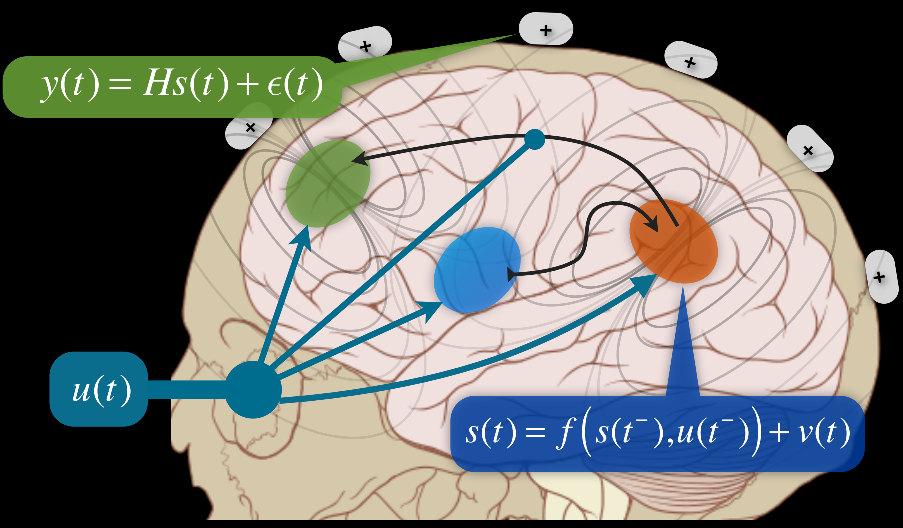
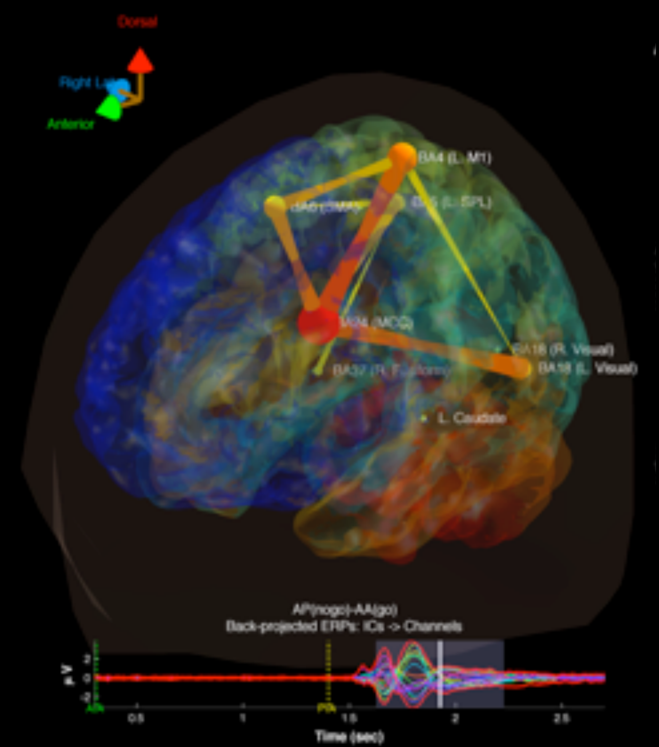
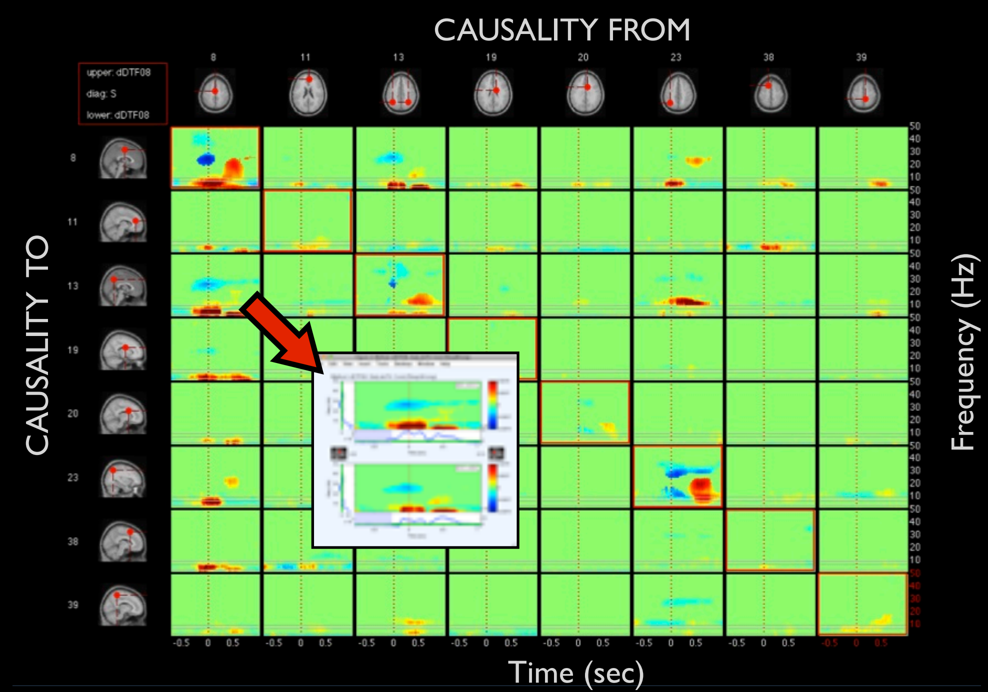
The Dynamic Brain: Modeling Neural Dynamics and Interactions from M/EEG





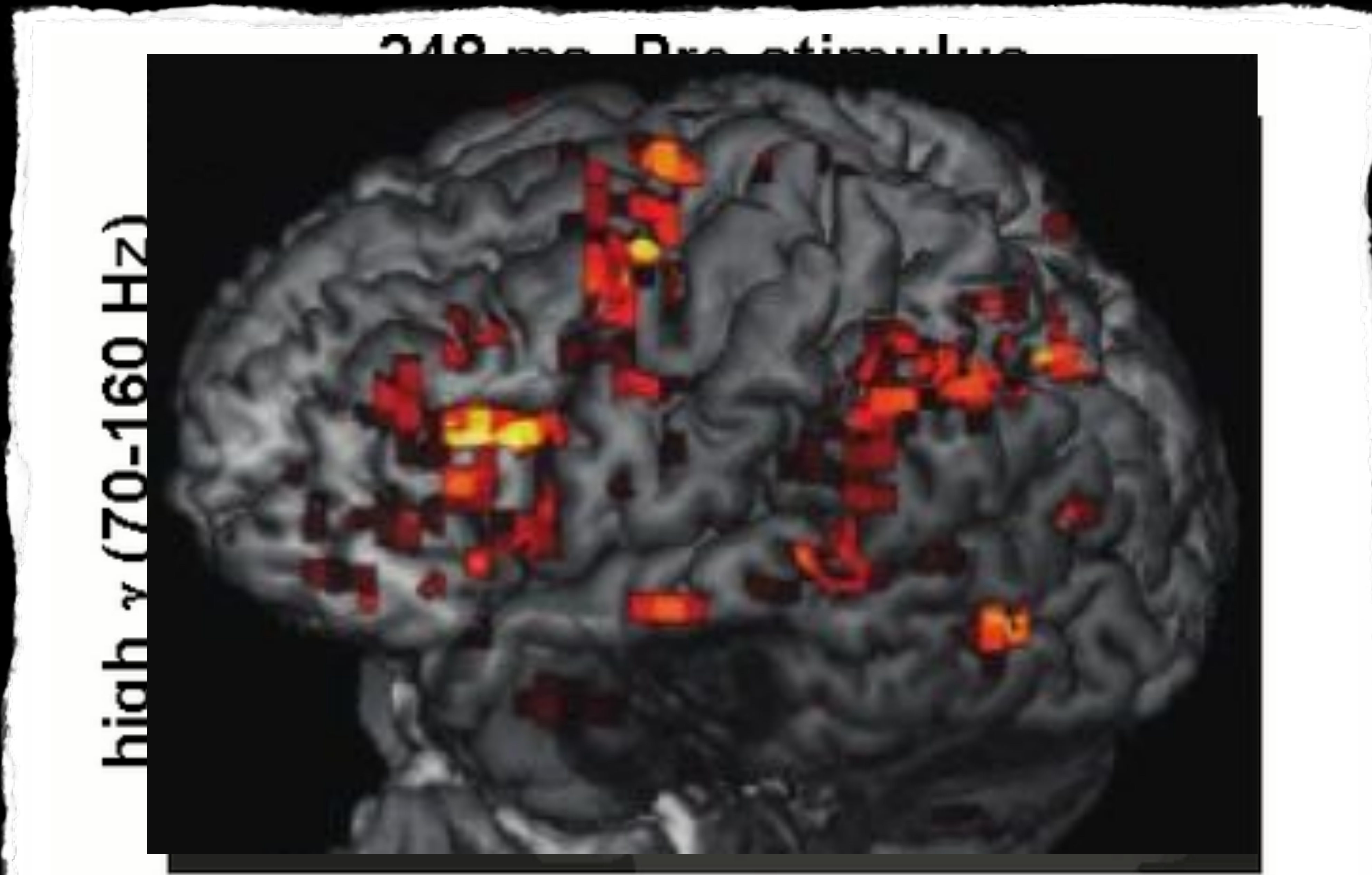


Tim Mullen

Swartz Center for Computational Neuroscience
Institute for Neural Computation
Dept. of Cognitive Science
UC San Diego

Outline

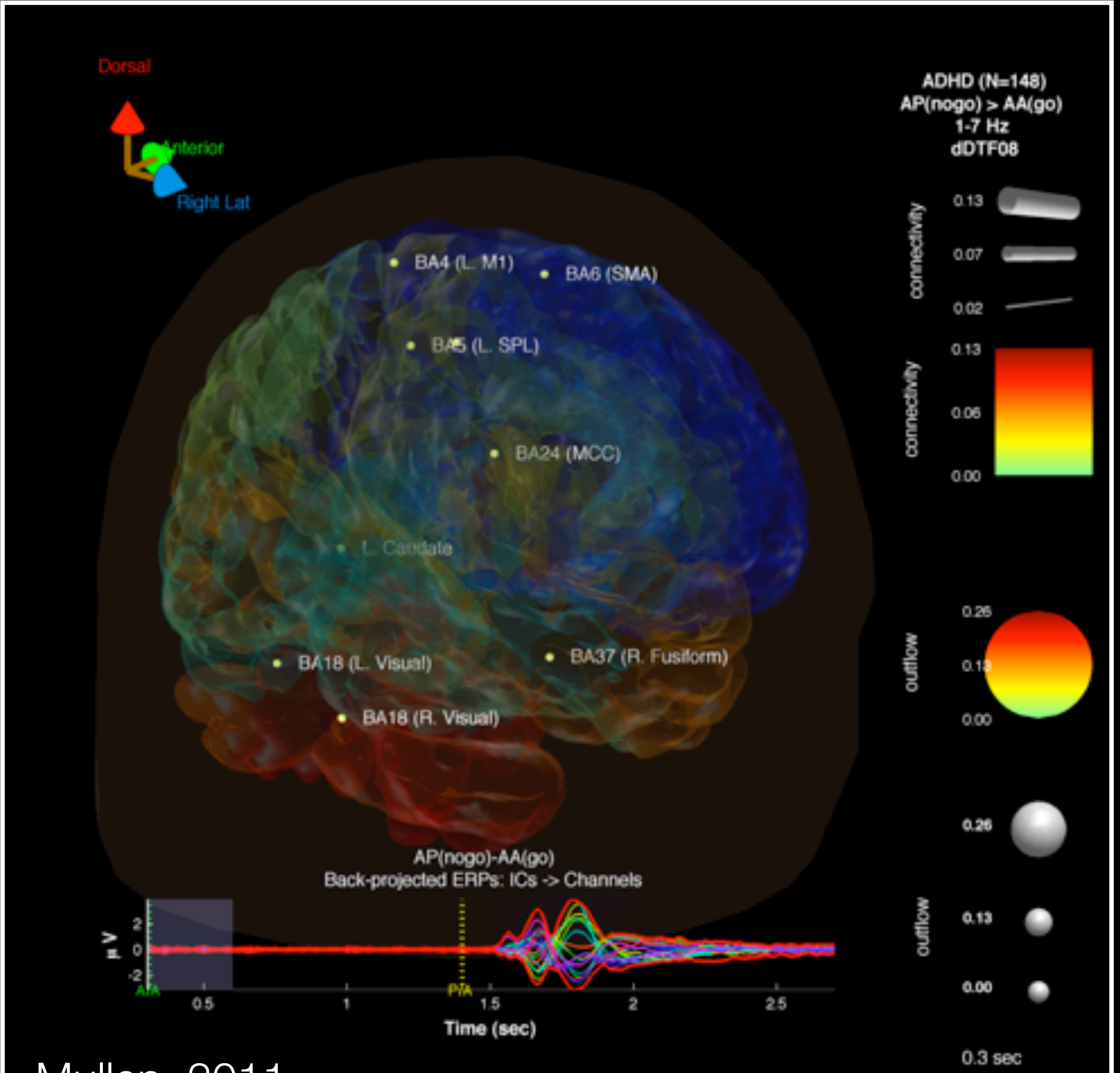
The Dynamic Brain



Edwards, Canolty, et al, 2007

The Dynamic Brain

- A key goal: To model temporal changes in neural dynamics and information flow that index and predict task-relevant changes in cognitive state and behavior
- **Open Challenges:**
  - Non-invasive measures (source inference)
  - Robustness and Validity (constraints & statistics)
  - Scalability (multivariate)
  - Temporal Specificity / Non-stationarity / Single-trial (dynamics)
  - Multi-subject Inference
  - Usability and Data Visualization (software)



Mullen, 2011

Modeling Brain Connectivity

- Model-based approaches mitigate the 'curse of dimensionality' by making some assumptions about the structure, dynamics, or statistics of the system under observation

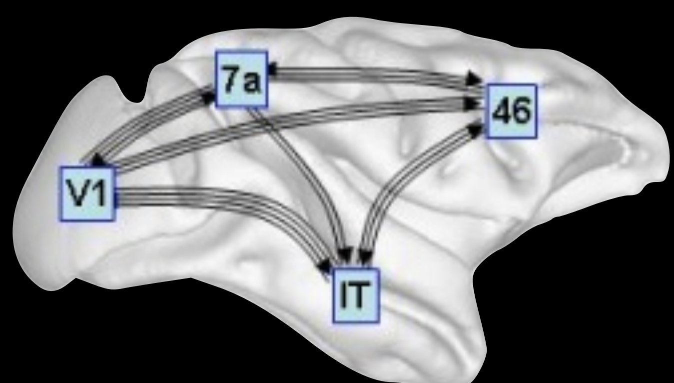
Box and Draper (1987):

> "Essentially, all models are wrong, but some are useful [...] the practical question is how wrong do they have to be to not be useful"

Categorizations of Large-Scale Brain Connectivity Analysis

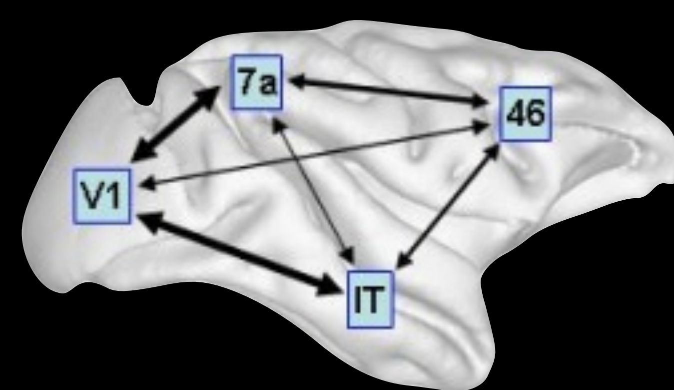
(Bullmore and Sporns, *Nature*, 2009)

Structural



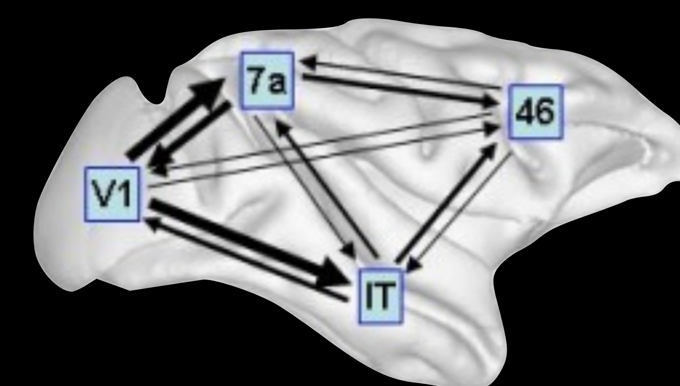
state-invariant, anatomical

Functional



dynamic, state-dependent, correlative, symmetric

Effective



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

Hours-Years **milliseconds-seconds**

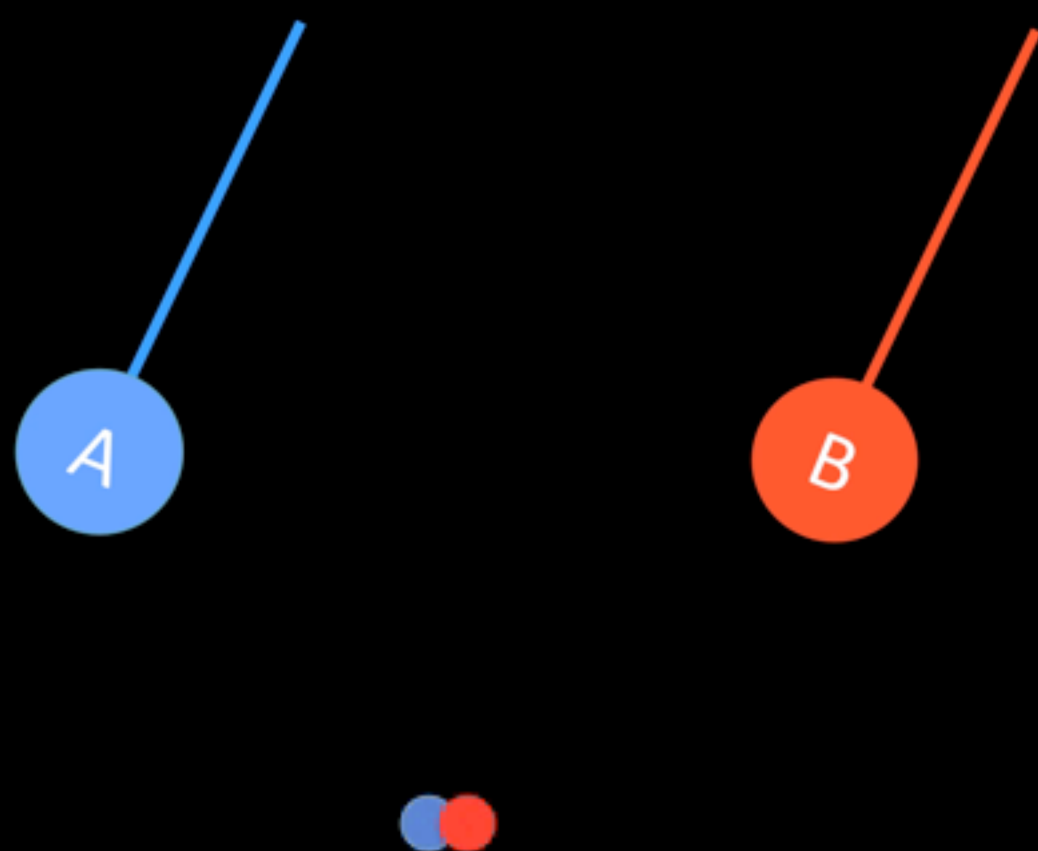
Temporal Scale

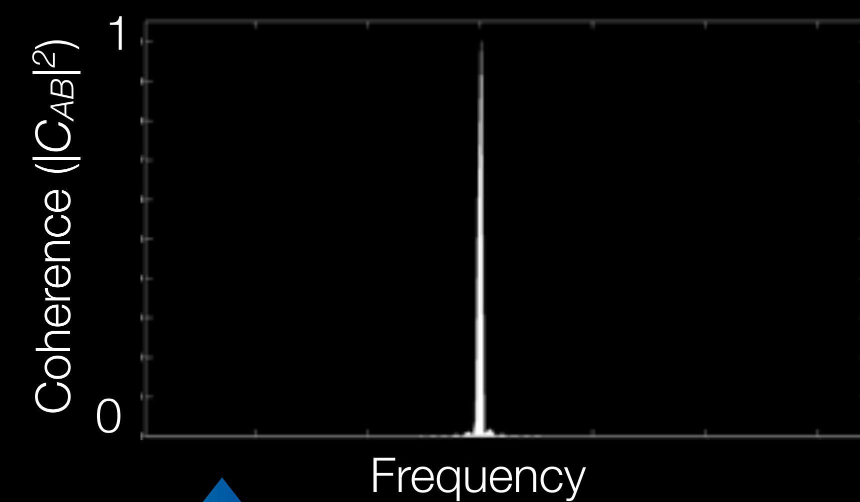
Intro | **Theory** | SIFT | Apps | To-Do | Fin

Estimating Functional Connectivity

Popular measures
Cross-Correlation
Coherence
Phase-Locking Value
Phase-amplitude coupling
...

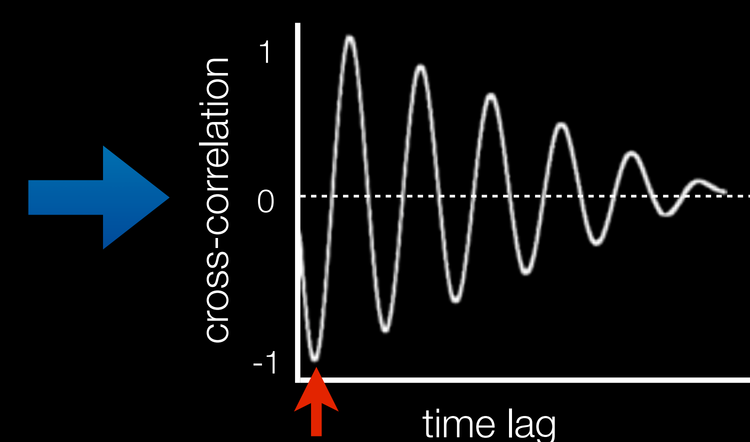
Cross-Correlation and Linear Coherence





DFT

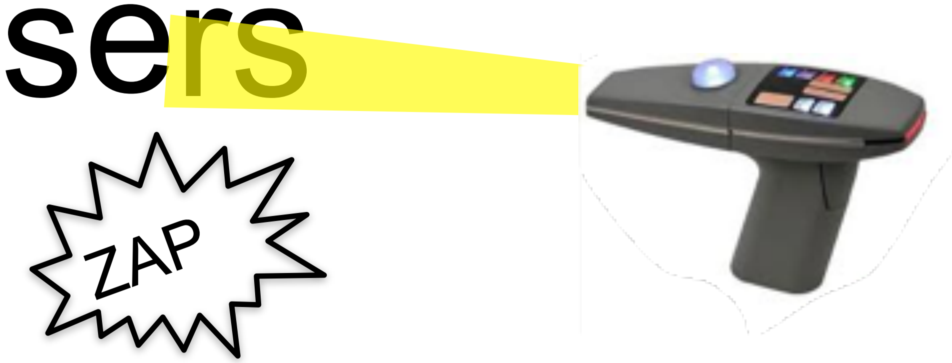
$$C_{AB}(f) = \sum_{k=0}^{p} \rho_{AB}(k) e^{-i2\pi fk} = \frac{S_{AB}(f)}{\sqrt{S_A(f) S_B(f)}}$$



$\rho_{AB}(k)$

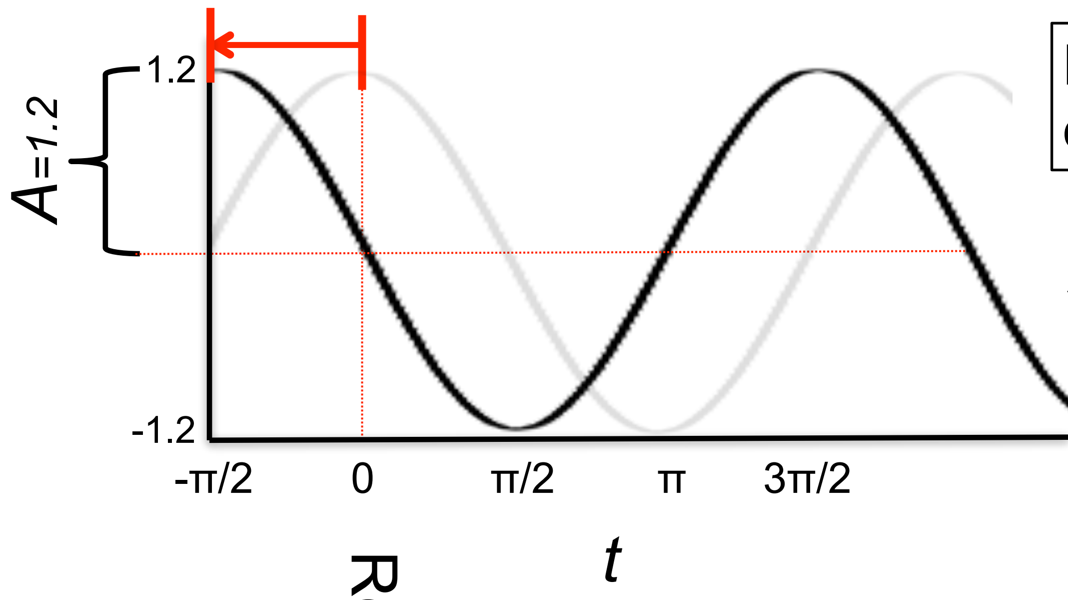
Issue: Linear coherence is biased by auto-power (just as the cross-correlation is biased by strong autocorrelation in individual time series)

Phasers



phase shift
$\phi = \pi / 2$

angular frequency
$\omega = 2\pi f = 2\pi$ rad/sec



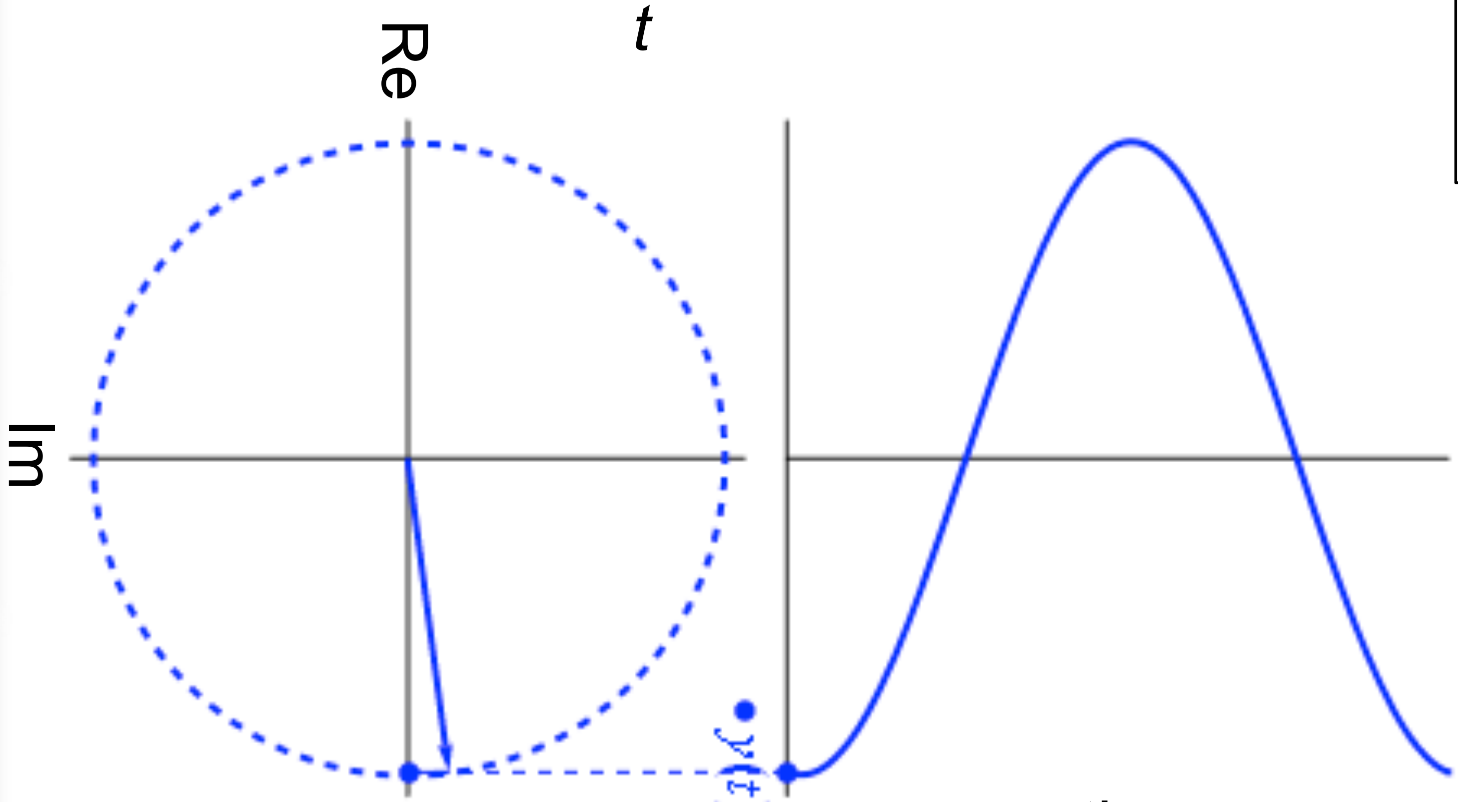
Euler's Formula tells us that any sinusoid can be expressed as the sum of two complex exponentials

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

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

... or (if real-valued) as the real part of a single complex exponential

instantaneous complex amplitude and phase



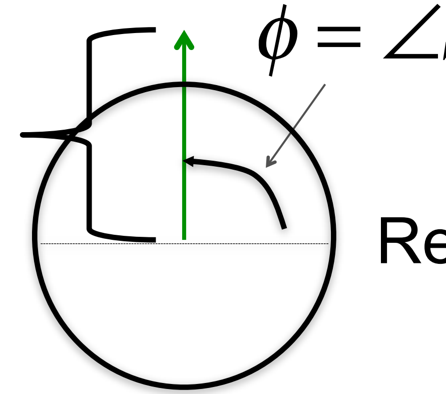
(Axis rotated 90º CCW for display)

Polar animation courtesy Wikipedia

Phasor

(Polar Coords)

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



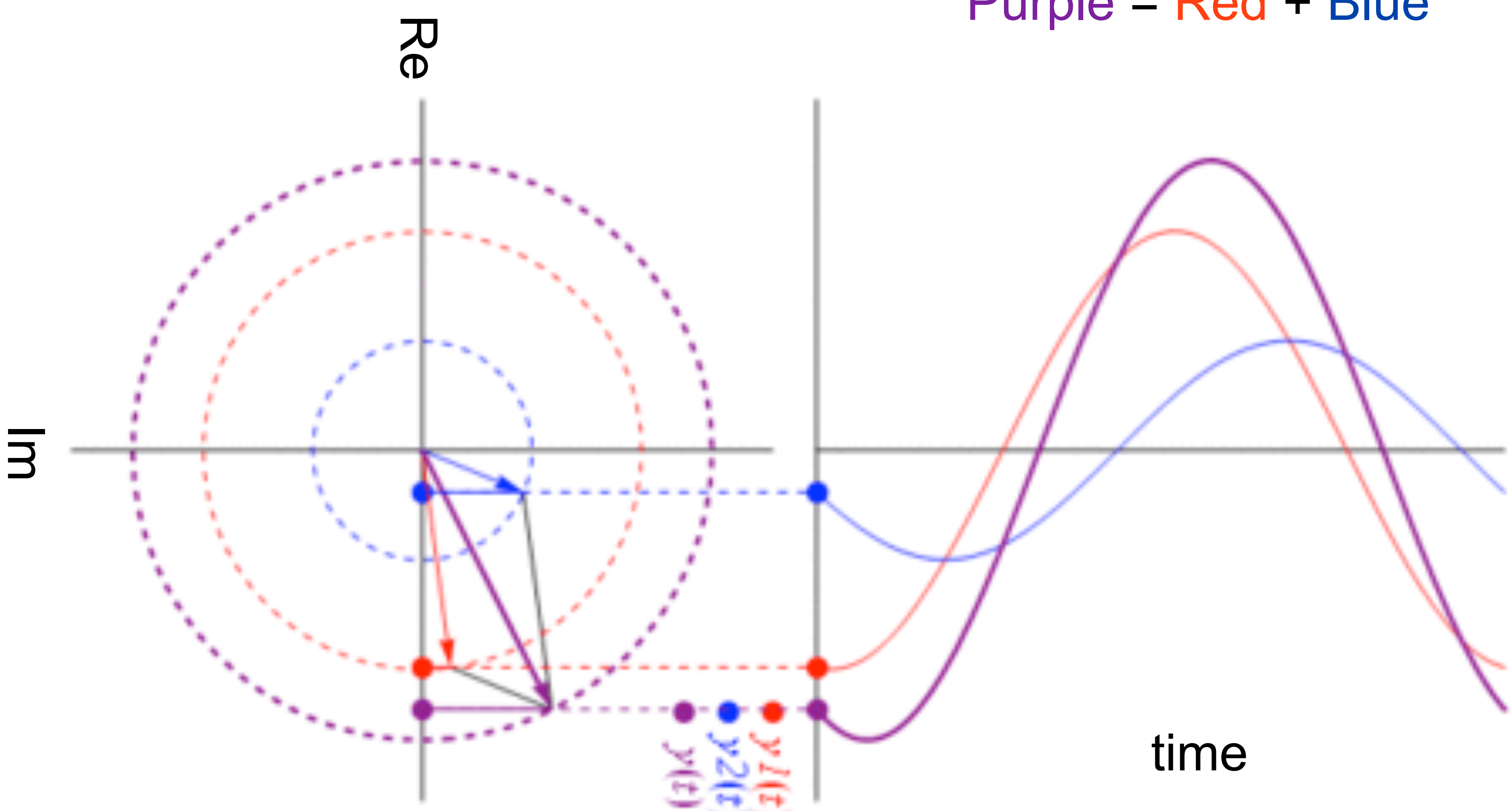
$\phi = \angle S(\omega, t)$

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

Phasors

If we want to examine oscillatory dynamics or relationships between oscillatory signals, analysis in the time domain (i.e. cartesian coordinates) is equivalent to (simpler) operations involving phasors in Fourier space (i.e. polar coordinates).

Purple = Red + Blue

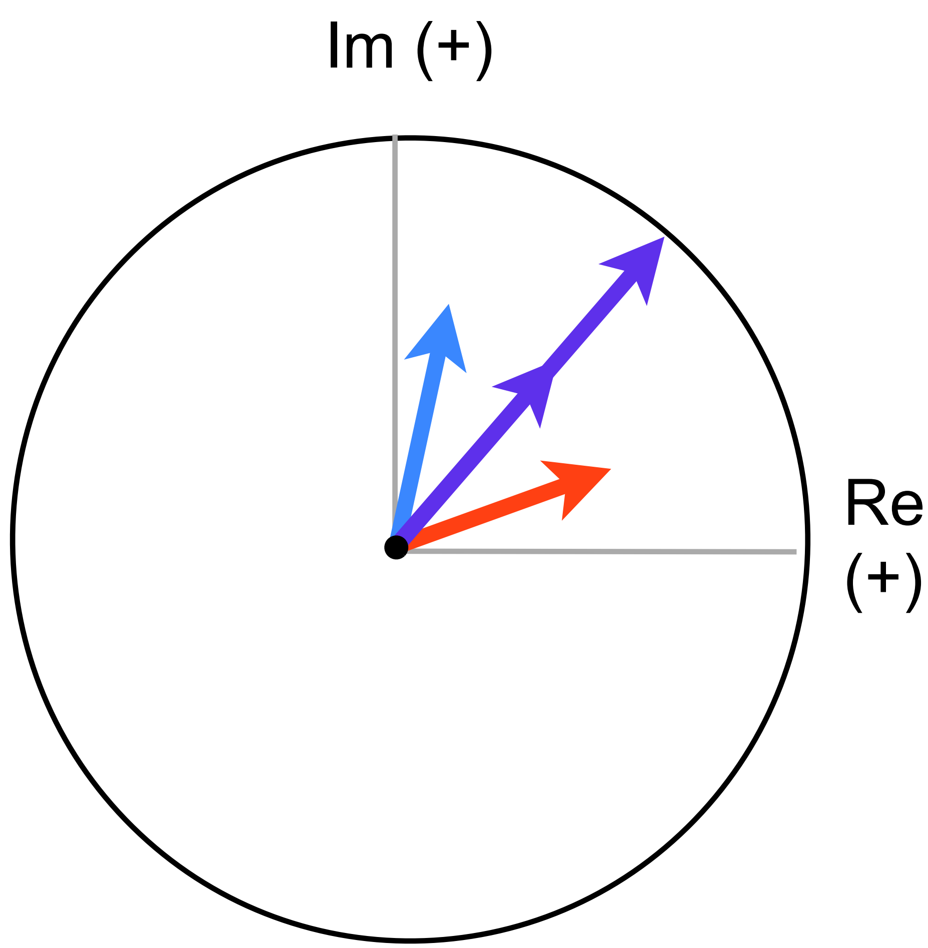


(Axis rotated 90° CCW for display)

Polar animation courtesy Wikipedia

The Mean Phasor

The average of *k* phasors is a new phasor constructed by adding up the original vectors and dividing the length of the resultant vector by *k.*



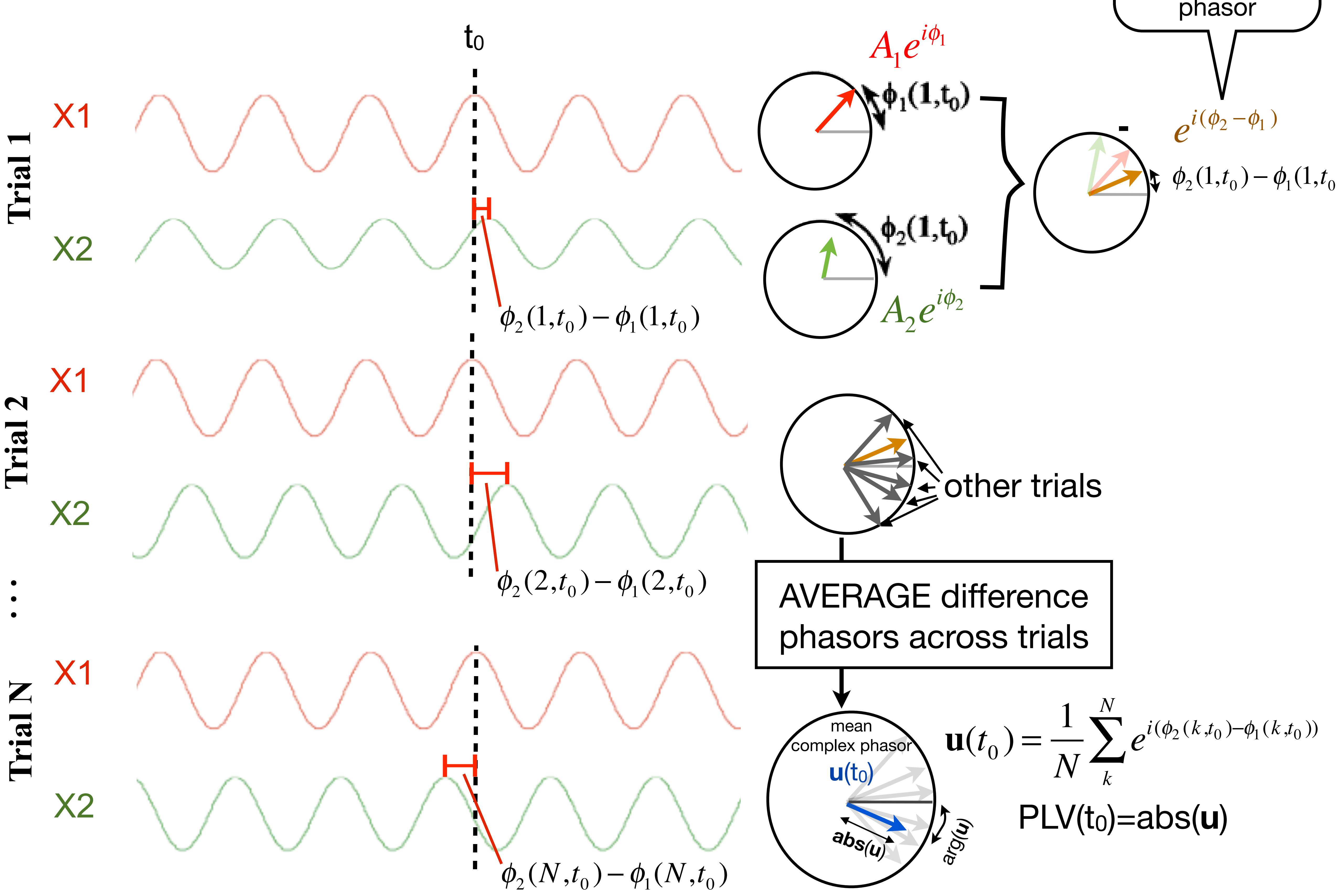
Purple = (Red + Blue) / 2

If all **phasors have similar angles**, then vectors will “point” in the same direction and the **length of the mean phasor** will be comparatively **large**.

