Measuring Interactions from Time Series

Information-theory-based techniques

measuring interactions **information-theory**

basic idea: interaction ⇔ information flow

- the "stronger" the information flow, the stronger the interaction
- information flow from one system to another indexes directionality

characterize information with Shannon entropy $H = -\sum_i p_i \log p_i$

p is the (normalized) probability for an event / state / amplitude /… to occur

estimate probability with $p_i = \lim_{N \to \infty} {}^{N_i}/_{N}$
where N is the total number of events / states / amplitudes /…

measuring interactions **information-theory**

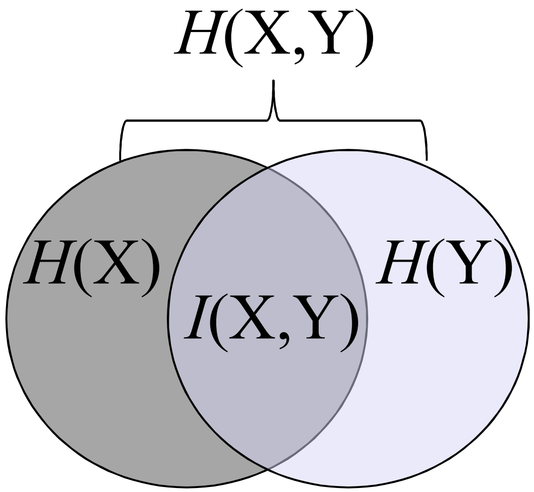
strength of interaction: **mutual information**

given systems X and Y, the mutual information is defined as:

$$I(\mathrm{X},\mathrm{Y}) = H(\mathrm{X}) + H(\mathrm{Y}) - H(\mathrm{X},\mathrm{Y})$$

information generated by system X is characterized by the Shannon entropy:

$$H(\mathrm{X}) = -\sum_i^N p_{\mathrm{X}}(i)\log p_{\mathrm{X}}(i)$$



joint information is characterized by the Shannon entropy:

$$H(\mathrm{X},\mathrm{Y}) = -\sum_{i,j}^N p_{\mathrm{X},\mathrm{Y}}(i,j)\log p_{\mathrm{X},\mathrm{Y}}(i,j)$$

measuring interactions **information-theory**

strength of interaction: **relative entropy**

relative entropy (also known as Kullback-Leibler divergence):

$$H_{\mathrm{rel}}(\mathrm{X}|\mathrm{Y}) = -\sum_i^N p_{\mathrm{X}}(i) \log \frac{p_{\mathrm{X}}(i)}{p_{\mathrm{Y}}(i)}$$

- characterizes the similarity between the probability distributions.
- relative entropy is asymmetric: $H_{\mathrm{rel}}(\mathrm{X}|\mathrm{Y}) \neq H_{\mathrm{rel}}(\mathrm{Y}|\mathrm{X})$
- in general H_{rel} is positive, and zero for identical systems

→ alternative definition of mutual information:

$$I_{\mathrm{rel}}(\mathrm{X}|\mathrm{Y}) = -\sum_{i,j}^N p_{\mathrm{X},\mathrm{Y}}(i,j) \log \frac{p_{\mathrm{X},\mathrm{Y}}(i,j)}{p_{\mathrm{X}}(i)p_{\mathrm{Y}}(j)}$$

characterizes *relative* difference between respective probability density distributions and the joint distribution density

S. Kullback, R.A. Leibler. On information and sufficiency. Ann. Math. Stat. ***22****, 79, 1951.*

measuring interactions **information-theory**

strength of interaction: **mutual information**

properties of mutual information:

- symmetric: $I\,(X, Y) = I\,(Y, X)$
- $I\,(X, Y) = 0$ for independent (non-interacting) systems
- $I\,(X, Y) = \max$ for identical (fully synchronized) systems
- $I\,(X, Y)$ increases monotonically with increasing coupling strength → data-driven estimator for strength of interaction

disadvantages:

- only considers (single/joint) probability density distributions
- no information about dynamics
- can not explicitly distinguish between information exchange and joint information (e.g. due to common input or joint past)

measuring interactions **information-theory**

strength of interaction: **mutual information**

extensions:

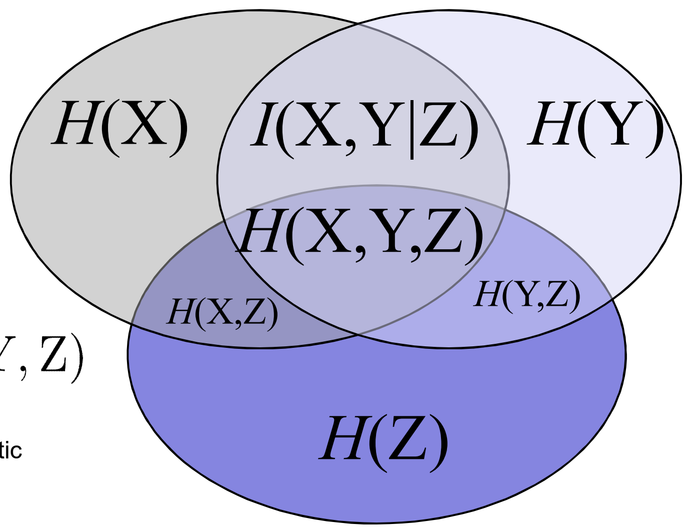
time-delayed mutual information

(Kaneko, Physica D 23, 436, 1986)

partial (or conditional) mutual information

(S. Frenzel & B. Pompe, PRL 99, 204101, 2007)

- part of mutual information of two random quantities that is not contained in a third one
- similar to partial correlation
- can also detect directionality*



$$I(\mathrm{X},\mathrm{Y}|\mathrm{Z}) = H(\mathrm{X},\mathrm{Z}) + H(\mathrm{Y},\mathrm{Z}) - H(\mathrm{Z}) - H(\mathrm{X},\mathrm{Y},\mathrm{Z})$$

* K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka & J. Bhattacharya. Causality detection based on information-theoretic approaches in time series analysis. Physics Reports, 441, 1-46, 2007

measuring interactions **information-theory**

direction of interaction: **transfer entropy**

aim: characterize flow of information between systems X and Y
idea: replace (static) probability density distributions by transition probability densities (cf. Entropies)

given: time series ***v***: $v_1, v_2, \ldots, v_N$ of some observable ***x*** and
time series ***w***: $w_1, w_2, \ldots, w_N$ of some observable ***y***

1) incorporate time-dependence by relating previous samples v_i and w_i to predict the next value v_{i+1} (cf. N. Wiener),

2) consider generalized Markov condition (p = transition probability density):

$$p(v_{i+1}|\mathbf{v}_i, \mathbf{w}_i) = p(v_{i+1}|\mathbf{v}_i)$$

T. Schreiber, Measuring information transfer. Phys. Rev. Lett. 85, 461, 2000

measuring interactions **information-theory**

direction of interaction: **transfer entropy**

3) if systems X and Y independent → Markov condition fulfilled

4) use relative entropy concept to quantify *incorrectness* of Markov condition; with this, transfer entropy is defined as:

$$T_{\mathrm{Y}\rightarrow\mathrm{X}} = \sum_i^N p\left(v_{i+1}, \mathbf{v}_i^{(k)}, \mathbf{w}_i^{(l)}\right) \log \frac{p\left(v_{i+1} | \mathbf{v}_i^{(k)}, \mathbf{w}_i^{(l)}\right)}{p\left(v_{i+1} | \mathbf{v}_i^{(k)}\right)}$$

(l,k) denote orders of Markov processes

$T_{\mathrm{X}\rightarrow\mathrm{Y}}$ defined in complete analogy

measuring interactions **information-theory**

direction of interaction: **transfer entropy**

properties of transfer entropy

- can detect direction of information flow since $T_{\mathrm{Y}\to\mathrm{X}} \neq T_{\mathrm{X}\to\mathrm{Y}}$
- unbounded, needs suitable definition of directionality, e.g.

$$T := T_{\mathrm{Y}\to\mathrm{X}} - T_{\mathrm{X}\to\mathrm{Y}} \begin{cases} > 0: & \text{Y drives X} \\ = 0: & \text{no or symmetric bidir. coupling} \\ < 0: & \text{X drives Y} \end{cases}$$

- depends on coupling strength → data-driven estimator for direction of interaction
- for Gaussian distributed data, transfer entropy equals Granger causality
- similar to conditional mutual information (replace system Z by e.g., past of system Y)

measuring interactions **information-theory**

direction of interaction: **transfer entropy**

extensions:

- multivariate (partial) transfer entropy
- various estimation techniques
- estimators for transient signals* and delay-systems**

* H. Dickten, K. Lehnertz K. Identifying delayed directional couplings with symbolic transfer entropy. Phys Rev E 90, 062706, 2014

** M. Martini, TA Kranz, T Wagner, K Lehnertz K. Inferring directional interactions from transient signals with symbolic transfer entropy. Phys Rev E 83, 011919, 2011

measuring interactions **information-theory**

strength and direction of interaction

how to estimate probability density distributions and the joint distribution densities from time series?

