Causal connectivity measures for pulse-output network reconstruction

PNAS

RESEARCH ARTICLE | NEUROSCIENCE
APPLIED MATHEMATICS

Causal connectivity measures for pulse-output network reconstruction: Analysis and applications

Zhong-qi K. Tian[a,b,c,1], Kai Chen[a,b,c,1], Songting Li[a,b,c,2], David W. McLaughlin[d,e,f,g,2], and Douglas Zhou[a,b,c,h,2]

Contributed by David W. McLaughlin; received April 3, 2023; accepted March 3, 2024; reviewed by David Hansel and Dario L. Ringach

presented by 张博涛

1. Background
2. Relations between Causality Measures(mathematical)
3. Verification in HH models
4. Application: reconstruction from physiological data
5. Discussion

Background

network connectivity

causal connectivity $\xrightarrow{?}$ structural connectivity

different causality measures may produce differnt results → in pulse-output networks they are related

different nature of ‘causal’ and ‘structural’
X->Y->Z, then there’s causal relations from X->Z

computation relying on the whole network is costly → pairwise reconstruction

Background

Four measures for Causal Interactions(connectivity):

TDCC: time-delayed correlation coefficient linear
TDMI: time-delayed mutual information non-linear
overlook historical effects → over-estimation
GC: Granger causality linear
TE: transfer entropy non-linear → but incur curse of dimensionality

Background

For each node X, a corresponding $\{x_n\}$, x_n is 1 iff X fires within the nth time window of $\triangle\, t$ (otherwise $x_n = 0$).

Suppose X, Y are two nodes, and for the following formulas, a positive value of 'm' indicates the calculation of causal value from Y to X.

- **TDCC**:

$$C(X, Y; m) = \frac{\mathrm{cov}(x_n, y_{n-m})}{\sigma_x \sigma_y}$$

  where cov is the covariance, σ_x and σ_y are the standard deviations of x_n and y_n, respectively.

- **TDMI**:

$$I(X, Y; m) = \sum_{x_n, y_{n-m}} p(x_n, y_{n-m}) \log \left(\frac{p(x_n, y_{n-m})}{p(x_n) p(y_{n-m})} \right)$$

  where $p(x_n, y_{n-m})$ is the joint probability distribution of x_n, y_{n-m}, and $p(\cdot)$ are the marginal probabilities.

- **GC**: Autoregression for X:

$$x_{n+1} = a_0 + \sum_{i=1}^{k} a_i x_{n+1-i} + \epsilon_{n+1}$$

Now we include the historical information of Y (message length l, time delay m):

$$x_{n+1} = \tilde{a_0} + \sum_{i=1}^{k} \tilde{a_i} x_{n+1-i} + \sum_{j=1}^{l} \tilde{b_j} y_{n+2-m-j} + \eta_{n+1}$$

GC is how much the regression is improved after incorporating the historical information of Y:

$$G_{Y \to X}(k, l; m) = \log \frac{\mathrm{Var}(\epsilon_{n+1})}{\mathrm{Var}(\eta_{n+1})}$$

- **TE**:

$$T_{Y \to X}(k, l; m) = \sum_{x_{n+1}, x_n^{(k)}, y_{n+1-m}^{(l)}} p(x_{n+1}, x_n^{(k)}, y_{n+1-m}^{(l)}) \log \left(\frac{p(x_{n+1} | x_n^{(k)}, y_{n+1-m}^{(l)})}{p(x_{n+1} | x_n^{(k)})} \right)$$