If **phasor angles are random**, then vectors will point in random directions and the **length of the mean phasor** will be close to **zero**

Phase-Locking Value (PLV)

Lachaux, J.P., et al (1999) *HBM*



Phase-Locking Value (PLV)

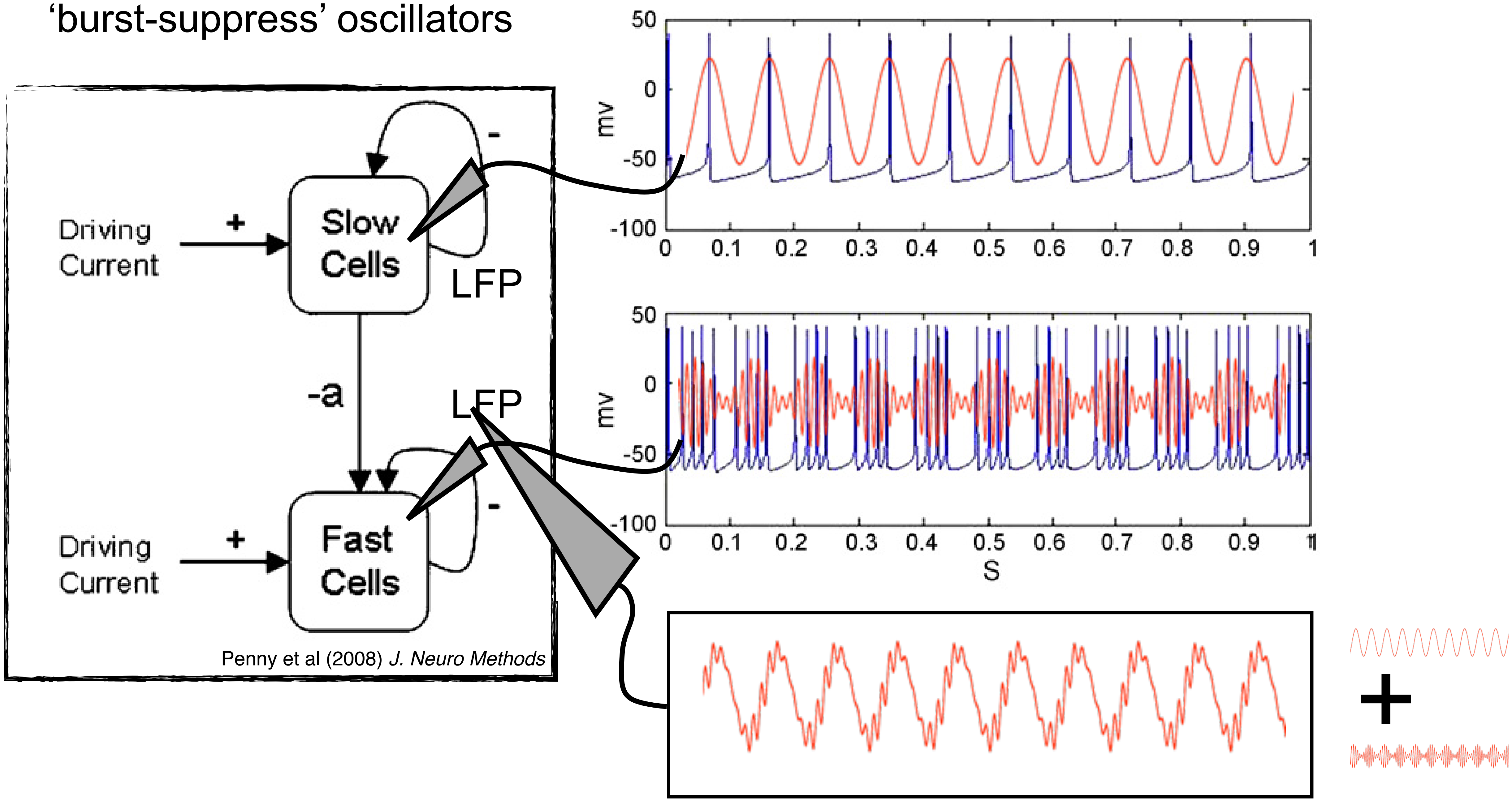
Lachaux, J.P., et al (1999) *HBM*

Computing PLV ("phase coherence") in EEGLAB:

```
pop_newcrossf(....,'type','phase')
```

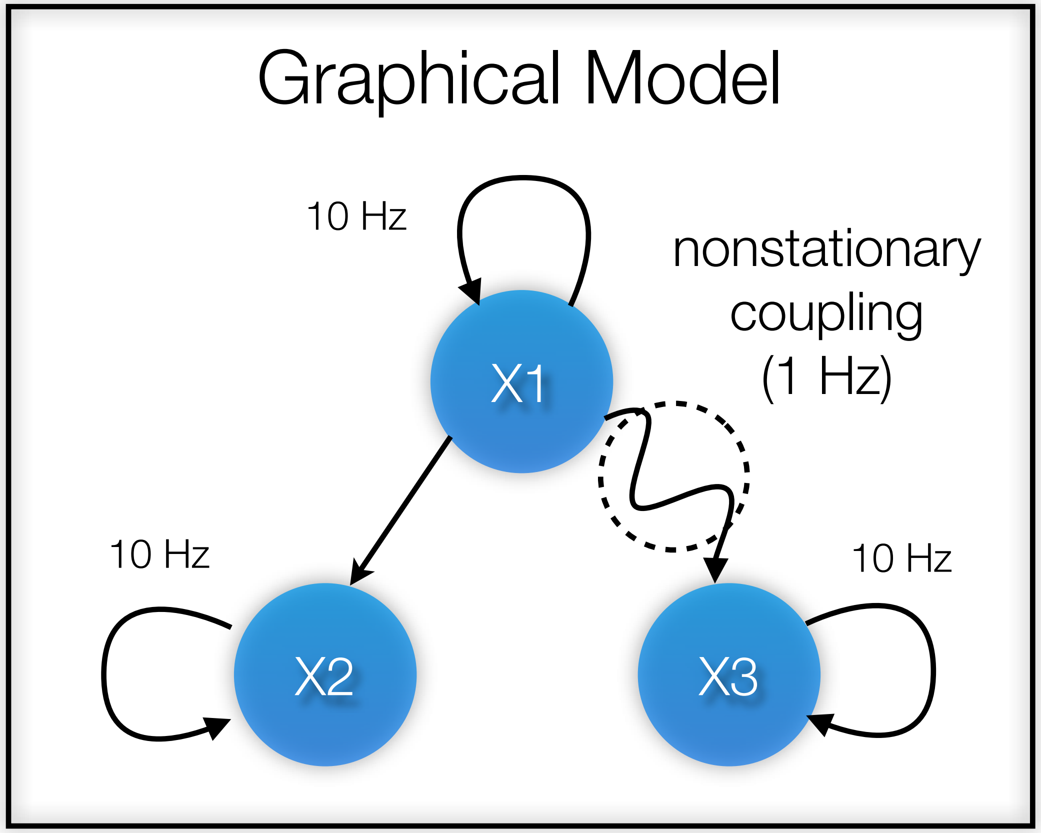
Phase-Amplitude Coupling

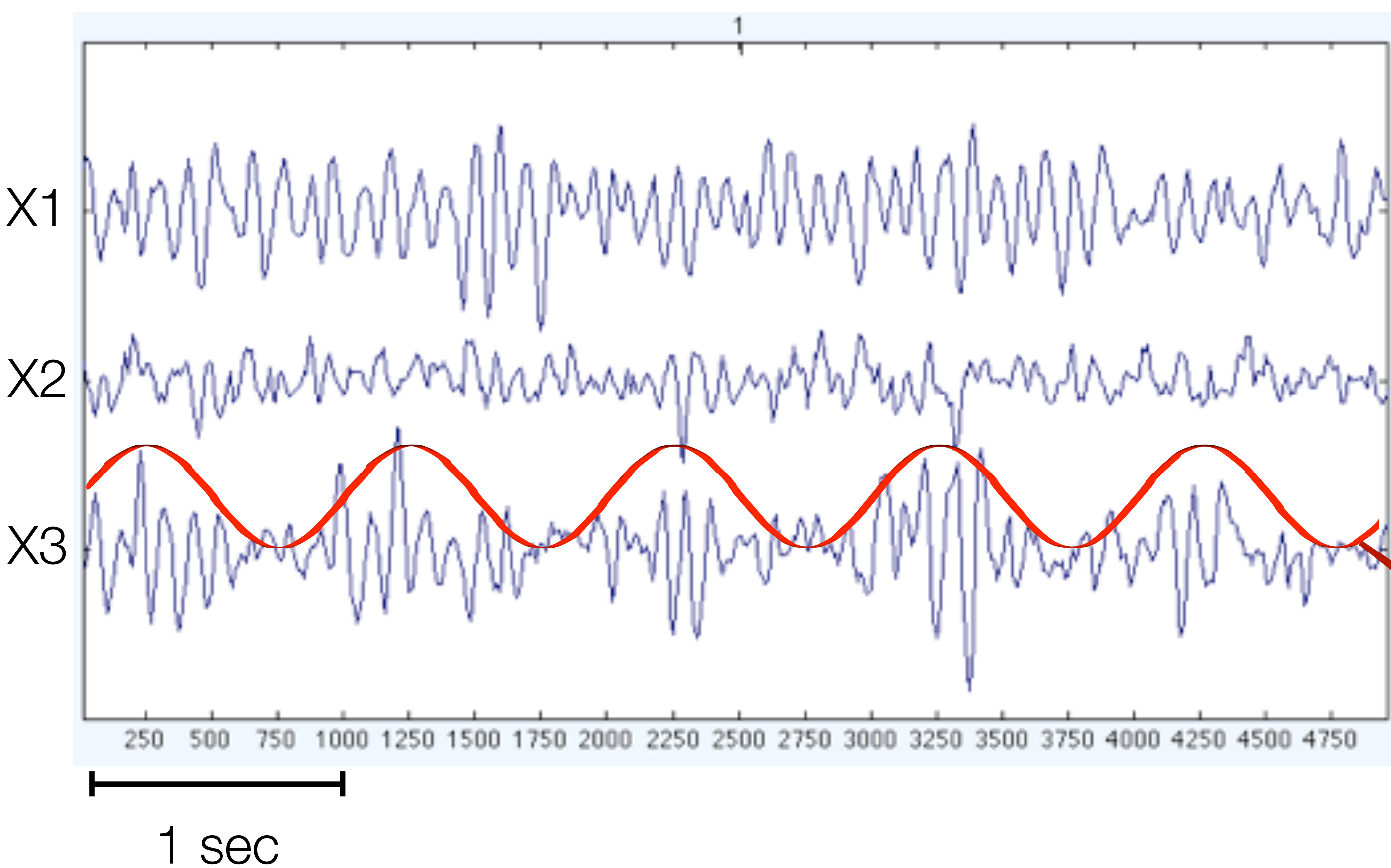
'burst-suppress' oscillators



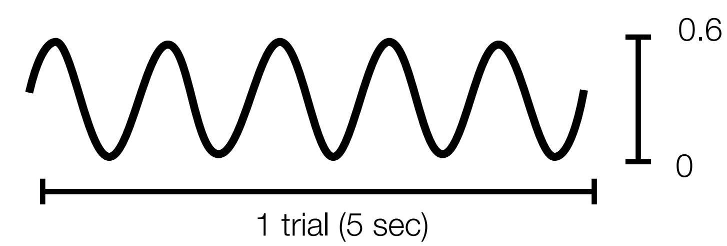
Local Field Potential (Slow + Fast cells)

PAC may reflect non-stationary or non-linear network dynamics





Time-varying X1→X3 coupling
(1 Hz modulation)



Amplitude Modulation
10Hz amplitude coupled to 1 Hz Phase

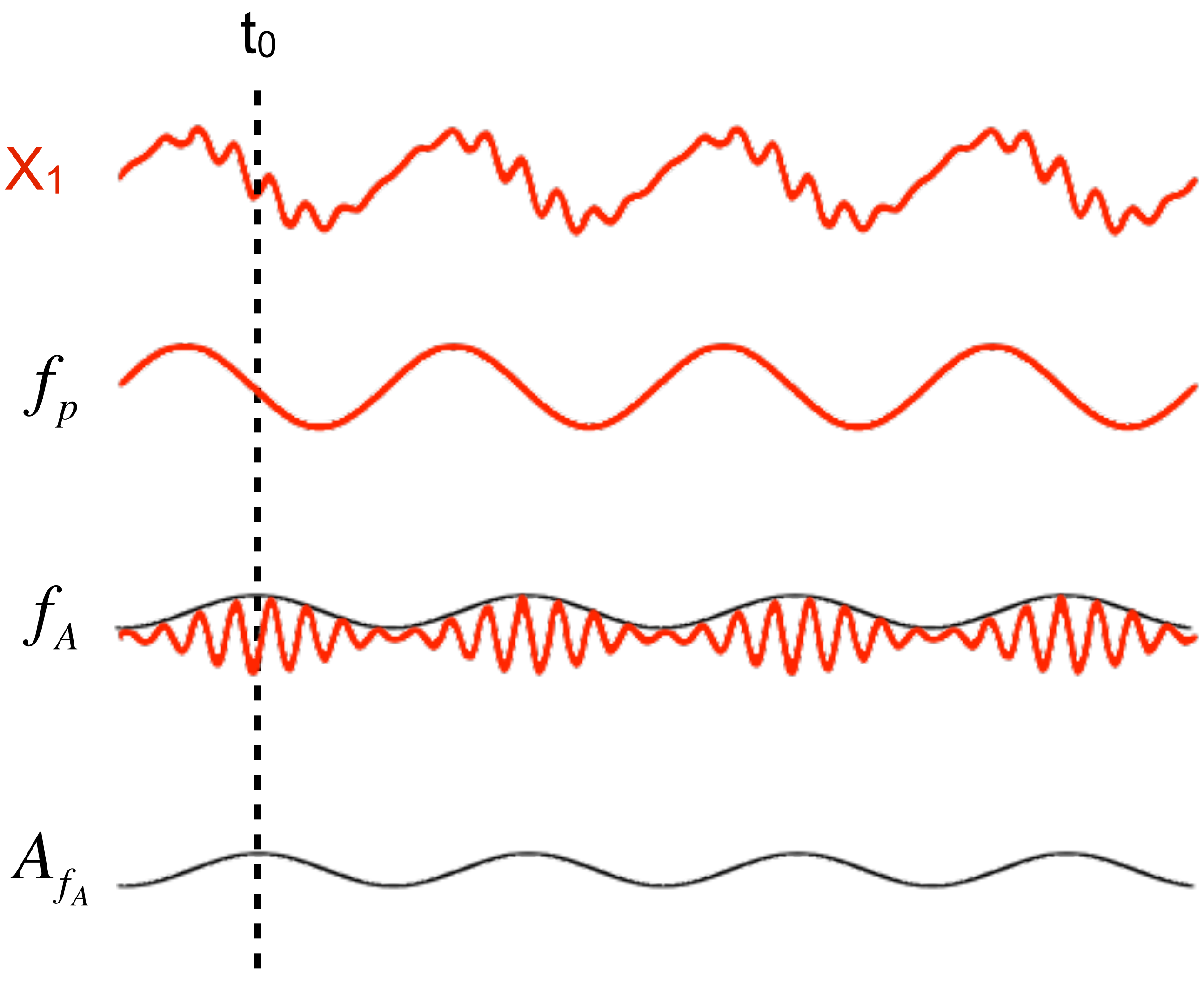
Phase-Amplitude Coupling

- 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

Vanhatalo, S et al (2004) *PNAS*



original raw signal

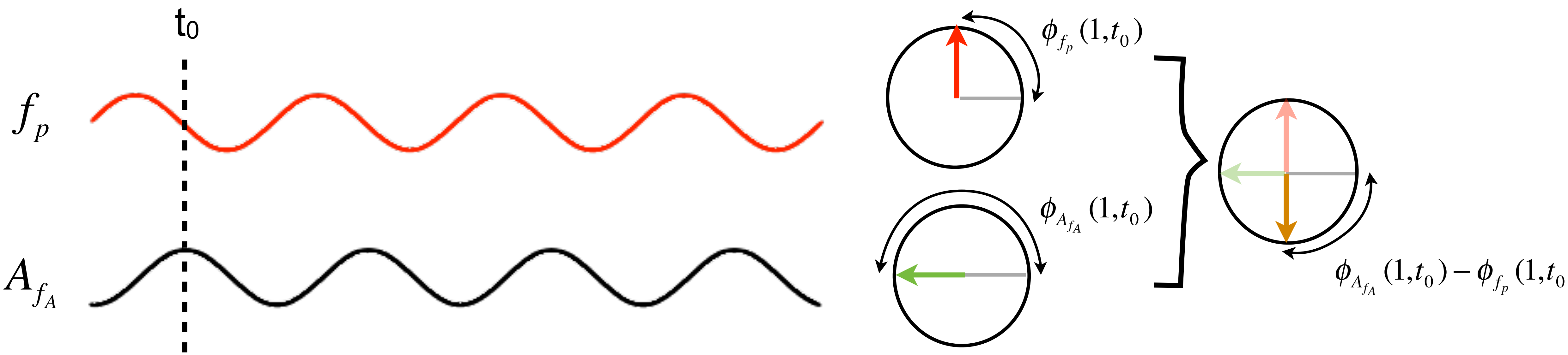
filter X_1 at LFO band (e.g. theta)

filter X_1 at HFO band (e.g. gamma)

get amplitude envelope of filtered signal

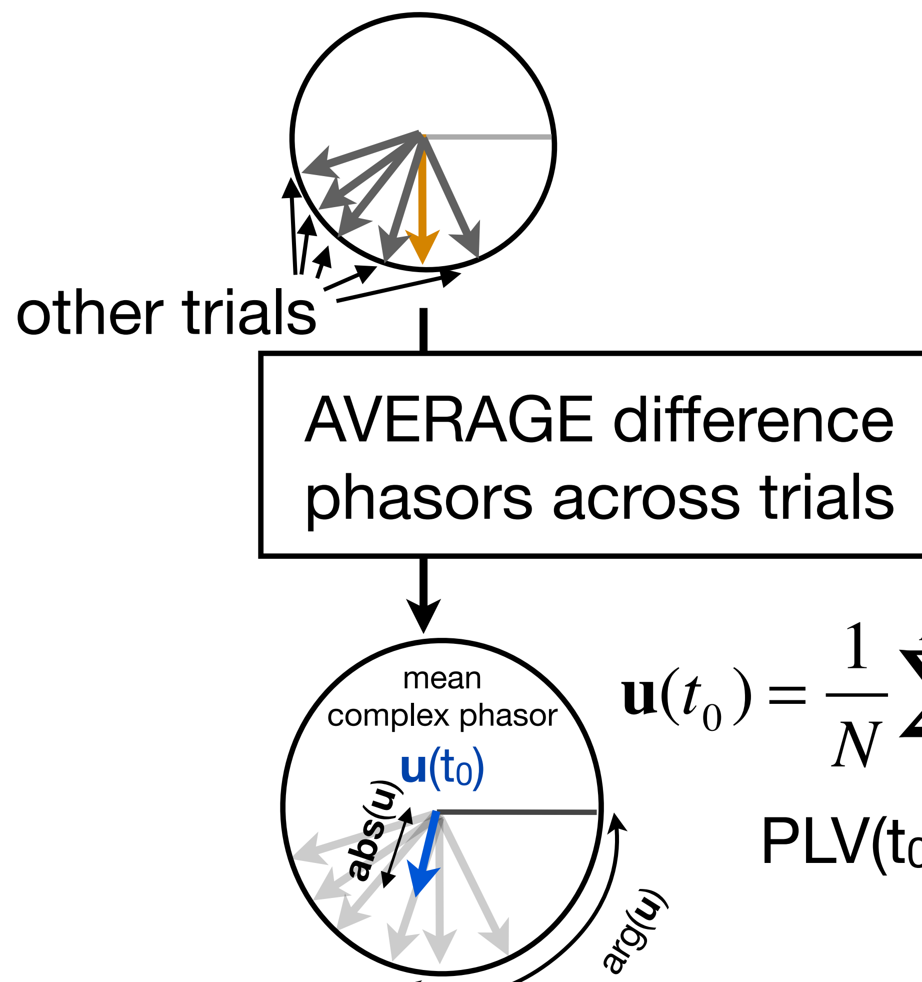
Phase-Amplitude Coupling: PLV Method

Vanhatalo, S et al (2004) *PNAS*



Compute PLV between LFO time-series (f_p) and amplitude envelope of HFO time-series (A_{fA}).

Significant PLV indicates that the central frequency of f_p modulates the amplitude of the central frequency of f_A



$$\mathbf{u}(t_0) = \frac{1}{N}\sum_{k}^{N} e^{i(\phi_{A_{f_A}}(k,t_0)-\phi_{f_p}(k,t_0))}$$

PLV(t_0)=abs(**u**)

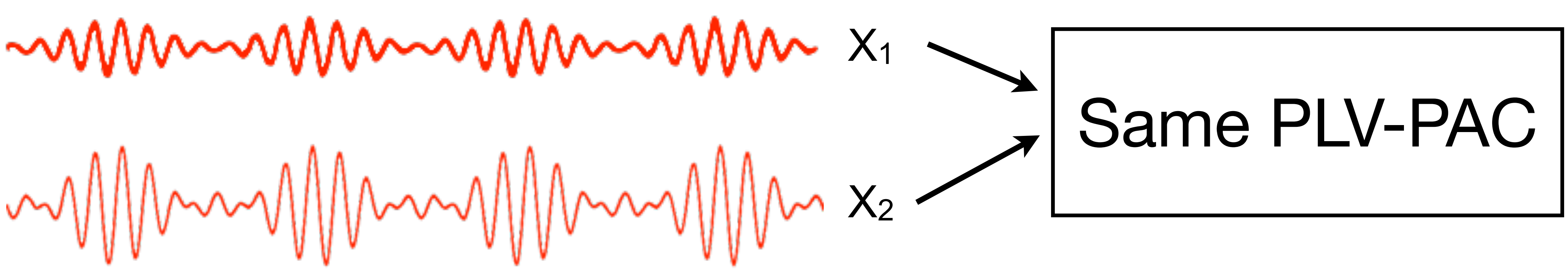
Phase-Amplitude Coupling: PLV Method

Vanhatalo, S et al (2004) *PNAS*

Problem:

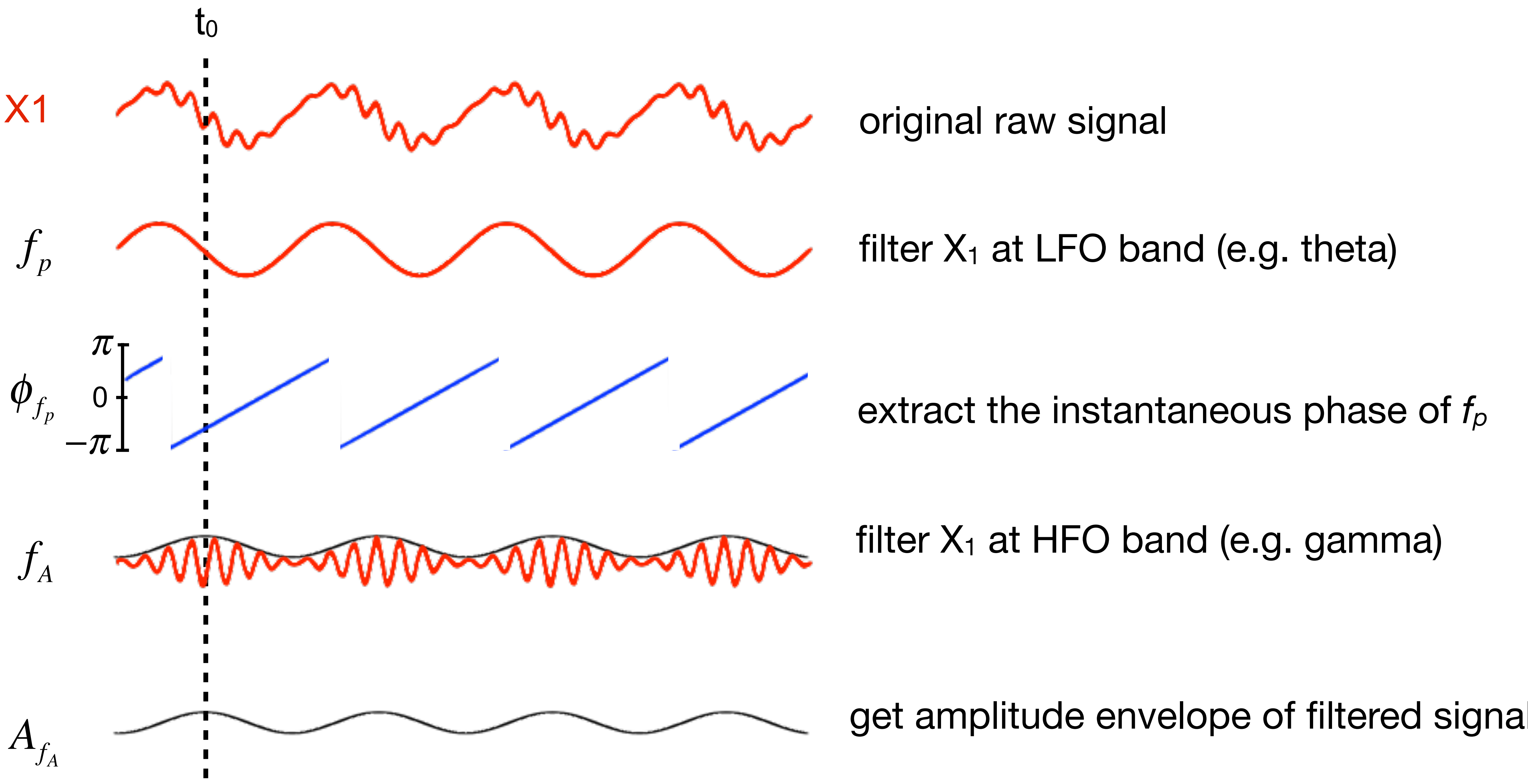
PLV is invariant to differences in amplitude between the two time-series (it only considers phase). Thus PLV-PAC doesn't take into account the *amplitude* of the co-modulation.

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

Canolty et al, (2006) *Science*

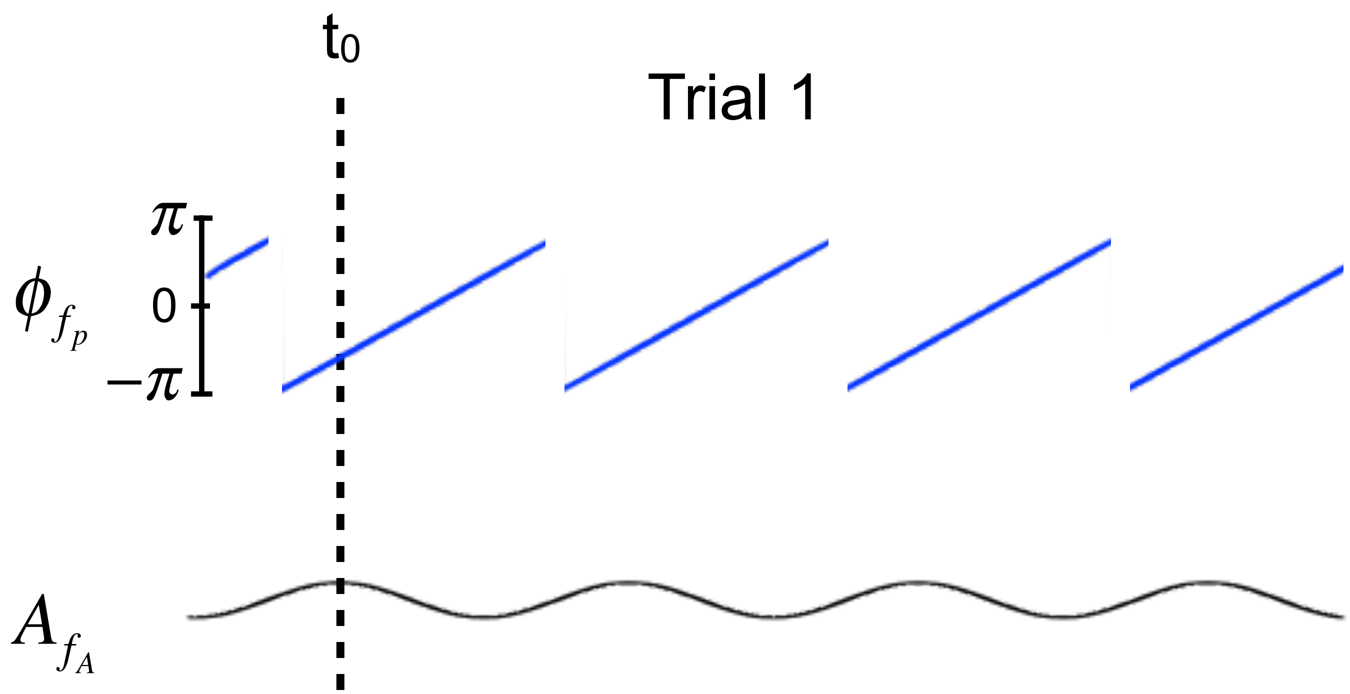


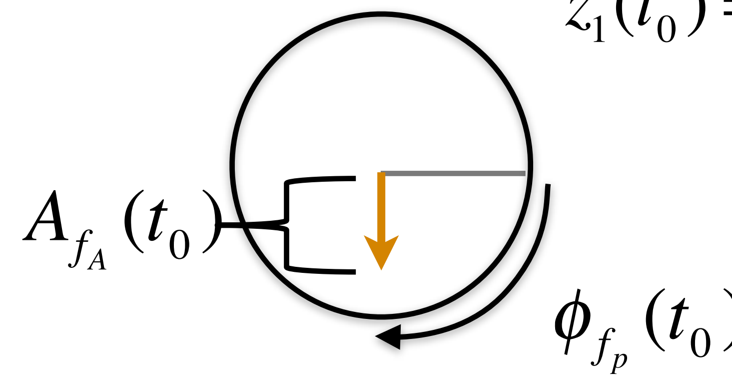
Phase-Amplitude Coupling: Modulation Index Method

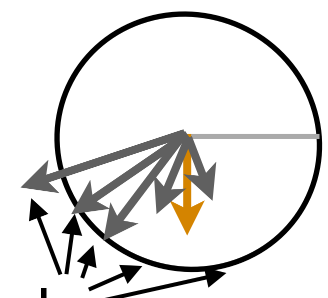
Canolty et al, (2006) *Science*

build complex phasor with instantaneous amplitude and phase

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

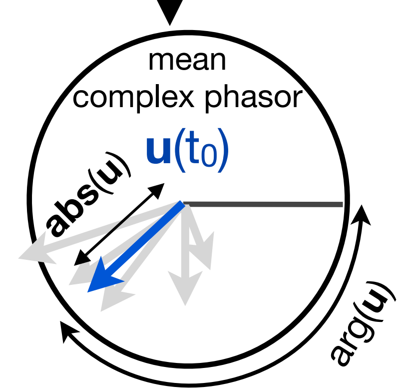






other trials

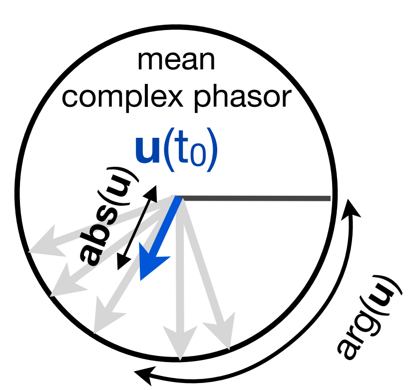
AVERAGE complex phasors across trials



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

PAC(t_0)=abs(**u**)

Comparison: PLV-PAC



Phase-Amplitude Coupling: Modulation Index Method

Canolty et al, (2006) *Science*

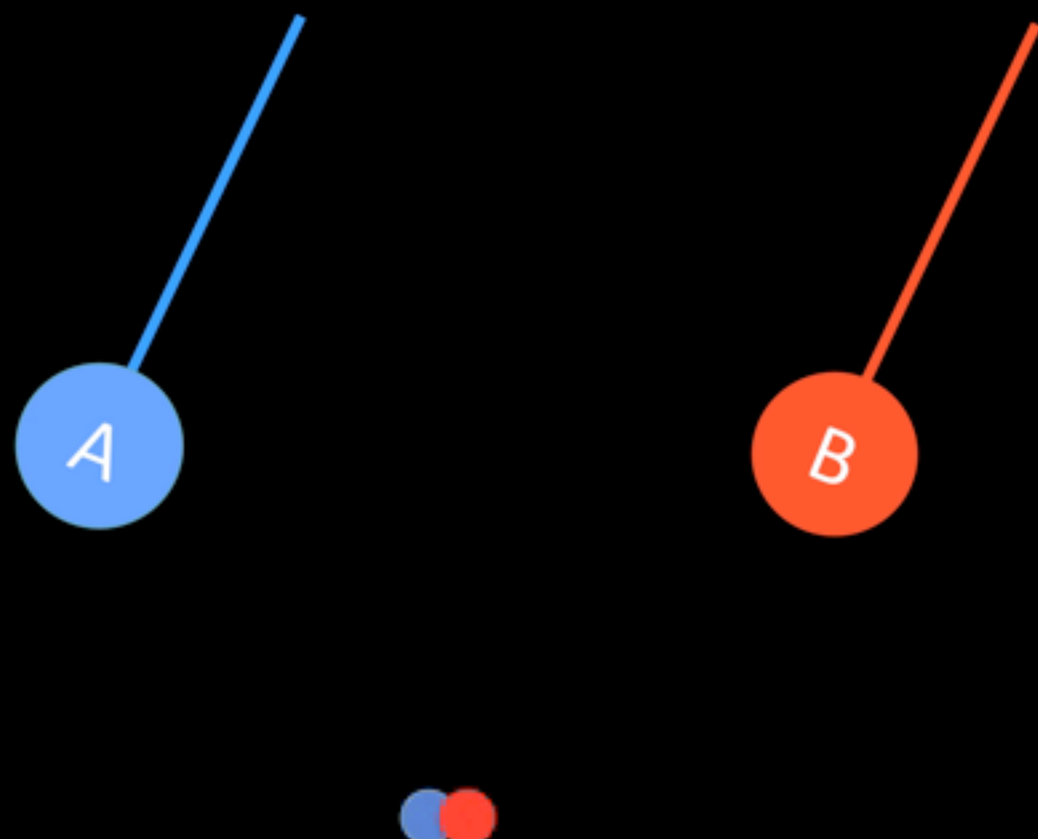
Computing PAC in EEGLAB:

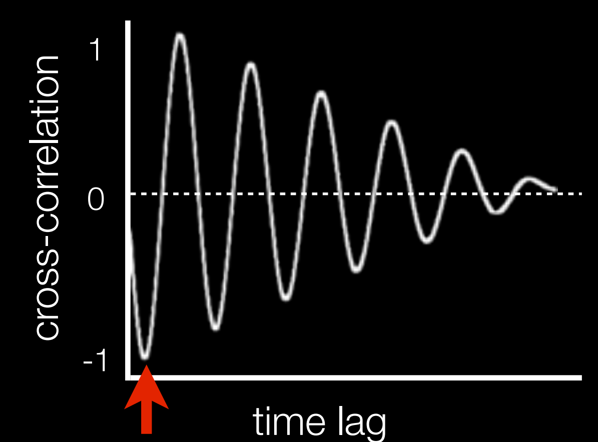
```
pac(IC1,IC2,...,'method','mod')
```

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





$$\rho_{AB}(k)$$

Coherence/CC/PLV indicate ***functional***, but not ***effective*** connectivity

Estimating Effective Connectivity

Non-Invasive

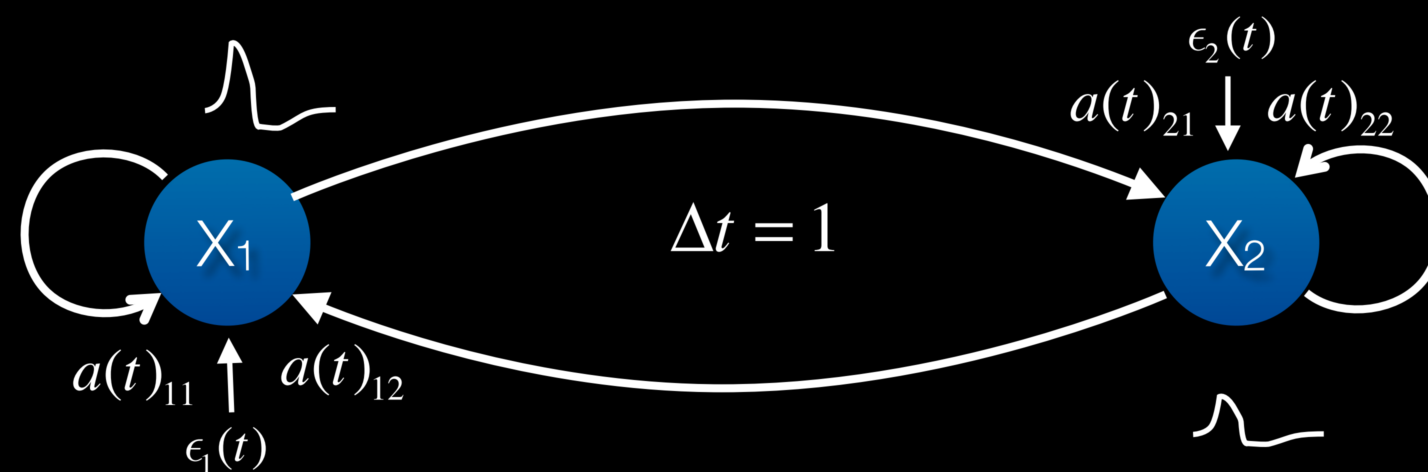
- *Post-hoc* analyses applied to measured neural activity
- Confirmatory
  - Dynamic Causal Models
  - Structural Equation Models
- Exploratory
  - Granger-Causal methods

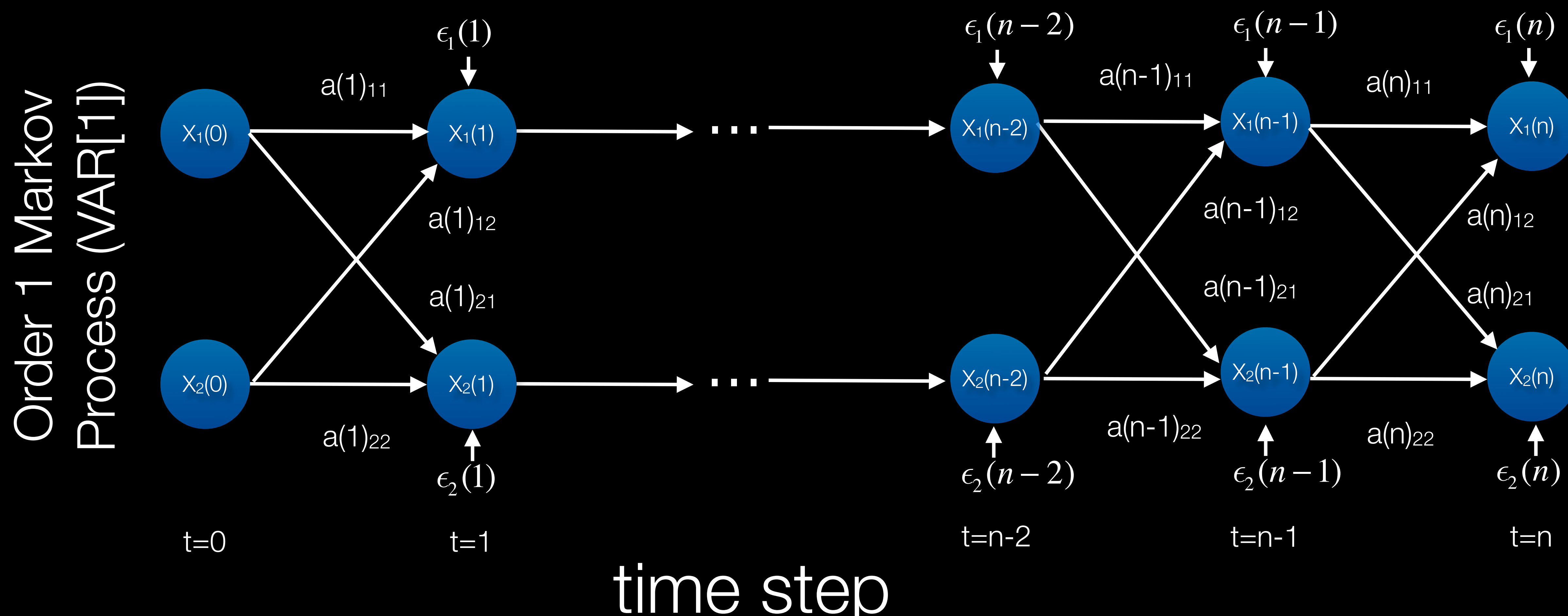
- Data-driven
- Rooted in *conditional predictability*
- Scalable (Valdes-Sosa, 2005)
- Extendable to nonlinear and/or non-stationary systems (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
- Extendable to non-parametric representations (Dhamala, 2009a,b)
- Can be (partially) controlled for (unobserved) 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

