2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

Reconstructing Networks from Dynamics

Emily S.C. Ching[1], Pik-Yin Lai[2,3], C.Y. Leung[1,4], and H.C. Tam[1]

[1]Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong
[2]Dept. of Physics and Center for Complex Systems, National Central University, Chungli, Taiwan 320, R.O.C.
[3]Physics Division, National Center for Theoretical Sciences, Kuang Fu Road 101, Hsinchu, Taiwan 300, R.O.C.
[4]Present address: School of Biology and School of Physics, Georgia Institute of Technology, Atlanta, GA, USA

Abstract—Many systems of interest in physics, biology or social science are complex networks. The links, their direction and weights or relative coupling strength of a network are important features that provide insights and fundamental understanding of the overall behavior and functionality of the network. Thus a method that can extract such information from measurements would be valuable. In this paper, we focus on weighted bidirectional networks, modeled by a dynamical system and subjected to a Gaussian white noise that mimics the effect of external disturbances. We show that general mathematical results relating the adjacency matrix of the network and the time-series measurements of the nodes can be obtained. Based on these results, we have developed a method that reconstructs both the links and their relative coupling strength using only the time-series measurements of node dynamics as input. We demonstrate that our method can give accurate results for unweighted and weighted random and scale-free networks with linear and nonlinear dynamics. We further show why relevance networks constructed using Pearson correlation coefficient and partial correlation coefficient can have significan deviations from the actual network.

1. Introduction

Many multi-component systems of interest in physics, biology or social science are complex networks with the components being the nodes or vertices and the interactions between components being the links or edges [1, 2, 3]. The links and their weights or relative coupling strength are important basic features of a network that provides insights and fundamental understanding of the overall behavior and functionality of the network. A vast amount of data has been measured for various networks of interest like gene regulatory network [4, 5] and brain network [6] but it remains a great challenge to reconstruct a network from measurements. All existing methods of network reconstruction have their limitations [7, 8]. In this paper, we focus on weighted bidirectional networks, modeled by a dynamical system and subjected to a Gaussian white noise that mimics the effect of external disturbances. We show that general mathematical results relating the adjacency matrix of the network and the time-series measurements of the nodes can be obtained. Based on these mathematical results, we have developed a method that reconstructs both the links and their relative coupling strength using only the time-series measurements of node dynamics as input. We demonstrate that our method can give accurate results for unweighted and weighted random and scale-free networks with linear and nonlinear dynamics. We further show why relevance networks constructed using Pearson correlation coefficient and partial correlation coefficient can have significan deviations from the actual network.

2. Our Method

We consider bidirectional weighted networks of N nodes, each with a variable $x_i(t)$, obeying the evolution equations

$$\dot{x}_i = f(x_i) + \sum_{j\neq i} g_{ij} A_{ij} h(x_i, x_j) + \eta_i \,, \qquad i = 1, 2, \ldots, N \,. \tag{1}$$

Here, an overdot denotes derivative with respect to time t, and f describes the intrinsic dynamics, which is taken to be identical for all nodes. The adjacency matrix element A_{ij} is 1 when node j is linked to node i by the coupling function $h(x_i, x_j)$ with strength g_{ij}; otherwise $A_{ij} = g_{ij} = 0$. The coupling is bidirectional such that $A_{ij} = A_{ji}$ and $g_{ij} = g_{ji}$. The effect of external disturbance is modeled by a Gaussian white noise η_i with zero mean and variance σ^2: $\overline{\eta_i(t)\eta_j(t')} = \sigma^2 \delta_{ij} \delta(t - t')$, where the overbar denotes an average over different realizations of the noise. We assume no self-loops such that $A_{ii} \equiv 0$. and focus on coupling function that satisfie

$$h(x, y) = h(z = y - x) \,; \;\; h(-z) = -h(z) \,; \;\; h'(0) > 0 \,. \tag{2}$$

With such a coupling function, the dynamics of the nodes tend to synchronize such that x_i's approach a stable fi ed point X_0 in the noise-free limit given by $f(X_0) = 0$ and $f'(X_0) < 0$. In the presence of weak noise, $\delta x_i = x_i - X_0$ is small and we have

$$\dot{\delta x_i} \approx -\sum_{j=1}^{N} \left[h'(0) \mathcal{L}_{ij} - f'(X_0) \delta_{ij} \right] \delta x_j + \eta_i \,, \tag{3}$$

where $\mathcal{L}$ is the Laplacian matrix of a weighted network given by

$$\mathcal{L}_{ij} = s_i - g_{ij} A_{ij}, \qquad s_i \equiv \sum_{j=1}^{N} g_{ij} A_{ij} \tag{4}$$

and s_i is the strength of node i.

