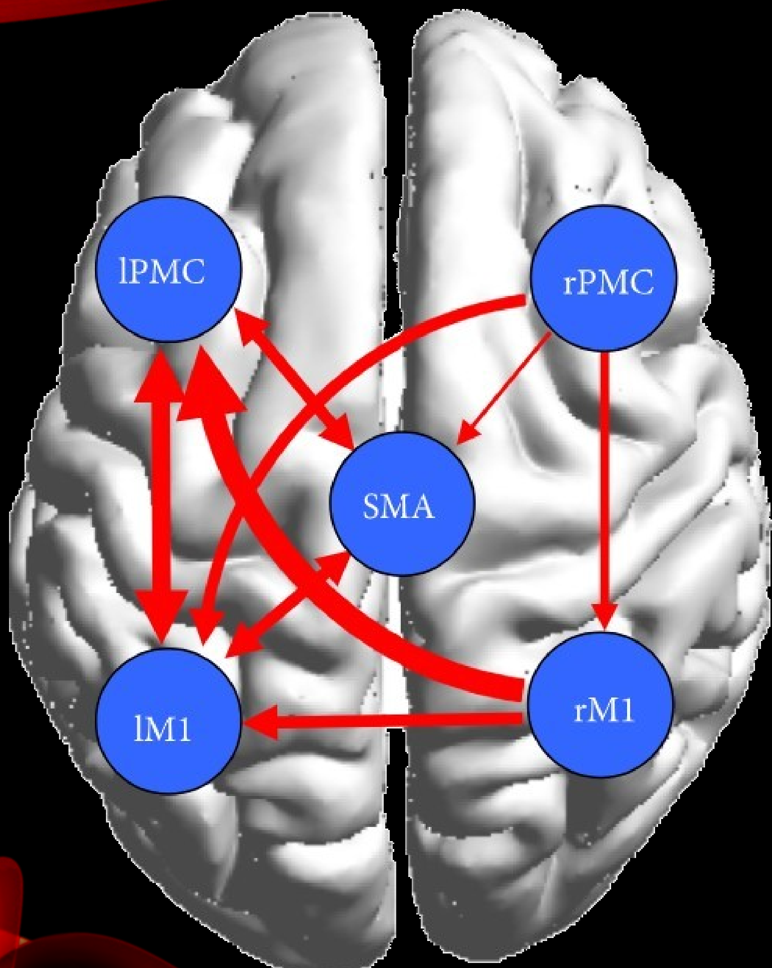


AGH
AGH University of Science and Technology

Machine Learning for Neuroimaging and Neuroscience

Lecture 5

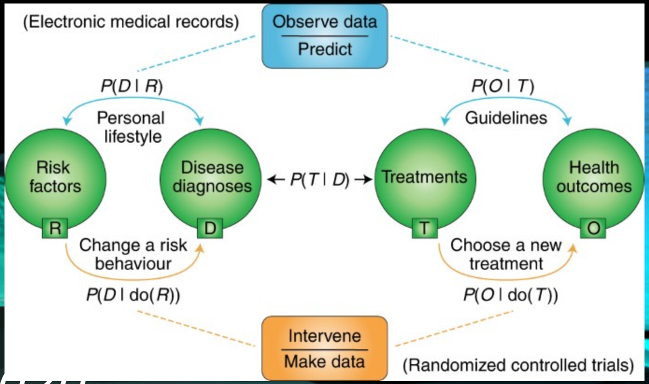
Effective Connectivity

AGENDA

- Introduction
- Brush up on Friston’s DCM, Pearl’s, Granger’s and Sugihara causality
- Structurally constrained effective connectivity (Granger version) (Crimi et al. Neuroimage 2021)
- Effective connectivity with Sugihara causality (Falco-Roget et al. PlosCompBio2023)

DISCLAIMER

- Causal inference provides a rigorous framework for understanding and quantifying cause-and-effect relationships, which has important applications in domains such as healthcare, economics, social sciences, and policy-making.
- We do not focus here on causal inference for machine learning
- Künzel, Sören R., et al. “Metalearners for Estimating Heterogeneous Treatment Effects Using Machine Learning.” PNAS 2019
- “Causal inference” J. Pearl
- CausalML, CausalPy, DoWhy,GCASTLE



Prosperi et al. 2020

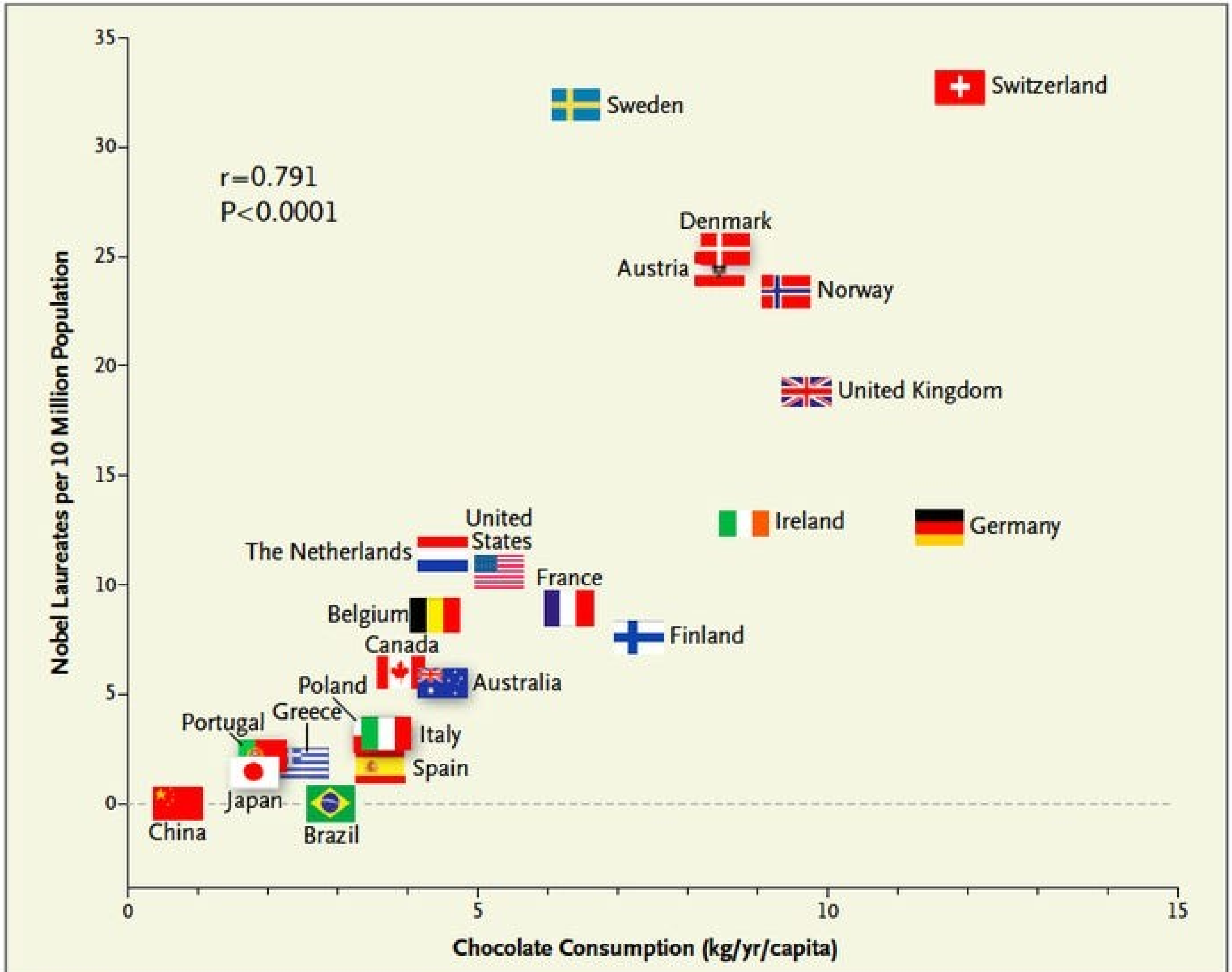


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Milk, chocolate and Nobel prizes. Linthwaite and Fuller, Practical Neurology 2013

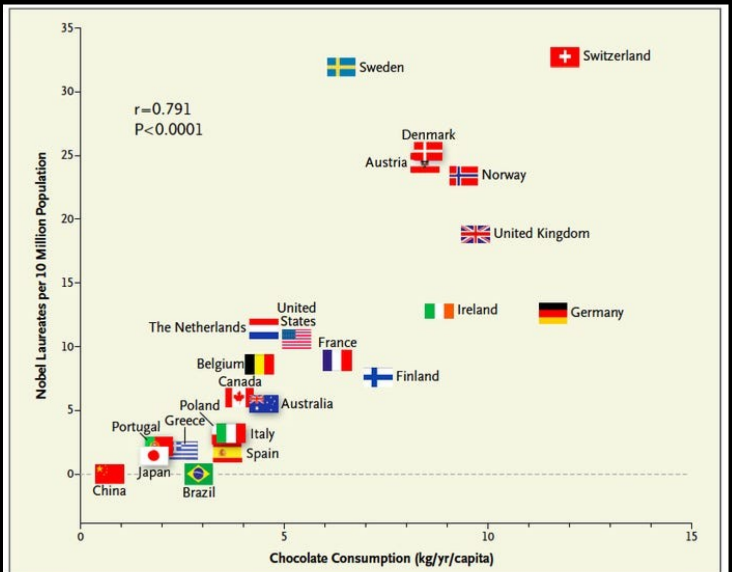
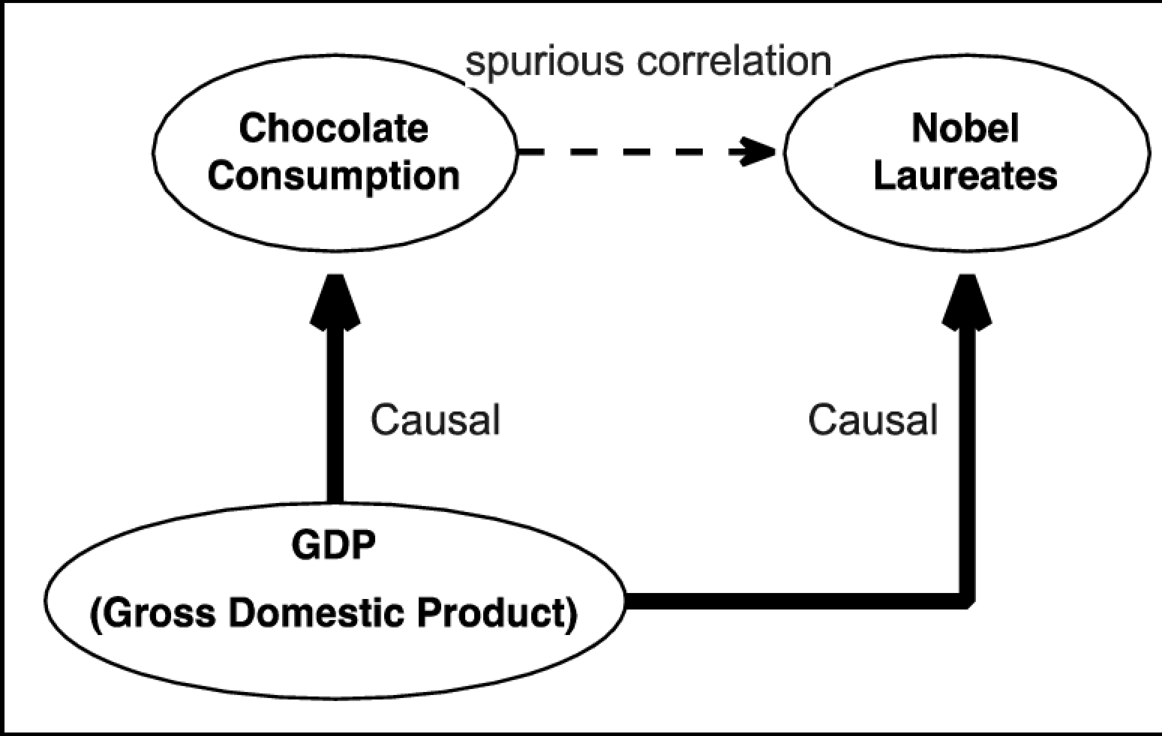
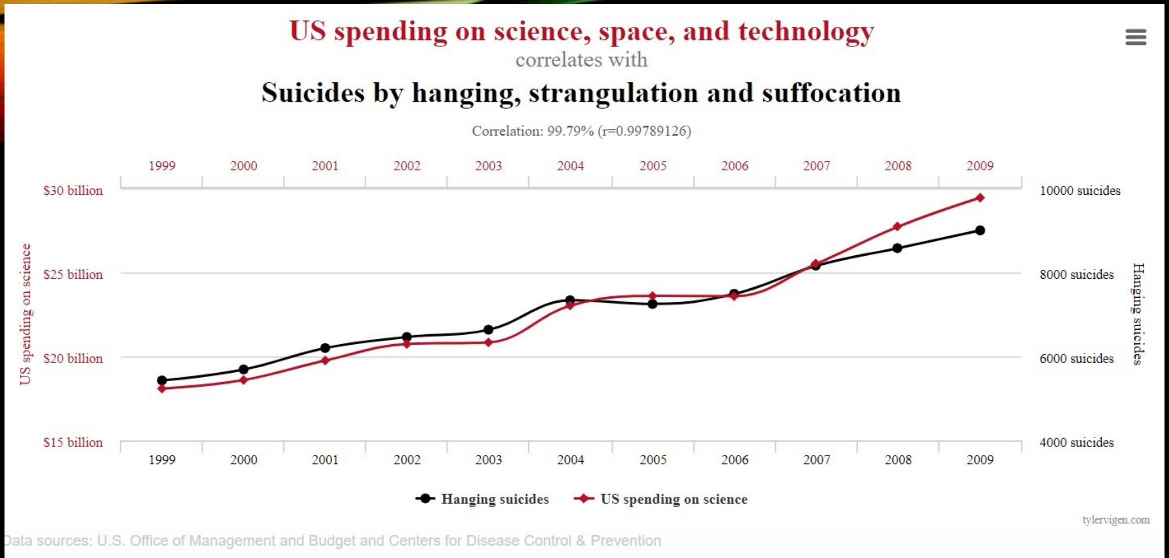
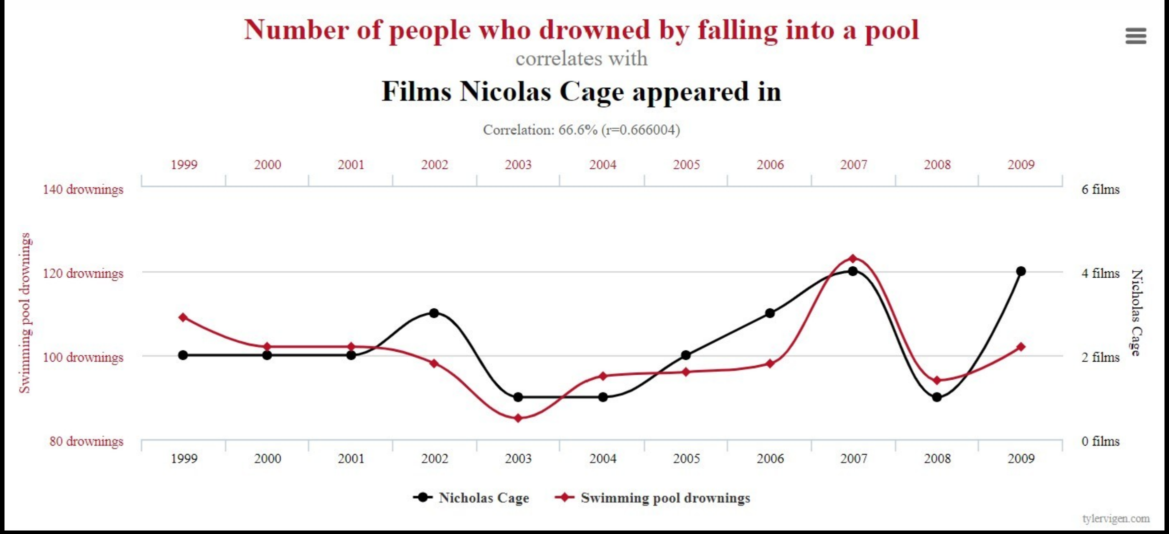


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.



