Mathematics, Physics, and Machine Learning Webinar Series
Instituto Superior TÉCNICO, University of Lisbon

Towards a deeper understanding of high-order interdependencies in complex systems

Fernando E. Rosas, PhD
Postdoctoral Researcher
Data Science Institute
Centre for Psychedelic Research
Centre for Complexity Science
Imperial College London

Work carried out in synergistic collaboration with:

Pedro A.M. Mediano
(University of Cambridge)

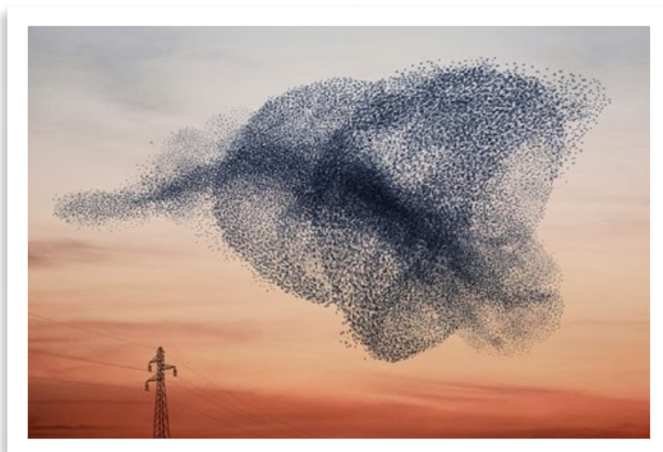
Andrea Luppi
(University of Cambridge)

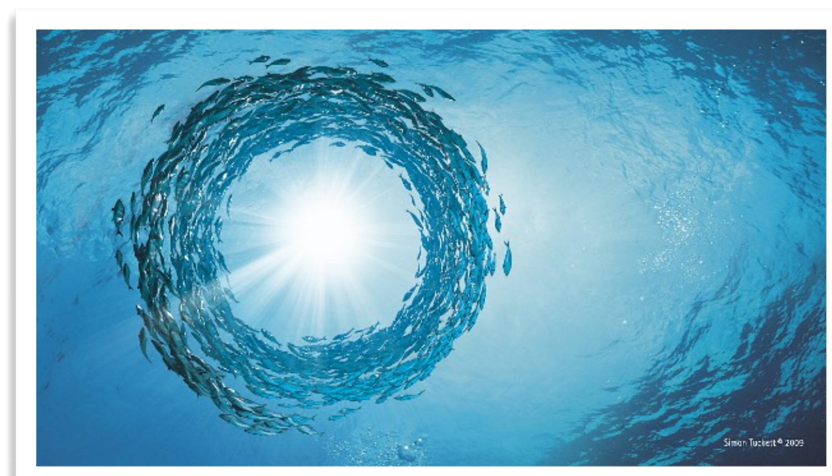
- Robin Carhart-Harris, Henrik Jensen (*Imperial College London*)
- Daniel Bor, Emmanuel Stamatakis (*University of Cambridge*)
- Anil Seth, Adam Barrett (*University of Sussex*)
- Borzoo Rassouli (*University of Essex*)
- Michael Gastpar (*EPFL*)
- Daniele Marinazzo (University of Gent)
- Sebastiano Estramaglia (Universitá degli Studi Aldo Moro)
- Rodrigo Cofre, Marilyn Gatica (Universidad de Valparaiso)

Imperial College London
UNIVERSITY OF CAMBRIDGE
US UNIVERSITY OF SUSSEX
University of Essex
EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
GHENT UNIVERSITY
UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
Universidad de Valparaíso CHILE

Synergy and emergence: what are they, and why they matter?

Informal definition: “*when the whole exhibit properties that the parts don’t*"





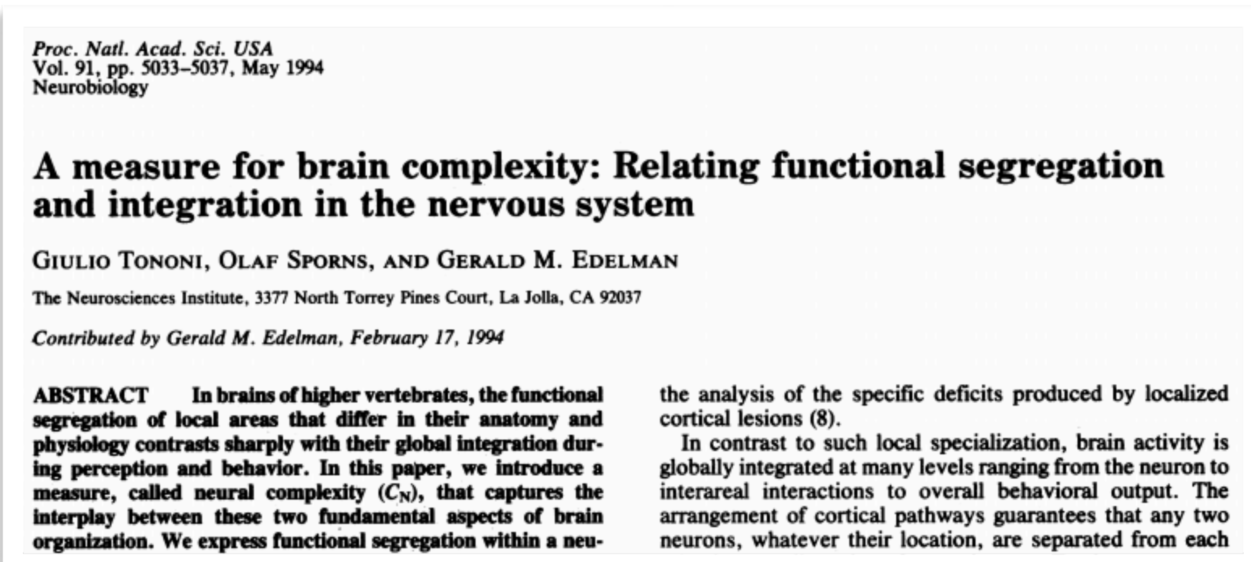
- Is **synergy** the same than **emergence**?
- Can they be **defined formally** and **measured from data**?
- What are they good for?

Today's menu
1. Synergy
2. Emergence
3. Ideas to take home

A hint from the brain…

High brain functions may be characterised for a coexistence of two "opposite" forces:

- *Differentiation (local independency)*: the parts of the brain has differentiated functions
- *Integration (global cohesion)*: these parts are integrated into high-level functions



Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 5033–5037, May 1994
Neurobiology

A measure for brain complexity: Relating functional segregation and integration in the nervous system

GIULIO TONONI, OLAF SPORNS, AND GERALD M. EDELMAN

The Neurosciences Institute, 3377 North Torrey Pines Court, La Jolla, CA 92037

Contributed by Gerald M. Edelman, February 17, 1994

ABSTRACT **In brains of higher vertebrates, the functional segregation of local areas that differ in their anatomy and physiology contrasts sharply with their global integration during perception and behavior. In this paper, we introduce a measure, called neural complexity (C_N), that captures the interplay between these two fundamental aspects of brain organization. We express functional segregation within a neu-**

the analysis of the specific deficits produced by localized cortical lesions (8).

In contrast to such local specialization, brain activity is globally integrated at many levels ranging from the neuron to interareal interactions to overall behavioral output. The arrangement of cortical pathways guarantees that any two neurons, whatever their location, are separated from each

G. Tononi

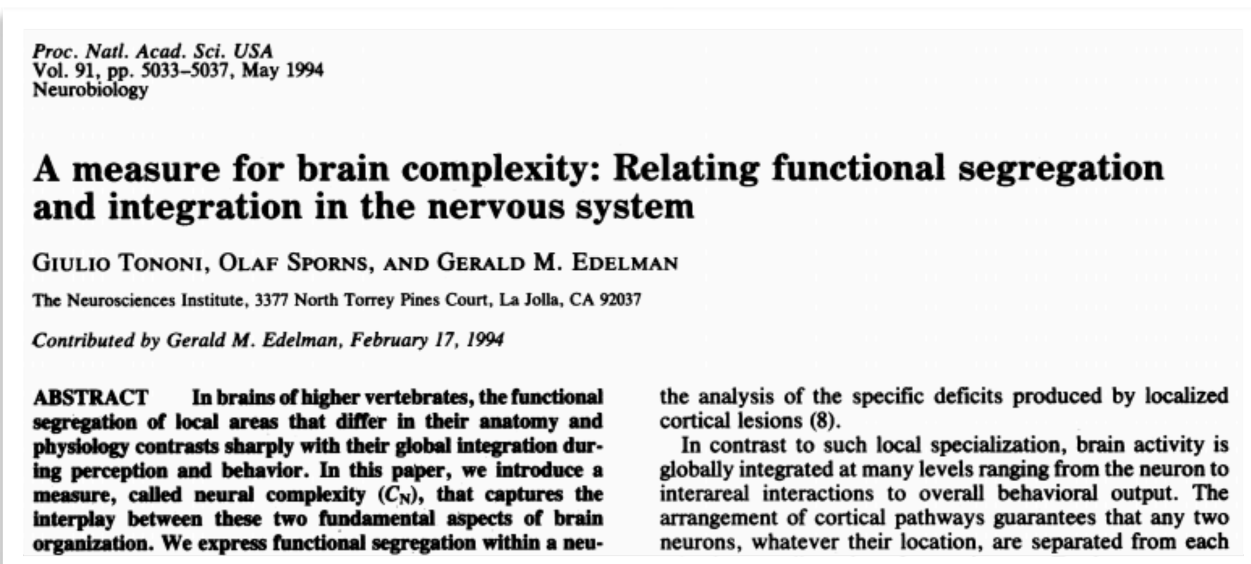
O. Sporns

G. Edelman

A hint from the brain…

High brain functions may be characterised for a coexistence of two “opposite” forces:

- *Differentiation (local independency)*: the parts of the brain has differentiated functions
- *Integration (global cohesion)*: these parts are integrated into high-level functions



Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 5033–5037, May 1994
Neurobiology

A measure for brain complexity: Relating functional segregation and integration in the nervous system

Giulio Tononi, Olaf Sporns, and Gerald M. Edelman

The Neurosciences Institute, 3377 North Torrey Pines Court, La Jolla, CA 92037

Contributed by Gerald M. Edelman, February 17, 1994

ABSTRACT In brains of higher vertebrates, the functional segregation of local areas that differ in their anatomy and physiology contrasts sharply with their global integration during perception and behavior. In this paper, we introduce a measure, called neural complexity (C_N), that captures the interplay between these two fundamental aspects of brain organization. We express functional segregation within a neu-

the analysis of the specific deficits produced by localized cortical lesions (8).

In contrast to such local specialization, brain activity is globally integrated at many levels ranging from the neuron to interareal interactions to overall behavioral output. The arrangement of cortical pathways guarantees that any two neurons, whatever their location, are separated from each

G. Tononi

O. Sporns

G. Edelman

Crucial insight:

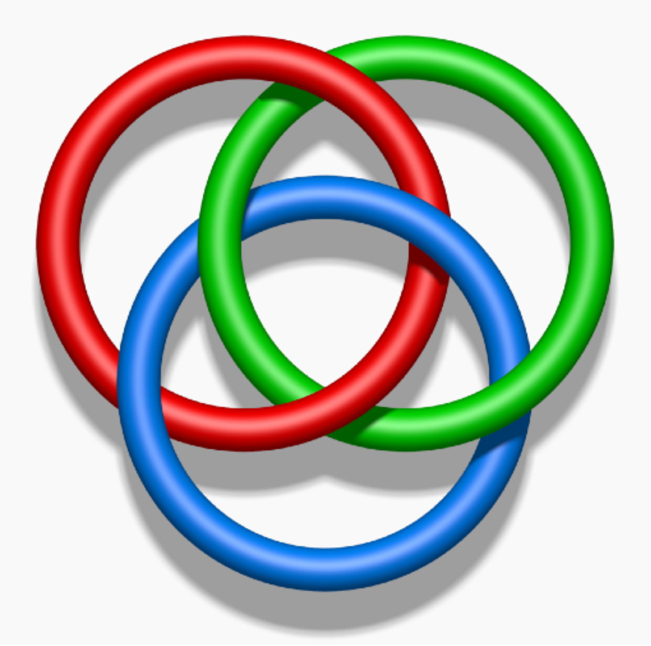
Integration and differentiation are not antithetical, but can coexist!

Synergy

A hint from the brain…

High brain functions may be characterised for a coexistence of two "opposite" forces:

- *Differentiation (local independency)*: the parts of the brain has differentiated functions
- *Integration (global cohesion)*: these parts are integrated into high-level functions

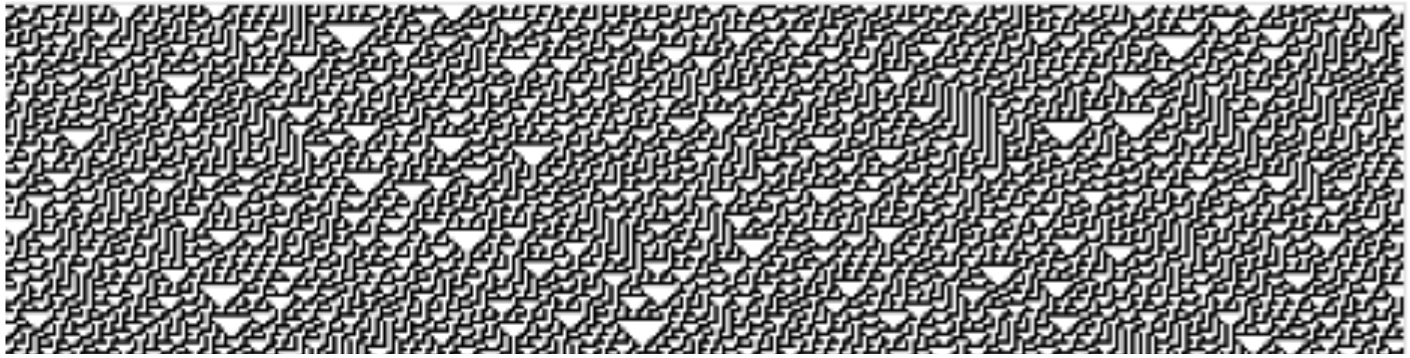


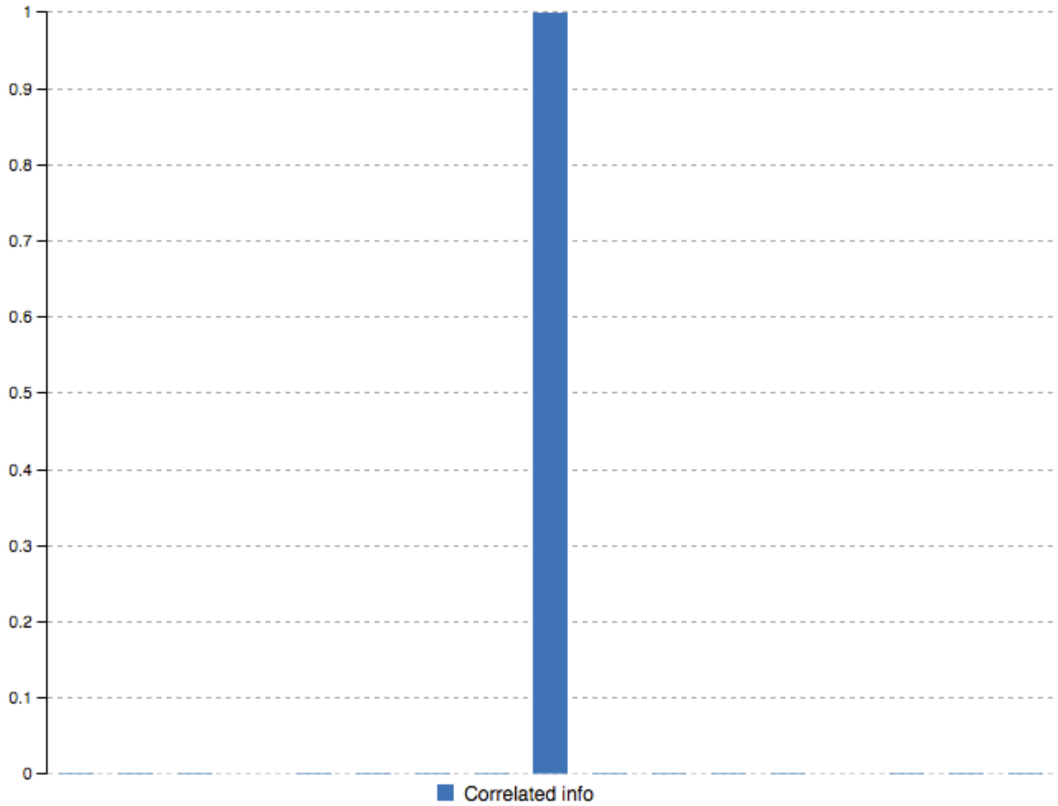
Crucial insight:

Integration and differentiation are not antithetical, but can coexist! ← ***Synergy***

High-order effects can make a difference

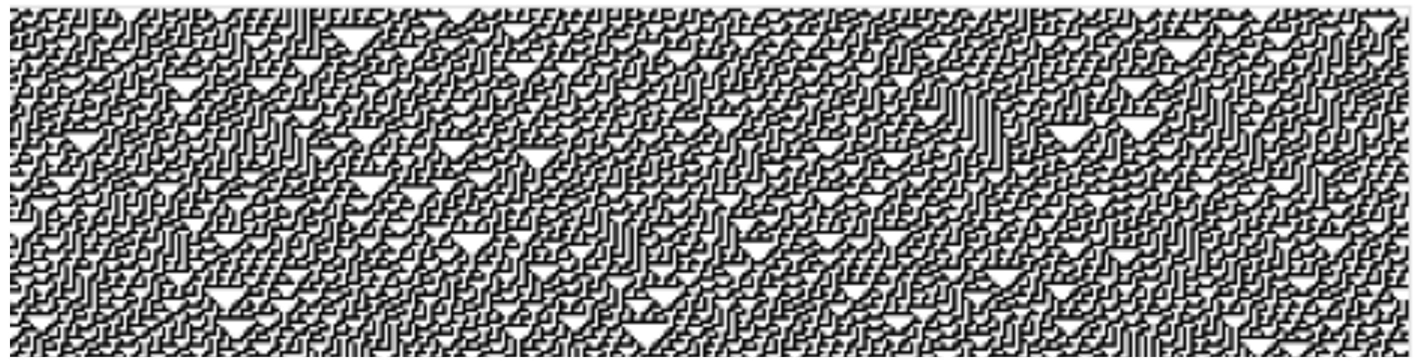
Rule 232

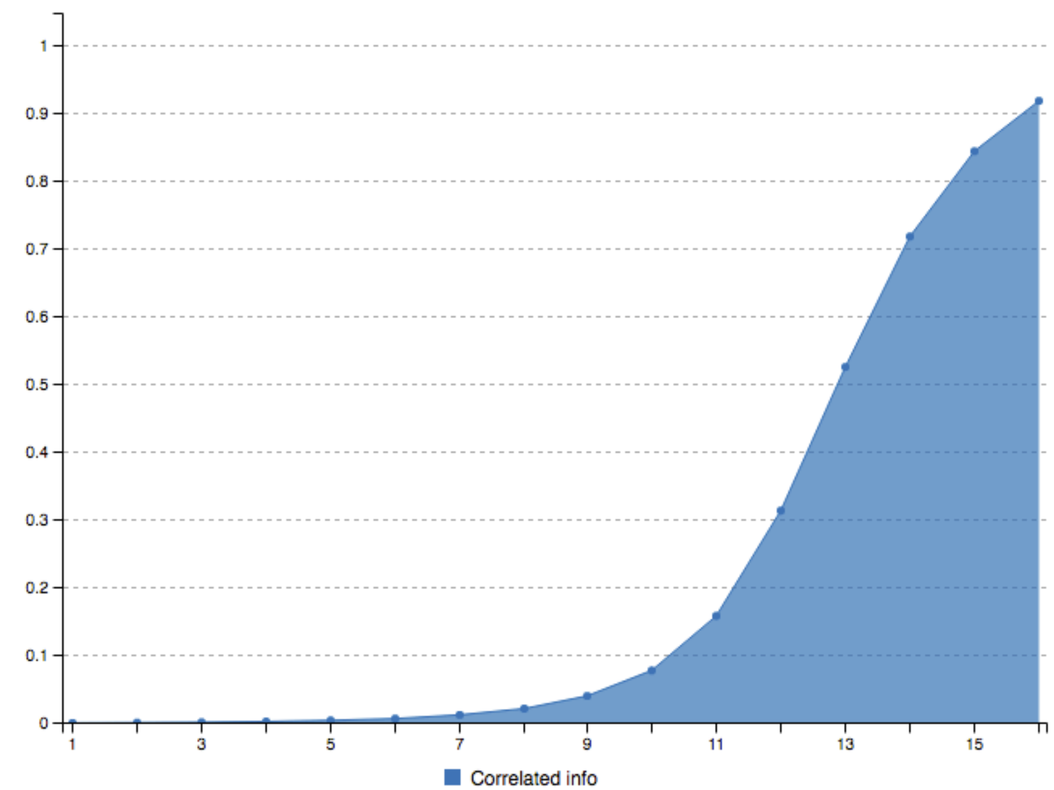




High-order effects can make a difference

Rule 232

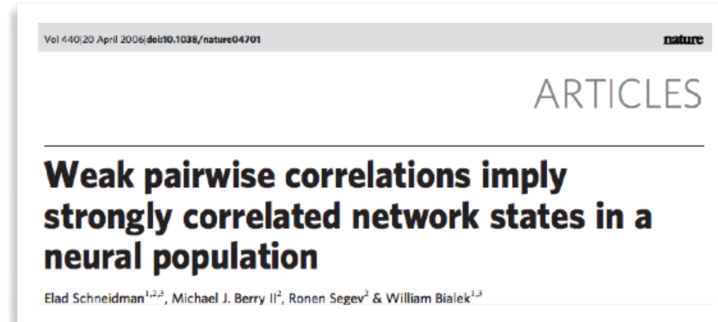




Rosas, F., Mediano, P. A., Ugarte, M., & Jensen, H. J. (2018). An information-theoretic approach to self-organisation: Emergence of complex interdependencies in coupled dynamical systems. Entropy, 20(10), 793.

