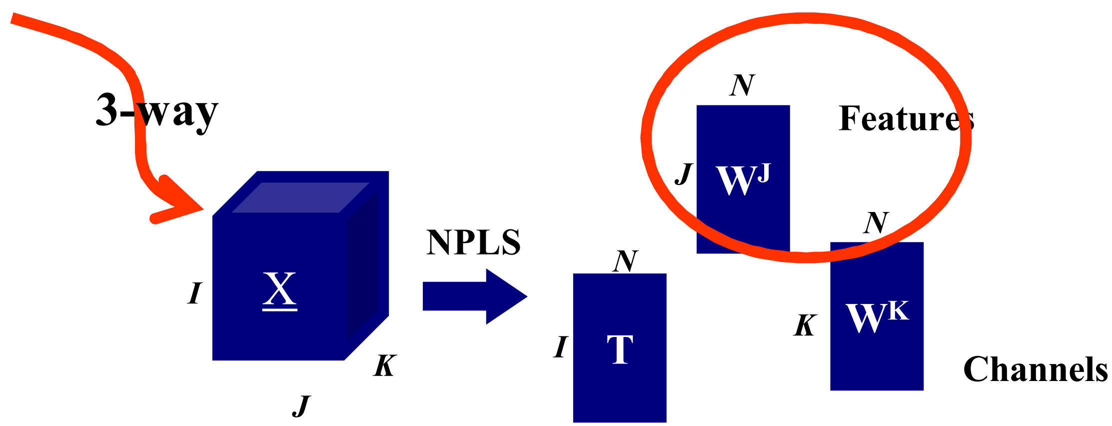
  Use the loadings to determine which variables are important (Variable Importance in Projection (VIP) (Word et *al.*'93))

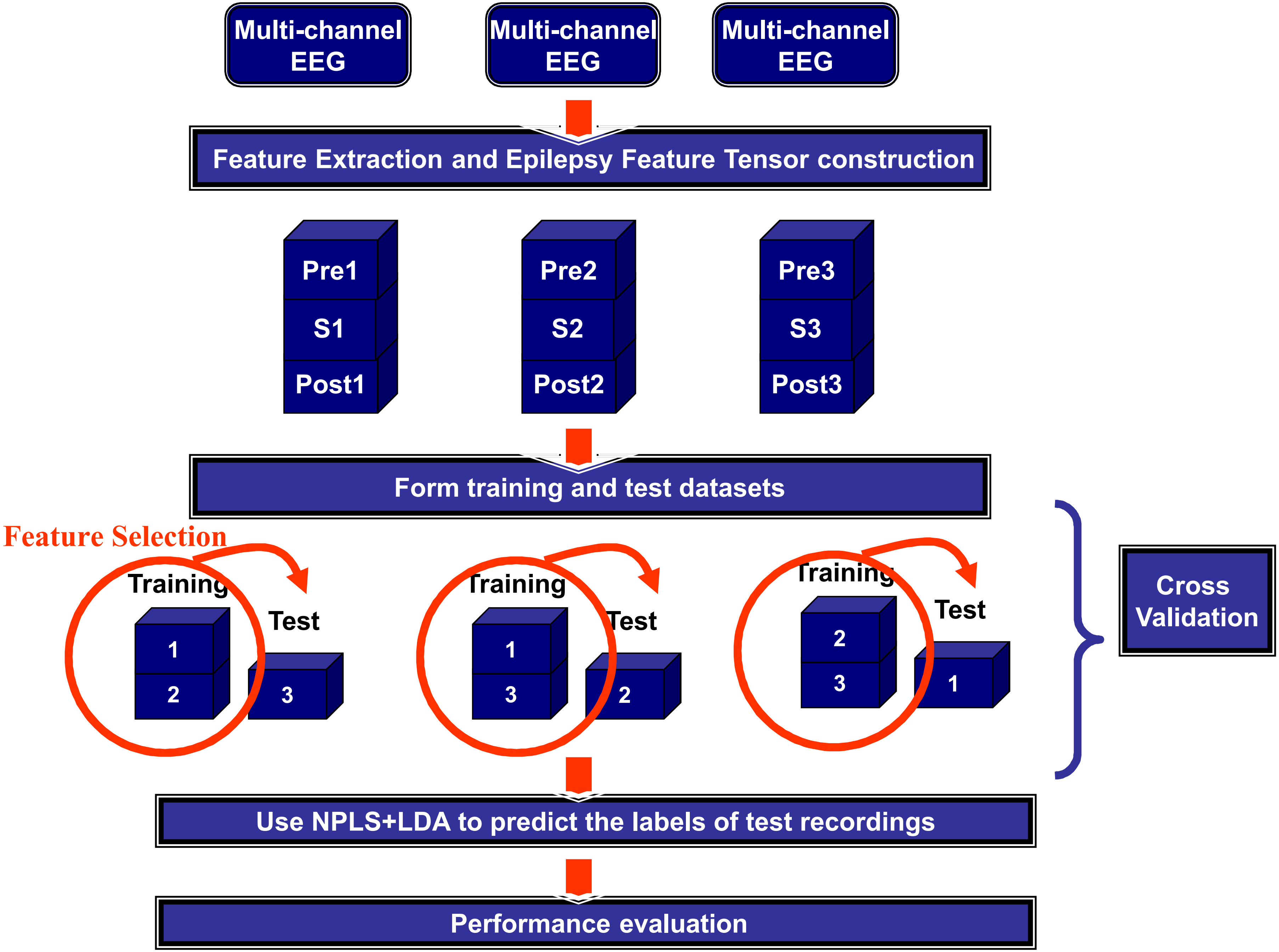
$$\mathbf{VIP}_j = \sqrt{J \times \frac{\sum_{k=1}^{N} b_k^2 t_k^T t_k \left(\frac{w_{jk}}{\|w_k\|}\right)^2}{\sum_{k=1}^{N} b_k^2 t_k^T t_k}}$$

