Reconstruction of Network Structure from Dynamics

Emily S.C. Ching
Department of Physics
The Chinese University of Hong Kong

Supported by Hong Kong Research Grants Council
(Grants No. CUHK 14300914 and 14304017)

Big Data Challenges for Predictive Modeling of Complex Systems
November 26, 2018
Outline

• Introduction

• Noise-induced Relations

• Implications
  – Why statistical covariance is not a good indicator of interaction
  – Reconstruction method for general networks

• Effects of Hidden Nodes

• Application to Real-world Networks

• Summary

Joey C.Y. Leung (former PhD student)
H.C. Tam (former PhD student)
P.Y. Lai (National Central University)

Benny P.H. Tam (former MPhil student)

K.C. Lin (former MPhil student)
Chumin Sun (PhD student)
Introduction

Many complex systems of interest in physics and biology can be represented as networks of a number of elementary units (nodes) that interact (link) with one another.

Yeast protein-protein interaction network

Lethality and centrality in protein networks
An important piece of information about a network is its connectivity structure, namely how its nodes are linked to one another, and the coupling strength of these links. Knowing the network structure is an important first step for understanding the functionality of behavior of the system.

Network biology: understanding the cell's functional organization
Albert-László Barabási & Zoltán N. Oltvai
Nature Reviews Genetics 5, 101-113 (February 2004)
A large amount of dynamical data of complex networks has been collected in many different disciplines. One “big data” challenge is to reconstruct the structure of a network from the measured data.
Another central issue is to understand the relation between the connectivity structure of the network and its dynamics.

Our research goal:
To infer the network connectivity structure from the dynamics of the nodes

**Complex brain networks: graph theoretical analysis of structural and functional systems**
Consider a network of $N$ nodes, each with a state variable $x_i(t)$, described by the following general model:

$$\dot{x}_i = f_i(x_i) + \sum_{j=1}^{N} g_{ij} A_{ij} h(x_i, x_j) + \eta_i \quad i = 1, 2, \ldots, N$$

- **Intrinsic dynamics**: $f_i(x_i)$
- **Coupling function**: $g_{ij} A_{ij} h(x_i, x_j)$
- **Adjacency matrix**: $A_{ij}$ giving the links of the network: $A_{ij} = 1$ if a link connects node to node $i$ or 0 otherwise.
- **Coupling strength**: $g_{ij}$ (weighted, bidirectional if $A_{ij} = A_{ji}$ and unweighted if $g_{ij} = g$)
- **Noise**: $\eta_i$ models external disturbance

In general, $A_{ij} \neq A_{ji}$ (directed) and non-uniform $g_{ij}$ (weighted)
Can we relate structure (\(A_{ij}\) and \(g_{ij}\)) to dynamics \([x_i(t), i=1, \ldots, N]\)?

It was found that the presence of noise gives rise to a relation between structure and dynamics.

The specific relation presented in this work applies only for bidirectional, unweighted networks with consensus dynamics: \(f_i=0\) and \(h(x_i, x_j)=x_j-x_i\) in the presence of white noise.
We focus on systems that have stationary fluctuations about a noise-free steady-state $X_i$ and study the linearized dynamics around $X_i$.

$$\frac{d}{dt} \delta x_i = \sum_{j=1}^{N} Q_{ij} \delta x_j + \eta_i \quad i = 1, 2, \ldots, N$$

$$\delta x_i = x_i - X_i$$

$$Q_{ij} = g_{ij} A_{ij} h_y (X_i, X_j) + \left[ \sum_{k \neq i}^{N} g_{ik} A_{ik} h_x (X_i, X_k) + f_i'(X_i) \right] \delta_{ij}$$

$Q_{ij}$ ($i \neq j$) contains information of the network connectivity
--- separate into 2 groups depending on $A_{ij} = 0$ or 1

Our approach is to derive mathematical results relating $Q$ and quantities that can be calculated using only $x_i(t)$ for general networks and noise and to use these relations to develop methods of network reconstruction.
correlated noise of correlation time $\tau_n$:

$$\eta_i(t) = 0 ; \quad \eta_i(t)\eta_j(t') = \frac{D_{ij}}{\tau_n} e^{-|t-t'|/\tau_n}$$

white noise:

$$\eta_i(t) = 0 ; \quad \eta_i(t)\eta_j(t') = 2D_{ij}\delta(t-t')$$
Noise-induced Relations

