Dynamical Causality

陳洛南 (Luonan Chen)
Chinese Academy of Sciences

Aristotle (BC384～322) developed “four causes” to explain change: material cause, formal cause, efficient cause (or moving cause), and final cause (or purpose cause)

Causal Model

- Statistical model (non-time-series data)

(Statistical learning, neural network, machine learning)

Criterion: correlation

- Causal model (non-time-series data)

(Neyman-Rubin structural causal model, Judea Pearl’s causal calculus)

Criterion: independence, do-calculus, OOD

- Dynamical model-1 (time-series data, statistical model)

(Granger causality, transfer entropy)

Criterion: prediction x(t) → y(t+1)

预测

- Dynamical model-2 (time-series data, dynamical model)

(embedding causality, embedding entropy)

Criterion: inverse mapping Y(t) → X(t-1)

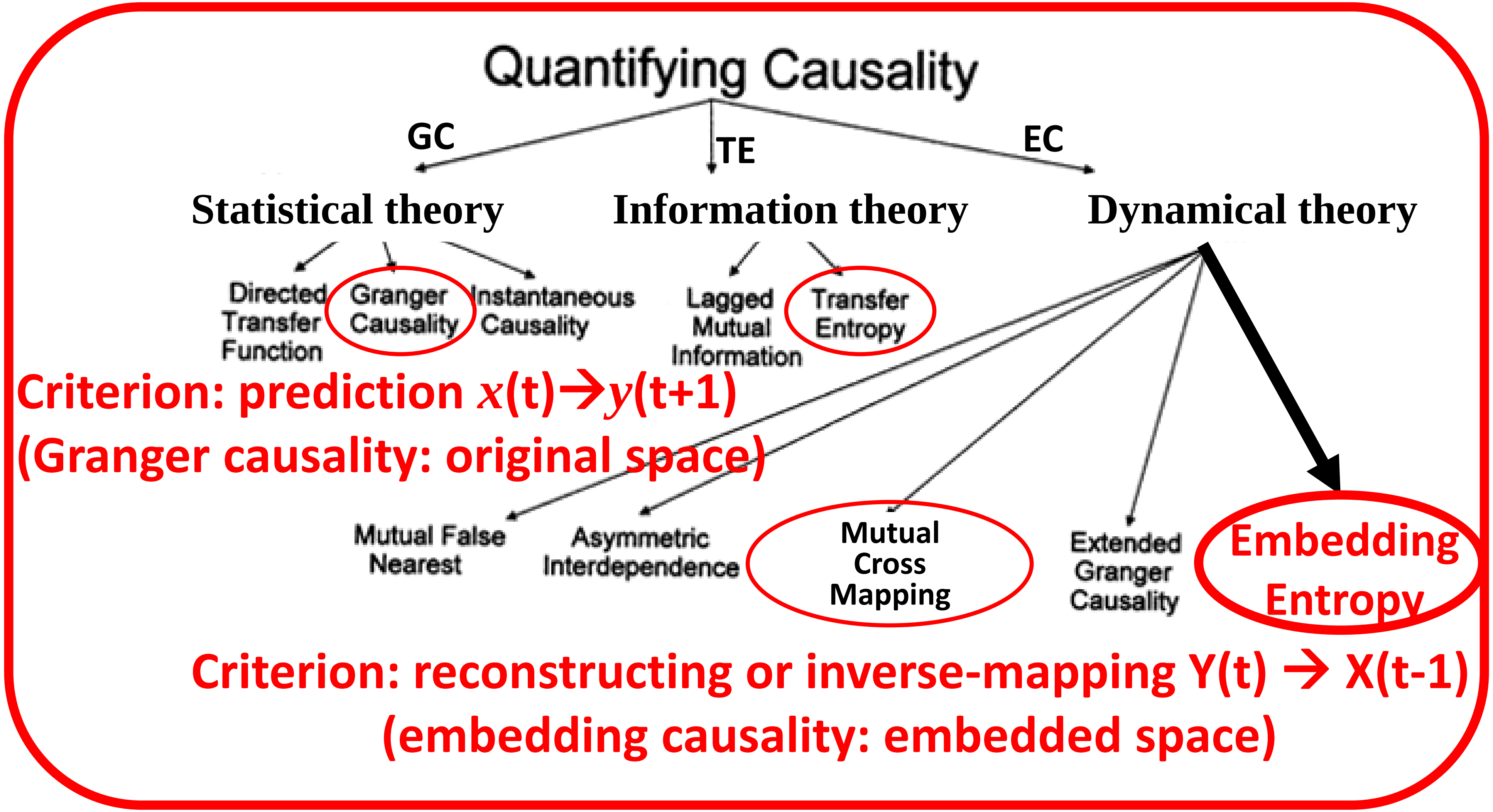
逆重构/逆映射

$$x_{t+1} = f(x_t)$$
$$y_{t+1} = g(x_t, y_t)$$

$X \rightarrow Y$

Existing models for time-course data

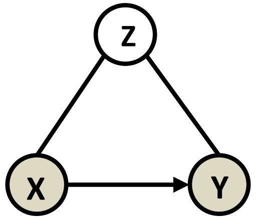
Dynamical Causality: uniform framework



Causality Definition

- Cause: X causes Y if and only if changing X leads to a change in Y while keeping everything else constant.

- Effect: causal effect is the magnitude by which Y is changed by one unit change in X.



New definition from dynamical system

J. R. Soc. Interface 2022

Dynamical Causality (DC)

$$x_t = f\left(x_{t-1}, \dots, x_{t-p}, \dots\right)$$

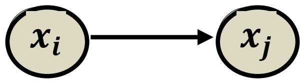
$$x_t = (x_{1,t}, x_{2,t}, \dots, x_{n,t});\ f = (f_1, f_2, \dots, f_n)$$

- **H_1 of DC: $x_i \rightarrow x_j$, dynamics of x_j depends on x_i, i.e.**

  $\exists k \in \{1, 2, \dots, p\}$, such that

  $$\frac{\partial f_j}{\partial x_{i,t-k}} \neq 0$$ **for almost any sampled t**

causal



$x_{j,t}=f(x_{t-1}, x_{i,t-1})$

- **H_0 of DC: dynamics of x_i does not depend on x_j, i.e.**

  $\forall k \in \{1, 2, \dots, p\}$, such that

  $$\frac{\partial f_j}{\partial x_{i,t-k}} = 0$$ **for almost any sampled t**

non-causal

$x_{j,t}=f(x_{t-1} \setminus x_{i,t-1})$

Direct DC : explicit dependence; Indirect DC: implicit dependence

DC includes Granger Causality (GC), Transfer Entropy (TE) and Embedding Causality (EC)

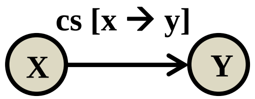
Causal Strength

- H_0: model error with no effect from x to y
- H_1: model error with an effect from x to y

Causal strength: $\mathbf{cs[x \rightarrow y]} = \boldsymbol{D}[\mathbf{H_1} \mid \mathbf{H_0}]$

- cs: causality strength
- *D*: distance function

$$x_{t+1} = f(x_t)$$
$$y_{t+1} = g(x_t, y_t)$$



$$y_{t+1} = f(x_t, y_t)$$

Granger Causality (GC)

如果使用X的信息，Y的预测误差减少：X → Y

Using X reduced the prediction error of Y：X → Y

If $$\sigma(y_{t+1} \mid \Lambda_t, x_t) < \sigma(y_{t+1} \mid \Lambda_t)$$

Λ_t={past terms of y_t , excluding x_t} ; x_t={past terms of x_t}

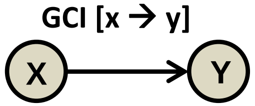
Then $X \rightarrow Y$

x(t) predicts y(t+1)

Nobel Prize

Granger Causal Effect

(original state space: *x*, *y*)



$$y_{t+1} = f(x_t, y_t)$$

- H0: $$y_{t+1} = \sum_{k=0}^{p} a_k y_{t-k} + \varepsilon_{2t}$$

- H1: $$y_{t+1} = \sum_{k=0}^{p} a_k y_{t-k} + \sum_{k=0}^{p} b_k x_{t-k} + \varepsilon_{1t}$$

- Causal strength: $CS_{x \rightarrow y} = \ln\varepsilon_2^2 - \ln\varepsilon_1^2$

cs[x$\longrightarrow$y] = D[H1 | H0]

requirement on separability of *X* and *Y;* cutoff error of *p*

Granger Causality (GC)

If ***X*** past values provide statistically significant information about ***Y*** future, ***X*** is said to Granger-cause ***Y***.

Questions ?

It is applicable for a strong associated network

– Linear causality

– **Non-separability of *X* and *Y* (false negative or H_0 underestimation)**

– **cutoff error for *p***

– Weak/moderate coupling

– indirect causality

Criterion: prediction

Transfer Entropy (TE)

(original state space: *x,y*)

$$X \rightarrow Y$$

$$y_{t+1} = f(x_t, y_t)$$

如果使用X的信息，Y的不确定性减少：X→Y

Using X reduced the uncertainty of Y：X→Y

- Granger causality: linear
- Transfer entropy: nonlinear

cs[x→y] = D[H1 | H0]

$$T_{X\to Y} = H\left(Y^t \mid Y^{t-1}\right) - H\left(Y^t \mid Y^{t-1}, X^{t-1}\right),$$

$$= MI(Y^t, X^{t-1} \mid Y^{t-1}) \quad > 0 \text{ (causal)}$$

Granger causality is the linear version of transfer entropy.

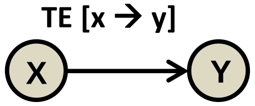
Question? 1. Separability (false negative or H0 underestimation),

2. Strong association problem,

3. Cutoff error of p

x(t) predicts *y*(t+1)

TE Causal Effect



$$\boldsymbol{y}_{t+1} = \boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{y}_t)$$

- H0: $H(\varepsilon_{2t}) = H(y_t | y_{t-1}, \dots, y_{t-p})$

- H1: $H(\varepsilon_{1t}) = H(y_t | y_{t-1}, \dots, y_{t-p}; x_{t-1}, \dots, x_{t-p})$

- Causal strength: $TE_{x \rightarrow y} = |\mathrm{H}(\varepsilon_1)\text{-}\mathrm{H}(\varepsilon_2)|$

separability of *X* and *Y* (false negative or H0 underestimation), strong association problems !
cutoff error problem of *p* !

Transfer entropy measures the decrease of uncertainty by considering y

Embedding Causality

(Delay embedding space: *X*,*Y*)

Y的邻居也是对应X空間的邻居: **X → Y**

- Original state space → Delay embedding space

Advantages: no requirement on separability, nonlinear, no cut-off

Criterion: Inverse-mapping (逆映射)

Y(t) inverse-maps to *X*(t-1)

从原空间到嵌入空间的时间序列

Delay Embedding

From original space x to embedded space X

Causal inference at embedded space, not at original space !

Original space in R : $x_1, x_2, \ldots, x_t, \ldots$

$$\downarrow \quad \downarrow \quad \quad \downarrow$$

Embedded space in R^L: $X_1, X_2, \ldots, X_t, \ldots$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_L \end{bmatrix}, \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_{L+1} \end{bmatrix}, \ldots, \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_{t+L-1} \end{bmatrix}, \ldots$$

X_t represents the state of whole system, including information of other variable y_t

PNAS

Randomly distributed embedding making short-term high-dimensional data predictable

Huanfei Ma[a], Siyang Leng[b,c,d], Kazuyuki Aihara[b,e,1], Wei Lin[c,d,f,g,h,1], and Luonan Chen[i,j,k,l,1]

[a]School of Mathematical Sciences, Soochow University, Suzhou 215006, China; [b]Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan; [c]School of Mathematical Sciences, Fudan University, Shanghai 200433, China; [d]Center for Computational Systems Biology, Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai 200433, China; [e]International Research Center for Neurointelligence, The University of Tokyo Institutes for Advanced Study, The University of Tokyo, Tokyo 113-0033, Japan; [f]Research Institute of Intelligent and Complex Systems, Fudan University, Shanghai 200433, China; [g]Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education, Shanghai 200433, China; [h]Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, Shanghai 200433, China; [i]Key Laboratory of Systems Biology, Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai 200031, China; [j]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China; [k]School of Life Science and Technology, ShanghaiTech University, Shanghai 200031, China; and [l]Shanghai Research Center for Brain Science and Brain-Inspired Intelligence, Shanghai 201210, China

One variable can possibly describe another variable !