Linear Dynamical Systems

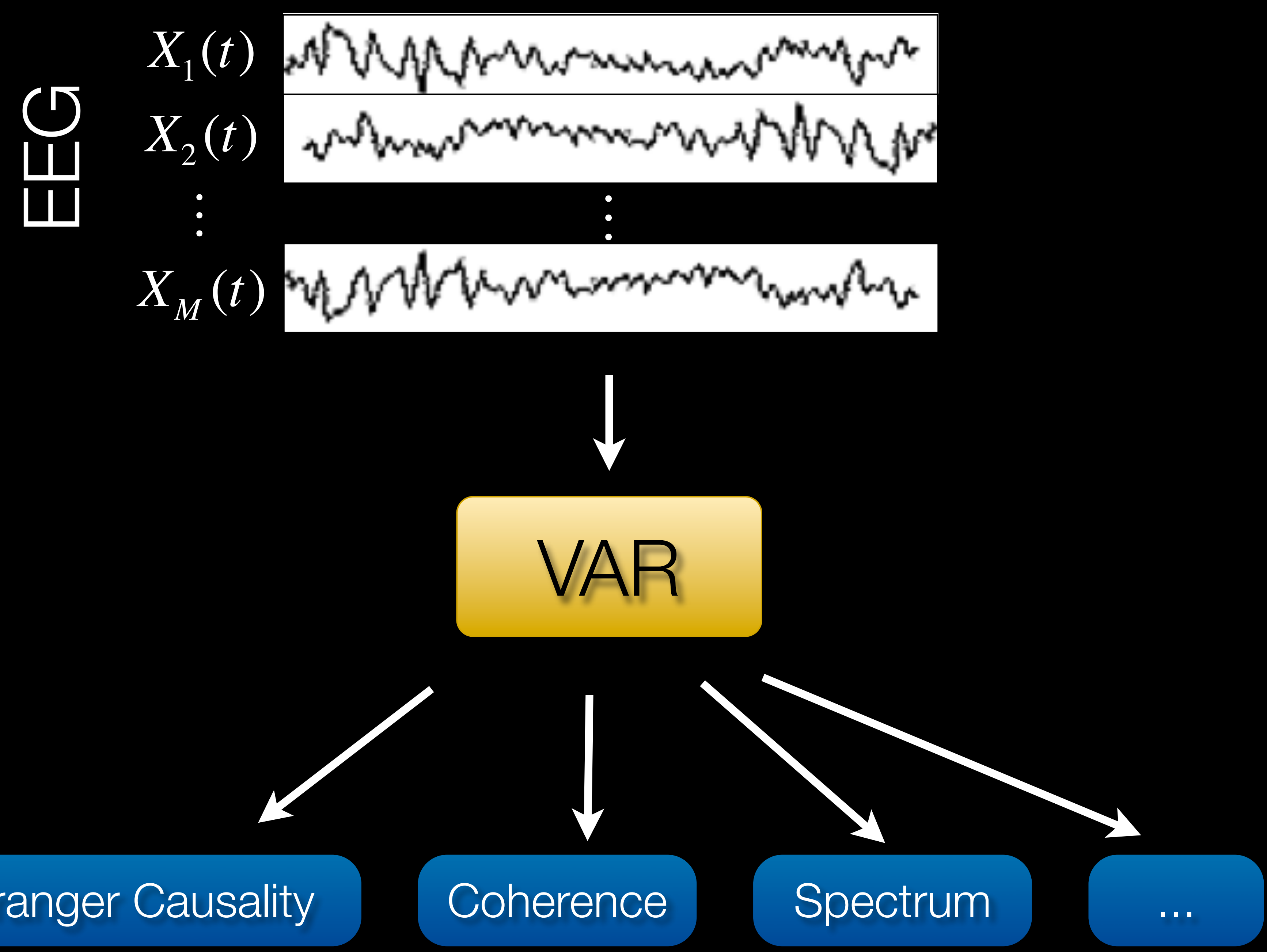
Stochastic Linear Dynamical System

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





Vector Autoregressive (VAR / MAR / MVAR) Modeling

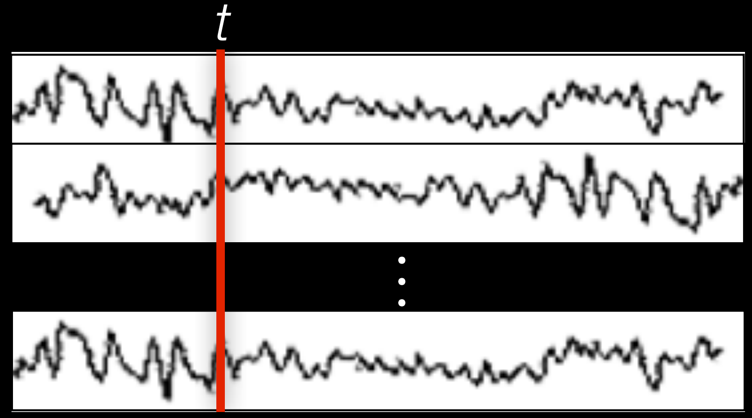


VAR Modeling: Assumptions

- **"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**
  - All eigenvalues of the system matrix are ≤1
  - A stable process will not "blow up" (diverge to infinity)
  - A stable model is always a stationary model (however, the converse is not necessarily true). If a stable model adequately fits the data (white residuals), then the data is likewise stationary

The Linear VAR Model

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



- Ordinary Least-Squares
- Lattice Filters
- Kalman Filtering
- Bayesian Methods
- Sparse methods
- ...

VAR[*p*] model

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

- model order: p
- random noise process: $\mathbf{E}(t)$
- M-channel data vector at current time *t*: $\mathbf{X}(t)$
- M x M matrix of (possibly time-varying) model coefficients indicating variable dependencies at lag *k*: $\mathbf{A}^{(k)}(t)$
- multichannel data *k* samples in the past: $\mathbf{X}(t-k)$

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

$$\mathbf{E}(t) = N(0, \mathbf{V})$$

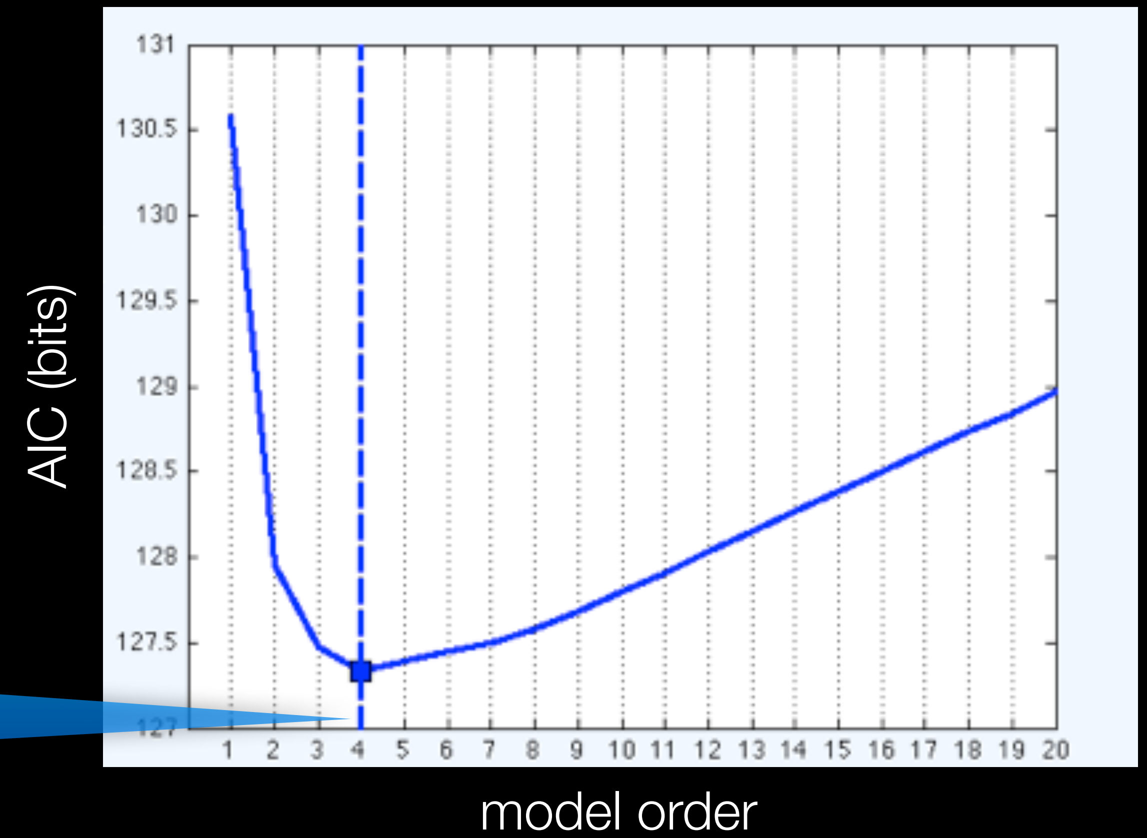
Selecting a VAR Model Order

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

$$\mathrm{AIC}(p) = 2\log(\det(\mathbf{V})) + M^2p/N$$

Penalizes high model orders (parsimony)

entropy rate (amount of prediction error)

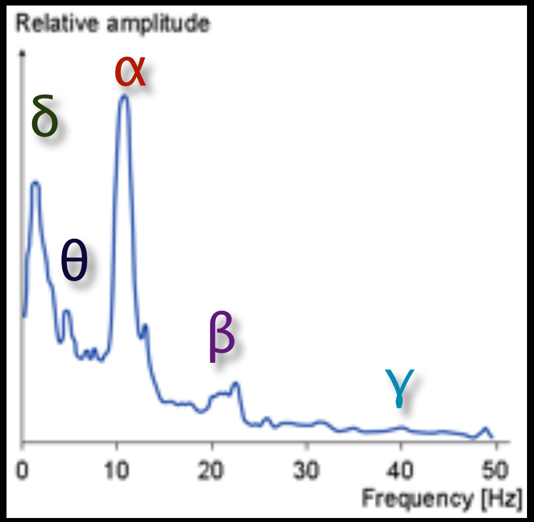


optimal order

Intro Theory SIFT Apps To-Do Fin

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 cross spectrum between each pair of variables)



  - Optimal model order depends on sampling rate. Higher sampling rate often requires higher model orders.

Model Validation

- If a model is poorly fit to data, then few, if any, inferences can be validly drawn from the model. There a number of criteria which we can use to determine whether we have appropriately fit our VAR model. Here are three commonly used categories of tests:
- **Whiteness Tests:** checking the residuals of the model for serial and cross-correlation
- **Consistency Test:** testing whether the model generates data with same correlation structure as the real data
- **Stability Test:** checking the stability/stationarity of the model.

Whiteness Tests

- We can regard the VAR[p] model coefficients $\mathbf{A}^{(k)}$ as a filter which transforms innovations (random white noise), **E**(*t*), into observed, structured data **X**(*t*):

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

- Consequently, for coefficient estimates $\hat{\mathbf{A}}^{(k)}$, we can obtain the residuals

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

- If we have adequately modeled the data, the residuals should be small and uncorrelated (white). Correlation structure in the residuals means there is still some correlation structure in the data that has not been described by our model.

- Checking the whiteness of residuals typically involves testing whether the residual **autocorrelation** coefficients up to some desired lag *h* are sufficiently small to ensure that we cannot reject the null hypothesis of white residuals at some desired significance level.

Whiteness Tests

$$\mathbf{E}(t) = N(0, \mathbf{V})$$

$$C_l = \left\langle \hat{\mathbf{E}}(t)\hat{\mathbf{E}}'(t-l) \right\rangle$$ autocovariance at lag *l* ...

$$R_l = D^{-1} C_l D^{-1}$$ with correponding autocorrelation *R*

$$D = diag\left(\sqrt{diag(C_0)}\right)$$

$$R_h = (R_1, \ldots, R_h)$$ set of autocorrelations up to lag *h*

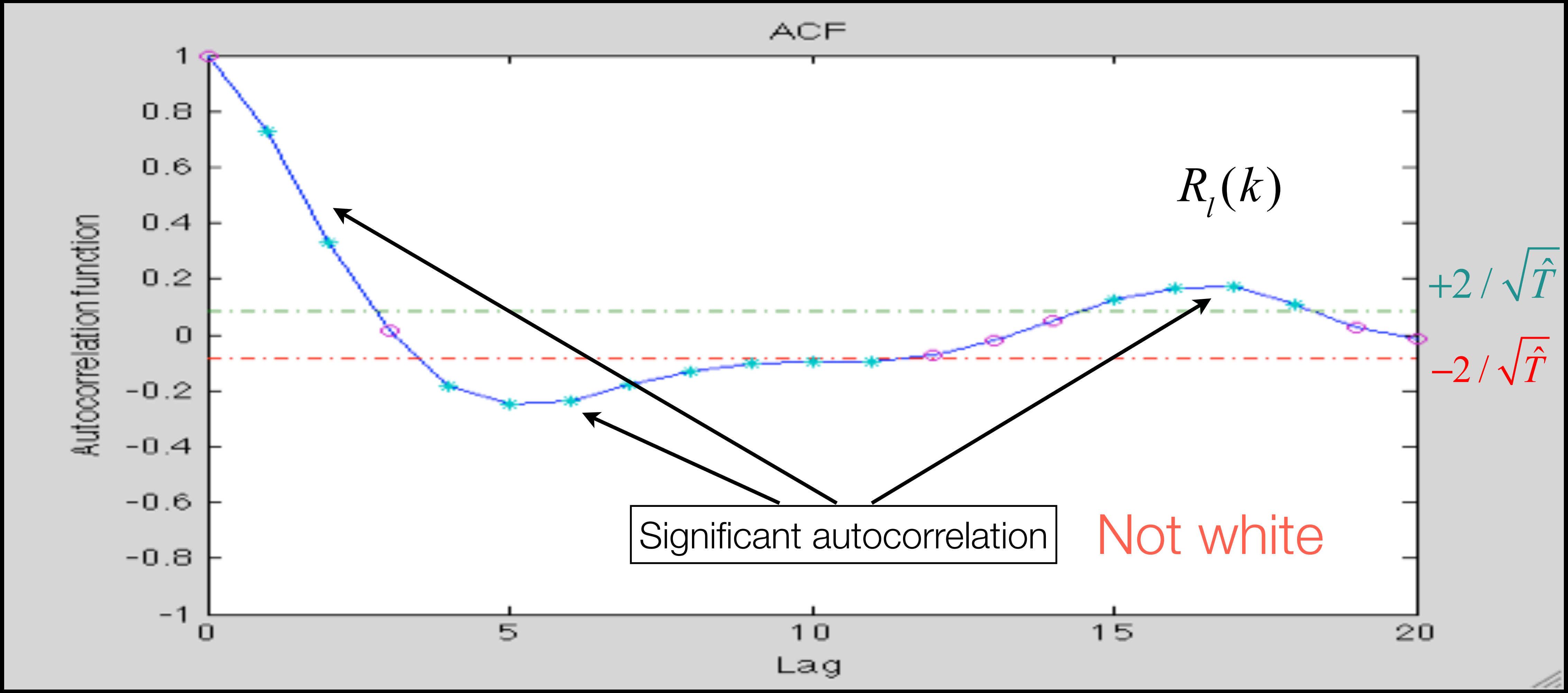
We want to test the null hypothesis $H_0 : \mathbf{R}_h = (R_1, \ldots, R_h) = 0$

against the alternative: $H_1 : \mathbf{R}_h \neq 0$

Two possible ways to do this:

- Autocorrelation function test
- Portmanteau tests

Whiteness Tests: ACF



Under the null hypothesis that **Ê**(*t*) is Gaussian white noise, we expect approximately 1/20=5% of a.c.f. coefficients to exceed the threshold $\pm 2/\sqrt{\hat{T}}$ This gives us a p-value for rejecting H_0

$$\rho = \frac{\text{count}\left(\left|\mathbf{R}_h\right| > 2/\sqrt{\hat{T}}\right)}{\text{count}\left(\mathbf{R}_h\right)} = \frac{\text{count}\left(\left|\mathbf{R}_h\right| > 2/\sqrt{\hat{T}}\right)}{M^2(h+1) - M}$$

If $p<0.05$ then we cannot reject H_0 at the 5% level and we accept that residuals **Ê**(*t*) are white

Whiteness Tests: Portmanteau

Table 3. Popular portmanteau tests for whiteness of residuals, implemented in SIFT. Here $\hat{T} = TN$ is the total number of samples used to estimate the covariance

Portmanteau Test	Formula (Test Statistic)	Notes
Box-Pierce (BPP)	$Q_h := \hat{T}\sum_{l=1}^{h}\mathrm{tr}\left(C_l' C_0^{-1} C_l C_0^{-1}\right)$	The original portmanteau test. Potentially overly-conservative. Poor small-sample properties.
Ljung-Box (LBP)	$Q_h := \hat{T}(\hat{T}+2)\sum_{l=1}^{h}(\hat{T}-l)^{-1}\mathrm{tr}\left(C_l' C_0^{-1} C_l C_0^{-1}\right)$	Modification of BPP to improve small-sample properties. Potentially inflates the variance of the test statistic. Slightly less conservative than LMP with slightly higher (but nearly identical) statistical power.
Li-McLeod (LMP)	$Q_h := \hat{T}\sum_{l=1}^{h}\mathrm{tr}\left(C_l' C_0^{-1} C_l C_0^{-1}\right) + \frac{M^2 h(h+1)}{2\hat{T}}$	Further modification of BPP to improve small-sample properties without variance inflation. Slightly more conservative than LBP. Probably the best choice in most conditions.

Mullen, 2010 (SIFT Manual)

These test statistics are asymptotically χ^2 distributed with $M^2(h-p)$ d.f.

Consistency Tests

- A well-fit model should be able to generate data that has the same correlation structure as the original data.
- One test of this is *percent consistency* (Ding et al, 2000)
- Here we generate simulated data from our fitted model (feeding it white noise) and calculate auto- and cross-correlations up to a fixed lag for both simulated data ($\mathbf{R}_s$) and real data ($\mathbf{R}_r$).
- The percent consistency (PC) is then given by

$$\mathrm{PC} = \left(1 - \frac{\left\|\mathbf{R}_s - \mathbf{R}_r\right\|_2}{\left\|\mathbf{R}_r\right\|_2}\right) \times 100$$

- A PC value near 100% indicates that the model is able to generate data that has a nearly identical correlation structure as the original data. A PC value near 0% indicates a complete failure to model the data.

Granger Causality

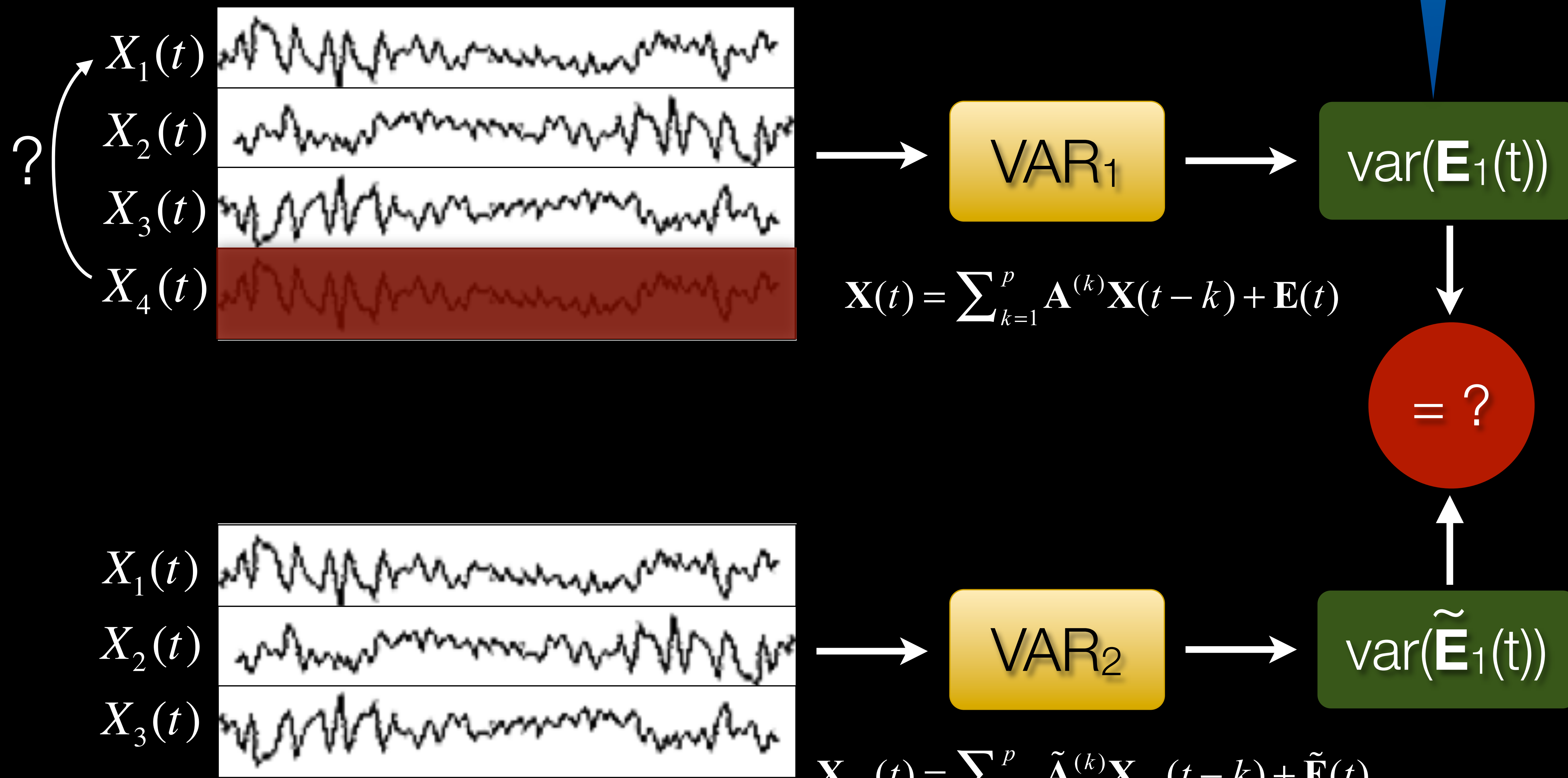
- First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- Relies on two assumptions:

Granger Causality Axioms
1. Causes should precede their effects in time (Temporal Precedence)
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)

Granger Causality

Does $\mathbf{X}_4$ granger-cause $\mathbf{X}_1$?
(conditioned on $\mathbf{X}_2$, $\mathbf{X}_3$)

prediction error for X_1
(variance of residuals E_1)



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

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

Granger Causality

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

reduced model

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

full model

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

Definition 1

X_j granger-causes X_i conditioned on all other variables in $\mathbf{X}$

if and only if $\mathbf{A}_{ij}(k) >> 0$ for some lag $k \in \{1, \ldots, p\}$

Granger Causality Quiz

- Example: 2-channel VAR process of order 1

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

$$X_1(t) = -0.5 X_1(t-1) + 0 X_2(t-1) + E_1(t)$$

$$X_2(t) = 0.7 X_1(t-1) + 0.2 X_2(t-1) + E_2(t)$$

Which causal structure does this model correspond to?

a) 1 ⟶ 2 b) 1 ⟵ 2 c) 1 ⟷ 2

Granger Causality – Frequency Domain

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

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

Likewise, **X**(*f*) and **E**(*f*) correspond to the fourier transforms of the data and residuals, respectively

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

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

$$\mathbf{X}(f) = \mathbf{A}(f)^{-1}\mathbf{E}(f) = \mathbf{H}(f)\mathbf{E}(f)$$

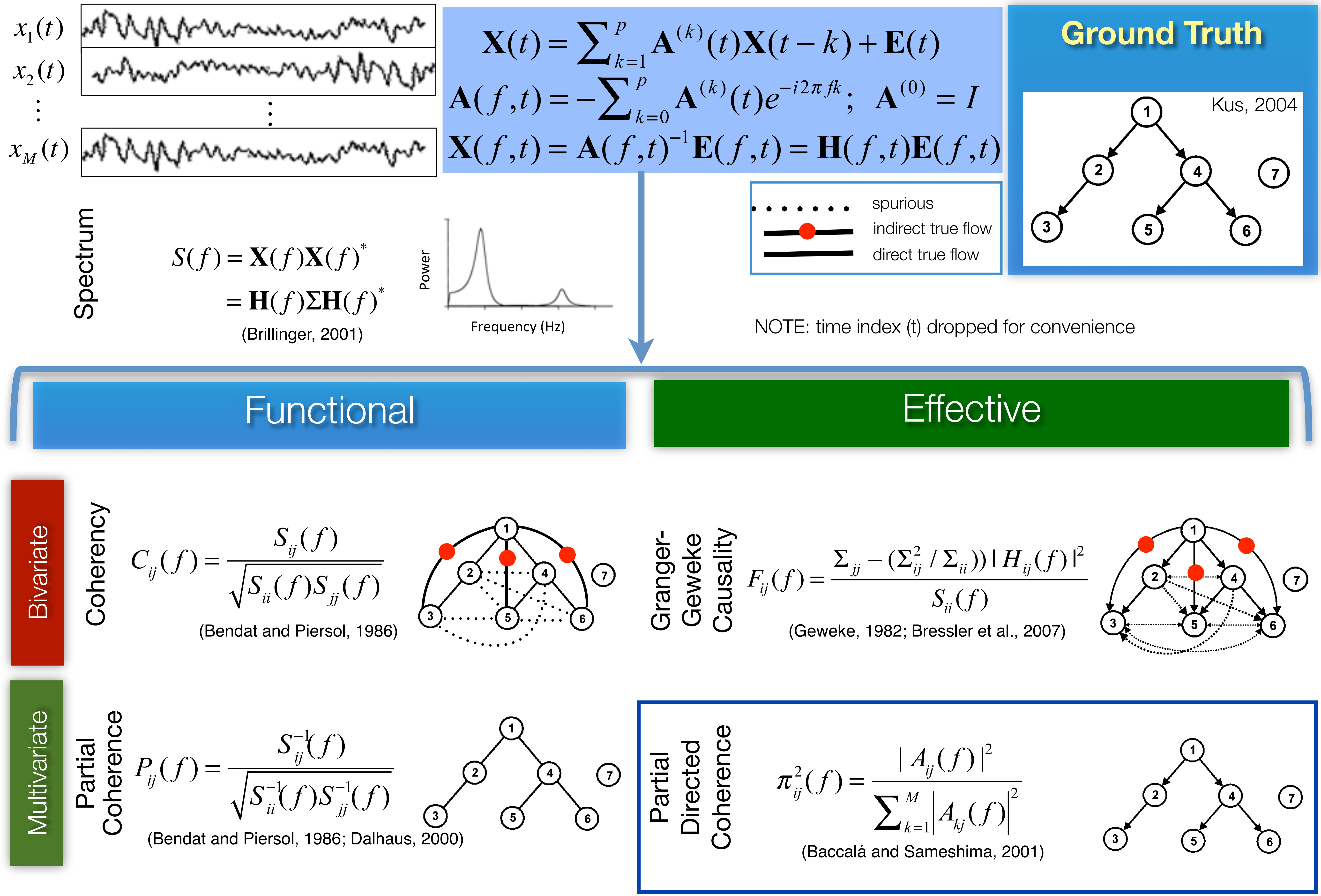
Where **H**(*f*) is the *transfer matrix* of the system.

Definition 2

X_j granger-causes X_i conditioned on all other variables in **X**

if and only if $|\mathbf{A}_{ij}(f)| >> 0$ for some frequency f

leads to PDC



$x_1(t)$
$x_2(t)$
$x_M(t)$
$\mathbf{X}(t)=\sum_{k=1}^{p}\mathbf{A}^{(k)}(t)\mathbf{X}(t-k)+\mathbf{E}(t)$
$\mathbf{A}(f,t)=-\sum_{k=0}^{p}\mathbf{A}^{(k)}(t)e^{-i2\pi fk};\quad \mathbf{A}^{(0)}=I$
$\mathbf{X}(f,t)=\mathbf{A}(f,t)^{-1}\mathbf{E}(f,t)=\mathbf{H}(f,t)\mathbf{E}(f,t)$
Ground Truth
Kus, 2004
spurious
indirect true flow
direct true flow
Spectrum
$S(f)=\mathbf{X}(f)\mathbf{X}(f)^{*}$
$=\mathbf{H}(f)\Sigma\mathbf{H}(f)^{*}$
(Brillinger, 2001)
Power
Frequency (Hz)
NOTE: time index (t) dropped for convenience
Functional
Effective
Bivariate
Coherency
$C_{ij}(f)=\frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$
(Bendat and Piersol, 1986)
Granger-Geweke Causality
$F_{ij}(f)=\frac{\Sigma_{jj}-(\Sigma_{ij}^{2}/\Sigma_{ii}))\,|H_{ij}(f)|^{2}}{S_{ii}(f)}$
(Geweke, 1982; Bressler et al., 2007)
Multivariate
Partial Coherence
$P_{ij}(f)=\frac{S_{ij}^{-1}(f)}{\sqrt{S_{ii}^{-1}(f)S_{jj}^{-1}(f)}}$
(Bendat and Piersol, 1986; Dalhaus, 2000)
Partial Directed Coherence
$\pi_{ij}^{2}(f)=\frac{|A_{ij}(f)|^{2}}{\sum_{k=1}^{M}\left|A_{kj}(f)\right|^{2}}$
(Baccalá and Sameshima, 2001)

	Estimator	Formula
Spectral M.	Spectral Density Matrix	$S(f)=X(f)X(f)^* = H(f)\Sigma H(f)^*$
Coherence Measures	Coherency	$C_{ij}(f)=\dfrac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$ $0\le\lvert C_{ij}(f)\rvert^2\le 1$
Coherence Measures	Imaginary Coherence (iCoh)	$iCoh_{ij}(f)=\mathrm{Im}(C_{ij}(f))$
Coherence Measures	Partial Coherence (pCoh)	$P_{ij}(f)=\dfrac{\hat{S}_{ij}(f)}{\sqrt{\hat{S}_{ii}(f)\hat{S}_{jj}(f)}}$ $\hat{S}(f)=S(f)^{-1}$ $0\le\lvert P_{ij}(f)\rvert^2\le 1$
Coherence Measures	Multiple Coherence (mCoh)	$G_i(f)=\sqrt{1-\dfrac{\det(S(f))}{S_{ii}(f)\mathbf{M}_{ii}(f)}}$ $\mathbf{M}_{ii}(f)$ is the **minor** of *S(f)* obtained by removing the i[th] row and column of *S(f)* and returning the determinant.

	Estimator	Formula
Partial Directed Coherence Measures	Normalized Partial Directed Coherence (PDC)	$\pi_{ij}(f)=\dfrac{A_{ij}(f)}{\sqrt{\sum_{k=1}^{M}\lvert A_{kj}(f)\rvert^2}}$ $0\le\lvert\pi_{ij}(f)\rvert^2\le 1$ $\sum_{j=1}^{M}\lvert\pi_{ij}(f)\rvert^2=1$
Partial Directed Coherence Measures	Generalized PDC (GPDC)	$\bar{\pi}_{ij}(f)=\dfrac{\frac{1}{\Sigma_{ii}}A_{ij}(f)}{\sqrt{\sum_{k=1}^{M}\frac{1}{\Sigma_{ii}^2}\lvert A_{kj}(f)\rvert^2}}$ $0\le\lvert\bar{\pi}_{ij}(f)\rvert^2\le 1$ $\sum_{j=1}^{M}\lvert\bar{\pi}_{ij}(f)\rvert^2=1$
Partial Directed Coherence Measures	Renormalized PDC (rPDC)	$\lambda_{ij}(f)=Q_{ij}(f)^* V_{ij}(f)^{-1}Q_{ij}(f)$ where $Q_{ij}(f)=\begin{pmatrix}\mathrm{Re}[A_{ij}(f)]\\ \mathrm{Im}[A_{ij}(f)]\end{pmatrix}$ and $V_{ij}(f)=\sum_{k,l=1}^{p}R_{jj}^{-1}(k,l)\Sigma_{ii}Z(2\pi f,k,l)$ $Z(\omega,k,l)$ $=\begin{pmatrix}\cos(\omega k)\cos(\omega l) & \cos(\omega k)\sin(\omega l)\\ \sin(\omega k)\cos(\omega l) & \sin(\omega k)\sin(\omega l)\end{pmatrix}$ ***R*** is the $[(Mp)^2 \times (Mp)^2]$ covariance matrix of the VAR[*p*] process (Lütkepohl, 2006)
Granger-Geweke	Granger-Geweke Causality (GGC)	$F_{ij}(f)=\dfrac{(\Sigma_{jj}-(\Sigma_{ij}^2/\Sigma_{ii}))\lvert H_{ij}(f)\rvert^2}{S_{ii}(f)}$

	Estimator	Formula
Directed Transfer Function Measures	Normalized Directed Transfer Function (DTF)	$\gamma_{ij}(f)=\dfrac{H_{ij}(f)}{\sqrt{\sum_{k=1}^{M}\lvert H_{ik}(f)\rvert^2}}$ $0\le\lvert\gamma_{ij}(f)\rvert^2\le 1$ $\sum_{j=1}^{M}\lvert\gamma_{ij}(f)\rvert^2=1$
Directed Transfer Function Measures	Full-Frequency DTF (ffDTF)	$\eta_{ij}^2(f)=\dfrac{\lvert H_{ij}(f)\rvert^2}{\sum_f\sum_{k=1}^{M}\lvert H_{ik}(f)\rvert^2}$
Directed Transfer Function Measures	Direct DTF (dDTF)	$\delta_{ij}^2(f)=\eta_{ij}^2(f)P_{ij}^2(f)$

$$\mathbf{X}(t)=\sum_{k=1}^{p}\mathbf{A}^{(k)}(t)\mathbf{X}(t-k)+\mathbf{E}(t)$$
$$\mathbf{A}(f,t)=-\sum_{k=0}^{p}\mathbf{A}^{(k)}(t)e^{-i2\pi fk};\quad \mathbf{A}^{(0)}=I$$
$$\mathbf{X}(f,t)=\mathbf{A}(f,t)^{-1}\mathbf{E}(f,t)=\mathbf{H}(f,t)\mathbf{E}(f,t)$$

$H(f)$ Transfer Function
$A(f)$ System Matrix
Σ Noise Covariance Matrix

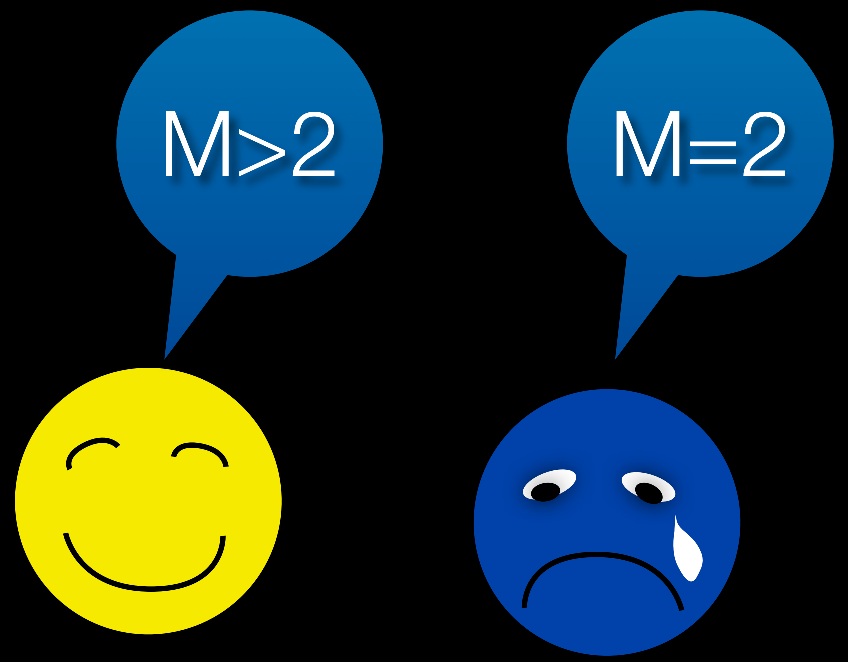
Variance Stabilization

For additional details, see SIFT Handbook (sccn.ucsd.edu/wiki/SIFT)

INTERMISSION

Multivariate versus Bivariate

- Exclusion of processes that may exert causal influence on modeled processes increases the risk of causal mis-identification. (c.f. Pearl, *Causality: Models, Inference and Reasoning,* 2009)
- Multivariate approaches are generally superior to bivariate approaches
  - allow detection of direct versus indirect dependence, reducing false positives
  - allow us to partially control for exogenous/ unobserved causes (e.g. Guo, et al., J. Neuro. Methods, 2008)



- In the absence of *a priori* knowledge concerning causal structure, it is advisable to include as many processes as possible in a causal model *(within data/modeling limitations)*

Multivariate Models: Limitations

- However, multivariate methods come with a cost:
  - More parameters + limited data = higher risk of over-fitting or worse yet....
  - ...the problem becomes *ill-posed* or *under-determined*. There are insufficient observations to uniquely determine a solution to the system of equations defining our model.

Multivariate Models: Limitations

How many samples do we need?

- N = number of samples required
- M = number of variables/sources
- T = number of trials/realizations
- p = model order
- We have M^2p model coefficients to estimate. So our ordinary least-squares solution requires a *minimum* of M^2p samples.

$$N = O(M^2p)$$

- Back-of-envelope: M=20, p=10, T=1 We need 20^2x10 = 4000 samples -- 20 second epoch at sampling rate of 200Hz!

Ensemble aggregation (T > 1)?

- M=20, p=10, T=50: 4000/50 samples/trial → 20/50 = 0.4 sec epoch

Multivariate Models: Constraints

Solutions?

Make assumptions (impose constraints)

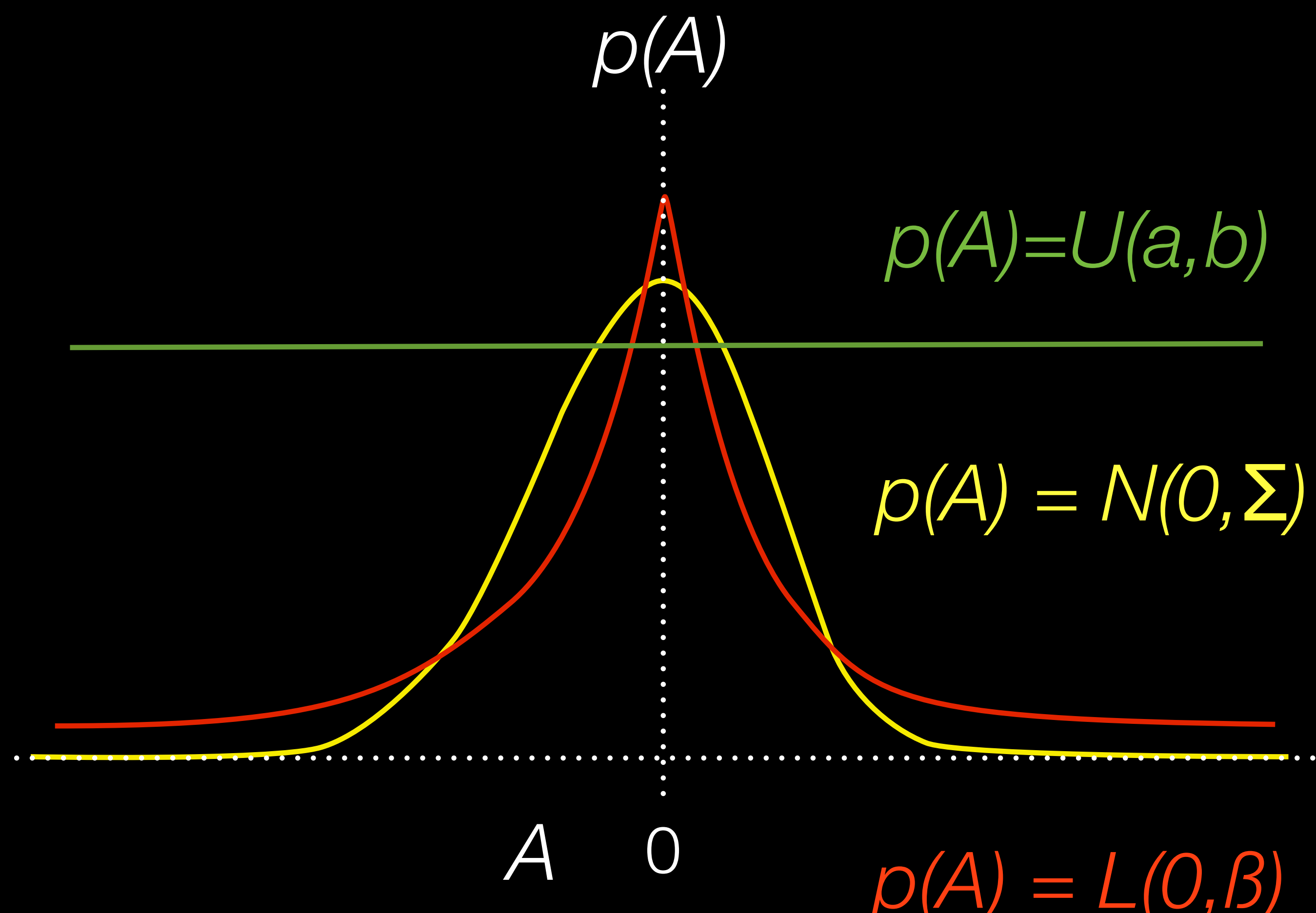
We want to *a priori* restrict the range of allowable values for our parameters -- transforming the problem from one with infinite number of solutions in the original parameter space to one with a unique (“best”) solution in the new parameter space

In a Bayesian context, this corresponds to making assumptions about the *prior distribution* of the parameters (Gaussian, Laplacian, ...)