Define \( (K_\tau)_{ij} = \frac{(x_i(t + \tau) - \bar{x}_i(t + \tau))(x_j(t) - \bar{x}_j(t))}{\langle\delta x_i(t + \tau) - \delta x_i(t + \tau)\rangle\langle\delta x_j(t) - \delta x_j(t)\rangle} \)

\[= \frac{(\delta x_i(t + \tau) - \delta x_i(t + \tau))(\delta x_j(t) - \delta x_j(t))}{\langle\delta x_i(t + \tau) - \delta x_i(t + \tau)\rangle\langle\delta x_j(t) - \delta x_j(t)\rangle} \]

\([K_0\text{ is the usual covariance matrix}]\)

Then solving the linearized stochastic dynamics for \(\delta x_i(t)\) we have obtained two mathematical results (which are good approximations in the weak-noise limit) relating \(K_0, K_\tau, Q\) and \(D\) (noise covariance matrix).
Result 1:
\[
QK_0 + K_0Q^T + D(I - \tau_n Q^T)^{-1} + (I - \tau_n Q)^{-1}D = 0
\]

Result 2:
\[
K_\tau = e^{\tau Q}K_0 + (e^{\tau Q} - e^{-\tau/\tau_n I})U + \tau e^{\tau Q}V
\]
where
\[
U \equiv \tau_n RD(I - \tau_n Q^T)^{-1}
\]
\[
V \equiv [I - (I - \tau_n Q)R]D(I - \tau_n Q^T)^{-1}
\]
\[
R = (I + \tau_n Q)^{-1}
\]
or \[
R = P(I + \tau_n \Lambda)^+ P^{-1}
\] for singular \( I + \tau_n Q \) with diagonalizable \( Q = P\Lambda P^{-1} \)
Implications

A common practice infers a link between nodes $i$ and $j$ when the Pearson correlation coefficient of $x_i(t)$ and $x_j(t)$ is larger than some threshold value (“large correlation indicates interaction”)

Using Result 1, we can understand theoretically why (i) statistical covariance is not a good indicator for interaction and (ii) strongly interacting nodes can have weak correlations.

Weak pairwise correlations imply strongly correlated network states in a neural population
Elad Schneidman, Michael J. Berry, II, Ronen Segev & William Bialek Nature 440, 1007-1012 (20 April 2006)
Result 1:

\[ QK_0 + K_0 Q^T + D(I - \tau_n Q^T)^{-1} + (I - \tau_n Q)^{-1}D = 0 \]

• This result involves only the covariance matrix \( K_0 \). Since the left hand side is a symmetric matrix, there are only \( N(N+1)/2 \) independent equations.

• Hence \( K_0 \) alone contains insufficient information for the reconstruction of an asymmetric \( Q \) or a general directed network (different \( Q \) can give same \( K_0 \)).

Bidirectional Networks

Result 1 can be simplified when $Q$ is symmetric and noise is uniform and uncorrelated for different nodes:

$$\frac{\sigma^2}{2}K_0^{-1} = -Q + \tau_n Q^2$$

when $Q_{ij} = Q_{ji}$ and $D_{ij} = (\sigma^2 / 2) \delta_{ij}$

This is the case for bidirectional networks with $f_i = f$ and diffusive-like coupling $[h(x,y) = h(z), z = y - x$, and $h(-z) = -h(z), h'(0) > 0]$.

Network structure is given by the inverse of the covariance matrix and not the covariance matrix itself. This explains why strongly interacting nodes can have weak correlations.
Circles and lines represent nodes and links; size of circle is proportional to degree of the node

(a) Actual scale-free network
(b) Reconstructed using $K_0^{-1}$ (for small $\tau_n$)
(c) Reconstructed using Pearson correlation coefficient
\[ \Pi_{ij} \equiv \frac{(K_0)_{ij}}{\sqrt{(K_0)_{ii}(K_0)_{jj}}} \]
(d) Reconstructed using partial correlation coefficient
\[ \rho_{ij} \equiv -\frac{(K_0^{-1})_{ij}}{\sqrt{(K_0^{-1})_{ii}(K_0^{-1})_{jj}}} \]
Reconstruction of General Networks

For reconstruction of general networks, need relation between $K_\tau$ and $K_0$. Result 2 provides such a relation.