(*Y may include the information of X*)

Cross Mapping

Application to many areas, i.e. economics, ecological systems, bio-systems, AI

H0 is not appropriate in GC and TE

- One variable y can topologically construct whole dynamics, and thus also includes the information of x, from the embedding theorem.

H0 : assume no effect from x to y

$$y_{t+1} = \sum_{k=0}^{p} a_k y_{t-k} + \varepsilon_{2t}$$

incorrect due to its past time points !

$$X \longrightarrow Y$$

$$y_{t+1} = f(x_t, y_t)$$

The separability of *X* and *Y* is generally not satisfied !

Problem-1 ?

Non-separability problem for GC and TE

- The information of one variable could be generically entangled with the whole system.
- H_1 data should not be used to construct H_0 model.

GC and TE

H_1 hypothesis with H_1 data

$$x_{t+1} = \sum_{k=0}^{p} c_k x_{t-k} + \varepsilon_{2t}$$

$$y_{t+1} = \sum_{k=0}^{p} a_k y_{t-k} + \sum_{k=0}^{p} b_k x_{t-k} + \varepsilon_{1t}$$

$\Rightarrow$

H_0 hypothesis but with H_1 data

$$x_{t+1} = \sum_{k=0}^{p} c_k x_{t-k} + \varepsilon_{2t}$$

$$y_{t+1} = \sum_{k=0}^{p} a_k y_{t-k} + \varepsilon_{1t}$$

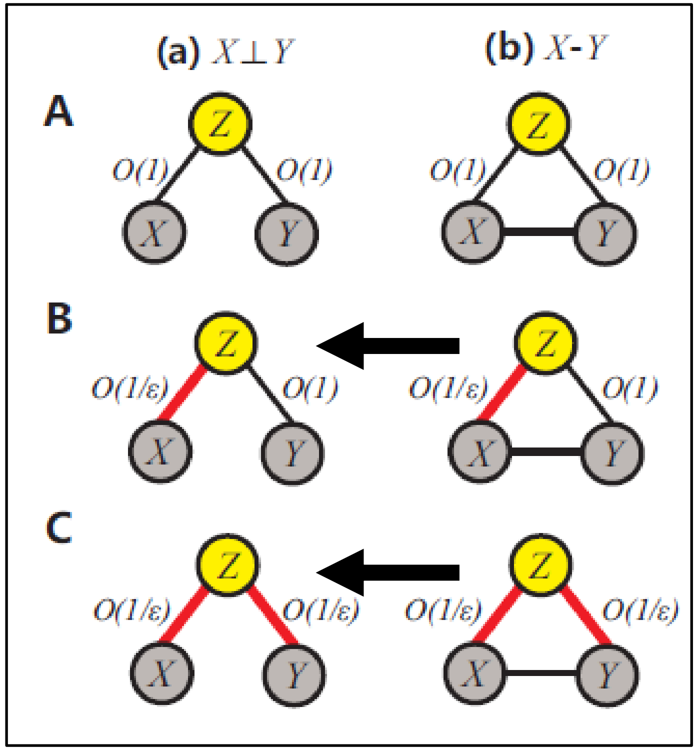
Summary of non-separability problem:

- Causal variables could be both explicit and implicit in the dynamics of the effect variable, and thus only removing the explicit causal variable to construct a H_0 model is improper and actually influence the dynamical behaviors of the whole system.
- The data sampled from H_1 should not be used to approximate the H_0 model because their dynamics could be different.

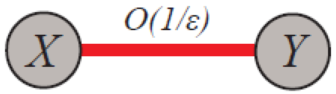
False negative problem!

Problem-2 ?

Strong association problem for TE



Strong associated network



$$p(x, \boldsymbol{y}) = \frac{1}{Z_\epsilon} \exp\Big(-\frac{1}{\epsilon}\phi(x, \boldsymbol{y}) - \psi(x, \boldsymbol{y}) \Big)$$

$$CMI(X, Y|\boldsymbol{Z}) = CMI_0 + \epsilon CMI_1 + \epsilon^2 CMI_2 + O(\epsilon^3)$$

A: $\mathrm{CMI}(X, Y|Z) \sim O(1)$

TE

B: $\mathrm{CMI}(X, Y|Z) \sim O(\varepsilon)$

C: $\mathrm{CMI}(X, Y|Z) \sim O(\varepsilon^2)$

Scale error

False negative problem!

Zhao et al. ***PNAS*** *2016*
Shi et al. ***IEEE/ACM TCBB*** *2018*
Shi et al. ***Sci China Math*** *2018*

F1000Prime

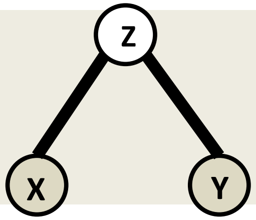
Part mutual information for quantifying direct associations in networks

Zhao et al. PNAS 2016

Juan Zhao[a,1], Yiwei Zhou[a,b,1], Xiujun Zhang[a], and Luonan Chen[a,b,c,2]

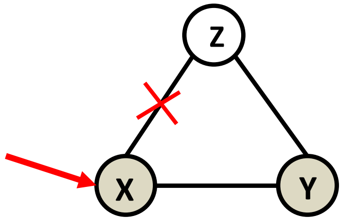
[a]Key Laboratory of Systems Biology, Innovation Center for Cell Signaling Network, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, University of the Chinese Academy of Sciences, Shanghai 200031, China; [b]School of Life Science and Technology, ShanghaiTech University, Shanghai 200031, China; and [c]Collaborative Research Center for Innovative Mathematical Modelling, Institute of Industrial Science, University of Tokyo, Tokyo 113-8654, Japan

Zhao et al. ***PNAS*** *2016 (one side)*
Shi et al. ***IEEE/ACM TCBB*** *2018 (two sides)*
Shi et al. ***Sci China Math*** *2018 (two sides)*



PMI algorithm and kernel PMI algorithm for network inference

Do运算: 后门规则
Back-door criterion



$$p^*(y|z) = \sum_{y} p(y|z,x)p(x)$$

Statistical model cannot mathematically prove that y changes accordingly if x changes, and thus validation experiment is generally required

Problem-3 ?

Cutoff error problem for p for GC and TE

$$\boldsymbol{x}_t = \boldsymbol{f}(\boldsymbol{x}_{t-1}, \dots, \boldsymbol{x}_{t-p}) + \epsilon_{x,t}$$
$$\boldsymbol{y}_t = \boldsymbol{g}(\boldsymbol{x}_{t-1}, \dots, \boldsymbol{x}_{t-p}, \boldsymbol{y}_{t-1}, \dots, \boldsymbol{y}_{t-p}) + \epsilon_{y,t}$$

Cutoff errors : ϵ_x, ϵ_y

Dynamical Causality (DC)

Dynamical System: $$x_t = f(x_{t-1}, \dots, x_{t-p}, \dots)$$

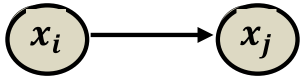
$$\text{with } x_t = (x_{1,t}, x_{2,t}, \dots, x_{n,t}); f = (f_1, f_2, \dots, f_n)$$

- **Dynamical Causality: $x_i \to x_j$, dynamics of x_j depends on x_i, i.e.**

  $\exists k \in \{1, 2, \dots, p\}$, such that

  $$\frac{\partial f_j}{\partial x_{i,t-k}} \neq 0$$ **for almost any sampled t**

causal



$$x_{j,t} = f(x_{t-1}, x_{i,t-1})$$

- **Non-Dynamical Causality: dynamics of x_i does not depend on x_j, i.e.**

  $\forall k \in \{1, 2, \dots, p\}$, such that

  $$\frac{\partial f_j}{\partial x_{i,t-k}} = 0$$ **for almost any sampled t**

non-causal

$$x_{j,t} = f(x_{t-1} \setminus x_{i,t-1})$$

Direct DC: explicit dependence; Indirect DC: implicit dependence

DC includes Granger Causality (GC), Transfer Entropy (TE) and Embedding Causality (EC)

Causal representation

$X \rightarrow Y$

$$y_t = f(y_{t-1}, \dots, y_{t-p}, x_{t-1}, \dots, x_{t-p}; \varepsilon_y)$$
$$x_t = g(x_{t-1}, \dots, x_{t-p}; \varepsilon_x)$$

where ε_y is all remaining terms of x_{t-i}, y_{t-j}; ε_x is all remaining terms of x_{t-i}

$\Downarrow$

$$0 = -y_t + f(y_{t-1}, \dots, y_{t-p}, x_{t-1}, \dots x_{t-p}; \varepsilon_y) = F(y_t, \dots, y_{t-p}, x_{t-1}, \dots x_{t-p}; \varepsilon_y)$$

hence → $0 = F(Y_t, X_{t-1}; \varepsilon_y)$

where $\boldsymbol{Y_t} = (\boldsymbol{y_t}, \dots, \boldsymbol{y_{t-p}})$; $\boldsymbol{X_{t-1}} = (\boldsymbol{x_{t-1}}, \dots, \boldsymbol{x_{t-p}})$

(x,y)观测空间2p+1, 流形自由度d, 所以约束2p+1-d; 当 2p+1-d> p (y的个数)时可用隐函数定理

Thus, from implict function theorem for F: $X_{t-1} = H(Y_t, \varepsilon_y)$ with $\varepsilon_y \neq 0$

y是果，包含x的信息，可重构(x,y)全空间，所以当p>2d时，Y可以表征X的空间

Further, from Takens' embedding theorem for f: $X_{t-1} = H(Y_t)$ with $\varepsilon_y = 0$

(Eliminating $\boldsymbol{\varepsilon_y}$ if $p+1 > 2d$; note generally $Y_t \neq G(X_{t-1})$)

$\Downarrow$

Therefore, we have $X_{t-1} = H(Y_t)$ for numerical causal inference

(functional continuity of H → neighbor cross-mapping)

Current effect reconstructs past cause (现在的果重建过去的因)

Key observations

Past cause reconstructs current effect (过去的因重建现在的果)

Prediction function by time-series data in original space (x_t, y_t)　　$y_t = H(y_{t-1}, x_{t-1})$ 预测

is equivalent to

Current effect reconstructs past cause (现在的果重建过去的因)

Inverse-mapping function by cross-mapping data in delay embedding space (X_{t-1}, Y_t)　　$X_{t-1} = H(Y_t)$ 逆重构

$$X_{t-1} = (x_{t-1}, x_{t-2}, \ldots, x_{t-p});\ Y_t = (y_t, y_{t-1}, \ldots, y_{t-p})$$

Benefits: solving non-separability, strong association, cut-off problems !

One-dimensional case

An example of inverse-mapping

仅用Y能描述X,则X→Y

Model

$$x(t+1) = c\,x(t) + d$$
$$y(t+1) = a\,y(t) + b\,x(t)$$

$$x(t) = y(t+1)/b - ay(t)/b \rightarrow x(t-1) = x(t)/c - d/c = c\,y(t+1)/b - ac\,y(t)/b - d/c$$

or $\mathbf{X(t-1)} = [x(t), x(t-1)] = [1/b, -a/b\ ; c/b, -ac/b]\ [y(t+1), y(t)] + [0, -d/c] = \mathbf{AY(t)} + [0, -d/c]$

We can represent x(t+1) in terms of y !

