

# Modeling Distributed Brain Network Dynamics

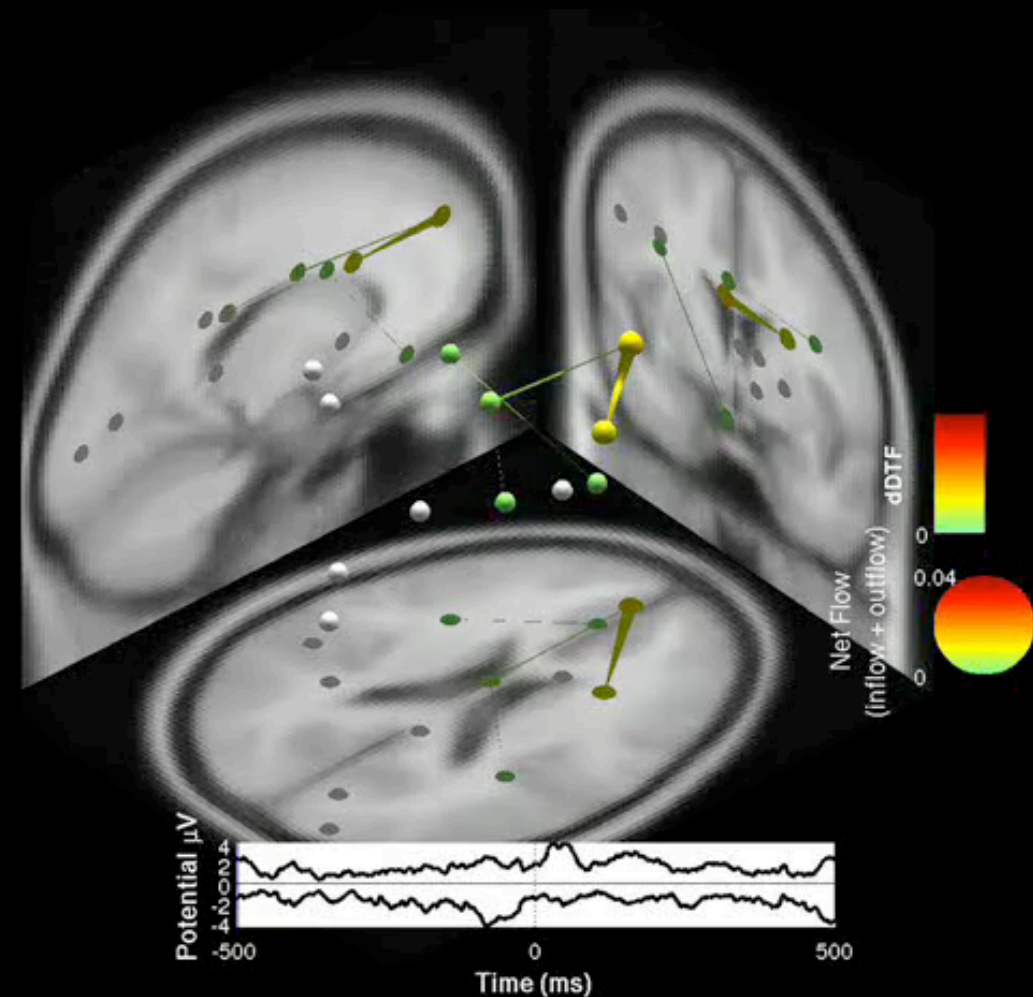
**Tim Mullen**

12th EEGLAB Workshop

UC San Diego

La Jolla, CA

November 20, 2010

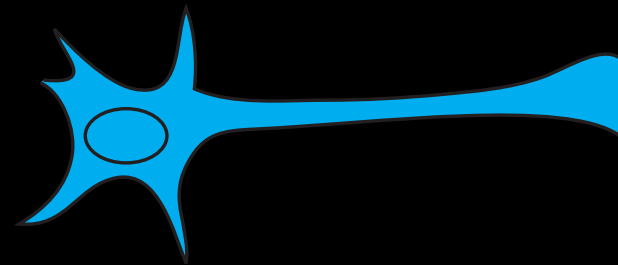


# Different types of connectivity

# Different types of connectivity

- ✦ **Structural Connectivity**

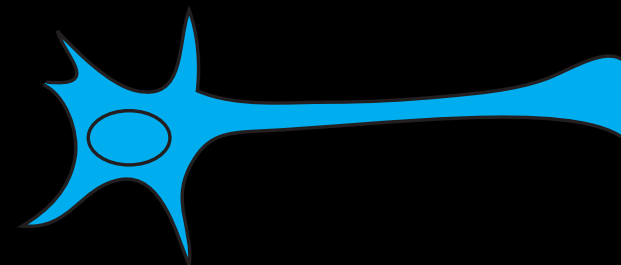
- ✦ anatomical



# Different types of connectivity

- **Structural Connectivity**

- anatomical



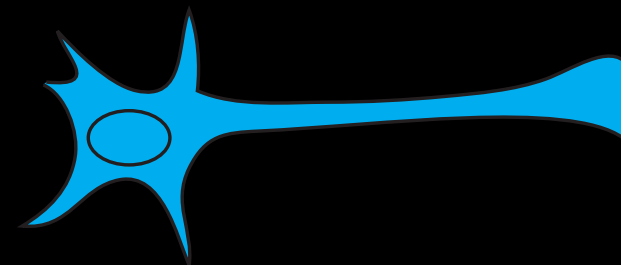
DTI



# Different types of connectivity

- ✦ **Structural Connectivity**

- ✦ anatomical



DTI

- ✦ **Functional Connectivity**

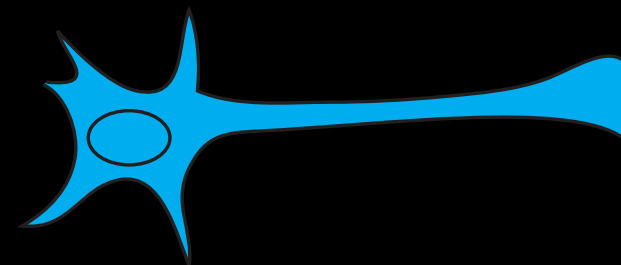
- ✦ symmetric, correlative



# Different types of connectivity

- ✦ **Structural Connectivity**

- ✦ anatomical



DTI

- ✦ **Functional Connectivity**

- ✦ symmetric, correlative

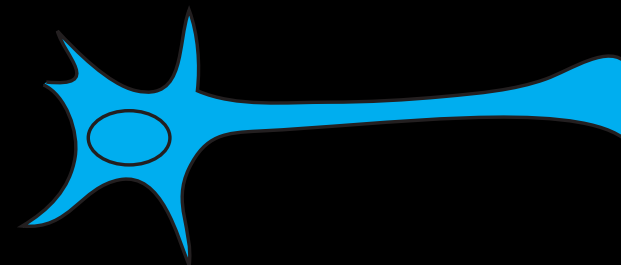


MEG/EEG  
fMRI

# Different types of connectivity

- ✦ **Structural Connectivity**

- ✦ anatomical



DTI

- ✦ **Functional Connectivity**

- ✦ symmetric, correlative



MEG/EEG  
fMRI

- ✦ **Effective Connectivity**

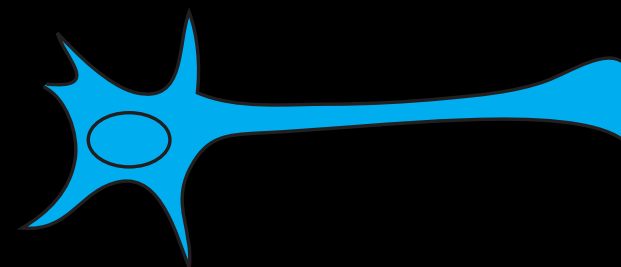
- ✦ asymmetric, causal, information flow



# Different types of connectivity

- ✦ **Structural Connectivity**

- ✦ anatomical



DTI

- ✦ **Functional Connectivity**

- ✦ symmetric, correlative



MEG/EEG  
fMRI

- ✦ **Effective Connectivity**

- ✦ asymmetric, causal, information flow

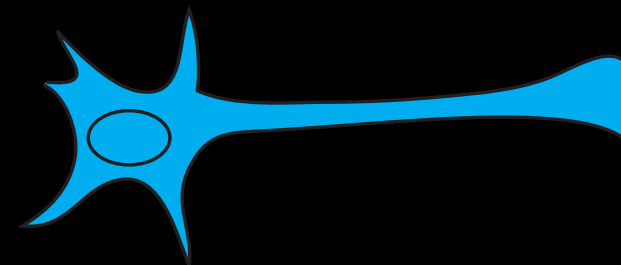


MEG/EEG  
fMRI?

# Different types of connectivity

- ✦ **Structural Connectivity**

- ✦ anatomical



DTI

- ✦ **Functional Connectivity**

- ✦ symmetric, correlative



MEG/EEG  
fMRI

- ✦ **Effective Connectivity**

- ✦ asymmetric, causal, information flow



MEG/EEG  
fMRI?

# Many ways to model effective connectivity in EEG

- ✦ Coherence, Phase-locking value
- ✦ Cross-correlation
- ✦ Transfer Entropy
- ✦ Dynamic Causal Models
- ✦ Structural Equation Models
- ✦ Granger-Causal methods

...

# Many ways to model effective connectivity in EEG

- ✦ Coherence, Phase-locking value
- ✦ Cross-correlation
- ✦ Transfer Entropy
- ✦ Dynamic Causal Models
- ✦ Structural Equation Models
- ✦ Granger-Causal methods

...

# Granger Causality

- ✦ First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- ✦ Relies on two assumptions:
  1. causes should precede their effects in time
  2. information in a cause's past should improve the prediction of the effect, above and beyond the information contained in the effect's own past.

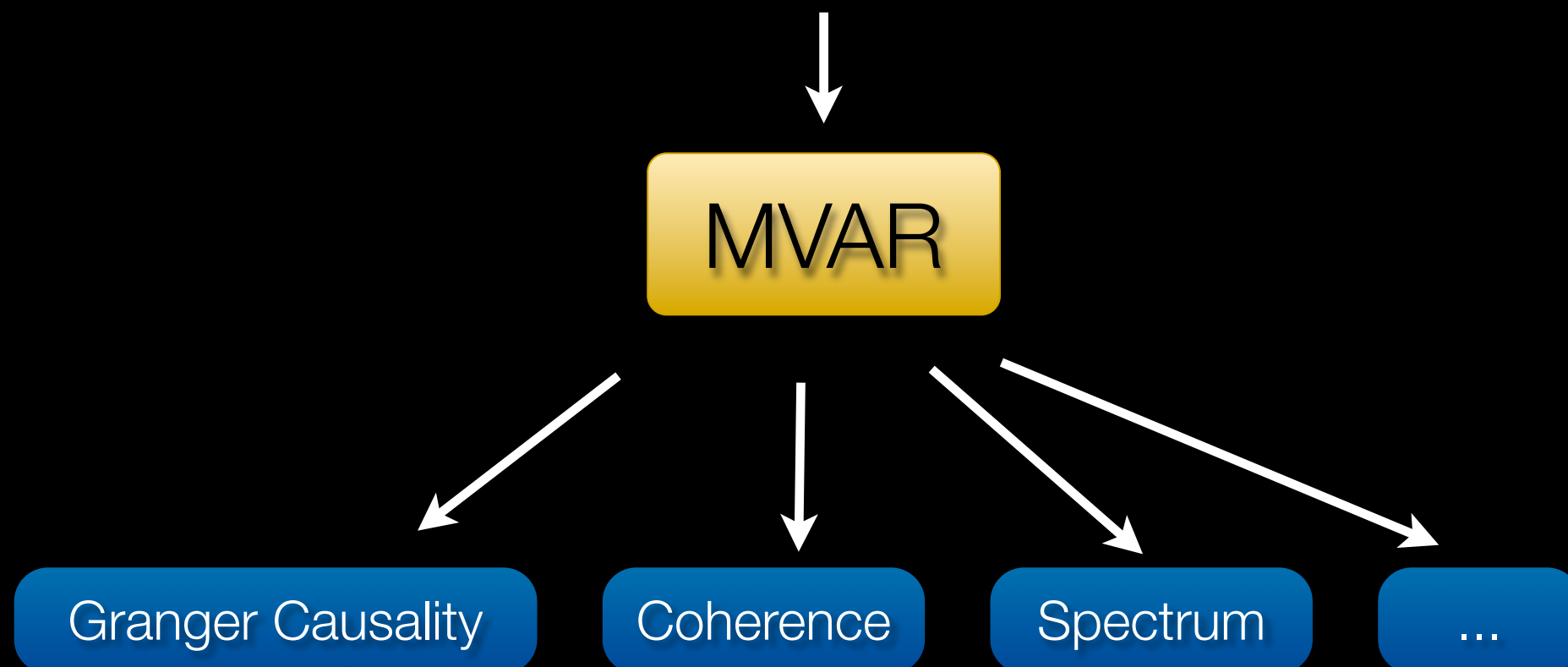
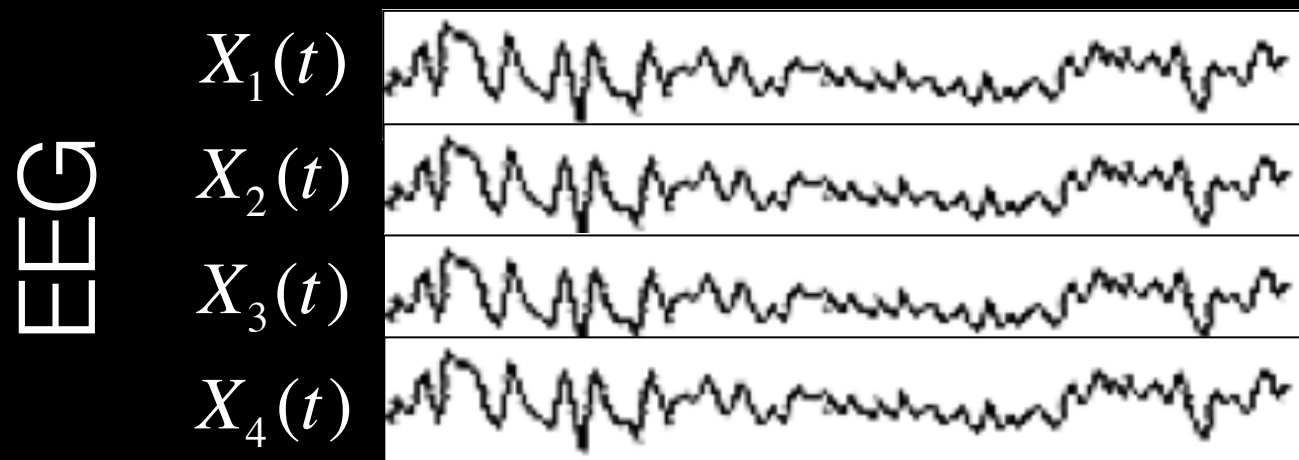


# Granger Causality

- ✦ First introduced by Wiener (1958). Later reformulated by Granger (1969) in the context of linear stochastic autoregressive models
- ✦ Relies on two assumptions:
  1. causes should precede their effects in time
  2. information in a cause's past should improve the prediction of the effect, above and beyond the information contained in the effect's own past.

This is **not** the same as (cross-)correlation!

# Multivariate Autoregressive (MVAR) Modeling



# Multivariate Autoregressive (MVAR) Modeling

- We have M variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

# Multivariate Autoregressive (MVAR) Modeling

- We have M variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

MVAR  
model

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

# Multivariate Autoregressive (MVAR) Modeling

- We have  $M$  variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

MVAR  
model

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

multichannel data  
vector at current time  $t$

# Multivariate Autoregressive (MVAR) Modeling

- We have  $M$  variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

MVAR  
model

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

multichannel data  
vector at current time  $t$

$M \times M$  matrix of model coefficients  
indicating variable dependencies at lag  $k$

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

# Multivariate Autoregressive (MVAR) Modeling

- We have  $M$  variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

MVAR  
model

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

multichannel data  
vector at current time  $t$

$M \times M$  matrix of model coefficients  
indicating variable dependencies at lag  $k$

multichannel data  $k$   
samples in the past

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

# Multivariate Autoregressive (MVAR) Modeling

- We have  $M$  variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

MVAR  
model

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

model order

multichannel data  
vector at current time  $t$

$M \times M$  matrix of model coefficients  
indicating variable dependencies at lag  $k$

multichannel data  $k$   
samples in the past

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



# Multivariate Autoregressive (MVAR) Modeling

- We have  $M$  variables (e.g., EEG channels or source activations):  
 $\mathbf{X}(t) = [\mathbf{X}_1(t), \mathbf{X}_2(t), \dots, \mathbf{X}_M(t)]^T$

MVAR  
model

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

model order

random noise process

multichannel data  
vector at current time  $t$

$M \times M$  matrix of model coefficients  
indicating variable dependencies at lag  $k$

multichannel data  $k$   
samples in the past

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

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

# Multivariate Autoregressive (MVAR) Modeling

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

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

# Multivariate Autoregressive (MVAR) Modeling

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

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



entropy rate (amount of prediction error)

# Multivariate Autoregressive (MVAR) Modeling

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

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

Penalizes high model orders (parsimony)

entropy rate (amount of prediction error)

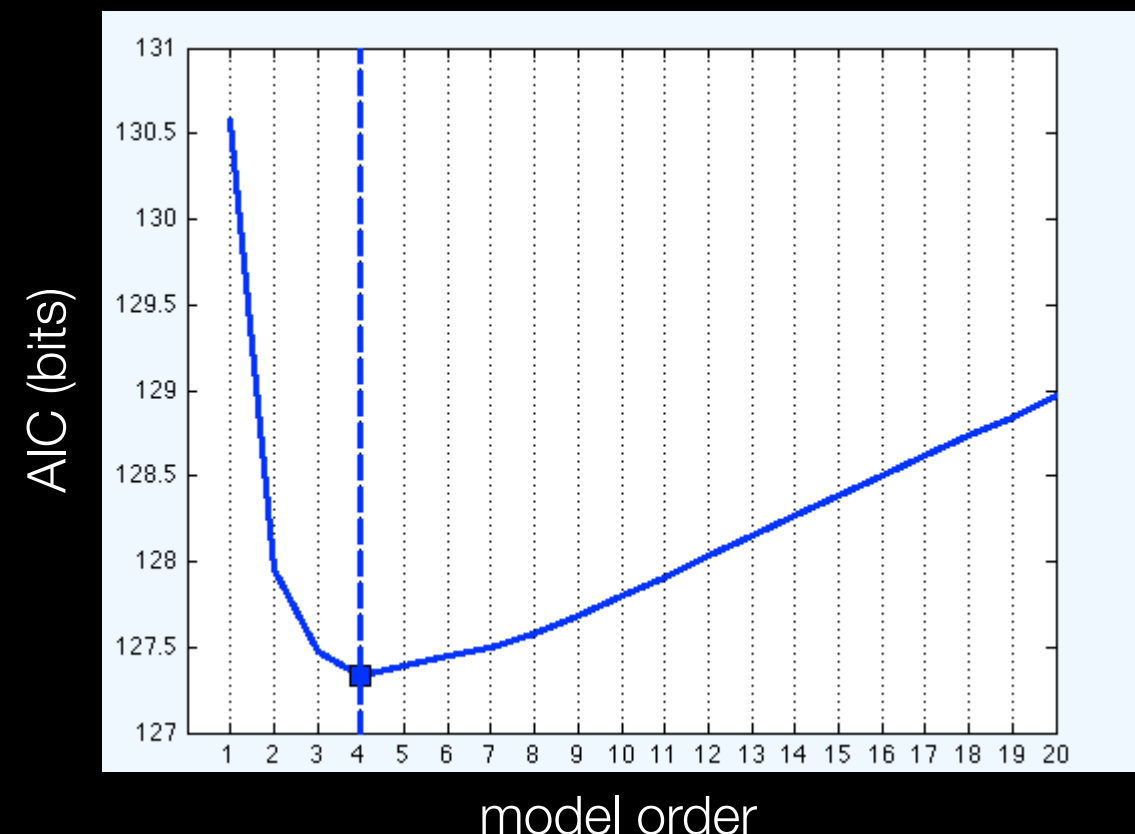
# Multivariate Autoregressive (MVAR) Modeling

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

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

Penalizes high model orders (parsimony)

entropy rate (amount of prediction error)



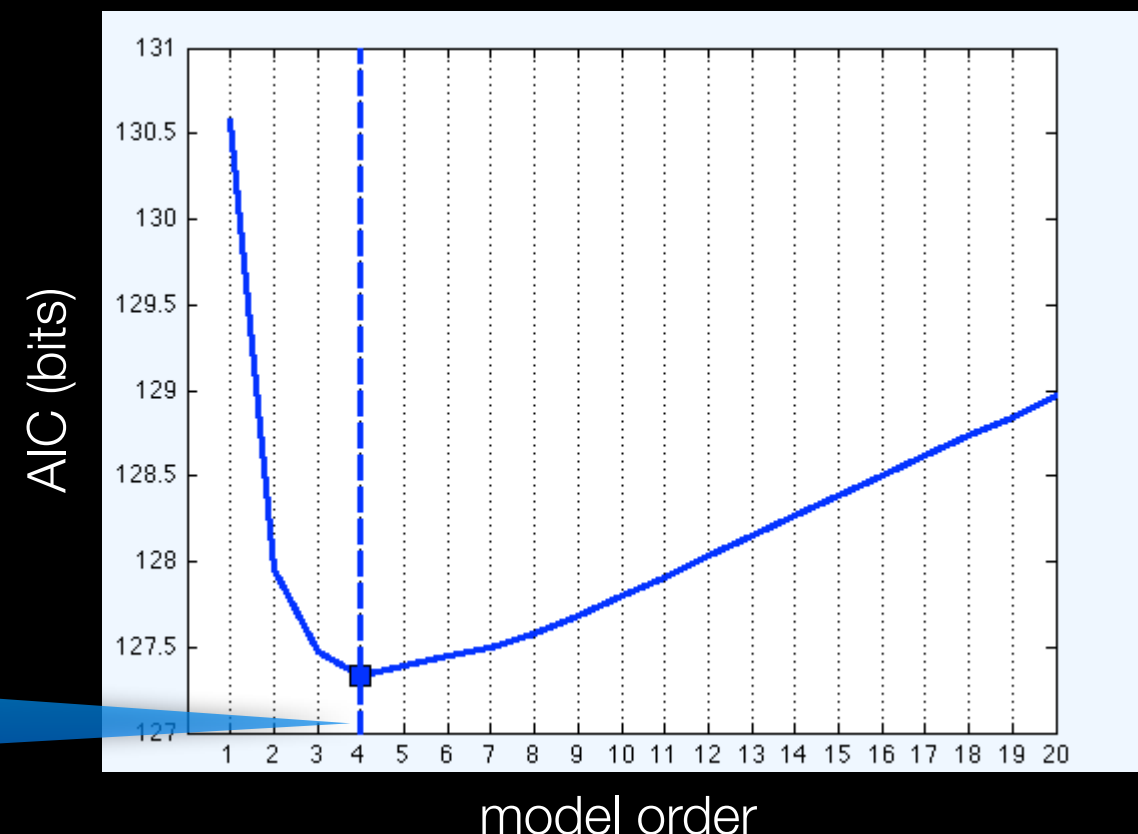
# Multivariate Autoregressive (MVAR) Modeling

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

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

Penalizes high model orders (parsimony)

entropy rate (amount of prediction error)



optimal order

# MVAR Modeling: Assumptions

# MVAR Modeling: Assumptions

- **“Weak” stationarity of the data**



# MVAR Modeling: Assumptions

- ✦ **“Weak” stationarity of the data**
  - ✦ mean and variance do not change with time

# MVAR Modeling: Assumptions

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

# MVAR Modeling: Assumptions

- **“Weak” stationarity of the data**

- mean and variance do not change with time

- An EEG trace containing ERPs is a classic example of a non-stationary time-series

- **Stability**

# MVAR Modeling: Assumptions

- **“Weak” stationarity of the data**

- mean and variance do not change with time
- An EEG trace containing ERPs is a classic example of a non-stationary time-series

- **Stability**

- Technically, an MVAR process is stable if the reverse characteristic polynomial of the process has all roots outside the complex unit circle (all eigenvalues of  $\mathbf{A}$  have modulus less than 1)

# MVAR Modeling: Assumptions

- **“Weak” stationarity of the data**

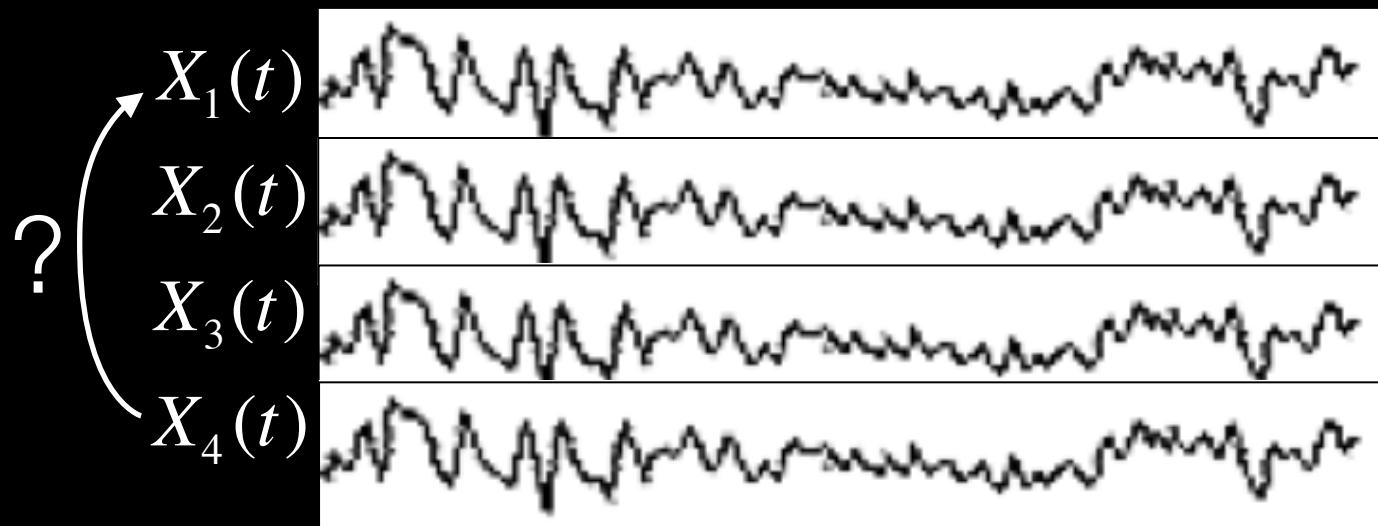
- mean and variance do not change with time
- An EEG trace containing ERPs is a classic example of a non-stationary time-series

- **Stability**

- Technically, an MVAR process is stable if the reverse characteristic polynomial of the process has all roots outside the complex unit circle (all eigenvalues of  $\mathbf{A}$  have modulus less than 1)
- Importantly, stability implies stationarity and SIFT provides you techniques for verifying the stability

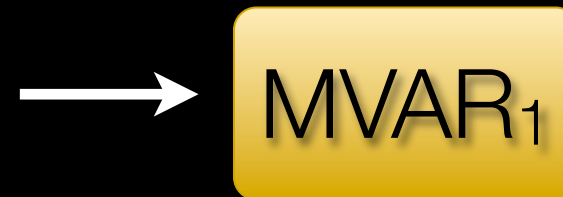
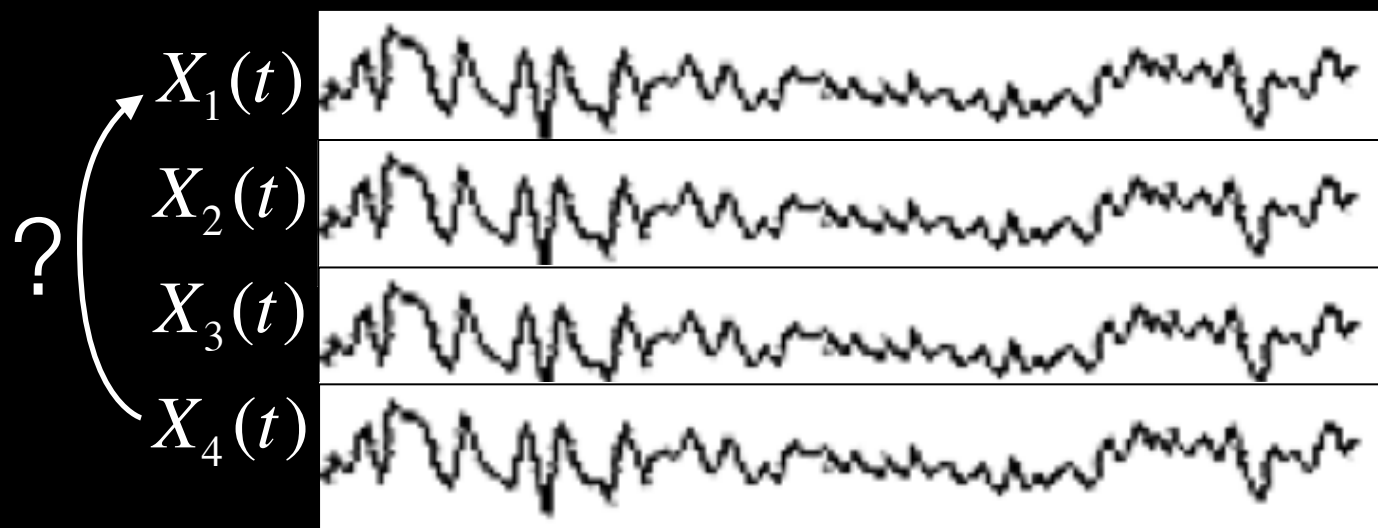
# Granger Causality

Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



# Granger Causality

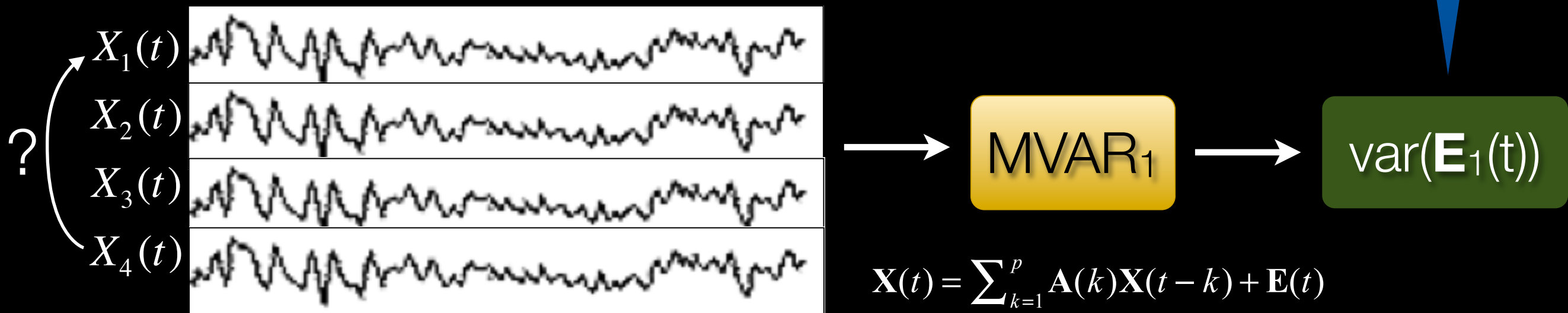
Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



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

# Granger Causality

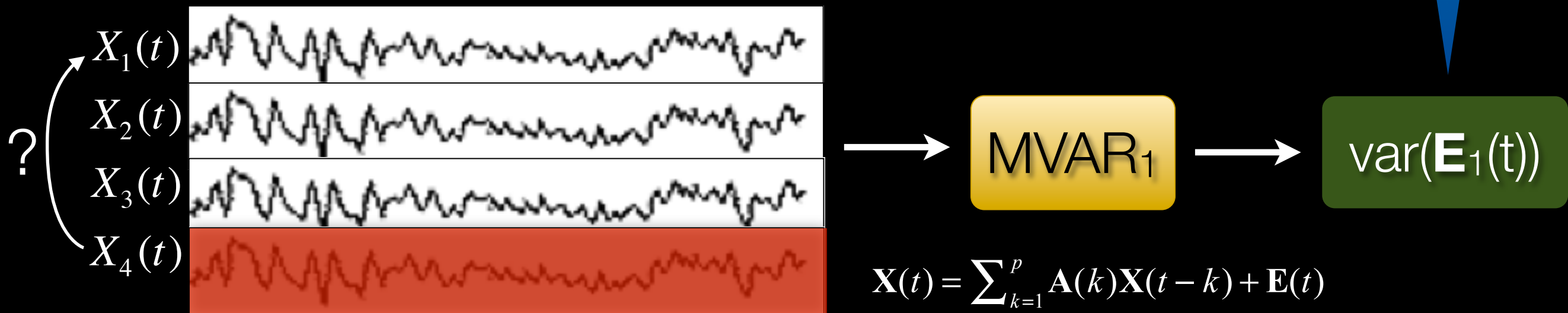
Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?