2-way

$$\mathbf{X} = \mathbf{T}\mathbf{W}^{\mathrm{T}} + \mathbf{E}$$

$$\mathbf{y} = \mathbf{T}\mathbf{b} + \mathbf{e}$$




Multi-channel EEG
Multi-channel EEG
Multi-channel EEG
Feature Extraction and Epilepsy Feature Tensor construction
Pre1
S1
Post1
Pre2
S2
Post2
Pre3
S3
Post3
Form training and test datasets
Feature Selection
Training
1
2
Test
3
Training
1
3
Test
2
Training
2
3
Test
1
Cross Validation
Use NPLS+LDA to predict the labels of test recordings
Performance evaluation

DATA (Patient-Specific Seizure Recognition)

- Collected in the epilepsy monitoring unit of
  - Yeditepe University Hospital (Istanbul, Turkey)
  - Albany Medical College (NY, USA)
- 32 seizures from 9 patients with different seizure origins (at least 3 seizures/patient)
- Seven Features: Activity, Complexity, Mobility, Mean Absolute Slope, Spatial Information, Spectral Entropy, Median Frequency.
- Preprocessing:
  - Filter at 50Hz/60Hz
  - Scaling within features mode

Patient Id	Seizure Id	Size of an Epilepsy Feature Tensor
1	1	302 x 7 x 18
	2	386 x 7 x 18
	3	320 x 7 x 18
	4	398 x 7 x 18
	5	444 x 7 x 18
2	1	878 x 7 x 18
	2	866 x 7 x 18
	3	902 x 7 x 18
	4	986 x 7 x 18
	5	998 x 7 x 18
3	1	790 x 7 x 18
	2	746 x 7 x 18
	3	1034 x 7 x 18
4	1	1174 x 7 x 18
	2	1346 x 7 x 18
	3	1170 x 7 x 18

