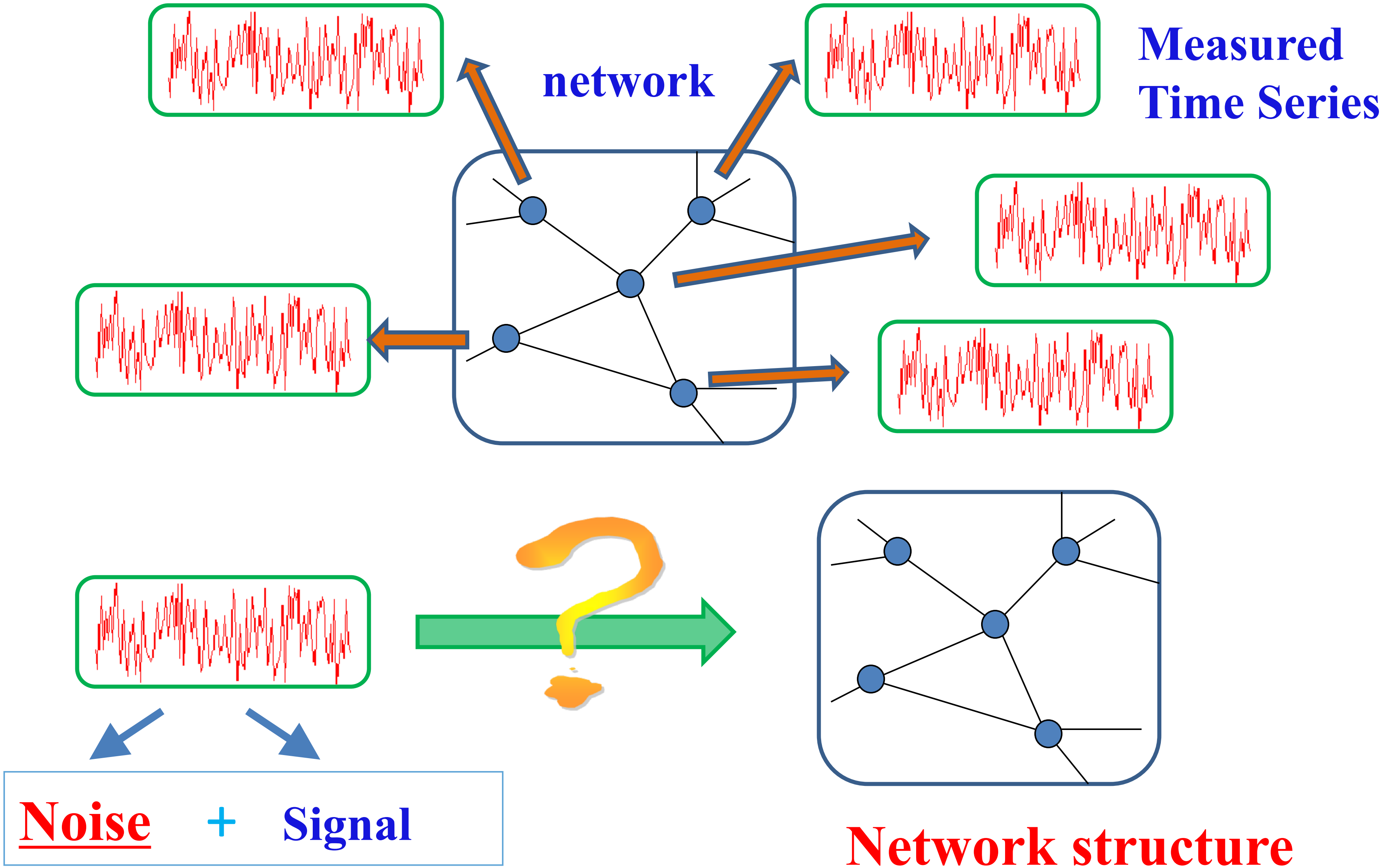
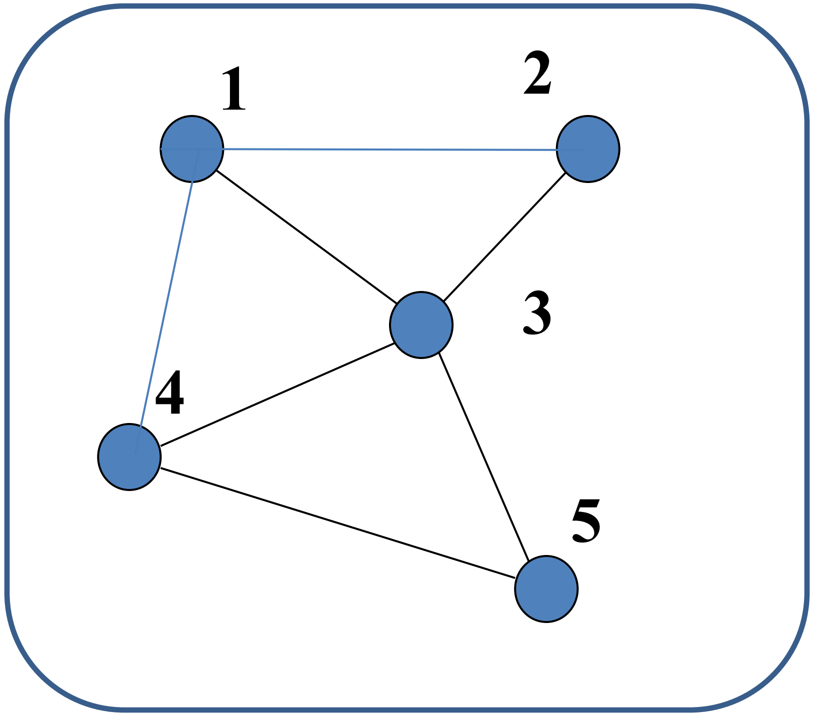
Network Prediction: Inverse Problem

Is it possible to infer network structure from time-series measurements?