Relations between Causality Measures(mathematical)

$$I(X,Y;m)=\frac{C^2(X,Y;m)}{2}+O(\Delta t^2\Delta p_m^3),\quad \Delta p_m=\frac{p(x_n=1,y_{n-m}=1)}{p(x_n=1)p(y_{n-m}=1)}-1$$

$$G_{Y\to X}(k,l;m)=\sum_{i=m}^{m+l-1}C^2(X,Y;i)+O(\Delta t^3\Delta p_m^2).$$

$$T_{Y\to X}(k,l;m)=\sum_{i=m}^{m+l-1}I(X,Y;i)+O(\Delta t^3\Delta p_m^2).$$

with the assumption that length of historical information used is shorter than consecutive spikes' time interval

proof: Taylor expand

Relations between Causality Measures---Mechanism

Y→W→X

id-connected pair (Y,X) might be mis-inferred as d-conn pair by causality measures

but no problem for pulse-output network:

$$\delta p_{Y\to X} = p(x_n = 1 \mid y_{n-m} = 1) - p(x_n = 1 \mid y_{n-m} = 0)$$

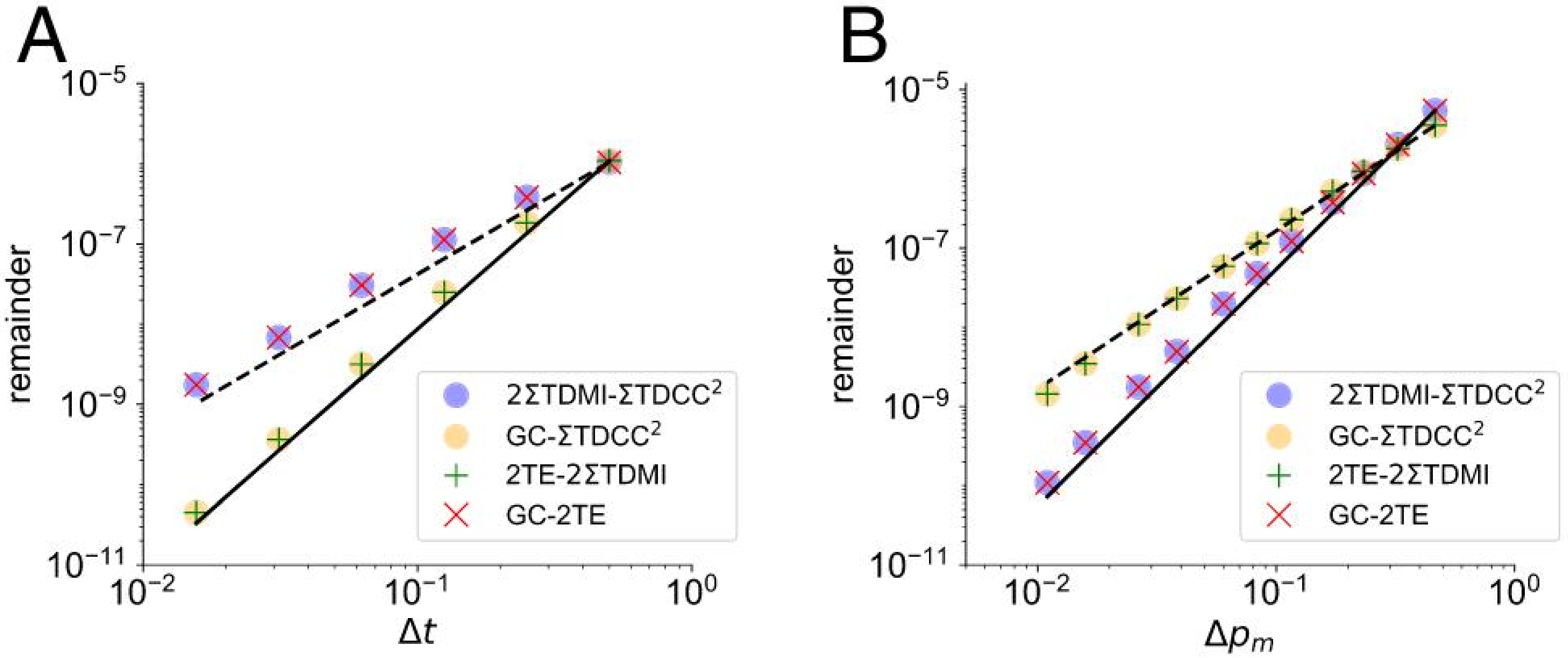
$$C(X,Y;m) = \delta p_{Y\to X}\sqrt{\frac{p_y - \hat{p}_y^2}{p_x - \hat{p}_x^2}}$$

