

专题：统计物理和复杂系统

复杂系统重构*

张海峰¹⁾ 王文旭^{2)†}

1) (安徽大学数学科学学院, 合肥 230601)

2) (北京师范大学系统科学学院, 认知神经科学与学习国家重点实验室, IDG/麦戈文脑研究院, 北京 100875)

(2020年1月2日收到; 2020年3月19日收到修改稿)

远离平衡态的开放复杂系统遍及自然、社会和技术领域, 是复杂性科学的主要研究对象. 通过与外界的能量和物质交换, 复杂系统通过自组织形成了多种多样的内在结构、秩序和规律, 对认识和预测复杂系统提出了艰巨的挑战. 随着实验技术的提高和科技的进步, 反映和体现各种复杂系统机理的数据呈指数增长, 为研究复杂系统提供了新的机遇. 通过系统行为表象数据, 揭示复杂系统结构和动力学属于物理领域的反问题, 是认识复杂系统的基础, 是预测系统状态演化的前提, 对于实现系统状态的调控必不可少. 然而, 复杂系统的多样性和复杂性给解决这一反问题造成了极大的困难. 因此, 需要开阔思路, 借助多学科的交叉与融合, 充分挖掘数据中隐藏的知识和深层次机理. 本文综述了近年来复杂系统, 特别是复杂结构重构和推断方面的研究成果, 希望能够启发复杂系统反问题方面的创新. 同时, 也希望呼吁各领域学者都能关注复杂系统反问题, 推动自然、社会、经济、生物、科技领域的交叉与融合, 解决大家共同面对的科学问题.

关键词: 统计物理, 复杂系统, 反问题, 网络重构**PACS:** 89.75.Hc, 02.50.Le, 05.45.Tp, 89.75.Fb**DOI:** 10.7498/aps.69.20200001

1 引言

小到微生物群落, 大到宇宙星系, 现实世界中的大部分系统属于远离平衡态的开放复杂系统. 这些系统通过与外界进行物质与能量的交换, 对抗内部熵的增加, 从而形成耗散结构和自组织现象, 这是复杂系统自发产生秩序和复杂性的根源^[1]. 与近平衡态不同, 在远离平衡态的系统中, 由于存在涨落决定的分叉和相变, 系统中的规律由特定的作用机制决定, 因此, 没有统一的定律, 这是统计物理与复杂系统领域所面对的最大挑战^[4]. BZ 化学振荡反应是典型的远离平衡态的物质获得新特性的例子, 表现出了远离平衡态条件下的长程相关性^[2]. 生物系统也是一大类远离平衡态的稳定系统. 通过从外界吸收能量和物质, 生物系统维持其内部各种

各样的非平衡状态和有序结构^[3], 并且受到自然选择等竞争的作用, 形成了地球上极大的生物多样性和复杂的生态环境^[4,5].

另一方面, 低维确定性混沌^[6]的发现和元胞自动机^[7]的出现, 从根本上改变了人们对复杂性的认识: 确定性的低维非线性系统也可以产生不可预测的复杂性^[8]. 复杂可以源于简单和复杂性的涌现导致传统还原论的局限性^[9], 迫切需要统计物理发展适用于复杂系统的理论与方法. 美国圣塔菲 (Santa Fe) 研究所在这一历史背景下应运而生, 旨在通过多学科的交叉与融合, 研究各种复杂系统的内在机理和演化规律等跨学科问题. 此后, 由于计算机技术的发展和大量数据的产生, 关于复杂系统的研究成果呈指数律增长^[10,11].

复杂系统具备一些普遍的特征和产生复杂性的因素, 主要包括单元 (个体) 动力学复杂、相互作

* 国家自然科学基金 (批准号: 61973001, 11975049, 71631002) 资助的课题.

† 通信作者. E-mail: wenxuwang@bnu.edu.cn

用结构和模式复杂和自适应演化等. 从基因到人脑, 不同层次和尺度的生物系统都具有高度的复杂性. 细胞受到基因和微环境的共同影响, 表现出协作、分裂、凋亡、融合、迁移、突变等复杂行为^[12]. 人的复杂行为源于人脑的复杂性和高级认知功能. 但是, 人并不具有经济学中的完美理性. 人的决策受到各种认知偏见、刻板印象和群体规范的影响. 对人类经济决策行为的深入研究催生了行为经济学^[13], 旨在通过实验和数据分析发现人类有限理性经济行为的内在动机和成因. 对于复杂相互作用结构的研究催生了复杂网络这一研究方向^[14]. 相关研究揭示了复杂系统共有的结构特征, 比如小世界效应、层次结构、社团结构、异质性和多样性等^[15]. 研究复杂结构如何影响其上的动力学对于理解同步、传播、级联失效、合作、协同、集群等群体行为有重要科学价值^[16,17]; 同时, 网络社团检测方法、网络控制方法、结构推断和相变等方法对传统方法进行了拓展^[18,19]. 与此同时, 研究人员逐渐认识到网络结构本身作用的局限性. 事实上, 复杂系统的动力学和群体行为由个体、相互作用模式、相互作用结构和环境共同决定. 复杂系统研究需要与相关学科紧密结合, 才能真正打破学科壁垒和促进学科的交叉与融合, 为深入理解社会、经济、金融和生物复杂性提供有效的理论工具和方法.

远离平衡态的开放复杂系统的耗散结构、自组织有序性和依赖于特定机制的规律, 为认识和预测复杂系统的状态演化提出了艰巨的挑战. 由于实验技术的提高和科技的发展, 反映和体现复杂系统现象和机制的数据呈指数增加, 为研究复杂系统提供了新的基础和肥沃的土壤. 通过复杂系统行为表象数据, 重构和推断复杂系统的结构和动力学, 属于物理科学中的反问题. 重构和推断复杂系统是复杂性研究的基础, 是通过动力学建模预测系统演化的

前提, 是有效调控系统状态的必要条件. 但是, 由于复杂系统机制多样性、表象的复杂性、动态适应性和极大的随机性等因素, 通过可获得的数据解决复杂系统的反问题比研究正问题难度更大, 需要开阔思路, 借助多学科知识的交叉与融合, 针对各类复杂系统的特性, 提出有效的复杂系统重构理论与方法. 本文将综述近年来复杂系统, 特别是复杂网络结构重构方向的代表性成果, 包括基于压缩感知的重构方法、基于微扰响应的推断方法等. 希望能够启发相关的后续研究, 特别是与近些年兴起的机器学习和人工神经网络方法的结合. 21世纪是复杂的世纪, 而复杂系统重构和推断方法是研究复杂现象的基础, 因此, 有不可替代的重要意义. 我们也希望通过本文呼吁各领域的学者都能够关注复杂系统重构问题, 集思广益, 推动多学科的交叉与融合, 在学科边界产生原始创新. 这是统计物理与复杂系统领域新的机遇和挑战.

