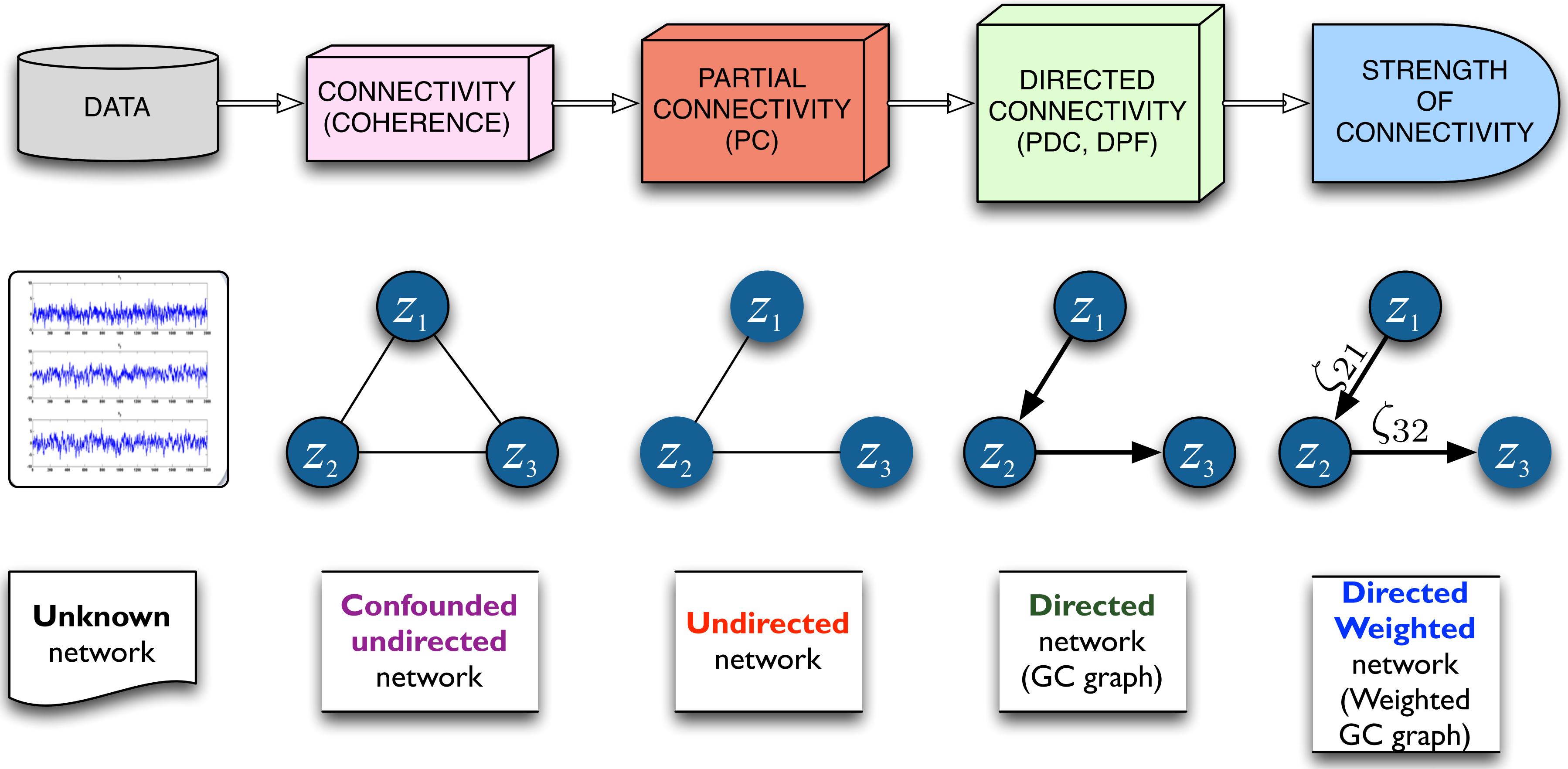
Reconstructing Causal Networks From Data - Overview and Developments (Part II)

Arun K. Tangirala

Process Systems Engineering & Data Sciences Group
Department of Chemical Engineering, IIT-M
Robert Bosch Centre for Data Sciences and Artificial Intelligence, IIT-M
Chennai, TN India

Systematic Network Reconstruction



- Stochastic DGP
- Jointly stationary
- Spectral densities exist

- Correlations
- Cross-spectral densities (CSD)
- Coherence functions

- Partial correlations
- Partial CSD
- Partial coherence (conditional)

- Partial directed correlation / coherence
- Direct pathway function
- …

- Direct pathway function
- Partial directed correlation
- …

Moving to Frequency Domain

Eigenvalues of the (infinite-dim) variance-covariance matrix constitute the spectral density function of a stochastic process over $\boldsymbol{\omega} \in [-\pi, \pi)$

Cross-spectral density

$$S_{ij}(\omega) = \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \sigma_{ij}[l] e^{-j\omega l}$$

(Squared) Coherence

$$C_{ij}(\omega) = |\kappa_{ij}(\omega)|^2 = \frac{|S_{ij}(\omega)|^2}{S_{ii}(\omega) S_{jj}(\omega)}$$

Partial cross-spectral density

$$S_{ij.\mathcal{W}}(\omega) = S_{ij}(\omega) - S_{i\mathcal{W}}(\omega) S^{-1}_{\mathcal{W}\mathcal{W}}(\omega) S_{\mathcal{W}j}(\omega)$$

(Squared) Partial Coherence

$$PC_{ij}(\omega) = \frac{|S_{ij.\mathcal{W}}(\omega)|^2}{S_{ii.\mathcal{W}}(\omega) S_{jj.\mathcal{W}}(\omega)}$$

Advantages of frequency-domain analysis:

- *Filtering characteristics* of the connectivity (system) can be computed.
- *Good separation of signal and noise* components in the spectral domain
- Can easily ***detect lagged relations*** (without knowing the lag)
- Fairly robust methods exist for estimating spectral densities
- …

WIENER-GRANGER CAUSALITY

WGC is a "prediction-based" causality measure.

Given a multivariate time-series $\{z_i[k]\}$, $i = 1, 2, \ldots, M$ and $k = 0, 1, \ldots, N-1$

1. Construct two predictions for $z_j[k]$ - one based on exclusion of past of $z_i[k]$ and the other including it.
2. If the prediction after inclusion is improved, then $z_i[k]$ Granger-causes $z_j[k]$.

Comparison of predictions is usually based on prediction-error variance

- A signal $z_i[k]$ Granger causing another signal $z_j[k]$ **does not necessarily imply physical causality.**
- The causal relations detected, of course, is subject to change if the variable set is changed (exclusion or inclusion of variables).

TESTS FOR GC

There exists several tests for GC:

- Variance-based tests
  - *Time-domain ratios-of-variance and spectral WGC tests by Geweke (1980s)*
  - *Granger-Sargent and Granger-Wald tests*
  - *Granger causality index (GCI)*
- Model-based tests (significance of parameter estimates of vector TS models)
- Partial directed correlation (with significance tests)
- Frequency-domain methods (with appropriate hypothesis tests)
  - *Partial directed coherence*
  - *Direct pathway functions*

PARAMETRIC IMPLEMENTATION: VAR MODELS

- Vector auto-regressive (VAR) models are often used for forecasting
  - *Easy to estimate and interpret*
- **VAR(*p*)** Model:

$$\mathbf{z}[k] = \sum_{r=1}^{p} \mathbf{A}_r \mathbf{z}[k-r] + \mathbf{e}[k]$$

$$\mathbf{A}(q^{-1})\mathbf{z}[k] = \mathbf{e}[k], \quad \mathbf{A}(q^{-1}) = \mathbf{I} - \sum_{r=1}^{p} \mathbf{A}_r q^{-r}$$

where $\mathbf{e}[k]$ is a (possibly spatially correlated) vector WN process.

- Assumptions:
  - *Joint stationarity, spectral density matrix is factorizable.*

z_i does not Granger cause z_j if and only if $a_{ji,r} = 0, \forall r = 1, \cdots, P$

Frequency-Domain: Partial directed coherence (PDC)

PDC (Definition)

$$\pi_{ij}(\omega) = \frac{\bar{a}_{ij}(\omega)}{\sqrt{\bar{\mathbf{a}}^*_{.j}(\omega)\bar{\mathbf{a}}_{.j}(\omega)}} = \frac{\bar{a}_{ij}(\omega)}{\sqrt{\sum_{i=1}^{m} |\bar{a}_{ij}(\omega)|^2}}$$

$$\bar{\mathbf{A}}(\omega) = \mathbf{I} - \mathbf{A}(\omega) = \mathbf{I} - \sum_{r=1}^{p} \mathbf{A}_r e^{-j\omega r}$$

z_i does not Granger cause z_j if and only if $|\pi_{ji}(\omega)|^2 = 0, \ \forall \omega$

- Bounded and normalized measure

$$0 \leq |\pi_{ij}(\omega)|^2 \leq 1 \text{ and } \sum_{i=1}^{m} |\pi_{ij}(\omega)|^2 = 1$$

- Factorizes partial coherency:

$$PC(\omega) = \Pi^*(\omega)\Sigma_e^{-1}\Pi(\omega)$$

VMA Representations (Total Causality)

- Dual of VAR models (represent variables as linear mix of innovations)

$$\mathbf{z}[k] = \sum_{l=1}^{Q} \mathbf{H}_l \mathbf{e}[k-l] + \mathbf{e}[k]$$

$\mathbf{H}_r$: MA coefficient matrix at lag l

  - *The element $h_{ji}[l]$ contains the **total** causal influence from e_i to z_j at lag l.*

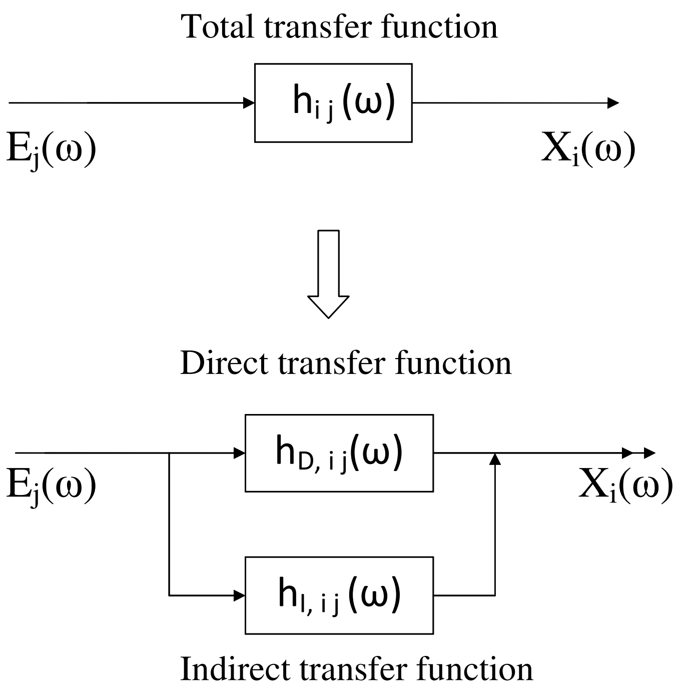
- Normalized frequency-domain **H**(ω) is the Directed Transfer Function (DTF)
  - *Measures Sims or **total** causality*