$$\delta p_{Y\to X} = O(\delta p_{Y\to W}\cdot \delta p_{W\to X})$$

$$C(X,Y;m) = O\left(C(W,Y;m)\cdot C(X,W;m)\right)$$

Y←W→X confounder is resolved the same way

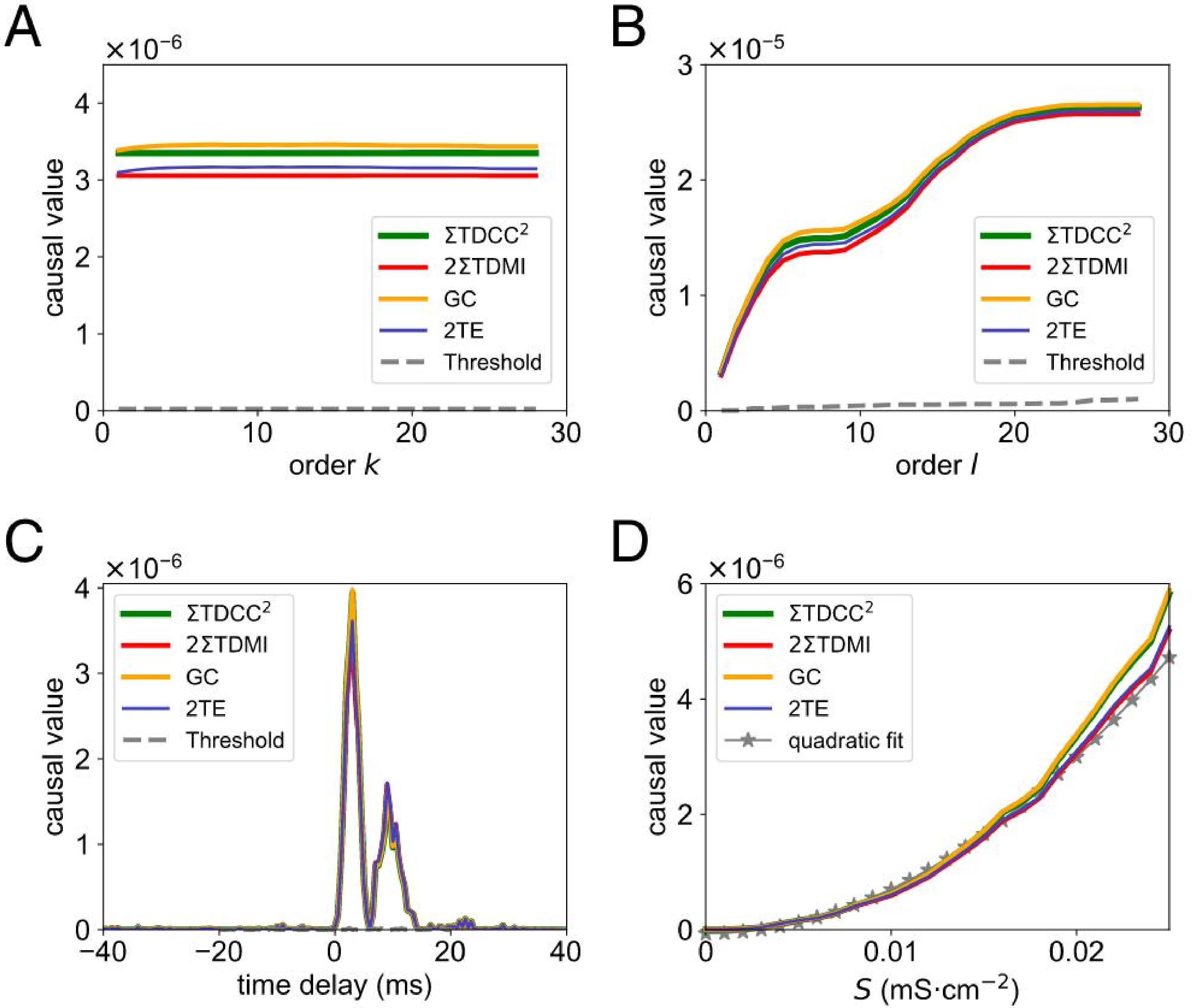
Verification in HH models



d-conn Y→X (in HH with 10 excitatory neurons)

