International Conference on
Basic and Clinical Multimodal Imaging
University of Geneva, Geneva, Switzerland,
September 8, 2013

Difficulties and pitfalls in estimating brain connectivity in source space

Kensuke Sekihara[1] and Srikantan S. Nagarajan[2]

[1]Department of Systems Design & Engineering
Tokyo Metropolitan University

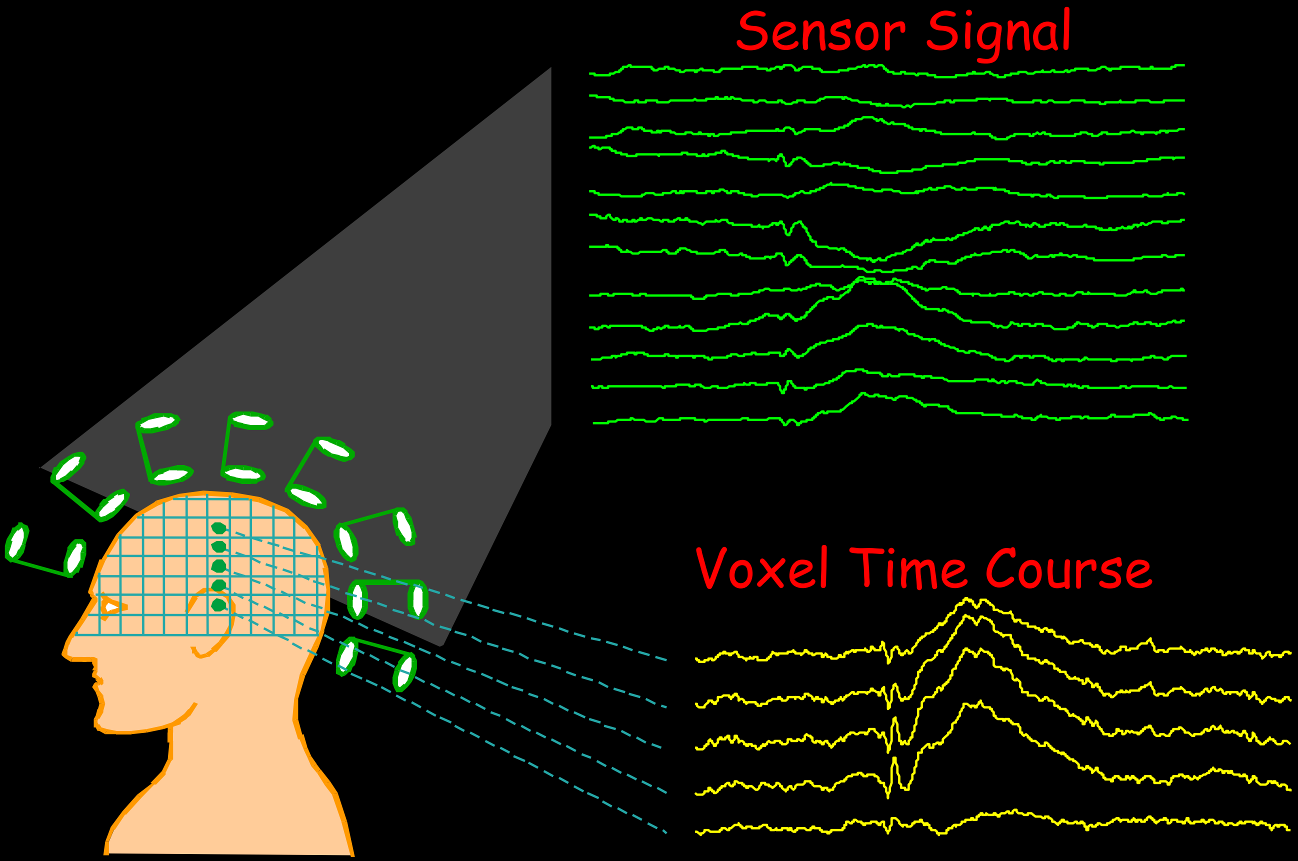
[2]Biomagnetic Imaging Laboratory,
University of California, San Francisco

In my talk today,

- Influence of imaging algorithm’s leakage on the source-space phase analysis.

- Problems in estimating causal relationships using Granger-causality-based measures.

Source space vs. sensor space



Particularly using MEG, the sensor space analysis does not give accurate spatial information, so the source space analysis is preferable.

However, the source space analysis has its own problems, primarily due to the leakage of the imaging algorithm.

What is "leakage" of imaging algorithm

Estimated source time course at $\boldsymbol{r}$:(sources are located at $\boldsymbol{r}_1,\ \boldsymbol{r}_2,\ldots,\boldsymbol{r}_Q$)

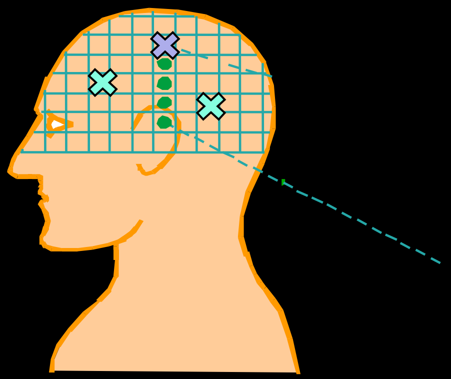
$$\begin{aligned}\hat{s}(t,\boldsymbol{r}) &= \boldsymbol{w}(\boldsymbol{r})^T\boldsymbol{b}(t)\\ &= s_1(t)\boldsymbol{w}(\boldsymbol{r})^T\boldsymbol{l}(\boldsymbol{r}_1)+s_2(t)\boldsymbol{w}(\boldsymbol{r})^T\boldsymbol{l}(\boldsymbol{r}_2)+\cdots+s_Q(t)\boldsymbol{w}(\boldsymbol{r})^T\boldsymbol{l}(\boldsymbol{r}_Q)\\ &= d_1s_1(t)+d_2s_2(t)+\cdots+d_Qs_Q(t)\end{aligned}$$

Ideally $d_k = \boldsymbol{w}(\boldsymbol{r})^T\boldsymbol{l}(\boldsymbol{r}_k) = \begin{cases}1 & \boldsymbol{r}=\boldsymbol{r}_k\\ 0 & \boldsymbol{r}\neq\boldsymbol{r}_k\end{cases}$

Actually $d_k = \boldsymbol{w}(\boldsymbol{r})^T\boldsymbol{l}(\boldsymbol{r}_k) \neq 0$, for $\boldsymbol{r}\neq\boldsymbol{r}_k$

Time course at any voxel is a mixture of time courses of all sources within a brain.

An example



Time course at this target voxel:

$$\underbrace{s_T(t) = d_1 s_1(t) + d_2 s_2(t) + d_3 s_3(t)}_{\text{Linear mixture of three source activities}}$$

- If the target voxel is not at one of the three sources, d_1, d_2, d_3 are generally small and reasonably good source image can be obtained.
- If the source reconstruction is our final purpose, this leakage causes a blur, but does not cause any other serious problems.

However, if we compute phase relationships of voxel time courses, we normalize the spectrum with its amplitude, and then …..

To estimate connectivity, phase-related measures are commonly used.

Phase-related measures

voxel spectrum: $\sigma_k(f) = A_k(f)e^{i\theta(f)}$

Coherence:

$$\eta_{k,j} = \frac{\left\langle A_k A_j e^{i(\theta_k - \theta_j)} \right\rangle}{\sqrt{\left\langle A_k{}^2 \right\rangle \left\langle A_j{}^2 \right\rangle}}$$

Phase coherence:

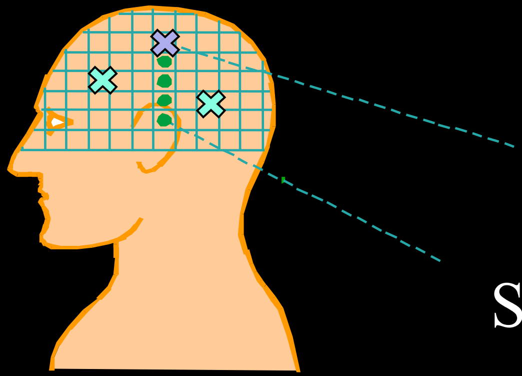
$$\rho_{k,j} = \left\langle e^{i(\theta_k - \theta_j)} \right\rangle$$

PLI:

$$\phi_{k,j} = \left\langle \mathrm{sgn}(\theta_k - \theta_j) \right\rangle$$

To extract the phase relationship, the voxel spectrum is normalized by its amplitude, and this normalization tends to introduce “spurious” phase relationship.