- counting (cumbersome)
- various binning techniques
- nearest neighbor estimators (e.g. Kozachenko-Leonenko)
- correlation sum (via phase-space embeddings)
- symbolization (e.g. based or permutation entropy*)

* C. Bandt, B. Pompe. Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett., 88, 174102, 2002

measuring interactions **information-theory**

strength and direction of interaction ***an example***

estimators based on the concept of symbolic dynamics and on symbolization

symbolic dynamics: modeling a smooth dynamical system by a *discrete space* consisting of infinite *sequences of symbols*, each of which corresponds to a state of the system, with the dynamics (evolution) given by the shift operator

symbolization: generate symbols via delay embedding

$$\mathbf{s}_i := \left(v_{i+(j_1-1)\tau}, v_{i+(j_2-1)\tau}, \ldots, v_{i+(j_m-1)\tau}\right)$$

where

$$v_{i+(j_1-1)\tau} \leq v_{i+(j_2-1)\tau} \leq \ldots \leq v_{i+(j_m-1)\tau}$$

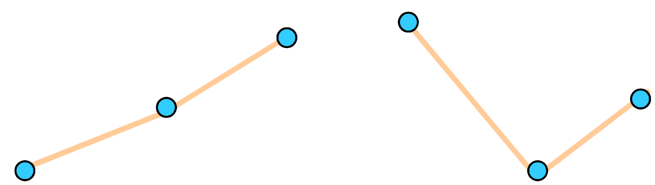
→ symbol $\hat{s}_i := \left(j_1, j_2, \ldots, j_m\right)$

* C. Bandt, B. Pompe. Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett., 88, 174102, 2002

measuring interactions **information-theory**

strength and direction of interaction ***an example***

embedded data: (3, 5, 9) (10, 1, 6)



symbols $\Rightarrow$ (1, 2, 3) (1, 3, 2)

permutation entropy: $H(m) = -\sum_{i=1}^{m!} \hat{s}_i \log \hat{s}_i$

normalization: $0 \leq H = \frac{H(m)}{\log(m!)} \leq 1$

$H \rightarrow$ 0 for deterministic systems, $H \rightarrow$ 1 for stochastic systems

measuring interactions information-theory

strength and direction of interaction *an example*

given time series of systems X and Y:

- estimate permutation entropy from windowed data
- investigate changing tendency of permutation entropies

$$S(w_i) = \begin{cases} +1 : & \text{if } H(w_i) < H(w_{i+1}) \\ -1 : & \text{else} \end{cases}$$

- characterize in-step behavior of pairs of permutation entropies

$$\gamma := \sum_{i=1}^{N_w} S_{\mathrm{X}}(w_i) S_{\mathrm{Y}}(w_i)$$

- $\gamma = 0$ for independent systems; $\gamma \rightarrow 1$ for synchronized systems; γ increase monotonically with increasing coupling strength → data-driven estimator for strength of interaction

Z.H. Liu. Measuring the degree of synchronization from time series data. Europhys. Lett., 68: 19, 2004

measuring interactions **information-theory**

strength and direction of interaction ***an example***

given time series of systems X and Y:

- for estimating probability density distributions and the joint distribution densities:
  - replace probabilities of data with probabilities of symbols
  - count symbols (→ very fast)
- symbolic transfer entropy:

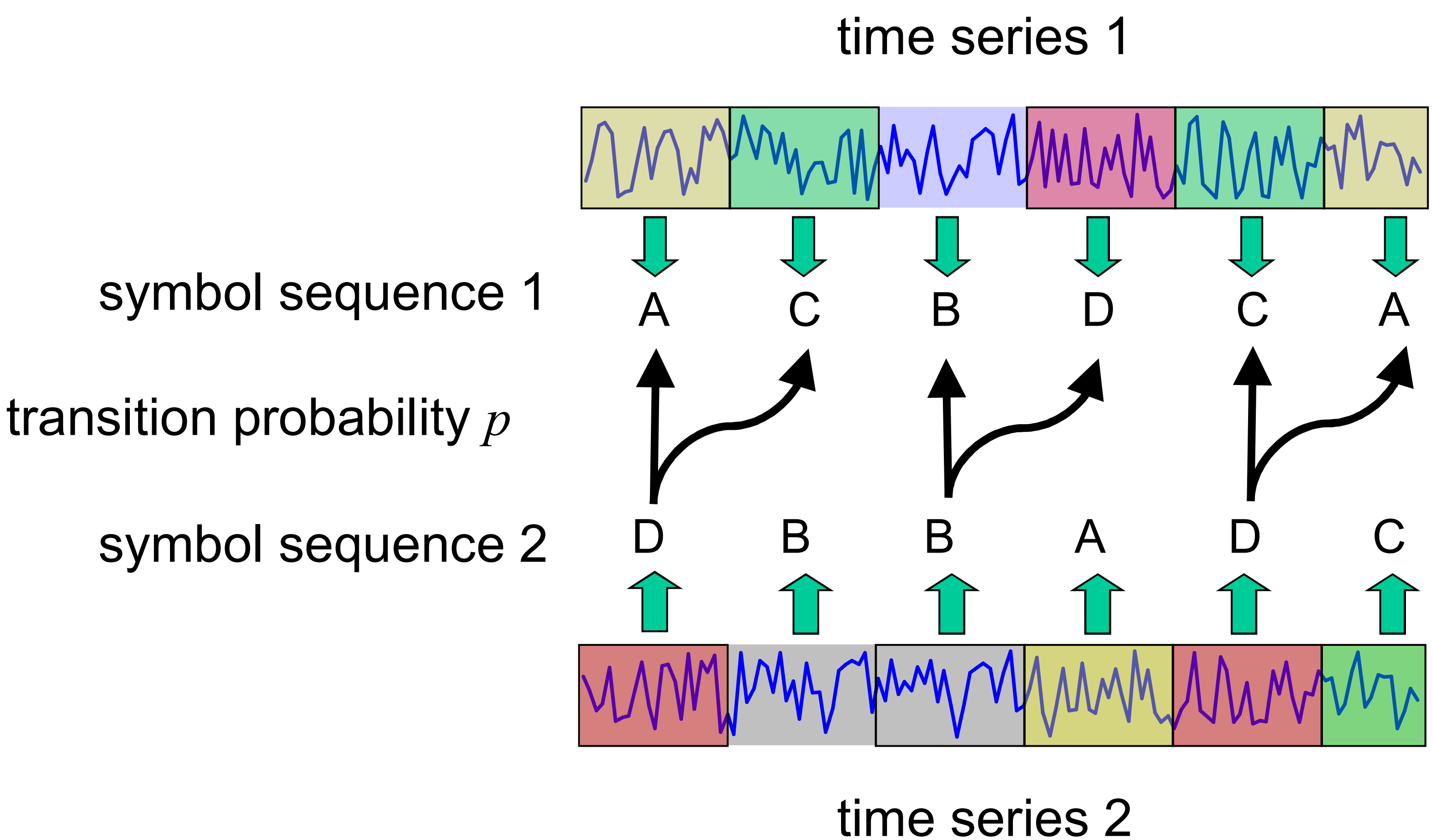
$$T^{\mathrm{S}}_{\mathrm{Y}\rightarrow\mathrm{X}} = \sum_{\hat{v}_{i+1},\hat{\mathbf{v}}_i^{(k)},\hat{\mathbf{w}}_i^{(l)}} p\left(\hat{v}_{i+1},\hat{\mathbf{v}}_i^{(k)},\hat{\mathbf{w}}_i^{(l)}\right) \log \frac{p\left(\hat{v}_{i+1}|\hat{\mathbf{v}}_i^{(k)},\hat{\mathbf{w}}_i^{(l)}\right)}{p\left(\hat{v}_{i+1}|\hat{\mathbf{v}}_i^{(k)}\right)}$$

- see properties of transfer entropy
- easy-to-use data driven estimator for direction of interaction