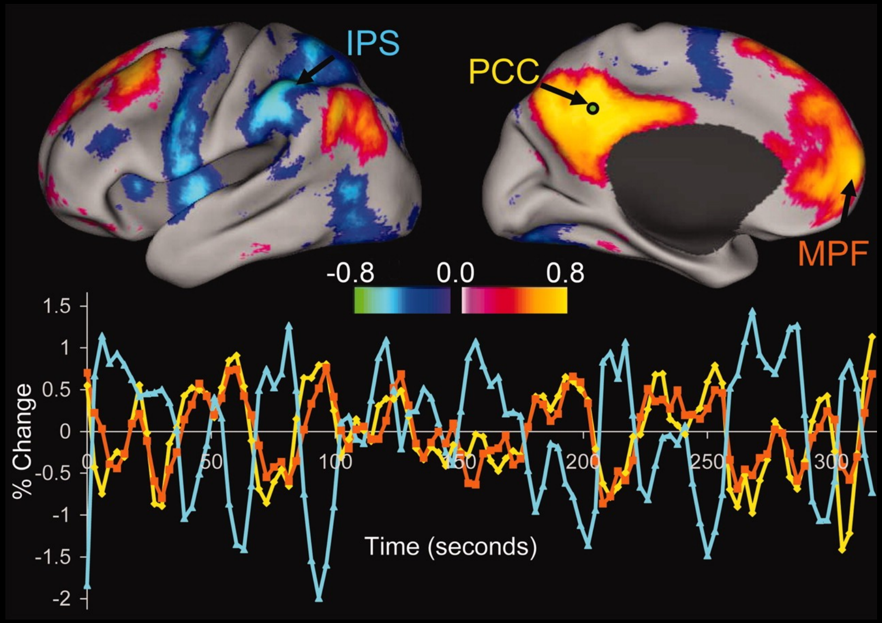
Milk, chocolate and Nobel prizes. Linthwaite and Fuller, Practical Neurology 2013





SPURIOUS CORRELATIONS (WWW.TYLERVIGEN.COM)

FUNCTIONAL CONNECTIVITY AND CAUSALITY



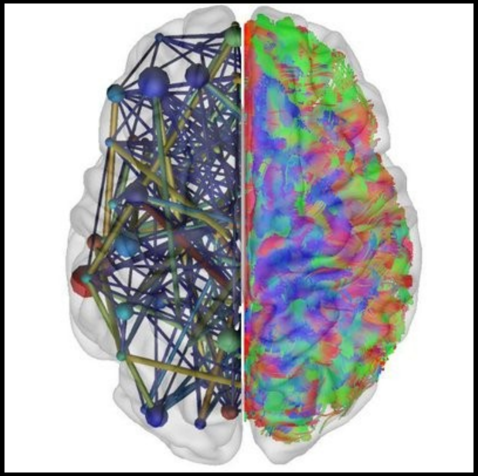
Fox et al. PNAS 2005

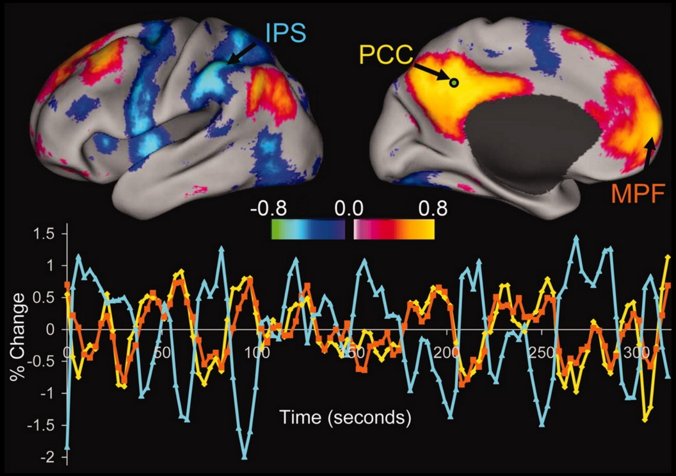
Research questions:

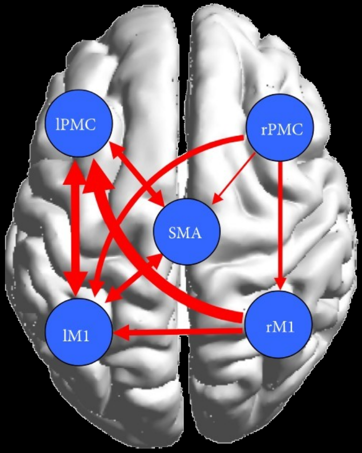
1. Can we assure causality (effective connectivity)?

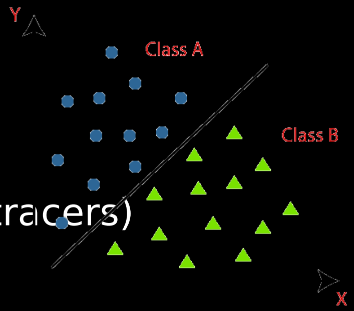
2. What would be the state-of-art even with limitations?

STRUCTURAL, FUNCTIONAL & EFFECTIVE CONNECTIVITY



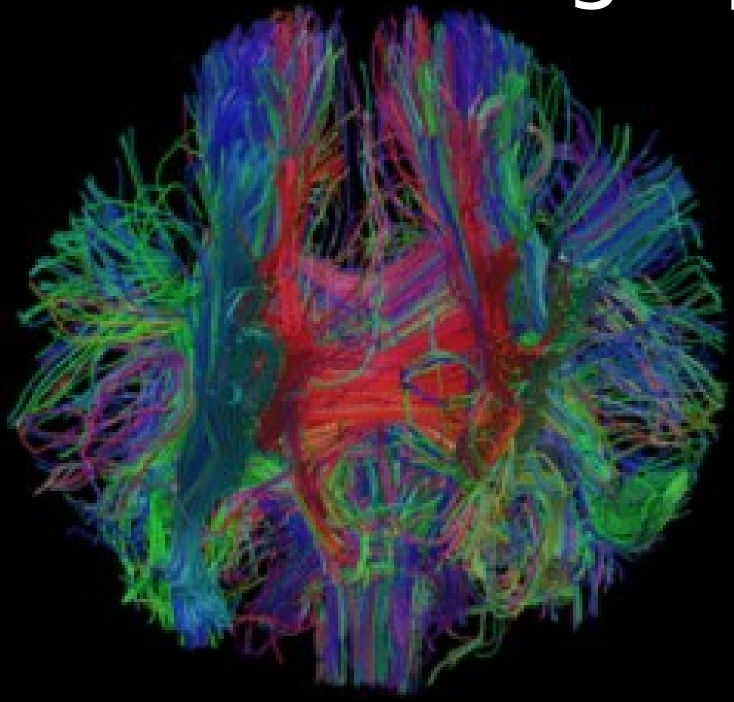




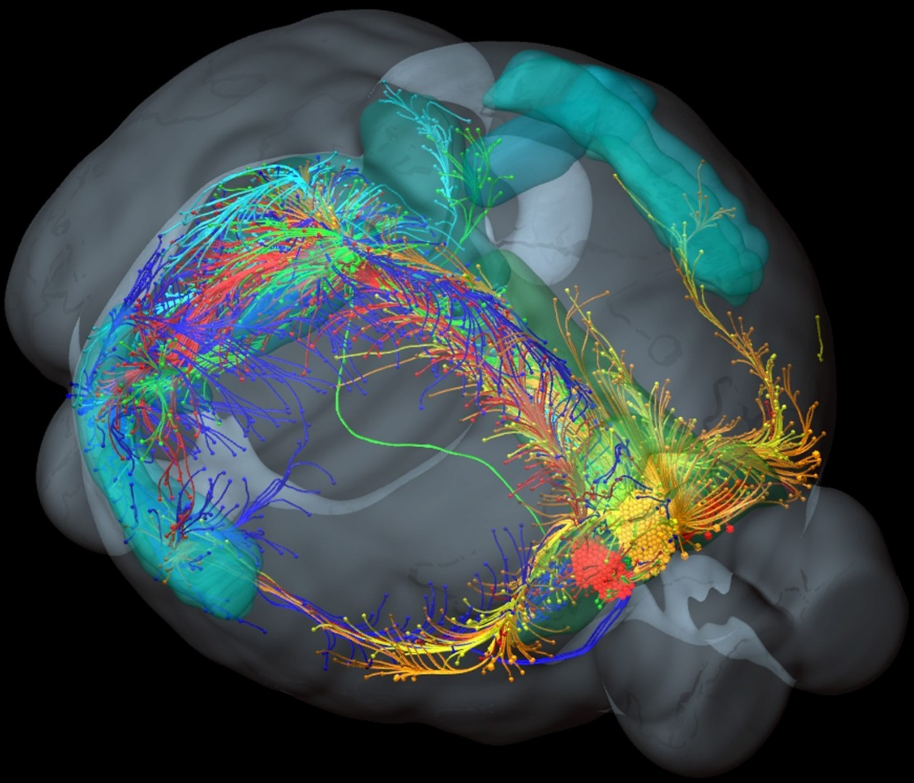


- **Structural/anatomical connectivity**
  = presence of axonal connections / white matter tracks (eg, DWI, AAV tracers)
- **Functional connectivity**
  = statistical dependencies between regional time series (eg, Pearson correlation, ICA,...)
- **Effective connectivity**
  = causal (directed) influences between neuronal populations (eg, DCM, Granger C., etc)
- Nobody cares about **Morphological connectivity...**

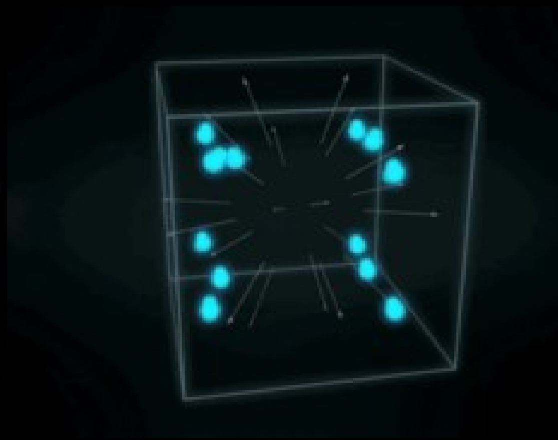
What is tractography



www.dipy.org

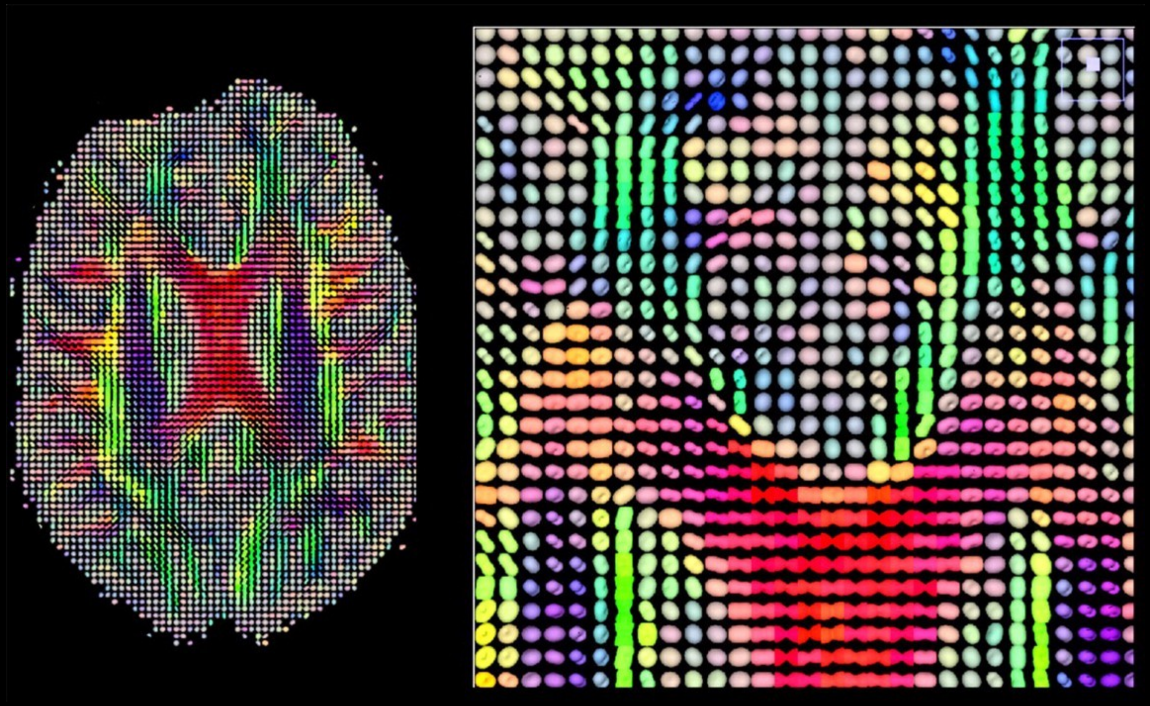


Oh. et al. 2014

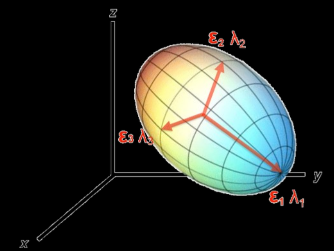


Credit: MaxPlanck Society

diffusion weighted images

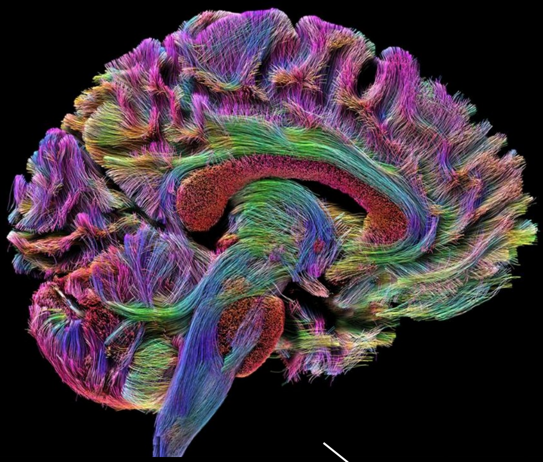


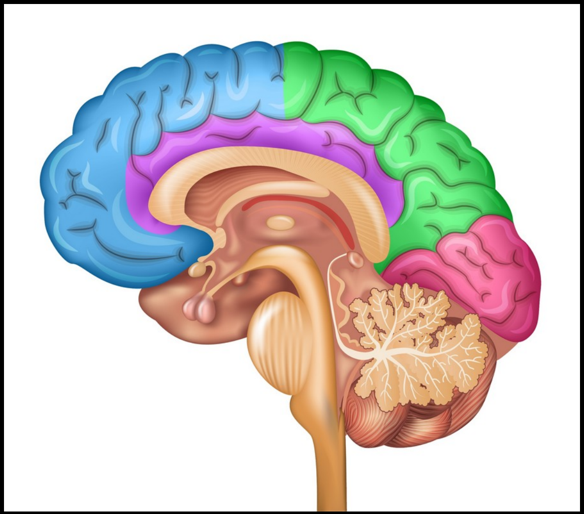
$$\bar{D} = \begin{vmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{vmatrix}$$

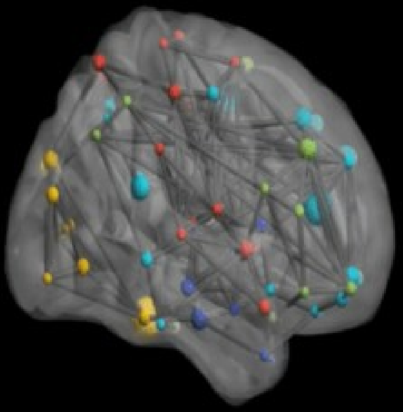


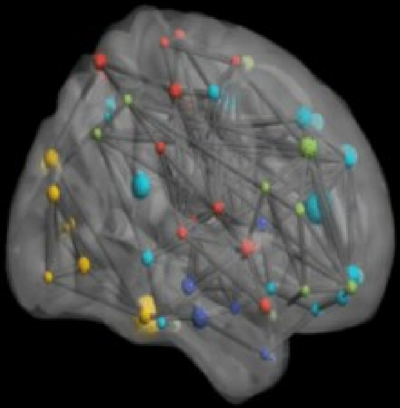
$$FA = \sqrt{\frac{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_1-\lambda_3)^2}{2(\lambda_1{}^2+\lambda_2{}^2+\lambda_3{}^2)}}$$