The same example



Seed spectrum: $\sigma_1(f)$

Spectrum at this target voxel:

$$\sigma_T(f) = d_1\sigma_1(f) + d_2\sigma_2(f) + d_3\sigma_3(f)$$

If $d_2 \gg d_1, d_3,$ $$\frac{\sigma_1\sigma_T^*}{|\sigma_1||\sigma_T|} \approx \frac{A_1e^{j\theta_1}d_2A_2e^{-j\theta_2}}{A_1d_2A_2} = e^{j(\theta_1-\theta_2)}$$

$\sigma_1(f)$ and $\sigma_T(f)$ have a phase relationship, in spite that there is no source at the target voxel.

The fact that d_1, d_2, d_3 are small does not help to avoid such a spurious phase relationship because these leakage constants are normalized out.

Phase-related measures

Coherence:
$$\eta_{k,j} = \frac{\left\langle A_k A_j e^{i(\theta_k - \theta_j)} \right\rangle}{\sqrt{\left\langle A_k^{\,2} \right\rangle \left\langle A_j^{\,2} \right\rangle}}$$

Phase coherence:
$$\rho_{k,j} = \left\langle e^{i(\theta_k - \theta_j)} \right\rangle$$

PLI:
$$\phi_{k,j} = \left\langle \mathrm{sgn}(\theta_k - \theta_j) \right\rangle$$

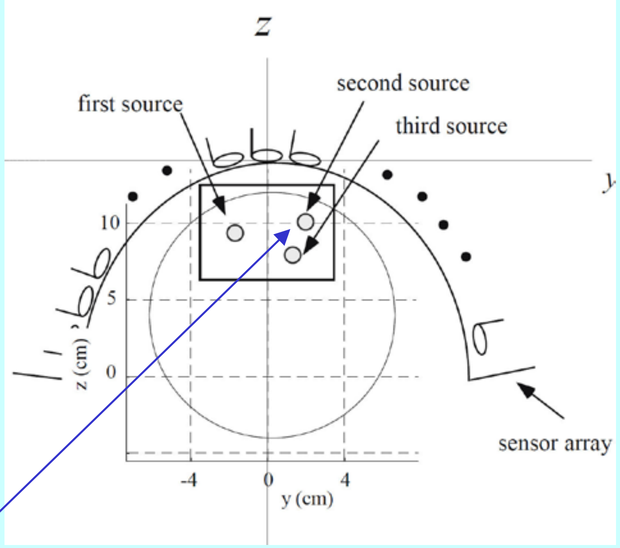
Our empirical finding for their stability:

Coherence > Phase coherence > PLI

This may be related to the order of the normalization and averaging.

Using phase relationship in source space analysis is tricky and caution is needed.

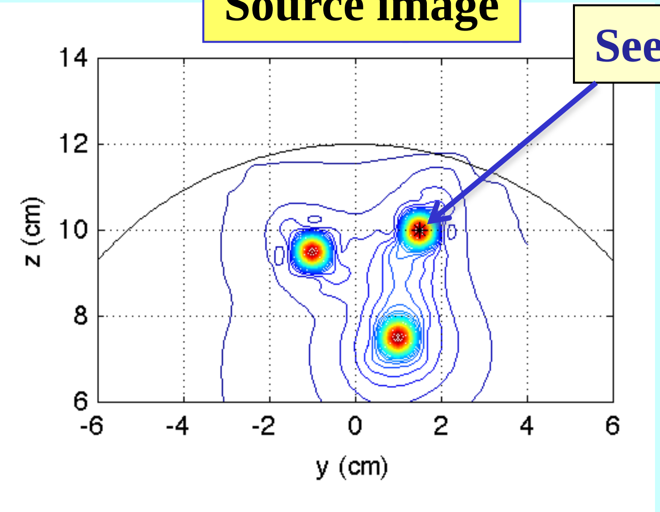
Phase related measures – high SNR computer simulation



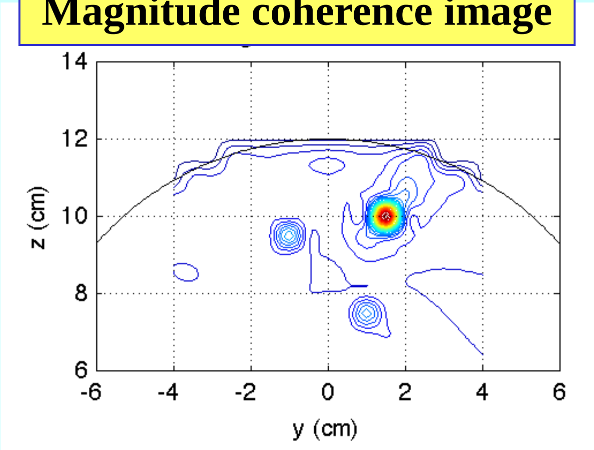
The second source interacts with the other two sources.