网络重构 (network reconstruction), 又称网络推断 (network inference), 研究的是基于观测的数据 (如图 1(a) 和图 1(b)) 去推断节点之间的连边关系^[20–23] (如图 1(c)). 一个复杂的系统, 其个体的动力学行为不仅仅只取决于个体本身, 还依赖于和其他个体之间的交互, 这些交互就构成了系统的结构, 个体之间的交互形成了网络. 因此, 网络重构问题本质上属于数据驱动的系统辨识范畴^[24], 是对哪些个体之间存在交互的辨识. 但鉴于网络结构的复杂性、网络节点动力学的非线性以及结构的稀疏性等性质^[25], 一方面需要我们发展系统辨识中的一些经典方法, 如极大似然估计的方法, 另一方面, 需要根据问题的特有属性提出一些新的方法, 如根据网络规模大而稀疏的特点, 我们提出了基于压缩感知的推断方法等. 以下将对近些年在网络重构方面取得的研究进展进行部分总结与展望.

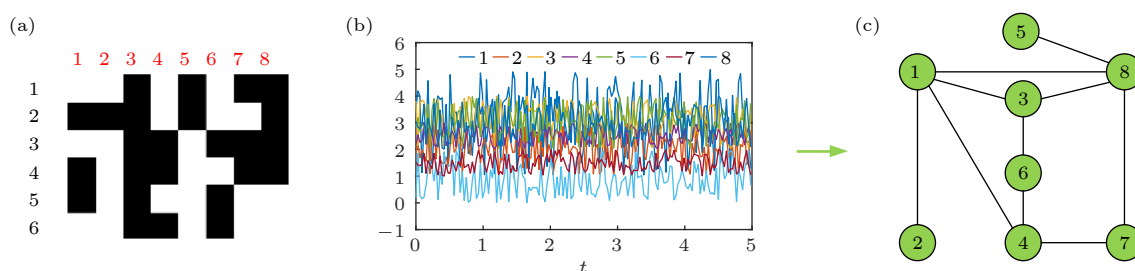


图 1 网络重构示意图 (a) 通过离散的数据; (b) 连续的数据; (c) 推断网络结构

Fig. 1. Illustration of network reconstruction: (a) By using the discrete data; (b) the continuous data; (c) reconstruct network.

2 基于压缩感知的网络重构方法

2.1 压缩感知理论

压缩感知理论是在信号稀疏的情况下, 通过少量的数据采集可以重构原始信号 [26]. 给定一个测量矩阵 $\mathbf{A} \in \mathbb{R}^{M \times N}$, 以及观测值 $\mathbf{Y} \in \mathbb{R}^M$, 可以通过下面公式来重构信号 $\mathbf{X} \in \mathbb{R}^N$:

$$\mathbf{A}\mathbf{X} = \mathbf{Y}. \quad (1)$$

压缩感知的思想是当 \mathbf{X} 是稀疏的时候, 只需要少量的观察数据 ($M \ll N$) 即可重构 \mathbf{X} . 可以通过求解下述凸优化问题 [26–30] 准确得到稀疏信号 \mathbf{X} :

$$\begin{aligned} \min \|\mathbf{X}\|_0 \\ \text{s.t. } \mathbf{A}\mathbf{X} = \mathbf{Y}. \end{aligned} \quad (2)$$

上述问题已经证明了是 NP-hard. 但是在一定条件下最小 L_1 范数下的结果是等价于最小 L_0 范数结果的, 所以有

$$\begin{aligned} \min \|\mathbf{X}\|_1 \\ \text{s.t. } \mathbf{A}\mathbf{X} = \mathbf{Y}. \end{aligned} \quad (3)$$

(3) 式是一个凸优化问题, 已有很成熟的算法作为参考 [27–30]. 因为网络数据具有天然的稀疏性, 所以可以通过压缩感知的方法对其进行网络重构. 如图 2 所示, 就是利用压缩感知方法重构出 Karate 网络的第 4 个节点的邻居, 可以看到求出的 \mathbf{X} (颜色越偏向蓝色, 其值越接近 0) 反映了第 4 个节点

的邻居结构.

2.2 基于耦合振子系统的网络重构

由于描述物理网络中的动力学函数未知, 可以应用幂级数来表示. 又因为级数的高阶项比较多, 因此估计其系数非常困难. 考虑到这些系数非零项很少, 比较稀疏, 而且网络的结构也是稀疏的. 所以可以通过少量的观察数据应用压缩感知的方法重构网络 [31].

一个复杂的振子网络可以通过下面节点动力学描述:

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + \sum_{j=1, j \neq i}^N \mathbf{C}_{ij} \cdot (\mathbf{x}_j - \mathbf{x}_i), \quad (4)$$

其中 $\mathbf{x}_i \in \mathbb{R}^D$ 是节点的状态量, $\mathbf{f}_i(\mathbf{x}_i)$ 为节点自身动力学, 函数形式未知, \mathbf{C}_{ij} 是节点 i 与 j 的耦合矩阵:

$$\mathbf{C}_{ij} = \begin{bmatrix} C_{ij}^{1,1} & C_{ij}^{1,2} & \dots & C_{ij}^{1,D} \\ C_{ij}^{2,1} & C_{ij}^{2,2} & \dots & C_{ij}^{2,D} \\ \dots & \dots & \dots & \dots \\ C_{ij}^{D,1} & C_{ij}^{D,2} & \dots & C_{ij}^{D,D} \end{bmatrix}, \quad (5)$$

其中 $C_{ij}^{k,l}$ 表示 i 节点状态的第 k 个分量与 j 节点状态的第 l 个分量的耦合. 如果矩阵 \mathbf{C}_{ij} 中有一个非零值, 则 i 节点与 k 节点有连边, 如果全为零, 则没有连边. 因此可以通过时间序列推断 \mathbf{C}_{ij} 即可重构网络. 令

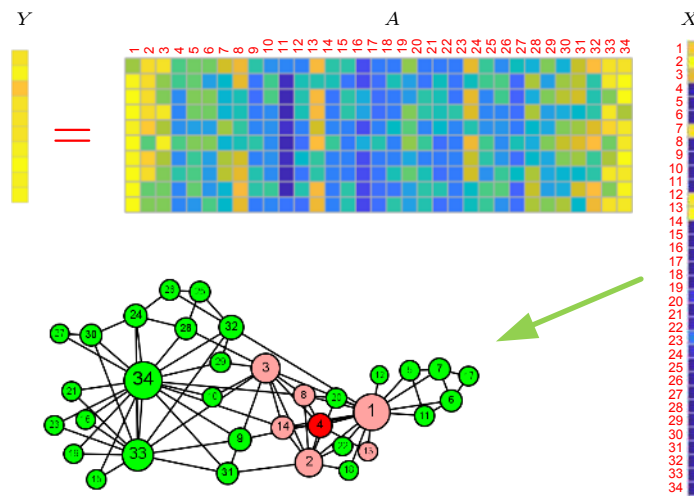


图 2 基于压缩感知方法重构 Karate 网络中 4 号节点的邻居 (重构方法见 2.4 节)

Fig. 2. Reconstructing of node 4 in the Karate network based on compressive sensing framework (the reconstruction method is introduced in Subsec. 2.4).

$$\mathbf{F}_i(\mathbf{x}_i) = \mathbf{f}_i(\mathbf{x}_i) - \sum_{j=1, j \neq i}^N \mathbf{C}_{ij} \cdot \mathbf{x}_j. \quad (6)$$

