Signal Processing for Diagnosis and Treatment of Brain Disorders

Justin Dauwels

www.dauwels.com

Laboratory for Information and Decision Systems (LIDS)
Massachusetts Institute of Technology (MIT)

March 18, 2009
“With all the tools available to modern medicine—the blood tests, M.R.I.’s and endoscopes—you might think that misdiagnosis has become a rare thing. But you would be wrong.”

Proposed: Signal processing aided decision making
Proposed: Signal processing aided decision making

- Signal processing “microscope”
  discovery of **features/patterns** in data associated with certain **disorders**

- Algorithmic decision making
  combine **evidence** “optimally” → **diagnosis** and **treatment**

- More **reliable** and **earlier** diagnosis
- More **effective** treatment (clinical outcome, time and cost)
Synchrony of brain signals is an important feature

- **Loss** of synchrony of brain signals $\rightarrow$ **brain disorder**
- **Normal** synchrony of brain signals
- **Increased** synchrony of brain signals $\rightarrow$ **brain disorder**

Often **subtle** effects, requires **signal processing**
Signal Processing Aided Diagnosis

Loss in EEG synchrony $\rightarrow$ diagnosis of early-Alzheimer’s

Signal Processing Aided Treatment

Increased EEG synchrony $\rightarrow$ localization of epileptic brain tissue

Future Work

How do we plan to improve our methods?

Electroencephalogram (EEG)
$=$ electrical activity along the scalp or brain
Alzheimer’s disease is complex, devastating, and common

- **Symptoms**: memory loss, language breakdown, loss of motor control, apathy, …
- **Causes**: not well understood, complex molecular mechanisms
- **2-5%** of people over **65** years old
Our goal: early diagnosis of AD using EEG

Early diagnosis of AD is important

- medication and other therapies most useful when used early

More precisely: diagnosis of MCI and mild AD

- predementia (a.k.a. mild cognitive impairment, MCI)
- mild–moderate–severe AD
Our goal: early diagnosis of AD using EEG

Standard approach

mental testing, MR, CT, SPECT, PET, corticospinal fluid

Scalp EEG is attractive complementary technology

- simple, affordable, mobile
- useful for screening

scalp EEG → signs of AD? → Yes → MR, CT, …
EEG signals of AD patients are less coherent

- **Loss** of neurons → weaker brain **connectivity**

  Healthy subject  
  ![Healthy brain image]

  AD patient  
  ![AD brain image]

- EEG from different brain areas is **less correlated** in AD patients

- Loss in EEG synchrony → loss in brain connectivity → AD
- How to **detect** loss in EEG synchrony?
How to measure EEG synchrony?

Classical techniques

- correlation coefficient, coherence
- phase synchrony
- Granger causality
- information-theoretic measures
- ...

*Few* of them applied on *single* data sets of early-AD patients
How to measure EEG synchrony?

Classical techniques

- correlation coefficient, coherence
- phase synchrony
- Granger causality
- information-theoretic measures
- ...

Few of them applied on single data sets of early-AD patients

Systematic study

30+ measures on two data sets (and more to follow)
How to measure EEG synchrony?

New approach: “Stochastic Event Synchrony (SES)”

- Extracts “events” from EEG signals and aligns those events
- Provides complementary information about synchrony
- Promising results for diagnosing AD
- General formalism for similarity of point processes:
  - pairs of point processes and $n > 2$ point processes
  - one-dimensional and multi-dimensional point processes
  - time-varying similarity

In this talk: pairs of two-dimensional point processes (EEG)
Local synchronous activity leads to “events” in EEG

Non-synchronized activity

Synchronized activity

Neuroscience: Exploring the Brain, M. Bear et al.
Events can clearly be observed in time-frequency map.

