Reconstructing weighted networks from dynamics

Emily S. C. Ching,1,2 Pik-Yin Lai,2,3,4 and C. Y. Leung1,4

1Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong
2Department of Physics and Center for Complex Systems, National Central University, Chungli, Taiwan 320, Republic of China
3Physics Division, National Center for Theoretical Sciences, Kuang Fu Road 101, Hsinchu, Taiwan 300, Republic of China

(Received 12 June 2014; revised manuscript received 6 February 2015; published 24 March 2015)

We present a method that reconstructs both the links and their relative coupling strength of bidirectional weighted networks. Our method requires only measurements of nodal dynamics as input. Using several examples, we demonstrate that our method can give accurate results for weighted random and weighted scale-free networks with both linear and nonlinear dynamics.

DOI: 10.1103/PhysRevE.91.030801 PACS number(s): 89.75.Hc, 05.45.Tp, 05.45.Xt

Many multicomponent 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 [1–3]. 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 such as gene regulatory [4,5] and brain networks [6] but it remains a challenge to reconstruct a network, i.e., finding the links and their relative coupling strength, from these measurements. All existing network reconstruction methods have their limitations [7,8]. Many existing methods are statistical in nature, inferring the links from statistical correlations [9] and dependence [10] of the measurements of the nodes or employing Bayesian graphical models [4,11,12] to find the network that best matches the measured statistics. Statistical correlation and dependence, however, does not necessarily follow from direct connections, and the matching problem is underdetermined as the number of possible answers far exceeds the number of available measurements. It has been suggested [13] that the topology of a network controls its dynamics and thus information about the network connectivity can be uncovered from their dynamics. We have recently developed [14] a method that reconstructs the links of bidirectional networks with uniform strength of interaction. Our method uses only measurements of nodal dynamics as input and is shown to give accurate results for various networks with both linear and nonlinear dynamics. Earlier methods [15] either assume linearity [16,17] or require additional information [18–22]. In realistic networks, the strength of interaction often differs for different pairs of nodes, and these networks are weighted [23]. Reconstruction of weighted networks is even more challenging.

In this Rapid Communication, we present a method that reconstructs both the links and their weights of bidirectional weighted networks. Our method requires only measurements of nodal dynamics as input. Using various examples, we demonstrate that our method gives accurate results for weighted random and weighted scale-free networks with both linear and nonlinear dynamics.

We consider bidirectional weighted networks of N nodes, each with a variable \( x_i(t), i = 1, 2, \ldots, N \), and

\[
\dot{x}_i = f(x_i) + \sum_{j \neq i} g_{ij} A_{ij} h(x_i, x_j) + \eta_i.
\] (1)

Here the 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 \( g_{ij} = g_{ji} \). We assume that the graphs of the networks have one connected part and no self-loops such that \( A_{ii} = 0 \). We model external disturbances by a Gaussian white noise with zero mean and variance \( \sigma^2: \langle \eta_i(t) \eta_j(T) \rangle = \sigma^2 \delta_{ij} \delta(t - T) \), where the overbar is an average over different realizations of the noise. Our goal is to reconstruct \( A_{ij} \) and \( g_{ij} \) using only \( x_i(t) \).

We define the matrix \( M \) by

\[
M_{ij} = \frac{s_i}{\langle g \rangle} \delta_{ij} - \frac{g_{ij} A_{ij}}{\langle g \rangle}, \quad s_i \equiv \sum_{j=1}^{N} g_{ij} A_{ij},
\] (2)

where \( \langle g \rangle \equiv \sum_{i,j} g_{ij} A_{ij} / \sum_{i,j} A_{ij} \) is the average coupling strength and \( s_i \) is the strength of node \( i \). \( M \) is the normalized Laplacian matrix of a weighted network and contains all the information of \( A_{ij} \) and \( g_{ij} \). Using \( x_i(t) \), we calculate the dynamical covariance matrix \( C \) with

\[
C_{ij} = \langle [x_i(t) - X(t)][x_j(t) - X(t)]^T \rangle_T,
\] (3)