Verification in HH models

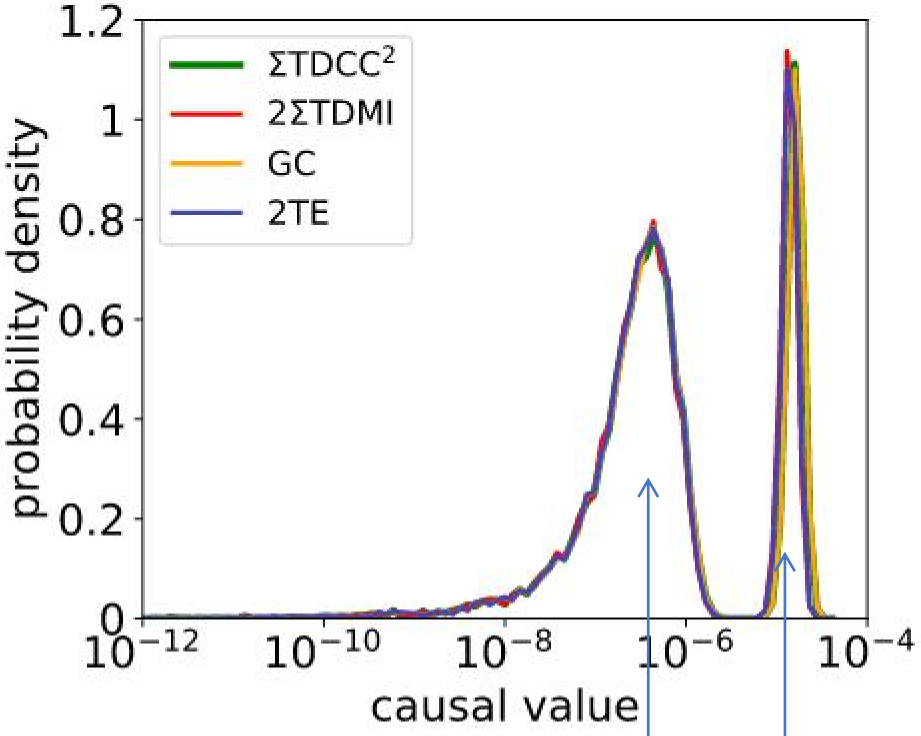
d-conn Y→X (in HH with 10 excitatory neurons)