Formulation of Problem



Levels of Prediction: Node Degrees and Network Topology



	1	2	3	4	5
Degree	**3**	**2**	**4**	**3**	**2**

Network topology – Laplacian matrix

	1	2	3	4	5
1	**3**	**-1**	**-1**	**-1**	**0**
2	**-1**	**2**	**-1**	**0**	**0**
3	**-1**	**-1**	**4**	**-1**	**-1**
4	**-1**	**0**	**-1**	**3**	**-1**
5	**0**	**0**	**-1**	**-1**	**2**

Results

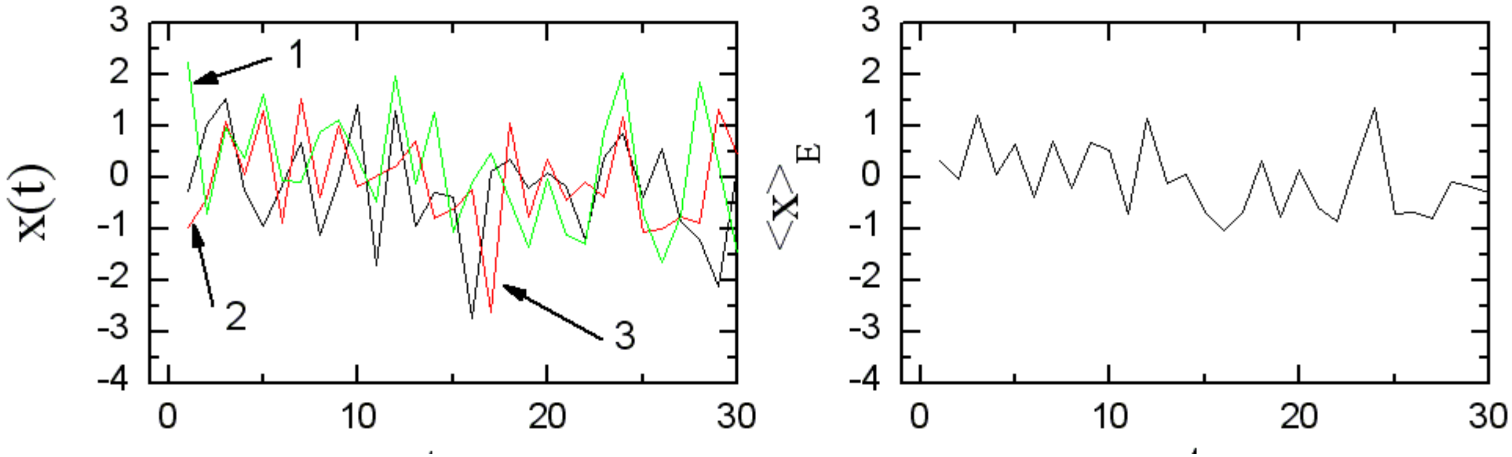
- **Universal scaling law relating noise-induced fluctuations to node degrees (Previous Lecture)**
- **A general method for time-series based prediction of FULL TOPOLOGY of network**

(Theory and numerical verification)

Key idea

Noise-induced fluctuations in observed signals

Noise-Induced Fluctuations



Mean field: $\langle x \rangle_E$ - average value of x(t) over all nodes at any given time.

Average squared fluctuation about mean field:

$$\Delta x_j^2 \equiv \left\langle \left[x_j(t) - \langle x \rangle_E \right]^2 \right\rangle_T$$

Dynamical Correlation and Network Topology

- We discovered a general one-to-one correspondence between **dynamical correlation** and topology

$$\dot{\mathbf{x}}_i = \mathbf{F}_i(\mathbf{x}_i) - c\sum_{j=1}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j) + \eta_i,$$