Source image

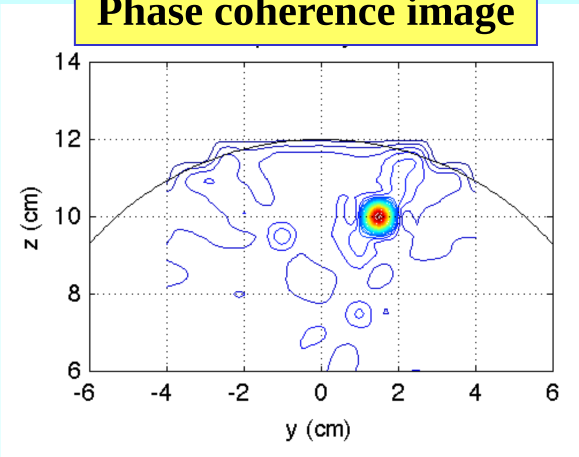
Seed point



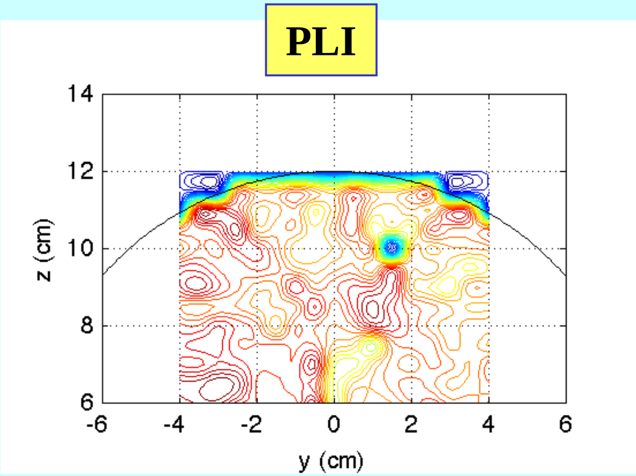
Magnitude coherence image



Phase coherence image

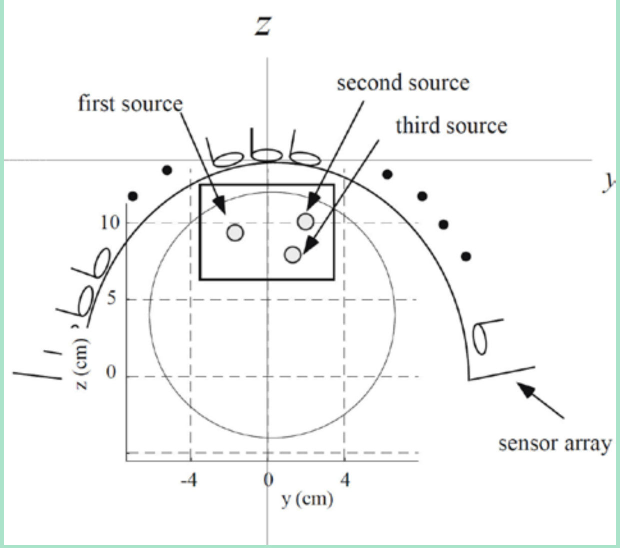


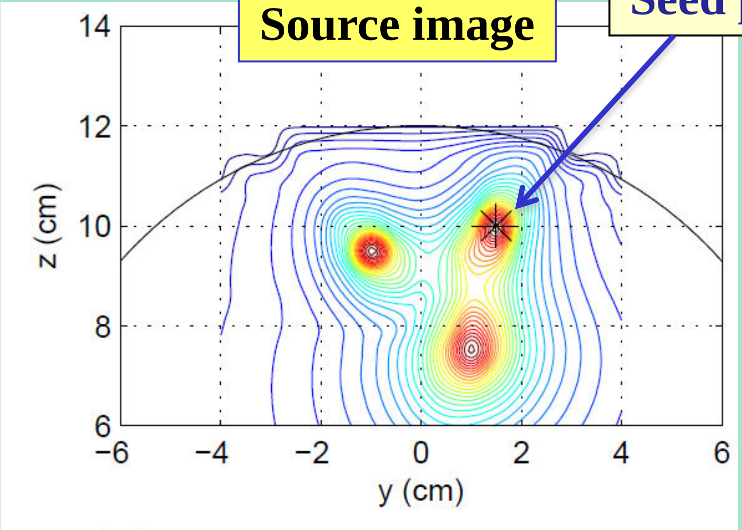
PLI



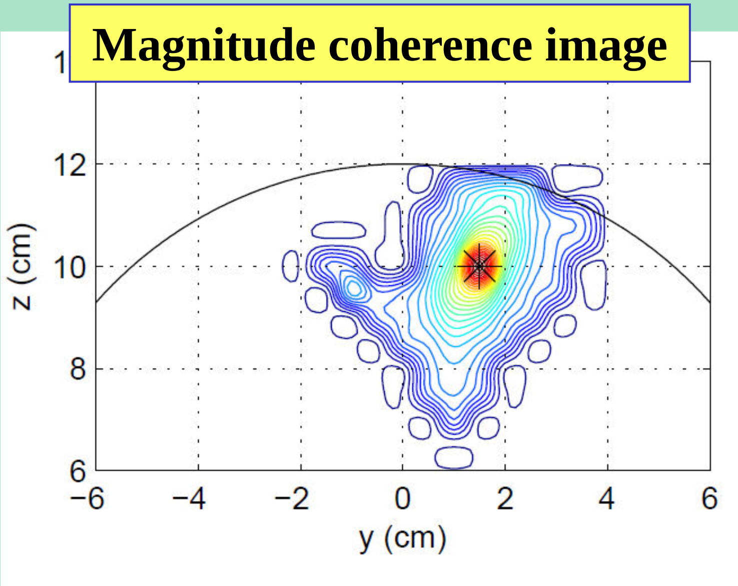
Leakage of imaging algorithm causes spurious connectivity, called seed blur.

Low SIR (SIR=0.5) computer simulation




Source image

Seed point


Magnitude coherence image

- **Magnitude coherence image shows a spurious coherence peak caused by the blur of imaging algorithm.**
- **The seed-blur is so high that it obscures the interacting sources in this low SNR simulation.**

Seed blur – analysis

Estimated seed voxel spectrum:

$$\hat{\sigma}_S = \sigma_S + d_1 \sigma_T$$

Estimated target voxel spectrum:

$$\hat{\sigma}_T = \sigma_T + d_2 \sigma_S$$

Voxel magnitude coherence

$$\left|\hat{\eta}\right| = \left|\left\langle \hat{\sigma}_T \hat{\sigma}_S^* \right\rangle\right| \Big/ \sqrt{\left\langle \left|\hat{\sigma}_T\right|^2 \right\rangle \left\langle \left|\hat{\sigma}_S\right|^2 \right\rangle}$$

$$= \frac{\left| \left\langle \sigma_T \sigma_S^* \right\rangle + d_1 \left\langle \left|\sigma_T\right|^2 \right\rangle + d_2 \left\langle \left|\sigma_S\right|^2 \right\rangle + d_1 d_2 \left\langle \sigma_S \sigma_T^* \right\rangle \right|}{\sqrt{\left[\left\langle \left|\sigma_T\right|^2 \right\rangle + d_2 \left\langle \left|\sigma_S\right|^2 \right\rangle + 2 d_2 \Re\left(\left\langle \sigma_T \sigma_S^* \right\rangle \right) \right] \left[\left\langle \left|\sigma_S\right|^2 \right\rangle + d_1 \left\langle \left|\sigma_T\right|^2 \right\rangle + 2 d_1 \Re\left(\left\langle \sigma_T \sigma_S^* \right\rangle \right) \right]}}$$

When no brain interaction exists, i.e., $\left\langle \sigma_T \sigma_S^* \right\rangle = 0$, $\left|\hat{\eta}\right| \neq 0$.

Imaginary part of coherence

$$\Im(\hat{\eta}) = \frac{(1-d_1 d_2)}{\sqrt{\left[1+d_2\frac{\langle|\sigma_S|^2\rangle}{\langle|\sigma_T|^2\rangle}+2d_2\frac{\Re(\langle\sigma_T\sigma_S^*\rangle)}{\langle|\sigma_T|^2\rangle}\right]\left[1+d_1\frac{\langle|\sigma_T|^2\rangle}{\langle|\sigma_S|^2\rangle}+2d_1\frac{\Re(\langle\sigma_T\sigma_S^*\rangle)}{\langle|\sigma_S|^2\rangle}\right]}}\Im(\eta)$$

When the phase= $\frac{\pi}{2}$, $\Im(\eta) = 0$ ➡

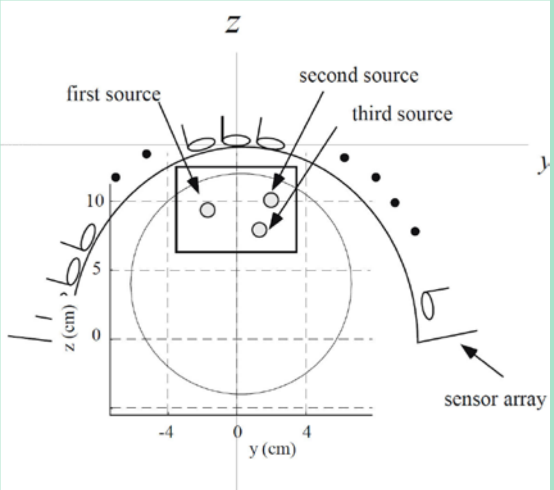
Amplitude of imaginary part may not necessarily indicate the strength of the connectivity.

Corrected imaginary coherence

$$\tilde{\eta} = \frac{\Im(\hat{\eta})}{\sqrt{\left[1-[\Re(\hat{\eta})]^2\right]}} = \frac{\Im(\eta)}{\sqrt{\left[1-[\Re(\eta)]^2\right]}}$$

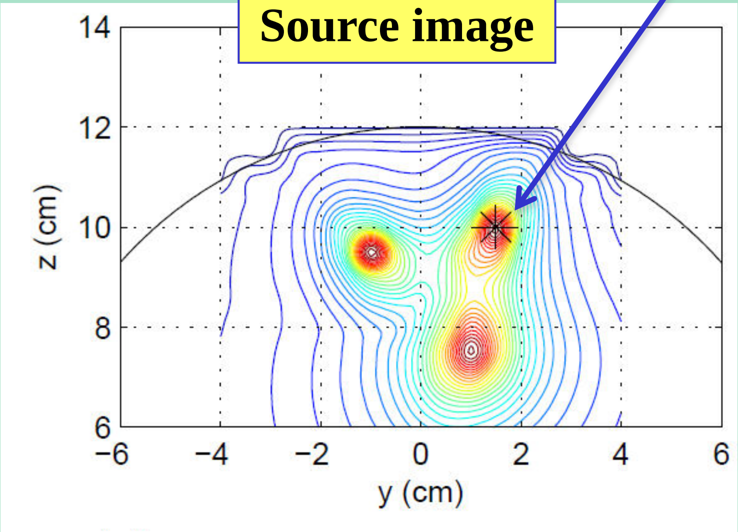
G. Nolte *et al. "Identifying true brain interaction from EEG data using the imaginary part of coherency," Clin. Neurophysiol.*