where \( X(t) \equiv (1/N) \sum_{i=1}^{N} x_i(t) \) and \( \langle \cdots \rangle_T \) is an average over a time interval \( T \) of the measurements. We first show an approximate relation between the pseudoinverse of \( C \), denoted as \( C^+ \), and \( M \) for networks with positive semidefinite \( M \) and \( h(x,y) \) satisfying

\[
h(x,y) = h(z), \quad h(-z) = -h(z), \quad h'(0) > 0,
\] (4)

where \( z = y - x \). With such a coupling function, the dynamics of the nodes tend to synchronize such that \( x_i \)’s approach a stable fixed 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
\[ \delta x_i \approx -(g_j) h'(0) \sum_{j=1}^{N} \left( M_{ij} - \frac{f'(X_0)}{(g_j) h'(0)} \delta j \right) \delta x_j + \eta_i. \]  
(5)

Following the ideas of the proof given in [14], we derive an exact result between \( C^+ \) and \( M \) in the limit \( T \to \infty \) for the linearized system (5), which is an approximation for the system (1) with (4):
\[ \sigma^2 \left( \frac{g_j}{2(g_j) h'(0)} \right)^{T \to \infty} \frac{\delta C^+_{ij}}{\delta t} \approx M_{ij} - f' (X_0) \left( \frac{\delta j}{g_j} \right) M_{ij}, \]  
(6)
with the factor \( \delta j = 1/N \) coming from \( \sum_i M_{ij} M_{ji} \). For consensus dynamics [24], defined by \( f = 0 \) and \( h(z) = z \), relation (6) becomes \( \sigma^2 / (2(g)) \lim_{T \to \infty} C^+_{ij} = M_{ij} \) and is exact [25].

Guided by relation (6), we devise a method to reconstruct \( A_{ij} \) and \( g_{ij} \). Approximating \( \lim_{T \to \infty} \frac{\delta C^+_{ij}}{\delta t} \) by \( C^+_{ij} \) with a finite \( T = T_\infty \) and using Eq. (2), we get
\[ r_{ij} \equiv \frac{C^+_{ij}}{C_{ii}} \approx \left\{ \begin{array}{ll} \frac{-g_{ij}}{s_i - f(X_0)/h(0)}, & A_{ij} = 1 \\ 0, & A_{ij} = 0 \end{array} \right. \]  
(7)
for \( i \neq j \), where we have neglected the \( 1/T \) term for large networks. For each node \( i \), \( r_{ij} \) for nodes \( j \) connected to \( i \) would have gaps among themselves because of different \( g_{ij} \) and also have a gap from those unconnected to node \( i \). By identifying the latter gap [26], we obtain the reconstructed \( A_{ij}^{(c)} \). Moreover, Eq. (6) implies
\[ -\left( \frac{\sigma^2}{2} \right) C^+_{ij} \approx h'(0) g_{ij}, \quad i \neq j, \]  
(8)
which further implies
\[ G_{ij} \equiv \frac{g_{ij}}{(g)} \approx \frac{C^+_{ij} + \sum_{l=1+n}^{k_{ij}} C^{(c)}_{il} \equiv G_{ij}^{(c)}}{\sum_{l=1+n}^{k_{ij}} C^{(c)}_{il}} \]  
(9)
and
\[ S_i \equiv \frac{s_i}{\langle k \rangle} = \left( \frac{1}{N} \right) \sum_{j=1}^{N} A_{ij} \approx \frac{N \sum_{j=1}^{N} C^+_{ij} - \sum_{l=1+n}^{k_{ij}} C^{(c)}_{il}}{\sum_{l=1+n}^{k_{ij}} C^{(c)}_{il}} \equiv S_i^{(c)}. \]  
(10)
where \( \langle k \rangle \equiv \sum_i k_i/N, k_i = \sum_{j=1}^{N} A_{ij} \) is the degree of node \( i \), and \( \langle k \rangle \equiv \sum_i k_i/N, k_i^{(c)} = \sum_{j=1}^{N} A_{ij}^{(c)} \) is the reconstructed degree of node \( i \), and \( \sum_{l=1+n}^{k_{ij}} \) represents a sum over nodes \( l \) that are reconstructed to be connected to node \( n \). We use Eqs. (9) and (10) to obtain the reconstructed relative coupling strength \( G_{ij}^{(c)} \) and relative strength of nodes \( S_i^{(c)} \).