(finite terms of y → cross-map to x ; *x*(t+1) = *f*(*y*(t+1),*y*(t),…))

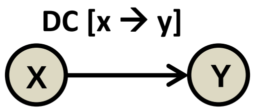
But we cannot represent y(t+1) in terms of x !

(finite terms of x ≠ cross-map to y ; *y*(t+1) ≠ *f*(*x*(t+1),*x*(t),…))

y has all information on x

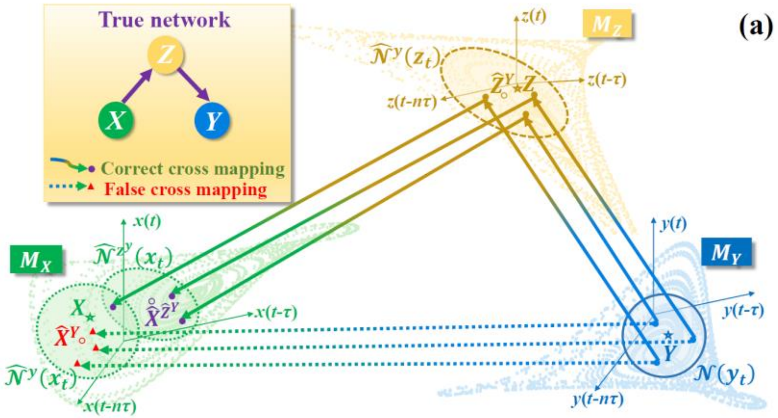
but x has no information on extra dynamics of y (i.e. a y(t))

Dynamical Causal Effect



- H0: Corr(X(t),X^Y(t)|H0)=0

- H1: Corr(X(t),X^Y(t)|H1)=1



- Causal strength: Corr(X(t),X^Y(t)|X^{Z^Y}(t))

$$\varrho_D = \left|\mathrm{Pcc}(X, \hat{X}^Y | \hat{X}^{\hat{Z}^Y})\right|$$

No-requirement on separability of *X* and *Y*

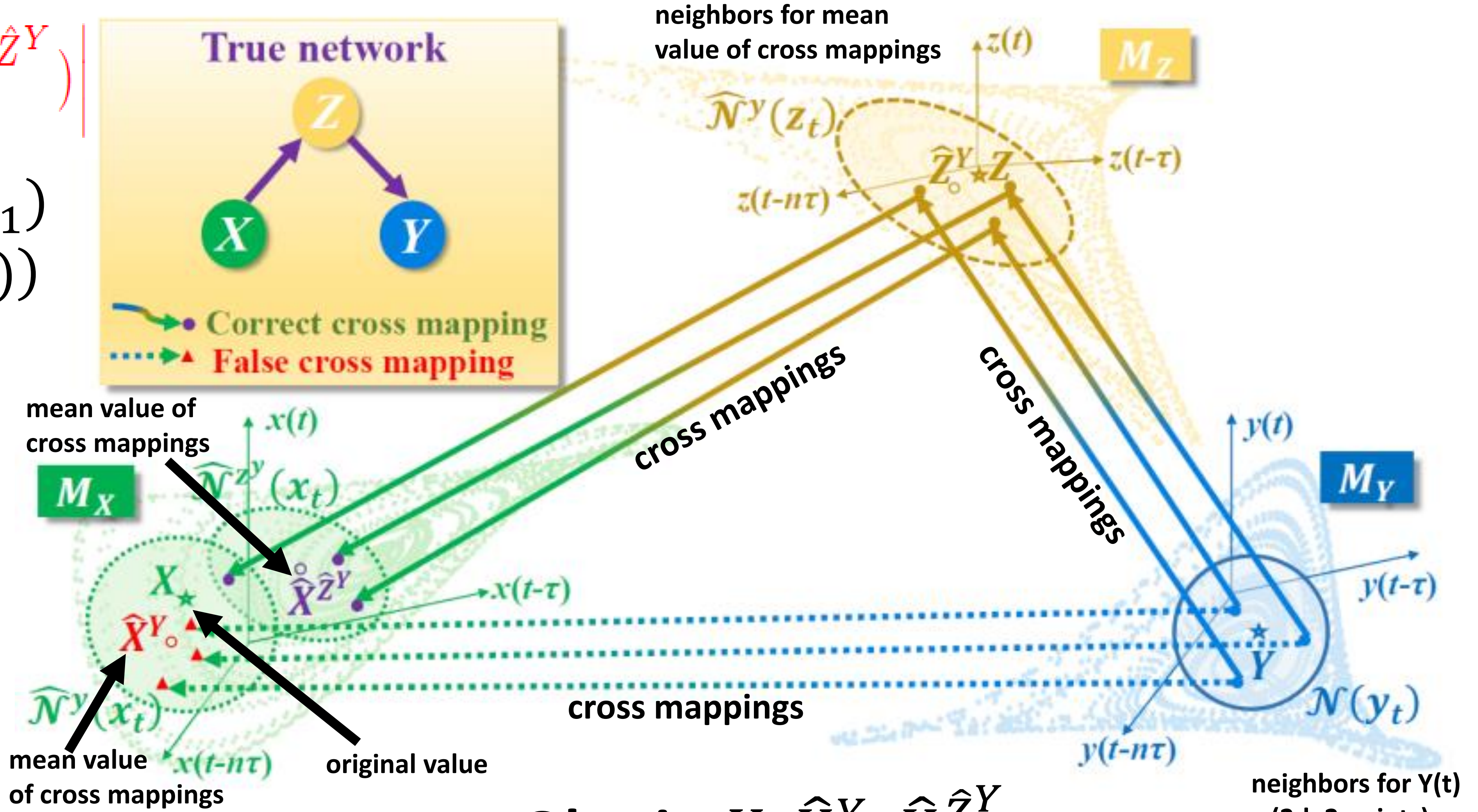
Quantify direct causality of X→Y|Z

Partial Cross Mapping (PCM)

$$\varrho_D = \left|\mathrm{Pcc}(X, \hat{X}^Y | \hat{X}^{\hat{Z}^Y})\right|$$

$$X_{t-1} = H(Y_t; Z_{t-1})$$
$$= H(Y_t; Z(Y_t))$$

Note $Z_{t-1} = G(Y_t)$



Obtain $X, \hat{X}^Y, \hat{X}^{\hat{Z}^Y}$

Indirect and direct causality

- MCM: mutual cross mapping (indirect causality)

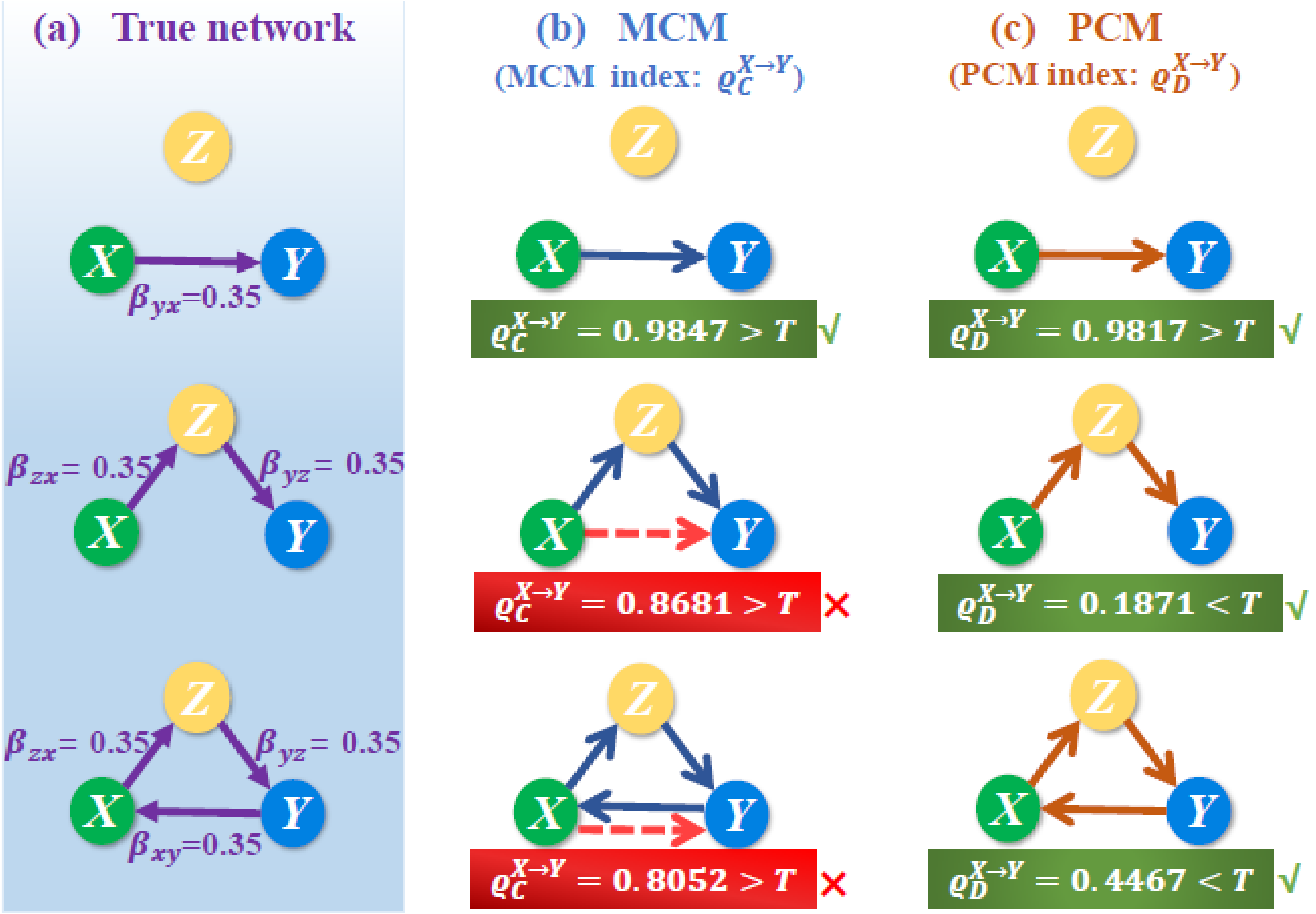
$$\varrho_C = \left|\mathrm{Corr}(X, \hat{X}^Y)\right|$$