Multivariate Models: Constraints

posterior likelihood prior

$$\hat{A} = \arg\max_{A} \left\{ p(A|D) \cong p(D|A)\, \boxed{p(A)} \right\}$$



Unconstrained (all values equally probable). e.g. Uniform distribution

Smoothness constraints
- large differences in values unlikely
- small (non-zero) values most probable. e.g. Normal (gaussian) prior.

Sparsity constraint
- many values small or exactly zero with occasional large values e.g. Laplacian prior

Smoothness Constraints

- Standard least-squares solution

prediction error

$$A(t) = \arg\min_{\hat{A}} \left(\left\| Y - Z\tilde{A} \right\|_2^2 \right)$$

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

$$\tilde{A} = [A^{(1)}(t), \ldots, A^{(p)}(t)]^T$$

$$X_k = [X(p+1-k), \ldots, X(N-k)]^T$$

$$Z = [X_1, \ldots, X_p]$$

$$Y = X_0$$

Rewrite VAR[p] as VAR[1]

Smoothness Constraints

- Ridge Regression
(Tikhonov Regularization, Minimum-(L_2)-Norm Estimation, ...)

prediction error

penalty term,
enforces smoothness

$$A(t) = \arg\min_{\hat{A}} \left(\left\| Y - Z\tilde{A} \right\|_2^2 + \lambda \left\| \tilde{A} \right\|_2^2 \right)$$

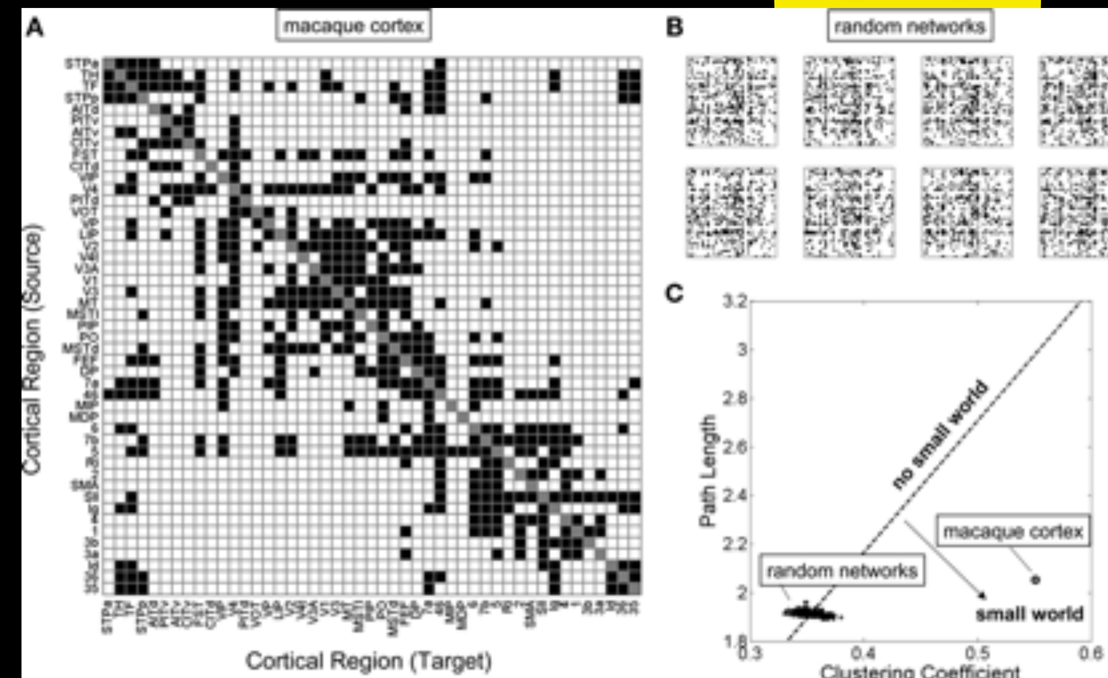
regularization

- Equivalent to assuming a Gaussian prior with variance determined by λ
- Large values of *A* are penalized. The range of allowable values for coefficients is restricted, reducing the *effective* degrees of freedom and allowing us to estimate VAR coefficients with fewer observations.

Sparsity Constraints

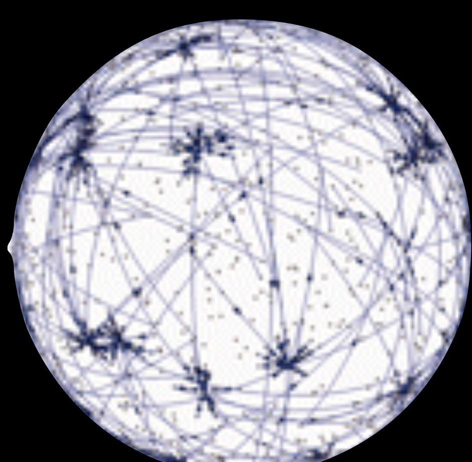
Sparsity

- Relatively low probability of a *direct* connection between any two anatomical functional units. This probability decreases with distance

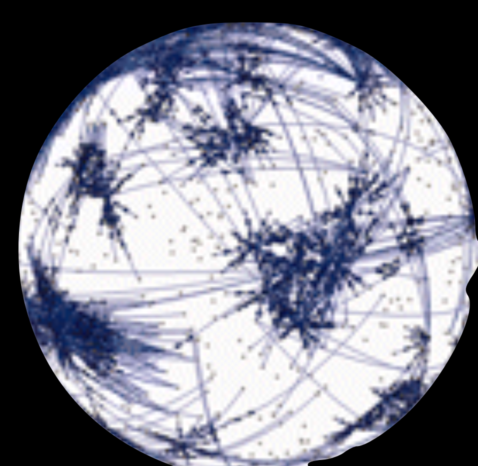


Sporns, *Frontiers in Computational Neuroscience*, 2011

It's a small world...

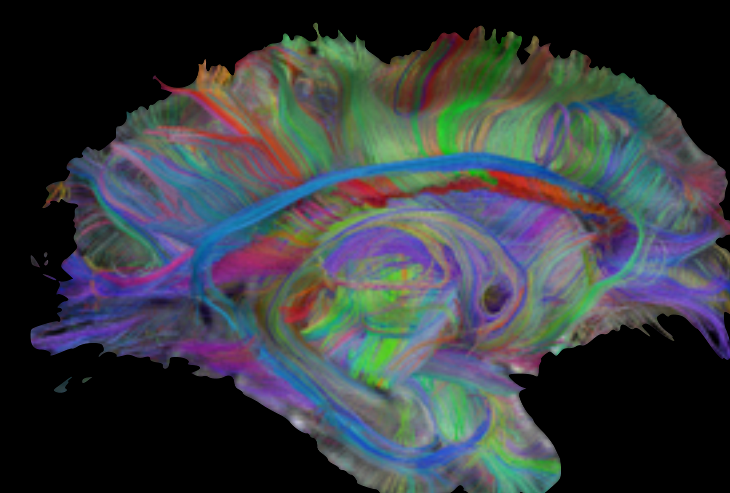


structural network



functional network

Sporns and Honey, *PNAS*, 2006



Structural Connectivity

Sparsity Constraints

- Standard least-squares solution

prediction error

$$A(t) = \arg\min_{\hat{A}} \left(\left\| Y - Z\tilde{A} \right\|_2^2 \right)$$

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

$$\tilde{A} = [A^{(1)}(t), \ldots, A^{(p)}(t)]^T$$

$$X_k = [X(p+1-k), \ldots, X(N-k)]^T$$

$$Z = [X_1, \ldots, X_p]$$

$$Y = X_0$$

Rewrite VAR[p] as VAR[1]

Sparsity Constraints

- Group Lasso ($L_{1,2}$ norm)

prediction error

smoothness (L2)
(preserves spectrum)

$$A(t) = \arg\min_{\hat{A}} \left(\left\| Y - Z\tilde{A} \right\|_2^2 + \lambda \sum_{ij} \left\| \tilde{A}_{ij}^{(1)}, \ldots, \tilde{A}_{ij}^{(p)} \right\|_2^2 \right)$$

regularization

group sparsity (L1)

ADMM
DAL

- Equivalent to assuming a Gaussian prior over coefficients within groups and a Laplacian prior over the groups themselves
- Entire groups of coefficients are jointly pruned (set *exactly* to zero) while remaining groups Allowing us to estimate VAR coefficients with fewer observations.

Sparsity Constraints

Compressive Sensing

- The process of acquiring and reconstructing a quantity that is underdetermined but known to be sparse (compressible) in some basis

How many samples do we need?

- N = number of samples required
- *M = number of variables/sources, p = model order*

$$N = O\left(K \log(M^2 p / K)\right) \approx O\left(\log M^2 p\right)$$

$N = O(M^2p)$
(unconstrained)

Constraints Improve Estimation (if prior assumptions are correct)

- Significant improvements using smoothness or sparsity assumptions
- (e.g. Haufe et al, 2009, Valdez-Sosa et al, 2009)

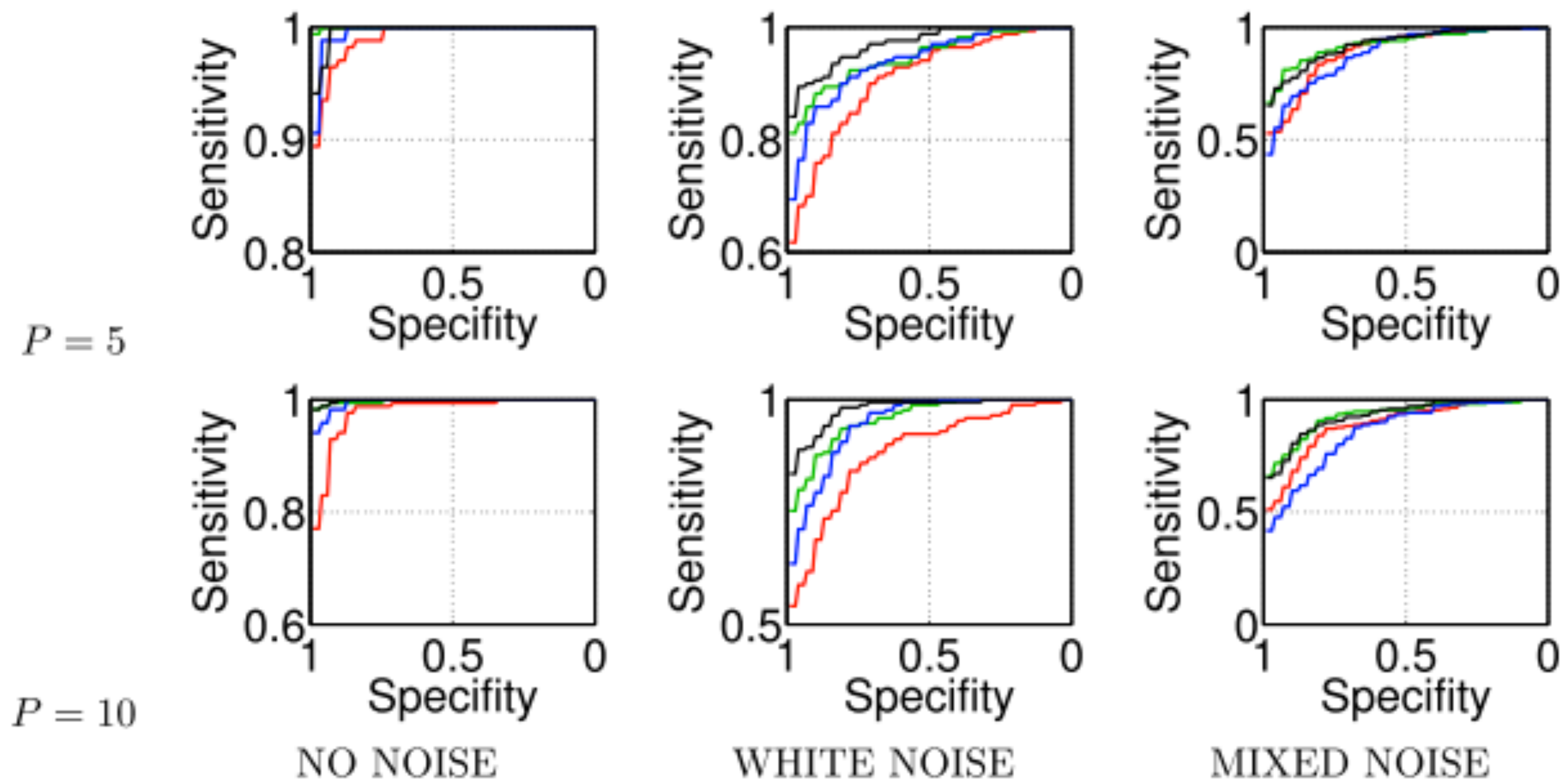
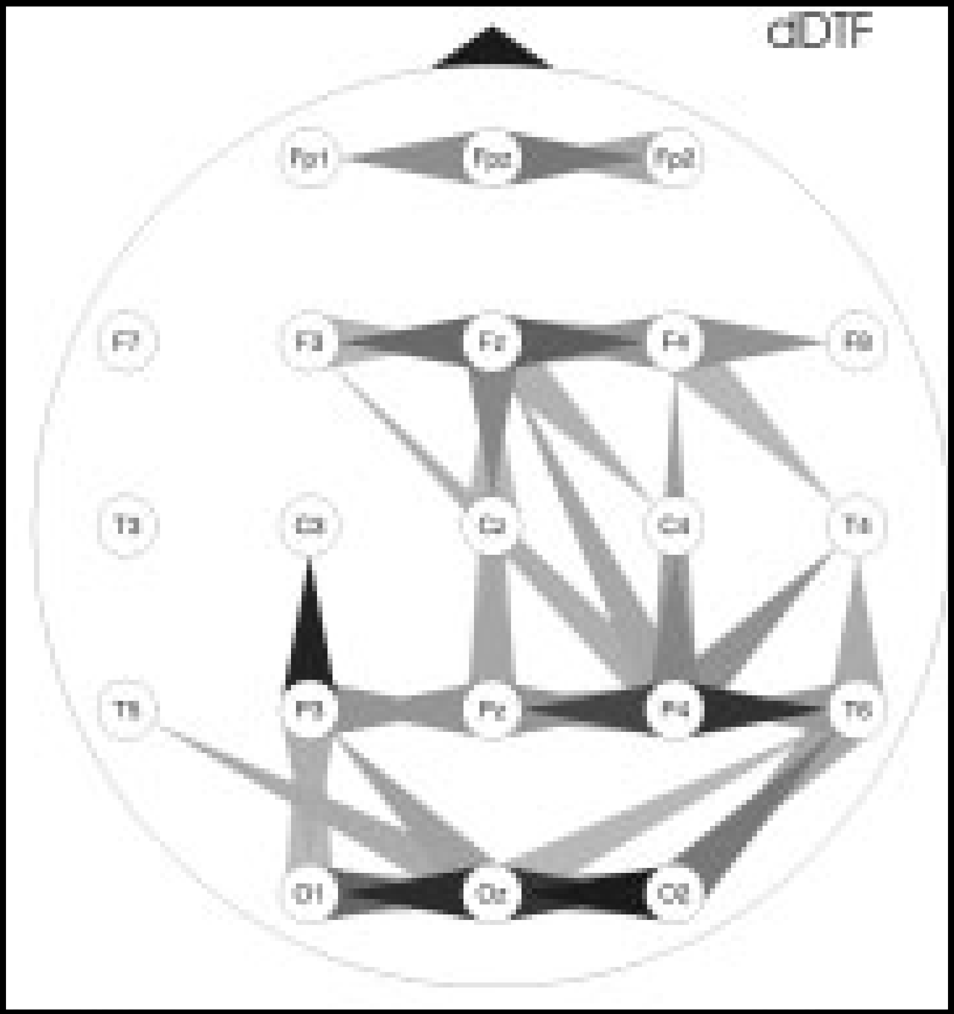


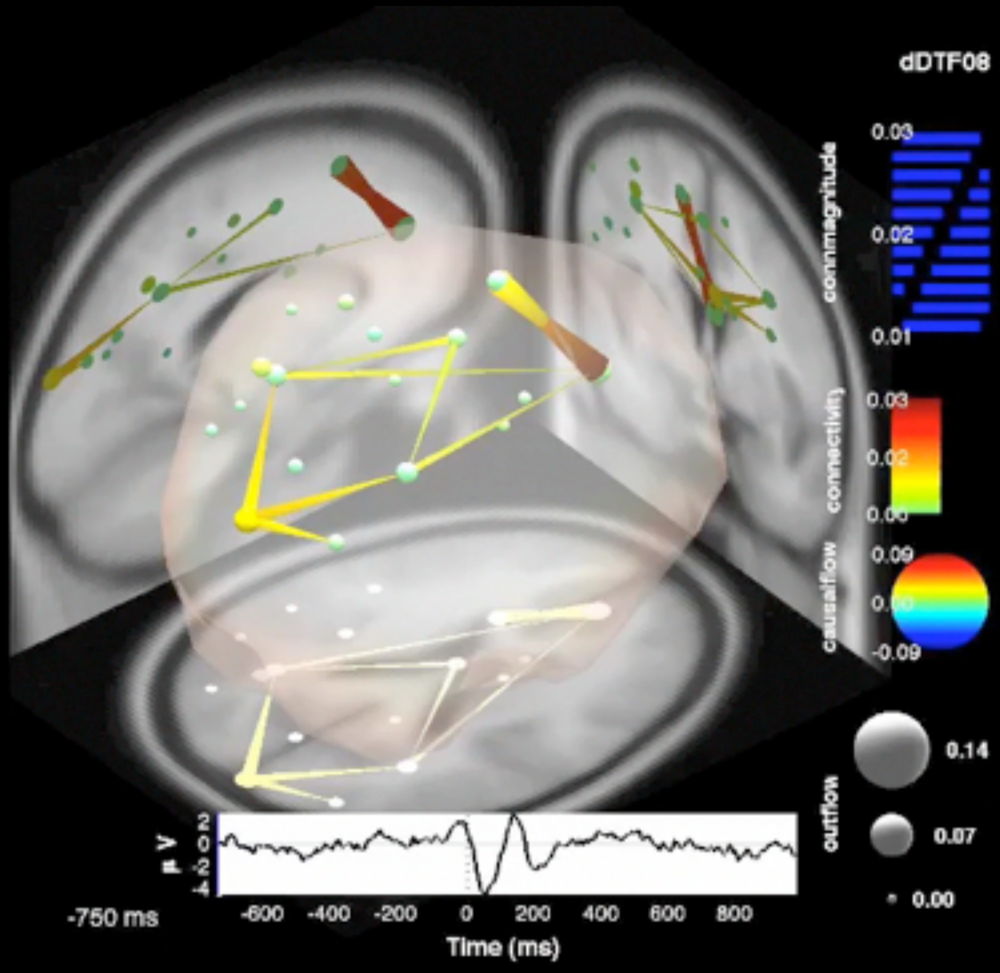
Figure 2: Average ROC curves of Granger Causality (red), Ridge Regression (green), Lasso (blue) and Group Lasso (black) in three different noise conditions and for two different model orders.

Haufe, 2009

Scalp or Source?



or

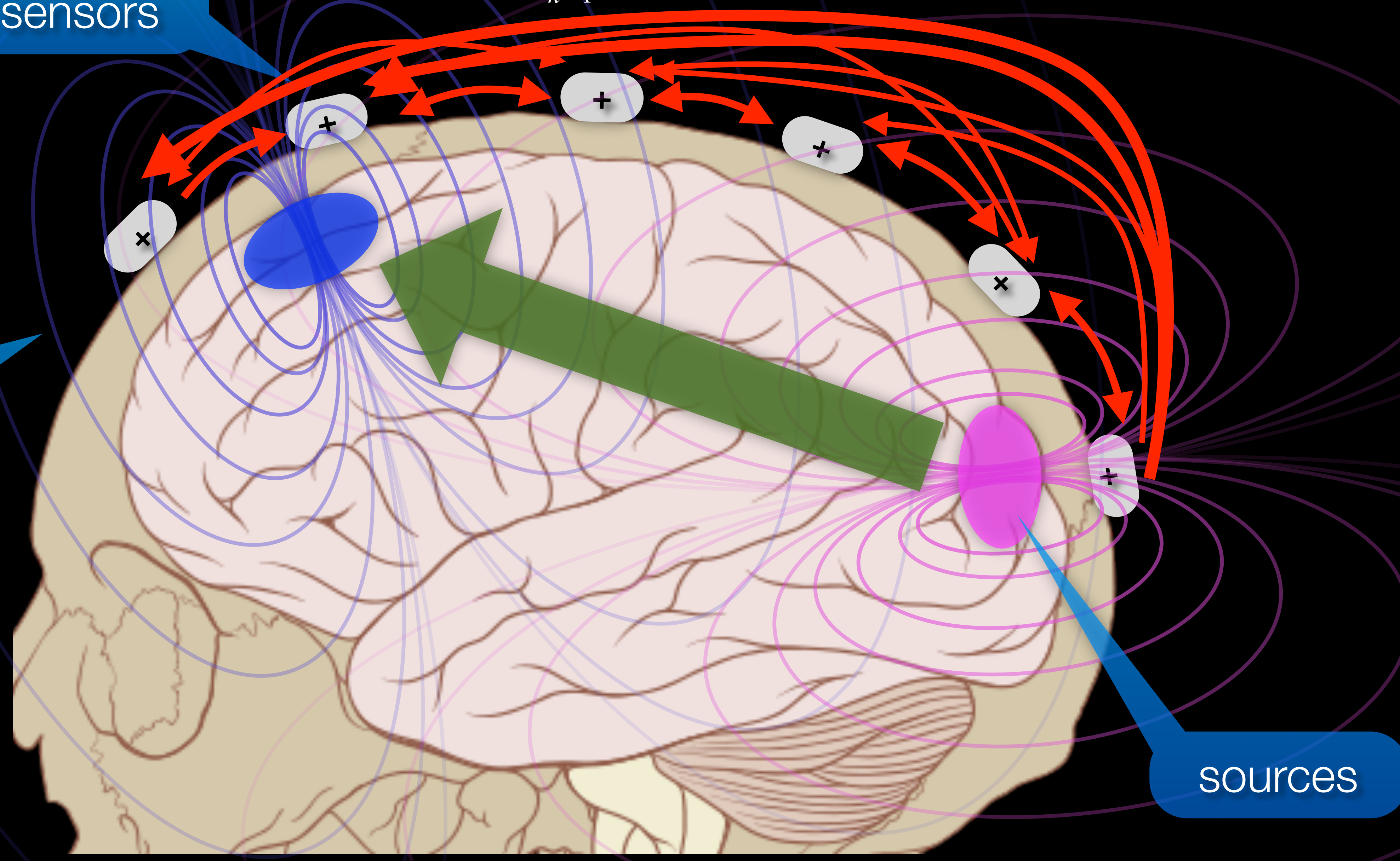


Scalp or Source?

$$X(t) = HS(t) = \sum_{k=1}^{p} HA^{(k)}(t)H^{-1}X(t-k) + HE(t)$$

$$X(t) = HS(t)$$

$$H^{-1}$$

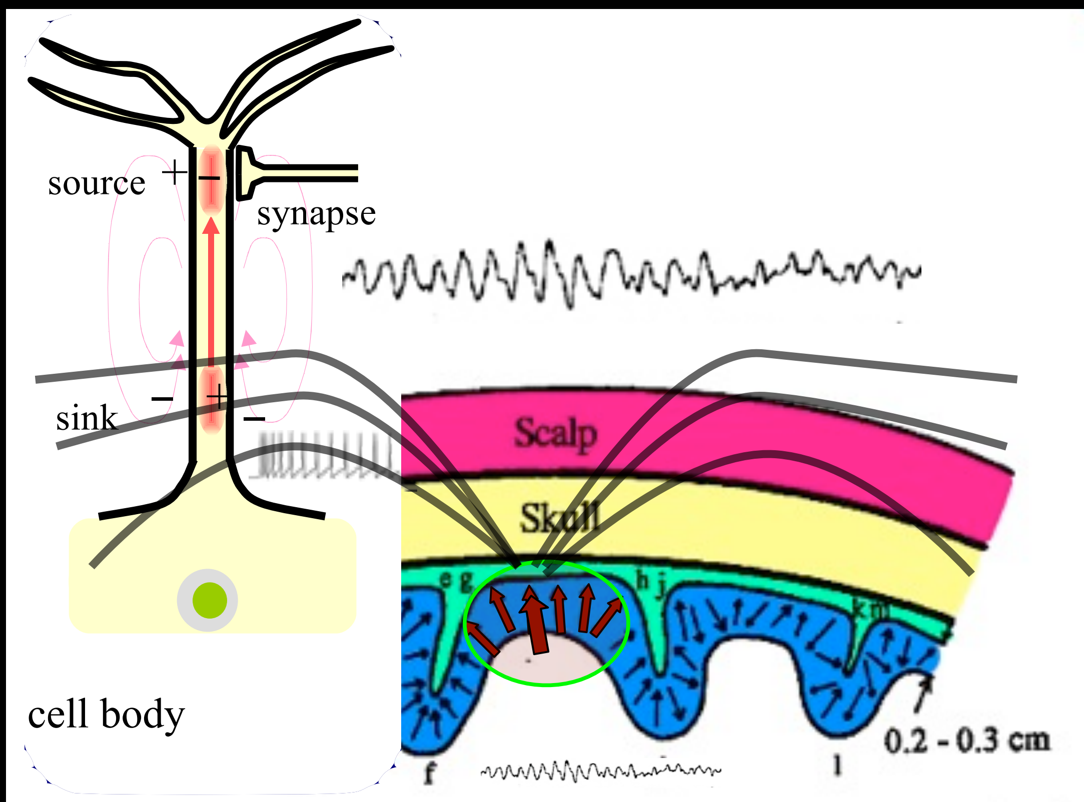


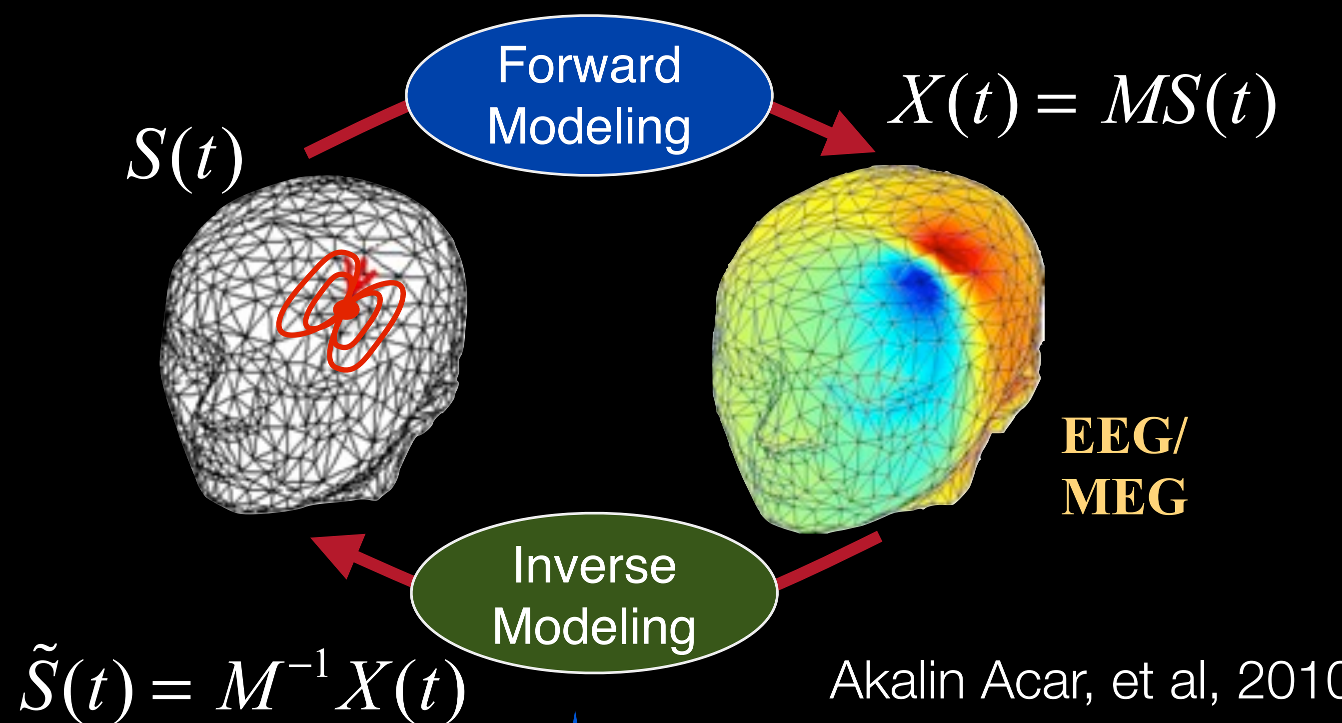
ICA
SBL
Beamforming
Minimum-norm
...

Solution? Source Separation

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

Forward/Inverse Modeling





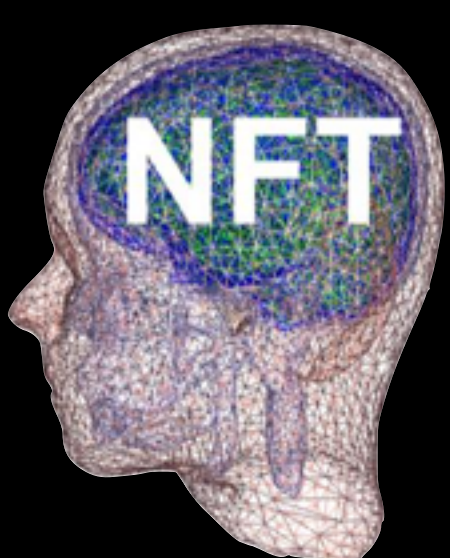
Akalin Acar, et al, 2010

ill-posed!

solutions?

impose constraints!

sparse/smooth
independence
anatomy
...



A Recipe for Reducing Errors:

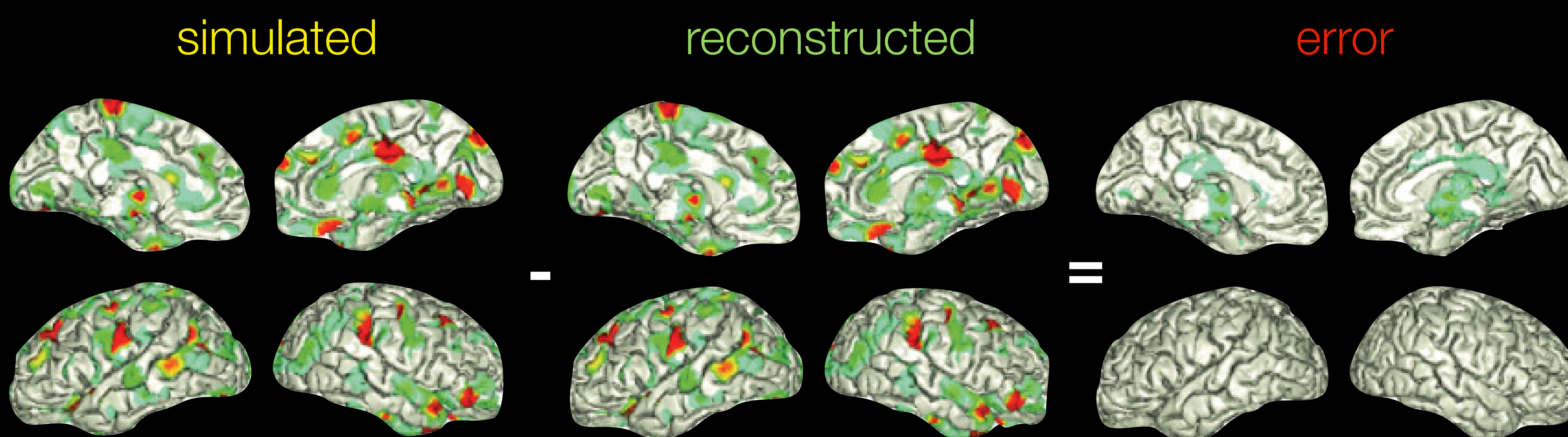
- Realistic Forward Model
- Appropriately Constrained Inverse Model

Akalin Acar and Makeig, 2009

Forward/Inverse Modeling

Method	Smoothness	Sparsity	Independence/Orthogonality
MNE	X		
LORETA	X		
dSPM	X		
Beamforming			X
Sparse Bayesian Learning	X	X	
S-FLEX	X	X	
FOCUSS		X	
ICA/PCA/SOBI			X

Source reconstruction with ICA+SBL



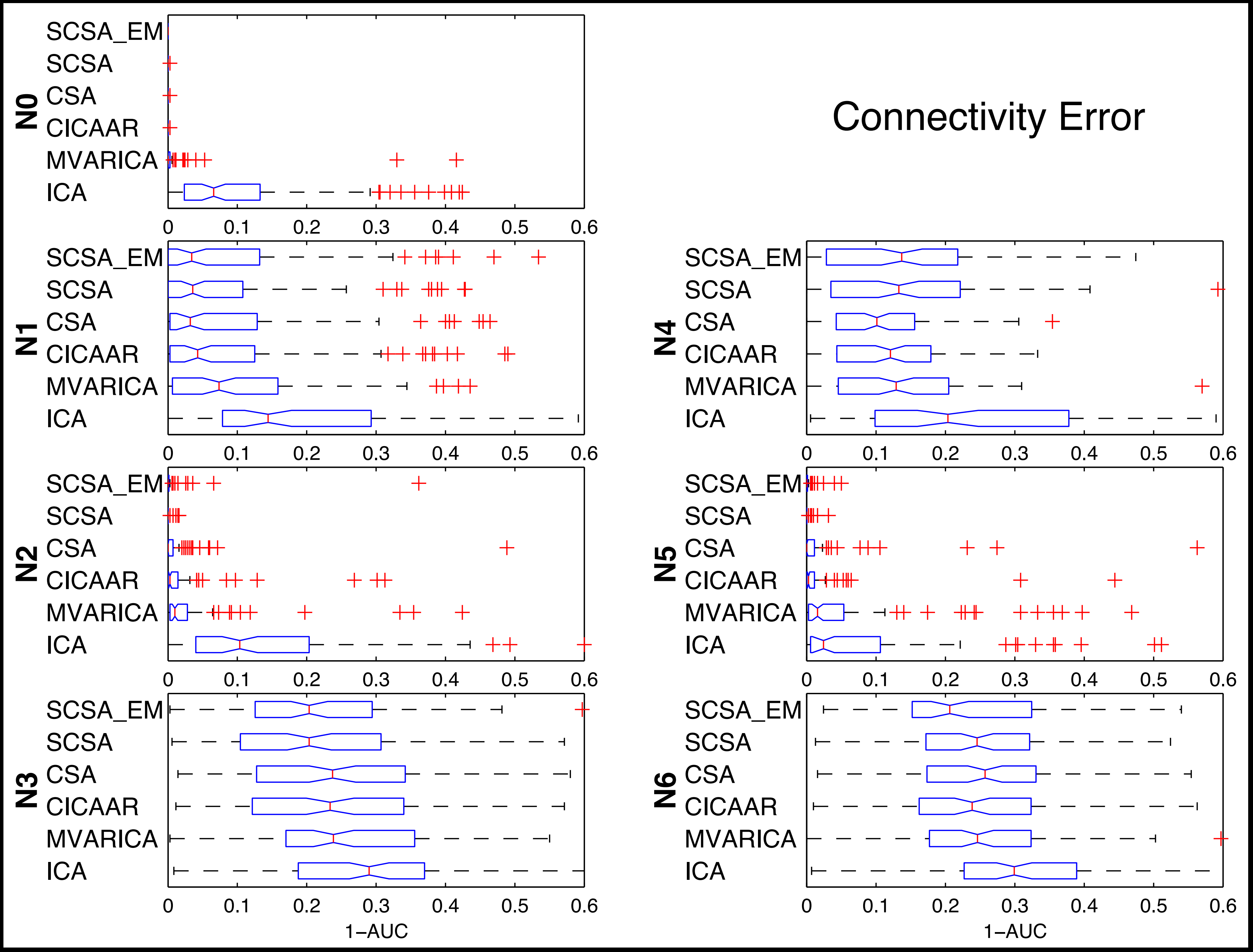
Makeig, Ramirez, Weber, Wipf, Dale, Simpson, *15th Inter. Conf on Biomagnetism* (2006)

Estimating Dependency of Independent Components ?

- Isn't it a contradiction to examine dependence between Independent/ Uncorrelated Components?
- Instantaneous (e.g., Infomax) ICA only explicitly seeks to maximize *instantaneous* independence. Time-delayed dependencies may be preserved.
- Infomax ICA seeks to maximize *global* independence (over entire recording session), transient dependencies may be preserved.
- Independence is a very strict criterion that cannot be achieved *in general* by a linear transformation (such as ICA). Instead, dependent variables will form a **dependent subspace**.

However, the *best* approach is to use an inverse model that explicitly preserves time-delayed dependencies or *jointly* estimates sources (de-mixing matrix) and connectivity (VAR parameters). See Haufe, 2008 IEEE TBME for a good treatment (coming soon to SIFT).

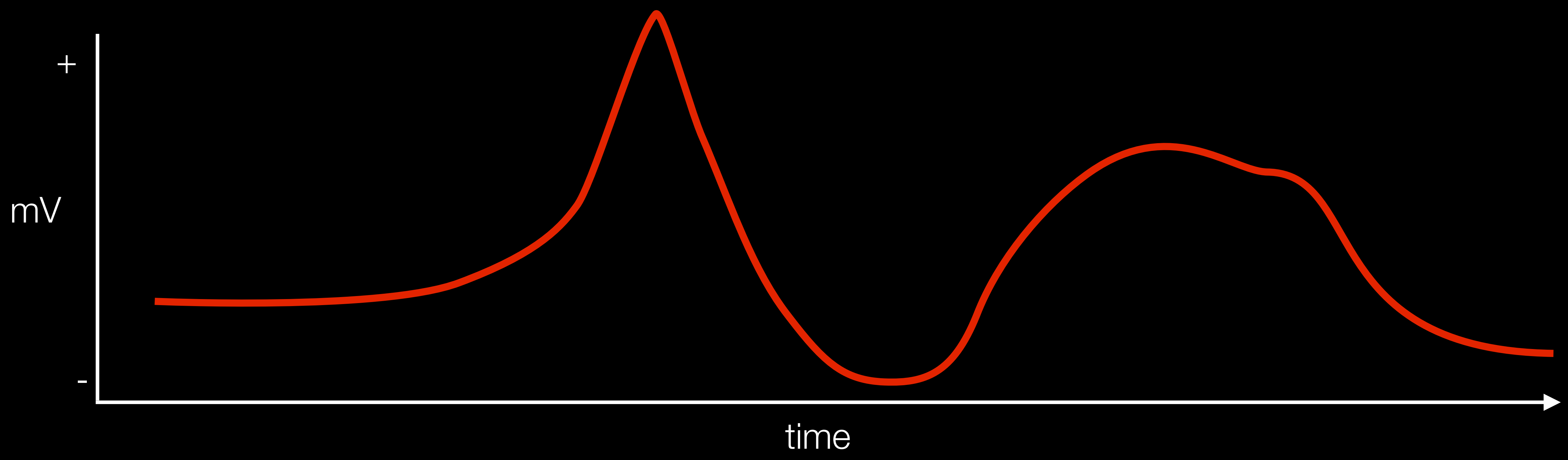
Estimating Dependency of Independent Components ?