Imaginary coherence– computer simulation

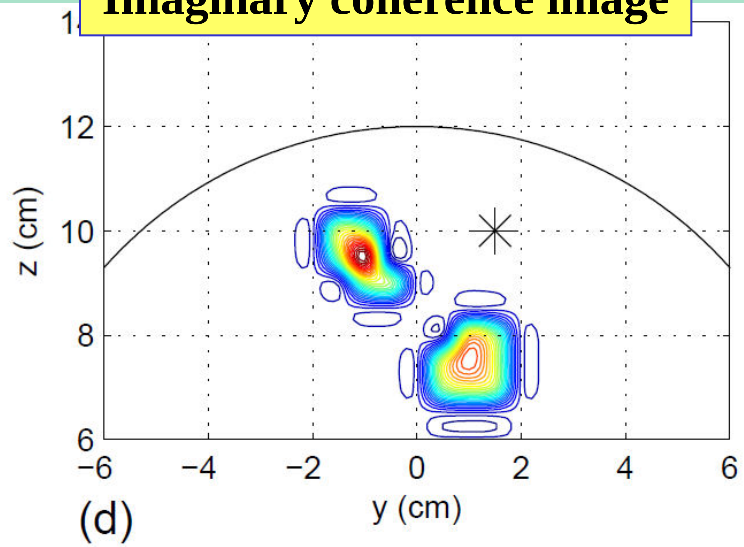


Source image

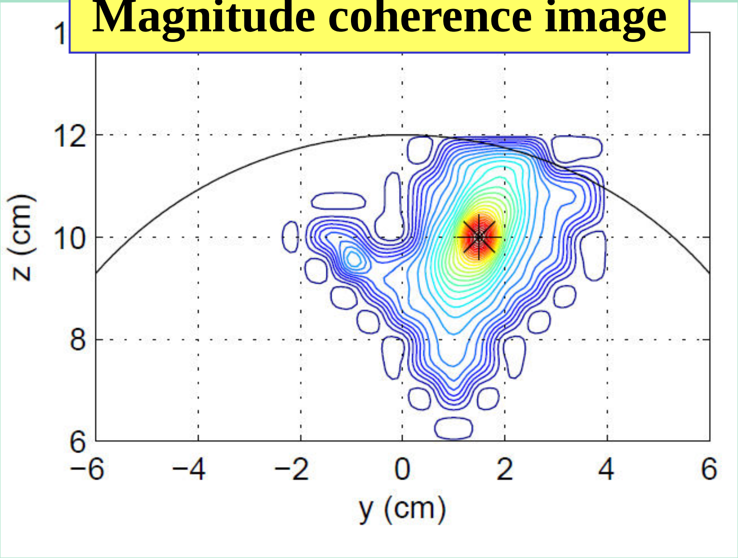
Seed point



Imaginary coherence image

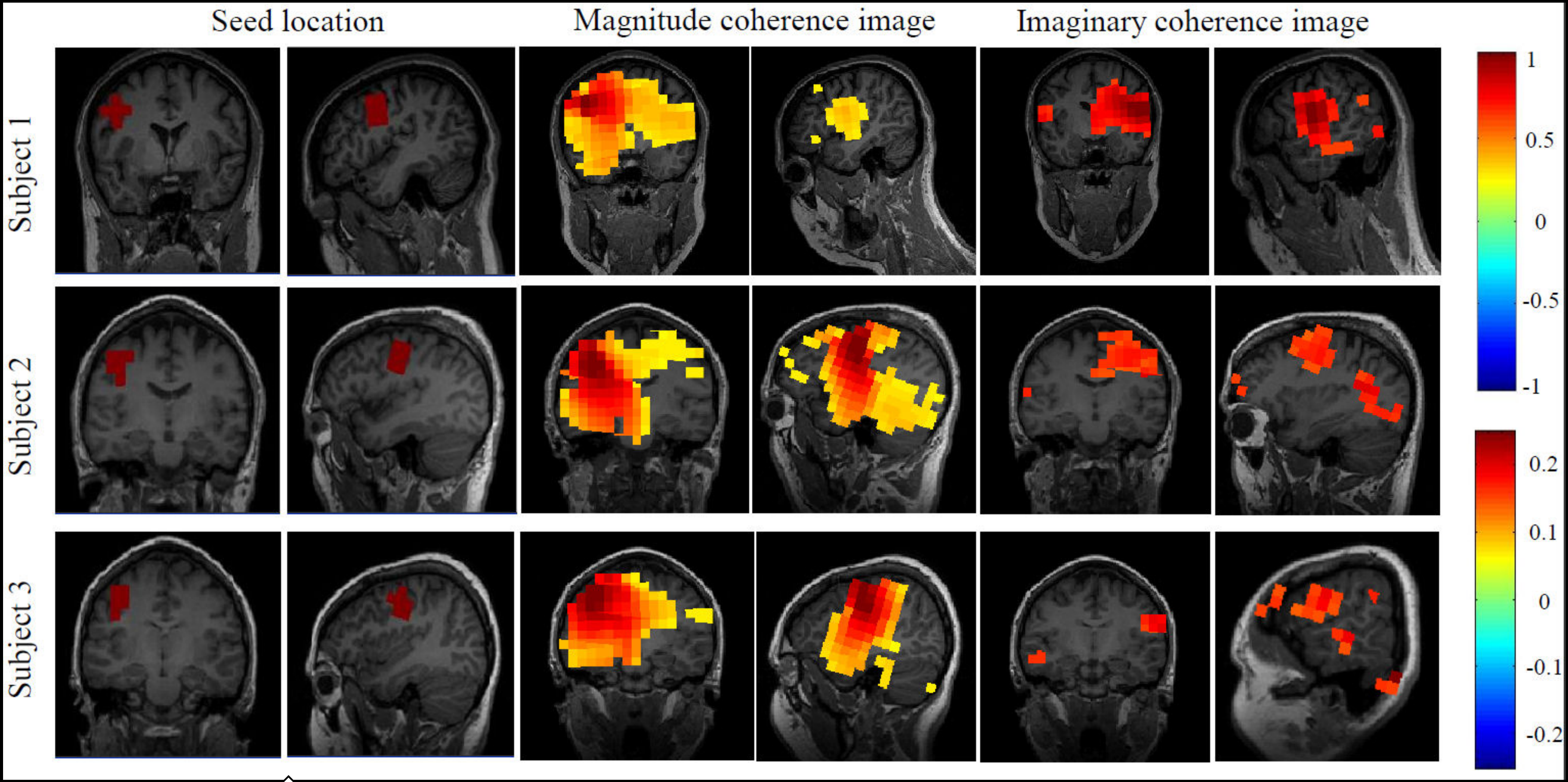


Magnitude coherence image



Surrogate-data method is used for thresholding.