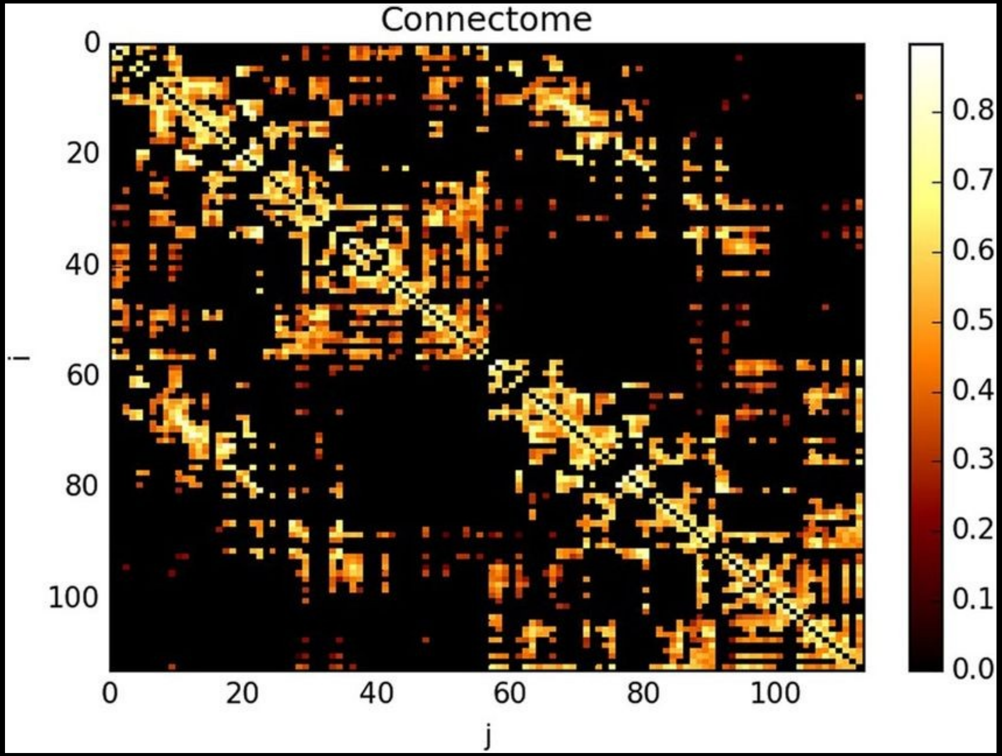
(Alger 2012 J.Neurocience)







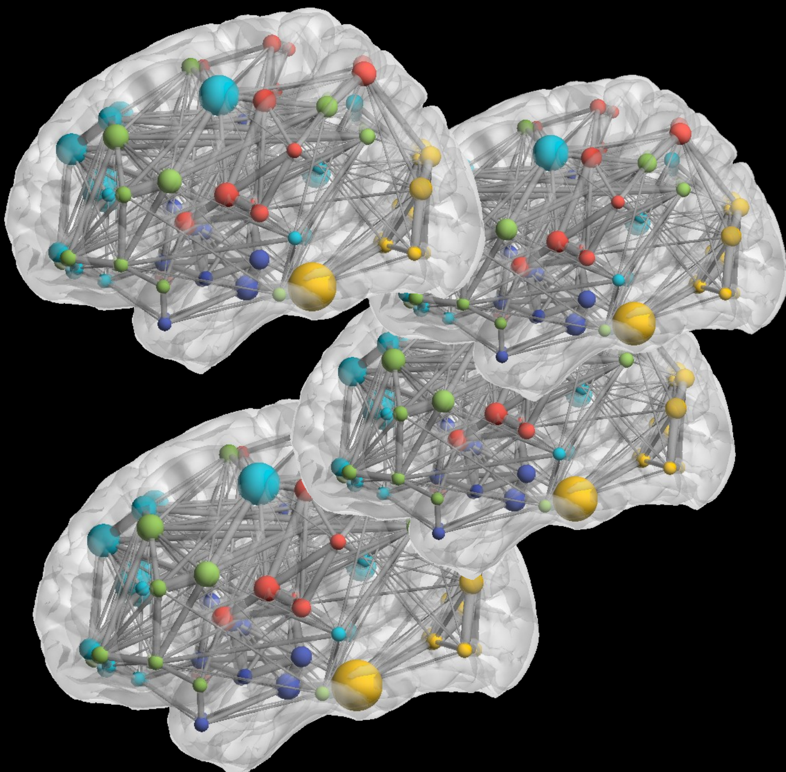


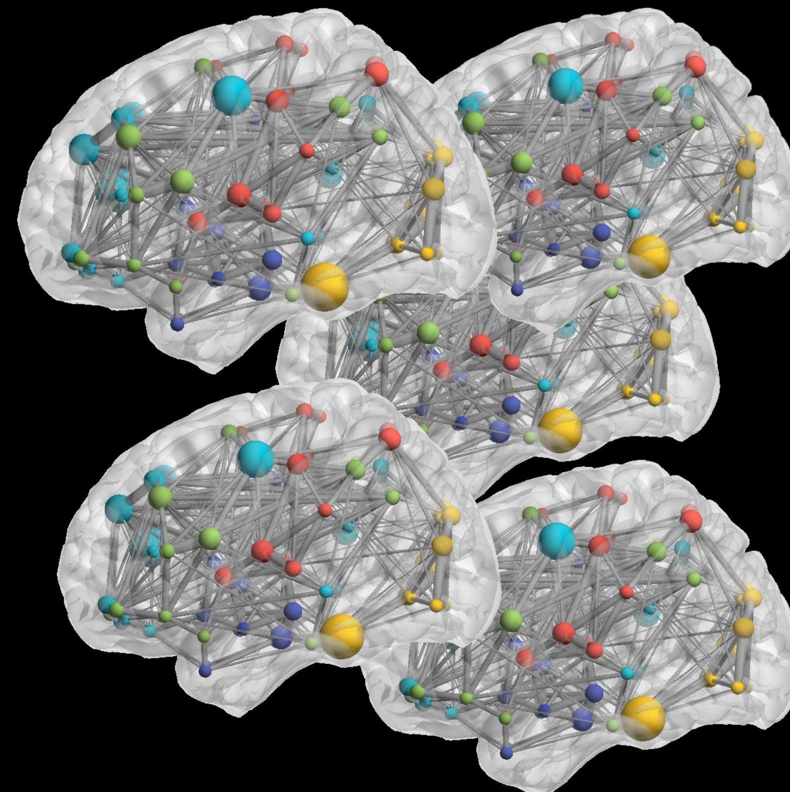


Connectome
0
20
40
60
80
100
i
j
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

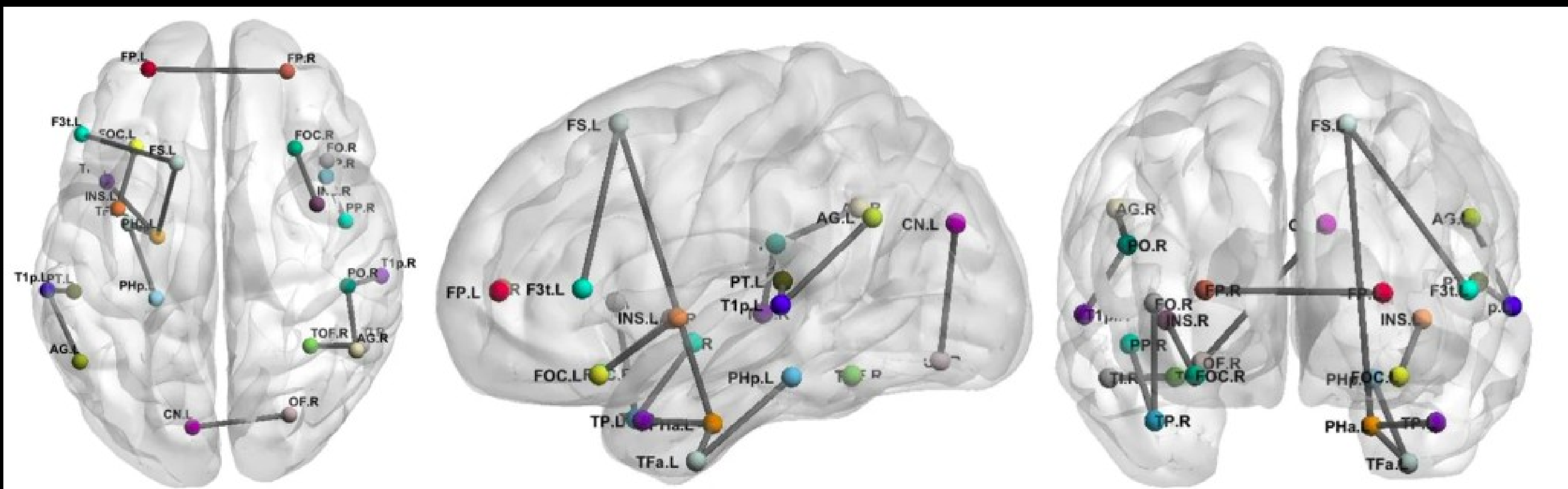
Alzheimer, Schizophrenia ...
Matched healthy control

Healthy vs Alzheimer





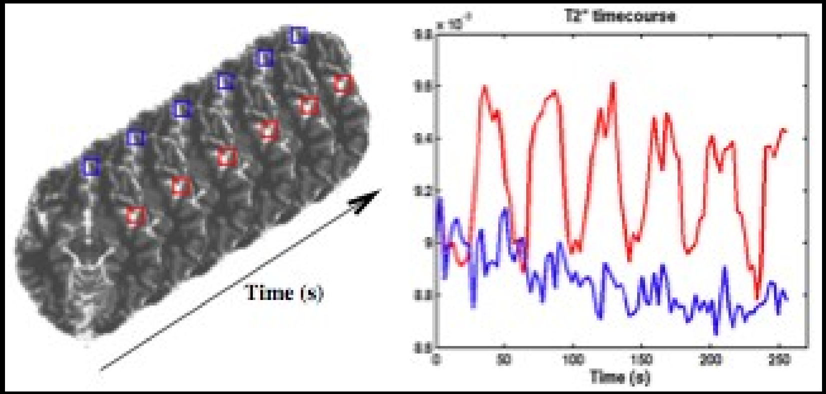
Difference between healthy and Alzheimers



(Crimi et al., Nature Scientific Reports 2019)

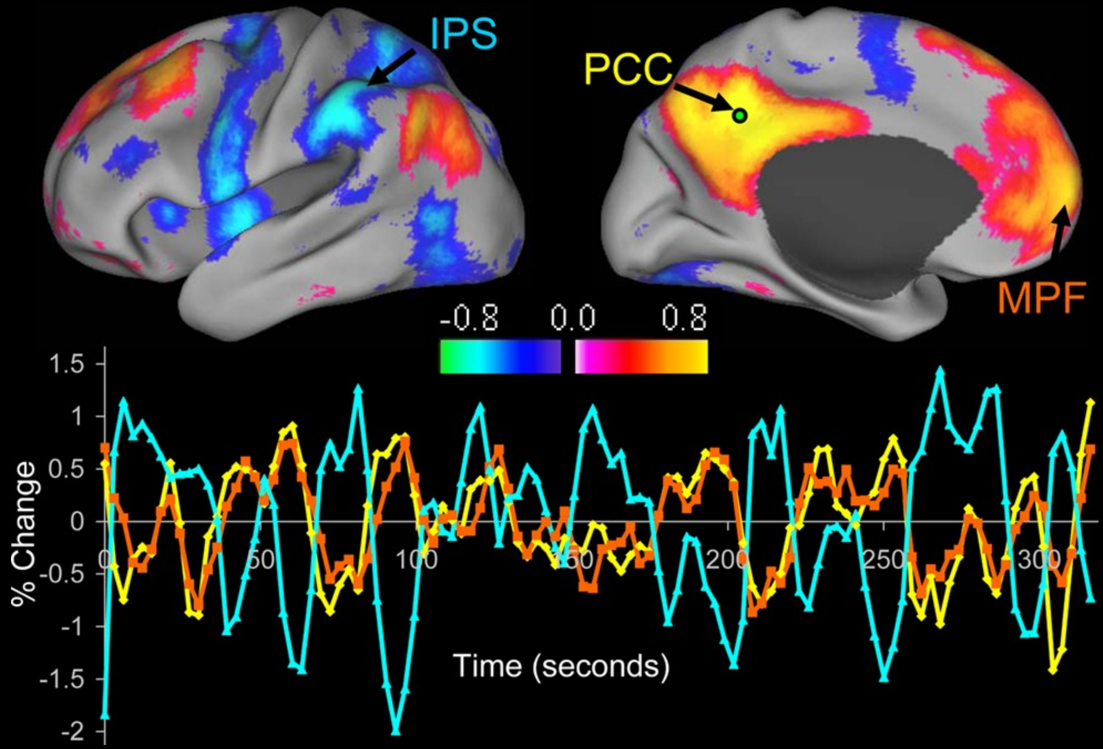
FUNCTIONAL MRI

- Hemodynamic response: blood releases oxygen to firing neurons at a greater rate than to inactive neurons.
- This causes a change of the relative levels of oxyhemoglobin and deoxyhemoglobin (oxygenated or deoxygenated blood) that can be detected on the basis of their differential magnetic susceptibility.



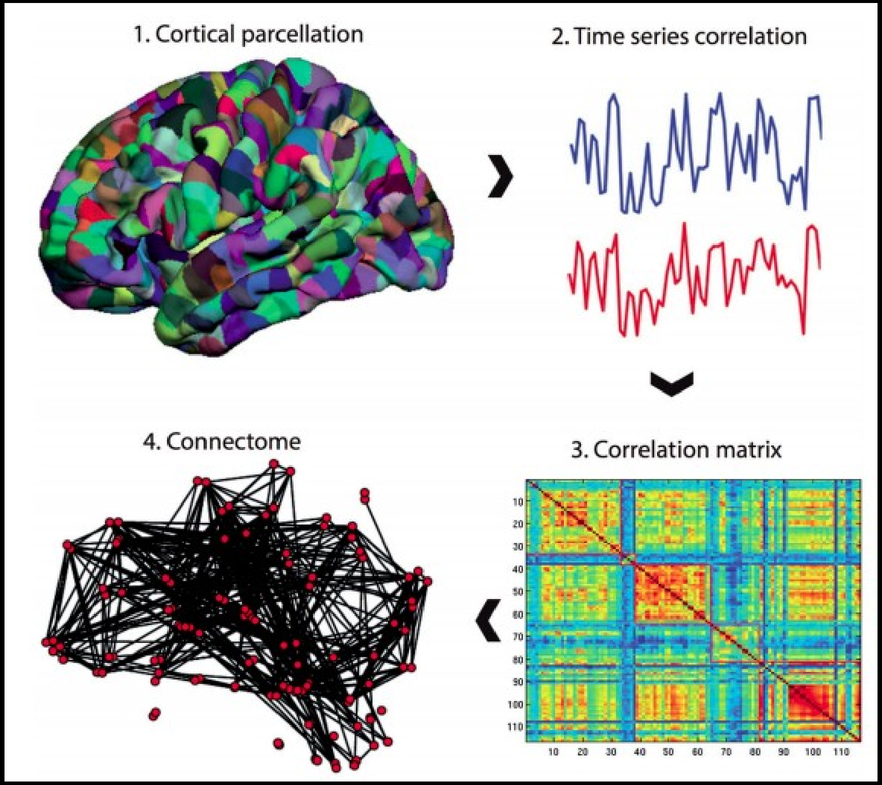
Gordon et al. (Cerebral Cortex 2014)