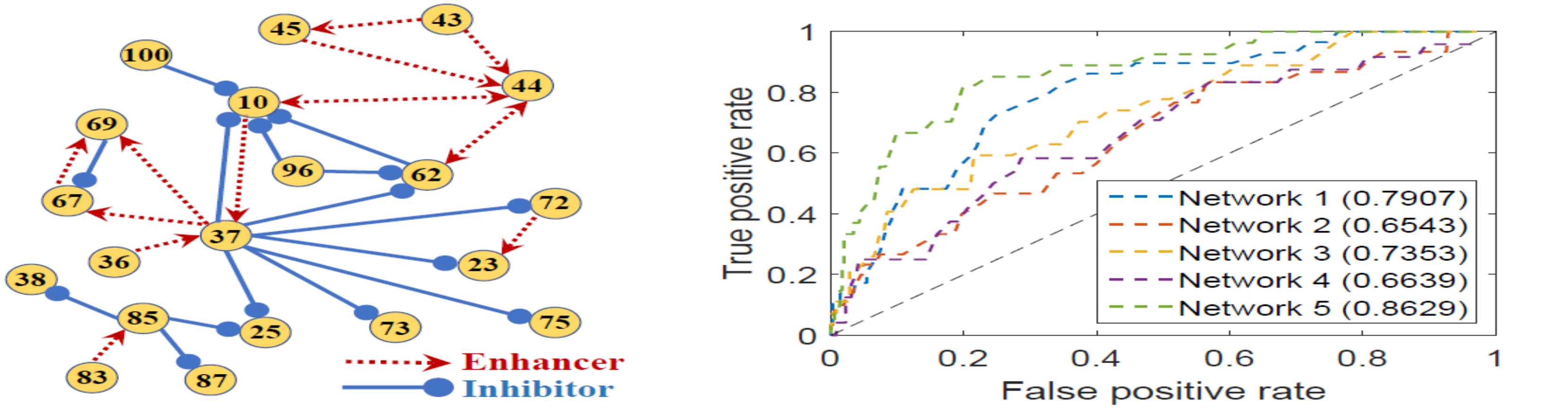
- PCM: partial cross mapping (direct causality)

$$\varrho_D = \left|\mathrm{Pcc}(X, \hat{X}^Y | \hat{X}^{\hat{Z}^Y})\right|$$

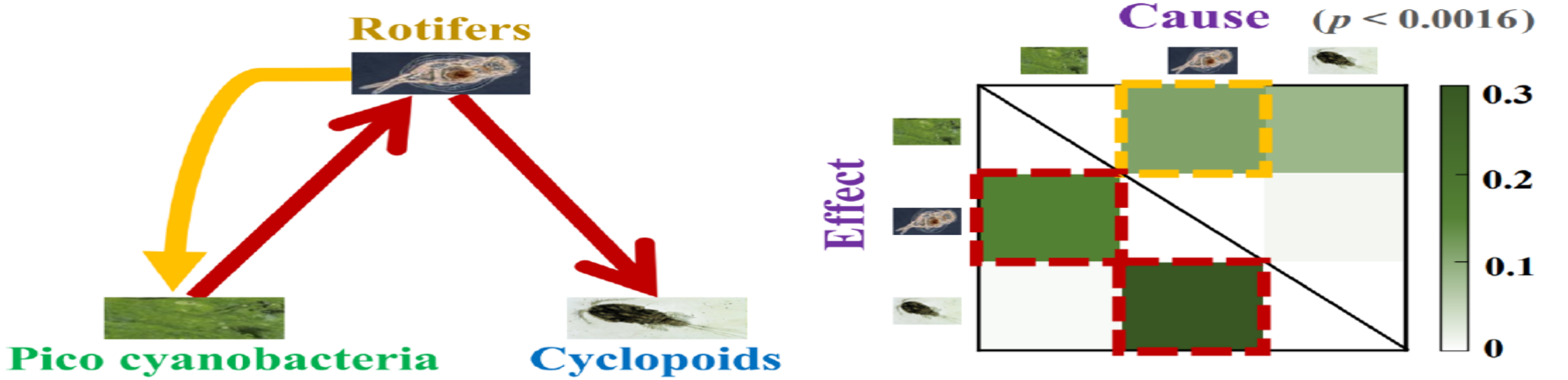
Quantify direct causality of X→Y|Z

Detection of causal links from X to Y

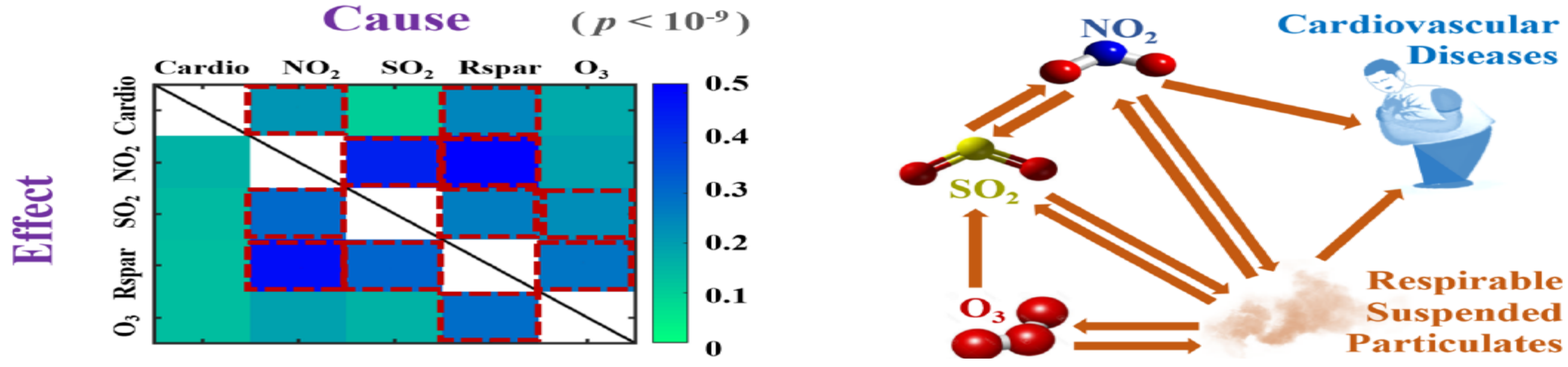




(a) **Gene regulatory networks** (b)

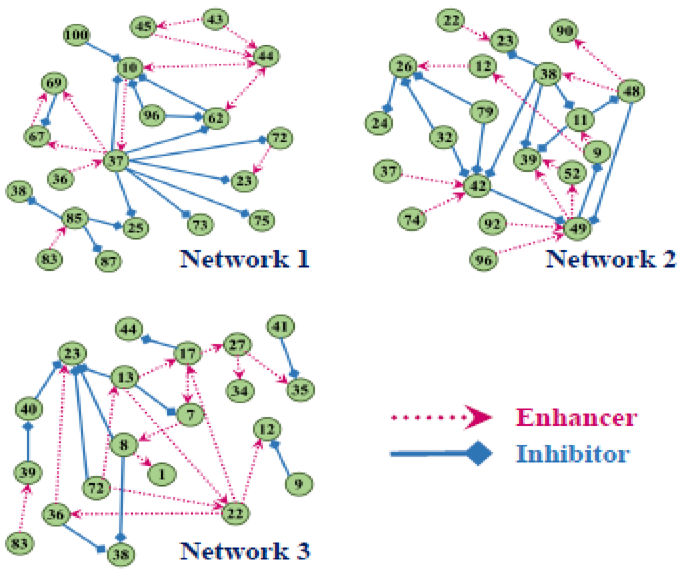


(c) **A food chain network of three plankton species** (d)

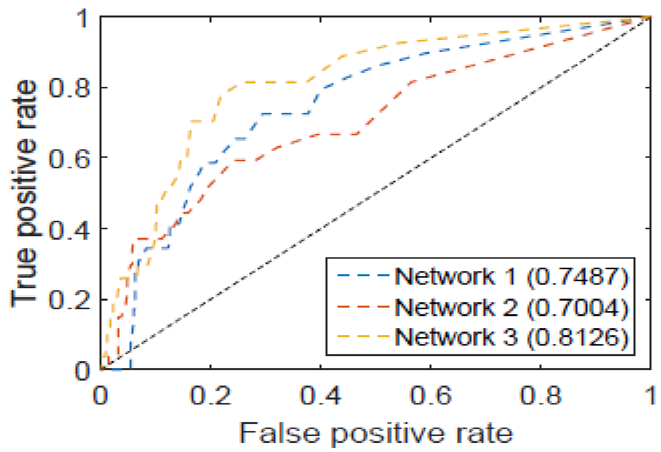


Interactions between air pollutants and cardiovascular diseases (Hong Kong)

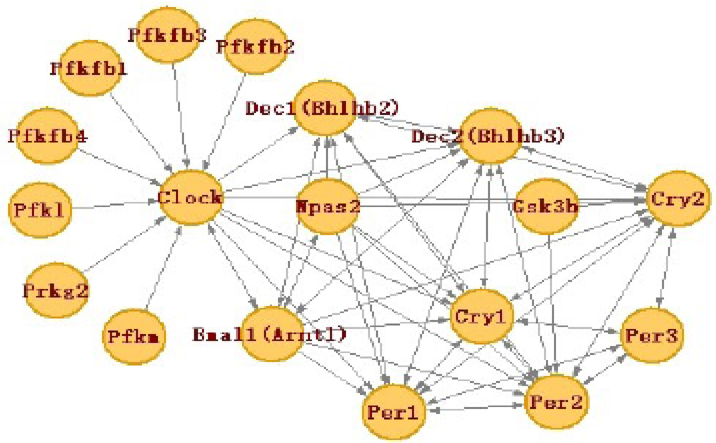
Results in real gene regulatory networks (circadian rhythms)



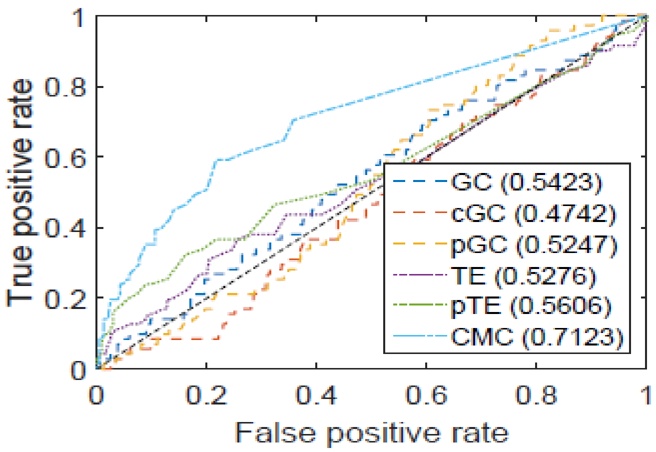
(a)



(b)



(c)



(d)

nature COMMUNICATIONS

Dynamical model can mathematically prove that Y changes accordingly if X changes

ARTICLE

Check for updates

https://doi.org/10.1038/s41467-020-16238-0 OPEN

Partial cross mapping eliminates indirect causal influences

Siyang Leng[1,2,3], Huanfei Ma[4], Jürgen Kurths[5,6], Ying-Cheng Lai[7], Wei Lin[1,2,8]✉, Kazuyuki Aihara[3,9]✉ & Luonan Chen[10,11,12,13]✉

CM → MCM → PCM

Linear combination problem w_i ?

动力学因果性和嵌入熵

Dynamical Causality by Embedding Entropy

Luonan Chen
Chinese Academy of Sciences

New Criterion for Dynamical Causality

Transfer entropy(TE) – Embedding entropy(EE)

Embedding Entropy (EE)

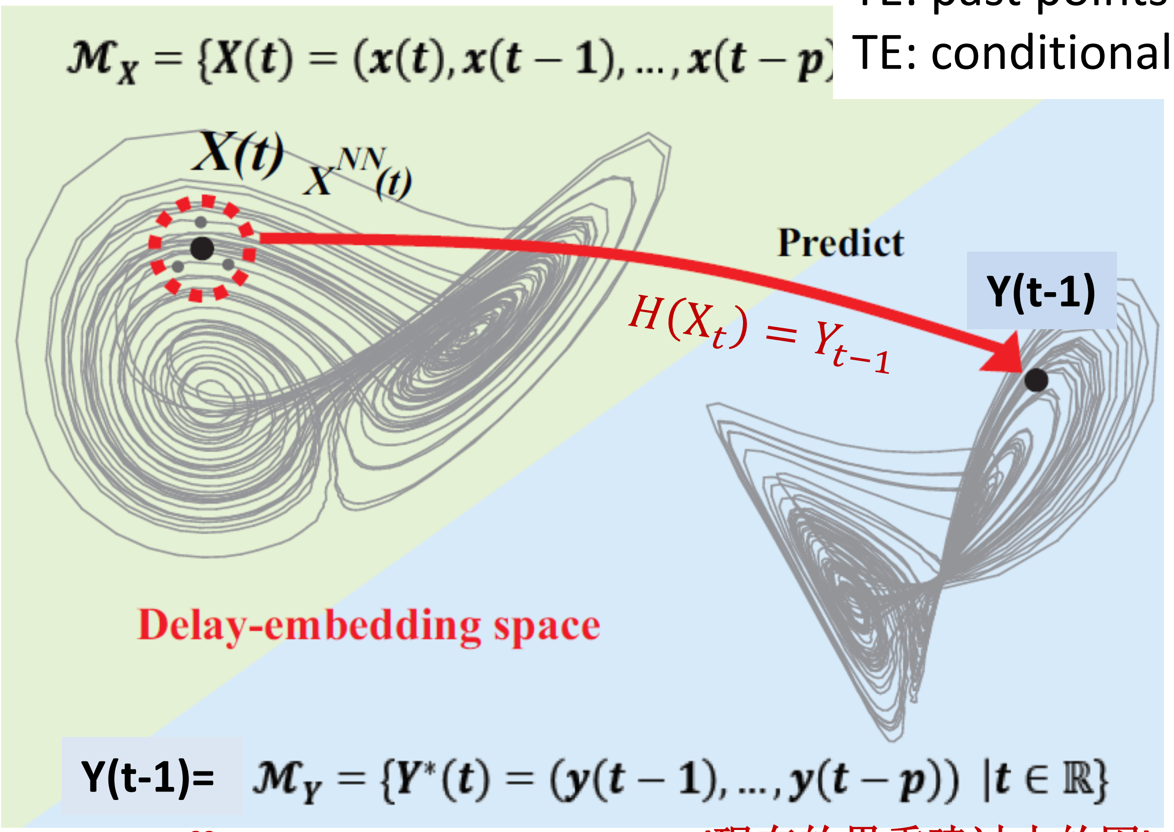
$TE[y \rightarrow x] = H(x_t | x_{t-1}) - H(x_t | y_{t-1}, x_{t-1}) = MI(x_t , y_{t-1} | x_{t-1})$

