Analyzing Oscillatory EEG Source Dynamics and Interactions using SIFT

Tim Mullen
14th EEGLAB Workshop
Mallorca, Spain (ICON XI)
Sept 22-25, 2011
Categorizations of Large-Scale Brain Connectivity Analysis

(Bullmore and Sporns, *Nature*, 2009)
Categorizations of Large-Scale Brain Connectivity Analysis

(Bullmore and Sporns, Nature, 2009)

**Structural**

state-invariant, anatomical

**Temporal Scale**

| Hours-Years | milliseconds-seconds |

Saturday, September 24, 2011
Categorizations of Large-Scale Brain Connectivity Analysis

(Bullmore and Sporns, Nature, 2009)

**Structural**
- state-invariant, anatomical

**Functional**
- dynamic, state-dependent, correlative, symmetric

**Temporal Scale**
- Hours-Years
- milliseconds-seconds
Categorizations of Large-Scale Brain Connectivity Analysis

(Bullmore and Sporns, *Nature*, 2009)

**Structural**
- state-invariant, anatomical

**Functional**
- dynamic, state-dependent, correlative, symmetric

**Effective**
- dynamic, state-dependent, asymmetric, causal, information flow

**Temporal Scale**
- Hours-Years
- milliseconds-seconds
Categorizations of Large-Scale Brain Connectivity Analysis

(Bullmore and Sporns, *Nature*, 2009)

- **Structural**: state-invariant, anatomical
- **Functional**: dynamic, state-dependent, correlative, symmetric
- **Effective**: dynamic, state-dependent, asymmetric, causal, information flow

Temporal Scale:
- Hours-Years
-Milliseconds-seconds
Estimating Functional Connectivity

- Cross-Correlation
- Coherence
- Phase-Locking Value
- Phase-amplitude coupling
- ...
Cross-Correlation and Linear Coherence
Cross-Correlation and Linear Coherence
Cross-Correlation and Linear Coherence
Cross-Correlation and Linear Coherence

\[
C_{AB}(f) = \sum_{k=0}^{P} \rho_{AB}(k)e^{-i2\pi fk}
\]

\[
\rho_{AB}(k) = \frac{S_{AB}(f)}{\sqrt{S_A(f)S_B(f)}}
\]
Cross-Correlation and Linear Coherence

Cross-correlation:

\[ \rho_{AB}(k) \]

Frequency:

\[ C_{AB}(f) = \sum_{k=0}^{P} \rho_{AB}(k) e^{-i2\pi fk} \]

\[ = \frac{S_{AB}(f)}{\sqrt{S_A(f)S_B(f)}} \]

Coherence:

\[ |C_{AB}(f)|^2 \]

Time lag:

\[ \rho_{AB}(k) \]
Theory: Euler’s Formula

Any sinusoid can be expressed as the sum of two complex numbers...

\[
A \cdot \cos(\omega t + \phi) = \frac{A}{2} e^{i(\omega t + \phi)} + \frac{A}{2} e^{-i(\omega t + \phi)}
\]

\[
= \text{Re}\{Ae^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\}
\]
Theory: Euler’s Formula

Phase shift \( \phi = \pi / 2 \)

Angular frequency \( \omega = 2\pi f = 2\pi \text{ rad/sec} \)

Any sinusoid can be expressed as the sum of two complex numbers...

\[
A \cdot \cos(\omega t + \phi) = \frac{A}{2} e^{i(\omega t + \phi)} + \frac{A}{2} e^{-i(\omega t + \phi)}
\]

\[
= \text{Re}\{A e^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\}
\]
**Theory: Euler’s Formula**

- **Phase shift**\( \phi = \frac{\pi}{2} \)
- **Angular frequency**\( \omega = 2\pi f = 2\pi \) rad/sec

Any sinusoid can be expressed as the sum of two complex numbers...

\[
A \cdot \cos(\omega t + \phi) = \frac{A}{2} e^{i(\omega t + \phi)} + \frac{A}{2} e^{-i(\omega t + \phi)}
\]

\[
= \text{Re}\{A e^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\}
\]

\[
\phi = \angle S(\omega, t)
\]

\[
|S(\omega, t)| = |A|
\]

- **Spectral Power**

- **Phasor!**

\( A=1.2 \)

\(-\pi/2\) 0 \(\pi/2\) \(\pi\) 3\(\pi/2\) 2\(\pi\)

\(t\)
Theory: Euler’s Formula

Any sinusoid can be expressed as the sum of two complex numbers...

\[ A \cdot \cos(\omega t + \phi) = \frac{A}{2} e^{i(\omega t + \phi)} + \frac{A}{2} e^{-i(\omega t + \phi)} \]

\[ = \text{Re}\{Ae^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\} \]

Another version:

\[ e^{i(\omega t + \phi)} = \cos(\omega t + \phi) + i \sin(\omega t + \phi) \]

\[ |S(\omega, t)| = |A| \]

\[ = \pi / 2 \]

\[ \phi = \angle S(\omega, t) \]

Real part
Cosine component

Imaginary part
Sine component

Spectral Power

Phasor!
Phasors

\[ A \cdot \cos(\omega t + \phi) = \text{Re}\{Ae^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\} \]

\[ \phi = \angle S(\omega, t) = \pi / 2 \]

\[ |S(\omega, t)| = |A| \]

Phasor!
Phasors

\[ A \cdot \cos(\omega t + \phi) = \text{Re}\{A e^{i(\omega t + \phi)}\} \]
\[ = \text{Re}\{S(\omega, t)\} \]

\[ \phi = \angle S(\omega, t) \]
\[ = \pi / 2 \]

\[ |S(\omega, t)| = |A| \]

Phasor!
Phasors

Rotation velocity (Rad/S; Hz) = (angular) frequency ($\omega; f$)

$A \cdot \cos(\omega t + \phi) = \text{Re}\{A e^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\}$

$\phi = \angle S(\omega, t)$

$|S(\omega, t)| = |A| = \pi / 2$

Polar animations courtesy Wikipedia

Phasor!
Phasors

Rotation velocity (Rad/S; Hz) = (angular) frequency ($w; f$)

$$A \cdot \cos(\omega t + \phi) = \text{Re}\{A e^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\}$$

$$\phi = \angle S(\omega, t) = \frac{\pi}{2}$$

$$|S(\omega, t)| = |A|$$

Polar animations courtesy Wikipedia
Phasors

Rotation velocity (Rad/S; Hz) = (angular) frequency ($w$; $f$)

$$A \cdot \cos(\omega t + \phi) = \text{Re}\{Ae^{i(\omega t + \phi)}\} = \text{Re}\{S(\omega, t)\}$$

$$\phi = \angle S(\omega, t)$$

$$|S(\omega, t)| = |A| = \pi / 2$$

Shorthand phasor notation: $Ae^{i\phi}$

Polar animations courtesy Wikipedia
Phase-Locking Value (PLV)


The main advantage of this approach is that it does not require any a priori hypothesis on the signals. We test the hypothesis that the two series of phase values are independent. For this purpose, we generate 200 new series of variables, constructed by shuffling the trials of one of the electrodes (see text for details). By averaging these phase differences across the trials, we obtain a complex value \( u \) (for each latency \( t \)), which amplitude (\( |u| \)) is the phase-locking value. Right: Surrogate data are constructed by shuffling the trials of one of the electrodes (see text for details).

Figure 2.

For each surrogate series \( y \)(\( shuffle (n) \)), where \( y(i) \) is the signal recorded at electrode 2 during trial \( i \) (Fig. 2). The proportion of surrogate values higher than the original PLV (between \( x \) and \( y \)) is the phase-locking statistics (PLS). It measures the probability of having false positives for a given level of significance. In this study, we used a criterion of 5% (PLS \( \text{HBM} \)).
Phase-Locking Value (PLV)


The main advantage of this approach is that it does not require any a priori hypothesis on the signals. We use a statistical test based on randomization and is adapted to our particular set of data. When the sampling distribution of a statistic is unknown, one must rely on recent techniques of randomization, or bootstrap [Fisher, 1993]. Our statistical test is constructed by shuffling the trials of one of the electrodes (see text for details). By averaging these phase differences across the trials, we obtain a complex value \( u \) (for each latency \( t \)), which amplitude (abs \( \phi_1 - \phi_2 \)) is the phase-locking value. Right: Surrogate data are constructed by shuffling the trials of one of the electrodes (see text for details). By averaging these phase differences across the trials, we obtain a complex value \( u \) (for each latency \( t \)), which amplitude (abs \( \phi_1 - \phi_2 \)) is the phase-locking value.
The main advantage of this approach is that it does not require any a priori hypothesis on the signals. We test the hypothesis that the two series of phase values higher than the original PLV (between x and y) would detect a phase-locking between the groups.

For each surrogate series $y^{(n)}$, we measure the maximum between x and y, and we define a criterion of 5% (PLS$_{5%}$) to characterize significant differences. The proportion of surrogate series $y^{(n)}$ such that

$$
\phi_2(1, t_0) - \phi_1(1, t_0) > \phi_2(2, t_0) - \phi_1(2, t_0)
$$

measures the probability of having false positives for a given level of significance. In this study, we used a criterion of 5% (PLS$_{5%}$). The requirement of significance in the context of the phase-locking value $\phi$ is a function of the number of trials and the number of surrogate series.

Estimation of phase-locking value. Left: Our synchrony index $\phi$ is directly related to the intertrial variability of the phase differences. By averaging these phase differences across the trials, we obtain a complex value $u$ (for each latency $t$), which corresponds to the following $\text{PLV} = \frac{1}{N} \sum_{i=1}^{N} \phi_i$. Right: Surrogate data are constructed by shuffling the trials of one of the electrodes (see text for details). We generate 200 new series of variables, $y^{(n)}$, which have the same characteristics as the original signals coming from electrode 2, except that we build them to be independent of the signals coming from electrode 1. These series are created by shuffling the trials within the measures of electrode 2 to make new series $y^{(n)}$, where $(y(i))$ is the signal recorded at electrode 2 during trial $i$ (Fig. 2).
Phase-Locking Value (PLV)


$$A_1 e^{i\phi_1}$$

$$A_2 e^{i\phi_2}$$

$$e^{i(\phi_2 - \phi_1)}$$

$$\phi_1(1,t_0)$$

$$\phi_2(1,t_0)$$

$$\phi_2(1,t_0) - \phi_1(1,t_0)$$

$$\phi_2(2,t_0) - \phi_1(2,t_0)$$

$$\phi_2(N,t_0) - \phi_1(N,t_0)$$
Phase-Locking Value (PLV)


The main advantage of this approach is that it does not require any a priori hypothesis on the signals. We test the hypothesis that the two series of phase values higher than the original PLV (between x and y) are independent. For this purpose, we generate 200 new series of variables, which have the same characteristics as the original signals being studied. Our method is related to an approach proposed by Friston et al. [1997] to quantify the synchrony, but this is, of course, a function of the required rigor of significance in the context of the experimental setup.

