



38th Annual International Conference of the *IEEE Engineering in Medicine and Biology Society*

*Disney's Contemporary Resort, Orlando, FL
August 16-20, 2016*

A Hands-on Approach to Neural Connectivity Inference Methods

Luiz A. Baccalá, Koichi Sameshima, Laura Astolfi



Escola Politécnica
Universidade de São Paulo

Program Overview

Morning

- ✦ Introduction and Overview - L.A. Baccalá
- ✦ Applications of Granger Causality to Neuroscience - Mingzhou Ding
- ✦ Statistical and Software Applications - K. Sameshima
- ✦ Power User Applications - L. Astolfi

Afternoon

Data Analysis Challenges

Challenge Resolution and Discussion

Overview

- ✦ Introduction - Historical Perspective
- ✦ Correlation based Connectivity and its Inadequacy
- ✦ Granger Causality based Connectivity
- ✦ Partial Directed Coherence
- ✦ Directed Transfer Function
- ✦ Model Diagnostics and Interpretation

Key Technology: Multichannel Neural Signal Recording

Modalities

EEG →

LFPs



voxel
based

MEG

fMRI

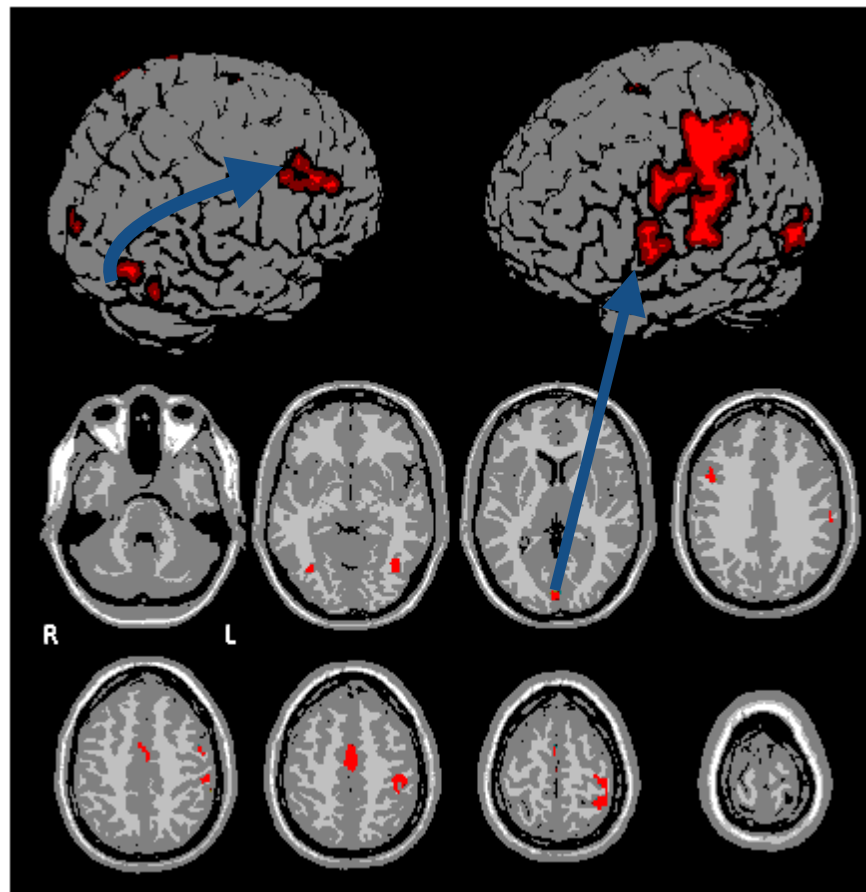
Multi-Single Unit Data



Paradigm Change

Active Areas x Interrelations among areas

neo-phrenology \longrightarrow connectivity

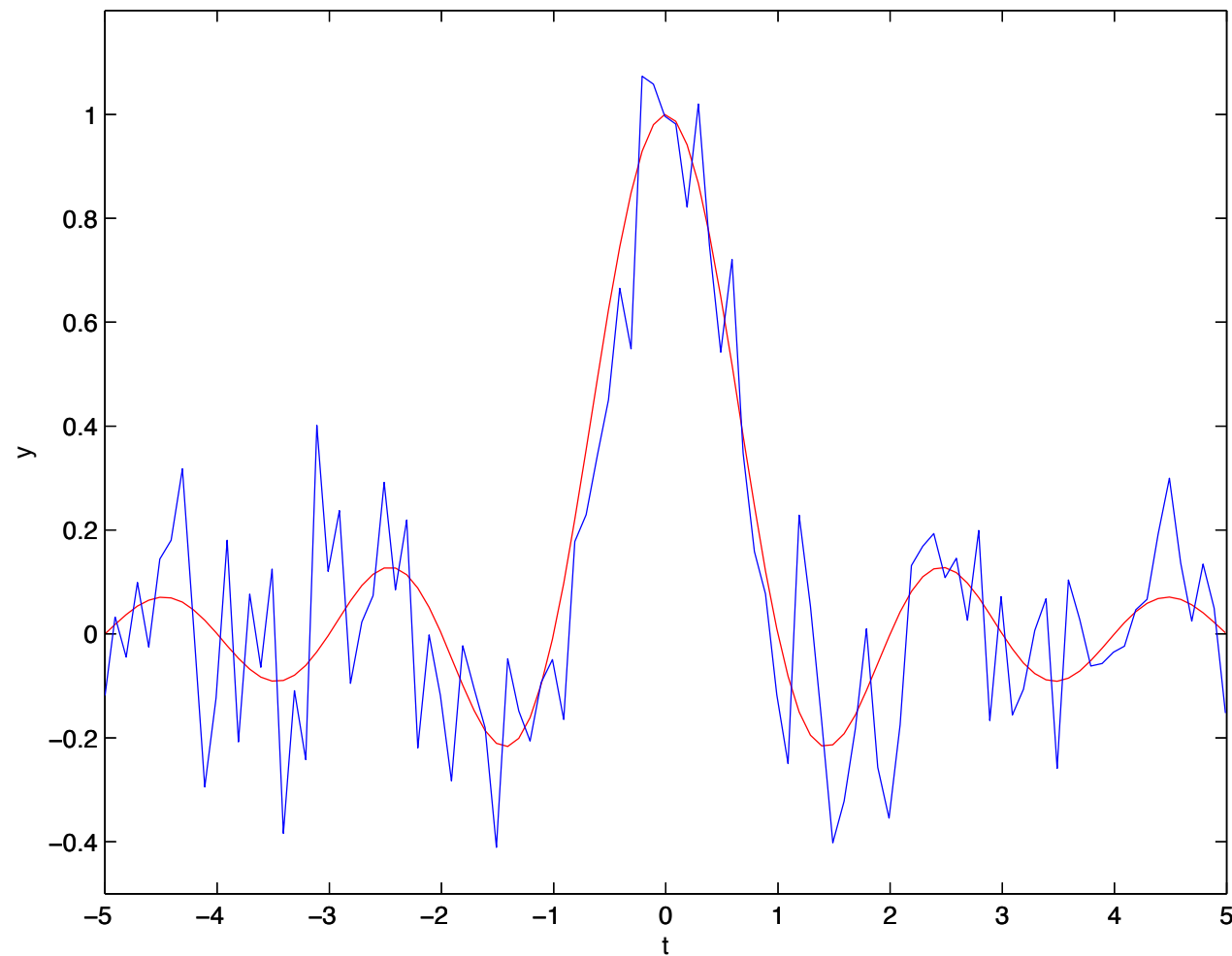


Investigate Information Flow

Goal: Recover the active dependence structure between
neural observations

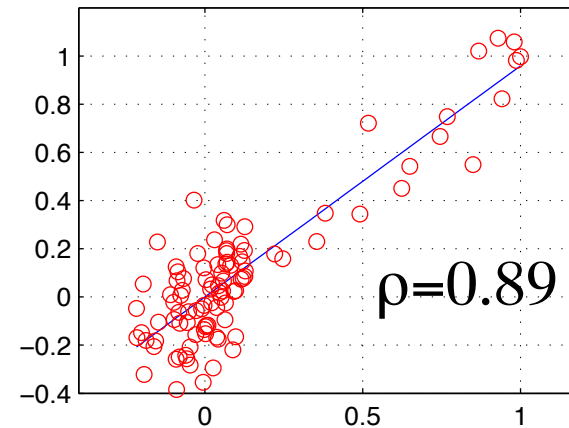
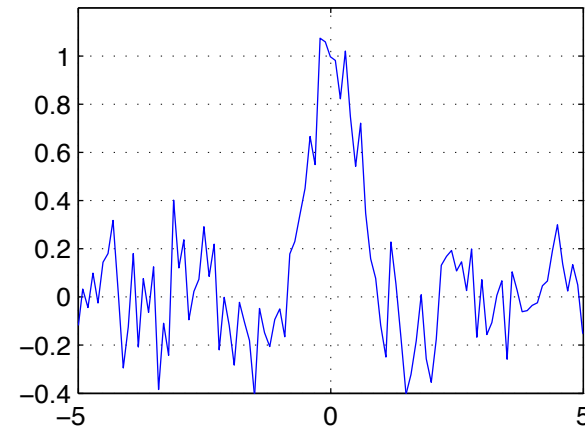
Correlation Ideas

Karl Pearson - dependence between variables



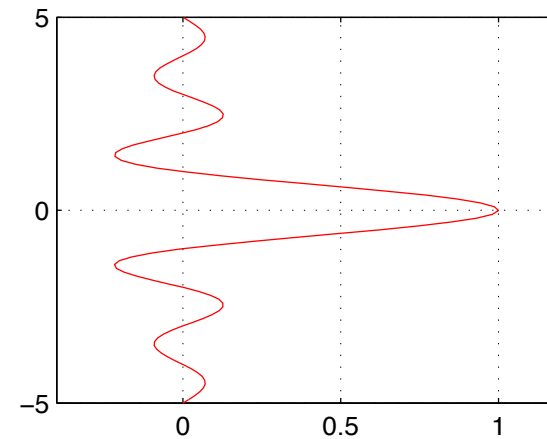
Correlation Ideas

Karl Pearson - dependence between variables (1896)



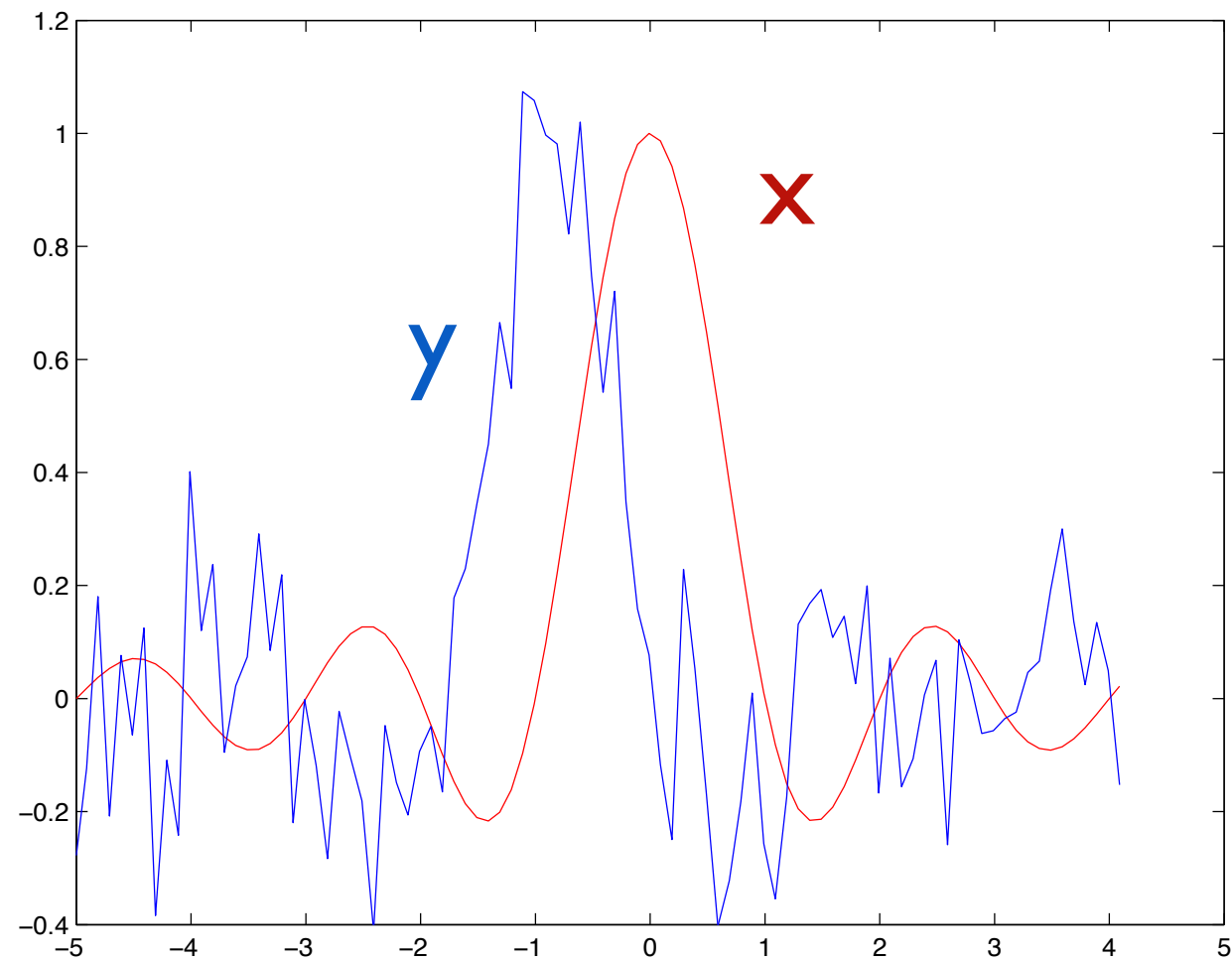
Correlation Coefficient

$$0 \leq |\rho| \leq 1$$

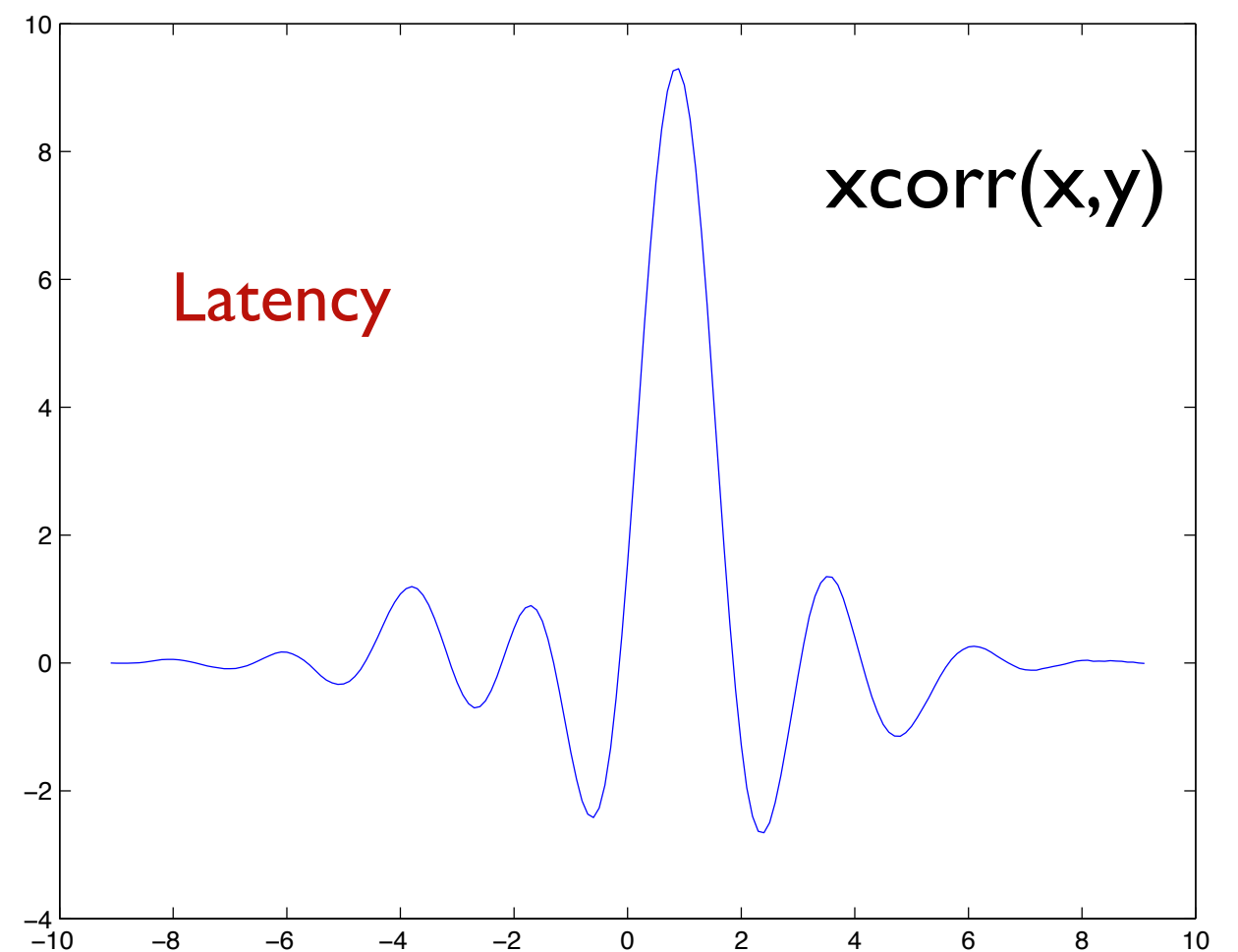


Degree of geometrical similarity

Delayed Waveforms

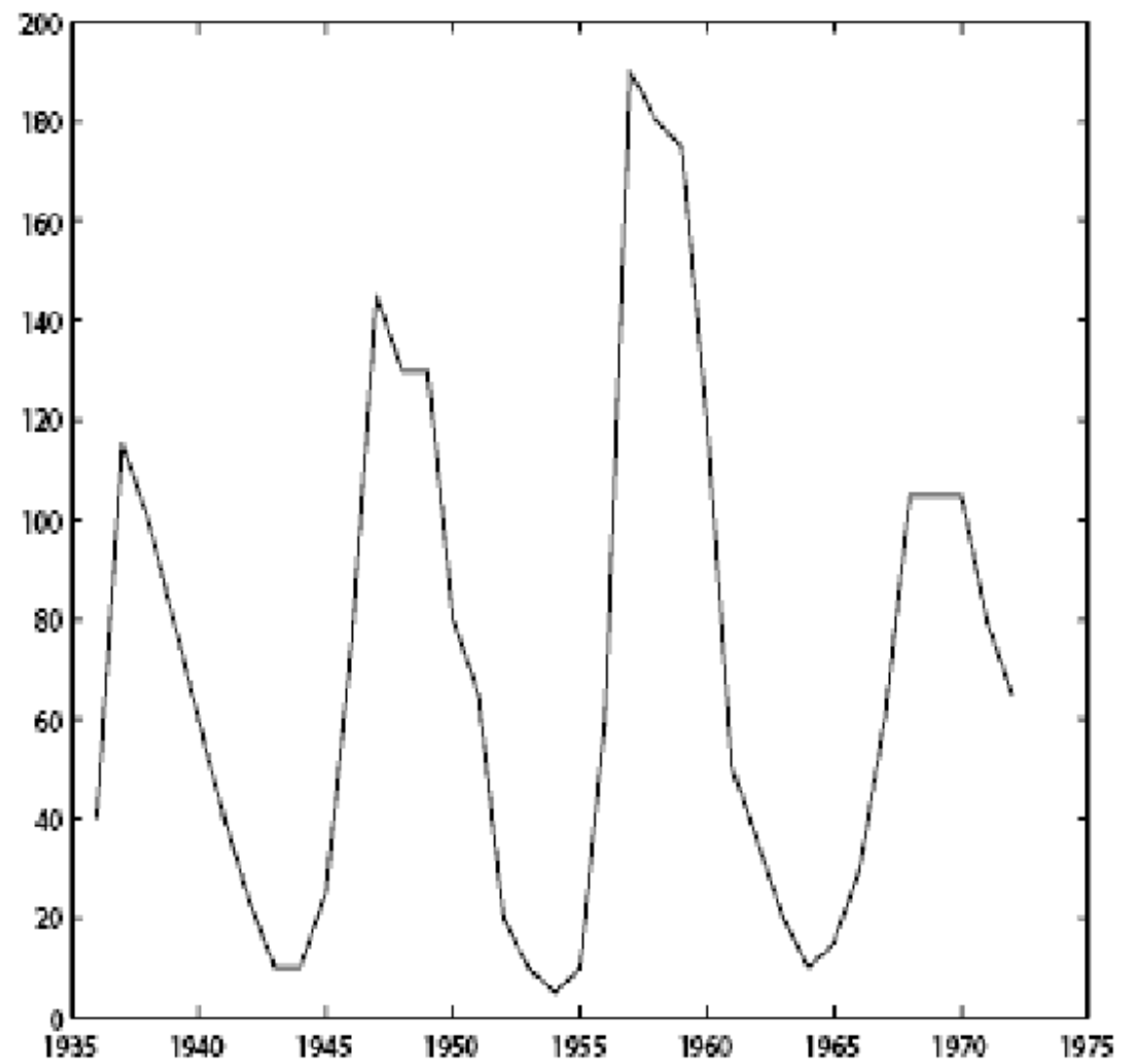


Cross Correlation Function



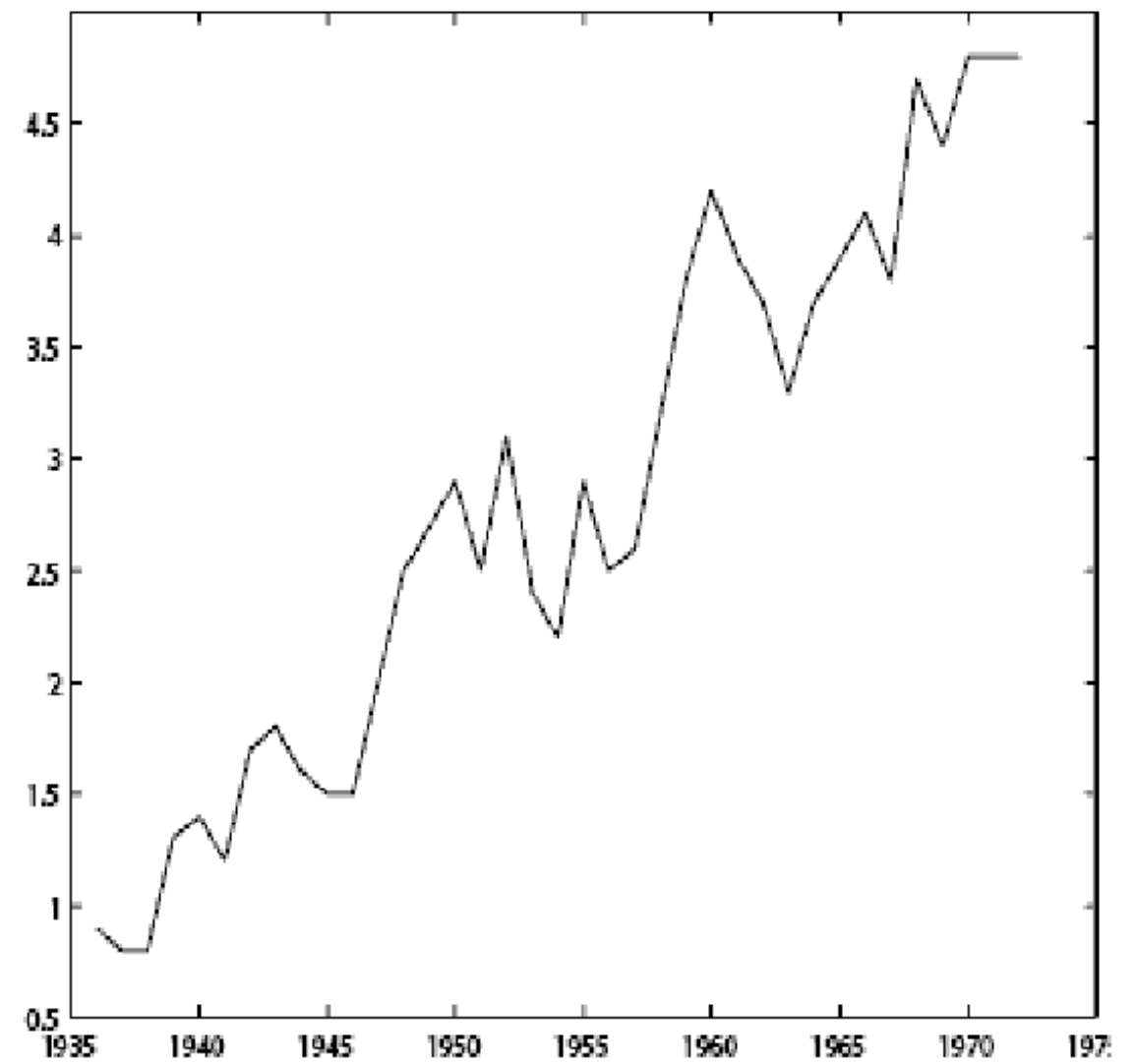
Delay of x with respect to y

Wölfer Sunspots



(a)

