Reconstructing Network connectivity and weights from Noisy Dynamical Data

Pik-Yin Lai (黎璧賢)
Dept. of Physics, & Center for Complex Systems, National Central University, Chung-Li, Taiwan 320
Email: pylai@phy.ncu.edu.tw

• Introduction
• Noisy perturbations reveal network interactions
• Reconstruction formulae
• Reconstructing the Adjacency Matrix, relative weights,…
• Applications & Outlook
"Big data" era
huge amount of dynamical data in many different disciplines are all waiting to be mined with novel and reliable methods of network reconstruction.
Network Reconstruction

Difficult/impossible to directly measure the network topology or coupling strengths.

How to infer the wiring diagram and reconstruct a network from measurements?

Passive measurement of the dynamics (without perturbing the system)

Challenging Inverse Problem!
• Cross correlation: **Pearson correlation**, partial correlation,…
• Information flow: mutual information, transfer entropy,…
• Granger Causality, Cross-convergence map,…

Constructing a gene co-expression network, S. Mohammad H. Oloomi, Wikipedia
infer links from correlation of dynamical measurements, with a higher correlation interpreted as a higher probability of a link

\[ c_{ij} = \frac{1}{N_{ij}} \int \Delta I_i(t) \Delta I_j(t) dt \]

“Functional connectivity”: i and j are connected if \( C_{ij} > \text{threshold} \)

**Firing activity in cultured neuronal network:**

Statistical correlation or dependence does not necessarily result from direct interaction between nodes. 

Indirect connections → correlations → wrongly inferred links

Noisy data are helpful for Network Reconstruction

Noise is ubiquitous in natural and technological systems

Noise provide small and sustained perturbations to the network

Relaxation of from such (independent) noisy perturbations reveals the underlying interaction topology of the network

\[ \dot{x}_i = f(x_i) + \sum_{j \neq i} g_{ij} A_{ij} h(x_i, x_j) + \eta_i \]

Fluctuation from the stable fixed point $X_0$:
\[ \delta x_i \equiv x_i - X_0 \]
\[ \dot{\delta x_i} \approx - \sum_{j=1}^{N} \left[ h'(0) L_{ij} - f'(X_0) \delta_{ij} \right] \delta x_j + \eta_i \]

Weighted Laplacian matrix:
\[ L_{ij} = s_i \delta_{ij} - g_{ij} A_{ij}, \quad s_i \equiv \sum_{j=1}^{N} g_{ij} A_{ij} \]

Dynamic correlation function:
\[ C_{ij} = \langle [x_i(t) - X(t)][x_j(t) - X(t)] \rangle_T \]
\[ X(t) \equiv (1/N) \sum_{i=1}^{N} x_i(t). \]

**Network Reconstruction formula:** (weighted undirected network)

\[
\frac{\sigma^2}{2} \lim_{T \to \infty} \overline{C}_{ij}^+ = h'(0)\mathcal{L}_{ij} - f'(X_0) \left( \delta_{ij} - \frac{1}{N} \right)
\]

Pseudo-inverse of \( C \)

Generate known network of known dynamics to verify:

\[-(\sigma^2/2)C_{ij}^+ \approx h'(0)g_{ij}, \quad i \neq j,\]

Weighted random network with Gaussian distributed \( g_{ij} \)
Weighted random network with Logistic dynamics: \( f(x) = 10x(1-x) \)

\[
\frac{\sigma^2}{2} \lim_{T \to \infty} \overline{C}_{ij}^+ = h'(0)\mathcal{L}_{ij} - f'(X_0) \left( \delta_{ij} - \frac{1}{N} \right)
\]

\[
\frac{\sigma^2}{2} C_{ii}^+ \simeq h'(0)s_i - f'(X_0)
\]

\[ s_i \equiv \sum_{j=1}^{N} g_{ij}A_{ij} \]

Logistic (1D)

Node strength/weighted degree

Generate known network of known dynamics verifying the reconstruction formula.
Reconstructing $A_{ij}$ solely from time-series data:

Values of $r_{ij}$ form two groups, by identifying these two groups $\Rightarrow$ reconstruction of $A_{ij}$.