We apply our method to two weighted random (WR) networks of \( N = 100 \) with a connection probability of 0.2 and two different weighted scale-free (WSF) networks of \( N = 1000 \). The two WR networks have \( g_{ij} \) taken from two Gaussian distributions of different mean \( \mu \) and standard deviation \( \gamma: \mu = 5 \) and \( \gamma = 2 \) (WR1) and \( \mu = 10 \) and \( \gamma = 2 \) (WR2). The WSF networks are two different extensions [27,28] of the (unweighted) Barabasi-Albert SF network [29], and are generated by starting with \( n = 5 \) nodes and adding a new node to \( m = 5 \) different existing nodes at each step. The probability of connecting to an existing node \( i \) is either proportional to \( k_i \) [27] (WSF1) or \( s_i \) [28] (WSF2). For WSF1, the coupling strength of a new link is \( g_{ij} = g_0 k_i / \sum_i k_i \), where \( g_0 = 10 \) and \( \sum_i \) is over the existing nodes \( i \) that the new node \( j \) is connected to. For WSF2, the new link changes the coupling strength \( g_{ij} \) of all the other links of the existing node \( i \) to \( g_{ij}(1 + \delta/s_i) \) with \( \delta = 2 \). The degree distribution \( P(k) \) is a power law with an exponent \(-3 \) (WSF1) or \(-(4 + 3)/(28 + 1) \) (WSF2). The relative coupling strength distribution \( P(g) \) has a peak and is skewed for WSF1 but is a power law for WSF2.

To test the general applicability of the method, we study not only the linear consensus dynamics and nonlinear intrinsic logistic dynamics \( f(x) = rx(1 - x) \) with \( r = 10 \) and \( h(z) = z \) (logistic) for which (6) is expected to hold, but also nonlinear coupling function that cannot be linearized: \( f = 0 \) and \( h(z) = z^3 \) (cubic), and higher-dimensional dynamics with different coupling. We study the FitzHugh-Nagumo (FHN) dynamics [30]:
\[ \dot{x}_i = (x_i - x_i^3/3 - y_i)/\epsilon + \sum_{j \neq i} g_{ij} A_{ij} h(x_i, x_j) + \eta_i, \]  
(11)
\[ \dot{y}_i = x_i + \alpha \]  
(12)
with both \( h(x, y) = h(z) = z \) and \( \alpha = 1.05 \) (FHN1) or \( \alpha = 0.95 \) (FHN2) and synapticlike coupling \( h(x, y) = (1 + tanh[\beta(\gamma - x_0)])/2 \) with \( \beta = 2 \) and \( x_0 = 1 \) and \( \alpha = 1.05 \) (FHN3), together with \( \epsilon = 0.01 \), as well as the Rössler dynamics [31]:
\[ \dot{x}_i = -y_i - z_i + \sum_{j \neq i} g_{ij} A_{ij} \tanh([x_j - x_i]) + \eta_i, \]  
(13)
\[ \dot{y}_i = x_i + ay_i + \sum_{j \neq i} g_{ij} A_{ij} \tanh(y_j - y_i), \]  
(14)
\[ \dot{z}_i = b + z_i(x_i - c) + \sum_{j \neq i} g_{ij} A_{ij} \tanh(z_j - z_i). \]  
(15)

