

**ISINP 2017** Como - 28.7.2017



## **Information Theory for Network Physiology**

Multivariate and Multiscale Methods to Dissect the Information Content of Brain, Cardiovascular and Muscular Networks

# Luca Faes

IRCS-HTA Laboratory, Bruno Kessler Foundation, Trento, Italy BIOtech, University of Trento, Italy





### **NETWORK PHYSIOLOGY:** A NEW FIELD IN SYSTEM MEDICINE AND BIOLOGY



Introduction

"The human organism is an integrated network where complex physiologic systems, each with its own regulatory mechanisms, continuously interact, and where failure of one system can trigger a breakdown of the entire network" [A. Bashan et al., Nature Communications 2012]

A new field, **Network Physiology**, is needed to probe the interactions among diverse physiologic systems

### Different organ systems dynamically interact to accomplish vital functions

- Traditional, "reductionist" approach To study the function of single organ in isolation
- New approach, fostered by Network Physiology

To look simultaneously at multiple organs

Each organ system is seen as a node of a complex network of physiological interactions



Introduction



## How to extract valuable information from physiological signals?





Introduction



### Network Physiology has high potential but is still in large part unexplored



- Challenging problems for physicists, engineers and physiologists
- We face these issues with the unifying framework of INFORMATION DYNAMICS





# **INFORMATION DYNAMICS: THEORY**

- Information-theoretic analysis of dynamical systems
- Information Measures
- Information Decomposition
- Information Storage
- Information Transfer
- Information Modification



- Investigation of Statistical dependencies:
  - ✓ SELF effects:
      $Y_n^- \to Y_n$  → Information storage

     ✓ CAUSAL effects:
      $X_n^- \to Y_n$  → Information transfer

     ✓ INTERACTION effects:
      $(X_{1,n}^- \leftrightarrow X_{2,n}^-) \to Y_n$  → Information modification





### THE FRAMEWORK OF INFORMATION DYNAMICS

Decomposition of the "information" contained in the target process



• Computation: basic information theoretic measures





Applications to Network Physiolog



### TARGET INFORMATION DECOMPOSITION



- Present Information about Y : H<sub>Y</sub> = H(Y<sub>n</sub>)
   Information contained in the present of the process Y
- **Predictive Information** about  $Y : P_Y = I(Y_n; Y_n^-, X_n^-)$ Information contained in the past of *S*=(*X*,*Y*) that can be used to predict the present of the target *Y*
- New information about Y :  $N_Y = H(Y_n | Y_n^-, X_n^-)$ Information contained in the present of Y that cannot be

Information contained in the present of *Y* that cannot be predicted from the past of S=(X,Y)

Uncertainty about the present state of the target

*Predictability of the target given the past network states* 



*Information generated in the target by the state transition* 



Applications to Network Physiolog



### PREDICTIVE INFORMATION DECOMPOSITION



- **Predictive Information** about Y :  $P_Y = I(Y_n; Y_n^-, X_n^-)$ Information contained in the past of *S*=(*X*, *Y*) that can be used to predict the present of the target *Y*
- **Information Storage** in Y :  $S_Y = I(Y_n; Y_n^-)$

Information contained in the past of *Y* that can be used to predict its present

• Information transfer from X to Y :  $T_{X \rightarrow Y} = I(Y_n; X_n^- | Y_n^-)$ 

Information contained in the past of X that can be used to predict the present of Y above and beyond the information contained in the past of Y

- Predictability of the target given the network past states
  - *Predictability of the target from its own past states* 
    - Causal interactions from all sources to the target



Applications to Network Physiolog<sup>y</sup>



### **INFORMATION TRANSFER DECOMPOSITION**





• Information transfer from X to Y :  $T_{X \rightarrow Y} = I(Y_n; X_n^- | Y_n^-)$ 

Information contained in the past of X that can be used to predict the present of Y above and beyond the information contained in the past of Y

Causal interactions from all sources to the target

Transfer

• Conditional information transfer:  $T_{X_1 \rightarrow Y|X_2} = I(Y_n; X_{1,n} | Y_n, X_{2,n})$ 

Information contained in the past of  $X_1$  that can be used to predict the present of *Y* above and beyond the information contained in the past of *Y* and  $X_2$ 

• Interaction information transfer:  $I_{X_1;X_2|Y}^Y = I(Y_n;X_{1,n}^-;X_{2,n}^-|Y_n^-)$ 

Information contained in the past of  $X_1$  and  $X_2$  that can be used to predict the present of Y when  $X_1$  and  $X_2$  are taken individually but not when they are taken together

*Causal interactions from one source to the target* 

Redundant or synergistic interactions contributing to transfer





### **INFORMATION MODIFICATION: REDUNDANCY AND SYNERGY**

• Interpretation of Information Modification:  $I_{X_1:X_2}^Y = T_{X_1 \to Y} + T_{X_2 \to Y} - T_{X_1,X_2 \to Y}$ 



Interaction information can be negative: synergy!