M. Staniek, K. Lehnertz. Symbolic Transfer Entropy. Phys. Rev. Lett. 100, 158101, 2008

measuring interactions **information-theory**

strength and direction of interaction ***an example***



M. Staniek, K. Lehnertz. Symbolic Transfer Entropy. Phys. Rev. Lett. 100, 158101, 2008

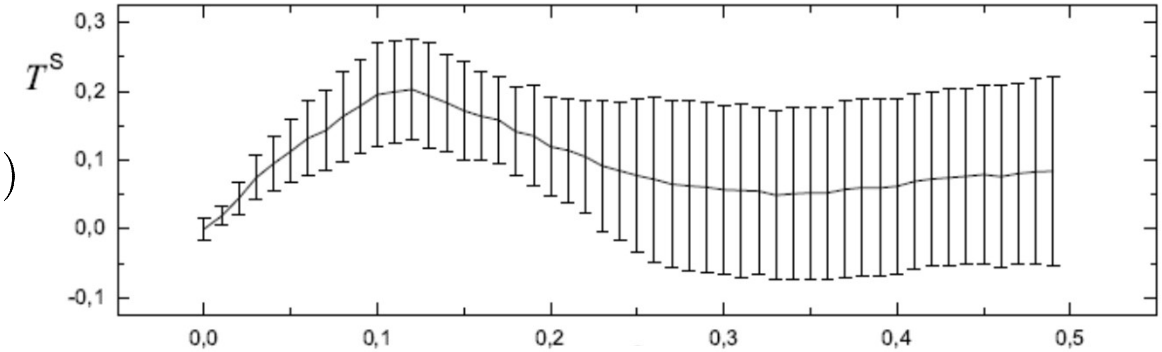
measuring interactions — information-theory

strength and direction of interaction — an example

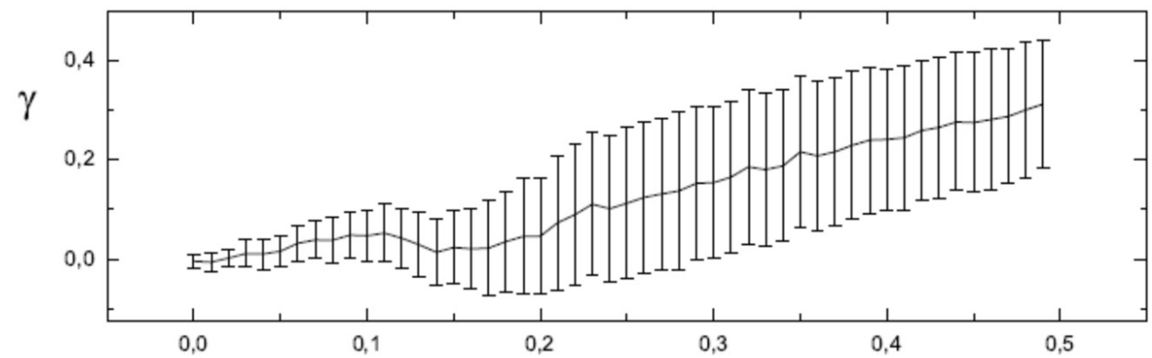
diffusively coupled Rössler oscillators
(100 realizations)

$$
\begin{aligned}
\dot{x}_{1,2} &= -\Omega_{1,2} y_{1,2} - z_{1,2} + c_{1,2}(x_{2,1} - x_{1,2}) \\
\dot{y}_{1,2} &= \Omega_{1,2} x_{1,2} + 0.165 y_{1,2} \\
\dot{z}_{1,2} &= 0.2 + z_{1,2}(x_{1,2} - 10) \\
\Omega_{1,2} &\in \mathcal{N}(0.89; 0.1) \\
c_1 &= 0 \\
c_2 &= c
\end{aligned}
$$

