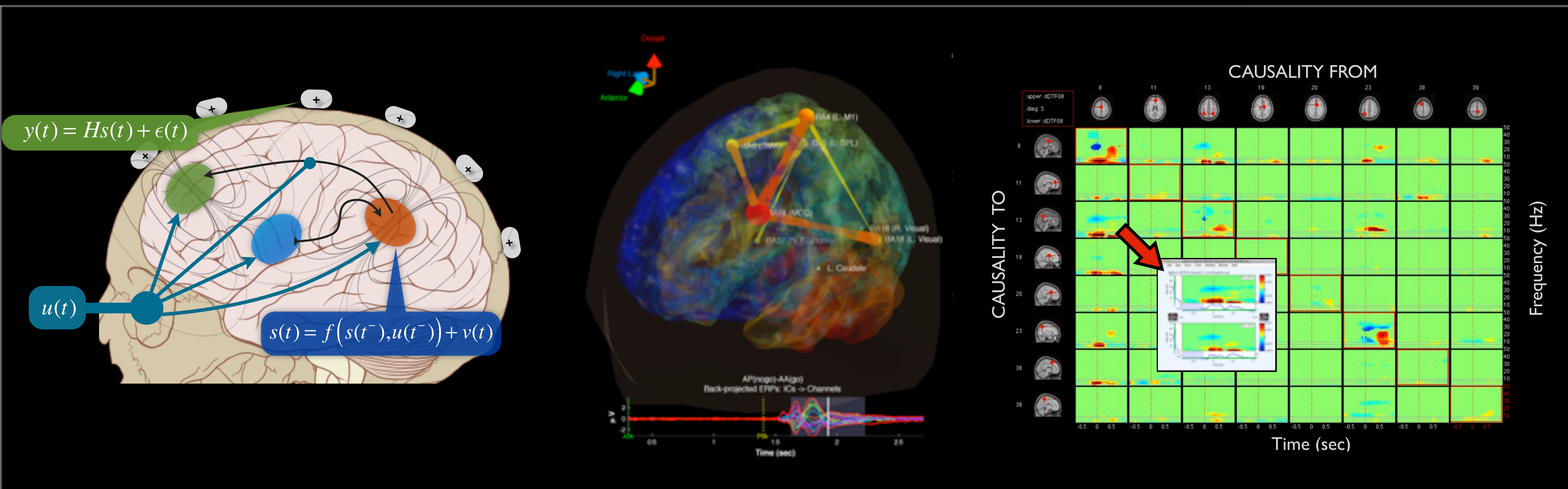
# The Dynamic Brain I: Modeling Neural Dynamics and Interactions from M/EEG




## Tim Mullen, PhD

Swartz Center for Computational Neuroscience
Institute for Neural Computation
UC San Diego
and Qusp Labs

Tim Mullen

# Outline

## Theoretical Foundations I

- Functional Connectivity Measures (PLV, PAC, Coherence)
- Linear Dynamical Systems and the VAR model
- Granger Causality and Effective Connectivity Measures
- Scalp versus Source
- Adapting to Time-Varying Dynamics

## Practicum: Hands-On Walkthrough of SIFT

Tim Mullen

# Preview Outline (Sunday)

## Theoretical Foundations II

- Model Validation
- Multivariate vs. Bivariate
- Imposing Constraints
- Single-trial Estimation and State-Space Models
- Statistical Testing

## Practicum: Hands-On Simulation-based training

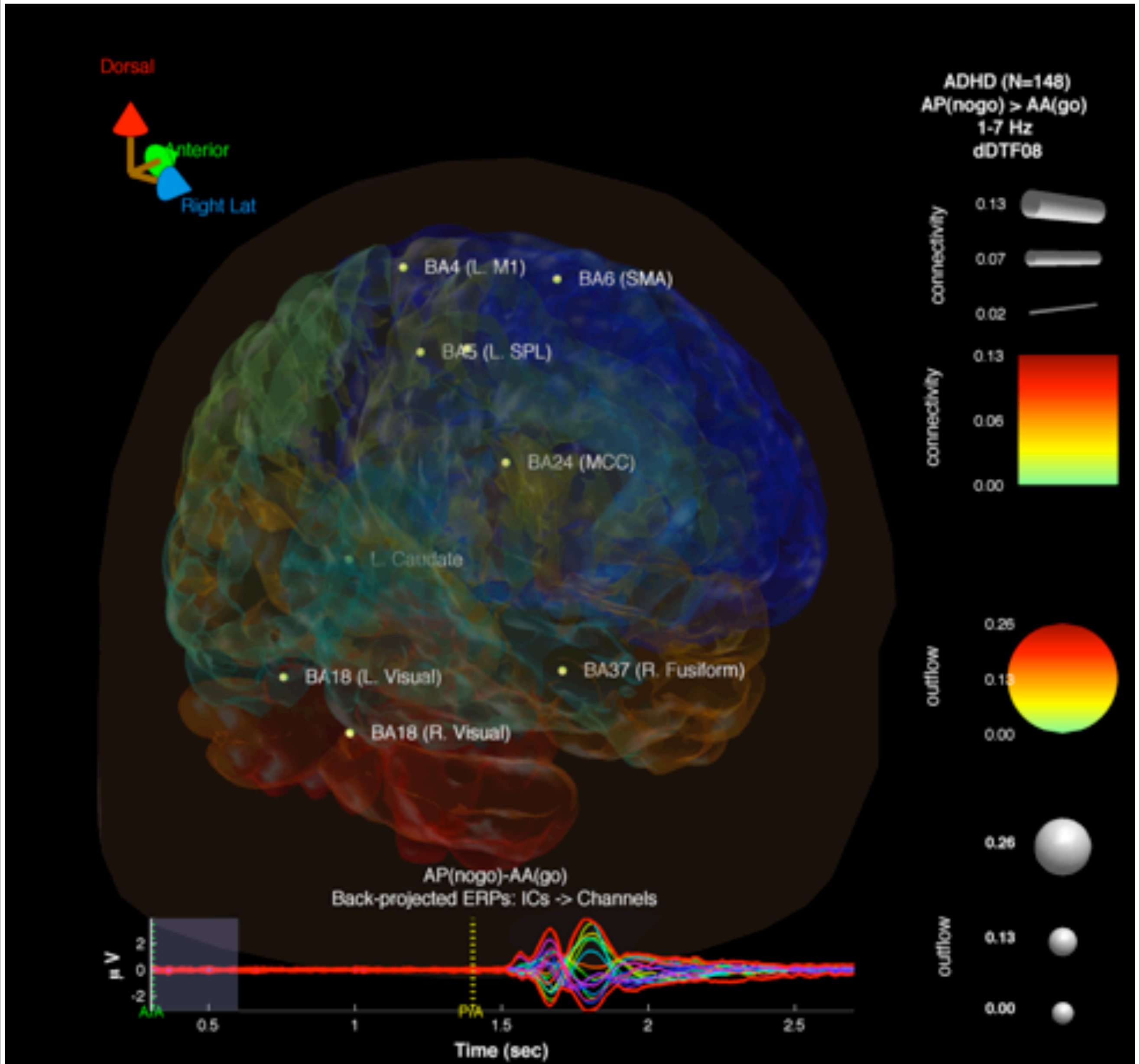
Intro Theory SIFT Apps To-Do Fin

Tim Mullen

# 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

Tim Mullen

# 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):

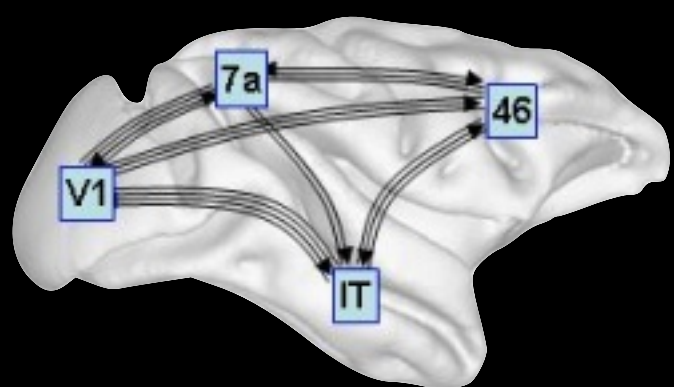
> “Essentially, all models are wrong, but some are useful [...] the practical question is how wrong do they have to be to not be useful”

Tim Mullen

# Categorizations of Large-Scale Brain Connectivity Analysis

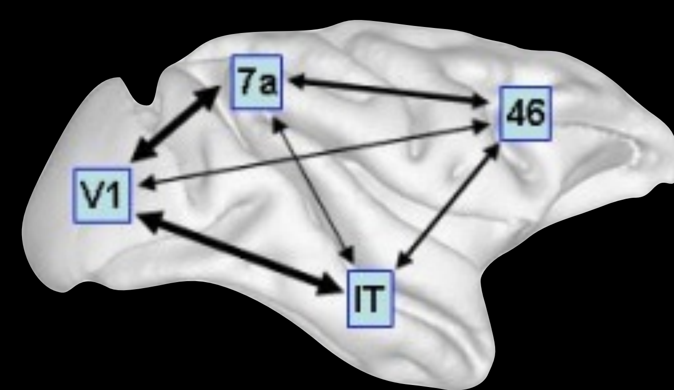
(Bullmore and Sporns, *Nature*, 2009)

**Structural**



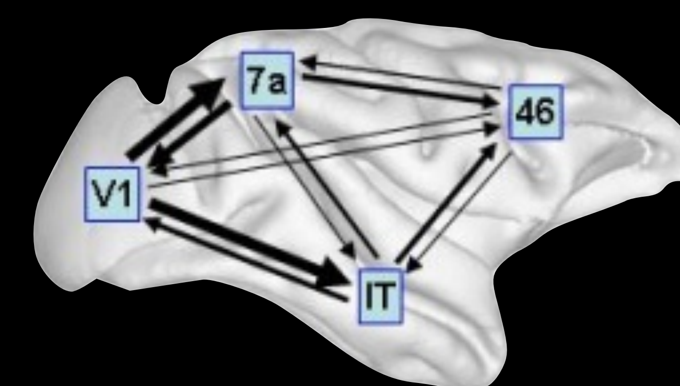

state-invariant, anatomical

**Functional**




dynamic, state-dependent, correlative, symmetric

**Effective**




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

**Hours-Years** **milliseconds-seconds**

Temporal Scale

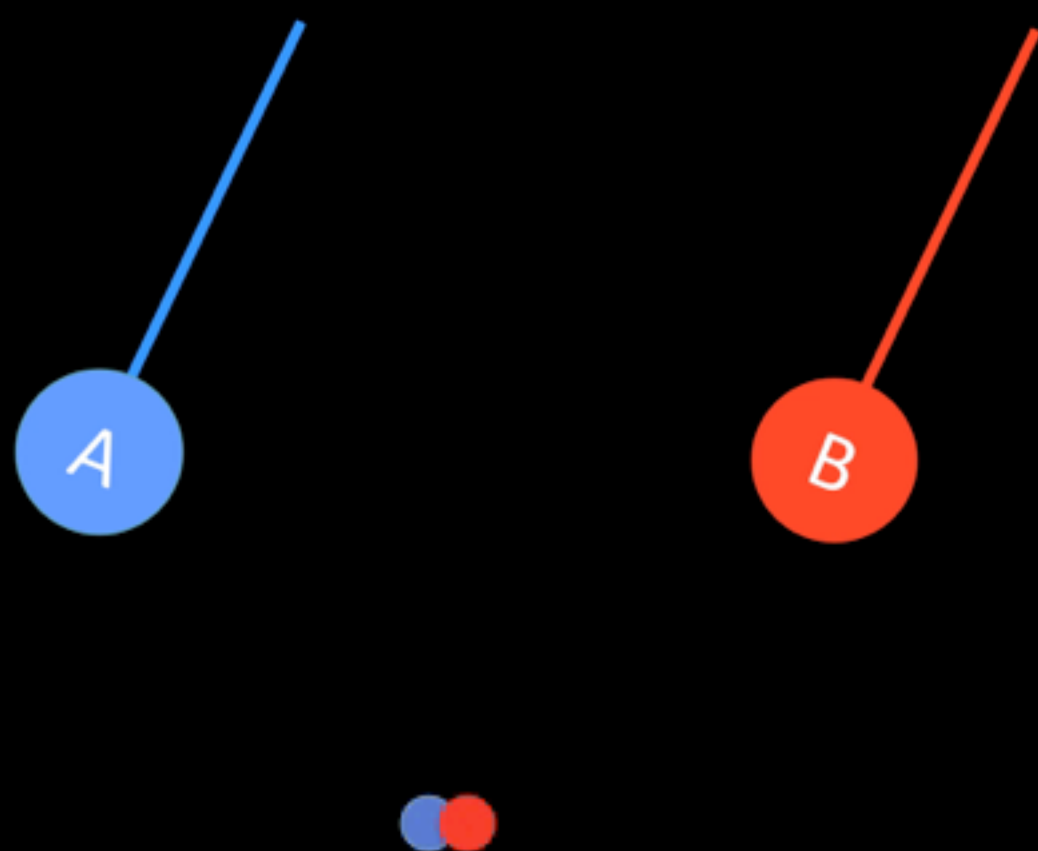
# Estimating Functional Connectivity

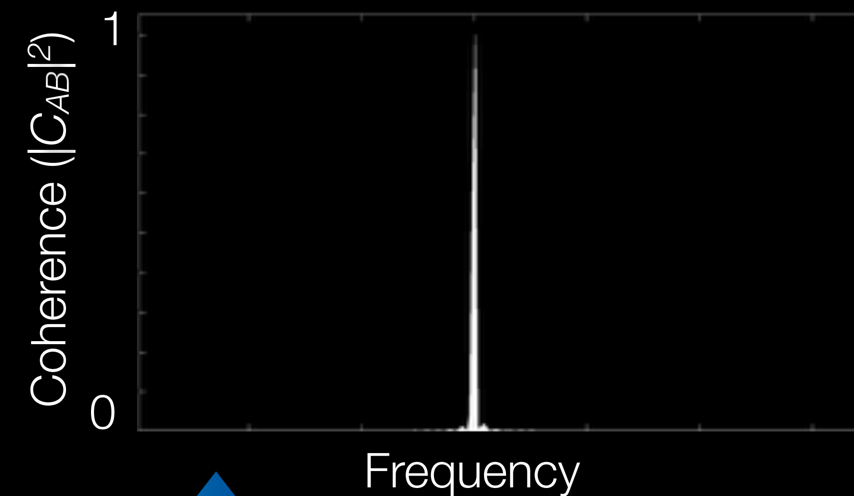
Tim Mullen

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

# Cross-Correlation and Linear Coherence

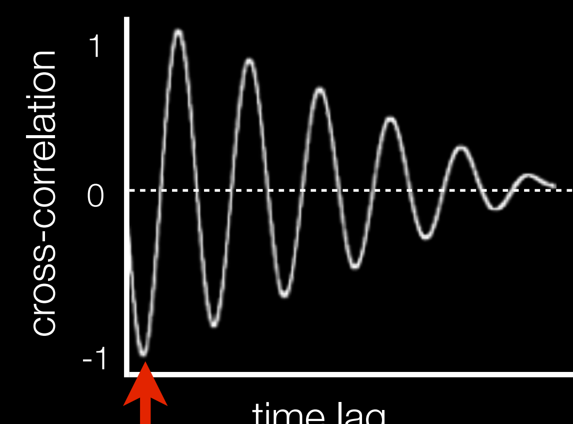
Tim Mullen




DFT

$$C_{AB}(f) = \sum_{k=0}^{p} r_{AB}(k)e^{-i2p fk}$$
$$= \frac{S_{AB}(f)}{\sqrt{S_A(f)S_B(f)}}$$