BRUSH UP ON FUNCTIONAL CONNECTIVITY

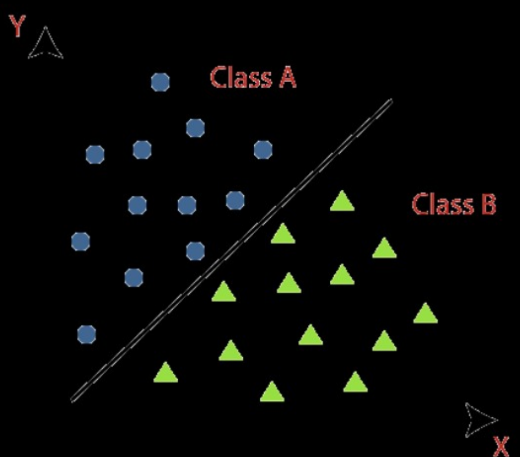


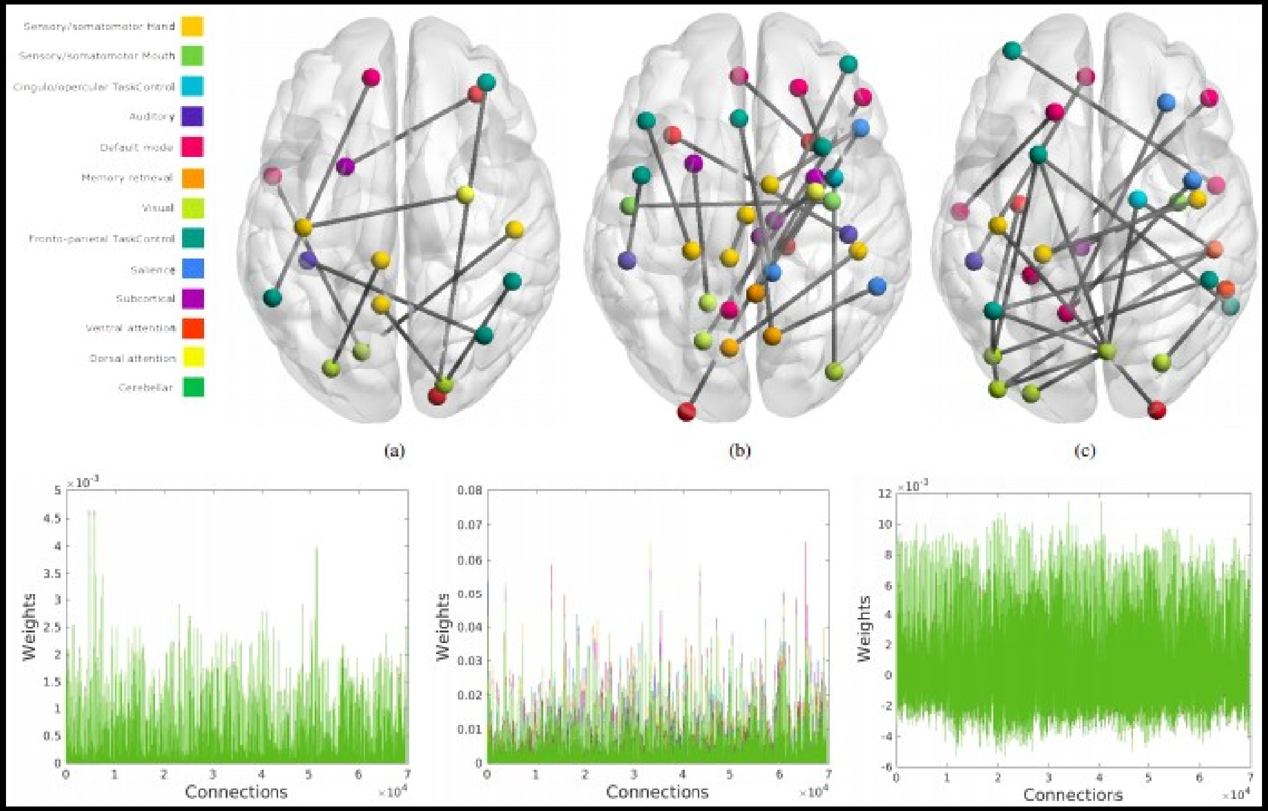
Fox & Greicius 2010

FUNCTIONAL CONNECTIVITY



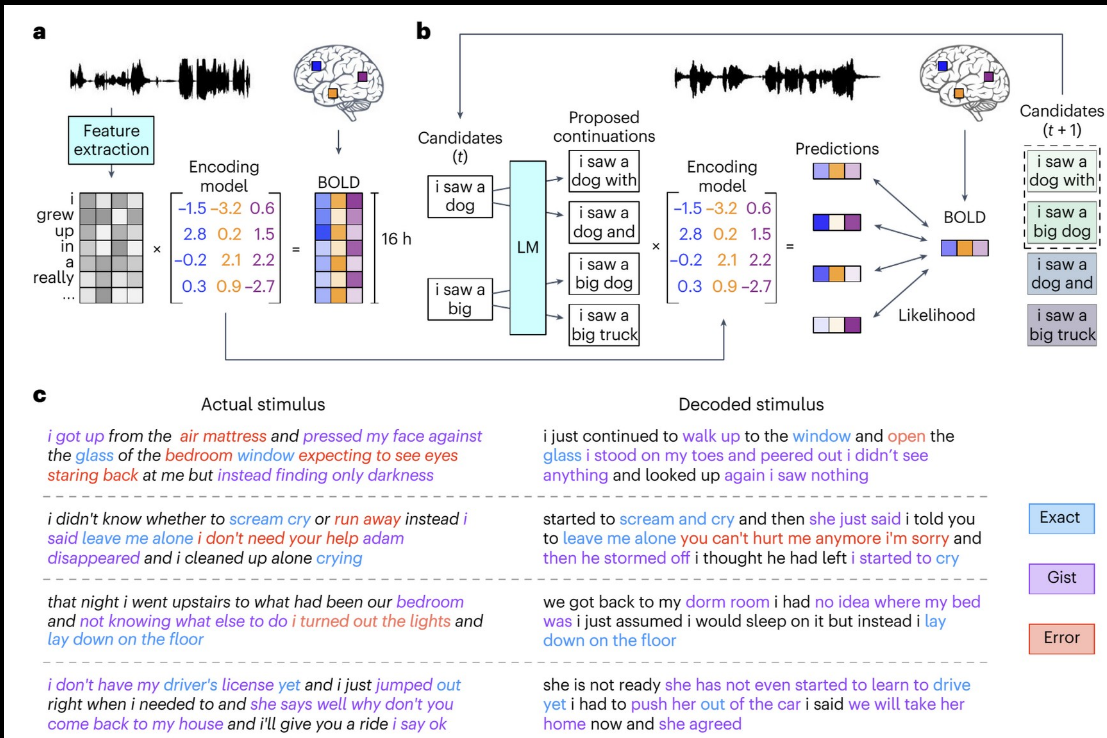
USING SUPPORT VECTOR MACHINE

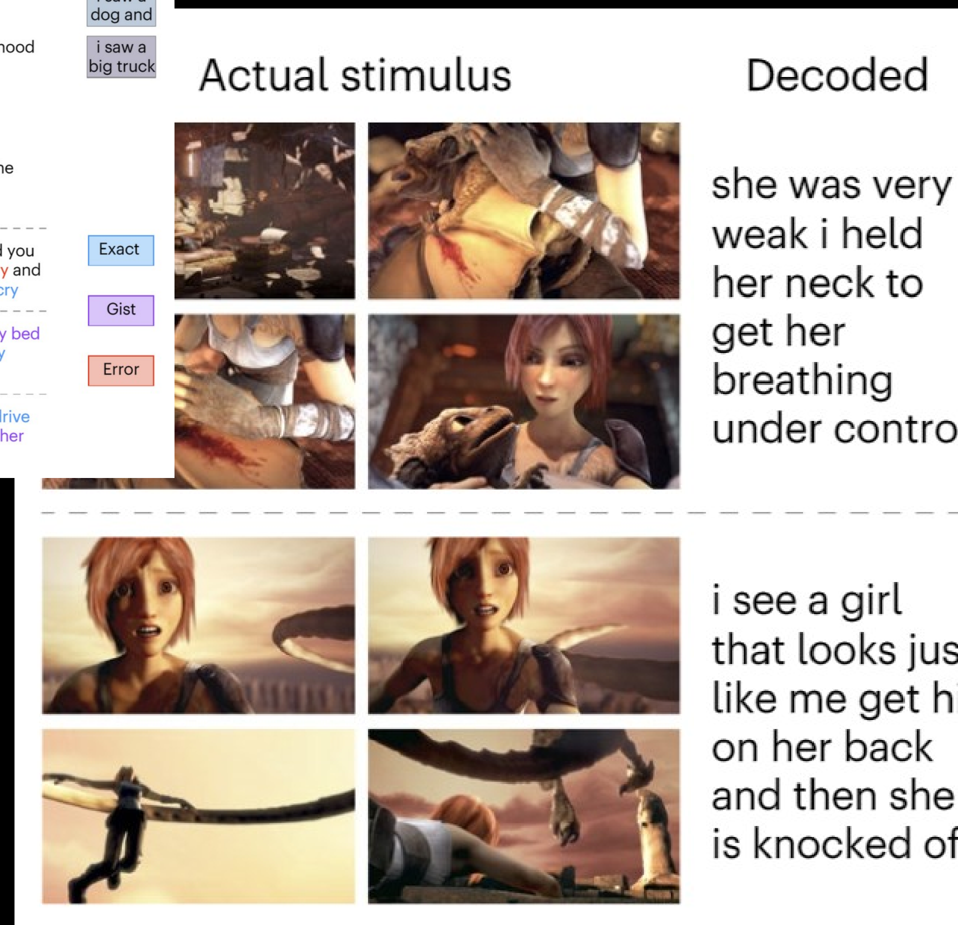


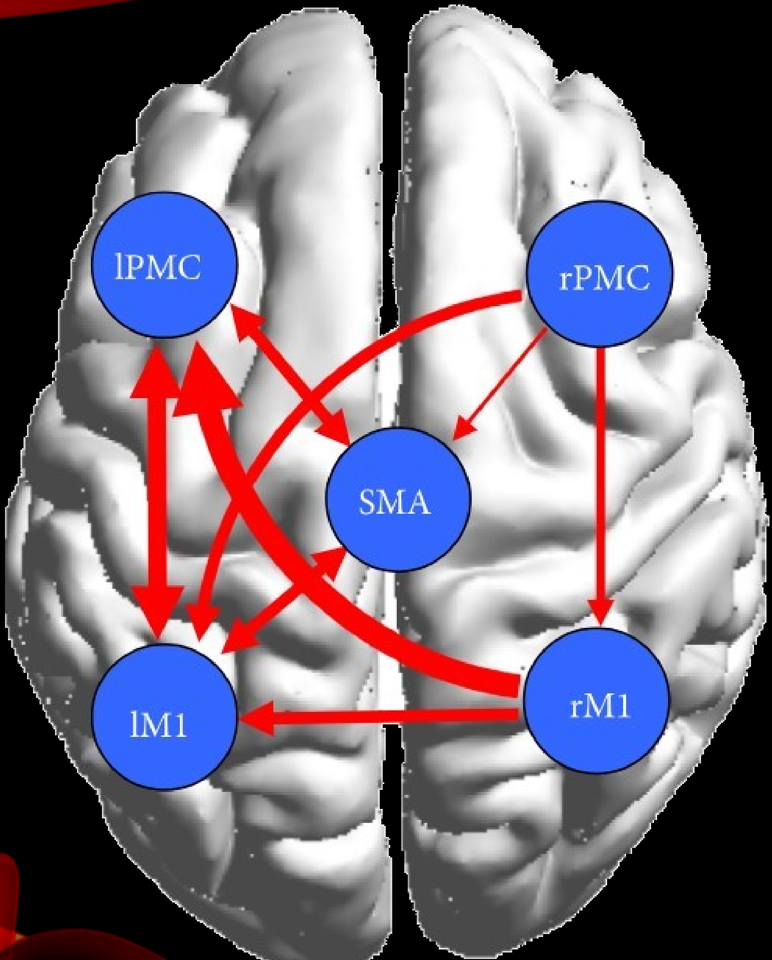


Crimi et al. Neuroimage 2021 (work done in 2014/2015)

"SEMANTIC RECONSTRUCTION OF CONTINUOUS LANGUAGE FROM NON-INVASIVE BRAIN RECORDINGS" TANG ET AL. NATURE NEUROSCIENCE 2023



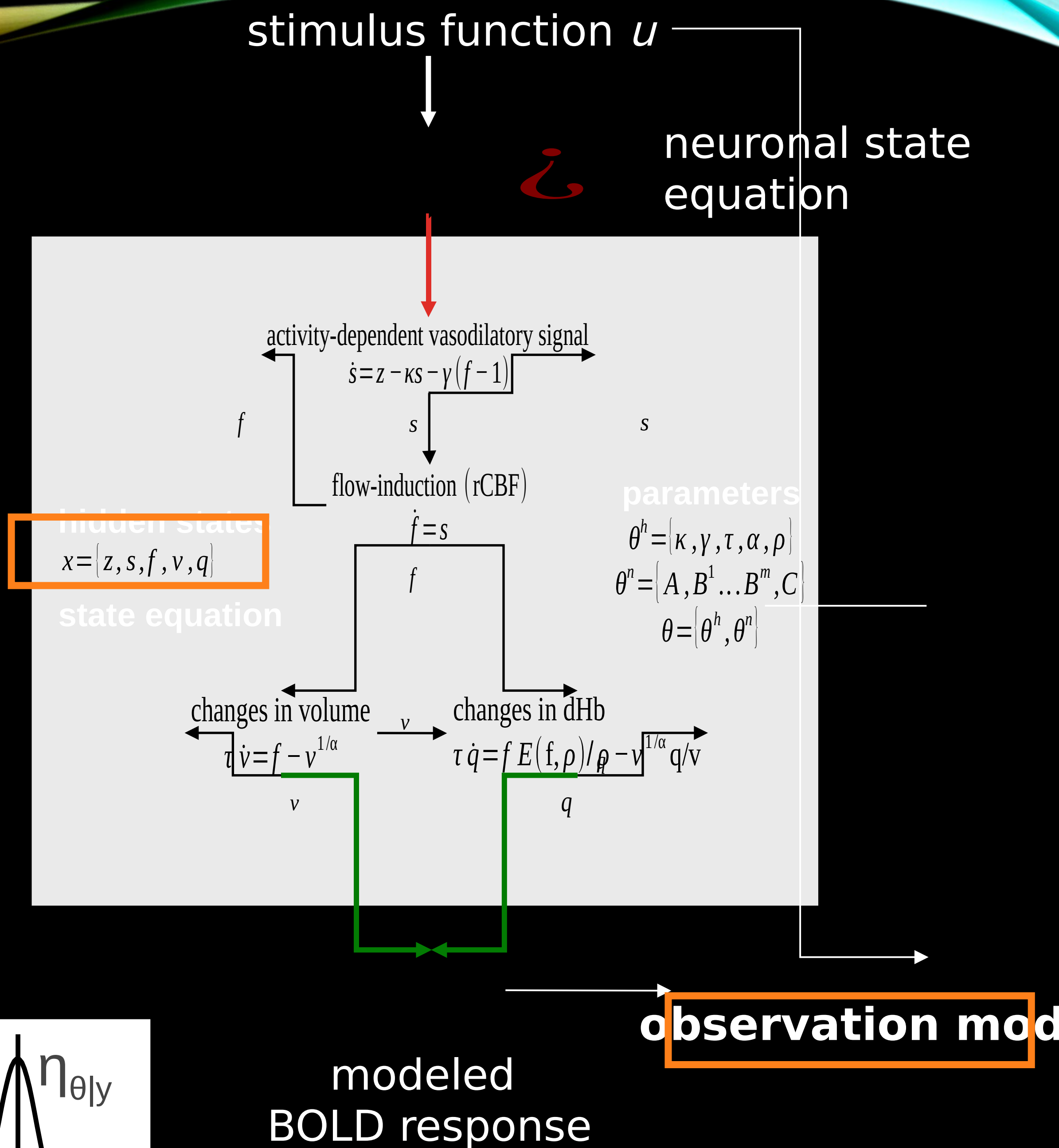


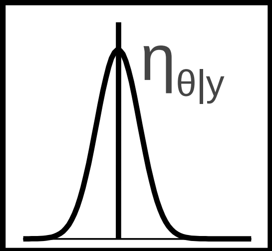


STRUCTURALLY CONSTRAINED EFFECTIVE CONNECTIVITY

DCM Overview: parameter estimation

- Specify model (neuronal and haemodynamic level)
- Make it an observation model by adding measurement error *e* and confounds *X* (e.g. drift).
- Bayesian parameter estimation using expectation-maximization.
- Result: (Normal) posterior parameter distributions, given by mean $\eta_{\theta|y}$ and Covariance $C_{\theta|y}$.