High-order effects can make a difference

The unreasonable effectiveness of pairwise correlations…


Vol 440|20 April 2006|doi:10.1038/nature04701 nature

ARTICLES

Weak pairwise correlations imply strongly correlated network states in a neural population

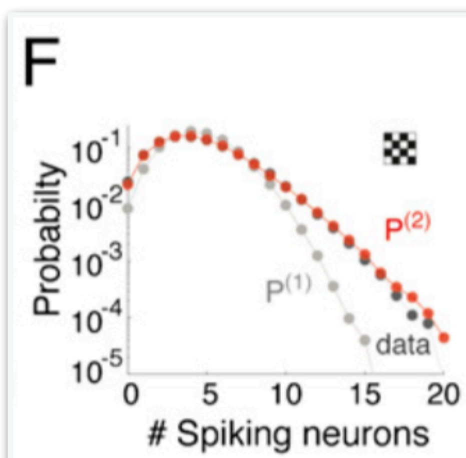
Elad Schneidman[1,2,3], Michael J. Berry II[2], Ronen Segev[2] & William Bialek[1,3]

…might depend on the stimuli!

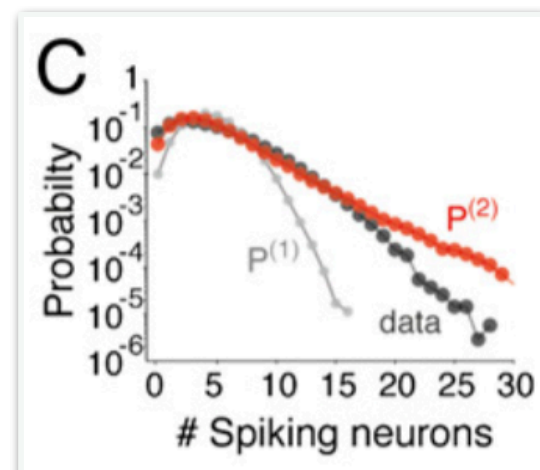

Sparse low-order interaction network underlies a highly correlated and learnable neural population code

Elad Ganmor[a], Ronen Segev[b,1], and Elad Schneidman[a,1]

[a]Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel; and [b]Department of Life Sciences and The Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel



retina exposed to white noise stimuli



retina exposed to natural images

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. Nature, 440(7087), 1007-1012.

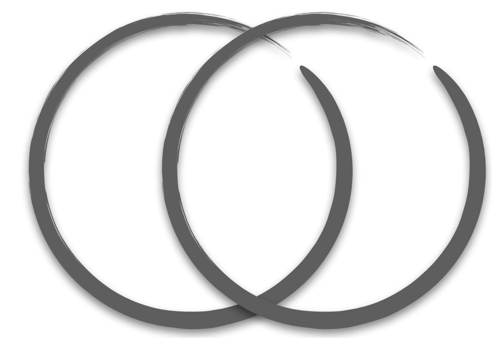
Ganmor, E., Segev, R., & Schneidman, E. (2011). Sparse low-order interaction network underlies a highly correlated and learnable neural population code. Proceedings of the National Academy of sciences, 108(23), 9679-9684.

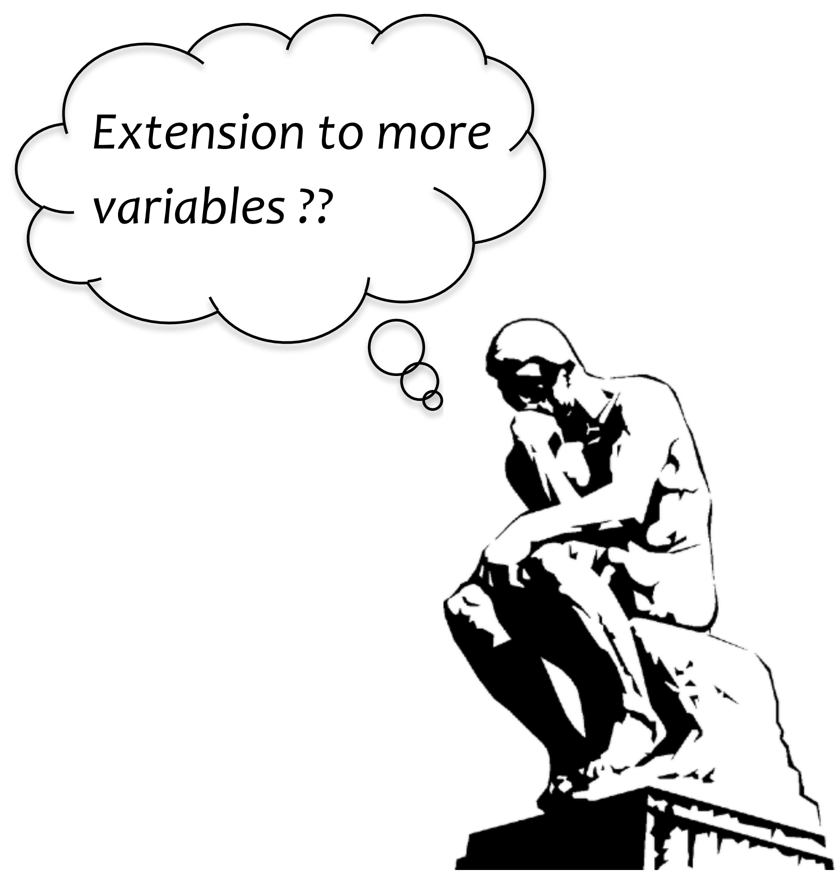
The three faces of the extended mutual information

Consider a system described by the random vector $\boldsymbol{X}^n = (X_1, \ldots, X_n)$, where each variable takes values over an alphabet $\mathcal{X}_j$.

Mutual information (*Shannon and Weaver*, 1949):

$$I(X_1;X_2) = H(X_1) + H(X_2) - H(X_1,X_2)$$
$$= H(X_1,X_2) - H(X_1|X_2) - H(X_2|X_1)$$



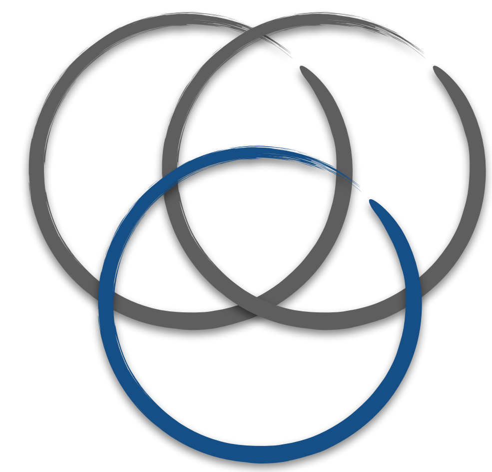


The three faces of the extended mutual information

Consider a system described by the random vector $\boldsymbol{X}^n = (X_1, \ldots, X_n)$, where each variable takes values over an alphabet $\mathcal{X}_j$.

Mutual information (*Shannon and Weaver*, 1949):

$$I(X_1; X_2) = H(X_1) + H(X_2) - H(X_1, X_2)$$
$$= H(X_1, X_2) - H(X_1|X_2) - H(X_2|X_1)$$



Three roads for defining a multivariate extension:

1. **Co-information** (*McGill* 1954): $I(X_1; X_2; X_3) = \sum_i H(X_i) - \sum_{j,k} H(X_j X_k) + H(X_1 X_2 X_3)$
   Topology

Baudot, P., & Bennequin, D. (2015). The homological nature of entropy. Entropy, 17(5), 3253-3318.

The three faces of the extended mutual information

Consider a system described by the random vector $\boldsymbol{X}^n = (X_1, \ldots, X_n)$, where each variable takes values over an alphabet $\mathcal{X}_j$.

Mutual information (*Shannon and Weaver*, 1949):

$$I(X_1; X_2) = H(X_1) + H(X_2) - H(X_1, X_2)$$
$$= H(X_1, X_2) - H(X_1|X_2) - H(X_2|X_1)$$

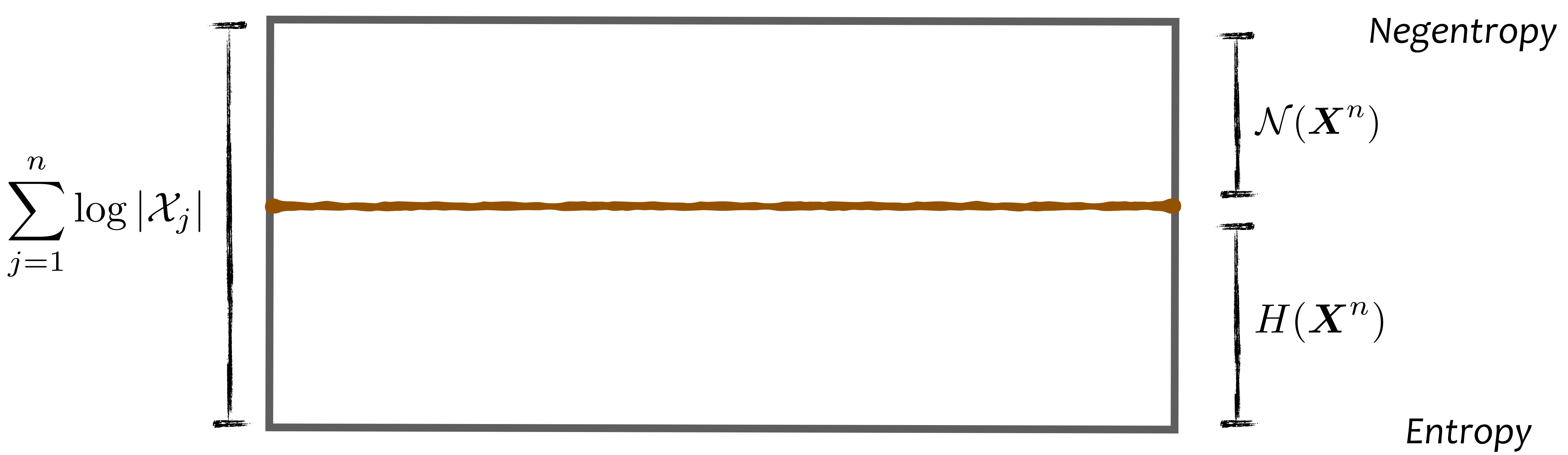
Three roads for defining a multivariate extension:

1. **Co-information** (*McGill* 1954): $I(X_1; X_2; X_3) = \sum_i H(X_i) - \sum_{j,k} H(X_j X_k) + H(X_1 X_2 X_3)$
   Topology

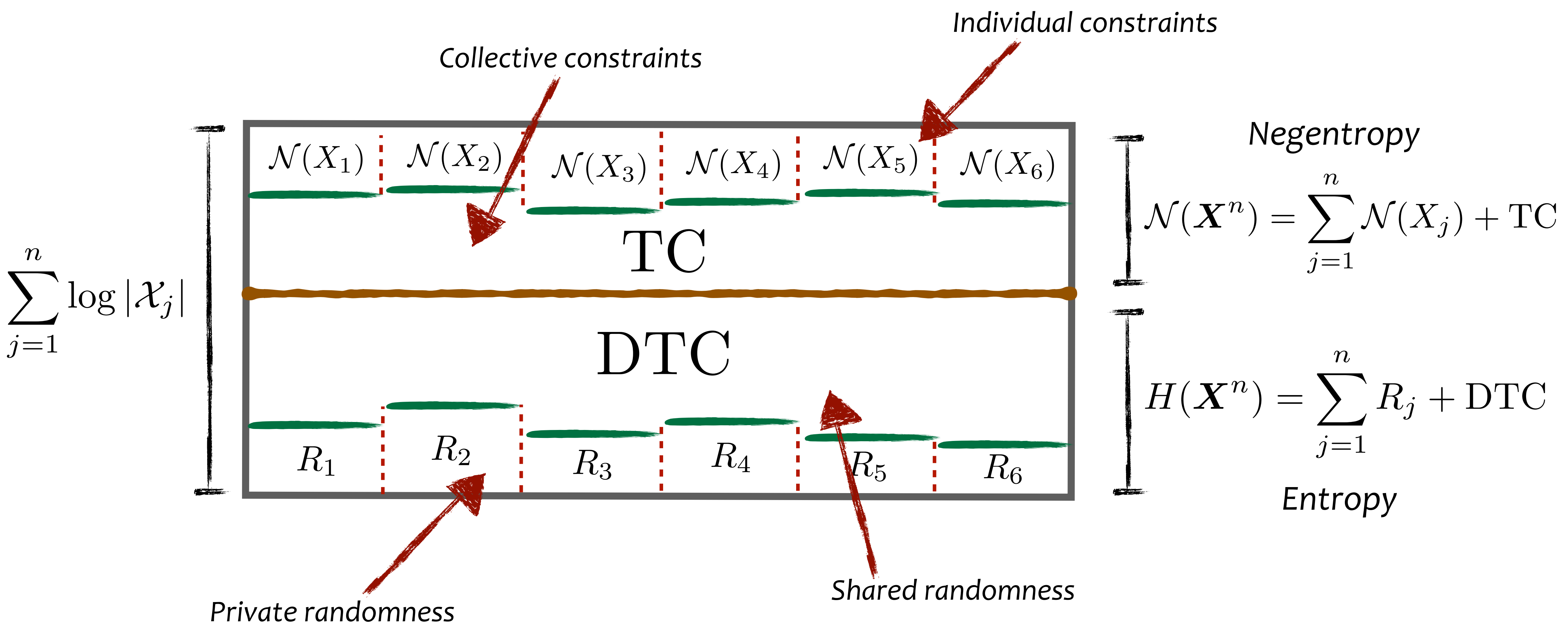
2. ***Total correlation*** (*Watanabe* 1960): $\mathrm{TC}(\boldsymbol{X}^n) = \sum_{j=1}^{n} H(X_j) - H(\boldsymbol{X}^n)$
   "External correlations"

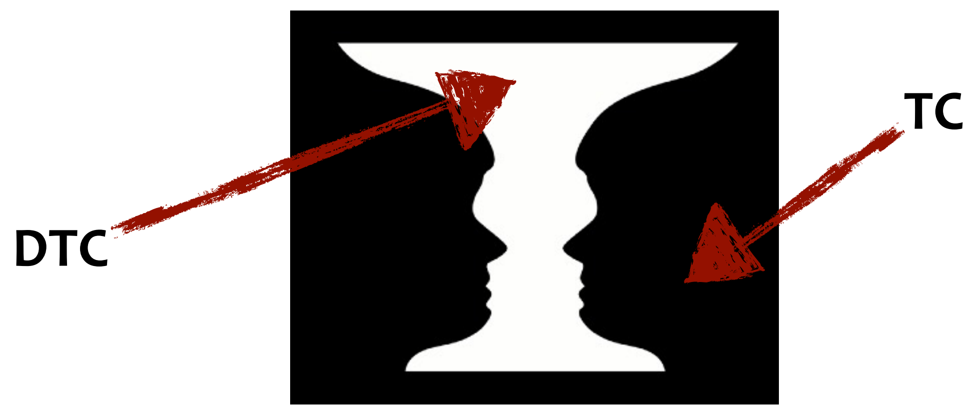
3. **Dual total correlation** (*Sun* 1978): $\mathrm{DTC}(\boldsymbol{X}^n) = H(\boldsymbol{X}^n) - \sum_{j=1}^{n} H(X_j | \boldsymbol{X}^n_{-j})$
   "Internal decomposition"

Putting the TC and DTC together…



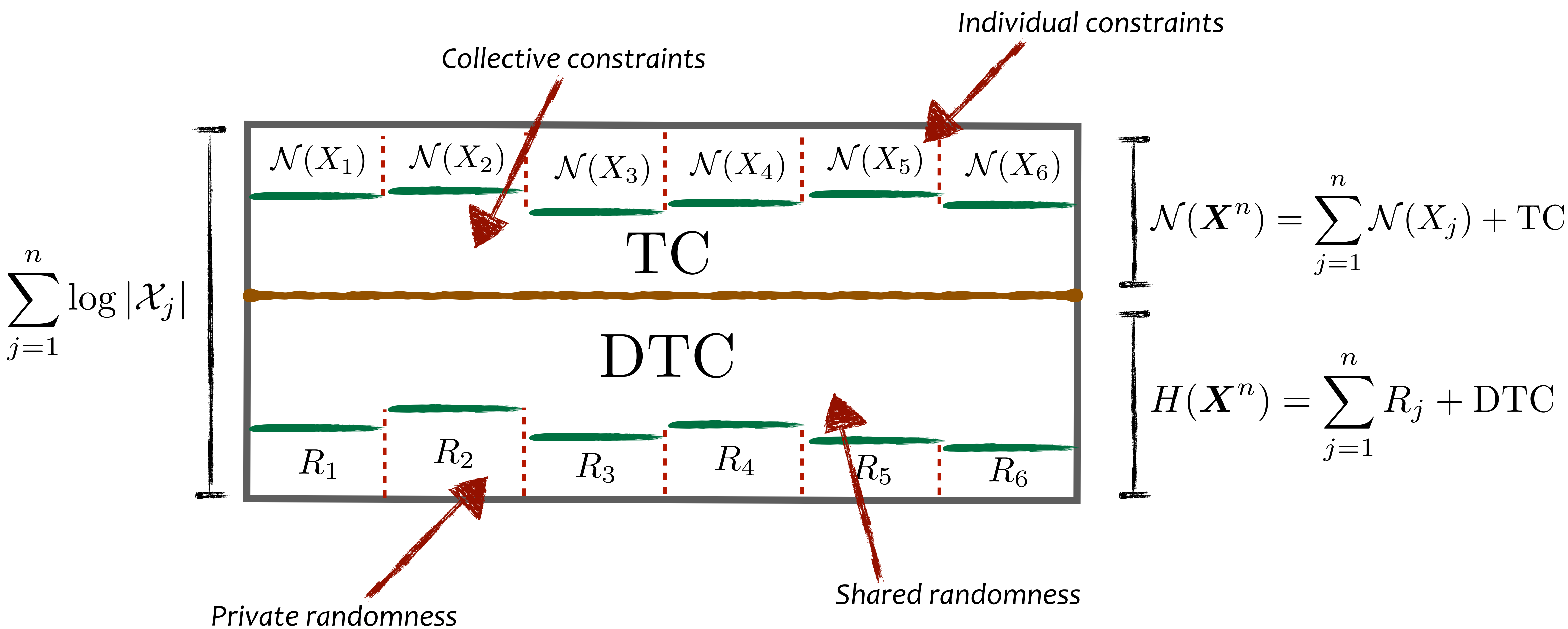
Putting the TC and DTC together…





F. E., Mediano, P. A., Gastpar, M., & Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E, 100(3), 032305.

Putting the TC and DTC together…



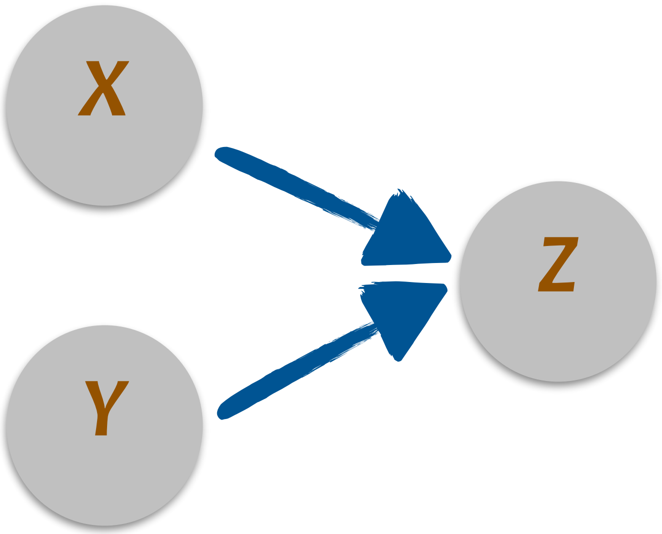
Example:

$X_1 = X_2 = X_3$	$X_3 = X_1 \oplus X_2$
$\mathrm{TC} = 2 > \mathrm{DTC} = 1$	$\mathrm{TC} = 1 < \mathrm{DTC} = 2$
redundancy!	***synergy!***

F. E., Mediano, P. A., Gastpar, M., & Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E, 100(3), 032305.

Making sense of synergistic effects

Making sense of synergistic effects



X and **Y** are used to predict **Z.**

Total predictability: $I(X, Y; Z)$

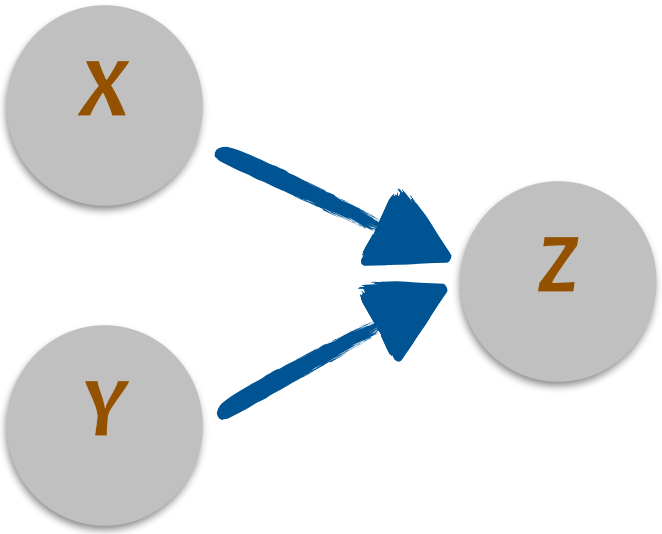
Information provided by **X** or **Y** : $I(X; Z)$

$I(Y; Z)$

However, in some situations

$$\underbrace{I(X, Y; Z)}_{\text{“the whole”}} > \underbrace{I(X; Z) + I(Y; Z)}_{\text{“the sum of the parts”}}$$