- **Time-frequency transform**: $|X(t,f)|^2$
- **Wavelet transform**: $X(t,f)$
- **Time-frequency events**: "bumps"
Oscillatory events extracted from time-frequency map

- Bump models are **point processes** on time-frequency plane
- **5 bump parameters:**
  - timing, frequency, width in time and frequency, amplitude
- **Lossy compression:** “background” activity neglected

F. Vialatte et al., Neural Networks 2007
We determine EEG synchrony from bump models

1. **Group** electrodes in regions
2. **Extract** bump model for each region
3. **Determine** similarity of each pair of bump models
4. **Average** similarity over all pairs

Misalignment of bump models $\rightarrow$ loss of brain connectivity $\rightarrow$ AD
Stochastic Event Synchrony

Bump models of two EEG signals

Coincident bumps

1. How many bumps coincide?
2. How well do the coincident bumps align?
Stochastic Event Synchrony \((\rho, \delta_t, \delta_f, s_t, s_f)\)

**Bump models of two EEG signals**

**Coincident bumps**

1. **How many bumps coincide?**
   - fraction of non-coincident bumps \(\rho\)

2. **How well do the coincident bumps align?**
   - average \(t\) and \(f\) offset of coincident bumps \((\delta_t\) and \(\delta_f)\)
   - \(t\) and \(f\) jitter of coincident bumps \((s_t\) and \(s_f)\)
Stochastic Event Synchrony \((\rho, \delta_t, \delta_f, s_t, s_f)\)

1. How many bumps coincide?
   - fraction of non-coincident bumps \(\rho\)

2. How well do the coincident bumps align?
   - average \(t\) and \(f\) offset of coincident bumps \((\delta_t\) and \(\delta_f)\)
   - \(t\) and \(f\) jitter of coincident bumps \((s_t\) and \(s_f)\)

It is natural to iterate Step 1 and 2.
How can we derive such iterative procedure?

Wish list

- Should give the “best” alignment
- Should converge
- Should be easily extendable
SES parameters computed by statistical inference

<table>
<thead>
<tr>
<th>Maximum a posteriori estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(c^<em>, \theta^</em>) = \arg\max_{c, \theta} p(e, e', c; \theta)$</td>
</tr>
</tbody>
</table>

- $e$ and $e'$ are the given bump models
- $c_{kk'} = 1$ if $(e_k, e'_k)$ coincident; $c_{kk'} = 0$ otherwise
- $\theta = (\delta_t, \delta_f, s_t, s_f)$, $\rho$ follows from $c$

<table>
<thead>
<tr>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the model $p$?</td>
</tr>
<tr>
<td>How do carry out MAP estimation practically?</td>
</tr>
</tbody>
</table>
Statistical model $p$ relates two bump models

Two given bump models are generated from “mother” model

1. “Mother” bumps uniformly at random within rectangle
Statistical model $p$ relates two bump models

Two given bump models are generated from “mother” model

1. “Mother” bumps uniformly at random within rectangle
2. Generate two copies of each “mother” bump
3. Perturb bump positions randomly: Gaussian RVs with mean $(\pm \delta_t/2, \pm \delta_f/2)$ and var $(s_t/2, s_f/2)$
Statistical model $p$ relates two bump models

Two given bump models are generated from “mother” model

1. “Mother” bumps uniformly at random within rectangle
2. Generate two copies of each “mother” bump
3. Perturb bump positions randomly: Gaussian RVs with mean $(\pm \delta_t / 2, \pm \delta_f / 2)$ and var $(s_t / 2, s_f / 2)$
4. Delete bumps at random $\rightarrow$ non-coincident bumps ($\rho$)
We alternate bump alignment and parameter estimation

**Problem**

Given two bump models $e$ and $e'$, infer $\rho$ and $\theta = (\delta_t, \delta_f, s_t, s_f)$

**Approach: Maximum a posteriori estimation**

$$(c^*, \theta^*) = \arg\max_{c, \theta} p(e, e', c; \theta)$$

**Iterative algorithm: coordinate descent**

$$\hat{c}^{(k+1)} = \arg\max_c p(e, e', c; \hat{\theta}^{(k)})$$  \hspace{1cm} \text{Bump Alignment}

$$\hat{\theta}^{(k+1)} = \arg\max_{\theta} p(e, e', \hat{c}^{(k+1)}; \theta)$$  \hspace{1cm} \text{Parameter Estimation}