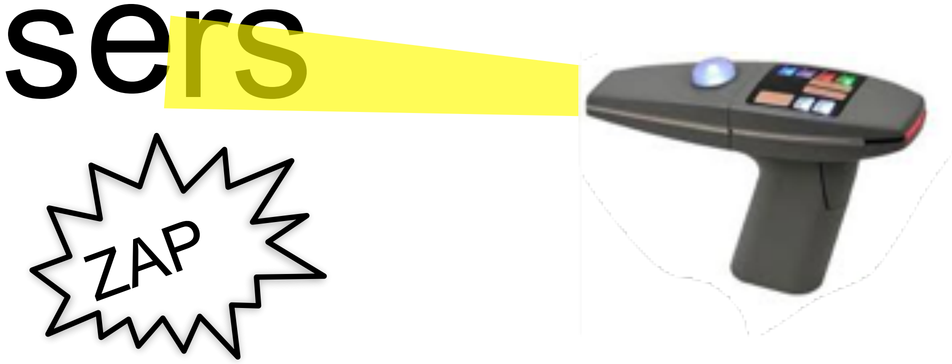



$r_{AB}(k)$

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

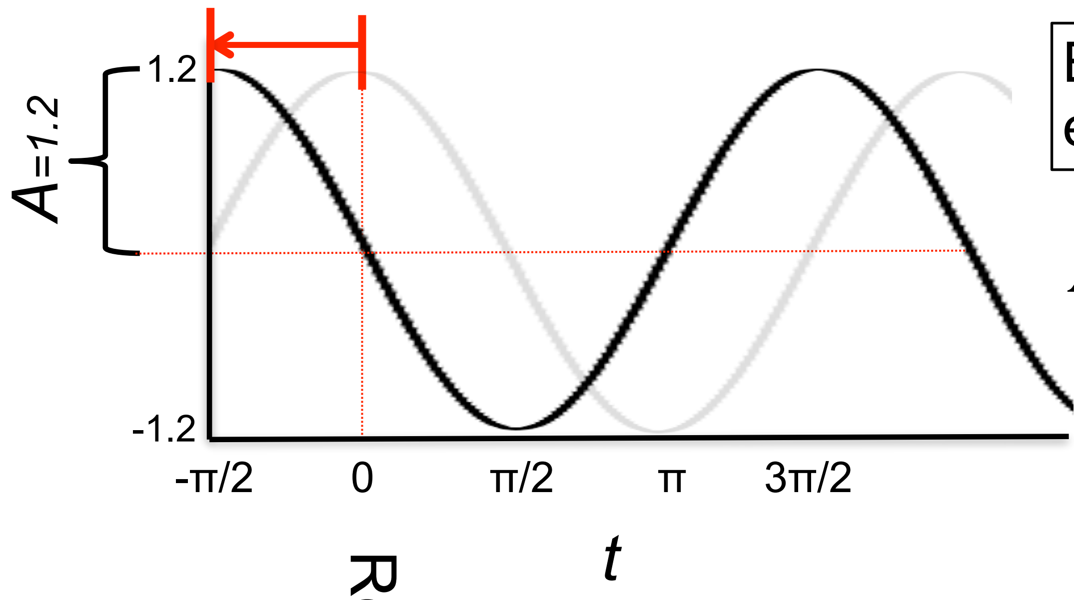
Tim Mullen

# Phasers




phase shift $f = p / 2$

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

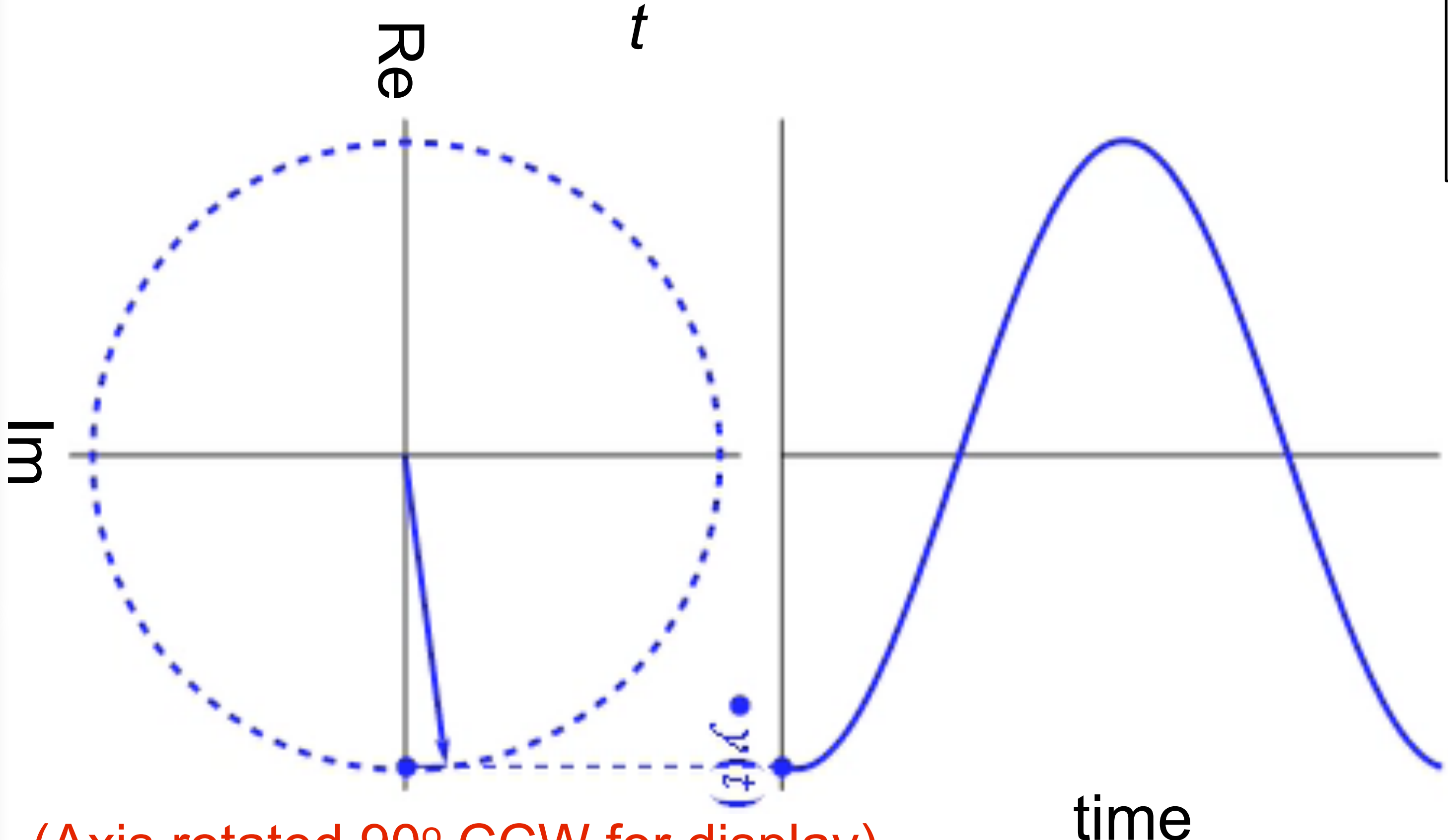



Euler's Formula tells us that any sinusoid can be expressed as the sum of two complex exponentials

$$A \cdot \cos(wt + f) = \frac{A}{2} e^{i(wt+f)} + \frac{A}{2} e^{-i(wt+f)}$$
$$= \mathrm{Re}\{Ae^{i(wt+f)}\} = \mathrm{Re}\{S(w,t)\}$$

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

instantaneous complex amplitude and phase




(Axis rotated 90º CCW for display)

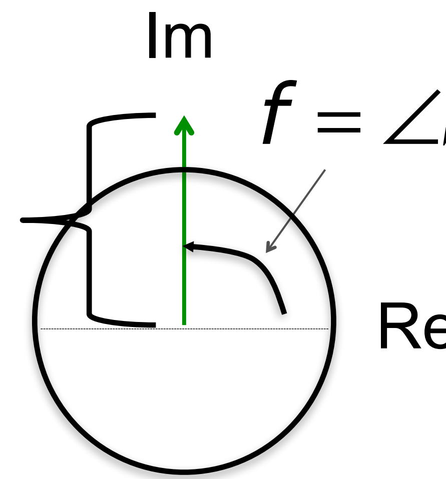
Polar animation courtesy Wikipedia

**Phasor**

*(Polar Coords)*

$f = \angle S(w,t)$

$|S(w,t)| = |A|$



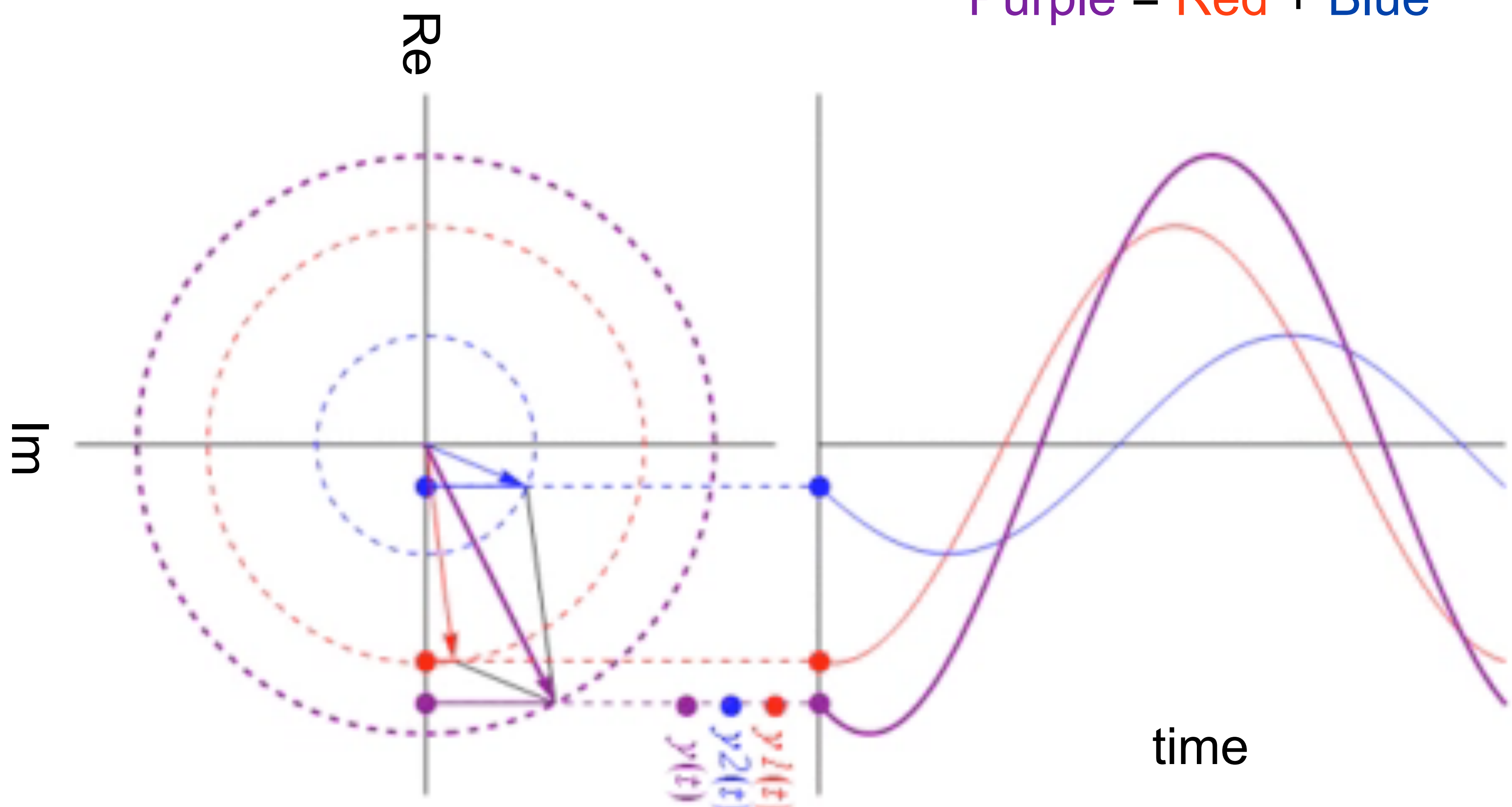

Shorthand notation: $Ae^{if}$

Tim Mullen

# 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



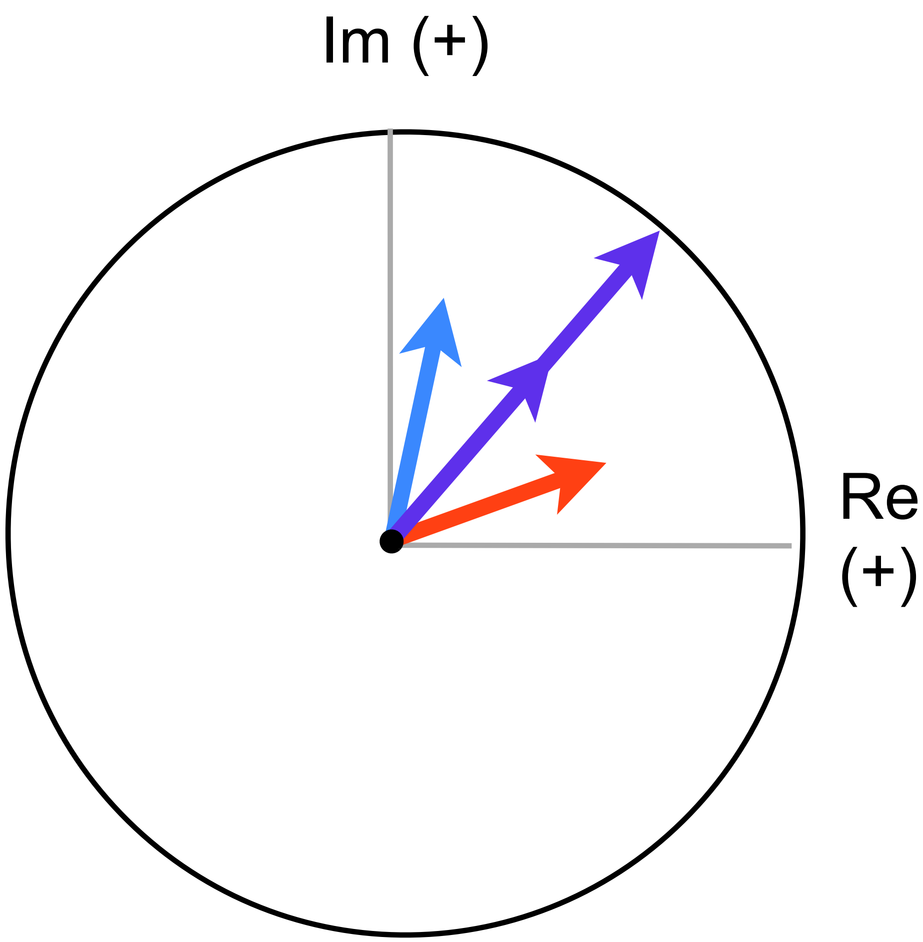

(Axis rotated 90° CCW for display)

Polar animation courtesy Wikipedia

Tim Mullen

# 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*.




$$\text{Purple} = \frac{\text{Red} + \text{Blue}}{2}$$

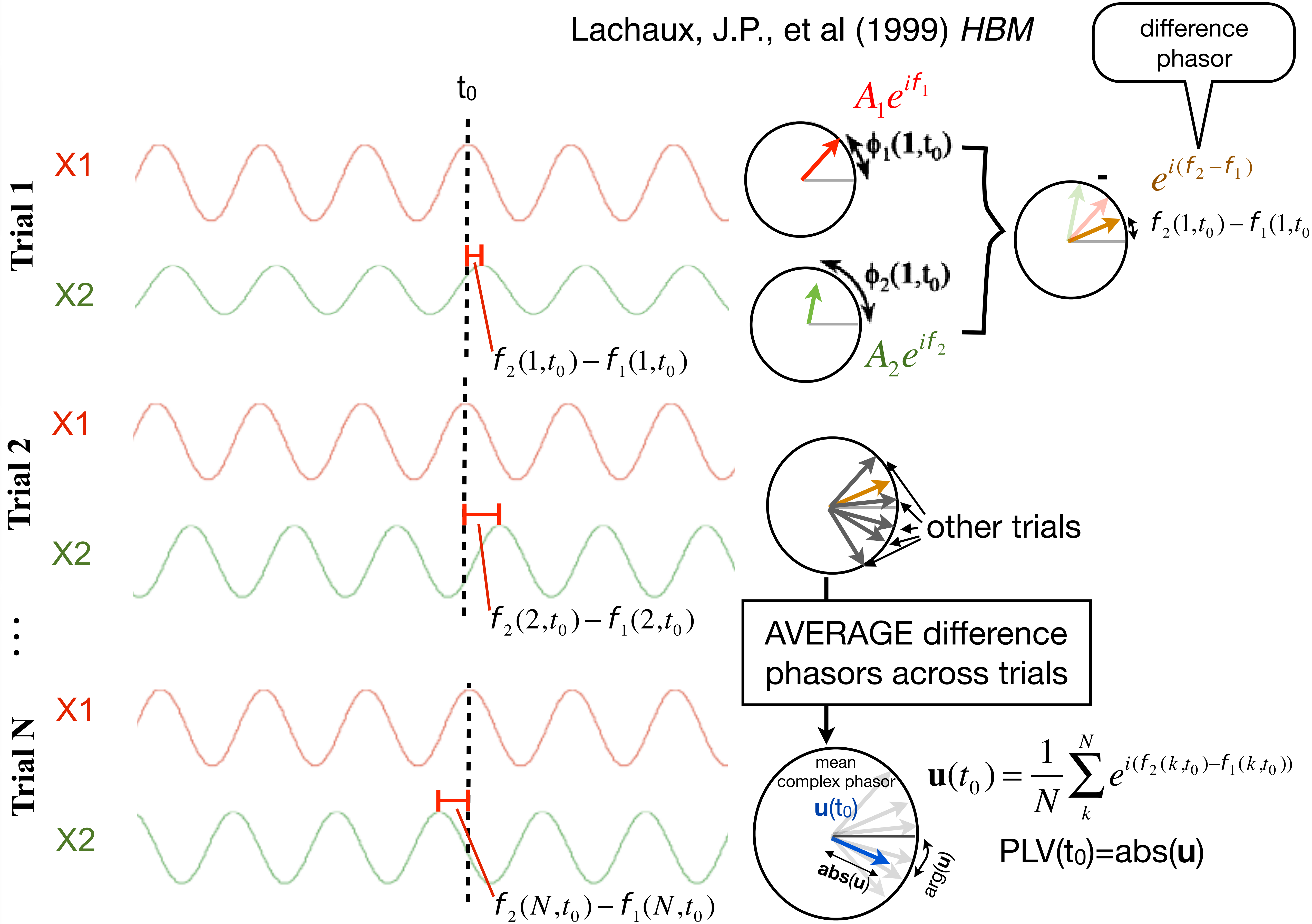
If all **phasors have similar angles**, then vectors will "point" in the same direction and the **length of the mean phasor** will be comparatively **large**.