则其第 k 个分量可以用 n 以下的幂级数的形式表示:

$$[\mathbf{F}_i(\mathbf{x}_i)]_k = \sum_{l_1=0}^n \sum_{l_2=0}^n \cdots \sum_{l_D=0}^n [(\alpha_i)_k]_{l_1, \dots, l_D} \times [(\mathbf{x}_i)_1]^{l_1} [(\mathbf{x}_i)_2]^{l_2} \cdots [(\mathbf{x}_i)_D]^{l_D}, \quad (7)$$

其中 $(\mathbf{x}_i)_k$ 表示第 i 个体状态的第 k 个分量, 可以通过数据观察得到. $[(\alpha_i)_k]_{l_1, \dots, l_D}$ 为幂级数的系数, 是未知量. 可以看出 (7) 式关于未知量是线性的. 给定一个时刻 t_m , 根据 (4) 式与 (7) 式有

$$\dot{\mathbf{x}}_i(t_m) = \mathbf{F}_i(\mathbf{x}_i(t_m)) + \sum_{j=1, j \neq i}^N \mathbf{C}_{ij} \cdot \mathbf{x}_j(t_m), \quad (8)$$

$(m = 1, 2, \dots, M),$

其中 $\dot{\mathbf{x}}_i(t_m)$ 为已知量; $\mathbf{F}_i(\mathbf{x}_i(t_m))$ 中只有 $[(\alpha_i)_k]_{l_1, \dots, l_D}$ 为未知量, 且是稀疏的, 线性的; $\sum_{j=1, j \neq i}^N \mathbf{C}_{ij} \cdot \mathbf{x}_j(t_m)$ 中只有 $C_{ij}^{k,l}$ 是未知量, 也是稀疏的, 线性的. 因此 (8) 式是关于未知量 $[(\alpha_i)_k]_{l_1, \dots, l_D}$ 与 $C_{ij}^{k,l}$ 的 D 个线性方程. 可以测量少量的时间序列 (M 个), 构造一个类似 (1) 式的线性方程组:

$$\mathbf{A}\mathbf{X} = \mathbf{Y}, \quad (9)$$

其中 \mathbf{X} 包含 $[(\alpha_i)_k]_{l_1, \dots, l_D}$ 与 $C_{ij}^{k,l}$, 是稀疏的. 应用压缩感知方法求解, 其中 $C_{ij}^{k,l}$ 可以用来揭示网络的结构 [31,32].

2.3 基于演化博弈系统的网络重构

在生物、社会科学和经济学中的许多复杂动力系统都可以用演化博弈建立数学模型 [33]. 在演化博弈试验中, 每一个人可以处于两种状态: 合作与背叛, 分别可以表示为 $\mathbf{S}(C) = [1, 0]^T$ 与 $\mathbf{S}(D) = [0, 1]^T$, 博弈中双方收益是由博弈双方的策略, 以及收益矩阵 \mathbf{P} (在囚徒困境博弈 [34] 中 $\mathbf{P} = \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix}$) 决定. 所以第 i 节点的收益为

$$P_i = \sum_{j=1, j \neq i}^N a_{ij} \mathbf{S}_i^T \cdot \mathbf{P} \cdot \mathbf{S}_j. \quad (10)$$

进行 M 轮演化博弈实验, 收集每个人状态与收益, 即有 M 个线性方程:

$$P_i(t_m) = \sum_{j=1, j \neq i}^N a_{ij} \mathbf{S}_i^T(t_m) \cdot \mathbf{P} \cdot \mathbf{S}_j(t_m), \quad (11)$$

$(m = 1, 2, \dots, M),$

式中只有 a_{ij} 为未知量. 所以上述公式可以写成以下形式:

$$\mathbf{G}_i \mathbf{A}_i = \mathbf{P}_i. \quad (12)$$

其中 $\mathbf{A}_i = [a_{i1}, \dots, a_{i,i-1}, a_{i,i+1}, \dots, a_{iN}]^T$, \mathbf{G}_i 与 \mathbf{P}_i 已知. 应用压缩感知理论即可求解 \mathbf{A}_i , 从而揭示网络的结构 [35–37], 该方法也可以推广到加权网络.

2.4 基于二值动力学系统的网络重构

二值动力学是复杂系统中常见的一种动力学形式 [38–41], 如疾病传播动力学中的感染态与易感态、演化博弈中的背叛与合作、Ising 动力学中的自旋向上和自旋向下等等. 对于疾病传播动力学, 可以应用 SIS(susceptible-infected-susceptible) 模型或 CP (contact process) 模型来模拟其传播过程. 文献 [42] 应用压缩感知的方法给出了详细的重构过程. 这里只简单介绍 SIS 模型的重构方法.

在 SIS 模型中, 如果 i 节点处于易感态 ($S_i^t = 0$), 且它与一个感染态节点 j 相连 ($S_j^t = 1$), 则会以 λ_i 的概率被感染, 感染态节点 j 会以 δ_j 的概率恢复成易感态. 因此 t 时刻易感状态的节点 i 被感染的概率 $P_i^{01}(t)$ 可以表示为

$$P_i^{01}(t) = 1 - (1 - \lambda_i)^{\sum_{j=1, j \neq i}^N a_{ij} S_j^t}, \quad (13)$$

两边取对数有

$$\ln [1 - P_i^{01}(t)] = \ln(1 - \lambda_i) \sum_{j=1, j \neq i}^N a_{ij} S_j^t. \quad (14)$$

通过一些方法 [42] 在时间序列中统计出 $P_i^{01}(t)$ 以及 S_j^t 构造 M 个线性方程, 类似

$$\mathbf{X}_i \mathbf{A}_i = \mathbf{Y}_i, \quad (15)$$

其中

$$\mathbf{A}_i = [\ln(1 - \lambda_i) a_{i1}, \dots, \ln(1 - \lambda_i) a_{i,i-1}, \ln(1 - \lambda_i) a_{i,i+1}, \dots, \ln(1 - \lambda_i) a_{iN}]^T$$

\mathbf{X}_i 与 \mathbf{Y}_i 已知. 应用压缩感知理论即可求解 \mathbf{A}_i , 从而揭示网络的结构 [42].

当二值动力学未知的情况, 可以记两种状态分别为激活态与未激活态, 假设一个个体 i 由未激活变为激活态的概率与处于激活态邻居的个数有关,

对其线性化^[43]:

$$P_i^{01}(t) \approx c_i \sum_{j=1, j \neq i}^N a_{ij} S_t^j + d_i. \quad (16)$$

类似地, 通过一些方法^[43], 在时间序列中统计出 $P_i^{01}(t)$ 以及 S_t^j 构造 M 个线性方程, 可以写成 (15) 式的形式, 进而通过压缩感知进行求解, 也可以通过 Lasso 进行求解^[43].

压缩感知在动力学系统重构与网络重构的应用还有很多, 如 Wang 等^[44] 利用压缩感知可以重构非线性动力学系统; Su 等^[45] 利用压缩感知重构具有空间地理信息的网络, 不仅可以重构网络结构还可以得到每个节点所在的地理位置; Su 等^[46] 还通过外部事件序列, 利用压缩感知探测隐藏节点; Tang 等^[47] 通过交通流量数据, 利用压缩感知对交通网络进行重构; Chen 和 Lai^[48] 通过玻尔兹曼机对动力学进行估计, 然后应用压缩感知方法重构网络; 最近压缩感知方法还被推广到了多层网络^[49–51]、加权网络的重构^[52] 等.

