



An information-theoretic framework to dissect multivariate and multiscale physiological interactions

Luca Faes

University of Palermo, Italy

**Multivariate
Interactions:**
*Information
Dynamics*

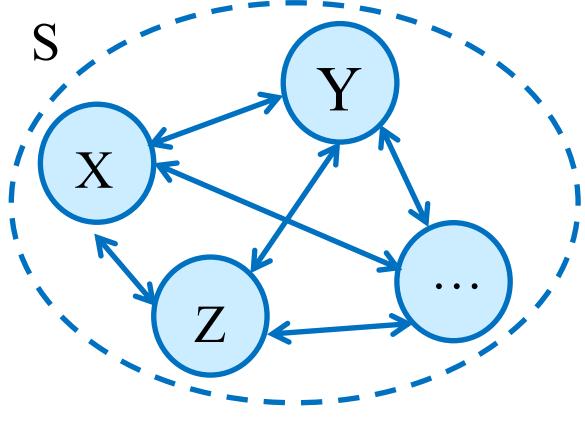
**Multiscale
Interactions:**
*State Space
Models*

Physiological Interactions:
Brain: epilepsy
Cardiovascular Physiology

INFORMATION DYNAMICS

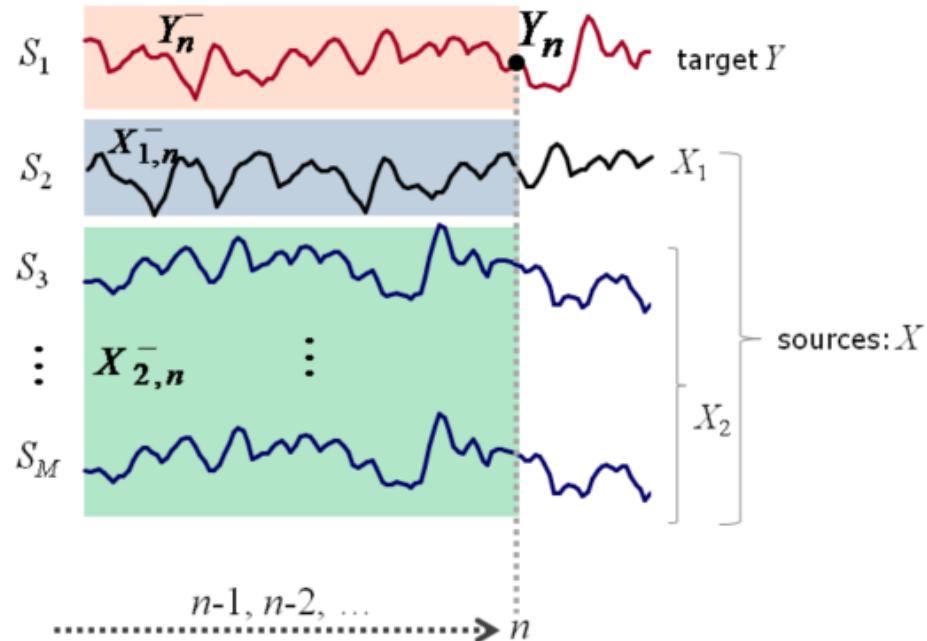
Physiological Networks

- Dynamic System $S = \{S_1, \dots, S_M\}$



Networks of Dynamical systems

- Dynamic Process S



- With reference to a target system Y :

$$X = \{X_1, \dots, X_{M-1}\} \rightarrow S = \{X_1, \dots, X_{M-1}, Y\} = \{X, Y\}$$

Investigation of Statistical dependencies:

✓ **SELF effects:**

$$Y_n^- \rightarrow Y_n$$

→ **Information storage**

✓ **CAUSAL effects:**

$$X_n^- \rightarrow Y_n$$

→ **Information transfer**

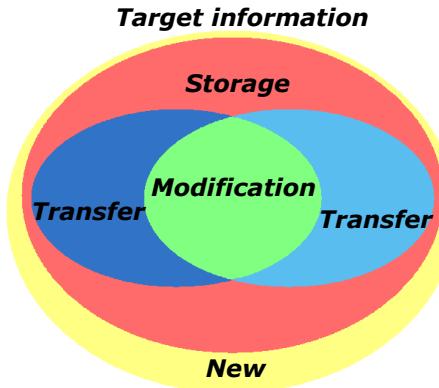
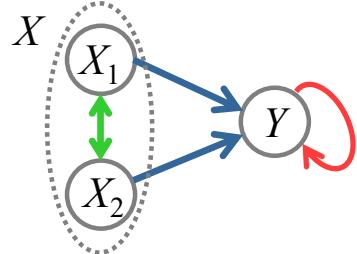
✓ **INTERACTION effects:** $(X_{1,n}^- \leftrightarrow X_{2,n}^-) \rightarrow Y_n$

→ **Information modification**

INFORMATION DECOMPOSITION

- Decomposition of the “information” contained in the target process

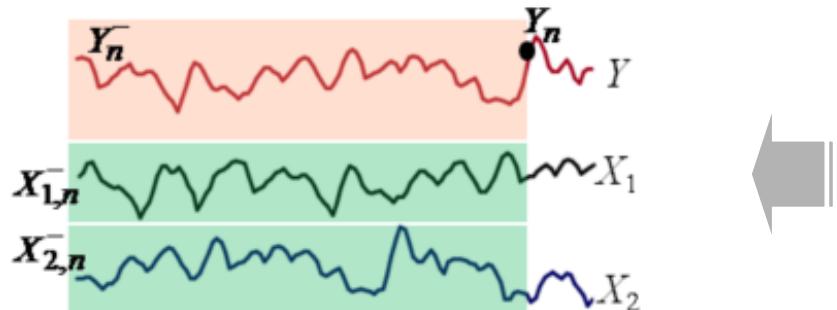
$$S = \{X, Y\} = \{X_1, X_2, Y\}$$



$$H_Y = S_Y + T_{X_1 \rightarrow Y} + T_{X_2 \rightarrow Y} + I_{X_1 X_2 \rightarrow Y} + N_Y$$

↓
Information
 ↓
Information Storage
 activity
 ↓
Information Transfer
 connectivity
 ↓
Information Modification
 interaction
 ↓
New Information
 non-predictable dynamics

- Computation: basic information theoretic measures



ENTROPY:

$$H(V) = -E[\log p(v)]$$

CONDITIONAL ENTROPY:

$$H(V|U) = H(V, U) - H(U)$$

MUTUAL INFORMATION:

$$I(V;U) = H(V) - H(V|U)$$

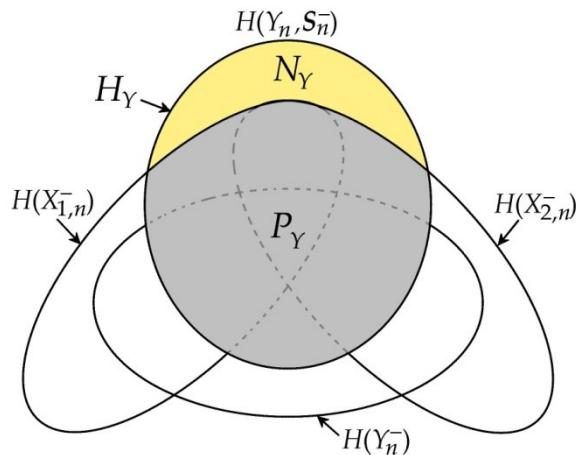
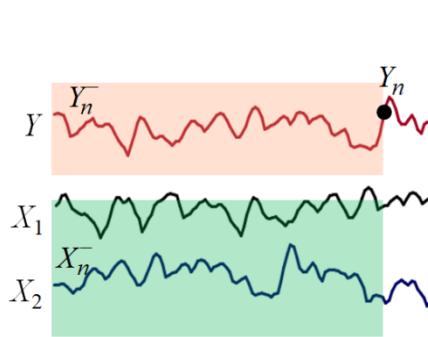
$$I(V;U|W) = H(V|W) - H(V|U,W)$$

INTERACTION INFORMATION:

$$I(V;U;W) = I(V;U) + I(V;W) - I(V;U,W)$$

$$I(V;U|W) = H(V|W) - H(V|U,W)$$

TARGET INFORMATION DECOMPOSITION



$$H(Y_n) = I(Y_n; \mathbf{X}_n^-, Y_n^-) + H(Y_n | \mathbf{X}_n^-, Y_n^-)$$

↓ H_Y ↓ P_Y ↓ N_Y
Information **Predictive Information** **New created Information**

- **Present Information** about $Y : H_Y = H(Y_n)$

Information contained in the present of the process Y



Uncertainty about the present state of the target

- **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



Predictability of the target given the past network states

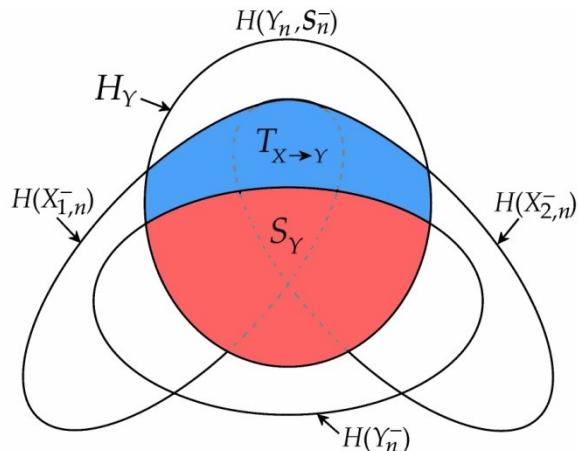
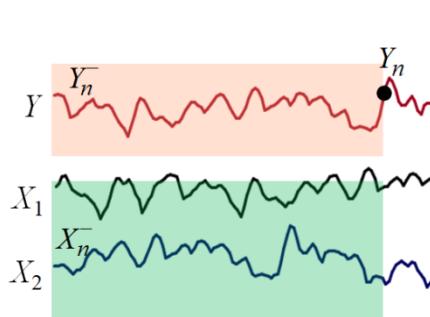
- **New information** about $Y : N_Y = H(Y_n | Y_n^-, X_n^-)$

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



Information generated in the target by the state transition

PREDICTIVE INFORMATION DECOMPOSITION



$$I(Y_n; \mathbf{X}_n^-, Y_n^-) = I(Y_n; Y_n^-) + I(Y_n; \mathbf{X}_n^- | Y_n^-)$$

\downarrow

$$P_Y$$

Predictive
Information

\downarrow

$$S_Y$$

Information
Storage

\downarrow

$$T_{X \rightarrow Y}$$

Information
Transfer

- **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



Predictability of the target given the network past states

- **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



Predictability of the target from its own past states

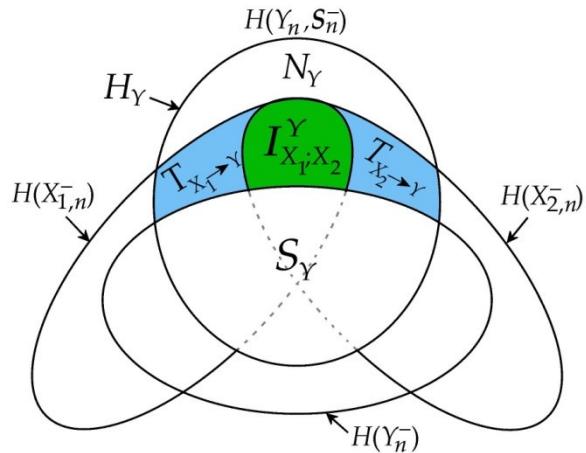
- **Information transfer** from X to Y : $T_{X \rightarrow Y} = I(Y_n; X_n^- | Y_n^-)$