Applications to Network Physiology



### THE FRAMEWORK OF INFORMATION DYNAMICS



**L Faes**, A Porta, G Nollo, 'Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics', **Entropy, special issue on "Entropy and** *Cardiac Physics"*, **2015**, 17:277-303.

**L Faes**, A Porta, G Nollo, M Javorka, 'Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks', *Entropy, special issue on Multivariate Entropy Measures and their applications*, 2017, 19(1), 5





# **INFORMATION DYNAMICS: ESTIMATION**

- Linear model-based estimator
- Nonlinear model-free estimators  $\rightarrow$
- Challenges of model-free estimation
- Binning
- Kernel
  - Nearest neighbor





### PRACTICAL COMPUTATION OF INFORMATION DYNAMICS

- All measures of Information dynamics are expressed in terms of measures of ٠ (conditional) entropy, (conditional) mutual information, or interaction information
- Estimation of entropy for variables with different dimension ٠
- **Example: Information Storage** •

$$Y_n^ Y_n^ Y_n^-$$
 target  $Y$ 

$$S_Y = I(Y_n; Y_n^-) = H(Y_n) - H(Y_n \mid Y_n^-) = H(Y_n) - H(Y_n, Y_n^-) + H(Y_n^-)$$

#### Approximation of the past history

$$Y_n^- = [Y_{n-1}Y_{n-2}\cdots] \longrightarrow Y_n^- \cong Y_n^L = [Y_{n-1}Y_{n-2}\cdots Y_{n-L}]$$

#### Computation

Discrete variables

$$H(Y_n) = -\sum_{y_n \in \Omega_Y} p(y_n) \log p(y_n) \qquad \qquad H(Y_n^-) \cong H(Y_n^L) = -\sum_{y_{n-1}} \cdots \sum_{y_{n-L}} p(y_n^L) \log p(y_n^L)$$

Continuous variables

$$H(Y_n) = -\int_{D_Y} p(y_n) \log p(y_n) dy_n \qquad \qquad H(Y_n^-) \cong H(Y_n^L) = -\int_{D_Y} \cdots \int_{D_Y} p(y_n^L) \log p(y_n^L) dy_{n-1} \dots dy_{n-L}$$





### **PARAMETRIC ESTIMATION: LINEAR METHOD**

• Exact Computation under the assumption of Gaussianity

✓ Entropy of 
$$Y_n$$
:  $H_Y = H(Y_n) = \frac{1}{2} \ln 2\pi e \cdot \sigma(Y_n)$ 

✓ Conditional Entropy of  $Y_n$  given  $Y_n^- \cong Y_n^L$ :  $H(Y_n | Y_n^-) \cong H(Y_n | Y_n^L) = \frac{1}{2} \ln 2\pi e \cdot \sigma(Y_n | Y_n^L)$  *linear regression of*  $Y_n$  on  $Y_{n-L}$ :  $Y_n = a_1 Y_{n-1} + \cdots + a_L Y_{n-L} + W_n$  $\sigma(Y_n | Y_n^L) = \sigma_W^2$  [Barnett et al, Phys Rev Lett 2009]

✓ Information Storage of Y :  $S_Y = I(Y_n; Y_n^-) \cong H(Y_n) - H(Y_n | Y_n^L) = \frac{1}{2} \ln \frac{\sigma_Y^2}{\sigma_W^2}$ 



**L Faes**, A Porta, G Nollo, M Javorka, 'Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks', *Entropy, special issue on Multivariate Entropy Measures and their applications*, 2017, 19(1), 5





## NONPARAMETRIC ESTIMATION: BINNING METHOD

- Discretization of continuous random variables using quantization levels
  - ✓ Target Information of Y :  $H_Y = H(Y_n)$
  - ✓ Information Storage of Y :  $S_Y = I(Y_n; Y_n^-) \cong I(Y_n; Y_n^L) = \hat{H}(Y_n) \hat{H}(Y_n, Y_n^L) + \hat{H}(Y_n^L)$



L Faes, A Porta, 'Conditional entropy-based evaluation of information dynamics in physiological systems', in *Directed Information Measures in Neuroscience*, R Vicente, M Wibral, J Lizier (eds), Springer-Verlag; 2014, pp. 61-86