3 重构非线性动力学系统网络

对于连续的非线性动力学系统, 一般情况下, N 个节点的网络动力学可以由以下常微分方程描述:

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + \sum_{j=1}^N J_{ij} \mathbf{g}_{ij}(\mathbf{x}_i, \mathbf{x}_j) + \mathbf{I}_i(t) + \boldsymbol{\eta}_i(t), \quad (17)$$

其中 $\mathbf{x}_i(t) = [x_{i1}(t), x_{i2}(t), \dots, x_{iD}(t)]^T \in \mathbb{R}^D$ 是节点的状态量, \mathbf{f}_i 表示节点自身动力学, \mathbf{g}_{ij} 表示节点之间的耦合函数, $\mathbf{I}_i(t)$ 表示外部驱动信号, $\boldsymbol{\eta}_i(t)$ 表示外部噪音. J_{ij} 表示耦合矩阵, 反映网络的拓扑结构, 简单情况下为网络的邻接矩阵 \mathbf{A} .

3.1 直接方法

给定网络动力学:

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + \sum_{j=1}^N J_{ij} \mathbf{g}_{ij}(\mathbf{x}_i, \mathbf{x}_j). \quad (18)$$

当其中的内部动力学与耦合函数已知的情况下, 可以通过记录时间序列中不同时刻的数据, 以此来重构网络^[53]. 第 i 个体的第 d 维动力学公式可以表示为

$$\begin{aligned} \dot{x}_i^{(d)}(t_m) &= f_i^{(d)}(\mathbf{x}_i(t_m)) \\ &+ \sum_{j=1}^N J_{ij} g_{ij}^{(d)}(x_i(t_m), \mathbf{x}_j(t_m)). \end{aligned} \quad (19)$$

如果可以观察 M 个时刻, 将会构造 M 个线性方程:

$$\dot{x}_{i,m}^{(d)} = f_{i,m}^{(d)} + \sum_{j=1}^N J_{ij} g_{ij,m}^{(d)}, \quad (20)$$

写成矩阵形式为

$$\mathbf{J}_i \mathbf{X}_i = \mathbf{Y}_i, \quad (21)$$

其中 $X_{i,m} = g_{ij,m}^{(d)}$, $Y_{i,m} = \dot{x}_{i,m}^{(d)} - f_{i,m}^{(d)}$. 求解 $\mathbf{J}_i = [J_{i1}, J_{i2}, \dots, J_{iN}]^T$, 可以得到关于 i 个体的连接情况. 对于 (21) 式的求解, 文献^[53] 中采用了最小化误差的方法求解. 该方法可以解决各种动力学系统, 当网络比较稀疏时, 可以应用 L_1 范数进行约束, 只需观察很少的时间序列即可重构整个网络^[54].

3.2 自适应同步方法

在动力学网络中的局部动力学与耦合函数已知的情况下, 可以根据原系统复制一个辅助系统, 然后通过不停迭代辅助系统中的网络结构使得复制系统与原系统同步. 则得到的辅助系统中的网络结构就是我们要重构的原始系统的结构^[55]. 考虑一个系统 (以一维为例):

$$\dot{x}_i = f_i(x_i) + \sum_{j=1}^N J_{ij} g_j(x_j), \quad (22)$$

假设 f_i 与 g_j 满足利普希茨连续条件. 给原始系统一个副本:

$$\dot{y}_i = f_i(y_i) + \sum_{j=1}^N K_{ij} g_j(y_j) + I_i, \quad (23)$$

其中 y_i 表示复制系统的状态, K_{ij} 表示复制系统的耦合强度, I_i 为可控制信号. 根据原系统和复制系统的状态, 可以通过不停迭代 I_i 与 K_{ij} 使得原系统与复制系统同步. 定义同步误差为

$$e_i = y_i - x_i, \quad (24)$$

副本中的耦合强度调整为

$$\dot{K}_{ij} = -\gamma_{ij} g_j(y_j) e_i, \quad (25)$$

可控信号设置为 $I_i = -\alpha e_i$. 在这里 $\gamma_{ij} > 0$, $\alpha > 0$. 文献^[55] 中证明了当 α 足够大的情况下, 两个系统的误差随着时间减少, 最终会收敛到同步状态. 此

时复制系统的拓扑结构与原系统的拓扑结构相同, 即 $K_{ij} \simeq J_{ij}$, 以此重构网络的局部结构.

对于自适应同步的方法, Zhou 和 Lu^[56] 把该方法推广到了加权网络; Liu 等^[57] 将这一方法推广到了含有耦合延迟的非自治复杂网络的拓扑识别; 更进一步, Wu 等^[58] 利用该方法研究了含时变耦合延迟和受随机扰动影响下的网络重构; Zhao 等^[59] 把这一方法推广到了多层网络, 等等^[60–64].

3.3 驱动-响应实验

上述方法都需要在已知节点的局部动力学以及耦合函数情况下重构网络结构, 下面将介绍一些在节点的局部动力学和耦合函数未知情况下的网络重构方法.

如果一个系统存在一个稳态, 当给系统一个微弱的、持续的扰动, 这个系统将趋向另外一个稳态, 且与第一个稳态相似. 两个稳态的差异不仅取决于驱动信号, 而且与网络的拓扑结构有关 (如图 3 所示). 因此可以通过多次不同的驱动-响应实验, 以此重构整个网络结构.

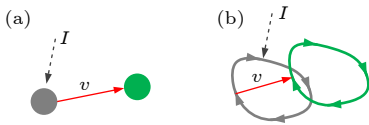


图 3 驱动-响应实验示意图. 对稳态系统施加 (稳态是一个稳定点 (a), 或者一个周期轨道 (b)) 一个持续驱动 I , 系统达到另外一个稳态. 两个稳态的差异 v 包含了网络的拓扑结构

Fig. 3. Driving-response experiments. System is shifted from one stable state (the stable state is a fixed point (a), or a periodical trajectory (b)) to another position by input a driving signal I . The difference of the trajectories contains information about the topology.

这种方法首先是为了解决生物网络上拓扑识别, 特别是基因调控网络^[65–67], 其动力学一般可以应用非线性动力系统描述. 当系统趋于稳态时, 这种系统可以近似为一个一阶线性微分方程:

$$\dot{x}_i = \sum_{j=1}^N \tilde{J}_{ij} x_j + I_i(t), \quad (26)$$

其中 x_i 表示第 i 个 RNA、蛋白质或代谢物的浓度; 和前面一样, \tilde{J}_{ij} 反映网络的拓扑结构, $I_i(t)$ 表示外部的扰动. 当给定一个持续的微弱的扰动 $I_i(t) = I_{i,m}$, 系统将趋向一个新的稳态 $x_{j,m}^*$. 当对

每个个体都进行 M 次扰动实验, 会得到一个线性方程组:

$$\sum_{j=1}^N \tilde{J}_{ij} x_{j,m}^* = -I_{i,m}. \quad (27)$$

通过求解此线性方程组即可推断网络的结构. 对于每个个体 i 都可以构造类似以下方程组:

$$\tilde{J}_i \mathbf{X}_i = \mathbf{Y}_i, \quad (28)$$

求解 $\tilde{J}_i = [\tilde{J}_{i1}, \tilde{J}_{i2}, \dots, \tilde{J}_{iN}]^T$, 可以得到关于 i 个体的连接情况.