Clustering analysis: Gaussian mixture model $\Rightarrow$ 2 groups of $A_{ij}=0$ or 1

No knowledge of $f$, $h$, $g$, noise is needed!
Reconstructing $A_{ij}$ solely from time-series data

$N=100$ random network with Gaussian weights, Logistic dynamics

No knowledge of $f, h, g, \text{noise, ...}$ is needed!
Reconstruction of $A_{ij}$

Sensitivity

$P_{\text{SEN}} \equiv \frac{100N_{TP}}{N_{TP} + N_{FN}} \%$, \quad $P_{\text{SPEC}} \equiv \frac{100N_{TN}}{N_{TN} + N_{FP}} \%$

<table>
<thead>
<tr>
<th>Case</th>
<th>Network</th>
<th>Dynamics</th>
<th>$P_{\text{SEN}}$</th>
<th>$P_{\text{SPEC}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>random; $g_{ij} = 1$</td>
<td>consensus</td>
<td>99.9</td>
<td>99.9</td>
</tr>
<tr>
<td>2</td>
<td>random; $g_{ij} = 10$</td>
<td>FHN</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>scale-free; $g_{ij} = 1$</td>
<td>consensus</td>
<td>100</td>
<td>100.00</td>
</tr>
<tr>
<td>4</td>
<td>scale-free; $g_{ij} = 10$</td>
<td>FHN</td>
<td>97.4</td>
<td>100.00</td>
</tr>
<tr>
<td>5</td>
<td>WR; $\mu = 5, \gamma = 2$</td>
<td>consensus</td>
<td>94.3</td>
<td>100.00</td>
</tr>
<tr>
<td>6</td>
<td>WR; $\mu = 10, \gamma = 2$</td>
<td>logistic</td>
<td>99.9</td>
<td>100.00</td>
</tr>
<tr>
<td>7</td>
<td>WR; $\mu = 10, \gamma = 2$</td>
<td>cubic</td>
<td>96.8</td>
<td>99.7</td>
</tr>
<tr>
<td>8</td>
<td>WR; $\mu = 10, \gamma = 2$</td>
<td>FHN</td>
<td>99.5</td>
<td>100.00</td>
</tr>
<tr>
<td>9</td>
<td>weighted scale-free</td>
<td>consensus</td>
<td>91.7</td>
<td>99.98</td>
</tr>
<tr>
<td>10</td>
<td>weighted scale-free</td>
<td>FHN</td>
<td>86.7</td>
<td>99.98</td>
</tr>
</tbody>
</table>

Table 1: Accuracy of our method measured by $P_{\text{SEN}}$ and $P_{\text{SPEC}}$ for various networks with consensus $[f = 0; h(z) = z]$, logistic $[f(x) = 10x(1 - x); h(z) = z]$, cubic $[f = 0; h(z) = z^3]$, and FHN dynamics.
Reconstructing relative weights & node strengths:

\[
G_{ij} \equiv \frac{g_{ij}}{\langle g \rangle} \approx \frac{C_{ij}^{+} \sum_{l} k_{l}^{(e)}}{\sum_{n,l \leftrightarrow n} C_{nl}^{+}} = G_{ij}^{(e)},
\]

\[
S_{i} \equiv \frac{s_{i}}{\langle s \rangle} = \frac{\sum_{j} G_{ij} A_{ij}}{\langle k \rangle} \approx \frac{N \sum_{j \leftrightarrow i} C_{ij}^{+}}{\sum_{n,l \leftrightarrow n} C_{nl}^{+}} = S_{i}^{(e)},
\]

\[
\langle g \rangle \equiv \frac{\sum_{i,j} g_{ij} A_{ij}}{\sum_{i,j} A_{ij}}
\]