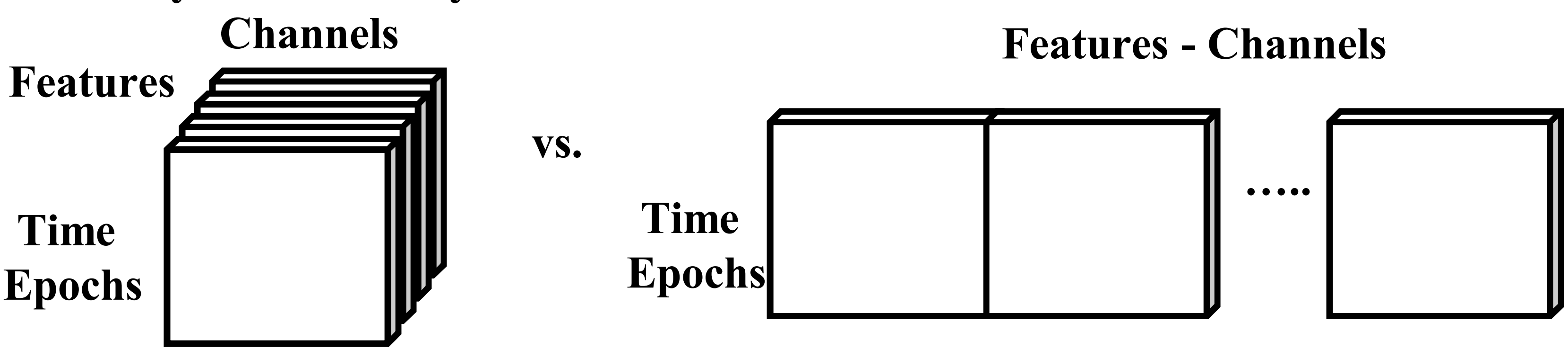
⋮

Performance Evaluation

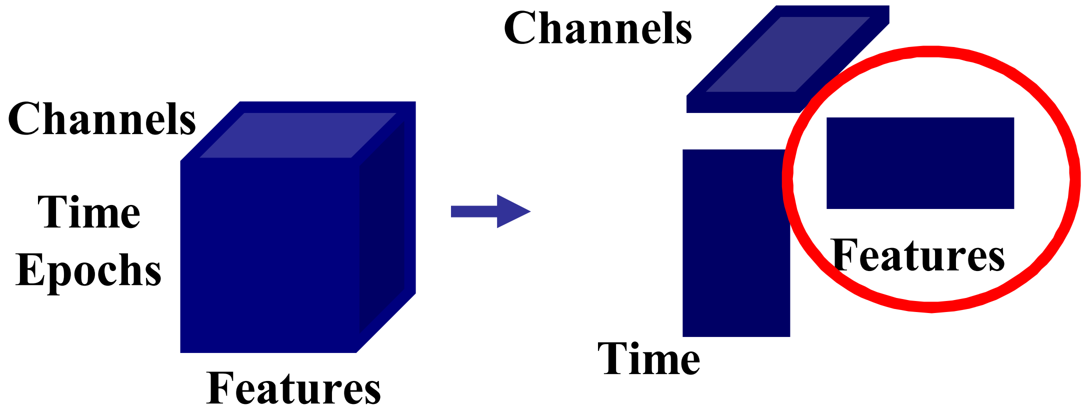
- **G-means:** $\sqrt{Sensitivity \times Specificity}$

	NPLS+LDA	NPLS+LDA (FS)
Patient1	85.3%	86.6%
Patient2	97.6%	96.7%
Patient3	91.3%	91.1%
Patient4	75.0%	**77.3%**
Patient5	28.6%	**83.1%**
Patient6	72.3%	**89.3%**
Patient7	97.0%	92.1%
Patient8	86.0%	78.4%
Patient9	84.5%	77.4%