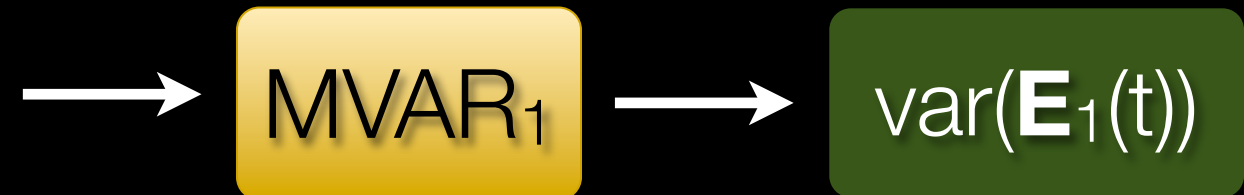
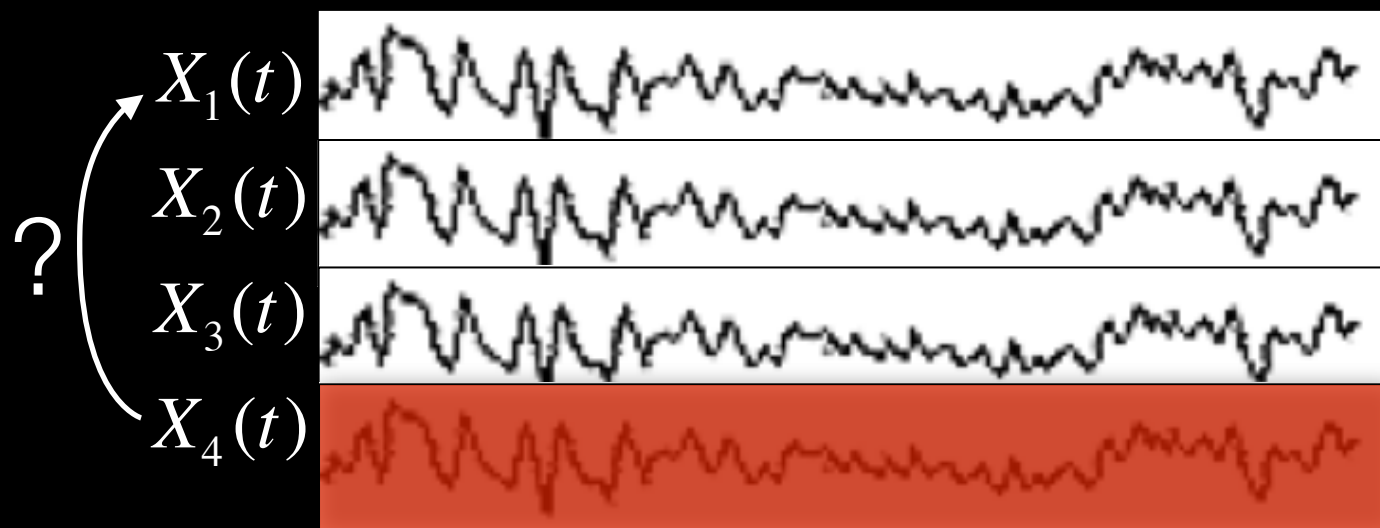
# Granger Causality

Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



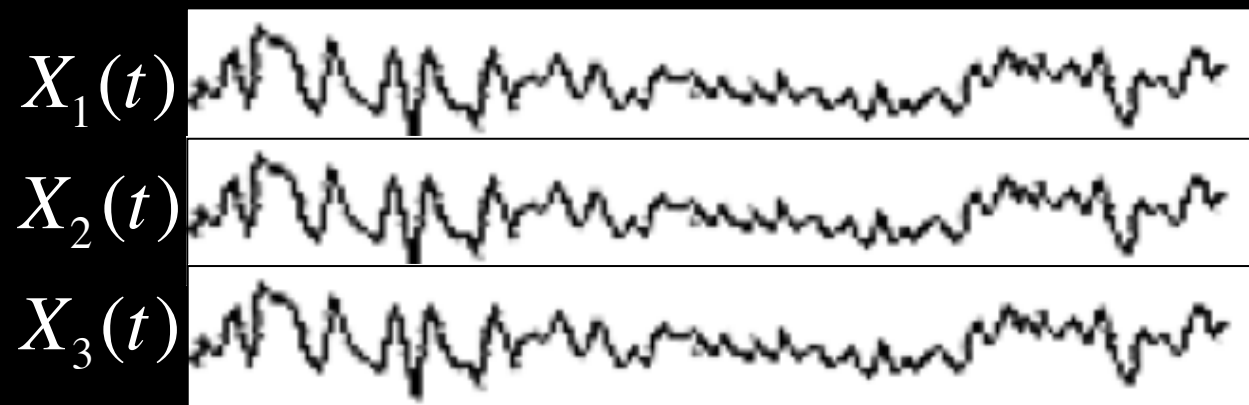
# Granger Causality

Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



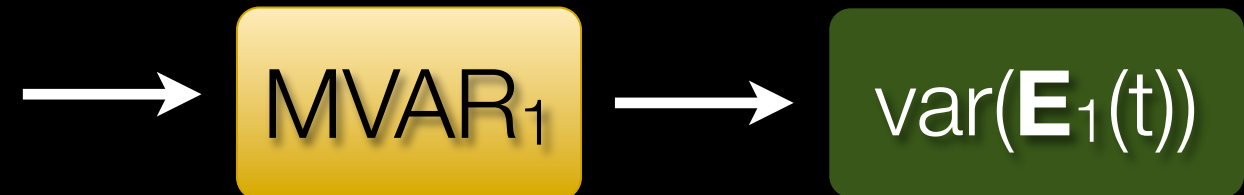
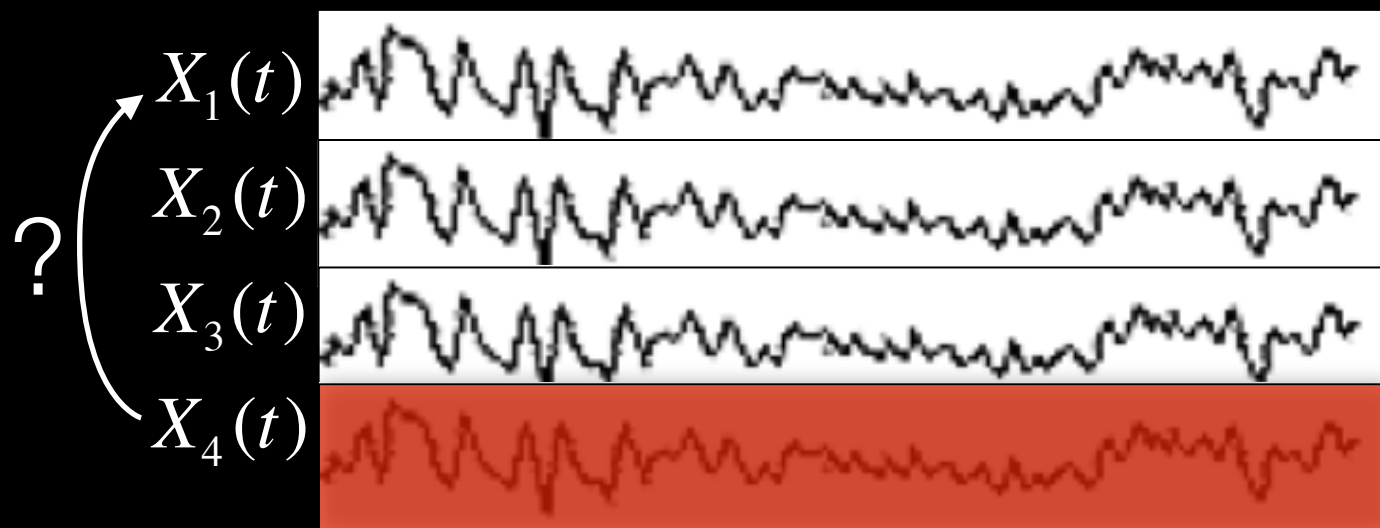
prediction  
error for  $X_1$

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



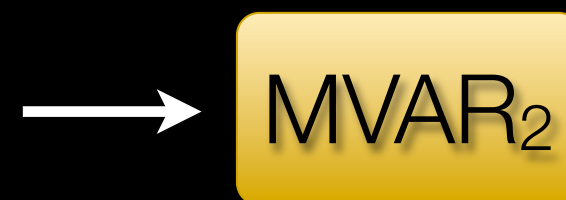
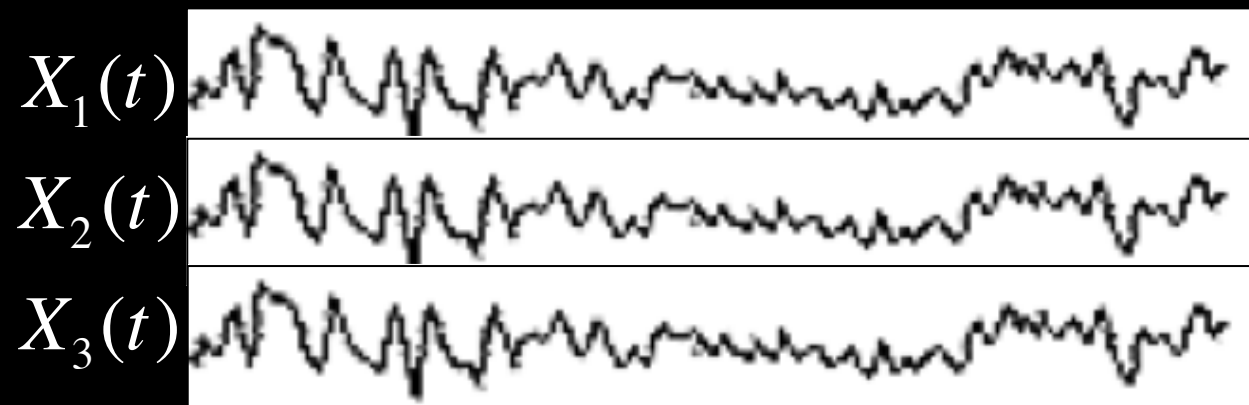
# Granger Causality

Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



prediction  
error for  $X_1$

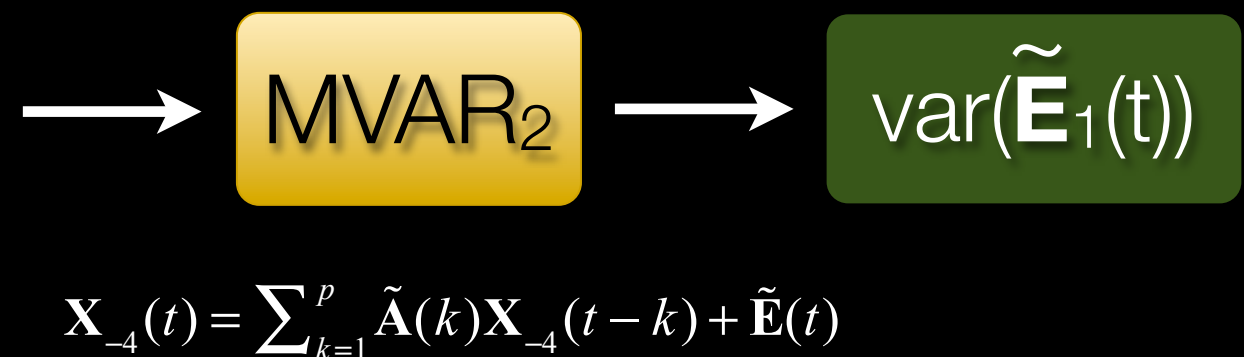
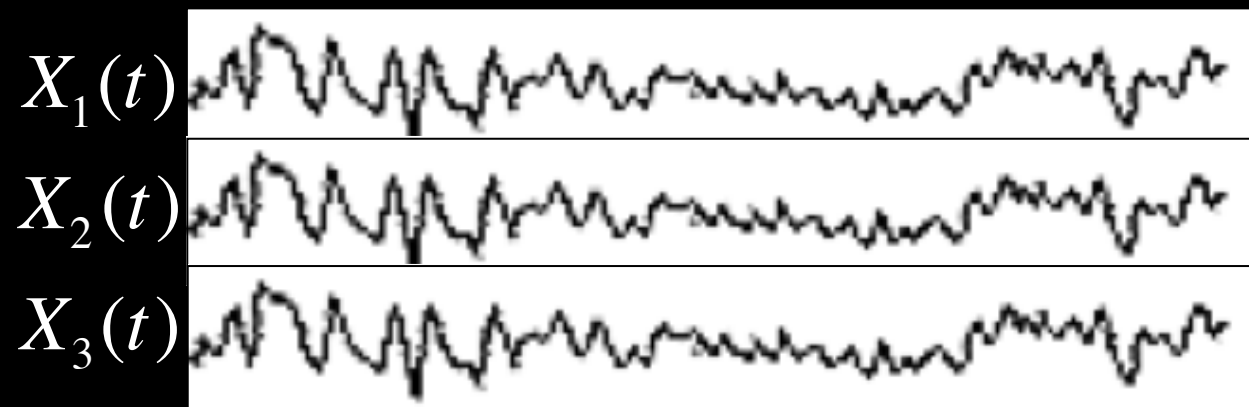
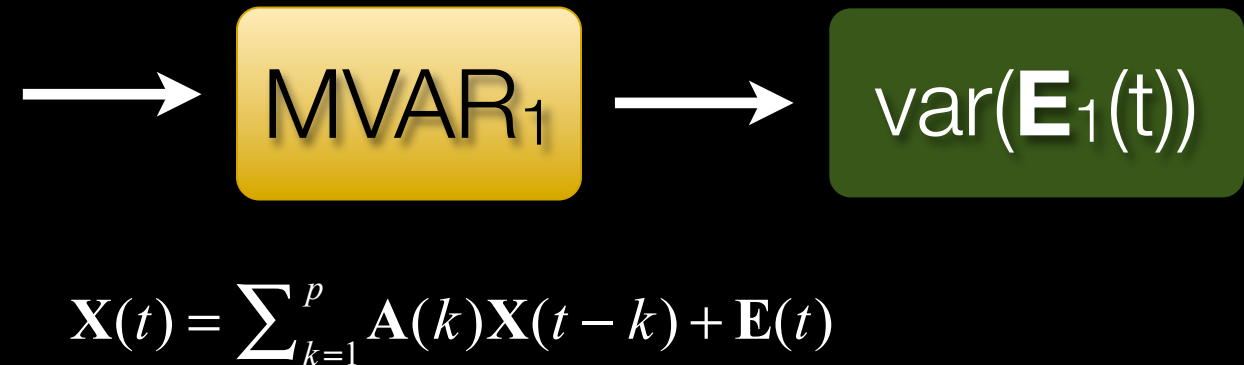
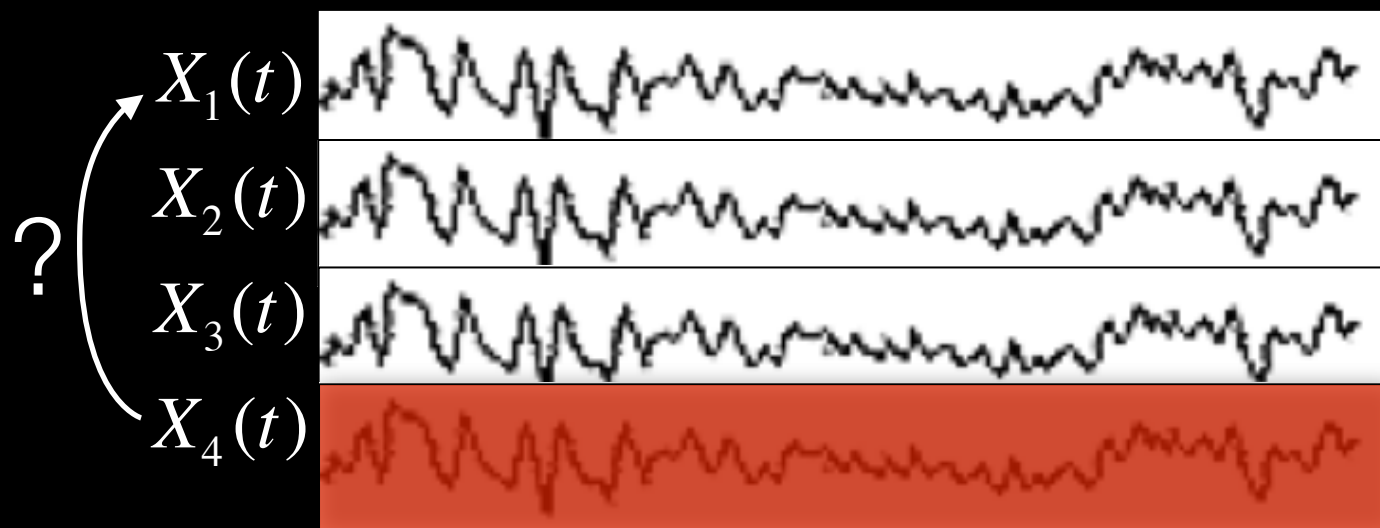
$$\mathbf{X}(t) = \sum_{k=1}^p \mathbf{A}(k) \mathbf{X}(t-k) + \mathbf{E}(t)$$



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

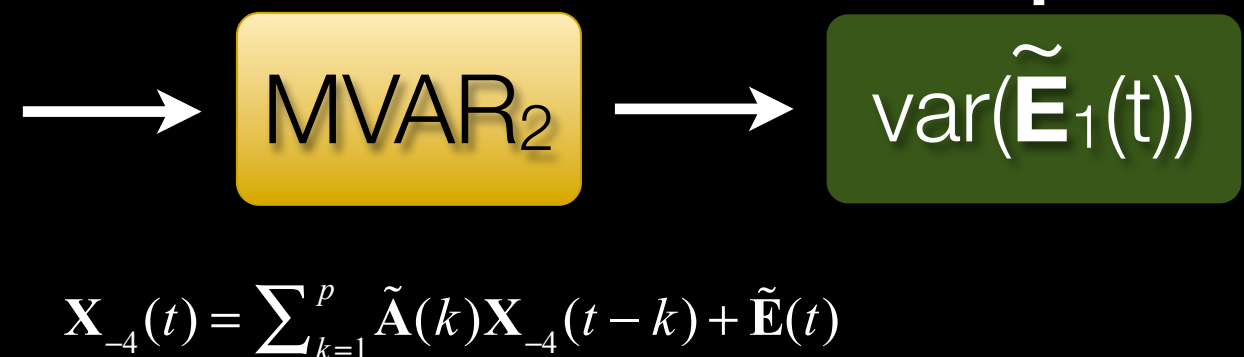
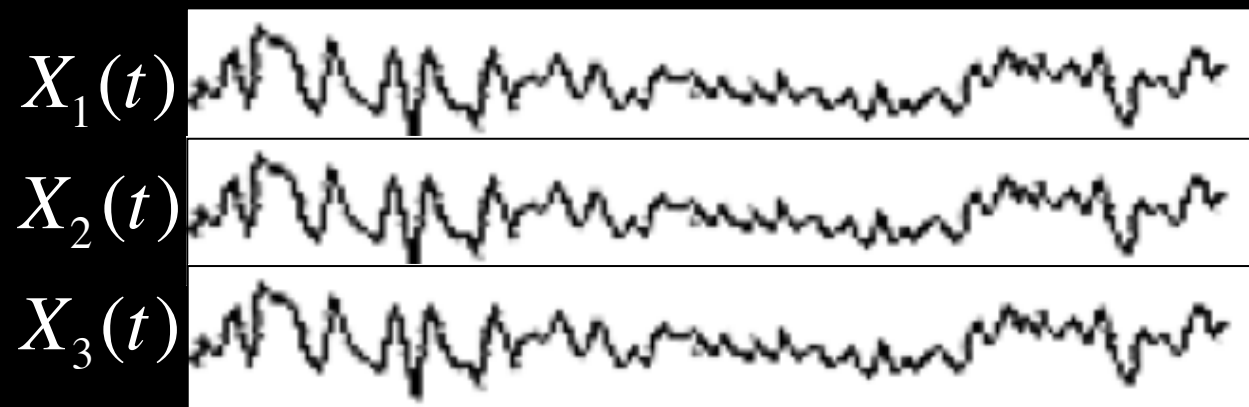
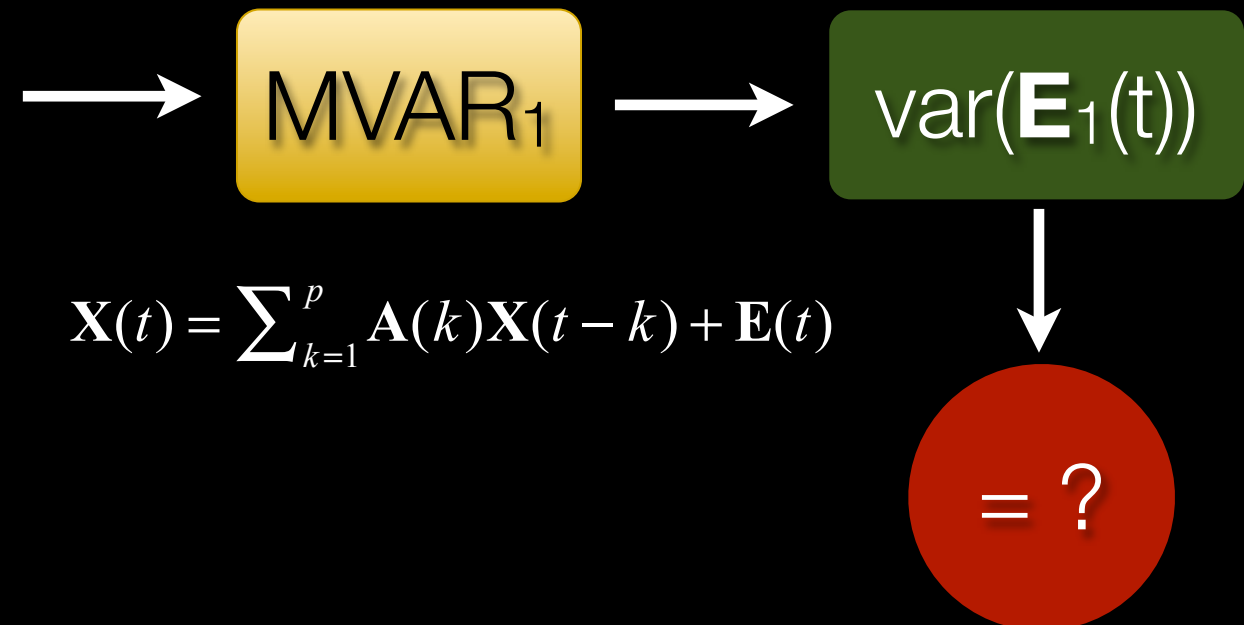
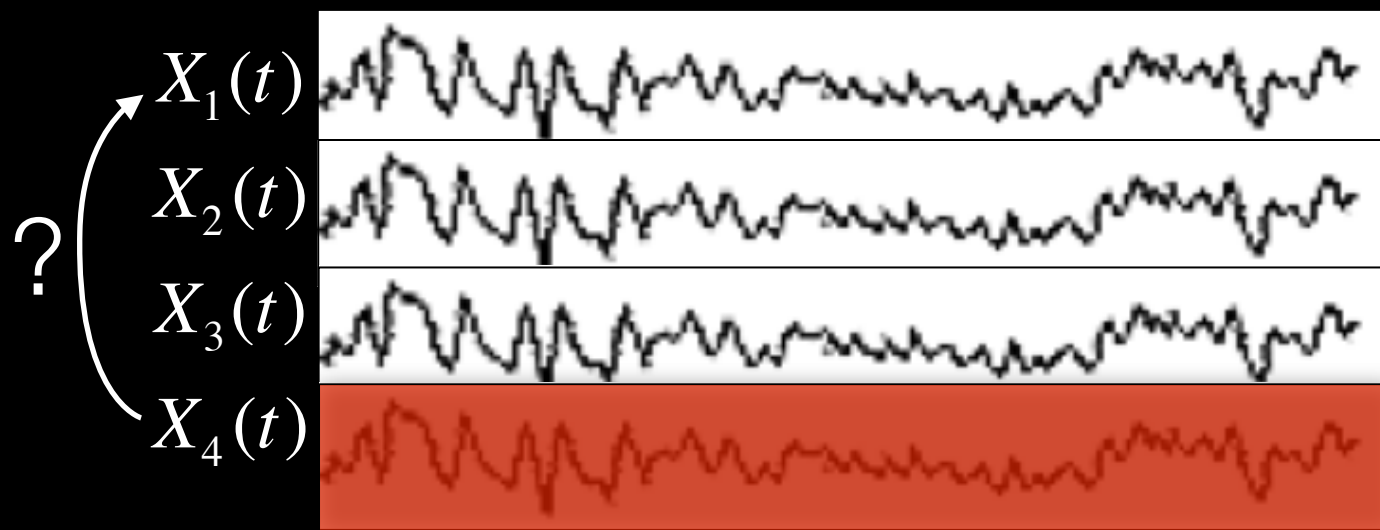
# Granger Causality

Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



# Granger Causality

Test: Does  $\mathbf{X}_4$  granger-cause  $\mathbf{X}_1$  ?



# Granger Causality

# Granger Causality

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

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

# Granger Causality

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

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

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

*$X_j$  granger-causes  $X_i$  condition on all other variables in  $\mathbf{X}$*   
if and only if  $\mathbf{A}_{ij}(k) \gg 0$  for some lag  $k \in \{1, \dots, p\}$



# 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)     b)     c) 

# 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?



# 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?



# Granger Causality Quiz

- Example: 2-channel MVAR process of order 1

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

The diagram illustrates the causal structure of the MVAR model. A red curved arrow with a large 'X' above it points from  $X_2(t-1)$  to  $X_1(t)$ , indicating that this path is not active (coefficient is 0). A green curved arrow points from  $X_1(t-1)$  to  $X_2(t)$ , indicating an active causal link (coefficient is 0.7). The terms  $0X_2(t-1)$  and  $0.7X_1(t-1)$  are highlighted with red and green boxes, respectively.

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

Which causal structure does this model correspond to?

a) 1  $\longrightarrow$  2

b) 1  $\longleftarrow$  2

c) 1  $\longleftrightarrow$  2

# Granger Causality – Frequency Domain

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

# 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

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

Likewise,  $\mathbf{X}(f)$  and  $\mathbf{E}(f)$  correspond to the fourier transforms of the data and residuals, respectively

# 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

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

Likewise,  $\mathbf{X}(f)$  and  $\mathbf{E}(f)$  correspond to the fourier transforms of the data and residuals, respectively

We can then define the spectral matrix  $\mathbf{X}(f)$  as follows:

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

Where  $\mathbf{H}(f)$  is the *transfer matrix* of the system.

# 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

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

Likewise,  $\mathbf{X}(f)$  and  $\mathbf{E}(f)$  correspond to the fourier transforms of the data and residuals, respectively

We can then define the spectral matrix  $\mathbf{X}(f)$  as follows:

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

Where  $\mathbf{H}(f)$  is the *transfer matrix* of the system.

**Definition:** If  $|\mathbf{A}_{ij}(f)|$  is significantly non-zero, then  $X_j$  *granger-causes*  $X_i$  (at frequency  $f$ ) conditioned on all other vars in  $\mathbf{X}$



# Granger Causality – Frequency Domain Estimators

✱ (some) Coherence measures

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

Coherence

---

$$P_{ij}(f) = \frac{\hat{S}_{ij}(f)}{\sqrt{\hat{S}_{ii}(f)\hat{S}_{jj}(f)}}$$

$$\hat{\mathbf{S}} = \mathbf{S}^{-1}$$

Partial coherence

---

$$G_i(f) = \sqrt{1 - \frac{\det(\mathbf{S}(f))}{S_{ii}(f)\mathbf{M}_{ii}(f)}}$$

Multiple coherence

$\mathbf{S}(f) = \mathbf{X}(f)\mathbf{X}(f)^*$  is the **spectral density matrix** of  $\mathbf{X}$

# Granger Causality – Frequency Domain Estimators

✱ (some) GC measures

$$\theta_{ij}^2(f) = |H_{ij}(f)|^2$$

(non-normalized) Directed Transfer Function (DTF)

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{k=1}^M |H_{ik}(f)|^2}$$

Normalized DTF

$$\delta_{ij}^2(f) = \eta_{ij}^2(f) P_{ij}^2(f) \quad \text{where} \quad \eta_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_f \sum_{k=1}^M |H_{ik}(f)|^2}$$

ffDTF

partial coherence

Direct DTF

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

Normalized Partial Directed Coherence (PDC)

# Time-Frequency GC

# Time-Frequency GC

- ✦ Brain network dynamics often change rapidly with time
  - ✦ event-related responses
  - ✦ transient network changes during sequential information processing

# Time-Frequency GC

- ✦ Brain network dynamics often change rapidly with time
  - ✦ event-related responses
  - ✦ transient network changes during sequential information processing
- ✦ Electrophysiological processes often exhibit oscillatory phenomena, making them well-suited for frequency-domain analysis

# Time-Frequency GC

- ✦ Brain network dynamics often change rapidly with time
  - ✦ event-related responses
  - ✦ transient network changes during sequential information processing
- ✦ Electrophysiological processes often exhibit oscillatory phenomena, making them well-suited for frequency-domain analysis
- ✦ 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**
  - ✦ Sliding-window adaptive multivariate autoregression (AMVAR)

# Time-Frequency GC

- ✦ **Many ways to do time-varying MVAR estimation**
  - ✦ Sliding-window adaptive multivariate autoregression (AMVAR)
  - ✦ Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)

# Time-Frequency GC

- ✦ **Many ways to do time-varying MVAR estimation**
  - ✦ Sliding-window adaptive multivariate autoregression (AMVAR)
  - ✦ Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
  - ✦ Kalman Filtering

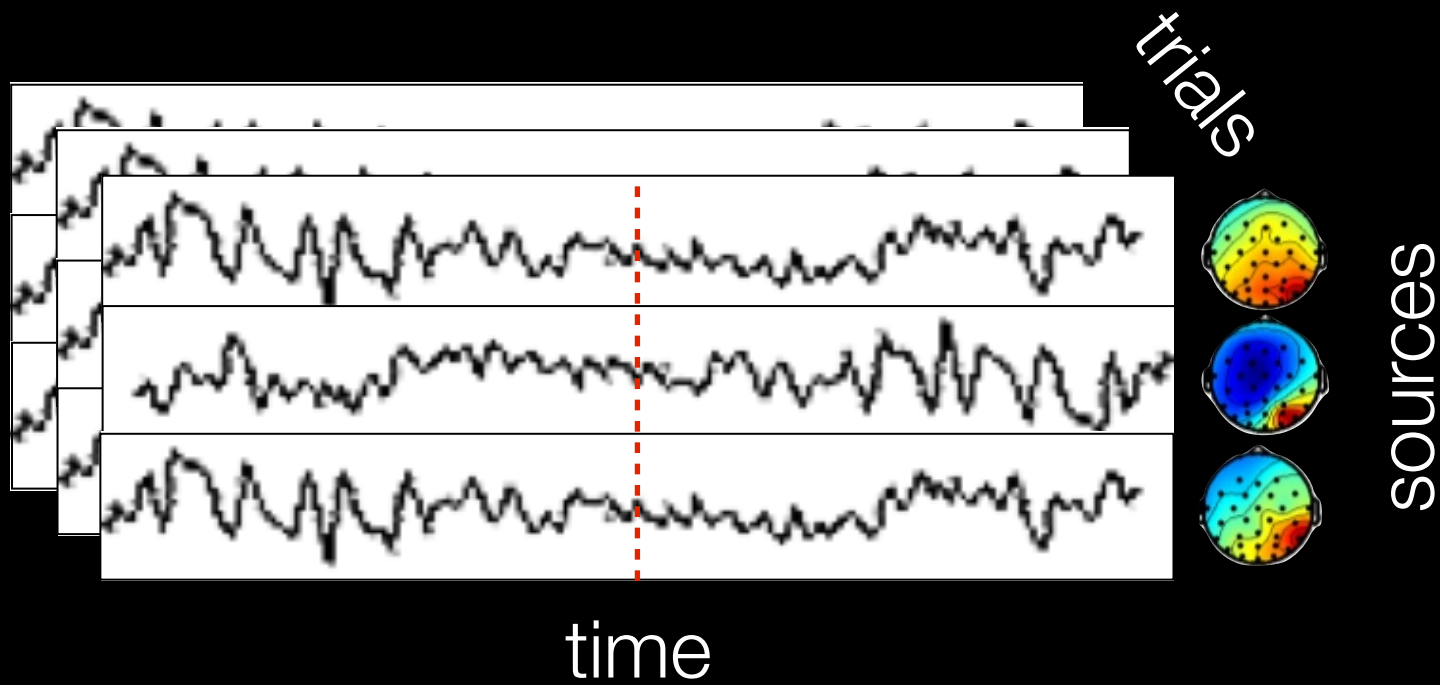
# Time-Frequency GC

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

# Time-Frequency GC

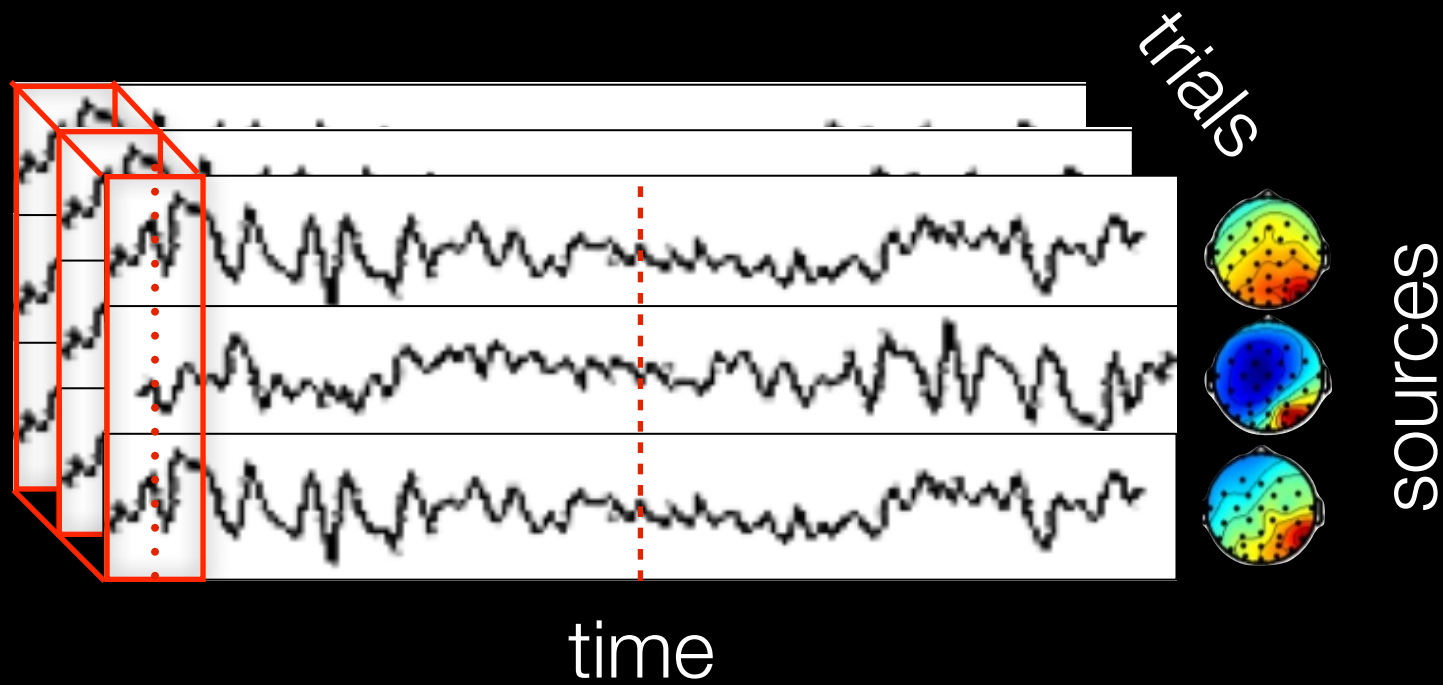
- ✦ **Many ways to do time-varying MVAR estimation**
  - ✦ Sliding-window adaptive multivariate autoregression (AMVAR)
  - ✦ Non-parametric MVAR estimation (minimum-phase spectral matrix factorization)
  - ✦ Kalman Filtering
  - ✦ ...