We defin the dynamical covariance matrix $\tilde{C}$ by

$$\tilde{C}_{ij}(t) \equiv \overline{[x_i(t) - X(t)][x_j(t) - X(t)]} \quad (5)$$

where $X(t) \equiv (1/N)\sum_{i=1}^{N} x_i(t)$. It is easy to see that $\tilde{C}_{ij}(t)$ is equal to $\overline{[\delta x_i(t) - \delta X(t)][\delta x_j(t) - \delta X(t)]}$ with $\delta X(t) \equiv (1/N)\sum_{i=1}^{N} \delta x_i(t)$. Using Eq. (3), we have derived [9, 10] an exact relation between $\tilde{C}^+$ and $\mathcal{L}$ for the linearized system:

$$\frac{\sigma^2}{2} \lim_{t\to\infty} \tilde{C}^+ = h'(0)\mathcal{L} - f'(X_0)\left(I - \frac{1}{N}J\right), \quad (6)$$

which is an approximation for the original system. Here, the superscript $^+$ denotes the pseudoinverse of a matrix and J is an $N \times N$ matrix whose elements are all equal to one. Based on Eq. (6), we are able to develop [9, 10] a method that reconstructs A_{ij} and the normalized coupling strength $G_{ij} \equiv g_{ij}/\langle g\rangle$, where $\langle g\rangle \equiv \sum_{i,j} g_{ij}A_{ij}/\sum_{i,j} A_{ij}$ is the average coupling strength. For stationary time-series measurements, we approximate the ensemble average over noise by a long-time average:

$$\lim_{t\to\infty} \tilde{C}_{ij}(t) \approx C_{ij}(T) \equiv \langle [x_i(t) - X(t)][x_j(t) - X(t)]\rangle_T , \quad (7)$$

where $\langle \ldots \rangle_T$ denotes an average over a time interval T of the measurements. We emphasize that $C_{ij}(T)$ can be calculated using only the time-series measurements $x_i(t)$, $i = 1, \ldots, N$. Using Eq. (6), we then obtain

$$r_{ij} \equiv \frac{C^+_{ij}}{C^+_{ii}} \approx \begin{cases} \dfrac{-g_{ij}}{s_i - f'(X_0)/h'(0)}, & A_{ij} = 1 \\ 0, & A_{ij} = 0 \end{cases}, \quad i \neq j \quad (8)$$

where we have neglected the $1/N$ term for large networks. For each node i, the values of r_{ij} form two groups, one for nodes j that are unconnected to node i, and the other for nodes j connected to node i. By identifying these two groups [10], we obtain the reconstructed $A^{(e)}_{ij}$. Moreover, Eq. (6) together with Eq. (7) imply that for $i \neq j$, $-(\sigma^2/2)C^+_{ij} \approx h'(0)g_{ij}A_{ij}$, which further implies

$$G_{ij} \approx \frac{C^+_{ij}\sum_l k^{(e)}_l}{\sum_{n,l\leftrightarrow n} C^+_{nl}} \equiv G^{(e)}_{ij} , \quad (9)$$

where $k^{(e)}_i = \sum_{j=1}^{N} A^{(e)}_{ij}$ is the reconstructed degree of node i, and $\sum_{l\leftrightarrow n}$ represents a sum over nodes l that are reconstructed to be linked to node n. Equation (9) gives the reconstructed relative coupling strength $G^{(e)}_{ij}$.