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)

Tim Mullen

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



Tim Mullen

# Phase-Locking Value (PLV)

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

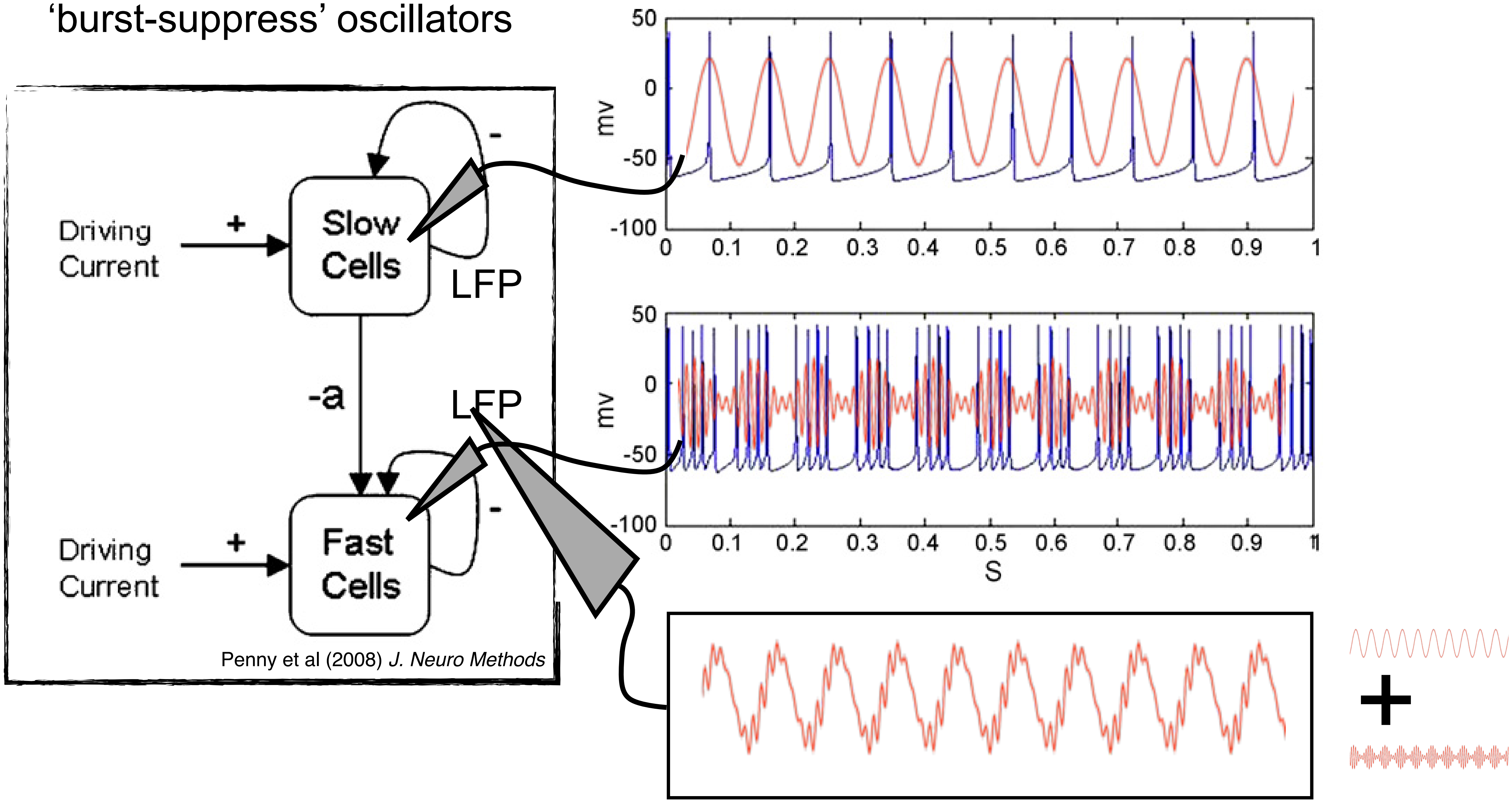
Computing PLV ("phase coherence") in EEGLAB:

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

# Phase-Amplitude Coupling

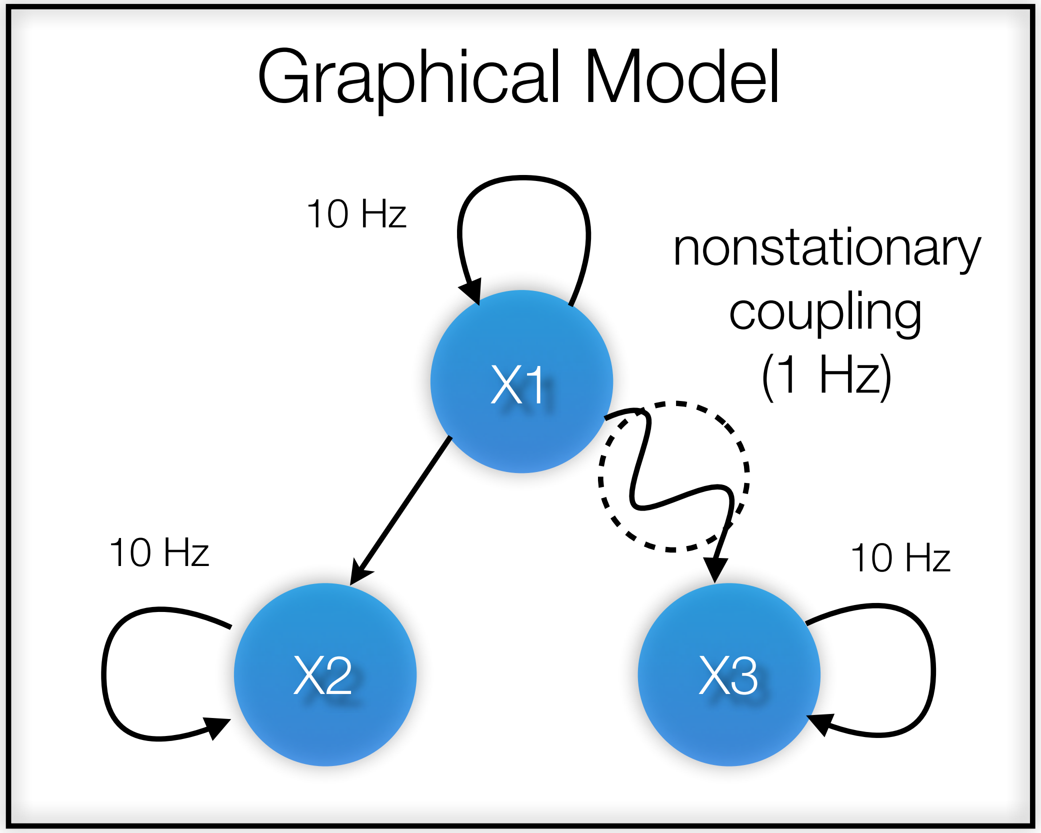
Tim Mullen

'burst-suppress' oscillators

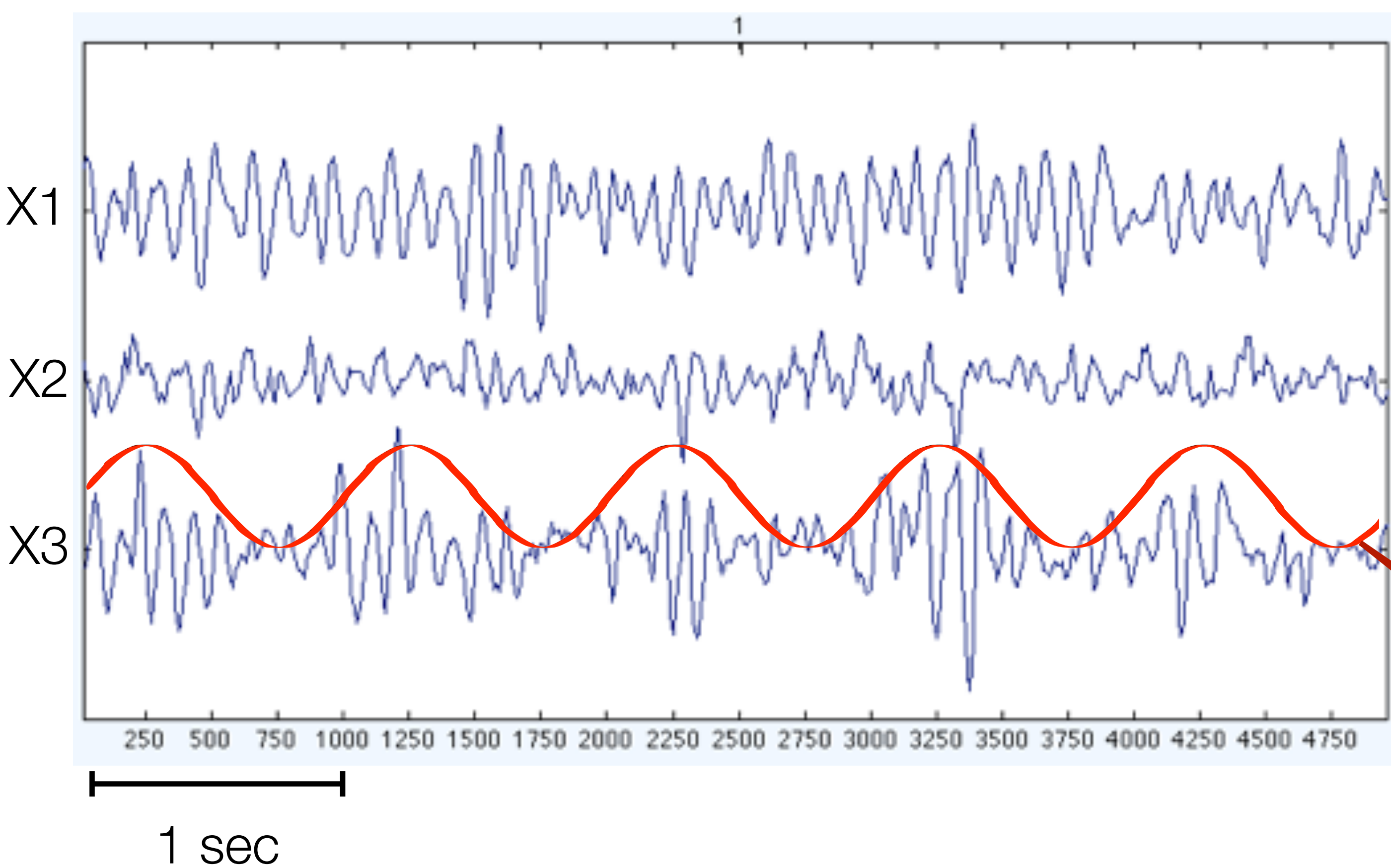



Local Field Potential (Slow + Fast cells)

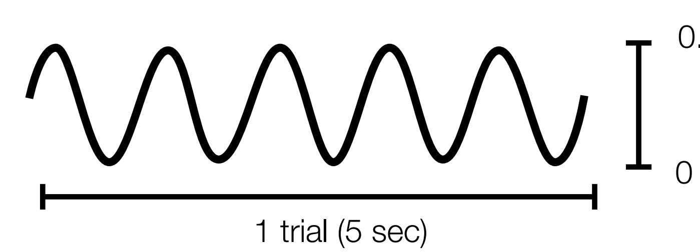
Tim Mullen




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



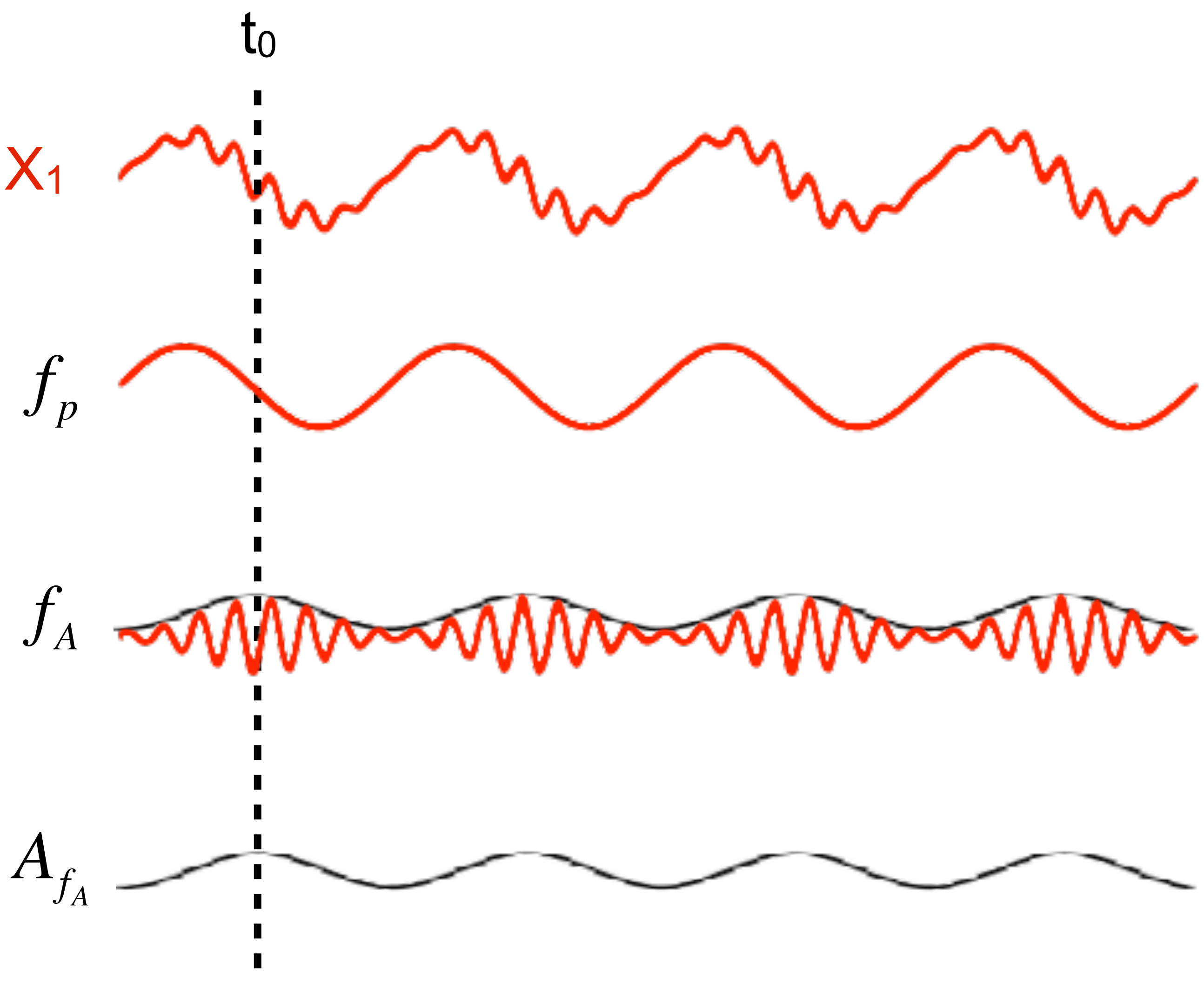
- May present a functional role in execution of cognitive functions (Axmacher et al. 2010; Cohen et al. 2009a,b; Lakatos et al. 2008; Tort et al. 2008, 2009).