Making sense of synergistic effects



X and **Y** are used to predict **Z.**

Total predictability: $I(X,Y;Z)$

Information provided by **X** or **Y** : $I(X;Z)$
$I(Y;Z)$

However, in some situations

$$I(X,Y;Z) > I(X;Z) + I(Y;Z)$$

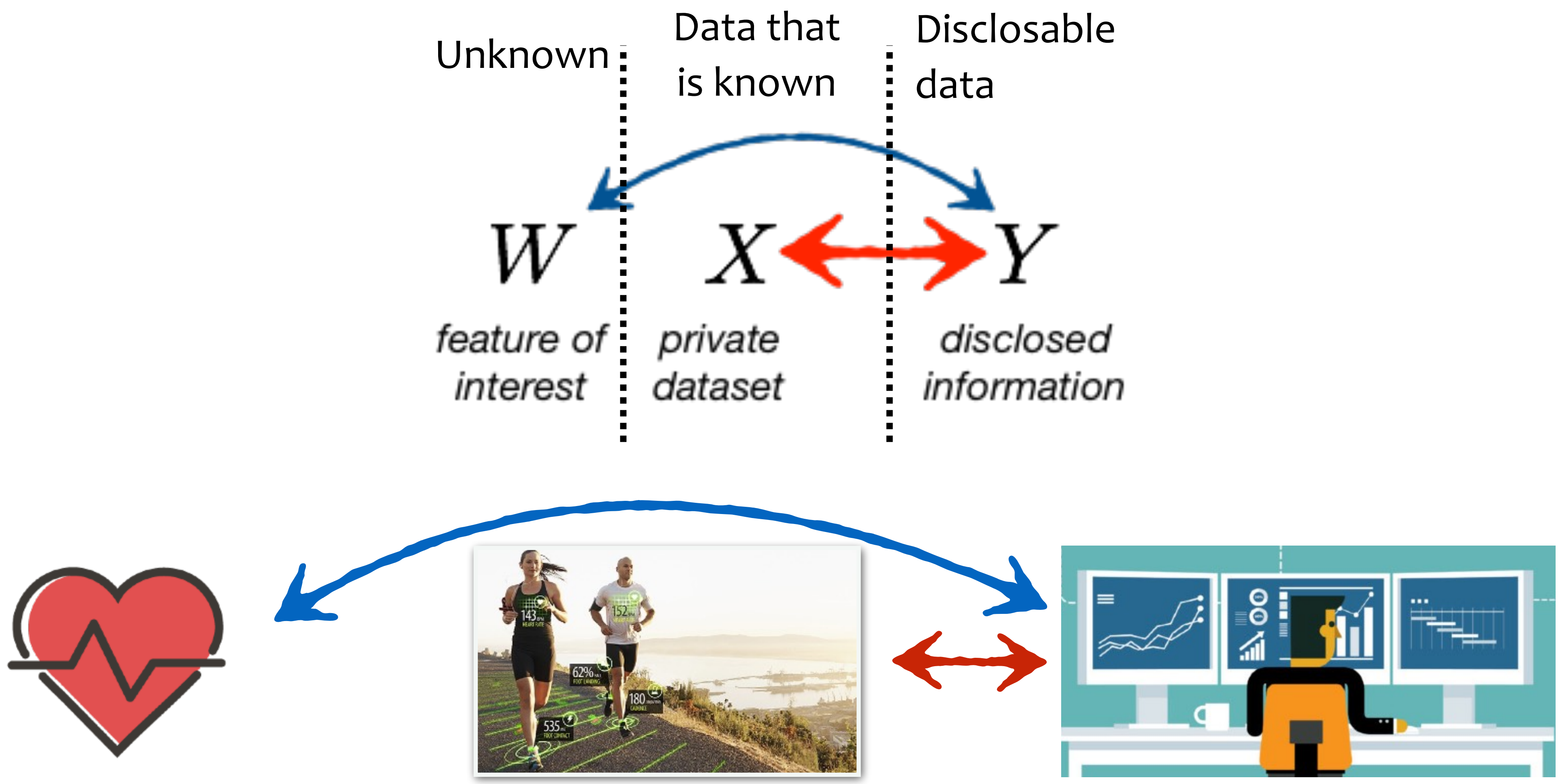
"the whole" *"the sum of the parts"*

"Partial Information decomposition" (PID):

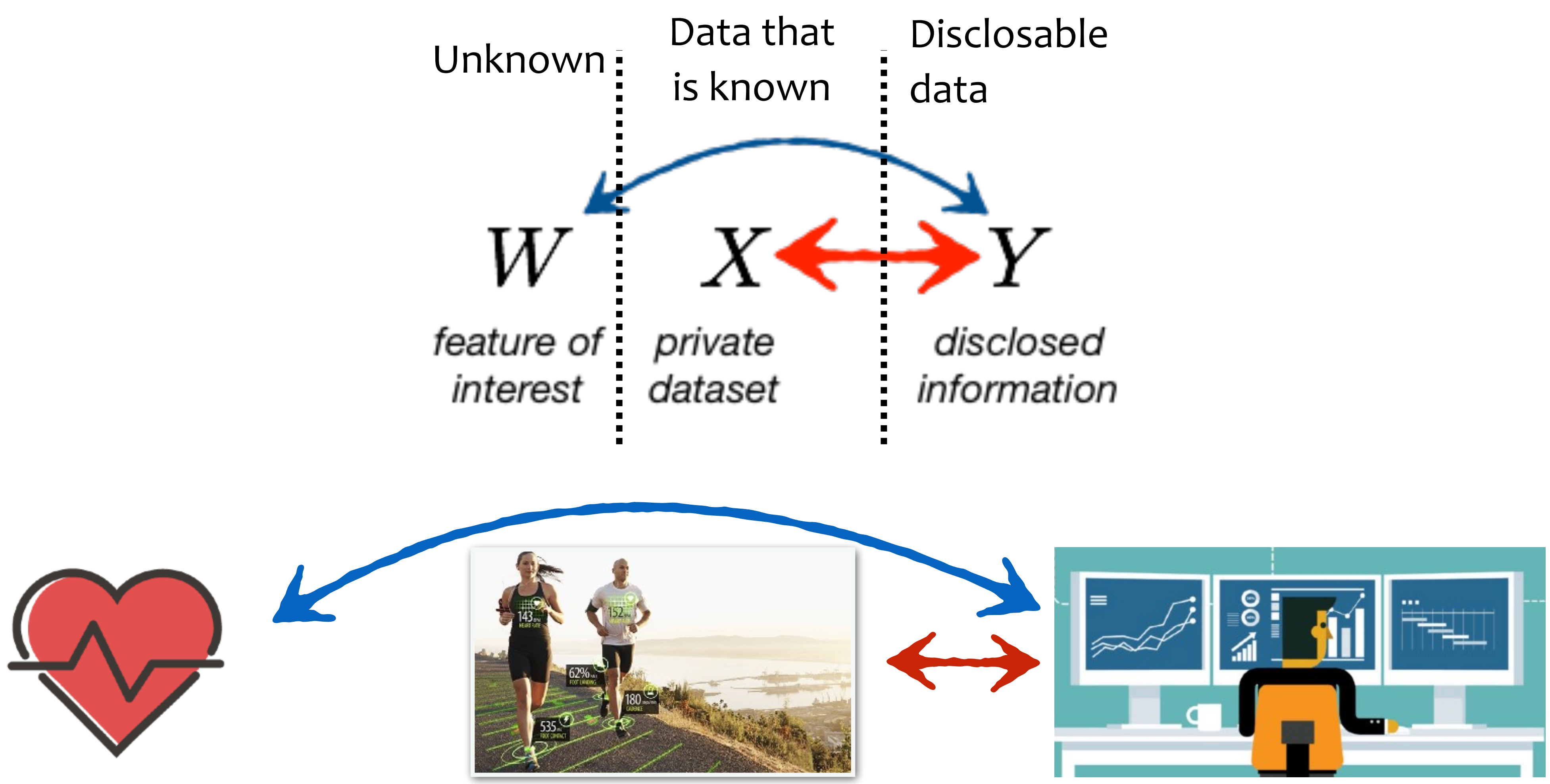
$$I(XY;Z) = \mathrm{Red}(XY;Z) + \mathrm{Un}(X;Z|Y) + \mathrm{Un}(Y;Z|X) + \mathrm{Syn}(XY;Z)$$

P. Williams and R. D. Beer. "Nonnegative decomposition of multivariate information." arXiv preprint arXiv:1004.2515 (2010).

Measuring synergy via information privacy



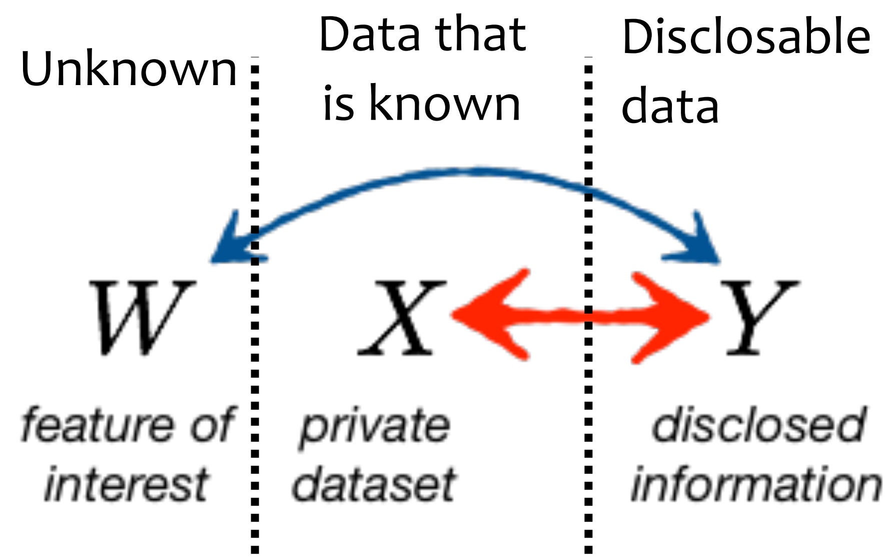
Measuring synergy via information privacy



The synergy/security paradox: *Could one learn something useful about a dataset, without learning anything about any individual sample?*

YES!!

Measuring synergy via information privacy



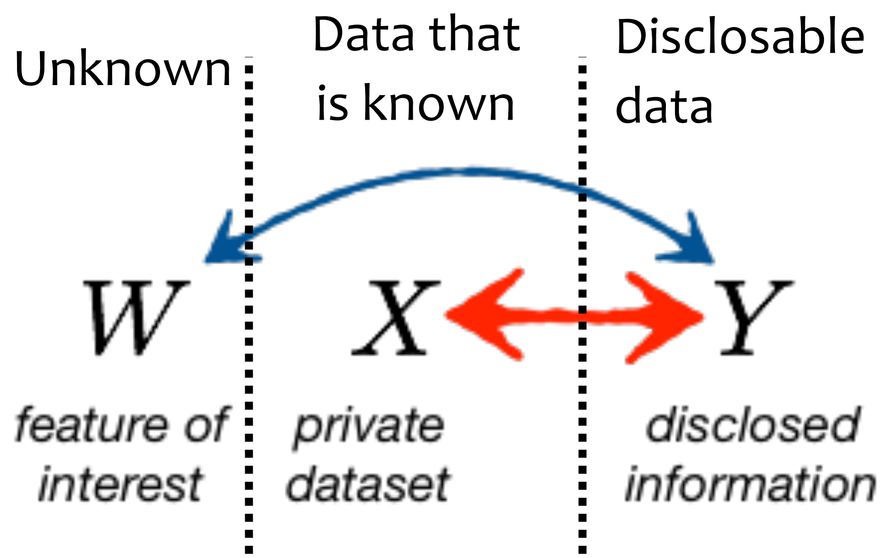
The synergy/security paradox: *Could one learn something useful about a dataset, without learning anything about any individual sample?*

YES!!

Example:

X = (A, B, C) → **Y = (A** (xor) **B, B** (xor) **C)**

Measuring synergy via information privacy



Definition of synergy based on sample privacy:

$$S^{\boldsymbol{\alpha}}(\boldsymbol{X} \to Y) := \sup_{\substack{p_{V|\boldsymbol{X}} \in \mathcal{C}(\boldsymbol{X};\boldsymbol{\alpha}): \\ V - \boldsymbol{X} - Y}} I(V;Y)$$

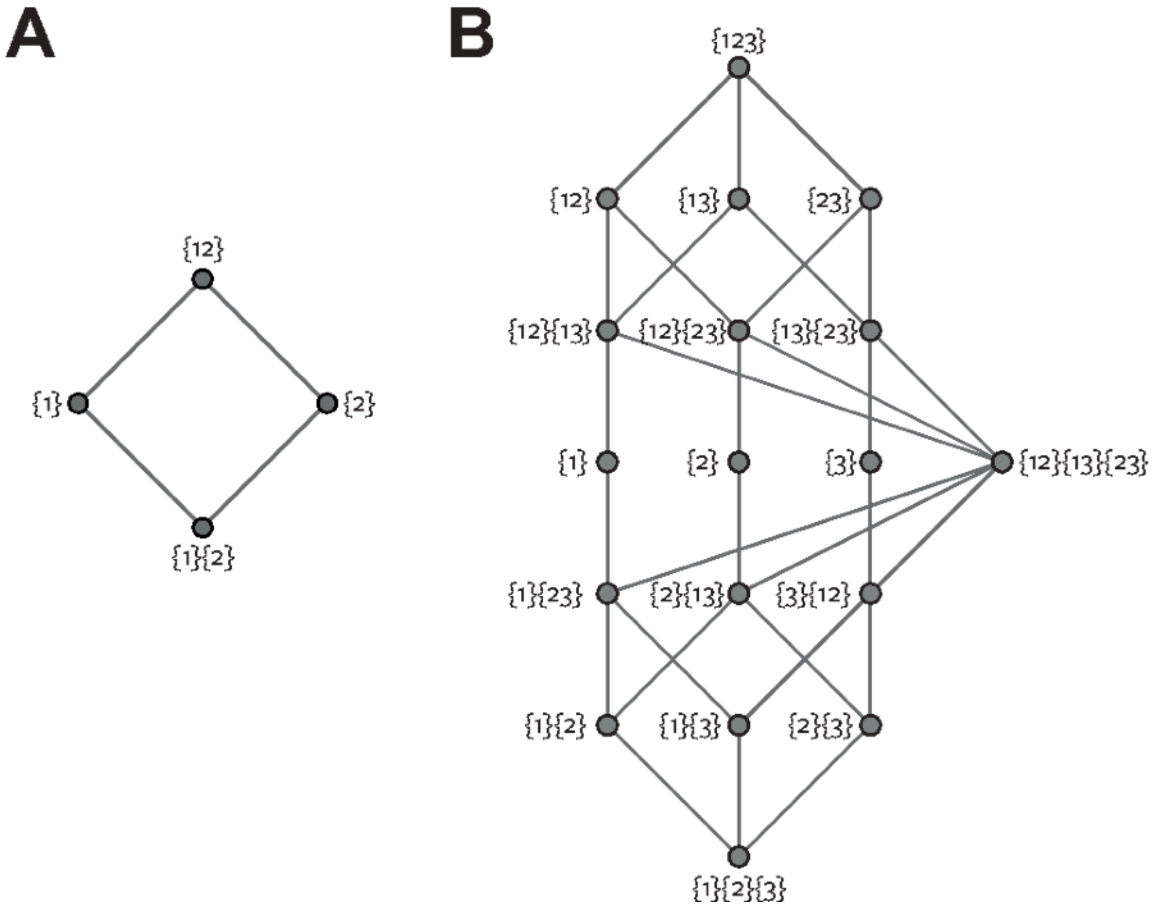
Rassouli, B., **Rosas, F.**, & Gündüz, D. (2018). Latent feature disclosure under perfect sample privacy. In 2018 IEEE International Workshop on Information Forensics and Security (WIFS) (pp. 1-7). IEEE.

Rassouli, B., **Rosas, F. E.**, & Gündüz, D. (2019). Data disclosure under perfect sample privacy. IEEE Transactions on Information Forensics and Security, 15, 2012-2025.

Rosas, F. E., Mediano, P. A., Rassouli, B., & Barrett, A. B. (2020). An operational information decomposition via synergistic disclosure. Journal of Physics A: Mathematical and Theoretical, 53(48), 485001.

Python toolbox: https://github.com/pmediano/syndisc

Making sense of synergistic effects

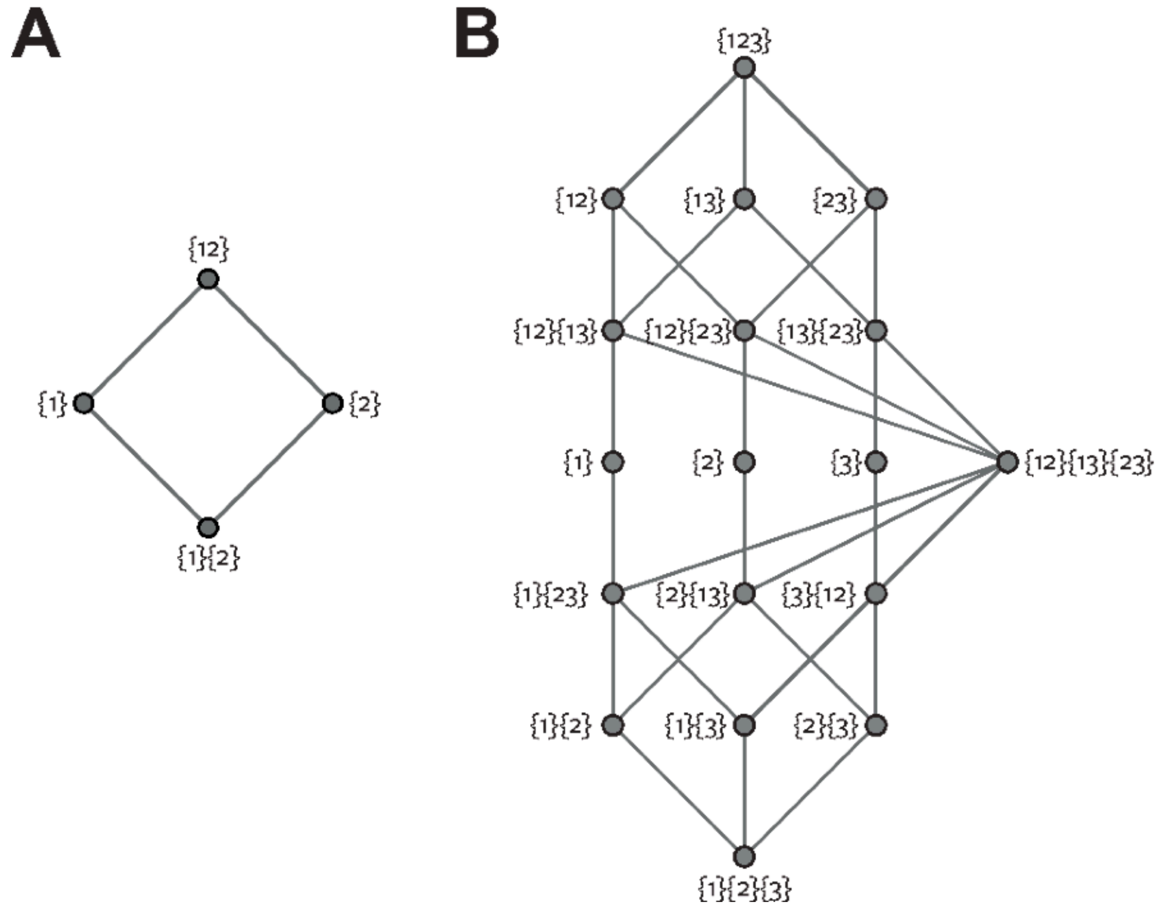


"Partial Information decomposition" (PID):

$$I(XY;Z) = \mathrm{Red}(XY;Z) + \mathrm{Un}(X;Z|Y) + \mathrm{Un}(Y;Z|X) + \mathrm{Syn}(XY;Z)$$

P. Williams and R. D. Beer. "Nonnegative decomposition of multivariate information." arXiv preprint arXiv:1004.2515 (2010).

Making sense of synergistic effects



"Partial Information decomposition" (PID):

$$I(XY;Z) = \mathrm{Red}(XY;Z) + \mathrm{Un}(X;Z|Y) + \mathrm{Un}(Y;Z|X) + \mathrm{Syn}(XY;Z)$$

Summary metric: "O-Information"

$$\begin{aligned} I(X;Y;Z) :=& I(X;Z) + I(Y;Z) - I(XY;Z) \\ =& \mathrm{Red}(XY;Z) - \mathrm{Syn}(XY;Z) \end{aligned}$$

O-information

Occam's raizor (*lex parsimoniae*): give preference to the simplest description!

Definition

O-information: $\Omega(\boldsymbol{X}^n) = \mathrm{TC}(\boldsymbol{X}^n) - \mathrm{DTC}(\boldsymbol{X}^n)$

- If $\Omega(\boldsymbol{X}^n) > 0$ it is shorter to describe the allowed states.
- If $\Omega(\boldsymbol{X}^n) < 0$ it is shorter shorter to describe the constraints.

F. E., Mediano, P. A., Gastpar, M., & Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E, 100(3), 032305.

O-information

Occam's raizor (*lex parsimoniae*): give preference to the simplest description!

Definition

O-information: $\Omega(\boldsymbol{X}^n) = \mathrm{TC}(\boldsymbol{X}^n) - \mathrm{DTC}(\boldsymbol{X}^n)$

- If $\Omega(\boldsymbol{X}^n) > 0$ it is shorter to describe the allowed states.
- If $\Omega(\boldsymbol{X}^n) < 0$ it is shorter shorter to describe the constraints.

Example:

$$X_1 = X_2 = \cdots = X_n$$
$$\mathrm{TC} = n - 1 > \mathrm{DTC} = 1$$

redundancy!

$$X_n = X_1 \oplus X_2 \oplus \cdots \oplus X_{n-1}$$
$$\mathrm{TC} = 1 < \mathrm{DTC} = n - 1$$

synergy!

O-information

Occam's raizor (*lex parsimoniae*): give preference to the simplest description!

Definition

o-information: $\Omega(\boldsymbol{X}^n) = \mathrm{TC}(\boldsymbol{X}^n) - \mathrm{DTC}(\boldsymbol{X}^n)$

- If $\Omega(\boldsymbol{X}^n) > 0$ it is shorter to describe the allowed states.
- If $\Omega(\boldsymbol{X}^n) < 0$ it is shorter shorter to describe the constraints.

Important cases:

1. **n=2:** $\Omega(X_1 X_2) = 0$, because $\mathrm{TC}(X_1 X_2) = \mathrm{DTC}(X_1 X_2) = I(X_1; X_2)$
   i.e. shared randomness is equal to predictability

O-information

Occam's raizor (*lex parsimoniae*): give preference to the simplest description!