- Alignment is equivalent to max-weight bipartite matching
- Can be solved efficiently (belongs to P)
- Overall algorithm is guaranteed to converge
- Easily extendable
We applied SES to EEG data

**Setup**

- **spontaneous EEG**, in rest with eyes closed
- **21 electrodes** (10-20 international system)
- electrodes grouped in **5 zones**
- **bandpass** filtered between 4 and 30Hz
- **20s of artifact-free EEG**

**2 data sets, different patients/hospitals**

- **22 MCI** and **38 Control** subjects
- **17 Mild AD** and **24 Control** subjects
Promising and consistent numerical results

- Applied 30+ synchrony measures, including SES
- Most indicate EEG synchrony loss, but weakly significant
- Only full-frequency direct transfer function (ffDTF) and $\rho$ strongly significant ($p < 0.001$), for BOTH data sets
- MCI vs. Control harder than Mild AD vs. Control
- EEG synchrony loss seems to be strong indicator of early AD
Promising and consistent numerical results

Observations about SES

- SES provides **complementary** information about synchrony
- SES is based on **oscillatory events** in EEG (local synchrony)
- SES **improves** diagnosis of AD (for those two data sets)
Want to know more? Want to try it out?

http://www.dauwels.com/SESToolbox/SES.html
## Signal Processing Aided Diagnosis

### Loss in EEG synchrony $\rightarrow$ diagnosis of early-Alzheimer’s

We needed to develop novel measures of synchrony.

## Signal Processing Aided Treatment

### Increased EEG synchrony $\rightarrow$ localization of epileptic brain tissue

Existing synchrony measures proved to be effective.

## Future Work

How do we plan to improve our methods?
Epilepsy is complex, devastating, and common

- **Heterogeneous group** of central nervous system disorders characterized by **recurrent unprovoked seizures**
- **Symptoms**: disturbance of consciousness or awareness, alterations of bodily movement, sensation or posture, . . .
- **Causes**: genetic/metabolic abnormalities, tumors, . . .
- 2.5 million patients in U.S., 300’000 in U.K., 50 million world
If medication alone fails, surgery may be solution

Medication is usual treatment but not always effective

- **Medication** is most common treatment
- However, medication alone not effective for 30% of patients
If medication alone fails, surgery may be solution

Medication is usual treatment but not always effective

- **Medication** is most common treatment
- However, medication alone not effective for **30%** of patients

Resection if medication alone fails

- **Resection** may be alternative for **location-related** epilepsy
- **Objective**: remove seizure onset zone
- **How to determine the seizure onset zone?**
  - from **clinical behavior** (e.g., left arm shaking) or scalp EEG
  - from **intracranial** EEG, if everything else fails
    - risky, costly, uncomfortable
If medication alone fails, surgery may be solution

Medication is usual treatment but not always effective

- **Medication** is most common treatment
- However, medication alone not effective for **30%** of patients

Resection if medication alone fails

- **Resection** may be alternative for **location-related** epilepsy
- **Objective**: remove seizure onset zone
- **How to determine the seizure onset zone?**
  - from **clinical behavior** (e.g., left arm shaking) or **scalp EEG**
  - from **intracranial** EEG, if everything else fails
    - risky, costly, uncomfortable

We wish to make intracranial recordings as **short** as possible
Standard procedure relies mostly on “seizure” EEG.
Our approach: exploit rest EEG to shorten hospitalization

Objective
Seizure onset zone from rest EEG only → shorter hospitalization

Hypothesis
Seizure onset zone is characterized by hypersynchronous activity even while at rest
Differences with diagnosis of AD project

- **Intracanial** EEG vs. **scalp** EEG
- Signal processing to guide surgery vs. diagnosis
- Increased EEG synchrony vs. EEG synchrony loss
Three main questions