<table>
<thead>
<tr>
<th>Network</th>
<th>Dynamics</th>
<th>( T_\infty )</th>
<th>( P_{\text{SEN}} ) (%)</th>
<th>( P_{\text{SPEC}} ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WR1</td>
<td>Consensus</td>
<td>1000</td>
<td>94.3</td>
<td>100.0</td>
</tr>
<tr>
<td>WR1</td>
<td>Rössler</td>
<td>1000</td>
<td>94.2</td>
<td>100.0</td>
</tr>
<tr>
<td>WR1</td>
<td>FHN3</td>
<td>1000</td>
<td>26.9</td>
<td>99.57</td>
</tr>
<tr>
<td>WR1</td>
<td>FHN3</td>
<td>5000</td>
<td>52.6</td>
<td>99.70</td>
</tr>
<tr>
<td>WR2</td>
<td>Logistc</td>
<td>1000</td>
<td>99.9</td>
<td>100.0</td>
</tr>
<tr>
<td>WR2</td>
<td>Cubic</td>
<td>1000</td>
<td>96.8</td>
<td>99.70</td>
</tr>
<tr>
<td>WR2</td>
<td>FHN1</td>
<td>1000</td>
<td>99.5</td>
<td>100.0</td>
</tr>
<tr>
<td>WR2</td>
<td>FHN2</td>
<td>1000</td>
<td>98.1</td>
<td>99.97</td>
</tr>
<tr>
<td>WSF1</td>
<td>Consensus</td>
<td>1000</td>
<td>88.3</td>
<td>99.98</td>
</tr>
<tr>
<td>WSF1</td>
<td>Rössler</td>
<td>1000</td>
<td>86.0</td>
<td>99.99</td>
</tr>
<tr>
<td>WSF1</td>
<td>FHN1</td>
<td>1000</td>
<td>56.8</td>
<td>99.83</td>
</tr>
<tr>
<td>WSF1</td>
<td>FHN1</td>
<td>5000</td>
<td>85.5</td>
<td>99.87</td>
</tr>
<tr>
<td>WSF1</td>
<td>FHN2</td>
<td>1000</td>
<td>61.4</td>
<td>99.80</td>
</tr>
<tr>
<td>WSF1</td>
<td>FHN2</td>
<td>5000</td>
<td>85.7</td>
<td>99.86</td>
</tr>
<tr>
<td>WSF2</td>
<td>Consensus</td>
<td>1000</td>
<td>91.7</td>
<td>99.98</td>
</tr>
<tr>
<td>WSF2</td>
<td>Rössler</td>
<td>1000</td>
<td>85.2</td>
<td>99.98</td>
</tr>
<tr>
<td>WSF2</td>
<td>FHN1</td>
<td>5000</td>
<td>34.9</td>
<td>99.69</td>
</tr>
<tr>
<td>WSF2</td>
<td>FHN1</td>
<td>5000</td>
<td>86.7</td>
<td>99.98</td>
</tr>
<tr>
<td>WSF2</td>
<td>FHN2</td>
<td>1000</td>
<td>37.5</td>
<td>99.73</td>
</tr>
<tr>
<td>WSF2</td>
<td>FHN2</td>
<td>5000</td>
<td>83.3</td>
<td>99.98</td>
</tr>
</tbody>
</table>
The parameters are \( a = b = 0.2 \) and \( c = 9 \), chosen such that the Rössler dynamics is chaotic without the coupling and noise. We take \( \sigma = 1 \) in all the cases studied and numerically integrate the equations of motion using either the Euler-Maruyama method or the weak second-order Runge-Kutta method [32]. For the FHN and Rössler dynamics, we reconstruct the networks using \( x_i(t) \) only. The sampling interval of \( x_i(t) \) is \( 5 \times 10^{-4} \) for all the cases.

We measure the accuracy of \( A_j^{(e)} \) by the percentages of correctly reconstructed links and nonexistent links, denoted by \( P_{\text{SEN}} \) and \( P_{\text{SPEC}} \), respectively (Table I). Using ten different realizations of the WR1 network, we check that the standard deviation of \( P_{\text{SEN}} \) and \( P_{\text{SPEC}} \) is less than 2%. As our identification of the gaps [26] tends to be more stringent, \( P_{\text{SEN}} \) is generally smaller than \( P_{\text{SPEC}} \). There are two general results: (i) the reconstruction is more accurate for the WR than the WSN networks with the same dynamics, and (ii) the accuracy is increased when \( T_w \) is increased. Our method gives rather accurate results: except for WR1 with FHN3 dynamics, both \( P_{\text{SEN}} \) and \( P_{\text{SPEC}} > 80% \) when \( T_w \) is at most 5000. In the WSN networks, small-degree nodes connected to large-degree nodes have a large disparity in \( g_{ij} \). The presence of links with dominantly large coupling strength makes it easier to miss the other links with moderate coupling strength in our method [26] and thus reduces \( P_{\text{SEN}} \). This explains the general result (i). Moreover, the number of errors for individual nodes increases as \( (\min_j g_{ij})/s_j \) decreases as expected from Eq. (7).