# Time-Frequency GC



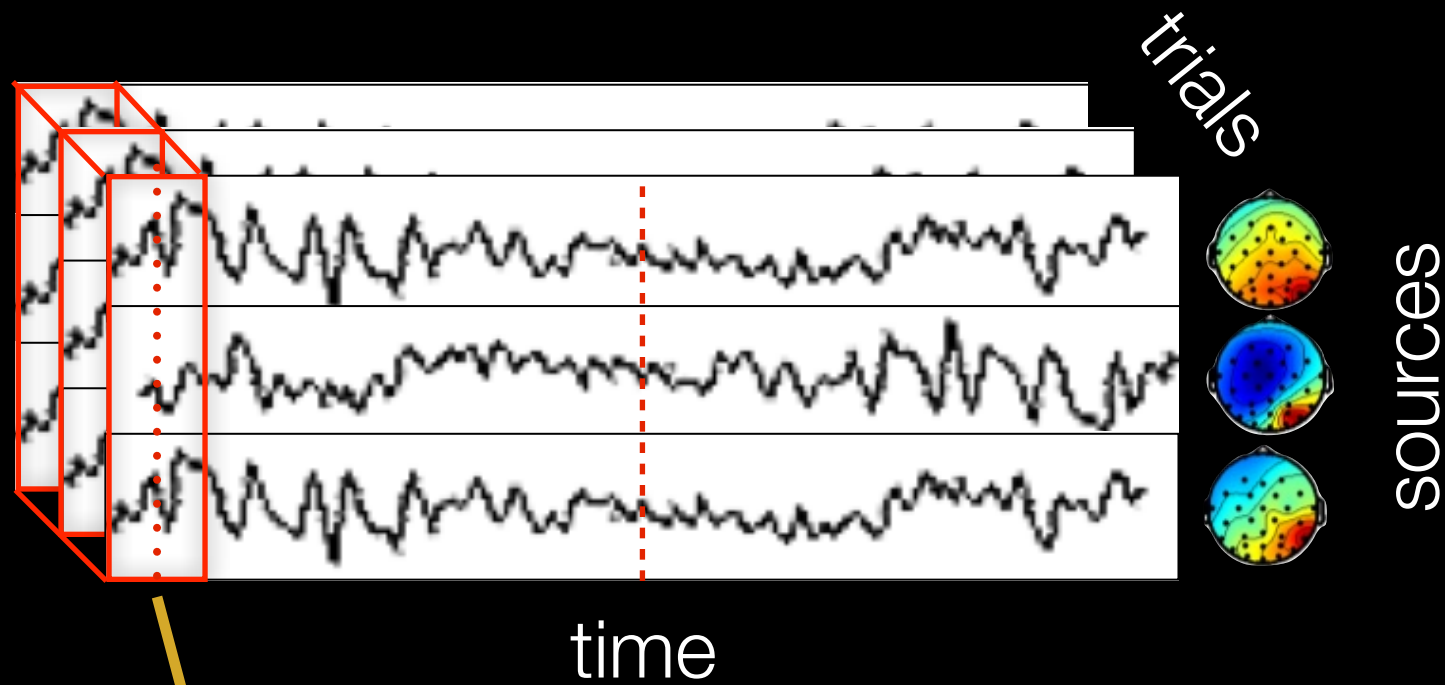
Analogous to short-time fourier transform

# Time-Frequency GC



Analogous to short-time fourier transform

# Time-Frequency GC



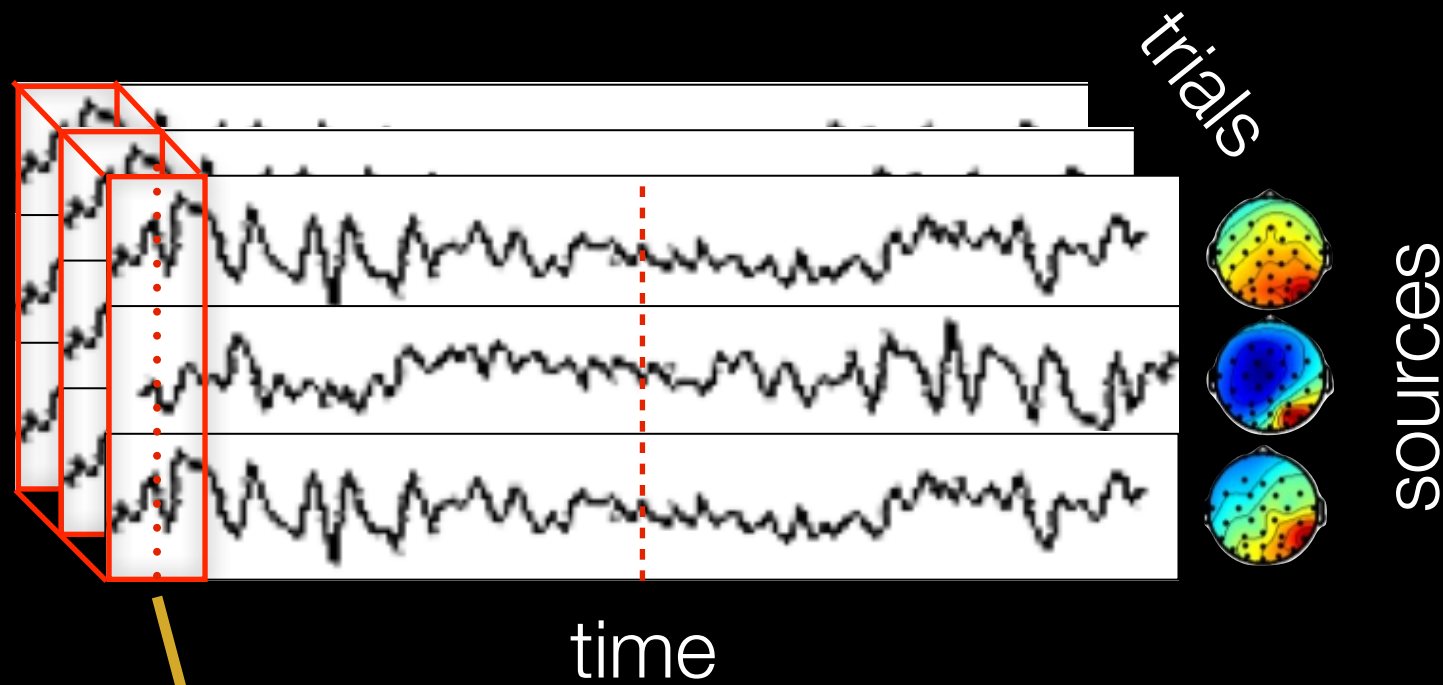
Analogous to short-time fourier transform

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

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



# Time-Frequency GC



Analogous to short-time fourier transform

ensemble normalization

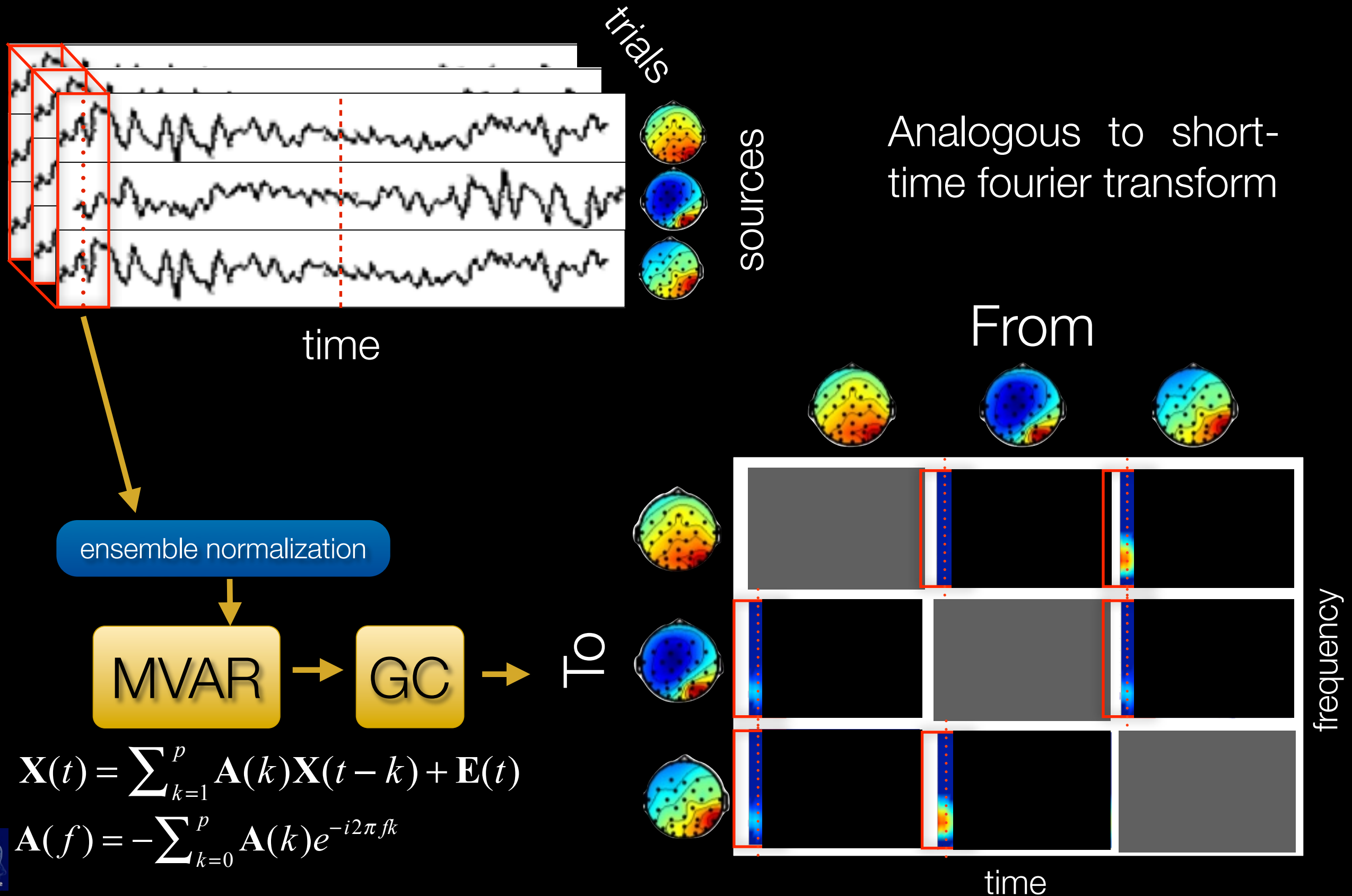
MVAR

GC

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

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

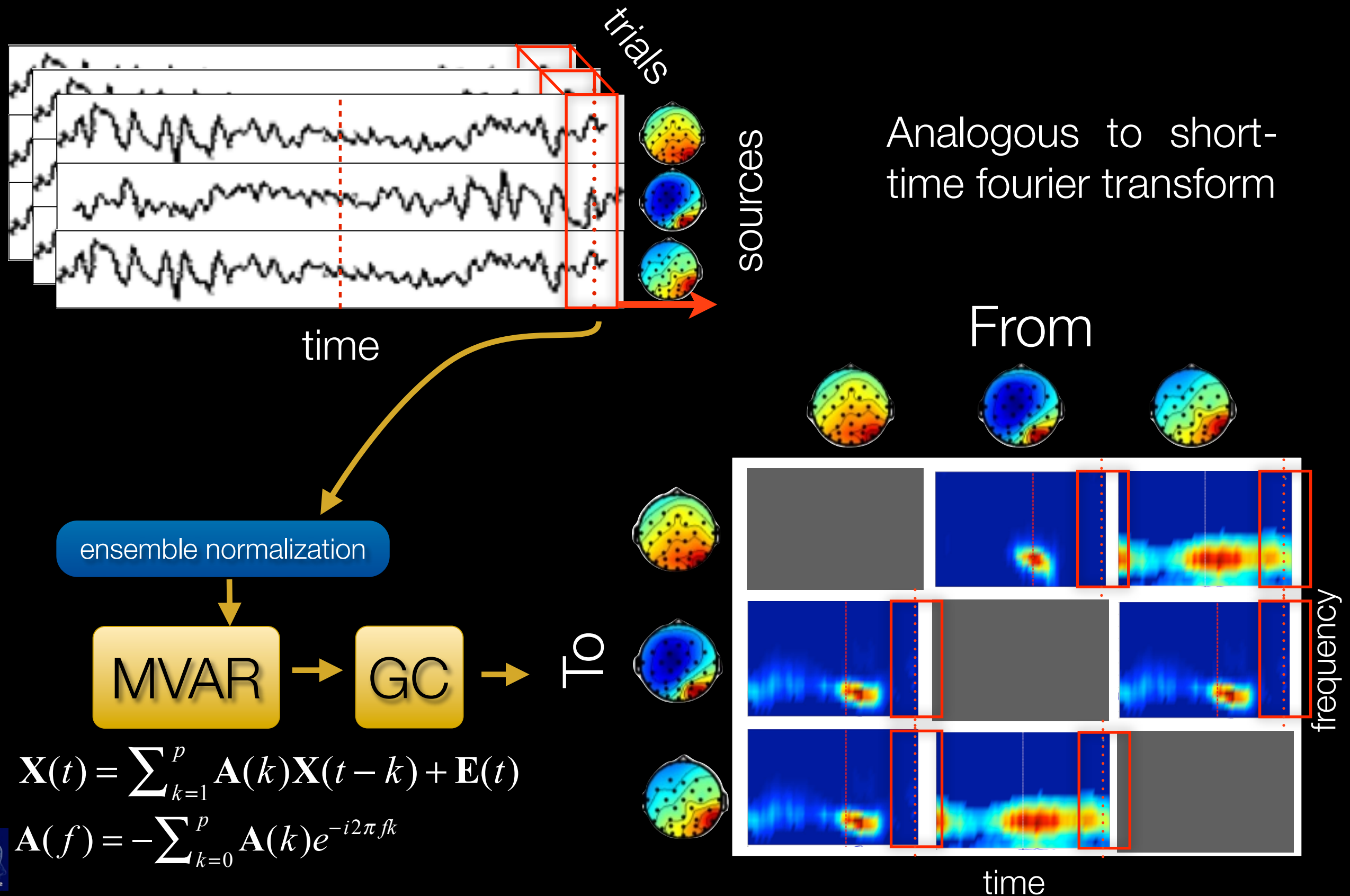
# Time-Frequency GC



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

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

# Time-Frequency GC



# 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

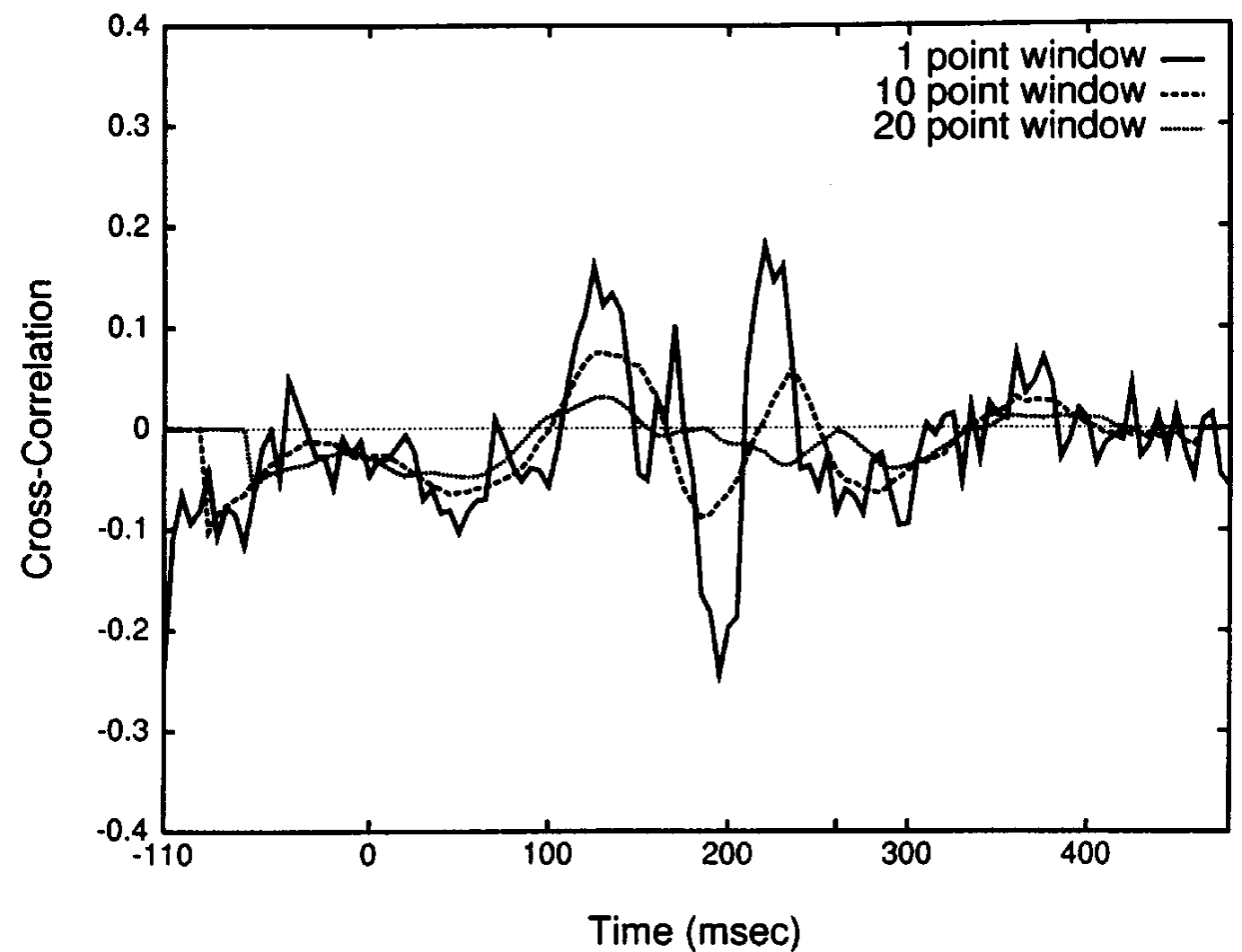
- ✧ **What is a good window length?**
- ✧ Considerations:
  - ✧ Temporal smoothing
  - ✧ Local stationarity
  - ✧ Sufficient amount of data

# Time-Frequency GC

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

# Time-Frequency GC

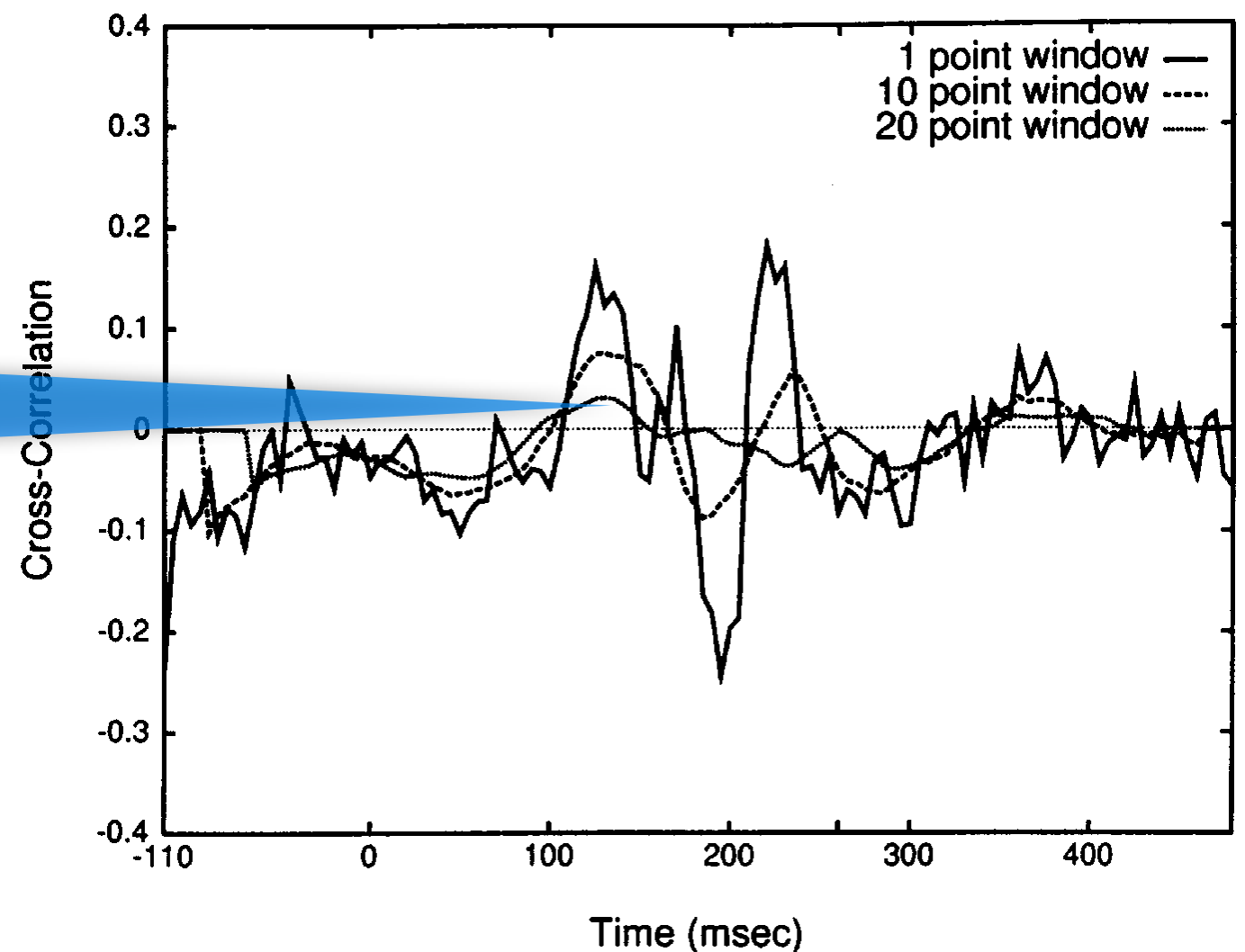
## Consideration: Temporal Smoothness



# Time-Frequency GC

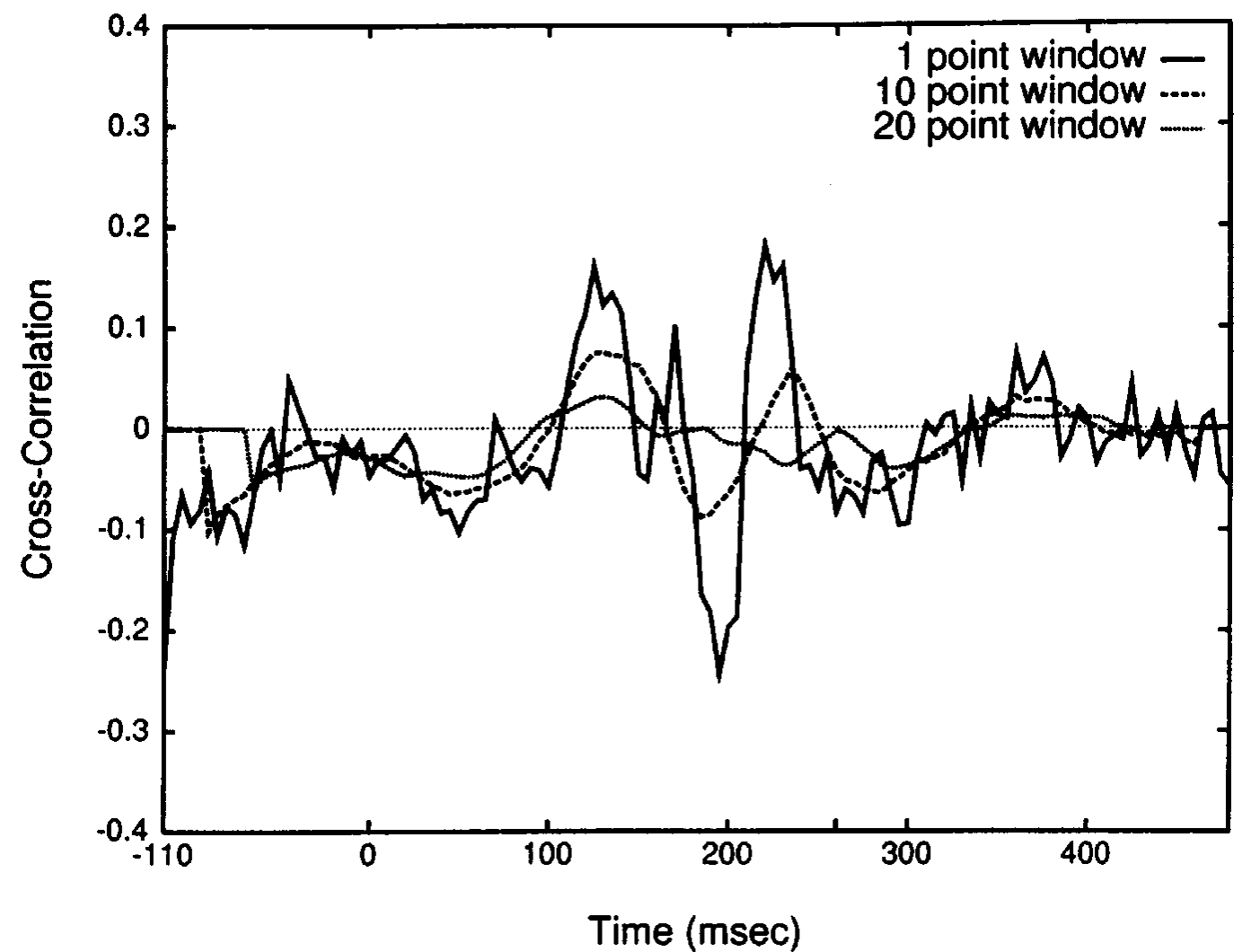
## Consideration: Temporal Smoothness

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



# Time-Frequency GC

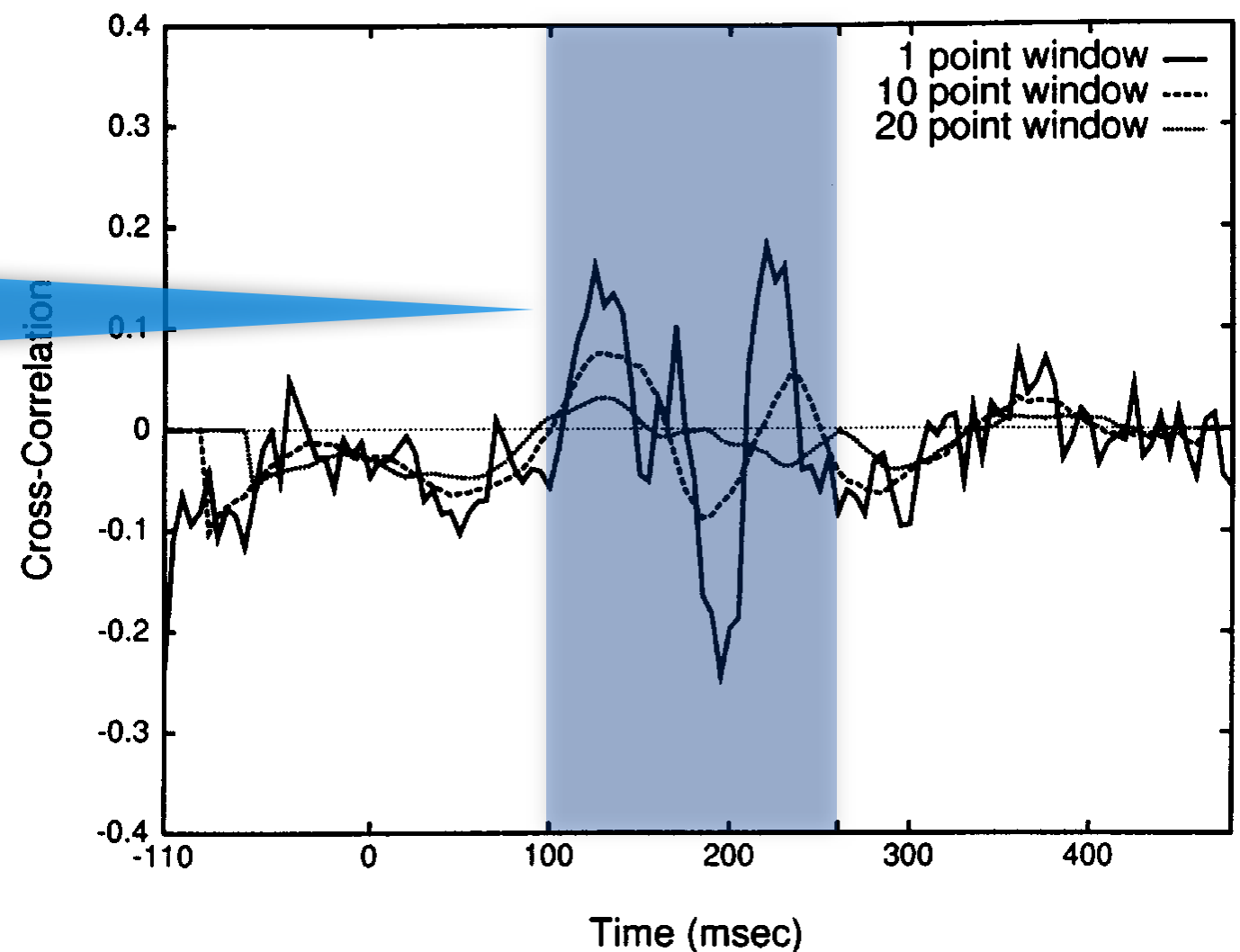
## Consideration: Local Stationarity



# Time-Frequency GC

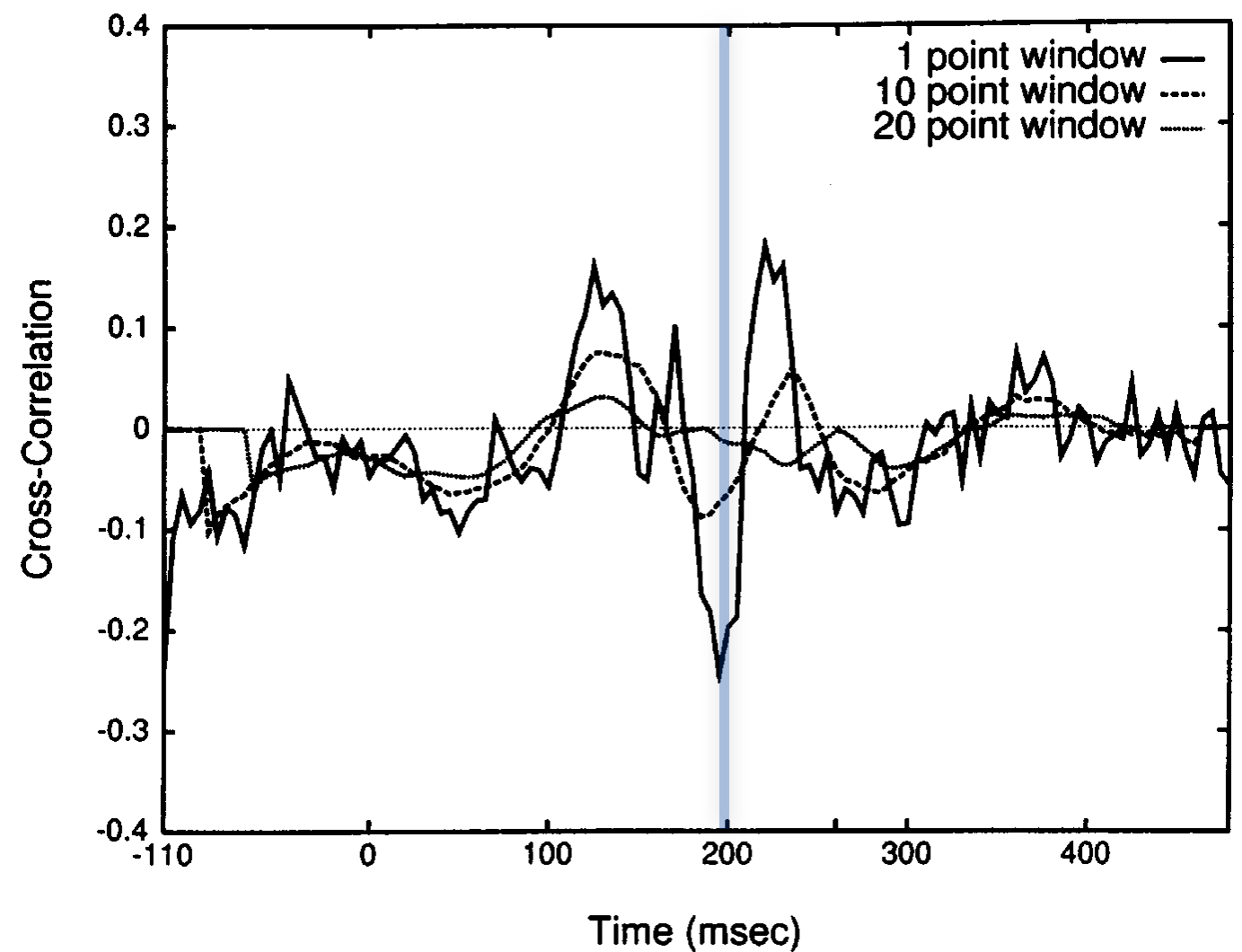
## Consideration: Local Stationarity

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**

$M$  = number of variables

# Time-Frequency GC

## **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

$M$  = number of variables

$p$  = model order

$N_{tr}$  = number of trials

$W$  = length of each window (sample points)

# Time-Frequency GC

## Consideration: Sufficient data

$M$  = number of variables

$p$  = model order

$N_{tr}$  = number of trials

$W$  = length of each window (sample points)

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

# 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)

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

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



# 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)

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  $M^2p \leq (1/10)N_{tr} W$ .

# 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)

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

10x more data points than parameters to estimate

# 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)

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

10x more data points than parameters to estimate

SIFT will let you know if your window length is not optimal

# Time-Frequency GC

# Time-Frequency GC

**Consideration: Process dynamics**

# Time-Frequency GC

## **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  $\sim 1$  cycle of the lowest frequency of interest

# 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  $\sim 1$  cycle of the lowest frequency of interest

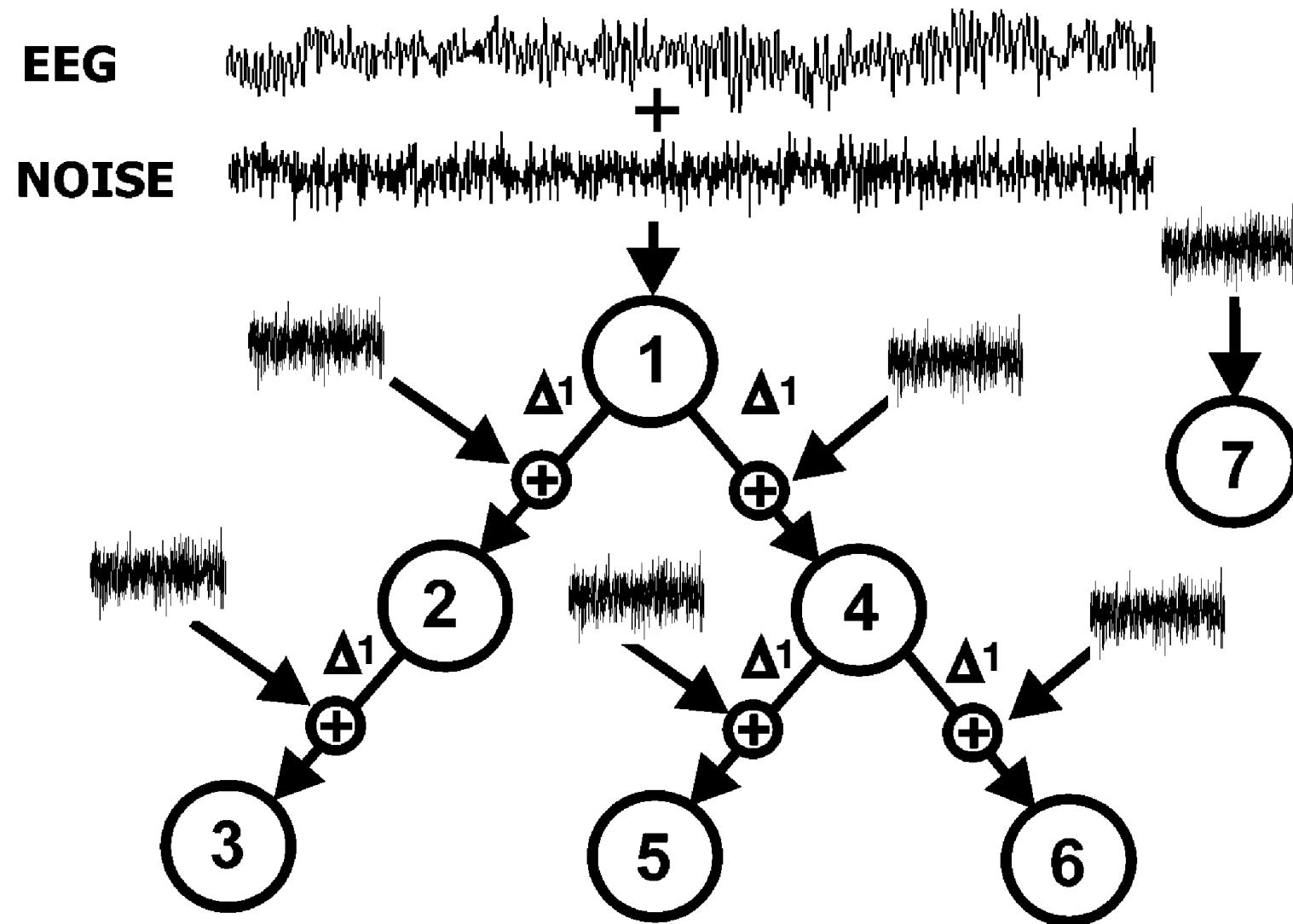


# Time-Frequency GC

## **Consideration: Process dynamics**

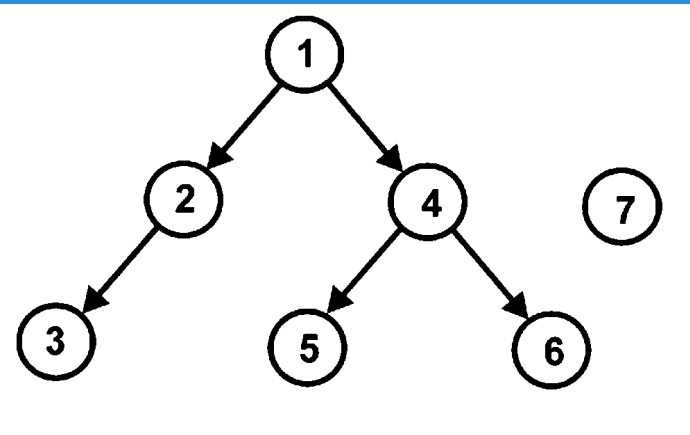
- Your window must be larger than the maximum expected interaction time lag between any two processes.
- Your window should be large enough to span  $\sim 1$  cycle of the lowest frequency of interest