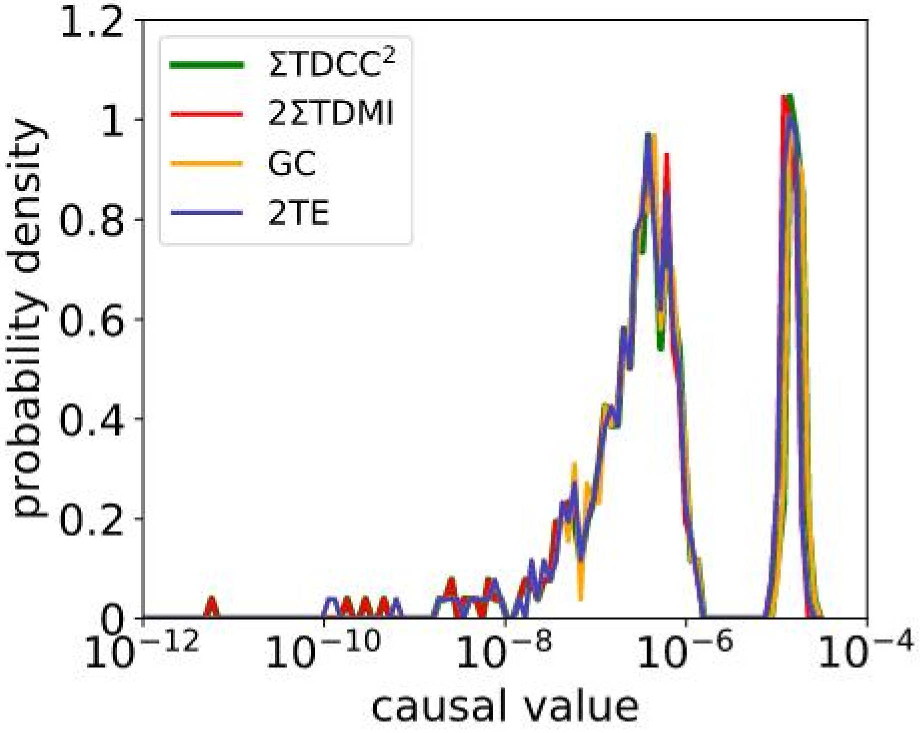
S: coupling strength, linearly related to Δp_m