For each surrogate series $y(n)$, we measure the maximum between $x$ and $y$ in time. These 200 values are used to estimate the significance of PLV between the original signals $x$ and $y$. The proportion of surrogate values higher than the original PLV is called phase-locking statistics (PLS).

Figure 2. MEK data. In fact, they propose to estimate the difference phasor $e^{i(\phi_2 - \phi_1)}$.

Details): By averaging these phase differences across the trials, we obtain a complex value $u$ (for each latency $t$), which amplitude (absolute value) corresponds to the phase-locking value. Right: Surrogate data are constructed by shuffling the trials of one of the electrodes (see text for details).
Phase-Locking Value (PLV)


For each surrogate series $y^{(n)}$, we measure the maximum between $x$ and $y$:

$$\phi_1(t_0) - \phi_1(t_0)$$

AVERAGE difference phasors across trials

$$u(t_0) = \frac{1}{N} \sum_{k=1}^{N} e^{i(\phi_2(k,t_0) - \phi_2(k,t_0))}$$

$$\text{PLV}(t_0) = \text{abs}(u)$$

**Figure 2.**
Phase-Locking Value (PLV)


Computing PLV ("phase coherence") in EEGLAB:
`pop_newcrossf(....,'type','phase')`
3.3. Sigmoidal coupling

Signals which provide control over the type of coupling produced, coupling. This examination takes place using a variety of synthetic fully and partially synchronized network dynamics.

significantly worse than the GLM, PLV or ESC measures for both

been smoothed for presentation purposes using a first-order bi-

Hz and a 'gamma' band from 50 to 130 Hz. The AUC curves have

in these curves was derived by generating 100 trials of PAC data

Fig. 5.

We first generate a stationary theta oscillation

$\sin(2\pi f t)$

We first generate a stationary theta oscillation

$\sin(2\pi f t)$

The observed theta oscillation is delayed by a number of samples

Fig. 7

The following subsections systematically examine the depen-

$\theta$ amplitude and frequency are set to

theta cycle. For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

Parameter that were not varied for

Fig. 6

6. The plots in the bottom row show the area under the curve (AUC) as a function of observation noise,

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.

For the simulations in this section we use the following

The gamma oscillation is given by

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$\sin(2\pi f t)$

$k$ to pause

$\theta$ cycle.
Phase-Amplitude Coupling

‘burst-suppress’ oscillators

Phase-Amplitude Coupling

'burst-suppress' oscillators

Phase-Amplitude Coupling

'burst-suppress' oscillators


Local Field Potential (Slow + Fast cells)
• May present a functional role in execution of cognitive functions (Axmacher et al. 2010; Cohen et al. 2009a,b; Lakatos et al. 2008; Tort et al. 2008, 2009).

• Suggested involvement in sensory signal detection (Handel and Haarmeier 2009), attentional selection (Schroeder and Lakatos 2009), and memory processes (Axmacher et al. 2010; Tort et al. 2009)
Phase-Amplitude Coupling: PLV Method


\[ X_1 \]

\[ t_0 \]

original raw signal
Phase-Amplitude Coupling: PLV Method  

\[ X_1 \quad t_0 \]

Original raw signal

\[ f_p \]

Filter \( X_1 \) at phase-modulation band (e.g. theta)
Phase-Amplitude Coupling: PLV Method


$X_1$ original raw signal

$f_p$ filter $X_1$ at phase-modulation band (e.g. theta)

$f_A$ filter $X_1$ at amp-modulation band (e.g. gamma)
Phase-Amplitude Coupling: PLV Method  

- $X_1$ (original raw signal)
- $f_p$ (filter $X_1$ at phase-modulation band (e.g. theta))
- $f_A$ (filter $X_1$ at amp-modulation band (e.g. gamma))
- Get amplitude envelope of filtered signal
Phase-Amplitude Coupling: PLV Method


$X_1$  

$\tau_0$  

original raw signal

$f_p$  

filter $X_1$ at phase-modulation band (e.g. theta)

$f_A$  

filter $X_1$ at amp-modulation band (e.g. gamma)

$A_{f_A}$  

get amplitude envelope of filtered signal
Phase-Amplitude Coupling: PLV Method  

Compute PLV between phase-modulation time-series ($f_p$) and amplitude envelope of amplitude modulation time-series ($A_{fA}$). Significant PLV indicates that the central frequency of $f_p$ modulates the amplitude of the central frequency of $f_A$.
Problem:

PLV is invariant to differences in amplitude between the two time-series (it only considers phase). Thus PLV-PAC doesn’t take into account the \textit{amplitude} of the co-modulation.

In the example below, $X_1$ and $X_2$ both would produce the same PAC, even though the high-frequency amplitude of $X_2$ clearly is more strongly modulated by the low-frequency rhythm.
Phase-Amplitude Coupling: Modulation Index Method


original raw signal
Phase-Amplitude Coupling: Modulation Index Method


Original raw signal

Filter $X_1$ at phase-modulation band (e.g. theta)
Phase-Amplitude Coupling: Modulation Index Method


$X_1$

$\hat{f}_p$

$\phi_{\hat{f}_p}$

original raw signal

filter $X_1$ at phase-modulation band (e.g. theta)

extract the instantaneous phase of $f_p$
Phase-Amplitude Coupling: Modulation Index Method

- $X_1$ (original raw signal)
- $f_p$ (filter $X_1$ at phase-modulation band, e.g. theta)
- $\phi_{f_p}$ (extract the instantaneous phase of $f_p$)
- $f_A$ (filter $X_1$ at amp-modulation band, e.g. gamma)
Phase-Amplitude Coupling: Modulation Index Method

- **$X_1$**: original raw signal
- **$f_p$**: filter $X_1$ at phase-modulation band (e.g. theta)
- **$\phi_{f_p}$**: extract the instantaneous phase of $f_p$
- **$f_A$**: filter $X_1$ at amp-modulation band (e.g. gamma)
- **$A_{f_A}$**: get amplitude envelope of filtered signal
Phase-Amplitude Coupling: Modulation Index Method

Phase-Amplitude Coupling: Modulation Index Method


\[ \phi_{f_p} \quad -\pi \quad 0 \quad \pi \]

\[ A_{f_A} \]

Trial 1
Phase-Amplitude Coupling: Modulation Index Method

$z_1(t_0) = A_{f_A} e^{i\phi_{fp}}$

$A_{f_A}(t_0)$

$\phi_{fp}(t_0)$

build complex phasor with instantaneous amplitude and phase
Phase-Amplitude Coupling: Modulation Index Method


\[
\chi_1(t_0) = A_{f_A} e^{i\phi_{fp}}
\]

\[
A_{f_A}(t_0)
\]

\[
\phi_{fp}(t_0)
\]

\[
\begin{align*}
A_{f_A} & \text{ (instantaneous amplitude)} \\
\phi_{fp} & \text{ (instantaneous phase)} \\
\end{align*}
\]

other trials

AVERAGE complex phasors across trials

\[
\mathbf{u}(t_0) = \frac{1}{N} \sum_{k=1}^{N} z_k(t_0)
\]

PAC(t_0) = \text{abs}(\mathbf{u})

build complex phasor with instantaneous amplitude and phase

Phase-Amplitude Coupling: Modulation Index Method


\[
\chi_1(t_0) = A_{f_A} e^{i\phi_{fp}}
\]

\[
A_{f_A}(t_0)
\]

\[
\phi_{fp}(t_0)
\]

\[
\begin{align*}
A_{f_A} & \text{ (instantaneous amplitude)} \\
\phi_{fp} & \text{ (instantaneous phase)} \\
\end{align*}
\]

other trials

AVERAGE complex phasors across trials

\[
\mathbf{u}(t_0) = \frac{1}{N} \sum_{k=1}^{N} z_k(t_0)
\]

PAC(t_0) = \text{abs}(\mathbf{u})
Phase-Amplitude Coupling: Modulation Index Method


\[ z_1(t_0) = A_{f_A} e^{i\phi_{f_p}} \]

Comparison: PLV-PAC

**AVERAGE complex phasors across trials**

\[ u(t_0) = \frac{1}{N} \sum_{k=1}^{N} z_k(t_0) \]

PAC(t_0) = \text{abs}(u)
Phase-Amplitude Coupling: Modulation Index Method

Computing PAC in EEGLAB:

\[ \text{pac(IC1,IC2,..., 'method', 'mod')} \]

PAC can also be applied between sources/channels (e.g. determine whether the phase of oscillation at freq. \( w_p \) in IC1 modulates the amplitude of oscillation at freq. \( w_A \) in IC2. This leads to a measure of cross-frequency (non-linear) functional connectivity.

For Modulation Index method (other modes also available)
(Cross)-Correlation ≠ Causation
(Cross)-Correlation ≠ Causation
(Cross)-Correlation ≠ Causation
(Cross)-Correlation ≠ Causation
(Cross)-Correlation ≠ Causation

Coherence/CC/PLV/PAC indicate functional, but not effective connectivity
Estimating Effective Connectivity

- **Non-Invasive**
  - *Post-hoc* analyses applied to measured neural activity
  - Confirmatory
    - Dynamic Causal Models
    - Structural Equation Models
  - Exploratory
    - *Granger-Causal methods*
Estimating Effective Connectivity

- Non-Invasive
  - *Post-hoc* analyses applied to measured neural activity
  - Confirmatory
    - Dynamic Causal Models
    - Structural Equation Models
  - Exploratory
    - Granger-Causal methods

• Data-driven
Estimating Effective Connectivity

Non-Invasive

- *Post-hoc* analyses applied to measured neural activity

Confirmatory

- Dynamic Causal Models
- Structural Equation Models

Exploratory

- *Granger-Causal methods*

- Data-driven
- Simple, but powerful

Saturday, September 24, 2011
Estimating Effective Connectivity

Non-Invasive

- Post-hoc analyses applied to measured neural activity
- Confirmatory
  - Dynamic Causal Models
  - Structural Equation Models
- Exploratory
  - Granger-Causal methods

- Data-driven
- Simple, but powerful
- Scalable (Valdes-Sosa, 2005)
Estimating Effective Connectivity

Non-Invasive

- *Post-hoc* analyses applied to measured neural activity
- Confirmatory
  - Dynamic Causal Models
  - Structural Equation Models
- Exploratory
  - Granger-Causal methods

- Data-driven
- Simple, but powerful
- Scalable (Valdes-Sosa, 2005)
- Extendable to nonlinear and/or non-stationary systems (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
Estimating Effective Connectivity

- **Non-Invasive**
  - *Post-hoc* analyses applied to measured neural activity
  - Confirmatory
    - Dynamic Causal Models
    - Structural Equation Models
  - Exploratory
    - Granger-Causal methods