# Which Measure to Use?

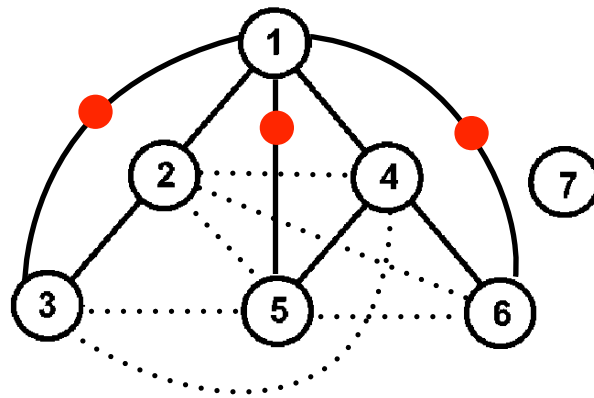


Kus et al, 2004

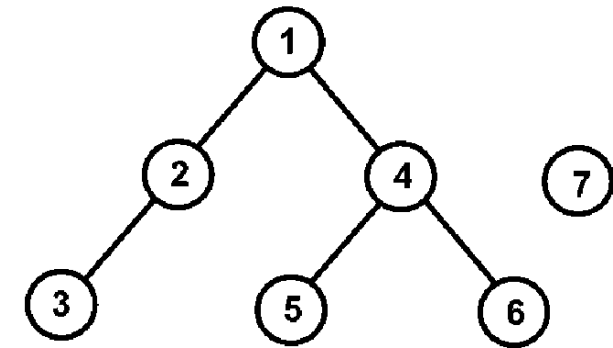
# Ground Truth



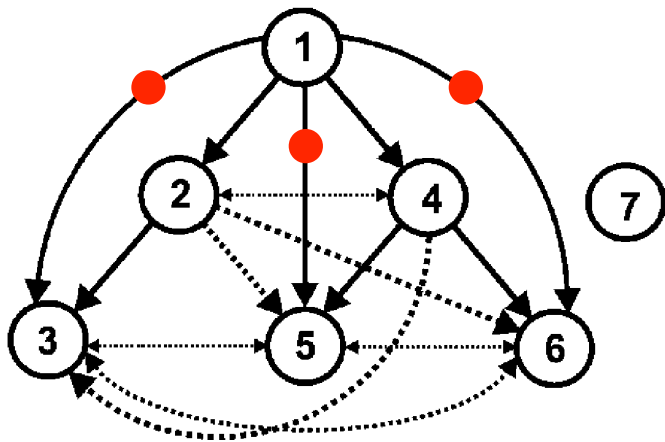
## Coherence



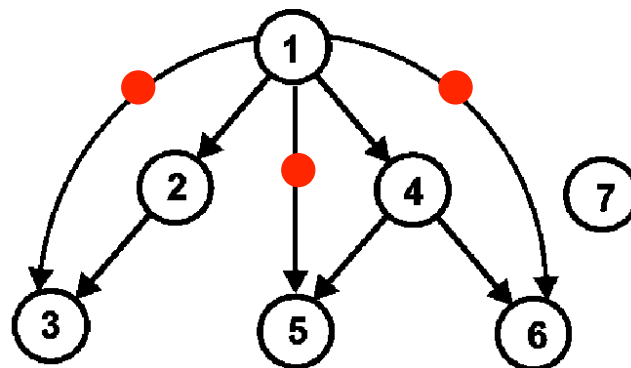
## Partial Coherence



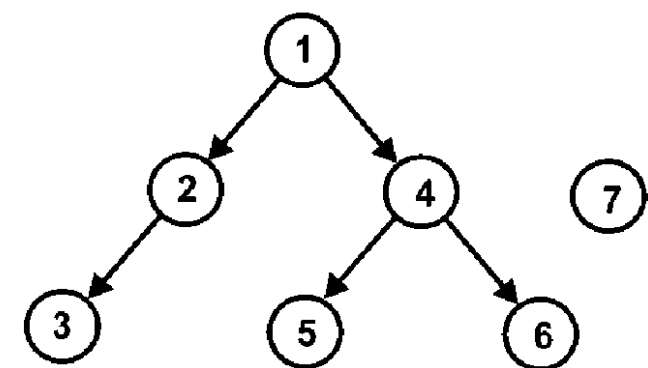
## Bivariate GC



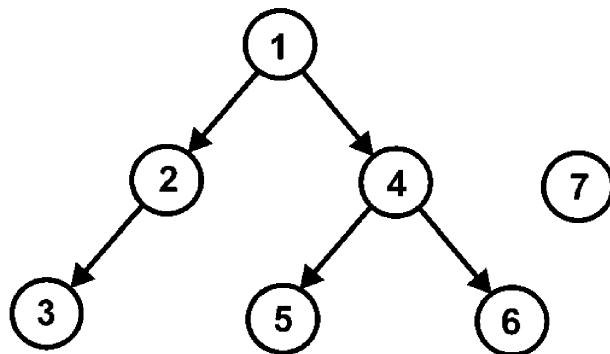
## DTF\*



## dDTF

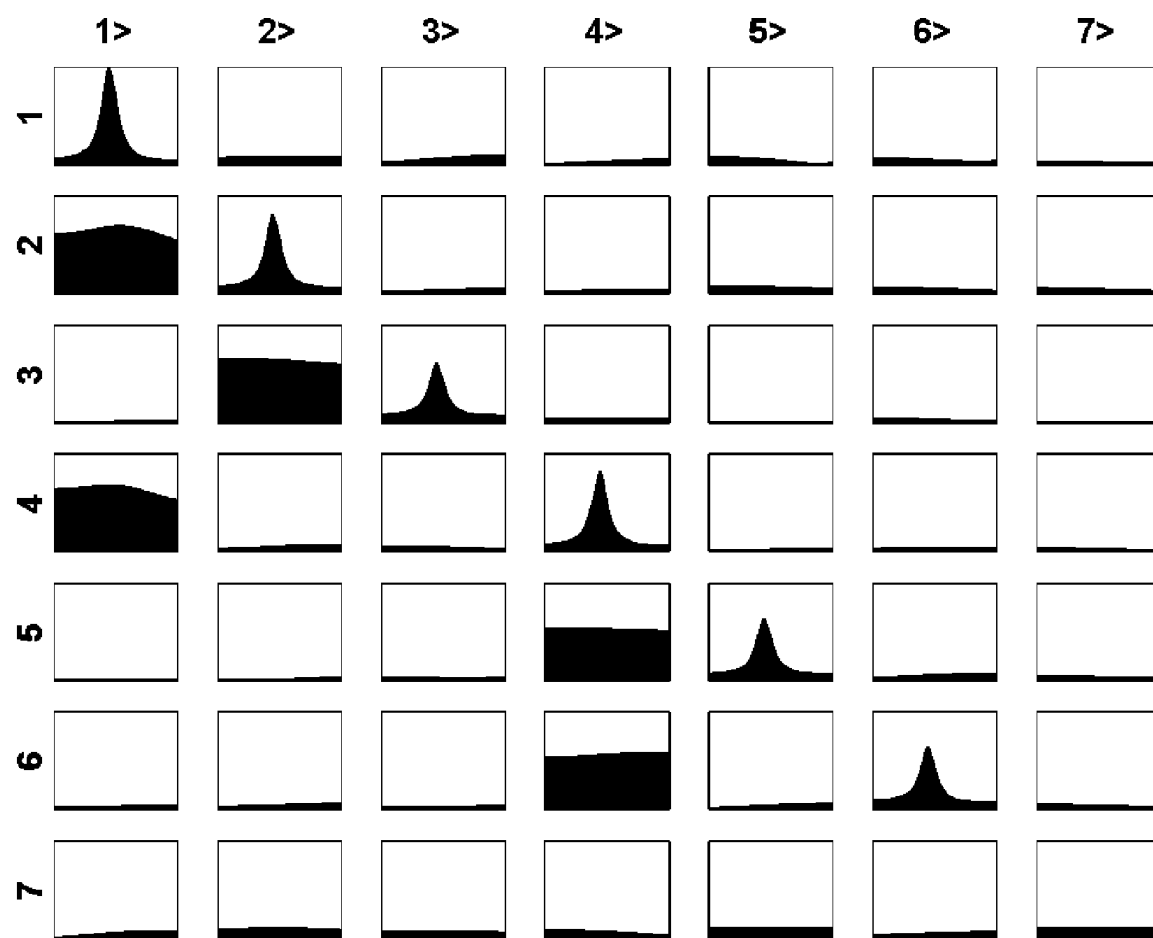


## PDC\*

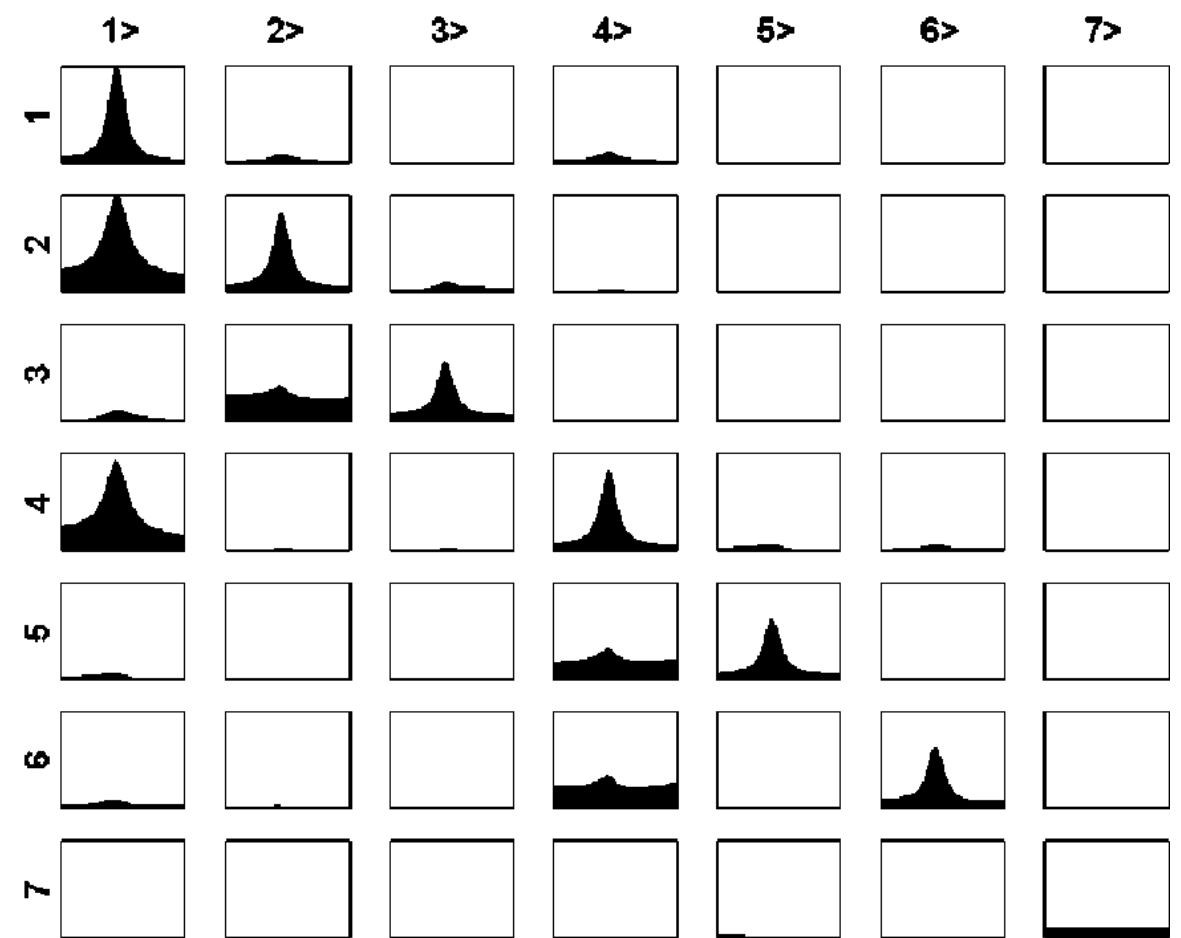


..... spurious  
 —●— indirect true flow  
 — direct true flow  
 \* non-normalized

# PDC versus DTF methods (spectral considerations)

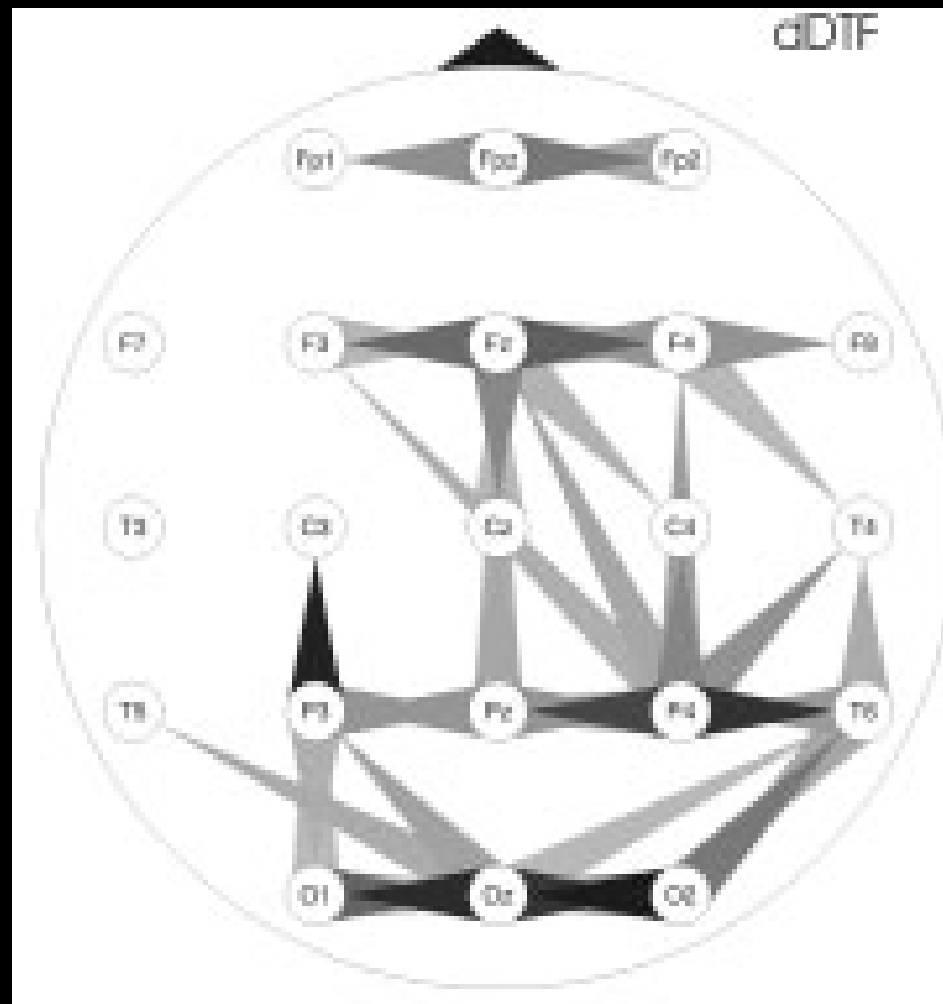


PDC

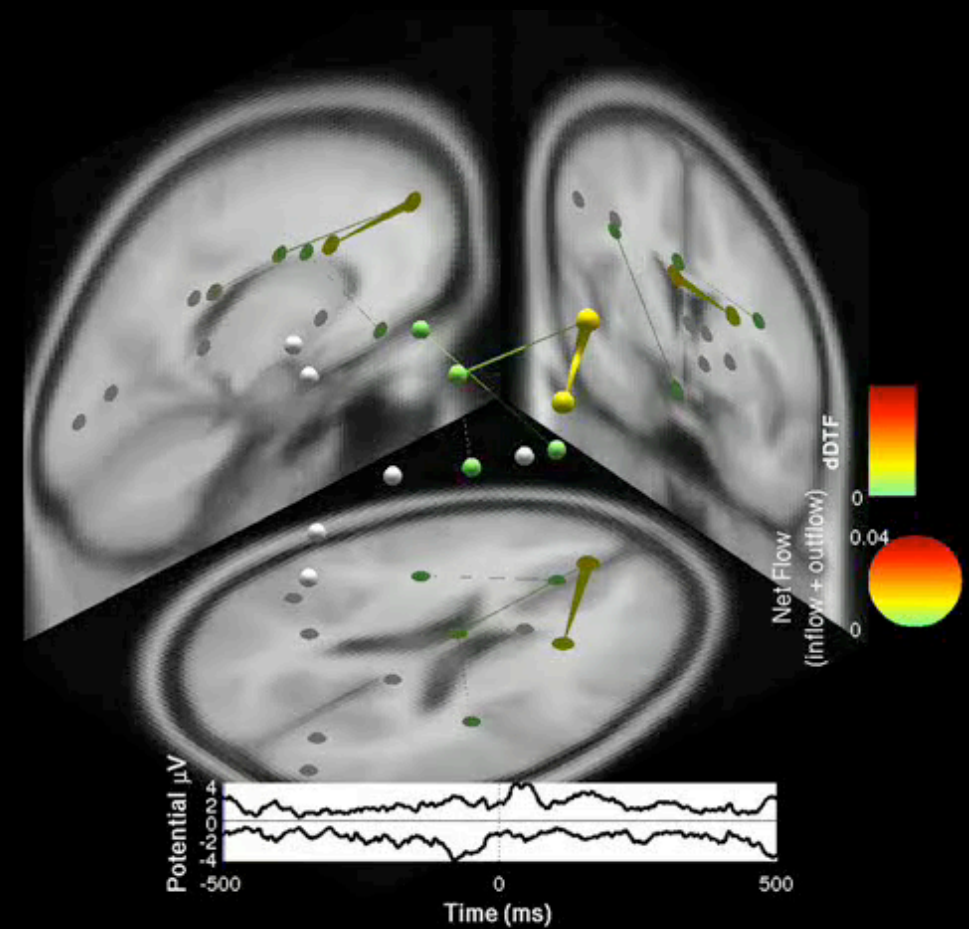


dDTF

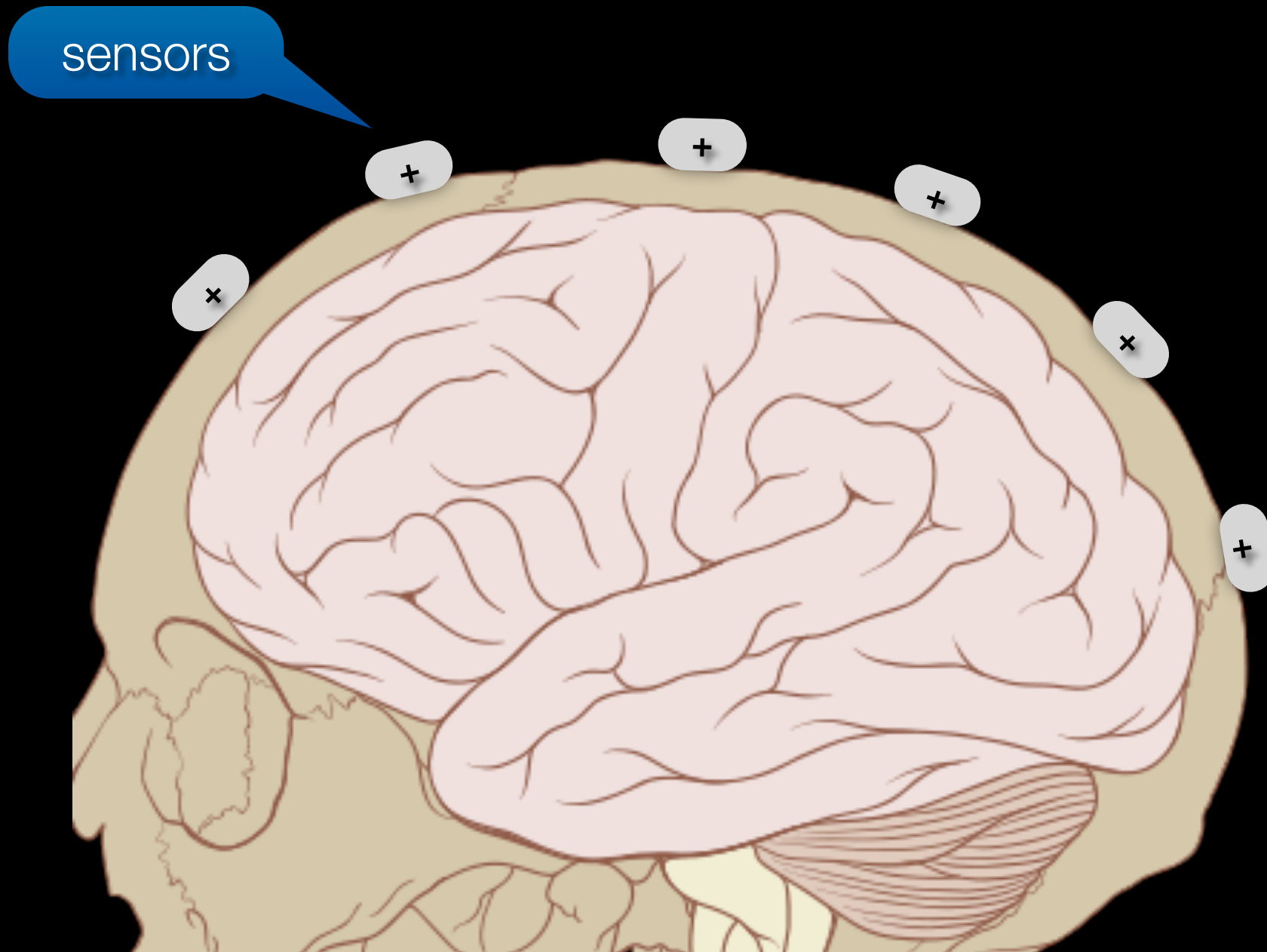
# Scalp or Source?



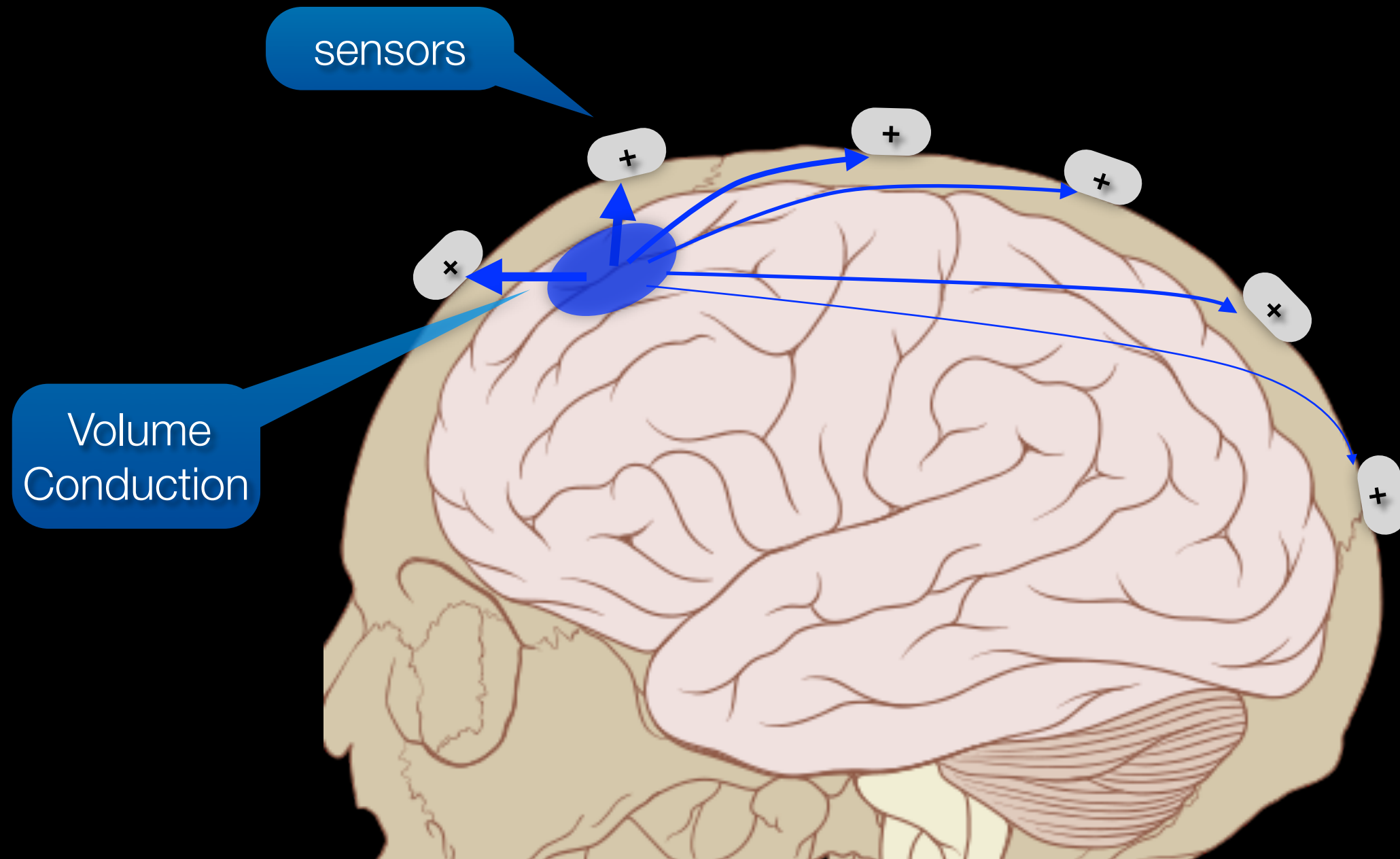
or



# Scalp or Source?

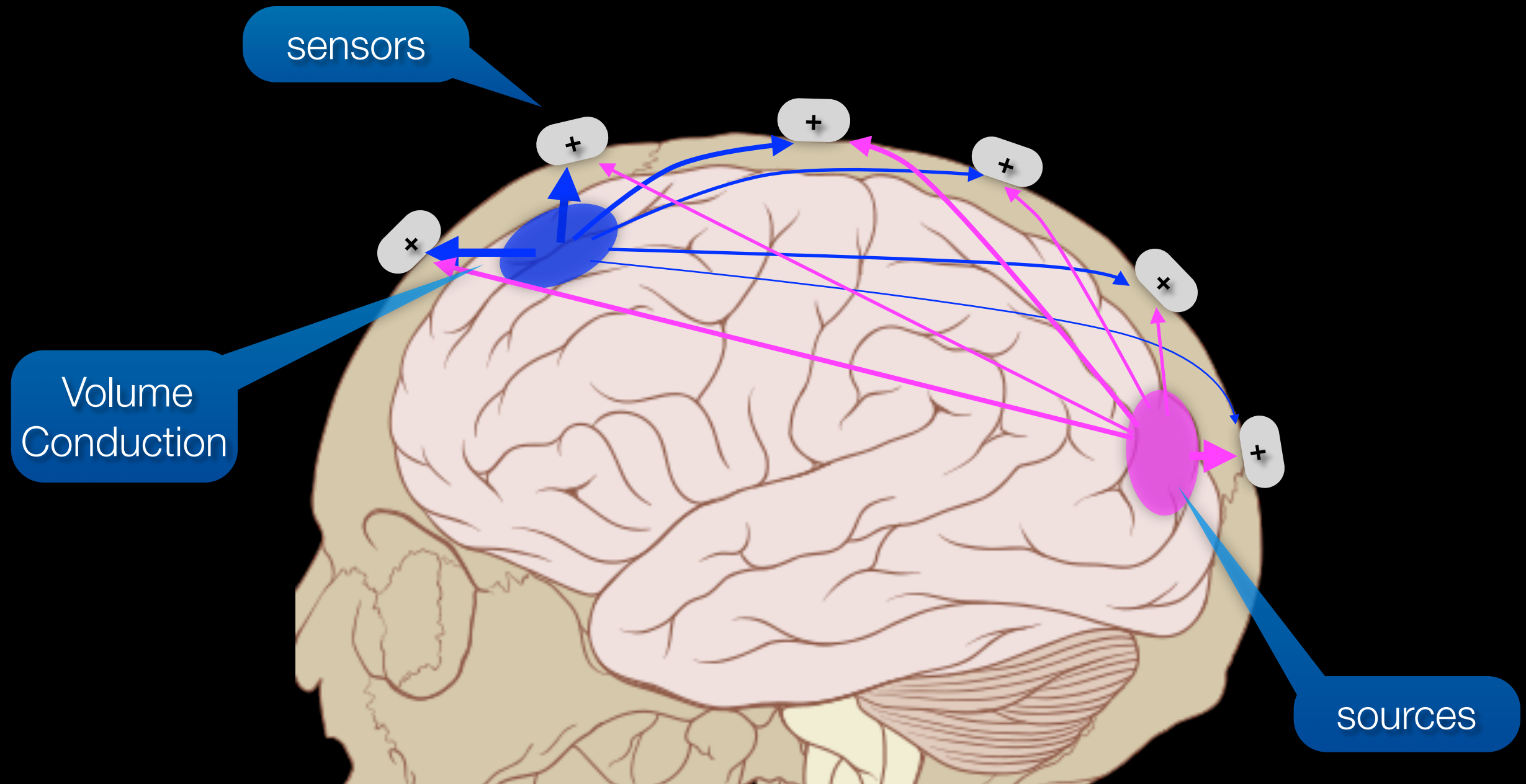


# Scalp or Source?



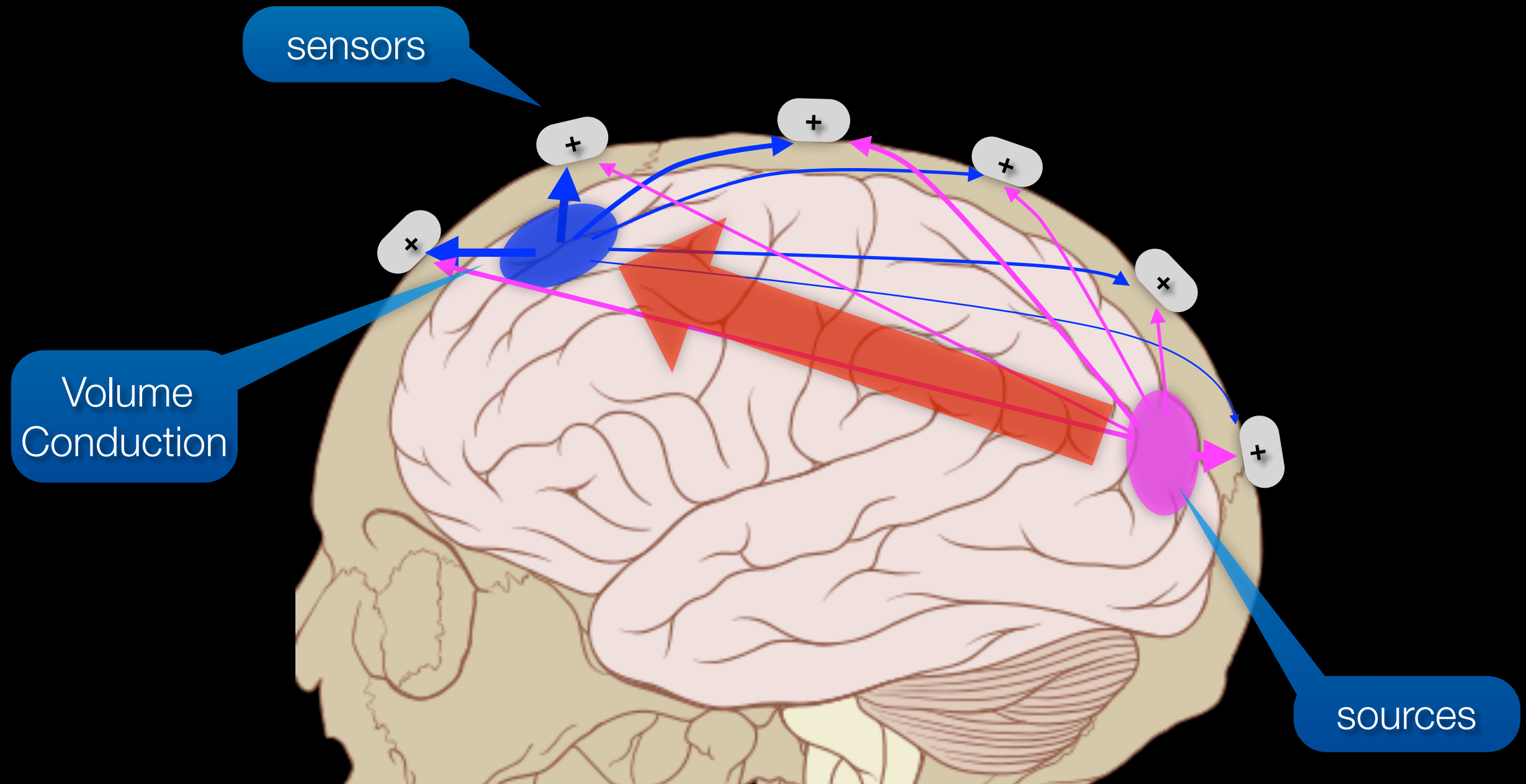


# Scalp or Source?

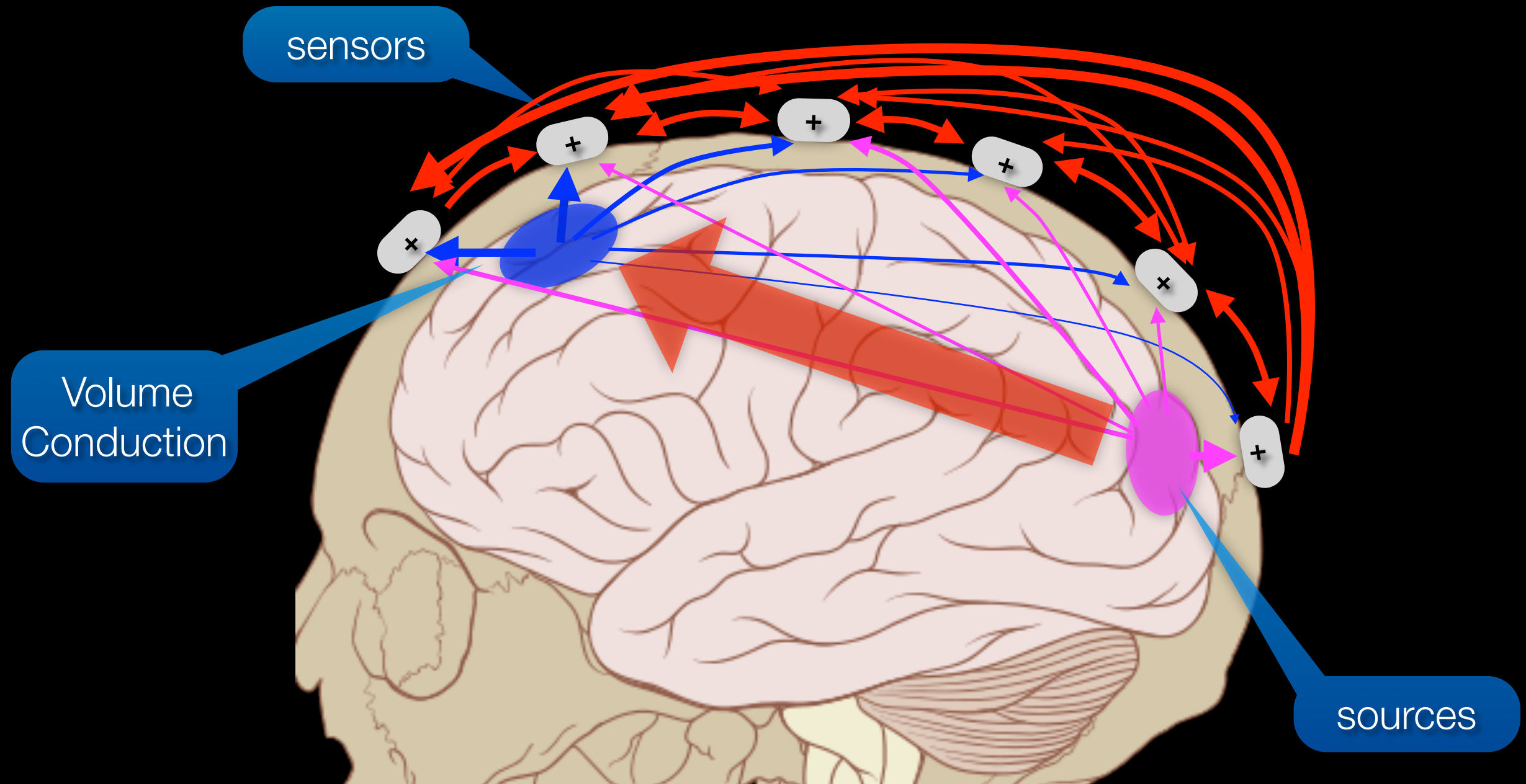




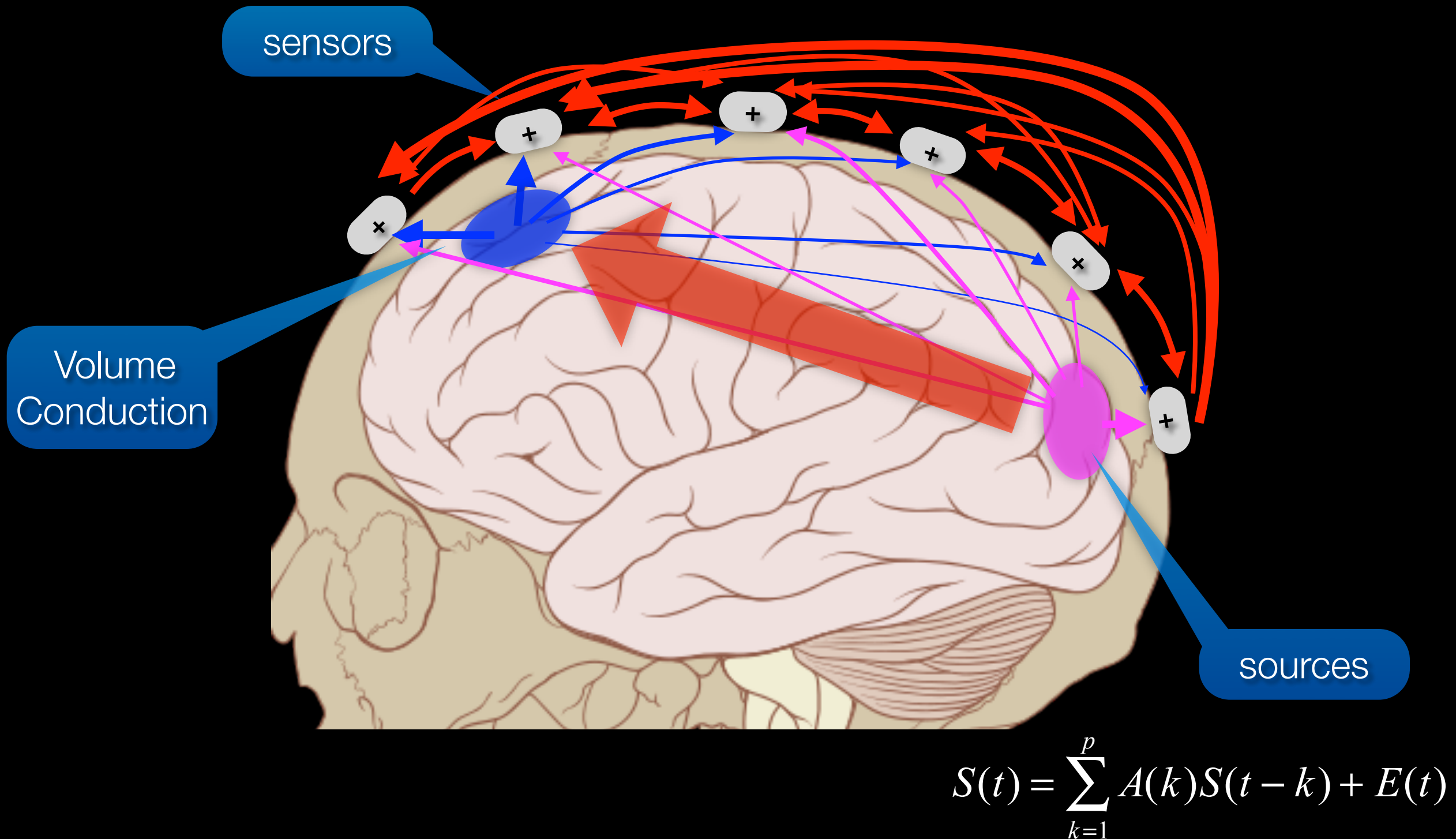
# Scalp or Source?