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

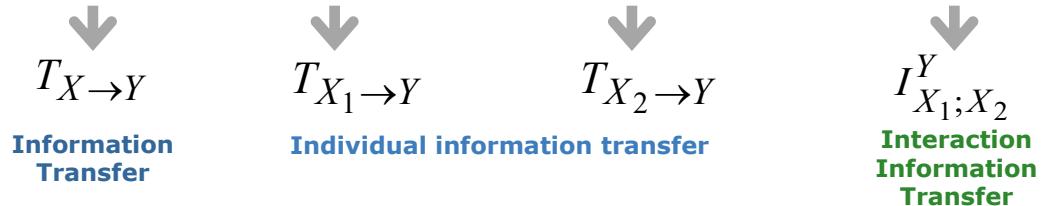


Causal interactions from all sources to the target

INFORMATION TRANSFER DECOMPOSITION



$$I(Y_n; \mathbf{X}_n^- | Y_n^-) = I(Y_n; X_{1,n}^- | Y_n^-) + I(Y_n; X_{2,n}^- | Y_n^-) + I(Y_n; X_{1,n}; X_{2,n}^- | Y_n^-)$$



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

Information contained in the past of X_1, X_2 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*

- **Individual information transfer:** $T_{X_1 \rightarrow Y} = I(Y_n; X_{1,n}^- | Y_n^-)$

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 of X_2)

→ *Causal interactions from one source to the target*

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

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

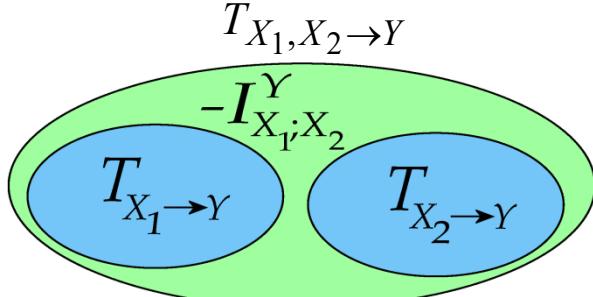
→ *Redundant or synergistic interactions contributing to transfer*

INFORMATION MODIFICATION: REDUNDANCY AND SYNERGY

- **Interpretation of Information Modification:** $I_{X_1;X_2}^Y = T_{X_1,X_2 \rightarrow Y} - (T_{X_1 \rightarrow Y} + T_{X_2 \rightarrow Y})$

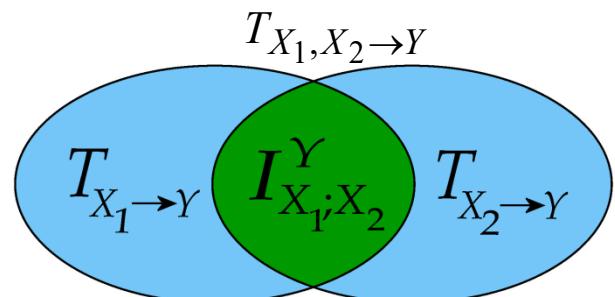
REDUNDANCY:

$$T_{X_1,X_2 \rightarrow Y} < T_{X_1 \rightarrow Y} + T_{X_2 \rightarrow Y} \rightarrow I_{X_1;X_2}^Y < 0$$



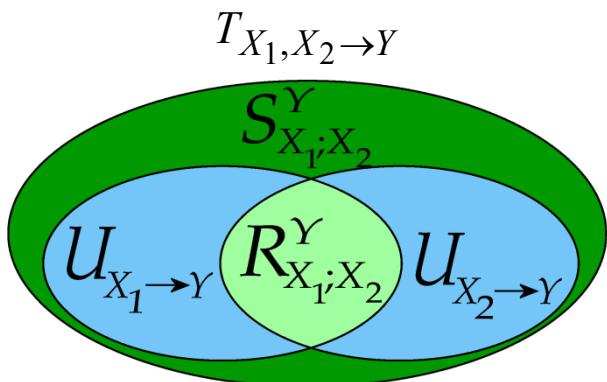
SYNERGY:

$$T_{X_1,X_2 \rightarrow Y} > T_{X_1 \rightarrow Y} + T_{X_2 \rightarrow Y} \rightarrow I_{X_1;X_2}^Y > 0$$



Interaction information can be positive or negative

- **PARTIAL INFORMATION DECOMPOSITION (PID)**



$$T_{X_1,X_2 \rightarrow Y} = U_{X_1 \rightarrow Y} + U_{X_2 \rightarrow Y} + R_{X_1;X_2}^Y + S_{X_1;X_2}^Y$$

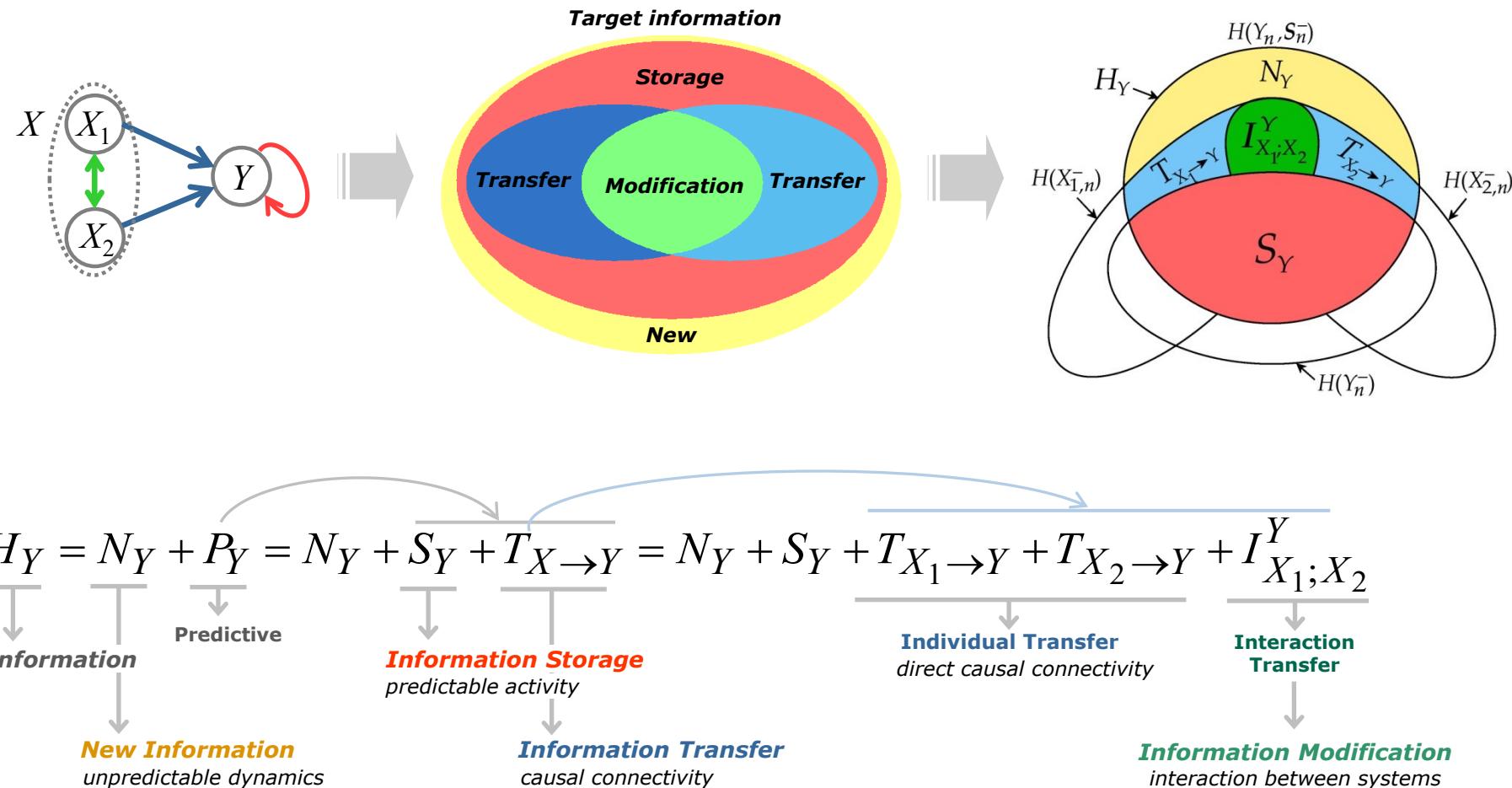
$$T_{X_1 \rightarrow Y} = U_{X_1 \rightarrow Y} + R_{X_1;X_2}^Y, \quad T_{X_2 \rightarrow Y} = U_{X_2 \rightarrow Y} + R_{X_1;X_2}^Y$$

- Minimum mutual information PID: $R_{X_1;X_2}^Y = \min\{T_{X_1 \rightarrow Y}, T_{X_2 \rightarrow Y}\}$
[A.B. Barrett, Phys. Rev. E 91, 2015]

Relation with interaction information:

$$I_{X_1;X_2}^Y = S_{X_1;X_2}^Y - R_{X_1;X_2}^Y$$

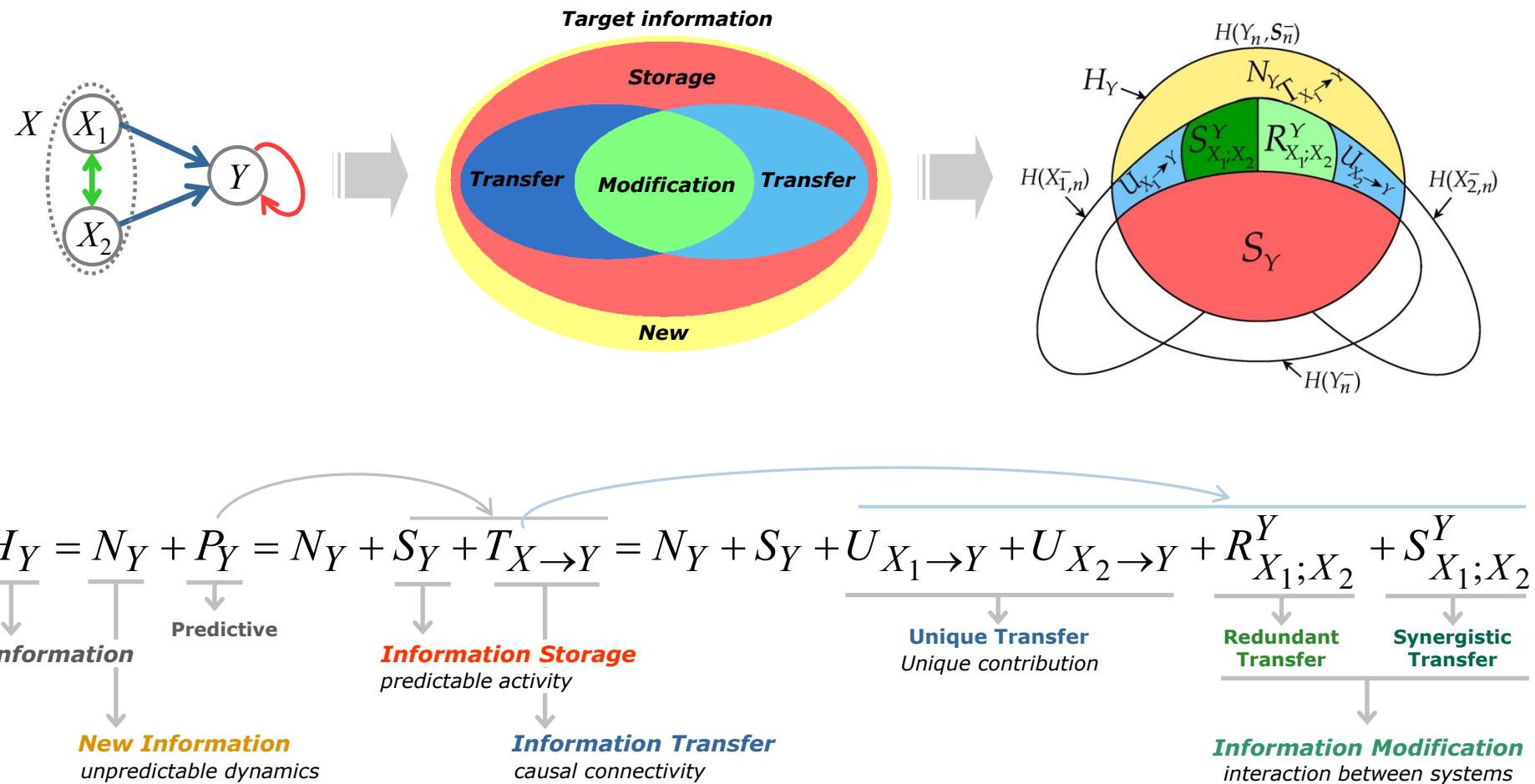
THE FRAMEWORK OF INFORMATION DYNAMICS



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

THE FRAMEWORK OF INFORMATION DYNAMICS

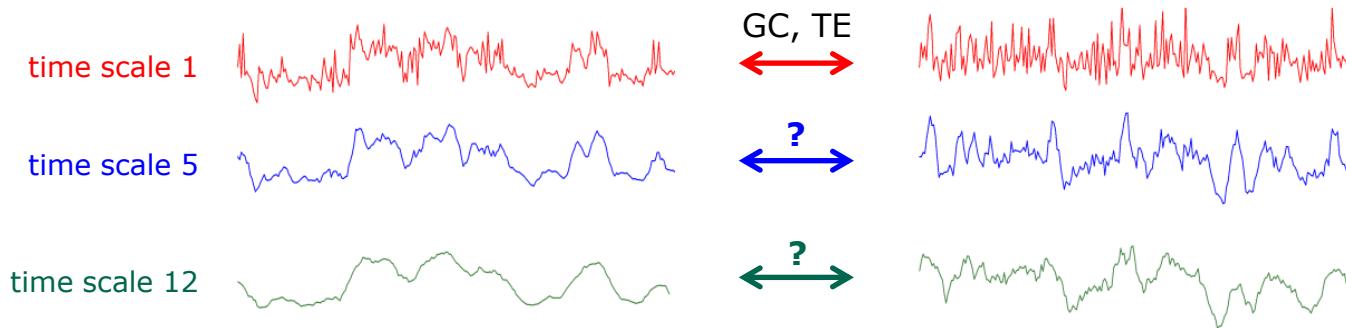


L Faes, A Porta, G Nollo, 'Information decomposition in bivariate systems: theory and application to cardiorespiratory dynamics', *Entropy, special issue on "Entropy and Cardiac Physics"*, 2015, 17:277-303.

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

MULTISCALE INFORMATION DYNAMICS

- Many physical processes exhibit dynamics spanning multiple temporal scales
- 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



- We propose a **formal extension of information decomposition to multiscale analysis** of jointly stationary multivariate linear processes (VAR)
- Formulations based on VARMA representation and state-space (SS) modeling, leading to **exact computation** of information dynamics from VAR parameters

L Faes, S Stramaglia, G Nollo, D Marinazzo 'Multiscale Granger causality', *Phys Rev E* **96**, 042150, 2017

L Faes, D Marinazzo , S Stramaglia 'Multiscale Information decomposition: exact computation for multivariate gaussian processes', *Entropy* **19**, 408, 2017

MULTISCALE INFORMATION DYNAMICS

MULTISCALE ANALYSIS OF TIME SERIES: CHANGE OF TIME SCALE

- Traditional procedure for rescaling

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

↓ RESCALING (scale factor τ):

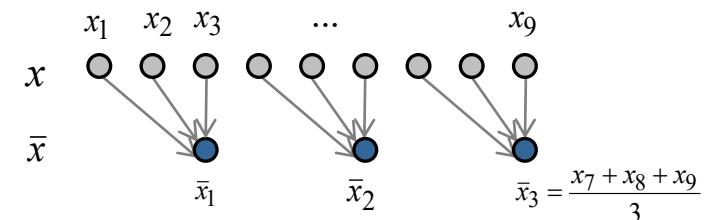
$$\bar{Y}_n = \{\bar{x}_n, \bar{y}_n\}, \quad n = 1, \dots, N/\tau$$

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

[M Costa et al, Phys. Rev. Lett. 89, 2002]

Example:

$$N = 9, \tau = 3$$



- Rescaling can be seen as a two-step procedure

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

↓ 1) AVERAGING (lowpass filtering)

$$\tilde{Y}_n = \frac{1}{\tau} \sum_{l=0}^{\tau-1} Y_{n-l}, \quad n = \tau, \dots, N$$

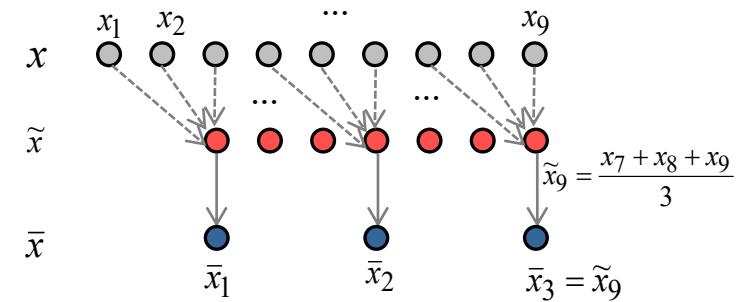
↓ 2) DOWNSAMPLING

$$\bar{Y}_n = \tilde{Y}_{n\tau}, \quad n = 1, \dots, N/\tau$$

[J. Valencia et al, IEEE Trans. Biomed Eng. 56, 2009]

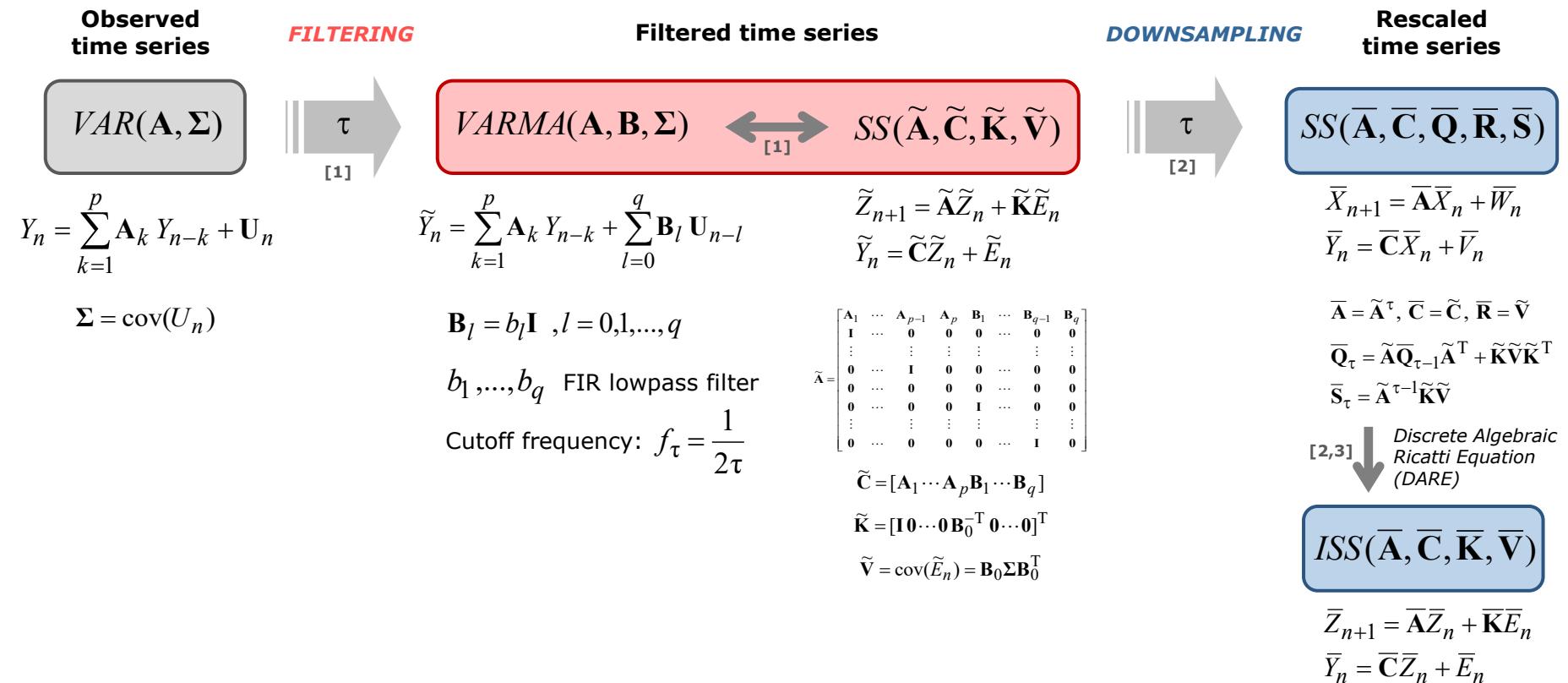
Example:

$$N = 9, \tau = 3$$



MULTISCALE INFORMATION DYNAMICS

MULTISCALE REPRESENTATION OF LINEAR PROCESSES USING STATE SPACE MODELS



¹[Aoki & Havenner, Econ. Rev. 10, 1991]

²[Solo, ArXiv 1501.04663, 2015]