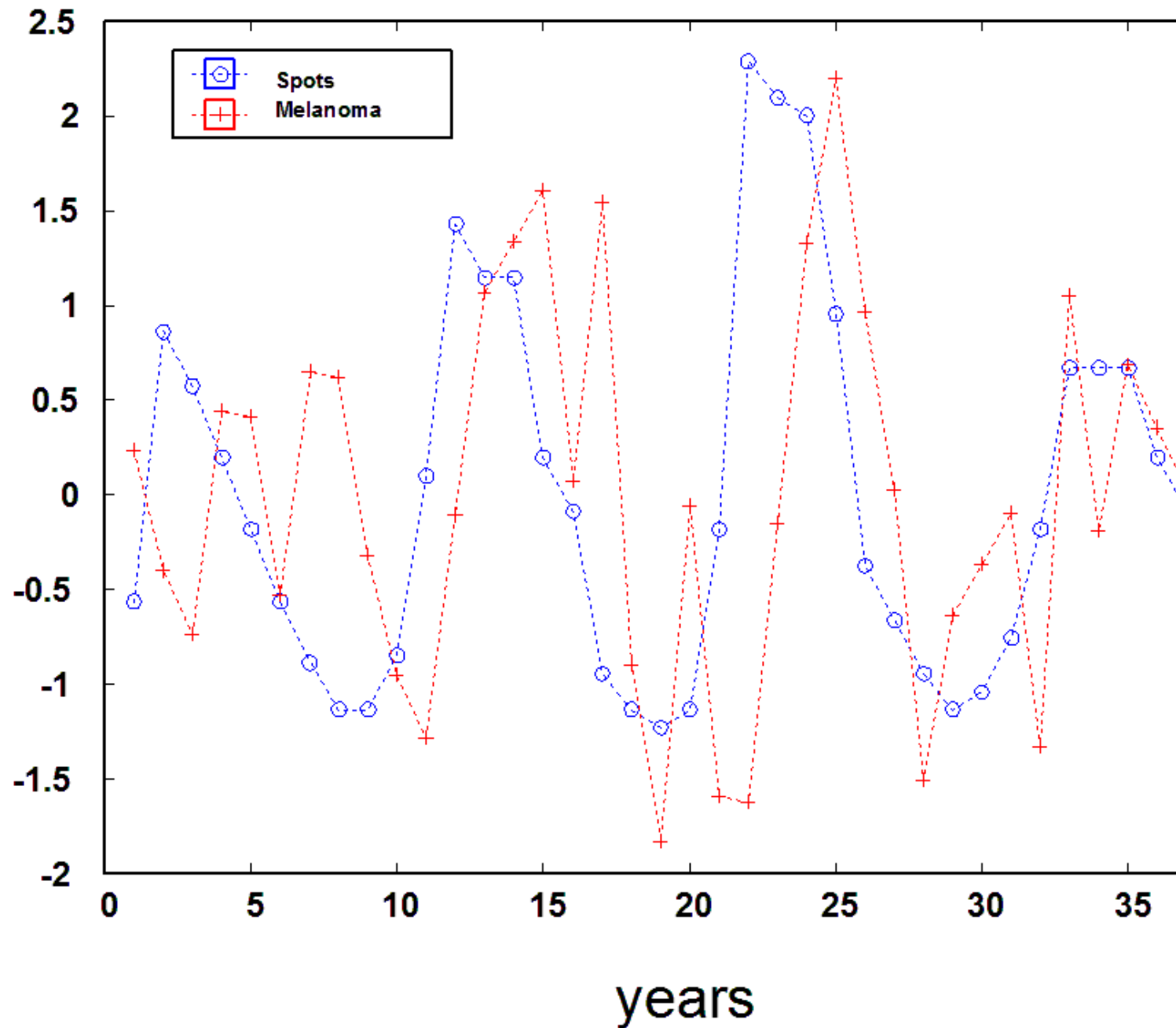
Melanoma Incidence (Connecticut, USA)



(b)

Time Series

Are Melanomas and Solar activity connected?

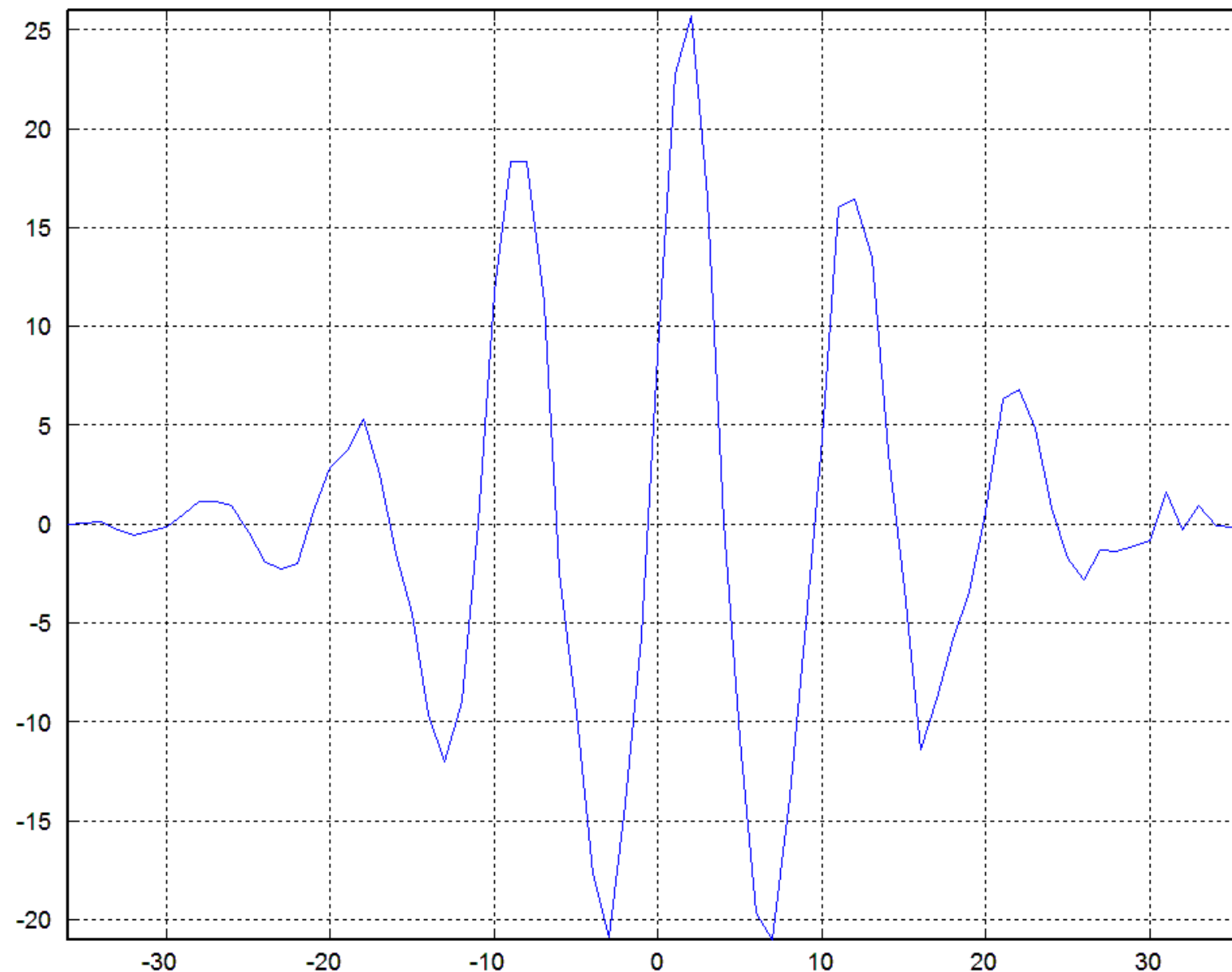


(after extracting means and trends and normalizing with respect to variance)

What is the meaning of cross-correlation?

Latency - when the correlation is maximum

(propagation delay) 3 years



**Melanoma
peaks are
late with
respect to
solar activity**

Logic

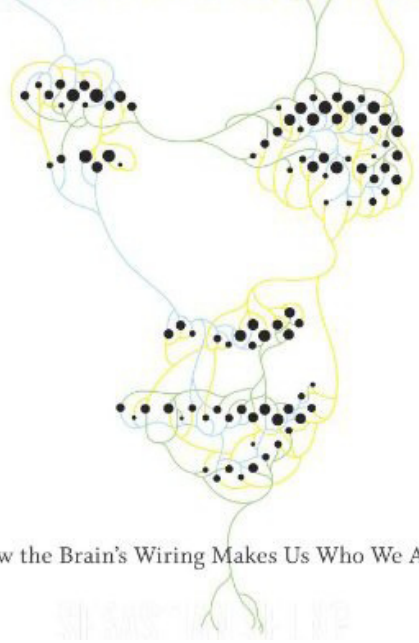
If **2** wave shapes are similar (accounting for delay) the physical processes are 'connected'

Minimal "Distortion" Implicit

Limitation: **Pairwise** Analysis

Popular Expositions

CONNECTOME

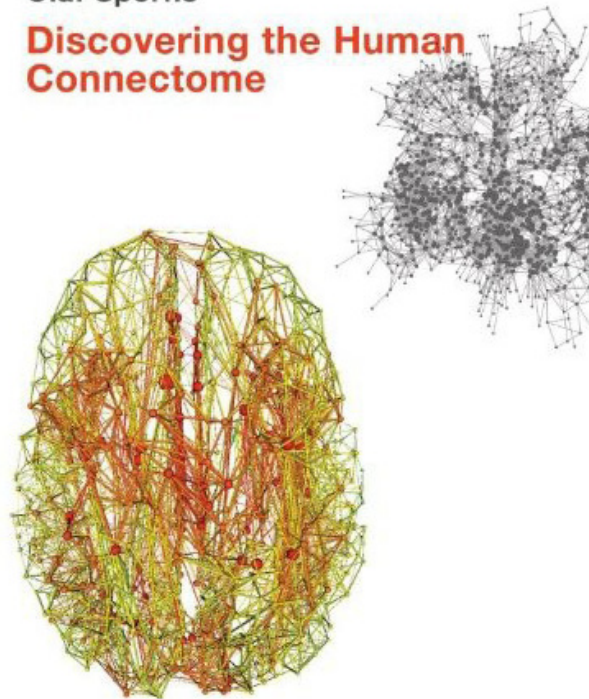


How the Brain's Wiring Makes Us Who We Are

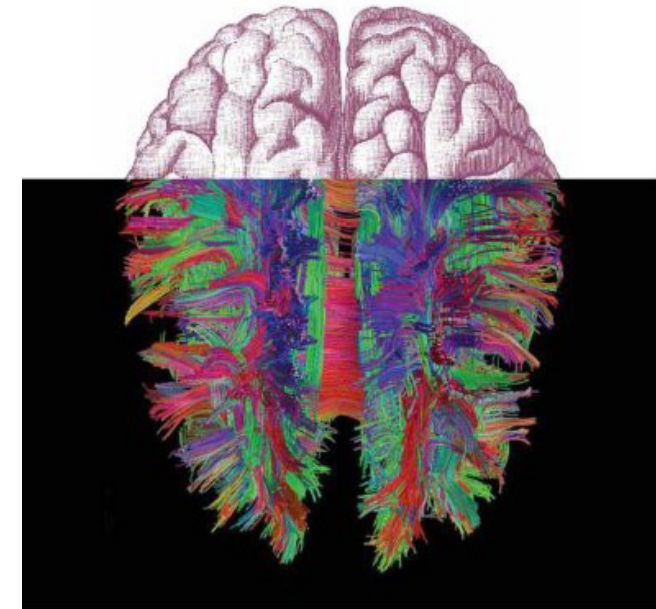
Correlation Centered

Olaf Sporns

**Discovering the Human
Connectome**



Networks of the Brain



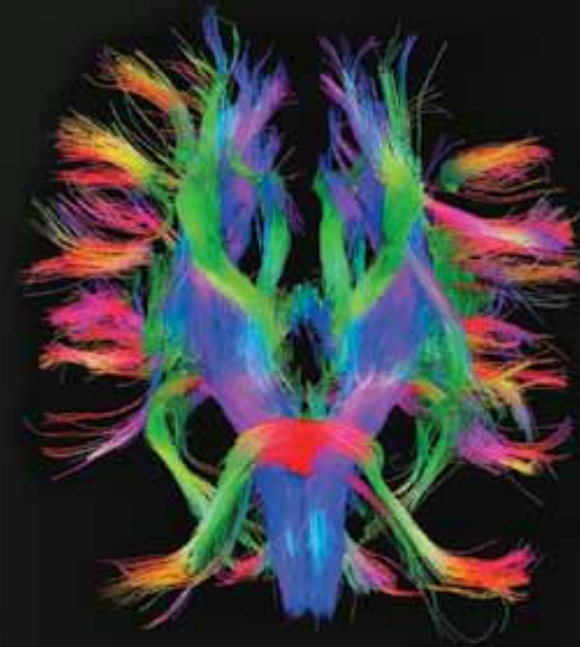
Olaf Sporns



This is Your Brain: Mapping the Connectome

It's been 20 years since Francis Crick and Edward Jones, in the midst of the so-called Decade of the Brain, lamented science's lack of even a basic understanding of human neuroanatomy. "Clearly what is needed for a modern human brain anatomy is the introduction of some radically new techniques," the pair wrote in 1993. Clearly, researchers were listening. Today, they are using novel technologies and automation to map neural circuitry with unparalleled resolution and completeness. The NIH has dedicated nearly \$40 million to chart the wiring of the human brain, and the Allen Brain Institute has poured in millions more to map the mouse brain. The data will take years to compile, and even longer to understand. But the results may reveal nothing less than the nature of human individuality. As MIT neuroscientist Sebastian Seung writes, "You are more than your genes. You are your connectome."

By Jeffrey M. Perkel



When Seung says in *Connectome: How the Brain's Wiring Makes Us Who We Are*, "You are your connectome," what he means is that neural connectivity is like a fingerprint. Each person has their own unique blend of genetics, environmental influences, and life experience. Those factors influence the detailed circuitry of the brain, such that even identical twins likely differ at the

"Genomes are child's play compared with connectomes."

Nevertheless, researchers are making a stab at the problem. From the so-called macroscale of magnetic resonance imaging, to the microscale of electron microscopy, the connectome is slowly coming into focus, one synapse at a time.

The Human Connectome Project

When thinking about the connectome, says Hongkui Zeng, senior director of research science at the **Allen Institute for Brain Science**, think Google Maps. Neuroscientists would like to navigate the brain in

Earlier

Limited to Comparing the activity of Pairs of Structures

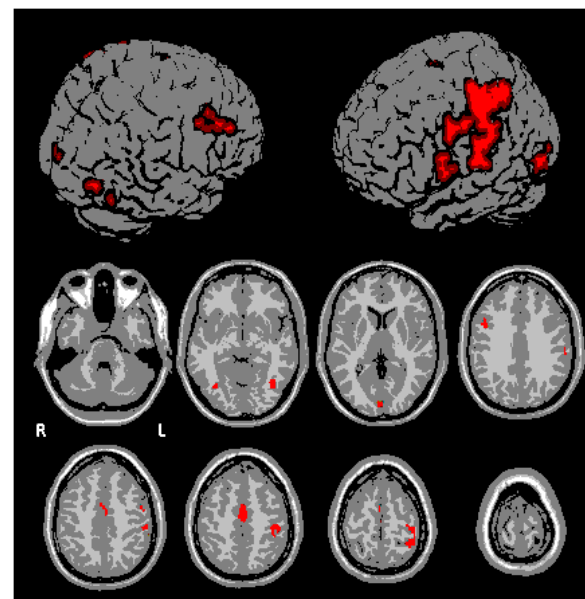


Correlation Analysis

(Coherence Analysis - Fourier Representation)

Search for Structures of Correlated/Coherent Activity

neo-phrenology



Frequency Domain

- Describe Phenomena important over 'bands'

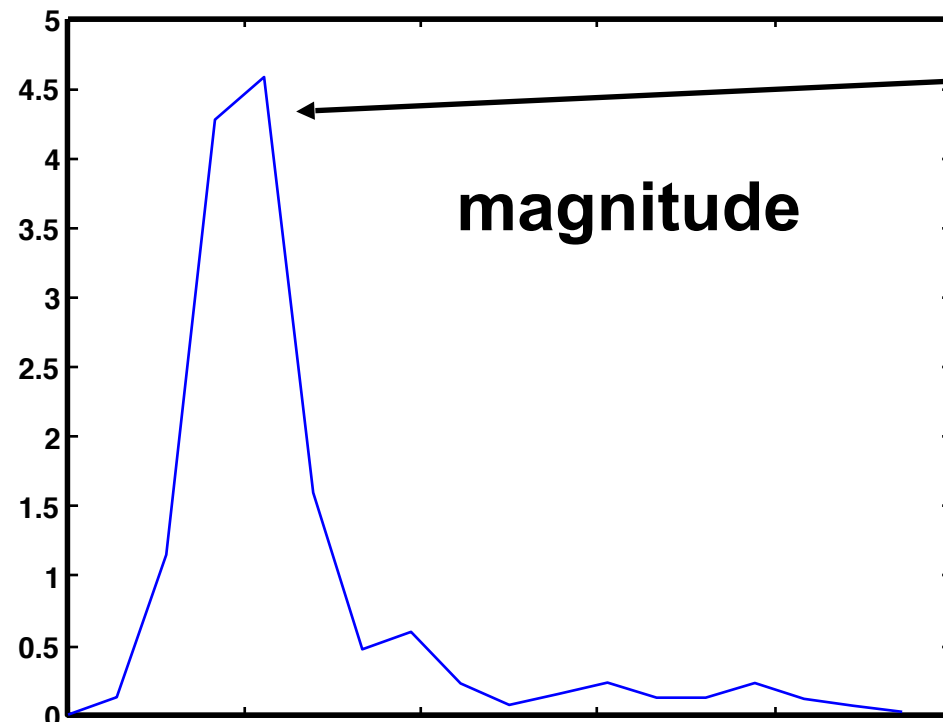
EEG α , β , γ , δ bands

- Robust against Linear Distortion

when different bands propagate at different speeds

- Fourier Transform of Cross Correlation Function

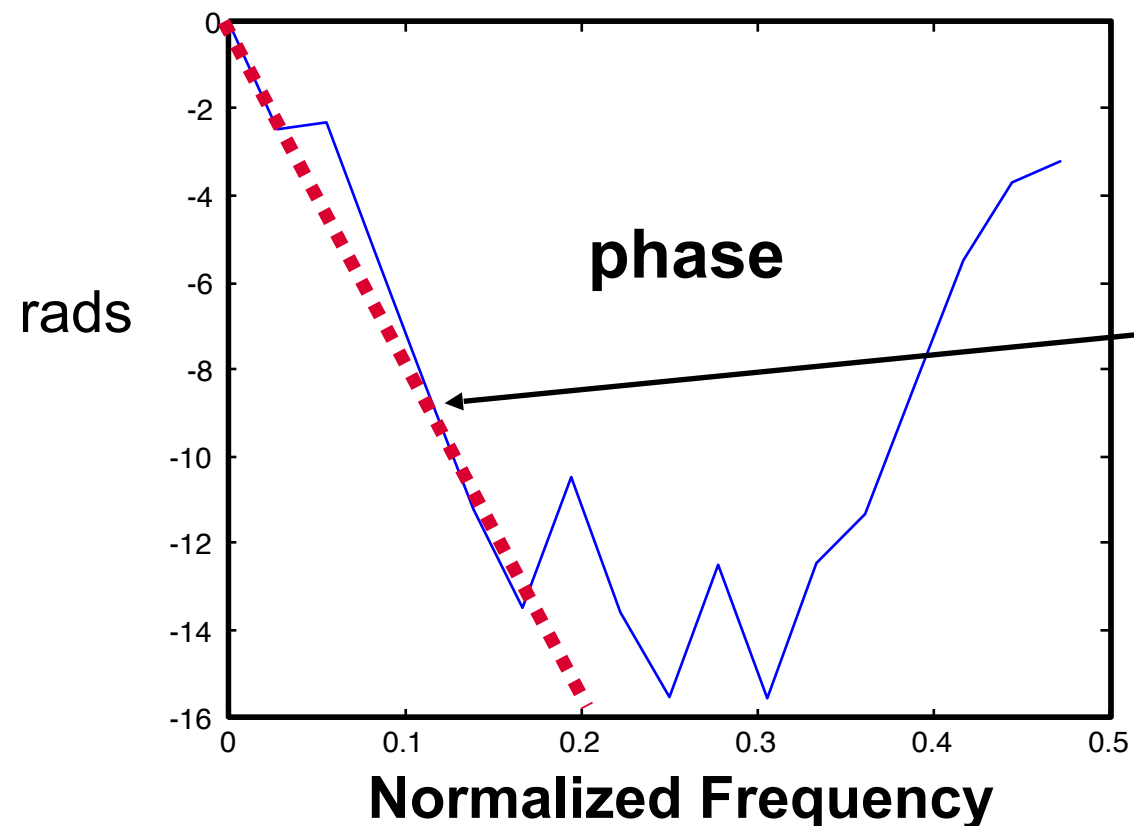
Cross Spectrum



Common frequency

magnitude

Same information as
cross-correlation

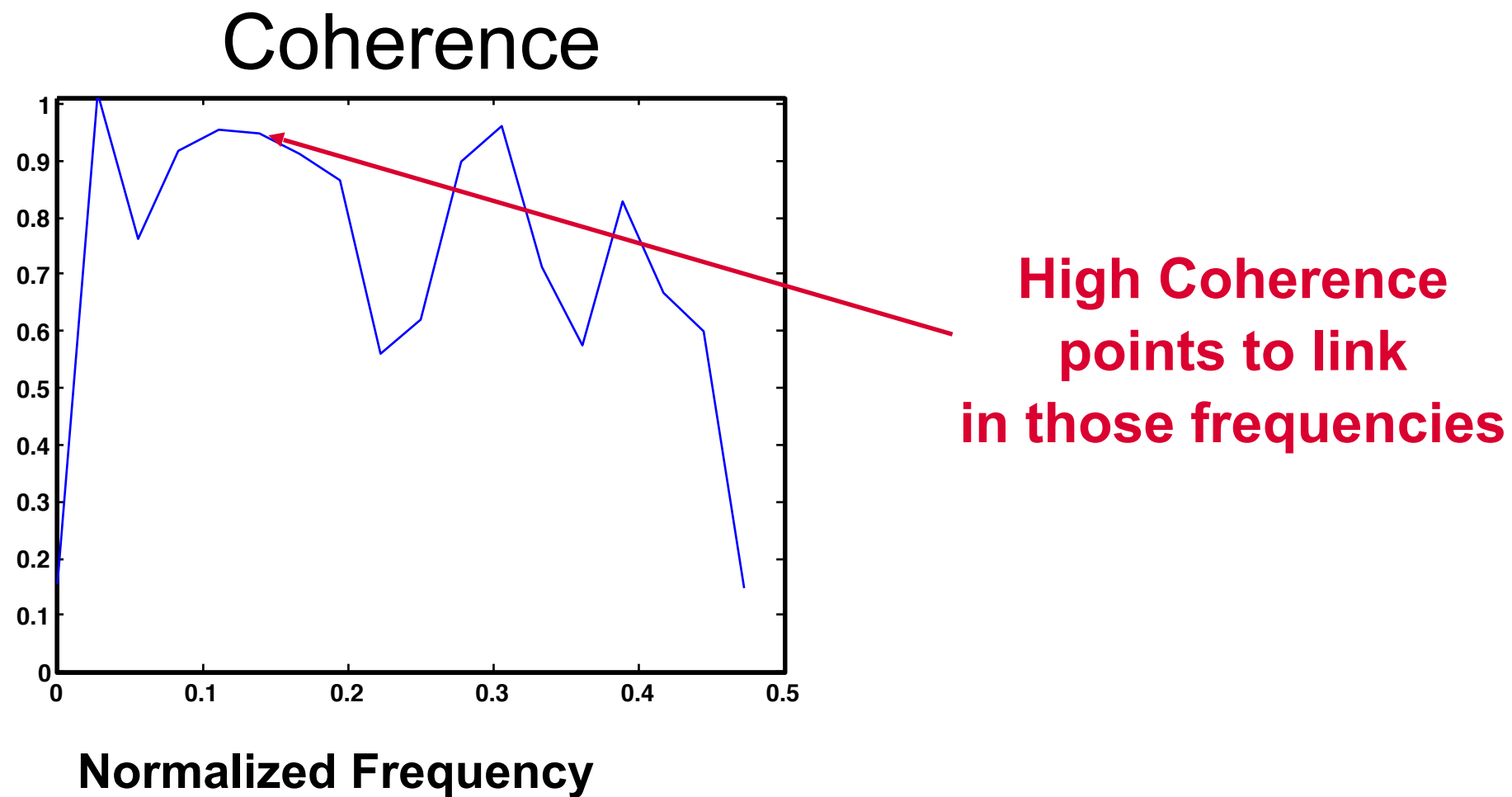


Phase is linear over
the region of common
variance

(Nonparametric Estimates)