Verification in HH models

A a whole HH of 100 excitatory neurons



B subnetwork of 20 neurons



id-conn

AUC -> 1

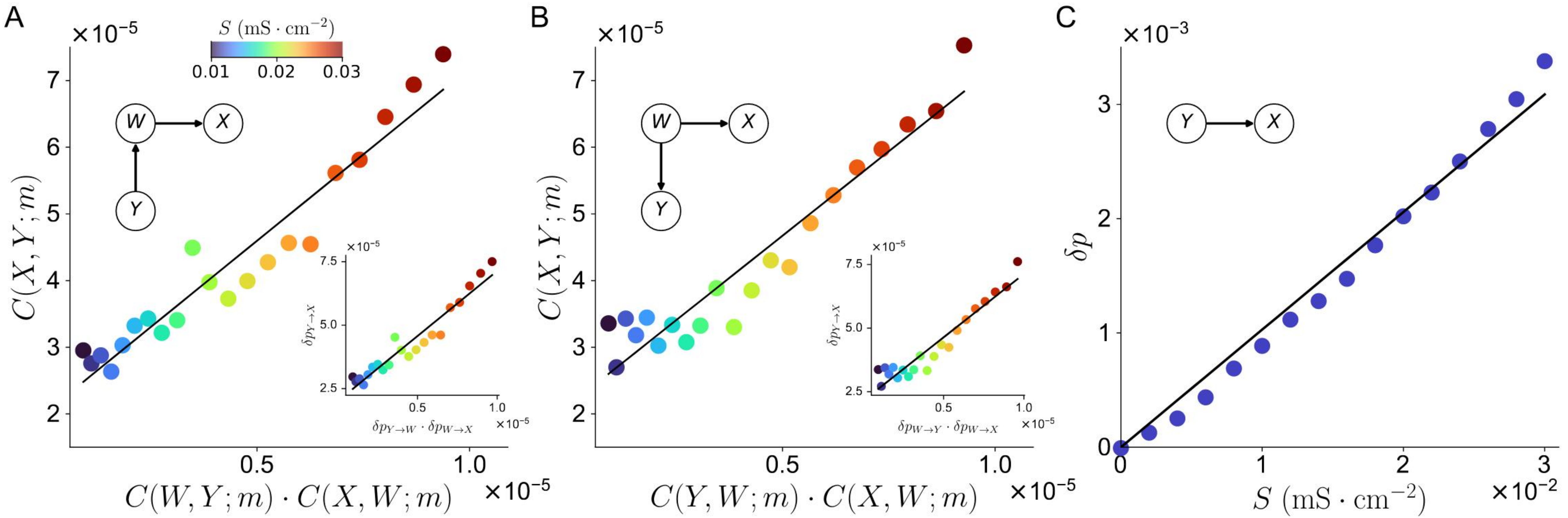
d-conn

Verification in HH models

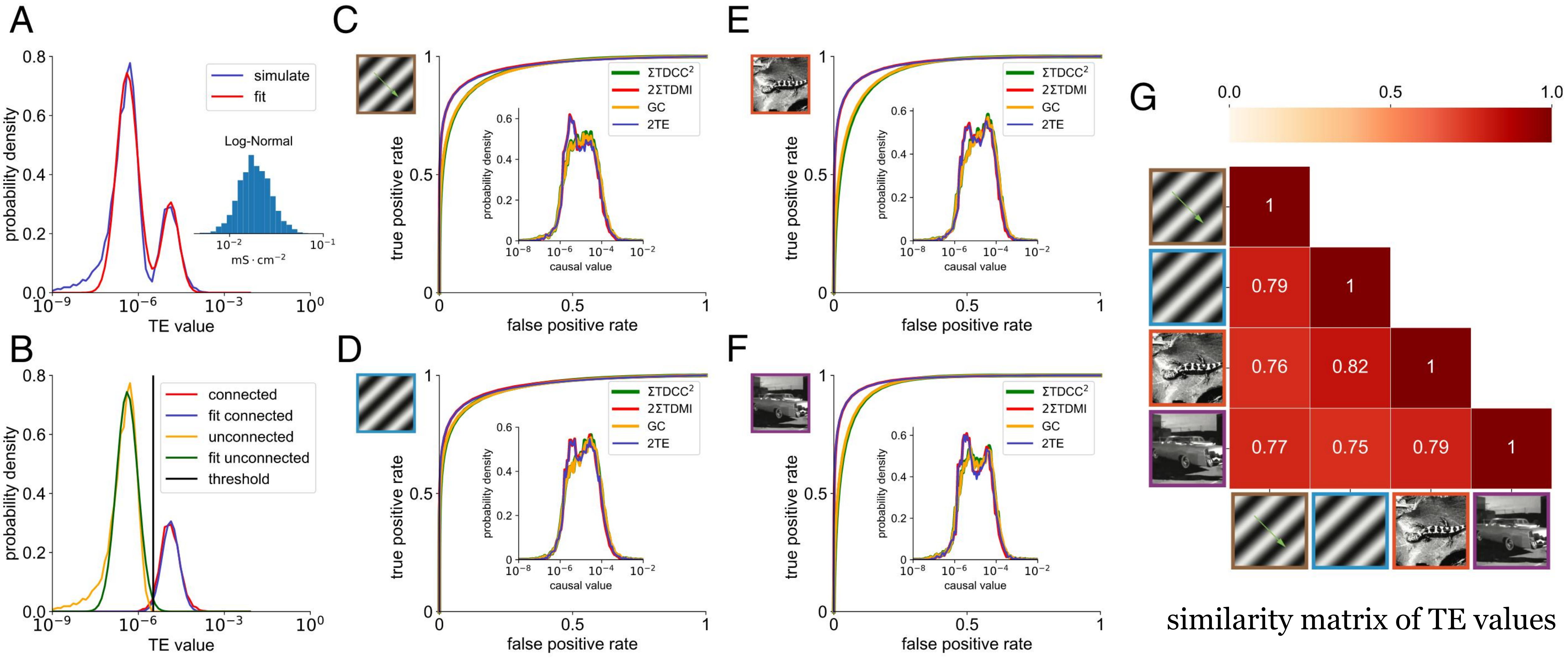
$$C(X,Y;m) = \delta p_{Y\to X}\sqrt{\frac{p_y - \hat{p}_y^2}{p_x - \hat{p}_x^2}}$$