The accuracy of the reconstructed adjacency matrix $A^{(e)}_{ij}$ can be measured by P_{SEN}, the percentage of correctly reconstructed links, and P_{SPEC}, the percentage of correctly reconstructed non-existent links. P_{SEN} and P_{SPEC} give respectively the sensitivity and specificit of the method in percentage. We present some of the results for unweighted and weighted random networks of $N = 100$ nodes and unweighted and weighted scale-free networks of $N = 1000$ nodes in Table 1. The weighted random (WR) networks used are random networks with g_{ij} taken from a Gaussian distribution of mean μ and standard deviation γ, and the weighted scale-free network used is the one extended [11] from the unweighted Barabasi-Albert scale-free network [12] and has a power-law distribution in g_{ij}. We use $\sigma = 1$ and $T \leq 5000$ for all the results presented in this paper. As shown in Table 1, our method gives accurate reconstruction with both P_{SEN} and P_{SPEC} larger than 85%. More importantly, we note that the applicability of our method goes beyond networks described by Eqs. (1) and (2); it is also applicable for networks with vector variables $\vec{x}_i(t)$ obeying higher-dimensional dynamics. In particular, for networks with $\vec{x}_i(t) = (x_i(t), y_i(t))$ obeying FitzHugh-Nagumo (FHN) dynamics, which is a good description for neurons [13]:

$$\dot{x}_i = (x_i - x_i^3/3 - y_i)/\epsilon + \sum_{j\neq i} g_{ij}A_{ij}(x_j - x_i) + \eta_i \quad (10)$$

$$\dot{y}_i = x_i + \alpha \quad (11)$$

with $\alpha = 1.05$ and $\epsilon = 0.01$, our method gives accurate reconstruction using only $x_i(t)$. In Figs. 1 and 2, we show respectively that our method can capture well the power-law distribution of the degree k_i of the unweighted scale-free network as well as the power-law distribution of the relative coupling strength G_{ij} of the weighted scale-free network.

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

Table 1: Accuracy of our method measured by P_{SEN} and P_{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.

3. Relevance networks constructed using statistical correlation

Using the time-series measurements, one can calculate the covariance matrix $\Sigma(T)$, define by

$$\Sigma_{ij}(T) = \langle [x_i(t) - \langle x_i(t)\rangle_T][x_j(t) - \langle x_j(t)\rangle_T]\rangle_T \quad (12)$$

and obtain the Pearson correlation coefficient

$$\Pi_{ij} \equiv \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}}\sqrt{\Sigma_{jj}}} \quad (13)$$

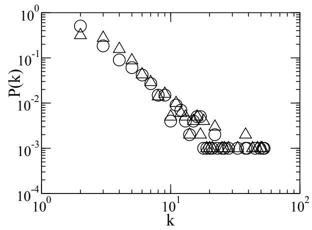

Figure 1: Comparison of the degree distribution $P(k)$ calculated using reconstructed $k_i^{(e)}$ (triangles) against the one calculated using the actual k_i (circles) for the unweighted scale-free network with FHN dynamics.

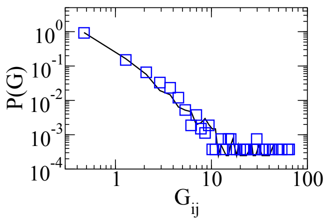

Figure 2: Comparison of $P(G)$ of the reconstructed $G_{ij}^{(e)}$ (squares) against the actual distribution (solid line) for weighted scale-free network with FHN dynamics.

of nodes i and j, which measures the linear correlation of the measurements $x_i(t)$ and $x_j(t)$. If $x_i(t)$ and $x_j(t)$ are uncorrelated, $\Pi_{ij} = 0$ whereas if $x_i(t)$ and $x_j(t)$ are linearly related, $|\Pi_{ij}| = 1$. It is commonly expected that nodes linked to one another would have correlated dynamics. Thus one common way to construct a relevance network for the network of interest is to infer a link between nodes i and j if Π_{ij} is larger than some threshold.