Next we study the dependence of \( -(\sigma^2/2)C_{ij}^{(e)} \) on \( g_{ij} \) for \( i \neq j \) (see Fig. 1). We find that the data points scatter around the line \( y = x \), confirming Eq. (8), not only for consensus and logistic dynamics with \( h(\theta) = 1 \) as expected but also for higher-dimensional Rössler dynamics for WR1 and WSNF1 and FHN1 dynamics for WR2. Our theoretical result, Eq. (6), for \( i = j \) is also confirmed for WR2 with logistic dynamics [see inset of Fig. 1(b)]. In the other cases, although Eq. (8) does not hold, \( -(\sigma^2/2)C_{ij}^{(e)} \) is approximately proportional to \( g_{ij} \) albeit with a much larger data scatter, particularly for \( g_{ij} = 0 \), except for the small number of very large \( g_{ij} \) for WSF2 with Rössler, FHN1, and FHN2 dynamics [see Fig. 1(d)]. These interesting observations suggest the approximate result

\[
-(\sigma^2/2)C_{ij}^{(e)} \approx B g_{ij}, \quad i \neq j \tag{16}
\]

general dynamics with some \( B > 0 \). When the data scatter for \( g_{ij} = 0 \) is large, \( C_{ij}^{(e)} \) for \( g_{ij} = 0 \) would overlap significantly with those for \( g_{ij} > 0 \), resulting in a large number of incorrectly reconstructed links and thus a low \( P_{\text{SEN}} \). This explains the low \( P_{\text{SEN}} \) for WR1 with FHN3 dynamics for which \( B \approx 0.3 \). The data scatter is decreased when \( T_w \) is increased [see Fig. 1(f)] leading to the general result (ii). Note that Eqs. (9) and (10) remain valid when Eq. (16) holds. We compare \( G_{ij}^{(e)} \) with the actual \( G_{ij} \). Except for the small number of very large \( G_{ij} \) for WSF2 with Rössler, FHN1, and FHN2 dynamics, \( G_{ij}^{(e)} \approx DG_{ij} \) (see Fig. 2) but \( G_{ij}^{(e)} \) can deviate from \( G_{ij} \) due to inaccurate \( k_i^{(e)} \) when \( P_{\text{SEN}} \) is low. Moreover, the undetected links are of relatively small \( g_{ij} \), thus
whose dynamics is described by Eq. (1) with coupling function

The dynamics with 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 $T_{\text{in}} = 5000$. 

The reconstructed average coupling strength is overestimated resulting in an underestimation of $G_{ij}$ and thus $D < 1$.