- **Three-way vs. Two-way:**



Features

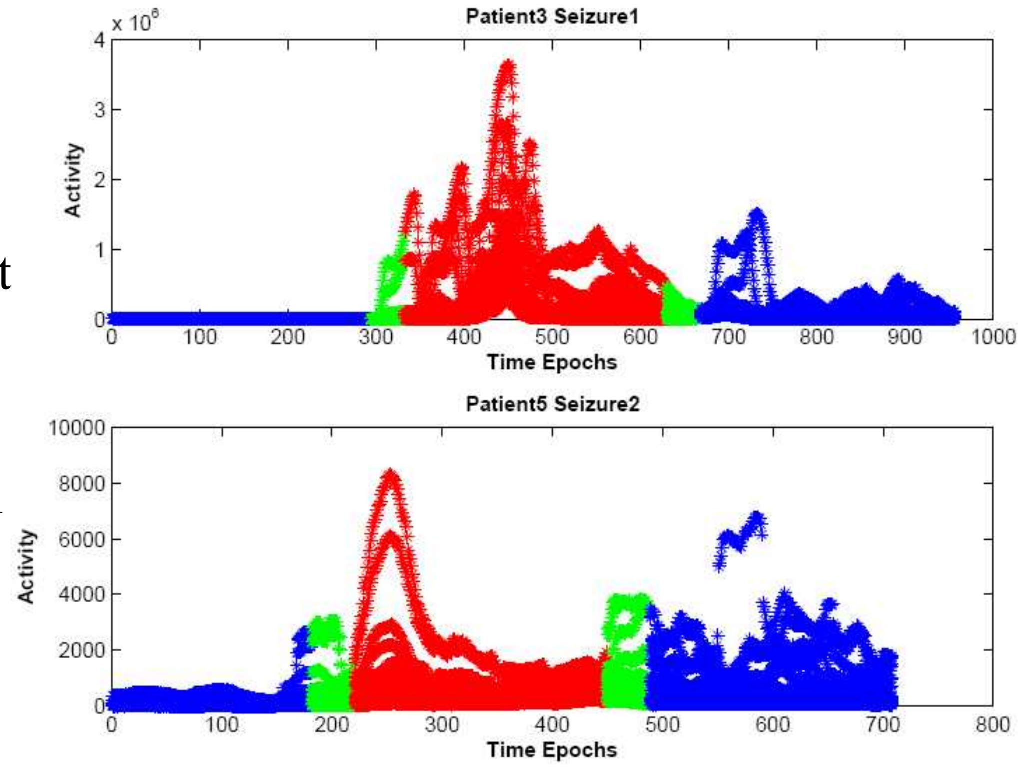


PID	Activity	Mobility	Complexity	Mean Abs. Slope	Spatial Info	Median Freq.	Spectral Entropy
1	✓	✓	✓	✓	✓	×	✓
2	✓	✓	✓	✓	✓	×	✓
3	✓	✓	✓	×	✓	×	✓
4	✓	✓	✓	×	✓	×	✓
5	✓	✓	✓	×	×	×	✓
6	✓	✓	✓	×	✓	×	×
7	✓	✓	✓	✓	✓	×	✓
8	✓	✓	✓	✓	✓	×	×
9	✓	✓	×	×	✓	×	×

DATA (Patient Non-Specific Seizure Recognition)

- Collected in the epilepsy monitoring unit of
  - Yeditepe University Hospital (Istanbul, Turkey)
  - Albany Medical College (NY, USA)
- 26 seizures from 9 patients (left or right temporal)
- Eleven Features: Activity, Complexity, Mobility, Mean Absolute Slope, Spatial Information, Spectral Entropy, Median Frequency, **Relative Energy**.
- Preprocessing:
  - Filter at 50Hz/60Hz
  - Scaling within features mode
  - **Log transform**

Patient Seizure Size of an Epilepsy



⋮

Performance Evaluation

- **G-means:** $\sqrt{Sensitivity \times Specificity}$

Patient ID	NPLS+LDA (v1)	NPLS+LDA (v2)	NPLS+LDA (v1+heuristic)
1	87.7%	81.4%	89.0%
2	88.8%	87.0%	90.2%
3	83.7%	88.1%	85.2%
4	41.9%	84.7%	39.8%
5	95.6%	94.9%	96.9%
6	95.6%	92.6%	95.6%
7	94.5%	88.1%	95.6%
8	69.6%	71.1%	69.3%
9	68.1%	72.8%	68.6%
MEAN	80.6%	84.5%	81.1%