Result 2: $K_\tau = e^{\tau Q}K_0 + (e^{\tau Q} - e^{-\tau/\tau_n}I)U + \tau e^{\tau Q}V$ implies

$K_{2\tau} = SK_{\tau} - TK_0 \quad (1)$

$K_{3\tau} = SK_{2\tau} - TK_{\tau} \quad (2)$

where $S = e^{\tau Q} + e^{-\tau/\tau_n}I$ and $T = e^{-\tau/\tau_n}e^{\tau Q}$

Solving (1) and (2) we obtain $S$ and $T$ and thus $\mu = e^{\tau/\tau_n}$ and $Q = \log (\mu T)/\tau$

Reconstructing networks from dynamics with correlated noise
White-noise limit:

For $\tau_n \to 0, T \to 0$

thus the previous method cannot give $T$ accurately

But in this white-noise limit, we have

$$K_\tau = e^{\tau Q} K_0$$

and $Q = \log \left( (K_\tau K_0^{-1})/\tau \right)$
We first check whether the white-noise limit holds by calculating

\[ \Delta = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{H_{ii} - F_{ii}}{F_{ii}} \right)^2 \]

which measures the relative deviation of \( F = \log(K_\tau K_0^{-1}) \) and \( H = \log(K_{2\tau} K_0^{-1})/2 \)

If \( \Delta < \) some threshold value, then we take the white-noise approximation and define \( M = \log (K_\tau K_0^{-1})/\tau \)

otherwise we define \( M = \log (\mu T)/\tau \)

Thus \( M \approx Q \) and \( M_{ij} \approx g_{ij} A_{ij} h_y(X_i, X_j) \) for \( i \neq j \)

We identify the two groups of \( M_{ij} \) \((i \neq j)\) for \( A_{ij} = 0 \) or \( 1 \) using clustering method based on a Gaussian mixture model
We can further reconstruct the relative coupling strength when \( h(x,y) \) has additional properties.

\[
M_{ij} \approx g_{ij}A_{ij}h_y(X_i, X_j)
\]

- \( h_y(x, y) \) is a constant
  
  \[ G_{ij} \equiv \frac{g_{ij}}{\langle g \rangle} \]
  
  \[ \hat{G}_{ij} \equiv \frac{M_{ij}\hat{k}_{tot}}{\sum_{n,n\leftrightarrow l}|M_{nl}|} \]

- \( h_y(x, y) \) depends on \( y \) only
  
  \[ G_{j}^{\text{out}}(i) \equiv \frac{g_{ji}}{\langle g \rangle_{\text{out}}(i)} \]
  
  \[ \hat{G}_{j}^{\text{out}}(i) \equiv \frac{M_{ji}\hat{k}_{\text{out}}(i)}{\sum_{k\leftarrow i}|M_{ki}|} \]

- \( h_y(x, y) \) depends on \( x \) only
  
  \[ G_{j}^{\text{in}}(i) \equiv \frac{g_{ij}}{\langle g \rangle_{\text{in}}(i)} \]
  
  \[ \hat{G}_{j}^{\text{in}}(i) \equiv \frac{M_{ij}\hat{k}_{\text{in}}(i)}{\sum_{k\rightarrow i}|M_{ik}|} \]
We check our theoretical results using numerical data generated from different networks and different dynamics.

<table>
<thead>
<tr>
<th>Network</th>
<th>$N$</th>
<th>$N_B$</th>
<th>$N_U$</th>
<th>$N_L$</th>
<th>$\rho$</th>
<th>$g_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WSF</td>
<td>1000</td>
<td>4985</td>
<td>0</td>
<td>9970</td>
<td>0.00998</td>
<td>$P(g_{ij}) \sim g_{ij}^{-6.6}$</td>
</tr>
<tr>
<td>DWR</td>
<td>100</td>
<td>186</td>
<td>1678</td>
<td>2050</td>
<td>0.207</td>
<td>$N(10, 2)$</td>
</tr>
<tr>
<td>DWSF</td>
<td>1000</td>
<td>3120</td>
<td>3730</td>
<td>9970</td>
<td>0.00998</td>
<td>$P(g_{ij}) \sim g_{ij}^{-6.6}$</td>
</tr>
</tbody>
</table>