DCM Estimation: Bayesian framework

Models of
- Haemodynamics in a single region
- Neuronal interactions

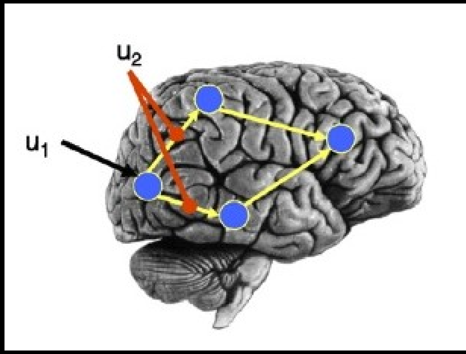
Constraints on
- Haemodynamic parameters
- Connections

likelihood

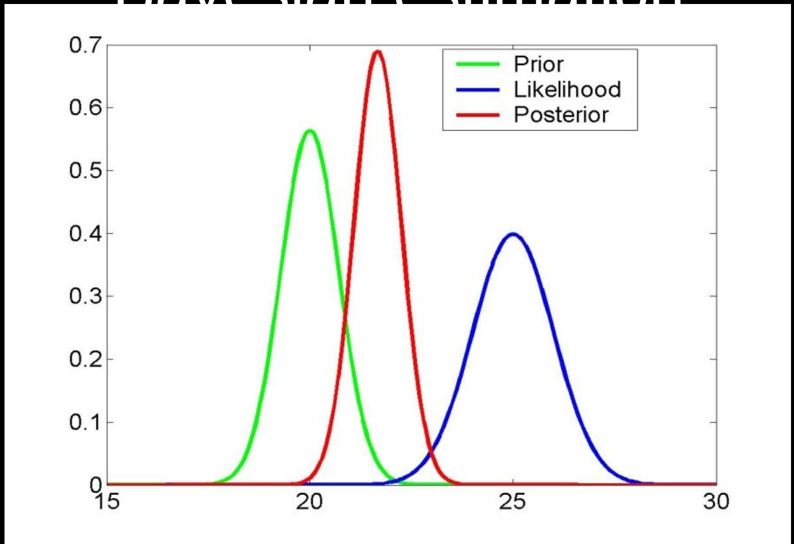
priors

posterior

Bayesian estimation

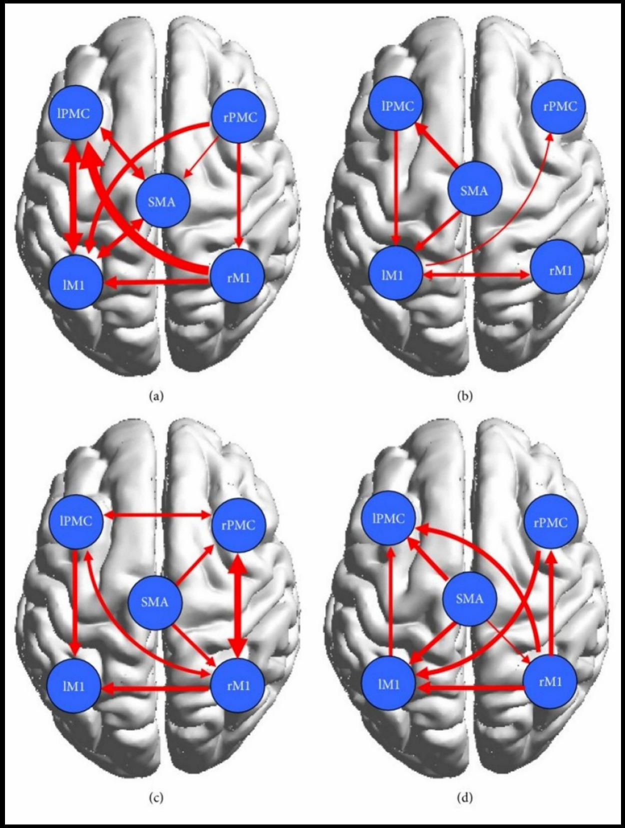


Friston et al.
Neuroimage 2003



Inferences on:
1. *Parameters*
2. *Models*

LEARN DCM AND PUT IT ASIDE



- DCM estimates neuronal interactions using a priori information about intrinsic physiology.
- High computational costs related to parameters and not scalable.
- Poor performance in resting-state (even using variations)
- Hypothesis driven to validate prior hypothesis (not really flexible and scalable).

Gowda et al.
Cures 2021

What is causality?

- **Pearl causality or Perturbation causality**(no complaint):

“Variable X causes variable Y if some externally applied perturbation of X at a point in time results in a perturbation in Y at a future or current time”

It strictly holds only in epigenetics and TMS, unpractical

- **Granger causality**:

A time series X causes a time series Y if given the past information in X predicts future Y better than without the history of X. Here X values provide statistically significant information about future values of Y.

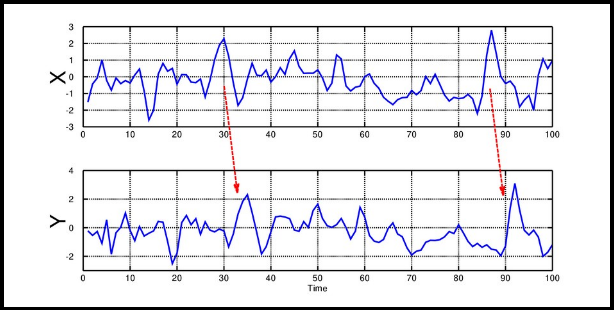
Criticized as working mostly in stationary signals, assuming separability.

- **Sugihara Causality and perturbation (later)**

What is causality?

- **Granger causality**:

A time series X causes a time series Y if given the past information in X predicts future Y better than without the history of X. Here X values provide statistically significant information about future values of Y.



$$X(t) = \sum_{\tau=1}^{L} A_\tau X(t-\tau) + \varepsilon(t),$$

Then tested through a series of t-tests and F-tests.

Under the assumption of using Gaussian variables, GC is computed as transfer entropy**: X Granger causes Y if and only if the transfer entropy from X to Y is nonzero (Barret et al. 2009 Physics review)**

Criticized as working mostly in stationary signals, **assuming separability**.

For non-stationary signals = introduce windowing

For indirect connections = introduce propagators

(Marinazzo, Roebroeck, Seth, etc)

Reichenbach Common Cause

"…If two events are correlated, then either there is a causal connection between the correlated events that is responsible for the correlation or there is a third event, a so called common cause, which brings about the correlation. In short we cannot assure that there is a real perturbation or just an unseen variable which is causing first one variable and then another one…" Reichenbach, 1956

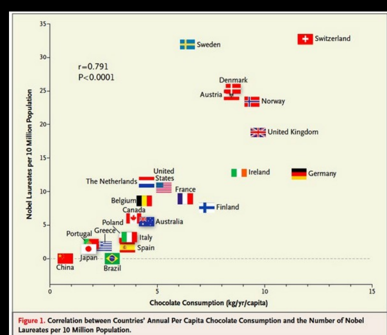
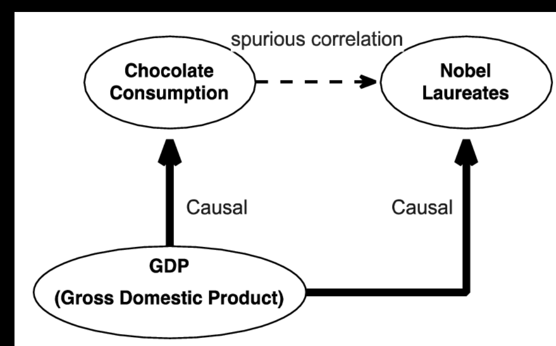
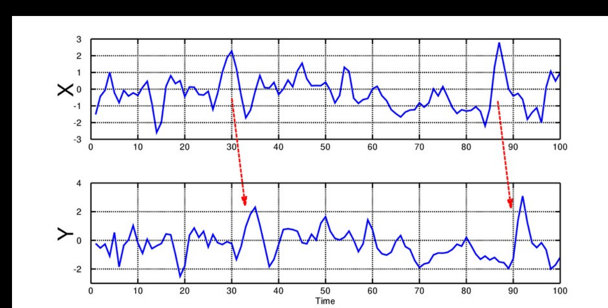


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.





Reichenbach Common Cause

"...If two events are correlated, then either there is a causal connection between the correlated events that is responsible for the correlation or there is a third event, a so called common cause, which brings about the correlation. In short we cannot assure that there is a real perturbation or just an unseen variable which is causing first one variable and then another one..." Reichenbach, 1956

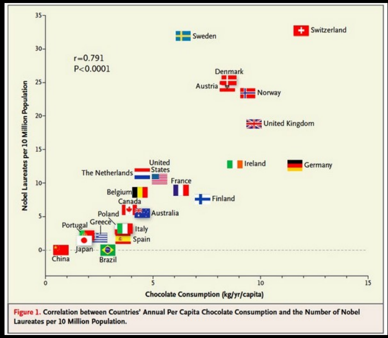
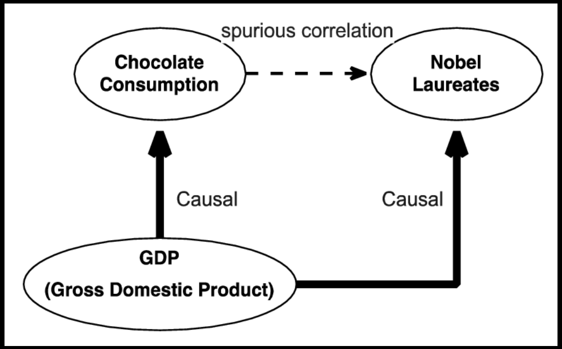
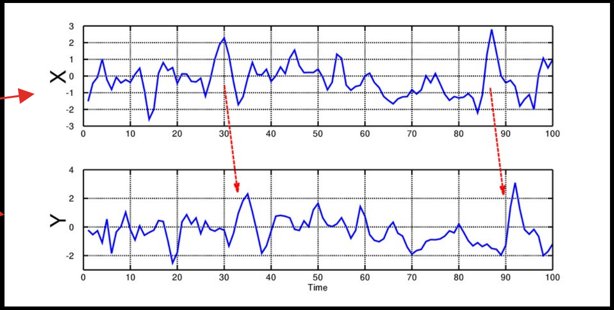


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.



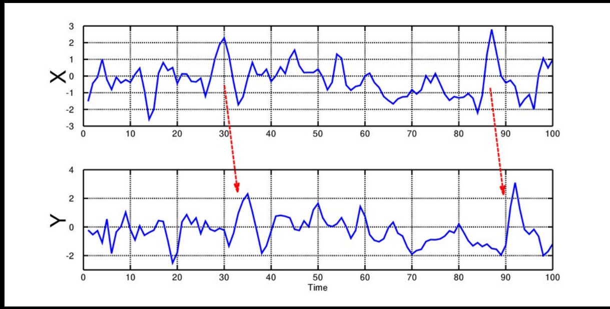
Z



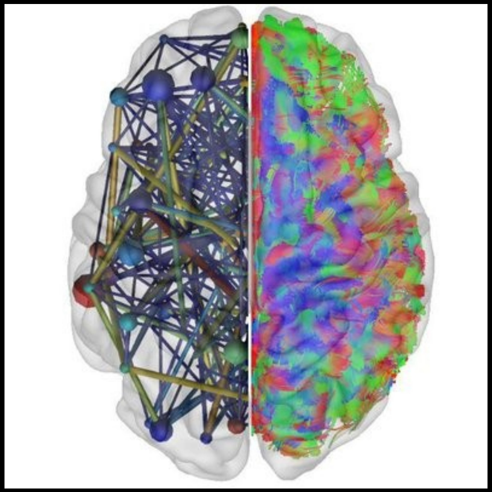
Structurally constrained **Granger causality**

Crimi et al. Neuroimage 2021

$$X(t) = \sum_{\tau=1}^{L} A_\tau X(t-\tau) + \varepsilon(t),$$



B=



Functional

Structural

$$\widehat{\mathbf{A}_\mathbf{i}} = \mathbf{A}_i \odot \mathbf{B} \tag{2}$$

where $\odot$ denotes the Hadamard or element-wise product, and $\mathbf{B}$ is an indicator matrix obtained by the structural connectivity $\mathbf{A}_{\text{init}}$, defined as

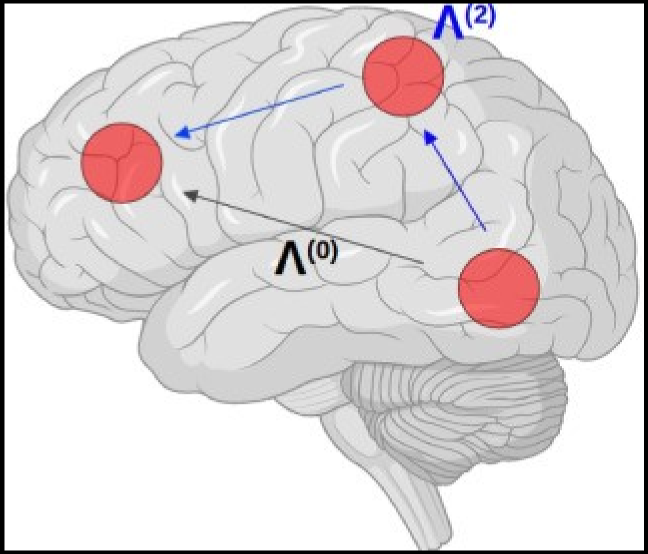
$$B_{uv} = \begin{cases} 0 & \text{if the} uv \text{element of} \mathbf{A}_{\text{init}} \text{is} 0 \\ 1 & \text{otherwise.} \end{cases}$$

Structurally constrained **Granger causality**

Crimi et al. Neuroimage 2021

Model can be expanded as
In the neural propagator of Robinson

$$X(t) = \sum_{\tau=1}^{L} A_\tau X(t-\tau) + \varepsilon(t),$$

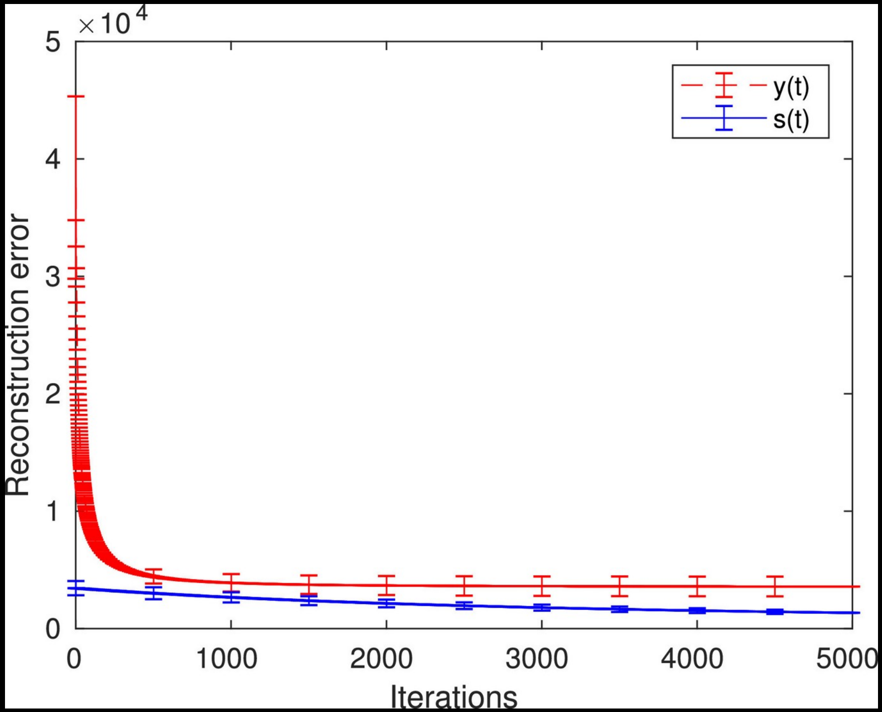


$$E = \frac{1}{2}\sum_{t=n}^{T} \left\| \mathbf{y}(t) - \sum_{i=1}^{n} \left(\mathbf{A}_i^{(0)} \odot \mathbf{B} + \mathbf{A}_\mathbf{i}^{(\mathbf{2})} \odot \mathbf{B}^{(2)} \right) \mathbf{y}(t-i) \right\|_2^2 .$$

Practical considerations I

1. BOLD signal is hemodynamics not neuronal signal (but we can use deconvolution to get closer to the real signal)
 Assuming a common HRF is shared across the various spontaneous point process events at a given voxel, the BOLD signal can be seen as the result of the convolution of neural states s(t) and HRF h(t):