- 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)

Tim Mullen

# Phase-Amplitude Coupling: PLV Method

Vanhatalo, S et al (2004) *PNAS*




original raw signal

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

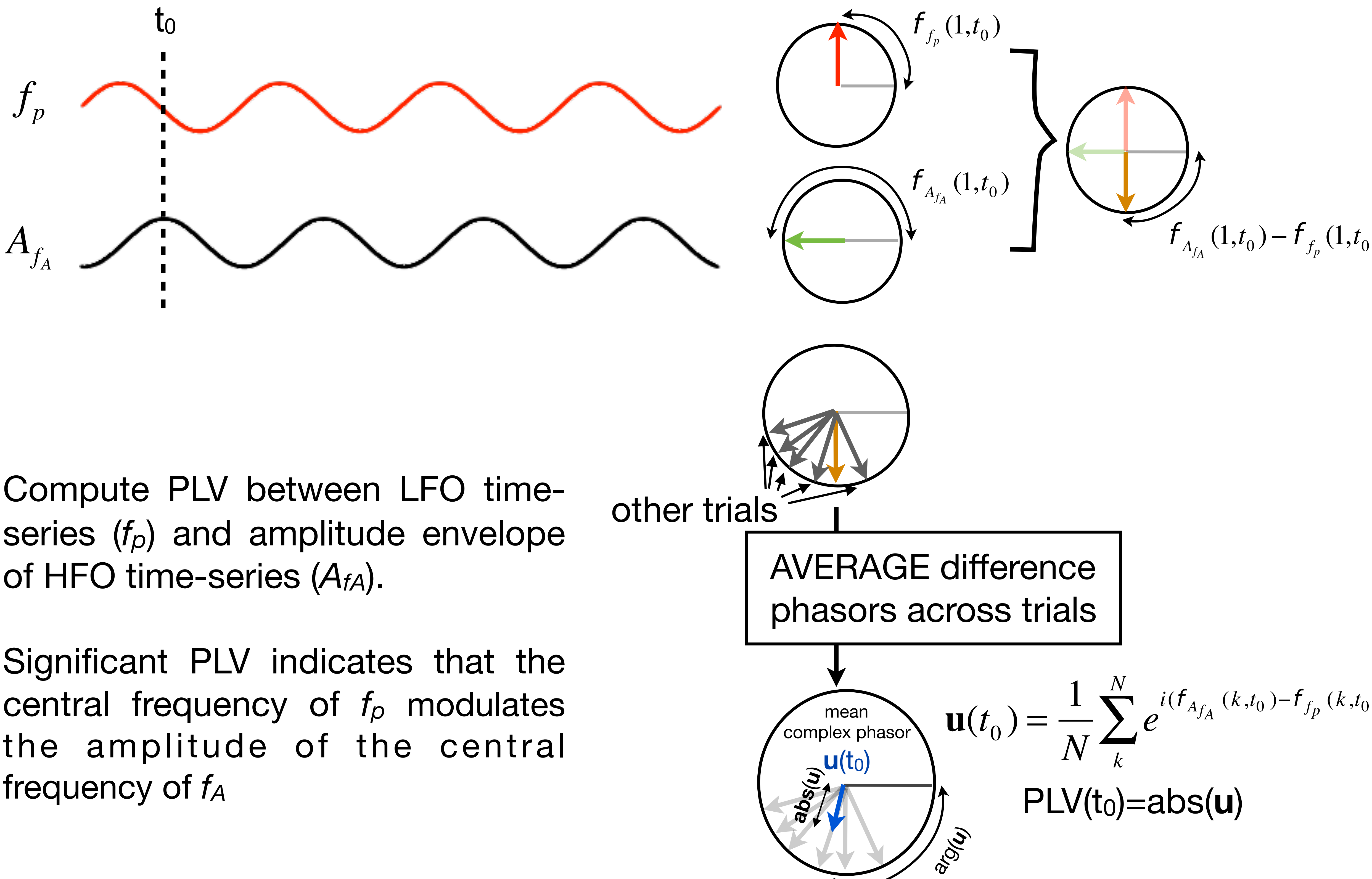
filter $X_1$ at HFO band (e.g. gamma)

get amplitude envelope of filtered signal

Tim Mullen

# Phase-Amplitude Coupling: PLV Method

Vanhatalo, S et al (2004) *PNAS*




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

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

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 $f_A$

Tim Mullen

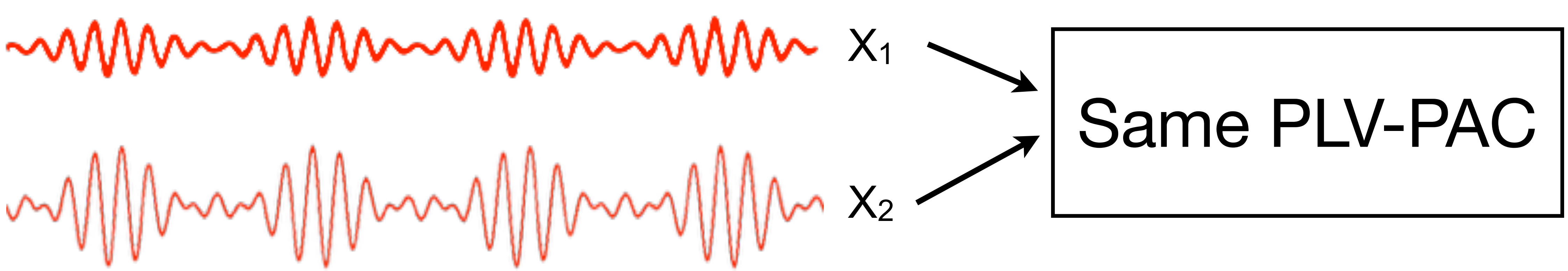
# Phase-Amplitude Coupling: PLV Method

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.

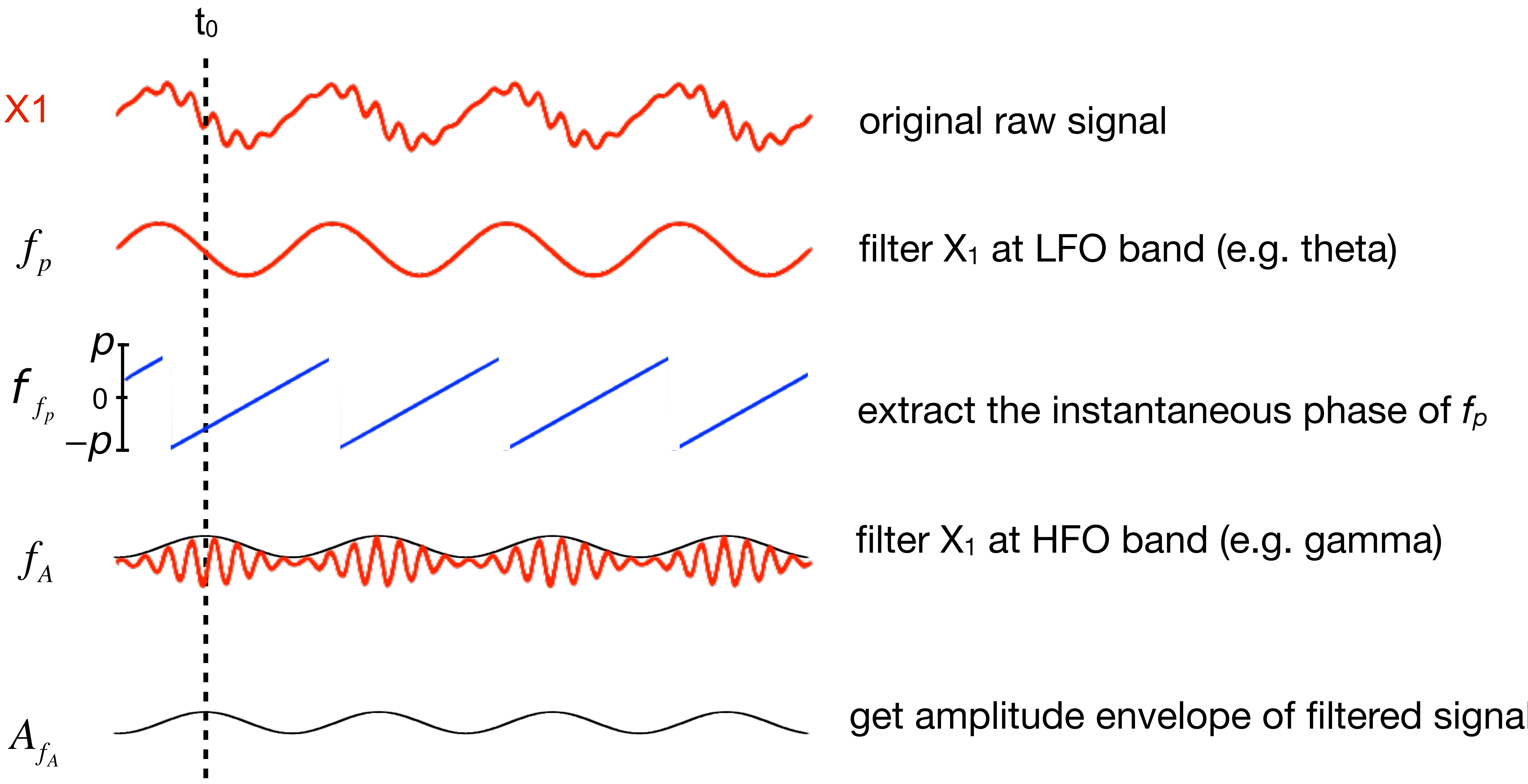
In the example below, $X_1$ and $X_2$ both would produce the same PAC, even though the high-frequency amplitude of $X_2$ clearly is more strongly modulated by the low-frequency rhythm.



# Phase-Amplitude Coupling: Modulation Index Method

Canolty et al, (2006) *Science*

Tim Mullen

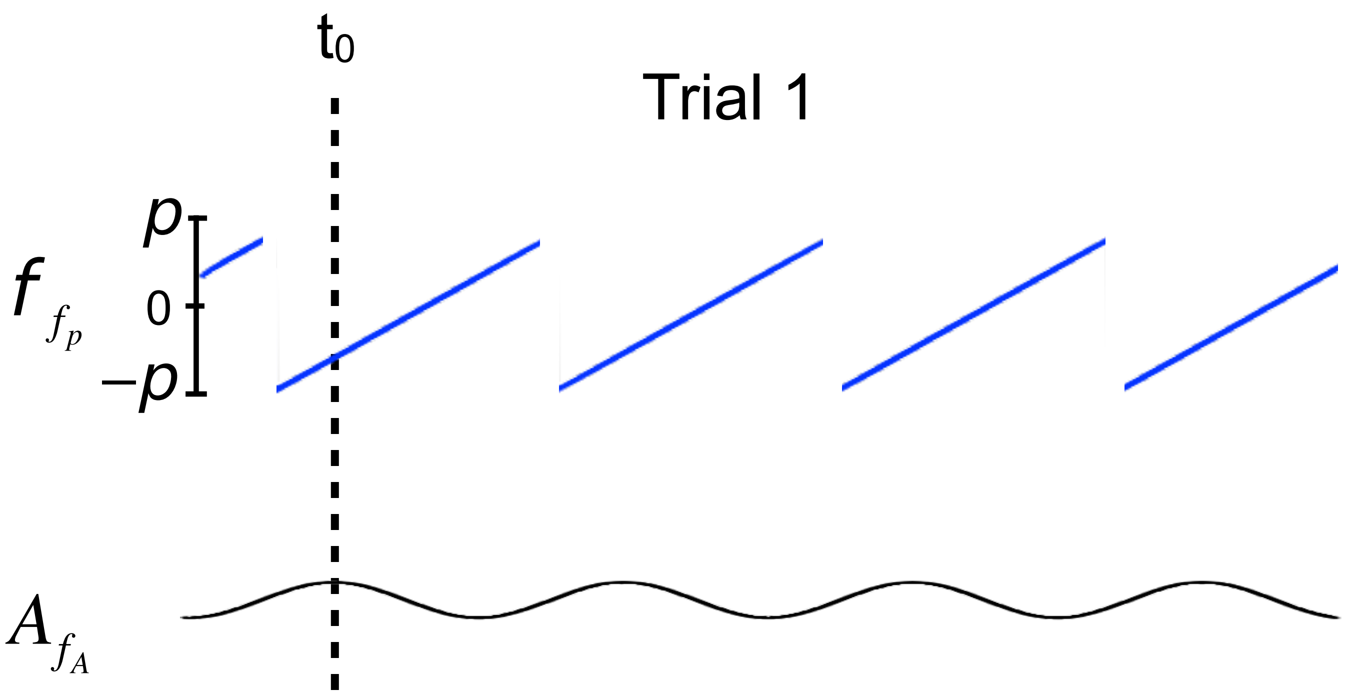


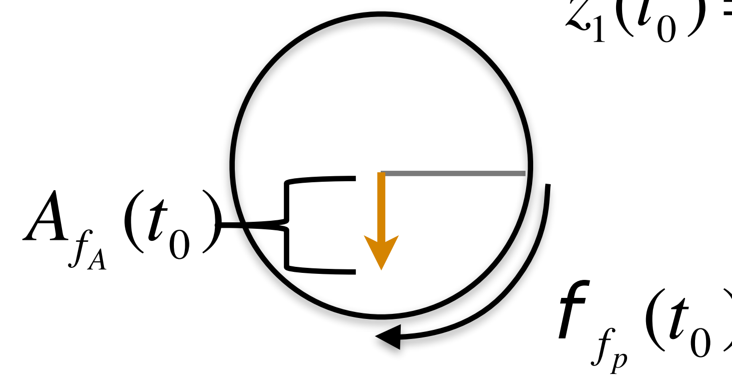
# Phase-Amplitude Coupling: Modulation Index Method

Canolty et al, (2006) *Science*

build complex phasor with instantaneous amplitude and phase

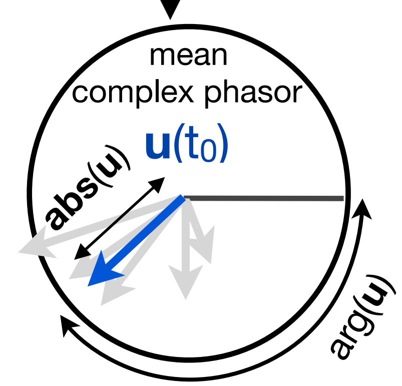
$$z_1(t_0) = A_{f_A} e^{i f_{f_p}}$$





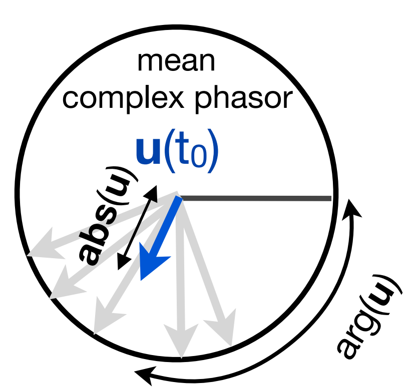
AVERAGE complex phasors across trials




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

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

Comparison:
PLV-PAC



Tim Mullen

# Phase-Amplitude Coupling: Modulation Index Method

Canolty et al, (2006) *Science*

## Computing PAC in EEGLAB:

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

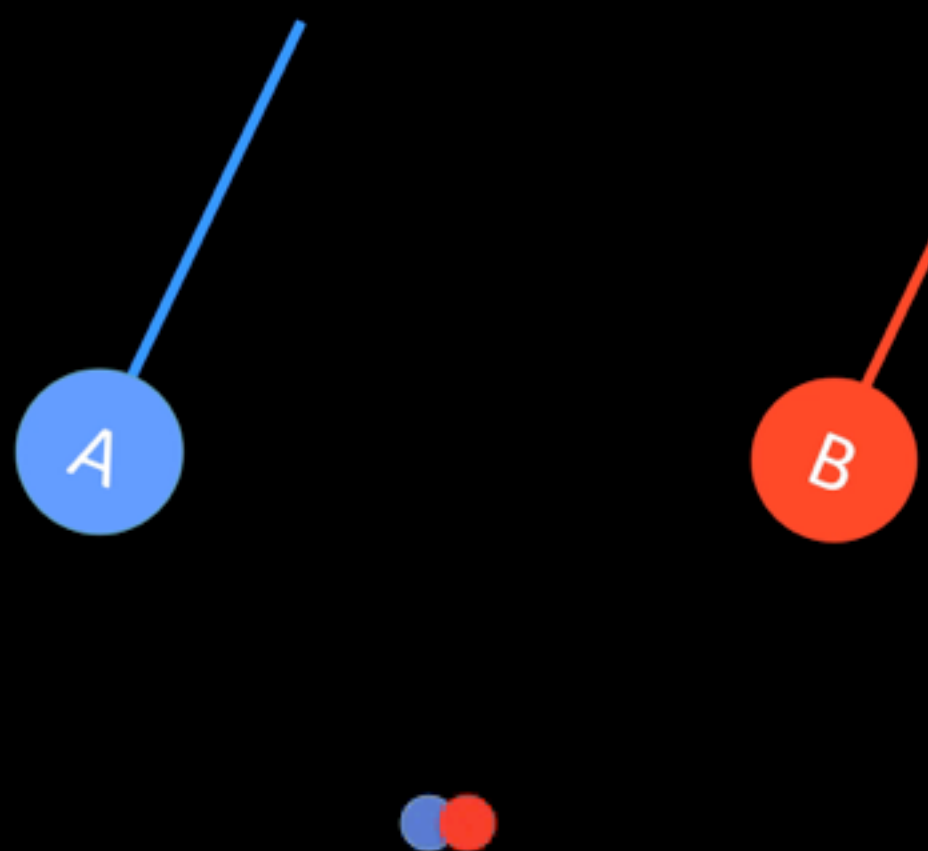
PAC can also be applied *between* sources/channels (e.g. determine whether the phase of oscillation at freq. $w_p$ in IC1 modulates the amplitude of oscillation at freq. $w_A$ *in IC2.* This leads to a measure of cross-frequency (non-linear) functional connectivity.