Definition

o-information: $\Omega(\boldsymbol{X}^n) = \mathrm{TC}(\boldsymbol{X}^n) - \mathrm{DTC}(\boldsymbol{X}^n)$

- If $\Omega(\boldsymbol{X}^n) > 0$ it is shorter to describe the allowed states.
- If $\Omega(\boldsymbol{X}^n) < 0$ it is shorter shorter to describe the constraints.

Important cases:

1. **n=2:** $\Omega(X_1 X_2) = 0$, because $\mathrm{TC}(X_1 X_2) = \mathrm{DTC}(X_1 X_2) = I(X_1; X_2)$
   i.e. shared randomness is equal to predictability

2. **n=3:** $\Omega(X_1 X_2 X_3) = I(X_1; X_2) - I(X_1; X_2 | X_3) =: I(X_1; X_2; X_3)$ —> co-information!

O-information

Occam's raizor (*lex parsimoniae*): give preference to the simplest description!

Definition

o-information: $\Omega(\boldsymbol{X}^n) = \mathrm{TC}(\boldsymbol{X}^n) - \mathrm{DTC}(\boldsymbol{X}^n)$

- If $\Omega(\boldsymbol{X}^n) > 0$ it is shorter to describe the allowed states.
- If $\Omega(\boldsymbol{X}^n) < 0$ it is shorter shorter to describe the constraints.

Important cases:

1. **n=2:** $\Omega(X_1 X_2) = 0$, because $\mathrm{TC}(X_1 X_2) = \mathrm{DTC}(X_1 X_2) = I(X_1; X_2)$
   i.e. shared randomness is equal to predictability

2. **n=3:** $\Omega(X_1 X_2 X_3) = I(X_1; X_2) - I(X_1; X_2 | X_3) =: I(X_1; X_2; X_3)$ —> co-information!

3. For any **n,** $\Omega(\boldsymbol{X}^n) = \sum_{k=2}^{n-1} I(X_k; \boldsymbol{X}^{k-1}; \boldsymbol{X}_{k+1}^n)$ —> redundancy minus synergy!

O-information

Occam's raizor (*lex parsimoniae*): give preference to the simplest description!

Definition

o-information: $\Omega(\boldsymbol{X}^n) = \mathrm{TC}(\boldsymbol{X}^n) - \mathrm{DTC}(\boldsymbol{X}^n)$

- If $\Omega(\boldsymbol{X}^n) > 0$ it is shorter to describe the allowed states.
- If $\Omega(\boldsymbol{X}^n) < 0$ it is shorter shorter to describe the constraints.

<u>Important cases:</u>

1. **n=2:** $\Omega(X_1 X_2) = 0$, because $\mathrm{TC}(X_1 X_2) = \mathrm{DTC}(X_1 X_2) = I(X_1; X_2)$
   i.e. shared randomness is equal to predictability

2. **n=3:** $\Omega(X_1 X_2 X_3) = I(X_1; X_2) - I(X_1; X_2 | X_3) =: I(X_1; X_2; X_3)$ —> co-information!

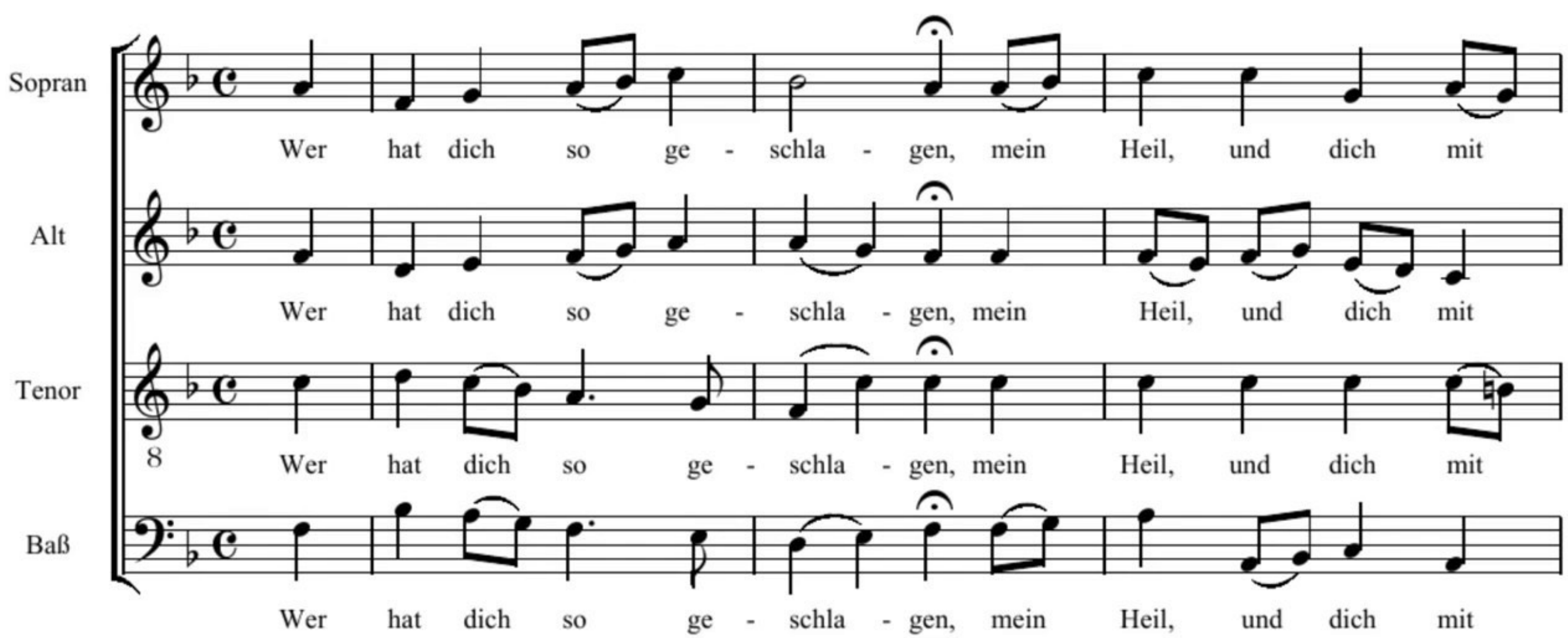
3. For any **n**, $\Omega(\boldsymbol{X}^n) = \sum_{k=2}^{n-1} I(X_k; \boldsymbol{X}^{k-1}; \boldsymbol{X}_{k+1}^n)$ —> redundancy minus synergy!

4. It is maximised by copied bits, and minimised by XoRs.

Case study on baroque music scores

Data analysis over music scores from the Baroque period (*Python, Music21 package*)

i) chorales for four voices by **J.S. Bach** (1685–1750)



F. E., Mediano, P. A., Gastpar, M., & Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E, 100(3), 032305.

Case study on baroque music scores

Data analysis over music scores from the Baroque period (*Python, Music21 package*)

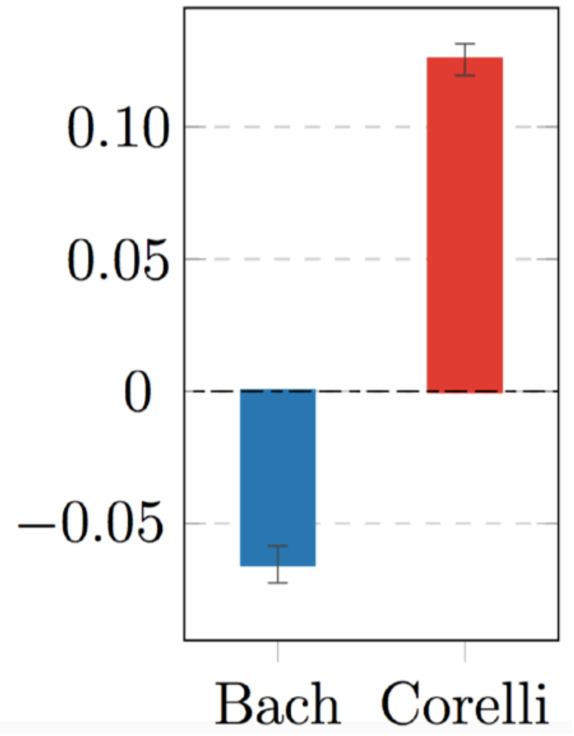
i) chorales for four voices by **J.S. Bach** (1685–1750)
(43k four-note chords)

ii) Op. 1, 3, 4, 5 and 6 of **A. Corelli** (1653–1713)
(80k four-note chords)

F. E., Mediano, P. A., Gastpar, M., & Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E, 100(3), 032305.

Case study on baroque music scores

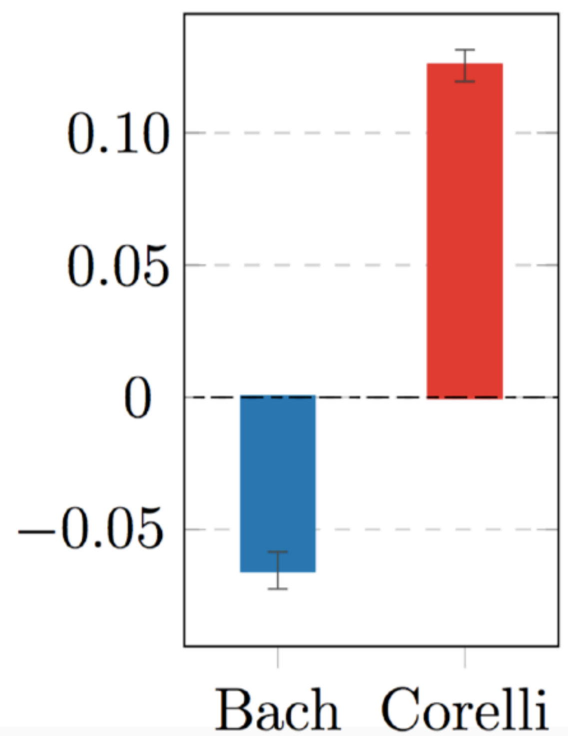
Data analysis over music scores from the Baroque period (*Python, Music21 package*)

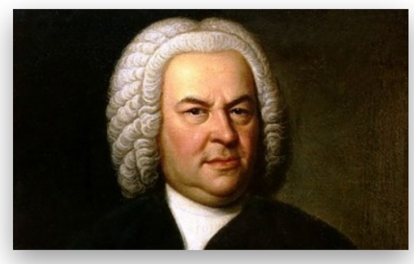


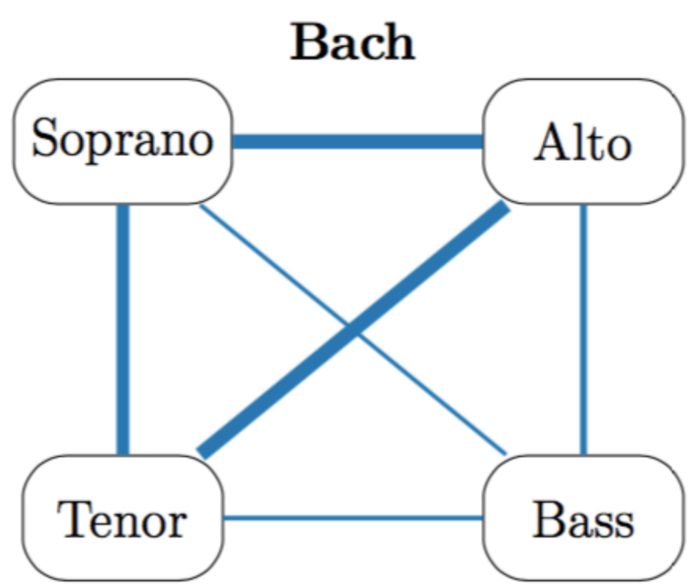
F. E., Mediano, P. A., Gastpar, M., & Jensen, H. J. (2019). Quantifying high-order interdependencies via multivariate extensions of the mutual information. Physical Review E, 100(3), 032305.

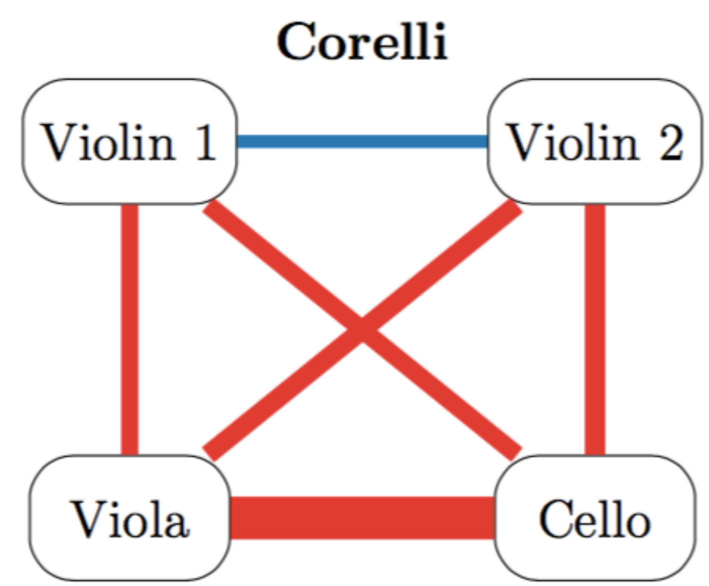
Case study on baroque music scores

Data analysis over music scores from the Baroque period (*Python, Music21 package*)







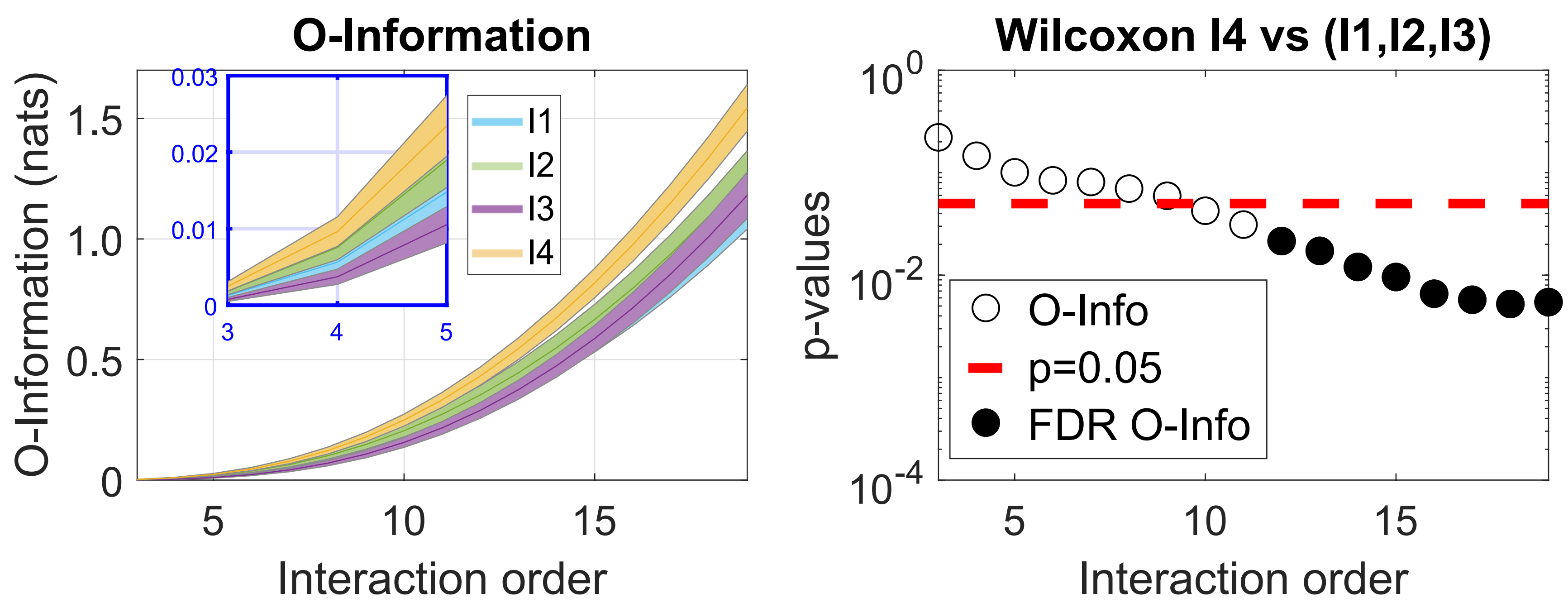


Synergy allows the coexistence of
local independency and **global harmony**

Case study on human resting-state fMRI data

We studied fMRI scans of 164 healthy volunteers divided in four groups

I1 (10-20), I2 (20-40), I3 (40-60), I4 (60-80).



Main finding:

brain loose its integration-differentiation balance with age, tending towards redundant interdependencies

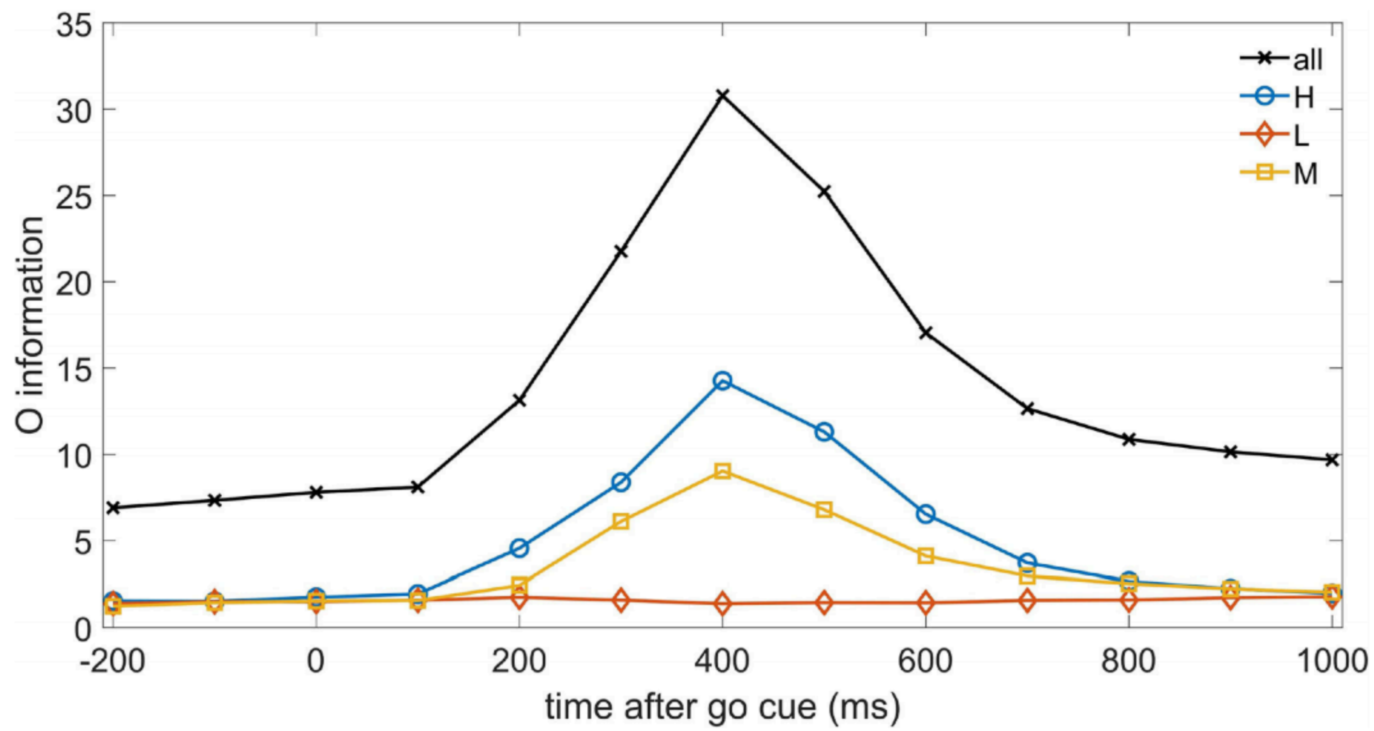
Gatica, M., Cofré, R., Mediano, P. A., Rosas, F. E., Orio, P., Diez, I., ... & Cortes, J. M. (2021). High-order interdependencies in the aging brain. Brain connectivity, 11(9), 734-744.

A dynamic O-information (dO-info)

One can build a "transfer entropy-like" O-information as follows:

$$d\Omega_n = (1-n)I(Y;\mathbf{X}|Y_0) + \sum_{j=1}^{n} I(Y;\mathbf{X}\setminus X_j|Y_0)$$

It can effectively discriminate between different types of macaque's neurons involved in a decision task.



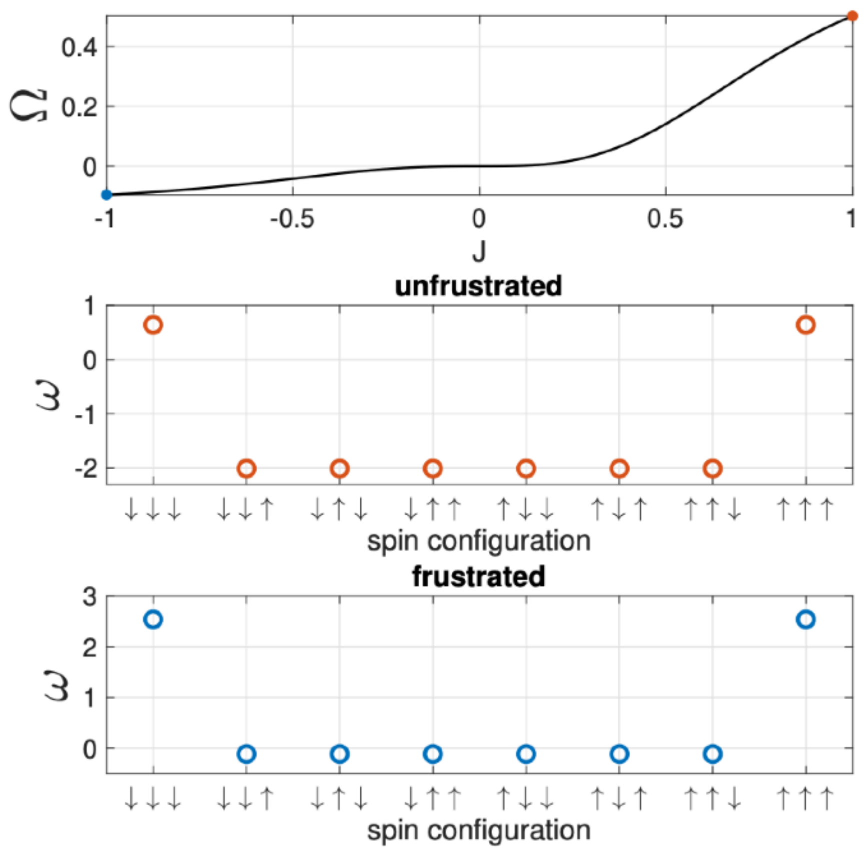
Stramaglia, S., Scagliarini, T., Daniels, B. C., & Marinazzo, D. (2021). Quantifying dynamical high-order interdependencies from the o-information: an application to neural spiking dynamics. Frontiers in Physiology, 11, 1784.