$EE[y \rightarrow x] = H(Y_{t-1}) - H(Y_{t-1} | X_t^{NN}) = MI(Y_{t-1} , X_t^{NN})$

TE: state space *x,y*; EE: embedding space *X,Y*

TE: past points y_{t-1}; EE: neighbor points X_t^{NN}

TE: conditional *y* to *x*; EE: conditional *X* to *Y*



Current effect reconstructs past cause (现在的果重建过去的因)

H_0 : $EE[y \rightarrow x] = 0$

H_1 : $EE[y \rightarrow x] = H(Y) - H(Y | X^{NN})$

$cs[x \rightarrow y] = D[H1 \,I\, H0]$

X^{NN}: neighbors of X

由X重构Y：加上X邻居的信息后，Y的不确定性减少（射影函数的连续性）

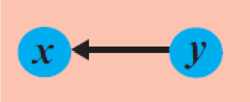
强关联问题！

x_t : original variable

X_t : embedded variable

$Y_{t-1} = H(X_t)$

Embedding entropy (EE)



$$
\begin{aligned}
x(t) &= \widehat{F}[x(t-\Delta t), x(t-2\Delta t), \ldots, x(t-p\Delta t)] + \epsilon_t \\
&= \widehat{F}[x_1(t-\Delta t), \ldots, x_1(t-p\Delta t), \ldots, x_n(t-\Delta t), \ldots, x_n(t-p\Delta t)] + \epsilon_t,
\end{aligned}
$$

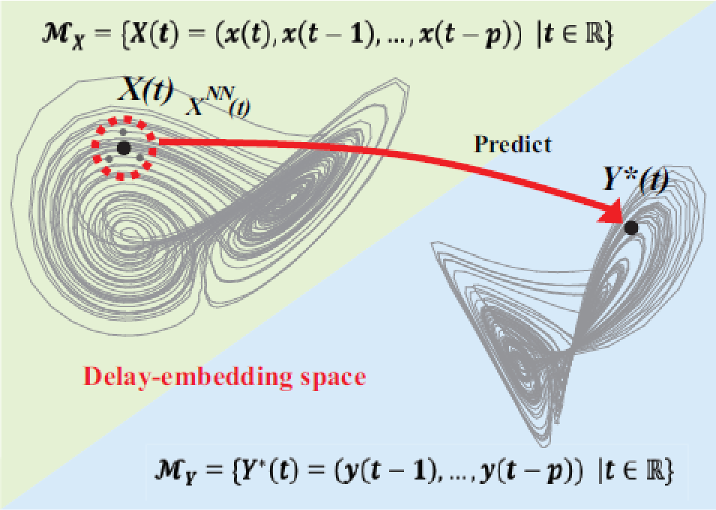
- Implicit function theorem (necessary condition)
- Takens' embedding theorem (sufficient condition)

$$\boxed{Y^*(t) = h(X(t), \epsilon_{x,t})}$$

$$X(t) = (x(t), x(t-\Delta t), \ldots, x(t-p\Delta t))^T \in \mathbb{R}^{p+1}$$

$$Y^*(t) = (y(t-\Delta t), \ldots, y(t-p\Delta t))^T \in \mathbb{R}^p$$

$$\boxed{\mathrm{EE}[y \to x] = \mathrm{MI}(Y^*, X^{NN}) = H(Y^*) - H(Y^*|X^{NN})}$$



X^{NN} denotes the $p+2$ NNs of X, i.e., $X^{NN}(t) = (X^{[1]}(t), X^{[2]}(t), \ldots, X^{[p+2]}(t))$

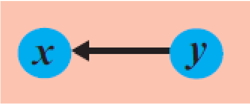
Solving linear problem of MCM, non-separability, false negative, cutoff problem !

$$H(Y_t) = -\sum p(Y_t) \log p(Y_t)$$

Shi et al. J. R. Soc. Interface 2022

Features of Embedding Entropy

$$EE[y \rightarrow x] = H(Y_{t-1}) - H(Y_{t-1} \mid X_t^{NN}) = MI(Y_{t-1}, X_t^{NN})$$



1. Use entropy to solve the **nonlinear problem**.
2. Use embedding mapping and implicit function from *Y* to *X*, to solve the **non-separability problem**.
3. Use X^{NN} instead of *X* in the delay embedding manifold to avoid the **strong association problem**.
4. Use implicit function theorems, solve **cutoff problem** of p.

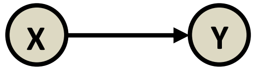
$$H(Y_t) = -\sum p(Y_t) \log p(Y_t)$$

Detecting Dynamical Causality by Cross Mapping

Hirata & Aihara. Phys. Rev. E 81. 016203, 2010 (CM)
Sugihara et al. Science 338, 496, 2012 (MCM)
Leng et al. Nature Communications 11, 2632, 2020 (PCM)
Shi et al. Journal of Royal Society Interface, 2022 (EE)

A neighborhood of M_Y is that of the corresponding M_X

Dynamical Causality from X to Y

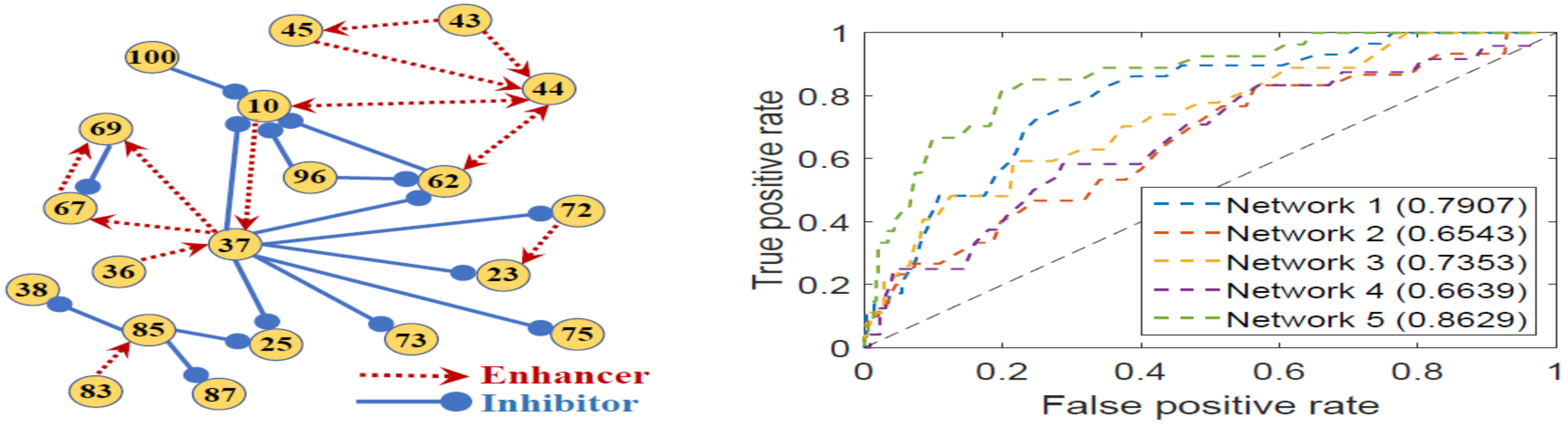


where M_X:delayed-coordinate reconstruction of X,
M_Y:delayed-coordinate reconstruction of Y.

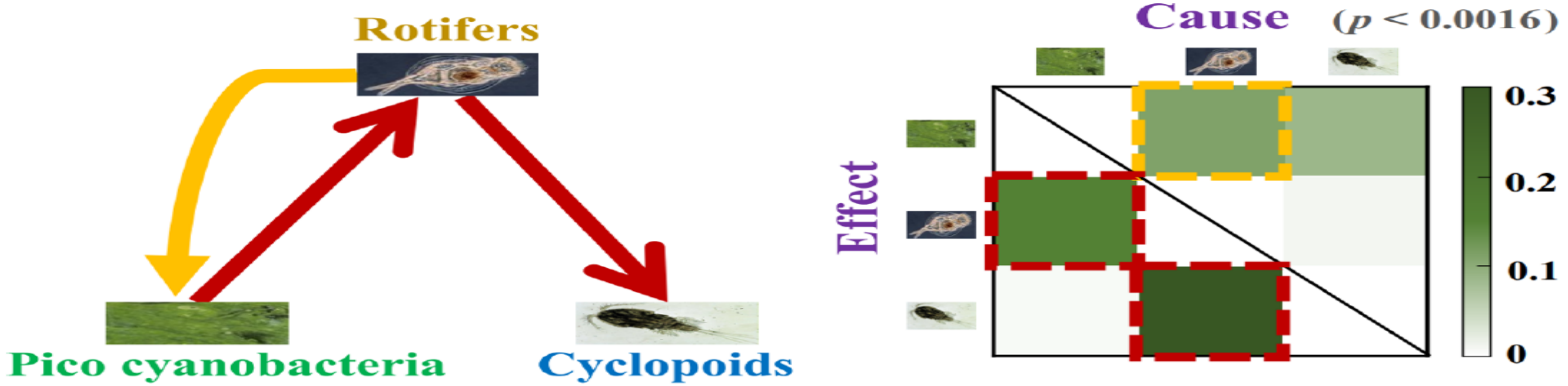
2010 **2012** **2020** **2022** **…**

cross-mapping (CM) → mutual cross-mapping (MCM) → partial cross-mapping (PCM) → embedding entropy (EE)