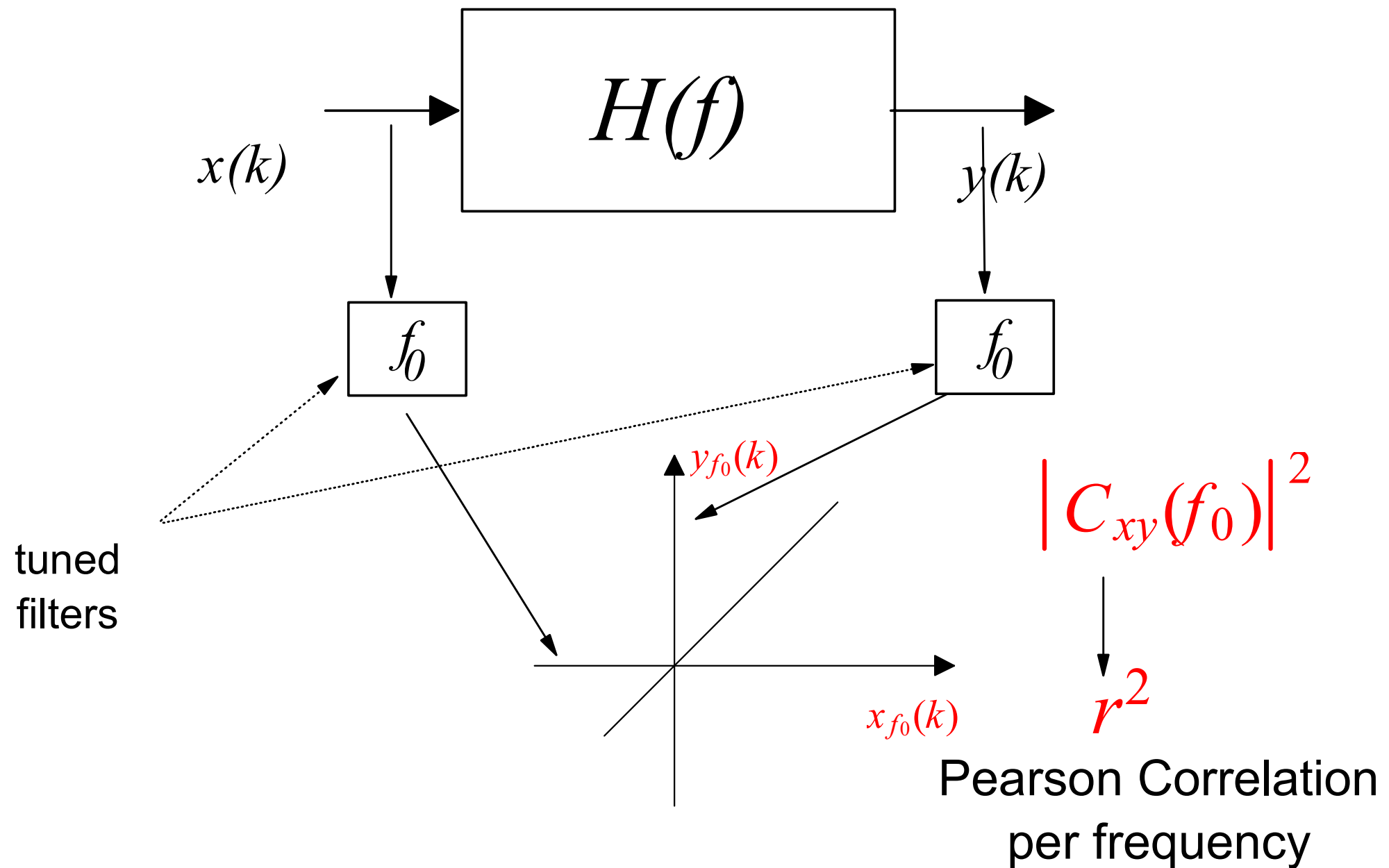
Coherence Function: normalized cross-spectrum

- Linear relationships
- Synchronization between structures



Nonparametric Estimate

Interpretation:

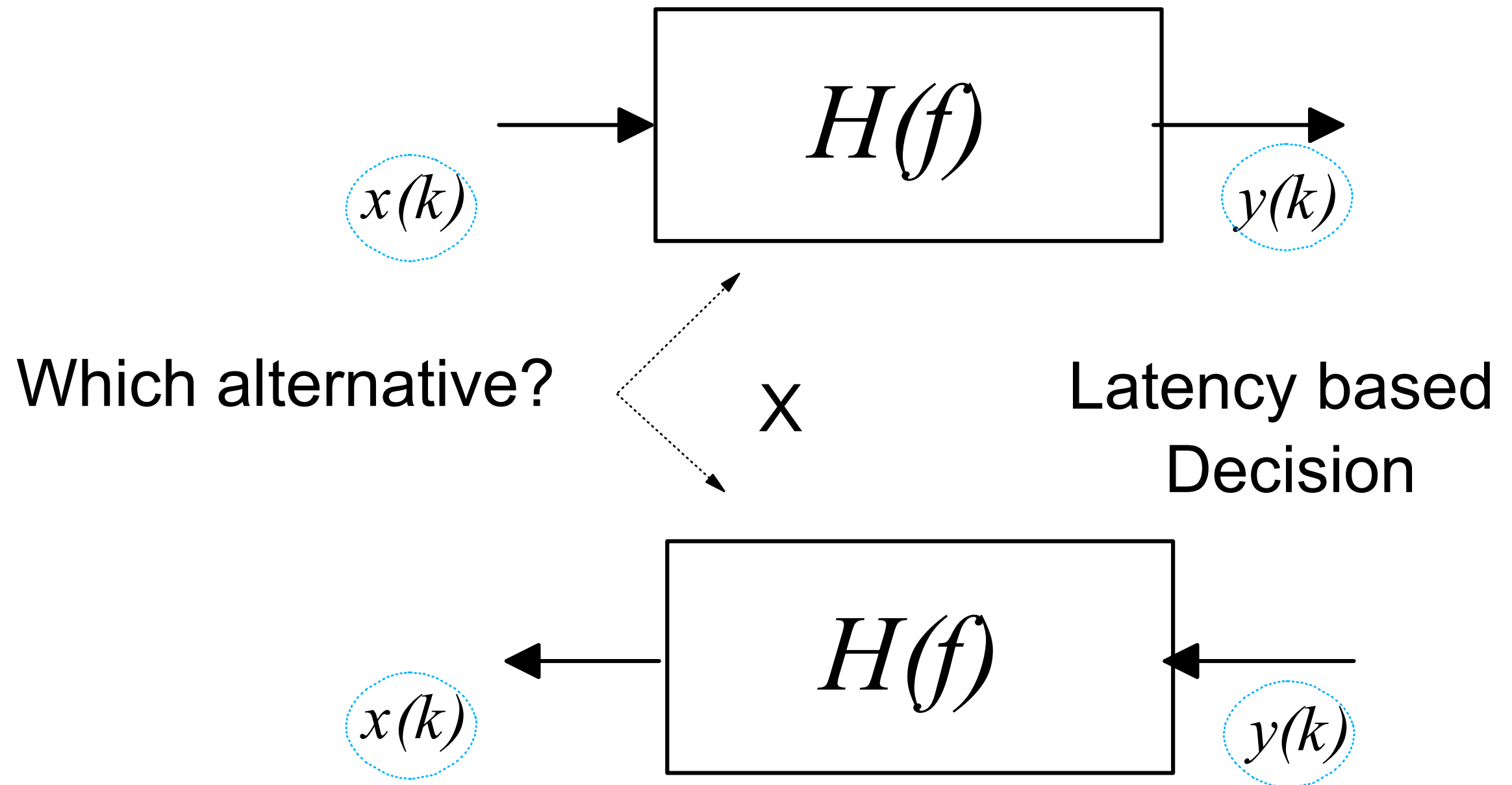


$$|C_{xy}(f)|^2 < 1$$

Remarks:

- nonlinear link
- uncorrelated additive noise

"Internal" $H(f)$ Structure



M.A.L. Nicolelis (Ed.)
Progress in Brain Research, Vol. 130
© 2001 Elsevier Science B.V. All rights reserved

CHAPTER 3

Overcoming the limitations of correlation analysis for many simultaneously processed neural structures

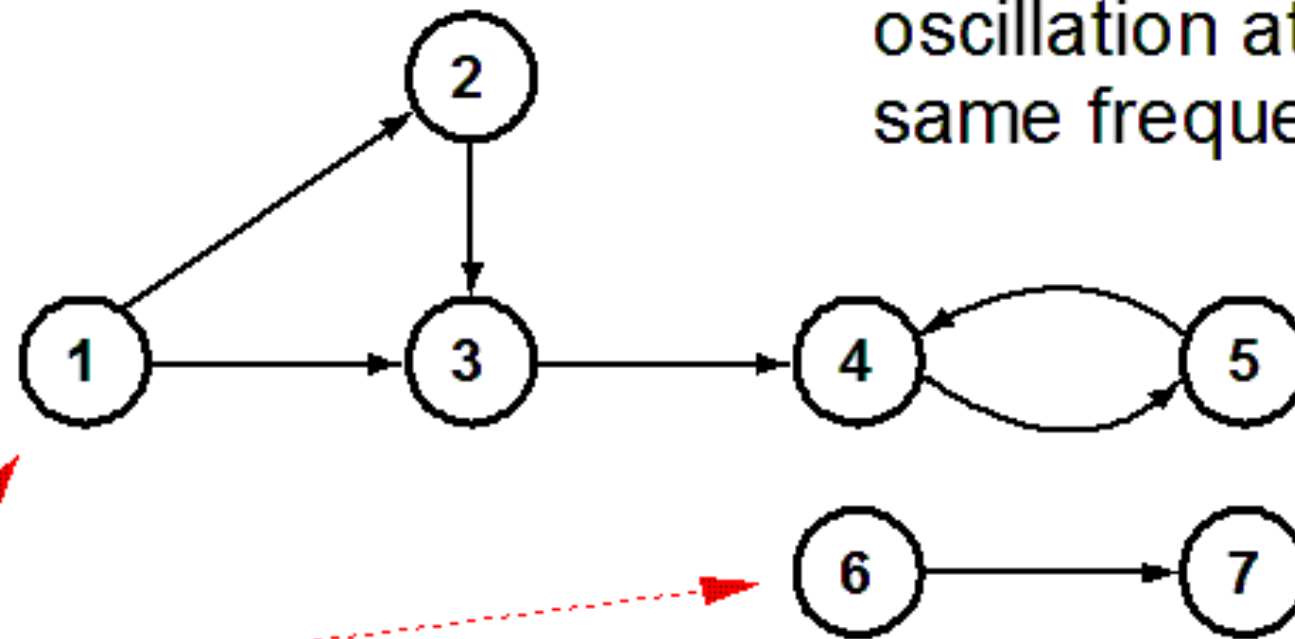
Luiz A. Baccalá^{1,*} and Koichi Sameshima²

¹ *Telecommunications and Control Engineering Department, Escola Politécnica, Av. Prof. Luciano Gualberto, Trav. 3, #158, University of São Paulo, São Paulo, SP, CEP 05508-900, Brazil*

² *Disc. Medical Informatics and Functional Neurosurgery Laboratory, School of Medicine, University of São Paulo, São Paulo, Brazil*