Results of Coherence Imaging

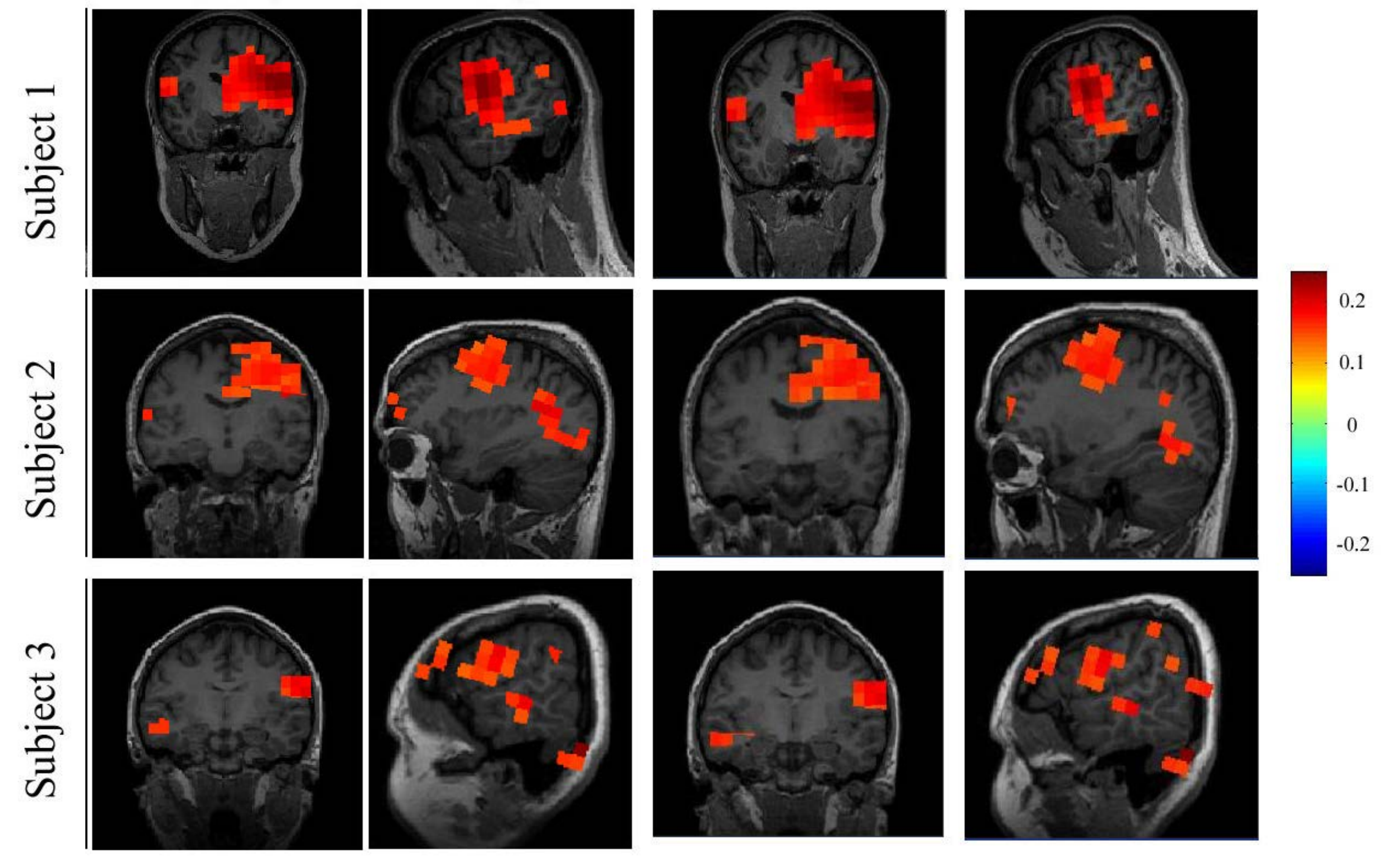


Voxels located within the left pre-central gyrus (left primary motor area) were selected as seed voxels.

Magnitude coherence image only shows seed blur, but imaginary coherence image shows activities near the right primary motor area

Imaginary coherence: $\Im\left(\hat{\eta}\right)$

Corrected imaginary coherence: $\tilde{\eta}$



Because imaginary coherence is small (<0.2) in this case, there is no clear difference between imaginary and corrected imaginary images.

Problems in estimating causal relationships using Granger-based measures

- Representative measures of estimating causal relationships are Granger-causality-based measures, including directed transfer function (DTF) and partial directed coherence (PDC).

- Estimating Granger-causality-based measures requires an accurate estimation of AR coefficients. However, various factors affect the accuracy of estimating the AR process, and cause spurious causal relationships.

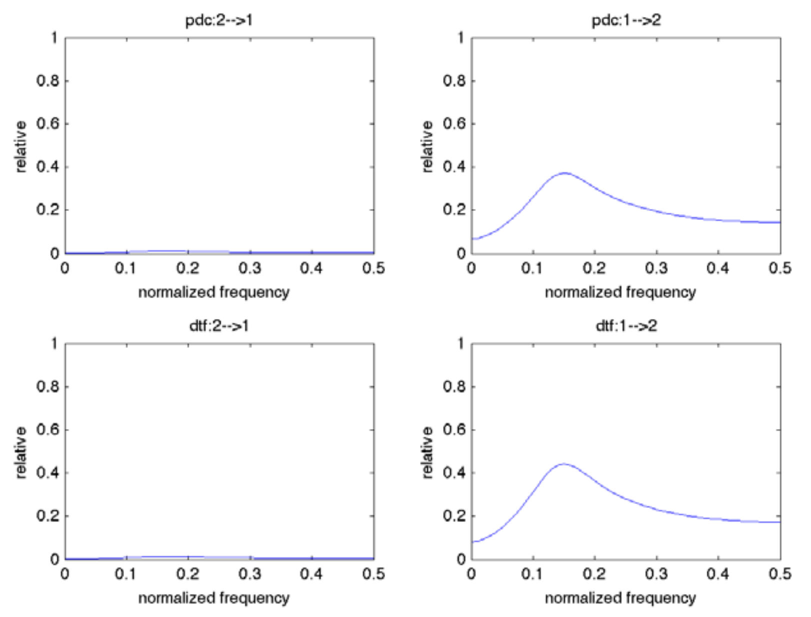
Simple two-channel computer simulation to cause *spurious* causal relationship.

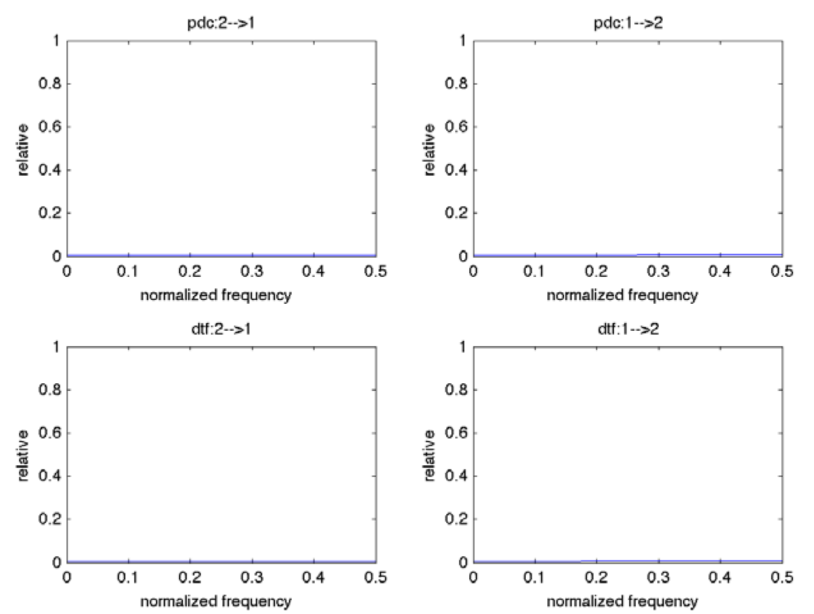
VAR process

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 0.9 & 0 \\ 0.16 & 0.8 \end{bmatrix} \begin{bmatrix} y_1(t-1) \\ y_2(t-1) \end{bmatrix} + \begin{bmatrix} -0.5 & 0 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} y_1(t-2) \\ y_2(t-2) \end{bmatrix} + \boldsymbol{e}(t)$$

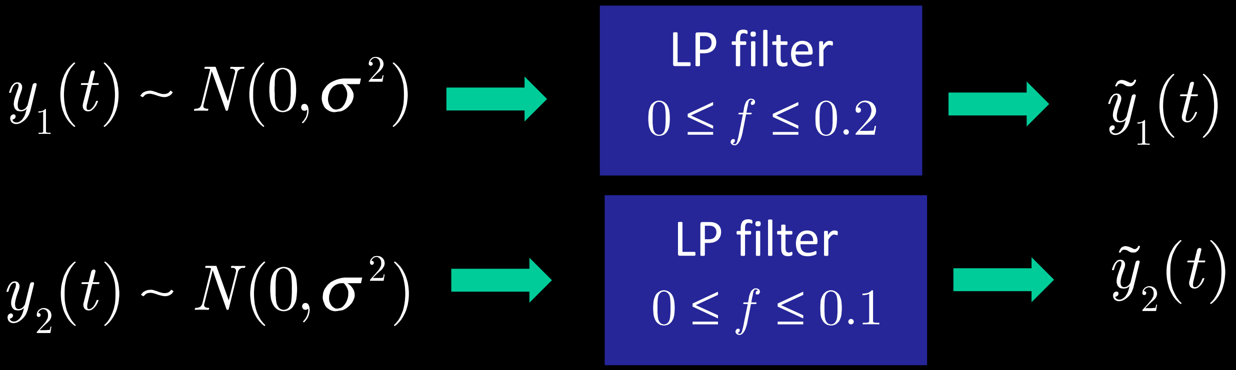
Gaussian random process.