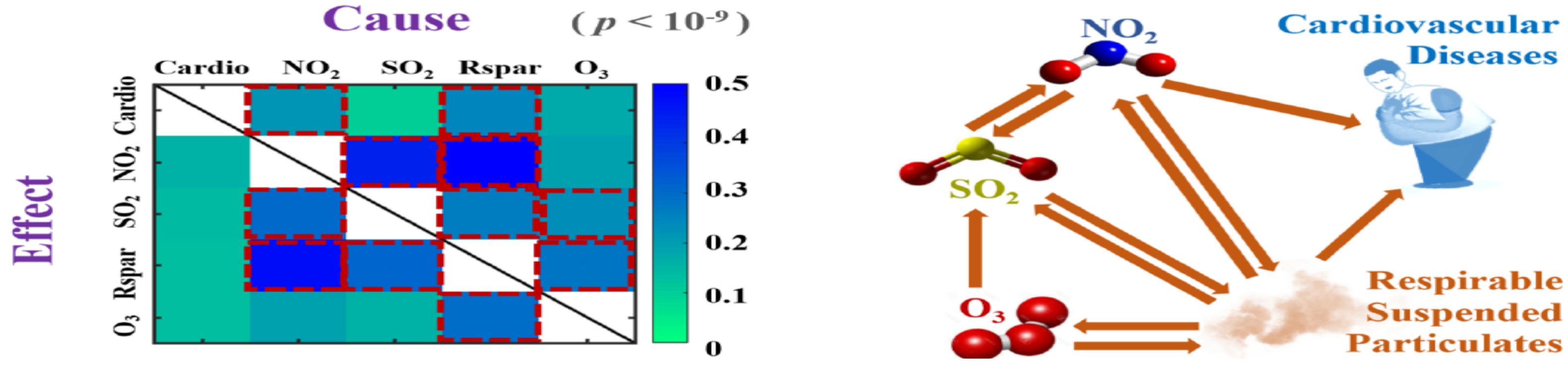
Dynamical causality
Inverse-mapping



(a) **Gene regulatory networks** (b)

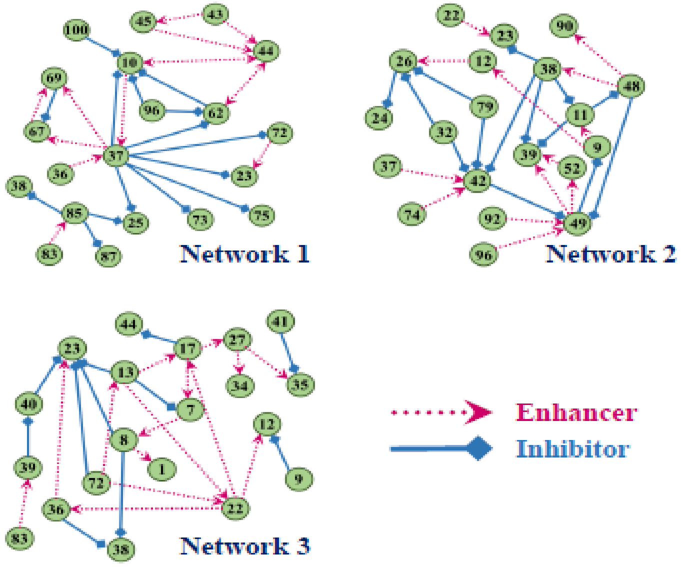


(c) **A food chain network of three plankton species** (d)

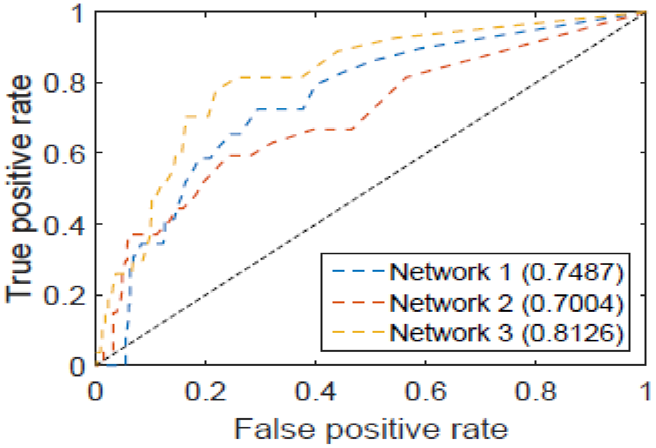


Interactions between air pollutants and cardiovascular diseases (Hong Kong)

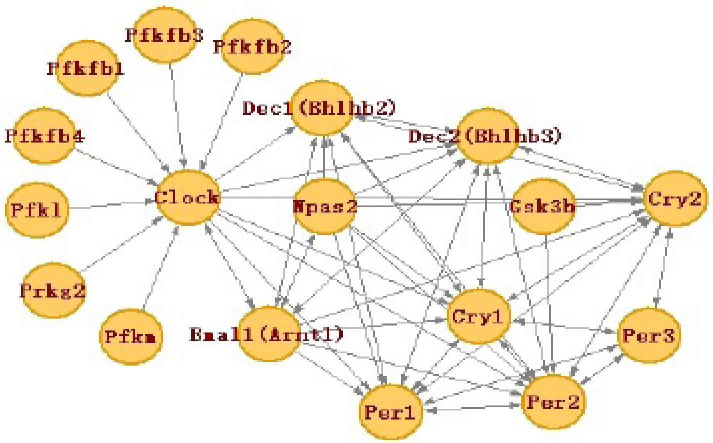
Results in real gene regulatory networks (circadian rhythms)



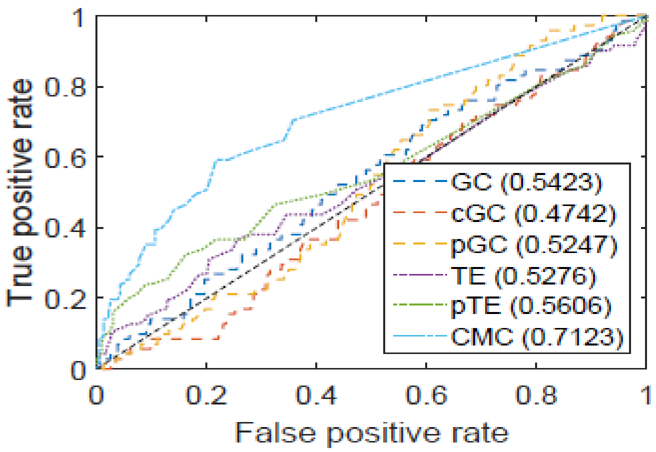
(a)



(b)



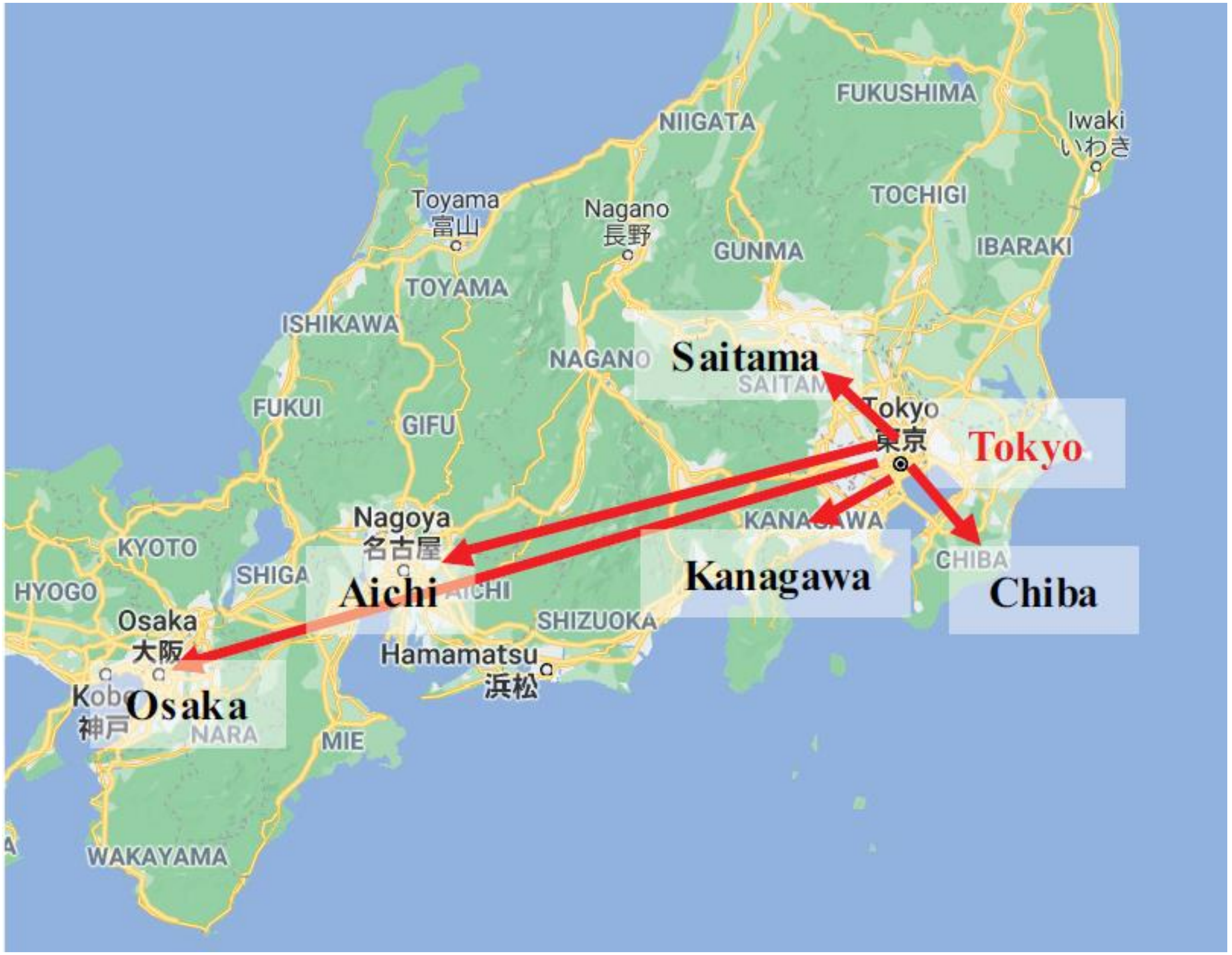
(c)



(d)

Shi et al. J. R. Soc. Interface 2022

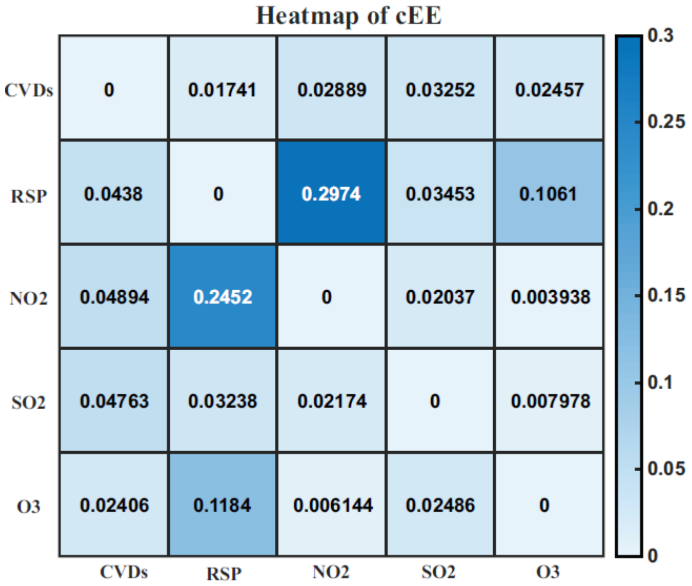
The 5 most affected prefectures by Tokyo

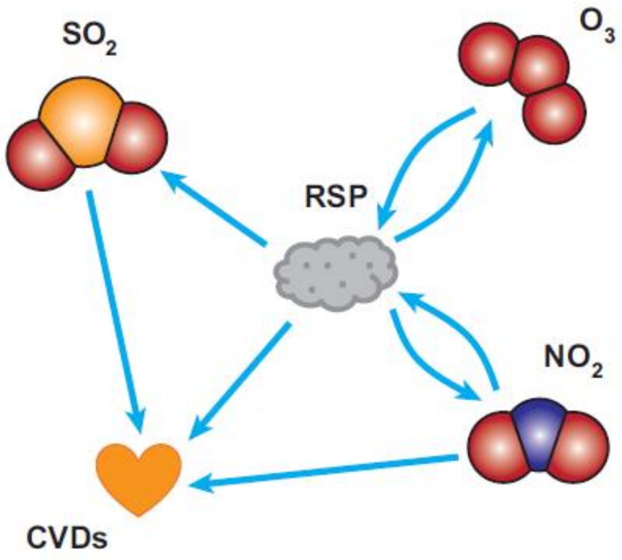


Real dataset 2 : Hong Kong air pollution and cardiovascular diseases(CVDs)