Sandia National Laboratories

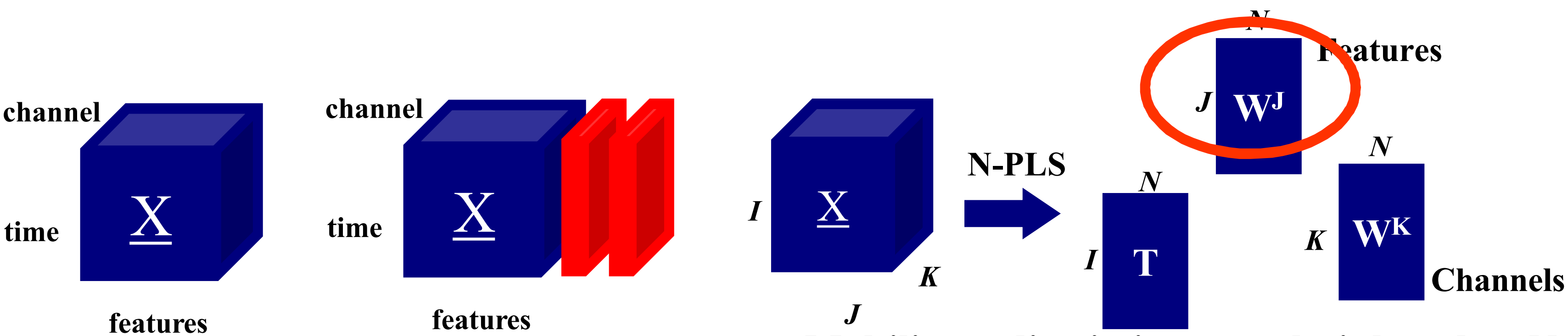
Results

- Average performance of patient-specific seizure recognition over 9 patients (32 seizures) is **85.8%.**
  - better performance using feature selection
  - performance comparable with SVMs and easier interpretation
  - feature selection may improve the understanding of epileptic seizures

- Average performance of patient non-specific seizure recognition over 9 patients (26 seizures) is **80.6%**.
  - heuristics to handle false-positives may improve the performance.

Summary

- **Epileptic Seizure Recognition:**



Epilepsy feature tensor

More features e.g., spatial information

- **Multilinear discriminant analysis based on N-PLS and LDA**
- **VIP extended to three-way**
- **Comparison with SVMs**
- **Patient specific and non-specific seizure recognition**

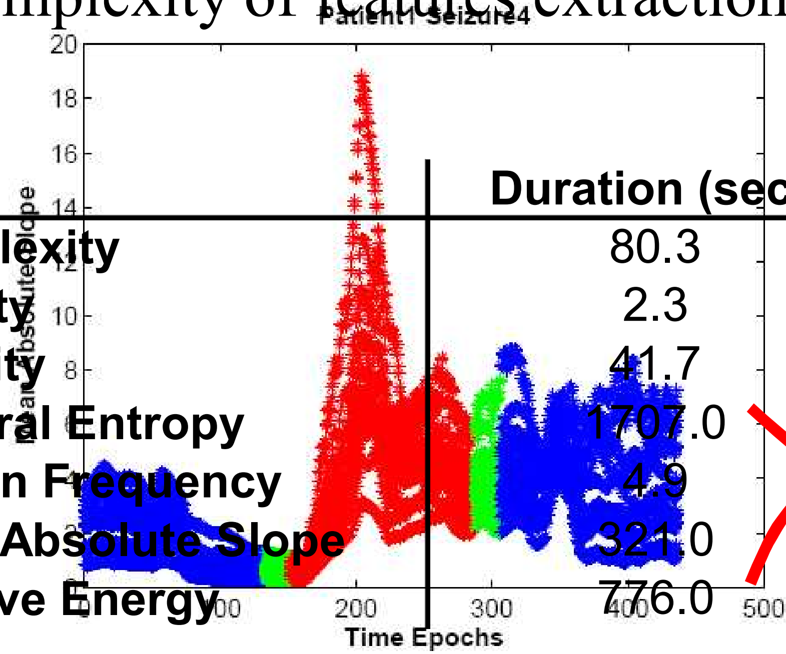
Overview

- **Part I: Tensor Basics**
- **Part II: Epileptic Seizure Recognition**
  - Related Work (e.g., feature extraction from epileptic EEG)
  - Methodology:
    - Construction of a third-order **Epilepsy Feature Tensor**
    - Seizure recognition using Multilinear PLS
- **Part III: Discussions**
  - Future Research Directions
  - Other studies on epileptic EEG signals

Epilepsy Feature Tensor Construction

- Features can distinguish between pre-seizure and post-seizure.
- Computational complexity of features extraction

Example:



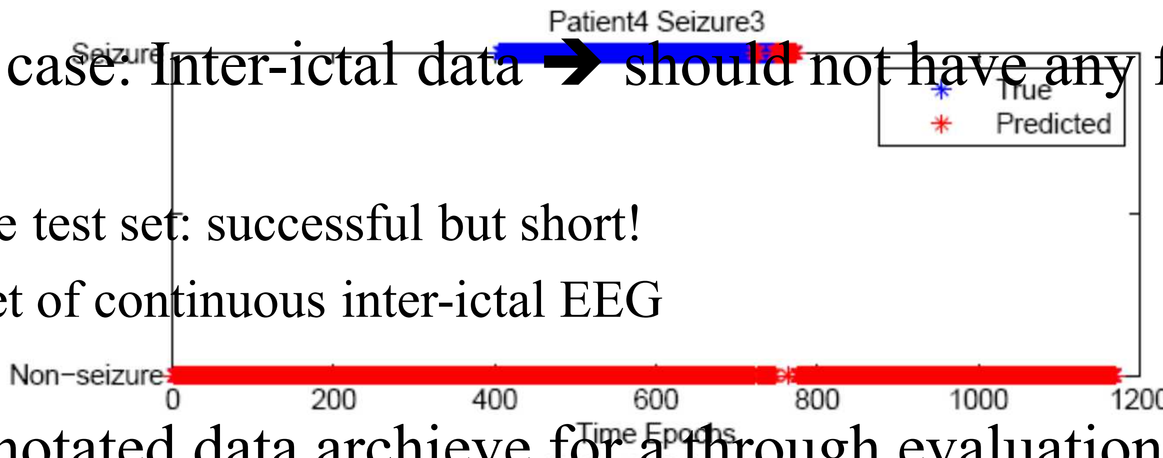
	Duration (sec.)
Complexity	80.3
Activity	2.3
Mobility	41.7
Spectral Entropy	1707.0
Median Frequency	4.9
Mean Absolute Slope	321.0
Relative Energy	776.0

O(nlogn)/scale

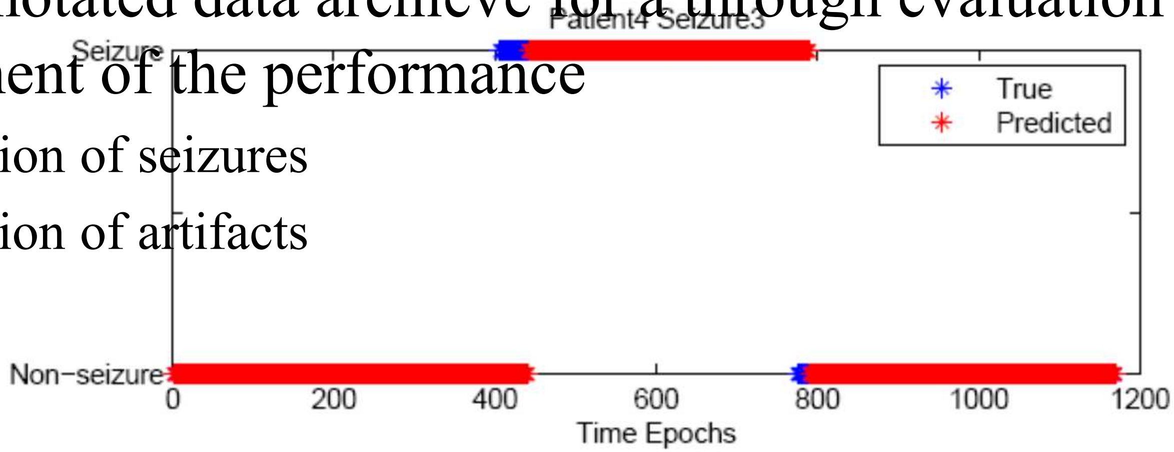
Sandia National Laboratories

Seizure Recognition

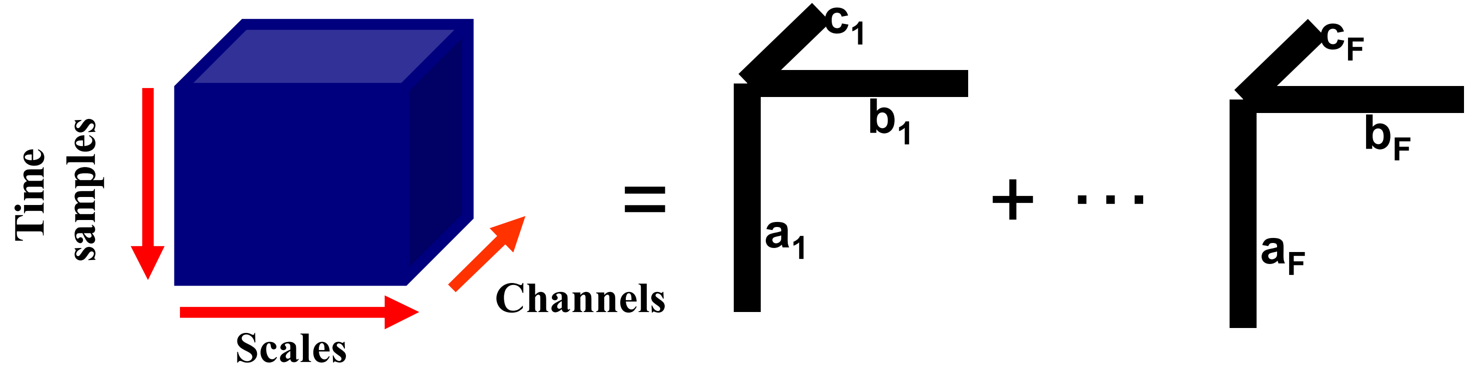
- Log transform to handle inter-patient variation not enough!
- Good test case: Inter-ictal data ➔ should not have any false positives!
  - 4-minute test set: successful but short!
  - Large set of continuous inter-ictal EEG