### **NONPARAMETRIC ESTIMATION: KERNEL METHOD**

- Entropy computation using kernel functions to weight distances between points
  - Probability of *d*-dimensional variable X:  $\hat{p}(x_n) = \frac{1}{N} \sum_{i=1}^{N} K(||x_n x_i||)$ Heaviside kernel function:  $K = \Theta(||z_n - z_i||) = \begin{cases} 1 & ||z_n - z_i|| \le r \\ 0 & ||z_n - z_i|| > r \end{cases}$
  - **Entropy:**  $H(X) = -E[\log p(x)] \cong -\log\langle \hat{p}(x) \rangle$

✓ Target Information of Y :  $H_Y = H(Y_n) = -\log \langle p(y_n) \rangle$ 

✓ Information Storage of Y :  $S_Y = I(Y_n; Y_n^-) \cong I(Y_n; Y_n^L) = \hat{H}(Y_n) - \hat{H}(Y_n, Y_n^L) + \hat{H}(Y_n^L)$  →  $S_Y = \log \left\langle \frac{p(y_n, y_n^L)}{p(y_n) p(y_n^L)} \right\rangle$ 

• Example: L=1  $p(y_n, y_{n-1}) \cong 2/7$   $\hat{H}(Y_n, Y_{n-1})$  $Y_n^L \cong Y_n^1 = Y_{n-1}$  $p(y_{n-1}) \cong 3/7 \quad \Longrightarrow \quad \hat{H}(Y_{n-1})$ Range search with fixed inreshold, r

W Xiong, L Faes, P Ch Ivanov, 'Entropy measures, entropy estimators and their performance in quantifying complex dynamics: effects of artifacts, nonstationarity and long-range correlations', Phys. Rev. E, in press, 2017.





### **NONPARAMETRIC ESTIMATION: NEAREST NEIGHBOR METHOD**

• Entropy computation from the statistics of distances between neighboring points in a multidimensional space

$$H(X) = -\mathbb{E}[\log p(x)] \cong -\psi(k) + \psi(N) + d\langle \log \varepsilon_n \rangle$$

- $\psi$ : Digamma function  $\psi(x) = \frac{d \log \Gamma(x)}{dx}$  $\mathcal{E}$ : 2·distance from  $x_n$  to its k-th neighbor N: number of outcomes of X
- Strategy for bias compensation in the estimation of entropies for variables of different dimension

$$\hat{H}(Y_n, Y_n^L) = -\psi(k) + \psi(N) + (L+1) \langle \log \varepsilon_n \rangle \longrightarrow \text{Neighbor Search}$$

$$distance from (y_n, y_n^L) \text{ to its } k \text{-th neighbor in the outcomes of } (Y_n, Y_n^L)$$

$$\hat{H}(Y_n) = - \langle \psi(N_{Y_n}) \rangle + \psi(N) + \langle \log \varepsilon_n \rangle \longrightarrow \text{Range Search}$$

$$\hat{H}(Y_n^L) = - \langle \psi(N_{Y_n^L}) \rangle + \psi(N) + L \langle \log \varepsilon_n \rangle \longrightarrow \text{Range Search}$$
number of outcomes of  $Y_n^L$  with distance to  $y_n^L$  strictly lower than  $\varepsilon_n/2$ 

$$S_Y = \hat{H}(Y_n) - \hat{H}(Y_n, Y_n^L) + \hat{H}(Y_n^L) = \psi(N) + \psi(k) - \left\langle \psi(N_{Y_n^L}) + \psi(N_{Y_n}) \right\rangle$$







### **ESTIMATION: SIMULATION EXAMPLE**

- Test on stochastic process with short-term dynamics and long-range correlations
- Fractionally-integrated autoregressive process:  $A(L)(1-L)^d Y_n = U_n$

Autoregressive polynomial: sets stochastic oscillation with amplitude  $\rho$  and frequency f  $A(L) = 1 - 2\rho \cos 2\pi f L - \rho^2 L^2$ 

**Fractional differencing**:  $(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)}$ sets long-range correlations depending on the differencing parameter d

• Estimation of Entropy, Conditional Entropy and Information Storage

 $S_Y = I(Y_n; Y_n^-) = H(Y_n) - H(Y_n | Y_n^-) \qquad L=2 \implies Y_n^- \cong Y_n^L = [Y_{n-1}Y_{n-2}]$ 

• Estimation: theoretical profiles and estimates, linear method



W Xiong, L Faes, P Ch Ivanov, 'Entropy measures, entropy estimators and their performance in quantifying complex dynamics: effects of artifacts, 19 nonstationarity and long-range correlations', Phys. Rev. E, in press, 2017.