In frequency domain, the relationship between VAR and VMA models is

$$\mathbf{H}(\omega) = \bar{\mathbf{A}}^{-1}(\omega), \quad \text{where } \bar{\mathbf{A}}(\omega) = I - \mathbf{A}(\omega)$$

DECOMPOSING TOTAL CAUSALITY

- The total transfer function $h_{ij}(\omega)$ contains contributions from both **direct and indirect pathways.**



$$h_{ij}(\omega) = h_{D,ij}(\omega) + h_{I,ij}(\omega)$$

$$h_{D,ij}(\omega) = \frac{-\bar{a}_{ij}(\omega)\det(\bar{M}_{ji}(\omega))}{\det(\bar{A}(\omega))}$$

$$h_{I,ij}(\omega) = h_{ij}(\omega) - h_{D,ij}(\omega)$$

$$|h_{ij}(\omega)|^2 = |h_{D,ij}(\omega)|^2 + |h_{I,ij}(\omega)|^2 + 2|h_{D,ij}(\omega)||h_{I,ij}(\omega)|\cos\phi_{ij}(\omega)$$

The direct transfer function $h_{D,ij}(.)$ can serve as a measure of Granger causality in frequency-domain

DIRECT PATHWAY FUNCTION

- The (direct) causal relation between two random signals is detected by

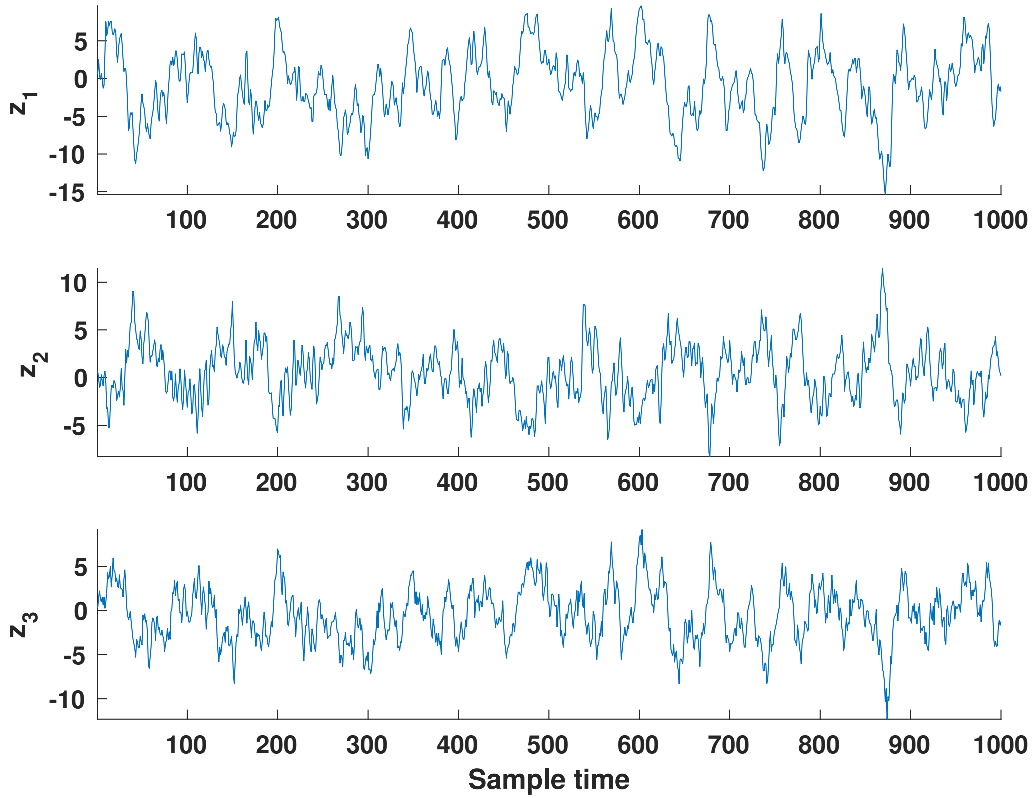
$$\psi_{ij}(\omega) \triangleq \frac{h_{D,ij}(\omega)}{\sqrt{\sum_{i=1}^{M} |h_{D,ij}(\omega)|^2}}$$

termed as the direct pathway function (DPF).

- The DPF is a measure of Granger causality between $z_j[k]$ and $z_i[k]$
  - *It characterises the direct pathway from the endogenous source $e_j[k]$ to the variable (effect) $z_i[k]$*

Causality structure and **strengths of connectivities** can both be obtained from the DPF $\psi_{Dij}(\omega)$

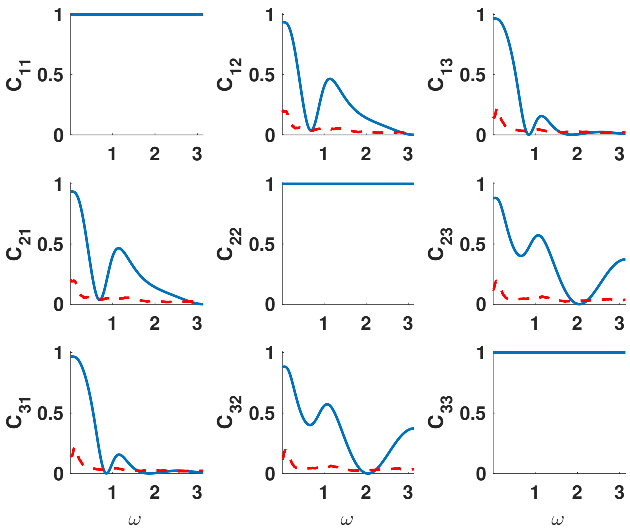
Simulation Case Study

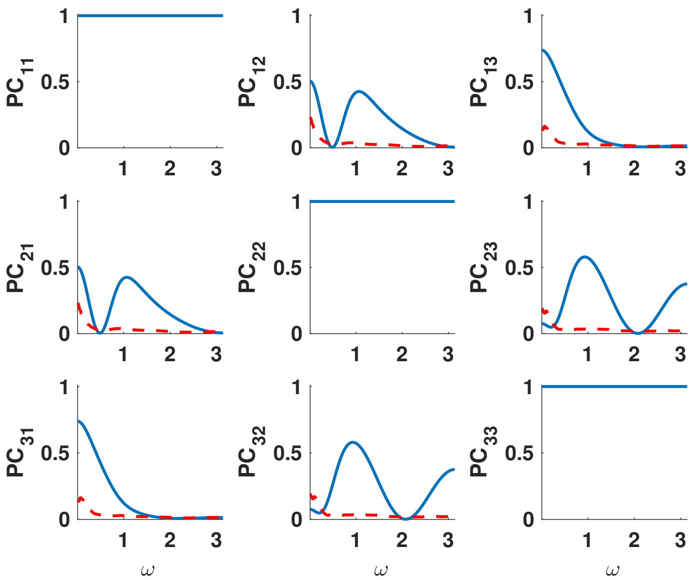


- N = 1000 observations of a 3-D stationary process
- **Goal:** Reconstruct a class of networks from correlation to the weighted causal network