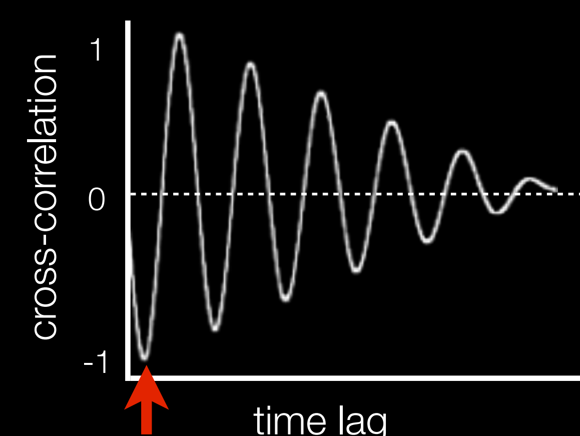
For Modulation Index method (other modes also available)

Also see PACT plugin for EEGLAB by Miyakoshi et al (http://sccn.ucsd.edu/wiki/PACT)

Tim Mullen

# (Cross)-Correlation ≠ Causation




$r_{AB}(k)$

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

Tim Mullen

# Estimating Effective Connectivity

## Non-Invasive

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

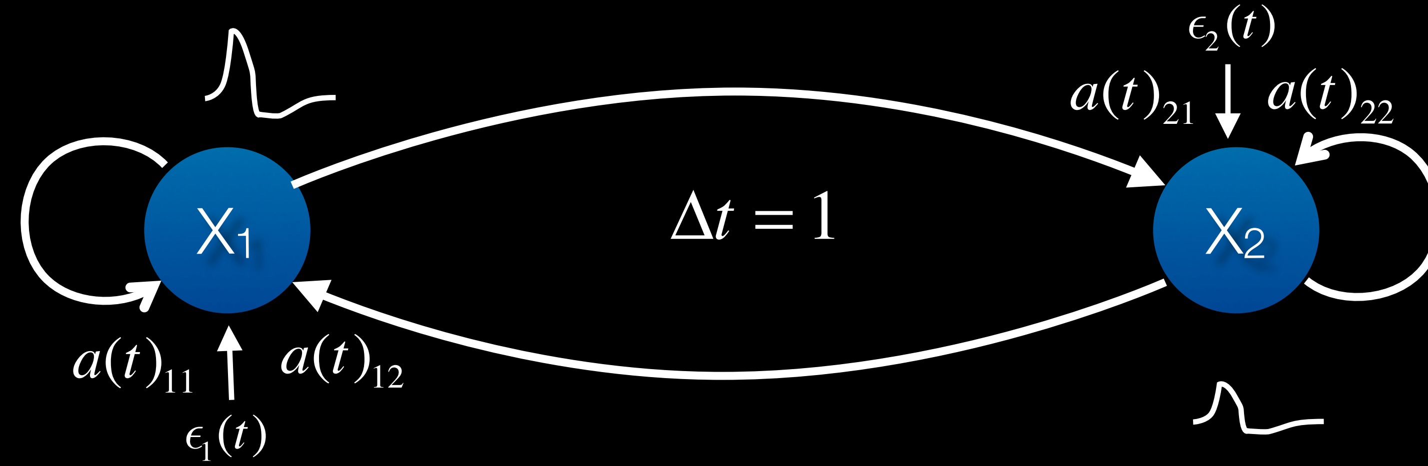
- Data-driven
- Rooted in *conditional predictability*
- Scalable (Valdes-Sosa, 2005)
- Extendable to nonlinear and/or non-stationary systems (Freiwald, 1999; Ding, 2001; Chen, 2004; Ge, 2009)
- Extendable to non-parametric representations (Dhamala, 2009a,b)
- Can be (partially) controlled for (unobserved)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

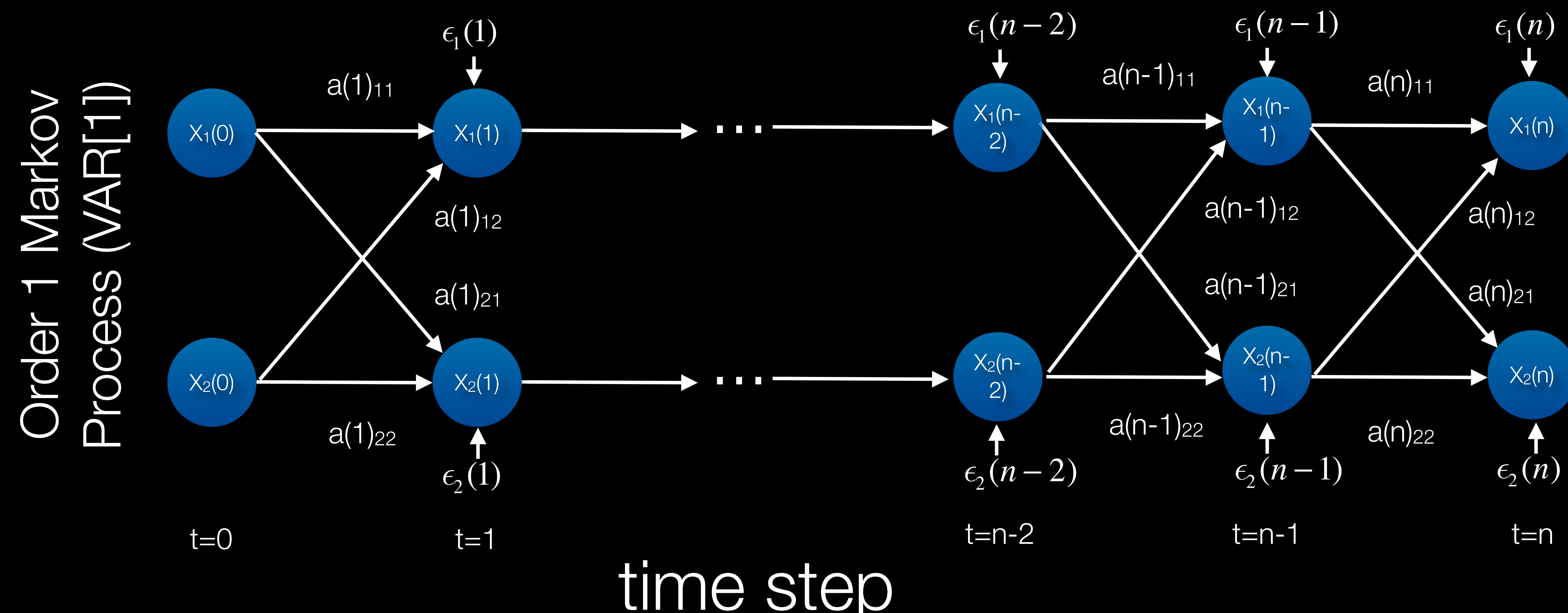
Tim Mullen

# Linear Dynamical Systems

## Stochastic Linear Dynamical System

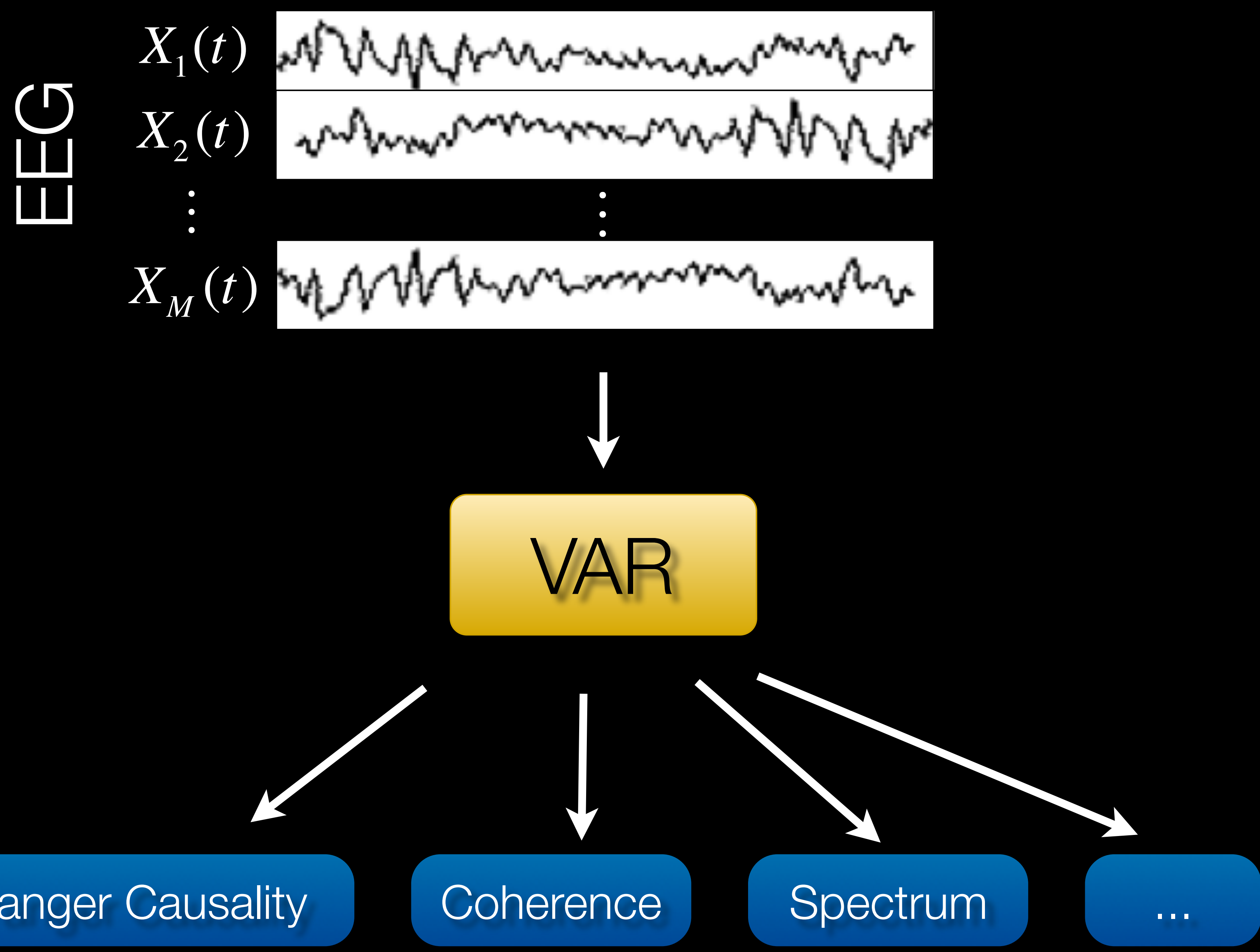
$$X_1(t) = a(t)_{11} X_1(t-1) + a(t)_{12} X_2(t-1) + \epsilon_1(t)$$
$$X_2(t) = a(t)_{22} X_2(t-1) + a(t)_{21} X_1(t-1) + \epsilon_2(t)$$



Tim Mullen

# Vector Autoregressive (VAR / MAR / MVAR) Modeling



Tim Mullen

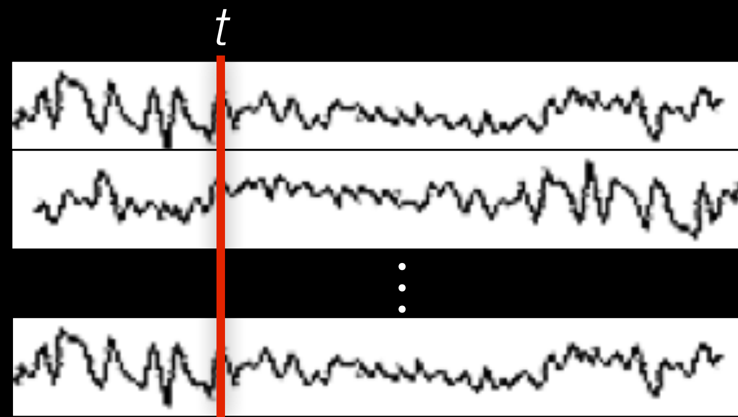
# VAR Modeling: Assumptions

- **"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 $\leq 1$
  - 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

Tim Mullen

# 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

model order

random noise process

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

M-channel data vector at current time *t*

M x M matrix of (possibly time-varying) model coefficients indicating variable dependencies at lag *k*

multichannel data *k* samples in the past

$$\mathbf{A}^{(k)}(t) = \begin{pmatrix} a_{11}^{(k)}(t) & \dots & a_{1M}^{(k)}(t) \\ \vdots & \ddots & \vdots \\ a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t) \end{pmatrix} \qquad \mathbf{E}(t) = N(0, \mathbf{V})$$

Tim Mullen

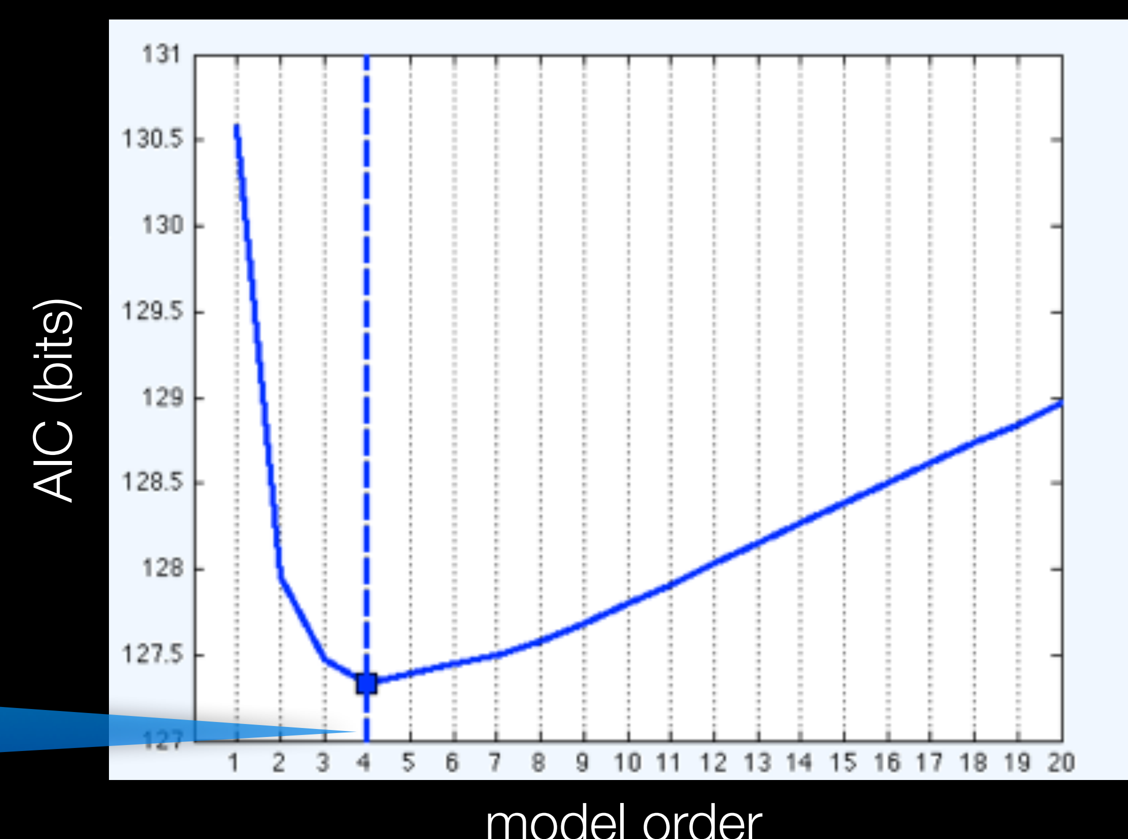
# Selecting a VAR Model Order

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

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

Penalizes high model orders (parsimony)

entropy rate (amount of prediction error)