$$y_1(t) \sim N(0, \sigma^2)$$
$$y_2(t) \sim N(0, \sigma^2)$$



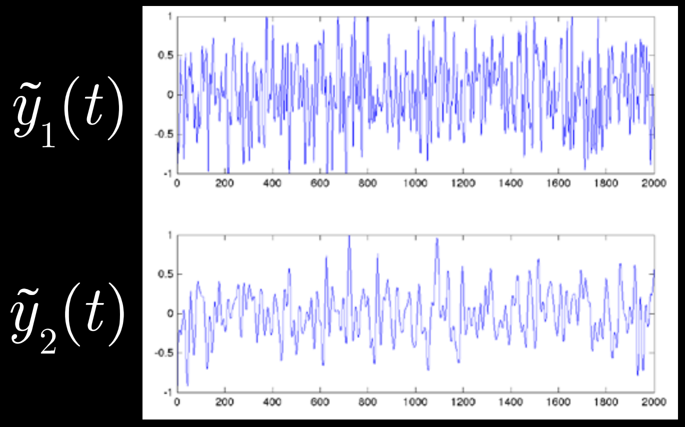


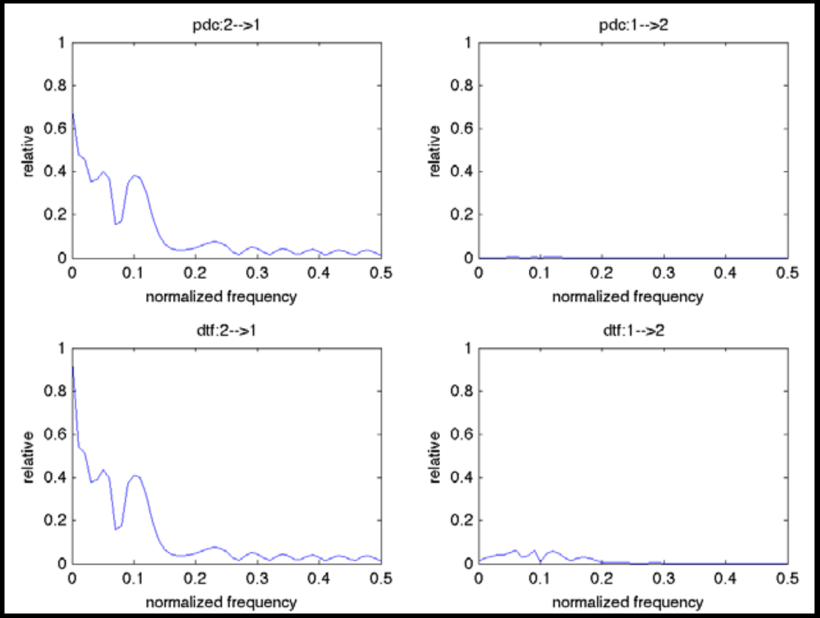
Low-pass filtered Gaussian random process.



$y_1(t) \sim N(0,\sigma^2) \rightarrow$ **LP filter** $0 \le f \le 0.2$ $\rightarrow \tilde{y}_1(t)$

$y_2(t) \sim N(0,\sigma^2) \rightarrow$ **LP filter** $0 \le f \le 0.1$ $\rightarrow \tilde{y}_2(t)$





Even the filtering with slightly different bandwidth causes spurious causal relationship!!

Why is such a spurious solution caused?

Voxel time course is modeled using MVAR process:

$$s_k(t) = \sum_{j=1}^{K} A_{k,j}(1)s_j(t-1) + \sum_{j=1}^{K} A_{k,j}(2)s_j(t-2) + \cdots + \sum_{j=1}^{K} A_{k,j}(P)s_j(t-P) + e_k(t)$$

AR coefficients that should be zero have small non-zero values.

How to avoid spurious causal relationship

It is probably true that only few AR coefficients should have non-zero values.

⬇

Imposing a sparsity constraint when estimating AR coefficients.

Sparse AR Modeling

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}$$

time course data AR coefficients

Prior probability: $f(\boldsymbol{x}) = N(\boldsymbol{x} \mid 0, \boldsymbol{\Phi})$

Noise: $\boldsymbol{e} \sim N(\boldsymbol{e} \mid 0, \Sigma_e) \Rightarrow f(\boldsymbol{y} \mid \boldsymbol{x}) = N(\boldsymbol{y} \mid \boldsymbol{H}\boldsymbol{x}, \boldsymbol{\Lambda})$

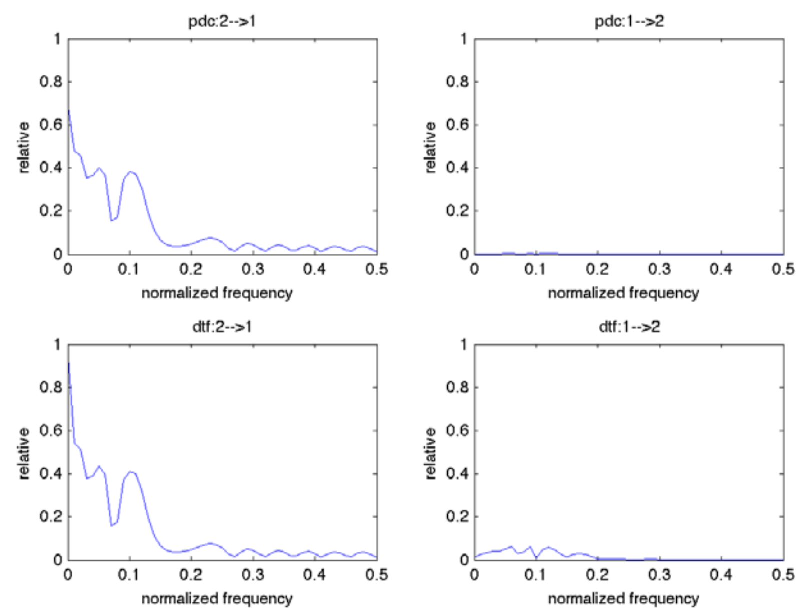
log evidence: $\log f(\boldsymbol{y} \mid \boldsymbol{\Phi}) = \log N(\boldsymbol{y} \mid 0, H\boldsymbol{\Phi}H^T + \Sigma_e)$

$$\hat{\boldsymbol{\Phi}} = \arg\min_{\boldsymbol{\Phi}} \left[\left| H\boldsymbol{\Phi}H^T + \boldsymbol{\Sigma}_e \right| + \boldsymbol{y}^T (H\boldsymbol{\Phi}H^T + \boldsymbol{\Sigma}_e)^{-1} \boldsymbol{y} \right]$$

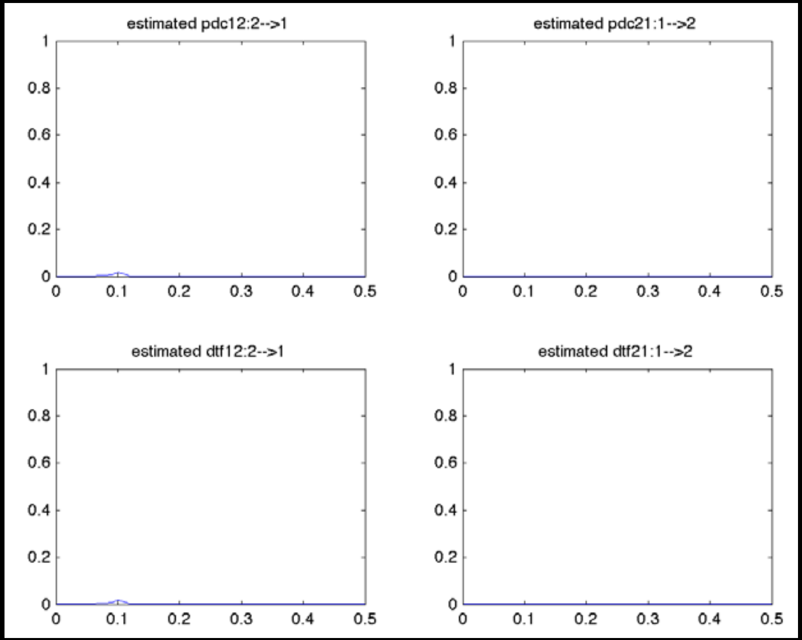
Once $\hat{\boldsymbol{\Phi}}$ is obtained, $\bar{\boldsymbol{x}} = \hat{\boldsymbol{\Phi}} H^T (H\boldsymbol{\Phi}H^T + \Sigma_e)^{-1} \boldsymbol{y}$

Applying sparse AR modeling to low-pass filtered Gaussian random process.