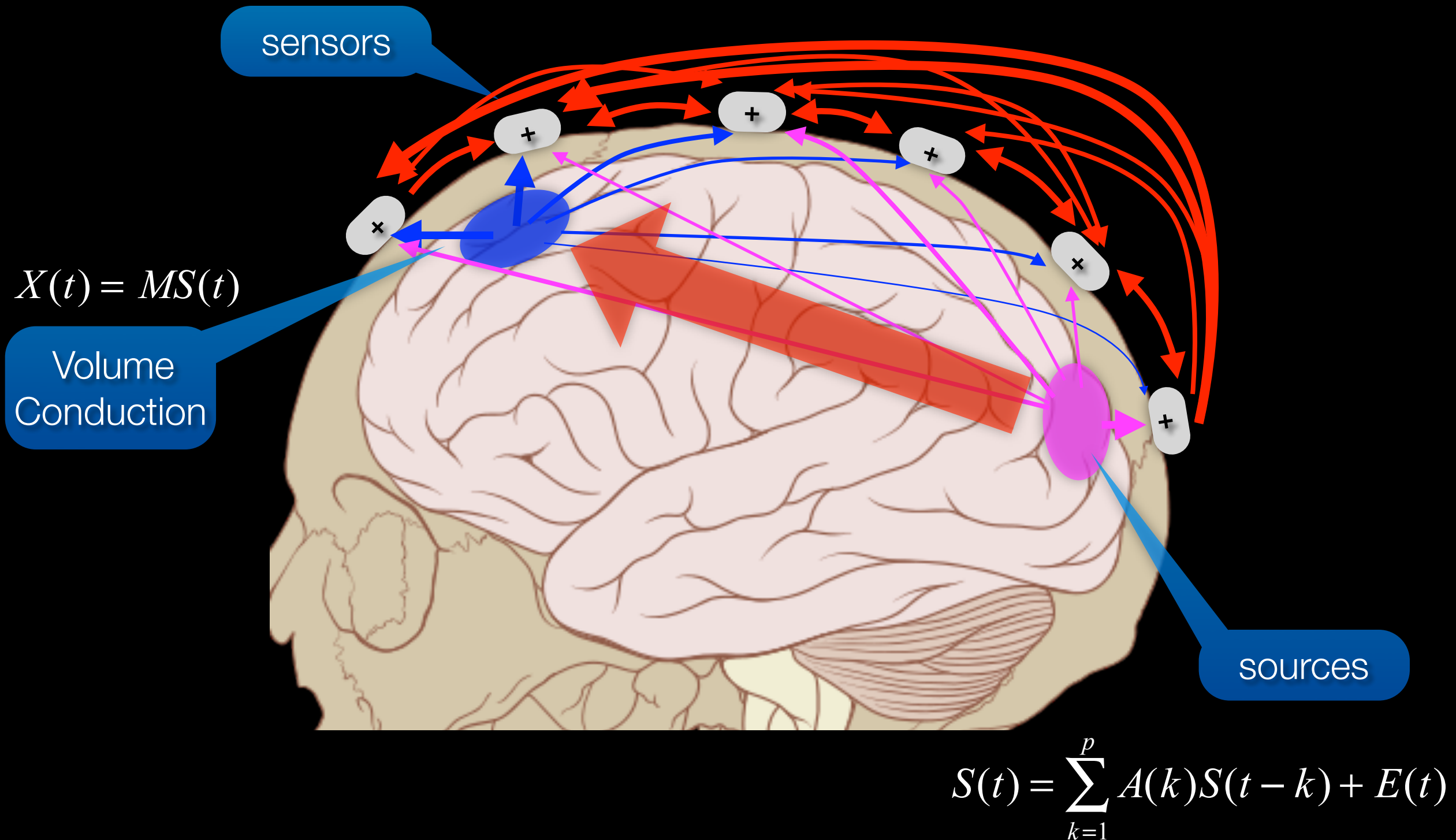
# Scalp or Source?



# Scalp or Source?



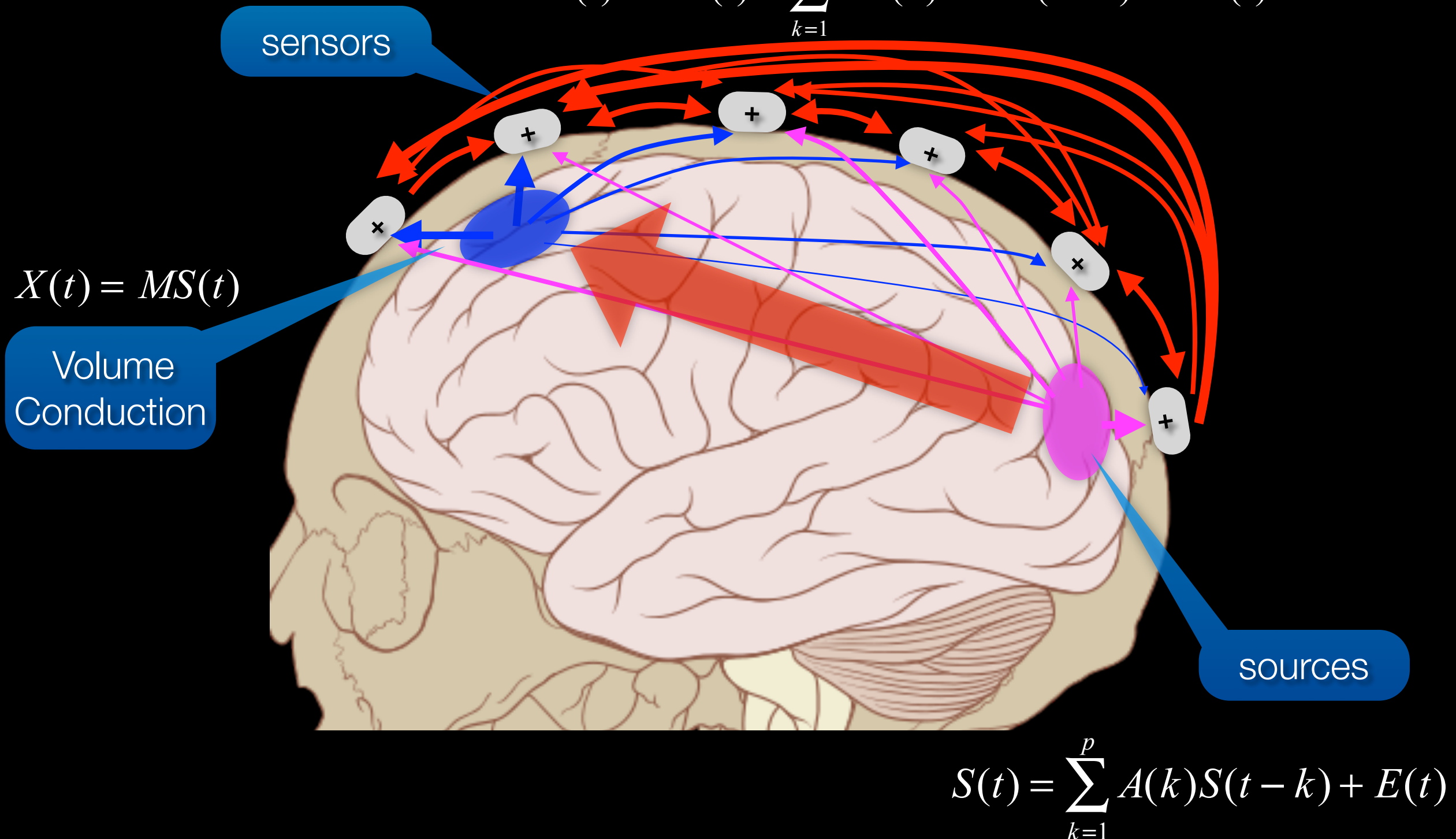
# Scalp or Source?





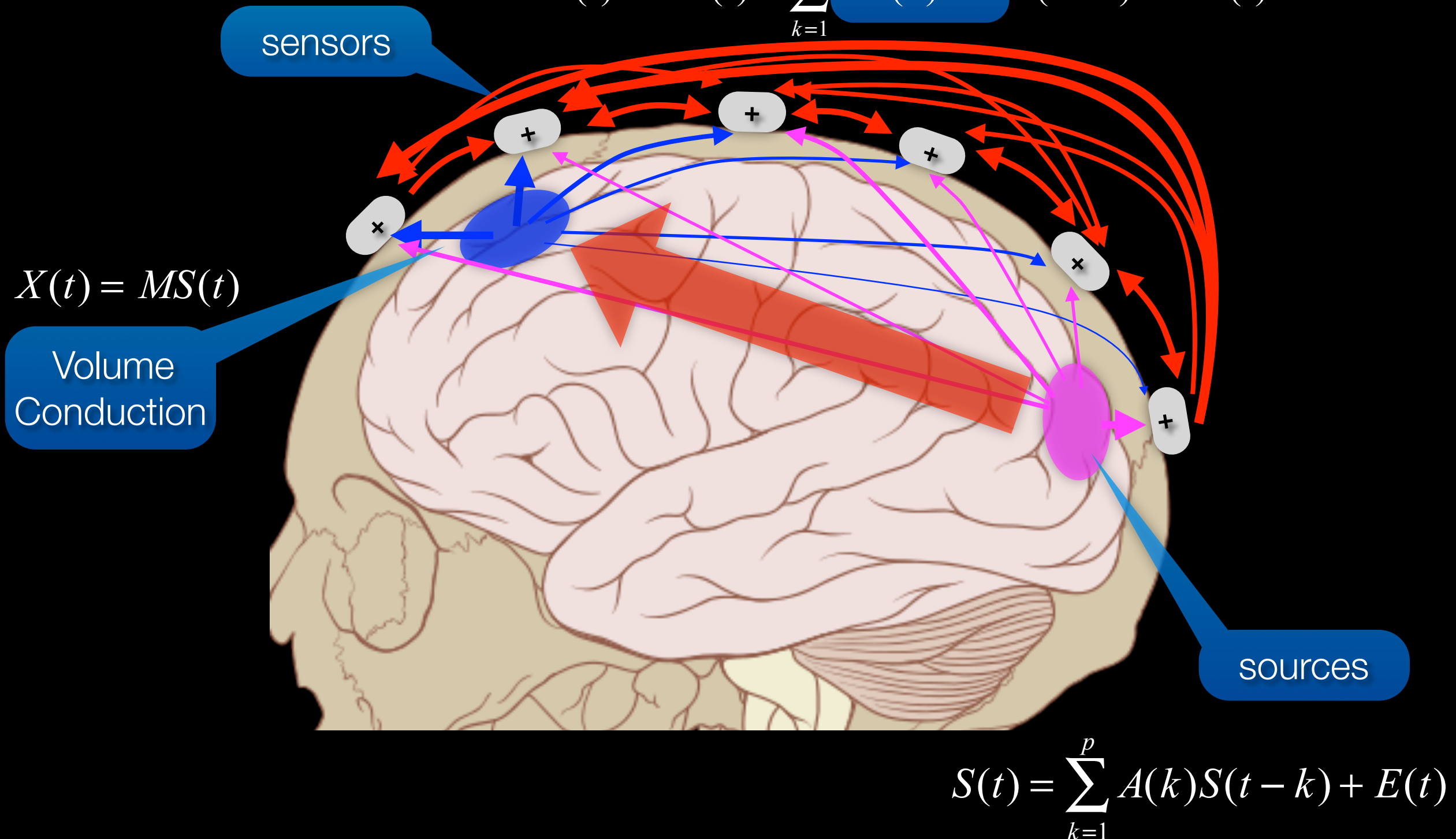
# Scalp or Source?

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



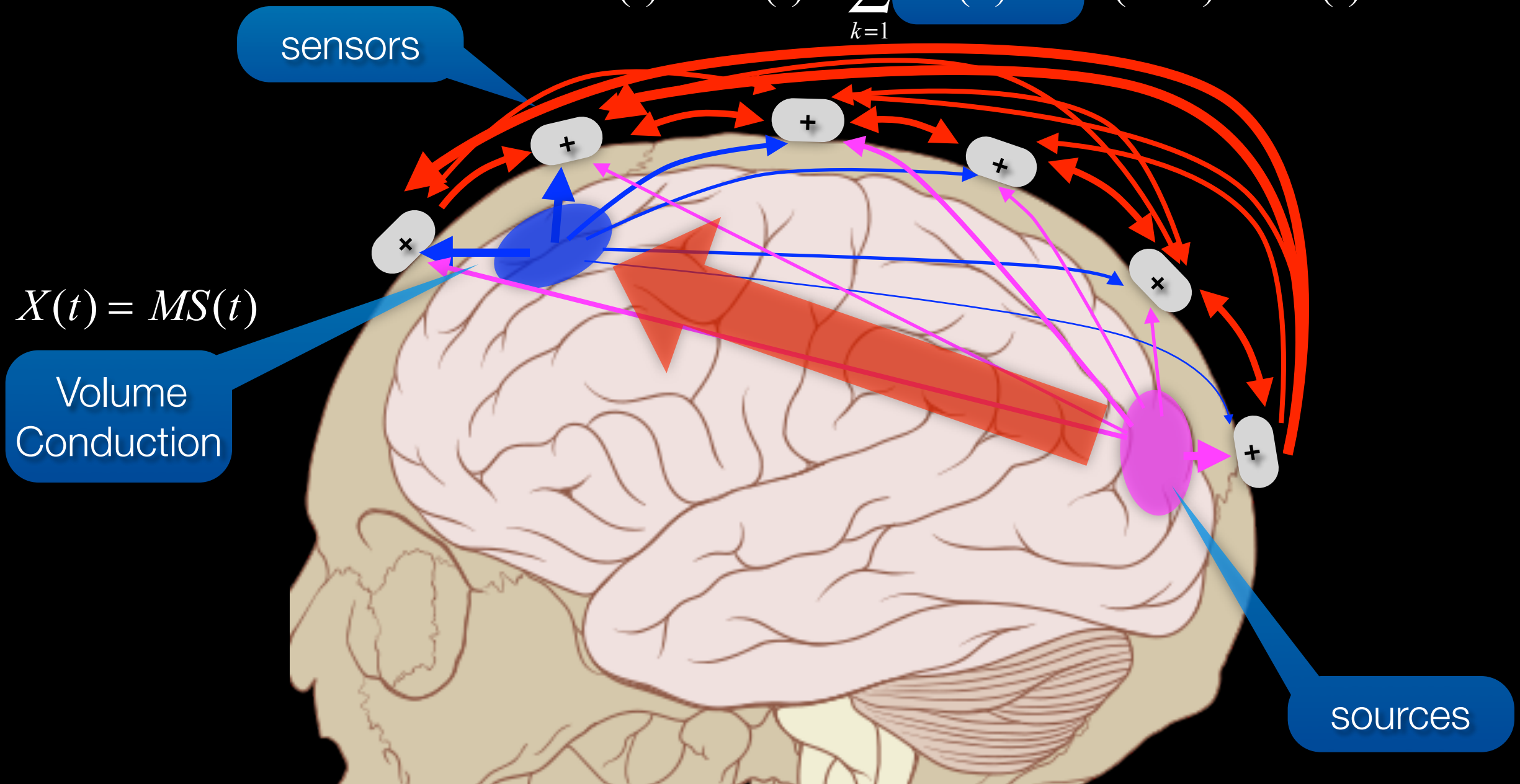
# Scalp or Source?

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



# Scalp or Source?

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



**Solution?** Source Separation

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

# Estimating Dependency of Independent Components ?



# Estimating Dependency of Independent Components ?

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

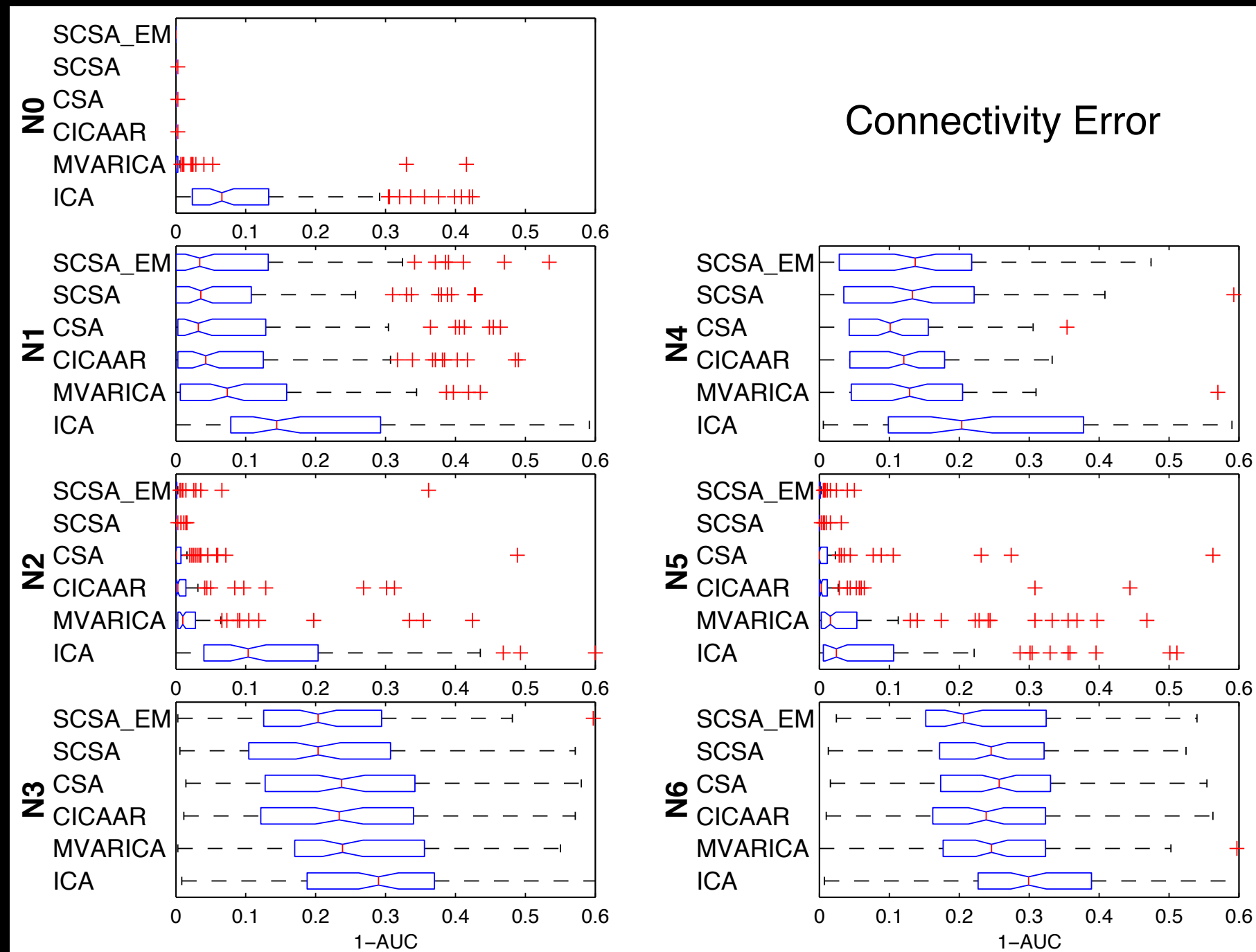
# Estimating Dependency of Independent Components ?

- ✦ 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

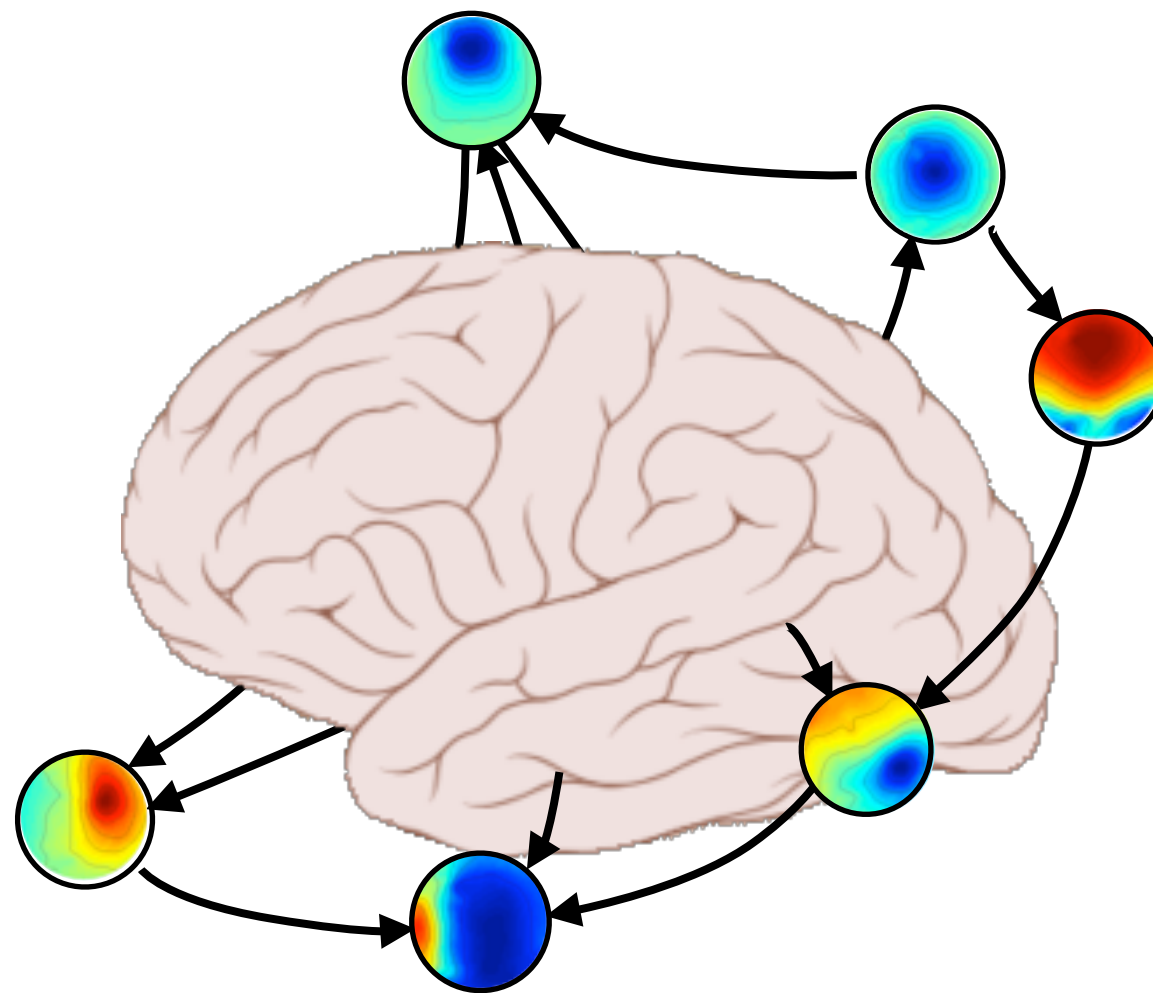
# Estimating Dependency of Independent Components ?

- ✦ 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 ?



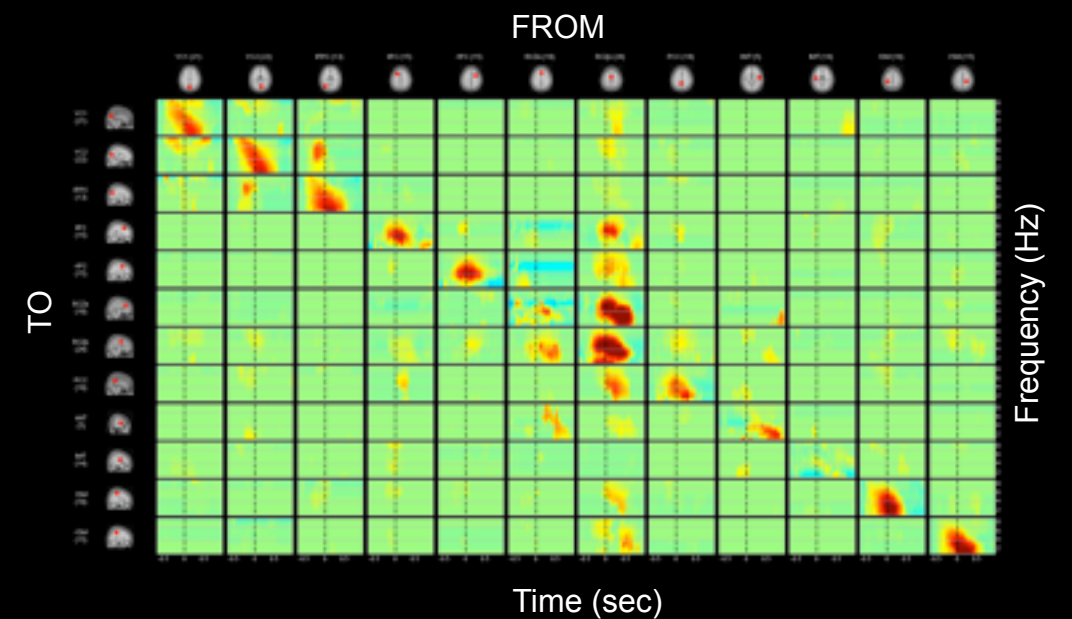
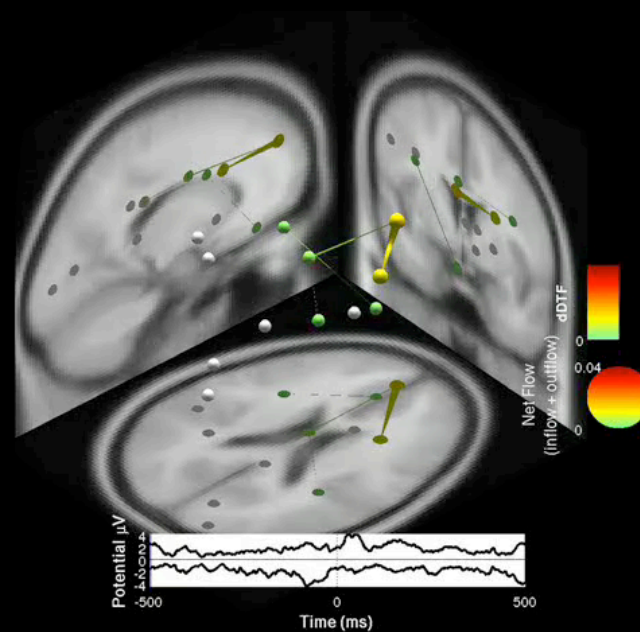
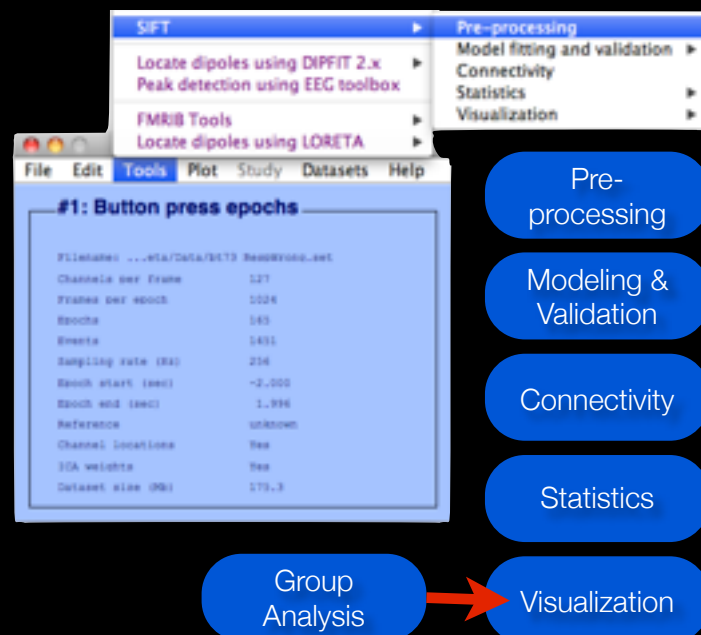
Haufe et al, 2008



# SIFT

Source Information Flow Toolbox

Version 0.1-Alpha



# Source Information Flow Toolbox (SIFT) 0.1-alpha

# Source Information Flow Toolbox (SIFT) 0.1-alpha

- 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) 0.1-alpha

- 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) 0.1-alpha

- 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 time-frequency domain approaches

# Source Information Flow Toolbox (SIFT) 0.1-alpha

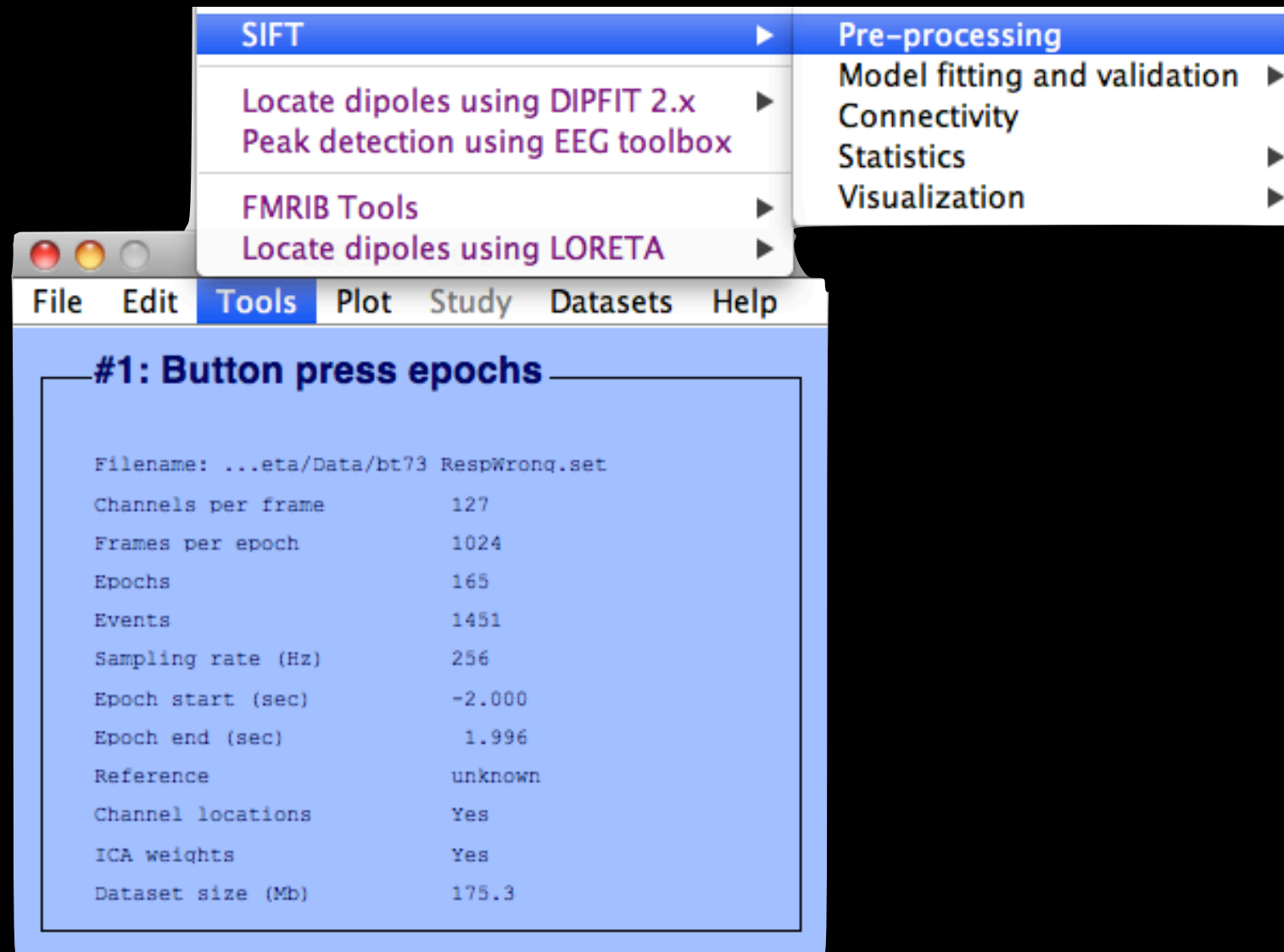
- A new (alpha) toolbox for source-space electrophysiological information flow and causality analysis (single-subject or group analysis) integrated into the EEGLAB software environment
- Modular architecture intended to support multiple modeling approaches
- Emphasis on time-frequency domain approaches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

# Source Information Flow Toolbox (SIFT) 0.1-alpha

- 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 time-frequency domain approaches
- Novel interactive visualization methods for exploratory analysis of connectivity across time, frequency, and spatial location
- Requirements: EEGLAB, MATLAB™ 2008b, Signal Processing Toolbox, Statistics Toolbox (the latter two dependencies may be removed in the future)

# SIFT: Acknowledgements

- ✦ Arnaud Delorme
- ✦ Scott Makeig
- ✦ Christian Kothe
- ✦ Nima Bigdely-Shamlo
- ✦ Wes Thompson
- ✦ SCCN