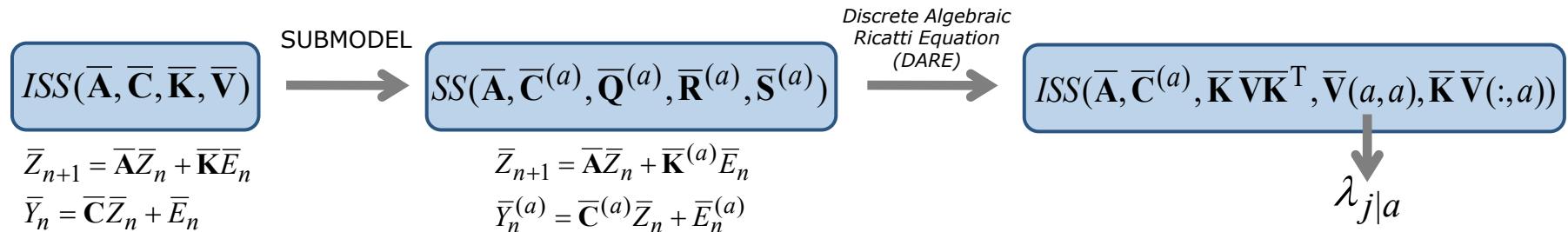
³[Barnett & Seth, Phys. Rev. E 91, 2015]

The State Space model defining the multivariate linear process after rescaling can be obtained from the original VAR parameters and the scale factor τ

MULTISCALE INFORMATION DYNAMICS

**EXACT COMPUTATION OF INFORMATION TRANSFER FOR LINEAR PROCESSES
BASED ON STATE SPACE MODELS**

- Computation of the partial variance of the target process y_j given a subset of processes Y_a



- Submodels containing only the target, the target and one source, and the target and both sources

$$\cdot \quad a=j \quad \Rightarrow \quad \lambda_{j|j}$$

$$\begin{array}{l} \cdot \quad a=j,i \\ \cdot \quad a=j,k \end{array} \quad \Rightarrow \quad \lambda_{j|ji}, \quad \lambda_{j|jk}$$

$$\cdot \quad a=j,i,k \quad \Rightarrow \quad \lambda_{j|jik}$$

$$S_j = \frac{1}{2} \ln \frac{\lambda_j}{\lambda_{j|j}}$$

$$T_{i \rightarrow j} = \frac{1}{2} \ln \frac{\lambda_{j|j}}{\lambda_{j|ji}}, \quad T_{k \rightarrow j} = \frac{1}{2} \ln \frac{\lambda_{j|j}}{\lambda_{j|jk}}$$

$$T_{ik \rightarrow j} = \frac{1}{2} \ln \frac{\lambda_{j|j}}{\lambda_{j|jik}}$$

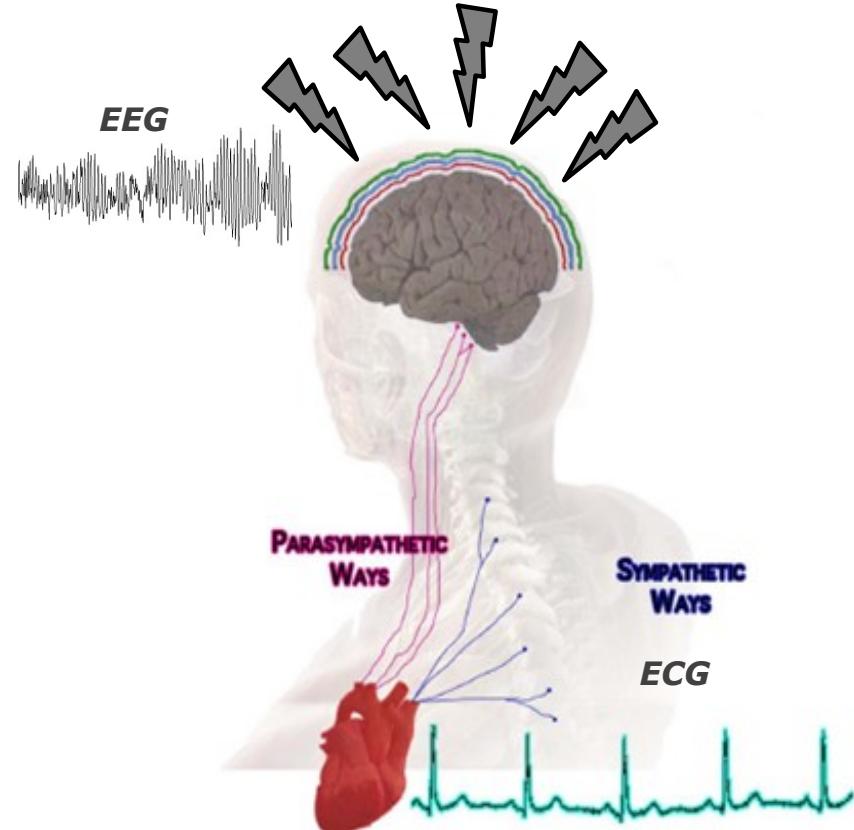
Information transfer in multivariate linear processes can be computed at any scale τ from the original VAR parameters

APPLICATIONS (1)

STUDY OF BRAIN AND PHYSIOLOGICAL NETWORKS IN EPILEPSY

- The cortical activity changes drastically during an epileptic seizure
- The study of brain networks during epilepsy may help in seizure prediction or detection
- Seizures influences also the peripheral ANS response
- Recent works studied the correlation between the epileptic neural network and the autonomic nervous system

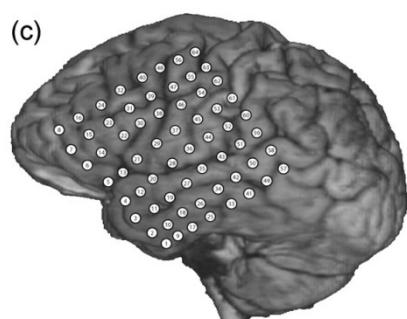
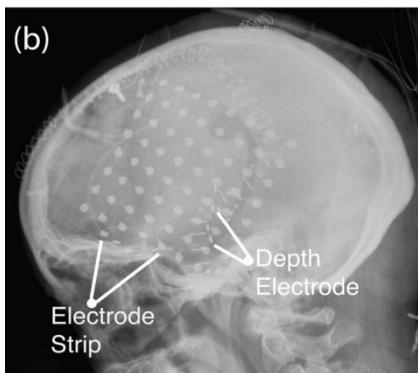
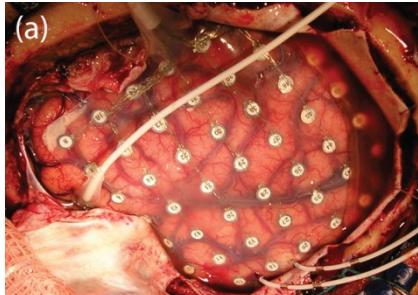
[K. Schiecke et al., IEEE Trans. Biomed Eng. 63, 2016]



*Application of multiscale partial information decomposition to
brain networks and networks of brain-heart interactions during epilepsy*

APPLICATIONS (1a)

MULTISCALE INFORMATION TRANSFER IN THE EPILEPTIC BRAIN



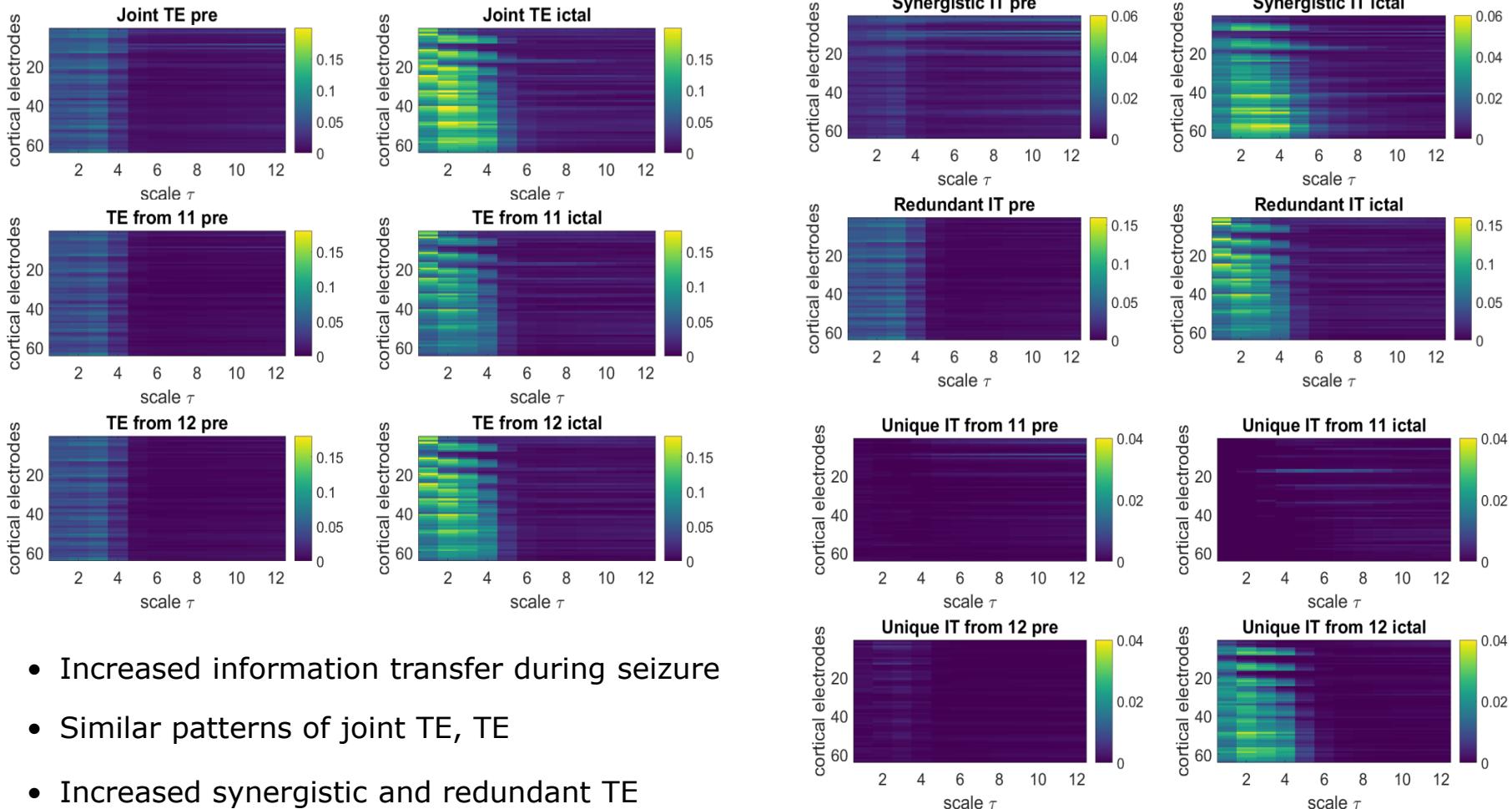
- **Intracranial EEG from a patient with intractable epilepsy (focal seizures, 8 episodes)**
[MA Kramer et al., Epilepsy Res. 79, 2008]
- **Grid of 8 x 8 cortical electrodes + 2 deep electrodes (left hippocampal region)**
- **10-sec time windows before seizure onset (pre-ictal) and during the seizure (ictal)**
- **Multiscale Partial Information Decomposition:**
 - model order: Bayesian Information Criterion (average $p=14$)
 - lowpass FIR filter with $q=12$ coeffs
 - Scale $\tau=1,\dots,12$ (cutoff freq. $f_\tau=200$ Hz, ..., 16.6 Hz)
 - Computation of PID transfer functions **from the two deep electrodes to each cortical electrode**

$$T_{ik \rightarrow j} = U_{i \rightarrow j} + U_{k \rightarrow j} + R_{ik \rightarrow j} + S_{ik \rightarrow j}$$

$$i = \text{deep } 11, k = \text{deep } 12; j = 1, \dots, 64$$

APPLICATIONS (1a)