Least-squares AR modeling

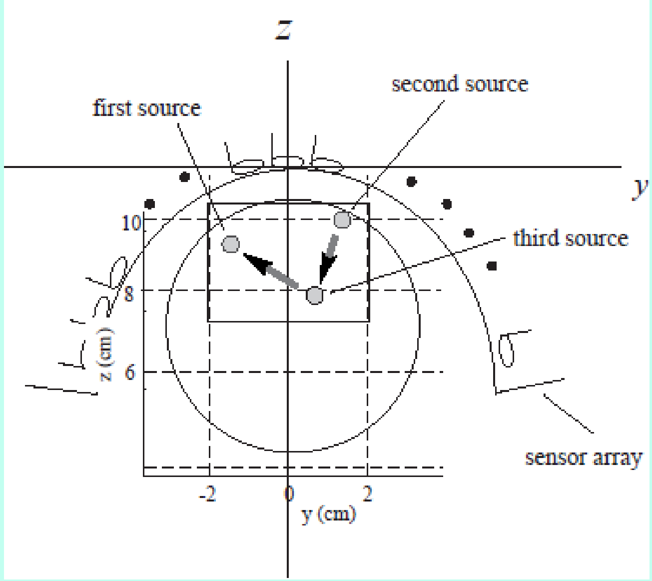


Sparse AR modeling



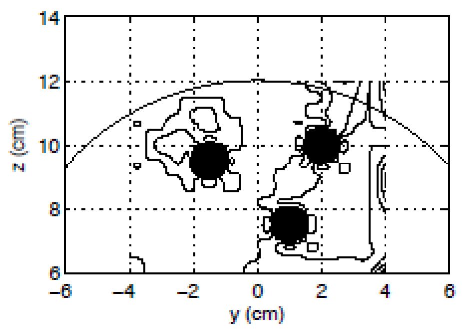
Sparse AR modeling works for the low-pass filtered data. Can it solve other problems?

Computer simulation of source-space causality analysis



- Three sources.
- Source time courses modeled with AR process.
- Brain noise added to simulated sensor data.
- Beamformer used to obtain voxel time courses.
- Voxel time courses at the three source locations used for estimating causal relationships.

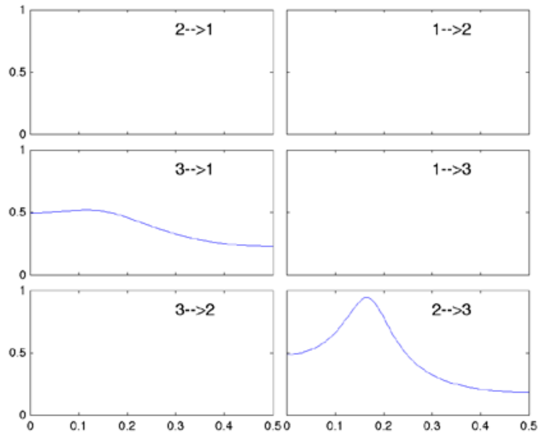
Beamformer source image



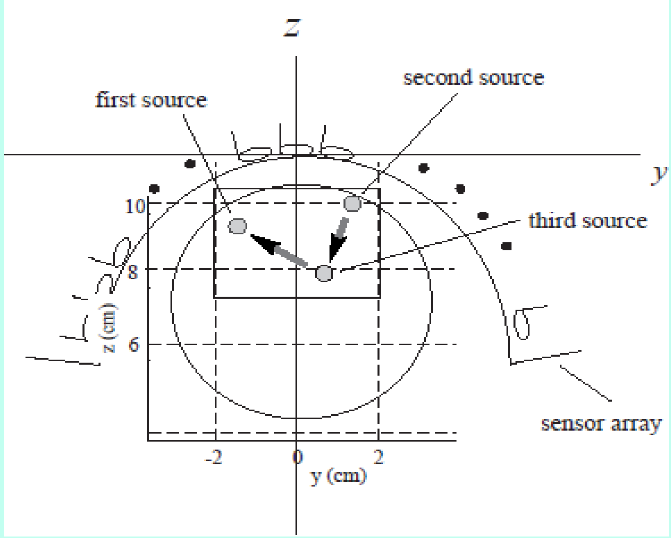
MVAR time courses

$$\begin{bmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0.4 \\ 0 & 0.9 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} s_1(t-1) \\ s_2(t-1) \\ s_3(t-1) \end{bmatrix} + \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & -0.8 & 0 \\ 0 & 0 & -0.2 \end{bmatrix} \begin{bmatrix} s_1(t-2) \\ s_2(t-2) \\ s_3(t-2) \end{bmatrix} + \boldsymbol{e}(t)$$

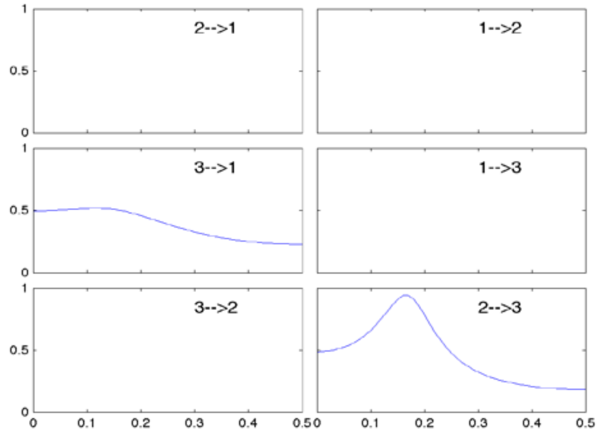
Assumed PDC



Results when SIR=4

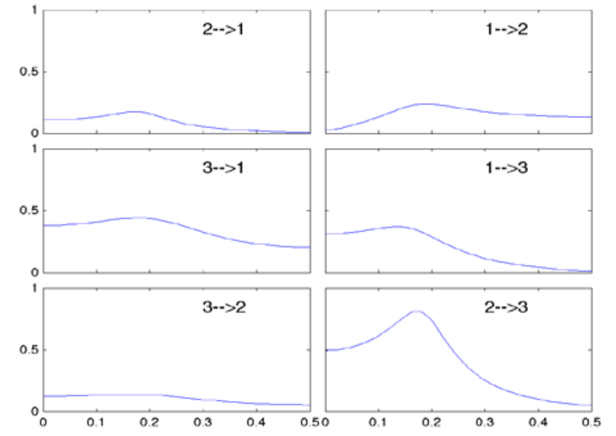


Assumed causal relationship

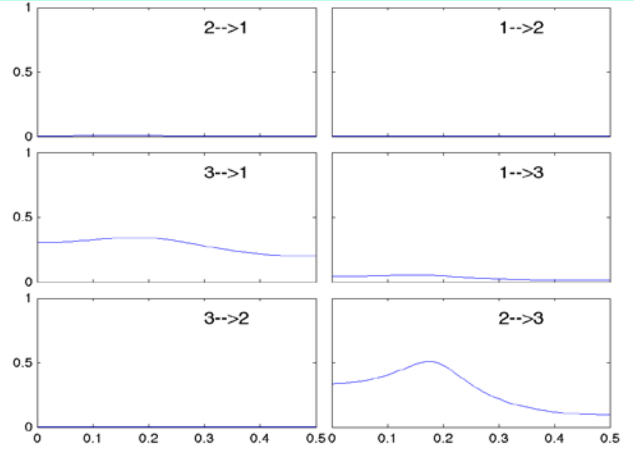


Estimated PDC

Least-squares AR modeling



Sparse AR modeling



We need some way to assess the goodness/badness of estimated results?

Statistical thresholding using surrogate-data bootstrapping

Surrogate data

$$s(t) \xrightarrow[FT]{} \sigma(f)\exp[-j\varepsilon] \xrightarrow[IFT]{} s^{*}(t)$$

Compute PDCs using many sets of surrogate data at each frequency bin

$$B \text{ surrogate data sets: } s^{1*}(t), s^{2*}(t), \ldots, s^{B*}(t)$$

$$\downarrow$$

$$B \text{ surrogate PDCs: } \psi^{1*}(f), \psi^{2*}(f), \ldots, \psi^{B*}(f)$$

Maximum statistics for multiple comparisons problem

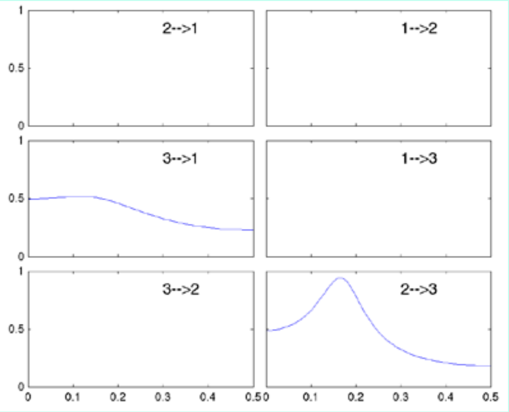
$$N(\# \text{ of fr. bin}) \text{ maximum PDCs: } \psi^{1}_{\max}(f), \psi^{2}_{\max}(f), \ldots, \psi^{N}_{\max}(f)$$