- **Data-driven**
- **Simple, but powerful**
- **Scalable** (Valdes-Sosa, 2005)
- **Extendable to nonlinear and/or non-stationary systems** (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
- **Extendable to non-parametric representations** (Dhamala, 2009a,b)

Saturday, September 24, 2011
Estimating Effective Connectivity

Non-Invasive

- Post-hoc analyses applied to measured neural activity
- Confirmatory
  - Dynamic Causal Models
  - Structural Equation Models
- Exploratory
  - Granger-Causal methods

- Data-driven
- Simple, but powerful
- Scalable (Valdes-Sosa, 2005)
- Extendable to nonlinear and/or non-stationary systems (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
- Extendable to non-parametric representations (Dhamala, 2009a,b)
- Can be (partially) controlled for (unobserved) exogenous causes (Guo, 2008a,b; Ge, 2009)
Estimating Effective Connectivity

Non-Invasive

- *Post-hoc* analyses applied to measured neural activity
- Confirmatory
  - Dynamic Causal Models
  - Structural Equation Models
- Exploratory
  - Granger-Causal methods

- Data-driven
- Simple, but powerful
- Scalable (Valdes-Sosa, 2005)
- Extendable to nonlinear and/or non-stationary systems (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
- Extendable to non-parametric representations (Dhamala, 2009a,b)
- Can be (partially) controlled for (unobserved) exogenous causes (Guo, 2008a,b; Ge, 2009)
- Equivalent to Transfer Entropy for Gaussian Variables (Seth, 2009)
Estimating Effective Connectivity

- **Non-Invasive**
  - *Post-hoc* analyses applied to measured neural activity
  - Confirmatory
    - Dynamic Causal Models
    - Structural Equation Models
  - Exploratory
    - Granger-Causal methods

- **Data-driven**
- **Simple, but powerful**
- **Scalable** (Valdes-Sosa, 2005)
- **Extendable to nonlinear and/or non-stationary systems** (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
- **Extendable to non-parametric representations** (Dhamala, 2009a,b)
- **Can be (partially) controlled for (unobserved) exogenous causes** (Guo, 2008a,b; Ge, 2009)
- **Equivalent to Transfer Entropy for Gaussian Variables** (Seth, 2009)
- **Flexibly allows us to examine time-varying (dynamic) multivariate causal relationships in either the time or frequency domain**
Granger Causality

- First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- Relies on two assumptions:
Granger Causality

- First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- Relies on two assumptions:

  Granger Causality Axioms
Granger Causality

- First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- Relies on two assumptions:

<table>
<thead>
<tr>
<th>Granger Causality Axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Causes should precede their effects in time (Temporal Precedence)</td>
</tr>
</tbody>
</table>
Granger Causality

- First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- Relies on two assumptions:

<table>
<thead>
<tr>
<th>Granger Causality Axioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Causes should precede their effects in time (Temporal Precedence)</td>
</tr>
<tr>
<td>2. Information in a cause’s past should improve the prediction of the effect, above and beyond the information contained in past of the effect (and other measured variables)</td>
</tr>
</tbody>
</table>
Multivariate (Vector) Autoregressive (VAR) Modeling

EEG

\[ X_1(t) \]
\[ X_2(t) \]

\[ \vdots \]  

\[ X_M(t) \]

\[ \text{VAR} \]

\[ \text{Granger Causality} \]

\[ \text{Coherence} \]

\[ \text{Spectrum} \]

\[ \ldots \]
The VAR Process

Stochastic Linear Dynamical System

\[ X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \]
\[ X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) \]
The VAR Process

Stochastic Linear Dynamical System

\[ X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \]

\[ X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) \]

\[ \Delta t = 1 \]
The VAR Process

**Stochastic Linear Dynamical System**

\[
X_1(t) = a(t)_{11}X_1(t-1) + a(t)_{12}X_2(t-1) + \epsilon_1(t)
\]

\[
X_2(t) = a(t)_{21}X_1(t-1) + a(t)_{22}X_2(t-1) + \epsilon_2(t)
\]

- **X_1(0)**
- **X_2(0)**
- \(t=0\)
- \(\Delta t = 1\)
The VAR Process

**Stochastic Linear Dynamical System**

\[
\begin{align*}
X_1(t) &= a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \\
X_2(t) &= a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t)
\end{align*}
\]
The VAR Process

Stochastic Linear Dynamical System

\[ X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \]

\[ X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) \]
The VAR Process

Stochastic Linear Dynamical System

\[ X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \]
\[ X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) \]

![Diagram of the VAR Process](image-url)
The VAR Process

Stochastic Linear Dynamical System

\[ X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \]
\[ X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) \]

\[
\begin{align*}
X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \\
X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) 
\end{align*}
\]

\[
\begin{align*}
X_1(0) &\rightarrow X_1(1) & X_1(n-2) &\rightarrow X_1(n-1) \\
X_2(0) &\rightarrow X_2(1) & X_2(n-2) &\rightarrow X_2(n-1) 
\end{align*}
\]

\[
\begin{align*}
\Delta t &= 1 \\
\epsilon_1(t) &\rightarrow a(t)_{11} \\
\epsilon_1(t) &\rightarrow a(t)_{12} \\
\epsilon_1(t) &\rightarrow a(t)_{21} \\
\epsilon_1(t) &\rightarrow a(t)_{22} \\
\epsilon_2(t) &\rightarrow a(t)_{11} \\
\epsilon_2(t) &\rightarrow a(t)_{12} \\
\epsilon_2(t) &\rightarrow a(t)_{21} \\
\epsilon_2(t) &\rightarrow a(t)_{22} 
\end{align*}
\]

\[
\begin{align*}
X_1(0) &\rightarrow X_1(1) & X_1(n-2) &\rightarrow X_1(n-1) \\
X_2(0) &\rightarrow X_2(1) & X_2(n-2) &\rightarrow X_2(n-1) 
\end{align*}
\]

\[
\begin{align*}
\Delta t &= 1 \\
\epsilon_1(t) &\rightarrow a(t)_{11} \\
\epsilon_1(t) &\rightarrow a(t)_{12} \\
\epsilon_1(t) &\rightarrow a(t)_{21} \\
\epsilon_1(t) &\rightarrow a(t)_{22} \\
\epsilon_2(t) &\rightarrow a(t)_{11} \\
\epsilon_2(t) &\rightarrow a(t)_{12} \\
\epsilon_2(t) &\rightarrow a(t)_{21} \\
\epsilon_2(t) &\rightarrow a(t)_{22} 
\end{align*}
\]

\[
\begin{align*}
X_1(0) &\rightarrow X_1(1) & X_1(n-2) &\rightarrow X_1(n-1) \\
X_2(0) &\rightarrow X_2(1) & X_2(n-2) &\rightarrow X_2(n-1) 
\end{align*}
\]

\[
\begin{align*}
\Delta t &= 1 \\
\epsilon_1(t) &\rightarrow a(t)_{11} \\
\epsilon_1(t) &\rightarrow a(t)_{12} \\
\epsilon_1(t) &\rightarrow a(t)_{21} \\
\epsilon_1(t) &\rightarrow a(t)_{22} \\
\epsilon_2(t) &\rightarrow a(t)_{11} \\
\epsilon_2(t) &\rightarrow a(t)_{12} \\
\epsilon_2(t) &\rightarrow a(t)_{21} \\
\epsilon_2(t) &\rightarrow a(t)_{22} 
\end{align*}
\]
The VAR Process

Stochastic Linear Dynamical System

\[ X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t) \]
\[ X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t) \]
**The VAR Process**

**Stochastic Linear Dynamical System**

\[ X_1(t) = a(1)_{11}X_1(t-1) + a(1)_{12}X_2(t-1) + \epsilon_1(t) \]

\[ X_2(t) = a(2)_{21}X_1(t-1) + a(1)_{22}X_2(t-1) + \epsilon_2(t) \]

**Order 1 VAR Process (VAR[1])**

\[ X_1(t) = a(1)_{11}X_1(t-1) + a(1)_{12}X_2(t-1) + \epsilon_1(t) \]

\[ X_2(t) = a(1)_{21}X_1(t-1) + a(1)_{22}X_2(t-1) + \epsilon_2(t) \]

\[ X_1(n-1) = a(n-1)_{11}X_1(n-2) + a(n-1)_{12}X_2(n-2) + \epsilon_1(n-1) \]

\[ X_2(n-1) = a(n-1)_{21}X_1(n-2) + a(n-1)_{22}X_2(n-2) + \epsilon_2(n-1) \]

\[ X_1(n) = a(n)_{11}X_1(n-2) + a(n)_{12}X_2(n-2) + \epsilon_1(n) \]

\[ X_2(n) = a(n)_{21}X_1(n-2) + a(n)_{22}X_2(n-2) + \epsilon_2(n) \]
The VAR Model
The VAR Model

\[ X(t) = x_1(t) \]
\[ x_2(t) \]
\[ \vdots \]
\[ x_M(t) \]
The VAR Model

\[ X(t) = x_1(t) \quad x_2(t) \quad \cdots \quad x_M(t) \]
The VAR Model

\[ X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix} \]

\[ X(t) = \sum_{k=1}^{p} A^{(k)}(t) X(t - k) + E(t) \]

M-channel data vector at current time \( t \)
The VAR Model

\[ X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix} \]

\[ \mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}(t) \mathbf{X}(t - k) + \mathbf{E}(t) \]

- M-channel data vector at current time \( t \)
- M x M matrix of (time-varying) model coefficients indicating variable dependencies at lag \( k \)

\[ \mathbf{A}^{(k)}(t) = \begin{pmatrix} a_{11}^{(k)}(t) & \cdots & a_{1M}^{(k)}(t) \\ \vdots & \ddots & \vdots \\ a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t) \end{pmatrix} \]
The VAR Model

VAR\([p]\) model

\(X(t) = \sum_{k=1}^{p} A^{(k)}(t)X(t - k) + E(t)\)

- M-channel data vector at current time \(t\)
- M x M matrix of (time-varying) model coefficients indicating variable dependencies at lag \(k\)
- Multichannel data \(k\) samples in the past

\[A^{(k)}(t) = \begin{pmatrix}
    a_{11}^{(k)}(t) & \cdots & a_{1M}^{(k)}(t) \\
    \vdots & \ddots & \vdots \\
    a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t)
\end{pmatrix}\]
The VAR Model

$$X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix}$$

The VAR Model is given by:

$$X(t) = \sum_{k=1}^{p} A^{(k)}(t)X(t-k) + E(t)$$

$$A^{(k)}(t) = \begin{pmatrix} a_{11}^{(k)}(t) & \cdots & a_{1M}^{(k)}(t) \\ \vdots & \ddots & \vdots \\ a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t) \end{pmatrix}$$