上述描述的系统稳态是趋向一个不动点, 然而在自然界还存在更多、更复杂的系统. 而另外一种简单的系统, 是稳态趋向一个周期轨道, 它经常以耦合振子的极限环形式出现. 对于此系统, 也可以应用对稳态系统微小地、持续地扰动来实现网络重构^[68]. 网络动力学可以给定:

$$\dot{\phi}_i = \omega_i + \sum_{j=1}^N J_{ij} g_{ij} (\phi_j - \phi_i) + I_{i,m}, \quad (29)$$

其中 $\phi_i(t)$ 表示振荡器 i 的相位, $\omega_i(t)$ 表示振荡器 i 的频率; 和前面一样, J_{ij} 反映网络的拓扑结构, $I_{i,m}$ 表示持续的外部扰动. 当外部没有驱动时, $m = 0$, 此时 $I_{i,0} = 0$.

对于驱动 $I_{i,m}$, 考虑稳态上面的锁相吸引子, 相位差可以表示为

$$\Delta_{ij,m} = \phi_{i,m} - \phi_{j,m}. \quad (30)$$

当对原始系统 ($I_{i,0}$) 的扰动 ($I_{i,m}$) 是微小的, 则会有 $|\Delta_{ij,m} - \Delta_{ij,0}| \ll 1$.

给定一个微小驱动 $I_{i,m}$, 集体的频率可以观察:

$$\Omega_m = \omega_i + \sum_{j=1}^N J_{ij} g_{ij} (\phi_{j,m} - \phi_{i,m}) + I_{i,m}. \quad (31)$$

令

$$D_{i,m} = \Omega_m - \Omega_0 - I_{i,m} = \sum_{j=1}^N J_{ij} [g_{ij} (\Delta_{ij,m}) - g_{ij} (\Delta_{ij,0})], \quad (32)$$

对 g_{ij} 在 $\Delta_{ij,0}$ 处进行线性展开可以得到包含拓扑结构 J_{ij} 的线性方程^[68], 当给定许多次扰动的情况下可以得到类似 (28) 式的线性方程组, 求解即可推断网络的拓扑结构.

上述两个方法都是在系统的稳态附近进行微小扰动, 下面将介绍一种将系统驱动到一个已知不

动点以此重构网络的方法^[69,70]. 考虑一个动力学网络:

$$\dot{x}_i = f_i(x_i) + \sum_{j=1}^N J_{ij}g_{ij}(x_i, x_j) + I_i, \quad (33)$$

参数如上述描述. 给定该系统一个持续的驱动:

$$I_i = -\theta(x_i - \hat{x}_i). \quad (34)$$

当 θ 足够大, 且 f_i 与 g_{ij} 满足利普希茨连续条件时, 该系统会被驱动到一个稳定点: $\mathbf{x}_s = [x_{1,s}, x_{2,s}, \dots, x_{N,s}]^T$, 它接近于 $\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_N]^T$. 即有

$$\theta(x_{i,s} - \hat{x}_i) = f_i(x_{i,s}) + \sum_{j=1}^N J_{ij}g_{ij}(x_{i,s}, x_{j,s}). \quad (35)$$

定义:

$$\Delta_i = f_i(\hat{x}_i) + \sum_{j=1}^N J_{ij}g_{ij}(\hat{x}_i, \hat{x}_j) - \left[f_i(x_{i,s}) + \sum_{j=1}^N J_{ij}g_{ij}(x_{i,s}, x_{j,s}) \right], \quad (36)$$

则有

$$\theta(x_{is} - \hat{x}_i) = f_i(\hat{x}_i) + \sum_{j=1}^N J_{ij}g_{ij}(\hat{x}_i, \hat{x}_j) - \Delta_i. \quad (37)$$

要判断 k 与 i 的关系, 取 $\hat{\mathbf{x}} = [0, \dots, \hat{x}_k, \dots, 0]^T$, 则

$$\theta_{x_{is}} = f_i(0) + \sum_{j \neq k}^N J_{ij}g_{ij}(0, 0) + J_{ik}g_{ik}(0, \hat{x}_k) - \Delta_{i,k}. \quad (38)$$

\hat{x}_k 进行两次实验, 分别取 $\hat{x}_{k,1}$ 与 $\hat{x}_{k,2}$, 然后两式取差有

$$\theta(x_{is,1} - x_{is,2}) = J_{ik} [g_{ik}(0, \hat{x}_{k,1}) - g_{ik}(0, \hat{x}_{k,2})] - [\Delta_{ik,1} - \Delta_{ik,2}], \quad (39)$$

于是可以写成

$$\theta_{s_{ik}} = \begin{cases} J_{ik}\eta_{ik} + \lambda_{ik}, & \text{if } a_{ik} = 1, \\ \lambda_{ik}, & \text{if } a_{ik} = 0. \end{cases} \quad (40)$$

因此固定节点 k , 对 $|\theta_{s_{ik}}|$ 进行排序处理即可得到 k 的连接情况^[69,70].

3.4 基于噪音的方法

当一个动力学网络趋向同步时, 在没有外部扰动的情况下, 整个系统的每个个体状态一致, 它们之间有效的相互作用消失, 从而无法提取其中的结构信息. 但是, 噪音会导致去同步化, 可以在时间序列中包含个体之间潜在的交互信息^[21].

因此可以通过噪音的扰动来揭示网络的结构^[71-74]. 考虑一个动力学网络:

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i) + c \sum_{j=1}^N L_{ij} \mathbf{H}(\mathbf{x}_j) + \eta_i(t), \quad (41)$$

其中 c 为耦合强度, \mathbf{H} 为耦合函数, $\hat{\mathbf{L}} = L_{ij}$ 为拉普拉斯矩阵, 其定义为: 如果 i 与 j 相连, 且 $i \neq j$, 则 $L_{ij} = -1$, 否则 $L_{ij} = 0$, L_{ii} 为节点 i 的度.

令 $\bar{\mathbf{x}}_i$ 表示 \mathbf{x}_i 无噪音时候的状态, 假设会有一个微小的偏差 ξ_i , 因此有 $\mathbf{x}_i = \bar{\mathbf{x}}_i + \xi_i$. 对上述公式在 $\bar{\mathbf{x}}_i$ 线性化处理, 有

$$\dot{\xi} = \left[D\hat{\mathbf{F}}(\bar{\mathbf{x}}) - c\hat{\mathbf{L}} \otimes D\hat{\mathbf{H}}(\bar{\mathbf{x}}) \right] \xi + \eta, \quad (42)$$