Coherence and Conditioned Networks

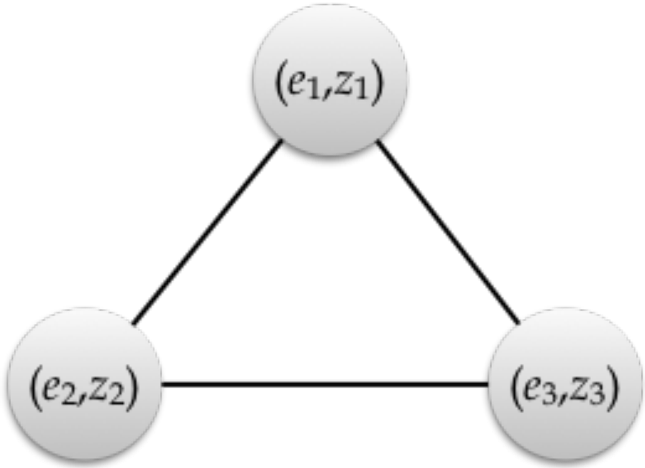
Coherence

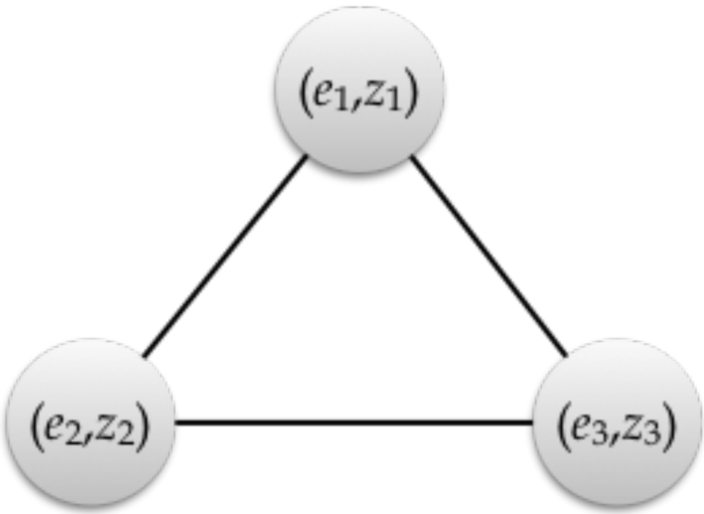




Partial Coherence

Undirected, confounded network

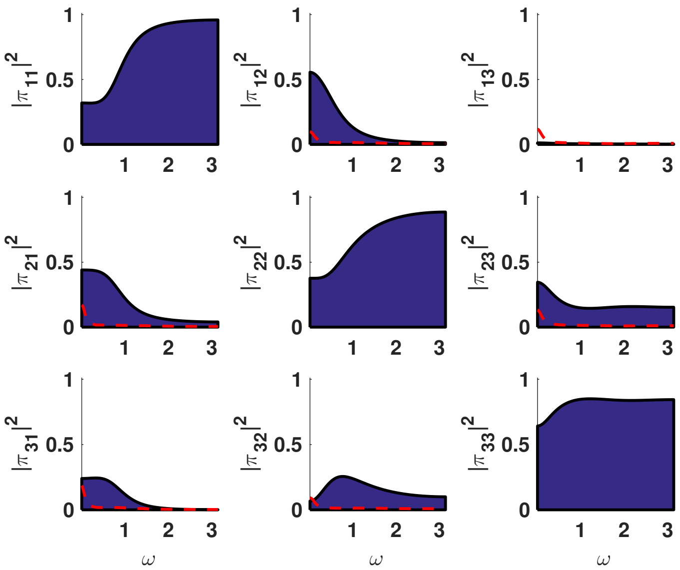


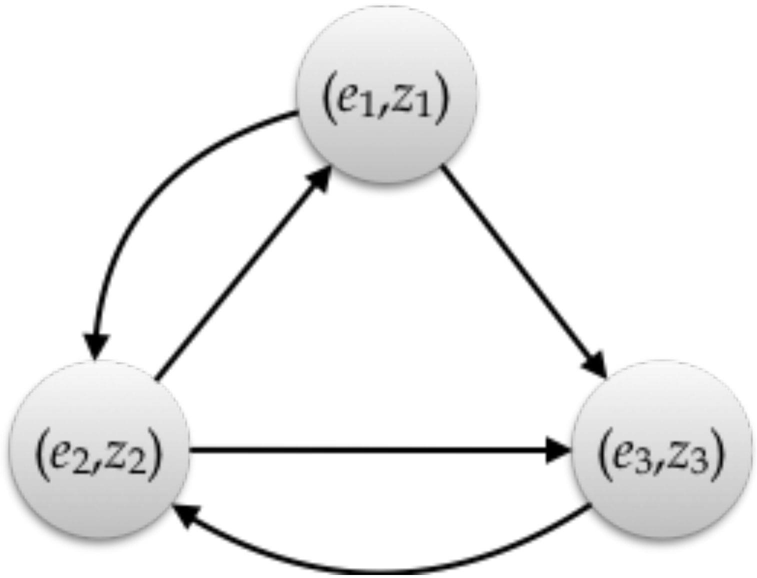


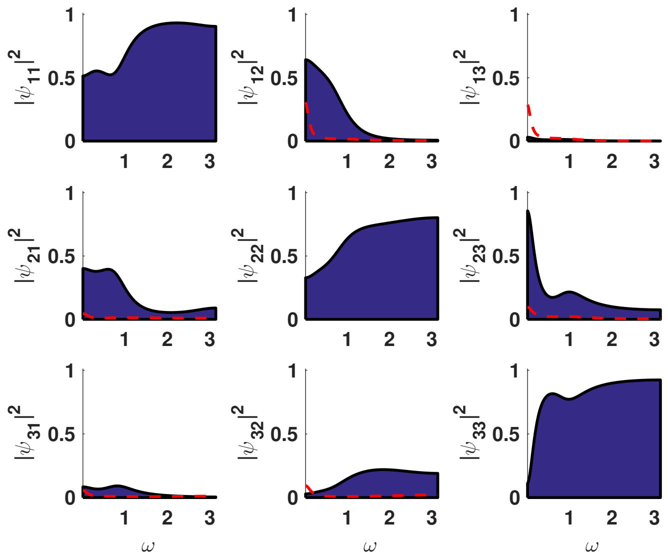
Undirected, direct network

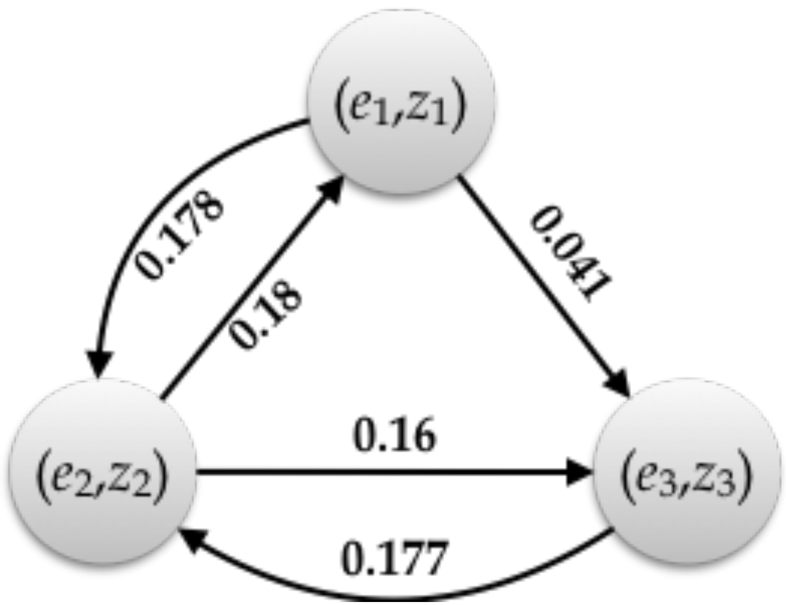
Directed (Causal) Weighted Networks

Partial directed coherence





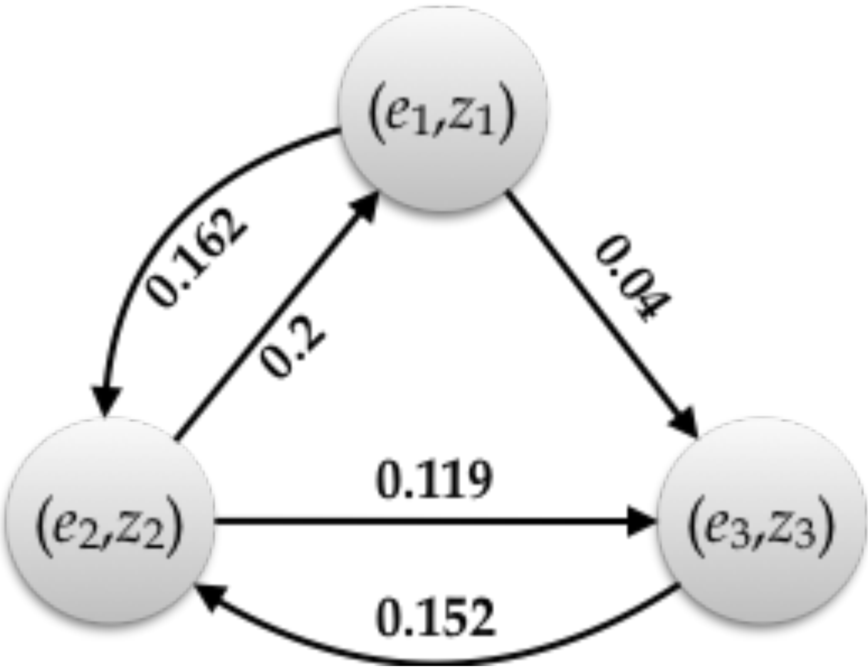




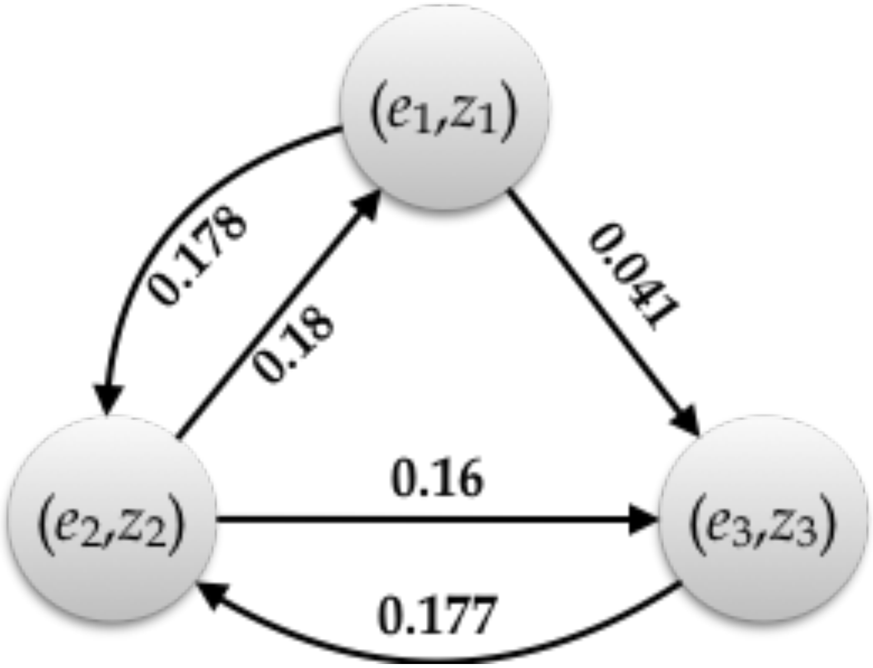
Direct Pathway function

True Process

$$z_1[k] = z_1[k-1] - 0.3z_1[k-2] + 0.4z_2[k-2] + e_1[k] + 0.3e_1[k-1]$$
$$z_2[k] = -0.4z_1[k-1] + 0.6z_2[k-1] + 0.3z_3[k-1] + e_2[k] + 0.4e_2[k-1] + 0.2e_2[k-2]$$
$$z_3[k] = 0.2z_1[k-1] - 0.4z_2[k-2] + 0.3z_2[k-3] + 0.6z_3[k-1] + e_3[k]$$