Hong Kong from March 1995 to November 1997 (1032 days in total)

- Air pollutants: NO_2, SO_2, respirable suspended particulate (RSP), and O_3 (in μg/m^3)
- CVDs: daily number of CVD admissions to major hospitals in Hong Kong



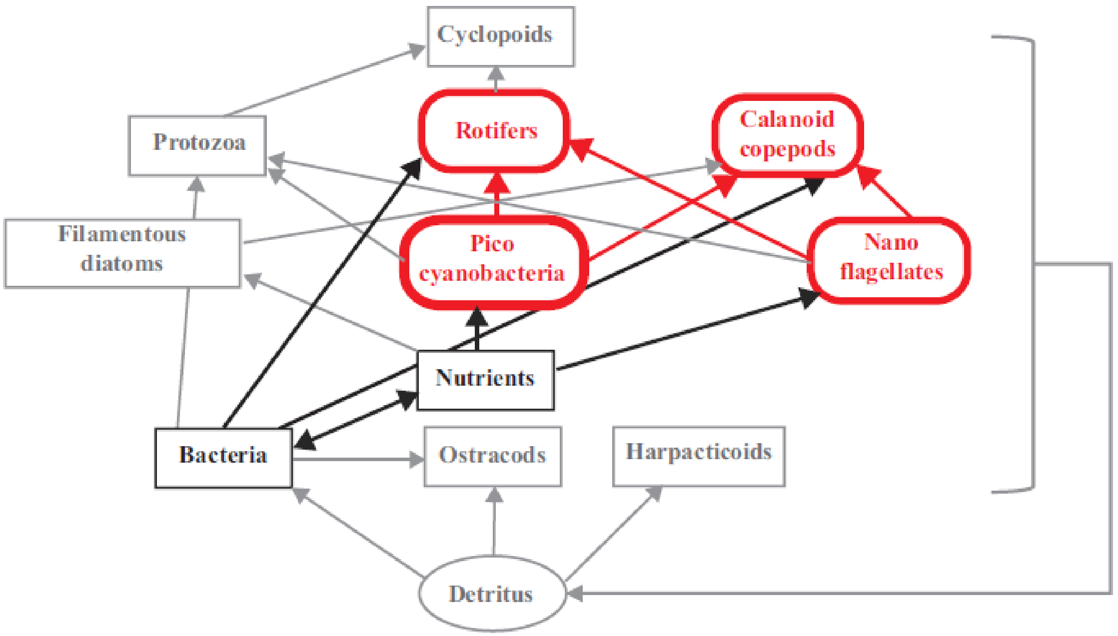


By cEE (65% quantile as the threshold value)

Real dataset 3 : Food chain network

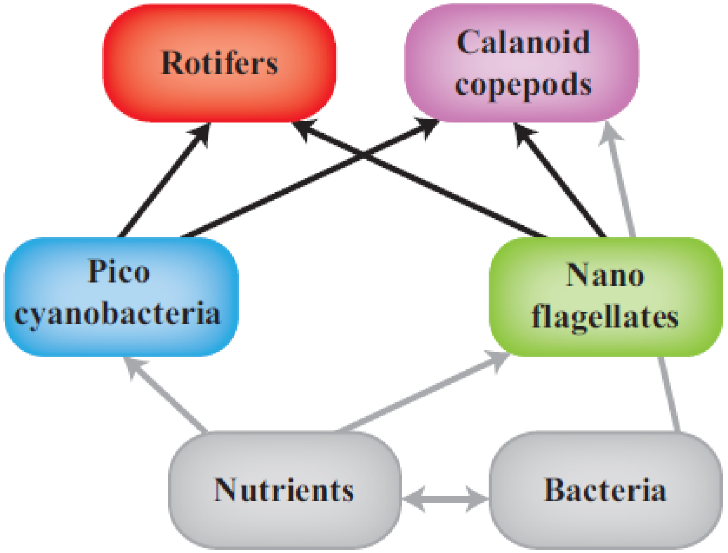
An 8-year mesocosm experiment of a plankton community isolated from the Baltic Sea (1032 days).

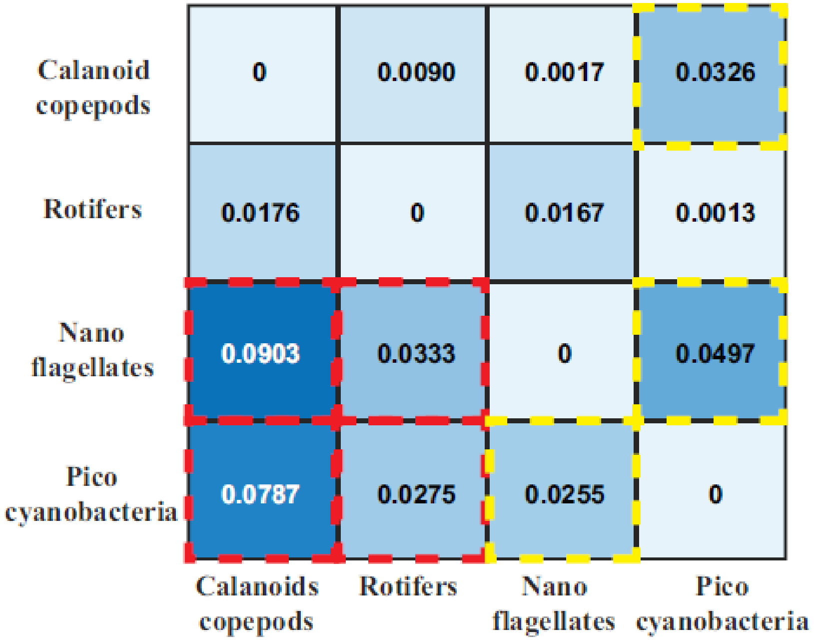
- Four species: calanoid copepods, rotifers, nanoagellates, and picocyanobacteria



Shi et al. J. R. Soc. Interface 2022

Result of cEE





Red square: True positive
Yellow square: False positive caused by hidden variables

EE for Inverse-Mapping

Causal Inference

- Cross-Mapping Cardinality (CMC)
  (time series data)
- Neighboring Cross-Mapping Entropy (NME)
  (non-time series)
- Counterfactual-Invariant Diffusion-based GNN Explainer (CIDER) (non-time series)

Luonan Chen (Chinese Academy of Sciences)