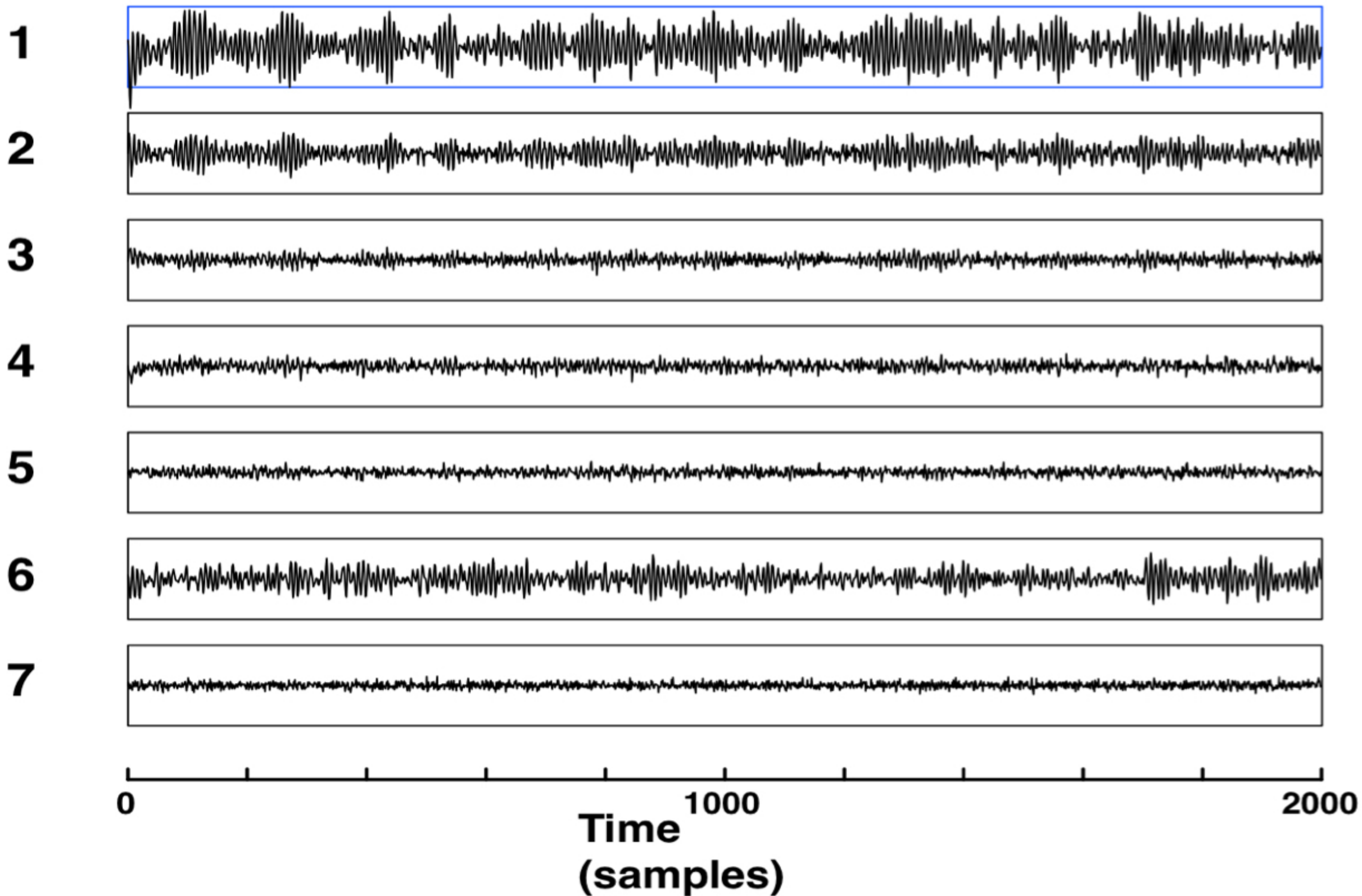
"Toy" Model

Feedback
produces
oscillation at the
same frequency

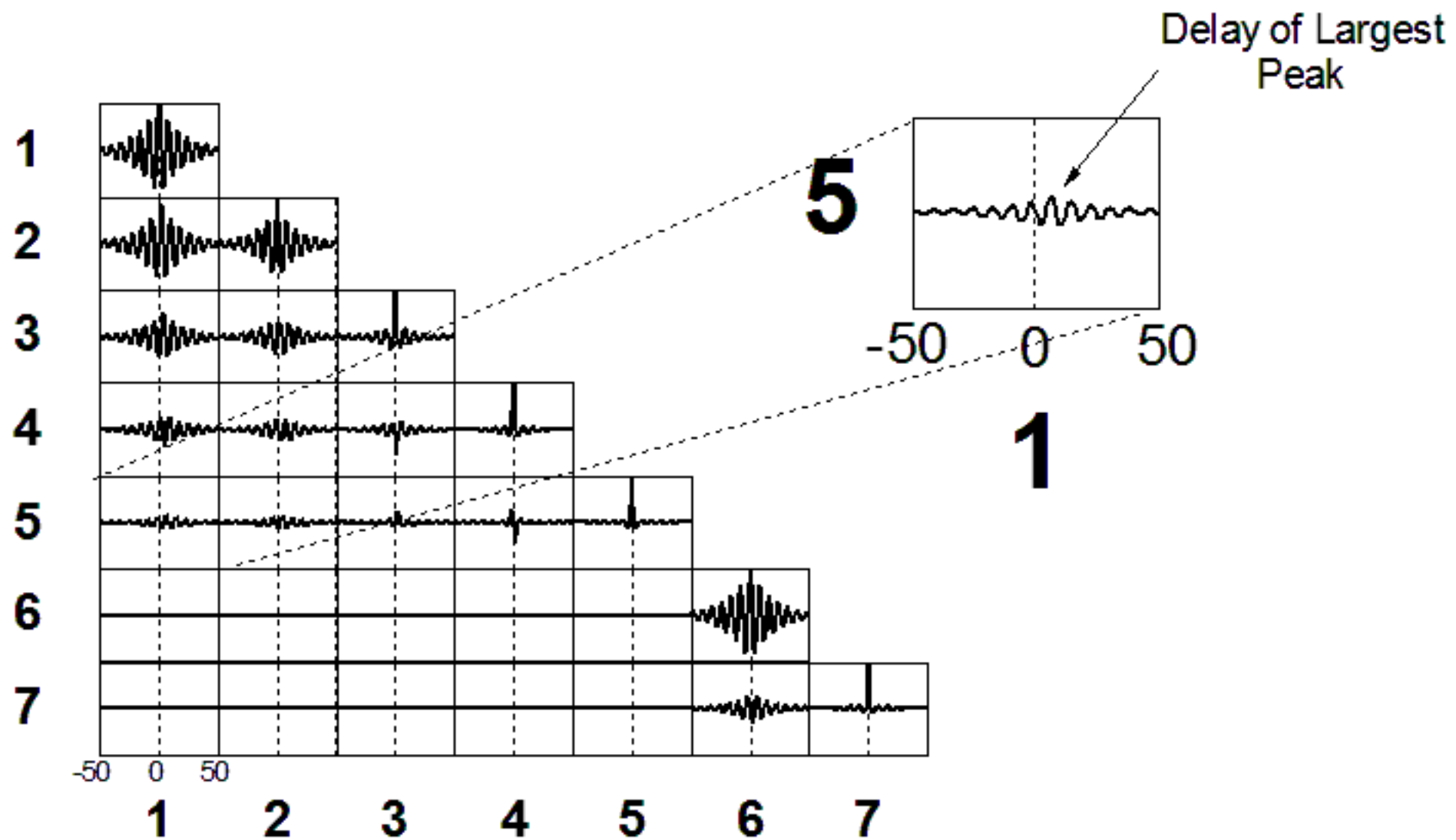


Same Frequency
oscillators

Toy Model Simulation Time Series

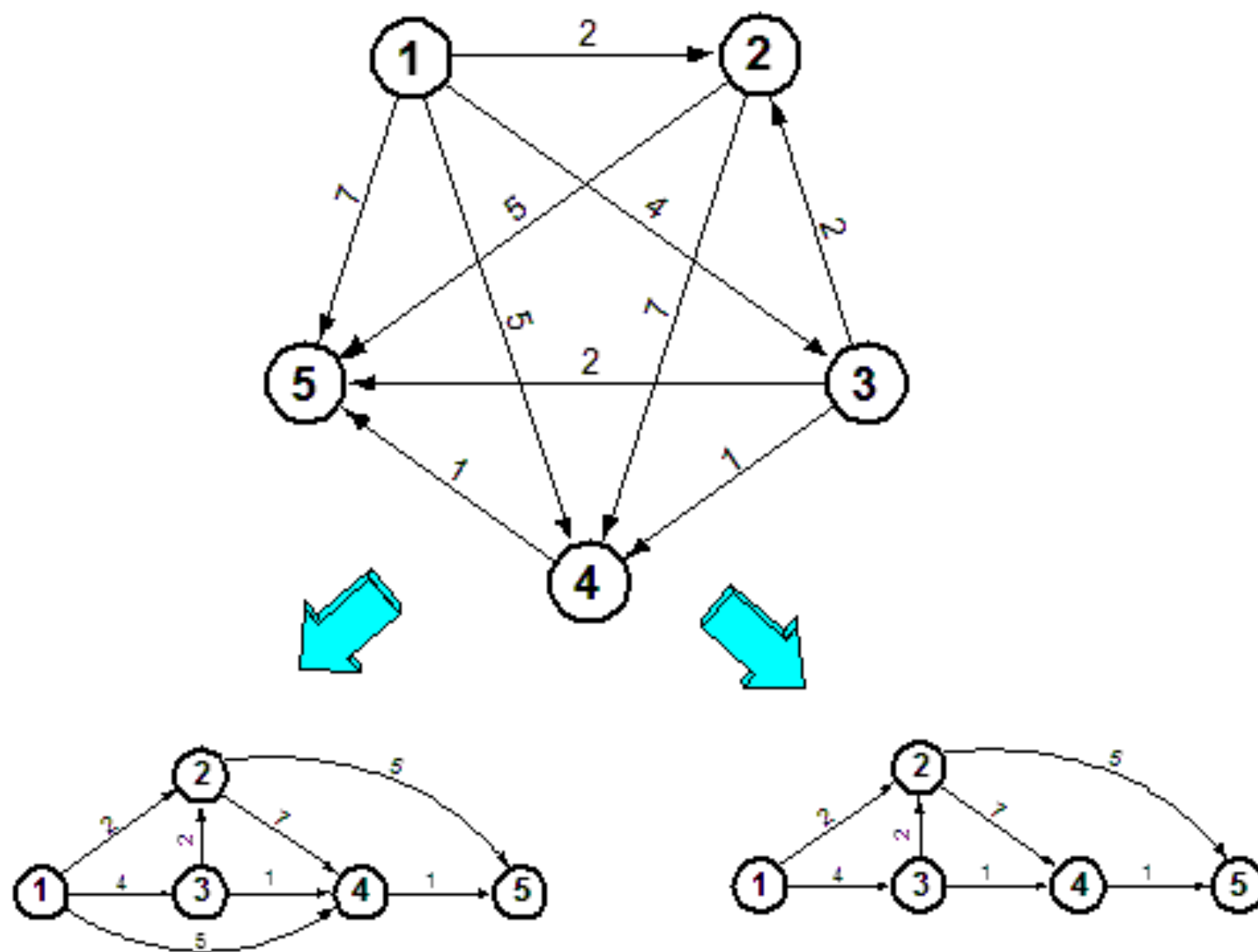


Latency Measurement



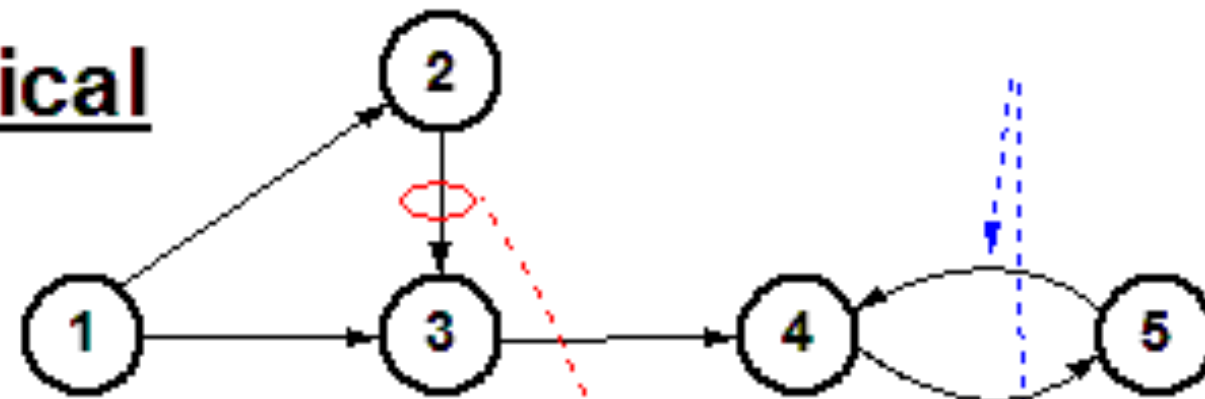
Cross-Correlation

Toy Model Latency Measurements



Comparison between theoretical and latency-based structural inference

Theoretical

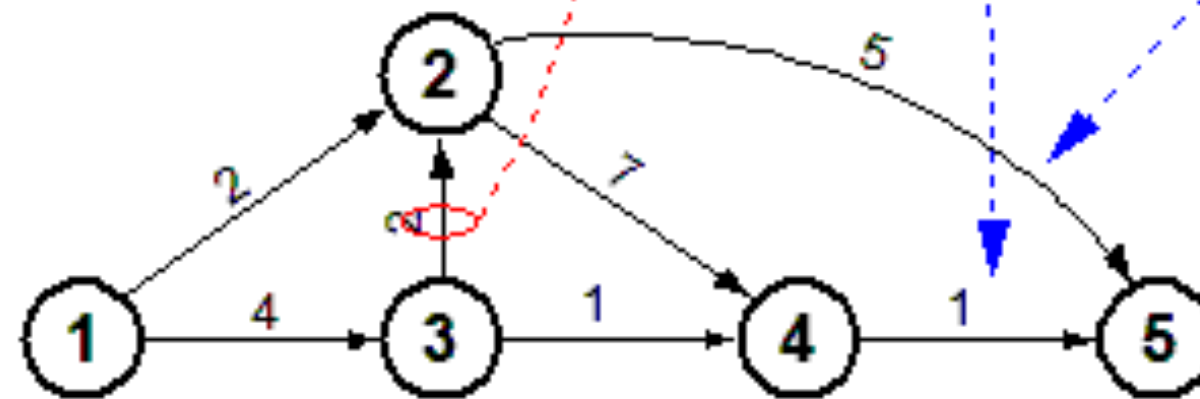


undetected feedback

inverted signal flow

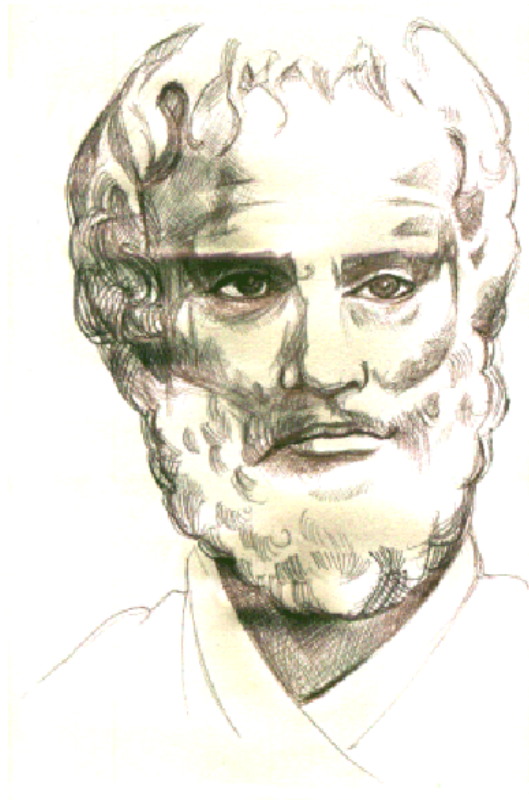
Nonexisting connection

Inferred



Granger-Causality Based Connectivity

Causality

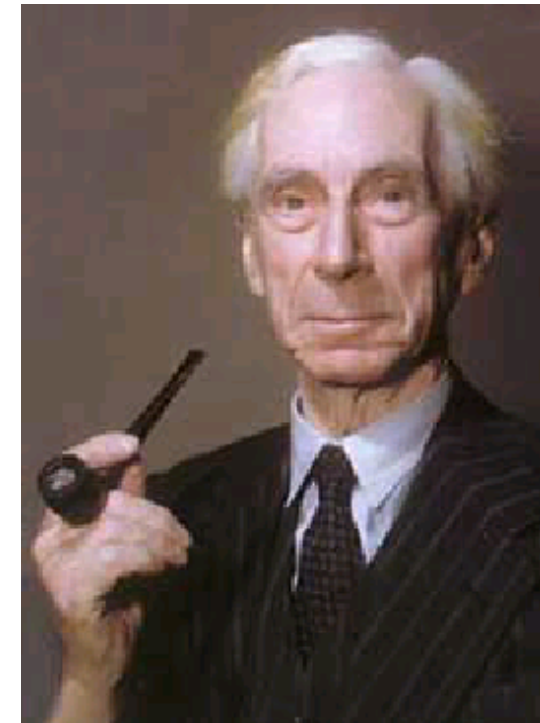


Aristotle
Physics II 3 and Metaphysics V 2.
~ 350 BC

**Predecessor is
responsible for
(generates)
successor**



Hume
An Enquiry Concerning Human
Understanding (1758)

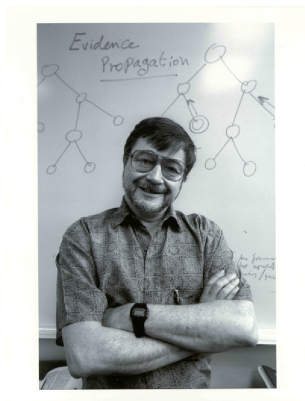


Bertrand Russell
On the Notion of Cause
(1913)

Causality



Bayes ~1763



Pearl, J., **Causality** - 1999

**Smoke and Cancer?
Probabilistic Relationship**

Bayesian Networks ~ 1988

**Physical Systems
have no response
prior to excitation**



**Kramers-
Kronig
~ 1920**



**Clive Granger
(~1969)**

Nobel Econ. 2003

Consistent Temporal Precedence

Prediction Improvement

Causality

Bayesian

Not necessarily temporal

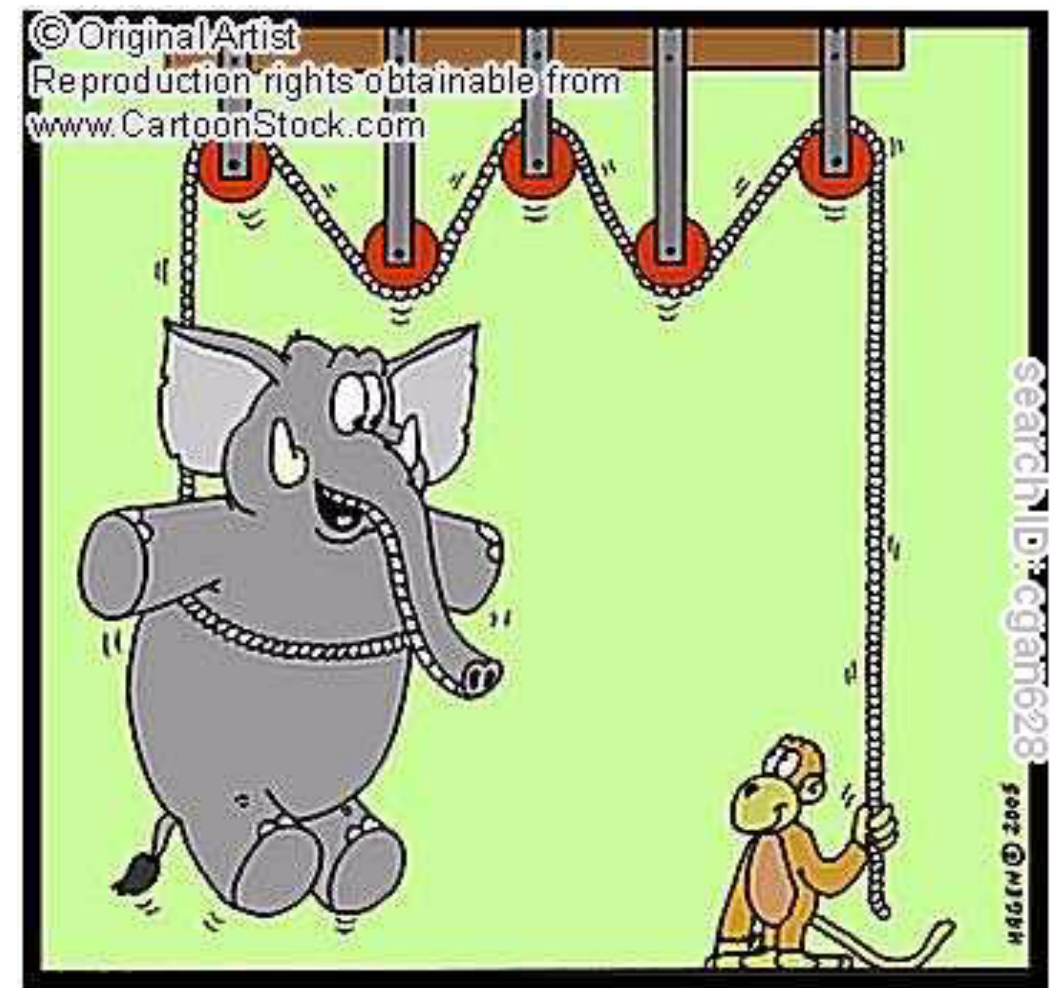
Many "samples" - population based

Physical

Intervention based

Granger

Observational
Temporal



Alright, alright you've won your bet:
You can lift me with one hand.

Causality

Sublata causa, tollitur effectus

Do away with cause, suppress its the effect

requires manipulation by the experimenter

a binary relation - who is the culprit?

Granger Causality

1969 (in econometrics)

Saito & Harashima (1981) (in neuroscience)

- Observational only - **non-interventional**
- Originally uses bivariate time series analysis
 - based on prediction improvement
- Coupling Directionality

$$x(n) = \underbrace{\sum_{k=1}^p a_{xx}(k)x(n-k)}_{x's\ past} + \underbrace{\sum_{k=1}^p a_{xy}(k)y(n-k)}_{y's\ past} + \underbrace{\varepsilon(n)}_{\text{prediction error}}$$

$a_{xy}(k) = 0$ No G-cause condition for all k

Consistent Temporal Precedence

2 Time Series

$$x(n) = \underbrace{\sum_{k=1}^p a_{xx}(k)x(n-k)}_{\substack{\text{past of} \\ x}} + \underbrace{\sum_{k=1}^p a_{xy}(k)y(n-k)}_{\substack{\text{past of} \\ y}} + \underbrace{\varepsilon(n)}_{\text{error}}$$

$a_{xy}(k) = 0$ means lack of connection

Y is "causal" to **X** only if its past significantly influences the present of **X**

(Granger causality)

Granger Causality is unreciprocal

Y is "causal" to X

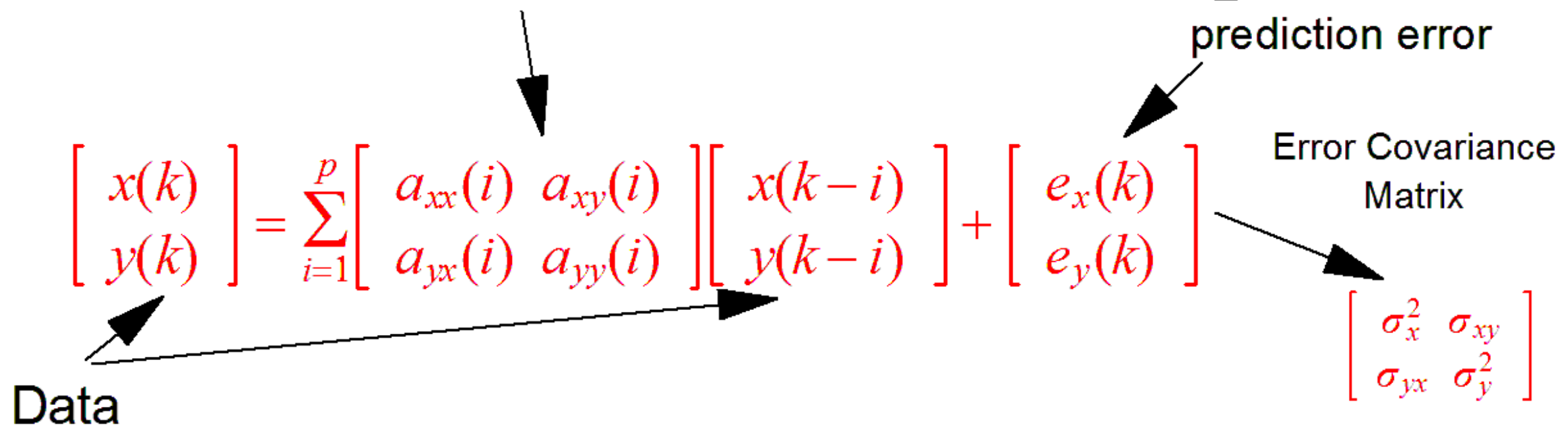
does not mean

X is "causal" to Y

Unlike correlation which is reciprocal

Two Time Series

Least Squares Adjustment



Granger causality = testing if prediction error reduction for the other time series when both are modelled compared to separate models

$$x(k) = \sum_{i=1}^p \beta_x(i) x(k-i) + \varepsilon_x(k)$$

Comparing

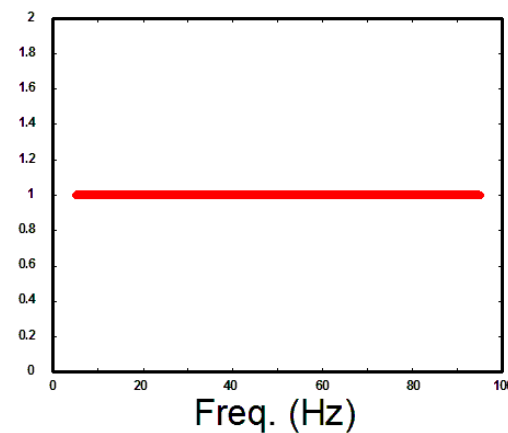
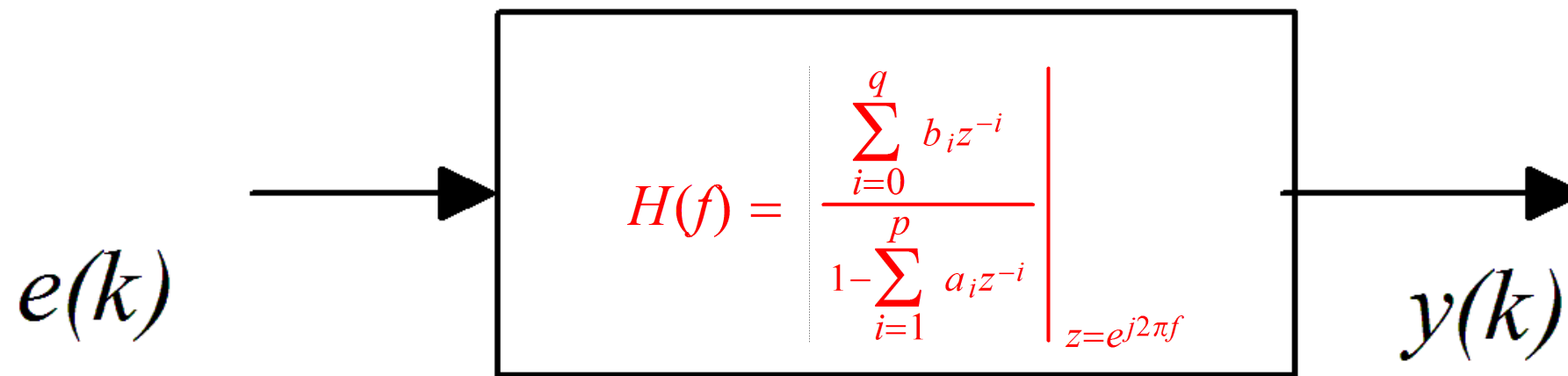
$$VAR(e_x) \ll VAR(\varepsilon_x) \Leftrightarrow a_{xy}(k) \neq 0$$

Frequency Domain Interpretation

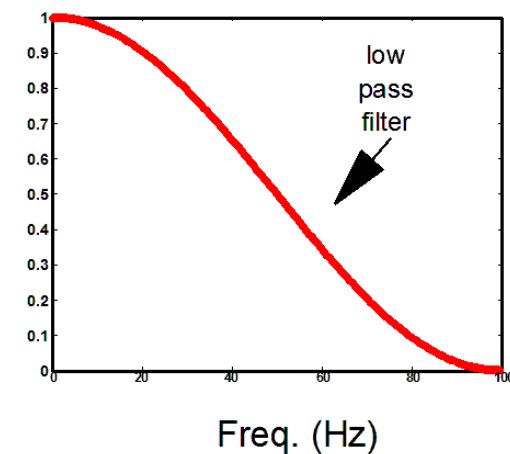
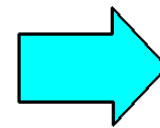
Univariate Parametric Representation

white noise (innovations)

$$y(k) = \underbrace{\sum_{i=1}^p a_i y(k-i)} + \underbrace{\sum_{i=0}^q b_i e(k-i)}$$



White Noise



Filtered White Noise

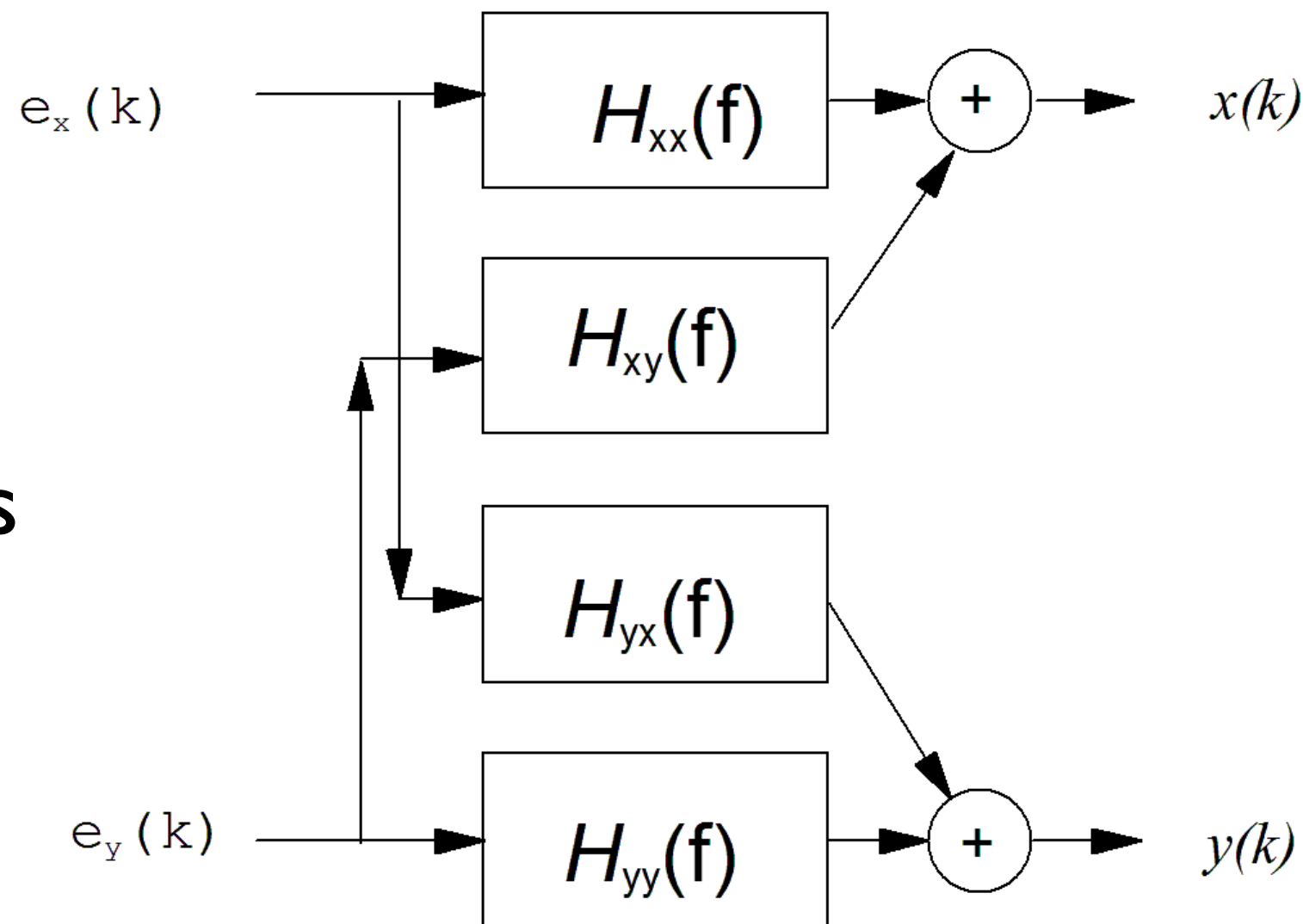
$$S_y(f) = |H(f)|^2 \sigma_e^2$$

Canonical Decomposition

Spectral Density Matrix

$$\begin{bmatrix} S_x(f) & S_{yx}(f) \\ S_{xy}(f) & S_y(f) \end{bmatrix} = \begin{bmatrix} H_{xx}(f) & H_{xy}(f) \\ H_{yx}(f) & H_{yy}(f) \end{bmatrix} \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix} \begin{bmatrix} H_{xx}(f) & H_{xy}(f) \\ H_{yx}(f) & H_{yy}(f) \end{bmatrix}^H$$