20

### **ESTIMATION: SIMULATION EXAMPLE**

• Estimation: theoretical profiles and estimates, non-parametric model-free methods







### **MODEL-FREE ESTIMATION: APPROXIMATION OF THE SYSTEM PAST**

• Uniform embedding (UE):  $X_n^- \approx [X_{n-m_x} \dots X_{n-L_xm_x}]$   $Y_n^- \approx [Y_{n-m_y} \dots Y_{n-L_ym_y}]$ 

UE introduces irrelevant and redundant components Curse of dimensionality

**Non-uniform embedding (NUE):** The embedding vector is formed progressively, including at ٠ each step the lagged variable better describing the target process

### • Sequential procedure:

(a) k=0: Initialization Set of initial candidate components (e.g.,  $\Omega = \{X_{n-1}, \dots, X_{n-L}, Y_{n-1}, \dots, Y_{n-L}\}$ ) Initial embedding vector:  $V_n^{(0)} = [\cdot]$ 

(b)  $k \ge 1$ : Selection – maximum relevance, minimum redundancy

Select the component  $W_n \in \Omega$  that maximizes  $I(Y_n, W_n | V_n^{(k-1)}) \longrightarrow V_n^{(k)} = [\hat{W}_n, V_n^{(k-1)}]$ 

(c) Termination – randomization test

Generate N surrogates of  $\hat{W}_n$  by sample shuffling:  $\hat{W}_n^{(S_1)}, \dots, \hat{W}_n^{(S_N)}$ ; Threshold for  $I(Y_n, \hat{W}_n | V_n^{(k-1)}) : I_{th}$ Stop if  $I(Y_n, \hat{W}_n | V_n^{(k-1)}) < I_{th}$ ; final set of components:  $V_n = V_n^{(k-1)}$ 

(d) After termination – embedding vector  $V_n = [V_n^X, V_n^Y] \longrightarrow X_n^- \approx V_n^X = Y_n^- \approx V_n^Y$ 

L Faes, G Nollo, A Porta: 'Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique', Physical Review E; 2011; 83(5 Pt 1):051112.





## INFORMATION DYNAMICS: APPLICATIONS TO NETWORK PHYSIOLOGY

- Short-term Cardiovascular, Cardiorespiratory, Cerebrovascular control
- Brain-heart and brain-brain interactions during sleep
- Brain networks (EEG, fMRI) and muscular networks (EMG)





### **PRACTICAL APPLICATIONS OF INFORMATION DYNAMICS**

### Network of cardiovascular, cardiorespiratory and cerebrovascular short-term physiological interactions







### **Applications: CARDIAC CONTROL**



• The dynamical complexity of short-term heart period variability decreases progressively with tilt-table angle

- Linear estimator
- Univariate analysis ٠ New Information  $N_V$ Information Storage  $S_V$







### Complexity assessed by linear model-based estimators significantly correlates with model-free estimates



A Porta, B De Maria, V Bari, A Marchi, L Faes, 'Are nonlinear model-free approaches for the assessment of the entropy-based complexity of the cardiac control superior to a linear model-based one?', IEEE Trans. Biomed. Eng., in press, 2017

### **Applications: CARDIOVASCULAR and CARDIORESPIRATORY INTERACTIONS**

Protocol: 61 young healthy subjects during head-up tilt and mental stress tasks



- Measured time series:
  - Heart period (H)
  - Systolic arterial pressure (S)
  - Respiration (R)

300 points in each condition

• Linear estimator

#### Network analysis

full information decomposition:

$$\begin{split} H_Y &= N_Y + S_Y + T_{X \to Y} \\ T_{X \to Y} &= T_{X_1 \to Y \mid X_2} + T_{X_2 \to Y \mid X_1} + I_{X_1;X_2 \mid Y}^Y \\ target: Y = H \end{split}$$







Information Transfer Decomposition



**L Faes**, A Porta, G Nollo, M Javorka, 'Information decomposition in multivariate systems: definitions, implementation and application to cardiovascular networks', *Entropy, special issue on Multivariate Entropy Measures and their applications*, 2017, 19(1), 5

### **Applications: CARDIOVASCULAR AND CEREBROVASCULAR INTERACTIONS**

Protocol: 10 subjects with postural-related syncope





• Signals and time series



- Binning estimator with NUE
- Bivariate analysis, **target HP or FV** Entropy decomposition:

$$H_Y = S_Y + T_{X \to Y} + N_Y$$

Information

Information Storage

Information Transfer

**New Information** 



L Faes, A Porta, G Rossato, A Adami, D Tonon, A Corica, G Nollo: 'Investigating the mechanisms of cardiovascular and cerebrovascular regulation in orthostatic syncope through an information decomposition strategy', *Autonomic Neuroscience* 2013; 178:76-82.