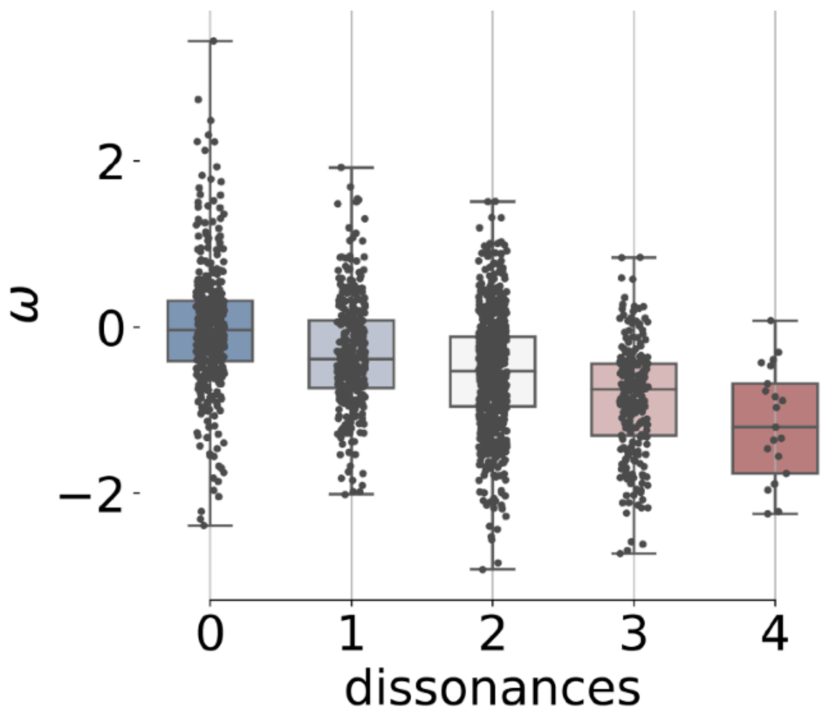
Measuring the synergy-vs-redundancy per pattern

Introducing a pattern-wise O-information:

$$\omega(\boldsymbol{X}^n) := (n-2)h(\boldsymbol{X}^n) + \sum_{j=1}^{n} \Big(h(x_j) - h(\boldsymbol{x}^n_{-j}) \Big).$$

Frustration in spin systems

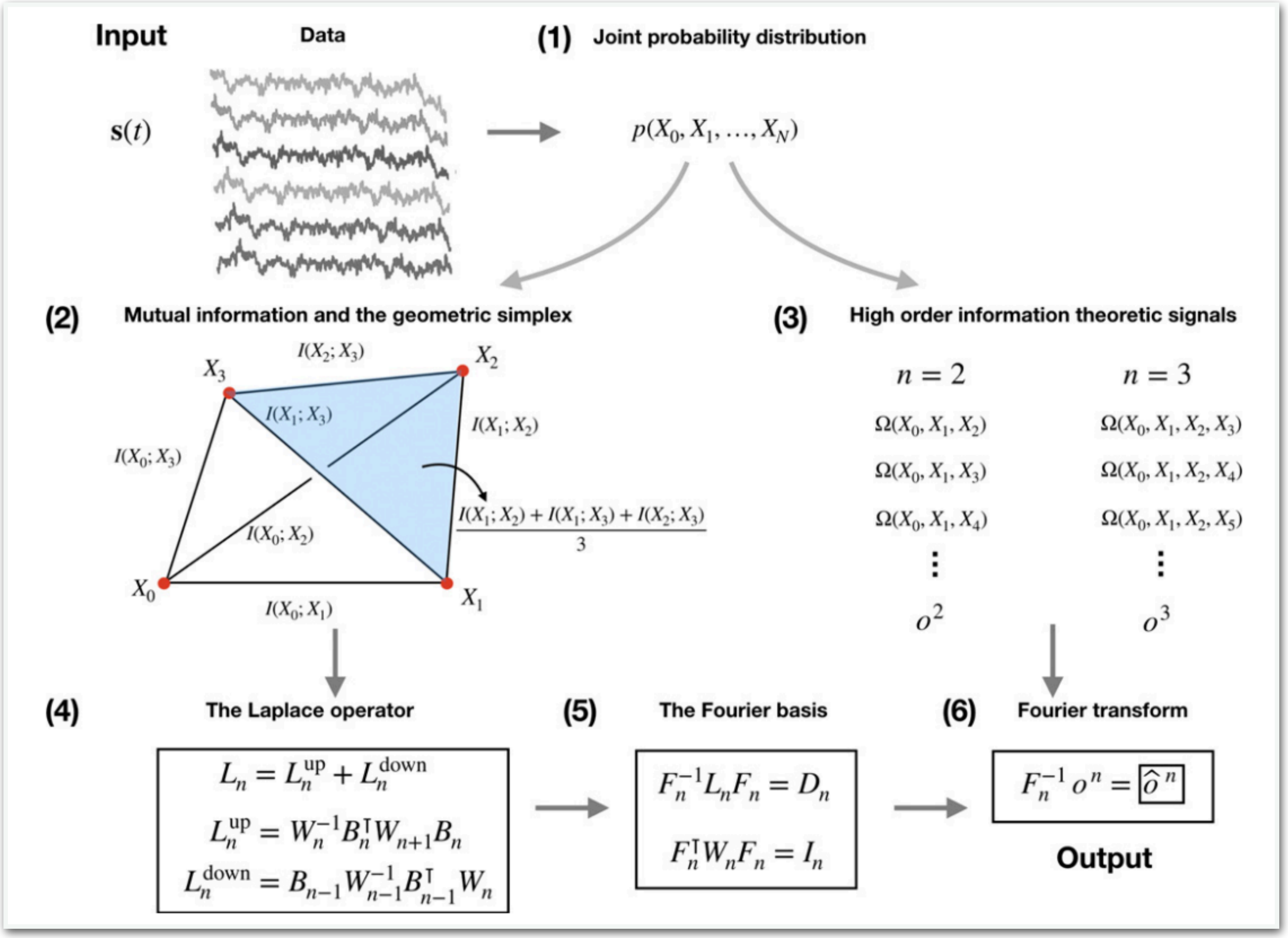




Scagliarini, T., Marinazzo, D., Guo, Y., Stramaglia, S., & Rosas, F. E. (2022). Quantifying high-order interdependencies on individual patterns via the local O-information: theory and applications to music analysis. Physical Review Research, 4(1), 013184.

Hyperharmonic decomposition of O-information

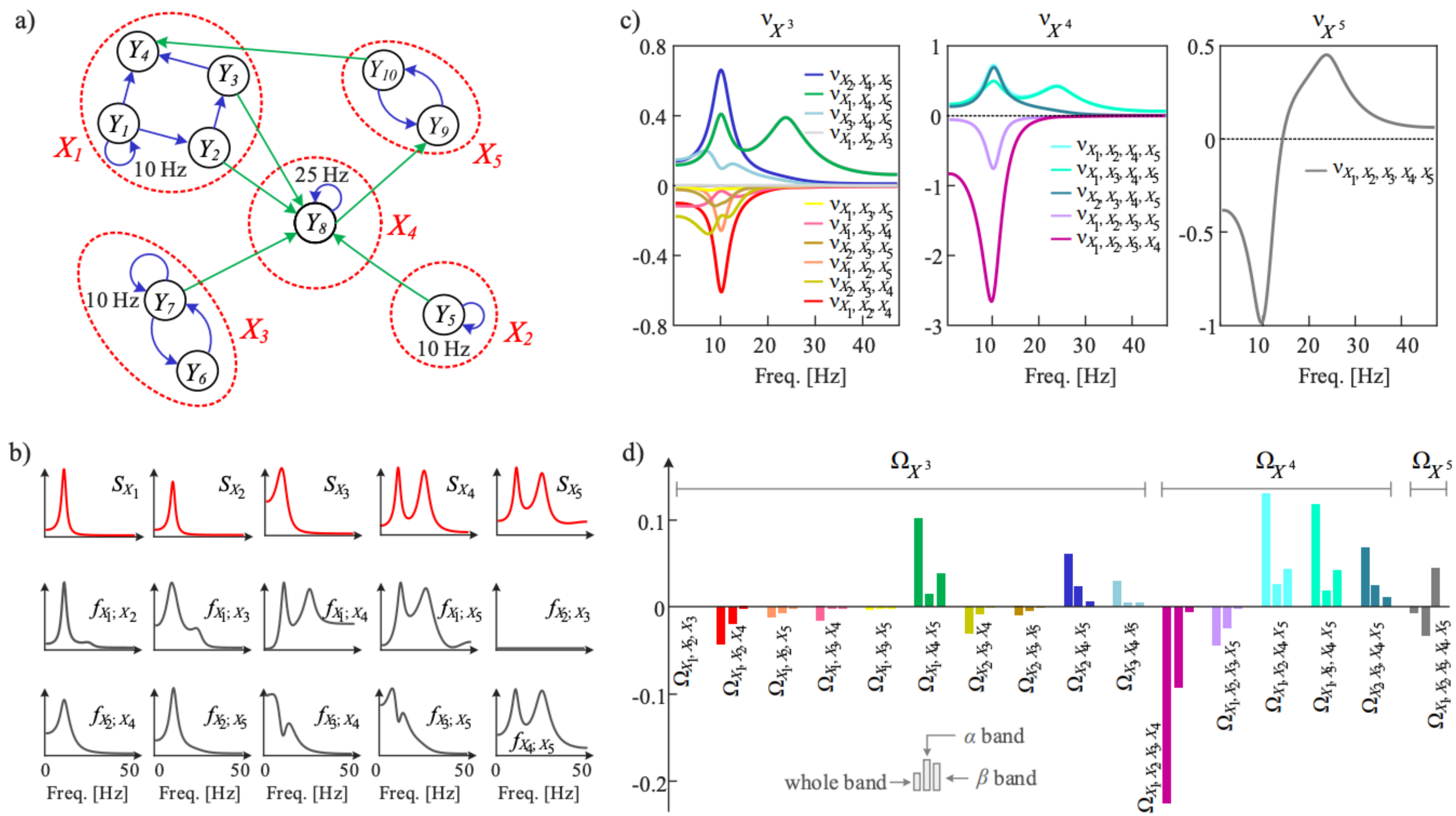
Introducing Fourier analysis over hyper-graphs!!



Medina, A. M., Rosas, F. E., Rodríguez, S. E., & Cofré, R. (2021). Hyperharmonic analysis for the study of high-order information-theoretic signals. Journal of Physics: Complexity, 2(3), 035009.

Spectral decomposition of the O-information rate

Leveraging the spectral representation of vector autoregressive and state-space models.



Faes, L., Mijatovic, G., Antonacci, Y., Pernice, R., Barà, C., Sparacino, L., ... & Stramaglia, S. (2022). A Framework for the Time-and Frequency-Domain Assessment of High-Order Interactions in Brain and Physiological Networks. arXiv preprint arXiv:2202.04179.

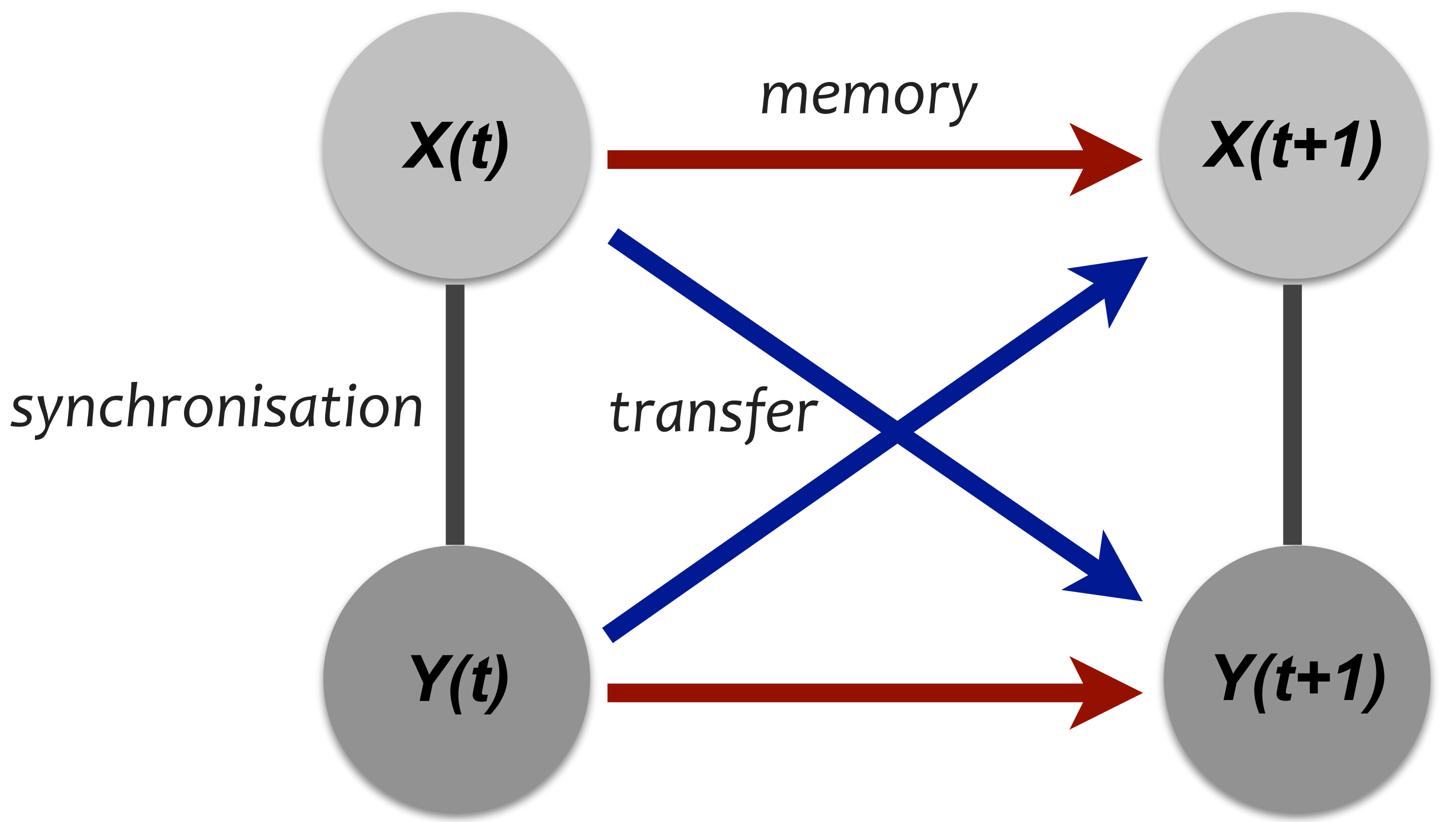
Today's menu
1. Synergy
2. Emergence
3. Ideas to take home

Emergence: what is it, and why it matters

Dynamical synergy?

Traditionally, complex dynamics are measured in terms of:

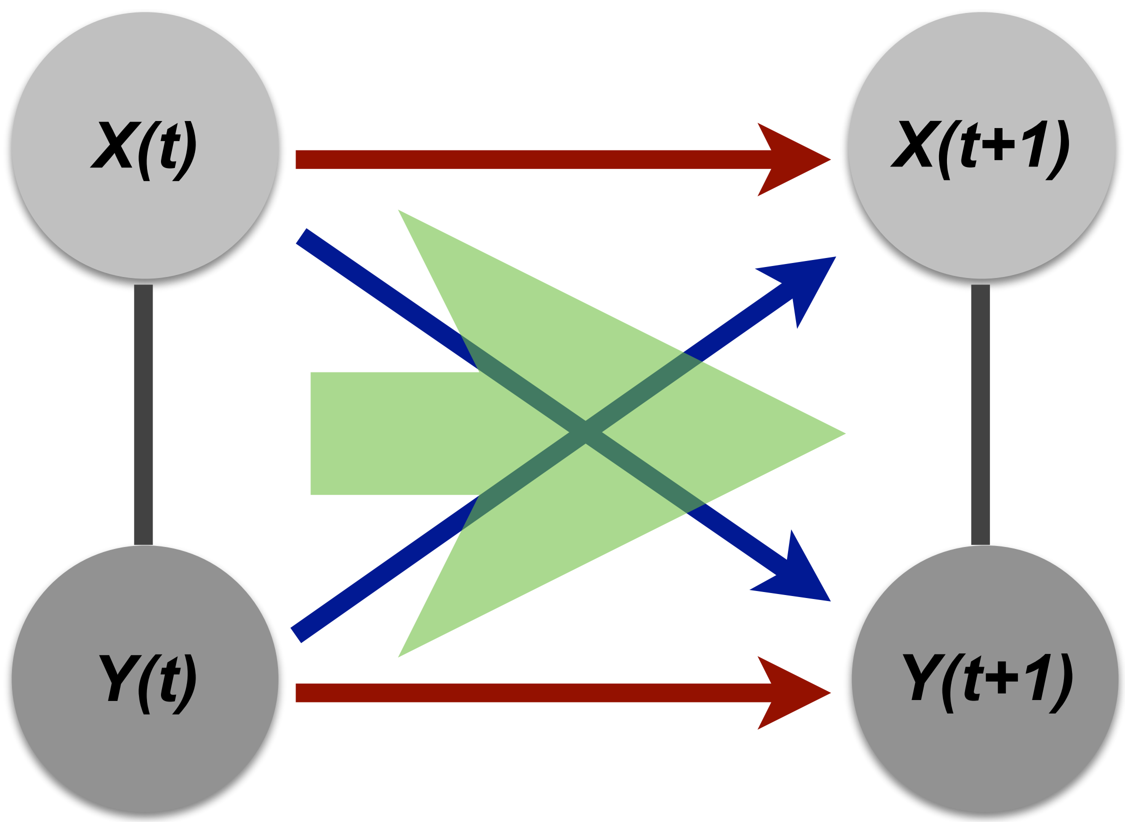
- **memory**: the persistence of information within one agent
- **transfer**: the co-influence of various agents



Dynamical synergy?

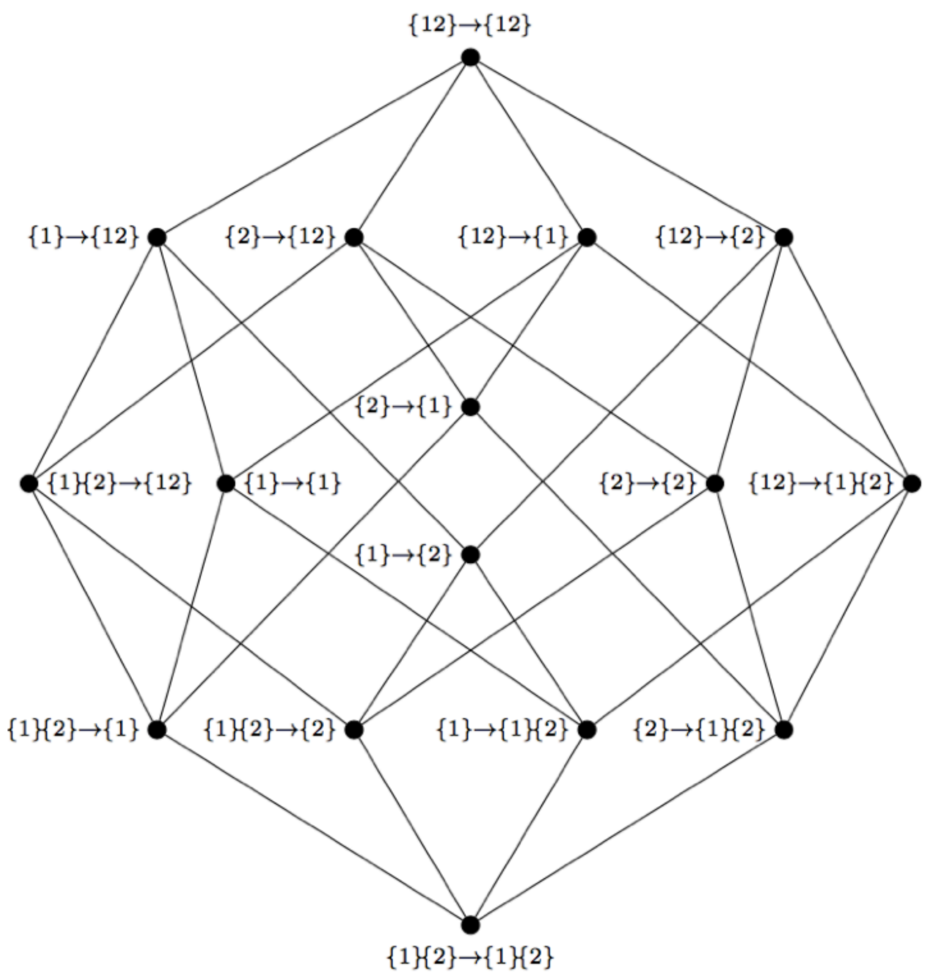
Traditionally, complex dynamics are measured in terms of:

- **memory**: the persistence of information within one agent
- **transfer**: the co-influence of various agents



… what about the dynamical high-order interactions ??