True causal network



Identified network

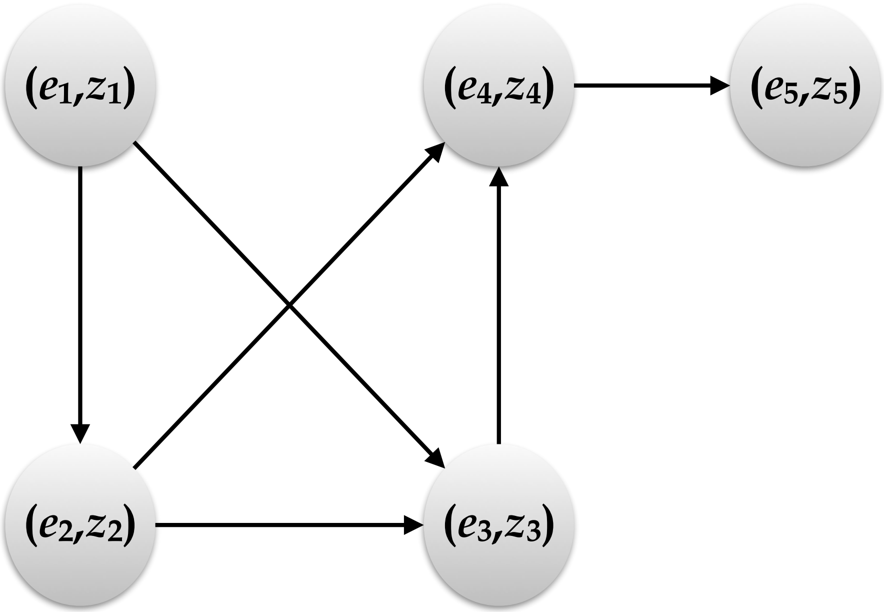
Issues and Challenges

- **Confounding**
  - *Spurious Causality*
- Instantaneous causality
- Model-process mismatch
- **Network (graphical) representations**
- Small sample scenario
- **Missing observations**
- Effects of measurement errors
- **Non-linearities**
- …

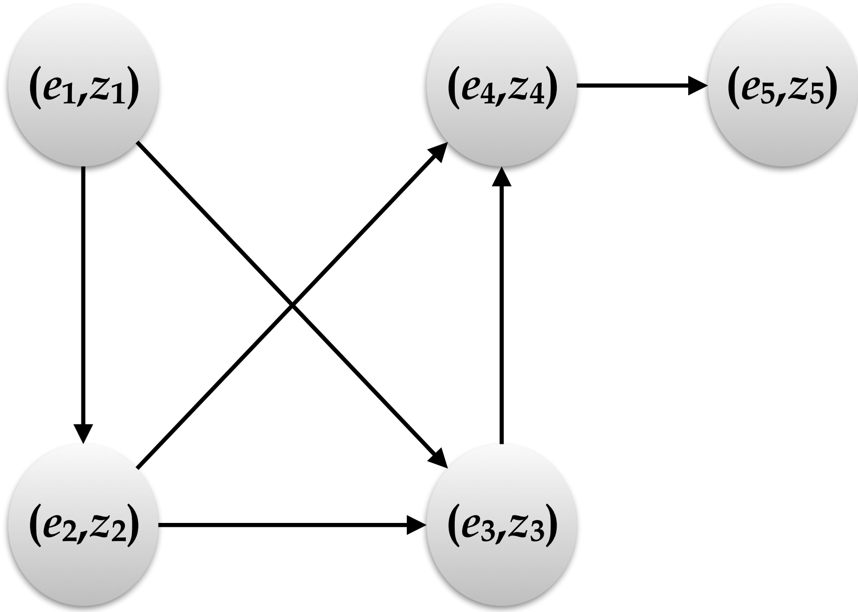
CONFOUNDING: EFFECT OF LATENT VARIABLES

- Consider a 5-D process:

$$
\begin{aligned}
z_1[k] &= 0.7z_1[k-1] + e_1[k] \\
z_2[k] &= 0.2z_1[k-2] + 0.9z_2[k-1] + e_2[k] \\
z_3[k] &= 0.2z_1[k-1] + 0.5z_2[k-1] + 0.8z_3[k-2] + e_3[k] \\
z_4[k] &= 0.5z_2[k-1] + 0.4z_3[k-1] + 0.6z_4[k-1] + e_4[k] \\
z_5[k] &= 0.7z_4[k-1] + 0.2z_5[k-1] + e_5[k]
\end{aligned}
$$

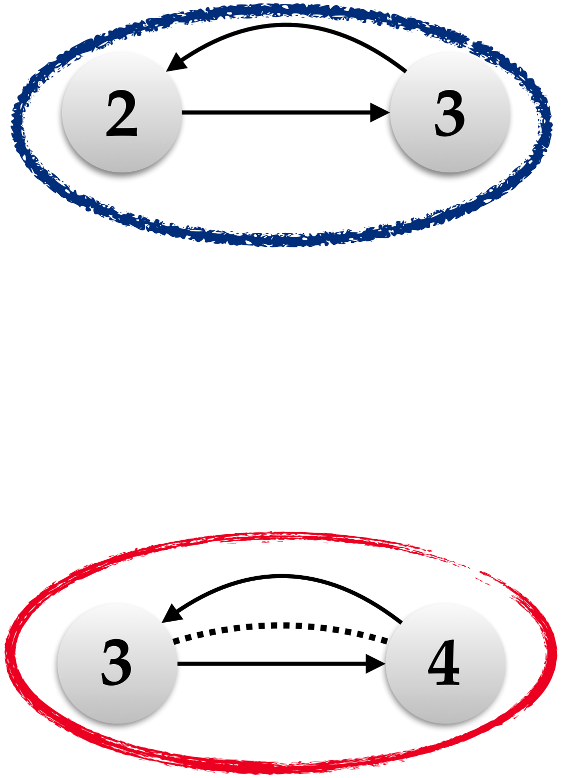


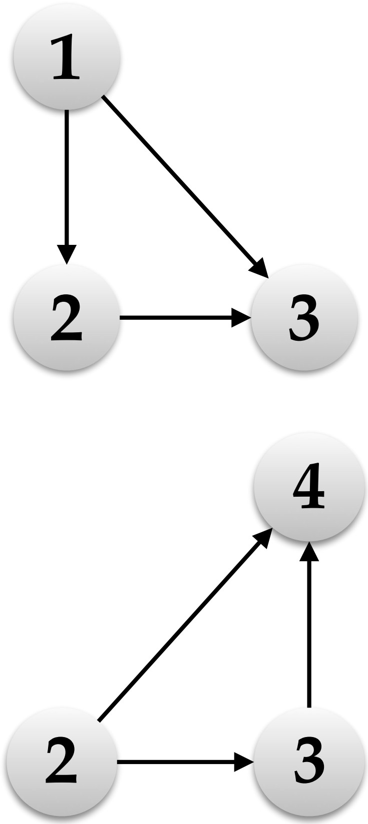
True GC graph



Reconstructed GC graph

Sub-Graphs (Obtained from Data)





Omission of latent variables could lead to bi-directional (pair 2-3) and /or "instantaneous" effects (pair 3-4)

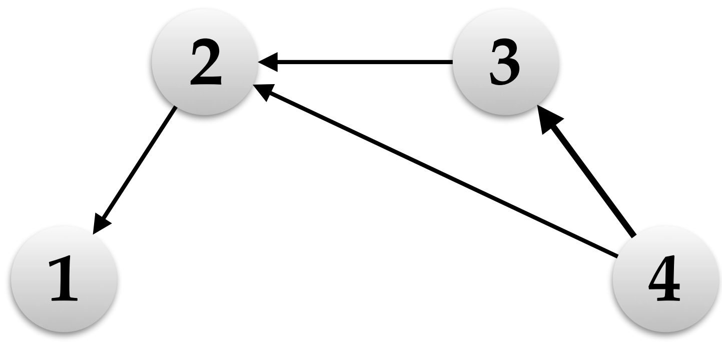
Sub-graphs derived from data do not necessarily tally with sub-graphs derived from the full-set reconstructed graph

This issue is addressed by constructing "mixed" graphs that are consistent with the data for all pairs. See Eichler (2005).

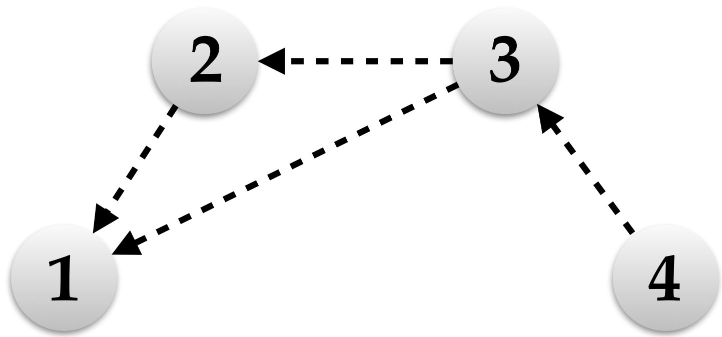
Mixed Graph: Example (Eichler, 2005)

$$z_1[k] = \alpha z_2[k-1] + e_1[k]$$
$$z_2[k] = \beta L[k-2] + e_2[k]$$
$$z_3[k] = \gamma L[k-1] + \delta z_4[k-1] + e_3[k]$$
$$z_4[k] = e_4[k]$$

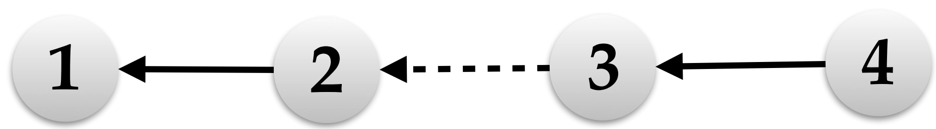
- *L* is a latent (unobserved) variable.
- Both path diagrams are not "consistent" with each other.
- A "mixed" graph based on the global Granger-causal Markov property is more appropriate.



Multivariate path diagram



Bivariate path diagram



Mixed path diagram

Missing Observations

FORMULATION

Measurements are obtained as:

$$\mathbf{y}_i = \mathbf{L}_i \mathbf{s}_i, \quad i = 1, 2, \ldots, P$$

where, $\mathbf{L}_i \in \mathbb{R}^{M \times N}$ is the sampling matrix.

Assume signal $\mathbf{s}_i$ has sparse representation in some "*dictionary*" $\mathbf{B} \in \mathbb{C}^{N \times N}$.

$$\mathbf{s}_i = \mathbf{B}\mathbf{x}_i$$

$$\mathbf{y}_i = \mathbf{L}_i \mathbf{B} \mathbf{x}_i = \mathbf{D}_i \mathbf{x}_i$$

$\mathbf{x}_i \in \mathbb{C}^{N \times 1}$ is the sparse representation of $\mathbf{s}_i$ and $\mathbf{D}_i \in \mathbb{C}^{M \times N}$ is known as over-complete dictionary in the compressed sensing literature.