Innovations



Spectral Density Matrix Factorization

Least squares fit

$$\begin{bmatrix} x(k) \\ y(k) \end{bmatrix} = \sum_{i=1}^p \begin{bmatrix} a_{xx}(i) & a_{xy}(i) \\ a_{yx}(i) & a_{yy}(i) \end{bmatrix} \begin{bmatrix} x(k-i) \\ y(k-i) \end{bmatrix} + \begin{bmatrix} e_x(k) \\ e_y(k) \end{bmatrix}$$

Residue covariance

$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix}$$

Data

$$\begin{bmatrix} H_{xx}(f) & H_{xy}(f) \\ H_{yx}(f) & H_{yy}(f) \end{bmatrix} = \begin{bmatrix} 1 - \sum_{i=1}^p a_{xx}(i)e^{-j2\pi fi} & -\sum_{i=1}^p a_{xy}(i)e^{-j2\pi fi} \\ -\sum_{i=1}^p a_{yx}(i)e^{-j2\pi fi} & 1 - \sum_{i=1}^p a_{yy}(i)e^{-j2\pi fi} \end{bmatrix}^{-1}$$

$$|f| < 1/2$$

Parametric Coherence Representation

(special case)

$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy} = 0 \\ \sigma_{yx} = 0 & \sigma_y^2 \end{bmatrix}$$

$S_{xy}(f)$



**Depends on the interaction
in both directions**

$$C_{xy}(f) = \frac{\sigma_x H_{xx}(f) H_{yx}^*(f) + \sigma_y H_{xy}(f) H_{yy}^*(f)}{\sqrt{S_x(f)} \sqrt{S_y(f)}}$$

Coerência Direcionada

(Saito & Harashima 1981)

$$\gamma_{xy}(f) = \frac{\sigma_{yy}H_{xy}(f)}{\sqrt{\sigma_{yy}^2|H_{xy}(f)|^2 + \sigma_{xx}^2|H_{xx}(f)|^2}} = \frac{\sigma_{yy}H_{xy}(f)}{\sqrt{S_x(f)}}$$

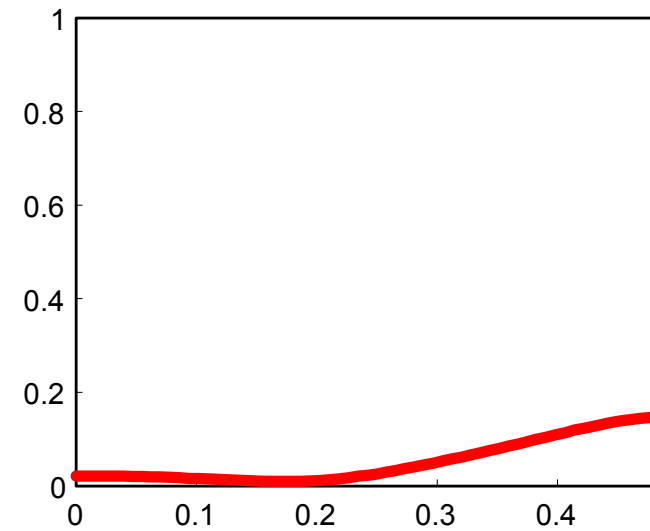
Reason for the name: a factor into which coherence can be decomposed

DTF: Directed Transfer Function (Franaczuc et. al 1985)

$$DTF_{xy}(f) = \frac{H_{xy}(f)}{\sqrt{|H_{xy}(f)|^2 + |H_{yy}(f)|^2}}$$

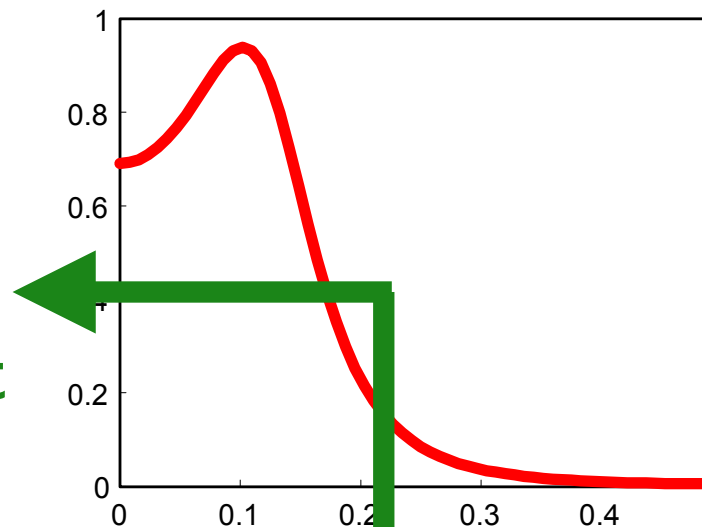
Causal Relations using directed coherence $\gamma_{ij}(f)$

Sun



Mel

target



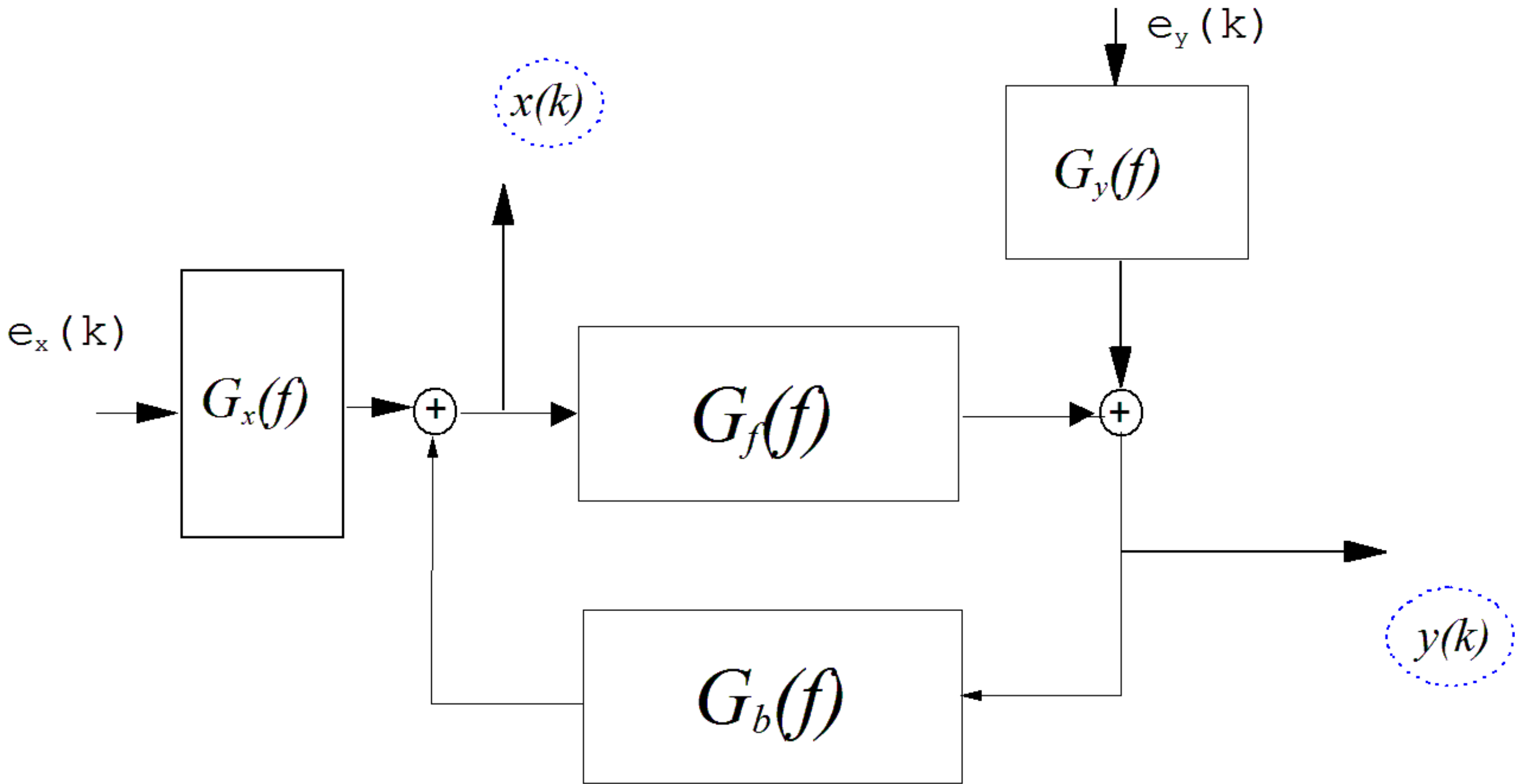
Sun

source



Mel

Equivalent Model



$$a_{xy}(r) = 0 \Leftrightarrow H_{xy}(f) = 0 \Leftrightarrow G_b(f) = 0 \quad (\text{Sims 1982})$$

Key Technology: Multichannel Neural Signal Recording

Modalities

EEG →

LFPs



voxel
based

MEG

fMRI

Multi-Single Unit Data



In the Multivariate Case

How does Granger causality generalize?

$$N > 2$$

The same data set has

2 different

possible representations

Time Series Representations

Autoregressive (AR)

frequency domain

PDC

$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \sum_{r=1}^{+\infty} \begin{bmatrix} a_{11}(r) & a_{12}(r) & \dots & \dots & a_{1N}(r) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & a_{ij}(r) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N1}(r) & \dots & \dots & \dots & a_{NN}(r) \end{bmatrix} \begin{bmatrix} x_1(k-r) \\ \vdots \\ x_N(k-r) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ \vdots \\ w_N(k) \end{bmatrix}$$

Moving Average (MA)

frequency domain

DTF

$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \sum_{r=0}^{+\infty} \begin{bmatrix} h_{11}(r) & h_{12}(r) & \dots & \dots & h_{1N}(r) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & h_{ij}(r) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{N1}(r) & \dots & \dots & \dots & h_{NN}(r) \end{bmatrix} \begin{bmatrix} w_1(k) \\ \vdots \\ w_N(k) \end{bmatrix}$$

Representation One

Moving Average (MA)

frequency domain

DTF

$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \sum_{r=0}^{+\infty} \begin{bmatrix} h_{11}(r) & h_{12}(r) & \dots & \dots & h_{1N}(r) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & h_{ij}(r) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{N1}(r) & \dots & \dots & \dots & h_{NN}(r) \end{bmatrix} \begin{bmatrix} w_1(k) \\ \vdots \\ w_N(k) \end{bmatrix}$$

Alternative to the multivariate Case

Spectral Density Matrix for more than 2 time series

Spectral Density Matrix \rightarrow

$$S(f) = H(f)\Sigma H^H(f)$$
$$H(f) = \begin{bmatrix} H_{11}(f) & H_{12}(f) & \dots & H_{1N}(f) \\ H_{21}(f) & H_{22}(f) & \dots & H_{2N}(f) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(f) & H_{N2}(f) & \dots & H_{NN}(f) \end{bmatrix}$$
$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22}^2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \dots & \dots & \sigma_{NN}^2 \end{bmatrix}$$

Directed Transfer Function (Kaminski & Blinowska, 1991)

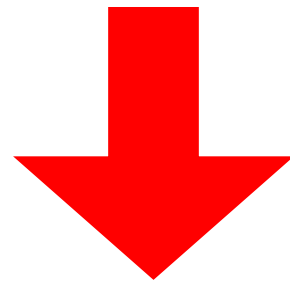
$$D TF_{ij}(f) = \frac{H_{ij}(f)}{\sqrt{\sum_{j=1}^N |H_{ij}(f)|^2}}$$

Directed Coherence (Baccala et al, 1998)

$$\gamma_{ij}(f) = \frac{\sigma_{jj} H_{ij}(f)}{\sqrt{\sum_{j=1}^N \sigma_{jj}^2 |H_{ij}(f)|^2}}$$

Multivariate Autoregressive Time Series Models

$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \sum_{r=1}^p \begin{bmatrix} a_{11}(r) & a_{12}(r) & \dots & \dots & a_{1N}(r) \\ \vdots & \vdots & \vdots & a_{ij}(r) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N1}(r) & \dots & \dots & \dots & a_{NN}(r) \end{bmatrix} \begin{bmatrix} x_1(k-r) \\ \vdots \\ x_N(k-r) \end{bmatrix}$$



$$\bar{A}_{ij}(f) = \begin{cases} 1 - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi fr}, & \text{if } i = j \\ - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi fr}, & \text{otherwise} \end{cases}$$

Obtain by model fitting

$$\mathbf{H}(f) = \begin{bmatrix} H_{11}(f) & H_{12}(f) & \dots & H_{1N}(f) \\ H_{21}(f) & H_{22}(f) & \dots & H_{2N}(f) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(f) & H_{N2}(f) & \dots & H_{NN}(f) \end{bmatrix}$$

Inversion Duality



$$\mathbf{A}(f) = \begin{bmatrix} A_{11}(f) & A_{12}(f) & \dots & A_{1N}(f) \\ A_{21}(f) & A_{22}(f) & \dots & A_{2N}(f) \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1}(f) & A_{N2}(f) & \dots & A_{NN}(f) \end{bmatrix}$$

$$\mathbf{H}(f) = [\bar{\mathbf{A}}(f)]^{-1}$$

Power Spectral Density Matrix

$$\mathbf{S}(f) = \mathbf{H}(f)\mathbf{\Sigma}\mathbf{H}^H(f)$$

$$\mathbf{H}(f) = \begin{bmatrix} H_{11}(f) & H_{12}(f) & \dots & H_{1N}(f) \\ H_{21}(f) & H_{22}(f) & \dots & H_{2N}(f) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(f) & H_{N2}(f) & \dots & H_{NN}(f) \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22}^2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \dots & \dots & \sigma_{NN}^2 \end{bmatrix}$$

Representation Two

Autoregressive (AR)

frequency domain



PDC

$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \sum_{r=1}^{+\infty} \begin{bmatrix} a_{11}(r) & a_{12}(r) & \dots & \dots & a_{1N}(r) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & a_{ij}(r) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N1}(r) & \dots & \dots & \dots & a_{NN}(r) \end{bmatrix} \begin{bmatrix} x_1(k-r) \\ \vdots \\ x_N(k-r) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ \vdots \\ w_N(k) \end{bmatrix}$$

Partial Coherence

$$|\kappa_{ij}(f)|^2 = \frac{|\mathbf{a}_i^H(f) \Sigma^{-1} \mathbf{a}_j(f)|^2}{(\mathbf{a}_i^H(f) \Sigma^{-1} \mathbf{a}_i(f))(\mathbf{a}_j^H(f) \Sigma^{-1} \mathbf{a}_j(f))}$$

matrix columns

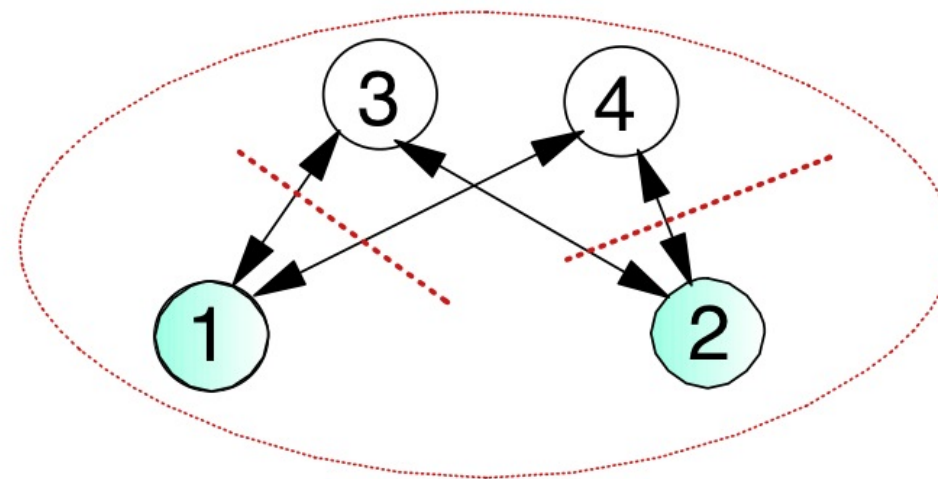
$$\mathbf{a}_i(f) = [\bar{A}_{ij}(f)]$$

$$\bar{A}_{ij}(f) = \begin{cases} 1 - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi fr}, & \text{if } i = j \\ - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi fr}, & \text{otherwise} \end{cases}$$

can be directly computed from
the multivariate autoregressive
model (Baccalá 2001)

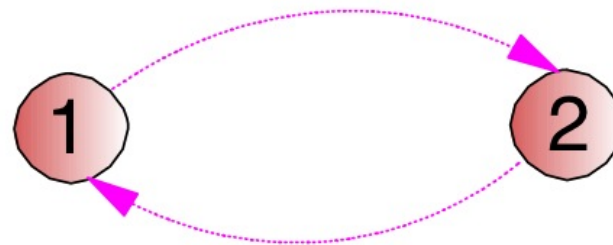
Partial Coherence

Coherence between the resulting series after optimal subtraction of the effect of other N-2 series



**Extraction
of the effect
between the
other series**

Residual Model



$$\begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} = \sum_{r=1}^p \begin{bmatrix} a_{11}(r) & a_{12}(r) & \dots & \dots & a_{1N}(r) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & a_{ij}(r) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N1}(r) & \dots & \dots & \dots & a_{NN}(r) \end{bmatrix} \begin{bmatrix} x_1(k-r) \\ \vdots \\ x_N(k-r) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ \vdots \\ w_N(k) \end{bmatrix}$$



Result:

Given a multivariate autoregressive of the data, the pairwise partial coherence may be written using

$$\bar{A}_{ij}(f) = \begin{cases} 1 - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi fr}, & \text{if } i = j \\ - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi fr}, & \text{otherwise} \end{cases}$$

PDC

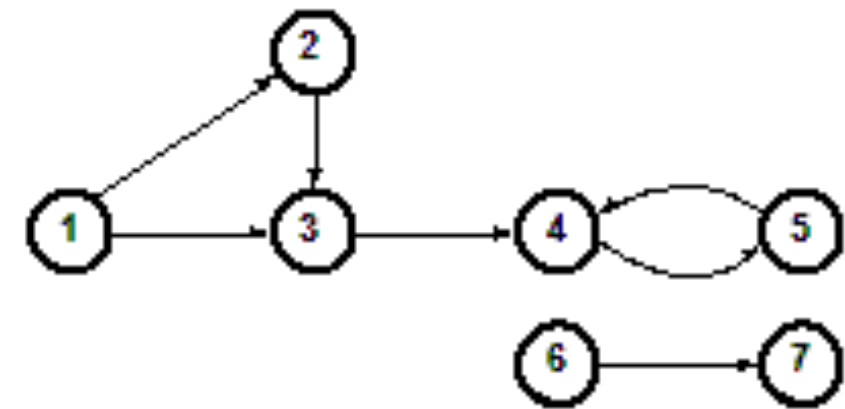
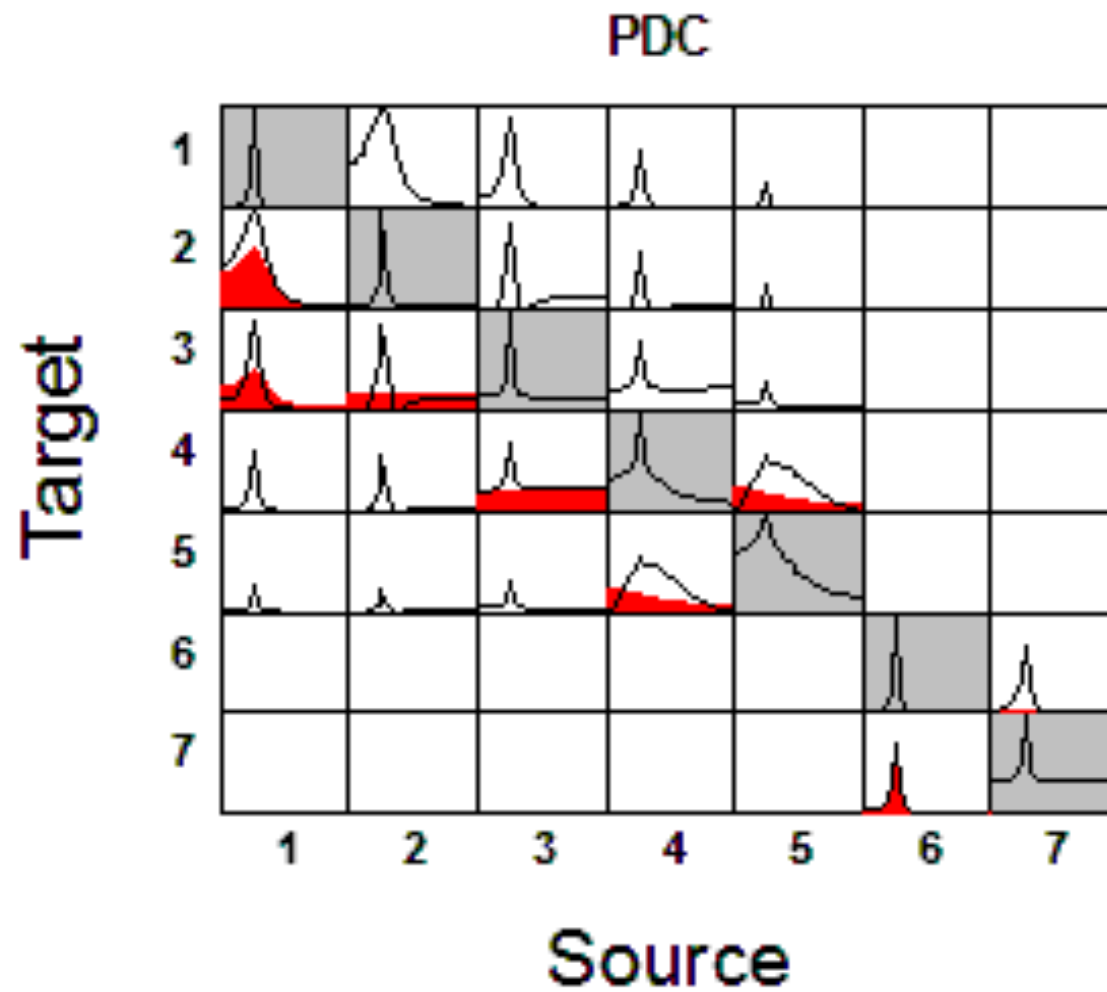
(Partial Directed Coherence - PDC)