其中 $\xi = [\xi_1, \xi_2, \dots, \xi_N]^T$, $\eta = [\eta_1, \eta_2, \dots, \eta_N]^T$, $D\hat{\mathbf{F}}(\bar{\mathbf{x}}) = \text{diag}[D\hat{\mathbf{F}}_1(\bar{\mathbf{x}}_1), D\hat{\mathbf{F}}_2(\bar{\mathbf{x}}_2), \dots, D\hat{\mathbf{F}}_N(\bar{\mathbf{x}}_N)]$, $D\hat{\mathbf{F}}_i$ 为 \mathbf{f}_i 的雅可比矩阵, 同理可以定义 $D\hat{\mathbf{H}}(\bar{\mathbf{x}})$, \otimes 表示点乘. 令 $\hat{\mathbf{A}} = -[D\hat{\mathbf{F}}(\bar{\mathbf{x}}) - c\hat{\mathbf{L}} \otimes D\hat{\mathbf{H}}(\bar{\mathbf{x}})]$, $\hat{\mathbf{C}} = \langle \xi \xi^T \rangle$, 其中 $\hat{\mathbf{A}}$ 中包含网络结构. 于是有

$$0 = \left\langle \frac{d(\xi \xi^T)}{dt} \right\rangle = -\hat{\mathbf{A}}\hat{\mathbf{C}} - \hat{\mathbf{C}}\hat{\mathbf{A}}^T + \langle \eta \xi^T \rangle + \langle \xi \eta^T \rangle, \quad (43)$$

设 $\hat{\mathbf{D}}$ 是噪音的协方差矩阵, 有 $\langle \eta \xi^T \rangle = \langle \xi \eta^T \rangle = \hat{\mathbf{D}}/2$ ^[71]. 于是

$$\hat{\mathbf{A}}\hat{\mathbf{C}} + \hat{\mathbf{C}}\hat{\mathbf{A}}^T = \hat{\mathbf{D}}, \quad (44)$$

其中 $\hat{\mathbf{C}}$ 与 $\hat{\mathbf{D}}$ 都可以通过时间序列观察得到, 求解 $\hat{\mathbf{A}}$ 中的 $\hat{\mathbf{L}}$ 即可重构网络的结构.

4 似然估计的方法重构网络

对于二值动力学, 当给定动力学过程以及网络, 会得到一系列时间序列. 但是, 当网络结构未知, 给定时间序列以后, 就变成了什么样的网络结构最有可能产生这种时间序列, 即应用最大似然估计的方法即可以得到^[75-81].

当动力学过程已知, 按照前面所述, P_i^{01} 与 P_i^{10} 可以用公式表示出来, 且是关于 $\sum_{j=1, j \neq i}^N a_{ij} S_t^j$ 的函数. 因此可以通过大量的时间序列构造似然函数, 但是由于 $\sum_{j=1, j \neq i}^N a_{ij} S_t^j$ 这一项在似然函数中往往都是非常复杂的非线性项(如指数函数), 因此需要通过平均场近似(即邻居中被激活个体的比例近似于整个网络中被激活个体的比例)的方法进行处理:

$$\frac{1}{k_i} \sum_{j=1, j \neq i}^N a_{ij} S_t^j \approx \frac{1}{N-1} \sum_{j=1, j \neq i}^N S_t^j, \quad (45)$$

然后对似然函数求微分取极值, 再将 $\sum_{j=1, j \neq i}^N a_{ij} S_t^j$ 在 $\frac{k_i}{N-1} \sum_{j=1, j \neq i}^N S_t^j$ 处一阶泰勒展开, 得到关于 a_{ij} 的线性方程组, 从而推断出网络结构 [75]. 该方法还可以推断出符号网络 [77] 与多层网络 [76]. 但是该方法有一个缺点, 就是需要知道动力学过程, 因此刘等 [78] 应用 Logistic 回归中的 Sigmoid 函数近似二值动力学过程, 然后再应用平均场近似的方法进行重构, 也得到了很好的效果, 尤为重要, 该方法还可以进一步复现原始动力学, 即二值动力学中的转移概率函数.

对于二值动力学, 假设一个节点由未激活态 (t 时刻) 变为激活态 ($t+1$ 时刻) 只受激活态的邻居的影响, 通过贝叶斯公式, 有

$$\begin{aligned} & P\left(S_{t+1}^j = 1, i \rightarrow j \mid S_t^i = 1, S_t^j = 0\right) \\ &= P\left(i \rightarrow j \mid S_t^i = 1, S_t^j = 0, S_{t+1}^j = 1\right) \\ &\quad \times P\left(S_{t+1}^j = 1 \mid S_t^i = 1, S_t^j = 0\right), \quad (46) \end{aligned}$$

其中 $i \rightarrow j$ 表示节点 i 对节点 j 的直接影响. 令 $P_{i \rightarrow j} = P\left(i \rightarrow j \mid S_t^i = 1, S_t^j = 0, S_{t+1}^j = 1\right)$, 显然 $P_{i \rightarrow j} > 0$, 表示节点 i 是节点 j 的邻居. 再令

$P_i^j = P\left(S_{t+1}^j = 1 \mid S_t^i = 1, S_t^j = 0\right)$, 可以通过数据统计出来, 这是一个已知量. 则节点 j 在第 t_m 时刻没被激活, $t_m + 1$ 时刻被激活的次数的期望可以表示为

$$E_j^{t_m+1} = \sum_{i(i \neq j)} P_{i \rightarrow j} P_i^j \Psi_i^{t_m} + \varepsilon_j, \quad (47)$$

其中 $\Psi_i^{t_m}$ 表示节点 i 在第 t_m 时刻被激活的次数, ε_j 表示重构的噪音项. 然后通过假设激活次数服从泊松分布, 利用时间序列可以写出似然函数, 最后应用 expectation-maximization (EM) 算法求解 $P_{i \rightarrow j}$, 从而推断出网络结构 [75,79,80], 图 4 为应用此方法推断出的 Karate 网络的第 33 个节点的结构. 另外, Ma 等 [81] 应用最大似然估计针对终态数据对网络重构进行了尝试性研究.

5 其他方法

网络重构的方法还有很多, 例如, Wu 等 [82,83] 将格兰杰因果关系检验运用到了网络重构当中, 并针对不同的情况, 提出了传统格兰杰因果关系检验、条件格兰杰因果关系检验、分段格兰杰因果关系检验、随机扰动分段格兰杰因果关系检验、偏相关格兰杰因果关系检验等不同的格兰杰因果关系检验识别方法; Li 和 Li [84] 通过时效网络扩散过程的到达时间数据, 提取时效交互过程的统计特征和

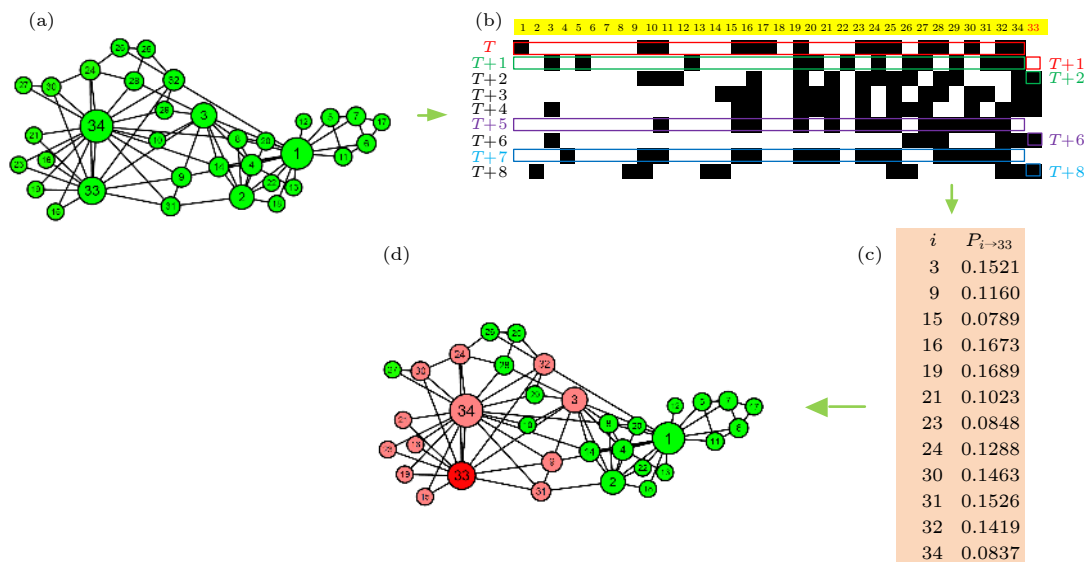


图 4 EM 算法推断 Karate 网络 33 号节点的结构 (a) 网络结构; (b) 二进制数据; (c) EM 算法推断出节点 33 的结构; (d) 真实网络 33 号节点的结构