**Hypothesis**

Seizure onset zone is characterized by hypersynchronous activity even while at rest

1. Does hypersynchronous activity occur in rest EEG?
2. Do hypersynchronous areas correlate with seizure onset zone?
3. How to determine seizure onset zone using hypersynchrony?
Technical aspects of data analysis

- **1 hour** segment of intracranial **rest** EEG
- **48 hours** between segment and seizures
- **bandpass** filtered between 4-30Hz, **no** further preprocessing
- **4 patients** with **grid electrodes**, 2 with **depth electrodes**
- **multiple** synchrony measures
Technical aspects of data analysis

- **1 hour** segment of intracranial **rest** EEG
- **48 hours** between segment and seizures
- **bandpass** filtered between 4-30Hz, no further preprocessing
- 4 patients with **grid electrodes**, 2 with **depth electrodes**
- **multiple** synchrony measures

Existing literature

- Most existing studies: **scalp** EEG
- **Few** studies: intracranial grid electrodes, **no depth** electrodes
- **Few** synchrony measures
Some brain areas are consistently hypersynchronous

**Strong local correlation in anterior temporal lobe**

Patient 1

![Brain scan with regions marked]

<table>
<thead>
<tr>
<th>correlation coeff $c$</th>
<th>histogram</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Correlation coefficient matrix" /></td>
<td><img src="image" alt="Histogram" /></td>
</tr>
</tbody>
</table>

Variability over time is small (less than 15%)
Some brain areas are consistently hypersynchronous

Strong local correlation in anterior temporal lobe

Patient 1

Correlation coeff \( c \)

Histogram

Very similar results for other synchrony measures

Phase Synchrony

Coherence

Granger causality (DTF)
Hypersynchrony correlates strongly with seizure onset zone

Strong local correlation in anterior temporal lobe

Patient 1

correlation coeff $c$

histogram

seizure onset

Very similar results for other synchrony measures

Phase Synchrony

Coherence

Granger causality (DTF)
Same phenomenon also occurs in other patients

**Patient 2** (left hemisphere)  
- [Image: Brain scan with labeled regions 1, 2, 3, 4]
- Correlation coefficient

**Patient 3** (right hemisphere)  
- [Image: Brain scan with labeled regions 1, 2, 3, 4]
- [Image: Heatmap with highlighted areas]

**Patient 4** (right hemisphere)  
- [Image: Brain scan with labeled regions 1, 2, 3, 4, 5]
- Sync (right - left)
Is this a counterexample?

**Patient 5 (left hemisphere)**

Considered as “counter example” in *earlier* studies
Depth electrodes provide crucial information . . .

Patient 5 (left hemisphere)

Considered as “counter example” in earlier studies
Depth electrode reveals actual seizure onset zone

seizure propagation

seizure onset

hypersynchrony
Three main questions

1. Does hypersynchronous activity occur in rest EEG?
   Yes! (in the 6 patients analyzed so far)

2. Do hypersynchronous areas correlate with seizure onset zone?
   Yes! (in the 6 patients analyzed so far)

3. How to determine seizure onset zone using hypersynchrony?
   Statistical decision making, using a graphical model
   Work in progress . . .

Signal processing of rest EEG has high potential
→ drastically shorter hospitalization
Topics

Signal Processing Aided Diagnosis

Loss in EEG synchrony $\rightarrow$ diagnosis of early-Alzheimer’s

We needed to develop novel measures of synchrony

Signal Processing Aided Treatment

Increased EEG synchrony $\rightarrow$ localization of epileptic brain tissue

Existing synchrony measures proved to be effective

Future Work

How do we plan to improve our methods?
Future work: Early diagnosis of AD

**EEG Data**
- **Analyze** additional data
- **Collect** additional data (MPI/UHospital Frankfurt)

**Signal Processing**
- **Extensions** of SES
- **Alternative approaches**
  - spectral properties
  - source reconstruction
  - non-negative tensor decomposition
  - ...
Future work: Treatment of epilepsy

- Additional features
  - high-frequency ripples (80–300Hz) [Worrell et al., Staba et al.]
  - slowing effect

- Additional data
  - Scalp EEG, MEG, SPECT, PET, CT, MR, DTI
  - Clinical behavior

- Real-time signal processing in Operating Room

Adaptive electrode placement/resection

- Fundamentals: seizure genesis/hypersynchrony/resection
  - biophysical models
  - animal models
Back to the big picture . . .