SIFT ▶

Locate dipoles using DIPFIT 2.x ▶  
Peak detection using EEG toolbox

FMRIB Tools ▶

Locate dipoles using LORETA ▶

Pre-processing

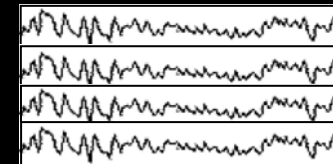
Model fitting and validation ▶  
Connectivity  
Statistics ▶  
Visualization ▶

File Edit **Tools** Plot Study Datasets Help

#1: Button press epochs

Filename: ...eta/Data/bt73 RespWronq.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



SIFT ▶  
Locate dipoles using DIPFIT 2.x ▶  
Peak detection using EEG toolbox ▶  
FMRIB Tools ▶  
Locate dipoles using LORETA ▶

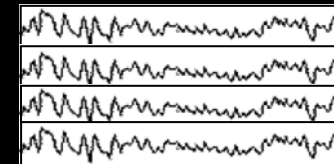
Pre-processing ▶  
Model fitting and validation ▶  
Connectivity ▶  
Statistics ▶  
Visualization ▶

File Edit **Tools** Plot Study Datasets Help

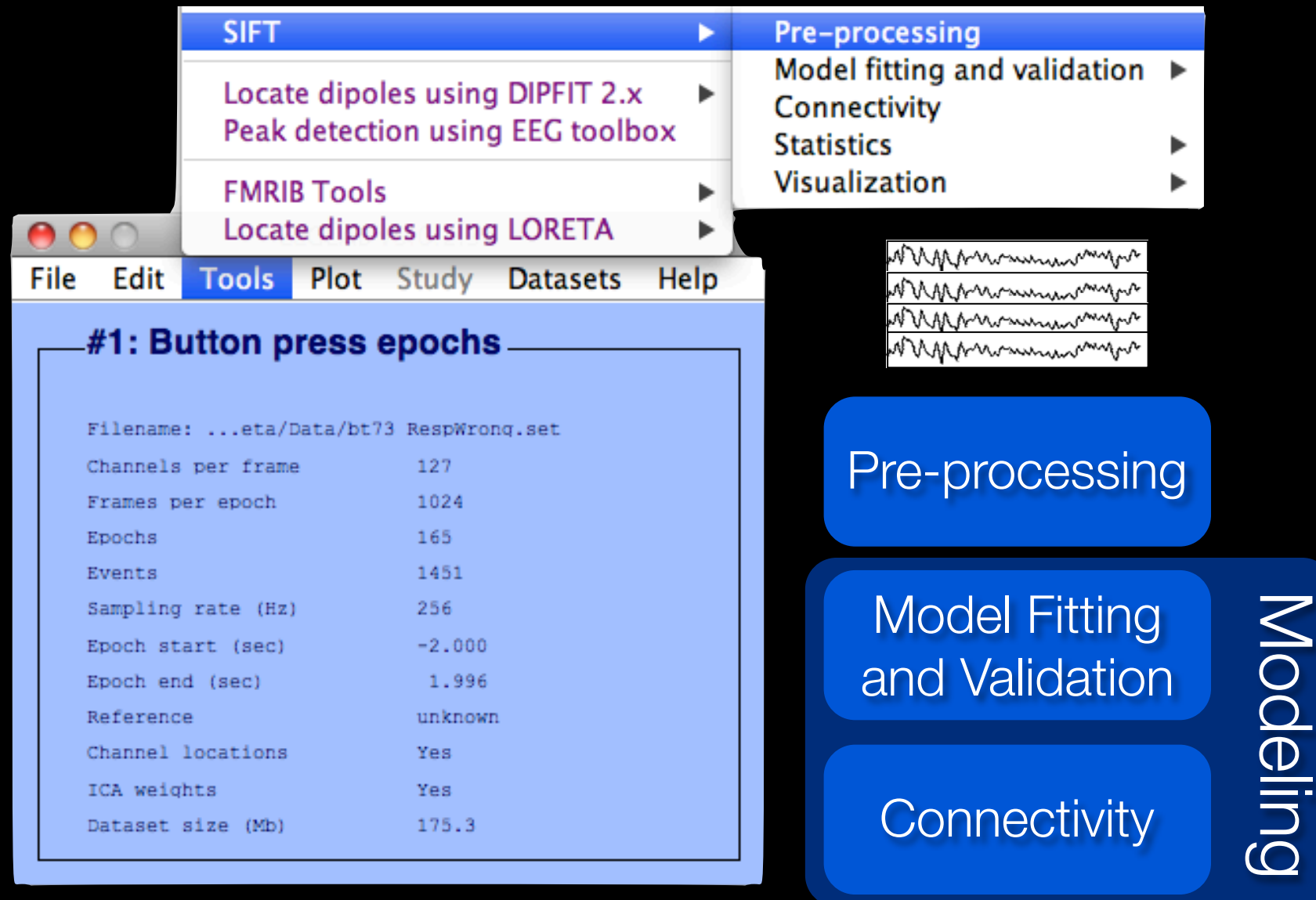
#1: Button press epochs

Filename: ...eta/Data/bt73 RespWronq.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

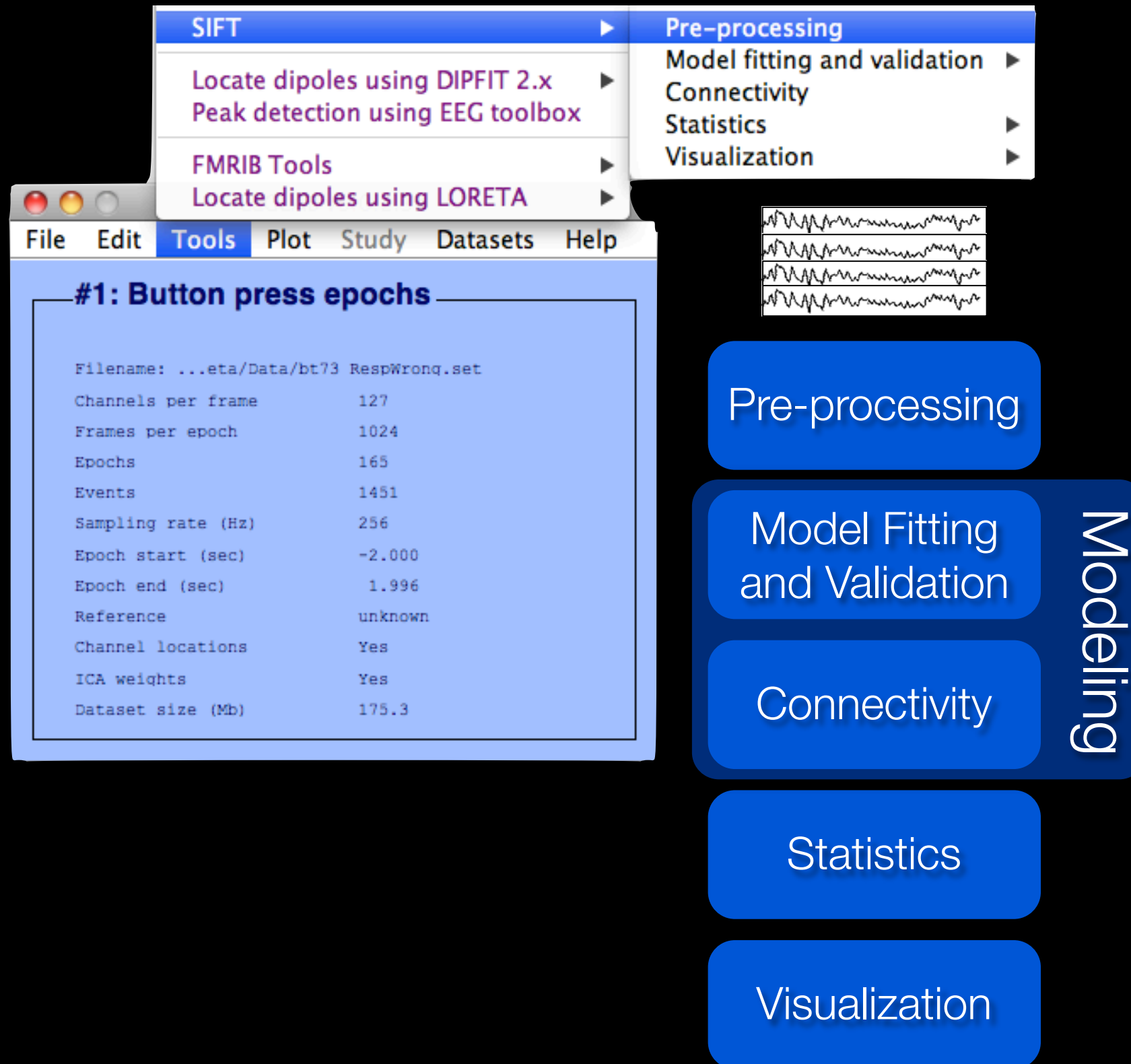


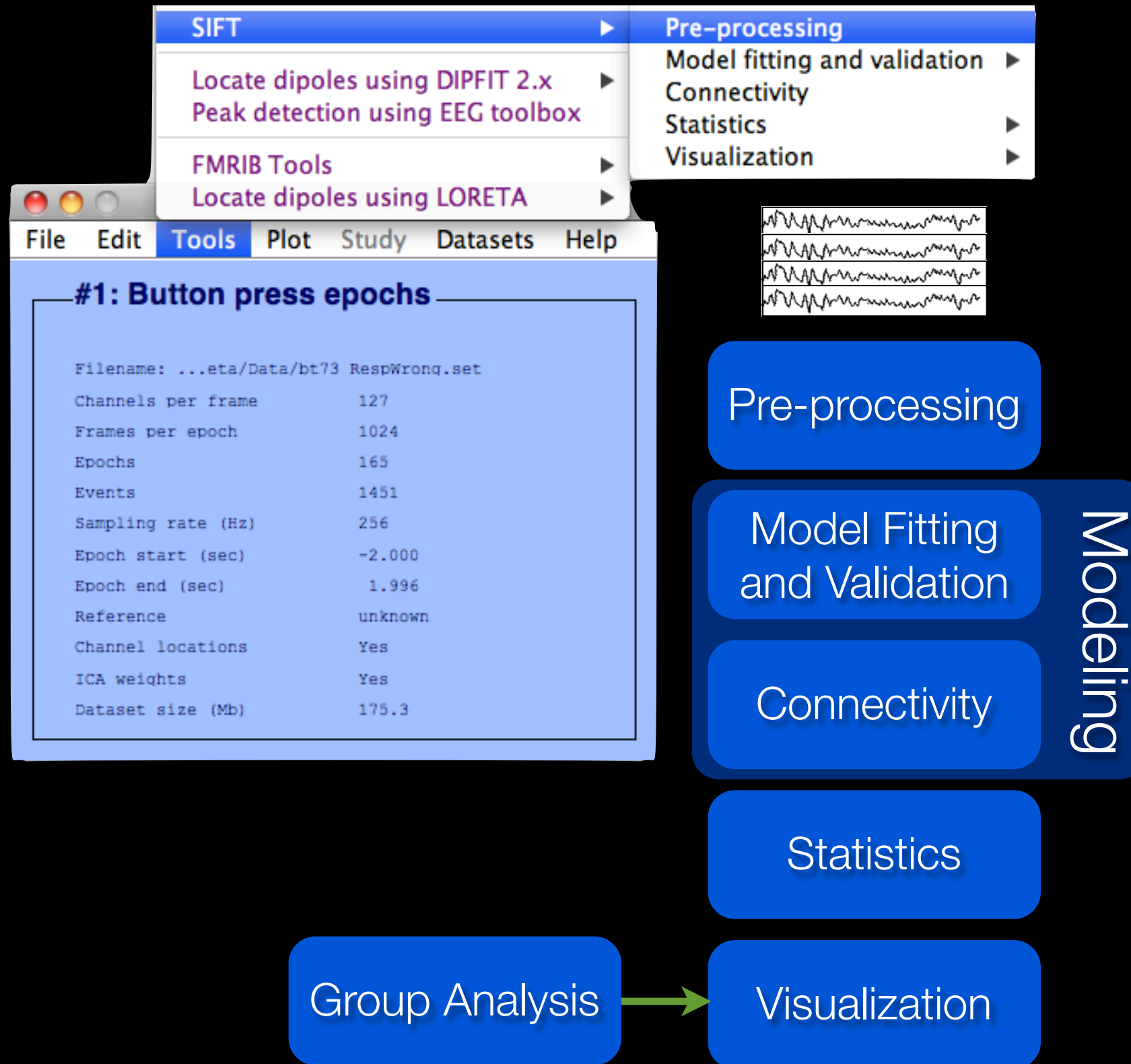
Pre-processing

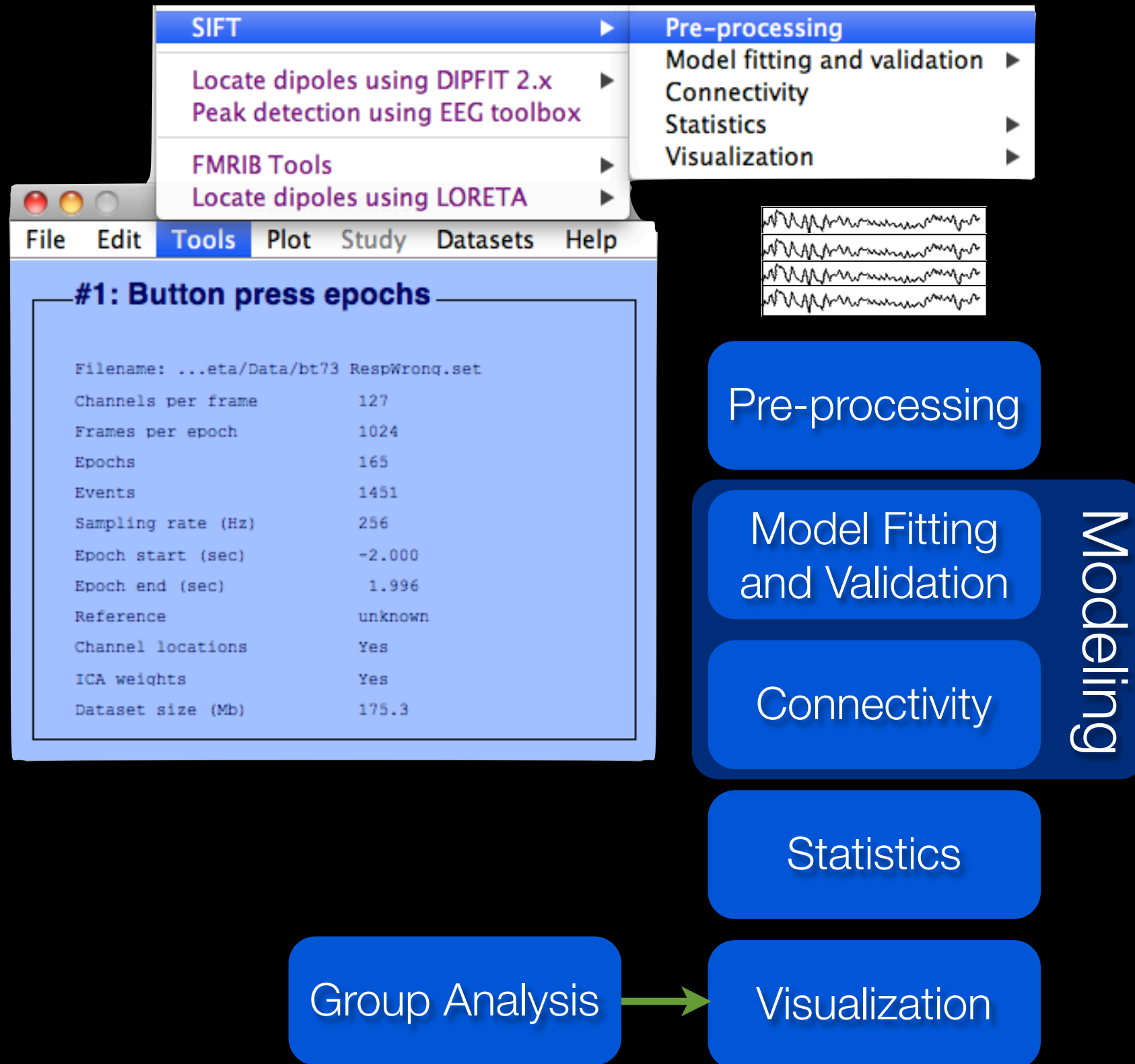












Preprocessing

Modeling

Statistics

Visualization

- ✦ **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)

# Preprocessing

# Modeling

# Statistics

# Visualization

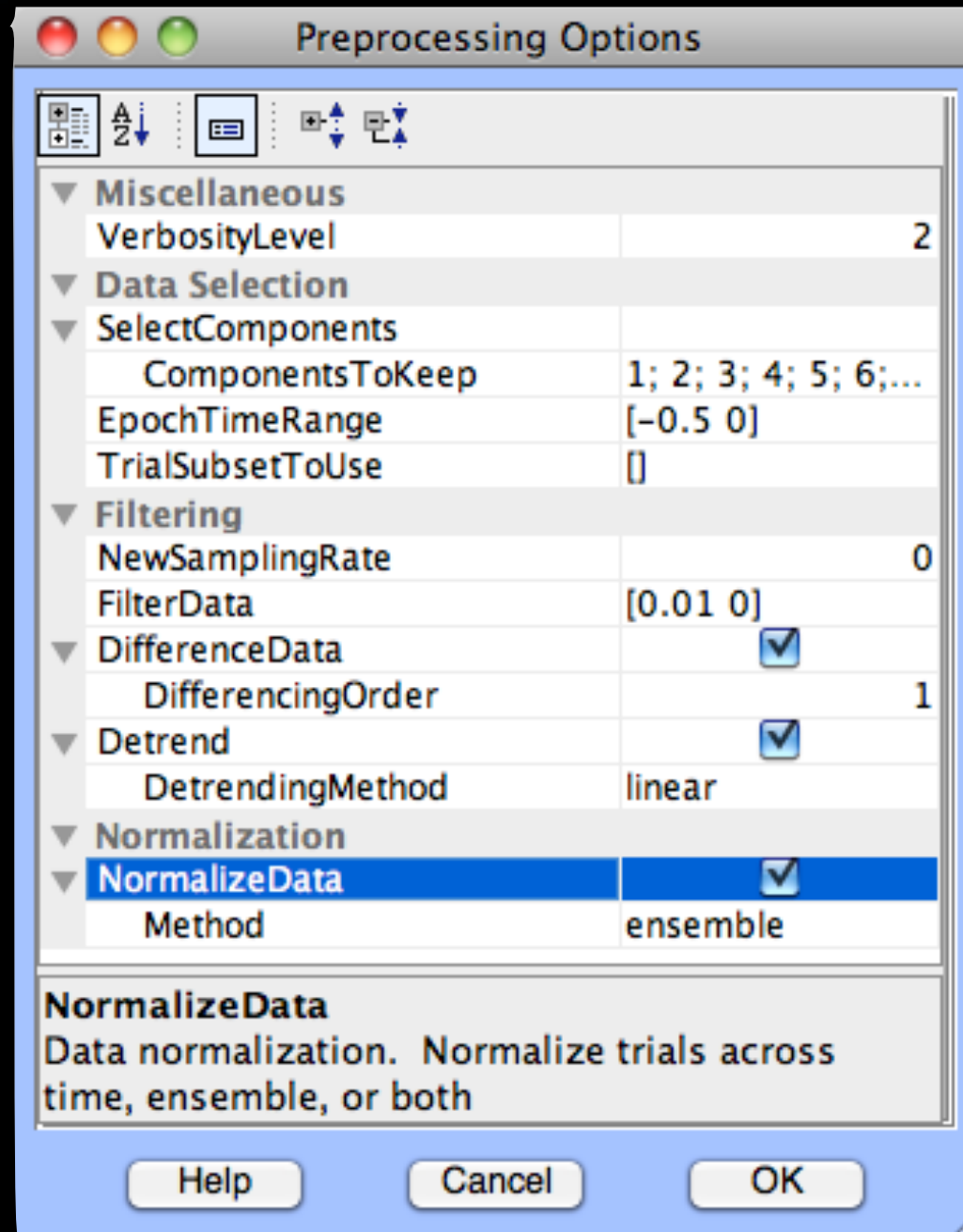
Pre-processing

Model fitting and validation ▶

Connectivity

Statistics ▶

Visualization ▶



Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity



Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

	Linear	Nonlinear/Nonstationary
Parametric	<div>MVAR Modeling</div> <div>Sparse MVAR</div> <div>Bayesian MVAR</div> <div>Kalman Filtering</div>	Dual Extended Kalman Filtering
Nonparametric	<div>Nonparametric MVAR (minimum-phase spectral factorization)</div> <div>Multivariate phase distribution</div>	Transfer Entropy



Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

Pre-processing

Model fitting and validation ▶

Connectivity

Statistics ▶

Visualization ▶

Fit AMVAR Model

Validate model

Preprocessing

Modeling

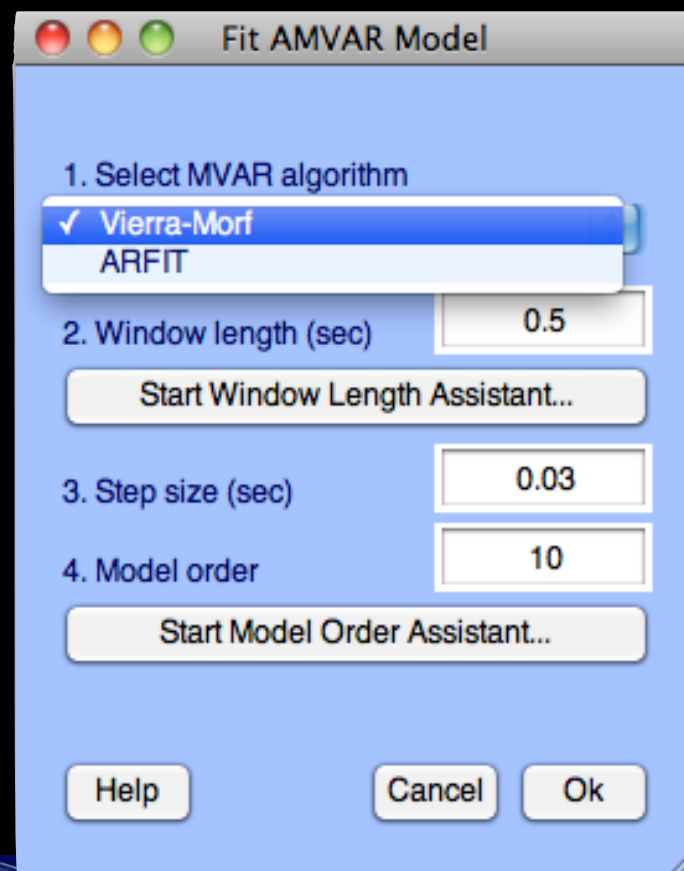
Statistics

Visualization

Model Fitting

Validation

Connectivity



Preprocessing

Modeling

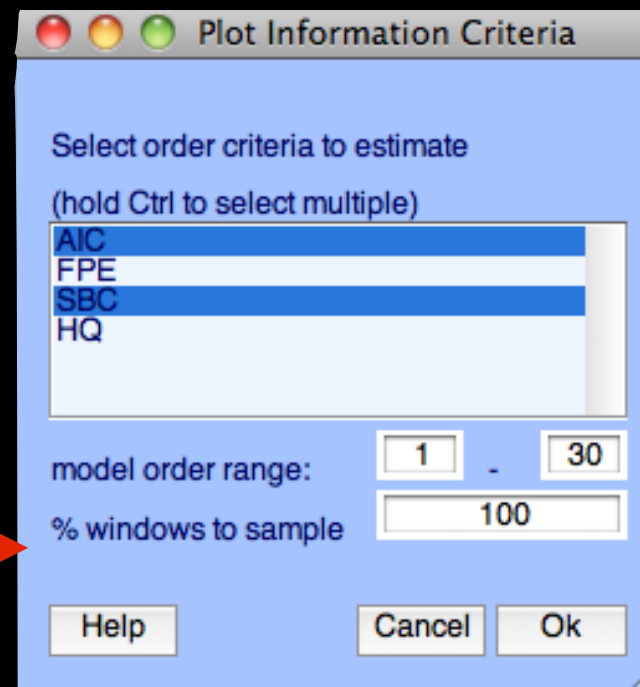
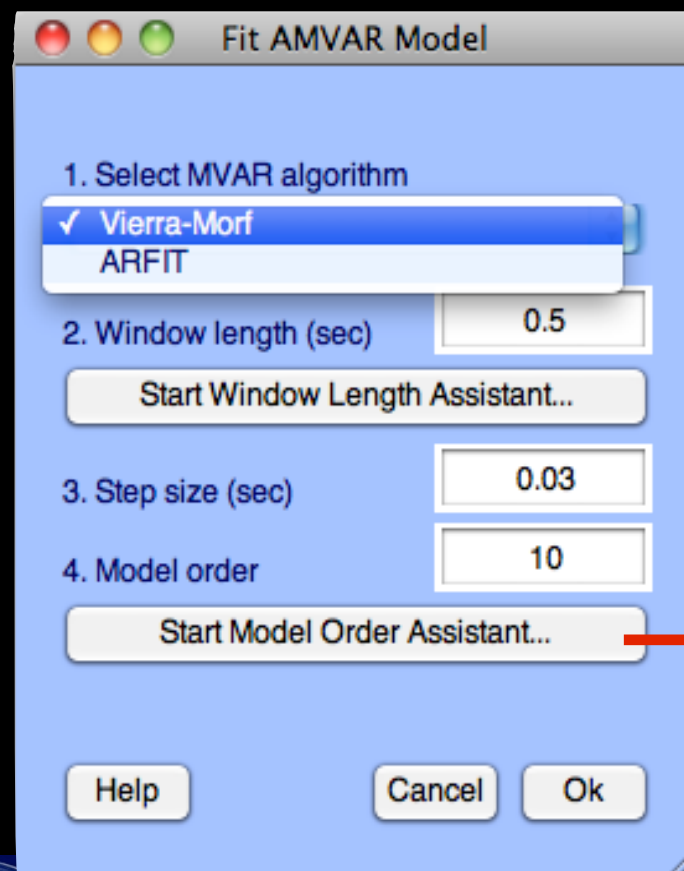
Statistics

Visualization

Model Fitting

Validation

Connectivity



Preprocessing

Modeling

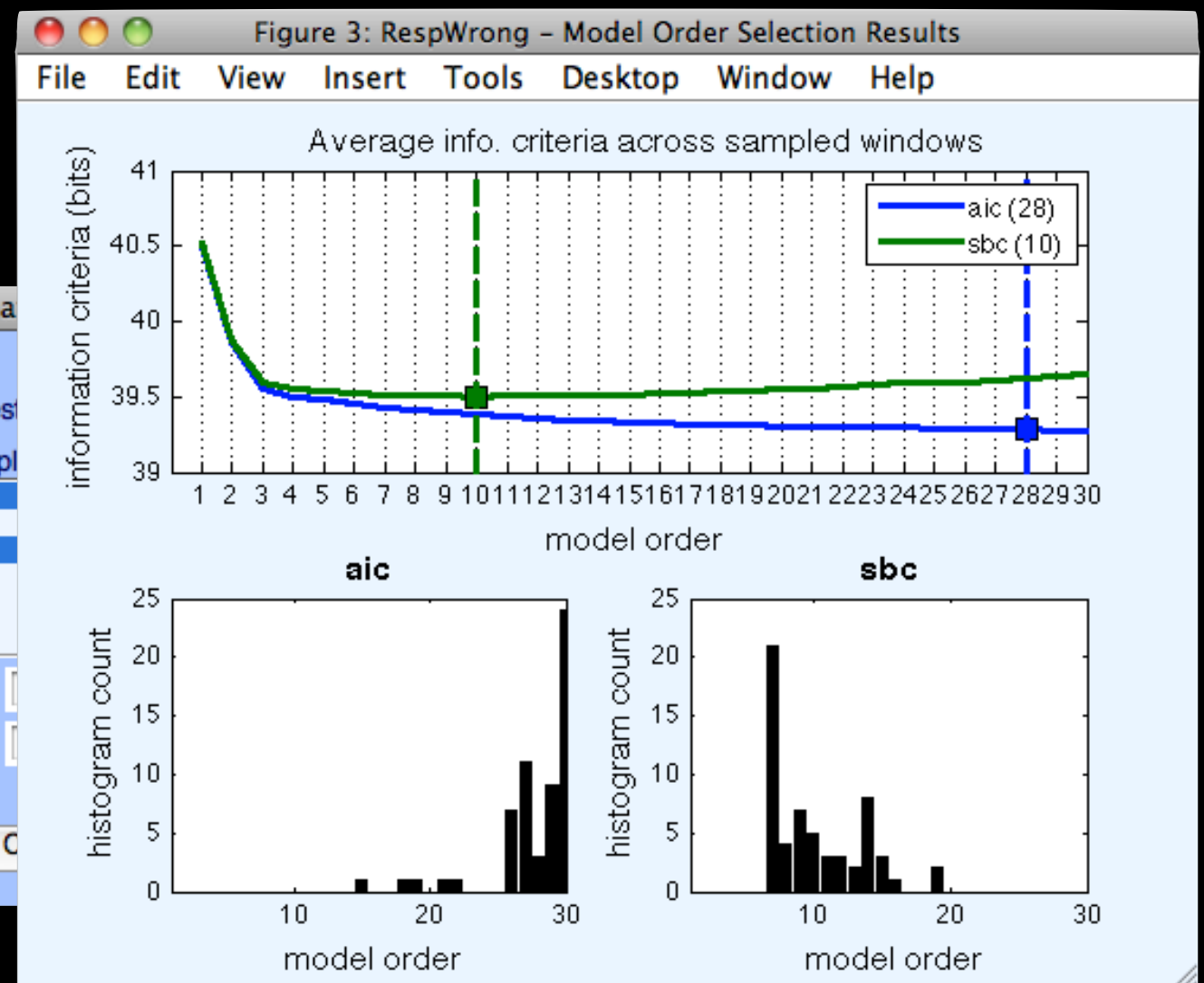
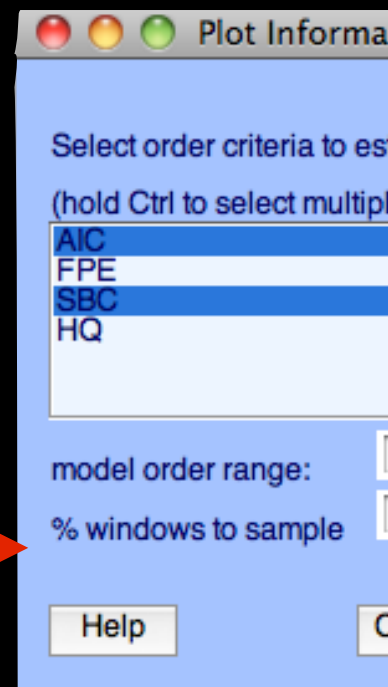
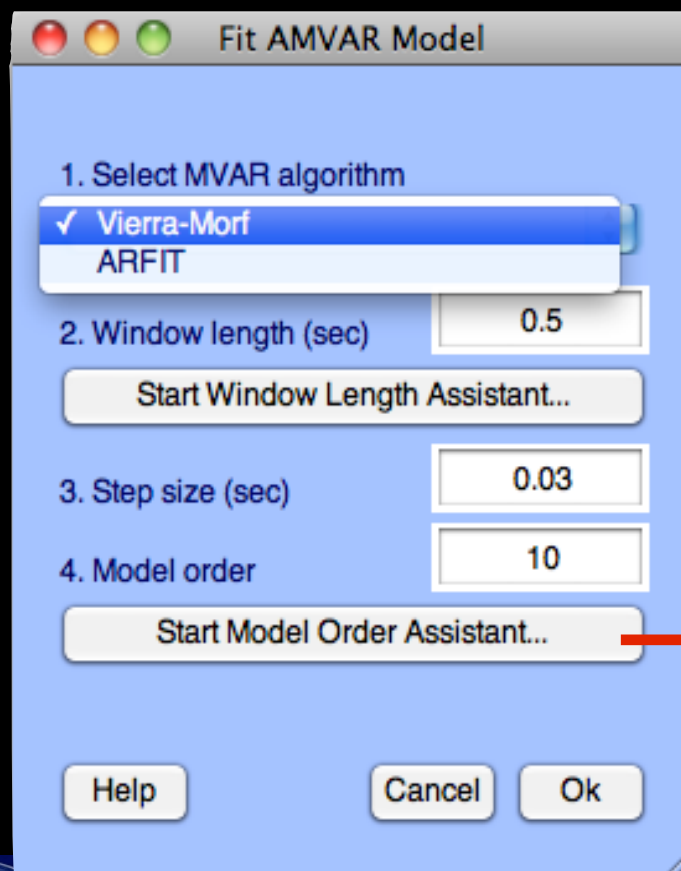
Statistics

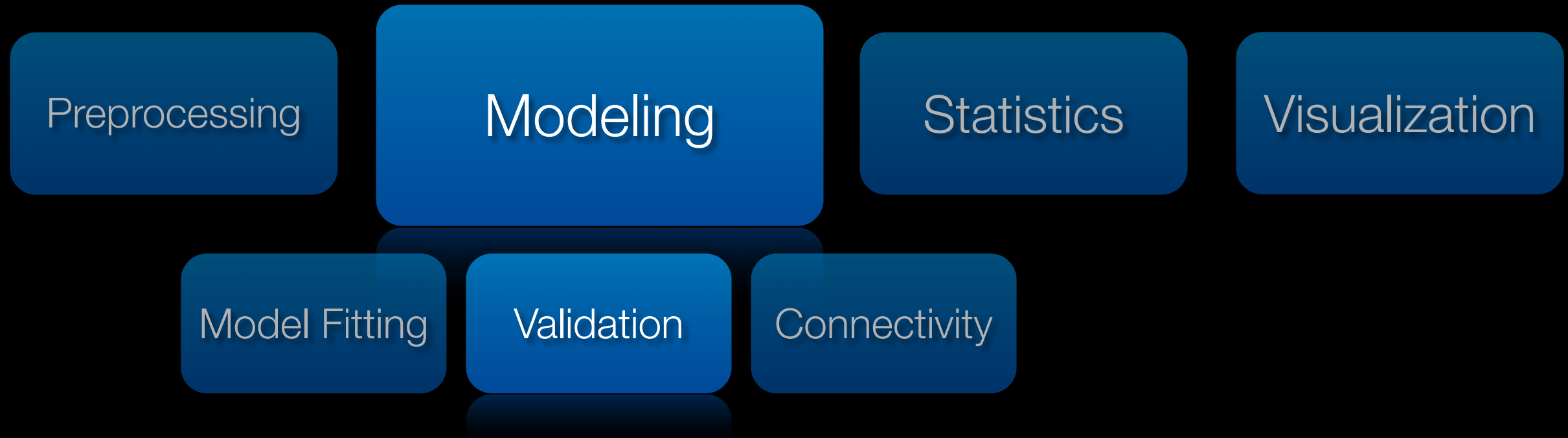
Visualization

Model Fitting

Validation

Connectivity





- ✖ Whiteness of Residuals
  - ✖ Portmanteau tests
  - ✖ Autocorrelation function
- ✖ Model Consistency
- ✖ Model Stability

Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

Pre-processing

Model fitting and validation ▶

Connectivity ▶

Statistics ▶

Visualization ▶

Fit AMVAR Model

Validate model



Preprocessing

Modeling

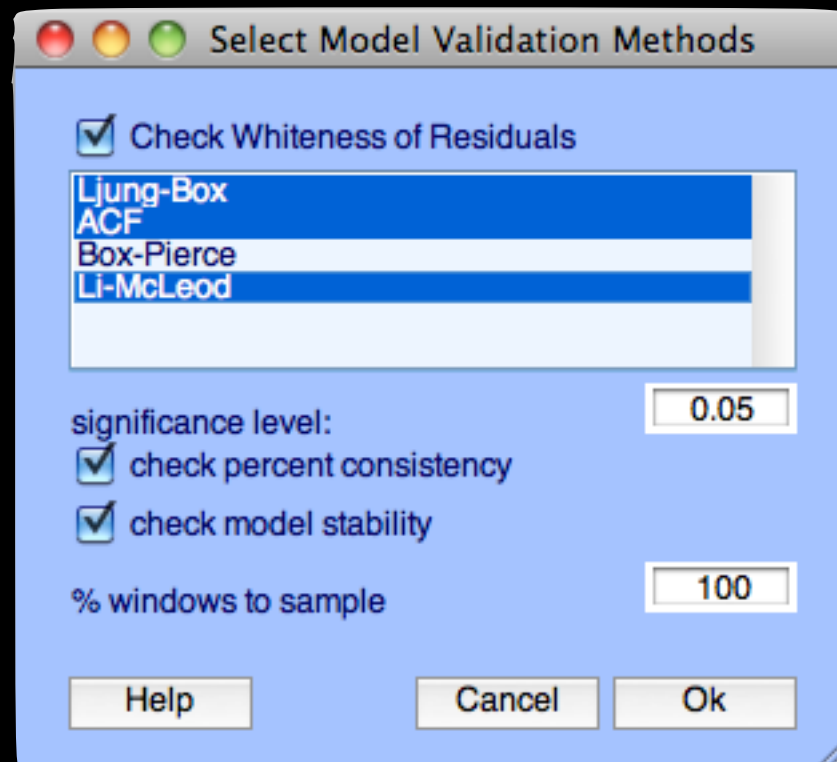
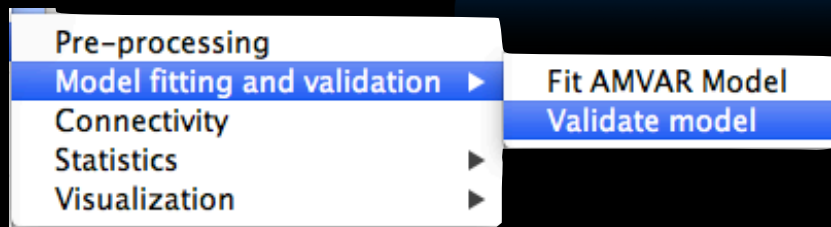
Statistics

Visualization

Model Fitting

Validation

Connectivity



Preprocessing

Modeling

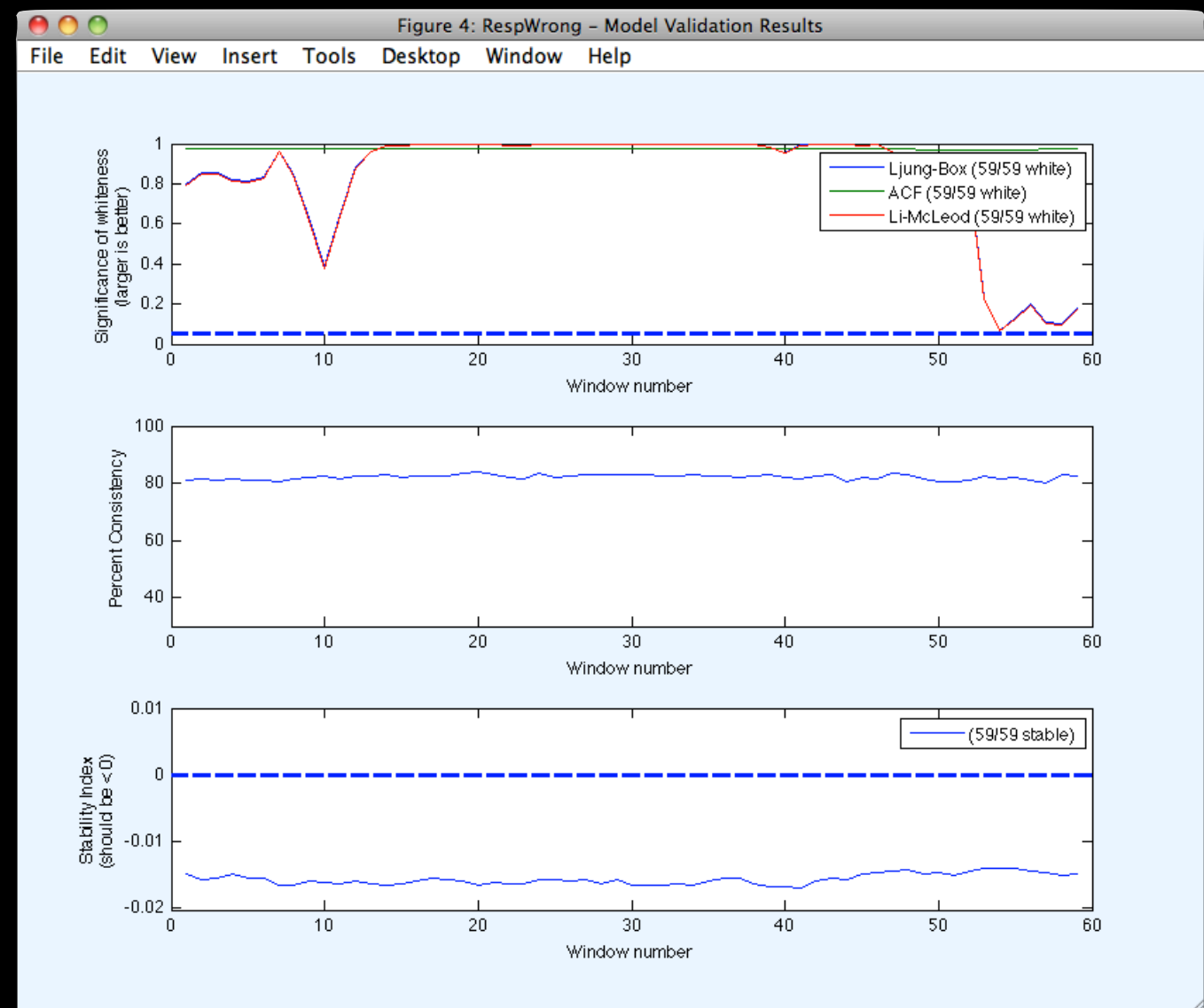
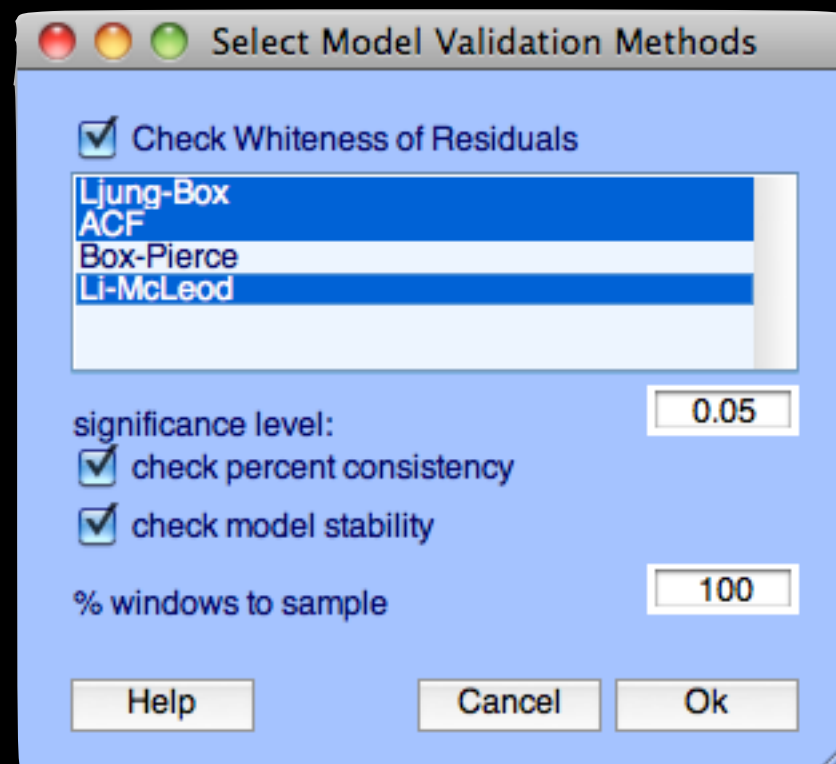
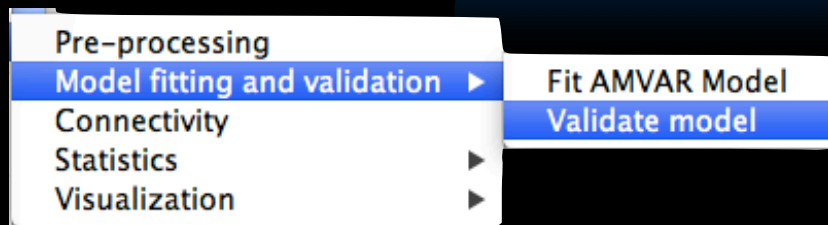
Statistics

Visualization

Model Fitting

Validation

Connectivity



Preprocessing

Modeling

Statistics

Visualization

Model Fitting

Validation

Connectivity

MVAR

- 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)
- Granger-Geweke Causality (GGC)
- Conditional GGC
- Blockwise GGC \*

Other

- Transfer Entropy \*
- Multivariate phase-locking value (mPLV) \*



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

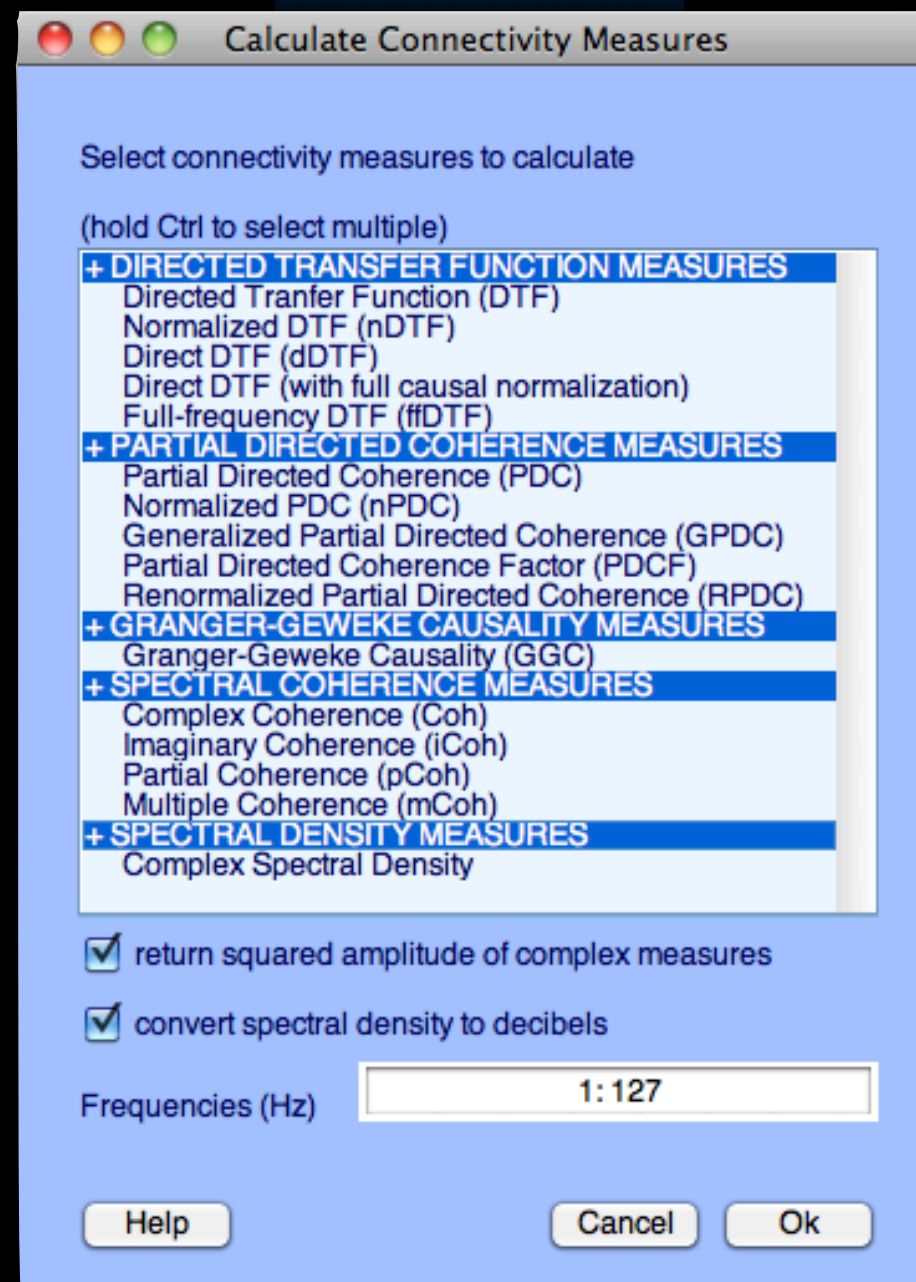
Visualization

Model Fitting

Validation

Connectivity

Pre-processing  
Model fitting and validation ▶  
**Connectivity**  
Statistics ▶  
Visualization ▶



Preprocessing

Modeling

Statistics

Visualization

Preprocessing

Modeling

Statistics

Visualization

## Parametric

Asymptotic analytic estimates of confidence intervals

Applies to: PDC, nPDC, DTF, nDTF, rPDC

Tests:  $H_{\text{null}}$ ,  $H_{\text{base}}$ ,  $H_{\text{AB}}$

Confidence intervals using thin-plate smoothing splines

Applies to: dDTF

Tests:  $H_{\text{base}}$ ,  $H_{\text{AB}}$

$$H_{\text{null}} : \mathbf{C}_{ij} \leq \mathbf{C}_{\text{null}}$$

$$H_{\text{base}} : \mathbf{C}_{ij} \leq \mathbf{C}_{\text{baseline}}$$

$$H_{\text{AB}} : \mathbf{C}_{ij}^{\text{A}} = \mathbf{C}_{ij}^{\text{B}}$$



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization

## Parametric

Asymptotic analytic estimates of confidence intervals

Applies to: PDC, nPDC, DTF, nDTF, rPDC

Tests:  $H_{\text{null}}$ ,  $H_{\text{base}}$ ,  $H_{\text{AB}}$

Confidence intervals using thin-plate smoothing splines

Applies to: dDTF

Tests:  $H_{\text{base}}$ ,  $H_{\text{AB}}$

## Non-parametric

Phase-randomization

Applies to: all

Tests:  $H_{\text{null}}$

Permutation Tests

Applies to: all

Tests:  $H_{\text{AB}}$ ,  $H_{\text{base}}$

Bootstrap and Jackknife

Applies to: all

Tests:  $H_{\text{AB}}$ ,  $H_{\text{base}}$

$$H_{\text{null}} : \mathbf{C}_{ij} \leq \mathbf{C}_{\text{null}}$$

$$H_{\text{base}} : \mathbf{C}_{ij} \leq \mathbf{C}_{\text{baseline}}$$

$$H_{\text{AB}} : \mathbf{C}_{ij}^{\text{A}} = \mathbf{C}_{ij}^{\text{B}}$$



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization



fully implemented



partially-developed



coming soon



Preprocessing

Modeling

Statistics

Visualization

Interactive Time-Frequency Grid



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization

Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization

Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie

Causal Density Movie



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization

Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie

Causal Density Movie

Directed Graphs on anatomicals (ECoG)



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization

Interactive Time-Frequency Grid

Interactive 3D Causal Brainmovie

Causal Density Movie

Directed Graphs on anatomicals (ECoG)

and more...



fully implemented



partially-developed



coming soon

Preprocessing

Modeling

Statistics

Visualization

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



fully implemented

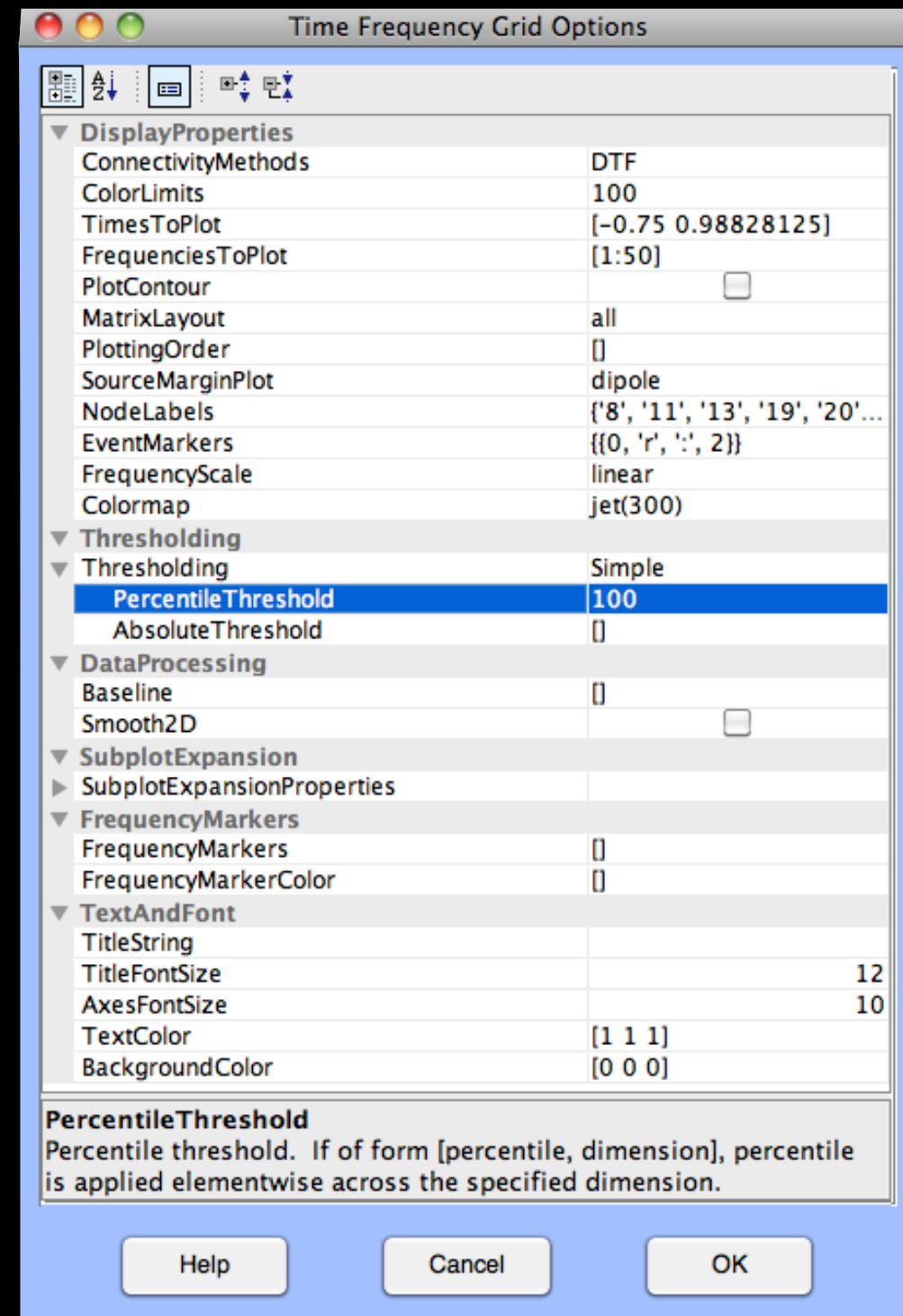
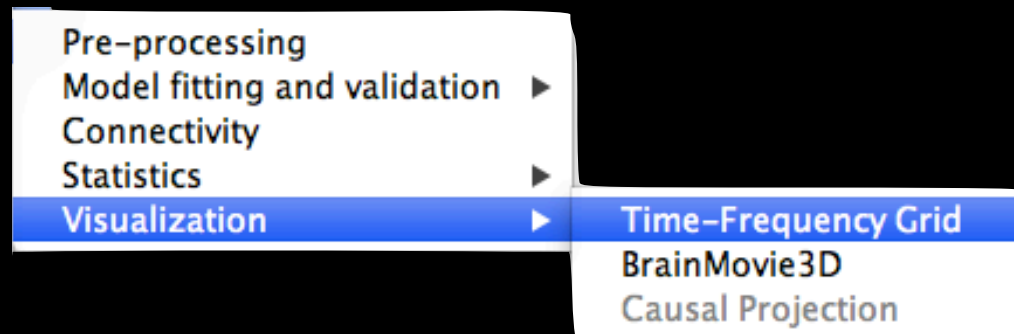


partially-developed

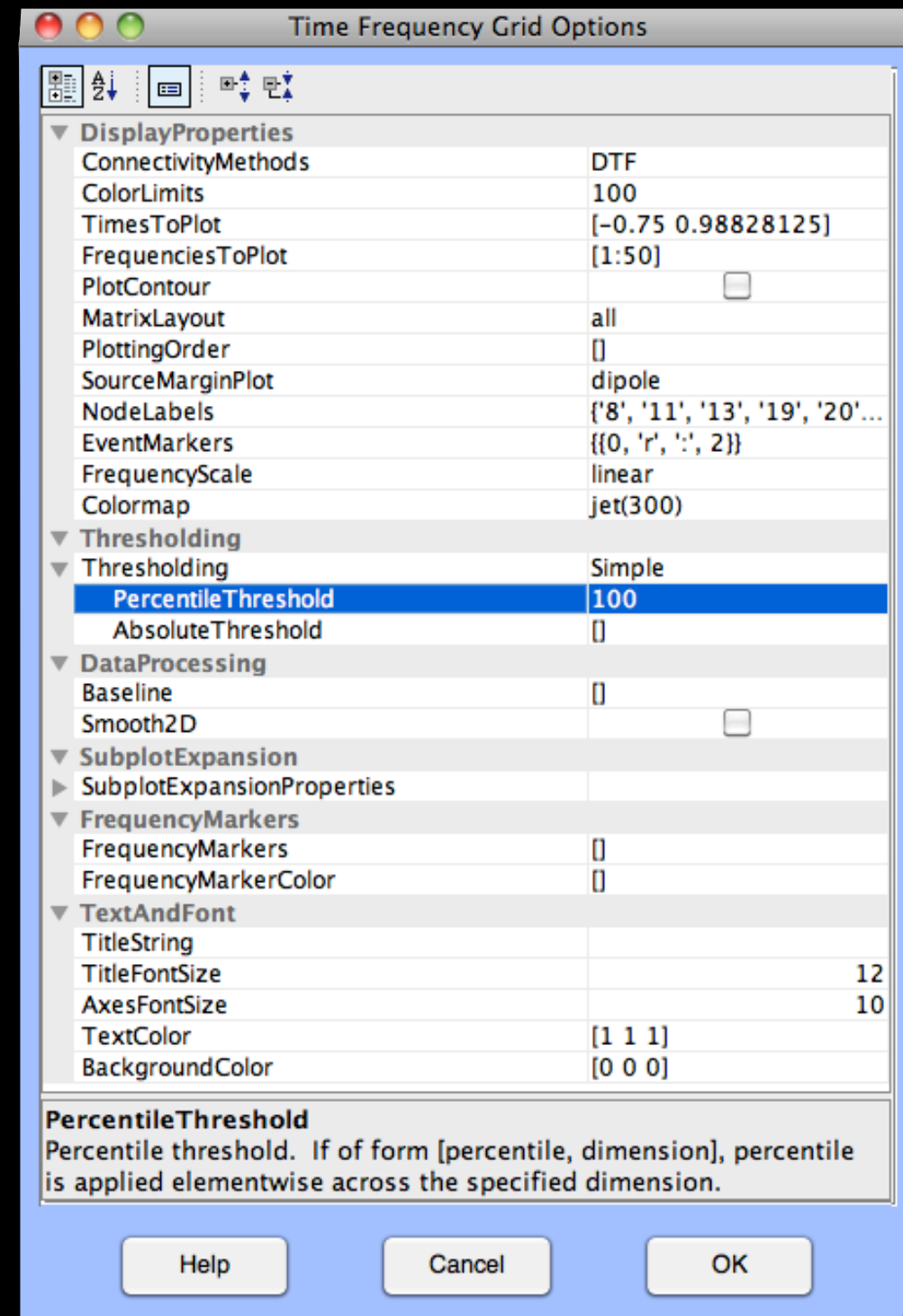
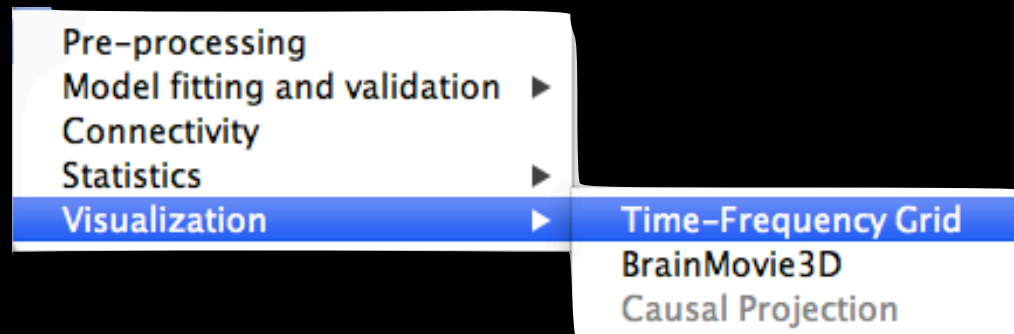


coming soon

# Interactive Time-Frequency Grid

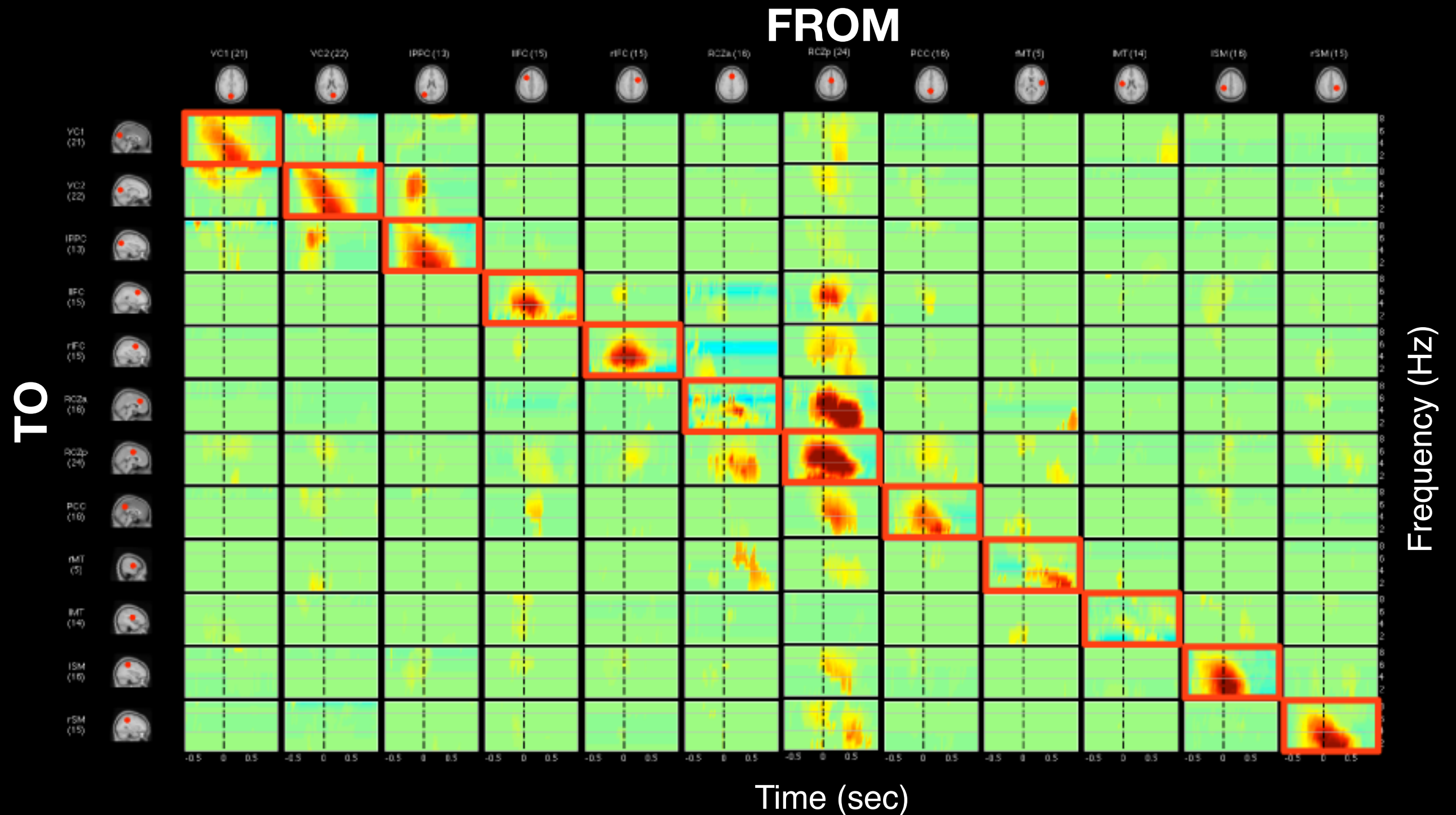


# Interactive Time-Frequency Grid

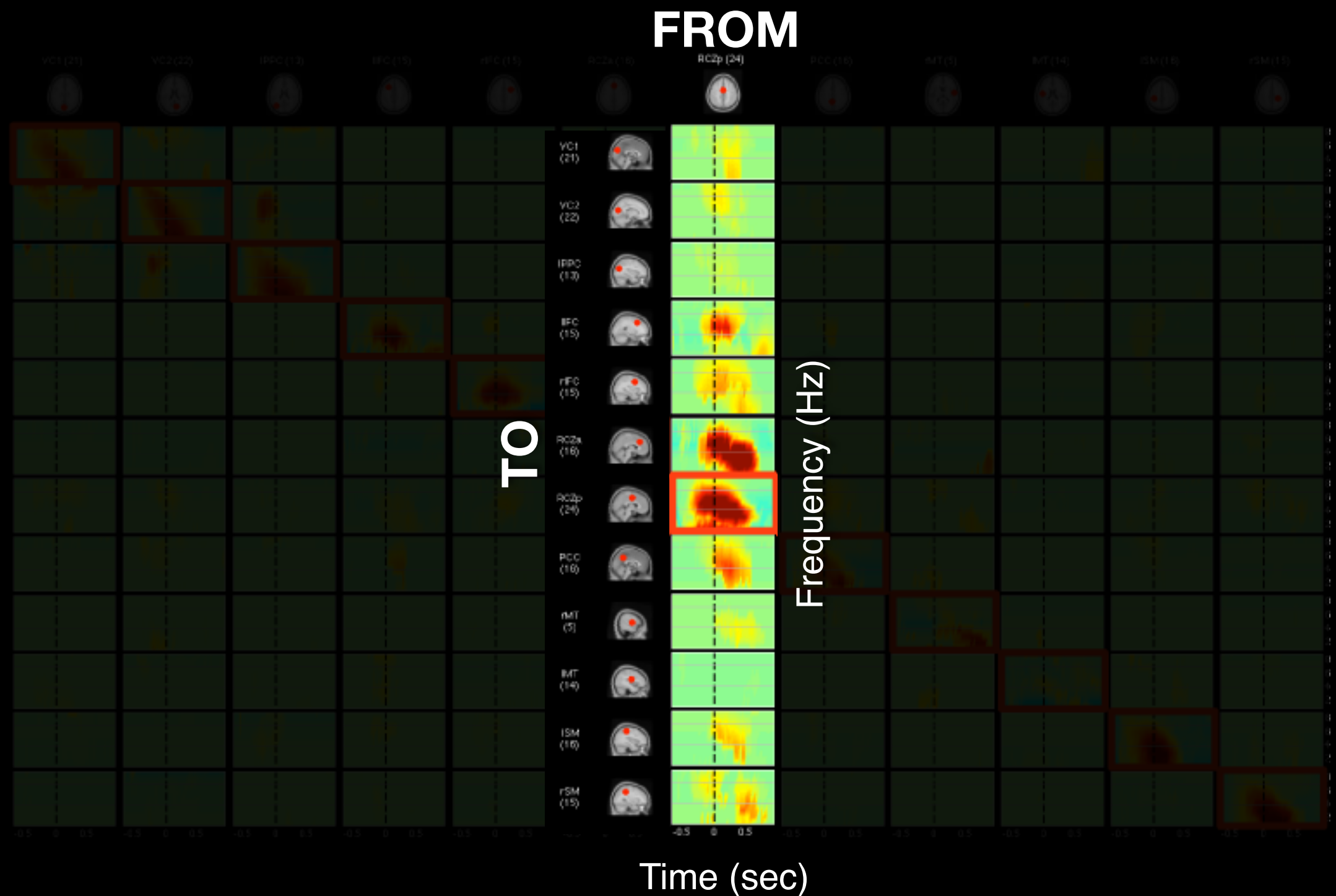




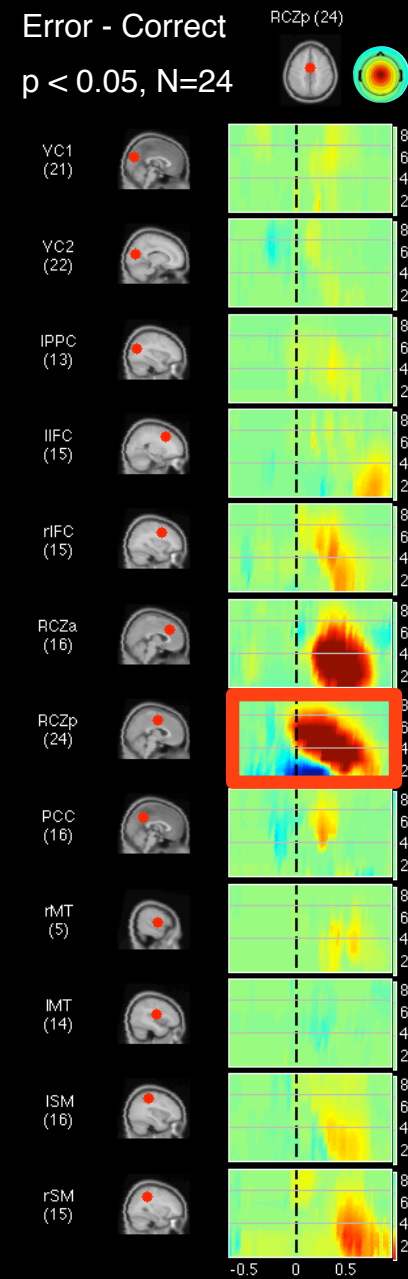
# Causal Time-Frequency Grid



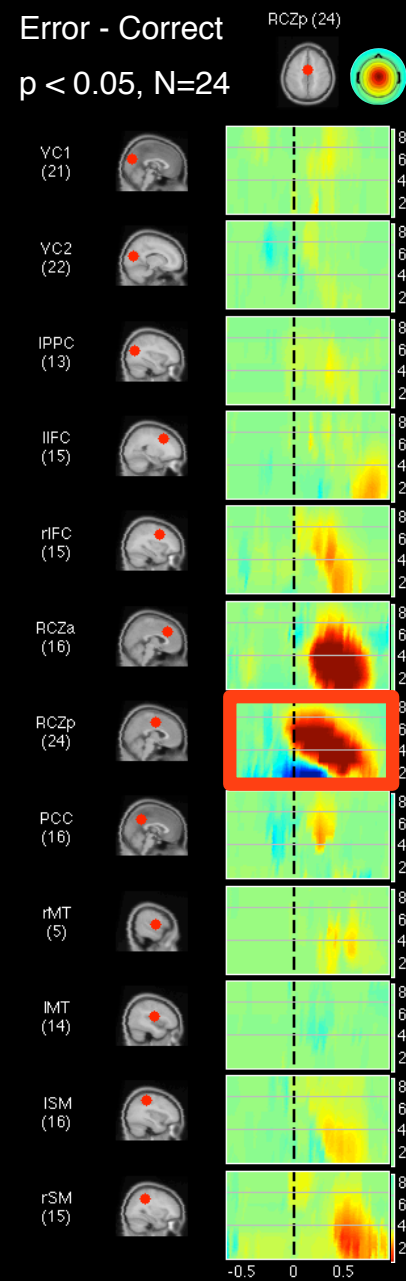
# Causal Time-Frequency Grid



# Causal Time-Frequency Grid



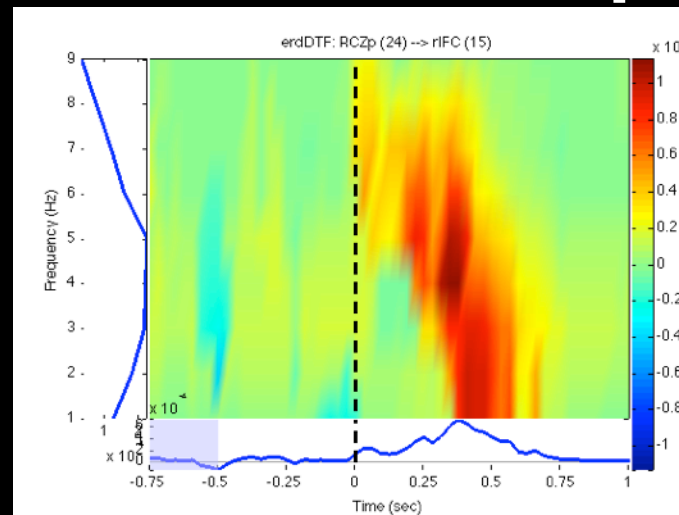
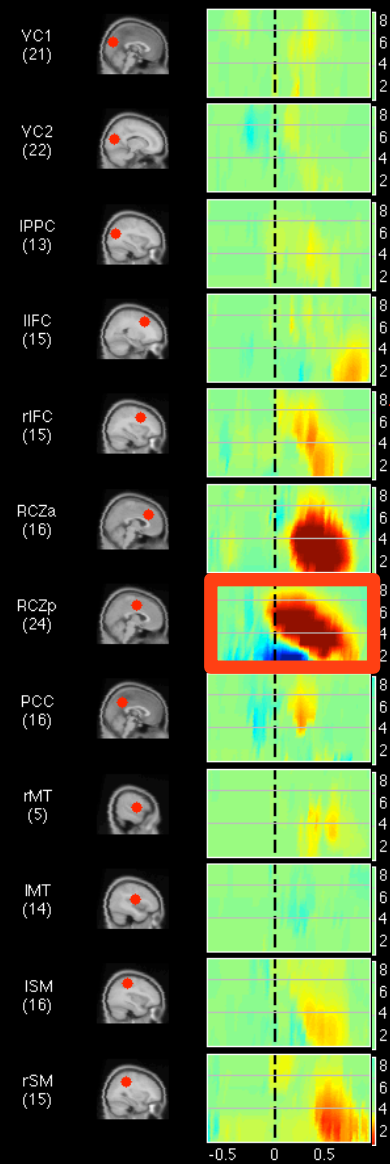
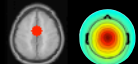
# Causal Time-Frequency Grid