optimal order

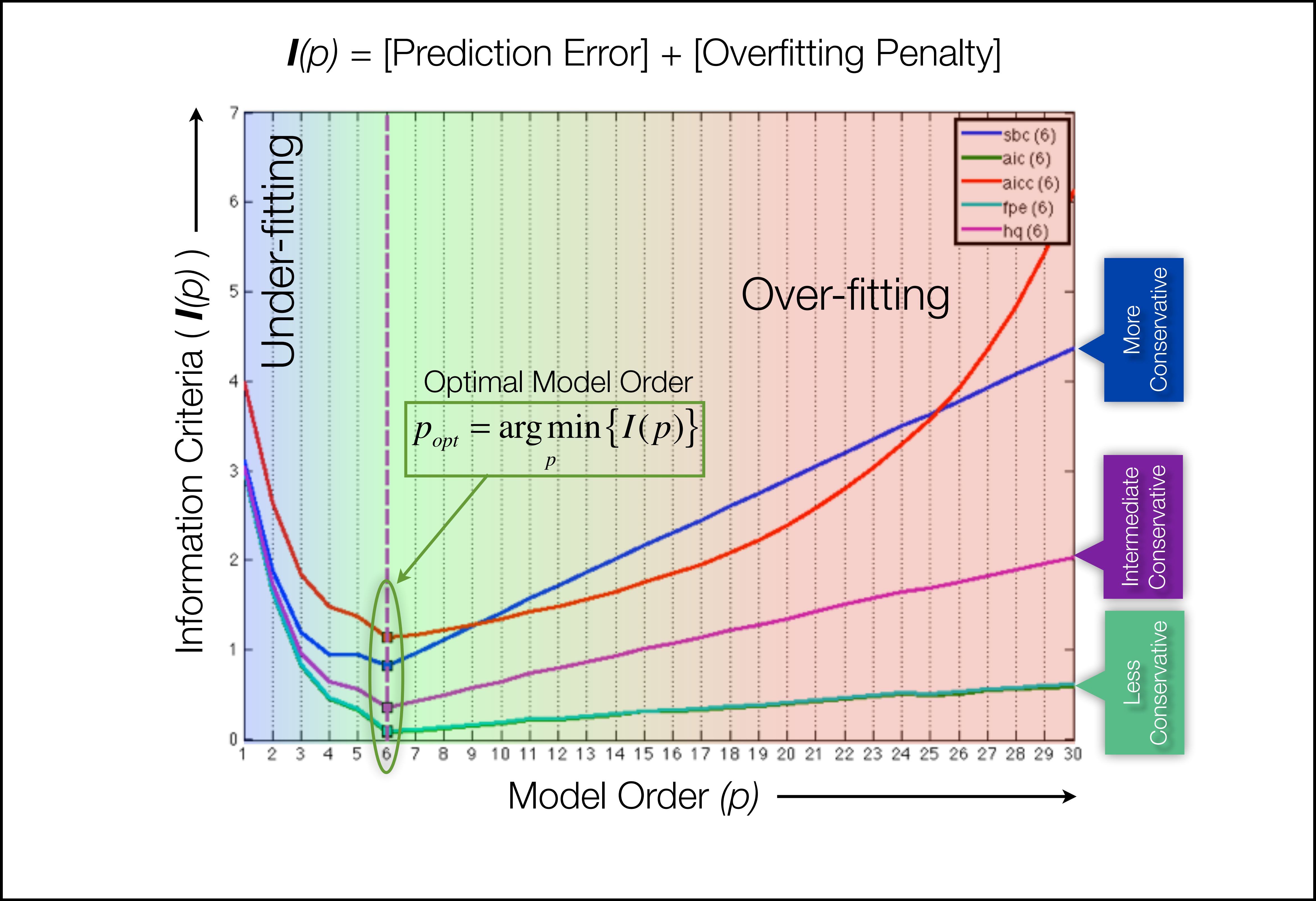
Swartz Center for Computational Neuroscience
Tim Mullen

# Model Order Selection Criteria

| | Estimator | Formula |
|---|---|---|
| More Conservative | **Schwarz-Bayes Criterion** (Bayesian Information Criterion) | $SBC(p) = ln\left\|\tilde{\Sigma}(p)\right\| + \frac{ln(\hat{T})}{\hat{T}}pM^2$ |
| | **Akaike Information Criterion** | $AIC(p) = ln\left\|\tilde{\Sigma}(p)\right\| + \frac{2}{\hat{T}}pM^2$ |
| Less Conservative | **Akaike's Final Prediction Error** | $FPE(p) = \left\|\tilde{\Sigma}(p)\right\| + \left(\frac{\hat{T}+Mp+1}{\hat{T}-Mp-1}\right)^M$ and its logarithm (used in SIFT) $ln(FPE(p)) = ln\left\|\tilde{\Sigma}(p)\right\| + M ln\left(\frac{\hat{T}+Mp+1}{\hat{T}-Mp-1}\right)$ |
| Intermediate Conservative | **Hannan-Quinn Criterion** | $HQ(p) = ln\left\|\tilde{\Sigma}(p)\right\| + \frac{2ln(ln(\hat{T}))}{\hat{T}}pM^2$ |

Tim Mullen

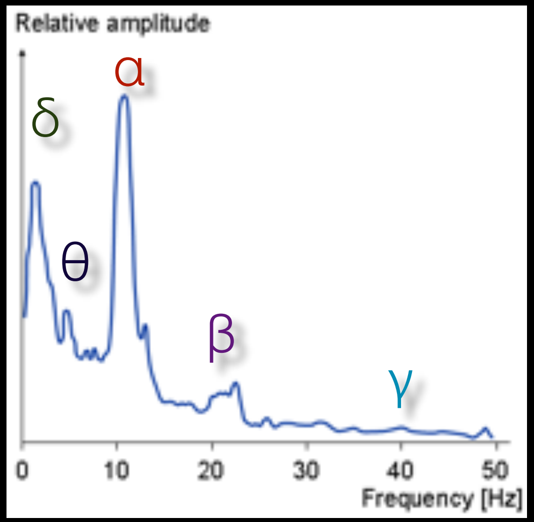
# Model Order Selection Criteria



Tim Mullen

# 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.

Tim Mullen

# 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.

We'll discuss these further in Part 2 (Sunday a.m.)

Tim Mullen

# 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) |

Tim Mullen

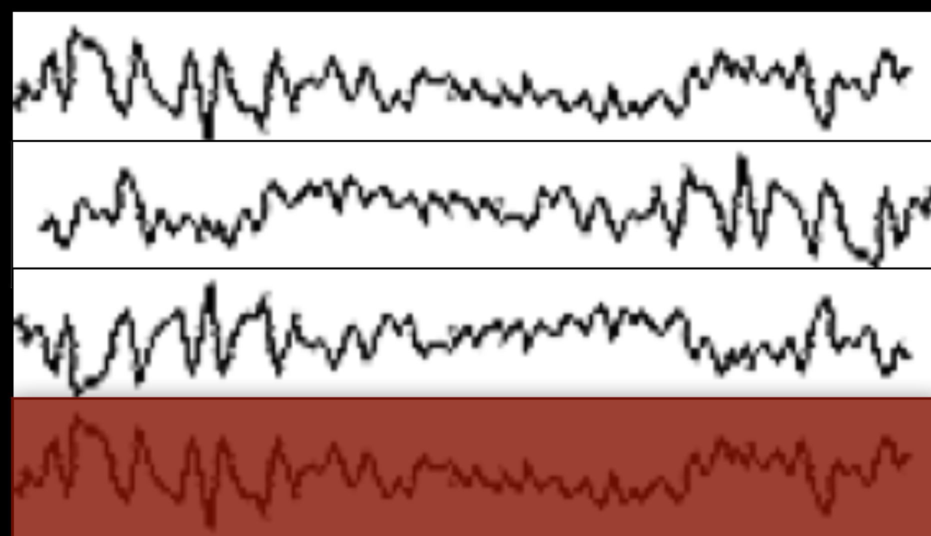
# Granger Causality

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 $E_1$)

?

$X_1(t)$
$X_2(t)$
$X_3(t)$
$X_4(t)$

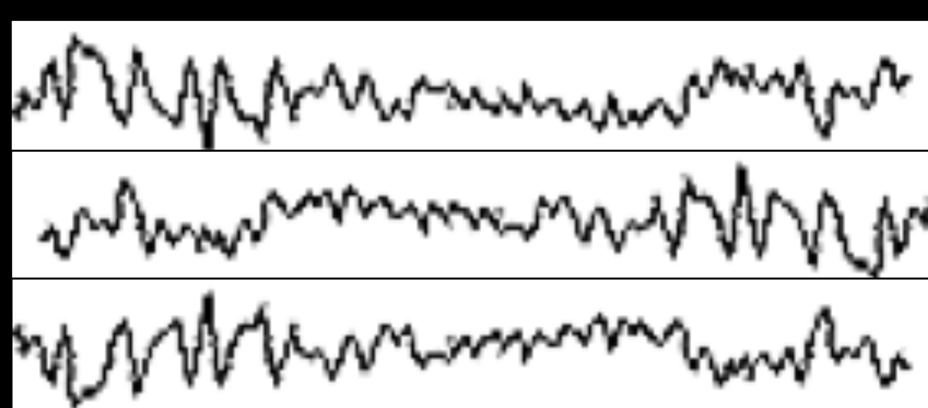


VAR$_1$ → var($\mathbf{E}_1$(t))

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

= ?

$X_1(t)$
$X_2(t)$
$X_3(t)$



VAR$_2$ → var($\tilde{\mathbf{E}}_1$(t))

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

Tim Mullen

# 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* $\mathbf{X}$

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

Tim Mullen

# Granger Causality Quiz

- Example: 2-channel VAR process of order 1

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

$$X_1(t) = -0.5X_1(t-1) + \boxed{0X_2(t-1)} + E_1(t)$$

$$X_2(t) = \boxed{0.7X_1(t-1)} + 0.2X_2(t-1) + E_2(t)$$

Which causal structure does this model correspond to?

a) 1 → 2 b) 1 ← 2 c) 1 ↔ 2

Tim Mullen

# 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:

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

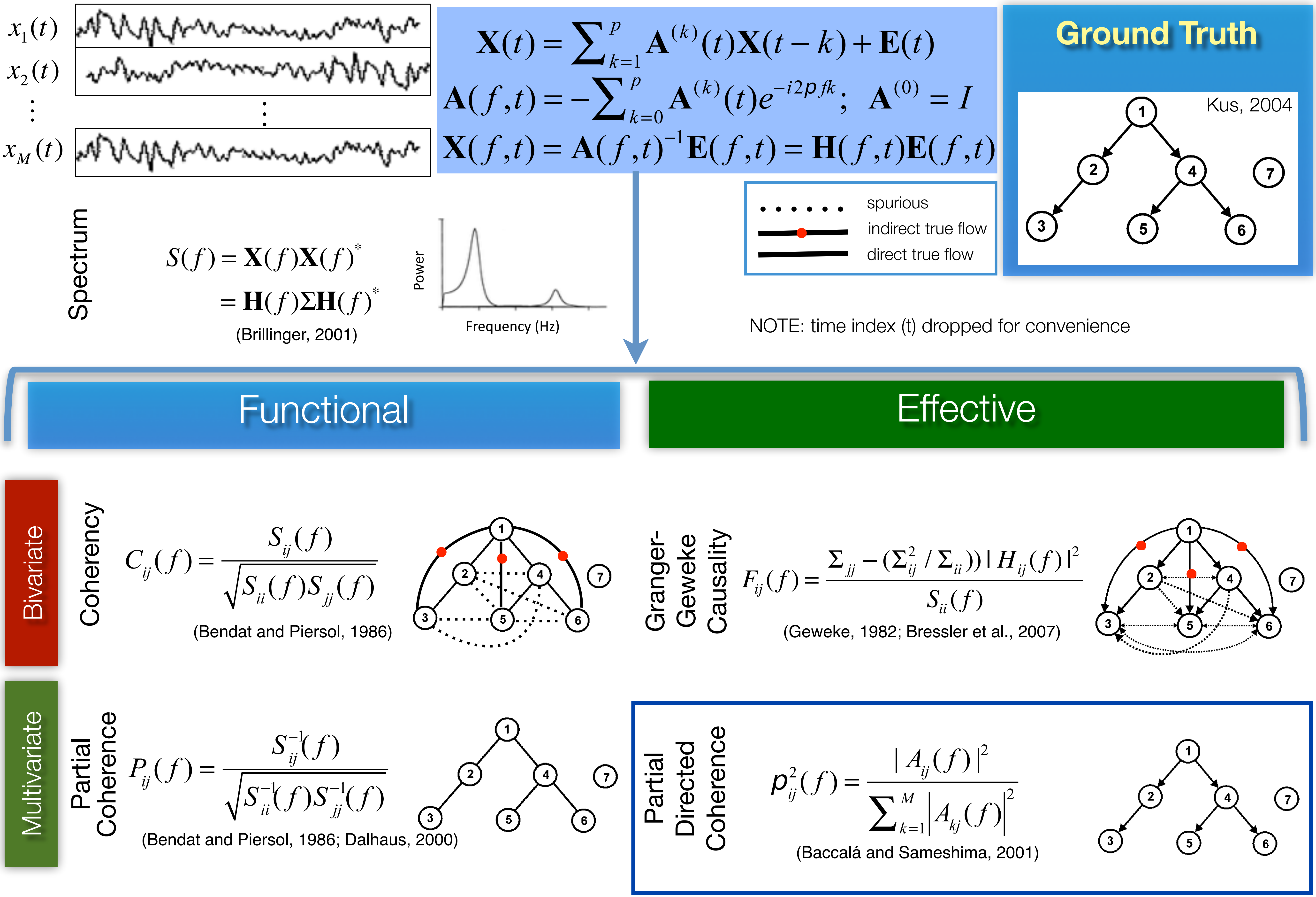
Where **H**(*f*) is the *transfer matrix* of the system.

## 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
$p_{ij}^2(f) = \frac{|A_{ij}(f)|^2}{\sum_{k=1}^{M}|A_{kj}(f)|^2}$
(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 | $C_{ij}(f) = \dfrac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$ <br> $0 \le \lvert C_{ij}(f)\rvert^2 \le 1$ |
| Coherence Measures | Imaginary Coherence (iCoh) | $iCoh_{ij}(f) = \mathrm{Im}(C_{ij}(f))$ |
| Coherence Measures | Partial Coherence (pCoh) | $P_{ij}(f) = \dfrac{\hat{S}_{ij}(f)}{\sqrt{\hat{S}_{ii}(f)\hat{S}_{jj}(f)}}$ <br> $\hat{S}(f) = S(f)^{-1}$ <br> $0 \le \lvert P_{ij}(f)\rvert^2 \le 1$ |
| Coherence Measures | Multiple Coherence (mCoh) | $G_i(f) = \sqrt{1 - \dfrac{\det(S(f))}{S_{ii}(f)\mathbf{M}_{ii}(f)}}$ <br> $\mathbf{M}_{ii}(f)$ 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) | $\pi_{ij}(f) = \dfrac{A_{ij}(f)}{\sqrt{\sum_{k=1}^{M}\lvert A_{kj}(f)\rvert^2}}$ <br> $0 \le \lvert \pi_{ij}(f)\rvert^2 \le 1$ <br> $\sum_{j=1}^{M}\lvert \pi_{ij}(f)\rvert^2 = 1$ |
| Partial Directed Coherence Measures | Generalized PDC (GPDC) | $\bar{\pi}_{ij}(f) = \dfrac{\frac{1}{\Sigma_{ii}}A_{ij}(f)}{\sqrt{\sum_{k=1}^{M}\frac{1}{\Sigma_{ii}^2}\lvert A_{kj}(f)\rvert^2}}$ <br> $0 \le \lvert \bar{\pi}_{ij}(f)\rvert^2 \le 1$ <br> $\sum_{j=1}^{M}\lvert \bar{\pi}_{ij}(f)\rvert^2 = 1$ |
| Partial Directed Coherence Measures | Renormalized PDC (rPDC) | $\lambda_{ij}(f) = Q_{ij}(f)^* V_{ij}(f)^{-1} Q_{ij}(f)$ <br> where <br> $Q_{ij}(f) = \begin{pmatrix}\mathrm{Re}[A_{ij}(f)] \\ \mathrm{Im}[A_{ij}(f)]\end{pmatrix}$ and <br> $V_{ij}(f) = \sum_{k,l=1}^{p} R_{jj}^{-1}(k,l)\Sigma_{ii} Z(2\pi f,k,l)$ <br> $Z(\omega,k,l) = \begin{pmatrix}\cos(\omega k)\cos(\omega l) & \cos(\omega k)\sin(\omega l) \\ \sin(\omega k)\cos(\omega l) & \sin(\omega k)\sin(\omega l)\end{pmatrix}$ <br> ***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) | $\gamma_{ij}(f) = \dfrac{H_{ij}(f)}{\sqrt{\sum_{k=1}^{M}\lvert H_{ik}(f)\rvert^2}}$ <br> $0 \le \lvert \gamma_{ij}(f)\rvert^2 \le 1$ <br> $\sum_{j=1}^{M}\lvert \gamma_{ij}(f)\rvert^2 = 1$ |
| 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