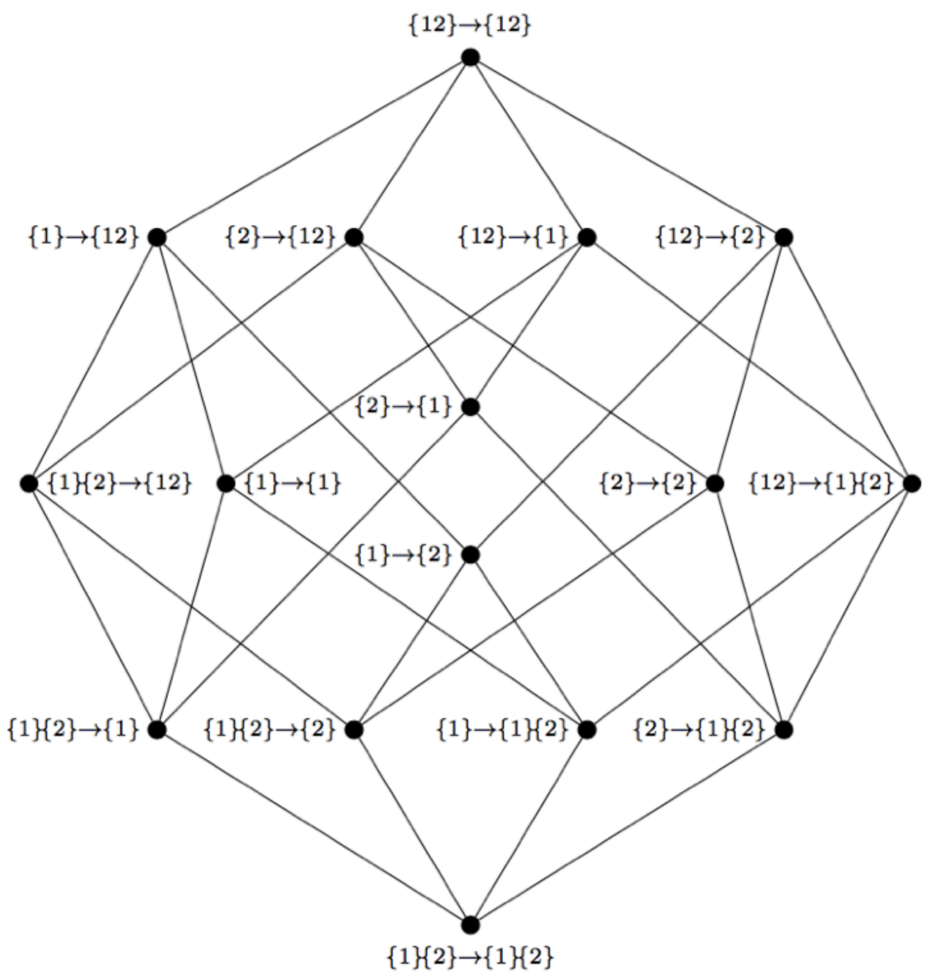
Decomposing dynamical synergy via ΦID



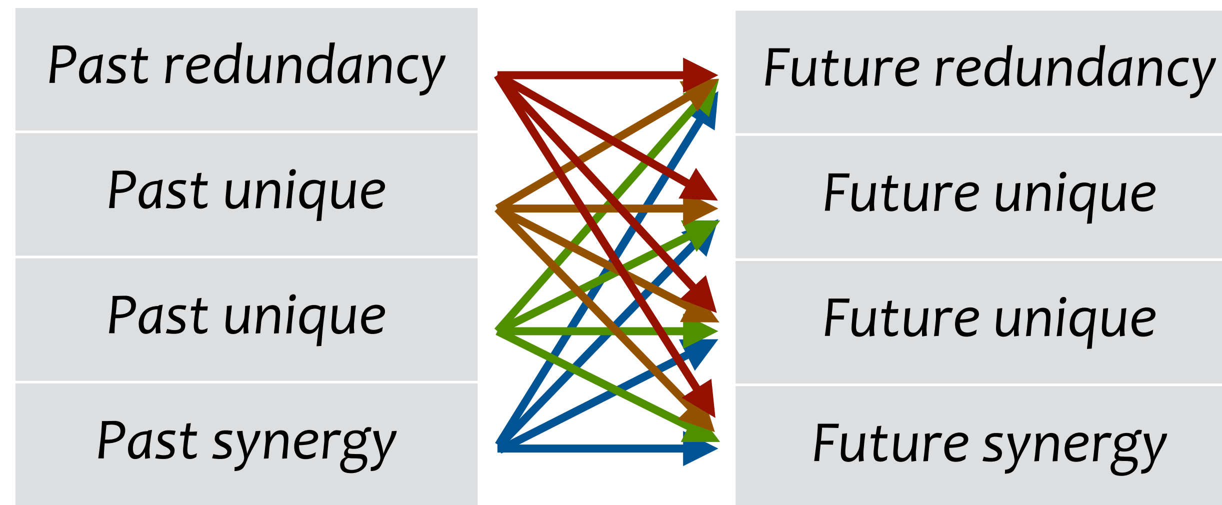
$$I(\boldsymbol{X}_t;\boldsymbol{X}_{t'}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{D}} I_\partial^{\boldsymbol{\alpha}\to\boldsymbol{\beta}}$$

Mediano, P. A., **Rosas, F.**, Carhart-Harris, R. L., Seth, A. K., & Barrett, A. B. (2019). Beyond integrated information: A taxonomy of information dynamics phenomena. arXiv preprint arXiv:1909.02297.

Decomposing dynamical synergy via ΦID



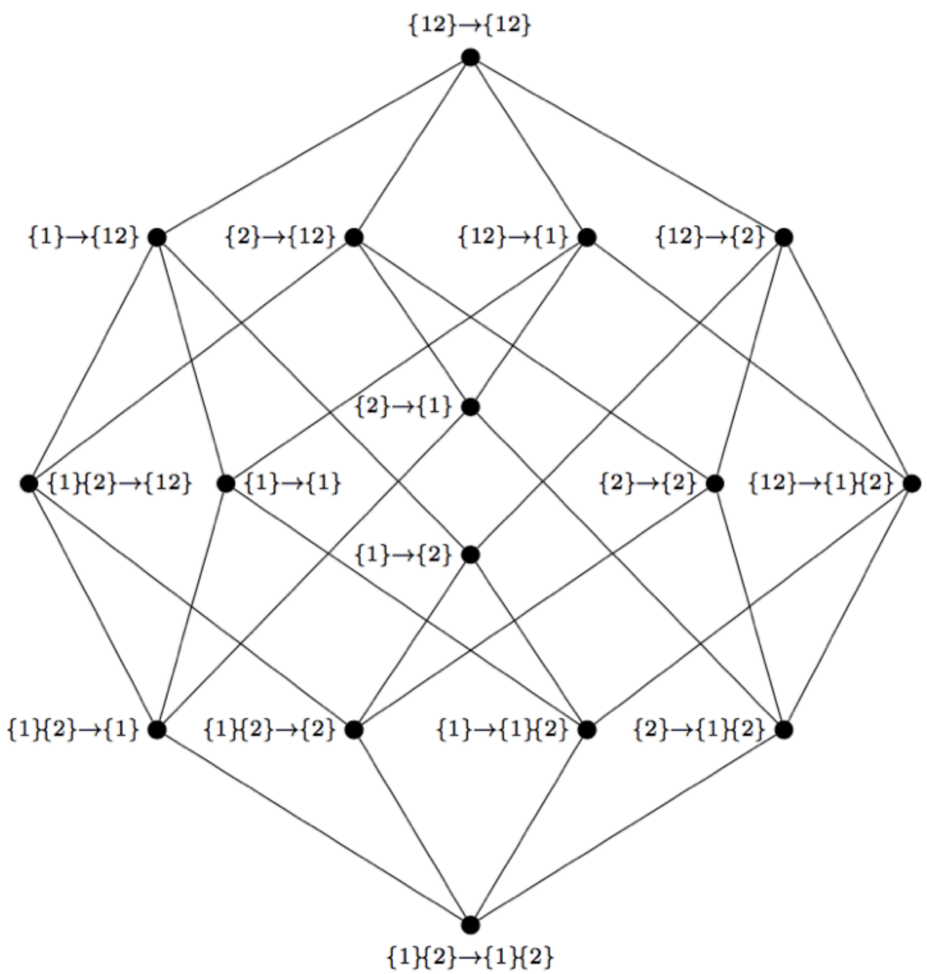
Forward X Backwards PIDs —> 16 terms



$$I(\boldsymbol{X}_t; \boldsymbol{X}_{t'}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathcal{D}} I_\partial^{\boldsymbol{\alpha} \to \boldsymbol{\beta}}$$

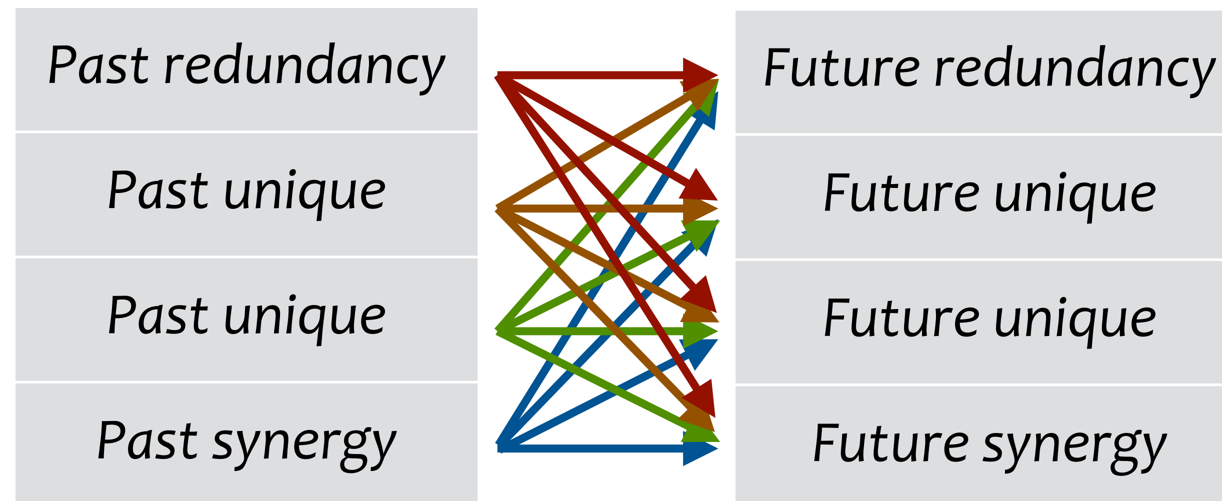
Mediano, P. A., **Rosas, F.**, Carhart-Harris, R. L., Seth, A. K., & Barrett, A. B. (2019). Beyond integrated information: A taxonomy of information dynamics phenomena. arXiv preprint arXiv:1909.02297.

Decomposing dynamical synergy via ΦID



$$I(\boldsymbol{X}_t;\boldsymbol{X}_{t'}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{D}} I_\partial^{\boldsymbol{\alpha}\to\boldsymbol{\beta}}$$

Forward X Backwards PIDs —> 16 terms



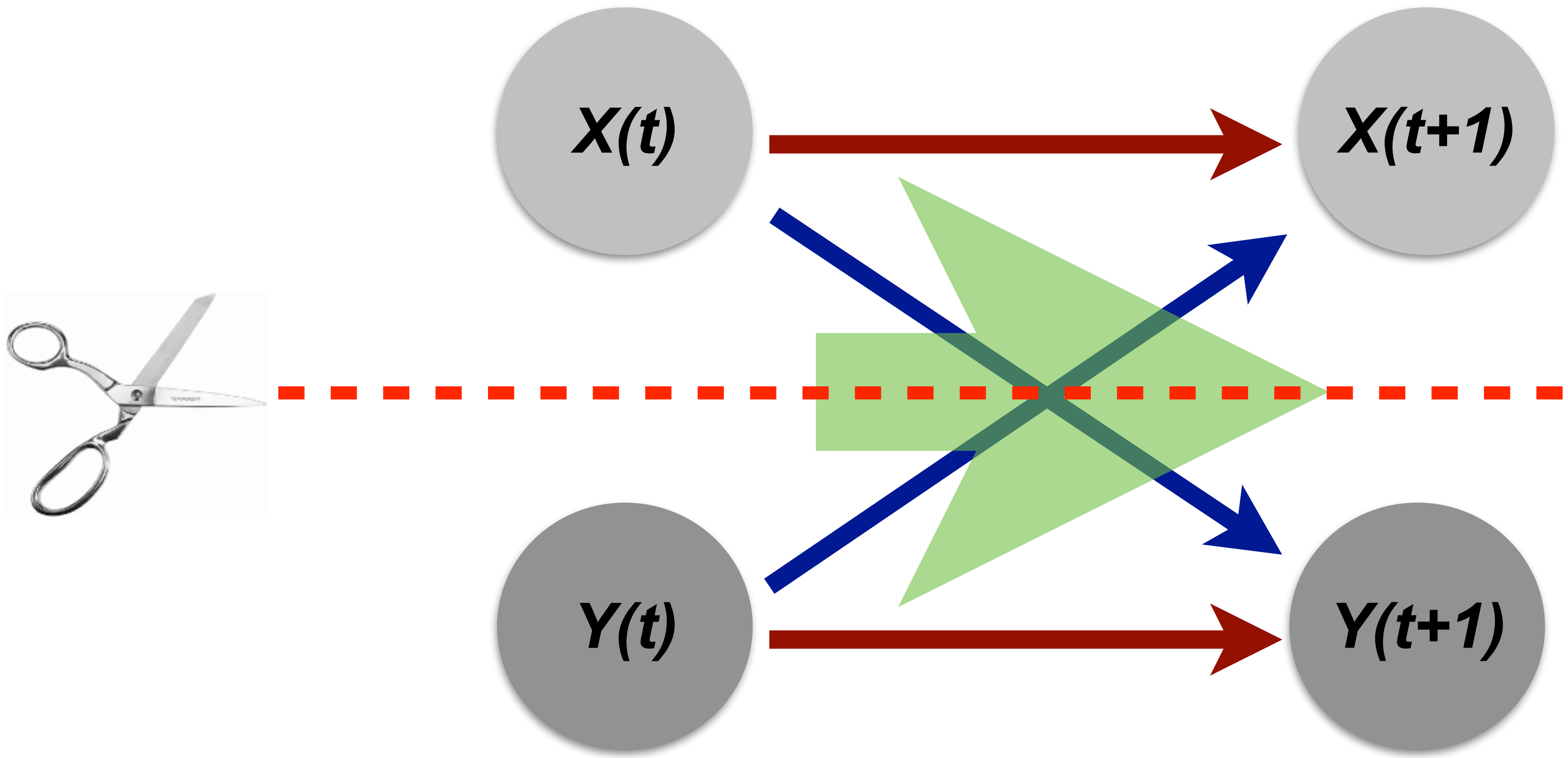
Fine **taxonomy** of *modes* of information dynamics:

- ***Storage*****:** stuff stays the way it is
- ***Copy:*** contents are duplicated
- **Transfer:** stuff moves
- **Erasure:** duplicated information is pruned
- **Downward causation:** collective properties affect individual futures

Mediano, P. A., **Rosas, F.**, Carhart-Harris, R. L., Seth, A. K., & Barrett, A. B. (2019). Beyond integrated information: A taxonomy of information dynamics phenomena. arXiv preprint arXiv:1909.02297.

Refining integrated information: Phi-R

ΦID allow us to decompose existent metrics into their constituents.

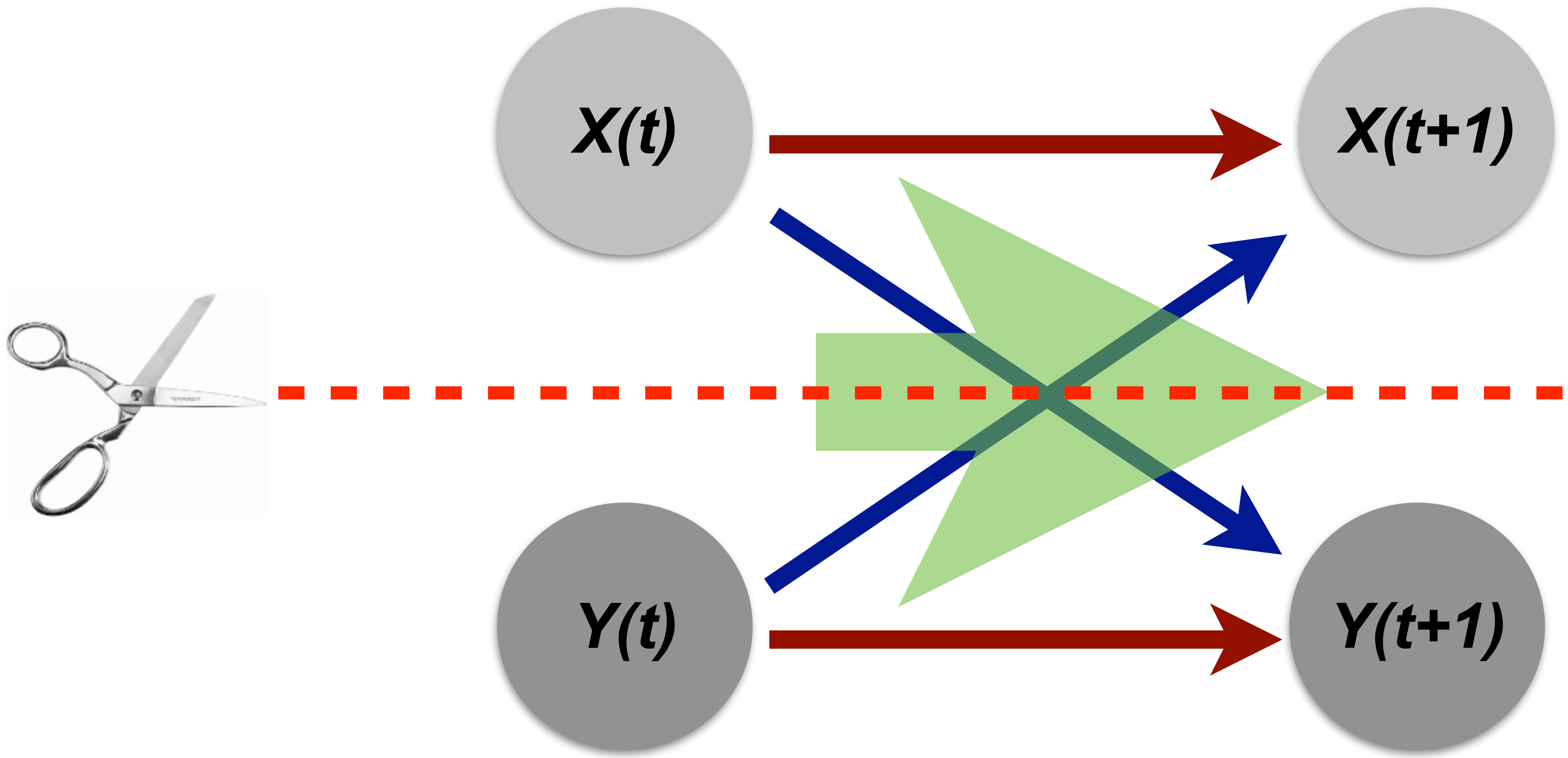


$$\Phi = I(X_t, Y_t; X_{t+1}, Y_{t+1}) - I(X_t, ; X_{t+1}) - I(Y_t, ; Y_{t+1})$$

Balduzzi D, Tononi G. Integrated information in discrete dynamical systems: motivation and theoretical framework. PLoS Comput Biol. 2008 Jun 13;4(6):e1000091.

Refining integrated information: Phi-R

ΦID allow us to decompose existent metrics into their constituents.



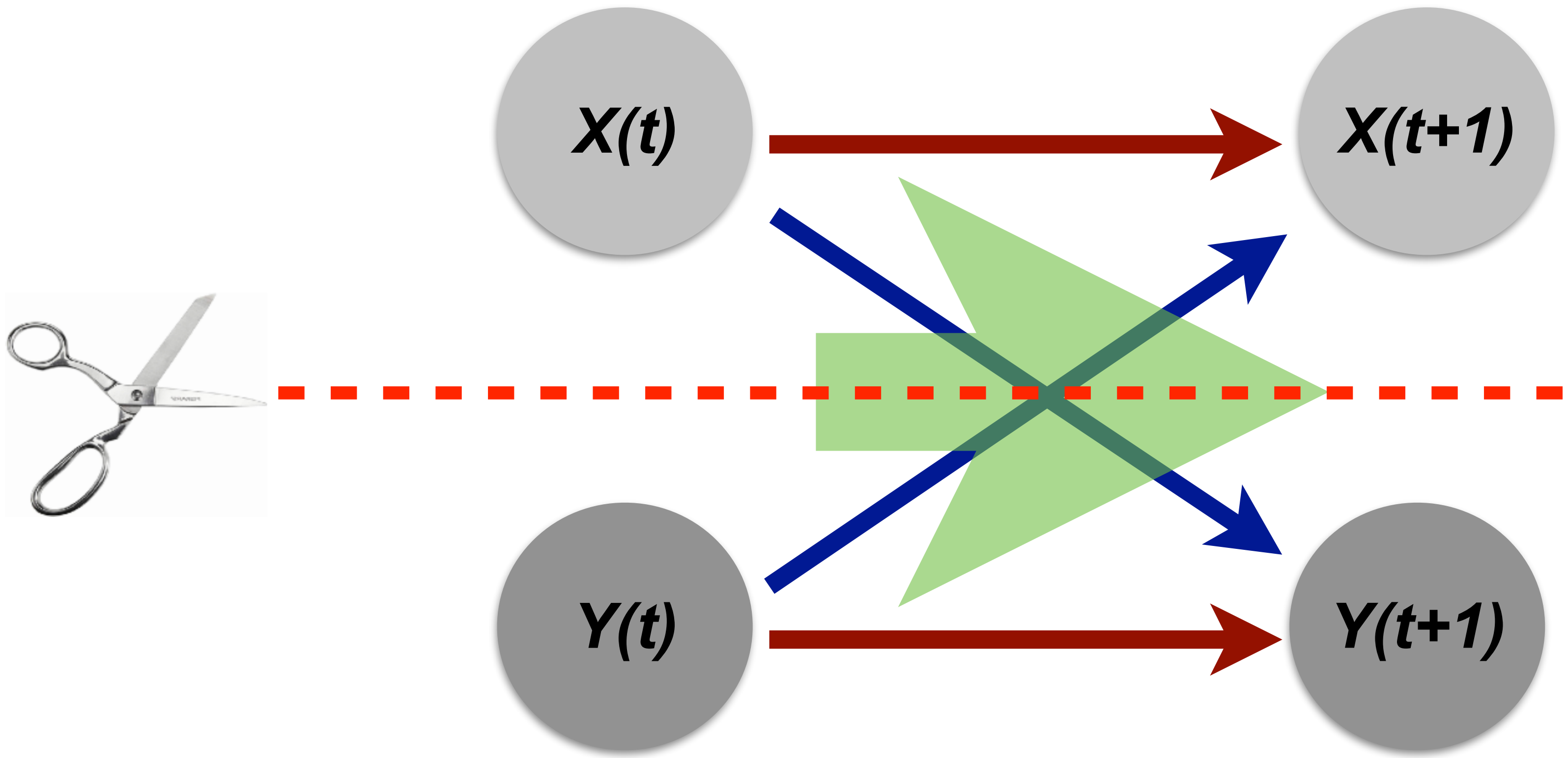
$$\Phi = I(X_t, Y_t; X_{t+1}, Y_{t+1}) - I(X_t, ; X_{t+1}) - I(Y_t, ; Y_{t+1})$$

$=$ Transfer + Synergy - Redundancy

Balduzzi D, Tononi G. Integrated information in discrete dynamical systems: motivation and theoretical framework. PLoS Comput Biol. 2008 Jun 13;4(6):e1000091.