### **PRACTICAL APPLICATIONS OF INFORMATION DYNAMICS**

### Network of brain-heart and brain-brain physiological interactions during sleep







### **Applications: BRAIN-BRAIN AND BRAIN-HEART INTERACTIONS**

- Protocol: full night polysomnography in 10 healthy subjects
- Linear estimator
- Network analysis Conditional information transfer + internal information Statistical significance assessed by F-test



Structured brain-heart and brain-brain network, with the EEG  $\beta$  wave acting as network hub The interaction network is sustained by the sleep stage transitions





### **Applications: BRAIN-HEART INTERACTIONS IN SLEEP APNEAS**

- ✓ 8 sleep apnoea-hypopnoea patients SAHS
- $\checkmark$  same patients after continuous positive airway pressure therapy CPAP
- ✓ 14 healthy controls CTRL





Redundancy is a feature of undisturbed sleep, lost in SAHS and recovered by treatment

**L Faes**, D Marinazzo, S Stramaglia, F Jurysta, A Porta, G Nollo, 'Predictability decomposition detects the impairment of brain-heart dynamical networks during sleep disorders and their recovery with treatment', *Phil. Trans. R. Soc. A* 2016; 374:20150177.



### **Applications: MULTISCALE BRAIN-HEART INTERACTIONS**



- Multiscale methods to study individual dynamics are well established [M Costa et al, Phys. Rev. Lett. 89, 2002]
- Multiscale computation of information transfer is non-trivial
- Exact computation of Information Dynamics for multivariate Gaussian processes

• Procedure for rescaling  
a (vector) time series  

$$Y_n = \{x_n, y_n\} \quad n = 1, ..., N$$

$$\bar{x}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} x_{n\tau-l} , \quad \bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} y_{n\tau-l}$$

$$\bar{x}$$

$$\bar{x}$$

$$\bar{x}$$

$$\bar{x}$$

$$\bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} x_{n\tau-l} , \quad \bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} y_{n\tau-l}$$

$$\bar{x}$$

$$\bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} x_{n\tau-l} , \quad n = \tau, ..., N$$

$$\bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} y_{n\tau-l}$$

$$\bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} y_{n\tau-l}$$

$$\bar{y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} y_{n\tau-l} , \quad n = 1, ..., N/\tau$$

#### • Multiscale representation of vector Autoregressive processes using state-space models



Information dynamics after rescaling can be obtained from the original VAR parameters and the scale factor au

L Faes, S Stramaglia, G Nollo, D Marinazzo 'Multiscale Granger causality', *Phys Rev E*; under revision, 2017.



L Faes, D Marinazzo, S Stramaglia, A Montalto, G Nollo, 'Multiscale information-theoretic analysis of coupled processes: theory and application to brainheart interactions', Brain Modes 2017; Bruxelles, Belgium, Dec 1-3 2016.

scale  $\tau$ 

10 11 12

60 s

0.015

0.01

0.005

5 s

12

60 s

0.1

0.05

0

5 s

scale  $\tau$ 





### **PRACTICAL APPLICATIONS OF INFORMATION DYNAMICS**

Physiological networks: EEG brain networks and muscular networks

• Study of networks formed by multichannel acquisitions of the same biomedical signal

### **Brain Networks**

Scalp multichannel EEG



### Muscular Networks

Whole-body multichannel EMG





33

### **Information Dynamics of Scalp EEG Networks**

- Protocol: scalp EEG in 21 healthy subjects during eyes open and eyes closed
- Nearest neighbor estimate of information transfer and conditional information transfer between all sensors



**L Faes**, D Marinazzo, G Nollo, A Porta 'An information-theoretic framework to map the spatio-temporal dynamics of the scalp electroencephalogram', *IEEE Trans. Biomed. Eng., special issue on Brain Connectivity*, 2016; 63(12):2488-2496





### **Information Dynamics of Muscle Networks**

- Protocol: multichannel EMG in 14 healthy subjects
- Conditions: standing and pointing to a target during normal altered stability





## **Network Physiology and Information Dynamics**



"The human organism is an integrated network where complex physiologic systems, each with its own regulatory mechanisms, continuously interact, and where failure of one system can trigger a breakdown of the entire network" [A. Bashan et al., Nature Communications 2012]

A new field, **Network Physiology**, is needed to probe the interactions among diverse physiologic systems







## **ITS Toolbox:** A Matlab toolbox for the practical computation of Information Dynamics



http://www.lucafaes.net/its.html

## faes.luca@gmail.com