Haufe et al, IEEE TBME 2008

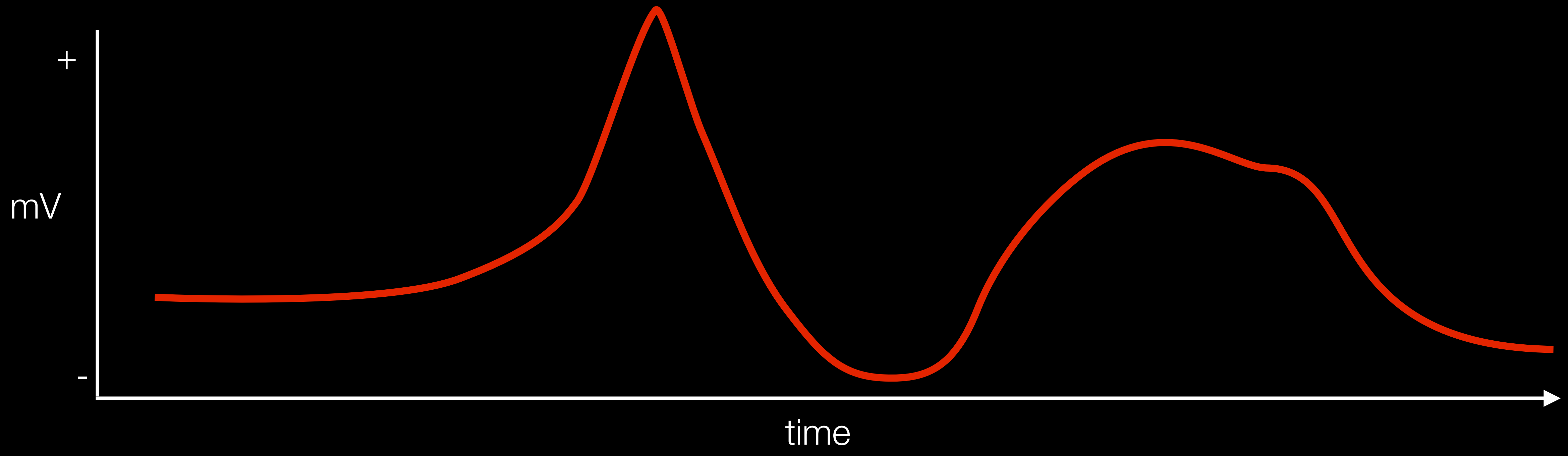
Adapting to Non-Stationarity

- The brain is a **dynamic system** and measured brain activity and coupling can change rapidly with time (non-stationarity)
  - event-related perturbations (ERSP, ERP, etc)
  - structural changes due to learning/feedback
- How can we adapt to non-stationarity?



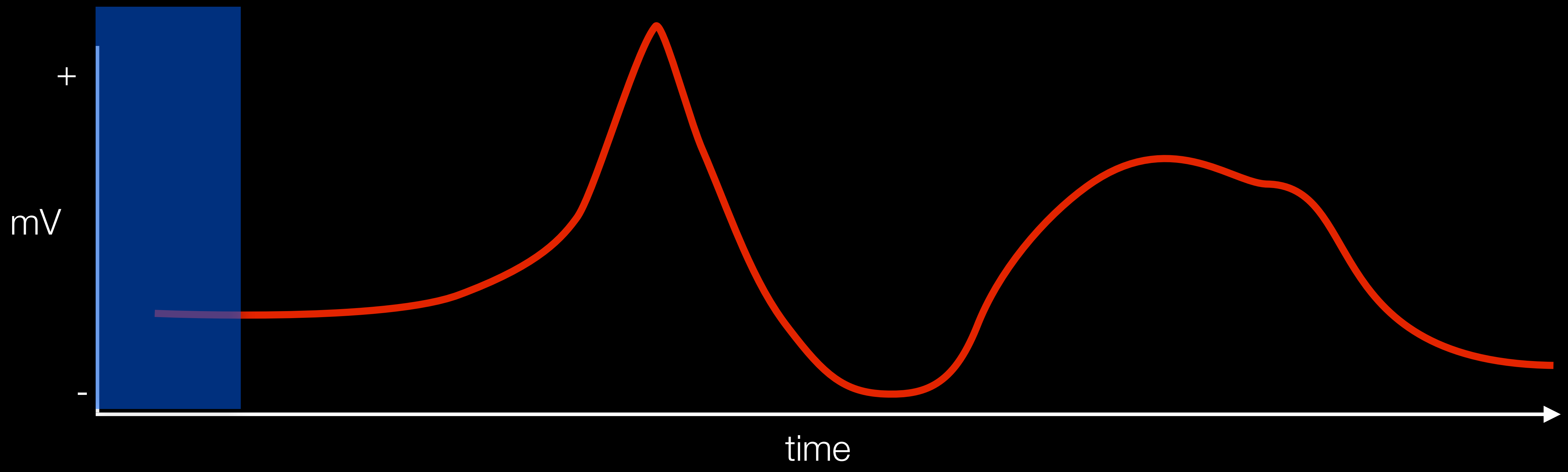
Adapting to Non-Stationarity

- **Many ways to do adaptive VAR estimation**
- Two popular approaches (adopted in SIFT):
  - Segmentation-based adaptive VAR estimation (assumes local stationarity)
  - State-Space Modeling



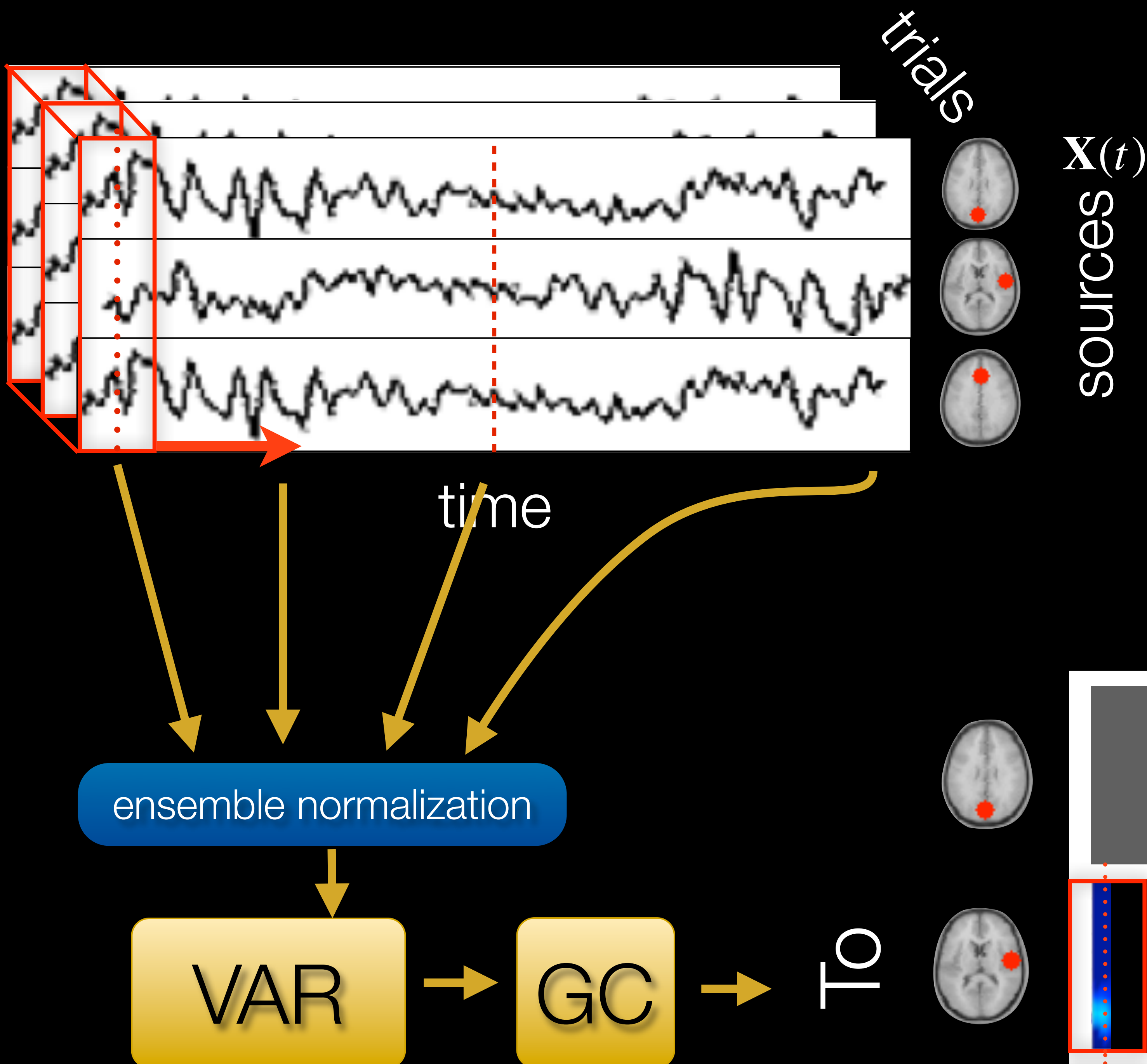
Adapting to Non-Stationarity

- **Many ways to do adaptive VAR estimation**
- Two popular approaches (adopted in SIFT):
  - **Segmentation-based adaptive VAR estimation (assumes local stationarity)**
  - State-Space Modeling

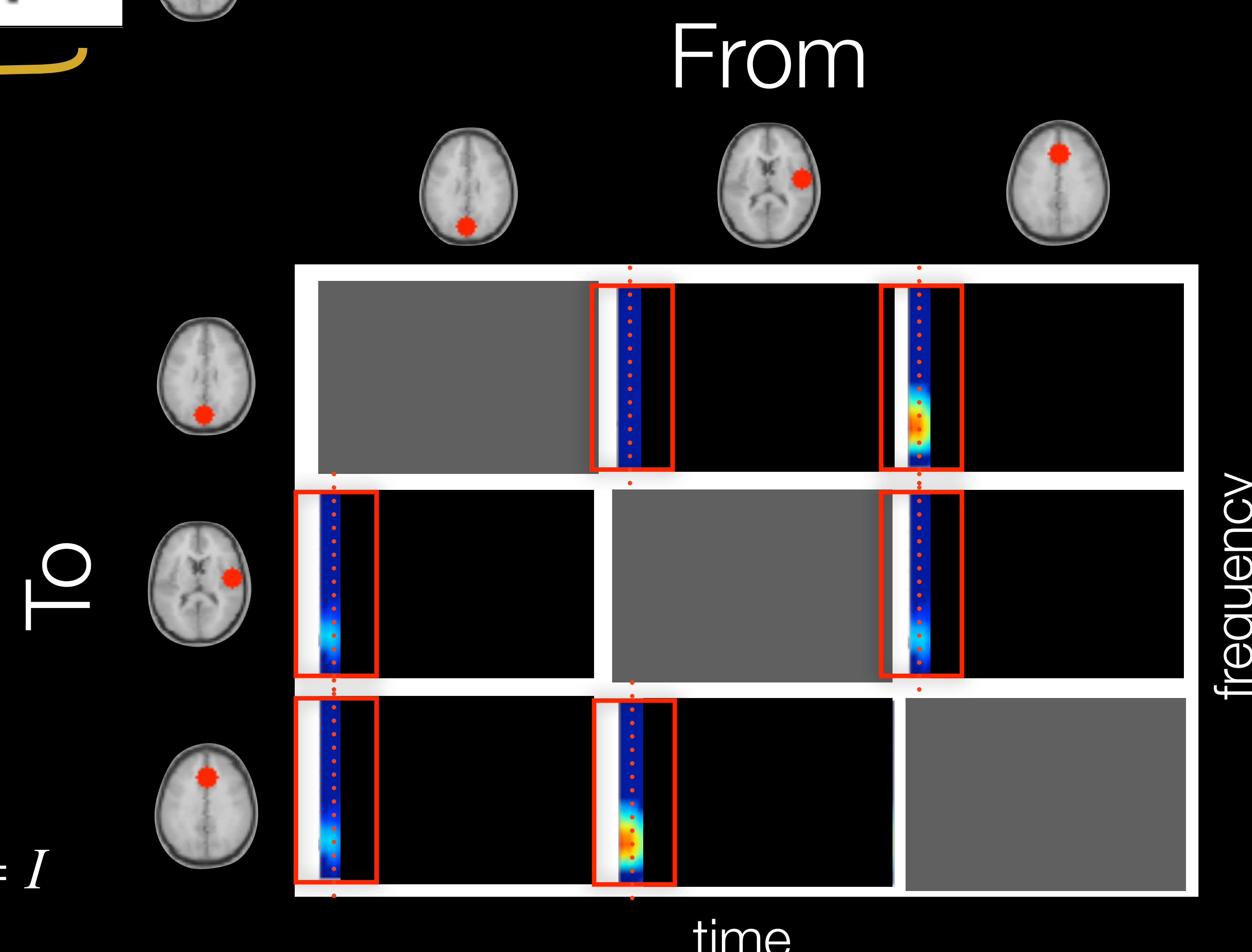


Segmentation-based VAR

(Jansen et al., 1981; Florian and Pfurtscheller, 1995; Ding et al, 2000)



Analogous to short-time Fourier transform



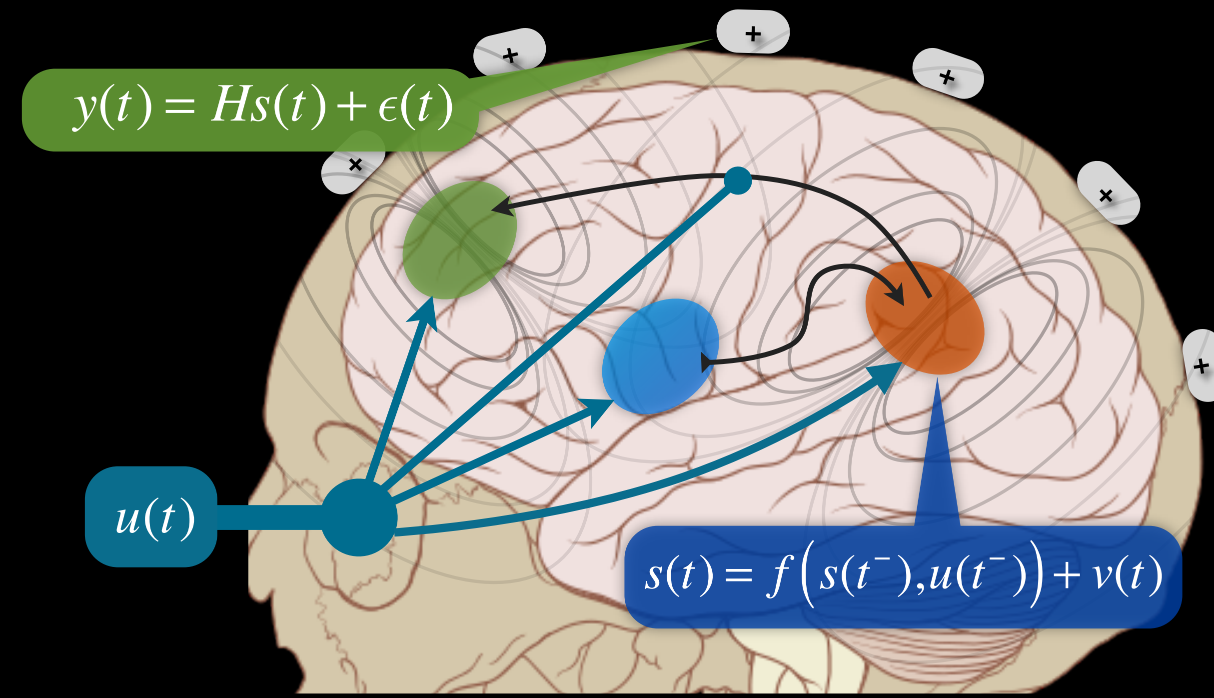
$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}(t)\mathbf{X}(t-k) + \mathbf{E}(t)$$

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

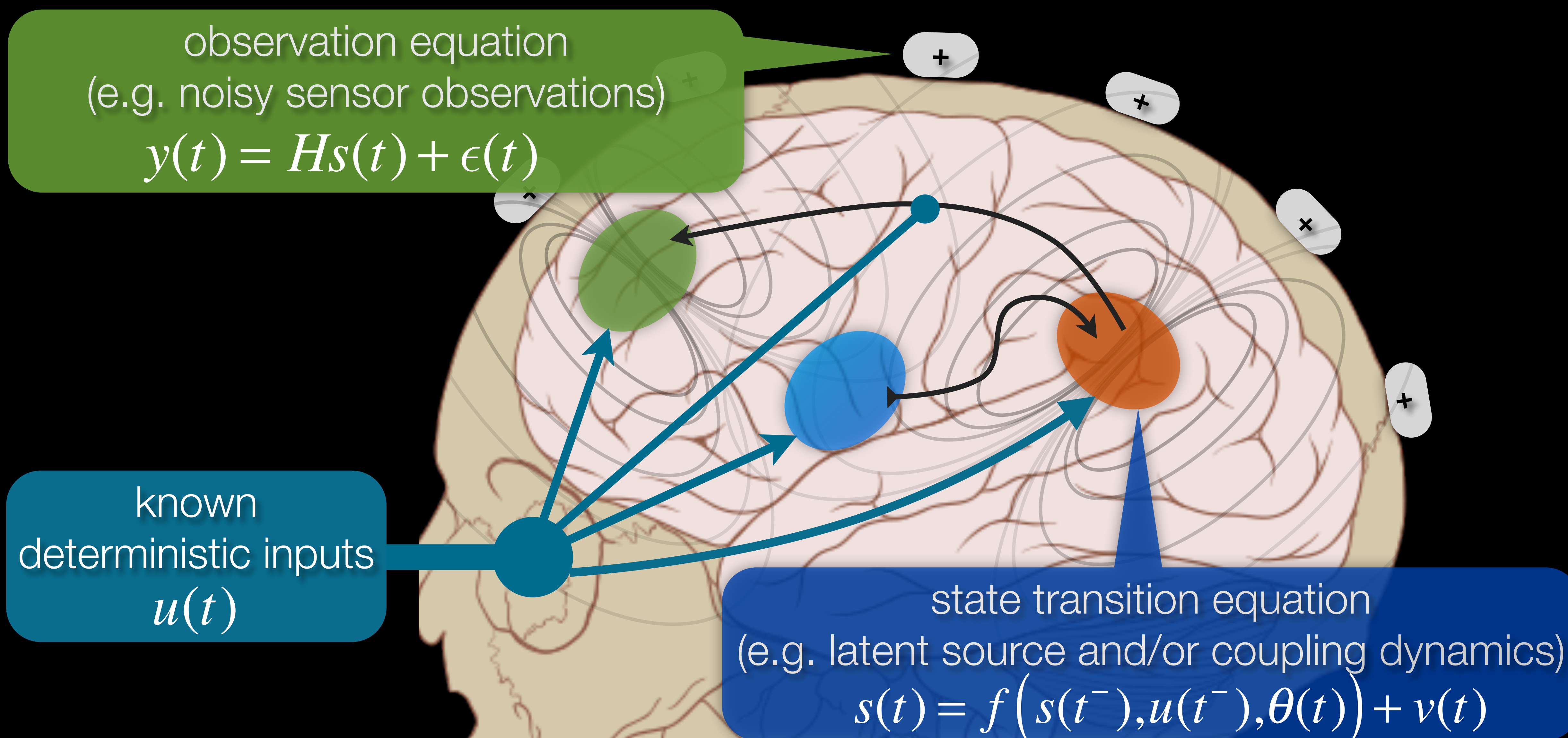
Adapting to Non-Stationarity

- **Many ways to do adaptive VAR estimation**
- Two popular approaches (adopted in SIFT):
  - Segmentation-based adaptive VAR estimation (assumes local stationarity)
  - **State-Space Modeling**

    Kalman Filtering and extensions



Discrete State-Space Model (SSM) for Electrophysiological Dynamics

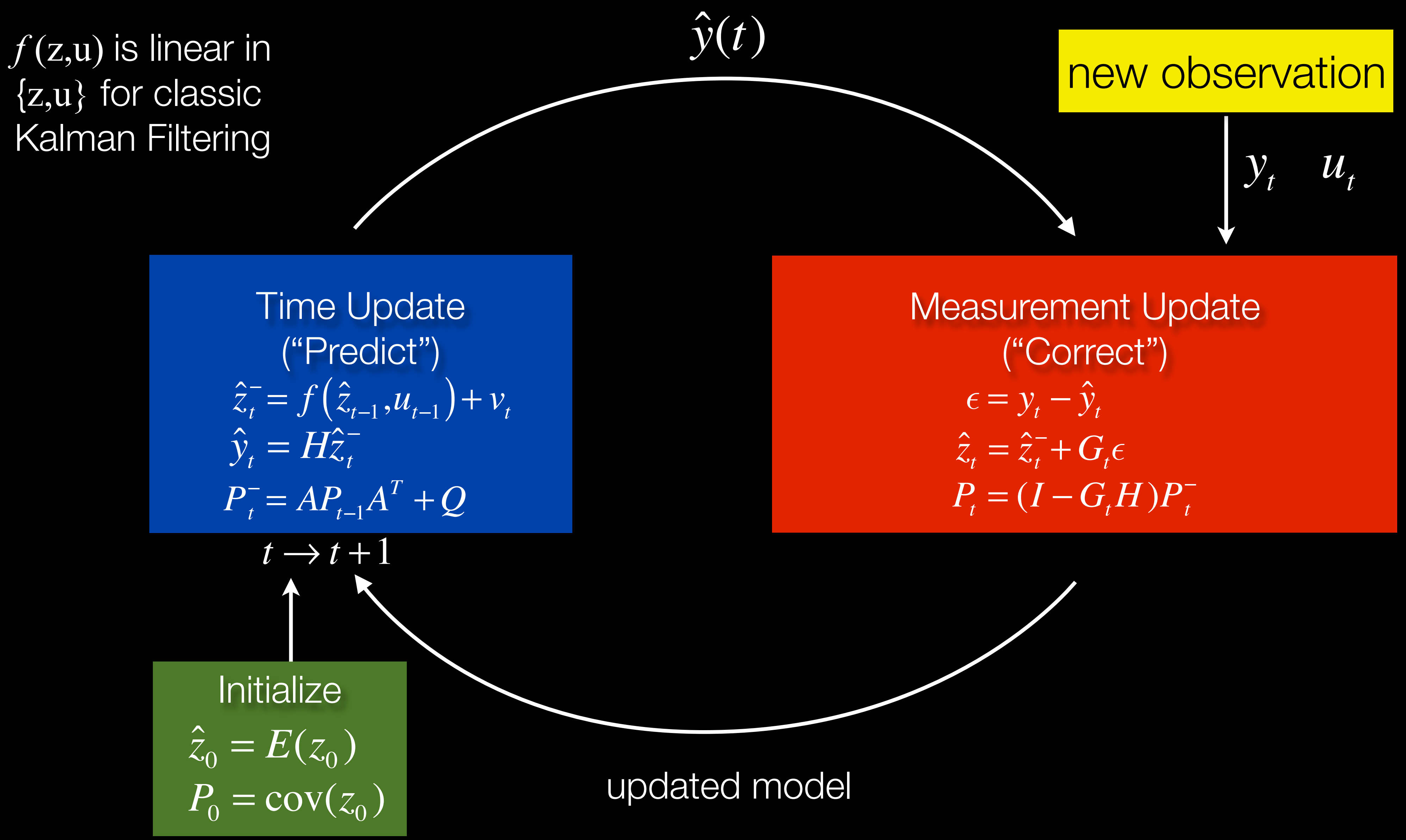


Linear VAR[1] $\;\; s(t) = \mathbf{A}(t)s(t-1) + v(t)$

- Dynamical system may be linear or nonlinear, dense or sparse, non-stationary, high-dimensional, partially-observed, and stochastic
- Subsumes discrete Delay Differential Equation (DDE) and Vector Autoregressive (VAR) methods and closely related to Dynamic Causal Modeling (DCM)

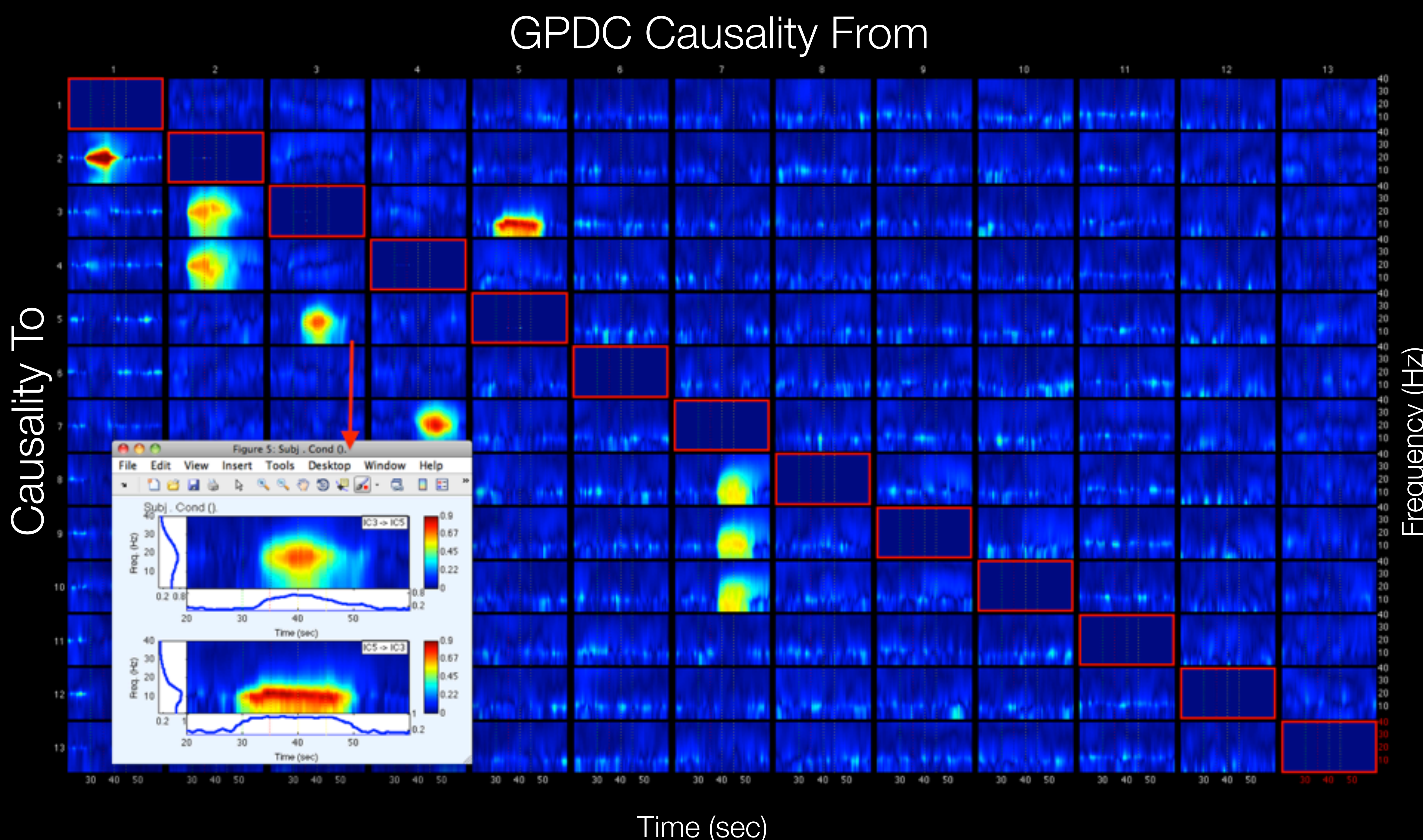
Kalman Filtering

optimal estimator (in terms of minimum variance) for the state of a linear dynamical system



z_t := unknown state vector at time *t*
e.g. delay-embedding of sources and/or coupling (VAR) parameters

Kalman Filtering



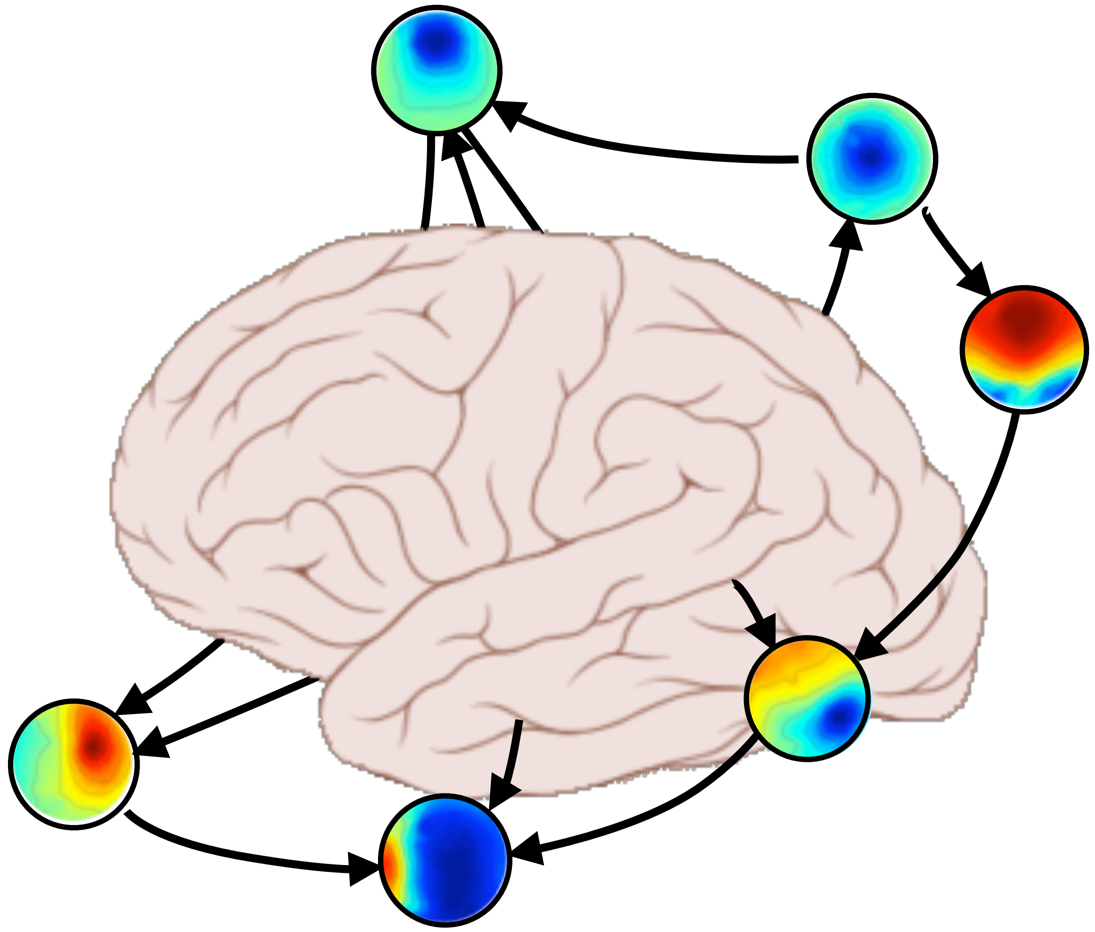
INTERMISSION

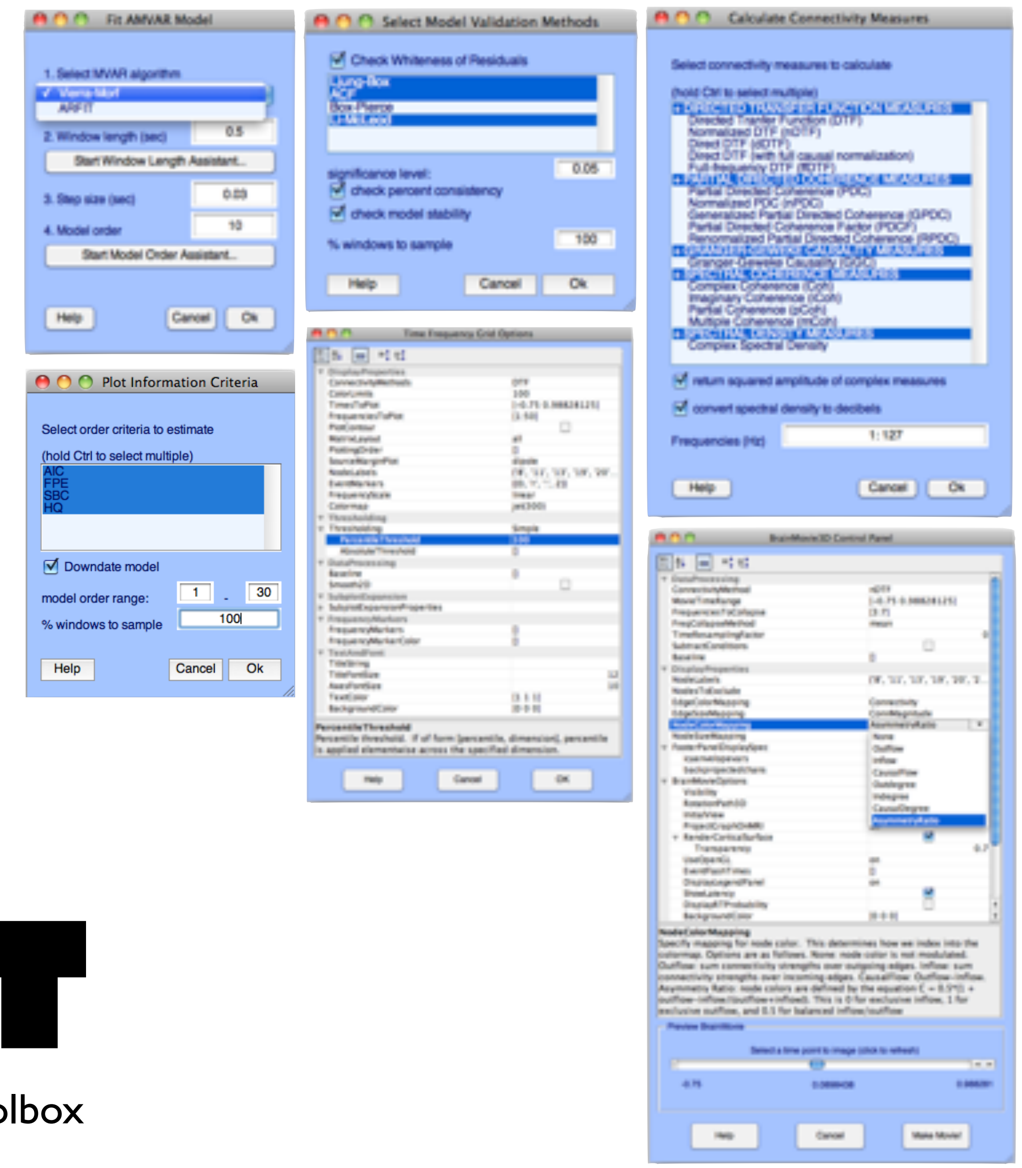
http://sccn.ucsd.edu/wiki/SIFT

Mullen, et al, *Journal of Neuroscience Methods* (in prep, 2012)

Mullen, et al, *Society for Neuroscience*, 2010

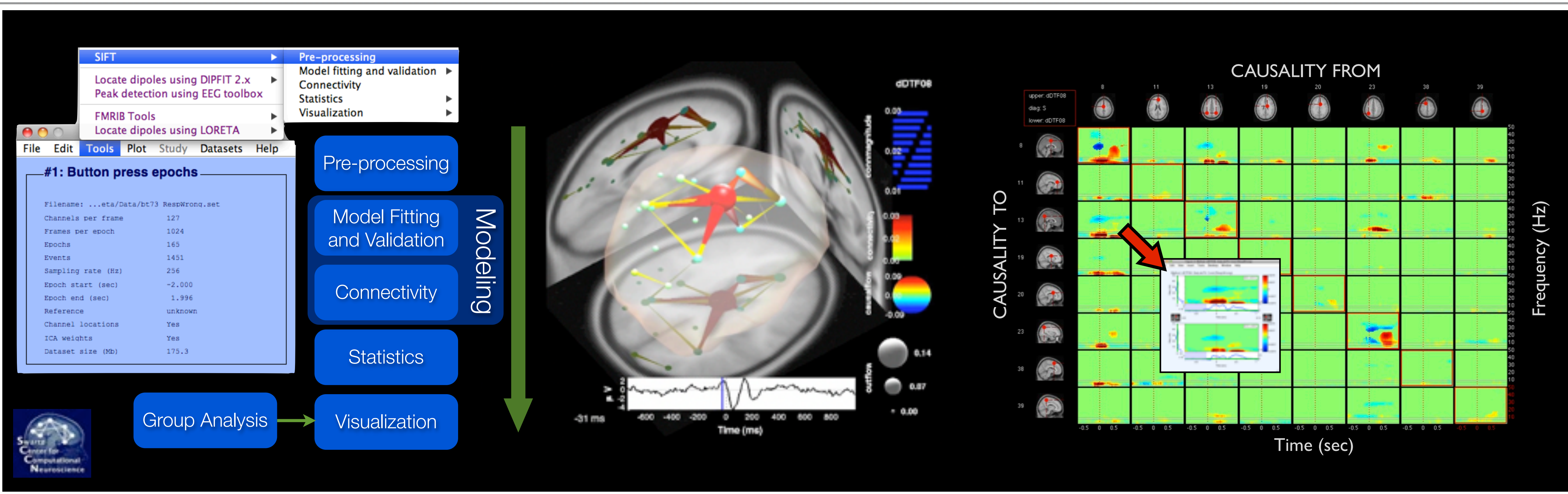
Delorme, Mullen, Kothe et al, *Computational Intelligence and Neuroscience*, vol 12, 2011