Statistical correlation between $x_i(t)$ and $x_j(t)$ can arise not only from direct interaction of nodes i and j but also from indirect interaction of the two nodes via other nodes. Thus an improved way to construct a relevance network is to use the partial correlation, which measures the linear correlation between $x_i(t)$ and $x_j(t)$ with the effect from $x_k(t)$, $k \neq i, j$, of the other nodes eliminated. Specificall , the partial correlation coefficient ρ_{ij} of nodes i and j is given in terms of Σ^{-1}:

$$\rho_{ij} = -\frac{\Sigma^{-1}_{ij}}{\sqrt{\Sigma^{-1}_{ii}}\sqrt{\Sigma^{-1}_{jj}}} \qquad (14)$$

Thus another common way to construct a relevance network is to infer a link between nodes i and j when ρ_{ij} is larger than some threshold.

In the following, we shall show that for networks described by Eqs. (1) and (2), the relevance networks constructed using either Π_{ij} or ρ_{ij} can have significan deviations from the actual network. Specificall , we defin

$$\tilde{\Sigma}_{ij}(t) \equiv \overline{[x_i(t) - \overline{x_i(t)}][x_j(t) - \overline{x_j(t)}]} \, . \qquad (15)$$

	Method 1		Method 2	
Case	P_{SEN}	P_{SPEC}	P_{SEN}	P_{SPEC}
1	39.0	82.9	99.8	99.9
2	49.8	86.1	100	100
3	8.97	99.6	91.1	99.97
4	34.6	99.7	88.8	99.95

Table 2: Comparison of the relevance network constructed using Pearson correlation coefficient (method 1) and partial correlation coefficient (method 2) with the actual (unweighted) network for cases 1 - 4 reported in Table 1.

Then we can derive an exact result similar to Eq. (6):

$$\lim_{t\to\infty} \tilde{\Sigma}^{-1}(t) = \frac{2}{\sigma^2}\left[h'(0)\mathcal{L} - f'(X_0)I\right] \, . \qquad (16)$$

For stationary time-series measurements, we again approximate an ensemble average over noise by a long-time average and approximate $\lim_{t\to\infty} \tilde{\Sigma}^{-1}(t)$ by $\Sigma^{-1}(T)$. Thus using Eq. (16), we obtain

$$\Sigma^{-1}_{ij} \approx -\frac{2h'(0)}{\sigma^2} g_{ij} A_{ij} \, , \qquad i \neq j \qquad (17)$$

$$\Sigma^{-1}_{ii} \approx \frac{2h'(0)}{\sigma^2} s_i - f'(X_0) \, . \qquad (18)$$

Equation (17) clearly shows that the adjacency matrix A_{ij}, giving network connectivity, is proportional to the *inverse* of the covariance matrix Σ^{-1}_{ij} thus relevance networks constructed using the Pearson correlation coefficient Π_{ij} are doomed to have little resemblance to the actual network. Since the denominator of ρ_{ij} depends on both Σ^{-1}_{ii} and Σ^{-1}_{jj}, ρ_{ij} would deviate from a simple multiple of Σ^{-1}_{ij} for networks whose nodes strength s_i varies a lot from node to node and, as a result, the relevance networks constructed using ρ_{ij} for such networks would have significan deviation from the actual networks. For unweighted networks, $s_i = k_i$ and thus scale-free networks with a power-law distribution in the degree k_i would be such an example, and we expect relevance network constructed using ρ_{ij} would deviate much from the actual scale-free network. In Table 2, we compare the relevance networks constructed using Π_{ij} and ρ_{ij} with the actual networks. We choose the threshold value such that the difference between the total degree of the relevance network and the total degree of the actual network is the smallest. Indeed, the relevance networks constructed using Π_{ij} miss many links in the actual networks and give very low values of P_{SEN}. In Fig. 3, we show the actual scale-free network, the relevance networks constructed by Π_{ij} and ρ_{ij}, and the network reconstructed using our method. Both the relevance networks constructed using Π_{ij} and ρ_{ij} deviate significantl from the actual network. On the other hand, the network reconstructed using our present method resembles the actual network very well.