Tao et al. Fundamental Research 2023

Reverse Mapping or Reverse Reconstruction

Y(t)=F(X(t+1)) at embedded space, but x(t+1)=f(x(t),y(t)) at original space

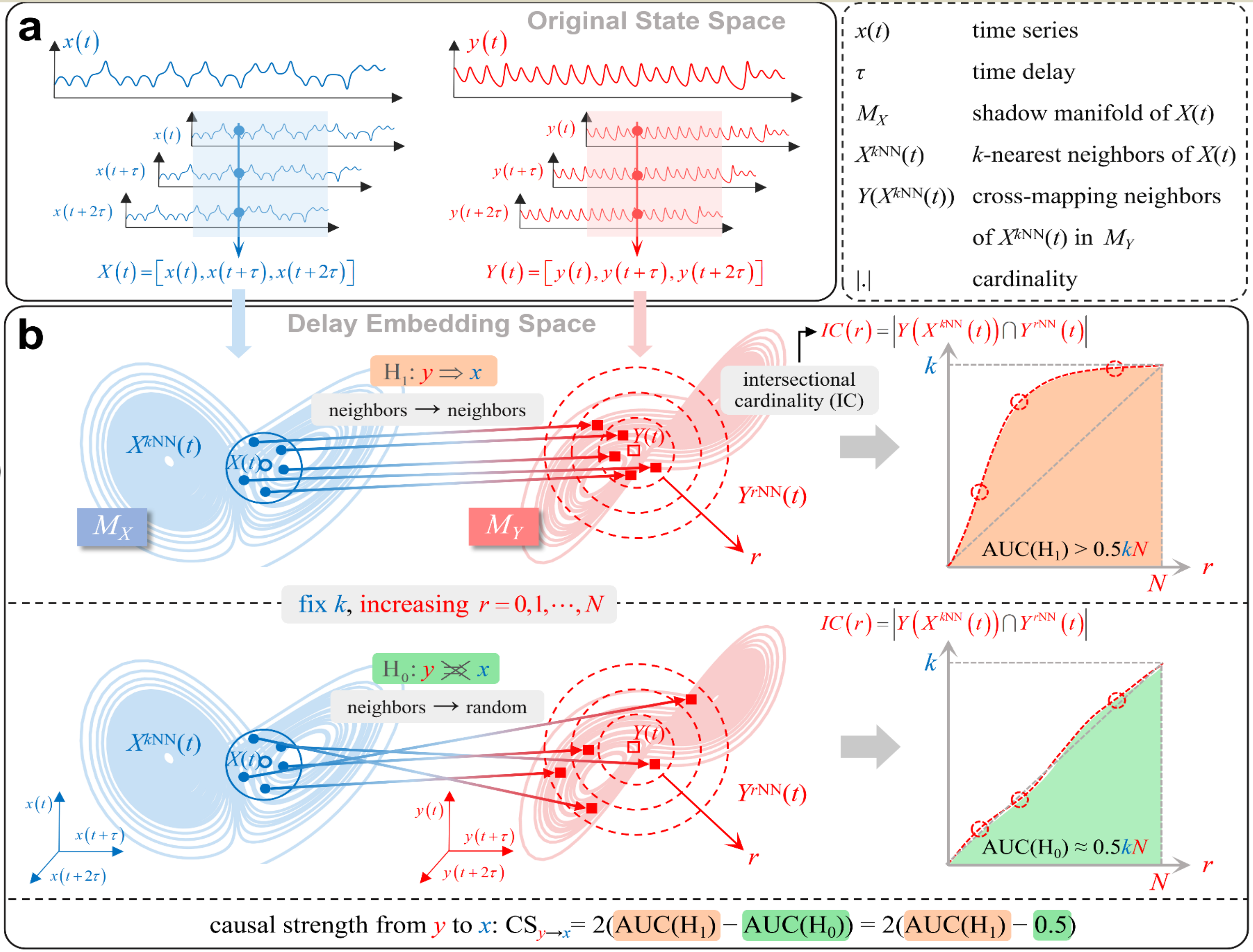
$x(t+1)=f(x(t), y(t))$
$y(t+1)=g(y(t))$

or $Y(t) = F(X(t+1))$

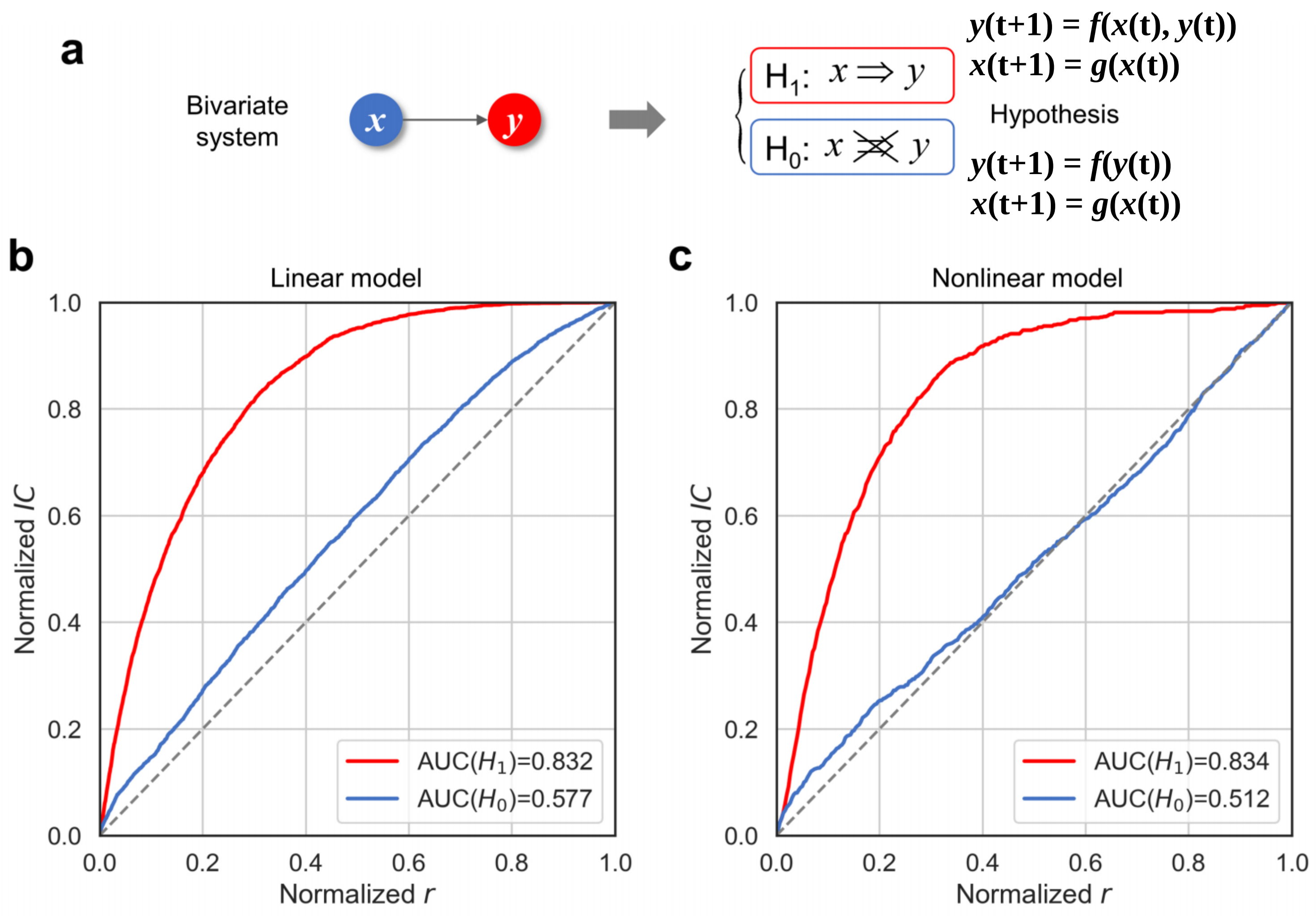
$x(t+1)=f(x(t))$
$y(t+1)=g(y(t))$

or $Y(t) \neq F(X(t+1))$

The corresponding X(Y) of Y neighbors are also X neighbors

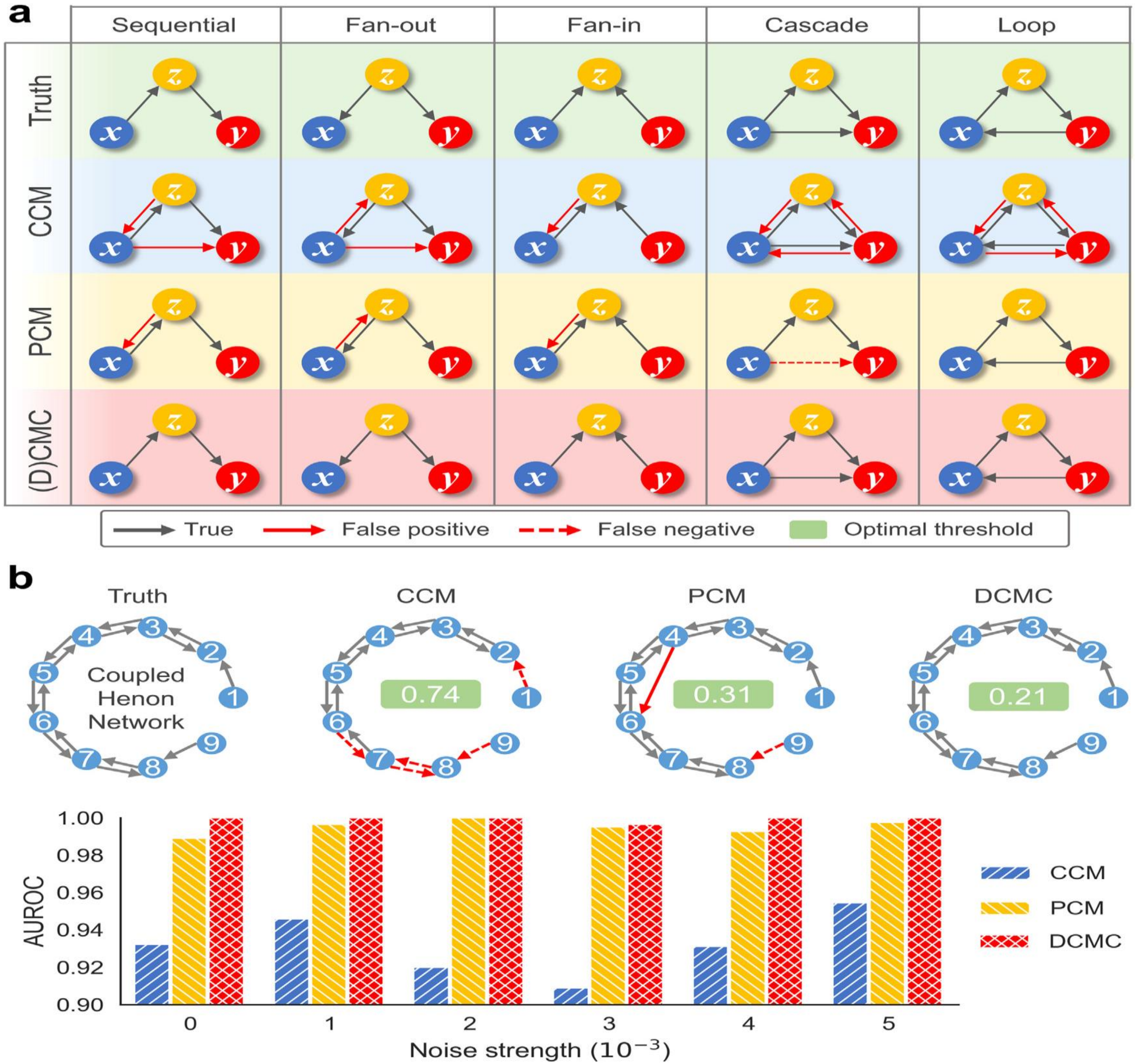


Results



$$\begin{cases} x(t+1) = 0.45x(t) + 0.45x(t-1) + \varepsilon_x(t) \\ y(t+1) = 0.45y(t) + 0.45y(t-1) + \beta_{yx}x(t) + \varepsilon_y(t) \end{cases},$$

$$\begin{cases} x(t+1) = x(t)\big(3.7 - 3.7x(t)\big) + \varepsilon(t) \\ y(t+1) = y(t)\big(3.7 - 3.7y(t) - \beta_{yx}x(t)\big) + \varepsilon(t) \end{cases},$$

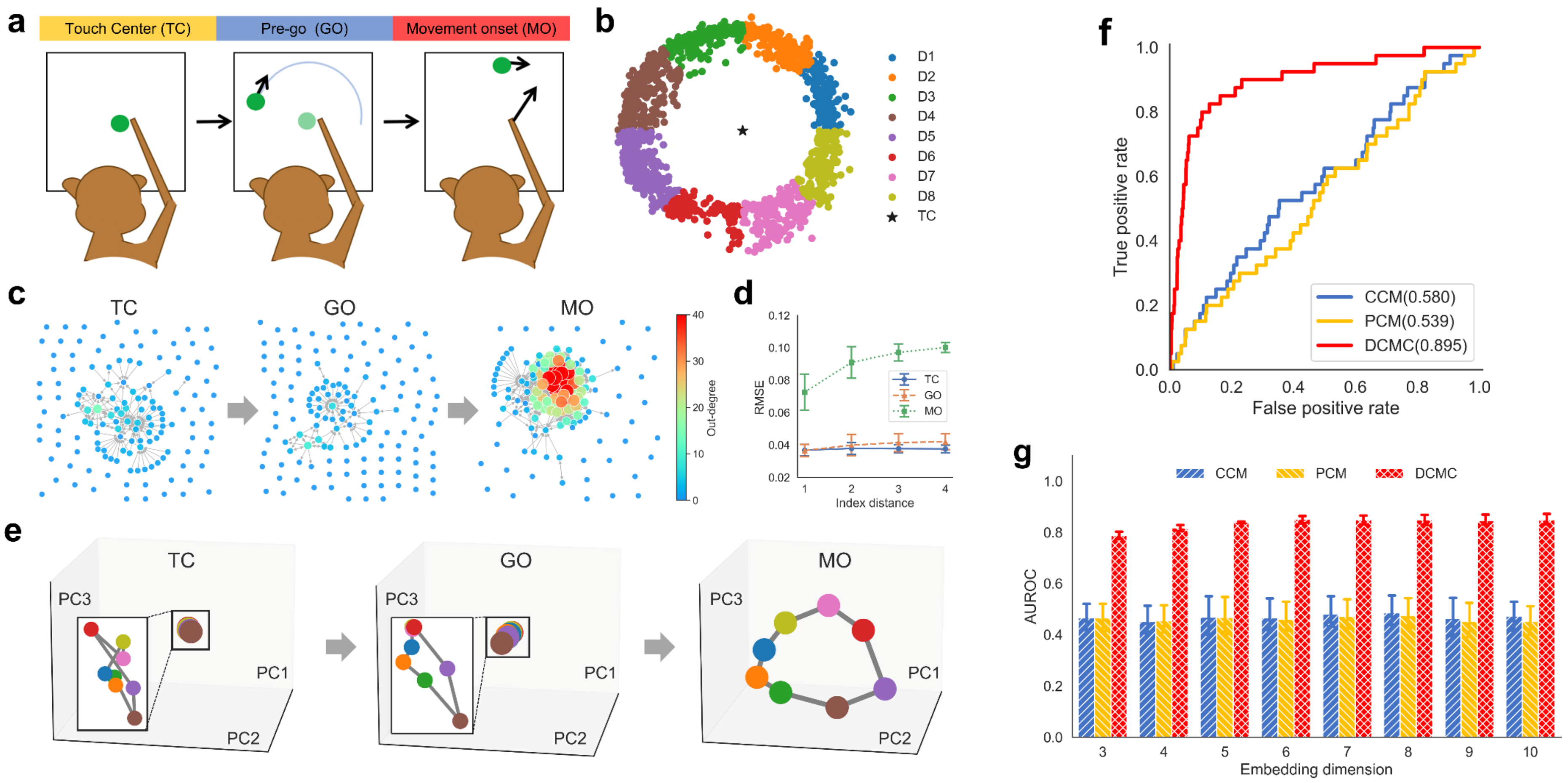


a
Sequential
Fan-out
Fan-in
Cascade
Loop
Truth
CCM
PCM
(D)CMC
True
False positive
False negative
Optimal threshold
b
Truth
CCM
PCM
DCMC
Coupled Henon Network
0.74
0.31
0.21
AUROC
1.00
0.98
0.96
0.94
0.92
0.90
0
1
2
3
4
5
Noise strength (10^{-3})
CCM
PCM
DCMC

Results

Tao et al. Fundamental Research 2023

Neural network of the rhesus monkey's motor cortex



Acknowledgments

- Jifan Shi, Siyang Leng, Kazuyuki Aihara,
  The University of Tokyo

- Huanfei Ma,
  Suzhou University

- Wei Lin,
  Fudan University

Part mutual information: Zhao et al. PNAS 2016
Randomly distributed embedding: Ma et al. PNAS 2018
Partial cross-mapping*:* Leng et al. Nature Communications 2020
Embedding entropy: Shi et al. J. R. Soc. Interface 2022