- **Signal processing “microscope”**
  discovery of **features/patterns** in data associated with certain **disorders**

- **Algorithmic decision making**
  combine **evidence** “optimally” → **diagnosis** and **treatment**

- **More reliable and earlier** diagnosis
- **More effective** treatment (clinical outcome, time and cost)
Acknowledgements

The SES Team

- François Vialatte (RIKEN)
- Theo Weber (MIT)
- Jordi Solé-Casals (U Vic)
- Andrzej Cichocki (RIKEN)

Diagnostic of AD

- Toshimitsu Musha (Tokyo Tech)
- Jaeseung Jeong (KAIST)
- Corinna Hänschel (MPI&U Frankfurt/U Bangor)
- Johannes Pantel (MPI&U Frankfurt)

SES: Applications beyond AD

- Monique Maurice (RIKEN)
- Tomek Rutkowski (RIKEN)
- Danilo Mandic (Imperial)
- Shin Ishii (Kyoto University)
- Yuichi Sakumura (NAIST)

Localization of seizure onset

- Sydney Cash (MGH/Harvard)
- Alan Willsky (MIT)
- Emery Brown (MIT/MGH/Harvard)
- Emad Eskandar (MGH/Harvard)
Signal Processing for Diagnosis and Treatment of Brain Disorders

Justin Dauwels

www.dauwels.com

Laboratory for Information and Decision Systems (LIDS)
Massachusetts Institute of Technology (MIT)

March 18, 2009
Analogy: Waiting for a train

- Train may not arrive (e.g., mechanical problem) = event reliability $\rho$
- Train may or may not be on time = timing precision $s_t$
Alignment is equivalent to bipartite max-weight matching
Alignment is solved efficiently using a graphical model
Extension of SES from 2 to $n > 2$ bump models

Clusters appear naturally ($n = 5$)

$n$ models similar if:
- few deletions $\rightarrow$ number of bumps/cluster close to $n$
- little jitter in $t$ and $f$ $\rightarrow$ clusters have small dispersion
Extension from 2 to $n > 2$ bump models

$n$ bump models are generated from “mother” model

- distribution $p_c$ of number of bumps per cluster
- average $t$ and $f$ offset ($\delta_{t,i}$ and $\delta_{f,i}$), $i = 1, \ldots, n$
- $t$ and $f$ jitter ($s_{t,i}$ and $s_{f,i}$), $i = 1, \ldots, n$
We alternate clustering and parameter estimation

**Problem**
Given \( n \) bump models \( e \), infer \( p_c \) and \( \theta = (\delta_t, \delta_f, s_t, s_f) \)

**Approach: Maximum a posteriori estimation**

\[
(c^*, \theta^*) = \arg\max_{c, \theta} p(e, c; \theta)
\]

**Iterative algorithm: coordinate descent**

\[
\hat{c}^{(k+1)} = \arg\max_c p(e, c; \hat{\theta}^{(k)}) \quad \text{Clustering}
\]

\[
\hat{\theta}^{(k+1)} = \arg\max_{\theta} p(e, \hat{c}^{(k+1)}; \theta) \quad \text{Parameter Estimation}
\]

- Clustering is equivalent to **max-weight k-partite matching**
- **NP-hard** in general, but can be solved **efficiently** in practice
- Overall algorithm is provably **convergent**
Various extensions of SES

Point processes in other domains: \((t, x), (x, y, z), \ldots\)

- Analysis of spike trains \((t)\)
- Study of cell migration using molecular imaging \((t, x, y)\)

Key: meaningful events in signal (spike, burst, peak, etc.)