$\Sigma$ Noise Covariance Matrix

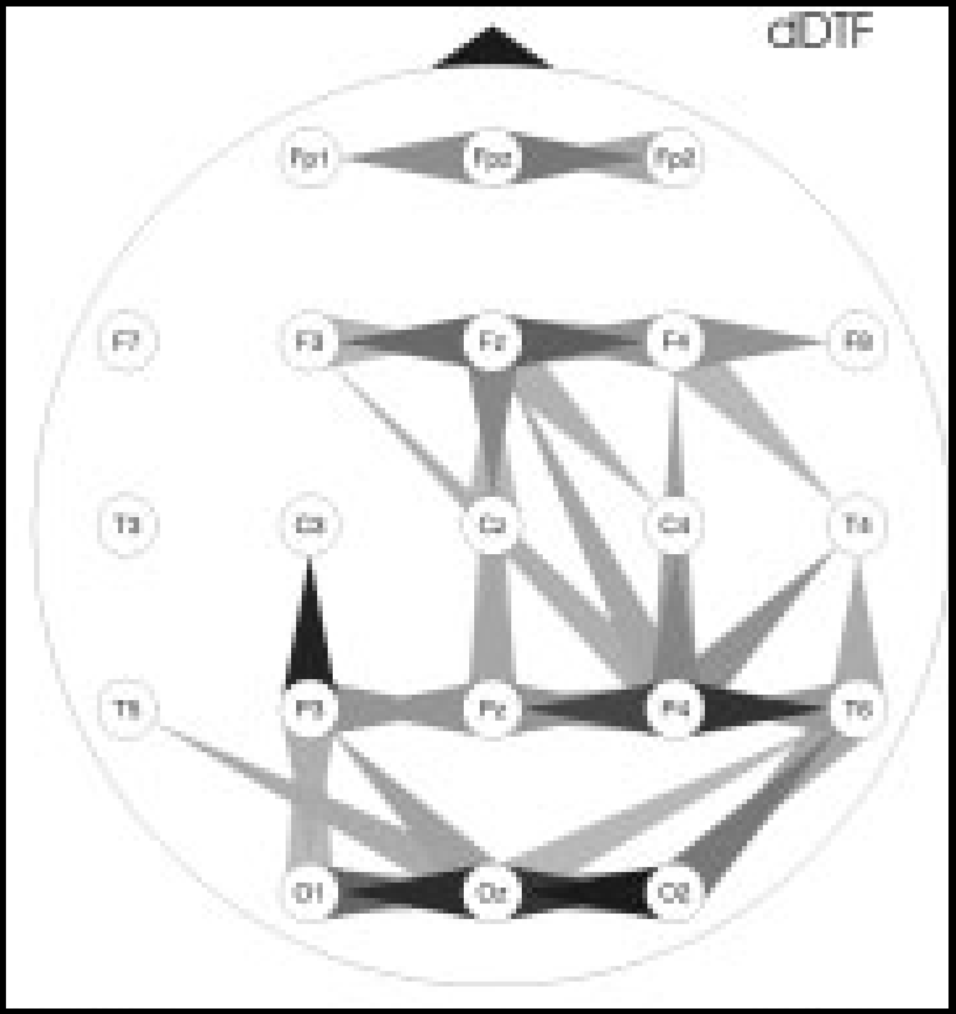
Variance Stabilization

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

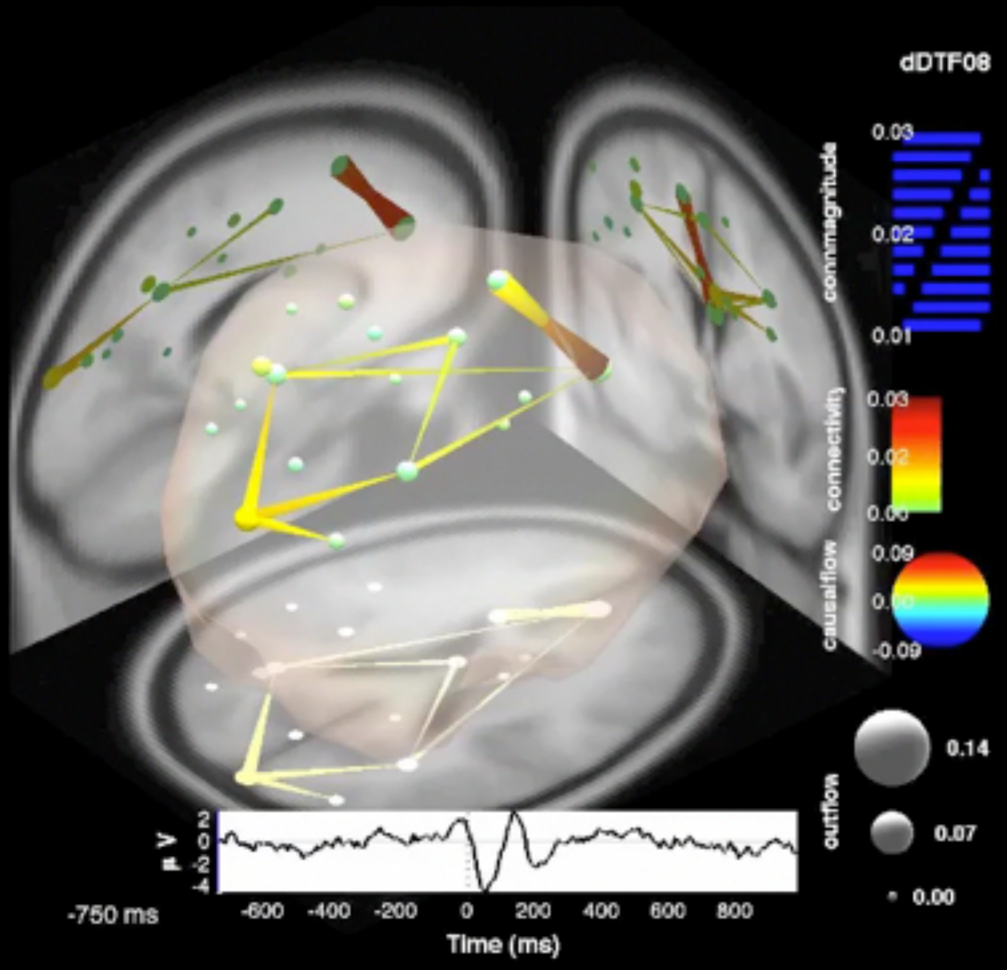
# INTERMISSION

Tim Mullen

# Scalp or Source?



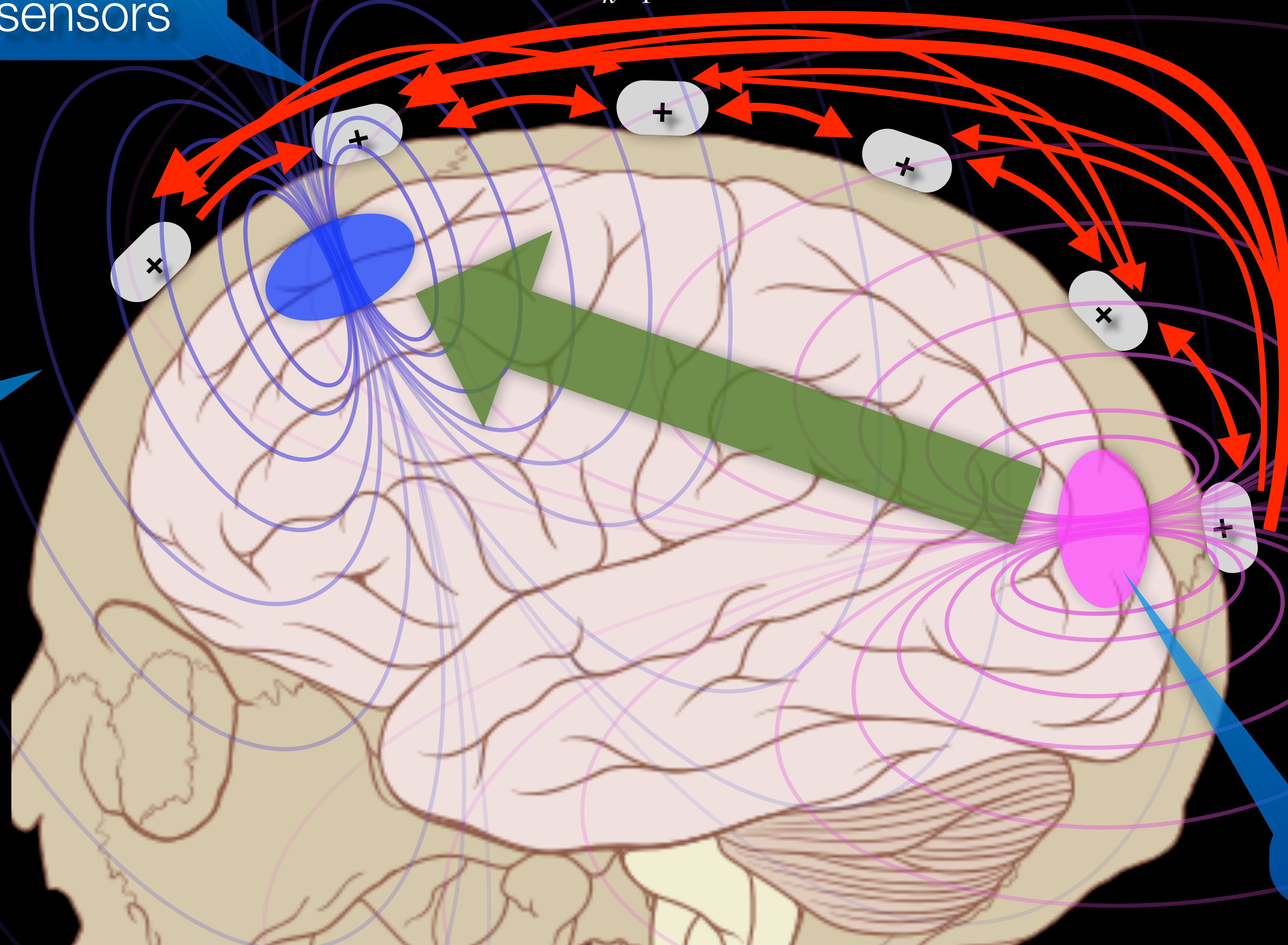

or



Tim Mullen

# Scalp or Source?

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




$$X(t) = \underbrace{H}_{H^{-1}}S(t)$$

Volume Conduction

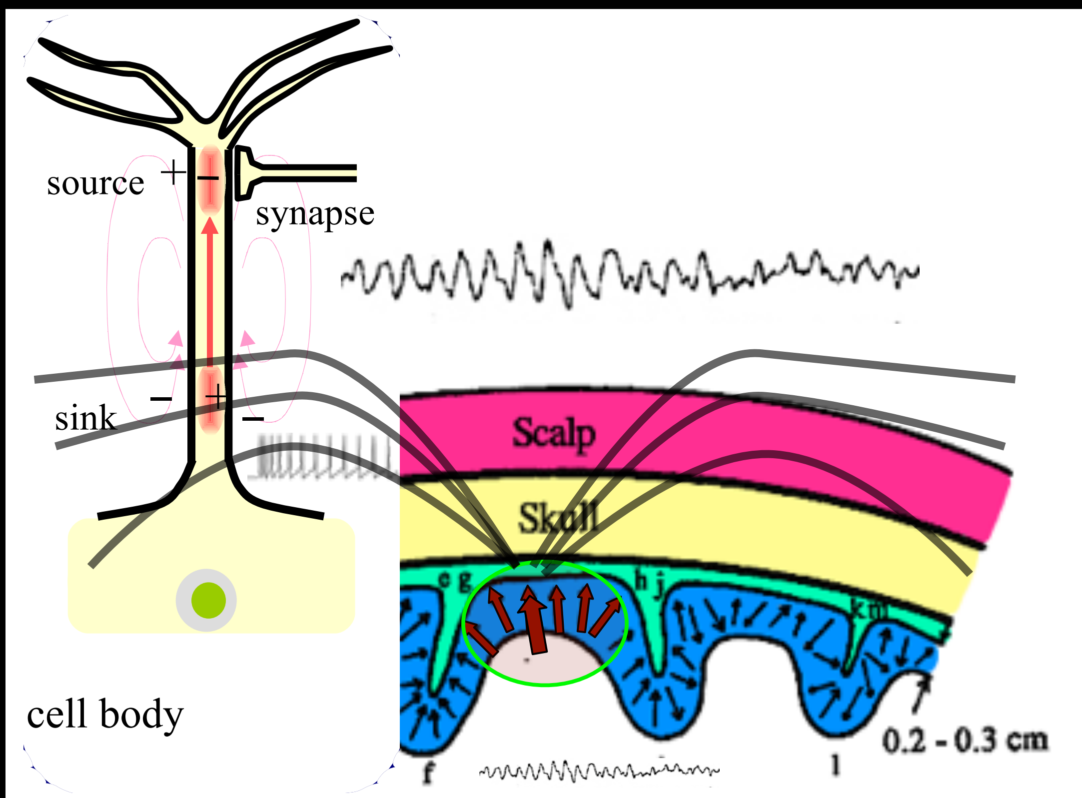
ICA
SBL
Beamforming
Minimum-norm
...

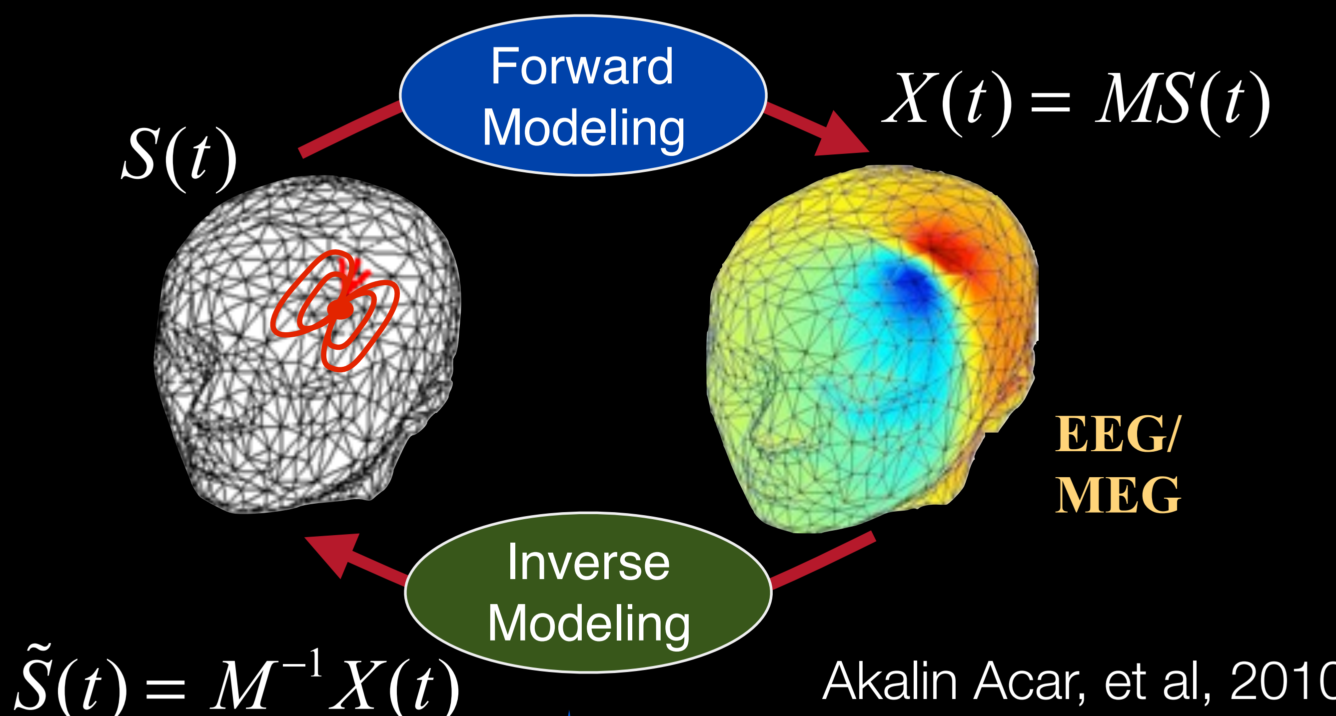
**Solution?** Source Separation

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

Tim Mullen

# Forward/Inverse Modeling



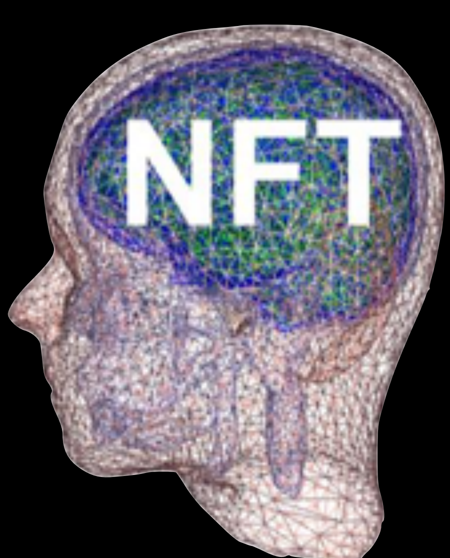




ill-posed!

solutions

sparse/smooth
independence
anatomy
...

impose constraints!