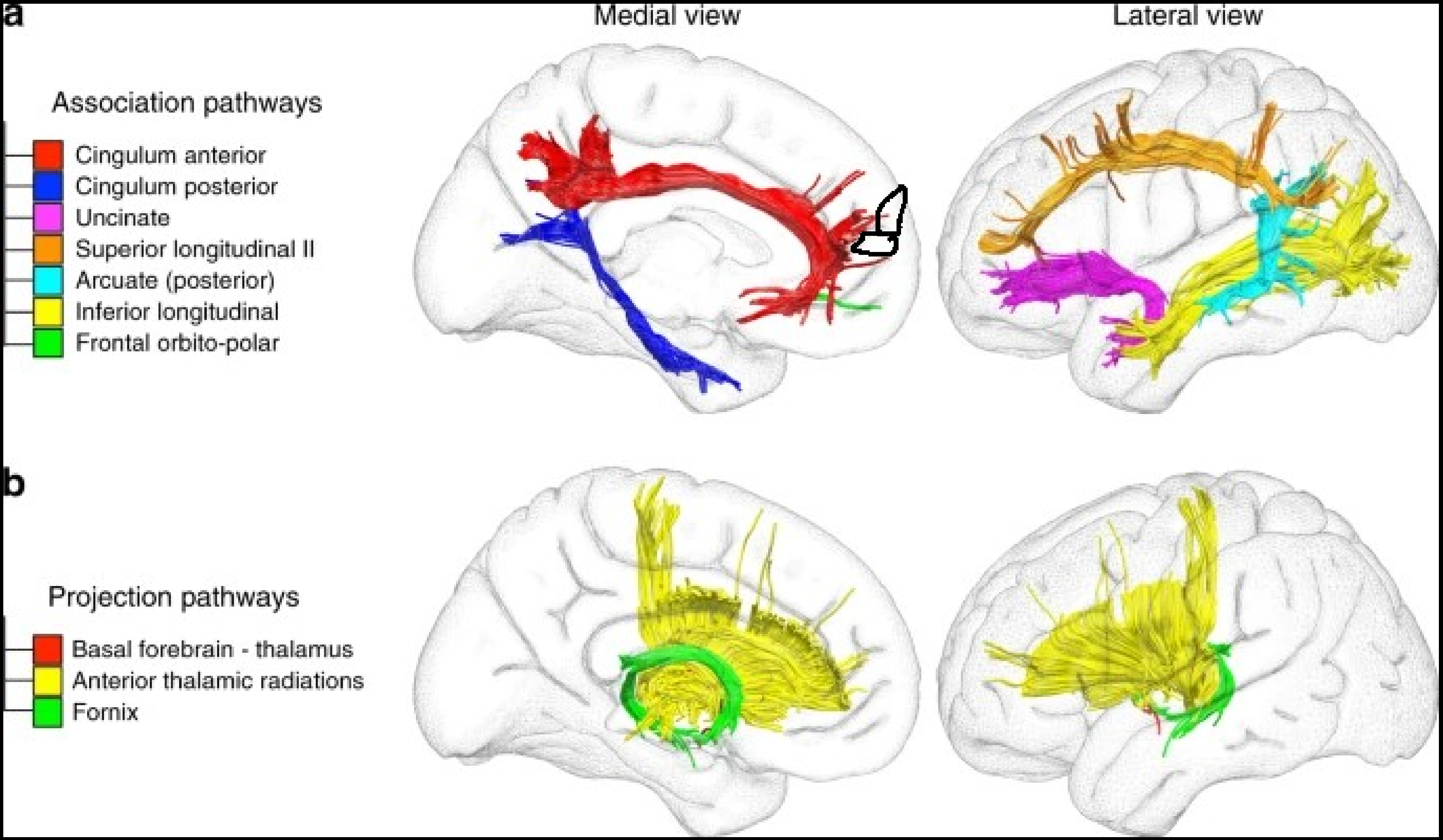
$$y(t) = s(t) \circledast h(t) + \epsilon(t)$$



Lesson learnt:
If you can work
with neural states
not BOLD signal

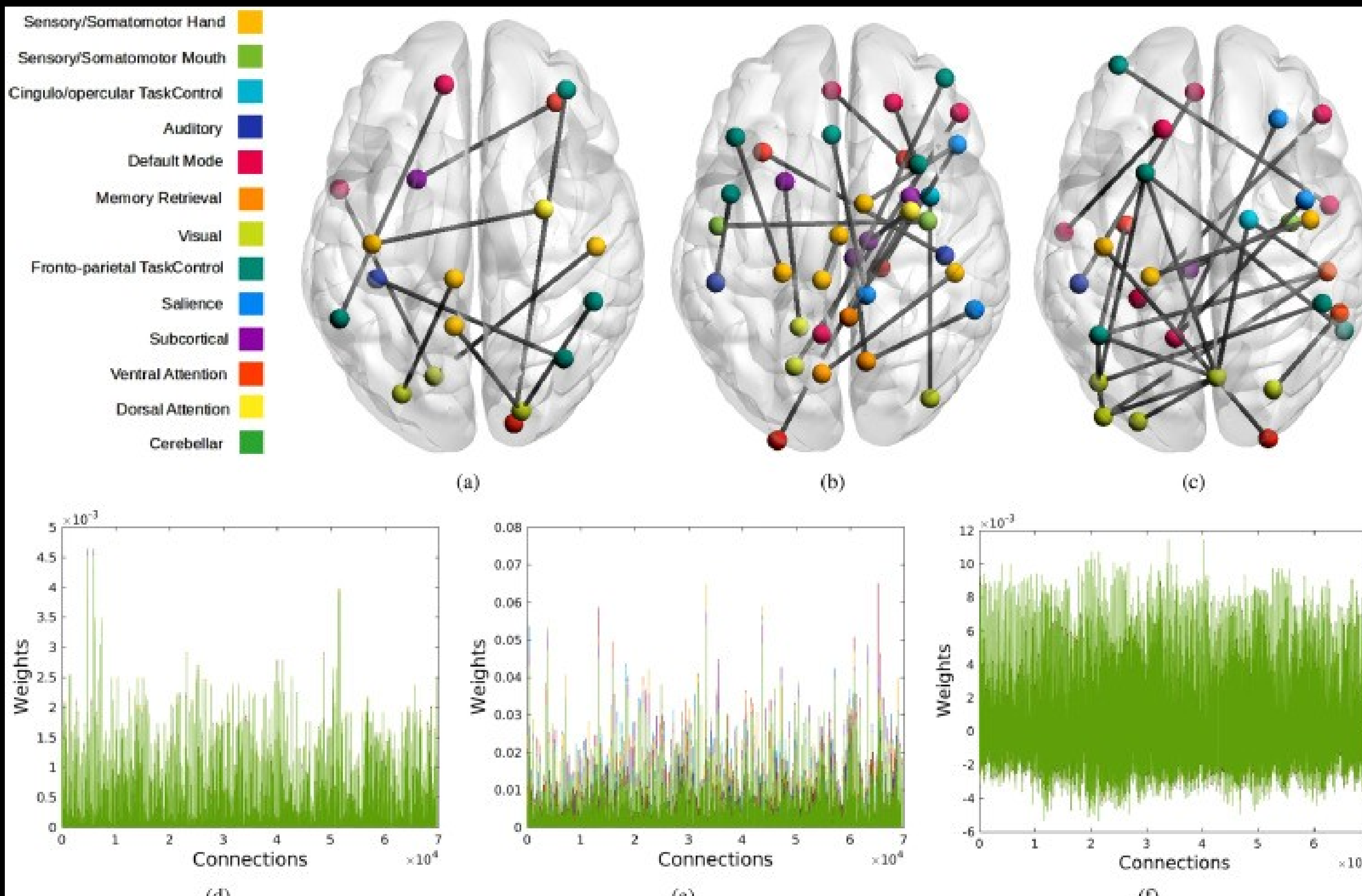
Practical considerations II

Functional ROIs are «far» from white-gray matter boundaries
Lesson learnt: dilate ROI functional Atlas or use tricks in tractography



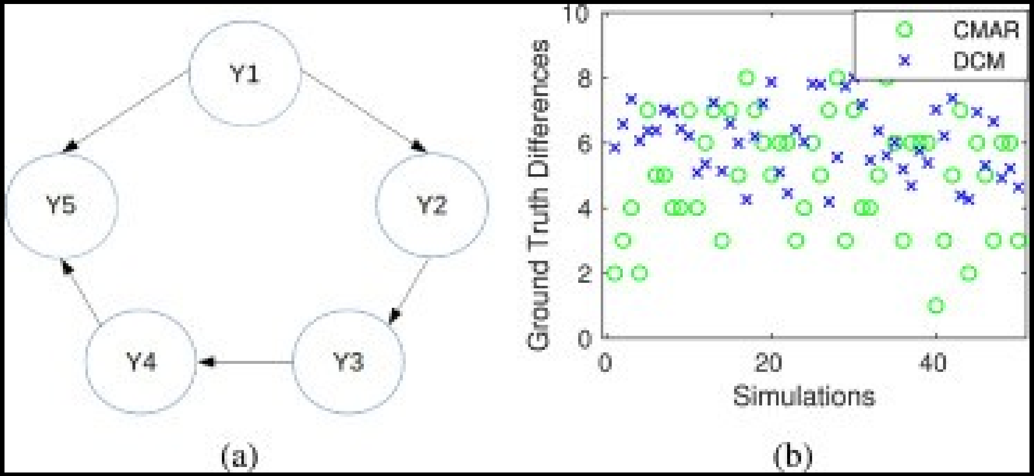
Results I

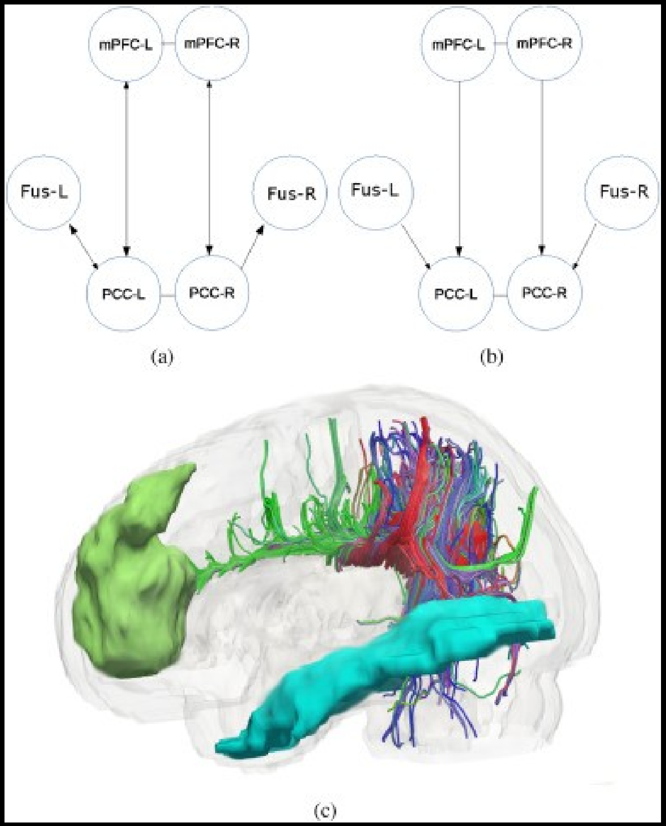
In a classification task (SVM) from ASD dataset better performance were obtained with eff.conn., then func. Con. and lastly structural con.
Discriminant func/eff connection were consistent with Yahata et al. Nat. Com 2016



Results II

More consistent than DCM (simulation and with DMN «ground-truth»)



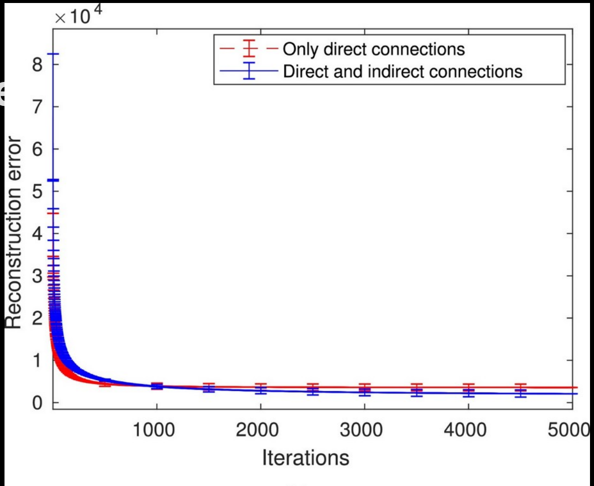


Let's complain about DCM: it is not Good for resting-state, you need to use a Variation of it (stochastic-DCM)

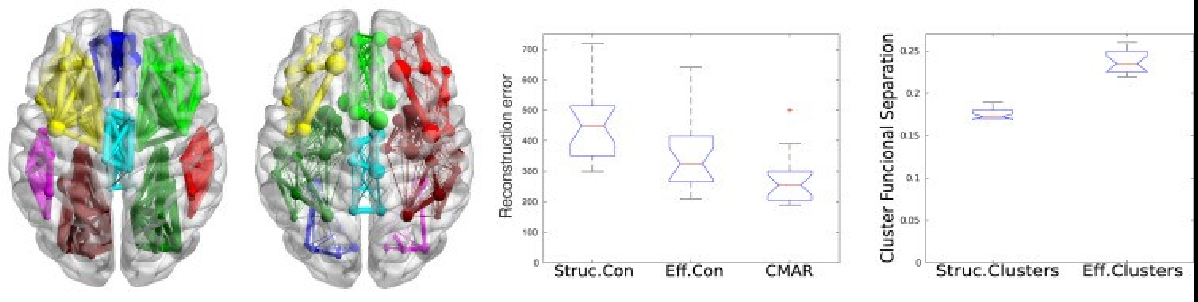
Results III

Lower reconstruction error using also indirect conne

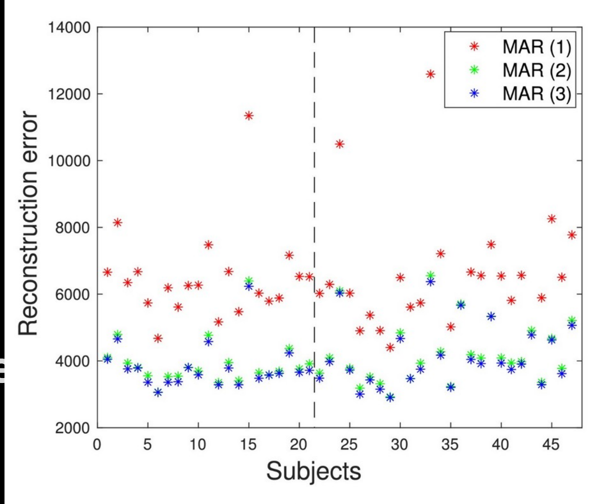
For our signal up to 2 time point back was enough



(a)



(a) (b) (c) (d)



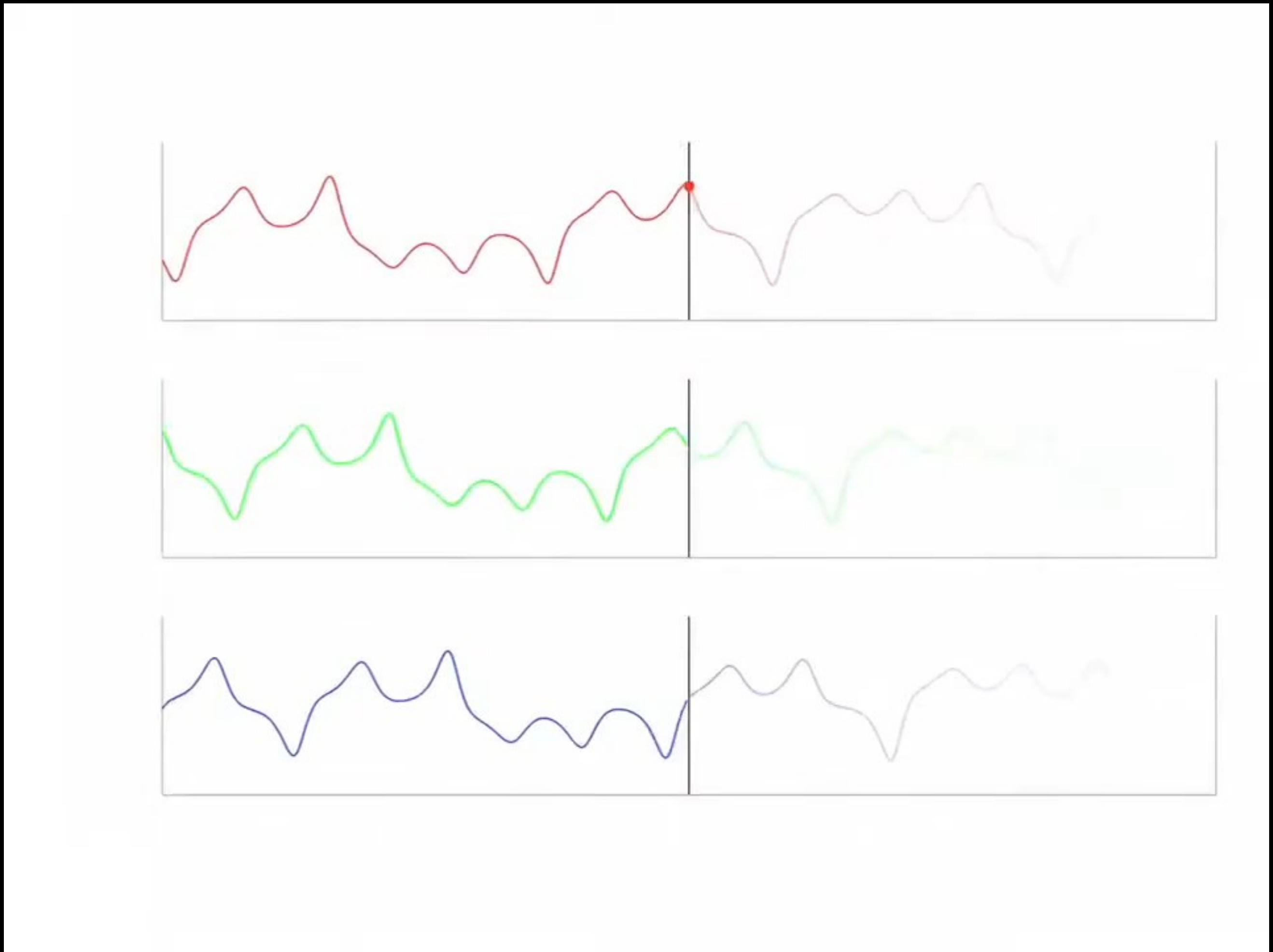
(b)

If we do clustering, clusters are obviously with le nodes

Sugihara Causality

- Perturbation causality cannot be tested in all cases.
- **However, let’s find something better than Granger causality (address the separability of X and Y)**
- if X drives Y, then a record of the X’s influence will be present in the history of Y, and therefore a continuous delay map from Y to X exist
- primarily suited for weakly coupled components of non-linear dynamic systems
- justified by Takens theorems : when two different variables represent different parts of the same dynamical system their shadow manifolds are diffeomorphic to the true attractor and therefore to each other: If x has a causal influence on the dynamics of Y, then X will influence the dynamics of Y. (Sugihara et al. Science 2012).
- This is not an hypothesis testing approach as DCM as it can easily scaled to brain-wide scenarios.