PDC from j to i

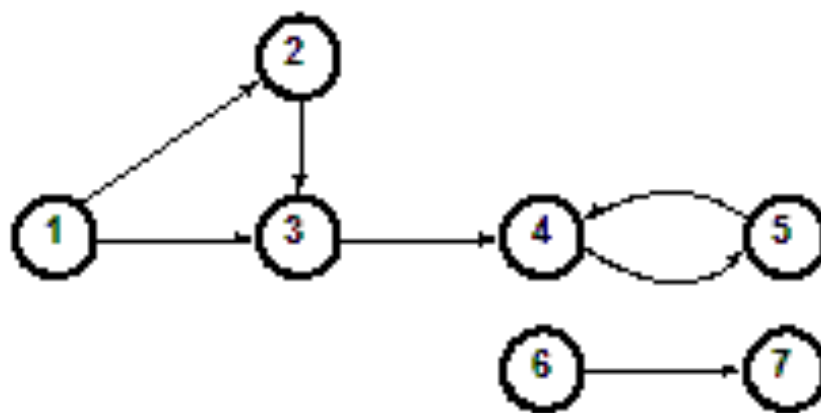
$$\pi_{ij}(f) \triangleq \frac{\bar{A}_{ij}(f)}{\sqrt{\sum_{k=1}^N \bar{A}_{kj}^*(f) \bar{A}_{kj}(f)}}$$

- derived from factoring partial coherence

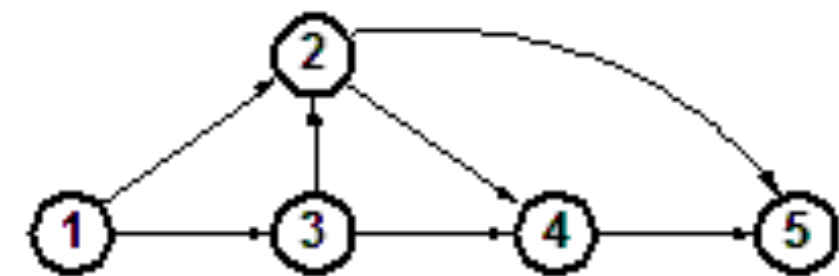
Partial Directed Coherence (PDC)



Deduced



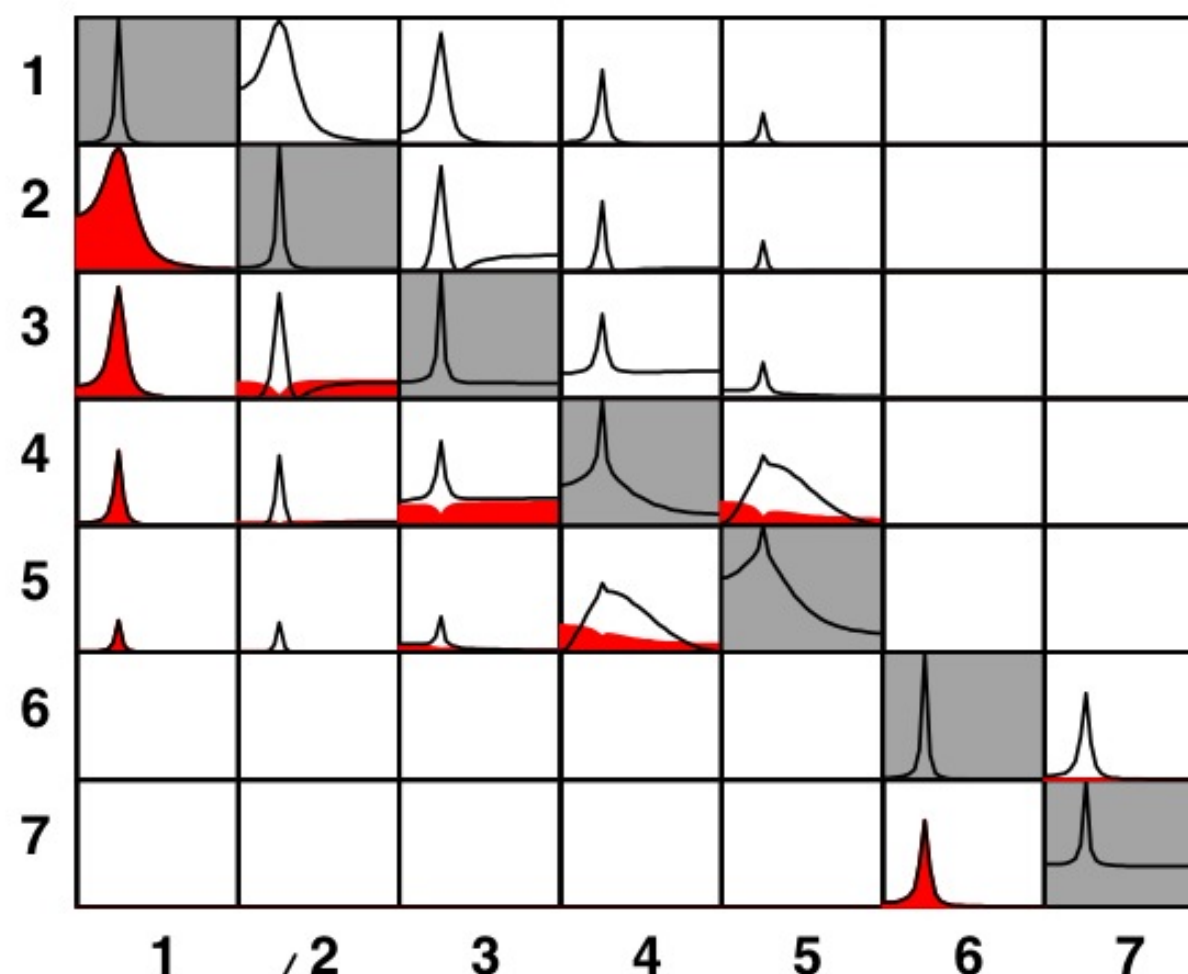
True



Correlation Based Picture

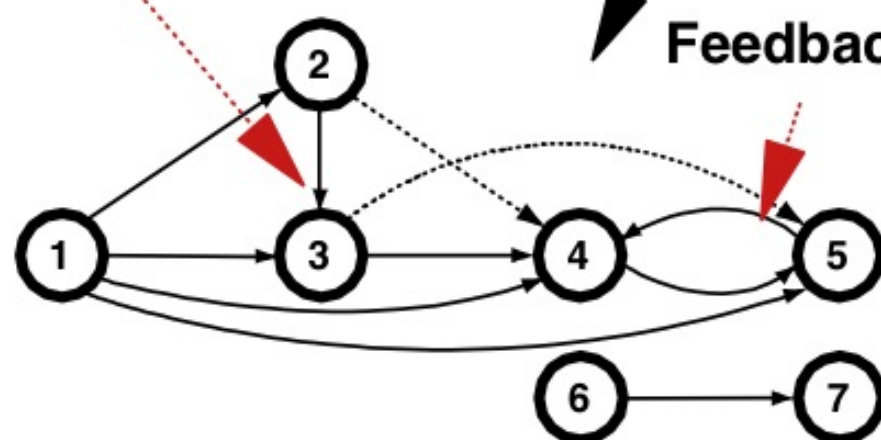
DTF

Target

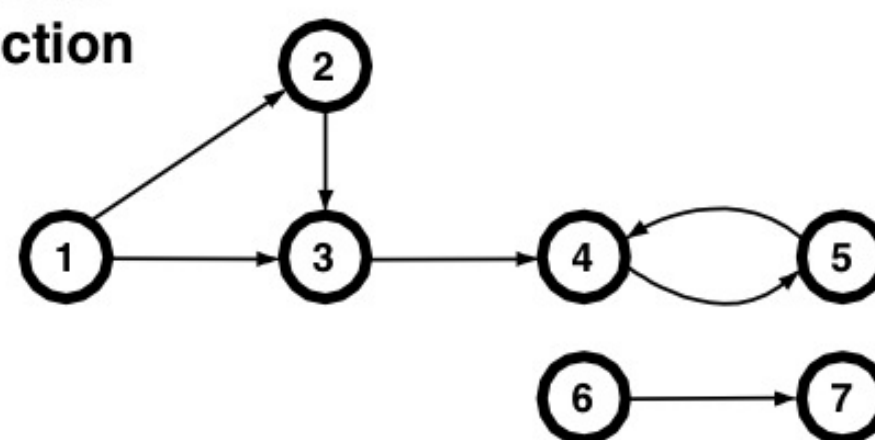


Correct Direction

Source
Feedback Detection



Inferred



True

$$DTF_{ij}(f) = \frac{H_{ij}(f)}{\sqrt{\sum_{j=1}^K |H_{ij}(f)|^2}}$$

Directed Transfer Function (Kaminski & Blinowska, 1991)

$$\gamma_{ij}(f) = \frac{\sigma_{jj} H_{ij}(f)}{\sqrt{\sum_{j=1}^K \sigma_{jj}^2 |H_{ij}(f)|^2}}$$

Directed Coherence
(Baccala et al, 1998)

Information DTF
(Takahashi et al, 2010)



$$iDTF_{ij}(\omega) = \frac{\sigma_{jj} H(\omega)}{\sqrt{h_j^H(\omega) \Sigma_w h_j(\omega)}}$$

$$\pi_{ij}(f) \triangleq \frac{\bar{A}_{ij}(f)}{\sqrt{\sum_{k=1}^N \bar{A}_{kj}^*(f) \bar{A}_{kj}(f)}} \quad \text{PDC 2001}$$

$${}_g\pi_{ij}(f) \triangleq \frac{\frac{1}{\sigma_{ii}} \bar{A}_{ij}(f)}{\sqrt{\sum_{k=1}^K \frac{1}{\sigma_{kk}^2} \bar{A}_{kj}^*(f) \bar{A}_{kj}(f)}} \quad \text{gPDC 2007}$$

Information PDC 2011 \longrightarrow ${}_{\text{IPDC}}\pi_{ij}(\omega) = \frac{\sigma_{ii}^{-1} \bar{A}_{ij}(\omega)}{\sqrt{a_j^H(\omega) \Sigma_w^{-1} a_j(\omega)}}$
iPDC

Multivariate Spectral Time Series

$$S(f) = H(f)\Sigma H^H(f)$$
$$H(f) = \begin{bmatrix} H_{11}(f) & H_{12}(f) & \dots & H_{1N}(f) \\ H_{21}(f) & H_{22}(f) & \dots & H_{2N}(f) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(f) & H_{N2}(f) & \dots & H_{NN}(f) \end{bmatrix}$$
$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22}^2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \dots & \dots & \sigma_{NN}^2 \end{bmatrix}$$

- Coherence (Counterpart to Correlation in Time)
 - Pairwise case
- Factorization
 - DTF
 - PDC

Apply to More than 2 time series

What do they reveal?

- Identical for time series pairs ($N=2$)
DTF/MA=PDC/AR=Granger Causality
- Expose distinct interaction aspects when $N>2$

DTF=PDC

=

Granger Causality

N=2 time series

Illustration of the Differences

DTF/MA

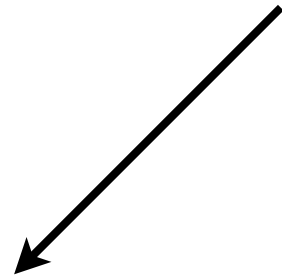
PDC/AR

Information

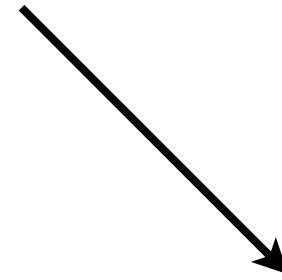
$${}_i\gamma_{ij}(\omega) = \frac{\sigma_{jj}H_{ij}(\omega)}{\sqrt{h_i^H(\omega)\Sigma_w h_i(\omega)}} \quad {}_i\pi_{ij}(\omega) = \frac{\sigma_{ii}^{-1}\bar{A}_{ij}(\omega)}{\sqrt{a_j^H(\omega)\Sigma_w^{-1} a_j(\omega)}}$$

(Takahashi et al, 2010)

The Connectivity Problem



Detection



Characterization

Prevalent Paradigm

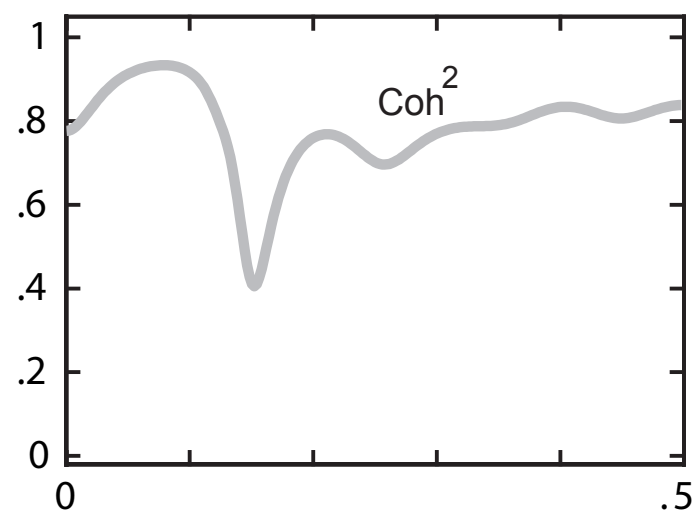
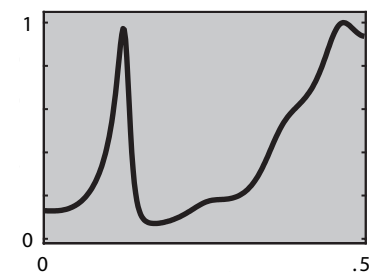
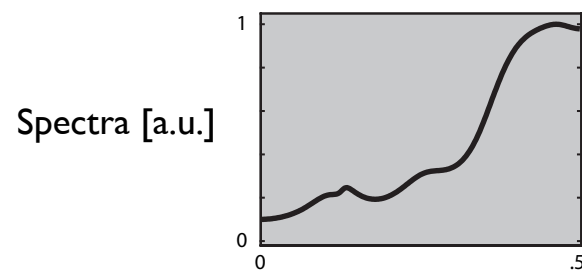
Amount of Correlation/Coherence

Toy Example I

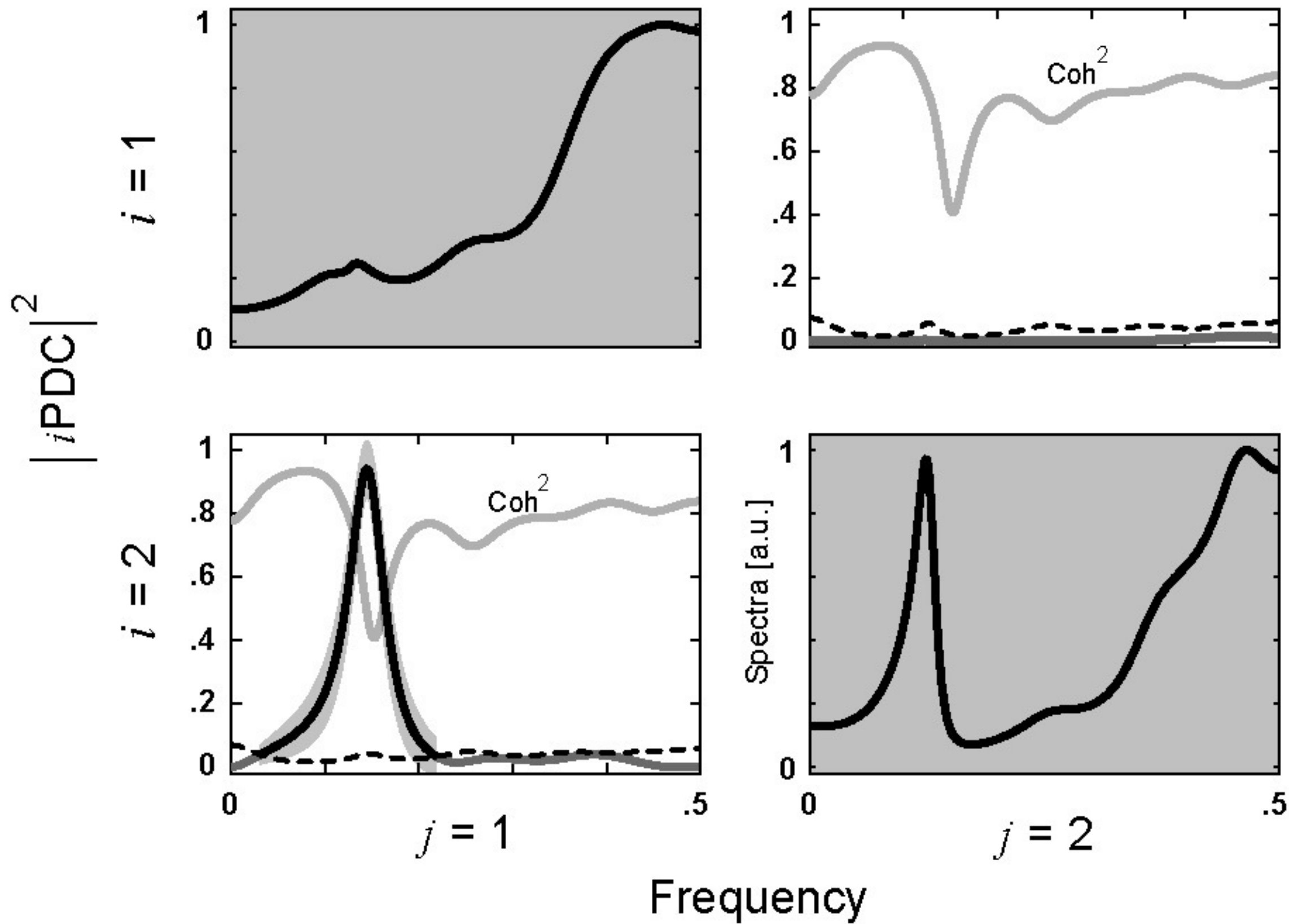
DTF=PDC=Granger Causality
2 time series

1

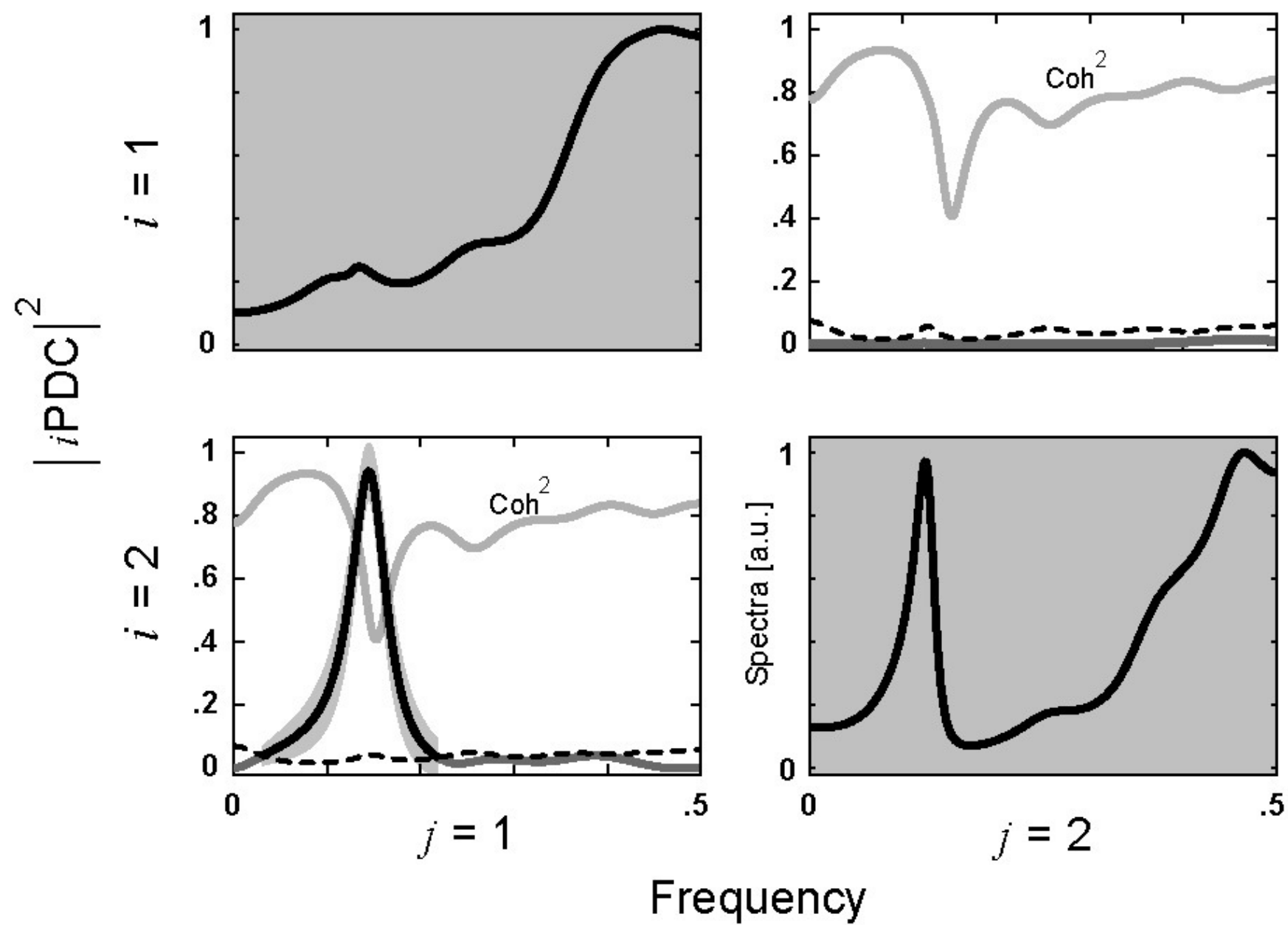
2



Normalized Freq. (x-axis)



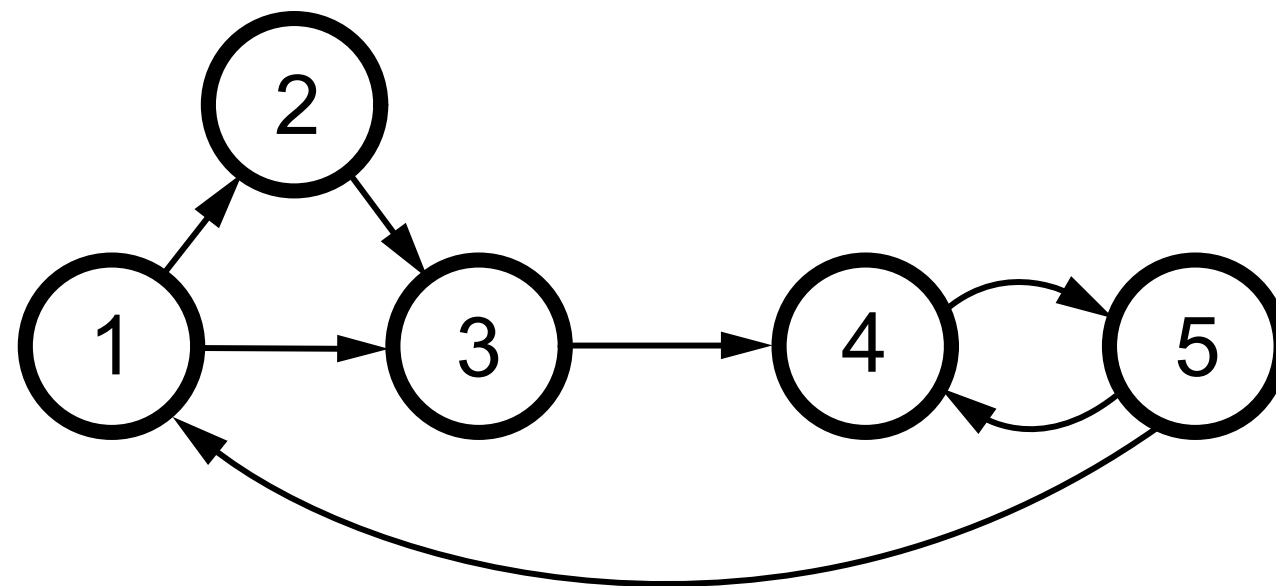
DTF=PDC
(Granger Causality)