**A Recipe for Reducing Errors:**

- Realistic Forward Model
- Appropriately Constrained Inverse Model

Akalin Acar and Makeig, 2009

Theory

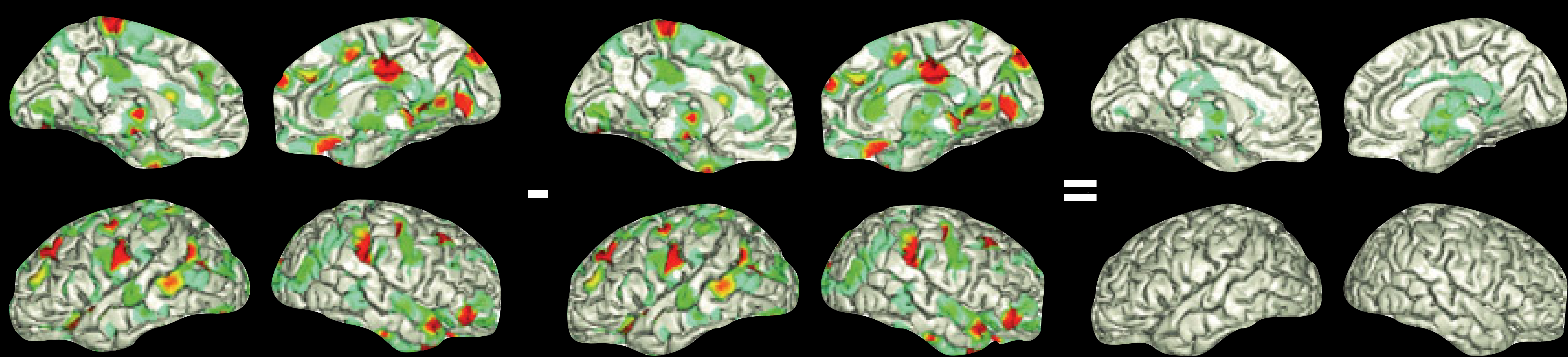
Tim Mullen

# 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

simulated reconstructed error




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

Tim Mullen

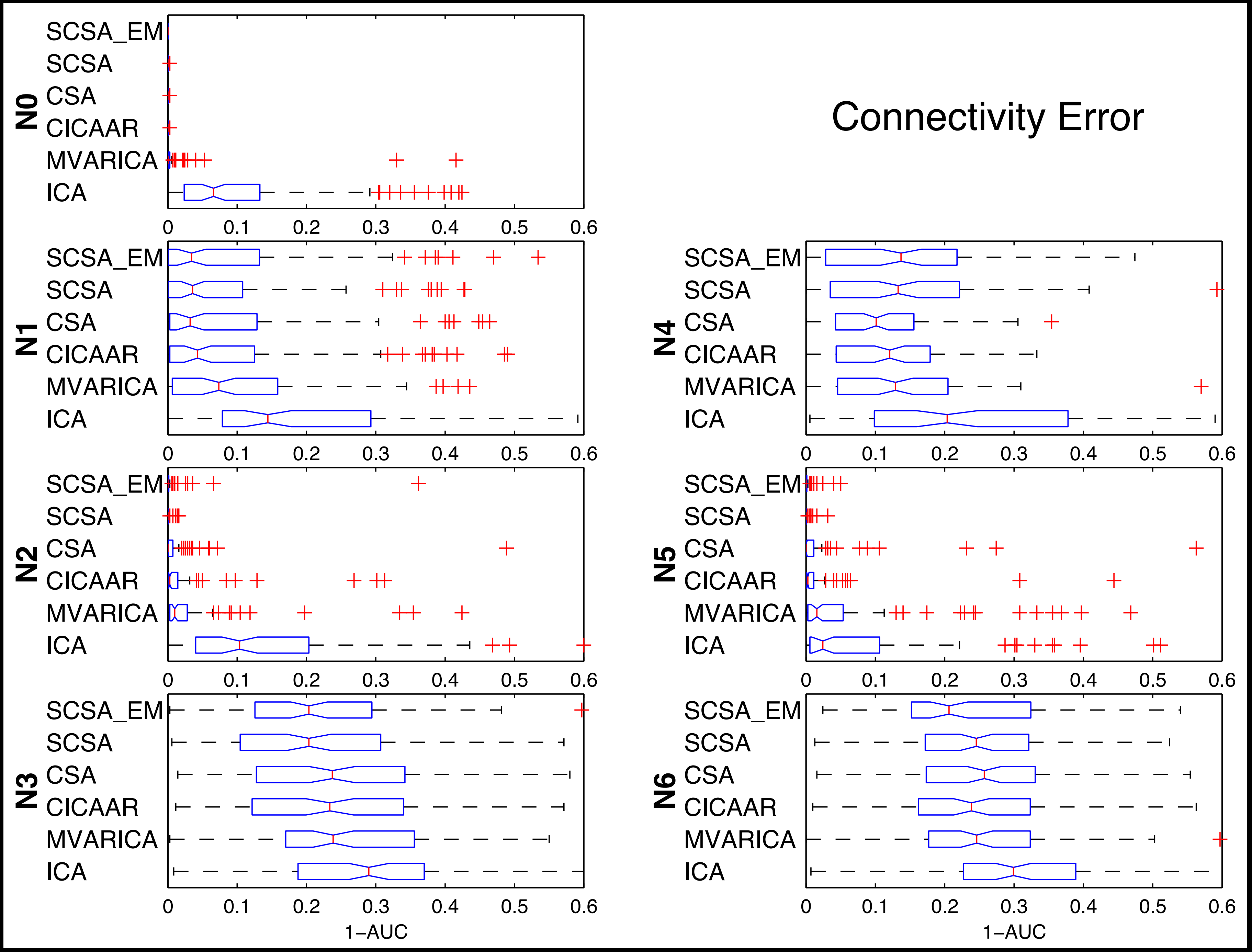
# Estimating Dependency of Independent Components ?

- 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).

Tim Mullen

# Estimating Dependency of Independent Components ?

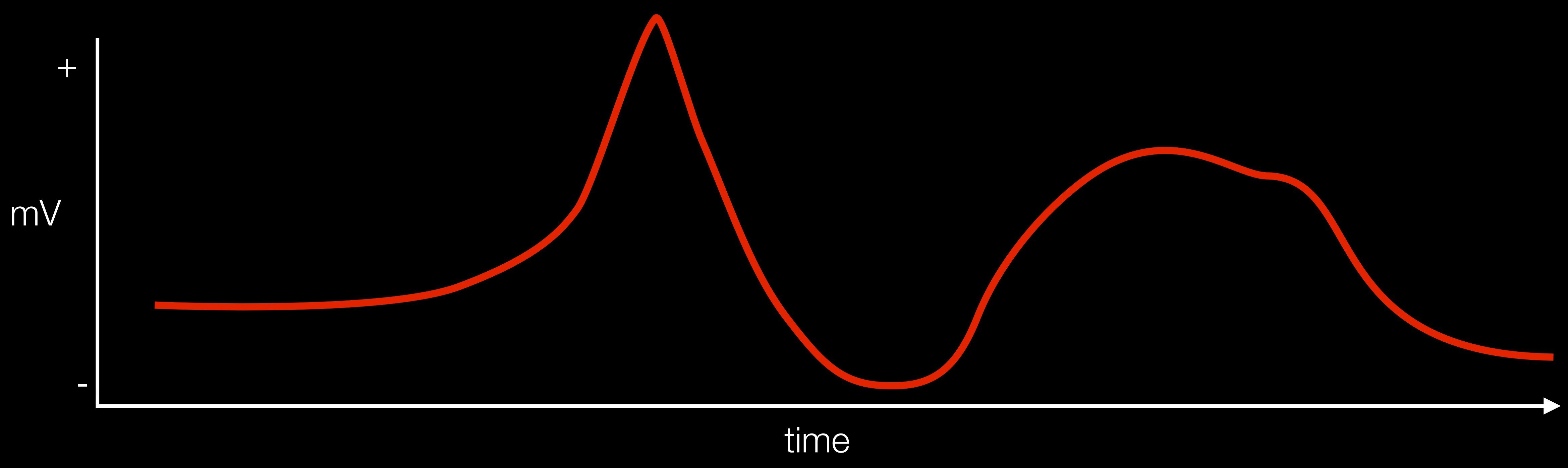



Haufe et al, IEEE TBME 2008

Tim Mullen

# Adapting to Non-Stationarity

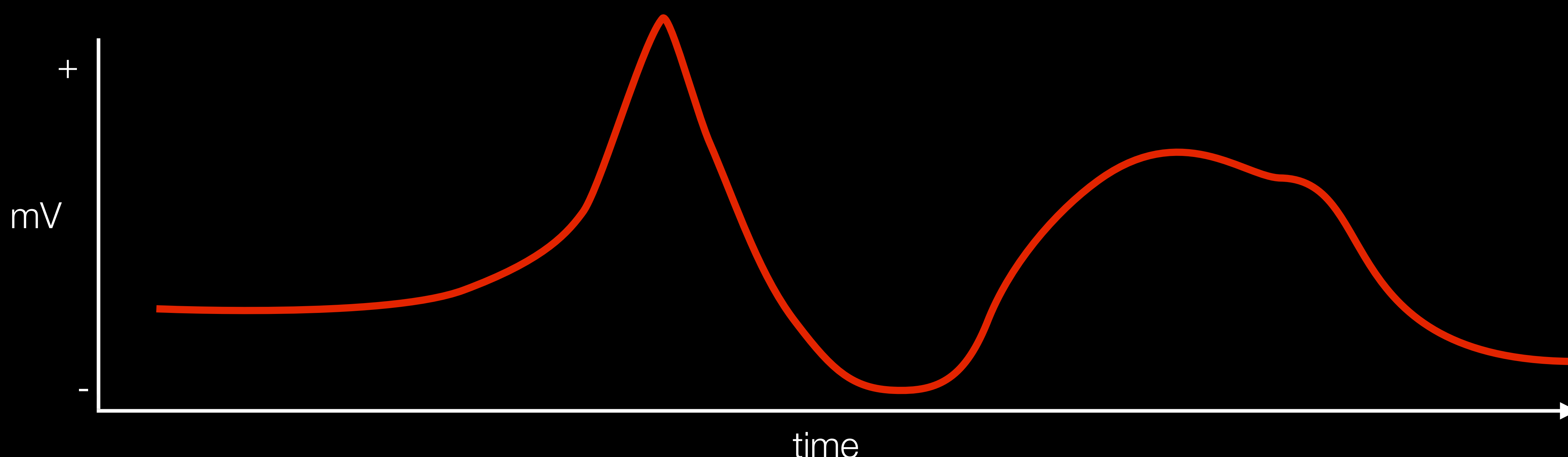
- The brain is a **dynamic system** and measured brain activity and coupling can change rapidly with time (non-stationarity)
  - event-related perturbations (ERSP, ERP, etc)
  - structural changes due to learning/feedback
- How can we adapt to non-stationarity?



Tim Mullen

# Adapting to Non-Stationarity

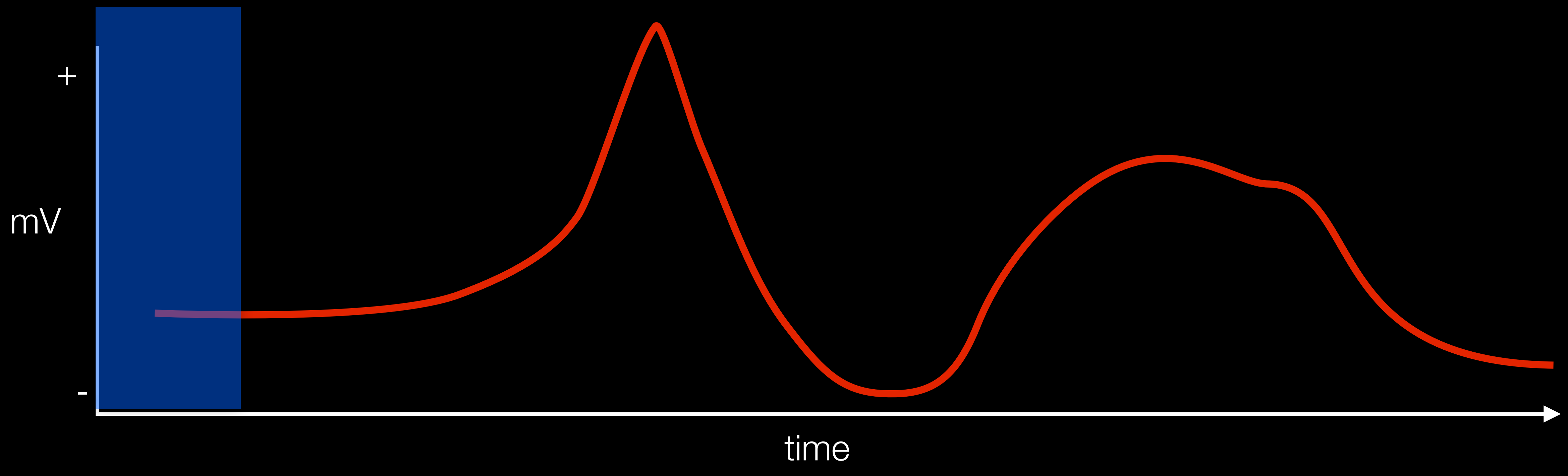
- **Many ways to do adaptive VAR estimation**
- Two popular approaches (adopted in SIFT):
  - Segmentation-based adaptive VAR estimation (assumes local stationarity)
  - State-Space Modeling



Tim Mullen

# 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

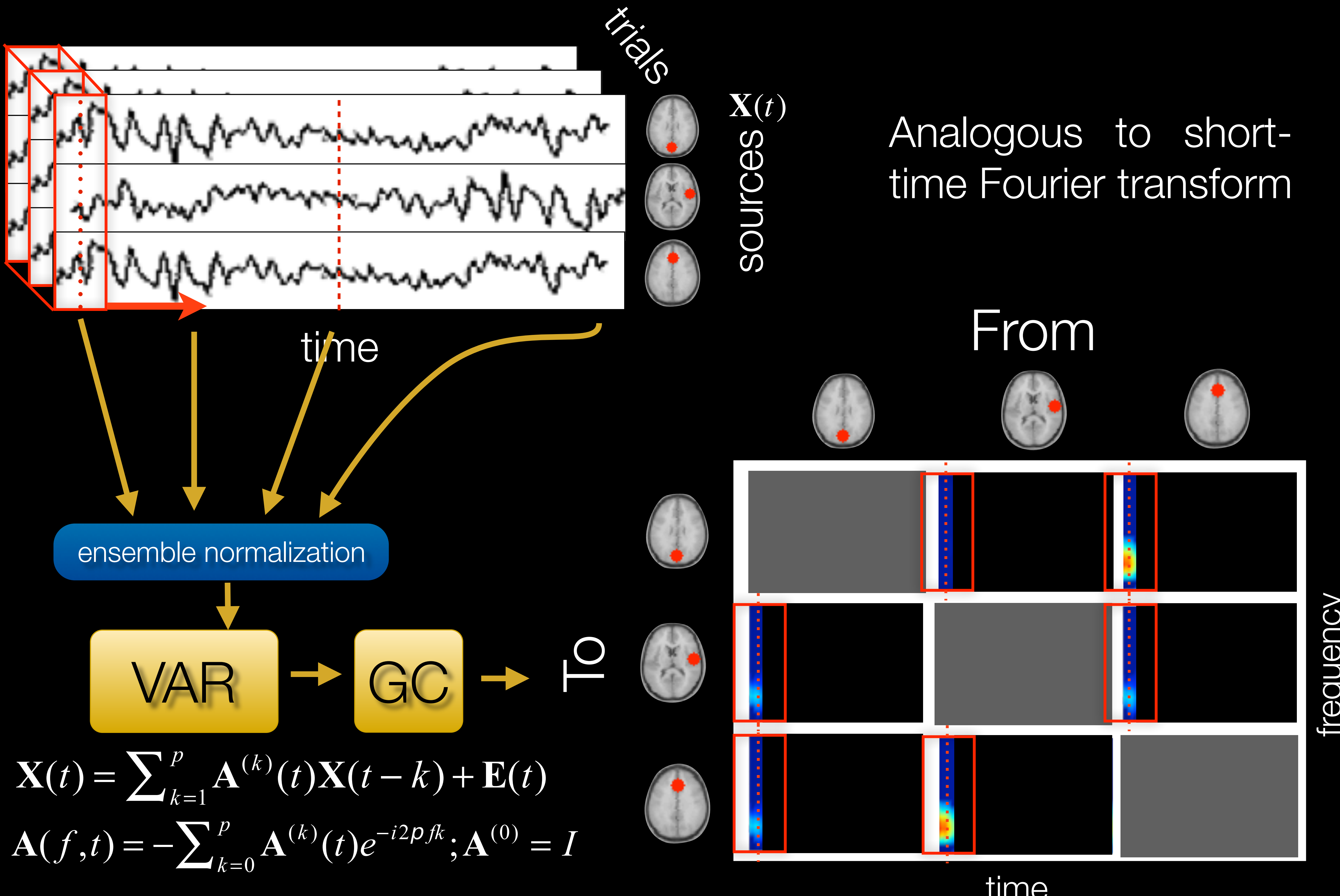


Swartz Center for Computational Neuroscience
Tim Mullen

# Segmentation-based VAR

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

Analogous to short-time Fourier transform




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

$$\mathbf{A}(f,t)=-\sum_{k=0}^{p}\mathbf{A}^{(k)}(t)e^{-i2pfk};\mathbf{A}^{(0)}=I$$

# INTERMISSION