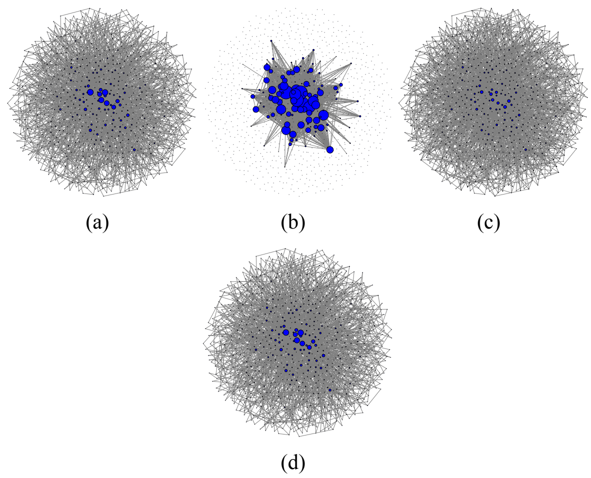

Figure 3: Comparison of (a) network reconstructed by our method using C^+, relevance networks constructed using (b) Π_{ij} and (c) using ρ_{ij} with (d) actual (unweighted) scale-free network of consensus dynamics. The circles and the lines represent the nodes and the links of the network. The size of the circle is proportional to the degree of the node.

4. Summary

In summary, we have developed a method that reconstructs both the links and their relative coupling strength for weighted bidirectional networks. The method is based on general mathematical result [Eq. (6)] derived for networks described by Eqs. (1) and (2) but its applicability has been shown to go beyond such networks and particularly to networks with vector variables obeying higher-dimensional dynamics such as the two-dimensional FitzHugh Nugamo dynamics. Using unweighted and weighted random and scale-free networks with various dynamics, we have demonstrated that our method can give accurate reconstruction that captures well the network topology, the degree distribution as well as the relative coupling strength distribution. Finally, using the general mathematical results obtained, we further show why relevance network constructed using Pearson correlation coefficient is doomed to have poor resemblance of the actual network and relevance network constructed using the partial correlation coefficient can have significan deviation from the actual network whose nodes have nonuniform node strength.

Acknowledgments

The work of ESCC, CYL and HCT has been supported by Hong Kong Research Grants Council under grant no. CUHK 14300914 while that of PYL by Ministry of Science and Technology of Taiwan under the grant no. 101-2112-M-008-004-MY3.

References

[1] S.H. Strogatz, "Exploring complex networks", Nature (London) **410**, 268 (2001).

[2] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks", Rev. Mod. Phys. **74**, 48 (2002).

[3] M.E.J. Newman, "The Structure and Function of Complex Networks", SIAM Rev. **45**, 167 (2003).

[4] F. Emmert-Streib, G.V. Glazko, G. Altay, and R.de Matos Simoes, "Statistical inference and reverse engineering of gene regulatory networks from observational expression data", Frontiers of Genetics **3**, 1 (2012).

[5] G. Karlebach and R. Shamir, "Modelling and analysis of gene regulatory networks", Nat. Rev. Mol. Cell Biol. **9**, 771 (2008).

[6] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems", Nat. Rev. Neurosci. **10**, 186 (2009).

[7] R. De Smet and K. Marchal, "Advantages and limitations of current network inference methods", Nat. Rev. Microbiol. **8**, 717 (2010).

[8] D. Marbach et. al, "Wisdom of crowds for robust gene network inference", Nature Methods **9**, 796 (2012).

[9] E.S.C. Ching, P.Y. Lai, and C.Y. Leung, "Extracting connectivity from dynamics of networks with uniform bidirectional coupling", Phys. Rev. E **88**, 042817 (2013); Erratum, Phys. Rev. E **89**, 029901(E) (2014).

[10] E.S.C. Ching, P.Y. Lai, and C.Y. Leung, "Reconstructing weighted networks from dynamics", Phys. Rev. E **91**, 030801(R) (2015).

[11] A. Barrat, M. Barthélemy, and A. Vespignani, "Weighted Evolving Networks: Coupling Topology and Weight Dynamics", Phys. Rev. Lett. **92**, 228701 (2004).

[12] A.-L. Barabási, R. Albert, "Emergence of scaling in random networks", Science **286**, 509 (1999).

[13] R. FitzHugh, "Impulses and Physiological States in Theoretical Models of Nerve Membrane", Biophys. J. **1**, 445 (1961).