direction of interaction



strength of interaction



coupling strength c

M. Staniek, K. Lehnertz. Symbolic Transfer Entropy. Phys. Rev. Lett. 100, 158101, 2008

measuring interactions information-theory

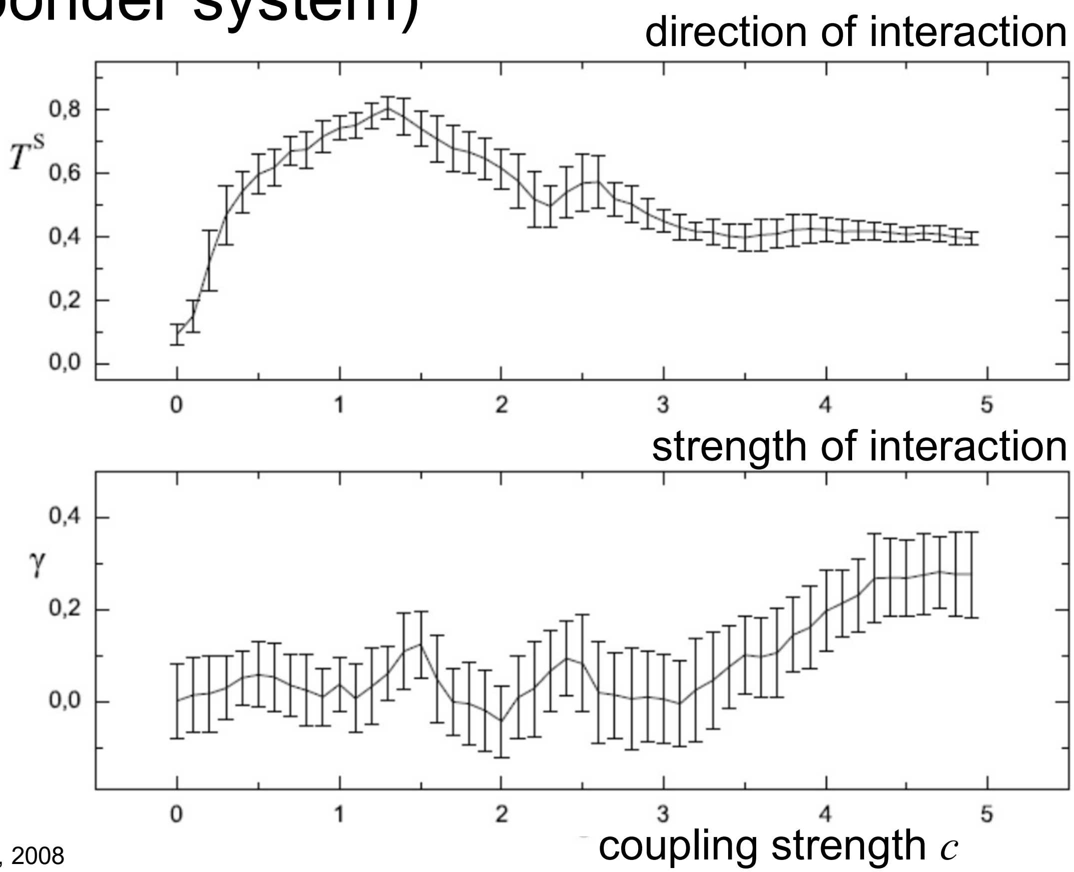
strength and direction of interaction *an example*

diffusively coupled Rössler-Lorenz oscillators
(100 realizations of driver-responder system)

$$\begin{aligned}
\dot{x}^{\mathrm{R}} &= -\Omega^{\mathrm{R}}(y^{\mathrm{R}} - z^{\mathrm{R}}) \\
\dot{y}^{\mathrm{R}} &= \Omega^{\mathrm{R}}(x^{\mathrm{R}} + 0.2y^{\mathrm{R}}) \\
\dot{z}^{\mathrm{R}} &= \Omega^{\mathrm{R}}\left(0.2 + z^{\mathrm{R}}(x^{\mathrm{R}} - 5.7)\right)
\end{aligned}$$

$$\begin{aligned}
\dot{x}^{\mathrm{L}} &= 10(y^{\mathrm{L}} - x^{\mathrm{L}}) \\
\dot{y}^{\mathrm{L}} &= \Omega^{\mathrm{L}}x^{\mathrm{L}} - y^{\mathrm{L}} - x^{\mathrm{L}}z^{\mathrm{L}} + c(y^{\mathrm{R}})^2 \\
\dot{z}^{\mathrm{L}} &= x^{\mathrm{L}}y^{\mathrm{L}} - \frac{8}{3}z^{\mathrm{L}}
\end{aligned}$$

$$\begin{aligned}
\Omega^{\mathrm{R}} &\in \mathcal{N}(6, 0.1) \\
\Omega^{\mathrm{L}} &\in \mathcal{N}(28, 1)
\end{aligned}$$