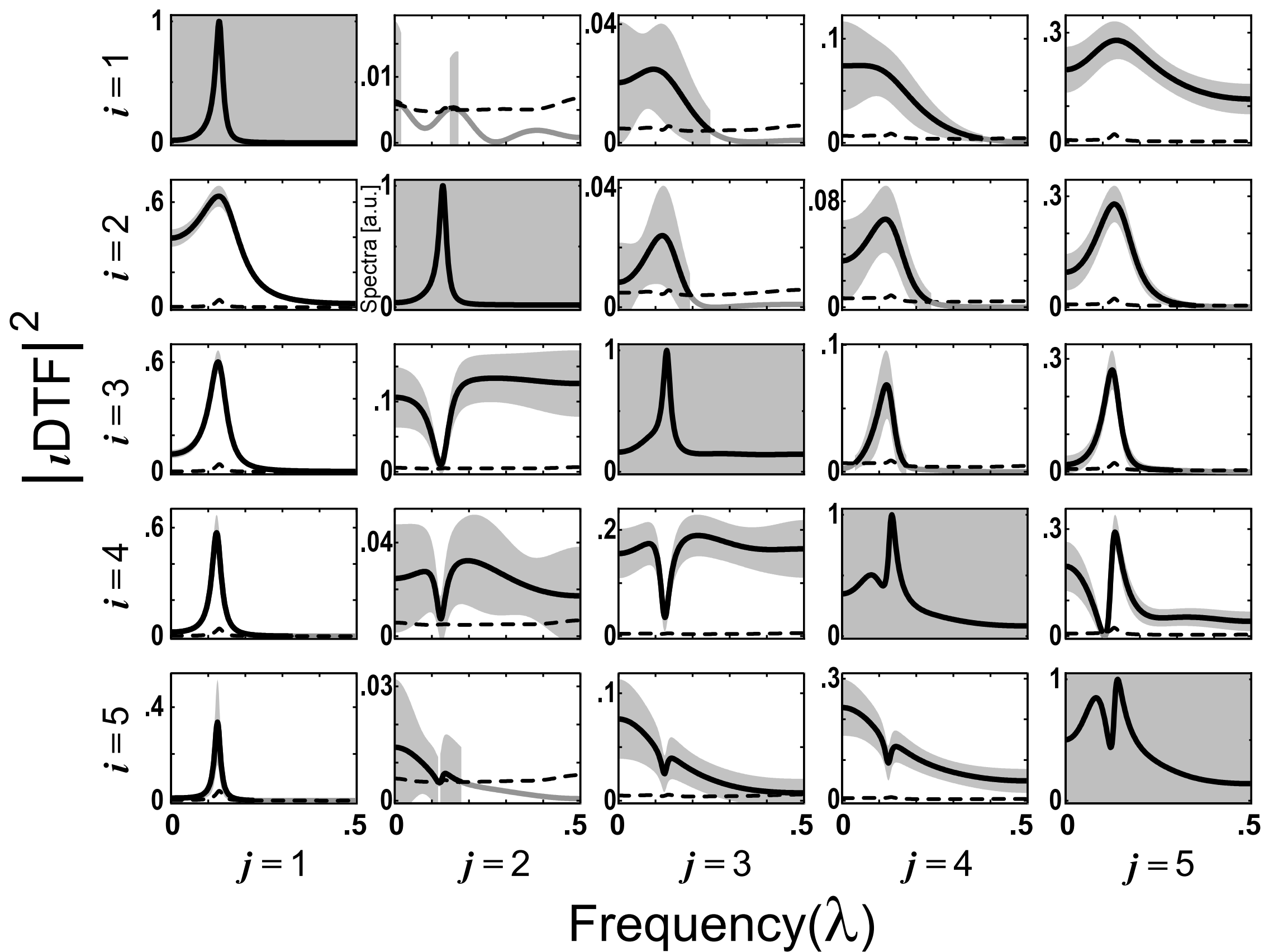


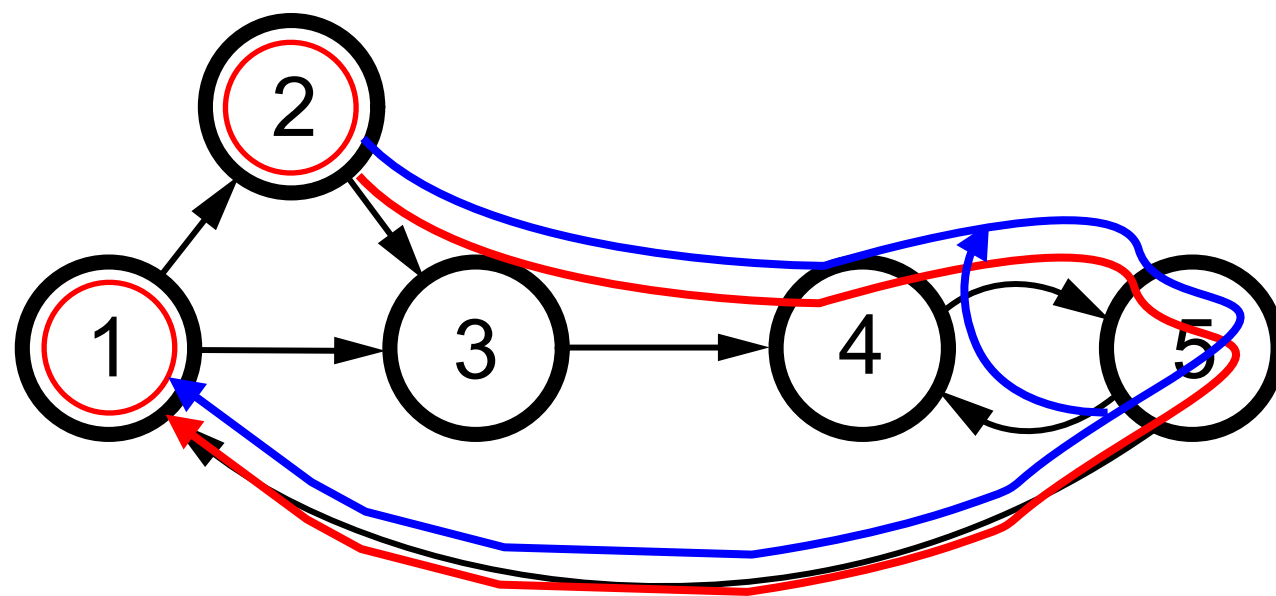
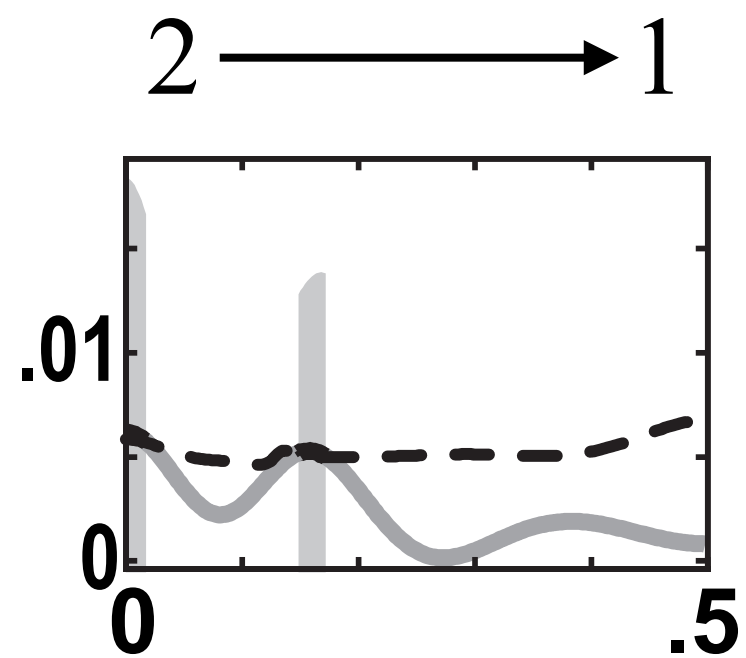
Toy Example II

DTF \neq PDC

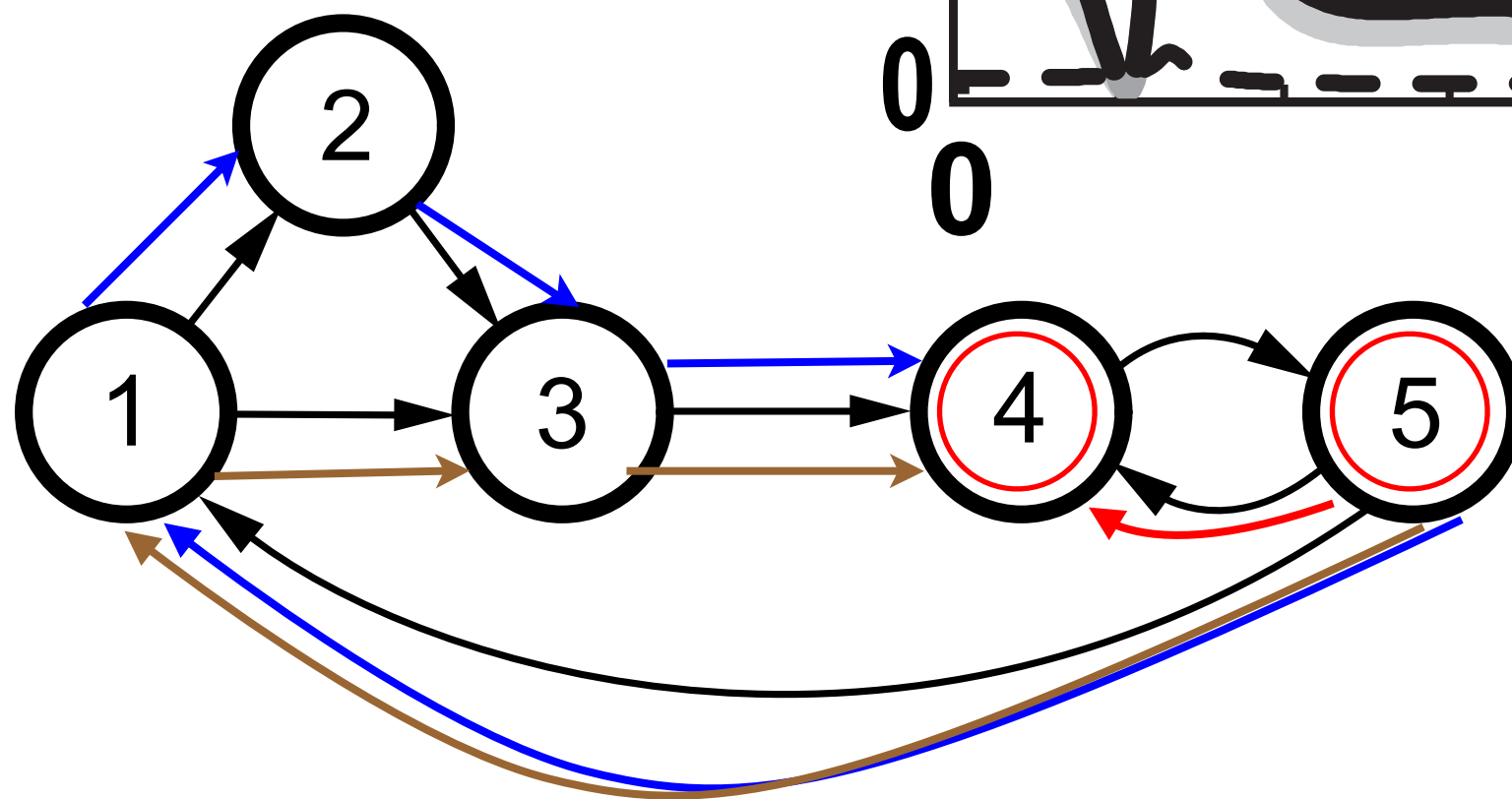
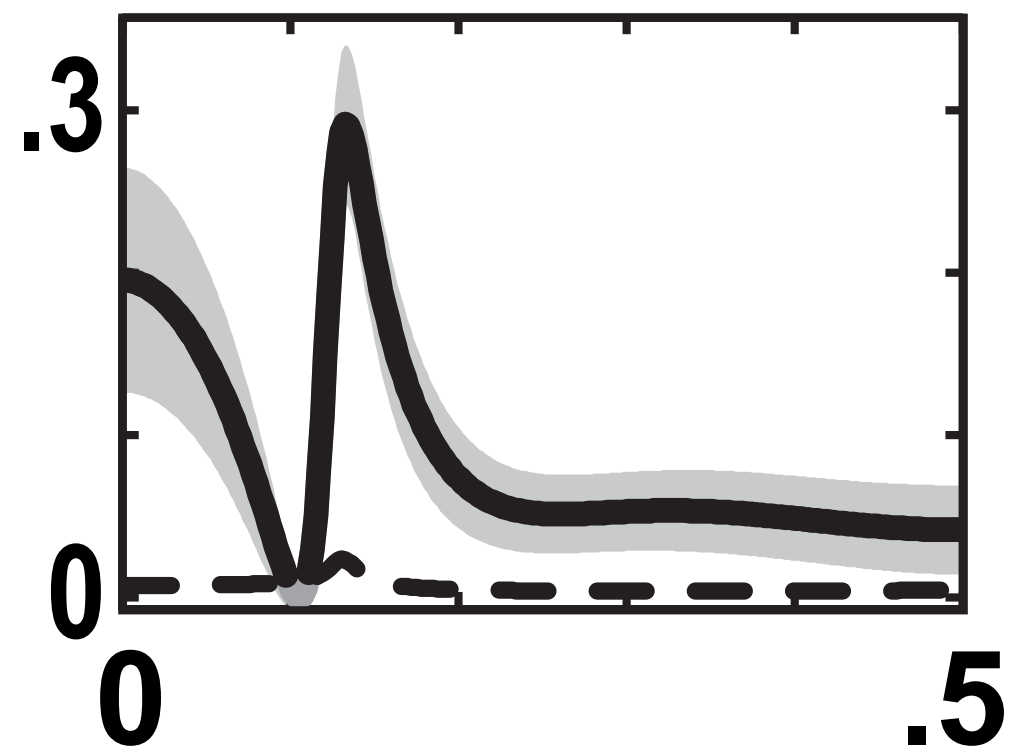
(More than 2 time series)







5 \longrightarrow 4



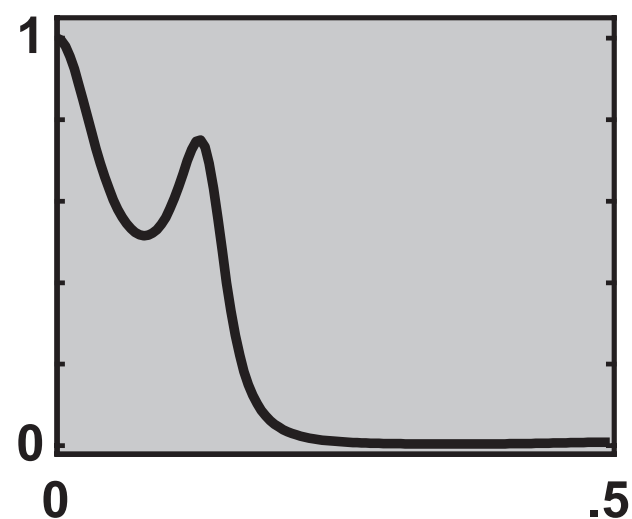
DTF

=

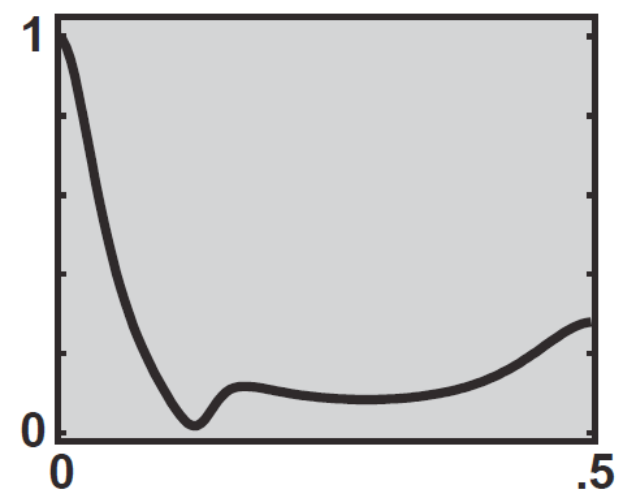
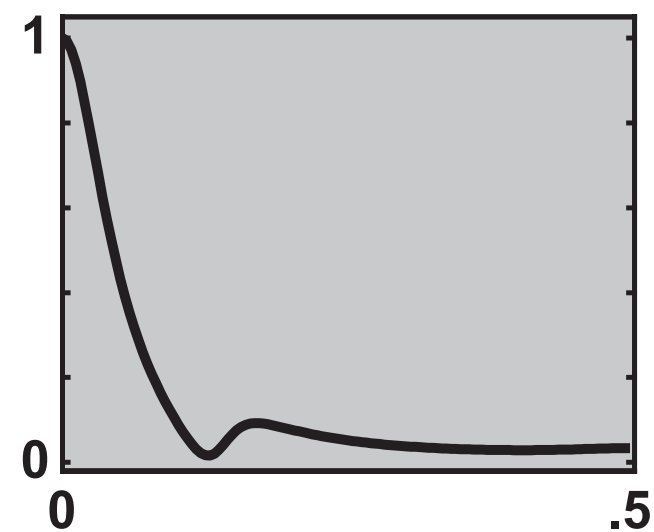
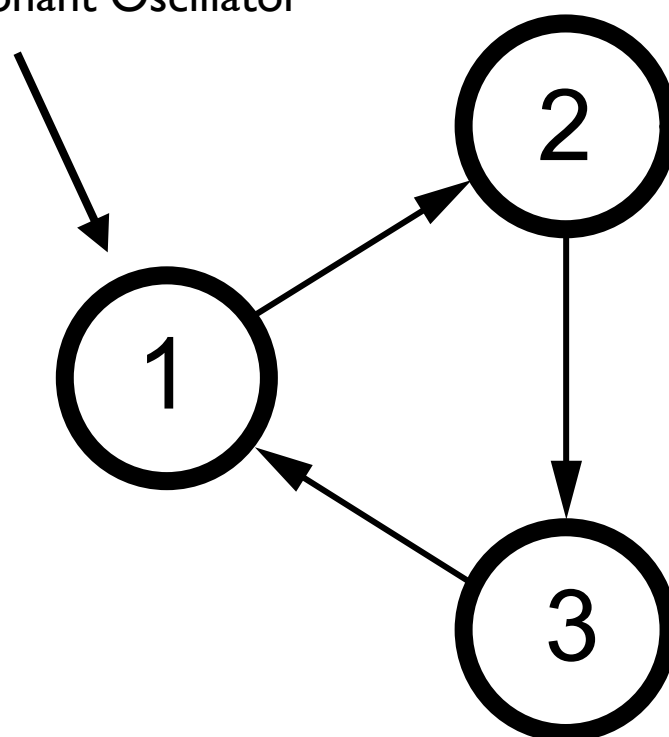
Sum-up of Influences Direct and Indirect

Toy Model III

(x 100)

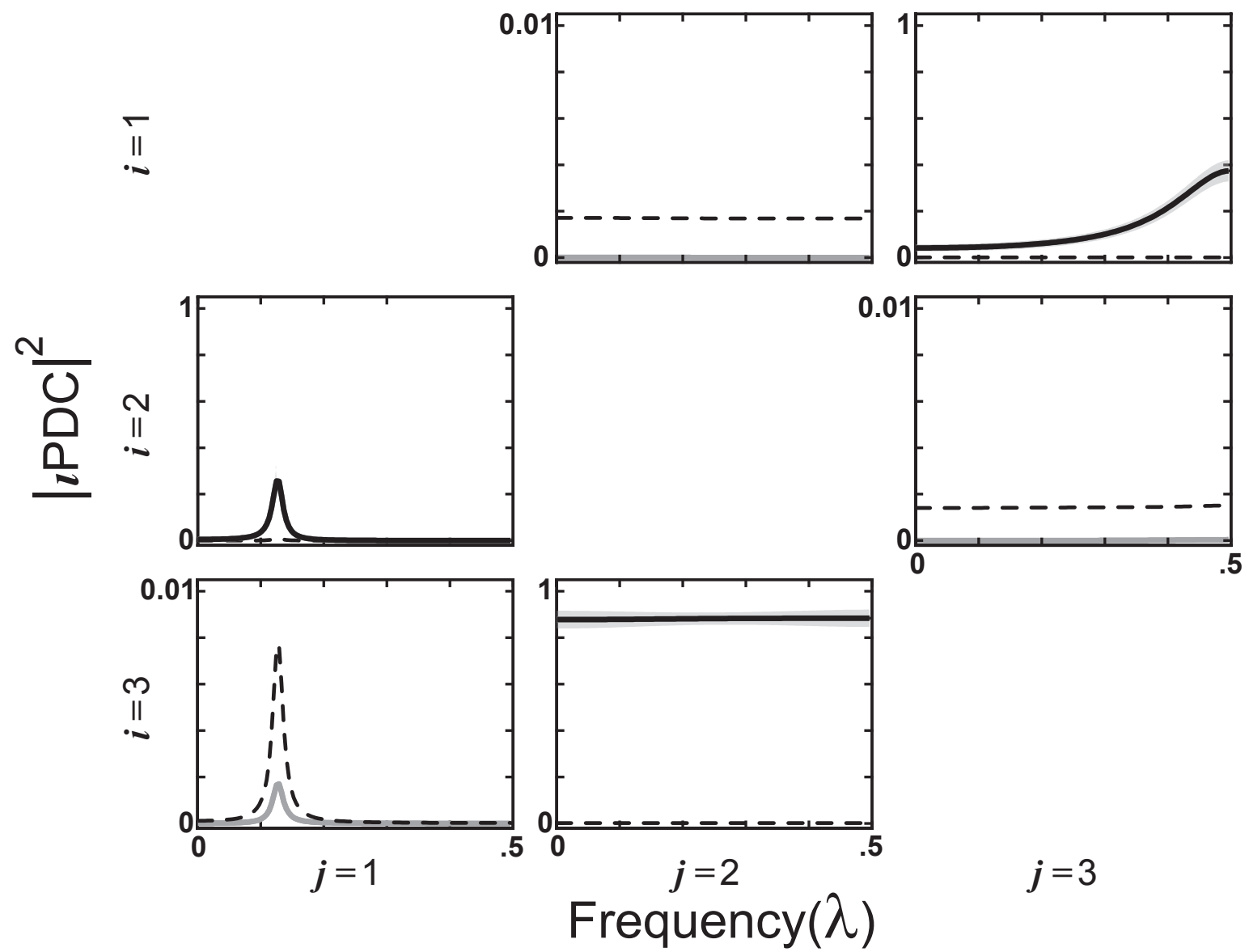
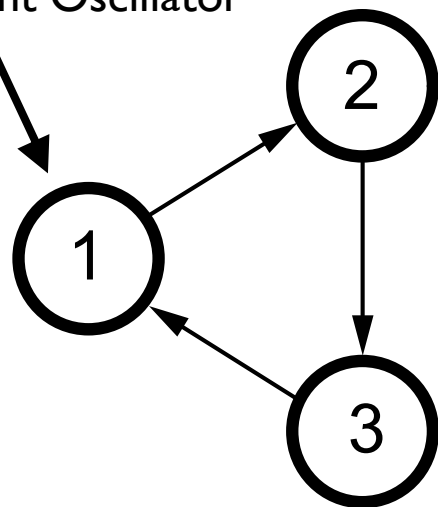