For the dynamics, we study nonlinear logistic function

$$f(x) = rx(1 - x)$$

and two coupling functions

$$h^{\text{diff}}(x, y) = y - x$$

$$h^{\text{syn}}(x, y) = \left(1/\beta_1\right)\left[1 + \tanh[\beta_2(y - y_0)]\right].$$
Investigation of general applicability of our method by studying additional cases for which our theoretical derivation does not directly apply

nonlinear FitzHugh–Nagumo (FHN) dynamics

\[
\dot{x}_i = \frac{\left(x_i - x_i^3/3 - y_i\right)}{\epsilon} + \sum_{j \neq i} g_{ij} A_{ij} (x_j - x_i) + \eta_i \\
\dot{y}_i = x_i + \alpha
\]

nonlinear Rössler dynamics

\[
\dot{x}_i = -y_i - z_i + \sum_{j \neq i} g_{ij} A_{ij} (x_j - x_i) + \eta_i \\
\dot{y}_i = x_i + ay_i + \sum_{j \neq i} g_{ij} A_{ij} (x_j - x_i) \\
\dot{z}_i = b + z_i (x_i - c) + \sum_{j \neq i} g_{ij} A_{ij} (x_j - x_i)
\]

We also study cases with node-dependent correlated noise \(\tau_n(i)\), taken from a uniform distribution with a certain range
\( FN/N_L \): ratio of missed links; \( FP/N_L \): ratio of incorrectly predicted links

\( N_L \) = total no of links

\( e_G \) = average % error of reconstructed relative coupling strength (excluding missed and incorrectly inferred links)