Given $\mathbf{D}_i$ (or alternatively given $\mathbf{B}$ and $\mathbf{L}_i$), we can recover $\mathbf{x}_i$ by solving the following optimization problem:

$$\arg\min_{\mathbf{x}_i} \; ||\mathbf{x}_i||_1 + \lambda ||\mathbf{x}_i||_2^2 \quad \text{s.t.} \quad \mathbf{y}_i = \mathbf{D}\mathbf{x}_i$$

The signal is reconstructed using:

$$\mathbf{s}_i = \mathbf{B}\hat{\mathbf{x}}_i$$

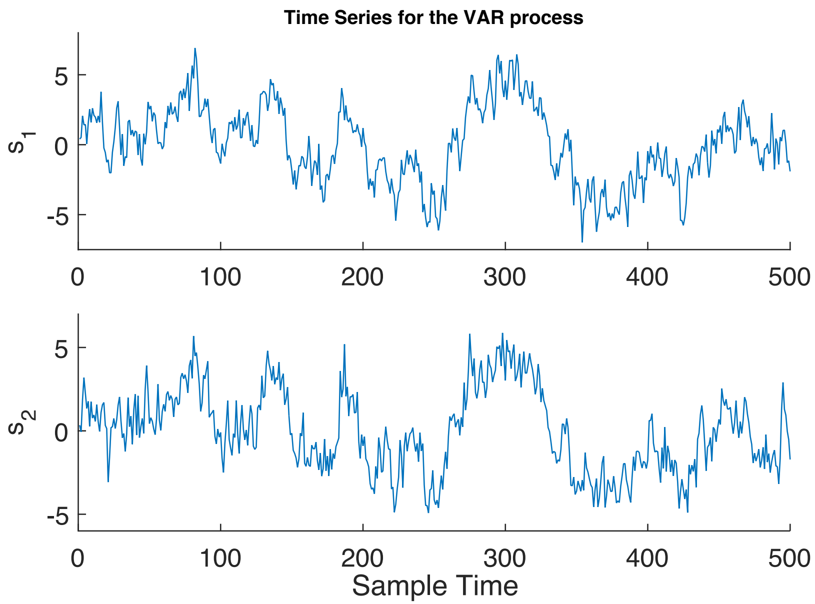
PROPOSED METHOD

1. Generate complete data set $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots \mathbf{s}_P]$ by incorporating labels for missing observations.
2. For each signal $\mathbf{s}_i$, execute the following sub-algorithm:
   (a) Set $\mathbf{B}$ to Discrete Fourier Transform (DFT) basis.
   (b) Obtain a preliminary estimate of the signals.
   $$\arg\min_{\mathbf{x}_i} \|\mathbf{x}_i\|_1 + \lambda\|\mathbf{x}_i\|_2^2 \quad \text{such that} \quad \mathbf{y}_i = \mathbf{D}_i\mathbf{x}_i$$
   (c) Update dictionary $\mathbf{B}$ using estimates obtained in the first step.
   $$\hat{\Phi} = \left[\hat{\mathbf{s}}_1[k-1] \; .. \; \hat{\mathbf{s}}_1[k-r] \; ... \; \hat{\mathbf{s}}_P[k-1] \; .. \; \hat{\mathbf{s}}_P[k-r]\right]$$
   $$\mathbf{B}_1 = \left[\mathbf{B} \quad \hat{\Phi}\right]$$
   (d) Normalize the columns of updated dictionary $\mathbf{B}_1$.
   (e) Re-estimate the signals using step 2(b) with the updated dictionary.

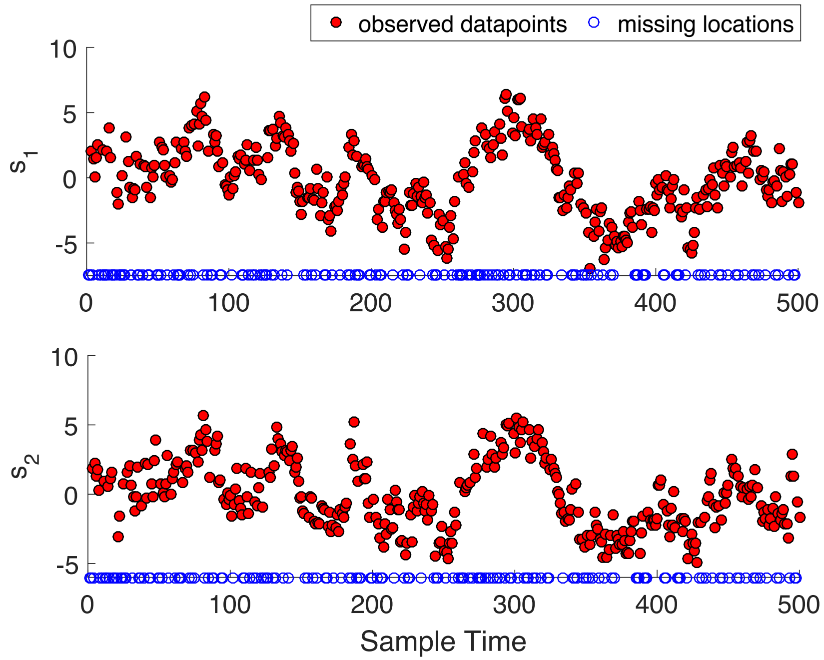
Case Study

Data generating process: VAR(1)

$$s_1[k] = 0.6s_1[k-1] + 0.4s_2[k-1] + e_1[k]$$
$$s_2[k] = 0.4s_1[k-1] + 0.5s_2[k-1] + e_2[k]$$

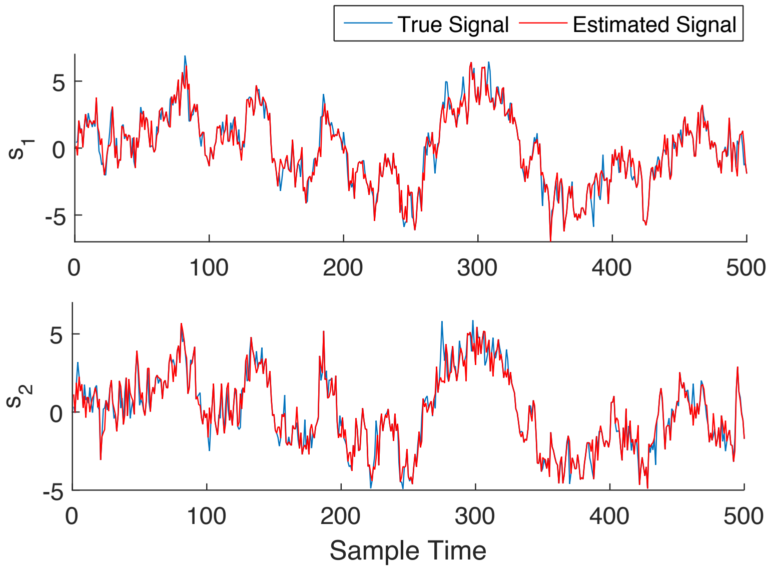


(a) Complete data record

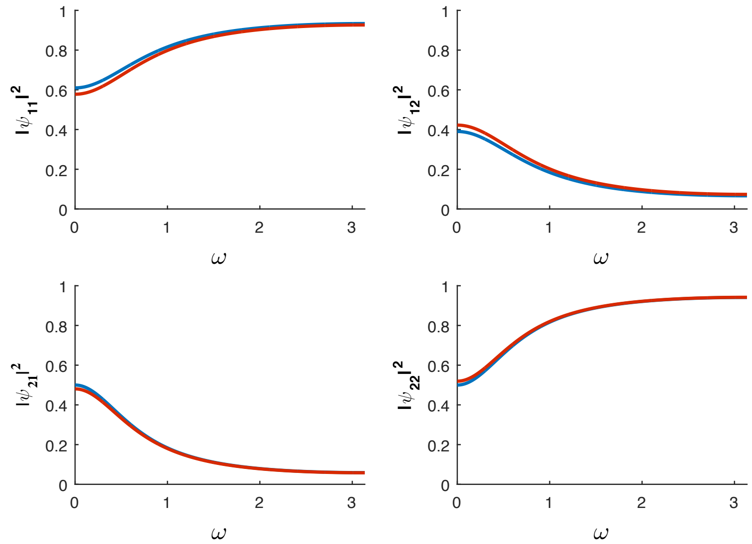


(b) Time-series with missing data

Case Study: Contd.



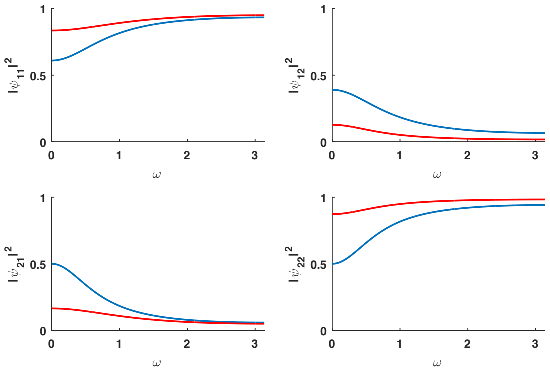
(a) Reconstructed Signal



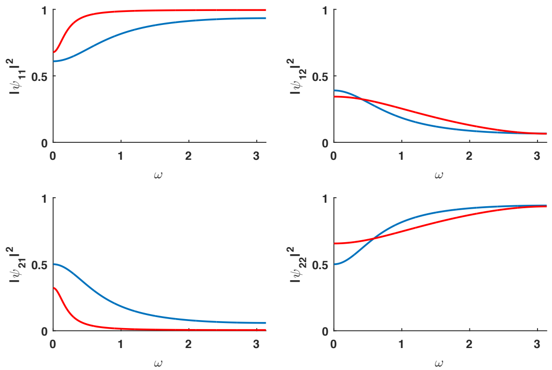
(b) Normalized DPT

	(5% missing)	(10% missing)	(20% missing)	(30% missing)
NMSE fit (averaged)	0.989, 0.987	0.977, 0.971	0.950, 0.940	0.931, 0.925
Strength of Connectivity (averaged)	$\mathbf{s}_1 \rightarrow \mathbf{s}_2 = 0.556$ $\mathbf{s}_2 \rightarrow \mathbf{s}_1 = 0.774$	$\mathbf{s}_1 \rightarrow \mathbf{s}_2 = 0.548$ $\mathbf{s}_2 \rightarrow \mathbf{s}_1 = 0.757$	$\mathbf{s}_1 \rightarrow \mathbf{s}_2 = 0.533$ $\mathbf{s}_2 \rightarrow \mathbf{s}_1 = 0.737$	$\mathbf{s}_1 \rightarrow \mathbf{s}_2 = 0.517$ $\mathbf{s}_2 \rightarrow \mathbf{s}_1 = 0.721$