Resonant Oscillator

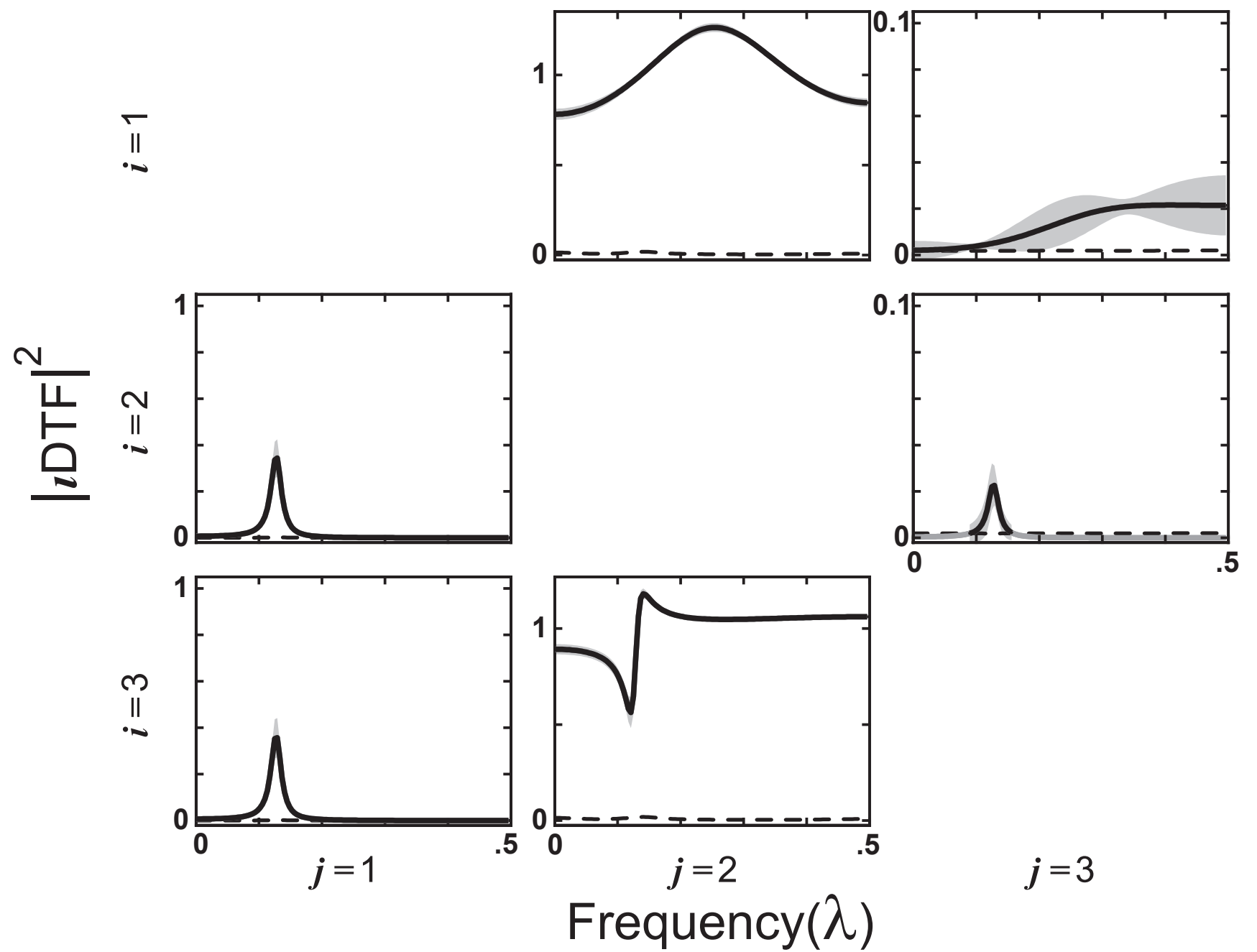
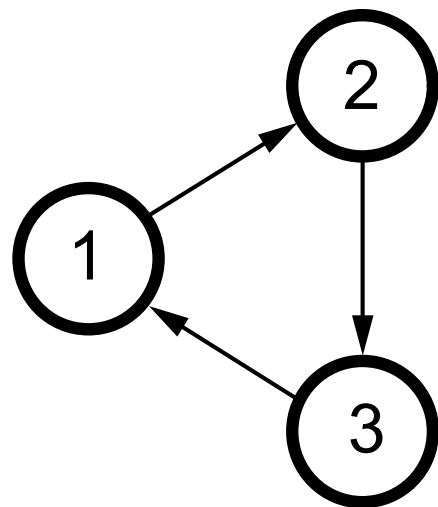


Toy Model III

Resonant Oscillator



Toy Model III



DTF and PDC

- Allow focusing on dynamic details of connections
- Work when coherence does not
- Give directional signal-flow information
- Are complementary for more than 2 time series
 - PDC represents immediate adjacent structure descriptions
 - DTF collects direct + indirect signal pathway flows

Effective Connectivity

the influence one neural system exerts over another

from Latin **effectivus**, from *efficere* 'accomplish'

Functional Connectivity

temporal correlations between remote neurophysiological events

(Friston, 1994)

from Latin *functionem* (nominative *functio*) "performance,

- relating to the way in which something works or operates: *there are important functional differences between left and right brain.*

Effective/Functional Classification

needs

UPGRADING!

Proposal

Link Centered Description

TABLE 13.1

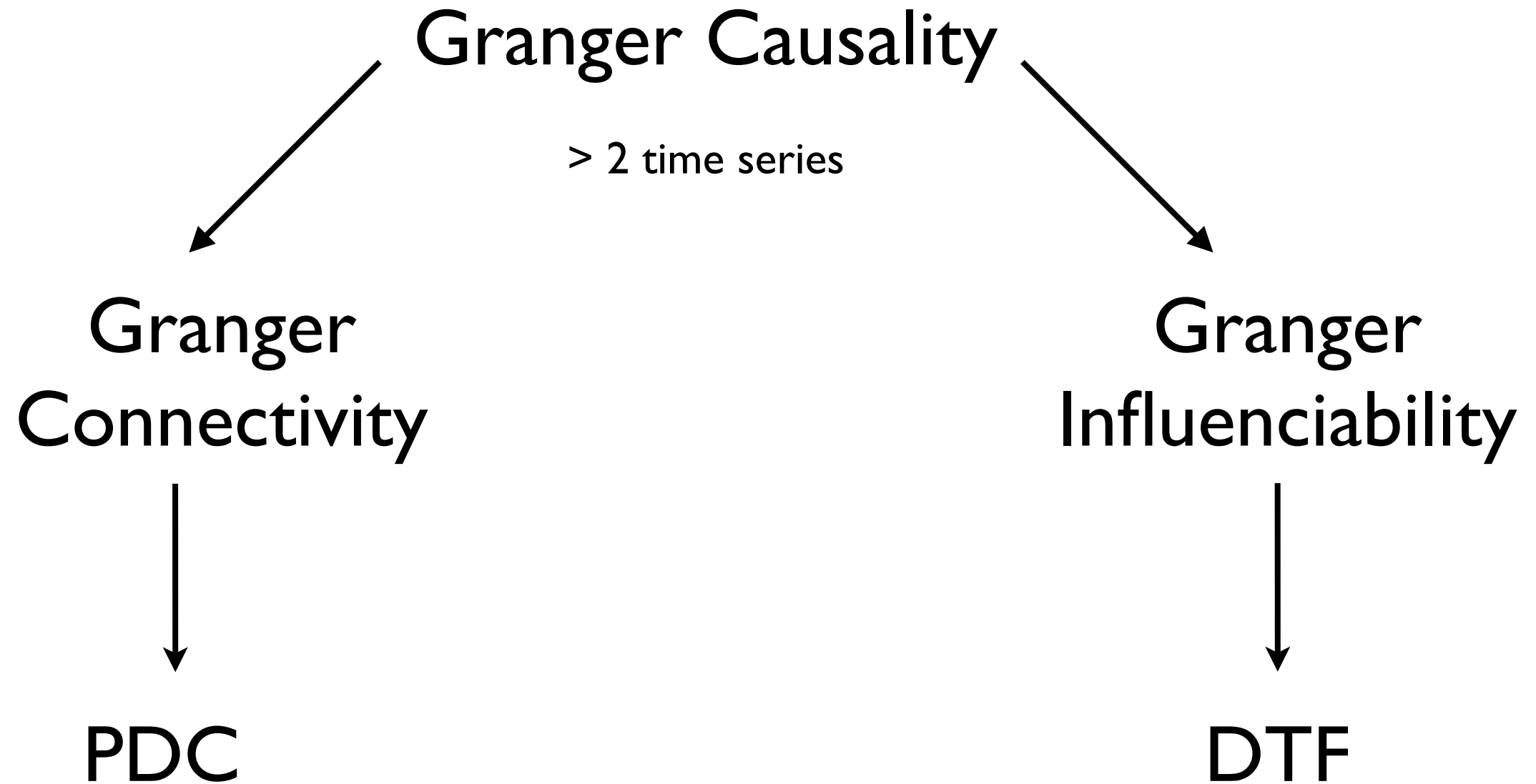
Key to the relationship between connectivity quantifiers and its classification

	<u>Direct</u>	<u>Indirect</u>
<u>Active</u>	$PDC \neq 0$	$PDC = 0$ and $DTF \neq 0$
<u>Inactive</u>	$PDC = 0$	$DTF = 0$

Baccalá LA, Sameshima K (2014b) Multivariate time series brain connectivity: a sum up. In: Sameshima K, Baccalá LA (eds) *Methods in brain connectivity inference through multivariate time series analysis*. Boca Raton: CRC Press, *in press*

DTF=PDC=Granger Causality

2 time series



Which one to use?

PDC if link description is desired

DTF to sum-up the net effect

The Connectivity Problem

Detection

Characterization

Host of Options

PDC/DTF

rPDC

GRANGER TIME DOMAIN MEASURES

Merely different units?

Downloaded from rsta.royalsocietypublishing.org on August 19, 2013

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

rsta.royalsocietypublishing.org

Research

Cite this article: Bacalá LA, de Brito CSN, Takahashi DY, Samehima K. 2013 Unified asymptotic theory for all partial directed coherence forms. *Phil Trans R Soc A* 371: 20120158.
<http://dx.doi.org/10.1098/rsta.2012.0158>

One contribution of 13 to a Theme Issue 'Assessing causality in brain dynamics and cardiovascular control'.

Subject Areas: biomedical engineering, statistics, applied mathematics

Keywords: partial directed coherence, Granger causality inference, asymptotic theory, connectivity detection

Author for correspondence: L. A. Bacalá
e-mail: bacala@ics.poli.usp.br

[†]Present address: Psychology Department, Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsta.2012.0158> or via <http://rsta.royalsocietypublishing.org>.

Unified asymptotic theory for all partial directed coherence forms

L. A. Bacalá¹, C. S. N. de Brito², D. Y. Takahashi^{2,3,4,†} and K. Samehima⁴

¹Telecommunications and Control Department, Escola Politécnica, ²Bioinformatics Graduate Program, ³Institute of Mathematics and Statistics, and ⁴Department of Radiology and Oncology, Faculdade de Medicina, University of São Paulo, São Paulo, Brazil

This paper presents a unified mathematical derivation of the asymptotic behaviour of the three main forms of partial directed coherence (PDC). Numerical examples are used to contrast PDC, gPDC (generalized PDC) and rPDC (information PDC) as to meaning, applicability and, more importantly, to show their essential statistical equivalence insofar as connectivity inference is concerned.

1. Introduction

Partial directed coherence (PDC), a multivariate time-series technique obtained from the factorization of partial coherence [1], has become popular for evaluating the connectivity between neural structures. In the frequency domain, it closely reflects the idea of Granger causality [2], i.e. a time series $x(k)$ is Granger-caused by $y(k)$ only if knowledge of $y(k)$'s past proves helpful in predicting $x(k)$, and it is, as such, an important measure given that many neuroscience research scenarios, such as sleep staging [3], have long been linked to typical neuroelectric oscillatory behaviour.

PDC has been finding increasing applications [4–6], most of which have been carried out by comparing sample connectivities between groups classified as presenting some known disorder against normal controls. Only recently have objective trial-by-trial criteria appeared that allow the evaluation of PDC's estimator asymptotic properties. This is the case with Takahashi *et al.* [7], who confirmed a previously available connectivity hypothesis test [8] by the addition of hitherto unavailable PDC confidence intervals.

Directed Transfer Function: Unified Asymptotic Theory and Some of its Implications

Luiz A. Bacalá, Member, IEEE, Daniel Y. Takahashi, Koichi Samehima, Member, IEEE

Abstract—Objective: To present a unified mathematical derivation of the frequency dependent asymptotic behaviour of the three main forms of directed transfer function (DTF). **Methods:** A synthesis of the results (proved in an extended Appendix) is followed by a series of Monte Carlo simulations of representative examples. **Results:** DTF estimators are asymptotically normal when the true value is different from zero. Under the null hypothesis of zero, the estimator is distributed as a linear combination of independent χ^2 variables. **Conclusions:** Null DTF rejection can be achieved with identical performance whether the DTF form is adopted. **Significance:** Together with recent unified partial directed coherence (PDC) results, this paper results up connectivity inference tools for a class of frequency domain connectivity estimators.

Index Terms—DTF asymptotics, partial directed coherence, DTF connectivity

I. INTRODUCTION

The recent realization of the potential role of connectivity analysis as a tool for understanding how the brain works has given rise to the appearance of a huge number of connectivity estimation methods that employ observed neurophysiological signals [1]–[6]. One common limitation of most methods is that their statistical behaviour is poorly understood and the existence of connectivity is often done without clear rigorous justification. One exception to this state of affairs is represented by [7] which discusses the rigorous characterization of *partial directed coherence* (PDC) [8] and its immediate variants [9], [10].

Here we carry out the statistical characterization along the lines presented in [7] for *directed transfer function* (DTF) [11], a frequency domain characterization of connectivity, that can be thought as a factorization of the coherence between pairs of observed time series [8].

PDC and DTF constitute dual connectivity description concepts. PDC captures *active* immediate directional coupling between structures whereas DTF portrays the existence of directional signal propagation even if it is only indirect, i.e. when signals may travel through intermediate structures rather than through an immediate direct causal influence path [12]. Thus DTF reflects signal '*reachability*' in a graph theoretical

L. A. Bacalá is with Telecommunications and Control Department, Escola Politécnica, University of São Paulo, Av. Prof. Luciano Gualberto, trav. 3,8 158, São Paulo, SP, 05508-900, Brazil. Tel.: +55-11-3091-5508, Fax: +55-11-3091-5119, email: bacala@ics.poli.usp.br

D. Y. Takahashi is with Psychology Department, Princeton University, Princeton, NJ 08540, USA, email: takahashi@princeton.edu

K. Samehima is with Department of Radiology and Oncology, Faculdade de Medicina, University of São Paulo, São Paulo, Brazil, email: ksamehima@uosp.br

[†]Copyright (c) 2016 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

Much like PDC, DTF has given rise to a number of variants, that were deemed appropriate to different situations, whose discussion from a historical perspective is provided by their authors in [14]. Since DTF's introduction, we examined two of its variants: (a) directed coherence (DC) [15] which is DTF's scale invariant form (and dual to generalized PDC (gPDC) [9]) and (b) information DTF (iDTF) which is an information theoretic generalization of DTF we introduced, dual to information PDC (iPDC), both of which provide accurate effect size information of the connectivity strength, i.e. of the amount of information flow [10], [16], [17].

Here we derive and illustrate inference results for the latter DTF forms from a unified perspective closely paralleling the inference results in [7] which were further illustrated in [18]. The importance of accurate asymptotics for DTF is that jointly DTF and PDC allow extending the current effective/functional classification connectivity concepts to a more accurate general and informative context [19], [20].

After briefly reviewing DTF's formulations (Sec. II) together with a summary of the unified asymptotic results (Sec. III), numerical illustrations (Sec. IV) discuss some implications of the results that are further elaborated in Sec. V. For clarity, mathematical details are left to the Appendix whose implementation is to appear in the next release of the AsympPDC package [7] which may be downloaded from <http://www.ics.poli.usp.br/~bacala/pdc/DTF.html> together with data and scripts used in the examples contained here.

II. BACKGROUND

The departure point for defining the present DTF variants is an adequately fitted multivariate autoregressive time series (i.e. vector time series $\mathbf{x}(n)$ made up by $x_k(n)$, $k = 1, \dots, K$) model:

$$\mathbf{x}(n) = \sum_{l=1}^p \mathbf{A}(l)\mathbf{x}(n-l) + \mathbf{w}(n), \quad (1)$$

where $\mathbf{w}(n)$ stands for a zero mean white innovations process of with $\Sigma_w = [\sigma_{ij}]$ as its covariance matrix and p is the model order. The $a_{ij}(l)$ coefficients composing each $\mathbf{A}(l)$ matrix describe the lagged effect of the j -th on the i -th series, wherefrom one can also define a frequency domain representation of (1) via the $\hat{\mathbf{A}}(\lambda)$ matrix whose entries are given by

$$\hat{A}_{ij}(\lambda) = \begin{cases} 1 - \sum_{l=1}^p a_{ij}(l)e^{-j\lambda l}, & \text{if } i = j \\ -\sum_{l=1}^p a_{ij}(l)e^{-j\lambda l}, & \text{otherwise,} \end{cases} \quad (2)$$

Connection Detection

H_0 : No PDC (DTF) exists at f can be rejected at α

Connection Characterization (if H_0 is rejected)

PDC (DTF) \sim Normally distributed

Information Flow Interpretation

DTF/MA

PDC/AR

Information

$${}_i\gamma_{ij}(\omega) = \frac{\sigma_{jj}H_{ij}(\omega)}{\sqrt{h_i^H(\omega)\Sigma_w h_i(\omega)}} \quad {}_i\pi_{ij}(\omega) = \frac{\sigma_{ii}^{-1}\bar{A}_{ij}(\omega)}{\sqrt{a_j^H(\omega)\Sigma_w^{-1} a_j(\omega)}}$$

(Takahashi et al, 2010)

Diagnostics and Interpretation

- Reliance on model fitting

Model quality appraisal

- “truthfulness” needs including “all” variables of interest

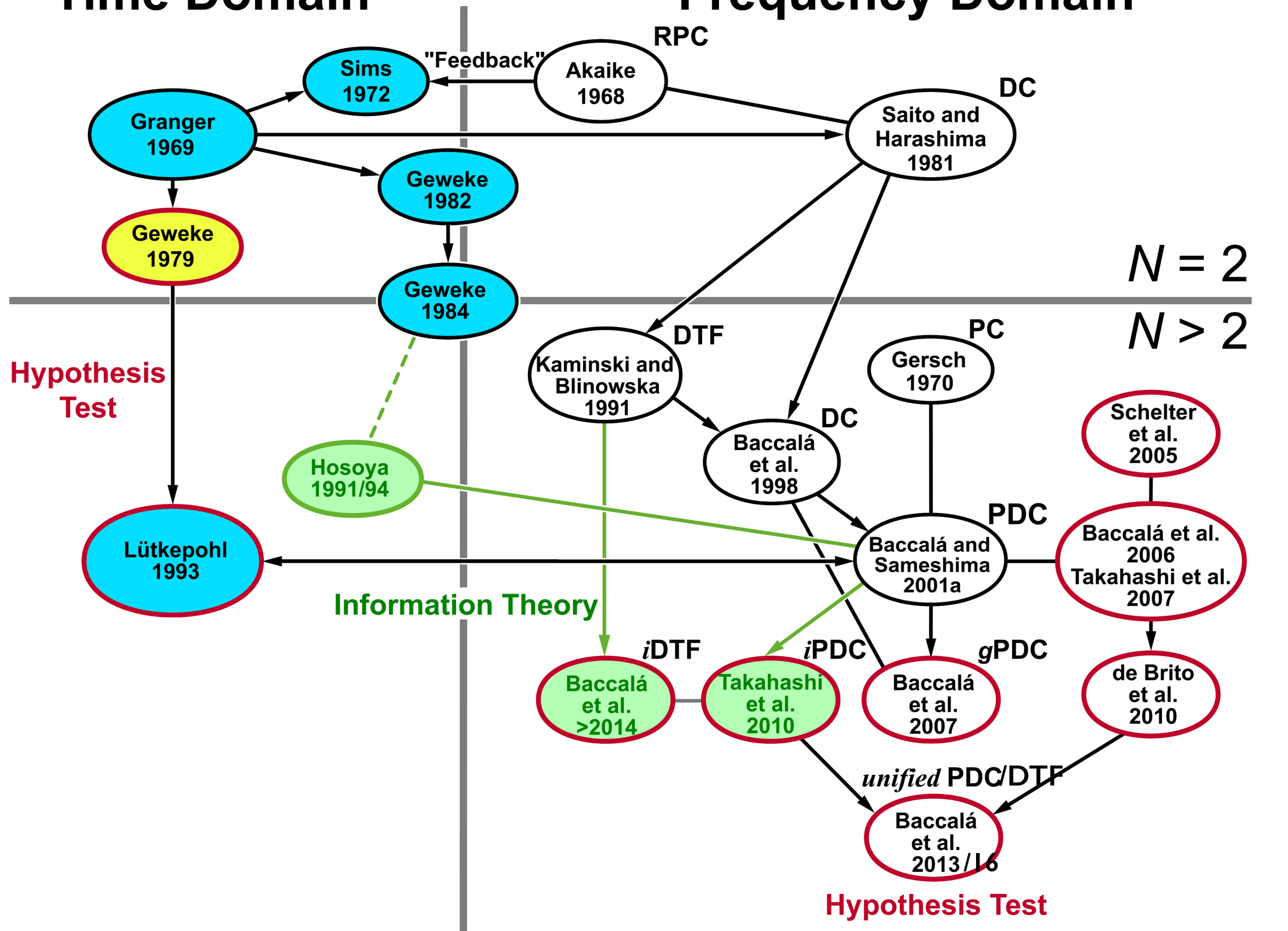
Implications on model size and signal length

- Limited to ‘linear models’

(as is correlation)

Time Domain

Frequency Domain



Some Developments at EMBC 2016

- Time Variant Linear Models for Short Segment ERPs
- Nonlinear Connectivity (Symposium)

Program Overview

Morning

- ✦ Introduction and Overview - L.A. Baccalá
- ✦ Applications of Granger Causality to Neuroscience - Mingzhou Ding
- ✦ Statistical and Software Applications - K. Sameshima
- ✦ Power User Applications - L. Astolfi

Afternoon

Data Analysis Challenges

Challenge Resolution and Discussion

Acknowledgments

My Co-authors

Daniel Yasumasa Takahashi

Koichi Sameshima

Web Page:

<http://www.lcs.poli.usp.br/~baccala/pdc>

Sponsors:



Thank You!

<http://www.lcs.poli.usp.br/~baccala/pdc>