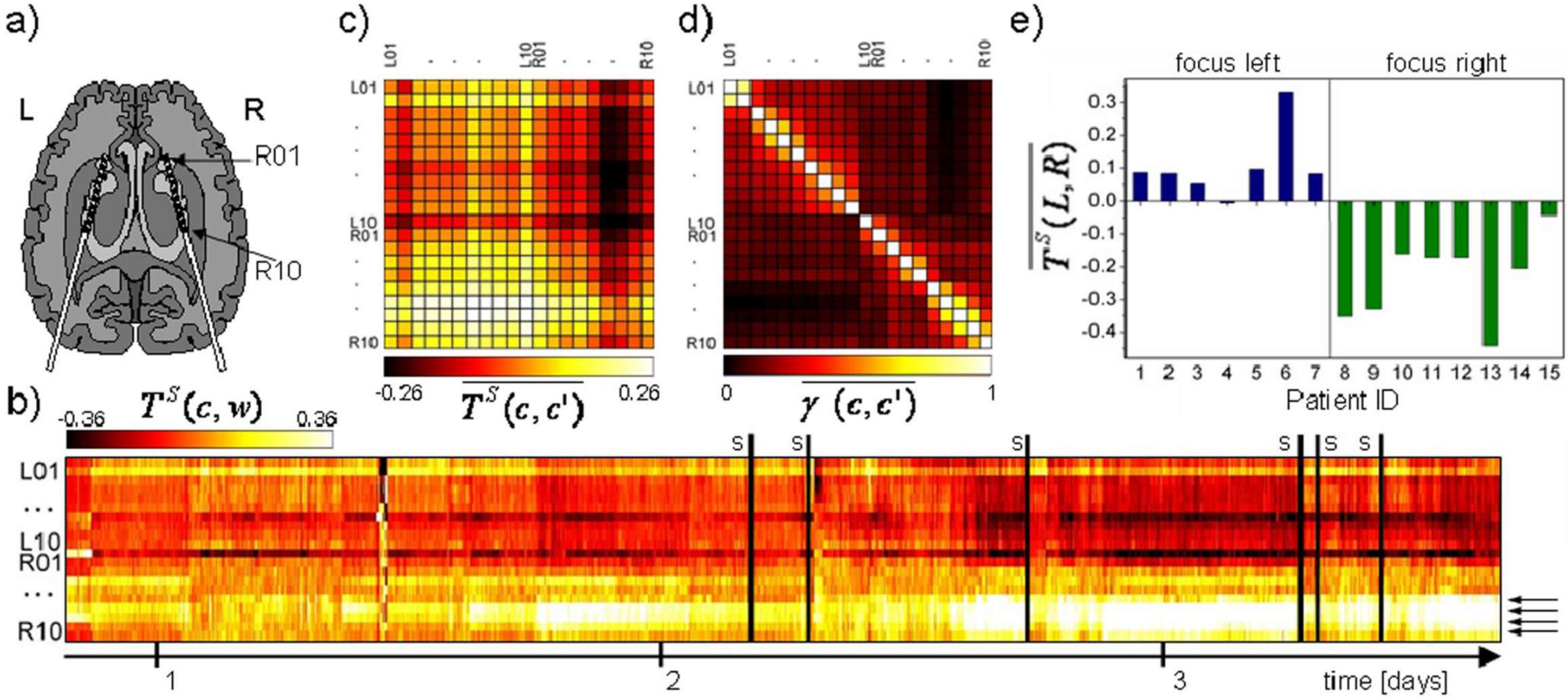


M. Staniek, K. Lehnertz. Symbolic Transfer Entropy. Phys. Rev. Lett. 100, 158101, 2008

measuring interactions information-theory

strength and direction of interaction an example

driver-responder relationships in EEG dynamics from epilepsy patients



M. Staniek, K. Lehnertz. Symbolic Transfer Entropy. Phys. Rev. Lett. 100, 158101, 2008

measuring interactions **information-theory**

strength and direction of interaction

permutation-entropy-based estimators

advantages

- easy-to-use, fast-to-calculate
- high robustness against noise (symbolization)

disadvantages

- symbolization may lead to loss of information
- require appropriate choice of embedding parameter
- choice of window-size, finiteness of available symbols
- "faster" system (eigen-frequency, noise) → driver (need reliable surrogate test for directionality)
- may be fooled by (unobserved) third system