# Causal Time-Frequency Grid

Error - Correct  
 $p < 0.05$ ,  $N=24$

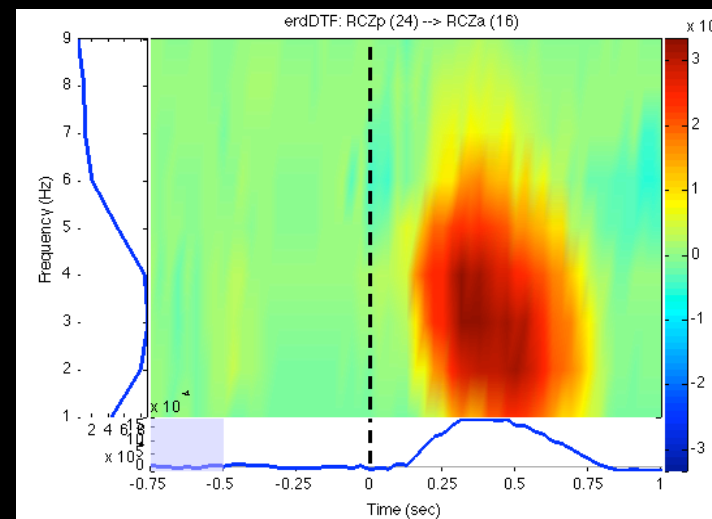
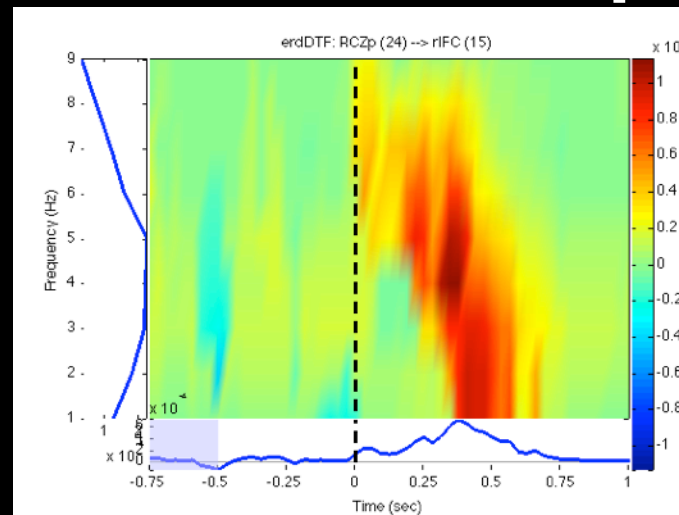
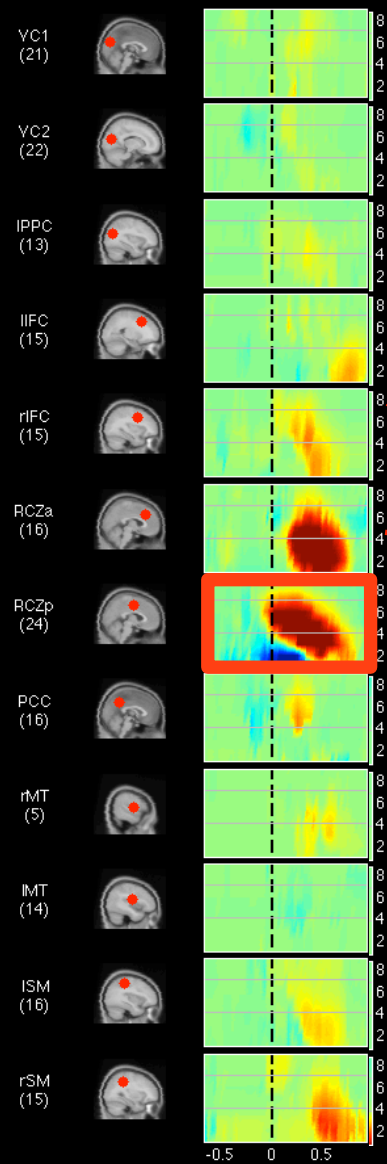
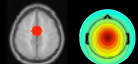
RCZp (24)



# Causal Time-Frequency Grid

Error - Correct  
 $p < 0.05$ ,  $N=24$

RCZp (24)

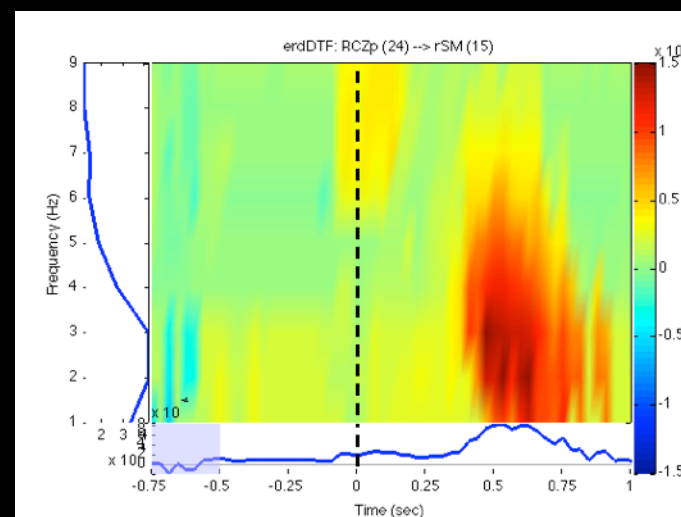
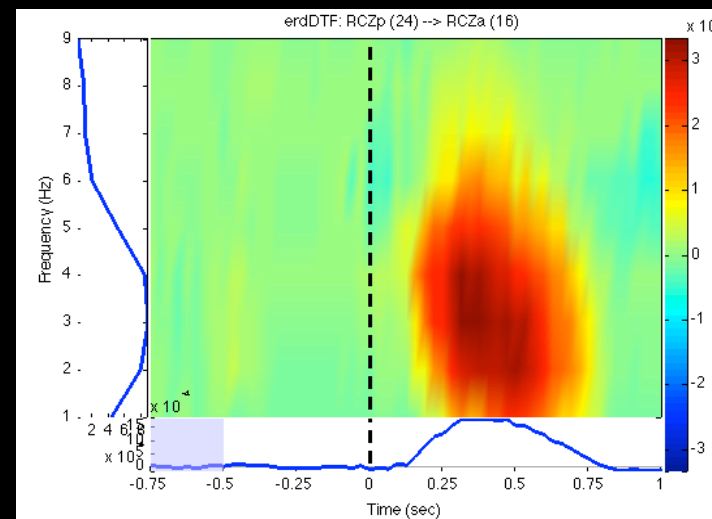
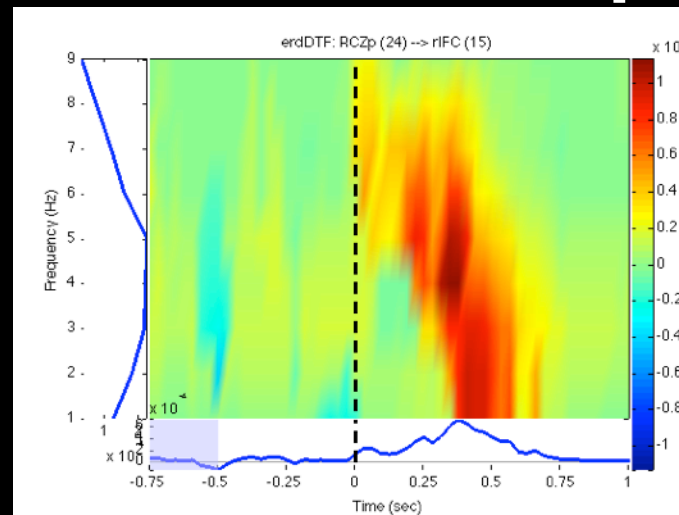
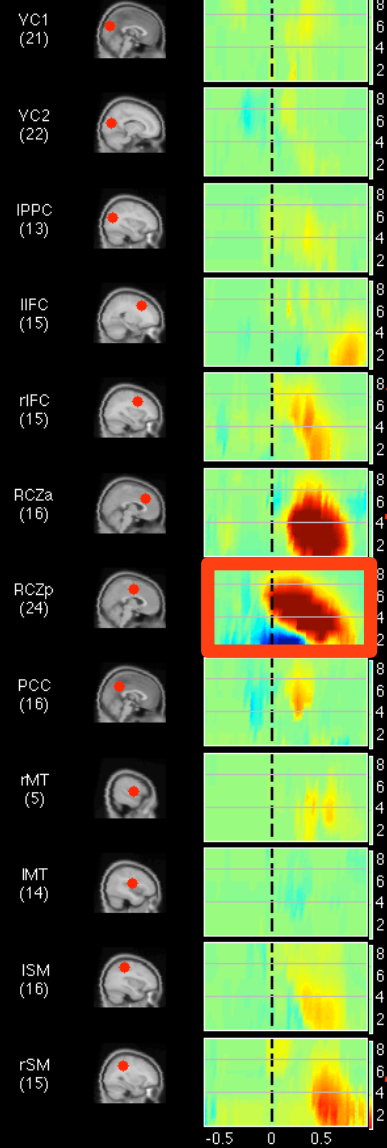
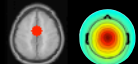




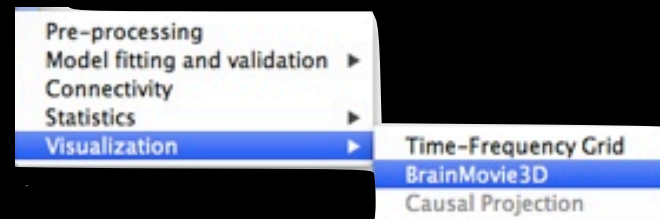
# Causal Time-Frequency Grid

Error - Correct  
 $p < 0.05$ ,  $N=24$

RCZp (24)

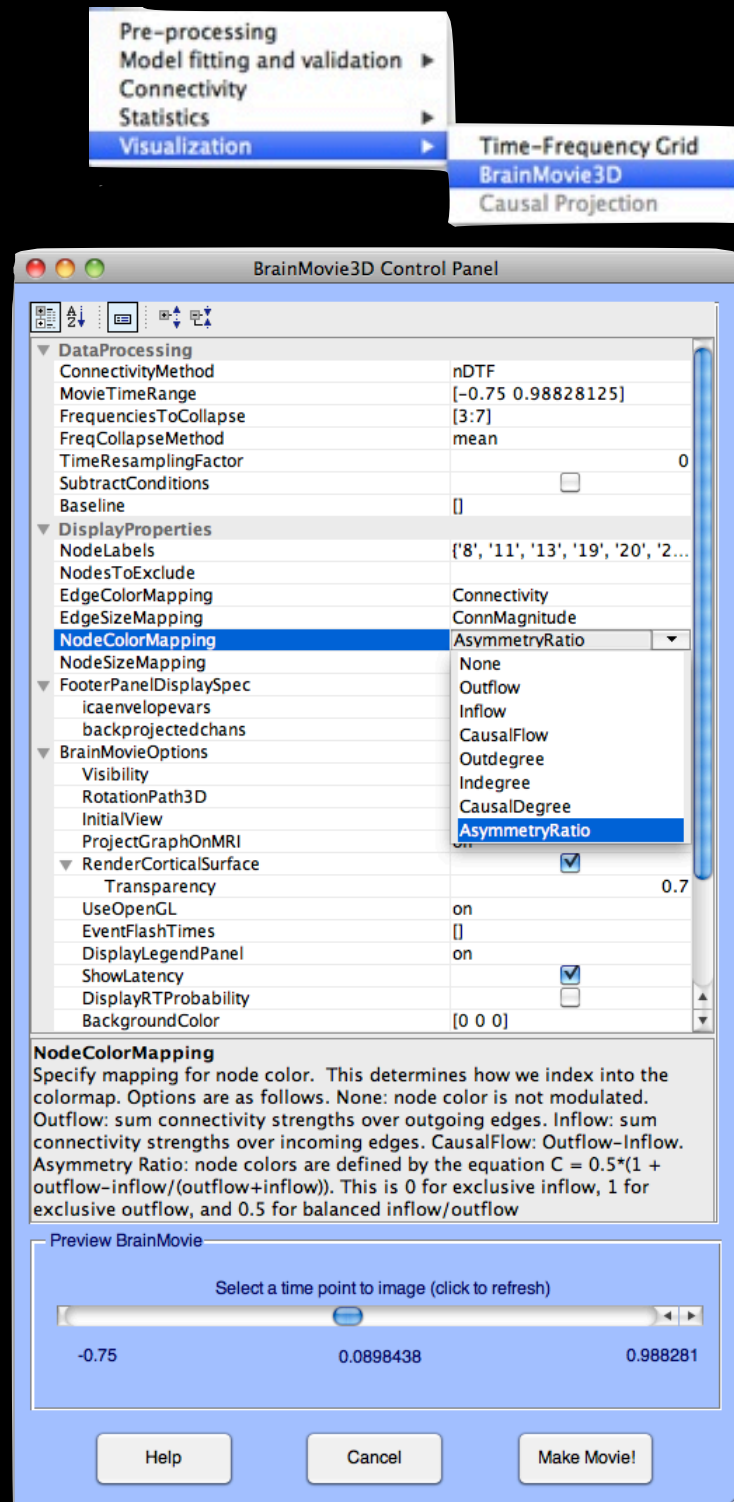


# Interactive BrainMovie3D

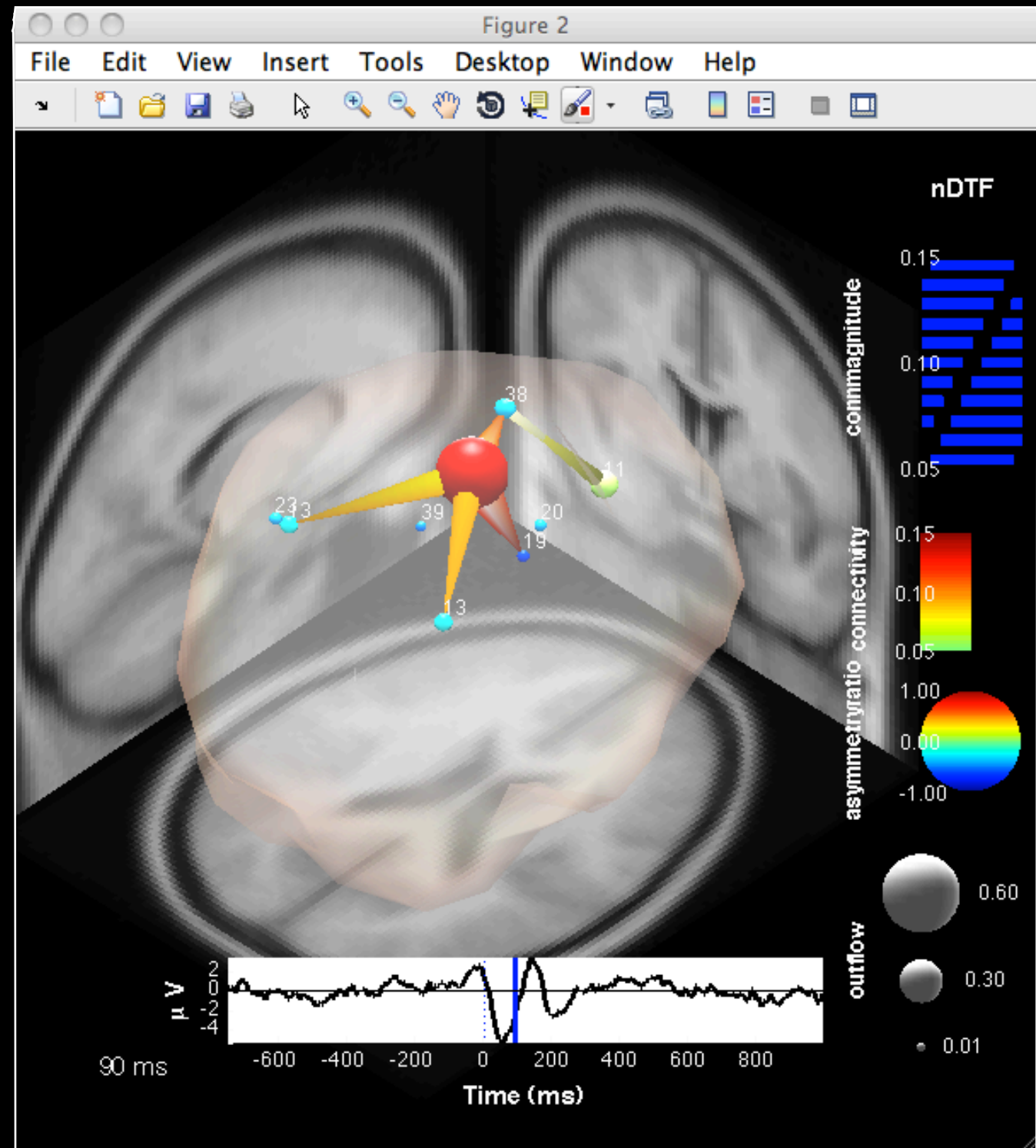
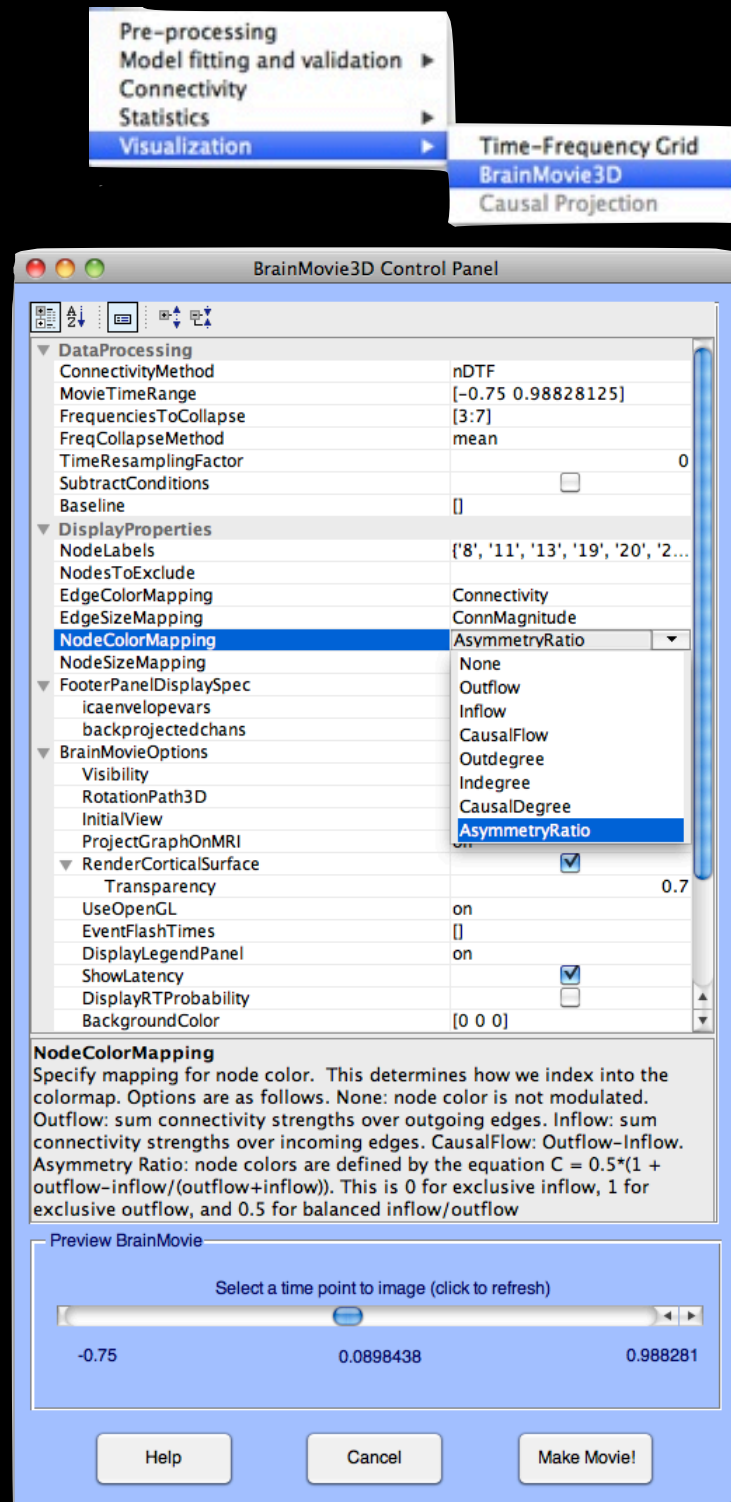




# Interactive BrainMovie3D



# Interactive BrainMovie3D

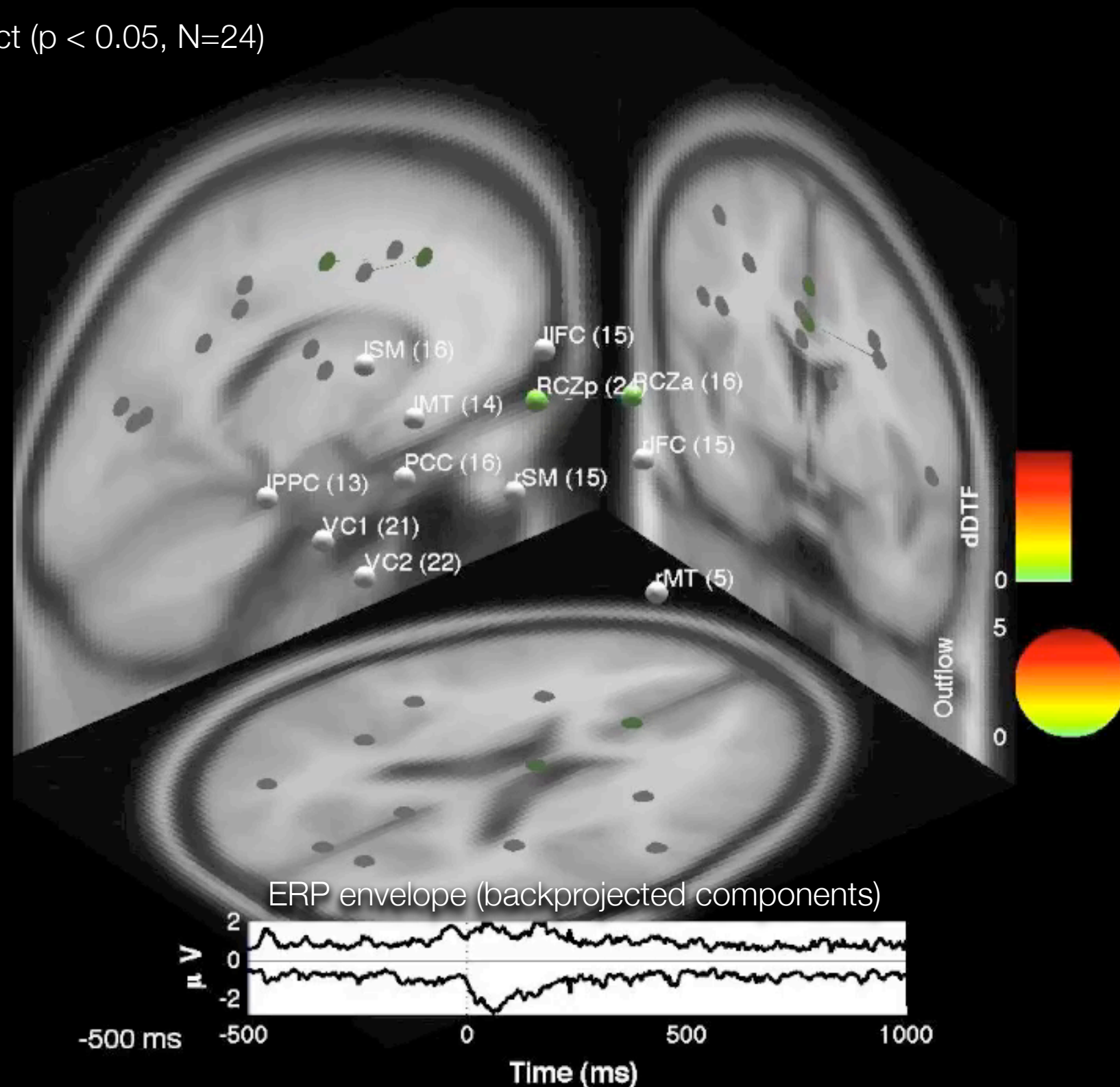


# Interactive BrainMovie3D

Error > Correct ( $p < 0.05$ ,  $N=24$ )

dDTF

3-7 Hz



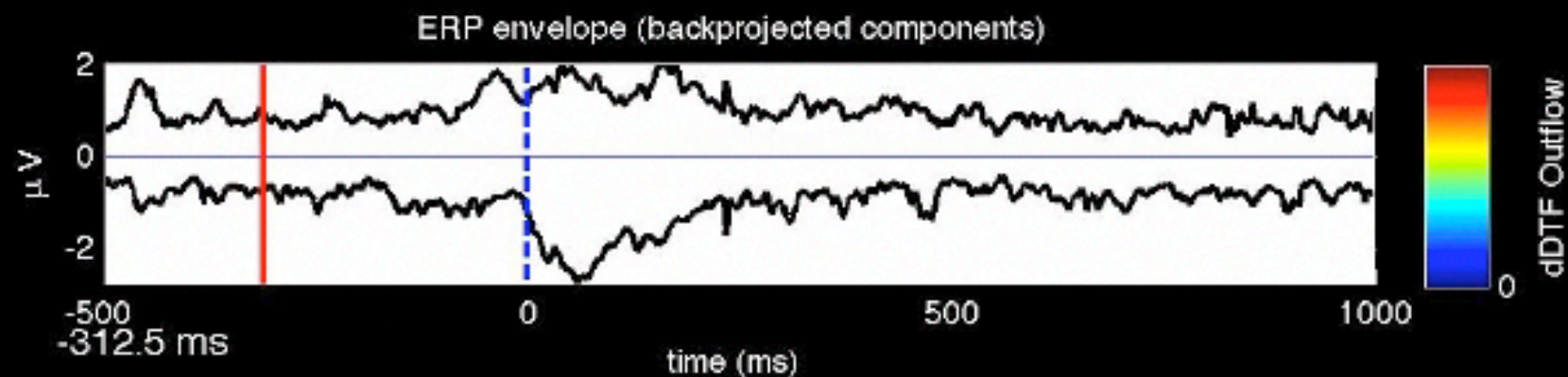
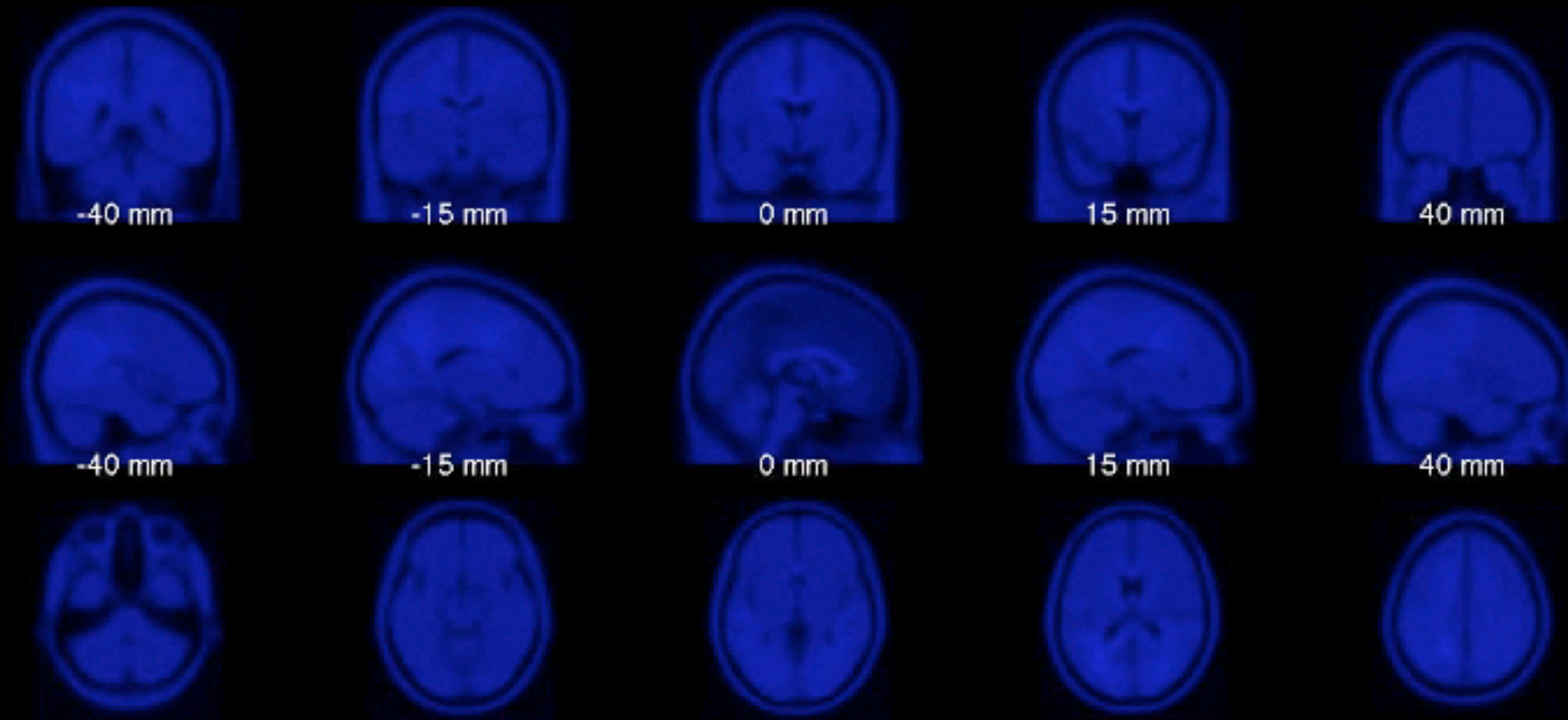


# Causal Density Movie

Error > Correct ( $p < 0.05$ ,  $N=24$ )

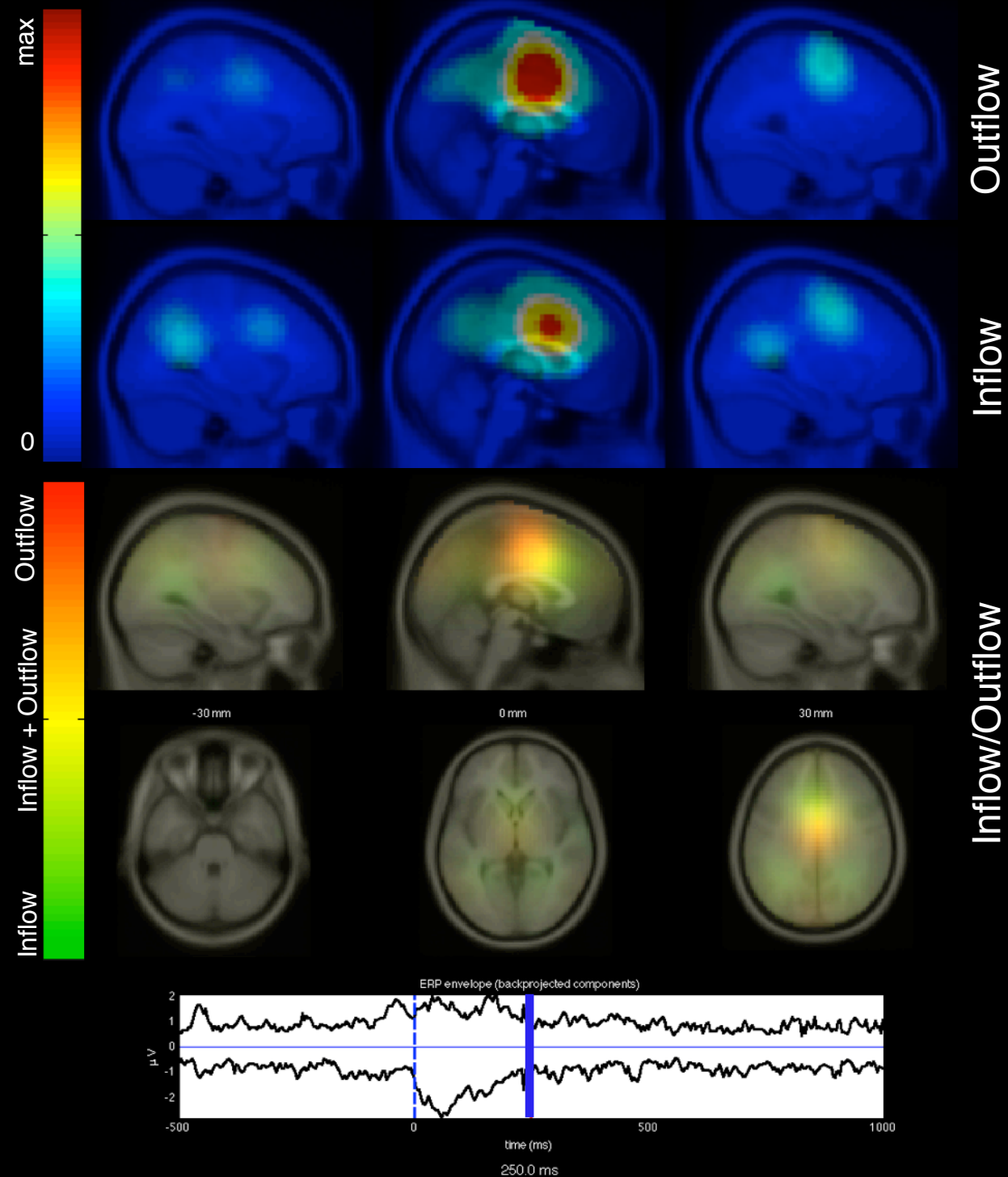
dDTF

3-7 Hz



# Causal Density Movie

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

# 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).

## Bayesian Mixture Model

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.

■ partially-developed



# Future Work

- ✦ Improvement of architecture, GUI, and EEGLAB integration
- ✦ Ongoing implementation/incorporation of state-of-the-art methods for effective connectivity analysis and visualization
- ✦ Improved group statistics
- ✦ Evaluation of relative suitability of various source-separation algorithms when combined with MVAR modeling