Then we compare the distribution $P(G)$ of the reconstructed $G_{ij}^{(e)}$ against the actual distribution. As shown in Fig. 3, our method captures the shape of $P(G)$ rather well and, in particular, reproduces both the peaked and skewed $P(G)$ of WSF1 and the power law in $P(G)$ of WSF2 even though it misses links of small $G_{ij}$ and for WSF2 with Rössler, FHN1, and FHN2 dynamics. Eq. (16) does not hold for a small number of very large $S_i$. The reconstructed $S_i^{(e)}$ is in very good agreement with the actual $S_i$ except for WSF1 with FHN1 and FHN2 dynamics and for large $S_i$ for WSF2 with Rössler, FHN1, and FHN2 dynamics (see Fig. 4). The fluctuation for large $S_i$ for WSF2 is a result of the deviation of Eq. (16) for large $S_i$. Using Eqs. (9) and (10), we obtain $S_i^{(e)} / S_i = [G_i^{(e)} / G_i] [(k_i^{(e)} / k_i)(k_i / k_i^{(e)})]$, where $\langle k_i^{(e)} \rangle = \sum_i k_i^{(e)} / N, G_i = \sum_j G_{ij} A_{ij} / k_i$, and $G_i^{(e)} = \sum_j G_{ij} A_{ij}^{(e)} / k_i^{(e)}$ is the average relative coupling strength of the links of node $i$, and $G_i^{(e)} \equiv \sum_j G_{ij}^{(e)} A_{ij}^{(e)} / k_i^{(e)}$ is the reconstructed value. By fitting $G_i^{(e)} = E G_i$ we get an estimate $S_i^{(e), \text{est}} = E[(k_i^{(e)} / k_i)(k_i / k_i^{(e)})]$, which captures well the observed $S_i^{(e)}$ [see Fig. 4(b)], showing that the deviation of $S_i^{(e)}$ from $S_i$ for WSF1 with FHN1 and FHN2 dynamics is due to the inaccuracy of $k_i^{(e)}$.

We have presented a method that reconstructs bidirectional weighted networks using only nodal dynamics. Equation (16) is the basis of why our method works. For networks whose dynamics is described by Eq. (1) with coupling function satisfying Eq. (4), we have derived Eq. (6) that leads to Eq. (8), a special case of Eq. (16) with $B = h'(0)$. Although we have no proof of how general Eq. (16) is, our numerical studies show that it holds also for some systems with higher-dimensional nonlinear FHN or Rössler dynamics and with nonlinear cubic or synapticle like coupling. Whenever Eq. (16) holds and the data scatter is small enough such that $C_i^{(e)}$ for $g_{ij} = 0$ would not overlap significantly with $C_i^{(e)}$ for $g_{ij} \neq 0$, our method can reconstruct accurately both the links and their relative coupling strength. Using several examples, we have demonstrated that our method gives accurate $P_{\text{SEN}}$ and $P_{\text{SPEC}}$ (>80%) and captures the shape of $P(G)$ for WR and WSF networks with both linear and nonlinear dynamics. We emphasize that the connectivity information of a network is contained in $C^+$ and not $C$, and $C^+$ is closely related to the inverse of the usual covariance matrix $\Sigma^{-1}$, where $\Sigma_{ij} = \langle (x_i(t) - \langle x_i(t) \rangle_T)(x_j(t) - \langle x_j(t) \rangle_T) \rangle_T$. Our work thus further explains why weakly correlated nodes can interact strongly [33]. Finally we note that when the interactions and their strengths are known, a feedback method that controls networks to desired dynamical states can be designed and implemented [34,35].

The work of E.S.C.C. and C.Y.L. has been supported by the Hong Kong Research Grants Council under Grant No. CUHK 14300914 while that of P.Y.L. has been supported by Ministry of Science and Technology of Taiwan under Grant No. NSC 101-2112-M-008-004-MY3.


[22] A related result between the Fourier transform of the cross-correlation function of the nodal dynamics and the Laplacian matrix for consensus dynamics has been derived in S. Shahrampour and V. M. Preciado, Reconstruction of Directed Networks from Consensus Dynamics, American Control Conference, 1685 (2013) but there is no discussion on how to use it for network reconstruction.

[23] For each node $i$, we arrange $r_{ij}$ in ascending order, denoted as $r_j(m)$ with $r_j(1)$ being the smallest and find $m_0(i)$, which is the largest value of $m$ such that $d_i(m) = r_i(m + 1) - r_i(m)$ exceeds some threshold (the mean value plus two standard deviations). Nodes $j$ with $r_j(i)$ belonging to $r_j(m)$ for $m \leq m_0(i)$ are reconstructed to be connected (unconnected) to node $i$.


[35] A related result between the Fourier transform of the cross-correlation function of the nodal dynamics and the Laplacian matrix for consensus dynamics has been derived in S. Shahrampour and V. M. Preciado, Reconstruction of Directed Networks from Consensus Dynamics, American Control Conference, 1685 (2013) but there is no discussion on how to use it for network reconstruction.