- **M-channel data vector at current time** $t$
- **M x M matrix of (time-varying) model coefficients indicating variable dependencies at lag** $k$
- **multichannel data $k$ samples in the past**
- **model order**
The VAR Model

\[ X(t) = \sum_{k=1}^{p} A^{(k)}(t)X(t - k) + E(t) \]

\[ A^{(k)}(t) = \begin{pmatrix}
    a_{11}^{(k)}(t) & \cdots & a_{1M}^{(k)}(t) \\
    \vdots & \ddots & \vdots \\
    a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t)
\end{pmatrix} \]

\[ E(t) = N(0, V) \]
The VAR Model

\[
X(t) = \begin{bmatrix}
x_1(t) \\
x_2(t) \\
\vdots \\
x_M(t)
\end{bmatrix}
\]

\[
X(t) = \sum_{k=1}^{p} A^{(k)}(t)X(t-k) + E(t)
\]

\[
A^{(k)}(t) = \begin{bmatrix}
a_{11}^{(k)}(t) & \cdots & a_{1M}^{(k)}(t) \\
\vdots & \ddots & \vdots \\
a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t)
\end{bmatrix}
\]

\[
E(t) = N(0, V)
\]
Selecting a VAR Model Order

- Model order is typically determined by minimizing information criteria such as Akaike Information Criterion (AIC) for varying model order (p):

\[ \text{AIC}(p) = 2 \log(\det(V)) + M^2 p / N \]
Selecting a VAR Model Order

- Model order is typically determined by minimizing information criteria such as Akaike Information Criterion (AIC) for varying model order (p):

\[
\text{AIC}(p) = 2\log(\det(V)) + M^2p/N
\]
Selecting a VAR Model Order

- Model order is typically determined by minimizing information criteria such as Akaike Information Criterion (AIC) for varying model order \( p \): 

\[
\text{AIC}(p) = 2 \log(\det(\mathbf{V})) + M^2 p / N
\]

Penalizes high model orders (parsimony)

Entropy rate (amount of prediction error)
Selecting a VAR Model Order

- Model order is typically determined by minimizing information criteria such as Akaike Information Criterion (AIC) for varying model order (p):

\[
\text{AIC}(p) = 2 \log(\det(V)) + \frac{M^2 p}{N}
\]

Penalizes high model orders (parsimony)

entropy rate (amount of prediction error)
Selecting a VAR Model Order

- Model order is typically determined by minimizing information criteria such as Akaike Information Criterion (AIC) for varying model order (p):

\[
\text{AIC}(p) = 2\log(\det(\mathbf{V})) + M^2 p/N
\]

- Entropy rate (amount of prediction error)
- Penalties high model orders (parsimony)
- Optimal order
Selecting a VAR Model Order

- **Other considerations:**
  
  - A $M$-dimensional VAR model of order $p$ has at most $Mp/2$ spectral peaks distributed amongst the $M$ variables. This means we can observe at most $p/2$ peaks in each variables’ spectrum (or in the causal spectrum between two each pair of variables)
  
  - Optimal model order depends on sampling rate (higher sampling rate often requires higher model orders)
Selecting a VAR Model Order
Selecting a VAR Model Order

- Jansen (1981) and Florian and Pfurtscheller (1995) demonstrated that a model order of 10 was generally quite adequate for describing EEG spectra.
Selecting a VAR Model Order

- Jansen (1981) and Florian and Pfurtscheller (1995) demonstrated that a model order of 10 was generally quite adequate for describing EEG spectra.
- VAR model is an “all-pole” filter well-suited for modeling oscillatory processes with “peaky” spectra (like EEG!).
VAR Modeling: Assumptions
VAR Modeling: Assumptions

- "Weak" stationarity of the data
VAR Modeling: Assumptions

- “Weak” stationarity of the data
- mean and variance do not change with time
VAR Modeling: Assumptions

- “Weak” stationarity of the data
  - mean and variance do not change with time
  - An EEG trace containing prominent evoked potentials is a classic example of a non-stationary time-series
VAR Modeling: Assumptions

- "Weak" stationarity of the data
  - mean and variance do not change with time
  - An EEG trace containing prominent evoked potentials is a classic example of a non-stationary time-series

- Stability
VAR Modeling: Assumptions

- "Weak" stationarity of the data
  - mean and variance do not change with time
  - An EEG trace containing prominent evoked potentials is a classic example of a non-stationary time-series

- Stability
  - A stable process will not "blow up" (diverge to infinity)
VAR Modeling: Assumptions

- **“Weak” stationarity of the data**
  - mean and variance do not change with time
  - An EEG trace containing prominent evoked potentials is a classic example of a non-stationary time-series

- **Stability**
  - A stable process will not “blow up” (diverge to infinity)
  - Importantly, stability implies stationarity and SIFT provides you techniques for verifying the stability
Granger Causality

Does $X_4$ granger-cause $X_1$?
(conditioned on $X_2, X_3$)
Granger Causality

Does $X_4$ granger-cause $X_1$?
(conditioned on $X_2, X_3$)

$x(t) = \sum_{k=1}^{p} A^{(k)} x(t-k) + e(t)$
Granger Causality

Does $X_4$ granger-cause $X_1$?
(conditioned on $X_2, X_3$)

$X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t)$

prediction error for $X_1$
(variance of residuals $E_1$)

Saturday, September 24, 2011
Granger Causality

Does \( X_4 \) granger-cause \( X_1 \)?
(conditioned on \( X_2, X_3 \))

\[
X(t) = \sum_{k=1}^{p} A^{(k)}X(t - k) + E(t)
\]

\( \text{VAR}_1 \)

Prediction error for \( X_1 \)
(variance of residuals \( E_1 \))

Saturday, September 24, 2011
Granger Causality

Does $X_4$ granger-cause $X_1$?
(conditioned on $X_2, X_3$)

$X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t)$

prediction error for $X_1$
(variance of residuals $E_1$)

Saturday, September 24, 2011
Granger Causality

Does $X_4$ granger-cause $X_1$?

(conditional on $X_2, X_3$)

$X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t)$

$X_{-4}(t) = \sum_{k=1}^{p} \tilde{A}^{(k)} X_{-4}(t - k) + \tilde{E}(t)$

prediction error for $X_1$
(variance of residuals $E_1$)
Granger Causality

Does $X_4$ granger-cause $X_1$?
(conditioned on $X_2, X_3$)

$X(t) = \sum_{k=1}^{p} A^{(k)}X(t-k) + E(t)$

Independence of $X_4(t)$ and $E_1(t)$ (variance of residuals $E_1$)

$X_{-4}(t) = \sum_{k=1}^{p} \tilde{A}^{(k)}X_{-4}(t-k) + \tilde{E}(t)$
Granger Causality

Does $X_4$ granger-cause $X_1$?
(conditioned on $X_2, X_3$)

$X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t)$

Prediction error for $X_1$ (variance of residuals $E_1$)

$X_{-4}(t) = \sum_{k=1}^{p} \tilde{A}^{(k)} X_{-4}(t - k) + \tilde{E}(t)$
Granger Causality
Granger Causality

Granger (1969) quantified this definition for **bivariate** processes in the form of an F-ratio:

\[
F_{X_1 \leftarrow X_2} = \ln \left( \frac{\text{var}(\tilde{E}_1)}{\text{var}(E_1)} \right) = \ln \left( \frac{\text{var}(X_1(t) \mid X_1(\cdot))}{\text{var}(X_1(t) \mid X_1(\cdot), X_2(\cdot))} \right)
\]
Granger Causality

Granger (1969) quantified this definition for **bivariate** processes in the form of an F-ratio:

\[
F_{X_1 \leftarrow X_2} = \ln \left( \frac{\text{var}(\tilde{E}_1)}{\text{var}(E_1)} \right) = \ln \left( \frac{\text{var}(X_1(t) \mid X_1(\cdot))}{\text{var}(X_1(t) \mid X_1(\cdot), X_2(\cdot))} \right)
\]

Alternately, for a **multivariate interpretation** we can fit a single MVAR model to all channels and apply the following definition:

**Definition 1**

\(X_j\) granger-causes \(X_i\) conditioned on all other variables in \(X\) if and only if \(A_{ij}(k) > 0\) for some lag \(k \in \{1, \ldots, p\}\)
Granger Causality Quiz

Example: 2-channel MVAR process of order 1

\[
\begin{bmatrix}
X_1(t) \\
X_2(t)
\end{bmatrix} =
\begin{bmatrix}
-0.5 & 0 \\
0.7 & 0.2
\end{bmatrix}
\begin{bmatrix}
X_1(t-1) \\
X_2(t-1)
\end{bmatrix} +
\begin{bmatrix}
E_1(t) \\
E_2(t)
\end{bmatrix}
\]

\[
X_1(t) = -0.5X_1(t-1) + 0X_2(t-1) + E_1(t)
\]

\[
X_2(t) = 0.7X_1(t-1) + 0.2X_2(t-1) + E_2(t)
\]

Which causal structure does this model correspond to?

a) 1 → 2  
b) 1 ↔ 2  
c) 1 ↔ 2
Granger Causality Quiz

- Example: 2-channel MVAR process of order 1

\[
\begin{pmatrix}
X_1(t) \\
X_2(t)
\end{pmatrix} =
\begin{pmatrix}
-0.5 & 0 \\
0.7 & 0.2
\end{pmatrix}
\begin{pmatrix}
X_1(t-1) \\
X_2(t-1)
\end{pmatrix} +
\begin{pmatrix}
E_1(t) \\
E_2(t)
\end{pmatrix}
\]

\[
X_1(t) = -0.5X_1(t-1) + 0X_2(t-1) + E_1(t)
\]

\[
X_2(t) = 0.7X_1(t-1) + 0.2X_2(t-1) + E_2(t)
\]

Which causal structure does this model correspond to?

a) 1 → 2  
b) 1 ← 2  
c) 1 ↔ 2
Granger Causality Quiz

- Example: 2-channel MVAR process of order 1

\[
\begin{pmatrix}
X_1(t) \\
X_2(t)
\end{pmatrix} =
\begin{pmatrix}
-0.5 & 0 \\
0.7 & 0.2
\end{pmatrix}
\begin{pmatrix}
X_1(t-1) \\
X_2(t-1)
\end{pmatrix} +
\begin{pmatrix}
E_1(t) \\
E_2(t)
\end{pmatrix}
\]

\[
X_1(t) = -0.5X_1(t-1) + 0X_2(t-1) + E_1(t)
\]

\[
X_2(t) = 0.7X_1(t-1) + 0.2X_2(t-1) + E_2(t)
\]

Which causal structure does this model correspond to?