MULTISCALE INFORMATION TRANSFER IN THE EPILEPTIC BRAIN



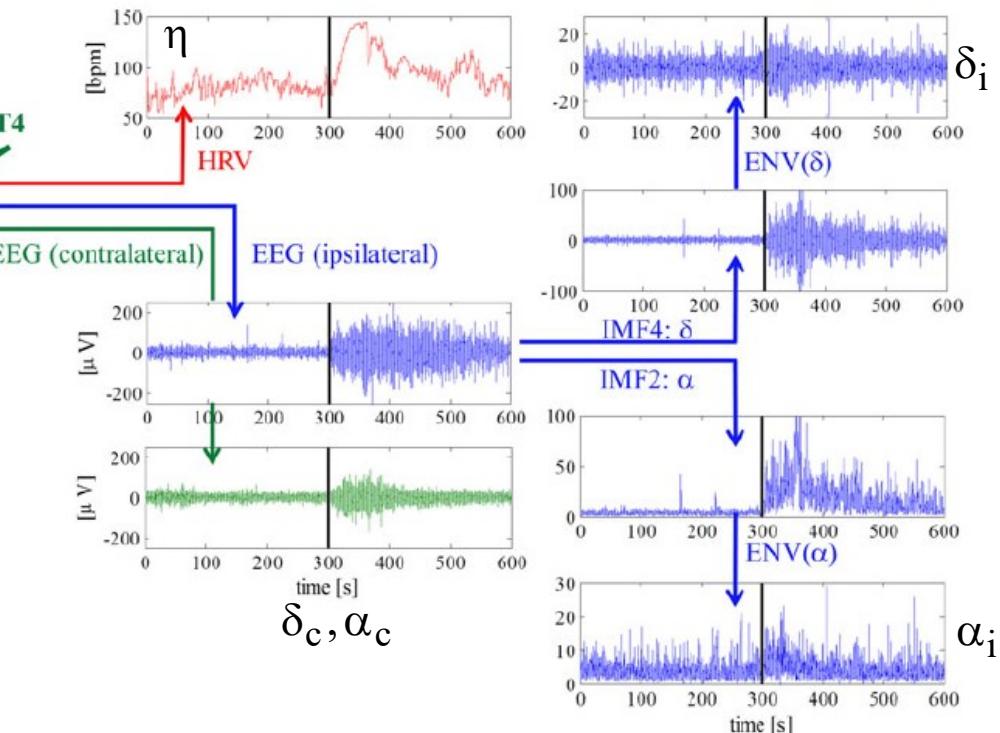
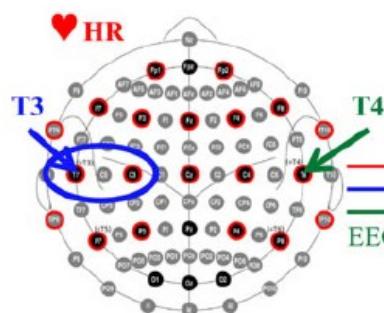
- Increased information transfer during seizure
- Similar patterns of joint TE, TE
- Increased synergistic and redundant TE
- Increased unique TE **only from deep electrode 12** → **useful for seizure localization**

APPLICATIONS (1b)

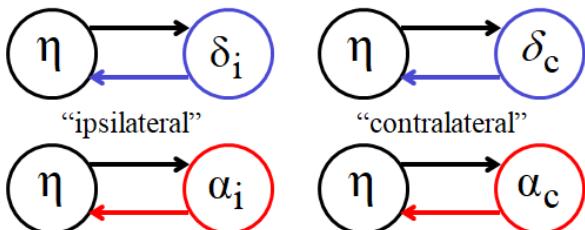
PARTIAL INFORMATION DECOMPOSITION IN EPILEPTIC BRAIN-HEART INTERACTIONS

- PROTOCOL:

- ✓ 18 children with **temporal lobe epilepsy**
- ✓ **Pre-ictal** (5 min)
- ✓ **Ictal** (~ 1.5 min)
- ✓ **Post-ictal** (~ 4.5 min)
- ✓ ECG → HRV
- ✓ EEG:
 - Selection of ipsilateral and contralateral temporal lobe electrodes
 - Extraction of δ and α EEG envelopes



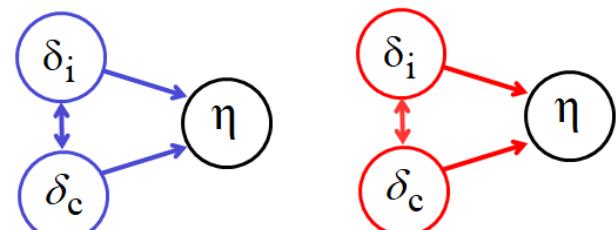
- Bivariate Information Transfer



$$T_{\delta \rightarrow \eta} = I(\eta_n; \delta_n^- | \eta_n^-), \quad T_{\eta \rightarrow \delta} = I(\delta_n; \eta_n^- | \delta_n^-)$$

$$T_{\alpha \rightarrow \eta} = I(\eta_n; \alpha_n^- | \eta_n^-), \quad T_{\eta \rightarrow \alpha} = I(\alpha_n; \eta_n^- | \alpha_n^-)$$

- Partial Information Decomposition



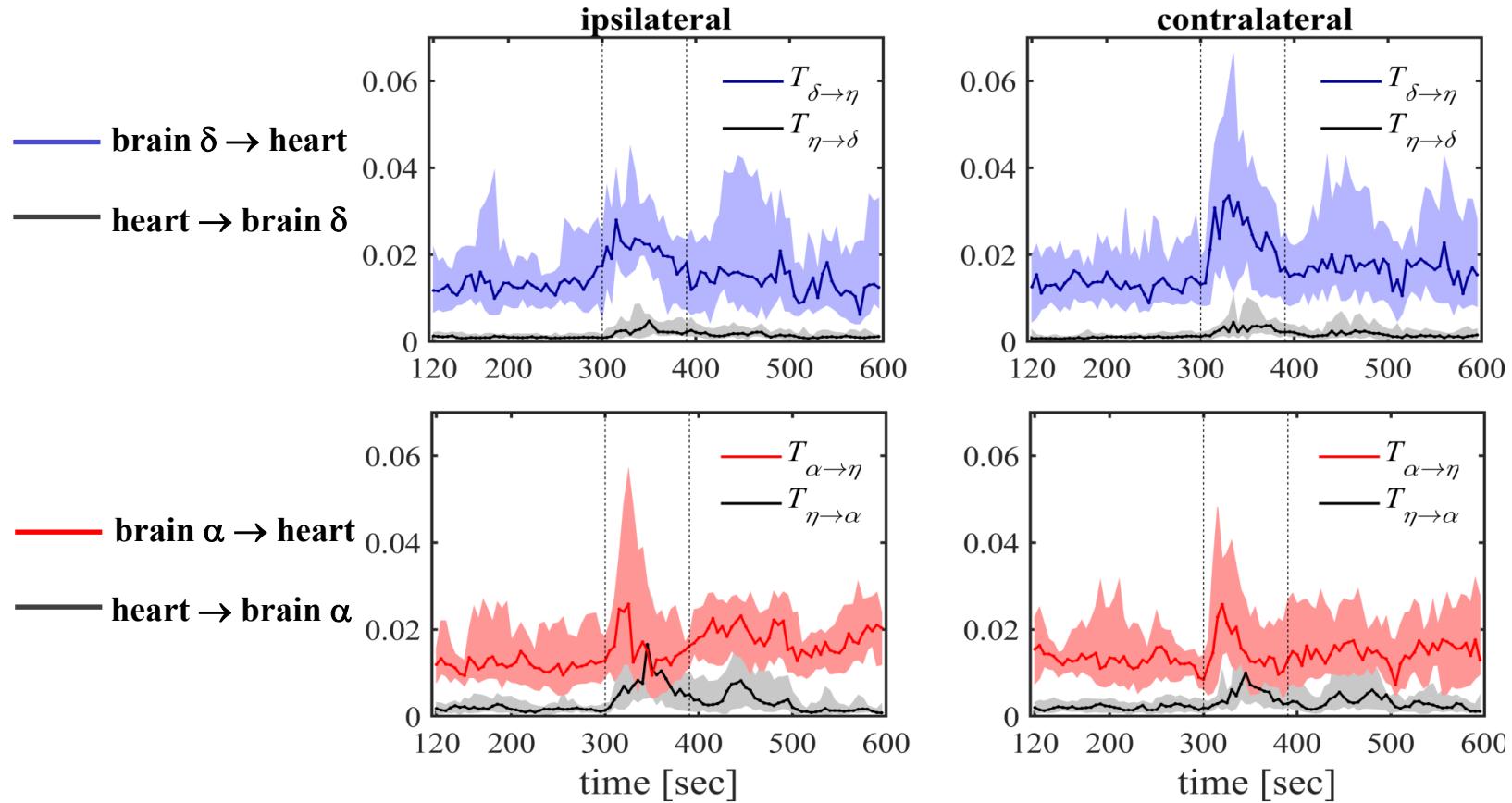
$$T_{\delta_i, \delta_c \rightarrow \eta} = U_{\delta_i \rightarrow \eta} + U_{\delta_c \rightarrow \eta} + R_{\delta_i; \delta_c}^{\eta} + S_{\delta_i; \delta_c}^{\eta}$$

$$T_{\alpha_i, \alpha_c \rightarrow \eta} = U_{\alpha_i \rightarrow \eta} + U_{\alpha_c \rightarrow \eta} + R_{\alpha_i; \alpha_c}^{\eta} + S_{\alpha_i; \alpha_c}^{\eta}$$

APPLICATIONS (1b)