<table>
<thead>
<tr>
<th>Case</th>
<th>Network</th>
<th>Dynamics</th>
<th>( \tau_n )</th>
<th>( \Delta )</th>
<th>( \tilde{\tau}_n )</th>
<th>FN/N_L (%)</th>
<th>FP/N_L (%)</th>
<th>e_G (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} )</td>
<td>10(^{-3})</td>
<td>0.50</td>
<td>9.994 \times 10^{-4}</td>
<td>1.02</td>
<td>0.68</td>
<td>11.8</td>
</tr>
<tr>
<td>2</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} )</td>
<td>10(^{-2})</td>
<td>5.8</td>
<td>9.96 \times 10^{-3}</td>
<td>0</td>
<td>0</td>
<td>8.0</td>
</tr>
<tr>
<td>3</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} )</td>
<td>10(^{-1})</td>
<td>15</td>
<td>0.0989</td>
<td>0</td>
<td>0</td>
<td>6.4</td>
</tr>
<tr>
<td>4</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} )</td>
<td>1</td>
<td>23</td>
<td>0.906</td>
<td>0</td>
<td>0.15</td>
<td>6.8</td>
</tr>
<tr>
<td>5</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} )</td>
<td>5</td>
<td>2.9</td>
<td>3.07</td>
<td>0</td>
<td>0</td>
<td>6.4</td>
</tr>
<tr>
<td>6</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} )</td>
<td>10</td>
<td>[0.37]</td>
<td>[0.989]</td>
<td>0</td>
<td>[0.49]</td>
<td>[7.4]</td>
</tr>
<tr>
<td>7</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} ); non-diagonal D</td>
<td>10(^{-1})</td>
<td>240</td>
<td>4.09</td>
<td>0</td>
<td>0</td>
<td>6.4</td>
</tr>
<tr>
<td>8</td>
<td>DWR</td>
<td>( r = 10; h^{\text{syn}} )</td>
<td>10(^{-1})</td>
<td>620</td>
<td>0.0989</td>
<td>0</td>
<td>0.20</td>
<td>6.6</td>
</tr>
<tr>
<td>9</td>
<td>DWSF</td>
<td>( r = 100; h^{\text{diff}} )</td>
<td>10(^{-3})</td>
<td>19</td>
<td>0.0990</td>
<td>5.76</td>
<td>2.83</td>
<td>15.6</td>
</tr>
<tr>
<td>10</td>
<td>DWSF</td>
<td>( r = 100; h^{\text{diff}} )</td>
<td>10(^{-2})</td>
<td>0.50</td>
<td>9.86 \times 10^{-4}</td>
<td>4.78</td>
<td>4.47</td>
<td>12.3</td>
</tr>
<tr>
<td>11</td>
<td>DWSF</td>
<td>( r = 100; h^{\text{diff}} )</td>
<td>10(^{-1})</td>
<td>0.88</td>
<td>9.85 \times 10^{-3}</td>
<td>5.89</td>
<td>2.85</td>
<td>9.5</td>
</tr>
<tr>
<td>12</td>
<td>DWSF</td>
<td>( r = 100; h^{\text{diff}} )</td>
<td>1</td>
<td>1.1</td>
<td>0.0901</td>
<td>4.89</td>
<td>3.25</td>
<td>8.4</td>
</tr>
<tr>
<td>13</td>
<td>DWSF</td>
<td>( r = 100; h^{\text{syn}} )</td>
<td>10(^{-1})</td>
<td>1100</td>
<td>0.409</td>
<td>4.77</td>
<td>3.06</td>
<td>8.3</td>
</tr>
<tr>
<td>14</td>
<td>WSF</td>
<td>( r = 1; h^{\text{diff}} )</td>
<td>10(^{-2})</td>
<td>0.88</td>
<td>9.79 \times 10^{-3}</td>
<td>2.83</td>
<td>1.98</td>
<td>7.6</td>
</tr>
<tr>
<td>15</td>
<td>WSF</td>
<td>( r = 1; h^{\text{diff}} )</td>
<td>1</td>
<td>2200</td>
<td>0.399</td>
<td>2.38</td>
<td>1.48</td>
<td>6.1</td>
</tr>
<tr>
<td>Ia</td>
<td>DWR</td>
<td>Rössler</td>
<td>10(^{-1})</td>
<td>15</td>
<td>0.0987</td>
<td>0</td>
<td>0</td>
<td>6.3</td>
</tr>
<tr>
<td>Ib</td>
<td>DWR</td>
<td>FHN</td>
<td>10(^{-1})</td>
<td>86</td>
<td>0.0943</td>
<td>0</td>
<td>0</td>
<td>6.4</td>
</tr>
<tr>
<td>IIa</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} ); non-diagonal D</td>
<td>(( \tau_n )) = 0.0743</td>
<td>22</td>
<td>0.07</td>
<td>0</td>
<td>0.54</td>
<td>6.9</td>
</tr>
<tr>
<td>IIb</td>
<td>DWR</td>
<td>( r = 10; h^{\text{diff}} ); non-diagonal D</td>
<td>(( \tau_n )) = 0.2686</td>
<td>4.6</td>
<td>0.175</td>
<td>0.24</td>
<td>47.46</td>
<td>29.8</td>
</tr>
</tbody>
</table>

Our method gives accurate results with low error rates (including additional cases Ia, Ib, IIa, IIb)
Accurate reconstruction of relative coupling strength

Fig. 6. Comparison of reconstructed $\hat{G}_{ij}$ and $\hat{G}_{j}^{\text{out}(i)}$ with the actual values for (a) case 7; (b) case 8; (c) case 9 with $N_{\text{data}} = 10^7$; (d) case 13 with $N_{\text{data}} = 6 \times 10^6$; (e) case 14; (f) case la; (g) case lb and (h) case lla. Missed links (red) have $G_{ij} \neq 0$ or $G_{j}^{\text{out}(i)} \neq 0$ but $\hat{G}_{ij} = 0$ or $\hat{G}_{j}^{\text{out}(i)} = 0$ while incorrectly inferred links (blue) have $G_{ij} = 0$ or $G_{j}^{\text{out}(i)} = 0$ but $\hat{G}_{ij} \neq 0$ or $\hat{G}_{j}^{\text{out}(i)} \neq 0$. Dashed line is $y = x$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Accurate reconstruction of noise covariance matrix