Fig. 4. Reconstructing the neighbors of node 33 in Karate network: (a) The real structure of the Karate network; (b) the binary state data; (c) inferring the neighbors of node 33 based on EM algorithm; (d) the real neighbors of node 33.

推断出时效网络的拓扑结构, 并严格证明了推断结构的渐近一致性; Casadiego 等^[85] 通过引入显式依赖矩阵把每个个体的动力学分解成与其他个体两点、三点或更高阶的相互作用, 从而可以通过数据揭示网络 (两点) 和超网络 (三点) 的交互作用.

6 未来展望

网络重构发展到现在, 虽然已经取得了一些成就, 但是还有很多问题需要解决. 现有的方法主要还是针对简单的动力学网络. 对于多层网络的重构、符号网络的重构、时效网络的重构、多体动力学网络的重构等等, 虽然有一些文献已经对这些问题有所涉及, 但是其重要性还没有受到应有的关注, 也没有很好地解决. 而且现在的网络重构针对的主要还是小规模网络, 如何快速地、精确地重构大规模网络也是以后需要解决的问题.

参考文献

- [1] Prigogine I, Hiebert E N 1982 *Phys. Today* **35** 69
- [2] Haken H 2006 *Information and Self-organization: A Macroscopic Approach to Complex Systems* (Berlin: Springer)
- [3] Schrödinger E 1944 *What is Life? The Physical Aspect of the Living Cell and Mind* (Cambridge: Cambridge University Press)
- [4] Wilson E O 1992 *The Diversity of Life* (Boston: Belknap Press)
- [5] Wilson E O 2016 *Half Earth: Our Planet's Fight for Life* (London: Liveright)
- [6] Lorenz E N 1963 *J. Atmos. Sci.* **20** 130
- [7] Conway J H 2000 *On Numbers and Games* (Boca Raton: AK Peters/CRC Press)
- [8] Ott E 2002 *Chaos in Dynamical Systems* (Cambridge: Cambridge University Press)
- [9] Barabási A L 2012 *Nat. Phys.* **8** 14
- [10] Downey A 2018 *Think Complexity: Complexity Science and Computational Modeling* (Sebastopol: O'Reilly Media)
- [11] Johnson N 2009 *Simply Complexity: A Clear Guide to Complexity Theory* (Oxford: Oneworld Publications)
- [12] Alberts B, Bray D, Hopkin K, Johnson A D, Lewis J, Raff M, Roberts K, Walter P 2013 *Essential Cell Biology* (New York: Garland Science)
- [13] Cartwright E 2018 *Behavioral Economics* (London: Routledge)
- [14] Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
- [15] Newman M E J 2003 *SIAM Rev.* **45** 167
- [16] Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D U 2006 *Phys. Rep.* **424** 175
- [17] Pastor-Satorras R, Castellano C, van Mieghem P, Vespignani A 2015 *Rev. Mod. Phys.* **87** 925
- [18] Liu Y Y, Barabási A L 2016 *Rev. Mod. Phys.* **88** 035006
- [19] Stankovski T, Pereira T, McClintock P V E, Stefanovska A 2017 *Rev. Mod. Phys.* **89** 045001
- [20] Timme M, Casadiego J 2014 *J. Phys. A: Math. Theor.* **47** 343001
- [21] Wang W X, Lai Y C, Grebogi C 2016 *Phys. Rep.* **644** 1
- [22] Lu J A, Lu J H, Liu H, Chen J 2010 *Complex Systems and Complexity Science* **7** 63 (in Chinese) [陆君安, 吕金虎, 刘慧, 陈娟 2010 *复杂系统与复杂性科学* **7** 63]
- [23] Wang W X 2013 *Journal of Electronic Science and Technology* **42** 3 (in Chinese) [王文旭 2013 *电子科技大学学报* **42** 3]
- [24] Qin S J 2012 *Annu. Rev. Control* **36** 2
- [25] Zhang Z Y, Chen Y, Mi Y Y, Hu G 2020 *Sci. Sin.-Phys. Mech. Astron.* **1** 3 (in Chinese) [张朝阳, 陈阳, 弭元元, 胡岗 2020 *中国科学: 物理学 力学 天文学* **1** 3]
- [26] Candes E J, Tao T 2006 *IEEE Trans. Inf. Theory* **52** 5406
- [27] Romberg J 2008 *IEEE Signal Process. Mag.* **25** 14
- [28] Candes E J, Wakin M B 2008 *IEEE Signal Process. Mag.* **25** 21
- [29] Candes E J, Romberg J, Tao T 2006 *IEEE Trans. Inf. Theory* **52** 489
- [30] Baraniuk R G 2007 *IEEE Signal Process. Mag.* **24** 118
- [31] Wang W X, Yang R, Lai Y C, Kovanis V, Harrison M A F 2011 *EPL* **94** 48006
- [32] Su R Q, Ni X, Wang W X, Lai Y C 2012 *Phys. Rev. E* **85** 056220
- [33] Szabó G, Fátih G 2007 *Phys. Rep.* **446** 97
- [34] Nowak M A, May R M 1992 *Nature* **359** 826
- [35] Wang W X, Lai Y C, Grebogi C, Ye J P 2011 *Phys. Rev. X* **1** 290
- [36] Ma L, Han X, Shen Z S, Wang W X, Di Z R 2015 *PLoS ONE* **10** 0142837
- [37] Han X, Shen Z S, Wang W X, Lai Y C, Grebogi C 2016 *Sci. Rep.* **6** 30241
- [38] Dorogovtsev S N, Goltsev A V, Mendes J F F 2002 *Phys. Rev. E* **66** 16104
- [39] Pastor-Satorras R, Vespignani A 2001 *Phys. Rev. E* **63** 066117
- [40] Nowak M A, May R M 1993 *Int. J. Bifurcation Chaos* **3** 35
- [41] Wang Y, Xiao G, Liu J 2012 *New J. Phys.* **14** 13015
- [42] Shen Z S, Wang W X, Fan Y, Di Z R, Lai Y C 2014 *Nat. Commun.* **5** 4323
- [43] Li J, Shen Z S, Wang W X, Grebogi C, Lai Y C 2017 *Phys. Rev. E* **95** 032303
- [44] Wang W X, Yang R, Lai Y C, Kovanis V, Grebogi C 2011 *Phys. Rev. Lett.* **106** 154101
- [45] Su R Q, Wang W X, Wang X, Lai Y C 2016 *R. Soc. Open Sci.* **3** 150577
- [46] Su R Q, Wang W X, Lai Y C 2012 *Phys. Rev. E* **85** 065201
- [47] Tang S Q, Shen Z S, Wang W X, Di Z R 2015 *Eur. Phys. J. B* **88** 211
- [48] Chen Y Z, Lai Y C 2018 *Phys. Rev. E* **97** 032317
- [49] Li G J, Li N, Liu S H, Wu X Q 2019 *Chaos* **29** 53117
- [50] Mei G F, Wu X Q, Wang Y F, Hu M, Lu J A, Chen G R 2018 *IEEE Trans. Cybern.* **48** 754
- [51] Wang X, Lu J H, Wu X Q 2018 *IEEE Trans. Syst. Man Cybern. Part A Syst. Humans*
- [52] Liu J, Mei G F, Wu X Q, Lu J H 2018 *IEEE Trans. Circuits Syst. I* **65** 2970
- [53] Shandilya S G, Timme M 2011 *New J. Phys.* **13** 13004
- [54] Han X, Shen Z S, Wang W X, Di Z R 2015 *Phys. Rev. Lett.* **114** 028701
- [55] Yu D, Righero M, Kocarev L 2006 *Phys. Rev. Lett.* **97** 188701
- [56] Zhou J, Lu J A 2007 *Physica A* **386** 481