- a) 1 → 2
- b) 1 ← 2
- c) 1 ↔ 2

Saturday, September 24, 2011
Granger Causality Quiz

- Example: 2-channel MVAR process of order 1

\[
\begin{pmatrix}
X_1(t) \\
X_2(t)
\end{pmatrix}
= \begin{pmatrix}
-0.5 & 0 \\
0.7 & 0.2
\end{pmatrix}
\begin{pmatrix}
X_1(t-1) \\
X_2(t-1)
\end{pmatrix}
+ \begin{pmatrix}
E_1(t) \\
E_2(t)
\end{pmatrix}
\]

\[
X_1(t) = -0.5X_1(t-1) + 0X_2(t-1) + E_1(t)
\]
\[
X_2(t) = 0.7X_1(t-1) + 0.2X_2(t-1) + E_2(t)
\]

Which causal structure does this model correspond to?

a) 1 → 2  
b) 1 ← 2  
c) 1 ←→ 2
Granger Causality – Frequency Domain

\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]
Granger Causality – Frequency Domain

\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]

Fourier-transforming \( A^{(k)} \) we obtain

\[ A(f) = \sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk}; A^{(0)} = I \]

Likewise, \( X(f) \) and \( E(f) \) correspond to the fourier transforms of the data and residuals, respectively.
Granger Causality – Frequency Domain

\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t-k) + E(t) \]

Fourier-transforming \( A^{(k)} \) we obtain

\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; A^{(0)} = I \]

We can then define the spectral matrix \( X(f) \) as follows:

\[ X(f) = A(f)^{-1} E(f) = H(f) E(f) \]

Where \( H(f) \) is the transfer matrix of the system.
Granger Causality – Frequency Domain

\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]

Fourier-transforming \( A^{(k)} \) we obtain

\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk}; A^{(0)} = I \]

We can then define the spectral matrix \( X(f) \) as follows:

\[ X(f) = A(f)^{-1} E(f) = H(f)E(f) \]

Where \( H(f) \) is the transfer matrix of the system.

**Definition 2**

\( X_i \) granger-causes \( X_i \) conditioned on all other variables in \( X \)

if and only if \( |A_{ij}(f)| >> 0 \) for some frequency \( f \)
Granger Causality – Frequency Domain

\[ X(t) = \sum_{k=1}^{p} A^{(k)}X(t-k) + E(t) \]

Fourier-transforming \( A^{(k)} \) we obtain

\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk}; A^{(0)} = I \]

We can then define the spectral matrix \( X(f) \) as follows:

\[ X(f) = A(f)^{-1}E(f) = H(f)E(f) \]

Where \( H(f) \) is the transfer matrix of the system.

---

**Definition 2**

\( X_i \) granger-causes \( X_i \) conditioned on all other variables in \( X \) if and only if \( |A_{ij}(f)| \gg 0 \) for some frequency \( f \)
<table>
<thead>
<tr>
<th>$x_1(t)$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_2(t)$</td>
<td></td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$x_M(t)$</td>
<td></td>
</tr>
</tbody>
</table>
\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]
\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]
\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; \quad A^{(0)} = I \]
\[ X(f) = A(f)^{-1} E(f) = H(f)E(f) \]
\[
X(t) = \sum_{k=1}^{p} A^{(k)} X(t-k) + E(t)
\]

\[
A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; \quad A^{(0)} = I
\]

\[
X(f) = A(f)^{-1} E(f) = H(f) E(f)
\]

Ground Truth

\[\begin{array}{c}
1 \\
2 \\
3 \\
4 \\
5 \\
6 \\
7 \\
\end{array}\]

Kus, 2004
A. Surrogate Data

KUS

\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]

\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; \quad A^{(0)} = I \]

\[ X(f) = A(f)^{-1} E(f) = H(f)E(f) \]

Fig. 2. Pair-wise coherences and resulting flow scheme for simulation I.

- Dotted arrows: false flows found by the applied method.
- Spurious indirect true flow
- Direct true flow

Effective

Ground Truth

Kus, 2004

\[ \sum_{i=2}^{M} + \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]

\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; \quad A^{(0)} = I \]

\[ X(f) = A(f)^{-1} E(f) = H(f)E(f) \]

Spurious indirect true flow
Direct true flow

Functional
\( X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \)
\[
A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i 2\pi f k}; \quad A^{(0)} = I
\]
\( X(f) = A(f)^{-1} E(f) = H(f)E(f) \)

The signal from channel 1 was propagated to channel 3 through a pattern of flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows. In the case of pair-wise estimates, we have high causality measure calculated pair-wise. Again, we obtain too many flows.
\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]
\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk}; \quad A^{(0)} = I \]
\[ X(f) = A(f)^{-1} E(f) = H(f) E(f) \]

\[ S(f) = X(f)X(f)^* = H(f)\Sigma H(f)^* \]

\[ C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \]

(Strela, 2001)
The datasets generated in such a way should have the same statistical properties as the original data. This is obtained by transforming the randomized data back to the time domain. The datasets obtained in this way have the same dimensions as the original data and represent the same information content. In order to test the validity of the procedure of surrogate data generation, one can compare the statistics of the original data with those of the surrogate data. This can be done by calculating the mean and variance of the power spectra of the original and surrogate data. If the statistics of the original and surrogate data are the same, then the procedure of surrogate data generation is valid. This is because the surrogate data are generated from the original data in such a way that they have the same statistical properties as the original data. Therefore, if the statistics of the original and surrogate data are the same, then the procedure of surrogate data generation is valid. If the statistics of the original and surrogate data are different, then the procedure of surrogate data generation is invalid. This is because the surrogate data are not generated from the original data in such a way that they have the same statistical properties as the original data. Therefore, if the statistics of the original and surrogate data are different, then the procedure of surrogate data generation is invalid.
\[ X(t) = \sum_{k=1}^{p} A^{(k)} X(t - k) + E(t) \]
\[ A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; \quad A^{(0)} = I \]
\[ X(f) = A(f)^{-1} E(f) = H(f)E(f) \]

\[
S(f) = X(f)X(f)^* \\
= H(f)\Sigma H(f)^* \\
\text{(Brillinger, 2001)}
\]

\[
C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \\
\text{(Brillinger, 2001)}
\]

\[
P_{ij}(f) = \frac{S_{ij}^{-1}(f)}{\sqrt{S_{ii}^{-1}(f)S_{jj}^{-1}(f)}} \\
\text{(Brillinger, 2001)}
\]

\[
G_{ij}(f) = \sqrt{1 - \frac{\det(S(f))}{S_{ii}(f)M_{ii}(f)}} \\
\text{(Brillinger, 2001)}
\]
X(t) = \sum_{k=1}^{p} A^{(k)} X(t-k) + E(t) \\
A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; A^{(0)} = I \\
X(f) = A(f)^{-1} E(f) = H(f) E(f) 

\[ S(f) = X(f)X(f)^* = H(f) \Sigma H(f)^* \] (Brillinger, 2001)

\[ C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \] (Brillinger, 2001)

\[ P_{ij}(f) = \frac{S_{ij}^{-1}(f)}{\sqrt{S_{ii}^{-1}(f)S_{jj}^{-1}(f)}} \] (Brillinger, 2001)

\[ G_{ij}(f) = \sqrt{1 - \frac{\det(S(f))}{S_{ii}(f)M_{ii}(f)}} \] (Brillinger, 2001)

Granger-Geweke Causality

\[ F_{ij}(f) = \frac{\Sigma_{ij} - (\Sigma_{ij}^2 / \Sigma_{ii})}{\Sigma_{ii}(f)} | H_{ij}(f) |^2 \] (Geweke, 1982; Bressler et al., 2007)

\[ \text{Functional} \quad \text{Effective} \]

Ground Truth

spurious indirect true flow direct true flow
The nonnormalized DTFs (equivalent to the multivariate
power spectral density) of activity.

This kind of normalization is better when the number of electrodes
is small.

According to our experience, the use of dDTF may be
more appropriate in practice.

The results obtained by means of dDTF for the simulation
are of the order of magnitude of accuracy determined from the surrogate data
but it could occur for electrodes implanted in
specific brain structures.

To avoid the situation described above, which is unlikely for sur-
rogate data, we have chosen values corresponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-
sponding to the frequency of the maximum of amplitude spec-
tum of coherences, we have chosen values corre-

\[
\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)} \mathbf{X}(t-k) + \mathbf{E}(t)
\]

\[
\mathbf{A}(f) = -\sum_{k=0}^{p} \mathbf{A}^{(k)} e^{-i2\pi f k} \quad ; \quad \mathbf{A}^{(0)} = \mathbf{I}
\]

\[
\mathbf{X}(f) = \mathbf{A}(f)^{-1} \mathbf{E}(f) = \mathbf{H}(f) \mathbf{E}(f)
\]
The functional spectrum for each channel is defined as:

\[ S(f) = X(f)X(f)^* = H(f)\Sigma H(f)^* \]

(From Brillinger, 2001)

The coherency between two channels is given by:

\[ C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \]

(From Brillinger, 2001)

Partial coherence is defined as:

\[ P_{ij}(f) = \frac{S_{ij}^{-1}(f)}{\sqrt{S_{ii}^{-1}(f)S_{jj}^{-1}(f)}} \]

(From Brillinger, 2001)

Multiple coherence is calculated as:

\[ G_i(f) = \sqrt{1 - \frac{\det(S(f))}{S_{ii}(f)M_{ii}(f)}} \]

(From Brillinger, 2001)

Effective spectrum of coherences is defined in terms of Granger causality:

\[ F_{ij}(f) = \frac{\sum_{j} - (\sum_{j}^2 / \sum_{j}) |H_{ij}(f)|^2}{S_{ii}(f)} \]

(Geweke, 1982; Bressler et al., 2007)

Directed transfer function is given by:

\[ \eta_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{j} \sum_{k=1}^M |H_{jk}(f)|^2} \]

(Kaminski and Blinowska, 1991)

Partial directed coherence is defined as:

\[ \pi_{ij}^2(f) = \frac{|A_{ij}(f)|^2}{\sum_{k=1}^M |A_{kj}(f)|^2} \]

(Baccalà and Sameshima, 2001)

The equation for the system's output is:

\[ X(t) = \sum_{k=1}^p A^{(k)}X(t-k) + E(t) \]

\[ A(f) = -\sum_{k=0}^p A^{(k)} e^{-i2\pi fk} ; \ A^{(0)} = I \]

\[ X(f) = A(f)^{-1}E(f) = H(f)E(f) \]