$$\hat{\mathbf{C}} = \frac{\sigma^2}{2c}\hat{\mathbf{L}}^+$$

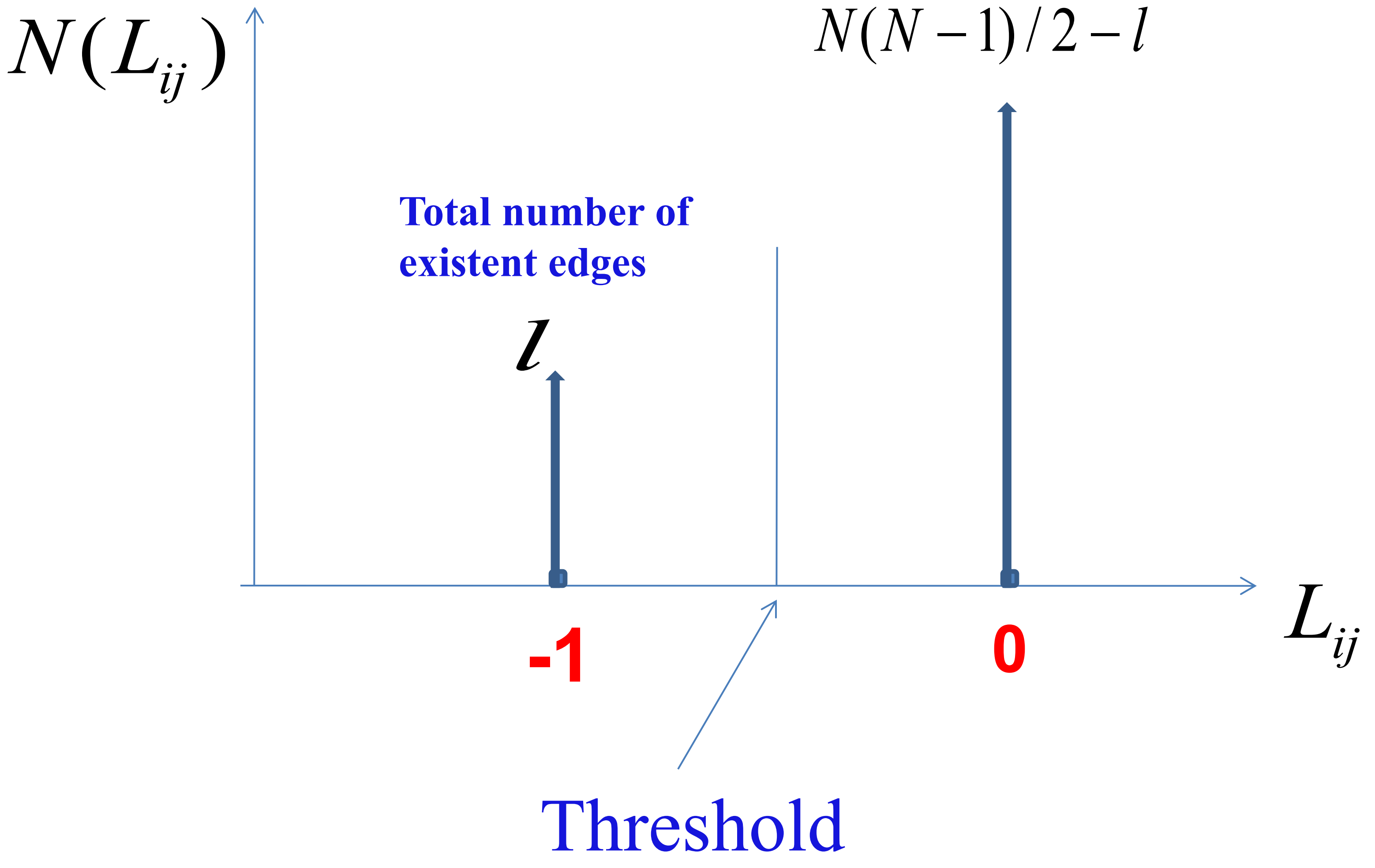
Time series | General Intrinsic dynamics | coupling | Noise

Dynamical correlation matrix ⟷ Network topology

$$C_{ij} = \langle [\mathbf{x}_i(t) - \bar{\mathbf{x}}(t)] \cdot [\mathbf{x}_j(t) - \bar{\mathbf{x}}(t)] \rangle$$