- [57] Liu H, Lu J A, Lü J H, Hill D J 2009 *Automatica* **45** 1799
- [58] Wu X Q, Zhao X Y, Lu J H, Tang L K, Lu J A 2016 *IEEE Trans. Control Netw. Syst.* **3** 379
- [59] Zhao X Y, Zhou J, Zhu S B, Ma C, Lu J A 2019 *IEEE Trans. Circuits Syst. II* **67** 290
- [60] Chen L, Lu J A, Tse C K 2009 *IEEE Trans. Circuits Syst. II* **56** 310
- [61] Zhou J, Yu W W, Li X M, Small M, Lu J A 2009 *IEEE Trans. Neural Networks* **20** 1679
- [62] Zhu S B, Zhou J, Chen G R, Lu J A 2019 *IEEE Trans. Cybern.*
- [63] Zhu S B, Zhou J, Lu J A 2018 *Chaos* **28** 43108
- [64] Yang P, Zheng Z G 2012 *Acta Phys. Sin.* **61** 120508 (in Chinese) [杨浦, 郑志刚 2012 物理学报 **61** 120508]
- [65] Gardner T S, Di Bernardo D, Lorenz D, Collins J J 2003 *Science* **301** 102
- [66] Tegner J, Yeung M K S, Hasty J, Collins J J 2003 *Proc. Natl. Acad. Sci.* **100** 5944
- [67] Yeung M K S, Tegnér J, Collins J J 2002 *Proc. Natl. Acad. Sci.* **99** 6163
- [68] Timme M 2007 *Phys. Rev. Lett.* **98** 224101
- [69] Yu D C 2010 *Automatica* **46** 2035
- [70] Yu D C, Parlitz U 2010 *Phys. Rev. E* **82** 026108
- [71] Ren J, Wang W X, Li B W, Lai Y C 2010 *Phys. Rev. Lett.* **104** 058701
- [72] Wang W X, Ren J, Lai Y C, Li B W 2012 *Chaos* **22** 33131
- [73] Zhang Z Y, Chen Y, Mi Y Y, Hu G 2019 *Phys. Rev. E* **99** 042311
- [74] Zhang Z Y, Zheng Z G, Niu H J, Mi Y Y, Wu S, Hu G 2015 *Phys. Rev. E* **91** 012814
- [75] Ma C 2019 *Ph. D. Dissertation* (Hefei: Anhui University) (in Chinese) [马闯 2019 博士学位论文 (合肥: 安徽大学)]
- [76] Ma C, Chen H S, Li X, Lai Y C, Zhang H F 2020 *SIAM J. Appl. Dyn.* **19** 124
- [77] Xiang B B, Ma C, Chen H S, Zhang H F 2018 *Chaos* **28** 123117
- [78] Liu Q M, Ma C, Xiang B B, Chen H S, Zhang H F 2019 *IEEE Trans. Syst. Man Cybern. Syst.*
- [79] Ma C, Chen H S, Lai Y C, Zhang H F 2018 *Phys. Rev. E* **97** 22301
- [80] Zhang H F, Xu F, Bao Z K, Ma C 2019 *IEEE Trans Circuits Syst. Regul Pap.* **66** 1608
- [81] Ma C, Zhang H F, Lai Y C 2017 *Phys. Rev. E* **96** 022320
- [82] Wu X Q, Wang W H, Zheng W X 2012 *Phys. Rev. E* **86** 046106
- [83] Wu X Q, Zhou C S, Chen G R, Lu J A 2011 *Chaos* **21** 43129
- [84] Li X, Li X 2017 *Nat. Commun.* **8** 15729
- [85] Casadiego J, Nitzan M, Hallerberg S, Timme M 2017 *Nat. Commun.* **8** 2192

SPECIAL TOPIC—Statistical physics and complex systems

Complex system reconstruction*Zhang Hai-Feng¹⁾ Wang Wen-Xu^{2)†}¹⁾ (*School of Mathematical Science, Anhui University, Hefei 230601, China*)²⁾ (*State Key Laboratory of Cognitive Neuroscience and Learning IDG/McGovern Institute for Brain & Research, School of Systems Science, Beijing Normal University, Beijing 100875, China*)

(Received 2 January 2020; revised manuscript received 19 March 2020)

Abstract

Open complex systems far from equilibrium widely exist in the nature and the fields of society and technology, which are the main research objects of complexity science. Through the exchange of energy and material with the outside world, complex systems can form a variety of internal structures, orders and laws by self-organization behaviors, which poses an arduous challenge to the understanding and predicting complex systems. With the improvement of experimental technology and the progress of science and technology, the data reflecting the mechanism of various complex systems are increasing exponentially, thereby providing new opportunities for studying complex systems. Revealing the structures and dynamics of complex systems from the measured data is an inverse problem in the field of physics, which is the premise of understanding complex systems, predicting the evolution of system state, and regulating system state. However, it is very difficult to solve this inverse problem due to the diversity and complexity of complex system. Therefore, we need to fully mine the hidden knowledge and deep mechanism in the data with the help of interdisciplinary integration. In this paper we briefly review the research results of complex system in recent years, especially the reconstruction of complex network structures, hoping to inspire the innovation to the inverse problem of complex systems. Meanwhile, we hope that researchers in different fields can pay much attention to the inverse problems of complex systems, promote the cross and integration of nature, society, economy, biology and technology, and solve the scientific problems that we are facing.

Keywords: statistical physics, complex systems, inverse problem, network reconstruction**PACS:** 89.75.Hc, 02.50.Le, 05.45.Tp, 89.75.Fb**DOI:** [10.7498/aps.69.20200001](https://doi.org/10.7498/aps.69.20200001)

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61973001, 11975049, 71631002).

† Corresponding author. E-mail: wenxuwang@bnu.edu.cn