Time-dependent SES \(\rightarrow\) dynamics of synchrony

More complex “parent” structure (future work)
Want to know more? Want to try it out?

SES Toolbox webpage

Welcome to the SES project homepage!

SES stands for Stochastic Event Synchrony, and is a family of similarity measure for point processes. SES can be applied to one-dimensional (e.g., spike trains) and multi-dimensional point processes (e.g., sparse time-frequency representations of electrophysiological signals such as matching-pursuit representations, chiplets, Hilbert-Huang transforms, and bump models).

SES tries to align events in the point processes; the better the alignment, the more similar the point processes are considered to be. More precisely, the similarity is quantified by the following parameters: time delay, variance of the timing jitter, fraction of "non-aligned" events, and average similarity of the aligned events.

The SES measures may be viewed as extensions of the cost-based metrics of Victor et al. SES gives a statistical interpretation of those metrics, and therefore, it is able to automatically infer the unit costs, in contrast to the cost-based metrics. The latter are so far only applicable to one-dimensional point processes, whereas SES is applicable to multi-dimensional point processes as well.

We have used SES to:

- quantify the firing reliability of Morris-Lecar neurons,
- detect loss of EEG synchrony in Mild Cognitive Impairment patients,
- analyze steady-state visually evoked potentials and EEG responses to auditory stimuli,
- investigate the causal relation between morphological and molecular signaling events in cell migration.

http://www.dauwels.com/SESToolbox/SES.html
SES will also be included in Spike Train Analysis Toolkit

Spike Train Analysis Toolkit
Version 1.0 Gamma (December 8, 2006)
Documentation

- Introduction ← Read this first!
- Download and installation
- Example datasets and toolkit demos
- Text-based input file format
- Input data structure
- Output data structures
- Options and parameters for information methods
- Options and parameters for entropy methods
- Function reference
- C API (under construction)
- Version history and release notes
- License
- Contributing the toolkit

The Spike Train Analysis Toolkit was developed by David H. Goldberg, Jonathan D. Victor, and Daniel Gardner of the Laboratory of Neuroinformatics at Weill Medical College of Cornell University.

Direct questions and feedback to David Goldberg at dhg2002 at med dot cornell dot edu. Also email David if you would like to be added to the Spike Train Analysis Toolkit mailing list.

Funding provided by the Human Brain Project-Neuroinformatics initiative via MH068012 from NIMH, NINDS, NIA, NIBIB, and NSF with additional support from MH/NS057153 from NIMH and NINDS.

We suggest that publications that include results obtained from this resource include an acknowledgement such as
"Information theoretic analyses in this study were conducted with the Spike Train Analysis Toolkit—a neuroinformatics resource funded by the NIH's Human Brain Project."

http://neuroanalysis.org/toolkit/
Standard invasive approach requires long hospitalization

- Implant electrodes in the brain
- Record EEG until sufficient seizures obtained (5 days–more)
- Determine seizure onset zone from “seizure” EEG

Critical issues
- risk (infection, prolonged seizures, damage to the brain)
- discomfort
- cost
Univariate measures normal in hypersynchronous areas
How to determine seizure onset area using hypersynchrony?

Adaptive threshold required

- **Patient 1**
- **correlation coeff** $c$
- **histogram**

Graphical model

$$h = \arg\max_h p(s, h)$$

- local evidence (synchrony $s$)
- binary variable $h$ (hypersynchronous?)
- spatial correlation
Future work: Framework for clinical decision making

- Develop **graphical model** for each source of information
  - Features obtained through **signal processing**
    e.g., hypersynchrony in rest EEG
  - **Clinical behavior**
  - **Biophysical** prior knowledge

- **Combine** those models for decision making
  
  \[ \text{decision } d = \arg\max_d \log p(i, d) - q(d) \]

  decision \( d \), “information” \( i \), statistical model \( p \), risk \( q \)

- **Graphical models** → efficient **optimization algorithms**