Sugihara Causality



Sugihara Causality

State Space Reconstruction:
Convergent Cross Mapping

A supplemental simulation and animation for
“Detecting Causality in Complex Ecosystems”

George Sugihara, Robert May, Hao Ye, Chih-hao Hsieh,
Ethan Deyle, Mike Fogarty, and Stephan Munch

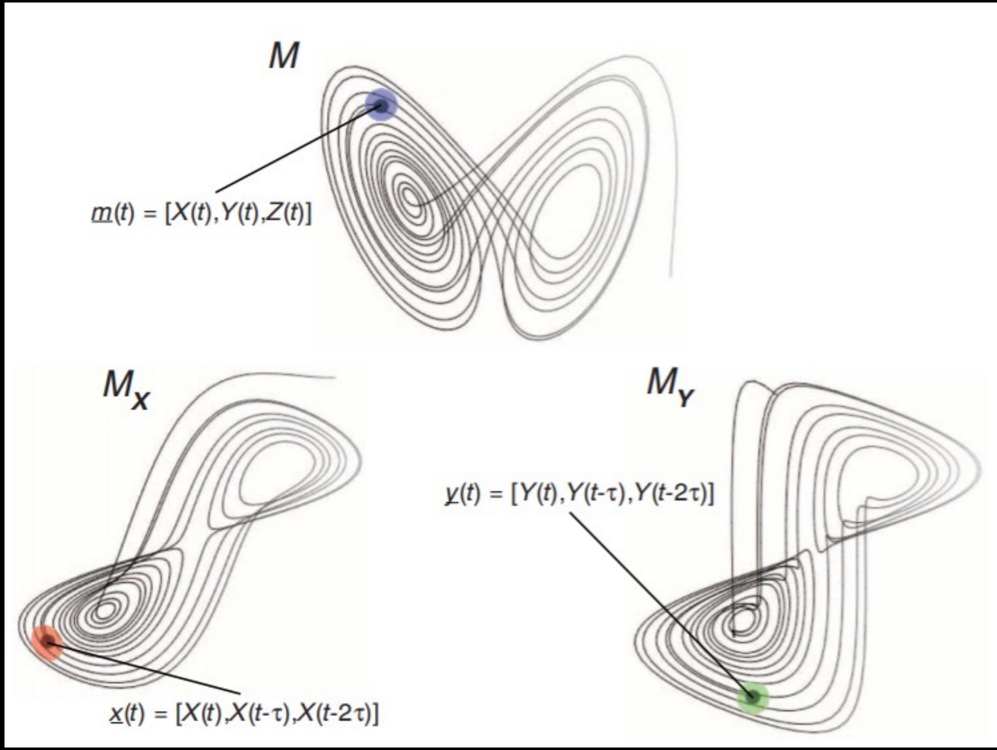
animation by: Hao Ye and George Sugihara

Sugihara Causality

$$\mathbf{x}_t = [x(t-(d-1)),\ \ldots,\ x(t-1),\ x(t)],$$

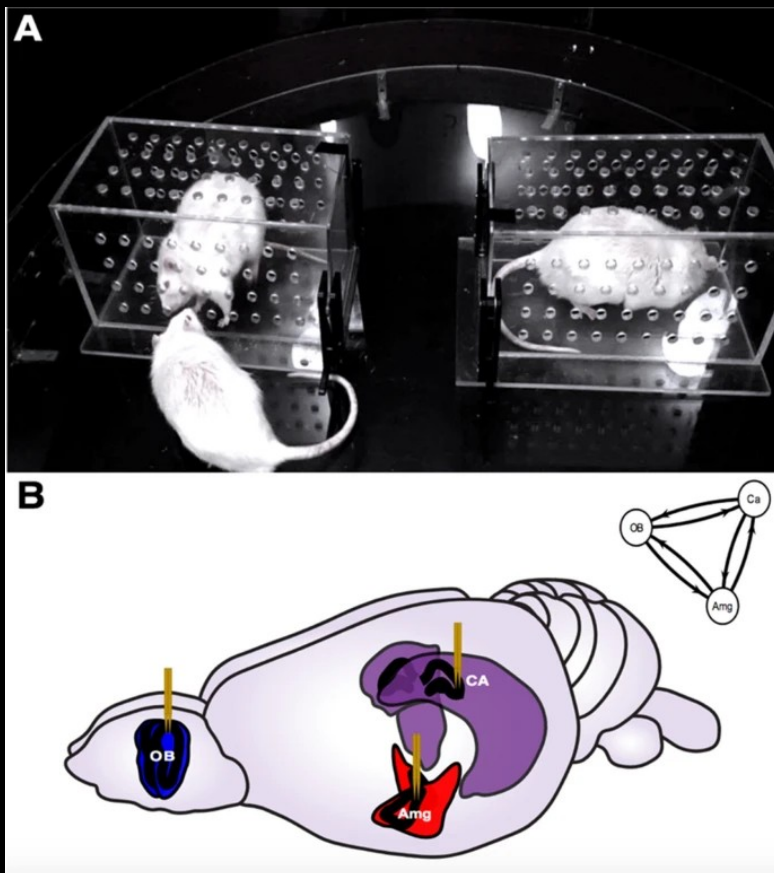
$$\hat{y}(t)|\boldsymbol{M}_{\mathbf{x}} = \sum_{i=1}^{d+1} w_i y(t_i)$$

$$w_i = \frac{e^{-\|\mathbf{x}_t - \mathbf{x}_{t_i}\|^2 / \|\mathbf{x}_t - \mathbf{x}_{t_1}\|^2}}{\sum_{j=1}^{d+1} e^{-\|\mathbf{x}_t - \mathbf{x}_{t_j}\|^2 / \|\mathbf{x}_t - \mathbf{x}_{t_1}\|^2}}$$



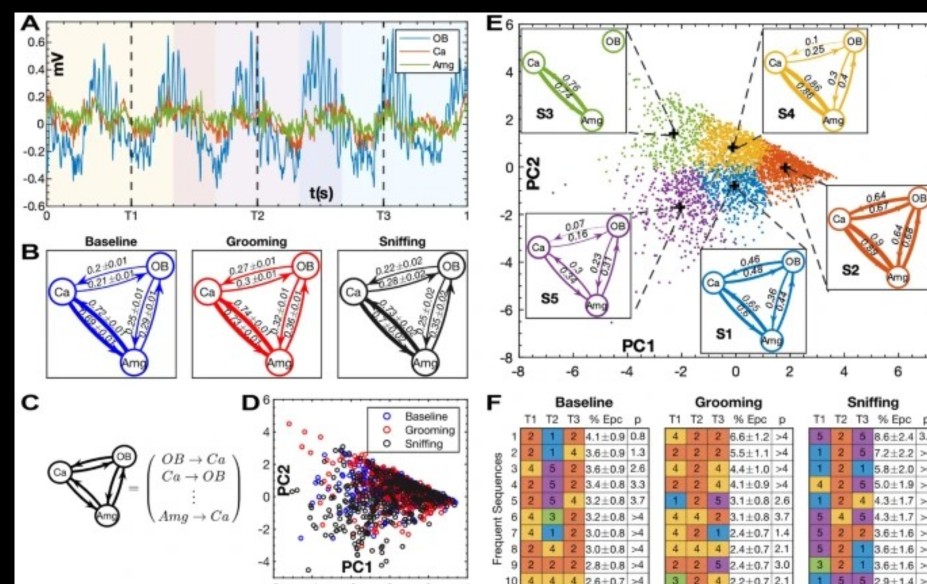
Sugihara et al. Science 2012

Sugihara Causality in Neuroimaging



Rats were surgically implanted with electrodes for electrophysiological recordings in the main olfactory bulb (OB), hippocampus (CA) and medial amygdala (Amg).

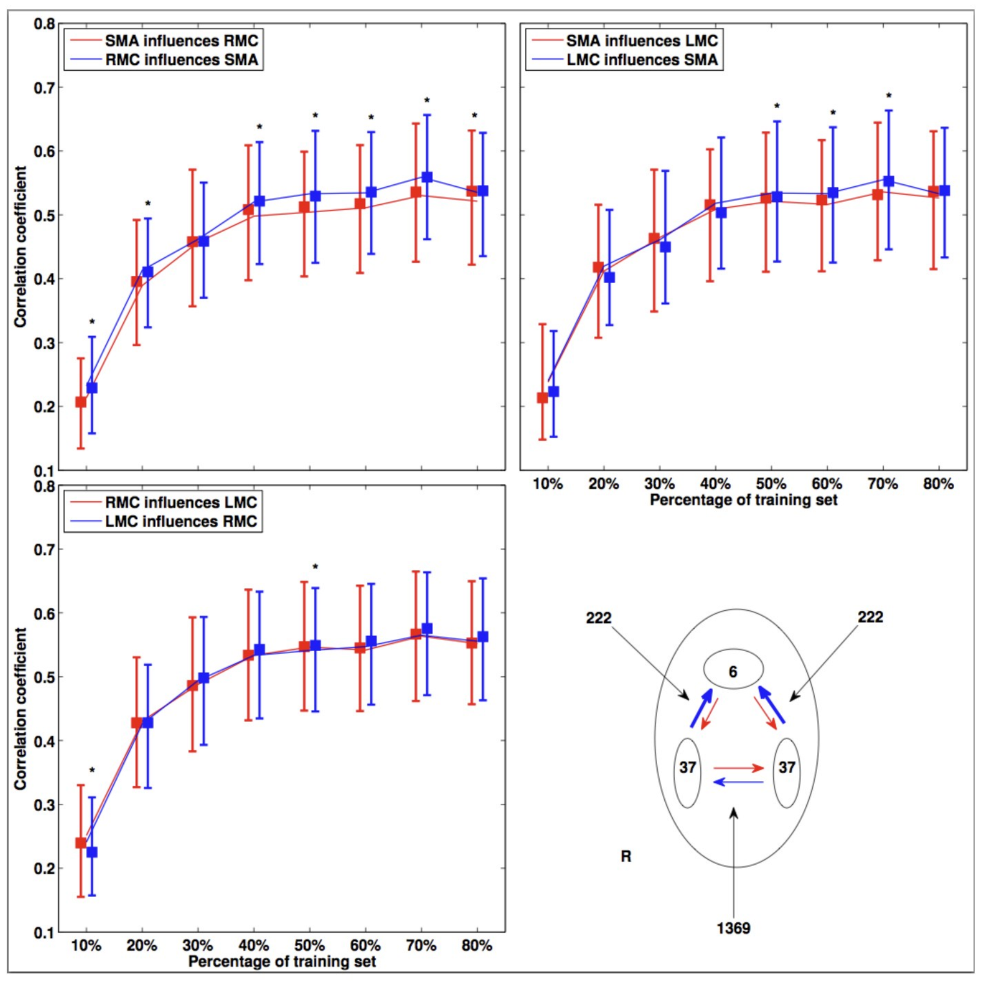
Breston et al. Nature SciRep 2021



Sugihara Causality in Neuroimaging

Tapping experiment

Wisemueller et al. 2019 Biorxiv



supplementary motor area (SMA)

Left somatomotor cortex (LMC)

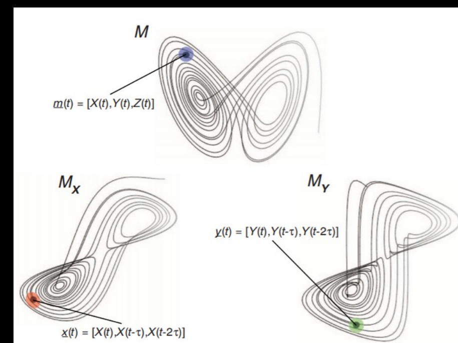
Right somatomotor cortex (RMC)

Structurally constrained Sugihara Causality

$$\mathbf{x}_t = [x(t-(d-1)),\ \ldots,\ x(t-1),\ x(t)],$$

$$\hat{y}(t)|\boldsymbol{M}_{\mathbf{x}} = \sum_{i=1}^{d+1} w_i\, y(t_i)$$

$$w_i = \frac{e^{-\|\mathbf{x}_t-\mathbf{x}_{t_i}\|^2/\|\mathbf{x}_t-\mathbf{x}_{t_1}\|^2}}{\sum_{j=1}^{d+1} e^{-\|\mathbf{x}_t-\mathbf{x}_{t_j}\|^2/\|\mathbf{x}_t-\mathbf{x}_{t_1}\|^2}}$$



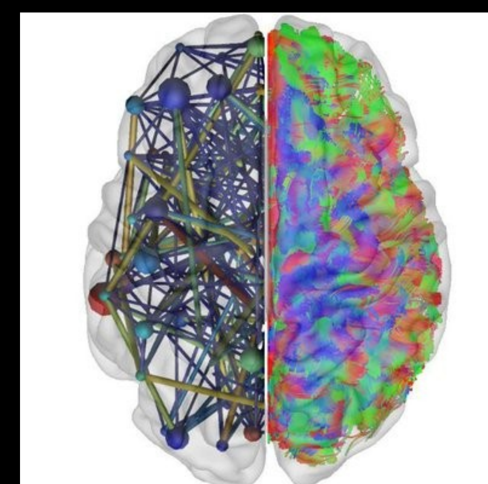
Pairwise but Brainwide

$$\widehat{\mathbf{A}_{\mathbf{i}}} = \mathbf{A}_i \odot \mathbf{B} \qquad (2)$$

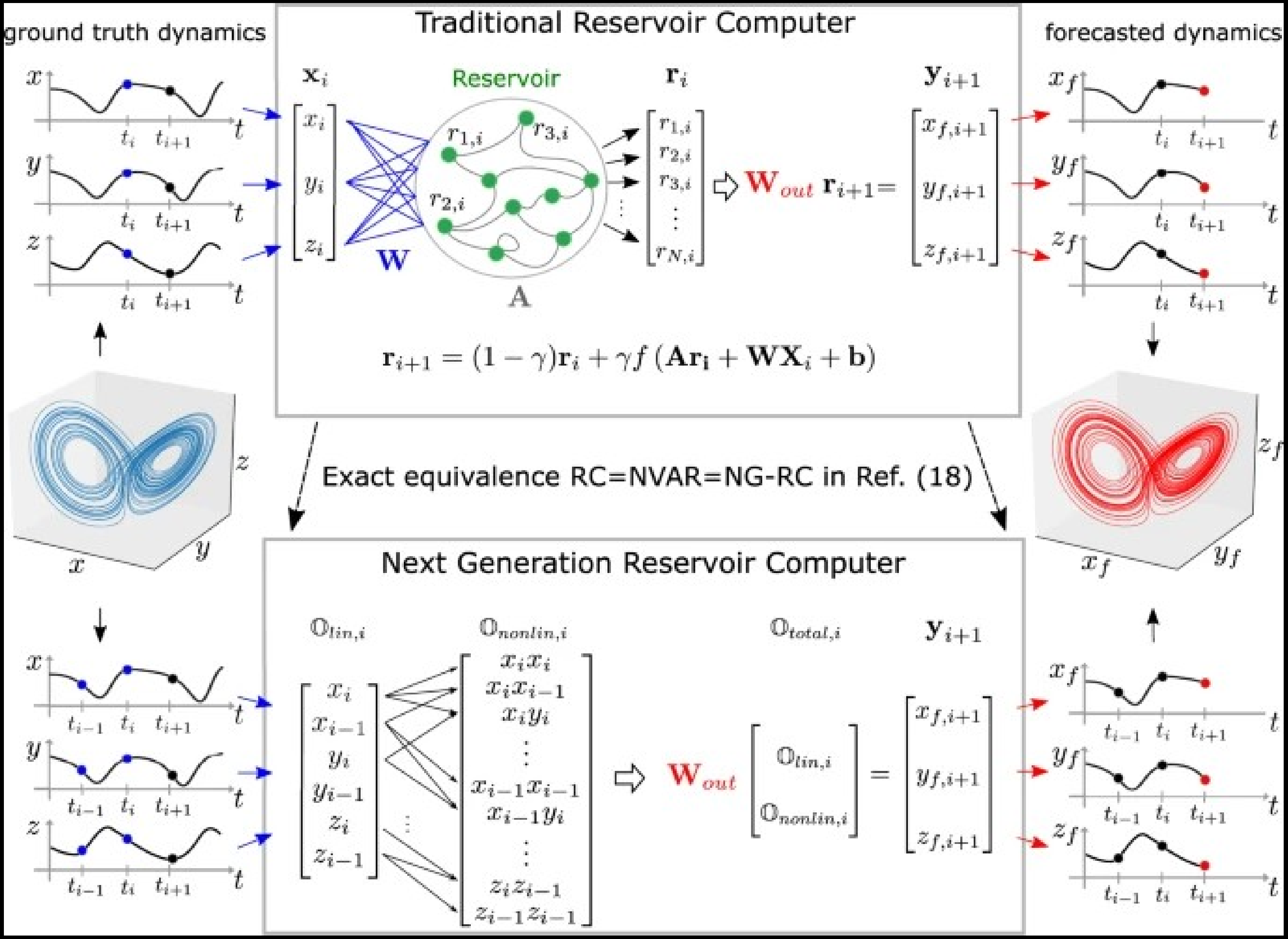
where $\odot$ denotes the Hadamard or element-wise product, and $\mathbf{B}$ is an indicator matrix obtained by the structural connectivity $\mathbf{A}_{\text{init}}$, defined as

$$B_{uv} = \begin{cases} 0 & \text{if the} uv \text{element of} \mathbf{A}_{\text{init}} \text{is} 0 \\ 1 & \text{otherwise.} \end{cases}$$