The ground truth diagram illustrates the flow patterns and interactions between different channels, with arrows indicating the direction of information transfer.
\[
\begin{align*}
    X(t) &= \sum_{k=1}^{p} A^{(k)} X(t-k) + E(t) \\
    A(f) &= -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk} ; \quad A^{(0)} = I \\
    X(f) &= A(f)^{-1} E(f) = H(f) E(f)
\end{align*}
\]
\[
X(t) = \sum_{k=1}^{p} A^{(k)} X(t-k) + E(t)
\]
\[
A(f) = -\sum_{k=0}^{p} A^{(k)} e^{-i2\pi fk}; \quad A^{(0)} = I
\]
\[
X(f) = A(f)^{-1}E(f) = H(f)E(f)
\]

**Functional Coherence Spectrum**

\[
S(f) = X(f)X(f)^* = H(f)\Sigma H(f)^*
\]
(Brillinger, 2001)

\[
C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}
\]
(Brillinger, 2001)

**Partial Coherence**

\[
P_{ij}(f) = \frac{S_{ij}^{-1}(f)}{\sqrt{S_{ii}^{-1}(f)S_{jj}^{-1}(f)}}
\]
(Brillinger, 2001)

**Multiple Coherence**

\[
G_i(f) = \sqrt{1 - \frac{\det(S(f))}{S_{ii}(f)M_{ii}(f)}}
\]
(Brillinger, 2001)

**Effective Directed Transfer Function**

\[
F_{ij}(f) = \frac{\Sigma_{ij} - (\Sigma_{ij}^2 / \Sigma_{ii}) |H_{ij}(f)|^2}{S_{ii}(f)}
\]
(Geweke, 1982; Bressler et al., 2007)

**Directed Transfer Function**

\[
\eta_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\Sigma_f \Sigma_k^M |H_{ik}(f)|^2}
\]
(Kaminski and Blinowska, 1991)

**Partial Directed Coherence**

\[
\pi_{ij}^2(f) = \frac{|A_{ij}(f)|^2}{\Sigma_k^M |A_{kj}(f)|^2}
\]
(Baccala and Sameshima, 2001)

**Direct DTF**

\[
\delta_{ij}^2(f) = \eta_{ij}^2(f)P_{ij}^2(f)
\]
(Korzeniewska, 2003)
PDC versus DTF methods (spectral considerations)
Time-Frequency GC
Time-Frequency GC

- Brain network dynamics often change rapidly with time (non-stationarity)
  - event-related responses
  - transient network changes during information processing
Time-Frequency GC

- Brain network dynamics often change rapidly with time (non-stationarity)
  - event-related responses
  - transient network changes during information processing
- How can we perform time-varying, frequency-domain analysis of network dynamics?
Time-Frequency GC
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
  - Short-Time adaptive multivariate autoregression (AMVAR)
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
  - Short-Time adaptive multivariate autoregression (AMVAR)
  - Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
  - Short-Time adaptive multivariate autoregression (AMVAR)
  - Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
  - Kalman Filtering
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
  - Short-Time adaptive multivariate autoregression (AMVAR)
  - Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
  - Kalman Filtering
  - ...

Saturday, September 24, 2011
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
  - Short-Time adaptive multivariate autoregression (AMVAR)
  - Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
  - Kalman Filtering
  - ...

Saturday, September 24, 2011
Short-Window Time-Frequency GC

(Ding et al, 2000)

Analogous to short-time Fourier transform

\[ X(t) \]

trials

sources
Short-Window Time-Frequency GC

(Ding et al, 2000)

Analogous to short-time Fourier transform
Short-Window Time-Frequency GC

(Ding et al, 2000)

Analogous to short-time Fourier transform

\[
X(t) = \sum_{k=1}^{p} A^{(k)}(t) X(t-k) + E(t)
\]

\[
A(f,t) = -\sum_{k=0}^{p} A^{(k)}(t) e^{-i2\pi fk}; A^{(0)} = I
\]
Short-Window Time-Frequency GC

(Ding et al, 2000)

Analogous to short-time Fourier transform

\[ X(t) = \sum_{k=1}^{p} A^{(k)}(t)X(t - k) + E(t) \]

\[ A(f, t) = -\sum_{k=0}^{p} A^{(k)}(t)e^{-i2\pi fk}; A^{(0)} = I \]
Short-Window Time-Frequency GC

(Ding et al, 2000)

Analogous to short-time Fourier transform

\[ X(t) = \sum_{k=1}^{p} A^{(k)}(t) X(t - k) + E(t) \]

\[ A(f,t) = -\sum_{k=0}^{p} A^{(k)}(t)e^{-i2\pi fk}; A^{(0)} = I \]

From sources to trials

ensemble normalization

VAR → GC → TO
Short-Window Time-Frequency GC

\[ X(t) = \sum_{k=1}^{p} A^{(k)}(t)X(t-k) + E(t) \]

\[ A(f,t) = -\sum_{k=0}^{p} A^{(k)}(t)e^{-i2\pi fk}; A^{(0)} = I \]

Analogous to short-time Fourier transform

(Ding et al, 2000)
Time-Frequency GC
Time-Frequency GC

• What is a good window length?
Time-Frequency GC

- What is a good window length?
- Considerations:
Time-Frequency GC

- What is a good window length?
- Considerations:
  - Temporal smoothing
Time-Frequency GC

- What is a good window length?
- Considerations:
  - Temporal smoothing
  - Local stationarity
Time-Frequency GC

- **What is a good window length?**
- Considerations:
  - Temporal smoothing
  - Local stationarity
  - Sufficient amount of data
Time-Frequency GC

- What is a good window length?
- Considerations:
  - Temporal smoothing
  - Local stationarity
  - Sufficient amount of data
  - Process dynamics
Time-Frequency GC

Consideration: Temporal Smoothness

Ding et al., 2000
Time-Frequency GC

Consideration: Temporal Smoothness

Too-large windows may smooth out interesting transient dynamic features.

Ding et al, 2000
Time-Frequency GC

Consideration: Local Stationarity
Time-Frequency GC

Consideration: Local Stationarity

Too-large windows may not be locally-stationary
Time-Frequency GC

Consideration: Local Stationarity
Time-Frequency GC
Time-Frequency GC

Consideration: Sufficient data
Time-Frequency GC

Consideration: Sufficient data

$M = \text{number of variables}$
Consideration: Sufficient data

\( M = \) number of variables
\( p = \) model order
Time-Frequency GC

Consideration: Sufficient data

M = number of variables
p = model order
N_{tr} = number of trials
Time-Frequency GC

Consideration: Sufficient data

M = number of variables
p = model order
N_{tr} = number of trials
W = length of each window (sample points)
Time-Frequency GC

Consideration: Sufficient data

\[ M = \text{number of variables} \]
\[ p = \text{model order} \]
\[ N_{tr} = \text{number of trials} \]
\[ W = \text{length of each window (sample points)} \]
**Consideration: Sufficient data**

\[ M = \text{number of variables} \]
\[ p = \text{model order} \]
\[ N_{tr} = \text{number of trials} \]
\[ W = \text{length of each window (sample points)} \]

We have \( M^2 p \) model coefficients to estimate. This requires a minimum of \( M^2 p \) independent samples.
Consideration: Sufficient data

\( M = \text{number of variables} \)
\( p = \text{model order} \)
\( N_{tr} = \text{number of trials} \)
\( W = \text{length of each window (sample points)} \)

We have \( M^2p \) model coefficients to estimate. This requires a minimum of \( M^2p \) independent samples.
So we have the constraint \( M^2p \leq N_{tr} W \).
Consideration: Sufficient data

\[ M = \text{number of variables} \]
\[ p = \text{model order} \]
\[ N_{tr} = \text{number of trials} \]
\[ W = \text{length of each window (sample points)} \]

We have \( M^2p \) model coefficients to estimate. This requires a minimum of \( M^2p \) independent samples. So we have the constraint \( M^2p \leq N_{tr}W \). In practice, however, a better heuristic is \( M^2p \leq (1/10)N_{tr}W \).
Consideration: Sufficient data

$M = \text{number of variables}$

$p = \text{model order}$

$N_{tr} = \text{number of trials}$

$W = \text{length of each window (sample points)}$

We have $M^2p$ model coefficients to estimate. This requires a minimum of $M^2p$ independent samples.

So we have the constraint $M^2p \leq N_{tr} W$.

In practice, however, a better heuristic is $M^2p \leq (1/10)N_{tr} W$.

Or: $W \geq 10(M^2p/N_{tr})$
Consideration: Sufficient data

\( M \) = number of variables
\( p \) = model order
\( N_{tr} \) = number of trials
\( W \) = length of each window (sample points)

We have \( M^2 p \) model coefficients to estimate. This requires a minimum of \( M^2 p \) independent samples.

So we have the constraint \( M^2 p \leq N_{tr} W \).

In practice, however, a better heuristic is \( M^2 p \leq (1/10)N_{tr} W \).

Or: \( W \geq 10(M^2 p/N_{tr}) \)

SIFT will let you know if your window length is not optimal.
Time-Frequency GC
Time-Frequency GC

Consideration: Process dynamics
Consideration: Process dynamics

- Your window must be larger than the maximum expected interaction time lag between any two processes.
Time-Frequency GC

Consideration: Process dynamics

• Your window must be larger than the maximum expected interaction time lag between any two processes.
• Your window should be large enough to span ~1 cycle of the lowest frequency of interest (remember the Heisenberg uncertainty principle)
Time-Frequency GC

Consideration: Process dynamics

• Your window must be larger than the maximum expected interaction time lag between any two processes.
• Your window should be large enough to span ~1 cycle of the lowest frequency of interest (remember the Heisenberg uncertainty principle)
Time-Frequency GC

**Consideration: Process dynamics**

- Your window must be larger than the maximum expected interaction time lag between any two processes.
- Your window should be large enough to span ~1 cycle of the lowest frequency of interest (remember the Heisenberg uncertainty principle)
Time-Frequency GC

- Many ways to do time-varying MVAR estimation
  - Short-Time adaptive multivariate autoregression (AMVAR)
  - Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
  - Kalman Filtering
  - ...
The State-Space Model

- Based on rich dynamical systems theory.
- Well-established state-space algorithms for tracking in non-stationary, high-dimensional, partially-observed, noisy systems
- Easily extendable to nonlinear systems
- Allows for the additional modeling of (known or inferred) exogenous inputs
- Allows for estimation of additional unknown sources (as additional states)
The State-Space Model

- Based on rich dynamical systems theory.
- Well-established state-space algorithms for tracking in non-stationary, high-dimensional, partially-observed, noisy systems
- Easily extendable to nonlinear systems
- Allows for the additional modeling of (known or inferred) exogenous inputs
- Allows for estimation of additional unknown sources (as additional states)
The State-Space Model