Brain-heart interactions in epilepsy: RESULTS

- **Brain-Heart Information Transfer**

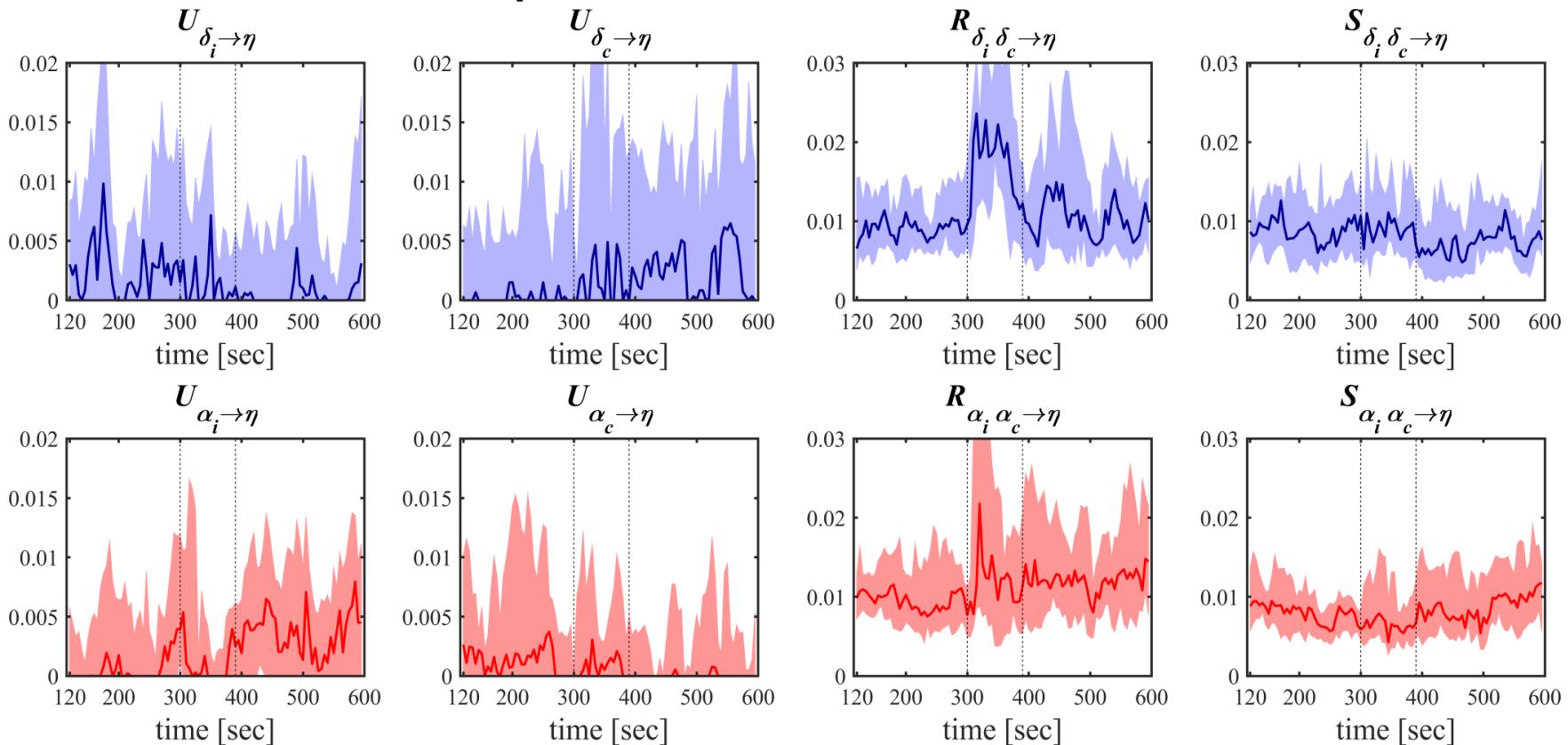


- *The information transfer is markedly higher along the brain \rightarrow heart direction*
- *No evident differences are observed between δ and α waves, pre-ictal and post-ictal phases, or contralateral and ipsilateral sites*

APPLICATIONS (1b)

Brain-heart interactions in epilepsy: RESULTS

- Partial information decomposition of brain→heart information transfer

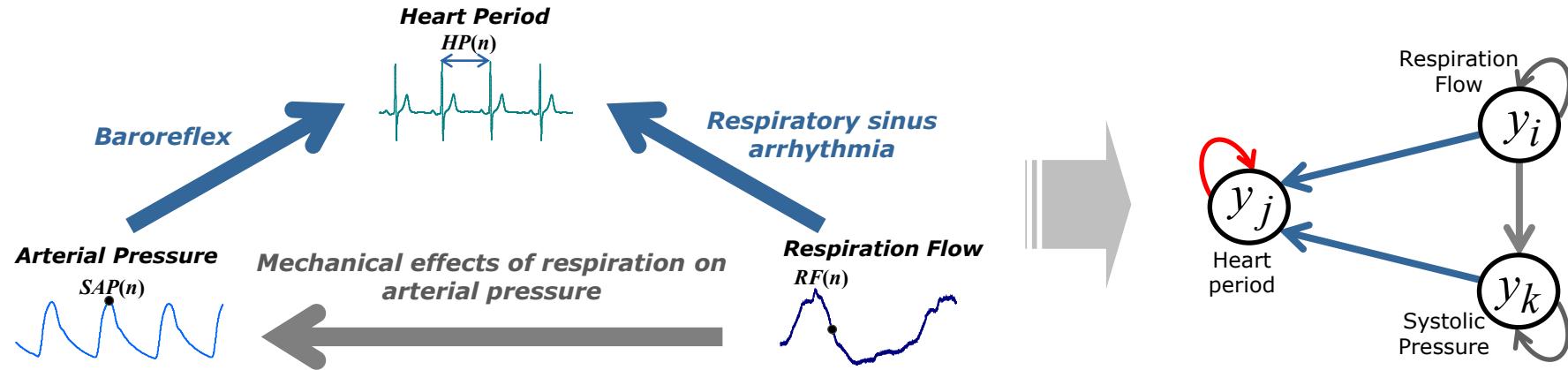


- *The unique information transfer $\delta \rightarrow \eta$ is mostly ipsilateral in the pre-ictal phase and contralateral during the seizure and in the post-ictal phase*
- *These findings document the importance of PID, which removes from the information transfer the redundancy between the EEG activity of the two hemispheres*

APPLICATIONS (2)

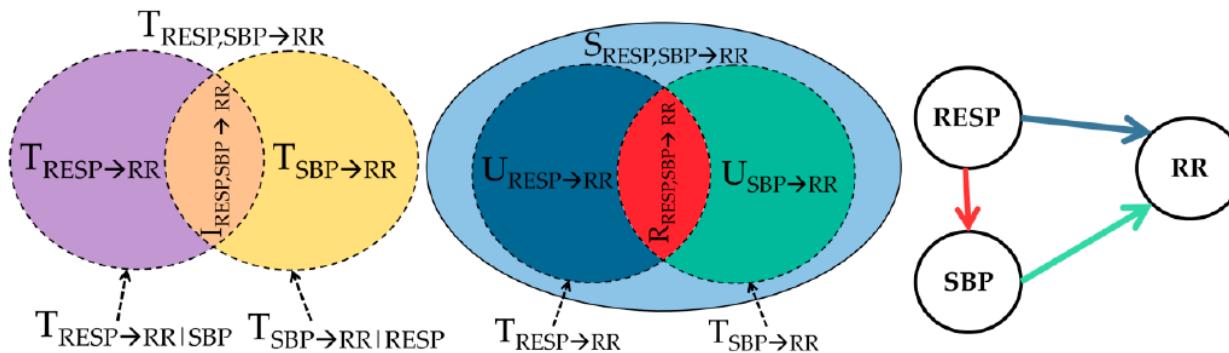
Applications: CARDIOVASCULAR and CARDIRESPIRATORY INTERACTIONS

- Cardiovascular regulatory physiology



- Multiscale Partial Information Decomposition:

Sympathetic and parasympathetic systems act at different time scales



$$T_{SAP, RESP \rightarrow HP} = U_{SAP \rightarrow HP} + U_{RESP \rightarrow HP} + R_{SAP, RESP \rightarrow HP} + S_{SAP, RESP \rightarrow HP}$$

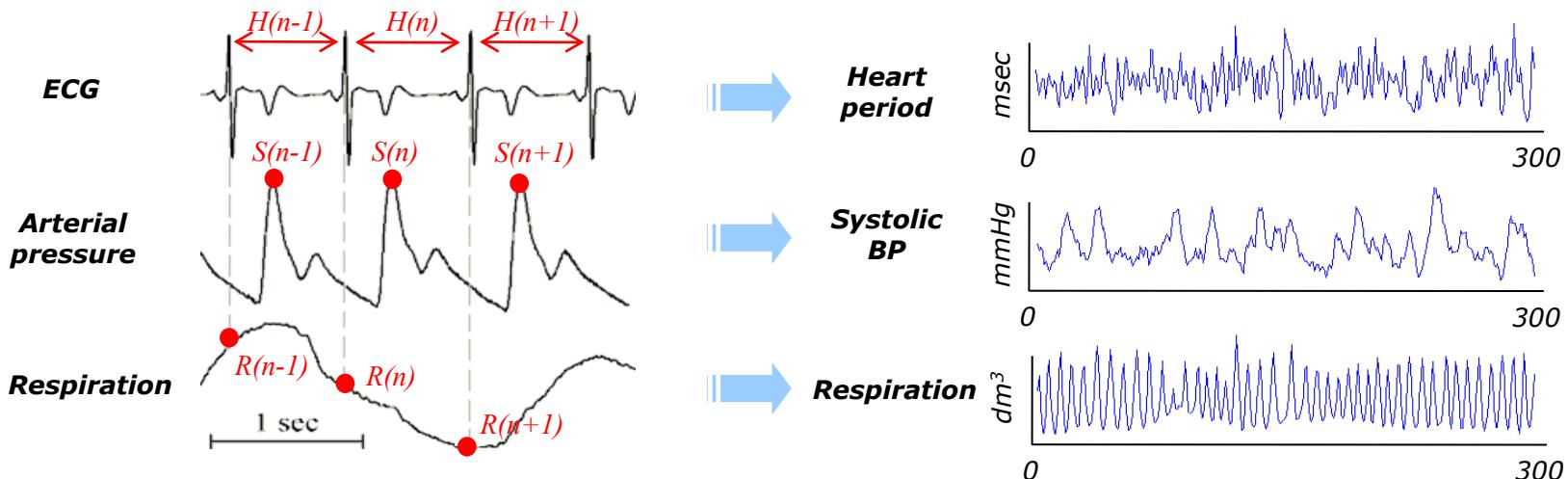
APPLICATIONS (2a)

MULTISCALE CARDIOVASCULAR INFORMATION DECOMPOSITION

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



- Signals and time series:



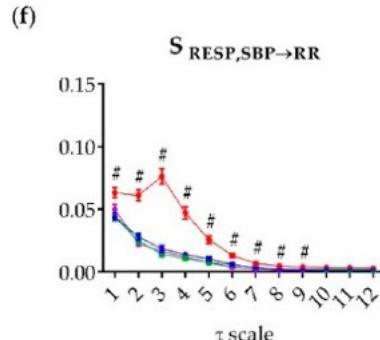
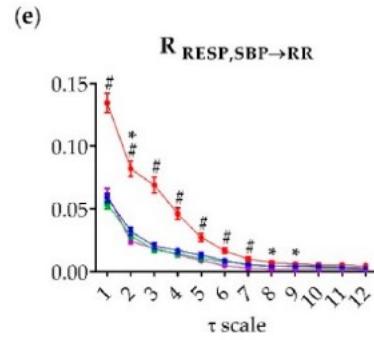
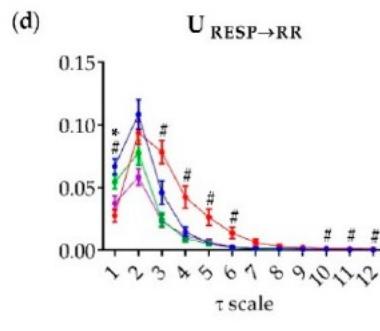
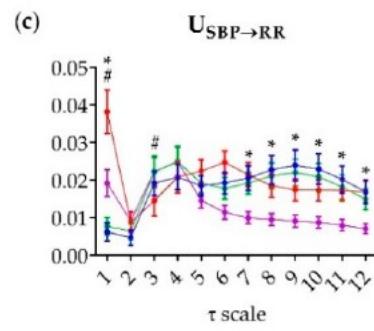
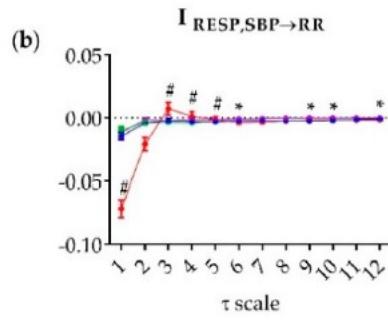
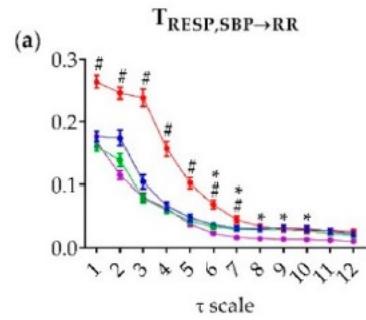
- Multiscale information decomposition:

$$T_{\text{SAP,RESP} \rightarrow \text{HP}} = U_{\text{SAP} \rightarrow \text{HP}} + U_{\text{RESP} \rightarrow \text{HP}} + R_{\text{SAP,RESP} \rightarrow \text{HP}} + S_{\text{SAP,RESP} \rightarrow \text{HP}}$$

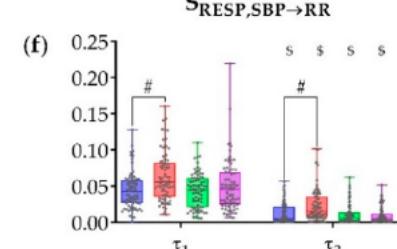
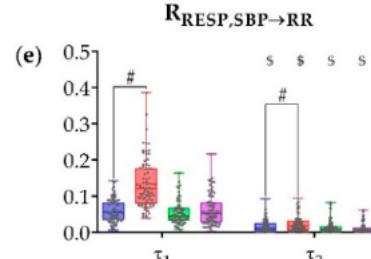
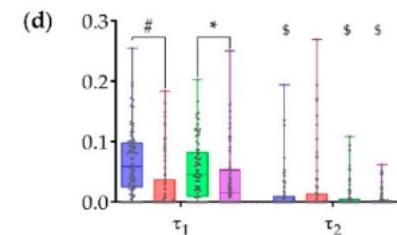
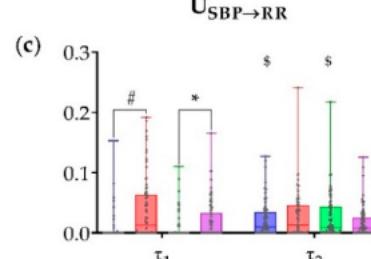
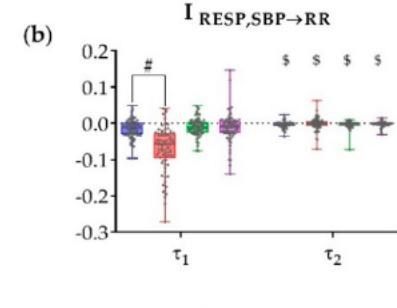
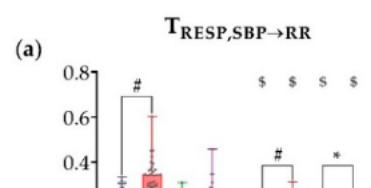
- model order: Bayesian Information Criterion (average $p=14$)
- lowpass FIR filter with $q=12$ coeffs
- Scale $\tau=1,\dots,12$

APPLICATIONS (2a)

MULTISCALE CARDIOVASCULAR INFORMATION DECOMPOSITION



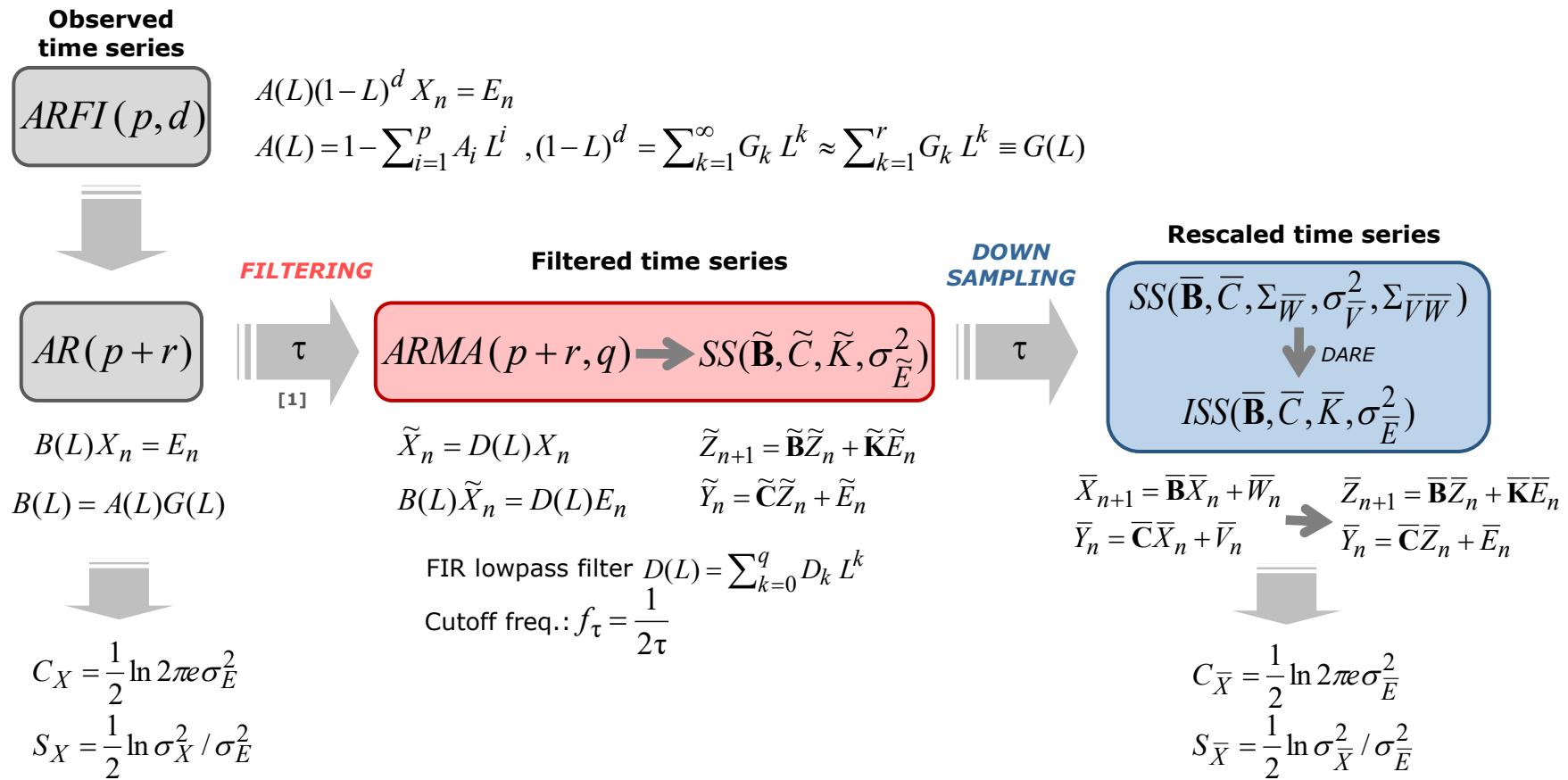
• Supine rest • HUT • Supine recovery • MA



■ Supine rest ■ HUT ■ Supine recovery ■ MA

APPLICATIONS (2b)

- **Limits of linear multiscale information dynamics**
 - The linear representation is restricted to AR processes
 - The model cannot account for long range correlations
- **Linear multiscale analysis based on fractionally integrated AR models**



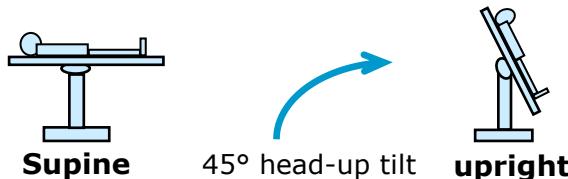
APPLICATIONS (2b)

Multiscale information storage in cardiovascular physiology

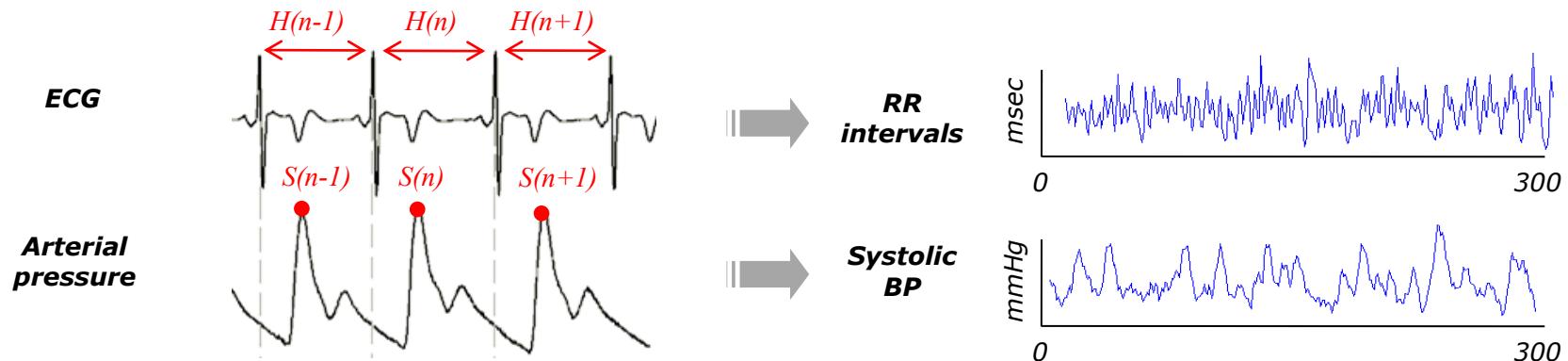
❖ Experimental Protocol

□ HEAD-UP TILT

61 Healthy subjects
(37 females, 17.5 ± 2.4 years)



❖ Construction of beat-to-beat variability series

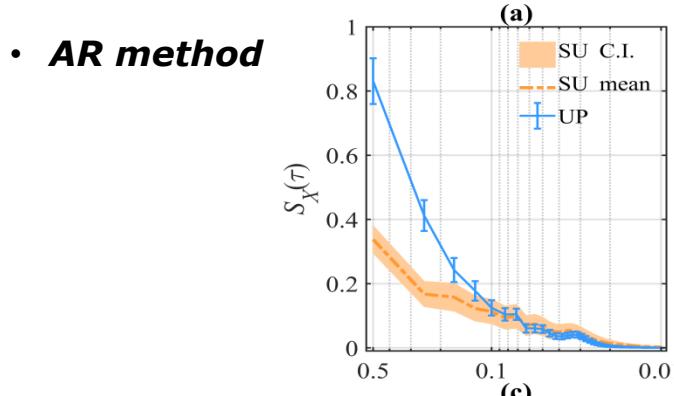


❖ Data analysis

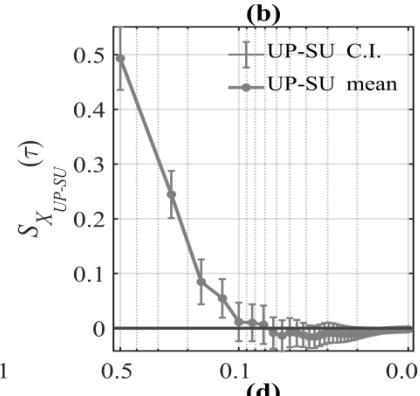
- Stationary windows of $N=300$ beats
- ARFI identification: computation of d with Whittle semiparametric estimator
- computation of $A(L)$ with least squares, order p with BIC criterion