Time series

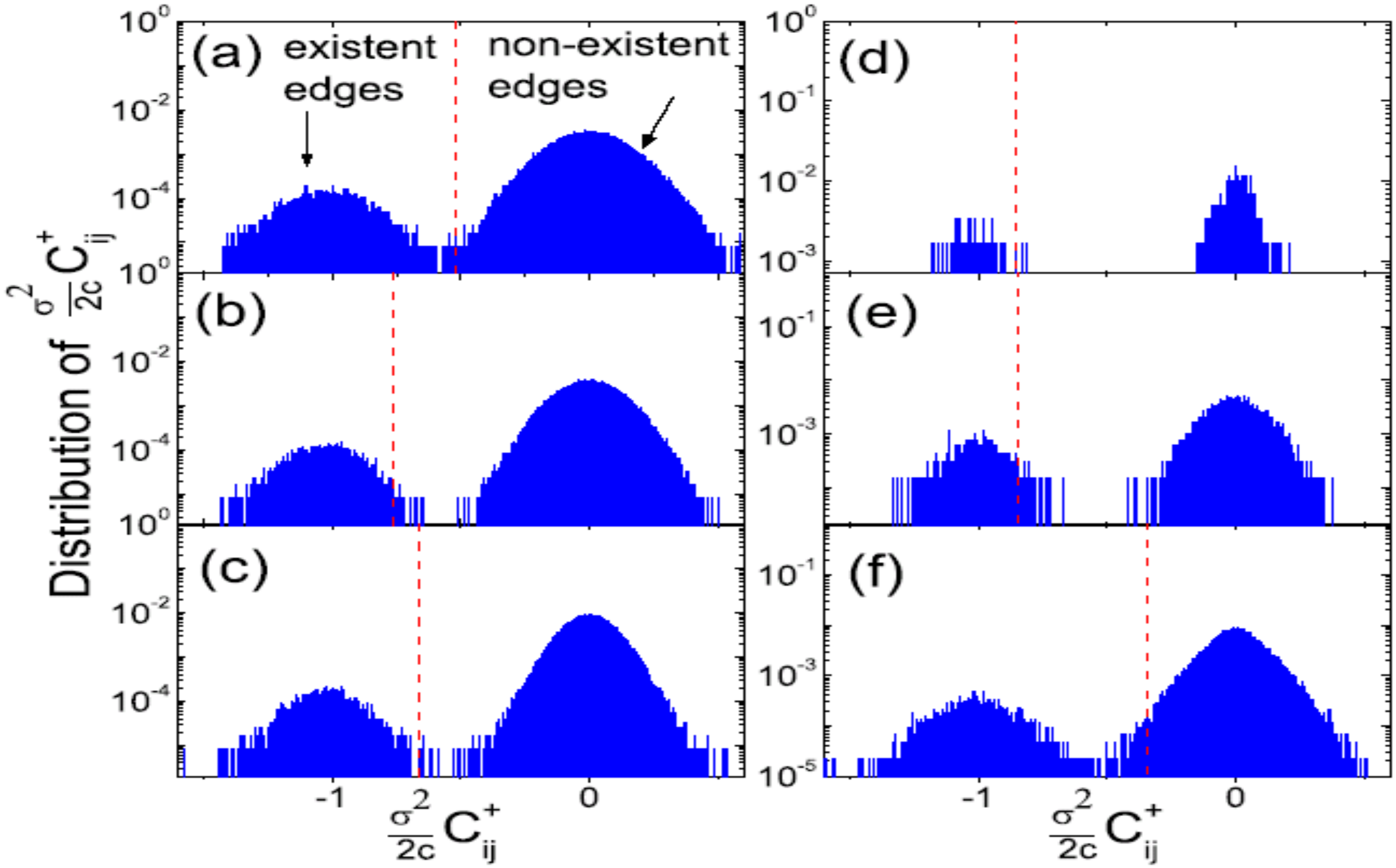
How does it work? – Ideal Case



In Reality – Two-Hump Distribution

$$\hat{\mathbf{C}} = \frac{\sigma^2}{2c}\hat{\mathbf{L}}^+$$

How to set threshold?



Predicting Total Number of Existent Edges

$$\hat{\mathbf{L}} = \frac{\sigma^2}{2c}\hat{\mathbf{C}}^+ \longrightarrow C_{ii} \approx \frac{\sigma^2}{2ck_i}\left(1 + \frac{1}{\langle k \rangle}\right)$$

$$S \equiv \sum_{i=1}^{N} \frac{1}{C_{ii}} = \frac{2cl^2}{\sigma^2 (N + l)}$$

Predicted total number of existent edges

$$l = \frac{S\sigma^2 + \sqrt{S^2\sigma^4 + 8cNS\sigma^2}}{4c}$$

Summing over $\frac{2c}{\sigma^2} C_{ij}$, starting from the most negative value until the sum is equal to l -- practical way of setting the threshold

Success Rates

Four types of dynamical systems

Model networks: Random, Small-world, Scale-free

Six real-world networks: Book, Karate, Football, Elec. Cir., Dolphins, C. Elegans

SREL/SRNL	consensus	I-Rössler	N-Rössler	Kuramoto
Random	1.00/1.00	1.00/1.00	0.995/1.00	0.977/0.999
Small-world	0.993/1.00	0.988/1.00	0.979/1.00	0.982/1.00
Scale-free	0.995/1.00	0.990/1.00	0.980/1.00	0.978/1.00
Book	0.971/1.00	0.977/1.00	0.964/1.00	0.967/1.00
Karate	0.962/1.00	0.962/1.00	0.936/1.00	0.949/1.00
Football	0.938/1.00	0.932/1.00	0.928/1.00	0.927/1.00
Elec. Cir.	0.976/1.00	0.973/1.00	0.971/1.00	0.965/1.00
Dolphins	0.984/1.00	0.981/1.00	0.984/1.00	0.973/1.00
C. Elegans	1.00/0.997	1.00/0.996	1.00/0.997	0.993/0.997

SREL – Success Rate of Existent Links
SRNL – Success Rate of Nonexistent Links

J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, "Noise bridges dynamical correlation and topology in complex oscillator networks," *Physical Review Letters* **104**, 058701 (2010).