WR network with Rössler dynamics
Reconstructing relative weights solely from time-series data

Original Gij

Gij predicted

N=100 random network with Gaussian weights, Logistic dynamics

Difference
Global network properties: eigenvalue spectrum of $A_{ij}$

\[
\rho(\lambda) = \begin{cases} 
\frac{\sqrt{4Np(1-p)} - \lambda^2}{2\pi Np(1-p)}, & |\lambda| < 2\sqrt{Np(1-p)}, \\
0, & \text{otherwise},
\end{cases}
\]

- FIG. 8. (Color online) Comparison of the actual eigenvalue spectrum (circles) for the random network with $N = 100$ with the extracted eigenvalue spectrum for consensus dynamics with $g = \sigma = 1$ (squares), FHN1 dynamics with $\sigma = 1$ and $g = 10$ (triangles), and FHN2 dynamics with $\sigma = 1$ and $g = 10$ (diamonds), respectively. The analytical result for the actual spectrum for $N \to \infty$ is also shown as the solid line.
Global network properties: weight distributions

Weighted random & scale-free networks with various intrinsic node dynamics

FIG. 3. (Color online) Comparison of $P(G)$ of the reconstructed $G_{ij}^{(e)}$ for consensus (circles), Rössler (triangles), FHN1 (squares), and FHN2 (diamonds) dynamics with $T_{av} = 1000$ against the actual distribution (solid line) for (a) WR1 and WR2, (b) WSF1, and (c) WSF2 networks. Dashed line in (b) and (c) is the result for FHN2 dynamics with $T_{av} = 5000$. 
Reconstruction using correlations doomed to fail:

Covariance matrix

\[ K_{ij} = \langle [x_i(t) - \langle x_i(t) \rangle_T][x_j(t) - \langle x_j(t) \rangle_T] \rangle_T \]

Reconstruction formula:

\[ \frac{\sigma^2}{2} K^{-1}_{ij} = -Q = h'(0) \mathcal{L} - f'(X_0) I \]

\[ K^{-1}_{ij} \approx -\frac{2h'(0)}{\sigma^2} g_{ij} A_{ij}, \quad i \neq j \]

\[ K^{-1}_{ii} \approx \frac{2h'(0)}{\sigma^2} s_i - f'(X_0). \]

Pearson correlation

\[ \frac{K_{ij}}{\sqrt{K_{ii}K_{jj}}} \]

Partial correlation

\[ -\frac{K^{-1}_{ij}}{\sqrt{K^{-1}_{ii}K^{-1}_{jj}}} \]

indirect interactions

\[ Q_{ij} = \sum_k Q_{ik} Q_{kj} \]
Reconstruction using correlations doomed to fail:

Covariance matrix \( K_{ij} = \langle [x_i(t) - \langle x_i(t) \rangle_T][x_j(t) - \langle x_j(t) \rangle_T] \rangle_T \)

\[
\frac{\sigma^2}{2} K_{ij}^{-1} = -Q \equiv h'(0) \mathcal{L} - f'(X_0) I
\]

Pearson correlation

\[
K_{ij} \approx \frac{2h'(0)}{\sigma^2} g_{ij} A_{ij}, \quad i \neq j
\]

Partial correlation

\[
K_{ij}^{-1} \approx \frac{2h'(0)}{\sigma^2} s_i - f'(X_0).
\]

\[
\frac{K_{ij}}{\sqrt{K_{ii}K_{jj}}} \quad \quad \quad \frac{-K_{ij}}{\sqrt{K_{ii}^{-1}K_{jj}^{-1}}}
\]