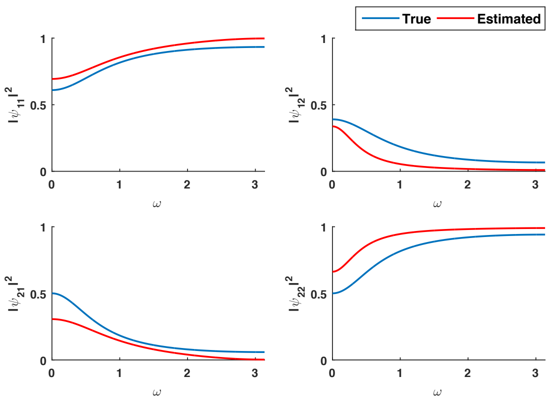
Comparison (30% Missing Data)



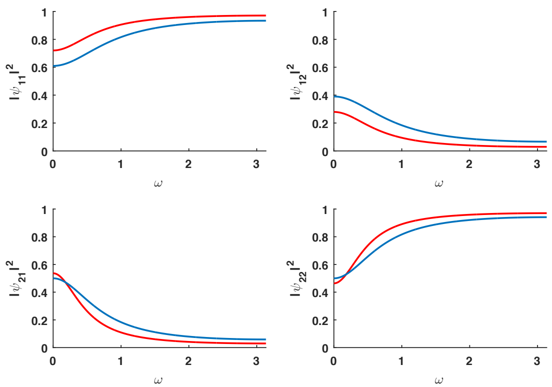
Mean imputation method



L-S Periodogram method



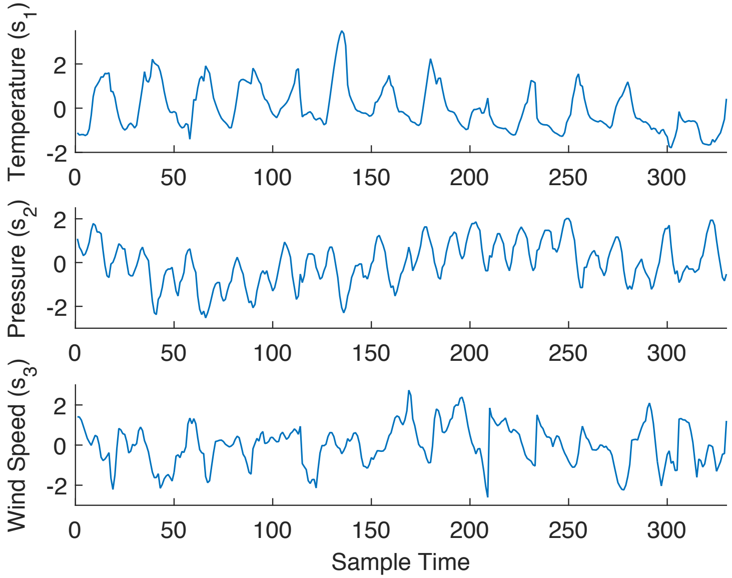
Modified L-S Periodogram method



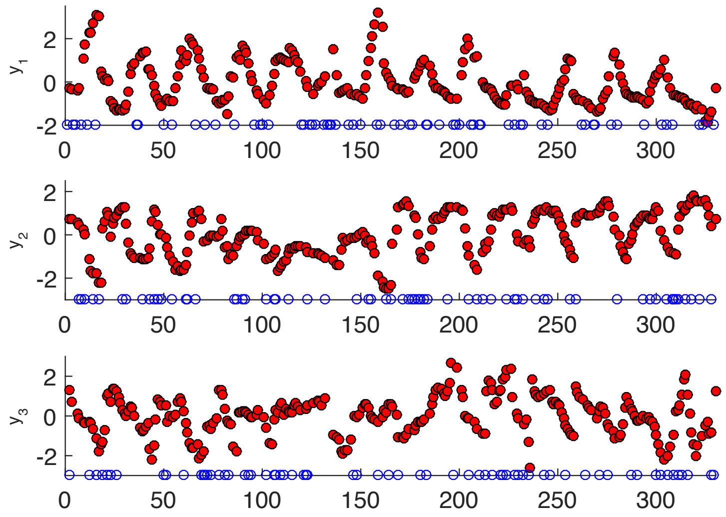
'MTSDI' method

	Mean imputation method	L-S Periodogram	Modified L-S Periodogram	'MTSDI' method
NMSE fit (averaged)	0.688, 0.689	0.139, 0.061	0.544, 0.522	0.912, 0.892

Climate Data



(a) Meteorological Time-series

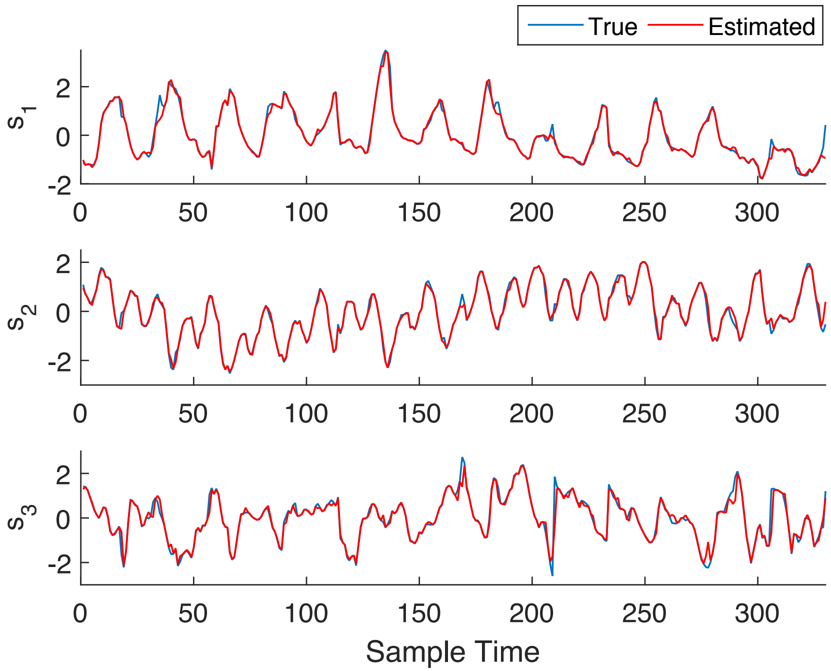


(b) Time-series with missing observations

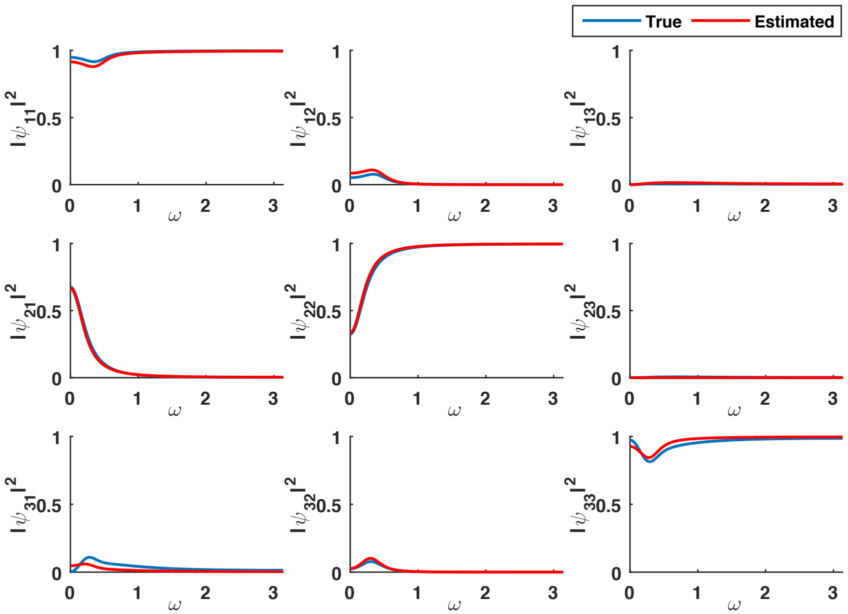
Complete and incomplete time-series (30% missing data)

Note: ***Missing data located at different locations for each signal***

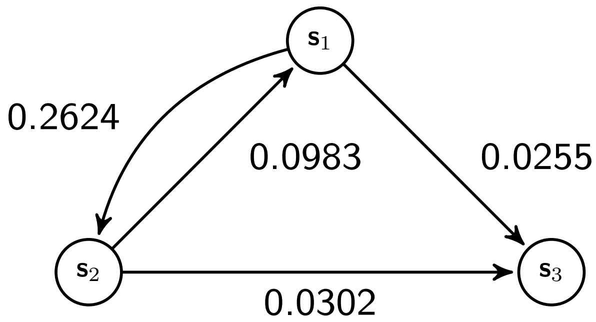
RESULTS



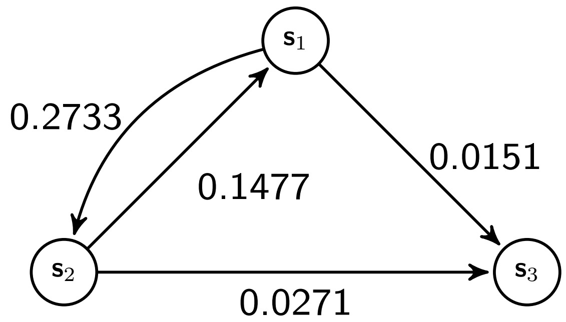
(a) Reconstructed Signal



(b) Normalized DPT



(a) Complete data case



(b) Incomplete data case

Non-Linearities

HANDLING NON-LINEAR PROCESSES

There exist several approaches to determining causality in non-linear processes.