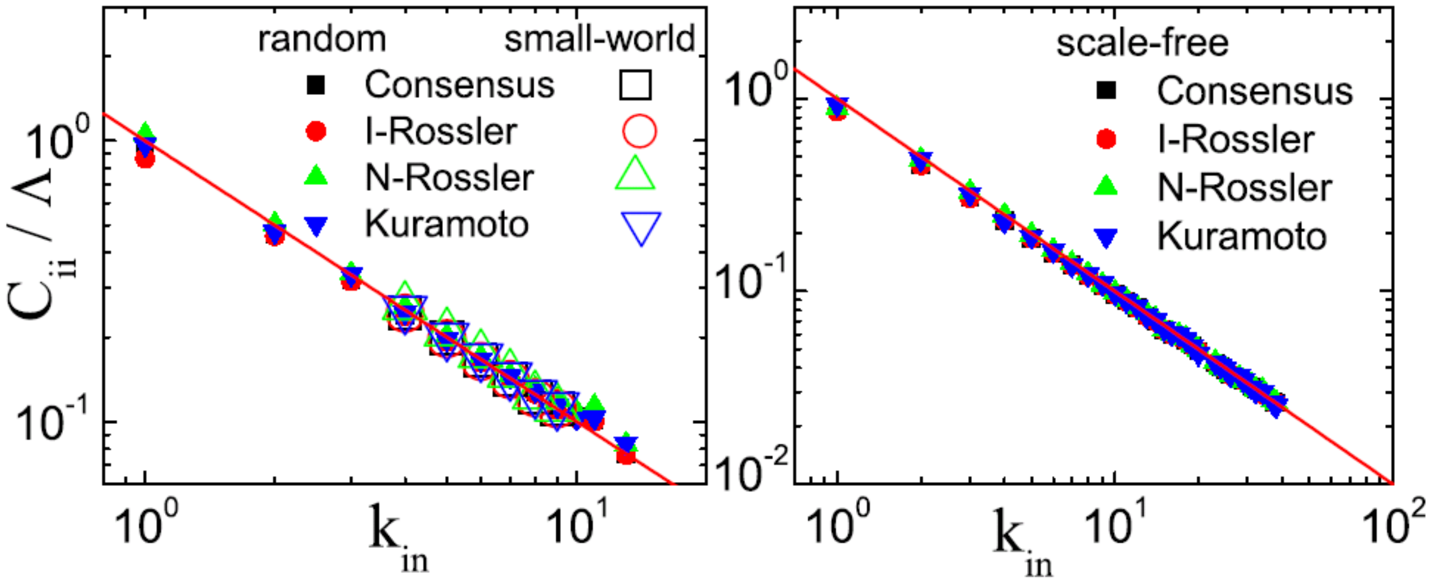
Success Rates versus Edge Density

SREL	consensus			N-Rössler		
$\langle k \rangle$	8	10	12	8	10	12
Random	0.986	0.993	0.996	0.975	0.984	0.989
Small-world	0.952	0.977	0.993	0.935	0.966	0.977
Scale-free	0.986	0.995	0.997	0.964	0.980	0.987

Success rates tend to increase with edge density

Scaling Law of Autocorrelation for Directed Networks

$$C_{ii} \approx \frac{\sigma^2}{2ck_{\text{in}}^i}\left(1+\frac{1}{\langle k_{\text{in}}\rangle}\right),$$ where k_{in}^i is the in - degree of node i



$$\Lambda = \frac{\sigma^2}{2c}\left(1+\frac{1}{\langle k_{\text{in}}\rangle}\right)$$ is a scaling parameter

Network Dynamics with Time Delay

$$\dot{\mathbf{x}}(t) = \mathbf{F}[\mathbf{x}(t)] - c\mathbf{L} \bullet \mathbf{x}(t-\tau) + \eta$$

$\mathbf{C}$: dynamical coupling matrix from time series

$\mathbf{L}$: Laplacian matrix,

$\mathbf{L}^{+}$: pseudo inverse of Laplacian matrix

$$\mathbf{C} = \frac{\sigma^2}{2c}(\mathbf{L}^{+} + c\tau \cdot \mathbf{I}),$$

$$\mathbf{L}'^{+} \equiv \mathbf{L}^{+} + c\tau \cdot \mathbf{I}, \quad \text{so}$$

$$\mathbf{C} = \frac{\sigma^2}{2c}\mathbf{L}'^{+} \Rightarrow \mathbf{L}' = \frac{\sigma^2}{2c}\mathbf{C}^{+}$$

$$\mathbf{L}'_{ij,(i\neq j)} = \mathbf{L}_{ij,(i\neq j)}$$

so the network is predicted by $\mathbf{L}$. We can then obtain the time delay τ as

$$\frac{2c}{\sigma^2}\mathbf{C} - \mathbf{L}^{+} = c\tau \cdot \mathbf{I},$$

$$\Rightarrow \tau \approx \frac{1}{Nc}\sum_{i=1}^{N}\left(\frac{2c}{\sigma^2}\mathbf{C} - \mathbf{L}^{+}\right)_{ii}$$