Refining integrated information: Phi-R

ΦID allow us to decompose existent metrics into their constituents.



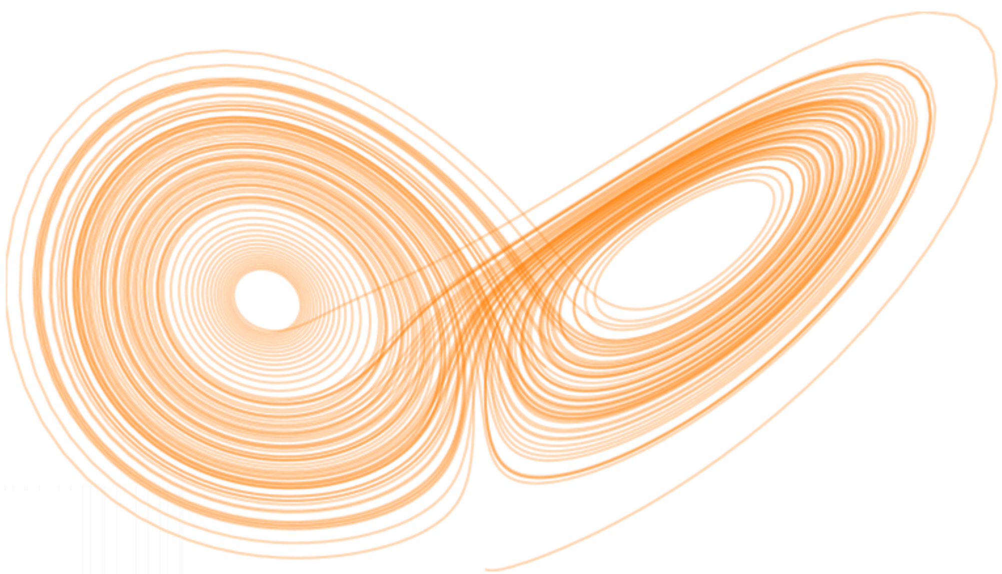
$$\Phi = I(X_t, Y_t; X_{t+1}, Y_{t+1}) - I(X_t, ; X_{t+1}) - I(Y_t, ; Y_{t+1})$$

$=$ Transfer + Synergy - Redundancy

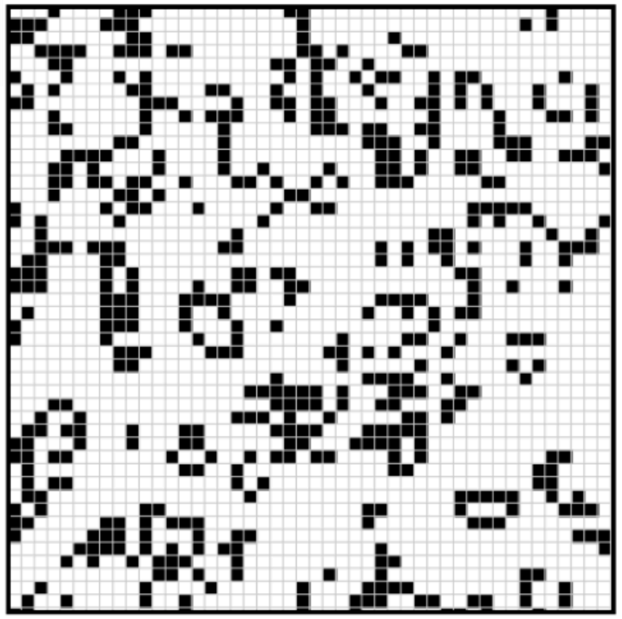
ΦID also let us introduce novel metrics tailored to specific components.

$\Phi_R = \Phi$ + Redundancy $=$ Transfer + Synergy

Phi-R as a transversal measure of dynamical complexity



Continuous dynamic systems

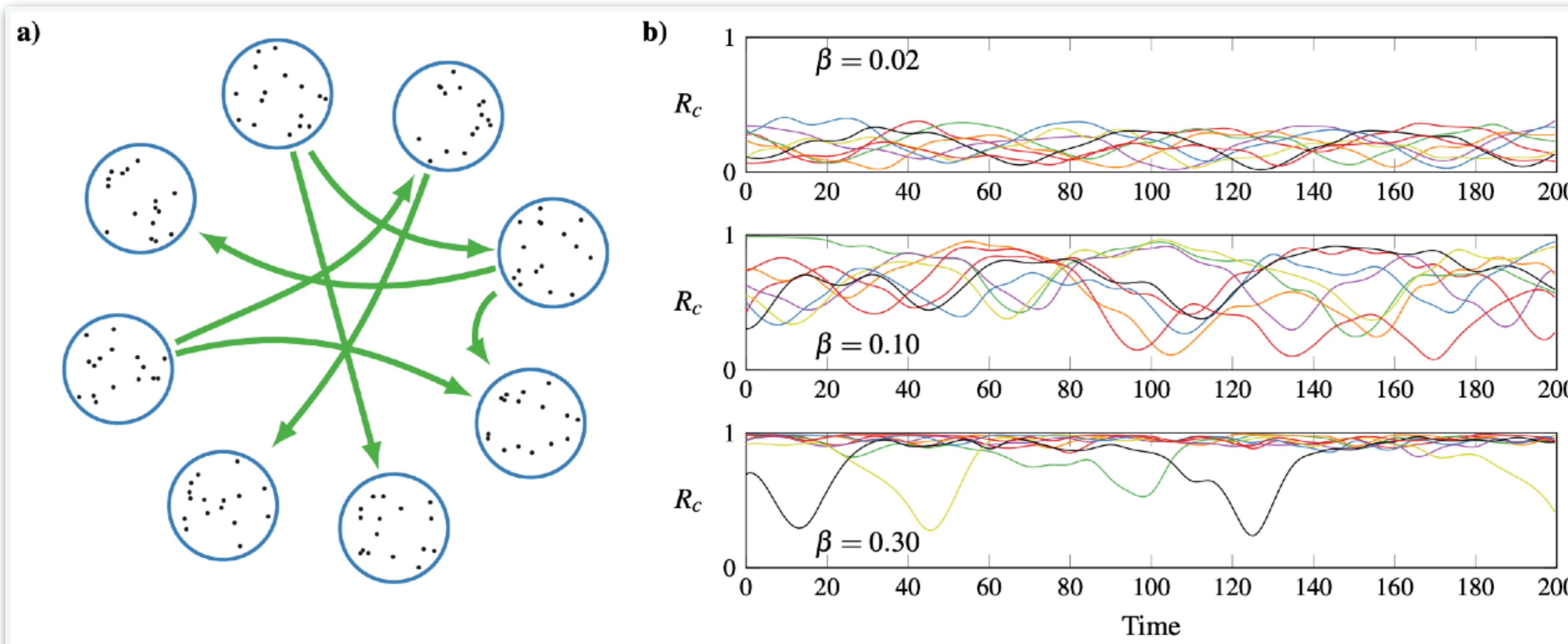


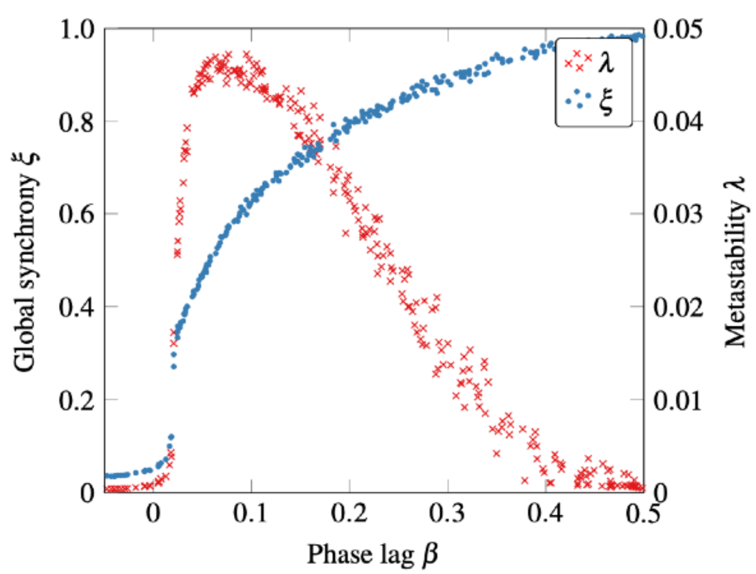
Discrete computational systems

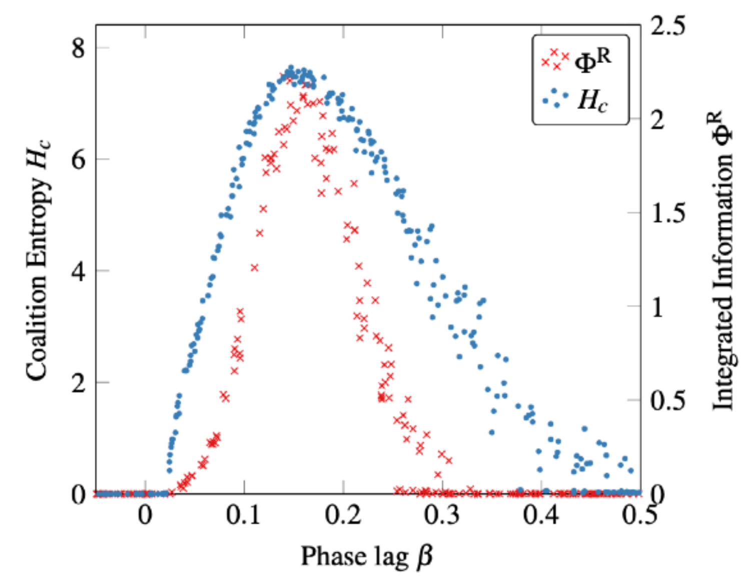
Mediano, P. A., Rosas, F. E., Farah, J. C., Shanahan, M., Bor, D., & Barrett, A. B. (2022). Integrated information as a common signature of dynamical and information-processing complexity. Chaos: An Interdisciplinary Journal of Nonlinear Science, 32(1), 013115.

Phi-R as a transversal measure of dynamical complexity

Phi-R detects phase transitions in systems of coupled oscillators



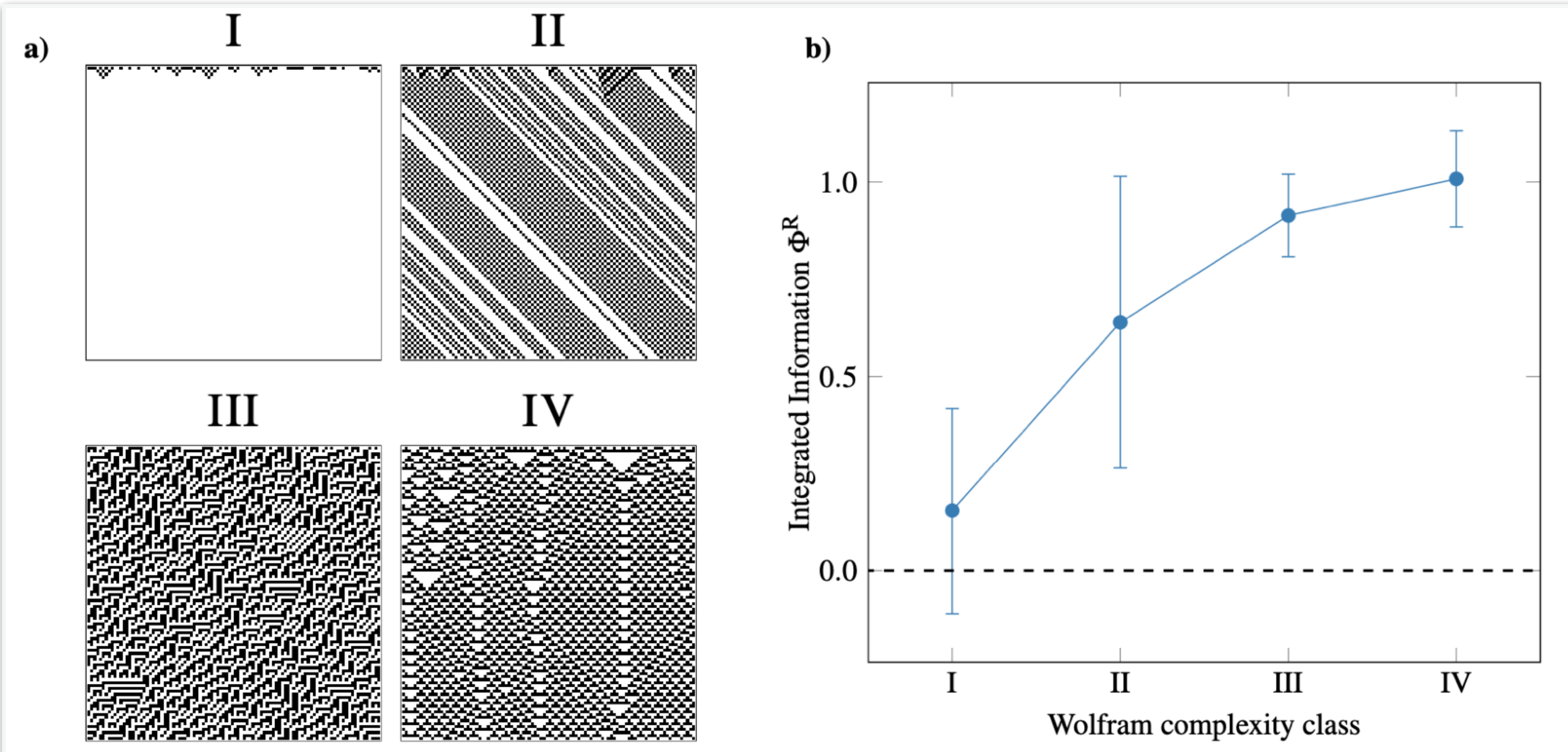




Mediano, P. A., Rosas, F. E., Farah, J. C., Shanahan, M., Bor, D., & Barrett, A. B. (2022). Integrated information as a common signature of dynamical and information-processing complexity. Chaos: An Interdisciplinary Journal of Nonlinear Science, 32(1), 013115.

Phi-R as a transversal measure of dynamical complexity

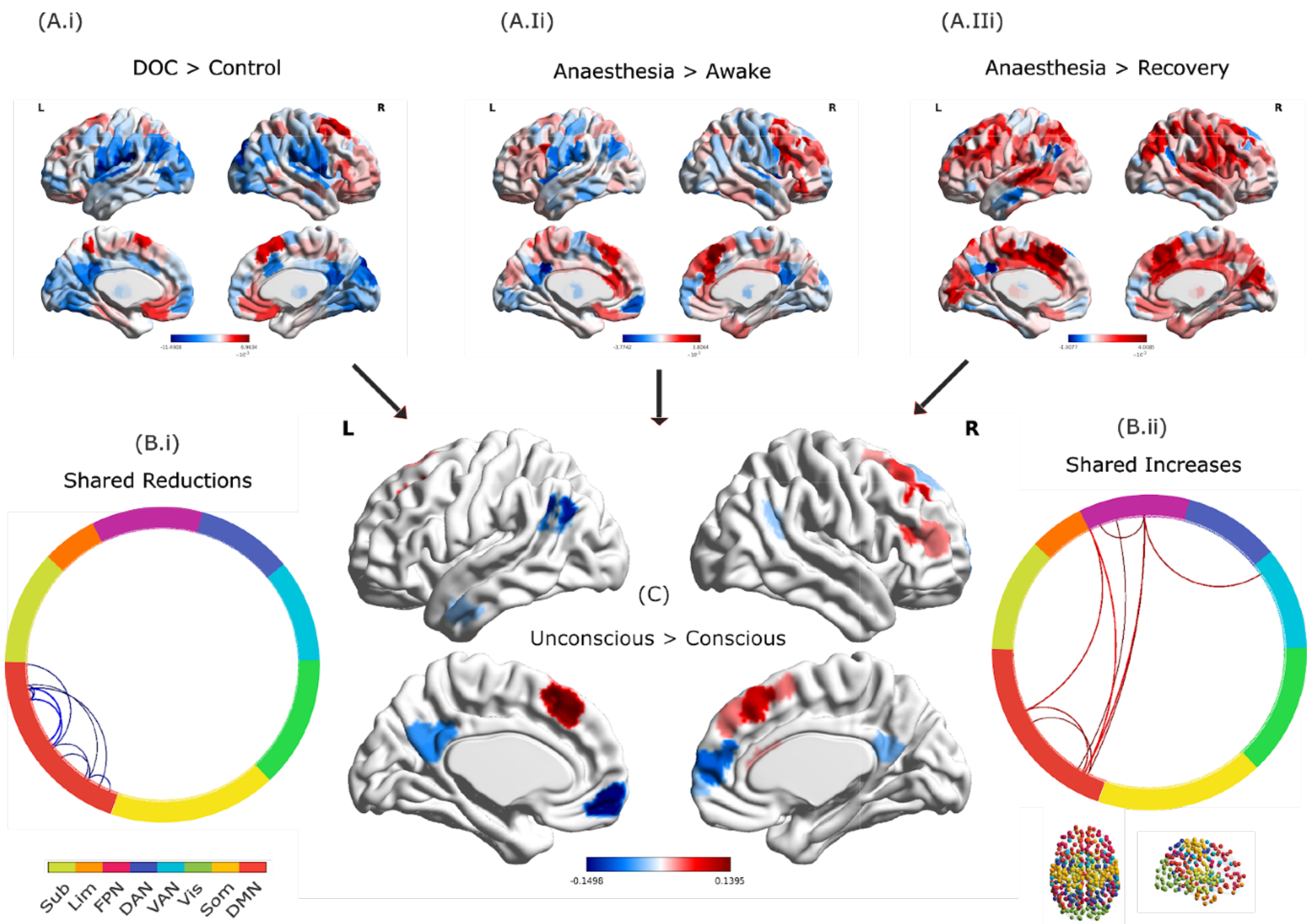
Phi-R increases with the class of computational complexity of elementary cellular automata



Mediano, P. A., Rosas, F. E., Farah, J. C., Shanahan, M., Bor, D., & Barrett, A. B. (2022). Integrated information as a common signature of dynamical and information-processing complexity. Chaos: An Interdisciplinary Journal of Nonlinear Science, 32(1), 013115.

Loss of consciousness disrupts DMN and FPN

Integrated information is most affected in default mode network and fronts-parietal areas.



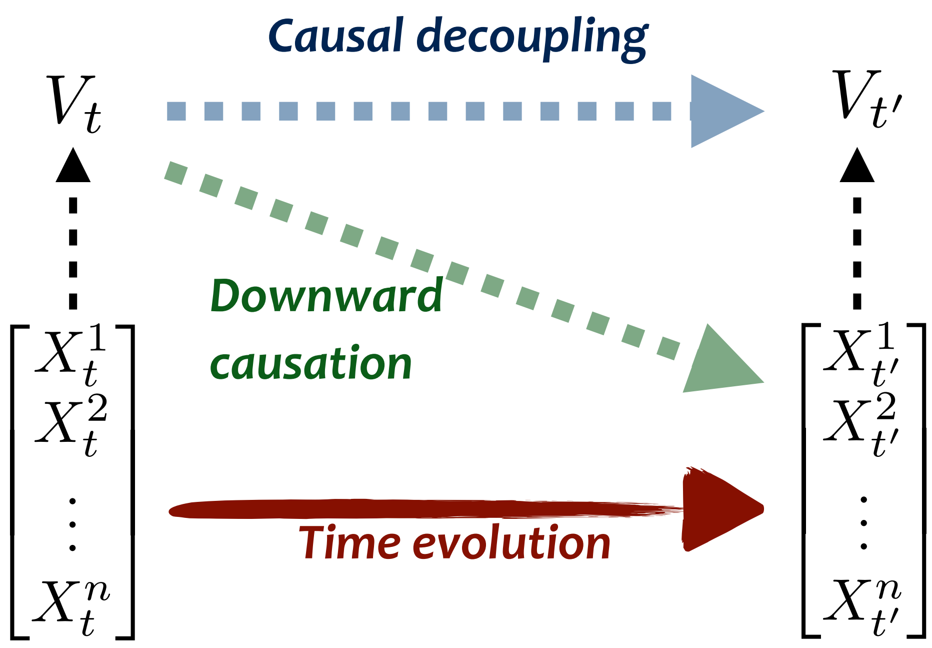
Luppi AI, Mediano PA, **Rosas FE**, Allanson J, Pickard JD, Carhart-Harris RL, Williams GB, Craig MM, Finoia P, Owen AM, Naci L. *A Synergistic Workspace for Human Consciousness Revealed by Integrated Information Decomposition*. bioRxiv. 2020 Jan 1.

How is dynamical synergy related with emergence?

What is emergence?

Scenario: a system composed by *n* sub-units (agents): $\boldsymbol{X}_t = (X_t^1, \ldots, X_t^n)$

Supervenient variable that is candidate of emergence: $V_t = F(\boldsymbol{X}_t^n)$



Informal definition: an emergent macro-variable can explain future stuff that single elements of the micro cannot.