- Transfer entropy
- Local linear predictors
- General information theoretic-based measures
- Correlation integrals
- Kernel-based ideas (recent)
- …

See *Schindler et al (2007). Causality detection based on information-theoretic approaches in time-series analysis, Physics Reports, 441, 1-46.*

Motivating Thoughts

- Second-order statistics measure linear (directed) relationships (useful for linear and/or Gaussian processes).
- Information-theoretic tools quantify / detect non-linear connections.
- Entropy / information-based techniques mainly require estimation of joint density functions -
  - *Computation for multivariable case (> 3 variables) becomes very cumbersome and demanding*

Idea: *Have the simplicity of a linear tool (correlation) and the capabilities of a non-linear measure (entropy-based)*

CORRENTROPY / GENERALIZED CORRELATION

Correntropy [Santamaría et al., 2006]

It is the covariance between two variables in a "kernel" space

$$V_{Z_j Z_i} = E(\kappa_\sigma(Z_j - Z_i)) \quad (1)$$

where $\kappa_\sigma(.)$ is a kernel function inducing a non-linear mapping to a RKHS

$$\kappa_\sigma(z_j - z_i) = \langle \phi(z_j), \phi(z_i) \rangle$$

A **translation-invariant** Gaussian kernel is commonly used

$$\kappa_\sigma(z_j - z_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{||z_j - z_i||_2^2}{2} \right)$$

Auto-Correntropy Function

Auto-correntropy function

Apply the correntropy in the same way as correlation, to measure lagged effects

$$
\begin{aligned}
V_{zz}(t_1, t_2) &= E[k_\sigma(z[t_1] - z[t_2])] \\
&= \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E\left[\|z[m] - z[m-l]\|^{2n}\right]
\end{aligned} \tag{2}
$$

Estimator

An unbiased and consistent estimator is given by

$$
\hat{V}[l] = \frac{1}{N - l + 1} \sum_{n=l}^{N} k(z[n] - z[n-l]) \tag{3}
$$

PROPERTIES

- Has the same interpretation of an auto-covariance function
- **Even moments** are considered, unlike all moments in an MI-based measure
- **Higher-order moments** allow it to distinguish between distributions
- **Linearity** makes it computationally elegant
- Non-linearities are captured by working in the **kernel space.**

  For an independent stationary processes $z_l[k]$, the auto-correntropy is independent of the lag:

$$\boxed{V[l] = V[0]}$$

EXAMPLE

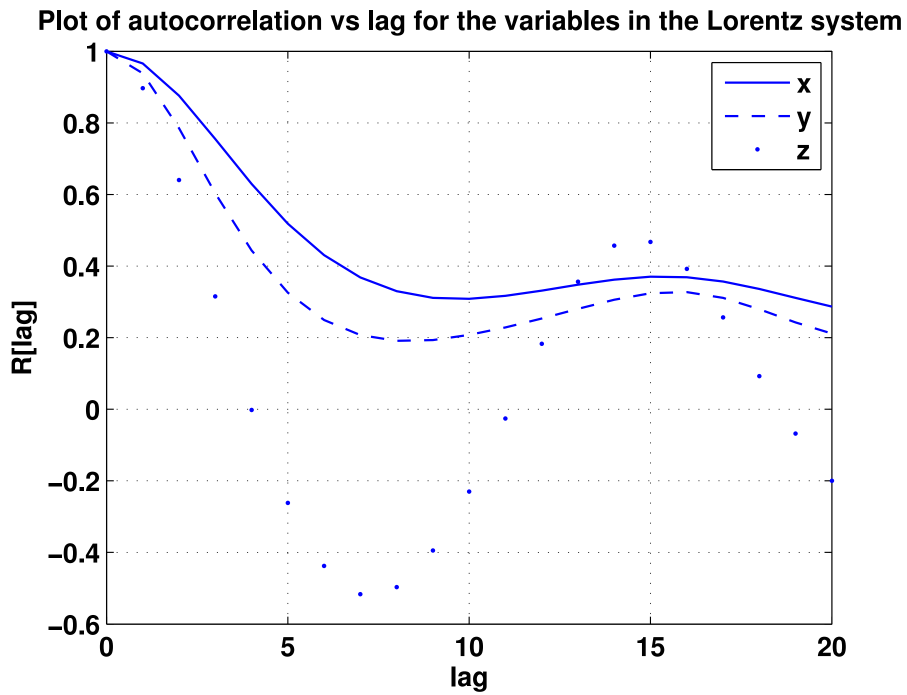
Chaotic Lorentz attractor system (Santamaria et al, 2006)

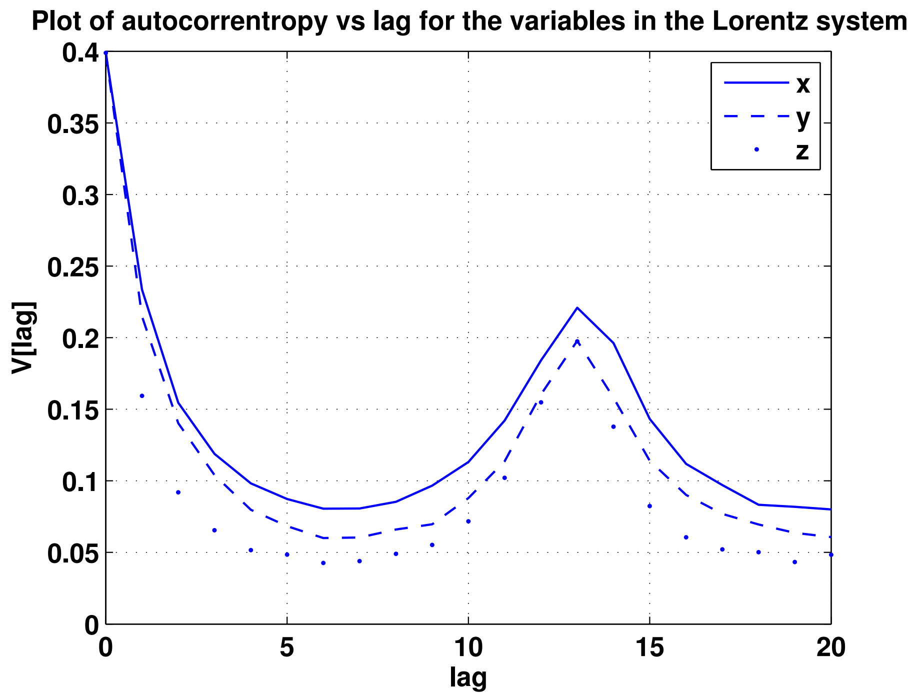
$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = -y - xz + Rx$$
$$\dot{z} = xy - bz$$

$R = 28$

$\sigma = 10$

$b = 10/3$





Kernel PDC

- **Idea:** Construct PDC in the kernel space using VAR models on kernel transformed data.

KPDC

$$\phi(\mathbf{z}[k]) = -\sum_{r=1}^{p} \mathbf{A}_r^{\phi}\, \phi(\mathbf{z}[k-r]) + \mathbf{v}[k]$$

$$\pi_{ij}^{\phi}(\omega) = \frac{\bar{a}_{ij}^{\phi}(\omega)}{(\bar{a}^{\phi})_{.j}^{*}(\omega)\bar{a}_{.j}^{\phi}(\omega)} = \frac{\bar{a}_{ij}^{\phi}(\omega)}{\sum_{i=1}^{m}\left|\bar{a}_{ij}^{\phi}(\omega)\right|^2}$$

where

$$\bar{\mathbf{A}}^{\phi}(\omega) = \mathbf{I} - \mathbf{A}^{\phi}(\omega); \qquad \mathbf{A}^{\phi}(\omega) = \sum_{r=1}^{p} A_r^{\phi} z^{-r}\Big|_{z=e^{-j\omega}} \tag{5}$$

Estimating KPDC

Idea: Build VAR models using Y-W method on correntropy estimates

- Transformation of variables not explicitly required.
- As in classical VAR modelling, only correntropy estimates are sufficient

For an M-dimensional multivariate process, denote

$$\mathbf{z} = \begin{bmatrix} z_1 & z_2 & \cdots & z_M \end{bmatrix}^T$$

The VAR model in the kernel space is given by

$$\mathbf{z}^{\phi}[k] = -\sum_{r=1}^{p} \mathbf{A}_r^{\phi} \, \mathbf{z}^{\phi}[k-r] + \mathbf{v}[k]$$

Build VAR Models in Kernel Space

Estimate $\mathbf{A}_r$ using the Y-W equations in the kernel space

$$\sum_{r=0}^{p} \mathbf{A}_r^{\phi}\, \mathcal{V}[r] = \mathbf{\Sigma}_v$$

$$\sum_{r=0}^{q-1} \mathbf{A}_r^{\phi}\, \mathcal{V}^{\mathsf{T}}[q-r] + \sum_{r'=q}^{p} \mathbf{A}_{r'}^{\phi}\, \mathcal{V}[r' - q] = \mathbf{0}, \;\; q = 1, \cdots, p \qquad (8)$$

Subsequently, estimate KPDC in a manner similar to estimating PDC

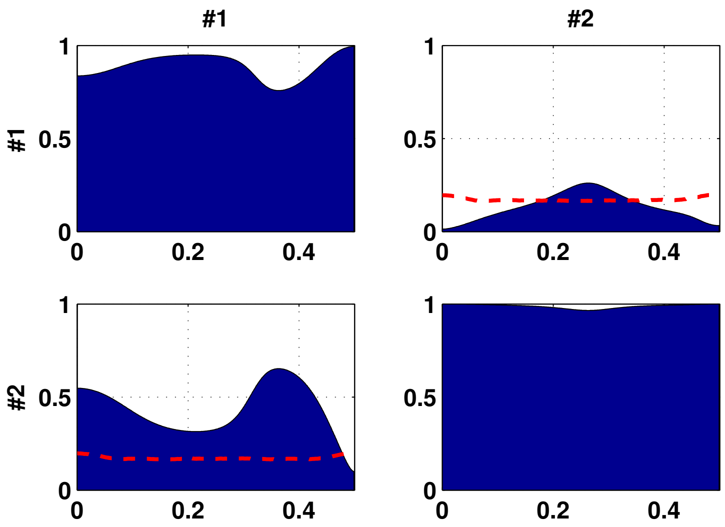
$$\mathbf{A}^{\phi}(\omega) = \sum_{r=1}^{p} A_r^{\phi} z^{-r}\Big|_{z=e^{-j\omega}} \qquad \bar{\mathbf{A}}^{\phi}(\omega) = \mathbf{I} - \mathbf{A}^{\phi}(\omega)$$

$$\Gamma_{ij}(\omega) = \frac{\bar{a}_{ij}^{\phi}(\omega)}{\bar{a}_{.j}^{\phi^*}(\omega)\bar{a}_{.j}^{\phi}(\omega)} = \frac{\bar{a}_{ij}^{\phi}(\omega)}{\sum_{i=1}^{m} \left|\bar{a}_{ij}^{\phi}(\omega)\right|^2} \qquad (8)$$