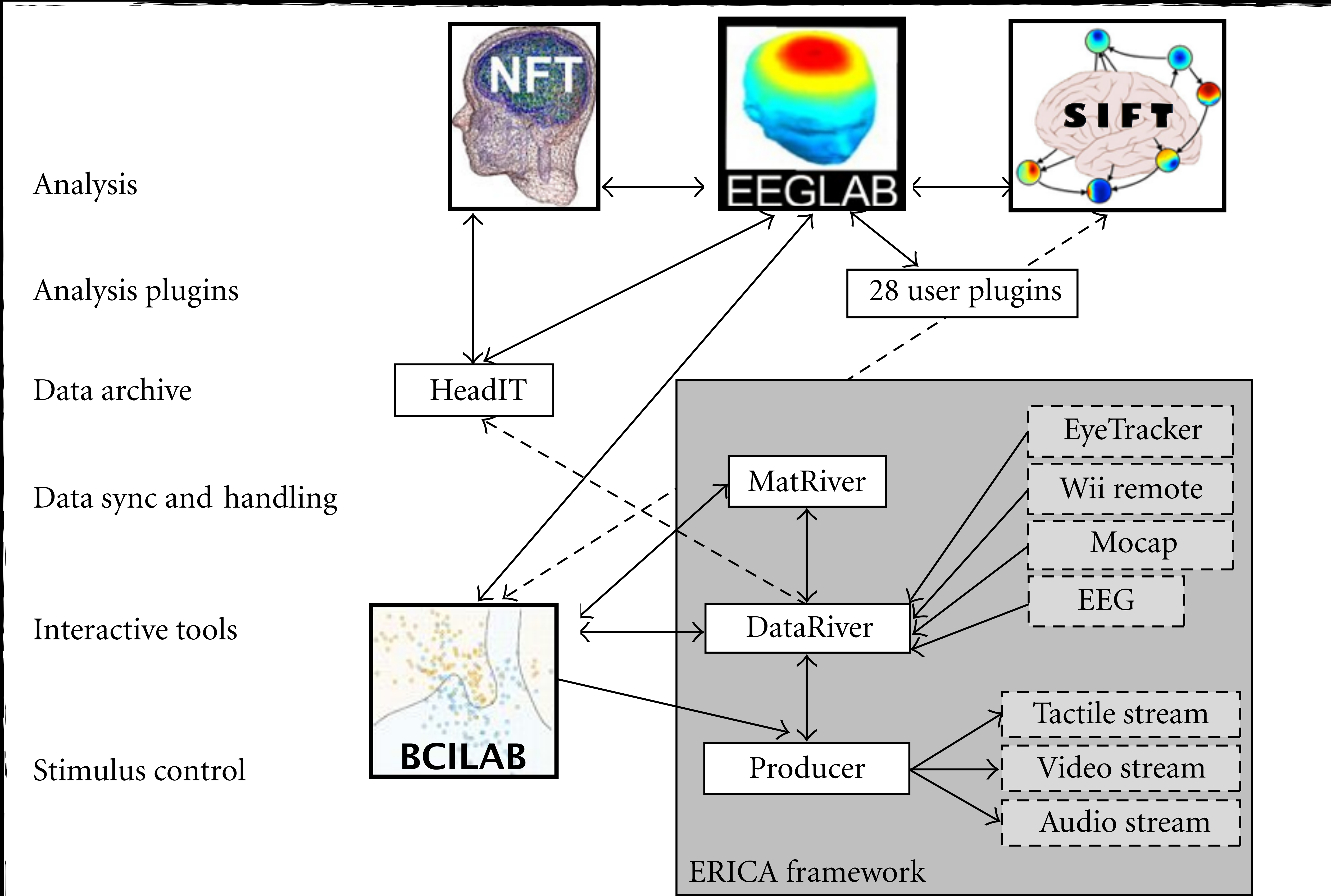


SIFT

Source Information Flow Toolbox



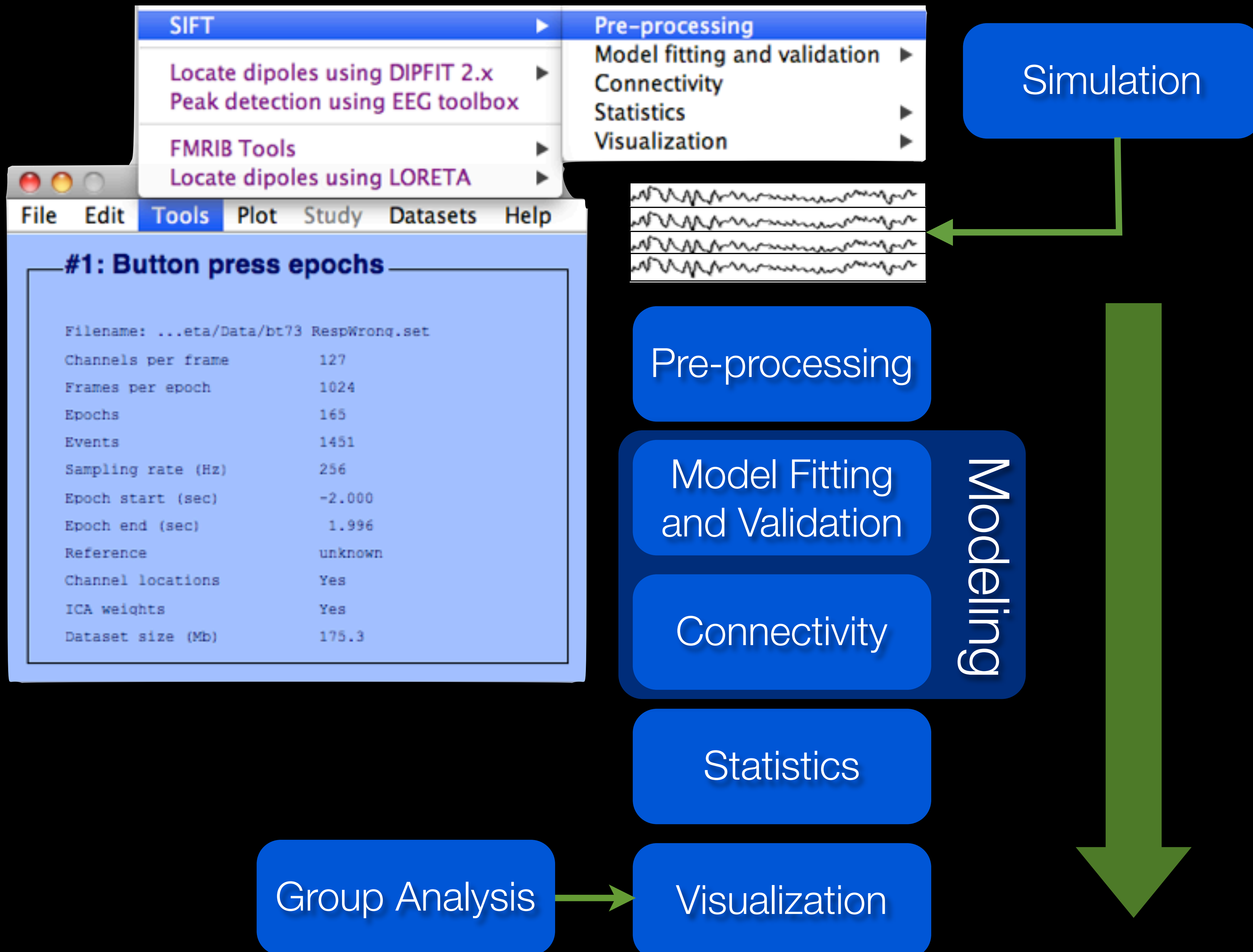
EEGLAB Software framework



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

Source Information Flow Toolbox (SIFT)

- A toolbox for (source-space) electrophysiological information flow and causality analysis (single- or multi-subject) integrated into the EEGLAB software environment.
- Modular architecture intended to support multiple modeling approaches
- Emphasis on vector autoregression and SSMs 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™ 2008a+, Signal Processing Toolbox, Statistics Toolbox (for some functions -- may be removed in the future)


SIFT
Locate dipoles using DIPFIT 2.x
Peak detection using EEG toolbox
FMRIB Tools
Locate dipoles using LORETA
Pre-processing
Model fitting and validation
Connectivity
Statistics
Visualization
File Edit Tools Plot Study Datasets Help
#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
Simulation
Pre-processing
Model Fitting and Validation
Connectivity
Modeling
Statistics
Group Analysis
Visualization

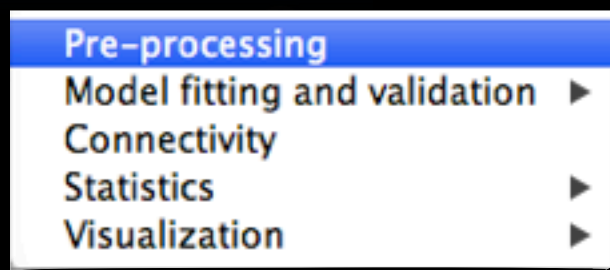
Preprocessing | Modeling | Statistics | Visualization

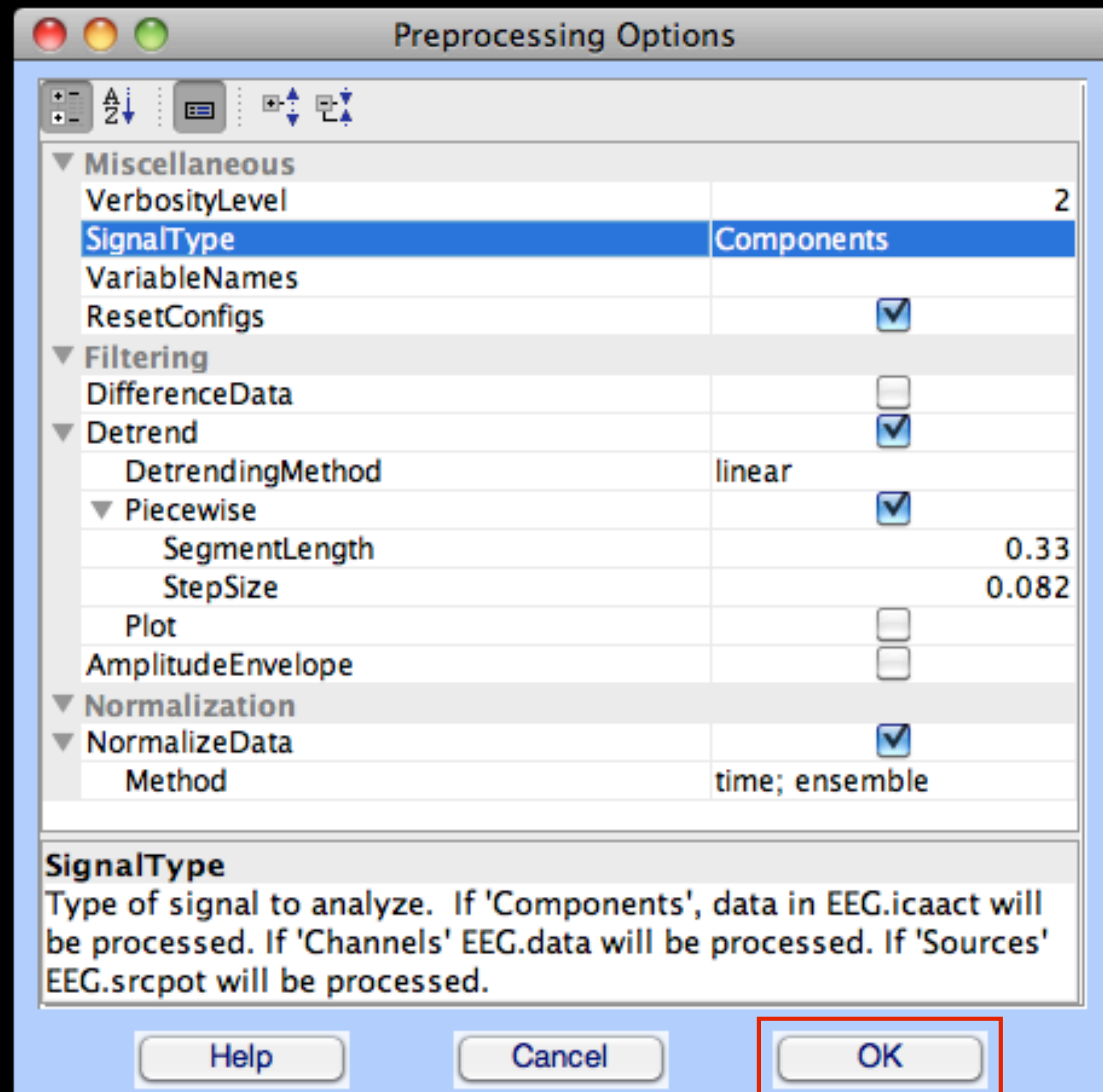
SIFT | Apps | To-Do | Fin

Source reconstruction (performed externally using EEGLAB or other toolboxes)
Filtering or Local Detrending
Downsampling
Differencing
Normalization (temporal or ensemble)
Trial balancing

Preprocessing | Modeling | Statistics | Visualization

SIFT | Apps | To-Do | Fin





Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

SIFT

Apps

To-Do

Fin

SIFT | Apps | To-Do | Fin

Preprocessing | **Modeling** | Statistics | Visualization

Model Fitting | Validation | Connectivity

Modeling Algorithm (1)	Linear	Nonlinear
Segmentation VAR (Sliding Window)		
Unconstrained		
Vieira-Morf	☑	
ARfit	☑	
Regularized		
Ridge Regression (L_2)	☑	
Group Lasso ($L_{1,2}$) ADMM, DAL	☑	
Elastic Net (L_1L_2)	☑	
Sparse Bayesian Learning (L_p) TMSBL, BSBL	☑	

fully implemented | alpha-testing | coming soon

SIFT
Apps
To-Do
Fin

Preprocessing | Modeling | Statistics | Visualization

Model Fitting | Validation | Connectivity

Modeling Algorithm (2)	Linear	Nonlinear
State-Space Modeling		
Linear Kalman Filtering	☑	
Dual Extended Kalman Filtering		☑
Cubature Kalman Filtering		☑
Sparsely Connected Components Analysis (SCSA)	☑	
Adaptive Mixture Impulse Response Analysis (AMIRA)	☑	
Nonparametric VAR Modeling		
Spectral Matrix Factorization	☑	

fully implemented (green) — alpha-testing (yellow) — coming soon (white)

Preprocessing

Modeling

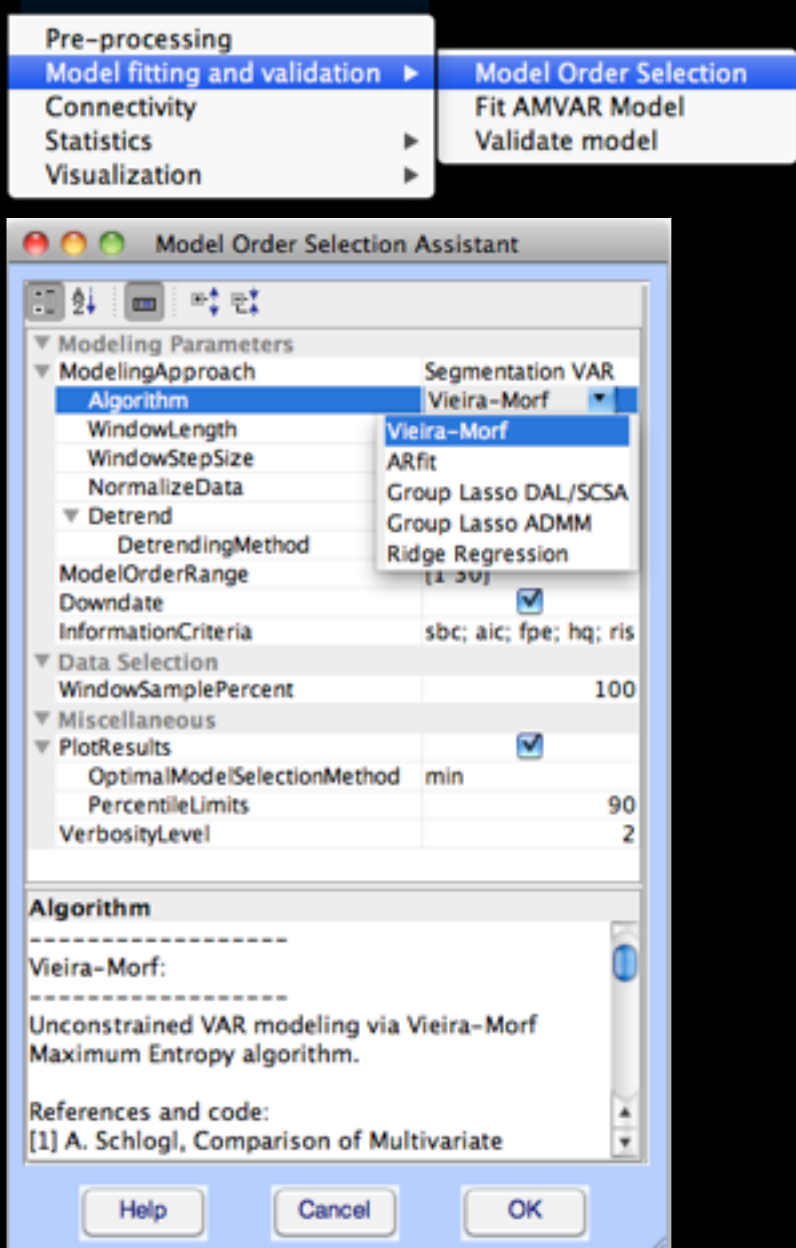
Statistics

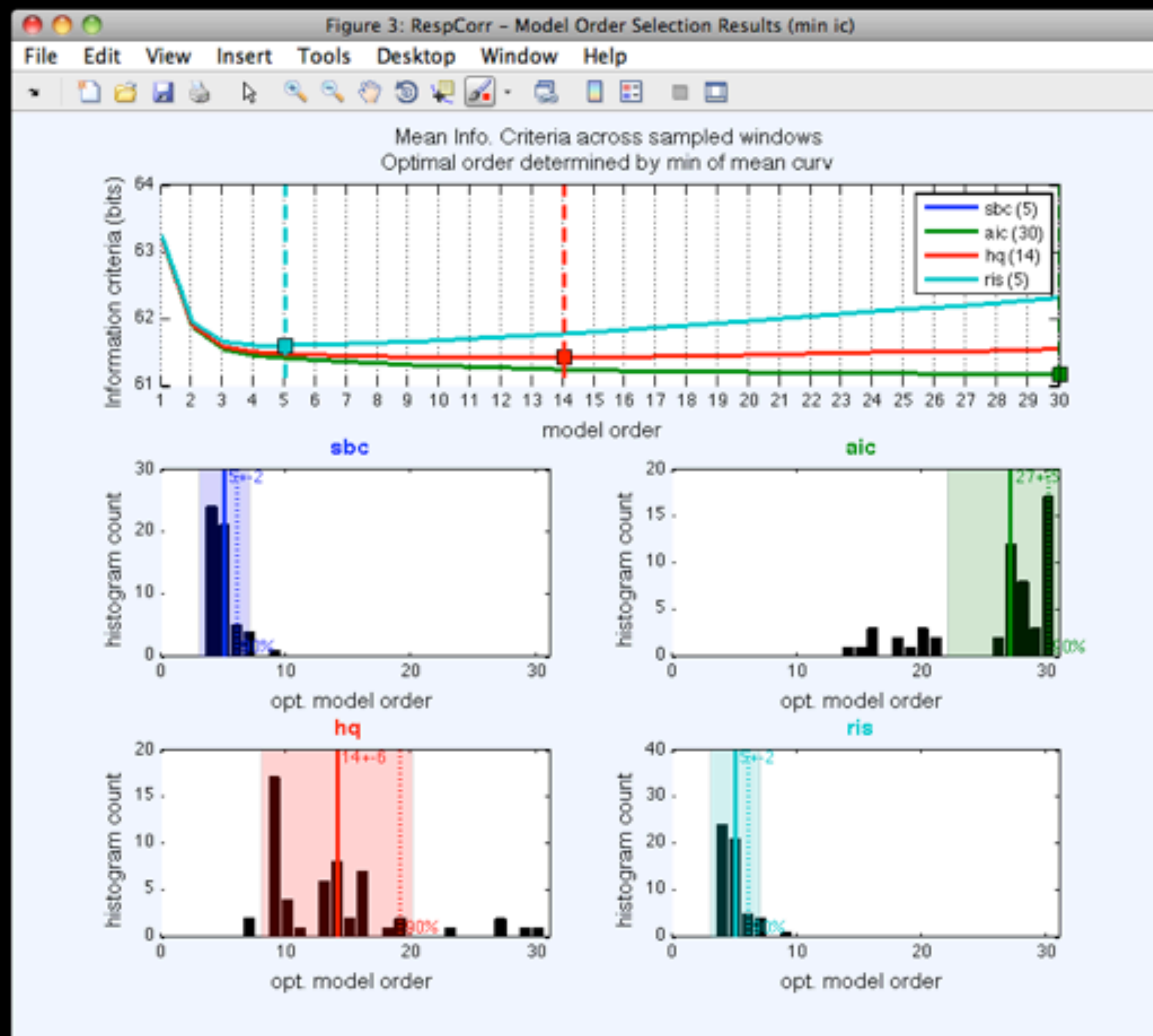
Visualization

Model Fitting

Validation

Connectivity





SIFT

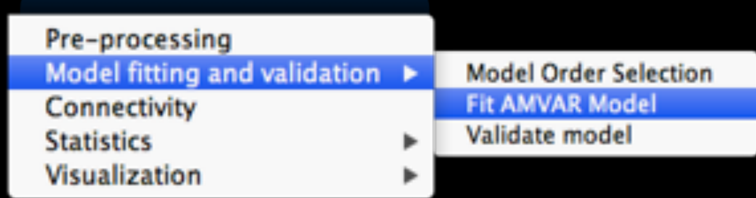
Apps

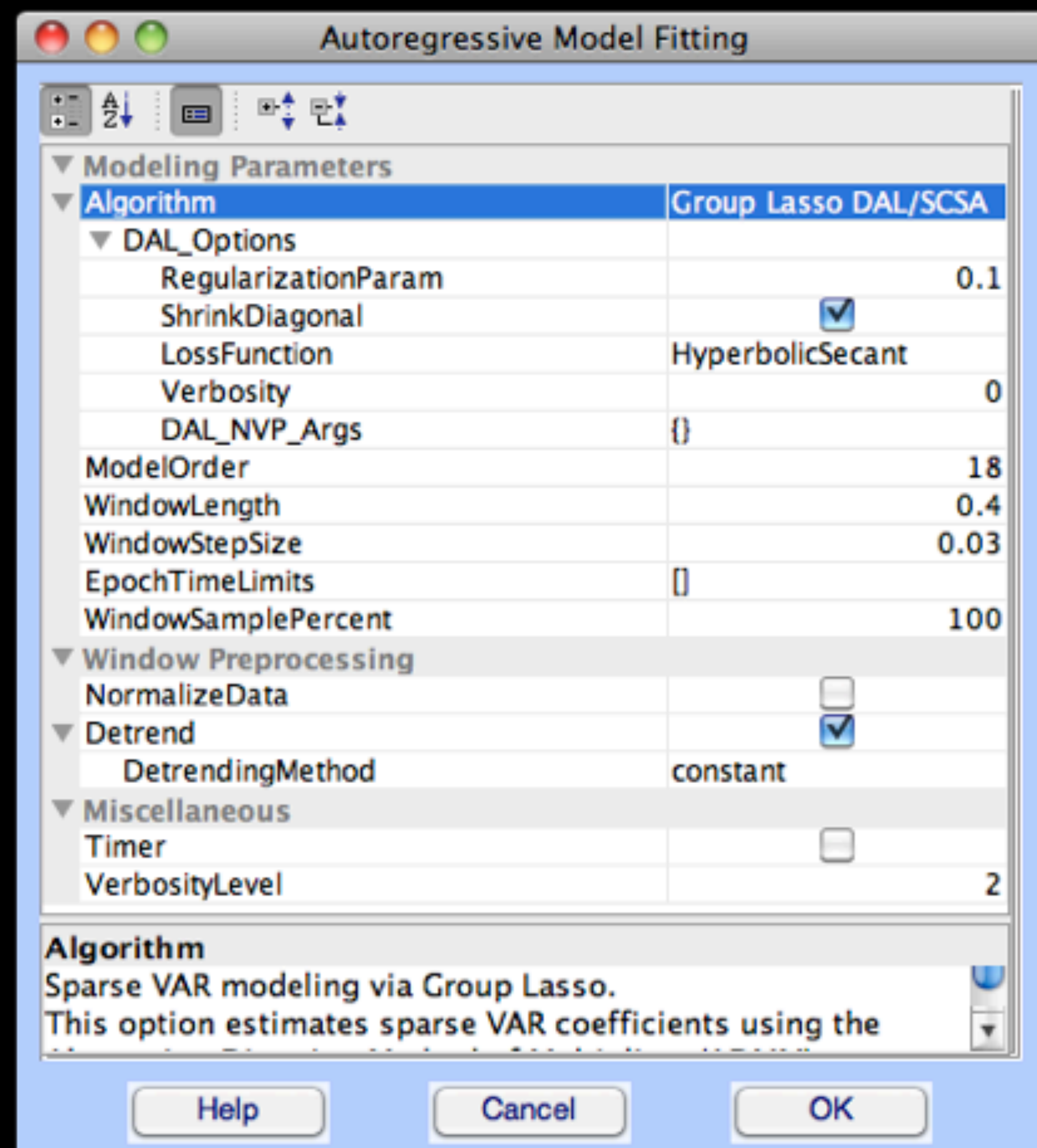
To-Do

Fin

Preprocessing | **Modeling** | Statistics | Visualization

Model Fitting | Validation | Connectivity





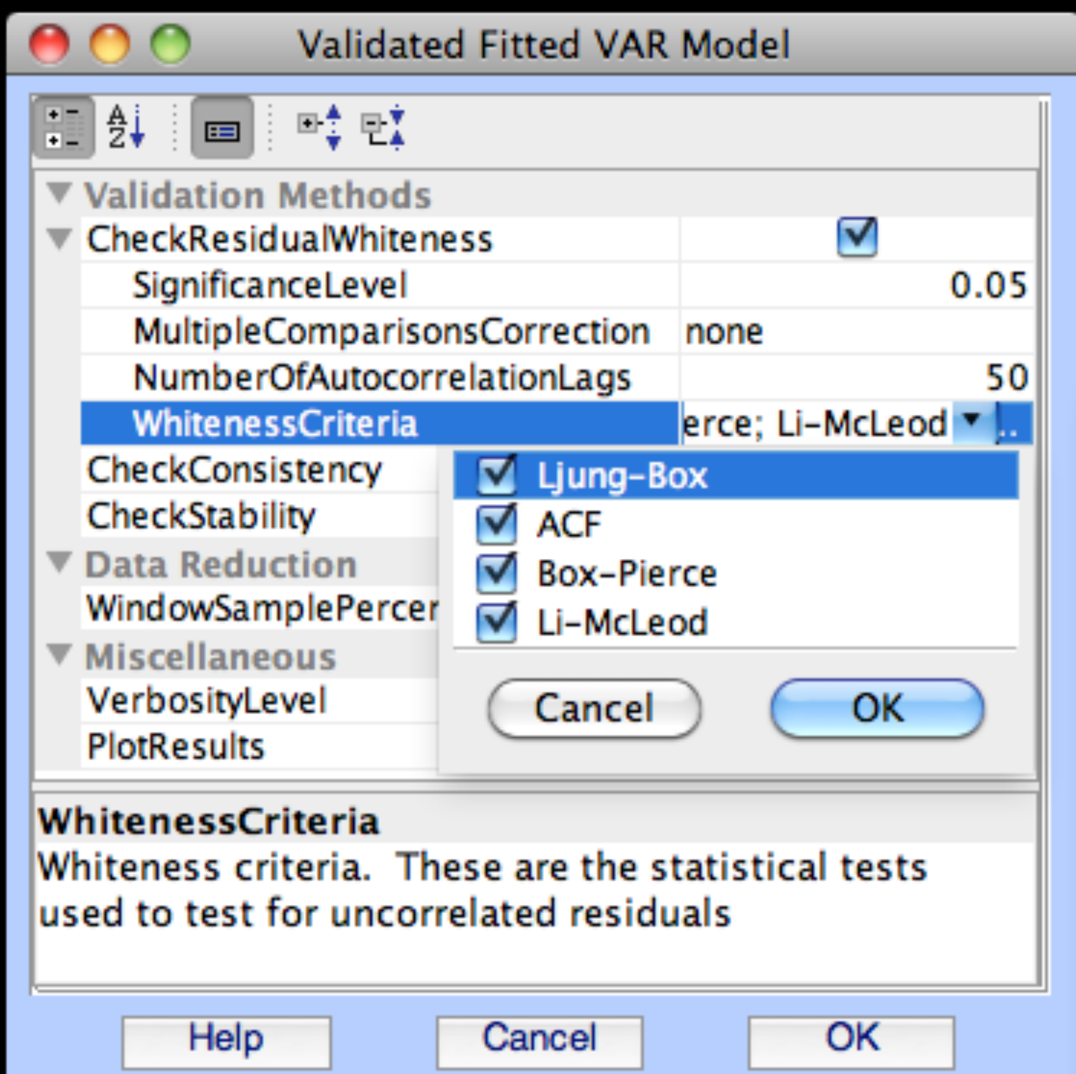
Preprocessing | **Modeling** | Statistics | Visualization

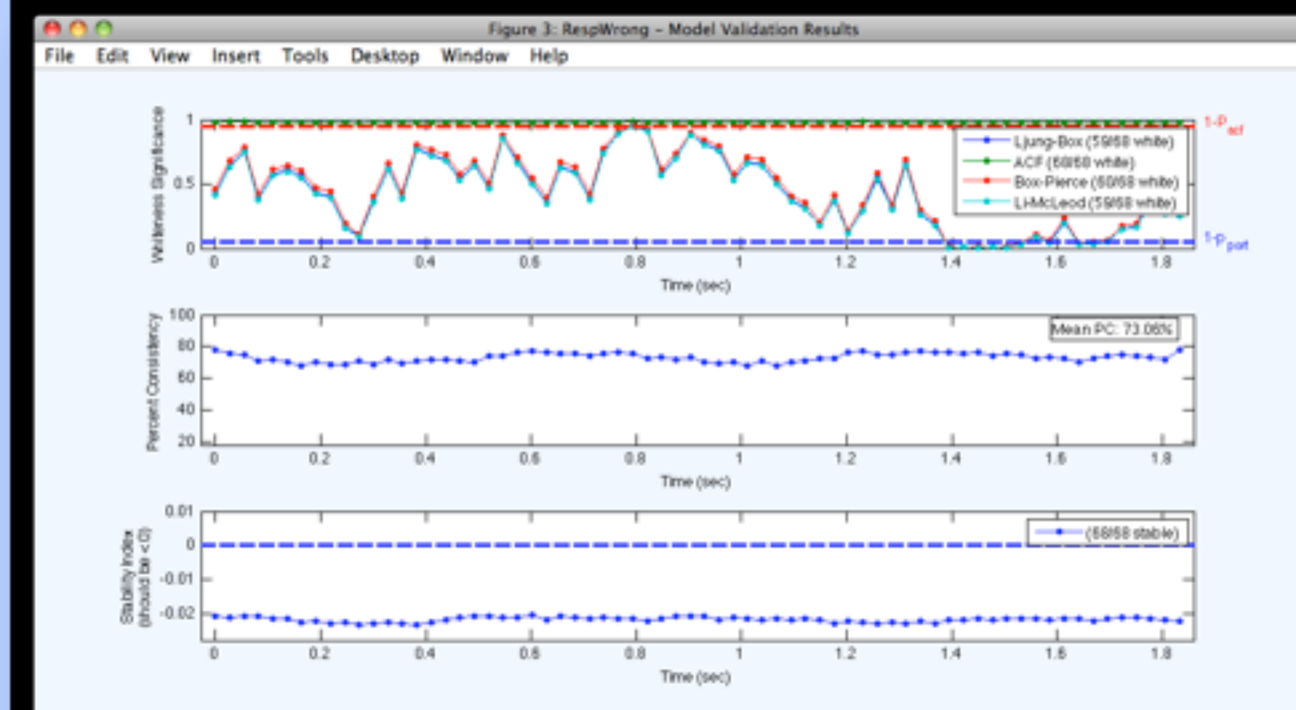
Model Fitting | **Validation** | Connectivity

SIFT | Apps | To-Do | Fin

VAR Model Validation
Residual ‘Whiteness’ Tests
Multivariate portmanteau tests
Residual autocorrelation probability test
Model Consistency
Model Stability
Nonparametric Spectral/Coherence Correlation

fully implemented — alpha-testing — coming soon





SIFT
Apps
To-Do
Fin

Preprocessing | Modeling | Statistics | Visualization

Model Fitting | Validation | Connectivity

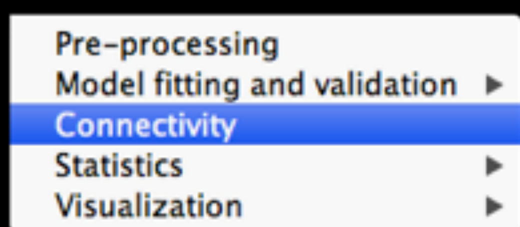
SIFT | Apps | To-Do | Fin

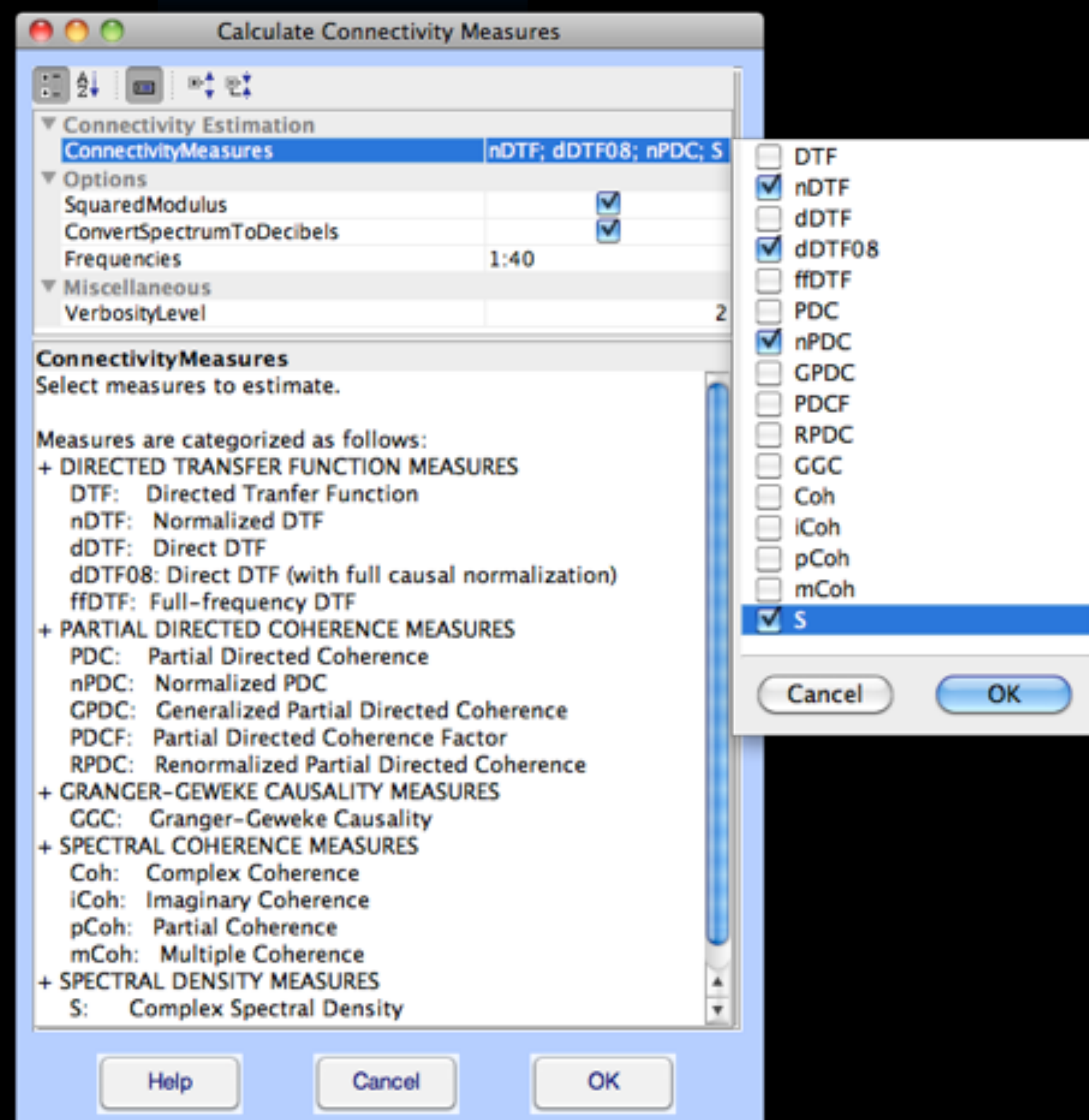
VAR-based Measures
Power spectrum (ERSP)
Coherence (Coh), Partial Coherence (pCoh), Multiple Coherence (mCoh)
Partial Directed Coherence (PDC)
Generalized PDC (GPDC)
Partial Directed Coherence Factor (PDCF)
Renormalized PDC (rPDC)
Directed Transfer Function (DTF)
Direct Directed Transfer Function (dDTF)
Bivariate Granger-Geweke Causality (GGC)
Conditional GGC
Blockwise GGC

fully implemented | alpha-testing | coming soon

Preprocessing | **Modeling** | Statistics | Visualization

Model Fitting | Validation | **Connectivity**





SIFT

Apps

To-Do

Fin

Preprocessing | Modeling | **Statistics** | Visualization

SIFT | Apps | To-Do | Fin

Statistical Approach	Test	Parametric	Nonparam.
Asymptotic analytic estimates of confidence intervals. Applies to: PDC, nPDC, DTF, nDTF, rPDC	H_{null}, H_{base}, H_{AB}	☑	
Theiler phase randomization Applies to: all	H_{null}		☑
Bootstrap, Jacknife, Cross-Validation Applies to: all	H_{AB}, H_{base}		☑
Confidence intervals using Bayesian smoothing splines Applies to: all	H_{base}, H_{AB}	☑	☑