Fig. 5. Comparison of the reconstructed $\hat{D}$ against the actual $D$ for (a) and (b) case 7 with constant $\tau_n = 10^{-1}$; (c) and (d) case Ila with $\langle \tau_n \rangle = 0.0743$ and $\delta \tau_n / \langle \tau_n \rangle = 34\%$ and (e) and (f) case IIb with $\langle \tau_n \rangle = 0.2686$ and $\delta \tau_n / \langle \tau_n \rangle = 84\%$. Dashed line is $y = x$. 
Effects of Hidden Nodes

Recall for bidirectional networks with diffusive-like coupling acted by a white noise, we have

\[
(\sigma^2 / 2)K_0^{-1} = -Q = h'(0)(L - aI); \quad a = f'(0) / h'(0)
\]

or \(K_0(L - aI) = \frac{\sigma^2}{2h'(0)}I; \quad L_{ij} = s_i \delta_{ij} - g_{ij} A_{ij}\)

Suppose the network has \(N\) nodes but only time-series measurements of \(n < N\) nodes are available. Denote the time-series of these \(n\) measured nodes by \(y_i(t), i = 1, \ldots, n\).

Can we still get the links among the measured nodes from \(y_i(t), i = 1, \ldots, n\), only?
\[ K(L - aI) = \frac{\sigma^2}{2h'(0)} I \quad (K = K_0) \]

\[ \Rightarrow \begin{pmatrix} K_m & U \\ U^T & K_h \end{pmatrix} \begin{pmatrix} L_m - aI_n & E \\ E^T & L_h - aI_{N-n} \end{pmatrix} = \frac{\sigma^2}{2h'(0)} \begin{pmatrix} I_n & 0 \\ 0 & I_{N-n} \end{pmatrix} \]

\[ \Rightarrow \frac{\sigma^2}{2h'(0)} K_m^{-1} = L_m - aI_n - E(L_h - aI_{N-n})^{-1} E^T \]

\[ \Rightarrow \begin{pmatrix} K_m^{-1} \end{pmatrix}_{ij} \propto g_{ij} A_{ij} + C_{ij} \quad i \neq j \quad \text{where } C = E(L_h - aI_{N-n})^{-1} E^T \]

\( C \) measures the effect of the hidden nodes on \( K_m^{-1} \);
\( C_{ij} \neq 0 \) only when there is a path connecting the measured nodes \( i \) and \( j \) via hidden nodes only
The magnitude of $C_{ij}$ depends on

- Number of paths connecting the measured nodes $i$ and $j$ via hidden nodes (number of nonzero terms in the sums)
- Strength of the hidden nodes
- Coupling strength of the links in these paths
Interesting observations:

• Hidden nodes with larger strength give rise to smaller corrections.
• When the three factors are comparable for the two groups of connected and unconnected nodes, information of $A_{ij}$ for the measured nodes can still be obtained even when a significant portion of nodes are hidden. Such situations arise when the hidden nodes are not preferentially linked to the measured nodes.

N=100, n=50

N=100, n=30
In a random network of $N=100$ nodes where the hidden nodes are randomly chosen:

\[ n_h = N - n \] is the number of hidden nodes.
In a network of $N=100$ nodes with $n_h=30$ hidden nodes that are preferentially linked to measured nodes that are unconnected:

$\frac{FN}{N_L} = 0.4\% \quad \frac{FP}{N_L} \approx 30\%$

The accuracy of the reconstruction deteriorates significantly but some useful information can still be uncovered.
Application to Real-World Networks (in-vitro neuronal networks)

Cortices from E17 Wistar rat embryos were plated on a multi-electrode array (MEA) Biocam 4096 system. Taken by Yu-Ting Huang of Dr. C.K. Chan’s group at Academia Sinica, Taiwan.

Each electrode receives input from a few nearby neurons.
Synaptic connections are directional.
Each electrode is taken as one node ---- 4095 nodes and apply our method to reconstruct the directed network from the voltage measurements.
Outgoing links are inferred to be either inhibitory or excitatory for each node $j$.