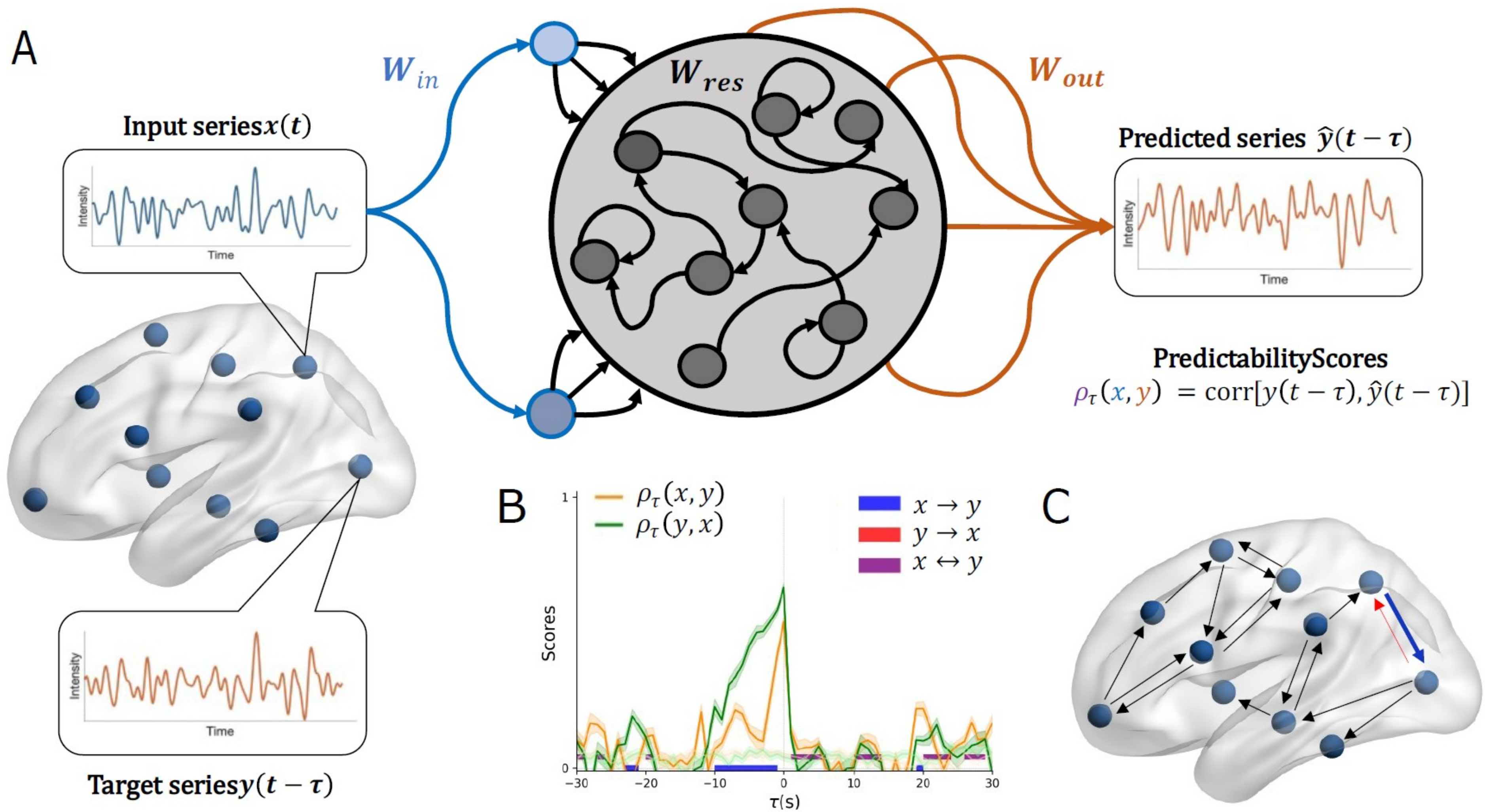
B=



Recurrent neural networks

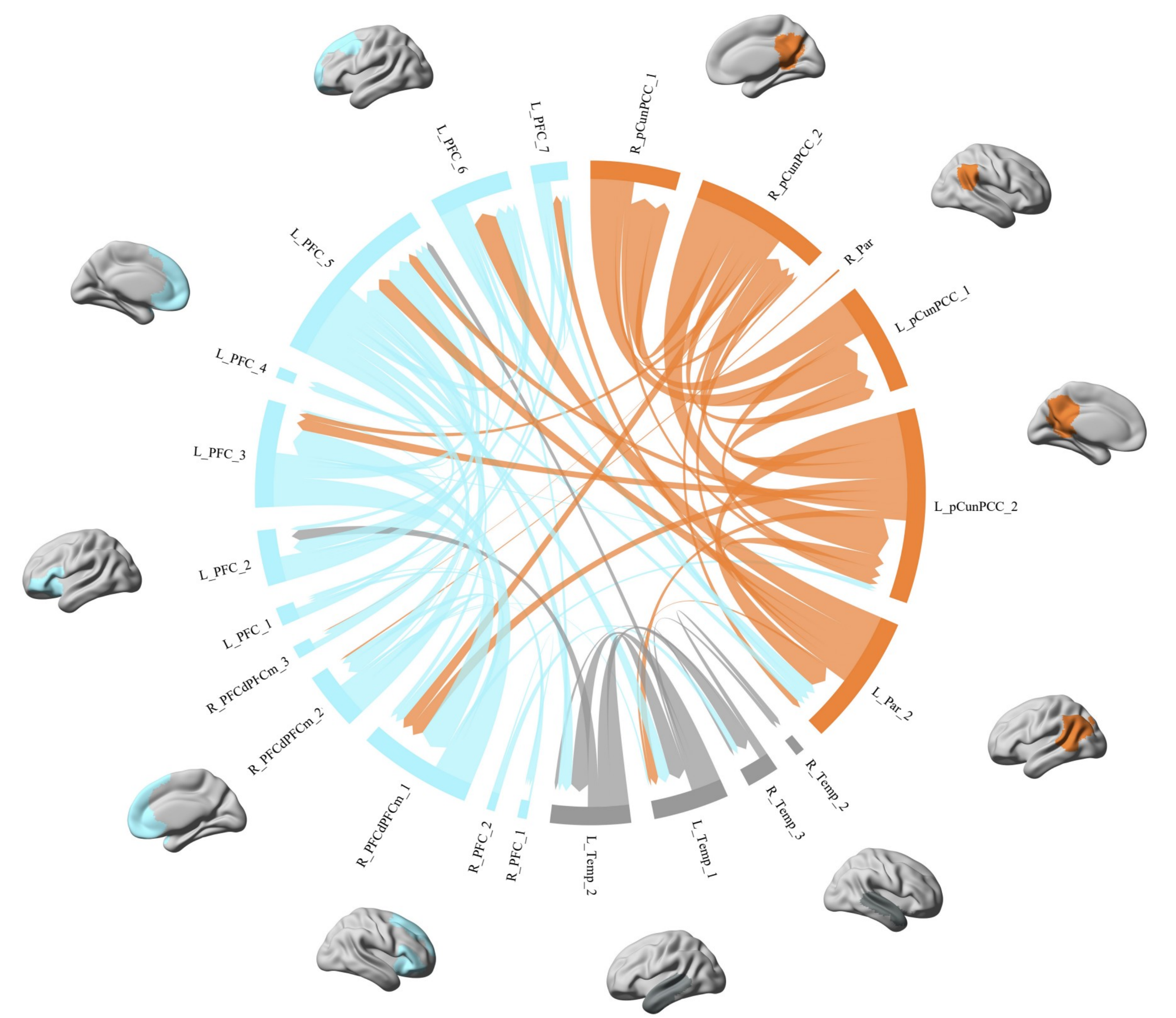


Causality reservoir computing



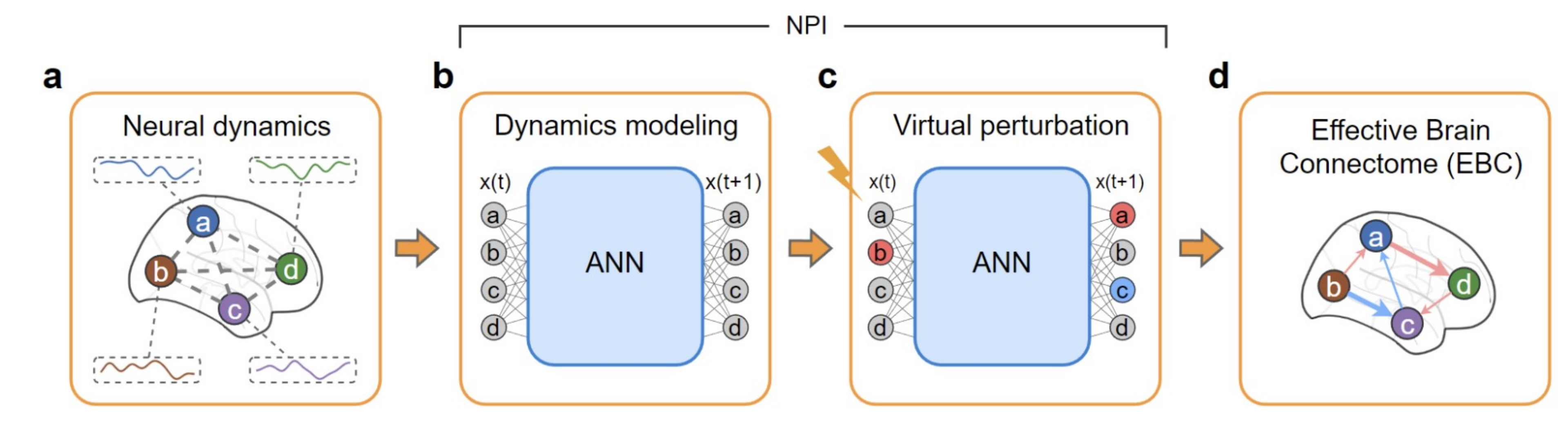
Falco-Roget, Onicas, Crimi Biorxiv 2023

Causality reservoir computing



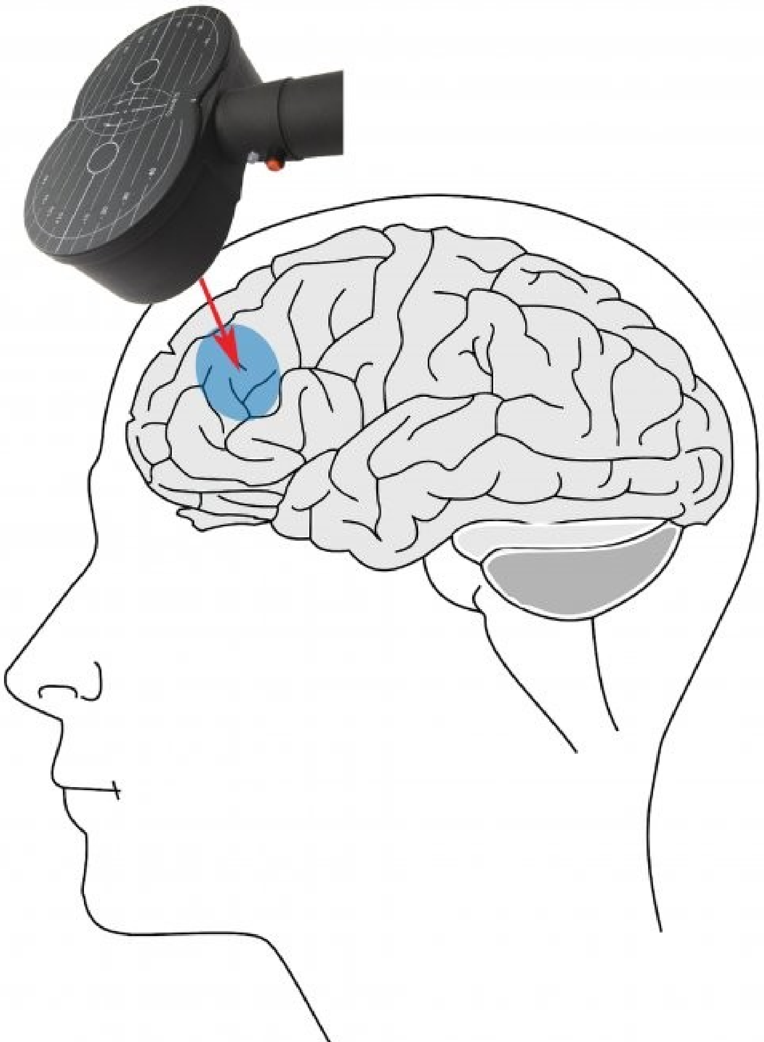
Falco-Roget, Onicas, Crimi Biorxiv 2023

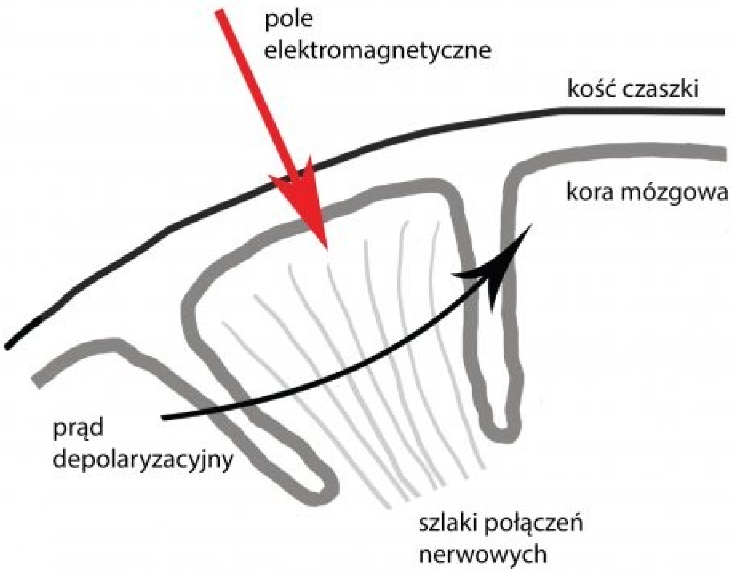
Neural perturbation inference



Luo et al. Arxiv 2022

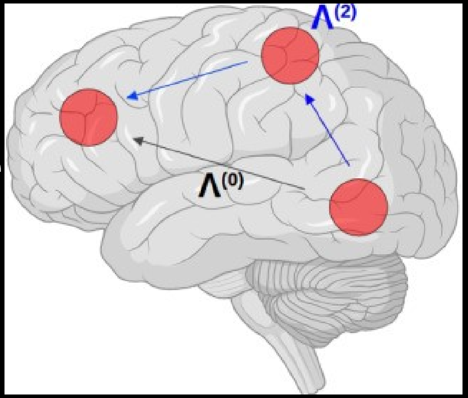
Real perturbation





Conclusions

- Eating chocolate won’t give you the business prize but living in a country with high GDP will raise the chances.
- Siguhara solves the separability issue of Granger causality, it is worth exploring, as there are many other study related to brain state as attractors.



- Reservoir computing address many things
  (though not resolving the unseen variables issue
- Pearl Causality is still holding the best position:
  Construct experiments using TMS or similar.
- Maybe we should give up our obsessions for effective connectivity and simply use functional connectivity biomarkers