- Based on rich dynamical systems theory.
- Well-established state-space algorithms for tracking in non-stationary, high-dimensional, partially-observed, noisy systems
- Easily extendable to nonlinear systems
- Allows for the additional modeling of (known or inferred) exogenous inputs
- Allows for estimation of additional unknown sources (as additional states)
The State-Space Model
The State-Space Model

\[
z(t) = \text{vec}\left([A^{(1)}(t), \ldots, A^{(p)}(t)]^T\right)_{[M^2 p \times 1]} \\
y(t) = X(t) \\
H(t) = I_M \otimes \text{vec}\left(\begin{bmatrix} X(t-1) & \ldots & X(t-p) \end{bmatrix}^T\right)^T_{[1 \times Mp]}
\]

unknown VAR parameters
The State-Space Model

\[ z(t) = \text{vec} \left( [A^{(1)}(t), \ldots, A^{(p)}(t)]^T \right)_{[M^2 p \times 1]} \]

\[ y(t) = X(t) \]

\[ H(t) = I_M \otimes \text{vec} \left( \begin{bmatrix} X(t-1) & \ldots & X(t-p) \end{bmatrix}^T \right)_{[1 \times M p]} \]

unknown VAR parameters
The State-Space Model

\[ z(t) = vec \left( [A^{(1)}(t), \ldots, A^{(p)}(t)]^T \right)_{[M^2p \times 1]} \]

\[ y(t) = X(t) \]

\[ H(t) = I_M \otimes vec \left( \begin{bmatrix} X(t-1) & \ldots & X(t-p) \end{bmatrix}^T \right)_{[1 \times M^2p]} \]

\[ z(t) = z(t-1) + \nu(t) \]

unknown VAR parameters

state transition equation (random walk)
The State-Space Model

State-Space Model

\[ z(t) = \text{vec} \left( [A^{(1)}(t), \ldots, A^{(p)}(t)]^T \right) \]
\[ y(t) = X(t) \]
\[ H(t) = I_M \otimes \text{vec} \left( \begin{bmatrix} X(t-1) & \ldots & X(t-p) \end{bmatrix}^T \right) \]

**state transition equation (random walk)**

**unknown VAR parameters**

**observation equation (VAR model)**

\[ z(t) = z(t-1) + v(t) \]
\[ y(t) = H(t)z(t) + \epsilon(t) \]
The State-Space Model

\[ z(t) = vec \left( [A(1)(t), \ldots, A(p)(t)]^T \right) \]
\[ y(t) = X(t) \]
\[ H(t) = I_M \otimes vec \left( \begin{bmatrix} X(t-1) & \ldots & X(t-p) \end{bmatrix} \right)^T \]

\[ z(t) = z(t-1) + v(t) \]
\[ y(t) = H(t)z(t) + \epsilon(t) \]

- How do we solve for the time-varying unknown states?
- Kalman Filtering (and extensions)
Kalman Filtering

GPDC Causality From

Causality To

Frequency (Hz)

Time (sec)
Scalp or Source?

Fig. 15.13: Direction of flows for 21-channel EEG (awake state eyes closed) obtained by means of different methods. The shade of gray of the arrow represents the strength of the connection (black = the strongest), for each method 40 strongest flows are shown. Reprinted from with permission [49] (© IEEE 2005).

A lot of activity flowing to the destination channels from the posterior electrodes, so the denominator in Eq. (15.6) is quite large, which diminishes the values of DTFs showing outflows from Fz. For Granger causality and DTF there is no propagation from the temporal electrodes. This is practically also the case for dDTF. The dDTF shows only direct flows, we can see that in this case the pattern of flows is consistent with anatomy, e.g., a lack of direct connection between Oz and Pz, Fz, and Fpz—locations where hemispheres are partitioned. The main sources of the activity—namely, electrodes P3, P4, O2, Oz, O1—are the same as for the other multivariate estimates.

Inspecting the results of application of the PDC function to the same data epoch we observe a different picture. One can notice that, unlike the results of dDTF, some channels became sinks. This is due to the normalization of PDC. In fact, we do not see the transmission, as is the case for dDTF, but the ratio between the flow to a given channel with respect to all the outflows from the considered channel. In this way, a channel propagating activity in all directions will show weaker flows than a channel propagating only in one direction. Therefore, the method is not suitable for identification of sources of EEG activity, but it may be useful when the destination channel is of primary interest.
Channel or Source?
Channel or Source?

sensors

Volume Conduction

Saturday, September 24, 2011
Channel or Source?

sensors

Volume Conduction

sources
Channel or Source?
Channel or Source?

Volume Conduction

sensors

sources
Channel or Source?

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t)S(t-k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) \]

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t)S(t - k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t - k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t - k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

Volume Conduction

\[ X(t) = MS(t) \]

sensors

sources

Solution? Source Separation

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t - k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

Volume Conduction

ICA
Sparse Bayesian Learning
Beamforming
Minimum-norm
...

Sources
sensors

Solution? Source Separation

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t - k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

\[ X(t) = MS(t) \]

Volume Conduction

ICA
Sparse Bayesian Learning
Beamforming
Minimum-norm
...

Solution? Source Separation

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t)S(t - k) + E(t) \]
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t-k) + ME(t) \]

\[ X(t) = MS(t) \]

Volume Conduction

ICA
Sparse Bayesian Learning
Beamforming
Minimum-norm...

Solution? Source Separation

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t-k) + E(t) \]

Saturday, September 24, 2011
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

\[ X(t) = \frac{M^{-1}}{MS(t)} \]

Volume Conduction

ICA
Sparse Bayesian Learning
Beamforming
Minimum-norm...

Solution? Source Separation

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t - k) + E(t) \]

Saturday, September 24, 2011
Channel or Source?

\[ X(t) = MS(t) = \sum_{k=1}^{p} MA^{(k)}(t) M^{-1} X(t - k) + ME(t) \]

Solution? Source Separation

\[ S(t) = \sum_{k=1}^{p} A^{(k)}(t) S(t - k) + E(t) \]
Forward/Inverse Modeling

\[ X(t) = MS(t) \]

anatomically-realistic

Akalin Acar

Saturday, September 24, 2011
Forward/Inverse Modeling

Forward Modeling

\[ X(t) = MS(t) \]

Inverse Modeling

\[ S(t) = M^{-1} X(t) \]

anatomically-realistic

Akalin Acar

EEG/ECoG
Forward/Inverse Modeling

Forward Modeling:

\[ X(t) = MS(t) \]

Inverse Modeling:

\[ \hat{S}(t) = M^{-1}X(t) \]

anatomically-realistic

ill-posed!

EEG/
ECoG

Akalin Acar

Saturday, September 24, 2011
Forward/Inverse Modeling

Forward Modeling

\[ X(t) = MS(t) \]

Inverse Modeling

\[ \hat{S}(t) = M^{-1} X(t) \]

anatomically-realistic

ill-posed!

solutions?
Forward/Inverse Modeling

Forward Modeling

\[ X(t) = MS(t) \]

Inverse Modeling

\[ \hat{S}(t) = M^{-1}X(t) \]

anatomically-realistic

EEG/ECOG

ill-posed!

impose constraints!

solutions?

Akalin Acar

Saturday, September 24, 2011
Forward/Inverse Modeling

\[ S(t) \rightarrow \text{Forward Modeling} \rightarrow X(t) = MS(t) \rightarrow \text{Inverse Modeling} \rightarrow \hat{S}(t) = M^{-1}X(t) \]

- \text{anatomically-realistic}
- \text{ill-posed!}
- \text{anatomy sparsity independence ...}
- impose constraints!
- solutions?

Forward Modeling

\[ \text{EEG/ECoG} \]
Forward/Inverse Modeling

\[ X(t) = MS(t) \]

\[ \hat{S}(t) = M^{-1} X(t) \]

A Recipe for Reducing Errors:
- Anatomically Realistic Forward Model
- Appropriately Constrained Inverse Model

Akalin Acar and Makeig, 2010
Estimating Dependency of Independent Components ?
Estimating Dependency of Independent Components?

Isn’t it a contradiction to examine dependence between Independent Components?
Estimating Dependency of Independent Components?

- Isn’t it a contradiction to examine dependence between Independent Components?

- Instantaneous (e.g., Infomax) ICA only explicitly enforces *instantaneous* independence. Time-delayed dependencies may be preserved.
Estimating Dependency of Independent Components?

- Isn’t it a contradiction to examine dependence between Independent Components?
- Instantaneous (e.g., Infomax) ICA only explicitly enforces *instantaneous* independence. Time-delayed dependencies may be preserved.
- ICA seeks to maximize *global* independence (over entire recording session), transient dependencies are often preserved.
Estimating Dependency of Independent Components?

**Connectivity Error**

<table>
<thead>
<tr>
<th>N0</th>
<th>N1</th>
<th>N2</th>
<th>N3</th>
<th>N4</th>
<th>N5</th>
<th>N6</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCSA_EM</td>
<td>SCSA</td>
<td>CSA</td>
<td>CICAAR</td>
<td>MVARICA</td>
<td>ICA</td>
<td>SCSA_EM</td>
</tr>
<tr>
<td>SCSA</td>
<td>SCSA</td>
<td>CSA</td>
<td>CICAAR</td>
<td>MVARICA</td>
<td>ICA</td>
<td>SCSA</td>
</tr>
<tr>
<td>N4</td>
<td>N5</td>
<td>N6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCSA_EM</td>
<td>SCSA</td>
<td>CSA</td>
<td>CICAAR</td>
<td>MVARICA</td>
<td>ICA</td>
<td>SCSA</td>
</tr>
<tr>
<td>SCSA</td>
<td>SCSA</td>
<td>CSA</td>
<td>CICAAR</td>
<td>MVARICA</td>
<td>ICA</td>
<td>SCSA</td>
</tr>
</tbody>
</table>

Haufe et al, 2008
Source Information Flow Toolbox (SIFT)

Mullen, Delorme, Kothe, Makeig. Society for Neuroscience, 2010
Delorme, Mullen, Kothe et al, Computational Intelligence and Neuroscience, vol 12, 2011
EEGLAB Software framework

Analysis

Analysis plugins

Data archive

Data sync and handling

Interactive tools

Stimulus control

28 user plugins

HeadIT

EEGLAB

NFT toolbox

SIFT

BCILAB

ERICA framework

EyeTracker

Wii remote

Mocap

EEG

Tactile stream

Video stream

Audio stream

Delorme, Mullen, Kothe, Akalin Acar, Bigdely-Shamlo, Vankov, Makeig, *Computational Intelligence and Neuroscience*, vol 12, 2011

Saturday, September 24, 2011
Source Information Flow Toolbox (SIFT)
Source Information Flow Toolbox (SIFT)

- A new (alpha) toolbox for source-space electrophysiological information flow and causality analysis (single-subject or group analysis) integrated into the EEGLAB software environment
Source Information Flow Toolbox (SIFT)

- A new (alpha) toolbox for source-space electrophysiological information flow and causality analysis (single-subject or group analysis) integrated into the EEGLAB software environment
- Modular architecture intended to support multiple modeling approaches
Source Information Flow Toolbox (SIFT)