Predicting Time Delay - Improved Method

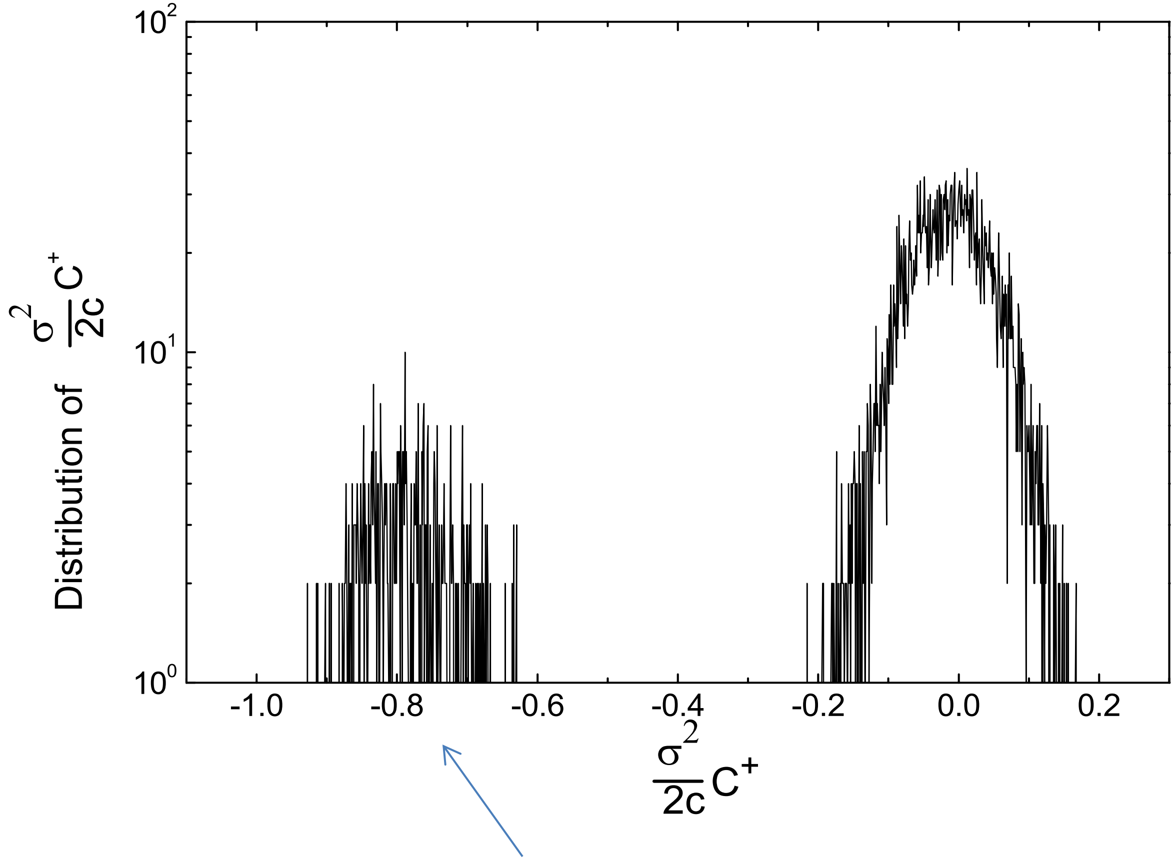
$$\mathbf{C}^{+} = \frac{\sigma^2}{2c}[\mathbf{L} - c\tau \cdot \mathbf{L}^2],$$

$$\Rightarrow \mathbf{L} - \frac{2c}{\sigma^2}\mathbf{C}^{+} = c\tau \cdot \mathbf{L}^2$$

$$\tau \approx \left\langle \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\left(\mathbf{L} - \frac{2c}{\sigma^2}\mathbf{C}\right)_{ij}}{c\mathbf{L}^2{}_{ij}} \right\rangle_{i \neq j, \mathbf{L}_{ij} \neq 0, \mathbf{L}^2{}_{ij} \neq 0}$$

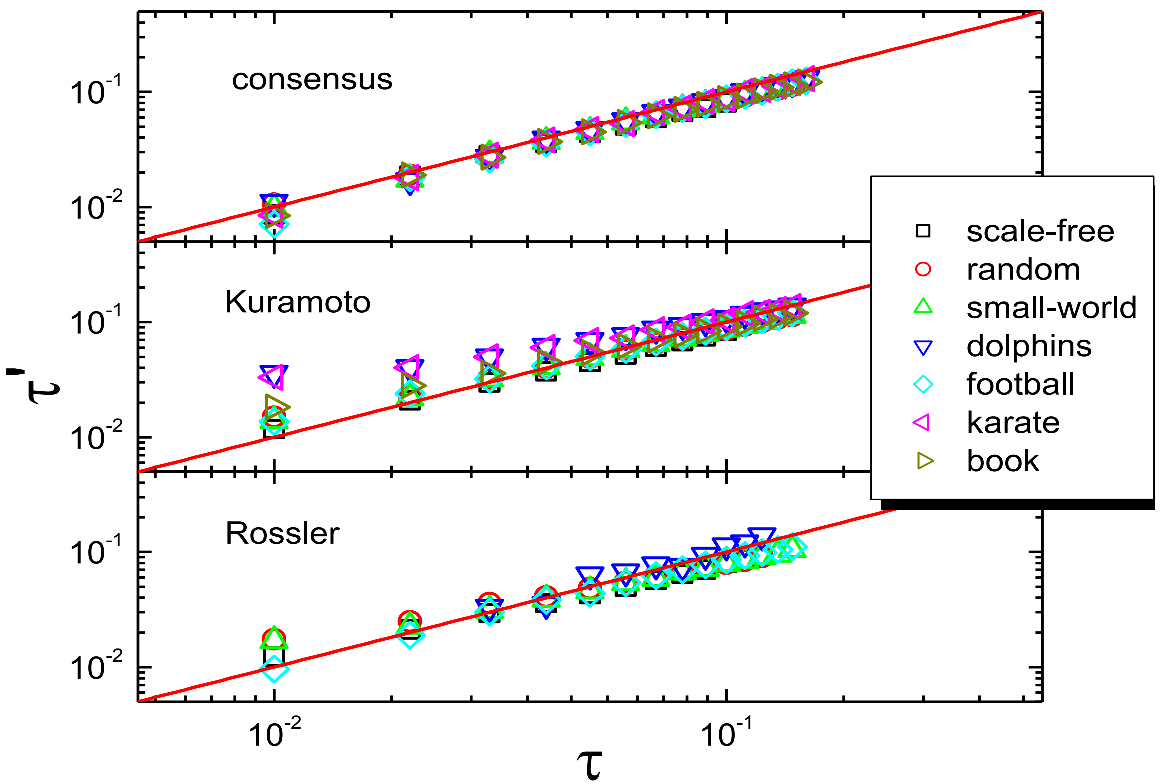
J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, "Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series," *Chaos* **22**, 033131 (2012).