linearly dependent on S

$$C(X,Y;m) = O\left(C(W,Y;m)\cdot C(X,W;m)\right)$$

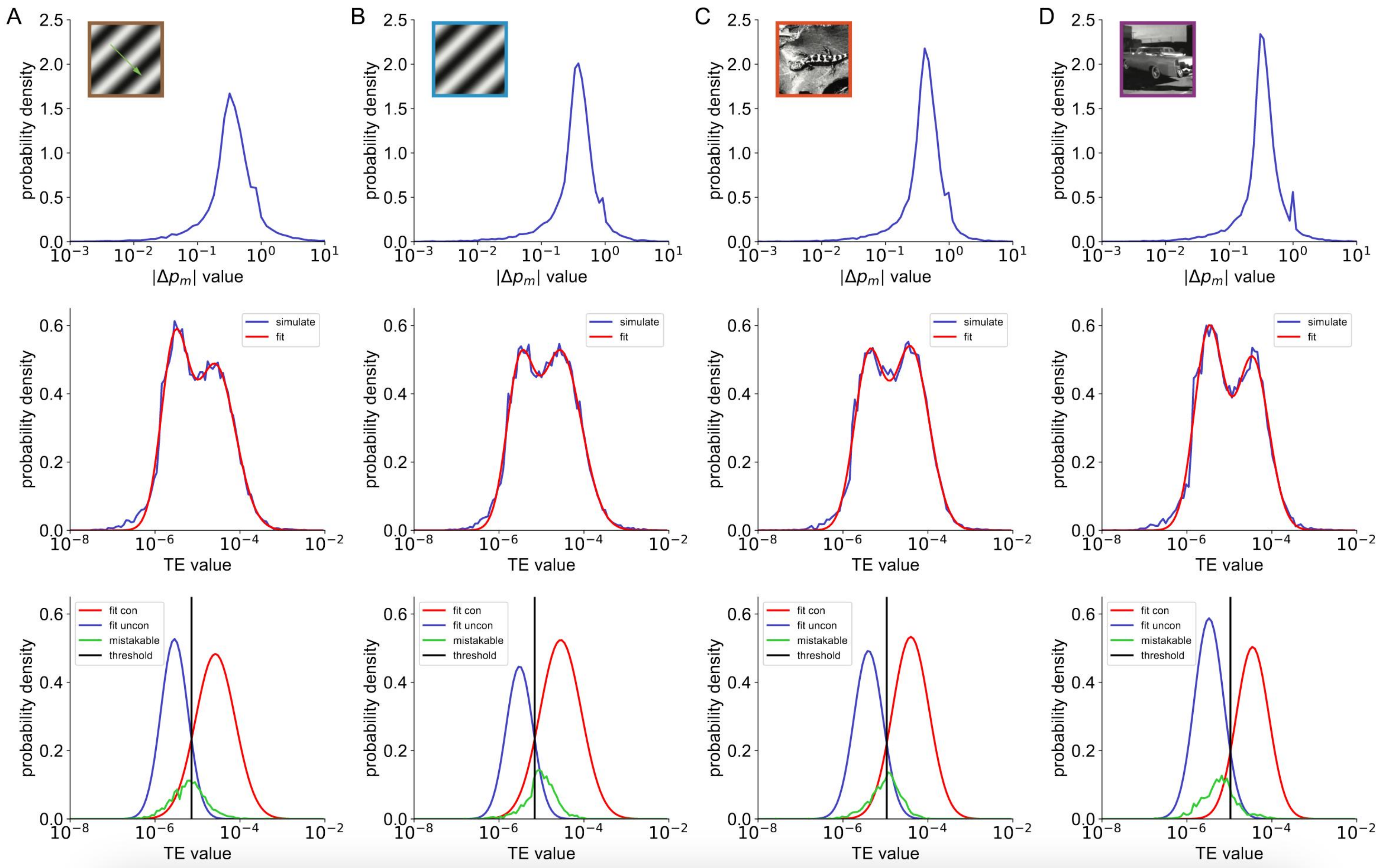


Application: reconstruction from physiological data

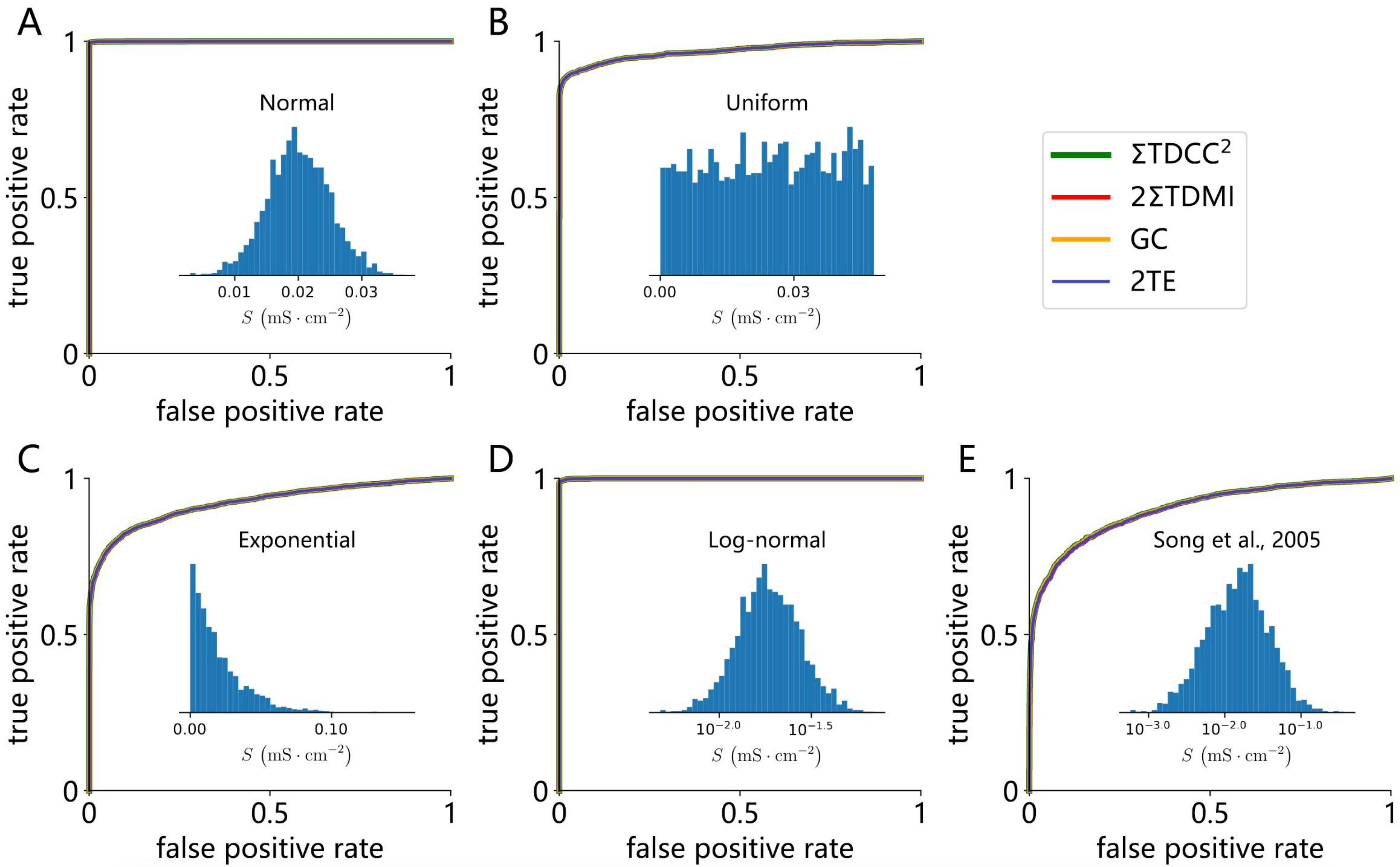


similarity matrix of TE values

Application: reconstruction from physiological data



Application: reconstruction from physiological data



Discussion

Why the method works for pulse-output networks?

Two key features of pulse-output networks:

i) short timescale of autocorrelation

ii) the weakness of indirect causalities