$M = \log[\mathbf{K}_\tau (\mathbf{K}_0)^{-1}] / \tau$

$M_{ij} \approx g_{ij} A_{ij} h_y(X_i, X_j)$ for $i \neq j$

(j=10) inhibitory node

(j=4080) excitatory node
Basic properties of the reconstructed neuronal networks

C. Elegans

<table>
<thead>
<tr>
<th>Days in vitro (DIV)</th>
<th>11</th>
<th>22</th>
<th>25</th>
<th>33</th>
<th>45</th>
<th>52</th>
<th>59</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>fraction of inhibitory nodes</td>
<td>0.27</td>
<td>0.13</td>
<td>0.15</td>
<td>0.18</td>
<td>0.21</td>
<td>0.24</td>
<td>0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>connection probability (%)</td>
<td>1.2</td>
<td>1.9</td>
<td>1.4</td>
<td>1.5</td>
<td>1.1</td>
<td>1.7</td>
<td>1.4</td>
<td>1.5</td>
</tr>
<tr>
<td>proportion of giant strongly connected component</td>
<td>0.88</td>
<td>0.96</td>
<td>0.98</td>
<td>0.92</td>
<td>0.81</td>
<td>0.90</td>
<td>0.77</td>
<td>0.83</td>
</tr>
<tr>
<td>average path length</td>
<td>4.0</td>
<td>3.7</td>
<td>3.7</td>
<td>3.9</td>
<td>4.1</td>
<td>3.7</td>
<td>4.0</td>
<td>3.7</td>
</tr>
<tr>
<td>clustering coefficient</td>
<td>0.26</td>
<td>0.36</td>
<td>0.38</td>
<td>0.30</td>
<td>0.25</td>
<td>0.28</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>small-world</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

measured fraction of inhibitory neurons ranges from 0.15-0.30 in various cortical regions in monkey [Hendry et al. J. Neuroscience 7, 1503-1519 (1987)]
Distribution of average outgoing coupling strength

average incoming/outgoing coupling strength $g_{\text{in}} / g_{\text{out}}$ defined by

$$g_{\text{in}}(i) = \sum_{j=1}^{N} g_{ij} / k_{\text{in}}(i)$$

$$g_{\text{out}}(i) = \sum_{j=1}^{N} g_{ji} / k_{\text{out}}(i)$$

The distribution of $g_{\text{out}}$ for excitatory nodes with $g_{\text{in}}(i) > 0$ is approximately lognormal.
• Lognormal distribution of synaptic connection strength [defined by the averaged peak excitatory postsynaptic potential (EPSP) amplitude] measured for layer 5 pyramidal neurons in the rat visual cortex using quadrupole whole-cell recordings technique.

The distribution of $\log |g_{\text{out}}|$ for excitatory nodes with $g_{\text{in}}(i) < 0$ and for inhibitory nodes is approximately exponential.
Relation between network structure and spike rate

We can predict whether a node has above-median spike rate or below-median spike rate using its average excitatory incoming coupling strength

$$g_{in}^+(i) = \frac{1}{k_{in}(i)} \sum_{j=1}^{N} g_{ij}$$

<table>
<thead>
<tr>
<th>DIV</th>
<th>11</th>
<th>22</th>
<th>25</th>
<th>33</th>
<th>45</th>
<th>52</th>
<th>59</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>sensitivity</td>
<td>0.83</td>
<td>0.75</td>
<td>0.74</td>
<td>0.79</td>
<td>0.81</td>
<td>0.80</td>
<td>0.70</td>
<td>0.78</td>
</tr>
<tr>
<td>specificity</td>
<td>0.80</td>
<td>0.70</td>
<td>0.72</td>
<td>0.81</td>
<td>0.80</td>
<td>0.78</td>
<td>0.73</td>
<td>0.80</td>
</tr>
</tbody>
</table>
Summary

• Using derived noise-induced relations between connectivity structure and dynamics for networks with stationary fluctuations around a noise-free steady-state, we have developed a method for reconstructing directed and weighted networks in the presence of white or correlated noise.

• Numerical tests show that our method gives good accurate results for networks with different nonlinear dynamics and coupling functions including networks that have non-steady noise-free state.

• We have shown that links among measured nodes of bidirectional networks can be reconstructed accurately when the hidden nodes are not preferentially linked with the measured nodes.

• We have applied our method to reconstruct in-vitro neuronal networks from experimental multi-electrode array measurements and obtained a number of interesting results (ongoing study).
Thank you!