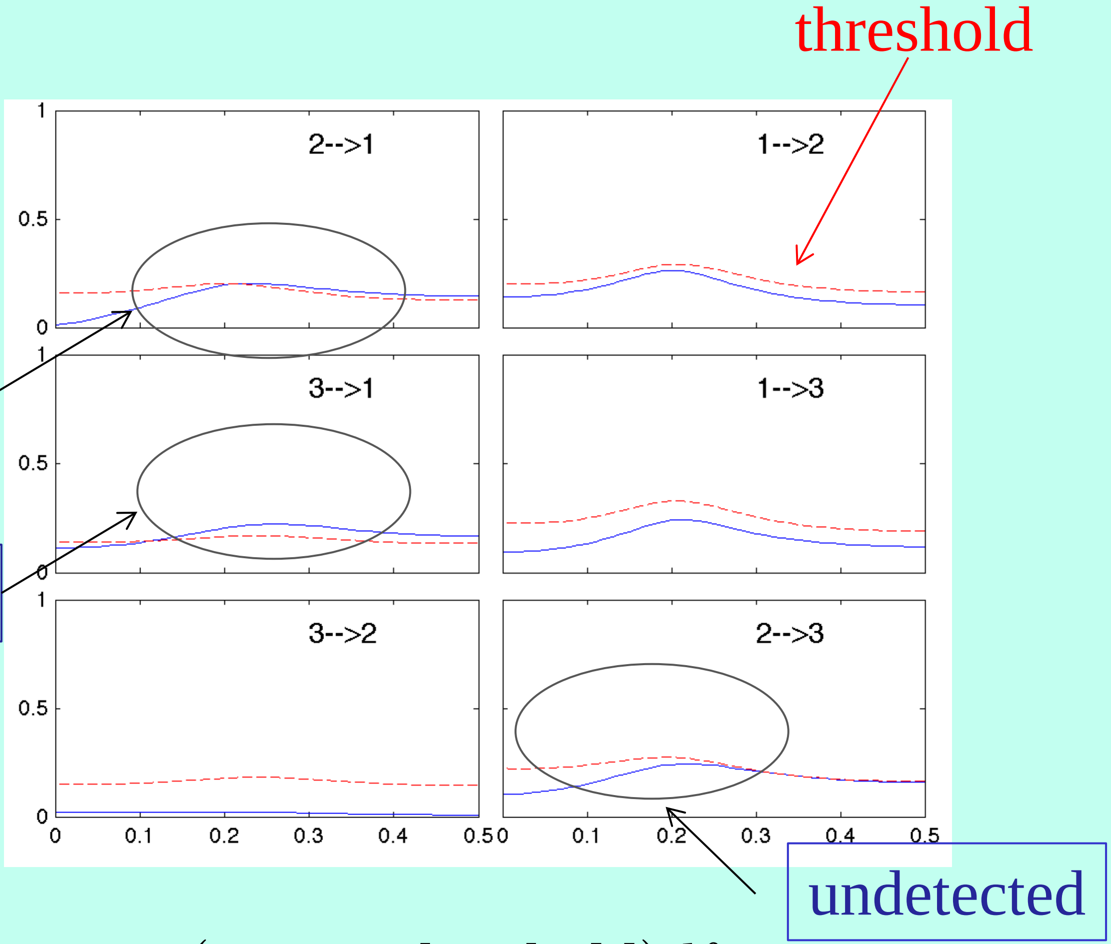
Setting the level of significance to α, choose the $(1-\alpha)N$th smallest $\psi^{*}_{\max}$ is obtained as the threshold.

J.Theiler, S.Eubank, A.Longtin, B.Galdrikian, and J.D. Farmer, Physica D, vol.58, 1992.

Results of thresholding

Assumed PDC





threshold

detected

detected

undetected

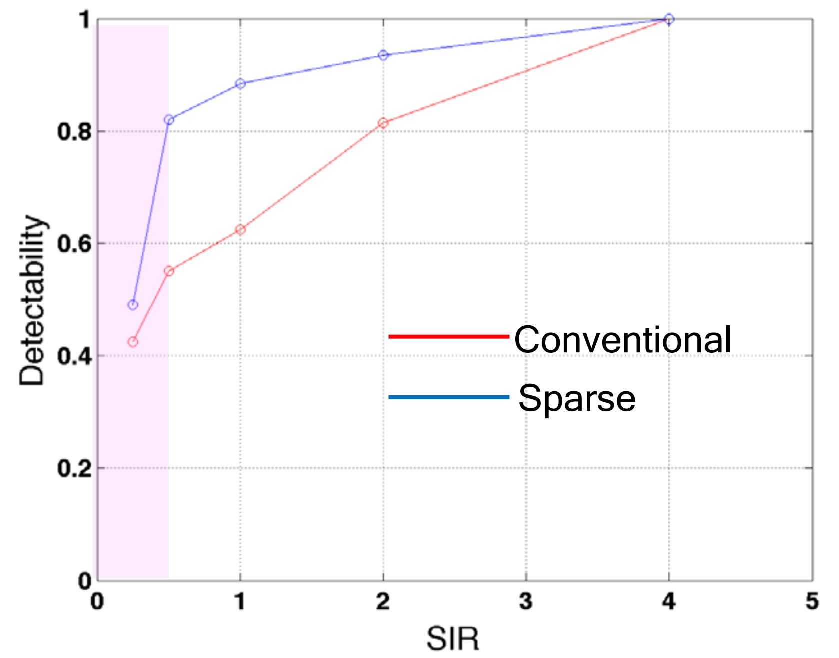
Compute Q : $Q = \int_{\mathrm{PDC} \geq \mathrm{threshold}} (\mathrm{PDC} - \mathrm{threshold}) df$

If Q> ε , the causal relationship is detected.
If Q< ε, the causal relationship is undetected.

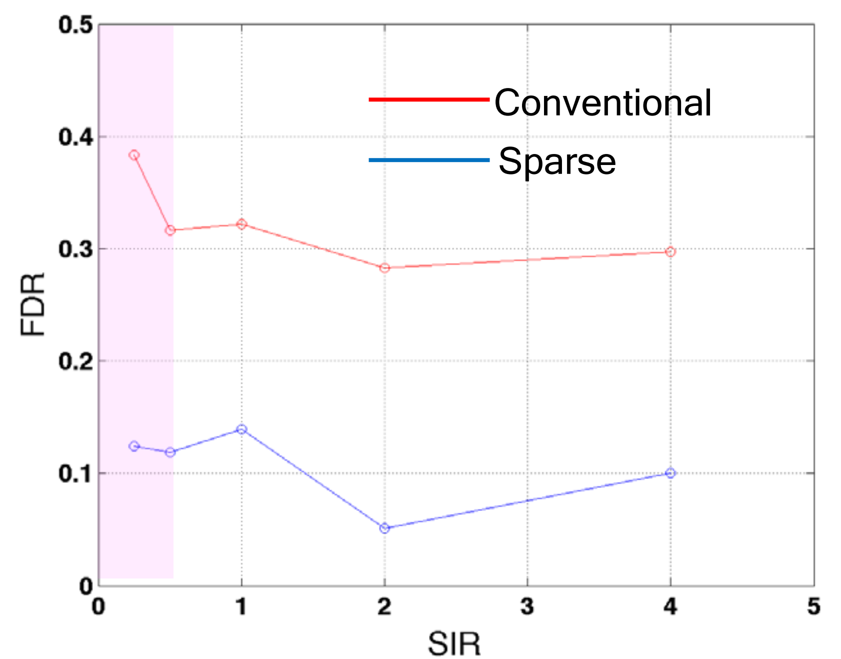
Results of Monte Carlo simulation

Each condition having 100 Monte Carlo trials

Detectability



False detection ratio



- Sparse AR estimation gives better causality estimates for high and medium SIR range.
- For low SIR range (SIR<0.5), detectability drops rapidly for both methods.

Even using the sparse method, the data should have high SIR (>1) to get reliable causality estimate!!

Thus, caution is needed when trying to estimate causal relationship using non-averaged raw trials, which we usually want to use for causality analysis.

Summary

- In source space analysis, a voxel time course is a mixture of time courses of all sources within a brain. Thus, the estimation of voxel phase relationships is tricky and needs caution.
- The leakage of an imaging algorithm causes the seed blur, which can be removed by using the imaginary coherence or corrected coherence.
- In Granger-causality-based analysis, an accurate estimation of AR coefficients is key to obtain reliable results.
- Sparse AR modeling gives better causality estimates for high and medium SIR range. For low SIR range (SIR<0.5), detectability drops rapidly even for the sparse method.

Thank you very much for your attention.