$H_{null} : \mathbf{C}_{ij} = 0$ $H_{base}: \mathbf{C}_{ij} = \mathbf{C}_{baseline}$ $H_{AB}: \mathbf{C}^{A}_{ij} = \mathbf{C}^{B}_{ij}$

fully implemented | alpha-testing | coming soon

Preprocessing

Modeling

Statistics

Visualization

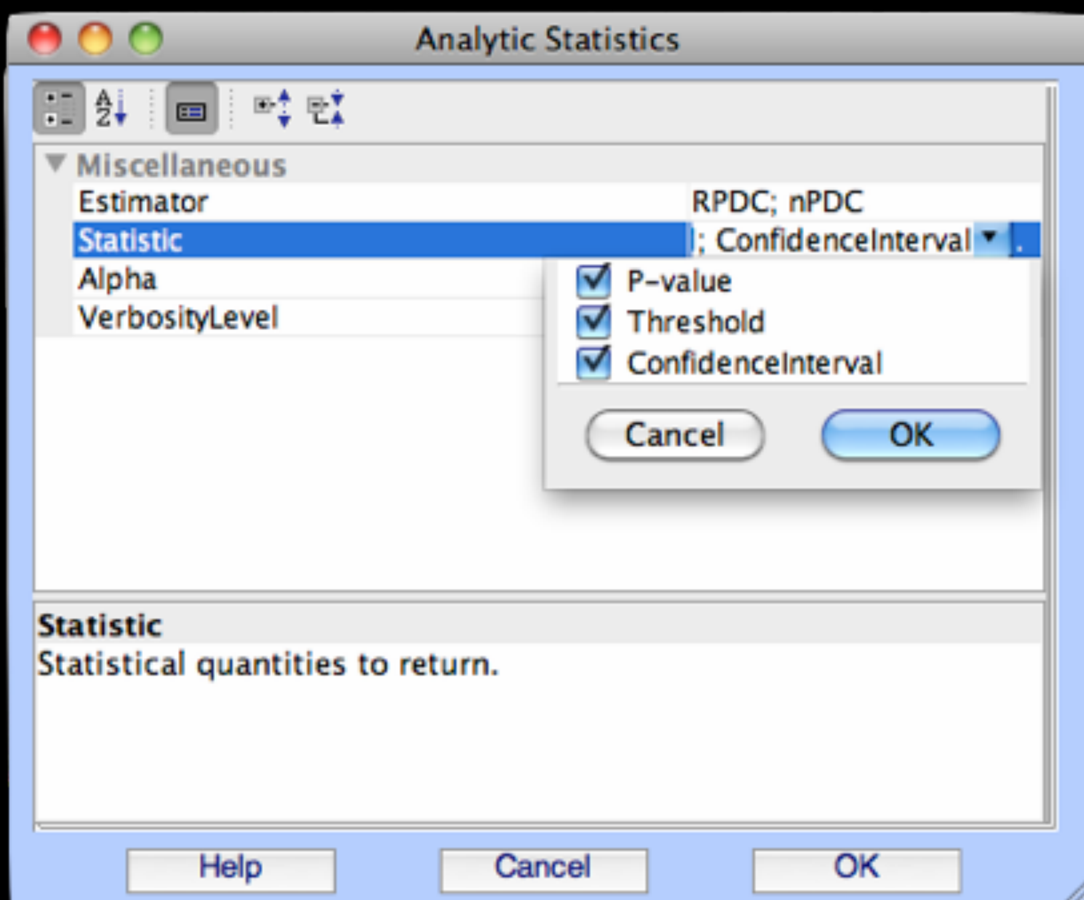
SIFT

Apps

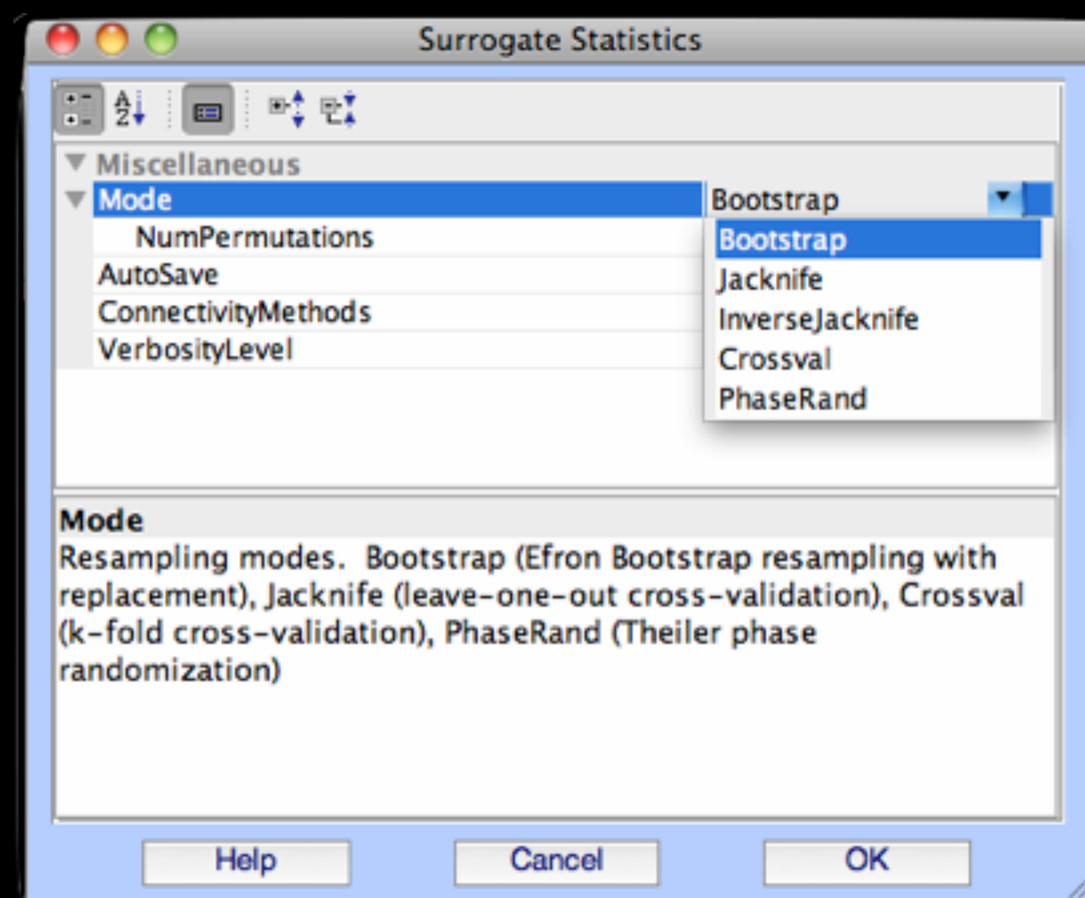
To-Do

Fin

Parametric



Non-parametric



fully implemented

alpha-testing

coming soon

Preprocessing | Modeling | Statistics | **Visualization**

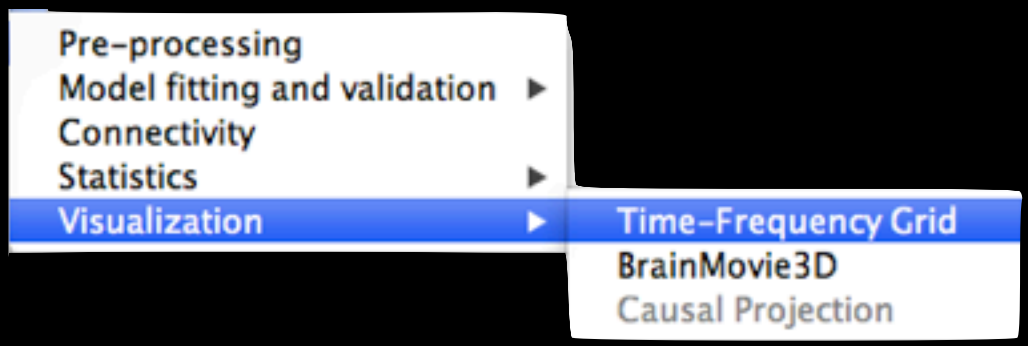
SIFT | Apps | To-Do | Fin

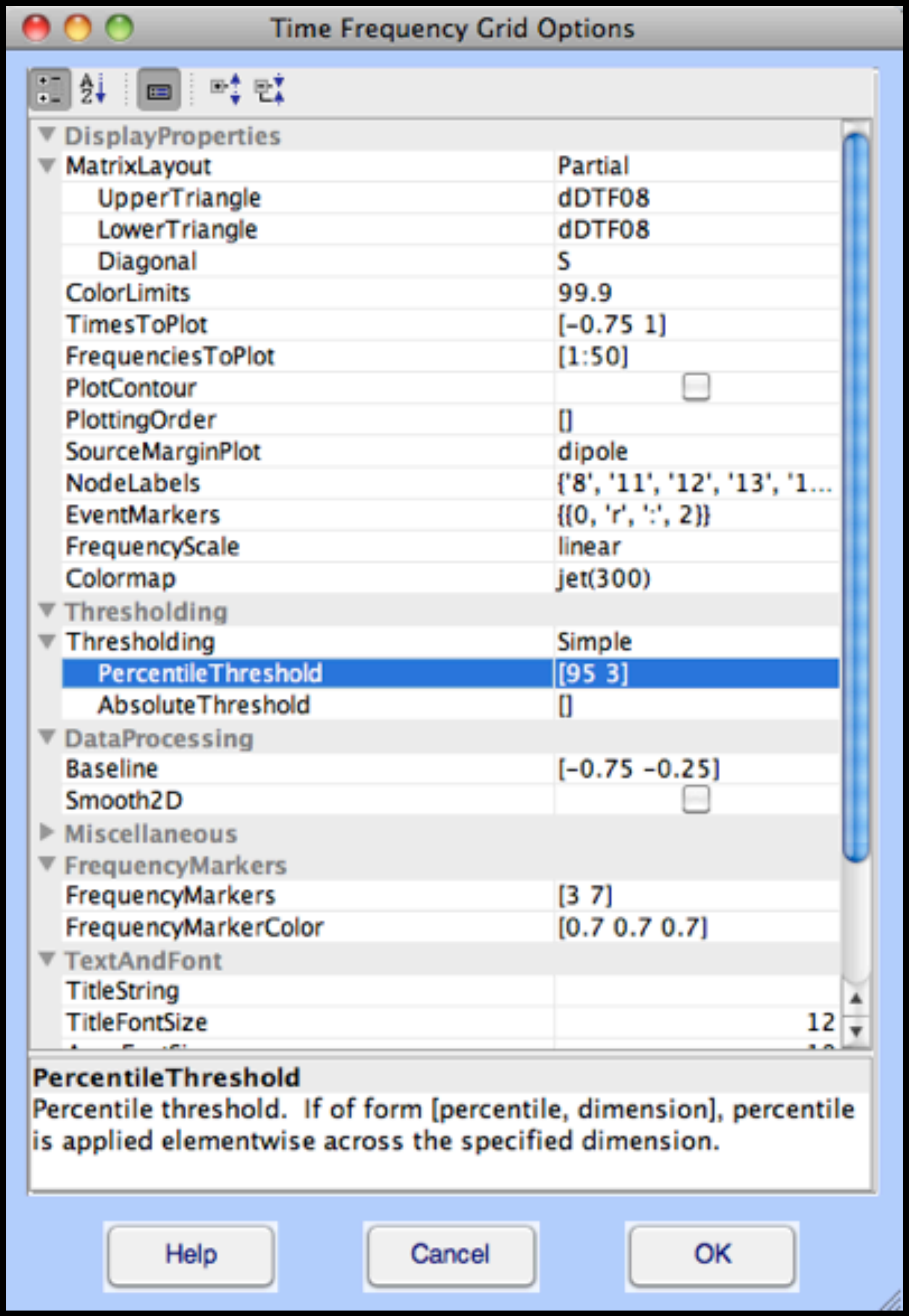
Interactive Visualizers

- Interactive Time-Frequency Grid
- Interactive 3D Causal Brainmovie
- Causal Projection Movie
- Directed Graphs and Graph Theoretic Analysis (Bioinformatics Toolbox Interface)
- and more ...

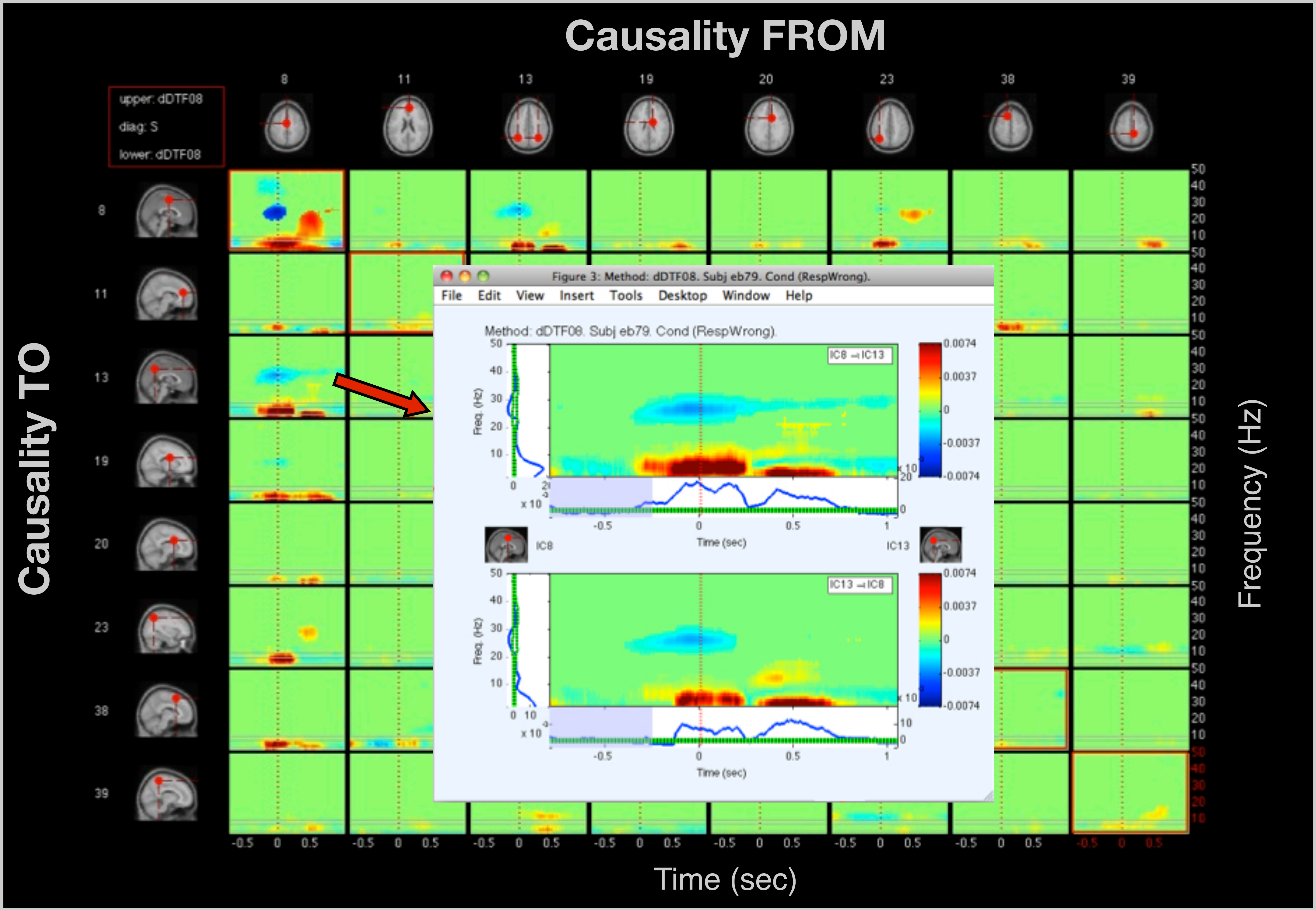
fully implemented (green) | alpha-testing (yellow) | coming soon (white)

Interactive Time-Frequency Grid

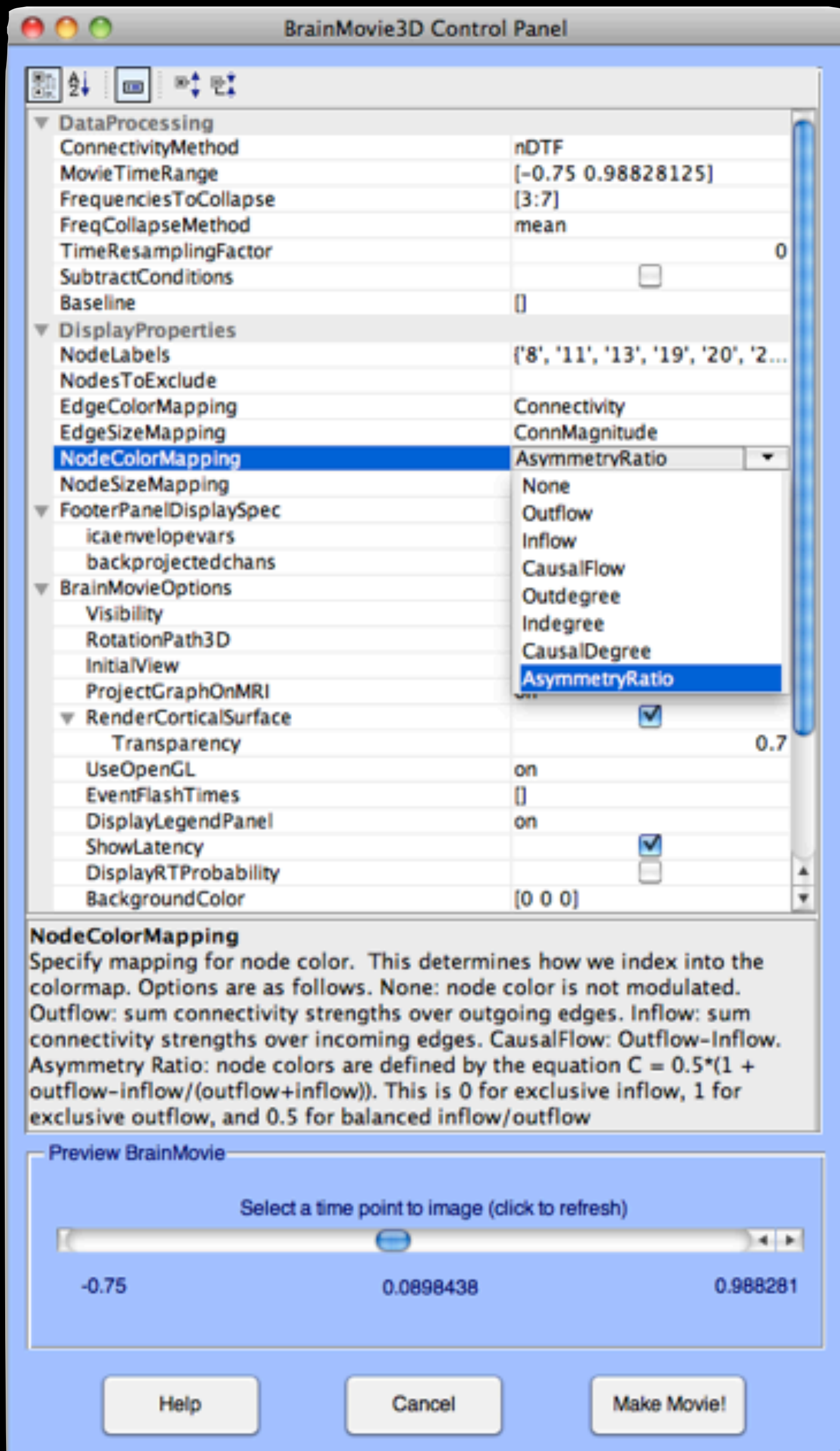


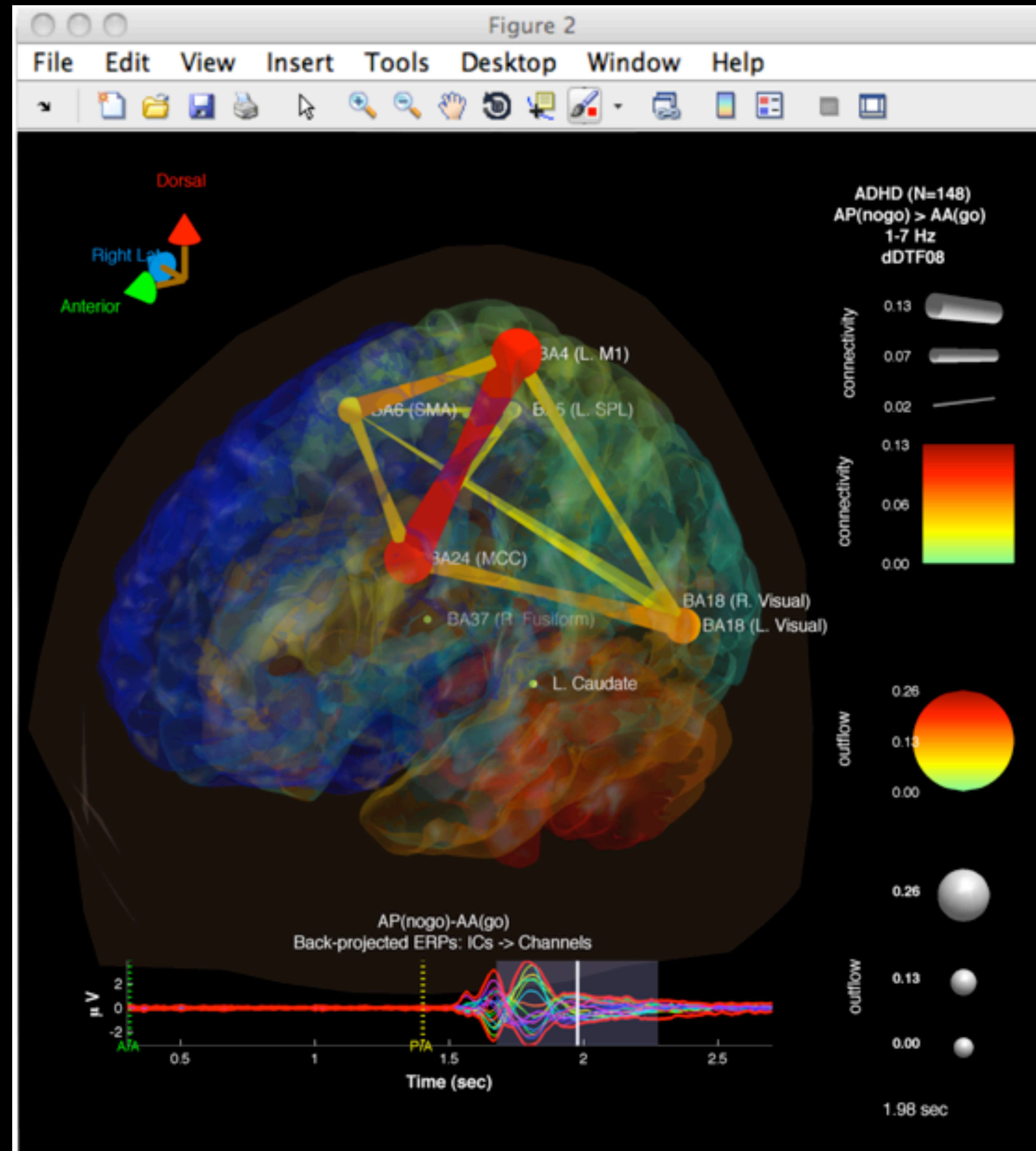


Interactive Time-Frequency Grid



Interactive Causal BrainMovie3D

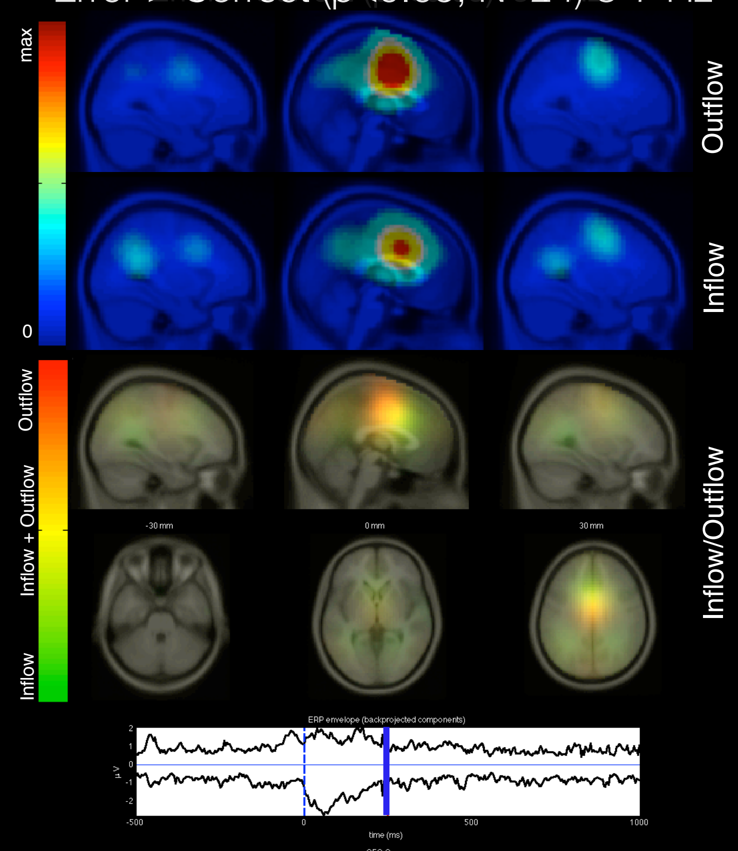




Group Analysis (in prep)

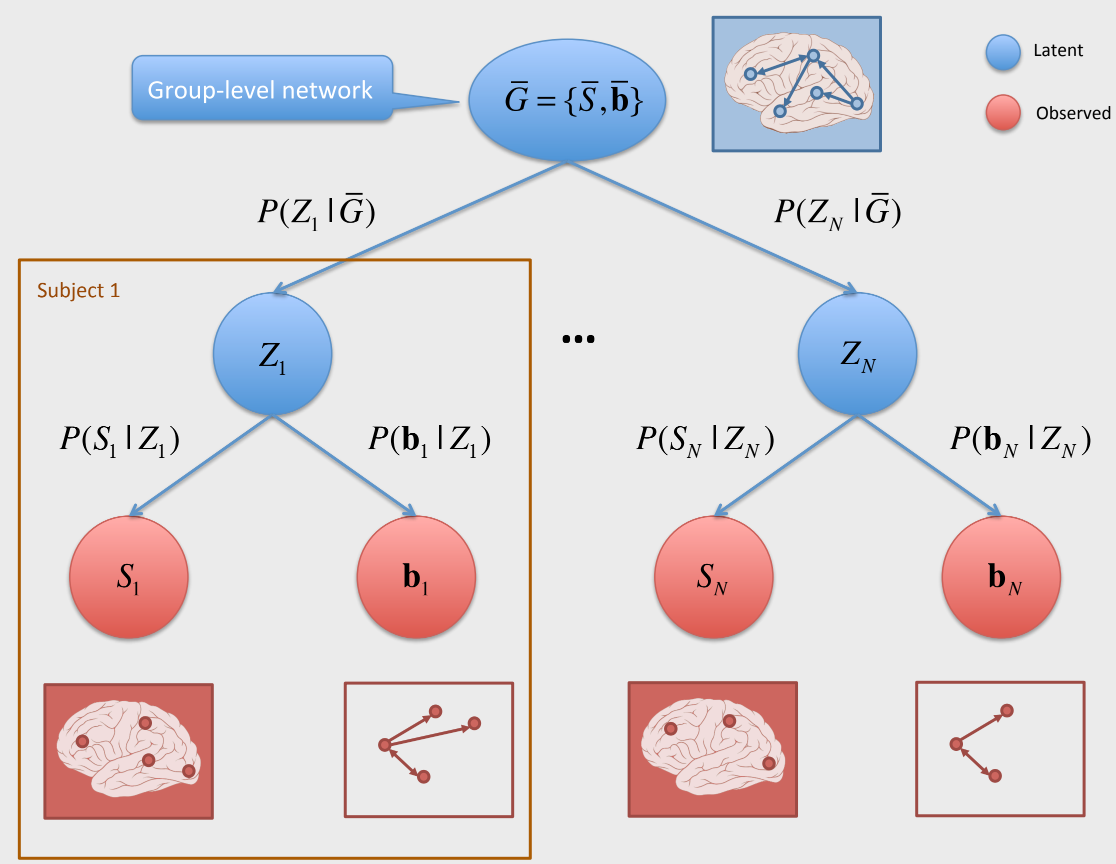
Causal/Measure Projection

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



Mullen, Onton, et al, 2010, HBM, Barcelona
Bigdely-Shamlo, Mullen, et al, 2012, *in review*

Bayesian Hierarchical Model

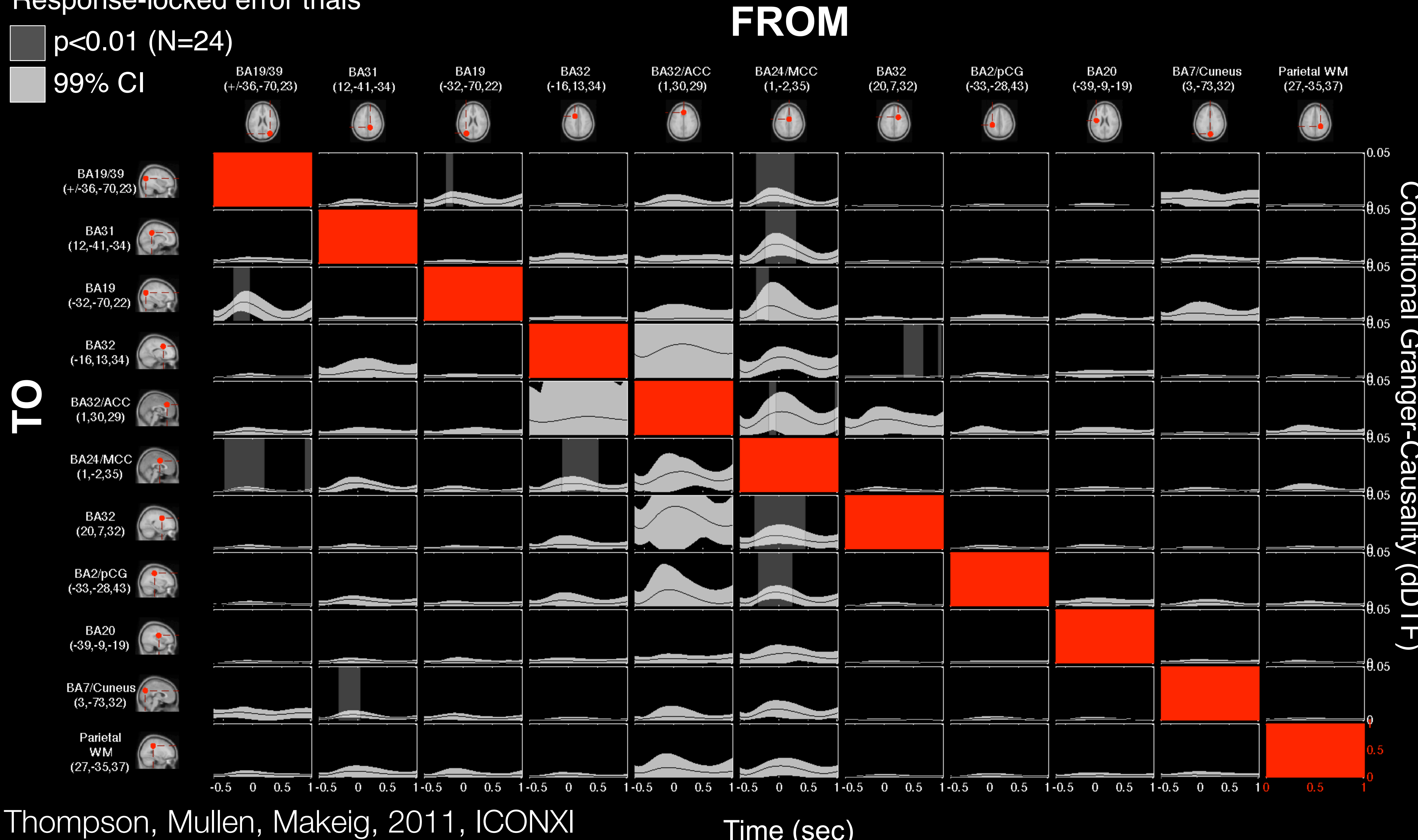


Thompson, Mullen, Makeig, 2011, ICONXI
Thompson, Mullen, Makeig, 2012, *in prep*

alpha-testing

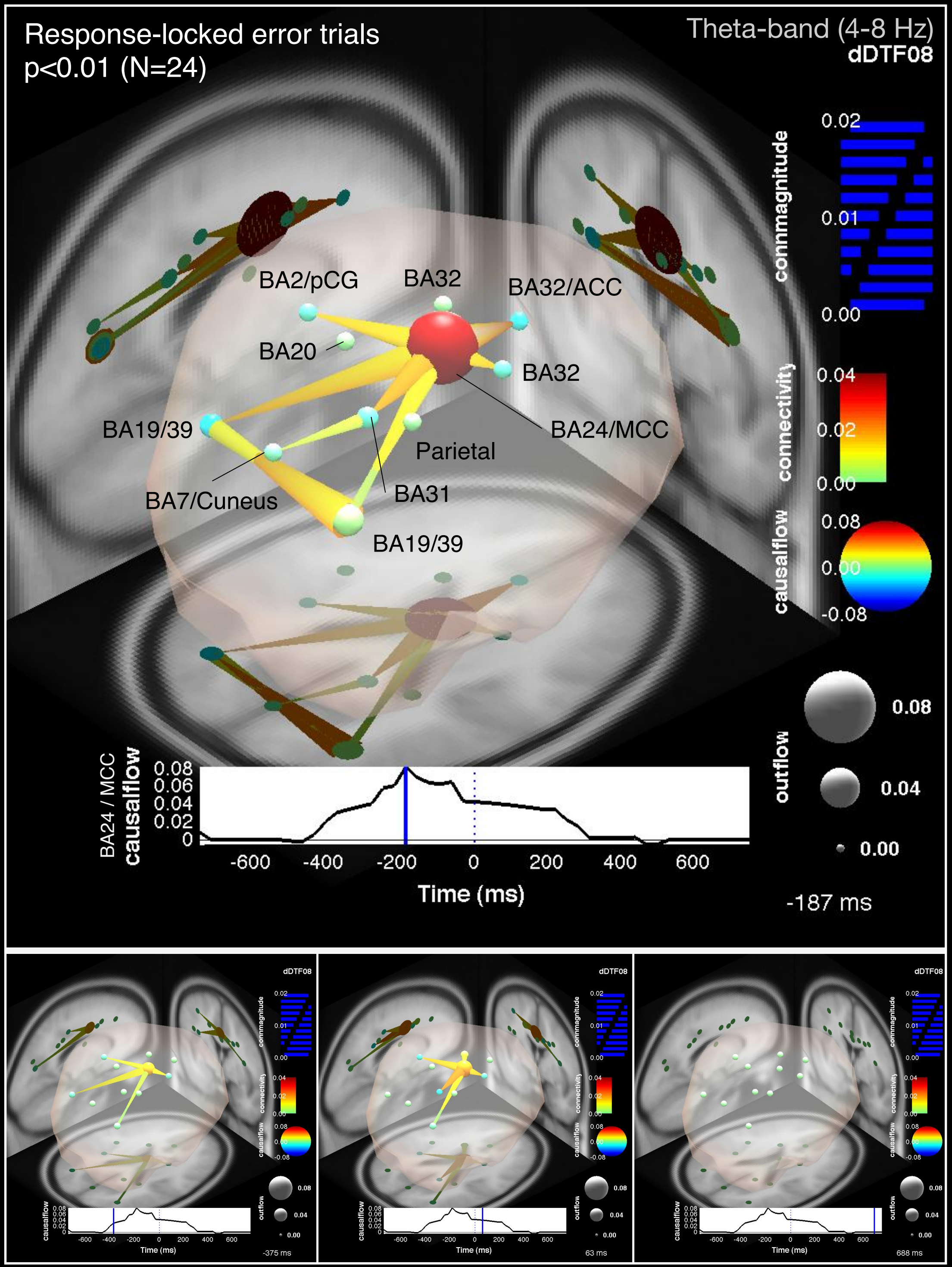
Bayesian Multi-Subject Inference

Theta-band (4-8 Hz) dDTF
Response-locked error trials
p<0.01 (N=24)
99% CI



Thompson, Mullen, Makeig, 2011, ICONXI
Thompson, Mullen, Makeig, 2012, *in prep*

Bayesian Multi-Subject Inference



Thompson, Mullen, Makeig, 2011, ICONXI
Thompson, Mullen, Makeig, 2012, *in prep*

Simulation

Dynamical System Simulation Workbench

Systems of linear stochastically-forced damped coupled oscillators

- Support for arbitrary time-varying (non-stationary) coupling dynamics
- Intuitive equation-based scripting environment
- Support for generalized gaussian or hyperbolic secant innovations

Nonlinear Dynamical Systems

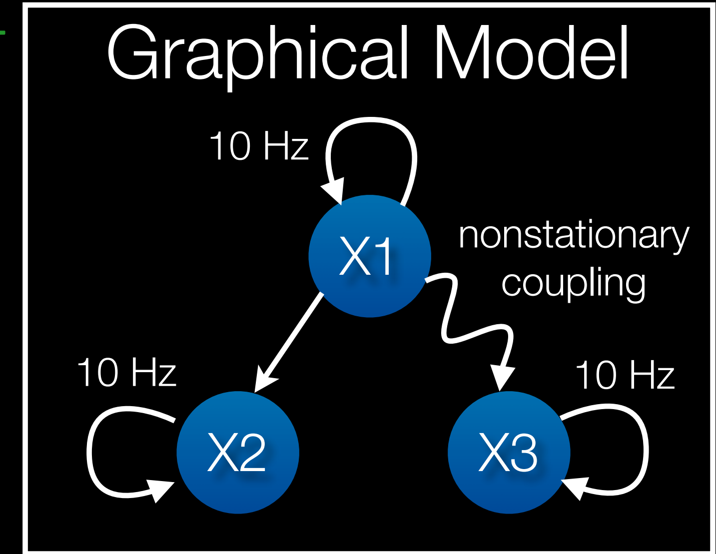
- Rössler and Lorenz Systems

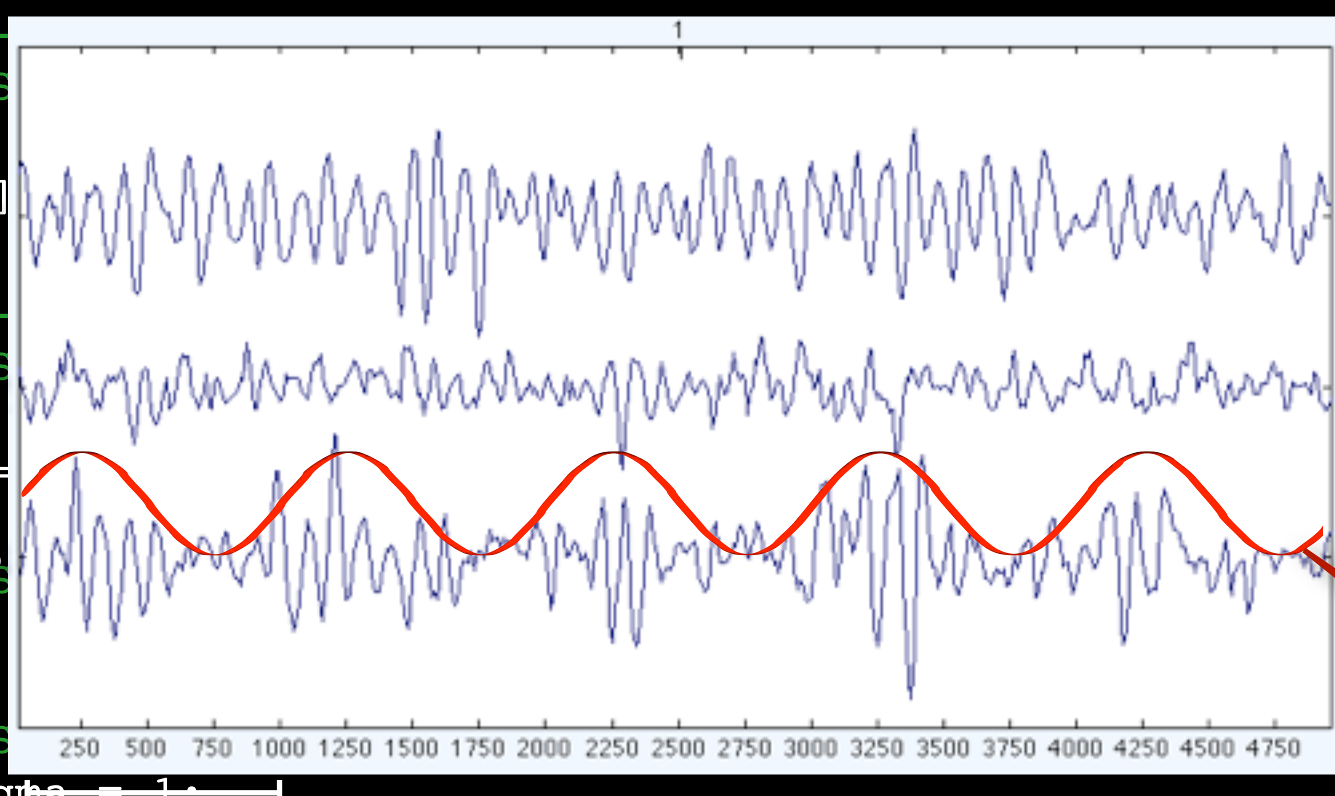
fully implemented

alpha-testing

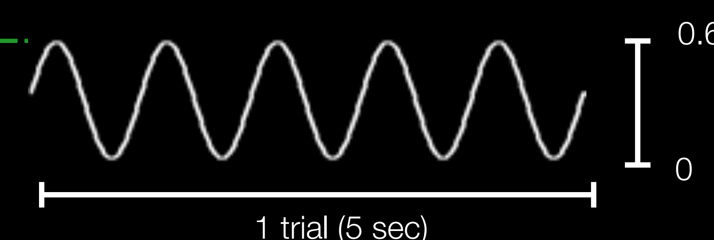
coming soon

```
% Example: Trivariate damped coupled oscillators with sinusoidally-modulated coupling
% ------------------------------------------------------------------------
% STEP 1: create prototype VAR structure

Fs = 100;                    % Sampling Rate (Hz)
Nl = 500;                    % length of each epoch (samples)
Nr = 100;                    % number of trials (realizations)
ndisc = 1000;                % number of startup samples to discard
ModelOrder = 2;              % model order
f0 = 10;                     % central oscillation frequency (Hz)

expr = {...
    ['x1(t) = ' sim_dampedOscillator(f0,9,Fs,1) '                                        + e1(t)'] ...
    ['x2(t) = ' sim_dampedOscillator(f0,2,Fs,2) ' + -0.1*x1(t-2)                          + e2(t)'] ...
    ['x3(t) = ' sim_dampedOscillator(f0,2,Fs,3) ' + {0.3*sin(2*pi*t/100)+0.3}*x1(t-2) + e3(t)'] ...
};

Aproto = sim_genVARModelFromEq(expr,ModelOrder);

% -
% S
[A]                                               Verbose',true);
% -
% S
M =
% S
Mu

% S
sigma = 1;
E = sigma*eye(M);

% generate simulated data with laplacian (supergaussian) innovations
data = sim_tvarsim(Mu,A,E,[Nl Nr],ndisc,1,1,'gengauss');
```



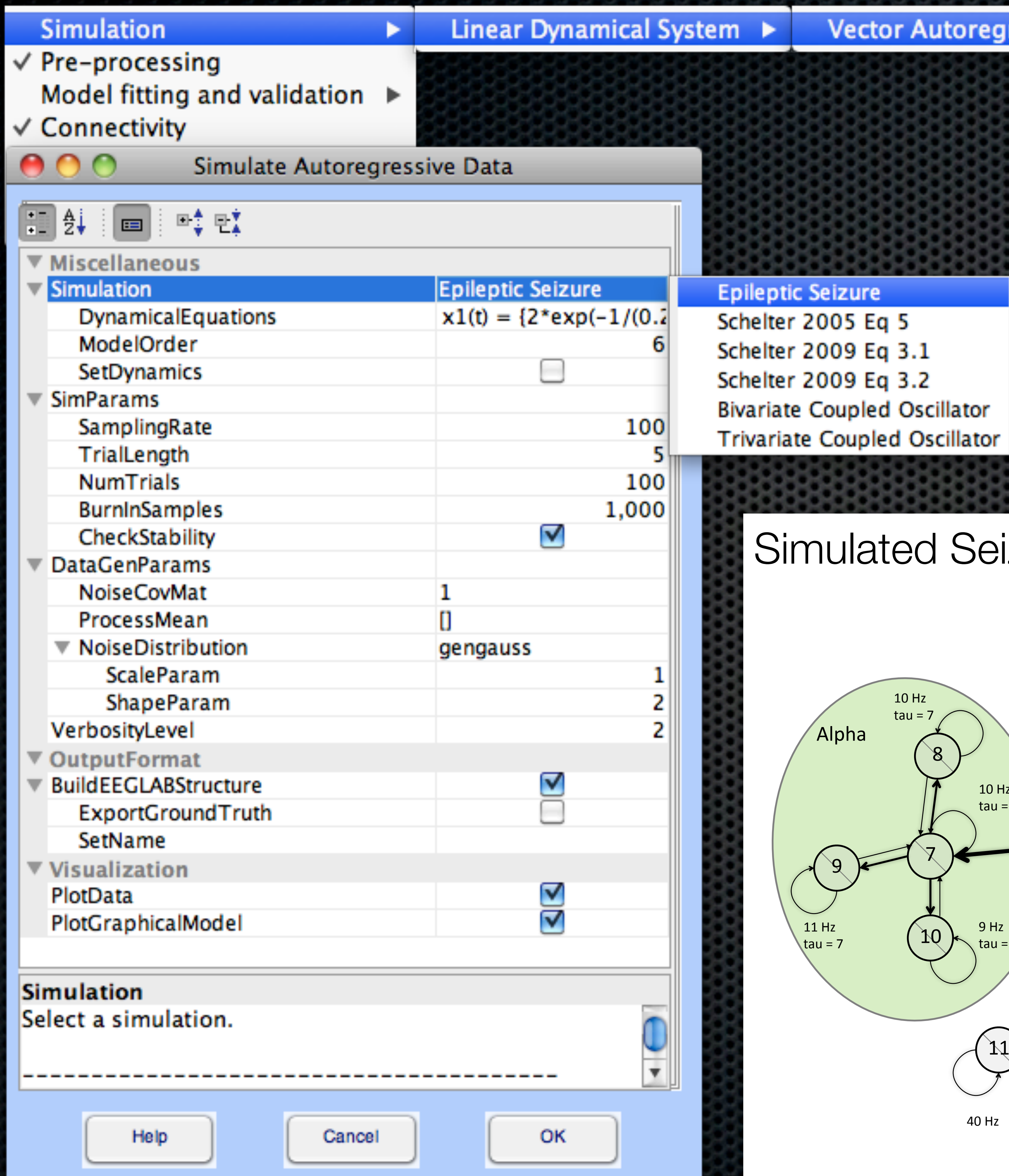


Time-varying X1→X3 coupling (1 Hz modulation)