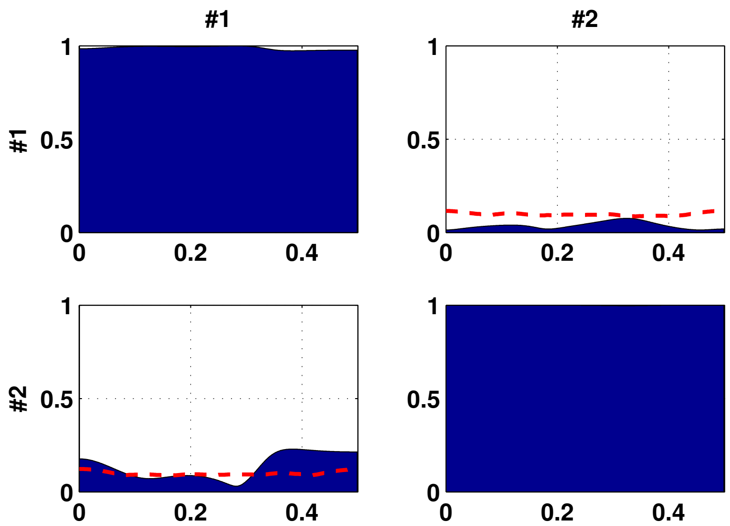
2-D Case Study

$$x_1[k] = 1 - ax_1^2[k-1] + se_1[k]$$
$$x_2[k] = (1-c)(1 - ax_2^2[k-1]) + c(1 - ax_1^2[k-1]) + se_2[k]$$

with $a = 1.8$, $s = 0.01$ and $c = 0.2$



Classical PDC



Kernel PDC using σ = 0.23

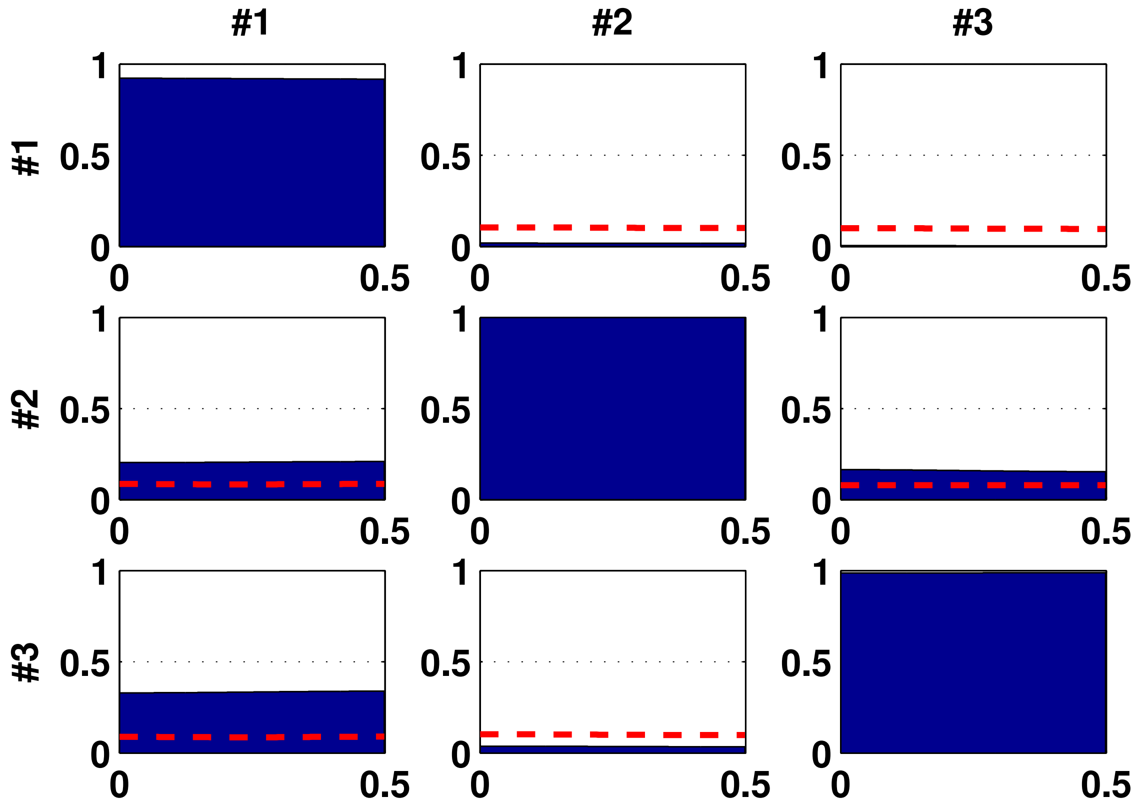
Significance limits are computed using bootstrapping

3-D Case Study

- Three-variable system with one of the variables having a non-Gaussian distribution [Duan et al., 2012]

$$y[k] = 5(z[k-1] + 7.2)^2 + 10\sqrt{|x[k-1]|} + v_1[k]$$
$$z[k] = 1 - 2\left|0.5 - (0.8\,x[k-1] + 0.4\sqrt{z[k-1]})\right| + v_2[k]$$

- x[k] ∈ *U*[4,5]
- First 3000 samples are discarded to generate stationarity of samples.
- KPDC correctly identifies the effective connectivity.



Case Study: Bioreactor System

A three-state bioreactor model [Enszer et al., 2008, Lin and Stadtherr, 2007a] with concentration of the cells (x_1), the substrate (x_2) and the product (x_3) as states:

$$\dot{x}_1 = (\mu - D)x_1$$

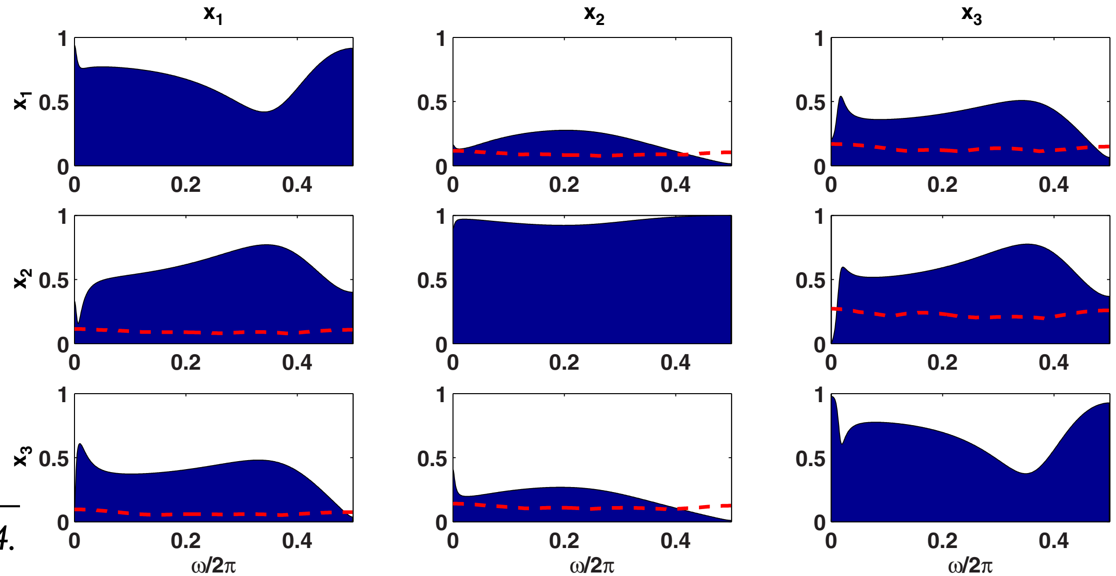
$$\dot{x}_2 = D(x_{2f} - x_2) - \frac{\mu x_1}{Y}$$

$$\dot{x}_3 = -Dx_3 + (\alpha\mu + \beta)x_1$$

where the growth rate μ is

$$\mu = \frac{\mu_{max}\left[1 - \frac{x_3}{x_{3m}}\right]x_2}{k_s + x_2}$$

with $x_1(0) = 6.50$ g/L, $\mu_{max} = 0.46$, $k_s = 1.1$, $x_2(0) = 5$ g/L, $x_3(0) = 15$ g/L, $Y = 0.4$ g/g, $\beta = 0.2\text{hr}^{-1}$, $D = 0.202\text{hr}^{-1}$, $\alpha = 2.2$ g/g, $x_{3m} = 50$ g/L and $x_{2f} = 20$ g/L.



R. Kannan and A.K. Tangirala (2014). Physical Reviews E, 89, 062144.

REFERENCES

- Baccala L, Sameshima K (2001) Partial directed coherence: a new concept in neural structure determination. Biological Cybernetics 84(6):463–474
- Carlo Berzuini, Philip Dawid and Luisa Bernardinelli (2012). Causality: Statistical Perspectives and Applications (Edited). John Wiley and Sons.
- Eichler M (2005). A graphical approach for evaluating effective connectivity in neural systems. Philosophical Transactions of the Royal Society of London B, 360, 953-967.
- Eichler M (2013) Causal inference with multiple time series: principles and problems. Philosophical Transactions of the Royal Society of London A, 371, DOI:10.1098/rsta.2011.0613
- Geweke JF (1984) Measures of conditional linear dependence and feedback between time series. Journal of the American Statistical Association 79(388):907–915,
- Granger C (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37(3):424–38
- Lutkepohl H (2005) New Introduction to Multiple Time Series Analysis. Springer
- Sims CA (1972) Money, income, and causality. The American Economic Review 62(4):540–552,
- Wei W (2005). Time Series Analysis: Univariate and Multivariate Methods. Pearson
- Wiener N (1956) The theory of prediction. In: Beckenbach E (ed) Modern mathematics for the engineer, ed. E. F. Beckenbach, McGraw-Hill, New York,, McGraw-Hill

REFERENCES

- Gigi S. and A.K. Tangirala (2010) Quantitative analysis of directional strengths in jointly stationary linear multivariate processes. Biological Cybernetics 103(2):119–133.
- S. Gigi and A.K. Tangirala (2013). Quantification of interaction in multiloop interacting systems using directed spectral decomposition. Automatica, 49(5), 1174-1183.
- S. Gigi (2013). Reconstruction and Quantification of Interaction Pathways in Multivariate Systems Using Directed Spectral Analysis. Ph.D. Thesis, IIT Madras.
- A. Garg (2013). Causality analysis for topology reconstruction and interaction assessment. MS Thesis, IIT Madras.
- R. Kannan and A.K. Tangirala (2014). Correntropy-based partial directed coherence for testing multivariate Granger causality in nonlinear processes. Physical Reviews E, 89, 062144.
- S. Kathari and A.K. Tangirala (2016). Estimation of network connectivity strengths in linear causal dynamic systems, ACODS 2016, pp. 77-82, NIT Tiruchi, India.
- P. Agarwal and A.K. Tangirala (2017). Reconstruction of missing data in multivariate processes with applications to causality analysis. Special Issue (on Data Sciences) of International Journal of Advances in Engineering Sciences and Applied Mathematics.
- A. Garg and A.K. Tangirala (2018). Metrics for interaction assessment in multivariable control systems using directional analysis. Industrial & Engineering Chemistry Research, 57 (3), 967-979. DOI: 10.1021/acs.iecr.7b03671