APPLICATIONS (2b)

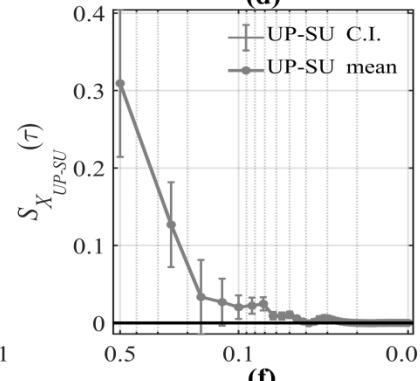
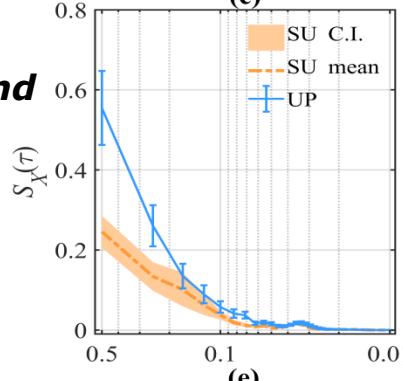
Multiscale information storage



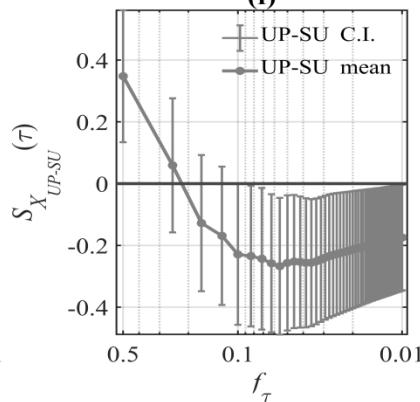
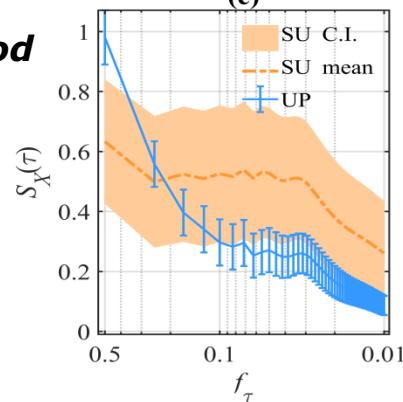
Difference upright – supine



- AR method after detrend**



- ARFI method**



- from supine to upright:
↑ S_x at short scales

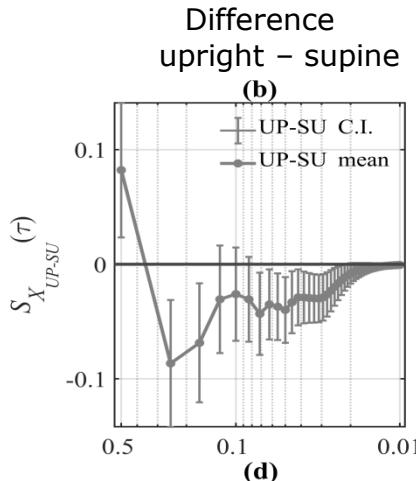
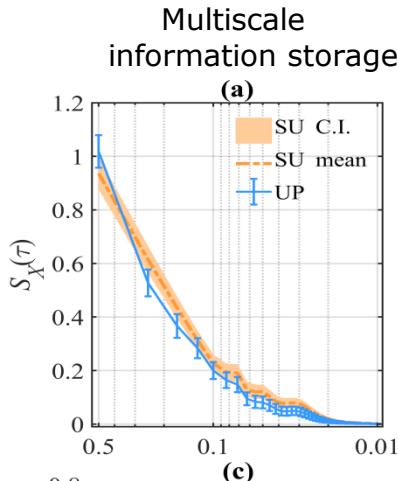
- from supine to upright:
↑ S_x at short scales
Increase of regularity of heart rate variability with tilt

- from supine to upright:
↑ S_x at short scales
↓ S_x at long scales

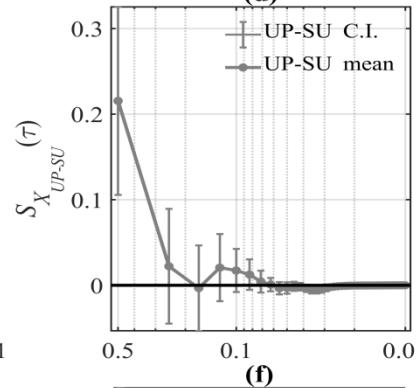
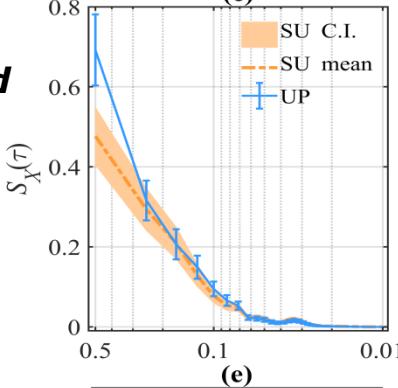
Higher complexity of heart rate variability with tilt, related to long-range correlations

APPLICATIONS (2b)

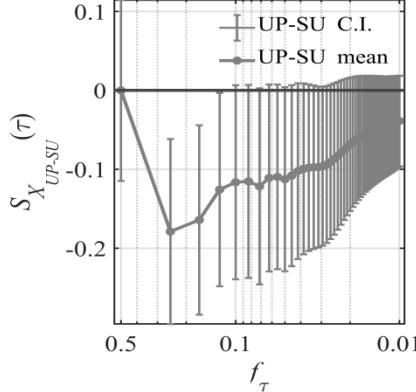
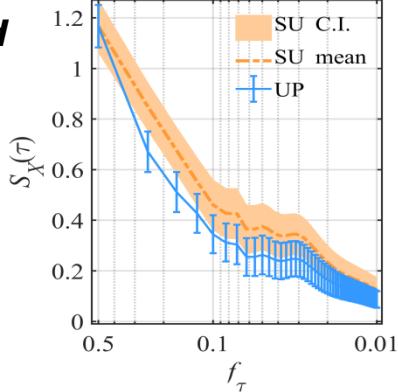
- **AR method**



- **AR method after detrend**



- **ARFI method**



- from supine to upright:

↑ S_x at scale 1

↓ S_x at scales >1

- from supine to upright:

↑ S_x at scale 1

↔ S_x at scales >1

Lower complexity of SAP associated with short term dynamics (respiratory?)

- from supine to upright:

↓ S_x at scales >1

↔ S_x at scale 1

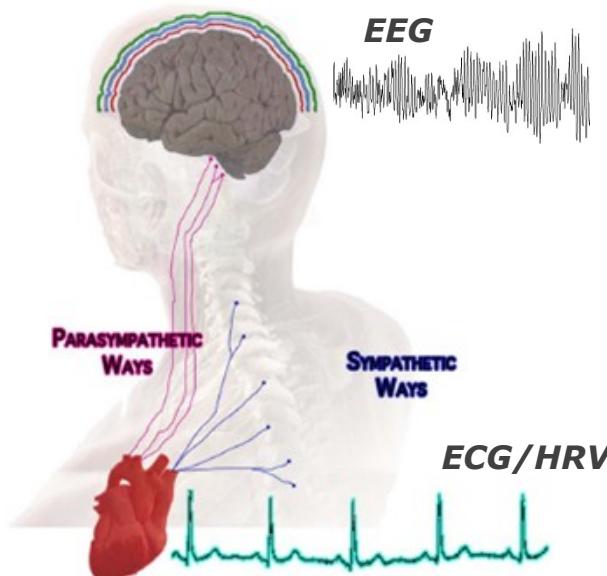
Higher complexity of SAP associated with slow oscillations (sympathetic?)

CONCLUSION

“An information-theoretic framework to dissect multivariate and multiscale physiological interactions”

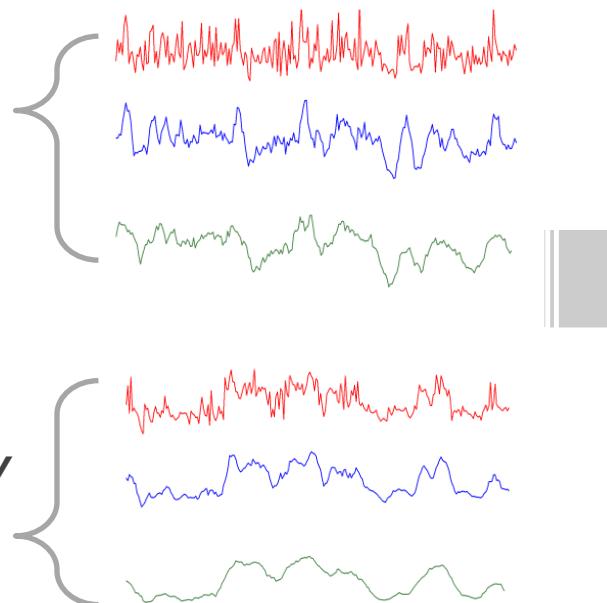
Multivariate analysis

different organ systems



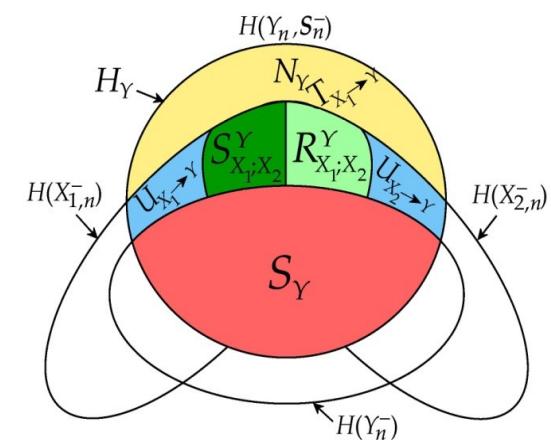
Multiscale analysis

different biological clocks



Information dynamics

Linear regression models



The ability to **handle multivariate and multiscale dynamics** and the **general applicability** should make the proposed tool useful in many contexts within the field of Network Physiology