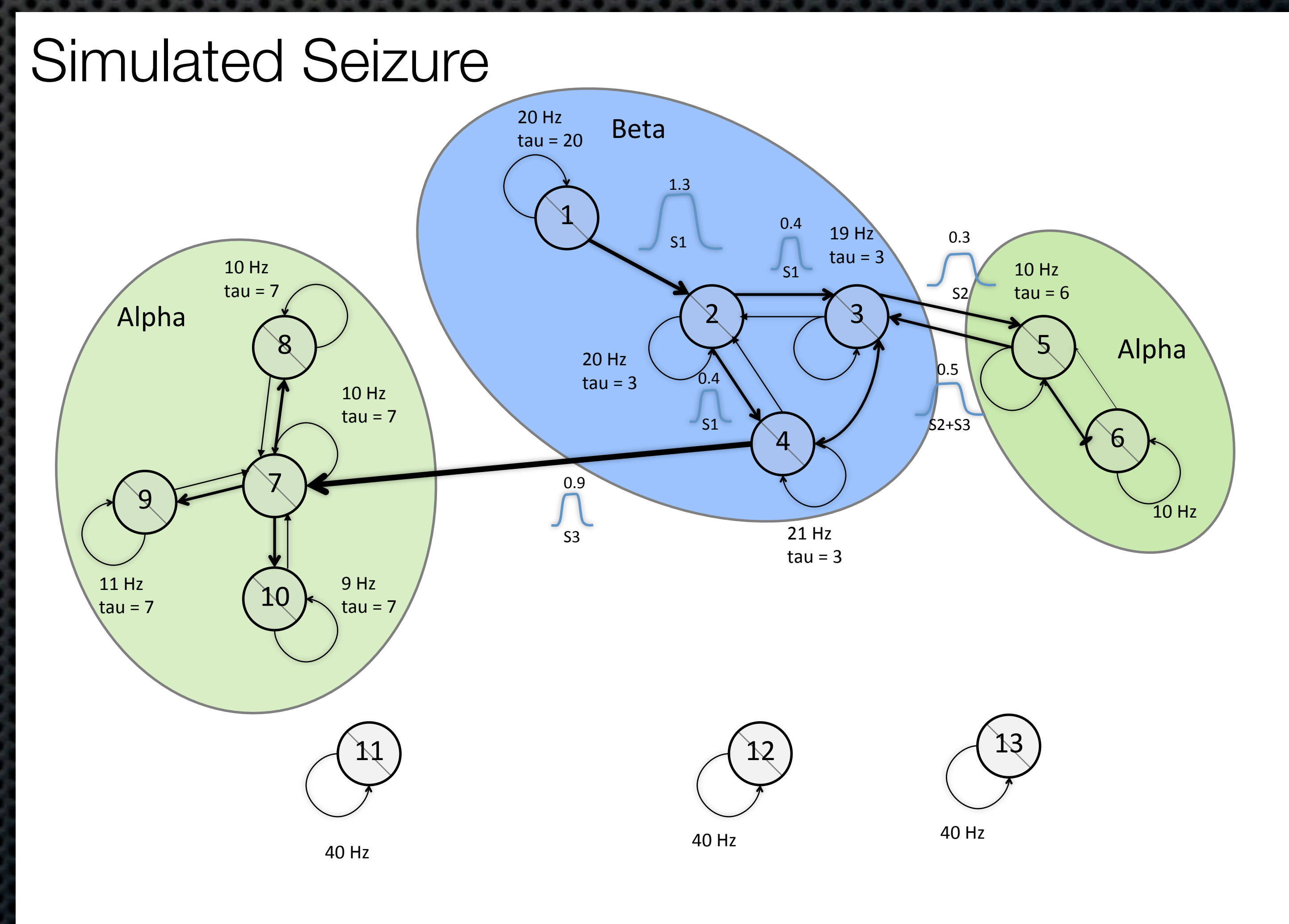


Amplitude Modulation (PAC)

Simulation

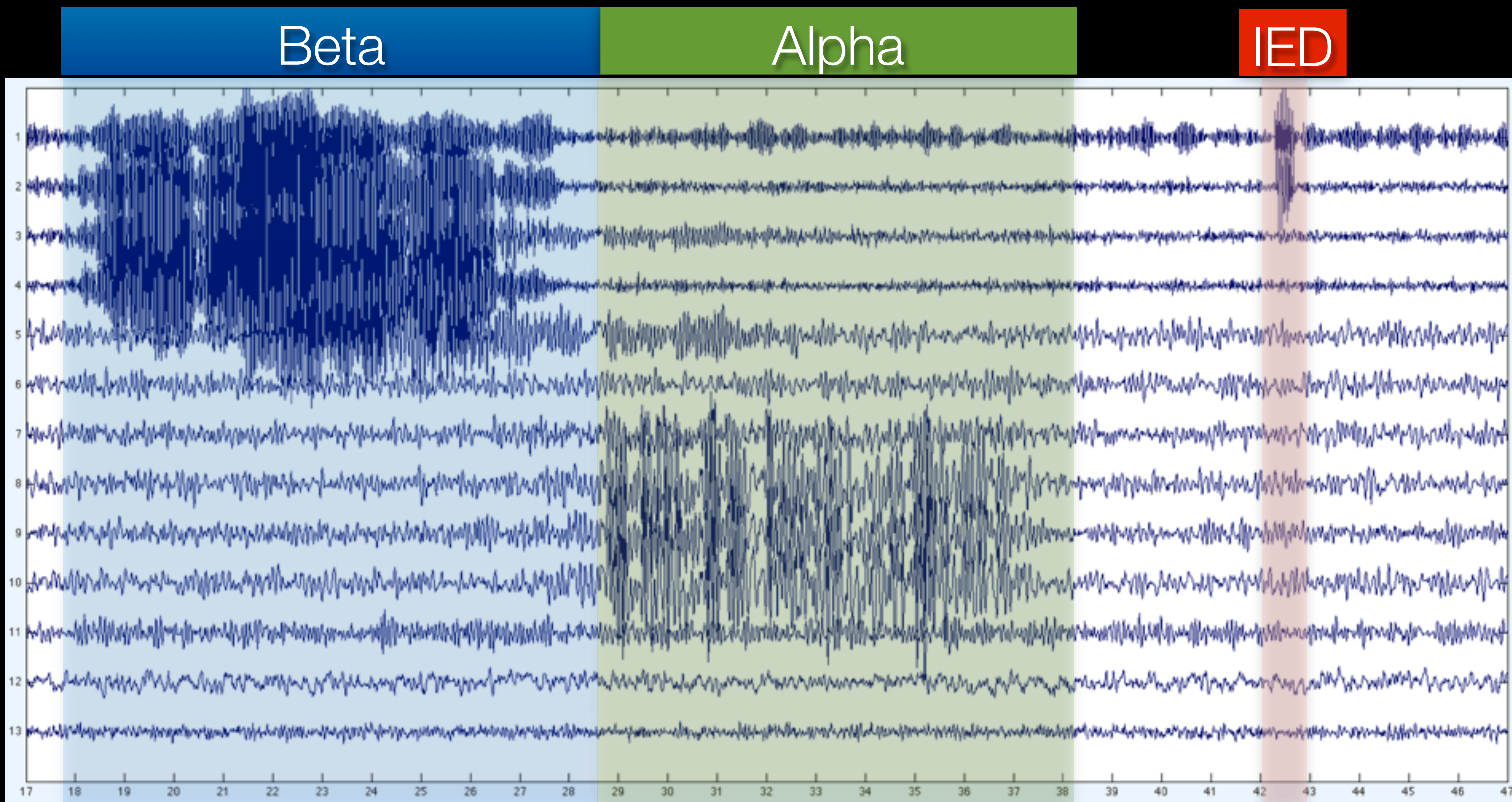


Simulation
Linear Dynamical System
Vector Autoregressive Process
Pre-processing
Model fitting and validation
Connectivity
Simulate Autoregressive Data
Miscellaneous
Simulation
Epileptic Seizure
DynamicalEquations
x1(t) = {2*exp(-1/(0.
ModelOrder
6
SetDynamics
SimParams
SamplingRate
100
TrialLength
5
NumTrials
100
BurnInSamples
1,000
CheckStability
DataGenParams
NoiseCovMat
1
ProcessMean
[]
NoiseDistribution
gengauss
ScaleParam
1
ShapeParam
2
VerbosityLevel
2
OutputFormat
BuildEEGLABStructure
ExportGroundTruth
SetName
Visualization
PlotData
PlotGraphicalModel
Simulation
Select a simulation.
Epileptic Seizure
Schelter 2005 Eq 5
Schelter 2009 Eq 3.1
Schelter 2009 Eq 3.2
Bivariate Coupled Oscillator
Trivariate Coupled Oscillator
Help
Cancel
OK



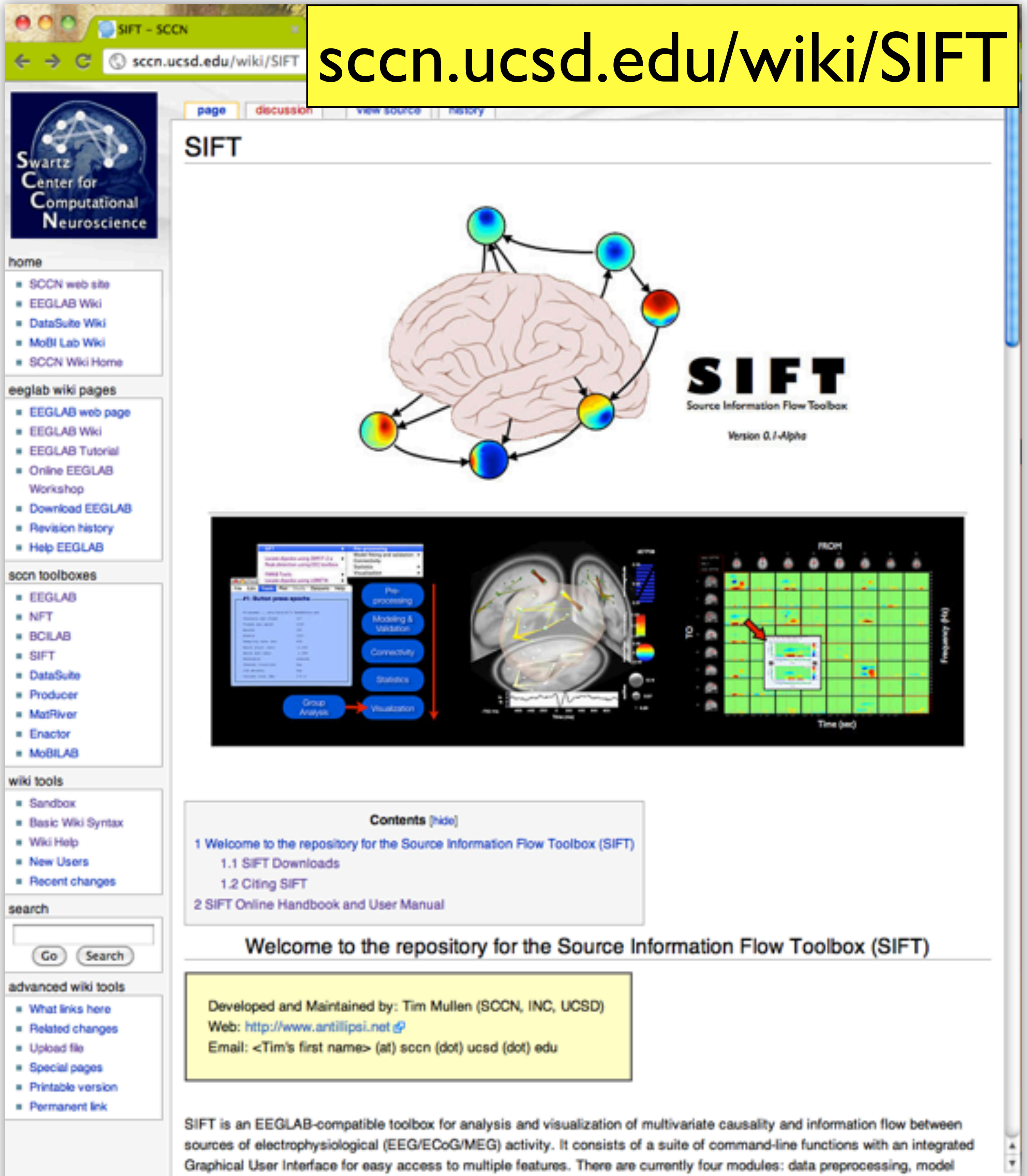
Simulated Seizure
Beta
20 Hz
tau = 20
1.3
S1
0.4
S1
19 Hz
tau = 3
0.3
S2
10 Hz
tau = 6
Alpha
20 Hz
tau = 3
0.4
S1
0.5
S2+S3
10 Hz
21 Hz
tau = 3
0.9
S3
Alpha
10 Hz
tau = 7
10 Hz
tau = 7
11 Hz
tau = 7
9 Hz
tau = 7
1
2
3
4
5
6
7
8
9
10
11
12
13
40 Hz
40 Hz
40 Hz

Simulated Seizure Sources



Where do I get SIFT?

sccn.ucsd.edu/wiki/SIFT


SIFT - SCCN
sccn.ucsd.edu/wiki/SIFT
page
discussion
view source
history

SIFT

home
- SCCN web site
- EEGLAB Wiki
- DataSuite Wiki
- MoBI Lab Wiki
- SCCN Wiki Home

eeglab wiki pages
- EEGLAB web page
- EEGLAB Wiki
- EEGLAB Tutorial
- Online EEGLAB Workshop
- Download EEGLAB
- Revision history
- Help EEGLAB

sccn toolboxes
- EEGLAB
- NFT
- BCILAB
- SIFT
- DataSuite
- Producer
- MatRiver
- Enactor
- MoBILAB

wiki tools
- Sandbox
- Basic Wiki Syntax
- Wiki Help
- New Users
- Recent changes

search
Go Search

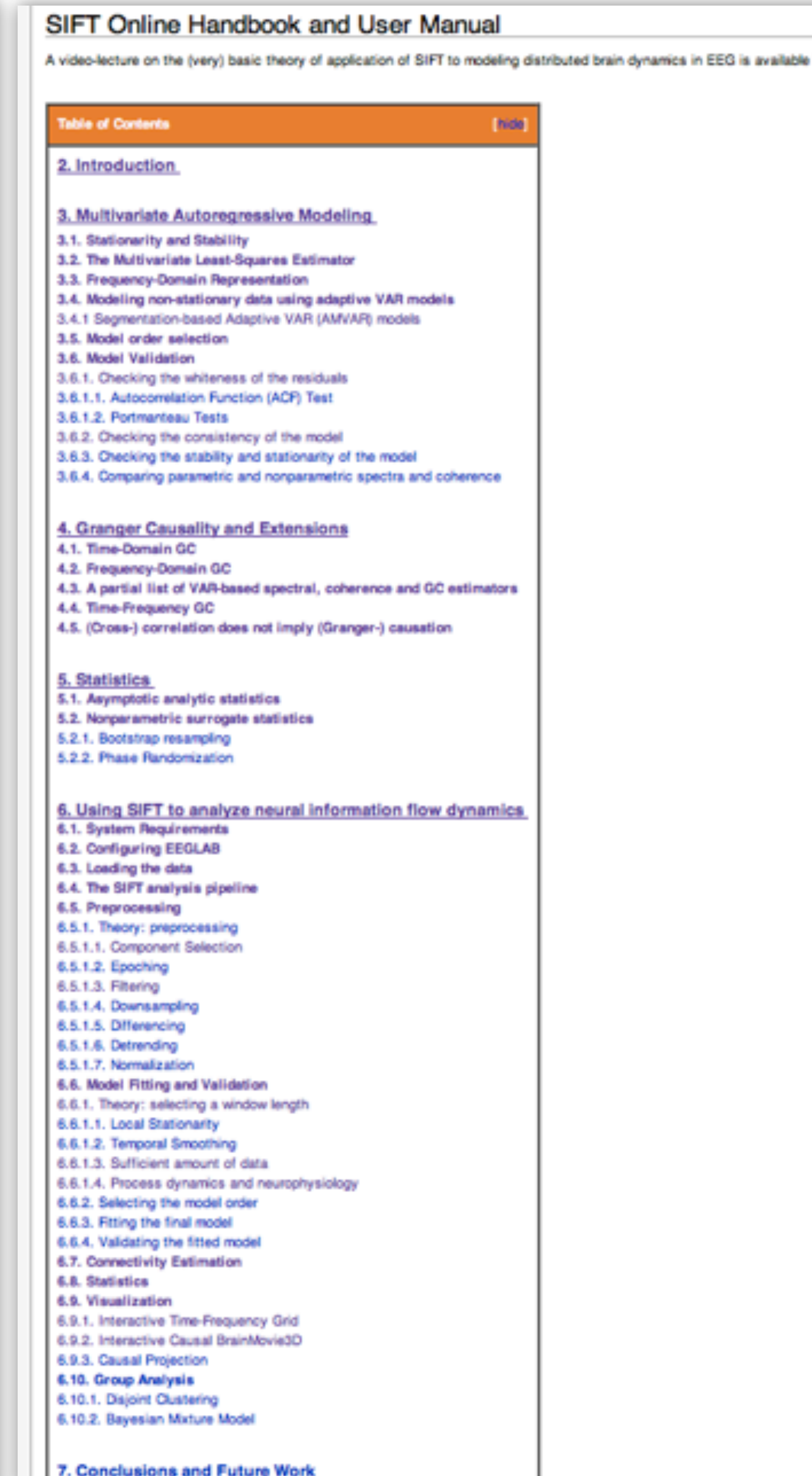
advanced wiki tools
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Contents [hide]

Welcome to the repository for the Source Information Flow Toolbox (SIFT)

Developed and Maintained by: Tim Mullen (SCCN, INC, UCSD)
Web: http://www.antillipsi.net
Email: <Tim's first name> (at) sccn (dot) ucsd (dot) edu

SIFT is an EEGLAB-compatible toolbox for analysis and visualization of multivariate causality and information flow between sources of electrophysiological (EEG/ECoG/MEG) activity. It consists of a suite of command-line functions with an integrated Graphical User Interface for easy access to multiple features. There are currently four modules: data preprocessing, model fitting and connectivity estimation, statistical analysis, and visualization.



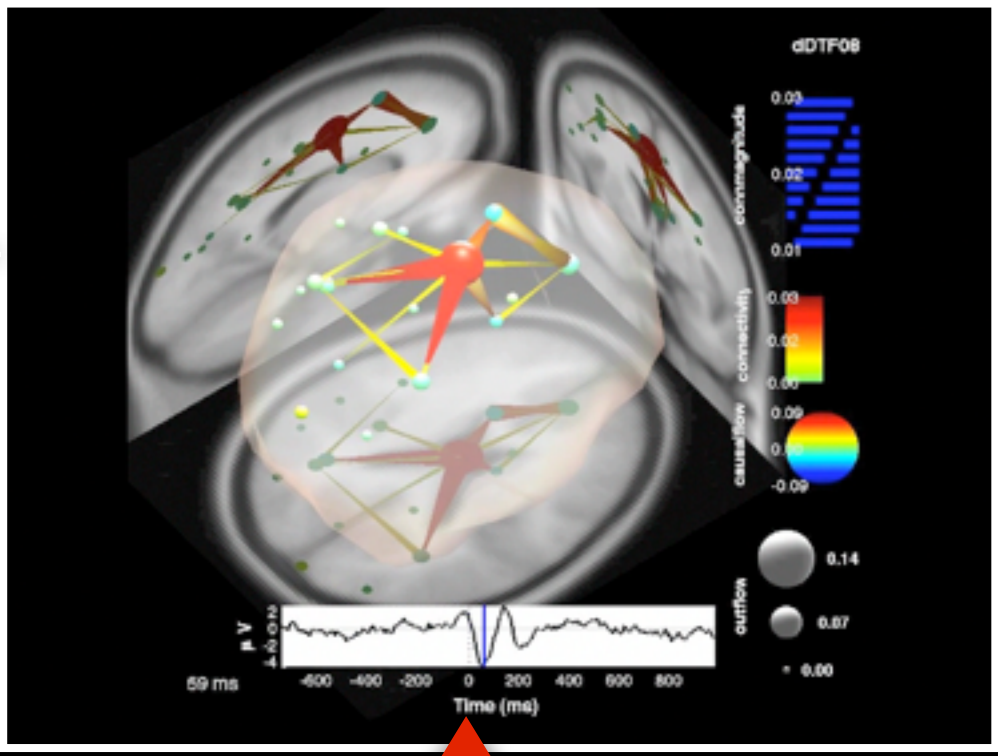
SIFT Online Handbook and User Manual

A video-lecture on the (very) basic theory of application of SIFT to modeling distributed brain dynamics in EEG is available he

Table of Contents [hide]

Some Applications of SIFT

Identification of event-related shifts in effective connectivity which index and predict behavior



error event

Mullen et al, *HBM*, Barcelona, 2010

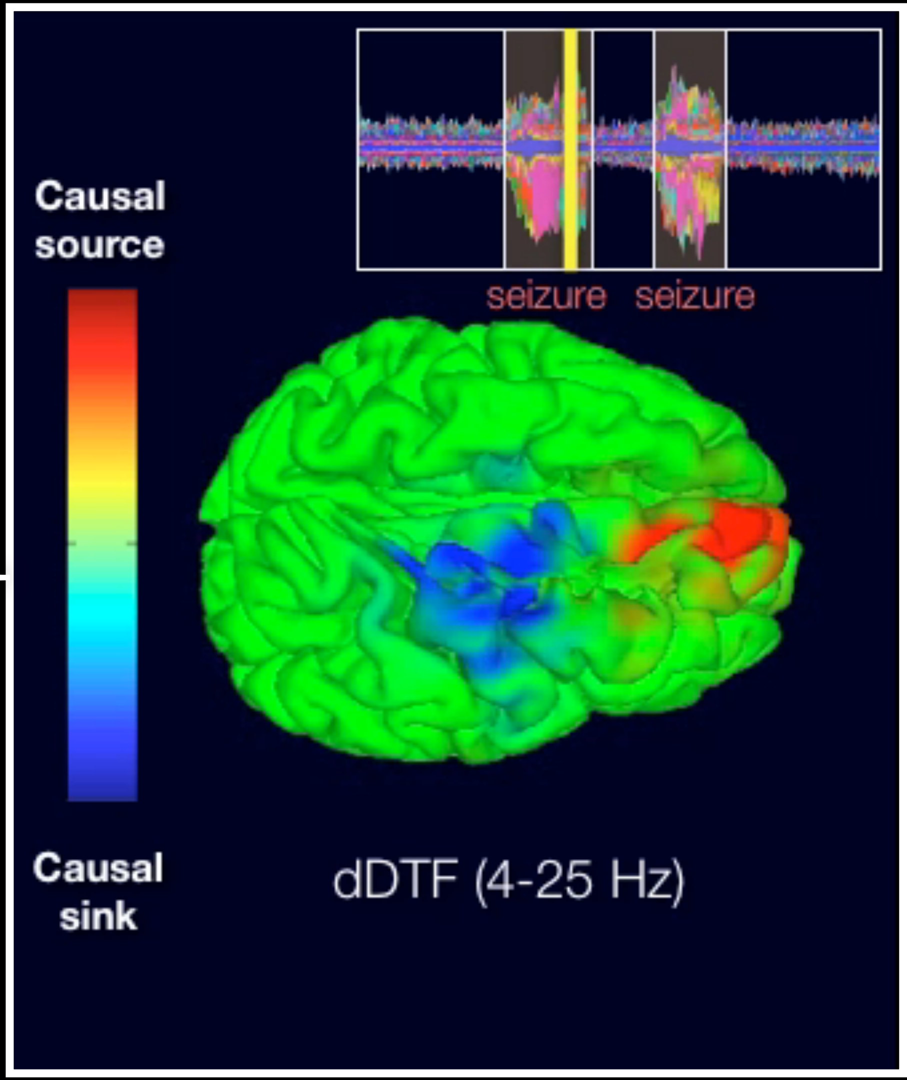
Single-trial spatiotemporal modeling of seizure propagation dynamics

Brain-Computer Interfaces
(Cognitive State Assessment)

Some Applications of SIFT

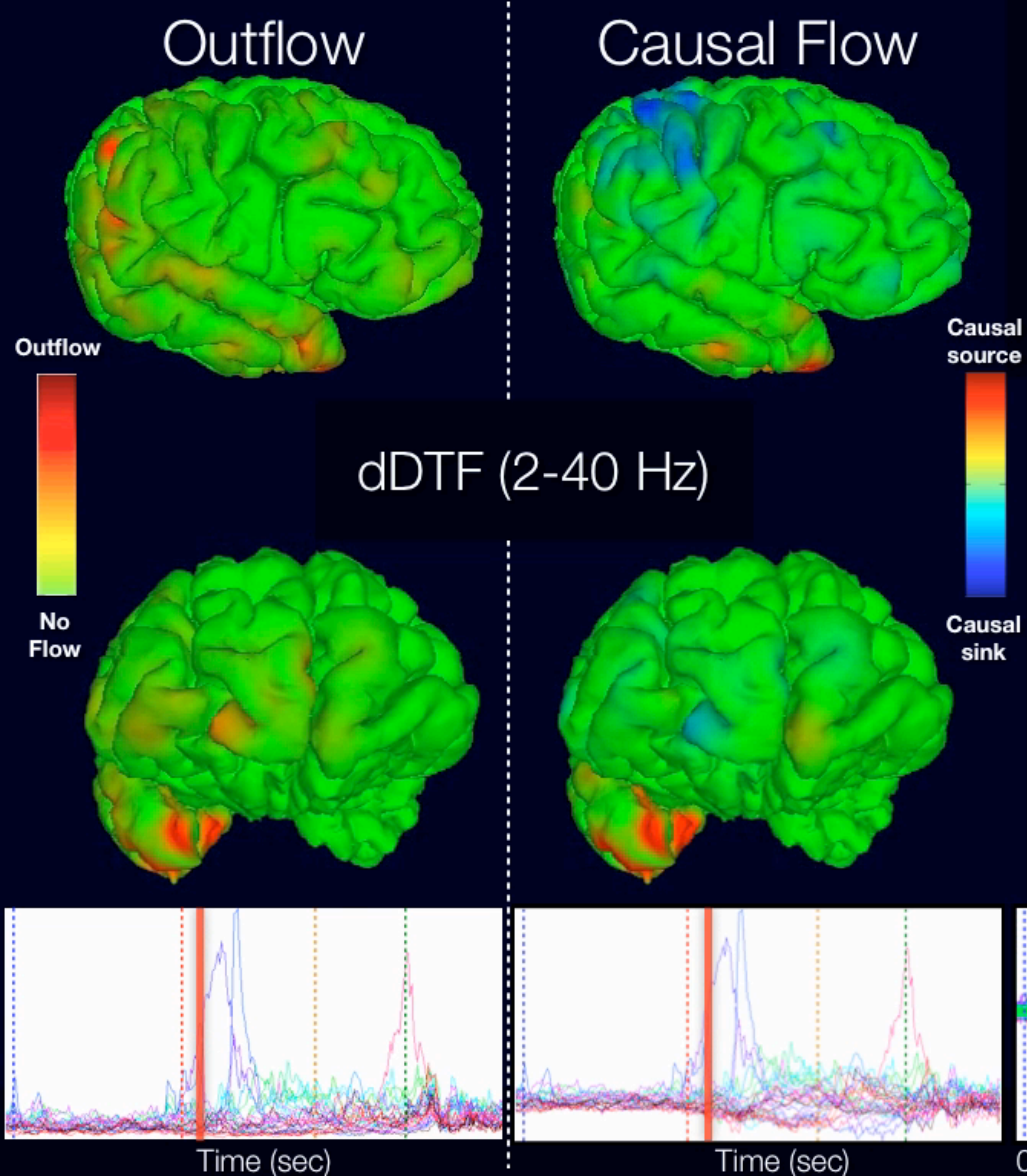
Identification of event-related shifts in effective connectivity which index and predict behavior

Single-trial spatiotemporal modeling of seizure propagation dynamics

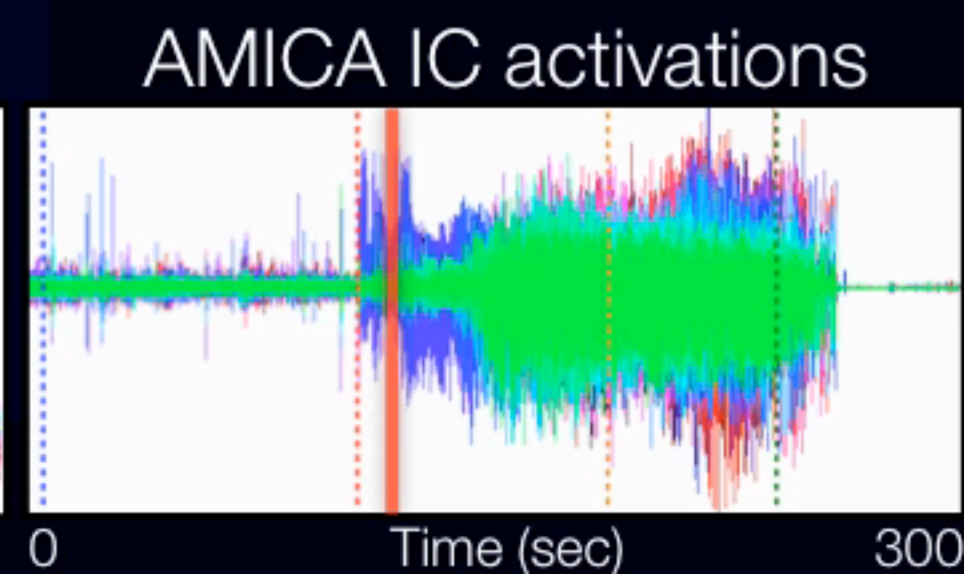


Mullen, Akalin Acar, et al *IEEE EMBC,* 2011

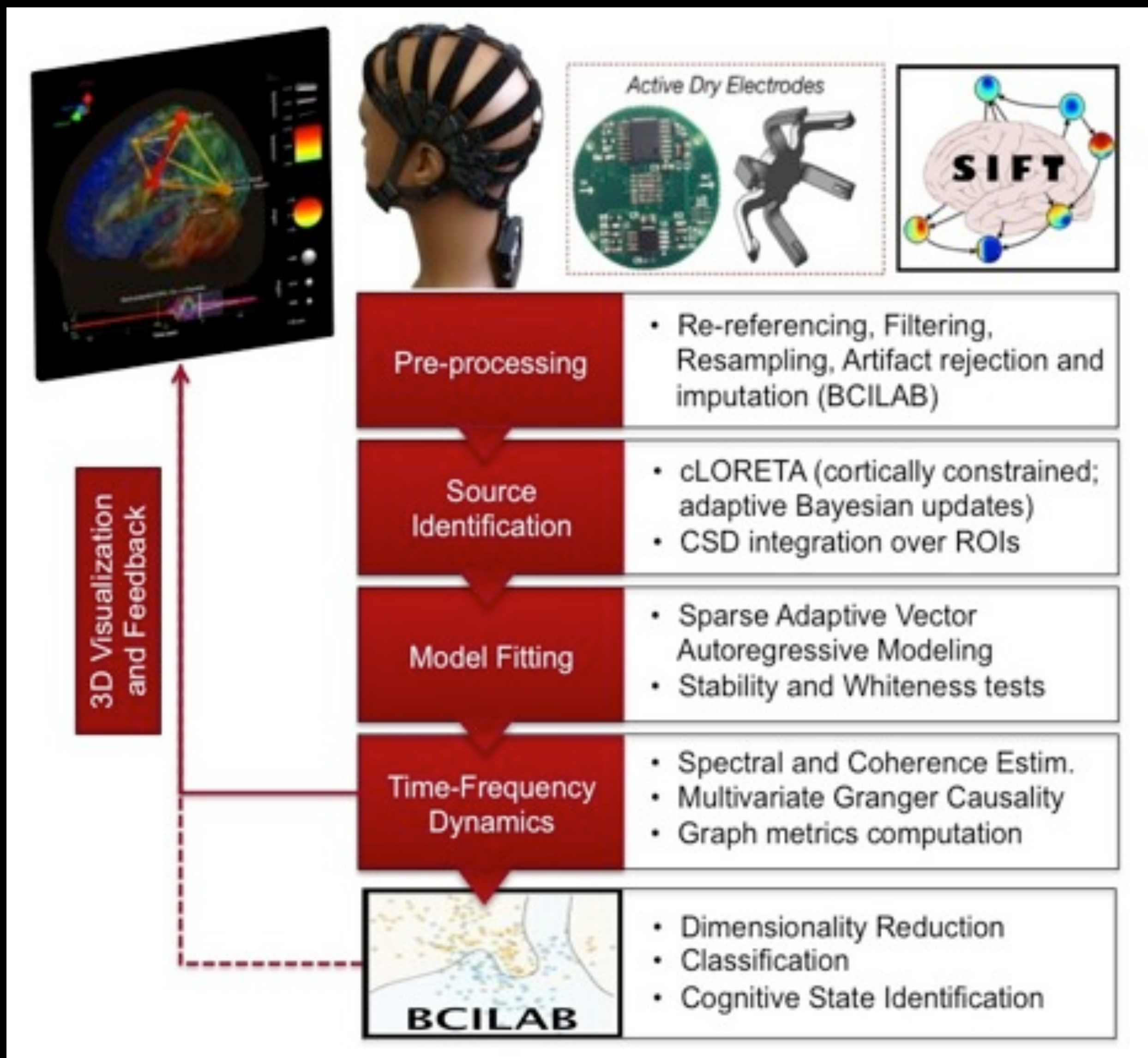
Brain-Computer Interfaces
(Cognitive State Assessment)

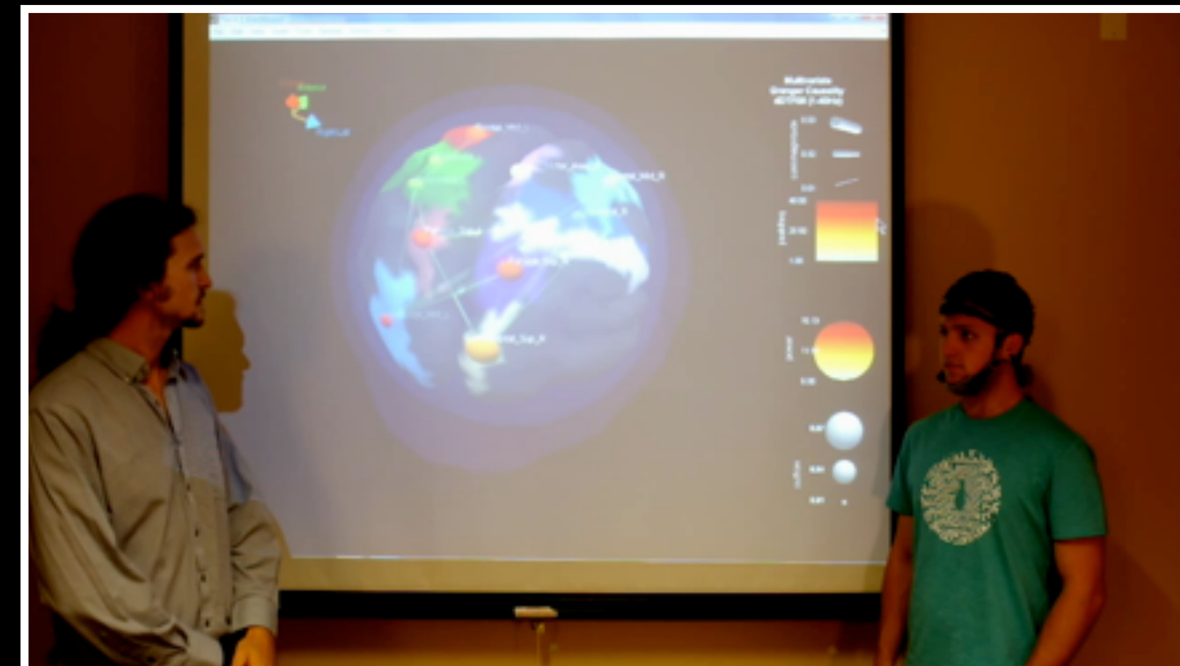


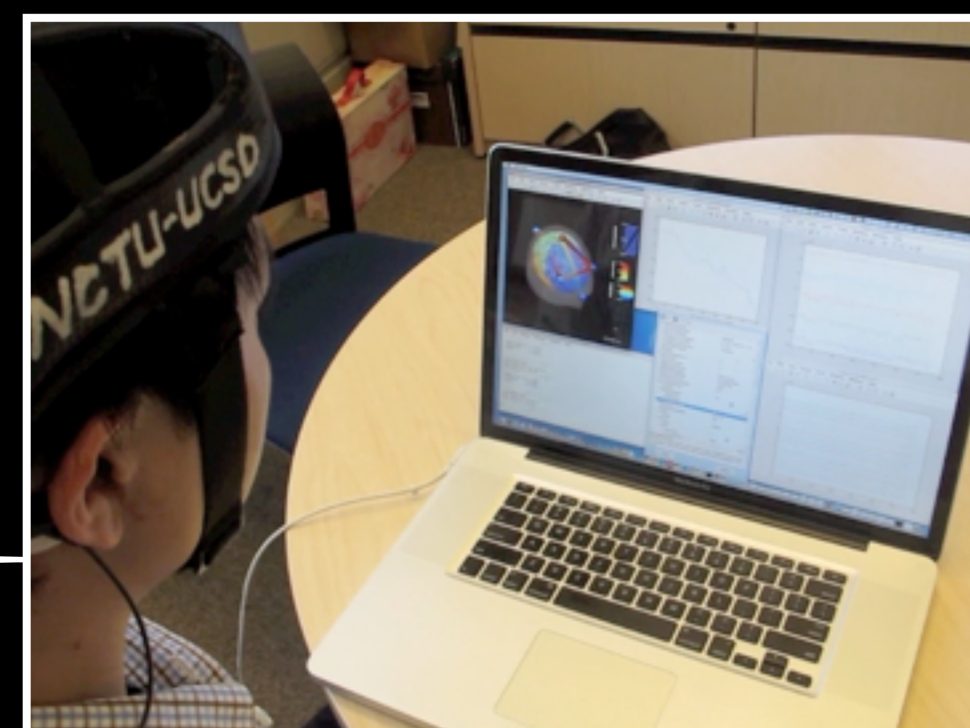
- **ECoG Data**
- 104 ECoG (subdural) electrodes
- Surgical Outcome: Negative
- Provided by Dr. Ashesh Mehta (Feinstein Institute for Clinical Research, NY)
- **Source Reconstruction**
- 5-Model **AMICA** decomposition
- Selection of globally dominant model
- Selection of **29 ICs** comprising independent subspaces
- Component localization via multiscale patch basis **Sparse Bayesian Learning**
- **Sparse VAR Model Fitting**
  - 29-dimensional VAR[12] model
  - Group Lasso (ADMM solver)
  - 5-second windows, 1-sec step
  - Model Order = 12
  - Adaptive regularization selection (step-down from 0.013)
  - direct DTF method of causality



Some Applications of SIFT







Brain-Computer Interfaces:
(Cognitive State Assessment)

Mullen, T., Kothe, C., Chi, Y.M., Ojeda, A., Makeig, S., Cauwenberghs, G., Jung, T-P. (2013). Real-Time Modeling and 3D Visualization of Source Dynamics and Connectivity Using Wearable EEG. *IEEE EMBC*

The Road Ahead

- Public release of new alpha-testing methods with updated online Handbook
- Ongoing incorporation/improvement of sparse VAR, and linear/nonlinear state-space models (Cubature Kalman Filter, EGCA, SCSA, AMIRA)
- Facilitate specification of constraints/priors on dynamic connectivity (e.g. from DTI, anatomy, etc)
- Release and further development of Group Analysis module with multi-subject Bayesian inference and comprehensive statistics (EEGLAB STUDY framework).
- Interfaces with other toolboxes: TRENTOOL (Transfer Entropy), SPM (Dynamic Causal Modeling), Fieldtrip, BCILAB (Brain-Computer Interfaces)
- Improved 2D/3D/4D interactive visualization suite