Distribution of Dynamical Correlation Matrix in the Presence of Time Delay



Center of distribution different from - 1 due to the term $c\tau \cdot \mathbf{I}$

Predicting Time delay
- Examples



J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, “Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series,” *Chaos* **22**, 033131 (2012).

Theory

Dynamical correlation ⟶ Network matrix

$$\dot{\mathrm{x}}_i = \mathbf{F}_i(\mathrm{x}_i) - c\sum_{j=1}^{N} L_{ij}\mathbf{H}(\mathrm{x}_j) + \eta_i, \qquad (1)$$

where

c - coupling strength, $\mathbf{H}$ - coupling function of oscillators, η_i - noise term,

$L_{ij} = -1$ if j connects to i (otherwise 0) for $i \neq j$ and $L_{ii} = -\sum_{j=1, j\neq i}^{N} L_{ij}$.

Let $\bar{\mathrm{x}}_i$ be the counterpart of x_i in the absence of noise, and assume a small perturbation ξ_i : $\mathrm{x}_i = \bar{\mathrm{x}}_i + \xi_i$.

Variational equation :

$$\dot{\xi} = [D\hat{\mathbf{F}}(\bar{\mathrm{x}}) - c\hat{\mathbf{L}} \otimes D\hat{\mathbf{H}}(\bar{\mathrm{x}})]\xi + \eta, \qquad (2)$$

where $\xi = [\xi_1, \xi_2, \cdots, \xi_N]^T$, $\eta = [\eta_1, \eta_2, \cdots, \eta_N]^{\mathrm{T}}$ is the noise vector, $\hat{\mathbf{L}}$ is the Laplacian matrix,

$D\hat{\mathbf{F}}(\bar{\mathrm{x}}) = \mathrm{diag}[D\hat{\mathbf{F}}(\bar{\mathrm{x}}_1), D\hat{\mathbf{F}}(\bar{\mathrm{x}}_2), \cdots, D\hat{\mathbf{F}}(\bar{\mathrm{x}}_{\mathrm{N}})]$, $\otimes$ denotes direct product, and

$D\hat{\mathbf{H}}$ is the Jacobian matrix of $\mathbf{H}$.

Let $\left\langle \xi\xi^T \right\rangle = \hat{\mathbf{C}},$ where $C_{ij} = \left\langle \xi_i \xi_j \right\rangle$ and $\langle . \rangle$ is time average. We have

$$0 = \left\langle d(\xi\xi^T)/dt \right\rangle = -\hat{\mathbf{B}}\hat{\mathbf{C}} - \hat{\mathbf{C}}\hat{\mathbf{B}} + \left\langle \eta\xi^T \right\rangle + \left\langle \xi\eta^T \right\rangle, \qquad (3)$$

where $\hat{\mathbf{B}} = -D\hat{\mathbf{F}}(\bar{\mathrm{x}}) - c\hat{\mathbf{L}} \otimes D\hat{\mathbf{H}}(\bar{\mathrm{x}})$. To obtain the expression of $\left\langle \eta\xi^T \right\rangle$ and $\left\langle \xi\eta^T \right\rangle$, we get the solution $\xi(t)$ from Eq. (2):

$$\xi(t) = \hat{\mathbf{G}}(t-t_0)\xi(t_0) + \int_{t_0}^{t} dt' \hat{\mathbf{G}}(t-t')\eta(t'),$$

where $\hat{\mathbf{G}} = \exp(-\hat{\mathbf{B}}t)$. In the absence of divergence of state variables, $\hat{\mathbf{G}}(\infty) = 0$. Setting $t_0 \to -\infty$, without loss of generality, we have

$$\xi(t) = \int_{t_0}^{t} dt' \hat{\mathbf{G}}(t-t')\eta(t').$$

Note that $\hat{\mathbf{G}}(0) = \hat{\mathbf{I}},$ we hence obtain

$$\left\langle \xi\eta^T \right\rangle = \int_{-\infty}^{t} \hat{\mathbf{G}}(t-t') \left\langle \eta(t)\eta^T(t') \right\rangle dt' = \int_{-\infty}^{t} \hat{\mathbf{G}}(t-t')\hat{\mathbf{D}}\delta(t-t')dt' = \frac{\hat{\mathbf{D}}}{2},$$

where $\hat{\mathbf{D}}$ is the covariance matrix of noise. Analogously, we have

$$\left\langle \eta\xi^T \right\rangle = \frac{\hat{\mathbf{D}}}{2}.$$

Therefore, Eq. (3) can be simplified to:

$$\hat{\mathbf{B}}\hat{\mathbf{C}} + \hat{\mathbf{C}}\hat{\mathbf{B}}^T = \hat{\mathbf{D}}. \qquad (4)$$