- A new (alpha) toolbox for source-space electrophysiological information flow and causality analysis (single-subject or group analysis) integrated into the EEGLAB software environment
- Modular architecture intended to support multiple modeling approaches
- Emphasis on vector autoregression and time-frequency domain approaches
Source Information Flow Toolbox (SIFT)

- A new (alpha) toolbox for source-space electrophysiological information flow and causality analysis (single-subject or group analysis) integrated into the EEGLAB software environment
- Modular architecture intended to support multiple modeling approaches
- Emphasis on vector autoregression and time-frequency domain approaches
- Standard and novel interactive visualization methods for exploratory analysis of connectivity across time, frequency, and spatial location
Source Information Flow Toolbox (SIFT)

- A new (alpha) toolbox for source-space electrophysiological information flow and causality analysis (single-subject or group analysis) integrated into the EEGLAB software environment
- Modular architecture intended to support multiple modeling approaches
- Emphasis on vector autoregression and time-frequency domain approaches
- Standard and novel interactive visualization methods for exploratory analysis of connectivity across time, frequency, and spatial location

**Requirements:** EEGLAB, MATLAB™ 2008b, Signal Processing Toolbox, Statistics Toolbox (for some functions -- may be removed in the future)
#1: Button press epochs

Filename: .../Data/bt73 RestWrong.set
Channels per frame  127
Frames per epoch  1024
Epochs  165
Events  1451
Sampling rate (Hz)  256
Epoch start (sec)  -2.000
Epoch end (sec)  1.996
Reference  unknown
Channel locations  Yes
ICA weights  Yes
Dataset size (Mb)  175.3
#1: Button press epochs

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filename: .../eta/Data/bt3 RestWrong.set</td>
<td></td>
</tr>
<tr>
<td>Channels per frame</td>
<td>127</td>
</tr>
<tr>
<td>Frames per epoch</td>
<td>1024</td>
</tr>
<tr>
<td>Epochs</td>
<td>165</td>
</tr>
<tr>
<td>Events</td>
<td>1451</td>
</tr>
<tr>
<td>Sampling rate (Hz)</td>
<td>256</td>
</tr>
<tr>
<td>Epoch start (sec)</td>
<td>-2.000</td>
</tr>
<tr>
<td>Epoch end (sec)</td>
<td>1.996</td>
</tr>
<tr>
<td>Reference</td>
<td>unknown</td>
</tr>
<tr>
<td>Channel locations</td>
<td>Yes</td>
</tr>
<tr>
<td>ICA weights</td>
<td>Yes</td>
</tr>
<tr>
<td>Dataset size (Mb)</td>
<td>175.3</td>
</tr>
</tbody>
</table>
#1: Button press epochs

Filename: ...eta/Data/bt73 RespWrong.set
Channels per frame  127
Frames per epoch  1024
Epochs  165
Events  1451
Sampling rate (Hz)  256
Epoch start (sec)  -2.000
Epoch end (sec)  1.996
Reference  unknown
Channel locations  Yes
ICA weights  Yes
Dataset size (Mb)  175.3
- Source-separation and localization
  (performed externally using EEGLAB or other toolboxes)

- Filtering/Detrending

- Downsampling

- Differencing

- Normalization (temporal or ensemble)

- Trial balancing

- Tests for stationarity of the data (linear methods)
<table>
<thead>
<tr>
<th></th>
<th>Linear</th>
<th>Nonlinear</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parametric</strong></td>
<td>MVAR Modeling, Sparse MVAR, Linear Kalman Filtering</td>
<td>Extended/Cubature Kalman Filtering</td>
</tr>
<tr>
<td><strong>Nonparametric</strong></td>
<td>Nonparametric MVAR (minimum-phase spectral factorization)</td>
<td>Transfer Entropy</td>
</tr>
<tr>
<td></td>
<td>Multivariate phase distribution</td>
<td></td>
</tr>
</tbody>
</table>

- **fully implemented**
- **partially-developed**
- **coming soon**
Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

Pre-processing

Model fitting and validation

Connectivity

Statistics

Visualization

Fit AMVAR Model

Start Window Length Assistant...

Start Model Order Assistant...

Help

Cancel

Ok

1. Select MVAR algorithm
   - Vierra-Morf
   - ARFIT

2. Window length (sec) 0.5

3. Step size (sec) 0.03

4. Model order 10
Preprocessing
Modeling
Statistics
Visualization
Model Fitting
Validation
Connectivity

Pre-processing
Model fitting and validation
Connectivity
Statistics
Visualization

Fit AMVAR Model

1. Select MVAR algorithm
   - Vierra-Morl
   - ARFIT
2. Window length (sec) 0.5
   - Start Window Length Assistant...
3. Step size (sec) 0.03
4. Model order 10
   - Start Model Order Assistant...

Plot Information Criteria

Select order criteria to estimate
(hold Ctrl to select multiple)
AIC, FPE, SBC, HQ

- Do not date model
- Model order range: 1 - 30
- % windows to sample: 10%
- Whiteness of Residuals
  - Portmanteau tests
  - Autocorrelation function
- Model Consistency
- Model Stability
<table>
<thead>
<tr>
<th>VAR</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Power spectrum (ERSP)</td>
<td>- Transfer Entropy *</td>
</tr>
<tr>
<td>- Coherence (Coh), Partial Coherence (pCoh), Multiple Coherence (mCoh)</td>
<td>- Multivariate phase-locking value (mPLV) *</td>
</tr>
<tr>
<td>- Partial Directed Coherence (PDC)</td>
<td></td>
</tr>
<tr>
<td>- Generalized PDC (GPDC)</td>
<td></td>
</tr>
<tr>
<td>- Partial Directed Coherence Factor (PDCF)</td>
<td></td>
</tr>
<tr>
<td>- Renormalized PDC (rPDC) *</td>
<td></td>
</tr>
<tr>
<td>- Directed Transfer Function (DTF)</td>
<td></td>
</tr>
<tr>
<td>- Direct Directed Transfer Function (dDTF)</td>
<td></td>
</tr>
<tr>
<td>- Granger-Geweke Causality (GGC)</td>
<td></td>
</tr>
<tr>
<td>- Conditional GGC</td>
<td></td>
</tr>
<tr>
<td>- Blockwise GGC *</td>
<td></td>
</tr>
</tbody>
</table>

**Fully implemented**

**Partially-developed**

**Coming soon**
Asymptotic analytic estimates of confidence intervals
Applies to: PDC, nPDC, DTF, nDTF, rPDC
Tests: $H_{\text{null}}$, $H_{\text{base}}$, $H_{AB}$

Confidence intervals using thin-plate smoothing splines
Applies to: dDTF
Tests: $H_{\text{base}}$, $H_{AB}$

$H_{\text{null}}$: $C_{ij} \leq C_{null}$

$H_{\text{base}}$: $C_{ij} \leq C_{\text{baseline}}$

$H_{AB}$: $C_{Aij} = C_{Bij}$

- **fully implemented**
- **partially-developed**
- **coming soon**
Asymptotic analytic estimates of confidence intervals
Applies to: PDC, nPDC, DTF, nDTF, rPDC
Tests: $H_{null}$, $H_{base}$, $H_{AB}$

Confidence intervals using thin-plate smoothing splines
Applies to: dDTF
Tests: $H_{base}$, $H_{AB}$

Phase-randomization
Applies to: all
Tests: $H_{null}$

Permutation Tests
Applies to: all
Tests: $H_{AB}$, $H_{base}$

Bootstrap and Jacknife
Applies to: all
Tests: $H_{AB}$, $H_{base}$
\( H_{null} : C_{ij} \leq C_{null} \)

\( H_{base} : C_{ij} \leq C_{baseline} \)

\( H_{AB} : C_{ij}^{A} = C_{ij}^{B} \)

- **Fully implemented**
- **Partially-developed**
- **Coming soon**
Interactive Time-Frequency Grid
Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie
Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie

Causal Density Movie
Interactive Time-Frequency Grid
Interactive 3D Causal Brainmovie
Causal Density Movie
Directed Graphs on anatomicals (ECoG)
Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie

Causal Density Movie

Directed Graphs on anatomicals (ECoG)

and more...
Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie

Causal Density Movie

Directed Graphs on anatomicals (ECoG)

and more...

All of these currently support single-subject or (in beta version) group analysis. ROI connectivity analysis can currently be performed using dipole clustering.
Interactive Time-Frequency Grid
Interactive Time-Frequency Grid
Causal Time-Frequency Grid

FROM

TO

Time (sec)

Frequency (Hz)

Saturday, September 24, 2011
Causal Time-Frequency Grid
Causal Time-Frequency Grid

Error - Correct
p < 0.05, N=24

Saturday, September 24, 2011
Causal Time-Frequency Grid

Error - Correct
p < 0.05, N=24

Saturday, September 24, 2011
Causal Time-Frequency Grid

Error - Correct
p < 0.05, N=24

Saturday, September 24, 2011
Causal Time-Frequency Grid

Error - Correct
p < 0.05, N=24

Saturday, September 24, 2011
Interactive BrainMovie3D
Interactive BrainMovie3D
Interactive BrainMovie3D
Causal Projection

Error > Correct (p < 0.05, N=24)
dDTF
3-7 Hz
Causal Projection

Error > Correct (p < 0.05) 3-7 Hz
Group Analysis

partially-developed
Group Analysis

Disjoint Clustering

This approach adopts a 3-stage process:

1. Identify K ROI’s (clusters) by affinity clustering of sources across subject population using EEGLAB’s Measure-Product clustering.

2. Average all incoming and outgoing statistically significant connections between each pair of ROIs to create a [ K X K [x freq x time ] ] group connectivity matrix.

3. Visualize the results using any of SIFTs visualization routines. This method suffers from low statistical power when subjects do not have high agreement in terms of source locations (missing variable problem).

partially-developed
This approach adopts a 3-stage process:

1. Identify K ROI’s (clusters) by affinity clustering of sources across subject population using EEGLAB’s Measure-Product clustering.
2. Average all incoming and outgoing statistically significant connections between each pair of ROIs to create a \( [K \times K \times f \times t] \) group connectivity matrix.
3. Visualize the results using any of SIFTs visualization routines. This method suffers from low statistical power when subjects do not have high agreement in terms of source locations (missing variable problem).

A more robust approach (in development with Wes Thompson and to be released in SIFT 1.0b) uses smoothing splines and Monte-Carlo methods for joint estimation of posterior probability (with confidence intervals) of cluster centroid location and between-cluster connectivity. This method takes into account the “missing variable” problem inherent to the disjoint clustering approach and provides robust group connectivity statistics.

See Thompson and Mullen et al (2011), *ICON XI*
Bayesian Group Inference

Error > Baseline (p < 0.01, N=24)
dDTF
3-7 Hz