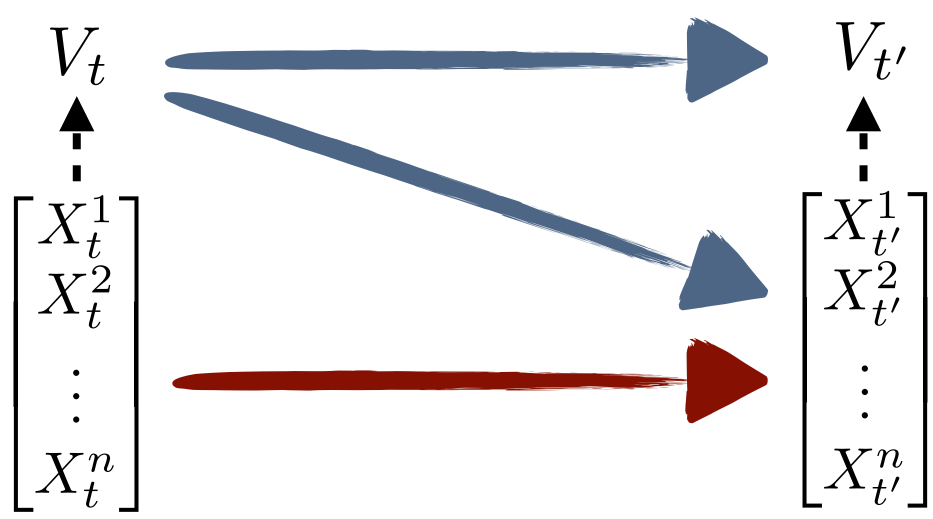
Two flavours of emergence: **downward causation** and **causal decoupling**.

What is emergence?

Scenario: a system composed by *n* sub-units (agents): $\boldsymbol{X}_t = (X_t^1, \ldots, X_t^n)$

Supervenient variable that is candidate of emergence: $V_t = F(\boldsymbol{X}_t^n)$

Definition (causal emergence):
A supervenient variable V_t exhibits *causal emergence* with respect to $\boldsymbol{X}_t$ if $\mathtt{Un}(V_t; \boldsymbol{X}_{t'} | X_t^1, \ldots, X_t^n) > 0$.



Rosas FE, Mediano PA, Jensen HJ, Seth AK, Barrett AB, Carhart-Harris RL, Bor D. *Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data.* PLOS Computational Biology. 2020 Dec 21;16(12):e1008289.

What is emergence?

Scenario: a system composed by *n* sub-units (agents): $\boldsymbol{X}_t = (X_t^1, \ldots, X_t^n)$

Supervenient variable that is candidate of emergence: $V_t = F(\boldsymbol{X}_t^n)$

Definition (causal emergence):
A supervenient variable V_t exhibits *causal emergence* with respect to $\boldsymbol{X}_t$ if $\mathtt{Un}(V_t; \boldsymbol{X}_{t'} | X_t^1, \ldots, X_t^n) > 0.$

Proposition (intrinsic criterium):
A system have features that exhibit causal emergence **iff** $\mathtt{Syn}(X_t^1, \ldots, X_t^n; \boldsymbol{X}_{t'}) > 0.$

Rosas FE, Mediano PA, Jensen HJ, Seth AK, Barrett AB, Carhart-Harris RL, Bor D. *Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data.* PLOS Computational Biology. 2020 Dec 21;16(12):e1008289.

Applications: practical criterion

Proposition:

Practical criterion

Theoretical criterion

$$I(V_t; V_{t'}) - \sum_{k=1}^{n} I(X_t^k; V_{t'}) > 0 \longrightarrow \mathtt{Un}(V_t; \boldsymbol{X}_{t'} | X_t^1, \ldots, X_t^n) > 0$$

Positive features of this criterion:

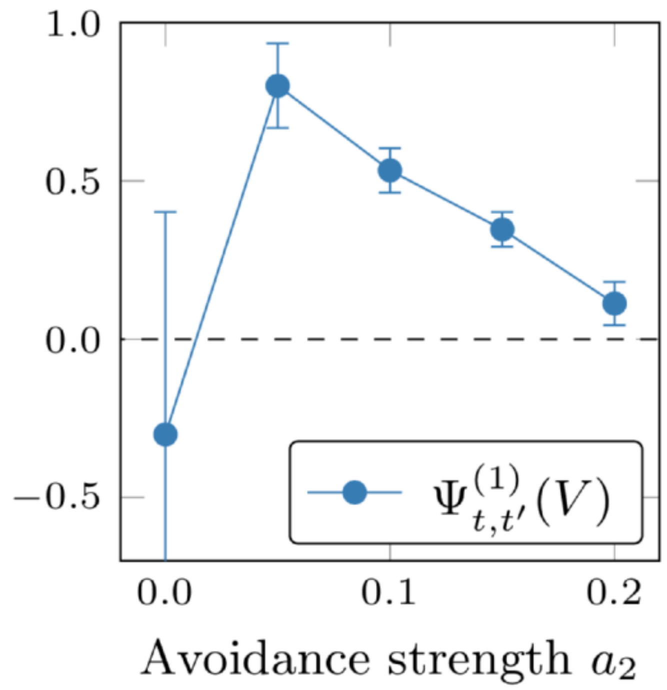
- Simple to compute.
- Avoids curse of dimensionality.
- Practical: might miss-detect but gives no false-positives.

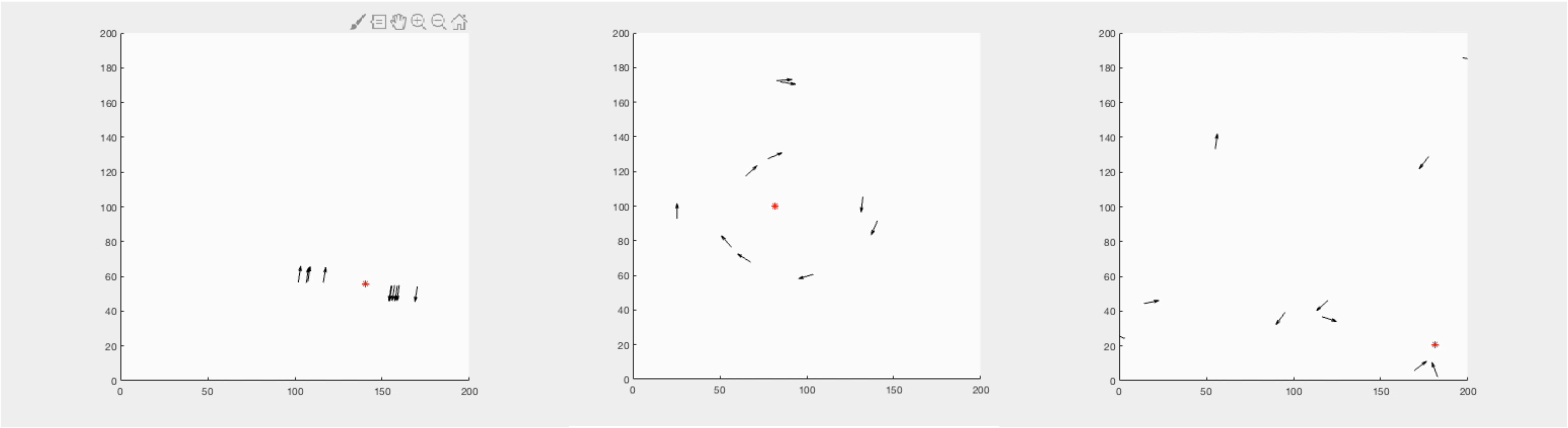
Example: flocking behaviour

Application: **Reynolds' flocking model**

Micro variables: position of each bird $\boldsymbol{X}_t \in \mathbb{R}^{2n}$

Emergent feature: center of mass $V_t \in \mathbb{R}^2$





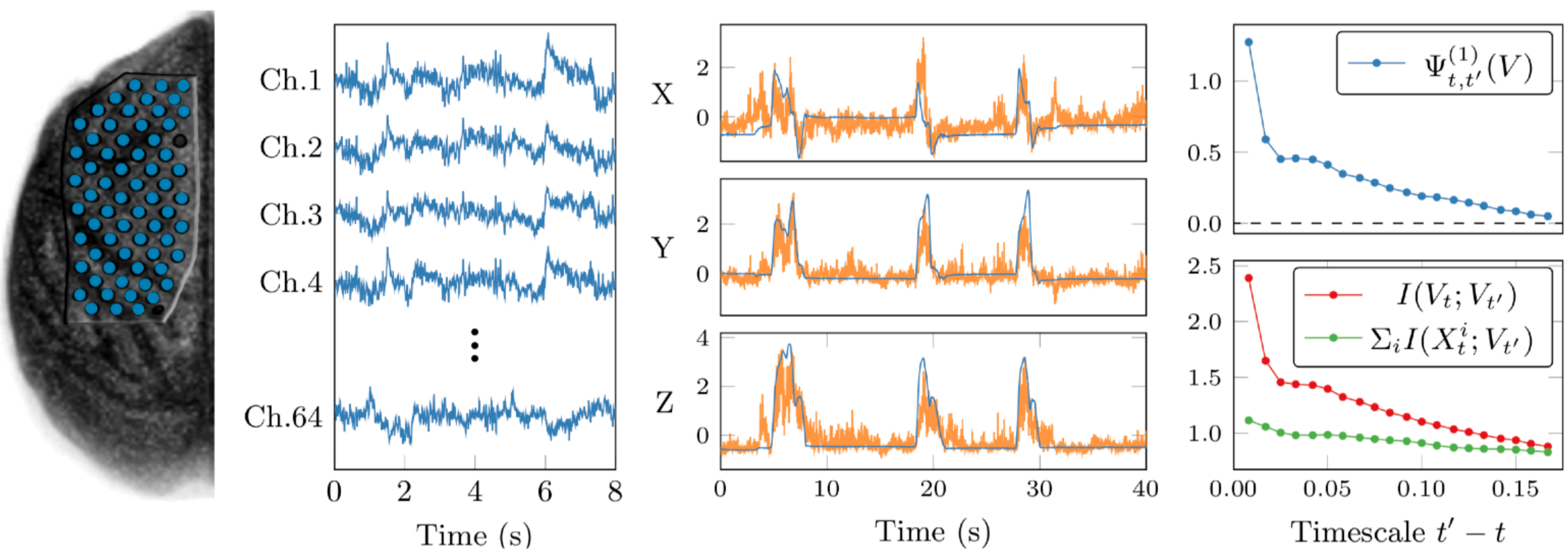
Applications: monkey neural activity

Application: **macaque's cortical activity involved in a food-grab task** (ECoG, *Neurotycho* dataset)

Micro variables: ECoG channels $\boldsymbol{X}_t \in \mathbb{R}^{64}$

Emergent feature: LPS-SVM estimator of monkey's wrist position $V_t = F(\boldsymbol{X}_t) \in \mathbb{R}^3$

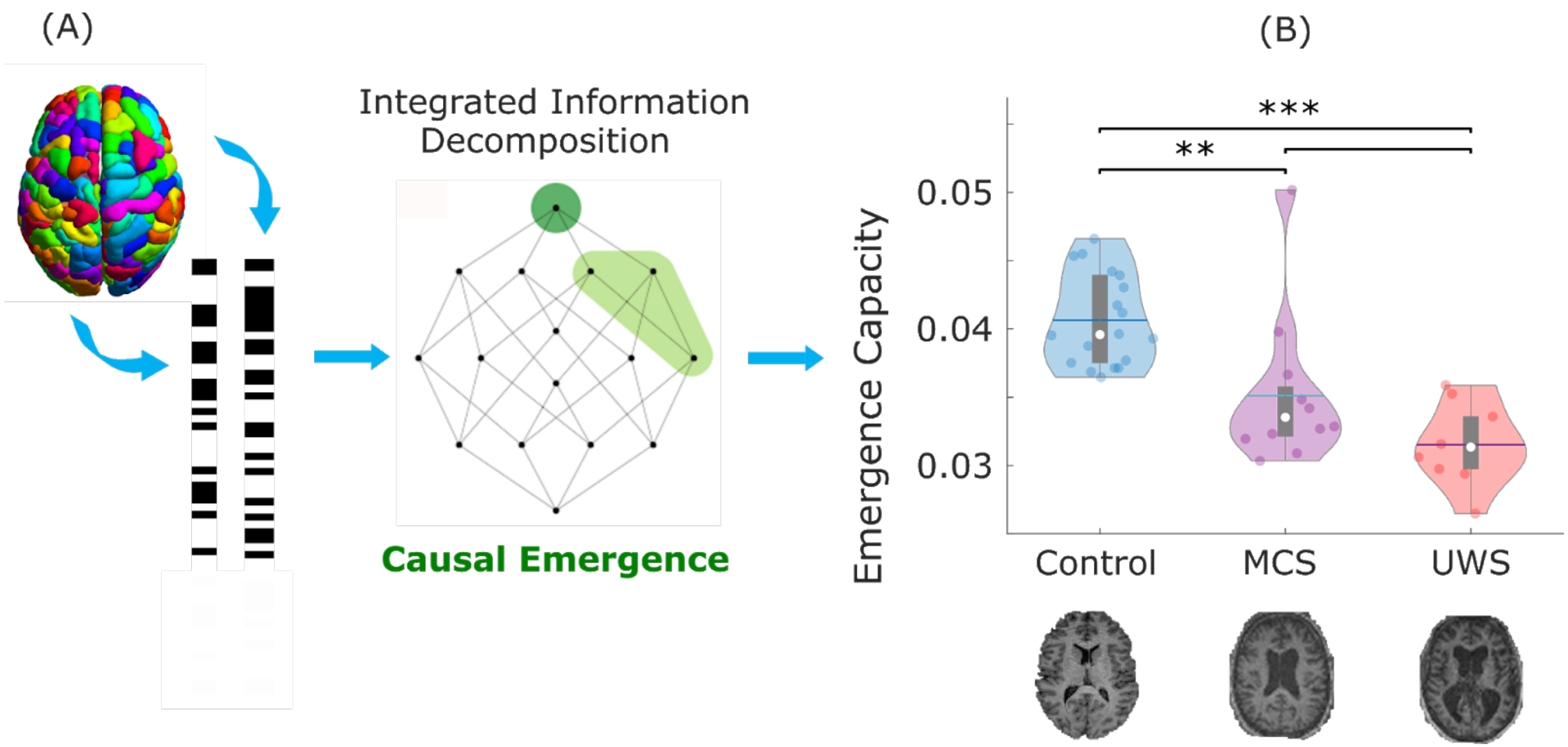
$$I(V_t; V_{t'}) - \sum_{k=1}^{n} I(X_t^k; V_{t'}) > 0$$



Z. Chao, Y. Nagasaka, and N. Fujii, "Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkey," Frontiers in Neuroengineering, vol. 3, p. 3, 2010

Loss of consciousness disrupts gateways

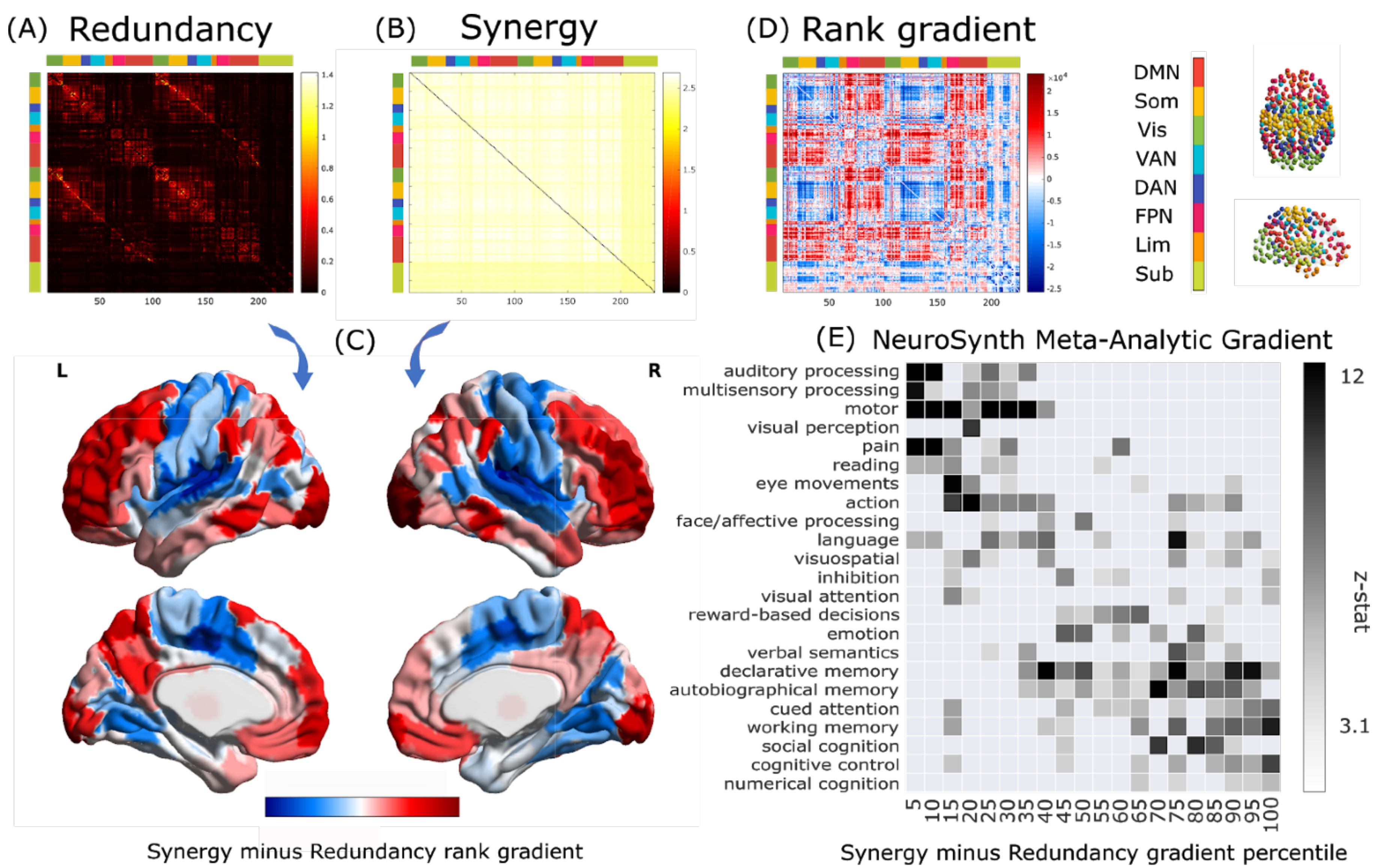
The emergence capacity is higher in control subjects than in minimally conscious or unresponsive wakeful subjects.



Luppi A., Mediano P., Rosas F., Allanson J., Pickard J., Williams G., Craig M., Finoi P., Peattie A., Coppola P., Menon D., Bor D., and Stamatakis E., "Reduced causal emergence in the brains of chronically unconscious patients: structural and functional contributions," in preparation.

Identifying a synergistic core in the human brain

A **gradient** of **synergy** minus **redundancy** in fMRI resting-state data (HCP) show a **synergistic core** located mainly in **DMN** and **FPN**!



Luppi, A. I., Mediano, P. A., **Rosas, F. E.**, Holland, N., Fryer, T. D., O'Brien, J. T., ... & Stamatakis, E. A. (2020). "*A synergistic core for human brain evolution and cognition.*" Accepted in Nature Neuroscience, to be published.

Today's menu
1. Synergy
2. Emergence
3. Ideas to take home

Ideas to take home

- High-order statistics have plentiful unexplored potential in practical data analyses.

Ideas to take home

- High-order statistics have plentiful unexplored potential in practical data analyses.

- Using synchronous coarse-grained metrics one can capture synergistic aspects of systems of interest.

Ideas to take home

- High-order statistics have plentiful unexplored potential in practical data analyses.

- Using synchronous coarse-grained metrics one can capture synergistic aspects of systems of interest.

- Using dynamic fine-grained measures one can build formalisms to quantify causal emergence from data.

Thanks to my synergistic collaborators:

Pedro A.M. Mediano

(University of Cambridge)

Andrea Luppi

(University of Cambridge)

- Robin Carhart-Harris, Henrik Jensen (*Imperial College London*)
- Daniel Bor, Emmanuel Stamatakis (*University of Cambridge*)
- Anil Seth, Adam Barrett (*University of Sussex*)
- Borzoo Rassouli (*University of Essex*)
- Michael Gastpar (*EPFL*)
- Daniele Marinazzo (University of Gent)
- Sebastiano Estramaglia (Universitá degli Studi Aldo Moro)
- Rodrigo Cofre, Marilyn Gatica (Univeersidad de Valparaiso)

Imperial College London

UNIVERSITY OF CAMBRIDGE

US UNIVERSITY OF SUSSEX

University of Essex

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

GHENT UNIVERSITY

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

Universidad de Valparaíso CHILE

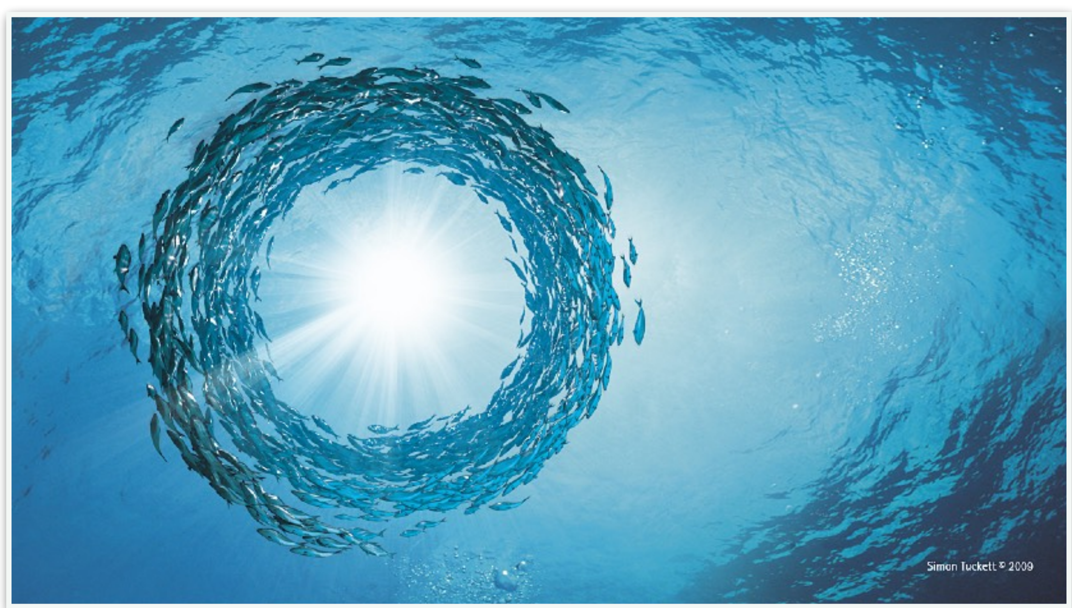
Data Science Institute

THE CENTRE FOR PSYCHEDELIC RESEARCH

Complexity Science

Imperial College London

Thank you!



Contact information:
Fernando E. Rosas
f.rosas@imperial.ac.uk