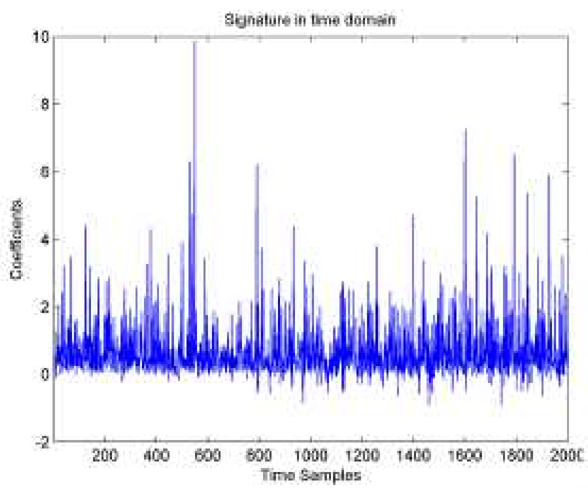
- Larger annotated data archieve for a through evaluation and improvement of the performance
  - Annotation of seizures
  - Annotation of artifacts

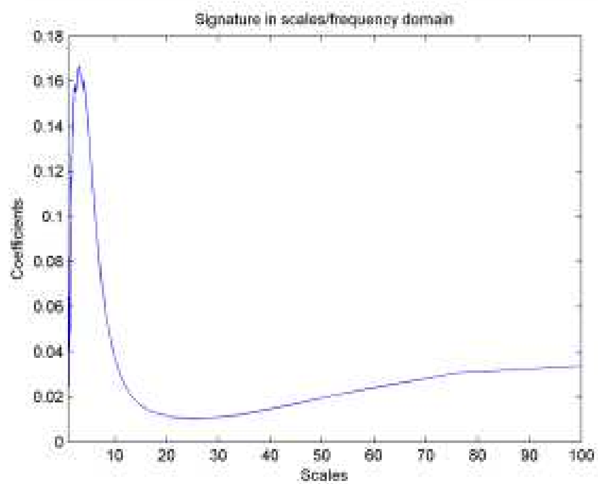


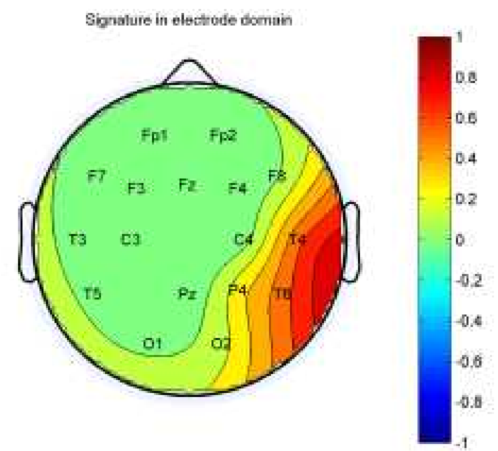
Epilepsy Focus Localization



$$\underline{X} = a_1 \circ b_1 \circ c_1 + a_2 \circ b_2 \circ c_2 + \underline{E}$$







Evrim Acar, Canan A. Bingol, Haluk Bingol, Rasmus Bro and Bulent Yener, **Multiway Analysis of Epilepsy Tensors,** *Bioinformatics*, 2007, 23(13): i10-i18.

Thank you!

Evrim Acar

eacarat@sandia.gov

MATLAB scripts: www.cs.rpi.edu/~acare/Epilepsy