<table>
<thead>
<tr>
<th>Case</th>
<th>Method 1</th>
<th>Method 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( P_{\text{SEN}} )</td>
<td>( P_{\text{SPEC}} )</td>
</tr>
<tr>
<td>1</td>
<td>39.0</td>
<td>82.9</td>
</tr>
<tr>
<td>2</td>
<td>49.8</td>
<td>86.1</td>
</tr>
<tr>
<td>3</td>
<td>8.97</td>
<td>99.6</td>
</tr>
<tr>
<td>4</td>
<td>34.6</td>
<td>99.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case</th>
<th>Network</th>
<th>Dynamics</th>
<th>( P_{\text{SEN}} )</th>
<th>( P_{\text{SPEC}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>random; ( g_{ij} = 1 )</td>
<td>consensus</td>
<td>99.9</td>
<td>99.9</td>
</tr>
<tr>
<td>2</td>
<td>random; ( g_{ij} = 10 )</td>
<td>FHN</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>scale-free; ( g_{ij} = 1 )</td>
<td>consensus</td>
<td>100</td>
<td>100.00</td>
</tr>
<tr>
<td>4</td>
<td>scale-free; ( g_{ii} = 10 )</td>
<td>FHN</td>
<td>97.4</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Comparison with different Reconstruction methods:

\[ \frac{K_{ij}}{\sqrt{K_{ii}K_{jj}}} \]

Pearson correlation

\[ \frac{K_{ij}^{-1}}{\sqrt{K_{ii}^{-1}K_{jj}^{-1}}} \]

partial correlation

\[ K_{ij}^{-1} \approx -\frac{2h'(0)}{\sigma^2}g_{ij}A_{ij}, \quad i \neq j \]

Actual scale free network
Reconstructing Neuronal Network in vitro (C.K. Chan’s Lab.)

Multi-electrode Array
Reconstructing Neuronal Network in vitro

- Multi-electrode Array BioCam 4096 systems: 64x64 nodes
- Synaptic connections are directed, gap-junctions are undirected
- Identify excitatory/inhibitory neurons and synapses
- 1 electrode gets contributions from spiking activities of several nearby neurons
- Reconstructing network of MEA nodes
Neuronal Spiking activity in vitro: (BioCam 4096 MEA) embryonic rat cortical neuronal culture

Waves $\rightarrow$ short range connections
Reconstructing Neuronal Network in vitro (MEA 4096 nodes)

Waves

Short-range connections with some distant connections

Reconstructed Adjacency matrix
Identify group of strongly interacting gang in a crowd.

Detect the hubs of adversarial social/electronic networks by monitoring a proper public area/net.
Network reconstruction from real-time dynamical data

Identify hub or special nodes/links to launch attacks or real-time feedback control to the desire dynamical state.
Reconstructing Biological Networks

- Inferring gene regulatory networks from observational expression data
- Network of Brain functional regions from fMRI data, EEG,…

Probing organism communication Network in a population

- Motion, flapping or flashing signals
- Existence of leadership
- Workload distribution
- Social organization in insect or animal populations
Summary

- Noise bridges correlations and network topology/coupling strengths.

- **Network Reconstruction Formulae derived: undirected network**

  \[
  \frac{\sigma^2}{2} \lim_{T \to \infty} \bar{C}_{ij}^+ = h'(0) \mathcal{L}_{ij} - f'(X_0) \left( \delta_{ij} - \frac{1}{N} \right)
  \]

  \[
  K_{ij}^{-1} \approx -\frac{2h'(0)}{\sigma^2} g_{ij} A_{ij} , \quad i \neq j
  \]

  \[
  K_{ii}^{-1} \approx \frac{2h'(0)}{\sigma^2} s_i - f'(X_0)
  \]

- Reconstruction scheme for Adjacency matrix and relative weights using only time-series data are proposed & verified for various intrinsic node dynamics, and different network topologies.

- **Model-free method for revealing interaction between nodes in many-body coupled network from dynamical measurements, in physical and biological systems**
Collaborators:
Prof. Emily S.C. Ching (CUHK)
C. Y. Leung, H.C. Tam

Neuron/MEA expt:
Prof. C. K. Chan (Academia Sinica/NCU)
Dr. Y. T. Huang, Dr. C.H. Huang


pylai@phy.ncu.edu.tw

Thank you for your attention