Since $\hat{\mathbf{B}} = -D\hat{\mathbf{F}}(\overline{\mathrm{x}}) + c\hat{\mathbf{L}} \otimes D\hat{\mathbf{H}}(\overline{\mathrm{x}})$, the above equality reveals a general relationship between the dynamical correlation $\hat{\mathbf{C}}$ and the connecting matrix $\hat{\mathbf{L}}$ in the presence of noise as characterized by $\hat{\mathbf{D}}$. The general solution of $\hat{\mathbf{C}}$ can be written as

$$\mathrm{vec}(\hat{\mathbf{C}}) = \frac{\mathrm{vec}(\hat{\mathbf{D}})}{\hat{\mathbf{I}} \otimes \hat{\mathbf{B}} + \hat{\mathbf{B}} \otimes \hat{\mathbf{I}}},$$

where $\mathrm{vec}(\hat{\mathbf{X}})$ is a vector containing all columns of matrix $\hat{\mathbf{X}}$.

Consider one-dimensional state variable and linear coupling such that $D\hat{\mathbf{H}} = 1$, with Gaussian white noise $\hat{\mathbf{D}} = \sigma^2\hat{\mathbf{I}}$, and further regard the intrinsic dynamics $D\hat{\mathbf{F}}$ as small perturbations. Then Eq. (4) can be simplified to

$$\hat{\mathbf{L}}\hat{\mathbf{C}} + \hat{\mathbf{C}}\hat{\mathbf{L}} = \frac{\sigma^2\hat{\mathbf{I}}}{c}.$$

For an undirected network with symmetric coupling matrix, the solution of $\hat{\mathbf{C}}$ can be expressed as

$$\hat{\mathbf{C}} = \frac{\sigma^2}{2c}\hat{\mathbf{L}}^+, \qquad (5)$$

where $\hat{\mathbf{L}}^+$ denotes the pseudo inverse of the Laplacian matrix.

Path integral → Local structures

We can decompose $\hat{\mathbf{L}}$ into two parts: $\hat{\mathbf{L}} = \hat{\mathbf{K}} - \hat{\mathbf{A}}$,

where $\hat{\mathbf{A}}$ is the adjacency matrix with $\mathrm{A}_{ij} = 1$ if node j connects i, otherwise 0, and

$\hat{\mathbf{K}} = \mathrm{diag}(k_1, \cdots, k_N)$, where k_i is the degree of node i.

The matrix $\hat{\mathbf{C}}$ can thus be expressed in a series:

$$\hat{\mathbf{C}} \sim (\hat{\mathbf{K}} - \hat{\mathbf{A}})^{-1} = \hat{\mathbf{K}}^{-1} + \hat{\mathbf{K}}^{-1}\hat{\mathbf{A}}\hat{\mathbf{K}}^{-1} + \hat{\mathbf{K}}^{-1}\hat{\mathbf{A}}\hat{\mathbf{K}}^{-1}\hat{\mathbf{A}}\hat{\mathbf{K}}^{-1} + \cdots.$$

For the second term $\hat{\mathbf{K}}^{-1}\hat{\mathbf{A}}\hat{\mathbf{K}}^{-1}$, if nodes j connects to i, its element (i, j) is $(k_i k_j)^{-1}$ and otherwise 0.

For the third term $\hat{\mathbf{K}}^{-1}\hat{\mathbf{A}}\hat{\mathbf{K}}^{-1}\hat{\mathbf{A}}\hat{\mathbf{K}}^{-1}$, if there are multiple two-step paths connecting j to i through node m_1 or $m_2, \cdots, m_r$, its element (i, j) is $k_i^{-1}(\sum_{q=m_1}^{m_r} k_q^{-1})k_j^{-1}$.

We thus have $$C_{ij} = \frac{\sigma^2}{2c}\sum_{\text{path}}\prod_{m\in\text{path}}\frac{1}{k_m}, \qquad (6)$$

where "path" means all paths from j to i, and m denotes the nodes on them.

This path-integral represetation directly reveals a relation between autocorrelation C_{ii} in the matrix $\hat{\mathbf{C}}$ and node degree k_i.

For nth-order approximation, we count all paths whose lengths are equal to or less than n. Under second-order approximation, we have

$$C_{ii} = \frac{\sigma^2}{2c}\left(\frac{1}{k_i} + \frac{1}{k_i^2}\sum_{q\in\Gamma_i}\frac{1}{k_q}\right) \approx \frac{\sigma^2}{2ck_i}\left(1 + \frac{1}{\langle k\rangle}\right), \qquad (7)$$

where mean-field approximation is applied and Γ_i denotes the neighbors of node i.

Conclusions

1. **In the presence of noise, dynamical correlation matrix reveals full topology of network.**

   J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, "Noise bridges dynamical correlation and topology in complex oscillator networks," *Physical Review Letters* **104**, 058701 (2010).

2. **The theory works even when there is small time delay**

   J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, "Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series," *Chaos* **22**, 033131 (2012).

3. **Compressive-sensing based methods for time-series based prediction of network topology and dynamics**

   Next Lecture